Supramolecular interactions of methylated amino acids: Investigations using small molecule aromatic cage mimics

by

Amanda Lee Whiting
B.Sc., University of Victoria, 2007

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Chemistry

Supervisory Committee

Supramolecular interactions of methylated amino acids: Investigations using small molecule aromatic cage mimics

by

Amanda Lee Whiting
B.Sc., University of Victoria, 2007

Supervisory Committee

Dr. Fraser Hof, Department of Chemistry
Supervisor

Dr. Jeremy Wulff, Department of Chemistry
Departmental Member

Dr. Robin Hicks, Department of Chemistry
Departmental Member

Dr. Alisdair Boraston, Department of Biochemistry and Microbiology
Outside Member

Abstract

Supervisory Committee
Dr. Fraser Hof, Department of Chemistry
Supervisor
Dr. Jeremy Wulff, Department of Chemistry
Departmental Member
Dr. Robin Hicks, Department of Chemistry
Departmental Member
Dr. Alisdair Boraston, Department of Biochemistry and Microbiology
Outside Member

The recognition of modified amino acids by reader proteins is governed by the competing interplay of weak, attractive, intermolecular forces and solvation effects. For the recognition of hydrophobic cations like methyl-lysines and methyl-arginines, native reader proteins utilize structural cages always containing multiple aromatic amino acids and sometimes an occasional acidic residue. Through the highly ordered arrangement of multiple aromatic surfaces, reader proteins can invoke the attractive forces of electrostatic, cation-pi, and in the case of arginine, pi-pi interactions. The hydrophobic effect can also significantly affect these binding events in aqueous environments.

In this thesis, a number of small molecule, synthetic cages containing significant aromatic surface area have been synthesized. Variation in both total host hydrophobicity and degree of flexibility were explored to determine what effect they have on the overall binding of methylated amino acids in water. Significant flexibility in the first generation of highly aromatic hosts was shown to be detrimental to binding. However, strong binding was observed for guests with significant hydrophobic character despite this flexibility. The cause of the strong affinities in this family of synthetic cages was shown to be due to the hydrophobic effect, rather than any attraction due to cation-pi interactions.

Synthetic efforts towards hosts with more rigid structures led to the use of Tröger's base as a structural building block. Hosts incorporating Tröger's bases into well-defined aromatic cavities were found to exhibit strong binding to both methyl-lysine and methyl-arginine derivatives in pure water. Differences in guest selectivity were due to the rigid altered host geometry introduced by the Tröger's base cleft.

Table of Contents

List of Tables

List of Figures

List of Schemes

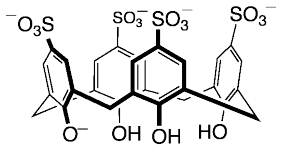
List of Abbreviations

AChCl	acetylcholine chloride
ACN	acetonitrile
aRme2	asymmetric dimethyl arginine, aDMA
ATR	attenuated total reflectance
B_2pin_2	bis(pinacolato)diboron
Boc	*tert*-butoxycarbonyl
CBX7	chromobox homolog 7
DCM	dichloromethane
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNMT3a	DNA (cytosine-5)-methyltransferase 3A
ESI	electrospray ionization
ESP	electrostatic potential
Fmoc	fluorenylmethyloxycarbonyl
HAT	histone aceyltransferase
HDAC	histone deacetylase
HDM	histone demethylase
HF	Hartree-Fock
HMT	histone methyltransferase
HP1	heterochromain protein 1
HPLC	high pressure liquid chromatography
HR-EIMS	high resoution electron impact mass spectrometry
HR-ESI-MS	high resolution electrospray ionization mass spectrometry
IR	infrared spectroscopy
JMJD2A	jumonji C domain-containing 2A
K	lysine
Kac	acetylated lysine
Kme	monomethyllysine
Kme2	dimethyllysine
Kme3	trimethyllysine
LAH	lithium aluminum hydride
LR-EIMS	low resolution electron impact mass spectrometry
MBT	malignant brain tumor
NMR	nuclear magnetic resonance spectroscopy
ORTEP	Oak Ridge Thermal-Ellipsoid Plot
PADI	protein arginine deiminase
PDB	protein data bank
PHD	plant homeobox domain
PM3	parameterized model number 3
ppm	parts per million
PRC2	polycomb repressive complex 2
PRMT	protein arginine methyltransferase
PSC	*para*-sulfonatocalix[4]arene
PTM	post-translational modification
R	arginine
RBF	round bottom flask

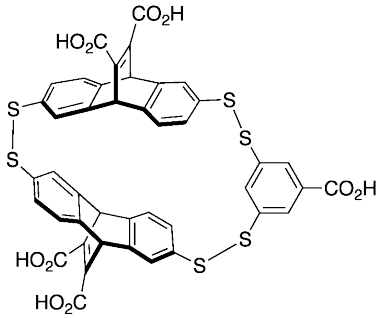
Rme	monomethyl arginine, MMA
ROESY	rotating frame Overhauser effect spectroscopy
SMN	survival of motor neuron
sRme2	symmetric dimethyl arginine, sDMA
TB	Tröger's base
TDRD3	Tudor domain-containing protein 3
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin layer chromatography
TMS	tetramethylsilane
UHRFI	ubiquitin-like, PHD and RING finger containing protein 1
WD40	tryptophan-aspartic acid repeat domains
WDRD5	WD repeat domain 5

List of Compounds

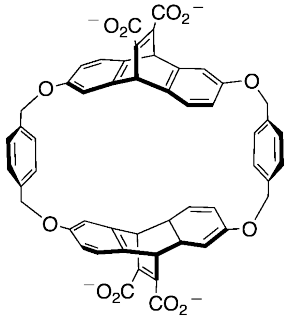
Compound **1.1**



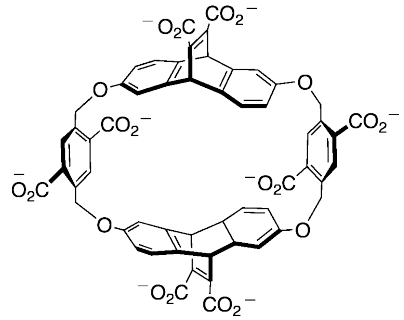
Compound **1.2**



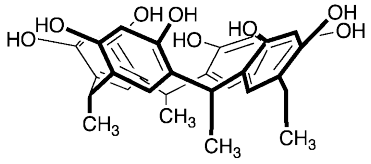
Compound **1.3**



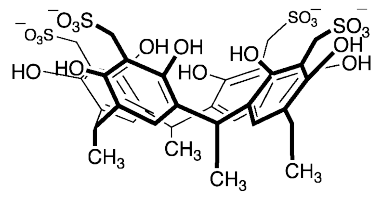
Compound **1.4**



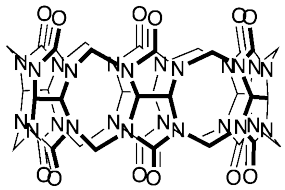
Compound **1.5**



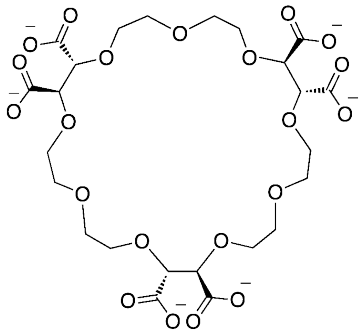
Compound **1.6**



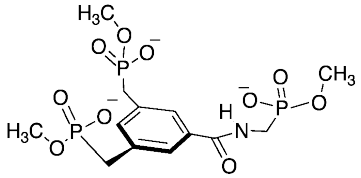
Compound **1.7**



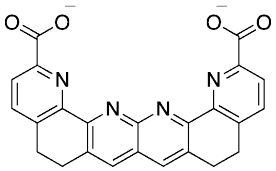
Compound **1.8**



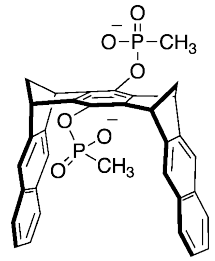
Compound **1.9**



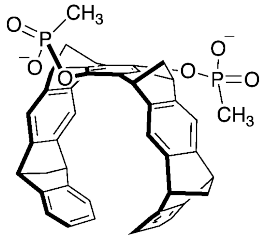
Compound **1.10**



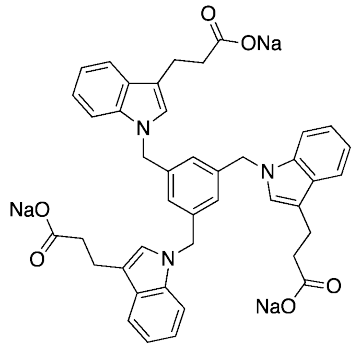
Compound **1.11**



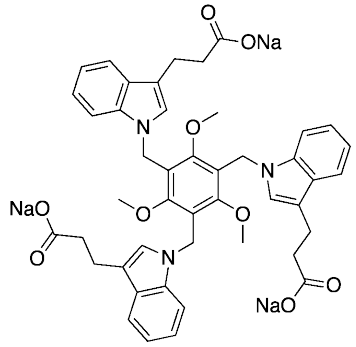
Compound **1.12**



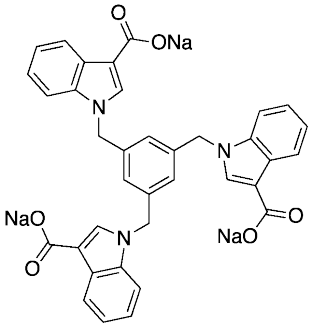
Compound **2.1**



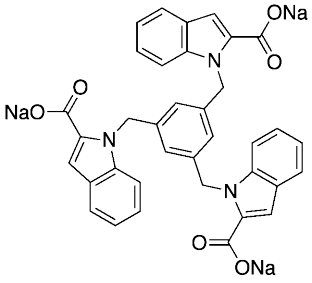
Compound **2.2**



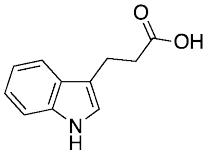
Compound **2.3**



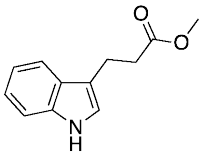
Compound **2.4**



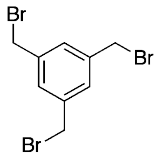
Compound **2.5**



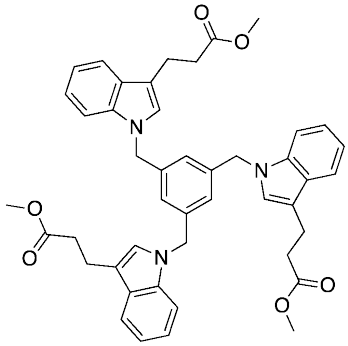
Compound **2.6**



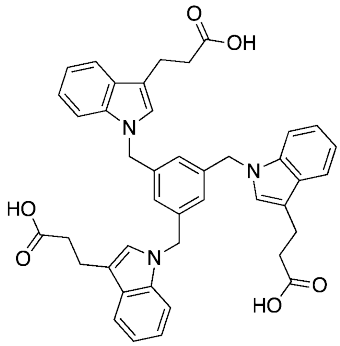
Compound **2.7**



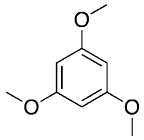
Compound **2.8**



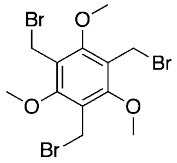
Compound **2.9**

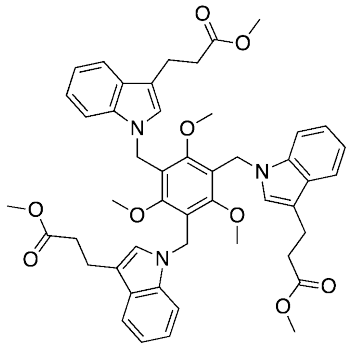


Compound **2.10**

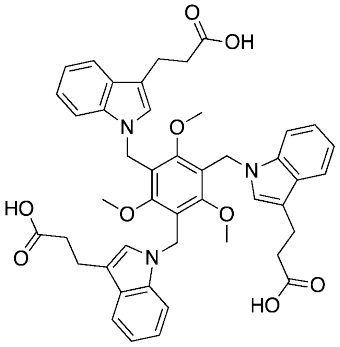


Compound **2.11**

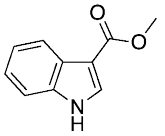




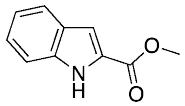
Compound **2.12**



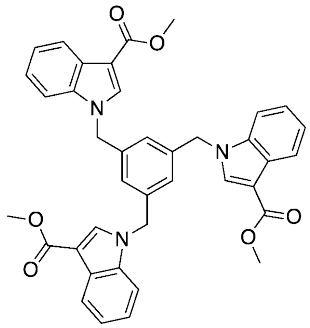
Compound **2.13**



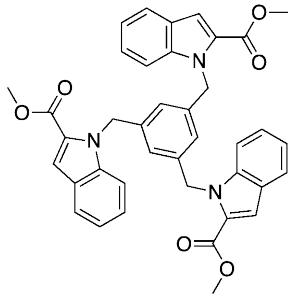
Compound **2.14**



Compound **2.15**

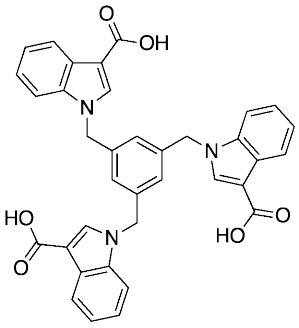


Compound **2.16**

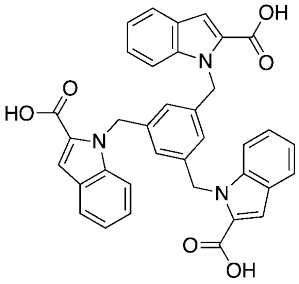


Compound **2.17**

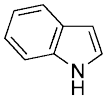
Compound **2.18**



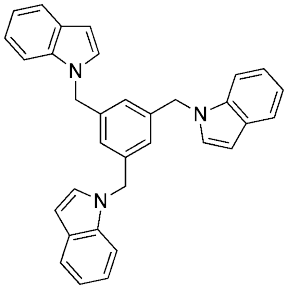
Compound **2.19**



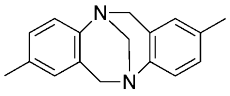
Compound **2.20**



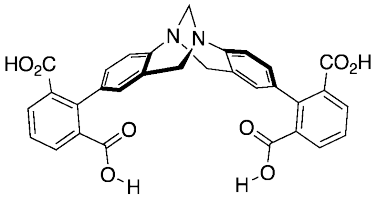
Compound **2.21**



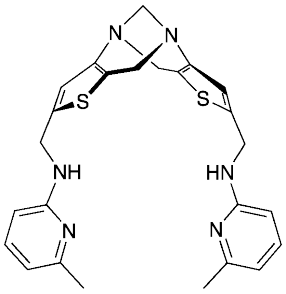
Compound **3.1**



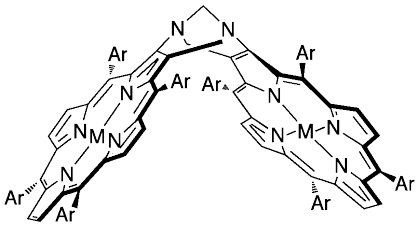
Compound **3.2**



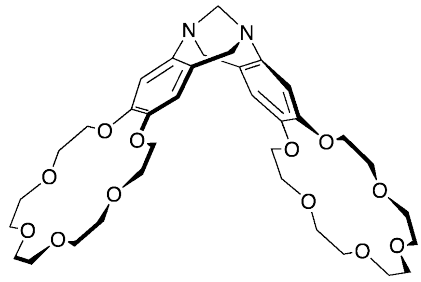
Compound **3.3**



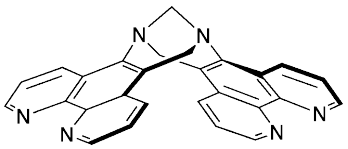
Compound **3.4**



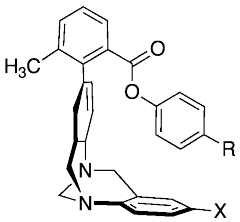
Compound **3.5**



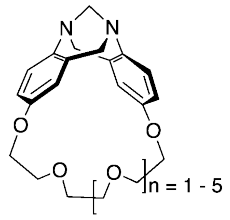
Compound **3.6**

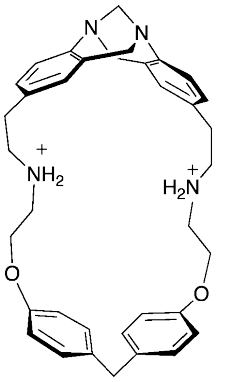


Compound **3.7**

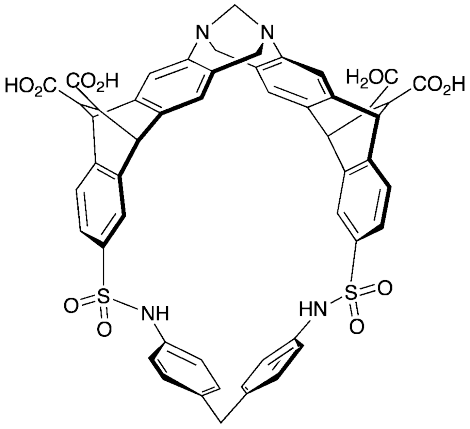


Compound **3.8**

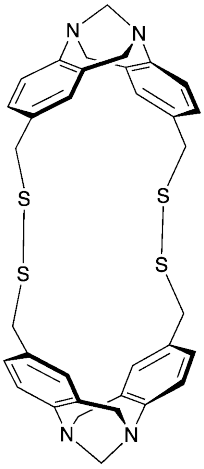




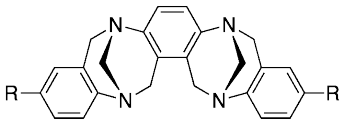
Compound **3.9**



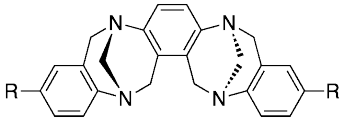
Compound **3.10**



Compound **3.11**

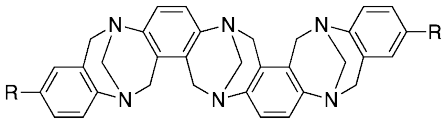


Compound **3.12**

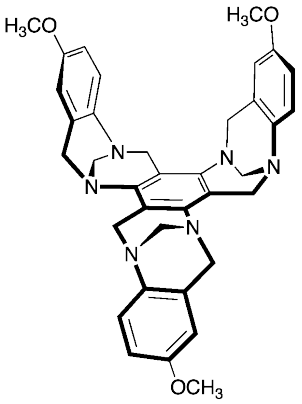


Compound **3.13**

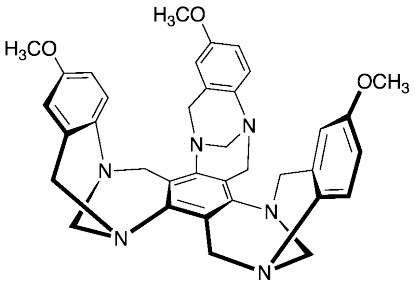
Compound **3.14**



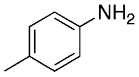
Compound **3.15**



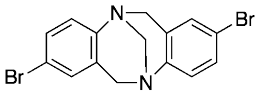
Compound **3.16**



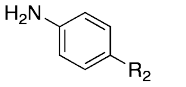
Compound **3.17**



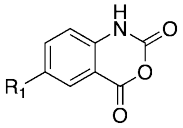
Compound **3.18**



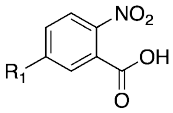
Compound **3.19**



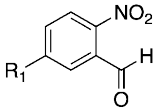
Compound **3.20**



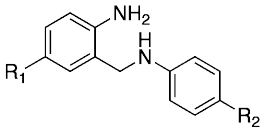
Compound **3.21**



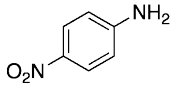
Compound **3.22**



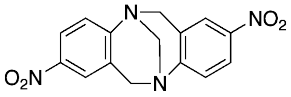
Compound **3.23**



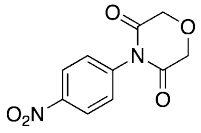
Compound **3.24**



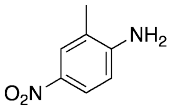
Compound **3.25**



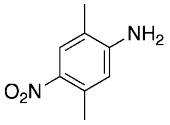
Compound **3.26**



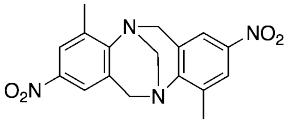
Compound **3.27**



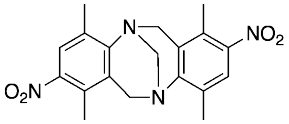
Compound **3.28**



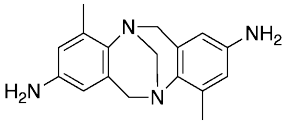
Compound **3.29**



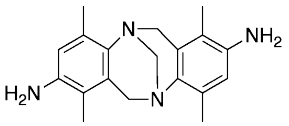
Compound **3.30**



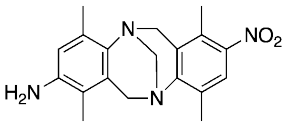
Compound **3.31**



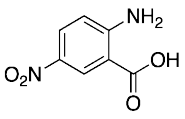
Compound **3.32**



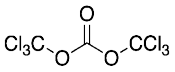
Compound **3.33**



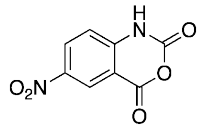
Compound **3.34**



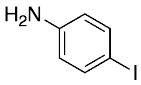
Compound **3.35**



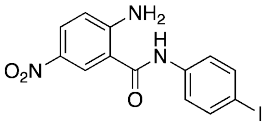
Compound **3.36**



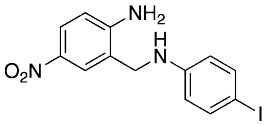
Compound **3.37**



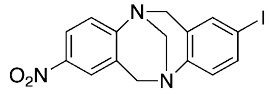
Compound **3.38**



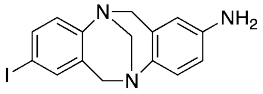
Compound **3.39**



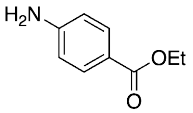
Compound **3.40**



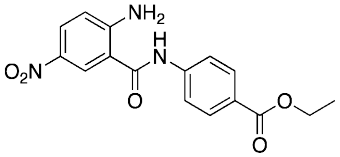
Compound **3.41**



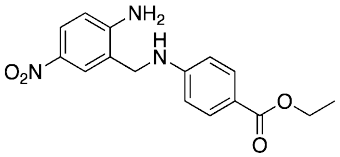
Compound **3.42**



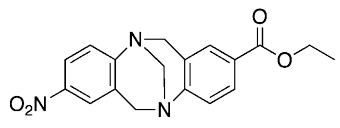
Compound **3.43**



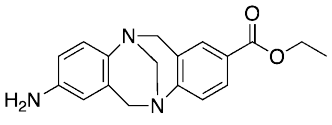
Compound **3.44**



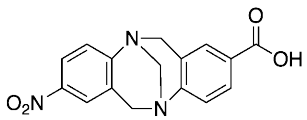
Compound **3.45**



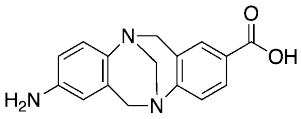
Compound **3.46**



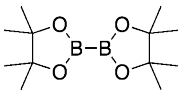
Compound **3.47**



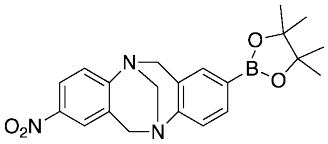
Compound **3.48**



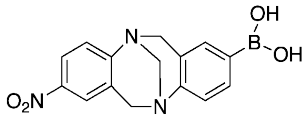
Compound **3.49**



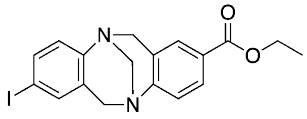
Compound **3.50**



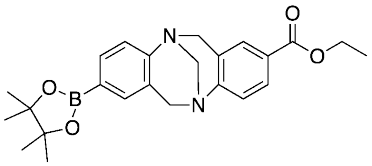
Compound **3.51**



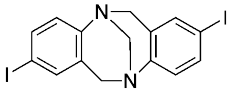
Compound **3.52**



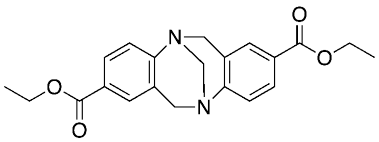
Compound **3.53**



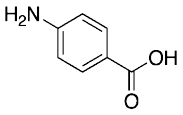
Compound **3.54**



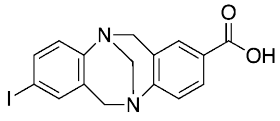
Compound **3.55**



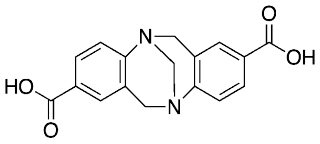
Compound **3.56**



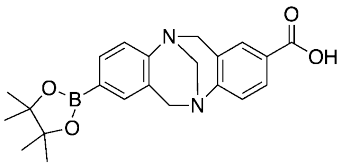
Compound **3.57**



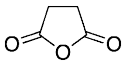
Compound **3.58**



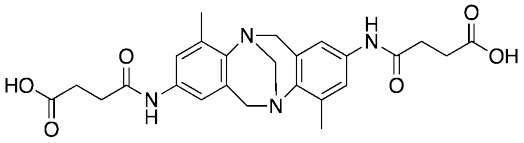
Compound **3.59**



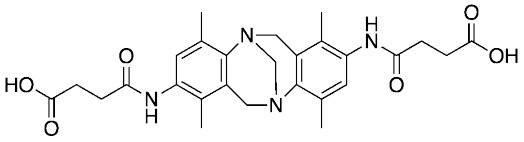
Compound **4.1**



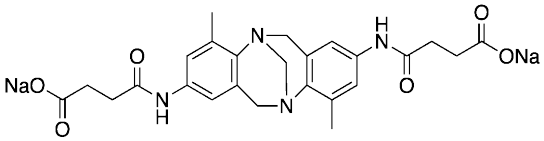
Compound **4.2**



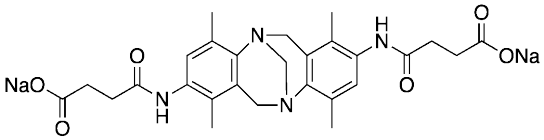
Compound **4.3**



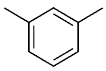
Compound **4.4**



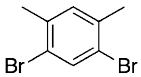
Compound **4.5**



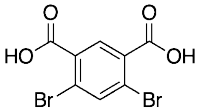
Compound **4.6**



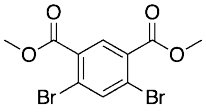
Compound **4.7**



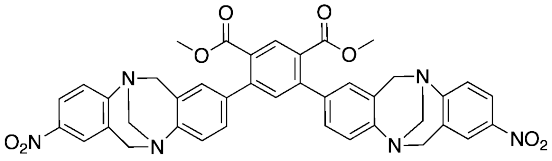
Compound **4.8**



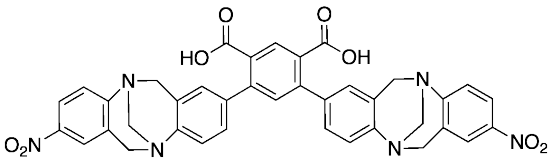
Compound **4.9**



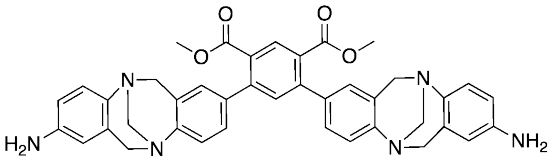
Compound **4.10**



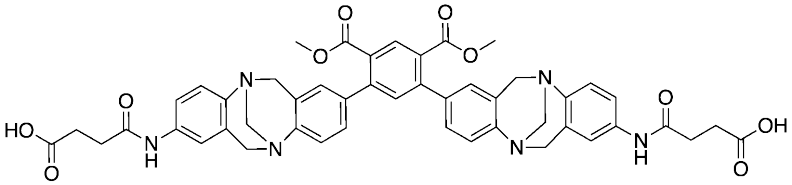
Compound **4.11**



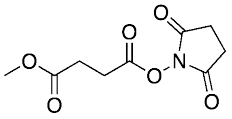
Compound **4.12**



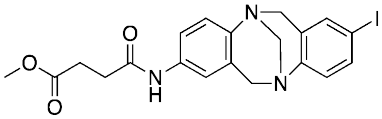
Compound **4.13**



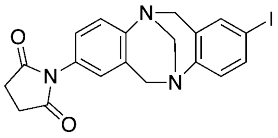
Compound **4.14**



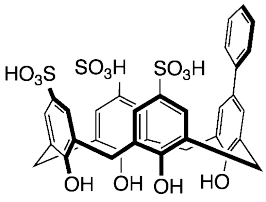
Compound **4.15**



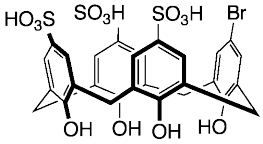
Compound **4.16**



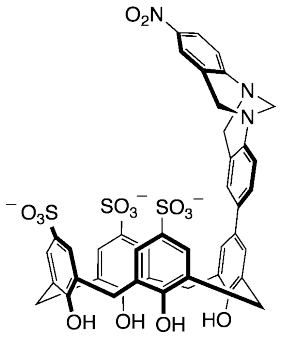
Compound **4.17**



Compound **4.18**



Compound **4.19**



Acknowledgments

A thesis is not produced solely by oneself, and I have many people to thank for their help, guidance, thoughts and support. Foremost, I must thank my supervisor, Fraser Hof. As an undergraduate and then a graduate student, he has guided and influenced my path as a scientist, combining my interests in the art of organic synthesis, the exactness of analytical determination, and the relevance of anything biological. I appreciate Fraser for the freedom I've had in his lab, and for the deliberate nudge when needed when I've had "too many irons in the fire" (which was more often than not). I am grateful for the opportunities I've had as a student in this lab: chances to travel, present and network at national and international conferences, as well as those more close to home (but fondly remembered nonetheless).

I am grateful for the technical support I have received in the Department of Chemistry. To Christine Greenwood and Chris Barr – for their conversation, advice and instruction in all things NMR. The basement was my second home over the course of this degree. To Ori Granot – for all the mass specs a girl could want. And to Sean Adams – for whom there was nothing I could break that he could not fix.

I am also indebted to my peers and fellow students, for advice, for moral support and for willing participants to head all the way across campus for good coffee. To Hof group members, past and present – I am honoured to have worked with you all and to have seen where our little group has gone. It will only get better from here. To the B-wing lab mates – I am as grateful to you for your time in conversation about current schemes and failed reactions, as I am for your seemingly endless supply of chemicals to test out new ideas. And to the undergraduates I have worked with or taught – it is always a pleasure, and there are no dumb questions.

And to those who have made this journey with me from the very beginning – the time to finish is now.

Dedication

~ To Chris ~

It was you.

Chapter 1 – Introduction

1.1 Prologue

Studying how and why two molecules interact with each other is a fundamental step required for understanding and influencing biological processes. Much of life as we know it results from the recognition, association and specific interaction of two or more molecules. These molecular systems – be they of cells, proteins, individual molecules, or a combination thereof – serve to regulate and drive biological systems and pathways. The science of molecular recognition involves studying the requirements of size, shape, and chemical complementarity of interacting partners. Recognition results from a fluid interplay of weak, reversible, and non-covalent interactions, balancing out in terms of the strength of the attractive vs. repulsive interaction(s) and other subtle factors (geometry, solvent, polarity, etc).

The study of molecular recognition in biological systems brings together the fields of biochemistry, synthetic chemistry and physical organic chemistry. At this junction, chemical knowledge of non-covalent interactions and thermodynamics can be used to investigate and understand fundamentals of biological processes. Supramolecular chemistry involves the study of systems of discrete molecules, and tries to explain how and why these molecules associate and interact with one another in specific ways. When supramolecular systems are studied under physiological conditions (in water at near neutral pH values), it becomes possible to ask and answer questions about biological systems. *The challenge then for supramolecular chemists is to design and build molecules that mimic and utilize aspects of molecular recognition from biology, and use these molecules to gain insight into the basic requirements of structure and function in biological settings*.

This thesis is primarily focused on applying the concepts of supramolecular recognition to the binding of specific, biologically important molecules – namely, methylated cationic amino acids. We were interested in what the primary driving forces of these interactions were, and what could be done to modify the strength, selectivity (among multiple analytes) and specificity (for a single analyte) of the “host” partner. In the following sections of this chapter, I have attempted to outline key concepts behind the studies described in chapters 2 to 4. These sections include discussions of a) the importance of post-translational amino acid modifications, b) the structure and function

of native binding partners for these modified amino acids, c) the non-covalent forces used by the native hosts, and d) current examples of synthetic partners for these guest molecules.

1.2 Post-translationally modified amino acids

There exists a great wealth of additional information available to a cell beyond the twenty naturally occurring amino acids that can be encoded by the transcription and translation of DNA into functional protein products. This information is contained in a number of post-translational modifications (PTMs) that can occur on proteins. At their simplest, protein modifications change the size, shape, connectivity, charge state and/or hydrophobicity of individual amino acid side chains. Consequently, the criteria for the molecular recognition of those amino acid motifs is also changed, leading to (potentially) different biological outcomes as different enzymes, co-factors, and/or other proteins are then recruited or repelled.

We were most interested in PTMs that occur on DNA-packaging histone proteins since they are directly involved in regulating transcription of the associated DNA. Histone proteins form the cores around which DNA is wound and packaged to form nucleosomes, which are subsequently bundled into chromatin fibers and condensed into chromosomes, allowing the entire length of DNA to fit within the nucleus of a cell. The core histone proteins are found as octamers, linked together by additional histone units which secure the DNA and prevent it from unravelling. Each octamer contains two copies each of histones from the H2A, H2B, H3 and H4 families. Histones H3 and H4 have long, unstructured peptide tails that extend outwards from the packaged DNA units into the solvent, making them accessible to other proteins. PTMs of specific amino acids along these tails are known to regulate gene expression and DNA repair, among others processes.[1,2] Prominent histone tail PTMs include methylation (of lysine and arginine), acetylation (of lysine), phosphorylation (of serine and threonine), ubiquitylation and SUMOylation (small ubiquitin-like modifier proteins) (of lysine), and amide bond cis-trans isomerization (of proline). A summary of PTMs for the H3 and H4 tails is shown in Figure 1.1.[3-5]

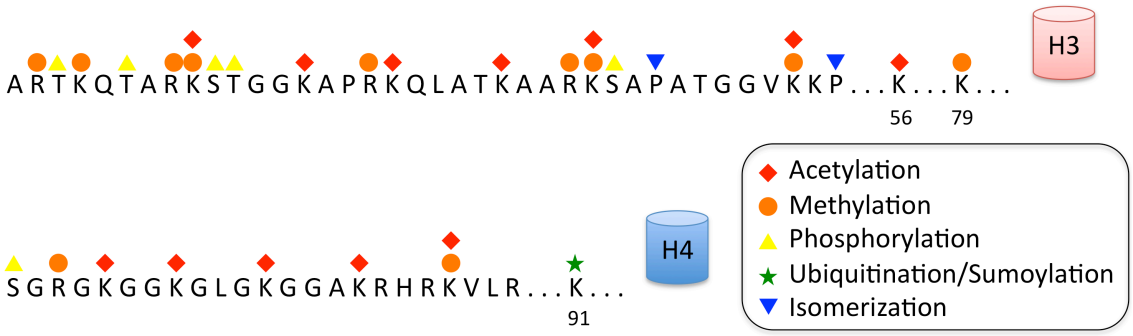


Figure 1.1 Select PTMs along Histone 3 (H3) and Histone 4 (H4) unstructured tails (indicating methylation, acetylation, phosphorylation, ubiquitylation/sumoylation and proline isomerization)

The major function of histone PTMs is either to create or obstruct sites for the binding of specific protein partners. These modifications alter the expression states of associated DNA, thus enabling gene up- or down-regulation. While individual histone PTMs (or marks) can cause specific outcomes, multiple modifications on a single histone or highly similar modifications in different positions along the histone can cause different, even opposite effects. This forms the basis of the "histone code hypothesis" whereby combinations of PTMs to histone proteins are indicators of different gene regulation outcomes.[6] To understand how these often modest structural modifications to specific amino acids can have such control over biomolecular events, it is critical to understand the underlying non-covalent interactions involved in their recognition.

1.3 Lysine and arginine: modified similarly, recognized similarly

For this thesis, I will focus on two specific amino acids: lysine (K) and arginine (R), and related cationic species. Both amino acids are cationic at physiological pH and can undergo covalent modifications that change the size, shape, charge and hydrophobicity of the exposed side chain. Importantly, both are present on histone tails and can be modified, contributing to a cell's gene expression.

1.3.1 Lysine PTMs

Common PTMs to histone lysines are the covalent additions of small chemical groups such as acetylation and methylation. Larger functionalities have also been found on histones lysines through biotinylation,[7,8] ribosylation,[9] and sumoylation/ubiquitination.[10,11]

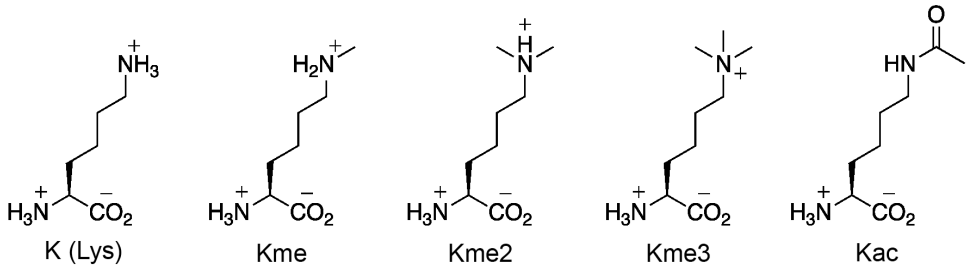


Figure 1.2 Methylation and acetylation states of lysine

One of the most well-studied histone modifications is lysine methylation. As shown in Figure 1.2, lysine can be methylated multiple times resulting in monomethyl- (Kme), dimethyl- (Kme2) and trimethyl-lysine (Kme3). The pKa of the ammonium side chain of lysine is 10.67;[12] it exists as a cationic ammonium ion at physiological pH. Methylation increases the size and hydrophobic character of the lysine head group. This does not change the cationic nature of the ammonium but does allow the charge to be spread over a larger area (Figure 1.3). The products of stepwise methylation are progressively more hydrophobic cations that are less well-solvated by water. Lysine can also undergo PTM via acetylation (to Kac), which neutralizes the cationic charge altogether.

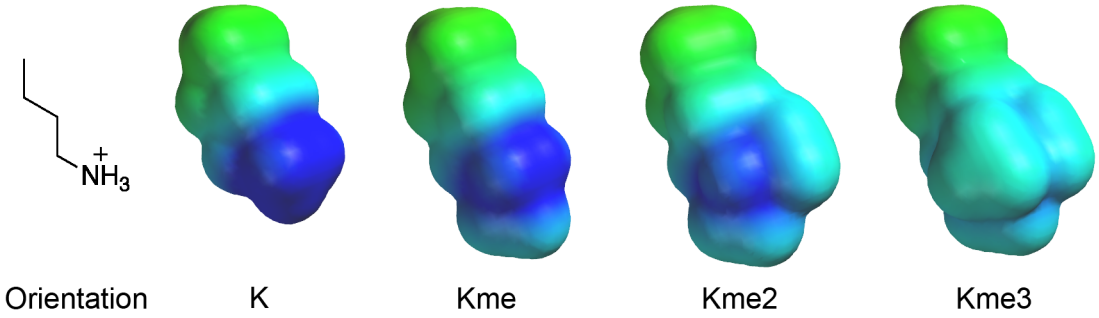


Figure 1.3 Electrostatic potential (ESP) maps of lysine side chain analogs with stepwise methylation. (Spartan '10: PM3; scale: -50 to 600 kJ mol^{-1}).[13] Chemdraw structure included to show the orientation of the side chain. Areas of red would indicate higher electron density (negative surface potential), while areas of blue indicate lower electron density (positive surface potential).

A key aspect of the "histone code" is the ability for these PTM marks to be installed *and* removed – reversibility allows for dynamic control over a cell's gene expression via on/off modifications. Enzymes that install the lysine methylation mark are known as histone lysine methyltransferases (HMTs). Histone methylation marks are highly specific – their downstream effect on the gene expression is dependent on both

the site and extent of methylation. Different enzymes are responsible for the extent of methylation (mono, di or tri) as well as which lysine is modified (H3K4, H3K9, H3K27, etc). Originally, lysine methylation was thought to be a terminal pathway – a mark which, once installed, required the histone to be exchanged or the tail cleaved before the signal could be removed.[14] However, the discovery of lysine demethylases (HDMs) in 2004[15] gave renewed interest to the idea that lysine methylation could form a part of a dynamic and reversible histone code. Currently, over thirty lysine methyltransferases and twenty demethylases are known.[4] While these enzymes represent a great number of potential targets themselves, for the purpose of this thesis, we will simply consider them the writers and erasers of an interesting PTM mark, methyl-lysine.

Similarly, lysine can be acetylated via lysine acyltransferases (or histone acyltransferases, HATs) and deacetylated via histone deacetylases (HDACs). Currently, nineteen lysine acyltransferases and eighteen histone deacetylases (HDACs) have been discovered,[4] further emphasizing the reversibility of histone lysine marks.

1.3.2 Arginine PTMs

Similar to the small covalent additions to lysine, arginine can undergo methylation as well as another PTM, citrullination (also called deimination). Arginine methylation can occur once (monomethyl arginine; MMA or Rme) or twice, both symmetrically with the methyl groups on opposite nitrogen atoms (symmetric dimethyl arginine; sDMA or sRme2) and asymmetrically on the same nitrogen (asymmetric dimethyl arginine; aDMA or aRme2) (Figure 1.4). The pKa of the arginine head group is 12.10,[12] which again results in a cationic side chain at physiological pH. As with lysine, arginine methylation results in a larger, more hydrophobic guest that retains its cationic charge. Citrullination (or deimination) removes the cationic charge, by converting the guanidinium head group into a neutral urea (Cit, Figure 1.4).

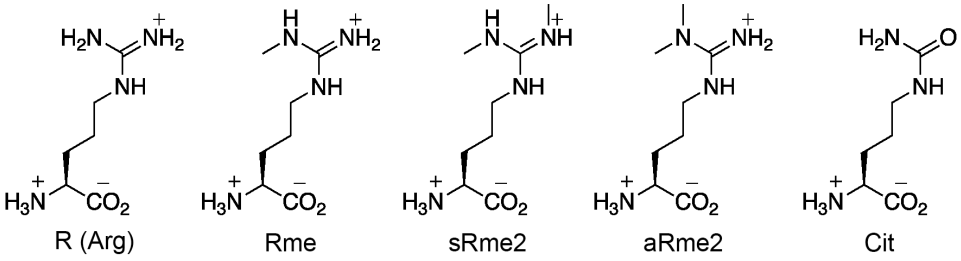


Figure 1.4 Methylation and deimination states of arginine

Although histone arginines were known to contain methyl groups as early as 1967,[16] interest in this PTM has only increased in the recent past.[17] As with other PTMs, the methylation of arginine residues in proteins creates binding sites for recognition by methyl-arginine-binding domains,[18] or sterically hinders other proteins from binding to neighbouring PTM sites,[19] and can thereby modulate protein–protein interactions and their physiological outcomes. Arginine methylation is catalyzed by a family of protein arginine N-methyltransferases (PRMTs). Genes encoding for PRMTs were identified in 1996,[20] and the first and only enzyme to demethylate histone arginines in humans was identified in 2007.[21] Protein arginine deiminases (PADIs) are a family of enzymes that remove arginines by deimination to produce citrulline,[22] It was originally suggested that PADIs could also convert monomethyl arginine to citrulline,[23,24] however more recent studies have shown that methylation of arginine residues blocks the conversion to citrulline.[25-27] While the existence of an eraser enzyme that directly removes the methyl marks down to a bare arginine remains disputed,[28] it is clear that there is biological machinery in place that signals via production and removal of methylated arginines. It is also clear that arginines have growing significance as residues that regulate, activate and/or suppress biological functions based on their dynamic methylation state.

While lysine and arginine have many similar properties – charge and the ability to be methylated, for example – there are chemical differences between them that should be noted – namely, their shape and, in the case of arginine, delocalized pi electronic character. Because of the ammonium head group, lysine can be thought of as a round, sphere-like cation, which presents a similar shape and charge density from many angles (Figure 1.5, a). In contrast, the guanidinium head group of arginine is a larger, flat cation with a delocalized pi system (Figure 1.5, b). The face of guanidinium is postulated to be slightly hydrophobic[29] relative to its peripheral hydrogen atoms, further influencing its behaviour as a binding partner. Synthetic hosts and proteins that recognize this flat, hydrophobic cation will have different properties than those optimized for recognition of the ball-like lysine cation.

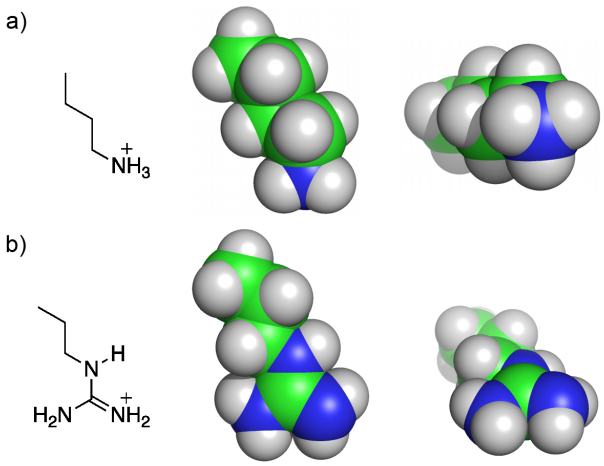


Figure 1.5 The shape and (lack of) delocalized character of the amino acid side chains of lysine and arginine are key differences that allow for their discrimination. Side chain analogs of a) lysine and b) arginine showing both a front and side view (Spartan '10: PM3).[13] Chemdraw structures included for orientation.

1.4 Natural binding partners of methylated lysine and arginine

Once the methylation marks described above have been installed on either lysine or arginine, the "code" that they comprise must be "read" in order to influence further cell processes. The proteins that recognize and bind to the PTM marks are often referred to as "readers". Below are a few examples of both methyl-lysine and methyl-arginine reader proteins. Further on, we will examine what structural aspects of these examples make them excellent binding partners for their respective PTMs.

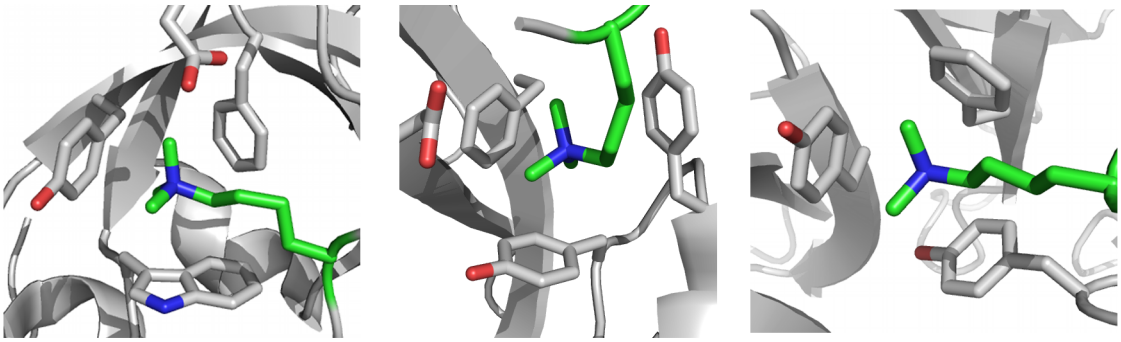
1.4.1 Methyl-lysine binders

As noted above, methyl-lysine can exist in three possible states: Kme, Kme2 and Kme3. Each of these methylated states are selectively recognized by a number of binding partners (both readers and erasers). Often the particular methylated lysine that is targeted depends on both the methylation state *and* on the state and identity of the amino acids in the surrounding area. A selection of histone methyl-lysine reader proteins is presented in Table 1.1.

Table 1.1 Select histone lysine binding proteins, their structural domains, and associated functions

Histone-binding protein	Binding site	Structural domain	Function
HP1	H3K9me3	Chromodomain	Gene silencing[30]
53BP1	H4K20me1/2	Tudor domain	DNA repair factor[31]
JMJD2A	H3K4me3 H4K20me3	Double Tudor domain	Histone demethylase[32]
L3MBTL1	K4K20me1/2	MBT protein	Transcriptional repressor[33]
ING2	H3K4me2/3	PHD protein	Modulates activity of histone deacetylases
PRC2	H3K27me3	WD40-repeat	Methyltransferase[34]
UHRF1	H3K9me3	Tandem Tudor domain	Associated with DNA methylation[35]
TAF1	H4 Kac	Bromodomain	Assembly of transcriptional machinery[36]

Methyl-lysine marks are recognized by protein structural domains that have evolved to have the right balance of characteristics to allow for molecular recognition of their individual targets. These domains can have affinity for more than one methylation state, or be specific for other lysine states such as Kac. We will first examine proteins that bind Kme3 marks – the most extensively methylated lysine. A selection of crystal structures of histone Kme3 peptides bound by reader proteins are shown below (Figure 1.6). From these structures, a common binding mode becomes apparent as the cationic lysine head group is repeatedly surrounded by two to four aromatic amino acids. This general "aromatic cage" motif is able to bind various Kme3-containing peptides with in vitro affinities ranging from 0.1 - 100 μM (10^4 -10^7 M^{-1}).[37-40]



2GFA 2L3R 3JZG

Figure 1.6 Histone Kme3 peptides bound by reader proteins. a) H3K4me3 peptide bound by JMJD2A (a histone demethylase)[32], b) H3K9me3 bound by UHRF1[35] and c) H3K27me3 bound by PRC2 (a methyltransferases).[34] Kme3 of histone tail peptide indicated in green. PDB codes: 2GFA, 2LR3, and 3JZG.

Aromatic cage motifs can also bind to other methyl-lysine states, though the degree to which they will depends on both the cation and the cage structure. In the case of Kme and Kme2, the cation is now smaller, with a less diffuse charge and retains the ability to make hydrogen bonds to water, unlike Kme3. In order to compete with water, cages which are selective for lower methylation states are modified and use acidic residues such as aspartic or glutamic acid to form salt bridges with the remaining hydrogen atoms on the guest.[33,37] The ability to make hydrogen bond contacts to the side chain's hydrogen bond donors (where present) is a critical feature in discrimination between Kme/Kme2 and Kme3.[41]

These cages also make use of steric interactions to bias binding towards smaller cations. Cages which are specific for the smaller Kme and Kme2 cations are often deep in cavities, rather than on the surface. The opening into the cavity containing the aromatic cage can be smaller, and use steric hindrance as a way to bias binding towards smaller cations.[31]

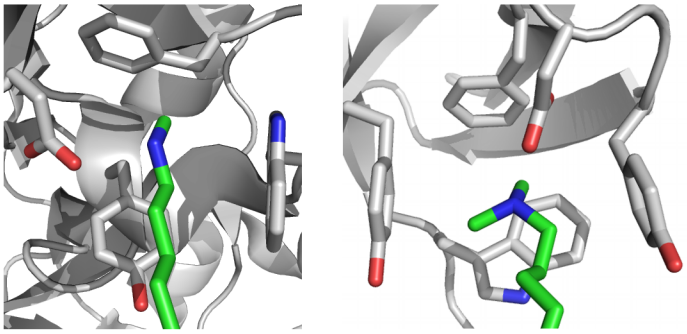


Figure 1.7 Histone Kme and Kme2 peptides bound by reader proteins. a) Kme1 bound by L3MBTL1[33] and b) H4K20me2 bound by 53BP1.[31] Lysine chain of histone peptide in green. PDB codes: 2RHY and 2IG0.

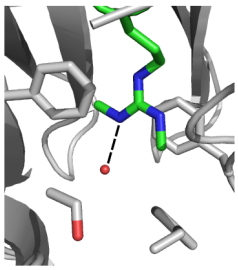
1.4.2 Methyl-arginine binders

Through methylation, arginine also exists in three possible methylated states: Rme, sRme2 and aRme2. Much less is known about histone methyl-arginine readers than their lysine counterparts. Known histone arginine binding proteins are presented below (Table 1.2). The first protein found to recognize arginine methylation marks on histones was described in 2010, when TDRD3 (Tudor domain-containing protein 3, a transcriptional cofactor) was found to bind to strongly to H3R17me2a and H4R3me2a and did not bind to H4R3me2s.[18]

Table 1.2 Histone arginine binding proteins, their structural domains, and associated functions

Histone-binding protein	Binding site	Structural domain	Function
TDRD3	H3R17me2a H4R3me2a	Tudor domain	Transcriptional cofactor[18]
WDRD5	H3R2me2s	WD40-repeat	Coactivator[42]
DNMT3a	H4R3me2s	PHD domain	Repressive, gene silencing[43]

Most recently, WDRD5 was shown to bind strongly to H3R2me2s through its WD40 structural domain. Symmetric dimethylation of H3R2 was shown to increase binding to WDRD5 to K_d = 0.1 μM from K_d = 5.6 μM for the unmethylated peptide.[42] Asymmetric methylation resulted in no detectible binding to WDRD5. This was also the first histone methyl-arginine crystallized with a protein binding partner. The methyl-arginine head group is sandwiched between two phenylalanine residues and participates in a hydrogen bonding network with one water molecule and a serine residue. Comparison to structures of WDRD5 with unmethylated H3 peptides show that one water molecule is displaced from the cavity and that the methyl group of arginine is located closer to an aromatic face.



4AJ7

Figure 1.8 Crystal structure of WDRD5 in complex with R2me2s-histone H3 peptide. Hydrogen bonding indicated by dashed line. (PDB: 4A7J)

More information on the structural requirements of binding both symmetric and asymmetric methyl-arginine can be found through non-histone proteins. Binding partners for both sRme2 and aRme2 were originally discovered in non-histone proteins. In 2001, survival of motor neurons (SMN) proteins, which contain a single Tudor domain, were discovered to bind to arginine-rich regions of proteins, and methylation of those arginines (to sRme2) was found to increase the interaction.[44] Binding of Tudor proteins to aRme2-containing proteins was noted in 2007.[45]

1.5 Dissecting the aromatic cage

The methyl-lysine and -arginine binding proteins described in Section 1.4 for the higher methylation states (di- and/or tri-methylated) share a common structural motif known as an "aromatic cage" due to the presence of multiple aromatic amino acids. The features of this motif allow the domains to facilitate binding and selectivity using many of the weak, intermolecular interactions that are accessible in an aqueous environment. These interactions include: cation-pi, charge-charge, pi-pi and van der Waals interactions. Favourable energetic contributions are also possible from the hydrophobic effect, solvation effects and host pre-organization. The following sections discuss each of these potential interactions from the biochemical and supramolecular points-of-view.

1.5.1 Cation-pi interactions

The primary interaction observed when looking at trimethyl-lysine (Kme3) bound to a typical aromatic cage is the cation-pi interaction. First observed in protein structures in 1978[46] and formally described in 1997,[47] the cation-pi interaction refers to the attractive interaction of a positively charged group (simple cations, ammonium ions, etc) with the quadrupole moment of an aromatic ring. The presence of electrons in the pi orbitals of an aromatic ring results in a build-up of partial negative charge above and below the faces of the aromatic ring. The area to the sides of the ring thus become reduced in electron density, resulting in a partial positive charge (see Figure 1.9) and an overall quadrupolar distribution of charge.

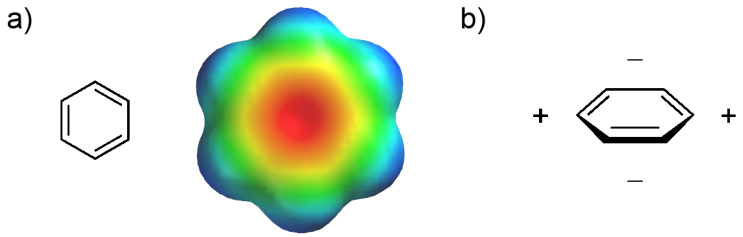


Figure 1.9 The quadrupole moment of benzene. a) The ESP map of benzene shows a region of high electron density (red) over the center of the ring, while the outer edges have low electron density (blue). b) This creates four "poles" to benzene, two positive and two negative. (Spartan '10: PM3; scale: -75 to 50 kJ mol^{-1}).[13]

One can imagine a cation like Kme3 (with its permanent positive charge) being attracted to the region of electron density present on the face of aromatic rings. The average calculated strength of cation-pi interactions with lysine and arginine in from protein structures are -13.8 ± 6.3 kJ/mol and -12.1 ± 5.9 kJ/mol, respectively.[48] Methylation of the cations creates a more diffuse positive charge, which weakens any

potential cation-pi interaction. From this perspective, it is understandable why multiple aromatic rings are present in aromatic cages to further multiply the potential attractive forces present. Of all the aromatic amino acids, tryptophan is overwhelmingly present in aromatic cages due to being the most electron-rich (Figure 1.10). Surveys of protein structures have shown that over 25% of all tryptophan side chains are in close contact with cationic neighbors.[48] The arginine side chain is also more likely than lysine to be found within close proximity of aromatic residues (tyrosine, phenylalanine, and especially tryptophan) to take advantage of cation-pi interactions.[48]

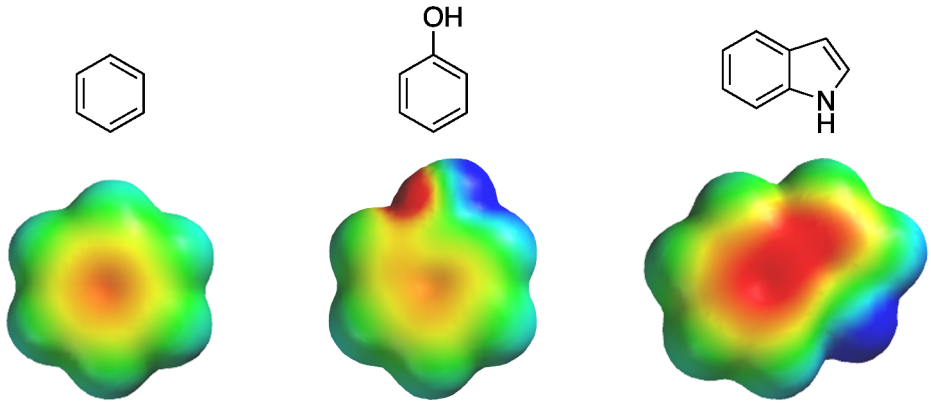


Figure 1.10 ESP maps of benzene (phenylalanine), phenol (tyrosine) and indole (tryptophan). (Spartan '10: PM3; scale: -100 to 100 kJ mol^{-1}).[13] Areas of red indicate high electron density, while areas of blue indicate low electron density.

Cation-pi interactions have been shown to be a critical component for binding Kme3 into an aromatic cage motif. Replacement of the cationic nitrogen in Kme3 with a neutral isosteric t-butyl group (same shape and hydrophobicity) was shown to abrogate binding to a tryptophan residue in a model hairpin pepide.[49]

1.5.2 Electrostatic interactions

Electrostatic interactions are another important weak interaction at play in aromatic cage binding motifs. One notable feature of aromatic cages is that they are often found with an acidic amino acid residue (such as aspartic or glutamic acid) located near or as a part of the cage. At their simplest, these negatively charged residues can participate in an attractive electrostatic interaction with the cationic head groups of lysine and arginine. While all methylated states of lysine and arginine can participate in these interactions, the acidic residues become critically important when protons are present on the cation being bound. Hydrogen bonding and salt bridges (*vide infra*) are key factors that allow the binding domains in Section 1.4 to be selective for the lower methylation

states of lysine (such as the MBT proteins),[33,37] as well as discriminate between the methylated and unmethylated states of arginine (like WDRD5).[42]

Hydrogen bonding is a non-covalent interaction between a hydrogen atom on a more electronegative atom and a source of electron density (such as lone pairs on an oxygen or nitrogen atom). The covalent bond between the hydrogen atom and its partner is polarized and results in a slight positive charge on the hydrogen. This positive atom can participate in attractive interactions with a negative source (Figure 1.11, a). Hydrogen bonding is one of the tools that supramolecular chemists use to build in attractive interactions in synthetic receptors when trying to bind anions, for example.[50-52] In biology, hydrogen bonds can stabilize protein structures and provide specificity for binding.[53]

A salt bridge is a combination of an electrostatic interaction with a hydrogen bond (Figure 1.11, b). All methylated and unmethylated states of arginine and lysine (except Kme3) can participate in salt bond formation. In both of these cases, the source of the polarized hydrogen bond is also a cation.

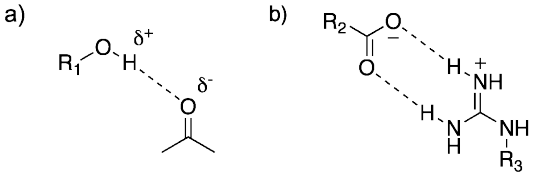


Figure 1.11 Electrostatic interactions. a) Hydrogen bonding and b) salt bridge formation are two attractive interactions used by reader proteins. Hydrogen bonds indicated by dashed lines.

1.5.3 pi-pi interactions with arginine

While arginine is more likely than lysine to be found in contact with aromatic amino acids,[48] this result is not solely due to cation-pi interactions. In addition to the forces described above, arginine can also take advantage of pi-pi interactions due to the delocalized pi electronic nature of the guanidinium group (Section 1.3.2). Pi-pi interactions or pi stacking occurs when the partial negative/electron-rich regions of an aromatic ring interact with the partial positive/electron-poor regions of a second aromatic ring (including examples where temporary induced charges between electrostatically identical partners play a role). Compared to most aromatic rings, guanidinium is an overall electron-poor delocalized pi system (Figure 1.12, a). This allows arginine to favourably stack directly on top of aromatic rings such as tryptophan. Over 50% of arginine side chains near aromatic amino acids have this parallel or stacked geometry.[54]

The stacked geometry is also present in the crystal structure of WDRD5, the only characterized histone methyl-arginine binding protein – the arginine ring is sandwiched between two phenylalanine rings (Figure 1.12, b).[42] The guanidinium motif can access both cation-pi and pi-pi interactions for favourable binding through this stacked geometry. Methylation of arginine residues have been shown to increase their affinity for aromatic surfaces, likely though a combination of solvation/hydrophobic and dispersive effects.[55]

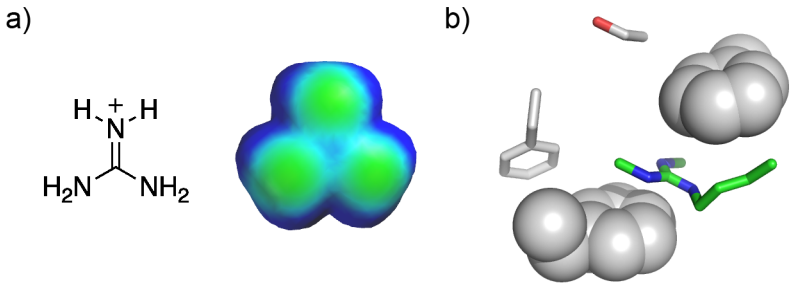


Figure 1.12 Arginine pi-stacking. a) ESP map of guanidinium. (Spartan '10: HF/3-21G; scale: 200 to 600 kJ mol^{-1}).[13] Areas of red would indicate higher electron density, while areas of blue indicate lower electron density. b) Pi-stacking in WDRD5 in complex with H3R2me2s (PDB: 4A7J).

1.5.4 Solvation and the hydrophobic effect

For any recognition that takes place in an aqueous environment, the hydrophobic effect and solvation/desolvation energetics play major roles. The hydrophobic effect refers to the tendency of polar and non-polar species to segregate from one another, much like oil and water. This separation results in a lower overall free energy and a more stable system, since the polar molecules are able to make more favourable interactions (such as hydrogen bonding in the case of water) with each other, than they are with the non-polar molecules. This is referred to as the "classical" hydrophobic effect, and the favourable energetics are entropy-driven by the restored degrees of freedom that result from release of ordered, low entropy water molecules at the non-polar interfacial surface to the higher entropy bulk state. The hydrophobic effect is an important contributor to protein folding, where polar amino acid residues orientate themselves to be solvent exposed while hydrophobic residues are buried towards the core of the protein.[56,57]

The more subtle energetic contributions from solvation and desolvation are also connected to the binding activities of reader protein domains. Histone proteins and their readers function in an aqueous environment, so at any given moment both host and guest are surrounded by water molecules. The degree to which they are solvated (or

stabilized by interactions with the solvent) depends on the type, strength and number of interactions that can be made with water. As a binding event occurs, both the host and the guest must be desolvated. The energy required to do so is reflected in the overall strength of binding: a well-solvated molecule will often bind more weakly to a given host since there is a high energetic penalty to pay for its desolvation. Conversely, a molecule that is poorly solvated will be more easily desolvated, and therefore bind more strongly.

In order to interact with an exposed residue on a histone tail, the aromatic cages of histone binding proteins are close to the protein surface as there is only a three or four carbon linker between the protein backbone and the cationic group. As such, these aromatic cages have multiple aromatic rings that are solvent exposed, but poorly solvated. These cages are static; they cannot collapse to reduce the exposed hydrophobic surface area, resulting in a higher overall energy. The water molecules lining the cavity are also high in energy – the tight, concave shape of the aromatic cages prevents the formation of a hydrogen bonding network between the water molecules.[58] Binding to a large, organic cation (like Kme3), allows the aromatic cage to make favourable contacts with a guest (through cation-pi, and van der Waals interactions) while at the same time displacing the water molecules that were poorly interacting with the aromatic surface, and freeing them to form better hydrogen bonds with water molecules in the bulk solvent. The result is an overall lower energy and favourable binding energetics, this time due to a large enthalpic contribution. Enthalpy-driven (de)hydration energetics is also known as the “non-classical” hydrophobic effect and is often observed for tight binding within concave, non-polar pockets in biological and synthetic receptors.[58]

Solvation effects can also affect the host selectivity for specific guests. Compared to unmethylated lysine, Kme3 is a larger, greasier cation that has no ability to hydrogen bond to surrounding water molecules (due to the absence of any N-H bonds). It is therefore poorly solvated and thus the energy required to remove those associated water molecules (and desolvate the guest) is low. Compare that to lysine which has three hydrogen bond donors and is very well solvated. The energy required for desolvation is high and represents an energetic penalty that must be paid in order to bind lysine. Methylation of lysine removes one of the potential hydrogen bond donors and thus weakens the guest-solvent interaction, making a guest-host interaction more favourable. Indeed, increased methylation is associated with stronger interactions with tryptophan/aromatic cages both in artificial and native systems.[59,60] Studies with HP1

(heterochromatin protein 1) and H3K9 peptides show that binding occurs whether lysine 9 is mono-, di-, or tri-methylated. However, the binding is strongest when lysine is tri-methylated[40] in spite of the fact that cation-pi interactions should be strongest in this series for the smaller, more compact unmethylated lysine side chain.

1.5.5 Host pre-organization

The final consideration of the aromatic cages is the degree of host pre-organization. It is well known in supramolecular chemistry that the less one partner has to adjust its configuration to accommodate the other, the smaller the energetic penalty to be paid upon binding.[61] Stronger affinities often result from pre-organized hosts.

The protein binding domains that target both methylated lysines and methylated arginines are highly pre-organized. Crystal structures show very little change overall in the aromatic cage structure both with and without the guest present (Figure 1.13) in both cases. Binding is therefore not because of an induced fit nor does the cage itself show any collapse in water despite its hydrophobic nature. This would indicate that the high degree of pre-organization present is a result of the surrounding protein structure.

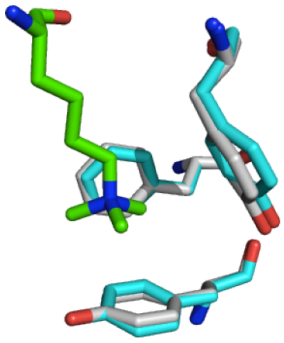


Figure 1.13 The affect of guest binding upon aromatic cage structure. Overlapping crystal structures of the aromatic cage of the WD40 binding domain of PRC2 both without (blue) and with (grey) H3K27me3 bound show little change in the amino acid positions. PDB: 3JZN (apo) and 3JZG (bound).[34]

The highly ordered structure exhibited by these binding domains can be understood when all of the potential favourable interactions described above are taken into account. In these rigid cages, cation-pi interactions are known to be essential.[49] This imposes a strict geometric requirement; in order to have a strong cation-pi interaction, the cation must be able to interact with the *face* of the aromatic rings (Section 1.5.1). As a result, the aromatic rings must be solvent exposed, despite the tendency for hydrophobic amino acids to bury themselves within a protein's core (Section 1.5.4). Hydrophobic surfaces exposed to water are poorly solvated, and the low

desolvation penalty (for both host and guest) results in stronger binding. Taken together, aromatic cage rigidity is crucial for all of these weak interactions to play a part in binding a cation. As such, pre-organization has been considered a hallmark of these aromatic cage motifs.[37]

1.6 Synthetic aqueous receptors for methyl-lysines

The supramolecular literature currently has few examples of water-soluble hosts that are designed for or have been specifically used to bind methyl-lysines from water. One of the earliest published examples for Kme3 was the very simple, water-soluble *p*-sulfonatocalix[4]arene **1.1** (PSC, Figure 1.14). Binding studies against free methyl-lysine and arginine amino acids showed that PSC bound strongest to Kme3 (37 000 ± 18 000 M^{-1}), with a 70-fold difference over unmodified K, and at least 30-fold stronger than any other amino acid tested.[62] Subsequent studies also showed that **1.1** could out-compete a native aromatic cage (CBX7 reader protein) for binding to Kme3 (in an H3K27me3 peptide).[63] A second example of a methyl-lysine binding small molecule was published by the Waters group. They used a dynamic combinatorial library to search out molecular combinations of aromatic monomers that were amplified in the presence of trimethyl-lysine, resulting in the isolation of cyclophane **1.2**.[39] Subsequent binding studies to H3K9me3 peptides again found protein-like affinity for **1.2** compared to the native HP1 chromodomain.

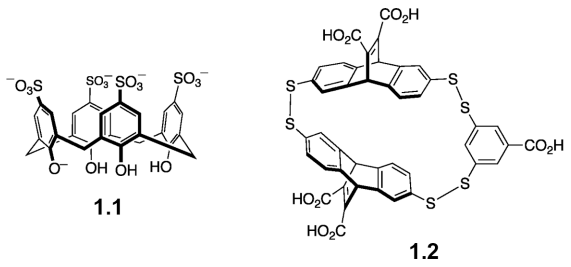


Figure 1.14 Kme3 binding molecules. a) PSC **1.1**, b) Waters' cyclophane **1.2**.

The receptors in Figure 1.14 make use of many of the traits present in native aromatic cage structures – multiple aromatic rings, potential for cation-pi and other electrostatic interactions, a defined geometry, etc. – in order to successfully bind Kme3. While few molecules are reported to directly bind to Kme3, the cationic head group target of the aromatic cage is simply a quaternary ammonium ion of the form R-NMe_3^+ where R = alkyl (such as acetylcholine, choline, and tetramethylammonium). Many more water-soluble receptors have been synthesized that are designed to be a size and shape

match for these cationic spherical alkyl ammonium ions. These hosts, though not directly used as methyl-lysine receptors, would be comparable in their behaviour.

Cyclophanes (aromatic units connected by linking atoms, usually carbon) such as **1.2** have often been used to bind quaternary ammonium ions. One of the more recognizable water-soluble cyclophanes was synthesized by Dougherty and co-workers during their early studies of cation-pi interactions.[64,65] Cyclophane **1.3** was originally synthesized as a host for water-soluble organic molecules by using the larger, hydrophobic cavity present. Tetramethyl ammonium groups were appended onto many aromatic rings and other aliphatic compounds simply as a means of rendering them water-soluble for study. Surprisingly, the quaternary ammonium groups themselves were found to be strongly encapsulated within the hydrophobic pocket, rather than extended out into the solvent. Alkyl-NMe_3^+ guests were bound strongly with affinities of approximately 10^7 - 10^8 M^{-1} (4.5 – 5 kcal/mol, borate buffered D_2O). Later modifications to add additional negative charges to the rim of the cavity (**1.4**) were found to improve binding to lower alkylation states such as unmethylated lysine (7 x 10^6 M^{-1}, 4 kcal/mol) and as well as unmethylated arginine (see section 1.7). Dougherty's ethanoanthracene skeleton was also the inspiration behind Waters' building blocks for **1.2**.

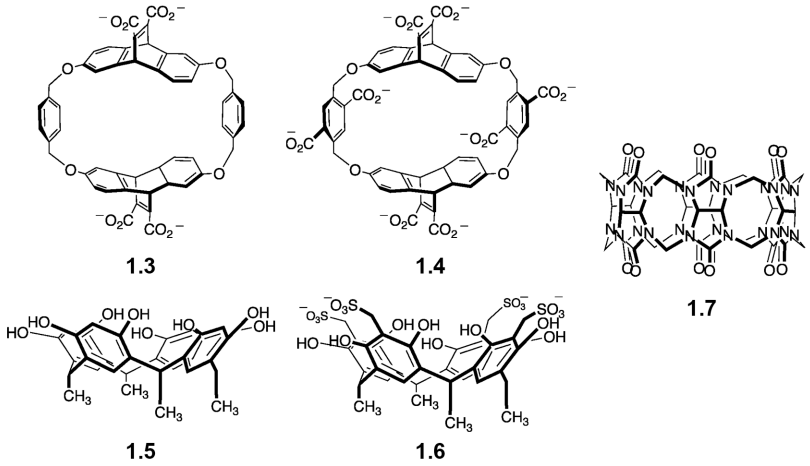


Figure 1.15 Other R-NMe_3^+ receptors

In addition to calixarenes and cyclophanes, resorcinarenes are another class of aromatic macrocycles that have been used to bind quaternary ammonium ions from water. Formed from the acid catalyzed condensation of resorcinol (1,3-benzenediol) and formaldehyde, the structure is highly similar to that of calixarenes like **1.1**. Strong binding of undecorated resocinarene **1.5** to R-Me_3^+ cations (such as NMe_4^+ and choline) on the order of 10^5 M^{-1} has also been observed in strongly basic solutions (0.01M NaOH).[66]

Under these conditions, four of the eight phenol groups are deprotonated, leading to a stronger electrostatic interaction. Sulfonation has also been used to make sulfonated resorcinarenes such as **1.6**, which are water soluble at neutral pH.[67] Tetraanionic **1.6** was found to bind tetramethylammonium with an association constant of 270 M^{-1} from unbuffered D_2O.[68] Cucurbit[7]uril (**1.7**) host-guests complexes with choline analogs have also shown high affinity in water.[69] This work was most recently extended to Kme3 derivatives.[70] Inclusion complexes with **1.7** do not include any cation-pi interactions, as the interior of the cucurbituril is hydrophobic but not aromatic.

1.7 Synthetic aqueous receptors for methylated arginines

While a number of receptors have been successfully designed, synthesized and shown to have strong aqueous binding affinities for trimethyl-lysine (and quaternary ammonium ions in general), fewer molecules are known as arginine binding hosts. Synthetic receptors for unmodified arginine/guanidinium ions are known. These receptors rely heavily on having multiple hydrogen bond- and/or salt bridge-forming functional groups present in a single plane to connect with the multiple, cationic hydrogen bond donors of arginine. *However, there are no known synthetic hosts that can strongly bind any form of methylated arginine in water, nor discriminate between methylation states.*

Water-soluble receptors for guanidinium and arginine have been around since the early days of supramolecular chemistry. Lehn and co-workers were able to bind guanidinium and arginine-like mono-substituted guanidinium ions in buffered water using functionalized [27]-crown-9 macrocycles **1.8** in 1979 (Figure 1.16).[71] Binding to this motif depended strongly on hydrogen bonding between the guanidinium N-Hs and the oxygen atoms in the crown ether. Any substitution of the guanidinium guest resulted in a decrease in binding by disrupting both the size and binding site match between host and guest.

Guanidinium binding in polar solvents was then not explored again until 1997 when Schrader and co-workers published the biomimetic bis-phosphonate tweezer scaffold.[72] This “arginine fork” also exploited the hydrogen bond/salt-bridge capacity of guanidinium but because of its non-macrocyclic nature, could accommodate a larger guest like methyl guanidinium. Schrader continued to expand on this bis-phosphonate scaffold by adding in an aromatic core for cation-pi and pi-pi interactions,[73] rigidifying the scaffold,[74] and adding additional phosphate groups (**1.9**) to generate a water-soluble

host.[75] While these molecules were eventually able to bind arginine-rich peptide sequences from water,[76] no exploration was done towards any form of post-translationally modified arginine residue. A similar arginine-binding motif based on extensive electrostatic interactions was also published around this time by Bell and co-workers.[77] This water-soluble "arginine cork" **1.10** was also able to bind arginine from water with a K_d of 1100 μM.

From these examples, it is clear that strong aqueous binding to methyl-arginine derivatives cannot be based on electrostatic interactions alone. As mentioned earlier, Dougherty was an early proponent of the importance of the cation-pi effect in biological systems. His highly aromatic cyclophane molecules (**1.4**), first explored for quaternary ammonium compounds, proved to be similarly useful when binding to arginine and guanidinium derivatives.[78] Arginine was found to be strongly bound from borate-buffered water (5.0 kcal/mol), while alkyl-guanidinium derivatives including per-methylated guanidinium were also strong binders (up to 6.7 kcal/mol). However, significant hydrophobic additions to guanidinium were required to induce any binding (such as per-methylation).

The final family of aromatic cage-style receptors for arginine derivatives comes from the work of Klärner and co-workers. They synthesized a water-soluble molecular clip **1.11** and explored binding with alkylpyridinium cations. These are similar to methyl-arginines in that they are also cationic, aromatic and flat. Binding constants of 5000 M^{-1} and higher were obtained with the N-alkyl pyridinium salts used.[79] However, in exploring the scope of binding to this molecular clip, they later found that other flat cations that are more strongly hydrated, such as N-H pyridinium, imidazolium and guanidinium cations (including methyl guanidinium and arginine) had non-existent affinity in water.[80] A later design of a molecular tweezer from similar synthetic methodology **1.12**, was successful at binding both lysine and arginine from buffered water, though spherical lysine was favored by 2.5 fold.[81]

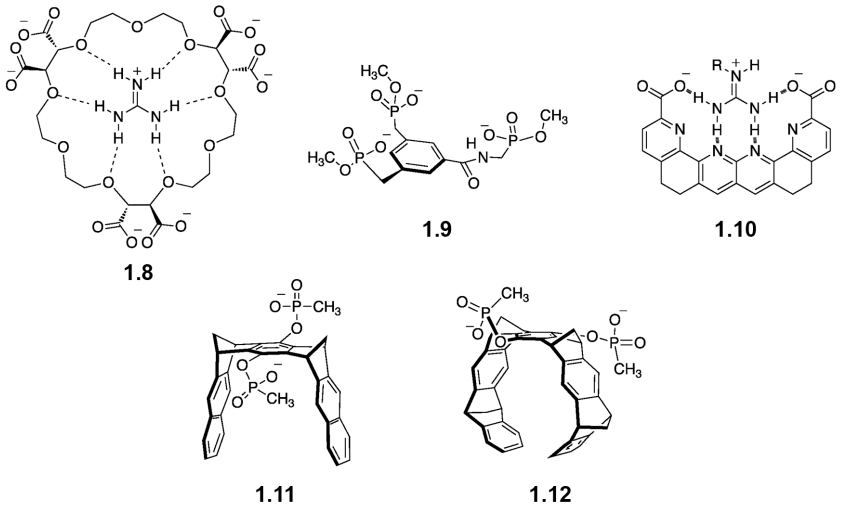


Figure 1.16 Examples of synthetic receptors for arginine binding

1.8 Summary and key questions

The study of histone tail PTMs is continually fascinating in that much of the body's regulation hinges on as little as the addition of one, two or three carbon atoms to key regulatory proteins. The subtlety that goes into being able to distinguish between each of those modifications, let alone distinguish the same modification at nearby sites on the same unstructured protein, or even begin to interpret an overall pattern of PTMs is incredible. While nature is well-equipped to be able to do just that, chemists and biochemists have not yet reached that level of expertise or understanding.

As detailed earlier, successful binding results from a synergistic interplay of many different forces and criteria. Charged interactions (such as an electrostatic salt bridge) are weakened in an aqueous (polar) medium but at the same time, contributions from hydrophobic interactions can now play a role in driving binding. Lysine methylation sequentially removes hydrogen bonding sites but also makes the cation more poorly solvated, and therefore a better binding partner for a hydrophobic host. The inclusion of a cation has been shown through experimentation to be a critical part of binding to aromatic cage motifs,[49] but so have the aromatic residues through mutation studies.[35,82,83] Even something as simple as swapping an acidic for an aromatic amino acid residue can have a profound effect on the specificity of the aromatic cage of interest.[41]

Even with choosing to focus on a single reader (the aromatic cage) of a single PTM (trimethyl-lysine), there are a number of areas for potential exploration. For this thesis, we have chosen to focus on some of the more subtle contributors to binding: the

intertwined energetic effects of host pre-organization and solvation. At first, we were interested in investigating how much of the binding strength of an aromatic cage comes from it being rigidly held open as shown by numerous crystal structures with- and without guests. Would a similarly electron-rich, aromatic, water-soluble but flexible host be able to bind quaternary ammonium ions (like Kme3) from water? (Chapter 2). From this study, we realized the importance of harnessing the power of hydrophobic contributions (both classical and non-classical) for our own gains, rather than letting them dictate structure and therefore, function. The need to enforce a stricter, pre-organized geometry led us to the synthetic exploration of a rigid building block, Tröger's base (Chapter 3). Finally, we were able to incorporate our building block into a number of water-soluble structures and examine what effect a hydrophobic, aromatic hinge would have on binding to our hydrophobic methylated cations (Chapter 4).

Chapter 2 – Binding trimethyllysine and other cationic guests in water with a series of indole-derived hosts: large differences in affinity from subtle changes in structure

Adapted from: Amanda L. Whiting,[1] Nicole M. Neufeld[1] and Fraser Hof[1]
Tetrahedron Letters (2009) *50*, 7035 – 7037.

and

Amanda L. Whiting[1] and Fraser Hof.[1]
Organic & Biomolecular Chemistry (2012), *10*, 6885 – 6892.

[1]Department of Chemistry, University of Victoria, Victoria, BC, Canada

ALW designed the research, performed the syntheses, collected and analyzed data, and wrote the manuscripts. NMN performed the first in-house synthesis of **2.6** using literature precedent.

2.1 Foreword

As described in Chapter 1, molecular recognition arises from the contributions of many kinds of weak, intermolecular forces, all interacting and influencing the strength and behaviour of each other. In order to design a successful receptor for a guest of interest, one must be able to predict what forces favour a binding event, how those forcers interact with each other, as well as how they behave in the specific system under study. A clear structure-function relationship that allows one to predict how a system will behave is critical in designing any synthetic supramolecular system.

While supramolecular structure-function relationships have become relatively easy to predict in organic solvents, they still present a challenge to the designer who wishes to create a system that functions well in pure water. In water, additional considerations such as solvation energies and the hydrophobic effect become increasingly important. While nature has perfected the art of encoding strong and selective binding in water, chemists still find it difficult to generate relatively simple synthetic receptors that can reproduce the biological level of mastery.

With this chapter, our goal was to explore the binding properties of a synthetically simple, highly aromatic but still water-soluble receptor. As our target molecule was the increasingly biologically important trimethyl-lysine mark (Kme3), we modeled our host on the binding motif native to its biological partners: the aromatic cage. We took note of what forces were important for the native affinity, and tried to build them into our receptor. For example, since the cation-pi effect has been shown to be a critical part of binding Kme3, our receptor was endowed with multiple aromatic rings to allow for multiple occasions of just such an interaction. Kme3 is an organic cation making it hydrophobic but still water-soluble; favourable solvation effects play into having a host that is also similarly hydrophobic, but water-soluble. The repeating aromatic units already present for the cation-pi interaction could also play positively into this requirement. Kme3 has no protons suitable for hydrogen bonding, nor can the cationic head be a hydrogen bond acceptor – as such, the host molecule does not include either option. Of the criteria listed in Chapter 1, the one pre-requisite that we left out of our design was strict host pre-organization. Instead of a macrocycle with a persistent hydrophobic cavity as is commonly used for Kme3-like cations, we choose to pursue a host based on a 1,3,5-substituted benzene scaffold. This scaffold is regularly used in the supramolecular literature as a way to position binding elements in close proximity to

interact with a single guest. While this was primarily a decision based on synthetic ease, it allowed us to ask questions: just how important is a pre-organized scaffold to this system? Would the other interactions present be sufficient to make up for any energetic penalty of reorganization? We synthesized a series of compounds each with small differences to complement our original design, and compared the results.

2.2 Abstract

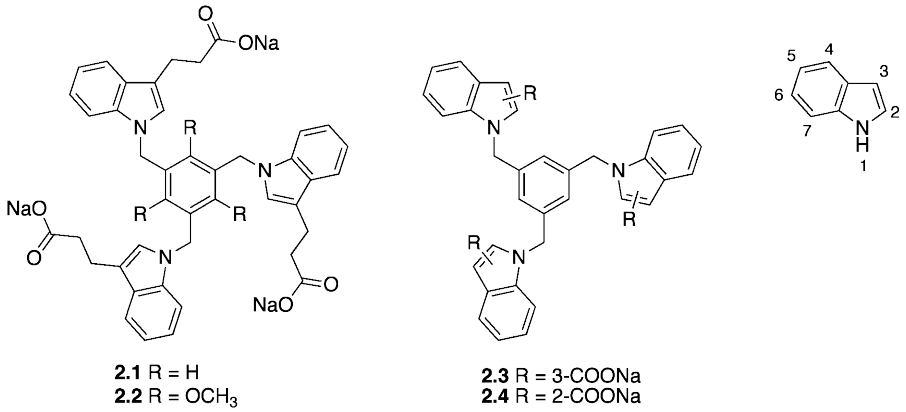
The binding of a series of indole-derived hosts to various ammonium cations in pure, buffered water is investigated using both solution phase ^{1}H NMR studies and computational modeling. These hosts can engage their targets via cation-pi interactions, electrostatic attraction, and the hydrophobic effect. Cation-pi interactions are thought to dominate in related biological systems; however, in this synthetic system, the hydrophobic effect is shown to have a dominant influence in the strength of the binding interactions, both in terms of the hydrophobicity of the host and of the guest. Significant decreases in binding are shown to result from small changes that reduce the host hydrophobic surface area without reducing either the number of negative charges or the amount of aromatic surface area. Additionally, the position of solubilizing charges is shown to influence the preferred host geometry and resulting binding constants.

2.3 Introduction

Indole rings, in the context of the amino acid tryptophan, have been identified as key players in a variety of biologically important recognition events. Compared to the other aromatic amino acids, tryptophan combines the most electron-rich pi surface with significant hydrophobic character (Figure 1.10). Surveys of protein structures have shown that over 25% of all tryptophan side chains are in close contact with cationic neighbors,[48] including a subset that are particularly important for protein-protein interactions.[54] Quaternary ammonium cations, such as post-translationally methylated lysine side chains,[30] as well as a variety of neurotransmitters[84] and phospholipid head groups,[85] constitute a specific class of cation that are routinely engaged by a family of protein binding pockets known as “aromatic cages.” The known examples of these pockets, expressed in a variety of evolutionarily distinct domains including plant homeobox domains (PHD), chromodomains, Tudor domains, etc, *all* contain at least one tryptophan indole ring alongside other aromatic residues.

Taking inspiration from the aromatic cages of biology, tris(indole) hosts **2.1 – 2.4** were designed while considering the general goals of rapid synthesis, the presentation

of multiple electron-rich indole rings, and water solubility. Host **2.1**, decorated with three indole rings appended with carboxylates, is a mimic of the tryptophan-rich protein binding pockets that have evolved to engage quaternary and tertiary ammonium ions such as acetylcholine (ACh) and trimethyllysine (Kme3).[32,63,86-92] We sought to explore the cation binding ability of this motif through a series of related hosts – each containing three indole rings and three carboxylic acids attached to a central benzene ring but with variation in the exact positioning of functionality from one host to another (Figure 2.1). All four hosts provide the same amount of aromatic surface area and the same number of negative charges to engage their cationic binding partners.



2.1 R = H
2.2 R = OCH_3

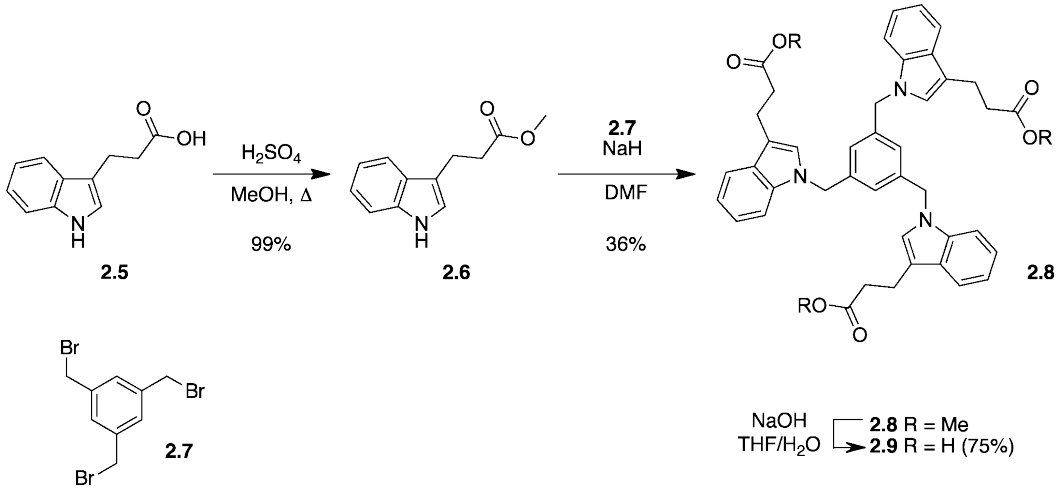
2.3 R = 3-COONa
2.4 R = 2-COONa

Figure 2.1 Hosts studied in this work (indole numbering guide for reference)

In host **2.2**, we chose to retain the 3-propionic carboxylate tails on the indoles and append methoxy groups to the core benzene ring. It was thought by us and others[65] that the electron-donating ability of the oxygen atoms could contribute to an increase in electron density on the core. This could potentially lead to an increase in binding due to an enhanced cation-pi effect at the central ring. Host **2.3** and **2.4** were obtained by shortening the propionate side chain of **2.1** to a simple carboxylate at the 3- and 2-position of the indoles, respectively, and were envisioned as less hydrophobic versions of **2.1**. The effect of removing the -CH_2CH_2- linker could be determined directly by comparing **2.1** and **2.3**, while altering the position of the carboxylate (host **2.4**) would allow us to probe a different type of geometric variation within this family. Our studies included both a detailed analysis of host conformations based on experimental and computational data, and NMR-based determinations of association constants for various ammonium ions of interest.

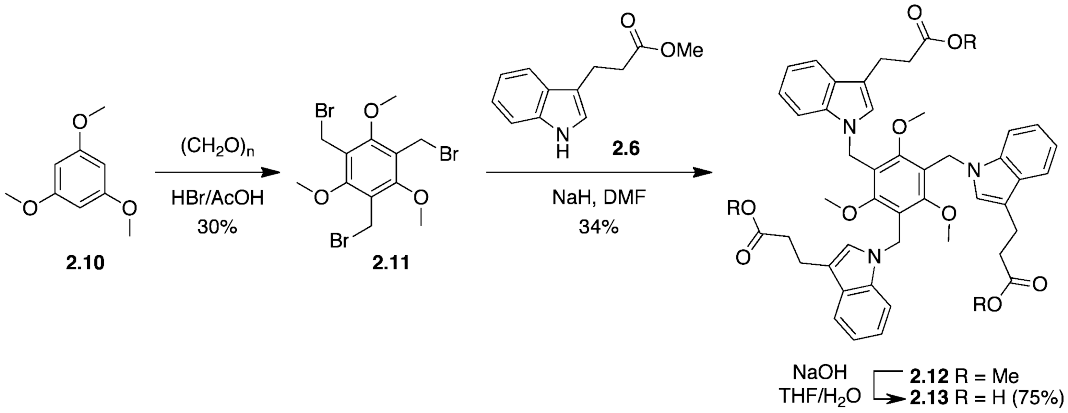
2.4 Synthesis of host molecules

The desired water-soluble indole host **2.1** was synthesized in three steps starting from commercially available indole-3-propionic acid **2.5** (Scheme 2.1). After protection as the methyl ester,[93] three equivalents of indole **2.6** were coupled to 1,3,5-tribromomethylbenzene **2.7** upon deprotonation with NaH in DMF to give **2.8** in 36% yield. Basic hydrolysis of methyl ester **2.8** was followed by acidification and isolation of the tri-acid product (**2.9**) in 27% overall yield. Quantitative conversion to the tri-sodium salt using stoichiometric NaOMe gave the final product **2.1**.



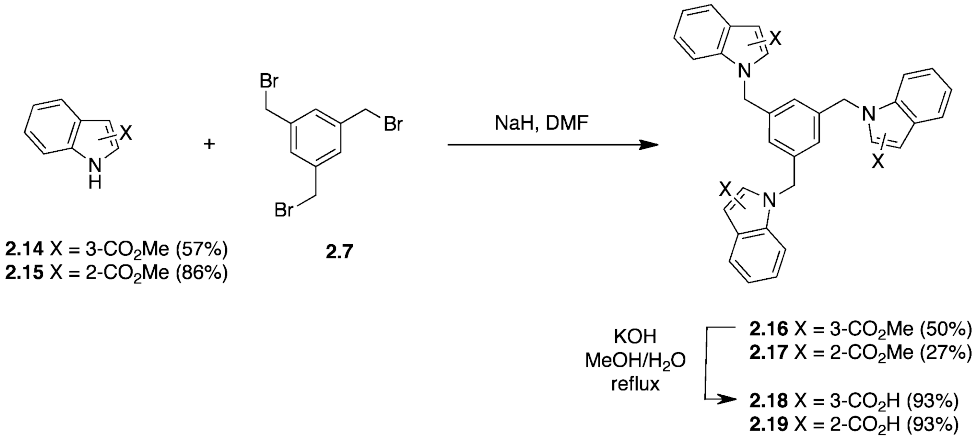
Scheme 2.1 Synthesis of **2.1**

Synthesis of host **2.2** began by treating 1,3,5-trimethoxybenzene **2.10** with paraformaldehyde and HBr to provide the 1,3,5-tris(bromomethyl)-2,4,6-trimethoxybenzene core **2.11** in 30% yield.[94] This methoxy core was then reacted with three equivalents of methyl indole-3-propionic acid **2.6** that had been pre-treated with NaH in DMF to give **2.12** in 34% yield. Deprotection of the methyl esters in basic aqueous solution gave the acid host **2.13** for an overall yield of 8%. **2.2** was obtained quantitatively upon treatment with NaOMe.



Scheme 2.2 Synthesis of **2.2** (Reproduced by permission of the Royal Society of Chemistry)

Hosts **2.3** and **2.4** were similarly synthesized starting with the protection of commercial 3- and 2-indole carboxylic acids to give methyl esters **2.14** and **2.15**. Coupling to 1,3,5-tris(bromomethyl)benzene **2.7** to give esters **2.16** and **2.17** followed by deprotection with KOH in MeOH gave hosts **2.18** and **2.19**, in 27% and 22% overall yields, respectively**.** Again, these hosts were isolated as carboxylic acids, and quantitatively converted to their tri-sodium salts by treatment with stoichiometric NaOMe prior to their dissolution in buffered water for use in NMR studies.



Scheme 2.3 Synthesis of **2.3** and **2.4** (Reproduced by permission of the Royal Society of Chemistry)

2.5 Solution-phase host geometries in organic solvent and water

Hosts **2.1 – 2.4** are inherently flexible due to the nature of the single bonds connecting the indole arms to the central benzene. As a result, they are likely able to adopt a number of conformations separated by low energetic barriers. This is especially true in organic solutions like dimethyl sulfoxide (DMSO) where the hydrophobic aromatic

elements of the hosts are well solvated. In water however, the geometries adopted will depend on the inherent bond rotational preferences of the host structures (as in DMSO) as well as aromatic clustering driven by the hydrophobic effect.[95,96] The preferred geometry prior to binding will not necessarily resemble an ideal open and bowl-like conformation.

2.5.1 ^{1}H NMR chemical shift changes and energy minimized host structures

We examined the aqueous host conformations without guests by comparing the chemical shifts of each host in DMSO-d_6 to those in buffered D_2O to see which protons on the hosts were most affected by the solvent change. The resulting spectra of the protons of each host are shown in Figure 2.2 to 2.5. Protons that were found to have ≥ 0.2 ppm change in chemical shift are highlighted in Figure 2.6. Protons that significantly shift upfield upon the move to D_2O indicate a closer association with aromatic surfaces. Complete chemical shift changes for all protons are indicated in Figure 2.7. As a control, indole-3-propionic acid **2.5** was also subjected to NMR in DMSO-d_6 and buffered D_2O. Each observable proton for **2.5** exhibited a downfield chemical shift going from the organic to aqueous solvent.

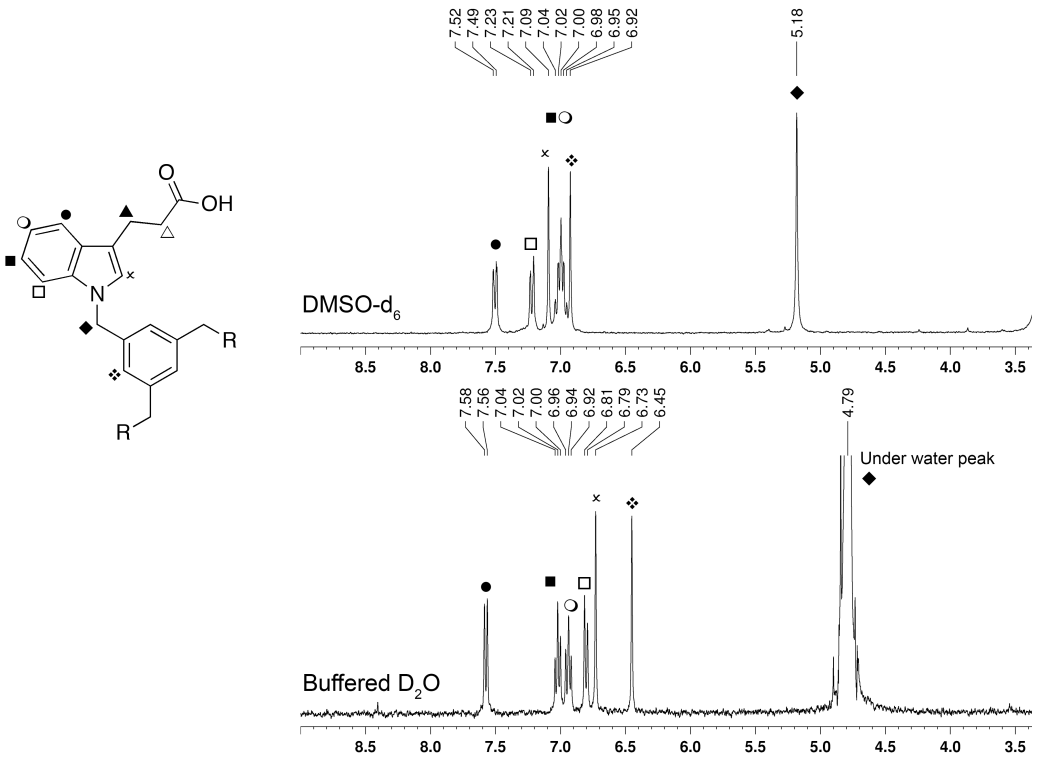


Figure 2.2 Protons of **2.1** (9.0 – 3.5 ppm) showing chemical shift upon solvent change in DMSO-d_6 (upper) and D_2O (lower). Proton identities as follows: 4-H (●), 7-H (□), 2-H (x), 6-H (■), 5-H (❍), Ar-H (of central benzene ring, ❖), and N-CH_2-Ar (◆). (Reproduced by permission of the Royal Society of Chemistry).

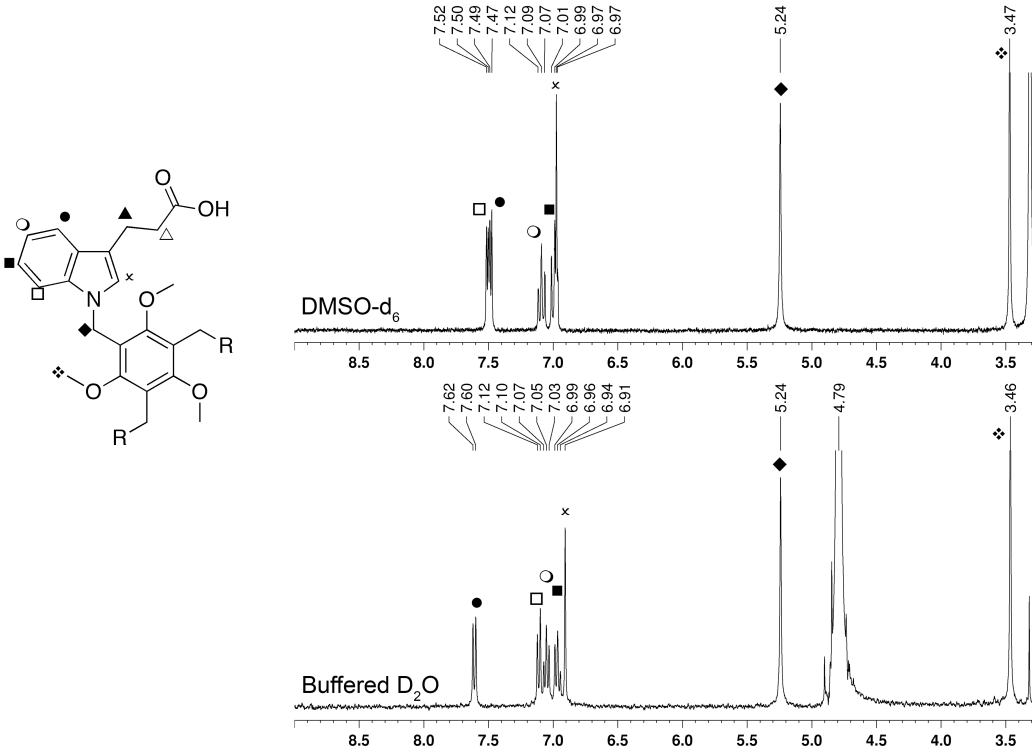


Figure 2.3 Protons of **2.2** (9.0 – 3.5 ppm) showing chemical shift upon solvent change in DMSO-d_6 (upper) and D_2O (lower). Proton identities as follows: 7-H (☐), 4-H (●), 5-H (❍), 6-H (■), 2-H (x), N-CH2-Ar (◆) and Ar-OCH_3 (of central benzene ring, ❖) (Reproduced by permission of The Royal Society of Chemistry).

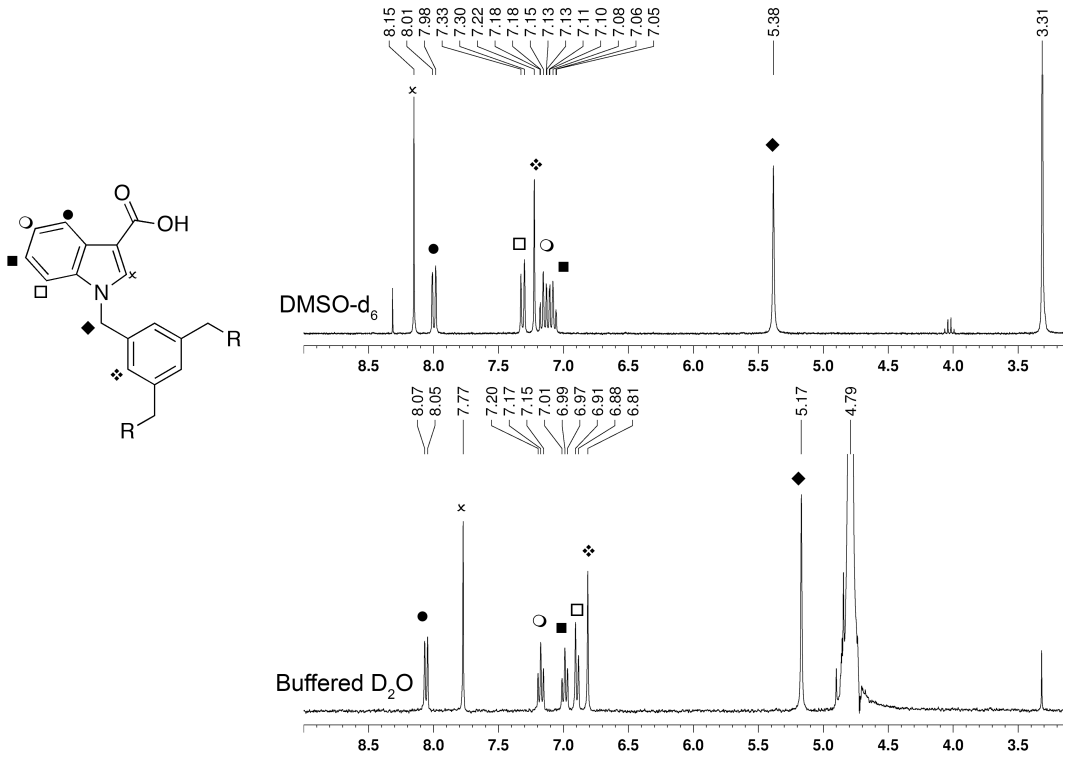


Figure 2.4 Protons of **2.3** (9.0 – 3.5 ppm) showing chemical shift upon solvent change in DMSO-d_6 (upper) and D_2O (lower). Proton identities as follows: 2-H (x), 4-H (●), 7-H (□), Ar-H (of central benzene ring, ❖), 5-H (❍), 6-H (■), and N-CH_2-Ar (◆) (Reproduced by permission of The Royal Society of Chemistry).

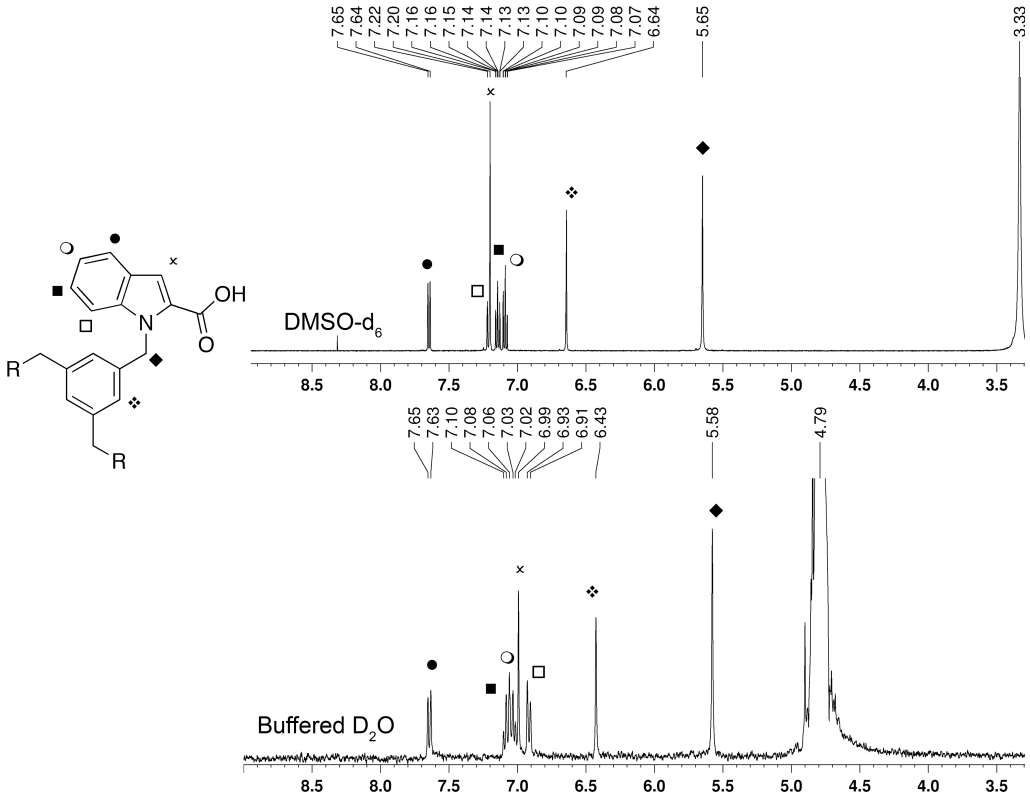


Figure 2.5 Protons of **2.4** (9.0 – 3.5 ppm) showing chemical shift upon solvent change in DMSO-d_6 (upper) and D_2O (lower). Proton identities as follows: 4-H (●), 7-H (□), 3-H (x), 6-H (■), 5-H (❍), Ar-H (of central benzene ring, ❖) and N-CH_2-Ar (◆) (Reproduced by permission of The Royal Society of Chemistry).

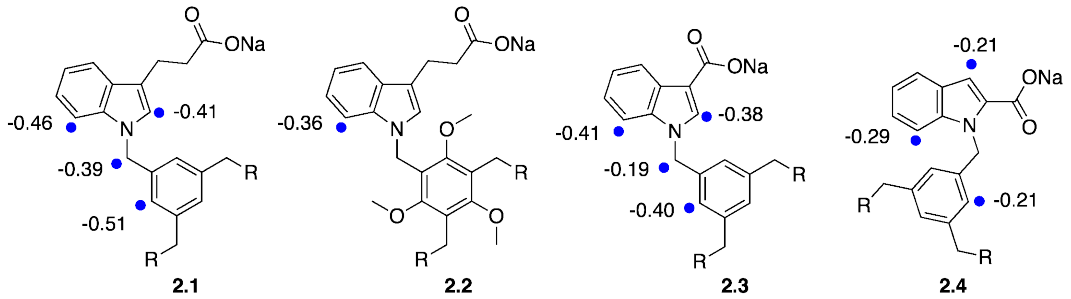


Figure 2.6 Significant chemical shift changes from DMSO to D_2O. Magnitude of shift is indicated (ppm).

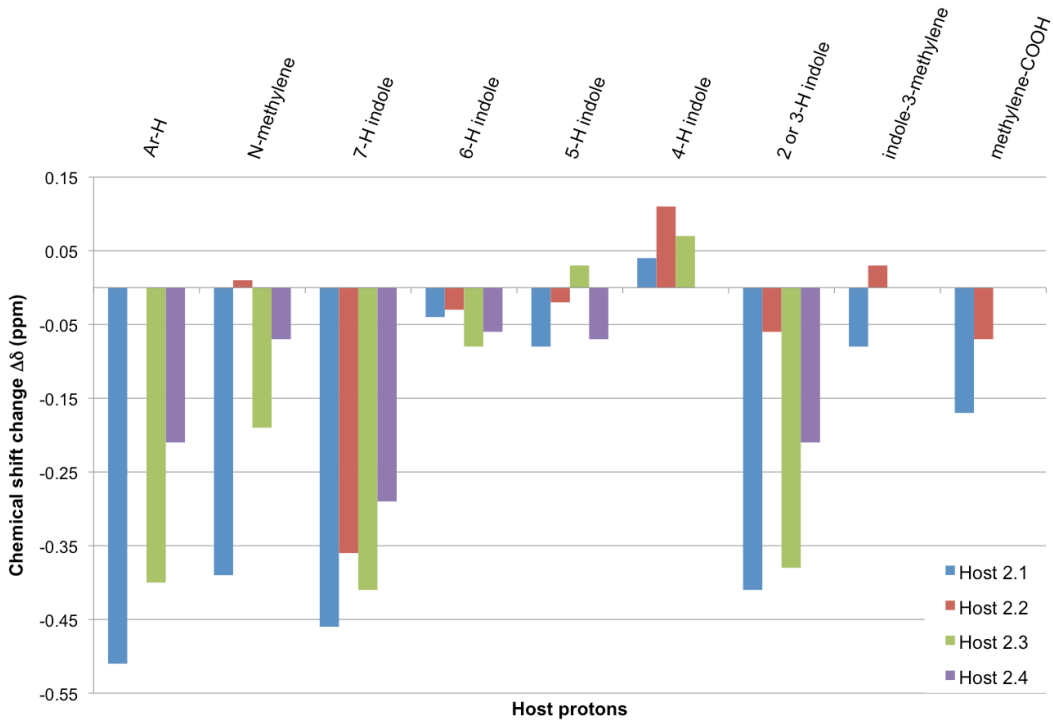


Figure 2.7 Complete NMR chemical shift changes between DMSO-d_6 (reference point) and D_2O for all hosts reported as $\Delta\delta$. $-OCH_3$ (from **2.2**) omitted due to negligible shifts. Data for 2- and 3-H protons are combined.

Unlike control indole **2.5**, where all protons shifted downfield, hosts **2.1** and **2.3** each have several protons that move upfield to the same degrees, indicating that a similar conformation is adopted for these two hosts in D_2O. Energy minimizations in implicit water (HF/6-31G* as implemented in Spartan '10)[13] suggest that the three indole rings are collapsed into a closed, propeller-like aromatic cluster, with each of the protons that are observed to be upfield-shifted in NMR data are located in the shielding cone of a neighboring aromatic ring (Figure 2.8).The methoxy-containing host **2.2** has only one proton that experiences significant solvent-induced shifts (7-H indole) with all other shifts being negligible. Energy minimizations again suggest a propeller-like conformation for **2.2** that is consistent with the (more limited) experimental data available for this host.

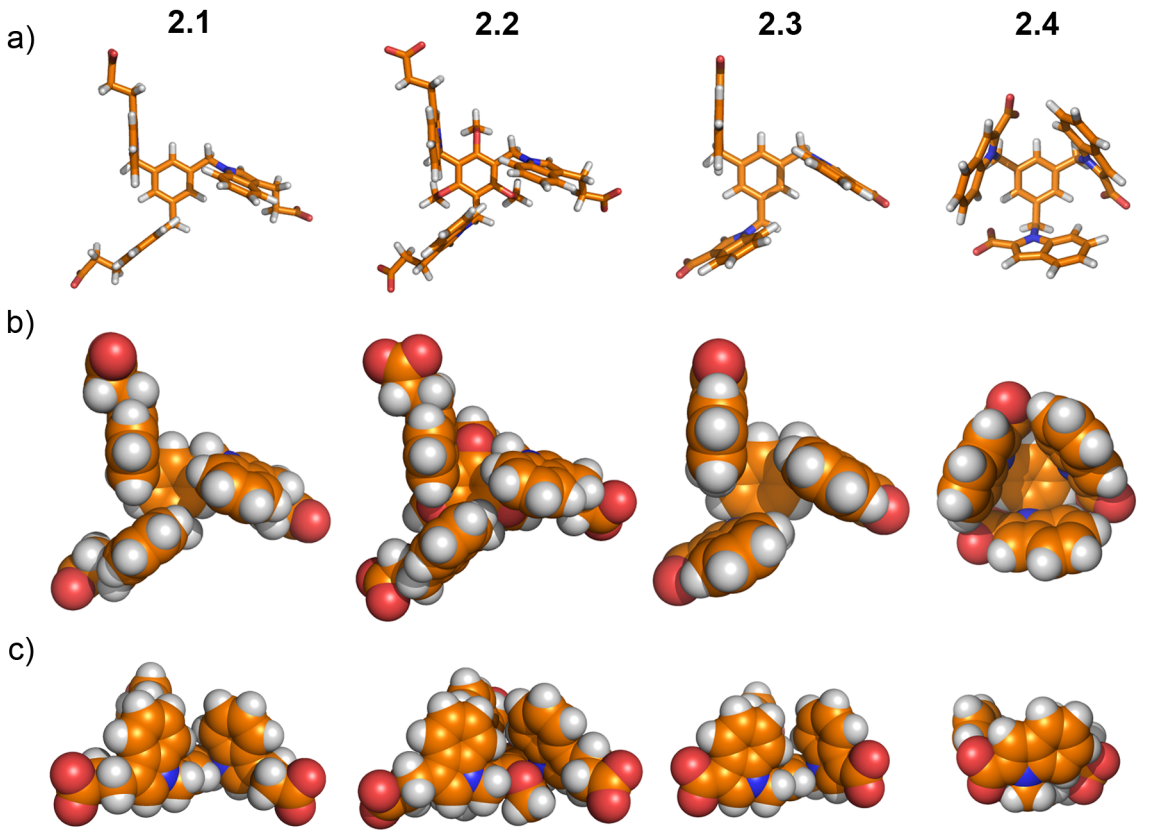


Figure 2.8 Equilibrium geometries in implicit water for hosts **2.1**, **2.2**, **2.3** and **2.4** (HF/6-31G* as implemented in Spartan '10).[13] a) Top view, ball and stick model; b) Top view, space-filling model and c) Side view, space-filling model.

Host **2.4**, with a 2-carboxy substituent, shows overall smaller changes in chemical shift than the other hosts, but follows a similar pattern to the shifts observed in hosts **2.1** and **2.3**, again indicating a closer association with aromatic surfaces. Interestingly, the energy minimized structure for **2.4** in implicit water suggests an open conformation, unlike the other hosts in this series. If accurate, the open conformation of **2.4** (Figure 2.8) would be expected to have downfield chemical shifts for the 6-H and 7-H indole protons arising from CH---O contacts with the carboxylates. Experimentally, this was observed for proton 6-H (the proton with the most direct carboxylate contact) on host **2.4** which, at 7.08 ppm in D_2O and 7.14 ppm in DMSO-d_6, is located furthest downfield of all hosts in both water and also DMSO.

2.5.2 ROESY evidence for structural comparisons

Further NMR evidence that helps us draw conclusions about the conformational changes experienced by all four hosts in DMSO and in buffered water comes from 1D ROESY experiments carried out in each solvent. In DMSO, key Rotating frame-of-

reference Nuclear Overhauser effects (ROEs) indicating through-space interactions are noted between the N-CH_2 methylene protons and indole H-7 for all hosts ("indole-out," Figure 2.9a), and between N-CH_2 and indole H-2 for hosts **2.1**, **2.2** and **2.3** ("indole-in," Figure 2.9b; host **2.4** has no H-2 proton). These two close contacts are mutually exclusive and occur in two different conformations that differ only by rotation about the indole N-CH_2 single bond. We conclude that the ROE data in DMSO are best explained by the rapid exchange between both conformations. In buffered water, hosts **2.1**, **2.2**, and **2.4** continue to show N-CH_2 to H-2 contacts that arise from the "indole-in" conformation shown in Figure 2.9b, but show no evidence of contacts between H-7 and N-CH_2 that would arise from the "indole-out" conformation shown in Figure 2.9a. Of note is that the "indole-in" rotamer, with H-7 closely packed over the central ring, is consistent with the calculated closed, propeller-like structures for hosts **2.1 – 2.3** shown in Figure 2.8. Host **2.4** maintains H-7 to N-CH_2 contacts in D_2O, as well as showing H-7 – ArH contacts that are consistent with the calculated open conformation for **2.4** shown in Figure 2.8.

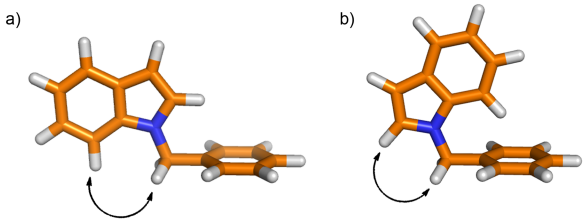


Figure 2.9 Key ROESY interactions in the a) "indole-out" rotamer (H-7 to CH_2 contact) and b) "indole-in" rotamer (H-2 to CH_2 contact). The "closed" indole rotamer is the same as that adopted by all three indoles of hosts **2.1 – 2.3**, as shown in Figure 2.8.

Overall, the chemical shift, computational, and ROE data suggest geometries for hosts **2.1 – 2.3** (not **2.4**) in which the aromatic rings are further compressed onto each other and the charged carboxylate arms are mostly directed outwards into solution. This is consistent with the expected influence of the hydrophobic effect on these radially amphiphilic molecules. The collapsed form in water also allows for stronger edge-to-face aromatic interactions between the rings than would be present in DMSO, where the hosts maintain significant mobility.

2.6 Solution-phase binding studies

The binding properties of hosts **2.1 – 2.4** were probed by NMR studies in phosphate-buffered D_2O (Table 2.1). Dilution titrations conducted on each host alone

gave concentration-dependent chemical shifts that were fit to determine a self-association constant for each host ($K_{dimerization}$). Subsequent titrations with various ammonium ions provided data that could be fit to a 1:1 host-guest binding isotherm to determine K_{assoc} while taking into account the effects of host homodimerization using HypNMR[97] (See Sections 2.9.3 and 2.9.4 for calculation and experimental details).

Table 2.1 Binding affinities for host **2.1** – **2.4** in phosphate-buffered D_2O[a]

		K_{assoc} (M^{-1})			
Entry	Guest	**2.1**	**2.2**	**2.3**	**2.4**
1	$K_{dimerization}$	330 ± 50	22 ± 6	22 ± 2	<1
2	NMe_4Cl	40 ± 15	27 ± 12	32 ± 14	22 ± 1
3	$BnNMe_3Cl$	47 ± 3	44 ± 1	41 ± 2	55 ± 1
4	nBuNMe$_3$I[b]	100 ± 20	35 ± 10	40 ± 10	22 ± 1
5	AChCl	120 ± 8	44 ± 1	30 ± 6	24 ± 1
6	Kme_3Cl	250 ± 9	22 ± 19	55 ± 27	24 ± 1
7	NEt_4Cl	180 ± 10	55 ± 15	80 ± 85	28 ± 1
8	NPr_4Cl	1100 ± 210	60 ± 20	70 ± 30	28 ± 1
9	NBu_4Cl	7060 ± 2100	90 ± 10	145 ± 55	23 ± 1
10	$HNMe_3Cl$	96 ± 10	8 ± 8	20 ± 10	80 ± 9
11	H_2NMe_2Cl	50 ± 20	n.b.[c]	26 ± 33	50 ± 20
12	H_3NMeCl	16 ± 27	n.b.	n.b.	n.b.

[a] Phosphate buffered D_2O (50 mM Na_2HPO_4/NaH_2PO_4) at pH 7.0 (pD 7.4). $K_{dimerization}$ is defined as the association constant (β_{HH}) for host dimerization while K_{assoc} is defined as β_{HG}, the association constant for the 1:1 host-guest complex. In all cases, $K_{dimerization}$ was first determined via dilution titrations of host alone, and then used as a constant in the fitting of host-guest titration data to determine the true value of K_{assoc}. [b] The iodide salt of this guest was used. The identity of counter ion for NMR studies in water has been shown to be negligible.[98,99] [c] n.b. = no binding. (Reproduced by permission of The Royal Society of Chemistry).

Similar trends in chemical shifts upon binding were observed across the hosts irrespective of guest. Host signals identified in Figure 2.6 as being most shielded (upfield) in water (Ar-H core, N-CH_2, 7-H indole and 2-H indole) were generally found to have downfield chemical shifts when a cationic guest was introduced. This suggested an opening of the aromatic rings of the indoles upon binding. The greatest downfield change in chemical shift ($\Delta\delta$) was observed for the protons closest to the benzene core: Ar-*H* (for hosts **2.1**, **2.3** and **2.4**) and Ar-OC*H*$_3$ (for host **2.2**). These shifts are consistent with a cationic guest positioned directly above the central ring, interacting both with it and the now face-on indole rings in an ideal bowl or “cage-like” conformation.

*2.6.1 Binding to **2.1** – cation-pi effect vs. hydrophobic driving forces*

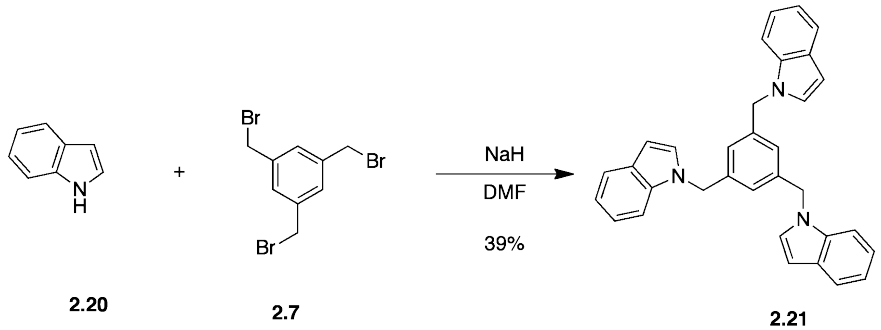
As shown in Table 2.1, host **2.1** exhibits a significant range in its association constants. For those guests containing a trimethyllysine-like R-NMe_3^+ group (entries 2-6) there is variation ranging from 40 M^{-1} to 250 M^{-1}. The simplest binding of a quaternary ammonium ion would be tetramethylammonium (entry 2); this could be considered a baseline of interaction with these types of hosts. For guests containing a single, simple alkyl substitution (*n*$BuNMe_3I$, entry 4), binding to host **2.1** in water becomes stronger as the length of the non-polar alkyl chain increases (from one carbon in NMe_4^+ to four carbons in *n*$BuNMe_3^+$). Given that the overall charge between these guests has not changed, this reflects an increased contribution from the hydrophobic effect. When the single alkyl substitution contains heteroatoms, such as the ester in AChCl (entry 5) and the amino acid head group in Kme3 (entry 6), the binding affinity to host **2.1** is again increased. While it is likely that there is an increased hydrophobic force associated with these larger guests, we cannot rule out additional polar interactions between host and guest. Binding to Kme3, for example, is significantly stronger than the other guests of this type, leading us to suggest additional favourable electrostatic interactions between the ammonium head group of Kme3 and the host. Binding to protein structures by other recognition elements (antibodies, or co-factors) is often sequence dependent, indicating that additional contacts from the surrounding area are important factors in both binding strength and specificity. Interestingly, the aromatic quaternary ammonium $BnNMe_3^+$ (entry 3), does not result in a significant increase in the binding constant compared to NMe_4^+. We hypothesize that there is a structural basis for this in that the short tether between the benzyl aromatic ring and then ammonium head group prevents a close contact between cation and benzene core, resulting in a smaller binding constant.

Binding to primary, secondary and tertiary ammonium ions, (entries 10 – 12) were included as models of binding to lysine, monomethyllysine and trimethyllysine. These guests showed weaker binding constants that diminished as methyl groups were removed. This was attributed to increased competition with the aqueous environment that could better solvate these guests as the number of hydrogen bonding NH's increased. This result could be expected given that it is often observed in protein binding events that more highly methylated ammonium ions bind more strongly to their aromatic partners (indeed that is the basis for selective signaling by trimethylated lysine residues[89]), and that trend is reproduced in the current model system.

We also examined the binding of larger quaternary ammonium ions (NR_4^+) by increasing the length of all four alkyl substituents across the series from methyl to *n*-butyl (entries 2, 7 – 9). Since the cation-pi effect has been shown to be a primary driving force for the interactions of indoles (in the form of tryptophan side chains) and quaternary ammonium ions,[49,100] we hypothesized that binding would decrease upon increasing alkylation. In general, larger cations should produce weaker cation–pi interactions because (a) the larger cations have a larger radius of interaction between pi surface and positive charge and (b) larger cations bear a more diffuse positive charge, creating a weaker electrostatic interaction with electron-rich pi surfaces. Unexpectedly, we observed a significant increase in binding, up to 175 times that of the baseline NMe_4^+ in the case of tetrabutyl ammonium. This result brought to light an assumption that we made regarding the weak forces at work in this binding interaction.

Host **2.1** was designed to present a significant amount of electron-rich pi surface with which to encourage binding via the cation-pi effect. However, this result clearly indicates that it is likely that cation-pi makes up only a small part of the driving force behind these interactions. There is no change in the electrostatics in this series, leading us to question the role of the hydrophobic effect in these binding effects.

To test whether hosts of the type of **2.1** bind using hydrophobic or cation-pi interactions, we synthesized host **2.21** (Scheme 2.4).[101] This host presents the same aromatic surface area as **2.1**, **2.3**, and **2.4** but without the carboxylate arms. We performed the same NMR titrations as **2.1** with the larger guests but this time in deuterated chloroform. $CDCl_3$ is an organic solvent that does not screen electrostatic interactions, unlike water. As a result, any cation-pi interactions present should appear strengthened in $CDCl_3$ compared to D_2O. At the same time, $CDCl_3$ can solvate nonpolar elements of both host and guest to nullify any contribution from the hydrophobic effect.



2.20 **2.7** **2.21**

Scheme 2.4 Chloroform-soluble host **2.21**

The results of the chloroform titrations are presented in Table 2.2. Note that AChCl was used in place of NMe_4^+ for solubility reasons. While binding in water had a marked upward trend as the alkyl chains became longer, binding in chloroform was weak to start and subsequently decreased. Taken together, these results are better explained by the surface-area-dependent hydrophobic effect. The complete lack of binding of NPr_4^+ and NBu_4^+ in $CDCl_3$ – where no hydrophobic effect can participate – confirms the importance of the hydrophobic effect for the binding of cations by **2.1** in water.

Table 2.2 Binding affinities compared between **2.1** in buffered D_2O and **2.21** in $CDCl_3$

		K_{assoc} (M^{-1})	
Entry	Guest	**2.1** D_2O	**2.21** $CDCl_3$
1	AChCl	120 ± 8	5 ± 5 (4 ± 5)[a]
2	NEt_4Cl	180 ± 10	2 ± 1
3	NPr_4Cl	1100 ± 210	n.b.
4	NBu_4Cl	7060 ± 2100	n.b.

[a] Performed as an "inverse" titration with host added to an existing solution of guest.

*2.6.2 Comparison of **2.1** vs. **2.2** – effect of methoxy substituents*

The addition of methoxy substituents to the benzene core in host **2.2** was intended to increase the electron-density available in the pi system and strengthen the resulting cation-pi interaction. However, the methoxy substituents were found to have significantly diminished binding relative to host **2.1**, especially with hydrophobic guests. A comparison of dimerization constants reveal **2.2** to be an order of magnitude less than **2.1** (entry 1). Presumably, the methoxy groups contribute to steric gearing[102] and can bias the indole arms of **2.2** to favor the closed "propeller" conformation (Figure 2.8) regardless of solvation effects or the presence of guests. This is consistent with the observation that host **2.2** has the fewest protons changing chemical shift upon moving from DMSO to water – its conformation in both solvents biased towards into the propeller shape. This also decreases the ability of **2.2** to associate with larger, hydrophobic guests (entries 7 – 9) that require an open and flexible host conformation. Similar diminished affinities (relative to **2.1**) are seen with almost all of the guests studied. It is clear that the added methoxy groups do not favour a conformation suitable for guests binding.

*2.6.3 Comparison of **2.1** vs. **2.3** – effect of methylenes in propionate chains*

Host **2.3**, with the water-solublizing carboxylates attached directly to the indole 3-position, was studied in order to observe the effect that the $-CH_2CH_2-$ units in the indole arms of **2.1** would have on the strength of its binding interactions. Guest binding to host

2.3 in all cases is diminished relative to **2.1** (Table 2.1), and the effects are largest for the most hydrophobic guests. This is consistent with the assertion that the hydrophobic effect plays a dominant role in driving cation binding to **2.1**.

Interestingly, host **2.3** binds all quaternary R-NMe_3^+ guests with approximately the same affinity (entries 2 – 6), regardless of their degree of hydrophobic character. Given the similarity in structure, the aromatic skeletons of hosts **2.1** and **2.3** should have very similar conformational preferences (as shown earlier) and abilities to bind guests through their identical aromatic rings. Thus, for guests that do not cause a large hydrophobic effect, the magnitude of binding should be similar. This is clearly evident for guests with little hydrophobic contribution such as NMe_4^+ (entry 2; 40 M^{-1} vs. 32 M^{-1}, for **2.1** and **2.3,** respectively). For greasier guests where the hydrophobic effect is likely to be larger, the disparity in binding between **2.1** and **2.3** increases to a 50-fold difference for the largest cation, NBu_4^+ (entry 9; 7060 M^{-1} vs. 145 M^{-1}). The largest guest, NBu_4^+, had the strongest effect on the host propionate chains, shifting the methylene protons 0.1 ppm downfield and indicating the strongest involvement of the alkyl chain in binding. These small shifts are most consistent with the small deshielding influence of a cationic alkylammonium binding partner.

The difference in surface area per -CH_2CH_2- group between **2.1** and **2.3** is calculated to be 41 $Å^2$. Assuming that *all* surface area of the CH_2CH_2 in **2.1** is buried upon complexation (which must be an over-estimate of the surface area actually involved), this would lead to a prediction of 10 kcal/mol difference in binding energy (3.3 kcal/mol x 3 -CH_2CH_2-) between the two,[103] which is larger than that observed for NBu_4^+ (2.3 kcal/mol). The small downfield shifts of -CH_2CH_2- protons upon binding NBu_4^+ of **2.1** are consistent with the formation of close contacts with the large cationic guest. Thus, although the deletion of the CH_2CH_2 groups seems intuitively to be a modest change of a peripheral group, the observed large changes in affinities are best interpreted as arising from the participation of those arms as important hydrophobic binding elements.

*2.6.4 Comparison of **2.3** vs. **2.4** – effect of carboxylate position*

We were also interested in what effect geometry would have on tripodal hosts of this nature. In all cases, the three indole rings need to adopt an open conformation to allow a cation to interact in a “face-on” manner with all four aromatic surfaces. Hosts **2.3** and **2.4** are structural isomers; they have the same number and identity of atoms and therefore the same degree of hydrophobicity. Their difference lies only in the position of

the carboxylate and thus, the preferred geometry of the molecule. Experimentally, host **2.4** has no tendency to self-associate in water and binds all quaternary cations with approximately the same value regardless of size, shape or hydrophobic character. All binding constants are approximately the same as for NMe_4^+, the baseline guest molecule. This differs from host **2.3**, which has a small dimerization constant and a weak but measurable preference for hydrophobic guests. For host **2.4**, binding with $BnNMe_3^+$ stands out as having the highest binding constant for guests with the R-NMe_3^+ motif.

In looking at Figure 2.8, carboxylates at the 2-positions of host **2.4** have the effect of drastically changing the geometric preferences of the indole arms without guest compared to host **2.3**. Host **2.4** also has the smallest change in proton chemical shifts of all the similar, unrestricted hosts (**2.1**, **2.3** and **2.4**) indicating that it likely adopts a geometry that is highly similar in aqueous and organic solution (which was also supported by ROE contacts). By moving the polar carboxylate groups to the indole-2 position, we have effectively excluded them from the binding pocket making them less able to participate in any electrostatic interaction with an incoming cation. For smaller R-NMe_3^+ guests that require little adjustment of the binding pocket, this translates into slightly smaller binding constants than for host **2.3** that could use its carboxylates in a favourable manner on the upper rim of the binding pocket. As we branch into larger quaternary ammonium cations, there is little change in binding as now host **2.4** must disrupt its preferred conformation and adjust to accommodate a larger guest. So while there is an increase in the hydrophobic association between host and the larger guests, the entropic penalty associated with opening the host binding pockets could negate any gain in binding affinity.

2.7 Calculated binding geometries

Energy minimizations in water were also used to evaluate the geometry of each host upon guest binding. For the typical guest NMe_4^+, the collapsed conformations of the empty hosts in each case convert to a more open conformation capable of making favourable contacts with guests by rotating the indole arms about the N-C single bond (Figure 2.10). The process of creating a pocket for the guest changes the chemical environment experienced by the host protons as shown in the solution phase data. Protons which were most shielded in the collapsed form are now deshielded as they are moved out of shielding regions (for example see host **2.1** in Figure 2.11). In addition, the aromatic rings are now involved in interacting with a cation. While this would shield any

protons attached to the cation, the aromatic rings themselves (and thus their protons) experience some deshielding (or a return to the more normal deshielded state).

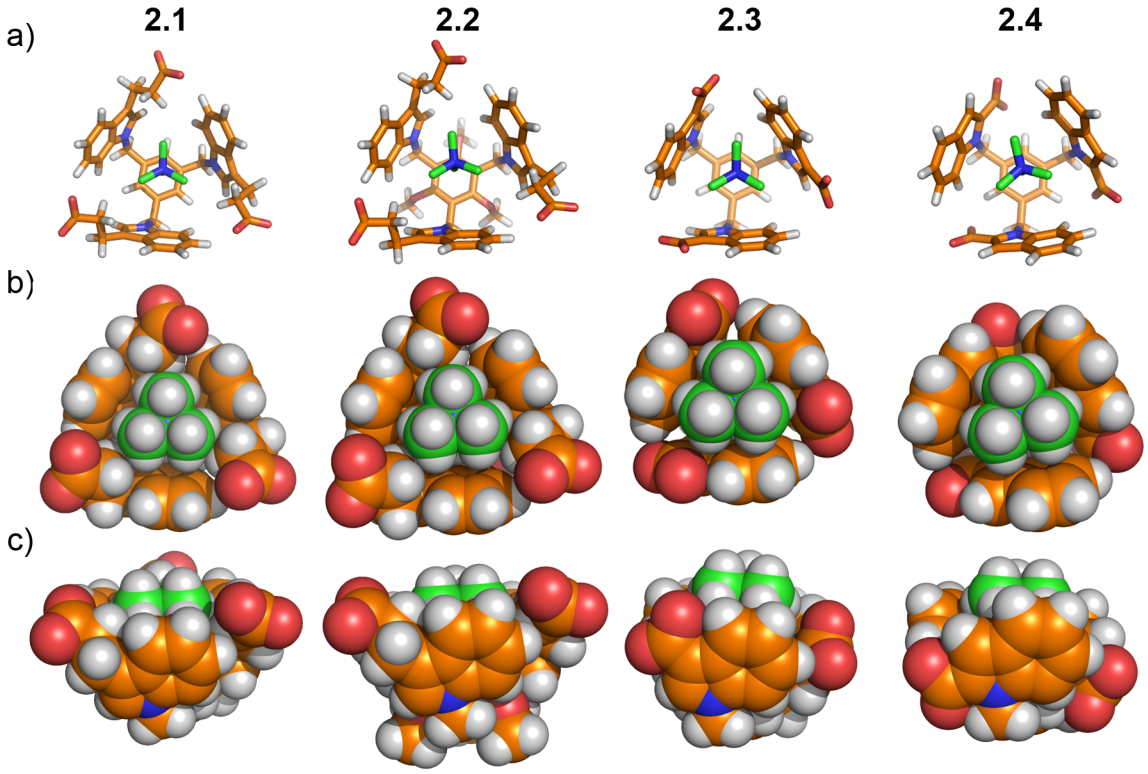


Figure 2.10 Equilibrium geometry in implicit water for **2.1**, **2.2**, **2.3** and **2.4** with NMe_4^+ (HF/6-31G* as implemented in Spartan '10).[13] a) Top view, ball and stick model; b) Top view, space-filling model and c) Side view, space-filling model.

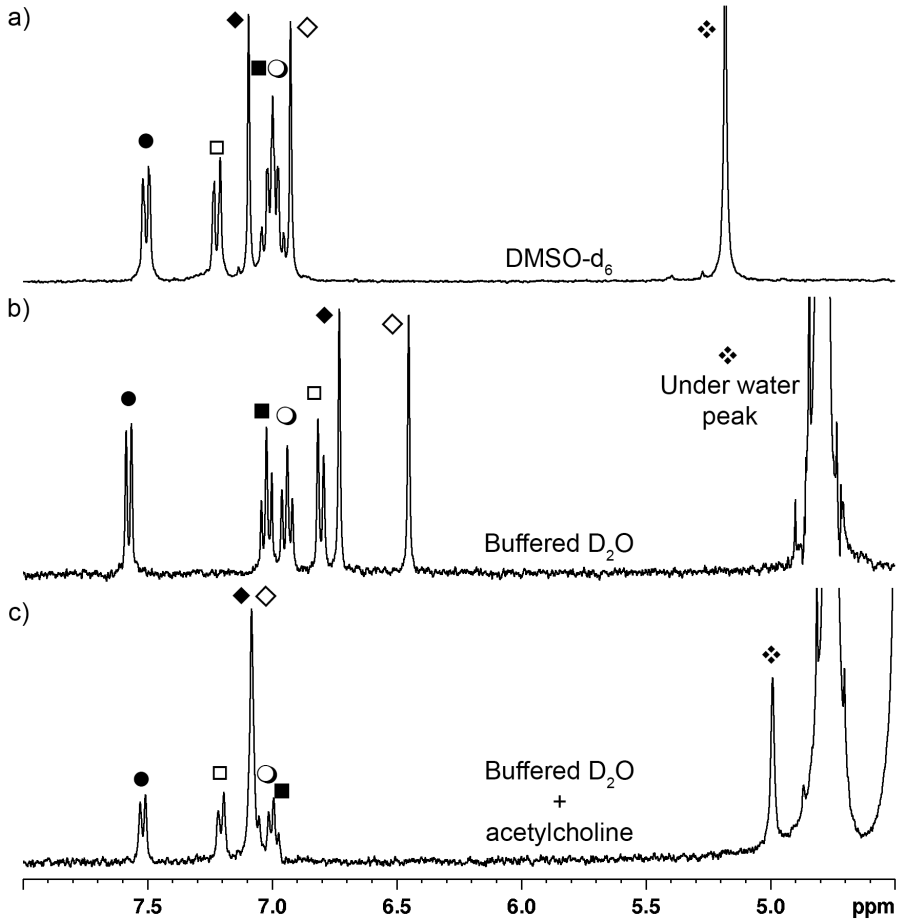


Figure 2.11 Chemical shift changes of host **2.1** in a) DMSO, b) D_2O (50 mM phosphate buffer), and c) D_2O (50 mM phosphate buffer) with 70 equivalents acetylcholine guest (1 mM host concentration).

The models of hosts **2.1** and **2.2** indicate that the methylene region of the propionate arms come into contact with cationic R-NMe_3^+ region of the guests. While the resulting change in NMR chemical shift might not be great for these protons, this would be an additional weak interaction that would give these hosts an advantage over the others. Indeed, the largest observed shift for these protons was only 0.1 ppm downfield and only for host **2.1** with NBu_4^+, the most hydrophobic guest. Still, the increased hydrophobic surface area was shown to play a significant role in the overall binding strength of these hosts given the differences in binding affinity, especially between hosts **2.1** and **2.3**. As well, these models show that the carboxylate groups of the 3-substituted hosts are associated with the upper rim of the host pocket, in position to interact favourably with cationic guests.

The models of hosts **2.3** and **2.4** demonstrate the physical difference that results from shifting the carboxylate from position 3 to 2. The carboxylates in **2.3** are more exposed, similar to those in hosts **2.1** and **2.2** and likely able to weakly participate in electrostatic interactions. In host **2.4**, those polar regions are buried in the base of the host cavity and less able to make contact with a cationic guest. Without the influence of electrostatics and only minimal hydrophobic contribution, host **2.4** could be considered a baseline for the magnitude of the cation-pi effect alone within this series of hosts in water.

2.8 Conclusions and future work

Within this series of highly similar indole-based hosts, two factors stand out as having the biggest influences on the strength of the host-guest interaction: 1) the amount of hydrophobic surface area available for binding (particularly for hydrophobic guests) and 2) the effect of small changes in functional group position on both geometric preference and binding affinity. The complementary, beneficial result of having additional hydrophobic regions on both host and guest is understandable as a consequence of a strong hydrophobic contribution within the aqueous environment. The more subtle effect of a small change in polar group position leading to a weaker electrostatic interaction was less expected and demonstrates our lack of ability to predict the outcome of such changes.

The importance of electron-rich tryptophan side chains (as an integral part of cation-pi interactions) in aromatic cages of been nicely shown through mutation studies of amino acids lining aromatic cage sites.[84] On the flip side, the importance of the cationic charge to those same interactions has also been investigated. Comparisons of *t*-butyl groups [R-C(Me_3)] and nearly isosteric trimethylammonium groups [R-N(Me_3)$^+$] have been elegantly used to demonstrate the importance of the positive charge in cation–pi interactions in protein–protein and drug-protein contacts.[49,104] Without the cationic charge present, weak to little binding is observed, and the authors of these studies conclude that the hydrophobic effect plays no role in the observed binding events.[49] Yet our studies of tris(indole) hosts **2.1**, **2.2** and **2.3** suggest that the hydrophobic effect can operate in this artificial tryptophan analog given a hydrophobic enough guest, and that the cation–pi interaction, if operative, is scarcely measurable. What differentiates the indoles of host **2.1** and those of proteins? Crystal structures of cation-binding proteins in bound and free states reveal almost no movement of the

aromatic cage side chains, suggestive of a highly rigid and preorganized binding pocket (Section 1.5.5).[105] It has been suggested that this rigidity is a hallmark of aromatic cage binding sites.[37] This rigidity is certainly lacking in these flexible hosts.

Though these hosts exhibit weak-to-moderately strong attraction to cations in water, there are important lessons here for future host designs targeting biologically relevant cations in water. Though we normally think of trying to build hosts that take advantage of as many attractive interactions as possible (hydrogen bonding, electrostatic and cation-pi interactions, for example), the complications that can be introduced by carrying out studies in pure water should not be underestimated. The current examples show that, in pure water, even subtle changes in pendant solublizing groups can have dramatic effects on host-guest affinities by affecting overall hydrophobicity and/or geometric preference, both of which could be loosely thought of as entropic considerations. While past studies have been concerned with the possible role of, for example, an anionic water-solublizing group vs. a neutral water-solublizing group,[78] the current results show that even apparently innocent changes in a small number of linker CH_2 groups can have a strong and determinant role in guest binding. Related structural elements in proteins, such as the methylene stretches of lysine, arginine, and glutamate side chains, are increasingly appreciated to be binding elements that play a strong role in determining the structures and interactions of folded proteins.[59,82,106-108] Designs for aqueous host-guest systems will need to continue to consider the hydrophobic aspects of small functional group changes as a key driver of their molecular recognition performance.

Moving forward from these results in this study of synthetic aromatic cage molecules, we recognize that both hydrophobic surface area and stronger pre-organization towards a cage-like geometry are important factors. Subsequent host scaffolds will need to incorporate both elements in order to have better success at strongly binding cations from water.

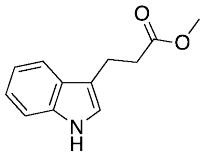
2.9 Experimental Section

2.9.1 General considerations

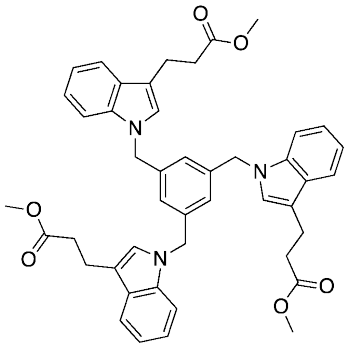
Solvents and reagents were used as obtained from Sigma-Aldrich. Proton (^{1}H) NMR and carbon (^{13}C) NMR spectra were recorded on a Brüker AC300 (300 MHz) spectrometer or Brüker Avance 500 (500 MHz) as indicated. 1D selective ROESY experiments were performed on a Brüker Avance 500 (500 MHz). Proton (^{1}H) NMR

spectra for NMR titration studies were recorded on a Brüker Avance 360 (360 MHz) spectrometer. Chemical shifts (δ) are given in parts per million (ppm) relative to TMS and referenced to residual protonated solvent ($CHCl_3$: δH 7.26 ppm, δC 77.16 ppm; DMSO δH 2.50 ppm, δC 39.52 ppm; DOH: δH 4.79 ppm).[109] *J* values are given in Hz. Abbreviations used are s (singlet), d (doublet), t (triplet), q (quartet), and m (multiplet). Infrared spectra were measured on a Thermo-Nicolet Nexus 670 FT-IR spectrometer with a resolution of 2 cm^{-1} using a Pike MIRacle attenuated total reflection (ATR) sampling accessory. Melting points were obtained using a Gallenkamp Melting Point Apparatus and are uncorrected. Low resolution electron impact mass spectra (LR-EIMS) and select high resolution electron impact mass spectra (HR-EIMS) were obtained on a double focusing Kratos Concept mass spectrometer. Remaining accurate mass measurements were done at the University of Victoria Chemistry mass spec facility on a Q-TOF II system by MicroMass. Samples of 1mg/ml diluted 1:100 in MeOH or CH_3CN were directly infused at a 5-10ul/min flow rate trough an ESI source. Spectra obtained at about 1 Hz scanning rate 100-2000 range with 60,000 resolution and less than 1ppm accuracy in most cases.

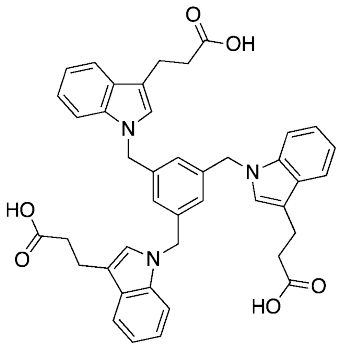
2.9.2 Synthetic procedures



Compound **2.6**. Synthesized using published procedure in 99% yield.[93] ^{1}H NMR (300 MHz; $CDCl_3$): δ 2.37 (t, *J* = 7.7, 2H, $CH_2CH_2CO_2Me$), 3.11 (t, *J* = 7.4, 2H, $CH_2CH_2CO_2Me$), 3.68 (s, 3H, CO_2Me), 6.98 (s, 1H, 2-H), 7.17 (m, 2H, indole-H), 7.34 (d, *J* = 7.8, 1H, indole-H), 7.61 (d, *J* = 7.8, 1H, indole-H), 8.00 (s br, 1H, N*H*).

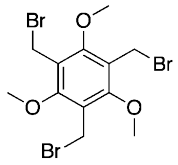


Compound **2.8**. NaH (0.24 g, 60% in oil, 6.1 mmol) and propionic indole **2.6** (1.2 g, 6.0 mmol) were suspended in anhydrous DMF (8 mL) and stirred under N_2. After 25 min, 1,3,5-tris(bromomethyl)benzene **2.7** (357 mg, 1.0 mmol) in DMF (2 mL) was added drop-wise and the reaction left to stir for 2 hours. After complete reaction of **2.7**, most solvent was removed via vacuum and the reaction was quenched with hexanes and wet THF. After filtration of solids, column chromatography (15 to 30% EtOAc/hexanes) yielded the product **2.8** (260 mg, 36%) as an off-white solid. Mp: 128-131°C. IR (neat), ν (cm^{-1}): 2951, 2926, 2855, 1741, 1737, 1728, 1469, 1251, 1172, 736. ^{1}H NMR (300 MHz; $CDCl_3$): δ 2.64 (t, *J* = 7.7, 6H, 3 x $CH_2CH_2CO_2Me$), 3.05 (t, *J* = 7.7, 6H, 3 x $CH_2CH_2CO_2Me$), 3.67 (s, 9H, 3 x CO_2Me), 5.09 (s, 6H, 3 x NCH_2Ph), 6.68 (s, 3H, 3 x 2-H), 6.79 (s, 3H, 3 x Ph*H*), 7.08 (m, 9H, 3 x indole-H), 7.57 (m, 3H, 3 x indole-H). ^{13}C NMR (75 MHz; $CDCl_3$): δ 20.8, 35.0, 49.7, 51.8, 109.8, 114.5, 119.2, 119.3, 122.1, 124.5, 125.6, 128.0, 136.7, 139.2, 174.0. LR-EIMS: *m/z* 723 (M^+, 8%), 552 (100), 536 (25), 203 (45), and 130 (45). HR-EIMS: *m/z* $[M]^+$ calculated for $C_{45}H_{45}N_3O_6$: 723.3308; found 723.3295.

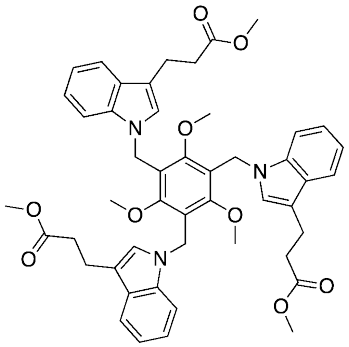


Compound **2.9**. Tri-substituted methyl ester indole host **2.8** (63 mg, 0.09 mmol) was dissolved in THF (2 mL) and distilled H_2O (2.5 mL). Excess $NaOH_{(s)}$ (91.5 mg, 2.3 mmol) was added and the reaction was stirred at room temperature for 21 hours under Ar. Reaction was diluted with an equal volume of 1 M $HCl_{(aq)}$, extracted with EtOAc (3 x 30 mL), and dried over $MgSO_4$ sulfate before concentrating under vacuum. Precipitation from a minimum of THF with hexanes and sonication gave the tri-acid **2.9** (53 mg, 89%) as a white solid. Mp: 156-160°C. IR (neat), ν (cm^{-1}): 3052, 2924, 1703, 1708, 1467, 1333, 1187, 740. ^{1}H NMR (500 MHz; $CDCl_3$): δ 2.71 (t, *J* = 7.1, 6H, 3 x $CH_2CH_2CO_2H$), 3.06 (t, *J* = 6.8, 6H, 3 x $CH_2CH_2CO_2H$), 5.06 (s, 6H, 3 x NCH_2Ph), 6.80 (s, 3H, 3 x 2-H), 6.87 (s, 3H, 3 x PhH), 7.10 (m, 9H, 3 x indole-H), 7.55 (m, 3H, 3 x indole-H). ^{13}C NMR (125 MHz; $CDCl_3$): 20.4, 34.6, 50.2, 109.9, 114.0, 119.0, 119.4, 122.1, 125.8, 126.2,

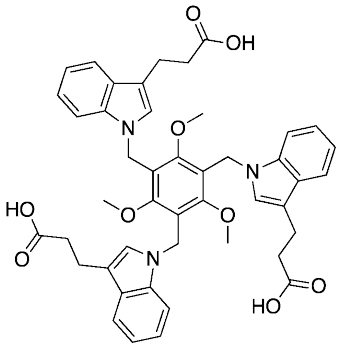
128.2, 136.8, 139.0, 179.5; LR-EIMS: m/z 681 (M^+, 30%), 189 (25), 130 (100), 83 (45). HR-EIMS: m/z $[M]^+$ calculated for $C_{42}H_{39}N_3O_6$: 681.2839; found 681.2801.



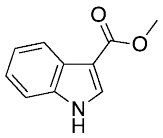
Compound **2.11**. Synthesized using published procedure in 30% yield.[110] ^1H NMR ($CDCl_3$, 300 MHz): δ 4.14 (s, 9H), 4.60 (s, 6H). ^{13}C NMR ($CDCl_3$, 75 MHz): δ 22.7, 62.9, 123.5, 160.3.



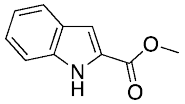
Compound **2.12**. 60% NaH (110 mg, 2.7 mmol) and indole ester **2.6** (550 mg, 2.7 mmol) were suspended in anhydrous DMF (5 mL) and stirred under N_2. After 40 minutes, 1,3,5-tris(bromomethyl)-2,4,6-trimethoxybenzene **2.11** (200 mg, 0.46 mmol) in DMF (2 mL) was added drop-wise and the reaction left to stir at room temperature for 17 hours. After complete reaction of **2.11**, most solvent was removed via vacuum and the reaction quenched with hexanes and wet THF. After salt removal via filtration, column chromatography (15 to 50 % EtOAc/hexanes) yielded **2.12** (127 mg, 34%) as a tan oil. IR (neat), ν (cm^{-1}): 3728, 3698, 3625, 3013, 2951, 2369, 2349, 2327, 1737, 1728, 1673, 1585 1467, 1193, 1107, 1093, 747. ^1H NMR ($CDCl_3$, 300 MHz): δ 2.73 (t, *J* = 7.7, 6H), 3.13 (t, *J* = 7.7, 6H), 3.45 (s, 9H), 3.68 (s, 9H), 5.28 (s, 6H), 6.97 (s, 3H), 7.16 (dt, *J* = 3.5, *J* = 0.6, 3H), 7.29 (dt, *J* = 4.0, *J* = 0.3, 3H), 7.61 (t, *J* = 8.9, 6H). ^{13}C NMR ($CDCl_3$, 75 MHz): δ 20.7, 35.1, 39.4, 51.5, 62.9, 109.8, 113.8, 118.9, 119.1, 121.4, 121.8, 125.1, 127.7, 136.6, 160.2, 173.7. HR-ESI-MS: *m/z* $[M+Na]^+$ calculated for $C_{48}H_{51}N_3O_9Na$: 836.3523; found 836.3519. (QTof).



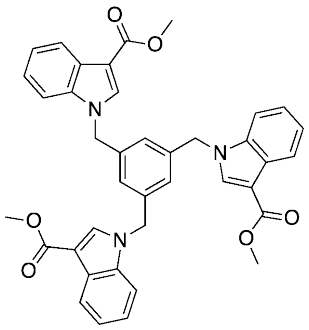
Compound **2.13.** Tri-substituted methyl ester indole **2.12** (127 mg, 0.16 mmol) was dissolved in THF (4 mL) and distilled H_2O (2 mL). Excess $NaOH_{(s)}$ (91.5 mg) was added and the reaction was stirred at room temperature for 18 hours under N_2. Reaction was diluted with an equal volume of 1 M $HCl_{(aq)}$, extracted with EtOAc (3 x 30 mL), and dried over $MgSO_4$ before concentrating under vacuum. Precipitation from a minimum of DCM with hexanes and sonication gave tri-acid **2.13** (90 mg, 75%) as an off-white solid. Mp: 117 – 120 °C (dec.). IR (neat), ν (cm^{-1}): 3024, 2957, 1708, 1581, 1463, 1249, 740. ^1H NMR ($CDCl_3$, 300 MHz): δ 2.71 (t, *J* = 6.5, 6H), 3.08 (t, *J* = 6.5, 6H), 3.20 (s, 9H), 5.24 (s, 6H), 6.90 (s, 3H), 7.11 (t, *J* = 7.2, 3H), 7.24 (t, *J* = 7.2, 3H), 7.48 (d, *J* = 8.2, 3H), 7.55 (d, *J* = 7.8, 3H). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 20.4, 34.7, 39.7, 62.7, 109.8, 113.4, 118.8, 119.3, 120.9, 122.0, 124.8, 127.8, 136.8, 160.8, 179.3. HR-ESI-MS: m/z $[M+Na]^+$ calculated for $C_{45}H_{45}N_3O_9Na$: calc: 774.3054; found, 774.3051. (QTof).



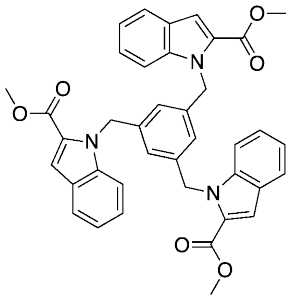
Compound **2.14**. Indole-3-carboxylic acid (2.00 g, 12.5 mmol) was dissolved in MeOH (23 mL) and concentrated H_2SO_4 (0.3 mL, 5.6 mmol) and heated to reflux for 3.5 hours. The reaction was cooled, poured into 75 mL ice, and extracted with DCM (3 x 60 mL). The combined organics were washed with saturated brine and saturated $NaHCO_{3(aq)}$, dried over $MgSO_4$, filtered and condensed. Column chromatography (20 to 25% EtOAc/hexanes) gave indole **2.14** (1.25 g, 57%) as a tan powder. Spectra matched known reference.[111] ^1H NMR (DMSO, 500 MHz): δ 3.80 (s, 3H), 7.19 (m, 2H), 7.47 (m, 1H), 7.99 (m, 1H), 8.07 (d, *J* = 3.0, 1H), 11.92 (br s, 1H). ^{13}C NMR (DMSO, 125 MHz): δ 50.6, 106.3, 112.3, 120.4, 121.2, 122.4, 125.6, 132.4, 136.4, 164.8.



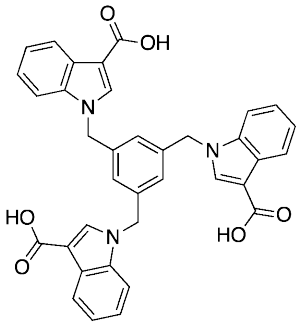
Compound **2.15**. Indole-2-carboxylic acid (3.00 g, 18.7 mmol) was dissolved in MeOH (30 mL) and concentrated H_2SO_4 (0.3 mL, 5.6 mmol) and heated to reflux for 15 hours. The reaction was cooled, diluted with saturated $NaHCO_{3(aq)}$ (50 mL), and extracted with EtOAc (2 x 40 mL). The combined organics were dried over Na_2SO_4, filtered and condensed. The crude solid was suspended in hexanes, sonicated, and then filtered to give indole **2.15** (2.80 g, 86% yield) as a tan powder. Spectra matched known reference.[111] ^{1}H NMR ($CDCl_3$, 300 MHz): δ 3.96 (s, 3H), 7.16 (t, *J* = 7.5, 1H), 7.23 (s, 1H), 7.33 (t, *J* = 7.6, 1H), 7.43 (d, *J* = 8.3, 1H), 7.70 (d, *J* = 8.0, 1H), 8.98 (br s, 1H). ^{13}C NMR ($CDCl_3$, 75 MHz): δ 52.2, 109.0, 112.1, 121.1, 122.9, 125.7, 127.3, 127.7, 137.1, 162.7.



Compound **2.16**. 60% NaH (245 mg, 6.1 mmol) and indole ester **2.14** (1.03 g, 5.9 mmol) were suspended in anhydrous DMF (8 mL) and stirred under N_2. After 25 minutes, 1,3,5-tris(bromomethyl)benzene **2.7** (359 mg, 1.0 mmol) in DMF (2 mL) was added drop-wise and the reaction left to stir at room temperature for 18 hours. Most solvent was removed via vacuum and the reaction quenched with hexanes and wet THF. After salt removal via filtration, column chromatography (15 to 50 % EtOAc/hexanes) yielded **2.16** (320 mg, 50% yield) as a light yellow solid. Mp: 165 – 167 °C (dec.). IR (neat), ν (cm^{-1}): 2925, 1703, 1692, 1530, 1536, 1243, 1180, 1092, 757, 748. ^{1}H NMR ($CDCl_3$, 300 MHz): δ 3.93 (s, 9H), 5.18 (s, 6H), 6.74 (s, 3H), 7.05 (d, *J* = 8.2, 3H), 7.18 (td, *J* = 4.1, *J* = 1.1, 3H), 7.27 (td, *J* = 7.5, *J* = 0.9, 3H), 7.73 (s, 3H), 8.18 (d, *J* = 7.9, 3H). ^{13}C NMR ($CDCl_3$, 75 MHz): δ 50.2, 51.0, 108.0, 110.0, 121.9, 122.2, 123.2, 124.9, 126.7, 134.2, 136.4, 138.1, 165.2. HR-ESI-MS: *m/z* $[M+Na]^+$ calculated for $C_{39}H_{33}N_3O_6Na$: 662.2267; found, 662.2264. (QTof).

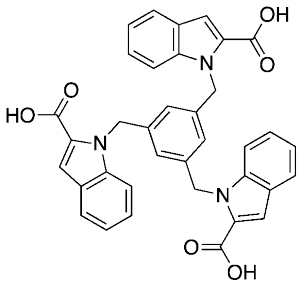


Compound **2.17**. 60% NaH (245 mg, 6.2 mmol) and indole ester **2.15** (1.05 g, 6.0 mmol) were suspended in anhydrous DMF (8 mL) and stirred under N_2. After 30 minutes, 1,3,5-tris(bromomethyl)benzene **2.7** (356 mg, 1.0 mmol) in DMF (2 mL) was added drop-wise and the reaction left to stir for 18 hours. Most solvent was removed via vacuum and the reaction quenched with hexanes and wet THF. After salt removal via filtration, column chromatography (50 to 80% DCM/hexanes) yielded **2.17** (172 mg, 27% yield) as a beige solid. Mp: 174 – 176°C (dec.) IR (neat), ν (cm^{-1}): 3050, 2948, 1712, 1709, 1249, 1197, 742. ^{1}H NMR ($CDCl_3$, 300 MHz): δ 3.65 (s, 9H), 5.52 (s, 6H), 6.44 (s, 3H), 6.96-7.14 (m, 9H), 7.17 (s, 3H), 7.57 (dd, *J* = 7.0, J = 1.1, 3H). ^{13}C NMR ($CDCl_3$, 75 MHz): δ 47.5, 51.5, 110.7, 111.2, 120.7, 122.6, 123.4, 125.2, 126.0, 127.1, 138.9, 139.3, 162.1. HR-ESI-MS: *m/z* $[M+Na]^+$ calculated for $C_{39}H_{33}N_3O_6Na$: 662.2267; found, 662.2263. (QTof).

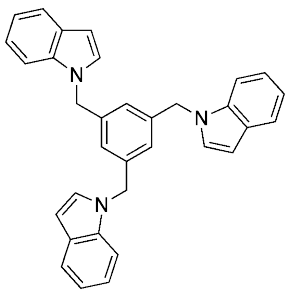


Compound **2.18**. Tri-substituted methyl ester indole **2.16** (122 mg, 0.19 mmol) was suspended in MeOH (12 mL). $KOH_{(s)}$ (221 mg) was dissolved in distilled H_2O (8 mL) and added drop-wise to the suspension. The reaction was heated to reflux for 16 hours until complete conversion observed by TLC (1:1 EtOAc/hexanes). Reaction was diluted with 1 M $HCl_{(aq)}$, extracted with EtOAc (3 x 30 mL), and dried over $MgSO_4$ before concentrating under vacuum. Solid was suspended in $CHCl_3$, filtered and air-dried to give tri-acid **2.18** (106 mg, 93% yield) as a tan solid. Mp: 261 – 263 °C (dec.) IR (neat), ν (cm^{-1}): 3055, 2938, 1665, 1659, 1536, 1531, 1278, 1253, 1190, 751. ^{1}H NMR (DMSO, 360 MHz): δ 5.38 (s, 6H), 7.08 (t, *J* = 7.6, 3H), 7.16 (t, *J* = 7.4, 3H), 7.23 (s, 3H), 7.32 (d,

J = 8.1, 3H), 8.00 (d, *J* = 7.9, 3H), 8.16 (s, 3H), 12.01 (br s, 3H). ^{13}C NMR (DMSO, 90 MHz): δ 49.4, 106.8, 111.0, 120.8, 121.3, 122.2, 126.4, 126.6, 135.4, 136.1, 138.1, 165.6. HR-ESI-MS: *m/z* $[M+Na]^+$ calculated for $C_{36}H_{27}N_3O_6Na$: 620.1797; found, 620.1797. (QTof).



Compound **2.19**. Tri-substituted methyl ester indole **2.17** (26 mg, 0.04 mmol) was suspended in MeOH (5 mL). $KOH_{(s)}$ (108 mg, 1.9 mmol) was dissolved in distilled H_2O (4 mL) and added drop-wise to the suspension. The reaction was heated to reflux for 17 hours until complete conversion observed by TLC (DCM). Reaction was diluted with 1 M $HCl_{(aq)}$, extracted with EtOAc (3 x 30 mL), and dried over $MgSO_4$ before concentrating under vacuum. Solid was suspended in $CHCl_3$, filtered and air-dried to give tri-acid **2.19** (23 mg, 93% yield) as a white solid. Mp: 262 – 264 °C (dec.) IR (neat), ν (cm^{-1}): 3035, 1721, 1687, 1678, 1673, 1519, 1267, 1200, 1171, 1136, 744. ^{1}H NMR (DMSO, 500 MHz): δ 5.65 (s, 6H), 6.64 (s, 3H), 7.10 (td, *J* = 3.9, *J* = 0.87, 3H), 7.15 (td, *J* = 3.5, *J* = 1.1, 3H), 7.19-7.23 (m, 6H), 7.65 (d, *J* = 7.9, 3H), 12.81 (br s, 3H). ^{13}C NMR (DMSO, 125 MHz): δ 46.8, 110.4, 111.0, 120.5, 122.3, 123.6, 124.7, 125.5, 127.7, 138.7, 139.0, 162.8. HR-ESI-MS: *m/z* $[M+Na]^+$ calculated for $C_{36}H_{27}N_3O_6Na$: 620.1797; found, 620.1795. (QTof).



Compound **2.21**. Synthesized using modification of published procedure.[101] 60% NaH (0.83 g, 20.8 mmol) and indole **2.20** (2.56 g, 22.0 mmol) were suspended in anhydrous DMF (30 mL) and stirred under Ar. After 45 min, 1,3,5-tris(bromomethyl)benzene **2.7** in

DMF (4 mL) was added drop-wise and the reaction left to stir 4 hours. After complete reaction of **2.7**, most solvent was removed via vacuum and the reaction quenched with hexanes and EtOAc. After filtration of solids, column chromatography (1:3 $CHCl_3$/hexanes) and precipitation in EtOH with cooling yielded the product **2.21** (653.3 mg, 39%) as an off-white solid. ^{1}H NMR (360 MHz, $CDCl_3$): δ 5.14 (s, 6H, 3 x NC*H*$_2$Ph), 6.52 (d, 3H, *J* = 3.1, 3 x 3-H), 6.72 (s, 3H, 3 x Ph*H*), 7.00 (d, *J* = 3.1, 3H, 3 x 2-H), 7.14 – 7.19 (m, 9H, indole-H), 7.64 (m, 3H, indole-H). ^{13}C (75 MHz, $CDCl_3$): δ 49.7, 101.9, 109.6, 119.6, 121.0, 121.8, 124.5, 125.1, 128.7, 136.1, 139.0.

2.9.3 K_{assoc} determination for 1:1 binding in theory

A crucial part of this chapter is the determination of solution-phase association (or binding) constants (K_{assoc}) by ^{1}H NMR spectroscopy. The supramolecular systems described in this chapter (and in the ones that follow) are dynamic equilibria; that is, they can form a complex between two (or more) components but are also free to dissociate reversibly. A generic equilibrium of compounds A and B is shown below. Compounds that have a higher affinity for one another will (at any given time) have more compound found in the bound or complexed form (AB). Complexes that have lower affinity for one another are likely to have less compound in a complexed state and more as the individual components (A and B).

$$A + B \rightleftharpoons AB$$

Figure 2.12 A simple equilibrium

One way to approach measuring association constants is to remember that K_{assoc} is simply a measure of the equilibrium between the two sides of the equation. Equilibrium constants can be determined from the equilibrium expression (Eq. 2.1), which is the ratio of product to reactant concentrations. [A], [B], and [AB] represent the concentration of free (unbound) A, the concentration of free B, and the concentration of the AB complex. K_{assoc} is also known as K_{11} since it represents a 1:1 complex of A and B. β_{11} is the overall binding constant for this binding event (in the 1:1 case, it is the same as K_{11}).

Eq. 2.1
$$K_{assoc} = \frac{[AB]}{[A][B]} = K_{11}$$

In this simple 1:1 system, each component (A or B) can only be found alone or in complex with each other. The concentrations in each state must sum to the total

concentration present ($[A]_t$) in order to follow the law of mass balance (Eq. 2.2) (likewise for $[B]_t$).

Eq. 2.2 $$[A]_t = [A]_{free} + [AB]$$

The ratios of the bound and unbound concentrations of A relative to total concentration present can be expressed as the fraction free and fraction bound.

Eq. 2.3 $$\text{fraction free} = f_{10} = \frac{[A]_{free}}{[A]_t}$$

Eq. 2.4 $$\text{fraction bound} = f_{11} = \frac{[AB]}{[A]_t}$$

Again, since A can only be found in two possible states (bound and unbound) the fraction unbound and fraction bound must add up to 100% of A as shown below.

Eq. 2.5 $f_{10} + f_{11} = 1$, then

Eq. 2.6 $f_{10} = 1 - f_{11}$

In a typical titration experiment, specific concentrations of B might be added into a solution of known concentration of A. As this occurs, the equilibrium redistributes the species present and the concentration of the complex AB increases while A decreases. The titration continues to a plateau where 100% AB has been formed. Experimentally, the signal that is proportional to the relative fraction of the product of the equilibrium (f_{11}) is plotted as the y-axis, while the x-axis is the concentration of B added.

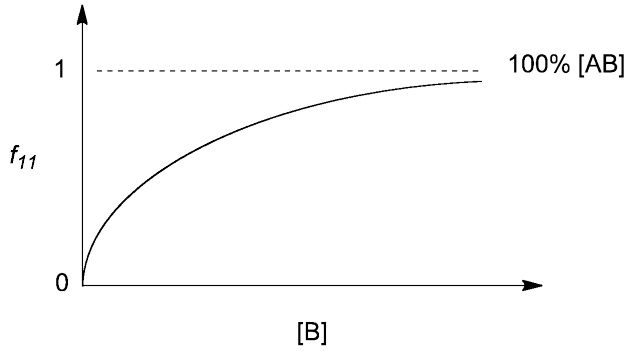


Figure 2.13 A hypothetical titration curve

In order to solve this curve to find K_{assoc}, we need is an expression for f_{11} that is dependent on B and includes the term we want to find, K_{assoc}. Recall that we know:

$$\text{fraction bound} = f_{11} = \frac{[\mathrm{AB}]}{[\mathrm{A}]_t}$$

Eq. 2.4

And from mass balance we know:

$$[A]_t = [A]_{free} + [AB]$$

Eq. 2.2

We can also rearrange the equilibrium expression from Eq. 2.1 to isolate [AB]:

$$[\mathrm{AB}] = K_{11}[\mathrm{A}][\mathrm{B}]$$

Eq. 2.7

Combining Eq. 2.4 first with Eq. 2.2, and then Eq. 2.7, generates a new expression for fraction bound (f_{11}), that now includes the association constant, K_{11}.

$$f_{11} = \frac{[\mathrm{AB}]}{[\mathrm{A}]_t} = \frac{[\mathrm{AB}]}{[\mathrm{A}]_{free} + [\mathrm{AB}]} = \frac{K_{11}[\mathrm{A}][\mathrm{B}]}{[\mathrm{A}]_{free} + K_{11}[\mathrm{A}][\mathrm{B}]}$$

Eq. 2.8

By cancellation of like terms, we arrive at the general form of a 1:1 binding isotherm (Eq. 2.9). A binding isotherm describes the theoretical change in the concentration of one component as a function of the concentration of another component at constant temperature.[112] This equation now describes the shape of the line shown Figure 2.13, our hypothetical titration curve: fraction bound, vs. the concentration of B. The only unknown in this expression now is K_{11}, the association constant we want to determine. We can also specifically include [AB] (from f_{11}) in the general isotherm as shown in Eq. 2.10.

$$f_{11} = \frac{K_{11}[\mathrm{B}]}{1 + K_{11}[\mathrm{B}]}$$

Eq. 2.9

$$[\mathrm{AB}] = \frac{[\mathrm{A}]_t K_{11}[\mathrm{B}]}{1 + K_{11}[\mathrm{B}]}$$

Eq. 2.10

A number of assumptions have gone into Eq. 2.10. One is that $[A]_t$ remains constant over the course of the titration; this can be done quite simply (and is how we perform the experiments). The second assumption is about [B]. We don't actually know [B] and approximate it to be $[B]_t$ at any given data point. This is only applicable in cases where $[B]_t >> [A]_t$. In that case, $[A]_t$ limits the amount of AB that can be formed, and very little B becomes AB (thus $[B] \approx [B]_t$).

Where the above approximations don't hold, we can get around the problem of uncertainty in the [B] by instead using a 1:1 binding isotherm that includes both $[A]_t$ and

$[B]_t$. This can be derived again starting from the equilibrium expression (Eq. 2.1) and using mass balance equations for both A and B.

$$K_{11} = \frac{[AB]}{[A][B]} = \frac{[AB]}{([A]_t - [AB])([B]_t - [AB])}$$

$$K_{11} = \frac{[AB]}{[A]_t[B]_t - [A]_t[AB] - [B]_t[AB] + [AB]^2}$$

Eq. 2.11

Inversion of the equation, multiplying by [AB]. moving all terms to one side, and collecting like terms leads to a quadratic equation, Eq. 2.15.

$$\frac{1}{K_{11}} = \frac{[A]_t[B]_t - [A]_t[AB] - [B]_t[AB] + [AB]^2}{[AB]}$$

Eq. 2.12

$$\frac{[AB]}{K_{11}} = [A]_t[B]_t - [A]_t[AB] - [B]_t[AB] + [AB]^2$$

Eq. 2.13

$$0 = [A]_t[B]_t - [A]_t[AB] - [B]_t[AB] + [AB]^2 - \frac{[AB]}{K_{11}}$$

Eq. 2.14

$$0 = [AB]^2 - \left([A]_t - [B]_t + \frac{1}{K_{11}}\right)[AB] + [A]_t[B]_t$$

Eq. 2.15

Solving Eq. 2.15 for [AB] using the quadratic formula leads to a general 1:1 binding isotherm where $[A]_t$ is constant, K_{11} is a variable, [AB] is measured and $[B]_t$ is known (Eq. 2.16). It is this equation that we will eventually use to solve for K_{11} in our systems.

$$[AB] = \frac{\left([A]_t - [B]_t + \frac{1}{K_{11}}\right) \pm \sqrt{\left([A]_t - [B]_t + \frac{1}{K_{11}}\right)^2 - 4[A]_t[B]_t}}{2}$$

Eq. 2.16

2.9.4 K_{assoc} determination for 1:1 binding in theory by ^{1}H NMR spectroscopy

For these studies, we have chosen to use ^{1}H NMR spectroscopy as our method of concentration determination. We assume that a binding event will change the chemical environment of particular protons, which can be observed as a change in chemical shift of those protons. What we need is a way to relate those observable changes in chemical shift to changes in the relative concentration of the species in solution.

Dynamic equilibria observed by NMR can be classified on whether the exchange between bound and unbound states is fast or slow compared to the NMR time scale of the nucleus under observation. The chemical shift of a given proton in the bound and unbound states is different, and those two shifts are separated by a given frequency ($\Delta\nu$, in Hz) depending on the nucleus and instrument being used. If the frequency at which the system samples the bound and unbound states is faster than $\Delta\nu$, then it is "faster" than the NMR timescale. Conversely, exchange rates slower than $\Delta\nu$ result in systems that are slow on the NMR timescale.[113] How fast a system exchanges depends on the energetic barrier between the states, and at what temperature the experiment is run.

In the case of slow exchange, at any given concentration of A and B, NMR signals for both the bound and unbound components are observed in the same spectra as distinct peaks (labelled δ_{bound} and δ_{free}, Figure 2.14). By integration of the NMR peaks, the molar ratio of each species can be determined. From there, the concentration of each species can be determined using mass balance equations and the original total concentration used. Finally, K_{assoc} can be calculated directly using Eq. 2.1. With this method, in theory only one data point is needed to calculate K_{assoc}.

$$\underset{(\delta_{free})}{A} + B \rightleftharpoons \underset{(\delta_{bound})}{AB}$$

Figure 2.14 NMR signals in a simple equilibrium

More commonly (and as is the case for our hosts), the system under study has fast exchange on the NMR time scale. At a given concentration of A and B, each NMR signal is coalesced into a single peak (δ_{obs}) that is a weighted average of the signals from both the bound and unbound states. We think of this as a continuum from 0 to 1, where the distance between the free and bound peaks represents how much of compound vs. complex is present. For example, in Figure 2.15, a peak for δ_{obs} is present at 60% of the distance between δ_{free} (0% bound) and δ_{bound} (100% bound). At this concentration of A and B, 60% of A would be present as AB (f_{11} = 0.6) and 40% would still be A (f_{10} = 0.4).

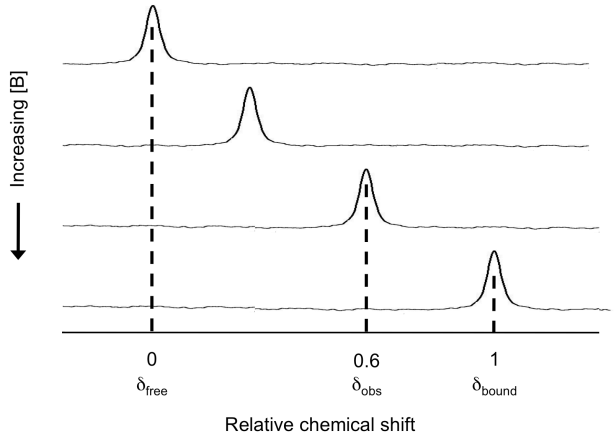


Figure 2.15 Measurement of relative fraction bound by peak position

We can therefore use the position of the observed peak relative to δ_{free} and δ_{bound} to determine the ratio of the total compound that is in the free and bound state. Observed chemical shift is expressed as a weighted average relative to the free and bound fractions (or mole ratios).

Eq. 2.17 $$\delta_{obs} = \left(f_{10} \times \delta_{free}\right) + \left(f_{11} \times \delta_{bound}\right)$$

If we substitute Eq. 2.6 into Eq. 2.17, we can see that an experimentally determined δ_{obs} can be used to find fraction bound (f_{11}), provided that δ_{free} and δ_{bound} are known (Eq. 2.18). δ_{free} can be determined by observing the chemical shift of a single component alone in solution. δ_{bound} can be estimated by observing the shape of the curve that is generated over the course of a titration. Theoretically, after a given concentration, the change in chemical shift will reach a plateau (100% bound) where no further change is possible. Experimentally, δ_{bound} is not often known exactly and becomes a variable to determine.

Eq. 2.18 $$\delta_{obs} = \left(1 - f_{11}\right)\left(\delta_{free}\right) + \left(f_{11}\right)\left(\delta_{bound}\right)$$

After expanding the brackets and collecting like terms, this becomes:

Eq. 2.19 $$\delta_{obs} = \delta_{free} - \left(f_{11}\right)\left(\delta_{free}\right) + \left(f_{11}\right)\left(\delta_{bound}\right)$$, then

Eq. 2.20 $$\delta_{obs} = \delta_{free} + \left(f_{11}\right)\left(\delta_{bound} - \delta_{free}\right)$$

Further manipulation can be used to define the δ measurements in terms of Δδ, the change in chemical shift rather than the absolute chemical shifts themselves. Rearrangement of Eq. 2.20 leads to:

Eq. 2.21 $$\delta_{obs} - \delta_{free} = f_{11}\left(\delta_{bound} - \delta_{free}\right)$$

We can define each Δ term as:

Eq. 2.22 $$\delta_{obs} - \delta_{free} = \Delta\delta_{obs}$$

Eq. 2.23 $$\delta_{bound} - \delta_{free} = \Delta\delta_{max}$$

And substitute both into Eq. 2.21:

Eq. 2.24 $\Delta\delta_{obs} = f_{11} \times \Delta\delta_{max}$, then

Eq. 2.25 $$f_{11} = \frac{\Delta\delta_{obs}}{\Delta\delta_{max}}$$

Again, we've arrived at a linear weighted average for the placement of the NMR signal. This expression intuitively makes sense when we look at the original interpretation of f_{11} in an NMR context. $\Delta\delta_{obs}$ is the position of the chemical shift signal between the maximum of two theoretical points – zero bound and 100% bound (Figure 2.15).

What we need to do now is relate what we can measure (chemical shift data, $\Delta\delta_{obs}$) and what we know (concentrations of A and B), to what we want to know (K_{assoc} or K_{11}). We can use the 1:1 binding isotherms determine previously. To make a binding isotherm specific to an NMR titration (as opposed to any other method to detect a change in the ratio of species concentration), we need to include the $\Delta\delta_{obs}$ variable. In Eq. 2.25, we just related fraction bound f_{11} to $\Delta\delta_{obs}$. Substituting Eq. 2.25 in to the first general 1:1 binding isotherm (Eq. 2.9) gives us Eq. 2.26. Further manipulation of Eq. 2.26 leads to Eq. 2.27, a 1:1 binding isotherm for NMR titrations.

Eq. 2.26 $$f_{11} = \frac{K_{11}[B]}{1 + K_{11}[B]} = \frac{\Delta\delta_{obs}}{\Delta\delta_{max}}$$

Eq. 2.27 $$\Delta\delta_{obs} = \frac{\Delta\delta_{max} K_{11}[B]}{1 + K_{11}[B]}$$

In this equation, we measure $\Delta\delta_{obs}$ while [B] is varied. $\Delta\delta_{max}$ and K_{11} are variables to be determined. This is a non-linear dependence of y ($\Delta\delta_{obs}$) upon x ([B]), so methods

have been developed to either a) make it linear to be solved (such as the Benesi-Hildebrand double reciprocal plot[114] or the Scatchard x-reciprocal plot)[115] or b) work with the non-linearity using modern computer fitting programs.

As stated earlier, the first 1:1 binding isotherm makes assumptions about [B]. In practice for our systems, we need to use the second isotherm (Eq. 2.16). This can again be converted to an NMR specific isotherm using our NMR equation for fraction bound (Eq. 2.25). Rearranging for [AB] and substituting into Eq. 2.16, gives us another 1:1 binding isotherm for NMR titrations (Eq. 2.28).

Eq. 2.25
$$f_{11} = \frac{\Delta\delta_{obs}}{\Delta\delta_{max}} = \frac{[AB]}{[A]_t}$$

Eq. 2.28
$$\Delta\delta_{obs} = \frac{\Delta\delta_{max}\left([A]_t - [B]_t + \frac{1}{K_{11}}\right) \pm \sqrt{\left([A]_t - [B]_t + \frac{1}{K_{11}}\right)^2 - 4[A]_t[B]_t}}{2[A]_t}$$

In practice, we work with the non-linear data using programs that have been developed to fit these such as HypNMR[97] or Solver in Excel. Using Eq. 2.28, each chemical shift data point (δ_{calc}) is calculated using the known (fixed) values for δ_{free}, $[A]_t$, $[B]_t$ (at each individual data point), as well as "guesses" or initial values for δ_{bound} and K_{11}. The difference between $\Delta\delta_{obs}$ and $\Delta\delta_{calc}$ at each data point is calculated and summed up as residuals. Through repeated iterations, the values for δ_{bound} and K_{11} are varied such that the sum of the least squares of the residuals becomes smaller and eventually stops changing.

This method can be used to calculate an association constant for each proton that experiences a change in chemical shift upon binding in a given titration. This could generate an individual K_{11} for each proton. Rather than average the value for each proton (giving equal weight to each), we use a global fitting method in HypNMR. In this, one value of K_{11} is determined using the data from every proton present with one overall error. This produces a more reliable K_{11}, as it corresponds to the entire data set.[64]

2.9.5 K_{assoc} determination in practice by ^{1}H NMR spectroscopy

Binding titrations in water were carried out in D_2O containing Na_2HPO_4/NaH_2PO_4 (total phosphate concentration of 50 mM) at pH 7.0 (pD 7.4). Acidity of buffered solutions in D_2O (pD) was measured using a saturated KCl pH electrode and corrected to pD by addition of 0.4 units to the measured pH reading.[116] All tri-acid hosts were prepared as

described in Section 2.9.2 and converted to their tri-sodium salts with 3 equivalents NaOMe prior to titration. pH variation of each host in 50 mM phosphate buffered D_2O from pH 5 – 10 (pD 5.4 – 10.4) showed no pH-induced chemical shift changes of protons expected to be adjacent to or affected by the protonation state of the carboxylates. This suggests that the four hosts studied remain in their deprotonated state as three carboxylates at the pH used for the following titrations. At the least, we can conclude that no protonation/deprotonation equilibria exist within ≥2 log units on either side of the experimental pH.

All guests were commercially available and used as received, with the exception of n-butyl trimethylammonium iodide ($BuNMe_3I$). This was synthesized from N,N-dimethylbutylamine and methyl iodide using a published procedure.[117] Compound spectra were a match to the known product.

The 1:1 binding situation described above in Section 2.9.4 becomes slightly more complicated when δ_{free} isn't δ_{free}; that is, when one component is able to bind or complex with itself, which changes the chemical environment of its protons. This is the case for host **2.1**, **2.2** and **2.3**. For these hosts, the NMR spectra showed measurable changes in proton chemical shifts with respect to host concentration. Consequently, NMR titrations with guest for these hosts would result in competing equilibria – one for host with guest, and one for host with host (Figure 2.16). A standard NMR titration will only show the concentration dependent chemical shift changes as a combination of the two processes. Consequently, host dimerization constants were determined first in the absence of guest in order to be included and accounted for in the linked equilibria for host-guest complexation.

$$HH \overset{K_{dimer}}{\rightleftharpoons} H + G \overset{K_{assoc}}{\rightleftharpoons} HG$$

Figure 2.16. Competing equilibrium in the case of host dimerization

Dimerization constants ($K_{dimerization}$) for each host were determined from the concentration-dependent host chemical shift changes during a ^1H NMR dilution titration. The chemical shift data was fitted to a dimerization model using HypNMR to give the dimerization constant.[97] With the determined dimerization constant and the prepared host concentration from the start of the titration, the actual concentration of free host can be determined (as well as the true value of δ_{free}). The pre-determined values for the host

self-dimerization, therefore, were included in the HypNMR calculations as fixed numbers to isolate the strength of the host-guest interaction from the host-host interactions.

Experimentally, dilution titrations were done by preparing a 10 – 15 mM solution of each host in buffered D_2O to create a stock solution. The stock solution was then added sequentially to an NMR tube originally containing 0.5 mL of blank, buffered D_2O in a "reverse" dilution titration (reverse because the concentration actually increase over the experiment, rather than decreases). Additions continued until the NMR tube contained a total of 1.5 mL. The concentrations observed in these studies ranged from 15 to 0.2 mM. Representative experimental data and curve fitting for a single proton on host **2.2** is shown in Figure 2.17. The host chemical shift for this proton were shown to move up-field with increasing concentration.

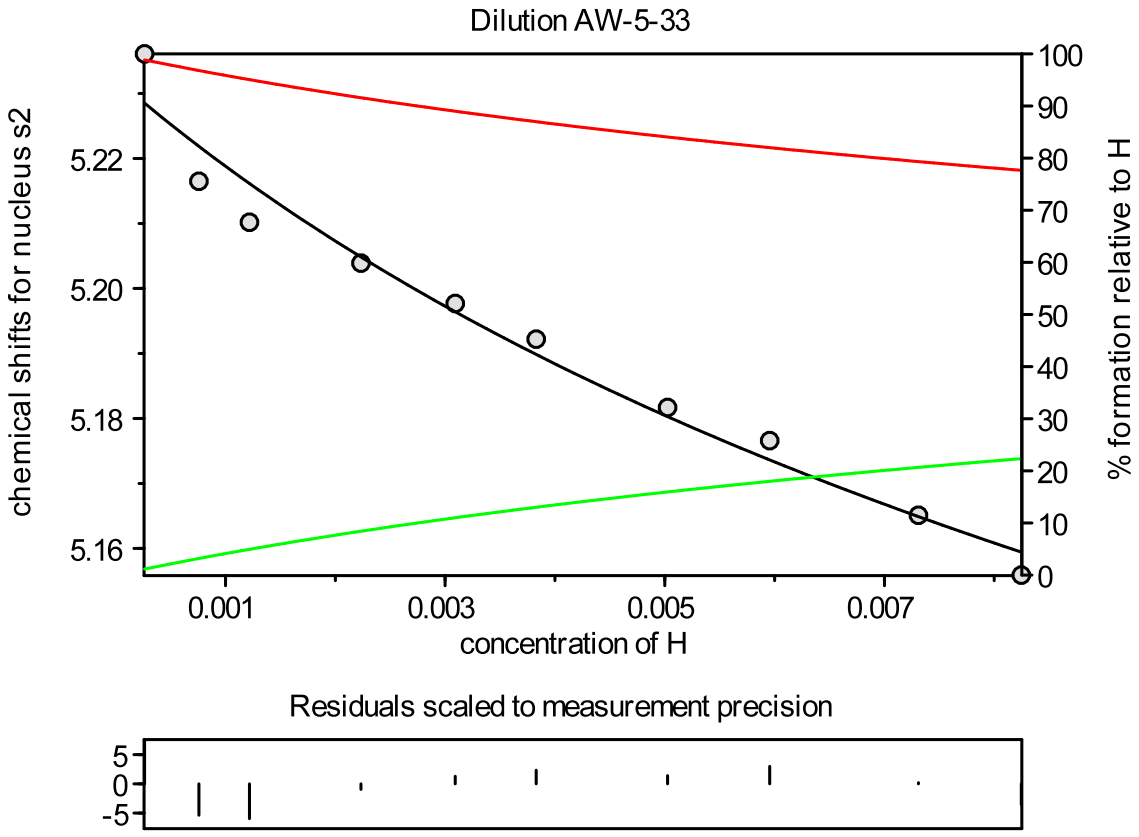
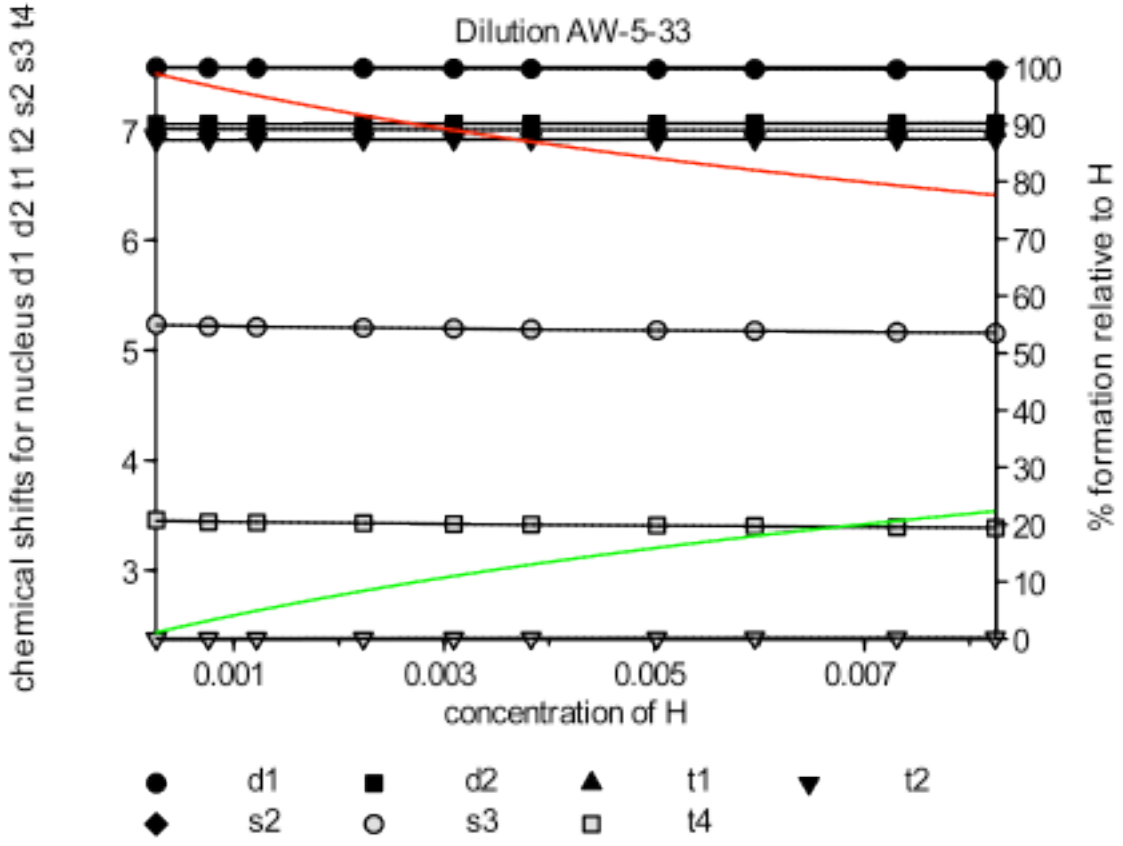


Figure 2.17. Dilution curve resulting (black line) for host **2.2**, signal s2 (singlet 2). Red and green lines represent concentration of free host and H2 dimer, respectively, relative to total host concentration (right axis). Residuals from best fit line shown in lower box.

To determine a dimerization constant, all host signals that show concentration dependent chemical shifts are taken into account. Complete representative experimental

data and curve fitting are presented in Figure 2.18. Non-linear curve fitting of the dilution data for each host in buffered D_2O solution was used to generate dimerization constants for each host using the HypNMR software package.[97] In the example shown in Figure 2.18, a dimerization constant (formation of H2) of 22 M^{-1} (log 1.3514) was determined for host **2.2** using the concentration-dependant chemical shifts of seven host proton signals. The upper solid line (red) represents the % of total host concentration that is present as free host. The lower solid line (green) represents the % of total host present has H2.



HypNMR2006
Refinement concluded at 4:40:36 PM on 10/27/2011
Data from D:\Personal_data\Amanda\1 Indole NMR Titrations with HypNMR\2 Methoxy tris indole\Dimerization AW-5-33\Dilution AW-5-33.hqd
Project title: Dilution AW-5-33

Converged in 1 iterations with sigma = 2.475903

	value	standard deviation	Comments
1 log beta(H2)	1.3514	0.1244	1.4(1)

Figure 2.18 Dilution curve results for host **2.2**. $K_{dimerization}$ determined using seven proton signals on host. (Reproduced by permission of The Royal Society of Chemistry).

Host-guest interactions were studied through concentration-dependent chemical shifts for each host upon the addition of guest as determined by ^{1}H NMR titrations. The association constants (K_{assoc} or K_{11}) of the host-guest interactions were determined by non-linear curve fitting for a 1:1 binding event using HypNMR2006 software. Host solutions (0.5 – 1 mM) were prepared fresh in buffered D_2O. A portion (0.5 mL) was removed to become the receiving solution in the NMR tube and a portion was used to make up guest solutions (2-250 mM) in order to ensure that the host concentration remained constant throughout the titration. The maximum concentration of guest solutions to be titrated was determined taking into account both guest solubility and buffer concentration. Guest species that could undergo a proton transfer (e.g. NMe_3H^+) were kept to concentrations under 50 mM in order to not overwhelm the buffer present. ^{1}H NMR binding titrations were carried out at 295 K on a Brüker Avance 360 spectrometer at 360 MHz. Concentration-dependant chemical shift data of multiple signals from each titration were fit to a 1:1 (H:G) binding isotherm using HypNMR. The final host- guest association constants were determined by an average of 2 – 4 titrations. The concentration dependent chemical shifts of one proton of host **2.2** upon titration with acetylcholine is shown below in Figure 2.19.

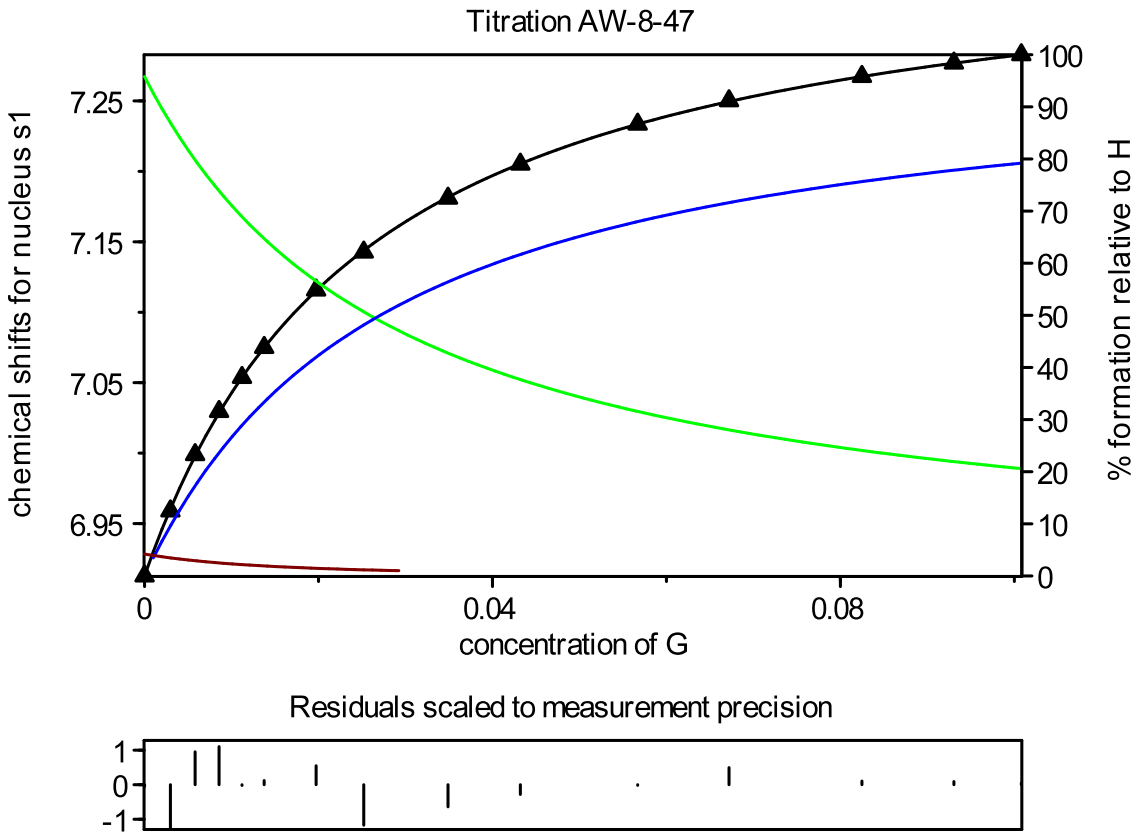
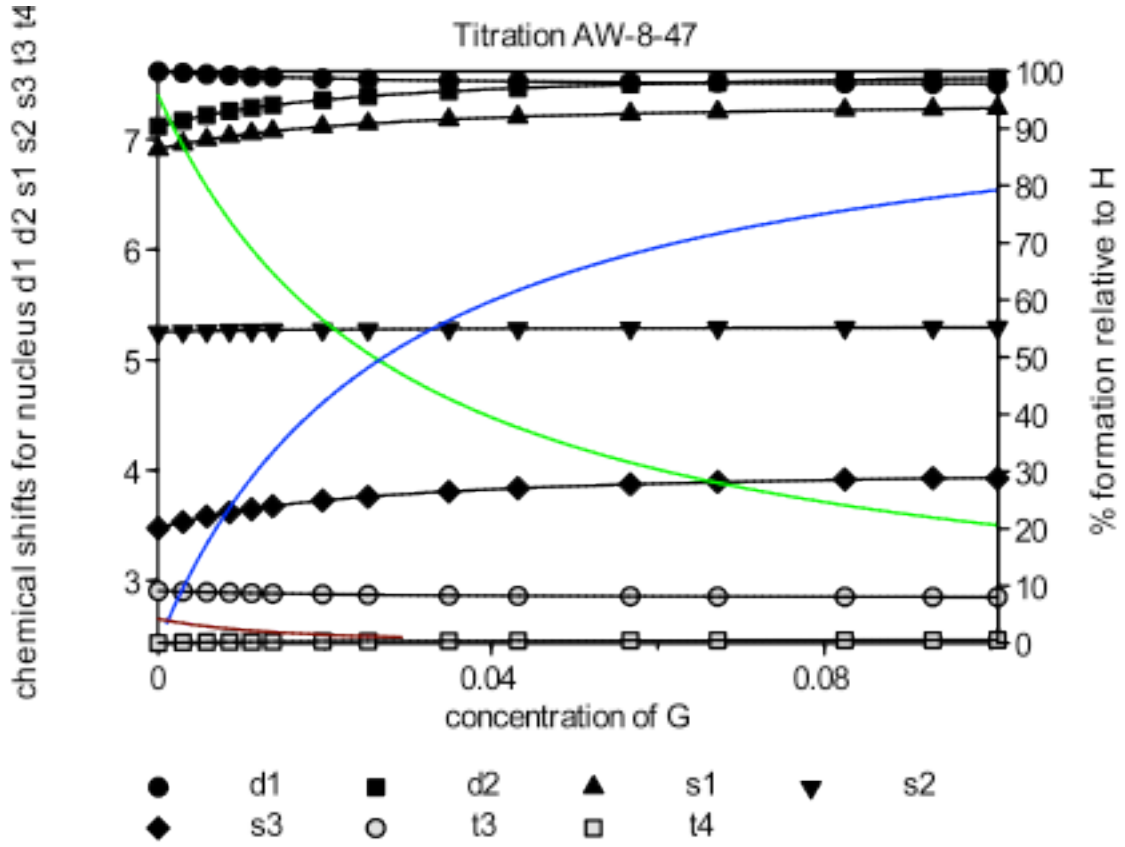


Figure 2.19. Titration curve result for a single proton on host **2.2** being titrated by AChCl. Red, green and blue lines represent the concentration of H2 dimer, free H, and HG complex, respectively, relative to total host present (right axis). Residuals from best fit line shown in lower box.

To correct for the known strength of the host dimerization determined earlier, the H2 equilibrium constant is included in the calculation for each guest. Representative fitted data for all the affected protons of host **2.2** is presented in Figure 2.20. In the example shown, an association constant of 38 M^{-1} (log 1.5849) for host **2.2** binding to acetylcholine chloride was determined using the concentration-dependant chemical shifts of seven host proton signals. The dimerization constant of 22 M^{-1} (log 1.3514) determined above was included as a fixed value in this calculation. The upper solid line (blue) represents the % of total host present as HG; one can see that this increases over the course of the titration, eventually reaching approximately 80% bound. The second solid line (green) again represents the % of total host present as free host; this decreases over the course of the titration. Notice that free host starts at approximately 95% of the total host before any guest is added. The remaining 5% is indicated by the third solid line (red) located to the bottom of the figure. This line represent the % of total

host present as H2. This quickly decrease as free host is consumed by binding to the guest and H2 feeds into the binding equilibrium.



HypNMR2008
Refinement concluded at 5:51:18 PM on 10/27/2011
Data from D:\Personal_data\Amanda\1 Indole NMR Titrations with HypNMR\2 Methoxy tris indole\AChCl\AW-8-47\Titration AW-8-47.hqd
Project title: Titration AW-8-47

Converged in 9 iterations with sigma = 1.389723

	value	standard deviation	Comments
1 log beta(GH)	1.5849	0.0743	1.58(7)
log beta(H2)	1.3514	fixed	

Figure 2.20 Titration curve results for host **2.2** being titrated by AChCl. Resulting K_{assoc} determined from seven proton signals on host that shifted during titration. (Reproduced by permission of The Royal Society of Chemistry).

Chapter 3 – Synthetic approaches to novel symmetric and dissymmetric Tröger's base molecules

3.1 Foreword

The hosts presented in Chapter 2 clearly demonstrated that the amount of hydrophobic surface area available is an important factor in the binding of the hydrophobic biological cations considered in these studies. However, even the strongest binding constant achieved in water by any of the tris-indoles hosts for any R-NMe_3^+ type guest (250 M^{-1} for **2.1** with Kme3) was significantly smaller than that of any native aromatic cage (for example, 400,000 M^{-1} (K_d 2.5 μM) for HP1 chromodomain with a H3K9me3 peptide).[89] Additionally, the three-fold symmetric structures used to position the aromatic moieties were ineffective at presenting them in a convergent, face-on manner as show by aromatic clustering in water to a collapsed state in the absence of guest. As introduced in Section 1.5, native aromatic cages for methylated lysines show very little movement between the bound and unbound states, suggesting that pre-organization of the hydrophobic surface also plays an important role. While native aromatic cages can achieve this pre-organized arrangement with the help of the surrounding protein structures, the synthetic hosts in Chapter 2 lacked this additional help or geometric preference.

Synthetic receptors for methylated lysines often incorporate macrocyclic structures to ensure the presentation of a well-defined, hydrophobic binding cavity and to prevent a detrimental collapse of that cavity in water (see Section 1.6). We hypothesized that improved binding constants could be obtained if the aromatic and hydrophobic surface area of the original hosts (indoles and other aromatic rings, for example) could be incorporated into a motif that a) had the correct spatial orientation (face-on) and b) was itself incapable of a hydrophobic collapse. Inspired by the ethanoanthracene building blocks used to create successful hydrophobic binding sites in both Dougherty's cyclophanes (**1.4**)[118] and Waters' macrocycle (**1.2**),[39] we sought to find a similar hydrophobic molecule that could be readily functionalized. To that end, we have explored the chemistry of Tröger's base.

3.2 Aims and Contributions

This chapter presents the synthesis and derivatization of a number of Tröger's base compounds and their intermediates. Specifically, the goals of this chapter are to show:

1. The synthesis of novel dissymmetric Tröger's base molecules using stepwise and direct approaches.
2. The incorporation of handles for further synthesis into Tröger's base molecules, specifically to facilitate the addition of water-solublizing functionalities or to allow incorporation into larger structures (via amide bond formation and/or cross coupling methodologies).

Synthesis of all symmetric Tröger's base molecules, as well as all nitro-iodo Tröger's base derivatives was performed by A. Whiting. This methodology was then followed by K. Dubicki for the synthesis of the nitro-ester Tröger's base, derivatization to amino-acid Tröger's base, and subsequent characterization. Derivatization of the nitro-ester Tröger's base and/or synthesis of the iodo-ester Tröger's base towards Suzuki coupling partners was completed by A. Whiting and K. Dubicki. Crystal structure data of **3.30** was collected by Dr. Lev Zakharov at the CAMCOR X-Ray Crystallography Facility at the University of Oregon. Characterization of all other molecules was done by A. Whiting.

3.3 Tröger's bases: rigid, aromatic, hydrophobic building blocks

The molecule we chose to pursue our rigid synthetic aromatic cages was the heterocyclic amine 2,8-dimethyl-6*H*,12*H*-5,11-methanodibenzo[*b,f*][1,5]diazocine, commonly known as Tröger's base, **3.1** (Figure 3.1). First synthesized by Julius Tröger in 1887 from *p*-toluidine and formaldehyde in acidic solution,[119] the actual structure was not completely elucidated until 1935 when Speilman established the correct connectivity.[120] The true structure of Tröger's base presents two aromatic rings at nearly perpendicular angles to each other connected by a bridged eight-membered ring containing two tertiary amines. A bridging CH_2 between the two nitrogen atoms fixes the orientation of the amines and prevents the normal rapid inversion, creating two stereogenic nitrogen centers. Resolution of Tröger's base into its pure enantiomeric forms was first accomplished by Prelog and Wieland in 1944[121] using liquid chromatography with a chiral D-lactose column, a relatively new technique. In fact, Tröger's base was the first chiral tertiary amine devoid of non-nitrogen stereogenic

centers to be resolved. The unique chirality of Tröger's base has since been exploited in the field of chromatography and frequently appears as a standard racemic mixture when developing a chiral separation method.[122-124]

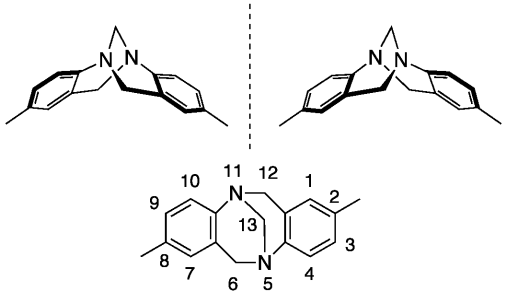


Figure 3.1 The enantiomers of Tröger's base **3.1**: the (5*R*,11*R*)-enantiomer and the (5*S*,11*S*)-enantiomer

3.3.1 Tröger's bases in molecular recognition

The unique properties of Tröger's bases have made them useful for applications in the world of supramolecular chemistry and molecular recognition. One of the most common uses for Tröger's bases in a supramolecular system is as a rigid spacer to position other functional groups at specific orientations. In this capacity, a number of Tröger's bases with specifically positioned hydrogen bonding units have been made including Wilcox's diacid **3.2** for binding to biotin and adenine derivatives,[125] and Kobayashi's thiophene derivative **3.3** bearing two pyridylamino groups for dicarboxylic acids.[126] In these cases, binding was studied in organic solvents (THF-d_8 and $CDCl_3$), as studies in water (or any polar solvent) would significantly weaken the strength of the hydrogen bonds.

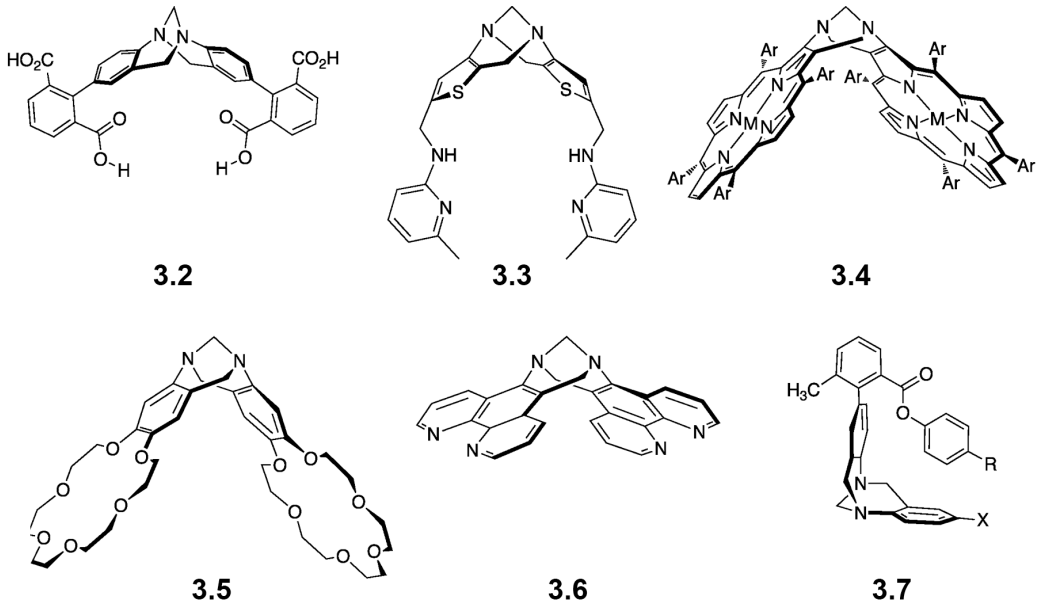


Figure 3.2 Tröger's bases in supramolecular structures

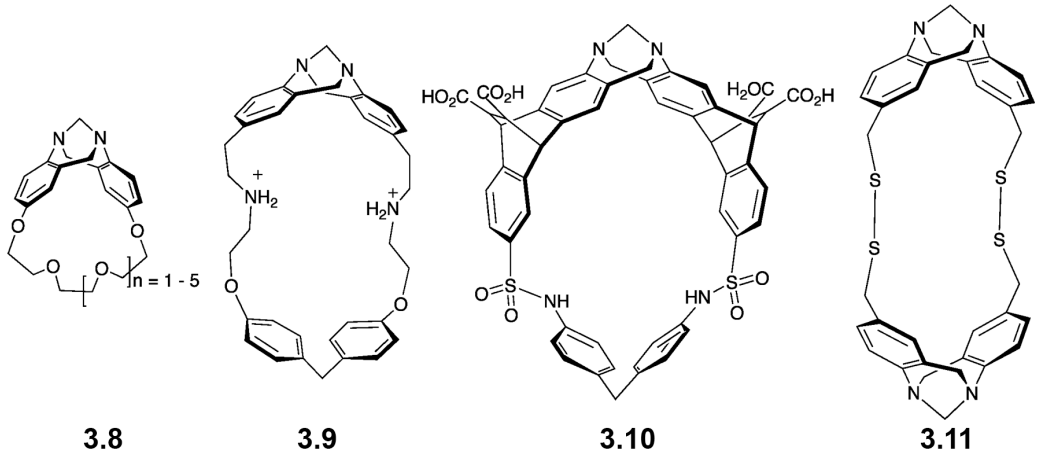
Other binding motifs have been held apart by Tröger's bases. Crossley and co-workers synthesized a porphyrin Tröger's base derivative (**3.4**) containing two metal binding sites occupied by zinc.[127] They have also bound palladium, cobalt, and copper,[128] as well as used for binding molecules divalently (such as both amines of free lysine) in toluene-d_8 and $CDCl_3$.[129] In similar fashion, bis(crown-ether) Tröger's base analogues (**3.5**) have been synthesized by Wärnmark and co-workers,[130] and used to divalently bind primary bisammonium salts from 1:1 mixtures of MeOD-d_4 and $CDCl_3$.

Tröger's base compounds have also been synthesized that contain various heterocycles, such as phenanthroline (**3.6**)[131] and similar acridine[132] "arms". One main interest in these complexes is their ability to intercalate and bind to DNA.[131,133] The ability of phenanthroline to act as a bidentate metal binding ligand has also led to its use in the preparation of a bimetallic ruthenium(II) complex.[134,135]

Tröger's base derivatives have been exploited in Wilcox's molecular torsion balance (**3.7**) as a tool to measure edge-to-face aromatic[136] and CH-pi interactions[137] depending on the substituents of the coupled aromatic ring. In this case, the Tröger's base aromatic ring is an integral part of the entire structure, unlike the other examples in this section.

More relevant to this work, Tröger's bases have been incorporated into a variety of macrocyclic structures. In these structures, the Tröger's base itself plays a role in binding rather than just in orienting other binding elements. Some of the simplest

examples are Tröger's base-crown ether molecules. Ibrahim *et al.*[138] and Manjula and Nagarajan[139] both published syntheses of Tröger's base crown ether composites of the type **3.8**. These scaffolds were intended for cation binding and were found to bind alkali metal and ammonium cations with high affinities (K_a on the order of 10^5 M^{-1}) using Cram's picrate extraction method[140] ($CHCl_3/H_2O$ water mixtures) but with little selectivity. Larger Tröger's base cyclophanes have also been synthesized and studied under aqueous conditions. Wilcox again made the first cyclophane **3.9**[141] that included two secondary ammonium groups. Under acidic conditions, this cyclophane was water soluble and found to bind simple substituted aromatic rings (phenols) with association constants ranging from 40 to 330 M^{-1}. A second generation Wilcox cyclophane **3.10** included additional rigid aromatic surface area in the form of ethanoanthracene units.[142] **3.10** was studied under basic aqueous conditions (pH 9) to deprotonate the tetracarboxylic acids for water solubility. Binding constants of up to 2000 M^{-1} were obtained for small aliphatic molecules such as menthol and isomenthol. Other, non-water soluble Tröger's base cyclophanes have been synthesized such as the thio-linked version **3.11** by Bag and von Kiedrowski.[143]



3.8 **3.9** **3.10** **3.11**

Figure 3.3 Tröger's bases as macrocyclic, supramolecular hosts

Other supramolecular architectures have been synthesized involving multiple or repeat Tröger's bases units. Because of the chiral nature of Tröger's bases, the combination of multiple Tröger's bases leads to the formation of diastereomers with subsequent unique geometries. Linear combinations of two Tröger's bases give rise to two bis-Tröger's base structures: the cleft-like *syn*-isomer **3.12** (boat) and its *anti*-isomer **3.13** (chair).[144] For steric reasons, the *anti*-isomer is more thermodynamically stable and often isolated in higher yield. Acidic conditions have been used to isolate *syn*-derivatives

through racemization of the higher yielding *trans*-derivatives.[121,145,146] Larger Tröger's base oligomers containing three repeating units (**3.14**) have also been synthesized, containing both symmetric[147,148] and dissymmetric substitutions.[145]

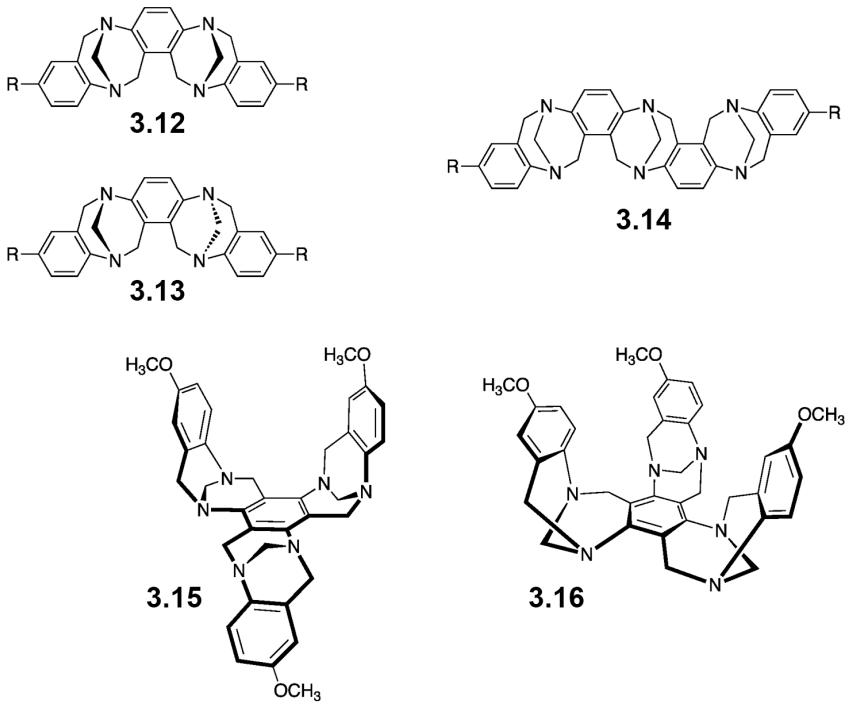


Figure 3.4 Larger, multiple Tröger's base superstructures

Finally, three-fold symmetric Tröger's base molecules (throne-**3.15** and calix-**3.16**) have been synthesized starting from 1,3,5-triaminobenzene. The throne-version (**3.15**) was isolated in 3% overall yield (with 6% yield during the final step forming twelve new bonds). Isomerization under acidic conditions was used to generate calix-**3.16** in 3% isolated yield..[146]

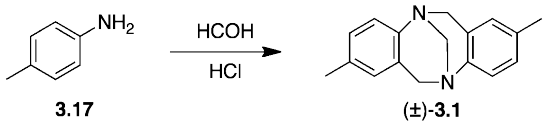
3.4 Synthetic approaches to Tröger's bases

In order to make use of Tröger's bases, one needs to be able to make them. While the original Tröger's base was a symmetric molecule (i.e. both ends were identical) made using *para*-substituted anilines, the examples above show that chemists have been successful at making a wide variety incorporating substituted aromatic rings,[130,142] heteroatoms,[126,127] and desymmetrized substitutions.[136,137]

3.4.1 Symmetric Tröger's bases

The path to symmetric Tröger's bases is straightforward and mimics the original method discovered by Tröger. The original reaction using *p*-toluidine **3.17** and

formaldehyde in the presence of HCl formed a racemic mixture of **3.1** in one step (Scheme 3.1). Since then, general Tröger's base forming conditions have expanded to include a number of different formaldehyde equivalents such as paraformaldehyde, hexamethylenetetraamine (HMT),[142] and dimethoxymethane.[149] More unusual conditions utilizing DMSO and HCl(g) as a formaldehyde equivalent have also been published.[150]



3.17 (±)-**3.1**

Scheme 3.1 One-step, symmetric Tröger's base synthesis

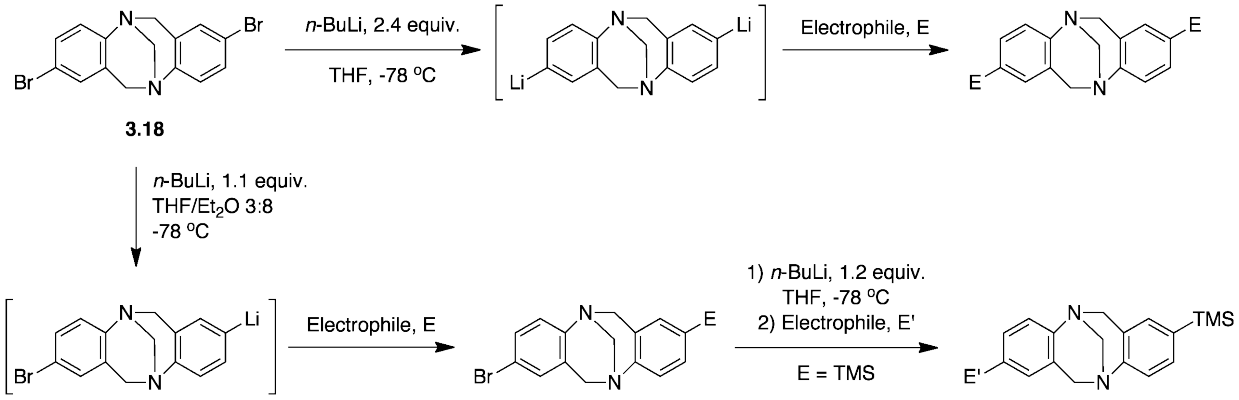
The mechanism of Tröger's base formation is believed to be electrophilic substitution, catalyzed by strongly acidic conditions. The finished product contains a total of two new C-C bonds and four new C-N bonds. With a number of reactive electrophilic centers present in the starting amines, undesired polymerization products are possible. In order to prevent this, the *para* position of the anilines are often blocked through substitution, leaving only the activated *ortho* position available for reaction. *Para*-substitution has been shown not to be absolutely necessary for Tröger's base formation, but drastically reduced yields are noted when they are absent (likely due to the formation of unproductive by-products).[151]

Although it has been reported in the literature that this mechanism does not work for deactivated, electron-deficient anilines, there is evidence to the contrary. *p*-Halo,[152-154] *p*-ester[155] and *p*-nitro substituted anilines[151] have been used with success to form Tröger's base analogs.

3.4.2 Desymmetrized Tröger's bases

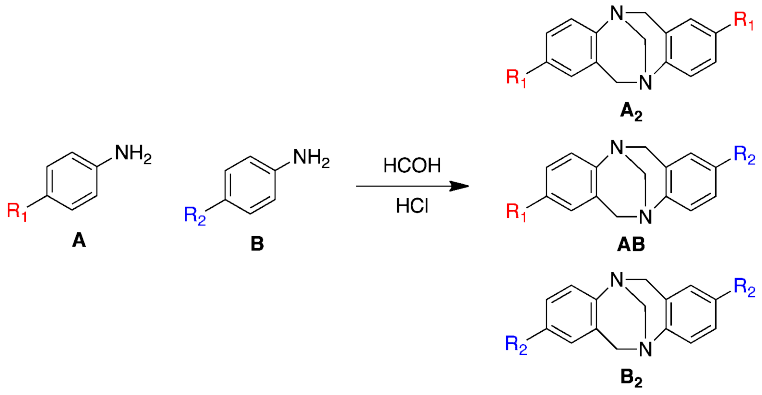
The synthesis of desymmetrized Tröger's bases presents more of a challenge. Desymmetrized Tröger's bases have been synthesized by taking symmetric Tröger's bases and reacting one side only with strict stoichiometric control, through halo-lithium exchange and cross-coupling reactions. Jensen and co-workers demonstrated that 2,8-dibromo-Tröger's bases **3.18** could be desymmetrized by single bromine-lithium exchange followed by reaction with select electrophiles (Scheme 3.2) with yields ranging from 60 to 80%.[156] A second bromine-lithium exchange reaction (where E = trimethylsilyl, TMS) introduced a second electrophile to the opposite end of the molecule. Similarly, Artacho, *et al*.[157] and Hof, *et al*.[158] demonstrated that cross-coupling

methodology could be used to generate desymmetrized Tröger's bases containing amino- and carbon-based functional groups from 2,8-diiodo-Tröger's base.



Scheme 3.2 Desymmetrization of 2,8-dibromo Tröger's base **3.18** via halo-lithium exchange

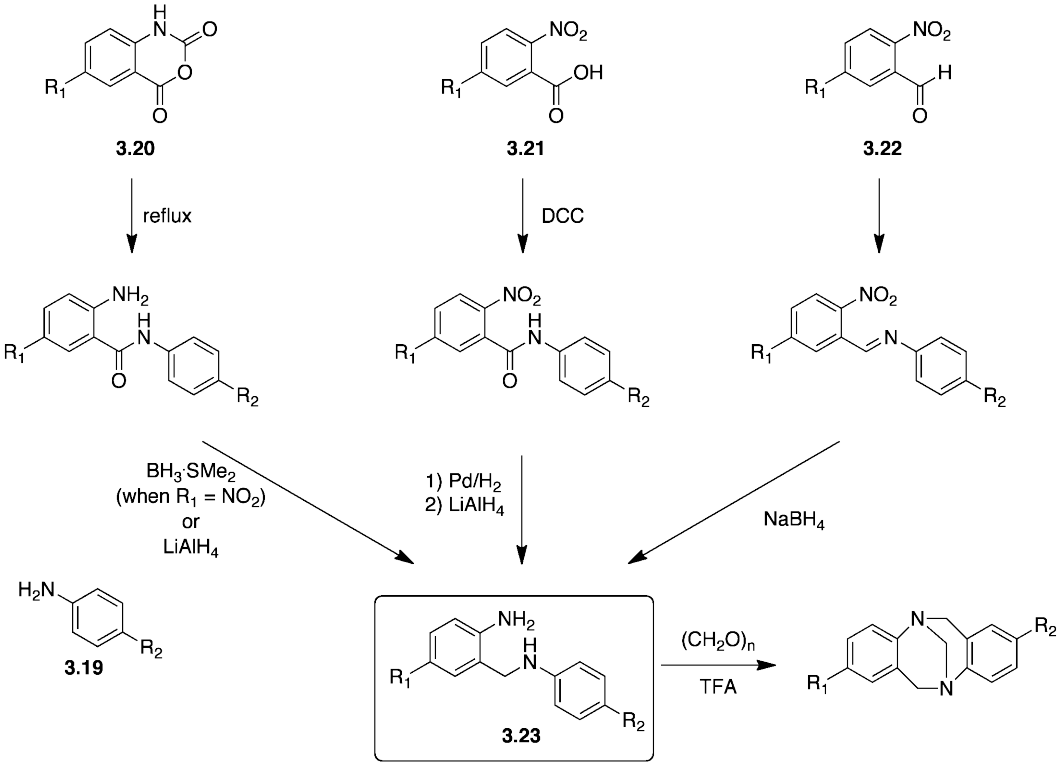
Desymmetrized Tröger's bases have also been synthesized using a 1:1 ratio of substituted amines,[159] generating a statistical mixture of A_2, AB (desired), and B_2 products that must be separated (Scheme 3.3). Assuming equal reactivity of each amine, this leads to a theoretical yield of 25% A_2, 50% AB and 25% B_2. The separation of these molecules can be problematic, as the desired product is found in between the polarities of the other two symmetric molecules. If the R_f values for A_2 and B_2 are not that different, separation can be difficult. Additionally, this approach is ineffective if one of the reacting partners is electron-poor and therefore less reactive than the other partner (such as the case for a nitro-substituted aniline).



Scheme 3.3 One-step synthesis of desymmetrized Tröger's bases

An alternative, stepwise synthetic method has also been used as a route to desymmetrized Tröger's bases (Scheme 3.4).[160] In this case, two different anilines are selectively coupled together by a simple amide bond before elaboration into a Tröger's

base ring system. Three examples are shown in the scheme below. The first involves reaction of an isatoic anhydride (**3.20**) with an aryl amine (**3.19**) to generate an amide bond, the second uses peptide coupling conditions to connect the amine with an anthraniclic acid derivative (**3.21**), while the third forms an imine from the amine and aryl aldehyde (**3.22**). The goal of each coupling is to obtain the amine-linked structure **3.23**, an intermediate of the Tröger's base-forming mechanism. From there, Tröger's base forming conditions are used to generate the desired final structure. This method is especially attractive when one wishes to include electron-withdrawing substituents (at R_1) as it allows the key carbon bond on the deactivated ring to be formed prior to Tröger's base formation.

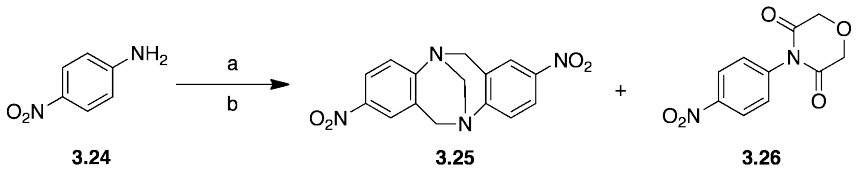


Scheme 3.4 Stepwise synthesis of desymmetrized Tröger's bases

3.5 Synthesis of symmetric 2,8-dinitro Tröger's bases

To begin our investigation into the utility of using Tröger's base molecules as building blocks for synthetic aromatic cages, we wanted to synthesize a simple, symmetric Tröger's base that would allow for later derivatization. To that end, we envisioned making a symmetric, dinitro Tröger's base with the intent of using the nitro groups as synthetic handles. As we (and others) have found, making nitro Tröger's bases is not trivial due to the electron-withdrawing nature of the -NO_2 group. It is difficult

to form the necessary aryl-methylene bond on a ring bearing a nitro group; the acid-catalyzed mechanism of formation requires an electron-rich aromatic ring to be a nucleophile which is hindered in this case preventing a direct reaction of 4-nitroanline **3.24** with paraformaldehyde.[151] Consequently, alternative conditions have been discovered that get around this intrinsic problem. We found a paper which claimed to be the first to cleanly make symmetric 2,8-dinitro Tröger's bases **3.25** from **3.24**.[161] The authors had reacted **3.24** with diglycolic acid (carbon source) and polyphosphoric acid (acidic conditions) and found that (by accident) they could only isolate **3.25**. Repetition of this procedure in our hands (Scheme 3.5, condition a) was successful in making **3.25**, as verified by ^{1}H NMR. However, cyclic imide **3.26** (4-(4-nitrophenyl)morpholine-3,5-dione), the original intended product of this procedure, was also obtained (as it co-eluted with **3.25**).[161] Similar results as ours were subsequently obtained and published by another group exploring functionalized di-nitro Tröger's bases synthesis.[151] Although they were unable to separate **3.25** and **3.26** by column chromatography, we isolated a fraction of **3.26** for characterization and future reference.

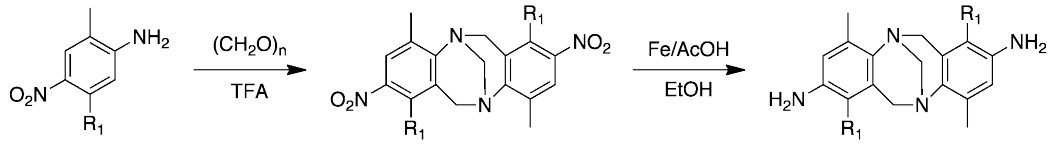


Scheme 3.5 Synthesis of 2,8-dinitro Tröger's base **3.25**. Conditions used a) diglycolic acid, polyphosphoric acid, 80^{o}C, 24h; b) DMSO and $HCl_{(g)}$.

A second method of making Tröger's base **3.25**, was found from a later procedure,[150] this time utilizing DMSO as the carbon source and HCl(g) and acetic acid as the acidic conditions (Scheme 3.5, conditions b). This successfully produced **3.25** directly in 17% yield following column chromatography. This result was comparable to the published 25% yield.

While the above efforts had shown that it was possible to make a symmetric di-nitro Tröger's base, the poor yields produced above did not encourage further efforts. We remained interested in using nitro functional groups as handles and so looked for alternative methods to incorporate them. An expanded literature search on Tröger's bases led us to a paper by Kiehne, et al.[162] These authors successfully synthesized di-nitro Tröger's base under standard conditions using *para*-nitroanilines when additional methyl substituents were present on the aromatic ring. The enhanced yields were thought to be due to increased solubility of the starting material in organic solvents.[151]

Encouraged, we synthesized the symmetric di-nitro Tröger's bases **3.29** and **3.30** from methyl-substituted anilines **3.27** and **3.28**, respectively, in high yield. The *ortho*-methyl substitution in both cases restricted the formation of more than one Tröger's base isomer by allowing only one site for reactivity. Single crystals of dinitro-Tröger's base **3.30** were also successfully grown from acetone by slow evaporation. The resulting crystal structure is shown in Figure 3.5 (ORTEP with 50% probability ellipsoids). Subsequent reduction with iron and acetic acid[162] gave the diamino-Tröger's bases **3.31** and **3.32**. These molecules were then readily set up for further modification using one of our connective methods, amide bond formation.



3.27 R_1 = H
3.28 R_1 = CH_3

3.29 R_1 = H (83%)
3.30 R_1 = CH_3 (76%)

3.31 R_1 = H (96%)
3.32 R_1 = CH_3 (98%)

Scheme 3.6 Synthesis of symmetric diamino-Tröger's bases, **3.31** and **3.32**

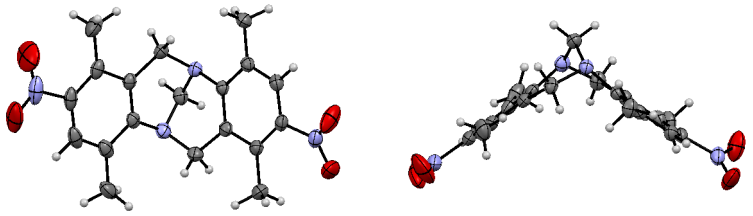
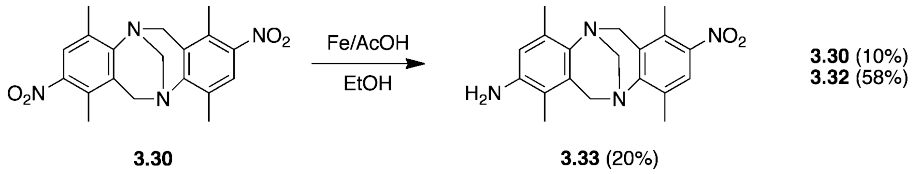


Figure 3.5 An ORTEP diagram of **3.30** with 50% probability ellipsoids.

3.6 Desymmetrization of a 2,8-dinitro Tröger's base

With the goal of incorporating Tröger's bases into more complicated structures, we were interested in producing a desymmetrized molecule that could be orthogonally coupled to other structures. One way that we envisioned doing this was through the desymmetrization of an already synthesized, symmetric Tröger's' base (as described in section 3.4.2). During our efforts to reduced dinitro Tröger's' base **3.30** to diamino Tröger's' base **3.32** (Scheme 3.6), we observed the mixture of products shown in Scheme 3.7, including the partially reduced Tröger's' base **3.33** in 20% yield. Though low yielding, this presented itself as a straightforward and convenient entry into desymmetrized amino Tröger's bases. Unfortunately, despite exact replication the previously successful conditions and attempts to vary the iron/acetic acid concentrations and/or reaction time, we were unable to generate mono-reduced **3.33** in any appreciably

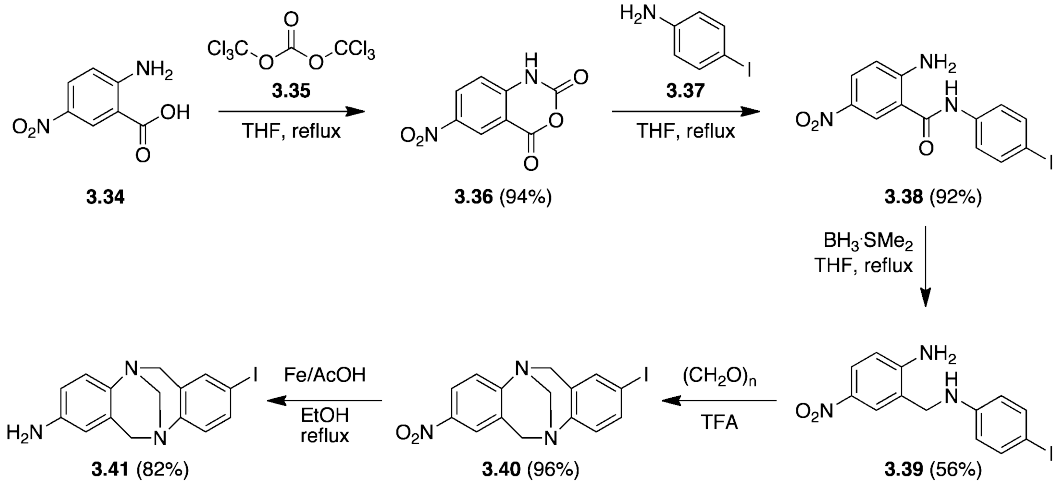
yield again. Partial hydrogenation as a route to desymmetrized Tröger's bases was thus abandoned.



Scheme 3.7 Partial reduction of **3.30** to produce nitro-amino-Tröger's base **3.33**

3.7 Stepwise synthesis of an amino-iodo Tröger's base

As a second route to monoamino-Tröger's bases, we chose to follow the step-wise route proposed by Wilcox[160] utilizing isatoic anhydrides. This method is especially useful when one of the two reacting amines is electron-deficient as it establishes the methylene link to the deficient ring prior to reaction under Tröger's base conditions. The successful route used is shown below.



Scheme 3.8 Stepwise synthesis of amino-iodo Tröger's' base **3.41**

Starting from 5-nitroanthranilic acid **3.34**, we reacted it with an excess of triphosgene **3.35** to produce 5-nitroisatoic anhydride **3.36** in 94% yield on gram scale. From there, the anhydride ring was opened using iodo-aniline **3.37** generating loss of CO_2 to drive the reaction and ensuring that the reaction only formed the desired amide. Heating at reflux in dry THF gave **3.38** in 92% yield. Again, isolation of this product was through filtration as the product was found to be insoluble in DCM.

The difficult step in this synthesis was the reduction of the amide bond to generate the desired amine-linked intermediate **3.39**. Due to the presence of the nitro group, strong reducing agents like lithium aluminum hydride (LAH) or more forcing

conditions such as catalytic hydrogenation (H_2/Pd under high pressure) were unsuitable. It was critical to prevent reduction of the nitro along with the amide, as multiple aryl amine sites would lead to product mixtures upon Tröger's base formation. We used borane as our selective reducing agent. Borane dimethyl sulfide was found to work better than borane-THF complex, as has been noted by others.[144,163] Despite multiple equivalents of borane per amide and high reaction temperatures, this reaction was consistently incomplete and low yielding. Product isolation via column chromatography provided both product and unreacted starting material. Optimized conditions led to a 56% yield of the desired product **3.39**.

Having reached the necessary amine-linked intermediate, the compound was subjected to Tröger's base forming conditions of paraformaldehyde in neat trifluoroacetic acid (TFA). Upon work up with base to neutralize the TFA present, the compound precipitated and was isolated by filtration. Heating at reflux in acetone, followed by cooling and isolation by filtration led to the desired compound **3.40** in 96% yield. Including the formation of the commercial isatoic anhydride this scheme gave **3.40** in 46% overall yield over four steps.

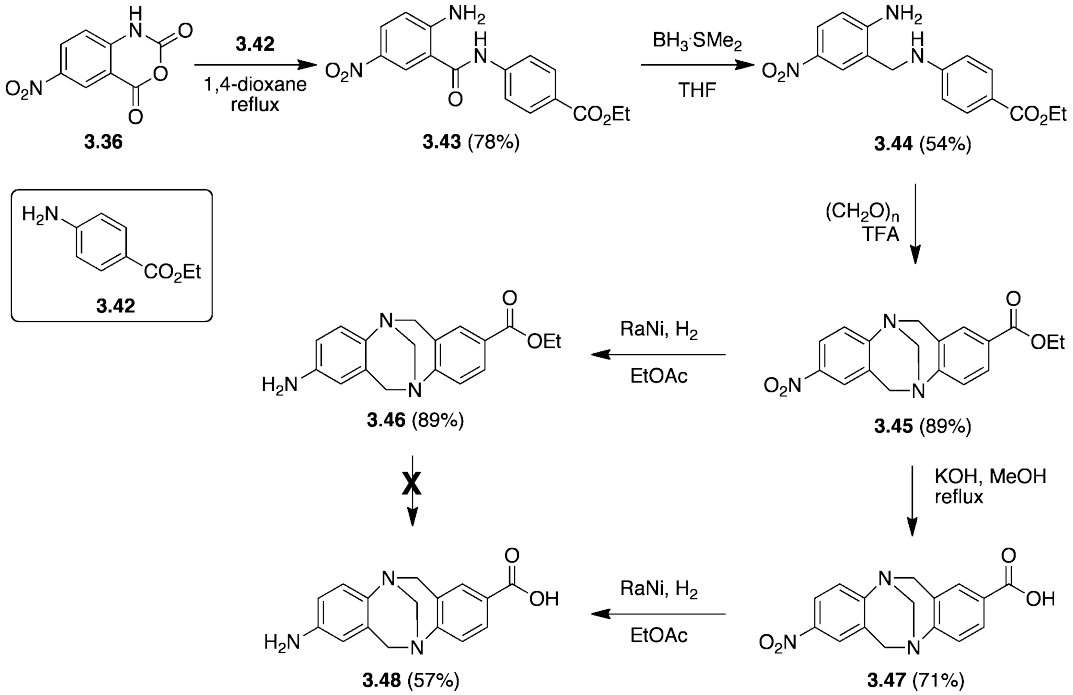
As a potential site of further attachment, the nitro group was included as an inert functional handle pending its reduction to amine. As a step to functionalizing this end of the molecule, we reduced the nitro group using iron and acetic acid as had been shown to be successful on other Tröger's base molecules.[162] While this was effective at producing the desired molecule **3.41** in 82%, the iron oxide (rust) by-product was frustrating to deal with and extractions were not completely straightforward. However, we were unable to use H_2/Pd reductions due to the undesired removal of the iodo group during hydrogenation. The Fe/AcOH combination was able to avoid this complication.

3.8 Stepwise synthesis of an amino-acid Tröger's base

The success of the methodology above in making amino-iodo Tröger's base **3.41** led us to consider making another potentially useful desymmetrized Tröger's base building block. An amino-acid Tröger's base could be incorporated directly into a peptide though amide bonds, thereby combining an aromatic cage-like motif with the possibilities of solid phase peptide synthesis. The 90^o constrained geometry of Tröger's bases lend themselves nicely to use as synthetic peptide hinges (similar to hairpin turns), or as hydrophobic aromatic binding elements themselves. A similar idea using symmetric α-amino acid adducts to Tröger's bases was previously published by Bew *et al*.[164]

However, their amine and acid functional groups are on long tethers connected to the Tröger's base, giving the overall molecule many degrees of freedom within the context of a peptide. Our molecules, as depicted in Scheme 3.9, would incorporate the amino acid handles directly on to the Tröger's base, in a much more conformationally rigid motif.

The synthesis of this amino-acid Tröger's base (Scheme 3.9) initially followed the method described in Section 3.7 for the iodo-nitro Tröger's base **3.40**. Starting again from 5-nitroisatoic anhydride **3.36**, the ring was opened using ethyl 4-aminobenzoate **3.42** to produce the amide-linked nitro-ethyl ester compound **3.43**. Through optimization of the reaction conditions, 1,4-dioxane was found to be a better solvent for this substrate than THF (as was used previously) and gave **3.43** in 78% yield.[165] As before, the amide proved amenable to isolation on gram scale by precipitation.



Scheme 3.9 Stepwise synthesis of amino-acid Tröger's base **3.48**

Again, difficulty in this synthetic pathway was encountered during the amide reduction, which in this case was in the presence of both a nitro group and an ethyl ester. $BH_3{\cdot}SMe_2$ was used as before to selectively target the amide carbonyl over both the nitro group and the ester, but with poor yields. As a precaution, we chose not to heat this reaction to reflux so as to avoid pushing the borane to reduce the ester as well. Optimization of the reduction was attempted by varying the number of equivalents, type of borane complex and overall reagent concentration. Scale was also not found to be

favourable with this reaction, with better yields being produced on smaller amounts (1 g reactions or less). The amine-linked structure **3.44** was eventually isolated in 54% optimized yield following column chromatography.

Once the amine-linked intermediate **3.44** was attained however, it was straight forward to subject this molecule to our standard Tröger's base forming conditions (TFA and paraformaldehyde). Nitro-ester Tröger's base **3.45** was formed in 89% yield and was isolated via column chromatography, as the product did not readily precipitate.

From nitro-ester **3.45**, one could imagine pursuing the amino-acid derivative through one of two ways – either by reduction of the amine followed by ester hydrolysis or vice versa. We pursued the amine reduction first and found it to proceed best using Raney nickel and hydrogen.[166] While iron and acetic acid was also successful at this reduction (49% yield), the ease of work up and a higher yielding reaction made the hydrogen method more attractive. The Raney nickel hydrogenation was successful in producing amino-ester Tröger's base **3.46** in 89% yield with minimal workup needed. Unfortunately, we found that hydrolysis of the ester at this stage was not productive as we were unable to isolate the resulting zwitterionic molecule from the aqueous acidic work up.

Alternatively, we reversed order of operations to first generate the nitro-acid Tröger's base **3.47**. We pursued the hydrolysis under the same conditions as attempted before and cleanly generate **3.47** in 71% yield. Acidification upon work up in this case produced a precipitate in MeOH that was isolated by filtration. Hydrogenation of **3.47** with Raney nickel and H_2 in EtOAc allowed us to avoid an aqueous work up and gave amino acid Tröger's base **3.48** in 57% isolated yield.

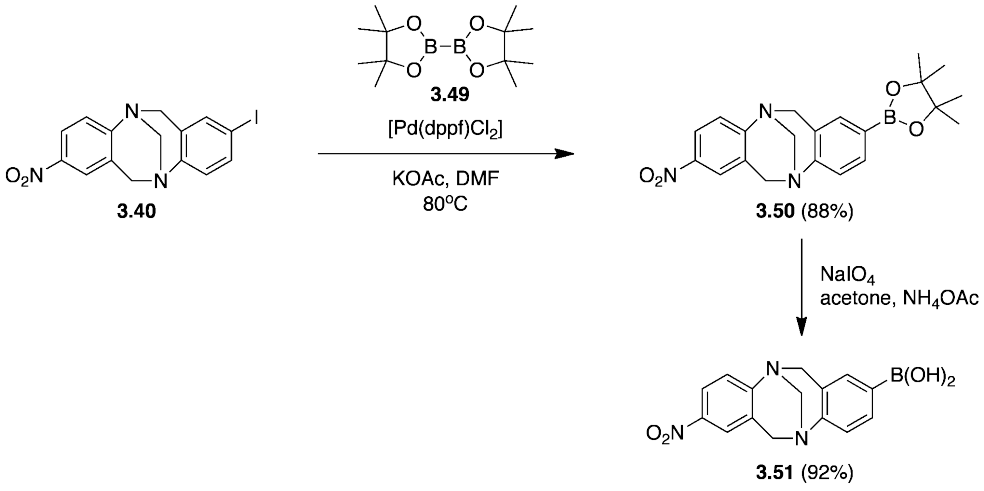
This amino-acid Tröger's base will require some further simple modifications in order to make it amenable to solid-phase peptide synthesis protocols. The amino functional group will need to be protected using either *t*-butyloxycarbonyl (Boc) or 9-fluorenylmethyloxycarbonyl (Fmoc) protecting groups. In addition, it would be preferable to have enantiomerically pure Tröger's base before incorporation into a peptide. Recent examples of chiral resolutions of Tröger's bases include using readily available chiral chromatography media (the Whelk O1 chiral stationary phase),[167] or chiral acids (such as dibenzoyl-L-tartaric acid) for diastereoselective co-crystallizations.[168]

3.9 Synthetic derivatization for use in palladium-catalyzed cross-couplings

While the molecules presented above could readily connect to other molecules using amide bonds (through the amino and/or acid functionalities present), we also wanted to incorporate handles for cross-coupling reactions. To this end, we endeavored to make Tröger's bases suitable for Pd-catalyzed Suzuki couplings by incorporating boronic ester and acid functional groups.

*3.9.1 Synthesis of nitro-boronic acid Tröger's base **3.51***

The iodo group of **3.40** already has the potential to be the aryl-halide partner in a Suzuki carbon-carbon bond forming reaction. Conversion of this iodo substitution to a boronic acid equivalent would also allow it to function as the organoboron partner of the same reaction. Fortuitously, there was already synthetic precedent for the synthesis of boronic ester Tröger's bases from iodo-Tröger's bases.[158] In our hands, reaction of the iodo group of **3.40** with bis(pinacolato)diboron (B_2pin_2) **3.49** and a palladium catalyst ([1,1'-bis(diphenylphosphino)ferrocene]dichloropalladium [$Pd(dppf)Cl_2$]), led to boronic ester **3.50** in 88% yield following column chromatography. The ester was deprotected to form the boronic acid **3.51** upon treatment with sodium periodate in acetone and aqueous ammonium acetate in 92% yield (isolation by extraction).[169] Both **3.40** and **3.50**/**3.51** would be possible partners in Suzuki-type coupling reactions, leading to products which still included a nitro group for further derivatization (see Chapter 4).

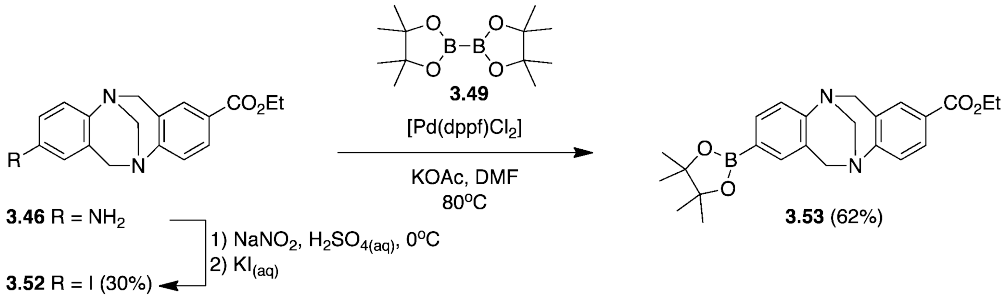


Scheme 3.10 Synthesis of nitro-boronic acid **3.51**

*3.9.2 Synthesis of ester-boronic ester Tröger's base **3.53***

We were also interested in making Suzuki reagent Tröger's bases that would included other functionality at the non-coupling end. **3.50** and **3.41** (above) retain a nitro

group with which one could further process to make amide bonds. We envisioned a similar Tröger's base but substituted with an acid group. The simplest route to this target, starting from molecules in hand, was through diazonium-intermediate Sandmeyer chemistry.[170] Amino-ester Tröger's base **3.46** was converted into iodo-ester Tröger's base **3.52** by diazotization with sodium nitrite followed by displacement of the diazonium species with potassium iodide (Scheme 3.11).[166] While successful, this reaction was low yielding at 30%. Conversion of iodo-ester **3.52** into boronated analogue **3.53** was done as before using B_2pin_2 and resulted in a 62% isolated yield. With **3.52** and **3.53**, we now had Tröger's base reagents with a protected acid group, which could participate in Suzuki reactions as either the aryl halide (**3.52**) or organoboron (**3.53**) partner.

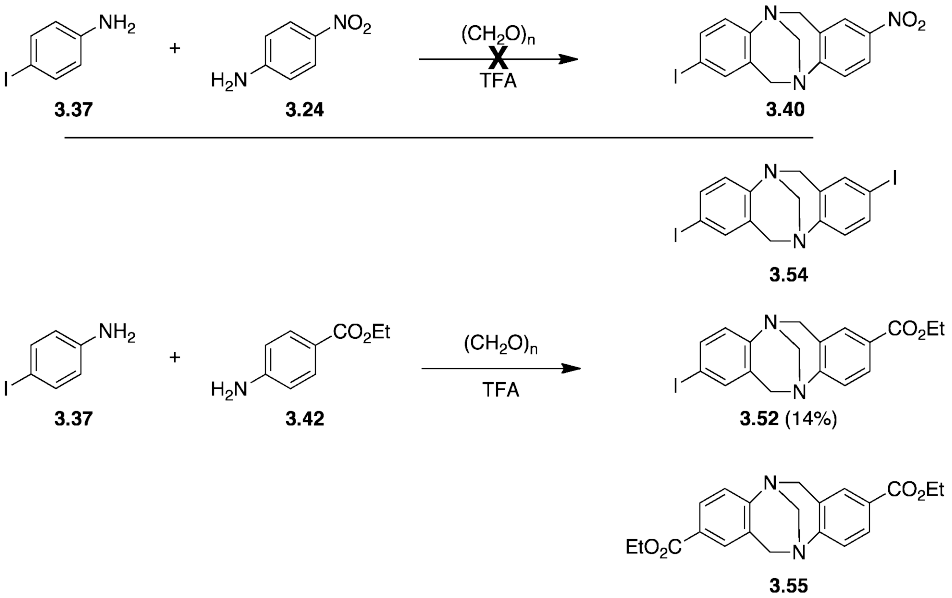


Scheme 3.11 Synthesis of ester-boronic ester Tröger's base **3.53**

While successful, the synthesis presented in Scheme 3.11 had drawbacks that would prevent its general use on larger scale. Specifically, we were concerned with the low yielding Sandmeyer step used to generate **3.52**, as well as the number of steps (and purifications) required to produce amino-ester Tröger's base **3.46** in the first place. Overall, the boronated Tröger's base **3.53** was produced in 6.2% yield over six steps starting from commercially available 5-nitroisatoic anhydride (**3.36**). Any method that could produce desymmetrized iodo-ester Tröger's base **3.52** in fewer steps and/or higher yield, would make these molecules more synthetically accessible and therefore, useful.

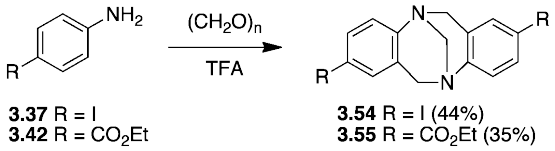
While less elegant, 1:1 mixtures of different *p*-substituted anilines have been shown to produce moderate yields of desymmetrized Tröger's bases after purification of the resulting mixtures of homo- and hetero-condensed products (Section 3.4.2). A reaction of this type was not open to us on earlier targets such as **3.40**, as one of the reacting anilines was *p*-nitro substituted and therefore too unreactive to form the desymmetrized products (Scheme 3.12, top). We decided to attempt this type of reaction for the iodo-ester Tröger's base **3.52**. This follows the precedent that 2-iodo-8-R hetero-

condensed Tröger's bases where R = -H, -CH_3, and -OCH_3 have been synthesized via this method in 15%, 31% and 46% yields, respectively.[159] Thus, iodo-ester Tröger's base **3.52** was synthesized by the reaction of iodo-aniline **3.37** and ethyl 4-aminobenzoate **3.42** (Scheme 3.12, bottom). After column chromatography to separate the symmetric 2,8-diiodo- and 2,8-diester Tröger's bases (**3.54** and **3.55**) from the desired product, **3.52** was obtained in 14% overall yield in one step. This represented an improvement in yield and time over the method used in Scheme 3.11, where **3.52** was obtained in 10% yield over five steps.



Scheme 3.12 One-pot synthesis of iodo-ester Tröger's base **3.52**

For comparison, both diiodo-Tröger's base **3.54** and diester-Tröger's base **3.55** were synthesized under standard Tröger's base conditions (Scheme 3.13).

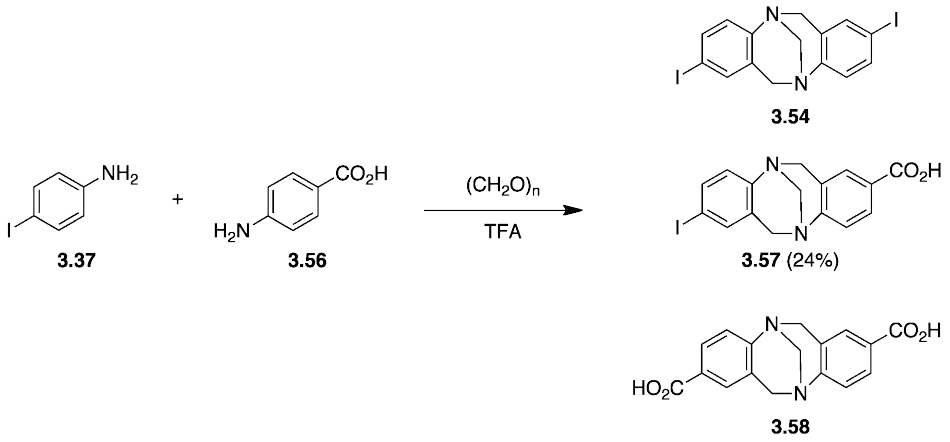


Scheme 3.13 Synthesis of symmetric Tröger's bases **3.54** and **3.55**

*3.9.3 Synthesis of acid-boronic ester Tröger's base **3.59***

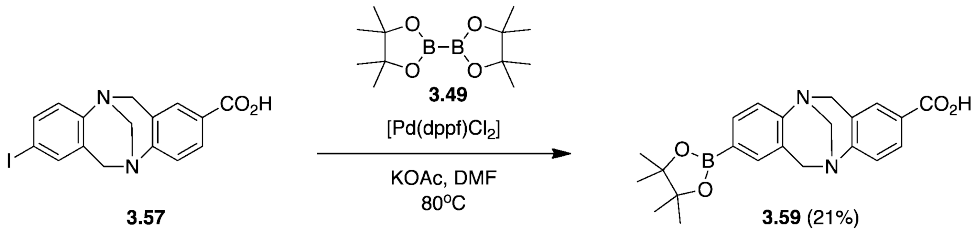
While the one-pot synthesis to produce iodo-ester Tröger's base **3.52** (Scheme 3.12) was again successful, it was also low yielding. This was partially due to purification issues from the similar R_f values of the three products. We thought that synthesis of an iodo-acid Tröger's base via the same method might produce a higher yield, as the R_f difference between symmetric diiodo- and diacid Tröger's base is greater. After reaction

under standard conditions and column chromatography, iodo-acid Tröger's base **3.57** was produced in 24% isolated yield.



Scheme 3.14 Direct synthesis of iodo-acid Tröger's base **3.57**

With **3.57** in hand, we again proceeded to convert it to an organoboron Tröger's base via reaction with B_2pin_2. This was successful in generating acid-boronic ester Tröger's base **3.59** in 21% yield.



Scheme 3.15 Synthesis of acid-boronic ester Tröger's base **3.59**

3.10 Tröger's base building blocks — summary

The synthetic efforts of this chapter have resulted in a number of Tröger's base molecules capable incorporation into other structures via amide bond formation, Suzuki coupling (as the aryl halide or organoboron partner) or both. We have also synthesized the first amino-acid Tröger's base derivative capable of direct incorporation into peptide structures. All of these building blocks presented here include properties that are desirable in the formation of synthetic aromatic cages: aromatic surfaces, hydrophobicity and structural rigidity.

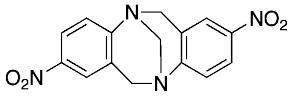
The next steps, partially described in Chapter 4, are to use these building blocks to form water-soluble host molecules and to evaluate their ability to bind and recognize our amino acid targets: methylated lysine and arginine.

3.11 Experimental Section

3.11.1 General considerations

Solvents and reagents were used as obtained from Sigma-Aldrich. Proton (^{1}H) NMR and carbon (^{13}C) NMR spectra were recorded on a Brüker AC300 (300 MHz), Brüker Avance 360 (360 MHz) or Brüker Avance 500 (500 MHz) spectrometer as indicated. Chemical shifts (δ) are given in parts per million (ppm) relative to TMS and referenced to residual protonated solvent ($CHCl_3$: δH 7.26 ppm, δC 77.16 ppm; DMSO δH 2.50 ppm, δC 39.52 ppm; DOH: δH 4.79 ppm).[109] *J* values are given in Hz. Abbreviations used are s (singlet), d (doublet), t (triplet), q (quartet), and m (multiplet). Infrared spectra (IR) were measured on a Perkin Elmer 1000 FT-IR spectrometer. Melting points were obtained using a Gallenkamp Melting Point Apparatus and are uncorrected. Accurate mass measurements were performed on two different ESI-MS instruments (HR-ESI-MS) and are labeled accordingly. Accurate mass measurements labeled with "QTof" were done at the University of Victoria Chemistry mass spectrometry facility on a Q-TOF II system by MicroMass. Samples of 1mg/ml diluted 1:100 in MeOH or ACN were directly infused at a 5-20ul/min flow rate through an ESI source. Spectra obtained at a scanning tine of 1 second, a 100-2000 range with 8,000 resolution and the higher of 3mDa or 5ppm accuracy. Accurate mass measurements labeled with "Orbitrap" were done at the UVic Proteomics Center on an Orbitrap Velos system by Thermo scientific. Samples of 1mg/ml diluted 1:100 in MeOH or CH_3CN were directly infused at a 5-10ul/min flow rate through an ESI source. Spectra were obtained at about 1Hz scanning rate 100-2000 range with 60,000 resolution and less than 1ppm accuracy in most cases.

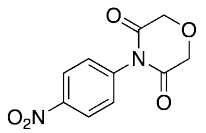
3.11.2 Synthetic procedures



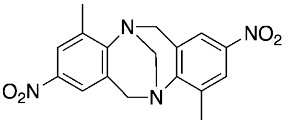
Compound **3.25**. *Via polyphosphoric acid*: 4-nitroaniline **3.24** (514 mg, 3.7 mmol), diglycolic acid (493 mg, 3.7 mmol) and polyphosphoric acid (6.9 g, 70 mmol) were combined in a round bottom flask and heated to 80^{o}C under Ar. After 24 hours, the reaction was cooled to room temperature and diluted with distilled H_2O (50 mL). The precipitated solids were vacuum filtered and washed with additional H_2O. The aqueous washings were further extracted with DCM, condensed, and combined with the solids. After purification by column chromatography (dry loaded, 2 – 10% acetone/toluene), the

product **3.25** was isolated as a mixture with **3.26** (260 mg). Exact conditions for this synthesis were from a published procedure.[161]

Via DMSO/$HCl_{(g)}$: 4-nitroaniline **3.24** (703 mg, 5.1 mmol) was added to a 1:1 mixture of DMSO and AcOH (10 mL total) in a lightly sealed 25 mL pressure tube and heated to 80°C. After 20 minutes at temperature, the flask was opened and $HCl_{(g)}$ bubbled through for 30 minutes using a reaction of H_2SO_4 dropping onto $NaCl_{(s)}$, before sealing the tube completely. After 24 hours sealed, the reaction was poured onto 30 mL ice to quench, followed by addition of sat. aq. Na_2CO_3. The solution was extracted with DCM (3 x 30 mL), and washed once each with distilled H_2O and sat. aq. NaCl. The combined organics were dried over Na_2SO_4, filtered and condensed. After purification by column chromatography (loaded in DCM, 75% Et_2O/hexanes) the product **3.25** (133 mg, 17% yield) was isolated as a yellow solid. Conditions for this synthesis were from a published procedure.[150] Characterization data was found to match published spectra for this known compound.[151] ^{1}H NMR ($CDCl_3$, 300 MHz): δ 4.09 (d, *J* = 16.8 Hz, 2H), 4.27 (s, 2H), 4.36 (d, *J* = 16.8 Hz, 2H), 6.90 (d, *J* = 8.4 Hz, 2H), 7.25 (s, 2H), 7.47 (d, *J* = 7.8 Hz, 2H). ^{13}C NMR ($CDCl_3$, 75 MHz): δ 58.8, 66.5, 123.1, 123.4, 125.8, 128.2, 144.1, 154.0.

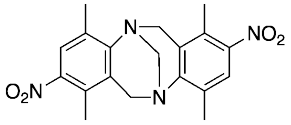


Compound **3.26**. Following procedure for **3.25**, cyclic imide **3.26** was obtained as the more polar, tailing by-product by column chromatography. No references for characterization data were found. Mp: 166 – 169 °C. IR (thin film on KBr) ν (cm^{-1}): 3434w, 3116w, 3086w, 2866w, 1703s, 1699s, 1525m, 1350s, 1262w, 1205m, 1146w, 973w, 855w, 693w, 567w. ^{1}H NMR (DMSO, 500 MHz): δ 4.57 (s, 4H), 7.59 (“d”, *J* = 9.0 Hz, 2H), 8.35 (“d”, *J* = 9.0 Hz, 2H). ^{13}C NMR (DMSO, 125 MHz): δ 67.3, 124.2, 130.4, 139.3, 147.3, 169.5.

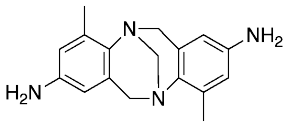


Compound **3.29**. 2-methyl-4-nitroaniline **3.27** (24.0 g, 158 mmol) and paraformaldehyde (10.0 g, 332 mmol) were stirred in TFA (320 mL) under Ar for 91 hours until complete by

TLC (40% EtOAc/hexanes). The reaction was then poured into distilled H_2O (800 mL) and basified to pH 10 with NaOH (10 M). A yellow precipitate was isolated by filtration and heated at reflux in acetone before re-filtering. Air-drying overnight gave compound **3.29** (22.2 g, 83%) as a yellow solid. Synthetic conditions and characterization data for this exact compound were found in a published souce.[171] ^{1}H NMR (DMSO, 500 MHz): δ 2.47 (s, 6H), 4.30 (d, *J* = 17.3 Hz, 2H), 4.35 (s, 2H), 4.67 (d, *J* = 17.2 Hz, 2H), 7.81 (d, *J* = 2.6 Hz, 2H), 7.96 (d, *J* = 2.6 Hz, 2H). ^{13}C NMR (DMSO, 125 MHz): δ 16.9, 54.0, 66.1, 120.3, 123.4, 129.4, 134.6, 142.7, 152.0.

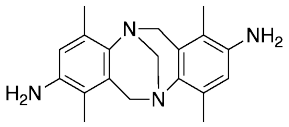


Compound **3.30**. 2,5-dimethyl-4-nitroaniline **3.28** (205 mg, 1.2 mmol) and paraformaldehyde (78 mg, 2.6 mmol) were stirred in TFA (2.5 mL) under Ar for 5 days until complete by TLC. The reaction was poured into distilled H_2O (6 mL) and basified to pH 10 with drop-wise addition of aq. NaOH (3.75 M). The precipitated solid was collected and heated at reflux in acetone (5 mL). Cooling overnight (in the freezer) and filtration of solids gave product **3.30** (172 mg, 76% yield) as a yellow solid. Characterization data was found to match the published spectra for this known compound.[151] Synthesis conditions were adapted from a similar compound.[171] ^{1}H NMR (DMSO, 500 MHz): δ 2.17 (s, 6H), 2.44 (s, 6H), 4.04 (d, *J* = 17.2 Hz, 2H), 4.21 (s, 2H), 4.57 (d, *J* = 18.5 Hz, 2H), 7.65 (s, 2H). ^{13}C NMR (DMSO, 125 MHz): δ 13.0, 16.6, 53.1, 64.5, 123.9, 127.6, 128.4, 131.9, 145.7, 150.0.

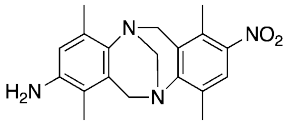


Compound **3.31**. Compound **3.29** (10.0 g, 29.4 mmol) and iron filings (22.3 g, 398 mmol) were suspended in EtOH (400 mL). After addition of AcOH (44 mL, 769 mmol), the reaction was heated to reflux under Ar for 108 hours. Once complete by TLC (75% EtOAc/hexanes), the reaction was cooled and filtered to remove unreacted solids. The filtrate was extracted with DCM (3 x 100 mL), and the combined organics washed with saturated aqueous $NaHCO_3$ (3 x 100 mL), dried over Na_2SO_4, filtered and condensed. After drying under high vacuum overnight, compound **3.31** (7.9 g, 96%) was obtained as

a beige solid. Synthetic conditions for this exact compound were found in a published procedure.[171] ^{1}H NMR (DMSO, 500 MHz): δ 2.18 (s, 6H), 3.62 (d, *J* = 16.7 Hz, 2H), 4.06 (s, 2H), 4.26 (d, *J* = 16.7 Hz, 2H), 4.58 (br s, 4H), 5.94 (d, *J* = 2.3 Hz, 2H), 6.20 (d, *J* = 2.3 Hz, 2H). ^{13}C NMR (DMSO, 125 MHz): δ 16.8, 55.0, 67.9, 108.8, 114.9, 128.5, 132.2, 135.4, 144.1.

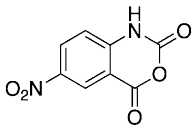


Compound **3.32**. Dinitro-Tröger's base **3.30** (753 mg, 2.0 mmol), iron powder (1.54 g, 27.6 mmol) and glacial AcOH (3.1 mL, 54.2 mmol) were combined in EtOH (30 mL) and heated to reflux under Ar for 20 hours. After cooling, the reaction was poured into distilled H_2O (70 mL) and filtered to remove unreacted metal. The aqueous solution was then extracted with DCM (3 x 20 mL) and the combined organics washed once with saturated aqueous $NaHCO_3$, dried over Na_2SO_4 and condensed. After drying, compound **3.32** (614 mg, 98% yield) was isolated as a pale yellow solid. Synthetic conditions were adapted from similar compound.[171] Mp: 206 – 208 °C (decomp.). IR (thin film on KBr) ν (cm^{-1}): 3351m, 3228m, 2944m, 2880m, 1622s, 1478s, 1412m, 1316w, 1223s, 945m, 734m. ^{1}H NMR (DMSO, 300 MHz): δ 1.73 (s, 6H), 2.22 (s, 6H), 3.68 (d, *J* = 16.7 Hz, 2H), 3.99 (s, 2H), 4.21 (d, *J* = 17.0 Hz, 2H), 4.37 (s, 4H), 6.38 (s, 2H). ^{13}C NMR (DMSO, 75 Hz): δ 10.8, 16.8, 54.1, 66.3, 114.8, 115.2, 126.4, 129.2, 136.4, 141.8. HR-EI-MS: *m/z* $[M+H]^+$ calculated for $C_{19}H_{25}N_4$: 309.20737; found 309.20509 (Orbitrap).

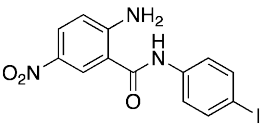


Compound **3.33.** Dinitro-Tröger's base **3.30** (502 mg, 1.4 mmol), iron powder (1.0 g, 18.3 mmol) and glacial AcOH (2.1 mL, 37 mmol) were combined in EtOH (30 mL) and heated to reflux under Ar for 17 hours. After cooling, the reaction was poured into distilled H_2O (60 mL) and filtered to remove unreacted metal. The aqueous solution was then extracted with DCM (3 x 20 mL) and the combined organics washed once with saturated aqueous $NaHCO_3$, dried over Na_2SO_4 and condensed. After column chromatography (10 – 100% EtOAc/hexanes), three spots were separated at R_f values of 0.9, 0.7 and 0.3 (TLC conditions: EtOAc). The spots were identified as starting

material **3.30** (52 mg, 10% yield), mono-reduced Tröger's base **3.33** (93 mg, 20% yield) and fully reduced **3.42** (244 mg, 58% yield). The original hydrogenation procedure was adapted from a similar compound.[171] Mp: 162 – 164 °C. IR (thin film on KBr) ν (cm^{-1}): 3367m, 2949m, 2926m, 2889w, 1706w, 1513s, 1484m, 1468m, 1343s, 1226s, 733s. ^{1}H NMR ($CDCl_3$, 300 MHz): δ 1.87 (s, 3H), 2.22 (s, 3H), 2.35 (s, 3H), 2.45 (s, 3H), 3.93 (d, *J* = 16.8 Hz, 1H), 3.98 (d, *J* = 16.5 Hz, 1H), 4.13 – 4.28 (m, 2H), 4.42 (d, *J* = 16.9 Hz, 1H), 4.47 (d, *J* = 16.7 Hz, 1H), 6.51 (s, 1H), 7.60 (s, 1H). ^{13}C NMR ($CDCl_3$, 90 MHz): δ 11.2, 13.7, 17.1, 17.3, 54.2, 54.6, 66.1, 116.5, 116.9, 124.8, 126.2, 128.6, 128.9, 131.3, 131.7, 137.6, 140.8, 145.9, 151.4. HR-EI-MS: *m/z* $[M+H]^+$ calculated for $C_{19}H_{23}N_4O_2$: 339.18155; found 339.18125 (Orbitrap).

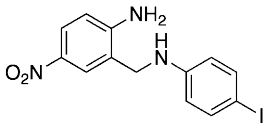


Compound **3.36**. Triphosgene **3.35** (9.23 g, 31.1 mmol) was sequentially added in pieces to a solution of 2-amino-5-nitrobenzoic acid **3.34** (8.02 g, 44.0 mmol) in anhydrous THF (150 mL) stirring at room temperature. The reaction was heated to reflux under Ar for 18 hours. A precipitate was noted after one hour. Once complete by TLC (75% EtOAc/hexanes), the reaction was cooled to room temperature and poured into ice/H_2O (300 mL). The precipitated solid was filtered and washed with H_2O (3 x 100 mL) while suspended in filter. The solid was additionally washed with a minimum of MeOH (2 x 25 mL) in filter and air-dried to give compound **3.36** (8.64 g, 94%) as an off-white solid. Characterization data was found to match published spectra for this known compound.[172] Synthetic conditions were adapted from a similar compound.[173] Product R_f = 0.6 (75% EtOAc/hexanes). ^{1}H NMR (DMSO, 300 MHz): δ 7.31 (d, *J* = 8.9, 1H), 8.53 (dd, *J* = 8.9, *J* = 2.5, 1H), 8.58 (d, *J* = 2.5, 1H), 12.34 (s, 1H, NH). ^{13}C NMR (DMSO, 125 MHz): δ 111.2, 116.7, 124.6, 131.3, 142.5, 146.1, 146.5, 158.6.

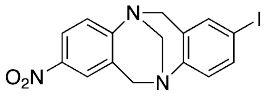


Compound **3.38**. 5-nitroisatoic anhydride **3.36** (1.50 g, 7.2 mmol) was dissolved in anhydrous THF (60 mL, 0.125 M in anhydride) at room temperature. *p*-Iodoaniline **3.37** (1.40g, 6.4 mmol) was added in portions before heating to reflux under Ar for 110 hours.

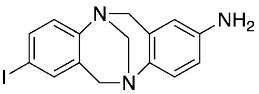
Once complete by TLC (30% EtOAc/hexanes), the reaction was cooled and solvent removed to give a green solid. Column chromatography (dry loading, 15 – 30% EtOAc/hexanes), followed by heating at reflux in DCM, and then filtration gave **3.38** (2.26 g, 92%) as a yellow solid. The amide-forming reaction with 5-nitroisatoic anhydride was adapted from published reference.[174] Mp: 228 – 232 °C (decomp.). IR (thin film on KBr) ν (cm^{-1}): 3474*w*, 3367*s*, 3288*s*, 2358*w*, 1647*s*, 1611*s*, 1587*s*, 1517*w*, 1496*w*, 1389*m*, 1307*s*, 1270*s*, 1130*s*, 918*w*, 816*m*. ^{1}H NMR (DMSO, 300 MHz): δ 6.85 (d, *J* = 9.3, 1H), 7.54 (m, 2H), 7.65 (br s, 2H, NH_2), 7.69 (m, 2H), 8.07 (dd, *J* = 9.3, *J* = 2.6, 1H), 8.60 (d, *J* = 2.6, 1H), 10.49 (s, 1H, NH). ^{13}C NMR (DMSO, 125 MHz): δ 87.7, 113.2, 116.0, 122.9, 126.4, 127.7, 134.9, 137.2, 138.6, 155.2, 166.0. HR-ESI-MS: *m/z* $[M+Na]^+$ calculated for $C_{13}H_{10}IN_3O_3Na$: 405.9665; found 405.9691 (QTof).



Compound **3.39**. Compound **3.38** (2.33 g, 6.1 mmol) was dissolved in dry THF (60 mL) in a flame-dried 250 mL round bottom flask and cooled to 0°C. $BH_3{\cdot}SMe_2$ (36.5 mL, 2.0 M in $PhCH_3$, 73 mmol) was added drop-wise. The reaction was warmed to room temperature before heating at reflux for 1 hour. After slowly quenching reaction on ice with aq. HCl (6 M), and stirring until all bubbling had stopped, aq. NaOH (6 M) was added to neutralize. The reaction was extracted with $CHCl_3$ (3 x 50 mL), washed with brine (1 x 100 mL), dried over $MgSO_4$, filtered and condensed. Column chromatography (dry loaded, 20 – 25% EtOAc/hexanes), followed by trituration in Et_2O and filtration gave **3.39** (1.25 g, 56%) as an orange powder. The reduction procedure was adapted from a related compound.[174] Product R_f: 0.1 (25% EtOAc/hexanes), SM R_f: 0.3 (25% EtOAc/hexanes). Mp: 143 – 145°C (decomp.). IR (thin film on KBr) ν (cm^{-1}): 3368*s*, 2363*w*, 1684*w*, 1623*s*, 1586*s*, 1506*s*, 1490*s*, 1437*m*, 1305*s*, 1280*s*, 1181*w*, 1152*w*, 1100*w*, 1058*w*, 995*w*, 811*m*, 751*w* . ^{1}H NMR (DMSO, 300 MHz): δ 4.06 (d, *J* = 5.9, 2H), 6.41 (d, *J* = 8.7, 2H), 6.46 (t, *J* = 5.9, 1H, NH), 6.61 (br s, 2H, NH_2), 6.69 (d, *J* = 8.8, 1H), 7.34 (d, *J* = 8.7, 2H), 7.89 (dd, *J* = 8.8, *J* = 2.6, 1H), 7.92 (d, *J* = 2.6, 1H). ^{13}C NMR (DMSO, 75 MHz): δ 42.4, 76.7, 113.4, 114.9, 121.2, 123.7, 124.5, 136.0, 137.2, 148.1, 153.1. HR-ESI-MS: *m/z* $[M+Na]^+$ calculated for $C_{13}H_{12}IN_3O_2Na$: 391.9872; found 391.9862 (QTof).

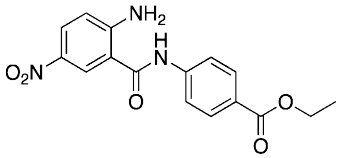


Compound **3.40**. Compound **3.39** (3.14 g, 8.5 mmol) and paraformaldehyde (1.02 g, 34.0 mmol) were dissolved in TFA (170 mL) and stirred at room temperature for 120 hours. The reaction was poured into ice (300 mL) forming a yellow precipitate. $NaOH_{(s)}$ (90 g, 2.25 mol) was added to basify the reaction (pH 9). The precipitated solid was isolated by filtration and heated at reflux in acetone briefly before filtering again. After air-drying overnight, compound **3.40** (3.19 g, 96%) was isolated as a yellow solid. If needed, **3.40** can also be purified by column chromatography (10% EtOAc/hexanes). Tröger's base-forming reaction conditions were adapted from reference.[171] Product R_f: 0.4 (1:1 EtOAc/hexanes). Mp: 156 – 162°C. IR (thin film on KBr) ν (cm^{-1}): 3399*w*, 3054*w*, 2910*w*, 2847*w*, 2358*m*, 1683*w*, 1610*m*, 1579*m*, 1511*s*, 1472*s*, 1436*w*, 1397*w*, 1338*s*, 1294*m*, 1209*s*, 1084*m*, 963*w*, 943*s*, 836*m*, 811*m*, 795*w*, 733*m*. ^{1}H NMR (DMSO, 500 MHz): δ 4.27 (m, 4H), 4.68 (d, *J* = 17.3, 2H), 6.94 (d, *J* = 8.5, 1H), 7.35 (m, 2H), 7.46 (dd, *J* = 8.5, *J* = 2.1, 1H), 7.94 (d, *J* = 2.7, 1H), 8.00 (dd, *J* = 8.9, *J* = 2.7,1H). ^{13}C NMR (DMSO, 125 MHz): δ 57.6, 57.8, 65.5, 87.6, 122.4, 122.9, 125.6, 127.2, 129.3, 130.7, 135.4, 135.8, 142.6, 147.2, 154.8. HR-ESI-MS: *m/z* $[M+Na]^+$ calculated for $C_{15}H_{12}IN_3O_2Na$: 415.9872; found 415.9868 (QTof).

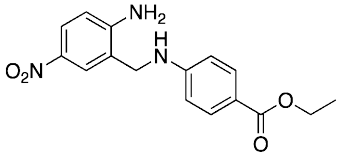


Compound **3.41**. To a solution of compound **3.40** (488 mg, 1.3 mmol) in EtOH (95%, 25 mL), was added iron filings (470 mg, 8.4 mmol) and AcOH (1.0 mL, 17.5 mmol). The reaction was then heated to reflux for 21 hours. After cooling to room temperature, the reaction was poured into a 1:1 mixture of saturated aq. $NaHCO_3$ and DCM before filtering to remove unreacted iron solids. The filtrate was extracted with DCM (3 x 50 mL) and washed with saturated aq. NaCl, dried over Na_2SO_4 and filtered. After solvent removal, **3.41** (370 mg, 82%) was isolated as a beige solid. Characterization data was found to match published spectra for this known compound.[157] Reduction conditions were adapted from reference.[175] Mp: 200 – 204°C. ^{1}H NMR (DMSO, 300 MHz): δ 3.91 (d, *J* = 16.6, 1H), 3.92 (d, *J* = 17.0, 1H), 4.07 (d, *J* = 13.0, 1H, AB system), 4.14 (d, *J* = 13.0, 1H, AB system), 4.45 (d, *J* = 17.0, 1H), 4.47 (d, *J* = 16.6, 1H), 4.71 (br s, 2H, NH_2), 6.07 (d, *J* = 2.5, 1H), 6.37 (dd, *J* = 8.5, *J* = 2.5, 1H), 6.73 (d, *J* = 8.5, 1H), 6.90 (d, *J* = 8.5, 1H), 7.29 (d, *J* = 2.0, 1H), 7.42 (dd, *J* = 8.5, *J* = 2.0, 1H). ^{13}C NMR (DMSO, 75

MHz): δ 57.5, 58.4, 66.4, 86.9, 110.7, 113.8, 125.1, 127.1, 127.8, 131.6, 135.3 (x 2), 136.6, 144.8, 148.4. HR-ESI-MS: *m/z* $[M+H]^+$ calculated for $C_{15}H_{14}IN_3H$: 364.0311; found 364.0287 (QTof).

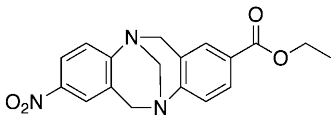


Compound **3.43**. 5-nitroisatoic anhydride **3.36** (15.0 g, 0.072 mol) and 4-ethyl aminobenzoate **3.42** (26.0 g, 0.158 mol) were measured into a 500 mL round bottom flask. 1,4-dioxane (87 mL) was added and the solution heated to reflux for 16 hours. The solution was then removed from heat and left to cool briefly while a 4:1 MeOH:H_2O mixture (218 mL) was heated. The warm mixture was then added to the cooling dioxane solution resulting in the formation of a yellow precipitate. The suspension was filtered, and the solid washed with distilled H_2O and Et_2O. The product was then air-dried to give **3.43** as a yellow solid (18.64 g, 78% yield). Synthetic conditions were adapted from a published procedure for a similar compound.[165] Mp.: 206 – 208 °C. IR (KBr pellet) ν (cm^{-1}): 3452*w*, 3377*s*, 3281*s*, 1653*s*, 1614*s*, 1568*s*, 1522*w*, 1502*w*, 1376*m*, 1310*s*, 1275*s*, 1128*s*, 912*w*, 816*m*. ^{1}H NMR (DMSO, 300MHz): δ 1.32 (t, *J* = 7.2, 3H), 4.30 (q, *J* = 7.2, 2H), 6.87 (d, *J* = 9.5, 1H), 7.67 (s, 1H), 7.87 (d, *J* = 8.8, 2H), 7.97 (d, *J* = 8.8, 2H), 8.09 (dd, *J* = 9.4, *J* = 2.3, 1H), 8.64 (d, *J* = 2.85, 1H), 10.71 (s, 1H). ^{13}C NMR (DMSO, 75MHz): δ 14.2, 60.5, 113.1, 116.1, 120.0, 124.7, 126.7, 127.9, 130.0, 134.9, 143.3, 155.2, 165.3, 166.4. ^{13}C DEPT-135 (DMSO, 75MHz): δ 14.19; 60.45; 116.02; 119.97 (x2); 126.63; 127.85; 129.94 (x2). HR-ESI-MS: *m/z* $[M+Na]^+$ calculated for $C_{16}H_{15}N_3O_5Na$: 352.0909; found 352.0905 (Orbitrap).

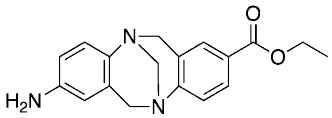


Compound **3.44**. Compound **3.43** (1.0 g, 3.04 mmol) was added to a flame dried 250 mL round bottom flask, followed by dry THF (35 mL) *via* syringe. The flask was cooled to 0°C and $BH_3{\cdot}SMe_2$ (12.0 mL, 2.0 M in THF, 24.0 mmol) was added drop-wise *via* syringe. The solution was left to stir at room temperature for 18 hours under N_2. After cooling to 0°C, the solution was quenched with MeOH (5 mL) and stirred at room

temperature until all bubbling ceased and a yellow colour had returned. Solvent was removed under reduced pressure until an orange film remained. Column chromatography (dry loaded, 35:65 EtOAc:hexanes) eluted the product **3.44** as a bright yellow solid (520 mg, 54% yield). Synthetic procedure adapted from published reference for a similar compound.[165] Mp.: 163 – 166 ^{o}C. IR (thin film on KBr) ν (cm^{-1}): 3387br, 2977w, 1674m, 1602s, 1585m, 1313s, 1278s, 1174m, 1104m, 772m. ^{1}H NMR (DMSO, 300MHz): δ 1.26 (t, J = 7.1, 3H), 4.19 (m, J_q = 7.1, 4H), 6.60 (d, J = 7.2, 2H), 6.65 (s, 2H), 6.71 (dd, J = 7.4, J = 1.8, 1H), 7.06 (t, J = 5.9, 1H), 7.69 (d, J = 8.8, 2H), 7.89 (s+d, J_d = 7.9, 2H). ^{13}C NMR (DMSO, 125MHz): δ 14.3, 42.0, 59.6, 111.3, 113.4, 116.8, 120.8, 123.7, 124.6, 131.0, 135.9, 152.4, 153.1, 165.8. ^{13}C DEPT-135 (DMSO, 125MHz): δ 14.3, 42.0, 59.6, 111.3, 113.4, 123.7, 124.7, 131.0. HR-ESI-MS: *m/z* [M+Na]$^{+}$ calculated for $C_{16}H_{17}N_3O_4Na$: 338.1117; found 338.1114 (Orbitrap).

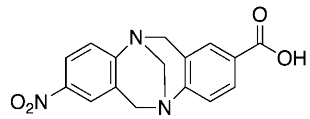


Compound **3.45**. Compound **3.44** (20 mg, 0.063 mmol) and paraformaldehyde (7.6 mg, 0.25 mmol) were added to a 100mL RBF. TFA (53 mL) was added and the solution was left to stir at room temperature for 48 hours. The solution was poured into a 250 mL Erlenmeyer flask containing 50 mL ice. $NaOH_{(aq)}$ (2.0 M) was added slowly while stirring until pH 14. The beige-yellow precipitate was then filtered and washed with distilled H_2O and diethyl ether and dried to give **3.45** (19.2 mg, 89% yield). Synthetic procedure was adapted from reference.[175] Mp.: 198 – 200 ^{o}C (decomp). IR (thin film on KBr) ν (cm^{-1}): 3417br, 1710s, 1611m, 1515m, 1340s, 1285s, 1209s, 950m. ^{1}H NMR (DMSO, 300MHz): δ 1.25 (t, J = 7.0, 3H), 4.23 (q, J = 7.0, 2H), 4.36 (m, 4H), 4.75 (dd, J = 16.8, J = 3.6, 2H), 7.23 (d, J = 8.4, 1H), 7.36 (d, J = 8.8, 1H), 7.59 (s, 1H), 7.71 (dd, J = 8.4, J = 1.7, 1H), 7.94 (d, J = 2.4, 1H), 7.99 (dd, J = 8.9, J = 2.6, 1H). ^{13}C NMR (DMSO, 75MHz): δ 14.1, 57.9, 58.0, 60.4, 65.5, 122.4, 122.8, 125.0, 125.6 (x2); 128.0, 128.4, 129.2, 142.6, 152.12, 154.7, 165.2. ^{13}C DEPT-135 (DMSO, 75MHz): δ 14.1, 57.9, 58.0, 60.4, 65.5, 122.4, 122.8, 125.0, 125.6 128.0, 128.4. HR-ESI-MS: *m/z* [M+Na]$^{+}$ calculated for $C_{18}H_{17}N_3O_4Na$: 362.1117; found 362.1116 (QTof).



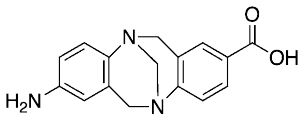
Compound **3.46**. *Via Fe/AcOH*: A suspension of compound **3.45** (100 mg, 0.30 mmol), iron powder (218 mg, 3.9 mmol), AcOH (0.44 mL, 7.7 mmol) and EtOH (5.8 mL, 0.05 M in compound **3.45**) was heated to reflux for 1.5 h. After completion was indicated by TLC (1:1 EtOAc/hexanes), the reaction was cooled and diluted with sat. aq. $NaHCO_3$. The solution was decanted and extracted with DCM (3 x 10 mL). The combined organic layers were washed with once each with sat. aq. $NaHCO_3$, H_2O and brine before being dried over Na_2SO_4, filtered and condensed. The crude oil was suspended in Et_2O to give a free flowing solid. Isolation by filtration yielded the product **3.46** (45 mg, 49% yield) as a beige solid. Reduction method was adapted from referenced procedure.[175]

Via Raney Nickel hydrogenation: Raney Nickel (approximately 1 mL, Raney® 2800 nickel, slurry in water) was washed with EtOAc (15 mL) and the EtOAc/H_2O solution decanted. To this was added a solution of nitro-ester Tröger's base **3.45** (80 mg, 0.2 mmol) in EtOAc (2 mL). The reaction was capped and H_2 was bubbled through the solution. After 1 hour, the reaction was filtered through Celite and condensed to give **3.46** (65 mg, 89% yield). Reduction conditions were adapted from published procedure.[166] Mp.: 198 – 200 ^oC (decomp.). IR (thin film on KBr) ν (cm^{-1}): 3418br, 1699m, 1683m, 1645m, 1615w, 1497s, 1288s, 1206s, 1181s, 1110m. ^1H NMR (DMSO, 300 MHz): δ 1.26 (t, *J* = 7.1, 3H), 4.02 (d, *J* = 16.8, 2H), 4.16 (d, *J* = 6.3, 2H), 4.23 (q, *J* = 7.1, 2H), 4.54 (d, *J* = 16.8, 2H), 4.71 (br s, 2H), 6.07 (d, *J* = 2.3, 1H), 6.37 (dd, *J* = 8.5, *J* = 2.3, 1H), 6.75 (d, *J* = 8.5, 1H), 7.20 (d, *J* = 8.4, 1H), 7.56 (d, *J* = 1.5, 1H), 7.70 (dd, *J* = 8.7, *J* = 1.5, 1H). ^{13}C NMR (DMSO, 90 MHz): δ 14.2, 57.9, 58.5, 60.3, 66.5, 110.6, 113.8, 124.4, 124.8, 125.2, 127.6, 127.8, 128.4, 128.7, 136.5, 144.8, 153.6, 165.4. HR-ESI-MS: *m/z* $[M+Na]^+$ calculated for $C_{18}H_{19}N_3O_2H$: 310.1555; found 310.1552 (Orbitrap).

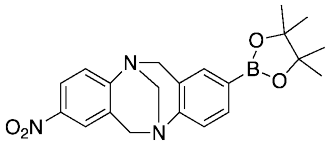


Compound **3.47**. Compound **3.45** (160 mg, 0.47 mmol) was suspended in MeOH (16 mL), and $KOH_{(aq)}$ (16 mL, 0.5 M) was added drop-wise. The suspension was heated at reflux for 48 hours. After cooling, the basic solution was washed with EtOAc (3 x 25 mL). The remaining aqueous solution was acidified with $HCl_{(aq)}$ (1 M), extracted with EtOAc (3

x 30 mL), dried over Na_2SO_4, filtered and condensed. The solid was dried overnight on high vacuum to give **3.47** (100 mg, 71% yield) as a yellow solid. Mp: >250 °C.IR (KBr pellet) ν (cm^{-1}): 3417br, 1705s, 1616m, 1520m, 1336s, 1289s, 1204s, 953m. ^{1}H NMR (DMSO, 300MHz): δ 4.28 – 4.39 (m, 4H), 4.74 (d, *J* = 17.0, 1H), 4.76 (d, J = 16.7, 1H), 7.21 (d, *J* = 8.4, 1H), 7.37 (d, *J* = 8.9, 1H), 7.58 (d, J = 2.3, 1H), 7.70 (dd, *J* = 8.4, *J* = 2.0, 1H), 7.95 (d, *J* = 2.6, 1H), 8.00 (dd, *J* = 8.8, *J* = 2.8, 1H), 12.71 (s, 1H). ^{13}C NMR (DMSO, 75 MHz): δ 57.9, 58.1, 65.6, 122.4, 122.9, 124.9, 125.6, 125.9, 127.9, 128.3, 128.6, 129.3, 142.6, 151.9, 154.8, 166.8. HR-ESI-MS: *m/z* $[M+H]^+$ calculated for $C_{16}H_{13}N_3O_4H$: 312.0979; found 312.09788 (Orbitrap).

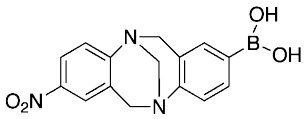


Compound **3.48**. Raney Nickel (approximately 1 mL, Raney® 2800 nickel, slurry in water) was washed with EtOAc (15 mL) and the EtOAc/H_2O mixture decanted. To this was added a solution of nitro-acid Tröger's base **3.47** (76 mg, 0.24 mmol) in EtOAc (12 mL). The reaction was capped and H_2 was bubbled through the solution. After 1 hour, the yellow colour suspension was replaced by a grey suspension. The reaction was diluted with excess MeOH to dissolve the solid, filtered through Celite, and condensed to give **3.48** (39 mg, 57% yield). Reduction conditions were adapted from published procedure for a similar compound.[166] Mp: >240 °C. IR (KBr pellet) ν (cm^{-1}): 3417br, 3245br, 1708s, 1616m, 1583m, 1336s, 1273s, 1216s, 946m. ^{1}H NMR (DMSO, 500 MHz): δ 3.98 (d, *J* = 16.6, 1H), 3.99 (d, *J* = 16.4, 1H), 4.10 – 4.20 (m, 2H), 4.52 (d, *J* = 16.6, 2H), 6.08 (s, 1H), 6.37 (dd, *J* = 8.4, *J* = 1.9, 1H), 6.75 (d, *J* = 8.4, 1H), 7.12 (d, *J* = 8.3, 1H), 7.51 (s, 1H), 7.66 (d, *J* = 8.1, 1H). ^{13}C NMR (DMSO, 125 MHz): δ 58.0, 58.5, 66.5, 110.7, 113.8, 124.3, 125.2, 127.8, 127.9 (x2), 128.0, 128.4, 136.7, 144.8, 152.3, 167.6. HR-ESI-MS: *m/z* $[M+H]^+$ calculated for $C_{16}H_{15}N_3O_2H$: 282.1237; found 212.1230 (Orbitrap).



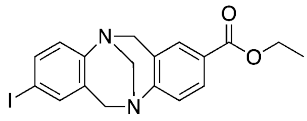
Compound **3.50**. Compound **3.40** (402 mg, 1.0 mmol), bis(pinacolato)diboron **3.49** (331 mg, 1.3 mmol), KOAc (303 mg, 3.1 mmol) and $[Pd(dppf)Cl_2]$

(bis(diphenylphosphino)ferrocene) (26 mg, 0.04 mmol) were added to a Schlenk tube, which was then evacuated and backfilled with N_2 (x 3). Anhydrous DMF (10 mL) was added via syringe and the reaction heated to 80°C for 23 hours. After cooling to room temperature, the reaction was diluted with benzene and portioned with sat. aq. NaCl. The separated organic layer was washed with H_2O (2 x 20 mL), dried over Na_2SO_4, and condensed. The residue was taken up in DCM and purified by column chromatography (DCM to 10% EtOAc/DCM). Further drying under high vacuum gave compound **3.50** (353 mg, 88%) cleanly as a shiny yellow solid. Synthetic conditions for coupling to iodo-Tröger's bases adapted from reference procedure.[158] Product R_f: 0.5 (1:1 EtOAc/hexanes). Mp: 200 – 204°C. IR (thin film on KBr) ν (cm^{-1}): 3435br, 2978w, 1610m, 1515m, 1361s, 1340s, 1211m, 1145m, 965m, 947m, 671w. ^{1}H NMR ($CDCl_3$, 300 MHz): δ 1.29 (s, 12H, 4 x CH3), 4.29 (m, 4H), 4.75 (dd, *J* = 16.7, *J* = 5.8, 2H), 7.13 (d, *J* = 8.0, 1H), 7.21 (d, *J* = 8.9, 1H), 7.40 (s, 1H), 7.62 (d, *J* = 7.7, 1H), 7.83 (d, *J* = 2.5, 1H), 8.00 (dd, *J* = 8.9, *J* = 2.5, 1H). ^{13}C NMR ($CDCl_3$, 75 MHz): δ 24.7, 24.8, 58.6, 58.9, 66.6, 83.8, 122.8, 123.0, 124.3, 125.5, 126.5, 128.7, 133.8, 134.2, 143.5, 150.2, 154.8. HR-ESI-MS: *m/z* $[M+H]^+$ calculated for $C_{21}H_{24}BN_3O_4H$: 394.193326; found 394.19392 (Orbitrap).



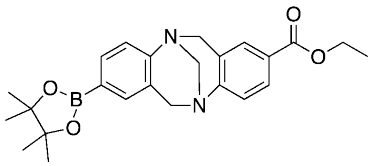
Compound **3.51**. Compound **3.50** (30 mg, 0.08 mmol) was suspended in acetone (2 mL) before adding NH_4OAc (0.1 M, 1.6 mL) and $NaIO_4$ (51 mg, 0.24 mmol) and stirring at room temperature for 30 hours. The acetone was removed by evaporation, and the residue diluted with $NaOH_{(aq)}$ (2 M) and washed with DCM (3 x 5 mL). The aqueous layer was acidified to pH 3 with $HCl_{(aq)}$ (1 M) followed by extraction with DCM (4 x 10 mL). The combined organics were dried over Na_2SO_4, filtered, and condensed to give **3.51** (23 mg, 92% yield) as a flaky yellow film. Hydrolysis conditions were adapted published reference.[169] Mp: 176 – 178 °C (decomp.). IR (thin film on KBr) ν (cm^{-1}): 3392br, 3055w, 2920w, 1610m, 1582w, 1514m, 1340s, 1265s, 1209m, 946m, 738s, 706m. ^{1}H NMR (DMSO, 300 MHz): δ 4.20 – 4.37 (m, 4H), 4.69 (d, *J* = 13.2, 1H), 4.74 (d, *J* = 12.5, 1H), 7.05 (d, *J* = 8.0, 1H), 7.36 (d, *J* = 0.8, 1H), 7.38 (d, *J* = 8.9, 1H), 7.55 (dd, *J* = 8.0, *J* = 1.0, 1H), 7.86 (s, 2H), 7.94 (d, *J* = 2.6, 1H), 7.98 (dd, *J* = 8.9, *J* = 2.7, 1H). ^{13}C NMR (DMSO, 90 MHz): δ 57.9, 58.4, 65.8, 122.2, 122.8, 123.7, 125.5, 126.5, 129.5, 133.1

(x2), 142.5, 149.2, 155.3 (missing one quaternary carbon signal). HR-ESI-MS: *m/z* $[M+H]^+$ calculated for $C_{15}H_{14}BN_3O_4H$: 312.11556; found 312.11565 (Orbitrap).

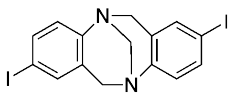


Compound **3.52**. *Via Sandmeyer reaction*: Amino-ester Tröger's base **3.46** (40mg, 0.13 mmol) was dissolved in $H_2SO_{4(aq)}$ (20% solution, 2 mL) and cooled to 0°C. A solution of $NaNO_2$ (9.7mg, 0.14 mmol in 0.5 mL H_2O) was added drop-wise and the solution stirred at 0°C. After 30 minutes, the reaction was added to a separate solution of KI (107 mg, 0.65 mmol) in distilled H_2O (4 mL) and stirred overnight. The resulting suspension was poured over ice (50 mL) and extracted with DCM (3 x 20 mL). The combined organics were washed with sat. aq. $NaHCO_3$ (2 x 30 mL) and brine (4 x 30 mL) before being dried over $MgSO_4$,filtered and condensed. Sonication in hexanes, followed by filtration yielded the product **3.52** (16 mg, 30% yield) as a white solid. Sandmeyer conditions were adapted from published reference.[166]

Via one-pot method: Ethyl-4-aminobenzoate **3.42** (960 mg, 5.8mmol), *p*-iodoaniline **3.37** (1.27 g, 5.8mmol) and paraformaldehyde (590mg, 18.6mmol) were dissolved in neat TFA (25 mL) and stirred for 6 days. The reaction was basified with 2:1 H_2O:NH_4OH solution (90 mL), followed by dilution with saturated $NaHCO_3$ (100 mL), and extraction with DCM (3 x 50 mL). The combined organic layers were washed with brine (3 x 50 mL), dried over Na_2SO_4, filtered and condensed. The crude organic was purified by column chromatography (DCM to 15% EtOAc/DCM). Compounds eluted in order of di-iodo TB, iodo-ester TB, and then di-ester TB. Column chromatography yielded the product **3.52** (340 mg, 14% yield) as a white solid. Synthetic procedure adapted from reference for similar compounds.[159] Mp: 80 – 84 °C. IR (thin film on KBr) ν (cm^{-1}): 3418br, 2378w, 2902w, 1708s, 1609m, 1474s, 1282s, 1208s, 1113m, 772w. ^{1}H NMR ($CDCl_3$, 300 MHz): δ 1.34 (t, *J* = 7.1, 3H), 4.17 (d, *J* = 16.7, 2H), 4.26 – 4.36 (m, 4H), 4.66 (d, *J* = 16.7, 1H), 4.70 (d, *J* = 16.7, 1H), 6.88 (d, *J* = 8.4, 1H), 7.15 (d, *J* = 8.4, 1H), 7.23 (d, *J* = 1.9, 1H), 7.45 (dd, *J* = 8.5, *J* = 1.9, 1H), 7.63 (d, *J* = 1.6, 1H), 7.84 (dd, *J* = 8.4, *J* = 1.9 Hz, 1H). ^{13}C NMR ($CDCl_3$, 75 MHz): δ 14.5, 58.3, 58.8, 60.9, 66.7, 87.7, 124.9, 126.3, 127.2, 127.6, 128.9, 129.0, 130.3, 135.8, 136.6, 147.7, 152.5, 166.3. HR-ESI-MS: *m/z* $[M+H]^+$ calculated for $C_{18}H_{17}IN_2O_2H$: 421.0408; found 421.0408 (Orbitrap).

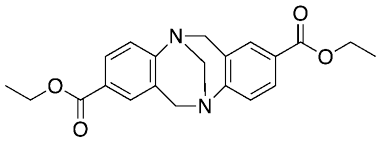


Compound **3.53**. Compound **3.52** (51 mg, 0.1 mmol), bis(pinacolato)diboron **3.49** (39 mg, 0.15 mmol), KOAc (38 mg, 0.4 mmol) and [Pd(dppf)Cl_2] (bis(diphenylphosphino)ferrocene) (3.3 mg, 0.005 mmol) were added to a Schlenk tube, which was then evacuated and backfilled with N_2 (x 3). Anhydrous DMF (1.5 mL) was added via syringe and the reaction heated to 80°C for 21 hours. After cooling to room temperature, the reaction was diluted with benzene and portioned with sat. aq. NaCl. The separated organic layer was washed with H_2O (2 x 50 mL), dried over Na_2SO_4, and condensed. The residue was taken up in DCM and purified by column chromatography (DCM to 50% EtOAc/DCM). Further drying under high vacuum gave compound **3.53** (31 mg, 62%) cleanly as an oil and eventually a solid. Synthetic conditions for coupling to iodo-Tröger's bases were adapted from published reference.[158] Mp: 119 – 122 °C (decomp). IR (thin film on KBr) ν (cm^{-1}): 3417br, 2978m, 2928w, 2852w, 1714s, 1610s, 1361s, 1284s, 1145m, 1105m, 965m, 731m, 671w. ^{1}H NMR ($CDCl_3$, 500 MHz): δ 1.28 (s, 12H), 1.33 (t, *J* = 7.1, 3H), 4.18 – 4.35 (m, 6H), 4.71 (d, *J* = 16.6, 2H), 7.12 (d, *J* = 8.2, 1H), 7.14 (d, *J* = 8.5, 1H), 7.38 (s, 1H), 7.57 – 7.64 (m, 2H), 7.81 (dd, *J* = 8.5, *J* = 1.7, 1H). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 14.5, 24.9, 25.0, 58.7, 58.8, 60.9, 66.9, 83.8, 124.5, 124.9, 126.0, 127.0, 127.8, 128.8, 128.9, 133.9, 134.0, 150.8, 152.8, 166.3. HR-EI-MS: *m/z* $[M+H]^+$ calculated for $C_{24}H_{29}BN_2O_4H$: 421.22986; found 421.22772 (Orbitrap).

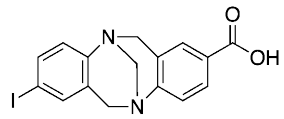


Compound **3.54**. 4-iodoaniline **3.57** (1.66 g, 7.6 mmol) and paraformaldehyde (454 mg, 15.1 mmol) were mixed as solids and added in portions to a vigorously stirred solution of TFA (12.5 mL) cooled to -5 ^{o}C. The solution was allowed to warm to room temperature and stirred for 44 hours under N_2. The reaction was slowly poured into a mixture of ice and concentrated NH_4OH to basify, followed by extraction with DCM (3 x 75 mL). The combined organic layers were dried over Na_2SO_4, filtered and condensed. Column chromatography (10% EtOAc/DCM) yielded the product **3.54** (795 mg, 44% yield) as a

yellow solid. Characterization data was found to match published spectra for this known compound.[176] ^{1}H NMR ($CDCl_3$, 300 MHz): δ 4.09 (d, *J* = 16.8, 2H), 4.27 (s, 2H), 4.64 (d, *J* = 16.8, 2H), 6.90 (d, *J* = 8.5, 2H), 7.25 (s, 2H), 7.47 (d, *J* = 7.7, 2H). ^{13}C NMR ($CDCl_3$, 75 MHz): δ 58.2, 66.6, 87.7, 127.1, 130.3, 135.8, 136.6, 147.7.

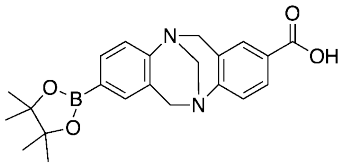


Compound **3.55**. Ethyl 4-aminobenzoate **3.42** (1.0 g, 6.1 mmol) and paraformaldehyde (382 mg, 12.7 mmol) were dissolved in neat TFA and stirred at room temperature. After 72 hours, the reaction was added drop-wise to distilled H_2O (50 mL), basified with $NaOH_{(aq)}$ (10%), and extracted with DCM (3 x 50 mL). The organic layer was washed once with brine, dried over Na_2SO_4, filtered and condensed to give a crude oil. Purification by column chromatography (25% to 50% EtOAc/hexanes), followed by rotary evaporation gave the product **3.55** (390 mg, 35% yield) as a solid. Characterization data was found to match published spectra for this known compound.[155] ^{1}H NMR ($CDCl_3$, 300 MHz): δ 1.33 (t, *J* = 7.1, 3H), 4.21 – 4.35 (m, 8H), 4.74 (d, *J* = 16.6, 2H), 7.16 (d, *J* = 8.4, 2H), 7.63 (d, *J* = 1.8, 2H), 7.83 (dd, *J* = 8.4, *J* = 2.0, 2H). ^{13}C NMR ($CDCl_3$, 90 MHz): δ 14.4, 58.8, 60.9, 66.8, 125.0, 126.2, 127.5, 128.9, 129.0, 152.5, 166.2.



Compound **3.57**. 4-iodoaniline **3.37** (1.3 g, 5.8 mmol), 4-aminobenzoic acid **3.56** (0.8 g, 5.8 mmol) and paraformaldehyde (558 mg, 18.6 mmol) were dissolved in neat TFA and stirred at room temperature. After 7 days, the reaction was basified with 2:1 H_2O:NH_4OH on ice and diluted with saturated $NaHCO_3$ (100 mL). The aqueous layer was extracted with EtOAc (3 x 50 mL), and the combined organics were washed with $NaOH_{(aq)}$ (2M), dried over Na_2SO_4, filtered and condensed to obtain a crude oil. Following column chromatography (EtOAc/DCM) the product **3.57** (547 mg, 24% yield) was isolated as a white solid. Spectral data a match to known compound.[158] ^{1}H NMR (DMSO, 300 MHz): δ 4.18 (d, *J* = 17.1, 1H), 4.19 (d, *J* = 17.5, 1H), 4.23 (s, 2H), 4.62 (d, *J* = 17.4, 1H), 4.64 (d, *J* = 16.9, 1H), 6.93 (d, *J* = 8.5, 1H), 7.18 (d, *J* = 8.4, 1H), 7.32 (d, *J* = 2.0, 1H), 7.45 (dd, *J* = 8.5, *J* = 2.0, 1H), 7.56 (d, *J* = 1.8, 1H), 7.70 (dd, *J* = 8.4, *J* = 2.0, 1H), 12.65 (br s,

1H). ^{13}C NMR (DMSO, 90 MHz): δ 57.5, 57.9, 65.7, 87.3, 124.7, 125.6, 127.1, 127.9, 128.1, 128.5, 131.0, 135.3, 135.6, 147.7, 152.4, 166.9. HR-ESI-MS: *m/z* $[M+H]^+$ calculated for $C_{16}H_{13}IN_2O_2H$: 393.0094; found 393.0094 (Orbitrap).



Compound **3.59**. Compound **3.57** (408 mg, 1.0 mmol), bis(pinacolato)diboron **3.49** (378 mg, 1.5 mmol), KOAc (343 mg, 3.5 mmol) and $[Pd(dppf)Cl_2]$ (bis(diphenylphosphino)ferrocene) (37 mg, 0.05 mmol) were added to a Schlenk tube, which was then evacuated and backfilled with N_2 (x 3). Anhydrous DMF (10 mL) was added via syringe and the reaction heated to 80°C for 24 hours. After cooling to room temperature, the reaction was diluted with benzene and portioned with sat. aq. NaCl. The separated organic layer was washed with H_2O (2 x 50 mL), dried over Na_2SO_4, and condensed. The residue was then purified by column chromatography (50% EtOAc/DCM to 100% EtOAc) and the resulting oil was sonicated in hexanes, filtered and air dried to give compound **3.59** (84 mg, 21%) cleanly as a solid. Mp: >240°C (decomp). IR (thin film on KBr) ν (cm^{-1}): 3370*br*, 2978*w*, 2923*w*, 2851*w*, 1694*m*, 1607*s*, 1361*s*, 1144*m*, 964*w*, 852*w*, 669*w*. ^{1}H NMR ($CDCl_3$, 500 MHz): δ 1.28 (s, 12H), 4.25 (d, *J* = 16.6, 2H), 4.31 (d, *J* = 13.0, 1H), 4.35 (d, *J* = 13.0, 1H), 4.73 (d, *J* = 16.6, 2H), 7.14 (d, *J* = 8.0, 1H), 7.18 (d, *J* = 8.4, 1H), 7.39 (apparent s, 1H), 7.61 (dd, *J* = 8.0, *J* = 1.0, 1H), 7.66 (d, *J* = 1.3, 1H), 7.85 (dd, *J* = 8.4, *J* = 1.4, 1H) (acid proton missing). ^{13}C NMR ($CDCl_3$, 125 MHz): d 24.9, 58.7, 58.8, 66.7, 83.9, 124.5, 124.8, 125.1, 125.2, 126.9, 127.9, 129.4, 129.7, 134.0, 134.1, 150.5, 153.4, 170.8. HR-ESI-MS: *m/z* $[M-H]^-$ calculated for $C_{22}H_{26}BN_2O_4$: 393.19856; found 393.19816 (Orbitrap).

Chapter 4 – Synthesis and study of water-soluble Tröger's base receptors as aromatic cage mimics

4.1 Foreword

The synthetic efforts of Chapter 3 resulted in a small collection of Tröger's base synthons, as well as methodology to pursue additional interesting derivatives. The Tröger's base units synthesized were designed to be incorporated into larger structures using amide bonds, and carbon-carbon bonds from Pd-catalyzed Suzuki coupling reactions. With these structural units, we envisioned using a modular approach to our synthesis and evaluation of synthetic aromatic cage molecules. Depending on the affinities achieved with the original unit, functional groups could be transformed or alternative building blocks (from Chapter 3) could be substituted into the synthesis, allowing us to quickly increase our knowledge of a particular synthetic cage.

4.2 Aims and contributions

This chapter presents our attempts to form water-soluble molecules from the Tröger's base building blocks of Chapter 3, and incorporate them into functional, pre-organized receptors. Specifically, this chapter will show:

1. The synthesis of water-soluble derivatives incorporating a single Tröger's base, and their binding in water to quaternary ammonium ions.
2. The incorporation of two Tröger's base molecules with a single, rigid spacing molecule, and attempts at water-solubilization.
3. The synthesis of a Tröger's base-calix[4]arene hybrid, and its selectivity towards the methylated amino acids of interest in water.

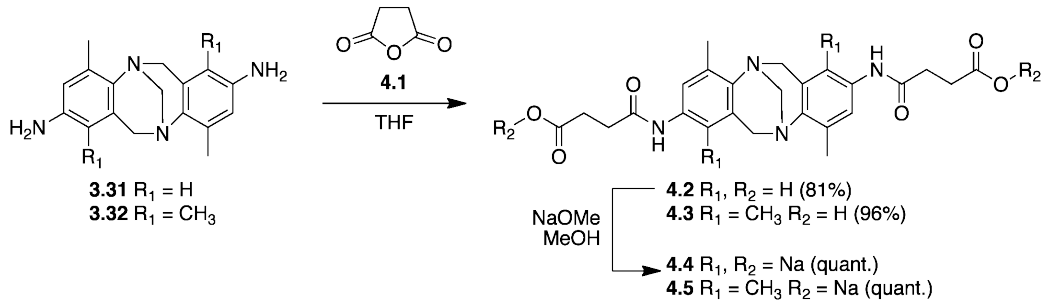
The synthesis and characterization of all molecules presented in this chapter were performed by A. Whiting. The protected monomethyl, mono-NHS ester of succinate **4.14** was synthesized by Dr. C. Beshara. The calixarene starting material, bromo-trisulfonated calix[4]arene **4.18**, was first synthesized by K. Daze[177] and provided for use in these studies by K. Daze, K. Allen, and G. Garnett. All NMR titrations, data analysis and interpretation were performed by A. Whiting.

4.3 Single, water-soluble Tröger's base hosts

4.3.1 Synthesis

Before venturing into dual or more complex Tröger's base cages, we sought to look at the chemical stability and binding properties of a single Tröger's base-containing

molecule in aqueous solution. Chapter 3 saw the synthesis of a number of symmetric, diamino-substituted Tröger's bases. To impart water-solubility to these scaffolds, we used the amino handles as nucleophiles to open succinic anhydride **4.1**. This resulted in two new amide bonds and two pendent carboxylic acids (**4.2** and **4.3**), reminiscent of the pendent groups of host **2.1** and **2.2** from Chapter 2. Upon treatment with two equivalents of sodium methoxide, **4.4** and **4.5** were isolated as their water-soluble sodium salts.



Scheme 4.1 Synthesis of mono-Tröger's base hosts **4.4** and **4.5**

4.3.2 Aqueous binding studies

We wanted to study Tröger's bases under aqueous conditions, but were curious about the stability of Tröger's bases to hydrolysis in water. While other water-soluble Tröger's bases had been presented in the literature (see Section 3.3.1), no mention was made of their stability. To test this for ourselves, we observed solutions of tetramethyl **4.5** in phosphate buffered D_2O at pH 4, 7 and 10 by ^1H NMR. Over a period of 7 days, we saw no change in the NMR spectra that would indicate degradation over time. With that information, we felt confident to proceed with binding studies at pH 7.

As an initial study, we looked at the binding of tetramethyl host **4.5** to AChCl. As **4.5** was the more hydrophobic of the two molecules, we suspected that whatever binding effect we would see would be strongest for **4.5**. As in Chapter 2, ^1H NMR titration was the main tool used to evaluate host performance. Titrations were carried out in 50 mM phosphate buffered D_2O at pH 7.0 (pD 7.4). The pKa of the tertiary amines of the Tröger's base scaffold (pKa_1 = 3.25)[167,178] is very different from those of other trialkylamines (for example, triethylammonium, pKa = 10.7).[179] We felt it reasonable to assume that at the pH of our study (pH 7), these amines would remain in an unprotonated state.

Host **4.5** was evaluated for its tendency to self-associate. Host dimerization constants were determined by reverse titrations using 20 mM stock solution. Over the concentration range of 0.3 – 14 mM, minimal concentration-dependent chemical shifts

were observed, indicating little to no host association present in this system (the calculated dimerization constant arising from curve fitting of these shifts was 1.0 ± 0.5 M^{-1}). In contrast, host **2.1** exhibited a much larger and significant dimerization constant (330 M^{-1}).

Host **4.5** was then titrated against AChCl as a representative of the R-NMe_3^+ motif. Similar to the dimerization case, the magnitude of the chemical shift changes upon guest addition were minimal, though a large excess of guest was added. The calculated binding constant for AChCl was 11.0 ± 0.5 M^{-1}. The minor chemical shift changes for this host are not necessarily unexpected, even if strong binding were present. In this case, $\Delta\delta_{obs}$ comes from the more subtle, electrostatic interaction of the incoming cation with an aromatic surface, rather than a large host conformational change (as seen in Chapter 2) or an observed proton involved directly in an interaction (such as in hydrogen bonding motifs). Further titrations with these types of rigid hosts should be host into guest, rather than guest into host. That way, binding would be apparent from changes on the guest protons, which should more strongly respond to encapsulation and/or a new chemical environment upon binding. The drawback of this method is that an excess of host molecule is needed and in particular, a large excess if weak binding is present.

4.3.3 Aromatic surface alone does not induce strong binding

While mono-Tröger's base **4.5** was able to remain soluble (not dimerized) in water and exhibited some binding behaviour towards AChCl, the fact remains that a single Tröger's base offers only a shallow aromatic surface rather than a well defined pocket. Similar weak binding results were observed by Dougherty in evaluating the "half-molecule" of his ethanoanthracene cyclophane.[64] It is clear that the binding of our amino acid targets requires a more defined space (such as a cage or cleft). With that, we set out to incorporate multiple Tröger's bases into a single host structure to better mimic a typical aromatic cage.

4.4 A rigid, dual Tröger's base structure

With one copy of a Tröger's base providing not enough of an aromatic/hydrophobic cavity to bind a cation such as trimethyllysine, we sought a design that would incorporate multiple copies of Tröger's bases. Ideally, they would be held separate from each other by a rigid linker that would prevent a hydrophobic collapse and enforce a cage-like structure. A similar idea to this was used in the synthesis of Wilcox's ethanoanthracene Tröger's base cyclophane (**3.10**). Our proposed host, shown in Figure

4.1, has two Tröger's bases linked through *meta* positions of an aromatic ring. Substitutions in positions *ortho* to the biaryl ring connections (R_1) were intended to enforce the perpendicular positioning of the connected aromatic rings through steric interference, acting as an internal push towards the desired conformation. Specific functionality, either on the Tröger's bases themselves (at R_2) or on the spacing ring, could be used to bring about water-solubility.

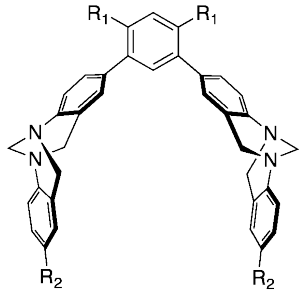
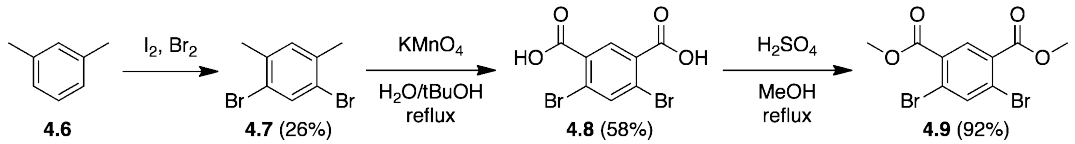


Figure 4.1 Generic rigid, dual Tröger's base host

4.4.1 Synthesis and water-solubilization attempts

With the Tröger's base building blocks suitable for cross-coupling reactions already in hand from Chapter 3, we turned to the synthesis of the linking aromatic ring. We envisioned using palladium-catalyzed Suzuki coupling methods to generate the C-C bond between our functionalized Tröger's bases and the central aromatic ring. As such, we needed an aromatic ring with the correct halogen substitution pattern to complement the boronic acid/ester Tröger's bases from Chapter 3. Additionally, we wanted to include carboxylates for solubility on this ring. Incorporating them at positions *ortho* to the halogen substitution would give us both steric interference for geometry, and charge for solubility following deprotonation.

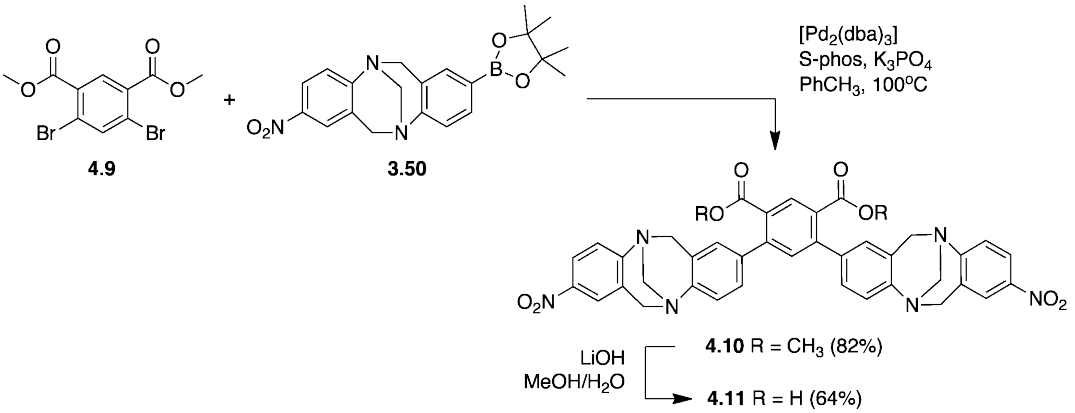
The synthesis was accomplished in three steps from commercially available *m*-xylene **4.6** (Scheme 4.2). Exact literature precedent was found for the dibromination of *m*-xylene that gave **4.7** in 26% yield (11 g) after recrystalization from absolute ethanol.[180] Conversion of the methyl groups into carboxylic acids was accomplished by oxidation with potassium permanganate yielding **4.8** in decent yield (58%). Literature procedure[181] (and advice from another research group)[182] for the oxidation on a similar compound were used. Finally, the acids were esterified to methyl esters under Fischer esterification conditions to generate **4.9**.



Scheme 4.2 Synthesis of aromatic core **4.9**

From this point, the aromatic spacer **4.9** was coupled to nitro-boronic pinacol ester Tröger's base **3.50** under Suzuki conditions using tris(dibenzylideneacetone)dipalladium(0) ($[Pd_2(dba)_3]$), the ligand 2-dicyclohexylphosphino-2',6'-dimethoxybiphenyl (also known as SPhos), and K_3PO_4 in toluene (Scheme 4.3).[158] Through optimization, it was found that this reaction worked best on small scale (100 mg at a time) and when the base had been freshly ground to a fine powder before use. We were able to obtain up to an 82% yield of **4.10**, including purification by column chromatography.

In our enthusiasm to study molecules of this type, we proceeded without further host functionalization. Basic hydrolysis with LiOH in MeOH/H_2O mixtures were sufficient to removed the methyl esters and leave us with the free carboxylic acid handles in **4.11**. Further treatment with two equivalents of NaOMe was used to convert both acids to the carboxylate salts for study.

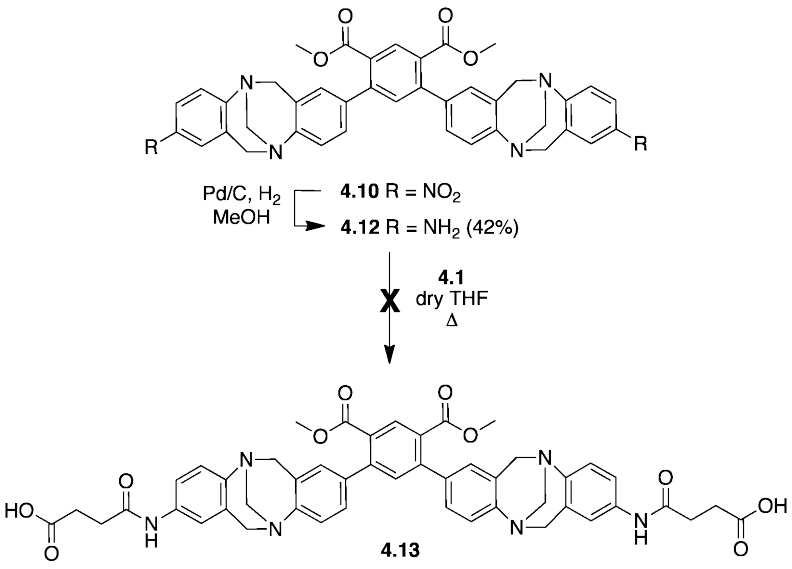


Scheme 4.3 Synthesis of dual Tröger's base host **4.11**

Unfortunately, the two negative charges installed on the linking spacer were insufficient at making **4.11** water soluble. This precluded us from being able to study this molecule under aqueous conditions, where weak interactions of interest (especially the hydrophobic effect) would play a large role. We performed some limited exploration in MeOD and MeOD/D_2O mixtures. Host chemical shift changes during ^1H NMR titrations with Kme3 in unbuffered MeOD resulted in likely proton transfer to the host Tröger's

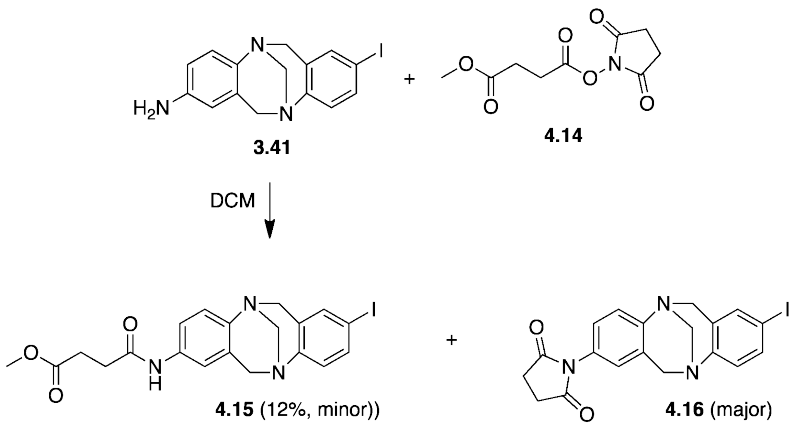
base rather than guest binding. To eliminate this, we repeated the titration using AChCl as a representative guest and were encouraged to see concentration dependent shift changes. The host signal that showed the greatest chemical shift was the spacing ring proton pointing directly into the ideal binding cavity. Subsequently, we were able to solublize **4.11** in a 3:1 mixture of MeOD/40 mM phosphate buffered D_2O. The resulting host chemical shift changes gave a binding constant of 30 ± 10 M^{-1}. While this was a start, the insolubility of this host in pure water limited the further use of this molecule.

Unwilling to abandon the scaffold at this stage, we sought to include additional negative charges on the Tröger's base itself as a way to increase the polar nature of the overall molecule. We originally thought to install pendent water-solublizing groups to the end of the Tröger's base, as had been done in Section 4.3, by reduction of the nitro groups followed by reaction with succinic anhydride to generate two additional free carboxylic acids. We first attempted reduction using the previously successful nitro conditions of $Fe_{(s)}$ and AcOH in refluxing EtOH.[162] While this produced a more polar spot by TLC of the reaction, upon work up we were unable to isolate or extract back that spot from the aqueous fractions. Reduction using Pd/C and H_2 in MeOH however, did produce diamino **4.12** in 42% yield (Scheme 4.4). We then attempted to functionalize **4.12** by reaction with succinic anhydride **4.1** in refluxing THF as had previously been done on **4.2** and **4.3**. Unfortunately, no conclusive evidence for the formation of product **4.13** was seen by ^{1}H NMR (addition of four methylenes) nor was TLC useful due to the high polarity of both starting material and product. This method was consequently abandoned.



Scheme 4.4 Attempted functionalization of Tröger's base host **4.10**

Our second thought was that the water-solublizing group could be introduced at an earlier stage of the Tröger's base synthesis, prior to coupling onto the aromatic core **4.9**. The functionalized Tröger's base would still have a number of further transformations for solubility and ease of purification, so it was preferable to have the pendent acids protected as esters until the final step, as was done with the hosts in Chapter 2. To that end, instead of opening succinic anhydride (which would leave free carboxylates which would need protecting), amino-iodo Tröger's base **3.41** was reacted with the methyl succinyl ester of succinic acid **4.14** in DCM (Scheme 4.5). Unfortunately, the aromatic amine proved nucleophilic enough to generate a mixture of products – both the desired structure **4.15** (in an isolated 12% yield), and a cyclized imide **4.16** (co-eluted with **4.15**). After the first amide bond formation, cyclization of the amide nitrogen onto the pendent methyl ester generated a more favourable five-membered cyclic imide in higher yield.



Scheme 4.5 Earlier introduction of water-soluble groups into Tröger's base

We thought that we might use the preference of this system to form this cyclic product to our advantage. Five-membered cyclic imides can be ring opened under mild basic conditions,[183] which would eventually generate our desired final structure. Major product **4.16** was successfully boronated, but further reaction to couple **4.16** to **4.9** proved unsuccessful. Because we were unable to install our desired water solubilizing groups, and because all of these synthetic routes were extremely long and low-yielding, efforts to make and derivatize this di-Tröger's scaffold were halted at this point.

4.4.2 Conclusions

Host **4.11** was designed to present aromatic surface area in a cage-like position for cation binding by a) restricting the host's ability to suffer from hydrophobic-driven aromatic clustering through the use of inflexible Tröger's bases and b) taking advantage of steric interference to further influence a cage preference.

What we found was that **4.11** binds only weakly to AChCl in MeOD/D_2O solutions. One reason for this lack of affinity would have to do with the fact that hydrophobic binding surfaces of host and guest are much better solvated in MeOD/D_2O mixtures than in pure D_2O. As such, there is no additional driving force from the hydrophobic effect to encourage binding (unlike in Chapter 2). Our attempts to add additional negative charges (through the addition of carboxylates) at the nitro groups of **4.10** proved synthetically unsuccessful.

A second reason proposed for poor binding constants (in the absence of a pure water solution) was structural. While we eliminated many degrees of freedom from this host, rotation about the C-C single bond connecting the biaryl rings was still possible. Through this, each Tröger's base arm could exist in positions ranging from the ideal

"Tröger's base-in" conformation to the opposite extreme of "Tröger's base-out" (Figure 4.2) while still satisfying the slightly offset preference of the biaryl bond. Both arms have this freedom of rotation leading to three possible conformations: in-in (the ideal cage), in-out (a half cage) and out-out (no cage or W-shaped). With this free rotation, it is not reasonable to expect that the lowest energy preferred geometry of this system would be one where both Tröger's bases are in. Instead, it is likely a mixture of all three conformations in solution, two of which do not have the preferred cage for binding. Converting to the ideal cage in those cases would require a large conformational change (180 degree rotation) through a sterically unfavorable conformation. One solution to this problem would be to tether the Tröger's bases together into a cyclophane such that they would be unable to twist into any "Tröger's base-out" shapes.

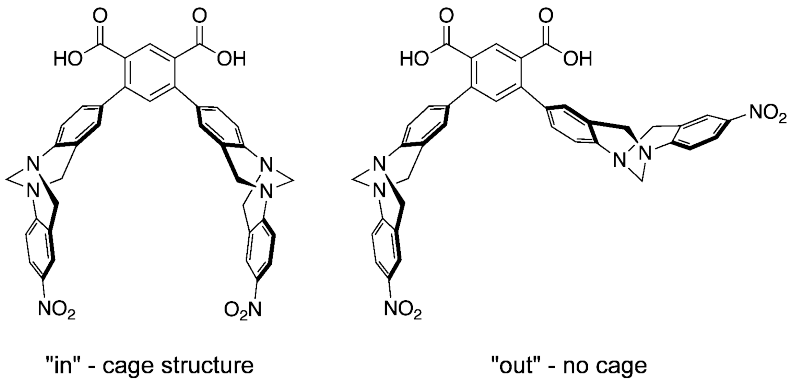


Figure 4.2 Possible dual-Tröger's base conformation

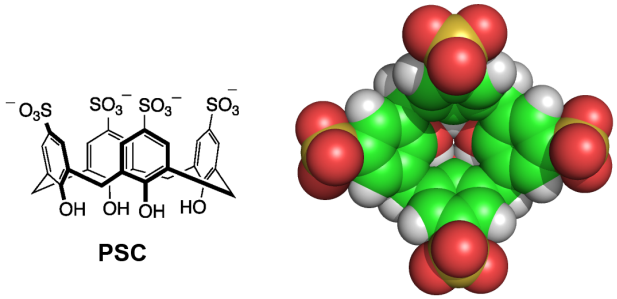
4.5 Tröger's base-calixarene hybrids

With the goal of trying to incorporate the rigid aromatic surface of Tröger's base into aromatic cage binding motifs, we looked to try to combine the Tröger's base building blocks from Chapter 3 with other known cage motifs. We decided to see how the addition of a Tröger's base would affect the trimethyllysine binding ability of *p*-sulfonatocalix[4]arene (PSC) **1.1**.

4.5.1 PSC is a strong binder of Kme3

As first mentioned in Chapter 1, PSC itself is a good host for methyl-lysines, reported recently to out-compete an intact native aromatic cage (chromodomain of CBX7) for binding to trimethyllysine (in an H3K27me3 peptide).[63] PSC is a successful host because it displays many of the properties necessary for an aromatic cage. It has four sulfonate groups on the upper rim that are important both for water solubility and as a source of negative charge for an electrostatic interaction with a cation. It is a

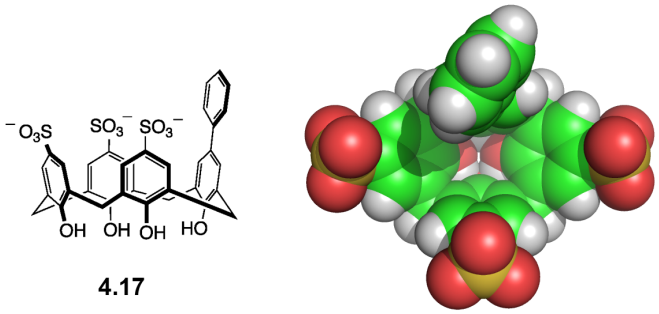
macrocycle, which helps to present the four aromatic rings face-on to an approaching cation. Under physiological conditions (pH 7.4), one of the four hydroxyl groups of the lower rim is expected to be deprotonated,[184] leading to a hydrogen bonding network that further helps to stabilize the open conformation cone conformation of the molecule.[185] It also has multiple aromatic rings, which present as an electron-rich, hydrophobic surface for both cation-pi interactions and a favourable hydrophobic effect.



PSC

Figure 4.3 PSC. Spartan molecules at PM3 level.[13]

The success of such a simple molecule to behave very similarly to a native aromatic cage has led to a number of attempts at synthetic modifications designed to improve upon the original binding ability of PSC. One of the most promising modifications was the discovery that substituting a single sulfonate of PSC with a phenyl ring (host **4.17**, Figure 4.4) doubled the binding affinity to trimethyllysine (to K_{assoc} 64000 M^{-1}, or K_d 16 μM), and increased selectivity over unmethylated lysine to 150 fold.[177] Though it removed a source of negative charge and thus decreased the electrostatic interaction between the cation and host, this substitution introduced a hydrophobic pi surface (and thus additional cation-pi and hydrophobic interactions), which far outperformed the "stronger" electrostatic interaction. Additional modifications of the phenyl ring that altered the electron-richness of the appended ring (*para*-substitutions of nitro and amino groups, for example) or that introduced additional negative charge (such as carboxylic acid), were all found to weaken the binding to Kme3. In addition, the mono-substitution of different functional groups (other than phenyl) designed to increase/decrease the electronic character of the base calixarene ring (and thus the cation-pi effect) were found to have no influence on the strength of the binding interaction.

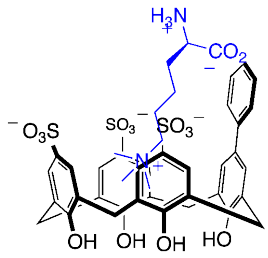


4.17

Figure 4.4 Phenyl modified PSC, **4.17**. Spartan molecules at PM3 level.[13]

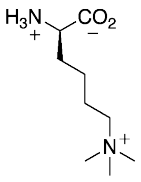
The beneficial result of adding an additional phenyl group to PSC was concluded to be primarily due to hydrophobic interactions. While the cationic head group of trimethyllysine is still engaged strongly by the calixarene portion of the molecule, it is believed that the lysine side chain methylenes are able to make contact with the face of the aromatic ring (Figure 4.5, a) incorporating both CH-pi interactions and additional hydrophobic contacts to improve binding. This is reminiscent of the conclusions of Chapter 2, in that stretches of hydrophobic regions (like the side chain methylenes of amino acids) are important to the binding and selectivity of protein-protein interactions. Further evidence of this is shown by the chemical shift changes of the protons of trimethyllysine upon binding to PSC vs. **4.17** (phenyl). Trimethyllysine clearly shows increased interaction of protons further away from the ammonium head group (such as α-CH) upon interaction with phenyl PSC **4.17**. (Figure 4.5, b).[62,177]

a)



4.17

b)



	PSC	4.17
α	0.03	0.31
ε	1.33	1.38
N-CH_3	2.24	2.16

Figure 4.5 Phenyl PSC **4.17** with trimethyllysine binding. a) model, b) chemical shift changes (in ppm).

With this in mind, we were curious as to what effect incorporating a Tröger's base onto the rim of a calixarene to form a Tröger's base calix-hybrid (Figure 4.6) would have on binding. The additional hydrophobic surface area provided by the lower ring of the Tröger's base would be similar to that provided by the single phenyl ring in **4.17**. In addition, the bent geometry of the Tröger's base could provide some interesting host

selectivity, especially in terms of the other methylation state of lysine (mono- and di-methylated) as well as the methylation states of the differently shaped amino acid arginine.

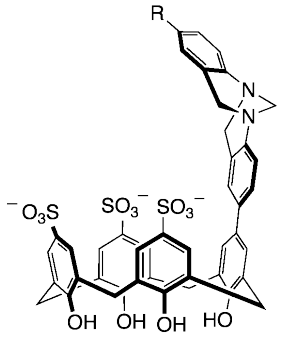
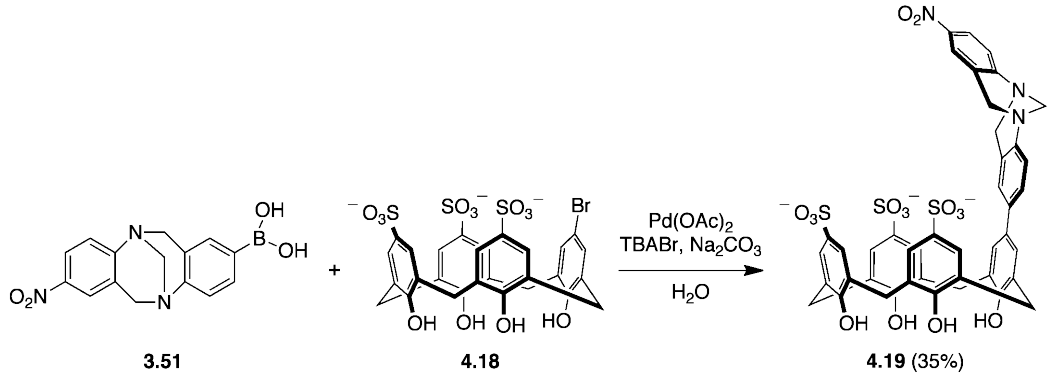


Figure 4.6 Generic Tröger's base-calix[4]arene hybrid

4.5.2 Synthesis of Tröger's base hybrid

The paper that presented phenyl-calixarene **4.17** also gave a straight-forward method for the synthesis of other aryl-substituted calixarenes. All revolved around a particular precursor, bromo-trisulfonate calixarene **4.18**, that could be used as the aryl halide partner in a Suzuki carbon-carbon bond forming reaction.[177] Microwave coupling conditions were used to make nitro-Tröger's base-calixarene **4.19** (Scheme 4.6) using nitro-boronic acid Tröger's base **3.51**. Synthesis using boronic ester version (**3.50**) was unsuccessful, due to the low water solublity of the boronic ester. Due to the highly water-soluble nature of the final host, purification of the crude microwave reaction was done by high pressure liquid chromatography (HPLC). Phenyl-calixarene **4.17** was also re-made by this method in 50% yield for use in further studies. Note that while Tröger's base is a chiral molecule, all synthesis was done using racemic compound mixtures, and the resulting products are also racemic.



3.51 **4.18** **4.19** (35%)

Scheme 4.6 Synthesis of nitro-Tröger's base-calixarene hybrid **4.19**

4.5.3 Solution studies - Binding to lysine via ^{1}H NMR titration

The synthesized hosts were all water soluble as isolated. While the initial studies of phenyl calixarene **4.17** focused strictly on the host selectivity of trimethyllysine over lysine,[177] we were curious to explore the other methylation states of lysine, as well as methylated states of arginine, and to do a more thorough analysis of the contributions of cation-pi, electrostatics, host preorganization and the hydrophobic effect. We performed ^{1}H NMR titrations in buffered D_2O (40 mM phosphate buffer, pD 7.4) using all lysine and arginine methylated compounds as guests. Titrations were performed in reverse with a 5 mM solution of host being added to a 0.5 mM solution of guest. Consequently, protons on the amino acid guests were followed including the N-methyl signal. All proton signals experienced upfield shifting, indicative of shielding of the cation head group within a highly aromatic region. Non-linear least squares fitting was used to fit the binding data to a 1:1 binding isotherm using HypNMR[97] as described in Chapter 2. Table 4.1 summarizes the binding constants determined in water for PSC, nitro-Tröger's base-calixarene **4.19** as well as bromo-calixarene **4.18** and phenyl-calixarene **4.17**.

Table 4.1 Binding studies of calixarene derivatives to methylation states of lysine

		K_{assoc} [a]			
Entry	Host	K	Kme	Kme2	Kme3
1	PSC[b]	520 ± 300	4000 ± 3000	16200 ± 4300	37000 ± 18000
2	**4.18** (bromo)	480 ± 15[c] (440 ± 110)	1020 ± 10	5070 ± 120	4000 ± 500[c] (3900 ± 100)
3	**4.17** (phenyl)	460 ± 6[c] (420 ± 50)	1360 ± 20	10600 ± 320	67500 ± 2700[c] (64000 ± 13000)
4	**4.19** (TB)	240 ± 10	1580 ± 50	12200 ± 600	13900 ± 1200

[a] Determined by ^{1}H NMR titrations in 40 mM phosphate buffered D_2O, pH 7.0 = pD 7.4. [b] Published data reproduced for comparison (determined by ^{1}H NMR titrations in 40 mM phosphate buffered D_2O, pH 7.0 = pD 7.4).[62] [c] Values determined via current method (HypNMR) from raw data for consistency. Published values in parentheses.[177]

While we were interested in comparing the binding affinities for the new hybrid calixarene **4.19** with the parent molecule PSC, we felt it would not be a direct "apples to apples" comparison. Multiple structural changes had been made that would affect the electrostatic interactions, hydrophobic contribution and molecular shape. To understand better the progression in binding, we also considered the binding of bromo-calixarene **4.18** (one less negative charge) and phenyl-calixarene **4.17** (one less negative plus additional hydrophobic surface area) as "intermediate" compounds (Figure 4.7).

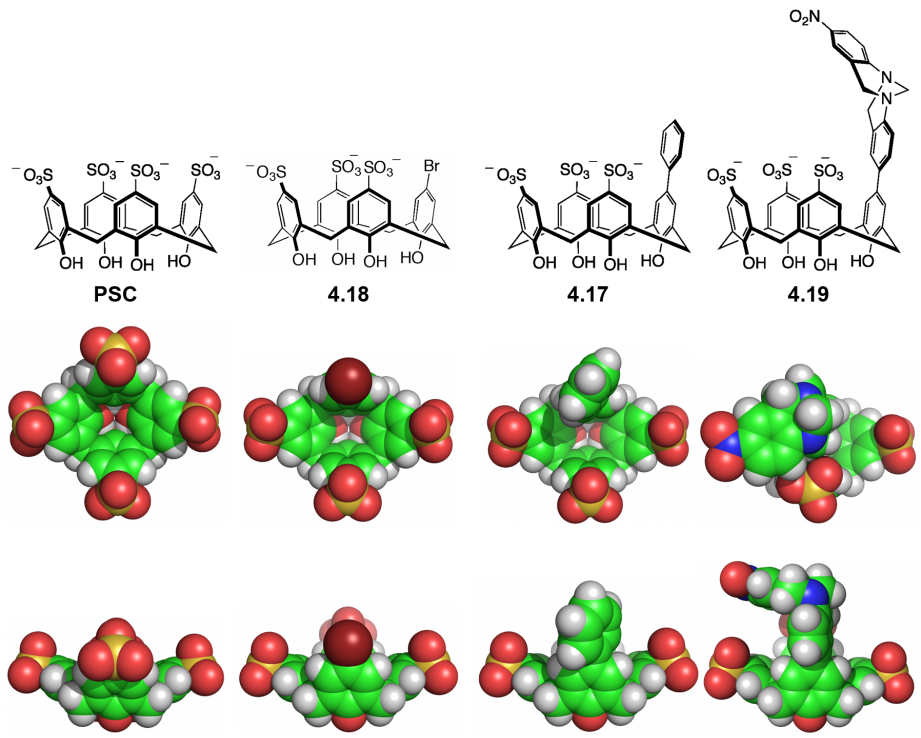


Figure 4.7 Calixarene hosts under study. Spartan molecules at PM3 level.[13]

Removing one sulfonate from the upper rim of the calixarene removes one negative charge from the interaction with an incoming cation. This was shown to be unfavorable for binding with all the lysines tested (Table 4.1, entries 1 and 2). The difference in binding was minimal for lysine, but became more pronounced with increasing step-wise methylation of the guest, leading to a nearly 10-fold decrease in binding with Kme3 compared to parent PSC. This result demonstrated that electrostatic interactions, even highly solvated ones such as those of these exposed sulfates, can have a significant influence when interacting with a poorly solvated cation in water.[78]

The second comparison we made was with replacement of the neutral bromo group with a hydrophobic phenyl (Table 4.1, entries 2 and 3). Little difference in binding was observed between the two hosts for K and Kme. However, significant differences were observed with the more hydrophobic Kme2 and Kme3. Binding of phenyl **4.17** to Kme2 was approximately double that of bromo **4.18**, but still less that parent PSC. With Kme3, the binding affinity increased to 15 times that of bromo **4.18**, *and* nearly double that of PSC. It appears that in the case of the very hydrophobic Kme3, the loss of an electrostatic interaction can be more than compensated for by the addition of a hydrophobic surface. This was not the case for the less hydrophobic cations. Binding to K and Kme was significantly affected by the sulfonate loss and their affinities remained

less than that of PSC. Kme2 appears on the edge, where the phenyl aromatic ring is able to partially rescue binding affinity, while Kme3 is significantly positively affected.

We were able to make a more direct comparison now between phenyl calix **4.17** and our host of interest, nitro-Tröger's base-calixarene **4.19**. Both **4.17** and **4.19** have free rotation about the single bond connecting the added aromatic groups to the calixarene structure and the preference of the rings would be to adopt a slightly perpendicular alignment to avoid steric clashes with *ortho* protons. In the case of phenyl **4.17**, rotation does not matter as either side will present the same surface and geometry. This is not the case for **4.19** – two extreme geometries are possible; one with the second Tröger's base ring over the calixarene cavity, and one where it is pointed outwards into solution (Figure 4.8). While outwards, the cavity-side aromatic surface presented would be much like that of phenyl **4.17**. While inward, a very different host shape would be present.

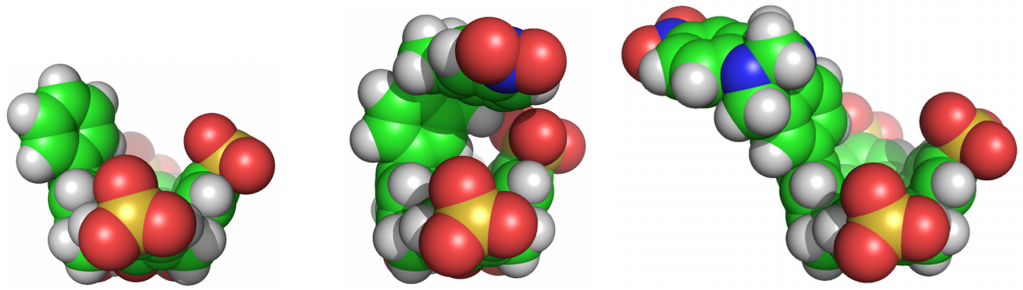


Figure 4.8 Comparison of possible host geometries, **4.17** and **4.19**. Spartan molecules at PM3 level.[13]

The solution binding data reveals that **4.17** and **4.19** have similar binding affinities for K, Kme and Kme2 (Table 4.1, entries 3 and 4). This would indicate that the unique structure of **4.19** did not significantly affect the ability of these cations to approach the binding pocket and bind in their preferred orientation. However, binding to Kme3 was significantly decreased compared to phenyl **4.17**. The large size of Kme3 allows it to completely fill the binding pocket of **4.17**, making as many favourable contacts as possible. The altered binding pocket of **4.19** was not able to optimally accommodate the size and shape of Kme3, and thus a decrease in binding compared to **4.17**.

Overall, host **4.19** proved to be selective for Kme3 over K by a factor of 58, with an absolute binding affinity five times less than phenyl PSC **4.17**. The binding cavity provided by **4.19** was not optimal for binding Kme3 (unlike **4.17**) but did prove somewhat

selective for Kme2/3 over K and Kme. As synthetic receptors that specifically target the lower methylation states of lysine currently do not exist, this is a good example of using geometric constraints as one tool in achieving selectivity. This is similar to what native cages do to be selective for Kme/Kme2 by sterically crowding out Kme3.[31,37,41]

4.5.4 Solution studies - Binding to arginine via ^{1}H NMR titration

We also took the opportunity to use the four hosts studied above and examine their binding to methyl-arginines, with the same electrostatic, hydrophobic and geometric considerations.. Titrations were again performed as host (5 mM) into guest (0.5 mM) which allowed us to track chemical shift changes on the guest. Strong upfield shifts of the N-methyl and lower residues protons again indicated that the cationic head of the guest was incorporated into the aromatic cages of the various hosts. These solution titrations are one of the first reported studies of methyl arginine binding by synthetic receptors.[62]

Table 4.2 Binding studies of calixarene derivatives to methylation states of arginine

			K_{assoc} [a]		
Entry	Host	R	Rme	s-Rme2	a-Rme2
1	PSC[b]	330 ± 260	760 ± 330	n.d.[c]	1100 ± 460
2	**4.18** (bromo)	320 ± 90	660 ± 15	610 ± 10	2240 ± 20
3	**4.17** (phenyl)	380 ± 5	1020 ± 15	840 ± 15	4200 ± 120
4	**4.19** (TB)	330 ± 10	1120 ± 30	900 ± 30	5800 ± 300

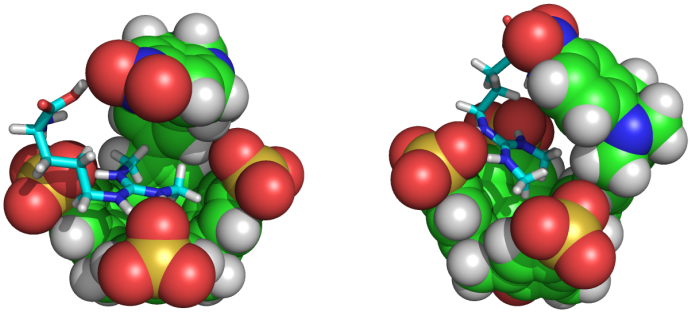
[a] Determined by ^{1}H NMR titrations in 40 mM phosphate buffered D_2O, pH 7.0 = pD 7.4. [b] Published data reproduced for comparison. [c] s-Rme2 gave a complex NMR curve that could not be fit to any simple 1:1, 2:1, or 1:2 binding isotherm.

Similar solution-phase binding constants for arginine were found with all four host studied (approximately 350 M^{-1}). Compared to the other guests in this series, unmethylated arginine is the most well-solvated, and the least likely to be influenced by the hydrophobic effect to make hydrophobic contacts with a poorly solvated cavity (such as the aromatic rich calixarenes). However, other water-soluble arginine-binding receptors are known, and our binding constants are low, relatively speaking. The aromatic cage-like hosts of Dougherty (**1.4**) [78] and Klarner (**1.12**),[81] presented in Chapter 1, have binding affinities for arginine of 5000 ± 200 M^{-1} (10 mM borate buffered D_2O, pH 8.6, H-Arg-NH_2) and 1800 ± 720 M^{-1} (25 mM phosphate buffered D_2O, pH 7.0, Ts-Arg-OEt), respectively. Other groups have recorded PSC to have an association constant of 1520 ± 90 M^{-1} with arginine by isothermal titration calorimetry (10 mM phosphate buffer, pH).[186] However, given that the interaction with arginine has a strong electrostatic

component, under our slightly more competitive experimental conditions (40 mM phosphate buffered D_2O, pH 7), a decreased relative affinity is reasonable.

As guest methylation increases from zero to one to two, we again see the general trend of increasing binding constants due to an increased role of the hydrophobic effect. There is a better match between the hydrophobicity of the guest and host, leading them to more strongly associate with each other (as seen in Chapter 2). For Rme and a-Rme2, there is a distinct difference between the hosts bearing hydrophobic area (**4.17** and **4.19**) to those without. One interesting observation is that despite the similarity between a- and s-Rme2, a-Rme2 is bound more strongly than s-Rme2 in all cases. There does not appear to be any significant solvation difference between the two guests[55,187] and both guests have three sites for potential hydrogen bonding (to the sulfonates, for example) though at different positions. The difference in binding must be due to the distinct shapes of the guests, and how they fit with each host.

These binding studies also show that nitro-Tröger's base-calixarene **4.19** is a better host for methylated arginines than the phenyl-only version **4.17**, unlike the case for lysine. We believe this is due to the different host shape offered by the perpendicular Tröger's base in **4.19**. Where as lysine is spherical, the cationic head group of arginine is planer. This flat cation could potentially not fit as well as lysine into the spherical aromatic cavity provided by the calix[4]arene hosts. Increased methylation of arginine extends this flat surface, making the guest larger and host-guest fit worse, while only minimally compensating for size through an increased hydrophobic effect. The inclusion of additional aromatic surface area in **4.17** allows the arginine guests to invoke another weak intermolecular contact (pi-pi interactions) and increase their binding affinity. However, this still leaves most of the hydrophobic guest exposed to solvent. The second aromatic ring introduced by the Tröger's base in **4.19** creates more of a flat cleft into which a flat guest (like arginine) could fit and still make pi-pi contacts with the second ring (Figure 4.9), similar to the phenylalanine clamp in WDRD5 (Figure 1.12). Arginine methylation in this case would increase the hydrophobic driving force to form such a complex, as well as not disrupt the binding (since the cation remains flat). Tröger's base **4.19** is better able to accommodate this shape, giving it a binding advantage.



4.19 (front) **4.19** (side)

Figure 4.9 The arginine cleft provided by nitro-Tröger's base-calixarene **4.19**. Spartan molecules at PM3 level.[13]

If the stacked binding motif of **4.19** is true for arginines, then binding affinities should be affected by changes to the electronic nature of the second ring unlike the phenyl PSC derivatives, where binding was mostly hydrophobically driven. For binding a cation, *p*-nitro substitution is not optimal as this removes electron density from the ring making it less attractive. However, modification of that position to include electron donating groups (such as $–NH_2$, -OMe) through modular use of other Tröger's base derivatives would explore the extent to which pi-pi stacking plays a role in the binding of methylated arginines.

4.6 Conclusions

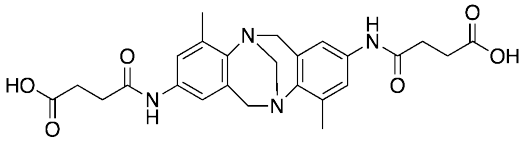
In this chapter, we explored three distinct ways that the rigid, aromatic nature of Tröger's bases could be exploited in the construction of synthetic aromatic cages. While our early attempts suffered from first a lack of overall cage structure, and second water solubility issues, these efforts highlight the difficulty in trying to design and synthesize molecules to perform a specific function. Combining lessons from both the mono- and di-Tröger's base molecules, we were able to synthesize a successful Tröger's base host that included both an aromatic cavity and the necessary water solubility. Binding studies with host **4.19** showed some expected behaviour of aromatic cages – such as increased binding with increased guest methylation – and also some new observations about the size and shape complement of host and guest. Further exploration of the derivatized aromatic ring of **4.19** will help to confirm the importance of pi-stacking and cation-pi interactions with the binding of methylated arginines especially.

4.7 Experimental Section

4.7.1 General considerations

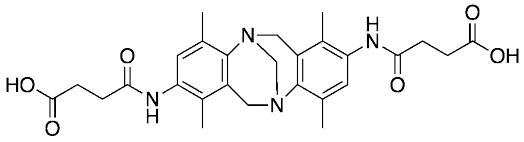
Solvents and reagents were used as obtained from Sigma-Aldrich. Proton (^{1}H) NMR and carbon (^{13}C) NMR spectra were recorded on a Brüker AC300 (300 MHz), Brüker Avance 360 (360 MHz) or Brüker Avance 500 (500 MHz) spectrometer as indicated. Chemical shifts (δ) are given in parts per million (ppm) relative to TMS and referenced to residual protonated solvent ($CHCl_3$: δH 7.26 ppm, δC 77.16 ppm; DMSO δH 2.50 ppm, δC 39.52 ppm; DOH: δH 4.79 ppm).[109] *J* values are given in Hz. Abbreviations used are s (singlet), d (doublet), t (triplet), q (quartet), and m (multiplet). Infrared spectra (IR) were measured on a Perkin Elmer 1000 FT-IR spectrometer. Melting points were obtained using a Gallenkamp Melting Point Apparatus and are uncorrected. Accurate mass measurements were performed on two different ESI-MS instruments (HR-ESI-MS) and are labeled accordingly. Accurate mass measurements labeled with "QTof" were done at the University of Victoria Chemistry mass spectrometry facility on a Q-TOF II system by MicroMass. Samples of 1mg/ml diluted 1:100 in MeOH or ACN were directly infused at a 5-20ul/min flow rate through an ESI source. Spectra obtained at a scanning tine of 1 second, a 100-2000 range with 8,000 resolution and the higher of 3mDa or 5ppm accuracy. Accurate mass measurements labeled with "Orbitrap" were done at the UVic Proteomics Center on an Orbitrap Velos system by Thermo scientific. Samples of 1mg/ml diluted 1:100 in MeOH or CH_3CN were directly infused at a 5-10ul/min flow rate through an ESI source. Spectra obtained at about 1Hz scanning rate 100-2000 range with 60,000 resolution and less than 1ppm accuracy in most cases.

4.7.2 Synthetic procedures

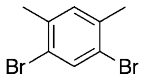


Compound **4.2**. Compound **3.31** (3.01 g, 10.7 mmol) and succinic anhydride **4.1** (2.14 g, 21.4 mmol) were dissolved in anhydrous THF (215 mL) and heated to reflux under Ar. After 17 hours, the reaction was considered complete by TLC (EtOAc) and evaporated to dryness. The resulting crude solid was heated to reflux in DCM and filtered to give **4.2** (4.15 g, 81% yield) as a free flowing solid. Mp: 190 – 196 °C (decomp.). IR (KBr pellet) ν (cm^{-1}): 3112br, 2954m, 2883m, 1715s, 1669s, 1610m, 1480s, 1413m, 1319w, 1232s, 734m. ^{1}H NMR (DMSO, 300 MHz): δ 2.30 (s, 6H), 2.46 (s, 8H), 3.83 (d, *J* = 17.0, 2H),

4.17 (s, 2H), 4.43 (d, *J* = 17.0, 2H), 7.09 (d, *J* = 2.1, 2H), 7.15 (d, *J* = 2.1, 2H), 9.72 (s, 2H), 12.07 (br s, 2H). ^{13}C NMR (DMSO, 90 MHz): δ 16.9, 28.9, 30.9, 54.6, 67.2, 114.7, 119.6, 128.1, 132.4, 134.7, 141.0, 169.6, 173.8. HR-EI-MS: *m/z* $[M+Na]^+$ calculated for $C_{25}H_{28}N_4O_6Na$: 503.19011; found 503.18996 (Orbitrap).

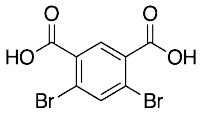


Compound **4.3**. Compound **3.32** (50 mg, 0.16 mmol) and succinic anhydride **4.1** (31 mg, 0.31 mmol) were heated to reflux under Ar in a 50 mL round bottom flask with dry THF (4 mL) for 18 hours until complete by TLC. After removal of the solvent by rotary evaporation, the remaining residue was suspended in DCM and filtered to isolate. The product **4.3** (76 mg, 96% yield) was obtained as a pale orange solid. Synthetic conditions for this reaction were adapted from published reference.[188] Mp: 180 – 184 °C (decomp.). IR (KBr pellet) ν (cm^{-1}): 3123br, 2944m, 2880m, 1712s, 1672s, 1612m, 1478s, 1412m, 1316w, 1223s, 945m, 734m. ^{1}H NMR (DMSO, 300 MHz): δ 1.85 (s, 6H), 2.32 (s, 6H), 3.84 (d, *J* = 17.1, 2H), 4.15 (s, 2H), 4.38 (d, *J* = 17.0, 2H), 6.95 (s, 2H), 9.21 (s, 2H, amides), 12.04 (br s, 2H, acids). (CH_2CH_2 signals missing – under DMSO). ^{1}H NMR (MeOH, 300 MHz): δ 1.91 (s, 6H), 2.39 (s, 6H), 2.64 (s, 8H), 3.95 (d, *J* = 17.0, 2H), 4.23 (s, 2H), 4.42 (d, *J* = 17.0, 2H), 6.95 (s, 2H). (amide and acid peaks exchanged away). ^{13}C NMR (DMSO, 90 MHz): δ 12.1, 16.6, 29.1, 30.3, 53.5, 65.6, 126.3, 126.4, 127.6, 129.3, 131.7, 143.3, 170.0, 173.8. HR-EI-MS: *m/z* $[M+Na]^+$ calculated for $C_{27}H_{32}N_4O_6Na$: 531.22141; found 531.21915 (Orbitrap).

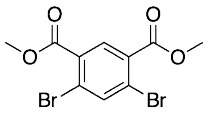


1,5-dibromo-2,4-dimethylbenzene **4.7**. Bromine (54.7 g, 0.34 mol) was added drop-wise over 1 hour in the absence of light to an ice-cooled solution of iodine (0.20 g, 0.79 mmol) in neat *m*-xylene **4.6** (16.95 g, 0.16 mol). After 16 hours at room temperature, $KOH_{(aq)}$ (20%, 100 mL) was added slowly. The mixture was stirred and prodded under slight warming until the yellow colour was gone from the reaction before being allowed to cool. The aqueous layer was decanted, and the remaining solids washed with water (2 x 100 mL). Recrystalization from absolute EtOH gave **4.7** (11.0 g, 26%) as a white solid. Synthetic conditions for exact compound were found in published procedure.[189]

Additional characterization data was found here.[190] Characterization data of synthesized sample was found to match published spectra for this known compound. ^{1}H NMR ($CDCl_3$, 300 MHz): δ 2.31 (s, 6H), 7.10 (s, 1H), 7.68 (s, 1H). ^{13}C NMR ($CDCl_3$, 75 MHz): δ 22.4, 122.2, 132.8, 135.1, 137.0.

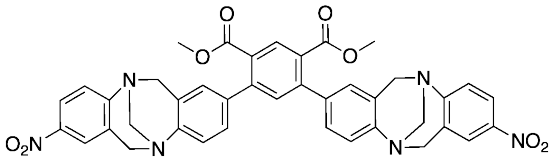


4,6-dibromo-1,3-benzenedicarboxylic acid **4.8**. Compound **4.7** (5.04 g, 19.0 mmol) was suspended in a 1:1 solution of H_2O and *t*-BuOH (100 mL total) and heated to reflux. $KMnO_4$ (30.0 g, 190 mmol) was added to the refluxing mixture in three portions, spaced 1 hour apart. After refluxing for 20 hours total, the reaction was complete by NMR. The reaction was filtered while hot to remove the $MnO_{2(s)}$ formed. The reaction was acidified to pH 1 using $HCl_{(aq)}$ (6M) and the product di-acid precipitated. The solid was filtered and air-dried to give **4.8** (3.59 g, 58%) as a white solid. Synthetic procedure adapted from reference for similar compound.[191] The advice of A. Bubar from S. Eisler's lab at the University of New Brunswick was also used during this synthesis.[182] Mp: 248 – 251°C (decomp.). ^{1}H NMR (MeOD, 500 MHz): δ 8.10 (s, 1H), 8.27 (s, 1H). ^{13}C NMR (MeOD, 125 MHz): δ 125.8, 133.5, 134.5, 140.5, 167.7. HR-ESI-MS: *m/z* $[M+Na]^+$ calculated for $C_8H_2Br_2O_4Na$: 344.8197; found 344.8197 (QTof).

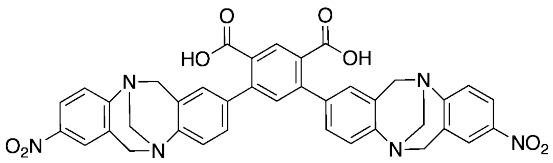


Dimethyl 4,6-dibromoisophthalate **4.9**. Concentrated H_2SO_4 (2.0 mL) was added drop-wise to a solution of **4.8** (3.0 g, 9.3 mmol) in MeOH (185 mL) which was then heated to reflux under N_2. After 48 hours, the reaction was complete by NMR. After cooling to room temperature, the reaction was diluted with H_2O (200 mL), brine (50 mL) and extracted with Et_2O (3 x 150 mL). The combined organics were washed with $NaOH_{(aq)}$ (2M, 200 mL), brine (150 mL), dried over Na_2SO_4 and condensed. Drying under high vacuum gave **4.9** (2.99 g, 92%) as a white solid. Compound **4.9** was made using the same procedure as starting material **4** in published reference.[192] No experimental characterization was given. Mp: 89 – 92°C. IR (thin film on KBr plate): 3093*w*, 3016*w*, 2957*w*, 1733*s*, 1709*m*, 1580*m*, 1543*w*, 1432*m*, 1426*m*, 1350*w*, 1286*s*, 1269*m*, 1237*s*, 1189*m*, 1118*s*, 1053*s*, 912*w*, 879*w*, 868*w*, 768*s*, 569*w*. ^{1}H NMR ($CDCl_3$, 300 MHz): δ

3.95 (s, 6H), 8.03 (s, 1H), 8.30 (s, 1H). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 53.0, 126.1, 130.9, 134.2, 140.0, 165.0. HR-ESI-MS: *m/z* $[M+Na]^+$ calculated for $C_{10}H_8Br_2O_4Na$: 374.8667; found, 374.8685 (QTof).

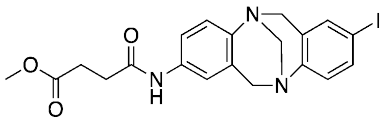


Compound **4.10**. Compound **4.9** (90 mg, 0.26 mmol), boronic ester Tröger's base **3.50** (300 mg, 0.76 mmol), anhydrous K_3PO_4 (327 mg, 1.5 mmol, freshly and quickly ground), S-phos (21 mg, 0.05 mmol), and $Pd_2(dba)_3$ (24 mg, 0.03 mmol) were added to an oven-dried Schlenk tube. The tube was then evacuated and backfilled with N_2 (x3), before adding anhydrous toluene through the top septum. After sealing, the reaction was heated to 100°C for 21 hours. The reaction was cooled, diluted with Et_2O and filtered through silica to remove any unreacted material with the Et_2O. The more polar product was then eluted from the silica with EtOAc, condensed, and precipitated from a minimum of EtOAc with hexanes. After air drying, this procedure yielded compound **4.10** (152 mg, 82% yield) as a solid. Synthetic conditions for the Tröger's base coupling were adopted from a published procedure.[158] Product Rf: 0.2 (20% EtOAc/DCM). Mp.: 174 – 176 °C. IR (thin film on KBr) ν (cm^{-1}): 3394br, 2952w, 1726s, 1611m, 1582m, 1515s, 1480m, 1339s, 1253m, 1210m, 1109m, 730m. ^{1}H NMR ($CDCl_3$, 360 MHz): δ 3.65 (s, 6H), 4.16 – 4.42 (m, 8H), 4.76 (d, *J* = 16.8, 2H), 4.77 (d, *J* = 16.6, 2H), 6.86 (s, 2H), 7.09 – 7.16 (m, 5H), 7.22 (d, *J* = 8.8, 2H), 7.87 (d, *J* = 2.5, 2H), 7.99 – 8.07 (m, 2H), 8.27 (s, 1H). ^{13}C NMR ($CDCl_3$, 90 MHz): δ 52.6, 58.9, 59.3, 67.0, 123.2, 123.4, 125.1, 125.9, 127.0, 127.4, 128.4, 129.1, 129.3, 132.5, 133.7, 136.4, 144.0, 145.0, 147.5, 155.2, 168.0. HR-ESI-MS: *m/z* $[M+H]^+$ calculated for $C_{40}H_{32}N_6O_8H$: 725.2360; found 725.2180 (QTof).

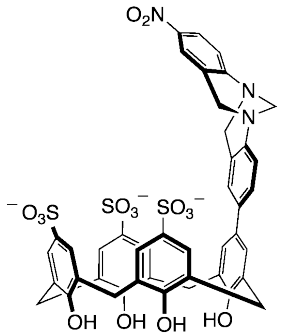


Compound **4.11**. $LiOH_{(s)}$ (32 mg, 1.3 mmol) and 4:1 MeOH/H_2O (1 mL) were combined in a 10 mL flask forming a suspension. Separately, **4.10** (30 mg, 0.04 mmol) and 4:1 MeOH/H_2O (1 mL) were combined and heated to 50°C before adding to the LiOH mixture. The resulting reaction was heated at reflux for 20 hours. Once complete, the

reaction was acidified with $HCl_{(aq)}$ (1M) to precipitate the product. After filtering, the product was dissolved with acetone and condensed to dryness. This procedure yielded **4.11** (19 mg, 64% yield) as a flaky, yellow solid. Mp.: 192 – 194 ^{o}C. IR (KBr pellet) ν (cm^{-1}): 3186br, 2948w, 1709s, 1614m, 1589m, 1515s, 1483m, 1337s, 1249m, 1213m, 1105m, 730m. ^{1}H NMR (acetone-d_6, 300 MHz): δ 4.40 – 4.71 (m, 8H), 4.90 (d, *J* = 17.0, 2H), 4.99 (d, *J* = 17.5, 2H), 7.13 – 7.22 (m, 3H), 7.28 – 7.34 (m, 2H), 7.36 – 7.43 (m, 2H), 7.49 (d, *J* = 8.6, 2H), 8.03 – 8.13 (m, 4H), 8.31 (s, 1H). ^{13}C NMR: Unavailable due to low amount of sample isolated. HR-ESI-MS: *m/z* $[M+Na]^+$ calculated for $C_{38}H_{28}N_6O_8Na$: 719.18611; found 719.18328 (Orbitrap).



Compound **4.15**. To a solution of methyl succinyl ester of succinic acid **4.14** (769 mg, 3.36 mmol) in DCM (95 mL), added amino-iodo Tröger's base **3.41** (1.28 g, 3.5 mmol) and stirred at room temperature under Ar for 19 hours. The solvent was evaporated, and the residue re-dissolved in EtOAc. The organic layer was washed with $HCl_{(aq)}$ (1 M), before being dried over Na_2SO_4, filtered and condensed. The resulting crude was suspended in Et_2O, filtered and air-dried. The collected solid was recrystalized from acetone to give compound **4.15** (189 mg, 12% yield) as a white solid. Mp: 198 – 200°C. IR (thin film from DCM on KBr plate): 3308w, 2948w, 2913w, 1737s, 1732s, 1495s, 1474s, 1207s, 831m. ^{1}H NMR (CD_2Cl_2, 500 MHz): δ 2.57 (t, *J* = 6.1, 2H), 2.66 (t, *J* = 6.1, 2H), 3.65 (s, 3H, OMe), 4.09 (d, *J* = 16.8, 1H), 4.11 (d, *J* = 16.7, 1H), 4.25 (d, *J* = 13.3, 1H), 4.31 (d, *J* = 13.0, 1H), 4.62 (d, *J* = 16.8, 1H), 4.64 (d, *J* = 16.7, 1H), 6.90 (d, *J* = 8.5, 1H), 7.08 (d, *J* = 8.6, 1H), 7.17 (dd, *J* = 8.6, *J* = 2.2, 1H), 7.22 (s (unresolved d), 1H), 7.26 (d, *J* = 2.0, 1H), 7.40 (br s, 1H, amide NH), 7.46 (dd, *J* = 8.5, *J* = 2.0, 1H). ^{13}C NMR (CD_2Cl_2, 125 MHz): δ 26.2, 29.7, 30.2, 32.4, 52.3, 58.6, 59.3, 67.4, 87.7, 118.5, 119.9, 125.8, 128.7, 130.8, 135.0, 136.4, 136.9, 170.1, 173.9. HR-ESI-MS: *m/z* $[M+Na]^+$ calculated for $C_{20}H_{20}IN_3O_3Na$: 500.0447; found 500.0434 (QTof).



Compound **4.19**. Bromo-calixarene **4.18** (26 mg, 0.036 mmol), nitro-Tröger's base-boronic acid **3.51** (11 mg, 0.034 mmol), NBu_4Br (6 mg, 0.02 mmol), $Pd(OAc)_2$ (1.1 mg, 0.005 mmol), and Na_2CO_3 (14 mg, 0.14 mmol) were combined in a heavy walled, sealed glass microwavable vial. Distilled H_2O (5.5 mL) was added and the reaction heated to 150°C for 10 minutes in a Biotage Initiator microwave (with stirring and cooling air on). The resulting suspension was filtered through Celite with distilled water and condensed. The residue was taken up in 90/10 H_2O/ACN with 1% TFA and purified by reverse phase HPLC using a 250 mm x 22 mm preparative C18 Alltech Apollo 10 µm column on a Shimadzu HPLC detecting at 280 nm. A gradient of 90 % H_2O (with 0.1% TFA)/10 % MeCN (with 0.1% TFA) to 90 % MeCN (with 0.1% TFA)/10 % H_2O (with 0.1% TFA) was used. After lyopholization of the isolated product fractions, the product **4.19** (11 mg, 35% yield) was obtained as a beige solid. Mp.: > 240 °C (decomp). IR (KBr pellet) ν (cm^{-1}): 3252s br, 3054w, 2910w, 2847w, 2358m, 1683w, 1610m, 1579m, 1511s, 1455s, 1338s, 1216s, 1149, 1114, 1041, 783w, 761w, 654,, 623m, 551m. ^{1}H NMR (MeOD-d_6, 500 MHz): δ 4.00 (br s, 8H), 4.40 – 4.53 (m, 2H), 5.05 – 5.17 (m, 2H), 7.34 (s, 1H), 7.35 – 7.41 (m, 3H), 7.44 (d, *J* = 8.8 Hz, 1H), 7.54 – 7.66 (m, 7H), 7.98 (s, 1H), 8.09 (d, *J* = 8.8 Hz, 1H). Missing -CH_2- from Tröger's base bridge under water peak in MeOD. ^{13}C NMR (MeOD-d_6, 125 MHz): δ 32.0, 32.1, 57.8, 58.4, 67.6, 124.1, 124.5, 125.0, 125.2, 126.8, 127.2, 128.0, 128.1, 128.2, 128.4, 128.8, 128.9, 129.2, 129.5 (x2), 129.8 (x2), 133.6, 135.1, 139.0, 143.9, 146.3, 150.9, 151.2, 152.8 (x2). HR-ESI-MS: *m/z* $[M-2H]^{2-}$ calculated for $C_{43}H_{35}N_3O_{15}S_3$-H_2: 436.55422; found 463.55445. (Orbitrap).

Chapter 5 – Concluding remarks

This thesis has evolved into an exploration of the properties that are critical to the design and construction of effective synthetic aromatic cages. The studies presented in Chapters 2 and 4 serve to highlight the two main observations to come out of this work. Seven hosts in four different series illustrate the strong influence of host preorganization on binding strength, and the unexpected power of the hydrophobic effect.

5.1 Host preorganization

Host preorganization and structure has conclusively been shown to be an important consideration in synthetic aromatic cage design. While the hosts of Chapter 2 (**2.1** – **2.4**) had ample aromatic surface with which to form a cage, the solution studies showed that the absence of a defined cavity was detrimental to the overall magnitude of binding. Similarly, concave host **4.3** also did not have a defined cavity, though its fixed aromatic rings did have a well defined surface. It was not until we combined fixed aromatic surface area with a defined cage in **4.19** that we achieved any significant measure of binding.

Host preorganization must also take into account the individual features of the primary guest target. Most of the molecules studied in this work were originally synthesized with Kme3 in mind. As such, hosts were designed to present aromatic surfaces in a concave manner to be well matched to the spherical cation of Kme3. However, when we changed guests (from lysine to arginine), we found a host that was poorly matched to Kme3 (**4.19**), was now more strongly suited to aRme2. Though the overall affinity constant of **4.19** for aRme2 is low by comparison to lysine, the change in selectivity resulting from a change in geometry highlights the affect that a specifically designed host can have on binding.

5.2 The hydrophobic effect

The more subtle and unexpected result that the hydrophobic effect has primacy over cation-pi interactions in these varied host families has also been demonstrated. Aromatic cages that were unable to be studied in water (such as **4.11**), exhibited very little binding partly due to the lack of any hydrophobic contribution from the non-aqueous solvent. Alternatively, even the extremely flexible, hydrophobic host **2.1** was shown to have strong binding in water due almost entirely to the hydrophobic effect. This was

repeatedly shown to be important when the guests themselves became more hydrophobic, either through increased methylation (K to Kme3, for example) or through increasing alkylation size (from Chapter 2). While the synthesis of water-soluble hosts with hydrophobic natures may be synthetically difficult, the advantages upon binding in water could be significant.

These lessons from synthetic systems raise the question of the relative importance of the hydrophobic effect and cation-pi interactions in proteins that have evolved to bind to methylated lysines and related quaternary ammonium cations like acetylcholine. The importance of the cation-pi interaction in proteins has been demonstrated by two types of studies. The first involve the use of neutral analogues of $RNMe_3^+$ groups — *t*-butyl, or $RCMe_3$ groups. The second type involve the use of fluorinated aromatic amino acids that reduce their electron-richness, and therefore the potency of their cation-pi interactions.

One example of the first kind of study was completed by Waters and co-workers.[49] Model β-hairpin peptides were synthesized containing either lysine ($R\text{-}NH_3^+$), trimethyllysine ($R\text{-}NMe_3^+$) or the neutral analog tert-butyl norleucine ($R\text{-}CMe_3$) in position to interact with a tryptophan residue on the opposing strand. The strength of the interaction was characterized by the amount of upfield shifting experienced by the side chain protons as they interacted with the aromatic face. As previously shown, the K to Kme3 substitution produced greater upfield shifting for the methylene chain protons (as also seen in Chapter 4). The neutral analog displayed no shifting, leading to the conclusion that the charged nature of Kme3 is critical to its interaction with tryptophan. A similar result was shown when the neutral analog was incorporated into H3 peptides at the K9 position and tested for binding to the native HP1 chromodomain against other H3K9 peptides. All of the methylated cationic peptides exhibited stronger binding than the neutral peptide, again enforcing the requirement of a cation to this interaction.

One example of the second kind of study was completed by Dougherty and co-workers through mutation studies in the nicotinic acetylcholine receptor (nAChR).[84] Four different tryptophan residues had been implicated as potential key residues in binding acetylcholine (ACh) in the active site through cation-pi interactions. Increasingly fluorinated tryptophan residues were substituted in at each site to evaluate the effect of binding to tryptophan with increasingly less electron density. Surprisingly, only one of the four residues shown any sensitivity to decreased electron density. This particular tryptophan, α149, was thus deemed to be the “key” residue involved in the cation-pi

interaction with ACh. The importance of the electron-richness of this amino acid to the cation-pi interaction is clear. However, less is noted about the role of the other aromatic rings present in the cage, if not to participate in a cation-pi interaction, given their very close contact with the cationic guest.

These studies, taken together with the results reported in this thesis, suggest strongly that the binding of alkylated ammonium ions by biomolecules is not simply governed by canonical cation-pi interactions. Instead, it seems that the cation-pi interactions can't function without significant binding energy arising from the hydrophobic effect. Strong solvation in water prevents "bare" cations (like lysine) from participating in cation-pi interactions, but can be counteracted through alkylation resulting in a decrease in the desolvation penalty. The positive charge of the cation appears necessary to properly attract and anchor the guest in a specific cage, but it is the contribution of the hydrophobic effect that is responsible for the overall strength of the interaction. Going forward, these lessons will be of utmost importance as the next generations of synthetic PTM-specific binding agents are developed.

Bibliography

(1) Greer, E. L.; Shi, Y. Histone methylation: a dynamic mark in health, disease and inheritance, *Nat. Rev. Genet.* **2012**, *13*, 343.

(2) Izzo, A.; Schneider, R. Chatting histone modifications in mammals, *Brief. Funct. Genomics* **2010**, *9*, 429.

(3) Zhang, Y.; Reinberg, D. Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails, *Genes Dev.* **2001**, *15*, 2343.

(4) Khare, S. P.; Habib, F.; Sharma, R.; Gadewal, N.; Gupta, S.; Galande, S. HIstome: a relational knowledgebase of human histone proteins and histone modifying enzymes, *Nucleic Acids Res.* **2011**.

(5) Wang, M.; Mok, M. W.; Harper, H.; Lee, W. H.; Min, J.; Knapp, S.; Oppermann, U.; Marsden, B.; Schapira, M. Structural Genomics of Histone Tail Recognition, *Bioinformatics* **2010**.

(6) Strahl, B. D.; Allis, C. D. The language of covalent histone modifications, *Nature* **2000**, *403*, 41.

(7) Filenko, N. A.; Kolar, C.; West, J. T.; Smith, S. A.; Hassan, Y. I.; Borgstahl, G. E. O.; Zempleni, J.; Lyubchenko, Y. L. The Role of Histone H4 Biotinylation in the Structure of Nucleosomes, *PLoS One* **2011**, *6*, e16299.

(8) Kobza, K.; Camporeale, G.; Rueckert, B.; Kueh, A.; Griffin, J. B.; Sarath, G.; Zempleni, J. K4, K9 and K18 in human histone H3 are targets for biotinylation by biotinidase, *FEBS J.* **2005**, *272*, 4249.

(9) Messner, S.; Altmeyer, M.; Zhao, H.; Pozivil, A.; Roschitzki, B.; Gehrig, P.; Rutishauser, D.; Huang, D.; Caflisch, A.; Hottiger, M. O. PARP1 ADP-ribosylates lysine residues of the core histone tails, *Nucleic Acids Res.* **2010**.

(10) Iñiguez-Lluhí, J. A. For a Healthy Histone Code, a Little SUMO in the Tail Keeps the Acetyl Away, *ACS Chem. Biol.* **2006**, *1*, 204.

(11) Jones, J. M.; Bhattacharyya, A.; Simkus, C.; Vallieres, B.; Veenstra, T. D.; Zhou, M. The RAG1 V(D)J recombinase/ubiquitin ligase promotes ubiquitylation of acetylated, phosphorylated histone 3.3, *Immunol. Lett.* **2011**, *136*, 156.

(12) *CRC Handbook of Chemistry and Physics*; 92nd Edition (2012 Internet Version) ed.; Haynes, W. M., Ed.; CRC Press/Taylor and Francis: Boca Raton, FL.

(13) Spartan 2006; Version 1.1.0 ed.; Wavefunction, Inc: Irvine, CA, 2010.

(14) Bannister, A. J.; Schneider, R.; Kouzarides, T. Histone Methylation: Dynamic or Static?, *Cell* **2002**, *109*, 801.

(15) Shi, Y.; Lan, F.; Matson, C.; Mulligan, P.; Whetstine, J. R.; Cole, P. A.; Casero, R. A.; Shi, Y. Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1, *Cell* **2004**, *119*, 941.

(16) Paik, W. K.; Kim, S. Enzymatic methylation of protein fractions from calf thymus nuclei, *Biochem. Biophys. Res. Commun.* **1967**, *29*, 14.

(17) Bedford, M. T.; Richard, S. Arginine Methylation: An Emergine Regulator of Protein Function, *Mol. Cell* **2005**, *18*, 263.

(18) Yang, Y.; Lu, Y.; Espejo, A.; Wu, J.; Xu, W.; Liang, S.; Bedford, M. T. TDRD3 Is an Effector Molecule for Arginine-Methylated Histone Marks, *Mol. Cell* **2010**, *40*, 1016.

(19) Kirmizis, A.; Santos-Rosa, H.; Penkett, C. J.; Singer, M. A.; Vermeulen, M.; Mann, M.; Bahler, J.; Green, R. D.; Kouzarides, T. Arginine methylation at histone H3R2 controls deposition of H3K4 trimethylation, *Nature* **2007**, *449*, 928.

(20) Lin, W.-J.; Gary, J. D.; Yang, M. C.; Clarke, S.; Herschman, H. R. The Mammalian Immediate-early TIS21 Protein and the Leukemia-associated BTG1 Protein Interact with a Protein-arginine N-methyltransferase, *J. Biol. Chem.* **1996**, *271*, 15034.

(21) Chang, B.; Chen, Y.; Zhao, Y.; Bruick, R. K. JMJD6 Is a Histone Arginine Demethylase, *Science* **2007**, *318*, 444.

(22) Thompson, P. R.; Fast, W. Histone Citrullination by Protein Arginine Deiminase: Is Arginine Methylation a Green Light or a Roadblock?, *ACS Chem. Biol.* **2006**, *1*, 433.

(23) Wang, Y.; Wysocka, J.; Sayegh, J.; Lee, Y.-H.; Perlin, J. R.; Leonelli, L.; Sonbuchner, L. S.; McDonald, C. H.; Cook, R. G.; Dou, Y.; Roeder, R. G.; Clarke, S.; Stallcup, M. R.; Allis, C. D.; Coonrod, S. A. Human PAD4 Regulates Histone Arginine Methylation Levels via Demethylimination, *Science* **2004**, *306*, 279.

(24) Cuthbert, G. L.; Daujat, S.; Snowden, A. W.; Erdjument-Bromage, H.; Hagiwara, T.; Yamada, M.; Schneider, R.; Gregory, P. D.; Tempst, P.; Bannister, A. J.; Kouzarides, T. Histone Deimination Antagonizes Arginine Methylation, *Cell* **2004**, *118*, 545.

(25) Kearney, P. L.; Bhatia, M.; Jones, N. G.; Yuan, L.; Glascock, M. C.; Catchings, K. L.; Yamada, M.; Thompson, P. R. Kinetic Characterization of Protein Arginine Deiminase 4: A Transcriptional Corepressor Implicated in the Onset and Progression of Rheumatoid Arthritis, *Biochemistry* **2005**, *44*, 10570.

(26) Hidaka, Y.; Hagiwara, T.; Yamada, M. Methylation of the guanidino group of arginine residues prevents citrullination by peptidylarginine deiminase IV, *FEBS Lett.* **2005**, *579*, 4088.

(27) Raijmakers, R.; Zendman, A. J. W.; Egberts, W. V.; Vossenaar, E. R.; Raats, J.; Soede-Huijbregts, C.; Rutjes, F. P. J. T.; van Veelen, P. A.; Drijfhout, J. W.; Pruijn, G. J. M. Methylation of Arginine Residues Interferes with Citrullination by Peptidylarginine Deiminases in vitro, *J. Mol. Biol.* **2007**, *367*, 1118.

(28) Webby, C. J.; Wolf, A.; Gromak, N.; Dreger, M.; Kramer, H.; Kessler, B.; Nielsen, M. L.; Schmitz, C.; Butler, D. S.; Yates, J. R.; Delahunty, C. M.; Hahn, P.; Lengeling, A.; Mann, M.; Proudfoot, N. J.; Schofield, C. J.; Böttger, A. Jmjd6 Catalyses Lysyl-Hydroxylation of U2AF65, a Protein Associated with RNA Splicing, *Science* **2009**, *325*, 90.

(29) Boudon, S.; Wipff, G.; Maigret, B. Monte Carlo simulations on the like-charged guanidinium-guanidinium ion pair in water, *J. Phys. Chem.* **2002**, *94*, 6056.

(30) Jacobs, S. A.; Taverna, S. D.; Zhang, Y.; Briggs, S. D.; Li, J.; Eissenberg, J. C.; Allis, C. D.; Khorasanizadeh, S. Specificity of the HP1 chromo domain for the methylated N-terminus of histone H3, *EMBO J.* **2001**, *20*, 5232.

(31) Botuyan, M. V.; Lee, J.; Ward, I. M.; Kim, J.-E.; Thompson, J. R.; Chen, J.; Mer, G. Structural Basis for the Methylation State-Specific Recognition of Histone H4-K20 by 53BP1 and Crb2 in DNA Repair, *Cell* **2006**, *127*, 1361.

(32) Huang, Y.; Fang, J.; Bedford, M. T.; Zhang, Y.; Xu, R.-M. Recognition of Histone H3 Lysine-4 Methylation by the Double Tudor Domain of JMJD2A, *Science* **2006**, *312*, 748.

(33) Li, H.; Fischle, W.; Wang, W.; Duncan, E. M.; Liang, L.; Murakami-Ishibe, S.; Allis, C. D.; Patel, D. J. Structural Basis for Lower Lysine Methylation State-Specific Readout by MBT Repeats of L3MBTL1 and an Engineered PHD Finger, *Mol. Cell* **2007**, *28*, 677.

(34) Xu, C.; Bian, C.; Yang, W.; Galka, M.; Ouyang, H.; Chen, C.; Qiu, W.; Liu, H.; Jones, A. E.; MacKenzie, F.; Pan, P.; Li, S. S.-C.; Wang, H.; Min, J. Binding of different histone marks differentially regulates the activity and specificity of polycomb repressive complex 2 (PRC2), *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 19266.

(35) Nady, N.; Lemak, A.; Walker, J. R.; Avvakumov, G. V.; Kareta, M. S.; Achour, M.; Xue, S.; Duan, S.; Allali-Hassani, A.; Zuo, X.; Wang, Y.-X.; Bronner, C.; Chédin, F.; Arrowsmith, C. H.; Dhe-Paganon, S. Recognition of Multivalent Histone States Associated with Heterochromatin by UHRF1 Protein, *J. Biol. Chem.* **2011**, *286*, 24300.

(36) Jacobson, R. H.; Ladurner, A. G.; King, D. S.; Tjian, R. Structure and Function of a Human TAFII250 Double Bromodomain Module, *Science* **2000**, *288*, 1422.

(37) Taverna, S. D.; Li, H.; Ruthenburg, A. J.; Allis, C. D.; Patel, D. J. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers, *Nat. Struct. Mol. Biol.* **2007**, *14*, 1025.

(38) Hansen, K. H.; Bracken, A. P.; Pasini, D.; Dietrich, N.; Gehani, S. S.; Monrad, A.; Rappsilber, J.; Lerdrup, M.; Helin, K. A model for transmission of the H3K27me3 epigenetic mark, *Nat. Cell Biol.* **2008**, *10*, 1291.

(39) Ingerman, L. A.; Cuellar, M. E.; Waters, M. L. A small molecule receptor that selectively recognizes trimethyl lysine in a histone peptide with native protein-like affinity, *Chem. Commun.* **2010**, *46*, 1839.

(40) Fischle, W.; Wang, Y.; Jacobs, S. A.; Kim, Y.; Allis, C. D.; Khorasanizadeh, S. Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains, *Genes & Development* **2003**, *17*, 1870.

(41) Eisert, R. J.; Waters, M. L. Tuning HP1α Chromodomain Selectivity for Di- and Trimethyllysine, *ChemBioChem* **2011**, *12*, 2786.

(42) Migliori, V.; Müller, J.; Phalke, S.; Low, D.; Bezzi, M.; Mok, W. C.; Sahu, S. K.; Gunaratne, J.; Capasso, P.; Bassi, C.; Cecatiello, V.; De Marco, A.; Blackstock, W.; Kuznetsov, V.; Amati, B.; Mapelli, M.; Guccione, E. Symmetric dimethylation of H3R2 is a newly identified histone mark that supports euchromatin maintenance, *Nat. Struct. Mol. Biol.* **2012**, *19*, 136.

(43) Zhao, Q.; Rank, G.; Tan, Y. T.; Li, H.; Moritz, R. L.; Simpson, R. J.; Cerruti, L.; Curtis, D. J.; Patel, D. J.; Allis, C. D.; Cunningham, J. M.; Jane, S. M. PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing, *Nat. Struct. Mol. Biol.* **2009**, *16*, 304.

(44) Friesen, W. J.; Massenet, S.; Paushkin, S.; Wyce, A.; Dreyfuss, G. SMN, the Product of the Spinal Muscular Atrophy Gene, Binds Preferentially to Dimethylarginine-Containing Protein Targets, *Mol. Cell* **2001**, *7*, 1111.

(45) Cheng, D.; Côté, J.; Shaaban, S.; Bedford, M. T. The Arginine Methyltransferase CARM1 Regulates the Coupling of Transcription and mRNA Processing, *Mol. Cell* **2007**, *25*, 71.

(46) Warme, P. K.; Morgan, R. S. A survey of amino acid side-chain interactions in 21 proteins, *J. Mol. Biol.* **1978**, *118*, 289.

(47) Ma, J. C.; Dougherty, D. A. The Cation-π Interaction, *Chem. Rev.* **1997**, *97*, 1303.

(48) Gallivan, J. P.; Dougherty, D. A. Cation-pi interactions in structural biology, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9459.

(49) Hughes, R. M.; Wiggins, K. R.; Khorasanizadeh, S.; Waters, M. L. Recognition of Trimethyllysine by a Chromodomain Is Not Driven by the Hydrophobic Effect, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 11184.

(50) Amendola, V.; Fabbrizzi, L.; Mosca, L. Anion recognition by hydrogen bonding: urea-based receptors, *Chem. Soc. Rev.* **2010**, *39*, 3889.

(51) Courtemanche, R. J. M.; Pinter, T.; Hof, F. Just add tetrazole: 5-(2-Pyrrolo)tetrazoles are simple, highly potent anion recognition elements, *Chem. Commun.* **2011**, *47*, 12688.

(52) Choi, K.; Hamilton, A. D. Macrocyclic anion receptors based on directed hydrogen bonding interactions, *Coord. Chem. Rev.*, *240*, 101.

(53) Hubbard, R. E.; Kamran Haider, M. Hydrogen Bonds in Proteins: Role and Strength, *eLS* **2001**.

(54) Crowley, P. B.; Golovin, A. Cation-pi interactions in protein-protein interfaces, *Proteins: Struct., Funct., Bioinf.* **2005**, *59*, 231.

(55) Hughes, R. M.; Waters, M. L. Arginine Methylation in a *β*-Hairpin Peptide: Implications for Arg-π Interactions, ΔC_p°, and the Cold Denatured State, *J. Am. Chem. Soc.* **2006**, *128*, 12735.

(56) Lesser, G. J.; Rose, G. D. Hydrophobicity of amino acid subgroups in proteins, *Proteins: Struct., Funct., Bioinf.* **1990**, *8*, 6.

(57) Pace, C. N.; Shirley, B. A.; McNutt, M.; Gajiwala, K. Forces contributing to the conformational stability of proteins, *FASEB J.* **1996**, *10*, 75.

(58) Meyer, E. A.; Castellano, R. K.; Diederich, F. Interactions with Aromatic Rings in Chemical and Biological Recognition, *Angew. Chem., Int. Ed.* **2003**, *42*, 1210.

(59) Hughes, R. M.; Waters, M. L. Influence of N-Methylation on a Cation-π Interaction Produces a Remarkably Stable *β*-Hairpin Peptide, *J. Am. Chem. Soc.* **2005**, *127*, 6518.

(60) Hughes, R. M.; Benshoff, M. L.; Waters, M. L. Effects of Chain Length and N-Methylation on a Cation–π Interaction in a β-Hairpin Peptide, *Chem. - Eur. J.* **2007**, *13*, 5753.

(61) Cram, D. J. Preorganization—From Solvents to Spherands, *Angew. Chem., Int. Ed.* **1986**, *25*, 1039.

(62) Beshara, C. S.; Jones, C. E.; Daze, K. D.; Lilgert, B. J.; Hof, F. A Simple Calixarene Recognizes Post-translationally Methylated Lysine, *ChemBioChem* **2010**, *11*, 63.

(63) Daze, K. D.; Pinter, T.; Beshara, C. S.; Ibraheem, A.; Minaker, S. A.; Ma, M. C. F.; Courtemanche, R. J. M.; Campbell, R. E.; Hof, F. Supramolecular hosts that recognize methyllysines and disrupt the interaction between a modified histone tail and its epigenetic reader protein, *Chem. Sci.* **2012**, *3*, 2695.

(64) Petti, M. A.; Shepodd, T. J.; Barrans, R. E.; Dougherty, D. A. "Hydrophobic" binding of water-soluble guests by high-symmetry, chiral hosts. An electron-rich receptor site with a general affinity for quaternary ammonium compounds and electron-deficient .pi. systems, *J. Am. Chem. Soc.* **1988**, *110*, 6825.

(65) Kearney, P. C.; Mizoue, L. S.; Kumpf, R. A.; Forman, J. E.; McCurdy, A.; Dougherty, D. A. Molecular recognition in aqueous media. New binding studies provide further insights into the cation-.pi. interaction and related phenomena, *J. Am. Chem. Soc.* **1993**, *115*, 9907.

(66) Biros, S. M.; Rebek, J. J. Structure and binding properties of water-soluble cavitands and capsules, *Chem. Soc. Rev.* **2007**, *36*, 93.

(67) Kazakova, E. K.; Makarova, N. A.; Ziganshina, A. U.; Muslinkina, L. A.; Muslinkin, A. A.; Habicher, W. D. Novel water-soluble tetrasulfonatomethylcalix[4]resorcinarenes, *Tetrahedron Lett.* **2000**, *41*, 10111.

(68) Mustafina, A. R.; Fedorenko, S. V.; Makarova, N. A.; Kazakova, E. K. H.; Bazhanova, Z. f. G.; Kataev, V. E.; Konovalov, A. I. The Inclusion Properties of a New Watersoluble Sulfonated Calix[4]resorcinarene towards Alkylammonium and N-Methylpyridinium Cations, *J. Inclusion Phenom. Macrocyclic Chem.* **2001**, *40*, 73.

(69) Wyman, I. W.; Macartney, D. H. Cucurbit[7]uril host-guest complexes of cholines and phosphonium cholines in aqueous solution, *Org. Biomol. Chem.* **2010**, *8*, 253.

(70) Gamal-Eldin, M.; Macartney, D. H. Cucurbituril complexes of steroidal neuromuscular blocking agents, *Canadian Chemistry Conference and Exhibition* **2012**.

(71) Lehn, J. M.; Vierling, P.; Hayward, R. C. Stable and selective guanidinium and imidazolium complexes of a macrocyclic receptor molecule, *J. Chem. Soc., Chem. Commun.* **1979**, 296.

(72) Schrader, T. Strong Binding of Alkylguanidinium Ions by Molecular Tweezers: An Artificial Selective Arginine Receptor Molecule with a Biomimetic Recognition Pattern, *Chem. - Eur. J.* **1997**, *3*, 1537.

(73) Schrader, T. H. Strong binding of arylguanidinium ions by benzylic bisphosphonates - Evidence for π-cation and π,π-interactions, *Tetrahedron Lett.* **1998**, *39*, 517.

(74) Wehner, M.; Schrader, T.; Finocchiaro, P.; Failla, S.; Consiglio, G. A Chiral Sensor for Arginine and Lysine, *Org. Lett.* **2000**, *2*, 605.

(75) Rensing, S.; Arendt, M.; Springer, A.; Grawe, T.; Schrader, T. Optimization of a Synthetic Arginine Receptor. Systematic Tuning of Noncovalent Interactions, *J. Org. Chem.* **2001**, *66*, 5814.

(76) Rensing, S.; Schrader, T. The First Synthetic Receptor for the RGD Sequence, *Org. Lett.* **2002**, *4*, 2161.

(77) Bell, T. W.; Khasanov, A. B.; Drew, M. G. B.; Filikov, A.; James, T. L. A Small-Molecule Guanidinium Receptor: The Arginine Cork, *Angew. Chem., Int. Ed.* **1999**, *38*, 2543.

(78) Ngola, S. M.; Kearney, P. C.; Mecozzi, S.; Russell, K.; Dougherty, D. A. A Selective Receptor for Arginine Derivatives in Aqueous Media. Energetic Consequences of Salt Bridges That Are Highly Exposed to Water, *J. Am. Chem. Soc.* **1999**, *121*, 1192.

(79) Jasper, C.; Schrader, T.; Panitzky, J.; Klärner, F.-G. Selective Complexation of N-Alkylpyridinium Salts: Recognition of NAD+ in Water, *Angew. Chem., Int. Ed.* **2002**, *41*, 1355.

(80) Fokkens, M.; Jasper, C.; Schrader, T.; Koziol, F.; Ochsenfeld, C.; Polkowska, J.; Lobert, M.; Kahlert, B.; Klärner, F.-G. Selective Complexation of N-Alkylpyridinium Salts: Binding of NAD+ in Water, *Chem. - Eur. J.* **2005**, *11*, 477.

(81) Fokkens, M.; Schrader, T.; Klärner, F.-G. A Molecular Tweezer for Lysine and Arginine, *J. Am. Chem. Soc.* **2005**, *127*, 14415.

(82) Tatko, C. D.; Waters, M. L. Investigation of the nature of the methionine–π interaction in β-hairpin peptide model systems, *Protein Sci.* **2004**, *13*, 2515.

(83) Ordentlich, A.; Barak, D.; Kronman, C.; Ariel, N.; Segall, Y.; Velan, B.; Shafferman, A. Contribution of Aromatic Moieties of Tyrosine 133 and of the Anionic Subsite Tryptophan 86 to Catalytic Efficiency and Allosteric Modulation of Acetylcholinesterase, *J. Biol. Chem.* **1995**, *270*, 2082.

(84) Zhong, W.; Gallivan, J. P.; Zhang, Y.; Li, L.; Lester, H. A.; Dougherty, D. A. From ab initio quantum mechanics to molecular neurobiology: A cation-π binding site in the nicotinic receptor, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12088.

(85) Martin, S. F.; Follows, B. C.; Hergenrother, P. J.; Trotter, B. K. The Choline Binding Site of Phospholipase C (*Bacillus cereus*): Insights into Substrate Specificity, *Biochemistry* **2000**, *39*, 3410.

(86) Beene, D. L.; Brandt, G. S.; Zhong, W.; Zacharias, N. M.; Lester, H. A.; Dougherty, D. A. Cation-π Interactions in Ligand Recognition by Serotonergic (5-HT3A) and Nicotinic Acetylcholine Receptors: The Anomalous Binding Properties of Nicotine, *Biochemistry* **2002**, *41*, 10262.

(87) Schmitt, J. D.; Sharples, C. G. V.; Caldwell, W. S. Molecular Recognition in Nicotinic Acetylcholine Receptors: The Importance of π-Cation Interactions, *J. Med. Chem.* **1999**, *42*, 3066.

(88) Shi, X.; Kachirskaia, I.; Walter, K. L.; Kuo, J.-H. A.; Lake, A.; Davrazou, F.; Chan, S. M.; Martin, D. G. E.; Fingerman, I. M.; Briggs, S. D.; Howe, L.; Utz, P. J.; Kutateladze, T. G.; Lugovskoy, A. A.; Bedford, M. T.; Gozani, O. Proteome-wide Analysis in Saccharomyces cerevisiae Identifies Several PHD Fingers as Novel Direct and Selective Binding Modules of Histone H3 Methylated at Either Lysine 4 or Lysine 36, *J. Biol. Chem.* **2007**, *282*, 2450.

(89) Jacobs, S. A.; Khorasanizadeh, S. Structure of HP1 Chromodomain Bound to a Lysine 9-Methylated Histone H3 Tail, *Science* **2002**, *295*, 2080.

(90) Flanagan, J. F.; Mi, L.-Z.; Chruszcz, M.; Cymborowski, M.; Clines, K. L.; Kim, Y.; Minor, W.; Rastinejad, F.; Khorasanizadeh, S. Double chromodomains cooperate to recognize the methylated histone H3 tail, *Nature* **2005**, *438*, 1181.

(91) Minaker, S. A.; Daze, K. D.; Ma, M. C. F.; Hof, F. Antibody-Free Reading of the Histone Code Using a Simple Chemical Sensor Array, *J. Am. Chem. Soc.* **2012**, *134*, 11674.

(92) Daze, K. D.; Hof, F. The Cation-π Interaction at Protein-Protein Interaction Interfaces: Developing and Learning from Synthetic Mimics of Proteins That Bind Methylated Lysines, *Acc. Chem. Res.* **2012**.

(93) Pedras, M. S. C.; Jha, M. Toward the control of Leptosphaeria maculans: Design, syntheses, biological activity, and metabolism of potential detoxification inhibitors of the crucifer phytoalexin brassinin, *Bioorg. Med. Chem.* **2006**, *14*, 4958.

(94) Li, H.; Homan, E. A.; Lampkins, A. J.; Ghiviriga, I.; Castellano, R. K. Synthesis and Self-Assembly of Functionalized Donor-Acceptor Molecules, *Org. Lett.* **2005**, *7*, 443.

(95) Martín-Gago, P.; Gomez-Caminals, M.; Ramón, R.; Verdaguer, X.; Martin-Malpartida, P.; Aragón, E.; Fernández-Carneado, J.; Ponsati, B.; López-Ruiz, P.; Cortes, M. A.; Colás, B.; Macias, M. J.; Riera, A. Fine-tuning the π–π Aromatic Interactions in Peptides: Somatostatin Analogues Containing Mesityl Alanine, *Angew. Chem., Int. Ed.* **2011**, 1820.

(96) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. Tryptophan zippers: Stable, monomeric *β*-hairpins, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 5578.

(97) Gans, P.; Sabatini, A.; Vacca, A.; HypNMR - http://www.hyperquad.co.uk/hypnmr.htm 2006.

(98) Gasa, T. B.; Valente, C.; Stoddart, J. F. Solution-phase counterion effects in supramolecular and mechanostereochemical systems, *Chem. Soc. Rev.* **2011**, *40*, 57.

(99) Whiting, A. L.; Neufeld, N. M.; Hof, F. A tryptophan-analog host whose interactions with ammonium ions in water are dominated by the hydrophobic effect, *Tetrahedron Lett.* **2009**, *50*, 7035.

(100) Zacharias, N.; Dougherty, D. A. Cation-[pi] interactions in ligand recognition and catalysis, *Trends Pharmacol. Sci.* **2002**, *23*, 281.

(101) Rajakumar, P.; Gayatri Swaroop, M. Synthesis of novel indole based cyclophanes and cylindrical cyclophanes by tandem alkylation methodology using NaH, *Tetrahedron Lett.* **2006**, *47*, 3019.

(102) Hennrich, G.; Lynch, V. M.; Anslyn, E. V. Novel C3-Symmetric Molecular Scaffolds with Potential Facial Differentiation, *Chem. - Eur. J.* **2002**, *8*, 2274.

(103) Mendel, D.; Ellman, J.; Chang, Z.; Veenstra, D.; Kollman, P.; Schultz, P. Probing protein stability with unnatural amino acids, *Science* **1992**, *256*, 1798.

(104) Schärer, K.; Morgenthaler, M.; Paulini, R.; Obst-Sander, U.; Banner, D. W.; Schlatter, D.; Benz, J.; Stihle, M.; Diederich, F. Quantification of Cation-pi Interactions in Protein-Ligand Complexes: Crystal-Structure Analysis of Factor Xa Bound to a Quaternary Ammonium Ion Ligand13, *Angew. Chem., Int. Ed.* **2005**, *44*, 4400.

(105) Li, H.; Ilin, S.; Wang, W.; Duncan, E. M.; Wysocka, J.; Allis, C. D.; Patel, D. J. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF, *Nature* **2006**, *442*, 91.

(106) Dyson, H. J.; Wright, P. E.; Scheraga, H. A. The role of hydrophobic interactions in initiation and propagation of protein folding, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 13057.

(107) Némethy, G.; Steinberg, I. Z.; Scheraga, H. A. Influence of water structure and of hydrophobic interactions on the strength of side-chain hydrogen bonds in proteins, *Biopolymers* **1963**, *1*, 43.

(108) Fernandez, A.; Scheraga, H. A. Insufficiently dehydrated hydrogen bonds as determinants of protein interactions, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 113.

(109) Fulmer, G. R.; Miller, A. J. M.; Sherden, N. H.; Gottlieb, H. E.; Nudelman, A.; Stoltz, B. M.; Bercaw, J. E.; Goldberg, K. I. NMR Chemical Shifts of Trace Impurities: Common Laboratory Solvents, Organics, and Gases in Deuterated Solvents Relevant to the Organometallic Chemist, *Organometallics* **2010**, *29*, 2176.

(110) Li, H.; Homan, E. A.; Lampkins, A. J.; Ghiviriga, I.; Castellano, R. K. Synthesis and Self-Assembly of Functionalized Donor,àíœÉ,àíAcceptor Molecules, *Org. Lett.* **2005**, *7*, 443.

(111) AIST; Spectral Database for Organic Compounds (SDBS) maintained by the National Institute of Advanced Industrial Science and Technology (Japan).

(112) Anslyn, E. V.; Dougherty, D. A. *Modern Physical Organic Chemistry*; University Science, 2006.

(113) Bryant, R. G. The NMR time scale, *J. Chem. Educ.* **1983**, *60*, 933.

(114) Wong, K. F.; Ng, S. On the use of the modified Benesi-Hildebrand equation to process NMR hydrogen bonding data, *Spectrochim. Acta, Part A.* **1976**, *32*, 455.

(115) Scatchard, G. The attractions of proteins for small molecules and ions, *Ann. N. Y. Acad. Sci.* **1949**, *51*, 660.

(116) Glasoe, P. K.; Long, F. A. Use of glass electrodes to measure acidities in deuterium oxide, *J. Phys. Chem.* **1960**, *64*, 188.

(117) Kim, Y. J.; Lek, M. T.; Schramm, M. P. pH Influenced molecular switching with micelle bound cavitands, *Chem. Commun.* **2011**, *47*, 9636.

(118) Petti, M. A.; Shepodd, T. J.; Dougherty, D. A. Design and synthesis of a new class of hydrophobic binding sites, *Tetrahedron Lett.* **1986**, *27*, 807.

(119) Tröger, J. Ueber einige mittelst nascirenden Formaldehydes entstehende Basen, *J. Prakt. Chem.* **1887**, *36*, 225.

(120) Spielman, M. A. The Structure of Troeger's Base, *J. Am. Chem. Soc.* **1935**, *57*, 583.

(121) Prelog, V.; Wieland, P. Über die Spaltung der Tröger'schen Base in optische Antipoden, ein Beitrag zur Stereochemie des dreiwertigen Stickstoffs, *Helv. Chim. Acta* **1944**, *27*, 1127.

(122) Kim, I. W.; Ryu, J. K.; Ahn, S. D.; Park, J. H.; Lee, K.-P.; Ryoo, J. J.; Hyun, M. H.; Okamoto, Y.; Yamamoto, C.; Carrf, P. W. Comparison of Chiral Separation on Amylose and Cellulose Tris(3,5-dimethylphenylcarbamate)-coated Zirconia in HPLC, *Bull. Korean Chem. Soc.* **2003**, *24*, 239.

(123) Zenoni, G.; Pedeferri, M.; Mazzotti, M.; Morbidelli, M. On-line monitoring of enantiomer concentration in chiral simulated moving bed chromatography, *J. Chromatogr., A* **2000**, *888*, 73.

(124) Umeda, S.; Satoh, T.; Saitoh, K.; Yokota, K.; Kakuchi, T. Synthesis of (1→6)-2,5-anhydro-D-glucitol through cyclopolymerization of 3,4-di-O-allyl-1,2 : 5,6-dianhydro-D-mannitol and optical resolution ability of its derivative in HPLC, *J. Polym. Sci., Part A: Polym. Chem.* **1998**, *36*, 901.

(125) Adrian, J. C.; Wilcox, C. S. Chemistry of synthetic receptors and functional group arrays. 10. Orderly functional group dyads. Recognition of biotin and adenine derivatives by a new synthetic host, *J. Am. Chem. Soc.* **1989**, *111*, 8055.

(126) Kobayashi, T.; Moriwaki, T. A new molecular receptor based on thiophene congener of Tröger's base: Selective binding with aliphatic and aromatic dicarboxylic acids, *Heterocycles* **2004**, *62*, 399.

(127) Crossley, M. J.; Hambley, T. W.; Mackay, L. G.; Try, A. C.; Walton, R. Porphyrin analogues of Troger's base: large chiral cavities with a bimetallic binding site, *J. Chem. Soc., Chem. Commun.* **1995**, 1077.

(128) Allen, P. R.; Reek, J. N. H.; Try, A. C.; Crossley, M. J. Resolution of a porphyrin analogue of Tröger's base by making use of ligand binding affinity differences of the enantiomers, *Tetrahedron: Asymmetry* **1997**, *8*, 1161.

(129) Crossley, M. J.; Mackay, L. G.; Try, A. C. Enantioselective recognition of histidine and lysine esters by porphyrin chiral clefts and detection of amino acid conformations in the bound state, *J. Chem. Soc., Chem. Commun.* **1995**, 1925.

(130) Hansson, A. P.; Norrby, P.-O.; Wärnmark, K. A bis(crown-ether) analogue of Tröger's base: Recognition of achiral and chiral primary bisammonium salts, *Tetrahedron Lett.* **1998**, *39*, 4565.

(131) Yashima, E.; Akashi, M.; Miyauchi, N. Chiral Bis(1,10-phenanthroline) with Tröger's Base Skeleton. Synthesis and Interaction with DNA, *Chem. Lett.* **1991**, *20*, 1017.

(132) Tatibouët, A.; Demeunynck, M.; Andraud, C.; Collet, A.; Lhomme, J. Synthesis and study of an acridine substituted Tröger's base: preferential binding of the (-)-isomer to B-DNA, *Chem. Commun.* **1999**, 161.

(133) Baldeyrou, B.; Tardy, C.; Bailly, C.; Colson, P.; Houssier, C.; Charmantray, F.; Demeunynck, M. Synthesis and DNA interaction of a mixed proflavine-phenanthroline Tröger base, *Eur. J. Med. Chem.* **2002**, *37*, 315.

(134) Van Gijte, O.; Tatibouët, A.; Demeunynck, M.; Lhomme, J.; Kirsch-De Mesmaeker, A. A phenanthroline analogue of Tröger's base as bridging ligand in the synthesis of a bimetallic ruthenium (II) complex, *Tetrahedron Lett.* **1997**, *38*, 1567.

(135) Bresson, C.; Luhmer, M.; Demeunynck, M.; Kirsch-De Mesmaeker, A.; Pierard, F. The diastereoisomeric forms of a mononuclear Ru(II) complex bearing a bis-phenanthroline Tröger's base, *Tetrahedron Lett.* **2004**, *45*, 2863.

(136) Paliwal, S.; Geib, S.; Wilcox, C. S. Molecular Torsion Balance for Weak Molecular Recognition Forces. Effects of "Tilted-T" Edge-to-Face Aromatic Interactions on Conformational Selection and Solid-State Structure, *J. Am. Chem. Soc.* **1994**, *116*, 4497.

(137) Kim, E.-i.; Paliwal, S.; Wilcox, C. S. Measurements of Molecular Electrostatic Field Effects in Edge-to-Face Aromatic Interactions and CH-π Interactions with Implications for Protein Folding and Molecular Recognition, *J. Am. Chem. Soc.* **1998**, *120*, 11192.

(138) Ibrahim, A. A.; Matsumoto, M.; Miyahara, Y.; Izumi, K.; Suenaga, M.; Shimizu, N.; Inazu, T. Synthesis and properties of a new series of trögerophanes, *J. Heterocycl. Chem.* **1998**, *35*, 209.

(139) Manjula, A.; Nagarajan, M. New supramolecular hosts: Synthesis and cation binding studies of novel Tröger's base-crown ether composites, *Tetrahedron* **1997**, *53*, 11859.

(140) Moore, S. S.; Tarnowski, T. L.; Newcomb, M.; Cram, D. J. Host-guest complexation. 4. Remote substituent effects on macrocyclic polyether binding to metal and ammonium ions, *J. Am. Chem. Soc.* **1977**, *99*, 6398.

(141) Cowart, M. D.; Sucholeiki, I.; Bukownik, R. R.; Wilcox, C. S. Molecular recognition in aqueous media. Conformationally restricted water-soluble cyclophanes derived from 6H,12H-5,11-methanodibenzo[b,f][1,5]diazocine, *J. Am. Chem. Soc.* **1988**, *110*, 6204.

(142) Webb, T. H.; Suh, H.; Wilcox, C. S. Chemistry of synthetic receptors and functional group arrays. 16. Enantioselective and diastereoselective molecular recognition of alicyclic substrates in aqueous media by a chiral, resolved synthetic receptor, *J. Am. Chem. Soc.* **1991**, *113*, 8554.

(143) Bag, B. G.; von Kiedrowski, G. Stepwise Replication of a Tröger's Base Analogue, *Angew. Chem., Int. Ed.* **1999**, *38*, 3713.

(144) Mas, T.; Pardo, C.; Salort, F.; Elguero, J.; Torres, M. R. A New Entry to Bis-Tröger's Bases, *Eur. J. Org. Chem.* **2004**, 1097.

(145) Dolenský, B.; Valík, M.; Sýkora, D.; Král, V. Synthetic Routes to Linear Oligo-Tröger's Bases, *Org. Lett.* **2004**, *7*, 67.

(146) Dolenský, B.; Valík, M.; Král, V. In *8th International Conference on Calixarenes (CALIX 2005)* Prague, Czech Republic, 2005.

(147) Artacho, J.; Nilsson, P.; Bergquist, K.-E.; Wendt, O. F.; Wärnmark, K. The Synthesis and Characterization of all Diastereomers of a Linear Symmetrically Fused Tris-Tröger's Base Analogue: New Chiral Cleft Compounds, *Chem. - Eur. J.* **2006**, *12*, 2692.

(148) Hansson, A.; Wixe, T.; Bergquist, K.-E.; Wärnmark, K. A Desymmetrization Route to Fused Tröger's Base Analogues: Synthesis, Isolation, and Characterization of the First Anti-Anti Diastereomer of a Fused Tris-Tröger's Base Analogue, *Org. Lett.* **2005**, *7*, 2019.

(149) Bag, B. G.; Maitra, U. A Convenient Method for the Synthesis Of Tröger's Base Analogues, *Synth. Commun.* **1995**, *25*, 1849.

(150) Li, Z.; Xu, X.; Peng, Y.; Jiang, Z.; Ding, C.; Qian, X. An Unusual Synthesis of Tröger's Bases Using DMSO/HCl as Formaldehyde Equivalent, *Synthesis* **2005**, *8*, 1228

(151) Bhuiyan, M. D. H.; Mahon, A. B.; Jensen, P.; Clegg, J. K.; Try, A. C. Synthesis of Symmetric Dinitro-Functionalised Tröger's Base Analogues, *European Journal of Organic Chemistry* **2009**, *2009*, 687.

(152) Jensen, J.; Wärnmark, K. Synthesis of Halogen Substituted Analogues of Tröger's Base, *Synthesis* **2001**, *2001*, 1873.

(153) Jensen, J.; Strozyk, M.; Wärnmark, K. Influence of scale, stoichiometry and temperature on the synthesis of 2,8-dihalo analogues of Tröger's base from the corresponding anilines and paraformaldehyde, *J. Heterocycl. Chem.* **2003**, *40*, 373.

(154) Hansson, A.; Jensen, J.; Wendt, Ola F.; Wärnmark, K. Synthesis of Dihalo-Substituted Analogues of Tröger's Base from ortho- and meta-Substituted Anilines, *Eur. J. Org. Chem.* **2003**, *2003*, 3179.

(155) Bhuiyan, M. D. H.; Zhu, K.-X.; Jensen, P.; Try, A. C. Synthesis of Symmetric Diester-Functionalised Tröger's Base Analogues, *Eur. J. Org. Chem.* **2010**, *2010*, 4662.

(156) Jensen, J.; Tejler, J.; Wärnmark, K. General Protocols for the Synthesis of C_2-Symmetric and Asymmetric 2,8-Disubstituted Analogues of Tröger's Base via Efficient Bromine-Lithium Exchanges of 2,8-Dibromo-6*H*,12*H*-5,11-methanodibenzo[*b,f*][1,5]diazocine, *J. Org. Chem.* **2002**, *67*, 6008.

(157) Artacho, J.; Wärnmark, K. On the Synthesis of Asymmetric and Dissymmetric Amino Analogues of Tröger's Base, *Synthesis* **2009**, *2009*, 3120.

(158) Hof, F.; Schär, M.; Scofield, D. M.; Fischer, F.; Diederich, F.; Sergeyev, S. Preparation of Tröger Base Derivatives by Cross-Coupling Methodologies, *Helv. Chim. Acta* **2005**, *88*, 2333.

(159) Faroughi, M.; Zhu, K.-X.; Jensen, P.; Craig, D. C.; Try, A. C. One-Step Synthesis of Tröger's Base Hybrids Containing at Least One Halogen Atom, *Eur. J. Org. Chem.* **2009**, *2009*, 4266.

(160) Webb, T. H.; Wilcox, C. S. Improved synthesis of symmetrical and unsymmetrical 5,11-methanodibenzo[b,f][1,5]diazocines. Readily available nanoscale structural units, *J. Org. Chem.* **1990**, *55*, 363.

(161) Mederski, W. W. K. R.; Baumgarth, M.; Germann, M.; Kux, D.; Weitzel, T. A convenient synthesis of 4-aminoaryl substituted cyclic imides, *Tetrahedron Lett.* **2003**, *44*, 2133.

(162) Kiehne, U.; Weilandt, T.; Lützen, A. Diastereoselective Self-Assembly of Double-Stranded Helicates from Tröger's Base Derivatives, *Org. Lett.* **2007**, *9*, 1283.

(163) Miyake, M.; Wilcox, C. S. Design and synthesis of a novel cyclophane host for aryl phosphate, *Heterocycles* **2002**, *57*, 515.

(164) Bew, S. P.; Legentil, L.; Scholier, V.; Sharma, S. V. α-Amino acid Troger base derivatives, possible conformationally restricted scaffolds?, *Chem. Commun.* **2007**, 389.

(165) Bhayana, B.; Wilcox, C. S. A Minimal Protein Folding Model To Measure Hydrophobic and CH–π Effects on Interactions between Nonpolar Surfaces in Water, *Angew. Chem. Int. Ed.* **2007**, *46*, 6833.

(166) Fischer, F. R.; Schweizer, W. B.; Diederich, F. Substituent effects on the aromatic edge-to-face interaction, *Chem. Commun.* **2008**, 4031.

(167) Didier, D.; Sergeyev, S. Thiourea derivatives of Tröger's base: synthesis, enantioseparation and evaluation in organocatalysis of Michael addition to nitroolefins, *ARKIVOC* **2009**, 124.

(168) Satishkumar, S.; Periasamy, M. A convenient method for the synthesis and resolution of Tröger base, *Tetrahedron: Asymmetry* **2006**, *17*, 1116.

(169) Coutts, S. J.; Adams, J.; Krolikowski, D.; Snow, R. J. Two efficient methods for the cleavage of pinanediol boronate esters yielding the free boronic acids, *Tetrahedron Lett.* **1994**, *35*, 5109.

(170) Sandmeyer, T. Ueber die Ersetzung der Amid-gruppe durch Chlor, Brom und Cyan in den aromatischen Substanzen, *Ber. Dtsch. Chem. Ges.* **1884**, *17*, 2650.

(171) Kiehne, U.; Weilandt, T.; Lutzen, A. Diastereoselective Self-Assembly of Double-Stranded Helicates from Troger's Base Derivatives, *Org. Lett.* **2007**, *9*, 1283.

(172) Das, J.; Sitaram Kumar, M.; Subrahmanyam, D.; Sastry, T. V. R. S.; Prasad Narasimhulu, C.; Laxman Rao, C. V.; Kannan, M.; Roshaiah, M.; Awasthi, R.; Patil, S. N.; Sarnaik, H. M.; Rao Mamidi, N. V. S.; Selvakumar, N.; Iqbal, J. Substituent activity relationship studies on new azolo benzoxazepinyl oxazolidinones, *Bioorg. Med. Chem.* **2006**, *14*, 8032.

(173) Hu, W.-P.; Wang, J.-J.; Lin, F.-L.; Lin, Y.-C.; Lin, S.-R.; Hsu, M.-H. An Efficient Synthesis of Pyrrolo[2,1-c][1,4]benzodiazepine. Synthesis of the Antibiotic DC-81, *J. Org. Chem.* **2001**, *66*, 2881.

(174) Mas, T.; Pardo, C.; Salort, F.; Elguero, J.; Torres, M. R. A New Entry to Bis-Tröger's Bases, *Eur. J. Org. Chem.* **2004**, *2004*, 1097.

(175) Kiehne, U.; Weilandt, T.; Lutzen, A. Diastereoselective Self-Assembly of Double-Stranded Helicates from Troger's Base Derivatives, *Org. Lett.* **2007**, *9*, 1283.

(176) Sergeyev, S.; Schär, M.; Seiler, P.; Lukoyanova, O.; Echegoyen, L.; Diederich, F. Synthesis of trans-1, trans-2, trans-3, and trans-4 Bisadducts of C60 by Regio- and Stereoselective Tether-Directed Remote Functionalization, *Chem. - Eur. J.* **2005**, *11*, 2284.

(177) Daze, K. D.; Ma, M. C. F.; Pineux, F.; Hof, F. Synthesis of New Trisulfonated Calix[4]arenes Functionalized at the Upper Rim, and Their Complexation with the Trimethyllysine Epigenetic Mark, *Org. Lett.* **2012**, *14*, 1512.

(178) Wepster, B. M. Steric effects on mesomerism: XI, The ultra-violet absorption spectrum of Tröger's base and the stereochemistry of aromatic amines, *Recl. Trav. Chim. Pays-Bas* **1953**, *72*, 661.

(179) R. Crampton, M.; A. Robotham, I. Acidities of Some Substituted Ammonium Ions in Dimethyl Sulfoxide, *J. Chem. Res., Synop.* **1997**, 22.

(180) Bonifacio, M. C.; Robertson, C. R.; Jung, J.-Y.; King, B. T. Polycyclic Aromatic Hydrocarbons by Ring-Closing Metathesis, *J. Org. Chem.* **2005**, *70*, 8522.

(181) Field, J. E.; Hill, T. J.; Venkataraman, D. Bridged Triarylamines: A New Class of Heterohelicenes, *J. Org. Chem.* **2003**, *68*, 6071.

(182) Personal communication, Alex Bubar and Prof. Sara Eisler, University of New Brunswick.

(183) Zhao, W.; Carreira, E. M. Solid-Phase Synthesis of Photochromic Spiropyrans, *Org. Lett.* **2005**, *7*, 1609.

(184) Shinkai, S.; Araki, K.; Koreishi, H.; Tsubaki, T.; Manabe, O. On the Acidity of the Hydroxyl Groups in Calix[4]arenes and the Dissociation-dependent Conformational Change, *Chem. Lett.* **1986**, *15*, 1351.

(185) Shinkai, S.; Araki, K.; Matsuda, T.; Manabe, O. NMR Determination of Association Constants for Aqueous Calixarene Complexes and Guest Template Effects on the Conformational Freedom, *Bull. Chem. Soc. Jpn.* **1989**, *62*, 3856.

(186) Douteau-Guevel, N.; W. Coleman, A.; Morel, J.-P.; Morel-Desrosiers, N. Complexation of the basic amino acids lysine and arginine by three sulfonatocalix[n]arenes (n = 4, 6 and 8) in water: microcalorimetric determination of the Gibbs energies, enthalpies and entropies of complexation, *J. Chem. Soc., Perkin Trans. 2* **1999**, 629.

(187) McQuinn, K.; McIndoe, J. S.; Hof, F. Insights into the Post-Translational Methylation of Arginine from Studies of Guanidinium–Water Nanodroplets, *Chem. - Eur. J.* **2008**, *14*, 6483.

(188) Jursic, B. S.; Patel, P. K. Cyclodextrin assisted enantiomeric recognition of benzo[de]isoquinoline-1,3-dione derived amino acids, *Tetrahedron* **2005**, *61*, 919.

(189) Bonifacio, M. C.; Robertson, C. R.; Jung, J.-Y.; King, B. T. Polycyclic Aromatic Hydrocarbons by Ring-Closing Metathesis, *J. Org. Chem.* **2005**, *70*, 8522.

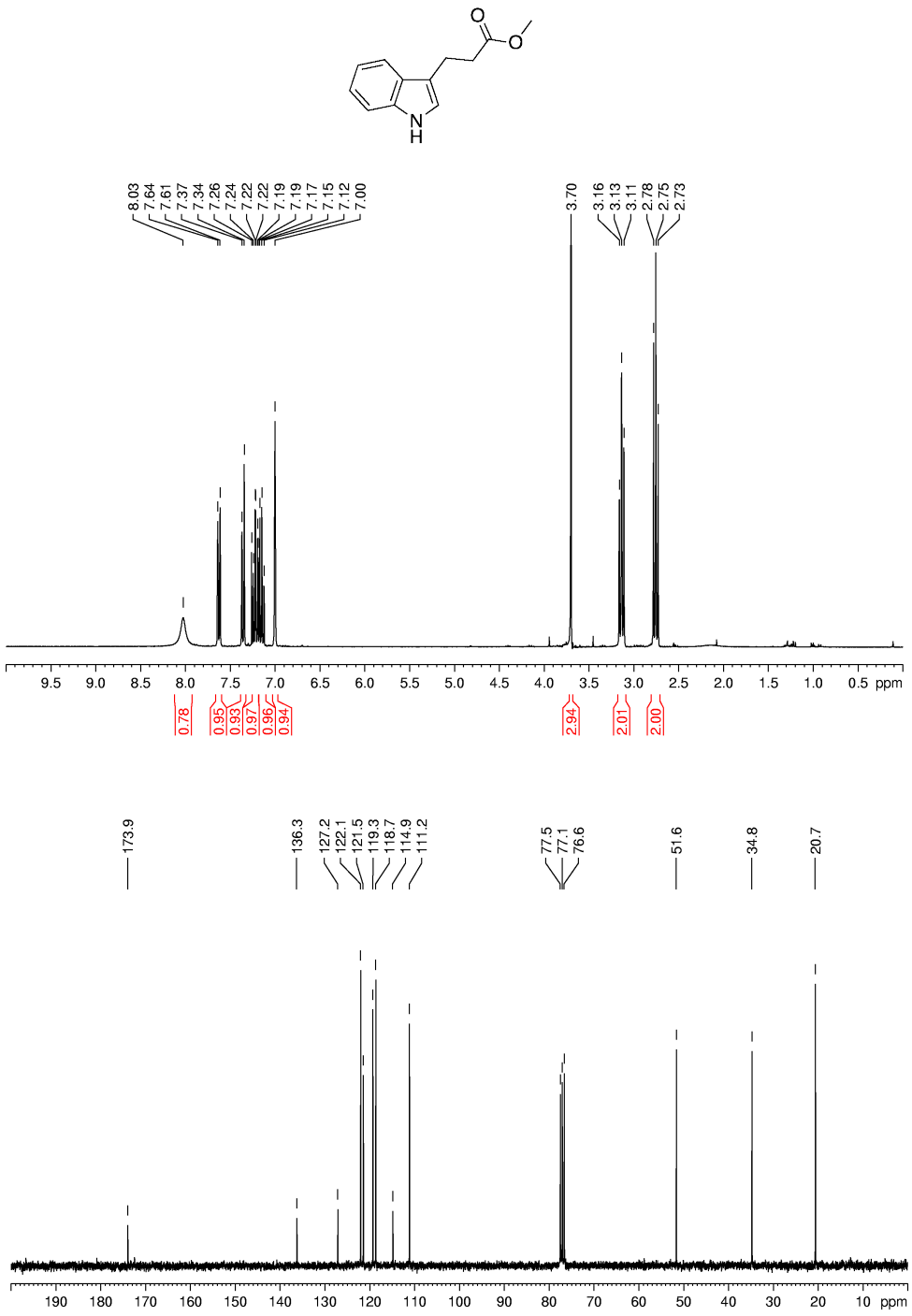
(190) Wilkinson, A. J.; Puschmann, H.; Howard, J. A. K.; Foster, C. E.; Williams, J. A. G. Luminescent Complexes of Iridium(III) Containing N-C-N-Coordinating Terdentate Ligands, *Inorg. Chem.* **2006**, *45*, 8685.

(191) Field, J. E.; Hill, T. J.; Venkataraman, D. Bridged Triarylamines:‚Äâ A New Class of Heterohelicenes, *J. Org. Chem.* **2003**, *68*, 6071.

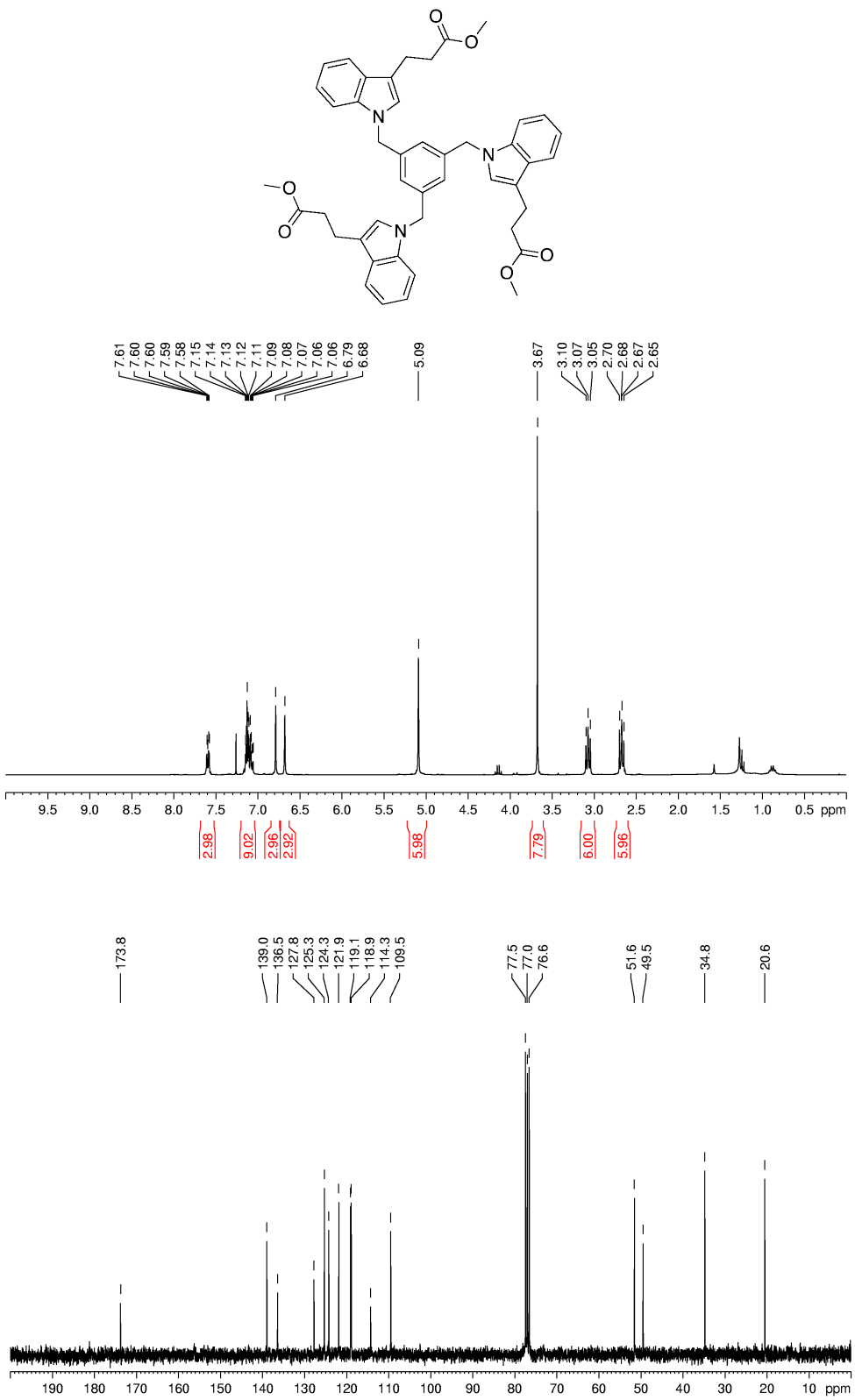
(192) Rajca, A.; Takahashi, M.; Pink, M.; Spagnol, G. l.; Rajca, S. Conformationally Constrained, Stable, Triplet Ground State (S = 1) Nitroxide Diradicals. Antiferromagnetic Chains of S = 1 Diradicals, *J. Am. Chem. Soc.* **2007**, *129*, 10159.

Appendices

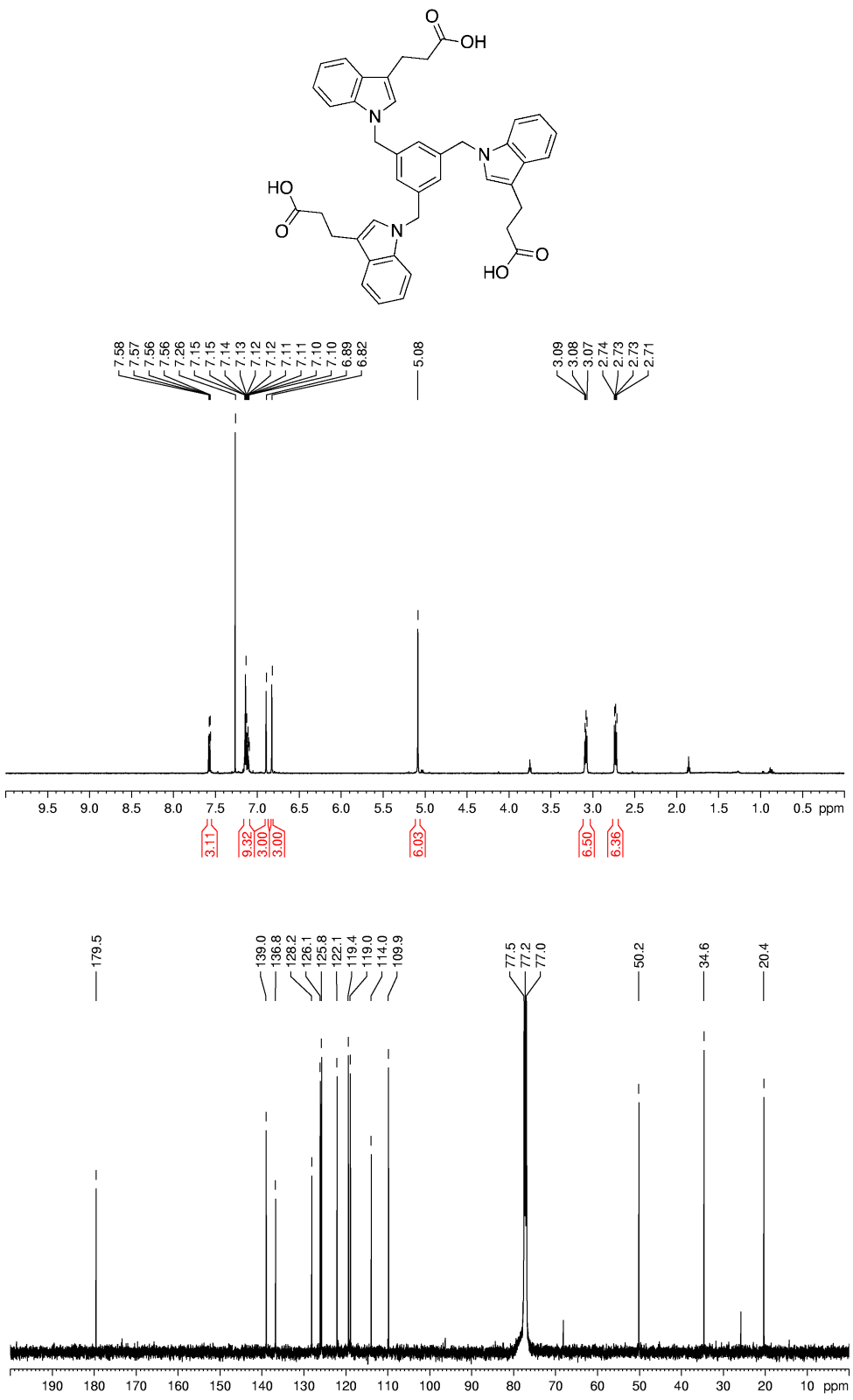
Appendix A – ^{1}H and ^{13}C NMR spectra



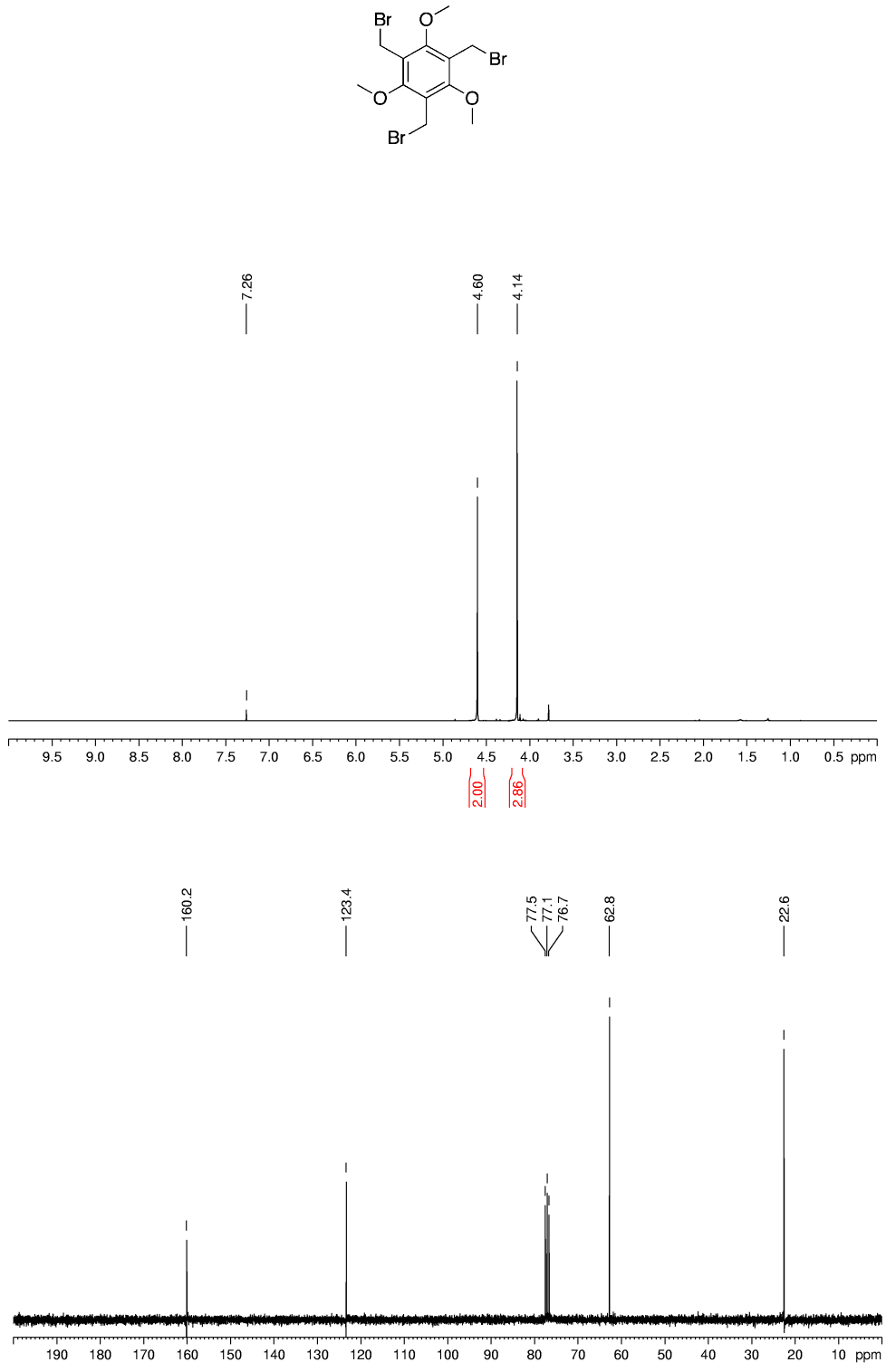
Spectra 1. Compound **2.6**



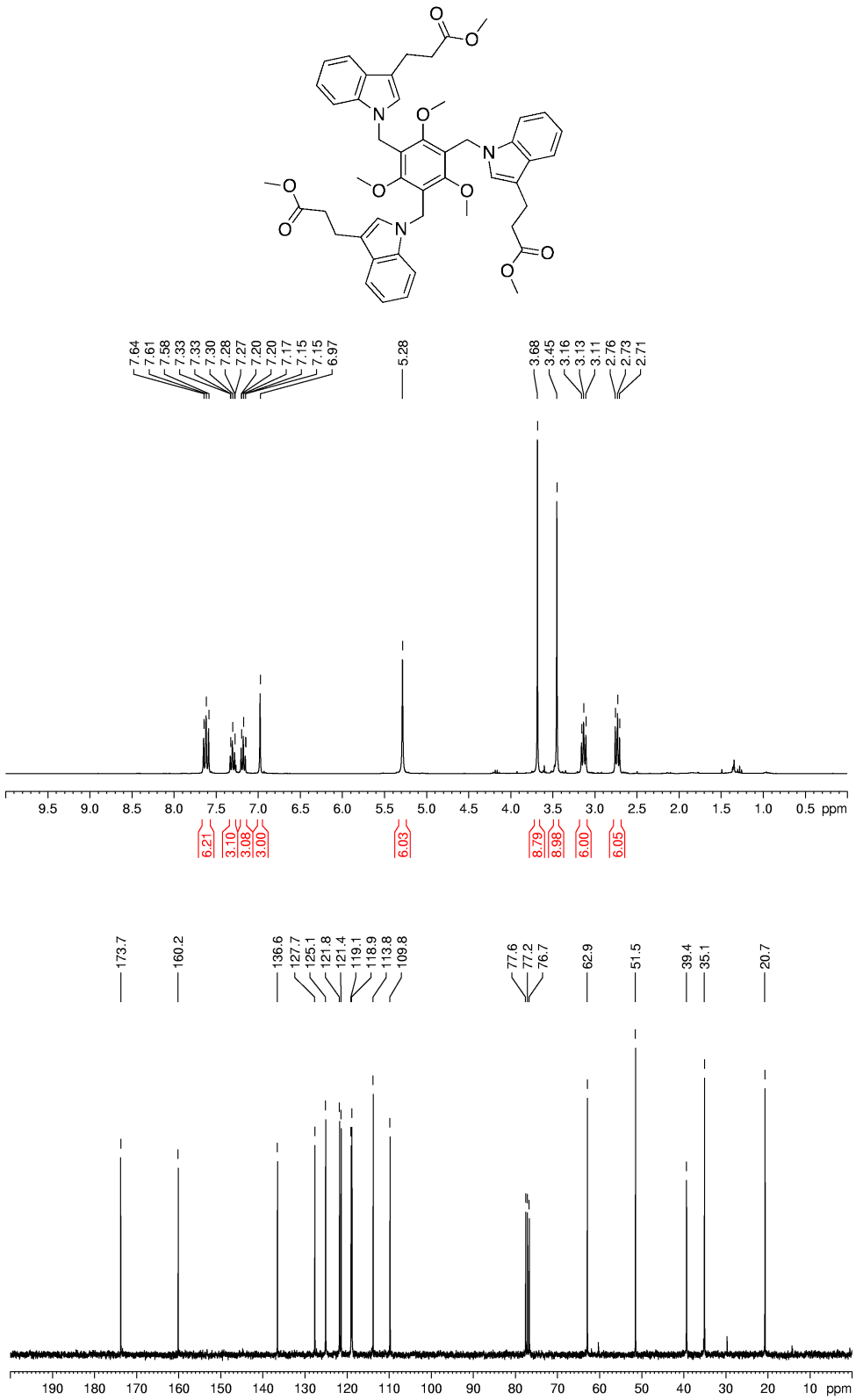
Spectra 2. Compound **2.8**



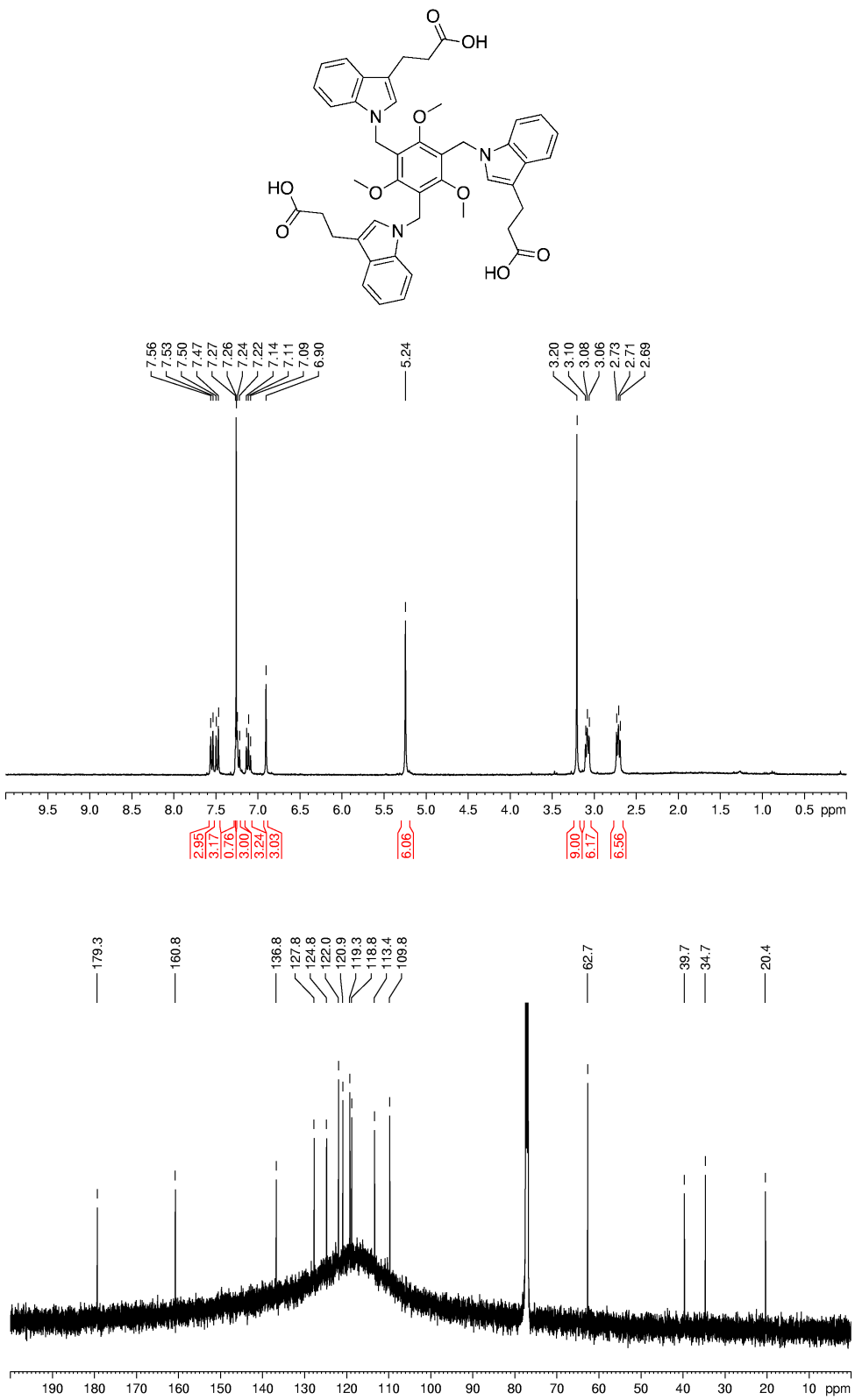
Spectra 3. Compound **2.9**



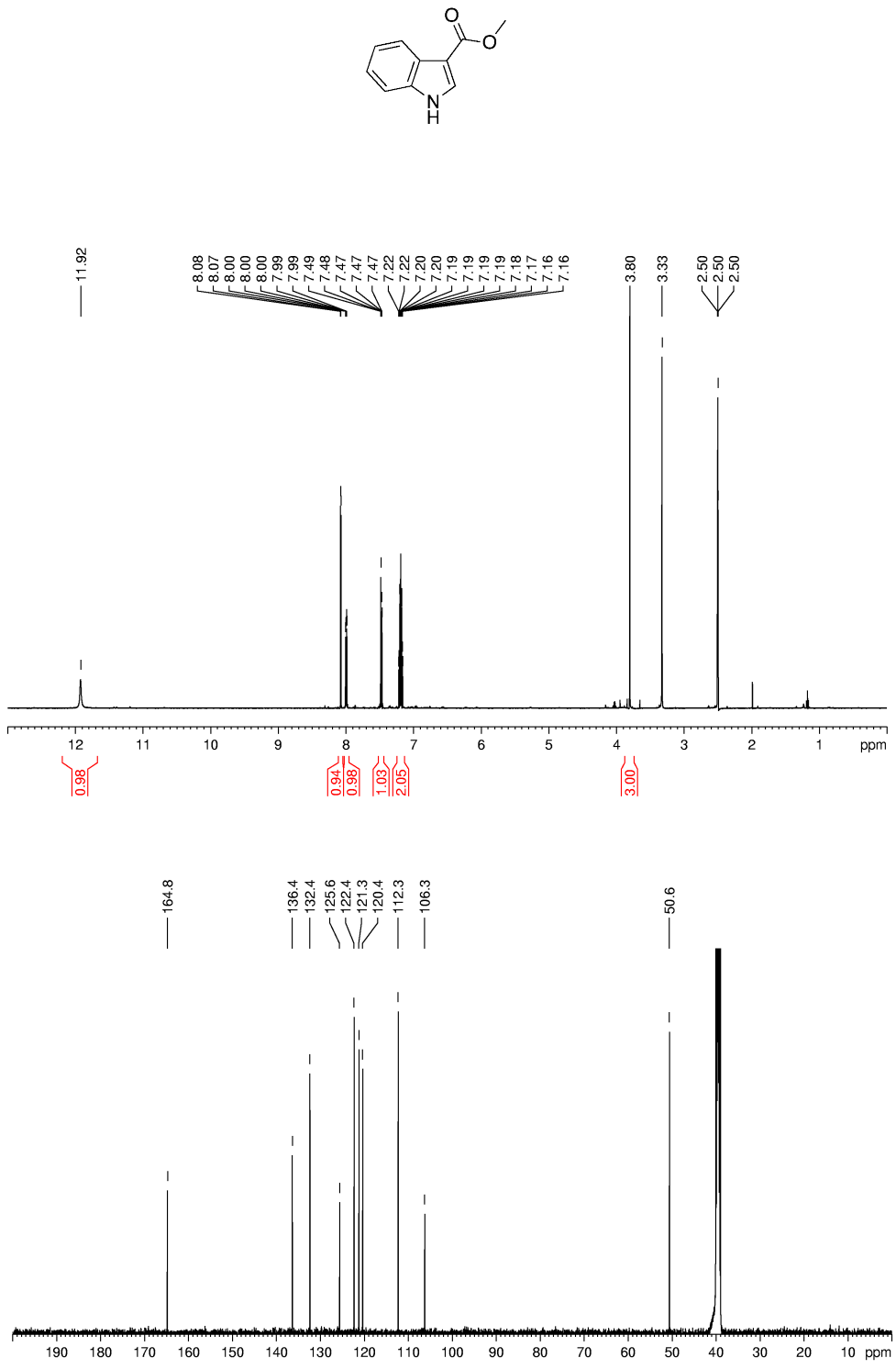
Spectra 4. Compound **2.11**



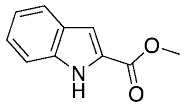
Spectra 5. Compound **2.12**

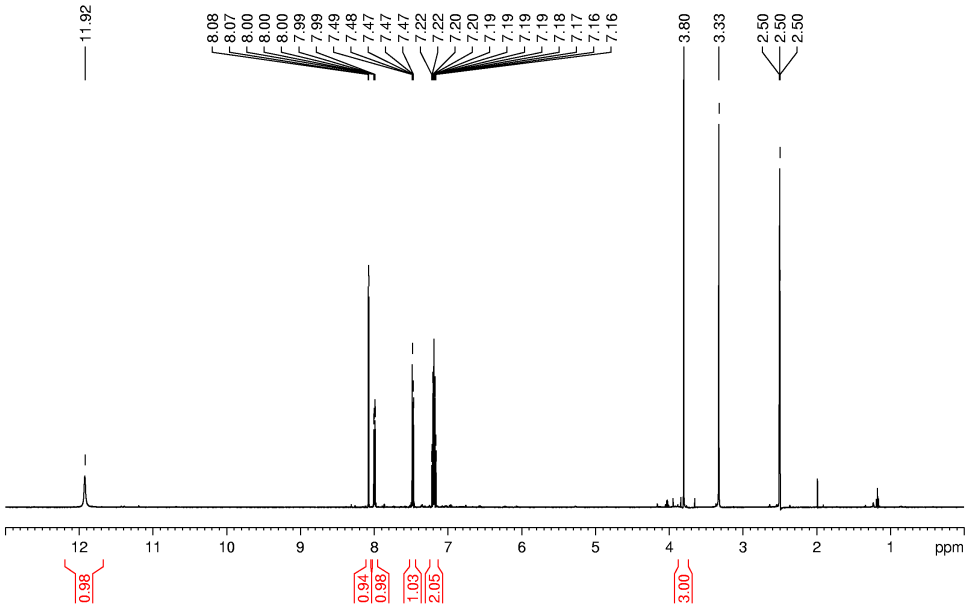


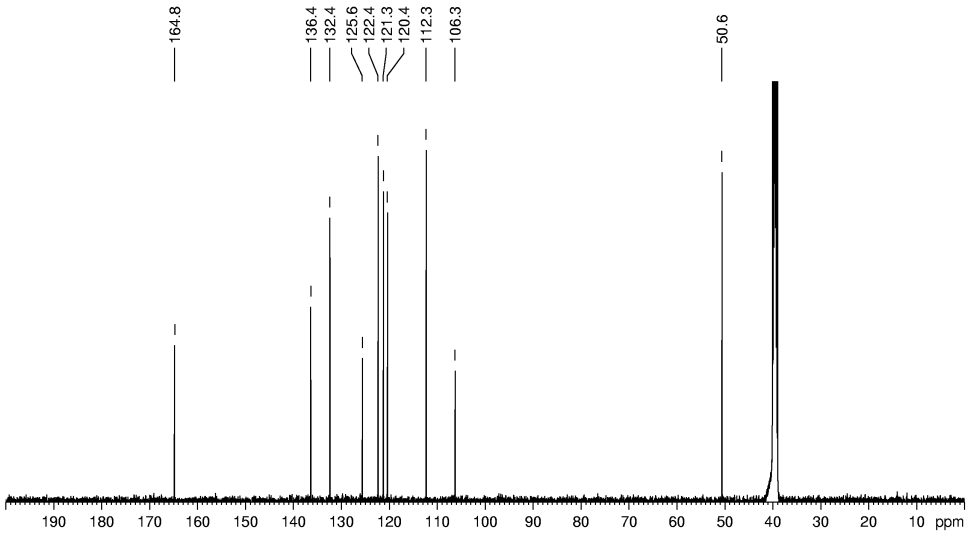
Spectra 6. Compound **2.13**



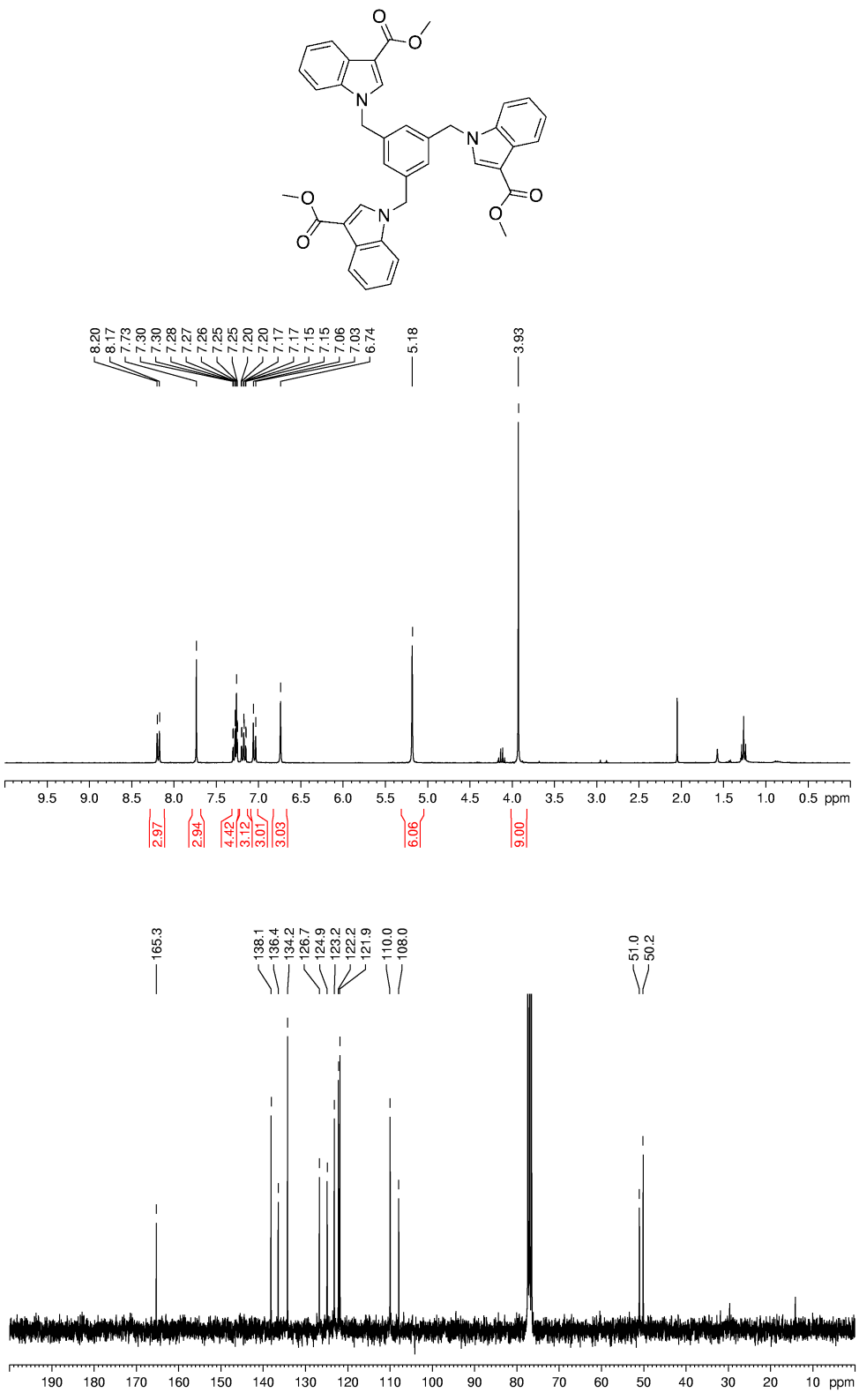
Spectra 7. Compound **2.14**



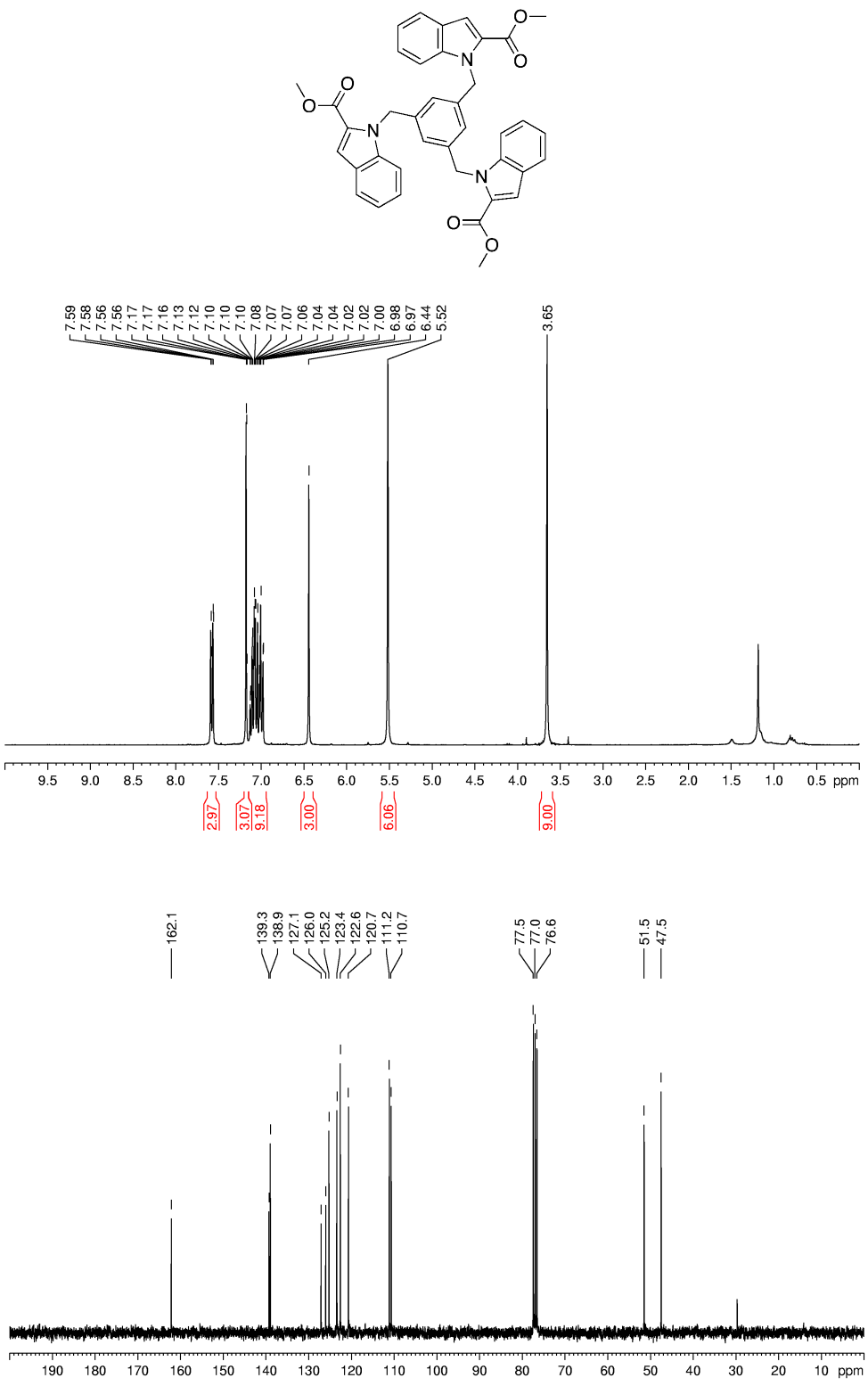




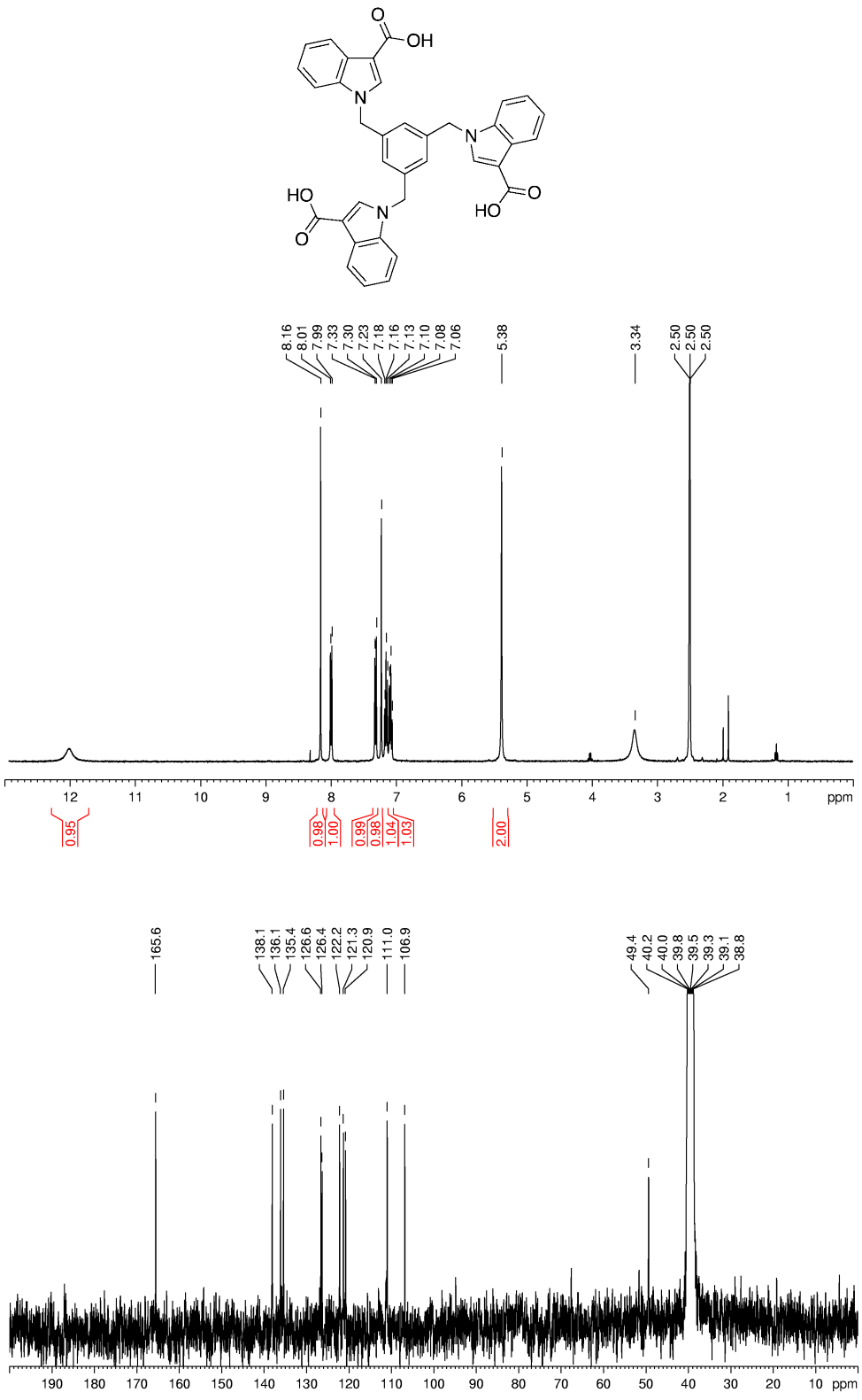
Spectra 8. Compound **2.15**



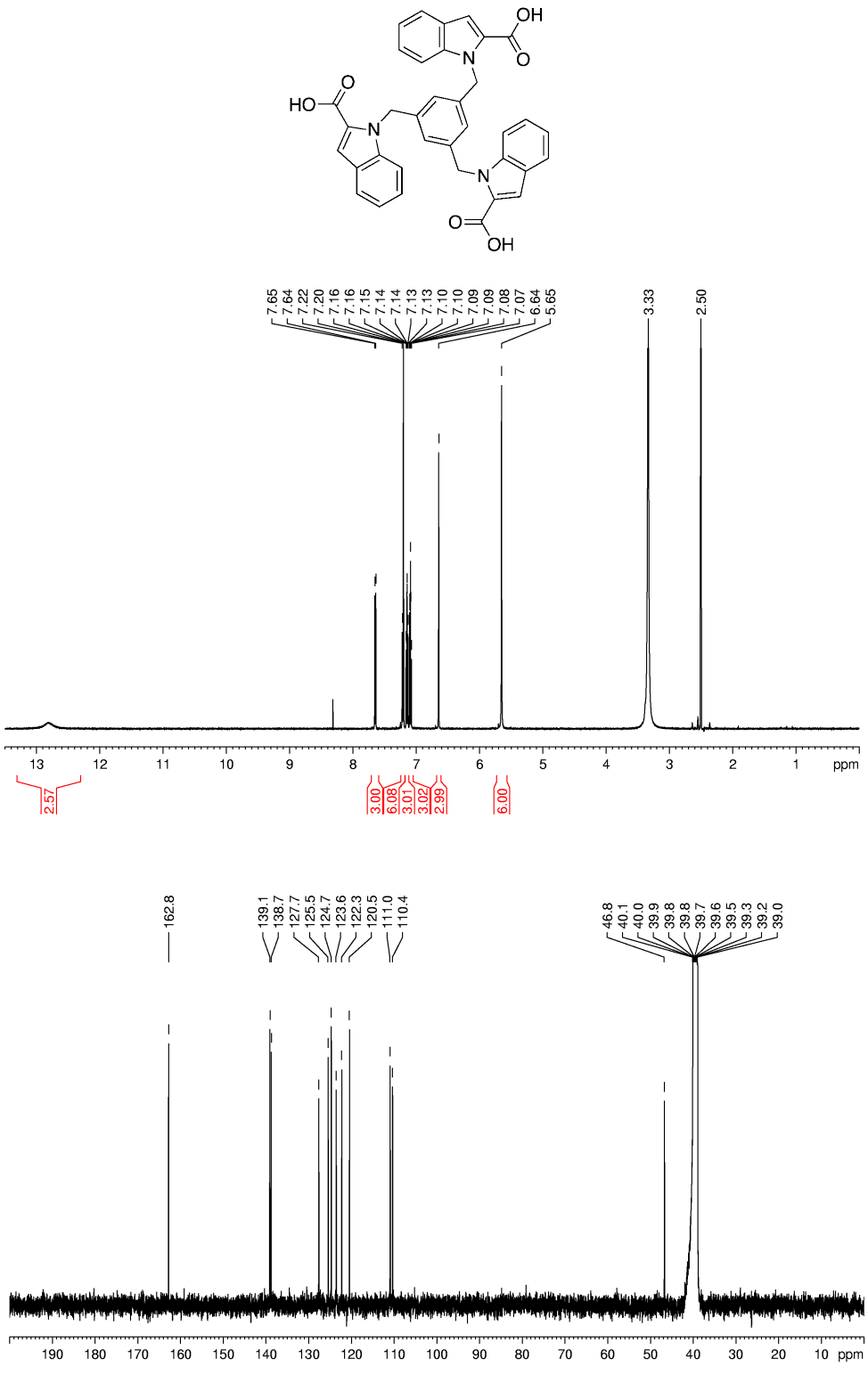
Spectra 9. Compound **2.16**



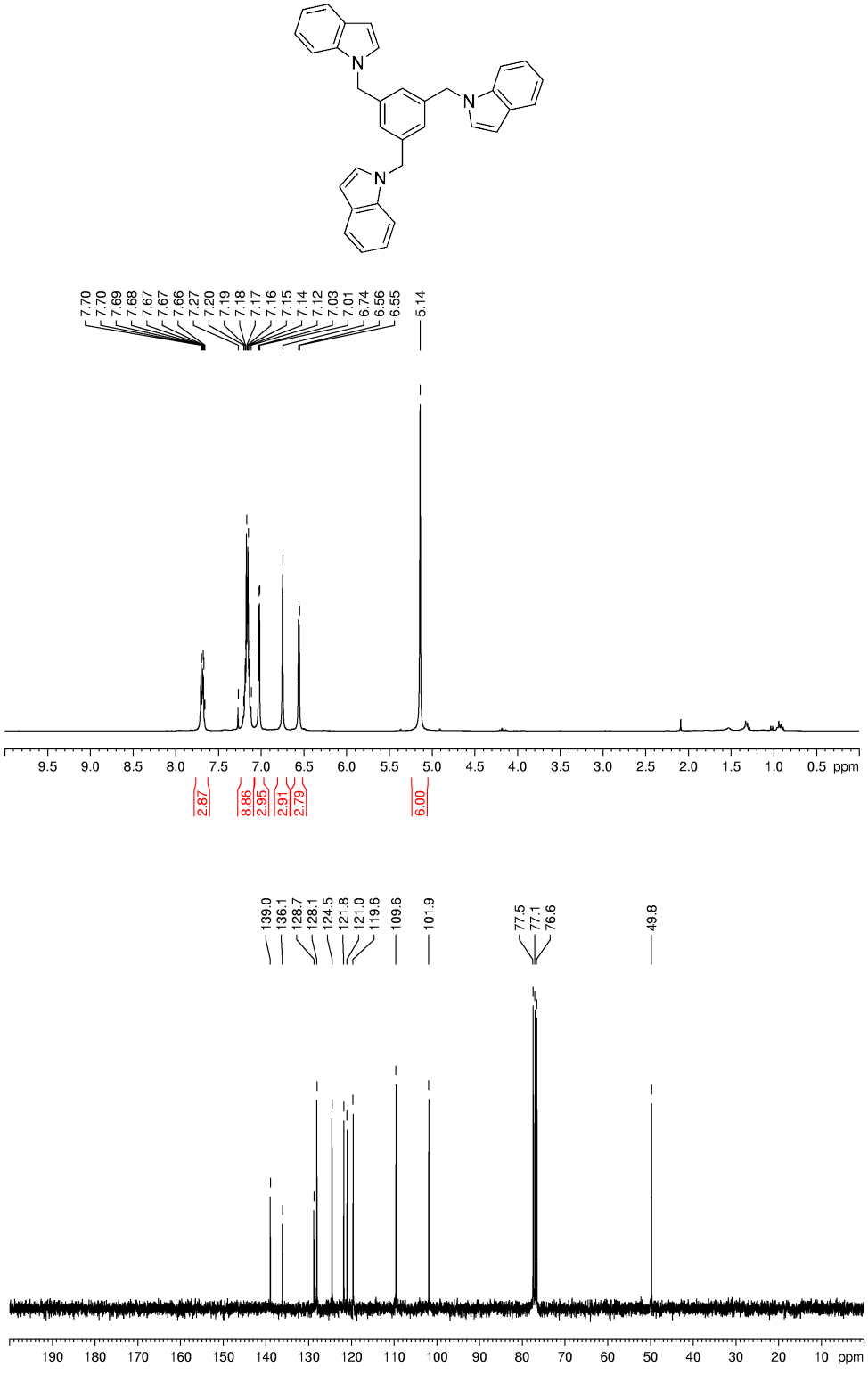
Spectra 10. Compound **2.17**



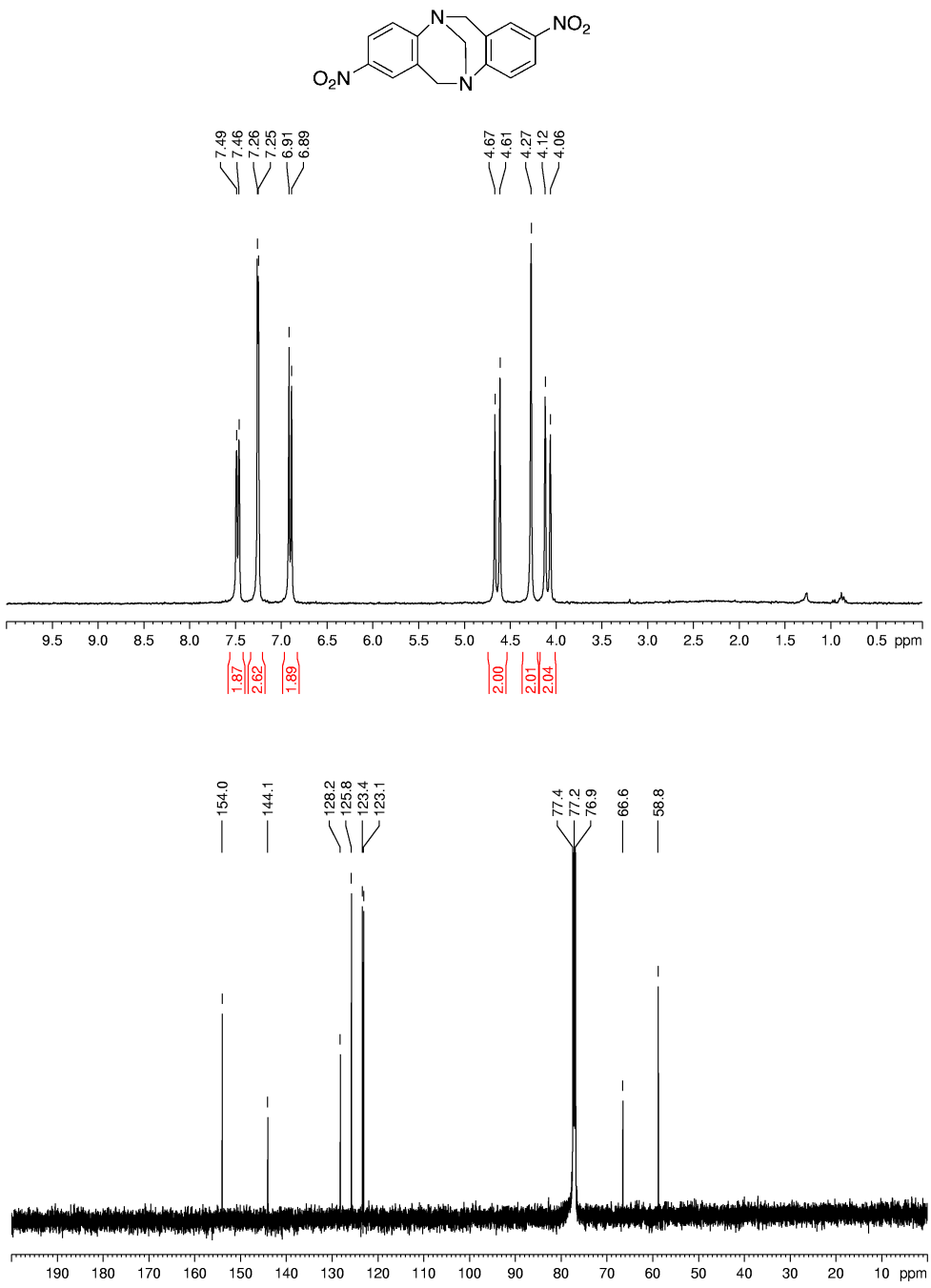
Spectra 11. Compound **2.18**



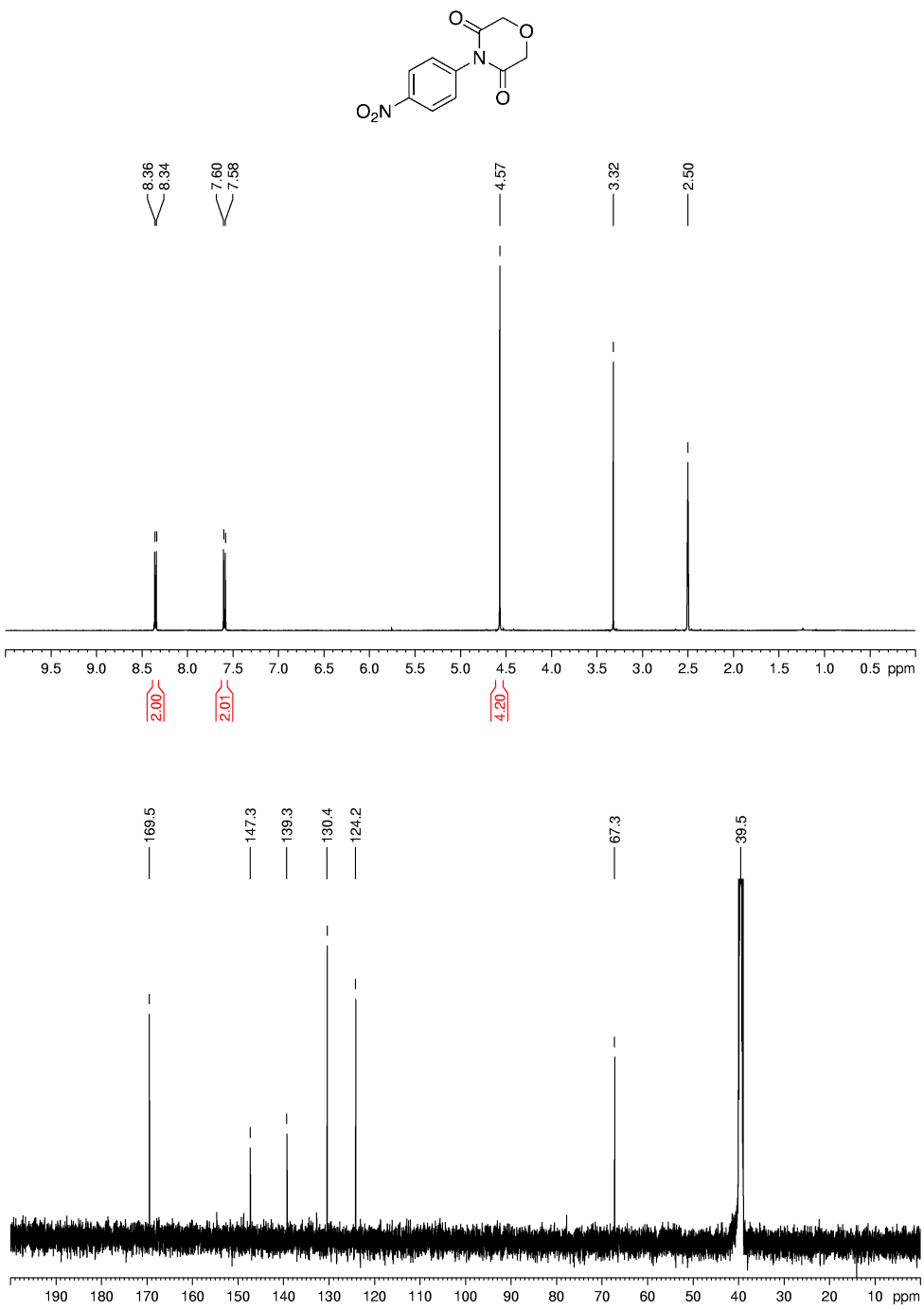
Spectra 12. Compound **2.19**



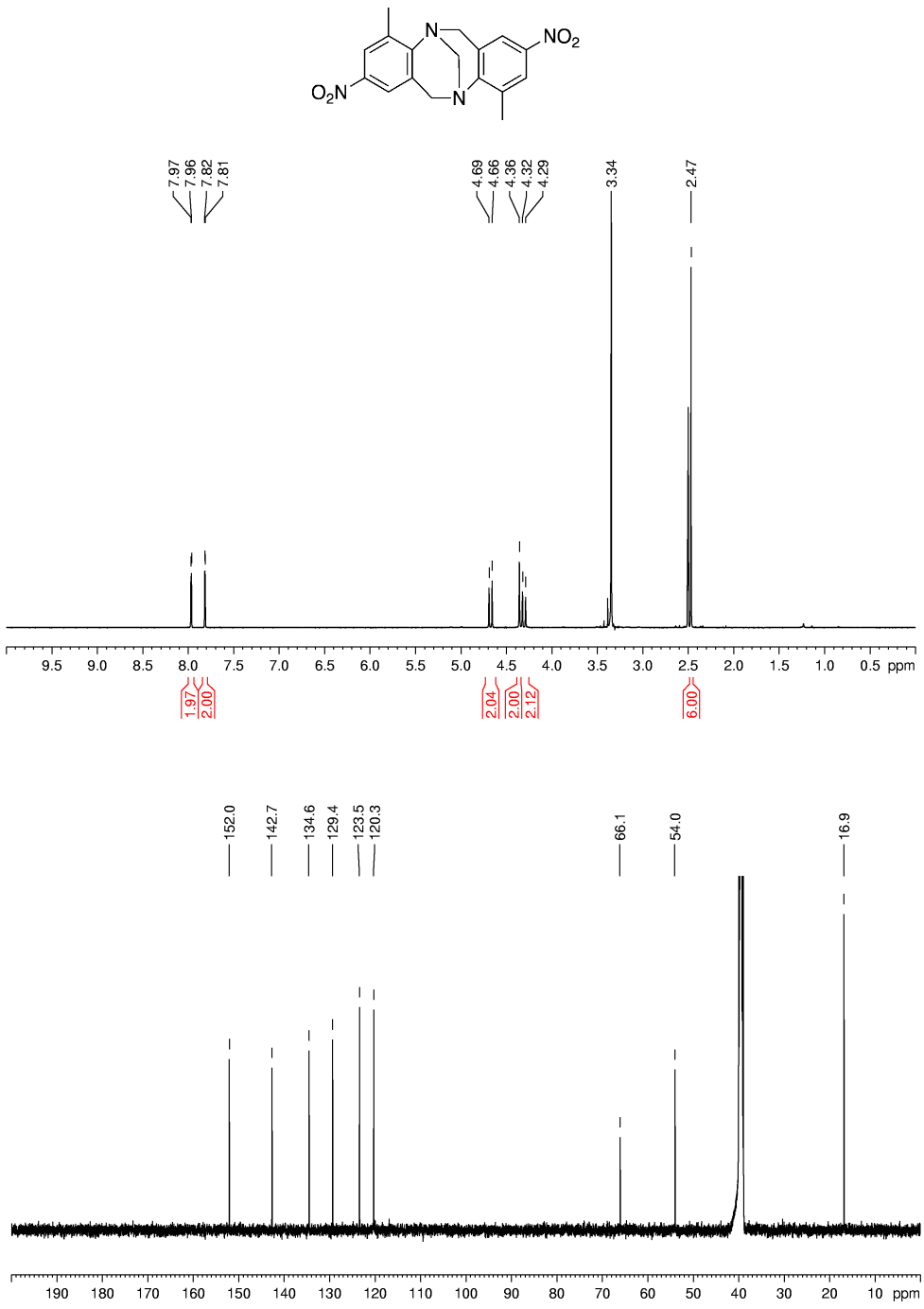
Spectra 13. Compound **2.20**



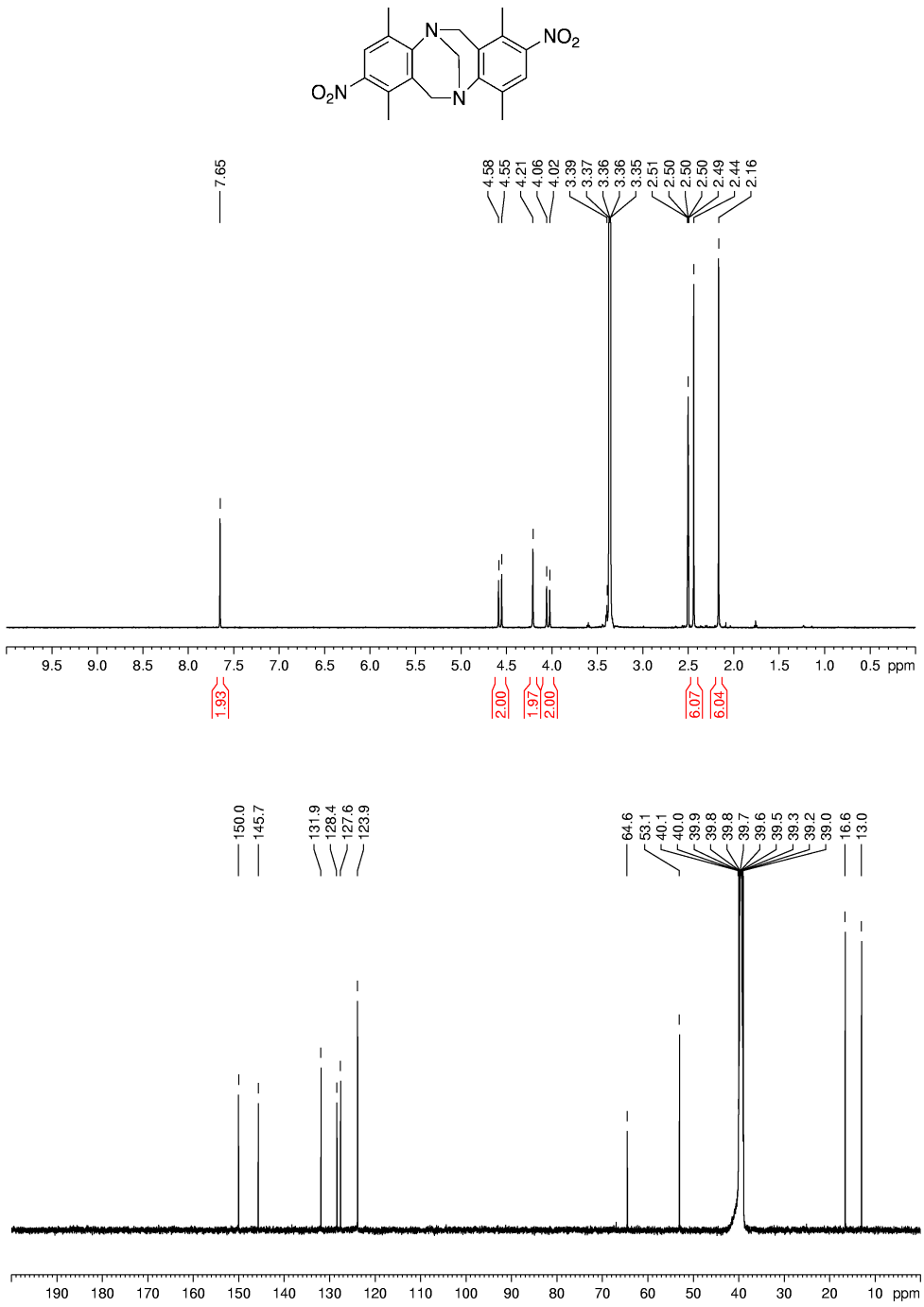
Spectra 14. Compound **3.25**



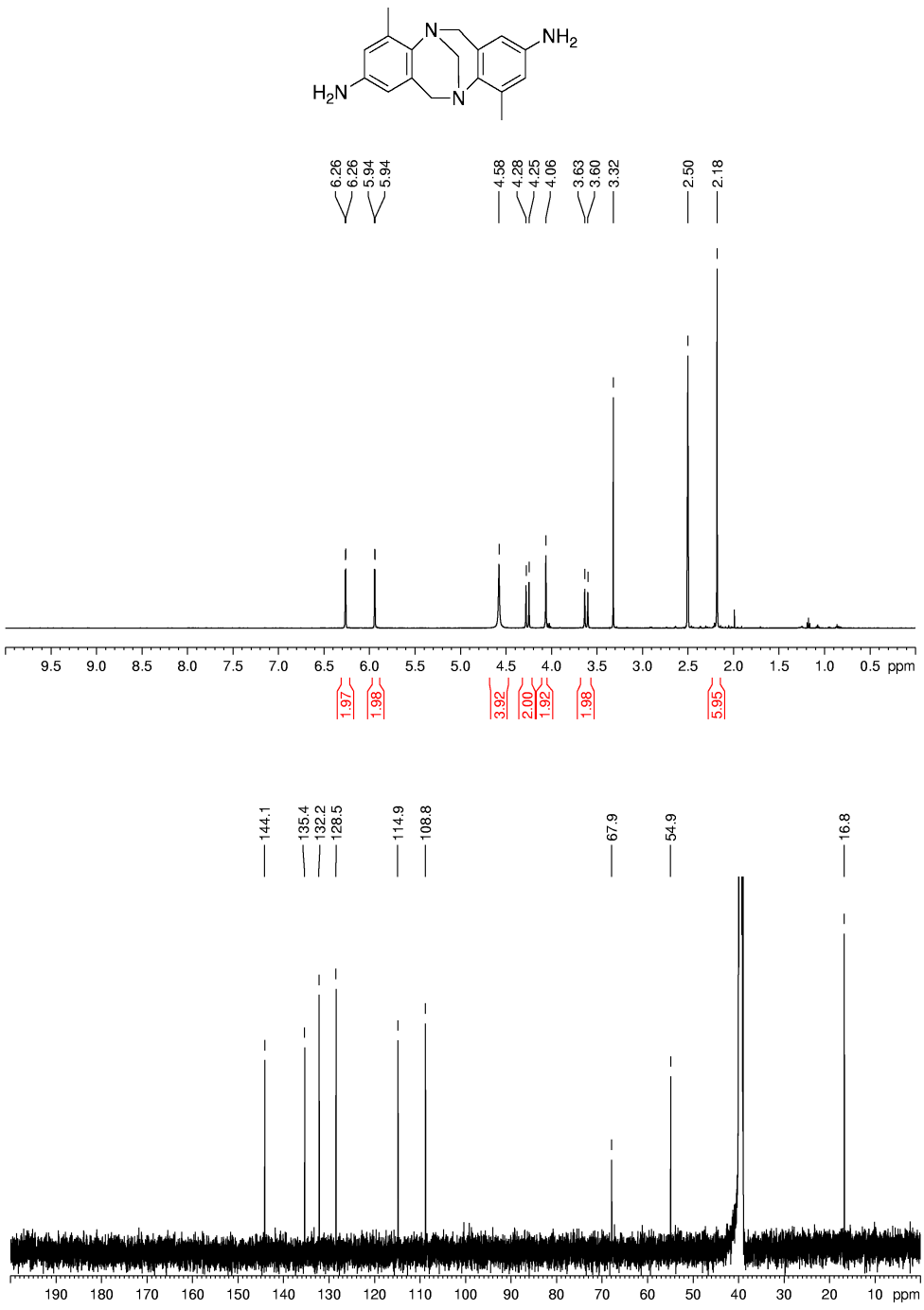
Spectra 15. Compound **3.26**



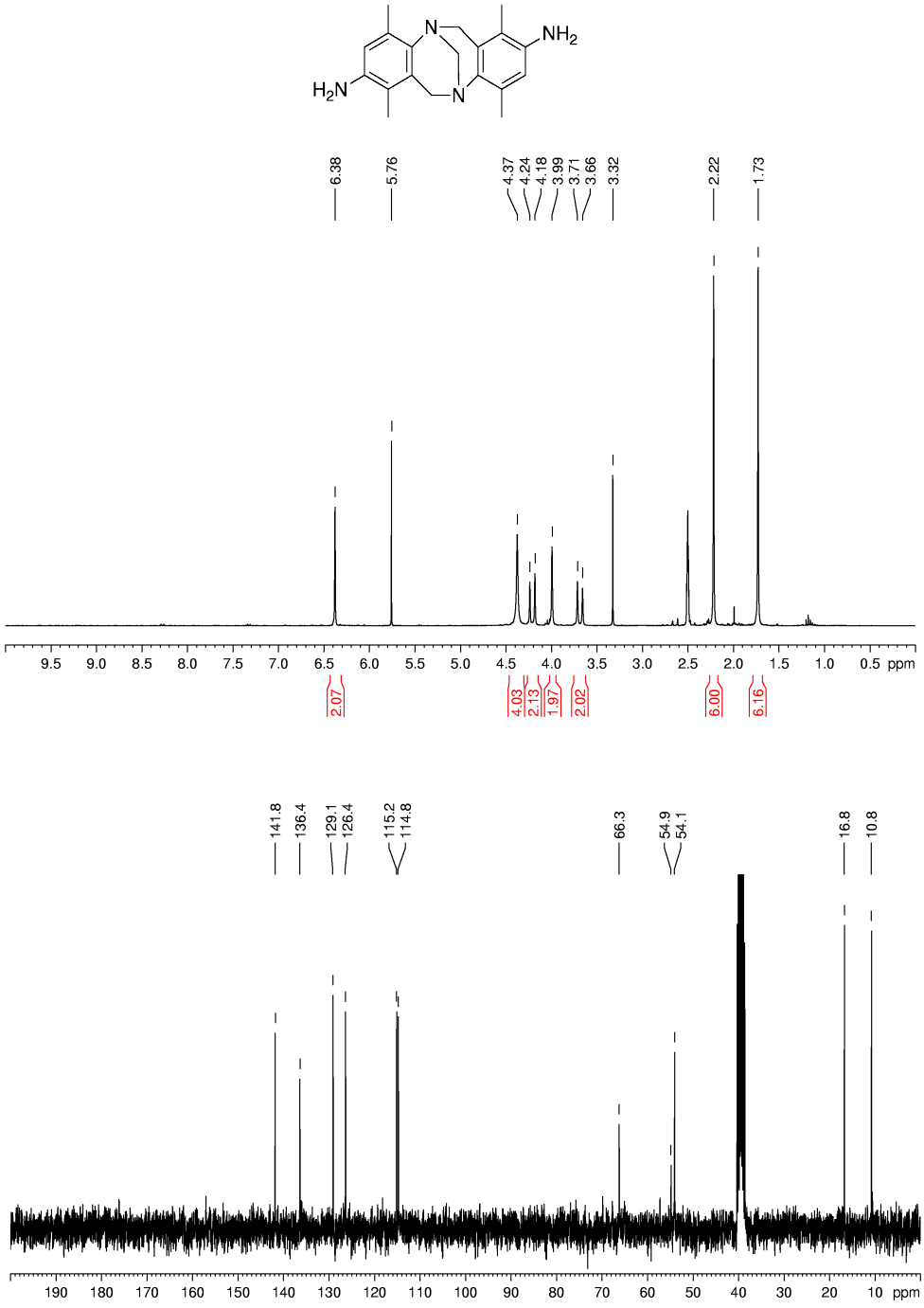
Spectra 16. Compound **3.29**



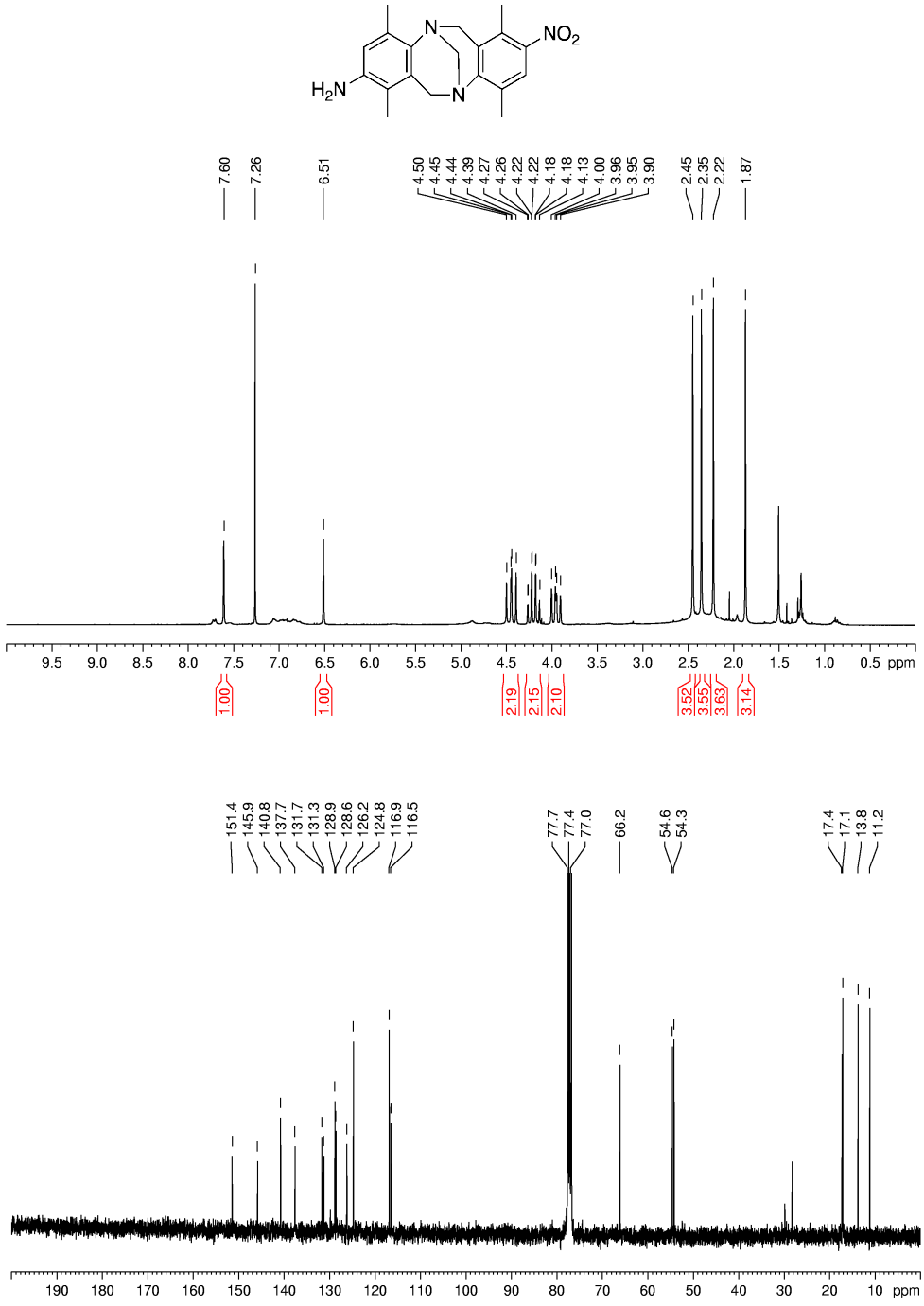
Spectra 17. Compound **3.30**



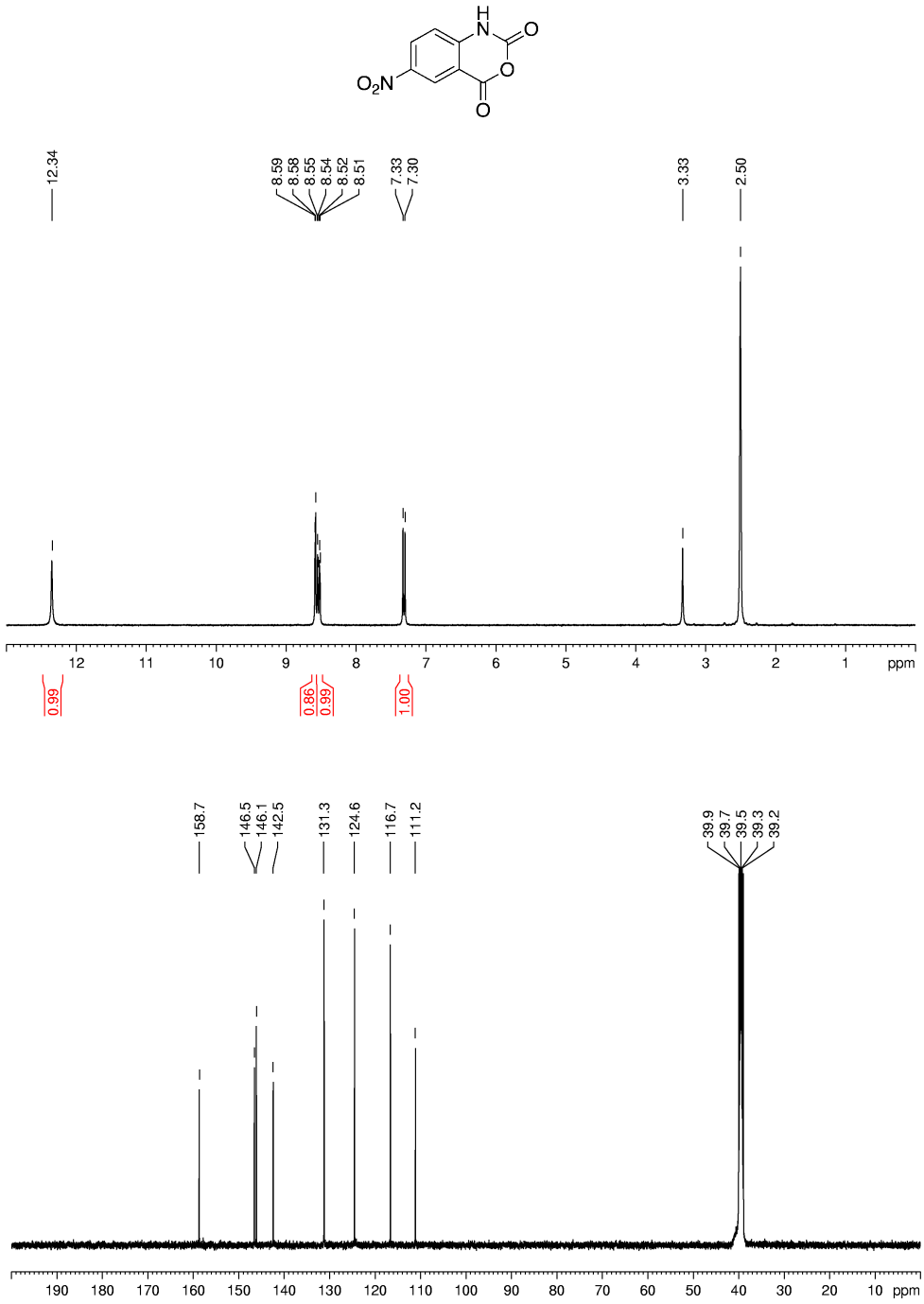
Spectra 18. Compound **3.31**



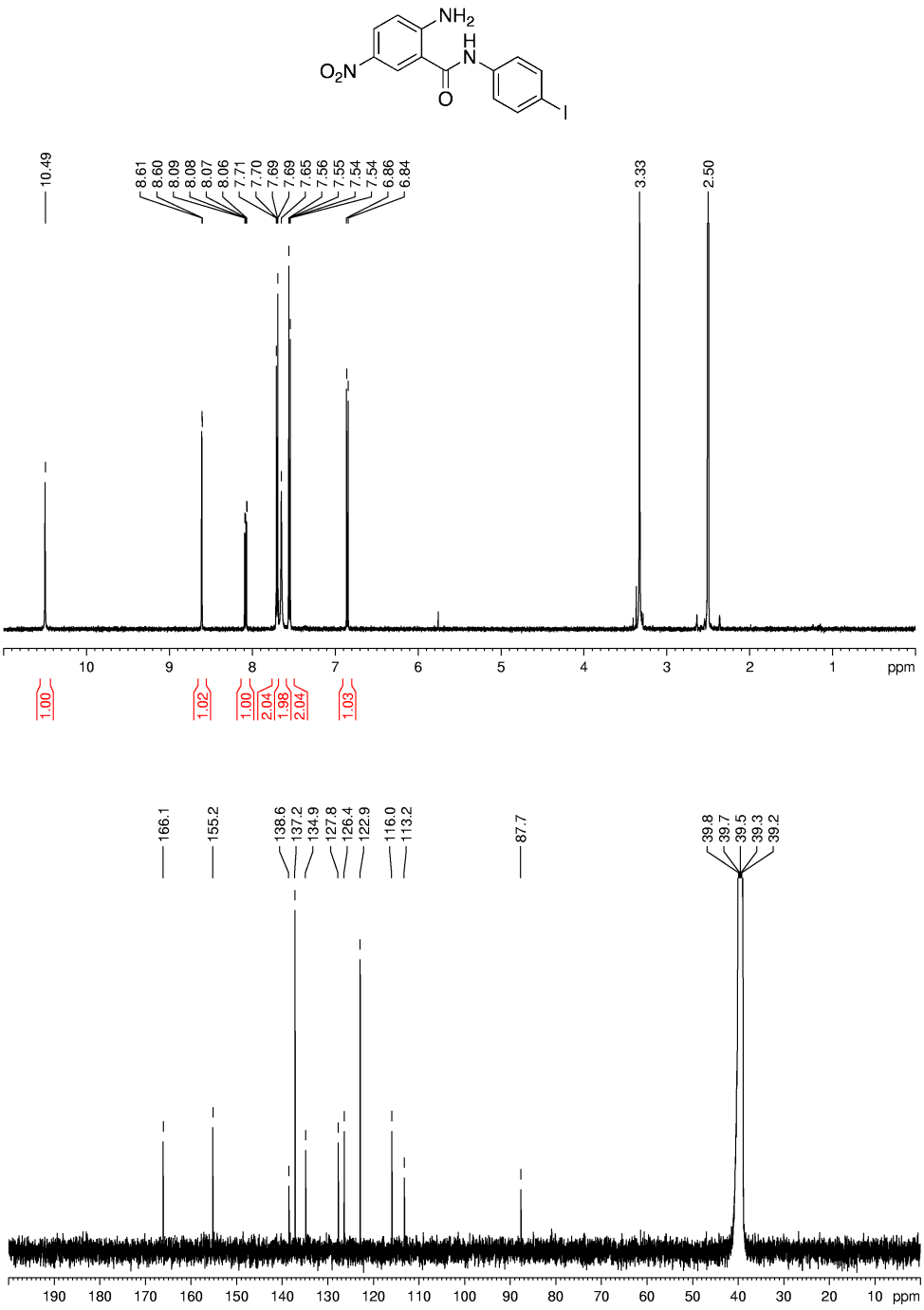
Spectra 19. Compound **3.32**



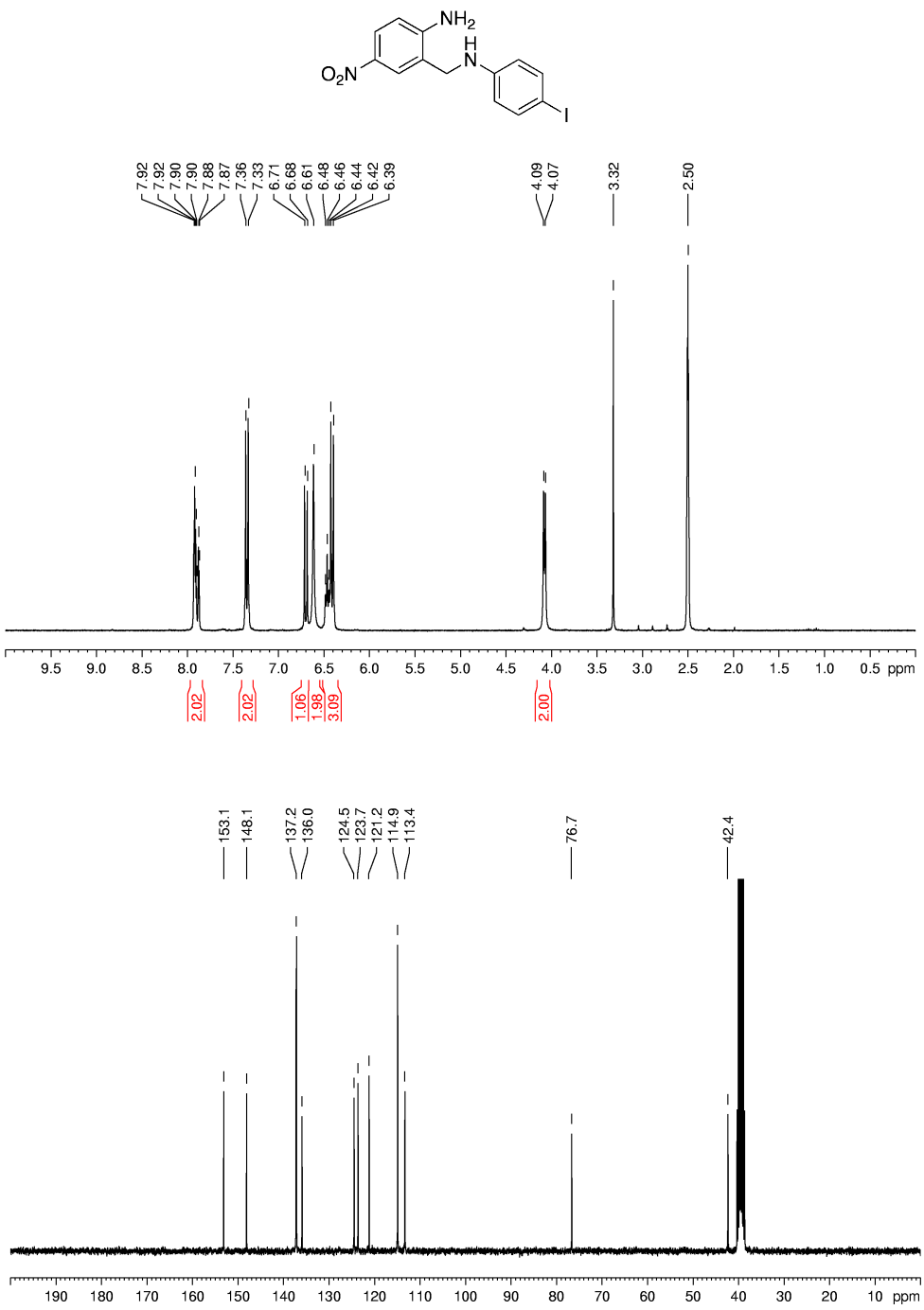
Spectra 20. Compound **3.33**



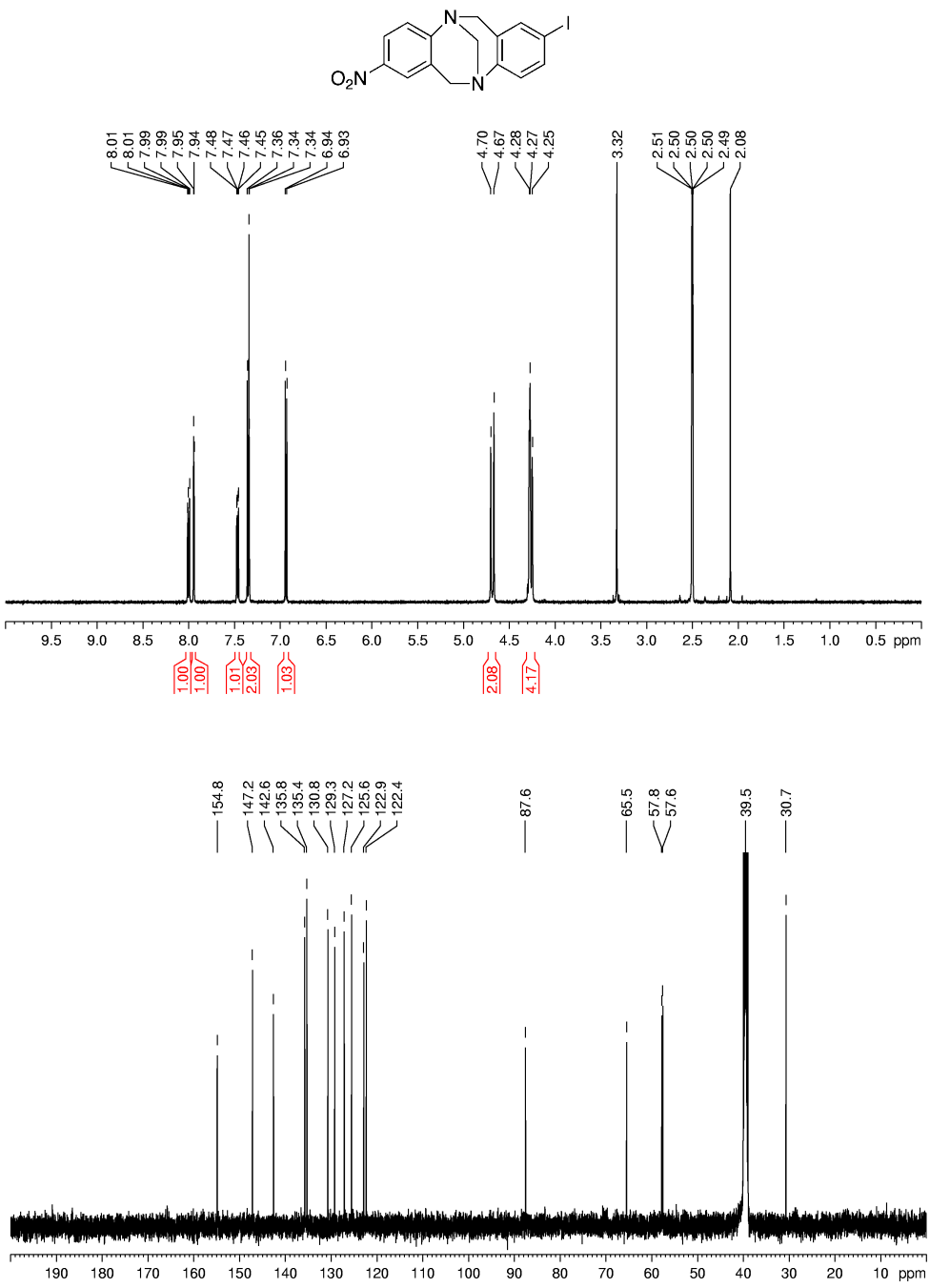
Spectra 21. Compound **3.36**



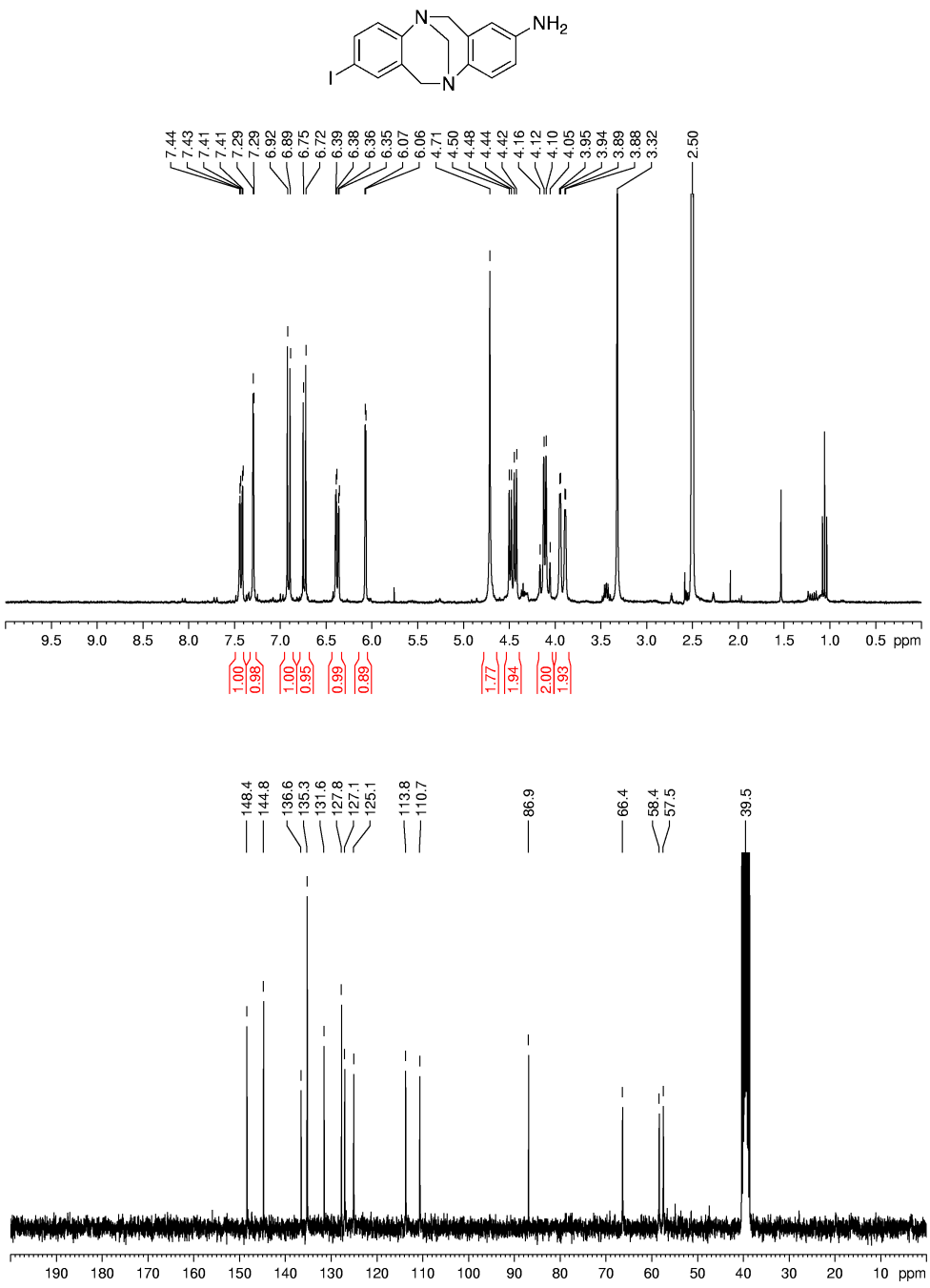
Spectra 22. Compound **3.38**



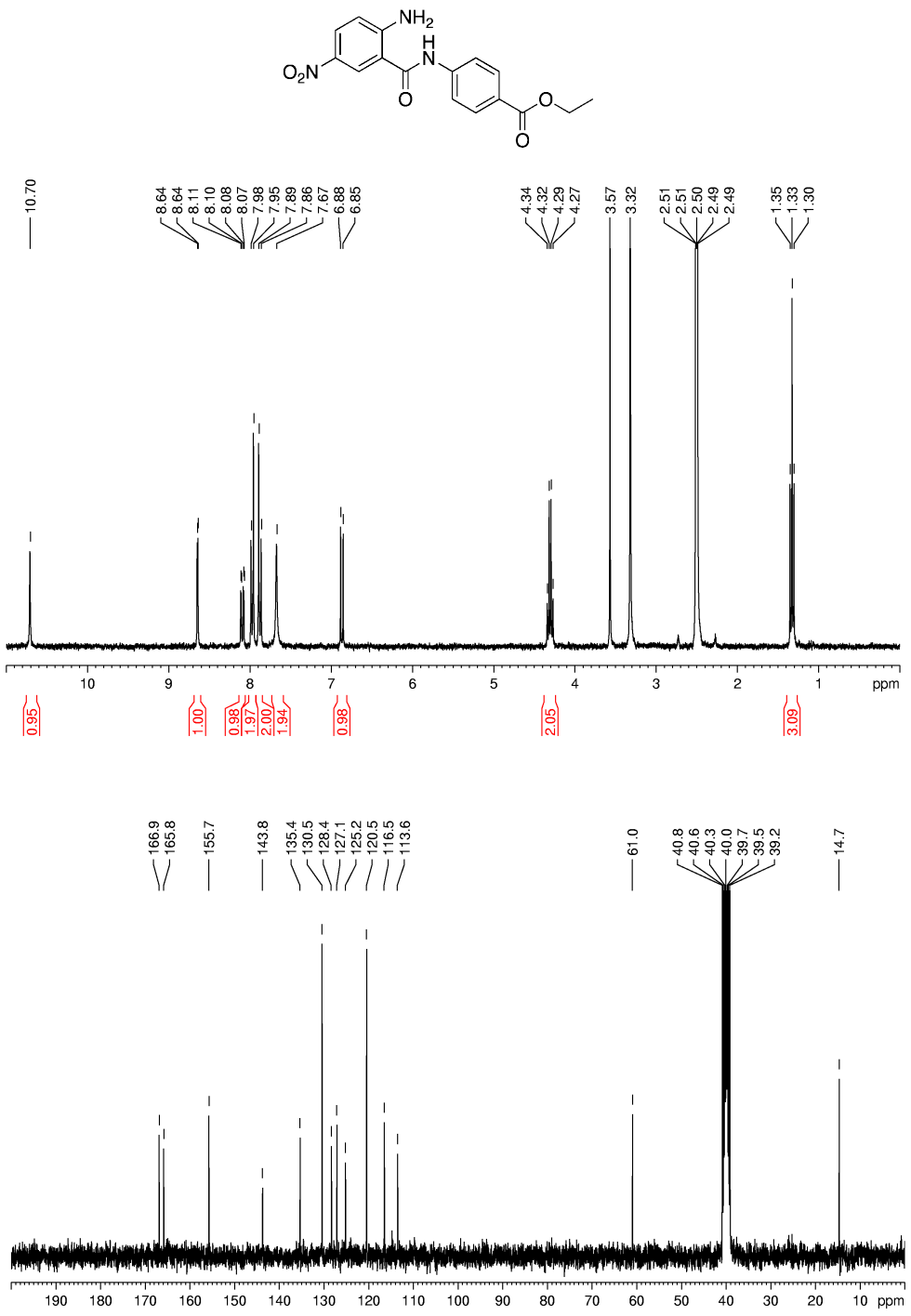
Spectra 23. Compound **3.39**



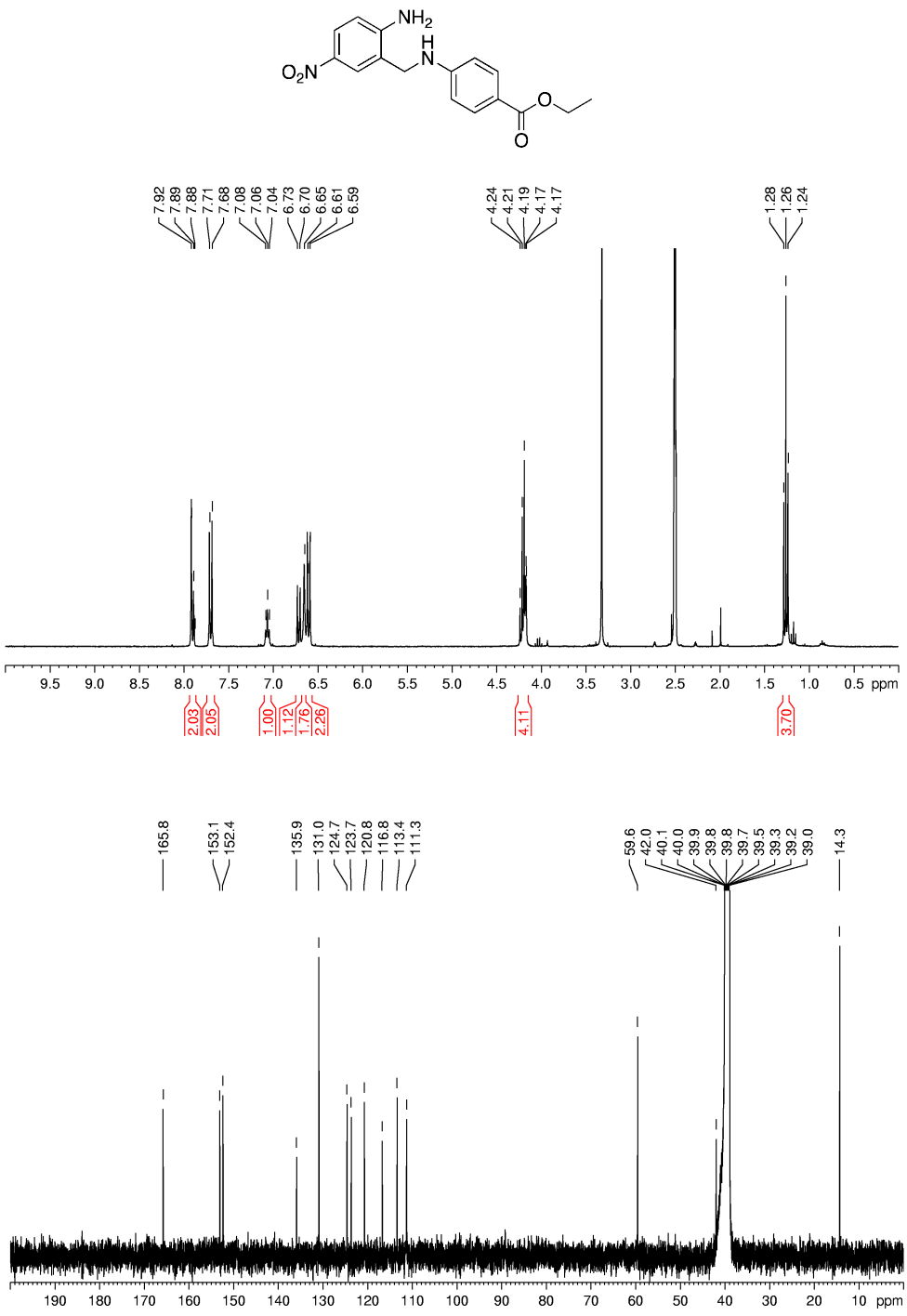
Spectra 24. Compound **3.40**



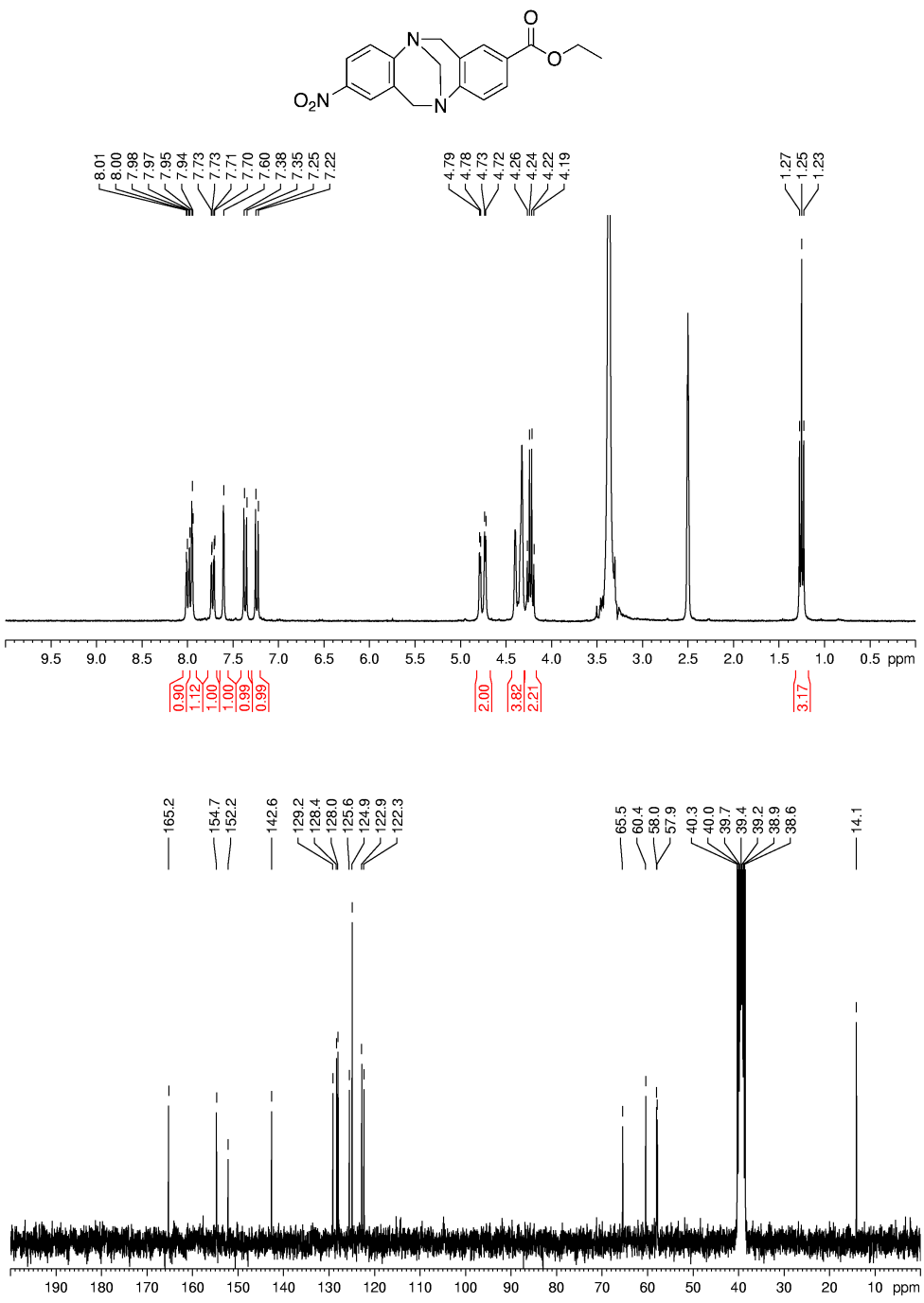
Spectra 25. Compound **3.41**



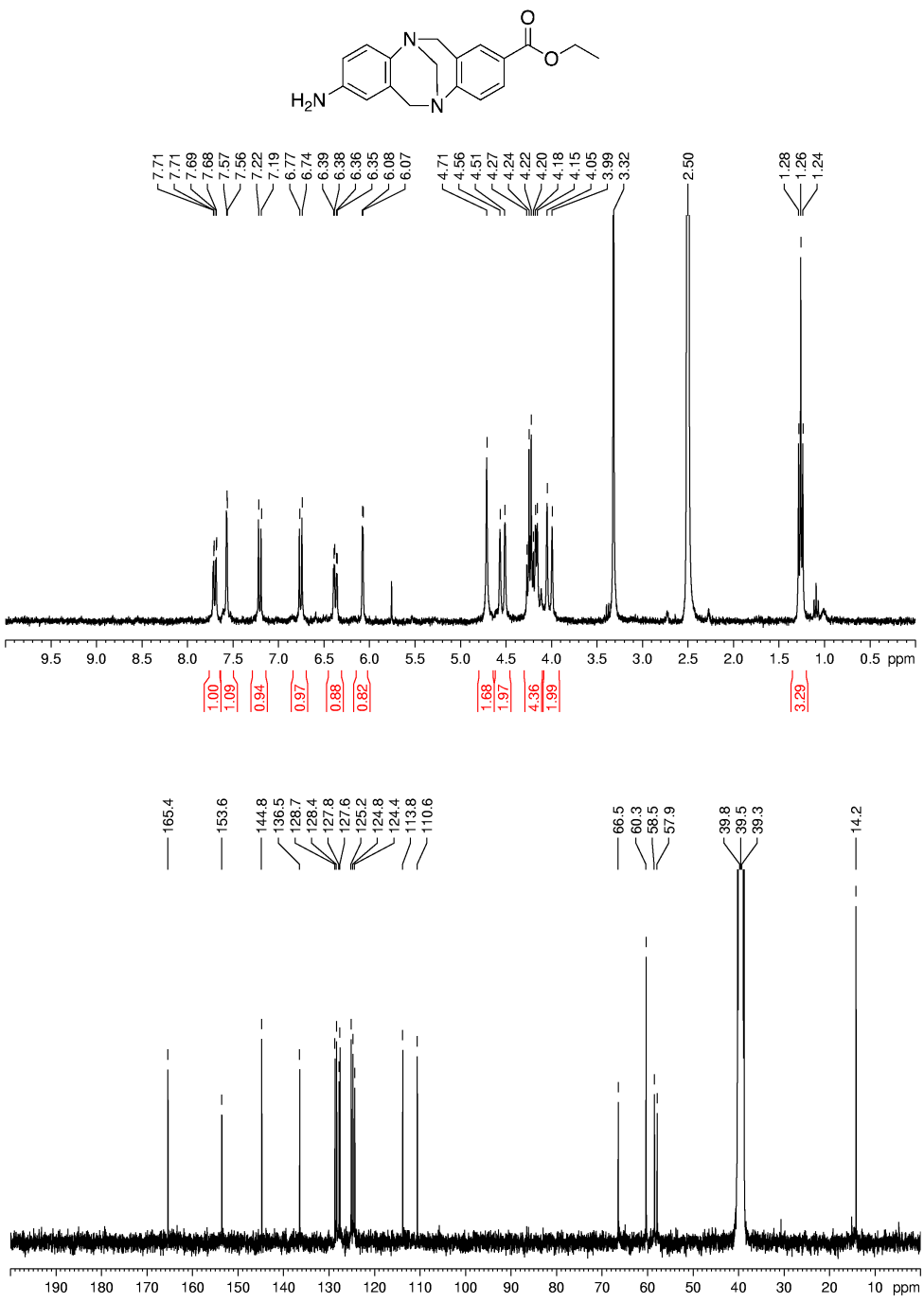
Spectra 26. Compound **3.43**



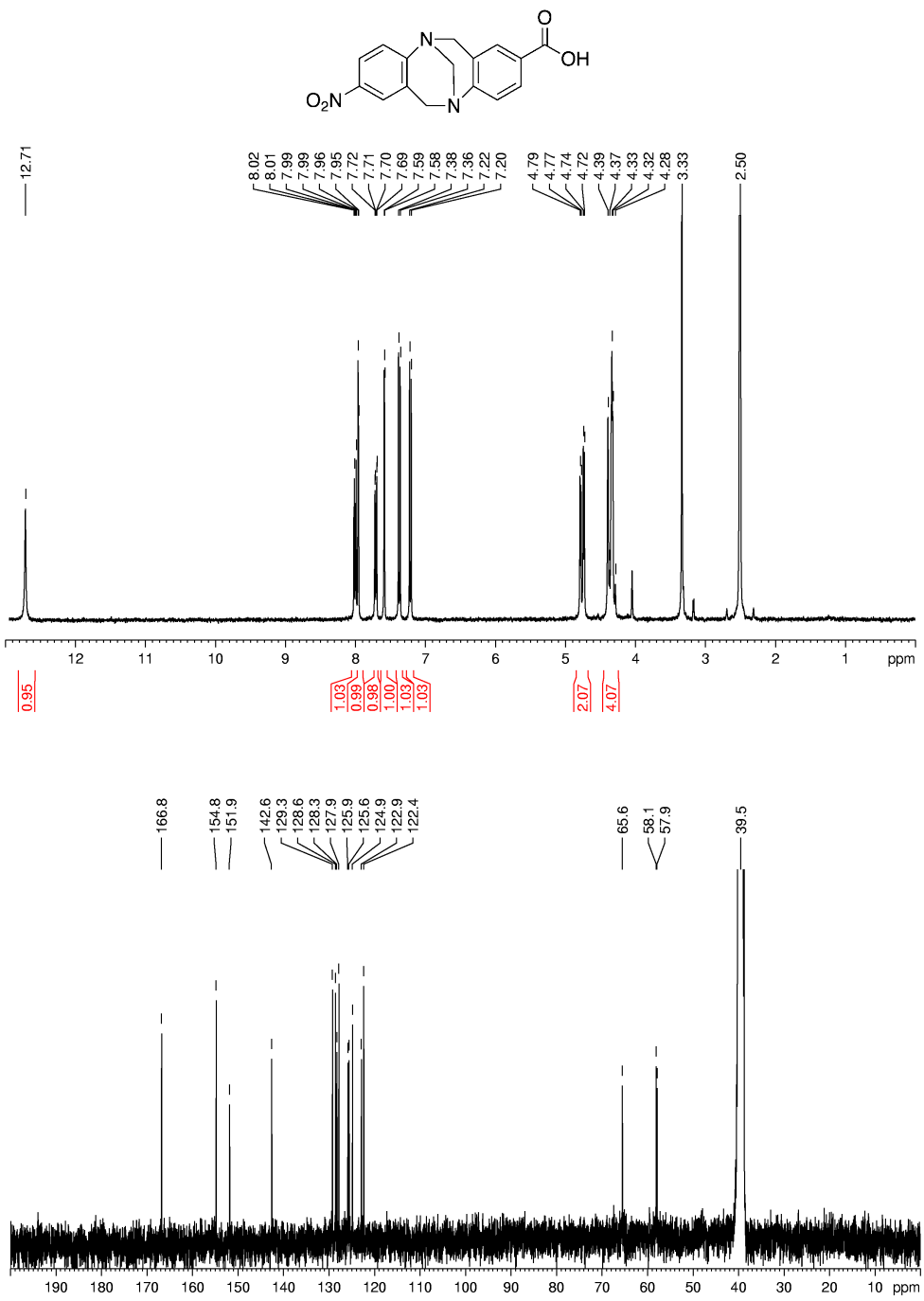
Spectra 27. Compound **3.44**



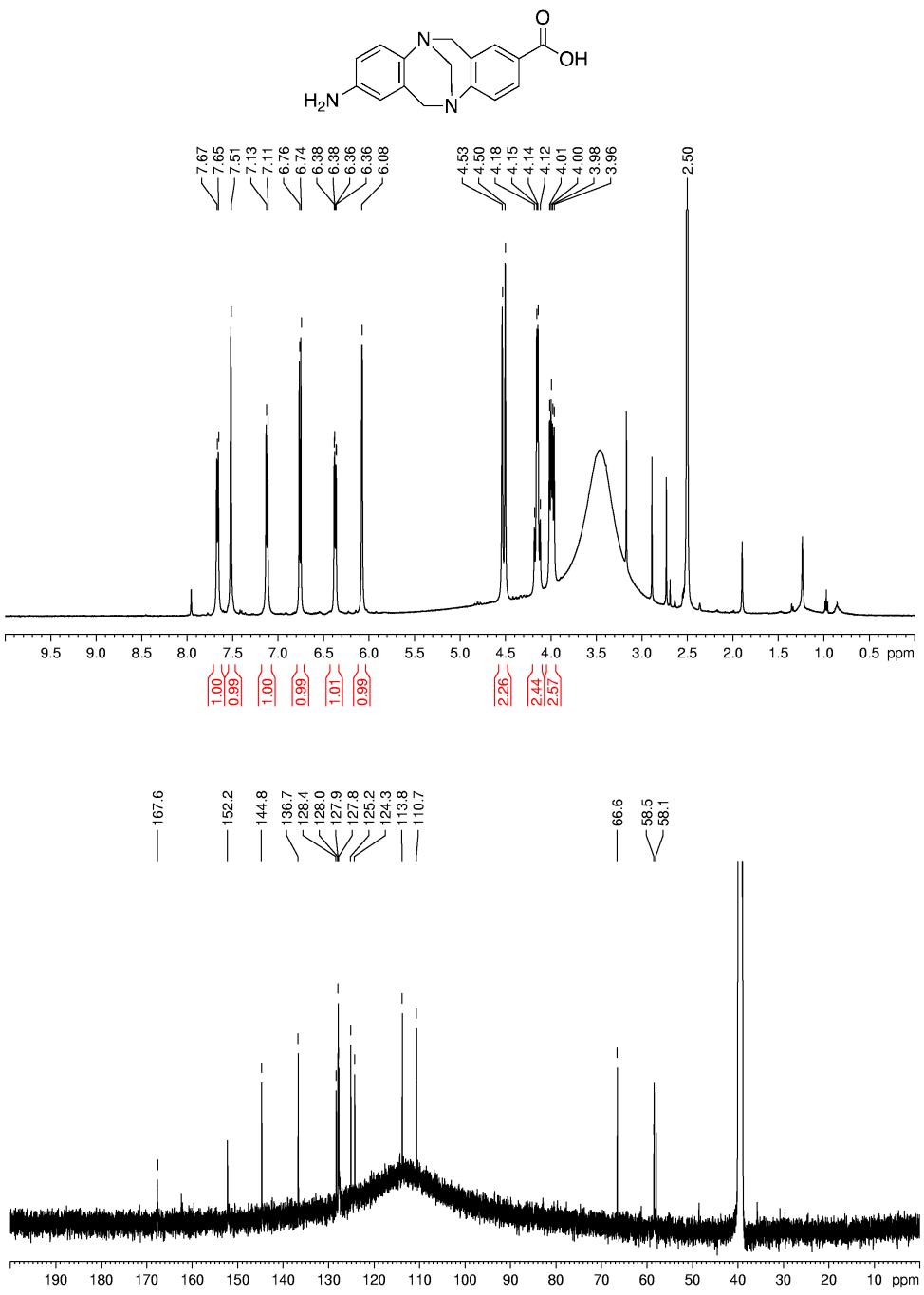
Spectra 28. Compound **3.45**



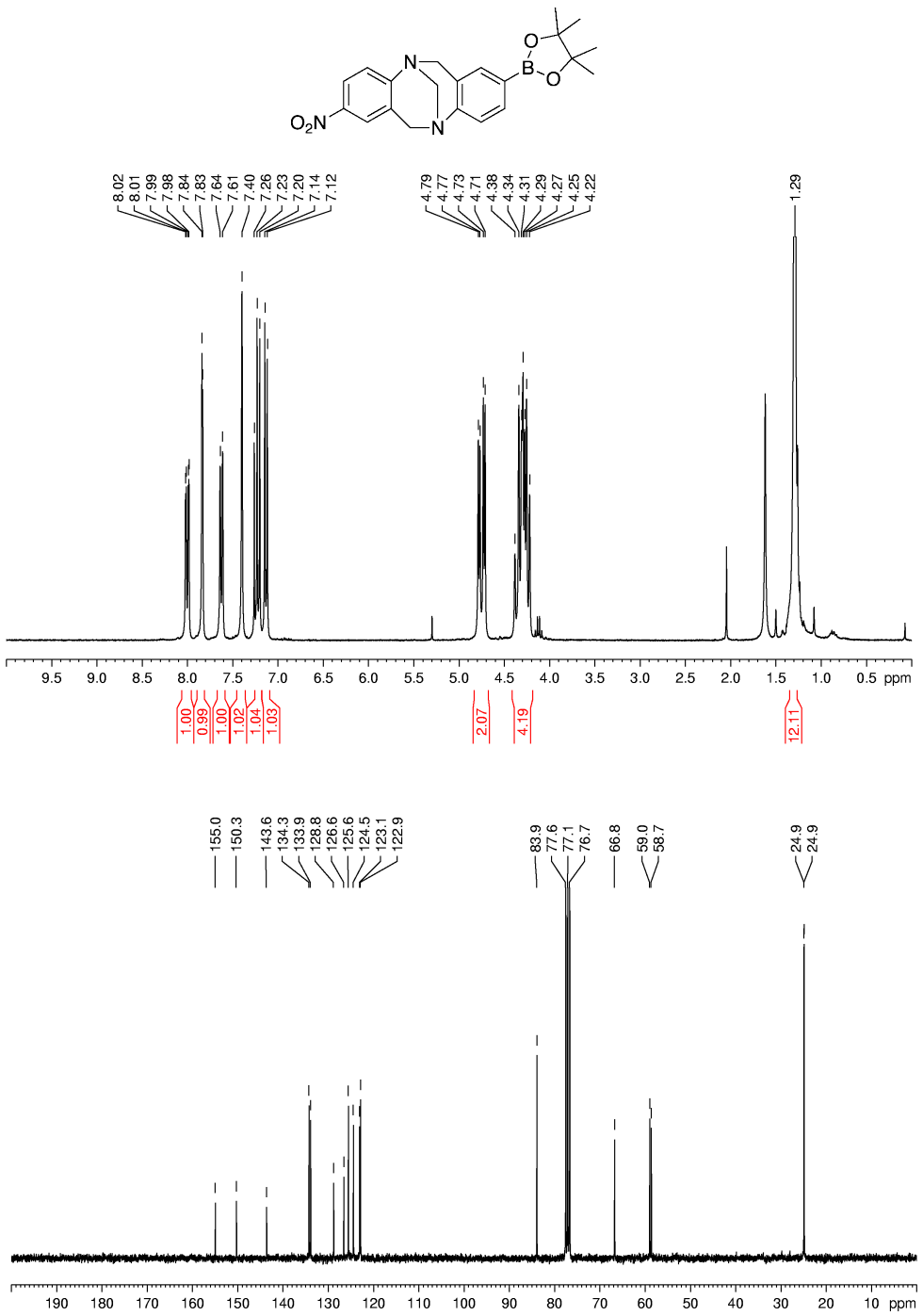
Spectra 29. Compound **3.46**



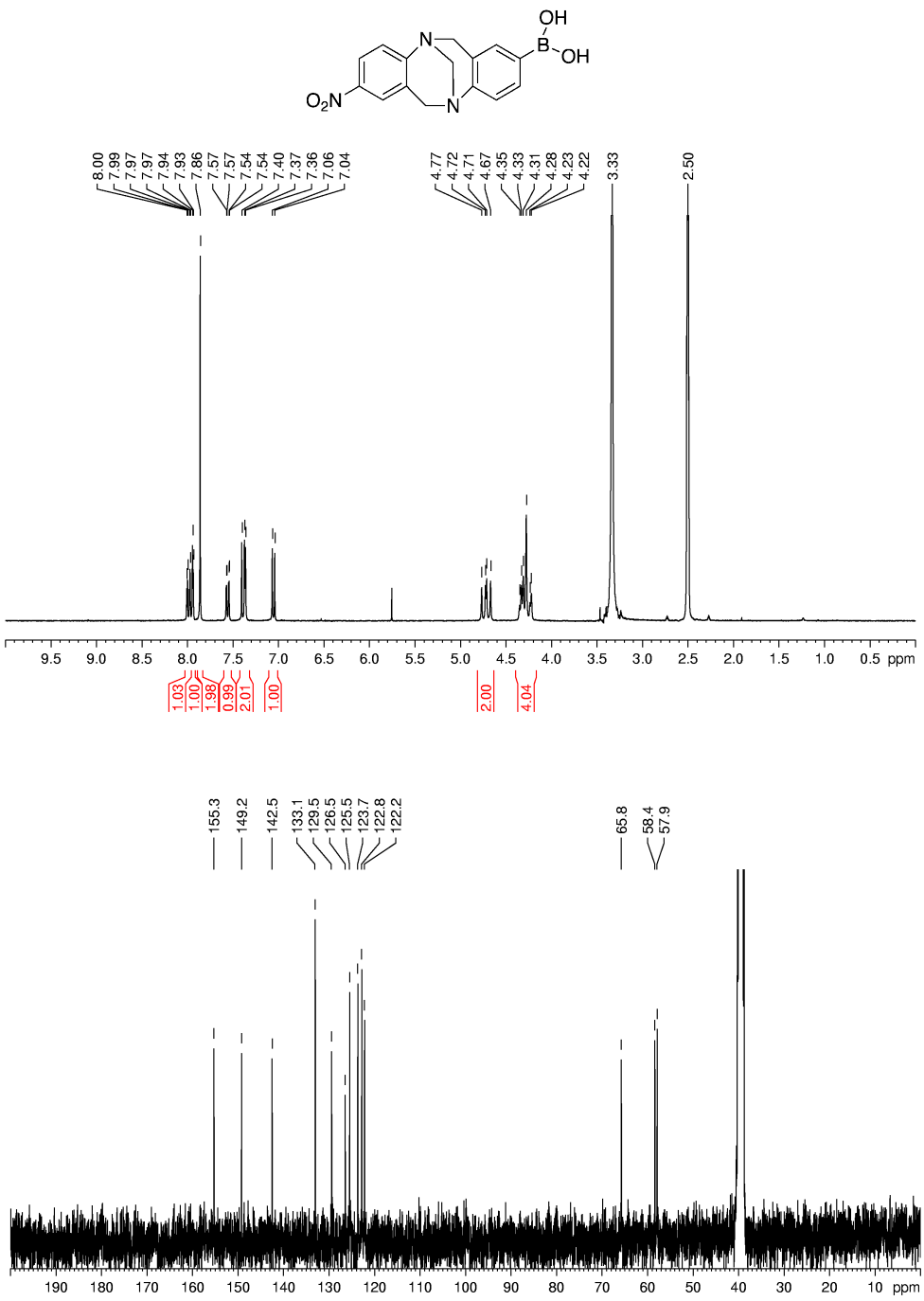
Spectra 30. Compound **3.47**



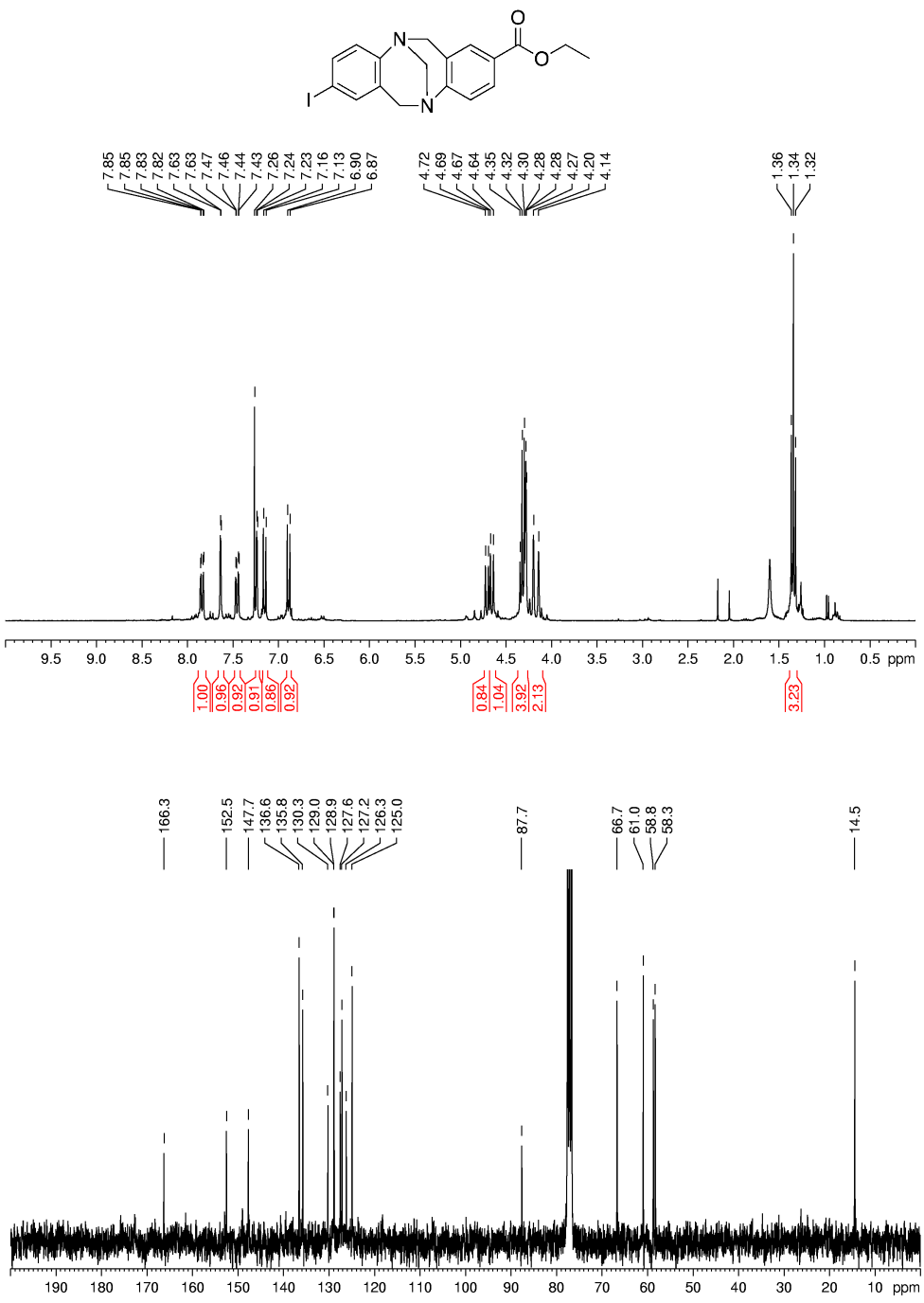
Spectra 31. Compound **3.48**



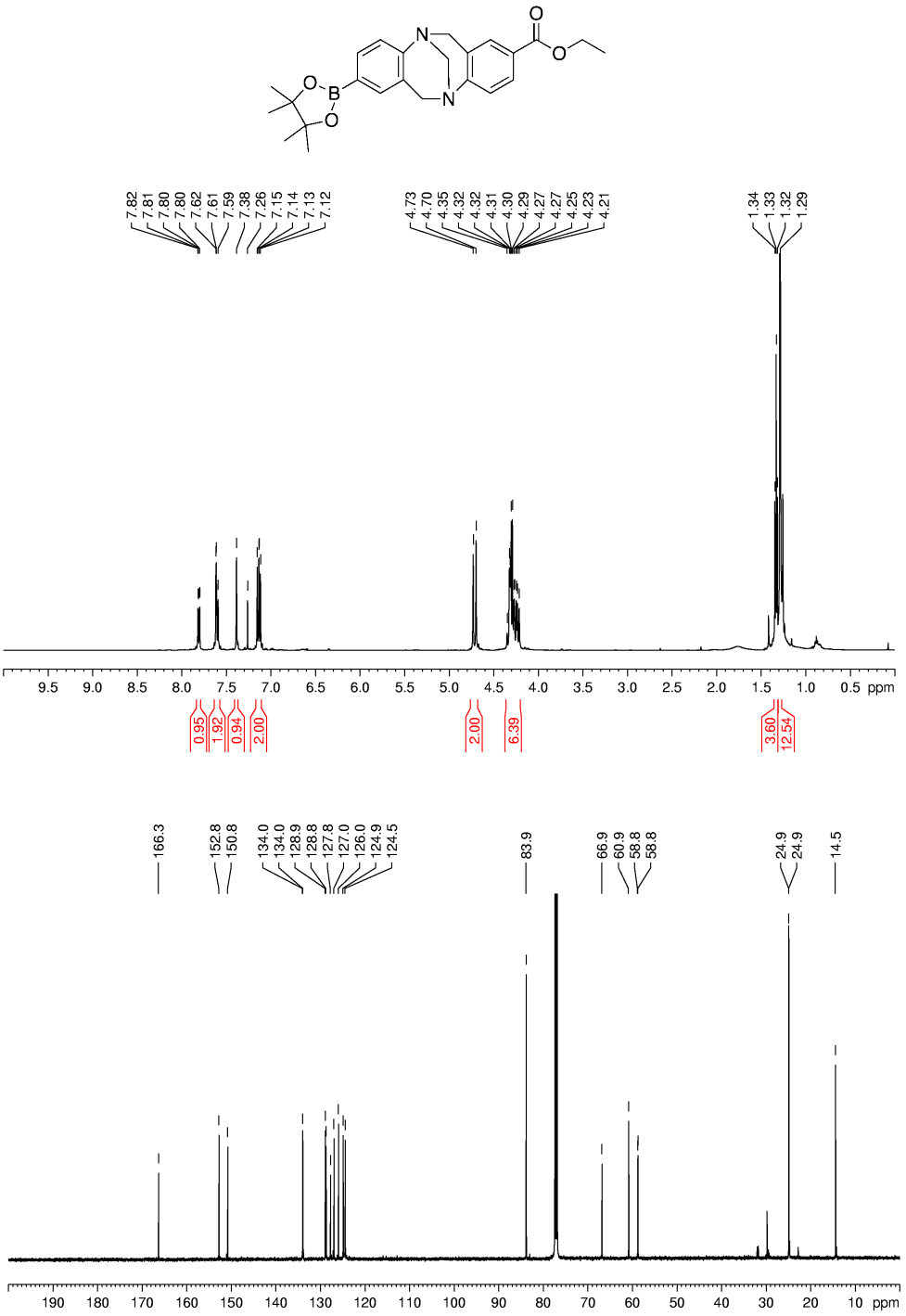
Spectra 32. Compound **3.50**



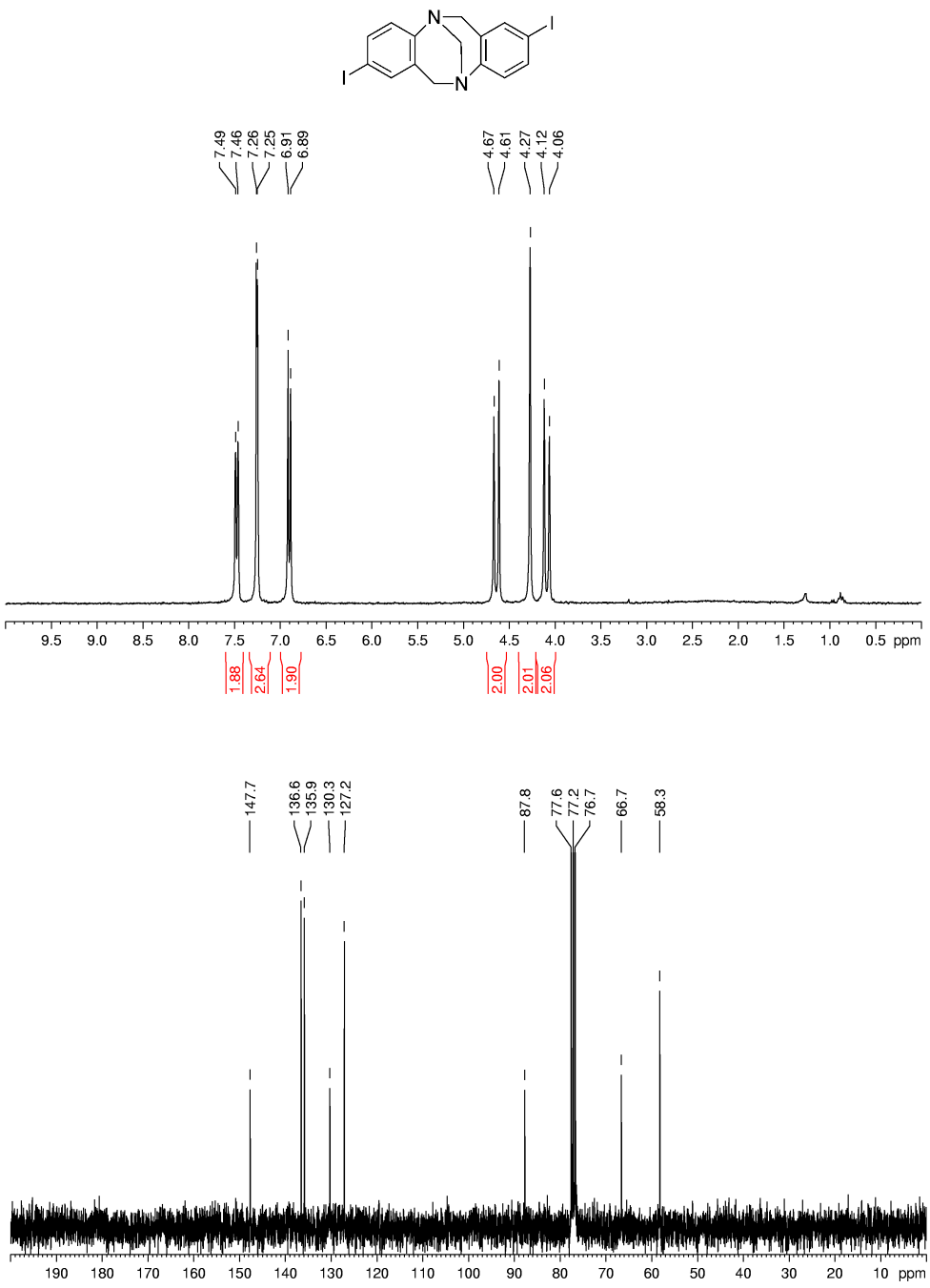
Spectra 33. Compound **3.51**



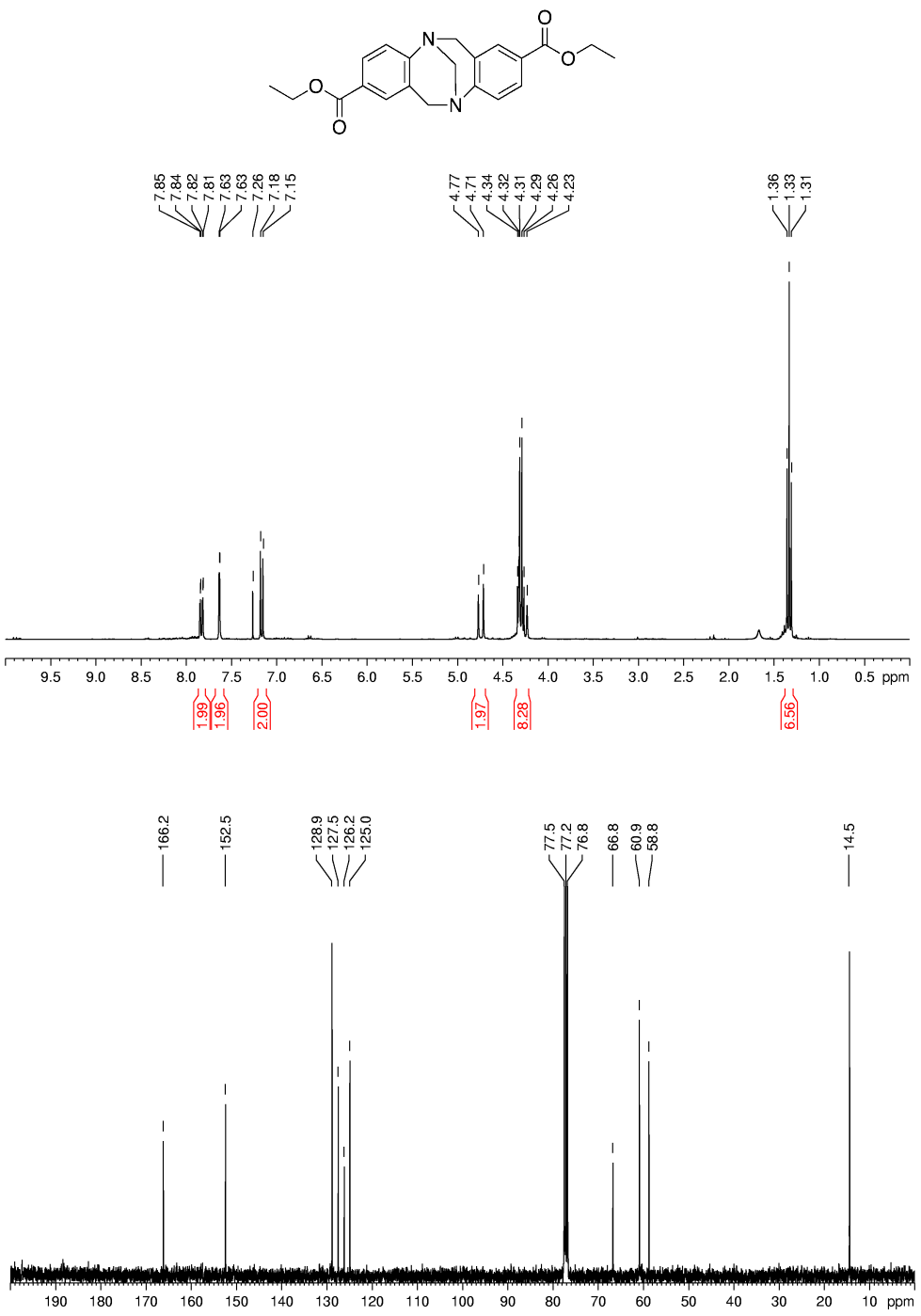
Spectra 34. Compound **3.52**



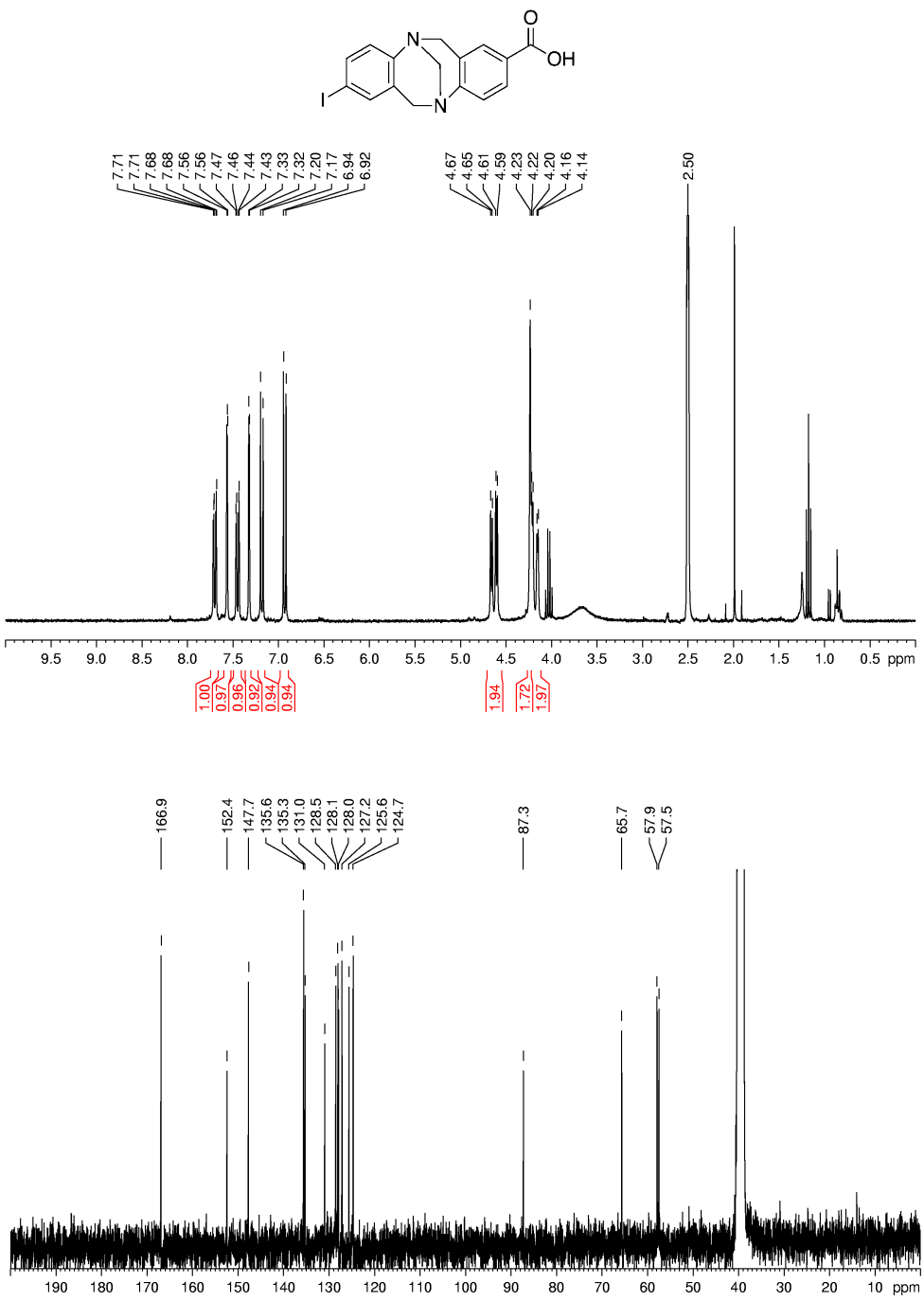
Spectra 35. Compound **3.53**



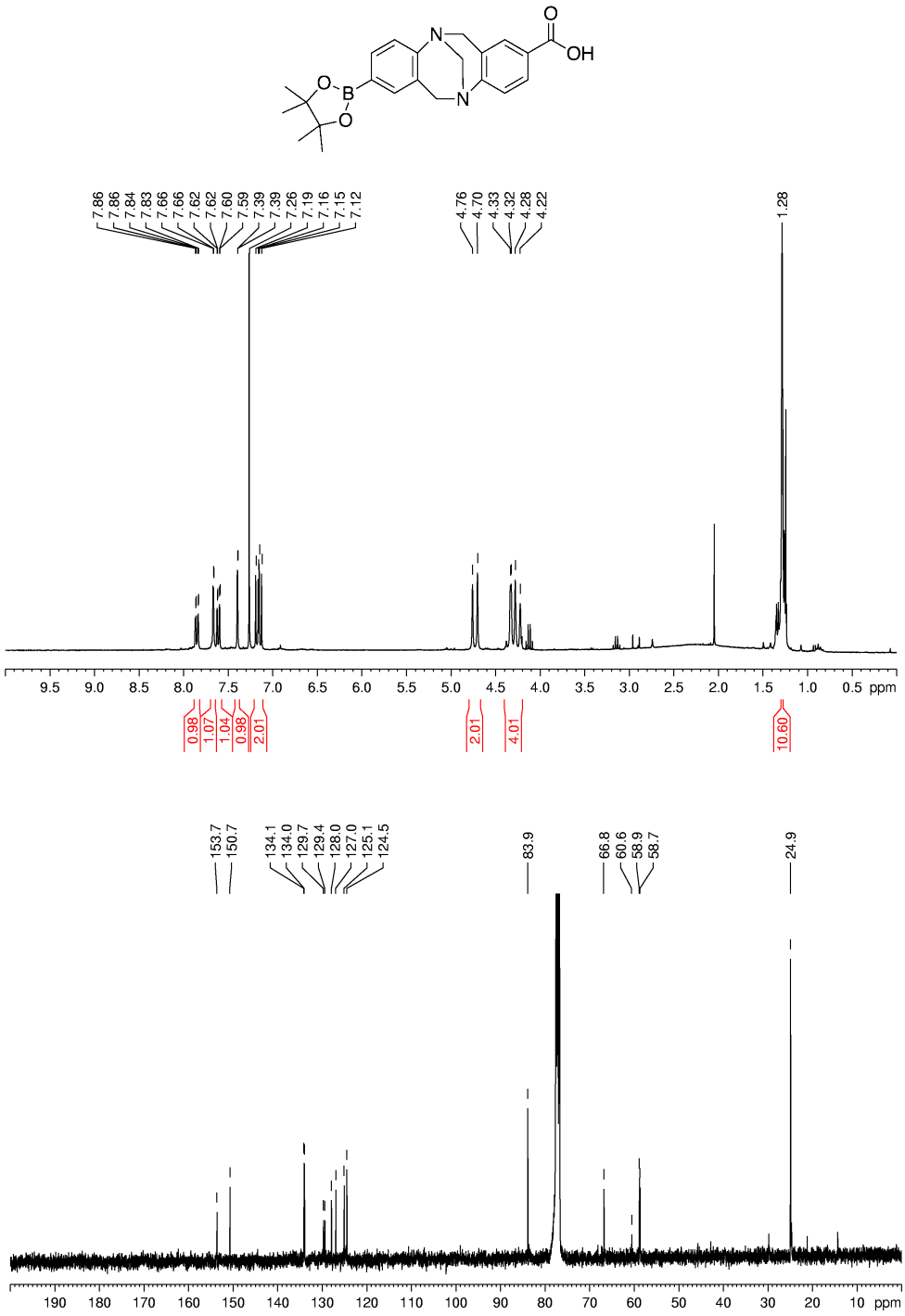
Spectra 36. Compound **3.54**



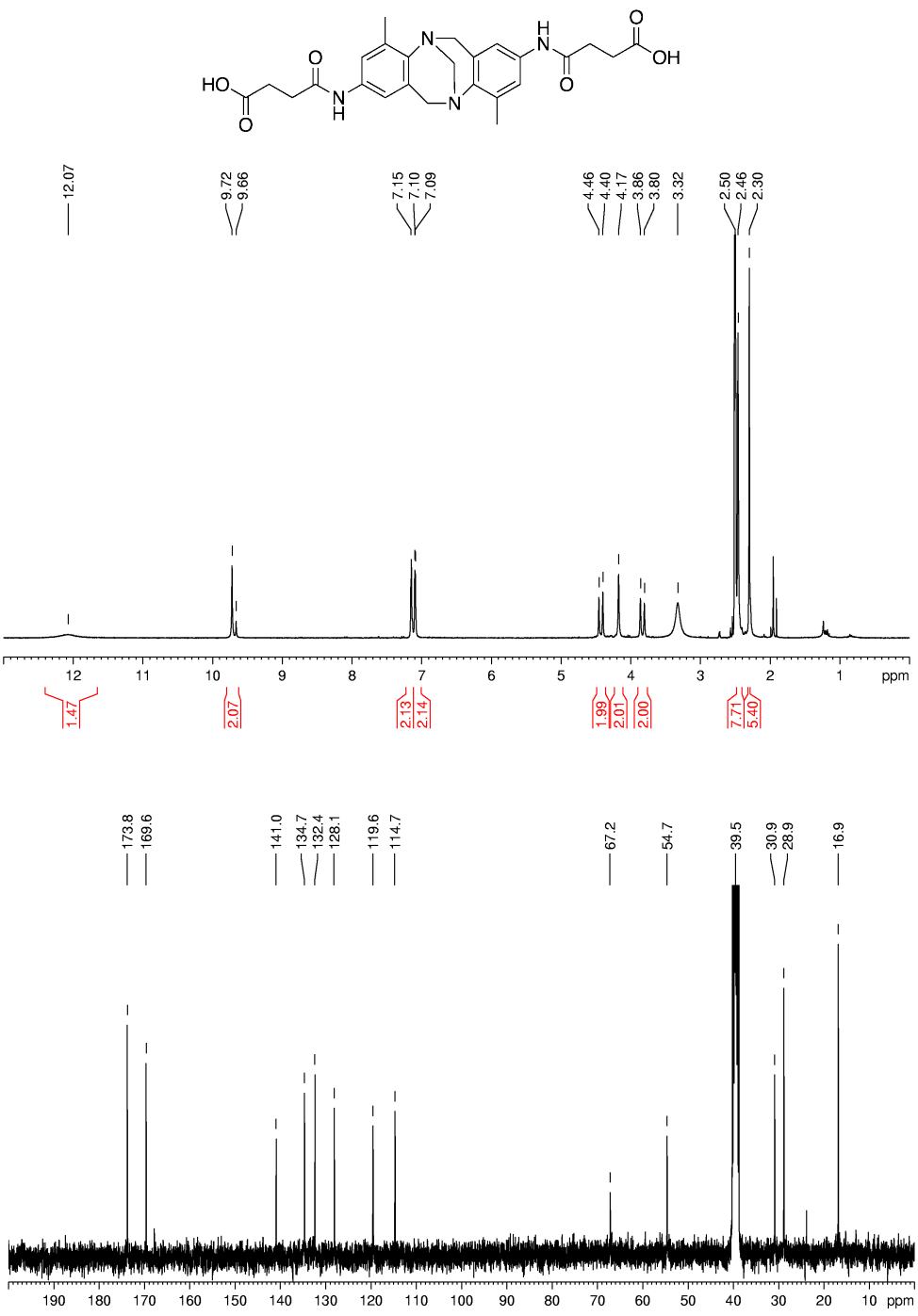
Spectra 37. Compound **3.55**



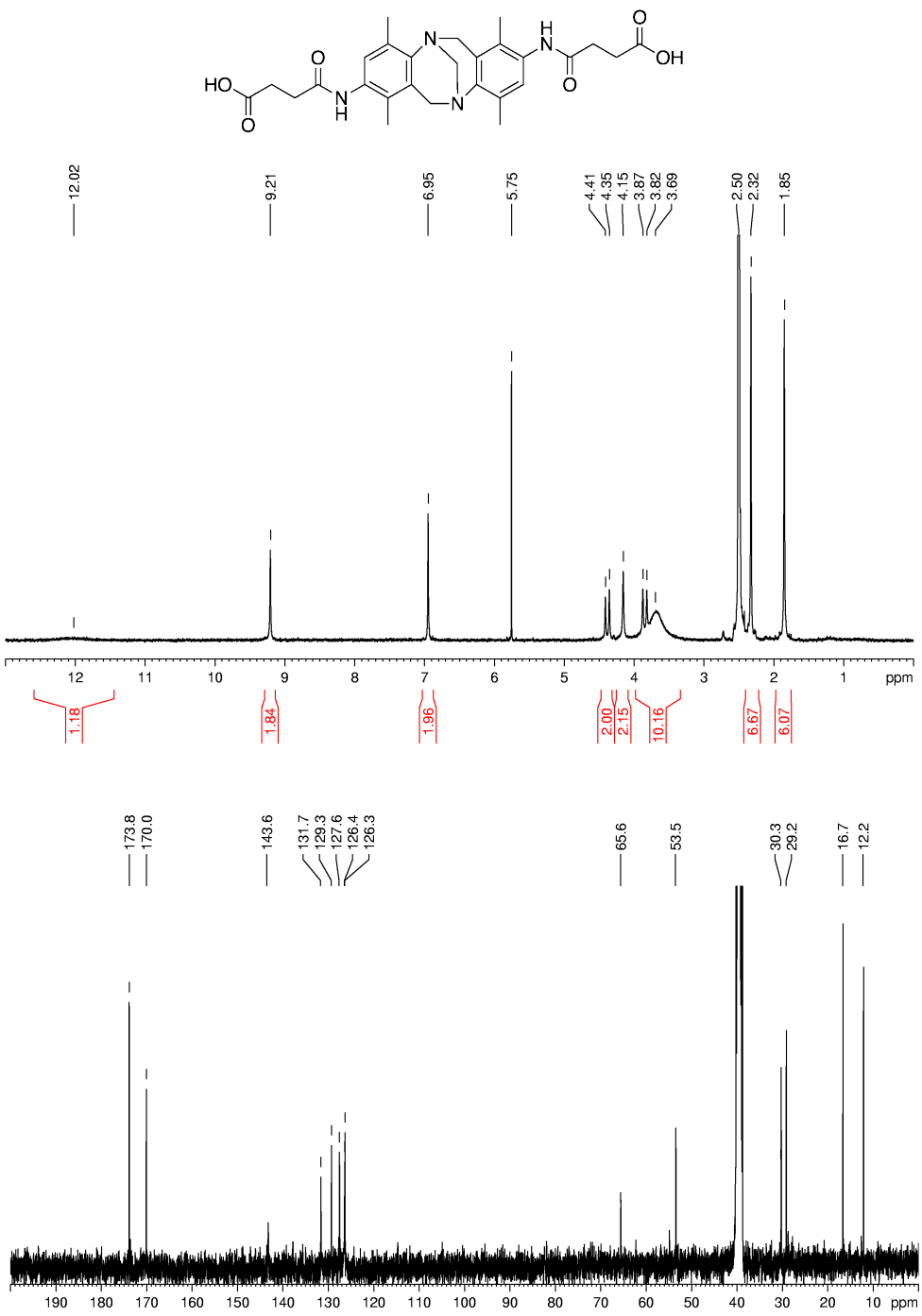
Spectra 38. Compound **3.57**



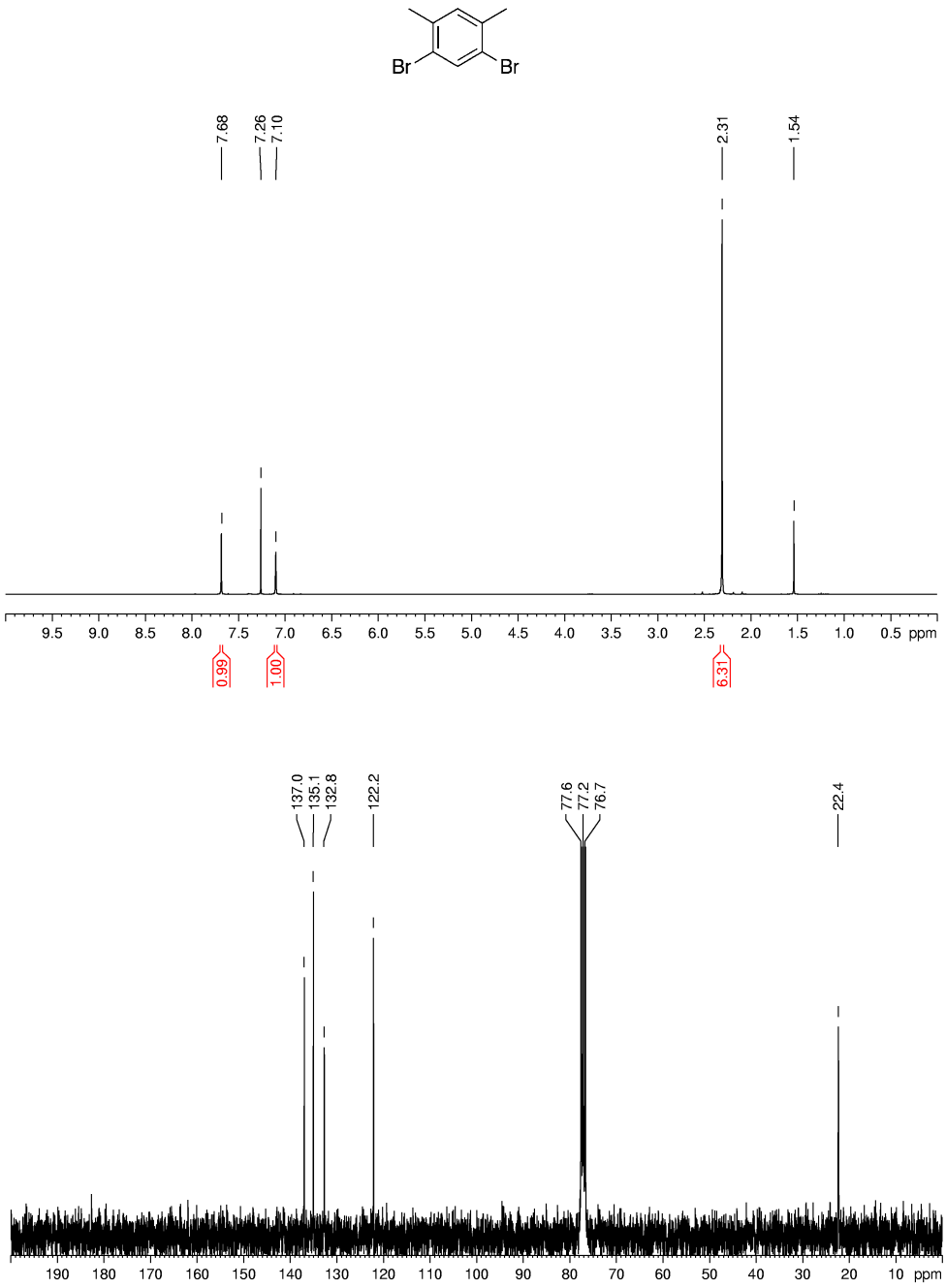
Spectra 39. Compound **3.59**



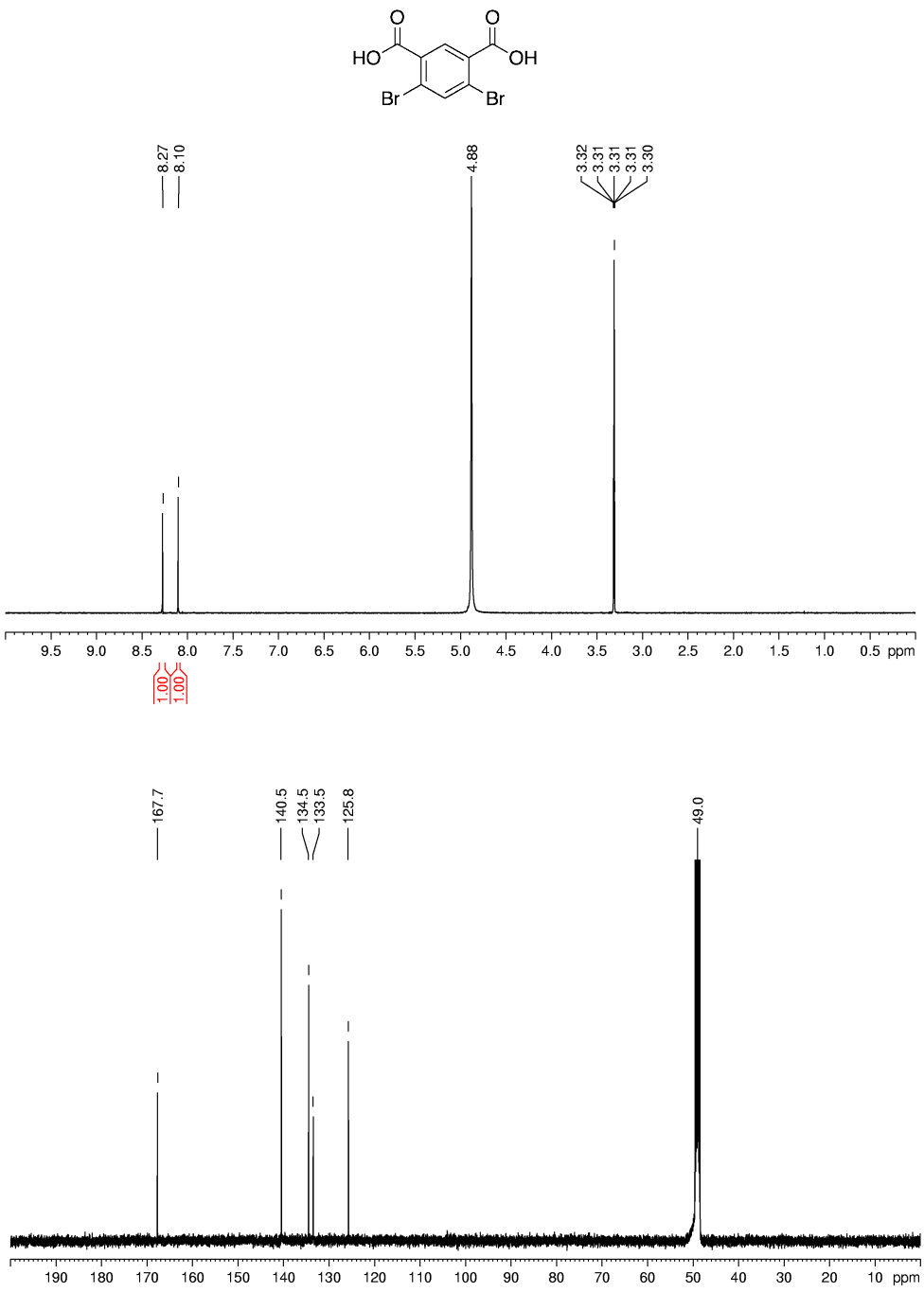
Spectra 40. Compound **4.2**



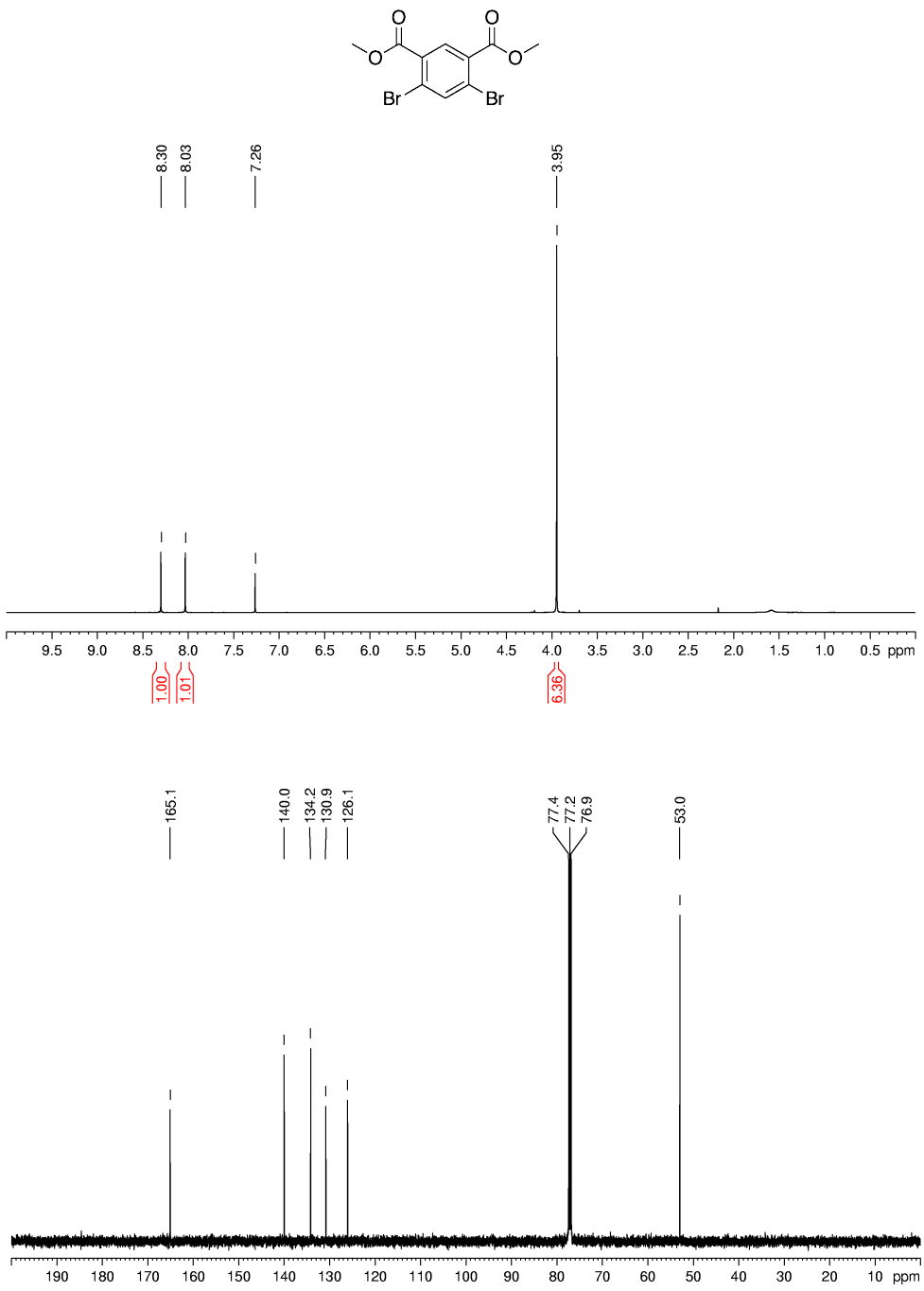
Spectra 41. Compound **4.3**



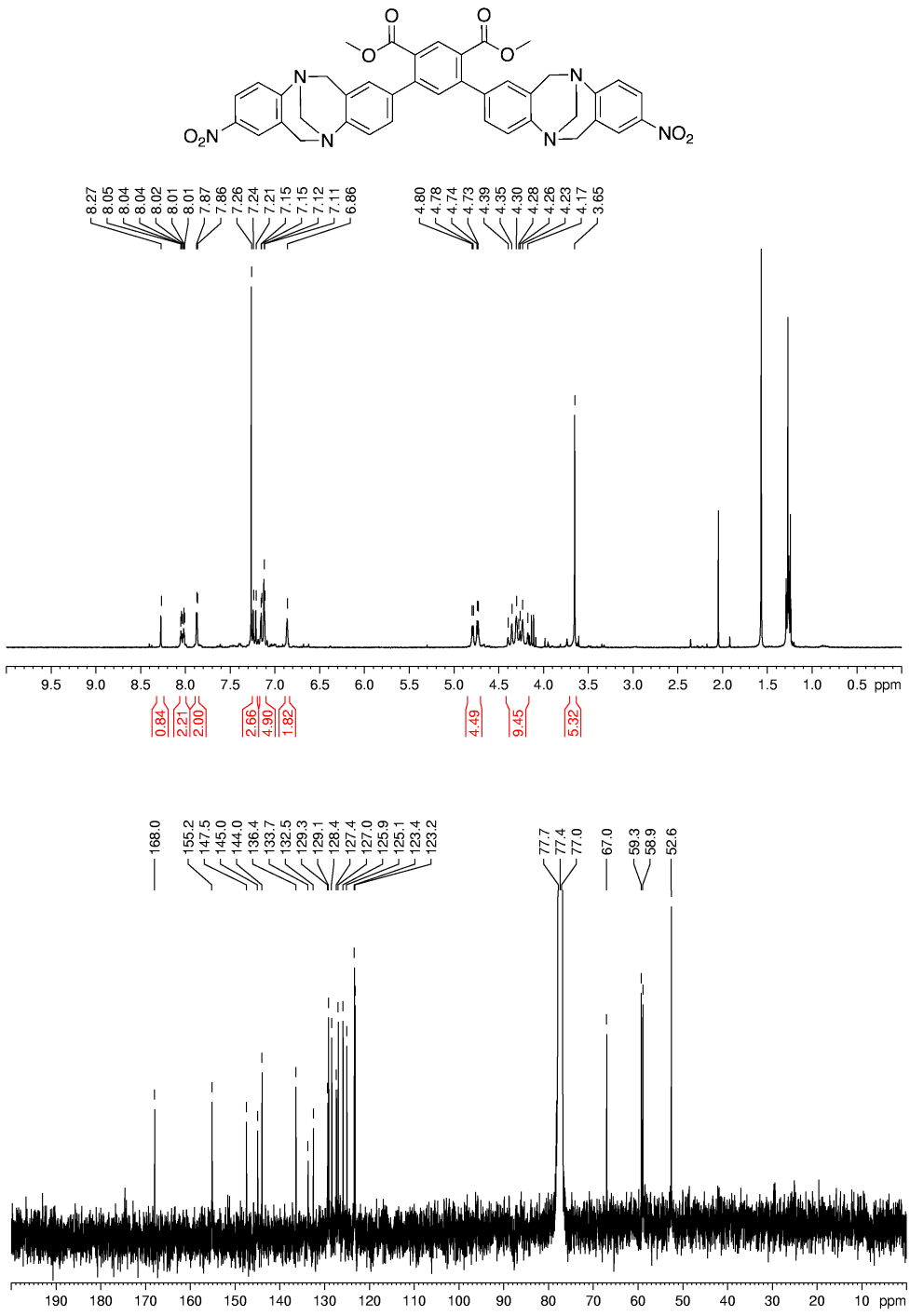
Spectra 42. Compound **4.7**



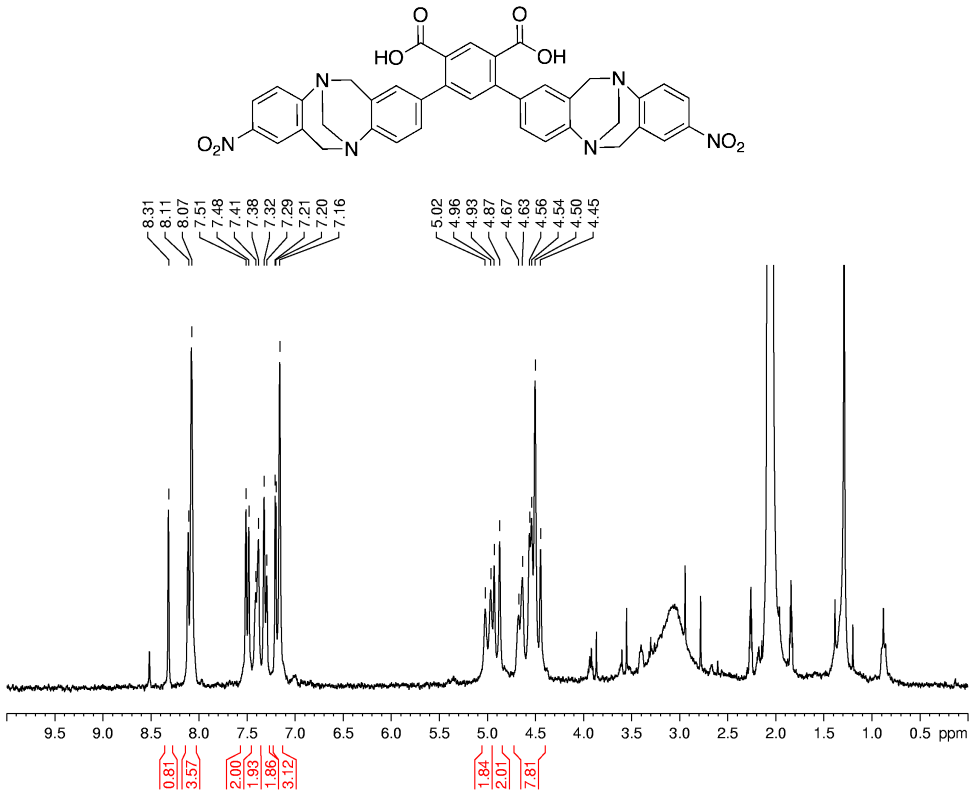
Spectra 43. Compound **4.8**



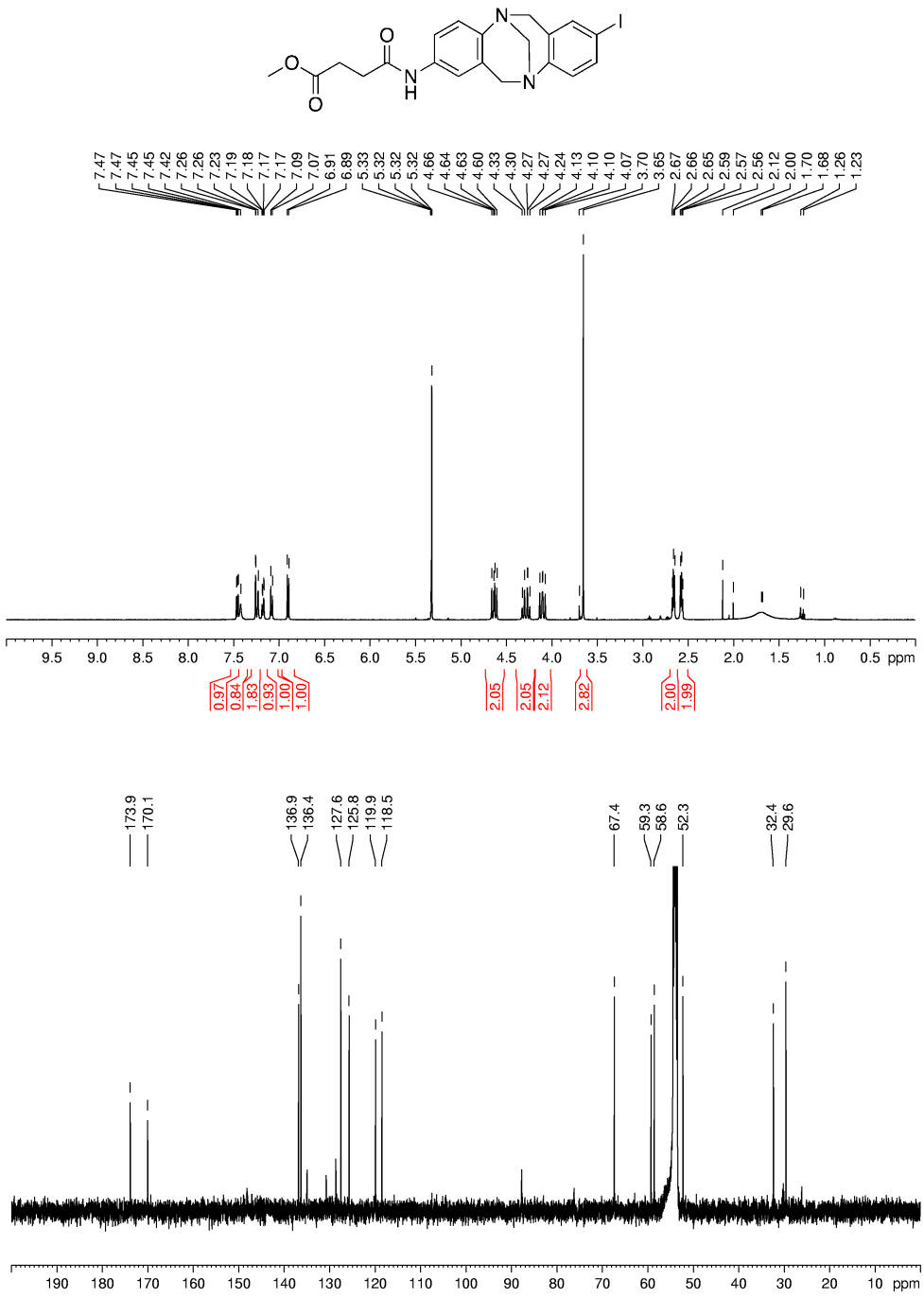
Spectra 44. Compound **4.9**



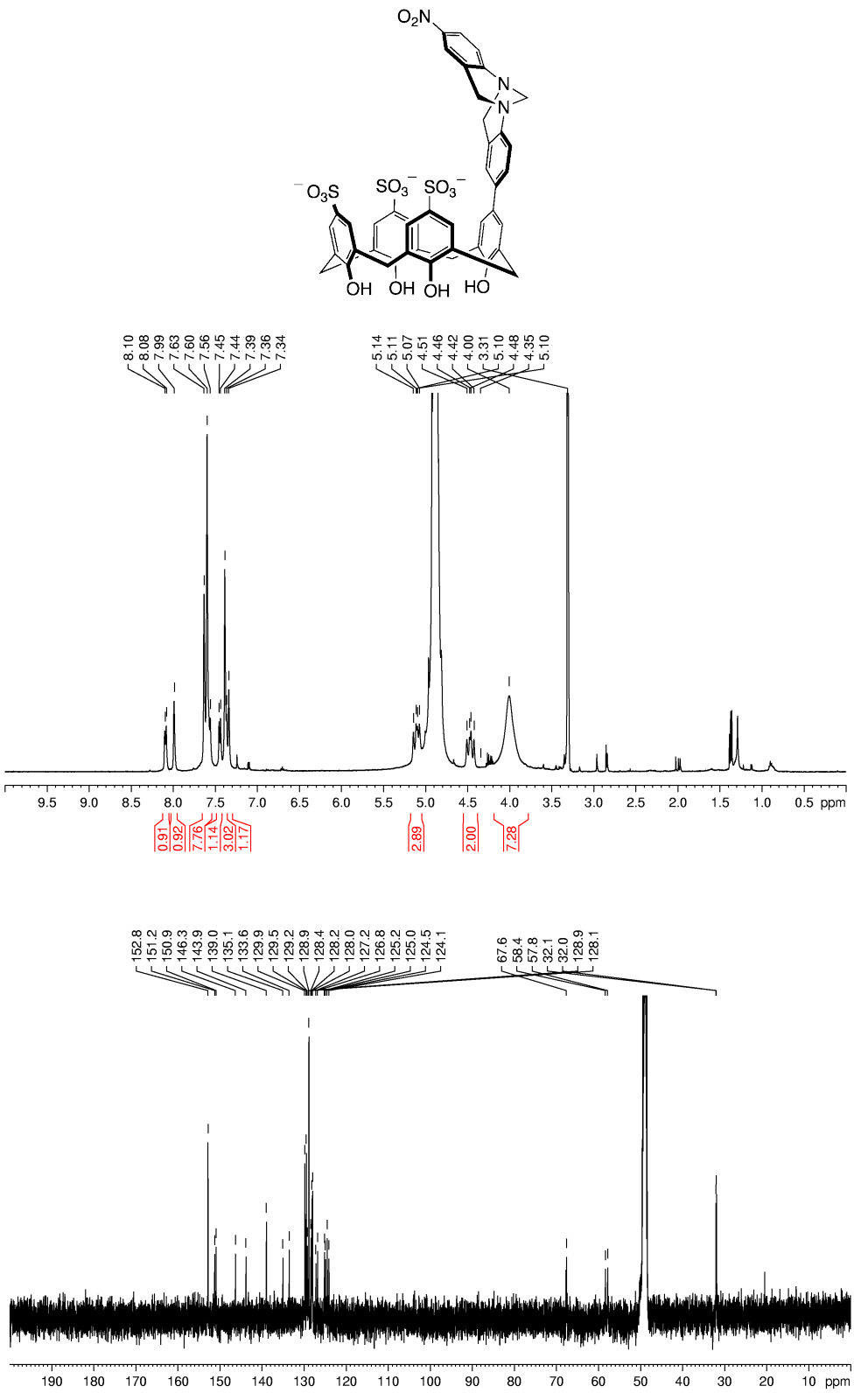
Spectra 45. Compound **4.10**



Spectra 46. Compound **4.11**



Spectra 47. Compound **4.15**



Spectra 48. Compound **4.19**