

Small Molecule Libraries Used in the Development of Structure Activity Relationships

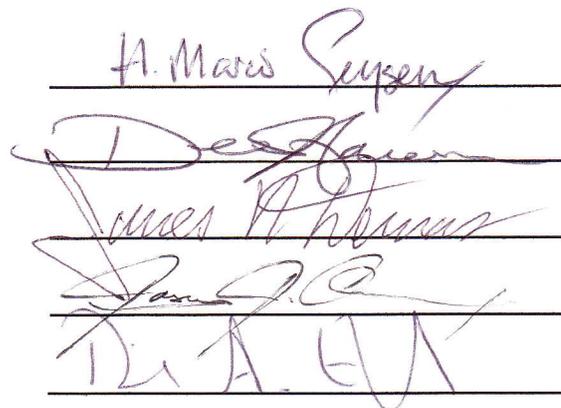
Joseph Jack Jablonski
Rochester, NY

B.S., Chemistry, St. Lawrence University, 2005

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Chemistry

University of Virginia
August, 2011


A. Marc Szymanski
D. J. H. Jones
James A. Thomas
D. A. H. H.

Abstract

With the constant need for the development of new therapeutics, library synthesis has become a staple of modern drug discovery. This technique allows for the synthesis of multiple compounds in parallel to gain insight to activity relationships associated with biological targets. Within this body of work is the development of structure-activity relationships (SAR) at two biologically relevant targets, DPP-IV and NS1.

DPP-IV has been shown to play a critical role in the regulation of incretin hormones responsible for stimulating insulin production. Two known inhibitors of DPP-IV (Vildagliptin and Saxagliptin) fall into the class of compounds known as cyanopyrrolidines and display large hydrophobic substituents at either the *N*-terminus or the α -carbon. Previous SAR studies have examined substitution patterns at the *N*-terminus, with little attention given to substitution at both the α -carbon and the *N*-terminus. A 42-compound library was designed to systematically examine steric restrictions within the DPP-IV binding pocket. Additionally, three compounds were synthesized to be used in the development of a functioning assay. Initial work suggested that the substrate Nle-Pro-AMC was participating in substrate inhibition with DPP-IV at concentrations above 50 μ M. Once this model of analysis was identified, a protocol was developed to screen compounds for biological activity.

The Influenza virus non-structural protein 1 (NS1) is a multifunctional protein which plays a critical role in viral replication. Recent research has identified NS1 as a viable target to treat influenza based on the highly conserved structure across different strains of influenza A and the important cellular interactions which are involved with viral replication. Following a large scale screen, an SAR was developed to design and

synthesize novel inhibitors of NS1. Initial studies identified key interactions with the left hand side of the molecule which may play a role in antiviral activity. The resulting compounds were highly potent while exhibiting little cytotoxicity. When amino acids were used as building blocks, antiviral activity was observed for multiple influenza A virus strains. With this result, a 96-compound library was designed and synthesized using 48 different carboxylic acids, amino acids, or dipeptides to further probe essential structural features required for antiviral activity.

Table of Contents

Abstract.....	ii
Acknowledgements.....	v
List of Abbreviations.....	vi
List of Figures.....	ix
List of Schemes.....	xi
List of Tables.....	xii
Chapter 1. Introduction to Library Synthesis	1
Chapter 2. Introduction to DPP-IV.....	18
Chapter 3. DPP-IV Results.....	46
Chapter 4. DPP-IV Discussion.....	64
Chapter 5. Introduction to NS1.....	71
Chapter 6. NS1 Results.....	94
Chapter 7. NS1 Discussion.....	130
Chapter 8. Experimental Protocols for Selected Compounds.....	142

Acknowledgements

I would like to thank Dr. H. Mario Geysen for all of his patience, guidance, and advice over the past six years. I'd also like to thank all of the past Geysen Lab group members: Dr. Cyrille Gineste, Leslie Willingham, Chuck Arrington, Dr. Catharine Prudom, Dr. Suzy 'the Suz' Arnett, Dr. Sara Firtzgerald, Robby Van Sciver, and Dr. Kalyani Jambunathan. I would also like to thank my collaborator Dr. Daniel A. Engel and his lab for all their work on our project. I would also like to thank Dr. Burnett for always helping me when I needed it.

I'd like to make a special thank you to Dr. Spencer McMinn. We lived together for five years and I don't think it would have been worth it without him.

The Chemistry Department has some wonderful people here who are always willing to help out when needed or go out and have a good time. I would like to thank all of my past and present friends: Robert Fehnel, Kiev Blasier, Dr. Edward Lee, Dr. Andrew Kennedy, Dr. Thomas Matthews, Dr. Albert DeBerardinis, Dr. Andrew Yeagley, Dr. James East, Dr. Jessica Chapman, and Lynn Mandeltort. There have been a number of friends both from Virginia and other places which have contributed keeping me sane, so I would like to thank: Pat and Tina McNelis, Chris and Beth Houlihan, Lauren McMinn, Laurie and Bob McMinn, Alie and Dennis Seese, Dr. Rob Klein, Dr. Erin Goken, Dr. Daniel Sunday, John Kupstas, and Liam de los Reyes.

I would like to thank my entire family. You have all been a huge inspiration to me and helped me along my path, I can't thank you enough.

Finally, I would like to thank my wonderful, loving, and ever patient wife Lori Jablonski. Thanks for putting up with me through this entire process.

List of Abbreviations

3-Pal	(3-pyridyl)alanine
Abu	aminobutyric acid
Ala	alanine
AMC	aminomethylcoumarin
Arg	arginine
BMS	BristolMyersSquibb
Boc	<i>tert</i> -butyloxycarbonyl
Boc ₂ O	Di- <i>tert</i> -butyl dicarbonate
Cha	cyclohexylalanine
Chg	cyclohexylglycine
cLeu	cycloleucine
CoMFA	Comparative Molecular Field Analysis
CPSF30	30kSa subunit of the polyadenylation specificity factor
DCM	dichloromethane
DIC	N,N'-diisopropylcarbodiimide
DIU	diisopropylurea
DMAP	4-dimethylaminopyridine
DPP-IV	Dipeptidyl Peptidase-4
DPP-8	Dipeptidyl Peptidase-8
DPP-9	Dipeptidyl Peptidase-9
Fmoc	9-fluorenylmethyloxycarbonyl
FDA	Food and Drug Administration

GIP	glucose-dependent insulinotropic peptide
GLP-1	glucagon-like peptide 1
Gly	glycine
HA	hemagglutinin
HPLC	high performance liquid chromatography
HOBt	hydroxybenztriazol
IFN	interferon
Ile	isoleucine
LAH	lithium aluminum hydride
Leu	leucine
M+H	molecular weight plus a proton
M2	matrix protein (influenza virus)
μM	micro molar concentration
μL	micro liter volume
μunit	micro unit measurement of DPP-IV enzyme
mol	mole(s)
mmol	milimoles
NA	neuraminidase protein
Nal	naphthylalanine
NCS	National Cancer Institute Diversity Set
Nle	norleucine
NMM	n-methylmorpholine
nm	nanometers

NS1	nonstructural protein 1
Nva	norvaline
OAS	2'-5'-oligo (A) synthase
Phe	phenylalanine
PKR	protein kinase R
RNA	ribonucleic acid
Pro	proline
cRNA	complimentary RNA
dsRNS	double stranded RNA
mRNA	messenger RNA
vRNA	viral RNA
SAR	Structure-Activity Relationship
S _N 2	bimolecular substitution reaction
<i>t_R</i>	retention time
TEF	triethyl orthoformate
TFA	trifluoroacetic acid
TFAA	trifluoroacetic anhydride
THF	tetrahydrofuran
UV/VIS	ultraviolet/visable
Val	valine

List of Figures

Figure 1.1	The structure of Taxol [®]	3
Figure 1.2	Marketed drugs resulting from successful scaffold hopping	9
Figure 1.3	Chemistry space as it related to the drug discovery process (Modified from Opera)	10
Figure 1.4	Compounds used to treat Type 2 Diabetes	12
Figure 1.5	Small molecules which have been identified as inhibitors of NS1	13
Figure 2.1	The sequence of Byetta [®] with the proposed site of metabolism. This figure was taken from Drauker and Nauk	24
Figure 2.2	The structure of Liraglutide. This figure was modified from Drauker and Nauk	26
Figure 2.3	Study showing the decrease in HbA1C percentage using DPP-IV inhibitor (Vildagliptin) vs. Placebo (PBO). Figure was taken from Campbell et al.	27
Figure 2.4	Structures of DPP-IV inhibitors	28
Figure 2.5	Figure demonstrating the binding pocket interactions within the DPP-IV binding pocket	29
Figure 2.6	Scaffolds used in previous SAR studies conducted by Villhauer	30
Figure 3.1	Three compounds synthesized to test for reproducibility	47
Figure 3.2	Sample CoMFA images showing the preference for hydrophobic side chains in the P2 region of the binding pocket	51
Figure 3.3	Raw experimental data showing the decrease in Intensity units over time with increasing substrate concentration (μM)	57
Figure 3.4	Initial rate versus substrate concentration which depicts substrate inhibition	58
Figure 3.5	Data for the three synthesized inhibitors of DPP-IV	59
Figure 5.1	Basic structure of the influenza A virus. Modified from Palese	74

Figure 5.2	Diagram of viral replication. Modified from Fodor <i>et al.</i>	76
Figure 5.3	Structure of Amantadine [®] and Rimantadine [®]	77
Figure 5.4	Marketed form of Tamiflu [®] shown with its active metabolite	79
Figure 5.5	Molecular structure of Relenza [®]	79
Figure 5.6	Dimer model of NS1A where red is RNA, yellow is the RNA binding domain and blue is the C-terminus effector domain. This figure was taken from Lin <i>et al.</i>	81
Figure 5.7	Structures of the compounds which showed anti-influenza activity in the initial screen	85
Figure 6.1	Plotted biological data used to make assessments of activity using K-values and Span	96
Figure 6.2	Compound from initial screen which was used as the parent molecule	97
Figure 6.3	Synthesized analogues based on parent compound 6.1	98
Figure 6.4	Analogues which were synthesized to optimize the carbon chain length	101
Figure 6.5	Molecule 6.12 broken up into three defined sections	101
Figure 6.6	Modifications made to 6.12 to optimize the hydroxyl group positioning	102
Figure 6.7	Analogues synthesized where the hydroxyl is substituted with alternative functional groups	103
Figure 6.8	Plotted data for halogenated compounds vs. parent molecule 6.12	104
Figure 6.9	Analogues with fused ring systems used to test steric restrictions at the binding site	105
Figure 6.10	Plotted data for naphthalene compounds vs. parent molecule 6.12	106
Figure 6.11	Data which compares three different fused ring systems	106
Figure 6.12	Synthetic analogues which incorporate a heterocycle into the left hand side	107

Figure 6.13	Plotted data comparing the pyridyl and quinolyl derivatives	108
Figure 6.14	Plotted data for compounds which compares pyridyl, amino, and cyano derivatives	108
Figure 6.15	Synthetic analogues which include additional functional groups on the left hand ring	109
Figure 6.16	Analogues which move the amide bond away from the hydroxyl group	111
Figure 6.17	Plotted data which compares ether based compounds to the amino analogue	114
Figure 6.18	Synthetic analogues which substitute an aliphatic chain for the aromatic ring with their corresponding biological data compared to 6.1 and 6.12	115
Figure 6.19	Analogues synthesized which examines the activity of previous building blocks with the aliphatic chain	116
Figure 6.20	Plotted data of the aliphatic halogen derivatives compared to 6.47	117
Figure 6.21	Plotted data comparing aliphatic amino and pyridyl derivatives	117
Figure 6.22	Aliphatic analogues which incorporate steric bulk	118
Figure 6.23	Plotted data for amino acid based compounds	121
Figure 6.24	Markush structure used for the design and synthesis of the library	123
Figure 7.1	Original active molecules from the NCS screen	133
Figure 7.2	Structure of the compounds containing fluorescein	137

List of Schemes

Scheme 2.1	Synthesis of compounds utilizing a solid phase approach	31
Scheme 2.2	Synthetic strategy for solution phase synthesis of desired products	32

Scheme 2.3	Formation of the cyclic imidate	33
Scheme 3.1	Synthetic scheme used to synthesis the first three compounds	49
Scheme 3.2	Synthetic scheme for the two methyl alcohol analogues	51
Scheme 3.3	Synthesis of the 42-membered library	52
Scheme 3.4	Proposed synthesis of oxazoline compounds	55
Scheme 3.5	The synthesis of the new substrate Gly-Pro-AMC	61
Scheme 4.1	Final synthetic procedures for the library	68
Scheme 6.1	Synthesis of imine analogues	99
Scheme 6.2	Synthesis of analogues with the removal of the hydrazine linker	100
Scheme 6.3	Synthesis of 6.41 using dihydrocoumarain	111
Scheme 6.4	Synthesis of the methylated compound	112
Scheme 6.5	Synthesis of analogues which incorporate an ether	113
Scheme 6.6	Synthesis of proline analogues with both the aromatic and aliphatic groups	120
Scheme 6.7	Synthesis of phenylalanine analogue	122
Scheme 6.8	Synthesis of tryptophan analogue	122
Scheme 6.9	Synthetic strategy used in the synthesis of the library	124

List of Tables

Table 1.1	Lipinski's 'Rule of Five' parameters	9
Table 2.1	Primary amino acid sequence of GIP and GLP-1	21
Table 3.1	Synthesis map used for synthesizing library compounds	54

Table 5.1	Current drugs on the market or in development for the treatment of influenza	77
Table 5.2	50% Inhibitory concentrations and selective indexes of anti-NS1 compounds on different influenza strains cultured in MDCK UK cells	85
Table 6.1	Log ₁₀ TCID ₅₀ values for synthesized compounds 6.2-6.7 and 6.11-6.19	98
Table 6.2	TCID ₅₀ values of compounds with additional functionality on the ring	110
Table 6.3	TCID ₅₀ values for compounds 6.56-6.58	118
Table 6.4	Library monomers used for the synthesis of the library. A '+' indicates an increase in activity when compared to 6.65 . A '-' indicates no increase in activity was observed. Blank boxes indicate the assay is ongoing.	126

Chapter 1:
Introduction to Library Synthesis

Pharmaceutical companies constantly need to develop new therapeutics and therapeutic technologies to stay competitive. Traditionally, there are three methods utilized in the process of drug discovery; natural product synthesis, computational analysis coupled with discrete synthesis, and the synthesis of chemical libraries. Each of these methods takes a completely different approach to discovering lead compounds and the way in which they are developed further. Although each of these methods can be effective, library synthesis has the ability to streamline the process of collecting and analyzing large amounts of data.

Natural Products

Natural product isolation and synthesis is one avenue currently utilized by the pharmaceutical industry for drug discovery. Often, there is therapeutic potential found in the extracts of certain plants or fungi. Although natural products can provide unique structural elements, there are some disadvantages to using this as a sole means to drug discovery.^{1,2} One of the problems associated with this style of drug discovery is the fact that natural products are often large and very complex molecules. Due to the complexity of these molecules, the synthesis is often not economically feasible. As an example, the natural product Taxol[®] (**Figure 1.1**) exhibits antitumor and antileukemic activity.³ One of the biggest obstacles to overcome is the scarcity of natural Taxol available. In one example, out of 1200Kg of western Yew bark obtained, only 10g of purified Taxol had been isolated.⁴

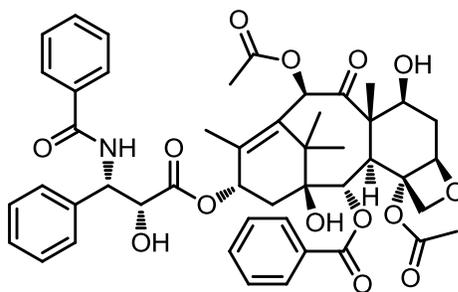


Figure 1.1 – The structure of Taxol

Taxol is isolated from the bark of the western yew tree *Taxus brevifolia*.³ Upon close inspection of the molecule, it is clearly complex, containing 11 stereocenters and multiple rings. During total synthesis, each of these stereocenters needs to be set (either through commercially available starting materials, or through stereospecific chemical reactions). One reported synthesis takes place in over 36 steps to yield the final product. Using a synthetic route with this many steps makes production difficult in sufficient quantities to be a practical drug candidate.^{5,6} Although Taxol is not commercially produced using total synthesis, this example illustrates the challenges to using total synthesis of natural products as a method of drug discovery.⁷ Also, each pharmaceutically desirable natural product is not as synthetically challenging as Taxol. However, this identifies a difficult problem facing the pharmaceutical industry: how can the process of drug discovery be made more efficient? With emerging technologies, as well as the development of high throughput screening, pharmaceutical companies chose to change their strategy in the hopes of gaining a competitive advantage.

Computational Methods

Computational methods can provide useful information about potential drug/target interactions. The types of analyses used for these studies can include virtual screening of compounds based on protein-ligand docking models and 3-D pharmacophore queries for lead compound identification.⁸ Traditionally, the compounds which are predicted to perform the best are selected for discrete synthesis. Occasionally, these predictions can be based on parameters which are predetermined, preprogrammed, or based on previously reported data. Although this can be a powerful tool, computers cannot always predict how the body will respond when foreign entities are introduced to it. As a result, potentially active compounds may be overlooked due to unfavorable computational outcomes. Additionally, the structures of the molecules are often built without regard to potential synthetic pathways. This can lead to complicated syntheses which can be inefficient.

More recently, computational methods have been developed to help facilitate library synthesis and exploration. Traditionally, structure searches have been used to compare molecules on a structure by structure basis. Recently, Fitzgerald and co-workers have developed the *Diversity Space* methodology which allows for comparisons to be made at the library level.^{9,10} It has been demonstrated by their research, that a total of 11 libraries can be synthesized to cover 50% of the diversity space of 698 previously synthesized libraries, indicating a repetition in the coverage space of libraries.¹⁰ Technology like *Diversity Space* has the ability to be used to facilitate designing libraries which can explore novel chemical space and potentially produce new leads in drug discovery.

Library Design and Synthesis

Library synthesis was introduced as a method to expedite the drug discovery process by gathering data faster and more comprehensively. This was accomplished through the use of large sets of chemical entities (small molecules) which are synthesized in parallel. Producing large sets of compounds quickly is important as the predicted number of small molecules in the universe is estimated to be between 10^{14} and 10^{30} molecules.¹¹ Knowing this, the activity of a biological target can be thought of as a large search space of all the possible molecules in which a single active molecule must be found.¹¹ This synthetic strategy utilizes large numbers of library compounds or members to systematically cover large amounts of chemical space in order to increase the probability of finding an active molecule.

During library synthesis, it is important to maximize the diversity of generated compounds to explore every possible interaction between the molecule and the chosen biological target. Once a library has been synthesized and tested for activity against the target, one can determine which type of compound should be the focus of the second, more specific chemical library. This structure activity relationship (SAR) is a common tool used to further understand the biological target. Thus, this type of drug discovery is also useful when attempting to discover a new target.

The process of library synthesis has seen dramatic changes over the past several decades. With the introduction of solid phase peptide synthesis by Merrifield in 1963,¹² the foundation was set for the ensuing explosion of library syntheses. In his seminal publication, Merrifield used insoluble polymer beads to synthesize a tetrapeptide.¹² Just

over 20 years later, H. Mario Geysen used the idea of a solid support to synthesize a peptide library using polyethylene rods on which linear chains of polymeric acrylic acid was grafted onto the surface using radiation.¹³ It is from this point that combinatorial chemistry began to grow as a distinct discipline and pharmaceutical companies started using library synthesis as a method of drug discovery.

As with the previously mentioned methods for drug discovery, there are both advantages and disadvantages to using library synthesis, depending on what method of synthesis is selected. Before synthesis begins, the appropriate resin must be selected. “Resin” typically refers to an insoluble polymer bead on which the chemistry will be performed. In some cases, these beads are made of chemically-derivatized polystyrene to allow for alternative functionalities to be attached to them. Most often the resin is selected based on two criteria, the chemistry which needs to be performed and the desired functional group that results upon cleaving from the resin. Each reaction is typically run in the presence of a large excess of reagents to force the reaction to go to completion in a relatively short time period. Furthermore, once the reaction is complete, all of the excess reagents and soluble side products can be removed easily by filtration, leaving a relatively pure product still attached to the resin.

Despite these distinct advantages, there are a couple disadvantages which must be taken into account. It is a common goal of any library to introduce a variety of functional groups to try and answer questions. However, the synthetic strategy must be generalized to accommodate a variety of building blocks. An added complication is the cleavage strategy of the final product from the solid support. Accordingly, the acceptable conditions for potential chemical reactions can become restricted to avoid premature

cleavage. Because of these restrictions placed on the chemistry, building blocks with an appropriate protection scheme must be used, thus increasing the number of steps and overall expense.

When using solution-phase methods, there are work-up steps (which include washing), purification steps, and analysis after each reaction. Accordingly, the time required to synthesize a library is exponentially greater in comparison to solid-supported strategies. With all of these extra steps, the opportunity to make a mistake greatly increases. Parallel reactions carried out on solid phase using a 96-well synthesis block are much easier to handle than 96 individual flasks. One advantage to using solution phase is the relative flexibility of the synthesis. With solid phase synthesis, synthetic options may be limited based upon final cleavage strategies. These types of limitations are not applicable to solution-phase chemistry.

When approaching library design, there are many important decisions which need to be made. One of the first decisions is the selection of the molecular scaffold. It is upon this skeletal structure that a diversity of functionality will be displayed. After making this decision, the next considerations are generally made with respect to one another; monomer selection and synthetic strategy. In general, the synthetic strategy is developed around what the desired monomers are. However, based on factors such as commercial availability and compatibility, the list of possible monomers can be narrowed.

Monomer/Scaffold Selection

A scaffold is a part of the molecule which remains the constant while the decorations or points of diversity, are changed. There are a number of different methods that can be used to select an appropriate scaffold. One such method is to use standard computational tools to identify a potentially active compound. Once a scaffold has been determined, monomers can be selected to append to the scaffold so as to display the desired functionalities. Although this can provide insight into the active site, as previously mentioned, computational tools are best used in tandem with other methods. There is no accurate way to predict how successful the compound will be in reality.

Another method commonly used is scaffold hopping. This process can be defined as the act of changing the core of a known active compound while retaining the key functionalities, to discover a structurally novel set of compounds.¹⁴ The basis for this methodology is to change small portions of a biologically active molecule enough to consider the new molecule structurally different, but still retain the essential features needed for activity. There are a number of different reasons one might consider scaffold hopping. One reason may be to substitute a metabolically unstable scaffold with one that is more resilient so as to improve upon the existing pharmacokinetic profile of the molecule. Changing the physical properties of molecules, such as replacing lipophilic side chains with more polar moieties can increase drug solubility. Furthermore, changes to the scaffold can help to avoid infringing upon established intellectual property.¹⁴

An example of scaffold hopping can be seen in the sequence of marketed drugs used for selective serotonin reuptake inhibitors (**Figure 1.2**). Each of the structures has an electron deficient aromatic ring with a basic nitrogen five bonds away and an

Although these rules have been widely accepted as appropriate guidelines, it is important to note that it is a statistical model and therefore not every drug molecule will follow these rules. This set of criteria was determined based on studies of drugs already on the market. If the aim of a program is to develop libraries to produce bioavailable, ‘drug-like’ molecules, then these parameters may be appropriate. However, if the overriding goal of the program is to create a library to find a lead compound which can be used for the basis to start a program, then these criteria will be limiting.¹⁶ It is with the rule of five that the definition of what makes up a drugable molecule has been narrowed. Furthermore, it makes the distinction that there is an overlapping difference in chemical space which makes up leadlike compounds versus druglike compounds (**Figure 1.3**).¹⁶

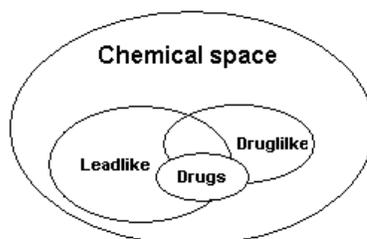


Figure 1.3 – Chemistry space as it relates to the drug discovery process (Modified from Opera)¹⁶

After Lipinski published his seminal paper on the ‘Rule of Five’ in 1997, other research groups have attempted to modify or add supplemental parameters to aid in library design. Veber and colleagues compiled a database of approximately 1100 drug candidates which were profiled at SmithKline Beecham. Through their studies they were able to conclude that another parameter which should be applied during the drug discovery process was the number of rotatable bonds. Reducing this number had the largest impact on bioavailability.¹⁷

In a study conducted by Vieth and coworkers, a total of 1729 compounds taken from marketed drugs, clinical trials, and related biologically active molecules were examined. The main focus of their research was to clarify the properties between molecules with favorable and unfavorable pharmacokinetic profiles. It was found that injectable drugs generally do not fall into the rule of five criteria, as expected. These drugs tend to have higher molecular weight, increased number of rings, hydrogen bond acceptors/donors, rotatable bonds, and lower ClogP values. Based on Lipinski's criteria and what was observed, this should lead to a drug which is not orally bioavailable. Furthermore, Vieth concluded that the core of the molecule might have the largest influence in the likelihood of obtaining acceptable levels of oral bioavailability.¹⁸

As mentioned previously, the 'Rule of Five' parameters, as determined by Lipinski, must be used within an appropriate context (i.e. optimization towards a druglike compound versus lead discovery). Applying the rule of five during exploratory library design essentially defeats the purpose of synthesizing the library. With any combinatorial library, once the scaffold has been selected, it is the aim to use as much diversity as possible. By doing so, the most possible information can be gathered about potential interactions between compounds and the desired target. By applying these filters too early, a number of monomers which might have been selected will be prematurely eliminated due to their potentially unfavorable characteristics. It is advantageous, therefore, to apply these filters after an active lead molecule has been identified. In this way, one is not limited in the monomers available for synthesis, but has the ability to use them to shape a molecule towards something more attractive to a drug discovery program. In summary, it is important to remember that the 'Rule of Five'

and similar filters are only general guidelines for developing orally available compounds and should be treated as such. Following these restrictions can narrow the focus of the medicinal chemist and does not allow for exploration in other parts of chemical space, which might be significantly relevant.

Within this current work, the tenants of library synthesis were used to discover possible pharmaceutical therapies for two prominent health concerns. In each instance, a different approach was taken to the traditional method of drug discovery. In the first project, the issue of Type 2 Diabetes Mellitus (T2DM) was examined. Of all the different treatment methods proposed for type 2 diabetes, the small molecule route is the best fit for library synthesis. Currently, there are two small molecules marketed to treat T2DM (Sitagliptin and Saxagliptin), with several more in clinical trials (**Figure 1.4**). Two of these compounds, which fall into the class of compounds known as cyanopyrrolidines (Saxagliptin and Vildagliptin), have a large amount of steric bulk substituted on the nitrogen or at the α -carbon.

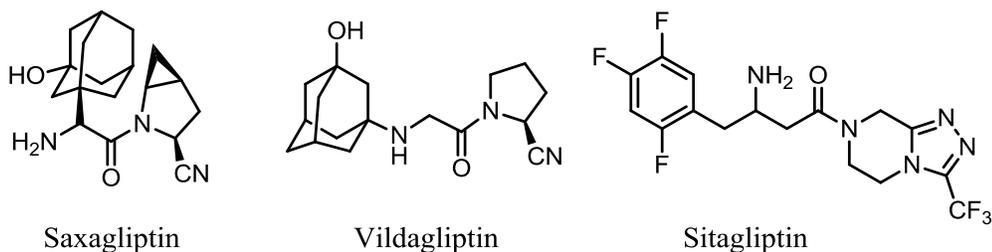


Figure 1.4 – Compounds used to treat Type 2 Diabetes

Some of these active compounds are the result of successful application of small molecule library development and screening.¹⁹ Despite these studies, further emphases could be placed on the optimum steric bulk substitution pattern. Through the use of

library design and synthesis, a relatively small set of compounds was synthesized to develop a greater structure activity relationship (SAR) with regards to the addition of steric bulk.

A second project discussed herein, which represents the primary focus of my efforts, is the development of novel Influenza A Nonstructural Protein 1 (NS1) inhibitors. This target was recently identified as a novel target to treat against infection by the influenza A virus. The actions of NS1 are well documented, but only now are steps being taken to advance influenza research through the use of small molecule inhibitors of this protein. In this instance, a library from the National Cancer Institute Diversity Set (NSC) diversity set was screened in which four compounds were identified as active at this protein (**Figure 1.5**).²⁰

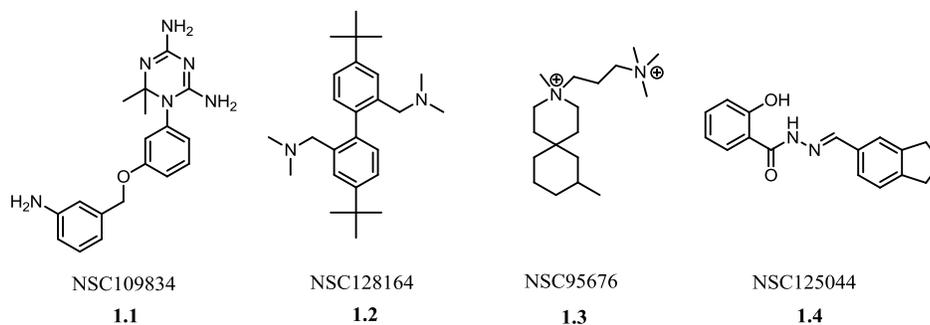


Figure 1.5 – Small molecules which have been identified as inhibitors of NS1

From the initial identification of these four molecules, **1.4** was selected to be the lead compound for the development of a SAR study. The overall goal was to develop a compound which is active against a variety of strains of the influenza A virus. Once a compound was identified with potential broad spectrum activity, a 96-compound library

was synthesized to further probe the essential interactions required for broad spectrum inhibition.

Chapter 1 References

1. Bindseil, K. U.; Jakupovic, J.; Wolf, D.; Lavayre, J.; Leboul, J.; van der Pyl, D., Pure compound libraries; a new perspective for natural product based drug discovery. *Drug Discovery Today* **2001**, *6* (16), 840-847.
2. Breinbauer, R.; Vetter, I. R.; Waldmann, H., From Protein Domains to Drug Candidates—Natural Products as Guiding Principles in the Design and Synthesis of Compound Libraries. *Angewandte Chemie International Edition* **2002**, *41* (16), 2878-2890.
3. Wani, M. C.; Taylor, H. L.; Wall, M. E.; Coggon, P.; McPhail, A. T., Plant antitumor agents. VI. Isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. *Journal of the American Chemical Society* **1971**, *93* (9), 2325-2327.
4. Goodman, J., Walsh, V., *The Story of Taxol: Nature and Politics in the Pursuit of an Anti-Cancer Drug*. Cambridge University Press: 2001; p 285.
5. Morihira, K.; Hara, R.; Kawahara, S.; Nishimori, T.; Nakamura, N.; Kusama, H.; Kuwajima, I., Enantioselective Total Synthesis of Taxol. *Journal of the American Chemical Society* **1998**, *120* (49), 12980-12981.
6. Doi, T.; Fuse, S.; Miyamoto, S.; Nakai, K.; Sasuga, D.; Takahashi, T., A Formal Total Synthesis of Taxol Aided by an Automated Synthesizer. *Chemistry – An Asian Journal* **2006**, *1* (3), 370-383.
7. Zhong, J.-J., Plant cell culture for production of paclitaxel and other taxanes. *Journal of Bioscience and Bioengineering* **2002**, *94* (6), 591-599.

8. Blake, J. F., Integrating cheminformatic analysis in combinatorial chemistry. *Current Opinion in Chemical Biology* **2004**, 8 (4), 407-411.
9. Fitzgerald, S. H.; Sabat, M.; Geysen, H. M., Diversity Space and Its Application to Library Selection and Design. *Journal of Chemical Information and Modeling* **2006**, 46 (4), 1588-1597.
10. Fitzgerald, S. H.; Sabat, M.; Geysen, H. M., Survey of the Diversity Space Coverage of Reported Combinatorial Libraries. *Journal of Combinatorial Chemistry* **2007**, 9 (4), 724-734.
11. Geysen, H. M.; Schoenen, F.; Wagner, D.; Wagner, R., Combinatorial compound libraries for drug discovery: an ongoing challenge. *Nat Rev Drug Discov* **2003**, 2 (3), 222-230.
12. Merrifield, R. B., Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *Journal of the American Chemical Society* **1963**, 85 (14), 2149-2154.
13. Geysen, H. M.; Meloen, R. H.; Barteling, S. J., Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid. *Proceedings of the National Academy of Sciences* **1984**, 81 (13), 3998-4002.
14. Böhm, H.-J.; Flohr, A.; Stahl, M., Scaffold hopping. *Drug Discovery Today: Technologies* **2004**, 1 (3), 217-224.
15. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews* **2001**, 46 (1-3), 3-26.

16. Oprea, T. I.; Davis, A. M.; Teague, S. J.; Leeson, P. D., Is There a Difference between Leads and Drugs? A Historical Perspective. *Journal of Chemical Information and Computer Sciences* **2001**, *41* (5), 1308-1315.
17. Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D., Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *Journal of Medicinal Chemistry* **2002**, *45* (12), 2615-2623.
18. Vieth, M.; Siegel, M. G.; Higgs, R. E.; Watson, I. A.; Robertson, D. H.; Savin, K. A.; Durst, G. L.; Hipskind, P. A., Characteristic Physical Properties and Structural Fragments of Marketed Oral Drugs. *Journal of Medicinal Chemistry* **2003**, *47* (1), 224-232.
19. Villhauer, E. B.; Brinkman, J. A.; Naderi, G. B.; Burkey, B. F.; Dunning, B. E.; Prasad, K.; Mangold, B. L.; Russell, M. E.; Hughes, T. E., 1-[[[3-Hydroxy-1-adamantyl)amino]acetyl]-2-cyano-(S)-pyrrolidine: A Potent, Selective, and Orally Bioavailable Dipeptidyl Peptidase IV Inhibitor with Antihyperglycemic Properties. *Journal of Medicinal Chemistry* **2003**, *46* (13), 2774-2789.
20. Basu, D.; Walkiewicz, M. P.; Frieman, M.; Baric, R. S.; Auble, D. T.; Engel, D. A., Novel Influenza Virus NS1 Antagonists Block Replication and Restore Innate Immune Function. *J. Virol.* **2009**, *83* (4), 1881-1891.

Chapter 2:
Introduction to DPP-IV

One of the largest health issues today is the continued rise in the number of diabetes cases. In 2000, it was estimated that more than 11 million Americans were diagnosed with either Type 1 or Type 2 diabetes mellitus.¹ It has been estimated that the number of individuals which are considered to be high risk will rise to approximately 66 million by the year 2031 and the number of diagnosed cases is estimated to be 19 million.² The approximate health care cost attributed to diabetes in 2007 was \$174 billion. Contained within this number is \$27 billion attributed to medical costs directly related to treatment, \$58 billion to treat chronic complications that are related to diabetes, and \$31 billion in excess general medical costs. Individuals who have been diagnosed with diabetes pay on average \$11,744 annually to deal with this problem.³

There are two different major forms of diabetes: Type 1 and Type 2. The former is characterized by insufficient production of insulin. Most of the Type 1 cases are due to the T-cell mediated destruction of pancreatic β -cells.⁴ It is believed that genetic and environmental factors contribute to the onset of Type 1 diabetes. Exact environmental stimuli which initiate autoimmune destruction of the β -cells are unknown. However, individuals who are considered at risk can be tested for diabetes associated autoantibodies, genetic markers and intravenous glucose tolerance testing to help identify themselves as Type 1 diabetics.^{5,6} Type 2 Diabetes (T2DM) can be classified by a number of different characteristics which include increased resistance to insulin in both muscle and adipose tissue, decreased insulin production by pancreatic β -cells, increased production of glucose by the liver, and decreased levels of incretins, which are gastrointestinal hormones normally secreted following the ingestion of food.⁷

Most of the newly diagnosed cases of diabetes in the United States are of type 2 diabetes (T2DM).⁸ T2DM is known as late-onset, or adult diabetes due to the disease being developed over time. Factors leading to the development of the disease include obesity and lack of physical activity. Aside from increases in morbidity and mortality, T2DM has been associated with eye, kidney and nerve disease. Furthermore, this debilitating disease promotes heart disease.⁹

Cells use glucose as one of the major sources of energy. In order to maintain acceptable blood sugar levels, the body releases a number of different hormones to aid in regulation. Among these hormones are insulin and glucagon. After the ingestion of food, pancreatic β -cells release insulin into the blood stream. This release stimulates glucose uptake from the blood. One method of removing glucose from the blood stream involves the GLUT4 protein. In muscle cells and adipose tissue, GLUT4 is contained within intracellular vesicles. Upon activation by insulin at the receptor, the GLUT4 protein is transported to the cell surface, which then facilitates the diffusion of glucose across the membrane.¹⁰ Once glucose enters the cells, it is used in the synthesis of glycogen and triacylglycerols.¹¹

Conversely, when blood glucose levels fall below normal, possibly due to exercise or fasting, the α -cells of the islets of Langerhans in the pancreas release another regulatory hormone, glucagon. Upon reaching the active site in the liver, glucagon stimulates the release of glucose through glycogenolysis.¹¹ This process continues until other hormones such as insulin trigger the shut-down of the active site.

There are a number of gastrointestinal hormones that also aid in the regulation of glucose in the body called incretins. The two most important of these hormones are the

Gastric Inhibitory Polypeptide (GIP) and Glucagon-Like Peptide 1 (GLP-1). There is supporting evidence that these two hormones are responsible for 50-70% of secreted insulin after oral glucose administration (**Table 2.1**).^{12,13} In addition to this, it was shown that oral glucose administration is associated with higher levels of plasma insulin than when administered intravenously. This observation has been labeled as the *incretin effect*.¹⁴

Table 2.1 – Primary amino acid sequence of GIP and GLP-1.¹³

Peptide	Sequence
GIP	YAEGTFISDYSIAMDKIHQQDFVNWLLAQKGGKNDWKHNITQ
GLP-1 _[7-36NH2]	HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR _{NH2}

GIP is a 42-amino acid peptide that exhibits a wide array of functions within the human body. Originally, it had been found that GIP was involved in the inhibition of gastric acid secretion and gastrointestinal (GI) motility in dogs. Further studies have shown the ability of GIP to exhibit glucose-dependent stimulatory effects on insulin secretion.¹⁵ A study conducted by Ross and Dupre showed that insulin secretion due to effects of GIP are glucose dependent.¹⁶ In addition to these effects, GIP was found to enhance insulin-stimulated incorporation of fatty acids into triglycerides, stimulate lipoprotein lipase activity, modulate fatty acid synthesis, and promote β -cell proliferation and cell survival.^{15,17,18}

The second major incretin important for insulin secretion is Glucagon-Like Peptide-1 (GLP-1). GLP-1 was initially discovered by Bell and coworkers during a cloning and characterization study of the proglucagon gene.¹⁹ Following this discovery, a

vast amount of research was undertaken to determine the function of this incretin. The results of these studies showed the ability of GLP-1 to stimulate glucose-dependent insulin secretion, suppress glucagon secretion, act as a satiety factor leading to reduced appetite and delay gastric emptying.²⁰

In normal circumstances, GIP and GLP-1 enter circulation in the blood stream shortly after the ingestion of food; GIP from K-cells and GLP-1 from L-cells in the intestine.²⁰⁻²¹ Upon reaching their respective active sites in the β -cells of the pancreas, the incretin hormones stimulate the production of insulin. It is now the understanding that in patients with T2DM, there is diminished incretin effect. Logically, some of the first types of treatments would center on the synthesis and production of these hormones. In 1993, a study by Nauk and coworkers compared the insulinotropic and glucagon lowering actions of GIP and GLP-1 in nine normal patients versus nine subjects with T2DM. When analyzing their data, they concluded that the diminished incretin effect was due to the decreased insulinotropic effect of GIP. It was their contention that because of these effects, GLP-1 analogues could be used to aid in lowering blood glucose levels.²²

Due to the continued ability of GLP-1 to stimulate insulin release, it became an attractive avenue for drug discovery. It has been hypothesized by numerous research groups that finding an alternative way to aid in the stimulation of insulin secretion could be an effective treatment for T2DM. However, there is one major obstacle to creating an effective means of treating T2DM; Dipeptidyl peptidase IV (DPP-IV). Once in circulation, both incretins are rapidly metabolized by the serine protease DPP-IV.^{23,24,25}

Because of this rapid degradation, the circulating concentration of each of these hormones is decreased, which results in a lower stimulation of the production of insulin.

DPP-IV

Dipeptidyl Peptidase IV, also known as CD26, is a membrane bound serine protease which also has a form that is soluble in plasma. It is a 766-amino acid residue protein which can be found in a wide array of tissues including vascular endothelial cells and immune-related cells (as a marker for activated T-cells).²⁶ DPP-IV appears in two distinct forms. First, as a membrane bound enzyme, it activates intracellular signaling associated with adenosine deaminase. Next, as both a membrane bound enzyme and in plasma, it has enzymatic activity related to the cleavage of substrates with proline in the P1 position and is tolerant to a wide variety of substitutions in P2.^{26,27}

Popular substrates for DPP-IV include the incretin hormones GLP-1 and GIP. Both of these hormones have their N-termini cleaved at the second residue rendering them inactive. For GIP, the residues cleaved are NH₂-GA and for GLP-1, the two are NH₂-HA. The major problem associated with the metabolism of these substrates is rate at which it happens. Once in circulation, GLP-1 is completely metabolized in less than five minutes.^{23,25} Even when GLP-1 is administered as an intravenous drug, it is still subject to the rapid degradation by DPP-IV. To complicate matters further, some recent studies demonstrated the ability of the GLP-1 metabolite to act as a possible inhibitor at the GLP-1 receptor site.²⁸

Current Therapies

Due to the large number of cases of T2DM, with the probability of more in the future, it becomes necessary to successfully combat this growing health threat. The most obvious mode of treatment is through the use of regular insulin injection therapy. Alternative methods are required, however, for patients who do not respond well to this type of treatment. Currently, there are two major therapeutic options under investigation. The first is incretin therapy, which involves treatment with incretins or incretin analogues. The second of these therapeutic options involves inhibition of the DPP-IV enzyme.

Incretin Therapy

Of the number of different ways to treat T2DM, one potentially efficient method is to use incretin mimetic analogues. There are two treatment strategies that are currently at the forefront of development. The first of these treatments is Exenatide[®] (Byetta[®]) (**Figure 2.1**). This 39-amino acid peptide was discovered in the saliva of the Gila Monster (*H. suspectum*) and was originally designated as exendin-4. Further studies showed that exendin-4 shares a 53% sequence homology with the incretin hormone GLP-1 and binds directly to the GLP-1 receptor (GLP-1R).^{29,30}



Figure 2.1 - The sequence of Byetta[®] with the proposed site of metabolism. This figure was taken from Drauker and Nauk.³¹

Early studies conducted by Young and coworkers showed the ability of exendin-4 to reduce plasma glucose levels in diabetic mice and monkeys.³² Moreover, the half-life of exendin-4 displayed a significant increase over that of GLP-1. Fifteen minutes after a single dose injection of Byetta[®], the drug was detectable in plasma and has an elimination half-life of 2.4 hours.^{33,34} Furthermore, the drug was still detectable in plasma samples 15 hours after injection. When compared to the 2-3 minute half-life of GLP-1, this is a dramatic increase.^{31,35} Placebo and comparator-controlled studies demonstrate that exenatide-4 improves glycemic control and reduced body weight in patients with T2DM.³⁰ Another positive discovery associated with Byetta[®] is the resistance to DPP-IV.

Although Byetta[®] has become the first incretin mimetic to be approved by the FDA as a viable treatment for T2DM, the major drawback to this is the poor oral bioavailability. Byetta[®] must be administered via injections. Initially injections were given twice a day, but now there have been reports of a weekly injection which use slow degrading polymers to release the drug in a controlled manner.³⁶

Alternatively, another GLP-1 mimetic was recently approved by the FDA, Liraglutide. This peptide drug has a 97% homology with GLP-1 and, like Byetta[®], is DPP-IV resistant.³⁷ The peptide portion of Liraglutide can be produced recombinantly in yeast. The sequence is identical to that of GLP-1 except for a substitution of an arginine for a lysine at position 34.³⁸ The most significant differences from most GLP-1 mimetic therapeutics is the inclusion of a fatty acid chain onto the lysine at position 26 through a glutamic acid spacer.³⁸ Studies have shown a retention of binding affinities at the GLP-1R despite these structural modifications (**Figure 2.2**).³⁹

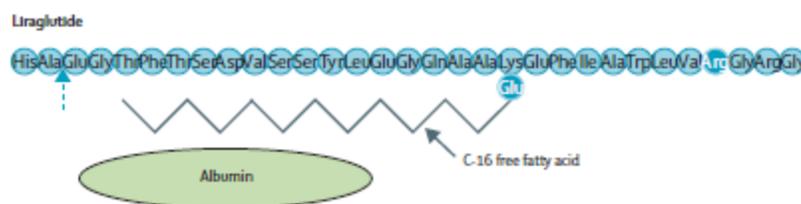


Figure 2.2 - The structure of Liraglutide. This figure was modified from Drucker and Nauk.³¹

The purpose of the C16 fatty acid chain is for the peptide to associate with albumin in the blood stream. This interaction allows for the peptide to travel longer in circulation without the effects of metabolism.³⁹ As a corollary to the extended circulation, Liraglutide has an improved half-life of about 13 hours.⁴⁰ This is a significant increase over the half-life of both endogenous GLP-1 and Byetta[®]. Despite this significant increase, the drug is still subject to poor oral bioavailability and must be injected. One of the few highlights is that Liraglutide need only be injected on a once-daily schedule as opposed to a twice-daily regimen.

DPP-IV Inhibitors

One of the biggest obstacles to overcome in incretin-based approaches to treating T2DM is the rapid degradation of GLP-1 by the enzyme DPP-IV. As an alternative to incretin mimetic therapy is the use of DPP-IV inhibitors. Recent studies demonstrated the ability of DPP-IV inhibitors to successfully decrease blood glucose levels. Recently, Ahrén and coworkers reported on a 12-week, randomized, double-blind, placebo control study testing the ability of DPP-IV inhibitors to lower blood glucose levels. The initial study lasted 12 weeks with one group of subjects receiving the DPP-IV inhibitor (n=56)

and the other receiving a placebo (n=51). The test was extended for an additional 40 weeks with those from the initial studies who wished to continue.⁴¹ From their research, they were able to show that patients who were on the DPP-IV inhibitor displayed an average decrease in HbA_{1C} (HbA_{1C} is the measure of plasma glucose through the measure of glycated hemoglobin) when compared to the placebo control (**Figure 2.3**). This was further confirmed when the study was extended for an extra 40 weeks.^{41,42}

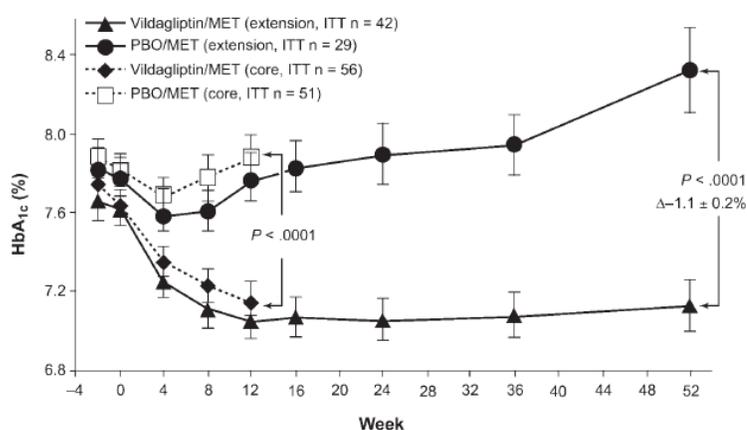


Figure 2.3 - Study showing the decrease in HbA_{1C} percentage using DPP-IV inhibitor (Vildagliptin) vs. Placebo (PBO). Figure taken from Campbell et al.⁴²

Currently, there are two FDA approved DPP-IV inhibitors on the market of T2DM. The first drug to gain approval was Sitagliptin (marketed as Januvia®). Sitagliptin (**Figure 2.4**) was put on the market in 2006 by Merck and Co. and is an orally available, potent, and selective DPP-IV inhibitor.⁴³ The approval of this drug is a large step forward in the treatment of T2DM. One of the biggest improvements is the ability to take the drug orally. Additionally, studies have shown that through dosing, there were improved levels of circulating GLP-1 and improved β -cell function.⁴⁴

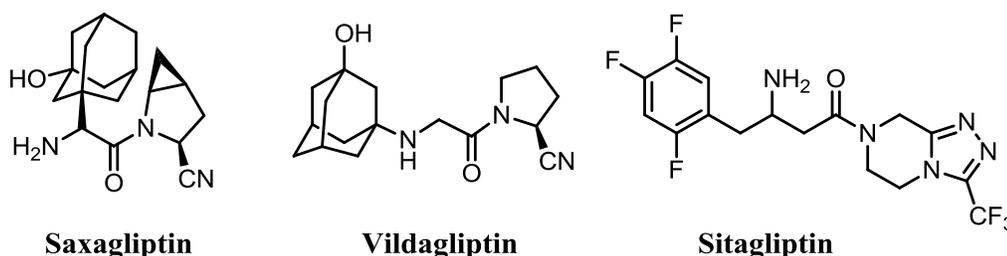


Figure 2.4 - Structures of DPP-IV inhibitors

The second DPP-IV inhibitor approved by the FDA is Saxagliptin (produced by Bristol-Myers Squibb and marketed as Onglyza[®]). This compound was given approval in 2009 as an effective treatment of T2DM. Saxagliptin has shown to be a potent, reversible, and selective inhibitor of DPP-IV. In studies, Saxagliptin has proved to be more potent than Sitagliptin. Although no direct comparison has been made, the trough concentrations were compared and it was shown that a 5-mg dose of Saxagliptin gave the equivalent efficiency to a 100mg dose of Sitagliptin.⁴⁵ Furthermore, it significantly improves the mean HbA_{1C} levels relative to the placebo in patients.⁴⁶ This drug falls into a class of compounds known as cyanopyrrolidines.

Cyanopyrrolidines have become of significant interest in the treatment of T2DM due to the interesting actions of these compounds in the binding pocket of DPP-IV. When DPP-IV acts on a substrate, it cleaves dipeptides that have a proline in the P1 position. The nitrile of the cyanopyrrolidine is higher binding affinities with DPP-IV. X-ray crystal structures have confirmed this by showing a strong, covalent-like interaction between the nitrile and Ser630 Of DPP-IV.⁴⁷ In this interaction, the nitrile loses its linear sp character and adapts a more imine like structure (**Figure 2.5**). The oxygen of Ser630 interacts with the nitrile carbon and bends the functional group so that the amine is now

participating in hydrogen bonding interactions with Tyr547.⁴⁸ This binding motif is a reversible interaction and is partially responsible for the slower dissociation kinetics.

Another key interaction in the binding of a substrate to DPP-IV involves the hydrogen bonding network, which includes the three residues Glu205, Glu206, and Tyr662. These residues interact with the primary or secondary amine commonly found on many DPP-IV inhibitors and mimics the peptide N-terminus recognized by DPP-IV.⁴⁸ It was shown that without the ability to create this hydrogen bonding system, there is a significant decrease in the ability of the compound to inhibit DPP-IV.⁴⁹ It can be concluded that the success of compounds such as Saxagliptin and Sitagliptin to inhibit DPP-IV, is primarily attributed to making these significant interactions in the binding pocket.

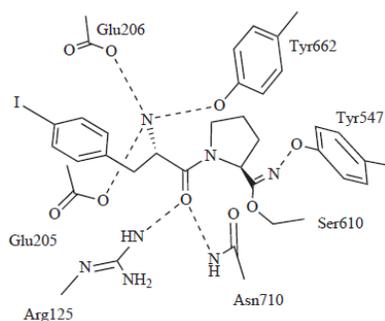


Figure 2.5 - Figure demonstrating the binding interactions within the DPP-IV binding pocket.⁴⁸

In addition to Saxagliptin, there is another cyanopyrrolidine which has been under investigation for some time now. In 2003, Villhauer and coworkers published a structure activity relationship (SAR) study regarding cyanopyrrolidines and their corresponding activities. In their study, they examined the necessity of steric bulk in the P2 region of

the binding site and briefly examined the connection between having the nitrile functionality and the improved half-life of the compounds. At the conclusion of their study, it was determined that 1-[[[3-Hydroxy-1-adamantyl)amino]acetyl]-2-cyano-(S)-pyrrolidine (also known as Vildagliptin) was a possible candidate to treat T2DM.⁵⁰ Vildagliptin was approved for the market in Europe; however, it was taken out of FDA clinical trials before release in the United States.

Previous SAR studies have been performed in the development of Vildagliptin by Villhauer and co-workers. A number of different scaffolds were used in their study (**Figure 2.6**), where three of the scaffolds use similar synthetic strategies. From this, a small library was synthesized to produce a SAR. A majority of the modifications made in this study surrounded substitutions made at the N-terminus; however, there were a select few compounds synthesized which explore the link between the P1 and P2 subunits.

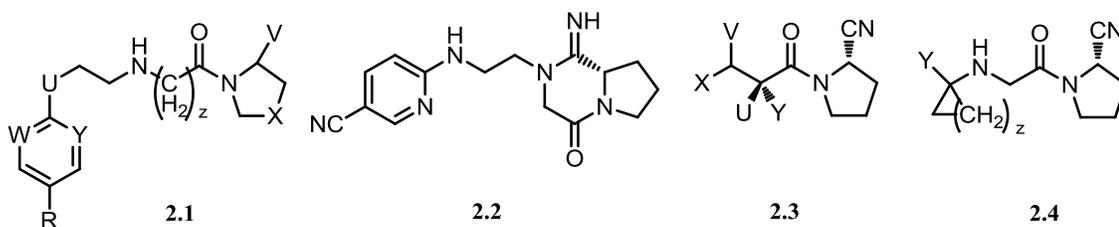
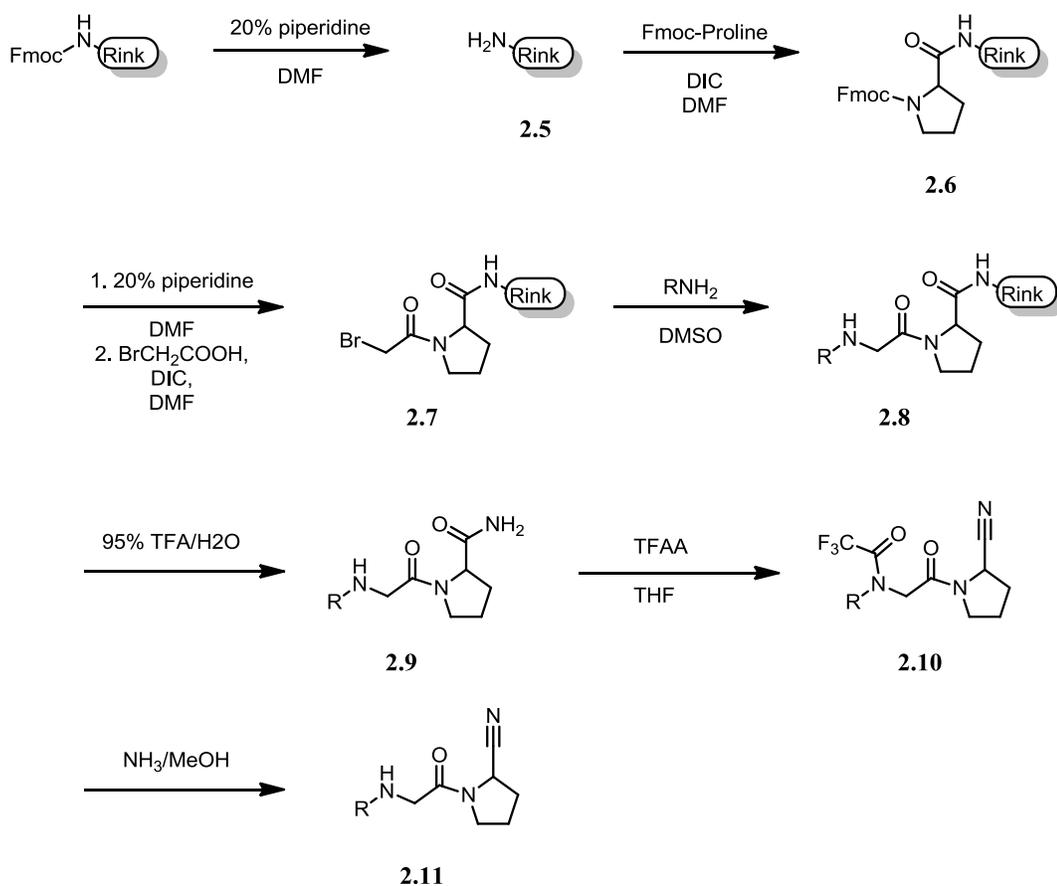


Figure 2.6 – Scaffolds used in previous SAR studies conducted by Villhauer⁵⁰

There are two different possible styles of synthesis which can be employed to construct these molecules: a solid phase approach and a solution phase approach. In the solid phase approach, Villhauer employed rink resin to synthesize the molecular library,

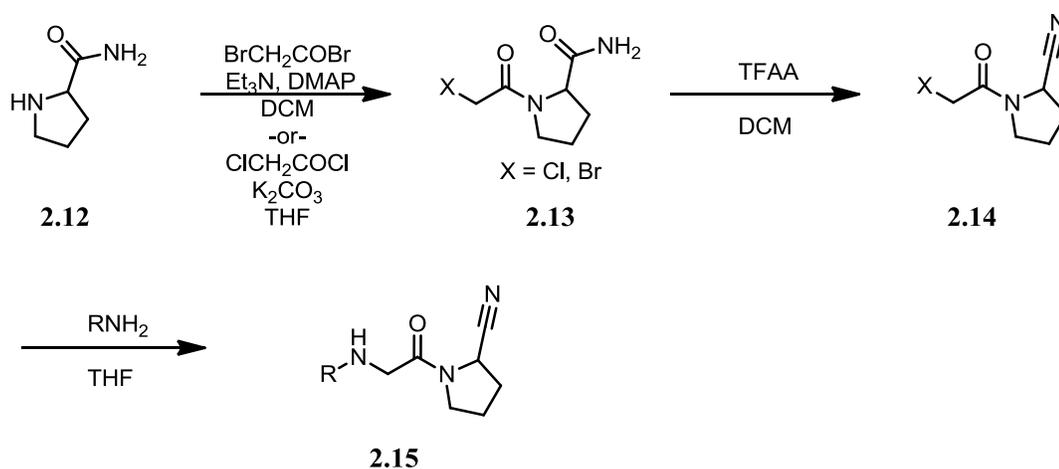
which upon cleavage, yielded precursors to the dehydration reaction to form the nitrile (**Scheme 2.1**).^{51,50} Following cleavage from the resin, the amide was converted to the nitrile using TFAA. Using this method, the issue of acetylating the amine was cause for concern. However, deacetylation of the amine is a relatively straight forward process using ammonia in methanol. This afforded the desired cyanopyrrolidine in relatively good yields.

Scheme 2.1 – Synthesis of compounds utilizing a solid phase approach

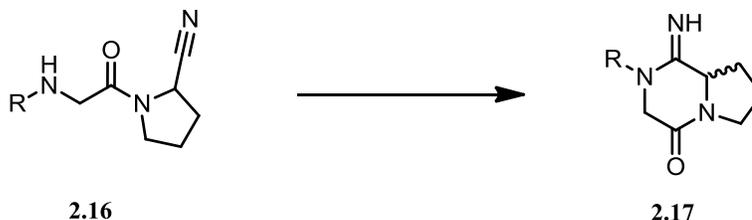


A possible synthetic strategy used to help avoid acetylation of the secondary amine is to perform the synthesis using solution phase protocols (**Scheme 2.2**). A synthetic pathway proposed by Villhauer started with coupling chloroacetyl chloride to pyrroliding-2-carboxamide. It is at this point in the synthetic protocol, that the amide was converted to the corresponding nitrile. By performing the S_N2 reaction with the amine after the dehydration, there is no way in which the amine could be acylated.

Scheme 2.2 – Synthetic strategy for solution phase synthesis of desired products



Studies have also shown that these cyanopyrrolidine compounds are relatively stable. It was established that if the compound is left to sit in a buffered solution at pH ~ 7.4, it will convert to the cyclic imidate with a $t_{1/2}$ of approximately 40h to >70 days (**Scheme 2.3**). The amine in the P2 region of the binding pocket can perform a nucleophilic attack at the electrophilic carbon of the nitrile.⁵² Further testing of this side product for biological activity showed that it was less active than its non-cyclic counterpart.

Scheme 2.3 – Formation of the cyclic imidate

From this study, the most active compounds were Vildagliptin and its non-hydroxylated analogue ($IC_{50} = 3.0 \pm 2$ nM and 3.5 ± 1.5 nM respectively). This study places a large, hydrophobic, tricyclic ring system at the N-terminus. This basic structure adheres to what has been determined to be significant for binding; having a nitrile to interact with Ser630 and a secondary amine which forms the salt bridge between Glu205/Glu206 in the binding pocket.

Out of all of the synthesized analogues in this study, only two were synthesized with substitution at both the α -carbon and the N-terminus. One of these analogues has a methyl group on the N-terminus and a cyclohexyl ring on the α -carbon. By doing this, the activity of the molecule decreased. The other molecule simply puts a cyclohexyl group at each position, which also decreases the activity at DPP-IV.

Aside from these two molecules, steric bulk was added at either location, but not at both together. Additionally, the size restriction was not fully examined at the α -carbon. However, at the N-terminus, the size restrictions and functionality were examined more extensively. In the initial studies, an increase in activity was seen by the addition of steric bulk through growing ring sizes. When compared to their linear, aliphatic counterparts, there was only a modest decrease in observed activity. Knowing

this, a series of cyclic, bicyclic, or tricyclic analogues were tested for their corresponding activity.

From these studies, a number of different conclusions can be drawn. First, the nitrile is essential for binding. Removal of the nitrile and replacement with a hydrogen displayed a significant decrease in binding with DPP-IV. Additionally, steric bulk added at either the N-terminus, or the α -carbon translated into an increase in inhibitory activity. When the amine was removed, the activity sharply decreased, indicating that the salt bridge created between Glu205 and Glu206 is essential for activity.

DPP-IV Inhibitors vs. Incretin Mimetic Therapy

Up to this point, two forms of treatment of T2DM have been presented. Although they are both capable in assisting the treatment of the disease, there are contrasting properties between the two. The biggest difference between the two styles of treatment is oral bioavailability. Due to the nature of incretin mimetics, these types of drugs must be given through injection. Conversely, DPP-IV inhibitors are able to be taken orally, which is generally preferred by most patients. Another major difference between the two therapies relates to levels of GLP-1. With the administration of an incretin mimetic, there can be an increased and sustained level of circulating GLP-1. With DPP-IV inhibitors, the levels of GLP-1 only increase at meal time, according to normal physiological responses.⁵³ Aside from these differences, most of the remaining contrasts are due to side effects of taking the drugs. Both are capable of reducing HbA_{1C}, but GLP-1 analogues are associated with weight loss as well.⁵³

Despite the preferential route of administration displayed by DPP-IV inhibitors, there are still concerns surrounding the safety of their use. With regards to GLP-1, the binding between it and GLP-1R is highly specific. GLP-1R is a G-Protein Coupled Receptor (GPCR), which is activated by Group B hormones. Targeting an enzyme with small molecule inhibitors proves to be a less daunting task than attempting to develop one for a GPCR. The two current drugs marketed to interact with GLP-1R (Byetta[®] and Liraglutide) have shown selectivity due to GLP-1 having a single known receptor. The selectivity of inhibitors for DPP-IV over other enzymes of the same class is what leads to questions concerning their safety.⁵⁴

At the onset of investigation of DPP-IV inhibitors, one of the biggest queries involved the ability of proposed compounds to selectively inhibit DPP-IV over other members of the same enzyme family. With enzymes sharing similar characteristic, there is the possibility for inhibitors of DPP-IV to have activity at these other locations which may lead to adverse biological effects.

DPP-8 and DPP-9 are proteins which belong to the same group of enzymes as DPP-IV. Studies have shown that inhibition of their activity is related to serious toxicity issues.⁵⁵ It is because of these issues that any inhibitor developed must be selective for DPP-IV. In studies of compounds and their selectivity, it was seen that drugs which selectively inhibit DPP-IV do not display serious toxicity issues like DPP-8 and DPP-9 inhibitors.^{56,55} With regards to future research, there should be an importance placed on the discovery of compounds that are selective for DPP-IV as opposed to other enzymes in the same class.

Object of This Study

Type 2 Diabetes is a growing problem worldwide. Aside from insulin therapy, there are two main methods of combating this disease; GLP-1 mimetics and DPP-IV inhibitors. GLP-1R agonists present a plausible method of treating the disease. However, due to poor bioavailability of these drugs, they must be administered through injection. Alternatively, DPP-IV inhibitors present a novel method of treating T2DM. Through inhibition of this enzyme, the half-life of circulating GLP-1 is extended which stimulates the release of insulin. Currently there are two drugs which have been approved by the FDA for use in the treatment of T2DM, Sitagliptin and Saxagliptin. Additionally, Vildagliptin has been approved for market in Europe. The last two of these compounds belong to the cyanopyrrolidine class of compounds. One of the most distinguishing differences between these two compounds is where the bulky substitution is added (at the α -carbon vs. N-substitution). As previously mentioned the primary or secondary amine is necessary to correlate with Glu205 and Glu6 and contributes to slow dissociation kinetics. In addition to this, Villhauer and coworkers have developed an SAR to study the optimum N-substitution. It can be concluded from their work that bulky substitutions on the nitrogen are well tolerated. With the development of Saxagliptin, which has the adamantyl motif on the α -carbon, the question becomes: what is the best location for bulky substitution? It is the goal of this study to develop a functional assay to accurately test compounds, and develop an SAR to systematically determine the optimum substitution of steric bulk. A list of natural and unnatural amino acids was selected to help display different types of steric bulk and functionality at the α -carbon with relatively minimal N-substituted steric bulk (compared to the adamantyl).

The compounds will be synthesized in a library format utilizing solid phase synthesis to aid in maintaining pure samples throughout the process.

Chapter 2 References

1. Alexander, G. C.; Sehgal, N. L.; Moloney, R. M.; Stafford, R. S., National Trends in Treatment of Type 2 Diabetes Mellitus, 1994-2007. *Arch Intern Med* **2008**, *168* (19), 2088-2094.
2. A. G. Mainous III, R. B., R. J. Koopman, S. Saxena, V. A. Diaz, C. J. Everett, and A. Majeed, Impact of the population at risk of diabetes on projections of diabetes burden in the United States: an epidemic on the way. *Diabetologia* **2007**, *50* (5), 934-40.
3. Economic Costs of Diabetes in the U.S. in 2007. *Diabetes Care* **2008**, *31* (3), 596-615.
4. Cooke, D. W.; Plotnick, L., Type 1 Diabetes Mellitus in Pediatrics. *Pediatrics in Review* **2008**, *29* (11), 374-385.
5. Craig, M. E.; Hattersley, A.; Donaghue, K. C., Definition, epidemiology and classification of diabetes in children and adolescents. *Pediatric Diabetes* **2009**, *10*, 3-12.
6. Hober, D.; Sauter, P., Pathogenesis of type 1 diabetes mellitus: interplay between enterovirus and host. *Nat Rev Endocrinol* **2010**, *6* (5), 279-289.
7. Stolar, M. W., et al., Managing type 2 diabetes: going beyond glycemic control. *J Manag Care Pharm* **2008**, *14*, 2-19.
8. Langley, A. K. S., T.J.; Jennings, H.R., Dipeptidyl Peptidase IV Inhibitors and the Incretin System in Type 2 Diabetes Mellitus. *Pharmacotherapy* **2007**, *27* (8), 1163-1180.
9. Hays, N. P.; Galassetti, P. R.; Coker, R. H., Prevention and treatment of type 2 diabetes: Current role of lifestyle, natural product, and pharmacological interventions. *Pharmacology & Therapeutics* **2008**, *118* (2), 181-191.

10. Watson, R. T.; Kanzaki, M.; Pessin, J. E., Regulated Membrane Trafficking of the Insulin-Responsive Glucose Transporter 4 in Adipocytes. *Endocr Rev* **2004**, *25* (2), 177-204.
11. Donald Voet, J. G. V., Charlotte W. Pratt, *Fundamentals of Biochemistry: Life at the Molecular Level*. 2nd ed.; John Wiley & Sons Hoboken, 2006; p 1254.
12. Vilsboll, T., Holst, J.J., Incretins, insulin secretion and Type 2 diabetes mellitus. *Diabetologia* **2004**, *47* (3), 357-366.
13. Manhart, S.; Hinke, S. A.; McIntosh, C. H. S.; Pederson, R. A.; Demuth, H.-U., Structure–Function Analysis of a Series of Novel GIP Analogues Containing Different Helical Length Linkers†. *Biochemistry* **2003**, *42* (10), 3081-3088.
14. Baggio, L. L.; Drucker, D. J., Biology of Incretins: GLP-1 and GIP. *Gastroenterology* **2007**, *132* (6), 2131-2157.
15. Kim, W.; Egan, J. M., The Role of Incretins in Glucose Homeostasis and Diabetes Treatment. *Pharmacological Reviews* **2008**, *60* (4), 470-512.
16. Ross, S. A.; Dupre, J., Effects of ingestion of triglyceride or galactose on secretion of gastric inhibitory polypeptide and on responses to intravenous glucose in normal and diabetic subjects. *Diabetes* **1978**, *27* (3), 327-333.
17. Yip, R. G. C.; Wolfe, M. M., GIF biology and fat metabolism. *Life Sciences* **1999**, *66* (2), 91-103.
18. Trumper, A.; Trumper, K.; Trusheim, H.; Arnold, R.; Goke, B.; Horsch, D., Glucose-Dependent Insulinotropic Polypeptide Is a Growth Factor for {beta} (INS-1) Cells by Pleiotropic Signaling. *Mol Endocrinol* **2001**, *15* (9), 1559-1570.

19. Bell, G. I.; Santerre, R. F.; Mullenbach, G. T., Hamster preproglucagon contains the sequence of glucagon and two related peptides. *Nature* **1983**, *302* (5910), 716-718.
20. Gromada, J.; Brock, B.; Schmitz, O.; Rorsman, P., Glucagon-Like Peptide-1: Regulation of Insulin Secretion and Therapeutic Potential. *Basic & Clinical Pharmacology & Toxicology* **2004**, *95* (6), 252-262.
21. Fujita, Y.; Wideman, R. D.; Asadi, A.; Yang, G. K.; Baker, R.; Webber, T.; Zhang, T.; Wang, R.; Ao, Z.; Warnock, G. L.; Kwok, Y. N.; Kieffer, T. J., Glucose-Dependent Insulinotropic Polypeptide Is Expressed in Pancreatic Islet [alpha]-Cells and Promotes Insulin Secretion. *Gastroenterology In Press, Uncorrected Proof*.
22. Nauck, M. A.; Heimesaat, M. M.; Orskov, C.; Holst, J. J.; Ebert, R.; Creutzfeldt, W., Preserved incretin activity of glucagon-like peptide 1 [7-36 amide] but not of synthetic human gastric inhibitory polypeptide in patients with type-2 diabetes mellitus. *The Journal of Clinical Investigation* **1993**, *91* (1), 301-307.
23. Mentlein, R., Mechanisms underlying the rapid degradation and elimination of the incretin hormones GLP-1 and GIP. *Best Practice & Research Clinical Endocrinology & Metabolism* **2009**, *23* (4), 443-452.
24. Vilsbøll, T.; Krarup, T.; Deacon, C. F.; Madsbad, S.; Holst, J. J., Reduced Postprandial Concentrations of Intact Biologically Active Glucagon-Like Peptide 1 in Type 2 Diabetic Patients. *Diabetes* **2001**, *50* (3), 609-613.
25. Mentlein, R.; Gallwitz, B.; Schmidt, W. E., Dipeptidyl-peptidase IV hydrolyses gastric inhibitory polypeptide, glucagon-like peptide-1(7-36)amide, peptide histidine methionine and is responsible for their degradation in human serum. *European Journal of Biochemistry* **1993**, *214* (3), 829-835.

26. Barnett, A., DPP-4 inhibitors and their potential role in the management of type 2 diabetes. *International Journal of Clinical Practice* **2006**, *60* (11), 1454-1470.
27. Rosenstock, J., Zinman, B, Dipeptidyl peptidase-4 inhibitors and the management of type 2 diabetes mellitus. *Current Opinion in Endocrinology, Diabetes & Obesity* **2007**, *14*, 98-107.
28. Knudsen, L. B.; Pridal, L., Glucagon-like peptide-1-(9-36) amide is a major metabolite of glucagon-like peptide-1-(7-36) amide after in vivo administration to dogs, and it acts as an antagonist on the pancreatic receptor. *European Journal of Pharmacology* **1996**, *318* (2-3), 429-435.
29. Eng, J.; Kleinman, W. A.; Singh, L.; Singh, G.; Raufman, J. P., Isolation and characterization of exendin-4, an exendin-3 analogue, from *Heloderma suspectum* venom. Further evidence for an exendin receptor on dispersed acini from guinea pig pancreas. *Journal of Biological Chemistry* **1992**, *267*: , 7402-7405.
30. Bunck, M. C.; Diamant, M.; Cornér, A.; Eliasson, B.; Malloy, J. L.; Shaginian, R. M.; Deng, W.; Kendall, D. M.; Taskinen, M.-R.; Smith, U.; Yki-Järvinen, H.; Heine, R. J., One-Year Treatment With Exenatide Improves β -Cell Function, Compared With Insulin Glargine, in Metformin-Treated Type 2 Diabetic Patients: A randomized, controlled trial. *Diabetes Care* **2009**, *32* (5), 762-768.
31. Drucker, D. J.; Nauck, M. A., The incretin system: glucagon-like peptide-1 receptor agonists and dipeptidyl peptidase-4 inhibitors in type 2 diabetes. *The Lancet* **2006**, *368* (9548), 1696-1705.
32. Young, A. A.; Gedulin, B. R.; Bhavsar, S.; Bodkin, N.; Jodka, C.; Hansen, B.; Denaro, M., Glucose-lowering and insulin-sensitizing actions of exendin-4: studies in

obese diabetic (ob/ob, db/db) mice, diabetic fatty Zucker rats, and diabetic rhesus monkeys (*Macaca mulatta*). *Diabetes* **1999**, *48* (5), 1026-1034.

33. Bond, A., Exenatide (Byetta) as a novel treatment option for type 2 diabetes mellitus. *Proc (Bayl Univ Med Cent)* **2006**, *19*, 281-284.

34. Brubaker, P. L., Incretin-based therapies: mimetics versus protease inhibitors. *Trends in Endocrinology & Metabolism* **2007**, *18* (6), 240-245.

35. Kolterman, O. G.; Kim, D. D.; Shen, L.; Ruggles, J. A.; Nielsen, L. L.; Fineman, M. S.; Baron, A. D., Pharmacokinetics, pharmacodynamics, and safety of exenatide in patients with type 2 diabetes mellitus. *American Journal of Health-System Pharmacy* **2005**, *62* (2), 173-181.

36. Drucker, D. J.; Buse, J. B.; Taylor, K.; Kendall, D. M.; Trautmann, M.; Zhuang, D.; Porter, L., Exenatide once weekly versus twice daily for the treatment of type 2 diabetes: a randomised, open-label, non-inferiority study. *The Lancet* **2008**, *372* (9645), 1240-1250.

37. Mudaliar, S.; Henry, R. R., Effects of Incretin Hormones on [beta]-Cell Mass and Function, Body Weight, and Hepatic and Myocardial Function. *The American Journal of Medicine* **2010**, *123* (3, Supplement 1), S19-S27.

38. Drucker, D. J.; Dritselis, A.; Kirkpatrick, P., Liraglutide. *Nat Rev Drug Discov* **2010**, *9* (4), 267-268.

39. Neumiller, J. J.; Campbell, R. K., Liraglutide: A Once-Daily Incretin Mimetic for the Treatment of Type 2 Diabetes Mellitus. *Ann Pharmacother* **2009**, *43* (9), 1433-1444.

40. Nauck, M.; Frid, A.; Hermansen, K.; Shah, N. S.; Tankova, T.; Mitha, I. H.; Zdravkovic, M.; Düring, M.; Matthews, D. R., Efficacy and Safety Comparison of

Liraglutide, Glimepiride, and Placebo, All in Combination With Metformin, in Type 2 Diabetes: The LEAD (Liraglutide Effect and Action in Diabetes)-2 study. *Diabetes Care* **2009**, *32* (1), 84-90.

41. Ahrén, B.; Gomis, R.; Standl, E.; Mills, D.; Schweizer, A., Twelve- and 52-Week Efficacy of the Dipeptidyl Peptidase IV Inhibitor LAF237 in Metformin-Treated Patients With Type 2 Diabetes. *Diabetes Care* **2004**, *27* (12), 2874-2880.

42. Campbell, R. K., Rationale for Dipeptidyl Peptidase 4 Inhibitors: A New Class of Oral Agents for the Treatment of Type 2 Diabetes Mellitus. *Ann Pharmacother* **2007**, *41* (1), 51-60.

43. Herman, G. A.; Bergman, A.; Liu, F.; Stevens, C.; Wang, A. Q.; Zeng, W.; Chen, L.; Snyder, K.; Hilliard, D.; Tanen, M.; Tanaka, W.; Meehan, A. G.; Lasseter, K.; Dilzer, S.; Blum, R.; Wagner, J. A., Pharmacokinetics and Pharmacodynamic Effects of the Oral DPP-4 Inhibitor Sitagliptin in Middle-Aged Obese Subjects. *J Clin Pharmacol* **2006**, *46* (8), 876-886.

44. Pospisilik, J. A.; Martin, J.; Doty, T.; Ehses, J. A.; Pamir, N.; Lynn, F. C.; Piteau, S.; Demuth, H.-U.; McIntosh, C. H. S.; Pederson, R. A., Dipeptidyl Peptidase IV Inhibitor Treatment Stimulates β -Cell Survival and Islet Neogenesis in Streptozotocin-Induced Diabetic Rats. *Diabetes* **2003**, *52* (3), 741-750.

45. Fura, A.; Khanna, A.; Vyas, V.; Koplowitz, B.; Chang, S.-Y.; Caporuscio, C.; Boulton, D. W.; Christopher, L. J.; Chadwick, K. D.; Hamann, L. G.; Humphreys, W. G.; Kirby, M., Pharmacokinetics of the Dipeptidyl Peptidase 4 Inhibitor Saxagliptin in Rats, Dogs, and Monkeys and Clinical Projections. *Drug Metabolism and Disposition* **2009**, *37* (6), 1164-1171.

46. Dhillon, S.; Weber, J., Saxagliptin. *Drugs* **2009**, *69* (15), 2103-2114
10.2165/11201170-000000000-00000.
47. Pei, Z.; Li, X.; Longenecker, K.; von Geldern, T. W.; Wiedeman, P. E.; Lubben, T. H.; Zinker, B. A.; Stewart, K.; Ballaron, S. J.; Stashko, M. A.; Mika, A. K.; Beno, D. W. A.; Long, M.; Wells, H.; Kempf-Grote, A. J.; Madar, D. J.; McDermott, T. S.; Bhagavatula, L.; Fickes, M. G.; Pireh, D.; Solomon, L. R.; Lake, M. R.; Edalji, R.; Fry, E. H.; Sham, H. L.; Trevillyan, J. M., Discovery, Structure–Activity Relationship, and Pharmacological Evaluation of (5-Substituted-pyrrolidinyl-2-carbonyl)-2-cyanopyrrolidines as Potent Dipeptidyl Peptidase IV Inhibitors. *Journal of Medicinal Chemistry* **2006**, *49* (12), 3520-3535.
48. Peters, J., 11 Years of Cyanopyrrolidines as DPP-IV Inhibitors. *Current Topics in Medicinal Chemistry* **2007**, *7*, 579-595.
49. Lautar, S. L.; Rojas, C.; Slusher, B. S.; Wozniak, K. M.; Wu, Y.; Thomas, A. G.; Waldon, D.; Li, W.; Ferraris, D.; Belyakov, S., DPP IV inhibitor blocks mescaline-induced scratching and amphetamine-induced hyperactivity in mice. *Brain Research* **2005**, *1048* (1-2), 177-184.
50. Villhauer, E. B.; Brinkman, J. A.; Naderi, G. B.; Burkey, B. F.; Dunning, B. E.; Prasad, K.; Mangold, B. L.; Russell, M. E.; Hughes, T. E., 1-[[[3-Hydroxy-1-adamantyl)amino]acetyl]-2-cyano-(S)-pyrrolidine: A Potent, Selective, and Orally Bioavailable Dipeptidyl Peptidase IV Inhibitor with Antihyperglycemic Properties. *Journal of Medicinal Chemistry* **2003**, *46* (13), 2774-2789.
51. Villhauer, E. B.; Brinkman, J. A.; Naderi, G. B.; Dunning, B. E.; Mangold, B. L.; Mone, M. D.; Russell, M. E.; Weldon, S. C.; Hughes, T. E., 1-[2-[(5-Cyanopyridin-2-

yl)amino]ethylamino]acetyl-2-(S)-pyrrolidinecarbonitrile: A Potent, Selective, and Orally Bioavailable Dipeptidyl Peptidase IV Inhibitor with Antihyperglycemic Properties. *Journal of Medicinal Chemistry* **2002**, *45* (12), 2362-2365.

52. Kondo, T.; Nekado, T.; Sugimoto, I.; Ochi, K.; Takai, S.; Kinoshita, A.; Tajima, Y.; Yamamoto, S.; Kawabata, K.; Nakai, H.; Toda, M., Design and synthesis of new potent dipeptidyl peptidase IV inhibitors with enhanced ex vivo duration. *Bioorganic & Medicinal Chemistry* **2007**, *15* (7), 2631-2650.

53. Boyle, P. J.; Freeman, J. S., Application of Incretin Mimetics and Dipeptidyl Peptidase IV Inhibitors in Managing Type 2 Diabetes Mellitus. *J Am Osteopath Assoc* **2007**, *107* (suppl_3), S10-16.

54. Drucker, D. J., Therapeutic potential of dipeptidyl peptidase IV inhibitors for the treatment of type 2 diabetes. *Expert opinion on investigational drugs* **2003**, *12* (1), 87-100.

55. Van der Veken, P.; Haemers, A.; Augustyns, K., Prolyl Peptidases Related to Dipeptidyl Peptidase IV: Potential of Specific Inhibitors in Drug Discovery. *Current Topics in Medicinal Chemistry* **2007**, *7* (6), 621-635.

56. Green, B. D.; Flatt, P. R.; Bailey, C. J., Dipeptidyl peptidase IV (DPP IV) inhibitors: a newly emerging drug class for the treatment of type 2 diabetes. *Diabetes and Vascular Disease Research* **2006**, *3* (3), 159-165.

Chapter 3:
DPP-IV Results

Synthesis of Analogues

In order to make reasonable conclusions regarding future testing and designing of molecules, a functional assay must be developed. The easiest way to do this was to reproduce previously recorded data. The first part of the project was to synthesize three compounds, two of which have already been tested (**3.1** and **3.2**), to aid in the development of an in-house assay. Decisions surrounding which compounds were to be synthesized first were based on an in-house Comparative Molecular Field Analysis (CoMFA) study as well as previously reported literature data.^{1,2} As a result, the three compounds synthesized first have an adamantyl, cyclohexyl, and methylcyclohexyl group attached to the *N*-terminus amino group (**Figure 3.1**). These molecules were synthesized so that synthetic protocols for a potential library could be explored, as well as assisting in the development of the assay.

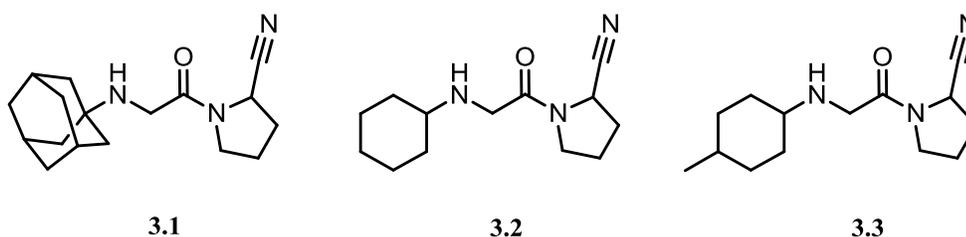
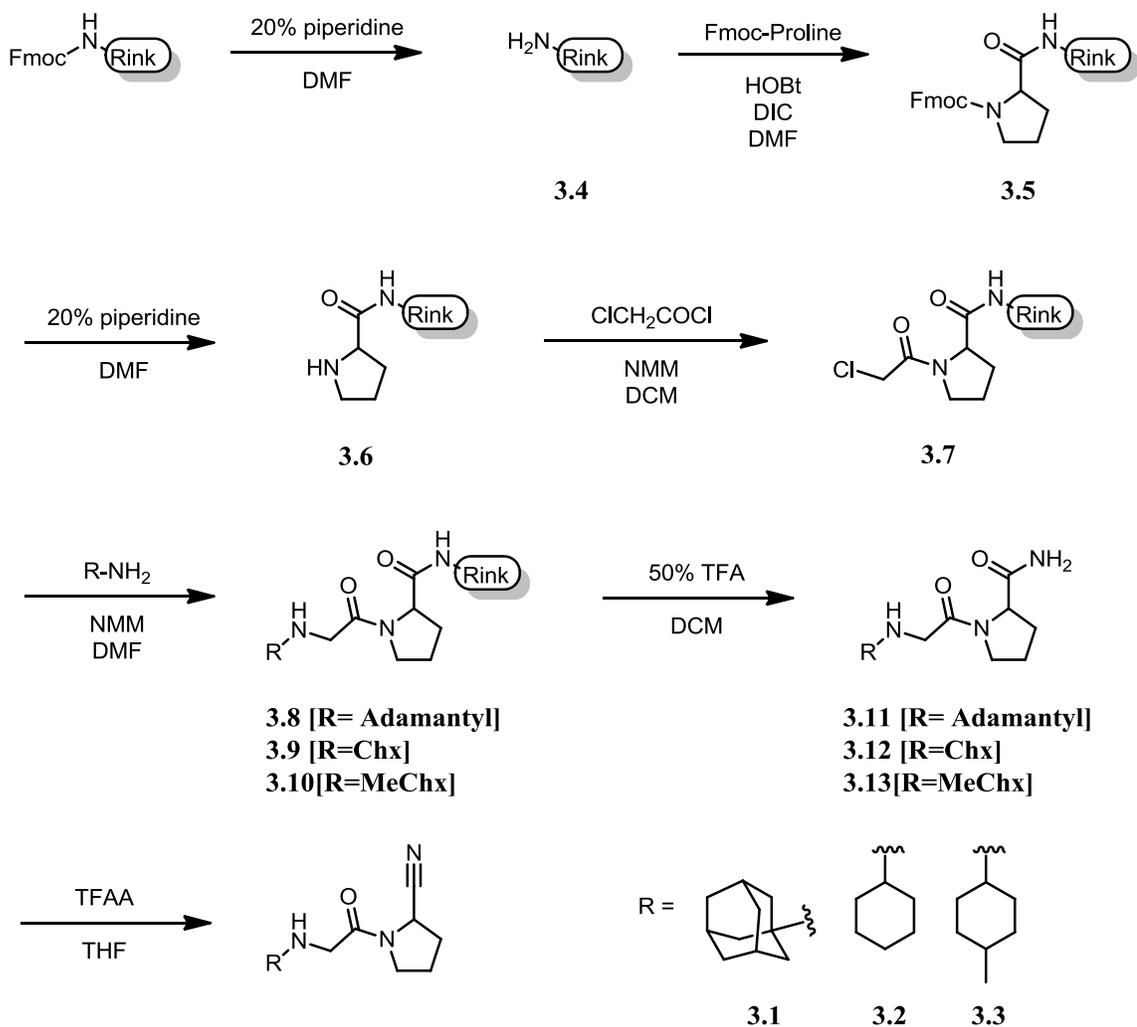


Figure 3.1 – Three compounds synthesized to test for reproducibility

For the synthesis of these molecules, a solid phase approach was employed similar to the one used by Villhauer and co-workers.² As stated previously, solid-phase synthesis allowed for a large excess of starting materials and reagents to be used to drive each step of the reaction to completion (**Scheme 3.1**). The synthesis began with coupling

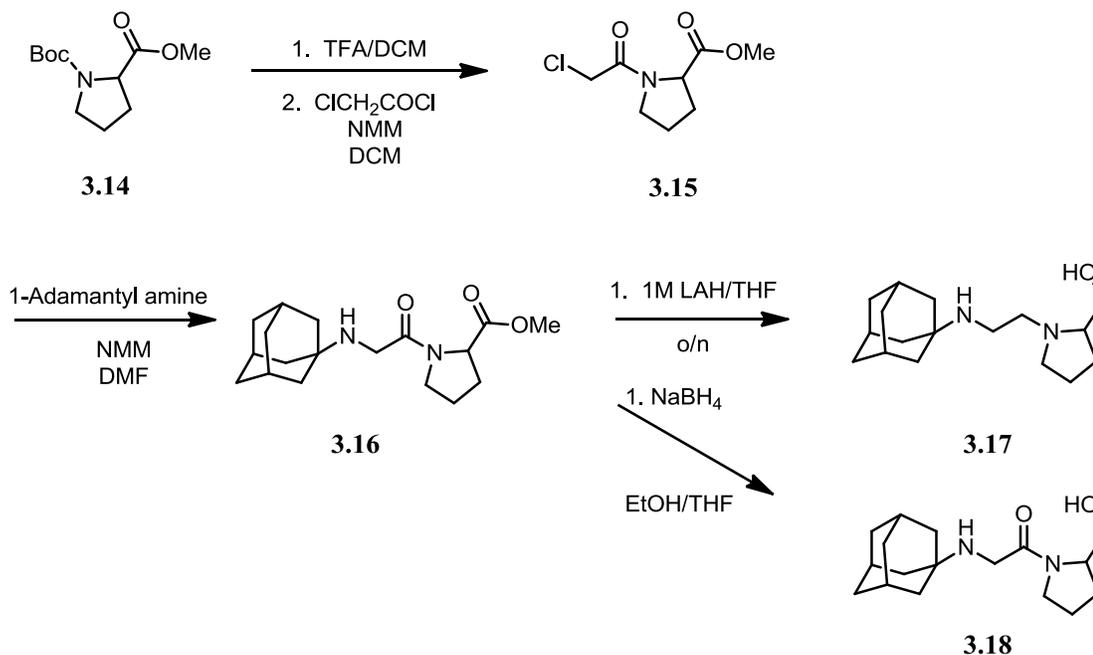
Fmoc-L-Proline to standard Rink linker resin using traditional hydroxybenzotriazole (HOBt) and *N,N'*-diisopropylcarbodiimide (DIC) coupling conditions in *N,N*-dimethylformamide (DMF). Following this step, the Fmoc protecting group was removed using 20% piperidine in DMF solution. After removal of the Fmoc group, chloroacetyl chloride was coupled to the secondary amine in DCM with an excess of the base *N*-methylmorpholine. Substitution of the chloride was achieved using a primary amine (adamantyl, cyclohexyl, methylcyclohexyl). Cleavage from the resin produced the primary amide which was readily converted to the nitrile mediated by trifluoroacetic anhydride (TFAA).

Scheme 3.1 – Synthetic scheme used to synthesize the first three compounds

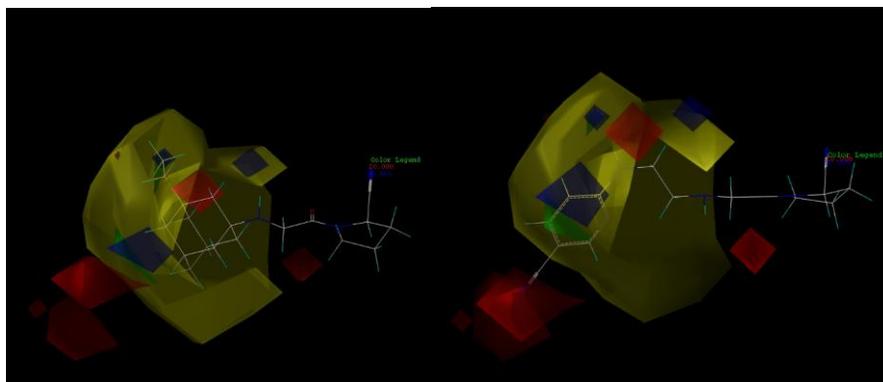
There was concern for acylating the amine during dehydration. In practice, there was little to no acylation observed when the adamantyl group was part of the molecule, most likely due to the large steric bulk surrounding the amine. The other two compounds produced a mixture of the acylated and non-acylated products. For the purposes of this study, these compounds were purified using preparative High Performance Liquid Chromatography (HPLC) for biological testing.

Two additional compounds were also synthesized for biological testing. In one of the compounds, the nitrile functionality was replaced with a methyl hydroxyl group to see how this modification would be tolerated in the enzyme binding pocket. This would take a functional group which has an electrophilic carbon and a hydrogen bond accepting nitrogen and replace it with a functional group that is not as electrophilic, but can participate in hydrogen bond accepting. The other compound also replaced the nitrile with a methyl alcohol, and additionally reduced the amide bond to a tertiary amine.

The synthesis was carried out using a solution phase protocol (**Scheme 3.2**). First, commercially available Boc-1-proline methyl ester was mixed with TFA in DCM to remove the Boc protecting group. This free secondary amine was coupled to chloroacetyl chloride in a basic environment. An S_N2 substitution using adamantylamine produced the secondary amine. From here, the product was split into two different batches and subjected to two different reduction conditions. The first set of conditions used sodium borohydride (NaBH_4) to selectively reduce the methyl ester to a methyl alcohol while leaving the amide bond untouched. The second set of conditions utilized lithium aluminum hydride (LAH) to reduce both the methyl ester and the amide to give the methyl alcohol and the tertiary amine.³ As before, these compounds were synthesized for the initial round of screening to determine if the nitrile could be removed.

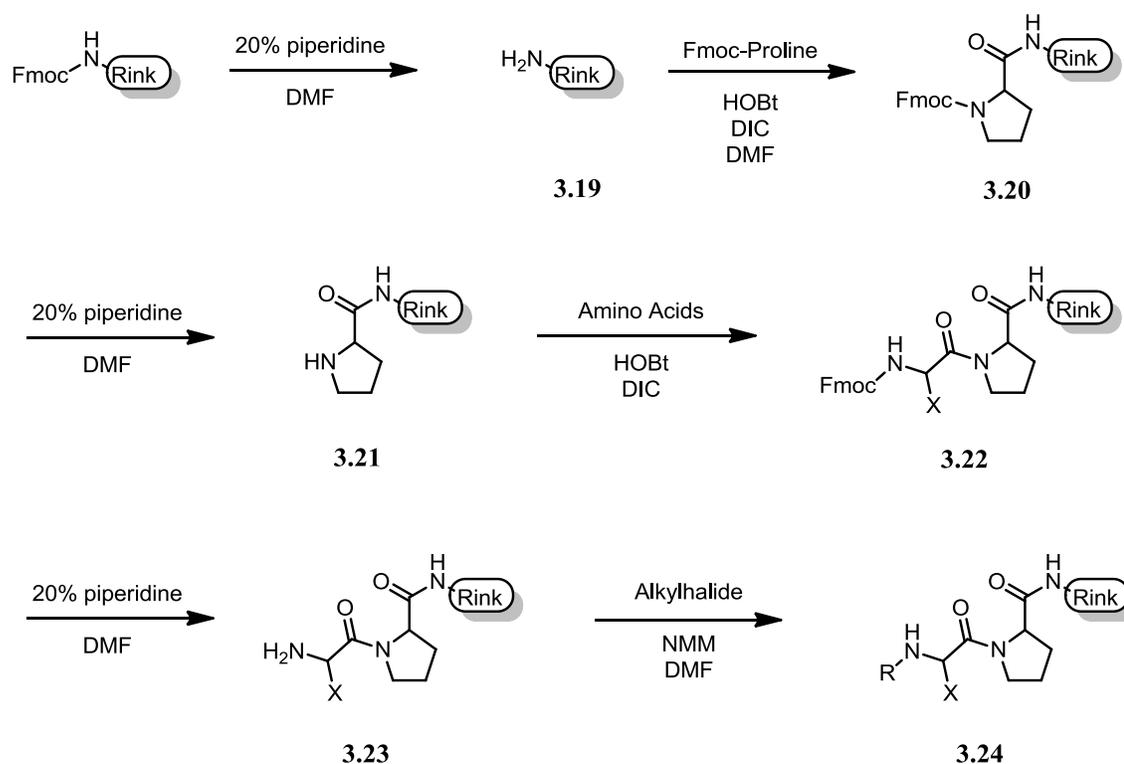
Scheme 3.2 – Synthetic scheme for the two methyl alcohol analogues*Library Synthesis*

The two compounds which were the basis for this work, utilize the cyanopyrrolidine scaffold while displaying large bulky hydrophobic side chains at either the α -carbon, or at the N-terminus of the P2 region of the binding pocket. In-house CoMFA studies confirm that steric bulk added at either of these locations should be well tolerated (**Figure 3.2**).

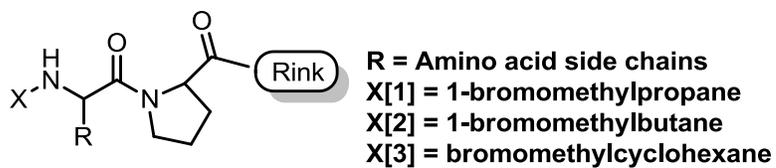
**Figure 3.2** – Sample CoMFA images showing the preference for hydrophobic side chains in the P2 region of the binding pocket

Through the SAR developed by Villhauer and co-workers at Novartis, and personal correspondence with BristolMyersSquibb, extensive studies were conducted on the tolerance of steric bulk at either the α -carbon or the N-terminus. However, little insight was given to molecules which display steric bulk at both locations. Villhauer demonstrated on a select number of compounds that bulk added to both places produced a reduction in activity. However, when examining their data, it was apparent that the substitutions they use were limited to just an isopropyl group and cyclohexyl group. Optimization of the steric bulk substitution and its distance from the amine could improve upon existing activity and was the basis for the design of the library.

Scheme 3.3 – Synthesis of the 42-membered library



The design library is considered a 3x14 compound library. There were fourteen normal and unusual amino acids used as one of the building blocks. To add substitution at the N-terminus, three different alkylhalides were selected (see synthesis map in **Table 3.1**). Following the deprotection of the Fmoc Rink resin, the synthesis began with the coupling of Fmoc-l-proline to the resin using standard HOBt/DIC coupling conditions (**Scheme 3.3**). Successful coupling was followed by removal of the Fmoc group from proline. Fourteen different natural and unusual Fmoc-l-amino acids were used to couple to the resin. Removal of the Fmoc group using 20% piperidine in DMF afforded the primary amine. This can be alkylated using standard S_N2 reaction conditions with three different alkylbromides. This reaction must be monitored carefully as to avoid over alkylation. From here, the library was cleaved to its corresponding amide. The synthesis was stopped at this point and the project was terminated before the amide could be converted to the nitrile.

Table 3.1 – Synthesis map used for synthesizing compounds.

Gly X [1]	Gly X [2]	Gly X [3]	Ala X [1]	Ala X [2]
Ala X [3]	Val X [1]	Val X [2]	Val X [3]	Ile X [1]
Ile X [2]	Ile X [3]	Leu X [1]	Leu X [2]	Leu X [3]
Nle X [1]	Nle X [2]	Nle X [3]	Nva X [1]	Nva X [2]
Nva X [3]	Phe X [1]	Phe X [2]	Phe X [3]	3-Pal X [1]
3-Pal X [2]	3-Pal X [3]	Nal X [1]	Nal X [2]	Nal X [3]
Chg X [1]	Chg X [2]	Chg X [3]	Cha X [1]	Cha X [2]
Cha X [3]	cLeu X [1]	cLeu X [2]	cLeu X [3]	Abu X [1]
Abu X [2]	Abu X [3]			

The proposed synthetic pathway would couple ethanolamine to Boc-l-Proline using standard HOBt/DIC coupling conditions. Following removal of the Boc group, chloroacetyl chloride could be coupled to the free secondary amine. Transformation of the alcohol to the mesylate can be achieved using methanesulfonyl chloride and the oxazoline is generated in situ.⁴ After generating this product, a simple S_N2 reaction would be able to add the adamantyl group to the molecule completing the synthesis.

Assay Development

The synthesized compounds were to be tested using a classic competition assay. Initial studies planned to have the assay completed using a UV/Vis Spectrometer. Using the substrate Nle-Pro-AMC (where AMC is 7-amino-4-methylcoumarin), the DPP-IV mediated release of AMC could be measured using wavelength absorption. However, due to only being able to run one sample at a time, it would be more efficient to simultaneously run multiple enzyme and substrate concentrations. Therefore, the experiments were moved to a plate reader where the DPP-IV mediated release of AMC was measured. The excitation wavelength used was 340nm and the emission wavelength measured was at 485nm. Literature values reported the K_M for the substrate Nle-Pro-AMC to be 34 ± 5 .⁵ In order to establish a proper assay protocol, experiments were run in an attempt to reproduce this number.

An appropriate concentration of DPP-IV enzyme had to be calculated in order to effectively test for compounds. An assay was set up using fixed substrate concentration (50μM and 100μM) and varied enzyme concentrations (1/20, 1/10, and 1/5 μunit enzyme concentration per well). By performing this experiment, it was determined that the optimum enzyme concentration was the 1/10 μunit/per well concentration.

Using the optimum enzyme concentration, an experiment was designed using a wide range of substrate concentrations to reproduce the literature substrate K_M values ($34 \pm 5 \mu\text{M}$). It was the expectation that the results produced by these experiments would follow traditional Michaelis-Menten kinetics for enzyme inhibition. However, it was observed that at higher concentrations of substrate, the output produced by the instrument started to decrease (**Figure 3.3**).

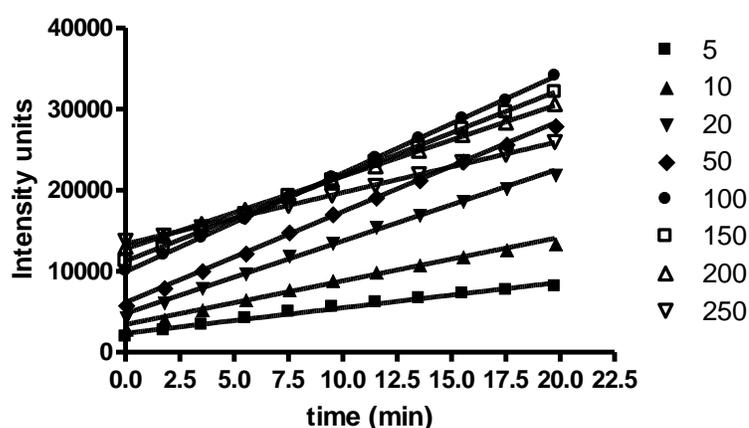


Figure 3.3 - Raw experimental data showing the decrease in Intensity units over time with increasing substrate concentration (μM).

When this data was further analyzed and the slope was plotted versus the substrate concentration, it was determined that at the higher concentrations, substrate inhibition was being observed.⁶ This occurs with about 20% of all enzymes and happens when two molecules of substrate can bind to the enzyme and ultimately block activity (**Figure 3.4**).⁶ Although this is not what was expected, by applying the equation $v = V_{\max} [S] / (K_M + [S](1 + ([S] / K_i)))$, a K_M value for the substrate could be obtained. The calculated K_M from this experiment was relatively close to what was reported in the literature ($K_M = 42.7 \pm 18.7 \mu\text{M}$). This model can be used for the testing of the

synthesized compounds. Also observed, was that at lower substrate concentrations, this data behaved more like a traditional Michaelis-Menten curve.

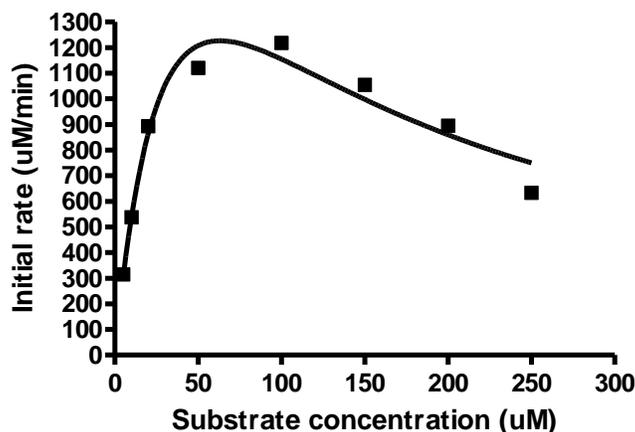
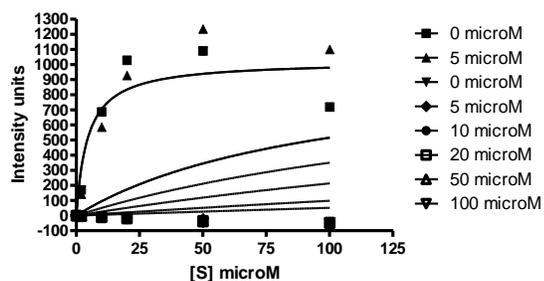


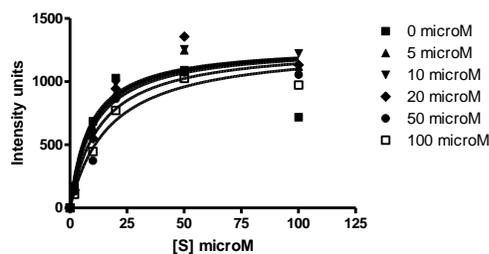
Figure 3.4 - Initial rate versus substrate concentration which depicts substrate inhibition

Previously, it has been reported in the literature, that all of the assay experiments were run at a constant temperature of 37 °C. It was with this in mind that the substrate inhibition assay was performed again to see if altering the temperature would have any effect substrate inhibition. Although at first it appeared as though this was a positive outcome, when the initial slopes were plotted against substrate concentration, above 100 μ M, substrate inhibition was still observed.

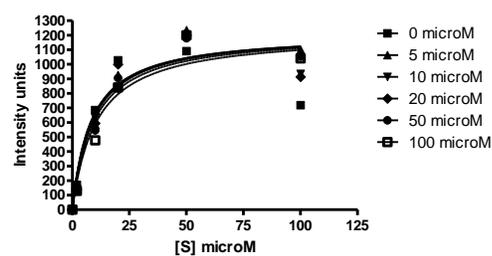
Although the substrate inhibition makes the analysis a little more challenging, the three synthesized compounds (**3.1**, **3.2**, and **3.3**) were tested for their corresponding activity (**Figure 3.5**). The results of these experiments represented previously reported data, where the molecule which contained the adamantyl group had the lowest K_M value of the three compounds ($4.5 \pm 3.5\mu\text{M}$ vs. 8.63 ± 1.9 and $8.20 \pm 1.8\mu\text{M}$).



3.1



3.2



3.3

Km	4.545
I	
Ki	0.2420
Vmax	1024
Std. Error	
Km	3.517
Ki	0.1847
Vmax	159.2
95% Confidence Intervals	
Km	-2.544 to 11.63
Ki	-0.1303 to 0.6143
Vmax	702.8 to 1344
Goodness of Fit	
Degrees of Freedom	45
R square	0.6112

Km	8.637
I	
Ki	99.45
Vmax	1291
Std. Error	
Km	1.864
Ki	55.70
Vmax	62.20
95% Confidence Intervals	
Km	4.842 to 12.43
Ki	-13.93 to 212.8
Vmax	1164 to 1417
Goodness of Fit	
Degrees of Freedom	33
R square	0.9301

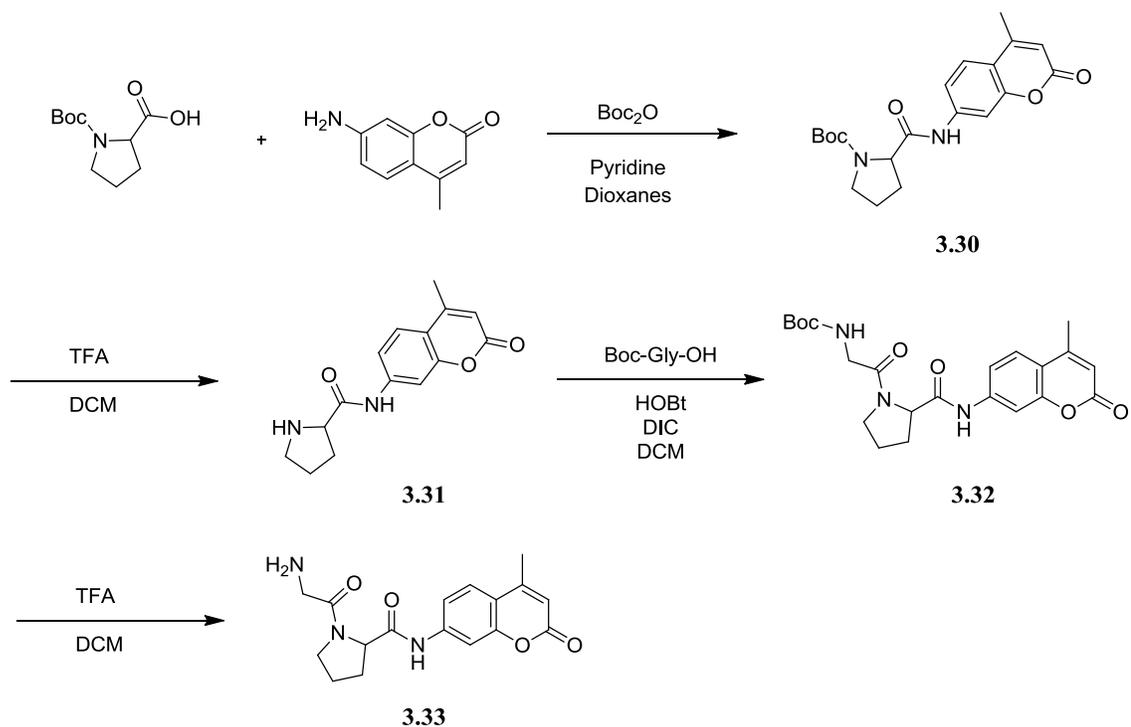
Km	8.207
I	
Ki	340.1
Vmax	1215
Std. Error	
Km	1.784
Ki	464.9
Vmax	57.74
95% Confidence Intervals	
Km	4.575 to 11.84
Ki	-606.4 to 1287
Vmax	1097 to 1333
Goodness of Fit	
Degrees of Freedom	33
R square	0.9232

Figure 3.5 - Data for the three synthesized inhibitors of DPP-IV

Although the assay appears to be functioning well, the issue of substrate inhibition at higher concentrations had not been eliminated. Initial screening of the library will be used to identify active molecules in a first pass assay. This screen may not be as rigorous as for compounds **3.1-3.3**, but higher concentrations of substrate would need to be used. Once a select number of compounds have been identified as having a desirable activity, they would be selected for further, more invasive testing. This would involve screening with a larger number of substrate and inhibitor concentrations and taking more data points over a longer period of time. As a way to limit the amount of substrate inhibition, a new substrate which would have a lower binding affinity should be selected. With this, Leiting and co-workers identified Gly-Pro-AMC as having a lower binding affinity of K_M of $63 \pm 12 \mu\text{M}$.

Synthesis of the new Substrate

The synthesis of the new substrate began with the coupling of 7-amino-4-methylcoumarin to Boc-L-Proline using pyridine, Boc_2O , and dioxanes as coupling conditions (**Scheme 3.5**).⁷ Following this, the Boc group was removed using a 50/50 mixture of TFA in DCM. The free secondary amine was coupled to Boc-L-Glycine using standard HOBt/DIC protocols. Removal of the Boc group yielded the final product Gly-Pro-AMC. This substrate was synthesized for the replacement of Nle-Pro-AMC.

Scheme 3.5 – The synthesis of the new substrate Gly-Pro-AMC

Following the synthesis of Gly-Pro-AMC, this substrate was also tested to determine if this substrate would act as an inhibitor at higher concentrations. Unfortunately, the outcome of this experiment showed that this new substrate followed the previously observed trends with substrate Nle-Pro-AMC. At concentrations above $50\mu\text{M}$, substrate inhibition was still observed. Before this issue could be resolved, the project was terminated and efforts were shifted to the new project explained in detail in Chapters 5-7.

Chapter 3 References

1. Villhauer, E. B.; Brinkman, J. A.; Naderi, G. B.; Burkey, B. F.; Dunning, B. E.; Prasad, K.; Mangold, B. L.; Russell, M. E.; Hughes, T. E., 1-[[3-Hydroxy-1-adamantyl)amino]acetyl]-2-cyano-(S)-pyrrolidine: A Potent, Selective, and Orally Bioavailable Dipeptidyl Peptidase IV Inhibitor with Antihyperglycemic Properties. *Journal of Medicinal Chemistry* **2003**, *46* (13), 2774-2789.
2. Villhauer, E. B.; Brinkman, J. A.; Naderi, G. B.; Dunning, B. E.; Mangold, B. L.; Mone, M. D.; Russell, M. E.; Weldon, S. C.; Hughes, T. E., 1-[2-[(5-Cyanopyridin-2-yl)amino]ethylamino]acetyl-2-(S)-pyrrolidinecarbonitrile: A Potent, Selective, and Orally Bioavailable Dipeptidyl Peptidase IV Inhibitor with Antihyperglycemic Properties. *Journal of Medicinal Chemistry* **2002**, *45* (12), 2362-2365.
3. Tuthill, P. A.; Seida, P. R.; Barker, W.; Cassel, J. A.; Belanger, S.; DeHaven, R. N.; Koblisch, M.; Gottshall, S. L.; Little, P. J.; DeHaven-Hudkins, D. L.; Dolle, R. E., Azepinone as a conformational constraint in the design of [kappa]-opioid receptor agonists. *Bioorganic & Medicinal Chemistry Letters* **2004**, *14* (22), 5693-5697.
4. Gilbertson, S. R.; Fu, Z., Chiral P,N-Ligands Based on Ketopinic Acid in the Asymmetric Heck Reaction. *Organic Letters* **2000**, *3* (2), 161-164.
5. Leiting, B.; Pryor, K. D.; Wu, J. K.; Marsilio, F.; Patel, R. A.; Craik, C. S.; Ellman, J. A.; Cummings, R. T.; Thornberry, N. A., Catalytic properties and inhibition of proline-specific dipeptidyl peptidases II, IV and VII. *Biochem. J.* **2003**, *371* (2), 525-532.
6. Copeland, R. A., *Enzymes*. 2nd ed.; Wiley: 2000.

7. Hamzé, A.; Martinez, J.; Hernandez, J.-F., Solid-Phase Synthesis of Arginine-Containing Peptides and Fluorogenic Substrates Using a Side-Chain Anchoring Approach. *The Journal of Organic Chemistry* **2004**, *69* (24), 8394-8402.

Chapter 4:
DPP-IV Discussion

It was the goal at the outset of this study to improve upon existing Biological activity by further developing a comprehensive structure activity relationship (SAR). Previous work has involved making substitutions at either the *N*-terminus or at the α -carbon associated with the P2 region of the binding pocket. Initial synthesis of analogues proceeded in a straight forward manner in accordance with what has been previously published.¹ Two of the first three compounds synthesized have previously been tested by Villhauer and co-workers at Novartis and were synthesized to aid in the development of a functioning assay. Further modifications were made to the parent molecule which included changing the nitrile to a methyl alcohol and the reduction of the amide to a tertiary amine. These were synthesized to determine how these changes affect overall activity.

Crystal structures have shown key interactions exists with cyanopyrrolidine based inhibitors between the primary or secondary amine and Glu205/Glu206 in the binding pocket.² By introducing too much steric bulk around the amine, this interaction could be prevented resulting in lower biological activity. The study by Villhauer put little emphasis on substitution at both locations. Optimization of this substitution pattern could lead to a more active molecule and was the basis for the design of the library. By removing the large adamantyl group and replacing it with two smaller groups may allow for the salt bridge interaction to take place, and still improve upon activity. The design of the library incorporated three smaller hydrophobic substitutions at the *N*-terminus and 14 different natural and unusual amino acids which were selected to display a variety of hydrophobic side chains at the α -position to systematically explore steric restrictions.

The synthesis of the library was carried out through the cleavage from the resin, however, before it could be completed, the project was terminated.

In addition to the synthesis associated with this project, an assay needed to be developed which can accurately test selected analogues. With initial experiments, it was the expectation that the results would follow the traditional Michaelis-Menten model for enzyme kinetics. However, at high substrate concentrations, results indicated that the substrate Nle-Pro-AMC may be participating in substrate inhibition. This result is not unusual, as this occurs with about 20% of all known enzymes.³ If two molecules of the substrate are able to interact at the binding site simultaneously, enzymatic activity can be further blocked. What was unusual about this is that there was no mention of substrate inhibition in the protocols for the assay as described by others.

Once the substrate inhibition model was applied, the K_M observed was relatively close to literature values ($K_M = 42.7 \pm 18.7\mu\text{M}$). Subsequent studies did show that at lower substrate concentrations, classical Michaelis-Menten kinetics was observed. With this model, three compounds were tested for their corresponding activity. From these experiments, it was shown that the adamantyl analogue (**3.1**) was the most active with a K_i of $0.242 \pm 1.28\mu\text{M}$. The cyclohexyl analogues was shown to be more active than the methylcyclohexy, but not as active as the adamantyl ($K_i = 99.5 \pm 55.7\mu\text{M}$ and $340.1\mu\text{M}$ respectively).

In an effort to minimize substrate inhibition, a new substrate was designed and synthesized. By selecting a new substrate with a lower binding affinity, the likelihood of observing substrate inhibition at higher concentrations should decrease. This would allow for a wider range of substrate concentrations to be used during assays. The

substrate Gly-Pro-AMC has a reported K_M of $63 \pm 12\mu\text{M}$ and was successfully synthesized.⁴ However, before it could be tested to see if it also participated in substrate inhibition, the project had been terminated.

Conclusions

Previously reported research has demonstrated that the binding pocket of DPP-IV tolerates large hydrophobic substituents on either the α -carbon or the *N*-terminal. However, little insight has been reported into developing optimum substitution of steric bulk with regards to each of these locations. With this, a 42-compound library was developed to explore the steric restrictions of the DPP-IV binding site.

In addition to the library, three additional compounds were synthesized to aid in the development of an assay protocol. During the development of the assay, it was observed that the substrate Nle-Pro-AMC participated in competitive inhibition of the DPP-IV enzyme. This was unusual as Nle-Pro-AMC is the standard substrate used for this assay. Once this had been identified, the appropriate model could be applied, and a K_M value for the substrate could be produced ($42.7 \pm 18.7\mu\text{M}$). After establishing this, compounds **3.1-3.3** could be assayed for their corresponding activity. Generally, the activity of the compounds followed the trends observed in the literature, with the adamantyl analogue being the most active.

Further molecules synthesized involved substituting a methyl alcohol for the nitrile. Additionally, in one of the methyl alcohol analogues, the amide bond was reduced to the tertiary amine. This modification would also introduce more flexibility to

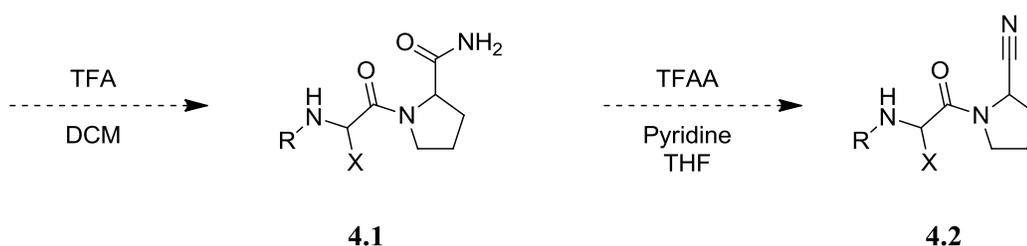
the molecule. Additionally, a molecule has been proposed which would substitute an oxazoline ring for the nitrile.

The proposed plan was to perform a simplified assay on the library in order to identify active compounds. In an effort to minimize the effects of substrate inhibition during this initial assay, a new substrate was synthesized. Gly-Pro-AMC was selected due to the reduced K_M value and successfully synthesized. Before the substrate could be examined and the library screened, the project was ultimately terminated.

Future Work

Although this project was terminated prematurely, there is much more work which can be accomplished with respect to both the biological and synthetic aspects of this project. First, the synthesis of the library needs to be completed. This will require dehydration of the primary amide to the nitrile mediated by TFAA and pyridine in THF (Scheme 4.1).

Scheme 4.1 – Final synthetic procedures of the library



Due to the decreased steric bulk surrounding the amine, the likelihood that the secondary amine is acetylated by TFAA is greatly increased. Therefore, following the dehydration of the amide, each compound can be treated with either ammonia or potassium carbonate to cleave the trifluoroacetyl group.

Development of a suitable assay protocol needs to be addressed. Finding a substrate where competitive inhibition is not observed would be ideal. Additionally, once the library synthesis has been completed, it can be tested for its corresponding activity and greater insight will be gained into the steric restrictions in the P2 binding pocket.

Chapter 4 References

1. Villhauer, E. B.; Brinkman, J. A.; Naderi, G. B.; Burkey, B. F.; Dunning, B. E.; Prasad, K.; Mangold, B. L.; Russell, M. E.; Hughes, T. E., 1-[[[3-Hydroxy-1-adamantyl)amino]acetyl]-2-cyano-(S)-pyrrolidine: A Potent, Selective, and Orally Bioavailable Dipeptidyl Peptidase IV Inhibitor with Antihyperglycemic Properties. *Journal of Medicinal Chemistry* **2003**, *46* (13), 2774-2789.
2. Peters, J., 11 Years of Cyanopyrrolidines as DPP-IV Inhibitors. *Current Topics in Medicinal Chemistry* **2007**, *7*, 579-595.
3. Copeland, R. A., *Enzymes*. 2nd ed.; Wiley: 2000.
4. Leiting, B.; Pryor, K. D.; Wu, J. K.; Marsilio, F.; Patel, R. A.; Craik, C. S.; Ellman, J. A.; Cummings, R. T.; Thornberry, N. A., Catalytic properties and inhibition of proline-specific dipeptidyl peptidases II, IV and VII. *Biochem. J.* **2003**, *371* (2), 525-532.

Chapter 5:
Introduction to NS1

Influenza is a continuing public health problem worldwide. The World Health Organization (WHO) estimates of seasonal influenza mortality ranges from 250,000 to 500,000, which includes 30,000 deaths and 200,000 hospitalizations in the United States.¹ Most common high risk groups include the elderly, the young, and those suffering from chronic illness.^{2,3,4} In recent history, there have been four recorded influenza pandemics, the most severe occurring in 1918. During an eight month period, the 1918 Spanish Influenza claimed the lives of between 20-40 million people.⁵ The subsequent pandemics of 1957 and 1968 were not as severe as in 1918, but shared the similar characteristic of having avian origin.⁶ Due to repeated outbreaks, there is concern that another pandemic is inevitable. This belief is perpetuated by the spread of the H5N1 influenza strain among avian species in Asia, Africa, and Europe, which periodically infect humans.^{7,8,9} Although the H5N1 strain of influenza has not demonstrated the ability to transmit from human to human, the small numbers who have been infected through direct avian contact displayed alarmingly high mortality rates, approximately 60%.^{10,11}

In April 2009, a novel strain of influenza emerged in the United States and Mexico. This new viral strain was able to rapidly spread throughout the world through human to human transmission. By early June, the virus had spread to 74 different countries and over 29,000 cases had been reported. On June, 11, 2009, the World Health organization had declared the H1N1 influenza virus a pandemic.¹² Up until this point, the seasonal influenza strain has primarily been composed of A/H3N2 virus variants which have been developed due to regular antigenic drift.¹³

Virus Structure

The influenza A virus belongs to the *Orthomyxoviridae* family. Other members of this group include influenza B and C viruses. The influenza A virus can further be classified according to the subtype of surface glycoproteins hemagglutinin (HA) and neuraminidase (NA). In total, there are 16 different subtypes of HA and 9 different subtypes of NA.¹⁴ When referencing specific strains of the virus, there is a widely accepted form, which includes identifying the strain of the virus (A, B, or C), the species it was isolated from (typically used if it was not isolated from humans), location of the species, the isolate number, and isolate year. Additionally, for strains of influenza A, the HA and NA subtypes are also given. As an example, A/Sydney/5/1997(H3N2) would indicate a strain of human influenza A was isolate number 5, found in Sydney, Australia in 1997. Furthermore, the virus had the HA subtype 3 and the NA subtype 2.^{14,15}

The influenza virus has one of two basic shapes, spherical or filamentous and has a diameter of roughly 100-300nm (**Figure 5.1**).¹⁴ It is also considered an envelope virus. There are a number of different proteins associated/found on the surface of the virus; the most well-known being the previously mentioned hemagglutinin and neuraminidase glycoproteins. Additional proteins include the surface matrix proteins 1 and 2 (M1 and M2).¹⁴ Each of these proteins is encoded by the 8 single stranded RNA segments of negative polarity, which are further wrapped in ribonucleic complexes.¹⁶ Additionally, these ribonucleoproteins are associated with the nucleoprotein and an RNA polymerase. This enzyme consists of three additional proteins: the polymerase basic protein 1 (PB1), polymerase basic protein 2 (PB2), and polymerase acidic protein (PA). All of these proteins have been shown to be important for viral replication.^{16,17} The RNP complex is

surrounded by the M1 protein which ‘envelopes’ the virus. Other proteins associated with the influenza virus are the M2 protein, the nuclear export protein (formally known as NS2), and the NS1 protein, which will be discussed later.¹⁸

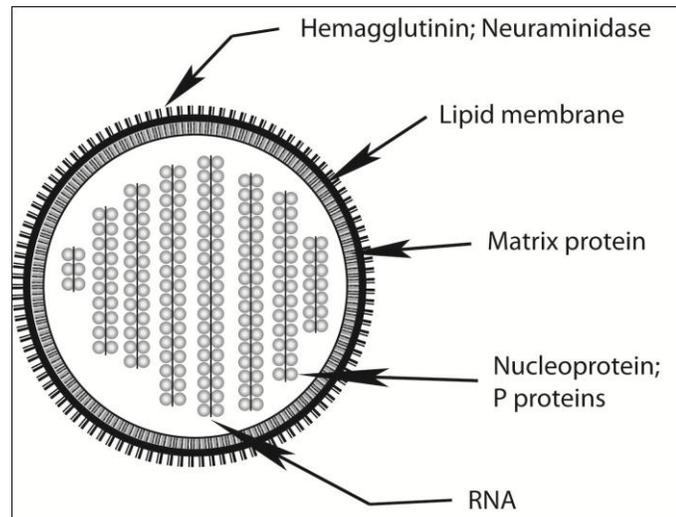


Figure 5.1 – Basic structure of the influenza A virus. Modified from Palese.¹⁹

Viral infection

Infection by the influenza virus is initiated by the recognition of sialic acid moieties on the host cell surface by the HA proteins of the virus. During viral replication, HA is cleaved during post-translational modifications into two different portions, HA1 and HA2 (though still connected through a disulfide bond). Each portion plays a specific part of the infection process. The HA1 portion is believed to be important for receptor binding and contains the antigenic sites, whereas the HA2 subunit is important to the fusion process between the virus envelope and the host cell membrane.²⁰ Once the virus has been internalized, the slightly acidic pH of the endosome (between 5 and 6), generates a conformational change in the HA which leads to the fusion of the viral membrane with the endosomal membrane allowing for release of the viral genome into

the host cell.²¹ Additionally, the M2 protein, which is an ion channel, pumps protons into the endosome to lower the pH which aids in the release and unpacking of the viral genome into the host cell.²²

Viral replication

Once the viral genome has been released into the cytoplasm, it is quickly transported to the nucleus of the host cell (**Figure 5.2**). The virus transports three different types of RNA into the nucleus which immediately start viral reproduction. The three types of RNA include mRNA, cRNA, and vRNA. Also transported into the nucleus is the vRNA dependent RNA polymerase. This enzyme uses the negative sense vRNA to synthesize mRNA and cRNA. Each of these types of RNA serves a specific purpose. The mRNA (which is in the positive sense), is used for the synthesis of viral proteins while the cRNA is used by the RNA polymerase enzyme to transcribe more copies of the negative sense viral genome.^{14,23}

Viral release

The final release of the virus from the host cell is a relatively straight forward process. Once the entire genome has been packed (all 8 negatively sensed strains), the process of budding occurs. It is believed that this process is initiated by an abundance of M1 proteins near the host cell wall. Once the process of budding has completed, the virus is still attached to the host cell via the HA and sialic acid interactions. It is at this point the neuraminidase protein cleaves this linkage and the virus is free to spread and

infect other healthy cells. It is the efficiency at which the influenza virus is able to evade common methods of treatment that makes it an annual problem.

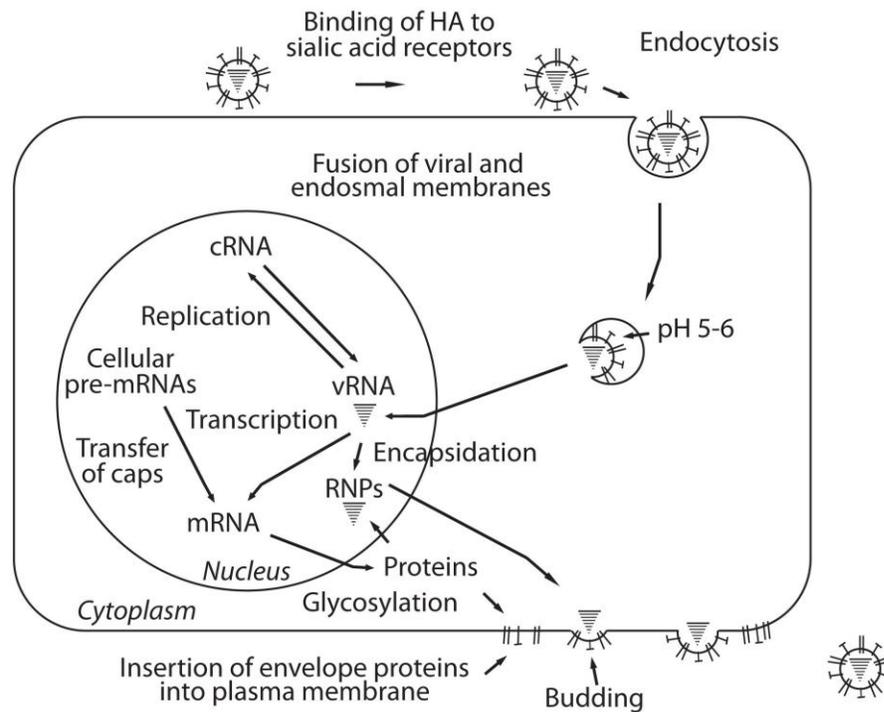


Figure 5.2 – Diagram of viral replication. Modified from Fodor *et al.*¹⁸

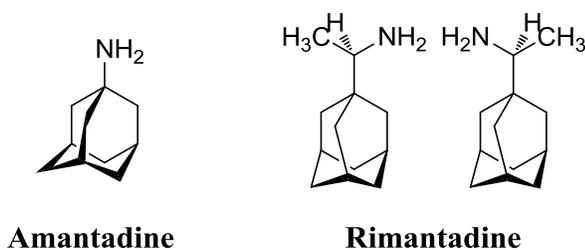
Previous treatment methods

Currently, the most common method of combating the virus is through vaccination. However, one of the biggest problems currently faced is the inability to fight off infections caused by pathogens which can escape from immunity induced by prior infection or vaccination.²⁴ Due to this antigenic drift, a new vaccine must be created every year to aid in combating the virus. Additionally, there are two classes of drugs on the market to help treat the flu: drugs that target the M2 ion channel and drugs which target neuraminidase proteins (**Table 5.1**).

Table 5.1 – Current drugs on the market or in development for the treatment of Influenza²⁵

Drug	Investigational Phase	Target
Oseltamivir	Market	Neuraminidase Enzyme
Zanamivir	Market	Neuraminidase Enzyme
Peramivir	Phase II-III	Neuraminidase Enzyme
Laninamivir	Phase III	Neuraminidase Enzyme
Amantadine	Market (CDC recommended against use)	M2 Ion Channel
Rimantadine	Market (CDC recommended against use)	M2 Ion Channel

In the mid-1960s, the US Food and Drug Administration (FDA) approved Amantadine[®] for the treatment and prevention of the Asian influenza strain.²⁶ This drug targets the M2 protein ion channel. M2 is a 97-residue, integral membrane protein which forms a pH-gated proton channel.²² The M2 ion channel is activated by the low pH within the endosome prior to hemagglutinin fusion. Once the proton channel is open, the influx of protons decreases the pH within the virus and subsequently dissociates the matrix protein from the viral nucleoproteins. This unpacking of the viral genome is essential for the promotion of viral replication.^{22,27} As previously mentioned, the use of Amantadine[®] was used as a M2 ion channel blocker (**Figure 5.3**) in an effort to treat influenza.

**Figure 5.3** – Structure of Amantadine[®] and Rimantadine[®]

Additionally, Rimantadine[®] was approved for use of treating influenza by the FDA in 1994. As with Amantadine[®], Rimantadine[®] targets the M2 ion channel. It was believed that by inhibiting this protein, that viral replication could be stopped. As with any elusive virus, a resistance was established and the virus became immune to the effects of both of these drugs. Amino acid mutation rendered the drugs ineffective while allowing the virus to act as normal.²⁸ Because of this growing resistance, the Center for Disease Control recommended against the use of these drugs to treat influenza.²⁵

In addition to drugs which target the M2 protein, another class of drugs targets the neuraminidase enzyme. The neuraminidase protein is one of the two most common surface glycoproteins found on the surface of the influenza virus. Furthermore, it accounts for 5-10% of the total viral proteins contained within the virus.²⁹ When a host cell is infected, the virus uses hemagglutinin to bind to sialic acid groups located on surface glycoproteins. In order for the virus to be released, neuraminidase must enzymatically cleave the connection between hemagglutinin and the sialic acid groups on the host glycoproteins.³⁰ Successful completion of this process allows for the influenza virus to be released and further infect healthy cells, thus ensuring the persistence of the virus. It was hypothesized that successful inhibition of the neuraminidase enzyme would prevent the release of the virus. In an effort to combat the virus in this fashion, a number of neuraminidase inhibitors were synthesized and put on the market.

One of the more well-known neuraminidase inhibitors is Oseltamivir (marketed as Tamiflu[®] by Roche). Tamiflu[®] in its marketed form is not active as a neuraminidase inhibitor. It is a prodrug which, when taken orally, is metabolized to its active form (**Figure 5.4**).

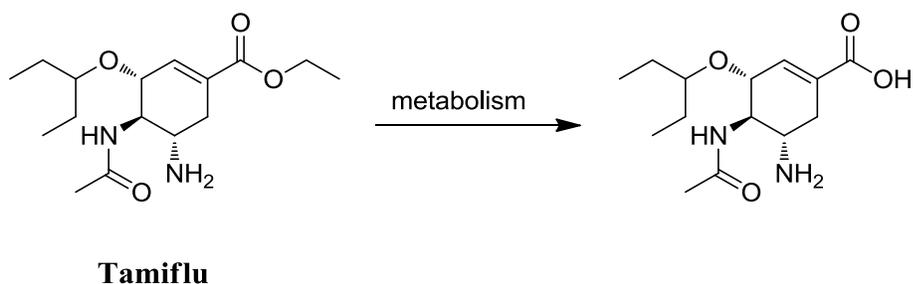


Figure 5.4 – Marketed form of Tamiflu[®] shown with its active metabolite.

Another commonly used neuraminidase inhibitor is Zanamivir (**Figure 5.5**). It is marketed by GlaxoSmithKline under the name Relenza[®]. Both Relenza[®] and Tamiflu[®] have shown to be an effective means of treating seasonal and 2009 pandemic strains of H1N1 influenza virus.³¹

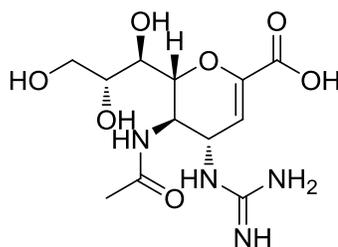


Figure 5.5 – Molecular structure of Relenza[®]

The main difference between Tamiflu[®] and its counterpart Relenza[®] is the method of administration. Where Tamiflu[®] is able to be taken orally; Relenza[®] is administered as a dry powder for inhalation. It can also be directly delivered to the lungs using an inhaler.³² Due to the prevalence of viral strains resistant to the M2 ion channel blockers, focus has been shifted towards the development of the neuraminidase inhibitors.³³ In addition to these current neuraminidase inhibitors, there are two potential drugs currently in Phase III clinical trials; Peramivir and Laninamivir. Peramivir had been developed

with the aim of being another orally available drug. However, despite promising *in vitro* and animal studies, phase III testing has demonstrated no statistically significant difference between treatment and placebo groups in the time required for resolving viral symptoms. It has been hypothesized that this is most likely due to the poor bioavailability in humans.²⁵ Further studies involving Laninamivir are also currently underway. Up to this point, neuraminidase inhibitors are required to be taken multiple times a week, sometimes multiple times a day. In an effort to make it easier for patients to comply with dosing responsibilities, a long acting inhibitor, Laninamivir, is currently in phase III clinical trials. It has shown to be effective as a once weekly treatment and preventative measure for influenza.³⁴

NS1 Protein

In order for the influenza virus to survive, it has to adapt and mutate in order to evade the immunological response from humans. Because of this ability, it becomes necessary to discover novel targets which can be used in the treatment of this virus. The Non Structural Protein 1 of Influenza A (NS1) has been identified as a new target at which to treat influenza. The NS1 protein has been observed to be the common link between all influenza A viruses as the way in which it blocks the host immune response.³⁵ In order to elude the host immune response, the NS1 contributes by participating in a wide array of actions. Included in these functions are: temporary regulation of RNA synthesis (which controls viral mRNA splicing and enhanced mRNA translation), blocking the host innate immune response, activation of phosphoinositide 3-kinase, and strain-dependent pathogenesis.^{35,36}

NS1 is a multifunctional protein which is approximately 230 amino acids long (may vary due to different strains) and weighs approximately 26kDa.³⁷ There are two important regions of this protein, the N-terminal domain which is widely accepted as the dsRNA binding domain and the C-terminal effector domain (**Figure 5.6**).

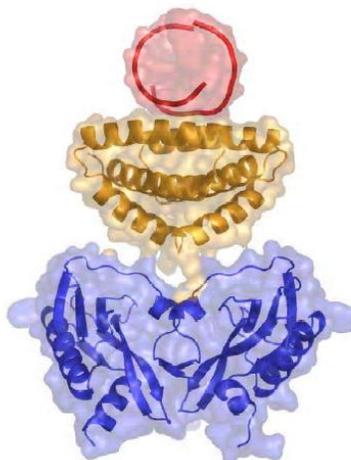


Figure 5.6 – Dimer model of NS1 where red is RNA, yellow is the RNA binding domain, and blue is the C-terminus effector domain. This figure was taken from Lin *et al.*³⁷

C-terminus and IFN

The 126 residue, C-terminal effector domain (residues 79-205) interacts with four different proteins; elongation initiation factor 4GI (eIF4DI), protein kinase R (PKR), poly(A)-binding protein II (PAB II) and more importantly the cleavage of the 30kDa subunit of the polyadenylation specificity factor (CPSF30).³⁸ It is now accepted that the interaction of NS1 with CPSF30 is responsible for the inhibition of the host cell to mature and export cellular antiviral mRNAs, which include those for interferon β .³⁹ When a cell becomes infected with a virus, the synthesis of specific cytokines called type I interferons (IFN) are produced. These cytokines are responsible for initiating the synthesis of antiviral proteins in neighboring cells which aid in preventing further

infection.⁴⁰ In addition to acting as an antiviral alarm, the production of interferons induces the maturation of dendritic cells. These cells are an important link between innate immunity and adaptive immunity which ultimately lead to the body clearing the virus.⁴⁰ Among other functions when the antigen is presented in the blood stream, the dendritic cells can take it up and recognize it as a foreign entity. Upon this recognition, a series of signaling cascades are initiated to alert the immune system to the presence of the virus. Actions of secreted type I interferons from virus infected cells enhance the antiviral response. If these actions are inhibited, then this can contribute to amplified virulence of the influenza virus. It is widely accepted that the NS1 protein of influenza A is crucial to evading the antiviral response. Therefore, inhibition of this protein should aid in the ability of type I interferons to effectively alert the immune system to the presence of the virus.

The RNA binding domain

The second domain associated with NS1 is the N-terminal 'RNA Binding' domain. This portion of NS1 exists as a symmetric homodimer with a unique six-helical fold.⁴¹ The RNA binding domain consists of residues 1-70 and is responsible for binding double-stranded RNA (dsRNA) in a non-sequence specific manner.⁴² Studies have demonstrated the binding pocket for RNA contains a number of basic residues which interact directly with the negatively charged phosphate backbone of dsRNA. These residues were identified and targeted for an alanine scan to determine which (if any) were essential for binding RNA. Wang and coworkers identified the arginine in position 38 as being 'absolutely essential' for RNA binding under all of their tested conditions.⁴³ Single

amino acid replacement of Arg38/Arg38' demonstrated high rates of attenuation, which indicates the binding of RNA to NS1 is essential for viral replication.^{41,43,44}

Previously, it was mentioned that interferons are essential cytokines involved in the antiviral response of the innate immune system. Additionally, the C-terminal effector domain was shown play a key role in preventing the export of antiviral mRNA for IFN- β . The N-terminal RNA binding domain, however, does not participate in the blocking of the IFN- β mRNA. Instead, it is responsible for the virus to evade the 'antiviral state' which is produced by IFN- β . The virus is able to survive in this 'antiviral state' by blocking the 2'-5'-oligo (A) synthase (OAS)/RNase L pathway which is induced by IFN- α/β .⁴¹ This pathway is a key part to the innate immune response to viral infection. This response to foreign entities is responsible for the degradation of viral and cellular RNA, which aid in preventing the spread of viral growth.⁴⁵ Furthermore, it was shown that cells that have been exposed to the IFN signaling pathway have elevated levels of OAS, adding to the 'antiviral state.' This state also includes the stimulation of the RNase L pathway. The purpose of this signaling pathway is to release ribonucleases which, when activated, destroy all viral RNA present within the cell.⁴⁶ The way in which NS1 evades the response from the innate immune response is by isolating viral dsRNA from the OAS, thus preventing activation of the RNase L pathway.⁴⁵ By blocking the action of NS1, these signaling pathways can occur, leading to the prevention of viral growth.

The NS1 Protein as a Novel Target

Based on previously reported data, the binding of RNA to the NS1 protein is essential for viral replication. Single amino acid replacement of Arg38 with an alanine

lead to high virus attenuation.⁴⁴ Furthermore, NS1 was shown to help evade the antiviral response by sequestering dsRNA from OAS which prevents its destruction by the RNase L Pathway. Given this, it reasons that inhibition of the interaction between dsRNA and NS1 provides a way to inhibit viral growth. Recently, it has been demonstrated that NS1 interacts with TRIM25. This interaction prevents the activation of retinoic acid-inducible gene I (RIG-I), which suppresses the stimulation and synthesis of cellular IFN.^{47,48,49} Therefore, the NS1 protein is a viable, novel target for the inhibition of growth by the Influenza A virus.

Due to the ability of the influenza virus to mutate and become resistant to current treatment options, there is a growing need for the development of a broad spectrum treatment. This way, regardless of whether or not the virus has mutated, the treatment should still be effective. The NS1 protein changes very little from mutated virus to mutated virus, which makes it attractive for a drug discovery program focused on finding an effective treatment for influenza.⁵⁰

Recent studies have demonstrated the viability of a novel yeast based assay to determine the activity of small molecules at NS1. This study included a screen of 2000 compounds from the National Cancer Institute Diversity Set (NCS) against NS1. From this screen, fifteen compounds were identified as active hits, and of these fifteen, four were carried through for further studies due to their ability to inhibit replication of the influenza virus (**Figure 5.7** and **Table 5.2**).¹

Table 5.2 – 50% Inhibitory concentrations and selective indexes of anti-NS1 compounds on different influenza strains cultured in MDCK-UK cells. (Modified from Basu *et al.*)

Anti-NS1 Compound	IC ₅₀ (μM)/SI		
	PR	HK	WSN
NSC128164	14/16.6	5/44.4	13/18.5
NSC109834	10/32.3	2/200.3	12/28.2
NSC95676	12/104.9	20/62.8	19/64.5
NSC125044	11/12.4	7/18.8	12/11.8

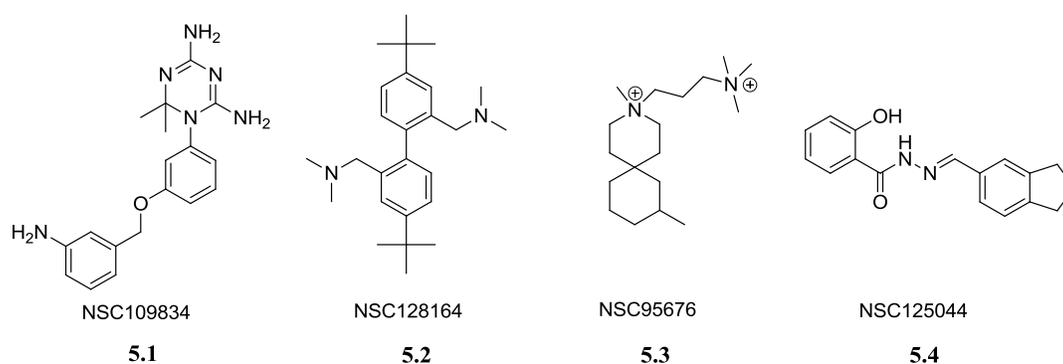


Figure 5.7 – Structures of the compounds which showed anti-influenza activity in the initial screen.¹

Object of this Study

The influenza A virus is responsible for each of the major pandemics over the last 100 years, including the 1918 Spanish Flu pandemic and the 2009 Swine flu pandemic. The most popular method of combating the virus is through vaccination. Other strategies include the use of antiviral drugs such as Tamiflu to treat infection. However, antigenic drift dictates new vaccines must be created each year, and emerging strains of the virus are becoming resistant to antiviral drugs. Because of this, there is a need to discover a target which changes very little from virus to virus.

The NS1 protein of influenza A has been identified as a novel target for the treatment of influenza. One of the main functions of this protein is to aid in evading the antiviral response. Furthermore, studies have shown that the sequence structure of NS1 changes very little from strain to strain, making it a desirable target to pursue.

It is the goal of this study to discover and develop molecule which can display biological activity at a number of strains of the influenza A virus. Four compounds which were identified from a library screen will be narrowed down to one lead compound. From here, simple modifications can be made to develop a comprehensive structure activity relationship (SAR) to determine key molecular features required for biological activity. Additionally, once a compound has been identified as displaying broad spectrum activity, the SAR can further be explored through the use of library synthesis.

Chapter 5 References

1. Basu, D.; Walkiewicz, M. P.; Frieman, M.; Baric, R. S.; Auble, D. T.; Engel, D. A., Novel Influenza Virus NS1 Antagonists Block Replication and Restore Innate Immune Function. *J. Virol.* **2009**, *83* (4), 1881-1891.
2. Thompson, W. W.; Shay, D. K.; Weintraub, E.; Brammer, L.; Cox, N.; Anderson, L. J.; Fukuda, K., Mortality Associated With Influenza and Respiratory Syncytial Virus in the United States. *JAMA: The Journal of the American Medical Association* **2003**, *289* (2), 179-186.
3. Thompson, W. W.; Shay, D. K.; Weintraub, E.; Brammer, L.; Bridges, C. B.; Cox, N. J.; Fukuda, K., Influenza-Associated Hospitalizations in the United States. *JAMA: The Journal of the American Medical Association* **2004**, *292* (11), 1333-1340.
4. Simonsen, L.; Reichert, T. A.; Viboud, C.; Blackwelder, W. C.; Taylor, R. J.; Miller, M. A., Impact of Influenza Vaccination on Seasonal Mortality in the US Elderly Population. *Arch Intern Med* **2005**, *165* (3), 265-272.
5. Reid, A. H.; Taubenberger, J. K.; Fanning, T. G., The 1918 Spanish influenza: integrating history and biology. *Microbes and Infection* **2001**, *3* (1), 81-87.
6. Vittorio Farina, J. D. B., Tamiflu: The Supply Problem. *Angewandte Chemie International Edition* **2006**, *45* (44), 7330-7334.
7. Chen, H.; Smith, G. J. D.; Li, K. S.; Wang, J.; Fan, X. H.; Rayner, J. M.; Vijaykrishna, D.; Zhang, J. X.; Zhang, L. J.; Guo, C. T.; Cheung, C. L.; Xu, K. M.; Duan, L.; Huang, K.; Qin, K.; Leung, Y. H. C.; Wu, W. L.; Lu, H. R.; Chen, Y.; Xia, N. S.; Naipospos, T. S. P.; Yuen, K. Y.; Hassan, S. S.; Bahri, S.; Nguyen, T. D.; Webster, R. G.; Peiris, J. S. M.; Guan, Y., Establishment of multiple sublineages of H5N1 influenza virus

in Asia: Implications for pandemic control. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103* (8), 2845-2850.

8. Ducatez, M. F.; Olinger, C. M.; Owoade, A. A.; De Landtsheer, S.; Ammerlaan, W.; Niesters, H. G. M.; Osterhaus, A. D. M. E.; Fouchier, R. A. M.; Muller, C. P., Avian Flu: Multiple introductions of H5N1 in Nigeria. *Nature* **2006**, *442* (7098), 37-37.

9. Olsen, B.; Munster, V. J.; Wallensten, A.; Waldenström, J.; Osterhaus, A. D. M. E.; Fouchier, R. A. M., Global Patterns of Influenza A Virus in Wild Birds. *Science* **2006**, *312* (5772), 384-388.

10. Hayman, A.; Comely, S.; Lackenby, A.; Hartgroves, L. C. S.; Goodbourn, S.; McCauley, J. W.; Barclay, W. S., NS1 Proteins of Avian Influenza A Viruses Can Act as Antagonists of the Human Alpha/Beta Interferon Response. *J. Virol.* **2007**, *81* (5), 2318-2327.

11. Abdel-Ghafar, A. N., T. Chotpitayasundh, Z. Gao, F. G. Hayden, D. H. Nguyen, M. D. de Jong, A. Naghdaliyev, J. S. Peiris, N. Shindo, S. Soerooso, T. M. Uyeki, Update on Avian Influenza A (H5N1) Virus Infection in Humans. *N Engl J Med* **2008**, *358* (3), 261-273.

12. Chang, L.-Y.; Shih, S.-R.; Shao, P.-L.; Huang, D. T.-N.; Huang, L.-M., Novel Swine-origin Influenza Virus A (H1N1): The First Pandemic of the 21st Century. *Journal of the Formosan Medical Association* **2009**, *108* (7), 526-532.

13. Girard, M. P.; Tam, J. S.; Assossou, O. M.; Kieny, M. P., The 2009 A (H1N1) influenza virus pandemic: A review. *Vaccine* **2010**, *28* (31), 4895-4902.

14. Bouvier, N. M.; Palese, P., The biology of influenza viruses. *Vaccine* **2008**, *26* (Supplement 4), D49-D53.

15. A J Hay, V. G., A R Douglas, and Y P Lin, The evolution of human influenza viruses. *Philos Trans R Soc Lond B Biol Sci* **2001**, 356 (1416), 1861-1870.
16. Ortega, J.; Martin-Benito, J.; Zurcher, T.; Valpuesta, J. M.; Carrascosa, J. L.; Ortin, J., Ultrastructural and Functional Analyses of Recombinant Influenza Virus Ribonucleoproteins Suggest Dimerization of Nucleoprotein during Virus Amplification. *J. Virol.* **2000**, 74 (1), 156-163.
17. Perales, B.; Ortin, J., The influenza A virus PB2 polymerase subunit is required for the replication of viral RNA. *J. Virol.* **1997**, 71 (2), 1381-1385.
18. Fodor, E.; Brownlee, G. G., Influenza virus replication. In *Perspectives in Medical Virology*, Potter, C. W., Ed. Elsevier: 2002; Vol. Volume 7, pp 1-29.
19. Palese, P., The genes of influenza virus. *Cell* **1977**, 10 (1), 1-10.
20. Steinhauer, D. A., Role of Hemagglutinin Cleavage for the Pathogenicity of Influenza Virus. *Virology* **1999**, 258 (1), 1-20.
21. Stegmann, T., Membrane Fusion Mechanisms: The Influenza Hemagglutinin Paradigm and its Implications for Intracellular Fusion. *Traffic* **2000**, 1 (8), 598-604.
22. Schnell, J. R.; Chou, J. J., Structure and mechanism of the M2 proton channel of influenza A virus. *Nature* **2008**, 451 (7178), 591-595.
23. Cros, J. F.; Palese, P., Trafficking of viral genomic RNA into and out of the nucleus: influenza, Thogoto and Borna disease viruses. *Virus Research* **2003**, 95 (1-2), 3-12.
24. Smith, D. J.; Lapedes, A. S.; de Jong, J. C.; Bestebroer, T. M.; Rimmelzwaan, G. F.; Osterhaus, A. D. M. E.; Fouchier, R. A. M., Mapping the Antigenic and Genetic Evolution of Influenza Virus. *Science* **2004**, 305 (5682), 371-376.

25. Boltz, D. A.; Aldridge, J. R. J.; Webster, R. G.; Govorkova, E. A., Drugs in Development for Influenza. *Drugs* **2010**, *70* (11), 1349-1362 10.2165/11537960-000000000-00000.
26. Maugh, T. H., Amantadine: An Alternative for Prevention of Influenza. *Science* **1976**, *192* (4235), 130-131.
27. Helenius, A., Unpacking the incoming influenza virus. *Cell* **1992**, *69* (4), 577-578.
28. Jing, X.; Ma, C.; Ohigashi, Y.; Oliveira, F. A.; Jardetzky, T. S.; Pinto, L. H.; Lamb, R. A., Functional studies indicate amantadine binds to the pore of the influenza A virus M2 proton-selective ion channel. *Proceedings of the National Academy of Sciences* **2008**, *105* (31), 10967-10972.
29. Bossart-Whitaker, P.; Carson, M.; Babu, Y. S.; Smith, C. D.; Laver, W. G.; Air, G. M., Three-dimensional Structure of Influenza A N9 Neuraminidase and Its Complex with the Inhibitor 2-Deoxy 2,3-Dehydro-N-Acetyl Neuraminic Acid. *Journal of Molecular Biology* **1993**, *232* (4), 1069-1083.
30. Huang, I. C. L., W.; Sui, J.; Marasco, W.; Choe, H.; Farzan, M., Influenza A Virus Neuraminidase Limits Viral Superinfection. *Journal of Virology* **2008**, *82* (10), 4834-4843.
31. Nguyen, H. T.; Sheu, T. G.; Mishin, V. P.; Klimov, A. I.; Gubareva, L. V., Assessment of pandemic and seasonal influenza A (H1N1) virus susceptibility to neuraminidase inhibitors in three enzyme activity inhibition assays. *Antimicrob. Agents Chemother.* **2010**, AAC.00581-10.

32. Moscona, A., Neuraminidase Inhibitors for Influenza. *New England Journal of Medicine* **2005**, 353 (13), 1363-1373.
33. Smee, D. F.; Hurst, B. L.; Wong, M.-H.; Tarbet, E. B.; Babu, Y. S.; Klumpp, K.; Morrey, J. D., Combinations of oseltamivir and peramivir for the treatment of influenza A (H1N1) virus infections in cell culture and in mice. *Antiviral Research In Press, Uncorrected Proof*.
34. Kubo, S.; Tomozawa, T.; Kakuta, M.; Tokumitsu, A.; Yamashita, M., Laninamivir Prodrug CS-8958, a Long-Acting Neuraminidase Inhibitor, Shows Superior Anti-Influenza Virus Activity after a Single Administration. *Antimicrob. Agents Chemother.* **2010**, 54 (3), 1256-1264.
35. Hale, B. G.; Randall, R. E.; Ortin, J.; Jackson, D., The multifunctional NS1 protein of influenza A viruses. *J Gen Virol* **2008**, 89 (10), 2359-2376.
36. García-Sastre, A.; Egorov, A.; Matassov, D.; Brandt, S.; Levy, D. E.; Durbin, J. E.; Palese, P.; Muster, T., Influenza A Virus Lacking the NS1 Gene Replicates in Interferon-Deficient Systems. *Virology* **1998**, 252 (2), 324-330.
37. Lin, D.; Lan, J.; Zhang, Z., Structure and Function of the NS1 Protein of Influenza A Virus. *Acta Biochim Biophys Sin* **2007**, 39 (3), 155-162.
38. Xia, S.; Monzingo, A. F.; Robertus, J. D., Structure of NS1A effector domain from the influenza A/Udorn/72 virus. *Acta Crystallographica Section D* **2009**, 65 (1), 11-17.
39. Xia, S.; Robertus, J. D., X-ray structures of NS1 effector domain mutants. *Archives of Biochemistry and Biophysics* **2010**, 494 (2), 198-204.

40. Haye, K.; Burmakina, S.; Moran, T.; Garcia-Sastre, A.; Fernandez-Sesma, A., The NS1 Protein of a Human Influenza Virus Inhibits Type I Interferon Production and the Induction of Antiviral Responses in Primary Human Dendritic and Respiratory Epithelial Cells. *J. Virol.* **2009**, *83* (13), 6849-6862.
41. Min, J.-Y.; Krug, R. M., The primary function of RNA binding by the influenza A virus NS1 protein in infected cells: Inhibiting the 2'-5' oligo (A) synthetase/RNase L pathway. *Proceedings of the National Academy of Sciences* **2006**, *103* (18), 7100-7105.
42. Das, K.; Aramini, J. M.; Ma, L.-C.; Krug, R. M.; Arnold, E., Structures of influenza A proteins and insights into antiviral drug targets. *Nat Struct Mol Biol* **2010**, *17* (5), 530-538.
43. Wang, W.; Riedel, K.; Lynch, P.; Chien, C. Y.; Montelione, G. T.; Krug, R. M., RNA binding by the novel helical domain of the influenza virus NS1 protein requires its dimer structure and a small number of specific basic amino acids. *RNA* **1999**, *5* (2), 195-205.
44. Cheng, A.; Wong, S. M.; Yuan, Y. A., Structural basis for dsRNA recognition by NS1 protein of influenza A virus. *Cell Res* **2009**, *19* (2), 187-195.
45. Silverman, R. H., Viral Encounters with 2',5'-Oligoadenylate Synthetase and RNase L during the Interferon Antiviral Response. *J. Virol.* **2007**, *81* (23), 12720-12729.
46. Li, X.-L.; Blackford, J. A.; Hassel, B. A., RNase L Mediates the Antiviral Effect of Interferon through a Selective Reduction in Viral RNA during Encephalomyocarditis Virus Infection. *J. Virol.* **1998**, *72* (4), 2752-2759.
47. Guo, Z.; Chen, L.-m.; Zeng, H.; Gomez, J. A.; Plowden, J.; Fujita, T.; Katz, J. M.; Donis, R. O.; Sambhara, S., NS1 Protein of Influenza A Virus Inhibits the Function of

Intracytoplasmic Pathogen Sensor, RIG-I. *Am. J. Respir. Cell Mol. Biol.* **2007**, *36* (3), 263-269.

48. Gack, M. U.; Albrecht, R. A.; Urano, T.; Inn, K.-S.; Huang, I. C.; Carnero, E.; Farzan, M.; Inoue, S.; Jung, J. U.; García-Sastre, A., Influenza A Virus NS1 Targets the Ubiquitin Ligase TRIM25 to Evade Recognition by the Host Viral RNA Sensor RIG-I. *Cell host & microbe* **2009**, *5* (5), 439-449.

49. Ludwig, S.; Wolff, T., Influenza A Virus TRIMs the Type I Interferon Response. *Cell host & microbe* **2009**, *5* (5), 420-421.

50. Shaw, M. W.; Lamon, E. W.; Compans, R. W., Immunologic studies on the influenza A virus nonstructural protein NS1. *The Journal of Experimental Medicine* **1982**, *156* (1), 243-254.

Chapter 6:
NS1 Results

Compound Assessment

Initial assessments of synthesized compounds were based on single point concentration comparison with the parent compound **6.1** in the form of LogTCID₅₀ to developing a structure activity relationship. In using this method, lower LogTCID₅₀ values indicated greater inhibition of virus replication. With the synthesis of more analogues, analysis moved toward using a dose-response method of assessing activity. It should be noted however, that this cell based assay is dependent upon the ability of the compound to enter the cell as well as the affinity of the compound for the target. Examination of the data obtained from the assay would suggest that the Log₁₀[Virus titre] versus concentration of compound approximated an exponential fall off to some residual virus replication at higher concentrations. To validate this as a viable way to make assessments surrounding activity, three compounds were selected and assayed in triplicate and typical results are shown in **Figure 6.1**. The data was plotted as a measure of concentration vs. Log₁₀[Virus titre] and was able to fit an exponential decay model. This result allowed for two different measures of activity to be established, a K-value which is a function of the slope, and the span, which represents the maximum reduction in virus replication by comparing control Log₁₀[Virus titre] values and those observed at high concentrations.

Generally, the K-value was a good measure to compare the activity of tested analogues. In some instances, however, when the data was plotted, compounds which displayed high K-values (good activity) did not exhibit a large span (overall reduction in virus titre). In these instances, both the K-value and the span were used as an indicator of relative activity to make decisions regarding future analogues. A good example of this

phenomena was observed for the halogenated analogues (**Figure 6.8**). Here the iodine substitution was more active than substitution by the other halogens at low concentrations, but gave a lower overall reduction in virus titre at than the chlorine and bromine compounds at higher concentrations.

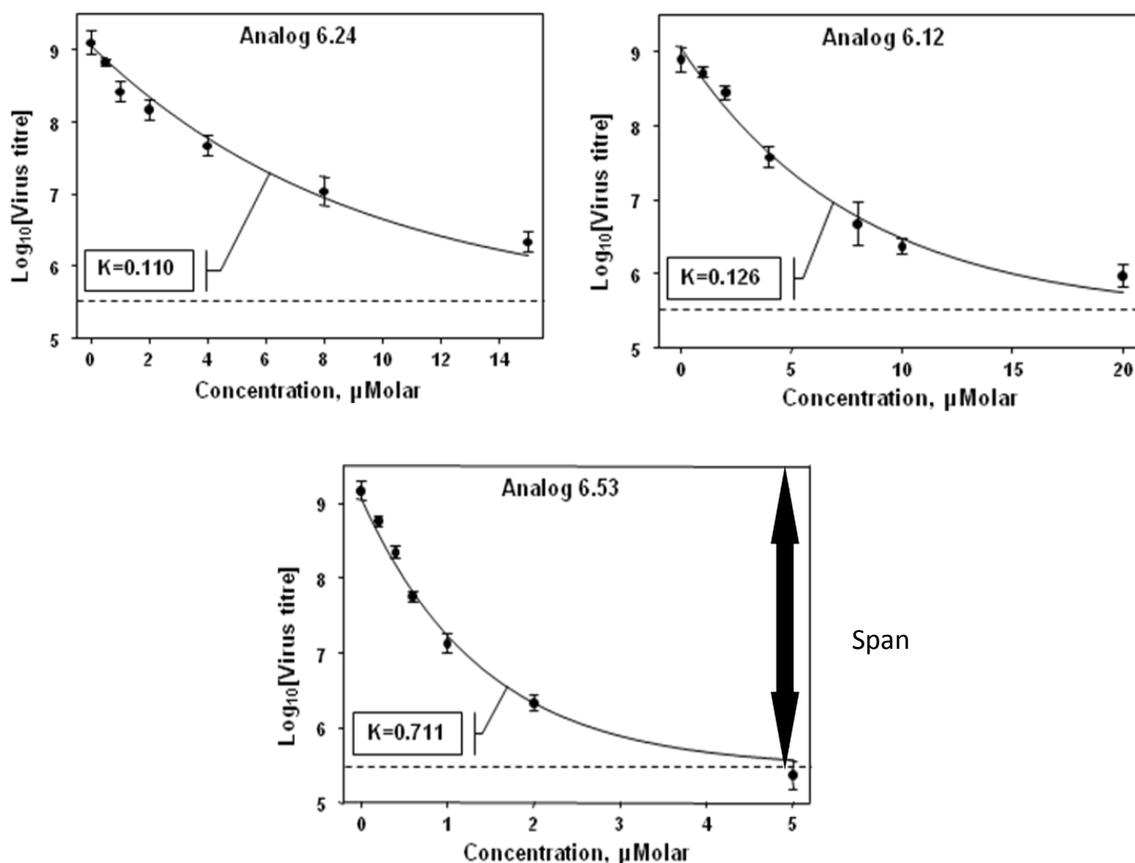
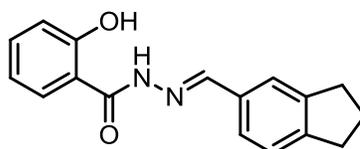


Figure 6.1 – Plotted biological data used to make assessments of activity using K-values and Span

Modification of the Indan Ring and Hydrazine Connector

It is the goal of this study to develop a comprehensive structure activity relationship (SAR) and to discover a compound which can be considered a viable

candidate for a drug discovery program. Upon closer inspection of the four compounds identified from the NSC screen, all of them contain undesirable properties and characteristics which would need to be addressed in order to develop a program around them.



NSC125044

6.1

Figure 6.2 – Compound from initial screen which was used as the parent molecule

Compound **6.1** has been selected lead for development of a comprehensive SAR (**Figure 6.2**). This molecule can be defined as a three component molecule, as there are three building blocks used for the synthesis. Initial studies examined if the indan ring system in the parent molecule is required for biological activity. This was identified as the least amenable to chemical manipulation and/or in the commercial availability of monomers. The first series of compounds was synthesized to examine the necessity of the indan ring to virus replication. This was accomplished through the use of benzaldehyde derivatives which display a number of different functional groups at the para position (**Figure 6.3**).

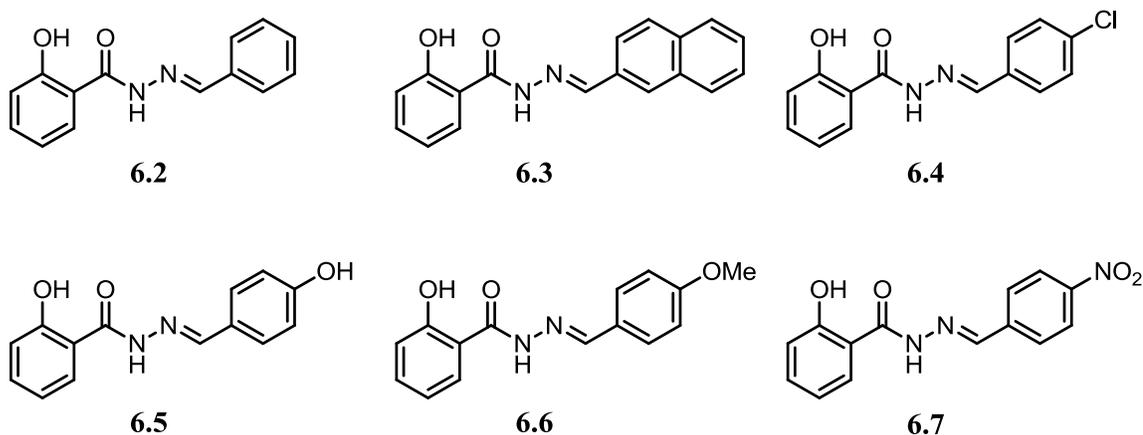
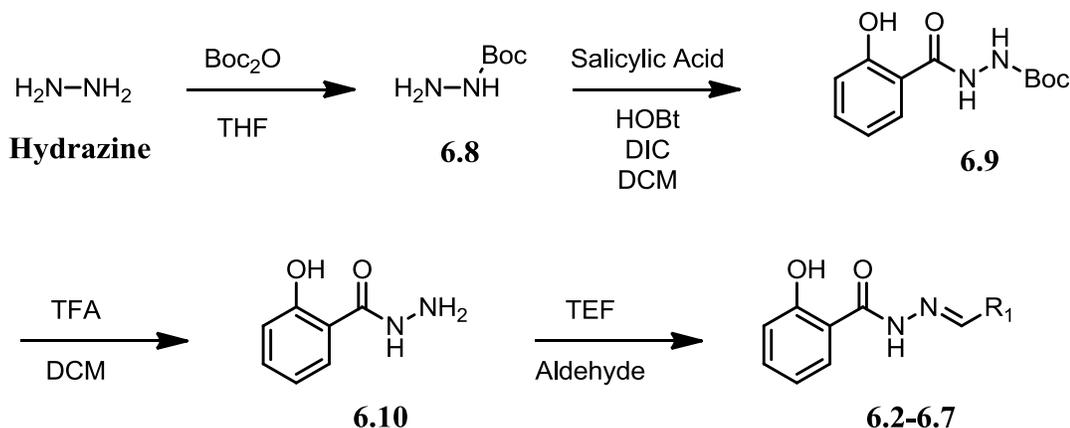


Figure 6.3 – Synthesized analogues based on parent compound **6.1**

Table 6.1 – LogTCID₅₀ values for synthesized compounds **6.2-6.7** and **6.11-6.19**

Compound	LogTCID ₅₀	Compound	LogTCID ₅₀
6.1 (Parent)	7.1	6.12	5.7
6.2	7.5	6.13	5.2
6.3	7.3	6.14	6.8
6.4	7.2	6.15	6.2
6.5	7.4	6.16	6.5
6.6	7.2	6.17	5.0
6.7	6.9	6.18	5.8
6.11	6.2	6.19	6.5

The synthesis of these analogues began by mono-Boc protecting hydrazine. Following this, the hydroxyl group was introduced to the molecule through a standard HOBt/DIC amide bond formation between the Boc-protected hydrazine and salicylic acid. Removal of the Boc protecting group was achieved through the use of trifluoroacetic acid (TFA). The last step in the synthesis, which introduces the imine and aromatic groups, was performed through the use of benzaldehyde derivatives and triethyl orthoformate as a dehydrating agent (**Scheme 6.1**).

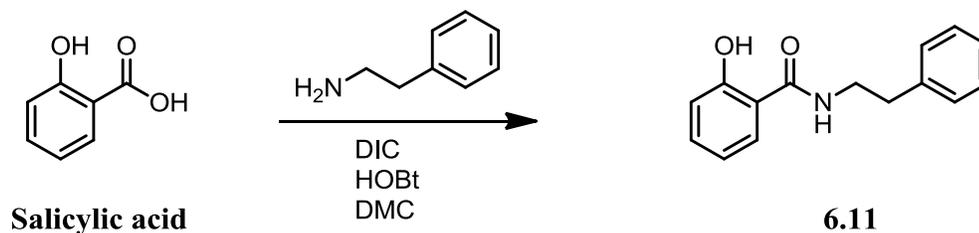
Scheme 6.1 – Synthesis of imine analogues

The way in which these initial compounds were analyzed was to use a single point concentration comparison with the parent compound **6.1**. Replacement of the five-membered ring with a variety of alternative functional groups including electron donating or withdrawing substituents, as well as hydrogen bond donating and accepting functional groups generally led to decreased inhibitory activity (**Table 6.1**). Despite this, it was established that the indan functionality was not a necessity for antiviral activity. By determining this, the available monomer sets available for synthesis drastically increased.

The least attractive part of the parent molecule is the hydrazine core. Due to the high reactive nature of this functional group, it is likely to contribute to a poor pharmacokinetic profile. Additionally, the hydrazine core adds structural rigidity to the molecule which at this point in compound development may be a positive or negative feature. In an attempt to simplify the linking region of the parent molecule, the hydrazine was replaced with an alkyl chain while maintaining the amide nitrogen. Phenethylamine was coupled to salicylic acid using standard HOBt/DIC coupling conditions (**Scheme**

6.2). This building block was used specifically due to the fact that it preserves the bond distance between the two aromatic rings.

Scheme 6.2 – Synthesis of analogues with the removal of the hydrazine linker



By substituting the carbon chain for the hydrazine imine, a small increase in activity was observed at high concentrations ($\text{LogTCID}_{50}[\mathbf{6.11}] = 6.2$). This result was a positive step forward in making the molecule a more attractive candidate by removing a reactive functional group. Subsequent modifications were made to optimize the length of the alkyl connector region (**Figure 6.4**). Lengths of three and four carbons each showed increased activity when compared to the two carbon version (**6.11**) ($\text{LogTCID}_{50}[\mathbf{6.12}] = 5.7$ and $\text{LogTCID}_{50}[\mathbf{6.13}] = 5.2$). With this result, the synthesis of molecules was streamlined from a four step process to a one step process. At this point the new parent molecule can be classified as a two component molecule with a defined left (hydroxyl) and right side (unsubstituted phenyl ring) to the molecule. Based on cost and commercial availability, **6.12** was used as the new parent molecule to base further studies on.

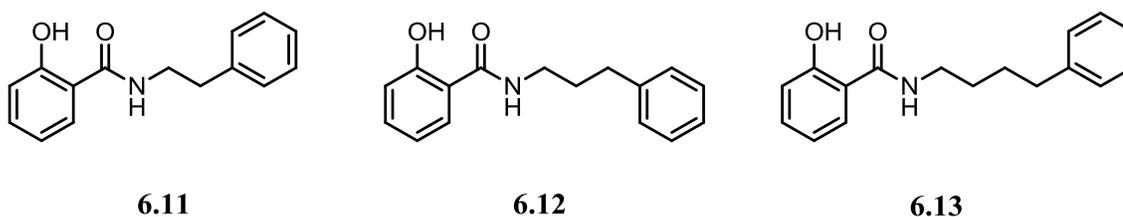


Figure 6.4 – Analogues which were synthesized to optimize the carbon chain length

In what will be called the left side of the molecule (**Figure 6.5**), the hydroxyl group which is ortho to the amide carbonyl was introduced using salicylic acid as a building block. The hydroxyl group can participate in a number of different interactions. It is an electron donating substituent which can participate in both hydrogen bond donating and accepting interactions. Both of these could be important for biological activity and needed to be explored further.

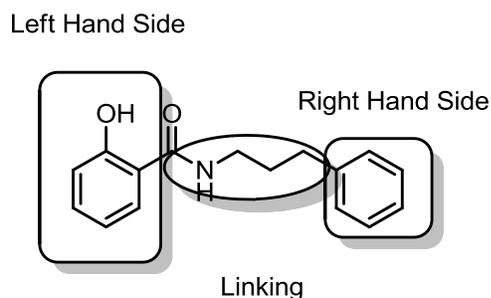


Figure 6.5 – Molecule 6.12 broken up into three defined sections.

Exploration of the Left Hand Side of the Molecule

Each of the modifications to the parent can be made by using benzoic acid and 3- and 4-hydroxy benzoic acid during the synthesis (**Figure 6.6**). By comparing the LogTCID₅₀ values, the activity resulting from these modifications decreased when compared to the parent molecule 6.12. This would indicate the position ortho to the

amide carbonyl may be involved with important interactions associated with antiviral activity.

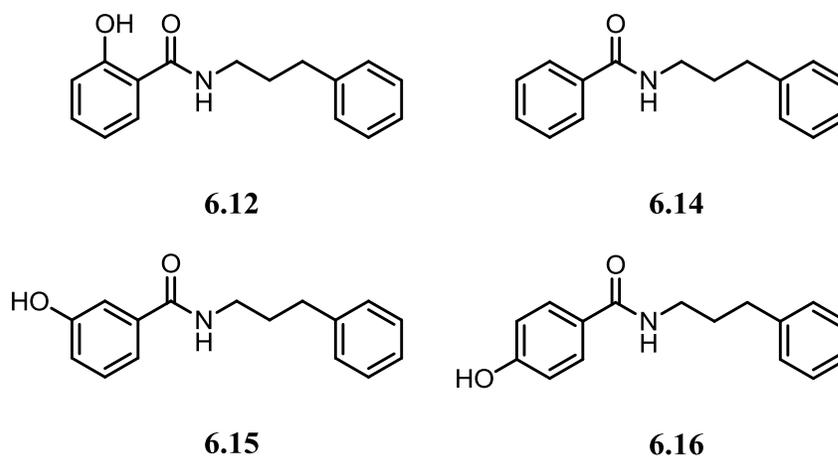


Figure 6.6 – Modifications made to **6.12** to optimize the hydroxyl group positioning

Once it had been determined that the optimum substitution position is ortho to the amide carbonyl, a series of ortho substituted benzoic acid derivatives were used in the synthesis of analogues **6.17-6.25** (**Figure 6.7**) to examine the requirement for the hydroxyl group. Contained within this set of analogues, are functional groups which have the ability to participate in H-bonding donating/accepting, as well as a range of strongly electron donating to withdrawing groups. Most of these compounds displayed a decrease in activity when compared to the parent molecule **6.12** ($K[\mathbf{6.12}] = 0.13$). With compound **6.21**, a significant increase in activity was observed ($K[\mathbf{6.21}] = 1.4$) when the hydroxyl group was substituted with an amino group. If the characteristics of the hydroxyl group and the amino group are compared, there are some similarities. Both of these functional groups are electron donating groups which affect the electronics of the phenyl ring as well as the chemistry of the amide. The hydroxyl group and the amino

group can participate in hydrogen bonding through donating and accepting interactions. Furthermore, the size difference between the amine and the hydroxyl group is negligible. This would indicate that the hydrogen bonding interactions may be important to antiviral activity. The main difference between the two is that the hydroxyl group is slightly acidic and the amino group is weakly basic.

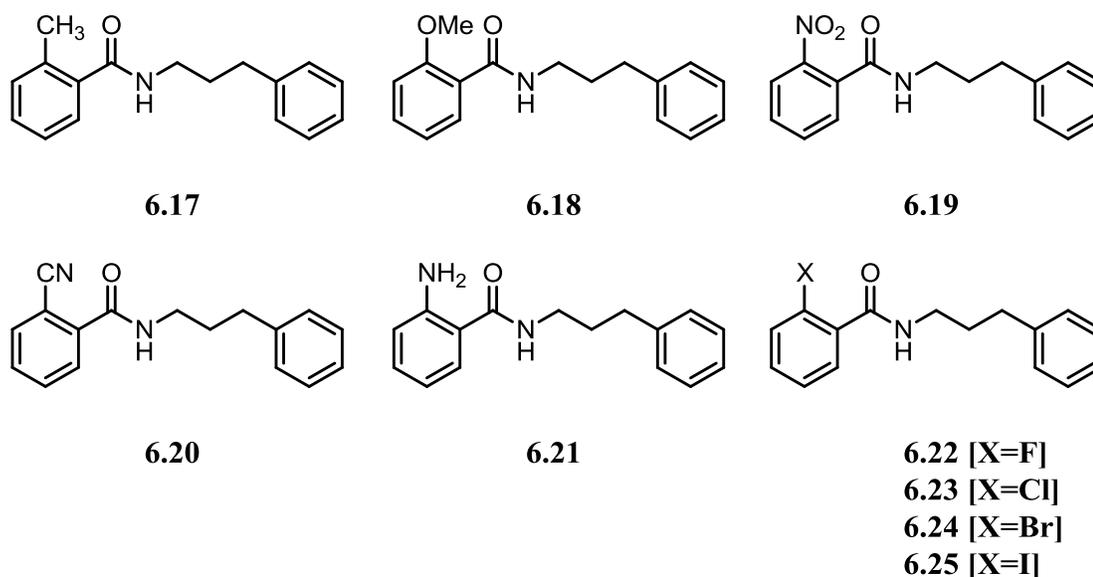


Figure 6.7 – Analogues synthesized where the hydroxyl is substituted with alternative functional groups

The activity profile based on K-values for the halogenated compounds would suggest that the iodinated compound was the most active, $K[6.25] = 0.86$. However, when the data was plotted, it is apparent that although the iodinated compound had the largest K-value, the chlorine derivative had a larger span and was therefore deemed more active (**Figure 6.8**). Because of results like this, assessing the activity of compounds needs to be made in the context of both the K-value and the span. Moreover, these results would seem to rule out hydrogen bonding, either as an acceptor or donator, as a

requirement for activity. Additionally, the results of the halogen series further probed the size restrictions associated with the ortho position, suggesting that medium sized functional groups such as chlorine or bromine were optimum for activity.

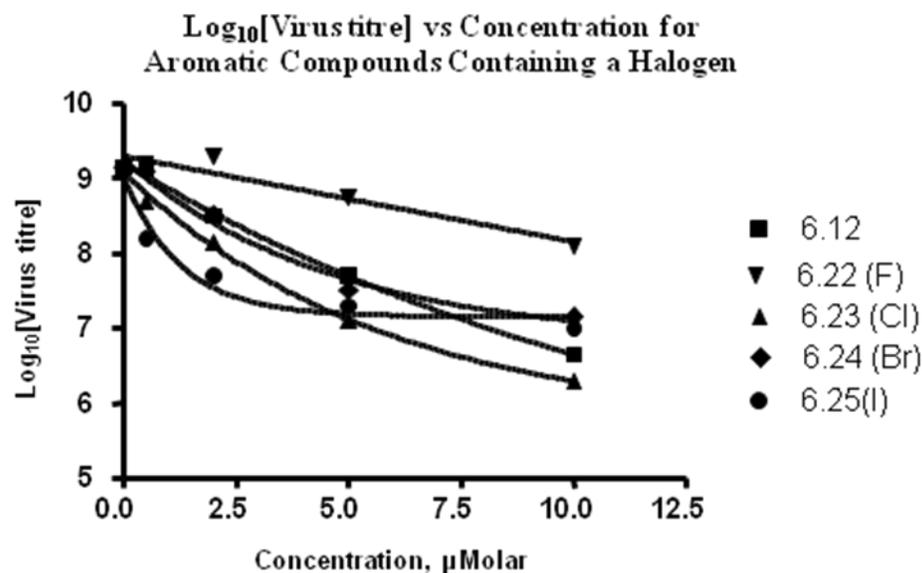


Figure 6.8 – Plotted data for halogenated compounds vs. parent molecule **6.12**

After establishing the significance of substitution ortho to the amide, exploration was moved towards examining size restrictions associated with the left hand side of the molecule. The three structural isomers of ortho hydroxy-naphthoic acid were used in the synthesis of the next set of compounds (**Figure 6.9**). These were selected because they did not change the position of the hydroxyl group, maintaining the consistency of the parent compound, and projected the second ring in different directions probing both size and shape of the target binding site. By testing these three compounds, it is ensured that all possible locations for the steric bulk added by naphthalene can be explored.

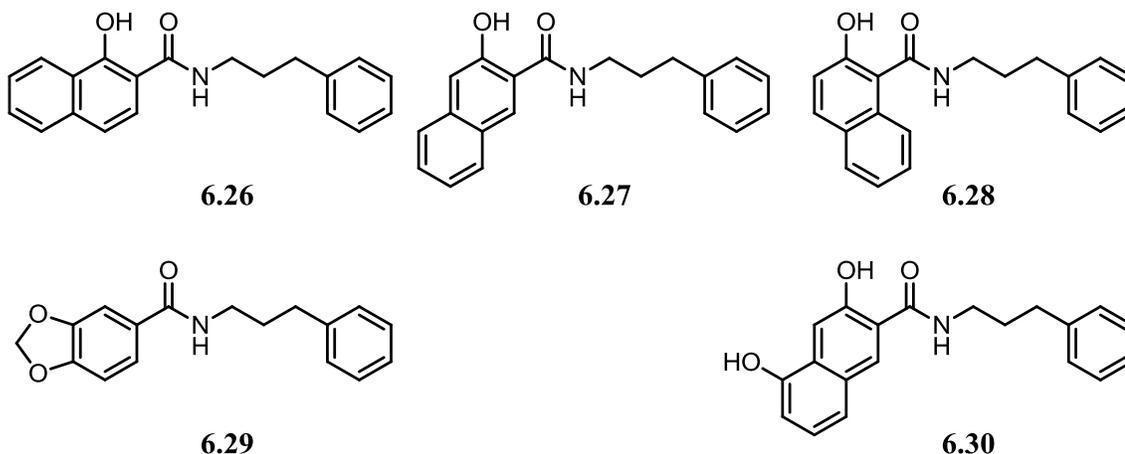


Figure 6.9 – Analogues with fused ring systems used to test steric restrictions at the binding site

Each of these modifications displayed increases in activity over the parent molecule **6.12** ($0.95 < K < 1.35$), the most active being **6.27** ($K[\mathbf{6.27}] = 1.35$). When looking at the plotted data, the naphthalene compounds were screened against the parent molecule, **6.12** ($K[\mathbf{6.12}] = 0.13$), and showed an increase in activity, although the plateau values were not significantly different from each other (**Figure 6.10**). Additionally, when the most active compound, **6.27**, was screened against two additional fused ring systems (**6.29** and **6.30**); it was observed that the span of **6.29** was smaller than that of **6.27** and therefore less active (**Figure 6.11**). Additionally, altered electronic effects can be observed in the context of examining possible limitations of steric bulk with **6.30** ($K[\mathbf{6.30}] = 2.84$). This data would suggest a large hydrophobic pocket at the binding site with respect to the left hand side of the molecule

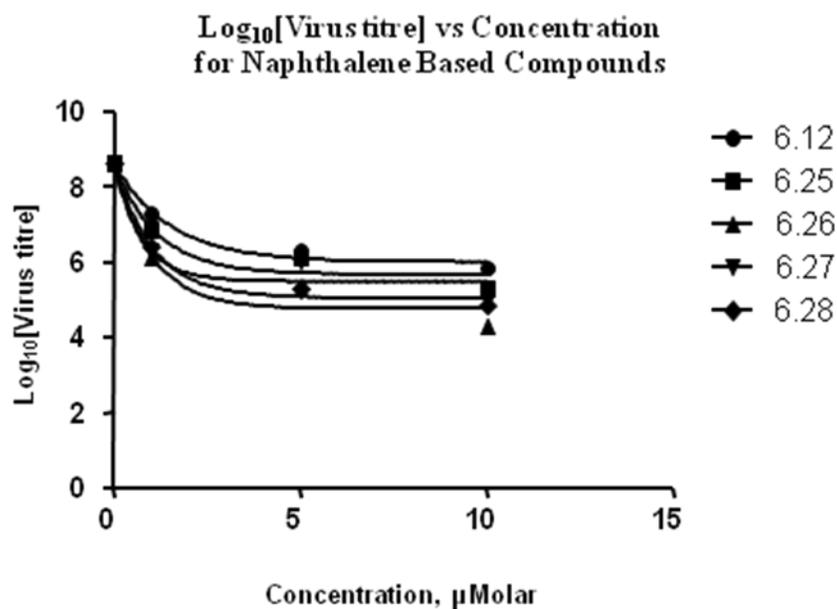


Figure 6.10 – Plotted data for naphthalene compounds vs. parent molecule **6.12**

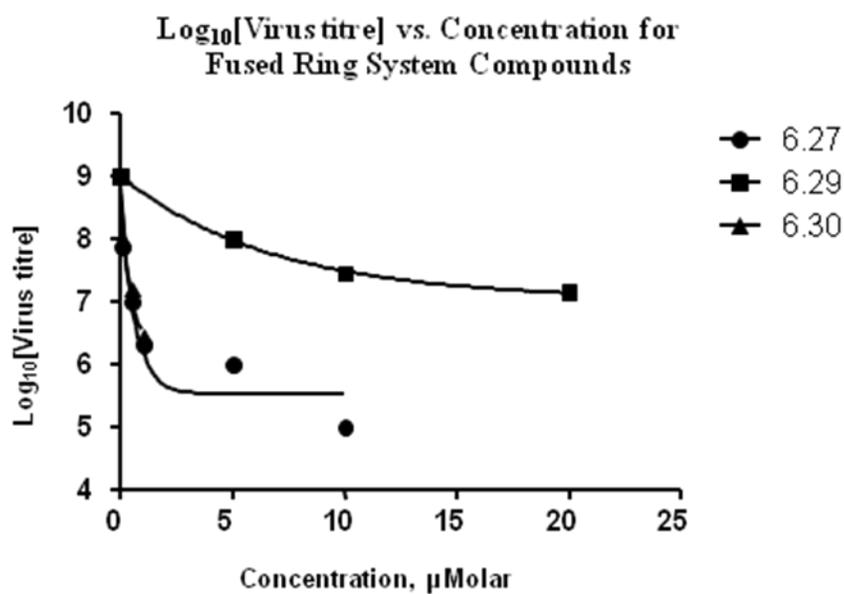


Figure 6.11 – Data which compares three different fused ring systems

Previously, it was demonstrated that replacing the hydroxyl group with a weakly basic amine was an acceptable modification and led to greater virus inhibition. An

alternative way to incorporate a weakly basic amine into the molecule on the left hand side is to synthesize an analogue which places the nitrogen within the ring itself. Using picolinic acid as the building block produced a compound (**6.31**) which has the basic center as part of the ring. This was generally well tolerated and was shown to be more active than the parent molecule **6.12** ($K[\mathbf{6.31}] = 1.1$). Additionally, this modification was also examined in the context of size restrictions (**Figure 6.12**). To do this, three quinolyl derivatives were used in the synthesis of **6.32-6.34** ($1.09 < K < 1.46$). All of the fused ring systems displayed moderate increases in activity with comparison to the pyridyl derivative (**Figure 6.13**). When all four of these compounds were compared to each other, it was observed that there was not a large difference in their activity based on span, again, indicating that this addition of steric bulk is well tolerated at the binding site. Furthermore, **6.31** ($K[\mathbf{6.31}] = 1.1$) was screened against **6.21** ($K[\mathbf{6.21}] = 1.4$) and was found to have similar span, indicating a similar level of activity (**Figure 6.14**).

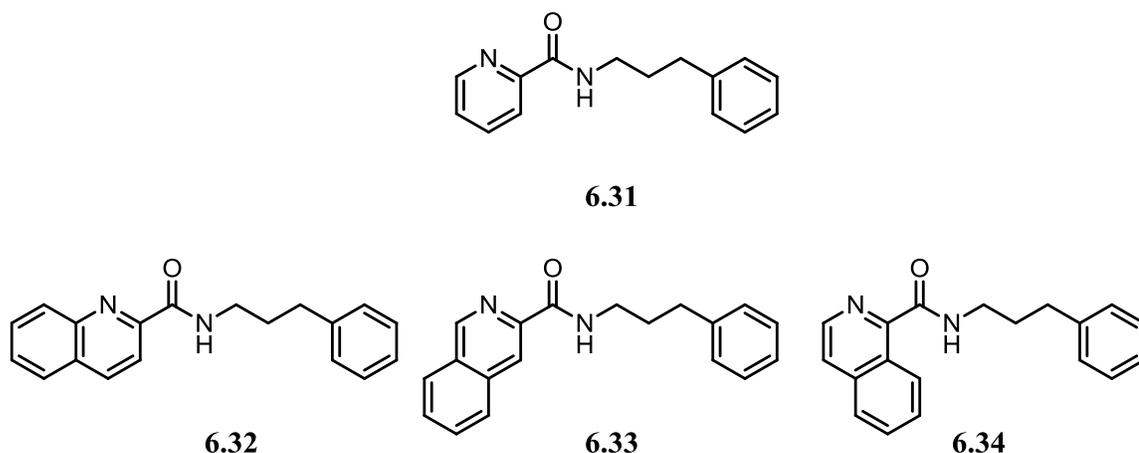


Figure 6.12 – Synthetic analogues which incorporate a heterocycle into the left hand side

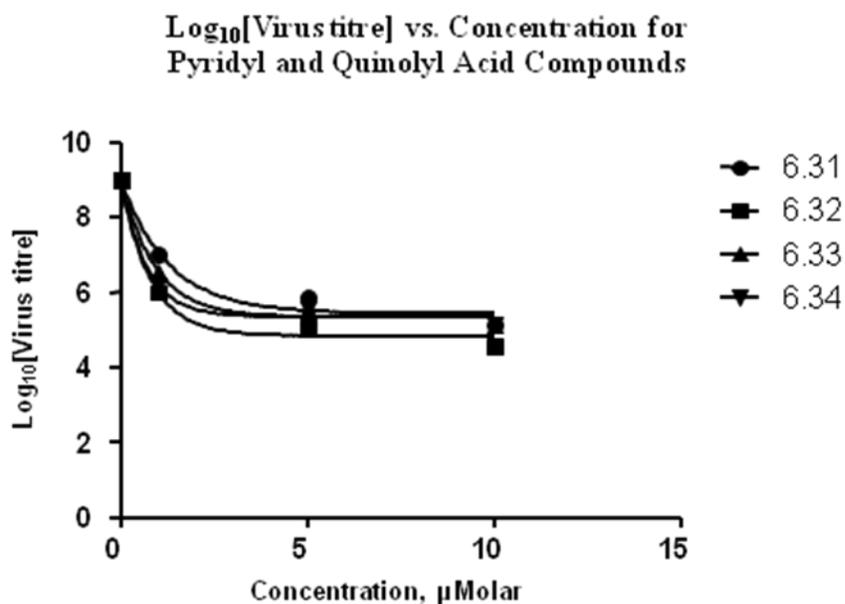


Figure 6.13 - Plotted data comparing the pyridyl, and quinoly derivatives.

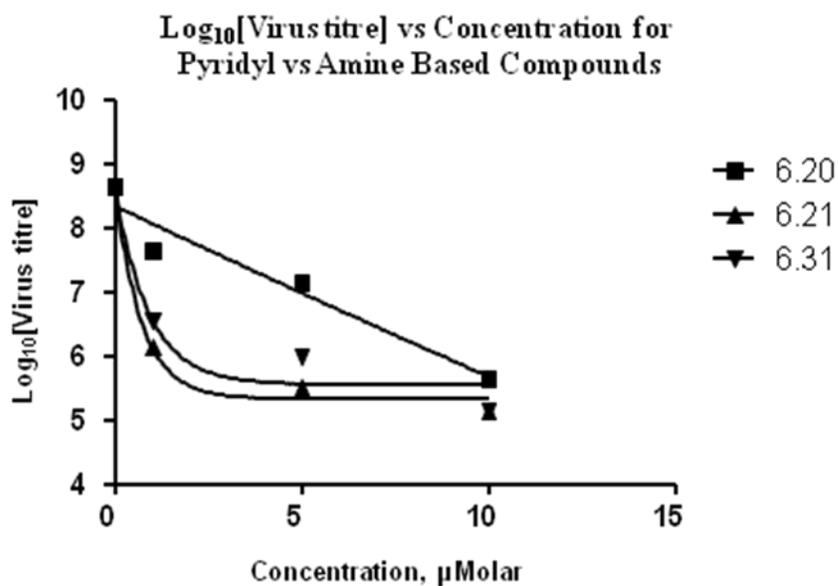


Figure 6.14 – Plotted data for compounds which compares pyridyl, amino, and cyano derivatives.

Both the hydroxyl and the amino functional group have shown to be acceptable substitutions at the ortho position. Additional functionality was added to the aromatic ring to examine any effect altering the electronics of the benzene ring would have on

antiviral activity (**Figure 6.15**). Compound **6.35** placed an amine, an electron donating substituent, para to the carbonyl group. This modification showed a decrease in activity with comparison to **6.21**, which would indicate that a strong electron donating group at the para position (which can also participate in hydrogen bonding) was not an appropriate substitution (**Table 6.2**).

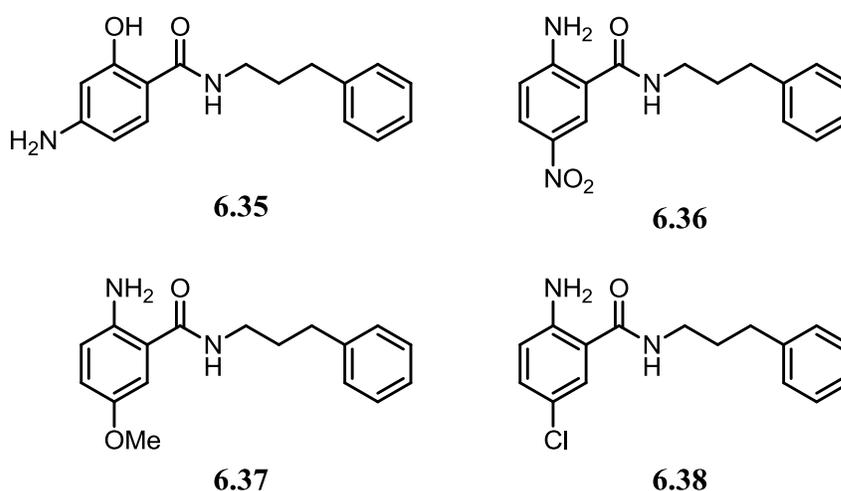


Figure 6.15 – Synthetic analogues which include additional functional groups on the left hand ring

Compounds **6.36-6.38** examined additional functional groups added to **6.21** (**Figure 6.15**). These groups included a strong electron donor, weak electron withdrawing, and a strong electron withdrawing group. Additionally, each of these groups can participate in hydrogen bond accepting interactions (weak to strong). In each instance, the activity at low concentrations decreased with the addition of subsequent functional groups to the ring when compared with **6.21** (**Table 6.2**). This would imply

that the interactions occurring with ortho substituent are being negatively affected regardless of the chemical nature of the substituent.

Table 6.2 - TCID₅₀ values of compounds with additional functionality on the ring.

Concentration (μ M)	6.21	6.35	6.36	6.37	6.38
0	9.0	9.0	9.0	9.0	9.0
1	6.6	7.9	6.9	7.0	7.5
5	6.0	7.0	6.0	6.1	6.5
10	5.2	6.0	5.0	5.0	5.3

Compounds were screened against the lab strain of the virus A/PR/8/34

Exploration of the Linking Region

Previously, it has been established that the preferred substituent for antiviral activity are either a hydroxyl or amino substituent. Given the proximity of these to the carbonyl of the amide, there is a possible hydrogen bonding interaction which may take place and be important for biological activity. A series of analogues were synthesized which explored within the linking region the requirement for a secondary amide, the position of the amide bond, and its orientation (**Figure 6.16**). Each of these subsequent analogues was synthesized using the HOBt/DIC protocol with their respective acids and amines. Compound **6.41** was synthesized according to **Scheme 6.3**. The synthesis was completed in one step by mixing dihydrocoumarin, dimethylaminopyridine (DMAP) and aniline to produce **6.41**.

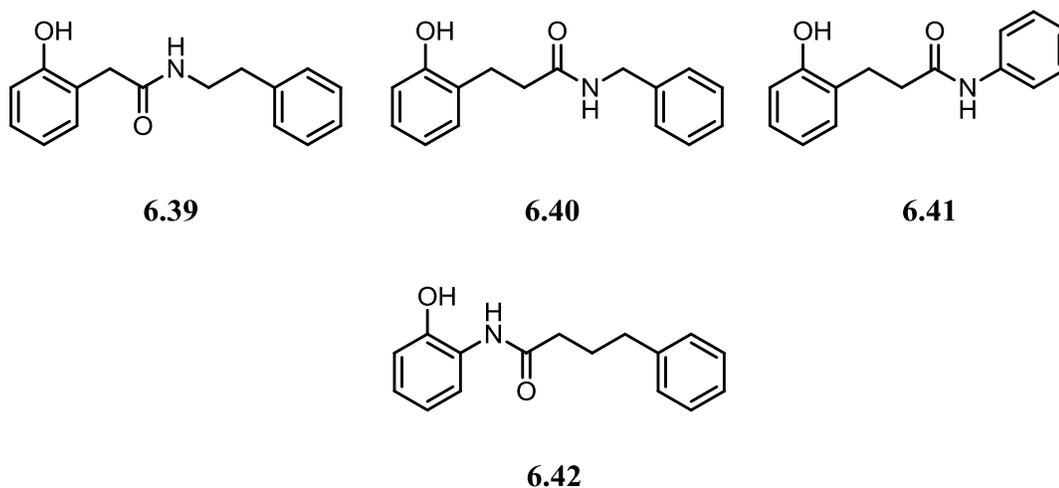
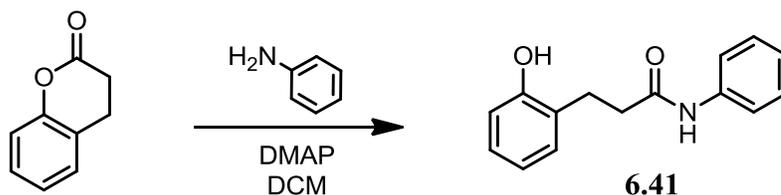


Figure 6.16 – Analogues which move the amide bond away from the hydroxyl group

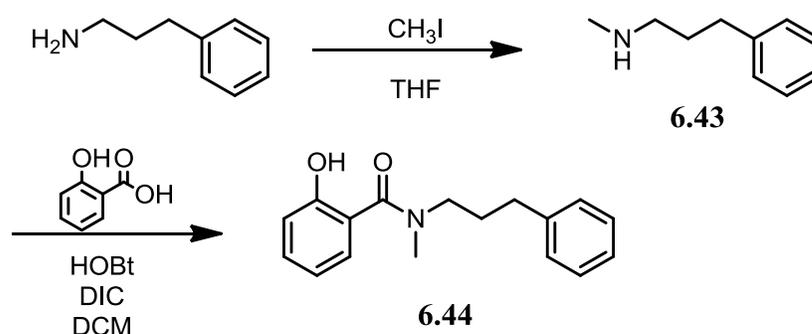
Scheme 6.3 – Synthesis of **41** using dihydrocoumarin



By simply reversing the amide (**6.42**), there was a sharp decrease in inhibitory activity. Additional evidence to support this was seen when the amide bond was moved down the linker in a step-wise fashion. By moving the amide further way from the hydroxyl group, the activity of the compounds decreased. This would indicate that proximity of the hydroxyl group to the amide carbonyl may play an important role in the activity of the compounds.

Besides the hydroxyl group, the amide nitrogen has the capacity to participate in hydrogen bonding as a donor. Up to this point, this possible interaction was ignored. By adding a methyl group to the amine, this interaction is taken away. The synthesis of this compound was achieved in two steps (**Scheme 6.4**). First, 3-phenylpropylamine was mixed with iodomethane at room temperature producing **6.43**. Then, the resulting secondary amine was reacted with salicylic acid to produce **6.44**. The subsequent biological studies showed that methylating the nitrogen had a negative effect on activity.

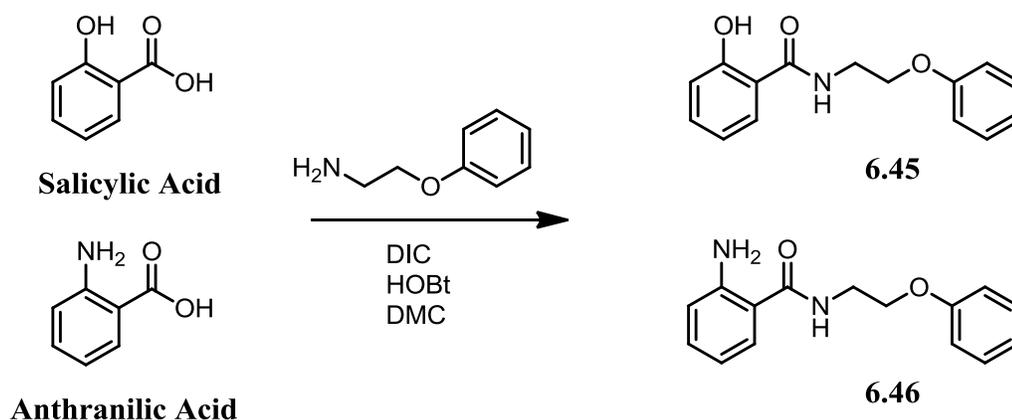
Scheme 6.4 – Synthesis of the methylated compound



Also included in this set of compounds were other versions of **6.12** and **6.21** (**6.45** and **6.46** respectively). If this modification is acceptable this will provide a direct avenue to synthesize analogues which will allow for the examination of important interactions associated with the right hand side of the molecule. These molecules replaced the methylene closest to the right hand phenyl ring with an ether oxygen. These molecules were synthesized with commercially available starting materials according to **Scheme**

6.5. Salicylic acid and anthranilic acid were coupled to 2-phenoxyethanamine using standard HOBt/DIC coupling conditions.

Scheme 6.5 – Synthesis of analogues which incorporated the ether



The inclusion of the ether modification was generally well accepted. The ortho amino analogue resulted in approximately equal inhibitory activity ($K[6.21] = 1.4$ and $K[6.46] = 1.5$). Additionally, the ortho-hydroxy ether analogue displayed an increase in activity when compared to compound **6.12** ($K[6.12] = .13$ and $K[6.45] = 1.15$). This would indicate that the inclusion of the ether had little to no affect on the biological activity. The plotted data would also indicate little difference in activity across the three compounds (**Figure 6.17**). This result provided a way to explore essential interactions with the right hand side of the molecule. Through the use of the Williamson ether synthesis, a variety of phenols can be used to introduce functionality in a synthetically efficient way.

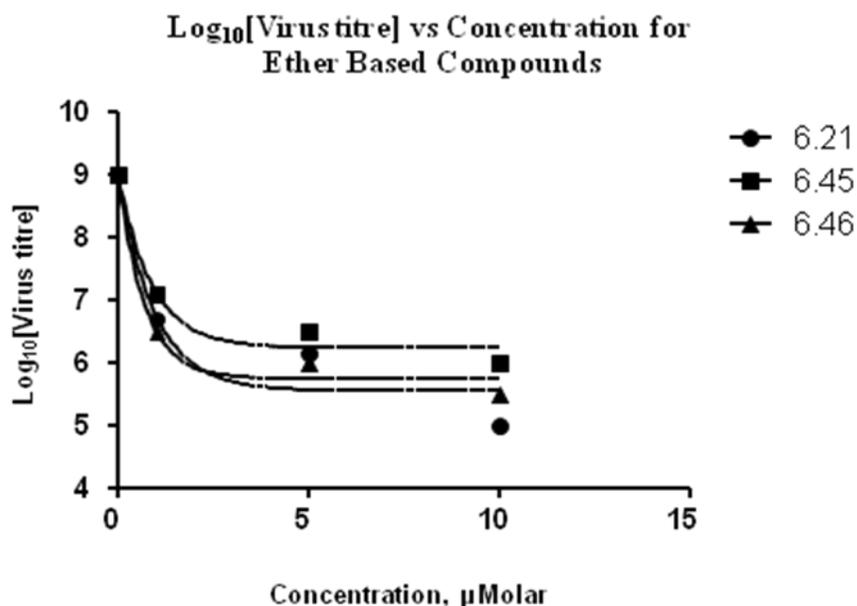
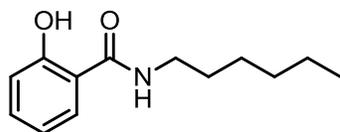
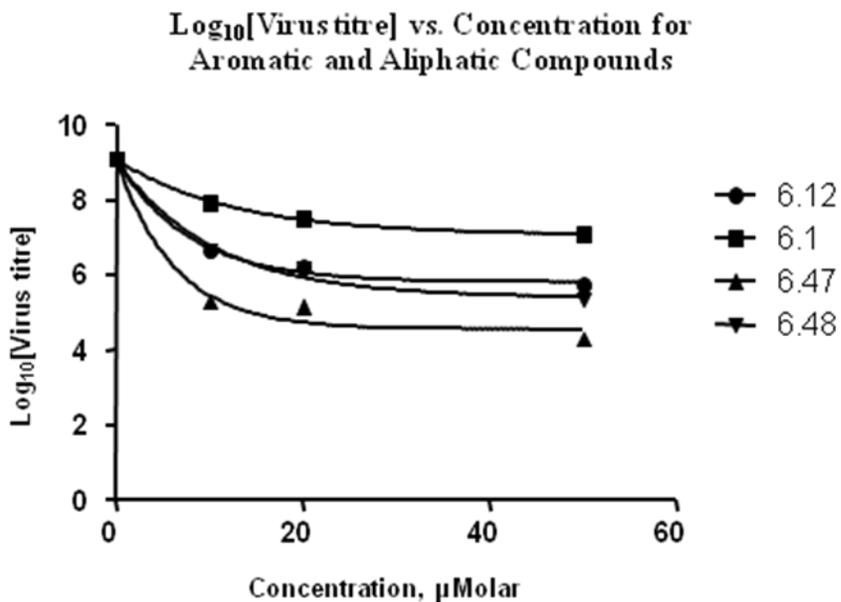


Figure 6.17 – Plotted data which compares ether based compounds to the amino analogue

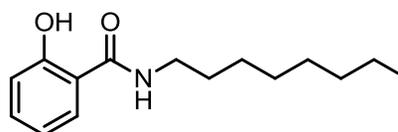
Modifications Associated with the Aliphatic Molecules

Up to this point, all of the molecules have been synthesized with an aromatic group on what is called the right hand side of the molecule primarily since the original molecule included one. Another way to introduce diversity into the molecule and examine if the aromatic ring is a requirement for activity is to use an alternative amine as the building block. Two compounds which were synthesized used n-hexylamine (**6.47**) and octylamine (**6.48**) coupled with salicylic acid using HOBt/DIC (**Figure 6.18**) to examine if removing the aromatic group would have an impact on antiviral activity. When these compounds were screened against the original parent **6.1** and **6.12**, they showed an increase in the ability to prevent viral replication. The most active of these compounds was **6.47**, which used n-hexylamine in the synthesis ($K[\mathbf{6.47}] = 1.68$). Given that the assay in which these compounds are tested is cell based, one possible explanation

for the increase in activity could be an increase in cell penetration. This will increase the concentration of inhibitor around the binding site giving the compounds the appearance of being more active.



6.47



6.48

Figure 6.18 – Synthetic analogues which substitute an aliphatic chain for the aromatic ring with their corresponding biological data compared to **6.1** and **6.12**

With this positive result, it is necessary to incorporate previously tested functionality to determine if activity trends seen in the aromatic versions of the molecule are also observed with this new monomer substitution. Initial experiments attempted to

move the hydroxyl group from the ortho position to the meta, and observe how this affected activity (**Figure 6.19**). As before, moving the hydroxyl group away from this position decreased the overall activity of the molecule, further supporting the idea that the two position is the optimum position for substitution.

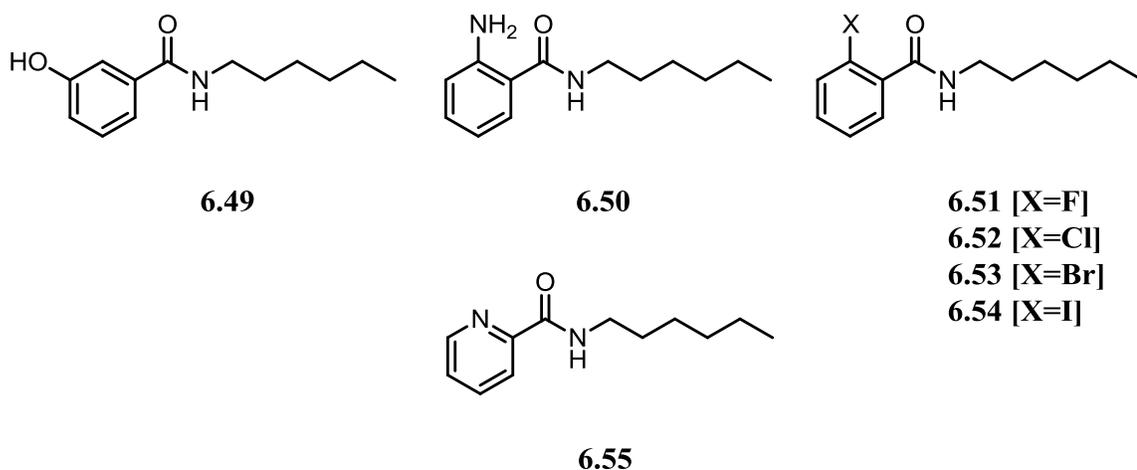


Figure 6.19 – Analogues synthesized which examines the activity of previous building blocks with the aliphatic chain.

The halogen series was also tested along with the amino and pyridyl derivatives (**Figure 6.19**). The results from the halogen screen were similar to the earlier observations with the aromatic group (**6.22-6.25**). Although the iodinated compound had a larger K-value ($K[6.54] = 1.24$), when the data was plotted, the bromine compound had a larger span, indicating it was ultimately, more active (**Figure 6.20**). The amino and pyridyl analogues ($K[6.50] = 9.63$ and $K[6.55] = 2.94$ respectively) were screened against each other and also showed to be very active, which coincided with previous findings that the inclusion of a weakly basic functional group improved upon existing activity (**Figure 6.21**). Although these K-values are drastically different, when

examining the raw data, there is not much difference in the span of these two compounds indicating similar levels activity.

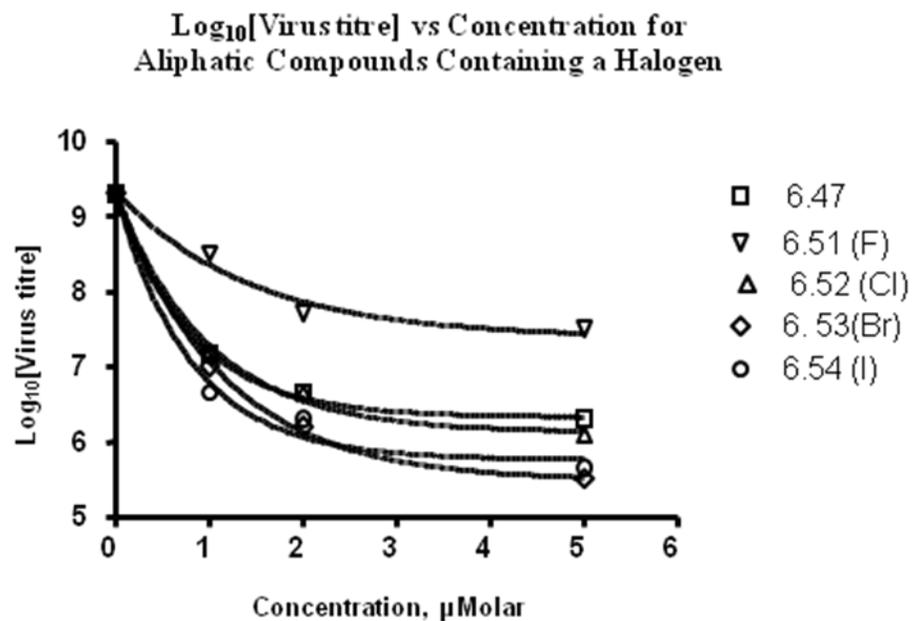


Figure 6.20 – Plotted data of the aliphatic halogen derivatives compared to **6.47**

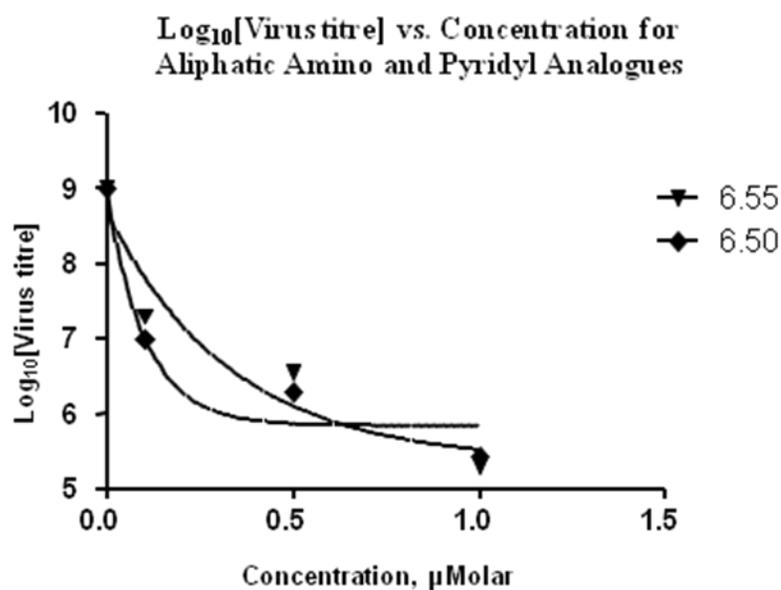


Figure 6.21 – Plotted data comparing aliphatic amino and pyridyl derivatives

Previously, it was determined that the binding site tolerated large hydrophobic bulk on the left hand side of the molecule. This structural motif was incorporated into the molecule using the aliphatic chain (**Figure 6.22**). These compounds were screened against **6.27** and much like before, each of these modifications was well tolerated, confirming earlier conclusions that the binding site is relatively large and hydrophobic (**Table 6.3**). Additionally, the 3,5-dihydroxy-2-naphthoic acid derivative was also synthesized which incorporated a second hydroxyl group on the naphthalene ring. This also appeared to suppress viral growth at acceptable levels.

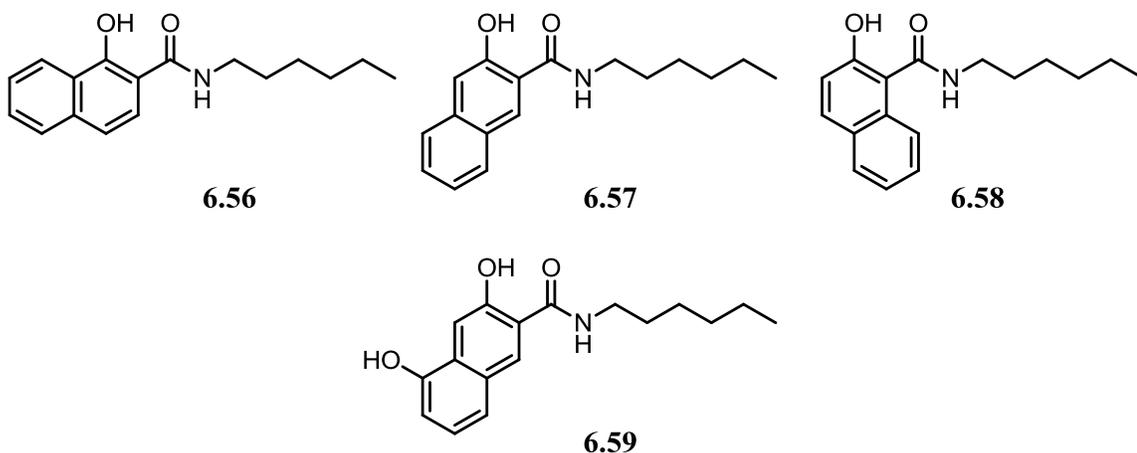


Figure 6.22 – Aliphatic analogues which incorporate steric bulk

Table 6.3 – TCID₅₀ values for compounds **6.56** – **6.58**

Concentration (μM)	6.27	6.56	6.57	6.58
0	9.3	9.3	9.3	9.3
1	6.0	5.7	5.5	6.0
5	5.3	5.0	5.3	5.2
10	5.0	4.7	4.9	5.0

Compounds were screened against the lab strain of the virus A/pr/8/34

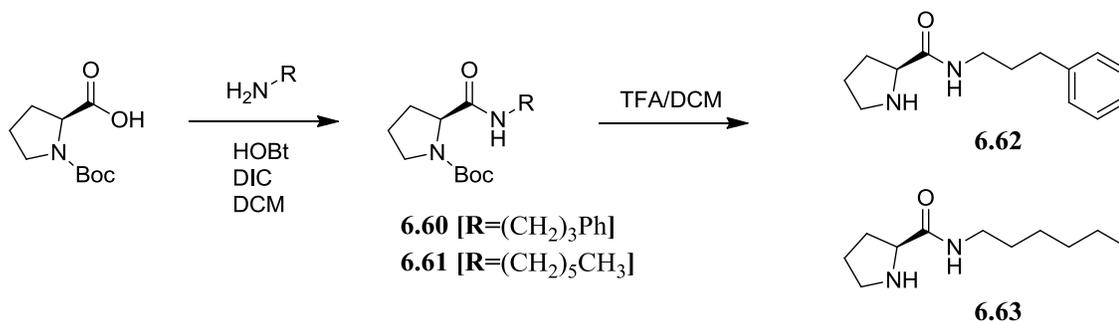
Thus far, each of the previously synthesized compounds were screened against a single strain of the virus, A/PR/8/34 and were shown to be extremely active. It was at this point that these molecules were screened against alternative strains of the influenza virus. Surprisingly, each of these molecules showed little to no activity when screened against these other strains. This was unusual seeing as the structure of NS1 is highly conserved across strains of the virus. With this result, it became necessary to develop a class of molecules in which broad-spectrum activity would be retained.

Using Amino Acids as Building Blocks

Up to this point, a molecule which contained unattractive characteristics was taken and simplified to produce two lead compounds. One of these incorporates an aromatic ring on the right hand side, and one which incorporates an aliphatic chain, with the weakly basic amino group being most active. Furthermore, when picolinic acid was used to introduce another weakly basic center, an increase in activity was observed when compared to **6.12**. A way to further probe the size and electronic restrictions of the left hand side of the molecule and introduce a basic center would be to use amino acids. This will allow for features already determined to be important to be retained while allowing for the introduction of a variety of functional groups with the left hand side of the molecule.

The first series of compounds synthesized used L-proline as the amino acid. For the first two analogues, Boc-l-Pro-OH was coupled to either 3-phenylpropyl amine or n-hexylamine using standard HOBt/DIC coupling conditions. After removal of the Boc group using TFA, the primary amine product was produced (**Scheme 6.6**).

Scheme 6.6 – Synthesis of proline analogues with both the aromatic and aliphatic groups



Based on the K-values, compounds **6.62** and **6.63** showed an increase in activity when compared to the parent molecule (K[**6.62**] = 2.13 and K[**6.63**] = 1.2). This result further supported earlier suspicions that as long as the basic center is preserved, the molecule should retain some biological activity. Using the twenty naturally occurring amino acids opens up another avenue for possible building blocks. Additionally, specificity at the binding site can be explored by using the L and D versions of the amino acids.

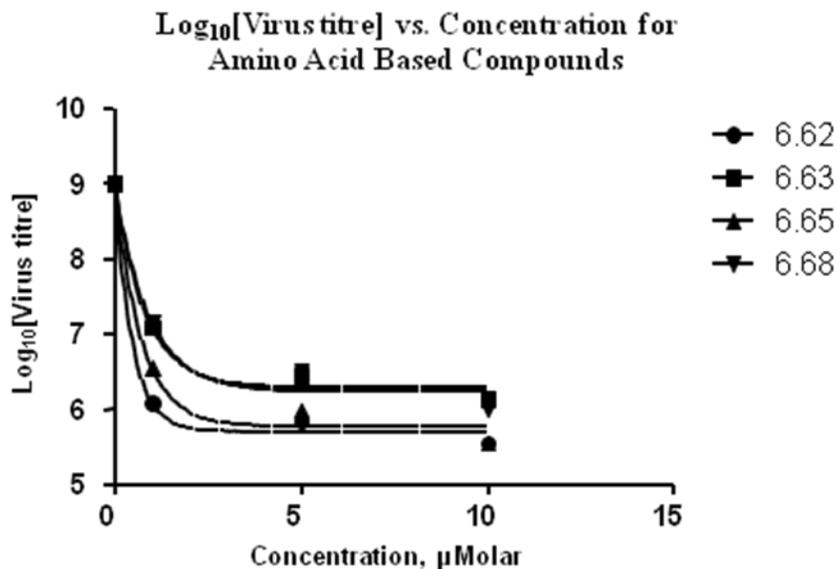
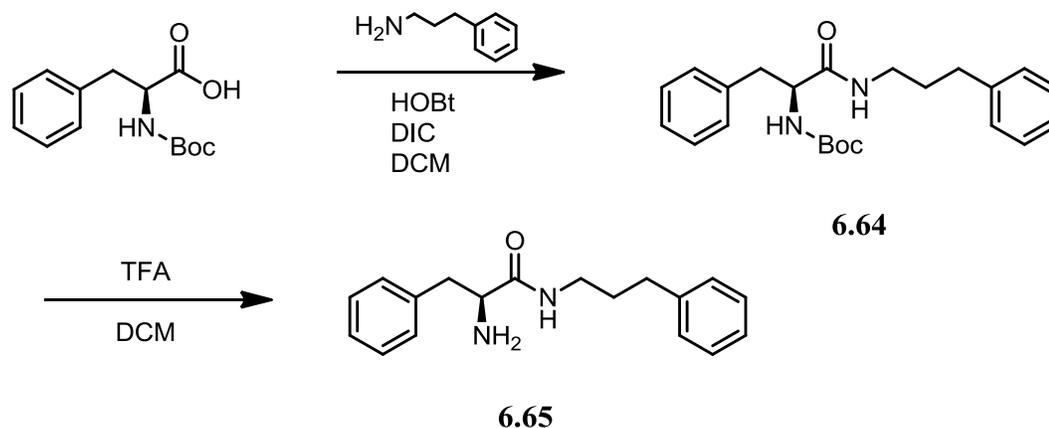
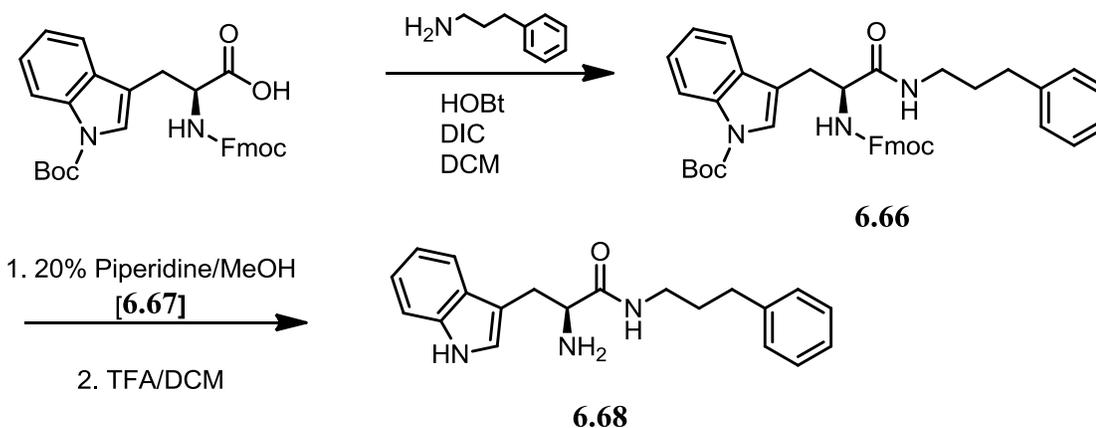


Figure 6.23 – Plotted data for amino acid based compounds.

To further explore the activity associated with amino acids, two more compounds were synthesized (**Scheme 6.7** and **Scheme 6.8**). These utilized phenylalanine and tryptophan as the building blocks. The synthesis proceeded through a straight forward amide bond formation. This was achieved through the use of standard HOBt/DIC coupling conditions between the amino acid and 3-phenylpropylamine. To produce **6.65**, the Boc protecting group was removed using TFA. For the tryptophan derivative (**6.68**), the Fmoc group was removed using 20% piperidine in methanol and the Boc group was subsequently removed using TFA in DCM.

Scheme 6.7 – Synthesis of phenylalanine analogue**Scheme 6.8** – Synthesis of tryptophan analogue

The testing of **6.65** and **6.68** revealed a slight decrease in activity when compared to the proline based derivative **6.62** ($K[6.65] = 1.41$ and $K[6.68] = 1.09$), but an increase in activity when compared to the parent **6.12**. When the $\text{Log}_{10}[\text{Virus titre}]$ was plotted against concentration for all four of the amino acid based compounds (**Figure 6.23**), the span for each compound was not that different indicating similar levels of antiviral

activity. This data confirms previous conclusions involving the necessity of having a weakly basic center as well as being able to tolerate larger hydrophobic substituents at the binding site.

As previously mentioned, many compounds which did not incorporate an amino acid were shown to be very active against the lab strain A/PR/8/34. However, when these compounds were screened against different strains of the virus, little broad-spectrum activity was observed. When the amino acid derivatives were screened against several strains of the influenza A virus, moderate broad spectrum activity was observed.

Library Synthesis

Given that moderate broad spectrum activity was observed using amino acids as building blocks, a library was synthesized to explore the binding site and to improve broad spectrum activity. Within this library, a wide variety of functional groups were examined. The markush structure for the library can be seen in **Figure 6.24**.

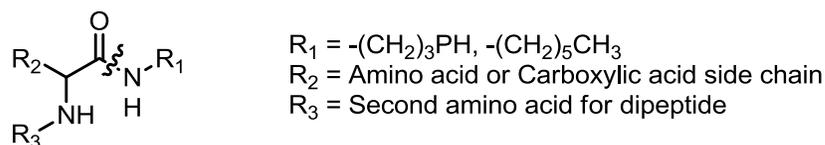
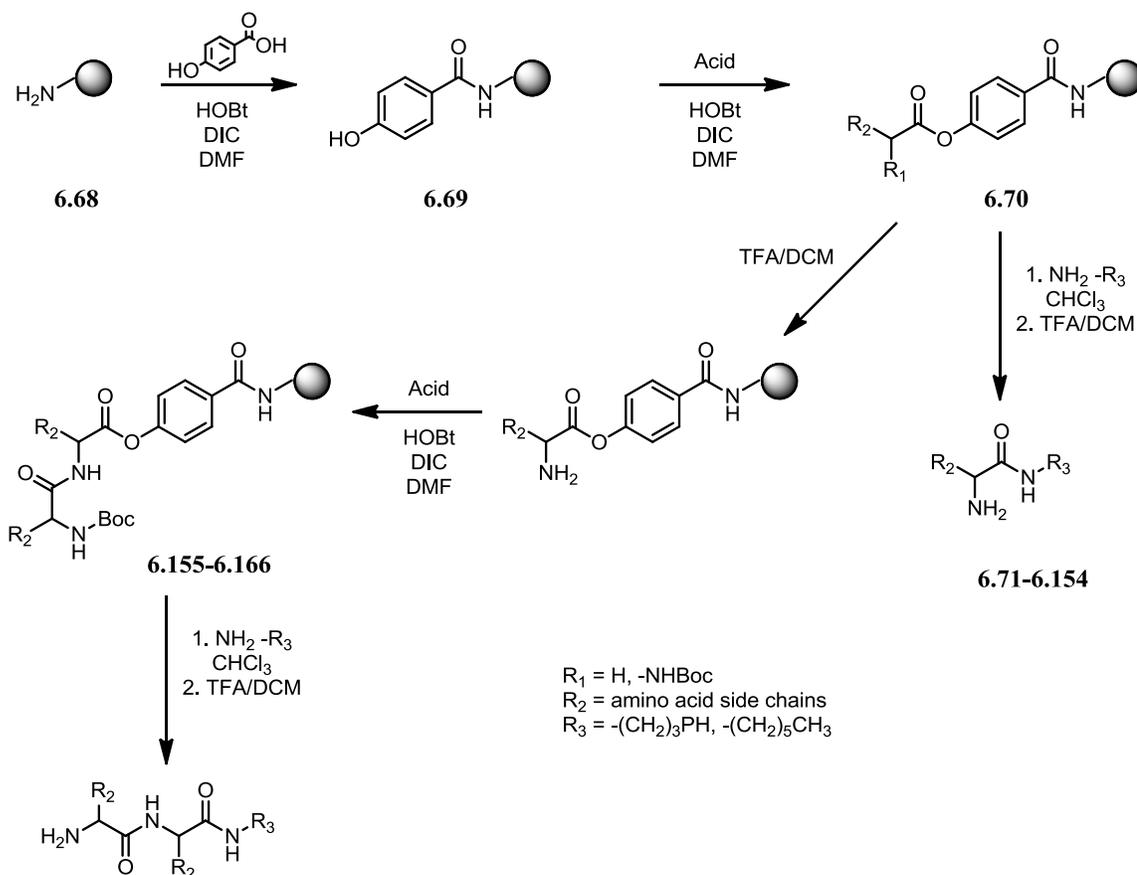


Figure 6.24 – Markush structure used for the design and synthesis of the library

The design of this library is considered a 2x48 compound library. Two different R_1 building blocks were used in the synthesis, 3-phenylpropylamine and n-hexylamine. Also used in this library were 48 different carboxylic acids, amino acids, or dipeptides (**Table 6.4**). The way in which the library was designed, the second monomer, which is a

primary amine (3-phenylpropyl amine or n-hexylamine), is used in a coupling and release step as part of the cleavage strategy (**Scheme 6.9**). In this method, the resin would be mixed with the primary amine and would utilize the active ester (**6.70**) to form the amide bond, releasing the final product from the resin. This strategy is analogous to resin which utilize a cyclize and release cleavage protocol.

Scheme 6.9 – Synthetic strategy used in the synthesis of the library.



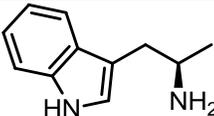
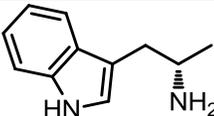
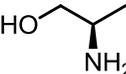
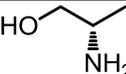
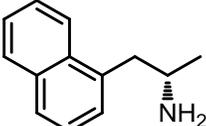
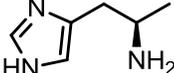
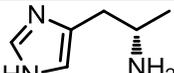
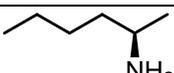
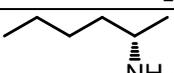
Synthesis of the library began by taking hydroxymethyl resin and coupling β -alanine mediated by DMAP and DIC. Following the removal of the Fmoc protecting

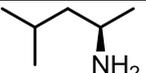
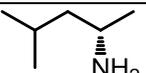
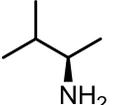
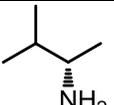
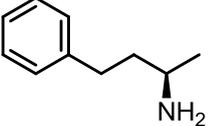
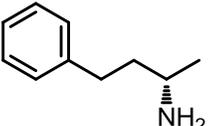
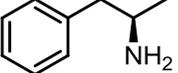
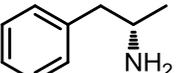
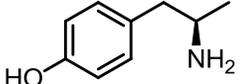
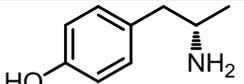
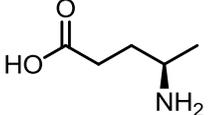
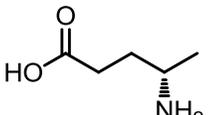
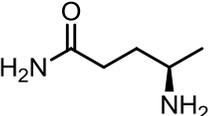
group, 4-hydroxybenzoic acid was coupled to the resin using standard HOBt/DIC coupling conditions. From here, an ester was formed with 48 different acids using DIC/HOBt conditions. Twelve of the wells had l-phenylalanine coupled as the acid. These twelve wells then were mixed with a 1:1 mixture of TFA and DCM to remove the Boc group. Following this, the excess TFA was removed by washing the resin with 10% *N*-methylmorpholine in DMF. A positive Kaiser test indicated a primary amine and a successful Boc deprotection. From here, six different amino acids were used and coupled to the free amine. In order to cleave the product from the resin, two primary amines were used as for the right hand side of the molecule. The molecule is built up as an activated ester and the use of a coupling and release strategy using a primary amine can form the amide bond and release final product from the resin. Following this, the final protecting groups were removed using a 25% TFA solution in CHCl_3/ACN . Some of the products will need to have the side chain deprotection performed under palladium mediated conditions. Following this, the final removal of the Boc group for these compounds can be performed using TFA in DCM.

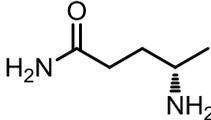
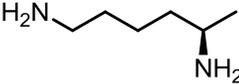
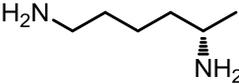
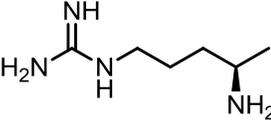
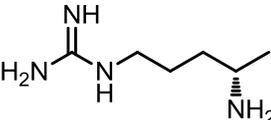
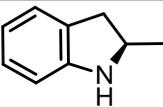
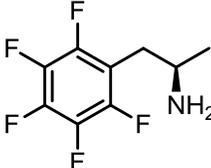
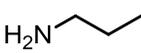
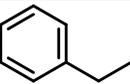
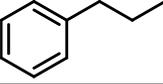
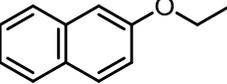
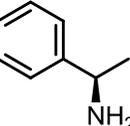
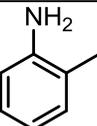
To date, approximately 50% of the library has been submitted for testing using a high throughput assay and final results will be forthcoming. This assay assigned a general '+' or '-' to each compound based on a bench mark activity. From what has been submitted in this first pass assay, 11 compounds have been identified as being more active than the parent compound **6.65**. Among these active compounds was the control compound **6.21**. Interestingly, the results of this initial screen have indicated a preference D-versions of amino acids. This would indicate potential specificity associated with the

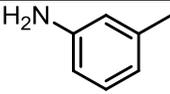
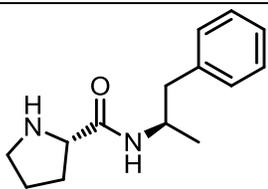
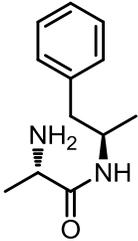
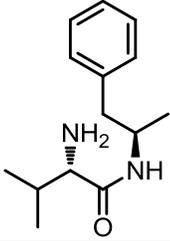
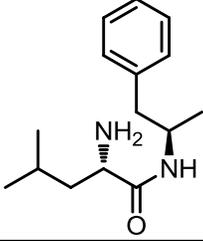
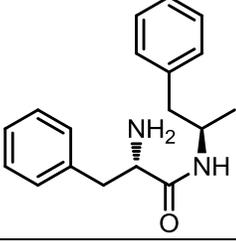
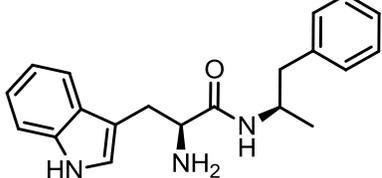
binding site. Each of these 11 identified compounds is being examined further to determine if these compounds retain broad-spectrum activity.

Table 6.4 – Library monomers used for the synthesis of the library. A '+' indicates an increase in activity when compared to **6.65**. A '-' indicates no increase in activity was observed. Blank boxes indicate the assay is ongoing.

R Group	Aromatic	Aliphatic
	-	-
	-	+
	-	-
	-	
		
		
		-
	-	-
		
		
	-	-
	+	

	-	-
	-	-
	-	-
	-	
	-	-
		+
	-	
	+	-
	-	-
	-	+
		
		
		
		
		

		
		
		
		
		
		
		+
		-
		-
		-
		+
		+
		
	+	-

	+	-
		
	-	-
		
		
	-	-
	+	-

Chapter 7:
NS1 Discussion

The original assay used in this project which was able to identify analogues as ‘hits’ was yeast based. The NS1 protein of influenza virus type A was expressed in yeast and resulted in a negative growth phenotype. Through inhibiting the actions of NS1, the yeast can continue to grow. However, the exact method for which the actions of NS1 were inhibited is not fully understood. When examining the activity of analogues in the context of this assay, one possibility is that the analogues interact directly with the NS1 protein and inhibit its function. This would lead to increased cell growth.

The simplest method to determine if there is direct binding to NS1 is to design a direct binding assay. To explore this, biotinylated compounds were synthesized, but were not shown to be active when assayed directly with NS1. This result indicated two different possibilities. First, the addition of biotin to the molecule could have directly led to the decreased binding affinity producing a negative outcome. Alternatively, the compound binds to NS1 while it is complex with a partner; either dsRNA or other proteins. In a direct binding assay this partner would be absent, and if this is the preferred method of interaction, than no activity should be observed.

Within the cell based assay, NS1 suppresses the growth of yeast through interactions with additional yeast proteins and RNA. One possible scenario is that the synthesized analogues are interacting with NS1 while it is in complex with either additional proteins or RNA. Crystal structures of NS1 have shown a conformational change when in complex with RNA.¹ This structural change may produce an allosteric binding site, which was not present before, where the analogues are binding. Additionally, the synthesized analogues may be binding directly in the RNA binding domain while in complex with RNA. Another alternative possibility is that the tested

analogues are interacting exclusively with a protein that NS1 interacts with and is essential for replication of the virus. By disrupting the ability of NS1 to participate in these interactions, viral replication should diminish. As a result, a greater emphasis should be placed on identifying in which model the analogues are inhibiting viral replication.

Developing an SAR

It was the goal of this study to develop a comprehensive structure activity relationship (SAR) and discover a compound which can be considered a viable candidate for a drug discovery program. Upon closer inspection of the four compounds identified in the NSC screen, all of them contain undesirable properties and characteristics which would need to be addressed in order to develop a program around them (**Figure 7.1**).² As an example, the symmetrical structure, presence of two tertiary amines, along with the trisubstituted ring system found in compound **7.2** presents a challenge when determining locations to modify or introduce functionality. Similarly, the synthesis of analogues of compound **7.3** would be cumbersome and lack diverse monomer sets to aid in the development of an SAR. Furthermore, compound **7.3** displays detergent-like characteristics with the two quaternary amines which make it an unattractive candidate for further development.

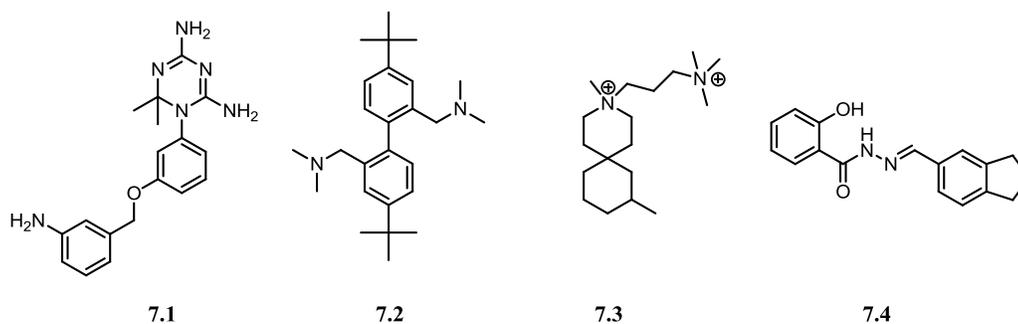


Figure 7.1 – Original active molecules from the NSC screen

Compound **7.1** was eliminated from further investigation primarily due to the triazol ring system. The nitrogen rich ring moiety is unattractive for a drug-like molecule and would likely contribute to a poor pharmacokinetic profile. Compound **7.4** contains a hydrazine core which would likely contribute to pharmacokinetic issues. However, out of all of the compounds, **7.4 (NSC125044)** was most suited for further development. It provided the most opportunity for derivation and has a direct synthetic pathway to desired analogues. Like all of the previous compounds, there are aspects of this molecule that needed to be changed in order to develop this molecule further.

The planned approach was to probe the modifications which simplified the structure of compound **7.4** such that activity was retained and maximal potential for further comprehensive SAR studies was made available. This was initiated by examining the two most unattractive parts of the molecule, the indan ring and the hydrazine core. The indan ring was identified as the least amenable to chemical manipulation, and there are a lack of monomer units available which could be used to further modify the molecule if this structural motif is a required for activity. At high concentrations (50 μ M) a marginal decrease in activity was generally seen across the synthesized analogues. This

discovery would indicate that this was not a required structural motif essential antiviral activity. Furthermore, the ability to use benzaldehyde derivatives drastically increased the number of monomers to select from.

The next portion of the molecule which was to be simplified was the hydrazine core. Each of the nitrogen atoms participated in two different types of bonds, amide and imine. The amide is relatively stable and should not pose any major issues moving the project forward. The hydrazine and imine, on the other hand, are unattractive due to their reactive nature. This part of the molecule would most likely contribute to a poor pharmacokinetic profile and should be removed. The imine also adds structural rigidity to the molecule, which may be desirable in late stage development, however is unwanted this early in development.

The most logical replacement for this is a simple alkyl chain. This removes the reactive nature and replaces it with something relatively inert. Through optimization of the alkyl chain length, (involving chain length and position of the amide bond), the activity was drastically increased to the point where the most active compound (**6.12**) was identified as the new parent molecule and the basis for further lead optimization. The two monomer units to construct this molecule are relatively inexpensive and the chemistry to connect them together is considered scalable. From a commercial standpoint, this modification is extremely desirable. Furthermore, this synthetic pathway allowed for a number of analogues to be synthesized in a relatively short amount of time.

Additional modifications to the parent molecule included replacing the aromatic group on the right hand side of the molecule with a linear aliphatic chain. This improved the observed antiviral activity of the molecule and presented two additional variables

which need to be considered. The first is actual activity at the binding site. The second is cell penetration. Molecules which contain the long aliphatic chain may be passing through the membrane more efficiently. This would allow for a higher concentration of compound to be 'available' to interact with the target and appear to be more active. This led to the conclusion that moving forward, the modifications made at the left hand side of the molecule needed to be examined in context of both the aromatic and aliphatic versions.

Of all the modifications made to the left hand side of the molecule, the most active was substituting an amine for the hydroxyl group. Initially it was suggested that since the hydroxyl group and the aniline amine could participate in both hydrogen bond accepting and donating interactions, this may be important for activity. However the positive result obtained from the testing of the halogen series would seem to rule this out as a requirement for activity. Furthermore, the inclusion of a weakly basic center using picolinic acid as a building block resulted in an increase in antiviral activity. This would indicate that the inclusion of a basic center may be important for biological activity. Additionally, it was concluded in the context of both parents, that the binding site tolerated a large amount of hydrophobic bulk through the screening of naphthalene derivatives.

One of the proposed hypotheses was that the tested compounds were acting at the RNA binding domain. By interrupting the ability of NS1 to bind dsRNA, the host immune system can actively fight off the spread of the virus. However, *in vitro* testing demonstrated there was no evidence to support the claim that compound **6.12** effected the binding interactions between NS1 and dsRNA. Despite this, it has been postulated that

6.12 could act allosterically to disrupt NS1 actions. This could produce downstream effects on the production and activity of IFN, which would ultimately effect the actions of NS1.³

Further studies were performed on **6.12** in an effort to examine how the compound acts in preventing the activity of NS1. From these studies, a number of observations were made. First, when cells are infected at a low m.o.i. (multiplicity of infection), without treatment of a drug, the NS1 protein inhibits the production of interferons (IFN). This allows the virus to spread easily. Second, when cells are infected at low m.o.i. and treated with **6.12**, they are able to successfully produce IFN, even with the actions of NS1. This treatment leads to an ‘antiviral’ state in neighboring cells which leads to resistance of infection and limits viral spread. Lastly, treatment of **6.12** in cells with a high multiplicity of infection demonstrated the ability to restore the IFN synthesis in infected cells, but was not shown to be able to slow the spread of the virus.³ This would indicate that this compound was capable of preventing the spread of the virus early on during infection, but would exhibit reduced effectiveness for later stage infection. Although the mechanism of action of **6.12** remains unknown, it was shown that the activity of the compound was dependent upon the presence of RNase L.³ One possibility is that the compound is interacting with the complex of NS1 and RNA. This could explain the observed biological interactions.³ Alternatively, the synthesized inhibitors may be disrupting the interactions between NS1 and other intracellular proteins required for efficient viral replication.

Despite these results, all of the synthesized analogues were only observed to be active at a single lab strain. When tested against a variety of seasonal influenza A strains,

the most active compounds had little effect on viral replication for each strain. This was an unusual result since NS1 is highly conserved from strain to strain.⁴ It is the overall goal to develop a drug which can be effective as an antiviral agent regardless of the strain of the virus. With this development, it becomes crucial that the binding site of the molecules is determined. To do this, compounds should be synthesized which allow for the ability to track them. One suggested method is to incorporate fluorescein into the molecule (**Figure 7.2**). In doing this, a number of studies will be able to be performed to answer questions regarding cell permeability and subcellular localization within the cell. Furthermore, compounds which include the fluorescein tag can be used in the development of a direct binding assay.

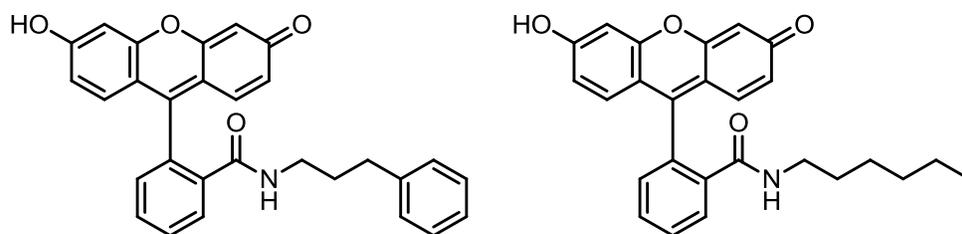


Figure 7.2 – Structure of the compounds containing fluorescein.

An alternative method to incorporate a basic center into the molecule is to use amino acids as building blocks. In doing this, a variety of diverse functional groups can be incorporated on the left hand side of the molecule. Furthermore, with the availability of L and D amino acids, specificity can also be examined at the binding site. The results of the initial amino acid compounds displayed an increase in activity when compared to

compound **6.12**. Upon further investigation of the amino acid analogues, it was shown that these exhibited mild broad spectrum activity. It was this result which led to the development of the library. In order to be thorough, 48 different carboxylic acids were used in the synthesis of the 96-compound library. Each of the carboxylic acids/amino acids was selected to probe the key structural features which may contribute to broad spectrum activity. Furthermore, with the introduction of stereochemistry, more selectivity may be observed based upon structural orientation.

Of the compounds screened in the library, 11 of them showed increased activity when compared to compound **6.65** in a first pass assay. When examining the structures of these molecules, it was noticed that 5 of the 11 compounds had the D-configuration, 1 had L-configuration, 4 had no stereochemistry, and 1 dipeptide with L-configuration for each amino acid. This may be an indication that specificity may play a role in activity. Moving forward with this result, these compounds should be tested to examine if these modifications contribute to broad-spectrum activity. Knowing this, better decisions can be made to further develop a comprehensive SAR and develop potential therapeutics.

Conclusions

At the outset of this project, it was the overriding goal to produce a structure activity relationship which could be applied to design an attractive drug candidate. Initial experiments were designed to simplify the indan ring system to make it more amenable to chemical modification. Additionally, successful modification of the linking region drastically simplified the molecule and as well as the synthetic pathway. Further

experiments involved the incorporation of an ether linkage to provide a method to explore important interactions associated with the right hand side of the molecule.

Initial exploration of the left hand side of the molecule indicated potential hydrogen bonding interactions might be important to the antiviral activity. However, the increased activity observed with the halogens seemed to rule hydrogen bonding. Two analogues which incorporated a weakly basic center resulted in larger increase in activity indicating a preference for a weakly basic substituent ortho to the amide carbonyl. Further experiments revealed that the binding site tolerated hydrophobic bulk on the left hand side of the molecule. This increase in activity was also observed in the context of including a weakly basic center into naphthalene derivatives. Additionally, when an aliphatic chain is used to replace the phenyl group at the right hand side of the molecule, an increase in activity was observed; however, this could be due to an increased ability to pass through the cell membrane.

Although each of these analogues was able to display high antiviral activity with relatively low cytotoxicity, these compounds were only shown to be active at the lab strain of the virus, A/PR/8/34. An alternative way to incorporate a basic center into the molecule was to use amino acids as building blocks. This allowed for a diverse set of building blocks to be used which also incorporates stereochemistry to test for specificity. The testing of these analogues revealed moderate broad-spectrum activity. With this result, 96-compound library was designed and synthesized to further explore essential structural features associated with broad spectrum activity. Also, with the introduction of stereochemistry, requirements relating to structural orientation can be explored.

To date, approximately 50% of the library has been screened for their corresponding activity. Analysis of these results revealed that 11 of the tested compounds revealed increased activity when compared to compounds **6.65**. Subsequent screening of these analogues should be performed to examine if broad spectrum activity was retained.

Continuing work should involve completing the screening of the library and subsequent screening of active compounds for broad spectrum activity. Subsets of compounds have been synthesized to determine structural key structural features which can be used in the design of additional analogues. The compounds which contain fluorescein should be used in studies to help develop a direct binding assay and examine cell permeability and subcellular localization within the cell.

Chapter 7 References

1. Cheng, A.; Wong, S. M.; Yuan, Y. A., Structural basis for dsRNA recognition by NS1 protein of influenza A virus. *Cell Res* **2009**, *19* (2), 187-195.
2. Basu, D.; Walkiewicz, M. P.; Frieman, M.; Baric, R. S.; Auble, D. T.; Engel, D. A., Novel Influenza Virus NS1 Antagonists Block Replication and Restore Innate Immune Function. *J. Virol.* **2009**, *83* (4), 1881-1891.
3. Walkiewicz, M. P.; Basu, D.; Jablonski, J. J.; Geysen, H. M.; Engel, D. A., Novel inhibitor of influenza non-structural protein 1 blocks multi-cycle replication in an RNase L-dependent manner. *J Gen Virol* **2011**, *92* (1), 60-70.
4. Shaw, M. W.; Lamon, E. W.; Compans, R. W., Immunologic studies on the influenza A virus nonstructural protein NS1. *The Journal of Experimental Medicine* **1982**, *156* (1), 243-254.

Chapter 8:
Experimental Protocols for Selected Compounds

Reagents and solvents used for synthesis were reagent grade and used as received. Rink resin, hydroxymethyl resin, HOBT, DIC and amino acids were purchased from Creosalus/AdvancedChemTech. All other reagents and solvents were purchased from Acros Organics, Fisher Scientific, or Sigma-Aldrich.

^1H NMR spectra were obtained using a 300MHz Varian instrument. Chemical shifts were reported in parts per million (δ) relative to the solvent as follows: CDCl_3 (δ 7.24) and $\text{DMSO-}d_6$ (δ 2.50). NMR spectra data are abbreviated as follows: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet.

All analytical LCMS were performed with using the Gilson Unipoint software and analytical equipment. This includes Gilson 215 Liquid Handler, Gilson 306 pumps, Gilson 811c Dynamic Mixer, Gilson 806 Manometric Module, and a Gilson 170 Diode Array Detector with deuterium lamp. The column used was a Thomson Instrument Inc. Advantage C18, 60A, 5u, 250x4.6mm column. In conjunction with the LC, the Mass Spectra were collected using a single quad PE SCIEX API 150EX mass spectrometer utilizing Analyst software. The solvent gradient started at 90% water/.01% TFA, 10% acetonitrile/.01% TFA and changed to 100% acetonitrile/.01% TFA over 10 minutes. The 100% acetonitrile/.01% TFA phase was held for 10 minutes.

Preparative HPLC was performed using a Gilson 215 Liquid Handler, Gilson 306 pumps, Gilson 811c Dynamic Mixer, Gilson 805 Manometric Module, and a Gilson UV/VIS-156 detector. A Thomson Instrument Inc. Advantage C18, 60A, 5u, 250x10.0mm column and a Thomson Instrument Inc. Advantage C18, 60A, 5u, 250x21.2mm column were used in the purification of samples. The solvent gradient started at 90% water/.01% TFA

10% acetonitrile/.01% TFA and changed to 100% acetonitrile/.01% TFA over 10 minutes. The 100% acetonitrile/.01% TFA phase was held for 10 minutes.

Biological assays were performed with a Packard Fusion plate reader using an excitation wavelength of 340nm and an emission filter of 485nm.

Experimental Protocols for Experiments in Chapter 3

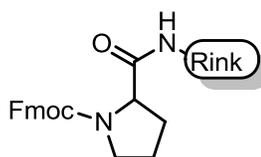
General Procedure A – Ninhydrin Colorimetric Test for Primary Amines - Take a small sample of beads (1-2mg) and place in a small test tube. To this, add 1-2 drops of 10% phenol solution in ethanol, 1-2 drops pyridine, and 1-2 drops of 10% Ninhydrin in Ethanol solution. Heat the mixture at 105°C for 45-60 seconds. If the beads turn blue/purple, this indicates a positive result for a primary amine. Clear beads indicate a negative result for a primary amine.

General Procedure B – Chloranil Colorimetric Test for Secondary Amines - Take a small sample of beads (1-2mg) and place them in a small test tube. To this, add 2 drops of a 20% Acetaldehyde in DMF solution and 2 drops of 20% of Chloranil in DMF solution. Mix the solution and allow it to sit for 2-5 minutes. Green beads indicate the presence of a secondary amine where as clear beads indicate a negative result for the presence of a secondary amine.

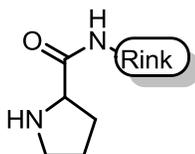
General Procedure C - Deprotecting Fmoc-Rink Resin – Rink resin (.7mmol/g) is added to a peptide shaker with a solution of 20% piperidine in DMF. The contents are mixed on a shaker for 20 minutes. Following this, the liquid is drained from the vessel and the resin is washed with DMF (2x), Methanol (1x), and DMF (1x).



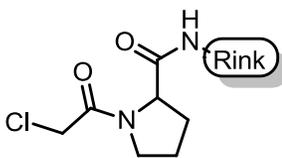
Deprotecting Fmoc-Rink Resin (3.4) – In a peptide shaker, 2g of Rink linker resin (loading = 0.7mmol/g) was deprotected using General procedure C. A positive Kaiser test indicated a successful deprotection.



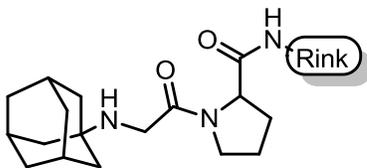
Coupling Fmoc-L-Pro-OH to Deprotected Rink Resin (3.5) – To the vessel of deprotected resin in DMF, add 3 equivalents of Fmoc-L-Pro-OH (1.415g), HOBt (642mg), and DIC (662.5 μL). Allow the contents of the vessel to mix for 2 hours. Upon completion of the reaction, drain/filter the resin, and wash with DMF (3x). A negative Kaiser test indicated a successful coupling.



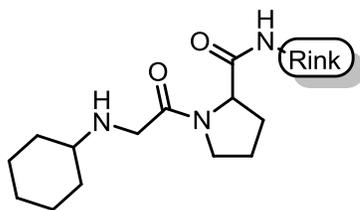
Deprotection of Fmoc-Pro-Resin (3.6) – To the peptide shaker containing the resin, a solution of 20% piperidine in DMF was added and mixed for 20 minutes. Following the deprotection, the resin was drained and washed with DMF (2x), Methanol (1x), and DMF (1x). A positive chloranil test indicates a successful deprotection and secondary amine.



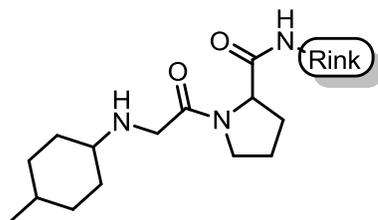
Coupling of Chloroacetyl Chloride to resin (3.7) – Approximately 1g of resin from **3.6** was taken up in DCM with 6 equivalents of N-methylmorpholine (NMM) to make sure the solution is basic. To this, 5 equivalents of chloroacetyl chloride were added. The reaction was allowed to mix for 3 hours. Following the reaction, drain and wash the resin with DCM (2x), MeOH (1x), DMF (1x). A negative chloranil test indicated a successful coupling



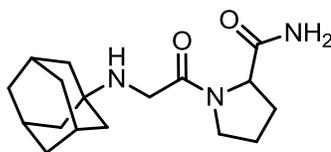
1-(2-((3s,5s,7s)-adamantan-1-ylamino)acetyl)pyrrolidine-2-carboxamide (3.8) – 170 mg of **3.7** resin was placed in a scintillation vial. To this, 4 equivalents of adamantylamine (71.8mgs) and 4 equivalents of N-methylmorpholine (52.33 μ L) were added. The reaction was allowed to mix overnight at room temperature. The next day, the liquid was drained from the vessel and the resin was washed with DMF (2x), Methanol (1x), and DCM (1x). A positive chloranil test indicated the presence of the secondary amine. $t_R = 6.23$ minutes. $M+H = 306.3$



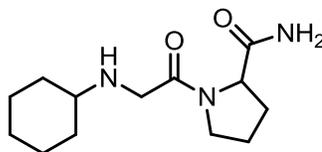
1-(2-(cyclohexylamino)acetyl)pyrrolidine-2-carboxamide (3.9) - 170 mg of **3.7** resin was placed in a scintillation vial. To this, 4 equivalents of cyclohexylamine (54.5 μL) and 4 equivalents of *N*-methylmorpholine (52.33 μL) were added. The reaction was allowed to mix overnight at room temperature. The next day, the liquid was drained from the vessel and the resin was washed with DMF (2x), Methanol (1x), and DCM (1x). A positive chloranil test indicated the presence of the secondary amine. $t_{\text{R}} = 5.89$ minutes. $\text{M}+\text{H} = 268.3$



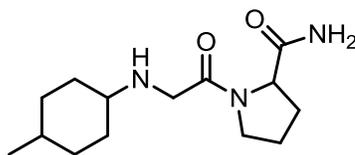
1-(2-((4-methylcyclohexyl)amino)acetyl)pyrrolidine-2-carboxamide (3.10) - 170 mg of **3.7** resin was placed in a scintillation vial. To this, 4 equivalents of methylcyclohexylamine (63.0 μL) and 4 equivalents of *N*-methylmorpholine (52.33 μL) were added. The reaction was allowed to mix overnight at room temperature. The next day, the liquid was drained from the vessel and the resin was washed with DMF (2x), Methanol (1x), and DCM (1x). A positive chloranil test indicated the presence of the secondary amine. $t_{\text{R}} = 6.23$ minutes. $\text{M}+\text{H} = 306.3$



Cleavage of 1-(2-((3s,5s,7s)-adamantan-1-ylamino)acetyl)pyrrolidine-2-carboxamide from resin (3.11) – The 170 mgs of resin of **3.8** was mixed in a 50/50 solution of Trifluoroacetic acid in DCM for 1 hr at room temperature. Following this, the solution was diluted with water and placed on the freeze dryer to remove the TFA and solvent. The resulting powder was carried on without further purification. Crude yield = 33.9mgs (87.8%). $t_R = 6.23$ minutes. $M+H = 306.3$

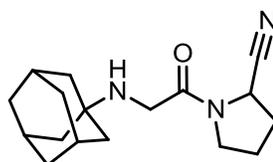


1-(2-(cyclohexylamino)acetyl)pyrrolidine-2-carboxamide (3.12) - The 170 mgs of resin (**3.9**) was mixed in a 50/50 solution of Trifluoroacetic acid in DCM for 1 hr at room temperature. Following this, the solution was diluted with water and placed on the freeze dryer to remove the TFA and solvent. The resulting powder was carried on without further purification Crude yield = 31.8mg (98%). $t_R = 3.89$ minutes. $M+H = 254.3$

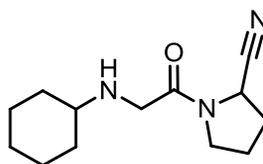


1-(2-((4-methylcyclohexyl)amino)acetyl)pyrrolidine-2-carboxamide (3.13) - The 170 mgs of resin (**3.10**) was mixed in a 50/50 solution of Trifluoroacetic acid in DCM for 1 hr

at room temperature. Following this, the solution was diluted with water and placed on the freeze dryer to remove the TFA and solvent. The resulting powder was carried on without further purification Crude yield = 27.8mg (82%). $t_R = 5.89$ minutes. M+H = 268.3

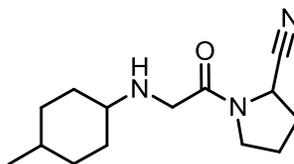


1-(2-((3s,5s,7s)-adamantan-1-ylamino)acetyl)pyrrolidine-2-carbonitrile (3.1) – The product of **3.11** was dissolved in DCM and cooled to 0°C. To this, 1.1 equivalents (.121mmol) of Trifluoroacetic anhydride were added dropwise. The reaction was allowed to warm to room temperature while mixing. The reaction was allowed to mix for one hour. Upon completion, the reaction was evaporated away and the crude oil was carried on to preparative HPLC without further work up. Recovery after purification = 11.5mg (33.9%). $t_R = 11.00$ minutes. M+H = 288.2

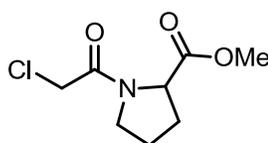


1-(2-(cyclohexylamino)acetyl)pyrrolidine-2-carbonitrile (3.2) - The product of **3.12** was dissolved in DCM and cooled to 0°C. To this, 1.1 equivalents (.121mmol) of Trifluoroacetic anhydride were added dropwise. The reaction was allowed to warm to room temperature while mixing. The reaction was allowed to mix for one hour. Upon completion, the reaction was evaporated away and the crude oil was carried on to

preparative HPLC without further work up. Recovery after purification = 5.0mg (15.7%). $t_R = 11.34$ minutes $M+H = 236.2$

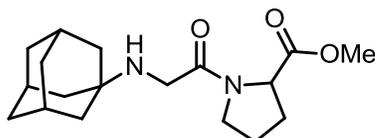


1-(2-((4-methylcyclohexyl)amino)acetyl)pyrrolidine-2-carbonitrile (3.3) - The product of **3.13** was dissolved in DCM and cooled to 0°C. To this, 1.1 equivalents (.121mmol) of Trifluoroacetic anhydride were added dropwise. The reaction was allowed to warm to room temperature while mixing. The reaction was allowed to mix for one hour. Upon completion, the reaction was evaporated away and the crude oil was carried on to preparative HPLC without further work up. Recovery after purification = 4.3mg (13.5%). $t_R = 11.71$ minutes. $M+H = 250.2$



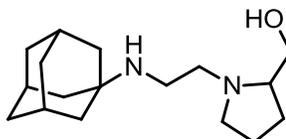
(S)-methyl 1-(2-chloroacetyl)pyrrolidine-2-carboxylate (3.15) - Commercially available Proline-methylester (9.2 mmol) was dissolved in DCM in a round bottom flask. To this, 1.1 eq (10.1 mmol) of chloroacetyl chloride was added dropwise. Additionally, 27.6 mmol of N-methylmorpholine was added to keep the solution basic (pH ~ 8). The reaction was mixed overnight at room temperature. The following day, the reaction was diluted with DCM and washed with equalvolume of 2M HCl (2x). The DCM was dried

over Na_2SO_4 and removed under reduced pressure. Crude yield = 8.4mmol (91.3%). $t_R = 6.90$ minutes. $M+H = 205.9$



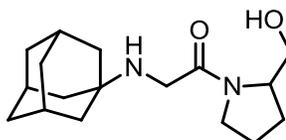
methyl 1-(2-((3s,5s,7s)-adamantan-1-ylamino)acetyl)pyrrolidine-2-carboxylate (3.16)

– The entire product of **3.15** was dissolved in THF in a round bottom flask. To this mixture, 1.1 equivalents (9.24 mmol) of adamantylamine were added. The pH of the solution was brought to a pH of 8 using N-methylmorpholine. The reaction was mixed overnight at room temperature. The following day, the reaction mixture was pumped down to a viscous oil and was carried on without further purification (95% crude yield). $t_R = 6.67$ minutes. $M+H = 321.2$



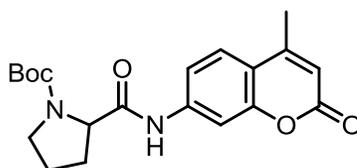
(1-(2-((3s,5s,7s)-adamantan-1-ylamino)ethyl)pyrrolidin-2-yl)methanol (3.17) – Based on 4 mmol of **3.16**, The methylester was dissolved in THF and cooled to 0°C in an ice bath. To this, 14 mmol of Lithium Aluminum Hydride was added portion-wise over 2 minutes to bring the concentration to 1M. The reaction was warmed to room temperature overnight. The next day, the reaction was cooled back down to 0°C on an ice bath and was diluted with DCM. The reaction was then quenched with water dropwise. The

reaction was then extracted with two more portions of equalvolume of DCM. The organic layer was dried over Na_2SO_4 and the DCM was removed under reduced pressure. HPLC showed complete conversion to the desired product. The crude reaction mixture was set aside for purification. $t_R = 4.11$ minutes. $M+H = 279.2$



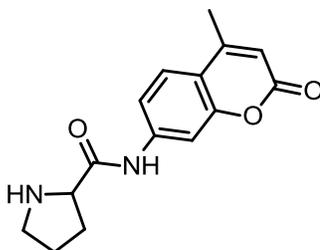
2-((3s,5s,7s)-adamantan-1-ylamino)-1-(2-(hydroxymethyl)pyrrolidin-1-yl)ethanone

(3.18) – The product of **3.16** (3.95mmol) was dissolved in a 1:2 mixture of Ethanol in THF and cooled on an ice bath. To this, 3.3 equivalents of NaBH_4 was added at 0°C . The reaction was mixed overnight while it was warmed to room temperature. The reaction was quenched with cold water and extracted three times with equalvolume of ethyl acetate. The combined organic layers were dried over Na_2SO_4 and the solvent was removed under reduced pressure. The crude product was carried on to preparative HPLC purification without further work up. Purified yield = 834.9mg (72.3%). $t_R = 6.23$ minutes. $M+H = 293.2$

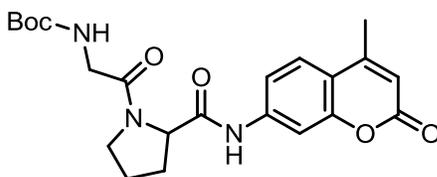


tert-butyl 2-((4-methyl-2-oxo-2H-chromen-7-yl)carbamoyl)pyrrolidine-1-carboxylate (3.30) – In a 20mL scintillation vial, 0.5mmol of Boc-L-Proline was dissolved in approximately 2mL of 1,4-Dioxanes. To this, 37 μL of pyridine was added

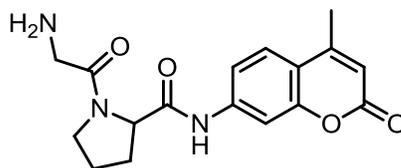
along with 134mgs of Boc-anhydride. This was allowed to mix for 10 minutes and then 107.86 mg of AMC. This was allowed to mix for 2.5 days at room temperature. Crude LCMS showed 50% conversion. Following completion of the reaction, the mixture was poured into a separatory funnel and diluted with DCM and washed with saturated bicarbonate solution (1x), water (1x), and brine (1x). The DCM was dried over anhydrous sodium anhydride and removed under reduced pressure resulting in a white solid. The crude material was carried on without further purification. $t_R = 10.79$ minutes. $M+H = 373.2$



N-(4-methyl-2-oxo-2H-chromen-7-yl)pyrrolidine-2-carboxamide (3.31) – The product **3.30** was dissolved in 4mL of a 1:1 mixture of TFA in DCM in a 20mL scintillation vial. The reaction was allowed to mix for 1.5 hours at room temperature. Upon completion of the reaction, the liquid was evaporated under reduced pressure. The remaining residue was cooled on dry ice and cold diethyl ether was used to precipitate the product as an off white solid (quant). The reaction was carried on without further purification. $t_R = 3.45$ minutes. $M+H = 273.2$



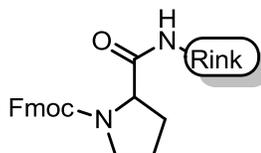
tert-butyl (2-(2-((4-methyl-2-oxo-2H-chromen-7-yl)carbamoyl)pyrrolidin-1-yl)-2-oxoethyl)carbamate (3.32) – The residue from **3.31** was dissolved in DCM and to this was added 1 equivalent of Boc-Glycine, 1.1 equivalents of HOBt, and 1.1 equivalents of DIC. Additionally, 3 equivalents of NMM were added to make sure the solution was basic. The reaction was mixed overnight at room temperature. Upon completion of the reaction, the DIU was filtered away and the DCM was washed with saturated bicarbonate solution (1x) and 2M HCl (1x). The DCM was then dried over anhydrous sodium sulfate and removed under reduced pressure resulting in an off white powder. The crude material was carried on without further work up. $t_R = 10.01$ minutes. $M+H = 430.4$



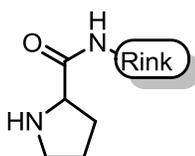
Gly-Pro-AMC (3.33) – The product of **3.32** was dissolved in 5mL solution of TFA in DCM (1:1). The reaction was mixed for 1.5 hours at room temperature. Following this, the solvent and TFA was removed under reduced pressure and the resulting oil was carried on to HPLC purification without work up (quant). $t_R = 3.11$ minutes. $M+H = 330.2$

Library Synthesis

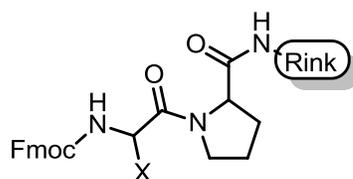
Deprotecting Fmoc-Rink Resin (3.19) – Rink resin is added to a peptide shaker with a solution of 20% piperidine in DMF. The contents are mixed on a shaker for 20 minutes. Following this, the liquid is drained from the vessel and the resin is washed with DMF (2x), Methanol (1x), and DMF (1x).



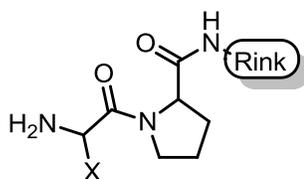
Coupling Fmoc-L-Pro-OH to deprotected Rink Resin (3.20) – To the vessel of deprotected resin in DMF, add 3 equivalents of Fmoc-L-Pro-OH, HOBt, and DIC. Allow the contents of the vessel to mix for 2 hours. Upon completion of the reaction, drain/filter the resin, and wash with DMF (3x).



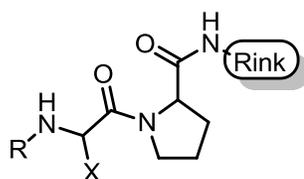
Deprotection of Fmoc-Pro-Resin (3.21) – To the peptide shaker containing the resin, a solution of 20% piperidine in DMF was added and mixed for 20 minutes. Following the deprotection, the resin was drained and washed with DMF (2x), Methanol (1x), and DMF (1x).



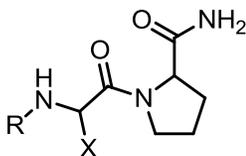
Coupling of amino acids to proline (3.22) – A selection of 14 normal and unusual L-amino acids were selected. Based on X mgs of Rink resin (.7mmol/g), 3 equivalents of amino acid were weighed out and distributed to the appropriate well of two 40 well synthesis blocks (39 wells in one block and 3 in another). HOBt (3 equivalents) was added to each well along with DMF. Following this, DIC was added to each well and the blocks were mixed overnight at room temperature. The next day, each block was drained and washed with DMF (3x). A negative chloranil test indicated the absence of a secondary amine. This would signify that each amino acid coupled successfully.



Deprotection of Amino Acid-Pro-Resin (3.23) – To each well of the synthesis block containing resin, a solution of 20% piperidine in DMF was added and mixed for 20 minutes. Following the deprotection, the block was drained and the resin was washed with DMF (2x), Methanol (1x), and DMF (1x).

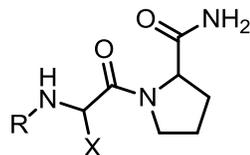


Alkylating the primary amine (3.24) – Three different amines were used in the synthesis of the alkylated amines. 1.1 equivalents of alkylbromides in DMF were added to the appropriate well in the synthesis block. Additionally, 2 equivalents of NMM were also added to each well. The reactions were allowed to mix for 24 hours. The following day, the block was drained and the resin was washed with DMF (3x). A positive chloranil test indicated a successful alkylation.

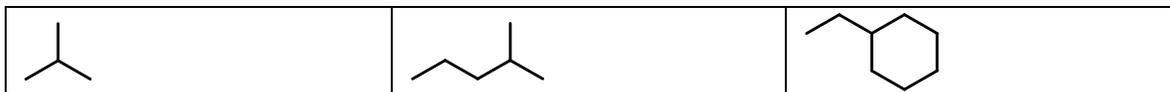


Cleavage from the resin to the primary amide – Each well was transferred to a 20mL scintillation vial and was mixed with a solution of 50/50 TFA/DCM. All 42 vials were placed on a shaker and mixed for 1-1.5 hours. Following this, the resin was filtered away and distilled water was added to each sample. Each sample was frozen and placed on the freeze dryer.

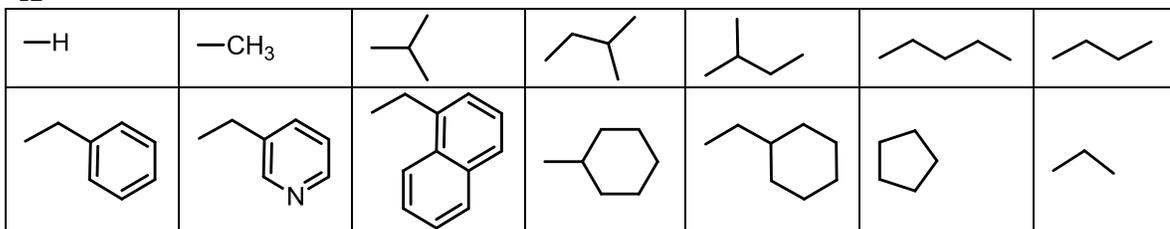
Library Monomers



R =



X =



Biological Protocols

Preparation of Buffer Solution – To make a 100mM solution of HEPES, 2.833 g of HEPES was dissolved in 100mL of distilled water. The pH was adjusted to 7.5 using 0.5M NaOH.

Preparation of Stock Solution of Substrate – In a 20mL scintillation vial, 6.7mg of Nle-Pro-AMC was dissolved in 5mL of the HEPES buffer solution to make a concentration of of 3.4mM.

Enzyme Concentration Determination Experiment – *Enzyme* preparation - Three enzyme concentrations were prepared by making a 1 μ unit/100 μ L stock solution. From

here, 20 μ L of solution was added to 80 μ L of buffer (total 100 μ L) for the 1/5 enzyme concentration. 10 μ L was placed in 90 μ L and for the 1/10 concentration, and 5 μ L enzyme stock was placed in 95 μ L of solution for the 1/20 concentration. Each of these enzyme concentrations were mixed with 50 μ M or 100 μ M substrate concentrations and readings were taken every 1.5 minutes for 2 hours and then every 10 minutes for the next 2hrs.

Substrate Inhibition Assay 1 – Substrate concentrations of 10, 20, 100, 200, 300, 400, and 500 μ M were prepared from the 3.4mM stock solution (Nle-Pro-AMC). 2 μ unit of DPP-IV enzyme was placed in 200 μ L of 100mM HEPES buffer. In each well, 20 μ L of enzyme solution was added along with 80 μ L of buffer. Then to each of the 7 wells, 100 μ L of each of the seven substrate concentrations were placed in a respective well. Readings were taken as often as possible for 1 hour. When the plate was not being used for readings, it was kept in an incubator at 37°C.

Substrate Inhibition Assay 2 - Substrate concentrations of 2, 4, 10, 20, 40, and 100 μ M were prepared from the 3.4mM stock solution (Nle-Pro-AMC). 2 μ unit of DPP-IV enzyme was placed in 200 μ L of 100mM HEPES buffer. In each well, 20 μ L of enzyme solution was added along with 80 μ L of buffer. Then to each of the six wells, 100 μ L of each of the six substrate concentrations were placed in a respective well. Readings were taken as often as possible for 1 hour. When the plate was not being used for readings, it was kept in an incubator at 37°C.

Assay of Compound 3.1 – Sample Preparation – Compound **3.1** (11.5mgs) was dissolved in 1.15mL of DMSO producing a 34800 μ M solution. From this, a 1000 μ M stock solution was prepared. The stock solution was used to prepare 10, 20, 40, 100, and 200 μ M solutions of **3.1**. *Substrate Preparation* – Substrate solutions of 8, 40, 80, 200, and 400 μ M were prepared from the 3400 μ M stock solution of Nle-Pro-AMC. *Enzyme Preparation* - 3 μ L of DPP-IV enzyme (2.84 μ unit/1 μ L) was dissolved in 4260 μ L of 100mM HEPES buffer solution (to be used for 3 experiments). *Assay* – A 5x5 matrix was set up to assay varying substrate and inhibitor concentrations with a constant enzyme concentration. In each well, 50 μ L of substrate, 100 μ L of inhibitor **3.1**, and 50 μ L enzyme solution were added. Readings were taken after 1 minute and then taken every 3.5 minutes. Final substrate concentrations were 2, 10, 20, 50, and 100 μ M. Final **3.1** inhibitor concentrations were 5, 10, 20, 50, 100 μ M.

Assay of Compound 3.2 – Sample Preparation – Compound **3.2** (5.0mgs) was dissolved in 0.5mL of DMSO producing a 42600 μ M solution. From this, a 1000 μ M stock solution was prepared. The stock solution was used to prepare 10, 20, 40, 100, and 200 μ M solutions of **3.2**. *Substrate Preparation* – Substrate solutions of 8, 40, 80, 200, and 400 μ M were prepared from the 3400 μ M stock solution of Nle-Pro-AMC. *Enzyme Preparation* - 3 μ L of DPP-IV enzyme (2.84 μ unit/1 μ L) was dissolved in 4260 μ L of 100mM HEPES buffer solution (to be used for 3 experiments). *Assay* – A 5x5 matrix was set up to assay varying substrate and inhibitor concentrations with a constant enzyme concentration. In each well, 50 μ L of substrate, 100 μ L of inhibitor **3.2**, and 50 μ L enzyme solution were added. Readings were taken after 1 minute and then taken every 3.5

minutes. Final substrate concentrations were 2, 10, 20, 50, and 100 μ M. Final **3.2** inhibitor concentrations were 5, 10, 20, 50, 100 μ M.

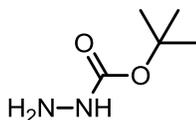
Assay of Compound 3.3 – Sample Preparation – Compound **3.3** (4.3mg) was dissolved in 0.5mL of DMSO producing a 40200 μ M solution. From this, a 1000 μ M stock solution was prepared. The stock solution was used to prepare 10, 20, 40, 100, and 200 μ M solutions of **3.3**. **Substrate Preparation** – Substrate solutions of 8, 40, 80, 200, and 400 μ M were prepared from the 3400 μ M stock solution of Nle-Pro-AMC. **Enzyme Preparation** - 3 μ L of DPP-IV enzyme (2.84 μ unit/1 μ L) was dissolved in 4260 μ L of 100mM HEPES buffer solution (to be used for 3 experiments). **Assay** – A 5x5 matrix was set up to assay varying substrate and inhibitor concentrations with a constant enzyme concentration. In each well, 50 μ L of substrate, 100 μ L of inhibitor **3.3**, and 50 μ L enzyme solution were added. Readings were taken after 1 minute and then taken every 3.5 minutes. Final substrate concentrations were 2, 10, 20, 50, and 100 μ M. Final **3.3** inhibitor concentrations were 5, 10, 20, 50, 100 μ M.

Substrate Inhibition Assay 3 - Substrate concentrations of 10, 20, 100, 200, 300, 400, and 500 μ M were prepared from the 3.4mM stock solution (Gly-Pro-AMC). 2 μ unit of DPP-IV enzyme was placed in 200 μ L of 100mM HEPES buffer. In each well, 20 μ L of enzyme solution was added along with 80 μ L of buffer. Then to each of the 7 wells, 100 μ L of each of the 7 buffers were placed in a respective well. Readings were taken as often as possible for 1 hour. When the plate was not being used for readings, it was kept in an incubator at 37°C.

Experimental Protocols for Experiments in Chapter 6

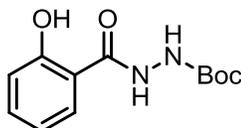
General procedure A for the final imine synthesis - The deprotected product of **6.10** was dissolved in a 1:1 mixture of triethyl orthoformate in DCM. To this, 1.1 equivalents of benzaldehyde and 2 equivalents of NMM were added at room temperature. The reactions were mixed overnight at room temperature. LCMS analysis of the reaction mixture indicates the successful production of product. The reaction mixture was pumped down and carried on crude to preparative HPLC purification.

General procedure B - amide bond formations which lead to final products - 1 mmol of the carboxylic acid was dissolved in DCM. To this, 1.1 equivalents of HOBT and the desired primary amine were added. Following this, add 1 eq of DIC was added to the mixture and the reaction was mixed overnight. Upon completion of the reaction, the DIU was filtered away and the DMC layer was washed with saturated bicarbonate solution (1x) and 2M HCl (1x). The organic layer was dried over anhydrous sodium sulfate and the solvent was then removed under reduced pressure. Typical yields for all amide bond formations were $\geq 80\%$ resulting in white powders. The resulting compounds were carried through and a small portion was used for HPLC purification.

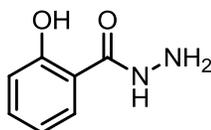


tert-butyl hydrazinecarboxylate (6.8) – To a round bottom flask equipped with a stir bar, 25mmol of hydrazine monohydrate was dissolved in THF at 0°C. Added dropwise to this was a solution of Boc₂O. The reaction was allowed to mix at 0°C for an additional

hour. The reaction was then diluted with ethyl acetate and washed with water (2x). The organic layer was dried over Na_2SO_4 and the solvent was removed under reduced pressure. 3.53mmol of desired product was collected (70.6% yield). $t_R = 7.23$ minutes. $M+H = 133.2$

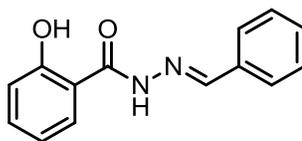


tert-butyl 2-(2-hydroxybenzoyl)hydrazinecarboxylate (6.9) – The 3.53mmol of **6.8** was dissolved in DCM in a round bottom flask with a stir bar. To this, 1.1 equivalents of HOBt and salicylic acid were added at room temperature. Following this, 1 equivalent of DIC was added to the flask. The reaction was mixed overnight at room temperature. Upon completion of the reaction, the DIU was filtered away and the resulting DCM was washed with saturated bicarbonate solution (1x) and 2M HCl (1x). The organic layer was dried over Na_2SO_4 and pumped down under reduced pressure to produce a white solid. This produced 3.34mmol of the desired product (94.6% yield). $t_R = 9.68$ minutes. $M+H = 253.2$

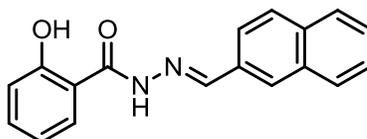


2-hydroxybenzohydrazide (6.10) – The 3.34mmol of **6.9** which was collected was dissolved in a 1:1 mixture of TFA in DCM in a round bottom flask. The reaction was mixed for one hour and then the excess TFA and DCM was pumped away for 1 hour

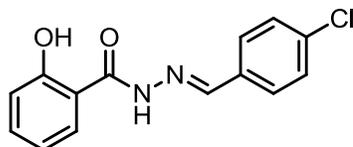
producing a clear oil (quant). The crude reaction material was carried on to the next step without further purification steps.



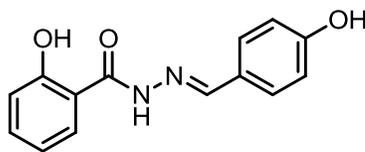
N'-benzylidene-2-hydroxybenzohydrazide (6.2) – General procedure A was used to form the imine bond using benzaldehyde to produce a white powder. ^1H NMR (300 MHz, DMSO) δ 11.92 (s, 1H), 8.45 (s, 1H), 7.87 (d, $J = 6.4$ Hz, 1H), 7.76 – 7.70 (m, 2H), 7.49 – 7.37 (m, 4H), 7.00 – 6.87 (m, 2H), 6.54 (s, 1H). $t_{\text{R}} = 10.45$ minutes. $\text{M}+\text{H} = 241.3$



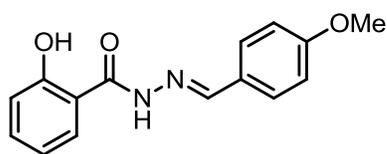
2-hydroxy-N'-(naphthalen-2-ylmethylene)benzohydrazide (6.3) - General procedure A was used to form the imine bond using naphthalaldehyde to produce a solid white powder. $t_{\text{R}} = 11.68$ minutes. $\text{M}+\text{H} = 291.0$



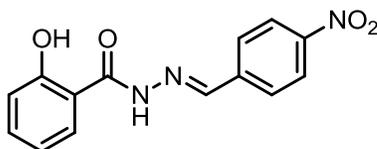
N'-(4-chlorobenzylidene)-2-hydroxybenzohydrazide (6.4) - General procedure A was used to form the imine bond using 4-chlorobenzaldehyde and produced a white powder. $t_{\text{R}} = 11.45$ minutes. $\text{M}+\text{H} = 274.9$



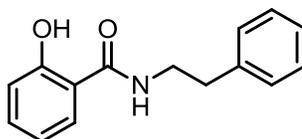
2-hydroxy-N'-(4-hydroxybenzylidene)benzohydrazide (6.5) - General procedure A was used to form the imine bond using 4-hydroxybenzaldehyde. A white powder was produced after HPLC purification. $t_R = 8.90$ minutes. $M+H = 257.1$



2-hydroxy-N'-(4-methoxybenzylidene)benzohydrazide (6.6) - General procedure A was used to form the imine bond using 4-methoxybenzaldehyde. A white powder was produced after purification. $t_R = 10.34$ minutes. $M+H = 271.2$

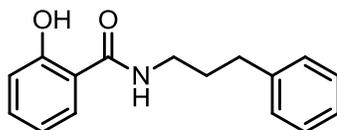


2-hydroxy-N'-(4-nitrobenzylidene)benzohydrazide (6.7) - General procedure A was used to form the imine bond using 4-nitrobenzaldehyde and after purification, a white powder was produced. $t_R = 10.68$ minutes. $M+H = 285.9$

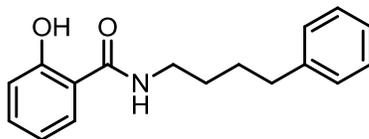


2-hydroxy-N-phenethylbenzamide (6.11) – General Procedure B was used to couple Salicylic acid to phenethylamine. The reaction was carried out at room temperature

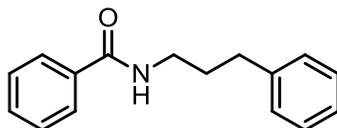
overnight and after HPLC purification, a white powder was produced. $t_R = 12.12$ minutes. $M+H = 242.3$



2-hydroxy-N-(3-phenylpropyl)benzamide (6.12) - General Procedure B was used to couple Salicylic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight. A white solid was produced after purification. ^1H NMR (300 MHz, CDCl_3) δ 12.42 (s, 1H), 7.44 – 7.17 (m, 6H), 7.12 (d, $J = 7.9$ Hz, 1H), 7.00 (d, $J = 8.4$ Hz, 1H), 6.82 (t, $J = 7.6$ Hz, 1H), 6.29 (s, 1H), 3.53 (dd, $J = 13.1, 6.4$ Hz, 2H), 2.77 (t, $J = 7.4$ Hz, 2H), 2.02 (t, 2H). $t_R = 13.90$ minutes. $M+H = 256.3$



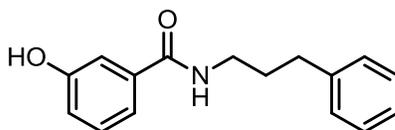
2-hydroxy-N-(4-phenylbutyl)benzamide (6.13) – General Procedure B was used to couple Salicylic acid to 4-phenylbutylamine. The reaction was carried out at room temperature overnight. $t_R = 13.90$ minutes. $M+H = 270.4$



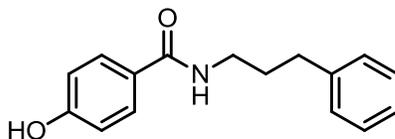
N-(3-phenylpropyl)benzamide (6.14) - General Procedure B was used to couple benzoic acid to 3-phenylpropylamine. The reaction was carried out at room temperature

overnight and produced a white powder after preparative HPLC. $t_R = 11.34$ minutes.

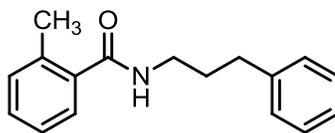
$M+H = 240.3$



3-hydroxy-N-(3-phenylpropyl)benzamide (6.15) - General Procedure B was used to couple 3-hydroxybenzoic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight and produced a white powder. $t_R = 9.90$ minutes. $M+H = 256.2$

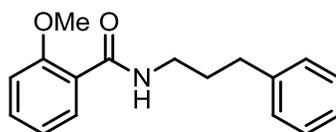


4-hydroxy-N-(3-phenylpropyl)benzamide (6.16) - General Procedure B was used to couple 4-hydroxybenzoic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight. A white solid was produced after purification. $t_R = 9.98$ minutes. $M+H = 256.3$

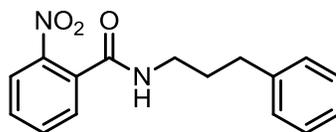


2-methyl-N-(3-phenylpropyl)benzamide (6.17) - General Procedure B was used to couple 2-methylbenzoic acid to 3-phenylpropylamine producing a white solid. The reaction was carried out at room temperature overnight and subsequently purified using HPLC. $^1\text{H NMR}$ (300 MHz, CDCl_3) δ 7.36 – 7.10 (m, 9H), 5.79 (s, 1H), 3.52 – 3.42 (m,

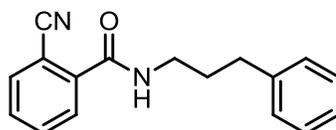
2H), 2.76 – 2.67 (t, 2H), 2.43 (s, 3H), 2.00 – 1.89 (m, 2H). $t_R = 11.68$ minutes. $M+H = 254.2$



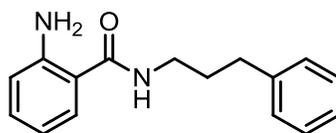
2-methoxy-N-(3-phenylpropyl)benzamide (6.18) - General Procedure B was used to couple 2-methoxybenzoic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight and produced a white solid. $t_R = 11.90$ minutes. $M+H = 270.3$



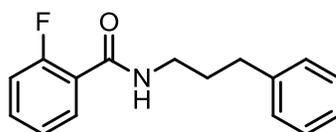
2-nitro-N-(3-phenylpropyl)benzamide (6.19) - General Procedure B was used to couple 2-nitrobenzoic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight which produced a white powder. $t_R = 11.01$ minutes. $M+H = 285.2$



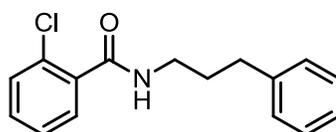
2-cyano-N-(3-phenylpropyl)benzamide (6.20) - General Procedure B was used to couple 2-cyanobenzoic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight. A white solid was produced after purification. $t_R = 11.90$ minutes. $M+H = 265.2$



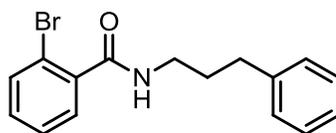
2-amino-N-(3-phenylpropyl)benzamide (6.21) - General Procedure B was used to couple 2-aminobenzoic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight and produced a white solid. $t_R = 12.01$ minutes. $M+H = 255.3$



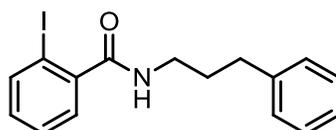
2-fluoro-N-(3-phenylpropyl)benzamide (6.22) - General Procedure B was used to couple 2-fluorobenzoic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight which produced a white solid after purification. $t_R = 12.34$ minutes. $M+H = 258.3$



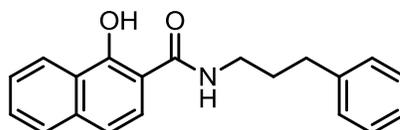
2-chloro-N-(3-phenylpropyl)benzamide (6.23) - General Procedure B was used to couple 2-chlorobenzoic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight. After HPLC purification, a white solid was produced. 1H NMR (300 MHz, $CDCl_3$) δ 7.67 – 7.54 (m, 1H), 7.42 – 7.10 (m, 8H), 6.24 (s, 1H), 3.50 (q, $J = 10.0$ Hz, 2H), 2.72 (t, $J = 17.7, 9.6$ Hz, 2H), 2.05 – 1.89 (m, 2H). $t_R = 11.57$ minutes. $M+H = 274.1$



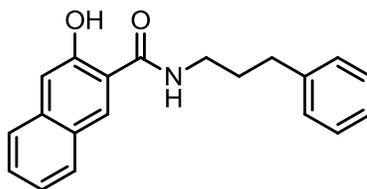
2-bromo-N-(3-phenylpropyl)benzamide (6.24) - General Procedure B was used to couple 2-bromobenzoic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight and produced a white solid. ^1H NMR (300 MHz, CDCl_3) δ 7.58 (d, 1H), 7.49 (dd, $J = 7.6, 1.8$ Hz, 1H), 7.39 – 7.15 (m, 7H), 6.02 (s, 1H), 3.50 (q, $J = 13.4, 6.7$ Hz, 2H), 2.75 (t, 2H), 2.03 – 1.89 (m, 2H). $t_{\text{R}} = 12.01$ minutes. $\text{M}+\text{H} = 320.1$



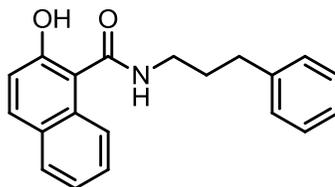
2-iodo-N-(3-phenylpropyl)benzamide (6.25) - General Procedure B was used to couple 2-iodobenzoic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight and produced a white solid after purification. ^1H NMR (300 MHz, CDCl_3) δ 7.85 (d, 1H), 7.40 – 7.03 (m, 8H), 5.77 (s, 1H), 3.50 (q, 2H), 2.76 (t, 2H), 2.05 – 1.92 (m, 2H). $t_{\text{R}} = 12.34$ minutes. $\text{M}+\text{H} = 366.1$



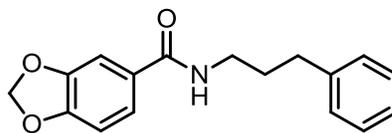
1-hydroxy-N-(3-phenylpropyl)-2-naphthamide (6.26) - General Procedure B was used to couple 1-hydroxy-2-naphthoic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight. After purification, a white solid was produced. $t_{\text{R}} = 17.02$ minutes. $\text{M}+\text{H} = 306.2$



3-hydroxy-N-(3-phenylpropyl)-2-naphthamide (6.27) - General Procedure B was used to couple 3-hydroxy-2-naphthoic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight. HPLC purification produced a white solid. $t_R = 15.57$ minutes. $M+H = 306.3$



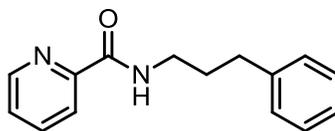
2-hydroxy-N-(3-phenylpropyl)-1-naphthamide (6.28) - General Procedure B was used to couple 2-hydroxy-1-naphthoic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight and produced a white solid after HPLC purification. $t_R = 13.12$ minutes. $M+H = 306.3$



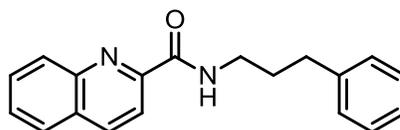
N-(3-phenylpropyl)benzo[d][1,3]dioxole-5-carboxamide (6.29) - General Procedure B was used to couple Piperonylic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight and produced a white solid. $t_R = 12.12$ minutes. $M+H = 284.2$



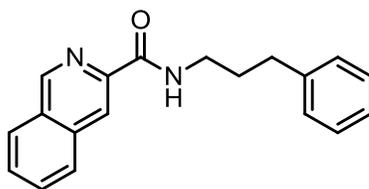
3,5-dihydroxy-N-(3-phenylpropyl)-2-naphthamide (6.30) - General Procedure B was used to couple 3,5-dihydroxy-2-naphthoic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight. A white solid was produced after HPLC purification. $t_R = 12.90$ minutes. $M+H = 322.3$



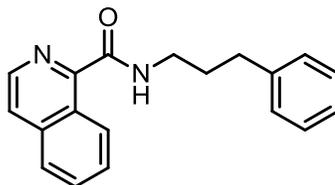
N-(3-phenylpropyl)picolinamide (6.31) – General Procedure B was used to couple picolinic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight. A white powder was produced after purification. $t_R = 12.57$ minutes. $M+H = 241.4$



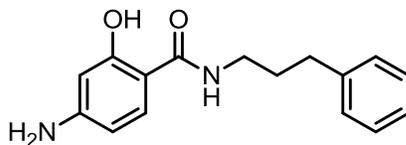
N-(3-phenylpropyl)quinoline-2-carboxamide (6.32) – General Procedure B was used to couple quinoline-2-carboxylic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight and produced a white solid after purification. $t_R = 14.90$ minutes. $M+H = 291.3$



N-(3-phenylpropyl)isoquinoline-3-carboxamide (6.33) – General Procedure B was used to couple isoquinoline-3-carboxylic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight. Following HPLC purification, a white powder was produced. $t_R = 14.12$ minutes. $M+H = 291.3$



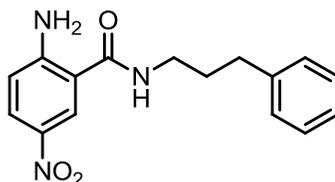
N-(3-phenylpropyl)isoquinoline-1-carboxamide (6.34) – General Procedure B was used to couple isoquinoline-1-carboxylic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight and produced a white solid after preparative HPLC purification. $t_R = 14.35$ minutes. $M+H = 291.3$



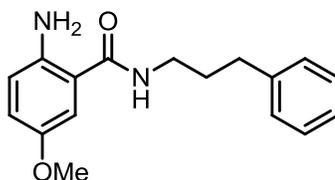
4-amino-2-hydroxy-N-(3-phenylpropyl)benzamide (6.35) - General Procedure B was used to couple isoquinoline-1-carboxylic acid to 3-phenylpropylamine. The reaction was

carried out at room temperature overnight and produced a white solid after purification.

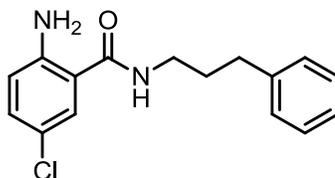
$t_R = 9.34$ minutes. $M+H = 271.3$



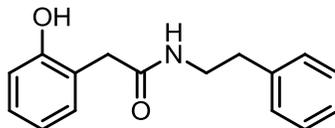
2-amino-5-nitro-N-(3-phenylpropyl)benzamide (6.36) - General Procedure B was used to couple 2-amino-5-nitrobenzoic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight and produced a white solid after purification. $t_R = 12.79$ minutes. $M+H = 300.4$



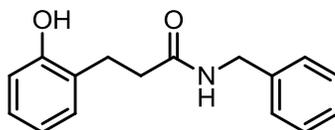
2-amino-5-methoxy-N-(3-phenylpropyl)benzamide (6.37) - General Procedure B was used to couple 2-amino-5-methoxybenzoic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight. A white solid was produced after purification. $t_R = 9.90$ minutes. $M+H = 285.3$



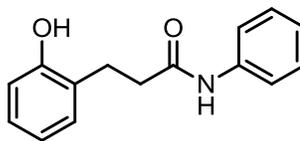
2-amino-5-chloro-N-(3-phenylpropyl)benzamide (6.38) - General Procedure B was used to couple 2-amino-5-chlorobenzoic acid to 3-phenylpropylamine. The reaction was carried out at room temperature overnight and produced a white solid. $t_R = 13.35$ minutes. $M+H = 289.3$



2-(2-hydroxyphenyl)-N-phenethylacetamide (6.39) - General Procedure B was used to couple 2-(2-hydroxyphenyl)acetic acid to 2-phenethylamine. The reaction was carried out at room temperature overnight. Following purification, a white solid was produced. $t_R = 11.23$ minutes. $M+H = 256.2$

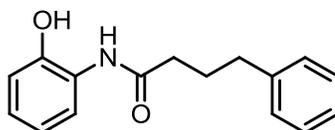


N-benzyl-3-(2-hydroxyphenyl)propanamide (6.40) - General Procedure B was used to couple 3-(2-hydroxyphenyl)propanoic acid to benzylamine. The reaction was carried out at room temperature overnight. HPLC purification produced a white solid. $t_R = 10.90$ minutes. $M+H = 256.3$

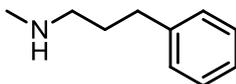


3-(2-hydroxyphenyl)-N-phenylpropanamide (6.41) - In a round bottom flask, 1mmol of dihydrocoumarin was dissolved in DCM. To this, 1.1mmol of aniline was added with

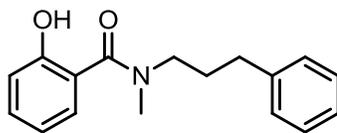
.25mmol of DMAP. The reaction was mixed overnight at room temperature. Following this, the reaction was heated for 8 hrs at 70°C. Upon completion of the reaction, the DCM was washed with saturated bicarbonate solution (1x) and with 2M HCl (1x). The DCM was dried over Na₂SO₄ and the solvent was removed under reduced pressure. The crude reaction material was carried on to preparative HPLC without further work-up and produced a white solid. $t_R = 10.79$ minutes. M+H = 242.4



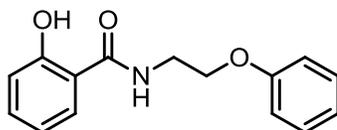
N-(2-hydroxyphenyl)-4-phenylbutanamide (6.42) - General Procedure B was used to couple 4-phenylbutanoic acid to 2-aminophenol. The reaction was carried out at room temperature overnight. This produced a white solid following HPLC purification. $t_R = 11.68$ minutes. M+H = 256.2



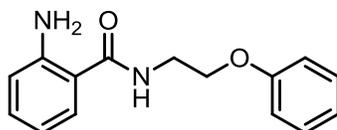
N-methyl-3-phenylpropan-1-amine (6.43) - 2mmol of 3-phenylpropylamine and 2.2mmol of N-methylmorpholine was dissolved in THF at room temperature in a round bottom flask. To this, iodomethane was added to the flask in one portion. The reaction was mixed overnight at room temperature. The next day, the precipitate was filtered away and the THF was evaporated away under reduced pressure. The crude mixture was carried on without any work-up. $t_R = 9.12$ minutes. M+H = 150.0



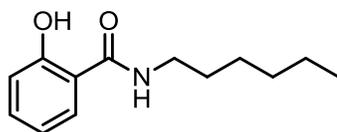
2-hydroxy-N-methyl-N-(3-phenylpropyl)benzamide (6.44) - General Procedure B was used to couple Salicylic acid to N-methyl-3-phenylpropan-1-amine. The reaction was carried out at room temperature overnight and produced a white solid. $t_R = 11.12$ minutes. $M+H = 270.3$



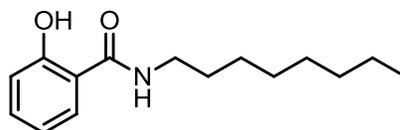
2-hydroxy-N-(2-phenoxyethyl)benzamide (6.45) - General Procedure B was used to couple salicylic acid to 2-phenoxyethanamine. The reaction was carried out at room temperature overnight. Following HPLC purification, a white solid was produced. $t_R = 13.12$ minutes. $M+H = 258.2$



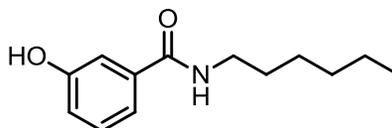
2-amino-N-(2-phenoxyethyl)benzamide (6.46) - General Procedure B was used to couple 2-aminobenzoic acid to 2-phenoxyethanamine. The reaction was carried out at room temperature overnight and produced a white solid following HPLC purification. $t_R = 12.01$ minutes. $M+H = 257.2$



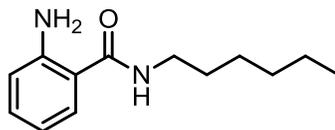
N-hexyl-2-hydroxybenzamide (6.47) - General Procedure B was used to couple Salicylic acid to n-hexylamine. The reaction was carried out at room temperature overnight. Following preparative HPLC purification, a white solid was produced. ^1H NMR (300 MHz, CDCl_3) δ 12.45 (s, 1H), 7.40 (m, 2H), 7.04 – 6.95 (m, 1H), 6.86 (ddd, J = 8.0, 7.3, 1.2 Hz, 1H), 6.39 (s, 1H), 3.51 – 3.41 (m, 2H), 1.71 – 1.56 (m, 6H), 1.47 – 1.26 (m, 2H), 0.92 (t, J = 6.9 Hz, 3H). t_{R} = 13.90 minutes. $\text{M}+\text{H}$ = 222.2



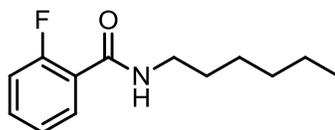
2-hydroxy-N-octylbenzamide (6.48) - General Procedure B was used to couple Salicylic acid to octylamine. The reaction was carried out at room temperature overnight and produced a white solid after HPLC purification. ^1H NMR (300 MHz, CDCl_3) δ 12.41 (s, 1H), 7.43 – 7.29 (m, 2H), 6.98 (dd, J = 8.4, 0.8 Hz, 1H), 6.84 (m, 1H), 6.29 (s, 1H), 3.44 (td, J = 7.2, 5.9 Hz, 2H), 1.69 – 1.53 (m, 2H), 1.29 (dd, J = 15.2, 11.6 Hz, 10H), 0.88 (t, J = 6.7 Hz, 3H). t_{R} = 16.68 minutes. $\text{M}+\text{H}$ = 250.4



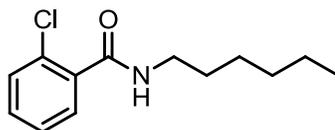
N-hexyl-3-hydroxybenzamide (6.49) - General Procedure B was used to couple 3-hydroxybenzoic acid to n-hexylamine. The reaction was carried out at room temperature overnight and produced a white solid following purification. t_{R} = 10.45 minutes. $\text{M}+\text{H}$ = 222.4



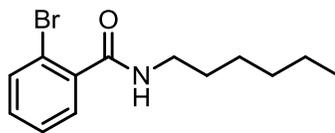
2-amino-N-hexylbenzamide (6.50) - General Procedure B was used to couple 2-aminobenzoic acid to n-hexylamine. The reaction was carried out at room temperature overnight. HPLC purification produced a white solid. $t_R = 12.68$ minutes. $M+H = 222.4$



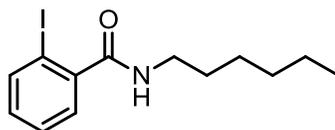
2-fluoro-N-hexylbenzamide (6.51) - General Procedure B was used to couple 2-fluorobenzoic acid to n-hexylamine. The reaction was carried out at room temperature overnight. A white solid was obtained after HPLC purification. $t_R = 12.90$ minutes. $M+H = 224.3$



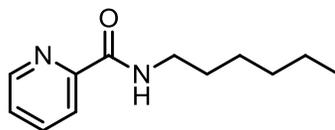
2-chloro-N-hexylbenzamide (6.52) - General Procedure B was used to couple 2-chlorobenzoic acid to n-hexylamine. The reaction was carried out at room temperature overnight. The reaction produced a white solid after HPLC purification. $t_R = 12.23$ minutes. $M+H = 224.3$



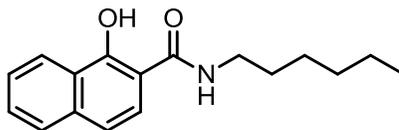
2-bromo-N-hexylbenzamide (6.53) - General Procedure B was used to couple 2-bromobenzoic acid to n-hexylamine. The reaction was carried out at room temperature overnight and produced a white solid after purification. ^1H NMR (300 MHz, CDCl_3) δ 7.60 – 7.49 (m, 2H), 7.39 – 7.21 (m, 2H), 5.96 (s, 1H), 3.54 – 3.37 (m, 2H), 1.69 – 1.54 (m, 2H), 1.47 – 1.27 (m, 5H), 0.89 (t, $J = 6.8$ Hz, 3H). $t_{\text{R}} = 12.68$ minutes. $\text{M}+\text{H} = 286.2$



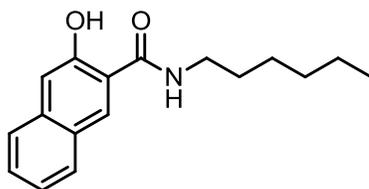
N-hexyl-2-iodobenzamide (6.54) - General Procedure B was used to couple 2-iodobenzoic acid to n-hexylamine. The reaction was carried out at room temperature overnight. HPLC purification produced the compound as a white solid. ^1H NMR (300 MHz, CDCl_3) δ 7.85 (m, 1H), 7.39 – 7.35 (m, 1H), 7.26 (t, $J = 0.8$ Hz, 1H), 7.14 – 7.00 (m, 1H), 3.44 (m, 2H), 1.62 (m, 2H), 1.37 (m, 5H), 0.90 (t, $J = 6.5$ Hz, 3H). $t_{\text{R}} = 13.01$ minutes. $\text{M}+\text{H} = 322.1$



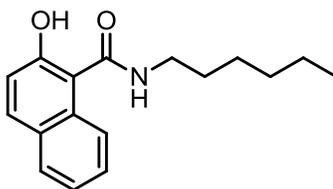
N-hexylpicolinamide (6.55) - General Procedure B was used to couple picolinic acid to n-hexylamine. The reaction was carried out at room temperature overnight. HPLC purification produced the compound as a white solid. $t_{\text{R}} = 13.46$ minutes. $\text{M}+\text{H} = 207.3$



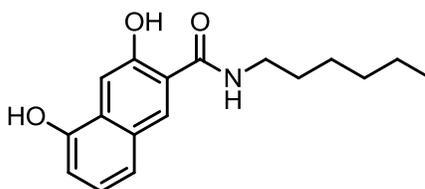
N-hexyl-1-hydroxy-2-naphthamide (6.56) - General Procedure B was used to couple 1-hydroxy-2-naphthoic acid to n-hexylamine. The reaction was carried out at room temperature overnight. A white solid was produced following purification of the compound. $t_R = 18.91$ minutes. $M+H = 272.4$



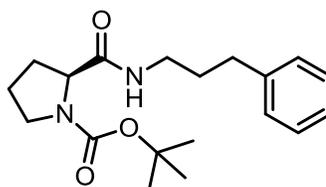
N-hexyl-3-hydroxy-2-naphthamide (6.57) - General Procedure B was used to couple 3-hydroxy-2-naphthoic acid to n-hexylamine. The reaction was carried out at room temperature overnight and produced a white solid following HPLC purification of the compound. $t_R = 17.24$ minutes. $M+H = 272.4$



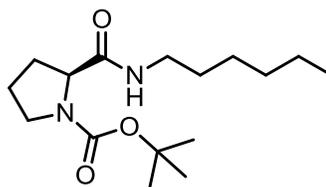
N-hexyl-2-hydroxy-1-naphthamide (6.58) - General Procedure B was used to couple 2-hydroxy-1-naphthoic acid to n-hexylamine. The reaction was carried out at room temperature overnight. HPLC purification of the compound produced a white solid. $t_R = 14.21$ minutes. $M+H = 272.4$



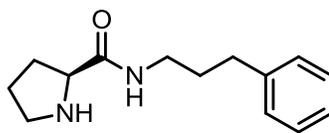
N-hexyl-3,5-dihydroxy-2-naphthamide (6.59) - General Procedure B was used to couple 2-hydroxy-1-naphthoic acid to n-hexylamine. The reaction was carried out at room temperature overnight. Following preparative HPLC purification, a white solid was produced. $t_R = 12.12$ minutes. $M+H = 288.4$



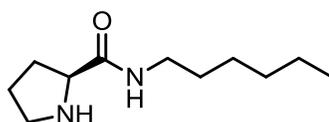
(S)-tert-butyl 2-((3-phenylpropyl)carbamoyl)pyrrolidine-1-carboxylate (6.60) - General Procedure B was used to couple Boc-L-Proline to 3-phenylpropylamine. The reaction was carried out at room temperature overnight. A white powder was produced after purification. $t_R = 12.68$ minutes. $M+H = 333.3$



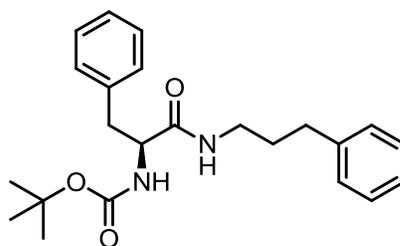
(S)-tert-butyl 2-(hexylcarbamoyl)pyrrolidine-1-carboxylate (6.61) - General Procedure B was used to couple Boc-L-Proline to n-hexylamine. The reaction was carried out at room temperature overnight and a white powder was produced after purification. $t_R = 13.23$ minutes. $M+H = 299.5$



(S)-N-(3-phenylpropyl)pyrrolidine-2-carboxamide (6.62) - 58 was dissolved in a 50/50 mixture of TFA/DCM in a round bottom flask. The reaction was mixed for room temperature for 1 hour. Upon completion, the TFA and DCM were pumped away on a rotary evaporator for 1 hour resulting in a clear oil. The crude material was carried on for HPLC purification without further work up. Following this, a white solid was produced. $t_R = 8.56$ minutes. $M+H = 233.3$

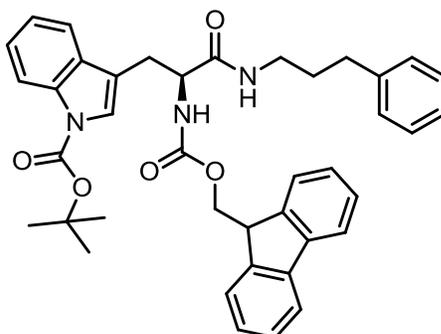


(S)-N-hexylpyrrolidine-2-carboxamide (6.63) - 59 was dissolved in a 50/50 mixture of TFA/DCM in a round bottom flask. The reaction was mixed for room temperature for 1 hour. Upon completion, the TFA and DCM were pumped away on a rotary evaporator for 1 hour resulting in a clear oil. The crude material was carried on for HPLC purification and produced a white powder. $t_R = 9.01$ minutes. $M+H = 199.3$

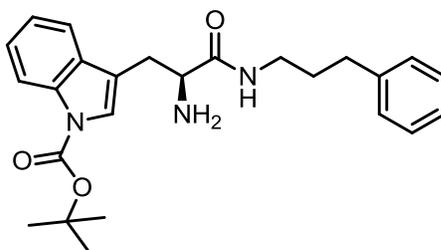


(S)-tert-butyl (1-oxo-3-phenyl-1-((3-phenylpropyl)amino)propan-2-yl)carbamate (6.66) - General Procedure B was used to couple Boc-L-Phenylalanine to 3-

phenylpropylamine. The reaction was carried out at room temperature overnight. The reaction produced an off-white solid. $t_R = 13.79$ minutes. $M+H = 383.3$

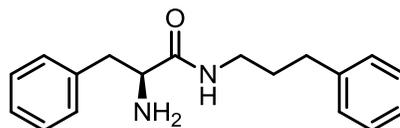


(S)-tert-butyl 3-(2-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-oxo-3-((3-phenylpropyl)amino)propyl-1H-indole-1-carboxylate (6.66) - General Procedure B was used to couple Fmoc-L-Tryptophan to 3-phenylpropylamine. The reaction was carried out at room temperature overnight. After removal of the solvent, a yellow solid was produced. $t_R = 9.68$ minutes. $M+H = 644.4$

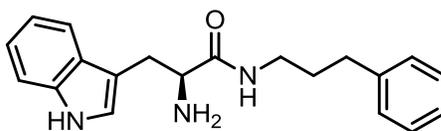


(S)-tert-butyl 3-(2-amino-3-oxo-3-((3-phenylpropyl)amino)propyl)-1H-indole-1-carboxylate (6.67) – The product **6.66** was dissolved in 10mL of 20% piperidine in Methanol. The reaction was mixed at room temperature for 3 hours. Upon completion, the solvent was evaporated and the resulting residue was taken back into DCM. The DCM was washed with approximately equal volume of saturated bicarbonate solution (1x). The organic layer was dried over Na_2SO_4 and evaporated under reduced pressure

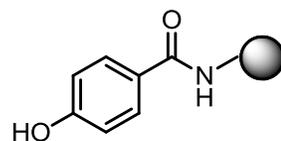
producing a light yellow residue. The resulting product was carried on without further purification. $t_R = 9.68$ minutes. $M+H = 422.4$



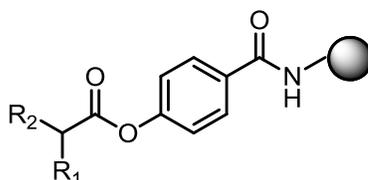
(S)-2-amino-3-phenyl-N-(3-phenylpropyl)propanamide (6.65) – 6.64 was dissolved in a 50/50 mixture of TFA/DCM in a round bottom flask. The reaction was mixed for room temperature for 1 hour. Upon completion, the TFA and DCM were pumped away on a rotary evaporator for 1 hour. The crude oil was carried on for HPLC purification without further work up. $t_R = 9.23$ minutes. Following purification, a white solid was obtained. $t_R = 9.23$ minutes. $M+H = 283.2$



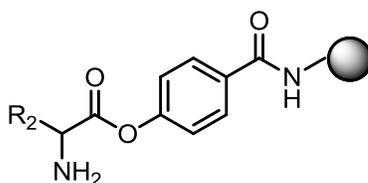
(S)-2-amino-3-(1H-indol-3-yl)-N-(3-phenylpropyl)propanamide (6.68) – 6.67 was dissolved in a 50/50 mixture of TFA/DCM in a round bottom flask. The reaction was mixed for room temperature for 1 hour. Upon completion, the TFA and DCM were pumped away on a rotary evaporator for 1 hour. The crude oil was carried on for HPLC purification without further work up. A white solid was obtained following purification. $t_R = 9.34$ minutes. $M+H = 322.3$



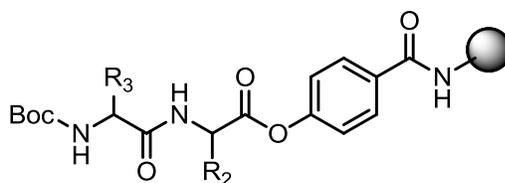
Attachment of 4-hydroxybenzoic acid to the resin (6.69) - Take 10g of hydroxyl methyl resin (which had β -alanine coupled to it; loading: 1mmol/g) and to this, add 3 equivalents of 4-hydroxybenzoic acid, 3 equivalents of HOBt, 3 equivalents of DIC in DMF. This was allowed to mix over night at room temperature. The next day, the resin was drained and washed with DMF (3x).



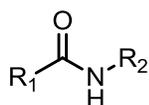
Ester Formation Using Carboxylic Acids (6.70) – Based on 100mg of resin per well, a series of 42 carboxylic acids (mostly L and D amino acids) were selected to attach to the resin in a 96-well synthesis block. Wells 1-42 each received a different carboxylic acid (5eq) and HOBt (5eq), which was dissolved in DMF. Wells 43-84 also received the same 42 carboxylic acid respectively. Wells 85-96 were treated with 5eq of phenylalanine and 5eq of HOBt. Each well also received 5 eq of DIC. After 1 hour, the wells were drained and another 5 equivalents of each respective carboxylic acid, HOBt, and DIC were added. The mixtures were shaken in a synthesis block overnight at room temperature. The following day, the block was drained and the wells were washed with DMF (5x).



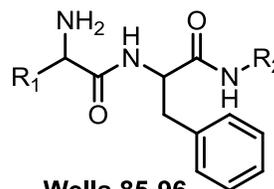
Deprotection of wells 85-96 – Wells 85-96 were collected in a peptide shaker vessel and to this, a solution of TFA in DCM (50:50) was added. The reaction was mixed at room temperature for 20 minutes and then drained. The resin was then washed with a 2% NMM in DCM solution (2x), Methanol (1x), and DCM (2x). A positive Kaiser test indicated the presence of a primary amine and a successful deprotection.



Addition of Second Carboxylic Acid to Wells 85-96 – Wells 85-90 and 91-96 each received the same amino acid respectively (5eq) along with 5eq of HOBt and DIC. The reactions were mixed at room temperature for one hour and then the wells were drained. Each reaction was set up a second time and ran for 2 hours. Upon completion of the reactions, the wells were drained and washed with DMF (5x).



Wells 1-84

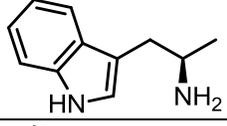
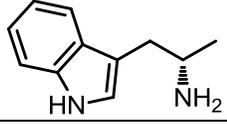
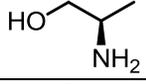
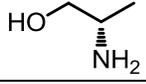
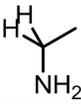
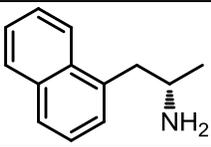


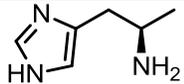
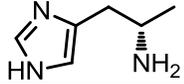
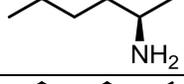
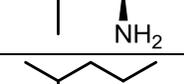
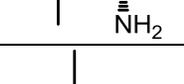
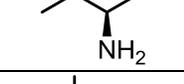
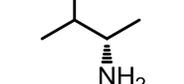
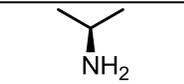
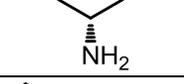
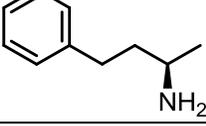
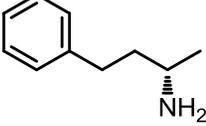
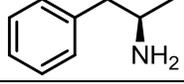
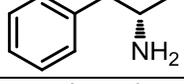
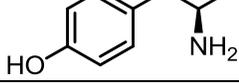
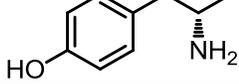
Wells 85-96

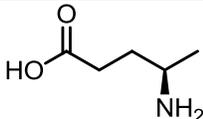
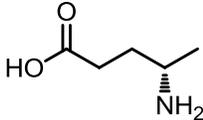
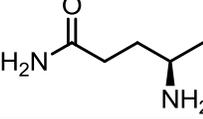
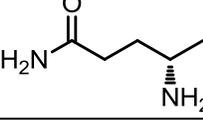
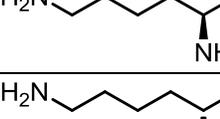
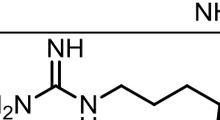
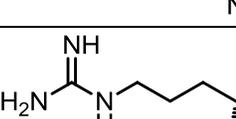
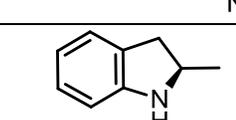
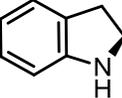
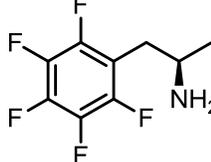
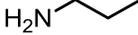
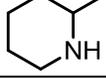
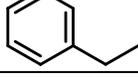
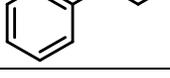
Cleavage of protected product from the resin – Each well was treated with 5 equivalents of the appropriate cleaving amine (wells 1-42 and 85-90 received 3-

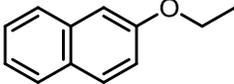
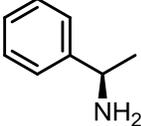
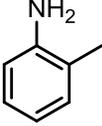
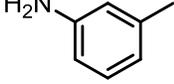
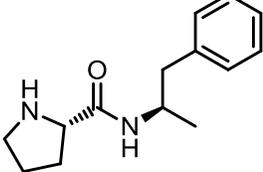
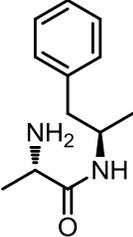
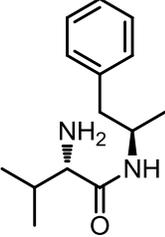
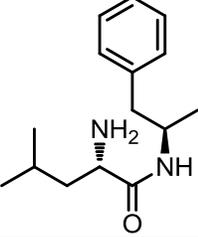
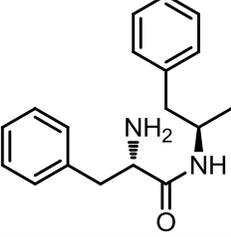
phenylpropylamine and wells 43-84 and 91-96 received n-hexylamine) in chloroform. The next day, the resin from each well was filtered through a pipette filled with silica gel using Acetonitrile. After allowing the chloroform and acetonitrile to evaporate, a solution of TFA/DCM (25/75) was added to each vial and allowed to mix over night. After letting the TFA and DCM evaporate away, each vial had distilled water added to it and freeze dried.

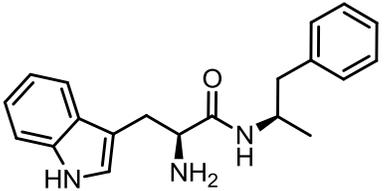
Library Monomers

R Group	Well	Cleaving amine	Well	Cleaving Amine
	1	Aromatic	43	Aliphatic
	2	Aromatic	44	Aliphatic
	3	Aromatic	45	Aliphatic
	4	Aromatic	46	Aliphatic
	5	Aromatic	47	Aliphatic
	6	Aromatic	48	Aliphatic
	7	Aromatic	49	Aliphatic
	8	Aromatic	50	Aliphatic

	9	Aromatic	51	Aliphatic
	10	Aromatic	52	Aliphatic
	11	Aromatic	53	Aliphatic
	12	Aromatic	54	Aliphatic
	13	Aromatic	55	Aliphatic
	14	Aromatic	56	Aliphatic
	15	Aromatic	57	Aliphatic
	16	Aromatic	58	Aliphatic
	17	Aromatic	59	Aliphatic
	18	Aromatic	60	Aliphatic
	19	Aromatic	61	Aliphatic
	20	Aromatic	62	Aliphatic
	21	Aromatic	63	Aliphatic
	22	Aromatic	64	Aliphatic
	23	Aromatic	65	Aliphatic
	24	Aromatic	66	Aliphatic

	25	Aromatic	67	Aliphatic
	26	Aromatic	68	Aliphatic
	27	Aromatic	69	Aliphatic
	28	Aromatic	70	Aliphatic
	29	Aromatic	71	Aliphatic
	30	Aromatic	72	Aliphatic
	31	Aromatic	73	Aliphatic
	32	Aromatic	74	Aliphatic
	33	Aromatic	75	Aliphatic
	34	Aromatic	76	Aliphatic
	35	Aromatic	77	Aliphatic
	36	Aromatic	78	Aliphatic
	37	Aromatic	79	Aliphatic
	38	Aromatic	80	Aliphatic

	39	Aromatic	81	Aliphatic
	40	Aromatic	82	Aliphatic
	41	Aromatic	83	Aliphatic
	42	Aromatic	84	Aliphatic
	85	Aromatic	91	Aliphatic
	86	Aromatic	92	Aliphatic
	87	Aromatic	93	Aliphatic
	88	Aromatic	94	Aliphatic
	89	Aromatic	95	Aliphatic

 <p>The chemical structure shows an indole ring system attached to a propan-2-amine chain. The amino group is on the second carbon, and the indole ring is attached to the first carbon. The propan-2-amine chain is further substituted with a benzyl group on the second carbon.</p> <chem>CC(N)CNC(=O)CCc1c[nH]c2ccccc12</chem>	90	Aromatic	96	Aliphatic
---	----	----------	----	-----------