8-16-2012 12:00 AM

The Development of Organometallic OBOC Peptide Libraries for use in Molecular Imaging

Dana R. Cruickshank, *The University of Western Ontario*

Supervisor: Leonard Luyt, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Chemistry

THE DEVELOPMENT OF ORGANOMETALLIC OBOC PEPTIDE LIBRARIES FOR USE IN MOLECULAR IMAGING

(Spine title: Organometallic OBOC Peptide Libraries for Molecular Imaging)

(Thesis format: Monograph)

by

Dana R. Cruickshank

Graduate Program in Chemistry

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

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

THE UNIVERSITY OF WESTERN ONTARIO SCHOOL OF GRADUATE AND POSTDOCTORAL STUDIES

CERTIFICATE OF EXAMINATION

Supervisor	Examiners
_____	_____
Dr. Leonard Luyt	Dr. Kim M. Baines
	_____
	Dr. Mike A. Kerr
	_____
	Dr. David W. Litchfield

The thesis by

Dana Roland Cruickshank

entitled:

The Development of Organometallic OBOC Peptide Libraries for use in Molecular Imaging

is accepted in partial fulfillment of the
requirements for the degree of
Master of Science

Date_____ _____

Chair of the Thesis Examination Board

Abstract

The discovery of peptide-based targeted imaging agents is hindered by the required addition of a radionuclide that can affect the binding properties of the peptide. Screening of peptide libraries is problematic because they do not take into account the presence of the imaging agent. Herein the development of an organometallic OBOC peptide library that contain a rhenium(I) tricarbonyl-based imaging entity surrogate incorporated directly onto the peptide chain, is reported. Additionally, MALDI tandem mass spectrometry is reported as a method to deconvolute organometallic peptides containing four different rhenium(I) tricarbonyl chelates on the N-terminus of octapeptides. Furthermore, an organometallic OBOC peptide library is created and screened against the GLP-1R (glucagon-like peptide 1 receptor) in order to discover potential imaging agents for diagnosis of diabetes and insulinomas. The methods reported herein can be applied to develop peptide-based imaging agents for a wide variety of relevant targets.

Keywords: Molecular Imaging, OBOC library, Rhenium(I) tricarbonyl, Peptide fragmentation, MALDI, GLP-1R

Statement of Co-Authorship

Chapter 2: All work in this chapter was completed solely by the author

Chapter 3: All work in this chapter was completed by the author with the exception of MALDI tandem mass spectrometry data acquisition. This was done by Dr. Christina Booker and Ms. Kristina Jurcic of the UWO MALDI MS Facility under the supervision of Dr. Ken Yeung, an associate professor at the University of Western Ontario jointly appointed to the Departments of Chemistry and Biochemistry.

Chapter 4: The creation of fluorescent CHO GLP-1R and CHO cells was done by the laboratory of Dr. Savita Dhanvantari, an assistant professor in Medical Biophysics at the University of Western Ontario. The work was done by Ms. Rebecca McGirr, a technician in the lab. Bead incubation was also done with their assistance.

Acknowledgements

I would like to thank everyone who supported me throughout graduate studies over the past two years.

Special thanks to Dr. Len Luyt, a great supervisor, who was always available to help, guide and advise me through my time in the lab.

I would like to thank Dr. Ken Yeung, and especially Dr. Christina Booker and Ms. Kristina Jurcic, members of his lab, for consistently providing timely access to mass spectrometry facilities as well as advice. Also, thanks to Dr. Savita Dhanvantari and Ms. Rebecca McGirr for their time and biological expertise.

Special thanks to all of the students and postdocs, past and present, in the Luyt Lab. Their aid and knowledge helped me greatly in completing my research, as well as always making it a great environment to work.

Also, thanks to my family and friends who probably had to endure too much "chemistry talk" over the course of my university studies; I couldn't have done it without you.

Table of Contents

List of Tables

List of Figures

List of Schemes

List of Abbreviations

AEEA	2-aminoethyoxy-2-ethoxyacetic acid
ANP	2-amino-3-(2-nitrophenyl)propanoic acid
Boc	*tert*-butoxycarbonyl
BPA	Bispicolylamino Acetic Acid
cAMP	cyclic Adenosine monophosphate
CID	collision-induced dissociation
$CDCl_3$	deuterated chloroform
CHCA	*trans*-α-cyano-4-hydroxycinnamic acid
CHO	Chinese hamster ovarian
DCM	dichloromethane
DIPEA	diisoproylethylamine
DMF	dimethyl formamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTPA	diethylenetriaminepentacetic acid
EDTA	ethylenediaminetetraacetic acid
ESI	electrospray ionization
Et	ethyl
EtOAc	ethyl acetate
EtOH	ethanol
FDA	U.S. Food and Drug Administration
FDG	fluorodeoxyglucose
Fmoc	fluorenylmethyloxycarbonyl
GFP	Green Fluorescent Protein
GHSR-1	Growth Hormone Secretagogue Receptor 1
GLP-1	Glucagon-Like Peptide 1
GLP-1R	Glucagon-Like Peptide 1 Receptor
GLP-2	Glucagon-Like Peptide 2
GPCR	G protein coupled receptor
HBSS	Hank's buffered saline solution
HBTU	O-Benzotriazole-N,N,N’,N’-tetramethyl-uronium-hexafluoro-phosphate
HCTU	2-(6-chloro-1H-benzotriazole-1-yl)- tetramethyl-uronium-hexafluoro-phosphate
HPLC	high performance liquid chromatography
IC_{50}	half maximal inhibitory concentration
MALDI	matrix assisted laser desorption/ionization
Me	methyl
MeCN	acetonitrile
MeOH	methanol
MIBI	methyoxyisobutyl iso cyanide
MRI	Magnetic Resonance Imaging
MS	mass spectrometry
MS/MS	tandem mass spectrometry

Nic-OSu	nicotinic acid hydroxysuccinimiydl ester
NMR	nuclear magnetic resonance
NTA	Nitrilotriacetic Acid
OBOC	one-bead, one-compound
OTf	trifluoromethanesulfonate
PADA	Picolylamino-N,N-Diacetic Acid
PBS	phosphate based buffer
PED	partial Edman degradation
PEG	polyethylene glycol
PET	Positron Emission Tomography
PHA	2-Picolylamino-N,N-2-Hydroxybenzyl Acetic Acid
QSAR	quantitative structure-activity relationship
RGD	arginine-glycine-aspartic acid
rpm	revolutions per minute
S_N2	bimolecular nucleophilic substitution
SPECT	Single Photon Emission Computed Tomography
tBu	*tert*-butyl
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TIS	triisopropylsilane
TOF	time-of-flight
TOF/TOF	tandem time-of-flight
UV	ultraviolet

Chapter 1: Introduction

1.1 Molecular Imaging

Molecular imaging allows for the non-invasive visualization of molecular pathways within the body. It can be defined as the visualization, characterization and quantification of biological processes at the cellular and sub-cellular level.[1] The use of molecular imaging has increasingly played a role in modern medicine because of the recognition of the importance of being able to detect, differentiate between, and treat diseases quickly and early in their development without using invasive procedures. As a result, the development of molecular imaging techniques is an area of intensive research.

There are a number of different molecular imaging modalities that find wide use in modern medicine such as MRI (magnetic resonance imaging), PET (positron emission tomography) and SPECT (single photon emission computed tomography). The images from these modalities are often combined with each other, or techniques such as CT scans (x-ray computed tomography) to provide information about the presence, development, progress and potential treatment for a disease.

PET and SPECT are both important molecular imaging techniques in nuclear medicine. The general definition of nuclear medicine is the injection of radioactive substances (radiopharmaceuticals) into a biological system for diagnostic or therapeutic purposes. More specifically, PET and SPECT rely on the introduction of radionuclides into the body and the subsequent detection of the radiation emitted. If it can be shown that the radiopharmaceutical preferentially moves to a specific part of the body or type of

tissue, these techniques can provide a wealth of information about development and disease.

While PET and SPECT imaging techniques both rely on radionuclides, there are a number of differences in the way in which radiation is detected. PET relies on the ability of certain radionuclides, such as carbon-11, flourine-18 and gallium-68, to undergo beta decay resulting in the emission of a positron. After a positron is emitted, it will move a very short distance until it comes into contact with an electron, which will result in an annihilation event. This produces two gamma-rays at an angle of 180^{o} from one another.[2] These gamma rays are then detected by two cameras, and the information obtained is used to create an image.

SPECT imaging detects gamma ray radiation emitted directly from radionuclides. This can occur by a number of mechanisms such as electron capture whereby a proton in the nucleus combines with an electron to form a neutron and a gamma ray. This is the process which allows the use of iodine-123 and indium-111 for imaging purposes. However, radionuclides such as technetium-99m undergo gamma emission whereby a nucleus in a high energy state relaxes over time via the emission of a gamma ray.[3] The electromagnetic radiation is then detected by a gamma camera and an image is formed.

Because of the difference in radionuclides necessary for PET and SPECT, the use of either to diagnose or treat a specific disease is dependent on the discovery of suitable radiopharmaceutical probes, as well as the presence of necessary imaging equipment and the availability of specific radionuclides.

1.2 PET and SPECT Molecular Imaging Agents

Despite the power of these two nuclear imaging techniques, the expanding use of SPECT and PET relies on the discovery and development of new radiopharmaceutical probes. There are two broad classes of imaging agents: targeted and non-targeted. Non-targeted imaging agents rely on general chemical properties such as size, charge and lipophilicity that determine the movement of the probe within the body.[3] Many examples of these types of imaging agents are currently available and widespread in the clinic, such as the PET agent ^{18}F-FDG (fluorodeoxyglucose), a glucose analogue which provides imaging based on sugar metabolism.[4] An example of a SPECT non-targeted imaging probe for use in heart imaging that has found widespread use is ^{99m}Tc-MIBI (methoxyisobutyl isocyanide).[5] The lipophilic cation proportionally distributes within the heart and allows for the effective imaging of blood flow.

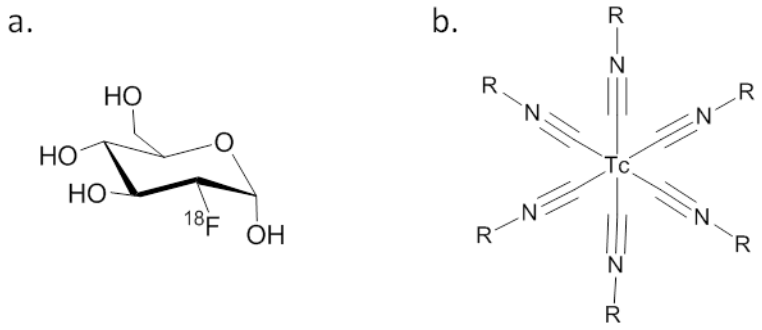


Figure 1.1 a) ^{18}F-fluorodeoxyglucose, a widely used analogue of glucose used often to image tumors due to either high uptake of sugars. b) Tc-MIBI, a non-specific imaging agent used especially for cardiac imaging.

Conversely, a targeted imaging probe usually contains two distinct parts (Figure 1.2). The first part is the imaging moiety, which for the purposes of SPECT and PET

must contain a radioactive isotope capable of emitting detectable radiation. The imaging moiety is responsible for providing radiation in order to facilitate the visualization of the imaging agent's location. The second part is a targeting moiety which is responsible for the imaging agent's distribution in a biological system, such as binding to a particular receptor or enzyme.

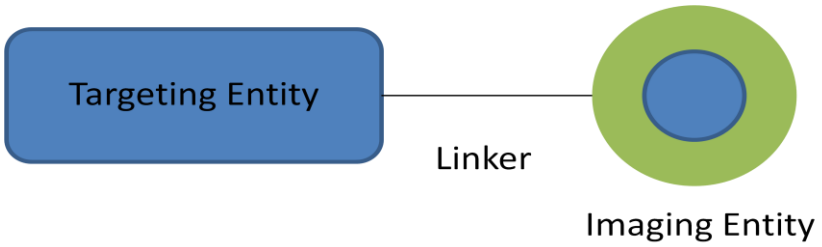


Figure 1.2 The general formulation of a targeted imaging agent, including a targeting entity, responsible for interactions with specific receptors, an imaging entity necessary for visualization and a linker sometimes included minimize interference between the two entities.

The number of useful imaging nuclei is a small fraction of the known radioactive elements due to a number of constraints necessary for medical use. Imaging radionuclides must have a short enough half-life to emit enough detectable radiation within a reasonable time frame, as it would be inefficient and result in an unsafe radiation dose. Further, the radionuclide must have a long enough half-life in order to be synthesized, incorporated into a probe and injected into an individual without significant decay. For example, the most common medical radionuclide, technetium-99m emits detectable gamma rays with a half-life of approximately six hours. Another example of a useful medical radionuclide is the positron emitting fluorine-18, which has a half life of approximately 110 minutes. The different chemistry of radionuclides also plays an

important role in the development of imaging entities. Radionuclides such as carbon-11 and fluorine-18 can be incorporated directly into the structure of an imaging agent. Conversely, when utilizing metallic radionuclides such as technetium-99m and gallium-68, a chelate must be formed in order for the metal to be incorporated into the radiopharmaceutical.

Table 1.1 Common positron emitting radionuclides for PET, showing their half-life and method of production.[6]

Isotope	Half-life (min)	Method of Synthesis
^{11}Carbon	20.3	cyclotron
^{13}Nitrogen	9.97	cyclotron
^{15}Oxygen	2.03	cyclotron
^{18}Fluorine	109.8	cyclotron
^{64}Copper	768	cyclotron
^{68}Gallium	67.7	generator
^{82}Rubidium	1.26	generator

Table 1.2 Common gamma ray emitting radionuclides for SPECT, showing their half-life and method of production.[7]

Isotope	Half-life (hours)	Method of Synthesis
^{67}Gallium	78.3	cyclotron
^{99m}Technetium	6	generator
^{111}Indium	67.9	cyclotron
^{123}Iodine	13.2	cyclotron
^{201}Thallium	73.1	generator

The second necessary part of an imaging probe is the targeting moiety. The responsibility of the targeting moiety is to move the imaging moiety to a desired part of the biological system. For example, a targeting moiety might be designed to move specifically to a particular type of cancer cell, such that when radiation is detected, it will reveal the location and size of tumors present throughout the body. There are also a

number of necessary preconditions for a targeting entity to be useful in nuclear medicine. The targeting entity must have the ability to preferentially go to a desired part of the body. The targeting entity must also be able to move quickly through the body to the area of interest, because of the time constraints imposed by using radionuclides. Additionally, the targeting moiety should not produce an immune response within the patient as well as being able to be cleared from the body after the imaging has been completed in order to minimize exposure to ionization radiation.

Because of the necessary preconditions imposed upon a useful targeting entity, the development of molecular imaging agents relies heavily on the discovery and development of new targeting entities.

1.3 Peptides as Targeting Entities

Targeting entities can be composed from a diverse array of compounds such as small molecules, antibodies and peptides. Small molecules, as a class of compounds, can be very useful targeting entities. They are usually low molecular weight, which in theory gives them access to a very large number of targets in the body, both intracellular and extracellular receptors and enzymes. In general, small molecules move quickly through the body and can be designed such that they have little trouble crossing biological barriers. Further, there is a high level of diversity associated with small molecule targeting entities which allows a wide array of modifications in order improve performance. In addition, small molecules are not readily degraded by enzymes in the body. Small molecules are also generally feasible for rapid and efficient addition of

small radionuclides such as fluorine-18 and carbon-11.[8] However, there are a number of drawbacks when investigating small molecules as targeting entities. The discovery process for small molecule targeting agents is difficult and leads to few useable imaging agents.[9] Additionally, because small molecules have low molecular weight, an acceptable imaging component rarely contains a metal radionuclide due to the overall mass and size of the chelate; this limits their use as targeted imaging agents.

Another type of targeting moiety are antibodies; y-shaped proteins which are naturally produced by the body and utilized by the immune system. They are very large (usually over 100 kDa) and bind to targets with extremely high affinity and specificity. They are also very stable to degradation in the body and relatively simple to synthesize biologically. There are a number of FDA approved antibody-based imaging agents;[10] however as a class of compounds they suffer from a number of problems. Because they are often recognized by the immune system of the target, a large number of undesired side effects are common, even in FDA approved radiopharmaceuticals.[11] Further, because of their large size, they have difficulty reaching targets across biological barriers, thereby limiting their use. Also, antibodies are not easily cleared from the body, possibly staying in the blood for weeks, resulting in poor image quality due to noise, and patients being exposed to high levels of unnecessary radiation.[12]

Peptides are an important class of targeting entities because of a number of beneficial properties. Peptides are larger than small molecules but still small in size when compared to antibodies. Peptides can possess high selectivity for specific targets as well as access to a large number of relevant receptors, enzymes and proteins based on their ability to cross biological barriers.[13] Peptides are also thought to be more useful

than small molecules in targeting enzymes or receptors which exhibit protein-protein interactions. This because some protein-protein interactions occur over a wide surface area with no obvious binding pockets.[14] These properties make peptides more promising imaging agents for these particular targets due to their ability to interact in a similar fashion. Peptides are also a very diverse class of compounds, with the potential to use a large number of natural and unnatural amino acids, alter the chain length, utilize both L and D-amino acids and are easily modified. When compared to antibodies, although peptides generally have lower affinity to targets, peptides rarely produce side effects[15] and have more favourable pharmacokinetic properties[16, 17] resulting in higher quality images. Also, peptides are more stable to the harsh chemical modification and radiolabelling conditions such as elevated temperatures or organic solvents.[18] The beneficial properties of peptide-based imaging agents can be shown by the widespread use of Octreotide, an ^{111}In-peptide imaging agent acting on the somatostatin receptor (Figure 1.3).[19] Because of these advantages, peptides as targeting entities have become important in the development of novel molecular imaging agents.

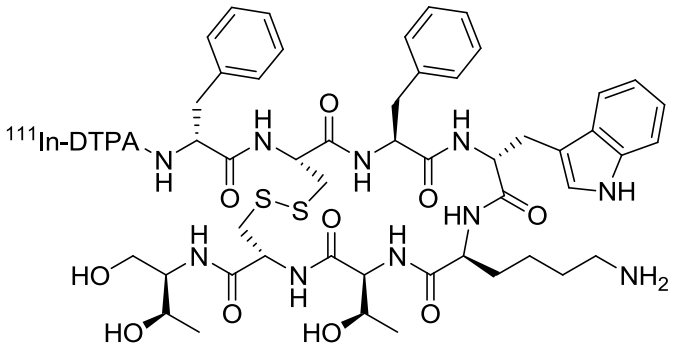


Figure 1.3 A peptide-based imaging agent containing a D-Phe and D-Trp, octreotide, acting on the somatostatin receptor. The peptide is cyclised via a disulfide bond and contains the imaging entity ^{111}In-DTPA (diethylenetriaminepentacetic acid).

1.4 Discovery and Development of Peptide-based Imaging Agents

The development of suitable peptide-based molecular imaging agents can broadly be divided into two categories (Figure 1.4). First, a rational design method can be taken. This approach relies on modifying peptide or peptide-like structures based on known information about the target. Second, a random library approach may be taken. In this method, one screens a large number of peptides (potentially millions) and determines which ligands show the most promise as imaging agents. There are a number of advantages and disadvantages associated with both techniques.

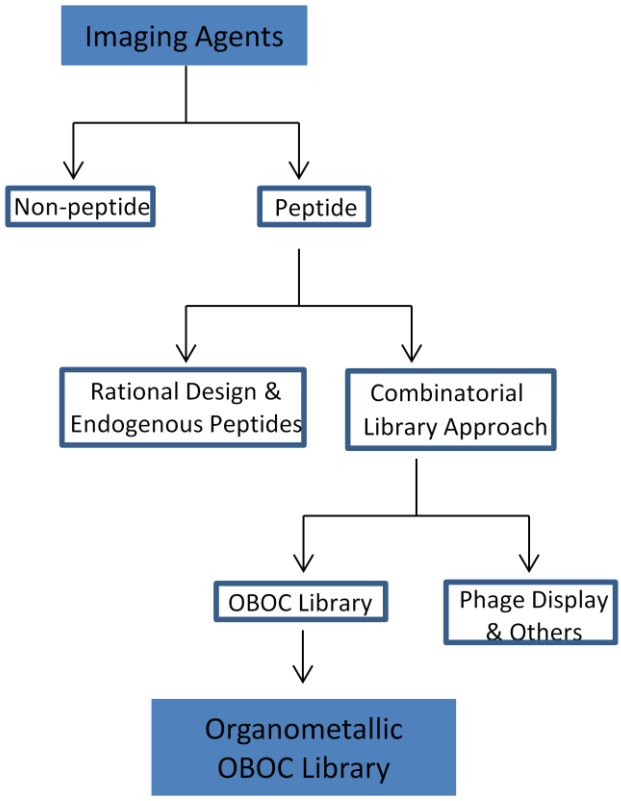


Figure 1.4 Flow chart of the available methods for developing imaging agents

Rational design has been a successful method in developing peptide-based therapeutic drugs. There are a number of different tools available in the design of peptides, such as the truncation and modification of known biological peptides. Endogenous peptides are rarely useful imaging agents due to their short half-life in the body. Rational modification of endogenous peptides can result in an increase in desired properties as well as maintenance of high selectivity and binding affinity. Amino acids can be altered, D-amino acids can be added to increase biological half-lives, unnatural amino acids can be added and lipophilicity can be changed with careful modifications. This provides a wide array of potential tools available in the rational design process.

Structural information can also be taken from the target of interest. If the desired binding site is known, it is possible to construct a peptide based on the properties of the target itself. Computational approaches to peptide-protein interaction using QSARs (quantitative structure-activity relationships) are useful, however the computational complexity of determining conformations of relatively large and flexible molecules with known protein structures has, despite steady advances in predictive ability, serious limitations.[20] High resolution crystallographic information derived from target-ligand structures remains the most useful tool in designing peptide ligands as they can provide a wealth of information about specific binding interactions and conformations of the binding site. Therefore, in utilizing a combination of the above approaches it is possible to design potential molecular imaging agents. As an example of the rational design process, the endogenous peptide hormone ghrelin(1-28), a 28 amino acid peptide with an octanoylated side chain, binds to the growth hormone secretagogue receptor 1 (GHSR-1),[21] and is a promising candidate for prostate cancer imaging. Rational modification of

this peptide chain has resulted in a significantly truncated amide peptide, ghrelin(1-5)amide with effective maintenance of binding affinity.[22] Additionally, modification of the octanoylated side chain with the radionuclides technetium-99m and fluorine-18[23] has produced promising imaging agent candidates.

However, there are a number of serious drawbacks in undertaking a rational design approach. The entire rational design method relies on the availability of information about the specific target of interest. If a rational design approach was to be undertaken to design a bioactive peptide sequence against a specific target, information about the endogenous ligand(s) and the binding site must already be known. In many cases, there is not enough data on the desired target. For example, obtaining crystal structures of protein active sites or protein-ligand interactions is a non-trivial exercise which results in few usable structures. Further, of the structures available, comparatively few protein crystal structures are of high value therapeutic or diagnostic targets, as membrane bound proteins are particularly difficult and time-consuming to crystallize.[24] All rational design processes rely on prior knowledge about the system of interest in order to develop imaging agents, and in many cases this information is not readily available. Because of this, other methods have been developed to aid in ligand discovery, in which a large precondition of target and ligand knowledge is not required.

The use of peptide libraries strives to overcome the drawbacks associated with the rational design process. In these methods, a large number of sequences are tested for binding affinity against a desired target. The ligands that have shown high affinity to the target are identified and then can be further developed as imaging agents. The advent of large scale economical automation has allowed these methods to become widely

available. There are a number of benefits to this process. First, the precondition of ligand and target information is dramatically lessened. Second, in comparison to a rational design process, a large number of peptides which are potential imaging agents can be discovered and developed simultaneously for one target. Third, the process is generally less time consuming than rational design, and potential peptides can be tested faster and more economically. Fourth, because proteins usually contain many potential binding sites, the random library approach does not distinguish between sites. As such, although subsequent tests might be necessary to determine where the peptides bind, this provides a particular advantage during imaging agent discovery because the peptide sequence is needed to locate the receptor, as opposed to induce or inhibit biological activity. Furthermore, combinatorial approaches can be used in conjunction with the rational design process to aid in the discovery of new peptide ligands.

Because of these benefits, a large variety of random libraries have been developed for the purpose of generating bioactive peptide ligands. Generally, peptide libraries involve the synthesis of random sequences of amino acids on a solid support of some kind. Random peptide libraries can further be divided in two categories, spatially and non-spatially addressable libraries. Because the creation of peptide libraries inherently results in random peptide sequences, one of the main concerns with random libraries is determining the identity of a peptide with desired properties.

In spatially addressable libraries, the identity of the peptide is known by its location on the solid support, such that if an interaction is observed, for example, at a particular spot on a peptide-coated plate, the location of the spot will allow determination of the peptide sequence. Two such representative spatially addressable libraries are

SPOT and light directed libraries. SPOT libraries were first reported by Frank in 1992[25] and utilize a cellulose solid support on which peptides are built.[26] With automation, activated amino acids can be added sequentially to specific locations such that the identity of the peptide at that location is known.[27] Light directed libraries created by Fodor *et al.* rely on a similar set of principles as SPOT libraries, by utilizing photomasks and photolabile protecting groups.[28] Lasers are used to deprotect photolabile protecting groups at specific locations and activated amino acids are added. When this process is repeated a number of times, many peptides can be built on the solid support and the identity of the peptides at each location will be known. One of the main drawbacks associated with these types of libraries is there is a limit to the number of peptide sequences which can be screened at any given time. Automation coupled with SPOT or light directed libraries can only result in tens of thousands of unique peptide sequences.[29] Although this is useful for small focused libraries in which only a few amino acids residues are altered while the others remain constant, large libraries containing perhaps millions of sequences are often required for drug or imaging agent discovery.

Non-spatially addressable random peptide libraries are defined by the inability to identify the peptide sequence based on its location. Two such examples of non-spatially addressable libraries are phage display libraries and OBOC (one-bead, one-compound) libraries. Both of these libraries require further manipulation of the library in order to identify the bioactive peptide sequence. However, it is possible to screen significantly more peptide sequences than in spatially addressable libraries, potentially numbering in the millions.

Phage display libraries are a broad classification of a type of biological library. The general process for creating phage display libraries (Figure 1.5) is as follows: phages, which are viruses infecting bacterial cells, are engineered with DNA coding for specific peptide sequences, which are then randomly mutated to create a large number of variations. After infection of *E. coli* cells with these phages, the peptide sequences are synthesized as a coat protein for the phage.[30] A phage display library has the potential to display billions of peptide sequences, of which bioactive peptides are identified based on DNA sequencing. Moreover, phage display libraries can be used to investigate relatively large polypeptides in comparison to other synthetic methodologies. Recently, Chakravarthy *et al.* utilized phage display libraries to discover a 174 amino acid polypeptide acting as an inhibitor to a protein kinase C enzyme,[31] a class of proteins that are important pharmaceutical targets.

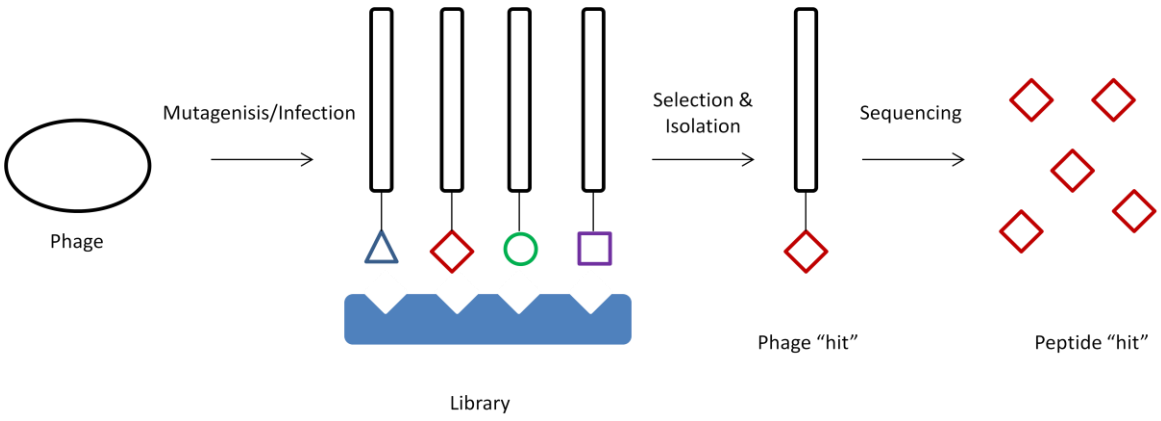


Figure 1.5 Overview of the methodology associated with utilizing phage display libraries for discovering peptide-protein interactions.

The problems associated with phage display libraries are mainly that there is a limitation to the types of peptides available to be tested. During the process of

transcription and translation of DNA into peptides, the biological machinery recognizes specific DNA sequences which correspond to specific peptide sequences. Because of this, the biological phage display library can only display peptides created from natural amino acids. Peptides containing only natural amino acids are prone to quick degradation within the body and generally need to be modified with unnatural or D-amino acids. Despite these disadvantages, phage display libraries are popular because of their unparalleled diversity, commercial availability and speed at which bioactive peptide sequences can be obtained.

OBOC libraries are another type of non-spatially addressable libraries. Unlike phage display libraries, OBOC libraries are not biological in nature. OBOC libraries involve the sequential synthesis of peptides on solid support beads. This technique utilizes standard solid-phase peptide synthesis to build amino acid chains on beads composed of an insoluble backbone. The creation of OBOC peptide libraries employs a split-mix methodology (Figure 1.6).

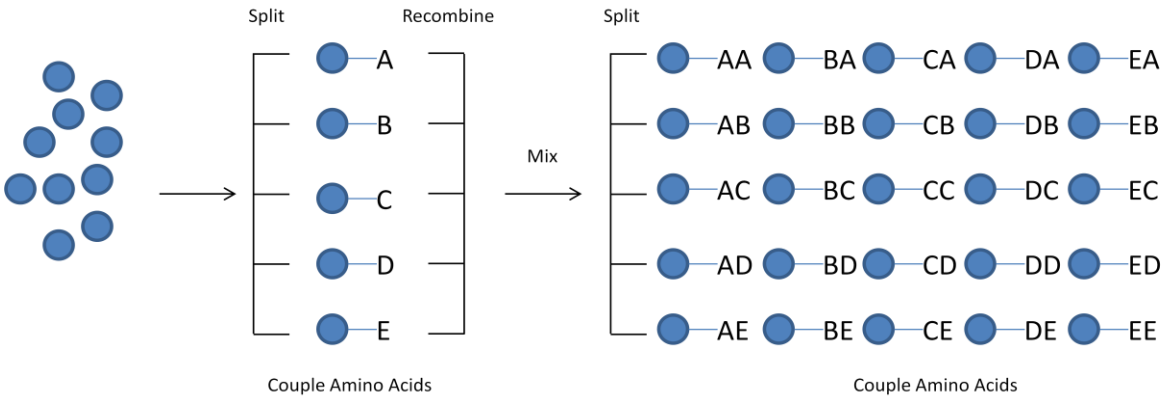


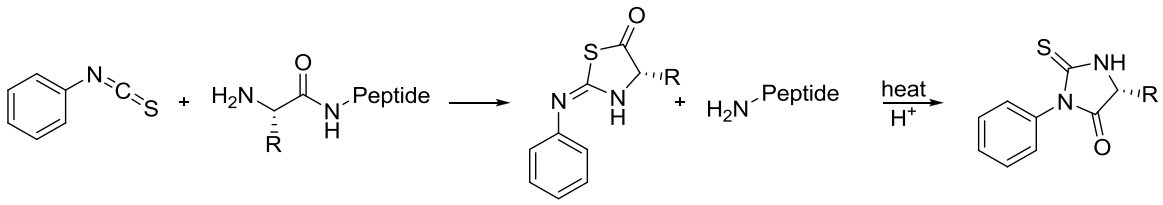
Figure 1.6 The synthesis of OBOC peptide libraries, where A-E are amino acids, using two mix-split steps. The number of possible peptide sequences increases exponentially with each step, following the formula $N = p^a$, where N is the number of sequences, p is the number of amino acids in a peptide, and a is the number of amino acids available.

A group of single beads is split into a number of discrete wells, which correspond to the number of amino acids used. The amino acids are then coupled to the beads. Subsequently, the beads from each well are then recombined and mixed together. This mix-split cycle is then repeated as many times as necessary to produce the desired length of peptide.[32] With every mix-split cycle, the number of possible sequences increases exponentially. In this way, each bead will contain many copies of unique peptide sequence. When compared to phage display libraries, OBOC peptide libraries are generally smaller however they potentially still contain millions of sequences.

However, OBOC peptides have the advantage of being able to contain unnatural amino acids, D-amino acids as well non-peptide structures. For drug or imaging agent development this is a powerful tool because, as previously mentioned, these modifications are usually necessary for a peptide to be stable in biological conditions. OBOC libraries have been used to discover compounds with high binding affinities to specific receptors; examples include cyclic peptides which bind to breast cancer cell lines,[33] cyclic peptides which bind in high affinity to the somatostatin receptor which is over-expressed in certain tumours,[34] and linear peptides which interact with RGD proteins.[35] These individual beads can then be tested against a specific target to determine which beads contain a bioactive peptide sequence. Like phage display libraries, the bioactive peptide sequences must be identified.

1.5 Deconvolution and Identification of Peptide Sequences

The two primary methods in order to identify unknown peptide sequences are Edman degradation and tandem mass spectrometry. Reported in 1950, Edman degradation involves the use of phenylisothiocyanate which reacts with a free amine present at the N-terminus of the peptide to form a cyclical derivative that contains the original terminal amino acid. This can then be removed from the N-terminus of the peptide without disrupting any other amino acid monomers to form a phenylthiohydantoin amino acid derivative.[36]



Scheme 1.1 Edman degradation utilizing phenylisothiocyanate to form a phenylthiohydantoin derivative.

Because the side chains are present in this derivative, original amino acid identity can be identified by chromatography and mass spectrometry. This method is useful as it requires a minimal amount of peptide, is capable of unambiguously identifying any natural amino acid based on their side chain and the ability to sequence peptides directly from a solid support. However, Edman degradation has a number of limitations. The process is time-consuming and prohibitively expensive when identifying a large number of peptides. Further, there is a requirement of an N-terminal amine which is not usually

desirable in imaging probe development due to a general lack of biological stability associated with peptides containing free N-terminus.

Tandem mass spectrometry relies on the ionization of the peptide, either through ESI (electrospray ionization) or MALDI (matrix-assisted laser desorption/ionization) techniques.[37, 38] A peptide ion is identified and subsequently fragmented and the resulting ions can then be used to identify the peptide sequence. This method is fast, economical and has few prerequisites in terms of the peptide's structure. For example, tandem mass spectrometry can be used to identify peptides with acetylated N-termini as well as non-peptide like motifs. Tandem mass spectrometry is also effective in the presence of impurities which are common in library screening.[39] However, in most cases prior to sequencing, peptides need to be cleaved from solid support. Further, the amino acids leucine/isoleucine and glutamine/lysine are isobaric and therefore indistinguishable via common mass spectrometry methods.[40] In addition, peptide fragmentation is largely sequence specific; individual peptides might not produce all required fragments for unambiguous sequence determination.[41] Mass spectrometry for peptide identification can be useful in screening large numbers of peptides for use in nuclear medicine because of the versatility allowed in the structure of the bioactive molecule.

There is a wide array of combinatorial tools available in order to find bioactive peptide sequences for use in discovering new targeting entities. These libraries can provide an alternative or complementary method to the rational design process.

1.6 Organometallic OBOC Peptide Libraries

Molecular imaging agents are not often developed via the combinatorial approach due to a number of difficulties associated with the technique. Regardless of the method employed, testing targeting entities for bioactivity fails to take into account the rest of the potential imaging agent. The targeting entity is only one portion of an imaging agent. In addition to the targeting entity, an imaging entity containing a radionuclide, a chelate or prosthetic group and occasionally a linker to attach the imaging entity to the targeting entity, is required. This presents the problem where after testing for bioactive peptide sequences and determining their identity, a subsequent modification is always necessary to add an imaging moiety. In the case of small peptides (<2000 Da) and small molecules, the addition of a metal chelate and possibly a linker can account for a significant portion of the overall mass of the potential imaging agent.[42] This can dramatically change the bioactivity of the targeting entity. [43, 44] A demonstration of this is shown by the reduction of the binding affinity of the 5-iodobenzovesamicol from an IC_{50} = 2 nM to an IC_{50} = 280 nM upon the addition of a technetium-99m chelate as shown in Figure 1.7.[45]

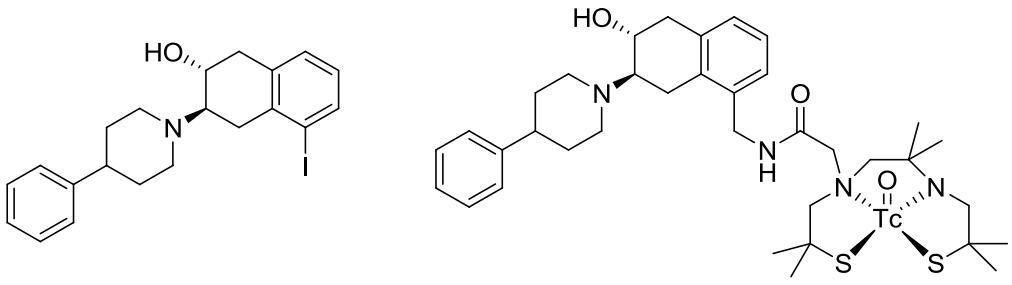


Figure 1.7 5-iodobenzovesamicol and the compound with a technetium chelate. The addition of the imaging moiety decreases binding affinity by over 100 fold.

In general, as targeting entities, either small molecules or peptides, become larger, the effect of the imaging entity on the binding affinity becomes less drastic. However, even moderately sized peptides have been shown to be influenced significantly by metal chelates. An example of the addition of rhenium(V) and technetium(V) complexes to a somatostatin receptor-binding peptide, containing nine amino acids, with an initial binding affinity of 1.8 nM results in the decrease in binding affinity by approximately eleven fold.[46] Also, the addition of a large chelation system can affect the lipophilictiy and electronic properties of the compound, as well as altering the conformations of the molecule. All of these can potentially have negative effects on the binding affinity such that the peptide sequence is no longer viable as a molecular imaging agent.

This is an inherently inefficient system because after identifying bioactive peptide sequences, the next step is to alter the properties of the compound. The addition of an imaging entity to a targeting entity is especially troublesome during high throughput library screening because knowledge of how the targeting entity binds to the target as well as its conformation in three dimensions is not usually known; this results in a general lack of understanding as to how the addition of an imaging entity might lower the initial peptide's binding affinity. Because of the problems associated with the addition of an imaging moiety, it is apparent that the development of imaging agents poses unique challenges of which new methods are required in order to facilitate their discovery.

Therefore, in order to counter these problems, the development of an organometallic OBOC peptide library is proposed. The introduction of a radionuclide surrogate into the peptide library prior to screening has the potential to overcome the problems associated with using traditional peptide libraries to discover molecular

imaging agents. This method has the potential to account for changes in binding affinity due to radionuclide incorporation; there also exists a possibility to screen for bioactive peptide sequences which only show activity when a radionuclide is present. This type of compound would not have been identified using traditional techniques.

Rhenium is used as a non-radioactive surrogate for the technetium-99m radionuclide. Technetium-99m is the most widely used radionuclide in modern nuclear medicine accounting for approximately 85% of radiopharmaceutical injections. Because non-radioactive isotopes of technetium do not exist, rhenium is often used to mimic the chemical properties of technetium. Rhenium is located directly below technetium in the periodic table, and due to the lanthanide contraction, has similar size, coordination properties and bond lengths.[47] As such, a rhenium(I) tricarbonyl containing OBOC peptide library is proposed for identification of potential technetium-99m molecular imaging agents. A non-radioactive isotope of rhenium is to be attached via a chelation system to the N-terminus of solid support-bound peptides in an OBOC library. The decision to utilize one-bead, one-compound libraries is the result of a need to create a large library for lead compound discovery in order to validate an organometallic library method. Phage display libraries are biological in nature and limited to natural amino acids, as such it is not compatible with the addition of organometallics. When synthesized, the organometallic OBOC peptide library can be tested against targets useful in disease diagnosis to develop new potential imaging agents.

The potential also exists to utilize this method on a wide variety of radionuclide imaging moieties. Because of the different size and coordination spheres of many different radionuclides, this method is not limited to the use of rhenium(I) tricarbonyl

metal cores, but can be expanded to include any radionuclide in any oxidation state. This method has the potential to be a versatile and universal way to develop molecular imaging agents. Furthermore, it would be possible to create an organometallic OBOC peptide library containing multiple different types of chelates or radionuclides, in order to investigate and account for differences such as size and charge while further increasing the diversity associated with OBOC peptide libraries.

The creation and utilization of organometallic OBOC peptide libraries would increase the efficiency in developing new peptide-based SPECT and PET imaging agents. However, there are a number of preliminary experiments and synthesis that are the first steps to create a testable organometallic OBOC peptide library.

1.7 Thesis Scope

Rhenium(I) tricarbonyl organometallic chelates are investigated as part of the development of novel organometallic OBOC peptide libraries. Rhenium is an especially attractive metal for this purpose. Firstly, the radioactive isotopes rhenium-186 and rhenium-188 show much promise in therapeutic nuclear medicine as β^- emitters of 1.07 MeV and 2.12 MeV electrons, respectively.[48] Secondly, rhenium has very similar properties to technetium, which is by far the most widely used and available diagnostic radionuclide in modern medicine. Thus, rhenium is commonly used as a non-radioactive surrogate to technetium-99m. Thirdly, the rhenium(I) tricarbonyl core is unique in that it requires comparatively small chelators. The use of large chelators might not be expected to produce many positive results in conjunction with relatively small targeting peptides due to the large affects it can exert. Fourthly, rhenium(I) tricarbonyl chelates can easily

hold a variety of charges depending on the chelation system, further allowing the investigation into imaging entities' effects on peptide-target affinity.

With this information, the initial research into organometallic OBOC peptide libraries involves the synthesis and characterization of rhenium(I) tricarbonyl chelates holding different charges and with the ability to be covalently bound to a peptide targeting entity.

MALDI tandem mass spectrometry will be evaluated as a method to identify and sequence peptides containing rhenium(I) tricarbonyl chelates. Based on the expense, presence of a chelate on the peptide N-terminus and the time-consuming nature of Edman degradation, developing a method of tandem mass spectrometry to sequence organometallic peptides is the most agreeable, efficient and unconstrained way to create a method for identifying organometallic peptides.

Ultimately, with rhenium-containing peptides and a method in order to identify unknown organometallic peptides, an OBOC library will be tested against a useful target in order to validate the method. A target for such an exercise is the GLP-1R (glucagon-like peptide 1 receptor). Finding a series of imaging agents selectively targeting this receptor would potentially allow for monitoring of type 2 diabetes as well as a way to diagnose and treat various insulinomas.

Chapter 2: Synthesis of Rhenium(I) Tricarbonyl Chelates as Imaging Entity Surrogates

2.1 Introduction

The first step in the creation of organometallic OBOC (one-bead, one-compound) peptide libraries is the development of the metal chelates that can be incorporated into peptide sequences. These metal chelates are to act as imaging agent surrogates. The surrogates will imitate the size, lipophilicity and charge of an imaging entity without emitting radiation. There are a number of requirements for these metal chelates that are imposed by their use as potential imaging agents. All chelates for targeted molecular imaging need to be bifunctional; they need to form complexes with metal radionuclides, as well as have a functional group that is able to be attached to the targeting entity. Further, the chelate needs to be inert in order to withstand the conditions of not only library screening but also be stable to *in vivo* conditions for extended periods of time. Despite these restrictions, there are a wide variety of useful imaging-related metal chelates which have been investigated.

Technetium-99m is the most popular medical isotope in nuclear medicine. It has a half-life that is convenient for imaging, as well as the gamma ray emission energy is approximately 140 keV which is ideal for SPECT.[49] Further, the technetium-99m radioisotope is readily available from generators, allowing for widespread use away from the cyclotron or nuclear reactor facilities used to create radioactive isotopes. Also, SPECT imaging facilities are much more widespread than the corresponding PET facilities, despite the superior properties of positron emission tomography. Because of these reasons, it is desirable to develop ^{99m}Tc-based imaging agents. Unfortunately,

technetium is an entirely synthetic element and does not possess non-radioactive isotopes. Therefore, rhenium is often used as a non-radioactive surrogate for development, characterization and testing of imaging agents. Rhenium and technetium have very similar coordination chemistry, bond lengths as well as having very similar size due to the lanthanide contraction. The atomic radius of rhenium is 137 pm and the corresponding radius of technetium is 136 pm.[50]

2.1.2 Development of Rhenium/Technetium(I) Tricarbonyl Cores

Rhenium tricarbonyl has increasingly become the subject of imaging-related research. The rhenium tricarbonyl core was first developed in 1994,[51] in order to further investigate the +1 oxidation state of rhenium and technetium. The rhenium(I) tricarbonyl core incorporates a reduction of size as well as an increased *in vivo* stability in comparison to complexes with other oxidation states of rhenium or technetium. However, discovery of useful $^{99m}Tc(CO)_3^+$ imaging agents has been hindered by the development of suitable methods of synthesis in a time frame imposed by imaging. When ^{99m}Tc is eluted from a generator at an imaging facility, it exists as pertechnate, $[^{99m}TcO_4]^-$, and needs to be reduced into the +1 oxidation state. The first methods utilized carbon monoxide gas under pressure and in the presence of a reducing agent such as $NaBH_4$.[52] These methods are not suitable for synthesis in the radiopharmacy as they are complicated, time-consuming, use pressurized gas and are unfamiliar. These problems were overcome with the development of fully aqueous preparation of $^{99m}Tc(CO)_3^+$ utilizing potassium boranocarbonate, $K_2BH_3CO_2$, which is a source of

carbonyl ligands as well as acting as a reductant.[53] This kit formulation allowed for a facile way to create radiolabelled compounds by simply mixing the pertechnate with a premade formulation in order to create technetium(I) tricarbonyl. This method does not require pressurized carbon monoxide and has increased the interest developing clinically relevant technetium(I) and rhenium(I) tricarbonyl molecular imaging agents.

2.1.3 The Properties of Rhenium(I) Tricarbonyl Complexes

The chemical stability of rhenium(I) tricarbonyl has been extensively investigated. Rhenium(I) tricarbonyl is a low spin d^6 metal that forms octahedral complexes and the metal is considered to be a soft Lewis acid. The metal centre is polarisable and prefers softer Lewis bases. However, the presence of the three carbonyl ligands, facially coordinated, significantly adds to the versatility of ligands available to the rhenium(I) oxidation state. Carbonyl ligands donate electron density through σ bonds to the metal core, thereby increasing the stability of electron deficient species. Further, the carbonyl ligands also can accept electron density from an electron abundant species through the process of π back bonding. Carbonyls are strong field ligands because they increase the magnitude of the Δ_{oct}. This is due to the π* anti-bonding orbitals of the carbonyl ligands with t_{2g} symmetry; the energy of the d^6 electrons is lowered and therefore produces a more stable complex (Figure 2.1).

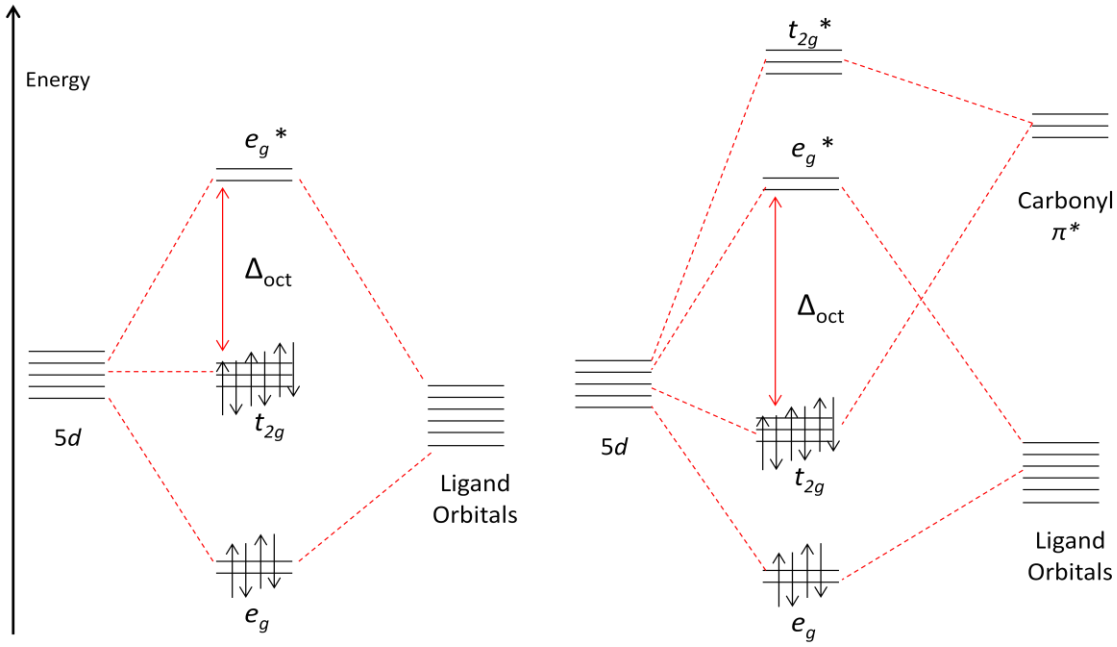


Figure 2.1 Partial molecular orbital diagrams for octahedral complexes of rhenium(I). The first diagram shows the splitting associated with ligands which do not accept or donate π electrons. The second diagram illustrates the increase in Δ_{oct} associated with having ligands which accept electron density, as is the case with carbonyl ligands. This occurs because the symmetry of the π* orbitals and three of the *d* orbitals of the rhenium are identical.

This creates molecular orbitals whereby the rhenium(I) tricarbonyl core can accept a wide variety of "harder" ligands then would be expected, and allows for a wide variety of chelation systems. The rhenium(I) tricarbonyl core has shown remarkable diversity in the ligands it can accept, forming complexes with amines,[54] aromatic amines (pyridine,[55] imidazole,[56] pyrazole[57]), phosphines,[58] thiols,[59] thioethers,[60] alcohols[61] and carboxylates,[62] among others.

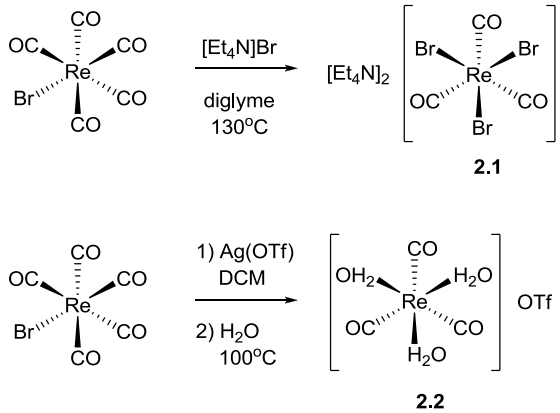
The non-reactivity of rhenium(I) tricarbonyl complexes in general is also due to their high kinetic inertness. The rhenium(I) tricarbonyl chelates are fully saturated complexes and the metal core is generally protected from attack by other ligands.

Investigations into the properties of mono-, bi- and tridentate ligands for rhenium(I) tricarbonyl cores have been extensive. Stability and *in vivo* studies have shown the tridentate chelators are the most useful for use in molecular imaging.[63] Most tridentate chelates have been shown to be stable to the atmosphere for long periods of time. Further, *in vivo* biodistribution studies of tridentate rhenium(I) tricarbonyl show complete clearance from the body 24 hours post-injection. This is a result of minimal substitution between the rhenium(I) tricarbonyl chelates and biomolecules in the body due to high thermodynamic stability and kinetic inertness of the coordination compound. Conversely, bidentate rhenium(I) tricarbonyl chelates show increased retention in the kidneys and liver post-injection. Despite very similar lipophilicity and size, the bidentate chelates act very differently. This is most likely due to the exchange of ligands in bidentate systems with water and proteins. As such, the *in vivo* kinetic stability of bidentate complexes is significantly compromised, leading to the exclusive use of rhenium(I) tricarbonyl tridentate chelators for imaging purposes.

2.2 Results and Discussion

Four rhenium(I) tricarbonyl bifunctional tridentate chelates have been synthesized for use in an organometallic OBOC peptide library. A selection of chelates, neutral, positively and negatively charged, were synthesized in order to obtain a complete assessment of the usefulness of rhenium(I) tricarbonyl chelates in OBOC peptide libraries, as there exist imaging agents that contain all of these charge states.

Two rhenium(I) tricarbonyl starting materials, $[Et_4N]_2[ReBr_3(CO)_3]$ and $[Re(H_2O)_3(CO)_3]OTf$, were initially synthesized from rhenium(I) bromopentacarbonyl (Scheme 2.1) These complexes contain bromo- or aqua- ligands which are labile and easily exchanged due to the trans-effect of the carbonyl ligands. The carbonyl ligands are able to accept electron density from the attack of a ligand opposite the position of the carbonyl, thereby stabilizing transition states.[64] The tribromo- complex is mildly soluble in organic solvents such as methanol, acetonitrile and DMF which allows for incorporation of the rhenium(I) tricarbonyl core into hydrophobic chelates. Conversely, the triaqua core is water-soluble and allows for facile incorporation water soluble chelators.

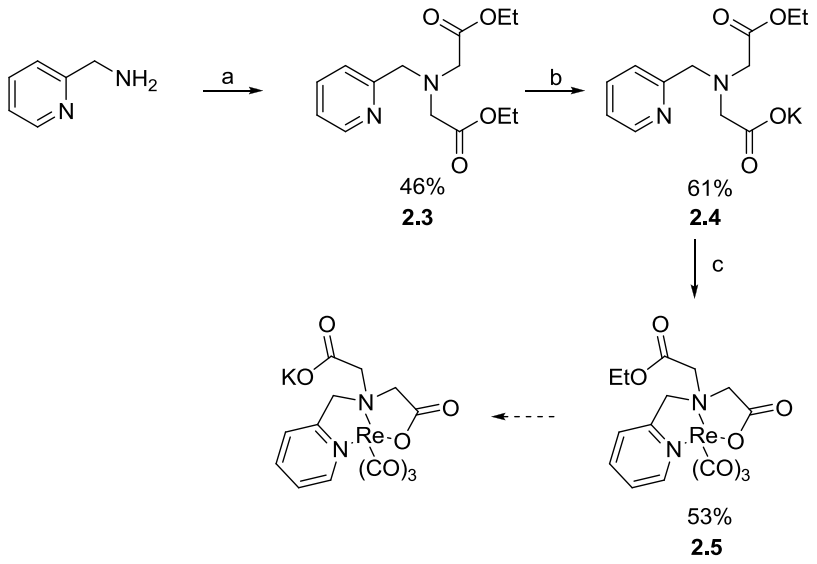


Scheme 2.1 The synthesis of rhenium(I) tricarbonyl starting materials. Compound **2.1** contains three labile bromide ligands which can be easily substituted as well as two tetraethylammonium counterions. Comparatively, **2.2** contains three labile aqua ligands and exists in an aqueous solution with a triflate counterion.

2.2.1 2-Picolylamino-N,N-Diacetic Acid (N,N,O) Rhenium(I) Tricarbonyl Chelate

The chelation system initially investigated was 2-picolylamino-N,N-diacetic acid (PADA). This is because it forms neutral complexes with rhenium(I) tricarbonyl compounds as well as provides three different unique electron donating species; an amine, a carboxylic acid and a pyridyl nitrogen. The $Re(CO)_3$-PADA complex has been previously shown to be kinetically stable, enduring acid reflux conditions for 24 hours.[52] The molecule is bifunctional as it contains functional groups that form a tridentate chelation system, as well as a free carboxylate that is used to attach the chelation system to a peptide sequence. Therefore, it is a good system to test the suitability of rhenium tricarbonyl cores in organometallic OBOC libraries.

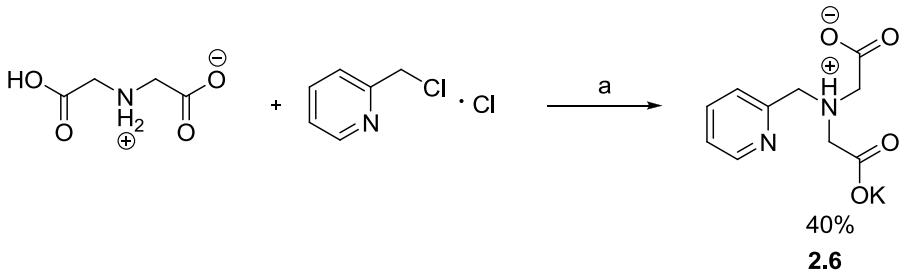
Initially, the PADA complex was synthesized via the use of ethyl ester protecting groups as shown below:



Scheme 2.2 Synthesis of $Re(CO)_3$-PADA. Reagents: a) Ethylbromoacetate, NEt_3, DMF, 0^oC. b) KOH, EtOH, 0^oC. c) $[Et_4N]_2[ReBr_3(CO)_3]$, MeOH

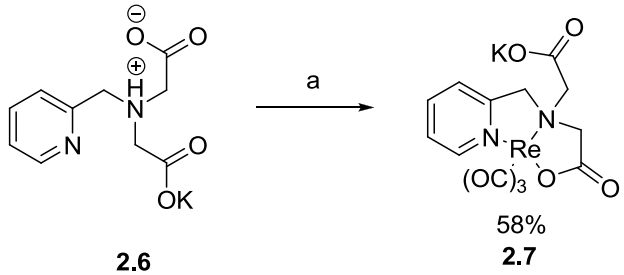
First, using the nucleophilic properties of primary amines, in an S_N2 reaction, **2.3** is formed from picolylamine and ethyl bromoacetate. A single ethyl ester group was saponified to form compound **2.4** utilizing potassium hydroxide in ethanol. The reaction involves a selective deprotection of one ethyl ester to form a carboxylate. The difficulty in this reaction is finding the correct conditions to minimize the instances of both ethyl esters of **2.3** undergoing saponification. The purpose of maintaining an ethyl protecting group on a single carboxylate is to stop potential coordination of the rhenium tricarbonyl core with two carboxylates, which would leave an undesired by-product. Although conditions were obtained where minimal saponification of both ethyl esters was observed, a significant portion of starting material was recovered. Subsequently, the coordination of rhenium through the use of $[Et_4N]_2[Re(CO)_3Br_3]$ resulted in the isolation of **2.5** by silica gel chromatography. The subsequent saponification of the remaining ethyl ester in compound **2.5** was unsuccessful due to decomposition of the coordination compound. It was decided that because of the difficulty of the selective saponification of one ethyl ester as well as the problems with deprotecting the ethyl ester from the rhenium coordinated species, a new synthetic method would be applied involving less synthetic steps.

In the second synthetic pathway, the number of synthetic steps required is reduced significantly. The synthesis of **2.6** involves the amine nucleophile of iminodiacetic acid, and picolyl chloride hydrochloride. With the addition of potassium hydroxide, **2.6** was synthesized in approximately 40 % yield after recrystallization from hot ethanol.



Scheme 2.3 Synthesis of PADA (**2.6**). Reagents: a) 1. KOH, MeOH, reflux 24 h. 2. $HCl_{(aq)}$

Rhenium is then coordinated to form the desired bifunctional PADA-rhenium chelate. The synthesis of **2.7** by this method significantly reduces the number of steps required to make large quantities of the desired chelation system. The pyridyl nitrogen is expected to coordinate preferentially to the rhenium(I) tricarbonyl core, resulting in the desired tridentate chelate. Although rhenium coordination to both carboxylate groups was observed, the desired neutral coordination compound was easily isolated. In methanol, the neutral coordination compound has low solubility and therefore precipitates out of solution, whereas the charged coordination compound does not. The product was then purified by recrystallization from hot methanol resulting in a 58% yield.



Scheme 2.4 Synthesis of $Re(CO)_3$-PADA (**2.7**). Reagents: a) MeOH, **2.1**

The product was characterized by ^{1}H NMR, ^{13}C NMR and HRMS. The ^{1}H NMR shows a clear downfield shift of the aromatic proton next to the pyridyl nitrogen due to the coordination of the Re(I) metal centre. Also, the most obvious coordination induced change in the ^{1}H NMR spectrum is the shift downfield and the splitting associated with all of the methylene protons. Compound **2.6** shows two singlets corresponding to the two chemically inequivalent methylene proton classes. However, due to the creation of a chiral centre at the nitrogen amine upon the coordination of the metal centre through the carboxylate and pyridyl nitrogen, three AB spin systems are observed (Figure 2.2).

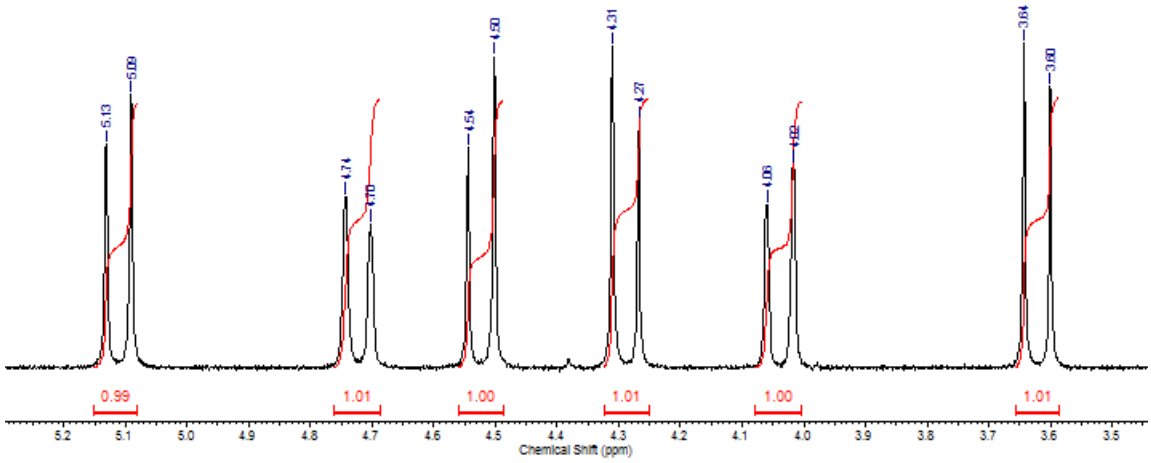


Figure 2.2 ^{1}H NMR spectrum showing the AB spin systems associated with the formation of a stereogenic centre upon $Re(CO)_3^+$ coordination of the PADA ligand, compound **2.7**. Each AB spin system represents two methylene protons, each in a different chemical environment.

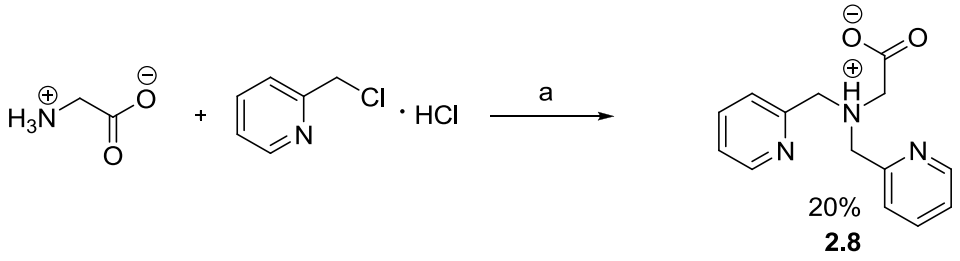
The ^{13}C NMR spectrum shows a downfield shift of the coordinated carboxylate carbon atom from approximately 171.3 ppm in MeOH-d_4 to 178.6 ppm in dmso-d_6 while the non-coordinated carboxylate is observed at 169.7 ppm, indicating two unique carboxylate environments within **2.7**. This can be rationalized as the loss of electron

density on the carbonyl group due to the coordination of a positively charged rhenium(I) tricarbonyl metal core. This provides strong evidence of the formation of the desired complex.

From this synthetic method, it is possible to acquire large quantities of bifunctional PADA chelator in order to be used in sequence deconvolution studies and peptide library construction.

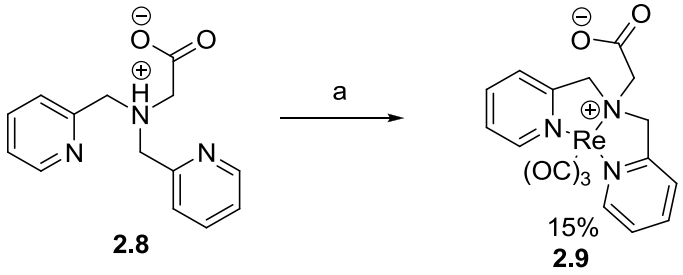
2.2.2 Bispicolylamino Acetic Acid (N,N,N) Rhenium(I) Tricarbonyl Chelate

A chelate holding an overall positive charge was also synthesized in order to investigate a wide variety of the charge states of rhenium(I) tricarbonyl complexes. Bispicolylamino Acetic Acid (BPA) has two pyridyl donating nitrogen atoms and an amine donating group. Like PADA, a free carboxylate is used as a functional group to attach the compound to a peptide sequence. The synthesis of BPA follows a similar methodology to the PADA ligand as shown in Scheme 2.5.



Scheme 2.5 Synthesis of BPA (**2.8**). Reagents: a) 1. NaOH, MeOH, reflux 24 h. 2. $HCl_{(aq)}$

Glycine and picolylchloride hydrochloride, via an S_N2 mechanism, form compound **2.8**. The resulting ligand is obtained via trituration with cold benzene. The product was difficult to crystallize due to high solubility in a variety of solvents from water to THF. Compound **2.8**, under all tested conditions, produced impure red oil. The product is subsequently coordinated with rhenium(I) tricarbonyl utilizing $[Et_4N]_2[ReBr_3(CO)_3]$ in methanol to produce **2.9**.



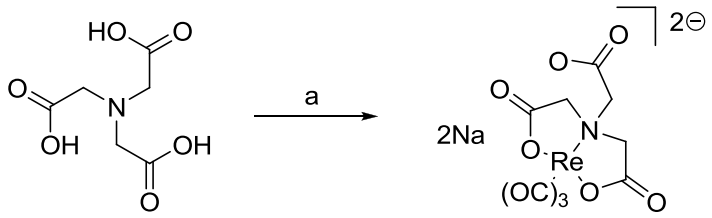
2.8 **2.9**

Scheme 2.6 Synthesis of $Re(CO)_3$-BPA (**2.9**). Reagents: a) **2.1**, MeOH, reflux.

The coordination of BPA with rhenium(I) tricarbonyl results in downfield shifts of the pyridyl methylene protons as well as splitting due to the loss of chemical equivalence in the ^{1}H NMR spectrum. Further, the free carboxylate methylene remains a singlet due to the formation of an achiral molecule. The pyridyl methylene protons produce one AB spin system because the protons above and below the plane of the macrocycle are not equivalent due their inability to freely rotate. The ^{1}H NMR, ^{13}C NMR spectra and high resolution mass spectrometry provide strong evidence for the desired coordination compound.

2.2.3 Nitrilotriacetic Acid (N,O,O) Rhenium(I) Tricarbonyl Chelate

The nitrilotriacetic acid (NTA)-rhenium(I) tricarbonyl complex holds an overall negative charge. The chelate allows for the investigation of negatively charged rhenium(I) tricarbonyl compounds for use in organometallic OBOC peptide libraries. The chelator contains two carboxylates and an amine nitrogen as potential ligands.



37%

2.10

Scheme 2.7 Synthesis of $Re(CO)_3$-NTA (**2.10**). Reagents a) $NaOH_{(aq)}$, **2.2**

The donation of two carboxylate oxygen atoms to the soft rhenium(I) metal centre is expected to be thermodynamically disfavourable due to the non-polarisable nature of the ligands. However, in the absence of any other more favourable donating ligands, the complex is formed in 37% yield from a reaction of nitrilotriacetic acid, $[Re(H_2O)_3(CO)_3]OTf$ and sodium hydroxide in water. Initially, the synthesis was attempted utilizing $[Et_4N]_2[ReBr_3(CO)_3]$. Although **2.10** was observed by ^1H NMR spectroscopy, there was a mixture of sodium and tetraethylammonium counterions present. The ^1H and ^{13}C NMR spectra show unambiguous evidence of coordination. High resolution mass spectrometry also shows the expected mass and isotopic signature.

2.2.4 2-Picolylamino-N,N-2-Hydroxybenzyl Acetic Acid (N,N,O) Rhenium(I) Tricarbonyl Chelate

The development of a tridentate rhenium(I) tricarbonyl chelate utilizing a phenol ligand was desirable due to the ligand's ability to hold a negative charge and therefore form a neutral complex in conjunction with an amine and pyridyl nitrogen as donating ligands. Phenol based chelators have not been utilized extensively in rhenium(I) tricarbonyl chemistry. The softer Lewis basicity of the phenol ligand, in comparison to carboxylates, would be expected to form thermodynamically stable complexes with the metal core, as well as a very lipophilic outer core due to the phenol aromatic systems involved, which is often desirable in imaging. The lipophilicity allows access to hydrophobic binding pockets that are common in enzymes and receptors, allowing for the rhenium chelate to be a necessary part of the binding affinity of the imaging probe.[23]

Recently, phenol, amine and aromatic nitrogen bifunctional systems have been investigated by Lim *et al.*, utilizing pendant alkene, cyano and amine functionalities as available methods to attach the chelate to a biomolecule with mixed results.[65] While complexes were successfully formed utilizing the alkene and cyano functionalities, most likely due to their poor coordination properties, the amine has been shown to preferentially coordinate over the phenol forming a positively charged complex.

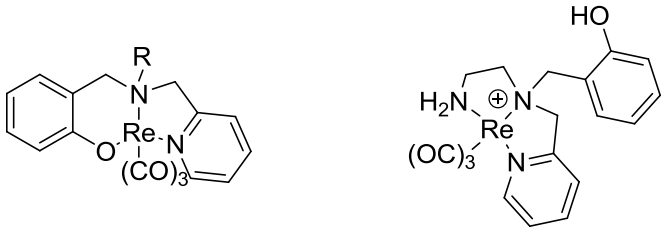
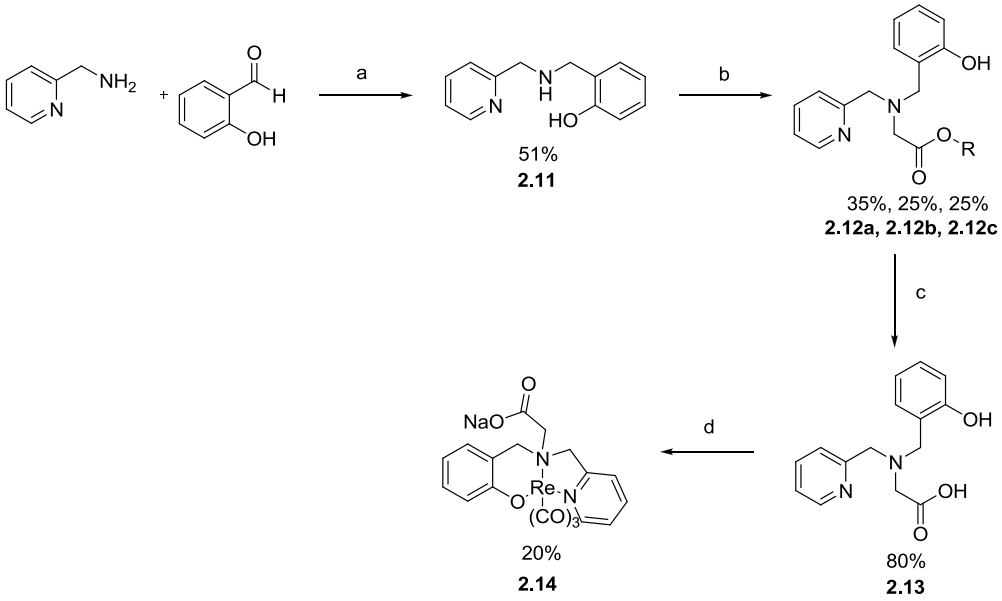


Figure 2.3 Phenol based (N,N,O) coordination of rhenium(I) tricarbonyl. R = H, -$CH_2CH{=}CH_2$, -$(CH_2)_2CN$. In these complexes, amines are shown to preferentially donate electrons to the rhenium(I) tricarbonyl in comparison to the phenol ligand.

However, utilizing a free carboxylate in a bifunctional chelator has not been investigated previously. One would suspect that a phenol as a Lewis base would be a more favourable ligand in comparison to a carboxylate due to the difference in thermodynamic stability. Thus, a synthetic method was developed in order to test this hypothesis.



Scheme 2.8 Synthesis of PHA (**2.13**) and $Re(CO)_3$-PHA(**2.14**). Reagents: a) 1. MeOH 2. $NaBH_4$. b) *tert*-butyl, benzyl, ethyl bromoacetate or benzyl bromoacetate, NEt_3, MeCN to form **2.12a**, **2.12b** or **2.12c**, respectively c) **2.12a**, TFA, DCM. d) **2.2**, H_2O, NaOH.

A reductive amination reaction utilizing salicaldehyde and 2-picolylamine creates **2.11** through a yellow imine intermediate. This compound is then reduced with the addition of three equivalents of sodium borohydride. Compound **2.11** was purified by column chromatography. The basic secondary amine in the presence of the acidic phenol results in a poor yield as the compound does not move cleanly on silica. Isolation and purification of the imine intermediate followed by reduction did not improve yields. A reaction with *tert*-butyl bromoacetate, benzyl bromoacetate or ethyl bromoacetate creates **2.12a**, **2.12b**, or **2.12c**. The low yields of the reactions can be explained by the presence of a nucleophilic oxygen atom of the phenol functionality. The *tert*-butyl ester of compound **2.12a** is then removed utilizing TFA/DCM to form compound **2.13** in 80% yield. Of note is the presence of the phenol peak in the ^{1}H NMR at a downfield position of 10-11 ppm for all of the previous compounds. A representative spectra is provided of compound **2.12a** in Figure 2.4.

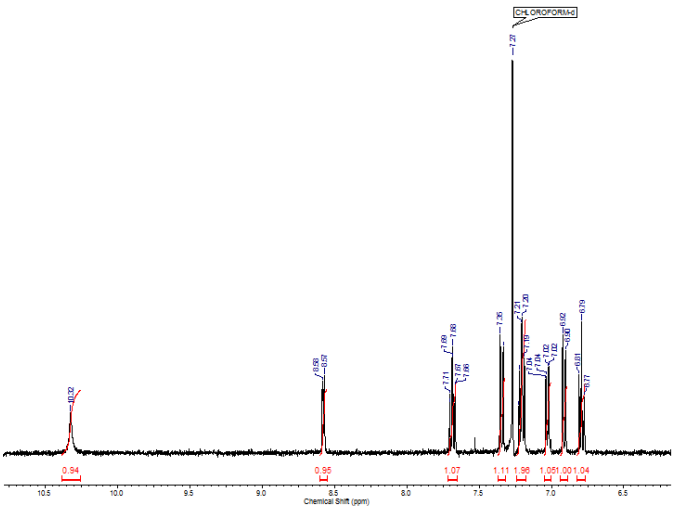


Figure 2.4 ^{1}H NMR spectrum of **2.12a** showing the region from approximately 6 ppm to 11 ppm in $CDCl_3$. The appropriate aromatic peaks are visible along with a peak at 10.3 ppm corresponding to the phenol proton.

The downfield shift of the phenol can be attributed to the unique chemical environment of the phenol proton as it may interact with the tertiary amine, producing downfield chemical shift for the phenolic hydrogen. This compound is then coordinated with rhenium(I) tricarbonyl in methanol. The expectation is that this results in product **2.14** (Scheme 2.8).

Compound **2.14** has been characterized by ^{1}H NMR, ^{13}C NMR and HRMS. However, this has led to inconclusive results about the nature of the coordination sphere. The spectra are unambiguous in that a tridentate complex has been formed due to the splitting of the methylene protons similarly observed in other chelates. However, evidence for coordination of the phenol or the carboxylate group to the metal centre is inconclusive. In figure 2.5, a comparison between the ^{1}H NMR spectra of compounds **2.13** and **2.14** shows an obvious downfield shift of all aromatic protons upon coordination. This provides evidence of tridentate coordination utilizing the phenol.

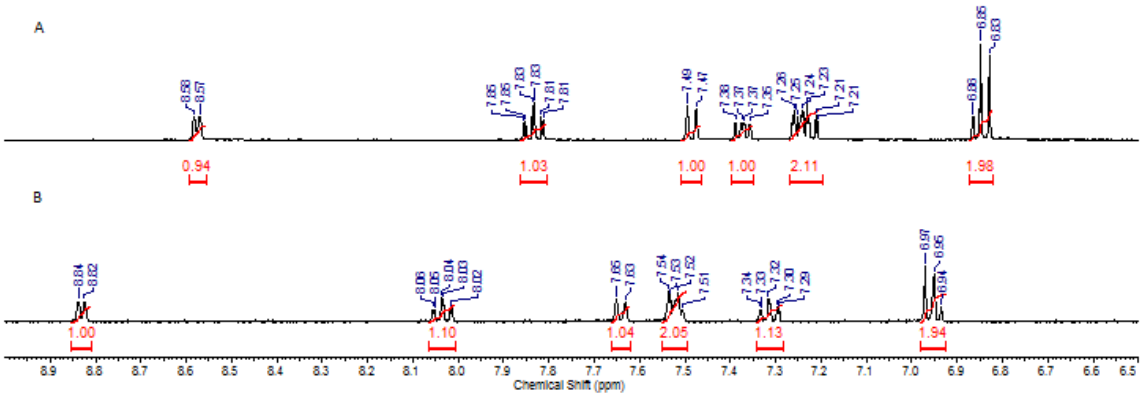


Figure 2.5 A comparison of the ^{1}H NMR spectra obtained from A) un-coordinated compound **2.13**, and B) coordinated compound **2.14**, in deuterated methanol. All of the aromatic peaks shift downfield as a result of the addition of the rhenium(I) tricarbonyl.

In addition, the coordinated complex **2.14** has been observed to produce a negative result during the qualitative ferric chloride test for phenol; a positive test results in a purple solution if a phenol group is present. These results lead to the conclusion that the phenol is a donating ligand in the coordination sphere. However, there is also a broad peak at 10-11 ppm in DMSO-d_6 which could correspond to an extremely downfield phenol peak or an upfield carboxylic acid peak. A downfield phenol peak, although easily explained in uncoordinated compounds due to the presence of a tertiary amine in close proximity, would not be easy to explain due to the absence of electrons on a nitrogen which is acting as a coordinating ligand. This further adds to the ambiguity of the ^{1}H NMR spectra obtained from this compound. Yet, the ^{13}C NMR spectrum provides evidence that shows a carboxylate shift of 178.5 in DMSO-d_6 which is consistent with coordination of the carboxylate ligand.

Because of the conflicting evidence, a protecting group strategy was utilized to attempt to unambiguously form a phenol-donating complex. **2.12a** was coordinated with rhenium(I) tricarbonyl in methanol in order to avoid the presence of a free carboxylate. Initially, this route seemed to produce the desired result, however upon further investigation by HPLC-ESI, the presence of masses associated with coordinated compounds containing methyl esters and free carboxylates and no compound containing the *tert*-butyl ester. The mechanism of this transformation is unknown. This reaction was also repeated using a benzyl protected carboxylic acid, **2.12b**, and the ethyl protected carboxylic acid, **2.12c**, to the same result. Verification of the formation of a methyl ester was obtained by ^{1}H NMR spectroscopy. The reaction was also repeated in DMF, in an

attempt to avoid the trans-esterification. However, the resulting reaction mixture contained a major portion of compound **2.14** by HPLC-MS.

Because of these results, it was decided to first attach compound **2.13** to a peptide chain, and coordinate rhenium(I) tricarbonyl after attachment to a peptide. This process will be discussed in the next chapter.

2.3 Conclusions

This chapter has outlined the successful synthesis and characterization of three rhenium(I) tricarbonyl chelates to be utilized in the investigation of organometallic OBOC peptide libraries. These chelates represent neutral, positive and negatively charged compounds that provide a representation of the entire spectrum of rhenium(I) tricarbonyl chelates. A fourth phenol-based chelate has also been investigated. However, the coordination of the phenol to the metal core has not been conclusively determined. This chelator (**2.13**) can be used, however, to investigate on resin coordination techniques for OBOC library creation.

2.4 Experimental

2.4.1 General Procedures

All chemicals and solvents were purchased from commercially available sources and used without further purification. Chemicals were purchased from Sigma Aldrich, Alfa Aesar or Strem. Solvents were purchased from Fisher Scientific. Silica gel

chromatography was performed using SiliFlash F60 (230-400 mesh) purchased from Silicycle. Flash silica gel chromatography was performed using a Biotage Isolera Prime Flash Purification System. Analytical HPLC was preformed utilizing a Waters Sunfire C18 column 4.6 x 150 mm, 5 µm. The gradient system utilized consists of acetonitrile + 0.1% TFA and water + 0.1% TFA. Ultraviolet absorbance measurements were obtained utilizing a Waters 2998 Photodiode Array Detector and mass spectra were obtained via electrospray ionization utilizing a Micromass Quatro Micro LCT mass spectrometer. NMR data were obtained on a Varian Inova 400 MHz spectrometer or a Varian Mercury VX 400 MHz spectrometer. Deuturated solvents were purchased from Sigma Aldrich. Chemical shifts were referenced in respect to residual solvent shifts. High resolution mass spectrometry, electron impact and electrospray, was done utilizing a Finnigan MAT 8200 or a PE-Sciex API 365, respectively.

2.4.2 Small Molecule Synthesis

$[Et_4N]_2[Re(CO)_3Br_3]$, **2.1**, was prepared according to a previously published literature procedure.[66] $[Re(H_2O)_3(CO)_3]OTf$, **2.2**, was prepared according to a previously published literature procedure.[67]

*2-Picolylamino-N,N-diethylacetate (**2.3**)*

To a solution of picolylamine (0.78 mL, 7.6 mmol) in DMF (9.8 mL) at 0 ^{o}C, triethylamine (2.48 mL, 17.7 mmol) was added and allowed to stir for 10 minutes. Ethyl bromoacetate (0.84 mL, 7.6 mmol) was added dropwise as the solution was allowed to

stir. The solution turned yellow, and a second portion of ethyl bromoacetate (0.84 mL, 7.6 mL) was added dropwise. The reaction was allowed to stir for two hours at 0 ^{o}C. Subsequently, the reaction was warmed to room temperature and the solvent was removed *in vacuo*. The remaining solid was dissolved in 20 mL of ethyl acetate and filtered. The ethyl acetate in the filtrate was removed *in vacuo*. The residue was purified by silica gel chromatography (EtOAc/NEt_3, 99:1). After purification, 0.96 g of product was isolated as yellow oil. (Yield: 45%). ^{1}H NMR ($CDCl_3$) δ 8.54 (m, 1H), 7.68 (m, 1H), 7.61 (m, 1H), 7.71 (m, 1H), 4.17 (q, 4H, J = 7 Hz), 4.07 (s, 2H), 3.62 (s, 4H), 1.27 (t, 6H, J = 7 Hz). ^{13}C NMR ($CDCl_3$) δ 171.3, 159.0, 149.2, 136.8, 123.2, 122.3, 60.7, 60.0, 55.0, 14.4. Mass (EI) (*m/z*) $[M]^+$ calc: 280.1423 obs: 280.1429.

*Potassium (2-picolylamino)-N-ethylacetate-N-acetate (**2.4**)*

To a solution of 0.5M KOH in EtOH (1.6 mL, 0.8 mmol) at 0 ^{o}C, **2.1** (0.23 g, 0.8 mmol) was added. The reaction was allowed to stir overnight. The solvent was removed *in vacuo* and the solid residue was triturated in diethyl ether. The remaining residue was dried *in vacuo*, resulting in 0.14 g of a yellow solid (Yield: 61%). ^{1}H NMR ($CDCl_3$) δ 8.58 (m, 1H), 7.54 (m, 1H), 7.12 (m, 1H), 7.08 (m, 1H), 4.00 (q, 2H, J = 7 Hz), 3.66 (s, 2H), 3.21 (s, 2H), 3.03 (s, 2H), 1.15 (t, 3H, J = 7 Hz).

*(2-picolylamino-N-acetate) (N,N,O) rhenium tricarbonyl-N-ethylacetate (**2.5**)*

To a solution of MeOH (10 mL) and $[Et_4N]_2[Re(CO)_3Br_3]$ (0.27 g, 0.35 mmol), **2.4** (0.1 g, 0.39 mmol) was added. The solution was allowed to stir overnight. The solvent was removed *in vacuo* to yield as solid residue. The residue purified by silica gel chromatography (EtOAc). This yields 0.10 g of white solid (Yield: 53%). ^{1}H NMR

($CDCl_3$) δ 8.84 (m, 1H), 7.96 (m, 1H), 7.51 (m, 1H), 7.46 (m, 1H), 5.28 (d, 1H, J = 15 Hz), 4.45 (d, 1H, J = 17 Hz), 4.37 (d, 1H, J = 15 Hz), 4.32 (d, 1H, J = 17 Hz) 4.31 (q, 2H, J = 7 Hz), 3.91 (d, 1H, J = 17 Hz), 3.69 (d, 1H, J = 17 Hz), 1.37 (t, 3H, J = 7 Hz)

*Potassium 2-picolylamino-N,N-diacetate (PADA) (**2.6**)*

To MeOH (20 mL) picolylchloride hydrochloride (0.68 g, 4.13 mmol) was added and allowed to stir for 5 minutes. Potassium hydroxide (0.23 g, 4.13 mmol) was added slowly. The solution begins to turn red and was stirred for 30 minutes; a solid forms. To a separate solution of MeOH (37.5 mL) and potassium hydroxide (0.65 g, 11.5 mmol), iminodiacetic acid (0.5 g, 3.75 mmol) was added. After 30 minutes of stirring, the picolylchloride hydrochloride solution was filtered and added dropwise to the iminodiacetic acid solution. The reaction mixture was refluxed overnight. The reaction was allowed to cool to room temperature and was filtered. The methanol of the filtrate was removed *in vacuo* and the residue was dissolved in water (10 mL) and washed with DCM (4 x 10mL). 1 M hydrochloric acid was added to the water solution until neutral pH was obtained. The water was then removed *in vacuo*. The remaining residue was redissolved in anhydrous methanol and filtered. The methanol was removed *in vacuo* and the residue was recrystallized in EtOH, yielding 0.45 g of product (Yield: 40%). ^{1}H NMR (MeOH-d_4) δ 8.59 (m, 1H), 7.85 (m, 1H), 7.84 (m, 1H), 7.39 (m, 1H), 4.48 (s, 2H), 3.71 (s, 4H). ^{13}C NMR (MeOH-d_4) δ 171.3, 153.2, 150.6, 139.3, 135.3, 135.2, 59.5, 58.0. Mass (ES-) $[M]^-$ *(m/z)* calc: 223.0719, obs: 223.0729.

*2-Picolylamino-N,N-diacetate rhenium(I) tricarbonyl (PADA-Re(CO)$_3$) (**2.7**)*

To a solution of MeOH (36 mL) and **2.6** (0.22 g, 0.75 mmol), $[Et_4N]_2[ReBr_3(CO)_3]$ (0.556 g, 0.75 mmol) was added slowly at room temperature. The reaction was stirred overnight, a white precipitate formed. The reaction mixture was filtered and dried, resulting in 0.233 g of product. (Yield: 58%) ^{1}H NMR (DMSO-d_6) δ 8.75 (m, 1H), 8.14 (m, 1H), 7.84 (m, 1H), 7.57 (m, 1H), 5.12 (AB, 1H, J = 16 Hz), 4.73 (AB, 1H, J = 16 Hz), 4.53 (AB, 1H, J = 17 Hz), 4.30 (AB, 1H, J = 17 Hz), 4.04 (AB, 1H, J = 17 Hz), 3.63 (AB, J = 17 Hz). ^{13}C NMR (DMSO-d_6): 197.0, 196.9, 196.7, 178.6, 169.7, 159.5, 151.7, 140.3, 125.7, 123.9, 67.8, 67.4, 61.5. Mass (ES+) $[M+H]^+$ *(m/z)* calc: 517.0022, obs: 517.0035.

*Bispicolylamino acetate (BPA) (**2.8**)*

To a suspension of 2-picolylchloride hydrochloride (0.6 g, 3.6 mmol) in water (10 mL), NaOH (0.26 g, 6.6 mmol) was added. The 2-picolylchloride hydrochloride dissolved and became red. Glycine (0.125 g, 1.6 mmol) was added. The reaction was refluxed overnight at 65 ^{o}C. The reaction mixture was cooled to room temperature and neutralized to pH = 7 using 1 M $HCl_{(aq)}$. The resulting solution was then washed with DCM. The aqueous layer was then removed *in vacuo*. The resulting residue was triturated with isopropyl alcohol and filtered. The filtrate was taken and the solvent removed *in vacuo*. The resulting residue was triturated with cold benzene and filtered. The filtrate was taken and the solvent removed *in vacuo*. This resulted in 0.08 g of brown powder (Yield: 20%). ^{1}H NMR (DMSO-d_6): δ 8.49 (m, 2H), 7.77 (m, 2H), 7.25 (m, 2H), 7.25 (m, 2H),

3.93 (s, 4H), 3.35 (s, 2H). ^{13}C NMR ($CDCl_3$): δ 174.0, 158.2, 148.6, 137.2, 122.8, 122.6, 59.9, 58.5. Mass (EI+) $[M+H]^+$ *(m/z)* calc: 257.1164, obs: 257.1159.

*Bispicolylamino acetate rhenium(I) tricarbonyl (BPA-Re(CO)$_3$) (**2.9**)*

To a solution of MeOH (20 mL), $[Et_4N]_2[ReBr_3(CO)_3]$ (0.10 g, 0.13 mmol) was added and stirred until it dissolved. **2.8** (0.04 g, 0.13 mmol) was dissolved in MeOH (2 mL) and added dropwise to the reaction mixture. The reaction was refluxed for 5 hours. Subsequently, the reaction was cooled to room temperature and the solvent was removed *in vacuo*. The resultant residue was recrystallized from methanol. This yielded 0.01 g product (Yield: 15 %). ^{1}H NMR (CD_3OD): δ 8.84 (m, 2H), 7.92 (m, 2H), 7.54 (m, 2H), 7.35 (m, 2H), 5.33 (AB, 2H, J = 17 Hz), 4.98 (AB, 2H, J = 17 Hz), 4.72 (s, 2H). ^{13}C NMR (CD_3OD): δ 212.7, 212.6, 171.2, 162.5, 153.2, 141.8, 127.1, 124.7, 69.4, 69.1. Mass (ES+) $[M]^+$ *(m/z)* calc: 527.0491, obs: 527.0481.

*Nitrilotriacetate rhenium(I) tricarbonyl (NTA-Re(CO)$_3$) (**2.10**)*

To a flask containing water (3 mL), nitrilotriacetic acid (0.06 g, 0.3 mmol) was added to form a suspension. 0.5 M $NaOH_{(aq)}$ (1.8 mL, 0.9 mmol) was added dropwise, resulting in a clear solution. 0.1 M $[Re(CO)_3(H_2O)_3]OTf_{(aq)}$ (3 mL, 0.3 mmol) was added dropwise and the reaction was allowed to stir overnight. The solvent was then removed *in vacuo*. The residue was then dissolved in IPA and filtered. The filtrate was taken and the solvent was removed *in vacuo*. The resulting white powder was triturated in DCM and filtered. The filtrate was dried, yielding 0.06 g of white powder (37% yield). ^{1}H NMR (CD_3OD): δ 4.21 (s, 2H), 4.03 (AB, 2H, J = 16 Hz), 3.83 (AB, 2H, J = 16 Hz). ^{13}C NMR (DMSO-

d_6): δ 198.4, 198.3, 178.3, 169.9, 68.0, 63.6. Mass (ES+) $[M+H]^+$ *(m/z)* calc: 505.9474, obs: 505.9470.

*2-[(2-Picolylamino)methyl]phenol (**2.11**)*

To a round bottom flask of MeOH (4.5 mL), salicaldehyde (0.213 mL, 2 mmol) was added. The reaction was stirred and 2-picolylamine was added dropwise (0.206 mL, 2 mmol). The reaction immediately became a yellow colour. After stirring for 2 hours, sodium borohydride (0.0756, 2 mmol) was added in four portions over the course of 1 hour. The yellow colour disappears and the reaction became clear. The solvent was removed *in vacuo*. The residue was dissolved in water and extracted into DCM. The DCM was removed *in vacuo* and the product was purified by silica gel chromatography (EtOAc/MeOH, 9:1). This yielded 0.22 g of a translucent oil (51% yield). ^{1}H NMR ($CDCl_3$): δ 8.60 (m, 1H), 7.68 (m, 1H), 7.21 (m, 3H), 6.98 (m, 1H), 6.87 (m, 1H), 6.79 (m, 1H), 4.02 (s, 2H), 3.94 (s, 2H). ^{13}C NMR ($CDCl_3$): δ 158.2, 157.8, 149.4, 136.7, 128.8, 128.6, 122.7, 122.4, 122.3, 119.0, 116.4, 53.1, 51.7. Mass (EI+) $[M+H]^+$ *(m/z)* calc: 214.1106, obs: 214.1106.

*2-Picolylamino-N,N-2-hydroxybenzyltert-butyl acetate (**2.12a**)*

To a stirred round bottom flask containing MeCN (4.5 mL), **2.11** (0.28 g, 1.3 mmol), *tert*-butyl bromoacetate (0.19 mL, 1.3 mmol) and triethylamine (0.18 mL, 1.3 mmol) were added. The reaction was allowed to stir overnight after which it became a dark yellow colour. The solvent was subsequently removed *in vacuo*, and the obtained residue was dissolved in EtOAc. The EtOAc was then filtered to remove an insoluble white powder and the solvent from the filtrate was removed *in vacuo*. The product was then purified by

silica gel chromatography [EtOAc]. The resulting fractions produced 0.13 g of product as a yellow oil (35% yield). ^{1}H NMR ($CDCl_3$): δ 10.32 (br, s, 1H) 8.58 (m, 1H), 7.68 (m, 1H), 7.35 (m, 1H), 7.20 (m, 2H), 7.02 (m, 1H), 6.91 (m, 1H), 6.79 (m, 1H), 3.96 (s, 2H), 3.89 (s, 2H), 3.32 (s, 2H), 1.48 (s, 9H). ^{13}C NMR ($CDCl_3$): δ 170.3, 157.7, 157.5, 149.0, 136.8, 129.7, 129.2, 123.3, 122.4, 122.1, 119.0, 116.5, 81.8, 58.9, 56.6, 55.0, 28.1. Mass (EI+) $[M+H]^+$ *(m/z)* calc: 328.1787, obs: 328.1775.

*2-Picolylamino-N,N-2-hydroxybenzyl, benzyl acetate (**2.12b**)*

This compound was synthesized in an identical fashion to **2.12a**, except for the utilization of benzyl bromoacetate in replacement of *tert*-butyl bromoacetate. The product was then purified by silica gel chromatography [EtOAc]. The resulting fractions produced 0.13 g of product as a yellow oil (25% yield). ^{1}H NMR ($CDCl_3$): δ 10.33 (br, s, 1H), 8.57 (m, 1H), 7.67 (m, 1H), 7.38-7.29 (m, 6H), 7.20 (m, 2H), 6.99 (m, 1H), 6.91 (m, 1H), 6.78 (m, 1H), 5.18 (s, 2H), 4.02 (s, 2H), 3.92 (s, 2H), 3.49 (s, 2H). Mass (EI+) $[M]^+$ *(m/z)* calc: 362.1630, obs: 362.1643.

*2-Picolylamino-N,N-2-hydroxybenzylethyl acetate (**2.12c**)*

This compound was synthesized in an identical fashion to **2.12a**, except for the utilization of ethyl bromoacetate in replacement of *tert*-butyl bromoacetate. The product was then purified by sildeica gel chromatography [EtOAc]. The resulting fractions produced 0.13 g of product as a yellow oil (25% yield). ^{1}H NMR ($CDCl_3$): δ 10.29 (br, s, 1H), 8.58 (m, 1H), 7.69 (m, 1H), 7.34 (m, 1H), 7.21 (m, 2H), 7.04 (m, 1H), 6.91 (m, 1H), 6.79 (m, 1H), 4.20 (q, 2H, J = 7 Hz), 3.99 (s, 2H), 3.91 (s, 2H), 3.43 (s, 2H), 1.28 (t, 3H, J = 7 Hz).

*2-Picolylamino-N,N-2-hydroxybenzyl acetate (**2.13**)*

To a flask containing DCM (5 mL), **2.12a** (0.48 g, 1.4 mmol) was added producing a yellow solution. TFA (5 mL) was added dropwise and immediately turned the solution clear. The reaction was stirred for 3 hours and the solvent was removed *in vacuo*. The resulting residue was purified by flash silica gel column chromatography (EtOAc/MeOH, 8:2). The pure fractions produced 0.31 g of product as a white powder (80% yield). ^{1}H NMR (CD_3OD): δ 8.58 (m, 1H), 7.83 (m, 1H), 7.48 (m, 1H), 7.38 (m, 1H), 7.24 (m, 2H), 6.85 (m, 2H), 4.31 (s, 2H), 4.29 (s, 2H), 3.56 (s, 2H). ^{13}C NMR (CD_3OD): δ 171.8, 158.0, 154.2, 150.5, 139.1, 133.0, 132.1, 125.2, 125.1, 121.1, 119.6, 116.9, 59.4, 57.1, 56.4. Mass (EI) $[M]^+$ *(m/z)* calc: 272.1161, obs: 272.1149

*2-Picolylamino-N,N-2-hydroxybenzyl acetate rhenium(I) tricarbonyl (PHA-Re(CO)$_3$) (**2.14**)*

To MeOH (20 mL) in a round bottom flask, $[Et_4N]_2[ReBr_3(CO)_3]$ (0.101 g, 0.13 mmol) was added. The solution was stirred for 10 minutes until the solution went clear. **2.13** (0.04 g, 0.13 mmol) was added. The solution was stirred under reflux for 5 hours. The solution was allowed to cool to room temperature and the solvent was removed *in vacuo*, producing a white powder. The powder was triturated in water and filtered. The resulting filtrand was dried *in vacuo* producing 0.02 g of a white powder (20% yield). ^{1}H NMR (DMSO-d_6): δ 10.17 (br s, 1H) 8.78 (m, 1H), 8.10 (m, 1H), 7.13 (m, 1H), 7.57 (m, 2H), 7.29 (m, 1H), 6.98 (m, 1H), 6.91 (m, 1H) 4.82 (AB, 1H, J = 16 Hz), 4.76 (AB, 1H, J = 13 Hz), 4.61 (AB, 1H, J = 13 Hz), 4.39 (AB, 1H, J = 16 Hz), 4.23 (AB, 1H, J = 17 Hz), 2.99 (AB, 1H, J = 17 Hz) . ^{13}C NMR (DMSO-d_6): δ 178.5, 159.3, 156.1, 151.8, 140.3,

133.4, 130.5, 125.7, 124.0, 119.1, 119.0, 115.9, 67.3, 65.6, 60.23. Mass (ES+) $[M+Na]^+$ *(m/z)* calc: 565.0385, obs: 565.0362.

Chapter 3: Investigation into the Fragmentation of Organometallic Peptides

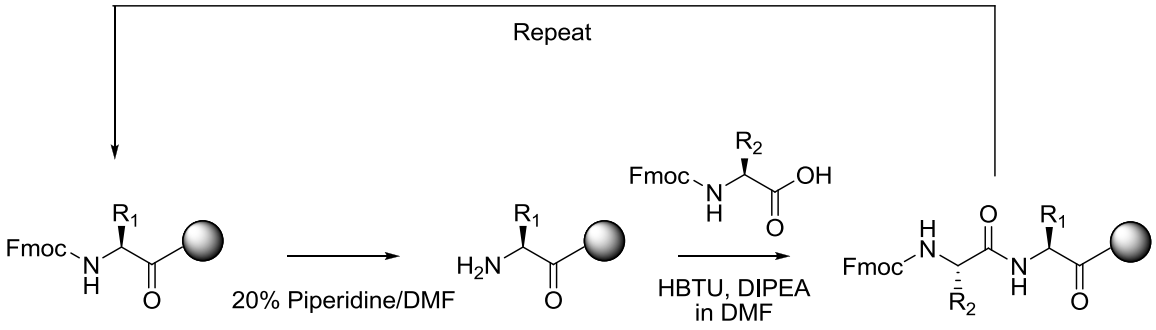
3.1 Introduction

In order to develop organometallic OBOC (one-bead, one-compound) peptide libraries, investigation of an amino acid sequencing method is necessary. Because of the nature of OBOC libraries, the identity of the compound present on a specific bead is unknown as OBOC libraries are non-spatially addressable. As such, the synthesis of organometallic peptides and the examination of their fragmentation patterns by to MALDI tandem mass spectrometry are investigated herein.

3.1.1 Peptide Synthesis

The synthesis of peptides often utilizes a well known methodology involving the use of orthogonal protecting groups and the sequential addition of amino acids on solid polymer support resin, known as solid phase peptide synthesis. This technique was initially reported by Merrifield in order to overcome difficult and time consuming solution phase peptide synthesis.[68] The most common technique in solid phase synthesis utilizes an Fmoc (fluorenylmethyloxycarbonyl) protecting group for the backbone amines of the amino acids. These groups are base labile and removed with piperidine in DMF. The side chains of the amino acids are protected with a variety of acid-labile groups, the identity of which depends on the nature of the amino acid. Common side-chain protecting groups include Boc (*tert*-butoxycarbonyl) for amines, *tert*-butyl ethers for alcohols and *tert*-butyl esters for carboxylic acids, among many others.

The synthesis of peptide sequences on a solid polymer support follows an iterative procedure of addition and deprotection of amino acids. First, an Fmoc-protected amino acid is added to the resin through a nucleophilic acyl substitution reaction. The free amine on the resin and the carboxylate of the Fmoc-amino acid react to form a peptide (amide) bond. The Fmoc group is necessary in order to avoid multiple additions of amino acids to the peptide chain. However, this reaction is slow due to the relative stability of carboxylates in comparison to amides and the moderate nucleophilicity of amines. As such, the addition of coupling agents is necessary in order to increase the rate of this reaction (Scheme 3.1).



Scheme 3.1 A diagram of the removal of an Fmoc group and the addition of a second amino acid to a resin bound amino acid. This process is repeated until the desirable peptide is synthesized.

After the addition of the Fmoc-amino acid, the removal of the Fmoc group from the peptide backbone is accomplished through the use of most commonly piperidine/DMF. These two steps can then be repeated with different Fmoc-amino acids such that the desired peptide sequence can be created.

After a peptide has been synthesized, the side-chain protecting groups are removed via the addition of acid, most commonly TFA (trifluoroacetic acid). The reagent used to cleave the acid-labile side chain protecting groups usually contains a small percentage of scavenging agents such as water, phenol or triisopropylsilane. The purpose of the addition of the scavenging agents is to react with electrophiles created from the cleavage of protecting groups, such as t-butyl cations. This is necessary in order to avoid the reaction of these electrophiles with the peptide chain.

There are a wide variety of solid-support resins available to be used. The general formulation of a solid-support involves the presence of an insoluble polystyrene cross-linked backbone on which PEG (polyethylene glycol) groups are attached. The PEG groups contain a functional group which a peptide can be built upon; the type of functional group depends on the type of resin used, and their purpose is primarily to provide a method to remove the peptide from the resin after complete synthesis.

Utilizing standard Fmoc solid phase peptide synthesis procedures, organometallic peptides can be synthesized and mass spectrometry can be used to investigate amino acid sequencing protocols for organometallic OBOC peptide libraries.

3.1.2 Mass Spectrometric OBOC Sequencing Methods

A number of mass spectrometric methods have been utilized in order to deconvolute compounds present in OBOC libraries in addition to Edman degradation. Ladder peptide synthesis, partial Edman degradation-MS (PED-MS), tagging techniques

and tandem mass spectrometry have all proven to be useful techniques for identification of bioactive compounds.

Ladder peptide synthesis involves an additional synthetic step during OBOC library synthesis. After each amino acid coupling, a small percentage of the peptide chains are "capped" such that subsequent amino acid additions do not react with the "capped" molecules.[69] After screening an OBOC library, mass spectrometry can identify all peptide sequences present on bead, and the difference between the peaks, corresponding to terminated peptide sequences has been shown to reveal the identity of the sequence.[70] However, the presence of truncated peptides, acting as contaminants during library screening, is not desirable.

PED-MS is similar to ladder peptide synthesis as a single OBOC bead will contain a series of truncated peptide sequences. Edman degradation is performed on a bead-bound peptide and with the addition of a percentage of capping agent. A small amount of peptide sequences will be terminated by the capping agent after each degradation step. Fmoc, Nic-OSu[71] and phenyl thiocyanate[72] groups have been shown to be useful for this strategy. The beads are then analyzed by mass spectrometry, and the difference in mass between truncated peptides allows for identification. The benefit of this technique over ladder synthesis is that all truncation is done after library screening. Recently, Thakkar *et al.* utilized PED-MS to sequence peptoid compounds from an OBOC library using Fmoc capping groups for sequence termination.[73] Additionally, PED-MS has been used to identify peptide-glycoprotein interactions from OBOC library screening.[74] The main drawback associated with PED-MS is the requirement of peptides to contain free N-terminal amines.

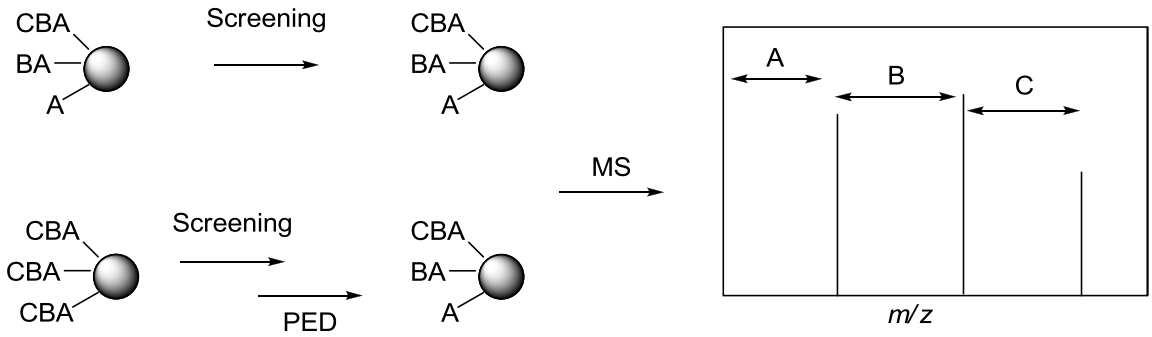


Figure 3.1 The general procedure of peptide deconvolution through the use of ladder synthesis and partial Edman degradation coupled with mass spectrometry.

Tagging methodologies are also popular techniques in screening OBOC libraries. A small percentage of the bead will contain a tag which can reveal the identity of the "hit" compound from an OBOC library. This technique is commonly used in small molecule OBOC libraries because of the difficulty in identifying a large number of small molecules as first reported by Liu *et al.*[75] Additions of amino acids on the bead corresponding to an addition of specific group to the small molecule, allow a small molecule to be identified by peptide mass spectrometry. Moreover, this technique has been utilized to test and identify cyclic peptides acting on breast cancer cell lines from OBOC libraries.[76] Additionally, the identification of isobaric amino acids is possible because the corresponding amino acid tag can be arbitrarily assigned to provide maximum possible information. The problem of interference between the tagging sequence and the bioactive compound has largely been overcome by the creation of biphasic solid phase support which allows for the synthesis of the tag on the interior of

the bead, and the bioactive compound on the exterior (Figure 3.2).[77] Limitations of such a procedure include an increase in synthetic steps and expense as well as the potential for decreased detectability of peptide due to the smaller amounts present on the bead.

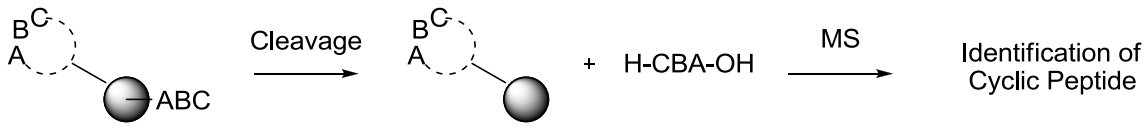


Figure 3.2 The utilization of linear peptides to code for bioactive cyclic peptides on a single bead. The inner core of the bead holds the tagging linear peptide and can be sequenced by MS methods in order to identify the cyclic tripeptide.

Tandem mass spectrometry is also an important technique that involves the ionization of peptides utilizing most commonly ESI (electrospray ionization) or MALDI (matrix-assisted laser desorption/ionization) methods. Electrospray ionization commonly produces multiply charged ions and can be used in conjunction with liquid chromatography,[78] while in utilizing MALDI the singly charged mass ion is observed. The ions formed are then separated based on their *m/z* ratio and subsequently fragmented. The ionic fragments that are produced can reveal information about the connectivity of the peptide. Tandem mass spectrometry is a very convenient method for OBOC peptide sequencing because it is able to quickly and economically provide sequences in a relatively high-throughput manner. The use of a UV MALDI laser to cleave peptides attached to beads via a photolinker,[79] as well as the sequencing of peptides through MALDI laser cleavage combined with tandem mass spectrometry[80] have been reported. Additionally, a combination of ladder synthesis and on-bead mass spectrometric

sequencing has also been reported.[81] However, sequencing of OBOC libraries containing octa- and pentapeptides via both on-bead and solution-phase MALDI tandem mass spectrometry has revealed an increase in spectrum quality via solution phase analysis.[82]

There are a wide-variety of OBOC sequencing methods available in the identification of bioactive compounds on "hit" beads. However, when developing an amino acid sequencing method applicable to organometallic peptides, various other considerations need to be examined. The properties of the imaging entity surrogate; the presence of rhenium(I) tricarbonyl chelates need to be taken into account as well as the position of the chelate on the N-terminal amine of the peptide chain. Tandem mass spectrometry was chosen as a tool to further investigate the deconvolution of organometallic OBOC peptide libraries.

3.1.3 Tandem Mass Spectrometry and Peptide Fragmentation

An investigation to observe the affects of the addition of the organometallic species affects the ionization or fragmentation of peptides is necessary in order to determine an appropriate method for amino acid sequencing. Further, if the chelate itself forms fragmentation products, these need to be identified. A variety of MALDI spectra of rhenium(I) tricarbonyl metal cores have previously been reported. In an investigation of porphyrin rhenium(I) tricarbonyl compounds by Toganohn *et al.*, MALDI results have shown a variety of prominent mass peaks in the spectrum of the precursor depending on the compound. The most common peak reported is the result of the loss of all three carbonyl ligands resulting in a $[M-3(CO)]^+$ peak.[83] Conversely, the MALDI results from

a vitamin B12 compound modified with a rhenium(I) tricarbonyl chelate, show the presence of primarily signals assigned to $[M\text{-}CO]^+$ and $[M\text{-}2(CO)]^+$, in addition to $[M]^+$ peak.[84] In these cases, carbonyls were readily discharged from the molecule, depending on the type of coordination environment. If this is observed for peptide sequencing applications, it could complicate organometallic peptide sequencing due to a number of low intensity peaks as well as removal of metallic carbonyls in fragmentation. Therefore, the identity of peaks might be ambiguous if it is unknown how many carbonyls the fragment contains. From these reported results, it would seem that the removal of carbonyls in MALDI mass spectrometry is highly dependent on the coordination environment of the metal core and therefore it is prudent to investigate the fragmentation properties of rhenium(I) tricarbonyl tridentate chelates.

Fragmentation though mass spectrometry for determining amino acid sequences in peptides is dependent on the ability of peptides to produce well defined series of peptide fragments which correspond to individual amino acids (Figure 3.3).

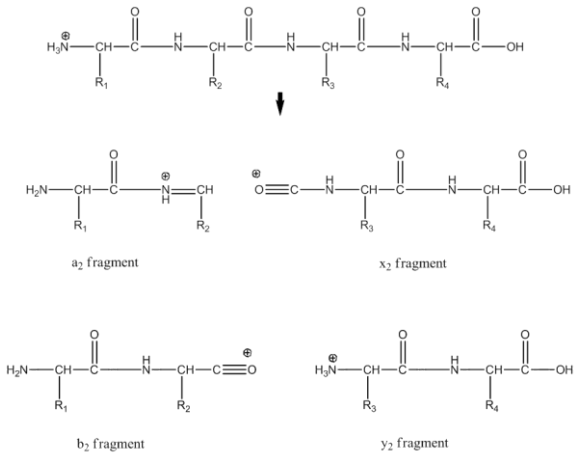


Figure 3.3 A generic tetrapeptide holding a positive charge on the N-terminus. This compound is fragmented through CID (collision-induced dissociation) The a_2 and b_2 fragments are N-terminal fragments while the x_2 and y_2 fragments are C-terminal.

The types of fragments shown above can be classified into two types of peptide fragments. The *a* and *b* fragments maintain a charge on the N-terminus of the peptide. Conversely, the *x* and *y* fragments maintain a charge on the C-terminus of the peptide. The fragmentation pathways are not fully understood, however it is known that *b* ions in have a protonated oxazolone structure[85] and are an intermediate in the creation of *a* fragments.[86] The *y* ions have been shown to be truncated peptide ions as shown in figure 3.3.[87] A peptide would be expected to produce a series of fragments corresponding to each peptide bond. The difference in mass between two fragments from the same series can be used to identify an amino acid based on the known mass of the side chain. This method is known as *de novo* peptide sequencing,[88] and has been shown to be highly applicable to small peptide sequences.[89]

Theoretically, this process is very simple and straight forward. However, a problem associated with peptide fragmentation is that the fragmentation patterns are sequence dependent. The source of the diversity in fragmentation patterns is that amino acids have a wide variety of side chains which have dramatically different abilities to accept and donate protons. The mobile proton theory is the most complete explanation of peptide fragmentation in tandem mass spectrometry. The theory states that during the initial ionization, basic residues such as amines accept protons from the matrix. During CID, the protons can move to various other parts of the peptide chain such as an amide nitrogen or oxygen, which leads to fragmentation.[90] Therefore, the fragmentation patterns depend on where and in which ratios the peptide initially becomes protonated, which residues are in close enough physical proximity to receive the proton after collisions with gas molecules and which are the most energetically favourable residues to

receive the proton. Thus, the process of peptide fragmentation involves a myriad of pathways and potential pathways which need to be understood if the presence or absence of a fragment is to be rationalized. Attempts to simplify *de novo* sequencing have been accomplished by the addition of isotopic signatures on the terminal of the peptide sequence in order to more easily identify N- and C-terminal fragments. A recent example of isotopic tagging involves the coupling of bromo pyridyl derivatives to the N-terminal of the peptide sequence and allows for identification of the N-terminal fragments via the bromine isotopic ratio.[91] Interestingly, this technique can also be applied in rhenium(I) tricarbonyl organometallic OBOC libraries because of the chelates are on the N-terminus and rhenium has a distinct isotopic ratio.

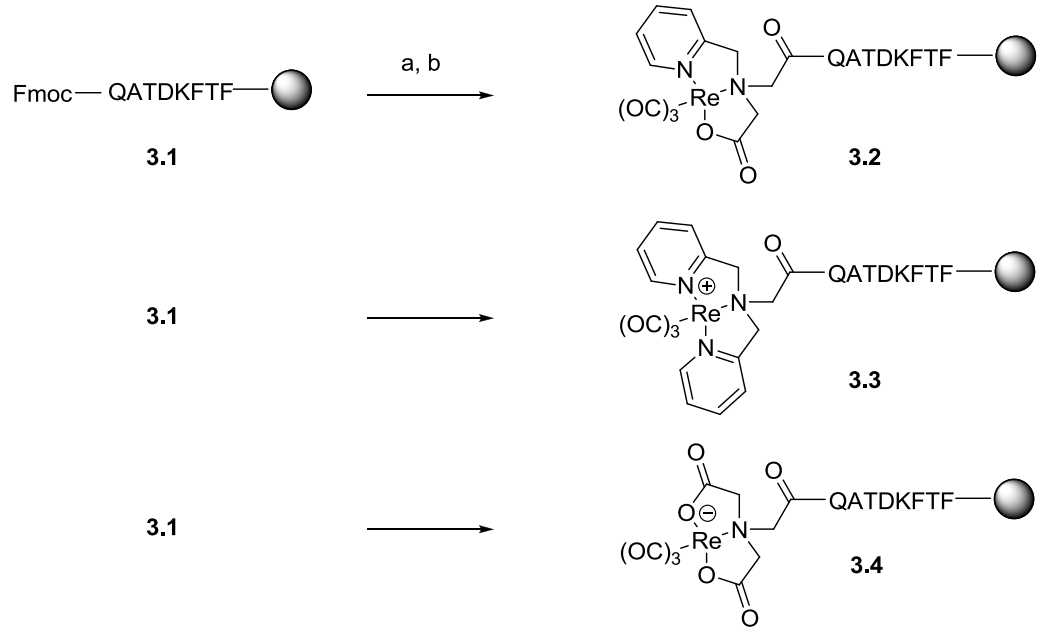
Because peptide fragmentation is not fully predictable, and the identity of the amino acids present as well as the three dimensional conformation of the peptide are all factors in obtaining mass spectra, the effect of the rhenium(I) tricarbonyl chelates also needs to be taken into account. It is possible that the presence of the organometallic may affect proton movement such that there are few favourable fragmentation pathways, and it is also possible that the type of chelate present can produce dramatically different fragments. Proteins and peptides are often deconvoluted with the assistance of protein databases, whereby fragmentation patterns are compared to thousands of known database entries.[92] However, due to the presence of rhenium(I) tricarbonyl chelates in the organometallic peptides, these techniques are not available. Therefore, variables associated with organometallic peptide fragmentation need to be addressed in order to utilize MALDI tandem mass spectrometry for sequencing organometallic peptides in OBOC libraries.

3.2 Results and Discussion

This chapter will outline the investigation of the synthesis and fragmentation patterns of peptides with N-terminal rhenium(I) tricarbonyl chelates. Firstly, the peptide synthesis and coupling of the organometallic will be discussed. Secondly, the results of MALDI TOF/TOF tandem mass spectrometry with organometallic peptides will be shown. Thirdly, the adaptation of these techniques to OBOC peptide library screening conditions is investigated herein.

3.2.1 Organometallic Peptide Synthesis

The synthesis of organometallic peptides proceeds through standard Fmoc solid phase peptide synthesis protocols, as discussed previously. The peptide that was chosen to test the influence of a rhenium(I) tricarbonyl chelate on peptide fragmentation, compound **3.1**, is a compound which has previously been reported to produce fully sequencable fragmentation patterns utilizing MALDI TOF/TOF mass spectrometry (QATDKFTF-NH_2). The coupling of the organometallic rhenium(I) tricarbonyl was shown for PADA, BPA, and NTA compounds (Scheme 3.2).

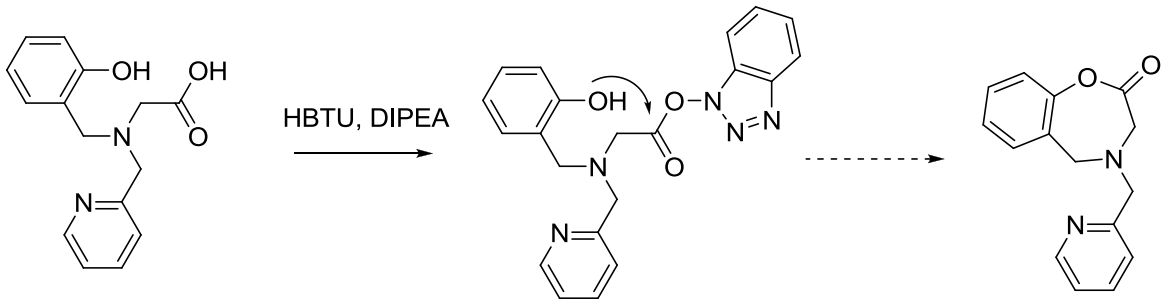


Scheme 3.2 Synthesis of organometallic peptides. Reagents: a) 20% piperidine in DMF b) HBTU, DIPEA, rhenium(I) tricarbonyl chelate in DMF. Rhenium(I) tricarbonyl chelates **2.7**, **2.9** and **2.10** were coupled onto the N-terminus of **3.1** to produce **3.2**, **3.3** and **3.4**, respectively.

With the peptides synthesized, the Fmoc is removed under standard conditions, and the organometallic chelate is coupled to the N-terminus utilizing standard amino acid coupling conditions of HBTU, DIPEA and the chelate in DMF. This is the benefit of utilizing bifunctional chelates with a free carboxylic acids; the chemistry of adding a free carboxylic acid to an amine is well-known and high yielding. This results in three organometallic peptides, **3.2**, **3.3** and **3.4**, which contain three different rhenium(I) tricarbonyl chelates on the N-terminus, neutral, positive and negative respectively. These compounds were purified by HPLC (high-performance liquid chromatography) and their identity confirmed by mass spectrometry.

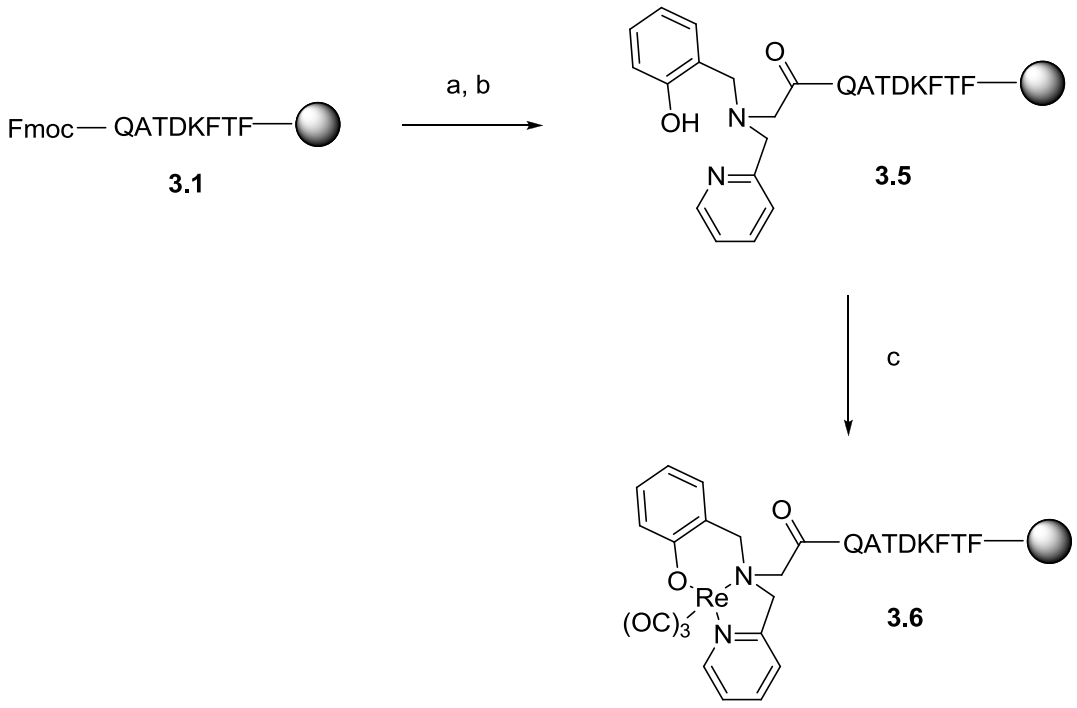
The last organometallic peptide to be synthesized involves the phenol-based chelator discussed in the previous chapter. The PHA chelate (**2.14**) proved to be of an ambiguous coordination sphere; it is unclear if the phenol or the carboxylate is donating electrons to the metal core. Therefore, an attempt to couple **2.14** to the N-terminus of **3.1** was made under identical conditions as shown in Scheme 3.2. HPLC analysis showed that no reaction had taken place, providing evidence for carboxylate-based coordination to the rhenium(I) tricarbonyl core. Therefore, an alternate method of synthesizing the PHA-organometallic peptide was undertaken utilizing on-resin coordination.

The uncoordinated PHA chelator (**2.13**) coupled to the peptide compound with a free N-terminus. Potential problems associated with this process are mainly the fact that the phenol group remains unprotected during this process. During the activation of the chelate carboxylic acid, there is potential that the phenol group, being a nucleophile, can attack the acyl position producing an ester.



Scheme 3.3 A potential reaction during the activation of the carboxylate of **2.13** with the coupling reagent HBTU. The phenol nucleophile might be expected to form a seven membered cyclic ring system through intramolecular nucleophylic acyl substitution of the activated ester.

However, the potential cyclic compound shown (Scheme 3.3), might also be susceptible to attack by a nucleophile by the amine of the peptide backbone. Regardless, the PHA chelator was coupled onto the N-terminus of compound. By HPLC, this reaction produced the desired product **3.5**, and subsequently $[Et_4N]_2[Re(CO)_3Br_3]$ was added in a DMF/MeOH mixture with DIPEA to produce the desired organometallic peptide shown below. This product was then purified by preparative HPLC and **3.6** was identified by mass spectrometry.



Scheme 3.4 Synthesis of **3.6**. Reagents: a) 20% piperidine in DMF. b) HBTU, DIPEA, **2.13** in DMF. c) $[Et_4N]_2[Re(CO)_3Br_3]$, DIPEA, in MeOH/DMF. Synthesis of $Re(CO)_3$-PHA-QATDKFTF (**3.6**) proceeds through on an on resin coordination methodology.

With four organometallic rhenium(I) tricarbonyl peptides with unique chelates synthesized, the peptides were removed from resin, as well as the side chain protecting

groups deprotected, utilizing Reagent K, a cleavage solution consisting of 95% TFA: 2.5% H_2O : 2.5% TIS over the course of three hours. This results in solution phase peptides containing a C-terminal amide due to the Rink Amide resin. The stability of the resultant cleaved and deprotected organometallic peptides was investigated by analytical HPLC. **3.2** ($Re(CO)_3$-PADA), **3.3** ($Re(CO)_3$-BPA) and **3.6** ($Re(CO)_3$-PHA) peptides show no decomposition of the chelate and high conversion rates of chelate coupling. However, **3.4** ($Re(CO)_3$-NTA) shows the desired product, the peptide with the chelator (loss of rhenium(I) tricarbonyl) and the peptide with a free N terminus (**3.1**) (Figure 3.4).

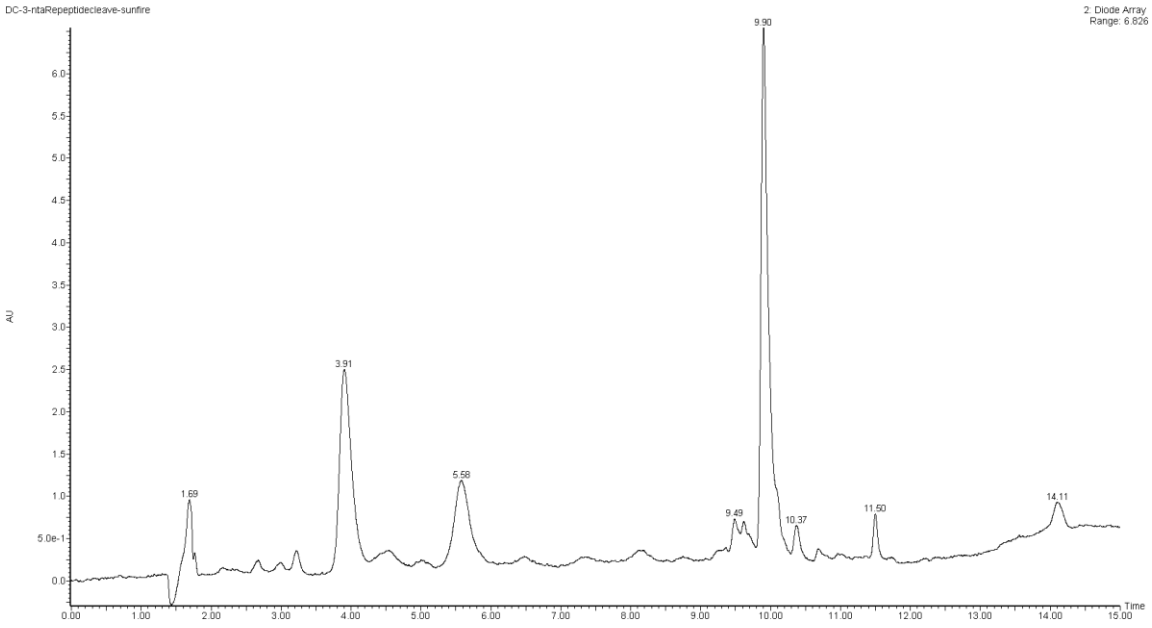


Figure 3.4 Crude HPLC UV trace of $Re(CO)_3$-NTA-QATDKFTF-NH_2 (**3.4**) after cleavage from resin. The major peak at 9.90 min represents the desired product with an *m/z* 1399. The two peaks at 3.91 and 5.58 min show an observed mass of *m/z* 956 and 1129 which correspond to the H-QATDKFTF-NH_2 and NTA-QATDKFTF-NH_2, respectively.

It is not surprising that the least stable chelate observed is the $Re(CO)_3$-NTA compound (**3.4**), which contains two donating carboxylates as well as formally holding a

negative charge. The increase in electron density as well as the decrease in steric hindrance associated with absence of an aromatic system might produce a metal centre with decreased stability towards highly acidic conditions. Therefore, because of the apparent decomposition of **3.4** as shown by HPLC, its use as a chelate in a library is not possible due to the absence of a viable purification method when screening a library. Nevertheless, all of the desired organometallic peptides were purified by semi-preparative HPLC, resulting in pure white solids as products. Desired products were identified by ES+ and MALDI mass spectrometry.

Through coupling rhenium(I) tricarbonyl chelates to the N-terminal amines of the peptide QATDKFTF, the organometallic PADA, BPA and NTA peptides were created. A second method of on-resin coordination was utilized to synthesize the organometallic PHA peptide. The fragmentation patterns of these organometallic peptides are then investigated utilizing MALDI TOF/TOF mass spectrometry.

3.2.2 MALDI Tandem Mass Spectrometry

In order to investigate the viability of MALDI tandem mass spectrometry as a method to deconvolute N-terminal organometallic peptides, initial MALDI studies were undertaken on purified organometallic peptides.

The organometallic peptides were dissolved in water and mixed with a solution of CHCA (*trans*-α-cyano-4-hydroxycinnamic acid). The solution is then spotted on a MALDI plate and is allowed to slowly evaporate resulting in the organometallic peptide

and the CHCA co-crystallized. The matrix absorbs UV light from a laser resulting in ablation and the transfer of a proton to the peptide, which can then be detected. There are many known matrices; however CHCA has been reported to be the most useful matrix for short peptides. The mass spectrum obtained shows the expected mass peaks for all of the organometallic peptides, with the associated rhenium-185/187 isotopic signature.

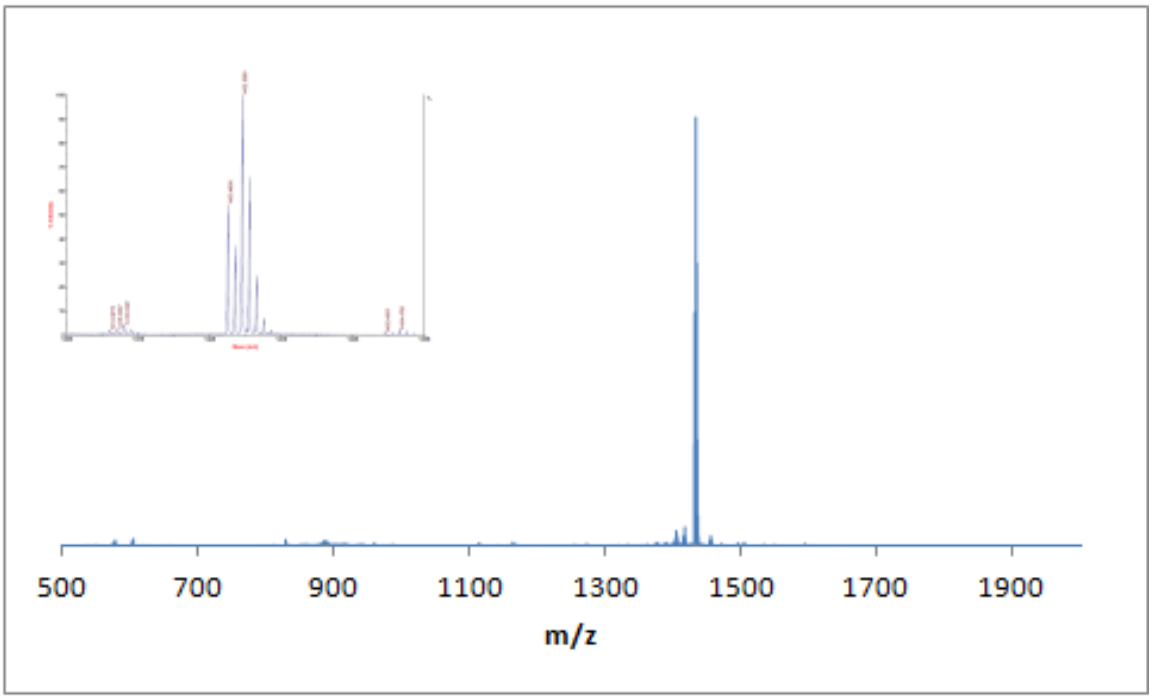


Figure 3.5 MALDI MS precursor ion spectrum for **3.2**. The organometallic peptide has the expected rhenium 185/187 isotope ratio. $[M+H]^+$ *m/z* calc: 1432.49 obs: 1432.50 (rhenium-187 peak).

The desired mass is then selected for further fragmentation. Both rhenium 185 and rhenium 187 peaks were selected for fragmentation. The results of these MS/MS spectra indicated that despite the increased intensity of the rhenium-187 peak, when the rhenium-185 peak is selected, the N-terminal fragments are easily and unambiguously identified due to the isotopic signature present in the MS/MS spectrum. This isotopic

"doublet" is not obvious in the rhenium-187 tandem mass spectrum. The product ion spectrum was obtained for the $Re(CO)_3$-PADA-QATDKFTF-NH_2 (**3.2**) (Figure 3.6).

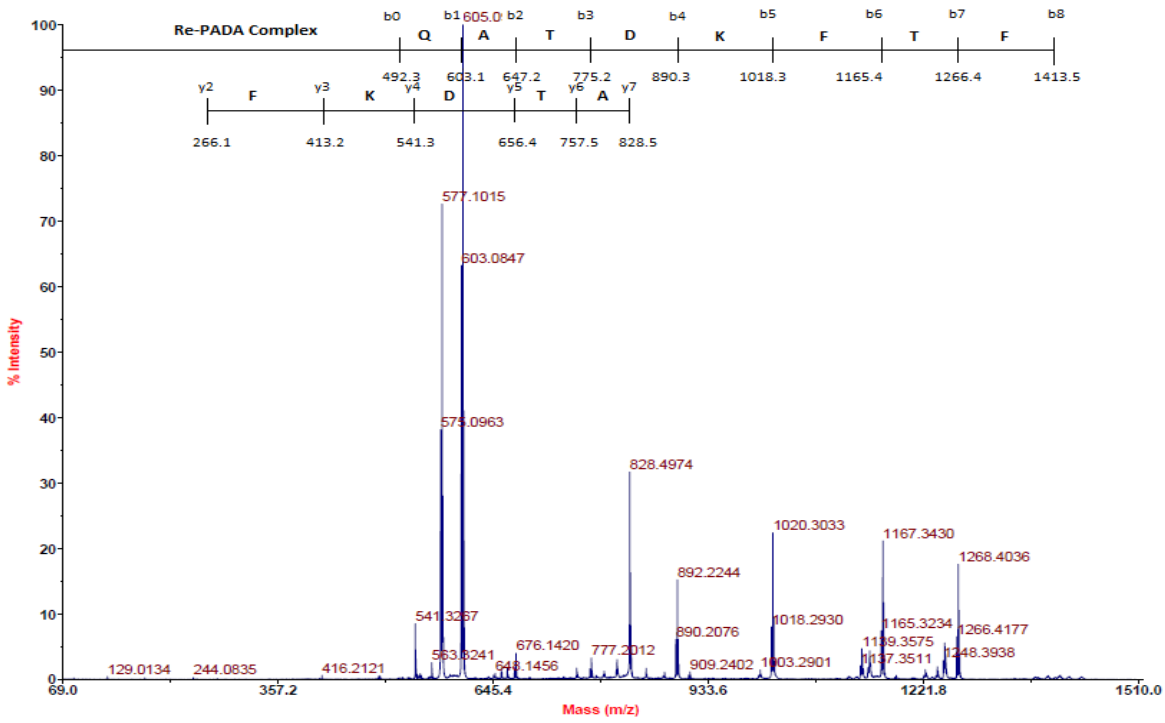


Figure 3.6 The product ion spectrum of compound **3.2**. The corresponding assignments of *b* and *y* series can be seen at the top of the figure. The identity of amino acids is determined based on the difference between the *m/z* in two adjacent peaks of the same series.

The rhenium-containing peptide chain gives a well defined fragmentation pattern that can be used to deconvolute the amino acid sequence. There are a number of identifiable fragment series present in the spectrum; the *b* and *y* series of fragments are the most abundant as they form a complete series where all amino acid fragments are present. The b_1 fragment is the most intense peak in the spectrum, making it the most favourable fragmentation pathway for this particular peptide. A table of the identifiable *b* and *y* fragments is shown below:

Table 3.1 Fragment assignments from series *b* and *y* of the product ion spectrum of compound **3.2**. $[M+H]^+$ *m/z* = 1430.5 (rhenium-185 peak)

Series	Label	Fragment	Peak (*m/z*)	Difference between peaks (*m/z*)	Amino Acid Assignment
b series	b_0	RE-PADA+NH_3	492.3		
	b_1	RE-PADA-Q	603.1	110.8	Glutamine
	b_2	RE-PADA-QA	674.2	71.1	Alanine
	b_3	RE-PADA-QAT	775.2	101.0	Threonine
	b_4	RE-PADA-QATD	890.3	115.1	Aspartic Acid
	b_5	RE-PADA-QATDK	1018.3	128.0	Lysine
	b_6	RE-PADA-QATDKF	1165.4	147.1	Phenylalanine
	b_7	RE-PADA-QATDKFT	1266.4	101.0	Threonine
	b_8	RE-PADA-QATDKFTF	1413.5	147.1	Phenylalanine
y series	y_1	NOT OBSERVED			
	y_2	TF	266.1		
	y_3	FTF	413.2	147.1	Phenylalanine
	y_4	KFTF	541.3	128.1	Lysine
	y_5	DKFTF	656.4	115.1	Aspartic Acid
	y_6	TDKFTF	757.5	101.1	Threonine
	y_7	ATDKFTF	828.5	71.0	Alanine
	y_8	NOT OBSERVED			

The peptide is sequenced by obtaining the difference between two adjacent peaks of the same series (Figure 3.7).

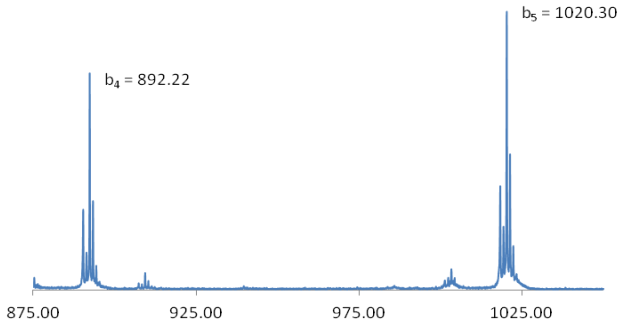


Figure 3.7 Two adjacent peaks, b_4 and b_5, from the product ion spectrum of compound **3.2**. The rhenium isotopic ratio can clearly be seen in each of the peaks, verifying the assignment of their identities as N-terminal fragments.

The difference of the two peaks corresponds to a mass of 128.1 Da which corresponds to a lysine residue. Figure 3.8 is also an illustration of the rhenium isotopic pattern present in the tandem mass spectrum.

The co-crystallization and MALDI tandem mass spectrometry procedures were repeated with all of the organometallic peptides and produced product ion spectra for all of the organometallic peptides (Figure 3.8).

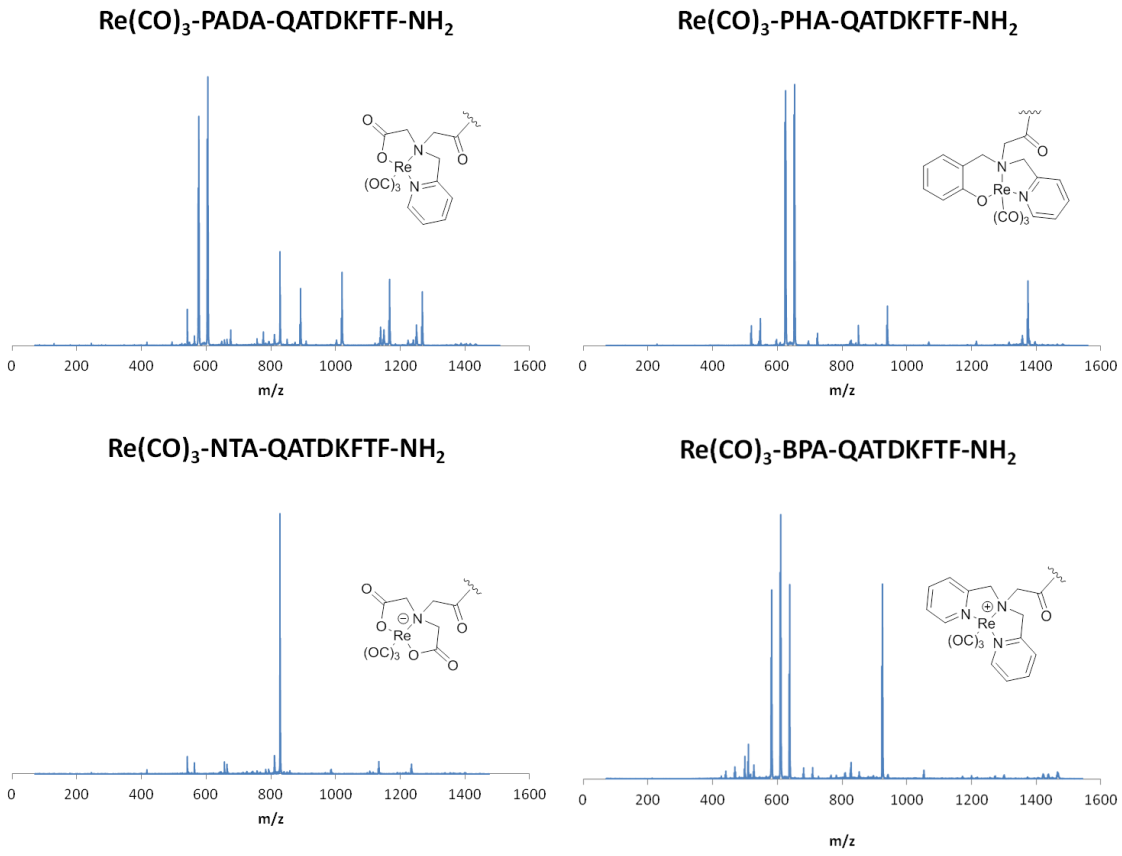


Figure 3.8 The product ion spectra of compounds **3.2** (PADA), **3.3** (BPA), **3.4** (NTA) and **3.6** (PHA). The effects of different rhenium(I) tricarbonyl chelates can be seen in the fragmentation patterns and intensities of each spectra.

From the resultant fragmentation patterns, the correct amino acid sequences of all of the peptides could be obtained. However, there is a considerable variation in the observed fragmentation patterns and intensities associated with different chelates on the N-terminus of the exact same peptide sequence.

It is apparent that the presence of a charged chelate has a large impact on the fragmentation patterns of the peptide. The $Re(CO)_3$-NTA peptide shows a significant decrease in observable fragments and intensity for all but the b_4 ion which has an *m/z* of 859.2. There are no complete series present in this spectrum, roughly half of the *b* and half of the *y* series are clearly distinguished (Table 3.2).

Table 3.2 Fragment assignments from series *b* and *y* of the product ion spectrum of compound **3.4**. The masses of the b_8 and y_8 ions can be determined because of the known mass of the peptide and the N-terminal chelate. $[M+2H]^+$ *m/z* = 1397.5 (rhenium-185 peak).

Series	Label	Fragment	Peak (*m/z*)	Difference between peaks (*m/z*)	Amino Acid Assignment
b series	b_0	NOT OBSERVED			
	b_1	NOT OBSERVED			
	b_2	NOT OBSERVED			
	b_3	RE-NTA-QAT	744.2		
	b_4	RE-NTA-QATD	859.2	115.0	Aspartic Acid
	b_5	RE-NTA-QATDK	987.3	128.1	Lysine
	b_6	RE-NTA-QATDKF	1134.3	147.0	Phenylalanine
	b_7	RE-NTA-QATDKFT	1235.3	101.0	Threonine
	b_8	NOT OBSERVED	1382.4	147.1	Phenylalanine
y series	y_1	NOT OBSERVED			
	y_2	NOT OBSERVED			
	y_3	NOT OBSERVED			
	y_4	KFTF	541.3		
	y_5	DKFTF	656.4	115.1	Aspartic Acid
	y_6	TDKFTF	757.4	101.0	Threonine
	y_7	ATDKFTF	828.5	71.1	Alanine
	y_8	NOT OBSERVED	956.5	128.0	Glutamine

From these two series it is possible to deconvolute the identity of the sequence. Despite the fact that the sequence was successfully obtained, the results cast doubt on the versatility of utilizing this chelate with a wide variety of unknown peptide sequences. The *de novo* sequencing of peptides is significantly simplified when the compound is of known connectivity, as one would expect. Furthermore, in the precursor mass spectrum, approximately 50% decomposition of the chelate is also observed. This is potentially the result of two mechanisms. Firstly, the compound might be unstable to the co-crystallization matrix, as CHCA is an acid and **3.4** shows limited stability to TFA. Secondly, this might be a result of the ionization process itself. Decomposition occurs with the transfer of protons to the organometallic peptide. Unfortunately, the combined lack of stability to TFA, the decomposition during MALDI and poor fragmentation suggest that these chelates are unsuitable for OBOC library screening. The negative chelate may be the cause of the poor fragmentation in a positive mass spectrometry experiment. A potential solution to this problem would be to utilize a negative mass spectrometry analysis. However, this was not pursued due to the fact that the matrix remains the same in the negative experiment and will lead to decomposition, in addition to the lack of TFA stability.

Conversely, **3.3** holds a positive charge and if the negative charge has a major contribution to the poor fragmentation the of $Re(CO)_3$-NTA peptide, it is expected that the positive charge would produce many high intensity fragments due to increased conversion of peptide to gas phase ion. A similar result was obtained recently by Cydzik *et al.* using positively charged tetraethylammonium N-terminal peptides to produce more intense fragments.[93] Indeed this is what is observed with the $Re(CO)_3$-BPA peptide. The

entire *b* and *a* series which both correspond to N-terminal charge species are present. This is not surprising because the N-terminus of the peptide contains a positive charge. There are also no observable *y* series fragments that correspond to a C-terminal charged fragment. This can be rationalized by the fact that *y* series fragments generally require the movement of the additional proton added by the matrix to the C-terminus. However, since $Re(CO)_3$-BPA-QATDKFTF-NH_2 is already positively charged it does not gain a proton from the matrix, as verified by the mass of the precursor peaks. The presence of every *a* series ion is somewhat unique in tandem mass spectrometry utilizing collision-induced dissociation and rarely reported (Table 3.3).

Table 3.3 Fragment assignments from series *b* and *a* of the product ion spectrum of compound **3.3**. The masses of the b_8 and a_8 ions can be determined because of the known mass of the peptide. The relationship between the *a* and *b* series is the loss of a carbonyl, *m/z* = 28. $[M]^+$ *m/z* = 1463.6 (rhenium-185 peak).

Series	Label	Fragment	Peak (*m/z*)	Difference between peaks (*m/z*)	Amino Acid Assignment
b series	b_0	RE-BPA+NH_2	527.1		
	b_1	RE-BPA-Q	638.2	111.1	Glutamine
	b_2	RE-BPA-QA	709.2	71.0	Alanine
	b_3	RE-BPA-QAT	810.3	101.1	Threonine
	b_4	RE-BPA-QATD	925.3	115.0	Aspartic Acid
	b_5	RE-BPA-QATDK	1053.4	128.1	Lysine
	b_6	RE-BPA-QATDKF	1200.4	147.0	Phenylalanine
	b_7	RE-BPA-QATDKFT	1301.5	101.1	Threonine
	b_8	NOT OBSERVED	1448.7	147.2	Phenylalanine
a series	a_0	NOT OBSERVED			
	a_1	RE-BPA-Q-CO	610.2		
	a_2	RE-BPA-QA-CO	681.2	71.0	Alanine
	a_3	RE-BPA-QAT-CO	782.2	101.0	Threonine
	a_4	RE-BPA-QATD-CO	897.3	115.1	Aspartic Acid
	a_5	RE-BPA-QATDK-CO	1025.4	128.1	Lysine
	a_6	RE-BPA-QATDKF-CO	1172.4	147.0	Phenylalanine
	a_7	RE-BPA-QATDKFT-CO	1273.4	101.0	Threonine
	a_8	NOT OBSERVED	1420.5	147.1	Phenylalanine

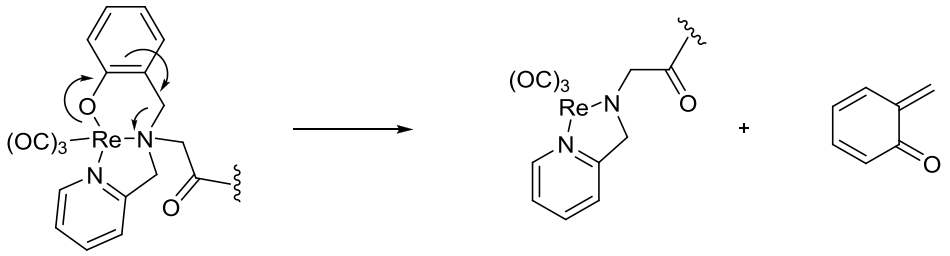
Because of the high intensity, complete fragment series and relative stability of BPA peptide (**3.3**) to acidic conditions, BPA is a candidate for use in OBOC peptide libraries. A potential problem in sequencing unknown peptides would be the large number of N-terminally charged fragments might make *de novo* sequencing difficult due to an increasing probability of inappropriate determinations of structure.

Like the $Re(CO)_3$-PADA peptide, $Re(CO)_3$-PHA-QATDKFTF-NH_2 (**3.6**), also forms neutral complexes with rhenium(I) tricarbonyl. As such, one would expect similar fragmentation patterns and peak intensities. Although the PADA and PHA chelates produce the most similar mass spectra of any of the chelates, there are a number of differences. While the b_1 peaks for both **3.2** and **3.6**, *m/z* 603.1 and 651.2, respectively, are both the base peaks for each spectrum, the relative intensities of peaks present in the PHA spectrum are generally lower in comparison to the dominant b_1 fragment,. There is a full complement of *b* series fragments that were used to successfully sequence the peptide. However, there was only one identifiable *y* peak, y_7 with an *m/z* = 828.5.

Table 3.4 Fragment assignments from series *b* of the product ion spectrum of compound **3.6**. The masses of the b_8 ion can be determined because of the known mass of the peptide. $[M+H]^+$ *m/z* = 1478.5 (rhenium-185 peak).

Series	Label	Fragment	Peak (*m/z*)	Difference between peaks (*m/z*)	Amino Acid Assignment
b series	b_0	RE-PHA-NH3	540.1		
	b_1	RE-PHA-Q	651.2	111.1	Glutamine
	b_2	RE-PHA-QA	722.2	71	Alanine
	b_3	RE-PHA-QAT	823.3	101.1	Threonine
	b_4	RE-PHA-QATD	938.3	115	Aspartic Acid
	b_5	RE-PHA-QATDK	1066.4	128.1	Lysine
	b_6	RE-PHA-QATDKF	1213.4	147	Phenylalanine
	b_7	RE-PHA-QATDKFT	1314.5	101.1	Threonine
	b_8	NOT OBSERVED	1461.5	147	Phenylalanine

Perhaps the most interesting aspect of the behaviour of PHA in the daughter ion spectrum is the presence of chelate fragmentation patterns. These fragments generally consist of a loss of an *m/z* of 106 when compared to *b* fragments. These peaks also contain the characteristic rhenium isotopic signature indicating the presence of rhenium in the fragment. This shows that these fragments must be N-terminal in nature. Further, these peaks, although not present in a full series, are found over a very large mass range. For example, a peak observed at *m/z* 1374.5 is 106.0 less than the $[M+H]^+$ of m/z 1480.5. Furthermore, there is a loss of 106.0 from the b_1 ion of *m/z* 651.2. Because of the presence of the mass loss throughout the entire series of *b* ions, as well as some of the *a* ions, the loss of 106.0 can only be explained by a fragmentation of the chelate (Scheme 3.5).



Scheme 3.5 The proposed fragmentation mechanism of the phenol-based chelate in compound **3.6**. The creation of an orthoquinone methide accounts for the loss of *m/z* = 106 observed in the mass spectrum.

This mechanism describes the creation of an orthoquinone methide. This explains the loss of the appropriate mass as well as the maintenance of the uncharged nature of the N-terminal chelate required for agreement with the observed fragmentation. The resultant bidentate chelate contains an amide functionality which balances the overall

positive oxidation state of the rhenium(I) tricarbonyl. There are four relatively intense fragments that correspond to the loss of a quinone methide (Table 3.5).

Table 3.5 The observed fragments in the product ion spectrum of compound **3.6** due to the loss of an *m/z* = 106 corresponding to the loss of an orthoquinone methide.

Series and Label	**Fragment**	**Observed *m/z***
$[M+H]^+$ - 106	$Re(CO)_3$-PHA-QATDKFTF-NH_3^+ - 106	1374.5
b_8 - 106	$Re(CO)_3$-PHA-QATDKFTF - 106	1357.5
b_1 - 106	$Re(CO)_3$-PHA-Q – 106	547.1
a_1 - 106	$Re(CO)_3$-PHA-Q – 106	519.1

From the fragmentation results of the PHA chelate organometallic peptide, it is clear that rhenium(I) tricarbonyl chelates are not universally stable to the fragmentation conditions. The additional information obtained from these experiments is important if a MALDI TOF/TOF sequencing method is to be used in the screening of organometallic OBOC libraries.

3.3 Conclusions

The results of these experiments into the fragmentation patterns of rhenium(I) tricarbonyl peptides has provided useful information into the effects of N-terminal organometallics on peptide fragmentation. These results have provided the first *de novo* sequencing of rhenium(I) tricarbonyl containing peptides which provides a foundation for further research into the identification of organometallic peptides for use in molecular imaging.

The same peptide sequence was used with the addition of four rhenium(I) tricarbonyl bifunctional chelates that have been directly attached to the peptide N-terminus. All of the resulting compounds were successfully sequenced by tandem mass spectrometric methods. The fragmentation patterns of neutral, positive and negative organometallics provide an insight into the different effects of the N-terminal chelates. The fragmentation patterns are unique between each charge state, most likely the result of different three dimensional conformations and mobile proton pathways during collision-induced dissociation. Furthermore, PADA and PHA, both neutral chelates, also produce different fragmentation patterns. The steric differences between a carboxylate and phenol ring produce unique fragmentation patterns and especially relative intensities between peaks.

Overall, the rhenium(I) tricarbonyl chelates have proven to be extremely stable to decarbonylation in MALDI mass spectrometry, in opposition to the literature reports. The chelates, in general, have also been shown to be stable to fragmentation conditions, with the notable exception of **3.6**, producing spectra that can be interpreted to fully sequence peptides. The results of these experiments have shown that MALDI TOF/TOF is an appropriate method for deconvolution of organometallic OBOC libraries.

3.4 Experimental

3.4.1 General procedures and materials

All chemicals and solvents were purchased from commercially available sources and used without further purification. Coupling agents, resin and Fmoc amino acids were

purchased from Peptides International (Louisville, KY) or Aapptec (Louisville, KY). Solvents and other chemicals were purchased from Fisher Scientific and Sigma Aldrich. Analytical HPLC was preformed utilizing a Waters Sunfire C18 column 4.6 x 150 mm, 5 µm. Semi-preparative HPLC was preformed utilizing a Waters Sunfire Prep C18 OBD column 19 x 150 mm, 5 µm or a Waters Atlantis T3 Prep C18 OBD column, 19 x 150 mm, 5 µm. The gradient system utilized consists of acetonitrile + 0.1% TFA (Solvent A) and water + 0.1% TFA (Solvent B). Ultraviolet absorbance measurements were obtained utilizing a Waters 2998 Photodiode Array Detector and Mass spectra were obtained via electrospray ionization utilizing a Micromass Quatro Micro LCT mass spectrometer. MALDI mass spectrometry experiments were done utilizing an Applied Biosystems 4700 PDS. This instrument is equipped with an Nd: YAG laser tuned to a wavelength of 355 nm, with a 200 Hz laser rate operating in reflectron mode and tandem time-of-flight ion detectors in series. The mass spectrometer was calibrated with a mixture of angiotensin I (m/z 1296.685), Glu-1-Fibrino-peptide B (m/z 1570.677) to a mass accuracy of ± 0.5 Da.

3.4.2 General Peptide Synthesis

Standard Fmoc solid phase peptide synthesis was performed in order to obtain the desired peptide with side chain protecting groups. Rink amide MHBA NH_2 resin (0.1 mmol, 0.28 meq/g) containing an Fmoc protecting group was utilized. The resin was swelled with DMF and DCM washes. The Fmoc was removed with a solution 20% piperidine in DMF (2x, 5 min and 20 min, respectively). The residual deprotection solution was removed from resin using DMF and DCM washes. Fmoc-protected amino

acids (0.3 mmol, 3 eq) were dissolved in DMF (0.5 mL) along with HBTU or HCTU (0.3 mmol, 3 eq) and DIPEA (0.6 mmol, 6 eq). The amino acids used were Fmoc-Phe-OH, Fmoc-Thr(tBu)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Asp(OtBu)-OH, Fmoc-Ala-OH and Fmoc-Gln(Trt)-OH. The solution was then combined with the resin and agitated for one hour. The resin was then washed with DMF and DCM. This process was repeated as second time for one and a half hours. This deprotection/coupling process was repeated for each amino acid. Chelates were added using similar procedures.

3.4.3 Organometallic Coupling Procedure

*Compounds **3.2**, **3.3**, **3.4***

The chelate (0.4 mmol, 4 eq.) was dissolved in DMF (2 mL/1 mL for PADA (**2.7**)/BPA (**2.9**) and NTA (**2.10**), respectively) with HBTU or HCTU (0.4 mmol, 4 eq.) and DIPEA (0.8 mmol, 8 eq.) and mixed with peptide containing Rink Amide resin (0.1 mmol). The resin and solution were agitated overnight and subsequently washed with DMF and DCM resulting in the desired organometallic peptide.

*Compound **3.6***

The Rink Amide resin (0.1 mmol) and **2.13**, the PHA chelator (0.4 mmol, 4 eq.) in 2 mL of DMF with HCTU (0.4 mmol, 4 eq.) and DIPEA (0.8 mmol, 8 eq.) and mixed overnight. The resin was washed with DMF and DCM resulting in compound **3.5**. Then, $[Et_4N]_2[Re(CO)_3Br_3]$ (0.2 mmol, 2 eq.) in 1 ml of a 50:50 mixture of MeOH/DMF (v/v) with DIPEA (0.2 mmol, 2 eq.) is added to the resin containing compound **3.5**. The resin

and solution was mixed overnight. The resin was then washed with DMF and DCM resulting in compound **3.6**.

3.4.4 Peptide Deprotection, Cleavage and Purification

The organometallic peptides on resin were mixed with a solution of 95% TFA, 2.5% H_2O and 2.5% TIS (1 mL) for three hours to remove the peptide from resin as well as remove side-chain protecting groups. The resulting solution was cooled on ice and cold *tert*-butyl methyl ether (TBME) (1 mL) was added dropwise to the solution until a white precipitate formed. The solution was then centrifuged at 0^oC at 3000 rpm for 10 minutes. The liquid was decanted and the resulting pellet was mixed with cold TBME (1 mL) and subsequently centrifuged again under identical conditions to obtain the crude peptide.

The crude peptides were dissolved in water and purified by HPLC and identified by mass spectrometry, resulting in white powders.

*$Re(CO)_3$-PADA-QATDKFTF-NH_2 (**3.2**)*

The compound was synthesized via standard solid phase peptide synthesis techniques and purified by preparative HPLC (15 min linear gradient 30% solvent A in solvent B to 70%). The purity and identity of the compound was assessed by analytical HPLC and mass spectrometry, respectively. Purity: 98%. MS (MALDI) $[M+H]^+$ *m/z* calc: 1432.5 obs: 1432.5. (Rhenium-187)

*$Re(CO)_3$-BPA-QATDKFTF-NH_2 (**3.3**)*

The compound was synthesized via standard solid phase peptide synthesis techniques and purified by preparative HPLC (15 min linear gradient 20% solvent A in solvent B to 90%). The purity and identity of the compound was assessed by analytical HPLC and mass spectrometry, respectively. Purity: 93%. MS (MALDI) $[M]^+$ *m/z* calc: 1465.5 obs: 1465.6. (Rhenium-187)

*$Re(CO)_3$-NTA-QATDKFTF-NH_2 (**3.4**)*

The compound was synthesized via standard solid phase peptide synthesis techniques and purified by preparative HPLC with a gradient (15 min linear gradient 20% solvent A in solvent B to 90%). The purity and identity of the compound was assessed by analytical HPLC and mass spectrometry, respectively. Purity: 90%. MS (MALDI) $[M+2H]^+$ *m/z* calc: 1399.5 obs: 1399.5. (Rhenium-187)

*$Re(CO)_3$-PHA-QATDKFTF-NH_2 (**3.6**)*

The compound was synthesized via standard solid phase peptide synthesis techniques and purified by preparative HPLC with a gradient (15 min linear gradient 25% solvent A in solvent B to 80%). The purity and identity of the compound was assessed by analytical HPLC and mass spectrometry, respectively. Purity: 95%. MS (MALDI) $[M+H]^+$ *m/z* calc: 1480.5 obs: 1480.5. (Rhenium-187)

3.4.5 MALDI Tandem Mass Spectrometry

Purified organometallic peptides were dissolved in water at a concentration of approximately 0.02 g/mL. A portion of the organometallic peptide solution (2 µL) and a CHCA solution (2 µL) were mixed for a few minutes. CHCA solution is a mixture of 50:50 water/acetonitrile with the addition of 0.1% TFA (v/v) at a concentration of 5 mg/mL. A portion of the mixture (0.7 µL) was then spotted onto a MALDI plate in duplicate for analysis. The solvent in the mixture was allowed for evaporate resulting in the co-crystallization of the matrix and peptide. The samples were then analysed via MALDI tandem mass spectrometry. The precursor ion spectra show rhenium-185/187 isotopic distributions. The rhenium-185 peak was chosen for fragmentation via collision induced dissociation with argon gas. Tandem mass spectra were obtained and peaks assigned via manual *de novo* peptide analysis.

Chapter 4: The Synthesis and Screening of Organometallic OBOC Peptide Libraries

4.1 Introduction

4.1.1 OBOC Libraries

OBOC libraries can be used to screen for bioactive molecules against particular receptors, enzymes or other targets. In the previous chapter, MALDI tandem mass spectrometry has been shown to provide a useful and efficient method to determine the amino acid sequences of organometallic peptides. Mass spectrometry can then be applied to the sequencing of organometallic OBOC peptide libraries. These libraries have the potential to be a powerful tool in developing targeted peptide-based imaging agents.

Traditionally, OBOC libraries have not been popular in developing imaging agents for a number of reasons. OBOC libraries contain a large number of beads, each single bead contains many copies of the exact same compound. Peptides are thought to sometimes be present in unusual conformations in OBOC libraries due to the high concentration of peptide on a single bead.[94] Also, the presence of the polystyrene backbone of the bead can also have a large effect on the conformations of peptides. Therefore, bioactive peptide hits need to be re-evaluated in solution after the library has been screened. Furthermore, due to the unnaturally high concentration of peptide on library beads, false positive hits also can be a problem because of interactions between peptides and the many macromolecules present on cell surfaces. In addition, the initial formulation of an OBOC screening method was heavily reliant on manual separation of beads for analysis.[95] This process was inherently slow and low-throughput, further discouraging its use. However, with the application of fluorescence sorting and the

development of high-throughput mass spectrometry these problems can largely be overcome in order to screen libraries efficiently and economically.

The main advantage to utilizing OBOC libraries is the ability to create a large diverse library in which many compounds can be tested against targets. Moreover, OBOC libraries are built through chemical synthesis, and as such, can include a large variety of chemical scaffolds. Aside from naturally occurring amino acids, there are a wide variety of unnatural amino acids, D-amino acids, linkers and non-peptide components available. Natural peptides, in general, have short biological half-lives, and the addition or replacement of natural amino acids with non-natural mimics increases their stability. Because OBOC peptide libraries can include these non-natural mimics, it is a very appealing method for developing peptide-based imaging agents. Recently, there has been an interest in furthering the use of OBOC libraries in developing non-peptide or unnatural libraries.

Aditya *et al.* recently reported the synthesis and sequencing, through MALDI MS/MS, of an OBOC library containing peptide-like compounds with unnatural heterocycle side-chains (N-substituted oligo-glycines)[96] (Figure 4.1). These types of compounds are promising for therapeutic and diagnostic applications because they have similar properties to peptides; however they are resistant to biological degradation. The chemistry associated with synthesizing these compounds on the solid-phase is not fully developed and results in incomplete and impure libraries

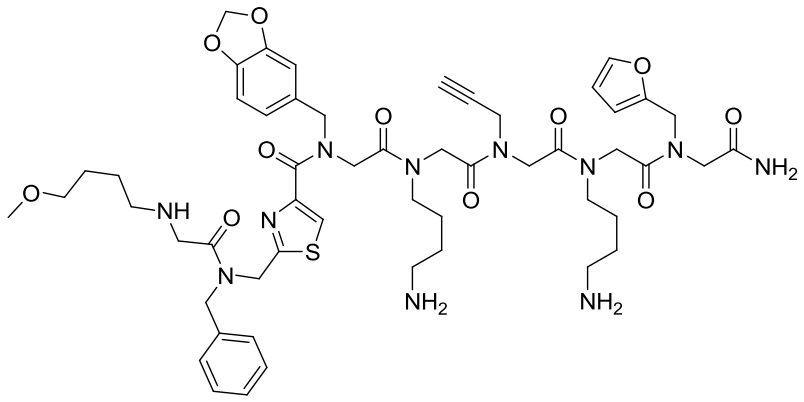


Figure 4.1 An example of the peptide-like compounds created in OBOC libraries by Aditya *et al.* These compounds are similar to peptides in structure, however they are expected to have long biological half-lives. These compounds were sequencing by MALDI TOF/TOF.

Another recent example of utilizing unnatural compounds has been shown by Goksel *et al.* in developing drugs against the androgen receptor. An OBOC peptide library containing unnatural amino acids was synthesized and tested for binding affinity. This resulted in a number of "hit" sequences deconvoluted through MALDI TOF ladder peptide synthesis. The best hit resulted in the peptide sequence SSRXESXXAGEKESRG, where X are β-(2-naphthyl)alanine, β-cyclohexylalanine and tryptophan, and the binding affinity measured is 0.47 μM.[97]

Recently, Bachor *et al.* have developed OBOC peptide libraries containing N-terminal quaternary ammonium salts. Peptides were reacted with iodoalkanes to form positively charged N-terminal residues. The purpose of these positively charged ammonium N-terminal residues is to provide increased sensitivity to mass spectrometric analysis of peptides,[98] and has resulted in the successful sequencing of these modified peptides. This methodology bears similarities to the development of organometallic

OBOC peptide libraries with rhenium(I) tricarbonyl chelates on the N-terminus of the peptide sequence.

Despite the increased use of OBOC libraries to develop peptide or peptide-like compounds, developing imaging agents via this method has been generally neglected due to the necessary modifications after screening. The proposed solution to this problem is to develop organometallic OBOC peptide libraries that take into account the effects of the imaging entity on the binding affinity of a peptide sequence.

4.1.2 Glucagon-Like Peptide 1 Receptor

The GLP-1R (glucagon-like peptide 1 receptor) is a promising target for developing imaging agents. The receptor is a member of GPCR (G Protein-Coupled Receptor) superfamily. This class of receptors is currently one of the most important drug targets in modern medicine as roughly half of all pharmaceuticals target this superfamily of receptors. The endogenous ligand of GLP-1R, glucagon-like peptide 1 (GLP-1), is a peptide hormone produced in the body which has roles in regulating blood glucose in response to nutrient digestion.[99] GLP-1, in its bioactive form, is a peptide containing 29 amino acids which is produced in intestinal L-cells. GLP-1 receptors are found largely in pancreatic beta cells which are responsible for insulin secretion. Because of this, the GLP-1R is an important target for treatment of type II diabetes. Furthermore, GLP-1R is also an important potential target for imaging of various insulinomas. These cancers often over-express the GLP-1R and a targeting imaging agent would be a very useful for diagnosis and monitoring.

Glucagon-like peptide 1 is produced through modification of the peptide prohormone proglucagon.[100] The products of the processing reactions result in a number of bioactive compounds, such as GLP-2 and glucagon, in addition to GLP-1 which usually exists as GLP-1(7-36) with a C-terminal amide. The binding interactions between GLP-1 and its corresponding receptor have been determined through alanine scans[101] and side-chain modifications[102] to identify the necessary amino acids for binding and the potential for imaging entity addition along the backbone, respectively. In addition, there have been reports of crystal structure determinations of the GLP-1R N-terminal extracellular domain with peptide antagonists at a resolution of 2.2 Å[103] and more recently 2.1 Å[104] to investigate peptide-receptor interactions. From these results, it can be shown the GLP-1R has two distinct binding motifs that GLP-1(7-36) amide interacts with simultaneously, with a binding affinity of approximately 0.16 nM.[105] This results in a large number of amino acids of GLP-1 being necessary for binding affinity to be maintained. The alanine scans, side chain modification studies and crystal structures have shown that not only are a large number of GLP-1(7-36) amino acids necessary for maintenance of binding affinity, these amino acids are also located along the entire length of the backbone; therefore, modifications are non-trivial exercises.

GLP-1, like most naturally occurring peptides, has been shown to have a short biological half-life of approximately two minutes. The poor biological stability of GLP-1(7-36) is due to the presence of L-alanine in position 8, which is susceptible to cleavage by the enzyme dipeptidyl peptidase-IV.[106] The replacement of the L-alanine with D-alanine or aminobutyric acid increases the stability of the peptide significantly, allowing GLP-1(7-36)amide to be a promising therapeutic and diagnostic pharmaceutical. A

number of GLP-1 analogues have or are currently undergoing clinical trials for treatment of diabetes, such as Victoza, which possesses a 97% sequence identity to GLP-1 with the addition of a fatty acid to the side chain of one amino acid.[107] In addition to GLP-1, exendin has also been shown to be a high affinity ligand of the GLP-1R. The exendin peptide hormone was isolated from the saliva of the lizard *Heloderma suspectum* in 1992.[108] The peptide possesses a 53% sequence identity to GLP-1,[109] has a significantly longer biological half-life and shows higher binding affinity to the GLP-1R. Currently, the peptide is used for treatment of diabetes through injections under the brand name Exenatide.

Despite the moderate success of GLP-1 and exendin analogues as therapeutic pharmaceuticals for treatment of diabetes, imaging agents have yet to be fully developed based on these peptide sequences. Initially, iodine-123 conjugates of GLP-1 and exendin-3 were developed as SPECT agents for detection of insulinomas. The results of these studies indicated that the accumulation of the molecular probe in the tumor was poor, suggesting that the iodine conjugates were not stable to biological conditions. Modification of the exendin-3 peptide at the lysine in the forty position with gallium-68 and indium-111 chelates led to the creation of [Lys^{40}(^{68}Ga-DOTA)]-exendin-3 and [Lys^{40}(^{111}In-DTPA)]-exendin-3. These imaging agents have shown selective uptake in subcutaneously insulinoma tumors implanted in mice.[110] However, there is also significant uptake in the kidneys most likely due to the positive charges present on lysine and arginine side chains. Furthermore, recent *in vitro* studies of a GLP-1 derivative with an AEEA (2-aminoethoxy-2-ethoxy acetic acid) linker, [Lys^{37}AEEA(^{111}In-DOTA)]GLP-

1 show high binding affinity to the GLP-1R.[111] However, similar problems with *in vivo* distribution of the imaging agent are observed.

Despite the development of promising imaging agents with high binding affinity to the GLP-1R developed based on the scaffolds of GLP-1 and exendin, there are problems associated with the biodistribution and specificity of the imaging agents. GLP-1 and exendin represent some common pharmacokinetic problems associated with using large peptides as targeting entities. Therefore, it is beneficial to develop a new class of peptide-based imaging agents for the GLP-1R that are not based on the structure of the endogenous ligands and have a significantly reduced size. The development of organometallic OBOC peptide libraries provides a potential method to overcome the limitations associated with endogenous ligand scaffolds in discovering small peptide molecules with high affinity.

The nature of the active site of the GLP-1R requires interactions between the ligand and the receptor at a variety of spatially distant points. This results in the GLP-1 having a large number of amino acids that are considered necessary for binding and that are located through the entire backbone of the peptide hormone. As such, it is unlikely that small peptide-based imaging agents will interact with the active site. Therefore, the library screening of small peptides might result in allosteric interactions between the ligand and the target. Family B of the GPCR superfamily, of which GLP-1R is a part, has been shown to be resistant to the development of compounds showing allosteric interactions.[112] There are few reports of compounds showing allosteric interactions with the GLP-1R, all of which are small molecules (Figure 4.2)

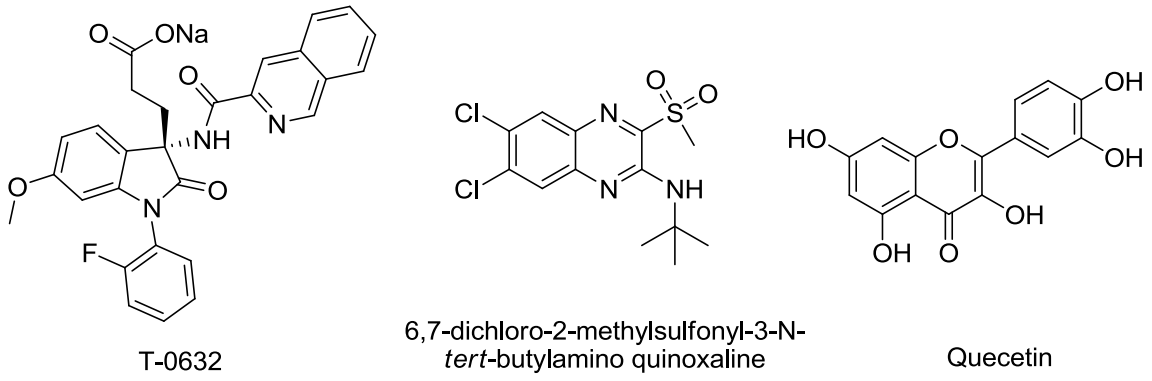


Figure 4.2 Allosteric small molecule antagonists of the GLP-1R.

The first molecule, T-0632, discovered to allosterically block peptide activity at the GLP-1R was reported in 2001.[113] The compound has an observed IC_{50} = 1.2 μm ± 0.4 μm. 6,7-dichloro-2-methylsulfonyl-3-N-*tert*-butylamino quinoxaline and Quecetin were investigated in 2010,[114] the former was previously found to be an allosteric modulator with a binding affinity of 0.1 μm ± 0.2 μm.[115] Their binding properties were investigated and found that they interact differently with the GLP-1R as they do not produce the same effects on cell pathways such as cAMP signal pathway stimulation and intracellular Ca^{2+} levels.

The binding interactions of these small molecule inhibitors are not well understood as the binding mechanism and binding site are not known. Nevertheless, these few compounds demonstrate the ability to discover compounds which can interact with GLP-1R in an allosteric fashion. Fortunately, knowledge of allosteric binding interactions is not a prerequisite to developing allosteric-based imaging agents. As long as a potential imaging agent binds to the GLP-1R with high binding affinity and specificity, the particular binding interactions do not necessarily need to be fully

understood, because imaging agents are introduced to the body in small quantities in comparison to therapeutic pharmaceuticals. Therefore, the affect of the binding of the imaging agent on the signalling pathway cascades are negligible.

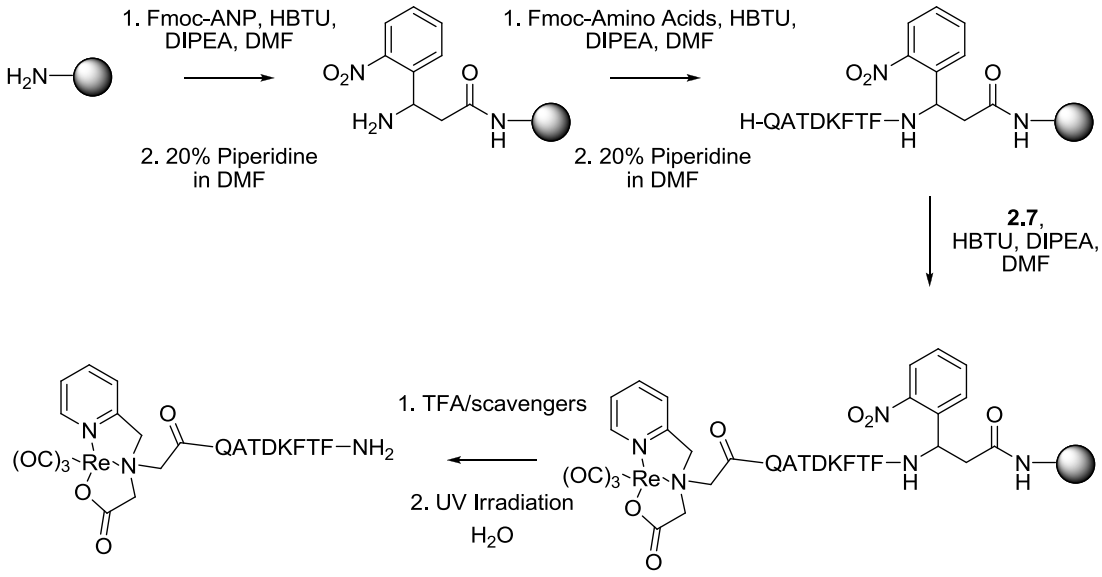
The development of allosteric GLP-1R imaging agents provides means to overcome the difficulty of utilizing the scaffolds of GLP-1 and exendin. These large peptides have poor pharmacokinetic properties; comparatively, small peptides, do not have these problems, in general. Because of the properties of the GLP-1R binding site, small peptides would be expected to interact through allosteric means. Therefore, the use of organometallic OBOC peptide libraries can be utilized to discover small peptides that interact with the GLP-1R.

4.2 Results and Discussion

4.2.1 Single bead MALDI Tandem Mass Spectrometry

The use of tandem time-of-flight mass spectrometry to sequence rhenium(I) tricarbonyl-containing peptides that were previously synthesized and purified in bulk quantities, has been shown to be an effective method for amino acid determination. However, these conditions are not indicative of the methods used in OBOC screening. Different resin, different cleavage techniques, lower quantities and lack of purification are all conditions of OBOC library screening that need to be addressed. Therefore, the first step in creating an organometallic OBOC peptide library is to verify the ability to sequence OBOC beads under library screening conditions.

The use of tentagel S NH_2 resin is necessary for OBOC peptide libraries. The difference between Rink Amide and tentagel resin is based on the functional groups to which the peptides are attached. Tentagel contains a free amine, and as such forms an amide bond when peptides are attached. Therefore, the acidic cleavage conditions which remove side-chain protecting groups, do not remove the peptide from the resin. This is important because the peptides need to be deprotected but still remain on the resin in order to be screened against GLP-1R. Then, in order to sequence the peptide through MALDI MS/MS, the peptide needs to be cleaved through the use of an ANP (2-amino-3-(2-nitrophenyl)propanoic acid) photolinker (Scheme 4.1).



Scheme 4.1 The synthesis of compound **3.2** on tentagel S NH_2 resin with a UV cleavable ANP photolinker. The organometallic peptide was cleaved from a single bead.

First the Fmoc-ANP photolinker was added to the resin under standard coupling conditions, and subsequently deprotected. In the absence of light, Fmoc-amino acids were added sequentially resulting in the peptide H-QATDKFTF. The organometallic was added under standard conditions. TFA/scavengers are then used to deprotect the side-chains of the amino acids with the peptide remaining on the resin. This was verified by the lack of observable peptide in the TFA/scavengers solution after three hours of side-chain cleavage. Subsequently, under a fluorescence microscope, a single bead was selected and placed in approximately 50 μL of water and cleaved through the use of UV light for three hours, resulting in a dilute solution of organometallic peptide. These conditions are identical to the conditions that an OBOC library would be screened. This produces a solution of compound **3.2** that is mixed with a CHCA solution and MALDI tandem mass spectrometry is then used to sequence the peptide. The resultant tandem mass spectrum is compared to the bulk-purified peptide results (Figure 4.3)

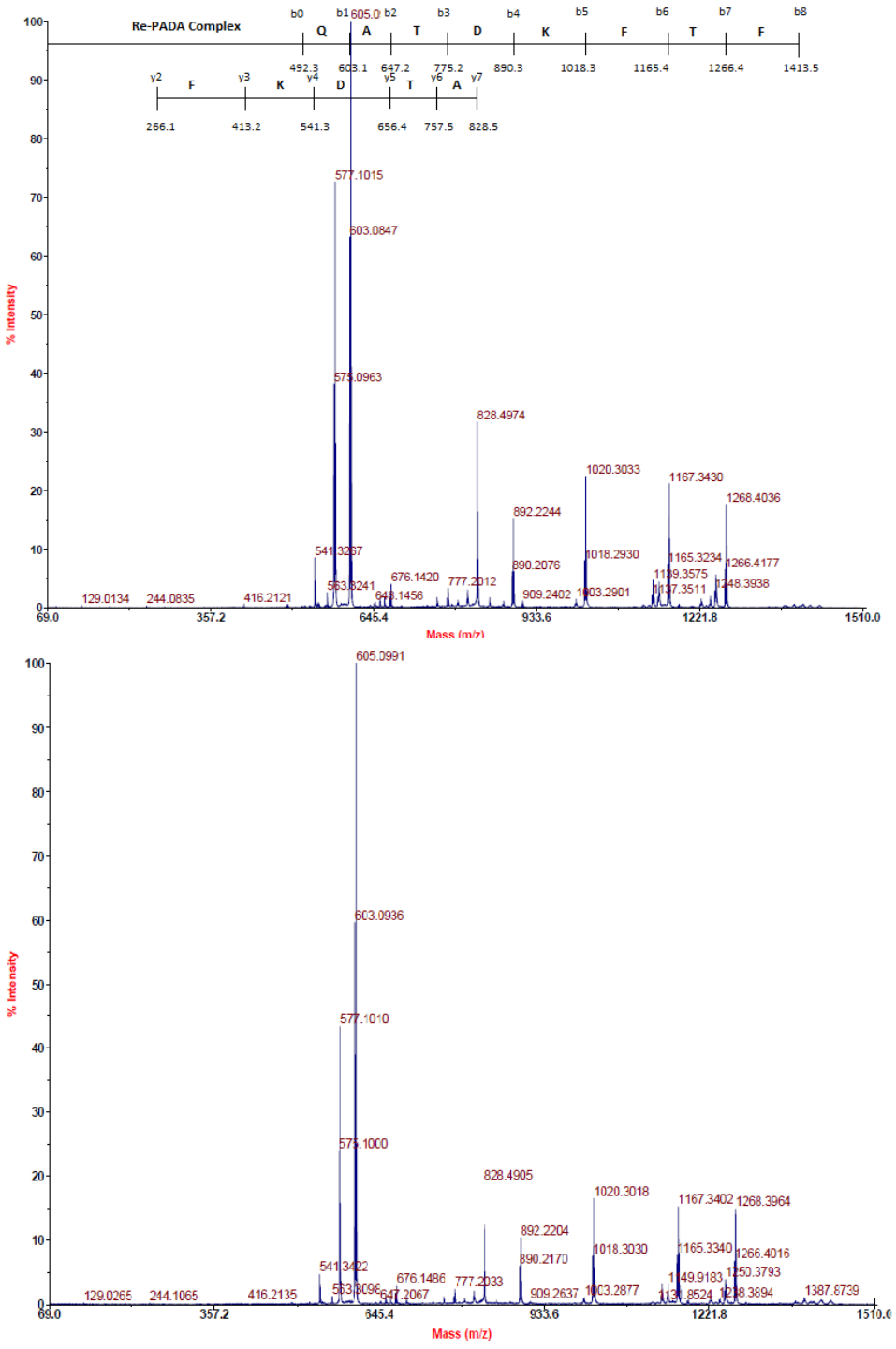


Figure 4.3 A comparison of the product ion spectra of compound **3.2** obtained by MALDI TOF/TOF mass spectrometry. a) The spectrum of the HPLC purified organometallic peptide. b) The spectrum of the organometallic peptide cleaved from a single 90 µm tentagel bead under OBOC screening conditions. The spectra are very similar and contain identical peaks.

The product ion mass spectrum obtained is almost identical to the purified peptide spectrum. All of the same peaks are present in the single-bead spectrum and the HPLC purified spectrum. This provides good evidence for the applicability of MALDI tandem mass spectrometry to organometallic OBOC peptide library screening.

4.2.2 Organometallic OBOC Peptide Library Synthesis

Because of the success of sequencing an organometallic peptide from a single tentagel bead with an ANP linker, a library of $Re(CO)_3$-PADA organometallic peptides was created to be screened against the GLP-1R. All natural amino acids were used, except cysteine, methionine and isoleucine. Cysteine forms disulfide bonds between peptide chains that would result in difficulty in sequencing procedures. Methionine is not stable to UV cleavage conditions and is prone to oxidation. Isoleucine is isobaric with leucine and they cannot be distinguished from each other by mass spectrometry. Therefore, seventeen amino acids are used to create peptides with eight amino acids. This results in a library of almost seven billion (17^8) possible peptide sequences. However, in order to create a library with that many peptide sequences, kilograms of resin would be required; this is neither economical nor practical. Only 1 g of beads is used in the library resulting in approximately 1.5 million unique peptide sequences.

Beads are coupled to Fmoc-ANP-OH and subsequently separated into 17 different wells, each corresponding to a different amino acid. Using an automatic peptide synthesizer, each bead is deprotected and a unique peptide is added to each well. The beads from each well are then recombined, mixed together, and then split back into the

wells. This iterative cycle is repeated eight times until a library of beads, each with a unique peptide sequence, was created. Initially, utilizing three equivalents of amino acid, the peptide coupling was observed to be incomplete, resulting in the observation of multiple different peptides present on a single bead. Increasing the proportion of amino acids to five equivalents, as well as the coupling time, resulted in complete coupling as observed by mass spectrometry. The beads are then recombined and then the $Re(CO)_3$-PADA complex is added to the N-terminus of the peptide. Four equivalents of the rhenium(I) tricarbonyl complex are coupled per peptide chain, resulting in complete conversion. Five millilitres of a mixture of TFA/H_2O/TIS was added to the beads in order to remove the side-chain protecting groups of the amino acids. The purity of the library is accessed through MALDI mass spectrometry, in which only one compound is observed for each single bead. The resultant beads contain a unique organometallic peptide connected to the bead through an ANP photolinker. The organometallic peptide contains no protecting groups and can be screened against the GLP-1R target.

4.2.3 Single Bead Organometallic OBOC Peptide Deconvolution

Single beads from the library were selected from the library in order to determine the ability of MALDI MS/MS to sequence unknown organometallic octapeptides. Previously, rhenium(I) tricarbonyl peptide sequences have been sequenced through this method, however the identity of the peptide is already known. The *de novo* sequencing of unknown organometallic peptides is significantly more difficult in comparison because peaks cannot be expected to be present. As such, single beads from the library were

individually selected via the use of a fluorescence microscope. Then, beads were irradiated with UV light for three hours in approximately 50 µL of water. The resultant organometallic peptide sequences were co-crystallized with CHCA and MALDI tandem spectra were obtained and sequenced (Table 4.1).

Table 4.1 Deconvoluted PADA-$Re(CO)_3$ peptides from an organometallic OBOC peptide library sequenced through manual *de novo* methods.

Compound	Deconvoluted Organometallic Peptide Sequence	Calculated Mass (*m/z*)	Observed Mass (*m/z*)
4.1	[PADA-$Re(CO)_3$]-VLYGVRPV-NH_2	1335.51	1335.49
4.2	[PADA-$Re(CO)_3$]-ALHEVGSW-NH_2	1373.46	1373.43
4.3	[PADA-$Re(CO)_3$]-HNES(K/Q)EGT-NH_2	1376.38	1376.43
4.4	[PADA-$Re(CO)_3$]-FEGGEESY-NH_2	1392.40	1392.37
4.5	[PADA-$Re(CO)_3$]-YARES(K/Q)AD-NH_2	1414.47	1414.45
4.6	[PADA-$Re(CO)_3$]-DHHLRAYV-NH_2	1443.48	1443.46
4.7	[PADA-$Re(CO)_3$]-TEHNP(K/Q)HE-NH_2	1466.44	1466.48

The observed organometallic peptides all contain eight amino acids, consistent with the library properties. The ability to sequence these peptides was aided by the rhenium isotopic signature present on N-terminal fragments, allowing the ability to distinguish between N and C-terminal fragments. Furthermore, the mass of the $Re(CO)_3$-PADA chelate is known and facilitates the sequencing by starting *de novo* sequencing at a universal *m/z* value. In addition, from the precursor spectrum, the total mass was also known. The resultant table of peptides shows calculated and observed mass in close agreement (Table 4.1). Lysine and glutamine are almost isobaric and therefore cannot be distinguished in mass spectrometry. Like $Re(CO)_3$-PADA-QATDKFTF-NH_2, (**3.2**), the most intense peak in almost all of the deconvoluted organometallic peptide sequences, with the exception of **4.5**, was the b_1 fragment, suggesting that this fragmentation pattern is in general the most favourable pathway for N-terminal PADA rhenium(I) tricarbonyl

chelates. Approximately 50% of beads sequenced by MALDI MS/MS resulted in complete assignments of amino acids. Many of the incomplete fragmentation patterns resulted in the majority of amino acids being assigned, with the omission of two amino acid identities. Despite the fact that not all of the peptide sequences could be identified, a fifty percent success rate is acceptable because the purpose of screening is to identify bioactive peptides and not to infer about structure-activity relationships.

4.2.4 Organometallic OBOC Peptide Library Screening Against the GLP-1R

In order to find bioactive peptide sequences that target the GLP-1R from the organometallic OBOC peptide library, fluorescent CHO (Chinese hamster ovarian) cells over-expressing the GLP-1R were created. These cells were then incubated with the OBOC library and the beads containing the bioactive organometallic peptide sequences had cells attached to them. In order to facilitate high throughput screening, after incubation, the beads are sorted utilizing a COPAS biosorter. The biosorter flows beads and lasers are used to measure the fluorescence, the size of the particle and the extinction (the amount of light which is blocked by the particle). The beads that have cells bound to them have a larger size and higher fluorescence. In this way, beads that have bioactive peptides can be separated from beads which do not have bioactive peptides due to the lack of peptide-cell interaction on the latter.

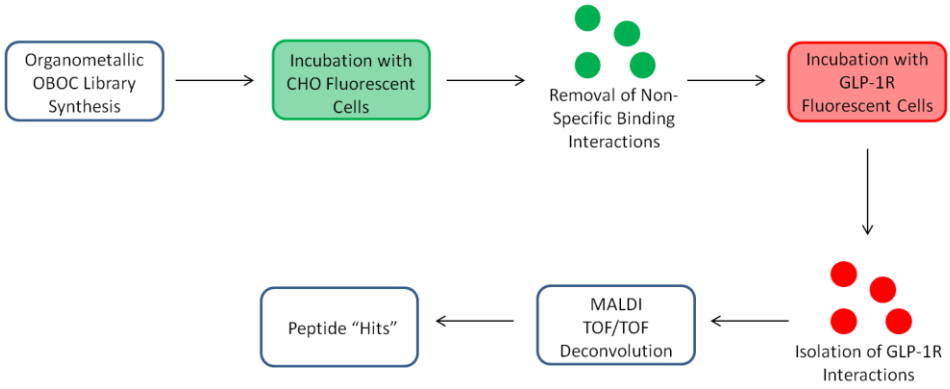


Figure 4.4 The procedure for screen an OBOC peptide library against cells. This technique accounts for non-specific binding interactions.

The process of screening an organometallic OBOC peptide library against the GLP-1R begins with the incubation of the library with CHO cells expressing the GFP (green fluorescent protein). These cells are green fluorescent and are used as a control to account for non-specific binding in the peptide library. Peptide sequences would be expected to interact with some of the thousands of biomolecules expressed on a cells' surface, and as such if the organometallic peptide sequences interact with these biomolecules they produce false positive results (Figure 4.5).

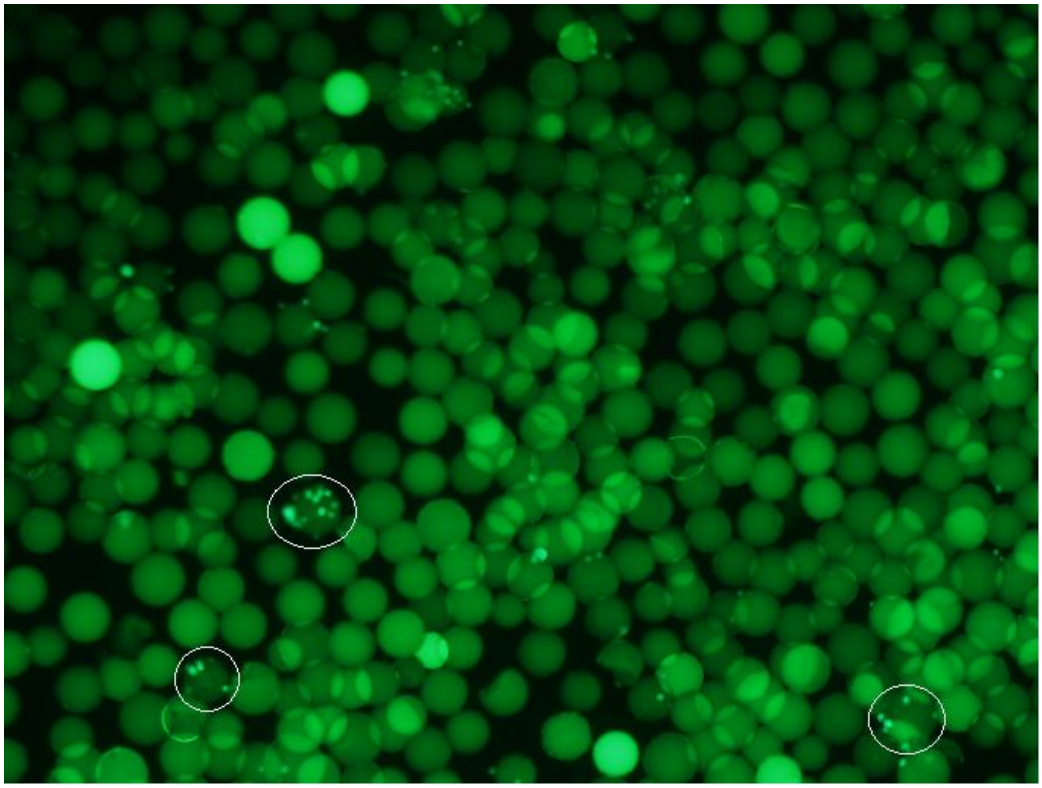


Figure 4.5 Organometallic OBOC peptide library after incubation with GFP expressing cells. There are a number of beads circled in white that show interaction with the cells while most do not. The intensity of the auto-fluorescence of the tentagel beads varies from bead to bead.

In order to account for these interactions, the incubation of the organometallic OBOC peptide library with CHO cells allows for the removal of the peptides showing these interactions. The beads that did not show interaction with the control green CHO cells were then incubated with red GLP-1R over-expressing CHO cells.

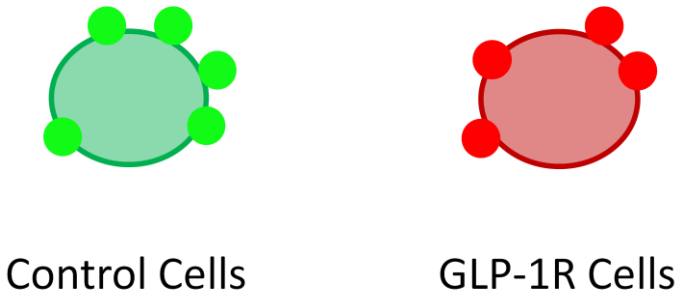


Figure 4.6 A diagram of CHO cells showing green and red fluorescence. The green cells express GFP (green fluorescent protein) and act a control for non-specific binding. The red cells express tdTomato and overexpress the GLP-1R.

The organometallic OBOC peptide library (1.5 million beads) was incubated with approximately 10 million CHO control cells suspended in 10 mL of PBS for one hour being mixed at 300 rpm at 37 ^{o}C. After incubation, formaldehyde was added to the solution in order to attach the cells to the beads permanently. The formaldehyde reacts with the biomolecules present on cell surface in order to create cross-linking. The beads were then rinsed with PBS three times in order to remove any of the unbound cells. The beads were then sorted via the COPAS biosorter. From the results, two distinct populations of the beads are shown in the library, the more fluorescent population contain control cells bound to the beads (Figure 4.7).

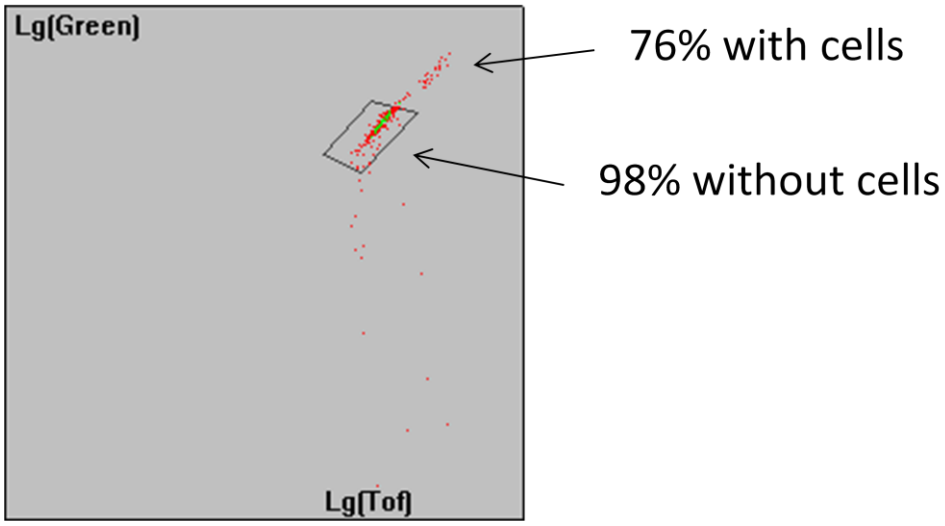


Figure 4.7 A graph from the COPAS biosorter showing two distinct populations of beads after incubation with green control cells. The y-axis is a measurement of green fluorescence and the x-axis is the time-of-flight or size. The top population shows increased fluorescence and contains GFP CHO cells with non-specific binding interactions. The bottom population contains no cells.

The population containing smaller, less fluorescent beads have not interacted with cells. After isolating these populations, investigation of the beads under a fluorescent microscope revealed that most of the bottom population have no cells bound to them and most of the more fluorescent population contain cell-bead interactions (Figure 4.8).

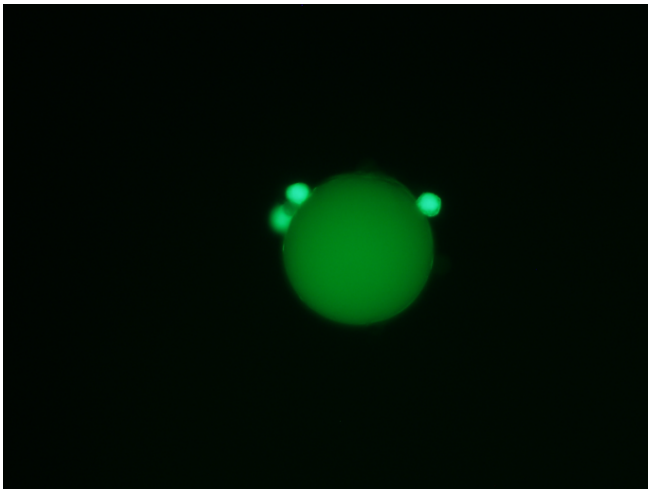


Figure 4.8 Tentagel bead interacting with three visible GFP expressing cells.

After the non-interacting bead population was isolated, the beads were incubated with red GLP-1R over-expressing cells utilizing the same procedure as outlined above. The biosorter was once again used to separate two populations of beads, those showing and not showing interactions with the GLP-1R. In this case, the beads showing interactions with the cells were to be isolated to find peptide sequences interacting with the GLP-1R. However, there were problems associated with bead sorting based on red fluorescence on the COPAS biosorter that could not be overcome to produce conditions to find two separate populations. The auto-fluorescence of the beads in the red is not as uniform as it is in green, and the variation in red fluorescence masks the bioactive beads.

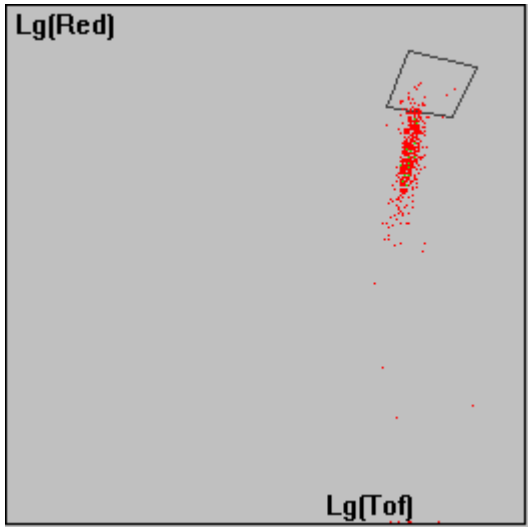


Figure 4.9 A graph from the COPAS biosorter with the y-axis is a measurement of red fluorescence and the x-axis is time-of-flight or size. This graph illustrates that there are not two distinct populations of beads containing red cells and those without.

Therefore, because of the difficulties associated with screening red fluorescence, manual separation under a fluorescence microscope was utilized. The beads were isolated from each other via the identification of red cells bound to beads and picking the beads out with a pipette. This process is time consuming, technically challenging and not inherently high throughput. Approximately 20 beads were isolated via this method. The organometallic peptides were cleaved from resin utilizing a UV light for three hours and MALDI was subsequently preformed on the resultant solution.

4.2.5 MALDI TOF/TOF Deconvolution of "Hit" Library Beads

The MALDI tandem mass spectrometry results of the 20 beads produced mass spectra with minimal signals. Most of the beads showed no signal in the corresponding

mass spectrum. It is suspected that the time-consuming isolation of beads under the fluorescent microscope resulted in the cleavage of much of the organometallic peptides from the resin. Light is shone on the beads in order to induce fluorescence, the ANP photolinker was cleaved and peptide removed from resin. Unfortunately, out of the 20 isolated beads, only two produce a precursor ion spectrum. Of these two peptide hits, only one of their corresponding product ion spectra could be fully deconvoluted via *de novo* sequencing. The organometallic peptide was identified as $Re(CO)_3$-PADA-RVTH(K/Q)(K/Q)(K/Q)E-NH_2 (**4.8**) through the use of *b* ion fragments (Figure 4.10).

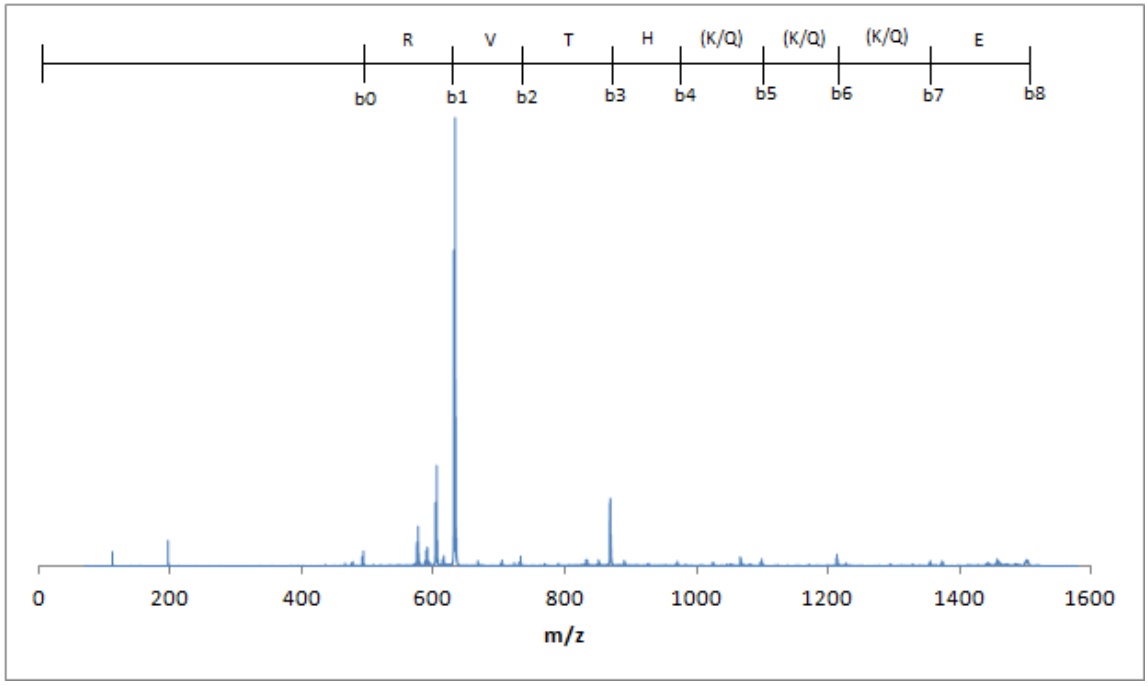


Figure 4.10 The product ion spectrum of the organometallic octapeptide **4.8** showing the *b* series fragments.

The product ion spectrum of compound **4.8** shows similar properties to the previously investigated $Re(CO)_3$-PADA organometallic peptides. The b_1 fragment is by

far the most intense, as was the case with every previous $Re(CO)_3$-PADA peptide except for compound **4.5**. The compound contains three positions which have been assigned either lysine or glutamine. As such, the actual identity of the peptide is one of eight possible peptides.

Table 4.2 Fragment assignments from series *b* of the product ion spectrum of compound **4.8**. The masses of the b_8 ion can be determined because of the known mass of the peptide and the N-terminal chelate. $[M+H]^+$ *m/z* = 1497.52 (rhenium-185 peak).

Label	**Fragment**	**Peak (*m/z*)**	**Difference between peaks (*m/z*)**	**Amino Acid Assignment**
b_0	RE-PADA-NH_3	492.0		
b_1	RE-PADA-R	631.1	156.1	Arginine
b_2	RE-PADA-RV	730.2	99.1	Valine
b_3	RE-PADA-RVT	831.3	101.1	Threonine
b_4	RE-PADA-RVTH	968.3	137.1	Histidine
b_5	RE-PADA-RVTH(K/Q)	1096.4	128.1	Lysine/Glutamine
b_6	RE-PADA-RVTH(K/Q)(K/Q)	1224.4	128.0	Lysine/Glutamine
b_7	RE-PADA-RVTH(K/Q)(K/Q)(K/Q)	1352.5	128.1	Lysine/Glutamine
b_8	NOT OBSERVED	1481.5	129.0	Glutamic Acid

The other peptide "hit" was partially sequenced via *de novo* methods. The identity of the peptide sequence is Re(CO)3-PADA-SWGxxLFS-NH_2 (**4.9**). The b_1 fragment is again the most intense peak in the spectrum (Figure 4.11). Two positions could not be identified due to lack of fragmentation at that position.

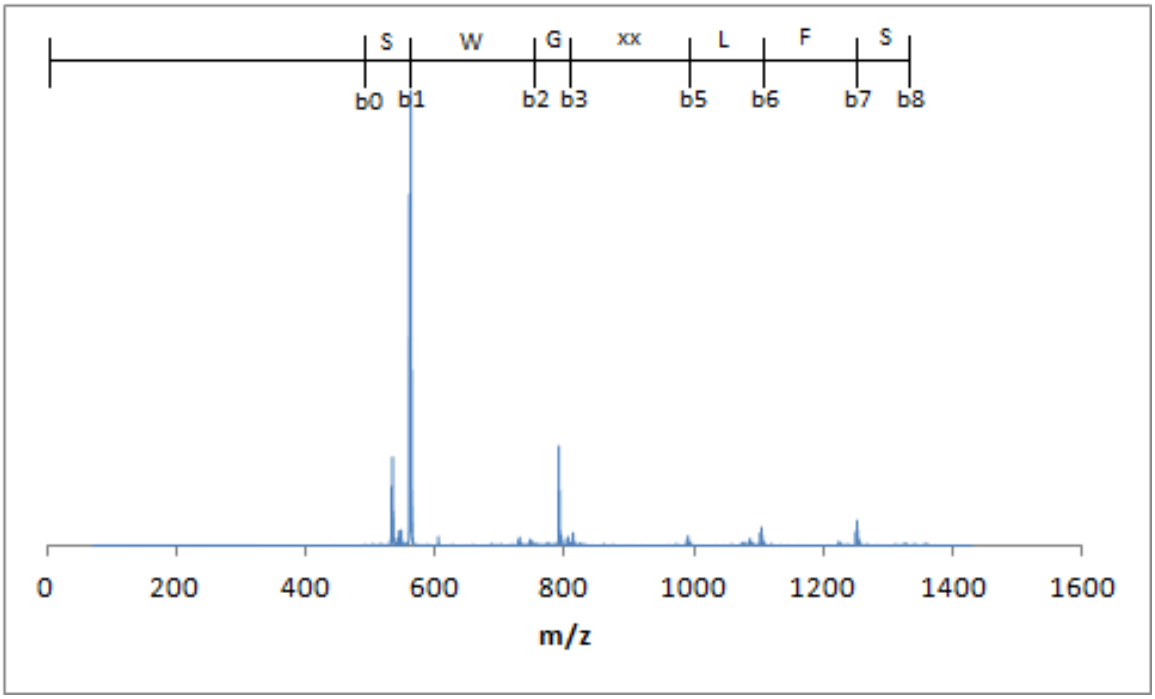


Figure 4.11 The product ion spectrum of the organometallic octapeptide **4.9** showing the *b* series fragments.

The mass allotted to the unknown dipeptide (xx) is approximately *m/z* 184. Based on the known masses of amino acids, the identity of these two amino acids could be either a combination of proline and serine or alanine and leucine. Additionally, the order of the two amino acids is not known. This highlights the limitations associated with *de novo* sequencing. Despite the fact that eight of nine *b* series fragments were determined, the identity of the peptide sequence could only be narrowed down to four potential sequences.

Table 4.3 Fragment assignments from series *b* and *y* of the product ion spectrum of compound **4.9**. The masses of the b_8 ion can be determined because of the known mass of the peptide and the N-terminal chelate. $[M+H]^+$ *m/z* = 1353.48 (rhenium-185 peak).

Series	Label	Fragment	Peak (*m/z*)	Difference between peaks (*m/z*)	Amino Acid Assignment
b series	b_0	NOT OBSERVED	475.0		
	b_1	RE-PADA-S	562.0	87.0	Serine
	b_2	RE-PADA-SW	748.1	186.1	Tryptophan
	b_3	RE-PADA-SWG	805.1	57.0	Glycine
	b_4	RE-PADA-SWGx			
	b_5	RE-PADA-SWGxx	989.3		
	b_6	RE-PADA-SWGxxL	1102.3	113.1	Leucine
	b_7	RE-PADA-SWGxxLF	1249.4	147.1	Phenylalanine
	b_8	NOT OBSERVED	1336.4	87.0	Serine

From the information given by the two organometallic peptide sequences, the two sequences do not appear to be similar. Compound **4.8** is a very positively charged peptide species, containing an arginine, histidine and potentially as many as three lysine residues, all of which are capable of holding a positive charge. Conversely, **4.9** is a largely hydrophobic peptide sequence containing phenylalanine, leucine, glycine and tryptophan with hydroxyl group of serine being the only polar functionality within the organometallic peptide. Therefore, it is difficult to draw any conclusions about the binding properties of organometallic octapeptides. In order to determine patterns among this class of bioactive molecule a larger sample of peptide "hit" sequences is necessary. Furthermore, compounds **4.8** and **4.9** need to be evaluated in solution phase in order to

determine their quantitative binding affinities in order to determine their true ability to interact with GLP-1R.

From the information given by the two organometallic peptide sequences, the two sequences do not appear to be similar. Compound **4.8** is a very positively charged peptide species, containing an arginine, histidine and potentially as many as three lysine residues, all of which are capable of holding a positive charge. Conversely, **4.9** is a largely hydrophobic peptide sequence containing phenylalanine, leucine, glycine and tryptophan with hydroxyl group of serine being the only polar functionality within the organometallic peptide. Therefore, it is difficult to draw any conclusions about the binding properties of organometallic octapeptides. In order to determine patterns among this class of bioactive molecule a larger sample of peptide "hit" sequences is necessary. Furthermore, compounds **4.8** and **4.9** need to be evaluated in solution phase in order to determine their quantitative binding affinities in order to determine their true ability to interact with GLP-1R.

4.3 Conclusions

In this chapter, an organometallic OBOC peptide library was synthesized containing octapeptides with an N-terminal PADA-rhenium(I) tricarbonyl complex. Random beads from the library were selected and MALDI tandem mass spectrometry was used to determine the identity of the peptide sequences on the bead. A fifty percent success rate was obtained in the sequencing of the unknown peptide sequences. These

results have demonstrated that rhenium(I) tricarbonyl-containing peptide sequences can be deconvoluted using MALDI tandem mass spectrometry.

The library was then screened against cells over-expressing the GLP-1R in order to identify bioactive organometallic peptide sequences. Two such peptide sequences were identified, although the use of a fluorescence microscope was resorted to as a method to identify fluorescent beads due to failure to find proper sorting conditions for red cells. This is the first reported example of an organometallic OBOC peptide library screened against a biological target for bioactive molecules. Two such sequences are not enough data to determine patterns and infer relationships between the structure of the ligand and its correlation with the binding affinity. These sequences need to be further evaluated in solution phase to determine their quantitative binding affinity.

4.4 Experimental

4.4.1 General Procedures

All chemicals and solvents were purchased from commercially available sources and used without further purification. Coupling agents, resin and Fmoc amino acids were purchased from Peptides International (Louisville, KY) or Aapptec (Louisville, KY). Solvents and other chemicals were purchased from Fisher Scientific and Sigma Aldrich. Automated peptide synthesis was done using an Advanced ChemTech Apex Multiple Peptide Synthesizer. Bead sorting was accomplished using a Union Biometrica COPAS BIOSORT instrument. Fluorescence was investigated using a fluorescence microscope. MALDI mass spectrometry experiments were done utilizing an Applied Biosystems 4700

PDS. This instrument is equipped with an Nd: YAG laser tuned to a wavelength of 355 nm, with a 200 Hz laser rate operating in reflectron mode and tandem time-of-flight ion detectors in series. The mass spectrometer was calibrated with a mixture of angiotensin I (m/z 1296.685), Glu-1-Fibrino-peptide B (m/z 1570.677) to a mass accuracy of ± 0.5 Da.

4.4.2 Single Bead Organometallic Peptide Synthesis

$Re(CO)_3$-QATDKFTF-NH_2 (**3.2**)

The peptide synthesis was carried out on tentagel S NH_2 resin (0.1 mmol) and using standard Fmoc peptide synthesis. Fmoc-ANP-OH (0.4 mmol, 4 eq.) was dissolved in DMF (1mL) and combined with HCTU (0.4 mmol, 4 eq.) and DIPEA (0.8 mmol, 8 eq.). The mixture was agitated for three hours and the process repeated once. The peptide was then deprotected using 20% piperidine in DMF (v/v) and the amino acids were coupled sequentially. **2.7** (0.4 mmol, 4 eq.) was dissolved in DMF (2 mL) with HCTU (0.4 mmol, 4 eq.) and DIPEA (0.8 mmol, 8 eq.) and mixed with peptide containing the resin. The resin and solution were agitated overnight and subsequently washed with DMF and DCM resulting in the desired organometallic peptide. The organometallic peptide was cleaved from resin by selecting a single bead under a fluorescence microscope and irradiated with UV light for three hours in 50 µL of water. The solution was then sequenced via MALDI TOF/TOF.

4.4.3 OBOC Library Synthesis

OBOC peptide libraries were synthesizer utilizing the Advanced ChemTech Multiple Peptide Synthesizer. The libraries were created with the seventeen natural amino acids with the exclusion of methionine, cysteine and isoleucine. Tentagel S NH_2 resin (1.30 g, 0.33 mmol) with a loading of 0.26 mmol/g. Fmoc-ANP-OH (1.42 g, 3.30 mmol, 10 eq.) was coupled to the resin using standard double coupling Fmoc peptide synthesis procedures. The resin was then kept in the dark for the rest of the procedure. The resin was divided into seventeen equal portions (0.07 g, 0.02 mmol) and placed in a single well. The Fmoc was removed using 20% piperidine in DMF (v/v) (800 µL/well) repeated twice and to each well a different amino acid (0.10 mmol, 5 eq.) was coupled using HCTU (0.04 g, 0.10 mmol, 5 eq.) and DIPEA (0.035 mL, 0.20 mmol, 10 eq.) in DMF (2.4 mL), rinsed and repeated. The resin was rinsed in DCM and recombined in a beaker. The resin was mixed thoroughly and divided into seventeen equal portions and placed back in the wells. The process of deprotection and coupling of amino acids was repeated seven more times creating a library of octapeptides. The resin was removed from the wells, the Fmoc was removed and coupled with compound **2.7**, $Re(CO)_3$-PADA (0.51 g, 1.0 mmol, 3 eq.), HCTU (0.41 g, 1.0 mmol, 3 eq) and DIPEA (0.350 mL, 2.0 mmol, 6 eq.) in DMF overnight. This results in an organometallic $Re(CO)_3$-PADA octapeptide OBOC library. The purity of the library was verified by MALDI MS.

4.4.4 Cell Growth

The CHO-GLP-1R/tdtomato cells were grown in F-12K media/10% FBS plus 200 µg/ml G418 and 1 mg/mL hygromycin in order to maintain the stable expression of GLP-1R and tdtomato, respectively. The CHO-EGFP cells were grown in F-12K media/10% FBS plus 1 mg/mL G418 in order to maintain stable EGFP expression.

4.4.5 Library Incubation

The library was washed with dichloromethane, methanol and three times with aqueous PBS (phosphate buffered saline). The cells were dissociated with 2 mM EDTA, collected, spun down and then resuspended in HBSS (Hank's buffered saline solution) at density of 2 x 10^6 cells/mL. 20 x 10^6 cells were plated in a 10 cm Petri dish and the beads (approximately 1.5 million beads) were added to the dish. The dish was agitated at 70 rpm for 1 hour, in a humidified 37 ^oC incubator. Paraformaldehyde was added to the dish to a final concentration of 4 % and incubated at room temperature for 10 minutes and placed in a 40 mL falcon tube. The beads settled at the bottom and the cell-bead pellet was washed twice with PBS.

4.4.6 Library Screening

The library, having previously been incubated with CHO cells, was screened using a Union Biometrica COPAS biosorter. The library was suspended in PBS solution

and mixed thoroughly. The biosorter measured the size and green fluorescence and separated the beads into two populations, one with fluorescent cells and one without. This was verified by investigation under the fluorescence microscope. The population without the cells was washed with methanol and PBS. The population was then incubated again with CHO cells over-expressing the GLP-1R and expressing the red fluorescent protein tdTomato. The incubation proceeded through an identical method outlined above. An inverted fluorescent microscope was used to separate the beads containing red fluorescent cells.

4.4.7 MALDI Tandem Mass Spectrometry

Beads were isolated via use of a fluorescent microscope and the beads were placed in 50 µL of water and then irradiated with UV light for three hours. An aliquot of the organometallic peptide solution (2 µL) and a CHCA solution (2 µL) were mixed for a few minutes. CHCA solution is a mixture of 50:50 water/acetonitrile (v/v) with the addition of 0.1% TFA at a concentration of 5 mg/mL. A portion of the mixture (0.7 µL) was then spotted onto a MALDI plate in duplicate for analysis. The solvent in the mixture was allowed to evaporate resulting in the co-crystallization of the matrix and peptide. The samples were then analyzed via MALDI tandem mass spectrometry. The precursor ion spectra show rhenium 185/187 isotopic distributions. The rhenium-185 peak was chosen for fragmentation via collision-induced dissociation with argon gas. Tandem mass spectra were obtained and peaks assigned via manual *de novo* analysis.

Chapter 5: Conclusions

Molecular imaging has become an important part of modern medicine due to its ability to obtain information assisting in diagnosis in a non-invasive manner.[1] The development of new nuclear imaging technology relies on developments from a wide variety of scientific and medical fields; the production and accessibility of radioisotopes, knowledge of potential biological targets and the development of new imaging agents, among many others, are all important in furthering the applicability and use of nuclear medicine.

The research in this dissertation utilizes OBOC (one-bead, one-compound) combinatorial peptide libraries to further the development of imaging agents by the incorporation of organometallic coordination compounds into the library. Synthesis of rhenium(I) tricarbonyl chelates as non-radioactive surrogates for technetium-99m, the most widely used medical radioisotope, is discussed as well as the incorporation of the chelates into peptide backbones. The ability to sequence rhenium(I) tricarbonyl-containing peptides through MALDI tandem mass spectrometry has been successful for all chelates investigated. Furthermore, an organometallic OBOC peptide library was applied to the discovery of new GLP-1R (glucagon-like peptide 1 receptor) candidate imaging agents for diabetes and insulinomas.

Four rhenium(I) tricarbonyl chelates have been synthesized, PADA, BPA, NTA and PHA, and were attached directly to the N-terminus of a peptide chain. These chelates were chosen because they provide a good representation of the variety of chelates associated with the rhenium(I) tricarbonyl metal centre. $Re(CO)_3$-PADA, is known to be

a stable chelate[52] that forms neutral complexes. $Re(CO)_3$-BPA and $Re(CO)_3$-NTA provides positively and negatively charged coordination compounds, while $Re(CO)_3$-PHA is a novel chelation system holding a neutral charge. An investigation into the ability to identify unknown organometallic amino acid sequences through the use of MALDI tandem mass spectrometry has been accomplished. The results of this study into organometallic peptide fragmentation patterns lead to the conclusion that a wide variety of peptide chelation systems can be successfully incorporated on the N-terminus of the peptide and these chains can be sequenced by MALDI TOF/TOF. All rhenium(I) tricarbonyl peptides were successfully sequenced by tandem mass spectrometry. $Re(CO)_3$-PADA and $Re(CO)_3$-BPA peptides provide intense and easily identifiable amino acid fragments. The fragmentation pattern of the $Re(CO)_3$-NTA peptide does not have complete fragmentation series, however was still successfully sequenced. The $Re(CO)_3$-PHA peptide is the first known example of a phenol based rhenium(I) tricarbonyl chelate bound to an amino acid chain. The fragmentation pattern of the $Re(CO)_3$-PHA peptide compound allowed for the successful identification of the peptide sequence, as well as the unique observation of the fragmentation of the PHA chelate resulting in the neutral loss of orthoquinone methides. The results of the MALDI TOF/TOF sequencing experiments indicate that a wide variety of rhenium(I) tricarbonyl chelates can be identified in this manner.

An N-terminal $Re(CO)_3$-PADA OBOC peptide library was synthesized containing approximately 1.5 million unique peptide sequences utilizing all natural amino acids with the exception of cystiene, isoleucine and methionine. The library was screened against fluorescent CHO (Chinese hamster ovarian) cells over-expressing the GLP-1 receptor in

order to identify imaging agent candidates for use in diabetes or insulinoma diagnosis and monitoring. Two organometallic peptide sequences, $Re(CO)_3$-PADA-RVTH(K/Q)(K/Q)(K/Q)E-NH_2 and $Re(CO)_3$-PADA-SWGxxLFS-NH_2, were isolated from the library and identified by tandem mass spectrometry. Cells were attached to the beads that contained these two organometallic peptides. Because only two "hit" sequences were isolated, inferences into the structure-activity relationships cannot be conclusively made.

Recent developments of OBOC peptide libraries have focused on the incorporation of unnatural and non-peptide-like structures into combinatorial libraries. OBOC libraries are robust to a wide variety of chemistry, which is one of their main advantages in comparison to other combinatorial libraries. Additions of unnatural amino acids[97] and quaternary ammonium salts[98] have been recently reported. Moreover, OBOC peptoid (N-substituted oligo-glycines) libraries[96] have been developed. Additionally, the identification of peptoid structures through tandem mass spectrometry has also been shown. Non-natural peptide-like structures are generally more stable to *in vivo* conditions, which is an important property of pharmaceuticals and imaging agents. The research in this dissertation utilizes the same benefits of OBOC libraries to further utilize non-peptide structures to improve the development of nuclear imaging agents. Rhenium(I) tricarbonyl-containing OBOC peptide libraries have unnatural radionuclide surrogates in order to account for their influence on receptor binding affinity. This advance in molecular imaging has the ability to alter the way in which peptide-based imaging agents are developed. For the discovery of novel imaging agents, the organometallic OBOC peptide methodology outlined in this dissertation is superior to the

traditional OBOC approach. Furthermore, organometallic OBOC peptide libraries could be used to investigate the influence of a large number of different imaging entity surrogates on known bioactive peptides.

In conclusion, novel organometallic OBOC peptide libraries have been synthesized through standard peptide chemistry, MALDI tandem mass spectrometry has been identified as a useful technique to deconvolute organometallic peptide sequences and a $Re(CO)_3$-PADA peptide library has been used to identify potential imaging agents for the GLP-1R. This technique has the potential to be applied to any number of radionuclide surrogates. However, fragmentation patterns and chelate stability under fragmentation conditions would need to be investigated. Organometallic OBOC peptide libraries provide a new methodology to expedite and economize the process of peptide-based imaging agent discovery.

Chapter 6: References

1. Massoud, T.F., Gambhir, S.S. *Genes & Development*. 2003, **17**, 545-580
2. Gambhir, S.S. *Nat. Rev. Cancer*. 2002, **2**, 683-93
3. Bartholoma, M.D., Louie, A.S., Valliant, J.F., Zubieta, J. *Chem. Rev.* 2010, **110**, 2903-20.
4. Alauddin, M.M. *J. Nucl. Med. Mol. Imaging*. 2012, **2**, 55-76
5. Prats, E., Aisa, F., Abos, M.D., Villavieja, L., Garcia-Lopez F., Asenjo, M.J., Razola, P., Banzo, J. *J. Nucl. Med.* 1999, **40**, 296-301
6. Nickels, Pham, W. *IFMBE Proceedings*. 2010, **27**, 25-38
7. Pimlott, S.L., Sutherland, A. *Chem. Soc. Rev.* 2011, **40**, 149-62
8. Reshef, A., Shirvan, A., Akselrod-Ballin, A., Wall, A., Ziv, I. *J. Nucl. Med.* 2010, **51**, 837-40
9. James, M.L., Gambhir, S.S. *Physiol. Rev.* 2012, **92**, 897-965
10. McCabe, K.E., Wu, A.M. *Cancer Biother. Radiopharm.* 2010, **25**, 253-61
11. Wu, A.M., Olafsen, T. Cancer J. 2008, 14, 191-7
12. Wu, A.M. *J. Nucl. Med.* 2009, **50**, 2-5
13. Okarvi, S.M. *Eur. J. Nucl. Med.* 2001, **28**, 929-38
14. Vanhee, P., van der Sloot, A.M., Verschueren, E., Serrano, L., Rousseau, F., Schymkowitz, J. *Trends in Biotechnology*. 2011, **29**, 231-9
15. Lee, S., Xie, J., Chen, X. *Chem. Rev.* 2010, **110**, 3087-111
16. Tweedle, M.F. *Acc. Chem. Res.* 2009, **42**, 958-68
17. Maecke, H.R., Hofmann, M., Haberkorn, U. *J. Nucl. Med.* 2005, **46**, 172S-8S

18. Okarvi, S.M. *Med. Res. Rev.* 2004, **24**, 357-97

19. Al-Nahhas, A., Fanti, S. *Eur. J. Nucl. Med. Mol. Imaging*. 2012, **39**, S1-3

20. Hemmatennejad, B., Miri, R., Elyasi, M. *J. Theor. Biol.* 2012, **305**, 37-44

21. Kojima, M., Hosoda, H., Date, Y., Nakazato, M., Matsuo, H., Kangawa, K. *Nature*. 1999, **402**, 656-60.

22. Van Craenenbroeck, M., Gregiore, F., De Neef, P., Robberecht, P., Perret, J. *Peptides*. 2004, 25, 959-65

23. Rosita, D., DeWit, M.A., Luyt, L.G. *J. Med. Chem.* 2009, **52**, 2196-203

24. Aloj, L., Morelli, G. *Curr. Pharm. Des.* 2004, **10**, 3009-31

25. Frank, R. *Tetrahedron*. 1992, **48**, 9217-32

26. Panicker, R.C., Huang, X., Yao, S.Q. *Comb. Chem. High T. Scr.* 2004, **7**, 547-56

27. Hilpert, K., Winkler, D.F.H., Hancock, R.E.W. *Nat. Protoc.* 2007, **2**, 1333-49

28. Fodor, S.P., Read, J.L., Pirrung, M.C., Stryer, L., Lu, A.T., Solas, D. *Science*. 1991, **251**, 767-73

29. Frank, R. *J. Immunol. Methods*. 2002, **267**, 13-26

30. Smith, G.P., Petrenko, V.A. *Chem. Rev.* 1997, **97**, 391-410

31. Chakravarthy, B., Menard, M., Brown, L., Atkinson, T., Whitfield, J. *Biochem. Biophys. Res. Comm.* 2012 (In Press)

32. Lam, K.S., Salmon, S.E., Hersh, E.M., Hruby, V.J., Kazmierski, W.M., Knapp, R.J. *Nature*. 1991, **354**, 82-4

33. Yao, N., Xiao, W., Wang, X., Marik, J., Park, S.H., Takada, T., Lam, K.S. *J. Med. Chem.* 2009, **52**, 126-33

34. Fridkin, G., Bonasera, T.A., Litman, P., Gilon, C. *Nucl. Med. Bio*. 2005, **32**, 39-50

35. Roof, R.A., Sobczyk-Kojiro, K., Turbiak, A.J., Roman, D.L., Pogozheva, I.D., Blazer, L.L., Neubig, R.R., Mosberg, H.I. *Chem. Biol. Drug. Des.* 2008, **72**, 111-9

36. Edman, P. *Acta Chem. Scand.* 1950, **4**, 283-93

37. Medzihradszky, K.F., Campbell, J.M., Baldwin, M.A., Falick, A.M., Juhasz, P., Vestal, M.L., Burlingame, A.L. *Anal. Chem.* 200, **72**, 522-58

38. Brown, B.B., Wagner, D.S., Geysen, H.M. *Mol. Divers.* 1995, **1**, 4-12

39. Chaurand, P., Luetzenkirchen, F., Spengler, B. *J. Am. Soc. Mass Spectrom.* 1999, **10**, 91-103

40. Lubec, G., Afjehi-Sadat, L. *Chem Rev.* 2007, **107**, 3568-84

41. Hernandez, P., Gras, R., Frey, J., Appel, R.D. *Proteomics*. 2003, **3**, 870-8

42. Liu, S., Edwards, D.S. *Bioconjugate Chem*. 2001, **12**, 7-34

43. Lee, S., Xie, J., Chen, X. *Biochemistry*. 2010, **49**, 1364-76

44. Eckelman, W.C. *Eur. J. Nucl. Med.* 1995, **22**, 249-63

45. Del Roasrio, R.B., Jung, Y.W., Baidoo, K.E., Lever, S.Z., Wieland, D.M. *Nucl. Med. Biol.* 1994, **21**, 197-203

46. Pearson, D.A., Lister-James, J., McBride, W.J., Wilson, D.M., Martel, L.J., Civitello, E.R., Taylor, J.E., Moyer, B.R., Dean, R.T. *J. Med. Chem.* 1996, **39**, 1361-71

47. Dilworth, J.R., Parrott, S.J. *Chem. Soc. Rev.* 1998, **27**, 43-55

48. Schibli, R., Schwarzbach, R., Alberto, R., Ortner, K., Schmalle, H., Dumas, C., Egli, A., Schubiger, P.A. *Bioconjugate Chem.* 2002, **13**, 750-6

49. Harper, P.V., Lathrop, K.A., Jiminez, F., Fink, R., Cottschalk, A. *Radiology*. 1965, **85**, 101-9

50. Dilworth, J.R., Parrott, S.J. *Chem. Soc. Rev.* 1998, **27**, 43-55

51. Alberto, R., Egli, A., Abram, U., Hegetschweiler, K., Gramlich, V., Schubiger, P.A. *J. Chem. Soc. Dalton Trans.* 1994, **19**, 2815-20

52. Alberto, R., Schibli, R., Egli, A., Schubiger, P.A., Abram, U., Kaden, T.A. *J. Am. Chem. Soc.* 1998, **120**, 7987-8

53. Dyszlewski, M.E., Bushgman, M.J., Alberto, R., Brodack, J.W., Knight, H., MacDonald, J., Chinen, L.K., Webb, E.G., Vries, E., Pattipawaej, M., Vanderheyden, J.-L. *J. Labelled Compds. Radiopharm.* 2001, **44**, S483-5

54. Wei, L., Babich, J., Zubieta, J. *Inorg. Chim. Acta*. 2005, **358**, 3691-700

55. Liu, G., Dou, S., He, J., Vanderheyden, J.L., Rusckowski, M., Hnatowich, D.J. *Bioconjugate Chem.* 2004, **15**, 1441-6

56. Simpson, E.J., Hickey, J.L., Breadner, D., Luyt, L.G. *Dalton Trans.*, 2012, **41**, 2950-8

57. Vitor, R.F., Alves, S., Correia, J.D.G., Paulo, A., Santos, I. *J. Organomet. Chem.* 2004, **689**, 4764-74

58. Zhang, J., Vittal, J.J., Henderson, W., Wheaton, J.R., Hall, I.H., Hor, T.S.A., Yan, Y.K. *J. Organomet. Chem.* 2002, **650**, 123-32

59. Kramer, D.J., Davison, A., Davis, W.M., Jones, A.G. *Inorg. Chem.* 2002, **41**, 6181-3

60. Karagiorgou, O., Patsis, G., Pelecanou, M., Raptopoulou, C.P., Terzis, A., Siatra-Papastaikoudi, T., Alberto, R., Pirmettis, I., Papadopoulos, M. *Inorg. Chem.* 2005, **44**, 4118-20

61. Pietzsch, H.J., Gupta, A., Reisgys, M., Drews, A., Seifert, S., Syhre, R., Spies, H., Alberto, R., Abram, U., Schubiger, A., Johannsen, B. *Bioconjugate Chem.* 2000, **11**, 414-24

62. Lipowska, M., Marzilli, L.G., Taylor, A.T. *J. Nucl. Med.* 2009, **50**, 454-60

63. Alberto, R., Pak, J.K., van Staveren, D., Mundwiler, S., Benny, P. *Biopolymers*. 2004, **76**, 324-33

64. Atwood, J.D., Wovkulich, M.J., Sonnenberger, D.C. *Acc. Chem. Res.* 1983, **16**, 350-5

65. Lim, N.C., Ewart, C.B., Bowen, M.L, Ferreira, C.L., Barta, C.A., Adam, M.J., Orvig, C. *Inorg. Chem.* 2008, **47**, 1337-45

66. Alberto, R., Egli, A., Abram, U., Hegetschweiler, K., Gramlich, V., Schubiger, P.A. *J. Chem. Soc. Dalton Trans.* 1994, **19**, 2815-2820

67. Lazarova, N., James, S., Babich, J., Zubieta, J. *Inorg. Chem. Commun.* 2004, **7**, 1023-6

68. Merrifield, R.B. *J. Am. Chem. Soc.* 1963, **8**, 255-73

69. Youngquist, R.S., Fuentees, G.R., Lacey, M.P., Keough, T., Baillie, T.A. *Rapid Commun. Mass Spectrom.* 1994, **8**, 77-81

70. Youngquist, R.S., Fuentes, G.R., Lacey, M.P., Keough, T. *J. Am. Chem. Soc.* 1995, **117**, 3900-6

71. Thakkar, A., Wavreille, A.-S., Pei, D. *Anal. Chem.* 2006, **78**, 5935-9

72. Sweeney, M.C., Pei, D. *J. Comb. Chem.* 2003, **5**, 218-22

73. Thakkar, A., Cohen, A.S., Connolly, M.D., Zuckermann, R.N., Pei, D. *J. Comb. Chem.* 2009, **11**, 294-302

74. Bicker, K.L., Sun, J., Lavigre, J.J., Thompson, P.R., *ACS Comb. Sci.* 2011, **13**, 232-43

75. Liu, R., Marik, J., Lam. K.S. *J. Am. Chem. Soc.* 2002, **124**, 7678-80

76. Yao, N., Xiao, W., Wang. X., Marik, J., Park, S.H., Takada, Y., Lam, K.S. *J. Med. Chem.* 2009, **52**, 126-33

77. Wang, X., Zhang, J., Song, A., Lebrilla, C.B., Lam, K.S. *J. Am. Chem. Soc.* 2004, **126**, 5740-9

78. Cramer, R., Corless, S. *Rapid Commun. Mass Spectrom.* 2001, **15**, 2058-66

79. Fitzgerald, M.C., Harris, K., Shevlin, C.G., Siuzdak, G. *Bioorg. Med. Chem. Lett.* 1996, **6**, 979-82

80. Semmler, A., Weber, R., Przybylski, M., Wittmann, V. *J. Am. Soc. Mass Spectrom.* 2010, **21**, 215-9

81. Maljaars, C.E., Halkes, K.M., de Oude, W.L., Haseley, S.R., Upton, P.J., McDonnell, M.B., Kamerling, J.P. *J. Comb. Chem.* 2006, **8**, 812-9

82. Amadei, G.A., Cho, C-F., Lewis, J.D., Luyt, L.G. *J. Mass Spectrom.* 2010, **45**, 241-51

83. Toganoh, M., Ikeda, S., Furuta, S. *Inorg. Chem.* 2007, **46**, 10003-15

84. Viola-Villegas, N., Rabideau, A.E., Bartholoma, M., Zubieta, J., Doyle, R.P. *J. Med. Chem.* 2009, **52**, 5253-61

85. Nold, M.J., Cerda, B.A., Wesdemiotis, C. *J. Am. Soc. Mass Spectrom.* 1999, **10**, 1-8

86. Grewal, R.N., El Aribi, H., Harrison, A.G., Siu, K.W.M., Hopkinson, A.C. *J. Phys. Chem. B.* 2004, **108**, 4899-908

87. Cordero, M.M., Houser, J.J., Wesdemoitis, C. *Anal. Chem.* 1993, **65**, 1594-1601

88. Steen, H., Mann, M. *Nat. Rev. Mol. Cell Biol.* 2004, **5**, 699-711

89. Yergey, A.L., Coorssen, J.R., Backlund Jr., P.S., Blank, P.S., Humphrey, G.A., Zimmerberg, J., Campbell, J.M., Vestal, M.L. *J. Am. Soc. Mass Spectrom.* 2002, **13**, 784-91

90. Paizs, B., Suhai, S. *Mass Spectrom. Rev.* 2005, **24**, 508-48

91. Kim, J.-S., Song, J.-S., Kim, Y., Park, S.B., Kim, H.-J. *Anal. Bioanal. Chem.* 2012, **402**, 1911-9

92. Rogaslki, J.C., Lin, M.S., Sniatynski, M.J., Youhnovski, N., Przybylski, M., Kast, J. *J. Am. Soc. Mass Spectrom.* 2005, **16**, 505-14

93. Cydzik, M., Rudowska, M., Stefanowicz, P., Szewczuk. *J. Pep. Sci.* 2011, **17**, 445-53

94. Aggarwal, S., Harden, J. L., Denmeade, S. R. *Bioconjugate Chem.* 2006, **17**, 335-40

95. Pei, D. *Chem. Biol.* 2010, **17**, 3-4

96. Aditya, A., Kodedek, T. *ACS Comb. Sci.* 2012, **14**, 164-9

97. Goksel, H., Wasserberg, D., Mocklinghoff, S., Araujo, B.V., Brunsveld, L. *Bioorg. Med. Chem.* 2011, **19**, 306-11

98. Bachor, R., Cydzik, M., Rudowska, M., Kluczyk, A., Stefanowicz, P., Szewczuk, Z. *Mol. Divers.* 2012, (Ahead of Print)

99. Nauck, M., Kleine, N., Orskov, C., Holst, J.J., Willms, B., Creutzfeldt, W. *Diabetologia*. 1993, **36**, 741-4

100. Dhanvantari, S., Brubaker, P.L. *Endocrinology*. 1998, **139**, 1630-7

101. Adelhorst, K., Hedegaard, B.B., Knudsen, L.B., Kirk, O. *J. Biol. Chem.* 1994, **269**, 6275-8

102. Leger, R., Thibaudeau, K., Robitaille, M., Quraishi, O., van Wyk, P., Bousquet-Gagnon, N., Carette, J., Castaigne, J.P., Bridon, D.P. *Bio. Med. Chem. Lett.* 2004, **14**, 4395-8

103. Runge, S., Thogersen, H., Madsen, K., lau, J., Rudolph, R. *J. Biol. Chem.* 2008, **283**, 11340-7

104. Underwood, C.R., Garibay, P., Knudsen, L.B., Hastrup, S., Peters, G.H., Rudolph, R., Reedtz-Runge, S. *J. Biol. Chem.* 2010, **285**, 723-30

105. Dakin, C.L., Gunn, I., Small, C.J., Eduads, C.M., Hay, D.L., Smith, D.M., Ghatei, M.A., Bloom, S.R. *Endocrinology*. 2001, **142**, 4244-50

106. Estall, J.L., Drucker, D.J. *Curr. Pharm. Des.* 2006, **12**, 1731-50

107. Jackson, S.H., Martin, T.S., Jones, J.D., Seal, D., Emanuel, F. *P. T.* 2010, **35**, 498-502

108. Eng, J., Kleinman, W.A., Singh, L., Singh, G., Raufman, J-P. *J. Biol. Chem.* 1992, **267**, 7402-5

109. Parkes, D., Jodka, C., Smith, P., Nayak, S., Rinehart, L., Gingerich, R., Chen, K., Young, A. *Drug Develop. Res.* 2001, **53**, 260-7

110. Brom, M., Oyen, W.J.G., Joosen, L., Gotthardt, M., Boerman, O.C. *Eur. J. Nucl. Med. Mol. Imaging*. 2010, **37**, 1345-55

111. Behnam Azad, B., Rota, V.A., Breadner, D., Dhanvantari, S., Luyt, L.G. *Bioorg. Med. Chem.* 2010, **18**, 1265-72

112. Raddatz, R., Schaffhauser, H., Marino, M.J. *Biochem. Pharmacol.* 2007, **74**, 383-91

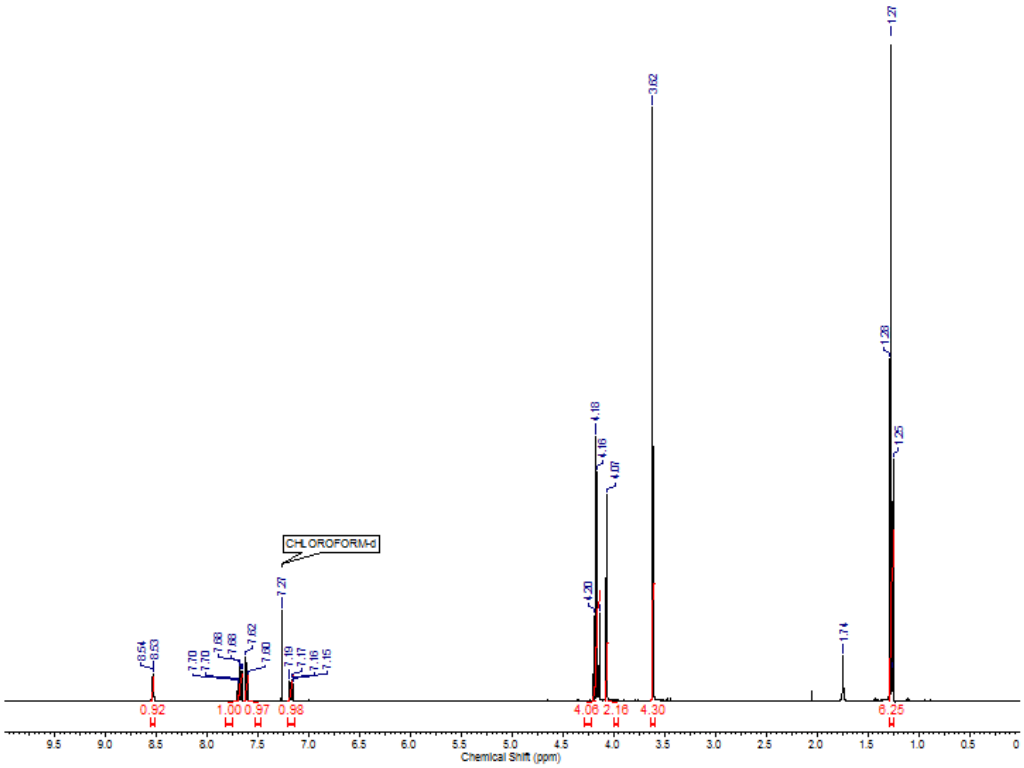
113. Tibaduiza, E.C., Chen, C., Beinborn, M. *J. Biol. Chem.* 2001, **276**, 37787-93

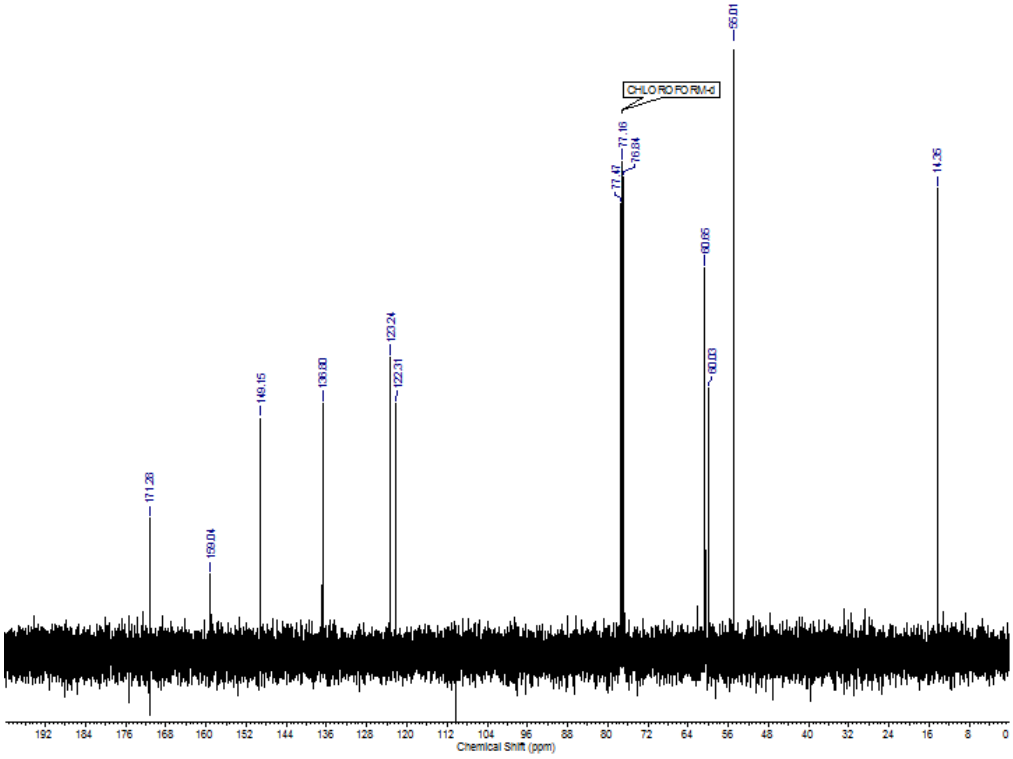
114. Koole, C., Wootten, D., Simms, J., Valant, C., Sridhar, R., Woodman, O.L., Miller, L.J., Summers, R.J., Christopoulos, A., Sexton, P.M. *Mol. Pharmacol.* 2010, **78**, 456-65

115. Knudsen, L.B., Kiel, D., Teng, M., Behrens, C., Bhumralkar, D., Kodra, J.T., Holst, J.J., Jeppesen, C.B., Johnson, M.D., de Jong, J.C., Jorgensen, A.S., Kercher, T., Kostrowicki, J., Madsen, P., Olesen, P.H., Petersen, J.S., Poulsen, F., Sidelmann, U.G., Sturis, J., Truesdale, L., May, J., Lau, J. *Proc. Nat. Acad. Sci. USA.* 2007, **104**, 937-42

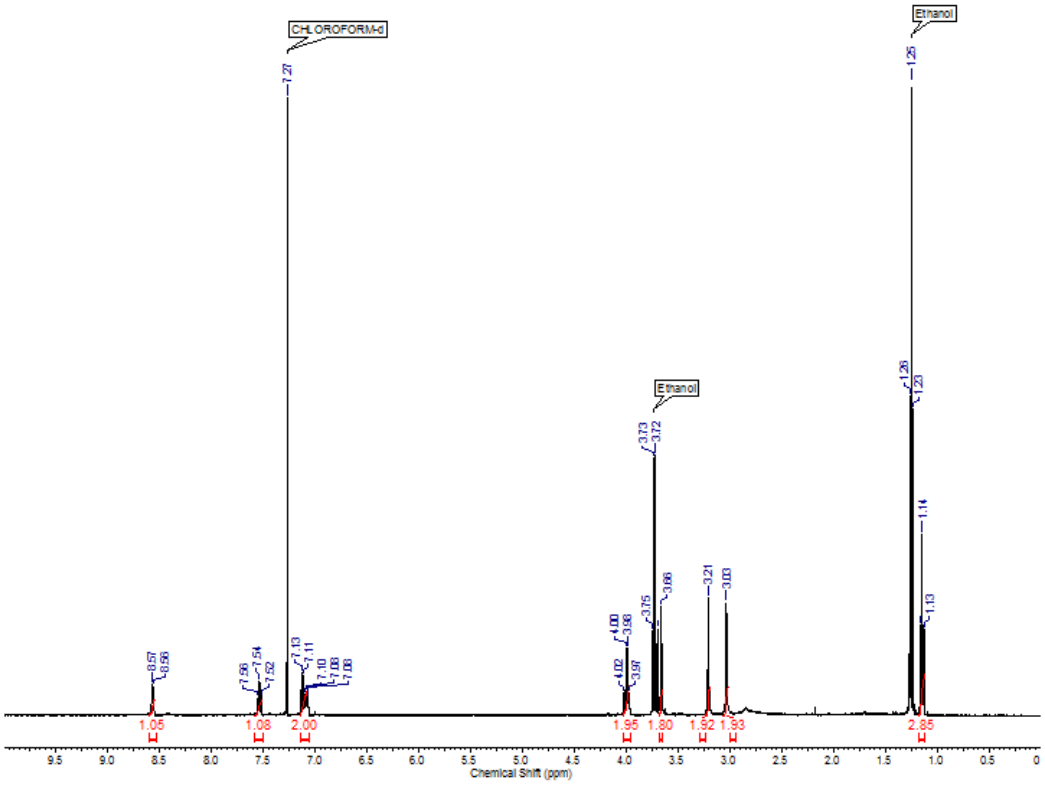
APPENDIX I – CHAPTER 2

^{1}H NMR and ^{13}C NMR spectra of **2.3**

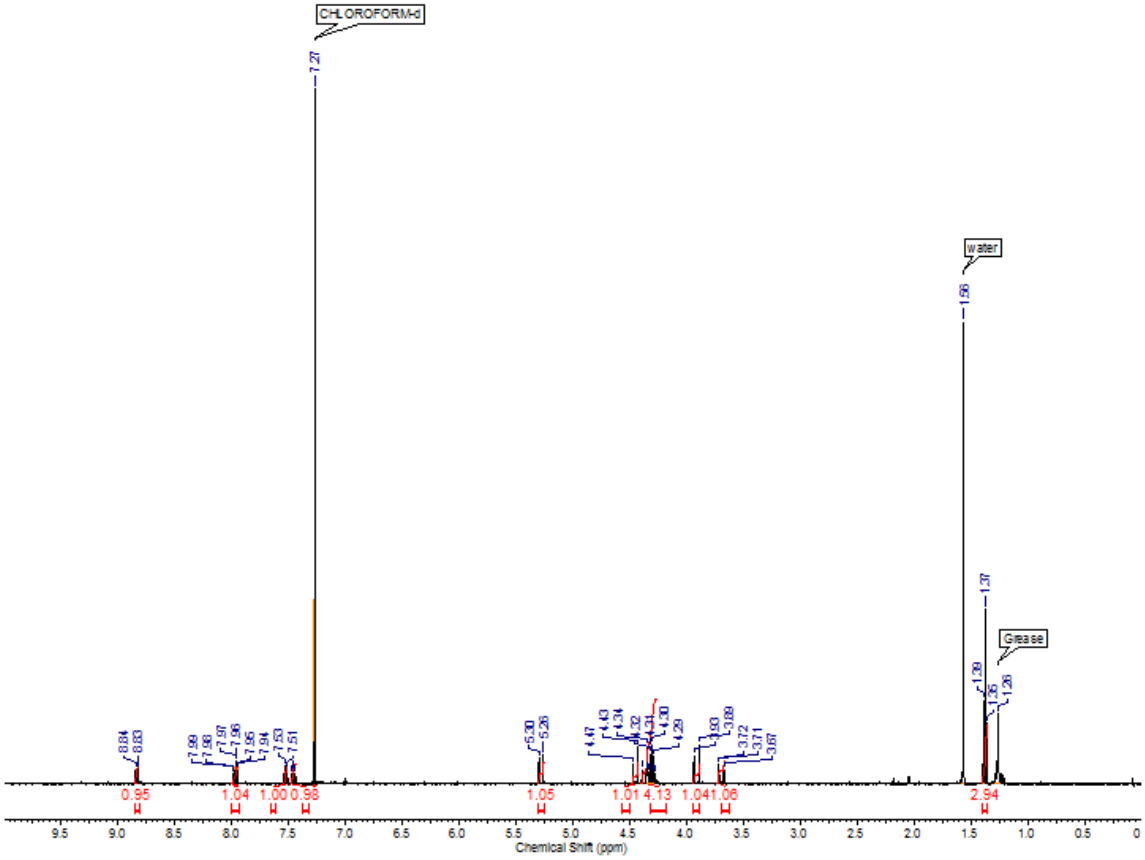




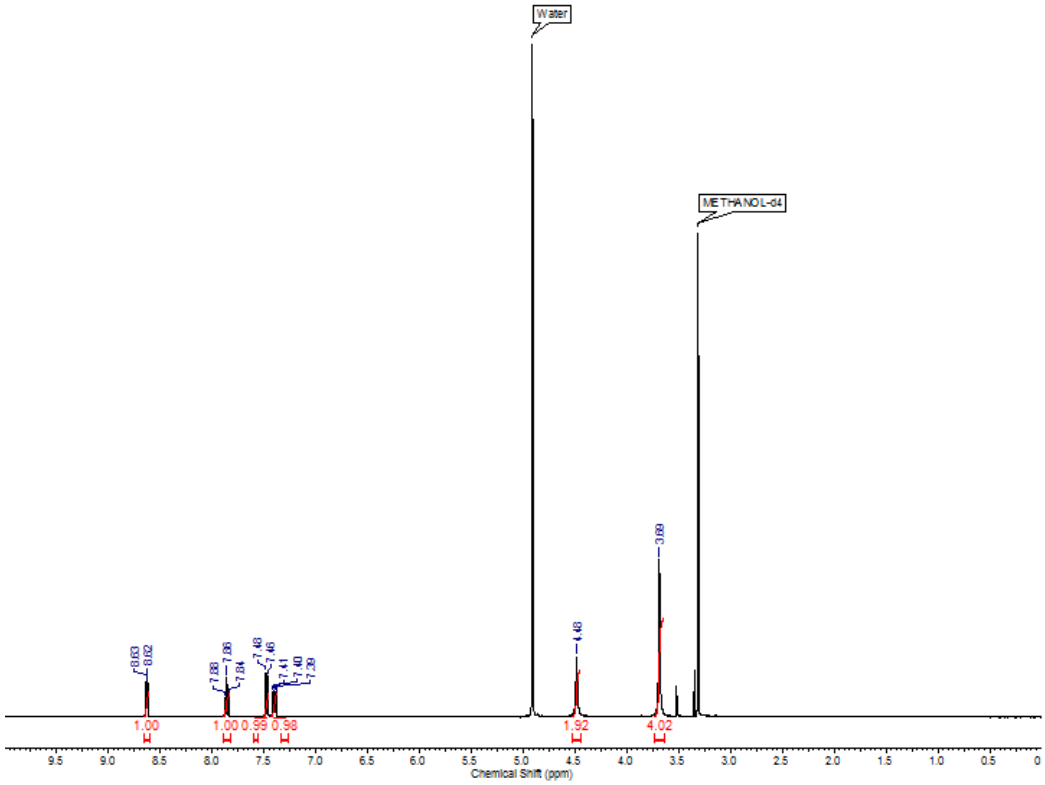
^{1}H NMR spectrum of **2.4**

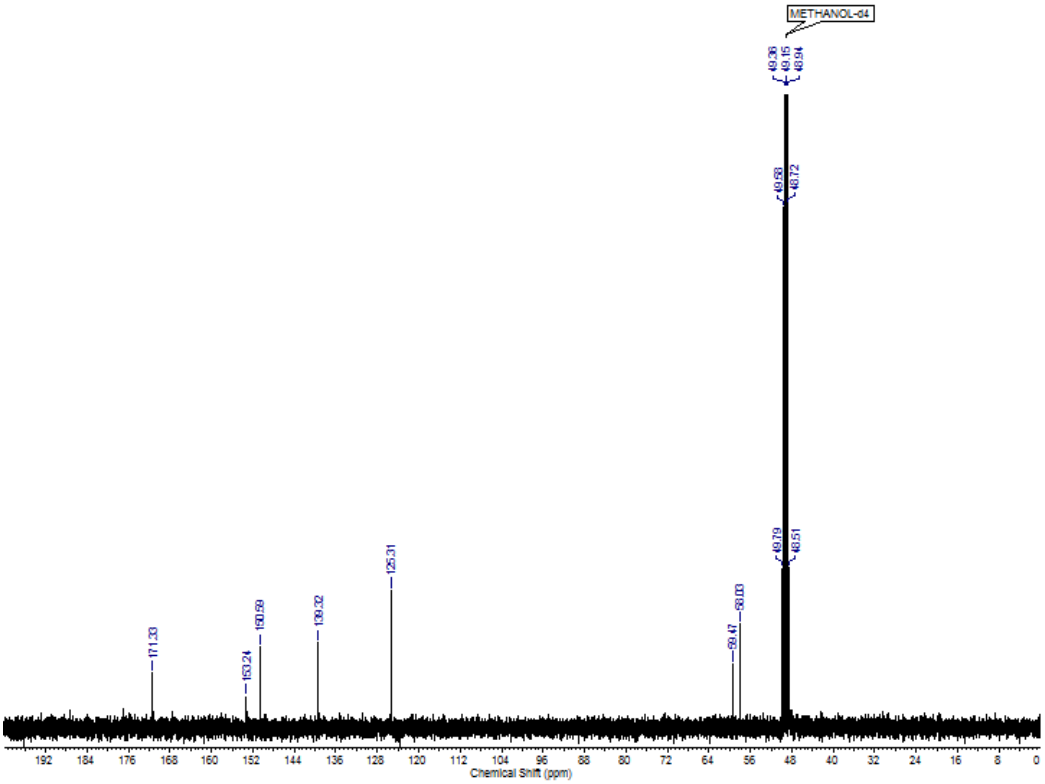


^{1}H NMR spectrum of **2.5**

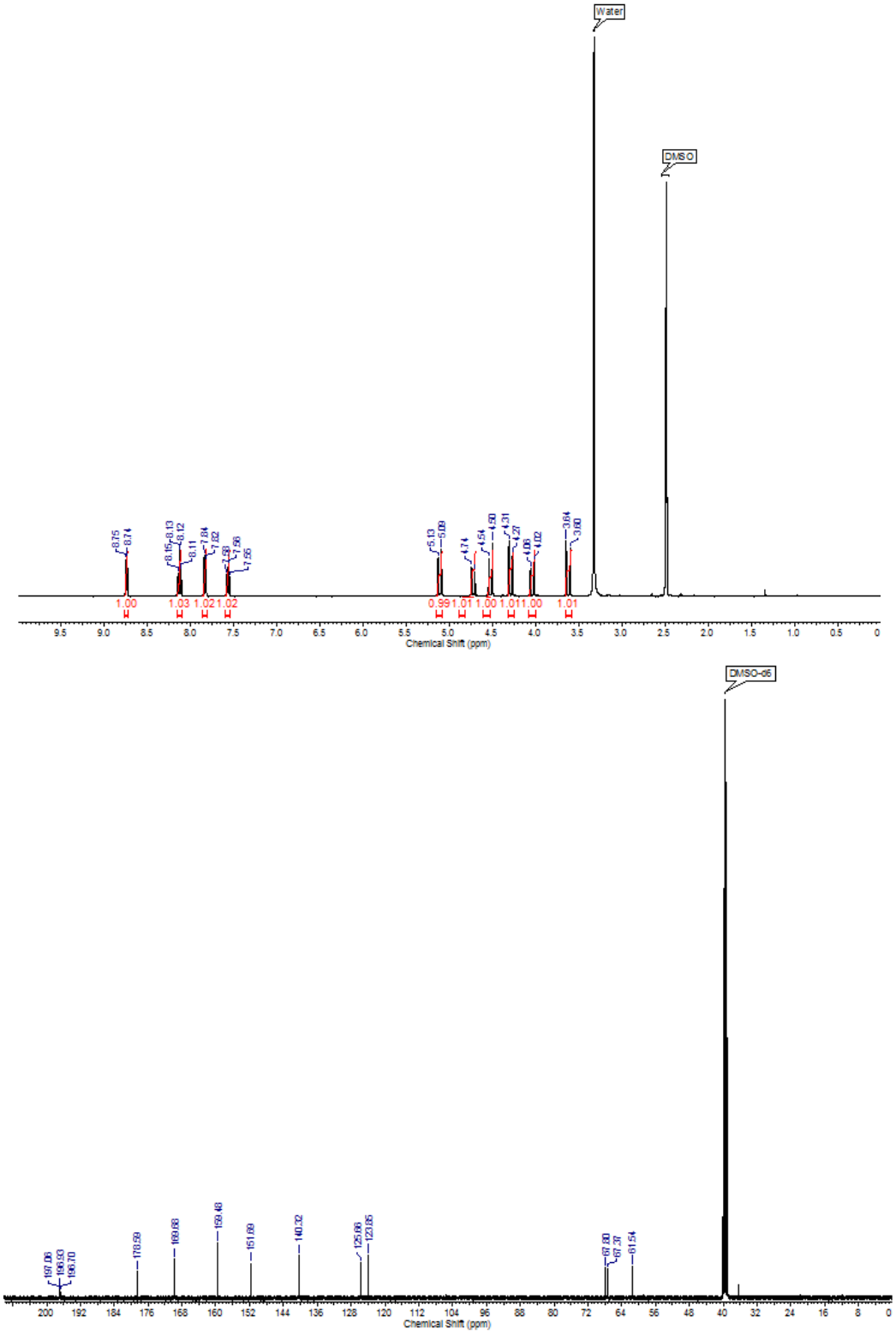


^{1}H NMR and ^{13}C NMR spectra of **2.6**

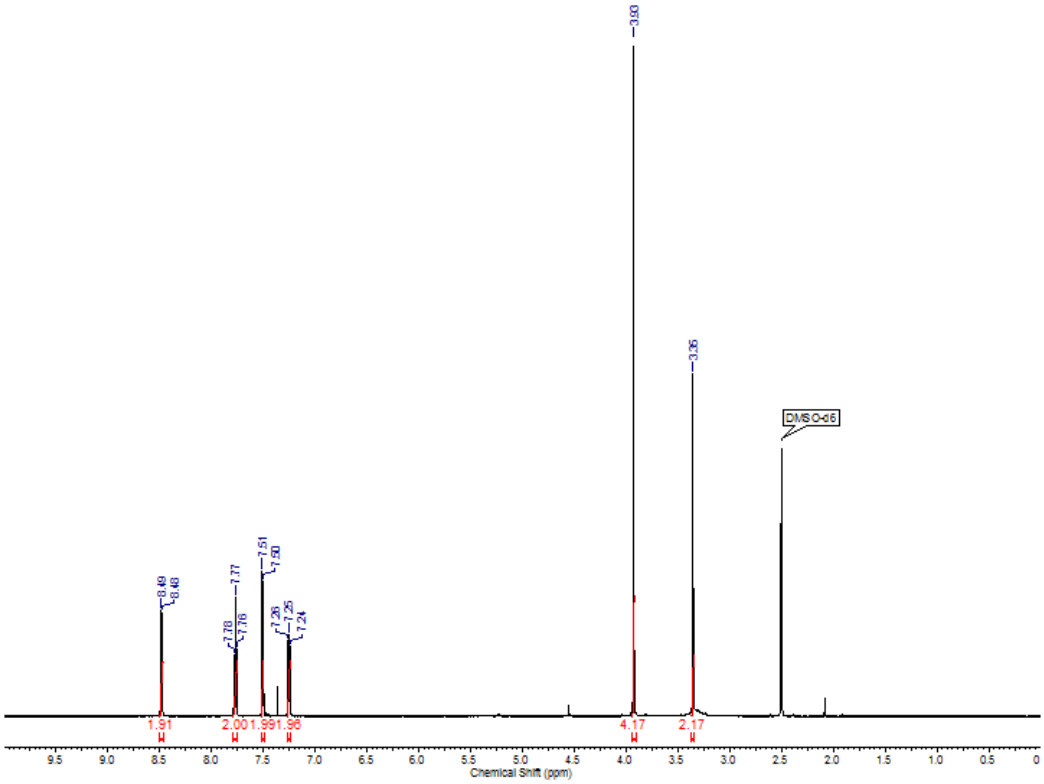


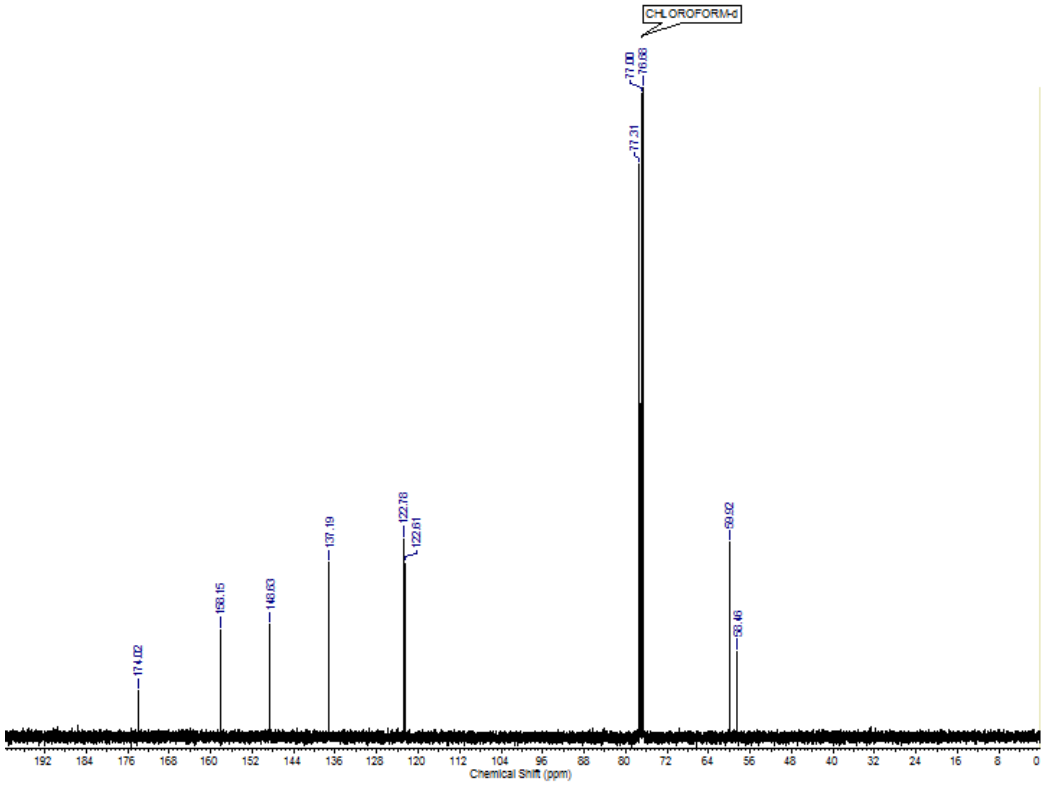


^{1}H NMR and ^{13}C NMR spectra of **2.7**

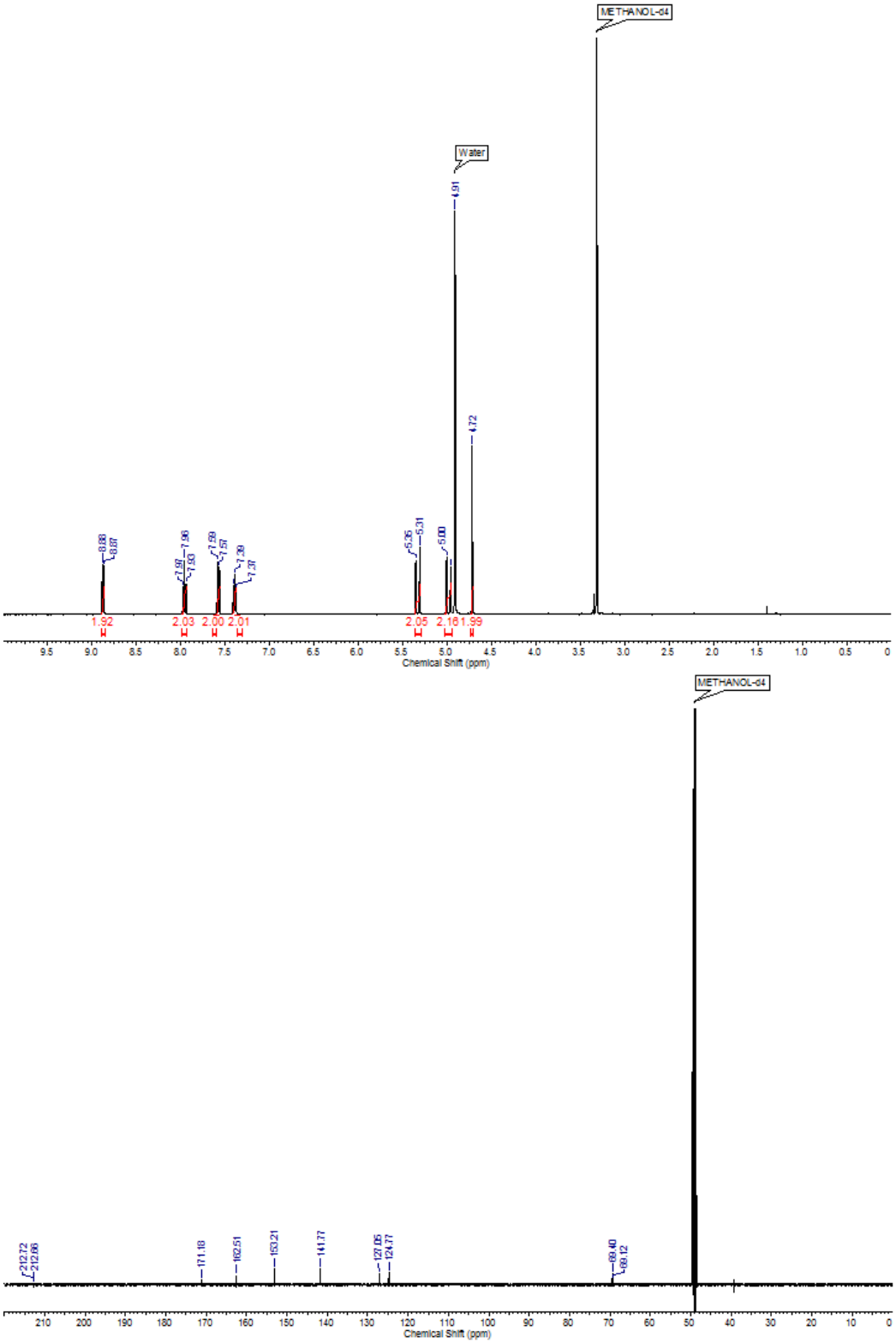


^{1}H NMR and ^{13}C NMR spectra of **2.8**

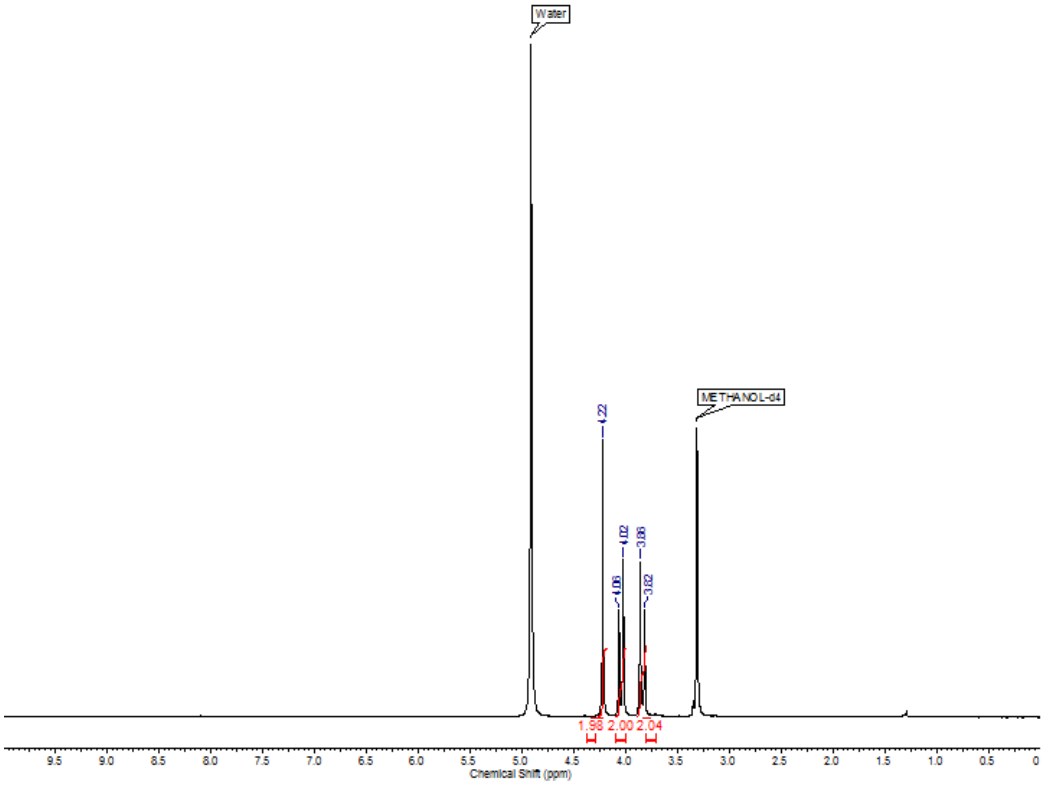


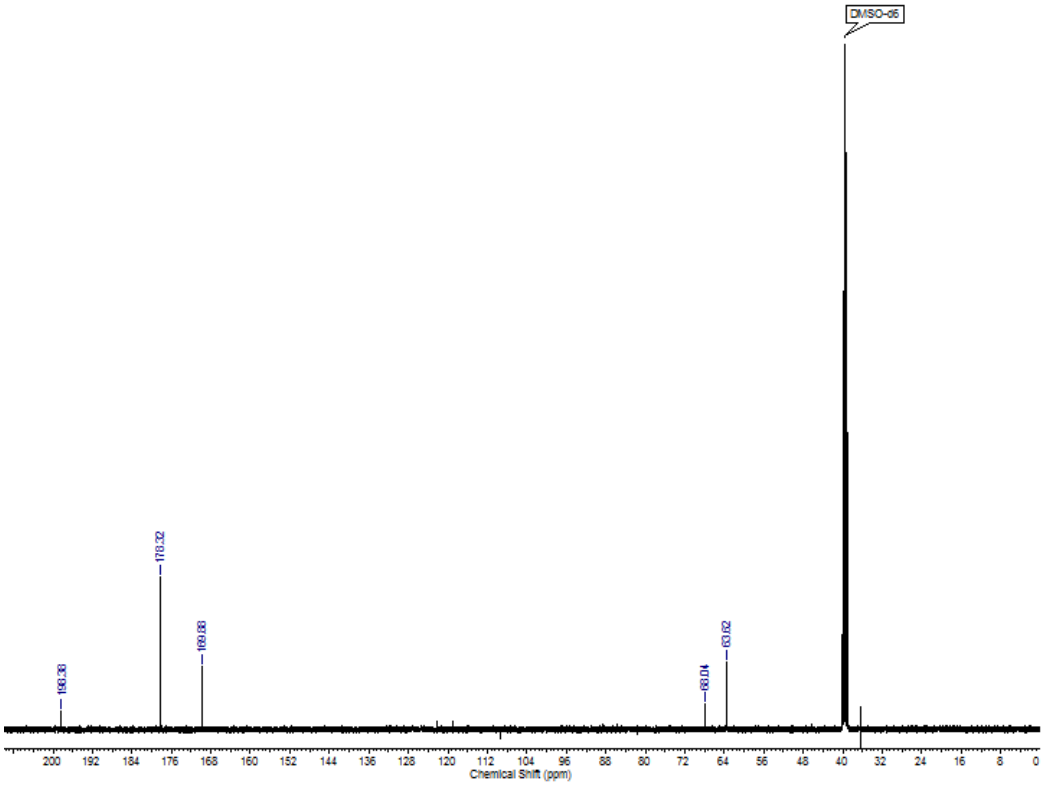


^{1}H NMR and ^{13}C NMR spectra of **2.9**

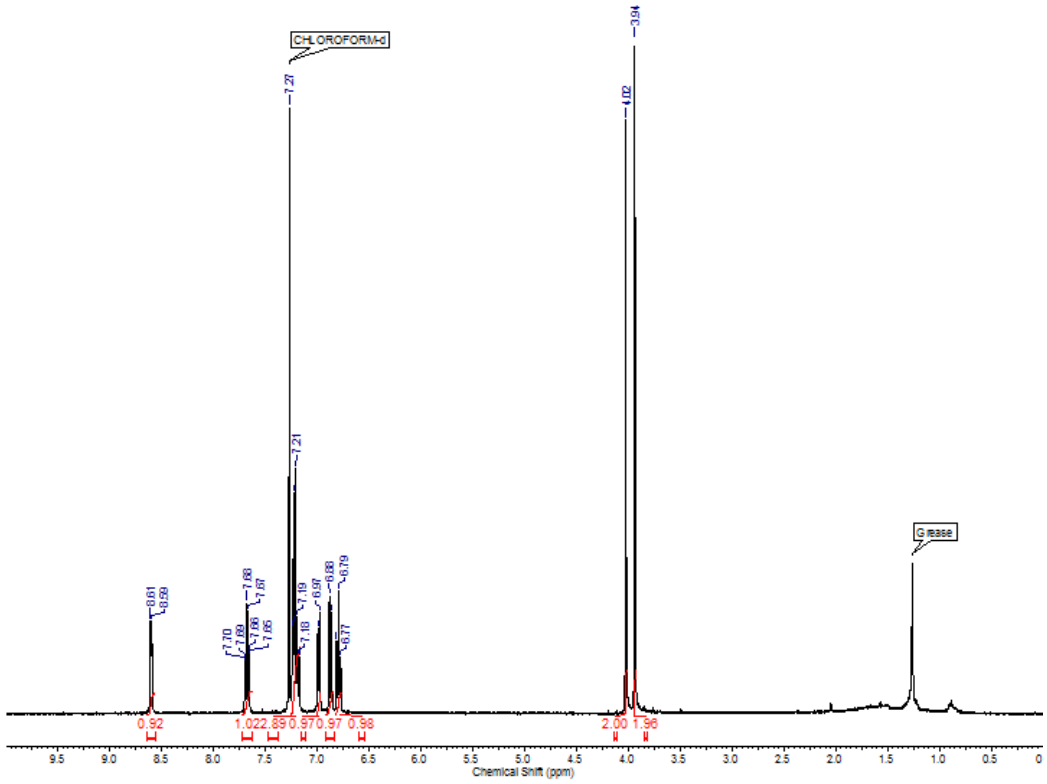


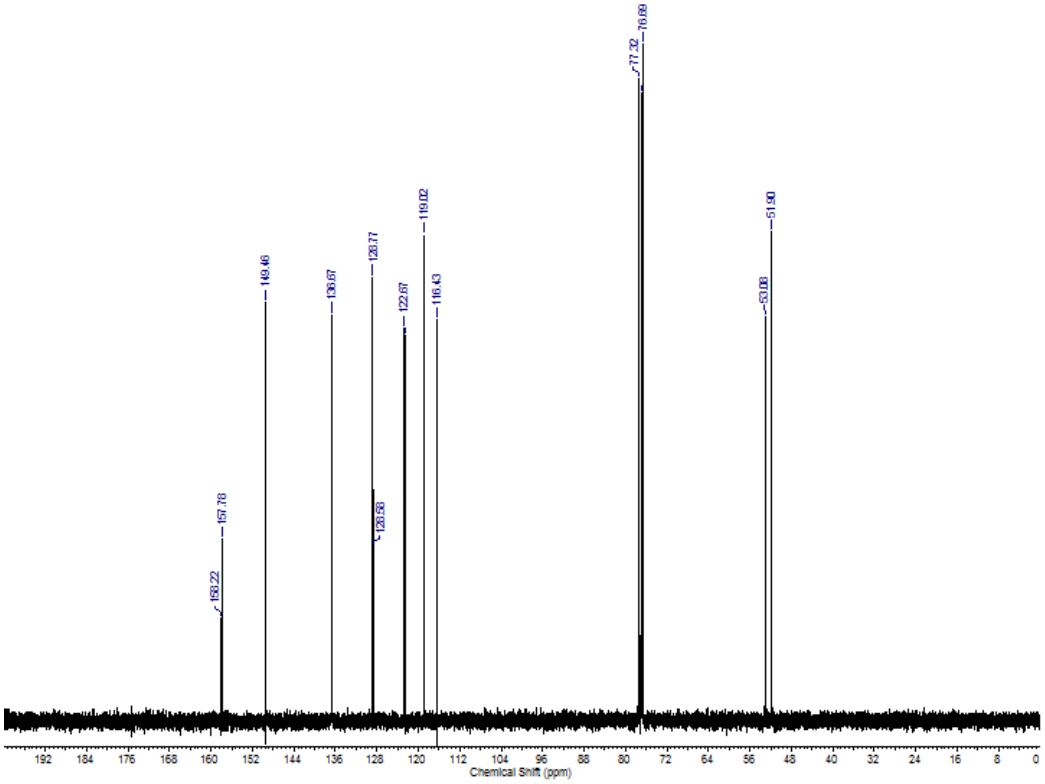
^{1}H NMR and ^{13}C NMR spectra of **2.10**



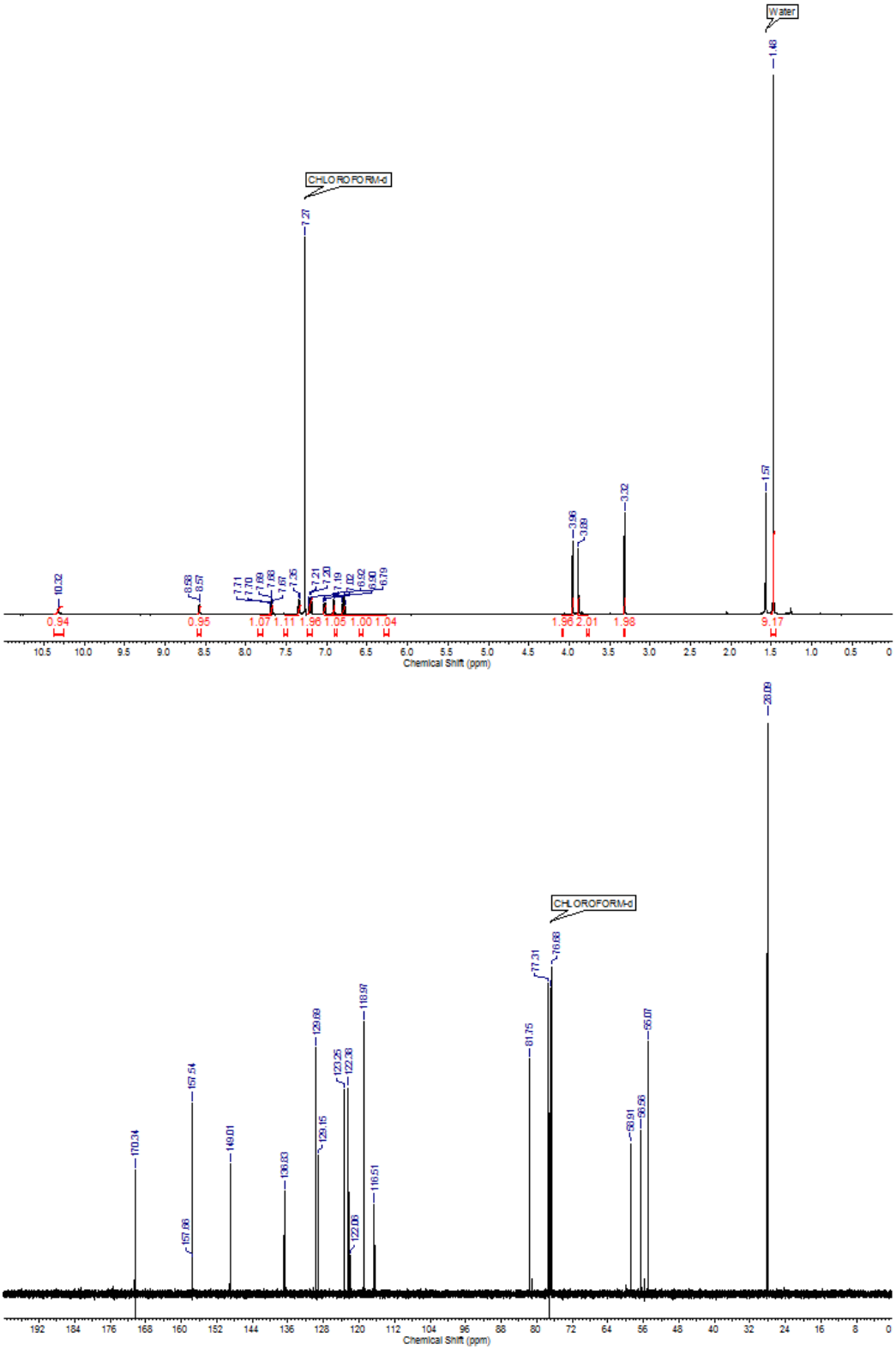


^{1}H NMR and ^{13}C NMR spectra of **2.11**

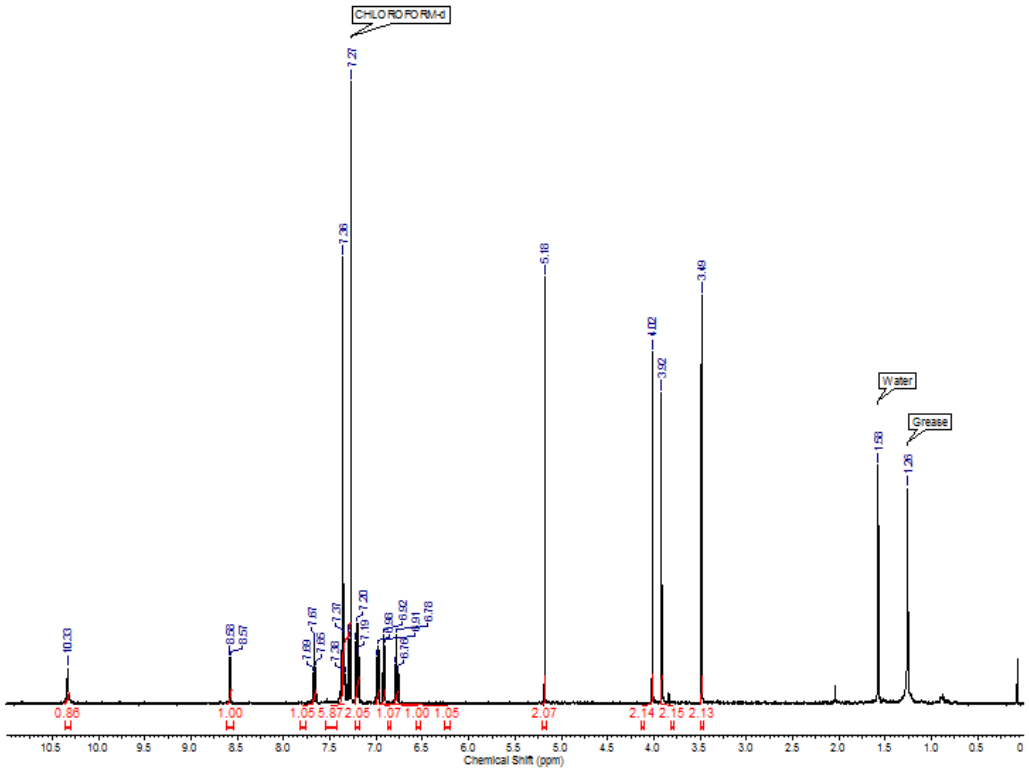




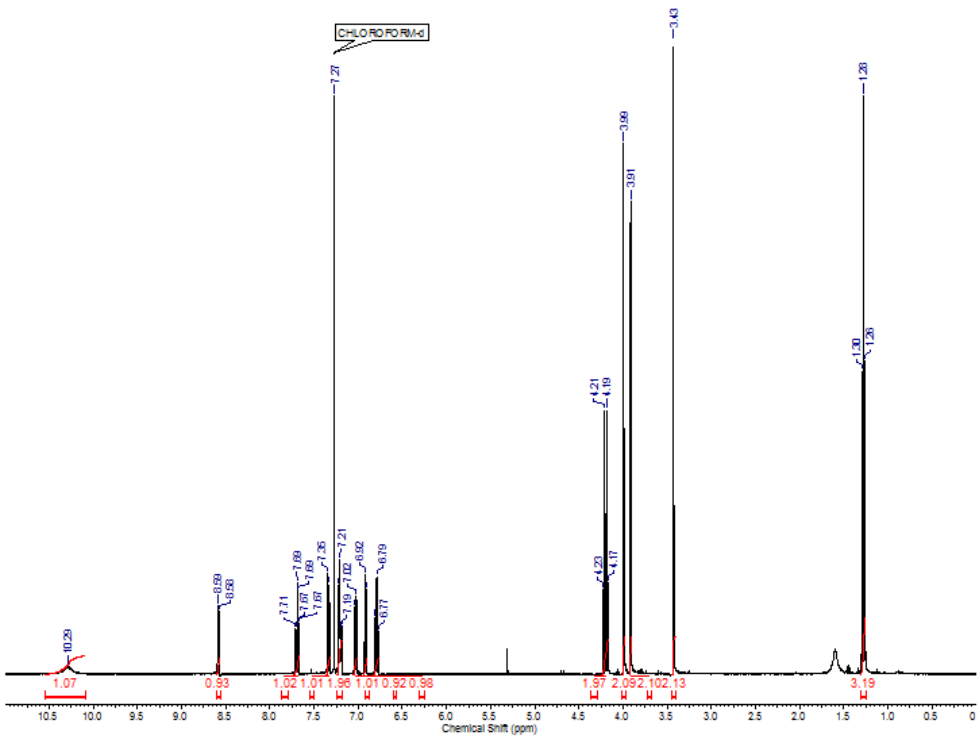
^{1}H NMR and ^{13}C NMR spectra of **2.12a**



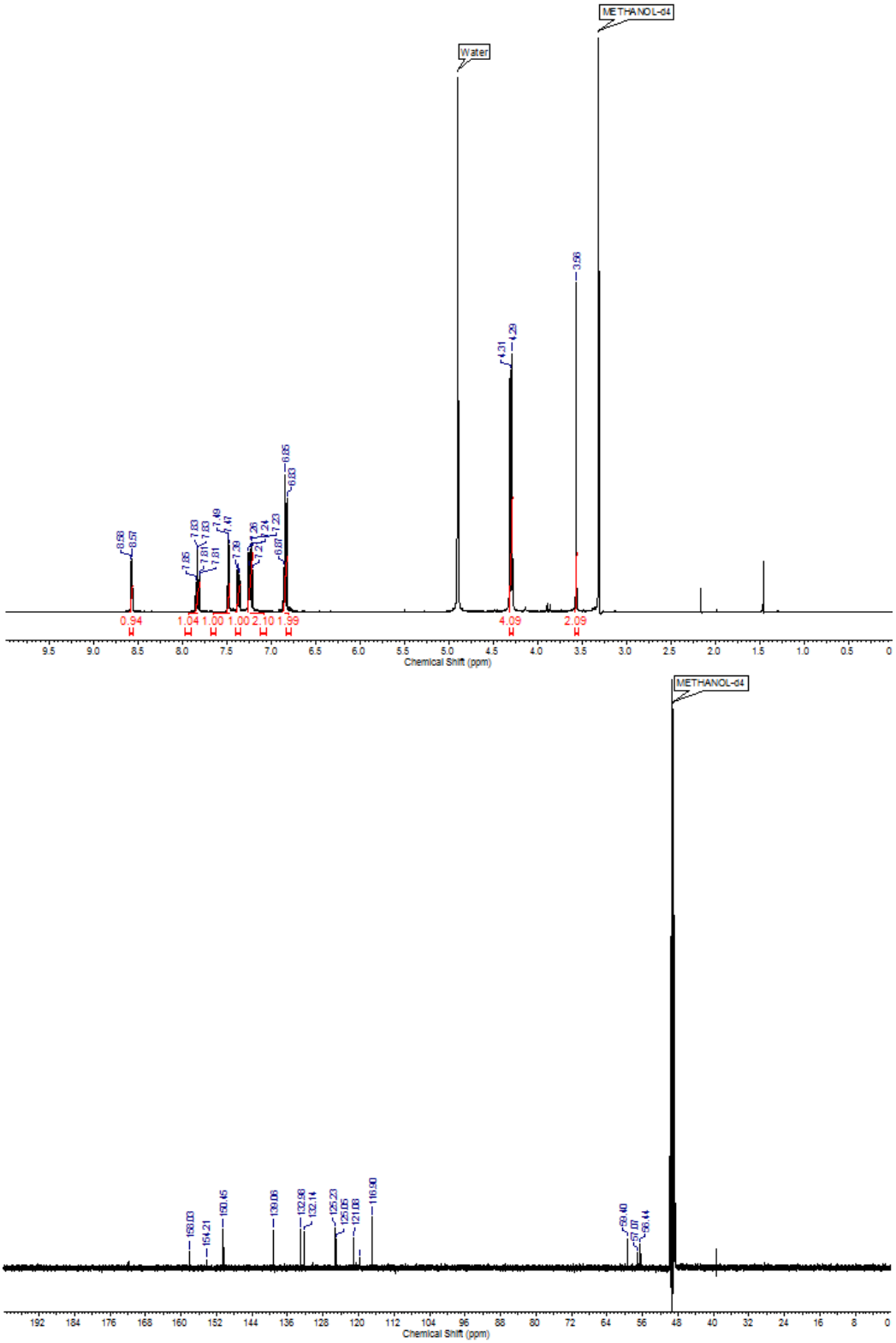
^{1}H NMR spectrum of **2.12b**



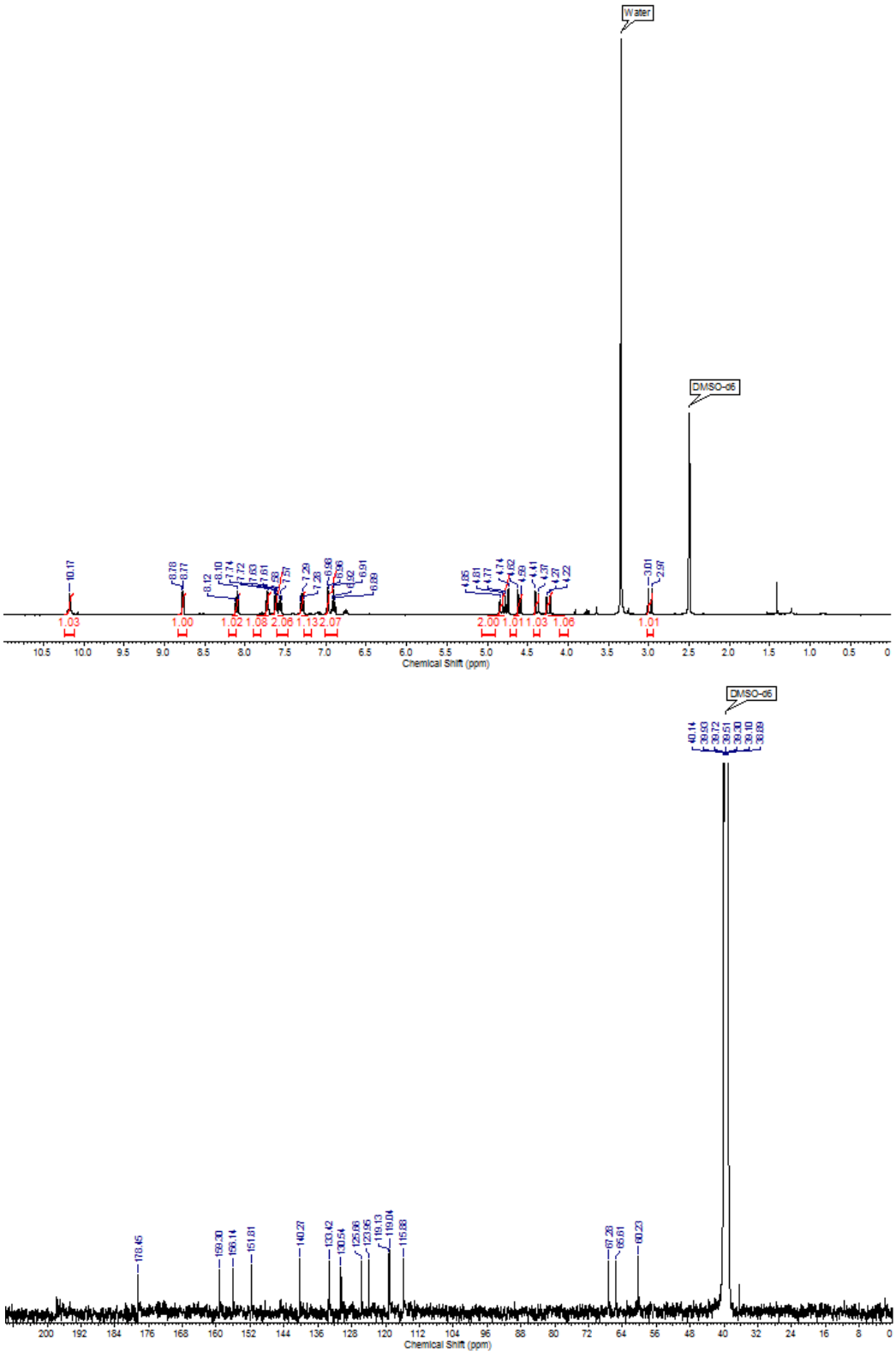
^{1}H NMR spectrum of **2.12c**



^{1}H NMR and ^{13}C NMR spectra of **2.13**

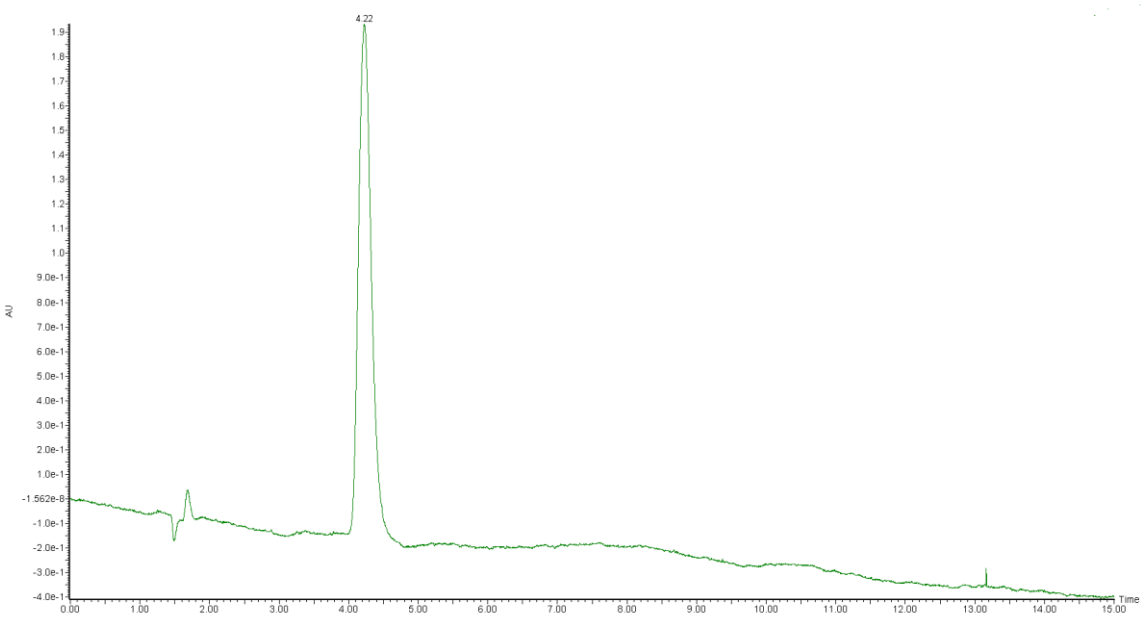


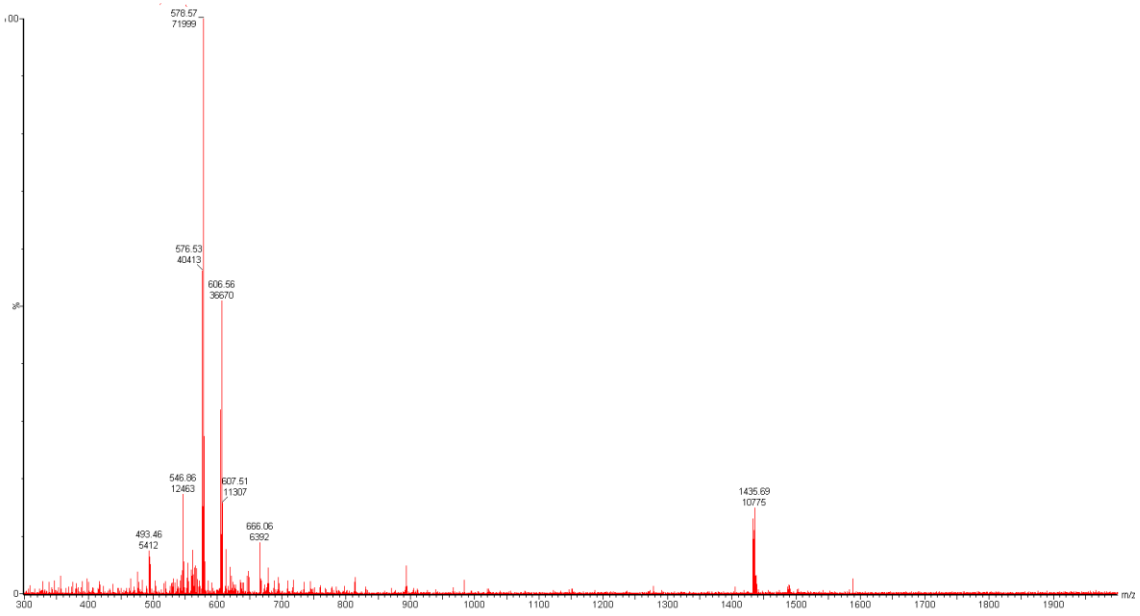
^{1}H NMR and ^{13}C NMR spectra of **2.14**



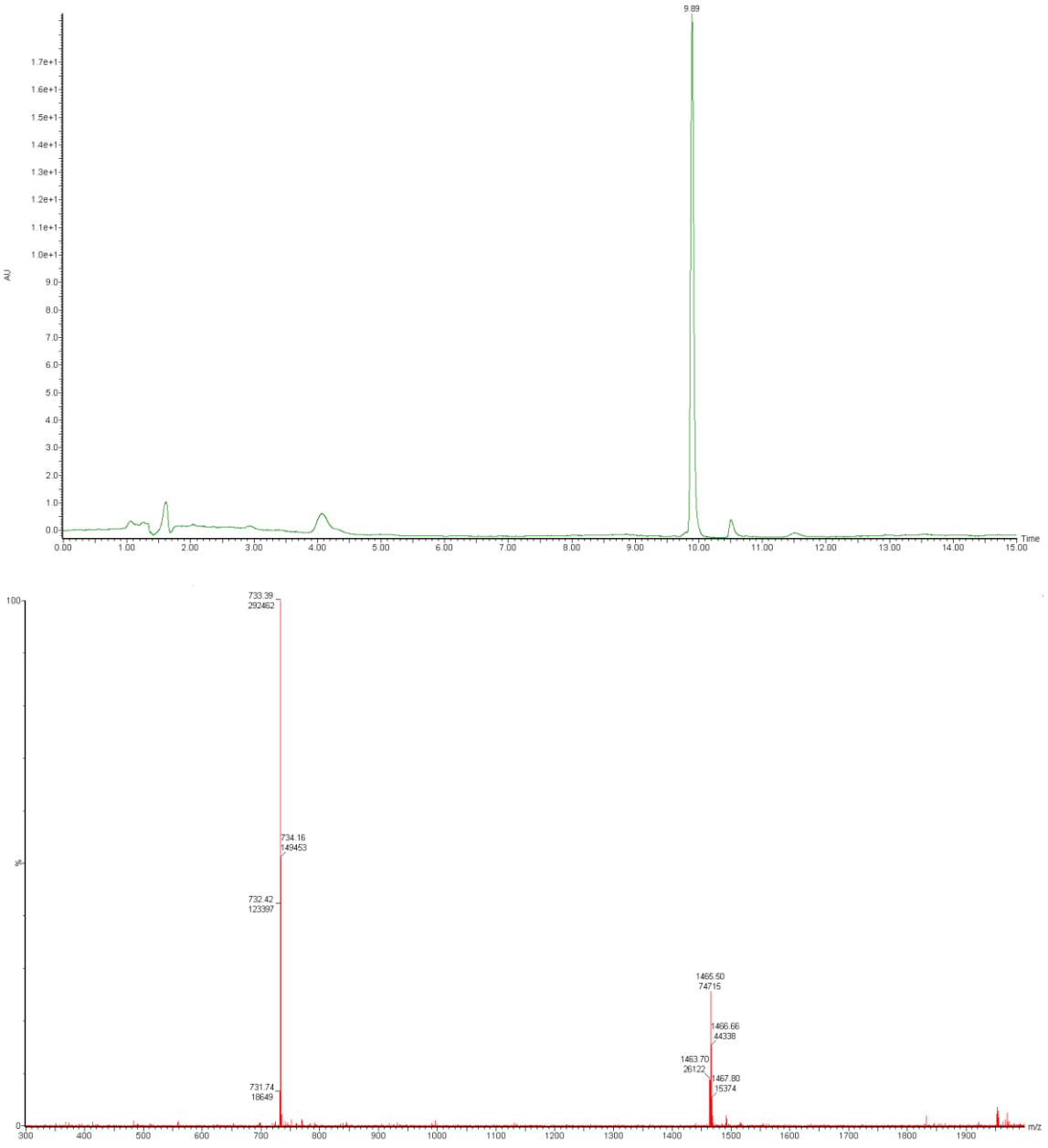
APPENDIX II – Chapter 3

HPLC trace and ES+ mass spectrum of **3.2**

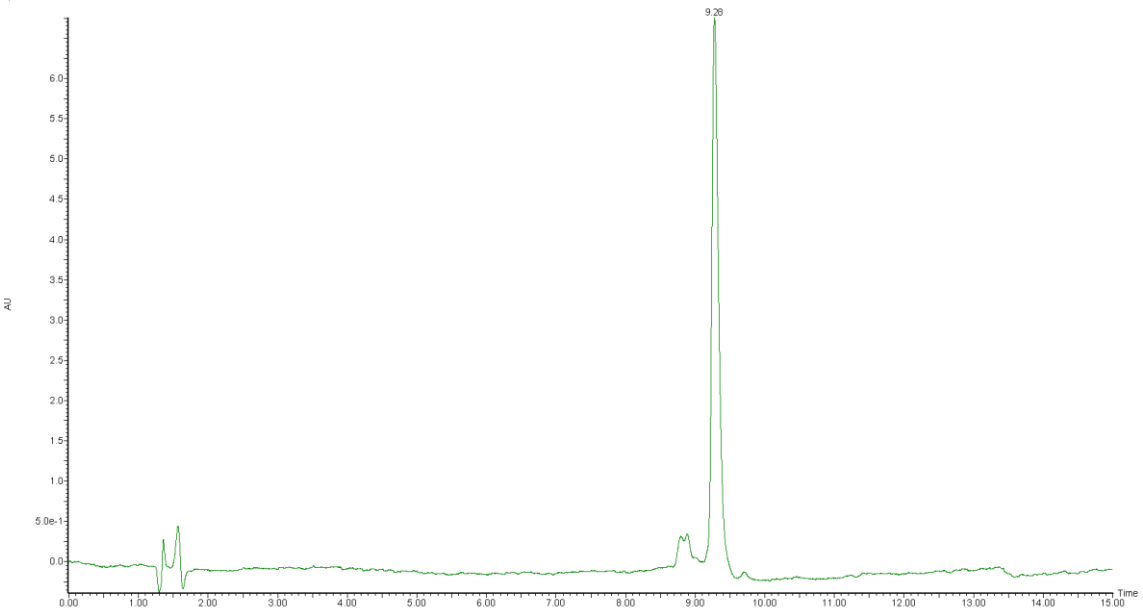


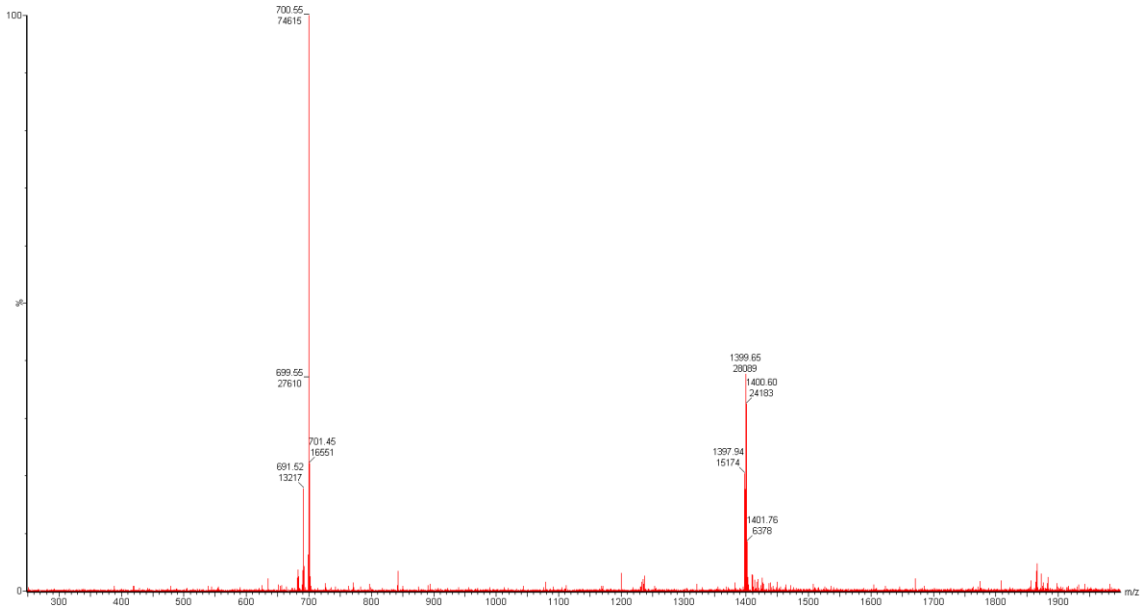


HPLC trace and ES+ mass spectrum of **3.3**

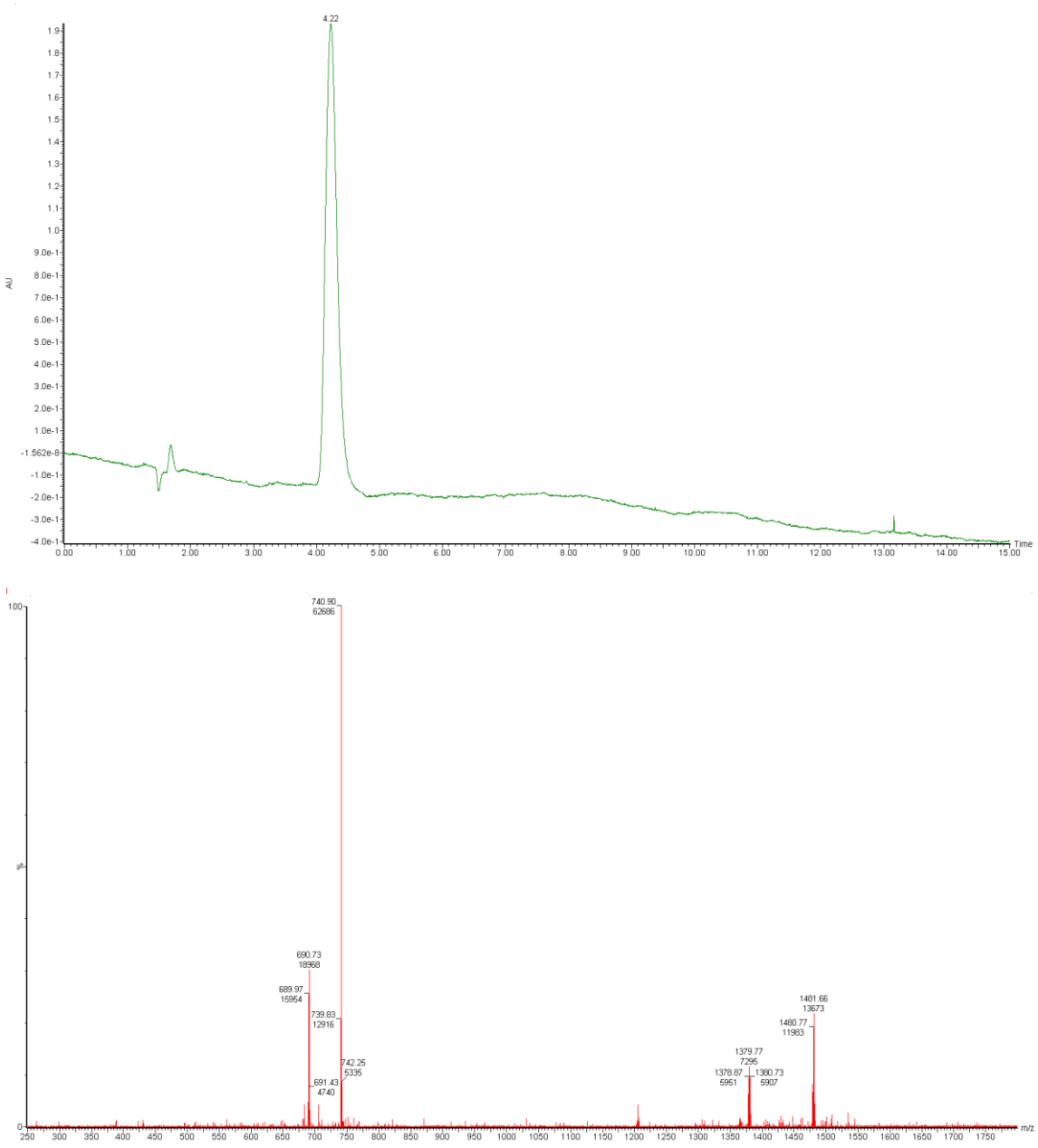


HPLC trace and ES+ mass spectrum of **3.4**





HPLC trace and ES+ mass spectrum of **3.6**

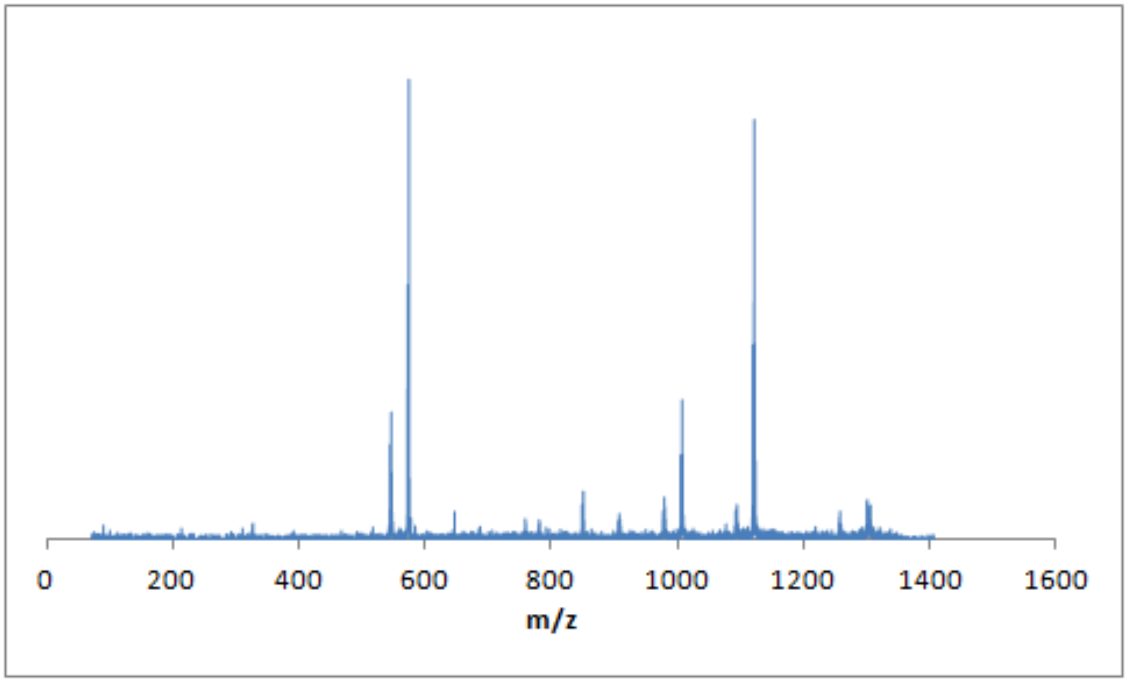


APENDIX III – Chapter 4

For all organometallic peptide product ion spectra, the rhenium-185 was selected for fragmentation

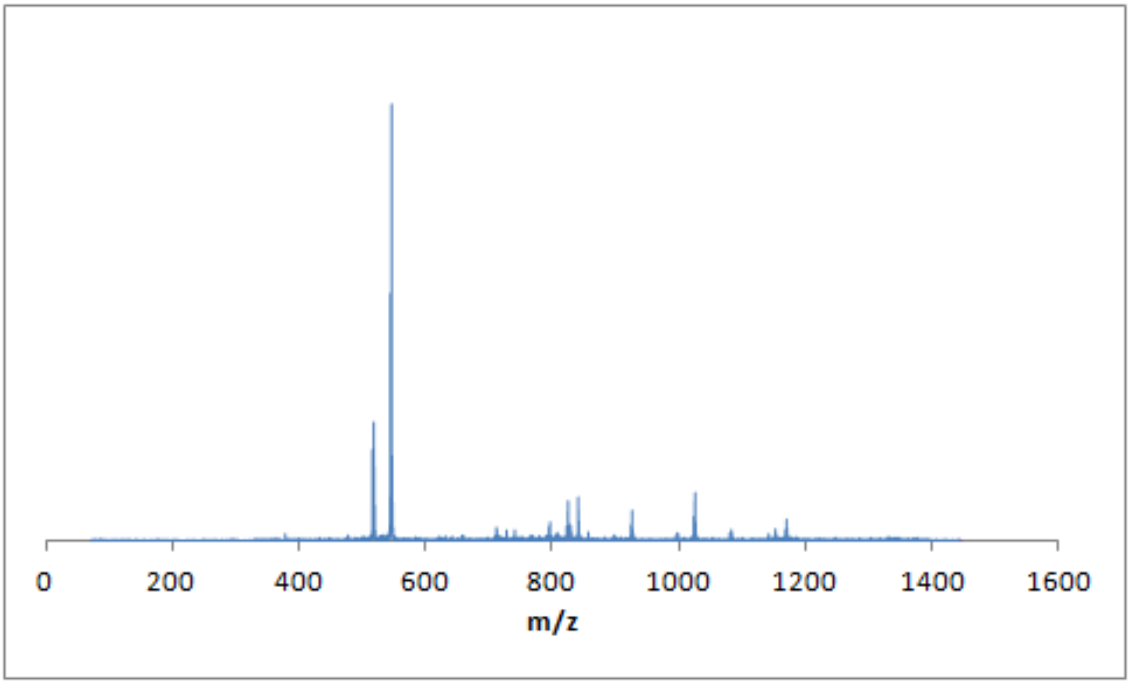
The MS/MS spectrum of compound [PADA-Re$(CO)_3$]-VLYGVRPV-NH_2 (**4.1**) and *b* series peak assignments
$[M+H]^+$ calc: 1335.51, obs: 1335.49 (rhenium-187 peak)

Series	Label	Fragment	Peak (*m/z*)	Difference between peaks (*m/z*)	Amino Acid Assignment
b series	b_0	NOT OBSERVED	475.0		
	b_1	RE-PADA-V	574.1	99.1	Valine
	b_2	RE-PADA-VL	687.2	113.1	Leucine
	b_3	RE-PADA-VLY	850.2	163.0	Tyrosine
	b_4	RE-PADA-VLYG	907.3	57.1	Glycine
	b_5	RE-PADA-VLYGV	1006.3	99.0	Valine
	b_6	RE-PADA-VLYGVR	1120.3	114.0	Asparginine
	b_7	RE-PADA-VLYGVRP	1217.4	97.1	Proline
	b_8	NOT OBSERVED	1316.5	99.0	Valine



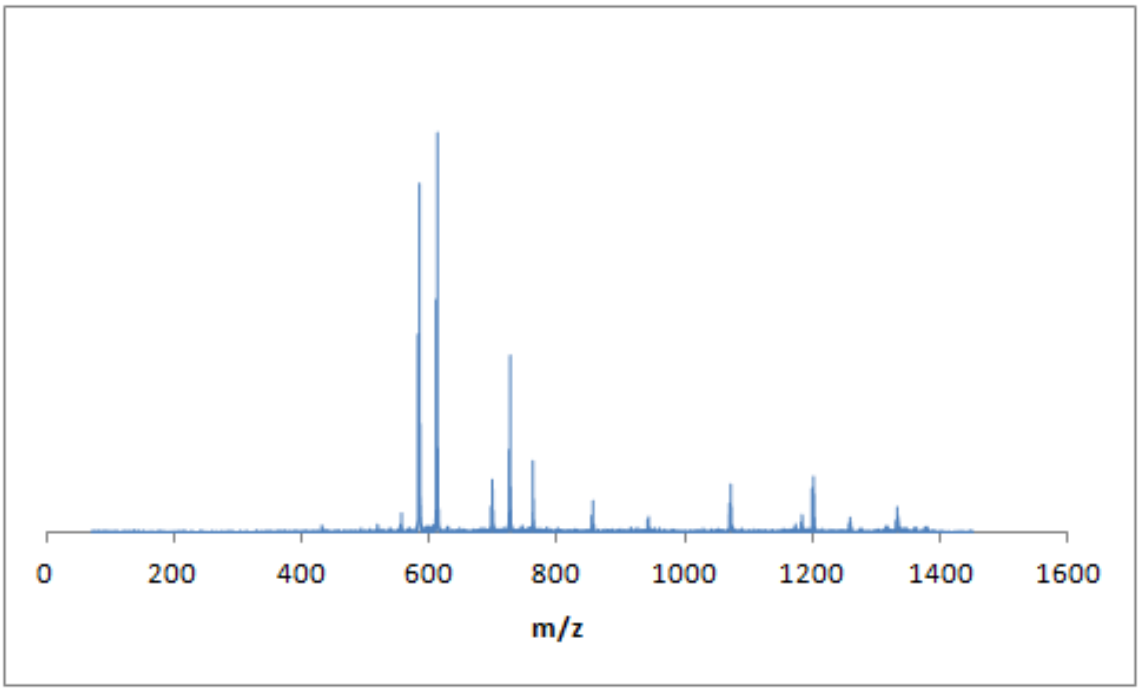
The MS/MS spectrum of compound [PADA-Re(CO)$_3$]-ALHEVGSW-NH$_2$ (**4.2**) and b series peak assignments
$[M+H]^+$ calc: 1373.46, obs: 1373.43 (rhenium-187 peak)

Series	Label	Fragment	Peak (*m/z*)	Difference between peaks (*m/z*)	Amino Acid Assignment
b series	b_0	NOT OBSERVED	475.0		
	b_1	RE-PADA-A	546.1	71.1	Alanine
	b_2	RE-PADA-AL	659.1	113.0	Leucine
	b_3	RE-PADA-ALH	796.2	137.1	Histidine
	b_4	RE-PADA-ALHE	925.3	129.1	Glutamic Acid
	b_5	RE-PADA-ALHEV	1024.3	99.0	Valine
	b_6	RE-PADA-ALHEVG	1081.3	57.0	Glycine
	b_7	RE-PADA-ALHEVGS	1168.3	87.0	Serine
	b_8	NOT OBSERVED	1354.4	186.1	Tryptophan



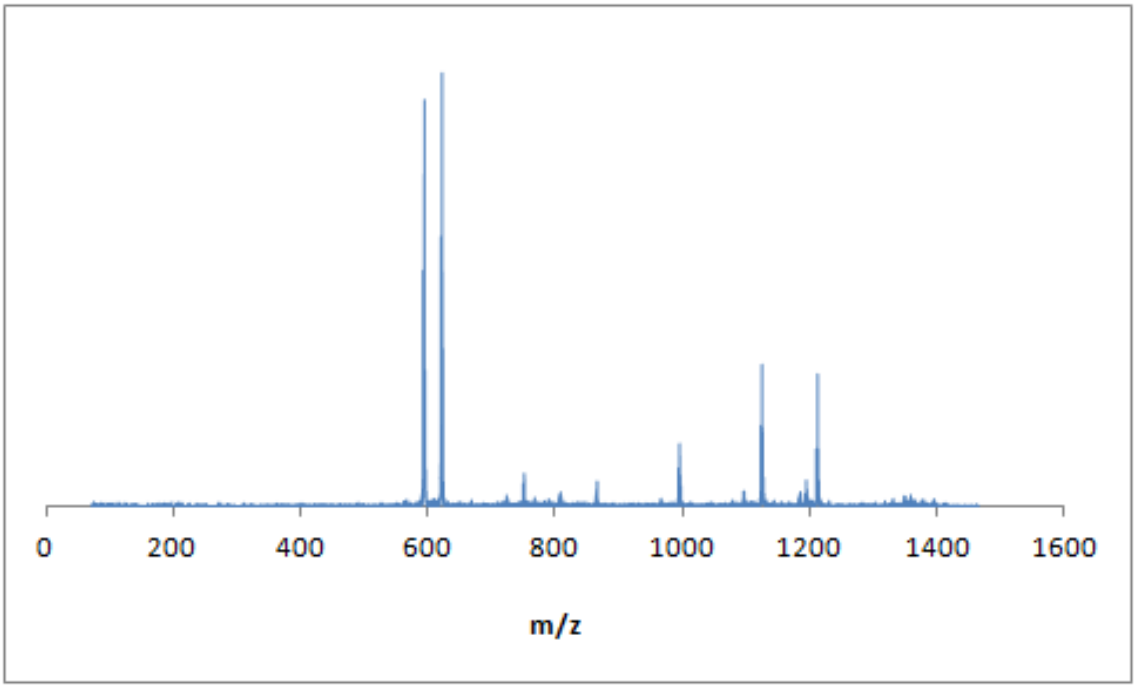
The MS/MS spectrum of compound [PADA-Re(CO)$_3$]-HNES(K/Q)EGT-NH$_2$ (**4.3**) and b series peak assignments
$[M+H]^+$ calc: 1376.38, obs: 1376.43 (rhenium-187 peak)

Series	Label	Fragment	Peak (*m/z*)	Difference between peaks (*m/z*)	Amino Acid Assignment
b series	b_0	NOT OBSERVED	475.0		
	b_1	RE-PADA-H	612.1	137.01	Histidine
	b_2	RE-PADA-HN	726.1	114	Asparginine
	b_3	RE-PADA-HNE	855.2	129.1	Glutamic Acid
	b_4	RE-PADA-HNES	942.2	87	Serine
	b_5	RE-PADA-HNES(K/Q)	1070.3	128.1	Lysine/Glutamine
	b_6	RE-PADA-HNES(K/Q)E	1199.3	129	Glutamic Acid
	b_7	RE-PADA-HNES(K/Q)EG	1256.3	57	Glycine
	b_8	NOT OBSERVED	1357.4	101.1	Threonine



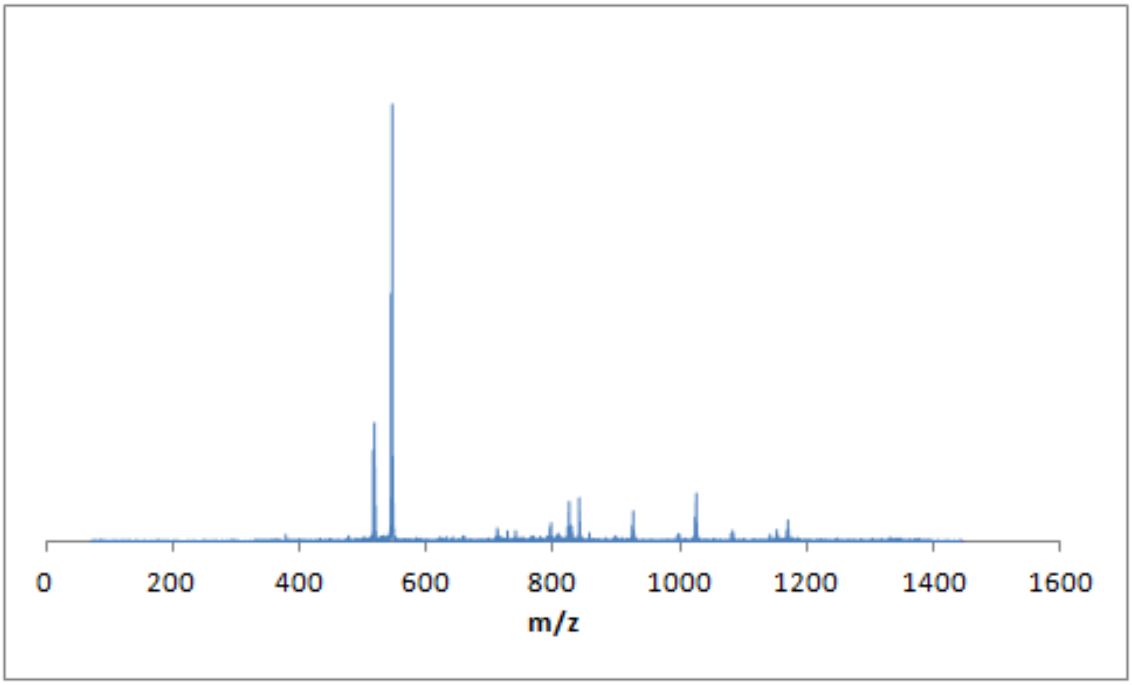
The MS/MS spectrum of compound [PADA-Re(CO)$_3$]-FEGGEESY-NH$_2$ (**4.4**) and b series peak assignments
$[M+H]^+$ calc: 1392.40, obs: 1392.37 (rhenium-187 peak)

Series	Label	Fragment	Peak (*m/z*)	Difference between peaks (*m/z*)	Amino Acid Assignment
b series	b_0	NOT OBSERVED	475.0		
	b_1	RE-PADA-F	622.1	147.1	Phenylalanine
	b_2	RE-PADA-FE	751.1	129	Glutamic Acid
	b_3	RE-PADA-FEG	808.2	57.1	Glycine
	b_4	RE-PADA-FEGG	865.2	57	Glycine
	b_5	RE-PADA-FEGGE	994.2	129	Glutamic Acid
	b_6	RE-PADA-FEGGEE	1123.2	129	Glutamic Acid
	b_7	RE-PADA-FEGGEES	1210.3	87.1	Serine
	b_8	NOT OBSERVED	1373.3	163	Tyrosine



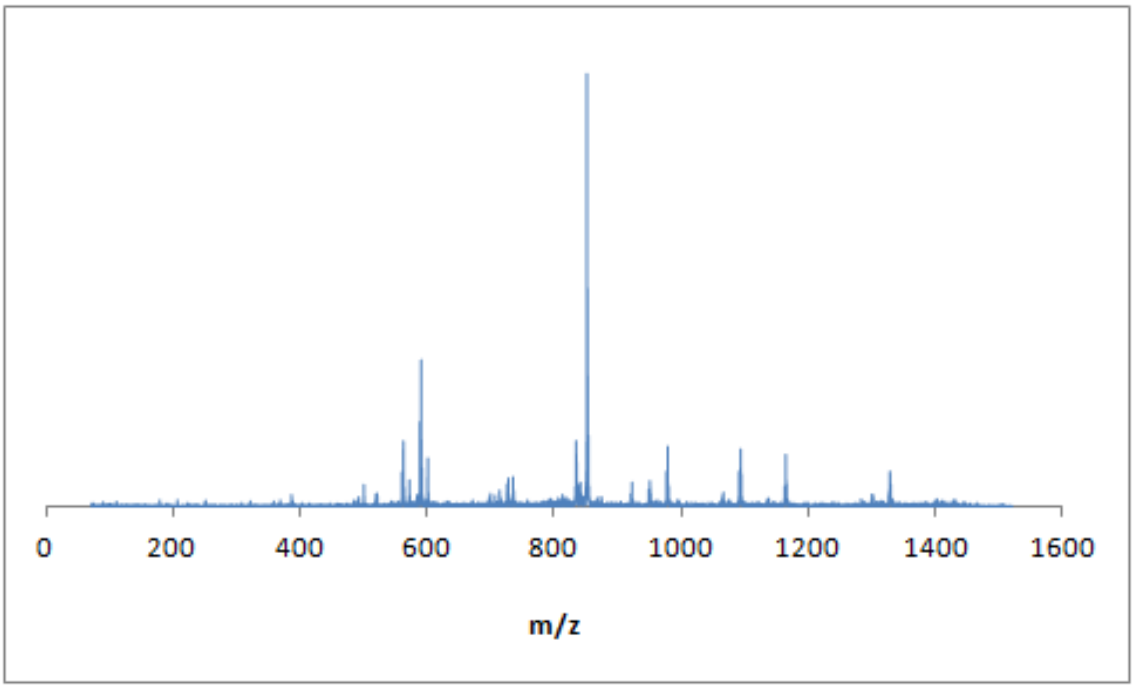
The MS/MS spectrum of compound [PADA-Re(CO)$_3$]-YARES(K/Q)AD-NH$_2$ (**4.5**) and b series peak assignments
$[M+H]^+$ calc: 1414.47, obs: 1414.45 (rhenium-187 peak)

Series	Label	Fragment	Peak (*m/z*)	Difference between peaks (*m/z*)	Amino Acid Assignment
b series	b_0	NOT OBSERVED	475.0		
	b_1	RE-PADA-Y	638.1	163.1	Tyrosine
	b_2	RE-PADA-YA	709.2	71.1	Alanine
	b_3	RE-PADA-YAR	865.3	156.1	Arginine
	b_4	RE-PADA-YARE	994.3	129.1	Glutamic Acid
	b_5	RE-PADA-YARES	1081.2	86.9	Serine
	b_6	RE-PADA-YARES(K/Q)	1209.4	128.1	Lysine/Glutamine
	b_7	RE-PADA-YARES(K/Q)A	1280.4	71.1	Alanine
	b_8	NOT OBSERVED	1395.4	115.0	Aspartic Acid



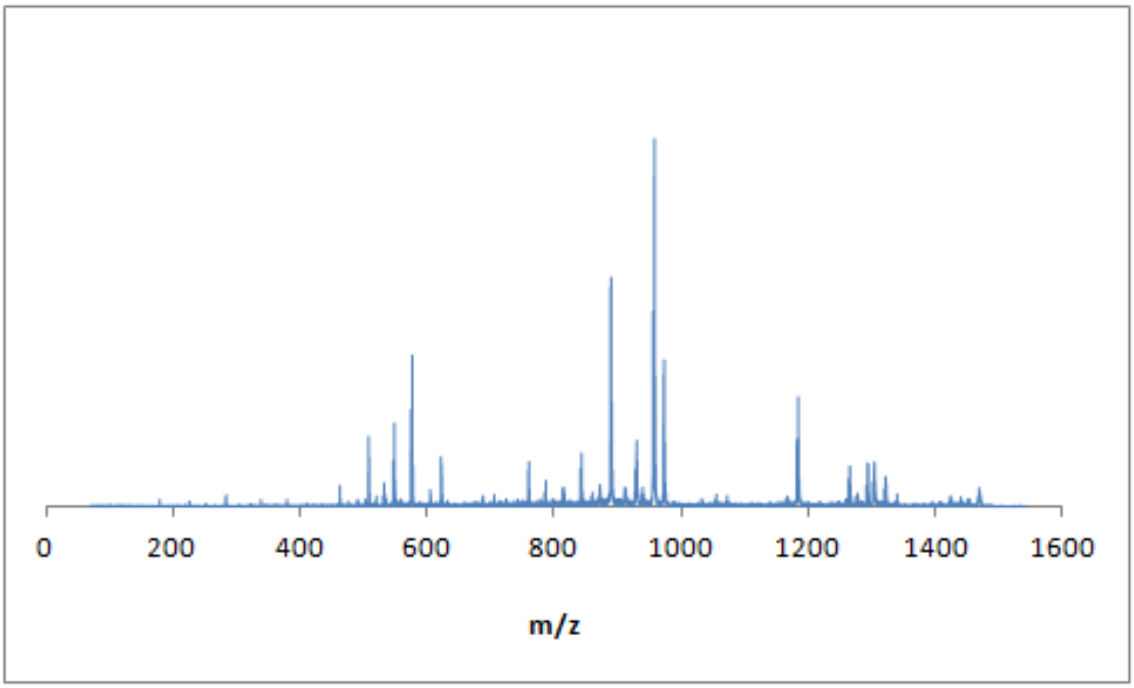
The MS/MS spectrum of compound [PADA-Re(CO)$_3$]-DHHLRAYV-NH$_2$ (**4.6**) and b series peak assignments
$[M+H]^+$ calc: 1443.48, obs: 1443.46 (rhenium-187 peak)

Series	Label	Fragment	Peak (*m/z*)	Difference between peaks (*m/z*)	Amino Acid Assignment
b series	b_0	NOT OBSERVED	475.0		
	b_1	RE-PADA-D	590.1	115.1	Aspartic Acid
	b_2	RE-PADA-DH	727.1	137.1	Histidine
	b_3	RE-PADA-DHH	864.6	137.5	Histidine
	b_4	RE-PADA-DHHL	977.3	112.7	Leucine
	b_5	RE-PADA-DHHLN	1091.3	114.0	Asparginine
	b_6	RE-PADA-DHHLNA	1162.3	71.0	Alanine
	b_7	RE-PADA-DHHLNAY	1325.5	163.1	Tyrosine
	b_8	NOT OBSERVED	1424.4	99.0	Valine



The MS/MS spectrum of compound [PADA-Re(CO)$_3$]-TEHNP(K/Q)HE-NH$_2$ (**4.7**) and b series peak assignments
[M+H]$^+$ calc: 1466.44, obs: 1446.48 (rhenium-187 peak)

Series	Label	Fragment	Peak (*m/z*)	Difference between peaks (*m/z*)	Amino Acid Assignment
b series	b_0	NOT OBSERVED	475.0		
	b_1	RE-PADA-T	576.1	101.1	Threonine
	b_2	RE-PADA-TE	705.2	129.1	Glutamic Acid
	b_3	RE-PADA-TEH	843.2	137	Histidine
	b_4	RE-PADA-TEHN	956.2	114	Asparginine
	b_5	RE-PADA-TEHNP	1053.3	97.1	Proline
	b_6	RE-PADA-TEHNP(K/Q)	1181.4	128.1	Lysine/Glutamine
	b_7	RE-PADA-TEHNP(K/Q)H	1318.4	137	Histidine
	b_8	NOT OBSERVED	1447.5	129.1	Glutamic Acid



Cirriculum Vitae

Dana Roland Cruickshank

Education

- The University of Western Ontario
  London, Ontario, Canada
  2010-2012 M.Sc. in Chemistry

- The University of Waterloo
  Waterloo, Ontario, Canada
  2006-2010 B.Sc. in Chemistry
  Dean's Honours

Work Experience

Teaching Assistant: Sept. 2010 – Dec. 2010
Sept 2011 – Apr. 2012

Scholarships and Awards

UWO Oncology Travel Award

CaRTT Trainee (2010-2011)

Western Graduate Research Scholarship - Chemistry

Queen Elizabeth II Aiming for the Top Scholarship

University of Waterloo President's Scholarship

Upper Year Chemistry Scholarship

Conference Presentations

Cruickshank, D.R., Luyt, L.G. "Organometallic OBOC Peptide Libraries for Molecular Imaging". *95th Canadian Chemistry Society Conference*, Calgary, Canada [Oral Presentation]