

# Guanidinium Compounds: Synthesis, Oxoanion Binding, and Cellular Delivery

by

Lindsey Orgren Calabretta

B.A. Biochemistry and Molecular Biology  
Hendrix College, 2015

M.S. Chemistry  
University of Wisconsin–Madison, 2017

SUBMITTED TO THE DEPARTMENT OF CHEMISTRY IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN CHEMISTRY  
AT THE  
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2022

©2022 Massachusetts Institute of Technology. All rights reserved.

Signature of Author: \_\_\_\_\_

Department of Chemistry  
August 5<sup>th</sup>, 2022

Certified by: \_\_\_\_\_

Ronald T. Raines  
Roger and Georges Firmenich Professor of Natural Products Chemistry  
Thesis Supervisor

Accepted by: \_\_\_\_\_

Adam Willard  
Associate Professor  
Graduate Officer

This doctoral thesis has been examined by a committee of the  
Department of Chemistry as follows:

Professor Bradley L. Pentelute \_\_\_\_\_  
Thesis Committee Chair  
Professor of Chemistry

Professor Ronald T. Raines \_\_\_\_\_  
Thesis Supervisor  
Roger and Georges Firmenich Professor of Natural Products Chemistry

Professor Laura L. Kiessling \_\_\_\_\_  
Thesis Committee Member  
Novartis Professor of Chemistry

# Guanidinium Compounds: Synthesis, Oxoanion Binding, and Cellular Delivery

by

Lindsey Orgren Calabretta

Submitted to the Department of Chemistry  
on August 5<sup>th</sup>, 2022 in Partial Fulfillment of the Requirements  
for the Degree of Doctor of Philosophy in Chemistry

## Abstract

The delivery of biological molecules into cells has been an issue of importance in both chemical biology and drug discovery. One method used to transport biologics into cells is the cell-penetrating peptide (CPP). This arginine-rich peptide forms strong interactions with the cell surface through bidentate guanidinium–oxoanion hydrogen bonds. Depending on conditions, this interaction guides the uptake of the CPP and its cargo through direct translocation or endocytosis.

In Chapter 1, I summarize literature that is relevant to this thesis.

In Chapter 2, I describe the synthesis and characterization of a small molecule, 1-guanidino-8-amino-2,7-diazacarbazole dichloride (**GADAC**), that displays a high binding affinity to a carboxylate, phosphate, and sulfate in water. **GADAC** is also fluorescent and displays an increase in quantum yield mediated by pH. The uptake and fluorescence of **GADAC** is observed in human melanoma cells via epifluorescent microscopy. Thus, the **GADAC** scaffold shows promise as a potential cell-uptake promoter and fluorescent reporter of biologics.

In Chapters 3 and 4, I explore alternative amino acids for use in CPPs. I studied the ability of canavanine, a  $\delta$ -oxa-analog of arginine, to partition into octanol in the presence of anionic lipids as a proxy for its cell-penetration ability. I observed that canavanine is worse at partitioning than arginine, indicating it may not be an effective CPP alternative.

In contrast, I synthesized and performed anion-mediated partitioning on  $N_{\alpha}$ -methylated arginine derivatives and observed increased octanol uptake compared to unmethylated arginine. This increased uptake is correlated with a decrease in topological polar surface area (TPSA) and indicates that an  $N_{\alpha}$ -methylated CPP could be a cell-uptake promoter with increased efficacy.

Lastly, in Chapter 4, I describe the synthesis of biaryl-bisguanidines. These guanidines are inspired by axially constrained organometallic catalyst ligands and have applications in oxoanion binding as dications and organometallic catalysts as dianions. I detail initial forays into determining the binding affinities of the guanidines to oxoanions through NMR titration experiments, which were hampered by changing ionic strength of the solutions.

Appendices describe the synthesis of photocaged phosphinothioesters for the traceless Staudinger ligation and attempts to install a diazo moiety site-selectively at the N-terminus of a peptide or protein.

Thesis supervisor: Ronald T. Raines

Title: Firmenich Professor of Chemistry

## Acknowledgements

I am very grateful to have had an incredible support system to help me through this journey. First, I have to thank my graduate research advisor, Prof. Ron Raines. I appreciate the independence he's given me to explore new directions in my projects, and I'm grateful he gave me the chance to develop as a scientist in a lab that does both chemistry and biology. Under his mentorship, I have grown in critical thinking and fortitude in ways I never thought possible.

I also have to thank my committee members. Prof. Laura Kiessling has followed my graduate career from the beginning and has been a great source of knowledge and advice on my projects. My committee chair, Prof. Brad Pentelute, has also been a great source of support and guidance.

I am grateful for the support of Dr. Walt Massefski, Dr. Mohan Kumar, and the other staff of the DCIF for always being available to help with experiments.

I am also grateful for the professors who helped me at Hendrix College, particularly Dr. Randall Kopper and Dr. Chris Marvin, who inspired my interest in organic chemistry and chemical biology and believed in my capabilities as a scientist.

I am thankful for all of the members of the Raines lab, both past and present. In particular, Jinyi Yang and Yana Petri have been wonderful to collaborate with, and Vienna Thomas was a lot of fun to mentor. I also have to thank Nile Abularrage, Dr. Brian Graham, Dr. Wen Chyan, Dr. Valerie Ressler, and Dr. Ian Windsor for their friendship and support.

I also have to thank my friends, who have been my support system for the past seven years. I am grateful to everyone who commiserated with me during classes, played board games, went on hikes, and generally helped keep me sane. I also appreciate everyone I have danced with for providing me with a much-needed creative outlet.

I am grateful to my family for all the support they have provided me over the years. I am incredibly grateful for my parents, who always believed I could achieve a Ph.D. even when I didn't. They encouraged me to work hard and persevere and have loved and supported me no matter what. I would not have been able to write this thesis without them.

And lastly, I have to thank my husband, Phil, who has been my rock throughout graduate school. He has been a constant calming presence and helped guide me through the toughest times — when I couldn't get reactions or titrations to work and when I didn't know if I would be able to finish. I am so grateful for his care and support. Graduate school was made so much easier by having a teammate who knew exactly what I was going through.

# Table of Contents

<b>Abstract</b>	<b>3</b>
<b>Acknowledgments</b>	<b>4</b>
<b>Table of Contents</b>	<b>5</b>
<b>List of Figures</b>	<b>10</b>
<b>List of Schemes</b>	<b>13</b>
<b>List of Tables</b>	<b>14</b>
<b>List of Abbreviations</b>	<b>15</b>
<b>Chapter 1</b>	
<b>The guanidinium–oxoanion binding interaction in supramolecular chemistry and cell penetration</b>	<b>18</b>
<b>Abstract</b>	<b>19</b>
<b>Introduction</b>	<b>20</b>
<b>Methods of delivering proteins into the cytoplasm</b>	<b>21</b>
Cell-penetrating peptides	23
<b>Quantifying binding interactions between guanidine and anions</b>	<b>25</b>
Monoguanidino oxoanion binders	29
Bisguanidino oxoanion binders	34
Trisguanidino oxoanion binders	38
<b>Abiotic guanidine-containing structures for cell penetration</b>	<b>40</b>
Peptoid transporters	40
Antibiotic transporters	42
Dendrimer transporters	43
Non-covalent transporters	45
Small molecule transporters	46
<b>Summary</b>	<b>49</b>

<b>Chapter 2</b>	
<b>A Fluorescent Guanidinium-Azacarbazole for Oxoanion Binding and Cell Penetration</b>	<b>50</b>
<b>Abstract</b>	<b>51</b>
<b>Introduction</b>	<b>52</b>
<b>Results and Discussion</b>	<b>55</b>
Synthesis of guanidino compounds	55
Evaluation of fluorescence	59
Fluorescence characterization of GAC	59
Fluorescence characterization of GADAC	61
GADAC–oxoanion binding affinities	63
Epifluorescent microscopy of GADAC in live mammalian cells	67
<b>Conclusions</b>	<b>69</b>
<b>Unsuccessful Routes</b>	<b>70</b>
Carbamate transfer during guanidinylation with thiourea	70
Pictet-Spengler route to 2,7-diazacarbazole	71
Attempted Suzuki conditions	72
Other metal-catalyzed coupling routes to 2,7-diazacarbazole 2.8	73
<b>Materials and Methods</b>	<b>75</b>
General	75
Synthesis	76
Fluorescence Spectra	89
pH Titrations	91
Extinction Coefficients	97
Quantum Yield	100
Anion Binding Titrations	101
cLog <i>P</i> and TPSA Calculation	107
Microscopy	108
NMR Spectra	111

<b>Chapter 3</b>	
<b>Canavanine versus arginine: Prospects for cell-penetrating peptides</b>	<b>128</b>
<b>Abstract</b>	<b>129</b>
<b>Introduction</b>	<b>130</b>
<b>Results and Discussion</b>	<b>131</b>
<b>Conclusions</b>	<b>135</b>
<b>Materials and Methods</b>	<b>136</b>
General	136
Octanol–water partitioning experiments	137
Synthesis	138
NMR Spectra	146
<b>Chapter 4</b>	
<b><i>N</i><sup>α</sup>-Methylation of Arginine: Implications for Cell-Penetrating Peptides</b>	<b>149</b>
<b>Abstract</b>	<b>150</b>
<b>Introduction</b>	<b>151</b>
<b>Results and Discussion</b>	<b>152</b>
<b>Conclusions</b>	<b>159</b>
<b>Materials and Methods</b>	<b>160</b>
General	160
Octanol–water partitioning experiments	161
Synthesis	162
NMR Spectra	177

<b>Chapter 5</b>	
<b>Synthesis of biaryl-bisguanidines and forays into <sup>1</sup>H NMR titration of charged species</b>	<b>186</b>
<b>Abstract</b>	<b>187</b>
<b>Introduction</b>	<b>188</b>
Guanidine-oxoanion binding in cell uptake	188
Biaryl and guanidino coordination in metal complexes	189
Unsubstituted biaryl-bisguanidines	191
<b>Results and Discussion</b>	<b>192</b>
Synthesis	192
pK <sub>a</sub> titrations via <sup>1</sup> H NMR	196
Binding affinity titrations via <sup>1</sup> H NMR in water	198
<i>Buffer optimization</i>	198
<i>Affinity titration results</i>	199
Constant ion concentration titrations via <sup>1</sup> H NMR in DMSO	205
<b>Conclusions</b>	<b>208</b>
<b>Materials and Methods</b>	<b>212</b>
General	212
pK <sub>a</sub> titrations	213
Original NMR titrations	213
Constant ion concentration NMR titrations	214
Synthesis	214
NMR Spectra	222
<b>Chapter 6</b>	
<b>Future Directions</b>	<b>235</b>
<b>Cell penetration with GADAC</b>	<b>236</b>
<b>Oxyguanidine in oxoanion binding</b>	<b>237</b>
<b>N<sup>α</sup>-methyl arginine oligomers as CPPs</b>	<b>238</b>
<b>Biaryl-2,2'-bisguanidines in oxoanion binding</b>	<b>238</b>



<b>Appendix A</b>	
<b>Efforts toward a photocleavable phosphine protecting group for the traceless Staudinger ligation</b>	<b>239</b>
<b>Abstract</b>	<b>240</b>
<b>Introduction</b>	<b>241</b>
<b>Results and Discussion</b>	<b>245</b>
<b>Future Directions</b>	<b>248</b>
<b>Epilogue</b>	<b>251</b>
<b>Synthesis</b>	<b>254</b>
<b>Appendix B</b>	
<b>Efforts Towards the Site-Selective Introduction of a Diazo Group at the N-Terminus of a Protein</b>	<b>256</b>
<b>Abstract</b>	<b>257</b>
<b>Introduction</b>	<b>258</b>
<b>Results and Discussion</b>	<b>259</b>
<b>Synthesis</b>	<b>261</b>
<b>Appendix C</b>	
<b>1-[3-(Diphenylphosphino)-propanoyl]-2,5-pyrrolidindione</b>	<b>264</b>
<b>Introduction</b>	<b>265</b>
<b>Azides to Diazo Compounds</b>	<b>265</b>
<b>Miscellaneous</b>	<b>268</b>
<b>References</b>	<b>269</b>

## List of Figures

<b>Figure 1.1.</b> The HIV-tat peptide binding to biological anions	24
<b>Figure 1.2.</b> The binding between a guanidinium ion and a sulfate or chloride ion	28
<b>Figure 1.3.</b> The association of two molecules of <b>1.3</b> to one molecule of sulfate	30
<b>Figure 1.4.</b> The binding of compound <b>1.13</b> to a phosphate	35
<b>Figure 1.5.</b> The perpendicular binding of compound <b>1.17</b> to a phosphate	36
<b>Figure 1.6.</b> Peptoid CPPs	41
<b>Figure 2.1.</b> Fluorescence emission spectra of $\beta$ -carboline, <b>GAC</b> , and <b>GADAC</b>	59
<b>Figure 2.2.</b> Fluorescence emission intensities of <b>GAC</b> with changing pH value	60
<b>Figure 2.3.</b> Fluorescence emission intensity and absorbance of <b>GADAC</b> with changing pH value	61
<b>Figure 2.4.</b> Representative fluorescence spectra of binding titrations of <b>GAC</b> and <b>GADAC</b>	65
<b>Figure 2.5.</b> Images of the uptake of <b>GADAC</b> into live M21 cells	68
<b>Figure 2.6.</b> The unexpected addition of Boc to amino-azaindole <b>2.1</b>	70
<b>Figure 2.7.</b> Absorbance/Emission spectra of <b>GAC</b> and <b>GADAC</b> in methanol	89
<b>Figure 2.8.</b> Absorbance/Emission spectra of <b>GAC</b> and <b>GADAC</b> in water	89
<b>Figure 2.9.</b> Absorbance/Emission spectra of <b>GADAC</b> in citrate and phosphate buffer	89
<b>Figure 2.10.</b> Excitation/Emission spectra of the 2-azacarbazole backbone compounds	90

<b>Figure 2.11.</b> Excitation/Emission spectra of the 2,7-azacarbazole backbone compounds	90
<b>Figure 2.12.</b> Fluorescence emission spectra of <b>GAC</b> with changing pH value, fitted data	93
<b>Figure 2.13.</b> Fluorescence emission spectra of <b>GADAC</b> with changing pH value, fitted data	94
<b>Figure 2.14.</b> Absorbance spectra of <b>GADAC</b> with changing pH value, fitted data	95
<b>Figure 2.15.</b> <sup>1</sup> H NMR spectra of <b>GAI</b> with changing pH value, fitted data	96
<b>Figure 2.16.</b> Extinction coefficient of <b>GADAC</b> in citrate and phosphate buffer	98
<b>Figure 2.17.</b> Extinction coefficient of <b>GAC</b> in water	99
<b>Figure 2.18.</b> Binding affinity titrations of <b>GP</b> , <b>GAC</b> , and <b>GADAC</b> and anions <b>2.12</b> , <b>2.13</b> , and <b>2.14</b> in isopropanol, methanol, or water.	102
<b>Figure 2.19.</b> The cLogP and TPSA for <b>GADAC</b> and <b>GCP</b>	107
<b>Figure 2.20.</b> Images of the uptake of <b>GADAC</b> into live M21 cells at different concentrations	110
<b>Figure 3.1.</b> Graph of the results of the octanol–water partitioning of <b>3.1·HCl</b> and <b>3.2·HCl</b> in the presence of anionic lipids <b>3.6</b> , <b>3.7</b> , or <b>3.8</b>	134
<b>Figure 3.2.</b> Representative <sup>1</sup> H NMR spectra from the octanol–water partitioning of <b>2·HCl</b> with lipid <b>3.7</b>	142
<b>Figure 3.3.</b> <sup>1</sup> H NMR spectra from the octanol–water partitioning of <b>3.1·HCl</b> in the presence of anionic lipids <b>3.6</b> , <b>3.7</b> , or <b>3.8</b>	143
<b>Figure 3.4.</b> <sup>1</sup> H NMR spectra from the octanol–water partitioning of <b>3.2·HCl</b> in the presence of anionic lipids <b>3.6</b> , <b>3.7</b> , or <b>3.8</b>	144
<b>Figure 3.5.</b> Calculation showing the electron density on atoms in Ac-Arg-NH <sub>2</sub> ·H <sup>+</sup> and Ac-Cav-NH <sub>2</sub> ·H <sup>+</sup>	145
<b>Figure 4.1.</b> Graph of the results of the octanol–water partitioning of <b>4.1–4.4·HCl</b> in the presence of sodium dodecanoate	155
<b>Figure 4.2.</b> A conformational model of <b>4.4</b> according to the NOESY correlations found	158

<b>Figure 4.3.</b> <sup>1</sup> H NMR spectra from the octanol–water partitioning of <b>4.1·HCl</b> in the presence of sodium dodecanoate	170
<b>Figure 4.4.</b> <sup>1</sup> H NMR spectra from the octanol–water partitioning of <b>4.2·HCl</b> in the presence of sodium dodecanoate	171
<b>Figure 4.5.</b> <sup>1</sup> H NMR spectra from the octanol–water partitioning of <b>4.3·HCl</b> in the presence of sodium dodecanoate	172
<b>Figure 4.6.</b> <sup>1</sup> H NMR spectra from the octanol–water partitioning of <b>4.4·HCl</b> in the presence of sodium dodecanoate	173
<b>Figure 4.7.</b> Selective 1D NOESY experiments of <b>4.4</b> in D <sub>2</sub> O	174
<b>Figure 5.1.</b> Biaryl-2,2'-bisguanidines	191
<b>Figure 5.2.</b> The biaryl-bisguanidines and diarginine <b>5.10</b> in order of increasing flexibility	192
<b>Figure 5.3.</b> Graphs of the pK <sub>a</sub> titration data of compounds <b>5.18</b> , <b>5.21</b> , and <b>5.6</b>	197
<b>Figure 5.4.</b> The buffering capacity of 2,6-lutidine	199
<b>Figure 5.5.</b> Binding affinity titration of bisguanidine <b>5.8</b> in D <sub>2</sub> O against methyl sulfate	200
<b>Figure 5.6.</b> The <sup>1</sup> H NMR assignments of bisguanidine <b>5.6</b>	200
<b>Figure 5.7.</b> Aberrant titration curves due to changes in ionic strength of the titration solutions	203
<b>Figure 5.8.</b> A comparison of the binding affinity curves using a traditional titrant addition method versus the constant ion concentration method.	206
<b>Figure 5.9.</b> Data resulting from a constant ion concentration titration of canavanine <b>3.2·HCl</b> against ammonium acetate	208
<b>Figure A.1.</b> Staudinger ligation mechanisms	242
<b>Figure A.2.</b> Initial diubiquitin synthetic route	243
<b>Figure A.3.</b> Proposed synthesis of diubiquitin using a photo-caged water-soluble phosphinothioester	244

<b>Figure A.4.</b> A model system for the traceless Staudinger ligation with a DMNB-photocaged phosphine	249
<b>Figure A.5.</b> The potential pathway for the degradation of the phosphinothioester	250
<b>Figure A.6.</b> Photo-caged phosphinothioester with potentially enhanced nucleophilicity	250

## List of Schemes

<b>Scheme 2.1.</b> Binding of <b>GCP</b> and the proposed binding of <b>GADAC</b> to acetate	54
<b>Scheme 2.2.</b> Synthesis of the model compound, 7-guanidino-6-azaindole ( <b>GAI</b> )	56
<b>Scheme 2.3.</b> Synthesis of 1-guanidino-2-azacarbazole ( <b>GAC</b> )	57
<b>Scheme 2.4.</b> Synthesis of 1-guanidino-8-amino-2,7-diazacarbazole ( <b>GADAC</b> )	58
<b>Scheme 2.5.</b> Compounds used in binding affinity titrations	64
<b>Scheme 2.6.</b> Synthetic route toward 2,7-diazacarbazole using the Pictet-Spengler method	71
<b>Scheme 2.7.</b> Cadogen cyclization route to <b>2.8</b>	73
<b>Scheme 2.8.</b> Ullmann coupling and Täuber synthesis route to <b>2.8</b>	74
<b>Scheme 2.9.</b> Buchwald-Hartwig and Ullmann coupling route to <b>2.8</b>	74
<b>Scheme 2.10.</b> Double Buchwald-Hartwig coupling route to <b>2.8</b>	74
<b>Scheme 3.1.</b> Structures of Ac-Arg-NH <sub>2</sub> ·HCl ( <b>3.1·HCl</b> ) and Ac-Cav-NH <sub>2</sub> ·HCl ( <b>3.2·HCl</b> )	132
<b>Scheme 3.2.</b> Synthetic route to Ac-Cav-NH <sub>2</sub> ·HCl ( <b>3.2·HCl</b> )	132
<b>Scheme 4.1.</b> Structures of <i>N</i> <sup>α</sup> -methylated arginine derivatives <b>4.1-4.4·HCl</b>	152
<b>Scheme 4.2.</b> Synthetic route to Ac-( <i>N</i> -Me)Arg-NMe <sub>2</sub> ( <b>4.4</b> ).	153

<b>Scheme 4.3.</b> Synthetic route to compounds <b>4.2·HCl</b> , <b>4.3·HCl</b> , and <b>4.4·HCl</b>	154
<b>Scheme 5.1.</b> Synthesis of the biaryl-bisguanidines <b>5.6</b> and <b>5.8</b>	193
<b>Scheme 5.2.</b> Synthesis of the biaryl-bismethyleneguanidines <b>5.7</b> and <b>5.9</b>	194
<b>Scheme 5.3.</b> Synthesis of the diarginine <b>5.10</b>	195
<b>Scheme A.1.</b> Synthesis of photocaged methoxyethoxymethylphosphinothioester <b>A.6</b> .	245
<b>Scheme A.2.</b> Synthesis of DMNB-protected acetylthiomethyl-diphenylphosphine <b>A.10</b> .	247
<b>Scheme A.3.</b> An anthracene-photocaged phosphine ( <b>A.14</b> ) reacts with azides upon exposure to UV light.	252
<b>Scheme B.1.</b> The synthesis of a water-soluble phosphinoester ( <b>B.4</b> ) for aqueous azide deimidogenation.	259
<b>Scheme B.2.</b> Synthesis of diazide <b>B.5</b> and conversion to diazo <b>B.8</b>	260

## List of Tables

<b>Table 2.1.</b> Photophysical properties and p <i>K</i> <sub>a</sub> values of <b>GADAC</b>	62
<b>Table 2.2.</b> Binding affinities (M <sup>-1</sup> ) of <b>GP</b> , <b>GAC</b> , and <b>GADAC</b> to anions <b>2.12</b> , <b>2.13</b> , and <b>2.14</b>	66
<b>Table 2.3.</b> Conditions explored for the Suzuki reaction to produce bipyridine <b>2.7</b>	72
<b>Table 4.1.</b> Values of cLog <i>P</i> and TPSA for arginine derivatives	157
<b>Table 4.2.</b> Values of cLog <i>P</i> and TPSA for <b>R<sub>9</sub></b> , <b>Me-R<sub>9</sub></b> , and <b>Narg<sub>9</sub></b>	157
<b>Table 5.1.</b> Binding affinities of compounds <b>5.6–5.10</b> and <b>5.18–5.22</b> against methyl sulfate in D <sub>2</sub> O	202
<b>Table 5.2.</b> Binding affinities compounds <b>5.18</b> , <b>5.19</b> , and <b>5.21</b> for acetate using the constant ion concentration method.	207

## List of Abbreviations

Ala	Alanine
Arg	Arginine
Boc	<i>tert</i> -Butyloxycarbonyl
Cav	Canavanine
cLog <i>P</i>	Calculated logarithm of the partition coefficient
COSY	Correlated spectroscopy
CPP	Cell-penetrating peptide
CsA	Cyclosporine A
DABCO	1,4-Diazabicyclo[2.2.2]octane
DAPI	4',6-Diamidino-2-phenylindole
DCM	Dichloromethane
DI	Deionized
DIC	Differential interference contrast
DMEM	Dulbecco's modified eagle medium
DMF	<i>N,N</i> -Dimethylformaldehyde
DMNB	Dimethoxynitrobenzyl
DMPG	Dimyristoyl phosphatidylglycerol
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DPBS	Dulbecco's phosphate-buffered saline
DPPF	1,1'-Bis(diphenylphosphino)ferrocene
EDC	3-(3-Dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
EtOAc	Ethyl acetate
EtOH	Ethanol
FBS	Fetal bovine serum
Fmoc	Fluorenylmethyloxycarbonyl
GAC	1-Guanidino-2-azacarbazole
GADAC	1-Guanidino-8-amino-2,7-diazacarbazole dichloride
GAI	7-Guanidino-6-azaindole

GCP	Guanidiniocarbonyl pyrrole
GFP	Green fluorescent protein
Gly	Glycine
GP	2-Guanidinopyridine
HATU	Hexafluorophosphate azabenzotriazole tetramethyl uronium
HIV	Human immunodeficiency virus
HMBC	Heteronuclear multiple bond correlation
HRMS	High resolution mass spectroscopy
HSPG	Heparan sulfate proteoglycan
HSQC	Heteronuclear single quantum coherence
ITC	Isothermal calorimetry
$K_a$	Equilibrium association constant
$K_d$	Equilibrium dissociation constant
LC-MS	Liquid chromatography-mass spectroscopy
LNP	Lipid nanoparticle
mAbs	Monoclonal antibodies
MeCN	Acetonitrile
MEM	Methoxyethoxymethyl
MeOH	Methanol
MES	2-( <i>N</i> -Morpholino)ethanesulfonic acid
mRNA	Messenger ribonucleic acid
$N_2(g)$	Nitrogen gas
NMR	Nuclear magnetic resonance
NOESY	Nuclear overhauser effect spectroscopy
Orn	Ornithine
PBS	Phosphate buffered saline
Pd/C	Palladium on carbon
PEG	Polyethylene glycol
PFA	Paraformaldehyde
pH	Negative of the logarithm of the hydrogen ion activity
$pK_a$	Negative of the logarithm of the acid-dissociation constant



PMB	<i>para</i> -Methoxybenzyl
PMB-NCS	<i>para</i> -Methoxybenzyl-isothiocyanate
POI	Protein of interest
PyBOP	Benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate
R	Arginine
RNA	Ribonucleic acid
SMoC	Small-molecule carriers
STR	Short tandem repeat
TEA	Triethylamine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TIPS	Triisopropyl silane
TLC	Thin layer chromatography
TNF	Tumor necrosis factor
TPSA	Topological polar surface area
Tris	2-Amino-2-(hydroxymethyl)propane-1,3-diol
UV	Ultraviolet
UV-Vis	Ultraviolet-Visible
VEGF	Vascular endothelial growth factor
$\epsilon$	Dielectric constant
$\epsilon$	Extinction coefficient
$\lambda_{\text{abs}}$	Maximum absorption wavelength
$\lambda_{\text{em}}$	Maximum emission wavelength
$\Phi$	Quantum yield

# **Chapter 1**

**The guanidinium–oxoanion binding interaction in supramolecular chemistry and cell penetration**

## Abstract

Lipinski's rule of five prescribes that only molecules with low molecular weight and a limited number of hydrogen bond donors and acceptors will permeate cells. Thus, classical drug discovery has focused on producing small-molecule medicines. Recently, biological molecules such as antibodies and mRNA have been investigated as medicines; however, these large, highly polar molecules certainly break Lipinski's rules. Thus, strategies to improve the cellular uptake of biologics are required. Cell-penetrating peptides (CPPs) are one such method.

CPPs can direct the cellular uptake of cargo to which they are attached through their unique interaction with the cellular surface. In arginine-rich CPPs, the guanidinium of arginine forms strong, bidentate hydrogen bonds with the cell surface, triggering cell uptake through direct translocation or endocytosis. To understand this binding interaction more thoroughly, I here describe many studies that have probed guanidinium-oxoanion binding using monoguanidine compounds or specially designed, preorganized bis- and trisguanidine compounds. These examples shine a light on the factors that may increase or impact this binding interaction.

Further, I present multiple strategies that have utilized the guanidinium-oxoanion binding interaction to increase the cellular uptake of protein cargo. These approaches use guanidine-containing dendrimers, lipids, carbohydrates, and small molecules. These examples highlight the utility of the guanidinium-oxoanion interaction in cell uptake and reveals that many guanidine-containing constructs have yet to be explored.

## Introduction

For over a century, pharmaceutical development has focused on developing small-molecule drugs. The majority of prescription and over-the-counter drugs are small molecules. These small molecules typically have a molecular weight lower than 500 Daltons. While they are often inspired by natural compounds – such as aspirin’s initial source in willow bark – they are typically synthesized by chemical processes. Small molecules are often preferred since they are easier to synthesize than biological molecules. In addition, small molecules can target intracellular proteins equally as well as circulating or membrane-bound proteins. The ability of small molecules to directly translocate into cells is due to their small size and relatively low hydrophilicity. In contrast, an increasing number of biologically inspired drugs are being investigated due to the greater understanding of genetics and biological mechanisms, though these moieties are typically unable to cross the cell membrane without modification. “Biologics” are pharmaceuticals that deploy biological molecules such as proteins, DNA and RNA, and even viruses as therapeutics.

The investigation of biologics has exploded in recent years. In 2013, there were 907 biologics in development.<sup>1</sup> In 2018, eight of the ten highest-grossing pharmaceuticals were biologics – specifically, monoclonal antibodies (mAbs).<sup>2</sup> These antibodies act as inhibitors or activators of crucial biological processes. Antibodies possess a greater surface area and structural diversity than do small molecules and can thus bind challenging protein targets more tightly and selectively. For example, mAbs Eylea and Avastin target the soluble protein vascular endothelial growth factor (VEGF). VEGF is a circulating protein whose overexpression results in the overgrowth of blood vessels. By inhibiting VEGF, these mAbs decrease blood vessel propagation, slowing tumor growth

in cancer and limiting excess retinal blood vessels in macular degeneration. The mAbs Humira and Remicade target the protein tumor necrosis factor (TNF). TNF is both a membrane-bound and soluble protein and, when activated, is the first protein triggered in a signal cascade that stimulates inflammation. Humira and Remicade inhibit TNF and thus reduce inflammation in autoimmune diseases such as arthritis, Chron's disease, and others. The mAbs developed to treat various cancers also target strictly membrane-bound proteins. Keytruda and Opdivo target the membrane-bound PD-1 on lymphocytes, Herceptin targets membrane-bound HER2 in breast cancer cells, and Rituxan binds to membrane-bound CD20 of cancerous B-cells.

These mAbs, while highly efficacious, target solely circulating and membrane-bound proteins. To date, no mAb has been developed to affect intracellular targets. Several other protein-based therapeutics besides mAbs have been developed; for instance, the fusion-protein Enbrel, which targets only membrane-bound TNF. New biological modalities are constantly being explored in the field of chemical biology, such as nanobodies, affibodies, DARPins, etc. While these entities are groundbreaking, they are limited in scope due to their inability to cross the cell membrane. The vast majority of metabolism occurs inside cells; therefore, a host of disease areas could be targeted if biologics could access the thousands of intracellular proteins and oligonucleotides. Thus, methods for delivering biologics through the cell membrane to targets within cells are necessary.

## **Methods of delivering proteins to the cytoplasm**

The effective delivery of proteins into cells is a challenge in both pharmaceutical development and chemical biology. The delivery of proteins is vital to pharmaceuticals,

and great leaps of understanding of basic science can be made by delivering protein-based probes and other biologics into cells *in vitro*. The methods developed for chemical biology could also prove fruitful for future drug delivery. A vast range of delivery methods have been studied, ranging from non-covalent carriers of proteins to direct chemical modifications.<sup>3</sup>

Ideally, protein-based probes and therapeutics could be delivered in their native form without extensive modifications that could render their function null. Some researchers prefer to use non-covalent methods to translocate their protein of interest (POI) into cells. These methods include simple techniques such as encasing the POI in a lipid nanoparticle (LNP) which encourages the particle to be endocytosed and protects the protein from endosomal degradation.<sup>3,4</sup> Other methods include synthetic pH-sensitive polymers and dendrimers, such as carboxymethyl chitosan-poly(amidoamine). This polymer encases a POI, encourages endocytosis of the structure, and subsequently releases the POI once the endosome matures into a low-pH lysosome.<sup>5</sup> Nanoparticles are also commonly used; for example, gold nanoparticles modified with cationic residues have successfully delivered functioning  $\beta$ -galactosidase into cells.<sup>6</sup>

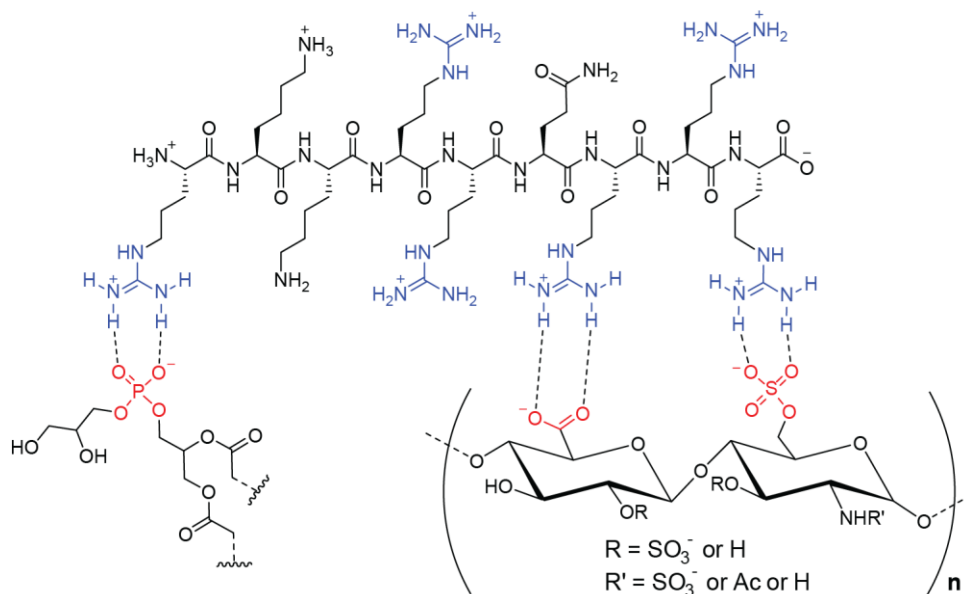
Using a similar strategy to non-covalent modifications, researchers have explored reversible covalent modifications that are cleaved once inside cells to release the native protein. Several previous studies from Raines lab take advantage of reversible modifications of proteins that are cleaved by endogenous cellular esterases. Andersen and coworkers have facilitated the translocation of GFP and RNase into cells using a benzoxaborole which binds to cell-surface glycans. The modification contains a trimethyl lock motif which self-cleaves after the benzoxaborole portion is cleaved by esterases.<sup>7</sup> Mix and coworkers have translocated GFP and RNase into cells by modifying carboxylates on

protein surfaces with neutral diazo moieties. These modifications therefore impart a greater positive charge on the protein by eliminating negative charges. Proteins with a higher positive charge are more readily internalized, allowing the modified protein to enter cells and subsequently return to its native state after esterases have cleaved the moieties.<sup>8,9</sup>

The Liu lab has developed a method to “supercharge” proteins to effect cell delivery. By mutating a high number of surface residues to lysine or arginine, they synthesized GFP(+36) and other supercharged enzymes. Due to the affinity between the highly positive protein and negative cell membrane, they found that these modified proteins could enter cells via endocytosis and perform their native enzymatic activity once inside cells.<sup>10,11</sup> Similarly, one of the most widely used transportation methods is the highly positive, arginine-rich cell-penetrating peptide.

**Cell Penetrating Peptides.** Discovered around thirty years ago, cell-penetrating peptides (CPPs) have become a mainstay in cellular delivery. One of the first CPPs was discovered from the tat protein in the human immunodeficiency virus (HIV). The isolated tat protein was found to enter cells and trans-activate the viral promoter without external influence.<sup>12</sup> Later, a highly basic portion of this protein, the HIV-tat peptide (Figure 1.1) was identified as the cause of this cell penetration.<sup>13</sup> Upon selective mutagenesis of Tat and other CPPs such as penetratin, arginine was identified as essential to its translocation ability.<sup>14,15</sup>

Subsequently, researchers have identified that a simple peptide of only arginine units is enough to effect translocation, with nona-arginine being the optimal length. This



**Figure 1.1.** The HIV-tat peptide binding to biological anions. The guanidine units of arginine form bidentate hydrogen bonds with phospholipids (left) and carboxylates and sulfate groups on heparan sulfate proteoglycan (right).

translocation is not based on charge alone, as nonamers of histidine and lysine do not show the same ability to translocate.<sup>15</sup> Arginine's unique translocation ability is due to its propensity to form strong, bidentate hydrogen bonds between its guanidine-containing side chain and anions found on the cell surface. The cell surface contains a variety of anions such as phosphates from phospholipids, carboxylates from carbohydrates and proteins, and sulfates from heparan sulfate proteoglycans (HSPGs). Researchers still debate which of these anions is most responsible for cell uptake. HSPGs make up the extracellular matrix and are the first thing a CPP-cargo complex would encounter, so it would stand to reason that they would play a role in CPP binding and uptake. Indeed, Raines and others have found that CPPs are not taken up by cell lines that do not express HSPG.<sup>16,17</sup> Some studies have found that the clustering of HSPG by cationic CPPs plays a role in cell uptake.<sup>18</sup> However, CPPs can form meaningful interactions with phospholipids, as proven by several groups studying the uptake of CPPs into phospholipid



micelle model systems.<sup>19,20</sup> In addition to the uncertainty over which anion binding interaction guides cell uptake, the transportation mechanism of CPPs is also disputed. CPP-linked cargo can be transported into cells through direct translocation or endocytosis. Frequently the uptake of cargo is guided by its size, with small cargo being directly translocated while macromolecules are endocytosed. There is also a link between CPP-cargo concentration and uptake mechanism; at very low concentrations, the CPP proceeds through direct translocation, at moderate concentrations, through endocytosis, and at high concentrations, through ceramide-mediated transduction, a process similar to direct translocation.<sup>21-23</sup>

The pathway through which CPPs perform their magic is murky. Various arginine–anion interactions guide CPPs to enter cells through various uptake mechanisms. One means to further elucidate how CPPs transport cargo into cells is to look closer at the guanidine–oxoanion binding interactions that mediate arginine’s affinity for the cell-surface.

## **Quantifying binding interactions between guanidine and anions**

Several studies have been performed to determine the binding affinity between CPPs and biological anions. Fuchs and Raines determined the binding between a fluorescently labeled R<sub>9</sub> and heparin to have a  $K_a$  of  $9.17 \times 10^6 \text{ M}^{-1}$ .<sup>16</sup> (This value, and all subsequent values originally reported as  $K_a$  or  $\log K$ , have been converted to  $K_a$  to aid in comparison between binding systems.) Ziegler and Seelig determined the binding between the HIV-tat peptide and heparin ( $(2.5 \pm 0.5) \times 10^5 \text{ M}^{-1}$ ), as well as heparan sulfate ( $(6.0 \pm 0.6) \times 10^5 \text{ M}^{-1}$ ).<sup>24</sup> This binding between these charged moieties is incredibly strong, and the high valency of the two polymers likely aids in the interaction.

Binding to biological phosphates is more challenging; phospholipids self-assemble in water, therefore, researchers must determine a CPPs binding affinity to vesicles instead of soluble monomers. For example, Ruzza *et al.* determined a surface partition coefficient for FAM-R<sub>10</sub> to a DMPG phospholipid vesicle to be  $K_p = 2.2 \times 10^8$ .<sup>25</sup> While not able to be directly compared to the binding of heparin and heparan sulfate, it is clear that this interaction plays a role in the interaction of CPPs with a cell surface.

These macromolecular binding studies attempt to mimic as closely as possible the conditions in which guanidine binds to anions on the cell surface. Conversely, many studies probe guanidine–anion binding in a non-biological context. For instance, Eric Anslyn, a leader in supramolecular chemistry, has studied several guanidine-containing preorganized molecules and their anion-binding abilities.<sup>26,27</sup> The binding affinities of these small molecules are determined in much the same way as with macromolecules – through isothermal calorimetry (ITC) and UV–Vis, fluorescence, and NMR spectroscopy. Given the ability to tune the small molecules through organic synthesis, the factors that regulate binding between guanidine and anions can be more precisely determined.

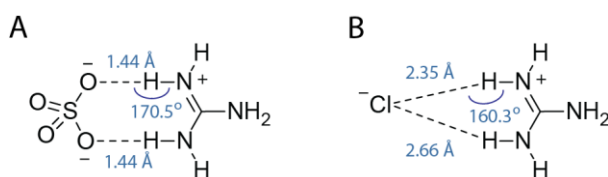
One of the main factors influencing binding affinity is the solvent in which the guanidine and anion are solvated. Clearly, the most biologically relevant solvent would be water, however, most titrations of guanidines are performed in organic solvents or organic/water mixtures. In solvents such as chloroform, the dielectric constant ( $\epsilon$ ) is very low ( $\epsilon = 4.81$ ), and the guanidine and anion have virtually no competing interactions with the solvent to impede their interaction. Conversely, in high dielectric constants such as DMSO ( $\epsilon = 46.7$ ), the two highly polar substrates are more solvated, leading to more interference in their association. In addition to having an extremely high dielectric

constant, water ( $\epsilon = 80.1$ ) also has two hydrogen bond donors and two hydrogen bond acceptors. In situations where hydrogen bonding guides binding affinity, as with guanidines and oxoanions, the presence of protic solvents such as water and methanol significantly interferes with binding. For instance, Berger and Smidtchen obtained binding affinities of their bicyclic guanidine with acetate in acetonitrile and DMSO but did not observe a heat of binding in methanol through ITC.<sup>28</sup>

In reviews regarding the binding between arginine and cell-surface anions, arginine is often touted as a strong binder due to its high  $pK_a$ . This argument states that because arginine ( $pK_a$  13.8)<sup>29</sup> is fully protonated at physiological pH, the positively charged guanidine can form strong interactions with negative anions. On a surface level, this argument is supported by the fact that peptides based on lysine ( $pK_a$  10.5) are less able to enter cells.<sup>15</sup> While arginine is indeed highly basic, its high  $pK_a$  is not the sole driver of the strength of guanidinium-oxoanion interactions. In fact, guanidine-containing compounds with lower  $pK_a$ 's display enhanced binding affinity. Carston Schmuck detailed this by comparing inorganic guanidine with acetylguanidine ( $pK_a$  7.6). When titrated into a solution of Ac-Ala-O<sup>-</sup> in 40% water/DMSO, inorganic guanidine induced no change in chemical shift of the carboxylate while acetylguanidine produced a  $K_a$  of  $50 \text{ M}^{-1}$ .<sup>30</sup> By installing an electron-withdrawing acyl group in conjugation with the guanidine, the terminal hydrogens receive an even greater partial positive charge and thus are better hydrogen bond donors. This strategy is used by many studying tightly binding guanidine-containing structures, as discussed later.

The unique structure of guanidine is the cause of its ability to bind oxoanions. The carbon and three nitrogens of guanidine take on a “Y” shape due to the  $sp^2$  character shared between all four atoms. This induces a nearly planar structure, with the terminal

hydrogens directed outward, parallel to each other. This positions them in an optimal location to form nearly 180° hydrogen bonds with the lone pairs on two oxygens of an oxoanion such as sulfate (Figure 1.2 A).<sup>31</sup> The guanidine NH's are effectively preorganized to bind to oxoanions, resulting in bidentate hydrogen bonding. While guanidine is still capable of hydrogen bonding with a monoatomic anion such as chloride, the interaction is weaker, as evidenced by the longer H-bond distances and worse bond angle (Figure 1.2 B). The preference of guanidinium for oxoanions over other anions is exemplified in work by Schmuck where their tri-guanidinium “molecular flytrap” was able to bind citrate with a  $K_a = 8.6 \times 10^4 \text{ M}^{-1}$  in water with a 1000-fold excess of chloride ions.<sup>32</sup> Additionally, the standard protein denaturing agent, guanidinium chloride, is ineffective upon the introduction of sulfate anions, indicating the preference of guanidinium for sulfate.<sup>33</sup> While the preference for halides is weaker than oxoanions, their affinity as common counterions to guanidinium-containing compounds is not to be discounted as it may still impact the resulting  $K_a$  from binding affinity studies.<sup>34</sup>

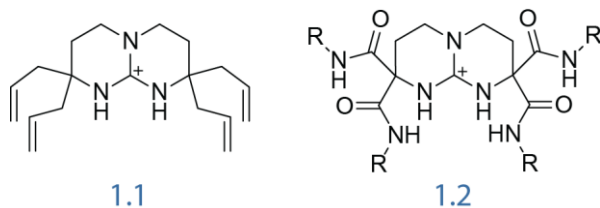


**Figure 1.2.** The binding between a guanidinium ion and a sulfate (a) or chloride (b) ion, as observed through a crystal structure and further calculated through NBO.<sup>31</sup> Hydrogen bond angles and lengths with chloride are less optimal than with sulfate.

Guanidine’s unique structure and propensity to bind to oxoanions have made it a moiety of interest to many research groups. As mentioned earlier, guanidine-containing biological polymers such as CPPs have been studied for their ability to bind to anion-containing polymers. However, the binding of oxoanions to synthetic guanidine-

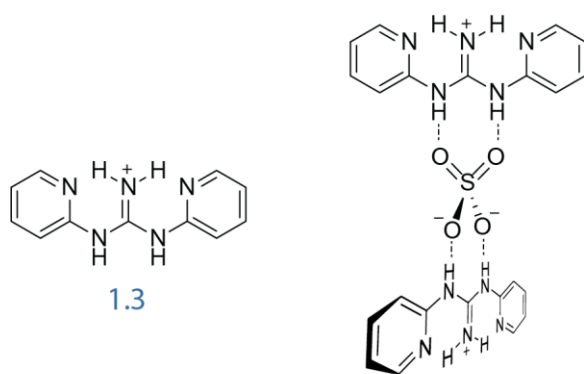
containing small molecules has been popular in supramolecular chemistry. This topic has been reviewed several times before.<sup>26,35-37</sup> Thus, the following sections represent a sampling of guanidine-containing structures and their binding to oxoanions that are most relevant for the work discussed in future chapters.

**Monoguanidino oxoanion binders.** Schmidtchen and coworkers were some of the first to study guanidinium-oxoanion binding using their 6,6-membered bicyclic guanidine.<sup>38</sup> They observed a bicyclic guanidine with *sec*-vinyl groups (**1.1**) to have binding affinities of  $10^4$  M<sup>-1</sup> for both the carboxylate *p*-nitrobenzoate and inorganic phosphate in acetonitrile. In contrast, they observed bicyclic guanidines with *sec*-amide groups (**1.2**) to have increased binding in the range of  $10^6$  M<sup>-1</sup> to *p*-nitrobenzoate and phosphate. They hypothesized that the adjacent amide groups increased the binding affinity of **1.2** due to the higher number of hydrogen bond donors. However, the authors determined by ITC that the greater binding energy of **1.2** was due to entropy instead of enthalpy due to hydrogen bonds. The authors also titrated an iteration of bicyclic guanidine (**1.2**) in methanol against various anions. In general, the binding of **1.2** to phosphate was weaker than of **1.2** to a variety of carboxylates. Inorganic sulfate did not produce enough of a response to be measured.



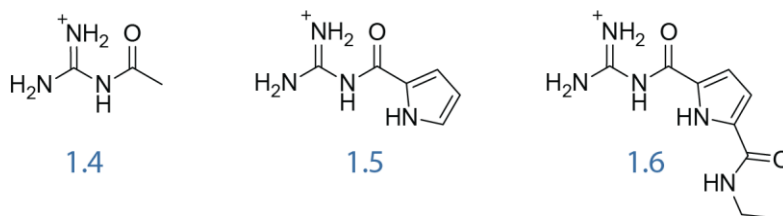
Like the bicyclic guanidines studied by the Schmidtchen lab, Seipp and coworkers have investigated a pseudo-bicyclic guanidine (**1.3**) that is conformationally restricted by

intramolecular hydrogen bonds.<sup>39,40</sup> The conformational restriction, along with the electron-withdrawing effect of two pyridine groups on the guanidine, allows this bis(2-pyridyl)guanidinium to demonstrate strong binding to anions in 10% water/methanol as measured by NMR spectroscopy. While the binding affinities to chloride ( $K_a < 0.3 \text{ M}^{-1}$ ) and nitrate ( $K_a \approx 1.3 \text{ M}^{-1}$ ) are weak, the compound shows significant binding to sulfate, even forming a 2:1 complex of guanidine to the tetrahedral sulfate dianion (Figure 1.3) ( $K_{a 1:1} \approx 6 \times 10^3 \text{ M}^{-1}$  and  $K_{a 2:1} \approx 125 \text{ M}^{-1}$ ).

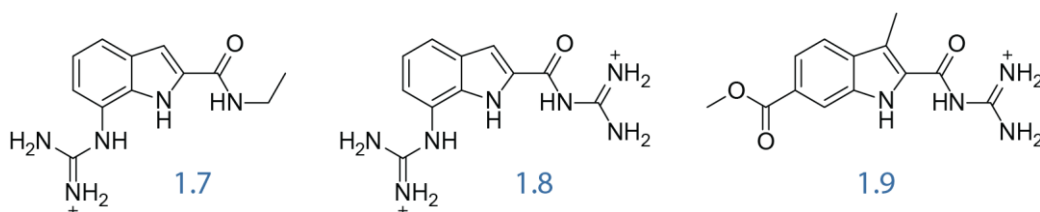


**Figure 1.3.** Compound **1.3** (left). The association of two molecules of **1.3** to one molecule of sulfate (right).

The Schmuck lab has explored enhancing the binding of structures with a single guanidine moiety by introducing additional hydrogen bond donors and tweaking the guanidinium  $pK_a$ . As mentioned earlier, the Schmuck lab has explored the increased binding affinity of a low- $pK_a$  acylguanidine as compared to an alkylguanidine. They have also explored increasing the number of hydrogen bond donors using amides and pyrroles, measuring their compounds against the  $\text{Ac-Ala-COO}^-$  in 40% water/DMSO by  $^1\text{H-NMR}$ .<sup>30</sup> While the addition of a pyrrole moiety (**1.5**) increased the  $K_a$  compared to acetylguanidine **1.4**, this increase was moderate ( $130 \text{ M}^{-1}$  vs  $50 \text{ M}^{-1}$ ). Instead, the further addition of various amide moieties in compounds such as **1.6** improved the  $K_a$  significantly ( $770 \text{ M}^{-1}$ ).



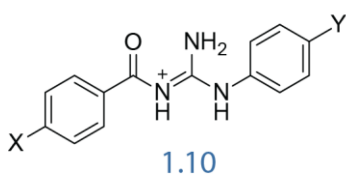
The Schmuck lab has also compared arylguanidine to acylguanidine. NMR binding studies of indole-based structures against Ac-Ala-COO<sup>-</sup> in DMSO demonstrated that acylguanidine **1.9** maintains a higher binding affinity ( $10^5$  M<sup>-1</sup>) than does arylguanidine **1.7** ( $10^3$  M<sup>-1</sup>). These results demonstrate the effect of p*K*<sub>a</sub> on binding affinity, considering phenylguanidine has a p*K*<sub>a</sub> of ~11 and acetylguanidine has a p*K*<sub>a</sub> of ~8. When both aryl- and acylguanidine moieties were incorporated into the same indole structure (**1.8**), the 1:1 and 1:2 binding affinities recapitulated the individual binding affinities.<sup>41</sup>



While installing electron-withdrawing groups on guanidines to lower their basicity will increase the binding affinity to oxoanions, at some point, the scales tip in favor of proton transfer over hydrogen bonding. Glasovac and coworkers have demonstrated that dramatically decreasing the p*K*<sub>a</sub> of a guanidine eventually leads to deprotonation by the oxoanion binding partner.<sup>42</sup> While attempting to devise a system to sense anions, the authors have developed a benzoylguanidine core (**1.10**) with benzoyl and aryl withdrawing groups, with or without additional *para* electron-donating or withdrawing groups (**1.10.1–1.10.6**). They have determined the p*K*<sub>a</sub>'s of their benzoylguanidines to be between 4.5 and 6.1 in 50% water/acetonitrile, which is comparable to those of pyridinium cations. They observed through UV–Vis titration in acetonitrile that acetate

( $pK_a = 4.8$ ), dihydrogen phosphate ( $pK_a = 2.1$ ), and fluoride ( $pK_a = 3.2$ ) showed some degree of deprotonation of all of the benzoylguanidines they studied.

Determining if changes in signal readout are due to deprotonation or binding complexation can be tricky. Deprotonation of the cation host by the anion can be teased apart from the binding affinity when the  $\Delta pK_a$  is close but not below  $\sim 1$ . Within one  $pK_a$  unit, the host and guest form a strong ion pair/salt bridge where hydrogen bonding and proton exchange are essentially the same. Above  $\Delta pK_a = 1$ , the deprotonation constant can be determined to obtain the  $pK_a$  of the host. Still, the binding affinity is difficult to separate from the deprotonation event. Perez-Casas and Yatsimirsky have determined a more complex fitting model that accounts for deprotonation and binding affinity. They acknowledge, however, that a typical 1:1 binding model gives an “apparent binding constant” that is a close-enough estimate to the true binding constant.<sup>43</sup> Glasovac and coworkers report that the “apparent” binding constants of their benzoylguanidines to acetate in 50% water/acetonitrile are  $10^4$ – $10^6$   $M^{-1}$ . Their data confirms the influence of guanidinium  $pK_a$  on binding – the benzoylguanidines with higher  $pK_a$ 's have lower “apparent” binding constants and vice versa – though the benzoylguanidine with the lowest  $pK_a$  (**1.10.5**) is merely reported as “mostly deprotonated.”



**1.10.1** X = H, Y = H

**1.10.2** X = H, Y = OCH<sub>3</sub>

**1.10.3** X = H, Y = COOCH<sub>3</sub>

**1.10.4** X = NO<sub>2</sub>, Y = H

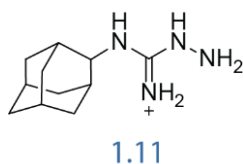
**1.10.5** X = H, Y = NO<sub>2</sub>

**1.10.6** X = NO<sub>2</sub>, Y = N(CH<sub>3</sub>)<sub>2</sub>

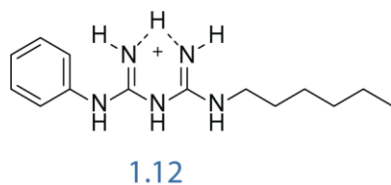
Related functional groups such as aminoguanidines and biguanidines have also been studied for their ability to interact with oxoanions. These functionalities have additional hydrogen bond donors that aid in the binding to oxoanions. Šekutor and



Mlinarić-Majerski have studied adamantyl aminoguanidines by NMR spectroscopy in DMSO.<sup>44</sup> They observed that the aminoguanidines participated in 1:2 binding with oxoanions. While inorganic phosphate's 1:1 binding was weakest, all three anions tested (inorganic phosphate, inorganic sulfate, and acetate) demonstrated the same 1:2 binding affinity to aminoguanidine **1.11** of approximately  $10^4 \text{ M}^{-1}$ . It has been previously demonstrated that the lone pair of the amino group does not participate in conjugation with the four atoms of the guanidine.<sup>45,46</sup> Therefore, the authors theorize that the amino group can also be protonated, forming a dication, which then allows the aminoguanidine to bind to two anions.



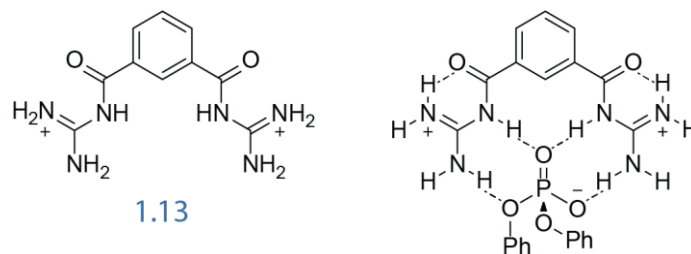
Pushina and Anzenbacher have studied the biguanidine **1.12** and its binding to various anions in 30% DMSO/chloroform by UV-Vis spectroscopy.<sup>47</sup> The authors observed via Job plots that the stoichiometry of the binding events to halides (fluoride) and carboxylates (acetate) was 1:2, yet they were unable to individually observe the 1:1 and 1:2 binding events. Thus, the authors report their binding constants as  $\text{M}^{-2}$  to reflect this stoichiometry. They found that the response to oxoanions was much greater than to halides, with fluoride having a  $K_a = 1.9 \times 10^4 \text{ M}^{-2}$  and chloride showing no response. In contrast, acetate displays a  $K_a = 1.2 \times 10^5 \text{ M}^{-2}$ . Inorganic phosphate also showed an appreciable  $K_a = 8.8 \times 10^4 \text{ M}^{-2}$ . The authors also looked at four dicarboxylates with mixed results; while oxalate showed the highest  $K_a$  of  $1.9 \times 10^5 \text{ M}^{-2}$ , phthalate showed the lowest  $K_a$  of  $8.6 \times 10^3 \text{ M}^{-2}$ , indicating no preference of a biguanidine to a dianion versus a monoanion.



**Bisguanidino oxoanion binders.** A common feature of compounds designed for molecular recognition is the incorporation of two or more binding functionalities in a preorganized and rigid structure to optimize binding to a guest molecule. Guanidine has been used as one such functionality in the molecular recognition of anions. Bisguanidines are advantageous because they can bind to tetrahedral oxoanions with a greater number of hydrogen bonds than monoguanidines.

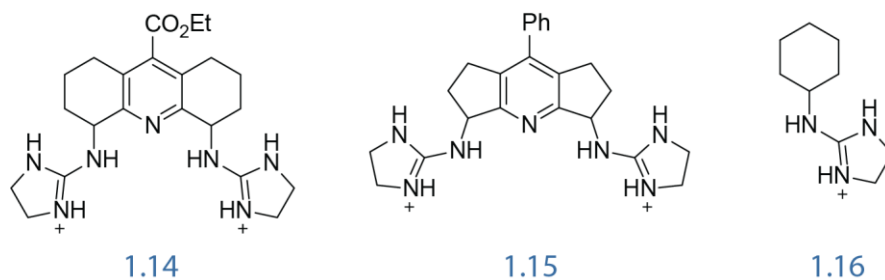
Hamilton and coworkers proposed one of the first bisguanidine molecules designed to bind oxoanions.<sup>48</sup> The *meta*-bis-acylguanidine **1.13** is elegant in its simplicity and demonstrated the binding of both guanidine units to diphenyl phosphate by NMR spectroscopy. Through dilution experiments of the 1:1 complex in acetonitrile, the authors determined the binding affinity to this phosphate to be  $4.6 \times 10^4 \text{ M}^{-1}$ , whereas a version with only one acylguanidine had a binding affinity of  $2.7 \times 10^3 \text{ M}^{-1}$ . Hutchings and coworkers also examined this structure through X-ray crystallography with inorganic sulfate.<sup>49</sup> They observed that the bis-acylguanidine is completely planar in part due to intramolecular hydrogen bonds between the guanidines and acyl oxygens. They also observed that each sulfate oxygen forms a hydrogen bond to a different guanidine hydrogen (Figure 1.4).

At the same time, Anslyn and coworkers were developing an alkyl bisguanidine.<sup>50</sup> They synthesized four different molecules with varying conformations (**1.14**, *meso* and *d,l*, and **1.15**, *meso* and *d,l*). With the guanidines placed on rings around a pyridine core,



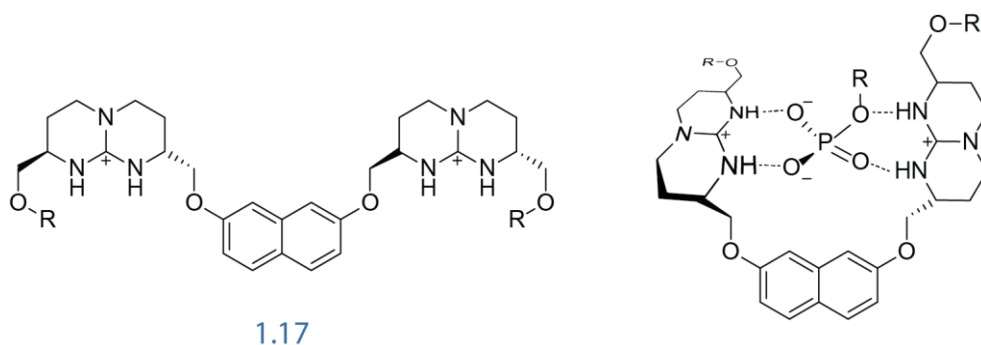
**Figure 1.4.** Compound **1.13** (left). The hydrogen bonding of both guanidides to one phosphate, and the intramolecular hydrogen bonds between guanidine and acyl carbonyls that preorganize the molecule (right).

the binding pocket size can be altered by respective ring size; a 6-membered ring (**1.14**) results in a smaller angle between the guanidines, whereas a 5-membered ring (**1.15**) results in an increased angle. The chirality of the guanidine centers – whether the guanidines are on the same face of the core (*meso*) or opposite sides (*d,l*) – also determines the shape of the binding pocket. The authors determined that all four structures bound more tightly to dibenzyl phosphate in DMSO by  $^{31}\text{P}$ -NMR spectroscopy than did the mono-guanidine (**1.16**) ( $350 \text{ M}^{-1}$ ).



Both *meso* forms of **1.14** and **1.15** show similar binding affinities to dibenzyl phosphate with a 1:1 binding in the  $10^3 \text{ M}^{-1}$  range and a 1:2 binding of  $10^1 \text{ M}^{-1}$ . The *d,l* form of **1.14** shows a lesser 1:1 binding affinity of  $10^2 \text{ M}^{-1}$  and a 1:2 binding of  $10^1 \text{ M}^{-1}$ , indicating that while this conformation is less optimal for binding, the flexibility of the six-membered rings still allows for cooperative binding of both guanidines to the phosphate in 1:1 binding. In contrast, the *d,l* form of **1.15** shows an equal binding of  $10^2 \text{ M}^{-1}$  for both

the 1:1 and 1:2 affinities, indicating that the guanidines are binding to dibenzyl phosphate entirely independent of each other. The authors also explored the effect of the solvent system on the binding interaction; upon increasing the percentage of water in the DMSO solution, the binding affinity dropped dramatically. At 5% water, the  $K_a$  of *meso*-**1.15** is  $\sim 10^3 \text{ M}^{-1}$ , between 10-25% the  $K_a$  decreases to  $10^2 \text{ M}^{-1}$ , with the  $K_a$  finally becoming immeasurable above 50% water. These studies show water's strong effect of disrupting hydrogen bond interactions and that fine-tuning the orientation of two guanidine moieties in a preorganized host is critical.



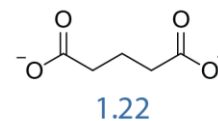
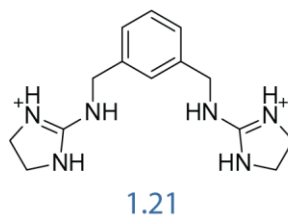
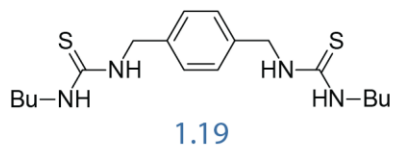
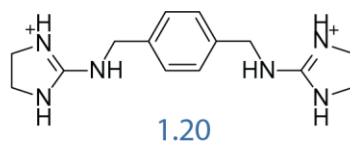
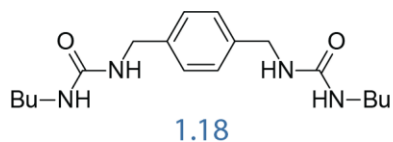
**Figure 1.5.** Compound **1.17** (left). The perpendicular binding of two bicyclic guanidinium moieties to one phosphate anion (right).

The Schmidtchen lab has incorporated the 6,6-bicyclic guanidine discussed earlier into a tweezer-like bisguanidine (**1.17**) capable of binding to phosphate in methanol and water.<sup>51</sup> In contrast to the rigid, preorganized structures described above, the two bicyclic guanidines are attached to a naphthalene core and can freely rotate open into an extended conformation. The structure is still well designed, however, as the chiralities of the bicyclic guanidines allow them to be perpendicular to one another in the bound state, allowing binding of both guanidines to all four phosphate oxygens (Figure 1.5). While the entropy cost of transitioning between the two conformers must be significant, the structure still manages to bind to a range of phosphate-containing anions with high

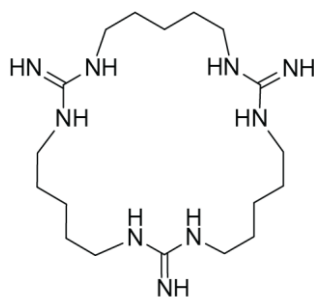
affinity when measured by  $^1\text{H-NMR}$  spectroscopy. The structure binds to inorganic phosphate with a  $K_a = 1.8 \times 10^4$  in methanol and a  $K_a = 9.7 \times 10^2$  in water.

As of yet, I have only discussed bisguanidines binding to tetrahedral oxoanions such as phosphate and sulfate. Bisguanidines have also been synthesized to bind to carboxylates; however, as carboxylates are not tetrahedral, they do not receive much extra benefit from two guanidines attempting to bind to a single carboxylate. Therefore, most bisguanidines designed to bind carboxylates target dicarboxylates. A later work of Hamilton and coworkers is one such example.<sup>52</sup>

These authors studied not only a bisguanidine but also compared their results to a bisurea and bithiourea. Neutral hydrogen bond donors such as ureas have been popular in supramolecular chemistry and work well in non-polar organic solvents. As this work shows, however, guanidines are far superior in polar protonated solvents. Bisurea **1.18** ( $K_a = 640 \text{ M}^{-1}$ ) and bithiourea **1.19** ( $K_a = 1 \times 10^4 \text{ M}^{-1}$ ) perform well against glutarate **1.22** in DMSO, but bisguanidines **1.20** and **1.21** outpace them by being too high to measure accurately by NMR titration ( $K_a > 5 \times 10^4 \text{ M}^{-1}$ ). In water/DMSO solutions, the bisguanidines **1.20** and **1.21** display a  $K_a$  of  $\sim 3 \times 10^3 \text{ M}^{-1}$  in 10% water/DMSO. The binding affinity can still be measured at  $10^2 \text{ M}^{-1}$  in 50% water/DMSO. By ITC, the authors observed that while the binding of bisurea **1.18** and bithiourea **1.19** in DMSO is enthalpically driven, the binding of bisguanidines **1.20** and **1.21** in methanol and methanol/water solutions is entropically driven. The ITC results indicate that in organic solvents, the enthalpy of hydrogen bonding drives the association. In protic solvents, however, the strength of the association is due to the increase in entropy from the release of the solvent molecules from the host and guest upon binding. This phenomenon is commonly observed in supramolecular chemistry.



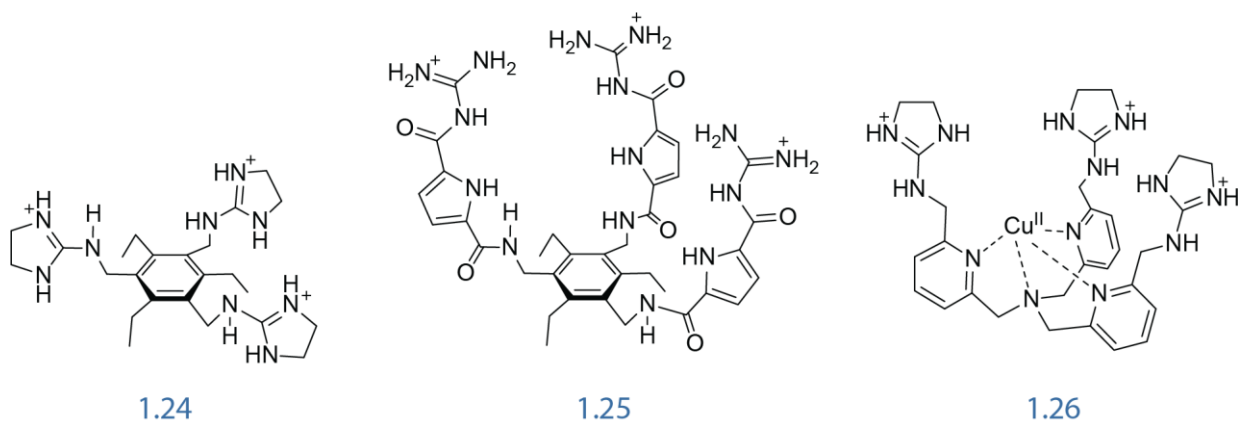
**Trisguanidino oxoanion binders.** One of the earliest guanidine-containing anion binders was proposed by Lehn and coworkers in the late 1970s.<sup>53</sup> Inspired by crown ethers that bind tightly to cations, the authors sought to achieve the reverse – a macrocycle containing multiple guanidines that could selectively bind anions. They successfully synthesized macrocycle **1.23** and observed its binding affinity by comparing pH titrations of inorganic phosphate with and without macrocycle present. They determined that macrocycle **1.23** had a stability constant (a parameter similar to an equilibrium constant) of  $K_s = 251$  in water and  $K_s = 1.9 \times 10^4$  in 10% water/methanol. The non-cyclic analog of **1.23** had a  $K_s$  slightly lower than the macrocyclic trisguanidine.



1.23

Anslyn and coworkers have developed many guanidine-containing anion binders, one such being a trisguanidine built off of a triethylbenzene core.<sup>54</sup> The three alternating

ethyl groups on trisguanidine **1.24** sterically enforce the three guanidine units onto the same face of the phenyl ring to preorganize them to bind to the biologically relevant tricarboxylate citrate. This structure binds to citrate in water with a  $K_a$  of  $7 \times 10^3$  by NMR titration. Analogous structures either lacking the ethyl groups or substituting the guanidine groups with amines led to citrate binding affinities of less than half of that of **1.24**. These results further prove that preorganization of the guanidine units can be crucial and reinforces that while guanidinium's charge is more diffuse than ammonium's, the bidentate hydrogen bonding of guanidinium is better suited for binding oxoanions. Trisguanidine **1.24** is so efficient at binding to citrate, even in the presence of other anions, that Anslyn and coworkers used it to detect citrate concentration in orange juice and sports drinks.<sup>55</sup> Sometime later, Schmuck and coworkers synthesized a similar structure (**1.25**) using their guanidiniocarbonyl pyrrole moiety. This so-called “molecular flytrap” boasts a  $K_a > 10^5$  binding affinity to citrate in water.<sup>32</sup>



Another notable trisguanidine from Anslyn and coworkers incorporates a copper cation to enforce preorganization and aid in binding.<sup>56</sup> UV–Vis titration with inorganic phosphate in 2% methanol/water reveals a binding affinity of trisguanidine (**1.26**) of  $K_a = 1.5 \times 10^4 \text{ M}^{-1}$ . This trisguanidine is also remarkably selective; titrations with other anions such as acetate, nitrate, and sulfate lead to binding affinities  $< 100 \text{ M}^{-1}$ . The authors

also titrated an analog of **1.26** in which the guanidines were substituted with azido groups. They observed a binding affinity to phosphate of  $4.0 \times 10^3 \text{ M}^{-1}$ . This affinity indicates that the copper cation plays a more significant role in binding phosphate than expected but that the guanidine groups are essential in increasing the binding and enforcing selectivity.

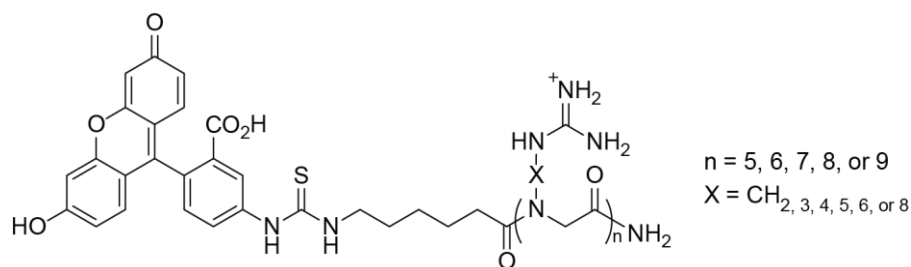
While many studies have been performed in the realm of supramolecular chemistry on guanidine-containing molecules' anion-binding abilities, few researchers have applied this knowledge to the subject of cellular uptake.

### **Abiotic guanidine-containing structures for cell penetration**

Since the discovery of the cell-penetrating capabilities of HIV-tat and penetratin, researchers have developed a plethora of guanidine-containing molecular transporters. Wender, a leader in guanidine-mediated transport, and coworkers have published an excellent review of the variety of molecules produced throughout the past twenty years.<sup>57</sup> As such, only select molecules will be reviewed here including recent work published after this review.

**Peptoid transporters.** After the discovery of HIV-tat and its ability to promote cell entry, Wender, Rothbard, and coworkers were the first to anoint the guanidino group as the primary contributor to this unique ability. The authors synthesized various fluorescently labeled peptide or peptoid oligomers and observed their uptake into Jurkat cells. They demonstrated that peptidic oligomers of lysine, histidine, and ornithine are much worse at promoting cellular uptake than oligomers of guanidine.<sup>58</sup>

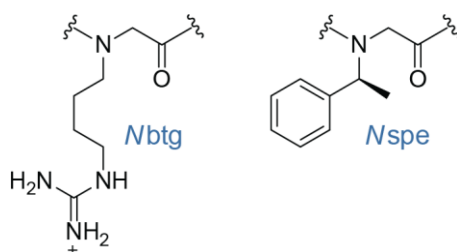




**Figure 1.6.** Wender and coworkers synthesized peptoid oligomers of varying length (5-9 units), and with peptoid monomers with different side chain spacers (2-8 carbons) between the backbone and guanidine residue.

They also synthesized a variety of peptoid oligomers (Figure 1.6). They found that, like peptide-based arginine oligomers, peptoid arginine oligomers enter cells with increasing efficacy based on length, with a nonamer exhibiting the highest cell uptake. They also observed that the distance between the backbone and guanidino group could be altered, with a six-methylene spacer exhibiting greater uptake than arginine or peptido-arginine with a three-methylene spacer. This paper solidly demonstrated guanidine's privileged ability to guide the translocation of cargo into cells.

Many years later, researchers are still studying peptoids for cellular transport. Barron and coworkers have recently explored fluorescently labeled amphipathic peptoids containing the guanidino monomer *N*btg and the lipophilic monomer *N*spe.<sup>59</sup> The authors observed that their peptoids entered cells for longer periods of time than their octa-arginine control, likely due to the increased stability of peptoids *in vivo*. Amphipathic peptoids containing a mix of *N*btg and *N*spe lead to increased cellular uptake compared to octa-arginine and octa-*N*btg.

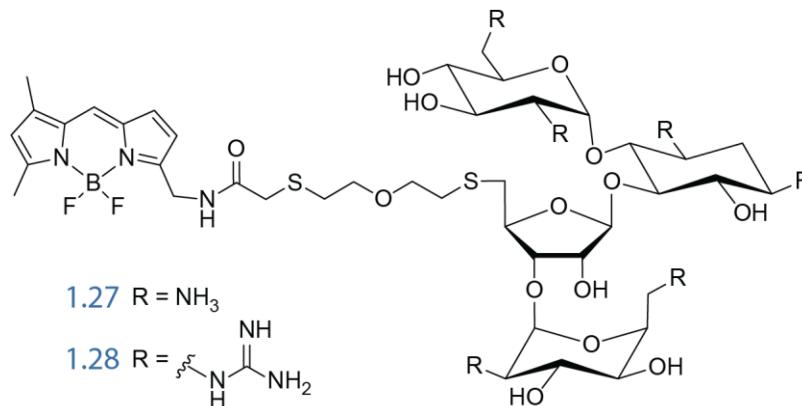


Interestingly, the amphipathic peptoid with a higher number of Nbtg residues resulted in less cytotoxicity; this is unexpected as molecular transporters with high positive charge lead to cytotoxicity.<sup>60–62</sup> Another vital observation from this study is that fluorescence microscopy studies of octa-arginine and octa-Nbtg showed diffuse staining indicative of direct translocation, while amphipathic peptoids showed punctate staining indicating endocytosis. The amphipathic peptoids also demonstrated limited cell uptake at 4 °C, indicating that the uptake of these peptides is controlled by energy-dependent pathways such as endocytosis. These results are reminiscent of the natural CPP penetratin, which is also amphipathic with a mix of arginines and tryptophans and also enters cells mainly via endocytosis.<sup>23</sup>

**Antibiotic transporters.** Sugars have proven to be another exciting scaffold for multi-guanidine-mediated cellular uptake. Chung and coworkers have utilized a dimeric structure of perguanidinylated inositol to observe cellular penetration of the blood–brain barrier.<sup>63</sup> Fascinatingly, Tor and coworkers have taken advantage of the aminoglycoside scaffolds of antibiotics to produce poly-guanidinylated transporters. The authors have shown that guanidinylated neomycin (**1.28**) accomplished a 20-fold increase in cellular fluorescence compared to unmodified neomycin (**1.27**). Guanidinylated neomycin even showed slightly increased cellular uptake compared to Arg<sub>9</sub>.<sup>64</sup>

Tor and coworkers later used these guanidinoglycosides to probe the role of heparan sulfate in cellular uptake. The authors produced structures containing two guanidinoglycosides connected through PEG linkers. These dimeric structures were able to recapitulate cellular uptake in cells with reduced heparan sulfate expression when monomeric structures were unable to do so. In addition, the authors used cell lines

incapable of forming heparan sulfated glycans and cell lines in which only sulfating enzymes were deficient, both of which were unable to take up the guanidinoglycosides. These results show that it is not the carbohydrates of the glycosaminoglycan structure that interact with guanidine-containing transporters, but specifically, the sulfate groups forming interactions with guanidines that enable the initiation of endocytosis.<sup>65</sup>

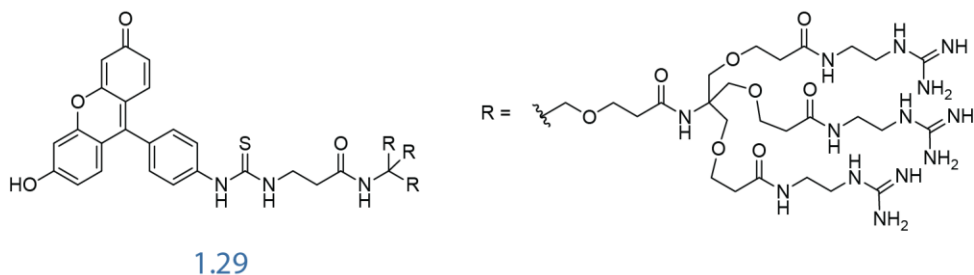


Tor and coworkers have further probed heparan sulfate's role in cell uptake through biotinylated guanidinoglycosides. After first coating the surface of cells in biotinylated guanidinoglycosides, the authors then introduced streptavidin, which in its tetrameric form pulls several of the biotinylated structures toward another. This clustering of guanidinoglycosides promoted endocytosis, providing further evidence that heparan sulfate clustering is a trigger for the endocytosis of guanidine-mediated transporters.<sup>66</sup> Lastly, Tor and coworkers have investigated cyclic peptide-based antibiotics as cellular transporters; however, they observed that both the unmodified and guanidinylated versions of polymyxins showed similar levels of uptake into cells.<sup>67</sup>

**Dendrimer transporters.** Dendrimers are high molecular weight, highly branched polymers that can display dozens of guanidine units per molecule. Bonduelle and Gillies

have reviewed the many dendrimers produced as cost-efficient and easy-to-synthesize analogs of CPPs.<sup>68</sup>

One of the first dendrimers produced for cellular uptake was that of Goodman and coworkers.<sup>69</sup> They observed that while all the dendrimers they synthesized could transport fluorescein into cells, dendrimers with 6, 9, and 12 guanidino groups had much higher uptake than dendrimers with 1 or 3 guanidines. They determined that the structure with nine guanidines (**1.29**) was nontoxic to HeLa cells. They also demonstrated that **1.29** conjugated to GFP could translocate into cells with the same efficiency as Tat-GFP, making dendrimers a promising tool for cellular delivery of high molecular weight molecules.



Bianco and coworkers have recently developed a new dendrimer that they have termed a HYDRAmer. These structures incorporate amino or guanidino groups linked through short PEG chains to an adamantane core. In this way, the authors can make a dendrimer with a trio of guanidino-capped branches, or a trio of trios with three additional adamantane foci creating nine guanidino-capped branches. The authors have shown the internalization of these fluorescently labeled dendrimers and proven them to be non-cytotoxic to HeLa cells and RAW 264.7 cells, a macrophage cell line.<sup>70</sup>

The authors performed studies to probe the pathways for cellular internalization by introducing a variety of endocytosis-inhibiting factors.<sup>71</sup> They found that the uptake of all HYDRAMers in RAW 264.7 cells decreased upon introduction of clathrin-inhibiting

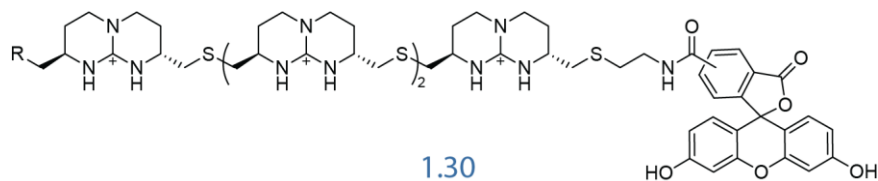
chlorpromazine, however, they still observed diffuse staining in cells under these conditions. Surprisingly, upon introducing micropinocytosis-inhibiting m $\beta$ CD, the amino-HYDRAMers showed decreased internalization while the nona-guanidino-HYDRamer showed *increased* internalization. The authors propose that their HYDRAMers can translocate both through endocytosis and passive diffusion, with the guanidino-HYDRAMers being particularly effective. The authors performed octanol/water partitioning experiments and found that while 99% of both the nona-amino-HYDRamer and nona-guanidino-HYDRamer remained in the water layer, with the addition of sodium dodecanoate, the nona-guanidino-HYDRamer partitioned into 51% into octanol. In comparison, the nona-amino-HYDRamer only partitioned 15% into octanol. These studies further demonstrate guanidine's superior ability to facilitate cellular transport through its interactions with oxoanions and elucidate some of the mechanisms through which guanidine-mediated transporters can enter cells.

**Non-covalent transporters.** Perhaps the most popular synthetic guanidine-containing transporters recently have been polymers capable of non-covalently ushering molecules, frequently genetic material, into cells.<sup>72-74</sup> Stenzel and coworkers have developed a zwitterionic guanidine-containing polymer that forms micelles capable of cell entry.<sup>75</sup> The authors developed block copolymers with a PEG core and arginine-based exterior in which both the side chain and carboxylate are free. The authors then tested the free carboxylate and methylated carboxylate versions of the guanidinylated micelle. They observed that while the methylated and non-methylated fluorescein-labeled micelles entered cells, the methylated micelles exhibit much higher cytotoxicity. The authors determined that this was due to the relative zeta potential of each micelle. The methylated

micelle is highly positively charged, while the zwitterionic micelle has an overall negative charge. It is a surprising result of this study that the overall charge of a transporter does not matter as much as its ability to form strong binding interactions with anions of the cell surface.

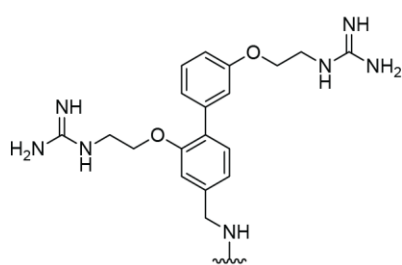
**Small molecule transporters.** While there has been much development of peptides, dendrimers, and other high molecular weight guanidine-containing transporters, less attention has been devoted to small-molecule transporters. High molecular weight transporters benefit from having a high number of guanidino groups displayed on their structures, resulting in higher valency interactions between the cell surface and the transporter. However, a more compact approach can be beneficial.

While not strictly a small molecule, Mendoza, Giralt, and coworkers have developed a tetramer of the 6,6-bicyclic guanidine small molecule discussed in Section 1.4.1 for cell uptake.<sup>76</sup> The fluorescently labeled tetraguanidinium **1.30** penetrates cells to a much greater extent than does Tat or penetratin. The authors determined that the structure was not cytotoxic at low concentrations (<5  $\mu$ M). The authors also observed that while Tat and penetratin localize in the nucleus, the tetraguanidinium appears exclusively in the cytosol, even colocalizing with mitochondria. The authors observed a decrease in uptake at 4 °C but rapid internalization within 5 min at 37 °C, indicating that the structure is internalized through both energy-dependant and independent pathways.

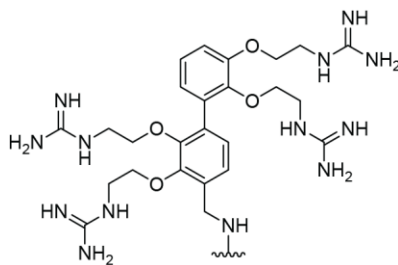


Subsequently, Mascareñas and coworkers have attached bicyclic guanidine oligomers to a transcription factor peptide fragment. They observed cellular uptake to an equal extent as the peptide-Arg<sub>3</sub> control. In addition, the bicyclic guanidine oligomer participated in the binding of DNA, but only in the presence of an A,T-rich region.<sup>77</sup>

Williams, Selwood, and coworkers first demonstrated the capacity of a truly small molecule to facilitate cell uptake, which they termed a small molecule carrier (SMoC).<sup>78</sup> The authors designed these biphenyl scaffolds to mimic the amphipathic  $\alpha$ -helices of penetratin, though whether their remarkable translocation efficacy is due to their mimicking of an  $\alpha$ -helix or merely due to the guanidine moieties is unclear. Initially, the authors determined that fluorescein-labeled SMoCs **1.31** and **1.32** entered cells with equal efficiency as Tat. Both **1.31** and **1.32** are found in the cytoplasm and nucleus, with **1.32** showing a slightly higher preference for the nucleus. They also found the fluorescently-labeled SMoCs to be nontoxic to the cell types tested.



1.31

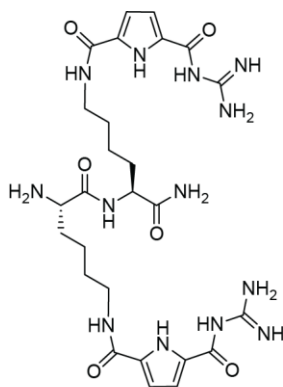


1.32

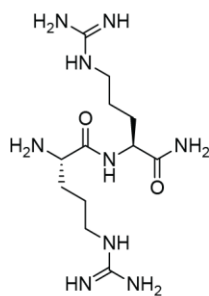
The authors then appended two units of **1.32** onto the fluorescently labeled protein geminin through disulfide bonds on native cysteines and observed translocation into cells. The **1.32**-labeled geminin also remains functional and inhibits the progression of G<sub>0</sub> phase to S phase in various cell types. The authors observe that both **1.32**-fluorescein and **1.32**-geminin demonstrate lesser cellular uptake at 4 °C. While the uptake of **1.32**-fluorescein is not affected by any endocytosis inhibitors, uptake of **1.32**-geminin is

partially inhibited by chlorpromazine, indicating clatherin-mediated endocytosis guides at least part of the uptake of geminin. This dichotomy likely indicates that the pathway of SMOc-guided cell uptake is dependent on the cargo size, similar to canonical CPPs.

Schmuck and coworkers have recently become the first to show that guanidinium compounds modified to have enhanced hydrogen bonding capability have increased cellular uptake efficacy.<sup>79</sup> The authors use the guanidiniocarbonyl pyrrole **1.6** discussed previously to synthesize an arginine analog. They demonstrate that while an arginine dipeptide **1.34** has insignificant binding to heparin by ITC, the guanidiniocarbonyl pyrrole dipeptide **1.33** binds to heparin with an affinity in the range of  $10^7$ . They also observe that rhodamine-labeled **1.33** enters cells, whereas rhodamine-labeled **1.34** does not. Both dipeptides were shown to be nontoxic. The authors also synthesized biotin-labeled **1.33** and **1.34**, which they appended to fluorescently labeled avidin, labeling the avidin with four units of **1.33** or **1.34**. They observed that again, avidin-**1.33** entered cells while avidin-**1.34** did not. The authors also observed that avidin-**1.33** had diminished uptake into CHO cells with reduced glycosaminoglycan expression, indicating that heparan sulfate is necessary to import the guanidiniocarbonyl pyrrole transporters the same as any guanidine-based transporter.



**1.33**



**1.34**



## Summary

The cellular delivery of proteins has challenged both pharmaceutical development and chemical biology. The facile delivery of proteins is necessary not only to probe phenotypes of cells but also to produce viable therapeutics. There are many methods to facilitate delivery into cells, but one of the most popular is the cell-penetrating peptide. CPPs have been used in translocation across the field of chemical biology, as well as investigated in therapeutics in clinical trials.<sup>80–82</sup> CPPs have also spawned a vast array of guanidine-containing transporters. These transporters include dendrimers, non-covalent guanidinium lipids, and small molecules. By studying the supramolecular chemistry of guanidinium-containing hosts and their anion guests, insights can be made about the mechanism of action of cell-penetrating peptides and methods to improve upon them.

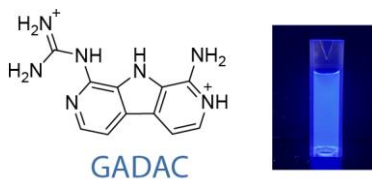
In the following chapters, I investigate preorganized guanidinium-containing structures. These studies in Chapters 2 and 5 elaborate on the monoguanidine and bisguanidine prior work discussed in this review. In addition, I investigate improvements to peptide transporters (Chapters 3 and 4) by incorporating factors that influence binding as illustrated in this review, such as altering the guanidinium  $pK_a$ .

## **Chapter 2**

### **A Fluorescent Guanidinium-Azacarbazole for Oxoanion Binding and Cell Penetration**

Contributions: Yana Petri performed the *in vitro* microscopy experiments. Dr. Luke Lavis performed the quantum yield determinations. All other work was performed by Lindsey O. Calabretta.

## Abstract



The transport of molecules into human cells can be effected by conjugation to a cationic moiety and probed with fluorophores. I report on a molecule that accomplishes both objectives. 1-Guanidino-8-amino-2,7-diazacarbazole dichloride (**GADAC**) binds to a carboxylate, phosphate, and sulfate in water with affinities of  $3.6 \times 10^4 \text{ M}^{-1}$ ,  $1.1 \times 10^3 \text{ M}^{-1}$ , and  $4.2 \times 10^3 \text{ M}^{-1}$ , respectively. Due to its diazacarbazole core, **GADAC** is fluorescent in water ( $\lambda_{\text{abs}} = 356 \text{ nm}$ ,  $\lambda_{\text{em}} = 403 \text{ nm}$ ,  $\epsilon = 13,400 \text{ cm}^{-1}\text{M}^{-1}$ ). The quantum yield of **GADAC** is pH-sensitive, increasing from  $\Phi = 0.12$  at pH 7.4 to  $\Phi = 0.53$  at pH 4.0 as a result of the protonation of the aminopyridine moiety. Due to the relative hydrophobicity of **GADAC** and its strong association with anions found on the cell-surface, **GADAC** can penetrate mammalian cells. The uptake of **GADAC** into M21 melanoma cells is detectable in the DAPI channel at single-digit micromolar concentrations. **GADAC**'s unique properties make it a promising candidate for a two-in-one protein modification system to simultaneously promote cell uptake and monitor cellular localization via fluorescence.

## Introduction

Small molecules have dominated the field of drug discovery for most of the last century.<sup>83</sup> While small molecules are advantageous in terms of their bioavailability and synthetic accessibility, some proteins of interest (POIs) remain “undruggable” by small molecules because they lack defined ligand-binding pockets. Biologics such as monoclonal antibodies possess a greater surface area and structural diversity than do small molecules and can thus bind such challenging protein targets more tightly and selectively. In 2018, eight of the top-ten selling drugs were biologics.<sup>84</sup> Biologics are, however, often restricted to targeting extracellular or membrane-bound proteins due to their large size and hydrophilicity. Approaches to improve the uptake of biologics into the cytosol of mammalian cells would unlock a greater range of metabolic processes that could be targeted by biologics.

Strategies have been used to promote the transport of biologics into the cytosol of cells, such as encapsulation into lipid nanoparticles or modification with cell-penetrating peptides (CPPs).<sup>3,4,82,85-90</sup> Arginine-rich CPPs are class of transporters that rely on the binding of guanidino groups to cell-surface anions. CPPs such as the HIV-tat peptide are rich in arginine residues, which, upon binding to the cell surface, are thought to promote cargo uptake via direct transduction through the plasma membrane, either through transient micellar pores or by clustering heparan sulfate proteoglycans, which triggers endocytosis and subsequent endosomal escape.<sup>22,91</sup> Although many groups have successfully used oligoarginines to promote cellular uptake, these modifications are bulky, and more compact small molecule modifications could be beneficial.

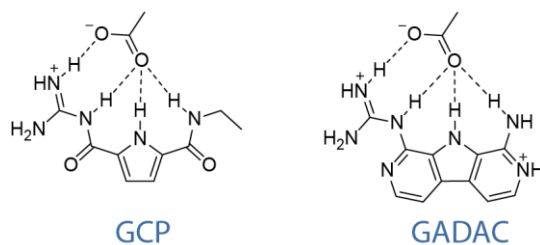
Few guanidine-containing small molecules have been explored as cell-uptake promoters. Selwood and coworkers have synthesized small-molecule carriers (SMoCs),

one of which contained four guanidino groups and mimicked the amphipathic  $\alpha$ -helical nature of some CPPs.<sup>78,92</sup> The researchers appended multiple SMOCs and a fluorescent label to the proteins geminin and Tev and observed their enhanced cellular uptake via microscopy.<sup>93</sup> In another study, Schmuck and coworkers appended multiple dimers of a guanidiniocarbonyl pyrrole (**GCP**) onto fluorescently labeled avidin and observed that the **GCP**-labeled avidin entered cells whereas avidin labeled with dimers of arginine did not.<sup>79</sup>

The ability of **GCP** to promote cellular uptake lies in its enhanced ability to bind to oxoanions in comparison to an unsubstituted guanidine. In a preliminary study, Schmuck determined that **GCP** was capable of binding to Ac-Ala-O<sup>-</sup> in 40% v/v water in DMSO with an affinity of 770 M<sup>-1</sup>, whereas guanidine showed an affinity of <10 M<sup>-1</sup>.<sup>30</sup> In addition, **GCP** was found to bind to acetate with an affinity of 2800 M<sup>-1</sup> (Scheme 2.1).<sup>94</sup> The ability of **GCP** to bind to oxoanions even in aqueous conditions is in part due to its low guanidino pK<sub>a</sub> of ~7 relative to that of arginine, which has a pK<sub>a</sub> of 13.8.<sup>29,30</sup> **GCP** also greatly benefits from two additional hydrogen bond donors on the pyrrole and amide. Schmuck was able to determine the effect of pK<sub>a</sub> and each hydrogen-bond donor through sequentially determining the binding affinity of acetylguanidine, **GCP** without the amide at the 4-position, and **GCP**.<sup>30</sup>

Whereas the binding of **GCP** to carboxylates is strong, anion affinity is hampered by the rotatable bonds between its acylguanidino and amido groups and its pyrrole ring. Schmuck and coworkers have computationally determined that **GCP** is more stable in the conformation in which the amide carbonyl oxygens are pointing inward toward the pyrrole nitrogen since this conformation balances the orientation of the amide and pyrrole dipoles.<sup>95</sup> Rotating the acylguanidino group into the binding conformation

requires  $2.5 \text{ kcal}\cdot\text{mol}^{-1}$ , and orienting both the acylguanidino and amido groups into the binding conformation requires  $5 \text{ kcal}\cdot\text{mol}^{-1}$ .



**Scheme 2.1.** Binding of Schmuck's guanidiniocarbonyl pyrrole (**GCP**) and the proposed binding of 1-guanidino-8-amino-2,7-diazacarbazole (**GADAC**) to acetate.

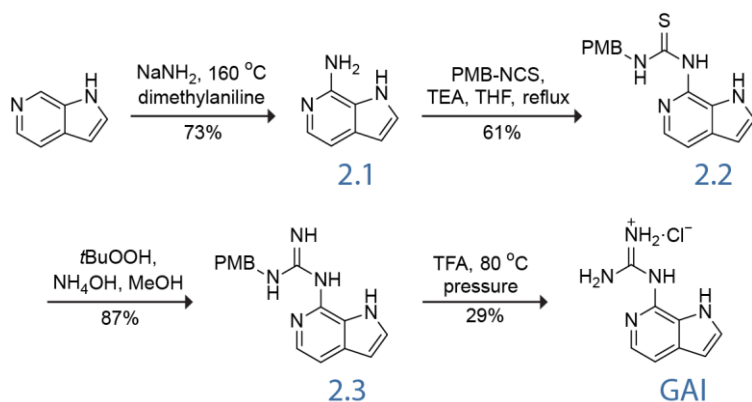
**GCP** is an effective molecular transporter due to its high affinity for oxoanions. I reasoned, however, that a preorganized structure could be even more effective. Towards that end, I conceived of 1-guanidino-8-amino-2,7-diazacarbazole (**GADAC**), which contains four hydrogen-bond donors in the same positions as **GCP** (Scheme 2.1). Due to its diazacarbazole core, the guanidino and amino groups of **GADAC** cannot rotate out of the conformation desirable for binding to an anion. A 2-guanidinopyridine has been previously observed through  $^1\text{H}$  NMR and X-ray crystallography to contain an intramolecular hydrogen bond between the pyridino nitrogen and guanidino hydrogen.<sup>96</sup> **GADAC** could also display this intramolecular hydrogen bond to enforce the planarity of the system and further preorganize the structure toward binding oxoanions. I expected the guanidino  $\text{p}K_{\text{a}}$  of **GADAC** to be lowered based on previous reports of a 2-guanidinopyridine having a  $\text{p}K_{\text{a}}$  of 9.4.<sup>97</sup> While the pyridino-guanidine may not have as low a  $\text{p}K_{\text{a}}$  as the acylguanidine of **GCP**, **GADAC** also contains a second protonation site that increases its overall positive charge and, thus, its affinity for anions.

Finally, **GADAC** contains an azacarbazole core. The heterocycle 2-azacarbazole, also known as  $\beta$ -carboline, is a natural product found in the carapace of scorpions and is

partially responsible for their ability to fluoresce under ultraviolet light.<sup>98</sup> Azacarbazoles and diazacarbazoles have been used as fluorescent probes<sup>99–101</sup> and organic light-emitting diodes.<sup>102–104</sup> I wondered what effects on fluorescence, if any, would result from adding a guanidino group to an azacarbazole or diazacarbazole core. I envisioned that if **GADAC** retains the fluorescence of the diazacarbazole core and has a high affinity for oxoanions, it could be useful as a two-in-one molecular transporter and fluorescent reporter.

## Results and Discussion

**Synthesis of guanidino compounds.** I began the synthesis of **GADAC** by first optimizing the late-stage guanidinylation conditions on a model compound, 7-guanidino-6-azaindole (**GAI**). (Scheme 2.2). I first synthesized aminoazaindole **2.1** from 6-azaindole through a Chichibabin reaction<sup>105,106</sup> in good yield, with exclusive selectivity for amination at the 7-position. I subsequently attempted to guanidylate this amine with the commonly used *N,N'*-di-Boc-methylisothiurea. Instead of guanidinylation, I observed that the Boc protecting group transferred to the amino group in the major product (see: Unsuccessful Routes, Figure 2.6). To avoid this unexpected byproduct, I instead converted the amino group into a thiourea. I selected a *p*-methoxybenzyl-isothiocyanate (PMB–NCS) instead of a more commonly used carbamate-protected isothiocyanate to effect this transformation, again to avoid the transfer of the carbamate onto the amino group. I successfully appended the thiourea to produce compound **2.2**, which I subsequently converted to PMB-guanidine **2.3** under mild conditions by treatment with ammonium hydroxide and *t*-butyl hydroperoxide.<sup>107</sup> The PMB group was removed with TFA at 80 °C to produce **GAI**.



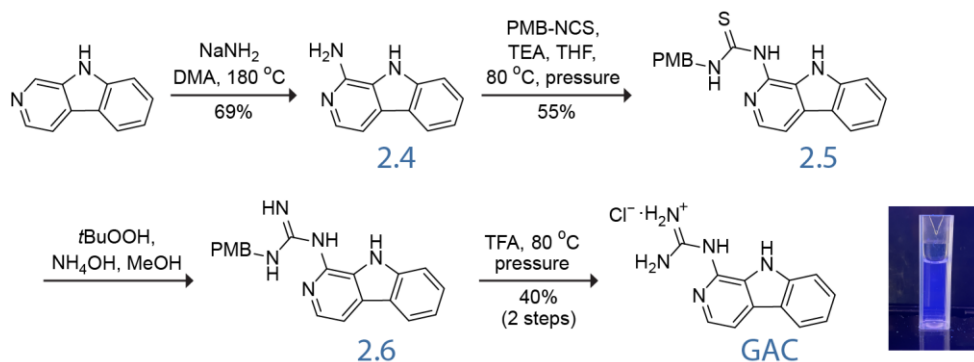
**Scheme 2.2.** Synthesis of the model compound, 7-guanidino-6-azaindole (**GAI**).

Before producing **GADAC**, I first applied this optimized synthetic route to the fluorescent natural product  $\beta$ -carboline to produce 1-guanidino-2-azacarbazole (**GAC**). I sought to compare the ability of 2-guanidinopyridine (**GP**), **GAC**, and **GADAC** to bind anions to assess the role of each hydrogen-bond donor in a stepwise fashion, in analogy to the experiments performed by Schmuck.<sup>30</sup> In addition, I wanted explore the effects of adding a guanidino group to an azacarbazole core on fluorescence.

The process was easily replicable, and I successfully produced amine **2.4**, followed by thiourea **2.5**, PMB-guanidine **2.6**, and, finally, **GAC** (Scheme 2.3). I noted that the guanidinylation of an azacarbazole core did not eliminate fluorescence—**GAC** glows violet under a 366-nm handheld UV lamp (Scheme 2.3). Unexpectedly, **GAC** is not appreciably soluble in water. In addition, I observed significant signal broadening of the carbon *ipso* to the guanidino group and, to a lesser extent, the other carbons of this ring through <sup>13</sup>C NMR spectroscopy in DMSO-*d*<sub>6</sub> (see: NMR Spectra, p. 118). **GAI** also showed broadening of the *ipso* carbon in DMSO-*d*<sub>6</sub> but did not show broadening in D<sub>2</sub>O. This spectrum likely indicates an intramolecular hydrogen bond occurring between a guanidino hydrogen and the pyridine nitrogen in DMSO. This hydrogen bond restricts the rotation of the



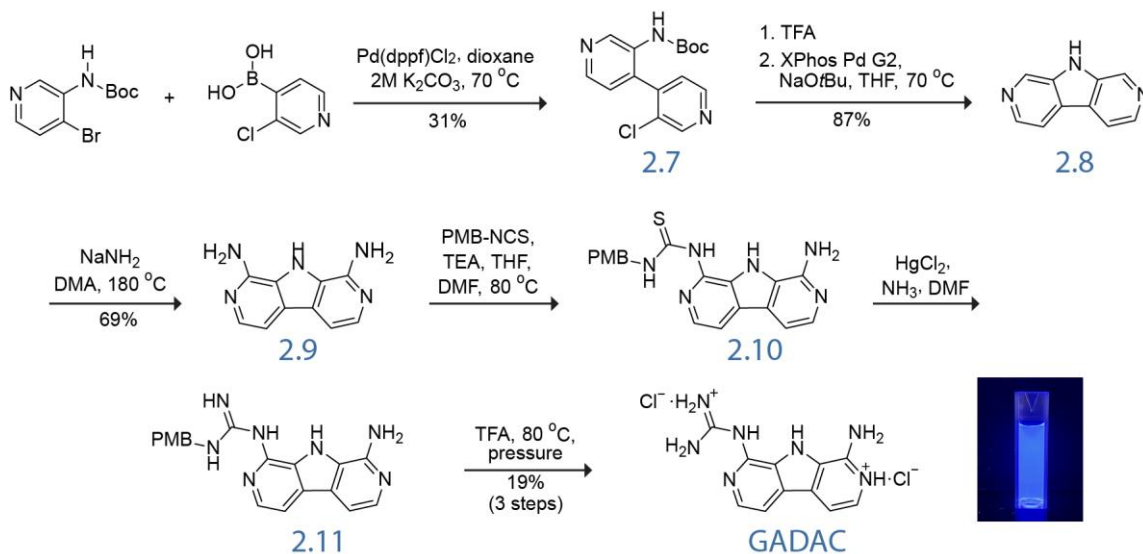
guanidine but its strength is diminished in a protic solvent such as water. A similar intramolecular hydrogen bond has been previously observed through  $^1\text{H}$  NMR spectroscopy and X-ray crystallography to stabilize a 2-guanidinopyridine.<sup>96</sup>



**Scheme 2.3.** Synthesis of 1-guanidino-2-azacarbazole (**GAC**) (center). **GAC** (300  $\mu\text{M}$  in ethanol) under a handheld 366-nm UV lamp (bottom right).

After optimizing the guanidinylation route with **GAI** and **GAC**, I sought to synthesize **GADAC**. To accomplish this goal, I first needed to synthesize 2,7-diazacarbazole **2.8** (Scheme 2.4). I attempted a Pictet–Spengler-inspired route using 6-azaindole. I successfully produced 6-azatryptamine (**2.18**), but the electronics of the ring system were not conducive to closing the third ring with formaldehyde (see: Unsuccessful Routes, Scheme 2.6). I then explored using organometallic couplings to form 3-amino-3'-chlorobipyridine **7**, which could be cyclized to 2,7-diazacarbazole **2.8** (Scheme 2.4). I reacted *N*-Boc-3-amino-4-bromopyridine and 3-chloro-4-pyridineboronic acid with a range of palladium catalysts (see: Unsuccessful Routes, Table 2.3). I expected this reaction to be of moderate yield since coupling reactions of two heterocycles are known to be difficult. In addition, (halopyridyl)boronic acids are susceptible to protodeboronation and side reactions.<sup>108</sup> Using  $\text{Pd}(\text{dppf})\text{Cl}_2$  as a catalyst, I was able to obtain a serviceable yield of product, and by using the Boc-protected amino

bromopyridine to aid in purification, I was able to recover up to 70% of the unreacted bromopyridine, allowing us to re-expose the bromopyridine to the reaction conditions to build a stock of bipyridine **2.7**.

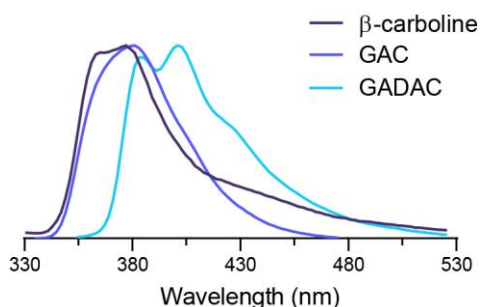


**Scheme 2.4.** Synthesis of 1-guanidino-8-amino-2,7-diazacarbazole (**GADAC**) (center). **GADAC** (5  $\mu$ M in Milli-Q water) under a handheld 366-nm UV lamp (bottom right).

The Boc group was subsequently cleaved with TFA, and the bipyridine was exposed to XPhos Pd G2 and sodium *tert*-butoxide to produce 2,7-diazacarbazole **2.8**. The Chichibabin amination was then performed on compound **2.8**, resulting in exclusively 1,8-diamino isomer **2.9**. Subsequently, diamine **2.9** was introduced to PMB–NCS to append the thiourea. Interestingly, even with a 10-fold excess of PMB–NCS, the reaction added only one thiourea moiety. This reaction must also be monitored carefully to avoid product degradation. I attempted the previously discussed peroxide-mediated guanidinylation with thiourea **2.10**, but this reaction was unsuccessful. Instead, I used the traditional mercury-catalyzed reaction to obtain PMB-guanidine **2.11**. The product was difficult to purify from the reaction mixture due to the fragility of thiourea **2.10** and insolubility of PMB-guanidine **2.11** and was therefore used directly. The PMB group was

subsequently cleaved with TFA and heat to produce guanidino-amino-diazacarbazole **GADAC**, which is also fluorescent under a 366-nm handheld UV lamp (Scheme 2.4).

**Evaluation of fluorescence.** I obtained the absorbance and emission spectra for **GAC** and **GADAC** in both methanol and water and were gratified to observe that both retained the fluorescence of the azacarbazole core (Figures 2.7-2.9). I observed that compounds **2.4**, **2.8**, and **2.9** were also fluorescent (Figures 2.10 and 2.11). Apparently, the addition of the guanidino group did not eliminate fluorescence. Interestingly, increasing functionality of the compound induced a redshift—the  $\lambda_{em}$  of **GAC** in methanol (380 nm) is slightly higher than that of  $\beta$ -carboline (375 nm), whereas the  $\lambda_{em}$  of **GADAC** (400 nm) is much higher (Figure 2.1).

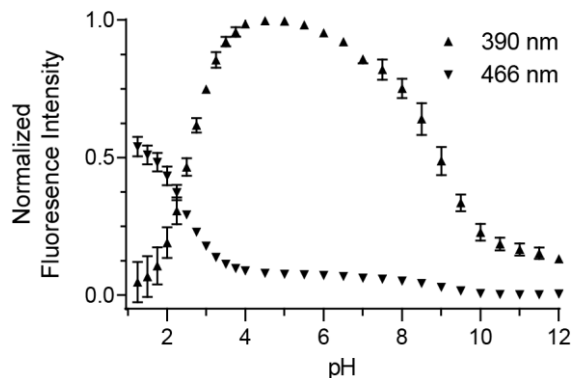


**Figure 2.1.** Normalized fluorescence emission spectra of  $\beta$ -carboline ( $\lambda_{em} = 375$  nm), **GAC** ( $\lambda_{em} = 380$  nm), and **GADAC** ( $\lambda_{em} = 400$  nm), in methanol, showing a redshift upon increasing functionalization of the azacarbazole core.

**Fluorescence characterization of GAC.** Given the number of basic nitrogens in **GAC** and **GADAC**, I was interested in observing their fluorescence in different protonation states. I monitored the fluorescence of these compounds in water at a range of pH values and observed that the fluorescence intensity was indeed related to the protonation state. For **GAC**, I observed that  $\lambda_{em} = 466$  nm at low pH, which shifted to  $\lambda_{em} = 390$  nm at higher

pH (Figure 2.2). This shift was expected, since  $\beta$ -carboline also shows a redshift when protonated at the pyridino nitrogen.<sup>109</sup> Thus, the  $\lambda_{em} = 466$  nm likely corresponds to the dication of **GAC** with protons on both the guanidino and pyridine groups. I believe the shift to  $\lambda_{em} = 390$  nm near neutral pH corresponds to the deprotonation of the pyridinium nitrogen and determined the  $pK_{a1}$  of **GAC** to be 2.57 (Figure 2.12).

The fluorescence at 390 nm is maintained through mid-range pH values but diminishes six-fold at pH 11.5 compared to pH 4.0. Previously, a 2-guanidinopyridine has been shown to have a  $pK_a$  of 9.4.<sup>97</sup> I was also able to determine the guanidinium deprotonation of **GAI** to have a  $pK_a$  of 9.43 via  $^1H$  NMR spectroscopy (Figure 2.15). Therefore, I believe the decrease of **GAC** fluorescence at high pH is due to the deprotonation of the guanidinium group and determined  $pK_{a2}$  of **GAC** to be 8.96 (Figure 2.12).

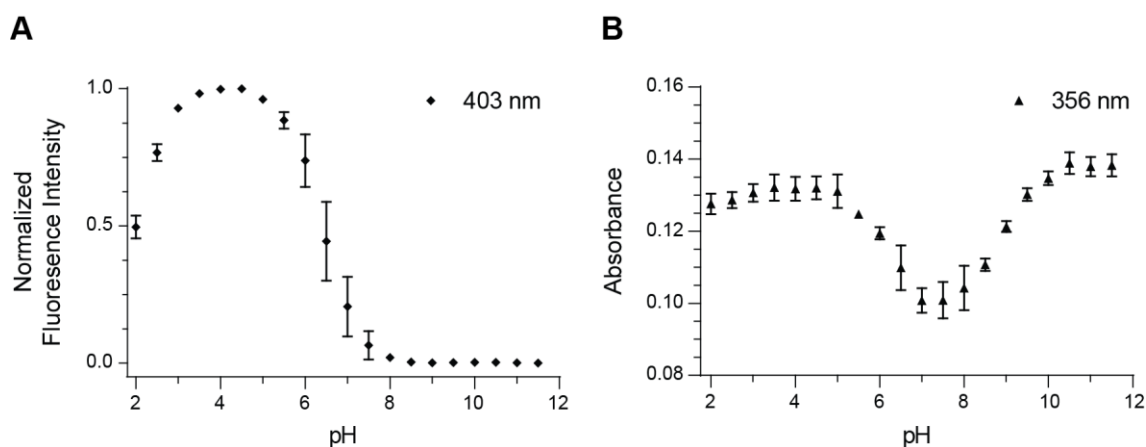


**Figure 2.2.** Fluorescence emission intensities (normalized, excitation 330 nm) of **GAC** in water containing NaCl (10 mg/mL), with changing pH value affecting fluorescence intensity and wavelength.

In addition to characterizing the response of the fluorescence of **GAC** to pH, I also determined the extinction coefficient of **GAC** in water to be  $\epsilon = 6400 \pm 30 M^{-1}\cdot\text{cm}^{-1}$  (Figure 2.17).

**Fluorescence characterization of GADAC.** After observing how pH affects the fluorescence of a guanidinium-containing azacarbazoles, I moved on to characterizing the fluorescence properties of **GADAC**. Similar to **GAC**, I observed a pH-dependent change in fluorescence at the  $\lambda_{em} = 403$  nm for **GADAC** (Figure 2.3).

As with **GAC**, the fluorescence intensity of **GADAC** trends downward at low pH. Yet, no shift in wavelength was observed, even at  $\text{pH} < 2$ . The decrease in fluorescence is likely caused by the formation of the trication from the protonation of both pyridine nitrogens and the guanidino group; however, the  $\text{p}K_{a1}$  of **GADAC** was too low to measure accurately. The fluorescence intensity also decreases by 30-fold from  $\text{pH} 4.0$  to  $11.5$ . From the curve obtained by plotting the fluorescence data, I determined the  $\text{p}K_{a2}$  of **GADAC** to be  $6.42$  (Table 2.1, Figure 2.13). The  $\text{p}K_{a2}$  of **GADAC** is likely from the protonation of the pyridine nitrogen on the ring that does not contain the guanidino group, since it is similar to the  $\text{p}K_a = 6.86$  of 2-aminopyridine.<sup>110</sup> Interestingly, no change in fluorescence was observed at higher pH that would correspond to guanidine protonation.



**Figure 2.3.** pH-Dependence of the photophysical properties of **GADAC** in water containing NaCl (10 mg/mL). A. Fluorescence emission intensities at  $\lambda_{em}$  (normalized, excitation 340 nm). B. Absorbance ( $\lambda_{abs}$ ).

To determine if this change in fluorescence intensity was due to an increase in the number of photons being emitted or absorbed, I determined the absorbance of **GADAC** at varying pH's. I found that there was a decrease in the absorbance at  $\lambda_{\text{abs}}$  upon moving from low pH to neutral pH, similar to the decrease in fluorescence emission. Yet, upon comparing neutral to high pH, the absorbance rebounds to slightly higher than the absorbance at low pH. The absorbance at the excitation wavelength for the fluorescence pH experiment (340 nm) displays the same trend, though to a lesser extent (Figure 2.14). This indicates that the changes in fluorescence intensity are not wholly due to changes in absorbance.

Given that two inflection points are observable, I used the pH-dependent absorbance data to calculate the  $\text{p}K_{\text{a}2}$  and  $\text{p}K_{\text{a}3}$  of **GADAC** (Figure 2.14). I determined the  $\text{p}K_{\text{a}2}$  to be 6.18, which is congruent with the value obtained from the pH-dependent fluorescence data (Table 2.1). I also determined the  $\text{p}K_{\text{a}3}$  to be 8.92, which is nearly identical to the  $\text{p}K_{\text{a}}$  of the guanidino group in **GAC**.

**Table 2.1.** Photophysical properties and  $\text{p}K_{\text{a}}$  values of **GADAC**. <sup>a</sup>Measured via fluorescence spectroscopy. <sup>b</sup>Measured via absorbance spectroscopy. <sup>c</sup>Citrate buffer (10 mM) with NaCl (150 mM). <sup>d</sup>Phosphate buffer (10 mM) with NaCl (150 mM).

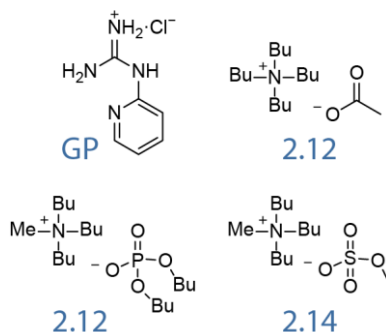
	$\text{p}K_{\text{a}}$		pH 4.0 <sup>c</sup>	pH 7.4 <sup>d</sup>
$\text{p}K_{\text{a}1}$	—	$\lambda_{\text{abs}}$	356	356
$\text{p}K_{\text{a}2}$	$6.42 \pm 0.04^{\text{a}}$ $6.20 \pm 0.1^{\text{b}}$	$\lambda_{\text{em}}$	403	428
$\text{p}K_{\text{a}3}$	$8.92 \pm 0.07^{\text{b}}$	$\epsilon$	$13,400 \pm 20 \text{ cm}^{-1}\text{M}^{-1}$	$12,200 \pm 40 \text{ cm}^{-1}\text{M}^{-1}$
		$\Phi$	$0.526 \pm 0.015$	$0.119 \pm 0.004$

To further quantify the relationship between absorbance and emission at different protonation states, the extinction coefficient (Figure 2.16) and quantum yield were

determined at pH 4.0 and pH 7.4, since pH 4.0 ensures nearly full protonation and pH 7.4 is most relevant for biological applications (Table 2.1). We determined that  $\epsilon = 12,200 \text{ cm}^{-1}\text{M}^{-1}$  and  $\Phi = 0.12$  at pH 7.4. These values are different than at pH 4.0, where  $\epsilon = 13,400 \text{ cm}^{-1}\text{M}^{-1}$  and  $\Phi$  increases to 0.53. Whereas the extinction coefficient at pH 4.0 is slightly higher than that at pH 7.4, it is not dramatically so, meaning that the increase in fluorescence seen at low pH is due to an increase in quantum yield.

The dramatic increase in the quantum yield of **GADAC** occurs upon protonation of the aminopyridine and is caused by the lone pair of the amino group coming into conjugation with the diazacarbazole ring system more so than in its neutral state. Any additional functionality that would block this protonation, such as a second guanidino moiety, would likely be disadvantageous.

**GADAC–oxoaion binding affinities.** I determined the binding affinity of **GADAC** to three commercially-available anions as a proxy for binding to anions on the cell surface, such as the phosphoryl groups of phospholipids or the carboxyl or sulfuryl groups of proteoglycans. I used commercially available acetate **2.12**, phosphate **2.13**, and sulfate **2.14** for these binding studies (Scheme 2.5). I also obtained **GP** for use as a guanidine with no additional hydrogen-bond donors. I hypothesized that the binding affinity for anions would increase in the order: **GP** < **GAC** < **GADAC**. Obtaining the binding affinity of **GAC** and comparing it to the binding of **GP** informs the extent to which the carbazole NH contributes to the binding of anions. I expected **GADAC** to have the highest affinity due to its additional hydrogen bond-donating amino group and ability to form a dication.

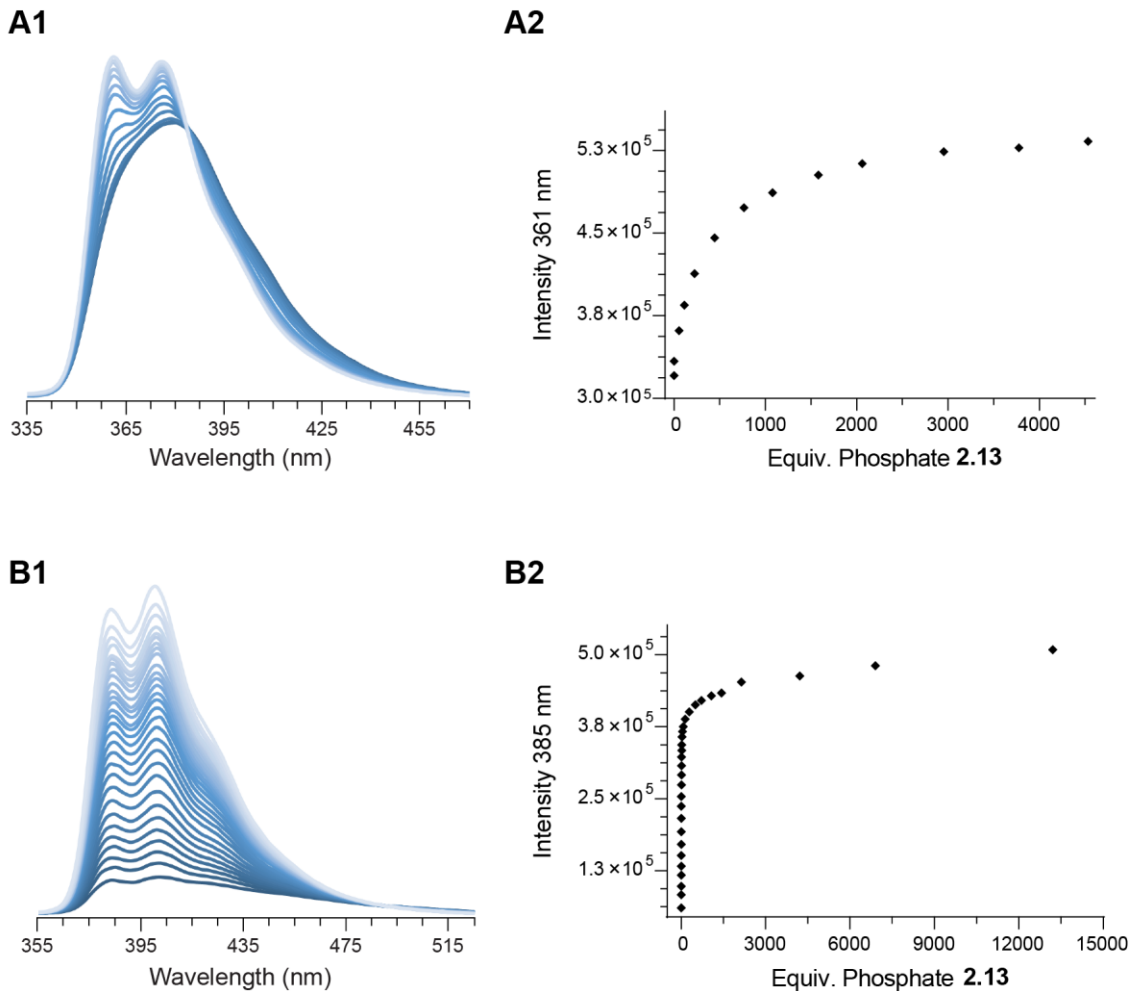


**Scheme 2.5.** Compounds used in binding affinity titrations.

To perform the binding affinity titrations, I introduced sequential amounts of stock solutions of the three anions to solutions of **GP**, **GAC**, and **GADAC** and measured their fluorescence spectra in protic solvents. Both **GP** and **GAC** are chloride salts, whereas **GADAC** was used as the dication and dichloride salt. I observed measurable changes in fluorescence correlated to anion concentration and determined binding affinities from these changes with [supramolecular.org](http://supramolecular.org) (Figure 2.4 and 2.18).<sup>111–113</sup> The fluorescence emission spectra of **GAC** and **GADAC** often shows two maxima. This property is also seen in  $\beta$ -carboline in the presence of acetic acid.<sup>114</sup>

We first explored the binding affinity of **GP** and **GAC** in isopropanol against all three anions (Table 2.2). I was unable to obtain a measurable binding affinity for **GP** to acetate **2.12** or sulfate **2.14**, and I observed only minimal binding to phosphate **2.13**. In contrast, **GAC** displayed a high binding affinity on the order of  $10^4 \text{ M}^{-1}$  with acetate **2.12** and sulfate **2.14**, and  $10^5 \text{ M}^{-1}$  with phosphate **2.13** in isopropanol. Thus, the additional carbazole hydrogen bond donor contributes significantly to the binding of oxoanions in **GAC**.





**Figure 2.4.** Representative fluorescence spectra of binding titrations of **GAC** and **GADAC** in methanol against phosphate **2.13**. Spectra for all binding titrations and links to open source BindFit data can be found in Figure 2.18. A1. Fluorescence spectra of **GAC** titrated with phosphate **2.13**. A2. Change in intensity of **GAC** at 361 nm, unfitted. B1. Fluorescence spectra of **GADAC** titrated with phosphate **2.13**. B2. Change in intensity of **GADAC** at 385 nm, unfitted.

We next compared the binding of **GAC** and **GADAC** with the assumption that the additional amino group of **GADAC** would enhance binding to oxoanions. I attempted to obtain the binding of **GADAC** to the anions in isopropanol but observed higher-order binding that impeded the ability to determine a value for the binding affinity. Instead, I switched to a more competitive solvent, methanol. I observed that the binding affinity of

**GAC** to acetate **2.12** and phosphate **2.13** decreased significantly, but the binding to sulfate **2.14** was nearly unchanged. I believe that this distinction might be due to its low basicity and soft ionic character. In other words, sulfate **2.14** might have a lower energy of desolvation than acetate **2.12** or phosphate **2.13**, allowing sulfate **2.14** to associate more freely with **GAC**, regardless of solvent.

**Table 2.2.** Binding affinities ( $M^{-1}$ ) of **GP**, **GAC**, and **GADAC** to anions **2.12**, **2.13**, and **2.14**. \*A non-competitive 1:2 binding model was used to fit the experimental data.

solvent	compound	acetate <b>2.12</b>	phosphate <b>2.13</b>	sulfate <b>2.14</b>
isopropanol	<b>GP</b>	—	$13 \pm 0.3$	—
	<b>GAC</b>	$(5.3 \pm 0.3) \times 10^4$	$(1.0 \pm 0.1) \times 10^5$	$(1.2 \pm 0.1) \times 10^4$
methanol	<b>GAC</b>	$(1.1 \pm 0.1) \times 10^3$	$680 \pm 30$	$(1.1 \pm 0.1) \times 10^4$
	<b>GADAC</b>	$(1.9 \pm 0.1) \times 10^5$ *	$(2.4 \pm 0.1) \times 10^5$	$(3.3 \pm 0.5) \times 10^5$
	$(K_{1:1}, K_{1:2})$		$288 \pm 70$	$(1.9 \pm 0.2) \times 10^4$
water	<b>GADAC</b>	$(3.6 \pm 0.1) \times 10^4$	$(1.1 \pm 0.0) \times 10^3$	$(4.2 \pm 0.3) \times 10^3$

We determined that **GADAC** was better at binding to oxoanions in methanol than **GAC**. With all three anions, **GADAC** has a  $K_{1:1}$  of  $10^5 M^{-1}$ , which is much greater than that of **GAC**. Interestingly, **GADAC** also exhibits a 1:2 binding mode with all three anions. The two binding modes are likely described by (1) binding of the oxoanion to the guanidine, carbazole, and amine hydrogens, and (2) binding of the oxoanion to the hydrogens of the amine and protonated pyridine. Because I observed the binding of **GADAC** in its dicationic state, the protonated aminopyridine mimics an amidinium group and can also form a bidentate hydrogen bond to oxoanions. The lower  $pK_a$  of an aminopyridine than a guanidino group enhances this interaction. Interestingly, the binding of the amidinium group to acetate **2.12** is strong enough that the 1:1 and 1:2 binding modes cannot be teased apart. I do see a lower 1:2 binding with sulfate **2.14** and

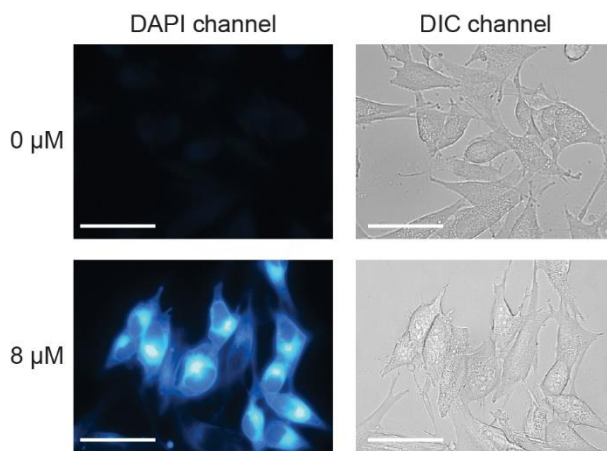
a much lower 1:2 binding with phosphate **2.13**, likely due to steric repulsion from the butyl groups preventing a second interaction.

Given the success of **GADAC** in binding to oxoanions in methanol, I next determined its ability to bind to oxoanions in water. I observed significant 1:1 binding of **GADAC** to all three anions, with **GADAC** having a binding affinity of  $3.6 \times 10^4 \text{ M}^{-1}$  to acetate **2.12** and bindings of  $1.1 \times 10^3 \text{ M}^{-1}$  and  $4.2 \times 10^3 \text{ M}^{-1}$  to phosphate **2.13** and sulfate **2.14**, respectively. The strong binding of **GADAC** to oxoanions in water is promising for its ability to bind to cell-surface anions in biological contexts.

Finally, I sought to compare **GADAC** against **GCP**. In Schmuck's initial paper introducing **GCP**, he determined the binding affinity of **GCP** to acetate (counterion not mentioned) in 40% water in DMSO to be  $2.8 \times 10^3 \text{ M}^{-1}$ .<sup>94</sup> I determined the binding of **GADAC** to acetate **2.12** in 40% water in DMSO to be  $(6.0 \pm 0.5) \times 10^4 \text{ M}^{-1}$ , a value approximately twenty times better than **GCP**. In addition to **GADAC**'s increased binding affinity, it also benefits from having a greater hydrophobicity than **GCP**. I determined the cLog *P* value and topological polar surface area (TPSA) of dicationic **GADAC** and a cationic, primary amide version of **GCP** for equal comparison of the structures through molinspiration.com (Figure 2.19).<sup>115</sup> I determined that **GCP** has a cLog *P* of  $-4.47$  and a TPSA of 140, whereas **GADAC** has a cLog *P* of  $-3.01$  and a TPSA of 132. Thus, even as a dication, **GADAC** is more lipophilic than **GCP**, which might improve its ability to diffuse through the phospholipid bilayer. Given its higher binding affinity to oxoanions and greater lipophilicity, I expect **GADAC** to be better than **GCP** at cell penetration.

**Epifluorescent microscopy of GADAC in live mammalian cells.** After determining that **GADAC** is fluorescent and binds strongly to oxoanions in water, we

tested the cell permeability of **GADAC** via epifluorescent microscopy. Specifically, we hoped to see that the fluorescence of **GADAC** could be detected inside cells with the commonly used DAPI filter. We incubated M21 melanoma cells in a range of concentrations of **GADAC** in serum-supplemented media for 1 h at 37 °C and then imaged the treated cells (Figure 2.5 and 2.20).



**Figure 2.5.** Images of the uptake of **GADAC** into live M21 cells after a 1-h incubation at 37 °C in serum-supplemented medium. DAPI channel (left) with 357/44 nm excitation and 447/60 nm emission; DIC channel (right). The DAPI channel epifluorescent images are normalized (taken at the same light intensity, exposure, and gain parameters). Scale bars, 50  $\mu\text{m}$ .

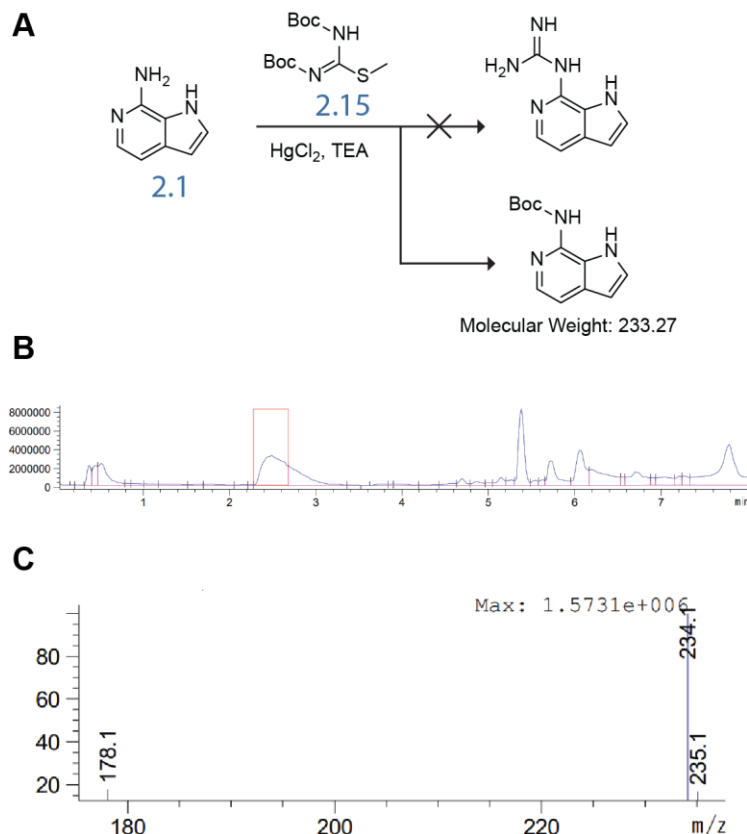
We observed that the fluorescence signal of **GADAC** was diffusely distributed in the cytosol and in the nucleus, indicating that the compound efficiently enters human cells even in the presence of serum. Notably, the fluorescence signal was particularly intense in the perinuclear compartment. We observed that **GADAC** enters cells and is detectable at concentrations as low as 1  $\mu\text{M}$  (Figure 2.20), which provides promise for biological applications.

## Conclusions

We have synthesized **GADAC**, a fluorescent guanidinium-containing diazacarbazole rationally designed to have a high affinity to cell-surface oxyanions. **GADAC** is capable of binding strongly to model oxoanions in water and can efficiently enter live mammalian cells, as observed through epifluorescent microscopy. I envision that **GADAC** could be used as a unique two-in-one system for transporting large molecules such as proteins into cells and for visualizing the cellular localization of said proteins without the need to append an additional fluorophore. In addition, its oxoanion binding and pH-dependent fluorescence intensity make **GADAC** a useful fluorescent reporter for non-biological applications.

## Unsuccessful Routes

### Carbamate transfer during guanidinylation with thiourea



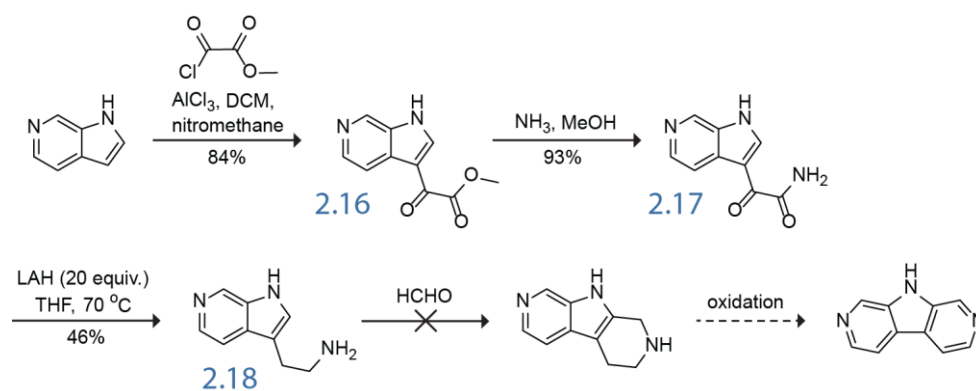
**Figure 2.6.** The guanidinylation of amino-azaindole **2.1** produces a product with a mass corresponding to a Boc addition. A. Reaction scheme showing the conditions and major undesired product. B. LC-MS chromatogram. C. Mass observed at the peak highlighted in the LC-MS chromatogram.

We attempted to convert the amino group into a guanidine through a mercury-aided reaction with a thiourea. Attempts at this reaction resulted in very low product yield and multiple byproducts as observed by LC-MS. The main byproduct of which corresponding to the addition of the Boc protecting group onto the azaindole (Figure 2.6). This is a surprising result given that many research groups have been successful at producing a guanidine through this method using a 2-aminopyridine.<sup>96,116</sup> I have also observed the di-Boc protected guanidine as the major product by LC-MS when reacting

2-aminopyridine with isothiurea (**2.15**). I hypothesize that the indole N-H is able to hydrogen bond with the carbamate carbonyl oxygen, increasing its electrophilicity and leading to the amine attacking the carbonyl through a 7-endo-trig cyclization.

Subsequently, other methods were explored to append the guanidine. Another common route is the reaction of the amine with an isothiocyanate to produce a thiourea. The most commonly used isothiocyanates are protected with carbamate groups, typically Cbz or Fmoc. I attempted reacting Cbz-NCS, Fmoc-NCS, as well as Bz-NCS with aminoazaindole (**2.1**), however I also observed addition of the carbamate or carbonyl group onto the amine. Bn-NCS was successful, however I was unable to remove the benzyl protecting group from the ensuing guanidine.

### Pictet-Spengler route to 2,7-diazacarbazole



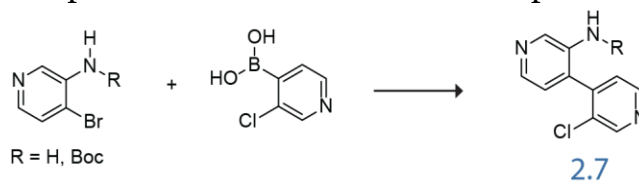
**Scheme 2.6.** Synthetic route toward 2,7-diazacarbazole using the Pictet-Spengler method.

To synthesize the 2,7-diazacarbazole, I sought to use the traditional Pictet-Spengler route, which has been used in the synthesis of azacarbazoles (Scheme 2.6).<sup>117</sup> This route is biomimetic and would traditionally involve the cyclization of tryptophan. In this case I used the 6-azaindole and reacted it with methyl oxalyl chloride to produce methyloxoacetate-azaindole (**2.16**), followed by exposure to ammonia to produce the

oxoacetamide-azaindole (**2.17**). I was able to reduce **2.17** to the alkyl chain forming azatryptamine (**2.18**), though in modest yields. Unfortunately, I was unsuccessful at closing the third ring upon exposure to formaldehyde in acidic or basic conditions. While I saw formation of the hemiaminal and the imine by LC-MS, the electronics of this ring structure are such that the 1-position is simply not nucleophilic enough to attack the imine.

### Attempted Suzuki conditions

**Table 2.3.** Conditions explored for the Suzuki reaction to produce bipyridine **2.7**.



<i>Catalysts</i>	Pd(PPh <sub>3</sub> ) <sub>2</sub> Cl <sub>2</sub>	Xphos Pd G3
	Pd(amphos)Cl <sub>2</sub>	Sphos Pd G3
	Pd(dppf)Cl <sub>2</sub>	cataCXium Pd G3
	Pd(DTBPF)Cl <sub>2</sub>	phosphaadamantane Pd G3
<i>Bases</i>	Na <sub>2</sub> CO <sub>3</sub>	
	K <sub>3</sub> PO <sub>4</sub>	
<i>Solvents</i>	Dimethoxyethane	
	Dioxane	
	Dimethylformamide	
<i>Temperatures</i>	70°	
	100°	
	150°	

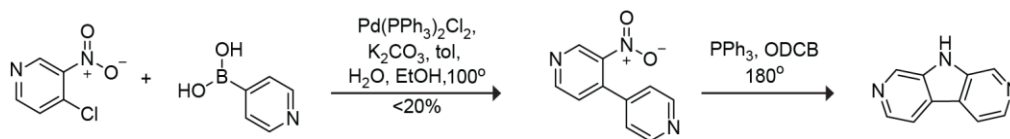
A range of Suzuki conditions were attempted to produce bipyridine **2.7** (Table 2.3). As the reactions were attempted *ad hoc* and not as part of a conditions screen, the various conditions attempted are displayed together in the above table for simplicity. Reactions with the unprotected bromopyridine resulted in yields at or below 20%, whereas the Boc-protected bromopyridine resulted in yields of up to 31%. In addition, while the



unprotected bromidopyridine is cheaper, this starting material suffered from worse purifications, and was found to be very unstable, likely undergoing  $S_NAr$  with itself.

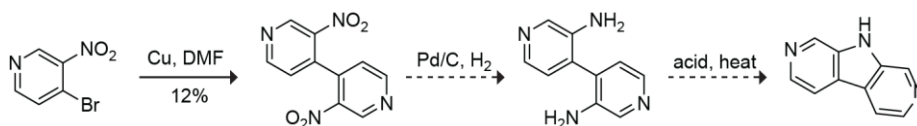
### Other metal-catalyzed coupling routes to 2,7-diazacarbazole **2.8**

Freeman and coworkers developed a method to produce carbazoles from 3-nitrobiphenyl compounds inspired by the Cadogen reaction.<sup>118</sup> Triphenyl phosphine is introduced under high heat to induce reductive deoxygenation of the nitro group, converting it to a nitrene. This nitrene undergoes a cyclization and subsequent hydride shift of the 3' hydrogen to produce a carbazole. I attempted to apply this reaction to produce 2,7-diazacarbazole **2.8** from a 3-nitrobipyridine (Scheme 2.7). While there is patent literature on this synthesis,<sup>119</sup> in my hands the Cadogen cyclization was messy and yielded minimal product along with the primary amine. In addition, a Suzuki reaction was necessary to produce the 3-nitrobipyridine, which was also low yielding.



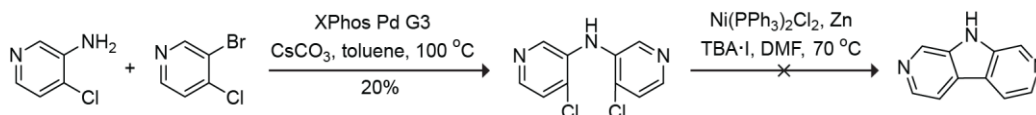
**Scheme 2.7.** Cadogen cyclization route to **2.8**.

I envisaged a route beginning with a copper-catalyzed Ullmann homocoupling to produce a 3,3'-dinitrobipyridine. The nitro groups could be reduced to amines, and the compound exposed to acid and heat to produce 2,7-diazacarbazole **2.8** through the Täuber carbazole synthesis (Scheme 2.8).<sup>120</sup> Unfortunately, the initial copper-catalyzed homocoupling was only minimally successful, even with freshly activated copper.



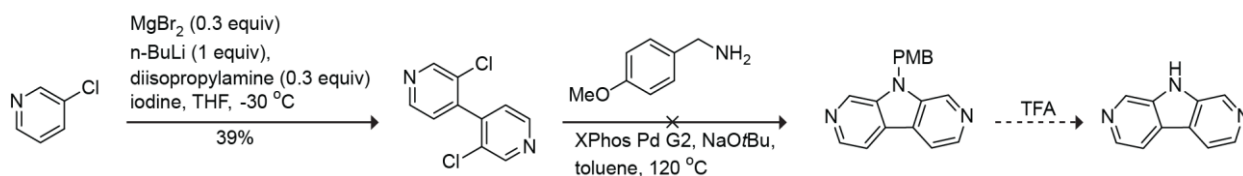
**Scheme 2.8.** Ullmann coupling and Täufer synthesis route to **2.8**.

I imagined flipping the route order to start with a Buchwald–Hartwig coupling to form the C–N bond and close the pyrrole ring via a nickel-mediated Ullmann coupling (Scheme 2.9).<sup>121</sup> The Buchwald–Hartwig coupling produced minimal product, and the Ullmann coupling was unsuccessful.



**Scheme 2.9.** Buchwald-Hartwig and Ullmann coupling route to **2.8**.

Mamane<sup>122</sup> and coworkers and Gong and coworkers<sup>104</sup> have synthesized substituted 2,7-diazacarbazoles from 3,3'-halobipyridines by performing a double Buchwald–Hartwig coupling with a substituted aniline. I synthesized a 3,3'-dichlorobipyridine according to Awad *et al.*,<sup>123</sup> but the subsequent double arylation with an aminobenzyl-PMB was unsuccessful. While Mamane and Gong have successfully synthesized 2,7-diazacarbazoles using this method, and the reaction likely could have been optimized, the deprotection of the PMB group adds an additional step to the route which is disadvantageous to the total yield.



**Scheme 2.10.** Double Buchwald–Hartwig coupling route to **2.8**.

Additional attempts to synthesize substituted 2,7-diazacarbazoles on route to a diguanidino analog of **GADAC** have been described by Christine Bradford.<sup>124</sup>

## Materials and Methods

### General

Materials. Commercial compounds were from Sigma–Aldrich (St. Louis, MO) and Combi-Blocks (San Diego, CA) and were used without further purification. **Note:** Sodium amide must be fresh and pure white to be fully active.

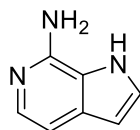
Conditions. All procedures were performed in air at ambient temperature (~22 °C) and pressure (1.0 atm) unless specified otherwise.

Solvent removal. The phrase “concentrated under reduced pressure” refers to the removal of solvents and other volatile materials using a rotary evaporator while maintaining a water-bath temperature at 40 °C. Residual solvent was removed from samples at high vacuum (<0.1 Torr), which refers to the vacuum achieved by a mechanical belt-drive oil pump, or through lyophilization (freeze-drying) using a Labconco FreeZone.

Chromatography. Chemical reactions were monitored by thin-layer chromatography (TLC) using EMD 250 µm silica gel 60-F<sub>254</sub> plates and visualization with UV-illumination or KMnO<sub>4</sub>-staining, or by LC–MS on an ESI Agilent 6125B mass spectrometer. Flash chromatography was performed with a Biotage Isolera automated purification system using prepacked and re-packed SNAP KP silica gel columns or SNAP KP C18 columns.

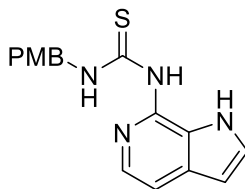
Instrumentation. <sup>1</sup>H-NMR and <sup>13</sup>C-NMR spectra for compound characterization were obtained with Bruker spectrometers, and ESI HRMS data were obtained with an Agilent 6545 Q-ToF mass spectrometer at the Department of Chemistry Instrumentation Facility at the Massachusetts Institute of Technology. Fluorescence spectra were obtained on a system from Photon Technology International running Felix fluorescence analysis software version 1.3. UV–Vis spectra were obtained with an Agilent Cary 60 UV–Vis spectrometer.

## Synthesis



**1H-pyrrolo[2,3-c]pyridin-7-amine (2.1).** 6-azaindole (590 mg, 5 mmol) was placed in a pressure-rated vial and dimethylaniline (20 mL, dried over mol. sieves) was added. The mixture was sonicated until fully dissolved. Sodium amide (780 mg, 20 mmol) was added, the vial was capped, and the solution stirred at 160 °C for five hours. The solution was cooled and vacuum filtered, and the solids rinsed with hexanes to remove the dimethylaniline. Water was then carefully added to the solid to quench the sodium amide. The aqueous mixture was then extracted with ethyl acetate three times, the organics combined and dried with sodium sulfate, filtered, and condensed under reduced pressure. The product was purified by silica chromatography in a gradient of DCM and 1% TEA in MeOH, eluting at 20% MeOH to produce a tan solid (442 mg, 66% yield).

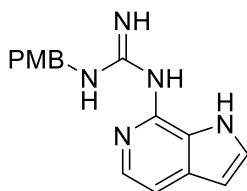
<sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>, δ): 10.99 (s, 1H), 7.44 (d, *J* = 5.7 Hz, 1H), 7.38 (d, *J* = 2.9 Hz, 1H), 6.76 (d, *J* = 5.7 Hz, 1H), 6.31 (d, *J* = 2.9 Hz, 1H), 5.90 (s, 2H). <sup>13</sup>C NMR (400 MHz, DMSO-*d*<sub>6</sub>, δ): 146.41, 135.89, 131.47, 126.16, 119.95, 105.53, 101.15. HRMS *m/z* calcd for C<sub>7</sub>H<sub>8</sub>N<sub>3</sub> [M+H]<sup>+</sup>, 134.0718; found 134.0711.



**1-(4-methoxybenzyl)-3-(1H-pyrrolo[2,3-c]pyridin-7-yl)thiourea (2.2).** 7-amino-6-azaindole **2.1** (442 mg, 3.3 mmol) was dissolved in dry THF (30 mL).

Triethylamine (1.37 mL, 9.9 mmol) and 4-methoxybenzyl isothiocyanate (656 mg, 3.66 mmol) were added and the solution was placed under a condenser connected to an N<sub>2</sub> line and refluxed overnight. The solution was condensed under reduced pressure. Ice cold isopropanol was added to precipitate the product, and the solids were vacuum filtered and rinsed with cold isopropanol to obtain a grey powder. The filtrate was left to evaporate in air and the aforementioned process was repeated to obtain further product (625 mg, 61% yield).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>, δ): 13.48 (t (br), *J* = 5.7 Hz, 1H), 11.12 (s, 1H), 10.68 (s, 1H), 7.68 (d, *J* = 5.7 Hz, 1H), 7.33 (d, *J* = 8.6 Hz, 2H), 7.23 (d, *J* = 5.7 Hz, 1H), 7.10 (t (br), *J* = 2.7 Hz, 1H), 6.87 (d, *J* = 8.7 Hz, 2H), 6.47 (br, *J* = 3.1, 1.5 Hz, 1H), 5.10 (d, *J* = 5.6 Hz, 2H), 3.81 (s, 3H). <sup>13</sup>C NMR (400 MHz, CDCl<sub>3</sub>, δ): 179.42, 159.14, 140.58, 134.68, 133.65, 129.61, 128.68, 128.17, 120.16, 114.31, 111.94, 102.46, 55.49, 48.76. HRMS *m/z* calcd for C<sub>16</sub>H<sub>17</sub>N<sub>4</sub>OS [M+H]<sup>+</sup>, 313.1123; found 313.1112.

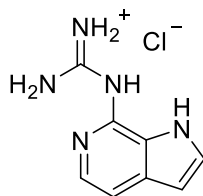


**1-(4-methoxybenzyl)-3-(1H-pyrrolo[2,3-c]pyridin-7-yl)guanidine (3).**

Thiourea **2.2** (194 mg, 0.62 mmol) was placed in a flask and dissolved in methanol (30 mL). Concentrated ammonium hydroxide (6 mL) and 5 M *tert*-butyl hydroperoxide in decane (2.48 mL, 12.4 mmol) were added and the solution was stirred at room temperature overnight. The mixture was subsequently condensed under reduced pressure, and resuspended in water. The solids were separated by vacuum filtration,

rinsing with water and ethyl acetate. The product was obtained as an orange solid. (161 mg, 87% yield).

**<sup>1</sup>H NMR** (500 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 12.23 (s, 2H), 10.99 (s, 1H), 7.75 (d,  $J$  = 5.7 Hz, 1H), 7.55 (t,  $J$  = 2.7 Hz, 1H), 7.34 (d,  $J$  = 5.8 Hz, 1H), 7.30 (d,  $J$  = 8.7 Hz, 2H), 6.86 (d,  $J$  = 8.7 Hz, 2H), 6.55 (dd,  $J$  = 2.9, 1.7 Hz, 1H), 4.51 (d,  $J$  = 5.8 Hz, 2H), 3.70 (s, 3H). **<sup>13</sup>C NMR** (500 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 158.67, 154.41, 139.13, 133.92, 133.46, 129.48, 128.57, 128.48, 121.16, 113.95, 111.62, 101.94, 55.03, 43.59. HRMS  $m/z$  calcd for C<sub>16</sub>H<sub>18</sub>N<sub>5</sub>O [M+H]<sup>+</sup>, 296.1511; found 296.1502.

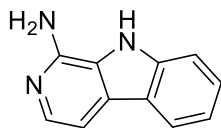


**1-(1H-pyrrolo[2,3-c]pyridin-7-yl)guanidine hydrochloride (GAI).** PMB-guanidine **2.3** (100 mg, 0.34 mmol) was placed in a pressure-rated vial. Trifluoroacetic acid (3 mL) and methanol (100  $\mu$  L) were added and the vial was sealed and stirred at 80 °C overnight. The solution was condensed under a stream of nitrogen, then dissolved in water and ethyl acetate. The mixture was passed through a syringe filter, and the organic layer was extracted with water two times. The aqueous fractions were condensed under reduced pressure, and then purified with reversed phase chromatography in acetonitrile and water with 0.1% TFA. The resulting product was dissolved in water and excess HCl was added, and the solution was lyophilized to produce a cream solid (17.4 mg, 29% yield).

**<sup>1</sup>H NMR** (500 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 12.23 (s, 1H), 8.40 (d,  $J$  = 135.9 Hz, 3H), 7.73 (d,  $J$  = 5.6 Hz, 1H), 7.59 (d,  $J$  = 2.9 Hz, 1H), 7.24 (d,  $J$  = 5.6 Hz, 1H), 6.51 (d,  $J$  = 2.9 Hz, 1H). **<sup>13</sup>C**

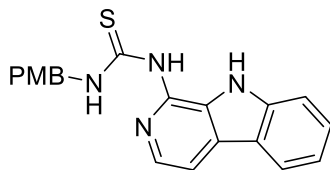
**NMR** (500 MHz, D<sub>2</sub>O,  $\delta$ ): 158.68, 141.41, 138.39, 135.73, 133.04, 127.70, 117.83, 106.81.

**HRMS**  $m/z$  calcd for C<sub>8</sub>H<sub>10</sub>N<sub>5</sub> [M+H]<sup>+</sup>, 176.0936; found 176.0938.



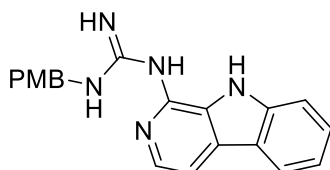
**9H-pyrido[3,4-b]indol-1-amine (2.4).**  $\beta$ -carboline (500 mg, 3 mmol) was placed in a pressure-rated vial and dimethylaniline (12 mL, dried over mol. sieves) was added. The mixture was heated until fully dissolved. Sodium amide (468 mg, 12 mmol) was added, the vial was capped, and the solution stirred at 180 °C for five hours. The solution was cooled and vacuum filtered, and the solids rinsed with hexanes to remove the dimethylaniline. Water was then carefully added to the solid to quench the sodium amide. The aqueous mixture was then extracted with ethyl acetate three times, the organics combined and dried with sodium sulfate, filtered, and condensed under reduced pressure. The product was purified by silica chromatography in a gradient of DCM and 1 M NH<sub>3</sub> in MeOH, eluting at 12% MeOH to produce a tan solid (377 mg, 69% yield).

**<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 11.01 (s, 1H), 8.06 (d,  $J$  = 7.8 Hz, 1H), 7.71 (d,  $J$  = 5.5 Hz, 1H), 7.58 (d,  $J$  = 8.2 Hz, 1H), 7.45 (t,  $J$  = 7.1 Hz, 1H), 7.29 (d,  $J$  = 5.5 Hz, 1H), 7.17 (t,  $J$  = 7.0 Hz, 1H), 6.06 (s, 2H). **<sup>13</sup>C NMR** (400 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 147.13, 139.49, 136.96, 127.13, 126.87, 123.29, 122.17, 121.66, 119.36, 112.38, 105.28. **HRMS**  $m/z$  calcd for C<sub>11</sub>H<sub>10</sub>N<sub>3</sub> [M+H]<sup>+</sup>, 184.0875; found 184.0866.



**1-(4-methoxybenzyl)-3-(9H-pyrido[3,4-b]indol-1-yl)thiourea (2.5).** 7-amino-6-azacarbazole **2.4** (174 mg, 0.95 mmol) was dissolved in dry THF (9 mL). Triethylamine (340  $\mu$ L, 2.85 mmol) and 4-methoxybenzyl isothiocyanate (340 mg, 1.9 mmol) were added and the solution was placed in a pressure-rated vial and stirred at 80  $^{\circ}$ C overnight. The solution was condensed under reduced pressure. Ice cold ethanol was added to precipitate the product, and the solids were vacuum filtered and rinsed with cold ethanol to obtain a grey powder (190 mg, 55% yield).

$^1\text{H NMR}$  (400 MHz, DMSO- $d_6$ ,  $\delta$ ): 12.69 (t,  $J$  = 5.6 Hz, 1H), 11.94 (s, 1H), 10.46 (s, 1H), 8.22 (d,  $J$  = 7.9 Hz, 1H), 7.92 (d,  $J$  = 5.5 Hz, 1H), 7.85 (d,  $J$  = 5.5 Hz, 1H), 7.68 (d,  $J$  = 8.3 Hz, 1H), 7.57 (t,  $J$  = 7.7 Hz, 1H), 7.37 (d,  $J$  = 8.1 Hz, 2H), 7.28 (t,  $J$  = 7.5 Hz, 1H), 6.94 (d,  $J$  = 8.2 Hz, 2H), 4.90 (d,  $J$  = 5.5 Hz, 2H), 3.74 (s, 3H).  $^{13}\text{C NMR}$  (400 MHz, DMSO- $d_6$ ,  $\delta$ ): 180.40, 158.94, 140.58, 140.13, 134.22, 130.56, 129.48, 129.26, 128.65, 123.12, 122.07, 121.30, 120.36, 114.42, 112.82, 110.95, 55.55, 47.98. **HRMS**  $m/z$  calcd for  $\text{C}_{20}\text{H}_{19}\text{N}_4\text{OS}$   $[\text{M}+\text{H}]^+$ , 363.1280; found 363.1262.

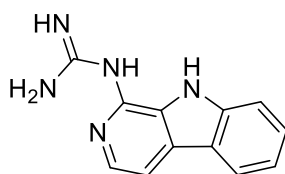


**1-(4-methoxybenzyl)-3-(9H-pyrido[3,4-b]indol-1-yl)guanidine (2.6).** Thiourea **2.5** (108 mg, 0.3 mmol) was placed in a flask and suspended in methanol (15 mL). Concentrated ammonium hydroxide (3 mL) and 5 M *tert*-butyl hydroperoxide in decane



(1.2 mL, 6 mmol) were added and the suspension was stirred at room temperature overnight. The mixture was subsequently condensed under reduced pressure and resuspended in water. The solids were separated by vacuum filtration, rinsing with water. The product was obtained as a grey solid that was used directly.

**HRMS**  $m/z$  calcd for  $C_{20}H_{20}N_5O$   $[M+H]^+$ , 346.1668; found 346.1661.

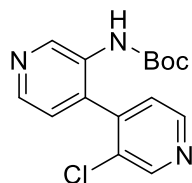


**1-(9H-pyrido[3,4-b]indol-1-yl)guanidine hydrochloride (GAC).** PMB-guanidine **2.6** (88 mg, 0.26 mmol) was placed in a pressure-rated vial. Trifluoroacetic acid (2.5 mL) and methanol (100  $\mu$ L) were added and the vial was sealed and stirred at 80 °C overnight. The solution was condensed under a stream of nitrogen, then dissolved in water and ethyl acetate. The aqueous solution was extracted with ethyl acetate three times, the organics combined with sodium sulfate, filtered, and condensed under reduced pressure. The crude product was purified by silica chromatography in a gradient of DCM and MeOH, eluting at 5% MeOH to produce a brown solid (23 mg, 40% yield). The product was converted to the HCl salt by suspending in DCM and adding one equivalent of 0.1 M HCl in MeOH at which point the compound fully dissolved. The product was condensed under reduced pressure to produce a pale pink solid.

**$^1H$  NMR** (500 MHz,  $DMSO-d_6$ ,  $\delta$ ): 11.58 (s, 1H), 11.47 (s, 1H), 8.35 (s, 3H), 8.26 (d,  $J$  = 7.9 Hz, 1H), 8.08 (d,  $J$  = 5.5 Hz, 1H), 7.98 (d,  $J$  = 5.4 Hz, 1H), 7.69 (d,  $J$  = 8.2 Hz, 1H), 7.60 (t,  $J$  = 7.6 Hz, 1H), 7.31 (t,  $J$  = 7.5 Hz, 1H).  **$^{13}C$  NMR** (600 MHz,  $DMSO-d_6$ ,  $\delta$ ): 155.95,

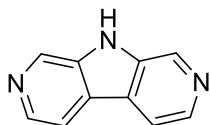
140.41, 138.42(br), 135.24, 130.08, 129.07, 125.36, 122.29, 121.31, 120.57, 112.80, 112.17.

**HRMS**  $m/z$  calcd for  $C_{12}H_{12}N_5$   $[M+H]^+$ , 226.1093; found 226.1079.



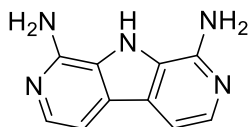
**tert-butyl (3'-chloro-[4,4'-bipyridin]-3-yl)carbamate (2.7).** N-(Boc)-3-amino-4-bromopyridine (1 g, 3.66 mmol), 3-chloropyridine-4-boronic acid (1150 mg, 7.33 mmol), and Pd(dppf)Cl<sub>2</sub> (270 mg, 0.37 mmol) were evenly divided into two pressure-rates vials. Dioxane (18 mL) and 2 M K<sub>2</sub>CO<sub>3</sub> (2.25 mL) were equally divided and added to the two vials. The vials were capped and the solutions stirred at 70 °C for 1.5-2 hours. The vials were removed from heat, combined, and water and ethyl acetate were added. The aqueous fraction was extracted with ethyl acetate four times, dried with sodium sulfate, filtered, and condensed under reduced pressure. The crude product was purified with silica chromatography in a gradient of ethyl acetate and hexanes, eluting at 100% ethyl acetate to produce a cream-color solid (366 mg, 33% yield). The N-(Boc)-3-amino-4-bromopyridine was also recovered as a white solid (446 mg, 68% recovered from theoretical remaining amount).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>,  $\delta$ ): 9.13 (s, 1H), 8.69 (s, 1H), 8.57 (d, J = 4.9 Hz, 1H), 8.42 (d, J = 4.9 Hz, 1H), 7.25 (d, J = 4.8 Hz, 1H), 7.10 (d, J = 4.9 Hz, 1H), 6.56 (s, 1H), 1.42 (s, 9H). **<sup>13</sup>C NMR** (400 MHz, CDCl<sub>3</sub>,  $\delta$ ): 152.60, 150.24, 148.17, 144.97, 142.81, 135.29, 131.87, 130.69, 125.12, 123.49, 81.46, 28.13. **HRMS**  $m/z$  calcd for  $C_{15}H_{17}ClN_3O_2$   $[M+H]^+$ , 306.1009; found 306.1000.



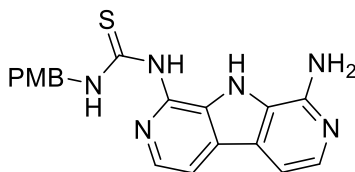
**9H-pyrrolo[2,3-c:5,4-c']dipyridine (2.8).** Bipyridine **2.7** (1.15 g, 3.77 mmol) was added to a flask and trifluoroacetic acid (5 mL) and methanol (100  $\mu$ L) were added. The orange solution was stirred for 30 minutes, after which time the solution was dried under a stream of nitrogen. The product was dissolved in ethyl acetate and 1 M NaOH was added. The aqueous fraction was extracted with ethyl acetate four times, dried with sodium sulfate, filtered, and condensed under reduced pressure. The deprotected bipyridine was then placed in three pressure-rated vials. XPhos Pd G2 (255 mg, 0.32 mmol) and sodium *tert*-butoxide (1.09 g, 11.3 mmol) was split evenly and added to the three vials, followed by THF (37 mL), split evenly. The vials were flushed with nitrogen, capped, and stirred at 70  $^{\circ}$ C overnight. The vials were pooled and methanol was added to fully dissolve the mixture, then silica was added and the mixture condensed under reduced pressure. The crude product was purified by silica chromatography in dichloromethane and methanol, eluting at 15% methanol to produce a cream colored solid (552 mg, 87% yield).

**$^1\text{H NMR}$**  (500 MHz, DMSO- $d_6$ ,  $\delta$ ): 11.98 (s, 1H), 9.06 (d,  $J = 1.2$  Hz, 2H), 8.44 (d,  $J = 5.3$  Hz, 2H), 8.24 (dd,  $J = 5.3, 1.2$  Hz, 2H).  **$^{13}\text{C NMR}$**  (500 MHz, DMSO- $d_6$ ,  $\delta$ ): 138.99, 136.88, 136.14, 126.18, 116.38. **HRMS**  $m/z$  calcd for  $\text{C}_{10}\text{H}_8\text{N}_3$   $[\text{M}+\text{H}]^+$ , 170.0718; found 170.0708.



**9H-pyrrolo[2,3-c:5,4-c']dipyridine-1,8-diamine (2.9).** Diazacarbazole **2.8** (338 mg, 2 mmol) was placed in a pressure-rated vial and dimethylaniline (10 mL, dried over mol. sieves) was added. The solution was stirred at 180 °C to dissolve **2.8**. Once dissolved, the solution was slightly cooled and sodium amide (624 mg, 16 mmol) was added. The vial was sealed and stirred at 180 °C for five hours. The solution was cooled and vacuum filtered, and the solids rinsed with hexanes to remove the dimethylaniline. Water was then carefully added to the solid to quench the sodium amide. The aqueous mixture was then extracted with ethyl acetate four times, the organics combined and dried with sodium sulfate, filtered, and condensed under reduced pressure. The product was purified by silica chromatography in a gradient of DCM and 1 M NH<sub>3</sub> in MeOH, eluting at 20% MeOH to produce a yellow solid (274 mg, 69% yield). The product could also be purified by reversed phase C18 chromatography in water and acetonitrile to avoid elution of silica gel from the highly polar solvent phase.

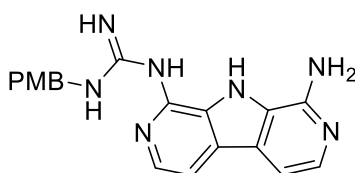
**<sup>1</sup>H NMR** (500 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 11.69 (s, 1H), 7.68 (d, *J* = 5.7 Hz, 2H), 7.30 (d, *J* = 5.7 Hz, 2H), 6.37 (s, 4H). **<sup>13</sup>C NMR** (600 MHz, MeOD,  $\delta$ ): 147.24, 133.59, 125.88, 125.05, 105.38. **HRMS** *m/z* calcd for C<sub>10</sub>H<sub>10</sub>N<sub>5</sub> [M+H]<sup>+</sup>, 200.09362; found 200.0927



**1-(8-amino-9H-pyrrolo[2,3-c:5,4-c']dipyridin-1-yl)-3-(4-methoxybenzyl)thiourea (2.10).** Diamine **2.9** (55 mg, 0.28 mmol) was placed in a

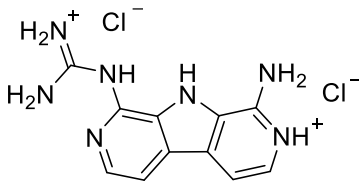
pressure-rated vial and dissolved in dry THF (11 mL) and dry DMF (3 mL). Triethylamine (115  $\mu$ L, 0.83 mmol) and *para*-methoxybenzyl isothiocyanate (500 mg, 2.8 mmol) were added. The headspace was flushed with nitrogen, and the vial was capped and stirred at 80  $^{\circ}$ C for five hours. The solution was removed from heat and concentrated under a stream of air overnight. Hexanes was added to the resulting orange oil and the mixture sonicated. The mixture was then vacuum filtered and rinsed with hexanes to produce 88.5 mg of orange powder which was used directly.

**HRMS**  $m/z$  calcd for  $C_{19}H_{19}N_6OS$   $[M+H]^+$ , 379.1341; found 379.1313.



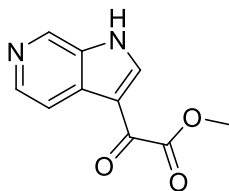
**1-(8-amino-9H-pyrrolo[2,3-c:5,4-c']dipyridin-1-yl)-3-(4-methoxybenzyl)guanidine (2.11)**. Half of the crude thiourea **2.10** (43 mg) was dissolved in dry DMF (11 mL) and cooled in an ice bath. Mercury(II) chloride (33 mg, 0.12 mmol) was added and the solution was stirred for five to ten minutes until cloudy. A 0.5 M solution of ammonium in dioxane (2.2 mL, 1.1 mmol) was added dropwise and the solution was stirred on ice for 15 minutes, then stirred at room temperature for one hour. The grey cloudy mixture was then filtered through celite and rinsed with ethyl acetate. The solution was condensed under a stream of nitrogen overnight and the resulting dark orange oil used directly.

**HRMS**  $m/z$  calcd for  $C_{19}H_{20}N_7O$   $[M+H]^+$ , 362.1729; found 362.1704.



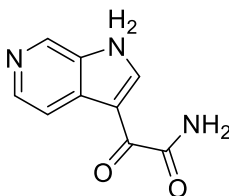
**1-(8-amino-9H-pyrrolo[2,3-c:5,4-c']dipyridin-1-yl)guanidine (GADAC).** PMB-guanidine **2.11** was dissolved in TFA (5 mL) and MeOH (100  $\mu$ L) in a pressure-rated vial and stirred at 80  $^{\circ}$ C for three hours. The solution was concentrated under a stream of nitrogen, then resuspended in ethyl acetate and 1 M NaOH. The aqueous layer was extracted with ethyl acetate five times, the organics combined, and concentrated under reduced pressure without drying. The solid was then resuspended in DI water with a few drops 1 M HCl and filtered through cotton. The aqueous solution was purified by reversed phase C18 chromatography in acetonitrile and 10 mM ammonium acetate in water, eluting at 30% acetonitrile. The resulting fractions were condensed and lyophilized to remove excess ammonium acetate. The product was then redissolved in DI water and a few drops of 1 M HCl were added and the product was lyophilized again to convert the product from the diacetate salt to the dichloride, resulting in a bright yellow solid (8.3 mg, 19% yield over three steps).

**$^1\text{H NMR}$**  (600 MHz,  $\text{D}_2\text{O}$ ,  $\delta$ ): 7.71 (d,  $J = 5.4$  Hz, 1H), 7.34 (d,  $J = 6.6$  Hz, 1H), 7.33 (d,  $J = 6.1$  Hz, 2H), 7.13 (d,  $J = 6.6$  Hz, 1H).  **$^{13}\text{C NMR}$**  (600 MHz,  $\text{D}_2\text{O}$ ,  $\delta$ ): 154.60, 143.73, 137.19, 135.95, 127.03, 126.47, 125.31, 124.48, 121.58, 112.19, 106.02. **HRMS**  $m/z$  calcd for  $\text{C}_{11}\text{H}_{12}\text{N}_7$   $[\text{M}+\text{H}]^+$ , 242.1154; found 242.1147.



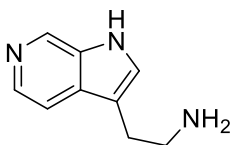
**methyl 2-oxo-2-(1H-pyrrolo[2,3-c]pyridin-3-yl)acetate (2.16):** 6-Azaindole (350 mg, 3 mmol) and aluminum chloride (2g, 15 mmol) were added to a dry flask. A mixture of DCM (11 mL) and nitromethane (4 mL) that had been dried over mol. sieves was added, followed by methyl chlorooxacetate (554  $\mu$ L, 6 mmol). The solution was stirred at room temperature under nitrogen for one hour. MeOH was added to quench the solution until the solution stopped bubbling. Sat. NaHCO<sub>3</sub> was added until the pH reached 7. The solution was then condensed under reduced pressure to remove the organic solvents. This aqueous solution was extracted four times with EtOAc, dried with sodium sulfate, filtered, and condensed under reduced pressure to produce a tan powder (513 mg, 84% yield).

**<sup>1</sup>H NMR** (500 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 12.84 (s, 1H), 8.91 (d, *J* = 1.2 Hz, 1H), 8.66 (s, 1H), 8.38 (d, *J* = 5.4 Hz, 1H), 8.06 (dd, *J* = 5.4, 1.2 Hz, 1H), 3.91 (s, 3H). **<sup>13</sup>C NMR** (500 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 179.05, 163.74, 142.21, 141.33, 136.02, 134.25, 131.16, 115.90, 112.39, 53.19. **HRMS** *m/z* calcd for C<sub>10</sub>H<sub>12</sub>N<sub>3</sub>O<sub>3</sub> [M+NH<sub>4</sub>]<sup>+</sup>, 222.0879; found 222.0875.



**2-oxo-2-(1H-pyrrolo[2,3-c]pyridin-3-yl)acetamide (2.17):** Methyloxoacetate-azaindole **2.16** (500 mg, 2.45 mmol) was dissolved in 7 N ammonia in methanol (15 mL) for 45 minutes. The solution was condensed under reduced pressure to produce a tan solid (430 mg, 93% yield).

**<sup>1</sup>H NMR** (500 MHz, DMSO-*d*<sub>6</sub>, δ): 12.60 (s, 1H), 8.88 (s, 1H), 8.86 (s, 1H), 8.36 (d, *J* = 5.4 Hz, 1H), 8.14 (s, 1H), 8.10 (dd, *J* = 5.4, 1.1 Hz, 1H), 7.80 (s, 1H). **<sup>13</sup>C NMR** (500 MHz, DMSO-*d*<sub>6</sub>, δ): 183.54, 165.89, 141.97, 141.10, 135.88, 133.85, 131.64, 115.99, 112.04. **HRMS** *m/z* calcd for C<sub>9</sub>H<sub>8</sub>N<sub>3</sub>O<sub>2</sub> [M+H]<sup>+</sup>, 190.0617; found 190.0609.

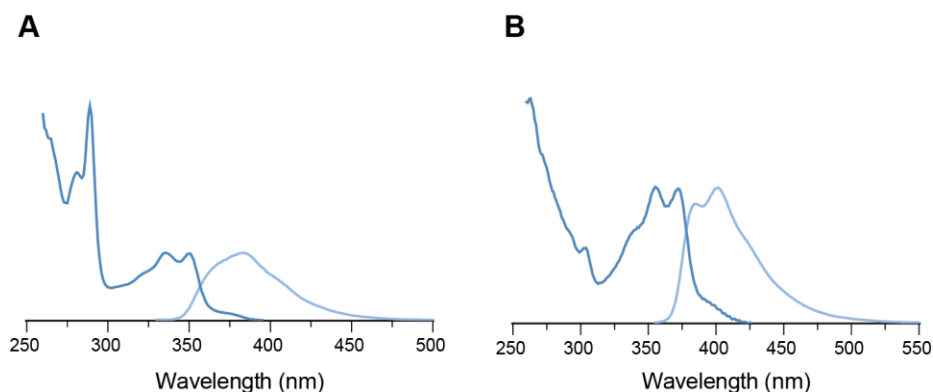


**2-(1H-pyrrolo[2,3-c]pyridin-3-yl)ethan-1-amine (2.18):** Oxoacetamide-azaindole **2.17** (150 mg, 0.79 mmol) was placed in a pressure-rated vial and dry THF (4 mL) added. The mixture was sonicated to suspend **2.17**, and lithium aluminum hydride (524 mg, 15.9 mmol) was added. The vial was flushed with nitrogen, capped, and stirred at 70 °C overnight. The solution was quenched with water, 2M NaOH, and more water, and then filtered. The filtrate was then purified via silica gel chromatography in a gradient of 7 N ammonia in MeOH and DCM to produce a yellow oil (58 mg, 46% yield).

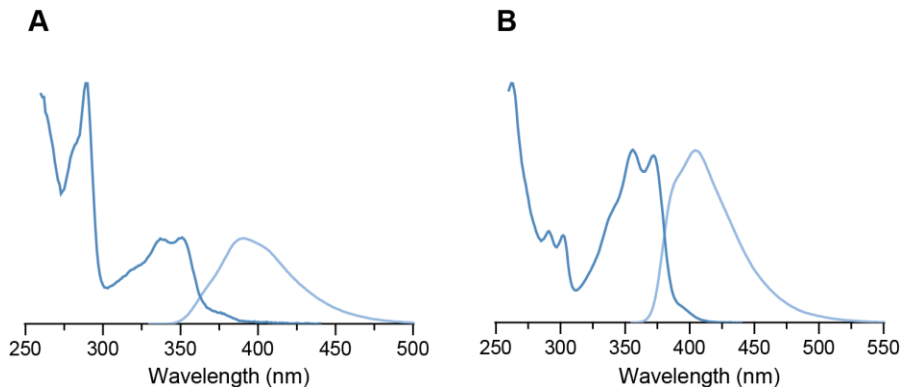
**<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>, δ): 11.31 (s, 1H), 8.68 (d, *J* = 1.1 Hz, 1H), 8.05 (d, *J* = 5.5 Hz, 1H), 7.50 (dd, *J* = 5.5, 1.1 Hz, 1H), 7.38 (s, 1H), 3.17 (s, 2H), 2.82 (dd, *J* = 7.9, 4.8 Hz, 2H), 2.76 (dd, *J* = 7.8, 4.8 Hz, 2H). **<sup>13</sup>C NMR** (400 MHz, DMSO-*d*<sub>6</sub>, δ): 137.09, 134.28, 133.48, 131.41, 126.84, 113.21, 112.27, 42.51, 28.75. **HRMS** *m/z* calcd for C<sub>9</sub>H<sub>12</sub>N<sub>3</sub> [M+H]<sup>+</sup>, 162.1031; found 162.1024.



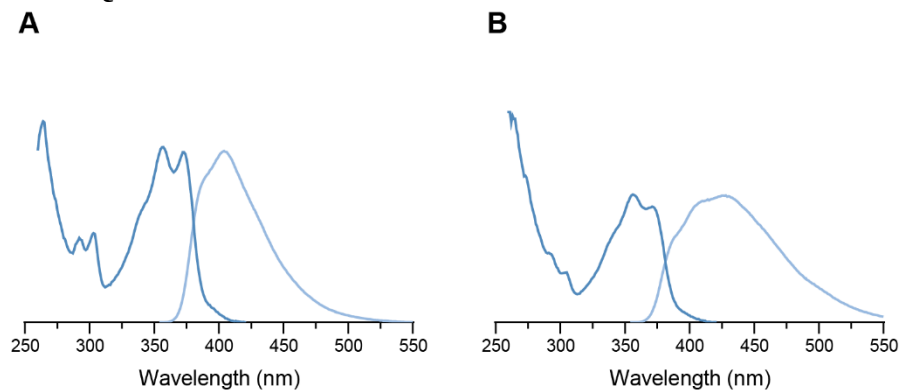
## Fluorescence Spectra



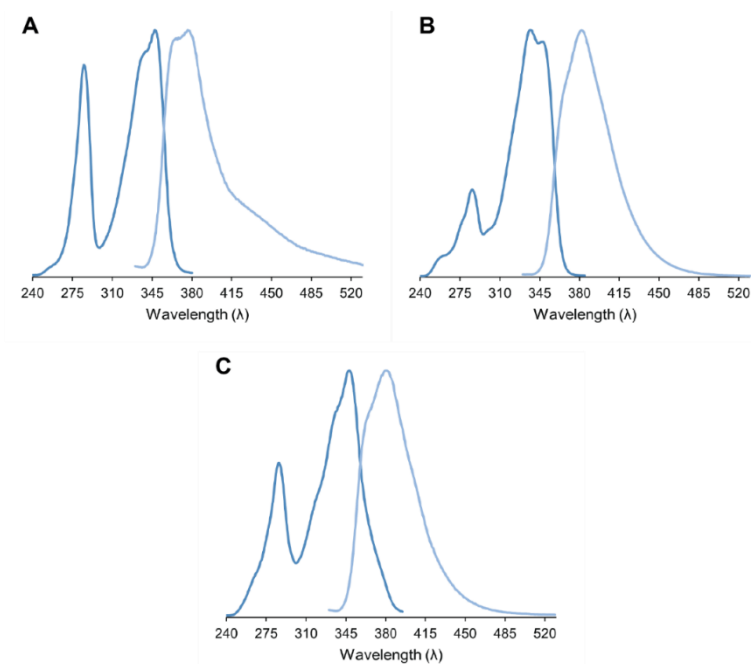
**Figure 2.7.** Absorbance/Emission spectra of the fluorescent guanidino-compounds performed in methanol. A. **GAC**. B. **GADAC**.



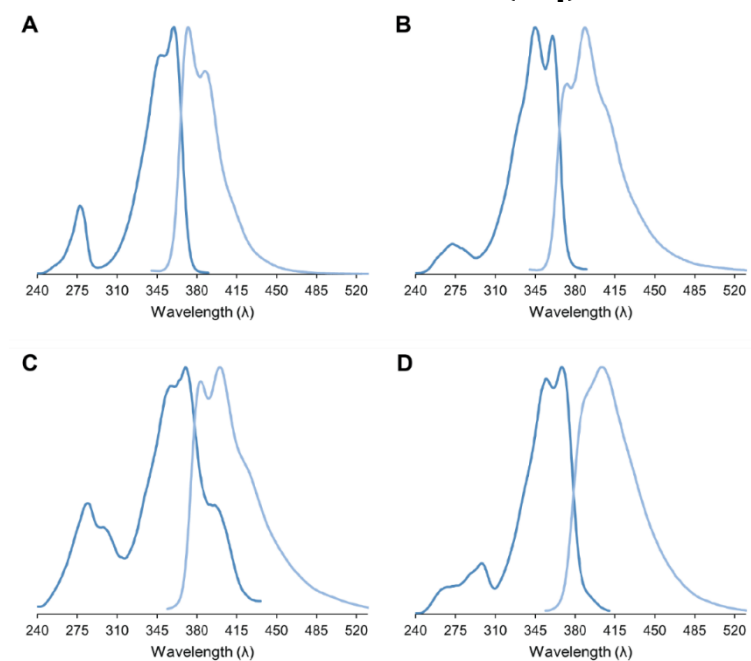
**Figure 2.8.** Absorbance/Emission spectra of the fluorescent guanidino-compounds performed in MilliQ water. A. **GAC**. B. **GADAC**.



**Figure 2.9.** Absorbance/Emission spectra of **GADAC** performed in buffer. A. 10 mM citrate buffer with 150 mM NaCl, pH 4. B. 10 mM phosphate buffer with 150 mM NaCl, pH 7.4.



**Figure 2.10.** Excitation/Emission spectra of the 2-azacarbazole backbone compounds in methanol. A.  $\beta$ -carboline. B. Amino-azacarbazole (**2.4**). C. **GAC**.



**Figure 2.11.** Excitation/Emission spectra of the 2,7-azacarbazole backbone compounds in methanol, unless otherwise stated. A. Diazacarbazole (**2.8**). B. Diamino-diazacarbazole (**2.9**). C. **GADAC**. D. **GADAC** in water.

## **pH Titrations**

### **Fluorescence pH Titration**

A 50  $\mu\text{L}$  aliquot of the stock **GAC** or **GADAC** isopropanol solutions was diluted in milli-Q water (9,950  $\mu\text{L}$ ) to produce a final concentration of **GAC** (4.42  $\mu\text{M}$ ), **GADAC** (0.685  $\mu\text{M}$ ). Sodium chloride (30 mg) was added to a cuvette and 3 mL of the aqueous guanidine solution was added. The excess chloride anions ensure that any change in fluorescence comes from changes in pH as opposed to additional chloride binding on addition or change in ionic strength of HCl. The solution was then adjusted to the desired pH's with aqueous NaOH or HCl. The emission spectra were acquired at each pH and the intensity of the selected wavelengths at each point determined and plotted in triplicate. The  $\text{p}K_{\text{a}}$ 's of each molecule were determined by GraphPad Prism 6.

### **Absorbance pH Titration**

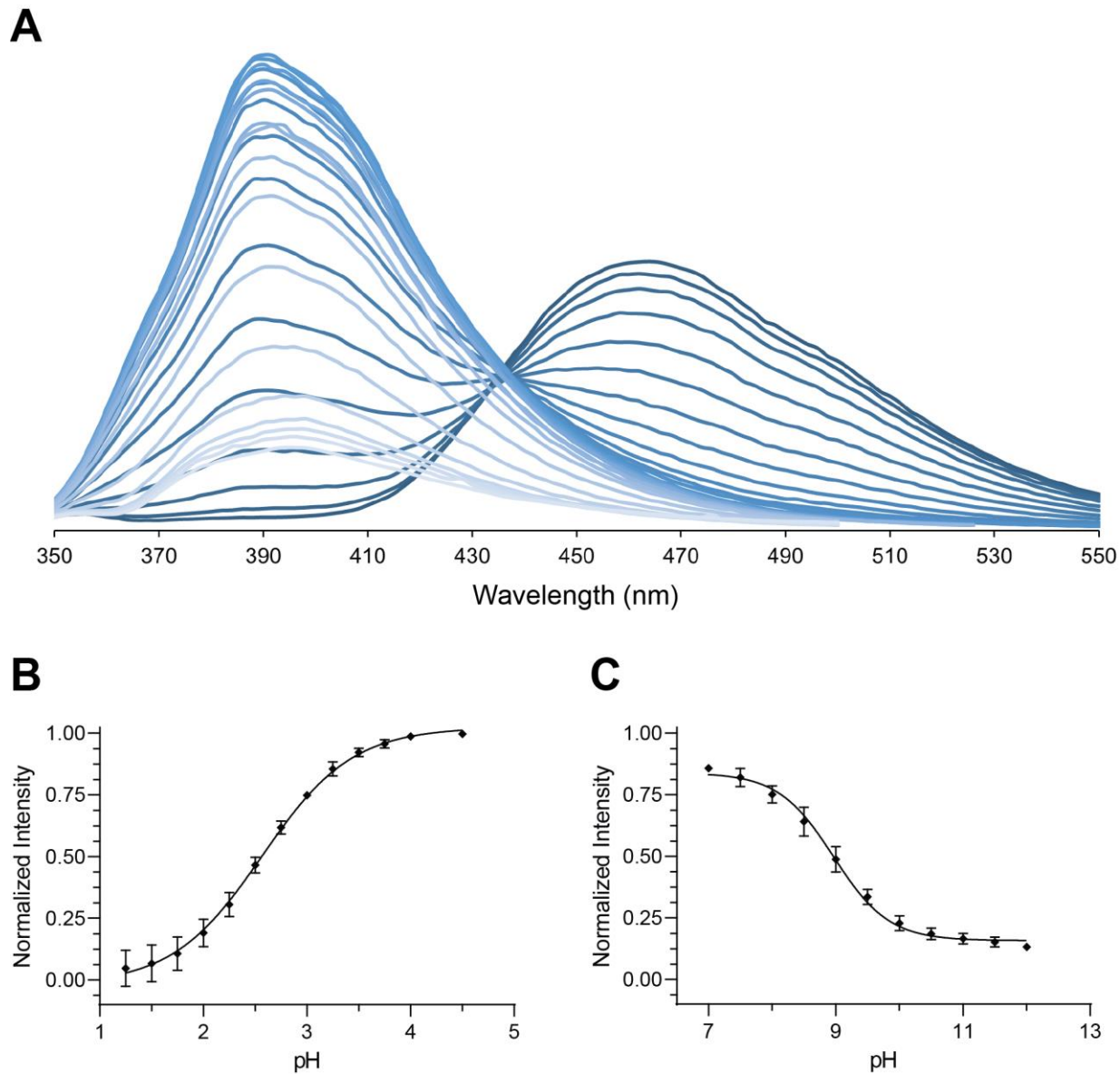
Sodium chloride (30 mg) was added to a cuvette and 2,940 mL MilliQ water and a 60  $\mu\text{L}$  aliquot of an aqueous stock of **GADAC** (0.51 mM) were added. The solution was then adjusted to the desired pH's with aqueous NaOH or HCl. The absorbance spectra were acquired at each pH and the absorbance of the selected wavelengths at each point determined and plotted in triplicate. The  $\text{p}K_{\text{a}}$ 's of each molecule were determined by GraphPad Prism 6.

### **$^1\text{H}$ NMR pH Titration**

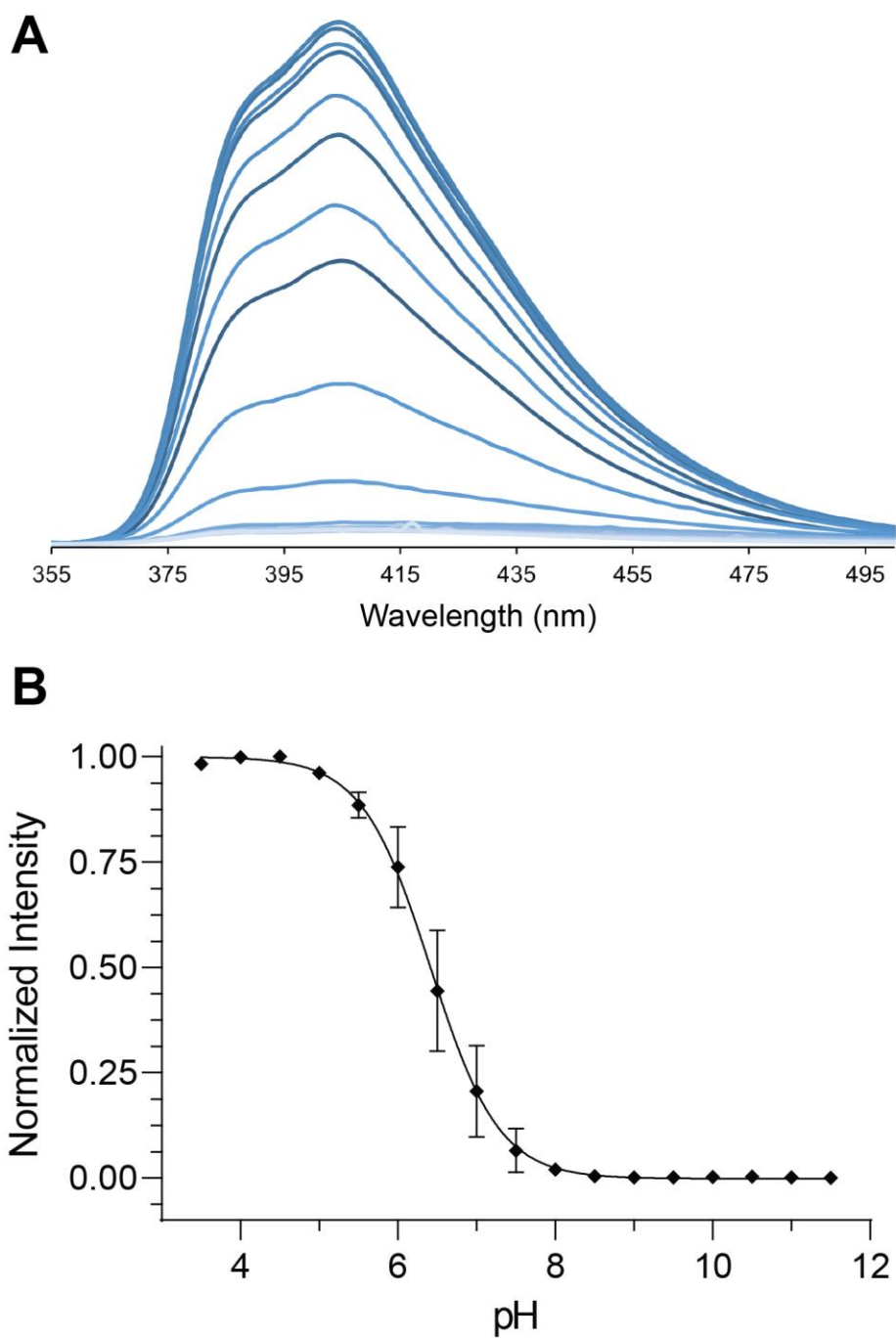
**GAI** (4.2 mg, 0.02 mmol) was dissolved in 15 mL  $\text{D}_2\text{O}$  in a vial with a stirbar. While stirring, a pH probe was used to measure the  $\text{pH}_{(\text{obs})}$  of the solution. The solution was

adjusted to the desired pH increments with NaOD and DCl. Aliquots were taken for  $^1\text{H}$ -NMR analysis at each increment and changes in chemical shifts plotted and fitted using GraphPad Prism. The experiment was repeated in triplicate. Since the titration was performed in  $\text{D}_2\text{O}$ , I also needed to adjust the observed titration inflection point values to account for the different exchange rates of deuterium vs hydrogen. I used the formula below described by Krężel and Bal<sup>125</sup> to do this adjustment (Eq. 2.1).

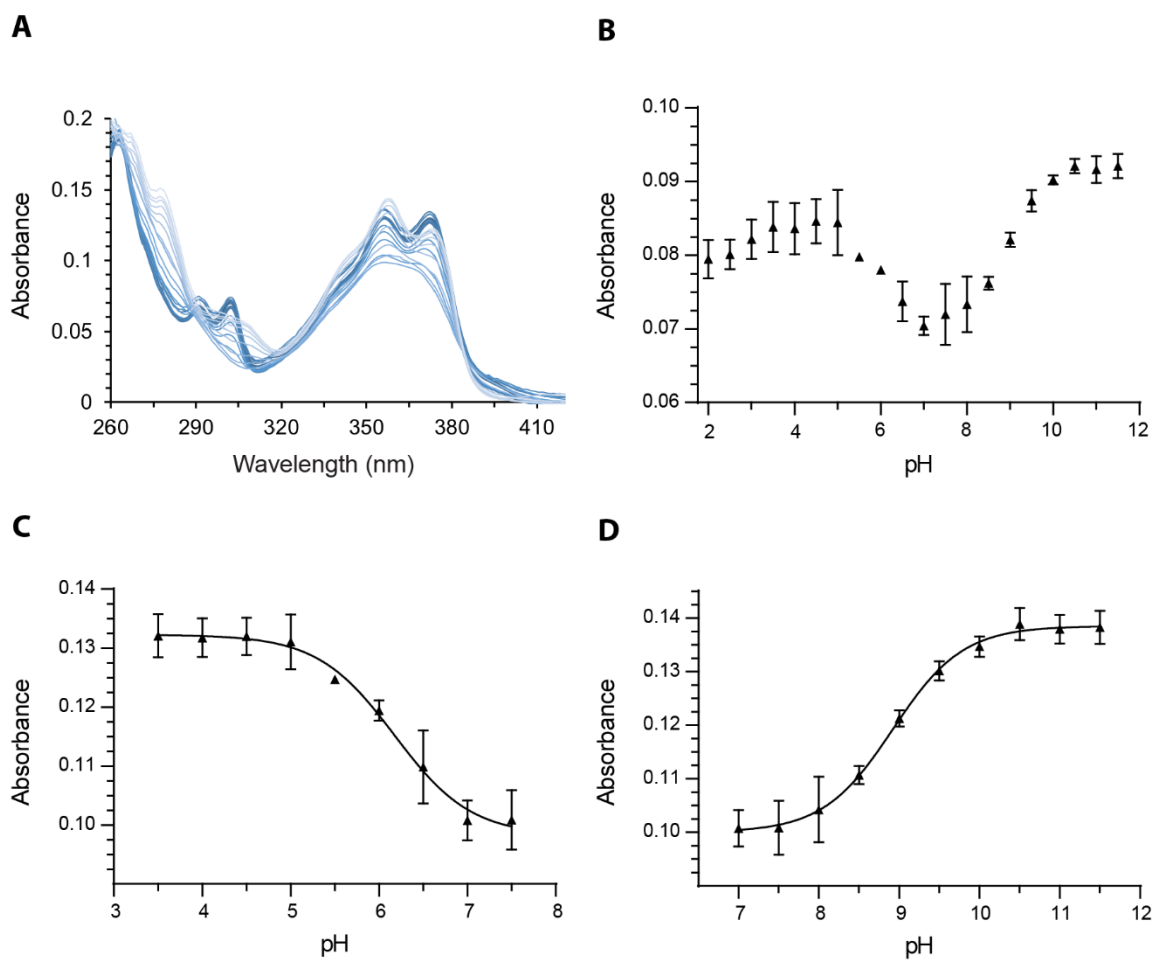
$$\text{pK}_a = 0.929\text{pK}_a(\text{obs}) + 0.42 \qquad \text{Eq. 2.1}$$



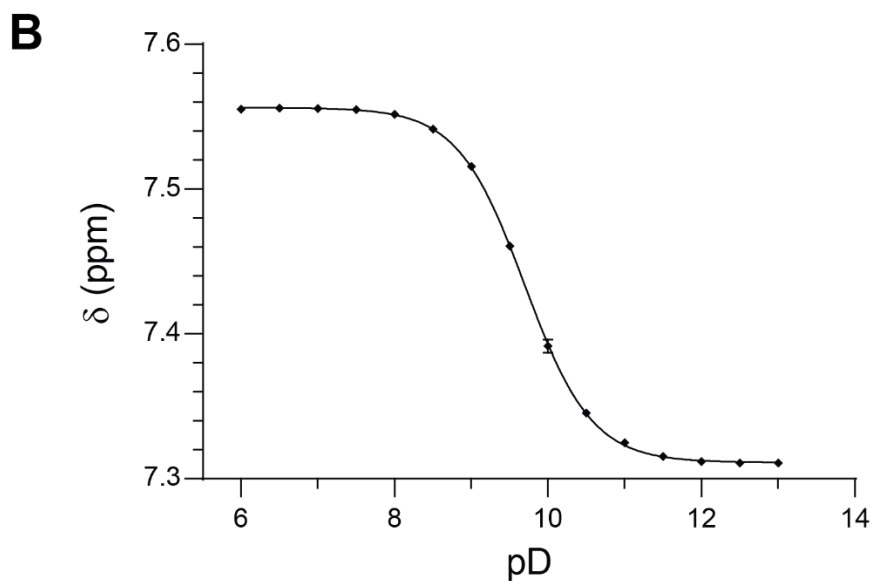
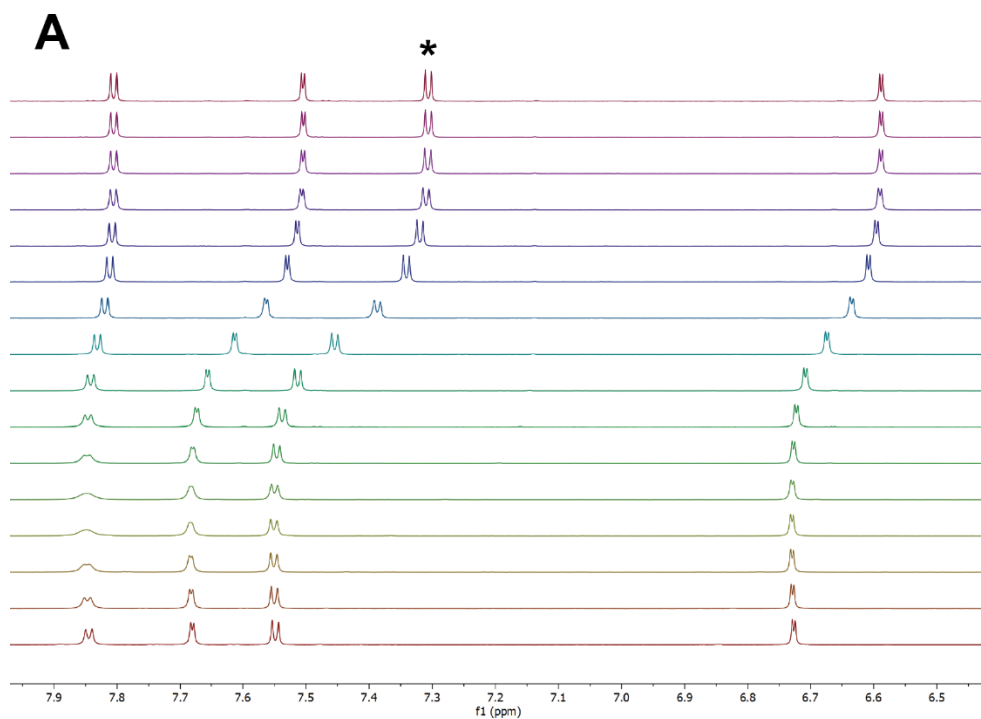
**Figure 2.12.** The fluorescence spectrum of **GAC** in water was measured at different pH values. **A.** Fluorescence spectra of one replicate. **B.** Fitted  $pK_a$  curve from the low pH values, data obtained at 390 nm, resulting in a  $pK_{a1}$  of  $2.57 \pm 0.03$ . **C.** Fitted  $pK_a$  curve from the high pH values, data obtained at 390 nm, resulting in a  $pK_{a2}$  of  $8.96 \pm 0.05$ .



**Figure 2.13.** The fluorescence spectrum of **GADAC** in water was measured at different pH values. A. Fluorescence spectra of one replicate. B. Fitted  $pK_a$  curve from the pH values, data obtained at 403 nm, resulting in a  $pH$  of  $6.42 \pm 0.04$ .



**Figure 2.14.** The absorbance of **GADAC** in water was measured at different pH values. A. Absorbance spectra of one replicate. B. The absorbance at 340 nm, the excitation wavelength for Figure 2.3A. C. Fitted  $pK_a$  curve from the low pH values, data obtained at 356 nm, resulting in a  $pH$  of  $6.18 \pm 0.1$ . D. Fitted  $pK_a$  curve from the high pH values, data obtained at 356 nm, resulting in a  $pH$  of  $8.92 \pm 0.07$ .



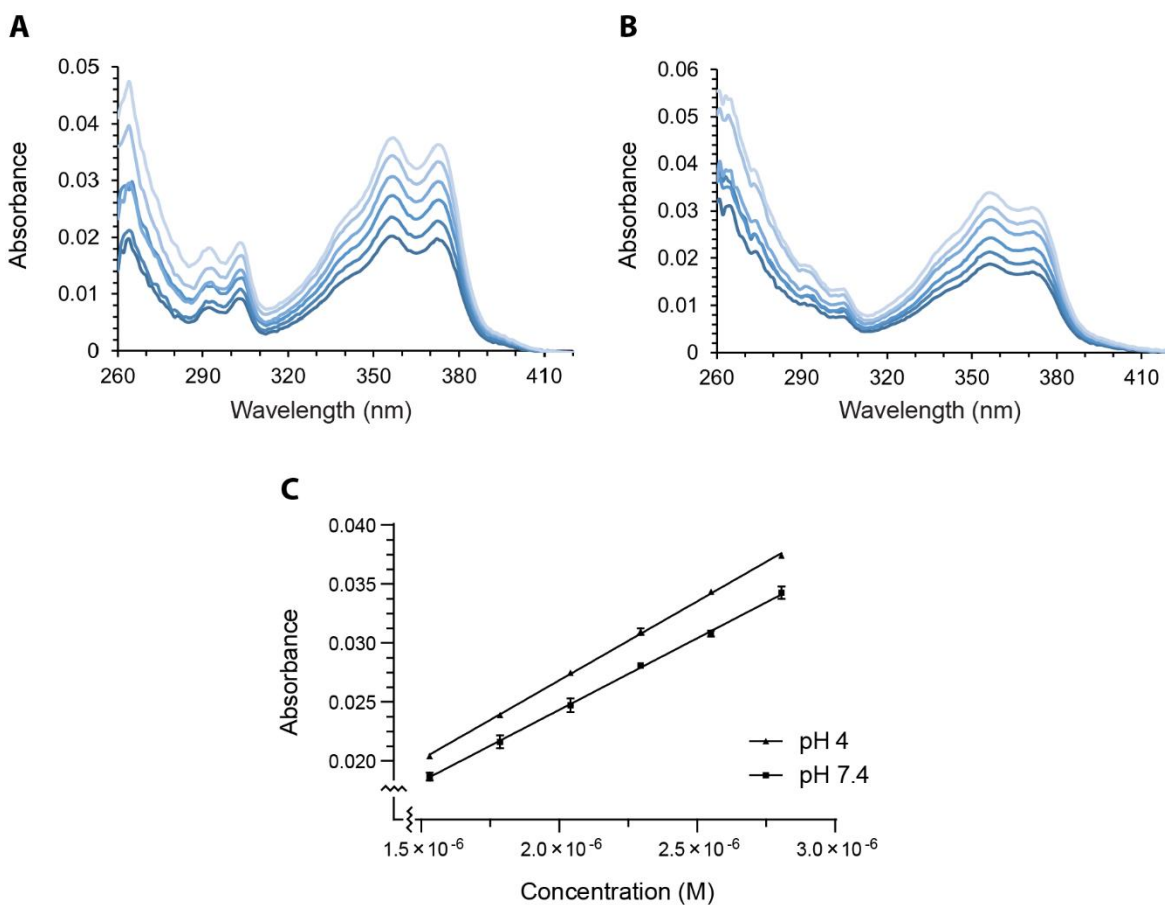
**Figure 2.15.** The <sup>1</sup>H NMR spectrum of **GAI** in D<sub>2</sub>O was measured at different pH values. A. <sup>1</sup>H NMR spectra of one replicate. B. Fitted pK<sub>a</sub> curve from the pH values, data obtained from the signal originating at 7.56, resulting in a pH of 9.43 ± 0.01.



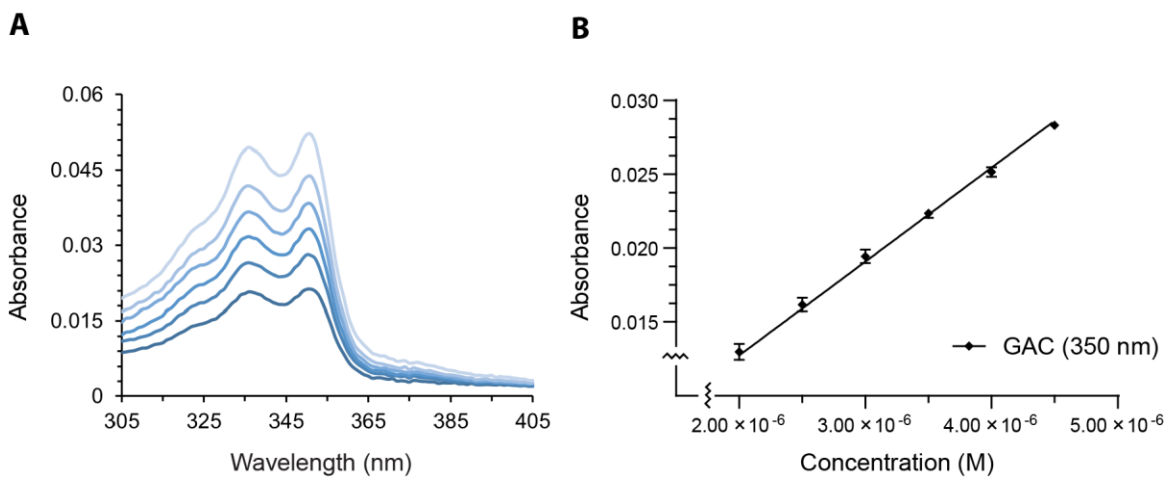
## Extinction Coefficient

A stock solution of **GADAC** (0.51 mM) in Milli-Q water was prepared. The concentration of this solution was confirmed by performing  $^1\text{H-NMR}$  on an aliquot spiked with an internal concentration reference (MeCN). Aliquots from these solutions were diluted into six cuvettes to a total of 3 mL buffer to produce a concentration gradient. The buffers used were a 10 mM citrate buffer with 150 mM NaCl, adjusted to pH 4 and a 10 mM phosphate buffer with 150 mM NaCl, adjusted to pH 7.4. The absorbance of each cuvette was obtained and normalized to a buffer-only blank. The concentration vs absorbance at  $\lambda_{\text{max}}$  was plotted in GraphPad Prism. The extinction coefficient was determined by performing a linear fit, forcing the X and Y intercept to be zero.

A stock solution of **GAC** (0.883 mM) in isopropanol was prepared. The concentration of this solution was confirmed by performing  $^1\text{H-NMR}$  on an aliquot spiked with an internal concentration reference (MeCN). Aliquots from this solution were diluted into six cuvettes to a total of 3 mL MilliQ water to produce a concentration gradient. The absorbance of each cuvette was obtained and normalized to a buffer-only blank. The concentration vs absorbance at  $\lambda_{\text{max}}$  was plotted in GraphPad Prism. The extinction coefficient was determined by performing a linear fit, forcing the X and Y intercept to be zero.



**Figure 2.16.** The extinction coefficient was determined for **GADAC** in aqueous buffer. A. The absorbance of **GADAC** at increasing concentrations in 10 mM citrate buffer with 150 mM NaCl, pH 4. B. The absorbance of **GADAC** at increasing concentrations in 10 mM phosphate buffer with 150 mM NaCl, pH 7.4.. C. The extinction coefficient was determined through a linear fit of concentration vs absorbance at  $\lambda_{\max}$  in GraphPad Prism:  $\epsilon_{\text{pH } 4} = 13,400 \pm 20 \text{ cm}^{-1}\text{M}^{-1}$ ,  $\epsilon_{\text{pH } 7.4} = 12,200 \pm 40 \text{ cm}^{-1}\text{M}^{-1}$ .



**Figure 2.17.** The extinction coefficient was determined for **GAC** in Milli-Q water. A. The absorbance of **GAC** at increasing concentrations. B. Linear fit of concentration vs absorbance at  $\lambda_{\max}$  in GraphPad Prism:  $\epsilon_{\text{GAC}} = 6400 \pm 30 \text{ cm}^{-1}\text{M}^{-1}$ .

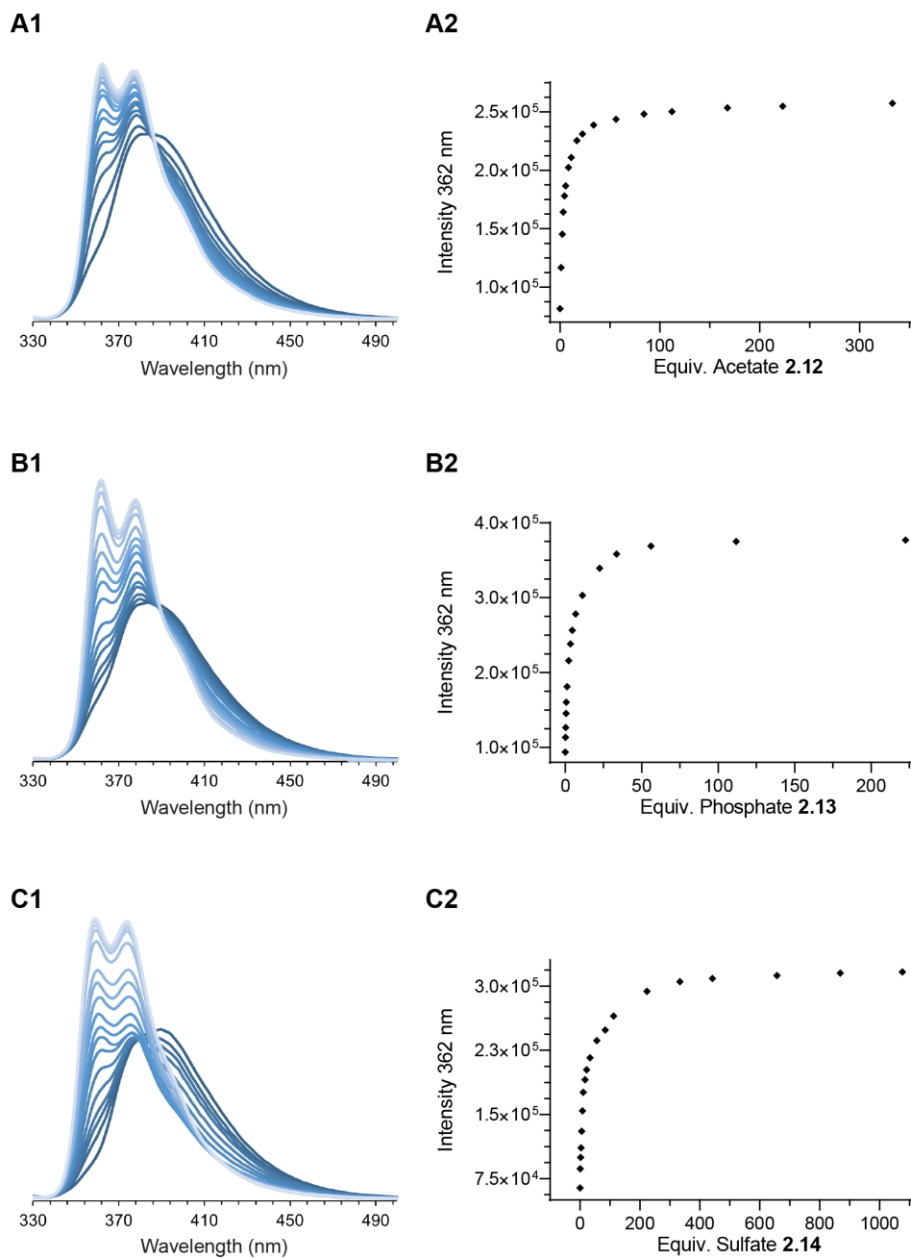
## Quantum Yield

Absolute quantum yields ( $\Phi$ ) were measured using a Quantaaurus-QY spectrometer (model C11374) from Hamamatsu. This instrument uses an integrating sphere to determine photons absorbed and emitted by a sample. Measurements were carried out using dilute samples ( $A < 0.1$ ) and self-absorption corrections were performed using the instrument software.<sup>126</sup> For quantum yield measurements, GADAC was first dissolved in DMSO to yield a 10 mM stock solution and then diluted 1:1000 (final concentration 10  $\mu\text{M}$ ) or 1:2000 (final concentration 5  $\mu\text{M}$ ) in either 10 mM citrate buffer containing 150 mM NaCl or 10 mM phosphate buffer containing 150 mM NaCl. Quantum yield values are averages of the values obtained from three separate experiments, two at 10  $\mu\text{M}$  and one at 5  $\mu\text{M}$ .

## Anion Binding Titrations

Bulk amounts of the three guanidine compounds were separated into many vials, and the resulting amount of guanidine per vial was calculated by obtaining an  $^1\text{H}$  NMR of the vial with an internal concentration reference: **GP** (1  $\mu\text{mol}$  per vial), **GAC** (0.883  $\mu\text{mol}$  per vial), **GADAC** (0.137  $\mu\text{mol}$  per vial). When a titration was performed, one mL of isopropanol was added to the chosen vial. A 100  $\mu\text{L}$  aliquot of this solution was then dissolved in 19,900  $\mu\text{L}$  of the desired solvent to obtain a final concentration of **GP** (5  $\mu\text{M}$ ), **GAC** (4.42  $\mu\text{M}$ ), and **GADAC** (0.685  $\mu\text{M}$ ). The desired anion, **12**, **13**, or **14**, (0.1 mmol) was then dissolved in 1 mL of guanidine stock solution to make a 100 mM solution of anion. The anion was serially diluted with more guanidine stock solution to make 10 mM and 1 mM solutions as necessary.

The guanidine solution (3 mL) was then added to a cuvette and a fluorescence emission scan was taken. The anion solution of choice was added sequentially to the desired equivalents and an emission scan acquired after each. Guanidine concentration was constant, but anion concentrations were back-calculated to account for dilution. An appropriate wavelength was chosen for each guanidine and the intensity of that wavelength at each concentration of anion was determined and plotted. These curves were obtained in triplicate and analyzed in one input in the web-hosted program Bindfit on [supramolecular.org](http://supramolecular.org). One-replicate examples of the full emission scans are shown below as well as the binding curves for the selected wavelengths. The open-source data for each titration is included as a [supramolecular.org](http://supramolecular.org) link.

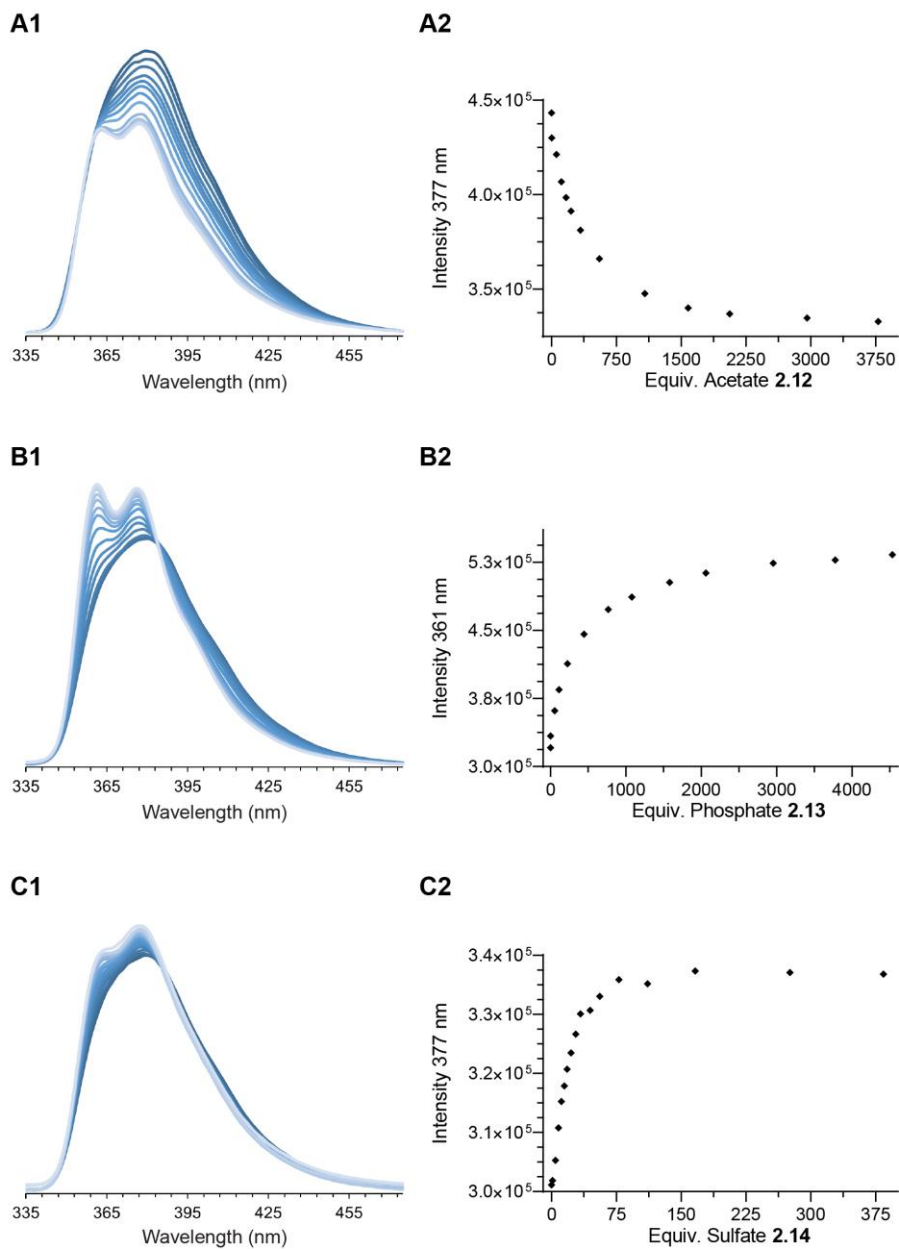


**Figure 2.18.1.** Binding titrations of **GAC** and anions in **isopropanol**. A1. Fluorescence spectra of one replicate with acetate **12**. B1. Fluorescence spectra of one replicate with phosphate **13**. C1. Fluorescence spectra of one replicate with sulfate **14**. A2-C2. Binding curve of one replicate with the same respective anion, not fitted. Fits can be found at the following links.

A. <http://app.supramolecular.org/bindfit/view/fb149c6f-55db-4f79-82e6-7a6227a9534c>

B. <http://app.supramolecular.org/bindfit/view/83244507-9aa8-4feb-8eae-a99e477d688c>

C. <http://app.supramolecular.org/bindfit/view/26ac45b7-f96c-46e0-9443-4a3209ebb8d1>

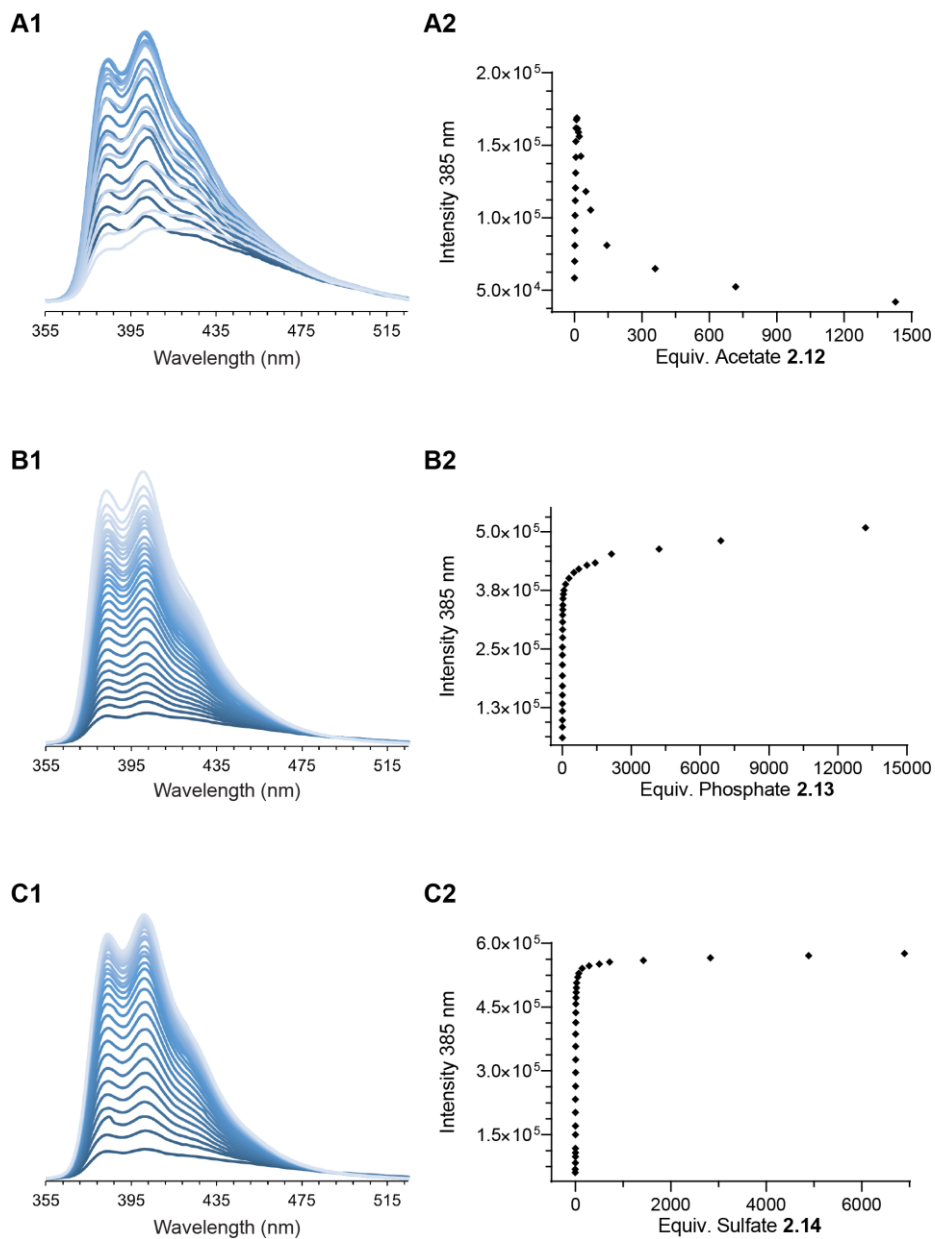


**Figure 2.18.2.** Binding titrations of **GAC** and anions in **methanol**. A1. Fluorescence spectra of one replicate with acetate **12**. B1. Fluorescence spectra of one replicate with phosphate **13**. C1. Fluorescence spectra of one replicate with sulfate **14**. A2-C2. Binding curve of one replicate with the same respective anion, not fitted. Fits can be found at the following links.

A. <http://app.supramolecular.org/bindfit/view/5f1bac0d-a54a-4a32-b2eb-a52ec468ee83>

B. <http://app.supramolecular.org/bindfit/view/165c9e03-172e-458d-89a2-3e6542b0da07>

C. <http://app.supramolecular.org/bindfit/view/ea0a2b47-7769-444b-96fb-a904fe6e3bf4>



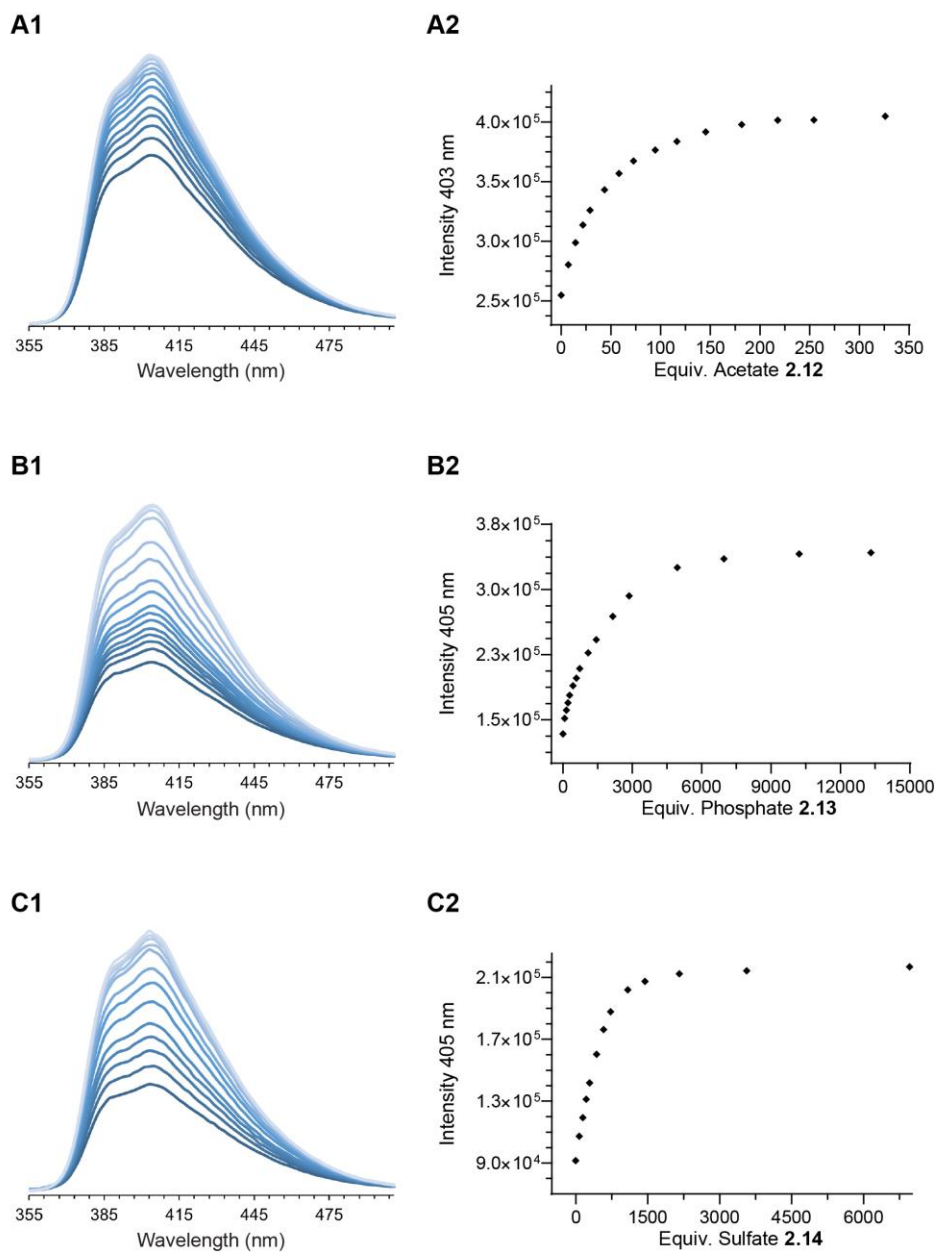
**Figure 2.18.3.** Binding titrations of **GADAC** and anions in **methanol**. A1. Fluorescence spectra of one replicate with acetate **12**. B1. Fluorescence spectra of one replicate with phosphate **13**. C1. Fluorescence spectra of one replicate with sulfate **14**. A2-C2. Binding curve of one replicate with the same respective anion, not fitted. Fits can be found at the following links.

A. <http://app.supramolecular.org/bindfit/view/dbab7eda-a2a1-4cb3-a4f0-6601e7aa97c2>

B. <http://app.supramolecular.org/bindfit/view/467f199e-c240-419d-a531-f718e66c12d4>

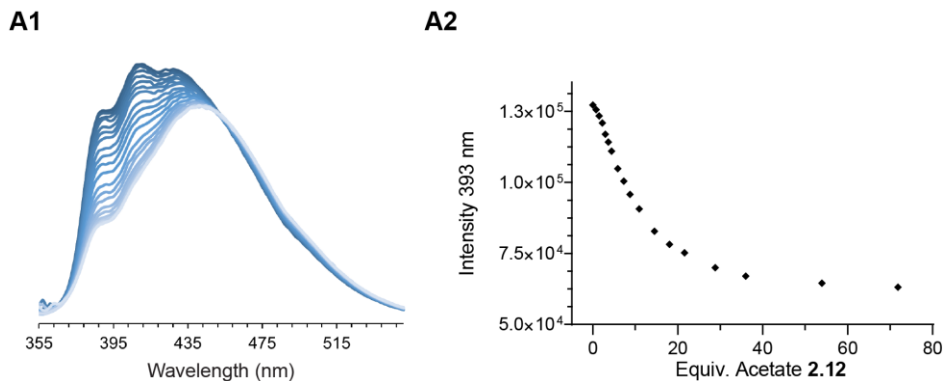
C. <http://app.supramolecular.org/bindfit/view/b82db046-d174-41d8-9708-6709c45bb904>



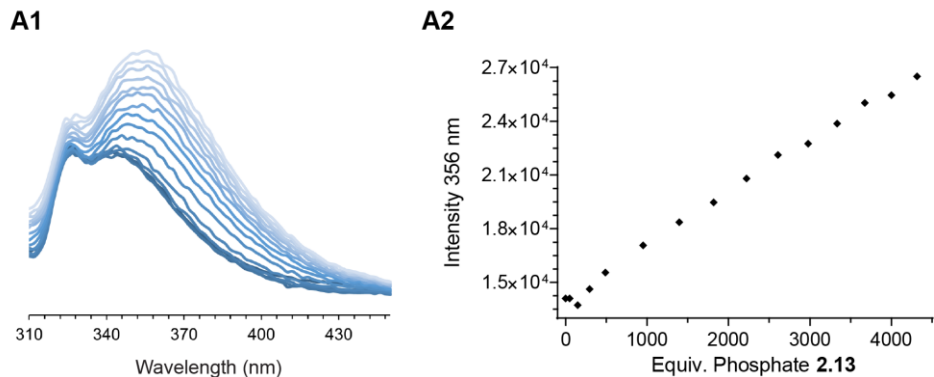


**Figure 2.18.4.** Binding titrations of **GADAC** and anions in **water**. A1. Fluorescence spectra of one replicate with acetate **12**. B1. Fluorescence spectra of one replicate with phosphate **13**. C1. Fluorescence spectra of one replicate with sulfate **14**. A2-C2. Binding curve of one replicate with the same respective anion, not fitted. Fits can be found at the following links.

- A. <http://app.supramolecular.org/bindfit/view/bbb829b9-f1df-4b52-8f63-21127053995d>  
 B. <http://app.supramolecular.org/bindfit/view/dc29a248-7d50-4ab6-b3c7-c02a1aa461d3>  
 C. <http://app.supramolecular.org/bindfit/view/9acaf0b2-de07-4e13-8a00-928ad2fe93fc>

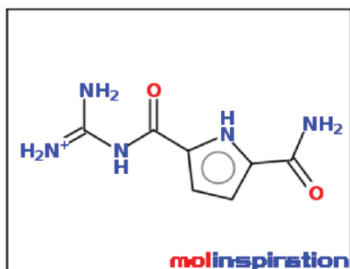


**Figure 2.18.5.** Binding titrations of **GADAC** and acetate **12** in **40% water in DMSO**. A1. Fluorescence spectra of one replicate. A2 Binding curve of one replicate, not fitted. Fits can be found at the following link.  
<http://app.supramolecular.org/bindfit/view/8aee4c70-c3ee-4ad6-8ecb-083d63e42c86>



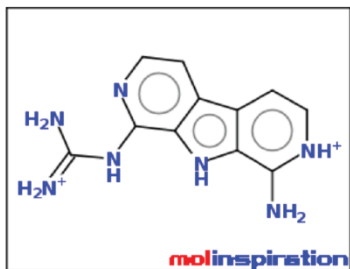
**Figure 2.18.6.** Binding titrations of **GP** and phosphate **13** in **isopropanol**. A1. Fluorescence spectra of one replicate. A2 Binding curve of one replicate, not fitted. Fits can be found at the following link.  
<http://app.supramolecular.org/bindfit/view/3d643da0-1c57-4434-a9ca-349f3453bcb9>

## cLog *P* and TPSA Calculations



[Molinspiration\\_property\\_engine](#) v2021.10

<a href="#">miLogP</a>	-4.47
<a href="#">TPSA</a>	139.60
<a href="#">natoms</a>	14
<a href="#">Mw</a>	196.19
<a href="#">nON</a>	7
<a href="#">nOHNH</a>	8
<a href="#">nviolations</a>	1
<a href="#">nrotb</a>	3
<a href="#">volume</a>	166.97



[Molinspiration\\_property\\_engine](#) v2021.10

<a href="#">miLogP</a>	-3.01
<a href="#">TPSA</a>	132.48
<a href="#">natoms</a>	18
<a href="#">Mw</a>	243.27
<a href="#">nON</a>	7
<a href="#">nOHNH</a>	9
<a href="#">nviolations</a>	1
<a href="#">nrotb</a>	2
<a href="#">volume</a>	211.53

**Figure 2.19.** The cLog *P* and TPSA for **GADAC** and **GCP** were calculated on using the web-hosted software on <https://www.molinspiration.com> (Molinspiration Cheminformatics Slovensky Grob, Slovakia).

## **Microscopy**

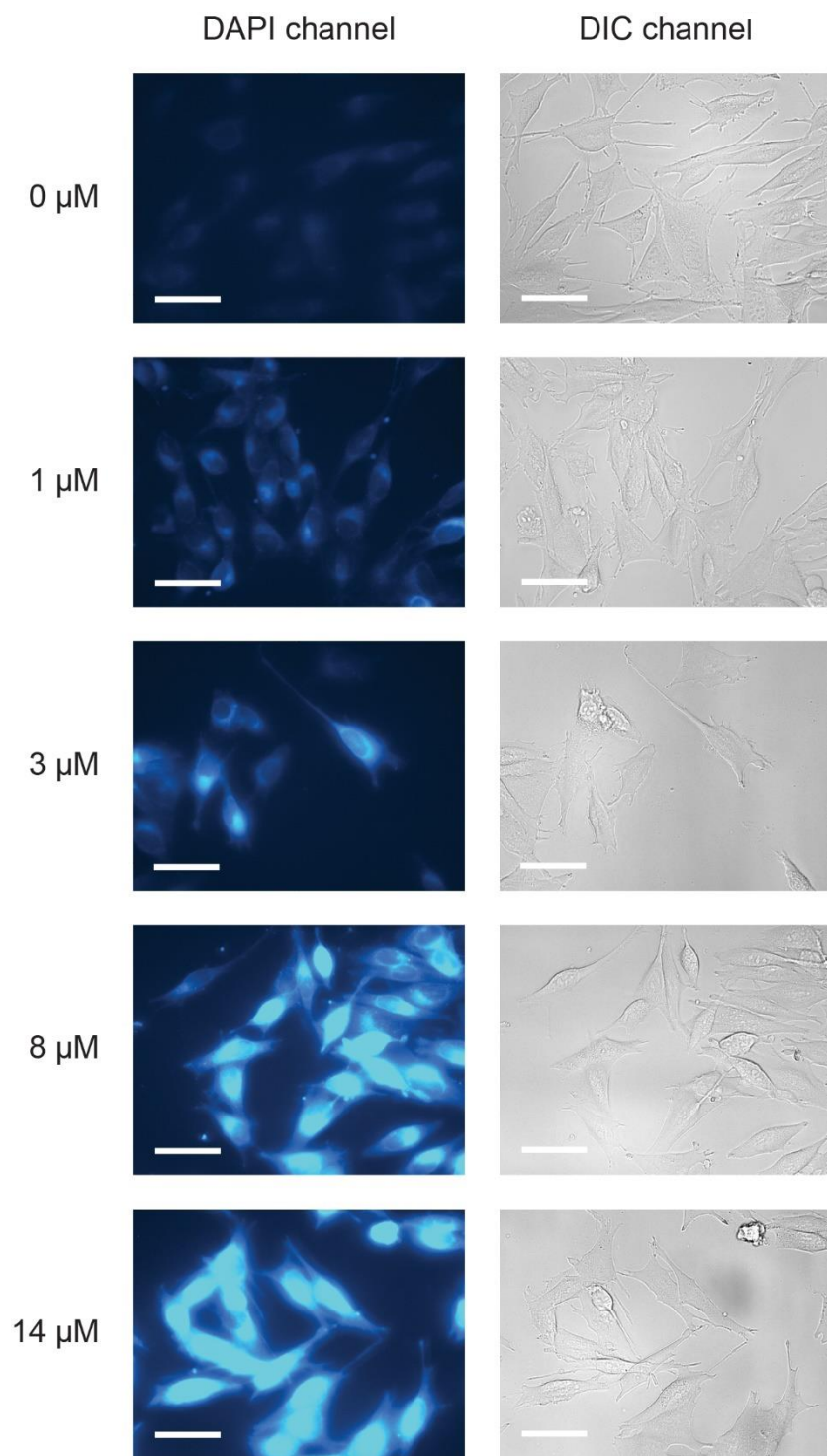
Biological Reagents, Supplies, and Instrumentation. Penicillin-streptomycin solution containing 10,000 units/mL of penicillin and 10,000 µg/mL of streptomycin (Catalog Number: 15140122) was from Thermo Fisher Scientific (Waltham, MA). DMEM, powder, high glucose (Catalog Number: 12100046) for M21 cells was from Thermo Fisher Scientific. Fetal Bovine Serum (FBS), Premium, US Sourced (Catalog Number: 45001-108) was from Corning (Corning, NY). Trypsin-EDTA (0.05%) with phenol red was from Thermo Fisher Scientific. Brightfield and fluorescent live cells images were acquired using an epifluorescent EVOS M7000 Imaging System (Catalog Number: AMF7000) from Thermo Fisher Scientific. µ-Slide 18-well chambered coverslips (Catalog Number: 81816, ibiTreat: #1.5 polymer coverslip, tissue culture treated, sterilized) from Ibi (Fitchburg, WI) were used for live cell imaging. DPBS with calcium and magnesium (Catalog Number: 14040141) was from Gibco (Waltham, MA). FluoroBrite DMEM (Catalog Number: A1896701) was from Thermo Fisher Scientific.

Cell line and cell culture conditions. The human melanoma M21 cell line<sup>127,128</sup> was a kind gift from Dr. Oscar Ortiz (Merck KGaA, Darmstadt, Germany). The cell line was tested negative for mycoplasma using the Lonza MycoAlert Plus kit. The M21 cell line was further authenticated by short tandem repeat profiling (STR) to validate the identity of the cell line and rule out intraspecies contamination. M21 cells were grown in sterile culture flasks in a cell culture incubator at 37 °C under CO<sub>2</sub> (5% v/v). Cells were counted to determine seeding density using a Countess II FL Automated Cell Counter from Thermo Fisher Scientific. To minimize genetic drift, thawed vials were used for fewer than twenty passages. M21 cells were grown in high-glucose DMEM medium (Catalog Number: 12100046) from Thermo Fisher Scientific supplemented with 1.5 g/L sodium

bicarbonate, 10% FBS, penicillin (100 U/mL), and streptomycin (100 µg/mL). The cells were passaged upon reaching 80% confluency with trypsin-EDTA (0.05%).

Preparation of GADAC stock solution for imaging. A stock solution of GADAC (137 µM) was prepared by dissolving 0.137 µmol of GADAC in 1 mL of Milli-Q water immediately before use in imaging experiments.

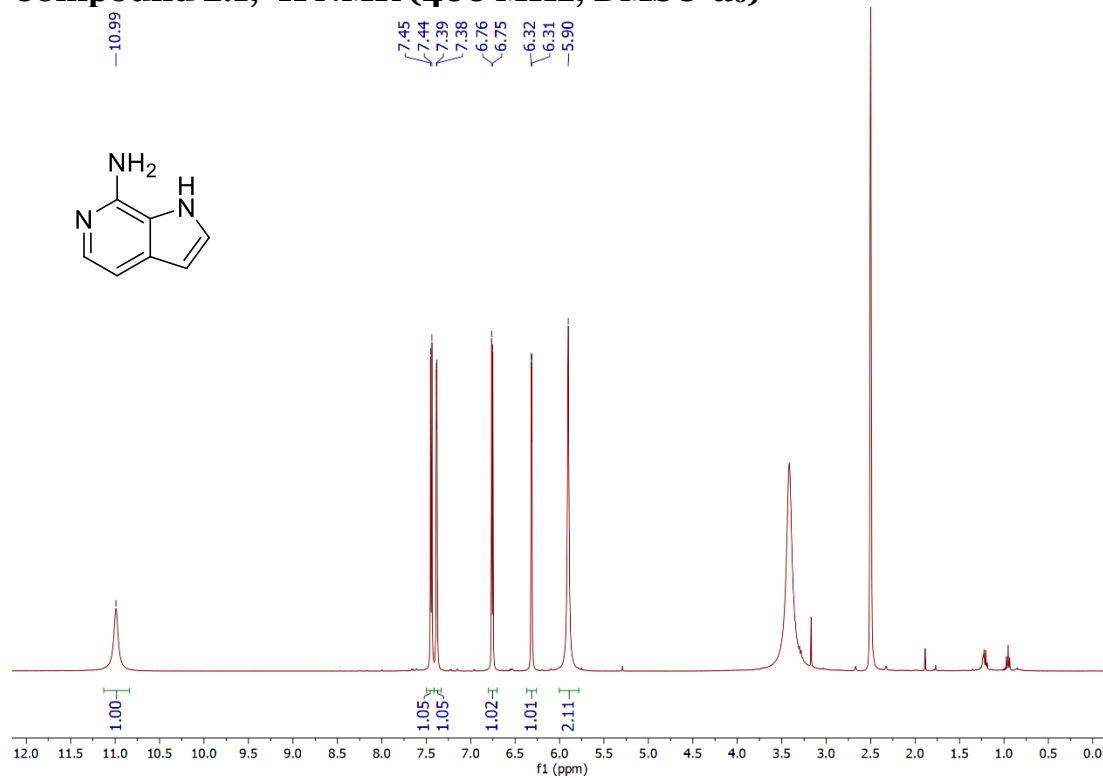
Imaging experiment. Before treatment with GADAC, M21 cells in FBS-supplemented DMEM (100 µL per well) were seeded at a density of 15,000 cells/well in a sterile 18-well plate from Ibidi. After 19 h, the cells were washed with DPBS (×2) and incubated with either FBS-supplemented DMEM alone (100 µL per well; negative control) or GADAC (100 µL per well; solutions of 1 µM, 3 µM, 8 µM, or 14 µM) in FBS-supplemented DMEM (≤10% v/v Milli-Q water) for 1 h at 37 °C. After the indicated time, the treatment medium was removed from cells and the cells were rinsed with DPBS (×3) and FluoroBrite DMEM (×3). The cells were then placed in FluoroBrite DMEM (100 µL per well) and examined using an epifluorescent EVOS M7000 microscope.



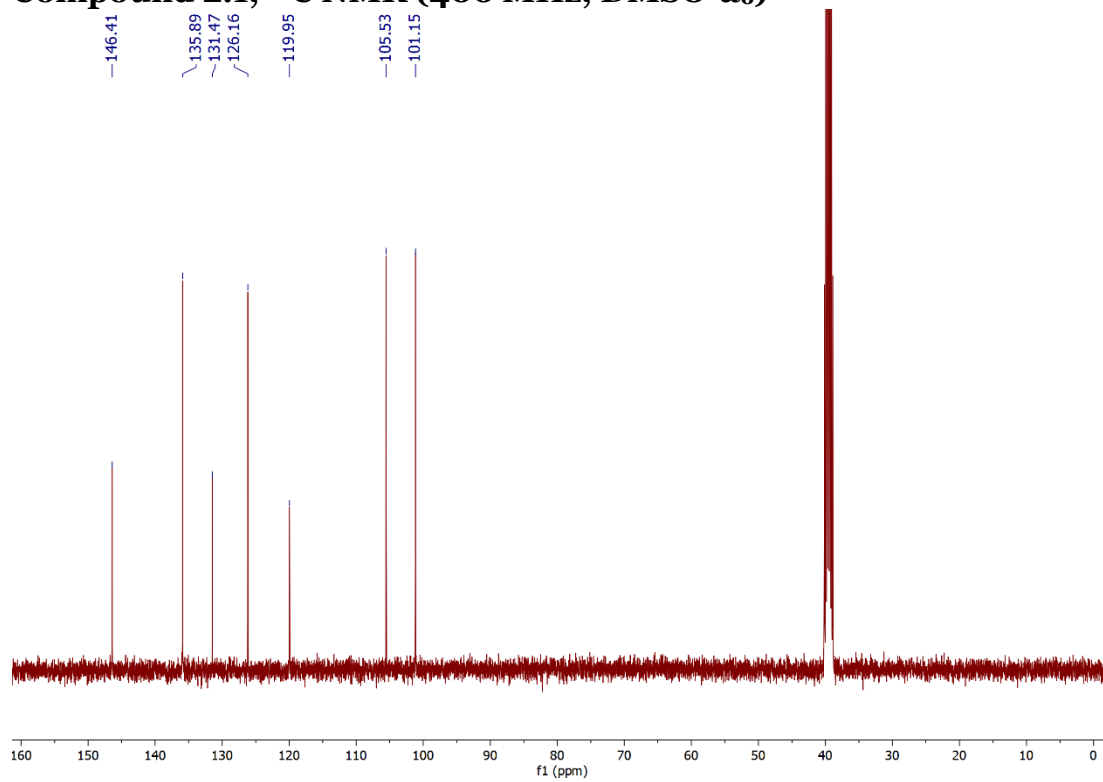
**Figure 2.20.** Images of the uptake of **GADAC** into live M21 cells over one hour at 37 °C in the presence of serum. DAPI channel (left) with 357/44 nm excitation and 447/60 nm emission; DIC channel (right). The DAPI channel epifluorescent images are normalized (were taken at the same LED intensity, exposure, and gain parameters). Scale bars correspond to 50  $\mu\text{m}$ .

# NMR Spectra

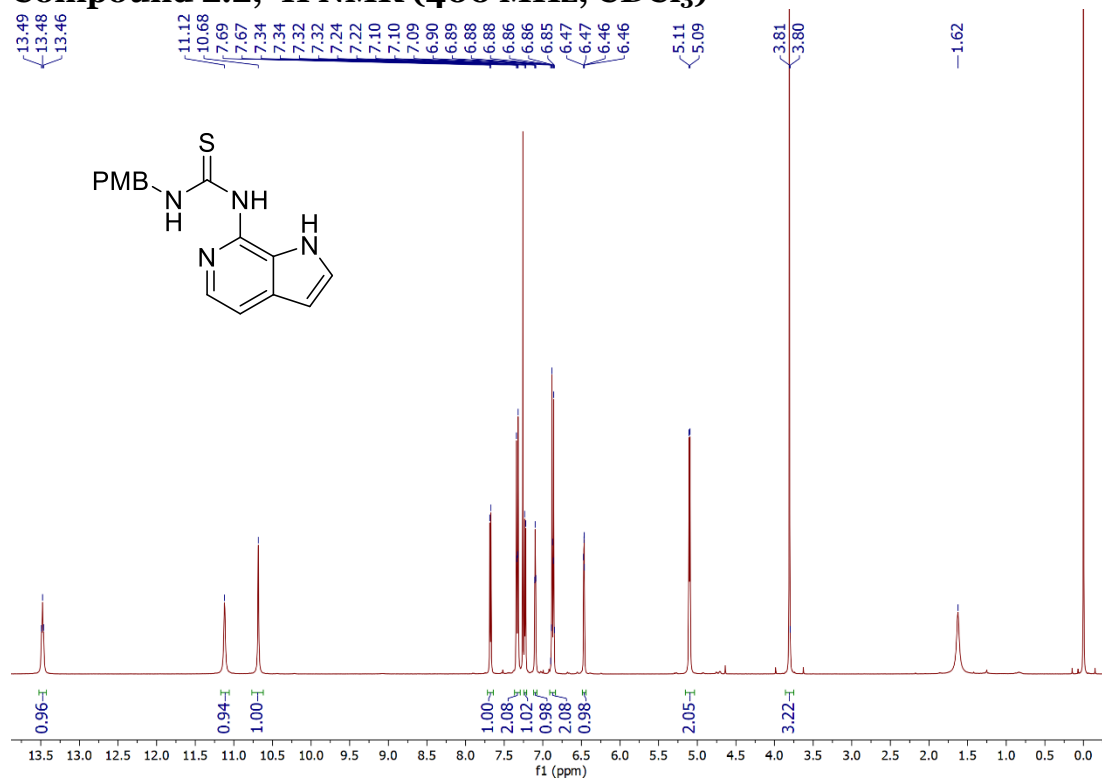
## Compound 2.1, $^1\text{H}$ NMR (400 MHz, $\text{DMSO-}d_6$ )



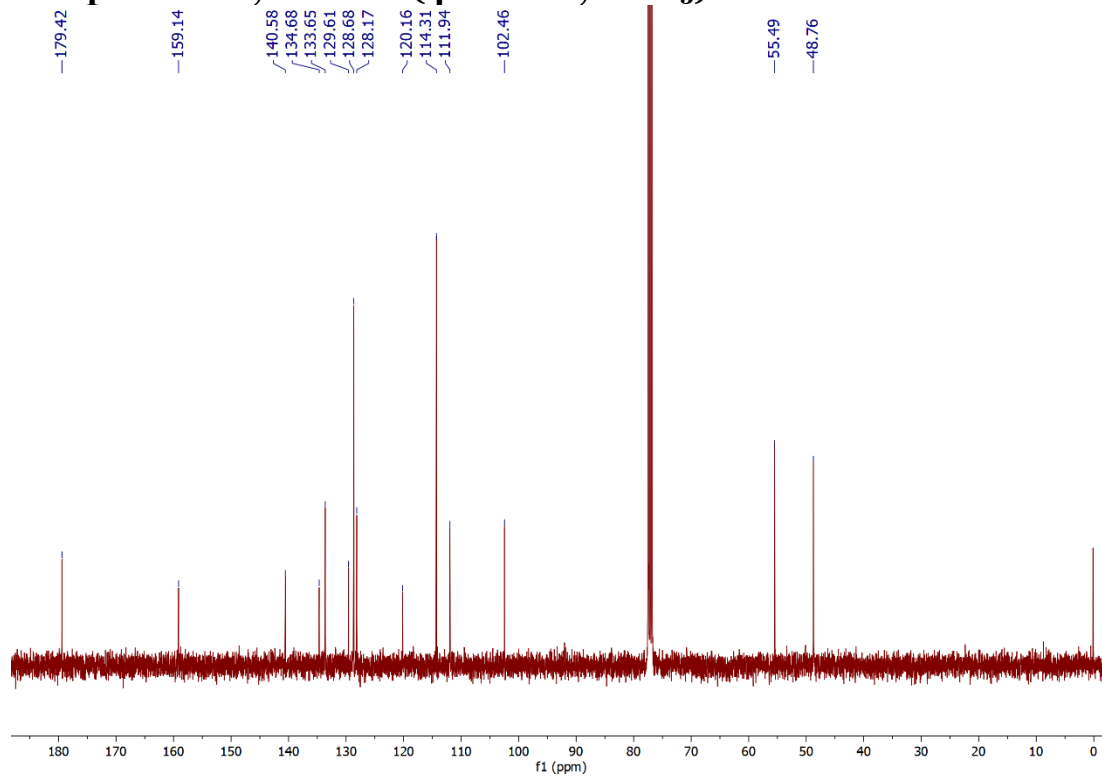
## Compound 2.1, $^{13}\text{C}$ NMR (400 MHz, $\text{DMSO-}d_6$ )



### Compound 2.2, <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)

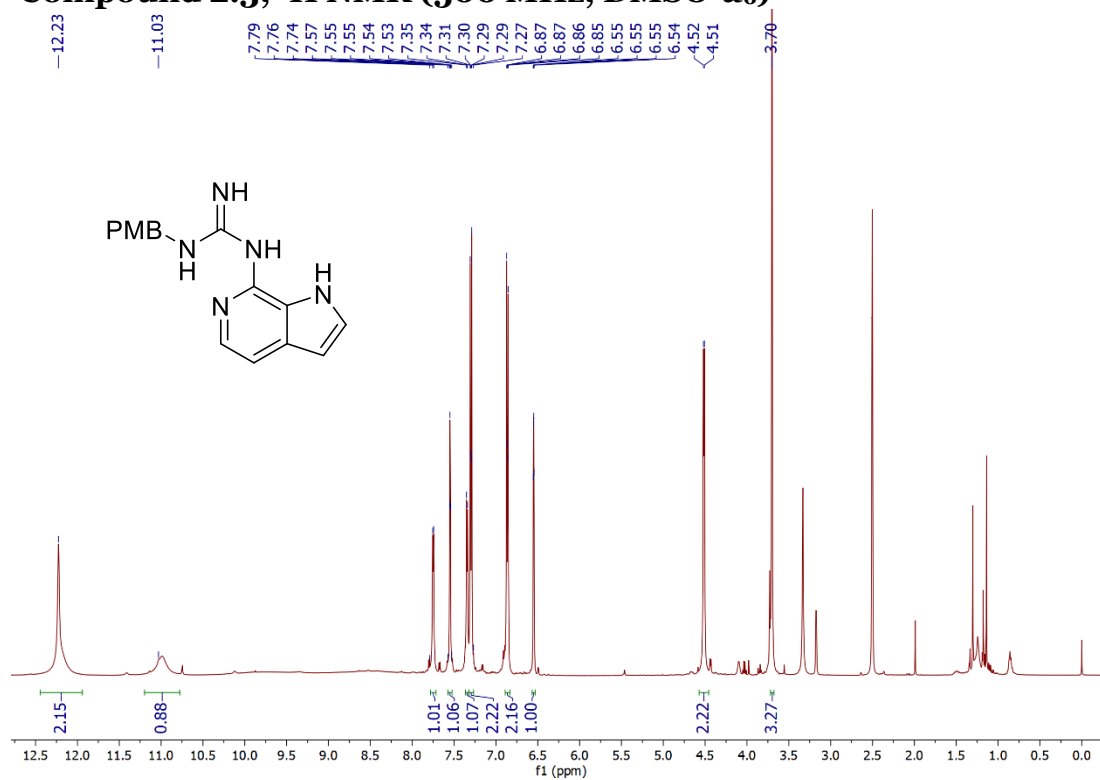


### Compound 2.2, <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)

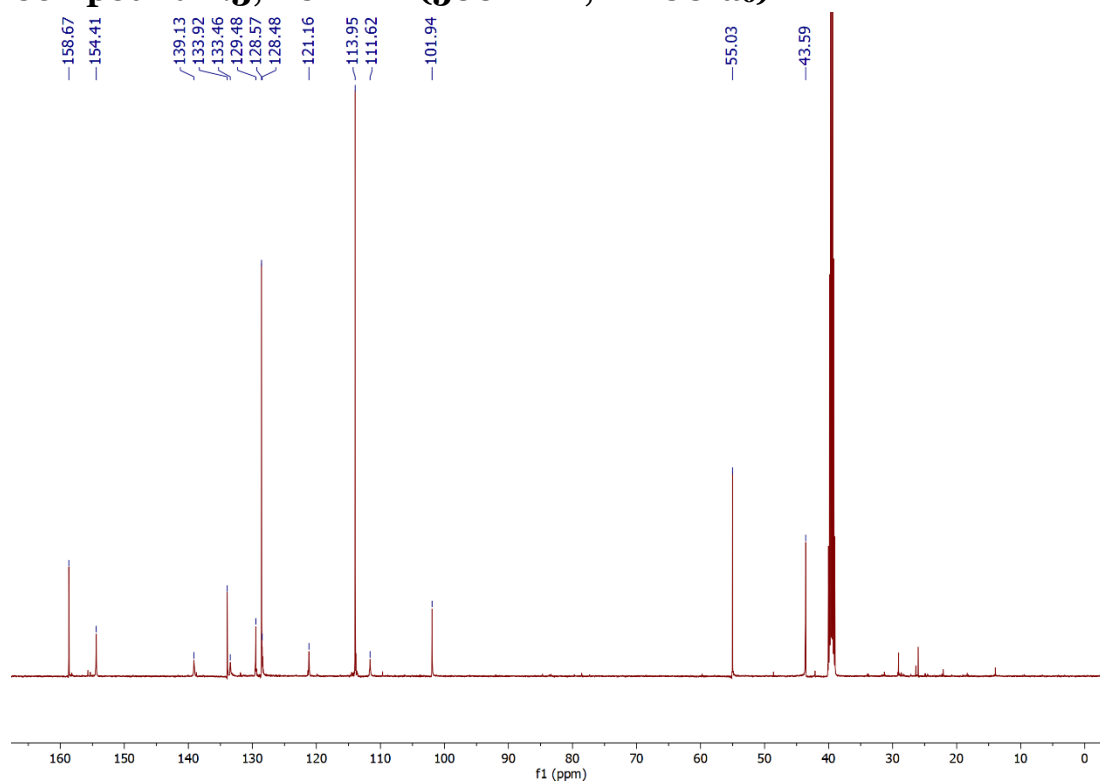




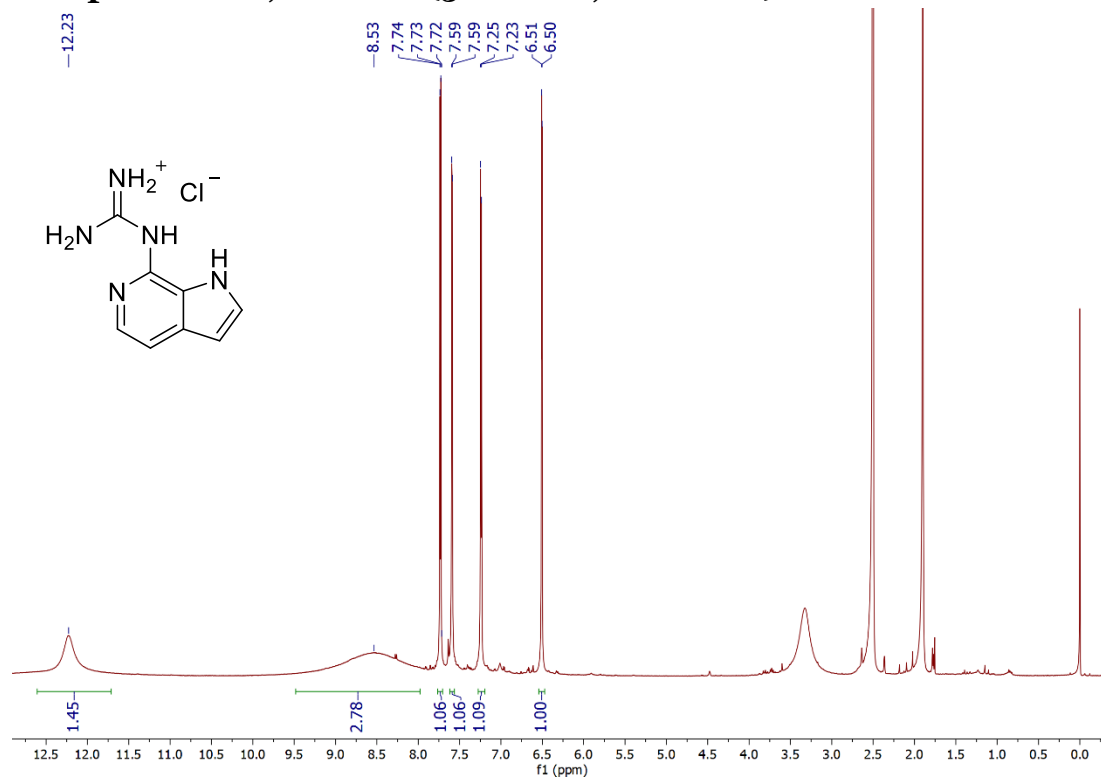
### Compound 2.3, <sup>1</sup>H NMR (500 MHz, DMSO-d<sub>6</sub>)



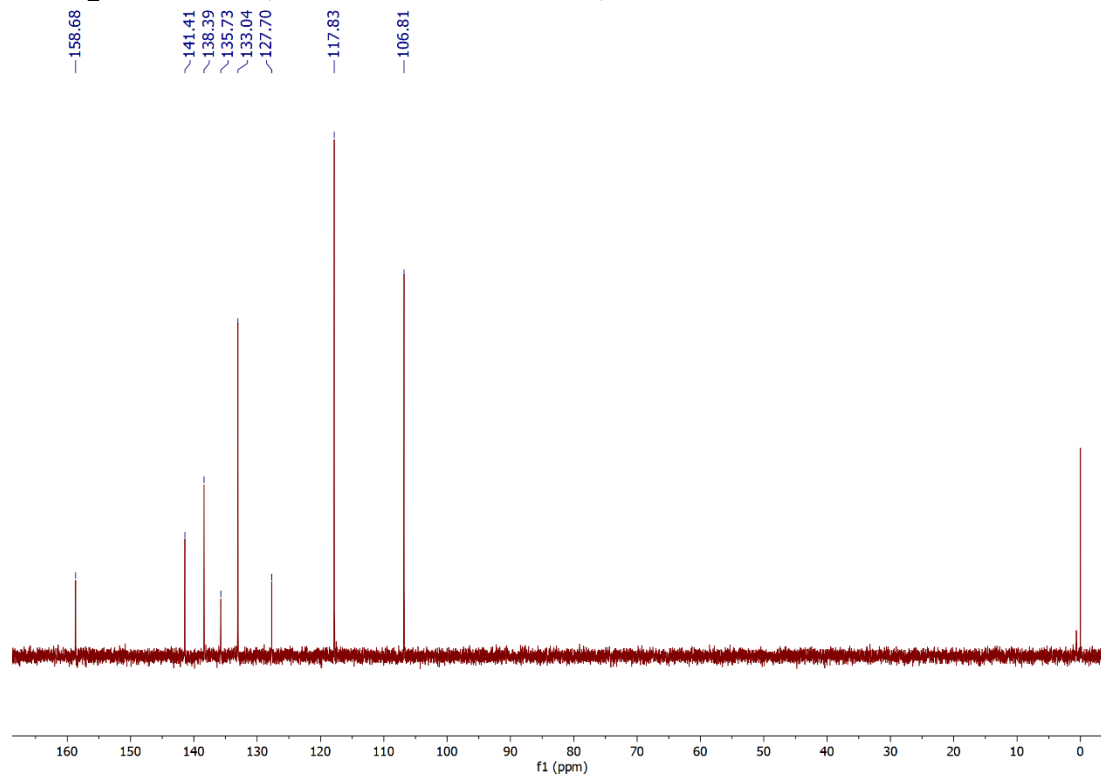
### Compound 2.3, <sup>13</sup>C NMR (500 MHz, DMSO-d<sub>6</sub>)



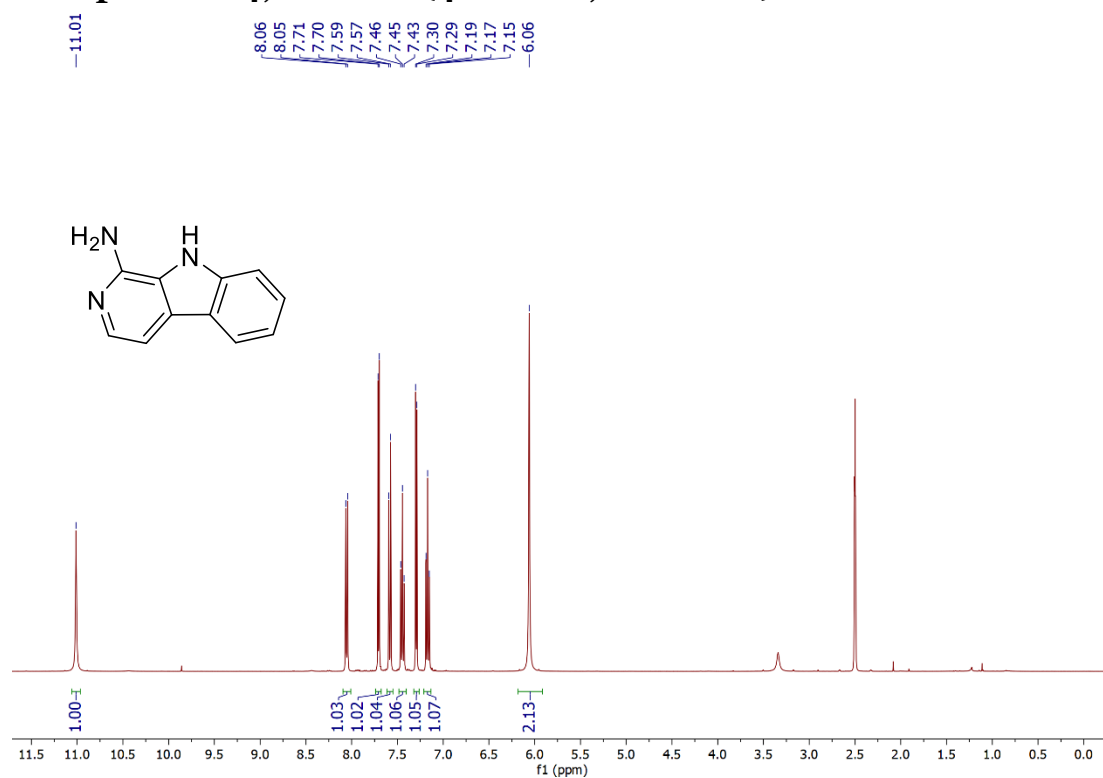
### Compound GAI, <sup>1</sup>H NMR (500 MHz, DMSO-d<sub>6</sub>)



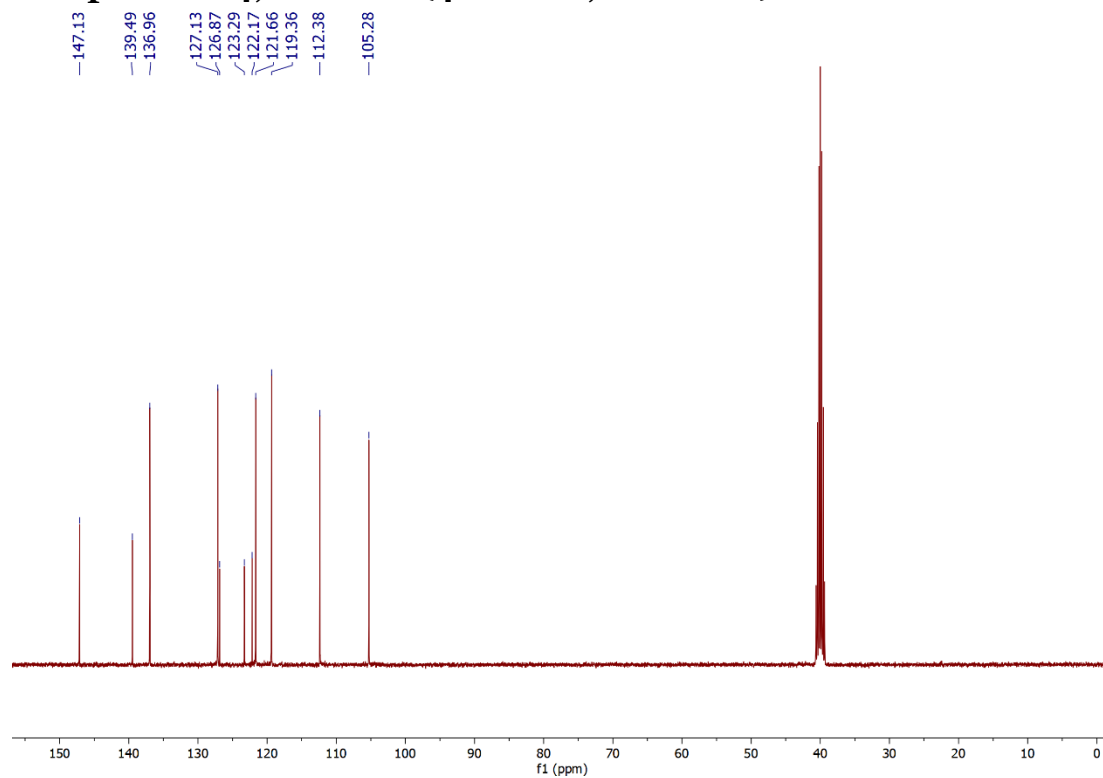
### Compound GAI, <sup>13</sup>C NMR (500 MHz, DMSO-d<sub>6</sub>)



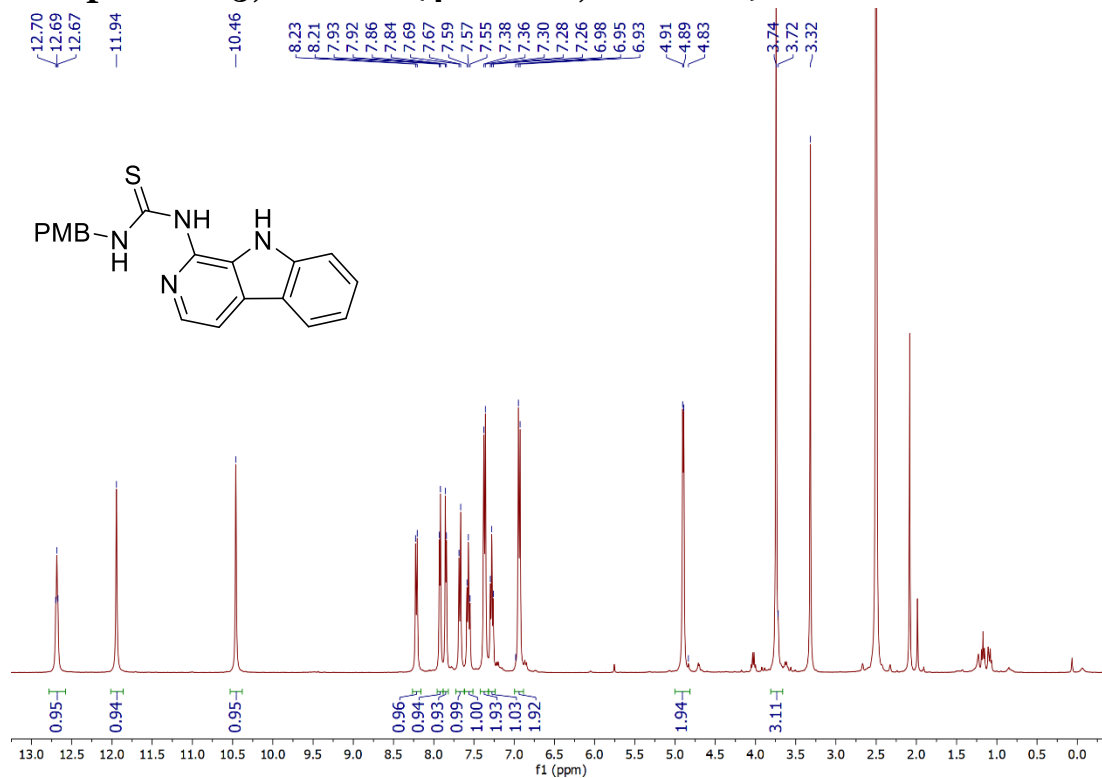
### Compound 2.4, $^1\text{H}$ NMR (400 MHz, $\text{DMSO-}d_6$ )



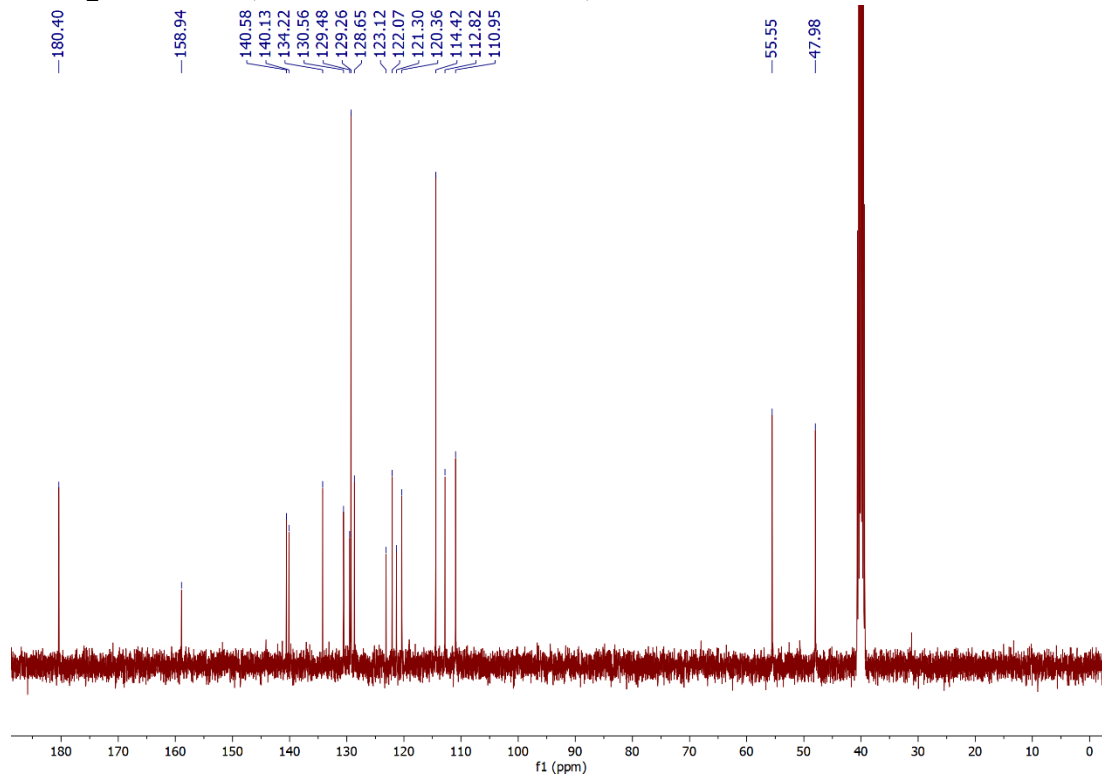
### Compound 2.4, $^{13}\text{C}$ NMR (400 MHz, $\text{DMSO-}d_6$ )



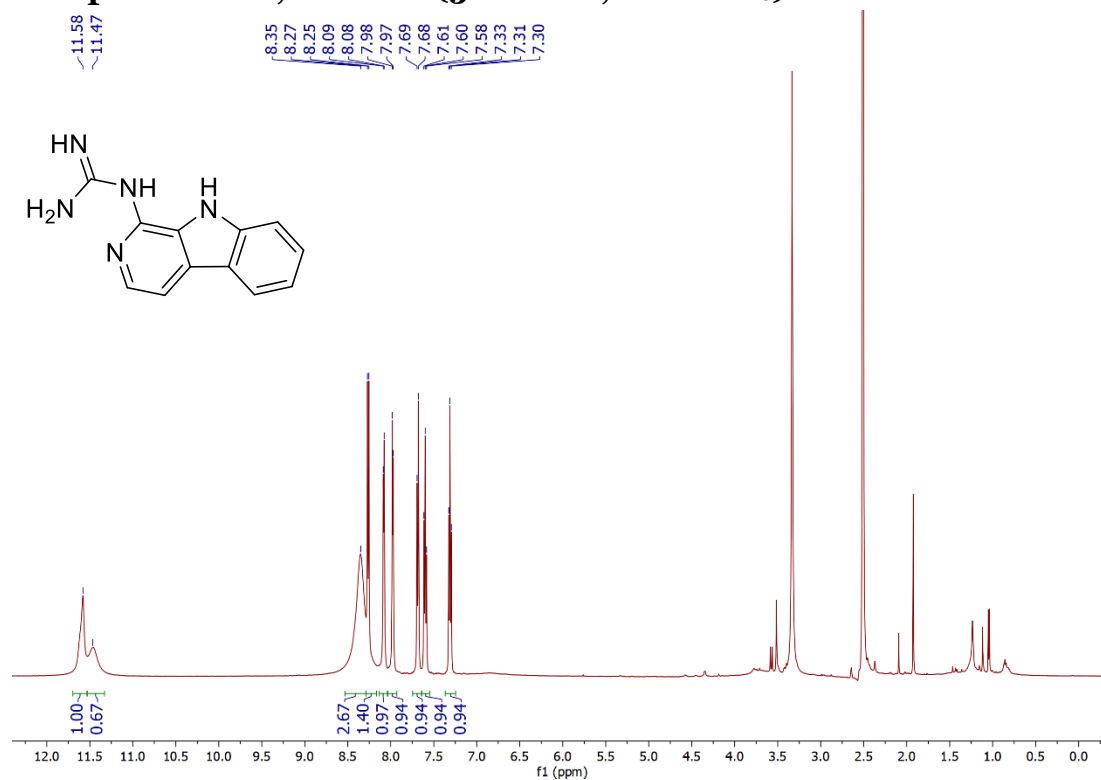
### Compound 2.5, <sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>)



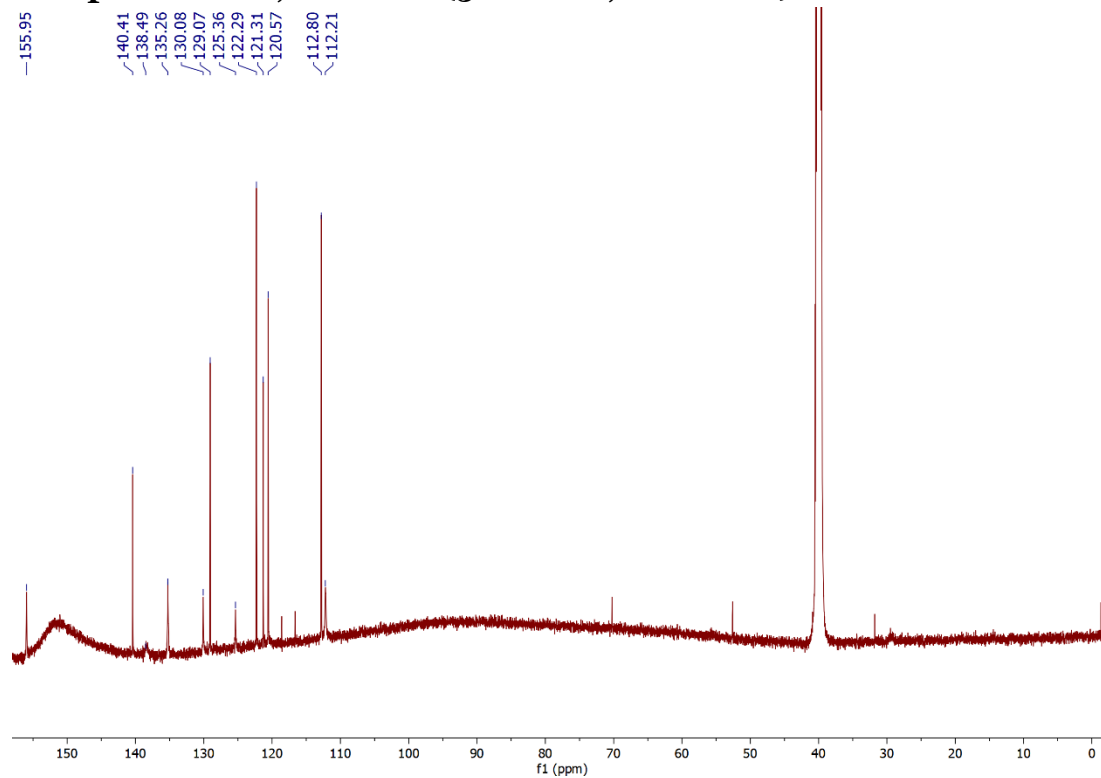
### Compound 2.5, <sup>13</sup>C NMR (400 MHz, DMSO-d<sub>6</sub>)



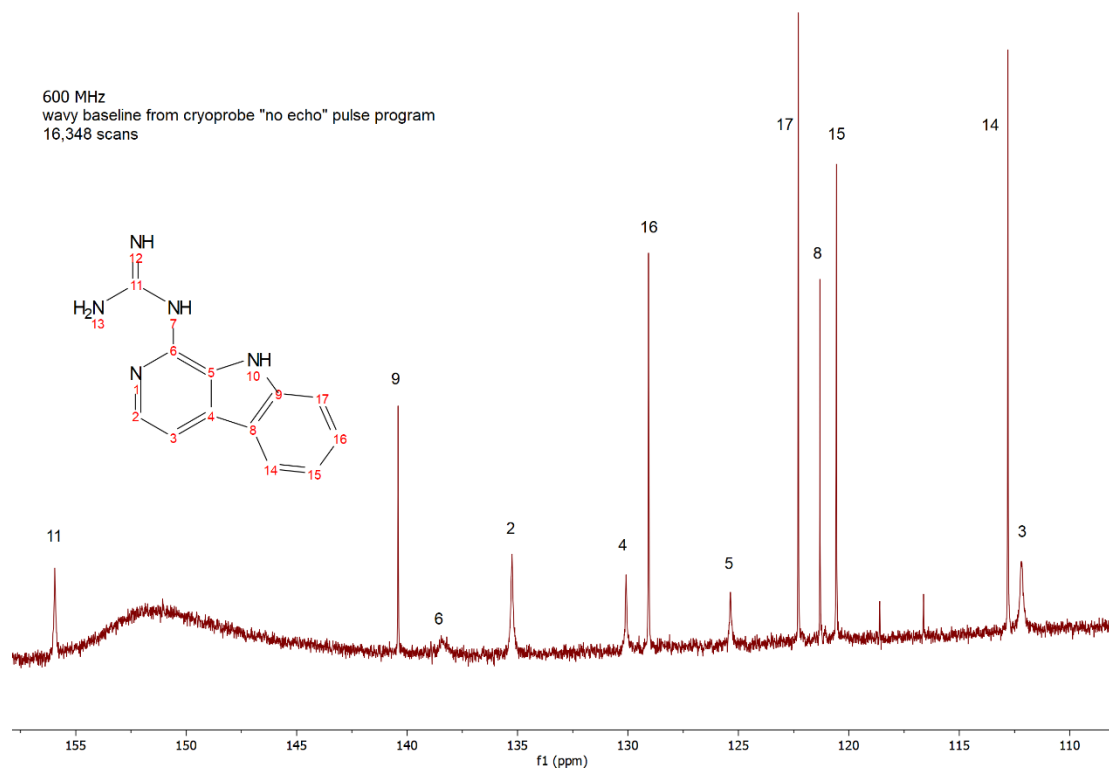
### Compound GAC, <sup>1</sup>H NMR (500 MHz, DMSO-d<sub>6</sub>)



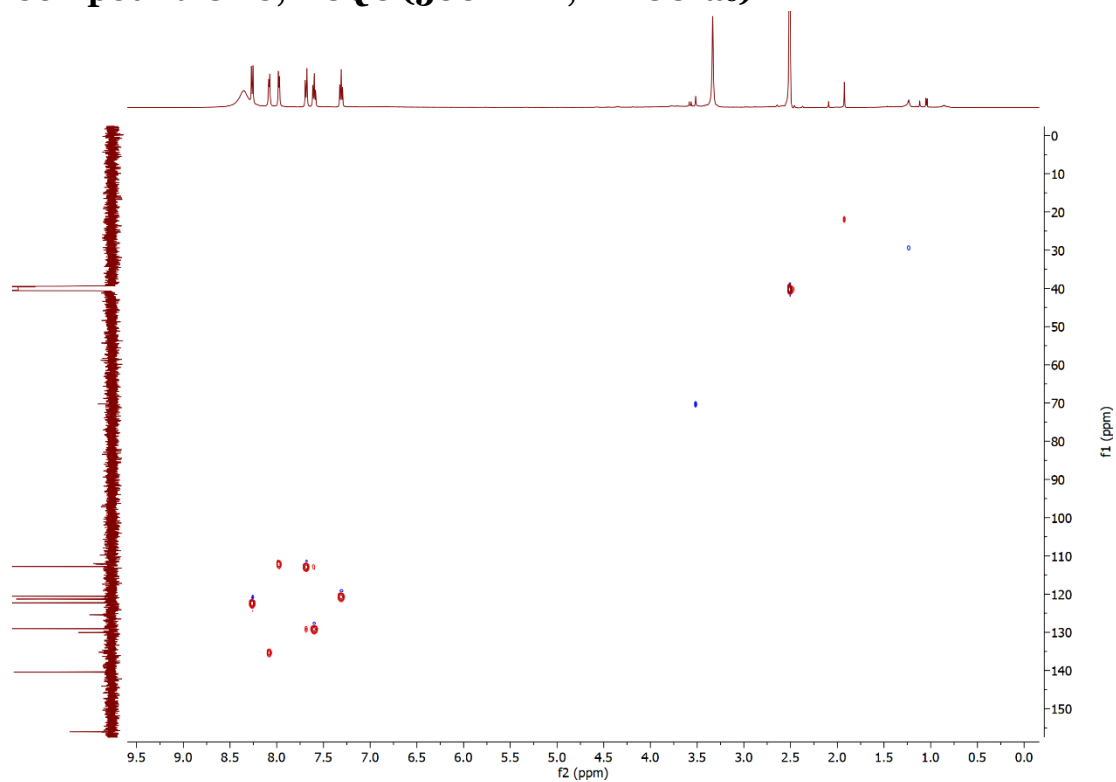
### Compound GAC, <sup>13</sup>C NMR (500 MHz, DMSO-d<sub>6</sub>)



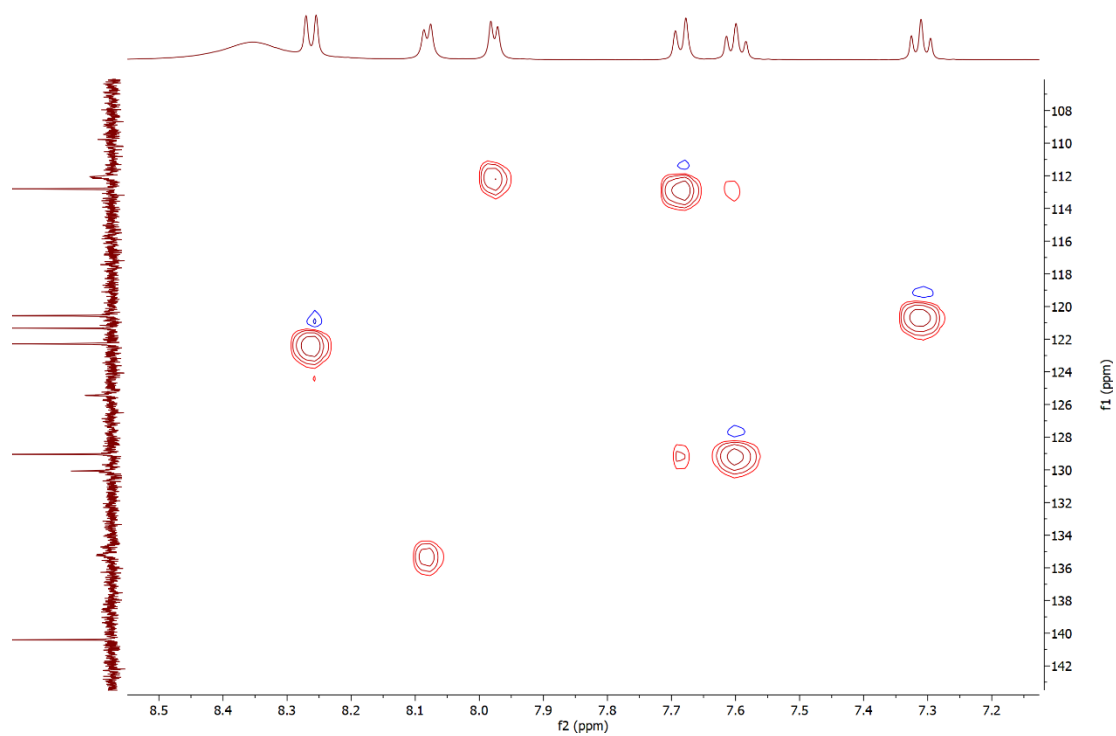
## Compound GAC, <sup>13</sup>C NMR Zoom (600 MHz, DMSO-*d*<sub>6</sub>)



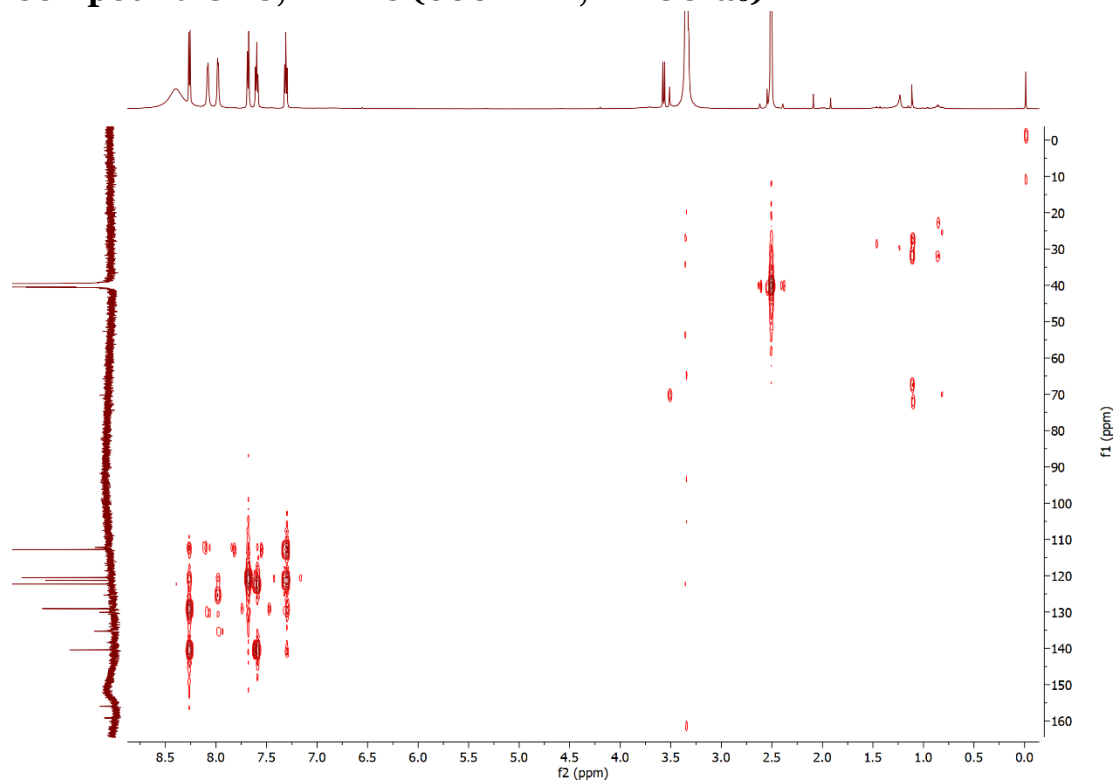
## Compound GAC, HSQC (500 MHz, DMSO-*d*<sub>6</sub>)



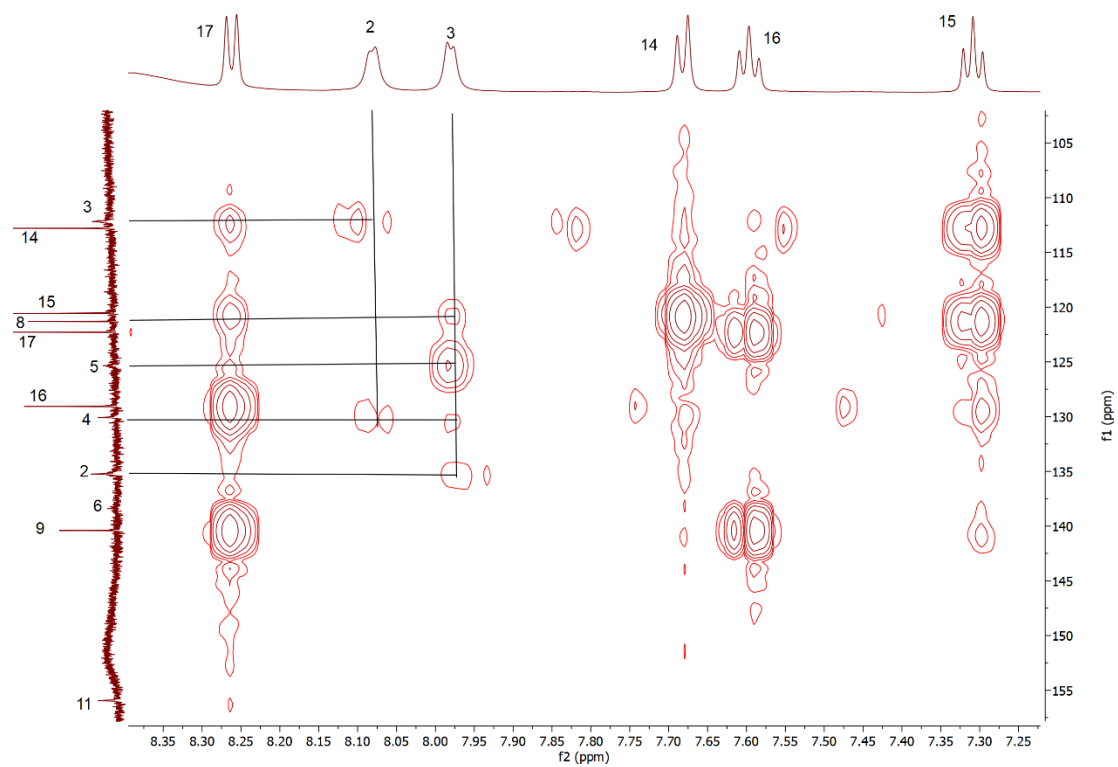
### Compound GAC, HSQC Zoom (500 MHz, DMSO-*d*<sub>6</sub>)



### Compound GAC, HMBC (600 MHz, DMSO-*d*<sub>6</sub>)

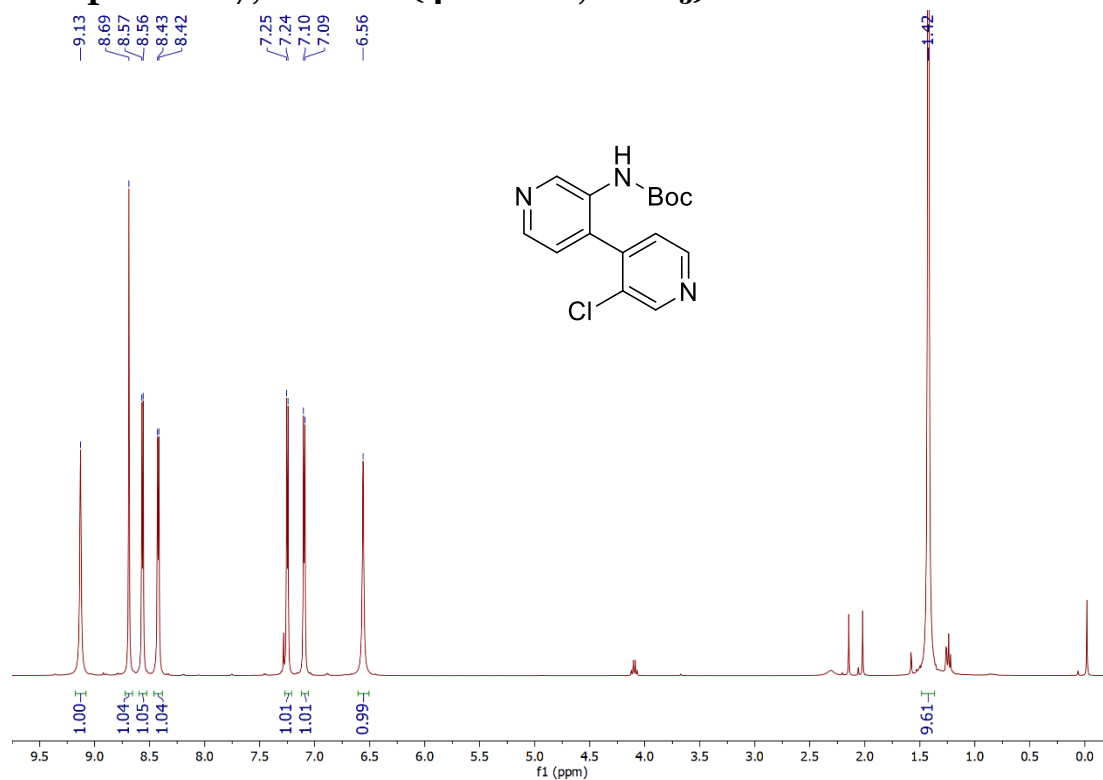


# Compound GAC, HMBC Zoom (600 MHz, DMSO-*d*<sub>6</sub>)

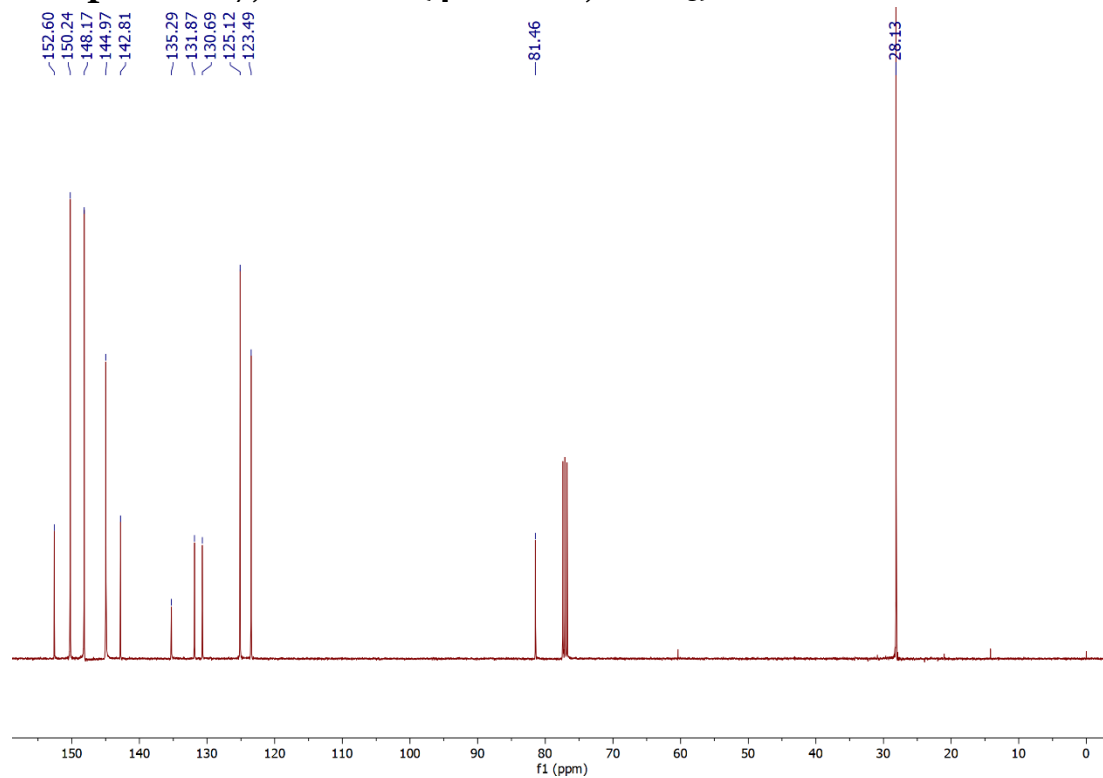




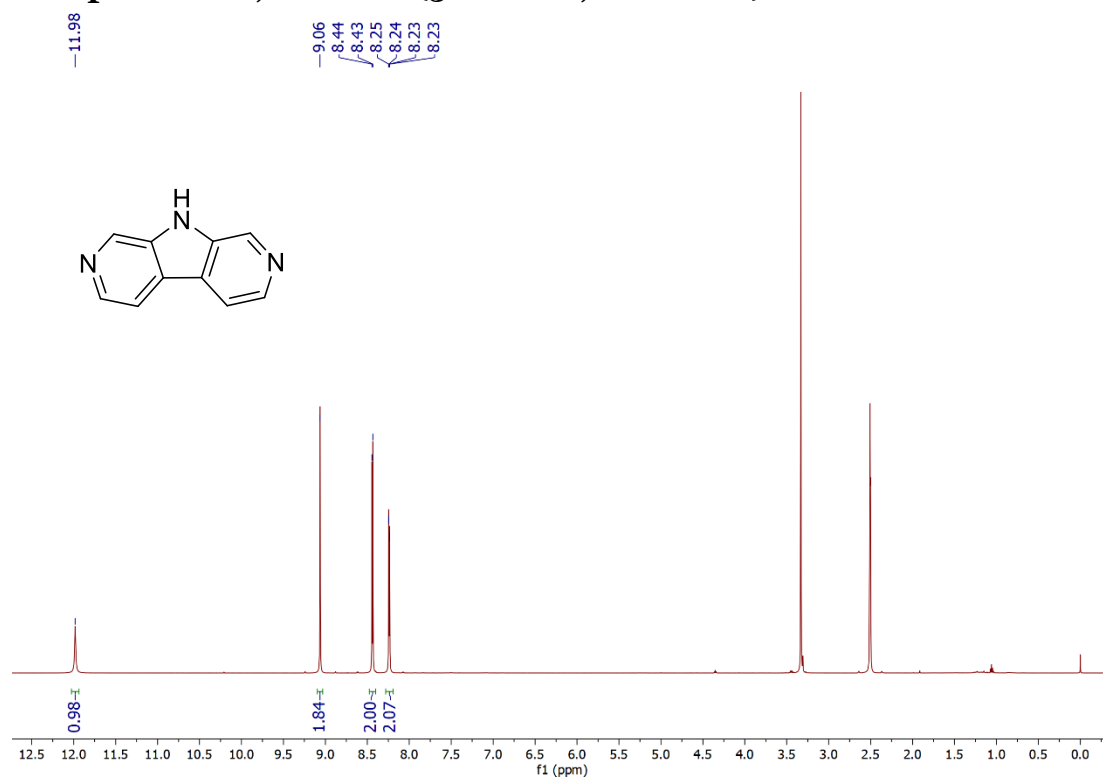
### Compound 2.7, <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)



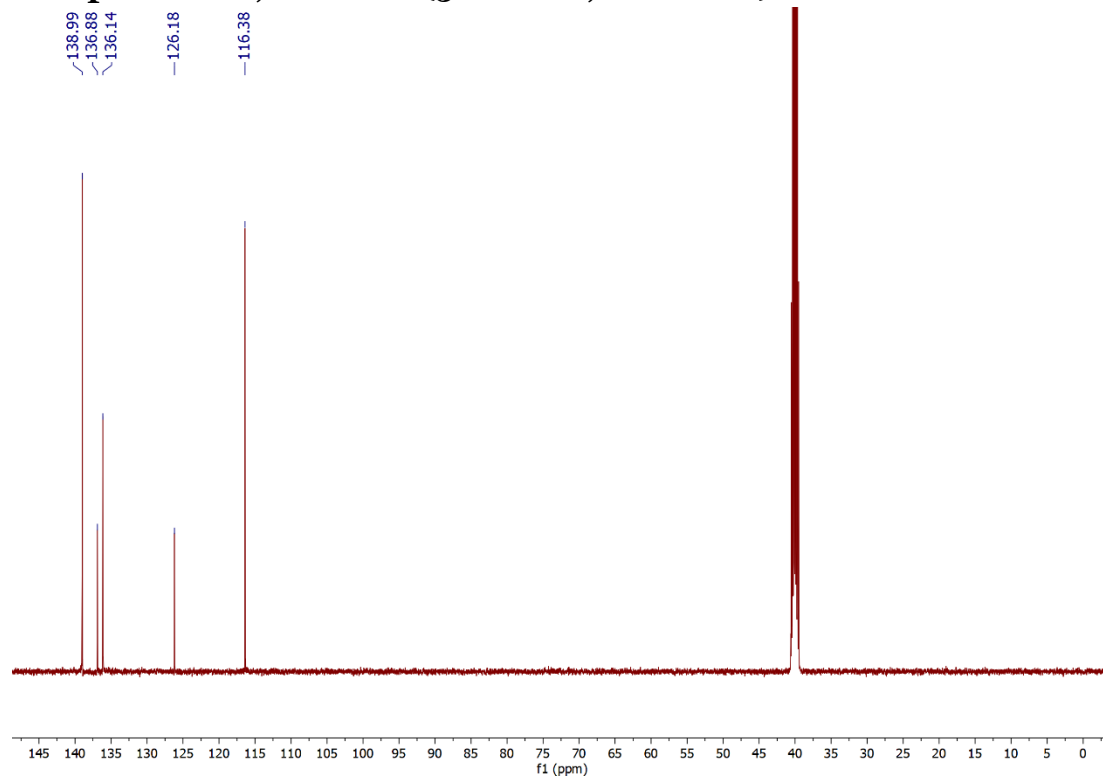
### Compound 2.7, <sup>13</sup>C NMR (400 MHz, CDCl<sub>3</sub>)



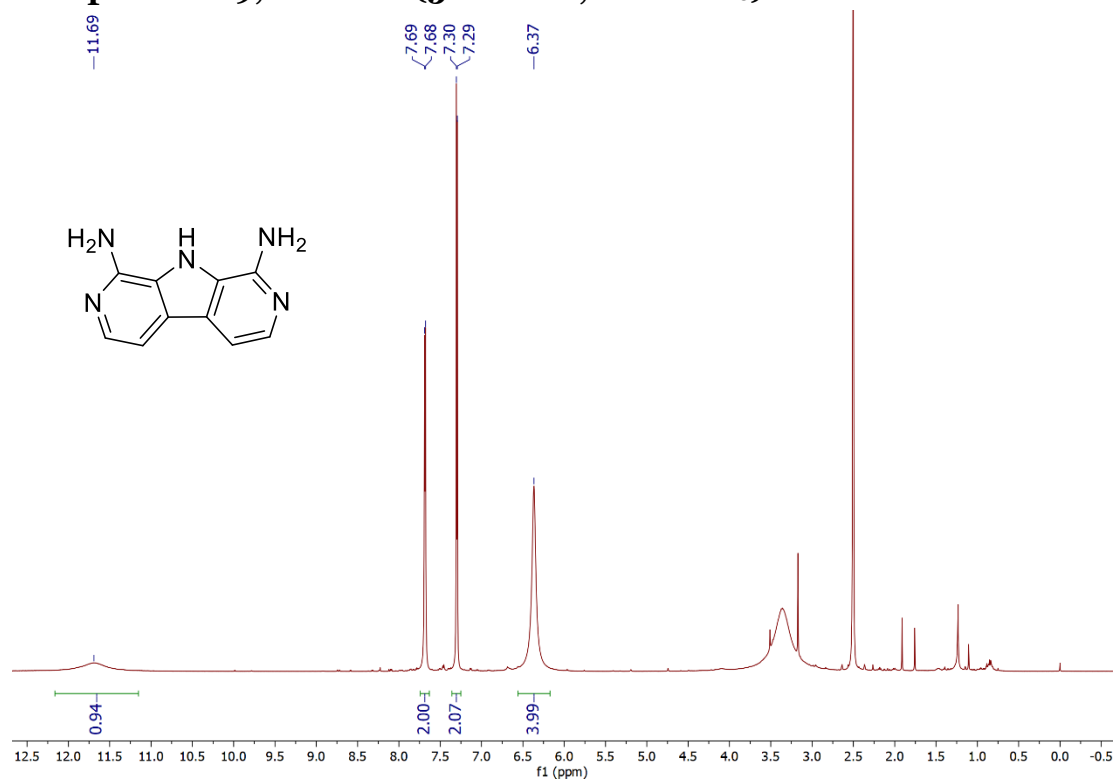
### Compound 2.8, <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)



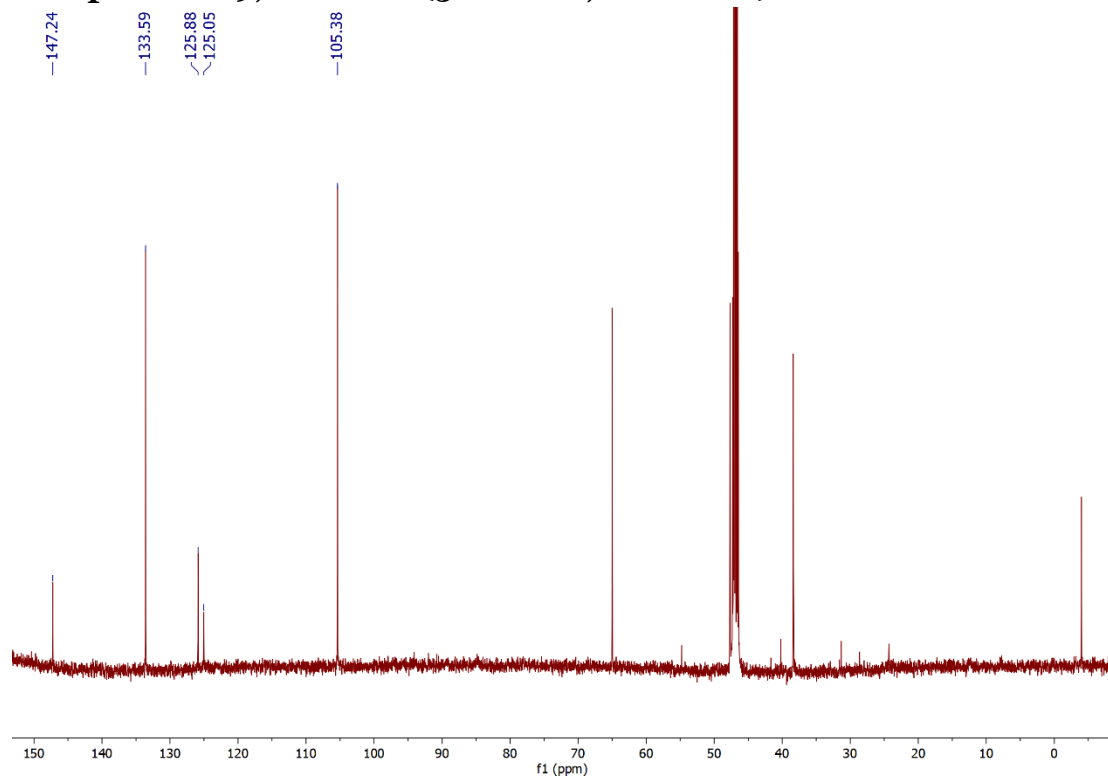
### Compound 2.8, <sup>13</sup>C NMR (500 MHz, DMSO-*d*<sub>6</sub>)



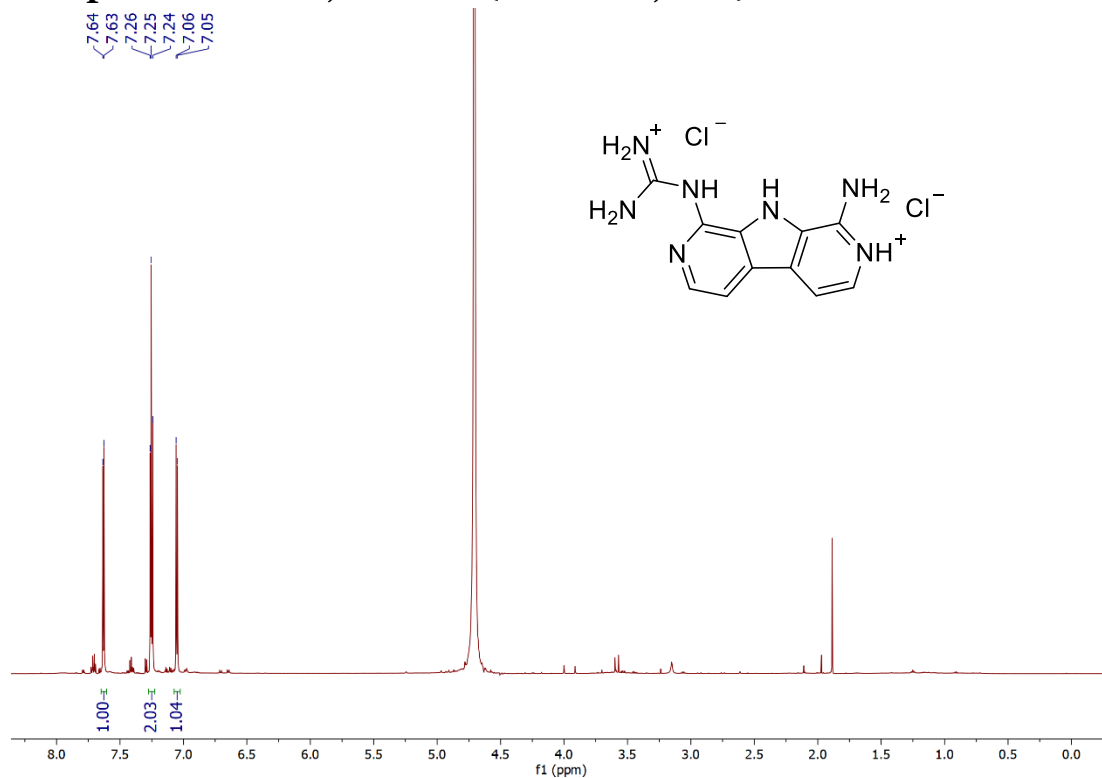
### Compound 2.9, <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)



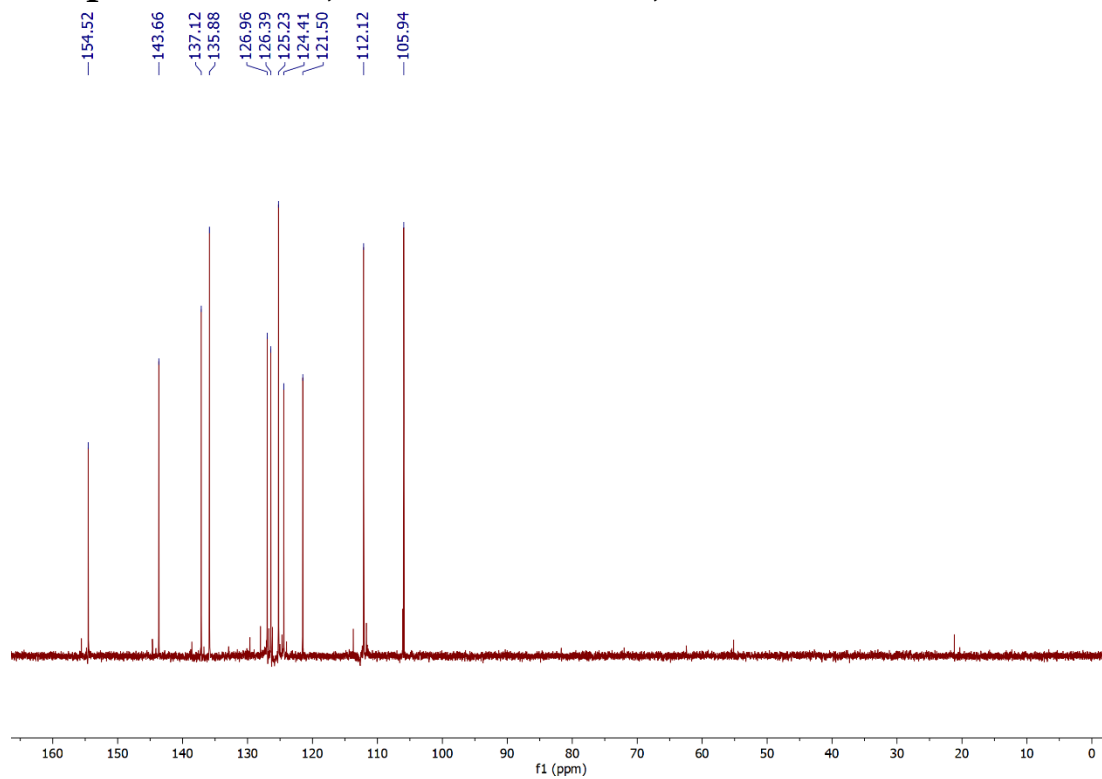
### Compound 2.9, <sup>13</sup>C NMR (500 MHz, DMSO-*d*<sub>6</sub>)



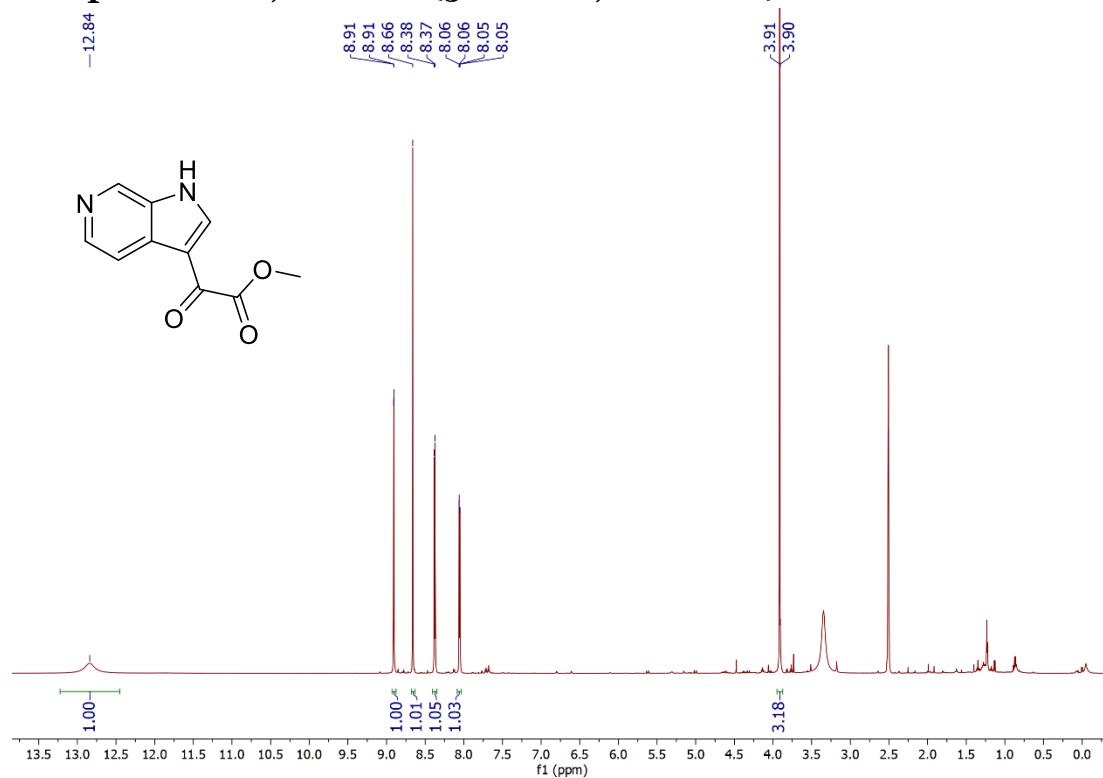
### Compound GADAC, <sup>1</sup>H NMR (600 MHz, D<sub>2</sub>O)



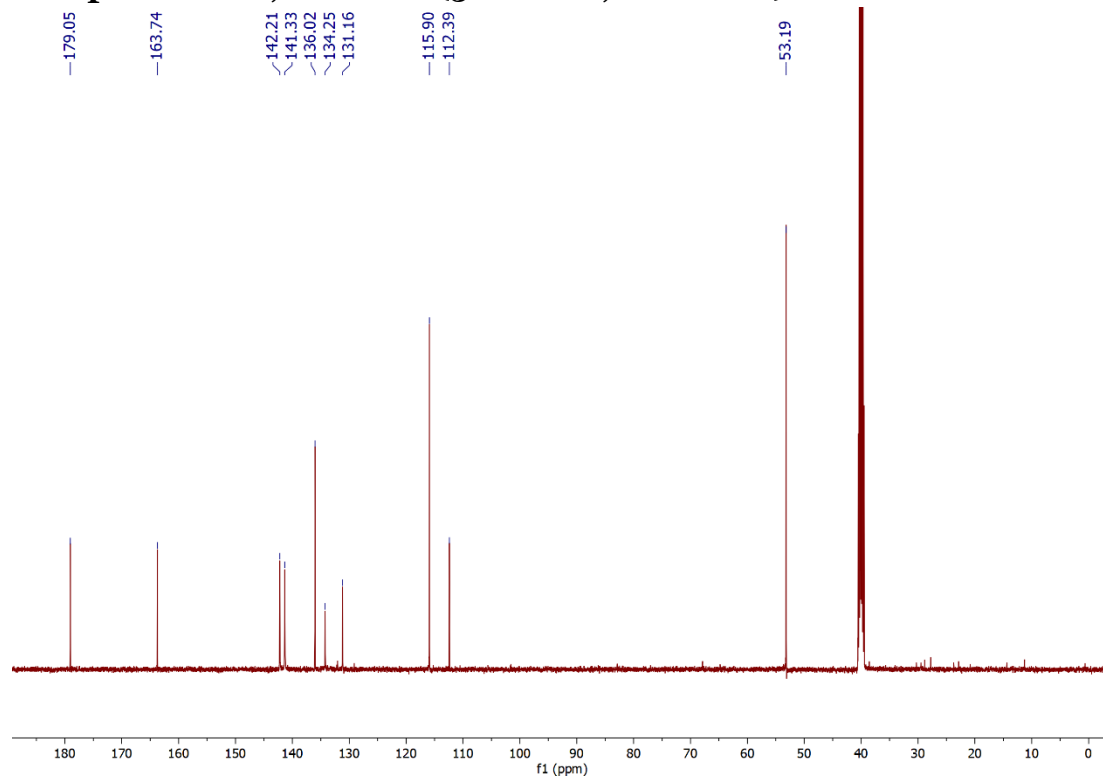
### Compound GADAC, <sup>13</sup>C NMR (600 MHz, D<sub>2</sub>O)



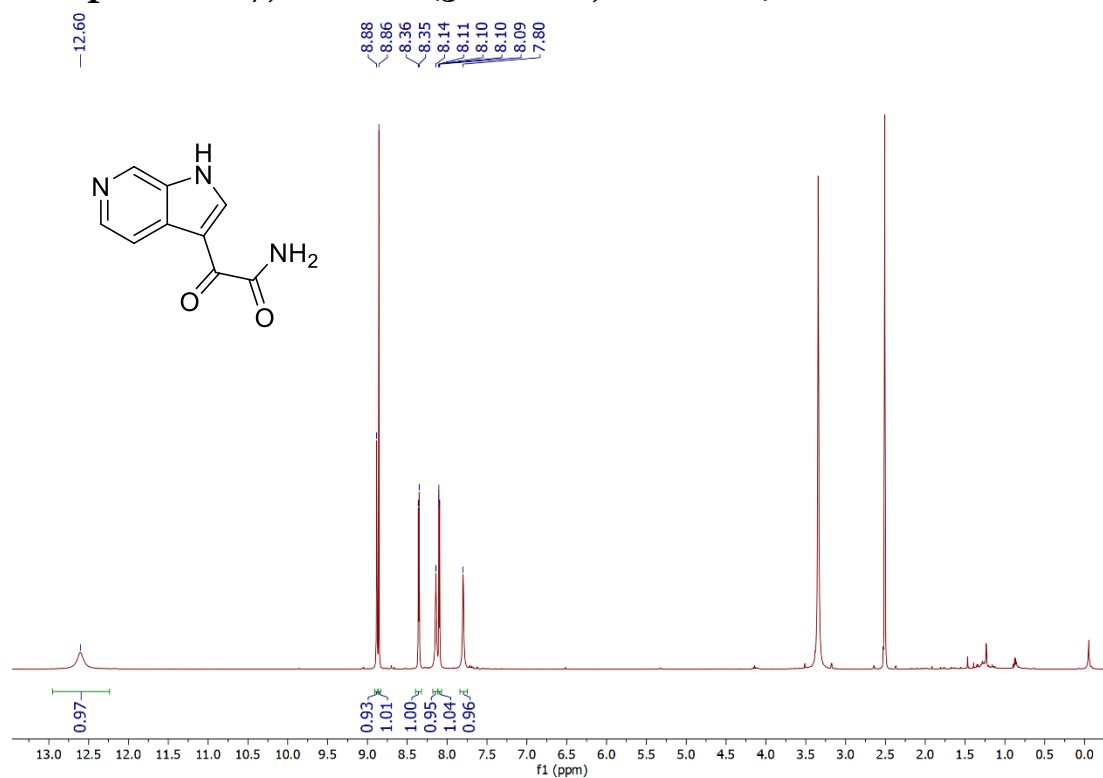
### Compound 2.16, <sup>1</sup>H NMR (500 MHz, DMSO-d<sub>6</sub>)



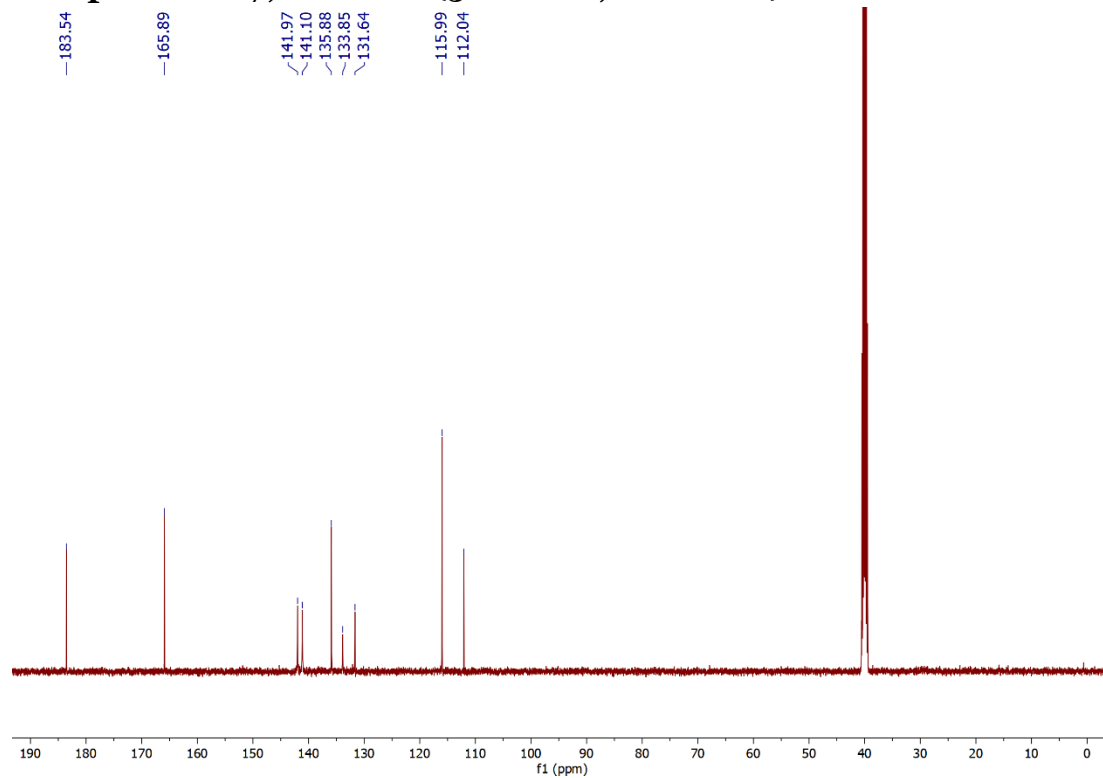
### Compound 2.16, <sup>13</sup>C NMR (500 MHz, DMSO-d<sub>6</sub>)



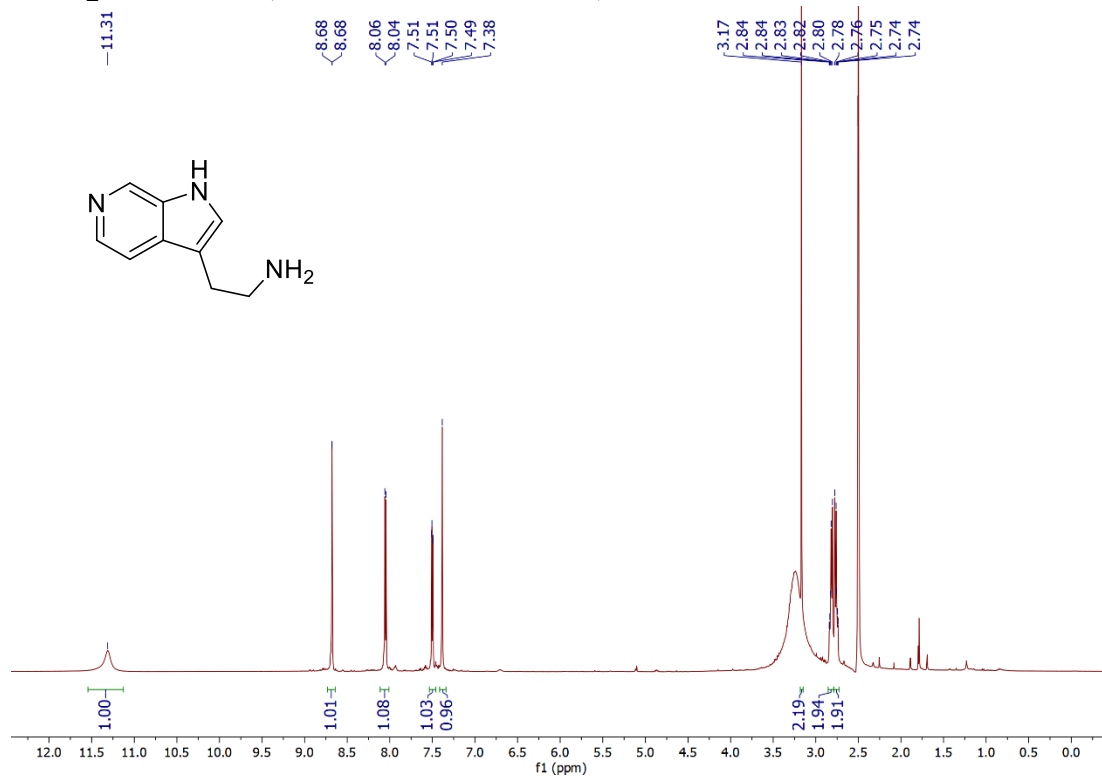
### Compound 2.17, <sup>1</sup>H NMR (500 MHz, DMSO-d<sub>6</sub>)



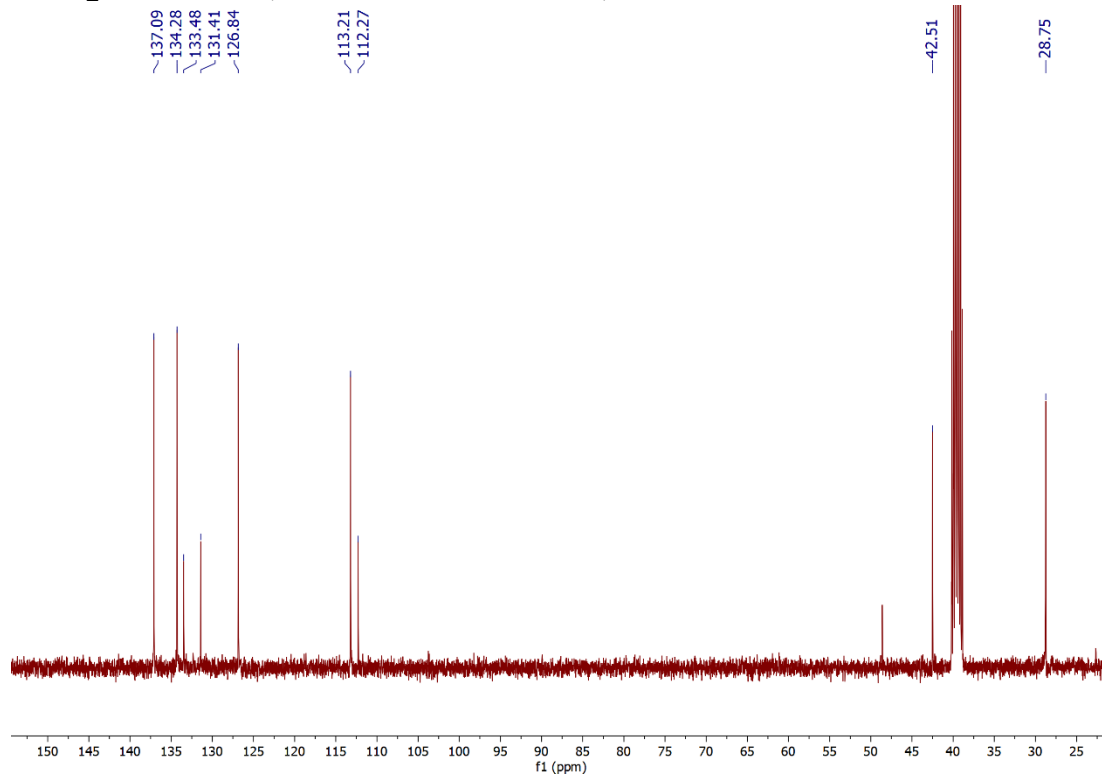
### Compound 2.17, <sup>13</sup>C NMR (500 MHz, DMSO-d<sub>6</sub>)



**Compound 2.18, <sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>)**



**Compound 2.18, <sup>13</sup>C NMR (500 MHz, DMSO-d<sub>6</sub>)**



## Chapter 3

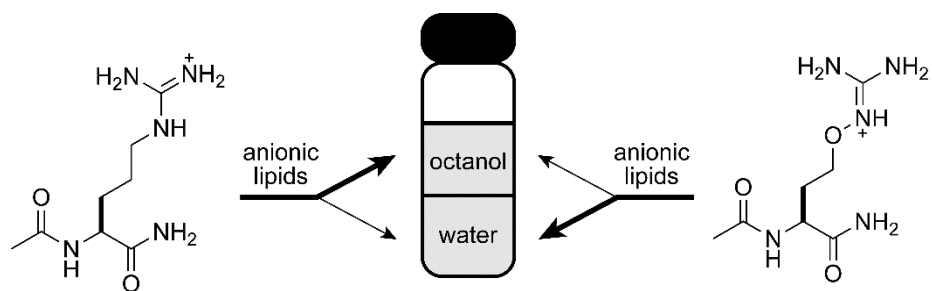
### **Canavanine versus arginine: Prospects for cell-penetrating peptides**

Reproduced from Calabretta, L. O., Thomas, V. M. & Raines, R. T. Canavanine versus arginine: Prospects for cell-penetrating peptides. *Tetrahedron Lett.*, **2022**, 99, 153848

Contributions: Synthesis of Ac-Cav-NH<sub>2</sub> was performed by Vienna M. Thomas and Lindsey O. Calabretta. All other work was performed by Lindsey O. Calabretta.



## Abstract



Octanol–water partitioning experiments in the presence of carboxylate-, phosphate-, and sulfate-containing anionic lipids revealed that Ac-Cav-NH<sub>2</sub> (where Cav refers to δ-oxa-arginine) partitions less into octanol than does Ac-Arg-NH<sub>2</sub>, suggesting that a cell-penetrating peptide based on canavanine would be relatively ineffective.

## Introduction

Since the discovery of HIV-tat peptide and penetratin, arginine-based cell-penetrating peptides (CPPs) have achieved broad use for transporting small molecules, proteins, nucleic acids, and nanoparticles into cells.<sup>82,85–88,90,129,130</sup> Wender, Rothbard, and coworkers were pioneers in this field, defining the role of peptide length, stereochemistry, and ability to form hydrogen bonds in cell-penetrating ability.<sup>15,58,131</sup> Subsequent approaches that deployed a variety of molecular architectures have extended the landscape.<sup>78,132–135</sup>

Recently, Schmuck and coworkers investigated the use of guanidiniocarbonyl-pyrroles (GCPs) in cell penetration and found that dimers of this moiety appended onto streptavidin induced uptake whereas dimers of arginine did not.<sup>79</sup> The GCP moiety has high affinity for oxoanions such as carboxylates, in part due to its low  $pK_a$  of  $\sim 7$ .<sup>30</sup> Guanidinium groups with lower  $pK_a$  values are likewise known to form stronger hydrogen bonds with oxoanions.<sup>41</sup> These findings inspired us to study a natural guanidinium group with a low  $pK_a$  value.

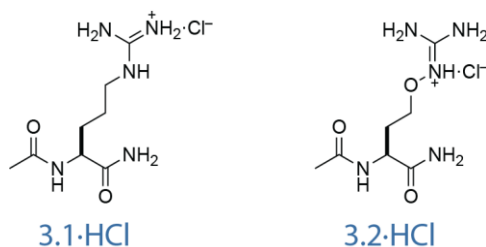
Canavanine (Cav), which is  $\delta$ -oxa-arginine, is a non-proteinogenic amino acid found in the seeds of leguminous plants. Herbivores are discouraged from consuming these seeds because the ribosomal misincorporation of canavanine residues in the place of arginine has deleterious consequences.<sup>136–139</sup> I was intrigued by canavanine because of an attribute that derives from its side-chain oxygen—a low guanidinium  $pK_a$ . The  $pK_a$  of the alkylguanidinium group of arginine is 13.8.<sup>29,140</sup> In contrast, the  $pK_a$  of the alkoxyguanidinium group of canavanine has been reported to be 7.01 and 7.40.<sup>140,141</sup> Arginine is the best of the canonical twenty amino-acid residues at facilitating the translocation of molecules into mammalian cells.<sup>58,85</sup> For two reasons, canavanine could

be better still. First, arginine is the most polar proteinogenic amino acid,<sup>142</sup> and the translocation of a cationic arginine residue across a nonpolar lipid bilayer is especially endergonic. In contrast, the cationic and neutral forms of canavanine have similar free energies at physiological pH, potentially facilitating membrane transversal. Secondly, stronger acids donate stronger hydrogen bonds.<sup>41,143–145</sup> Accordingly, the salt bridges formed by a cationic canavanine residue with cell-surface anionic groups could be stronger than those formed by a cationic arginine.

Wender, Rothbard, and coworkers demonstrated that octanol–water partitioning can report on desirable attributes of a CPP.<sup>58,131</sup> For example, they observed that fluorophore-labeled Arg<sub>8</sub> was transported into the octanol layer upon binding to an amphiphilic lipid, dodecanoate. In contrast, an 8-mer of ornithine (Orn) was less capable of binding to dodecanoate and was retained in the water layer. Because Arg<sub>8</sub> enters cells whereas Orns does not, an octanol–water partitioning experiment can serve as a proxy for determining cell-penetrating ability.<sup>58,131</sup>

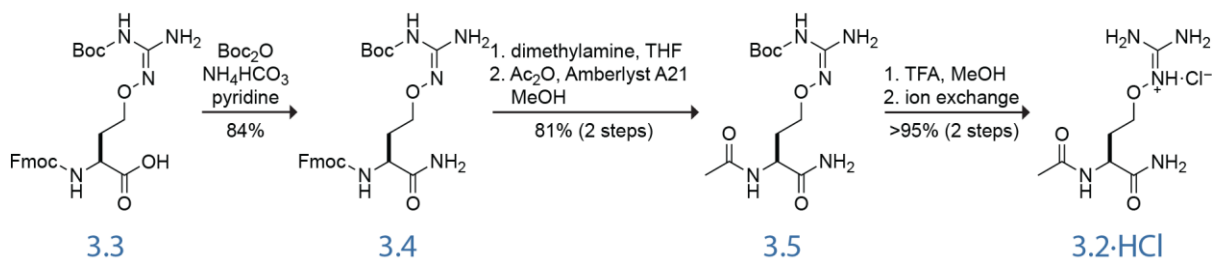
## Results and Discussion

How do the prospects of canavanine as a CPP compare to those of arginine? To answer this question, I sought to measure the partitioning of arginine and canavanine residues in the presence of carboxylate-, phosphate-, and sulfate-containing anionic lipids. To replicate the environment within a peptide, I amidated each amino acid on its N and C termini (Scheme 3.1). Amidated arginine Ac-Arg-NH<sub>2</sub> (**1**) was obtained from a commercial vendor as an acetic acid salt. Ac-Cav-NH<sub>2</sub> (**2**) was accessed by synthesis (Scheme 3.2).



**Scheme 3.1.** Structures of Ac-Arg-NH<sub>2</sub>·HCl (**3.1·HCl**) and Ac-Cav-NH<sub>2</sub>·HCl (**3.2·HCl**).

To synthesize Ac-Cav-NH<sub>2</sub>, the carboxylate of commercial Fmoc-Cav(Boc)-OH (**3.3**) was amidated by exposure to Boc<sub>2</sub>O and ammonium bicarbonate in pyridine to produce Fmoc-Cav(Boc)-NH<sub>2</sub> (**3.4**). Notably, using traditional solid-phase methods to produce this amide (*e.g.*, loading onto a Rink amide resin with PyBOP, *N*-acylation with acetic anhydride, and cleavage with TFA and TIPS) were unsuccessful. Subsequently, the Fmoc protecting group was removed in a THF solution of dimethylamine, which was easier to separate via evaporation than the traditional piperidine. Acetylation with acetic anhydride in the presence of the basic resin Amberlyst A-21 produced Ac-Cav(Boc)-NH<sub>2</sub> (**3.5**). The Boc group was removed by TFA to produce Ac-Cav-NH<sub>2</sub> (**3.2**) as its trifluoroacetic acid salt.



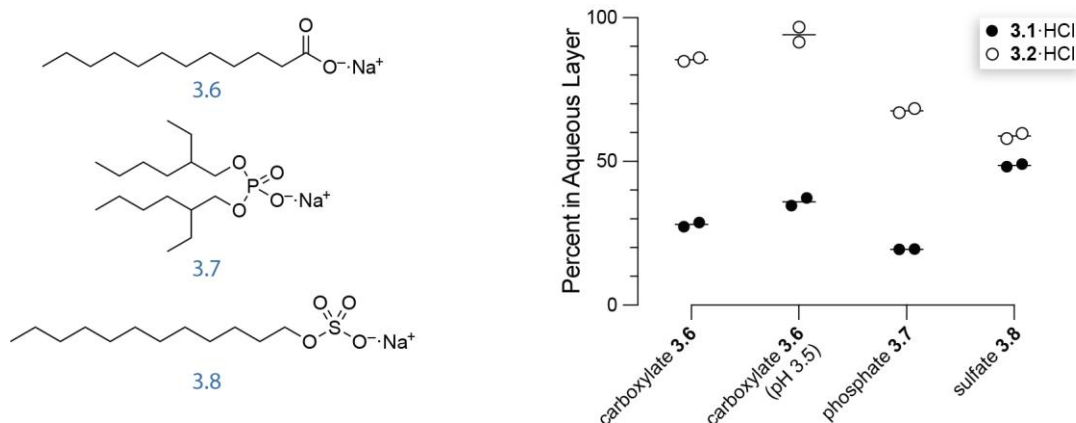
**Scheme 3.2.** Synthetic route to Ac-Cav-NH<sub>2</sub>·HCl (**3.2·HCl**).

A rigorous comparison of the ability of Ac-Arg-NH<sub>2</sub> (**3.1**) and Ac-Cav-NH<sub>2</sub> (**3.2**) to bind to oxoanions requires that both residues contain the same counterion. To remove the strongly associated acetate and trifluoroacetate counterions, we exposed both **3.1·HOAc** and **3.2·TFA** to excess HCl(aq) and lyophilization. Under these acidic conditions, we observed significant hydrolysis of the C-terminal amide in both residues. To avoid this decomposition, I synthesized a guanidine resin from a commercial 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) resin, charged the guanidine resin with HCl, and used the resulting guanidinium chloride resin to perform ion-exchange. This procedure was successful in replacing the oxoanions with chloride ions, as evidenced by the disappearance of the HOAc protons and TFA fluoro groups by <sup>1</sup>H-NMR and <sup>19</sup>F-NMR, respectively.

To compare the abilities of canavanine and arginine to bind to oxoanions, **3.1·HCl** and **3.2·HCl** were mixed with three oxoanion lipids that served as models for the functional groups found in membrane phospholipids and cell-surface glycans. Briefly, **3.1·HCl** and **3.2·HCl** were dissolved in D<sub>2</sub>O (pD 7.0; pH 7.4) and washed with octanol. An aliquot of the D<sub>2</sub>O layer was carefully excised, and the **3.1** and **3.2** content was analyzed by <sup>1</sup>H-NMR spectroscopy using an added standard (Figure 3.2). This procedure enabled us to quantify the octanol–water partitioning without installing a pendant fluorophore or other label, which could be perturbative.

We observed that all of the **3.1·HCl** and **3.2·HCl** remained in the aqueous layer after partitioning with octanol only. Next, I added sodium dodecanoate (**3.6**), bis(2-ethylhexyl) phosphate (**3.7**), or dodecylsulfate (**3.8**) to the octanol wash with the expectation that these oxoanions could bind to the guanidinium groups and, due to their amphipathic nature, draw **3.1·HCl** and **3.2·HCl** into the octanol layer.<sup>131</sup> For each combination of

amino acid and anionic lipid, I did indeed observe substantial partitioning of the amino acid into the octanol layer (Figure 3.1). I had hypothesized that canavanine (**3.2**), with its significantly lower  $pK_a$  value, would partition more than arginine (**3.1**) into the octanol layer. Surprisingly, with each lipid, *less* arginine than canavanine remained in the aqueous layer.



**Figure 3.1.** Graph showing the octanol–water partitioning of **3.1·HCl** and **3.2·HCl** in the presence of anionic lipids **3.6**, **3.7**, or **3.8** (2.5 equiv<sup>131</sup>) at pH 7.4 (unless indicated otherwise). Values were determined by <sup>1</sup>H-NMR spectroscopy in duplicate experiments (Figures 3.2–3.4).

Finally, I determined whether the lesser ability of canavanine to partition into octanol was due to its lower level of protonation at pH 7.4. To do so, I measured the octanol–water partitioning of arginine (**3.1·HCl**) and canavanine (**3.2·HCl**) at pH 3.5 with carboxylate **3.6**. I found that the partitioning of canavanine into octanol did not increase at low pH (Figure 3.1), suggesting that fully protonated canavanine was a less effective transporter than fully protonated arginine.

Why are anionic lipids relatively ineffective at pulling canavanine into octanol? One reason could be the location of its cationic charge, which resides largely on the bridging

$N^{\epsilon}$ -H group rather than the two terminal  $N^{\eta}$ -H groups (Figure 3.5). That location could engender steric hindrance in interactions with a carboxylate, phosphate diester, or sulfate monoester. In addition, the two lone pairs on the proximal  $\delta$  oxygen of canavanine could repel the oxygens of the anionic groups, weakening hydrogen bonding. Later, I determined that the topological polar surface area (TPSA) of **3.2**· $H^+$  is greater than **3.1**· $H^+$  (Table 4.1). TPSA appears to significantly influence the anion-mediated partitioning of amino acids, since  $N_{\alpha}$ -methylated arginine derivatives have lower TPSAs and are partitioned more into the octanol layer than non-methylated arginine (Chapter 4).

## Conclusions

We sought to assess the prospects of canavanine (**3.2**) in comparison to arginine (**3.1**) as an effective CPP. Although oxoanionic lipids draw both canavanine and arginine into the octanol layer during octanol–water partitioning, canavanine partitions significantly less extensively into octanol than does arginine. These data suggest that canavanine-based CPPs would be less capable of binding to cell-surface anions and mediating cell penetration than traditional arginine-based CPPs.

## Materials and Methods

### General

*Materials.* Commercial compounds were from Sigma–Aldrich (St. Louis, MO), Chem Impex (Wood Dale, IL), or Iris Biotech GmbH (Marktredwitz, Germany) and were used without further purification. Bis(2-ethylhexyl) hydrogen phosphate was converted to its sodium salt by adding NaOH (1 equiv) to an aqueous solution of the phosphate followed by lyophilization. Ac-Arg-NH<sub>2</sub>·HOAc (**3.1·HOAc**) was converted to Ac-Arg-NH<sub>2</sub>·HCl (**3.1·HCl**) by cation ion-exchange chromatography using the procedure that produced Ac-Cav-NH<sub>2</sub>·HCl (**3.2·HCl**), *vide infra*.

*Conditions.* All procedures were performed in air at ambient temperature (~22 °C) and pressure (1.0 atm) unless specified otherwise.

*Solvent removal.* The phrase “concentrated under reduced pressure” refers to the removal of solvents and other volatile materials using a rotary evaporator while maintaining a water-bath temperature at 40 °C. Residual solvent was removed from samples at high vacuum (<0.1 Torr), which refers to the vacuum achieved by a mechanical belt-drive oil pump, or through lyophilization (freeze-drying) using a Labconco FreeZone freeze dryer.

*Chromatography.* Chemical reactions were monitored by thin-layer chromatography (TLC) using EMD 250 μm silica gel 60-F<sub>254</sub> plates and visualization with UV-illumination or KMnO<sub>4</sub>-staining, or by LC–MS with an ESI Agilent 6125B mass spectrometer. Flash chromatography was performed with a Biotage Isolera automated purification system using prepacked and re-packed SNAP KP silica gel columns.

*Instrumentation.* <sup>1</sup>H-NMR and <sup>13</sup>C-NMR spectra for compound characterization were obtained with Bruker spectrometers, and HRMS data were obtained with an Agilent 6545



Q-ToF mass spectrometer at the Department of Chemistry Instrumentation Facility at the Massachusetts Institute of Technology.

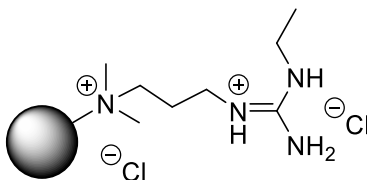
## **Octanol–water partitioning experiments**

Ac-Cav-NH<sub>2</sub>·HCl (5 mg, 0.02 mmol) was placed in four separate vials and dissolved in D<sub>2</sub>O (1.0 mL) to make 0.02 M solutions. The pH of the solutions was adjusted to 7.4 (pD 7.0) or 3.5 (pD 3.1) by the addition of highly concentrated NaOD and DCl solutions in 1- $\mu$ L increments. A 200- $\mu$ L aliquot was taken from each vial for NMR analysis. Subsequently, sodium bis(2-ethylhexyl) phosphate (13.8 mg, 0.04 mmol, 2.5 equiv of Ac-Cav-NH<sub>2</sub>·HCl) was added to two vials, and then 800  $\mu$ L of octanol was added to each vial. The vials were shaken vigorously for 5 min, and then subjected to centrifugation for 10–20 min. Once the layers had separated, a 200- $\mu$ L aliquot was removed carefully from the aqueous layers with a pipette and added to a new tube. A solution of 0.2 M pyridine in D<sub>2</sub>O (20  $\mu$ L) was added to each tube, the tubes were shaken, and the solutions were transferred to 3-mm NMR tubes for analysis. This procedure was repeated with sodium dodecylsulfate and sodium dodecanoate, as well as with Ac-Arg-NH<sub>2</sub>·HCl and each anionic lipid.

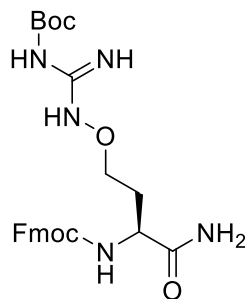
<sup>1</sup>H-NMR spectra were collected of the aliquots. Spectral integrations were referenced to the signal of the *para*-hydrogen of pyridine, and the integration of the signal from the H <sup>$\alpha$</sup>  proton of the amino acid residue was measured. The relative integrations pre- and post-wash were compared to each other. For both amino acid residues, the spectra following the octanol washes showed a higher integration. This increase was likely the result of an increase in the residual water signal from octanol protons exchanging with

D<sub>2</sub>O and influencing the integration of the H<sup>α</sup> proton by altering the local baseline (Figure 3.2). Because the amount of additional water in the post-octanol and octanol + lipid washes is likely to be the same, I evaluated the extent of partitioning by determining the difference between the octanol and octanol + lipid washes and averaging the values from two replicates.

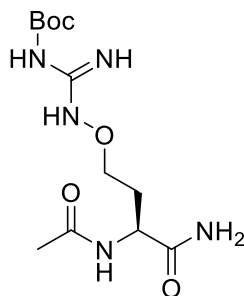
## Synthesis



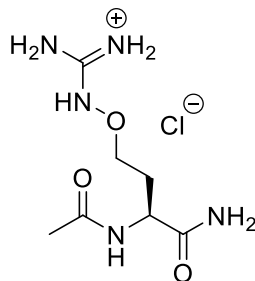
**Dimethylaminopropyl-ethylguanidine resin.** Polymer-bound 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide (EDC), 200–400 mesh, ~1.4 mmol/g loading (1.0 g) was suspended in a solution of 7 N NH<sub>4</sub>Cl in MeOH (6 mL). A solution of 4 M HCl in dioxane (375 μL, 1.5 mmol) was added, and the resulting suspension was stirred. As a solution-phase surrogate, EDC (267.4 mg, 1.4 mmol) was dissolved concurrently in 7 N NH<sub>4</sub>Cl in MeOH (6 mL). A solution of 4 M HCl in dioxane (375 μL, 1.5 mmol) was added, and the resulting solution was stirred. The progress of the latter reaction was monitored by LC–MS. When the signal for the guanidine product had appeared and the signal for the starting material had disappeared, the resin beads were filtered and dried (and the solution-phase reaction mixture was discarded). The resin was packed into a pipette column and rinsed with water (3×), 1 M HCl (3×), and water again (3×).



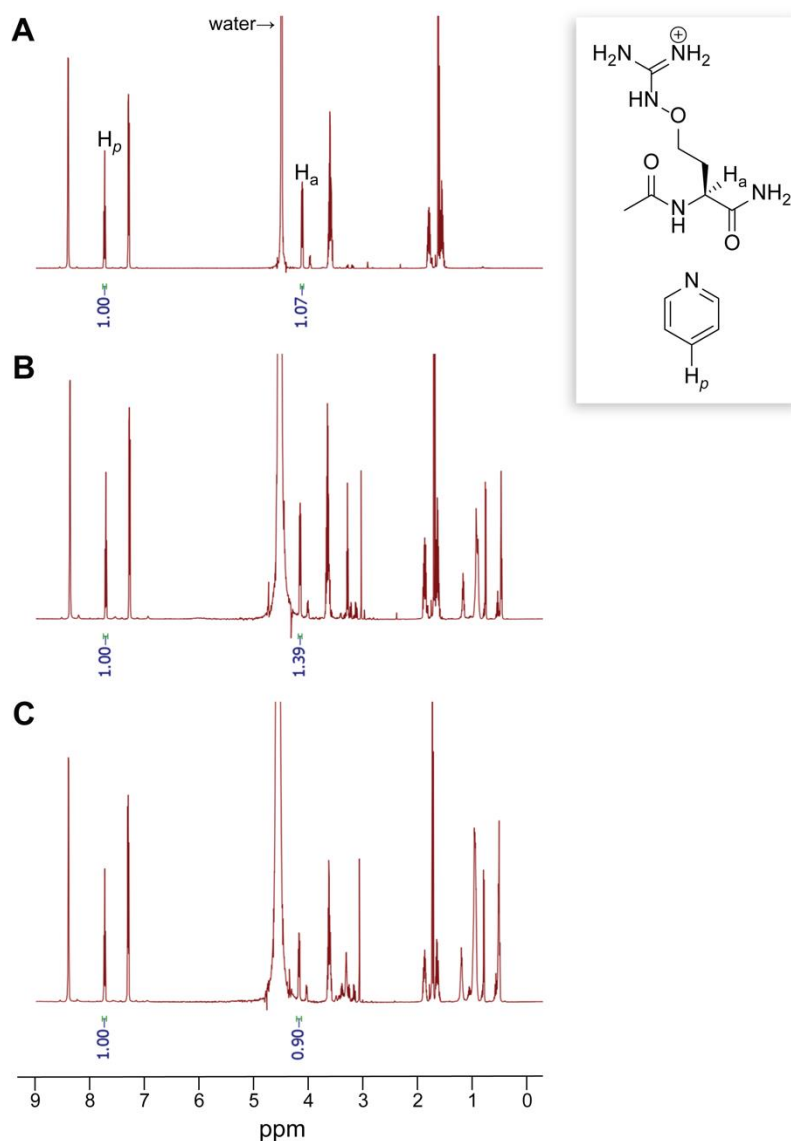
**Fmoc-Cav(Boc)-NH<sub>2</sub> (3.4).** Ammonium bicarbonate (277.66 mg, 3.51 mmol) and Boc<sub>2</sub>O (152.81 mg, 0.7 mmol) were dissolved in pyridine (2.5 mL, 0.20 mol). Fmoc-Cav(Boc)-OH (**3.3**) (250 mg, 0.5 mmol) was added, and the mixture was stirred for 6 h. The reaction mixture was left to dry under a stream of air overnight. The product was purified by column chromatography on silica gel with a gradient of 1–5% v/v MeOH in DCM to produce a white fluffy solid (209 mg, 84% yield). <sup>1</sup>H NMR (600 MHz, MeOD, δ): 7.81 (d, *J* = 7.5 Hz, 2H), 7.68 (t, *J* = 6.4 Hz, 2H), 7.40 (t, *J* = 7.5 Hz, 2H), 7.33 (t, *J* = 7.3 Hz, 2H), 4.45–4.35 (m, 2H), 4.30 (dd, *J* = 9.4, 4.8 Hz, 1H), 4.24 (t, *J* = 6.6 Hz, 1H), 3.89 (m, 2H), 2.18 (m, 1H), 1.91 (m, 1H), 1.48 (s, 9H). <sup>13</sup>C NMR (600 MHz, MeOD, δ): 176.16, 157.09, 153.81, 143.91, 143.84, 141.20, 127.38, 126.79, 124.82, 119.51, 80.90, 68.81, 66.57, 52.26, 47.04, 31.12, 27.05. HRMS *m/z* calcd for C<sub>25</sub>H<sub>31</sub>N<sub>5</sub>O<sub>6</sub> [M + H]<sup>+</sup>, 498.2353; found 498.2345.



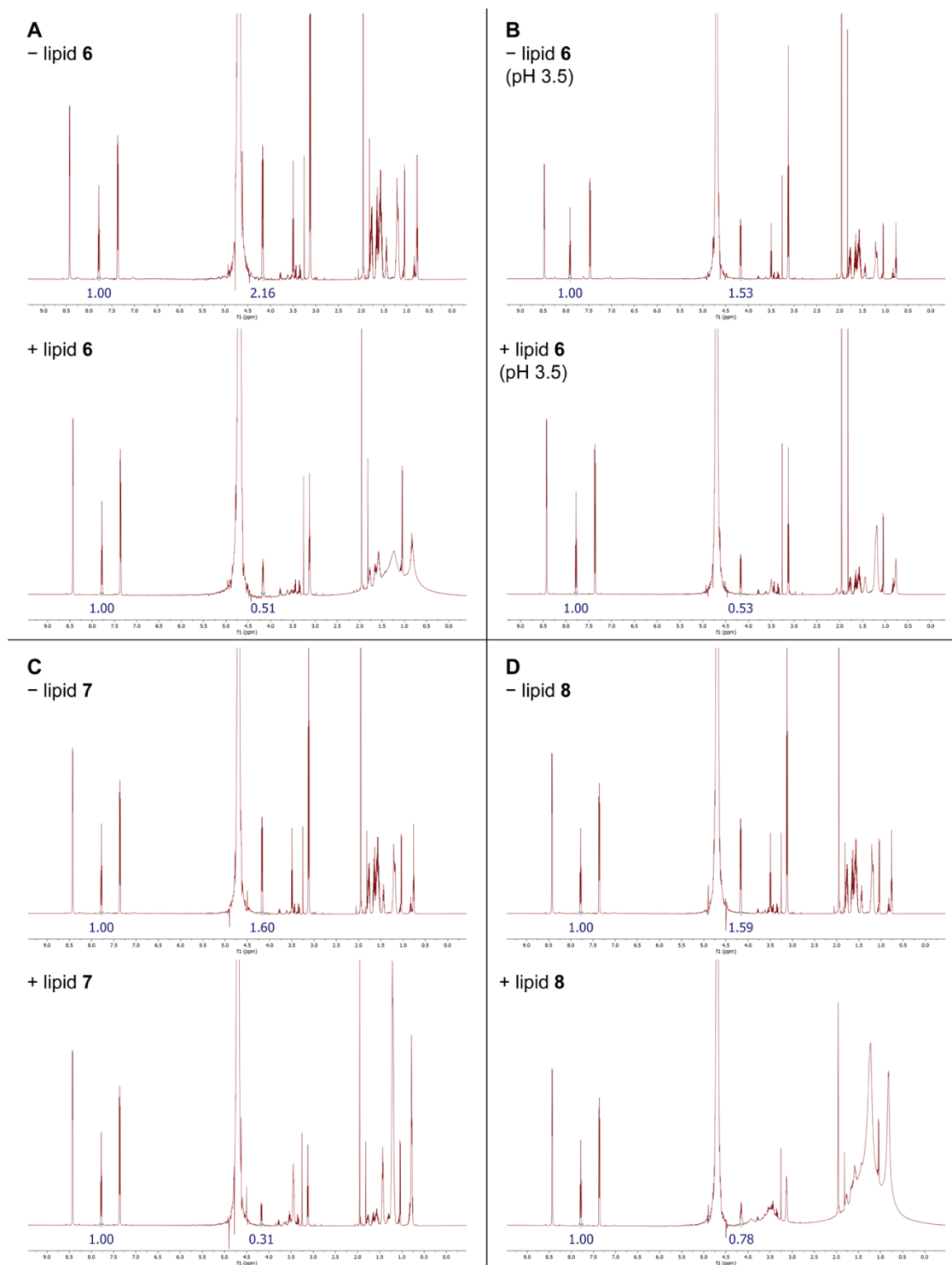
**Ac-Cav(Boc)-NH<sub>2</sub> (3.5).** Fmoc-Cav(Boc)-NH<sub>2</sub> (**3.4**) (280 mg, 0.56 mmol) was dissolved in 2.0 M dimethylamine in THF (3 mL), and the resulting solution was stirred for 2 h. The mixture was concentrated under reduced pressure before being suspended in water and vacuum-filtered to remove Fmoc byproducts. The filtrate was lyophilized to yield a white solid. The mixture was then dissolved in MeOH (5.5 mL), and Amberlyst-A21 tertiary amine resin (1 g) and acetic anhydride (529  $\mu$ L, 5.6 mmol) were added. The mixture was stirred for 3 h, filtered, concentrated under reduced pressure, and purified by column chromatography on silica gel in 20% v/v MeOH in DCM to produce a white foam (144 mg, 81% yield). <sup>1</sup>H NMR (600 MHz, MeOD,  $\delta$ ): 4.51 (dd,  $J$  = 8.9, 5.1 Hz, 1H), 3.93 (m, 2H), 2.19 (ddt,  $J$  = 10.7, 7.8, 5.2 Hz, 1H), 2.01 (s, 3H), 1.93 (ddt,  $J$  = 14.4, 9.0, 5.4 Hz, 1H), 1.50 (s, 9H). <sup>13</sup>C NMR (600 MHz, MeOD,  $\delta$ ): 176.81, 173.47, 154.27, 153.96, 83.63, 71.72, 51.70, 32.10, 28.32, 22.58. HRMS  $m/z$  calcd for C<sub>12</sub>H<sub>23</sub>N<sub>5</sub>O<sub>5</sub> [M + H]<sup>+</sup>, 318.1777; found, 318.1771.



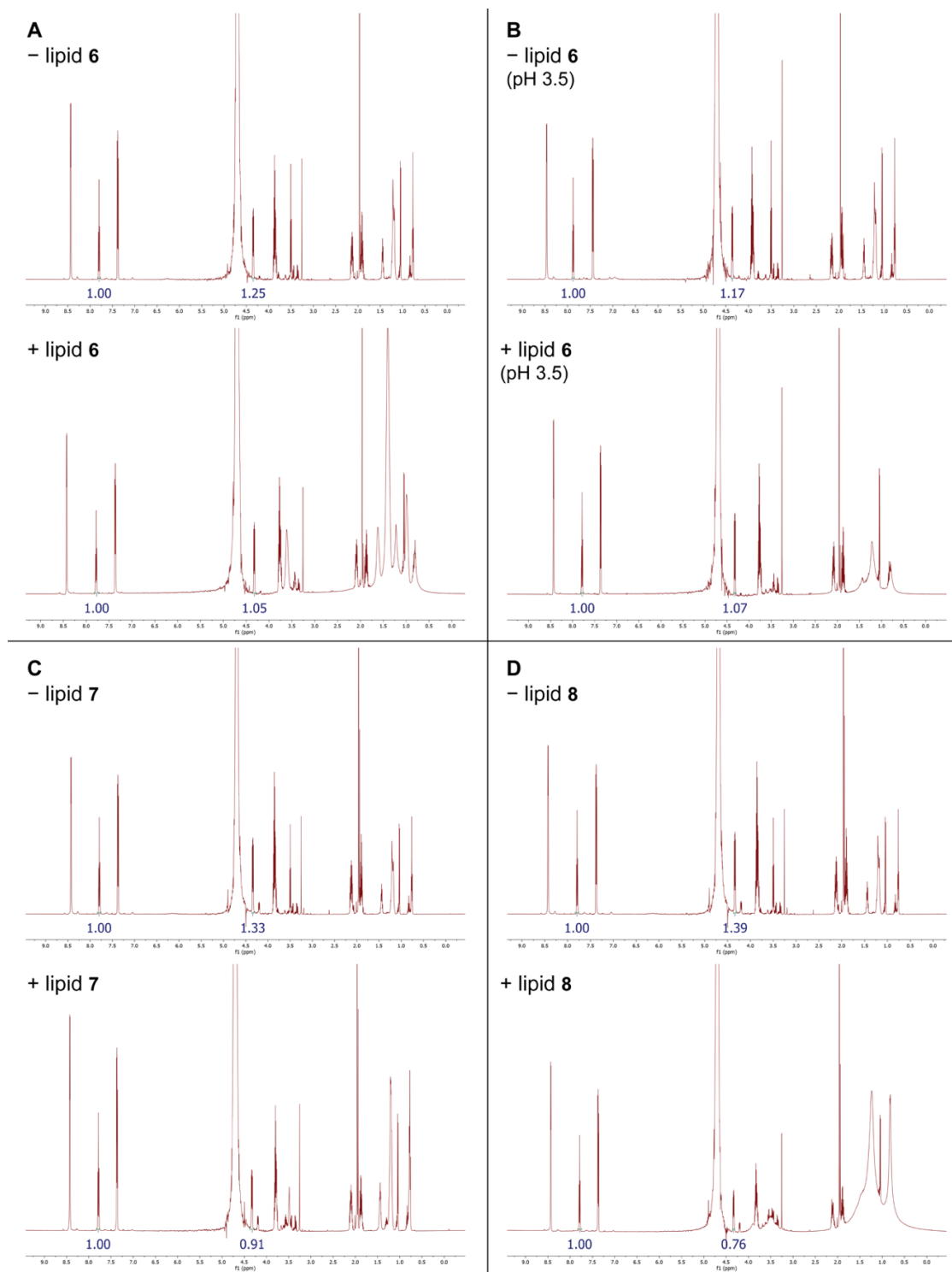
**Ac-Cav-NH<sub>2</sub>·HCl (3.2·HCl):** Ac-Cav(Boc)-NH<sub>2</sub> (**3.5**) (144 mg, 0.45 mmol) was dissolved in a solution of TFA (1.0 mL) and MeOH (50  $\mu$ L), and the resulting solution was stirred for 1 h. The mixture was dried under a stream of air and then under reduced pressure. The residue was dissolved in water, and the resulting solution was lyophilized to remove excess TFA. The residue was redissolved in water, and the resulting solution was flushed through a pipette column of dimethylaminopropyl-ethylguanidine resin that had been charged with HCl and lyophilized to produce a clear solid (97 mg, >95% yield). **<sup>1</sup>H NMR** (600 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 11.16 (s, 1H), 8.20 (d, *J* = 8.1 Hz, 1H), 7.70 (s, 4H), 7.55 (s, 1H), 7.13 (s, 1H), 4.29 (td, *J* = 8.7, 5.0 Hz, 1H), 3.80 (t, *J* = 6.5 Hz, 3H), 2.03 (dtd, *J* = 14.1, 7.0, 4.9 Hz, 1H), 1.86 (s, 3H), 1.80 (ddt, *J* = 14.7, 9.3, 5.9 Hz, 1H). **<sup>13</sup>C NMR** (600 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 174.19, 170.51, 158.79, 73.62, 49.86, 30.56, 22.94. **HRMS** *m/z* calcd for C<sub>7</sub>H<sub>15</sub>N<sub>5</sub>O<sub>3</sub> [M + H]<sup>+</sup>, 218.1253; found, 218.1245.



**Figure 3.2.** Representative  $^1\text{H}$  NMR spectra of the aqueous layer from the octanol–water partitioning of Ac-Cav-NH<sub>2</sub>·HCl (**3.2**·HCl) into octanol in the absence or presence of sodium bis(2-ethylhexyl) phosphate (lipid **3.7**), showing 65% (= 0.90/1.39) remaining in the aqueous layer post-octanol + lipid wash. A. Pre-wash. B. Post-wash with octanol. C. Post-wash with octanol containing lipid **3.7**.

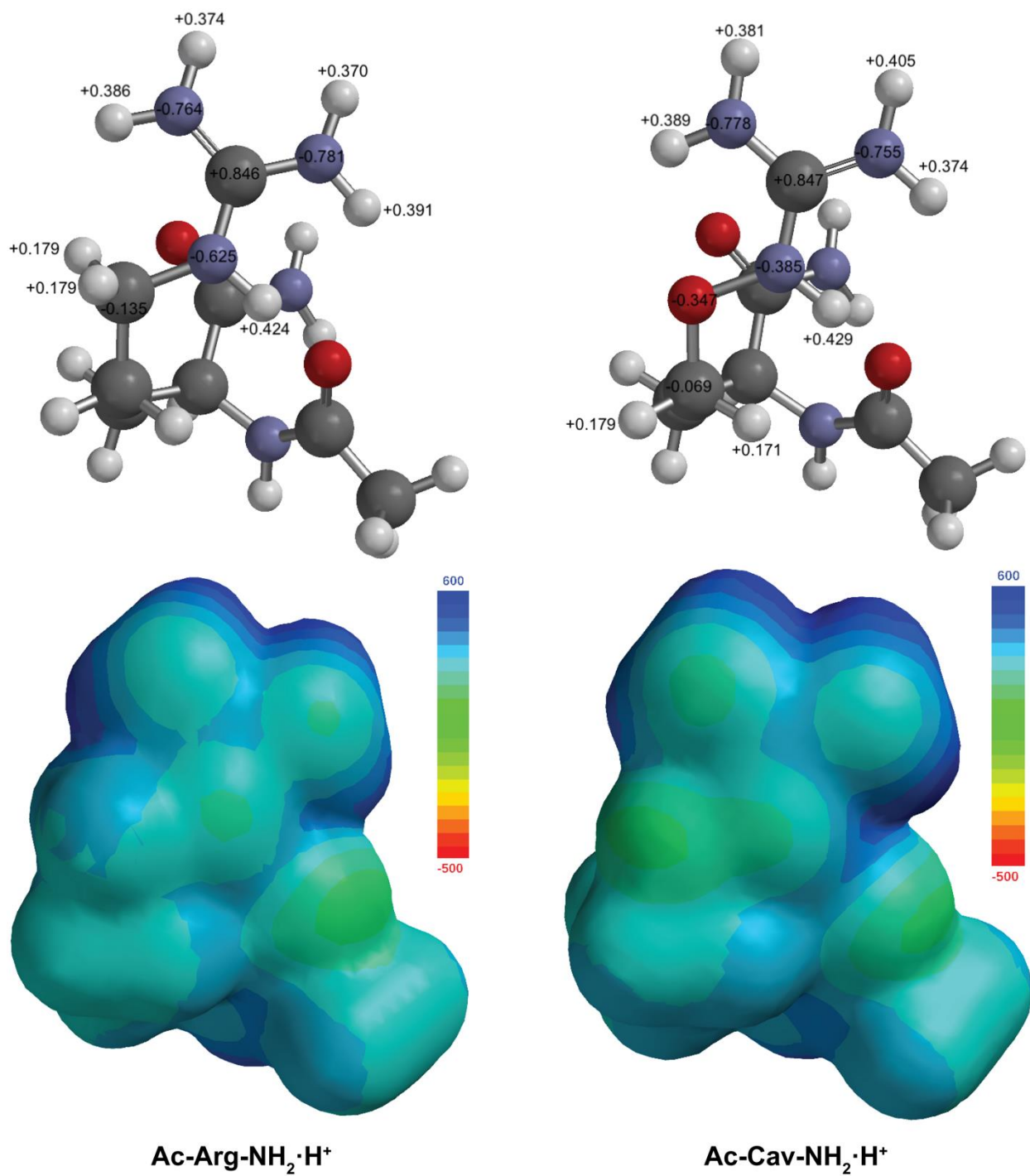


**Figure 3.3.**  $^1\text{H}$  NMR spectra from the partitioning of Ac-Arg-NH<sub>2</sub>·HCl (**3.1·HCl**) into octanol alone (top) and octanol with a lipid (bottom). A. **3.1·HCl** and lipid **3.6** (carboxylate). B. **3.1·HCl** and lipid **3.6** at pH 3.5. C. **3.1·HCl** and lipid **3.7** (phosphate). D. **3.1·HCl** and lipid **3.8** (sulfate). Experiments were performed in duplicate; one data set is shown.



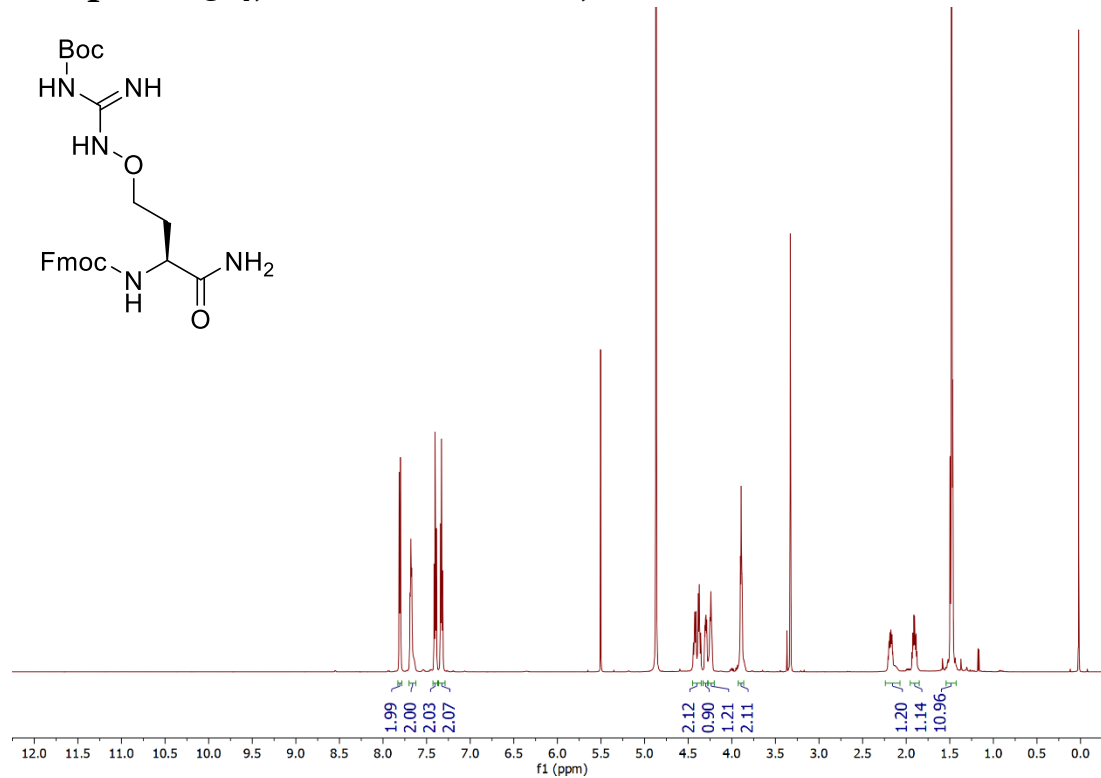
**Figure 3.4.**  $^1\text{H}$  NMR spectra from the partitioning of Ac-Cav-NH<sub>2</sub>·HCl (**3.2·HCl**) into octanol alone (top) and octanol with a lipid (bottom). A. **3.2·HCl** and lipid **3.6** (carboxylate). B. **3.2·HCl** and lipid **3.6** at pH 3.5. C. **3.2·HCl** and lipid **3.7** (phosphate). D. **3.2·HCl** and lipid **3.8** (sulfate). Experiments were performed in duplicate; one data set is shown.



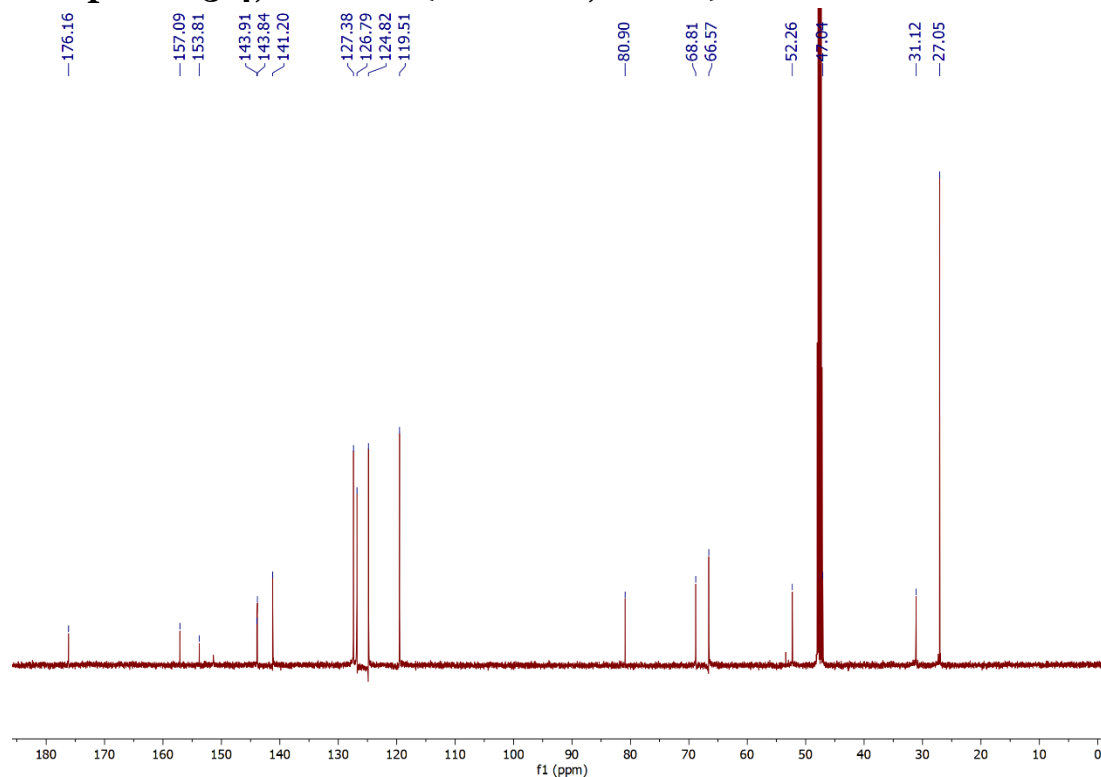


**Figure 3.5.** Calculation with Spartan '18 (Wavefunction, Irvine, CA) showing the electron density on atoms in Ac-Arg-NH<sub>2</sub>·H<sup>+</sup> (left) and Ac-Cav-NH<sub>2</sub>·H<sup>+</sup> (right).

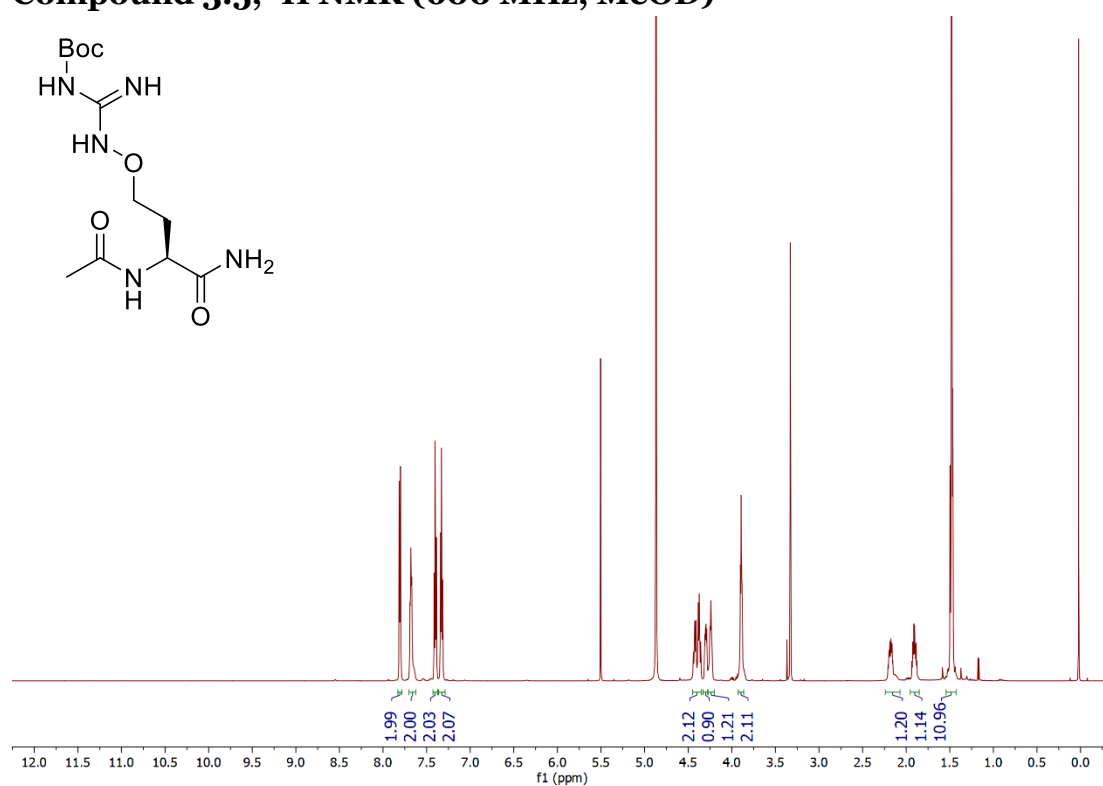
## NMR Spectra Compound 3.4, <sup>1</sup>H NMR (600 MHz, MeOD)



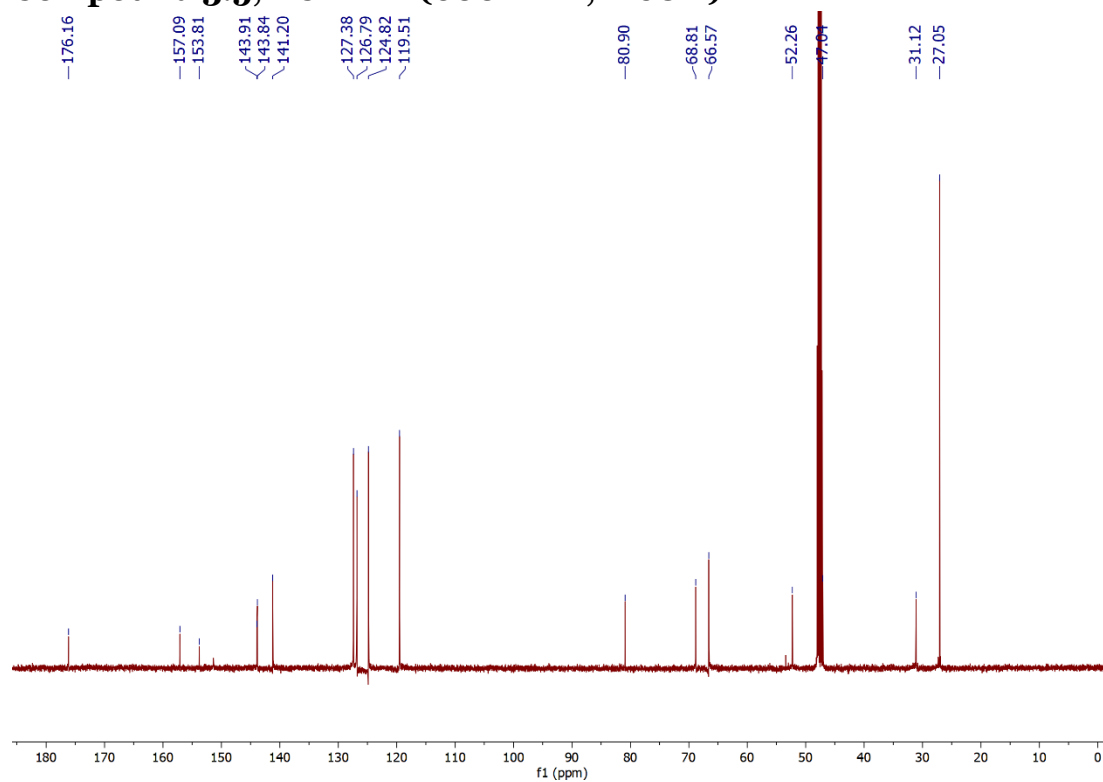
## Compound 3.4, <sup>13</sup>C NMR (600 MHz, MeOD)



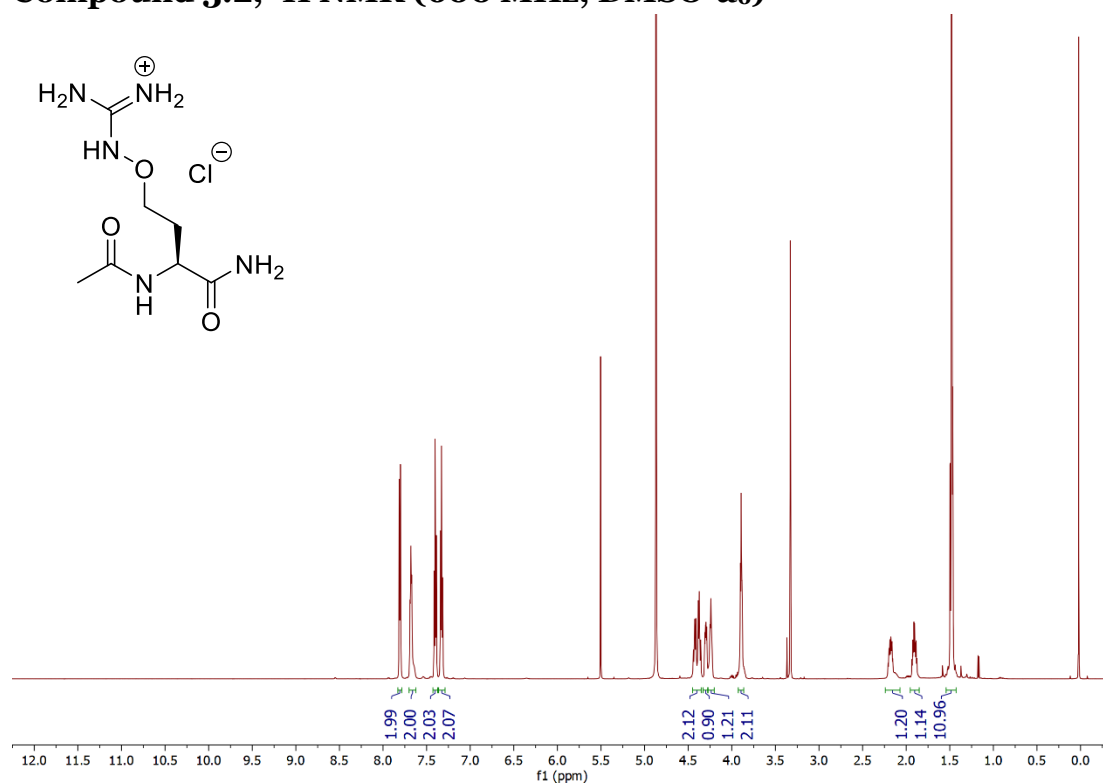
### Compound 3.5, <sup>1</sup>H NMR (600 MHz, MeOD)



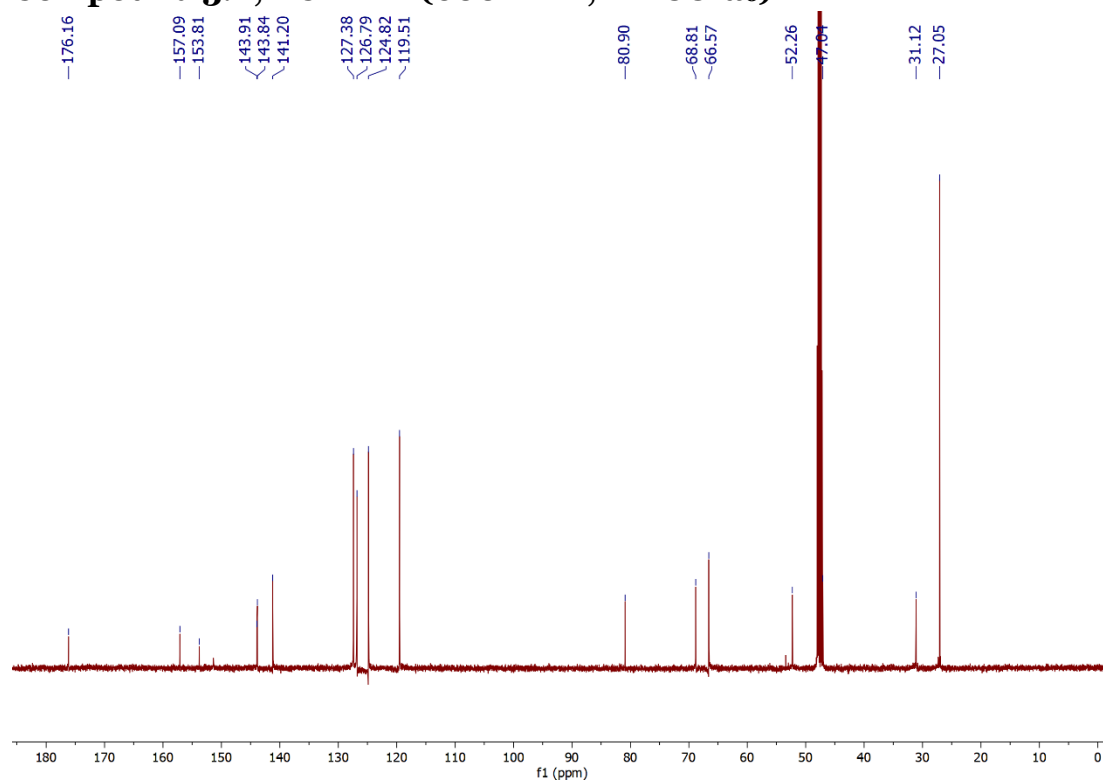
### Compound 3.5, <sup>13</sup>C NMR (600 MHz, MeOD)



### Compound 3.2, <sup>1</sup>H NMR (600 MHz, DMSO-d<sub>6</sub>)



### Compound 3.2, <sup>13</sup>C NMR (600 MHz, DMSO-d<sub>6</sub>)

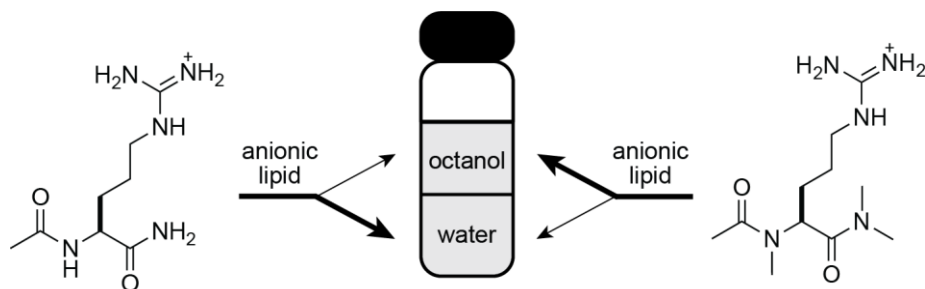


## **Chapter 4**

### ***N*<sup>α</sup>-Methylation of Arginine: Implications for Cell-Penetrating Peptides**

Contributions: All work was performed by Lindsey O. Calabretta.

## Abstract



The field of cell-penetrating peptides (CPPs) is dominated by the use of oligomers of arginine residues. Octanol–water partitioning in the presence of an anionic lipid is a validated proxy for cell-penetrative efficacy. Here, I added one, two, or three *N*-methyl groups to Ac-Arg-NH<sub>2</sub> and examine the effects on octanol–water partitioning in the presence of sodium dodecanoate. I found that increasing *N*-methylation correlates with increasing partitioning into octanol, which is predictive of higher cell-penetrative ability. These findings indicate that the ability of *N*<sup>α</sup>-methylated peptides to enter human cells is worthy of exploration.

## Introduction

The *N*-methylation of the amide backbone is a popular modification of peptides. Doig and others have used *N*<sup>α</sup>-methylated peptides as antagonists of β-sheet formation in the context of amyloidogenic peptides.<sup>146–151</sup> Still others have used main-chain *N*-methylation to improve the intestinal uptake, cell permeability, and metabolic stability of peptides and peptide-like therapeutic agents.<sup>152–156</sup>

The use of main-chain *N*-methylation as a method to increase cell permeability has taken inspiration from the natural product cyclosporine A (CsA). CsA is a cyclic peptide that is remarkably cell-permeable due, in part, to the *N*-methylation of seven amido groups. This replacement of N–H with N–CH<sub>3</sub> eliminates the ability to donate a hydrogen bond to solvent water molecules and thereby reduces the energetic cost of desolvation that is necessary for crossing a lipid bilayer.<sup>157</sup>

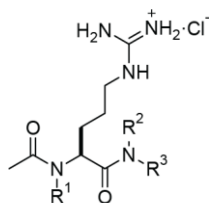
A distinct method to increase cellular uptake of cargo has been to append a cell-penetrating peptide (CPP).<sup>82,85–90</sup> Inspired by HIV-tat peptide and penetratin, CPPs are rich in arginine residues. The binding of guanidinium groups to cell-surface oxoanions can facilitate the transport of the peptide and pendant cargo into cells through endocytosis or direct translocation.

Wender and coworkers have explored CPPs based on peptoids, which are *N*-alkylated glycine oligomers. They found that peptoid-based CPPs are more cell-permeable than peptide-based CPPs.<sup>15</sup> Similarly, Kodadek and coworkers screened a large library of peptides and peptoids and found that peptoids are generally more cell-penetrating than peptides.<sup>158</sup> Although peptoids exhibit greater cell permeability than peptides, commercial monomers for the synthesis of peptoids are more costly than those for the synthesis of cognate *N*-methyl peptides.

To my knowledge, these two stratagems for cell penetration— $N^\alpha$ -methylation and CPPs—have yet to be used in combination. Here, I examine whether an  $N^\alpha$ -methylated arginine residue has the physicochemical attributes desirable for cell penetration.

## Results and Discussion

In this proof-of-concept study, I used four arginine derivatives: Ac-Arg-NH<sub>2</sub> (**4.1**) and its  $N$ -methylated analogs: Ac-Arg-NHMe (**4.2**), Ac-Arg-NMe<sub>2</sub> (**4.3**), and Ac-( $N$ -Me)Arg-NMe<sub>2</sub> (**4.4**) (Scheme 4.1). As a proxy for cell penetration, I used the method pioneered by Rothbard and Wender, who established that partitioning from water into octanol in the presence of an anionic lipid correlates with cell-penetration ability. For example, Arg<sub>9</sub> partitions into octanol in the presence of dodecanoate whereas Orn<sub>9</sub> (which is a nonamer of L-ornithine) does not; likewise, they also observed that Arg<sub>9</sub> is taken up into cells while Orn<sub>9</sub> is not.<sup>58,131</sup>



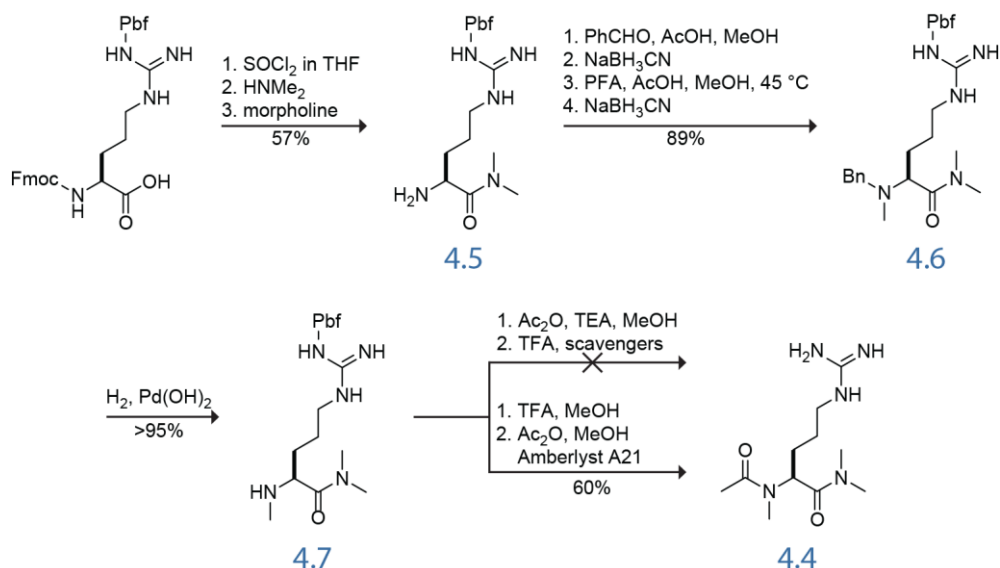
- 4.1-Cl R<sup>1,2,3</sup> = H  
 4.2-Cl R<sup>1,2</sup> = H, R<sup>3</sup> = Me  
 4.3-Cl R<sup>1</sup> = H, R<sup>2,3</sup> = Me  
 4.4-Cl R<sup>1,2,3</sup> = Me

**Scheme 4.1.** Structures of Ac-Arg-NH<sub>2</sub>·HCl (**4.1·HCl**), Ac-Arg-NHMe·HCl (**4.2·HCl**), Ac-Arg-NMe<sub>2</sub>·HCl (**4.3·HCl**), and Ac-( $N$ -Me)Arg-NMe<sub>2</sub>·HCl (**4.4·HCl**).

To access tri- $N$ -methylated analog Ac-( $N$ -Me)Arg-NMe<sub>2</sub> (**4.4**), I converted commercial Fmoc-Arg(Pbf)-OH into a dimethyl amide (**4.5**) (Scheme 4.2). My initial attempts to do so used traditional coupling methods, such as HATU in DMF. These methods resulted in low yields and difficulty in separating the polar product from the



coupling reagent. Instead, I converted the carboxyl group of Fmoc-Arg(Pbf)-OH into an acyl chloride with  $\text{SOCl}_2$  and produced dimethyl amide **4.5** with the addition of dimethylamine.<sup>159</sup> The dimethylamine was also partially successful at inciting a  $\beta$ -elimination to remove the Fmoc group, and morpholine was used to entice the elimination of remaining Fmoc groups. Next, I sought to append a single methyl group to the N-terminal amino group. Using an approach inspired by White and Konopelski,<sup>160</sup> I installed a benzyl and methyl group on the amino group through iterative reductive amination reactions with benzaldehyde and paraformaldehyde (PFA) to produce compound **4.6**. Subsequently, I removed the benzyl group by hydrogenolysis to produce compound **4.7**.

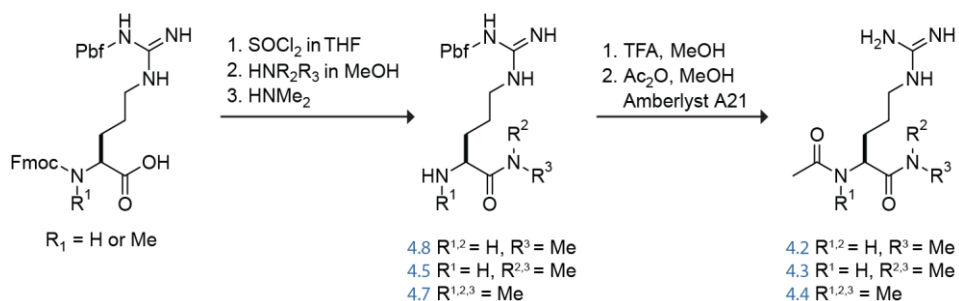


**Scheme 4.2.** Synthetic route to Ac-(N-Me)Arg-NMe<sub>2</sub> (**4.4**).

Initially, I attempted to install an acetyl group on compound **4.7** by using acetic anhydride, followed by removal of the Pbf group<sup>161</sup> with TFA in the presence of scavengers, such as ethanethiol and phenylsilane. Surprisingly, these conditions resulted

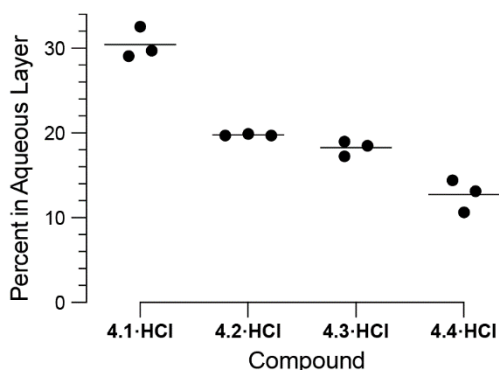
in the hydrolysis of the dimethyl amide along with the removal of the Pbf group. In hindsight, I realized that amides in close proximity to carboxyl groups can suffer rapid hydrolysis.<sup>162,163</sup> I hypothesize that, in acidic conditions, the acetyl oxygen attacks the protonated C-terminal amide to form a five-membered ring, which leads to the hydrolysis of the C-terminal amide. I obtained conformational data for **4.4** using selective 1D NOESY (discussed later) and found that the carbonyls are preorganized for this intramolecular attack (Figure 4.2). The electron-donating N-terminal methyl group also stabilizes this intermediate. To avoid this degradative route, I first cleaved the Pbf group with TFA and then appended the acetyl group by using acetic anhydride and the tertiary amino resin Amberlyst A21 to produce the desired compound, Ac-(*N*-Me)Arg-NMe<sub>2</sub> (**4.4**).

The mono- and di-*N*-methylated compounds, Ac-Arg-NHMe (**4.2**) and Ac-Arg-NMe<sub>2</sub> (**4.3**), were synthesized by the sequential use of thionyl chloride and acetic anhydride as in the route to compound **4.4** (Scheme 4.3). Later, I acquired Fmoc-*N*-Me-Arg(Pbf)-OH and used this same route (Scheme 4.3) to produce compound **4.4**. The unmethylated compound, Ac-Arg-NH<sub>2</sub> (**4.1**), was obtained from a commercial vendor as an acetate salt or was synthesized from H-Arg-NH<sub>2</sub>·2HCl using the acetylation conditions described above. Compounds **4.1–4.4** were converted to the desired chloride salt by using a guanidine resin charged with HCl, as described previously.<sup>164</sup>



**Scheme 4.3.** Synthetic route to compounds **4.2·HCl**, **4.3·HCl**, and **4.4·HCl**.

With compounds **4.1–4.4·HCl** in hand, I compared their ability to partition from water into octanol in the presence and absence of an anionic lipid, sodium dodecanoate. Previously, I used this anion-mediated octanol–water partitioning to compare the partitioning of L-canavanine (which is  $\delta$ -oxa-arginine) versus arginine.<sup>164</sup> Here, I used this method to compare the partitioning of compounds **4.1–4.4·HCl** (Figures 4.3-4.6). To do so, I prepared solutions of each amino acid in D<sub>2</sub>O and added octanol. I carefully extracted an aliquot of the D<sub>2</sub>O layer and spiked it with a known quantity of a reference compound as an internal standard. Then, I used <sup>1</sup>H-NMR spectroscopy to determine the concentration of compounds **4.1–4.4** that remained in the water layer after the wash with octanol. I found that all of the compounds remained entirely in the water layer after washing with octanol. Next, I exposed the aqueous solutions of **4.1–4.4·HCl** to octanol containing sodium dodecanoate. I found that dodecanoate transports all of the compounds into the octanol layer but to varying extents (Figure 4.1).



**Figure 4.1.** Graph showing the extent of octanol–water partitioning of compounds **4.1–4.4·HCl** in the presence of sodium dodecanoate (2.5 equiv)<sup>131,164</sup> were determined by <sup>1</sup>H-NMR spectroscopy.

We observed a general trend: increasing the extent of  $N^\alpha$ -methylation of Ac-Arg-NH<sub>2</sub>, decreased its concentration in the water layer after partitioning with octanol plus dodecanoate. The largest differential occurred between compounds with zero methyl groups (**4.1**) and one methyl group (**4.2**), with smaller differences between one methyl group (**4.2**) and two methyl groups (**4.3**) and between two methyl groups (**4.3**) and three methyl groups (**4.4**). Even so, I did observe a substantial increase in the partitioning into octanol of **4.4** compared to **4.2**. These two compounds most accurately mimic the environment in the middle of an unmodified peptide (**4.2**) and  $N^\alpha$ -methylated peptide (**4.4**).

To provide insight on the origin of anion-mediated partitioning upon  $N^\alpha$ -methylation, I calculated the cLog  $P$  and topological polar surface area (TPSA) values of all compounds (Table 4.1). The cLog  $P$  values of compounds **4.1–4.4**·H<sup>+</sup> show no correlation with the partitioning results. In contrast, the TPSA of each compound does correlate with the partitioning results. With decreasing TPSA, the partitioning of the compound into octanol increases. Moreover, the difference in TPSA between compounds **4.2** and **4.3** and between compounds **4.3** and **4.4** is 9 Å<sup>2</sup>. The difference in TPSA between compounds **4.1** and **4.2** is greater at 14 Å<sup>2</sup>, analogous to the greater difference in their octanol–water partitioning in the presence of sodium dodecanoate (Figure 4.1).

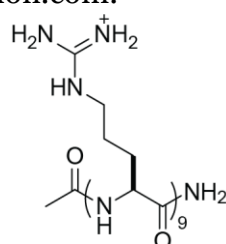
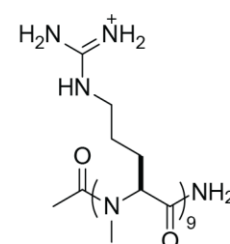
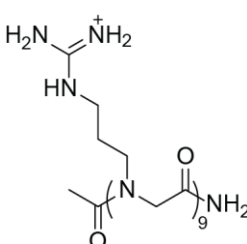
Previously, I observed that  $N$ -acetylated canavanine diamide (Ac-Cav-NH<sub>2</sub>) partitions into octanol to a lesser extent than does compound **4.1** in the presence of dodecanoate.<sup>131</sup> Whereas the TPSAs of **4.2–4.4**·H<sup>+</sup> are lower than that of **4.1**·H<sup>+</sup>, the TPSA of Ac-Cav-NH<sub>2</sub>·H<sup>+</sup> is higher than that of **4.1**·H<sup>+</sup>, suggesting that TPSA is predictive of anion-mediated octanol–water partitioning. Likewise, TPSA values correlate inversely with the permeability of small-molecule drugs.<sup>165,166</sup>

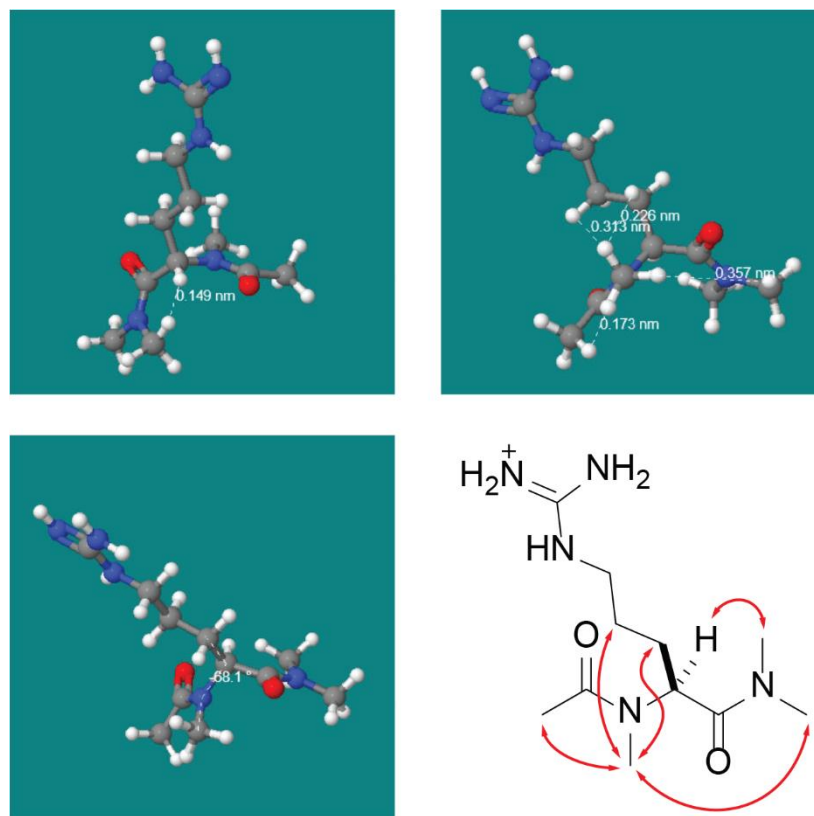
**Table 4.1.** Values of cLog *P* and TPSA for arginine derivatives as calculated with software from Molinspiration Cheminformatics (Slovenský Grob, Slovak Republic).

Compound	cLog <i>P</i>	TPSA (Å <sup>2</sup> )
Ac-Arg-NH <sub>2</sub> ·H <sup>+</sup> ( <b>4.1</b> ·H <sup>+</sup> )	-4.63	135.83
Ac-Arg-NHMe·H <sup>+</sup> ( <b>4.2</b> ·H <sup>+</sup> )	-4.72	121.84
Ac-Arg-NMe <sub>2</sub> ·H <sup>+</sup> ( <b>4.3</b> ·H <sup>+</sup> )	-4.81	113.05
Ac-(N-Me)Arg-NMe <sub>2</sub> ·H <sup>+</sup> ( <b>4.4</b> ·H <sup>+</sup> )	-4.70	104.26
Ac-Cav-NH <sub>2</sub> ·H <sup>+</sup>	-4.90	147.44

We extended this analysis by calculating the TPSA values of a well-known CPP, **R<sub>9</sub>**, and its *N*<sup>α</sup>-methylated congener, **Me-R<sub>9</sub>**. I found that **Me-R<sub>9</sub>** has a significantly lower TPSA value than does **R<sub>9</sub>** (Table 4.2). The TPSA calculated for the analogous peptoid, **Narg<sub>9</sub>**, is also lower than that of **R<sub>9</sub>** and, remarkably, results in exactly the same value as **Me-R<sub>9</sub>**. These calculated values correlate with the experimental data of Kodadek and coworkers, who observed that peptoids had, in general, a lower TPSA and higher cell permeability than do peptides and that peptoids with especially small TPSA values had enhanced cell permeability.<sup>158</sup>

**Table 4.2.** Values of cLog *P* and TPSA for **R<sub>9</sub>**, **Me-R<sub>9</sub>**, and **Narg<sub>9</sub>** as calculated with molinspiration.com.

	cLog <i>P</i>	TPSA (Å <sup>2</sup> )
	-6.70	877.62
	-6.54	798.52
	-6.74	798.52



**Figure 4.2.** A model of **4.4** was constructed in Chemagic Virtual Molecular Model Kit according to the NOESY correlations found. (Note: this is not a computationally optimized model.) Several angles of the model are shown with the distance between protons labeled (above) and the Me- $N$ - $C^\alpha$ - $CH_2^\beta$  dihedral angle labeled (bottom left). NOESY correlations can be observed up to 0.4-0.5 nm. The notable NOESY correlations are displayed on a 2D representation of **4.4** with red arrows (bottom right).

Finally,  $N^\alpha$ -methylation is known to restrict the conformational flexibility of peptides.<sup>167</sup> I noticed that in  $D_2O$  solutions, the signal for the  $\delta$  protons of amino acids **4.1-4.3** was a triplet, whereas the signal for **4.4** was two overlapping doublet of triplets with the doublet  $J \approx 20$  Hz, indicative of geminal coupling. This suggested that **4.4** is significantly more conformationally restricted than the other amino acids for the chirality of the  $\alpha$  position to induce different  $\delta$  proton environments several bonds away. I performed selective 1D NOESY on all protons of **4.4** (Figure 4.2, Figure 4.7). HSQC and 2D COSY aided in the assignment of the signals (see NMR Spectra, p. 182). I observed

through-space correlations that indicate a *trans* peptide bond of the *N*-terminal methylamide – the *C*-terminal amide, being symmetric, could not be said to be *cis* or *trans*.

Arvidsson and coworkers have obtained crystal structures of *all-N<sup>α</sup>*-methyl-peptides and determined that they adopt an all-*trans* extended conformation similar to a β-strand.<sup>168</sup> Peptides of *N<sup>α</sup>*-Me-Ala adopt two main-chain conformations, one of which has a Me-*N*-C<sup>α</sup>-Me<sup>β</sup> dihedral angle of -68°. The NOESY correlations I observed support this conformation (Figure 4.2). The secondary structure of CPPs is known to influence their efficiency and pathway of uptake.<sup>135</sup> Thus, the conformational rigidity in *N<sup>α</sup>*-methyl-CPPs such as **Me-R<sub>9</sub>** could prove to be critical to their cell-penetration capability.

## Conclusions

We have discovered that *N*-methylated arginine monomers partition more into octanol in the presence of an anionic lipid than does a non-methylated monomer. These findings illuminate a gap in the field of cell-penetration, as an *N*-methylated CPP could have increased cell-penetrating ability. In displaying a guanidinium group but not an N-H group, an *N*-methylated arginine residue is reminiscent of the γ-guanidinoproline residues of Wennemers and coworkers, which are superior to arginine residues at eliciting cell penetration.<sup>134</sup> Notably, *N*-methylated CPPs are highly accessible by solid-phase peptide synthesis because of the commercial availability of *N*-methylated amino acids. Finally, CPPs based on *N*-methylated amino acid residues might not only demonstrate enhanced cellular uptake, but also benefit from greater metabolic stability by avoiding proteolysis.

## Materials and Methods

### General

Materials. Commercial compounds were from Sigma–Aldrich (St. Louis, MO), Chem Impex (Wood Dale, IL) and Ambeed (Arlington Hts, IL) and were used without further purification.

Conditions. All procedures were performed in air at ambient temperature (~22 °C) and pressure (1.0 atm) unless specified otherwise.

Solvent removal. The phrase “concentrated under reduced pressure” refers to the removal of solvents and other volatile materials using a rotary evaporator while maintaining a water-bath temperature at 40 °C. Residual solvent was removed from samples at high vacuum (<0.1 Torr), which refers to the vacuum achieved by a mechanical belt-drive oil pump, or through lyophilization (freeze-drying) using a Labconco FreeZone freeze dryer.

Chromatography. Chemical reactions were monitored by thin-layer chromatography (TLC) using EMD 250 µm silica gel 60-F<sub>254</sub> plates and visualization with UV-illumination or KMnO<sub>4</sub>-staining, or by LC–MS with an ESI Agilent 6125B mass spectrometer. Flash chromatography was performed with a Biotage Isolera automated purification system using prepacked and re-packed SNAP KP silica gel columns and SNAP KP C18 columns.

Instrumentation. <sup>1</sup>H-NMR and <sup>13</sup>C-NMR spectra for compound characterization were obtained with Bruker spectrometers, and HRMS data were obtained with an Agilent 6545 Q-ToF mass spectrometer at the Department of Chemistry Instrumentation Facility at the Massachusetts Institute of Technology.



## Octanol–Water Partitioning Experiments

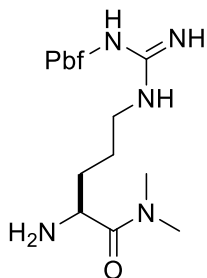
Procedures for the octanol–water partitioning experiment were similar to those reported previously<sup>164</sup> and are described in detail below.

Ac-Arg-NH<sub>2</sub>·HCl (**4.1·HCl**) (27.7 mg, 0.11 mmol) was dissolved in D<sub>2</sub>O (5.5 mL) to make a 0.02 M stock solution. Six Eppendorf tubes were prepared, three of which contained sodium dodecanoate (8.9 mg, 0.04 mmol, 2.5 equiv of **4.1·HCl**). An 800-μL aliquot of the arginine stock solution was pipetted into each of the six tubes. An 800-μL aliquot of octanol was then partitioned into each tube. The tubes were vortexed for 30 sec, then subjected to centrifugation for 15 min. Once the layers had separated, a 200-μL aliquot was removed carefully from the aqueous layers with a pipette and added to a new tube. A 200-μL aliquot of the unadulterated stock was also placed in a new tube. A solution of 0.2 M pyridine in D<sub>2</sub>O (20 μL) was added to each tube, the tubes were shaken, and the solutions were transferred to 3-mm NMR tubes for analysis. This procedure was repeated with **4.2·HCl**, **4.3·HCl**, and **4.4·HCl**.

<sup>1</sup>H-NMR spectra were collected of the aliquots using a 20-s relaxation delay (D1) to ensure quantitative integrations between the pyridine and amino acid residue. Spectral integrations were referenced to the signal of the *para*-hydrogen of pyridine, and the integration of the signal from the H<sup>δ</sup> protons of the amino acid residue were measured. The relative integrations pre- and post-wash were compared to each other. For all amino acid residues, the spectra following the octanol washes showed a slightly higher integration. This increase was likely the result of a small amount of octanol remaining in the water layer. Because the amount of additional integration in the post-octanol and octanol + lipid washes is likely to be the same, I evaluated the extent of partitioning by

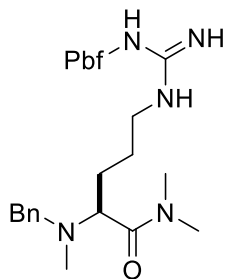
determining the difference between the octanol and octanol + lipid washes and averaging the values from the three replicates.

## Synthesis



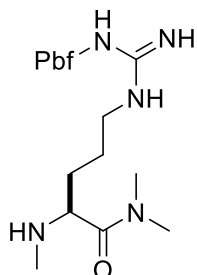
**H-Arg(Pbf)-NMe<sub>2</sub> (4.5).** Fmoc-Arg(Pbf)-OH (2 g, 3 mmol) was placed in a dry flask. Thionyl chloride (5 mL) was added, and the yellow solution was stirred under N<sub>2</sub>(g) for 45 min. The acyl chloride intermediate was concentrated under reduced pressure, dissolved in 5 mL of dry THF, and concentrated again to remove any remaining thionyl chloride. The resulting tan foam was dissolved in 5 mL of dry THF, and the solution was cooled under N<sub>2</sub>(g) in an ice bath. A 2 M solution of dimethylamine in THF (25 mL, 50 mmol) was added dropwise. After that addition, the resulting solution was stirred overnight. The solution was then vacuum-filtered to remove the dimethylamine salts and concentrated under reduced pressure. The resulting oil was then redissolved in morpholine, and the resulting solution was stirred for 1 h and again concentrated under reduced pressure. The crude oil was purified by chromatography on silica gel with a MeOH/DCM gradient, eluting at 20% v/v MeOH to produce compound **4.5** as a white foam (1.23 g, 52% yield). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>, δ): 6.60 (br, 1H), 6.45 (br, 2H), 3.76 (dd, *J* = 8.4, 3.8 Hz, 1H), 3.20 (br, 2H), 3.00 (s, 3H), 2.95 (s, 2H), 2.93 (s, 3H), 2.57 (s,

3H), 2.50 (s, 3H), 2.35 (br, 2H), 2.08 (s, 3H), 1.73–1.55 (m, 3H), 1.46 (s, 6H), 1.46 (m, 1H).  $^{13}\text{C}$  NMR (500 MHz,  $\text{CDCl}_3$ ,  $\delta$ ): 174.74, 158.64, 156.40, 138.24, 133.13, 132.19, 124.56, 117.42, 86.35, 50.69, 43.25, 40.81, 36.89, 35.90, 28.60, 28.59, 25.55, 19.27, 17.92, 12.48. HRMS  $m/z$  calcd for  $\text{C}_{21}\text{H}_{35}\text{N}_5\text{O}_4\text{S}$   $[\text{M} + \text{H}]^+$ , 454.2488; found 454.2483.



***N*-Bn-*N*-Me-Arg(Pbf)-NMe<sub>2</sub> (4.6).** H-Arg(Pbf)-NMe<sub>2</sub> (4.5) (200 mg, 0.44 mmol) was dissolved in methanol (5 mL). Benzaldehyde (48  $\mu\text{L}$ , 0.48 mmol) and acetic acid (25  $\mu\text{L}$ , 0.44 mmol) were added, and the resulting solution was stirred for 2 h. Sodium cyanoborohydride (30 mg, 0.48 mmol) was added, and the resulting solution was stirred overnight. Paraformaldehyde (14.4 mg, 0.48 mmol) and acetic acid (25  $\mu\text{L}$ , 0.44 mmol) were added, and the resulting solution was stirred at 40 °C for 1 h. Sodium cyanoborohydride (30 mg, 0.48 mmol) was dissolved in 1 mL of acetonitrile, and the resulting solution was added to the reaction mixture over 10 h via a syringe pump. After another 8 h, the solution was concentrated under reduced pressure. To remove excess cyanoborohydride, the crude product was dissolved in 5 mL of methanol and 1 mL of acetic acid, and the resulting solution was concentrated under reduced pressure, and then the process was repeated. The crude product was dissolved in water, and extracted three times with dichloromethane. The organic layers were combined, dried over  $\text{Na}_2\text{SO}_4(\text{s})$ , and filtered. The crude product was purified by chromatography on silica gel with a

MeOH/DCM gradient, eluting at 8% v/v MeOH to produce compound **4.6** as a white foam (219 g, 89% yield).  $^1\text{H NMR}$  (500 MHz,  $\text{CDCl}_3$ ,  $\delta$ ): 7.33–7.27 (m, 2H), 7.27–7.22 (m, 3H), 6.17 (br, 1H), 6.08 (br, 2H), 3.62 (d,  $J = 13.3$  Hz, 1H), 3.53 (d,  $J = 13.3$  Hz, 1H), 3.46 (dd,  $J = 10.2, 3.7$  Hz, 1H), 3.19 (br, 2H), 3.05 (s, 3H), 2.94 (s, 2H), 2.94 (s, 3H), 2.60 (s, 3H), 2.54 (s, 3H), 2.19 (s, 3H), 2.09 (s, 3H), 1.95–1.81 (m, 1H), 1.75–1.66 (m, 1H), 1.59–1.48 (m, 1H), 1.45 (s, 6H), 1.39–1.27 (m, 1H).  $^{13}\text{C NMR}$  (500 MHz,  $\text{CDCl}_3$ ,  $\delta$ ): 172.00, 158.60, 156.11, 138.74, 138.39, 133.22, 132.33, 128.82, 128.33, 127.22, 124.48, 117.37, 86.28, 63.35, 58.01, 43.26, 41.20, 37.85, 37.33, 36.04, 28.60, 28.59, 26.31, 19.27, 17.92, 12.47. **HRMS**  $m/z$  calcd for  $\text{C}_{29}\text{H}_{43}\text{N}_5\text{O}_4\text{S} [\text{M} + \text{H}]^+$ , 558.3114; found 558.3115.



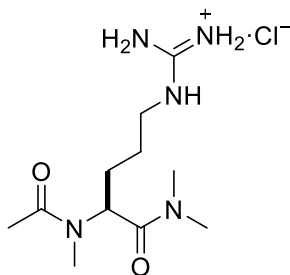
### **Me-Arg(Pbf)-NMe<sub>2</sub> (4.7)**

**Method 1.** *N*-Bn-*N*-Me-Arg(Pbf)-NMe<sub>2</sub> (**4.6**) (100 mg, 0.18 mmol) was dissolved in methanol (2 mL). The resulting solution was sparged with  $\text{N}_2(\text{g})$  for 10 min. Palladium hydroxide (15 mg, 15% wt) was added, and the resulting solution was sparged with  $\text{H}_2(\text{g})$  for 15 min. Then, the solution was stirred under a balloon of  $\text{H}_2(\text{g})$  overnight. The reaction mixture was filtered through celite and concentrated under reduced pressure. The product, as a white foam (76 mg, quantitative yield), was used directly.

**Method 2.** Fmoc-*N*-Me-Arg(Pbf)-OH (613 g, 0.92 mmol) was placed in a dry flask. Thionyl chloride (5 mL) was added, and the yellow solution was stirred under  $\text{N}_2(\text{g})$  for 1

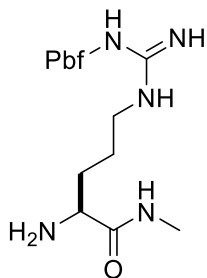
h. The acyl chloride intermediate was concentrated under reduced pressure, dissolved in 5 mL of dry THF, and concentrated again to remove any remaining thionyl chloride. The resulting tan foam was dissolved in 2 mL of dry THF, and the solution was cooled under N<sub>2</sub>(g) in an ice bath. A 2 M solution of dimethylamine in THF (5 mL, 10 mmol) was added dropwise and the solution was stirred for 1 h. The solution was then vacuum-filtered to remove the dimethylamine salts and concentrated under reduced pressure. The resulting foam was then redissolved in a 2 M solution of methylamine in MeOH (5 mL), and the resulting solution was stirred for 1 h and again concentrated under reduced pressure. The crude oil was purified by chromatography on silica gel with a MeOH/DCM gradient, eluting at 15% v/v MeOH to produce a white foam (454 mg, 97% yield.)

<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>, δ): 6.78 (br, 1H), 6.55 (br, 3H), 3.84 (m, 1H), 3.46 (s, 3H), 3.22 (t, *J* = 6.7 Hz, 2H), 3.07 (s, 3H), 2.99 (s, 3H), 2.95 (s, 3H), 2.57 (s, 3H), 2.50 (s, 3H), 2.08 (s, 3H), 1.87–1.76 (m, 1H), 1.76–1.55 (m, 3H), 1.46 (s, 6H). <sup>13</sup>C NMR (500 MHz, CDCl<sub>3</sub>, δ): 177.41, 158.66, 156.58, 138.25, 133.00, 132.19, 124.58, 117.46, 86.37, 59.04, 43.25, 40.50, 37.09, 35.96, 33.93, 28.78, 28.59, 24.94, 19.27, 17.94, 12.47. HRMS *m/z* calcd for C<sub>22</sub>H<sub>37</sub>N<sub>5</sub>O<sub>4</sub>S [M + H]<sup>+</sup>, 468.2645; found 468.2638.



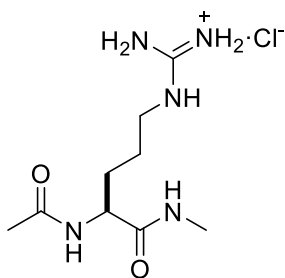
**N-Acetyl-N-Me-Arg-NMe<sub>2</sub>·HCl (4.4·HCl):** Me-Arg(Pbf)-NMe<sub>2</sub> (**4.7**) (532 mg, 1.14 mmol) was placed in a flask, followed by methanol (100 μL) and trifluoroacetic acid (3

mL). After 1.5 h, the trifluoroacetic acid was evaporated under a stream of N<sub>2</sub>(g). The deprotected product was then dissolved in water and extracted with DCM two times. The greenish organic fractions were discarded, and the aqueous fractions were combined and lyophilized. The crude intermediate was dissolved in methanol (12 mL), and Amberlyst A21 was added until the pH was above 7 (1.5 g). Acetic anhydride (323 μL, 3.4 mmol) was added, and the resulting solution was stirred for 3 h. The solution was filtered to remove the resin and concentrated under reduced pressure. The product was then purified by reversed-phase chromatography in H<sub>2</sub>O. The product was then flushed through a *N*-dimethylaminopropyl-*N'*-ethylguanidine resin pipette column charged with HCl and lyophilized to produce **4.2·HCl** as a white powder (200 mg, 60% yield). <sup>1</sup>H NMR (600 MHz, D<sub>2</sub>O δ): 5.24 (dd, *J* = 8.9, 6.0 Hz, 1H), 3.14 (dt, *J* = 20.4, *J* = 6.8, 1H), 3.12 (dt, *J* = 20.4, *J* = 6.8, 1H), 3.13–3.09 (m, 1H), 2.90 (s, 3H), 2.86 (s, 3H), 2.82 (s, 3H), 2.08 (s, 3H), 1.72 (dtd, *J* = 14.1, 8.0, 5.9 Hz, 1H), 1.64 (ddt, *J* = 14.1, 8.7, 7.2 Hz, 1H), 1.43 (p, *J* = 7.3 Hz, 2H). <sup>13</sup>C NMR (600 MHz, D<sub>2</sub>O δ): 174.28, 171.41, 156.70, 53.36, 40.66, 36.70, 35.87, 31.18, 25.19, 24.17, 20.87. HRMS *m/z* calcd for C<sub>11</sub>H<sub>23</sub>N<sub>5</sub>O<sub>2</sub> [M + H]<sup>+</sup>, 258.1930; found 258.1922.



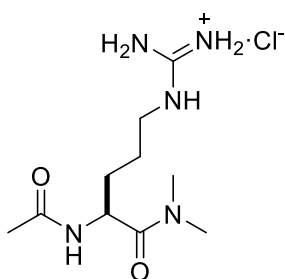
**H-Arg(Pbf)-NHMe (4.8).** Fmoc-Arg(Pbf)-OH (2 g, 3 mmol) was placed in a dry flask. Thionyl chloride (5 mL) was added, and the yellow solution was stirred under N<sub>2</sub>(g) for 1 h. The acyl chloride intermediate was concentrated under reduced pressure, dissolved in

5 mL of dry THF, and concentrated again to remove any remaining thionyl chloride. The resulting tan foam was dissolved in 5 mL of dry THF, and the solution was cooled under N<sub>2</sub>(g) in an ice bath. A 2 M solution of methylamine in THF (15 mL, 30 mmol) was added dropwise and the solution was stirred for 1 h. The solution was then vacuum-filtered to remove the dimethylamine salts and concentrated under reduced pressure. The resulting foam was then redissolved in a 2 M solution of methylamine in MeOH (10 mL), and the resulting solution was stirred for 1 h and again concentrated under reduced pressure. The crude oil was purified by chromatography on silica gel with a MeOH/DCM gradient, eluting at 15% v/v MeOH to produce a white foam (1.05 g, 80% yield). **<sup>1</sup>H NMR** (500 MHz, CDCl<sub>3</sub>, δ): 7.61 (d (br), *J* = 6.4 Hz, 1H), 6.51 (t (br), *J* = 5.6 Hz, 1H), 6.42 (s (br), 2H), 3.53 (t, *J* = 6.5 Hz, 1H), 3.22 (t, *J* = 6.4 Hz, 2H), 2.95 (s, 2H), 2.56 (s, 3H), 2.49 (s, 3H), 2.17 (s, 3H), 2.08 (s, 3H), 1.88–1.73 (m, 1H), 1.69–1.54 (m, 3H), 1.46 (s, 6H). **<sup>13</sup>C NMR** (500 MHz, CDCl<sub>3</sub>, δ): 175.14, 158.79, 156.50, 138.27, 132.73, 132.20, 124.69, 117.58, 86.45, 54.10, 43.24, 40.47, 31.65, 30.94, 28.60, 25.40, 25.34, 19.28, 17.94, 12.48. **HRMS** *m/z* calcd for C<sub>20</sub>H<sub>34</sub>N<sub>5</sub>O<sub>4</sub>S [M + H]<sup>+</sup>, 440.2332; found 440.2325.



***N*-Acetyl-Arg-NHMe·HCl (4.2·HCl).** H-Arg(Pbf)-NHMe (**4.8**) (440 mg, 1 mmol) was placed in a flask, followed by methanol (100 μL) and trifluoroacetic acid (3 mL). After 1 h, the trifluoroacetic acid was evaporated under a stream of N<sub>2</sub>(g). The deprotected

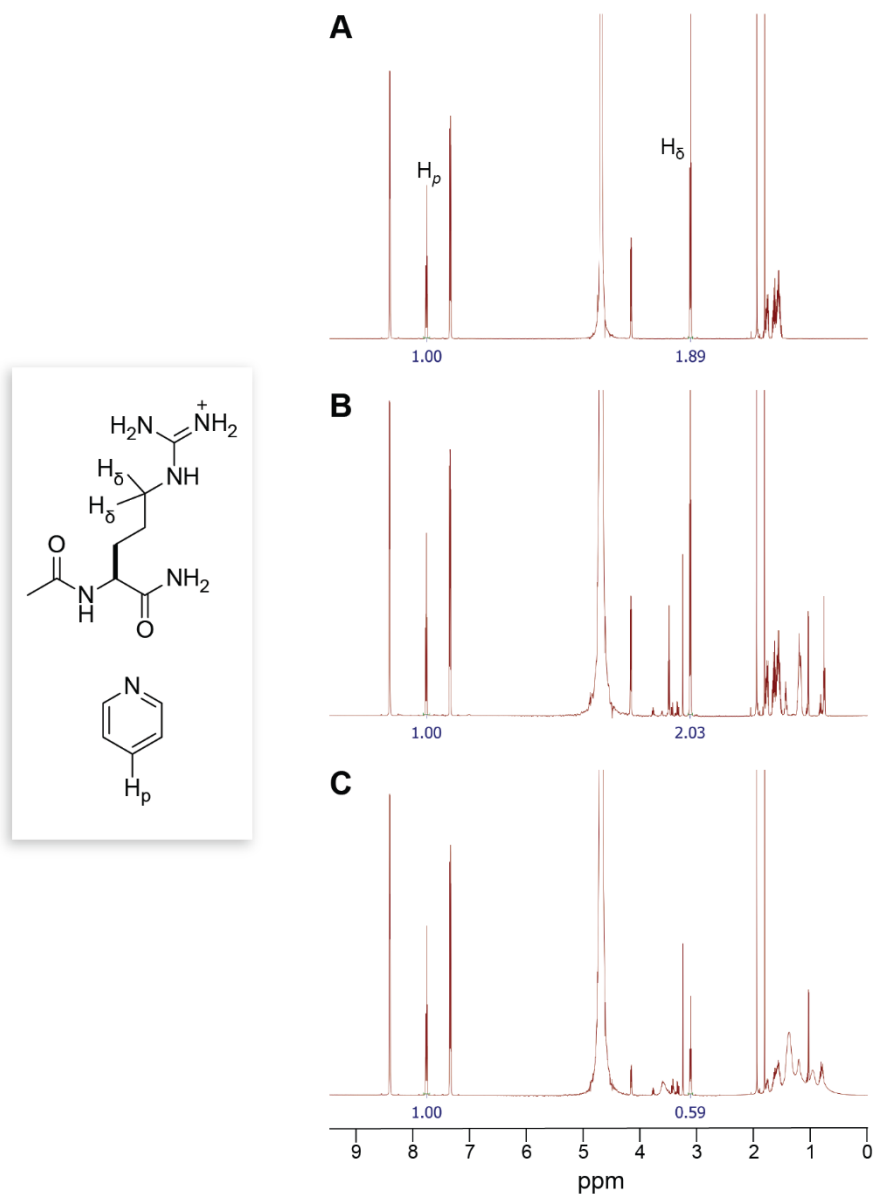
product was then dissolved in water and extracted with DCM two times. The greenish organic fractions were discarded, and the aqueous fractions were combined and evaporated under reduced pressure. The resulting oil was dissolved in MeOH (10 mL) and Amberlyst A21 free base (1.5 g) and acetic anhydride (284  $\mu$ L, 3 mmol) were added to the flask. The resulting solution was stirred for 3 h. The solution was filtered to remove the resin and concentrated under reduced pressure. The product was then purified by reversed-phase chromatography in H<sub>2</sub>O. The product was then flushed through a *N*-dimethylaminopropyl-*N'*-ethylguanidine resin pipette column charged with HCl and lyophilized to produce hard colorless foam (160 mg, 60% yield). **<sup>1</sup>H NMR** (600 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 8.10 (d, *J* = 8.1 Hz, 1H), 8.04 (s, 1H), 7.97 (d, *J* = 5.0 Hz, 1H), 7.28 (br, 4H), 4.18 (q, *J* = 7.7, 7.1 Hz, 1H), 3.08 (t, *J* = 6.3 Hz, 2H), 2.57 (d, *J* = 4.4 Hz, 3H), 1.86 (s, 3H), 1.71–1.61 (m, 1H), 1.56 – 1.34 (m, 3H). **<sup>13</sup>C NMR** (600 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 172.39, 169.79, 157.46, 52.54, 40.64, 29.56, 25.97, 25.62, 23.03. **HRMS** *m/z* calcd for C<sub>9</sub>H<sub>20</sub>N<sub>5</sub>O<sub>2</sub> [M + H]<sup>+</sup>, 230.1617; found 230.1608.



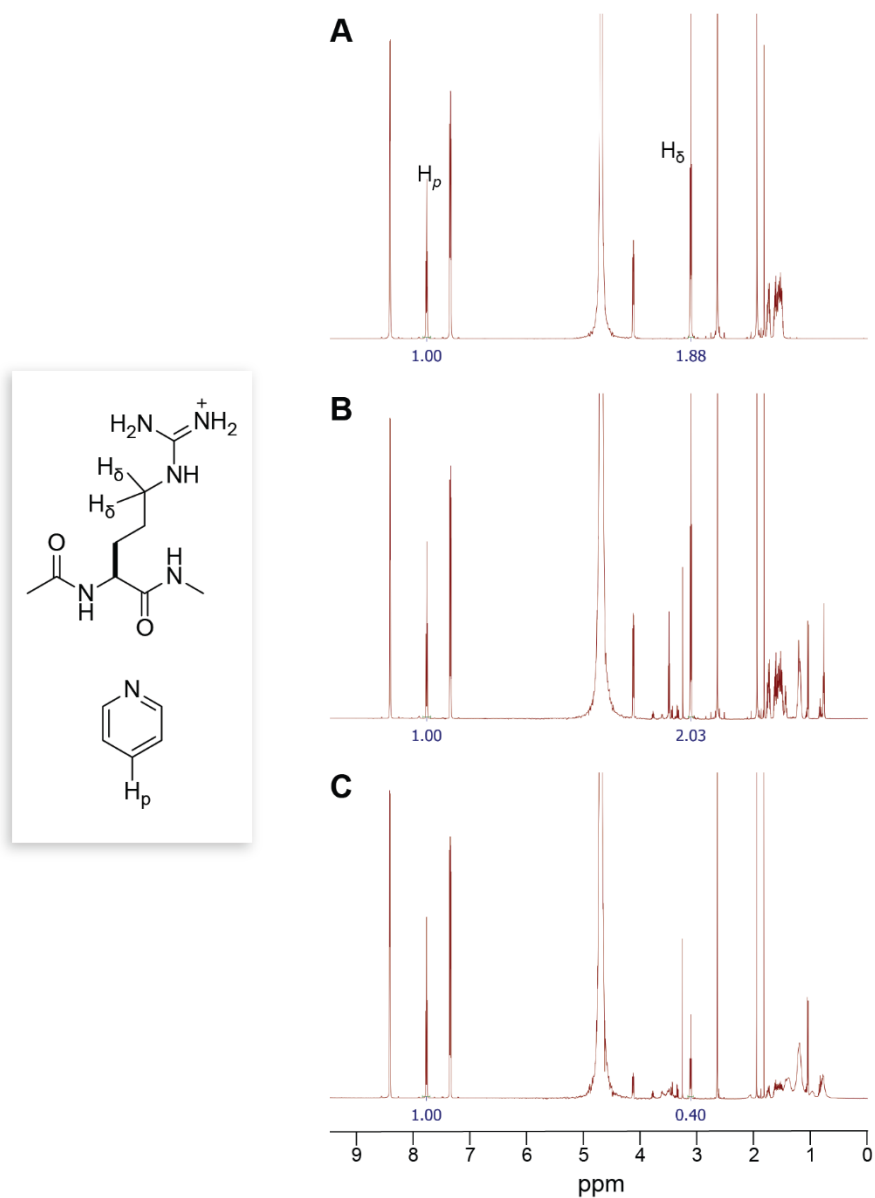
***N*-Acetyl-Arg-NMe<sub>2</sub>·HCl (4.3·HCl).** H-Arg(Pbf)-NMe<sub>2</sub> (**4.5**) (453 mg, 1 mmol) was placed in a flask, followed by methanol (100  $\mu$ L) and trifluoroacetic acid (3 mL). After 1 h, the trifluoroacetic acid was evaporated under a stream of N<sub>2</sub>(g). The deprotected product was then dissolved in water and extracted with DCM two times. The greenish



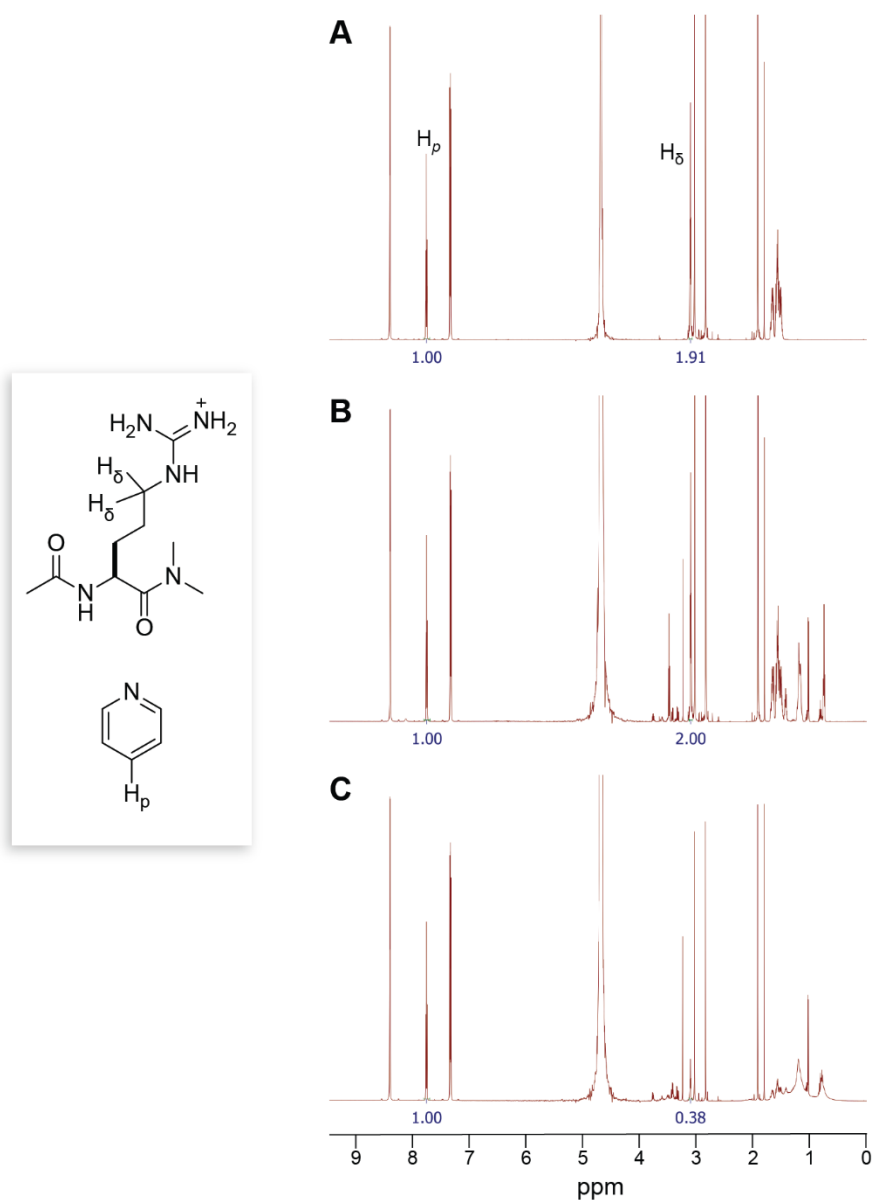
organic fractions were discarded, and the aqueous fractions were combined and evaporated under reduced pressure. The resulting oil was dissolved in MeOH (10 mL) and Amberlyst A21 free base (1.5 g) and acetic anhydride (284  $\mu$ L, 3 mmol) were added to the flask. The resulting solution was stirred for 3 h. The solution was filtered to remove the resin and concentrated under reduced pressure. The product was then purified by reversed-phase chromatography in H<sub>2</sub>O. The product was then flushed through a *N*-dimethylaminopropyl-*N'*-ethylguanidine resin pipette column charged with HCl and lyophilized to produce a white powder (132 mg, 47% yield). **<sup>1</sup>H NMR** (600 MHz, DMSO-*d*<sub>6</sub>  $\delta$ ): 8.14 (d, *J* = 8.3 Hz, 1H), 7.75 (t, *J* = 5.8 Hz, 1H), 7.39 (br, 2H), 6.95 (br, 2H), 4.69 (q, *J* = 6.7 Hz, 1H), 3.15–3.05 (m, 2H), 3.02 (s, 3H), 2.83 (s, 3H), 1.84 (s, 3H), 1.70–1.53 (m, 1H), 1.55–1.37 (m, 3H). **<sup>13</sup>C NMR** (600 MHz, DMSO-*d*<sub>6</sub>  $\delta$ ): 171.56, 169.46, 157.35, 48.26, 40.87, 37.07, 35.65, 29.16, 25.35, 22.77. **HRMS** *m/z* calcd for C<sub>10</sub>H<sub>21</sub>N<sub>5</sub>O<sub>2</sub> [M + H]<sup>+</sup>, 244.1774; found 244.1765.



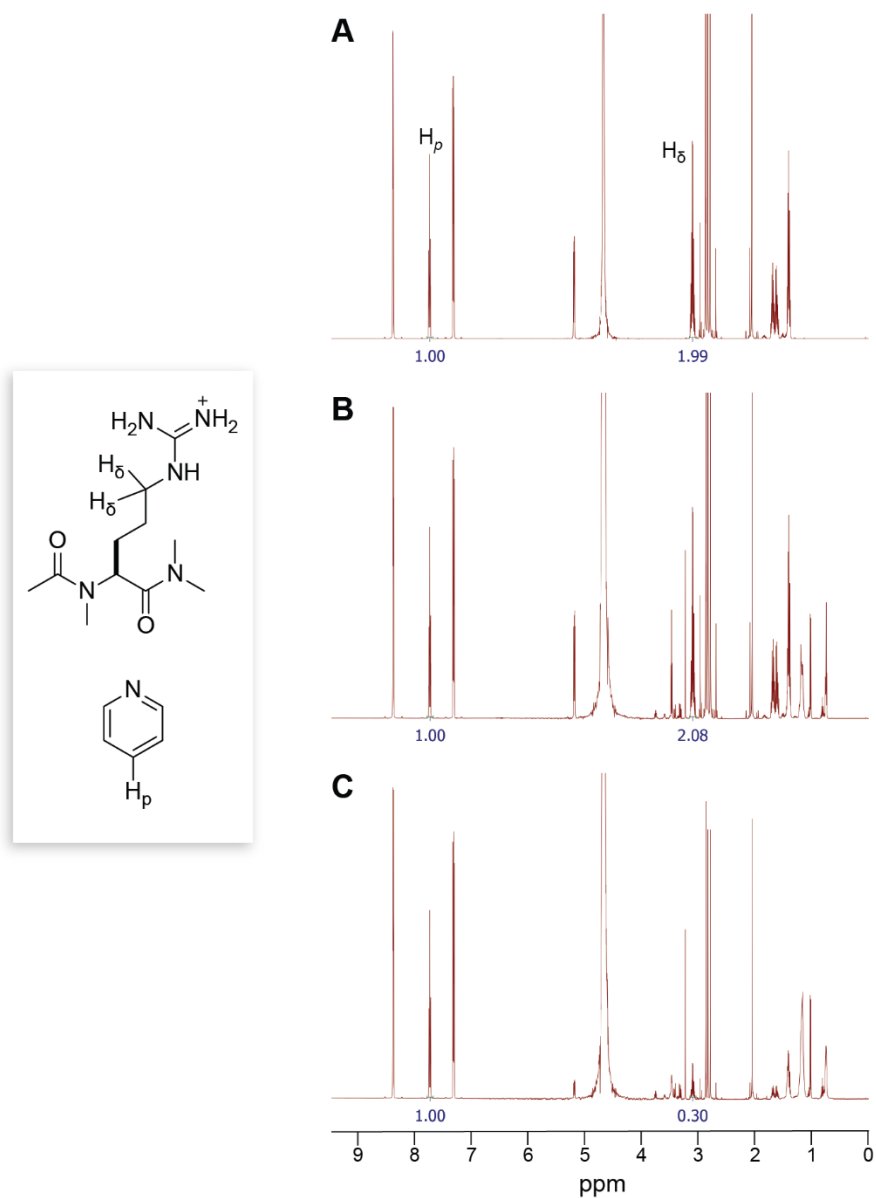
**Figure 4.3.** Representative  $^1\text{H}$  NMR spectra of the aqueous layer from the octanol–water partitioning of Ac-Arg-NH<sub>2</sub>·HCl (4.1·HCl) into octanol in the absence or presence of sodium dodecanoate showing 29% (= 0.59/2.03) remaining in the aqueous layer post-octanol + lipid wash. A. Pre-wash. B. Post-wash with octanol. C. Post-wash with octanol containing sodium dodecanoate.



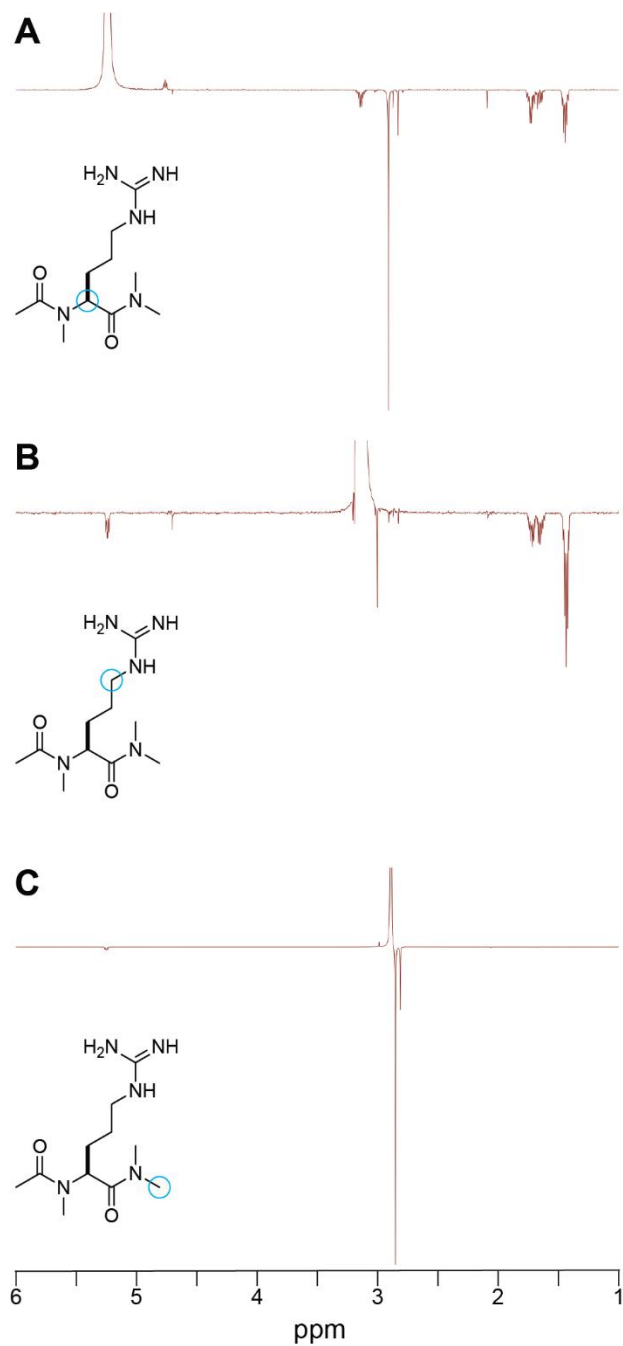
**Figure 4.4.** Representative  $^1\text{H}$  NMR spectra of the aqueous layer from the octanol–water partitioning of Ac-Arg-NHMe·HCl (**4.2·HCl**) into octanol in the absence or presence of sodium dodecanoate showing 20% (= 0.40/2.03) remaining in the aqueous layer post-octanol + lipid wash. A. Pre-wash. B. Post-wash with octanol. C. Post-wash with octanol containing sodium dodecanoate.



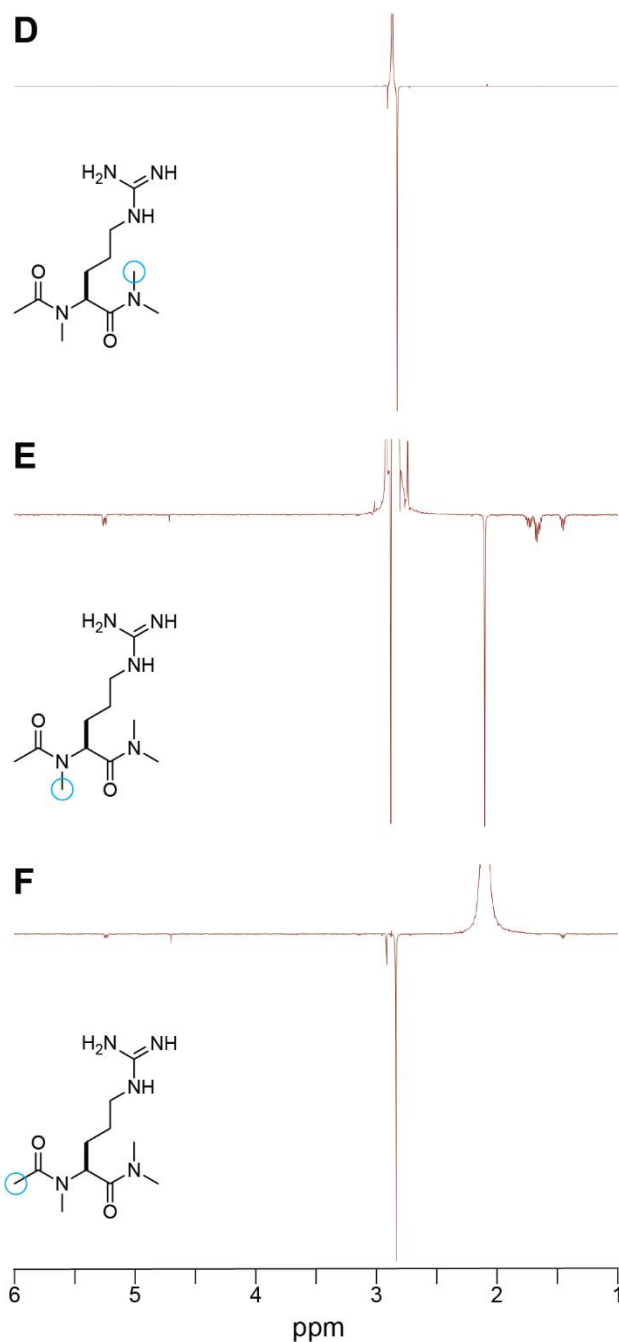
**Figure 4.5.** Representative  $^1\text{H}$  NMR spectra of the aqueous layer from the octanol–water partitioning of Ac-Arg-NMe<sub>2</sub>·HCl (**4.3**·HCl) into octanol in the absence or presence of sodium dodecanoate showing 19% (= 0.38/2.00) remaining in the aqueous layer post-octanol + lipid wash. The  $\alpha$  hydrogen of **4.3**·HCl is buried beneath the water signal. A. Pre-wash. B. Post-wash with octanol. C. Post-wash with octanol containing sodium dodecanoate.



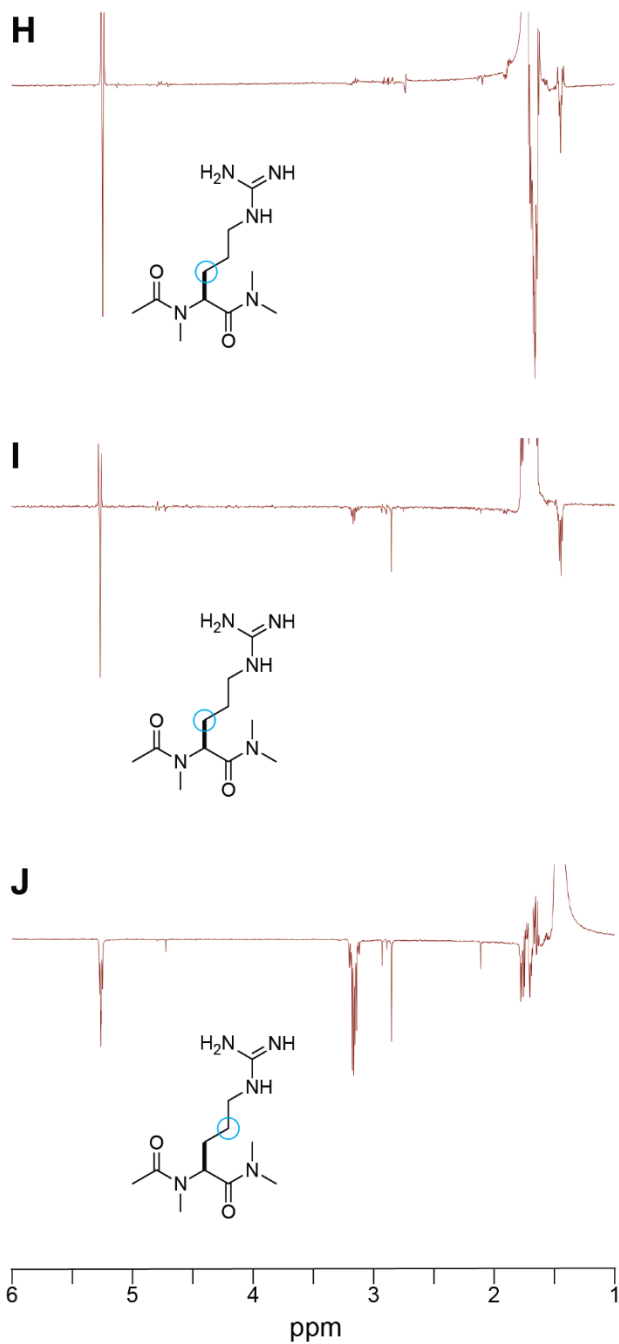
**Figure 4.6.** Representative  $^1\text{H}$  NMR spectra of the aqueous layer from the octanol–water partitioning of Ac-N-Me-Arg-NMe<sub>2</sub>·HCl (**4.4**·HCl) into octanol in the absence or presence of sodium dodecanoate showing 14% (= 0.30/2.08) remaining in the aqueous layer post-octanol + lipid wash. A. Pre-wash. B. Post-wash with octanol. C. Post-wash with octanol containing sodium dodecanoate.



**Figure 4.7, A-C.** Selective 1D NOESY experiments of **4.4** in D<sub>2</sub>O. Blue circles highlight the protons excited, phased positive in the spectra. Through-space correlations to the excited proton are shown phased negative. A.  $\alpha$ -proton. B.  $\delta$ -protons. C. One of the C-terminal methyl groups.



**Figure 4.7, D-F.** Selective 1D NOESY experiments of **4.4** in  $D_2O$ . Blue circles highlight the protons excited, phased positive in the spectra. Through-space correlations to the excited proton are shown phased negative. D. One of the C-terminal methyl groups. B. N-terminal methyl group. C. Acetyl methyl group.

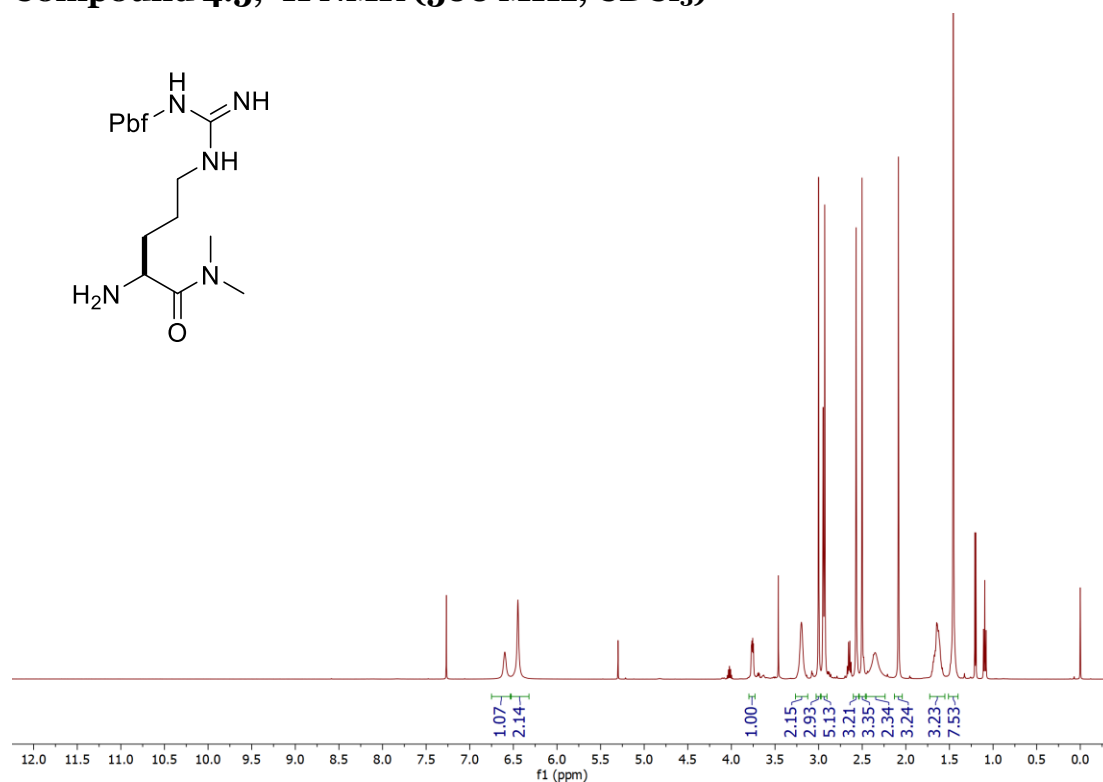


**Figure 4.7, D-F.** Selective 1D NOESY experiments of **4.4** in D<sub>2</sub>O. Blue circles highlight the protons excited, phased positive in the spectra. Through-space correlations to the excited proton are shown phased negative. D. One of the β-protons. B. One of the β-protons. C. γ-protons.

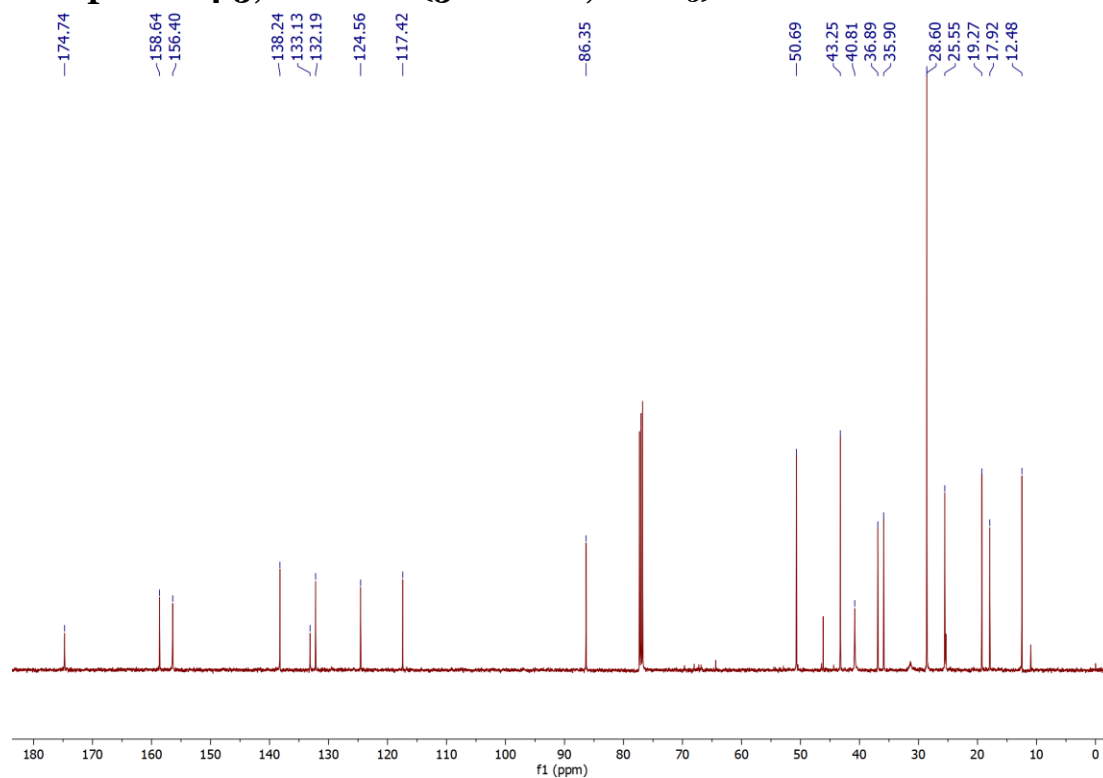


# NMR Spectra

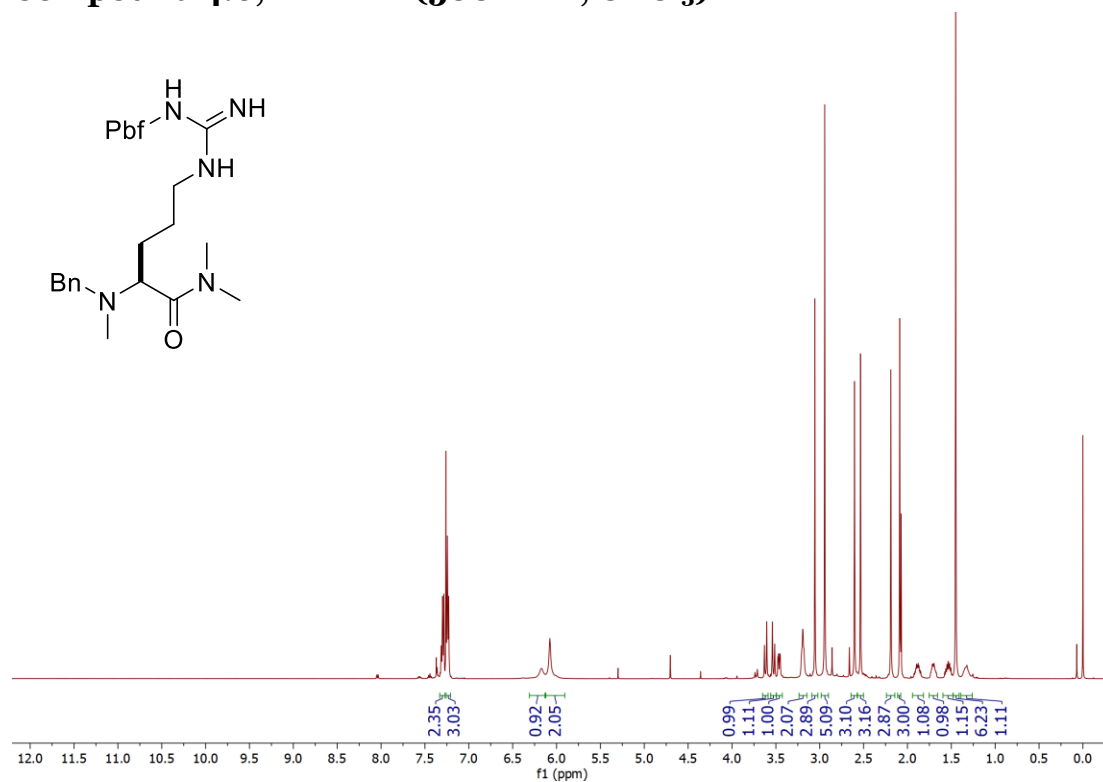
## Compound 4.5, <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)



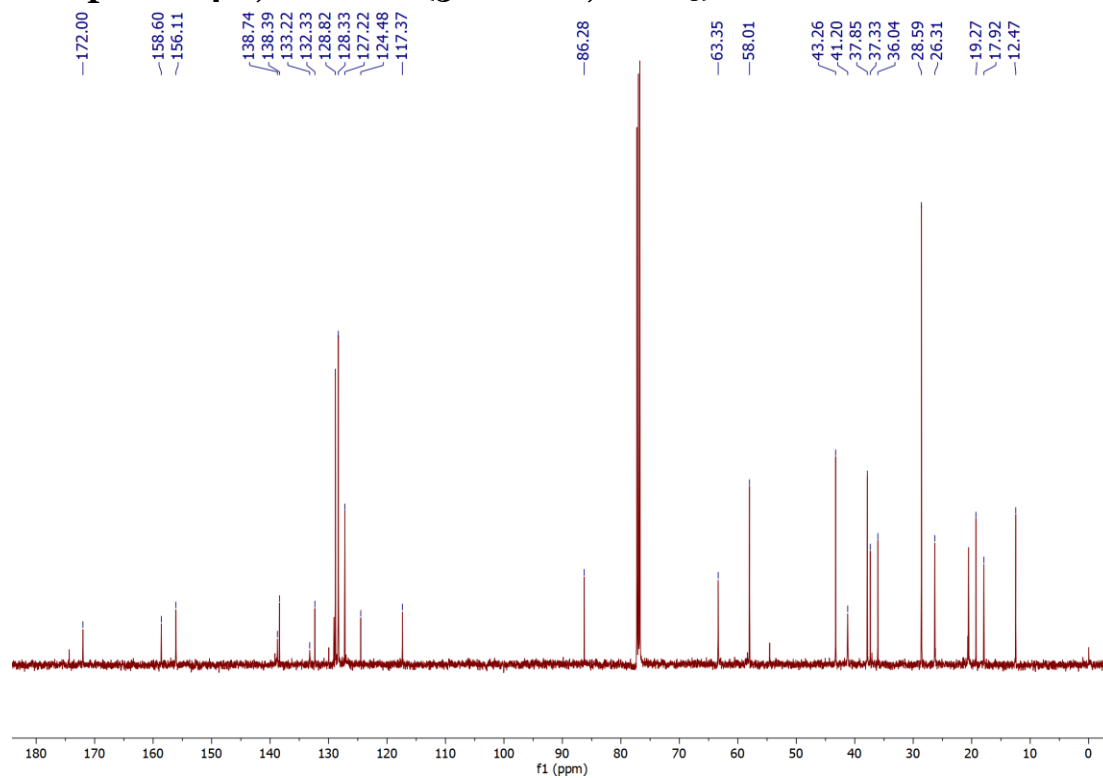
## Compound 4.5, <sup>13</sup>C NMR (500 MHz, CDCl<sub>3</sub>)



**Compound 4.6,  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )**

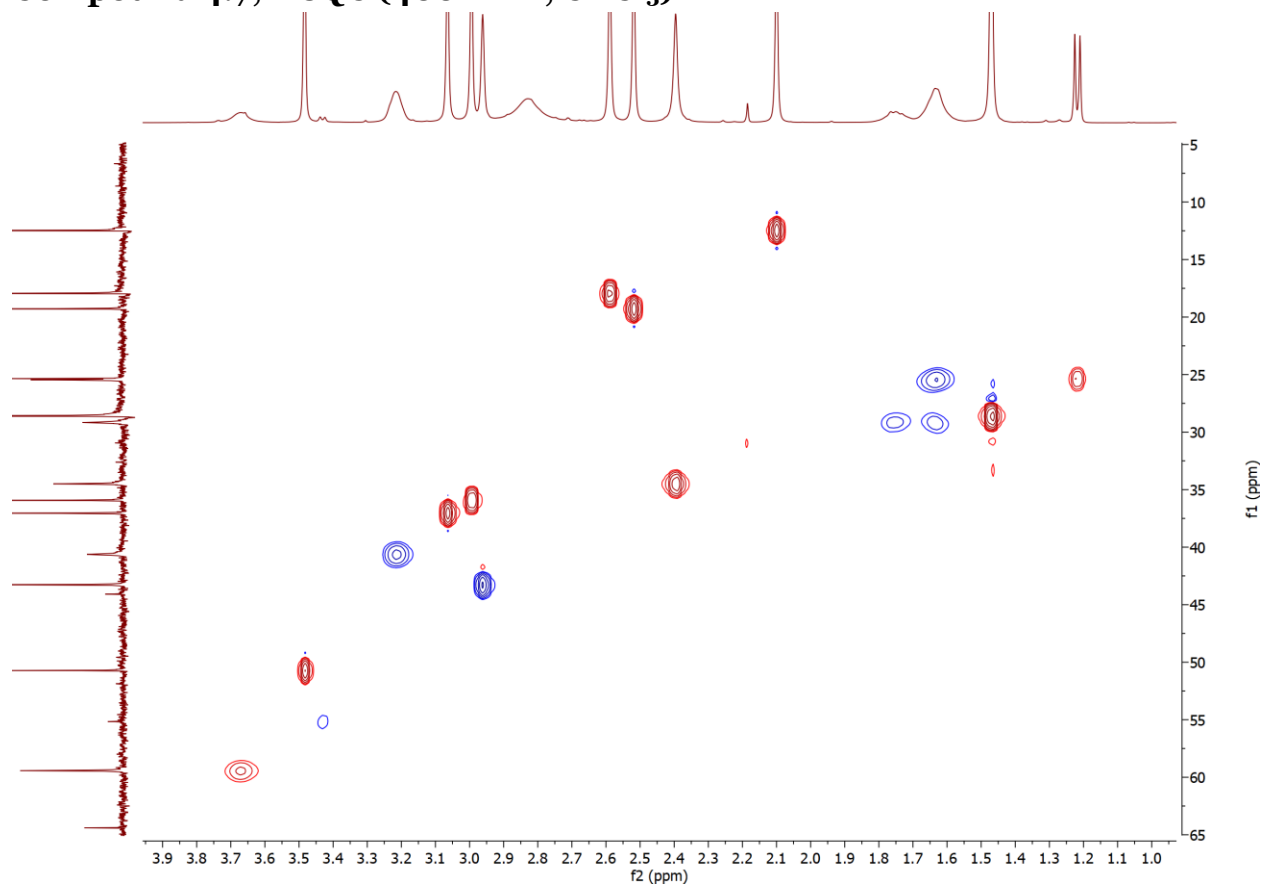


**Compound 4.6,  $^{13}\text{C}$  NMR (500 MHz,  $\text{CDCl}_3$ )**

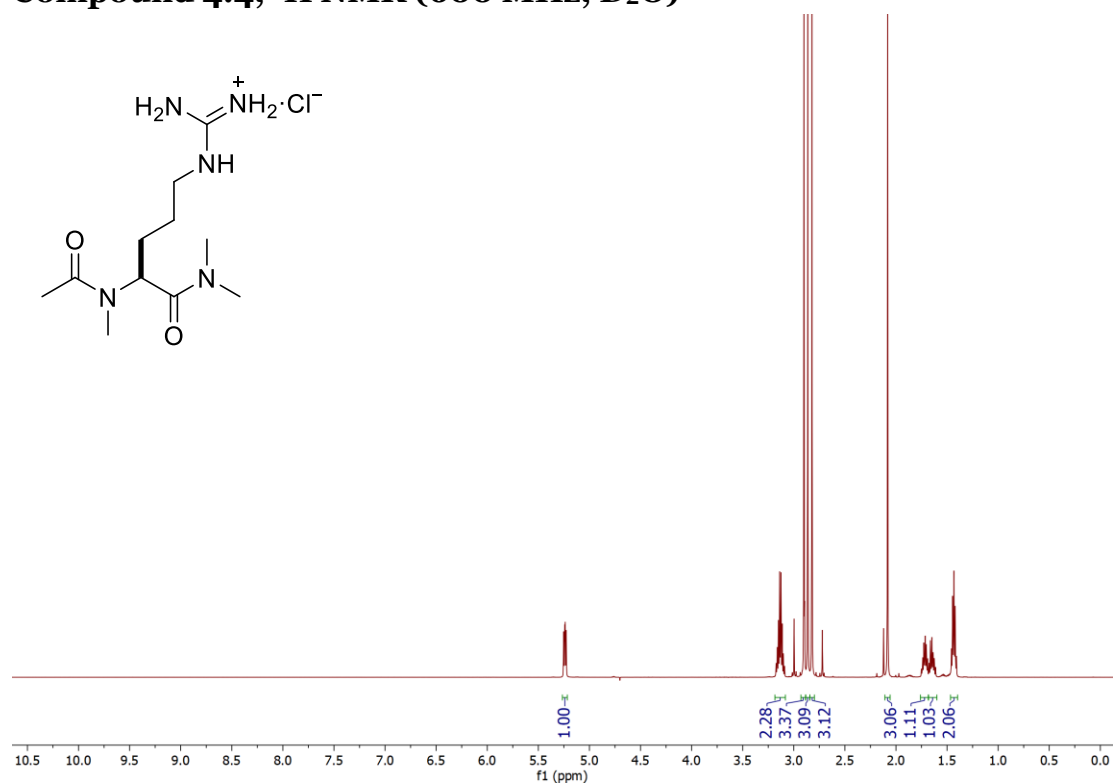




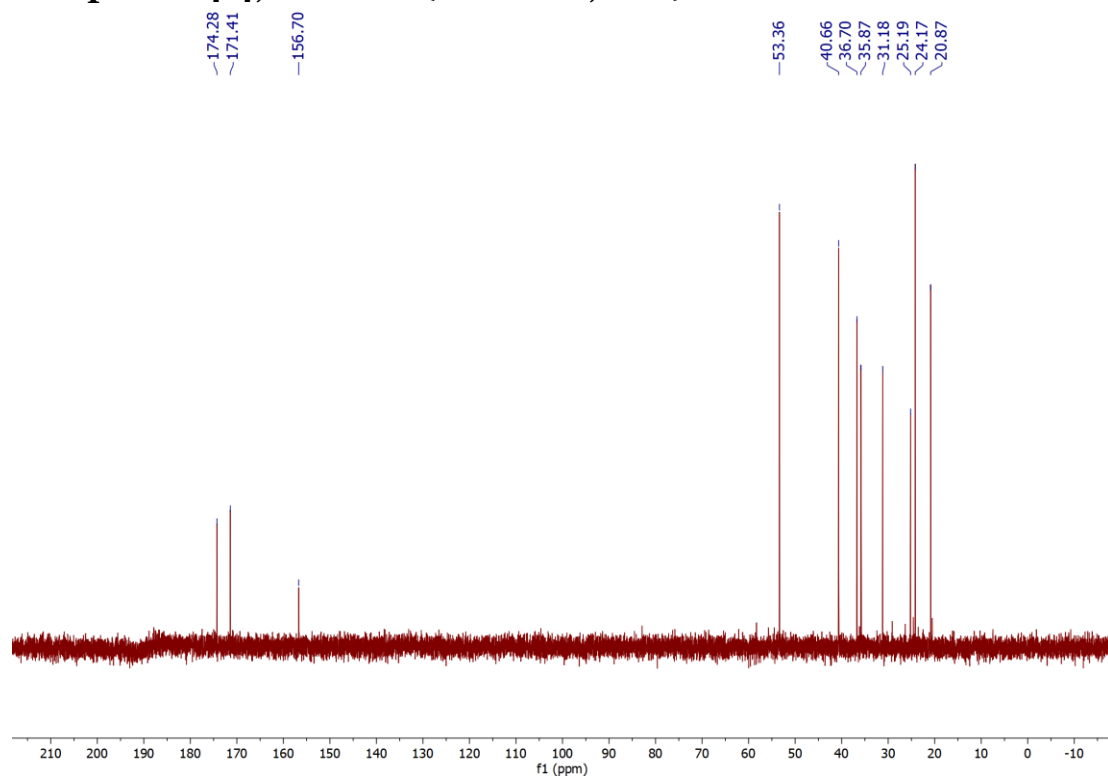
**Compound 4.7, HSQC (400 MHz, CDCl<sub>3</sub>)**



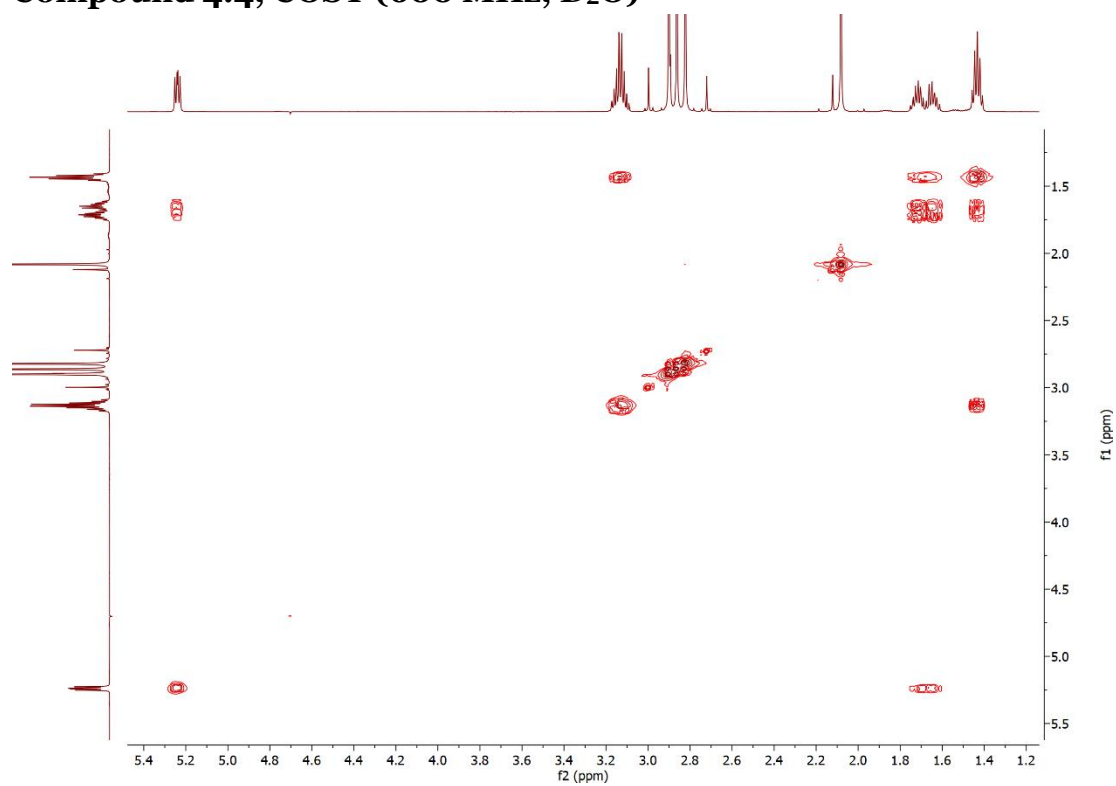
### Compound 4.4, <sup>1</sup>H NMR (600 MHz, D<sub>2</sub>O)



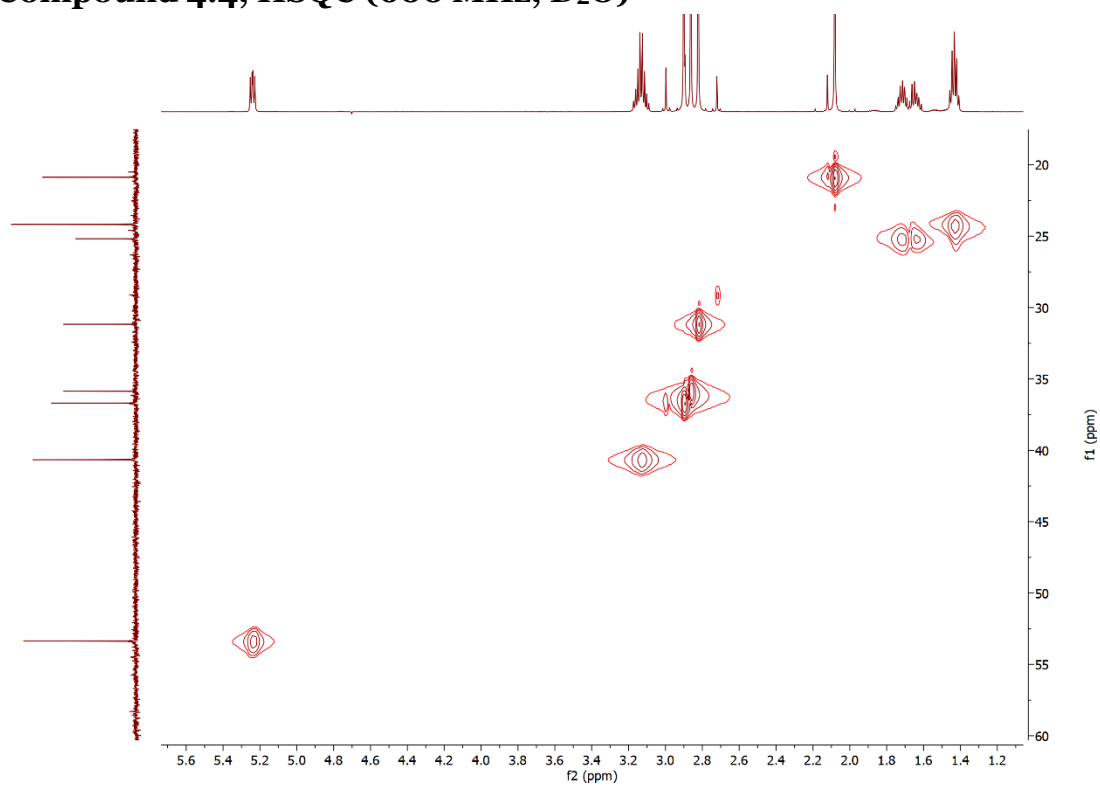
### Compound 4.4, <sup>13</sup>C NMR (600 MHz, D<sub>2</sub>O)



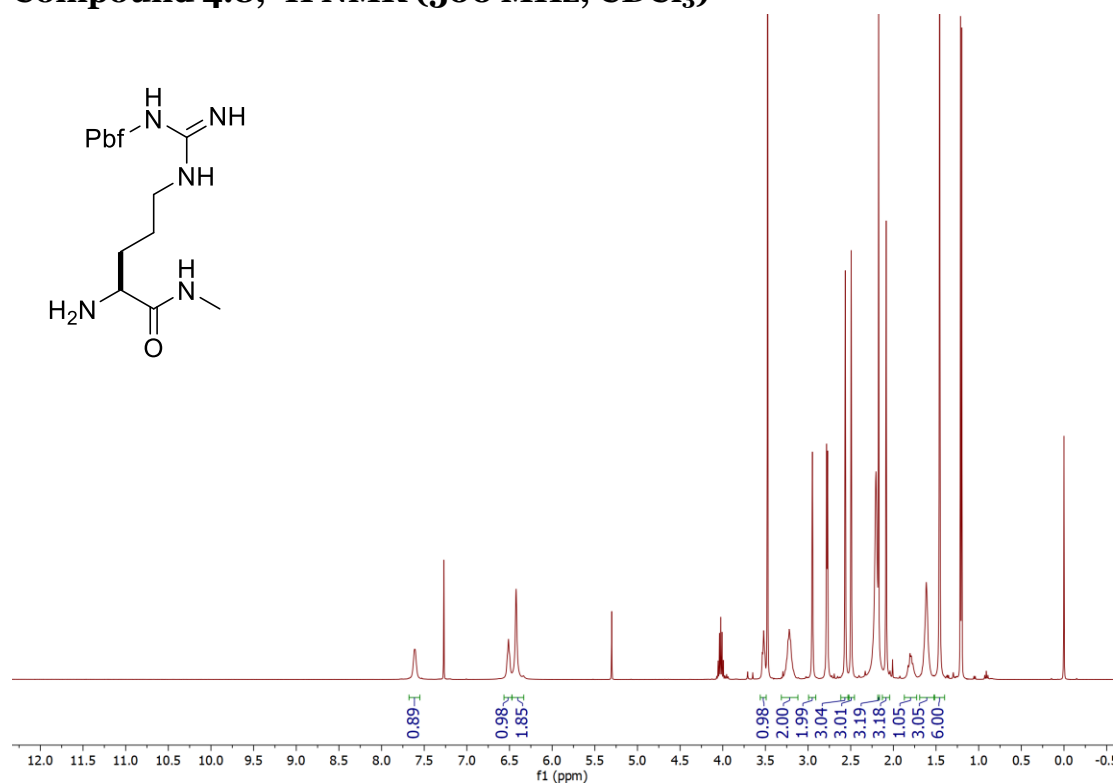
### Compound 4.4, COSY (600 MHz, D<sub>2</sub>O)



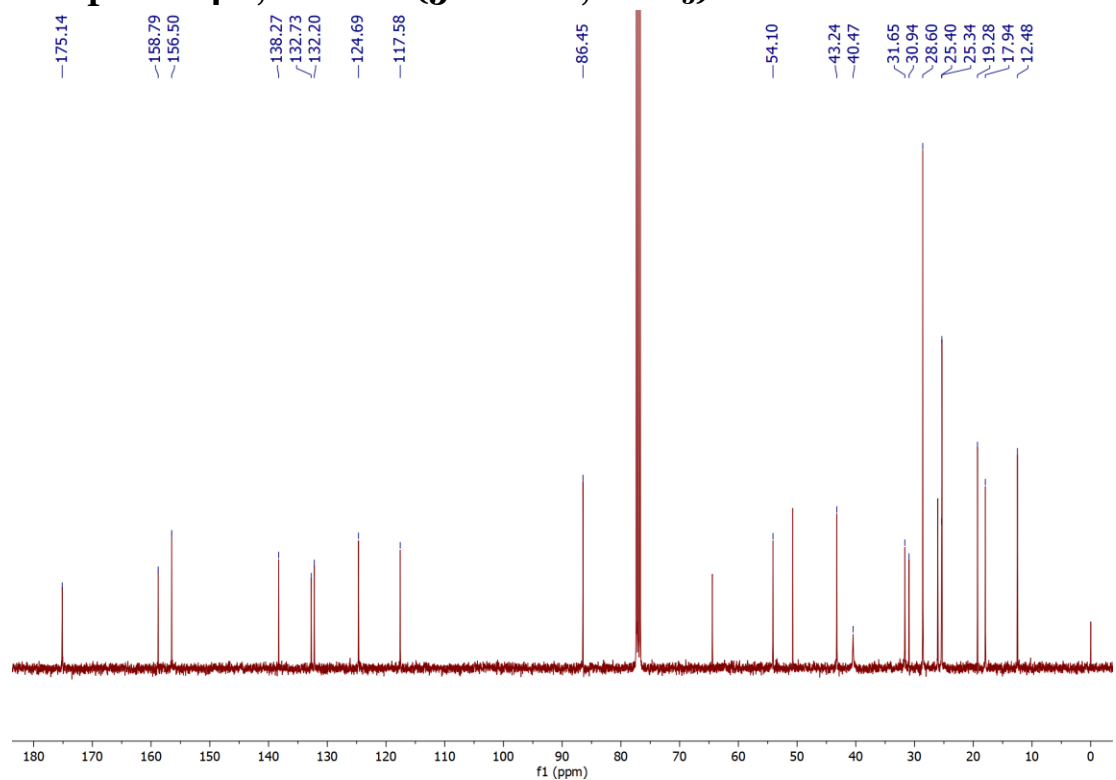
### Compound 4.4, HSQC (600 MHz, D<sub>2</sub>O)



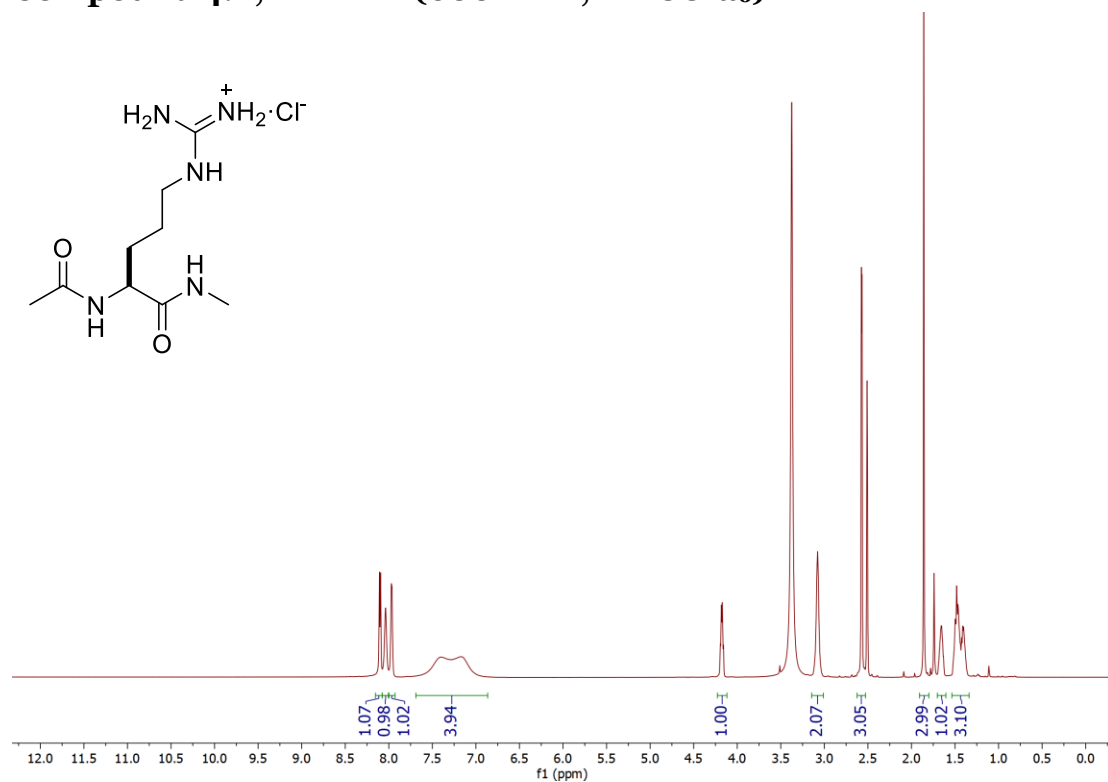
### Compound 4.8, $^1\text{H}$ NMR (500 MHz, $\text{CDCl}_3$ )



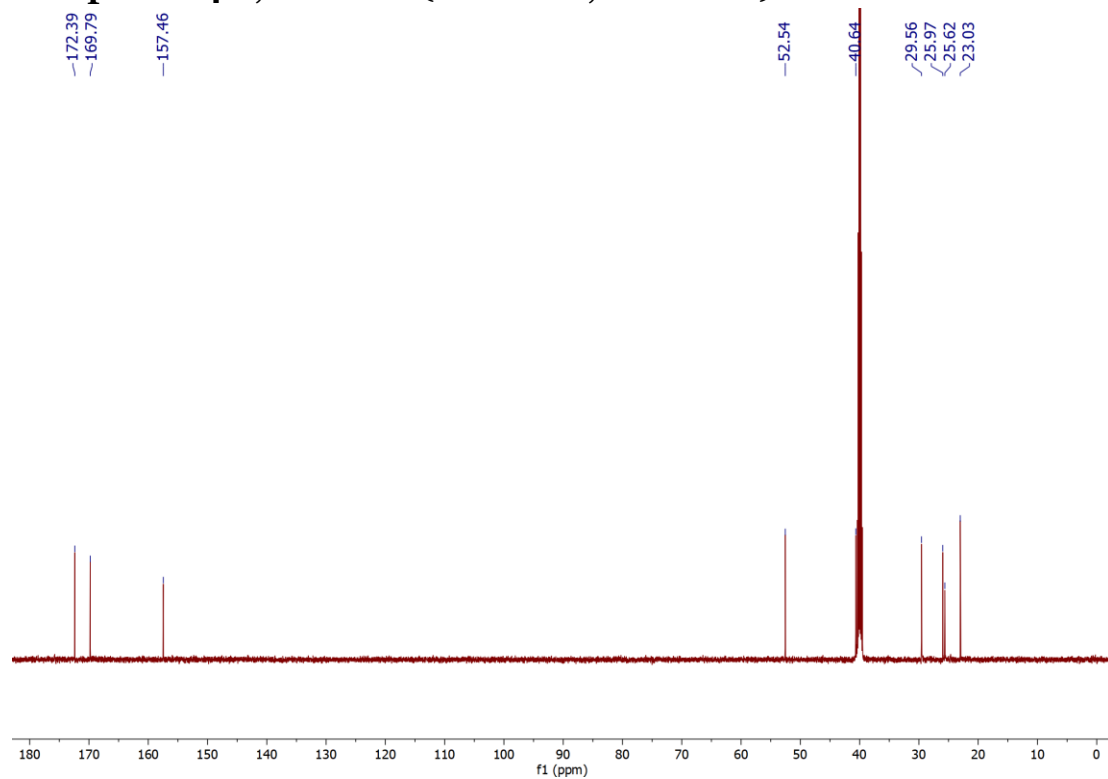
### Compound 4.8, $^{13}\text{C}$ NMR (500 MHz, $\text{CDCl}_3$ )



### Compound 4.2, <sup>1</sup>H NMR (600 MHz, DMSO-d<sub>6</sub>)

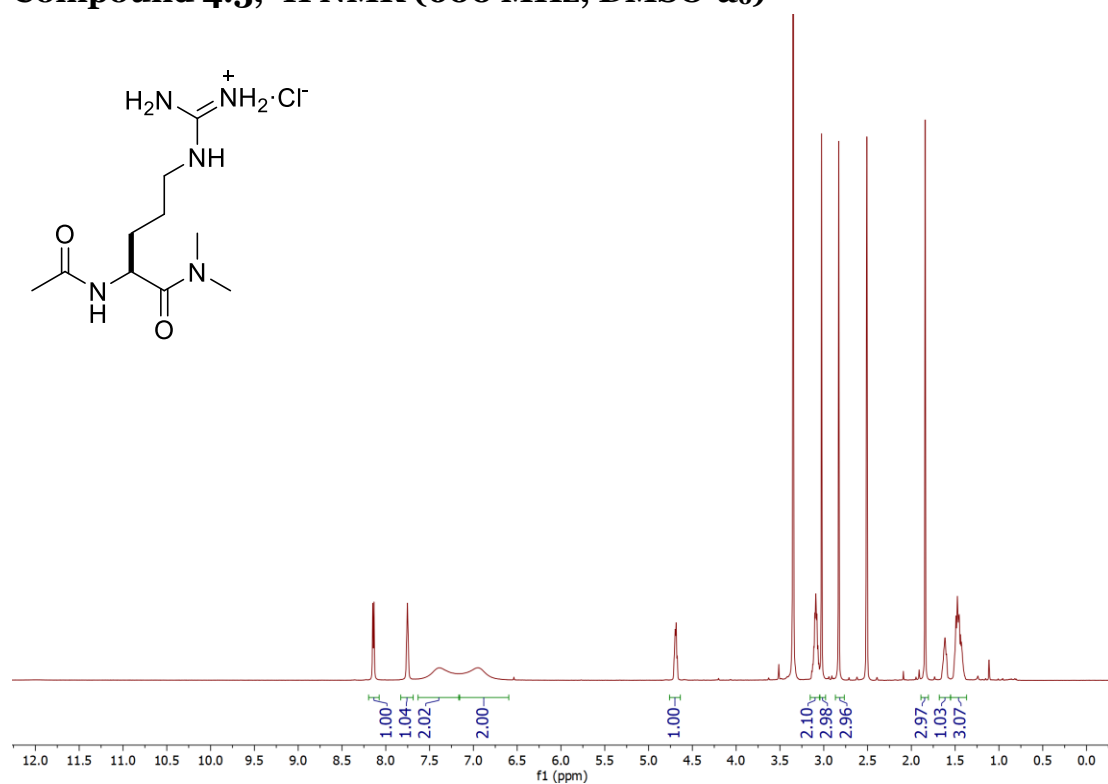


### Compound 4.2, <sup>13</sup>C NMR (600 MHz, DMSO-d<sub>6</sub>)

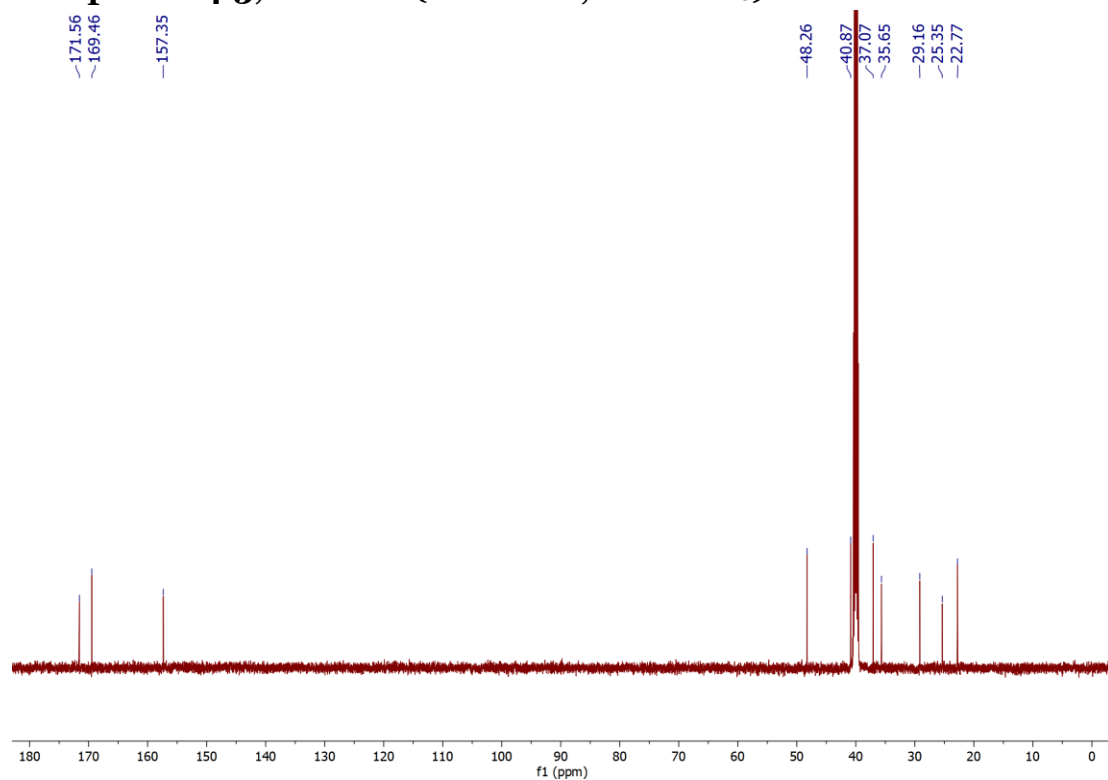




### Compound 4.3, $^1\text{H}$ NMR (600 MHz, $\text{DMSO-}d_6$ )



### Compound 4.3, $^{13}\text{C}$ NMR (600 MHz, $\text{DMSO-}d_6$ )

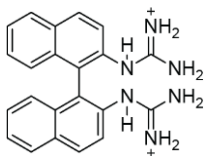


## **Chapter 5**

**Synthesis of biaryl-bisguanidines and forays into  $^1\text{H}$  NMR titration of charged species**

Contributions: All work was performed by Lindsey O. Calabretta.

## Abstract

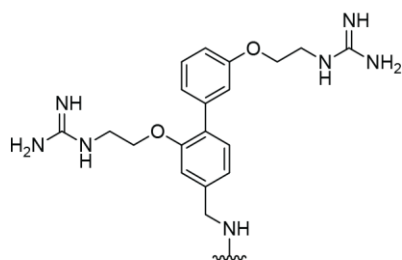


Since the discovery of the axially chiral ligand BINAP, 2,2'-substituted biaryls have been broadly utilized in metal coordination. Inspired by these structures, I designed four biaryl-2,2'-bisguanidines. I envisioned that these structures could be preorganized to bind to oxoanions in a pincer-like fashion. These structures contain either a binaphthyl or biphenyl core, with or without a methylene spacer between the guanidines and the core. These attributes could probe the effects of flexibility on oxoanion binding.

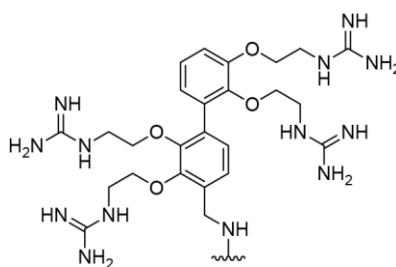
Though the synthesis of these molecules was successful, I encountered significant setbacks in determining their binding to oxoanions through  $^1\text{H}$  NMR binding titrations. Due to the high concentrations required to measure binding by NMR, the signal readouts from these titrations were heavily influenced by the changes in ionic strength imparted by the charged host and guest. Although the data obtained through these studies were unreliable, two future avenues of research were illuminated. For one, oxyguanidines appear to be promising anion binders due to their low  $pK_a$ , and have not yet been studied in supramolecular context. In addition, future studies on the oxoanion binding of the biaryl-2,2'-bisguanidines could be successful in solvents such as isopropanol and when observed at lower concentrations via UV-Vis spectrometry. The metal coordination capabilities of these structures could also be of interest.

## Introduction

**Guanidine-oxoanion binding in cell uptake.** As discussed in the previous chapters, cell-penetrating peptides (CPPs) are arginine-rich peptides capable of transporting themselves and pendant cargo into cells. This uptake is mediated by strong, bidentate hydrogen bonds between the guanidino group of arginine and the oxoanions of the cell surface. To increase the uptake of biologics, I have sought methods to improve the interaction between a guanidino-containing compound and oxoanions, hypothesizing that such an improvement would increase cell uptake efficiency.



1.31

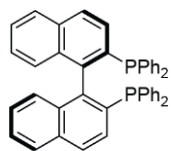


1.32

Schmuck and coworkers have synthesized a guanidine-containing compound with an increased binding affinity to carboxylate due in part to its additional hydrogen bond donors. They observed that it transports a protein, avidin, into cells more effectively than does arginine.<sup>79</sup> Williams, Selwood, and coworkers have studied compounds **1.31** and **1.32** as small molecule carriers capable of transporting protein cargo into cells.<sup>78,92</sup> They proposed that the conformation of these structures mimics the  $\alpha$ -helix of some CPPs due to the twist of the aryl–aryl bond. Although this structure has proven successful at transporting proteins into cells, the structure still relies on the oxoanion binding of single, unmodified guanidines. One could envision that a similar structure with fewer atoms between the 2,2'-guanidines and biphenyl core could lead to the structure being capable

of binding an oxoanion with two guanidines in a pincer fashion. As currently designed, the spacing between the guanidines and biphenyl core is simply too flexible to make this entropically likely.

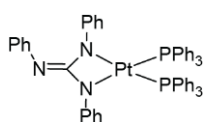
**Biaryl and guanidino coordination in metal complexes.** Biaryl structures have long been used to form coordination complexes due to their unique structure. In 1,1'-binaphthyl scaffolds, the steric repulsion between the naphthyl rings restricts rotation around the aryl-aryl bond, inducing axial chirality. The most notable binaphthyl organometallic ligand is **BINAP** (2,2'-bis(diphenylphosphino)-1,1'-binaphthyl). Noyori and coworkers were the first to synthesize **BINAP** and incorporate it as a ligand for metal ions, first exploring rhodium and later ruthenium catalysts.<sup>169-171</sup> In an early review, Tayaka and Noyori note that while the ligand is axially chiral, the strength of this ligand is its flexibility.<sup>171</sup> The aryl-aryl bond and the C-P bond can rotate slightly to optimize binding to the metal ion and accommodate a reaction substrate. **BINAP** has since been used as a ligand for a range of metal ions (Rh, Ru, Ni, Pd, etc.) and in various asymmetric reactions (hydrogenations, cycloadditions, Heck reactions, etc.).<sup>172</sup>



(R)-BINAP

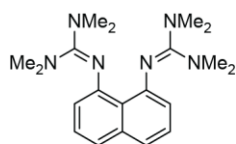
Since the discovery of **BINAP**, other axially chiral biaryl ligands have been explored. For instance, BINOL (1,1'-binaphthyl-2,2'-diol) and BINAM (2,2'-diamino-1,1'-binaphthyl) have been used in metal coordination and molecular recognition.<sup>173,174</sup> Himmel and coworkers have also studied biaryl-bisguanidino scaffolds as ligands of group 10 metals (Ni, Pd, and Pt).<sup>175-177</sup>

The guanidino group has been investigated for its ability to bind to transition metals in its neutral form, monoanionic form, or, more rarely, dianionic form.<sup>178</sup> Dianions of guanidine can chelate metal ions with two nitrogens, as exemplified by the platinum complex **5.1** synthesized by Henderson and coworkers.<sup>179</sup> Organoguanidates have been shown to coordinate metals better than the commonly used amidinate group due to their higher basicity.<sup>180,181</sup> In addition to monoguanidines, many bisguanidines have been explored for their coordinating capabilities.<sup>182</sup>

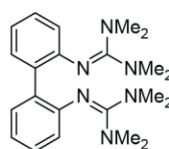


5.1

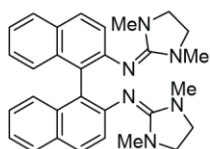
Himmel and coworkers began their journey into bisguanidino metal coordination by synthesizing 1,8-bisguanidinonaphthalene **5.2**.<sup>183</sup> Compound **5.2** was inspired by the prototypical "proton sponge" 1,8-bis(dimethylamino)naphthalene, a scaffold in which the lone pairs of two amines are forced into close proximity via the steric repulsion of the methyl groups.<sup>184</sup> This proximity results in a remarkably basic structure in which both amino groups contribute equally to the trapping of a proton. Compound **5.2** was also found to have proton sponge-like qualities.<sup>185</sup> Himmel and coworkers determined that **5.2** could also chelate Pd and Pt with both guanidino groups.<sup>183</sup>



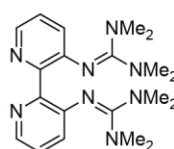
5.2



5.3



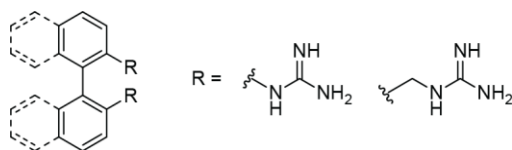
5.4



5.5

Himmel and coworkers subsequently studied the proton sponge and metal coordinating capabilities of biaryl-2,2'-bisguanidines.<sup>175-177</sup> They studied both biphenyl and binaphthyl cores with two *N,N,N',N'*-tetramethylguanidines or *N,N'*-dimethyl-*N,N'*-ethyleneguanidines. Compounds **5.3** and **5.4** are examples of the biaryls studied. They determined that these structures indeed functioned as proton sponges and were capable of forming complexes with group 10 metals using one or both guanidino groups, depending on the conditions. They also explored the bipyridine-bisguanidine **5.5** and found that while protonation occurs on the guanidino groups, compound **5.5** coordinates to metals via the pyridine nitrogens.<sup>177</sup>

**Unsubstituted biaryl-bisguanidines.** Inspired by Himmel and coworkers, I proposed a structure in which two unsubstituted guanidino moieties would be held in close proximity to each other through a rigid biaryl core. To optimize the size and flexibility of the binding pocket, I designed four molecules with increasing flexibility. The choice of biaryl core can control flexibility – a biphenyl compound has more freedom of rotation around its aryl–aryl bond than does a binaphthyl compound. In addition, flexibility can be introduced by the addition of a methylene spacer between the biaryl core and guanidino group (Figure 1).



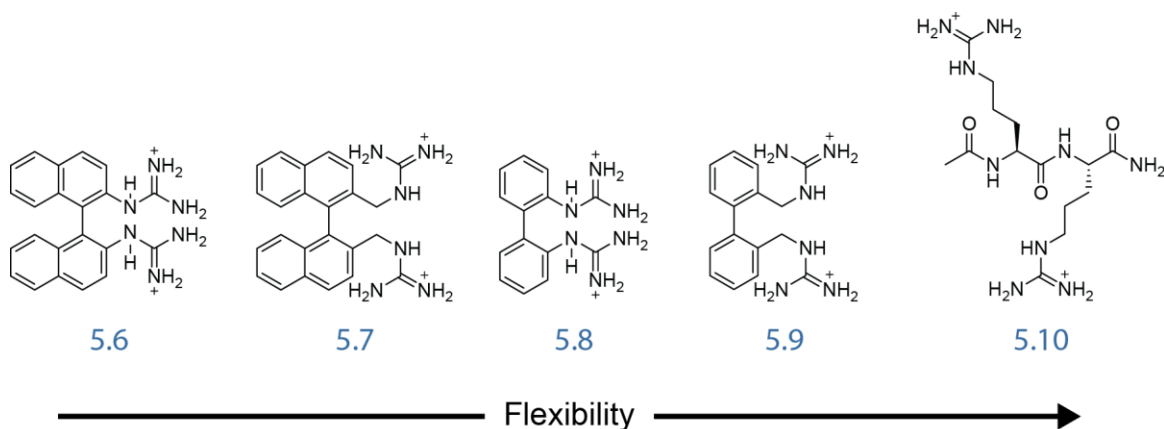
**Figure 5.1.** Biaryl-2,2'-bisguanidines were designed with variable flexibility through the choice of biaryl core and the use of a methylene spacer.

Compared to the per-substituted biaryl-bisguanidines produced by Himmel, this structure could be deprotonated to form guanidates that could chelate the metal ions

with more than two nitrogens or coordinate different metal oxidation states. In addition, these biaryls can interact with oxoanions through hydrogen bonding. In this work, my efforts focused on investigating whether these unsubstituted biaryl-bisguanidines demonstrate improved binding to oxoanions compared to arginine derivatives and monoguanidines.

## Results and Discussion

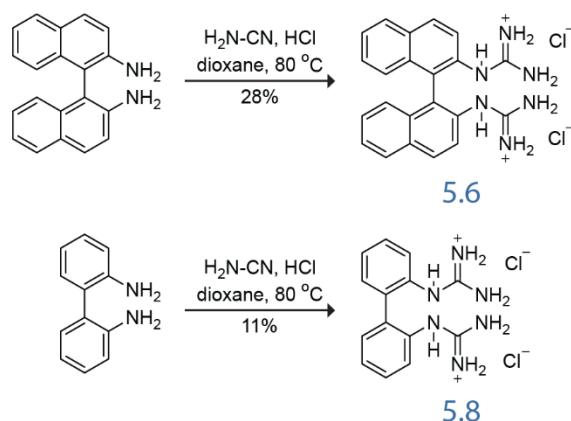
**Synthesis.** I synthesized four biaryl-bisguanidines (**5.6**, **5.7**, **5.8**, and **5.9**) that demonstrate increasing flexibility, with binaphthyl **5.6** as the least flexible and biphenyl-bismethyleneguanidine **5.9** as the most flexible (Figure 5.2). Diarginine **5.10** was also prepared as a bisguanidine analog to CPPs. While the optimal biaryl binding pocket was yet to be determined, I hypothesized that diarginine **5.10** would be significantly worse at oxoanion binding than the biaryls; the spacing and flexibility between the two guanidino groups are simply too great for it to simultaneously engage its two guanidino groups with one oxoanion without substantial entropic cost.



**Figure 5.2.** The biaryl-bisguanidines in order of increasing flexibility, including diarginine **5.10**, a CPP analog, as the most flexible.



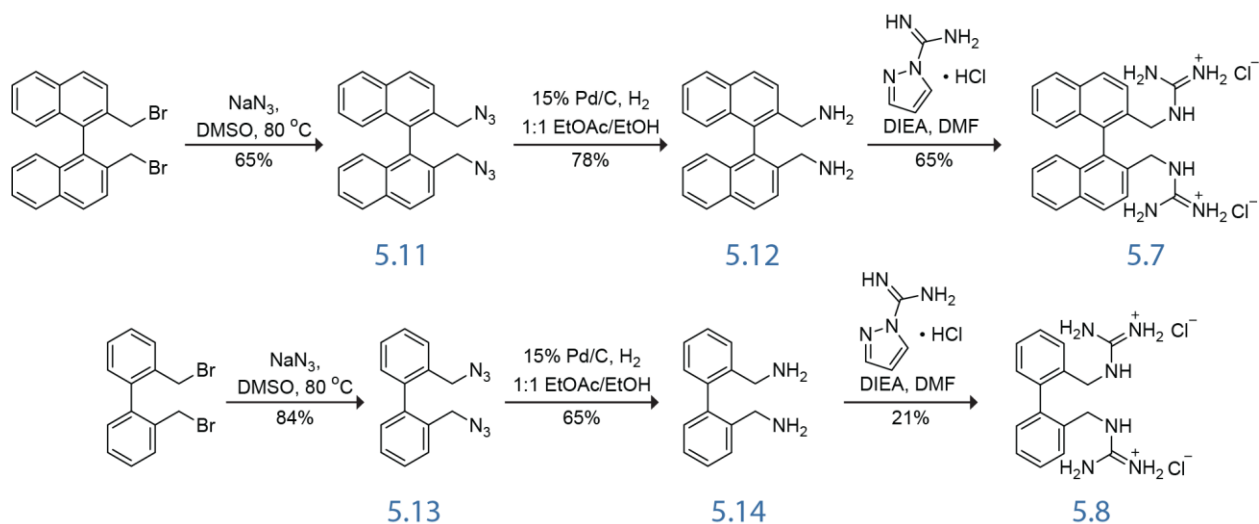
Bisguanidines **5.6** and **5.8** were synthesized by reacting commercially available racemic 2,2'-diamio-biaryls with excess cyanamide under acidic conditions (Scheme 5.1). The use of nitric acid in ethanol produced ethyl carbamimidate as a byproduct that could not be separated from the product. Upon switching to hydrochloric acid in dioxane, **5.6** and **5.8** were obtained in moderate yield with a weak counterion, chloride. Such a non-competitive counterion is desirable for future oxoanion binding applications. The partially reacted 2-amino-2'-guanidinobiaryl side products were more abundant than the desired bisguanidines, but **5.6** and **5.8** were purified successfully.



**Scheme 5.1.** Synthesis of the biaryl-bisguanidines **5.6** and **5.8**.

Diguanidines **5.7** and **5.9** were synthesized by first reacting commercially available 2,2'-bis(bromomethyl)-biaryl compounds with sodium azide to produce diazides **5.11** and **5.13** (Scheme 5.2). I then sought to reduce the azides to amines but encountered challenges with the traditional Staudinger reduction as well as other reduction methods. These methods produce byproducts with masses consistent with an intramolecular cyclization product; this byproduct likely results from the reduction of one azidomethylene to the amine, which subsequently attacks the opposite methylene to release azide. This 7-*exo-tet* cyclization is favored under Baldwin's rules. I found success

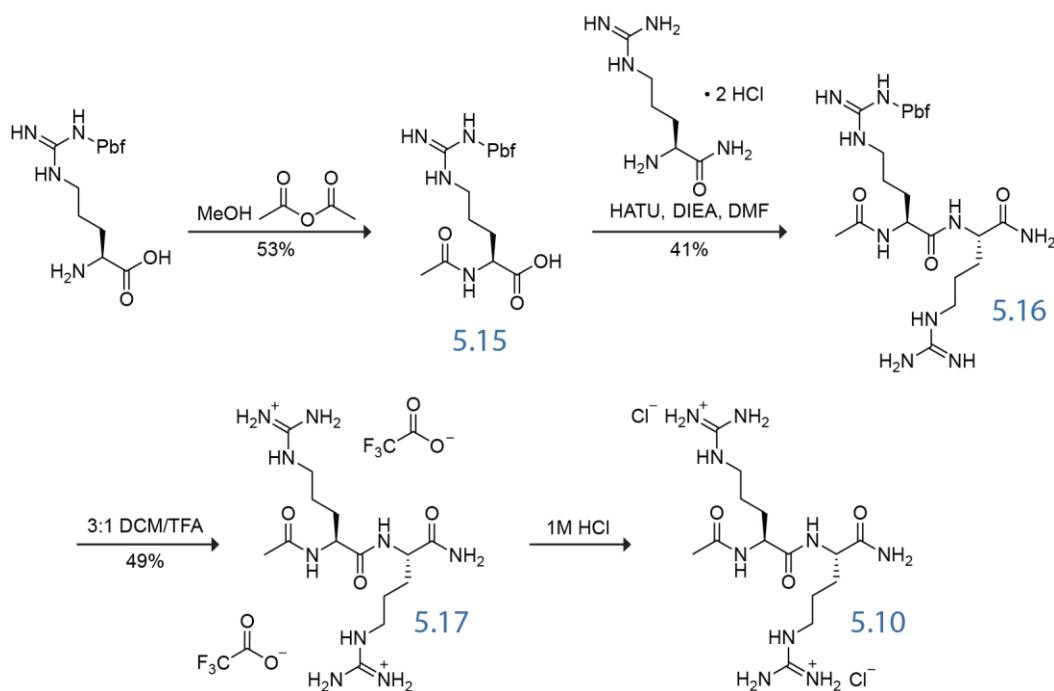
by utilizing hydrogenation conditions with palladium on charcoal, since the reaction proceeded quickly enough to avoid the cyclization byproduct. I subsequently attempted to guanidylate diamines **5.12** and **5.14** using the acidic conditions depicted in Scheme 1. However, the cyclization byproduct was again observed. The cyclization is likely due to the protonation of one aminomethyl group and the subsequent attack of this methylene carbon by the other amino group to release ammonia. I successfully produced diguanidines **5.7** and **5.9** by utilizing 1*H*-pyrazole-1-carboxamide in basic conditions.



**Scheme 5.2.** Synthesis of the biaryl-bismethyleneguanidines **5.7** and **5.9**.

Diarginine **5.10** was synthesized with the help of Vienna M. Thomas (Scheme 5.3). We first acetylated H-Arg(Pbf)-OH using acetic anhydride in methanol to produce Ac-Arg(Pbf)-OH (**5.15**). Compound **5.15** was subsequently coupled with H-Arg-NH<sub>2</sub>·2HCl using the traditional coupling reagent HATU to produce Ac-Arg(Pbf)-Arg-NH<sub>2</sub> (**5.16**). The Pbf group was subsequently cleaved with TFA, providing the desired diarginine as a TFA salt (**5.17**). This salt was converted to the dichloride salt by introducing excess aqueous HCl and lyophilizing to produce diarginine **5.10**. Notably, neither TFA nor HCl

promoted the amide bond cleavage that I observed in Chapters 3 and 4 with canavanine diamide or *N*<sup>α</sup>-methylated arginine diamide (see: p. 133, 153). The stability of diarginine **5.10** could result from its sterics preventing the conformational change required for the intramolecular attack of one carbonyl oxygen on an adjacent carbonyl carbon.



**Scheme 5.3.** Synthesis of the diarginine **5.10**.

In addition to the diguanidino compounds (**5.6–5.10**), I synthesized or purchased a range of monoguanidino compounds as control molecules for binding affinity experiments. Phenylguanidine **5.18** HCO<sub>3</sub><sup>-</sup> salt was purchased and converted to the Cl<sup>-</sup> salt by the addition of aq. HCl. Benzylguanidine **5.19** Cl<sup>-</sup> salt was purchased and used as-is. The 3-phenyl-1-propylguanidine **5.20** was synthesized from 3-phenyl-1-propylamine and 1*H*-pyrazole-1-carboxamide using conditions similar to those used to synthesize **5.7** and **5.9**. Benzyloxyguanidine **5.21** was synthesized from *O*-benzylhydroxylamine and cyanamide using conditions similar to those used to synthesize biaryl-bisguanidines **5.6**

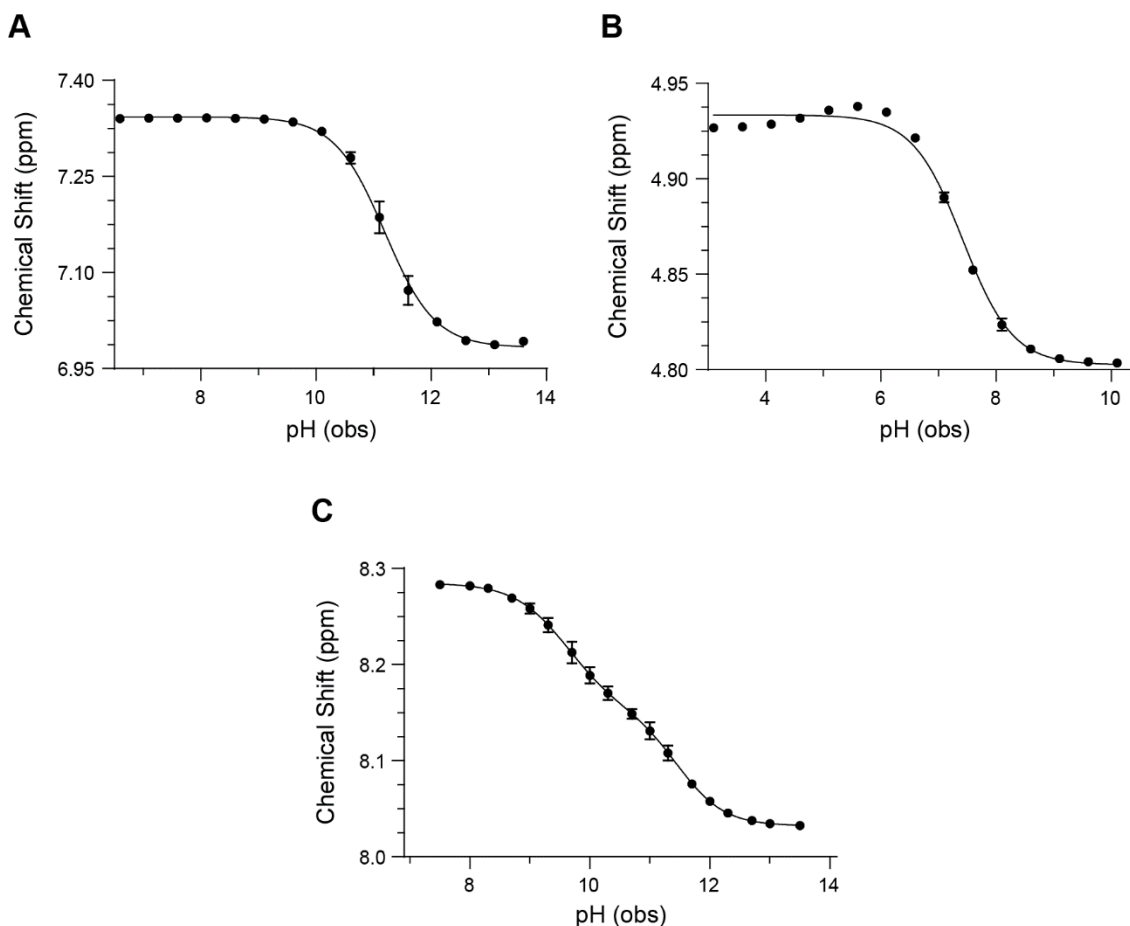
and **5.8**. Arginine diamide **5.22** was purchased and used as an acetate salt because it was prone to amide hydrolysis in aqueous HCl. (The guanidino resin used in Chapters 3 and 4 had not yet been produced.)

**pK<sub>a</sub> titrations via <sup>1</sup>H NMR.** As discussed in Chapters 1–3, a decrease in the pK<sub>a</sub> of a guanidino group can correlate to an increase in that guanidino compound's affinity for oxoanions. The literature pK<sub>a</sub> value of phenylguanidine is 10.88.<sup>110</sup> I hypothesized that the binding affinity of conjugated bisguanidines **5.6** and **5.8** may be enhanced by a lower pK<sub>a</sub> compared to that of the alkylguanidines of **5.7** and **5.9**. To further explore the effects of pK<sub>a</sub>, benzyloxyguanidine **5.21** was synthesized. I expected its pK<sub>a</sub> to be similar to that of canavanine (pK<sub>a</sub> 7.01, 7.40).<sup>140,141</sup>

The pK<sub>a</sub>'s of compounds **5.6** and **5.18–5.22** were determined by dissolving the guanidino compound in D<sub>2</sub>O and adjusting the pH of the solution using DCl and NaOD, taking aliquots at regular intervals. The inflection point was determined with the software GraphPad Prism. Performing the titration in D<sub>2</sub>O instead of water influences the observed pH and protonation (deuteration) rate; therefore, the resulting pK<sub>a</sub> values were corrected using the equation developed by Krężel and Bal (Figure 5.3).<sup>125</sup>

$$pK_a = 0.929pK_a(\text{obs}) + 0.42 \quad \text{Eq. 2.1}$$

The pK<sub>a</sub> of phenylguanidine **5.18** was found to be 10.82 ± 0.02, and the pK<sub>a</sub> of benzyloxyguanidine **5.21** was found to be 7.28 ± 0.03. The pK<sub>a</sub>'s of alkylguanidines **5.19**, **5.20**, and **5.22** were too high to measure using this method – I did not observe an asymptote at high pH that would indicate full deprotonation. Therefore, the pK<sub>a</sub>'s of these molecules are assumed to be greater than 13.



**Figure 5.3.** Graphs displaying the  $pK_a$  titration data and fits of aryl- and oxyguanidines. A. Phenylguanidine **5.18**. B. Benzyloxyguanidine **5.21**. C. Binaphthyl-bisguanidine **5.6**.

Since bisguanidine **5.6** has two basic functional groups, I expected to observe two  $pK_a$ 's. I hypothesized that the  $pK_{a1}$  would be significantly depressed because the Coulombic interactions between two guanidinium ions would be disfavored. Indeed, I determined that compound **5.6** has a  $pK_{a1}$  of  $9.36 \pm 0.1$  and a  $pK_{a2}$  of  $11.08 \pm 0.09$ . Interestingly, the  $pK_{a2}$  is slightly higher than the  $pK_a$  observed for phenylguanidine **5.18**. The increase in  $pK_{a2}$  could indicate that compound **5.6** is capable of acting as a proton sponge, based on the proton sponge capability of similar structures described by Himmel and coworkers.<sup>177</sup> Alternatively, the proximity of the two guanidino groups could be

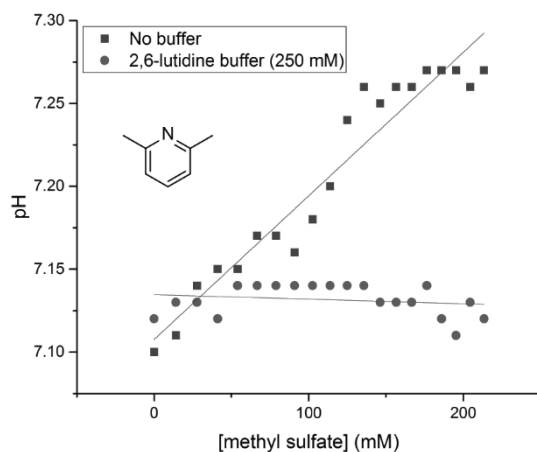
forcing them out of the plane of the naphthyl rings, disrupting the conjugation between the guanidine and naphthyl and hindering the electron-withdrawing effect that lowers  $pK_{a2}$ .

### **Binding affinity titrations via $^1\text{H}$ NMR in water.**

**Buffer optimization.** To analyze the binding affinity of the guanidino compounds to oxoanions, I planned to perform binding affinity titrations in which I would observe the change in  $^1\text{H}$  NMR signal from a guanidine host with increasing concentrations of anion guest. These titrations would be done in water ( $\text{D}_2\text{O}$ ) to be relevant to future biological applications.

Because the titrations would be done in water, a buffer was needed to ensure that any changes in chemical shift of the guanidino compound were the result of a binding event and not changes in pH. Classical buffers such as phosphate-buffered saline (PBS) and 2-(*N*-morpholino)ethanesulfonic acid (MES) would be inappropriate, given that they contain oxoanions that would compete with the guest. A buffer such as tris(hydroxymethyl)aminomethane (TRIS) would also interfere with the titration since a positively charged amine could compete with the host to bind to the guest.

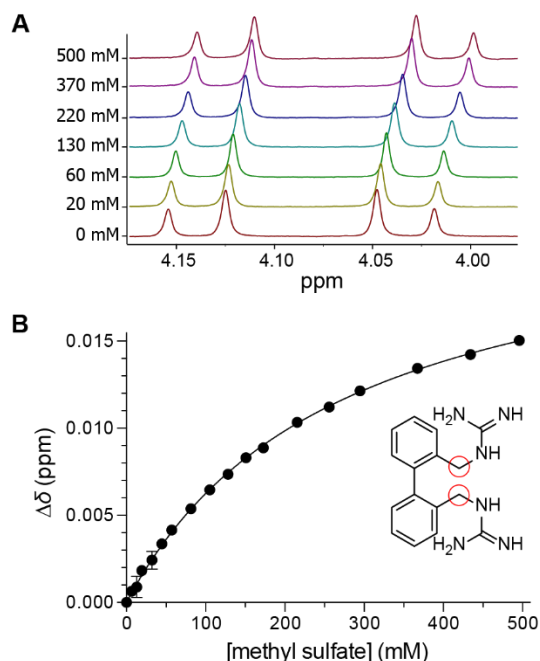
I identified 2,6-lutidine as an appropriate buffer for these experiments. This pyridine derivative has a  $pK_a$  of 6.72, offering buffering capacity at a biological pH of 7.4.<sup>186</sup> The two methyl groups *ortho* to the pyridino nitrogen also sterically prevent the competitive coordination of oxoanions. I tested the buffering capacity of 2,6-lutidine by preparing a 250 mM solution and incrementally introducing an oxoanion (methyl sulfate). While the pH changed in pure water upon addition of the oxoanion, 2,6-lutidine succeeded in maintaining a constant pH (Figure 5.4).



**Figure 5.4.** The buffering capacity of 2,6-lutidine was determined by observing the change in pH of aqueous solutions with and without 2,6-lutidine upon the addition of methyl sulfate.

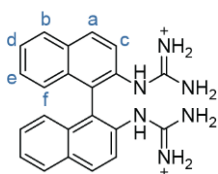
**Affinity titration results.** Binding titrations<sup>112</sup> were performed on compounds **5.6**–**5.10** and **5.18**–**5.22** against methyl sulfate in aqueous solutions. In this case, methyl sulfate (as a sodium salt) acts as an analog of heparan sulfate, a cell-surface proteoglycan implicated in the translocation mechanism of CPPs.<sup>16</sup> The guanidino compounds (5 mM) were dissolved in 2,6-lutidine buffer (500 mM). Stock solutions of methyl sulfate were made with the guanidine stock solution to maintain the concentration of guanidine throughout the experiment. The guanidine solution was placed into an NMR tube, aliquots of methyl sulfate (0 to 500 mM final concentration) were sequentially added, and a <sup>1</sup>H NMR spectra was acquired after each addition. The change in chemical shift of the observable protons of the molecules was determined and fitted to a one-site binding model in GraphPad Prism (Figure 5.5).

I observed that different signals from the same molecule would give different binding affinities. To further understand this difference, I performed COSY and selective 1D NOESY on compound **5.6** to fully assign the <sup>1</sup>H NMR peaks (see: p. 222–223).



**Figure 5.5.** Binding affinity titration of bisguanidine **5.8** in D<sub>2</sub>O against methyl sulfate. A. Selected spectra demonstrating the change in chemical shift of the highlighted protons of **5.8**. B. The changes in chemical shift fitted to a binding model.

I observed a general trend that peaks further from the binding site displayed a larger  $K_d$  value than those nearer the binding site (Peak A has a  $K_d$  of 320 mM while Peak D has a  $K_d$  of 450 mM). I decided to use the peak that resulted in the lowest binding affinity for all molecules for comparison.



**Figure 5.6.** The <sup>1</sup>H NMR assignments of bisguanidine **5.6**, corresponding to the signal labels of the <sup>1</sup>H NMR spectrum on p. 222.

The results of these binding titrations were inconclusive. These data were initially used to calculate the dissociation constant ( $K_d$ ), with a lower  $K_d$  indicating a stronger host-guest interaction. Taking the obtained  $K_d$  values at face value, compounds **5.6** and **5.8**

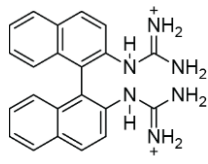
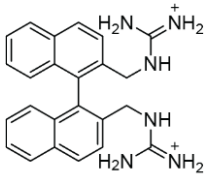
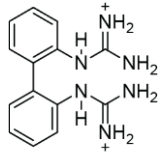
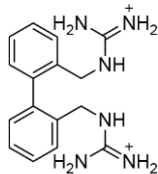
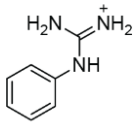
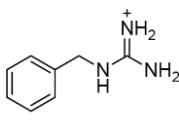
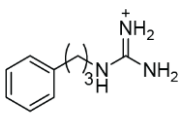
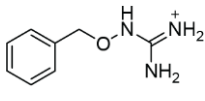
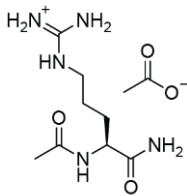
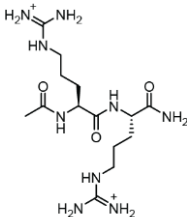
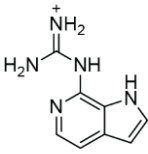


resulted in a  $K_d$  similar to phenylguanidine **5.18**. One would expect that if **5.6** and **5.8** were using two guanidines to bind to a single anion, then their  $K_d$  values should be lower than that of **5.18**. Thus, it appears that **5.6** and **5.8** could be too conformationally constrained to bind anions in a multi-dentate fashion. Similarly, the binaphthyl-bisguanidine with a methylene spacer, **5.7**, shows a binding affinity similar to benzylguanidine **5.19**. Only bisguanidine **5.9** is noticeably improved from its monoguanidine counterpart, **5.19**.

Oxyguanidine **5.21** also shows a noticeably lower  $K_d$  value than do the aryl and alkyl amines. Its lower  $pK_a$  can explain this disparity. However, arginine derivatives **5.22** and **5.10** also show a lower  $K_d$  value than do the aryl or alkyl guanidines. This result is unexpected, as one would predict them to demonstrate a binding affinity similar to that of an alkylguanidine, such as **5.19**. This result is not due to less steric interference between the amino acids and the methyl sulfate guest, since propylguanidine **5.20** contains similar spacing between the guanidine and phenyl 'backbone' as **5.22**, yet still displays a much worse binding affinity than **5.22**.

If the data are to be trusted, these results are mystifying. However, these affinity values may be too weak to be meaningful. In fragment-based drug-discovery, low mM  $K_d$  values for small molecule fragments to proteins are trustworthy enough to guide the drug-discovery process.<sup>187,188</sup> However, the binding energy difference between, say,  $K_d = 3$  mM and  $K_d = 300$  mM is drastic. To determine  $K_d$  values, I used the software GraphPad prism. Seeking to verify these results, I also ran the data through the web-hosted software on supramolecular.org, which uses more accurate binding models and reports binding affinities in  $K_a$ , the reciprocal of  $K_d$ .<sup>111–113</sup> The resulting  $K_a$  values of **5.6–5.9** and **5.18–5.20** were similar to one another and did not follow the same trends as observed in the

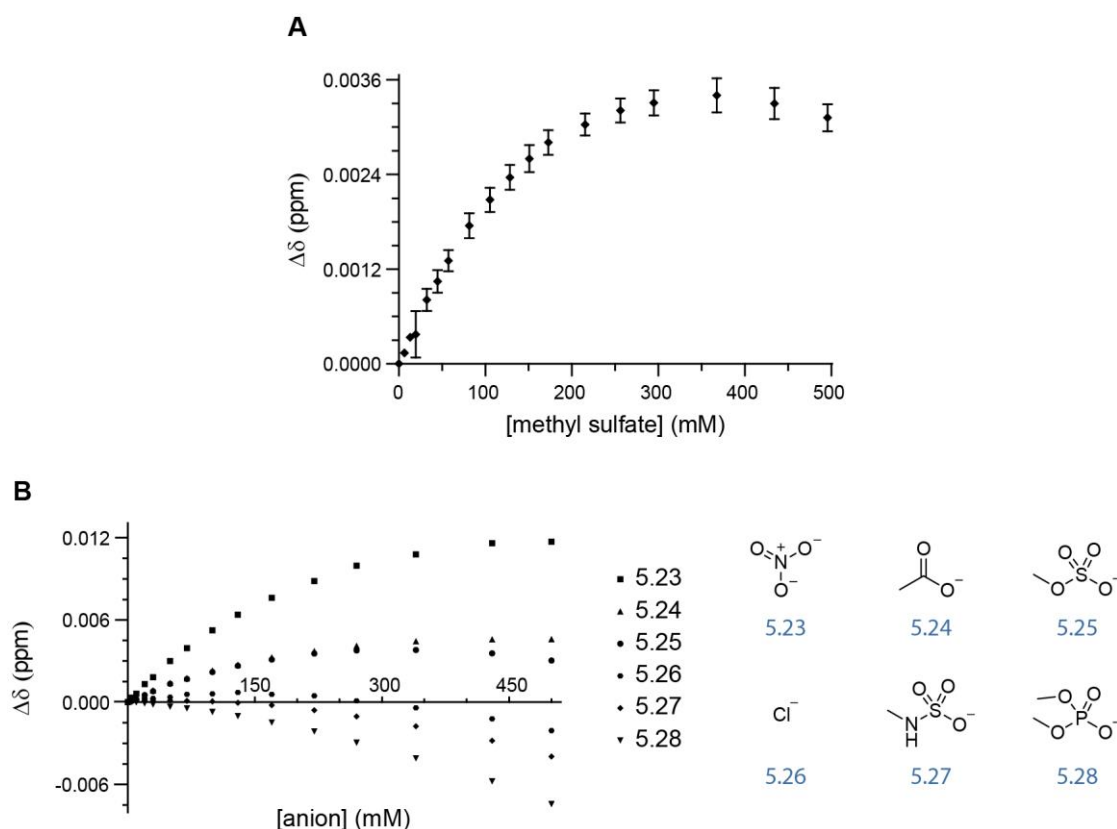
**Table 5.1.** Binding affinities of all compounds against methyl sulfate in D<sub>2</sub>O. Compounds are chloride salts if not otherwise shown. p*K*<sub>a</sub>'s of compounds were included if observed experimentally. *K*<sub>d</sub>'s were calculated using GraphPad Prism; data points were excluded if they reversed direction at high concentrations of methyl sulfate. *K*<sub>a</sub>'s were calculated with supramolecular.org;<sup>111–113</sup> no values were excluded.

				
	<b>5.6</b>	<b>5.7</b>	<b>5.8</b>	<b>5.9</b>
<i>K</i> <sub>d</sub>	330 ± 13 mM	350 ± 10 mM	330 ± 19 mM	270 ± 7 mM
<i>K</i> <sub>a</sub>	3.09 ± 0.06 M <sup>-1</sup>	3.06 ± 0.09 M <sup>-1</sup>	2.93 ± 0.03 M <sup>-1</sup>	3.76 ± 0.05 M <sup>-1</sup>
p <i>K</i> <sub>a</sub>	9.36 11.08	-	-	-
				
	<b>5.18</b>	<b>5.19</b>	<b>5.20</b>	<b>5.21</b>
<i>K</i> <sub>d</sub>	300 ± 22 mM	380 ± 8 mM	420 ± 32 mM	170 ± 18 mM
<i>K</i> <sub>a</sub>	3.38 ± 0.11 M <sup>-1</sup>	3.76 ± 0.05 M <sup>-1</sup>	2.44 ± 0.04 M <sup>-1</sup>	8.88 ± 0.80 M <sup>-1</sup>
p <i>K</i> <sub>a</sub>	10.82	<13	<13	7.28
				
	<b>5.22</b>	<b>5.10</b>	<b>GAI</b>	
<i>K</i> <sub>d</sub>	160 ± 13 mM	130 ± 13 mM	200 ± 22 mM	
<i>K</i> <sub>a</sub>	8.78 ± 0.56 M <sup>-1</sup>	11.70 ± 0.99 M <sup>-1</sup>	6.07 ± 0.34 M <sup>-1</sup>	
p <i>K</i> <sub>a</sub>	<13	-	9.43	

*K*<sub>d</sub> values. Still, oxyguanidine **5.21** and amino acids **5.10** and **5.22** display a stronger binding affinity than do the other compounds tested.

Additionally, the observed *K*<sub>a</sub> values observed were ~10 M<sup>-1</sup> or less. Frequently, if an interaction between a host and guest is too weak to measure, authors will express the

$K_a$  as  $<10\text{ M}^{-1}$ , suggesting a general consensus that values below  $10\text{ M}^{-1}$  are not accurately measurable.<sup>30,189</sup> And yet,  $K_a$  values below  $10\text{ M}^{-1}$  are still frequently reported. Houk and coworkers published a meta-study comparing the binding affinities of many host–guest and protein–ligand complexes.<sup>190</sup> They determined that synthetic cavitand hosts (macrocycles similar to cyclodextrins) have a statistical  $K_a$  value of  $10^{3.4\pm 1.6}\text{ M}^{-1}$  in aqueous solvents and  $10^{2.2\pm 1.6}\text{ M}^{-1}$  in organic solvents. Therefore, a fair number of  $K_a$  values less than  $10\text{ M}^{-1}$  have been reported. While the low  $K_a$  of compounds **5.6–5.10** and **5.18–5.22** does not necessarily condemn our data, the data would certainly be more believable if the  $K_a$  differences between the molecules were greater.



**Figure 5.7.** Changes in ionic strength of the titration solutions at high concentrations of anions results in unfittable curves. A. Arginine **5.22** against methyl sulfate in  $\text{D}_2\text{O}$ . The final two data points at the highest concentrations of anion reverse direction. B. Anions **5.23–5.28** (sodium salts) were screened against benzylguanidine **5.19**, with some anions displaying aberrant curves.

I also encountered a frustrating artifact in several titrations, particularly titrations that produced lower  $K_a$  values. In some experiments at the highest sulfate concentrations, I observed the change in chemical shift reverse (Figure 5.7A). Initially, I excluded these values from the binding fits performed in GraphPad Prism as I assumed they were an artifact that occurred post-saturation. I later performed a screen of benzylguanidine **5.19** against several different anions. Some anions behaved as expected (**5.23–5.25**), but some did not fit with any binding models explored (**5.26–5.28**) (Figure 5.7B).

These results were confounding, and I sought alternative explanations for the unusual chemical shifts changes. In 2018, Huber, Erdelyi, and coworkers reported aberrant curves resulting from a titration they performed to quantify a halogen bond between an imidazole cation host and bromide anion guest.<sup>191</sup> They initially observed the chemical shift of protons on the imidazole decrease in chemical shift, and then reverse direction and increase. Opposing chemical shift changes are not unheard of, often indicating a 1:2 host-guest binding stoichiometry. However, the authors did not believe their system would produce a 1:2 stoichiometry.

Instead, they believed the increasing ionic strength of the solution due to the rising concentration of ionic guest throughout the experiment influenced the observed chemical shifts of the host. Therefore, they redesigned their experiment to maintain a constant total concentration of ions in the system. To do this, they produced stock solutions of equimolar concentrations of the host and guest (both as salts with non-coordinating counterions). They then prepared titration samples of equal volume from these solutions but varied the ratio between the host and guest stock solution. From these experiments, they obtained data that resembled a binding curve with a typical 1:1 stereochemistry.

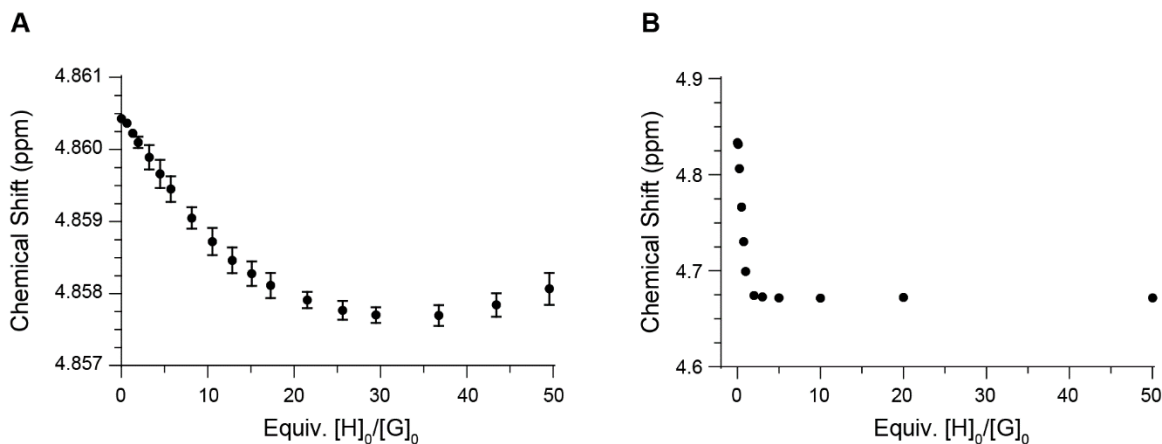
Due to the high concentrations required to obtain an NMR signal compared to other reporting methods such as fluorescence spectroscopy, NMR titration is particularly susceptible to interference by changes in ionic strength. In addition, the binding interactions between the guanidines studied in this section and methyl sulfate in aqueous conditions are weak, requiring even higher concentrations of ionic guest be introduced. In combination, these factors result in the data displayed in Table 1 being untrustworthy.

**Constant ion concentration titrations via  $^1\text{H}$  NMR in DMSO.** I subsequently attempted to improve the titration conditions using the constant ion concentration method developed by Huber, Erdelyi, and coworkers.<sup>191</sup> I initially optimized the process on molecules that were commercially available or easy to synthesize (compounds **5.18**, **5.19**, and **5.21**). I also pivoted from using methyl sulfate as the guest to using acetate. Due to acetate's higher  $\text{p}K_{\text{a}}$ , it should be a better hydrogen bond acceptor and demonstrate stronger binding to guanidines. Lastly, I switched solvents from  $\text{D}_2\text{O}$  to the less competitive  $\text{DMSO-}d_6$ . DMSO is aprotic and has a lower dielectric constant ( $\epsilon = 46$ ) than does water ( $\epsilon = 80$ ).

The binding titrations were performed by producing stock solutions of equimolar concentrations of the guanidine compound and acetate (ammonium or tetramethylammonium salt) in  $\text{DMSO-}d_6$ . Proportional volumes of each stock solution were used to create sample solutions with zero to fifty equivalents of guest to host. The data were analyzed with the software hosted by [supramolecular.org](http://supramolecular.org).<sup>111</sup>

I observed a dramatic improvement in data quality using this optimized method. As shown in Figure 5.8, the curve produced using the constant ion concentration method (B) looks much more like a traditional 1:1 binding stoichiometry compared to the original

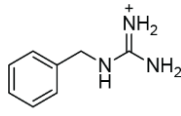
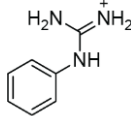
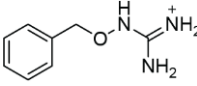
data (A). The change in chemical shift is also much greater. The resulting  $K_a$  of oxyguanidine **5.21** for acetate was much higher than the  $K_a$  obtained using the original method (Table 5.2).



**Figure 5.8.** A comparison of the binding affinity curves using a traditional titrant addition method versus the constant ion concentration method. A. Oxyguanidine **5.21** against methyl sulfate in  $D_2O$  using the traditional gradual addition of guest. B. Oxyguanidine **5.21** against ammonium acetate in  $DMSO-d_6$  using the constant ion concentration method. In both, the chemical shift of the methylene hydrogens is reported.

The data in Table 5.2 coincide with expectations that the binding affinity of guanidine to acetate increases with a decrease in guanidino  $pK_a$ . The data for the guanidino compounds against ammonium acetate shows that the  $K_a$  for **5.21** ( $pK_a = 7.28$ ) is greater than for **5.18** ( $pK_a = 10.82$ ), which is, in turn, greater than for **5.19** ( $pK_a > 13$ ). The same trend holds for the titrations against tetramethylammonium acetate. Interestingly, there is a significant difference in the magnitude of the  $K_a$ . I studied the acetate guest with two different counterions with the expectation that  $NMe_4^+$  would be a less competitive counterion than  $NH_4^+$  due to its inability to hydrogen bond to acetate, resulting in higher  $K_a$  values. Instead, experiments using  $NH_4^+$  as the counterion resulted in higher  $K_a$  values. I do not have an explanation for this anomaly.

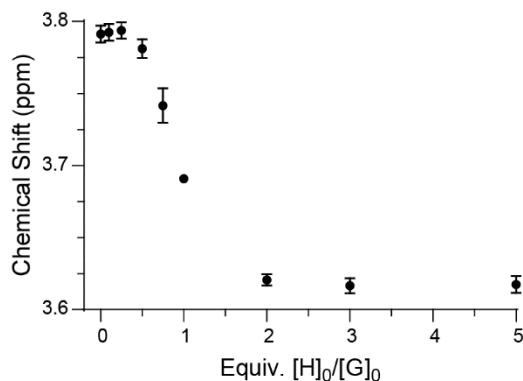
**Table 5.2.** Binding affinities obtained from titrations of guanidine compounds **5.18**, **5.19**, and **5.21** against ammonium acetate or tetramethylammonium acetate in DMSO-*d*<sub>6</sub> using the constant ion concentration method.

			
	5.19	5.18	5.21
$K_a$ (NH <sub>4</sub> CH <sub>3</sub> CO <sub>2</sub> )	< 6 x 10 <sup>-4</sup> M <sup>-1</sup>	3.44 ± 0.07 M <sup>-1</sup>	198 ± 37 M <sup>-1</sup>
$K_a$ (NMe <sub>4</sub> CH <sub>3</sub> CO <sub>2</sub> )	3.87 ± 0.14 M <sup>-1</sup>	5.77 ± 0.26 M <sup>-1</sup>	24.4 ± 2 M <sup>-1</sup>
p <i>K</i> <sub>a</sub>	>13	10.82	7.28

The improvements observed using the constant ion concentration method were promising, so I sought to apply the method to additional guanidino compounds. By this time, I had synthesized canavanine **3.2·HCl** (Scheme 3.2) and wanted to compare its binding affinity to oxoanions against that of arginine. Since oxyguanidine **5.21** demonstrates a higher  $K_a$  for acetate than an alkylguanidine (**5.19**), I expected canavanine **3.2·HCl** would have a stronger affinity for acetate than arginine **3.1·HCl** (the chloride salt of arginine **5.22**). I titrated **3.1·HCl** and **3.2·HCl** against ammonium acetate in DMSO-*d*<sub>6</sub> using the optimized conditions. Unfortunately, the curves observed from these titrations yet again did not conform to the classic 1:1, or even 1:2 binding stoichiometry models. Though I saw unusual curves with many proton signals from both amino acids, these results are best exemplified by the curve obtained from the  $\gamma$ -protons of canavanine **3.2·HCl** (Figure 5.9).

I made several more attempts to optimize the binding titrations to obtain meaningful results from arginine **3.1·HCl** and canavanine **3.2·HCl** but was unsuccessful. These attempts included keeping the concentration of guanidine compound constant and instead altering the ratios of ammonium acetate and NaCl to maintain the constant concentration of ions, which still yielded unfittable results. Thus, I sought

methods alternative to NMR titration to determine the anion-binding capability of arginine and canavanine, and settled on the octanol–water partitioning experiments described in Chapter 3. Work was not continued on biaryl-bisguanidines **5.6–5.9**.



**Figure 5.9.** Data resulting from a constant ion concentration titration of canavanine **3.2·HCl** against ammonium acetate in DMSO-*d*<sub>6</sub>. Chemical shifts of the  $\gamma$ -protons are shown.

## Conclusions

In this chapter, I have described the synthesis of biaryl-bisguanidines and attempts to determine their binding affinity to oxoanions through NMR titration. These structures contain either a binaphthyl or biphenyl core, with guanidino units that are either directly conjugated to the aryl backbone or further separated from the backbone using a methylene spacer. The four biaryl-bisguanidines offer a range of binding-pocket flexibilities. Attempts to determine the binding affinity of these structures, and other monoguanidines, to oxoanions via NMR titrations were fraught.

I determined that quantifying the binding affinity of the guanidino compounds studied in this chapter to oxoanions via NMR is untenable. The binding affinities to oxoanions are simply too weak. In Schmuck's original paper exploring his guanidiniocarbonyl pyrrole (GCP), he screened several compounds and found that



guanidinium chloride did not produce a measurable change in signal and acetylguanidine produced a minimal  $K_a$  via  $^1\text{H}$  NMR titrations with carboxylates in 40%  $\text{D}_2\text{O}/\text{DMSO-}d_6$ . Therefore, the weak binding of **5.6–5.10** and **5.18–5.22** requires exceptionally high concentrations of guest to measure. In neutral systems, this might be acceptable; however, with the host and guest being charged species, ionic interactions cannot be separated from the binding events.

These issues appear to be constrained to NMR spectroscopy. With observation techniques such as fluorescence spectroscopy, the host concentration can be in the sub-micromolar or nanomolar range, whereas with NMR spectroscopy, the host concentration is, at lowest, sub-millimolar.<sup>112</sup> Thus, changes in ionic strength of a solution at, say,  $[\text{H}]_0/[\text{G}]_0 = 50$  equivalents will be much more drastic at the host concentrations required for NMR versus for fluorescence. UV–Vis spectroscopy requires slightly higher concentrations than for fluorescence, though still drastically lower than NMR. I therefore recommend that any future researchers attempting to determine the binding affinity between two charged molecules use fluorescence or UV–Vis spectroscopy if possible and use the constant ion concentration method if they must use NMR spectroscopy.

While the data obtained from the experiments in this chapter are uninterpretable, they spark questions that could be studied further using more appropriate methods. One area of study that has yet to be explored in the field of guanidino supramolecular chemistry is the oxyguanidine. While the use of a conjugated carbonyl to lower the  $\text{p}K_a$  of a guanidine and increase its binding affinity has been widely studied,<sup>30,41,48</sup> the same cannot be said of oxyguanidine. The  $\text{p}K_a$  of acetylguanidine is 8.33,<sup>110</sup> while I have determined the  $\text{p}K_a$  of an oxyguanidine to be 7.28. Due to the simplicity of synthesis of an oxyguanidine, it is surprising that more researchers have not studied its utility in

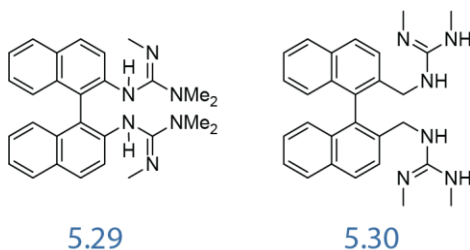
oxoanion binding. Indeed, while the results displayed in Table 5.2 do not align with expectations in regards to the acetate counterion, it is still evident that the binding between an oxyguanidine and acetate is stronger than with an alkyl- or arylguanidine.

It is surprising then that the results found in Chapter 3 indicate that the oxyguanidine-containing canavanine is less capable of partitioning into octanol in the presence of anions than arginine. Similarly, Bradford determined that an oxyguanidine-displaying polymer did not penetrate cells to a greater extent than the equivalent guanidino-polymer.<sup>124</sup> These results are likely caused by oxyguanidine imparting significantly more hydrophilicity than an alkylguanidine, as demonstrated by the topological polar surface area (TPSA) being greater for canavanine **3.2·HCl** than arginine **3.1·HCl** (Chapters 3 and 4). While oxyguanidine might not improve cell-uptake as initially hypothesized, its increased oxoanion-binding capability should be explored for its benefits to the basic science of supramolecular chemistry and for its potential use as an anion sensor.<sup>192</sup>

The biaryl-bisguanidines synthesized (**5.6–5.8**) may also be of use in various applications. While I did not return to studying the oxoanion binding capabilities of these structures, it is possible that they may produce more of a response in a less competitive solvent. I believe if they could bind to anions to the same extent as **GADAC** (see Chapter 1), I would likely have seen more of a response through NMR titration in water. However, I also performed the initial NMR titrations on the guanidino-azaindole **GAI** (Scheme 2.2) and saw a binding affinity in the same range as the biaryl-bisguanidines (Table 5.1). While I did not perform fluorescence titrations on **GAI**, I would expect its ability to bind oxoanions to be similar to that of **GAC** (see Chapter 1). Therefore, biaryl-bisguanidines **5.6–5.8** could display binding affinities high enough to be measured and compared in

solvents such as isopropanol and methanol. The original titrations in this chapter were performed using NMR partly because molecules such as diarginine **5.10** do not have strong chromophores and would not be amenable to UV–Vis. However, if one simply wanted to compare the binding affinities between the biaryls, UV–Vis would be an excellent choice of method.

In addition to oxoanion binding, biaryl-bisguanidines **5.6–5.8** may be capable of coordinating to metals, similar to the complexes studied by Himmel and coworkers.<sup>175–177</sup> As designed, these structures are unsubstituted and include labile hydrogens. Therefore, these structures could be converted into anions to coordinate with different metals or metal oxidation states. If the additional labile hydrogens of these guanidines interferes with metal coordination, the synthetic routes discussed in this chapter could be useful in the synthesis of bisguanidines with one labile hydrogen per guanidine, such as in compound **5.29**, or with two labile hydrogens per guanidine, such as in compound **5.30**.



Overall, the works described in this chapter are highly relevant to future researchers seeking to study guanidino compounds. Several future directions have been described, including further probing oxyguanidines' oxoanion binding and the metal coordination of biaryl-bisguanidines. Still, future researchers should be aware of the perils of determining the binding affinity of two charged species using NMR spectroscopy.

## Materials and Methods

### General

Materials. Commercial compounds were from Sigma-Aldrich (St. Louis, MO), Chem Impex (Wood Dale, IL), and Carbosynth (San Diego, CA) and were used without further purification.

Conditions. All procedures were performed in air at ambient temperature (~22 °C) and pressure (1.0 atm) unless specified otherwise.

Solvent removal. The phrase "concentrated under reduced pressure" refers to the removal of solvents and other volatile materials using a rotary evaporator while maintaining a water-bath temperature at 40 °C. Residual solvent was removed from samples at high vacuum (<0.1 Torr), which refers to the vacuum achieved by a mechanical belt-drive oil pump, or through lyophilization (freeze-drying) using a Labconco FreeZone.

Chromatography. Chemical reactions were monitored by thin-layer chromatography (TLC) using EMD 250 µm silica gel 60-F<sub>254</sub> plates and visualization with UV-illumination or KMnO<sub>4</sub>-staining, or by LC-MS on an ESI Agilent 6125B mass spectrometer. Flash chromatography was performed with a Biotage Isolera automated purification system using prepacked and re-packed SNAP KP silica gel columns or SNAP KP C18 columns.

Instrumentation. <sup>1</sup>H-NMR and <sup>13</sup>C-NMR spectra for compound characterization were obtained with Bruker or Varian spectrometers, and ESI HRMS data were obtained with an Agilent 6545 Q-ToF mass spectrometer at the Department of Chemistry Instrumentation Facility at the Massachusetts Institute of Technology. All <sup>1</sup>H-NMR spectra taken for titrations were performed on a Varian Inova-500 NMR, with the exception of compounds 15 and 19 which were performed on a Bruker Neo-500 when the Varian instrument was decommissioned.

## **pK<sub>a</sub> titrations**

Bisguanidine **5.6** (8.2 mg, 0.02 mmol) was dissolved in 15 mL D<sub>2</sub>O in a vial with a stirbar. While stirring, a pH probe was used to measure the pH<sub>(obs)</sub> of the solution. The solution was adjusted to the desired pH increments with NaOD and DCl. Aliquots were taken for <sup>1</sup>H-NMR analysis at each increment and changes in chemical shifts plotted and fitted using GraphPad Prism. The experiment was repeated in duplicate.

## **Original NMR titrations**

To prepare the buffer, 2,6-lutidine (1.16 mL, 10 mmol) was dissolved in D<sub>2</sub>O (18.84 mL, 0.05 wt% TSP) to produce a 500 mM lutidine buffer. Bisguanidines **5.6-5.10** (0.02 mmol) were dissolved in 4 mL buffer to produce a 5 mM stock solution; monoguanidine **5.18-5.22** and **GAI** (0.04 mmol) were dissolved in 4 mL buffer to produce a 10 mM stock solution. Solutions were prepared so that the total number of guanidine units was kept the same between molecules. Methyl sulfate sodium salt (268 mg, 2 mol) was dissolved in the guanidine stock solution (1 mL) to prepare a 2 M stock solution. Both the guanidine and sulfate stock solutions were then adjusted to pH<sub>(obs)</sub> = 7 using concentrated NaOD and DCl.

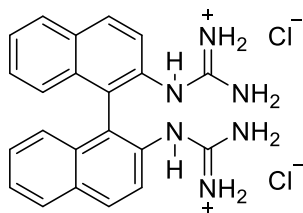
Three NMR tubes were prepared containing 575 μL of guanidine solution. Sulfate stock solution was added in 18 increments and an <sup>1</sup>H NMR was taken after each addition. The relevant peaks were picked in MestraNova and the plots stacked. <sup>1</sup>H NMR were referenced to the TSP peak using the autotune function in MestraNova to align the spectra. The data were exported by saving the stack plot and a peak list, opening the resulting file in a word processor, copying all the data, and pasting it into an MS Excel

file. The data were then processed and fitted in GraphPad Prism using a one-site binding model.

## Constant ion concentration NMR titrations

An aliquot of acetonitrile (7.8  $\mu\text{L}$ , 0.15 mmol) was added to DMSO- $d_6$  (3 mL) to produce a 50 mM solution. Monoguanidine **5.18**, **5.19**, **5.21**, **3.1·HCl**, or **3.2·HCl** (0.24 mmol) was dissolved in the acetonitrile-spiked DMSO- $d_6$  (1.2 mL) to make a 0.2 M stock solution of guanidine. Ammonium or tetrabutylammonium acetate (3 mmol) was dissolved in the acetonitrile-spiked DMSO- $d_6$  (1.5 mL) to make a 0.2 M stock solution of acetate. Proportional amounts of both stock solutions were added to 12 Eppendorf tubes to total 200  $\mu\text{L}$  each, ranging from ratios of zero to fifty equivalents [acetate] to [guanidine]. The solutions were then transferred to 3 mm NMR tubes using a pasture pipette.  $^1\text{H}$  NMR spectra were taken of all tubes, and the data was processed as described previously.

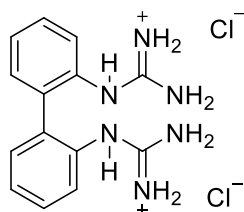
## Synthesis



### **1,1'-([1,1'-binaphthalene]-2,2'-diyl)diguandine dihydrogen chloride (5.6):**

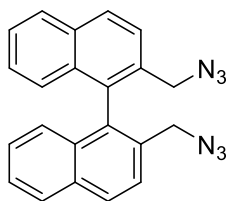
1,1'-binaphthyl-2,2'-diamine (50 mg, 0.18 mmol) and cyanamide (29.4 mg, 0.7 mmol) were dissolved in 1,4-dioxane (1.8 ml). A 4M HCl in dioxane solution (45  $\mu\text{L}$ , 0.18 mmol) was added, and the solution was heated at 80  $^{\circ}\text{C}$  overnight. The orange solution was then

concentrated under reduced pressure. The product was dissolved in a minimal amount of 1M HCl, and purified by reversed phase chromatography with a MeCN/H<sub>2</sub>O gradient, eluting between 5-10% MeCN. The product was lyophilized to yield **2** as a white powder (23 mg, 28% yield). **<sup>1</sup>H NMR** (400 MHz, D<sub>2</sub>O,  $\delta$ ): 8.29 (d,  $J$  = 8.8 Hz, 2H), 8.16 (d,  $J$  = 8.3 Hz, 2H), 7.75 (d,  $J$  = 8.8 Hz, 2H), 7.66 (dd,  $J$  = 8.1, 6.9 Hz, 2H), 7.39 (dd,  $J$  = 8.5, 6.9 Hz, 2H), 7.19 (d,  $J$  = 8.5 Hz, 2H). **<sup>13</sup>C NMR** (400 MHz, D<sub>2</sub>O,  $\delta$ ): 156.11, 132.87, 132.69, 131.16, 130.85, 129.69, 128.66, 127.83, 127.31, 125.62, 124.59. **HRMS**  $m/z$  calcd for C<sub>22</sub>H<sub>20</sub>N<sub>6</sub> [M + H]<sup>+</sup> 369.1827; found 369.1825.



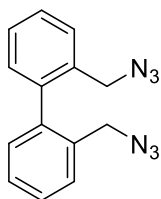
**1,1'-([1,1'-biphenyl]-2,2'-diyl)diguandine dihydrogen chloride (5.8):**

Bisguanidine **5.8** was prepared as described for bisguanidine **5.6**, producing a white powder (11% yield). **<sup>1</sup>H NMR** (500 MHz, D<sub>2</sub>O,  $\delta$ ): 7.59 (td,  $J$  = 7.5, 1.8 Hz, 2H), 7.56 (td,  $J$  = 7.5, 1.8 Hz, 2H), 7.48 (dd,  $J$  = 7.3, 1.8 Hz, 2H), 7.45 (dd,  $J$  = 7.3, 2.0 Hz, 2H). **HRMS**  $m/z$  calcd for C<sub>14</sub>H<sub>16</sub>N<sub>6</sub> [M + H]<sup>+</sup> 269.1514; found 269.1511.

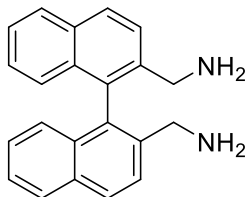


**2,2'-bis(azidomethyl)-1,1'-binaphthalene (5.11):** Sodium azide (174 mg, 2.68 mmol) and 2,2'-bis(bromomethyl)-1,1'-binaphthalene (472 mg, 1.07 mmol) were dissolved in DMSO (6 mL) and stirred at 80 °C overnight. The solution was diluted with water (20 mL) and extracted with diethyl ether (3 x 10 mL). The combined organics were

then washed with water (10 ml) and brine (10 mL), dried with Na<sub>2</sub>SO<sub>4</sub>, filtered, and concentrated under reduced pressure. The crude product was purified by silica gel chromatography with an EtOAc/Hex gradient, eluting at between 3–5% EtOAc to produce a clear oil (230 mg, 60% yield). **<sup>1</sup>H NMR** (500 MHz, CDCl<sub>3</sub>, δ): 8.06 (d, *J* = 8.6 Hz, 2H), 7.96 (d, *J* = 8.4, 2H), 7.72 (d, *J* = 8.7 Hz, 2H), 7.50 (t, *J* = 7.5 Hz, 2H), 7.29 (t, *J* = 7.6 Hz, 2H), 7.06 (d, *J* = 8.6 Hz, 2H), 4.08 (d, *J* = 2.5 Hz, 4H). **<sup>13</sup>C NMR** (500 MHz, CDCl<sub>3</sub>, δ): 134.72, 133.98, 133.41, 133.38, 129.91, 128.97, 127.79, 127.32, 126.87, 126.81, 53.67. **HRMS** *m/z* calcd for C<sub>22</sub>H<sub>16</sub>N<sub>6</sub> [M + Na]<sup>+</sup> 387.1334; found 387.1326.



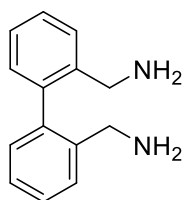
**2,2'-bis(azidomethyl)-1,1'-biphenyl (5.13)**: Diazide **5.13** was prepared as described for diazide **5.11**, producing a clear oil (75% yield). **<sup>1</sup>H NMR** (500 MHz, CDCl<sub>3</sub>, δ): 7.48 (dd, *J* = 7.4, 1.8 Hz, 2H), 7.44 (td, *J* = 7.3, 1.5 Hz, 2H), 7.41 (td, *J* = 7.3, 1.9 Hz, 2H), 7.22 (dd, *J* = 7.5, 1.5 Hz, 2H), 4.13 (d, *J* = 13.6 Hz, 2H), 4.07 (d, *J* = 13.5 Hz, 2H). **<sup>13</sup>C NMR** (500 MHz, CDCl<sub>3</sub>, δ): 139.68, 133.74, 130.31, 129.43, 128.54, 128.37, 52.69. **HRMS** *m/z* calcd for C<sub>14</sub>H<sub>12</sub>N<sub>6</sub> [M + H]<sup>+</sup> 287.1021; found 287.1019.



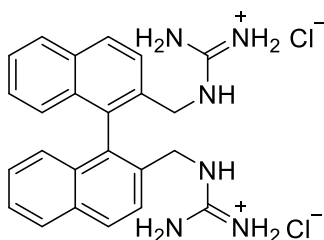
**[1,1'-binaphthalene]-2,2'-diyldimethanamine (5.12)**: Pd/C (35 mg, 15% wt.) was placed in a dry flask, to which diazide **5.11** (230 mg, 0.63 mmol) dissolved in degassed 1:1 EtOAc/EtOH (6 mL) was added. The solution was sparged with one balloon H<sub>2</sub> and left



to stir under H<sub>2</sub> pressure for 3 h. The suspension was then filtered through celite and concentrated under reduced pressure. The crude product was purified by silica gel chromatography with a 7N NH<sub>3</sub> in MeOH/DCM gradient, eluting at between 6-10% 7N NH<sub>3</sub> in MeOH to produce a light-yellow oil (197 mg, 78% yield). **<sup>1</sup>H NMR** (500 MHz, CDCl<sub>3</sub>, δ): 8.00 (d, *J* = 8.4 Hz, 2H), 7.92 (dd, *J* = 8.0, 1.2 Hz, 2H), 7.73 (d, *J* = 8.4 Hz, 2H), 7.43 (ddd, *J* = 8.0, 6.7, 1.1 Hz, 2H), 7.23 (ddd, *J* = 8.2, 6.8, 1.3 Hz, 2H), 7.06 (dd, *J* = 8.4, 1.1 Hz, 2H), 3.56 (d, *J* = 13.6 Hz, 2H), 3.48 (d, *J* = 13.7 Hz, 2H), 1.8 (s, br, 4H). **<sup>13</sup>C NMR** (500 MHz, CDCl<sub>3</sub>, δ): 139.14, 133.69, 133.18, 132.88, 128.75, 128.24, 126.61, 126.55, 126.11, 125.80, 44.51. **HRMS** *m/z* calcd for C<sub>22</sub>H<sub>20</sub>N<sub>2</sub> [M + H]<sup>+</sup> 313.1704; found 313.1702.



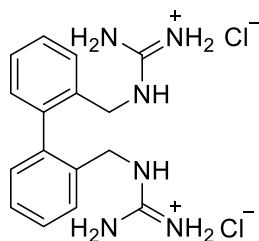
**[1,1'-biphenyl]-2,2'-diyl dimethanamine (5.14)**: Diamine **5.14** was prepared as described for diamine **5.12**, producing a light-yellow oil (65% yield). **<sup>1</sup>H NMR** (500 MHz, CDCl<sub>3</sub>, δ): 7.47 (d, *J* = 7.6 Hz, 2H), 7.38 (t, *J* = 7.5 Hz, 2H), 7.29 (t, *J* = 7.5 Hz, 2H), 7.14 (d, *J* = 7.5 Hz, 2H), 3.58 (s, 4H), 1.43 (s, br, 4H). **<sup>13</sup>C NMR** (500 MHz, CDCl<sub>3</sub>, δ): 141.25, 140.01, 130.10, 128.36, 128.31, 127.03, 44.51. **HRMS** *m/z* calcd for C<sub>14</sub>H<sub>16</sub>N<sub>2</sub> [M + H]<sup>+</sup> 213.1391; found 213.1387.



**1,1'-([1,1'-binaphthalene]-2,2'-diylbis(methylene))diguanidine**

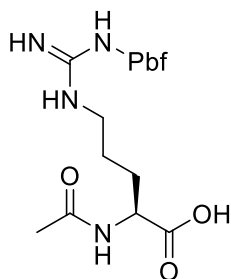
**dihydrochloride (5.7)**: Diamine **5.12** (65 mg, 0.2 mmol) and 1*H*-pyrazole-1-

carboxamidine (117 mg, 0.8 mmol) were added to a dry flask, followed by dry DMF (2 mL) and DIEA (174  $\mu$ L, 1 mmol). The solution was stirred at rt overnight, then concentrated under a stream of air. The product was dissolved in a minimal amount of 1M HCl, and purified by reversed phase chromatography with a MeCN/H<sub>2</sub>O gradient, eluting between 15-20% MeCN. The product was lyophilized to yield **5.7** as a white powder (61 mg, 65% yield). **<sup>1</sup>H NMR** (600 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 8.19 (d, *J* = 8.6 Hz, 1H), 8.07 (d, *J* = 8.2 Hz, 1H), 7.83 (t, *J* = 5.4 Hz, 1H (*NH*)), 7.70 (d, *J* = 8.6 Hz, 1H), 7.54 (dd, *J* = 8.2, 6.8 Hz, 1H), 7.32 (dd, *J* = 8.2, 6.8, 1H), 7.10 (s, br, 4H (*NH*)), 6.89 (d, *J* = 8.5 Hz, 1H), 4.03 (dd, *J* = 14.9 (*gem. CH*), 4.9 (*NH*) Hz, 1H), 3.96 (dd, *J* = 14.9 (*gem. CH*), 5.6 (*NH*) Hz, 1H). **<sup>13</sup>C NMR** (400 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 157.28, 133.51, 133.34, 133.15, 132.51, 129.37, 128.69, 127.51, 126.84, 126.08, 125.90, 43.50. **HRMS** *m/z* calcd for C<sub>24</sub>H<sub>24</sub>N<sub>6</sub> [M + H]<sup>+</sup> 397.2141; found 397.2139.

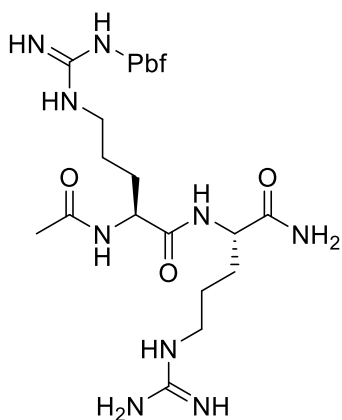


**1,1'-([1,1'-binaphthalene]-2,2'-diylbis(methylene))diguanidinium**

**dihydrochloride (5.9):** Diamine **5.9** was prepared as described for diamine **5.14**, producing a white powder (21% yield). **<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 7.88 (t, *J* = 5.5 Hz, 2H (*NH*)), 7.47 (t, *J* = 7.8 Hz, 2H), 7.44 (d, *J* = 9.3, 2H), 7.41 (t, *J* = 7.3, 2H), 7.23 (d, *J* = 7.2 Hz, 2H), 7.35 (s, br, 4H (*NH*)) 4.12 (dd, *J* = 15.0 (*gem. CH*), 5.5 Hz (*NH*), 2H), 4.06 (dd, *J* = 15.0 (*gem. CH*), 5.4 Hz (*NH*), 2H). **<sup>13</sup>C NMR** (400 MHz, DMSO-*d*<sub>6</sub>  $\delta$ ): 156.75, 138.81, 134.19, 129.72, 128.10, 128.01, 127.54, 42.86. **HRMS** *m/z* calcd for C<sub>16</sub>H<sub>20</sub>N<sub>6</sub> [M + H]<sup>+</sup> 297.1828; found 287.1823.

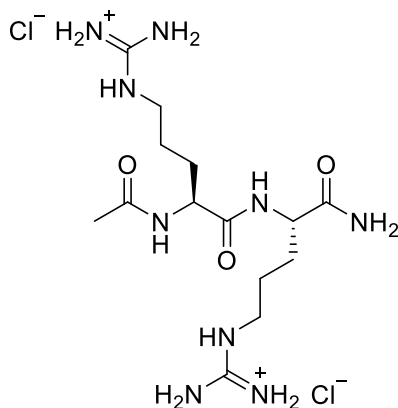


**N-Ac-Arg(Pbf)-H (5.15):** H-Arg(Pbf)-OH (250 mg, 0.59 mmol) was dissolved in 1:1 MeOH/acetic anhydride (2 mL) and stirred for 2 h at rt until solution became clear. The solution was concentrated under reduced pressure and purified by silica gel chromatography with a MeOH/DCM gradient to produce a white foam (268.2 mg, 97% yield).  $^1\text{H NMR}$  (500 MHz,  $\text{CDCl}_3$ ,  $\delta$ ): 4.33 (t,  $J = 6.7$  Hz, 1H), 3.15 (m, 2H), 2.90 (s, 2H), 2.47 (s, 3H), 2.43 (s, 3H), 2.03 (s, 3H), 2.00 (s, 3H), 1.86 (m, 1H), 1.74 (m, 1H), 1.59 (m, 2H), 1.42 (s, 6H). **HRMS**  $m/z$  calcd for  $\text{C}_{21}\text{H}_{32}\text{N}_4\text{O}_6\text{S}$   $[\text{M} + \text{H}]^+$  469.2121; found 469.2122.

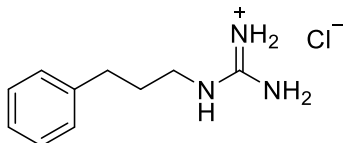


**N-Ac-Arg(Pbf)-Arg-NH<sub>2</sub> (5.16):** N-Ac-Arg(Pbf)-H **5.15** (134 mg, 0.28 mmol), arginine amide dihydrochloride (69 mg, 0.28 mmol), and HATU (106 mg, 0.28 mmol) were added to a flask, followed by DMF (3 mL) and DIEA (148  $\mu\text{L}$ , 0.85 mmol). The bright yellow solution was stirred at rt for 48 h, then concentrated under a stream of air. The product was purified by reversed phase chromatography with a MeCN/ $\text{H}_2\text{O}$  gradient, to produce **5.16** (94 mg, 54% yield).  $^1\text{H NMR}$  (500 MHz,  $\text{DMSO}-d_6$ ,  $\delta$ ): 8.16 (dd,  $J = 21.7$ ,

7.7 Hz, 1H), 8.05 (d,  $J = 7.7$  Hz, 1H), 7.96 (d,  $J = 8.1$  Hz, 1H), 7.42 (s, 1H), 7.33 (s, 1H), 7.14 (s, 1H), 7.09 (s, 1H), 6.68 (s, 3H), 6.37 (s, 2H), CH: 4.20 (q, 7.0 Hz, 1H), 4.19 (q, 7.0 Hz, 1H), 3.08 (t, br, 2H), 3.02 (q, br, 2H), 2.97 (s, 2H), 2.48 (s, 3H), 2.42 (s, 3H), 2.01 (s, 3H), 1.84 (s, 3H), 1.73–1.57 (m, 3H), 1.56–1.38 (m, 5H), 1.41 (s, 6H). **HRMS**  $m/z$  calcd for  $C_{27}H_{45}N_9O_6S$   $[M + H]^+$  624.3292; found 624.3296.

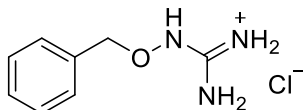


**N-Ac-Arg-Arg-NH<sub>2</sub> dihydrochloride (5.10):** N-Ac-Arg(Pbf)-Arg-NH<sub>2</sub> **5.17** (45 mg, 0.07 mmol) was dissolved in 3:1 DCM/TFA and stirred overnight. The solution was condensed under reduced pressure, dissolved in a minimal amount of 1M HCl, and purified by reversed phase chromatography with a MeCN/H<sub>2</sub>O gradient, eluting at 0% MeCN. The product was lyophilized to yield **5.10** as a white powder (32 mg, 49% yield). **<sup>1</sup>H NMR** (500 MHz, D<sub>2</sub>O,  $\delta$ ): 4.33 (dd,  $J = 9.0, 5.5$  Hz, 1H), 4.27 (dd,  $J = 8.2, 6.2$  Hz, 1H), 3.21 (t,  $J = 7.0$  Hz, 4H), 3.20 (q,  $J = 7.0$  Hz, 4H), 2.02 (s, 3H), 1.93 – 1.55 (m, 8H). **HRMS**  $m/z$  calcd for  $C_{14}H_{29}N_9O_3$   $[M + H]^+$  372.2472; found 372.2470.



**1-(3-phenylpropyl)guanidine hydrochloride (5.20):** 1-(3-phenylpropyl)amine (105  $\mu$ L, 0.74 mmol) and 1H-pyrazole-1-carboximidine (217 mg, 1.48 mmol) were added

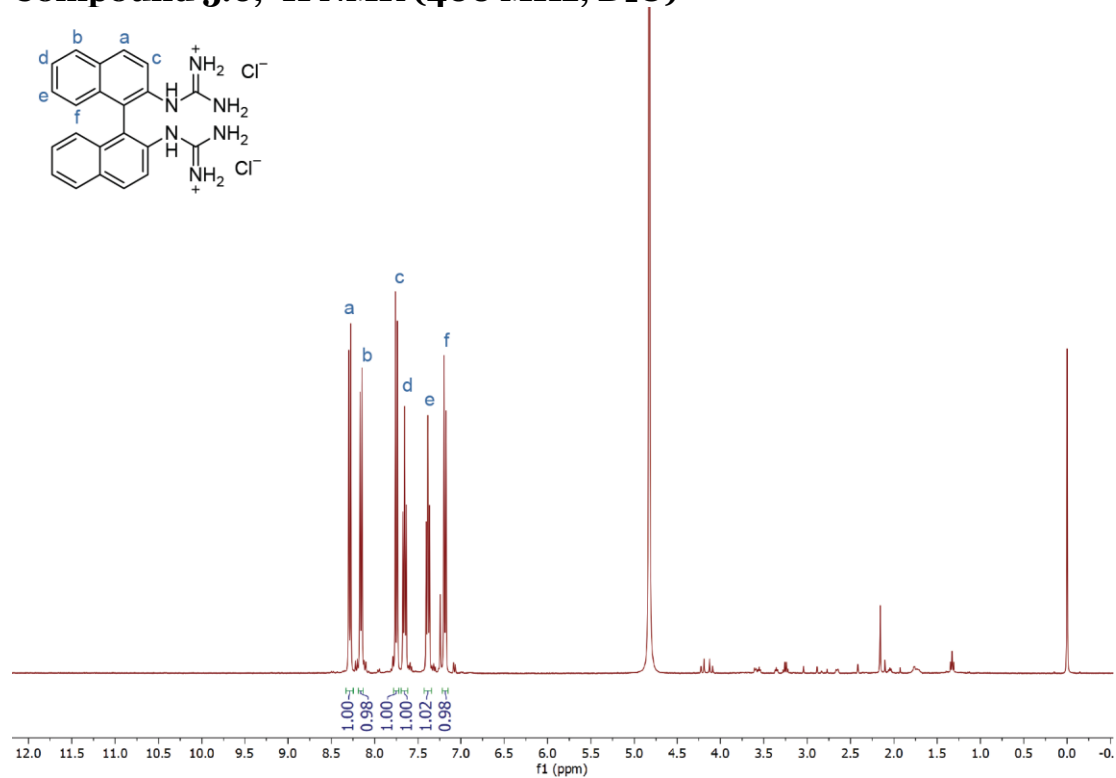
to a dry flask, followed by dry DMF (7.5 mL) and DIEA (193  $\mu$ L, 1.11 mmol). The solution was stirred at rt overnight, and concentrated under a stream of air. The product was dissolved in a minimal amount of 1M HCl, and purified by reversed phase chromatography with a MeCN/H<sub>2</sub>O gradient, eluting at 0% MeCN. The product was lyophilized to yield **5.20** as a white powder (127 mg, 80% yield). **<sup>1</sup>H NMR** (600 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 7.82 (t, *J* = 5.6 Hz, 1H (*NH*)), 7.50–6.77 (br, 4H (*NH*)), 7.29 (t, *J* = 7.6 Hz, 2H), 7.22 (d, *J* = 6.7 Hz, 2H), 7.19 (t, *J* = 5.6 Hz, 1H), 3.11 (q, *J* = 7.0 Hz, 2H), 2.62 (dd, *J* = 8.9, 6.7 Hz, 2H), 1.77 (dt, *J* = 14.8, 7.1 Hz, 2H). **<sup>13</sup>C NMR** (600 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 156.91, 141.09, 128.30, 128.21, 125.84, 40.19, 31.95, 30.17. **HRMS** *m/z* calcd for C<sub>10</sub>H<sub>15</sub>N<sub>3</sub> [M + H]<sup>+</sup> 178.1344; found 178.1339.



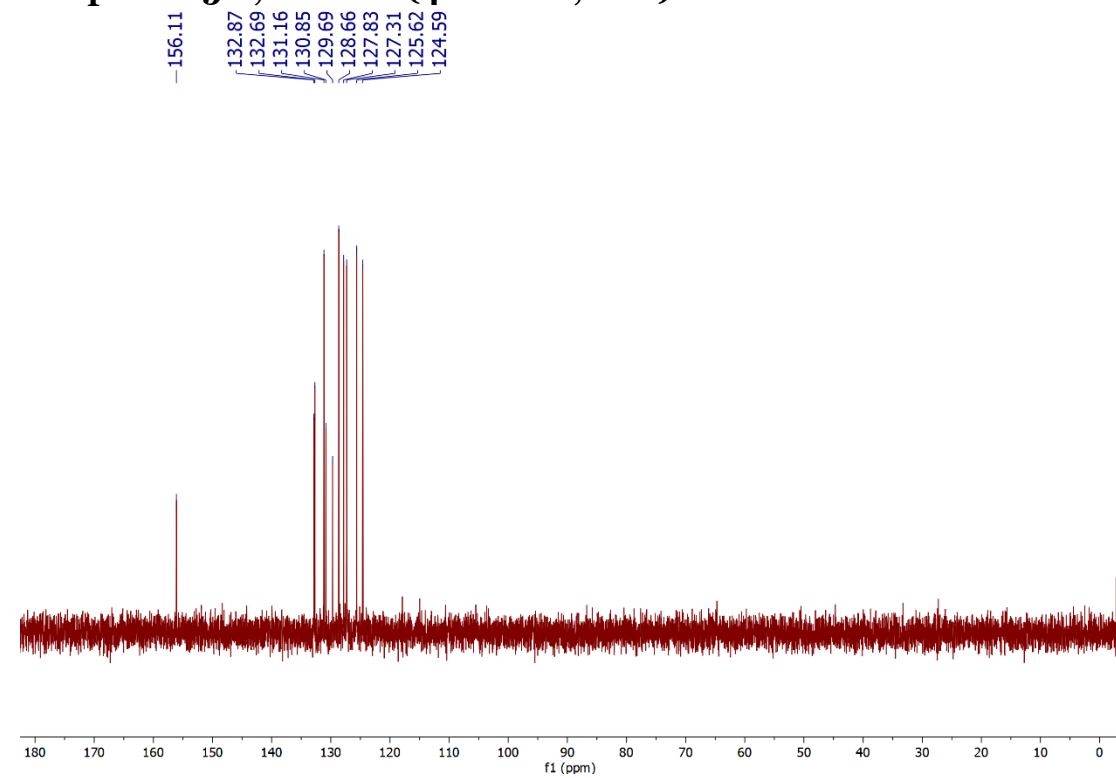
**O-Benzylhydroxylamine (5.21):** O-benzylhydroxylamine (100 mg, 0.62 mmol) and cyanamide (52 mg, 1.24 mmol) were dissolved in dioxane in a pressure sealed vial. Hydrochloric acid (4M in dioxane, 155  $\mu$ L, 0.62 mmol) was added, and the reaction was stirred at 80 °C overnight. The solution was condensed under reduced pressure, dissolved in a minimal amount of 1M HCl, and purified by reversed phase chromatography with a MeCN/H<sub>2</sub>O gradient, eluting at 0% MeCN. The product was lyophilized to yield **5.21** as a white powder (38 mg, 30% yield). **<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 11.02 (s, 1H), 7.69 (s, 3H), 7.50 – 7.35 (m, 5H), 4.83 (s, 2H). **<sup>13</sup>C NMR** (400 MHz, DMSO-*d*<sub>6</sub>,  $\delta$ ): 158.57, 135.17, 129.26, 128.62, 128.37, 78.01. **HRMS** *m/z* calcd for C<sub>8</sub>H<sub>11</sub>N<sub>3</sub>O [M + H]<sup>+</sup> 166.0980; found 166.0975.

## NMR Spectra

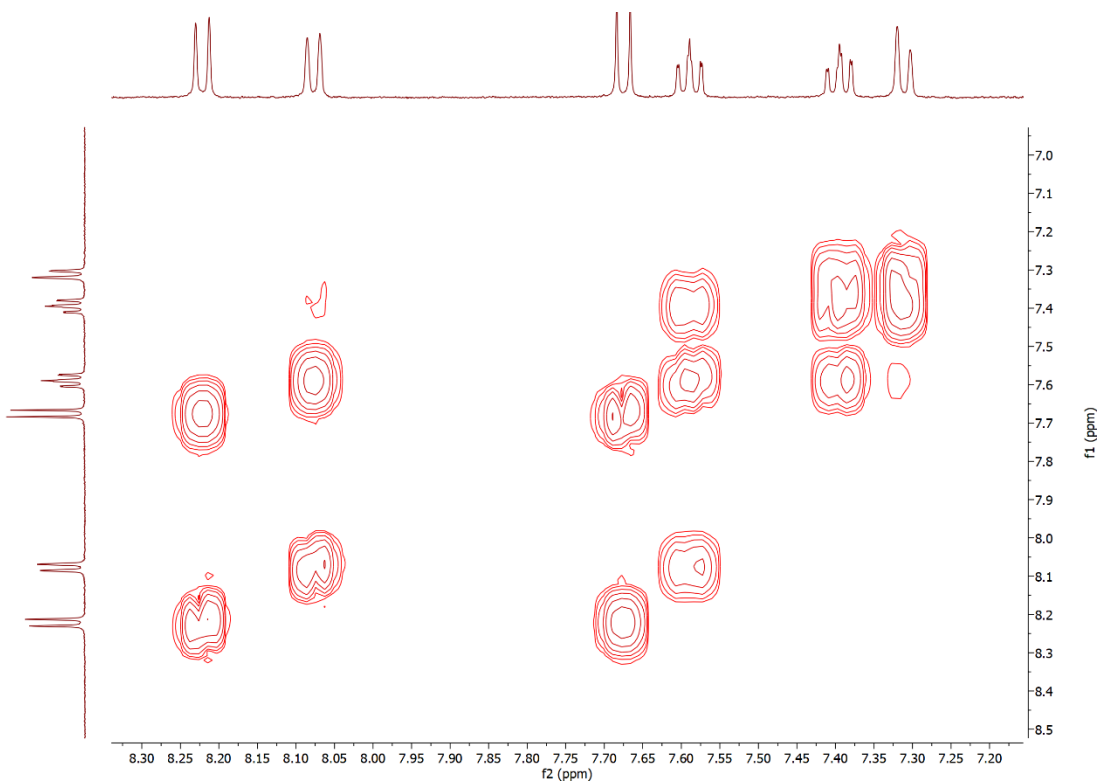
### Compound 5.6, $^1\text{H}$ NMR (400 MHz, $\text{D}_2\text{O}$ )



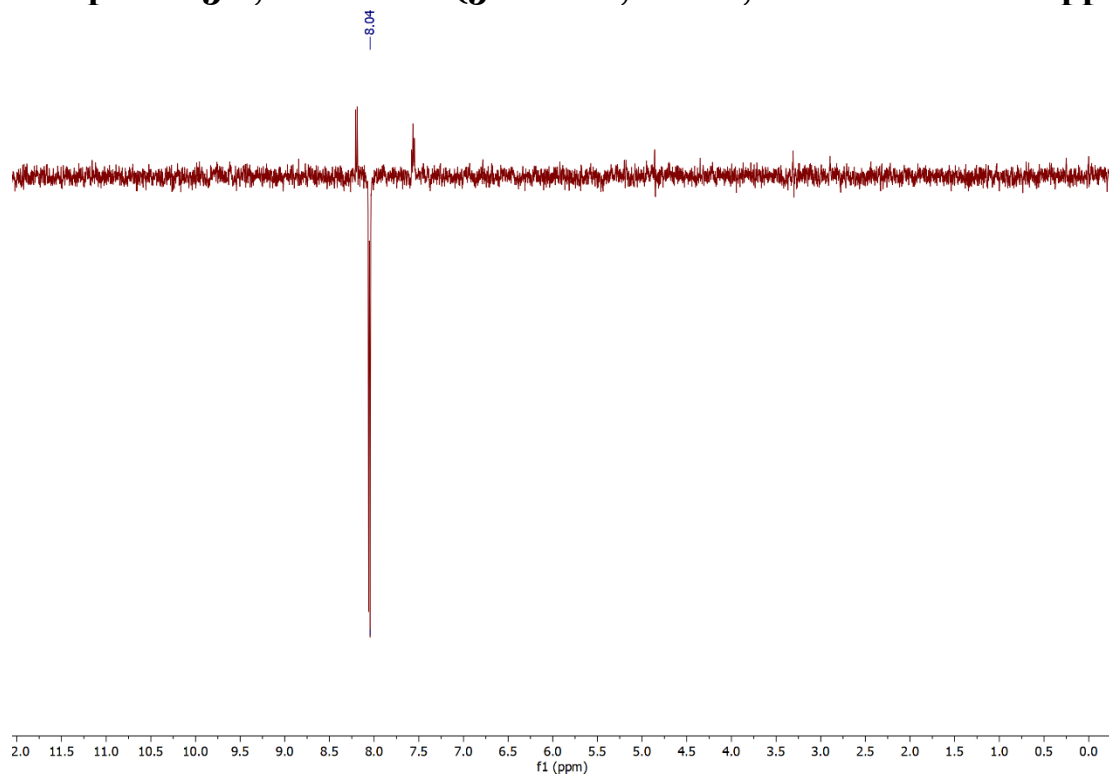
### Compound 5.6, $^{13}\text{C}$ NMR (400 MHz, $\text{D}_2\text{O}$ )



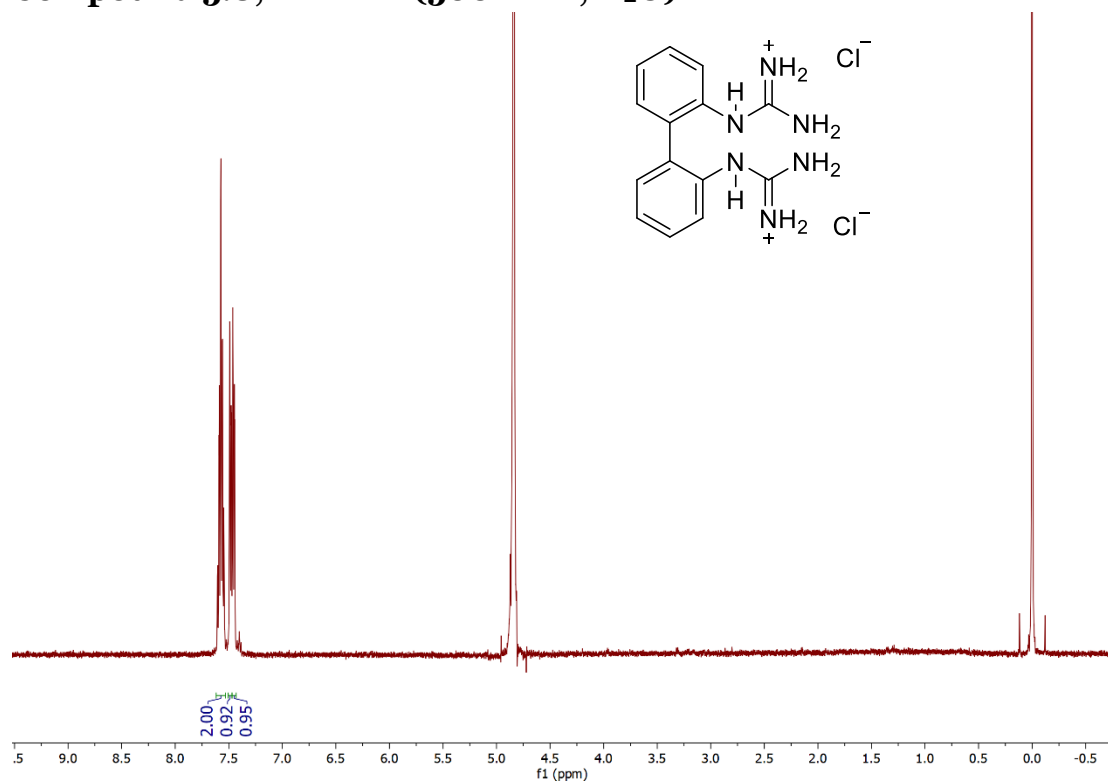
### Compound 5.6, COSY (500 MHz, MeOD)



**Compound 5.6, 1D-NOESY (500 MHz, MeOD, Irradiated at 8.20 ppm)**

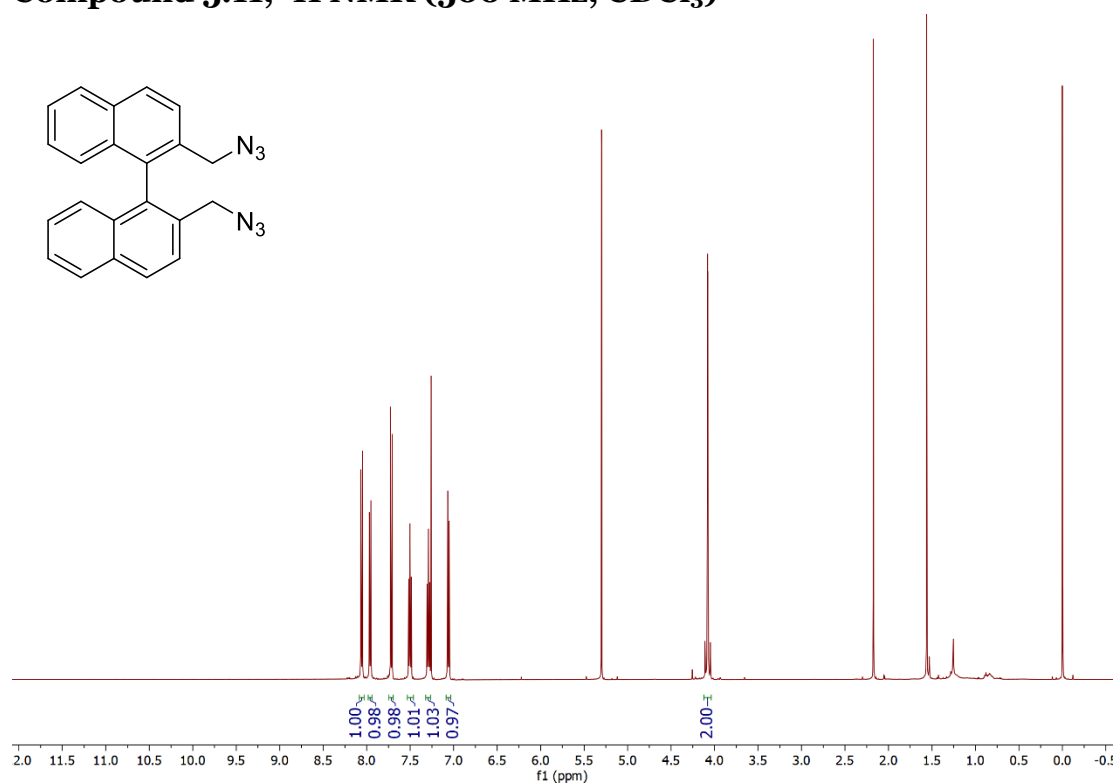


Compound 5.8, <sup>1</sup>H NMR (500 MHz, D<sub>2</sub>O)

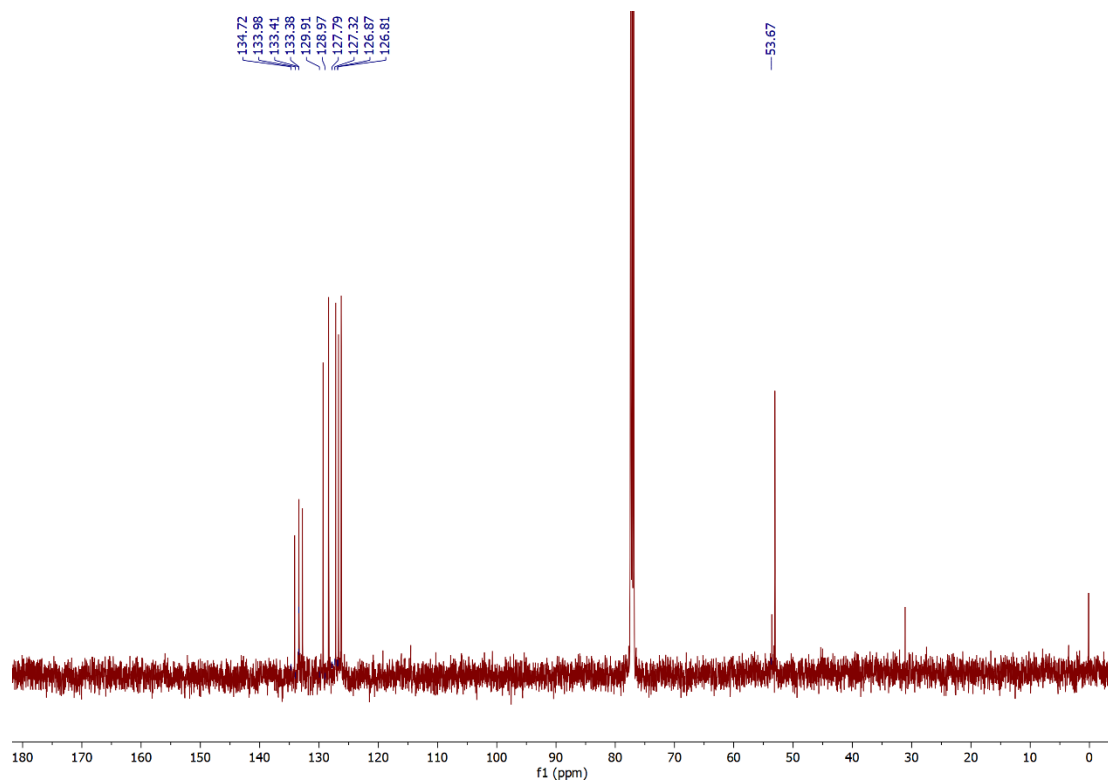




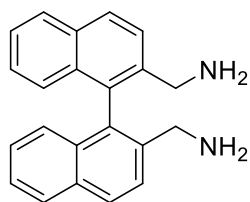
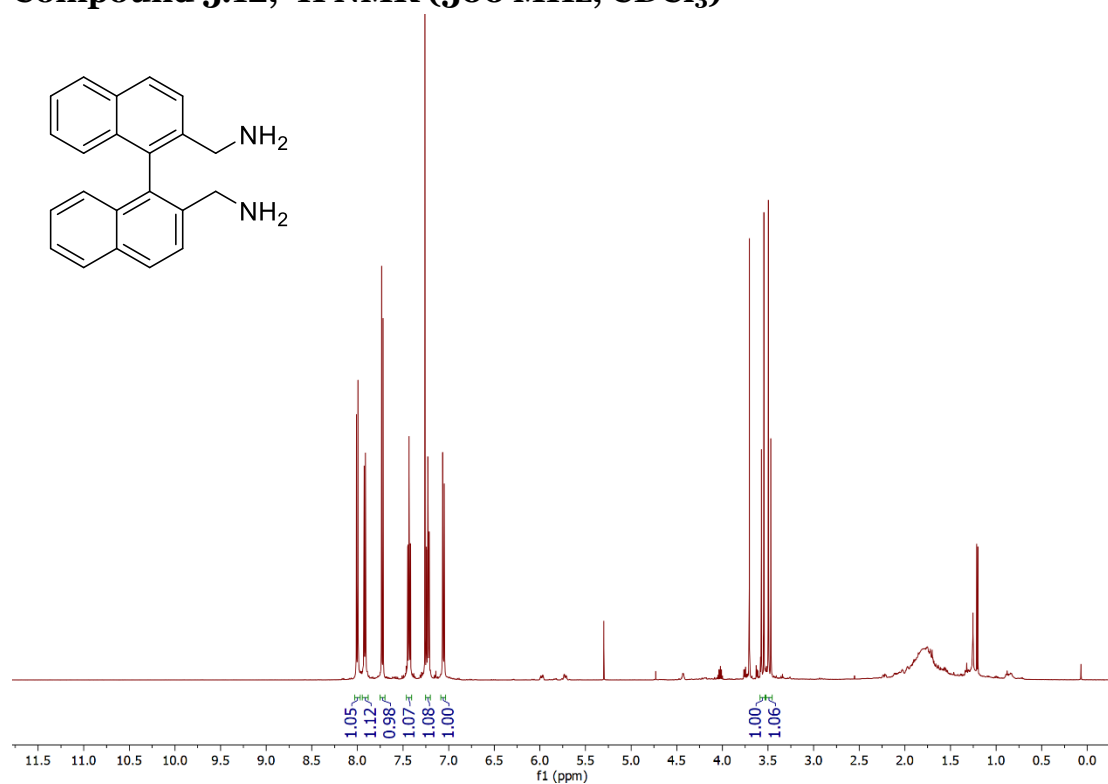
### Compound 5.11, $^1\text{H}$ NMR (500 MHz, $\text{CDCl}_3$ )



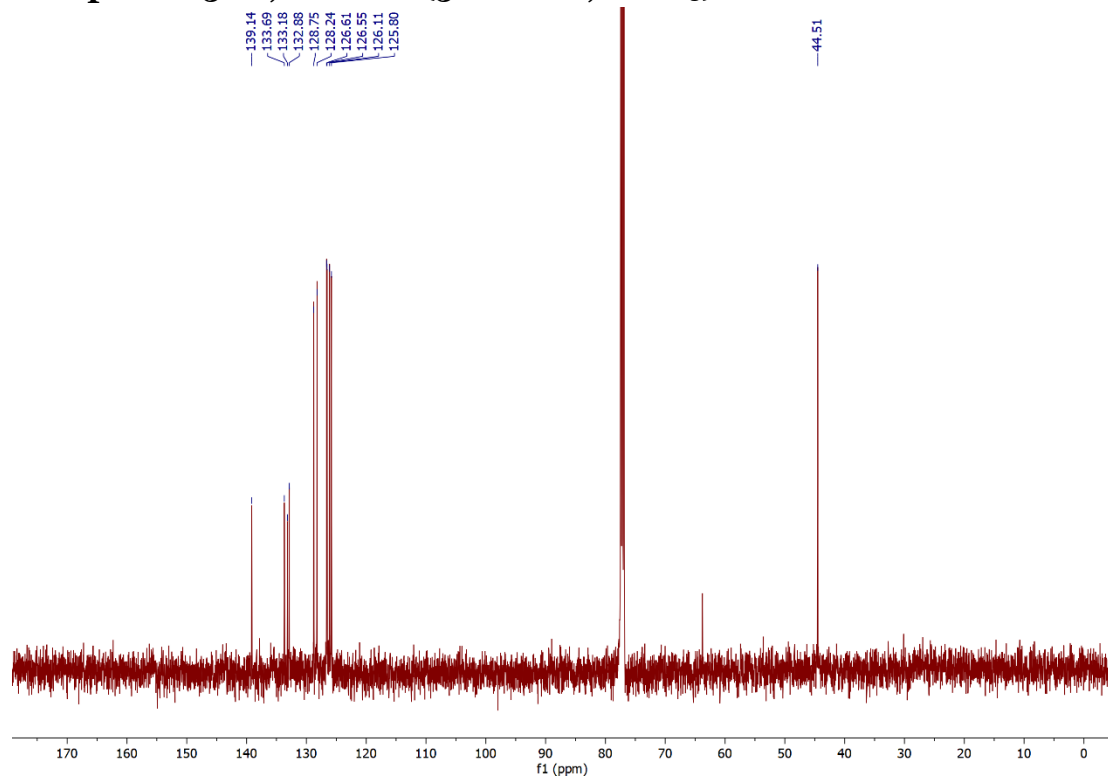
### Compound 5.11, $^{13}\text{C}$ NMR (500 MHz, $\text{CDCl}_3$ )



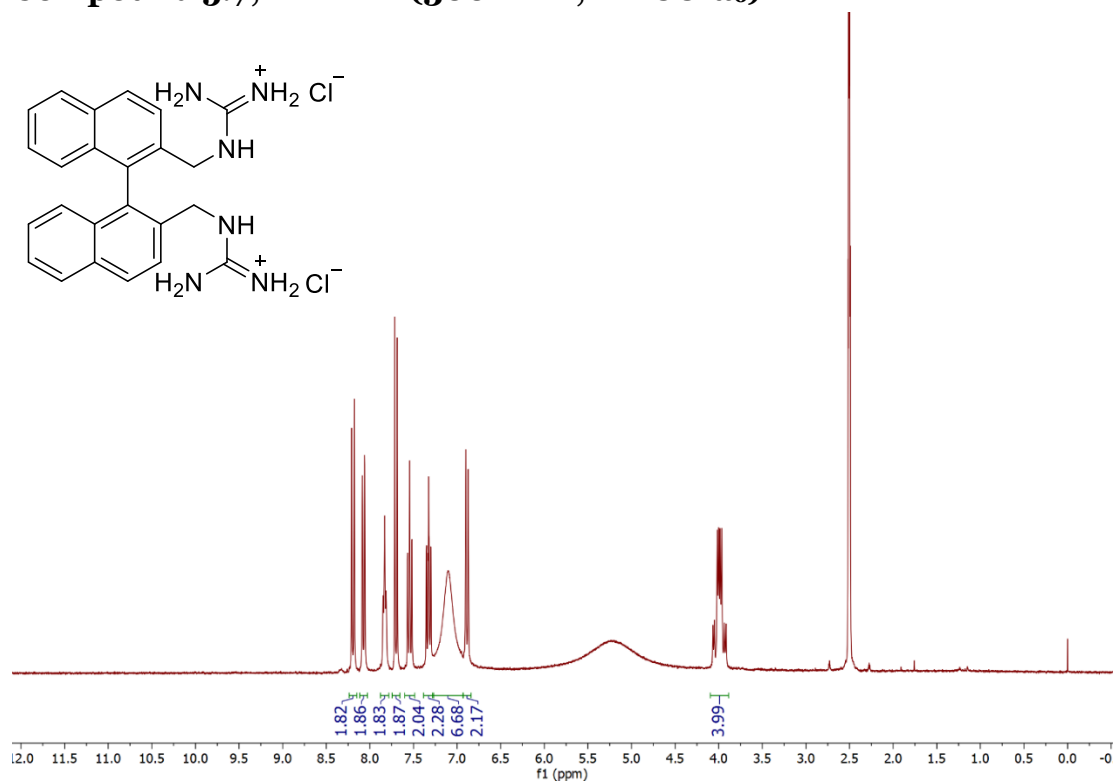
### Compound 5.12, <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)



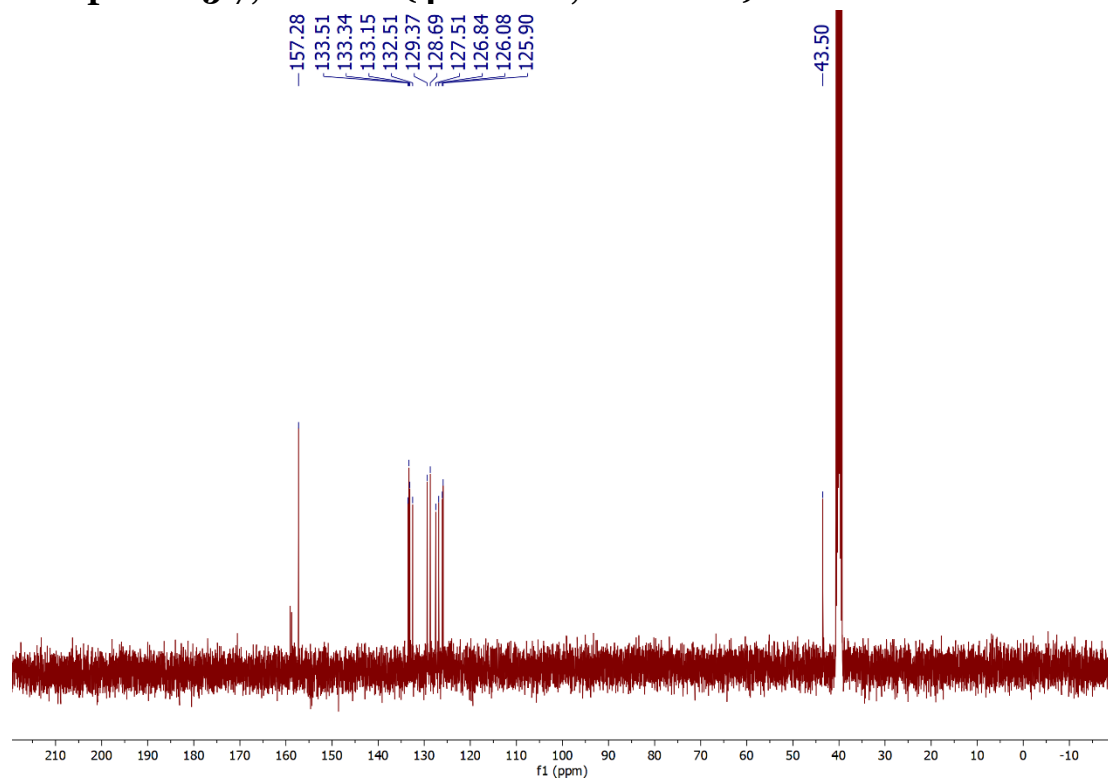
### Compound 5.12, <sup>13</sup>C NMR (500 MHz, CDCl<sub>3</sub>)



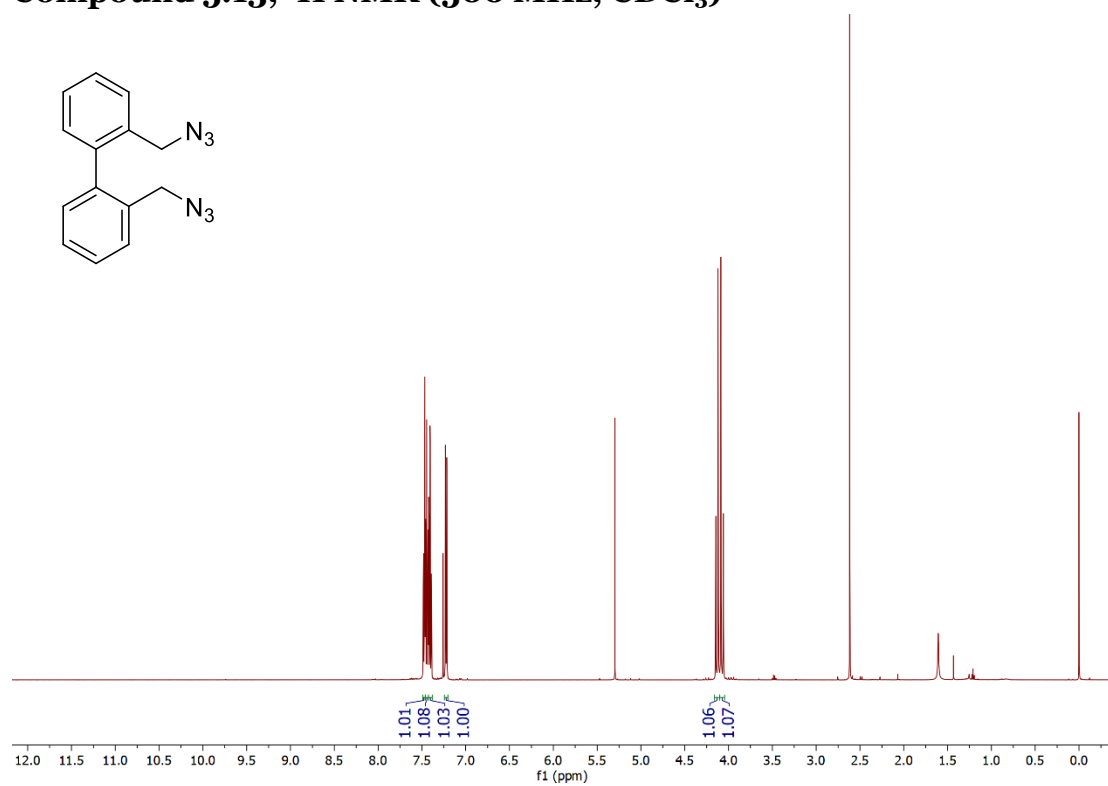
### Compound 5.7, $^1\text{H}$ NMR (300 MHz, $\text{DMSO-}d_6$ )



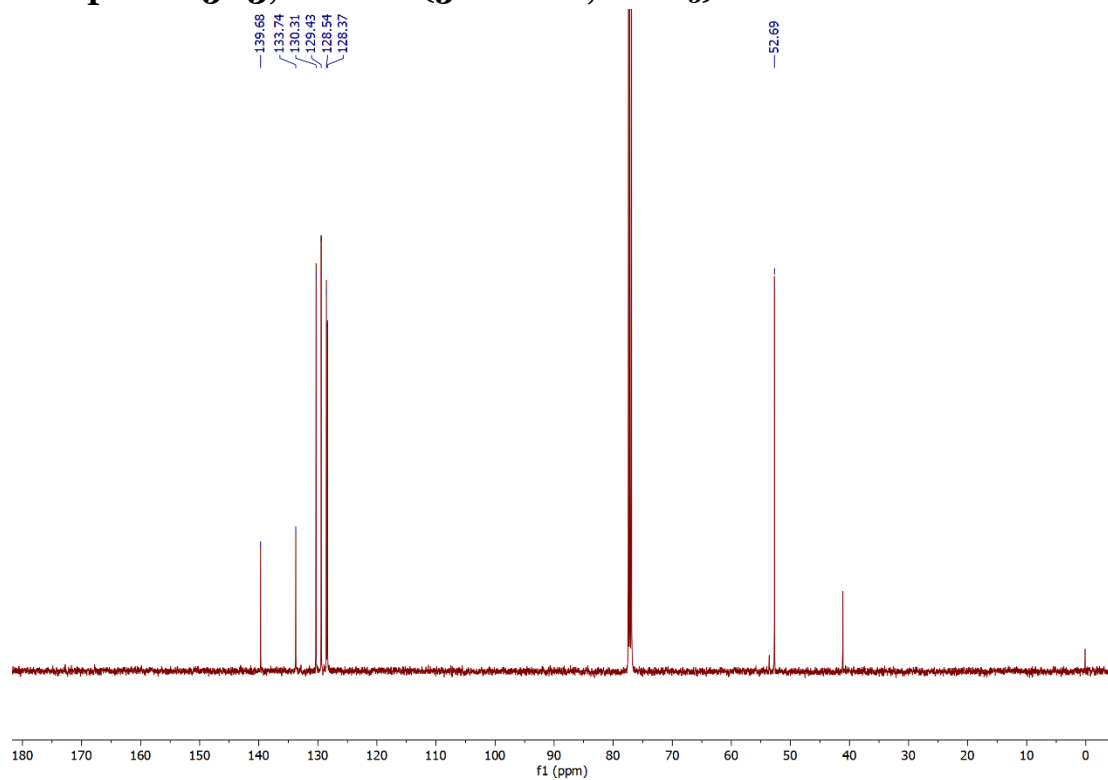
### Compound 5.7, $^{13}\text{C}$ NMR (400 MHz, $\text{DMSO-}d_6$ )



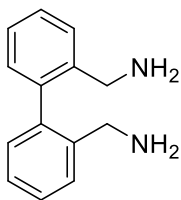
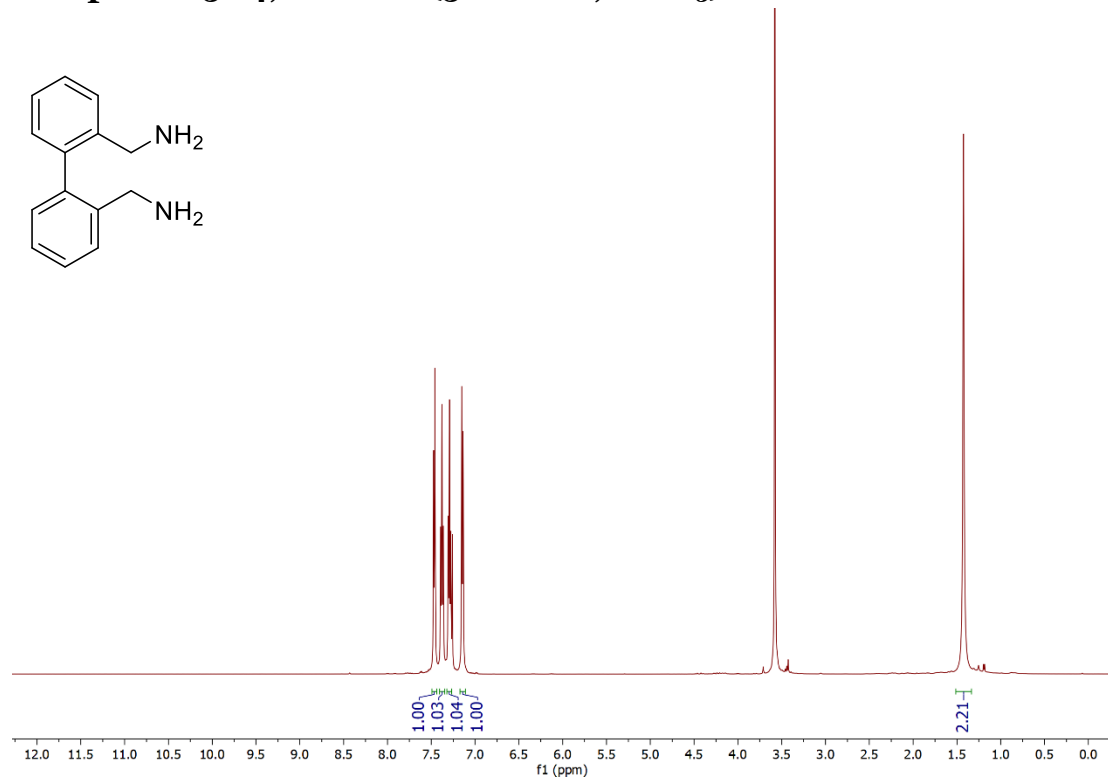
### Compound 5.13, $^1\text{H}$ NMR (500 MHz, $\text{CDCl}_3$ )



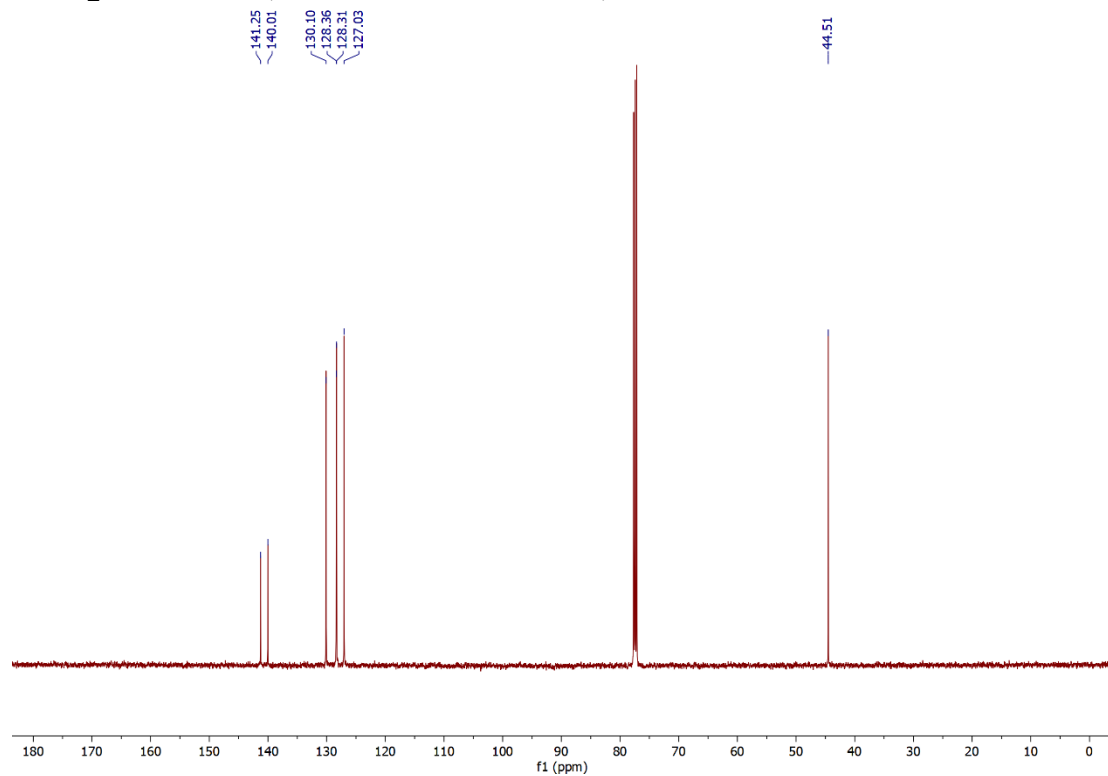
### Compound 5.13, $^{13}\text{C}$ NMR (500 MHz, $\text{CDCl}_3$ )



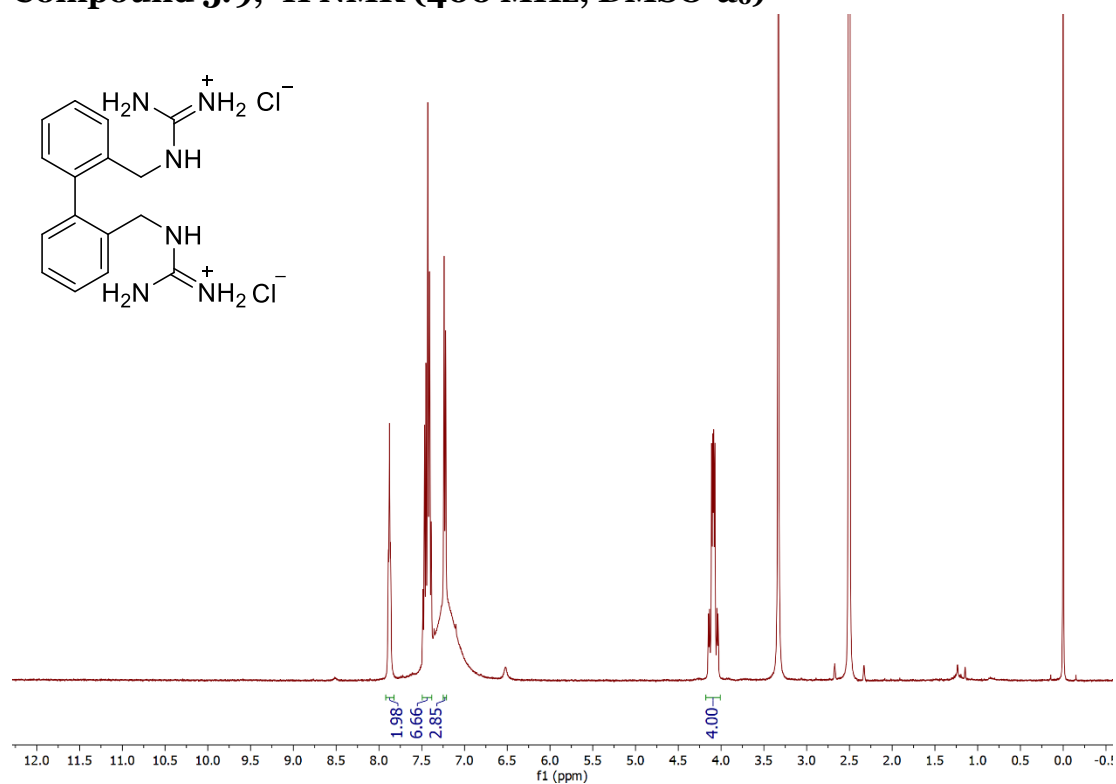
**Compound 5.14,  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )**



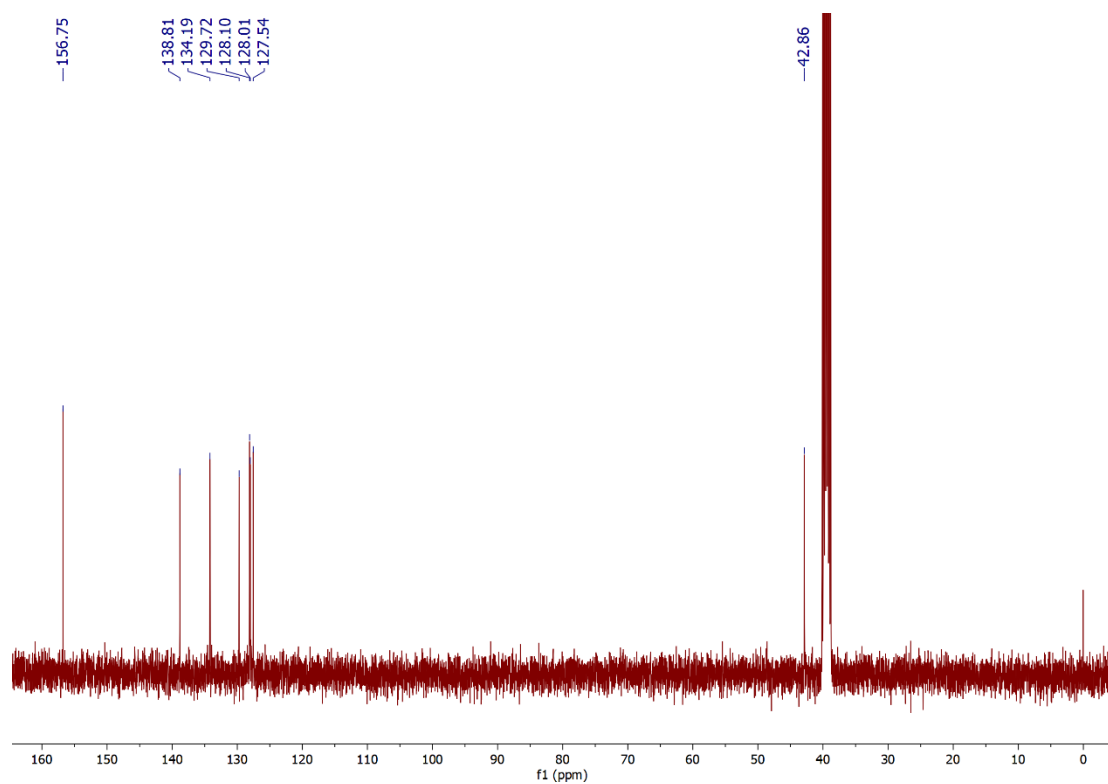
**Compound 5.14,  $^{13}\text{C}$  NMR (500 MHz,  $\text{CDCl}_3$ )**



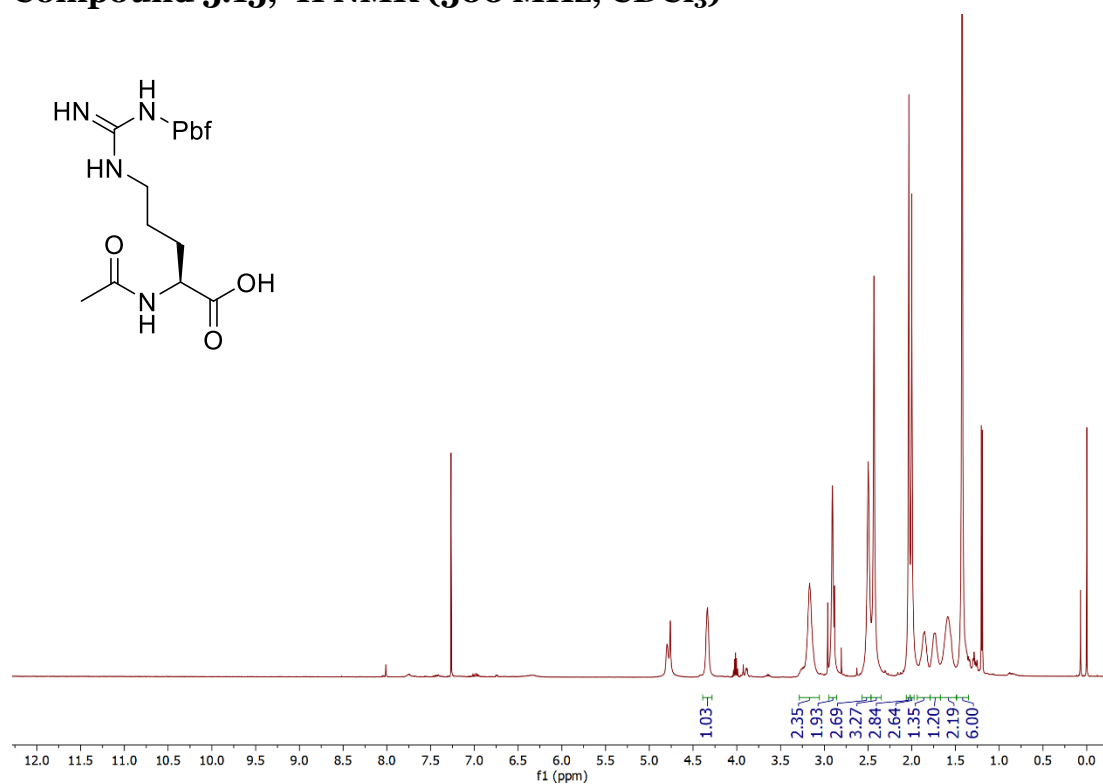
### Compound 5.9, <sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>)



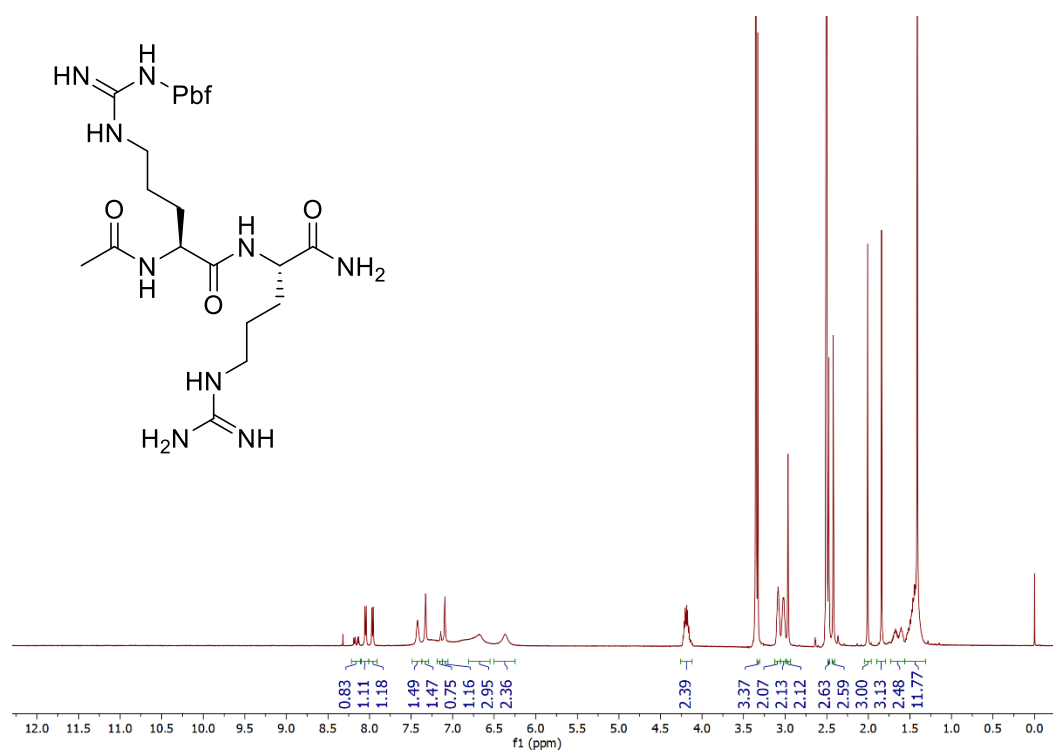
### Compound 5.9, <sup>13</sup>C NMR (400 MHz, DMSO-d<sub>6</sub>)



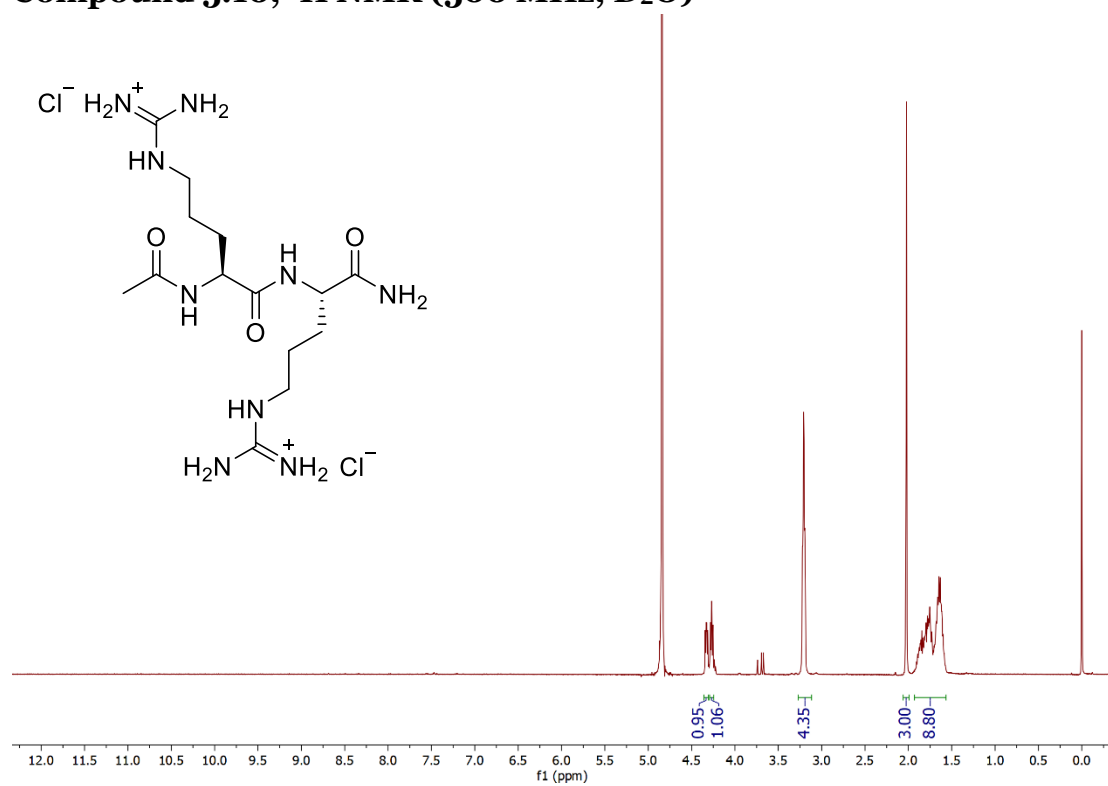
**Compound 5.15,  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )**



**Compound 5.16,  $^1\text{H}$  NMR (500 MHz,  $\text{DMSO}-d_6$ )**

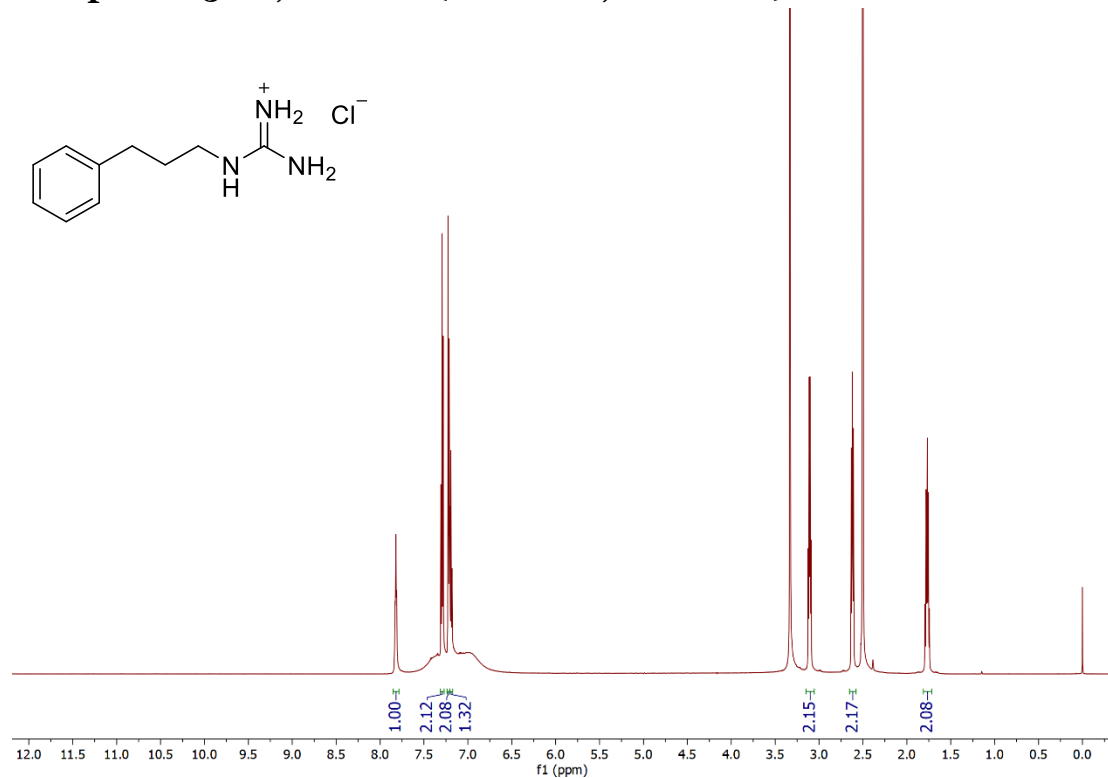


# Compound 5.10, <sup>1</sup>H NMR (500 MHz, D<sub>2</sub>O)

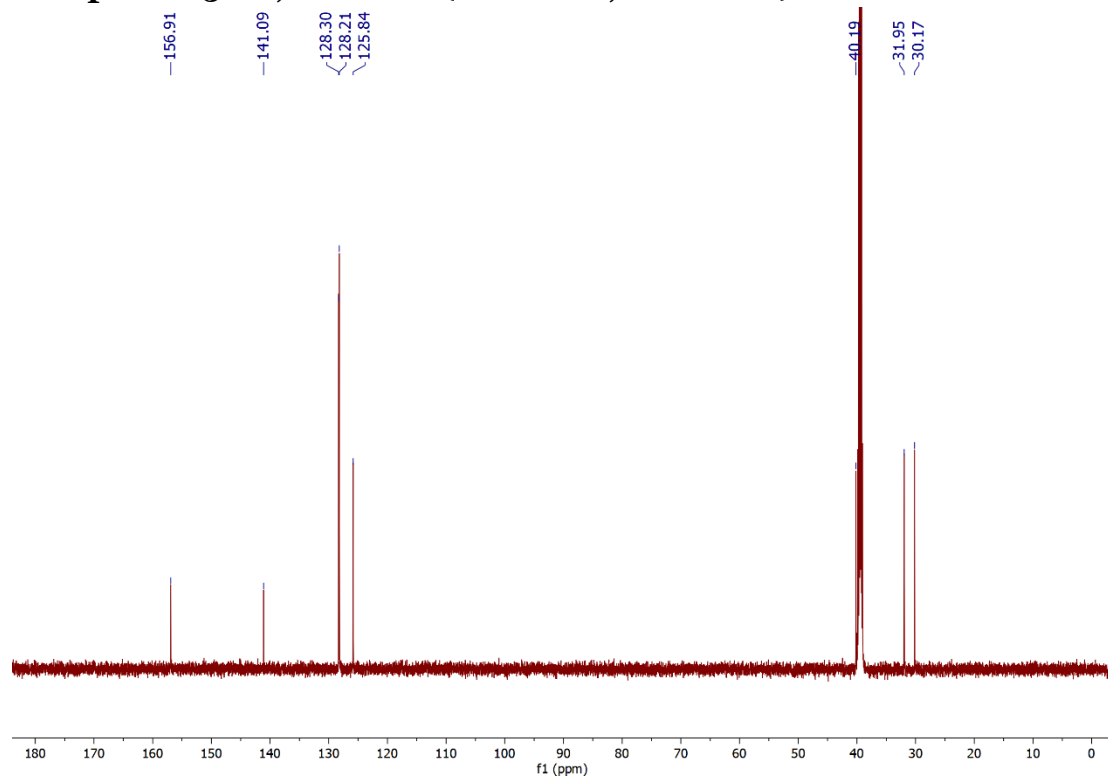




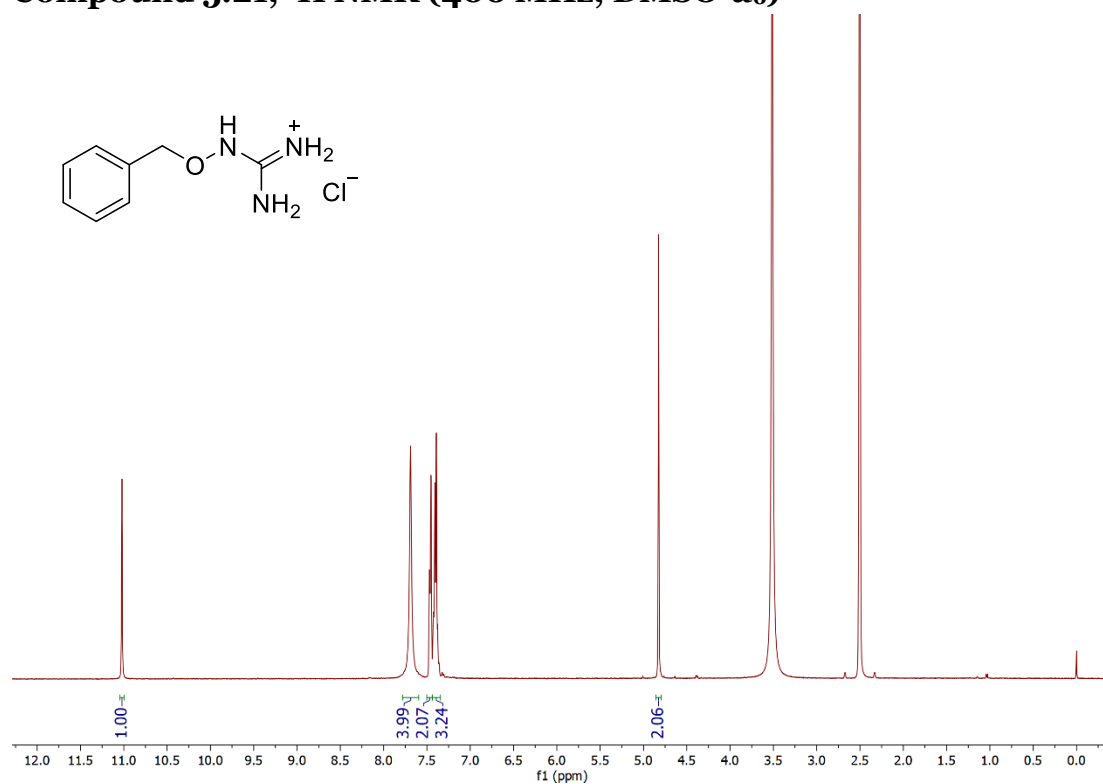
**Compound 5.20,  $^1\text{H}$  NMR (600 MHz,  $\text{DMSO-}d_6$ )**



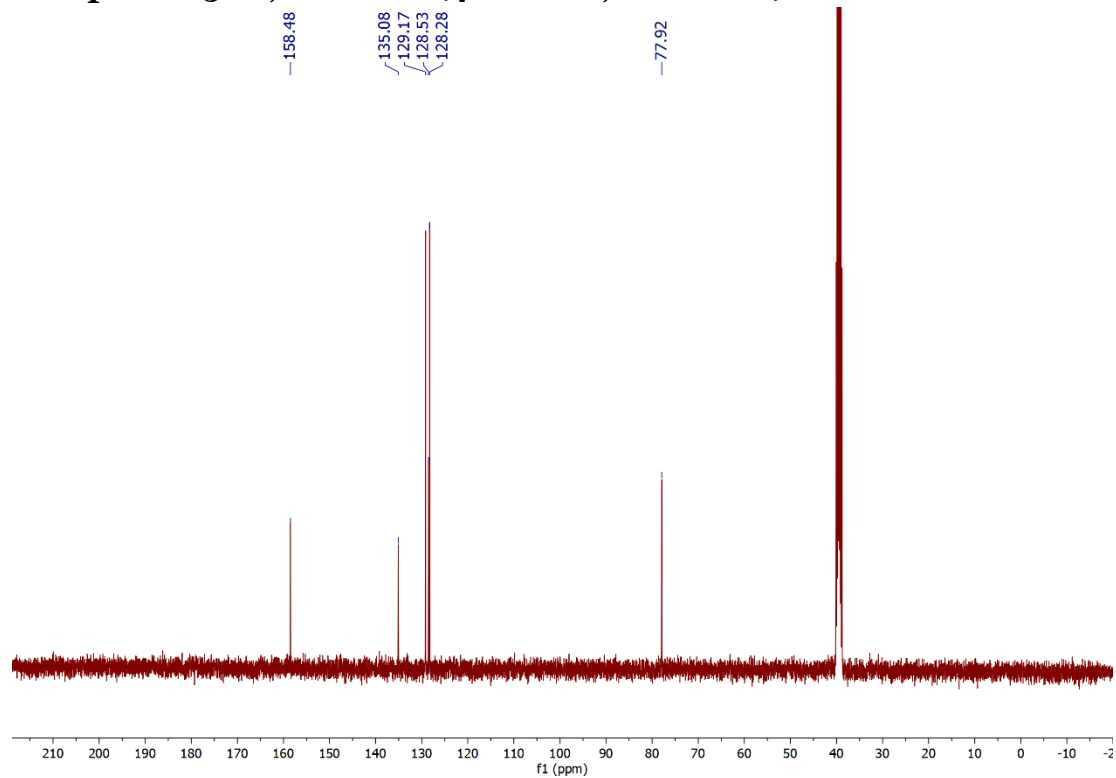
**Compound 5.20,  $^{13}\text{C}$  NMR (600 MHz,  $\text{DMSO-}d_6$ )**



**Compound 5.21,  $^1\text{H}$  NMR (400 MHz,  $\text{DMSO-}d_6$ )**



**Compound 5.21,  $^{13}\text{C}$  NMR (400 MHz,  $\text{DMSO-}d_6$ )**



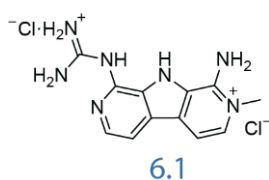
# **Chapter 6**

## **Future Directions**

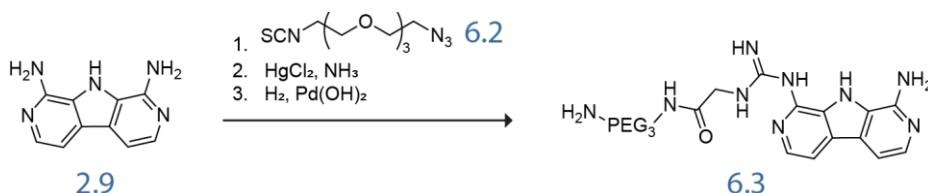
## Cell penetration with GADAC

As discussed in Chapter 2, the fluorescent guanidine **GADAC** has promise as a two-in-one system for cell delivery and fluorescence reporting of biologics. Two improvements to the structure of **GADAC** could be considered. Efforts could be made to force the addition of a second guanidinium group onto the amine of **GADAC** in an attempt to append additional hydrogen bond donors. This addition may result in a higher binding affinity to oxoanions at pH's in which the amidinium would be deprotonated. However, I was unable to append a second guanidine through many conditions screened. In addition, a second guanidine would likely eliminate the increase in quantum yield that was seen upon protonation of the amidinium group.

An improvement that could be valuable for **GADAC**'s fluorescence properties would be the addition of a methyl group to the amidine, as seen in compound **6.1**. This structure mimics the **GADAC** dication, but the amidinium is permanently cationic. The permanent charge could result in the dye retaining its quantum yield of  $\Phi \approx 0.5$  at high pH, whereas the quantum yield of **GADAC** decreases at high pH. One downside to this approach is that the additional amidinium binding site is eliminated. However, it is unknown whether this binding site is involved in oxoanion binding at biological pH, given that the amidinium is mostly deprotonated at pH 7.4. Binding studies on **GADAC** were performed in Milli-Q water, which typically has a pH  $\approx 6$ . Future work should include binding studies of **GADAC** and **6.1** to oxoanions performed at pH 7.4 to determine the effect of the amidinium on oxoanion binding in biological contexts.



In order to append **GADAC** to a biologic for transport, a linker must be installed. I envisaged that this linker could be installed during the formation of the isothiourea. In the original synthesis, an alkyl group, PMB, was used as the protecting group on the isothiocyanate. This PMB group is removed as the final step in the synthesis, and the removal conditions are harsh enough to impact the total yield of **GADAC**. Thus, thiourea formation would be an ideal step for the incorporation of a linker since the linker would replace the PMB group and eliminate the need for the deprotection step. Isothiocyanate **6.2** is commercially available and displays an azide at the end of a short PEG chain. Compound **6.2** can be used with **2.9** to produce a thiourea, and subsequent guanidinylation and reduction of the azide to the amine would result in compound **6.3**. The linker-displaying **GADAC 6.3** could be appended to proteins using the methods developed by Raines and coworkers.<sup>193</sup>



## Oxyguanidine in oxoanion binding

In Chapter 3, I describe the ability of canavanine, an oxyguanidine, to bind to oxanions. I measured this ability through anion-mediated octanol/water partitioning and found that canavanine partitions into octanol less than arginine. The cause of this difference is unclear. In Chapter 5, a benzyloxyguanidine exhibits a higher  $K_a$  than an alkylguanidine in DMSO. Thus, canavanine's preference for the water layer is likely due to its greater hydrophilicity, which is demonstrated by its higher TPSA than arginine

(Chapter 4). While oxyguanidines may not be beneficial in cellular transport, they may still demonstrate binding superior to alkylguanidines in a supramolecular chemistry context. Acylguanidines have low  $pK_a$ 's and have been shown to bind to oxoanions with a greater affinity than alkyl- or arylguanidines.<sup>30,41</sup> To date, oxyguanidines have not yet been explored as oxoanion binders. Thus, future work on oxyguanidine-containing preorganized hosts could be valuable for fields such as anion sensing.<sup>26</sup>

### ***N*<sup>α</sup>-methyl arginine oligomers as CPPs**

I have determined that *N*<sup>α</sup>-methylated arginine derivatives are more effective at partitioning into octanol in the presence of anionic lipids than unmodified arginine. Work is ongoing to synthesize fluorescently labeled *all-N*<sup>α</sup>-methyl arginine oligomers. These peptides, and their unmodified counterparts, will be introduced to human cells. The extent of uptake will be determined through microscopy and flow cytometry.

### **Biaryl-2,2'-bisguanidines in oxoanion binding**

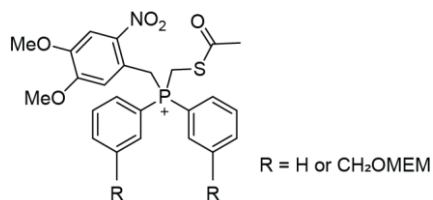
While initial studies to determine the binding affinity of biaryl-2,2'-bisguanidines to oxoanions were inconclusive, further studies on these molecules are warranted. The data presented in Chapter 5 indicate that these structures have minimal binding to oxoanions in water; however, observing strong oxoanion binding in aqueous conditions is rare. Experiments should be performed on these structures in less competitive solvents such as alcohols or non-protic solvents. In addition, UV-Vis spectroscopy should be evaluated as a more appropriate method to determine the binding affinities since the concentrations required are much lower, which would minimize the impact of changes in ionic strength.

# **Appendix A**

## **Efforts toward a photocleavable phosphine protecting group for the traceless Staudinger ligation**

Contributions: All work was performed by Lindsey O. Calabretta.

## Abstract



The traceless Staudinger reaction has been fruitful in its many applications in bioconjugation. Used to bioorthogonally produce amide bonds with no extraneous atoms, this methodology could be used to produce isopeptide linkages in ubiquitin chains. Previous Raines group members have investigated appending a phosphinothioester to the C-terminus of one ubiquitin and reacting it with a second azido-modified ubiquitin. This traceless Staudinger ligation was successful in producing diubiquitin, however, in very low yields. The reaction was hampered by oxidation of the free phosphinothioester, preventing the ligation from occurring. This appendix details efforts towards the synthesis of a protected phosphinothioester to prevent this oxidation. The photocleavable dimethoxynitrobenzyl (DMNB) protecting group was successfully appended onto a water-soluble phosphinothioester as well as commercially available acetylthiomethyl-diphenylphosphine borane complex. The DMNB group coincidentally imparted water-solubility to the diphenylphosphine compound, which is promising for future applications in the ligation of proteins. Future directions for the synthesis and use of this molecule are discussed.

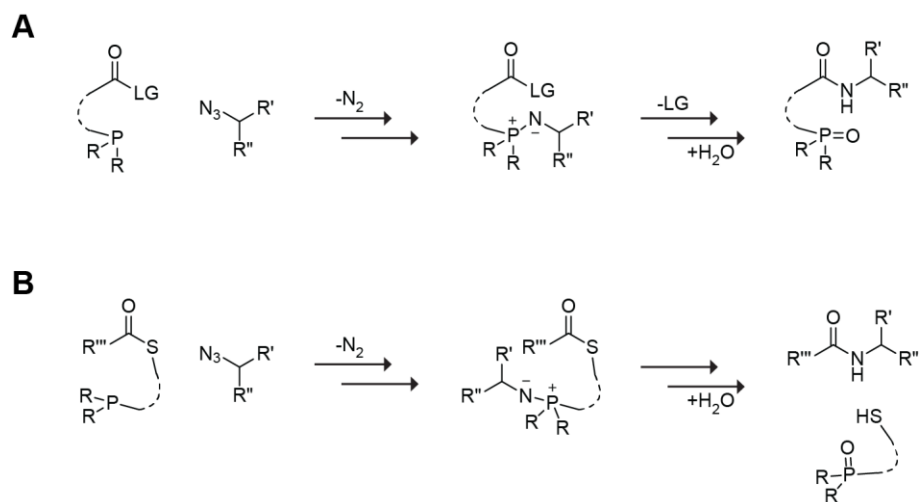


## Introduction

Since its development in 2000, the Staudinger ligation has become a powerful tool for chemical biology.<sup>194,195</sup> The reaction utilizes a phosphine and an azide, two functionalities that are unreactive to biological functional groups, making this reaction biorthogonal. The reaction relies on the phosphine reducing the azide to an iminophosphorane intermediate, upon which the nucleophilic nitrogen of the iminophosphorane becomes trapped by the electrophilic acyl group. Subsequent hydrolysis produces an amide (Figure A.1A). This reaction is useful for producing amide bonds; however, this method also adds extraneous atoms. The traceless Staudinger ligation, developed by the Raines and Bertozzi labs, is a powerful adaptation of this reaction capable of producing new amide bonds without residual atoms.<sup>196,197</sup> In this method, the phosphine is installed onto the acyl group of choice by thioesterification and is integrated as a part of the leaving group. Therefore, the entire phosphine oxide is released after the trapping of the iminophosphorane nitrogen by the acyl group and subsequent hydrolysis of the nitrogen-phosphorous bond, which allows the production of an amide with no extraneous moieties (Figure A.1B).

While non-traceless Staudinger ligation is adequate for uses such as bioconjugation, the traceless Staudinger ligation is useful when extraneous moieties are undesirable, such as in the formation of native peptide bonds. Raines and coworkers have used the traceless Staudinger ligation in the total synthesis of a complete protein, ribonuclease A (RNase A).<sup>198</sup> They synthesized two short peptide fragments via solid phase peptide synthesis, with one fragment containing a phosphinothioester at the C-terminus and a second fragment containing an azide at the N-terminus. These fragments were connected via traceless Staudinger ligation and then appended to a third fragment

(produced biosynthetically) through native chemical ligation to produce fully functional RNase A. This work helped to show the power of the traceless Staudinger ligation, but its full utility in chemical biology has yet to be explored.

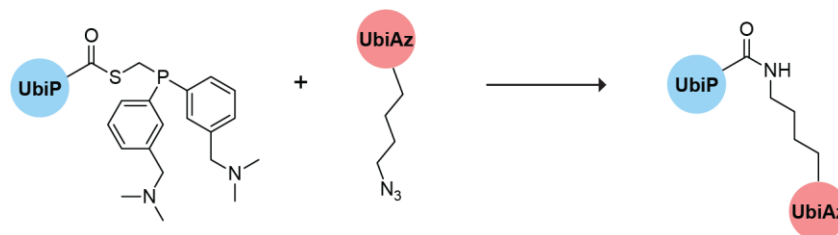


**Figure A.1** a) In non-traceless Staudinger ligation, the phosphine attached to the acyl group, leading to the phosphine oxide remaining on the molecule after amide bond formation. b) In traceless Staudinger ligation, the phosphine is part of the leaving group as a thioester, which is released upon amide bond formation.

Other applications of this chemistry include the formation of native isopeptide linkages between full proteins, such as those found in ubiquitin chains. Ubiquitin is a small, globular protein that is involved in the post-translational modification of other proteins. Mono- or poly-ubiquitin chains are appended on a protein via isopeptide bonds at any one of ubiquitin's seven lysine residues (positions 6, 11, 27, 29, 33, 48, and 63).<sup>199</sup> These modifications commonly signal for protein degradation but can signal other cellular processes depending on the linkage site.<sup>200</sup> The understanding of these cellular signaling pathways is incomplete since less abundant linkages (6, 27, 29, 33) are difficult

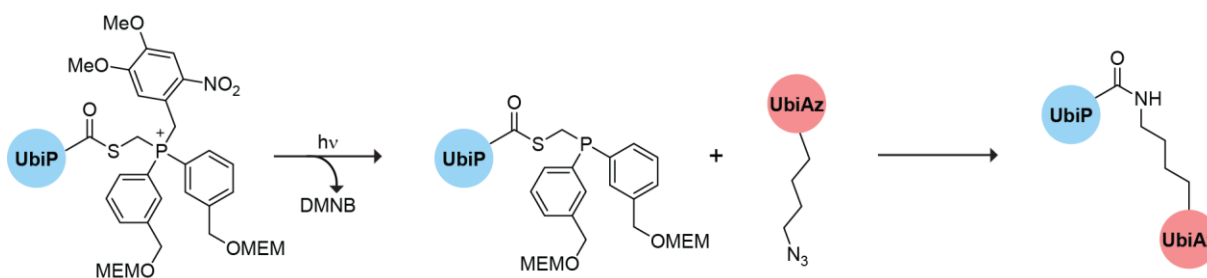
to study. Thus, chemical methods for synthesizing these rare chains in quantities large enough for study are desired.

Martin and Raines have reviewed many chemical syntheses of diubiquitin; however, these methods often require harsh conditions and result in low yields.<sup>201</sup> Previous Raines group members have attempted to apply the traceless Staudinger ligation to the synthesis of diubiquitin to produce diubiquitin in useable yields and under mild conditions. Andersen appended a water-soluble phosphinothioester at the C-terminus of one ubiquitin (UbiP) and performed a traceless Staudinger ligation with a second ubiquitin (UbiAz) that had been genetically modified to replace a lysine with azidonorleucine selectively through non-natural amino acid incorporation (Figure A.2).<sup>202</sup> Aronoff sought to improve this method by exploring modifications of the phosphinothioester phenyl rings with different functional groups and positions of functional groups.<sup>203</sup> While both studies successfully produced diubiquitin, they were confounded by the oxidation and subsequent hydrolysis of the phosphinothioester under the oxidative conditions present during the protein purification and reactions. This oxidation prevents ligation from occurring and results in lower yields. Thus, a method to protect the phosphine from oxidation is required.



**Figure A.2.** Initial diubiquitin synthetic route. Andersen synthesis of diubiquitin via installation of a water-soluble phosphinothioester at the C-terminus of ubiquitin to produce UbiP, and genetic incorporation of an azidonorleucine to produce UbiAz.

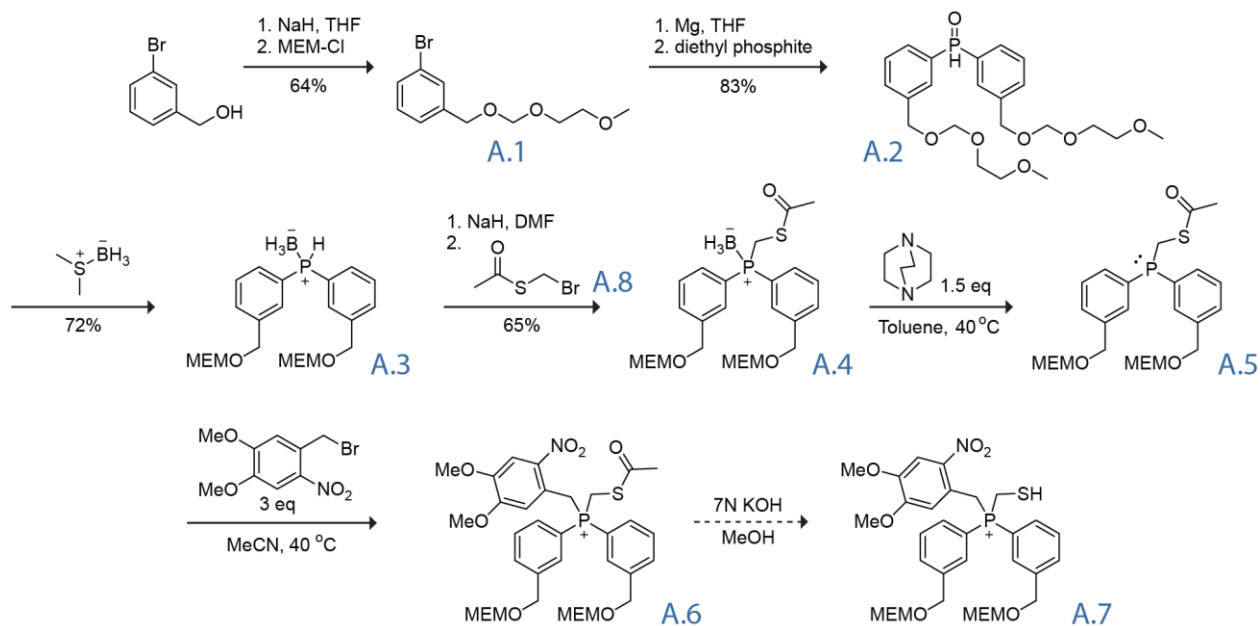
In a recent study, Carrico and coworkers utilized a photo-caged phosphine reagent to perform non-traceless Staudinger ligation *in vivo*.<sup>204</sup> Their 4,5-dimethoxy-2-nitrobenzyl (DMNB) photo-caged phosphine was resistant to oxidation and selectively reacted with azide-labeled glycoproteins in zebrafish larvae when exposed to UV light. Based on this study, I proposed that photo-caging the water-soluble phosphinothioester with DMNB would provide protection from oxidation throughout the lengthy purification process of UbiP. Exposure to light would then provide an efficient deprotection of the phosphinothioester, only once introduced to UbiAz in an oxygen-free environment (Figure A.3). Initially, I sought to incorporate benzyl methoxyethoxymethyl groups at the 3-position on the phenyl rings, as explored by Aronoff, to impart better water solubility and ease of synthesis.<sup>203</sup> I later synthesized the phosphinothioester from a commercially available reagent containing only phenyl moieties. This enhanced traceless Staudinger ligation could be applied to a broad scope of future systems requiring the chemical synthesis of native peptide or isopeptide bonds, and insights gained through the study of this photo-caged phosphinothioester could be broadly applicable to phosphine chemistry in general.



**Figure A.3.** Proposed synthesis of diubiquitin using a photo-caged water-soluble phosphinothioester appended to a UbiP and azidonorleucine-modified UbiAz. The photo-caged phosphinothioester should resist oxidation.

## Results and Discussion

Initially, I utilized a previously optimized route by Chou and Raines to synthesize **A.3**,<sup>205</sup> followed by conditions optimized by Aronoff to produce **A.4** (Scheme A.1).<sup>203</sup> Starting with a commercially available 3-bromobenzyl alcohol, the methoxyethoxymethyl (MEM) group is appended to the benzyl oxygen via a substitution reaction. This group typically serves as an alcohol protecting group, but in this case, it acts similarly to a short PEG chain and imparts water solubility. Next, bromobenzene **A.1** is introduced to magnesium to form a Grignard reagent. The Grignard reagent is then introduced to diethyl phosphite and replaces both ethoxide groups to produce the bisphenylphosphine oxide **A.2**. Borane dimethyl sulfide is then capable of reducing the phosphorous atom and subsequently protecting it to produce the borane phosphine complex **A.3**. A second substitution with the phosphine (deprotonated by NaH) and *S*-bromomethyl ethanethiolate (**A.8**) produces the borane phosphinothioester **A.4**.



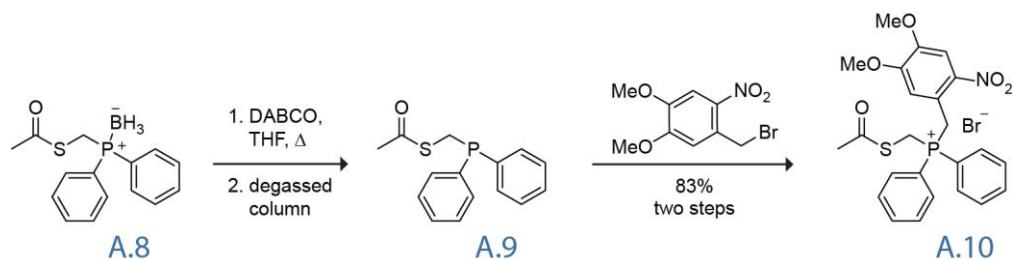
**Scheme A.1.** Synthesis of photocaged methoxyethoxymethyl-phosphinothioester **A.6**.

At this point, I wished to remove the borane protecting group from the phosphine without hydrolyzing the thioester since DMNB-phosphinothioester **A.6** would make for an excellent model system in small molecule tests. The nucleophilic amine 1,4-diazabicyclo[2.2. 2]octane (DABCO) was used to remove the borane from the phosphine by forming a borane amine complex. Typically, DABCO-BH<sub>3</sub> can be separated through chromatography. However, as the phosphine is very prone to oxidation, exposing the free phosphine to oxygen during purification would be unwise. Thus, the mixture of free the phosphinothioester **A.5** (confirmed via LC-MS) and DABCO-BH<sub>3</sub> was used directly in the next step.

I also explored a known protocol in which the borane phosphine complex is heated with methanol to produce the trimethyl borate which can be removed via evaporation (bp 68-60 °C).<sup>206</sup> This process would have been very useful since the free phosphinothioester **A.5** could be easily and quickly purified by evaporation before the subsequent reaction. Unfortunately, this process resulted in very low yields of free phosphine.

Once I had obtained the free phosphine via DABCO, I performed the photo-caging of the phosphine via a substitution reaction described by Carrico and coworkers.<sup>204</sup> As DABCO-BH<sub>3</sub> was still in the reaction mixture, an excess of 4,5-dimethoxy-2-nitrobenzyl bromide was used in the subsequent reaction in case the second free amine of DABCO reacted competitively. I observed the DMNB-phosphinothioester **A.6** to be the major product via LC-MS; however, the reaction is still quite messy and difficult to purify. Since the compound is positively charged, attempts to purify via silica gel chromatography resulted in the product sticking to the baseline. I had moderate success purifying **A.6** using reversed phase chromatography, though the product was not pure enough to obtain reliable NMR data.

In this semi-purified fraction, I observed masses by LC-MS corresponding to the free phosphinothiol **A.7** as well as the triaryl phosphine that would result from the elimination of the thiol as thioformaldehyde (Figure A.5). The crude reaction mixture was not immediately purified, so it is unknown if this degradation is a result of time or if the reaction conditions contribute to the formation of these degradation products.



**Scheme A.2.** Synthesis of DMNB-protected acetylthiomethyl-diphenylphosphine **A.10**.

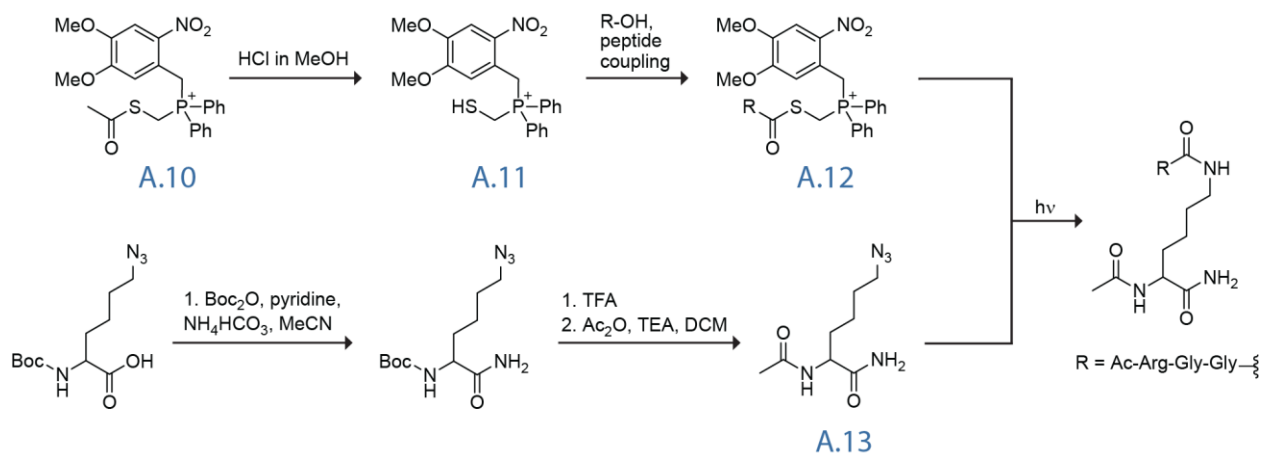
Given that the synthesis of **A.4** is arduous, I decided to attempt the DMNB protection of a simpler model phosphine to optimize the protection route. Acetylthiomethyl-diphenylphosphine borane complex **A.8** is commercially available for use in Staudinger ligations. I again successfully removed the borane group using DABCO; however, as mentioned earlier, I was concerned about having excess DABCO in the following step. I noticed the DABCO and DABCO-BH<sub>3</sub> stuck resolutely to the baseline upon TLC with ethyl acetate. Instead of performing a full column, I prepared a pipette column and flushed the silica with N<sub>2</sub>(g) for several hours. The reaction was filtered through the pipette column with degassed ethyl acetate, which resulted in the pure deprotected intermediate (**A.9**) with no apparent oxidation of the phosphine. The intermediate (**A.9**) was then exposed to 4,5-dimethoxy-2-nitrobenzyl bromide to produce the desired protected phosphine **A.10**. Coincidentally, DMNB-phosphine **A.10** is soluble

in water, so further attempts to synthesize methoxyethoxymethyl phosphine **A.6** were deemed unnecessary.

## Future Directions

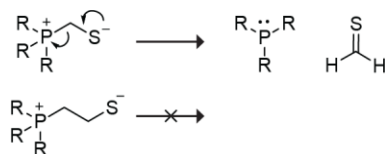
Before attempting the Staudinger ligation with the DMNB-phosphine on ubiquitin, it would be best to test the reaction on a model system. Performing a traceless Staudinger ligation on the thioacetate compound (**A.10**) and a small molecule azide to produce an acetamide would be the simplest method. However, once the DMNB group has been removed from **A.10**, the resulting phosphine might not be soluble enough in water. If attached to a protein, the solubility of the phosphine is negligible, but in a model system, it could be detrimental. The last three amino acids of the C-terminus of ubiquitin are Arg-Gly-Gly. If appended onto the thiol, this short peptide should impart enough solubility to the free phosphine. Initial tests to release the free thiol (**A1.11**) from the thioacetate (**A.10**) using 1.25 M HCl in MeOH were successful, though purification proved challenging. An Ac-Arg-Gly-Gly-OH peptide could be synthesized through traditional peptide synthesis and the thiol (**A.11**) could be appended using traditional coupling reagents. Azidonorleucine (**A.13**) could be synthesized using methods discussed in Chapters 3 and 4 and would make a suitable model compound for the proof-of-concept traceless Staudinger ligation. If the model system is effective at producing an isopeptide linkage between the tripeptide and azidonorleucine, the DMNB-protected phosphine (**A.11**) could be incorporated into the synthesis of diubiquitin as described by Andersen.<sup>202</sup>





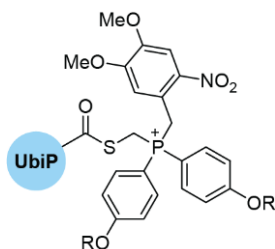
**Figure A.4.** As a model system, thiol (**A.11**) could be appended to a short peptide to impart water solubility and reacted with an azidonorleucine diamide (**A.13**).

There are some factors of this route that future researchers consider. The degradation of the phosphinothioester via elimination of thioformaldehyde could prove to be a confounding issue. An alternative route one could explore would be to simply add an extra methylene between the phosphine and sulfur atoms. The thioester 2-bromoethyl thioacetate can be synthesized by reacting potassium thioacetate and 1,2-dibromoethane,<sup>207</sup> and could replace the *S*-bromomethyl ethanethiolate (**A.8**) in the synthesis of alternative molecules. Thiols are mildly acidic ( $pK_a \sim 10$ ), and the free thiol in phosphinothioester (**A.11**) is likely made even more acidic by the adjacent positive charge. Thus, the sulfur is easily deprotonated, allowing electrons to form a double bond with the methylene carbon, with subsequent reduction of the positively charged phosphorous to release the tertiary phosphine (Figure A.5). By placing a methylene unit between the charged atoms, an elimination would be unlikely.



**Figure A.5.** The potential pathway for the degradation of the phosphinothioester. Elimination of thioformaldehyde is feasible from the quaternary phosphine found in (A.11) (above), while the additional methylene group makes this elimination unlikely (below).

By inserting an extra methylene group between the phosphine and the thioester, the traceless Staudinger ligation will occur through a six-membered ring transition state during the N→S acyl transfer step, as opposed to the five-membered ring transition state in the current system. In initial studies of the traceless Staudinger reaction, 2-phosphinobenzenethiols were explored.<sup>197,198</sup> These structures contain two carbons between the phosphine and thioester, and thus must go through a six-membered transition state during traceless Staudinger ligation. While ligation reactions with these compounds were successful, amide product yields were lower than those with structures that went through five-membered ring transition states, likely due to slower reaction rates. Thus, when exploring this system, one would have to determine whether the benefit of decreased degradation of the phosphinothiol out-weighs the potential of lower yields.



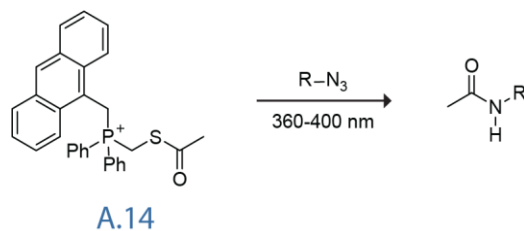
**Figure A.6.** Photo-caged phosphinothioester with potentially enhanced nucleophilicity, due to *para*, conjugated oxygens donating electron density.

If the protection strategy proves successful, one could investigate synthesizing a modified phosphinothioester that could enhance the rate of the traceless Staudinger ligation. In this modified version, substituents on the phenyl rings could be positioned *para* to the phosphine, with the oxygen conjugated with the ring (Figure A.6). The oxygen would then donate electron density through the ring to the phosphine. With the increased electron density, the nucleophilicity of the phosphorous would be greatly enhanced, which could potentially increase the rate of traceless Staudinger ligation. This phosphine would likely only be viable in the photo-caged system, as the phosphine would also have increased propensity for oxidation.

## Epilogue

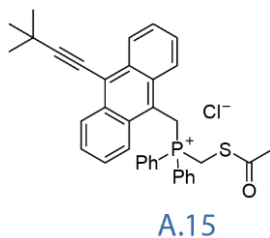
In 2016, Lam and coworkers developed anthracene-protected traceless Staudinger ligation reagent **A.14** by first synthesizing **A.9** and then reacting it with 9-(chloromethyl)anthracene.<sup>208</sup> They observed that the anthracene was cleavable by light, with a maximum quantum efficiency at 376 nm. Though, under broadband irradiation, the yield of deprotected phosphine was 58%. To probe the ability of **A.14** to undergo a traceless Staudinger ligation, they exposed it to three azide-containing compounds in 3:1 THF/water. They observed no formation of the amide Staudinger ligation product in dark conditions but did observe amide product yields of 45–55% when the reaction mixtures were exposed to 360–400 nm broadband irradiation (Scheme A.3).

In a subsequent paper published in 2018, Lam and coworkers explored the addition of substituents to the 10-position of anthracene to increase the  $\lambda_{\text{abs}}$  of the photoprotecting group into the visible light range.<sup>209</sup> This would be beneficial for *in cellulosa* or *in vivo* experiments, as UV light is harmful to life.



**Scheme A.3.** An anthracene-photocaged phosphine (**A.14**) reacts with azides upon exposure to UV light.

The authors sought to alter the absorption wavelength by directly conjugating either a phenyl ring or an alkynyl group to the 10-position. They observed minimal change in absorbance with the addition of the phenyl rings (containing a *para* trifluoromethyl, methoxy, or hydrogen), likely due to the steric strain of the compound forcing the phenyl ring to be perpendicular to the anthracene and out of conjugation. They saw a significant red-shift in absorbance with the alkynyl groups, and focused their efforts on compound **A.15** since it gave the fastest photo-uncaging and highest yield of the deprotected phosphine.



The authors successfully used compound **A.15** in traceless Staudinger ligation reactions under broadband visible light that was filtered to be  $>420$  nm. They observed yields of  $\sim 50\%$  for small molecule azides. The free thiol of **A.15** was then loaded with an amino acid monomer or dimer and reacted with azido-modified amino acid di-, tri-, and tetramers to produce oligomers in yields of  $\sim 30-40\%$ .

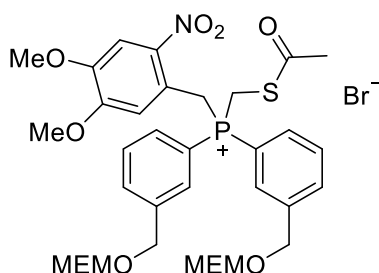
While an important addition to the field of photo-caged biorthogonal reactions, these studies contain many drawbacks. Notably, the reactions performed in these studies were performed in 3:1 THF/water<sup>208</sup> or 3:1 THF/buffer.<sup>209</sup> The authors claim in both studies that **A.14** and **A.15** are soluble in water, yet give no evidence under what conditions or concentrations. One wonders if the molecules were truly soluble in water, why the traceless Staudinger ligation reactions were performed in organic/aqueous mixtures when 100% aqueous buffer would be the most biologically relevant solvent. In addition, the oligomers Lam and coworkers produced contain exclusively hydrophobic amino acids with no heteroatoms, perhaps due to solubility issues when running the reactions in a majority organic solvent. Thus, the results could have limited scope as for traceless Staudinger ligations using **A.15** with compositionally diverse peptides and proteins.

Finally, while the photo-triggered ligation using **A.15** with small-molecule azides produced amide yields of ~50%, ligations using the unprotected phosphine **A.9** produced amide yields of ~75%. As mentioned earlier, the photocleavage of **A.14** was only 58% efficient. The authors discuss that incomplete cleavage of the anthracene could be responsible for the lower ligation product yield when compared to reactions with **A.9**. They also hypothesize that byproducts of the photocleavage event may be interfering with the ligation reaction.

Given this evidence, I believe a DMNB-protected phosphine may still be worth pursuing. The DMNB group is assured to impart more water solubility than a highly hydrophobic anthracene. In addition, while DMNB photo-cleavage occurs best under UV light, photo-cleavage at 420 nm has been shown to proceed at reasonable rates.<sup>210</sup> Carrico and coworkers have also used broadband UV light to photo-cleave a DMNB-phosphine

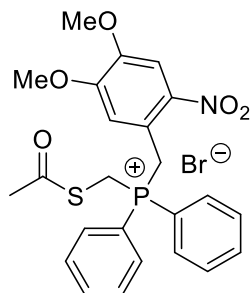
and trigger a Staudinger ligation in live cells and life zebrafish.<sup>204</sup> Therefore, a focus on visible-light cleavable photoprotecting groups may be unnecessary. These studies, however, are still foundational and informative for future work. Any future researchers studying DMNB-protected phosphines should be aware of photocleavage efficiency and potential byproducts interfering with the desired traceless Staudinger ligation.

## Synthesis



**((Acetylthio)methyl)(4,5-dimethoxy-2-nitrobenzyl)bis(3-(((2-methoxyethoxy)methoxy)methyl)phenyl)phosphonium bromide (A.6)** The borane phosphine thioester complex **A.4** (50 mg, 0.10 mmol) and 1,4-diazabicyclo[2.2.2]octane (17 mg, 0.15 mmol) were dissolved in toluene (1 mL), and the resulting solution stirred under N<sub>2</sub>(g) at 40 °C for 6 h. The reaction was concentrated under reduced pressure and used directly. The crude phosphine thioester **A.5** (17 mg, 0.15 mmol) and 4,5-dimethoxy-3-nitrobenzene (82.8 mg, 0.30 mmol) were dissolved in acetonitrile (1 mL), and the resulting solution stirred for 4 h at 40 °C under N<sub>2</sub>(g), protected from light with foil. The product was purified by reversed phase chromatography, but the resulting fractions were still too impure to quantify yield or analyze by NMR.

**LC-MS** *m/z* calcd for C<sub>34</sub>H<sub>45</sub>NO<sub>11</sub>PS [M + H]<sup>+</sup>, 706.24; found 706.20.



**((Acetylthio)methyl)(4,5-dimethoxy-2-nitrobenzyl)diphenylphosphonium**

**bromide (A.10)** Acetylthiomethyl-diphenylphosphine borane complex (20 mg, 0.07 mmol) and DABCO (11.7 mg, 0.1 mmol) were added to a dry flask and placed under a dry condenser. THF (2 mL) was added, and the solution was refluxed under N<sub>2</sub>(g). Concurrently, a silica pipette column was flushed with N<sub>2</sub>(g) and EtOAc was sparged with N<sub>2</sub>(g). After two hours, TLC in 20% EtOAc/Hex showed the absence of starting material. The reaction was concentrated under reduced pressure, and redissolved in degassed EtOAc. The solution was then filtered through the degassed pipette column into a new sealed container (the tip of the pipette was pierced through a septa), flushing with EtOAc and maintaining pressure with N<sub>2</sub>(g). The solution was concentrated under reduced pressure and redissolved in THF (2 mL). 4,5-Dimethoxy-3-nitrobenzene (27.5 mg, 0.1 mmol) was added, and the solution was returned to reflux under N<sub>2</sub> for three hours. The solution was removed from heat and placed in an ice bath. The solution was then filtered and rinsed with chilled ether to produce a yellow solid (32 mg, 83% yield). In some cases, filtration was not enough to ensure pure product, in which case the material was purified through a reversed phase chromatography gradient, eluting at 40% MeCN/water.

<sup>1</sup>H NMR (300 MHz, D<sub>2</sub>O,  $\delta$ ): 7.94–7.83 (m, 2H), 7.82–7.62 (m, 10H), 6.71 (d,  $J$  = 2.5 Hz, 1H), 4.95 (d,  $J$  = 14.6 Hz, 2H), 4.46 (d,  $J$  = 8.1 Hz, 1H), 3.91 (s, 3H), 3.73 (s, 3H), 2.25 (s, 3H).

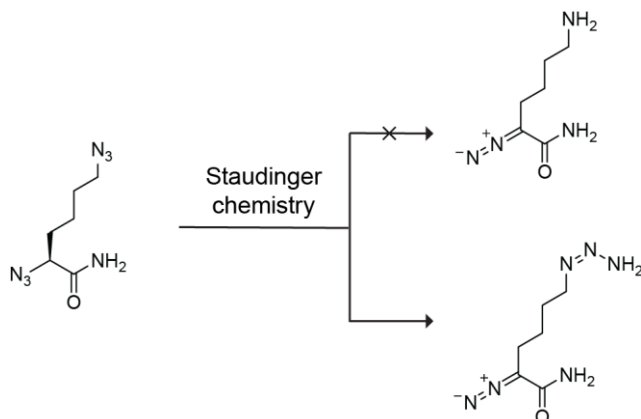
## **Appendix B**

### **Efforts Towards the Site-Selective Introduction of a Diazo Group at the N-Terminus of a Protein**

Contributions: All work was performed by Lindsey O. Calabretta.



## Abstract



The ability to selectively install a biorthogonal linker onto a protein of interest is a common goal in chemical biology. Often methods utilize non-natural amino acid incorporation to site-specifically modify a protein; however, these methods require genetic modification techniques and the biosynthesis of full proteins. Other attempts to introduce biorthogonal linkers are often effective, but as they target amino acid residues that might be numerous on a protein, they are often not selective. I explored a system to impart selective modification of a protein at the N-terminus in which all amines of a protein (lysines and the N-terminus) could display azido groups. Subsequently, through Staudinger chemistry, the lysine azides could be reduced back to amines, and the N-terminus could be converted selectively into a diazo group. Initial tests on lysine amide afforded a diazo group at the N-terminal site, but the lysine  $\epsilon$ -position was converted to a triazene, a functional group not easily reduced under mild biologically tolerant conditions.

## Introduction

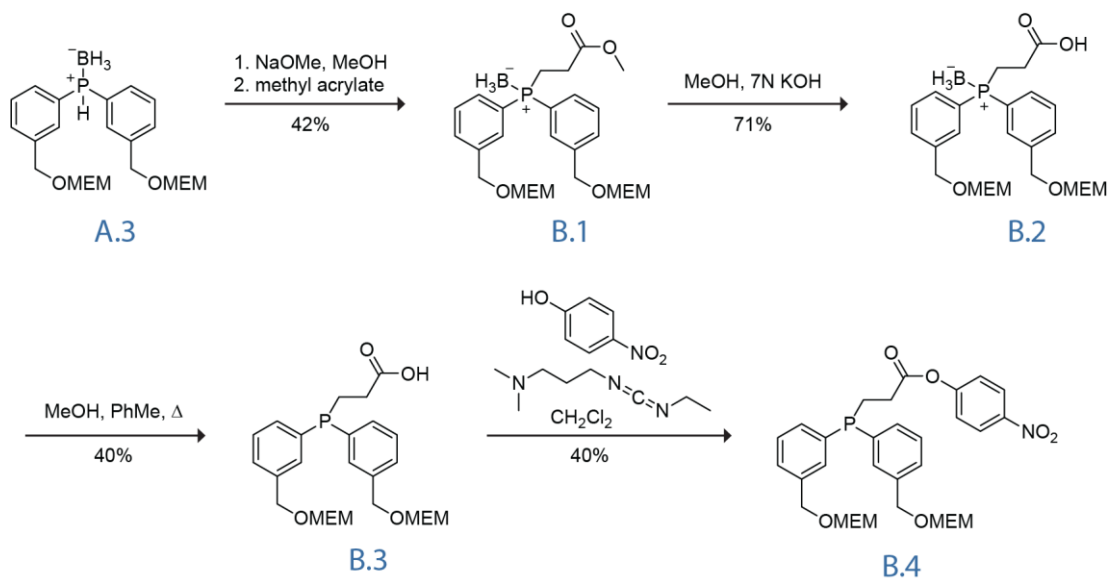
Methods to introduce site-specific bioconjugations are highly sought after in chemical biology. Bioconjugation can be useful in many areas, from observing proteins *in vivo* by conjugating fluorophores to producing therapeutic agents, such as antibody–drug conjugates. Lysine and cysteine residues on proteins are commonly utilized as nucleophiles to react with electrophile conjugates. Tyrosine and tryptophan residues can also be utilized through transition metal catalysis or electrophilic aromatic substitution.<sup>211</sup> However, since there could be several residues of a single type within a protein, this can lead to non-specific conjugation. Site-specific modifications can be achieved through non-natural amino acid incorporation; however, this requires the use of genetic modification and *de novo* protein synthesis.<sup>212</sup>

Another method to site-specifically modify proteins is to target the C- or N-terminus. The N-terminus has been under-utilized as a target, with only a few modifications having been explored at this site.<sup>213,214</sup> In particular, Francis and coworkers have taken advantage of the acidity of the  $\alpha$ -proton at the N-terminus by using pyridoxal-5-phosphate to produce a reactive carbonyl, which can be further conjugated with a hydroxylamine to form an oxime.<sup>215</sup> Myers and Raines have also utilized the acidity of the  $\alpha$ -proton to convert  $\alpha$ -azido carbonyls to diazo compounds through a process called azide deimidogenation.<sup>216</sup> This reaction utilizes a phosphinoester, and is in many ways related to the Staudinger reaction, a reaction that reduces azides to amines. Thus, I reasoned that one could convert the N-terminus to a diazo group by first transforming all amino groups of a protein or peptide to azides. Exposing these azides to a water-soluble phosphinothioester could reduce all lysine side-chain azides back to amines through the

Staudinger reaction, while the N-terminal azide is converted to a diazo through deimidogenation due to the acidity of the  $\alpha$ -proton. In this way, one could install a biorthogonal linker site-specifically at the N-terminus.

## Results and Discussion

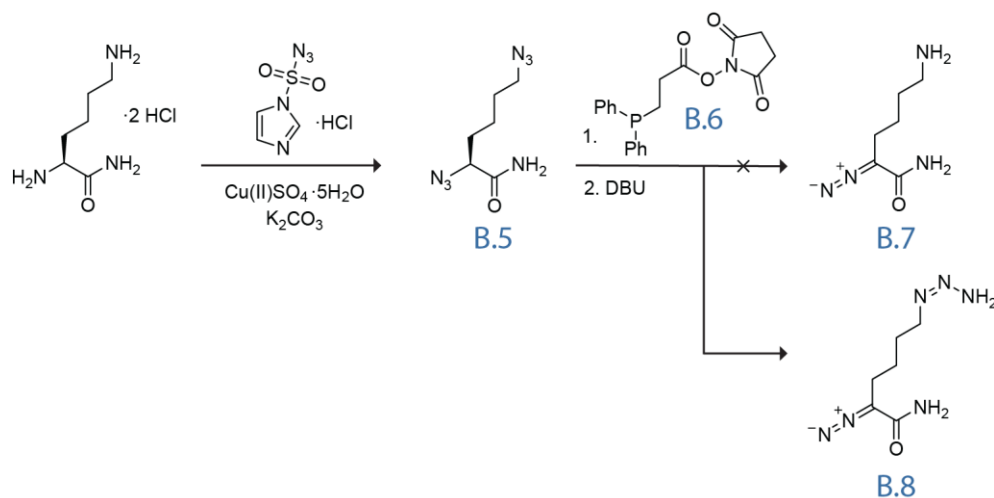
The synthesis of the water-soluble phosphinoester was initially developed by Chou and Raines, and overlapped nicely with the synthesis of the phosphinothioester used in Appendix A.<sup>205</sup> Starting from phosphine-borane complex **A.3**, methyl acrylate is used in a Michael addition-type reaction to append an ester. The ester is then hydrolyzed in base, and the phosphine is deprotected by refluxing in methanol to remove the borane as trimethyl borate.<sup>206</sup> Lastly, a diimide-catalyzed transesterification with *p*-nitrophenol produces phosphinoester **B.1** (Scheme B.1). Chou and Raines have shown deimidogenation reactions with this phosphinoester to be successful in aqueous conditions in neutral pH. Reactions with **B.4** also have the added benefit of being chromogenic upon release of *p*-nitrophenol.



**Scheme B.1.** The synthesis of a water-soluble phosphinoester (**B.4**) for aqueous azide

deimidogenation. Procedures outlined by Chou and Raines<sup>205</sup> were successfully replicated.

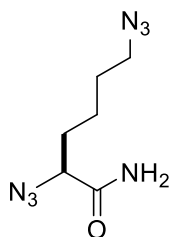
Concurrently, I investigated the initial diazo transfer reaction using lysine amide as a model (Scheme B2). Using the diazo transfer reagent 1*H*-Imidazole-1-sulfonyl azide, I successfully produced di-azide lysine amide **B.5** in both organic<sup>217</sup> and aqueous<sup>218</sup> conditions, though aqueous conditions resulted in slightly lower yields. Next, I tested the deimidogenation reaction in organic conditions with a commercially available phosphinoester (**B.6**) to determine the reaction's feasibility. The results of the deimidogenation were puzzling. By <sup>1</sup>H-NMR and <sup>13</sup>C-NMR I had produced a diazo compound; the  $\alpha$ -proton was no longer seen by <sup>1</sup>H-NMR, and the signal corresponding to the  $\alpha$ -carbon had shifted upfield, similar to the shifts of this carbon in other work.<sup>205</sup> However, the product's mass was 184, much higher than that of the desired product (**B.7**) and two units higher than the expected mass for a product with a diazo group at the  $\alpha$ -position and an azido group at the  $\epsilon$ -position.



**Scheme B.2.** A diazo transfer with 1*H*-Imidazole-1-sulfonyl azide can be done in methanol or water to produce di-azide **B.5**. Deimidogenation with a commercially available phosphinoester (**B.6**) and subsequent hydrolysis produces compound **B.8**.

At this point, I encountered work by Carvalho *et al.* detailing the basic hydrolysis of acyltriazenes to triazenes.<sup>219</sup> As an acyltriazene is an intermediate in the deimidogenation pathway, I concluded the  $\epsilon$ -position must be reacting with the phosphinoester in a process similar to the first steps of deimidogenation. However, as the adjacent proton is not acidic enough to continue reacting toward the diazo upon addition of base, the acyltriazene is hydrolyzed, producing triazene **B.8**. As no mild reactions were found to either convert the triazene to an amine or back to an azide in conditions tolerable to a protein, no further work on this project was pursued.

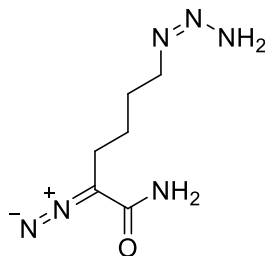
## Synthesis



### (S)-2,6-diazidohexanamide (B.5)

*Organic solvent conditions:* L-Lysine amide dihydrochloride (150.0 mg, 0.68 mmol), potassium carbonate (294.2 mg, 2.12 mmol), and copper(II) sulfate pentahydrate (34.2 mg, 0.14 mmol) were dissolved in methanol (3.5 mL) with vigorous stirring. 1*H*-Imidazole-1-sulfonyl azide hydrochloride (356.8 mg, 1.65 mmol) was added, and the solution changed from dark purple to light blue. After 40 h, 15 ml of DI water was added, and the color of solution was extracted into DCM (3 x 10 mL), dried with Na<sub>2</sub>SO<sub>4</sub>(s), filtered, and concentrated under reduced pressure. The product was purified via silica gel chromatography, eluting in a gradient of 1:4 to 1:0 EtOAc/hexanes to produce **A2.6** as a white solid (118.2 mg, 87% yield).

*Aqueous solvent conditions:* L-Lysine amide dihydrochloride (20.0 mg, 0.09 mmol) and copper(II) sulfate pentahydrate (5.0 mg, 0.02 mmol) were dissolved in H<sub>2</sub>O (10 mL). A 1 M solution of sodium carbonate was added until the pH was approximately 11. 1*H*-Imidazole-1-sulfonyl azide hydrochloride (69.4 mg, 0.32 mmol) was added, and the solution changed from purple to light blue. After reacting overnight, the solution was extracted into DCM (3 x 10 mL), dried with Na<sub>2</sub>SO<sub>4</sub>(s), filtered, and concentrated under reduced pressure. The product was purified via silica gel chromatography, eluting in a gradient of 1:4 to 1:0 EtOAc/hexanes to produce **B.5** as a white solid (10.9 mg, 55% yield). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>, δ): 6.34 (s, broad, 1H), 5.91 (s, broad, 1H), 4.01 (dd, J = 7.0, 1.7 Hz, 1H), 3.31 (t, J = 6.7 Hz, 2H), 1.93 (m, 2H), 1.65 (m, 2H), 1.54 (m, 2H). <sup>13</sup>C NMR (400 MHz, CDCl<sub>3</sub>, δ): 171.37, 63.5, 50.74, 31.23, 28.13, 22.17.



**(Z)-2-diazo-6-(triaz-1-en-1-yl)hexanamide (B.8)** Azide (**B.5**) (118.1 mg, 0.60 mmol) was dissolved in dry THF (6 mL) and H<sub>2</sub>O (0.9 mL) that had been sparged with N<sub>2</sub>. *N*-Succinimidyl 3-(diphenylphosphino)propionate (468.4 mg, 1.32 mmol) was added, and the light yellow solution was stirred under N<sub>2</sub>(g). After 3.5 h, 1,8-Diazabicyclo(5.4.0)undec-7-ene (270 μL, 1.78 mmol) was added, upon which the solution turned bright orange and was allowed to stir overnight. The solution was diluted with brine (20 mL), extracted into DCM (3 x 10 mL), dried with Na<sub>2</sub>SO<sub>4</sub>(s), filtered, and

concentrated under reduced pressure. The product was purified via silica gel chromatography, eluted in EtOAc to produce **B.8** as a yellow solid (20.6 mg, 19% yield).

**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>,  $\delta$ ): 5.34 (s, broad, 2H), 3.33 (t, J = 6.0 Hz, 2H), 2.36 (t, J = 7.2 Hz, 2H), 1.65 (m, 4H). **<sup>13</sup>C NMR** (400 MHz, CDCl<sub>3</sub>,  $\delta$ ): 168.13, 50.88, 42.03, 27.80, 24.80, 23.32. **HRMS**  $m/z$  calcd for C<sub>6</sub>H<sub>12</sub>N<sub>6</sub>O [M + H]<sup>+</sup>, 185.1151; found 185.1142.

## Appendix C

### **1-[3-(Diphenylphosphino)-propanoyl]-2,5-pyrrolidindione**

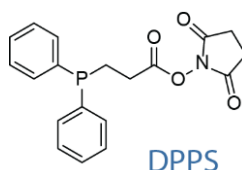
Reproduced from Sharma, I., Orgren, L. R. & Raines, R. T. 1-[3-(Diphenylphosphino)-propanoyl]-2,5-pyrrolidindione. *eEROS*, **2017**, 1–3

Contributions: Lindsey O. Calabretta prepared this manuscript.



## Introduction

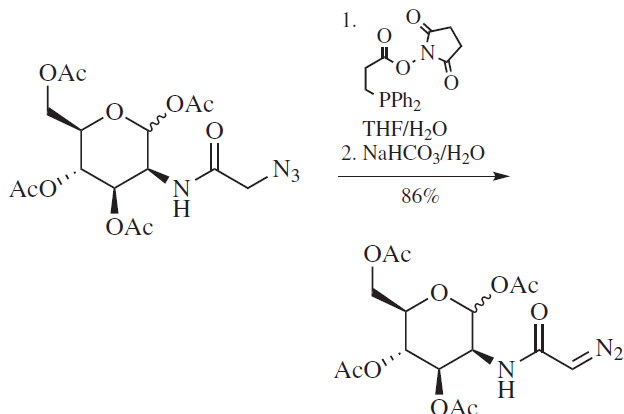
Raines and coworkers have explored many avenues in phosphine-mediated Staudinger chemistry, one of which being the use of 1-[3-(diphenylphosphino)propanoyl]-2,5-pyrrolidindione (**B.6** or **DPPS**) in deimidogenation reactions to convert azides into diazo compounds (see Appendix B).<sup>216</sup> In 2016, the database “Encyclopedia of Reagents for Organic Synthesis” requested that we write an addition to the existing article on **DPPS** with updates on recent literature reactions that employed this compound. The update we provided is reproduced below.



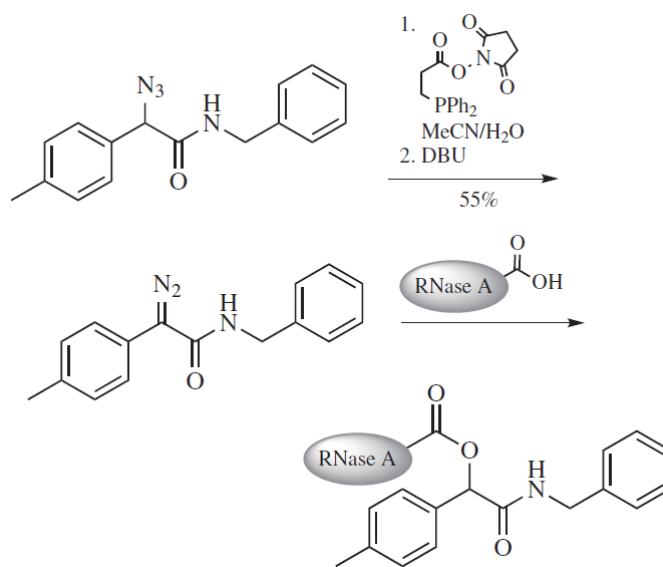
## Azides to Diazo Compounds

Deimidogenation of azides with DPPS was used to produce a variety of diazo compounds, extending the scope of the seminal paper.<sup>220,221</sup>

Andersen et al. established the biological stability of a diazo group by using **DPPS** to produce a diazo-functionalized N-acetylmannosamine (eq 1), which was found to endure cellular metabolism and label the surface of mammalian cells.<sup>222</sup>

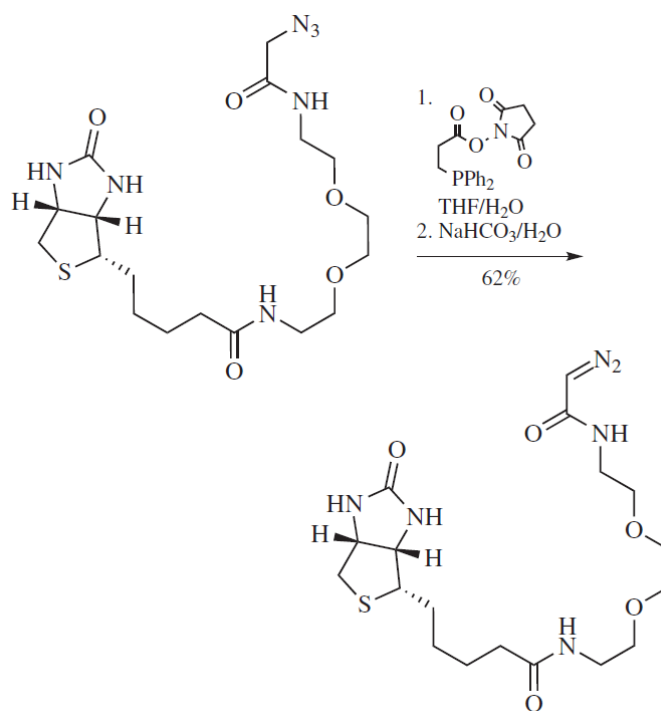


Mix and Raines used **DPDS** to generate diazo compounds that could label proteins by esterification. Through a Hammett analysis, a diazo-functionalized (p-methylphenyl)-glycinamide was identified as an optimal compound for the esterification of protein carboxyl groups (eq 2).<sup>8</sup>



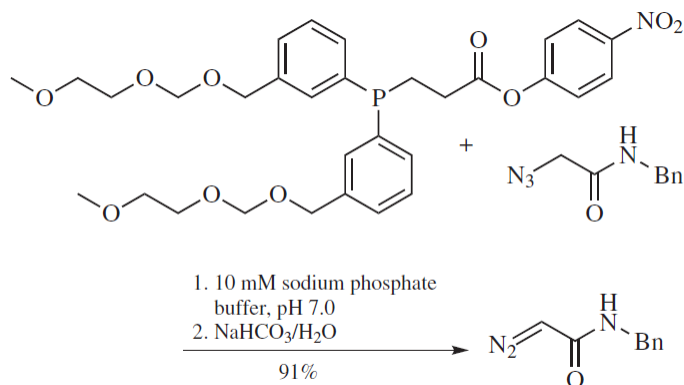
**Eq. 2**

Aronoff et al. investigated a means to establish chemoselective 1,3-dipolar cycloadditions of a diazo group generated from an azido group with **DPDS**. By tuning the electronics of the dipolarophile, cycloaddition with a diazo compound was obtained in the presence of the azide precursor. These workers also employed **DPDS**-mediated deimidogenation on an azidoacetamide–biotin conjugate to produce a diazoacetamide–biotin conjugate (eq 3) that undergoes cycloaddition with dehydroalanine residues on nisin, unlike the azidoacetamide–biotin precursor.<sup>223</sup>



**Eq. 3**

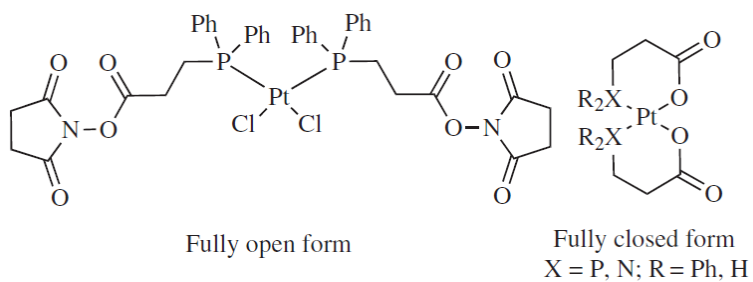
Chou and Raines developed a water-soluble analog of **DPPS** that is capable of performing the deimidogenation reaction in an aqueous environment and is tolerant of biological nucleophiles (eq 4). The chromogenic nature of this reagent also enables monitoring of reaction progress.<sup>205</sup>



**Eq. 4**

## Miscellaneous

Ravera et al. investigated the antiproliferative activity of cisplatin-like Pt(II)-phosphane complexes in tumor cell lines. **DPSS** was used as a ligand that would retain its “fully opened” form, incapable of undergoing an intramolecular rearrangement into the closed form, as did unprotected compounds (eq 5). Only the open form manifested antiproliferative activity.<sup>224</sup>



**Eq. 5**

## References

- (1) PhRMA. Medicines in Development for Biologics 2013 Report. *retrieved from <https://phrma.org/resource-center/Topics/Medicines-in-Development/Medicines-in-Development-for-Biologics-2013-Report> 2013.*
- (2) Lindsley, C. W. New 2016 Data and Statistics for Global Pharmaceutical Products and Projections through 2017. *ACS Chem. Neurosci.* **2017**, *8*, 1635–1636. <https://doi.org/10.1021/acschemneuro.7b00253>.
- (3) Horn, J. M.; Obermeyer, A. C. Genetic and Covalent Protein Modification Strategies to Facilitate Intracellular Delivery. *Biomacromolecules* **2021**, *22*, 4883–4904. <https://doi.org/10.1021/acs.biomac.1c00745>.
- (4) Li, Y.; Ye, Z.; Yang, H.; Xu, Q. Tailoring Combinatorial Lipid Nanoparticles for Intracellular Delivery of Nucleic Acids, Proteins, and Drugs. *Acta Pharm. Sin. B* **2022**, *12*, 2624–2639. <https://doi.org/https://doi.org/10.1016/j.apsb.2022.04.013>.
- (5) Zhang, X.; Zhao, J.; Wen, Y.; Zhu, C.; Yang, J.; Yao, F. Carboxymethyl Chitosan-Poly(Amidoamine) Dendrimer Core–Shell Nanoparticles for Intracellular Lysozyme Delivery. *Carbohydr. Polym.* **2013**, *98*, 1326–1334. <https://doi.org/https://doi.org/10.1016/j.carbpol.2013.08.005>.
- (6) Ghosh, P.; Yang, X.; Arvizo, R.; Zhu, Z.-J.; Agasti, S. S.; Mo, Z.; Rotello, V. M. Intracellular Delivery of a Membrane-Impermeable Enzyme in Active Form Using Functionalized Gold Nanoparticles. *J. Am. Chem. Soc.* **2010**, *132*, 2642–2645. <https://doi.org/10.1021/ja907887z>.
- (7) Andersen, K. A.; Smith, T. P.; Lomax, J. E.; Raines, R. T. Boronic Acid for the Traceless Delivery of Proteins into Cells. *ACS Chem. Biol.* **2016**, *11*, 319–323. <https://doi.org/10.1021/acschembio.5b00966>.
- (8) Mix, K. A.; Raines, R. T. Optimized Diazo Scaffold for Protein Esterification. *Org. Lett.* **2015**, *17*, 2358–2361. <https://doi.org/10.1021/acs.orglett.5b00840>.
- (9) Ressler, V. T.; Mix, K. A.; Raines, R. T. Esterification Delivers a Functional Enzyme into a Human Cell. *ACS Chem. Biol.* **2019**, *14*, 599–602. <https://doi.org/10.1021/acschembio.9b00033>.
- (10) Lawrence, M. S.; Phillips, K. J.; Liu, D. R. Supercharging Proteins Can Impart Unusual Resilience. *J. Am. Chem. Soc.* **2007**, *129*, 10110–10112. <https://doi.org/10.1021/ja071641y>.
- (11) McNaughton, B. R.; Cronican, J. J.; Thompson, D. B.; Liu, D. R. Mammalian Cell Penetration, siRNA Transfection, and DNA Transfection by Supercharged Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 6111–6116. <https://doi.org/10.1073/pnas.0807883106>.
- (12) Frankel, A. D.; Pabo, C. O. Cellular Uptake of the Tat Protein from Human Immunodeficiency Virus. *Cell* **1988**, *55*, 1189–1193. [https://doi.org/10.1016/0092-8674\(88\)90263-2](https://doi.org/10.1016/0092-8674(88)90263-2).
- (13) Vivès, E.; Brodin, P.; Lebleu, B. A Truncated HIV-1 Tat Protein Basic Domain Rapidly Translocates through the Plasma Membrane and Accumulates in the Cell Nucleus. *J. Biol. Chem.* **1997**, *272*, 16010–16017. <https://doi.org/10.1074/jbc.272.25.16010>.
- (14) Derossi, D.; Joliot, A. H.; Chassaing, G.; Prochiantz, A. The Third Helix of the Antennapedia Homeodomain Translocates through Biological Membranes. *J. Biol. Chem.* **1994**, *269*, 10444–10450. [https://doi.org/https://doi.org/10.1016/S0021-9258\(17\)34080-2](https://doi.org/https://doi.org/10.1016/S0021-9258(17)34080-2).
- (15) A., W. P.; J., M. D.; Kanaka, P.; T., P. E.; Lawrence, S.; B., R. J. The Design, Synthesis, and Evaluation of Molecules That Enable or Enhance Cellular Uptake: Peptoid Molecular Transporters. *Proc. Natl.*

- Acad. Sci. U.S.A.* **2000**, *97*, 13003–13008. <https://doi.org/10.1073/pnas.97.24.13003>.
- (16) Fuchs, S. M.; Raines, R. T. Pathway for Polyarginine Entry into Mammalian Cells. *Biochemistry* **2004**, *43*, 2438–2444. <https://doi.org/10.1021/bi035933x>.
- (17) Tyagi, M.; Rusnati, M.; Presta, M.; Giacca, M. Internalization of HIV-1 Tat Requires Cell Surface Heparan Sulfate Proteoglycans. *J. Biol. Chem.* **2001**, *276*, 3254–3261. <https://doi.org/10.1074/jbc.M006701200>.
- (18) Ziegler, A.; Seelig, J. Binding and Clustering of Glycosaminoglycans: A Common Property of Mono- and Multivalent Cell-Penetrating Compounds. *Biophys. J.* **2008**, *94*, 2142–2149. <https://doi.org/10.1529/biophysj.107.113472>.
- (19) Christoph, A.; Aniket, M.; Piotr, J.; Katarína, B.; Matti, J.; E., M. P.; Radek, Š.; Marek, C.; Martin, H.; Dominik, H.; Veronika, H.; Reinhard, R.; M., Z. C.; Adam, S.; Pavel, J. Arginine-Rich Cell-Penetrating Peptides Induce Membrane Multilamellarity and Subsequently Enter via Formation of a Fusion Pore. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, 11923–11928. <https://doi.org/10.1073/pnas.1811520115>.
- (20) Barba-Bon, A.; Pan, Y.-C.; Biedermann, F.; Guo, D.-S.; Nau, W. M.; Hennig, A. Fluorescence Monitoring of Peptide Transport Pathways into Large and Giant Vesicles by Supramolecular Host–Dye Reporter Pairs. *J. Am. Chem. Soc.* **2019**, *141*, 20137–20145. <https://doi.org/10.1021/jacs.9b09563>.
- (21) Brock, R. The Uptake of Arginine-Rich Cell-Penetrating Peptides: Putting the Puzzle Together. *Bioconjug. Chem.* **2014**, *25*, 863–868. <https://doi.org/10.1021/bc500017t>.
- (22) Herce, H. D.; Garcia, A. E.; Cardoso, M. C. Fundamental Molecular Mechanism for the Cellular Uptake of Guanidinium-Rich Molecules. *J. Am. Chem. Soc.* **2014**, *136*, 17459–17467. <https://doi.org/10.1021/ja507790z>.
- (23) Patel, S. G.; Sayers, E. J.; He, L.; Narayan, R.; Williams, T. L.; Mills, E. M.; Allemann, R. K.; Luk, L. Y. P.; Jones, A. T.; Tsai, Y.-H. Cell-Penetrating Peptide Sequence and Modification Dependent Uptake and Subcellular Distribution of Green Florescent Protein in Different Cell Lines. *Sci. Rep.* **2019**, *9*, 6298. <https://doi.org/10.1038/s41598-019-42456-8>.
- (24) Ziegler, A.; Seelig, J. Interaction of the Protein Transduction Domain of HIV-1 TAT with Heparan Sulfate: Binding Mechanism and Thermodynamic Parameters. *Biophys. J.* **2004**, *86*, 254–263. [https://doi.org/10.1016/S0006-3495\(04\)74101-6](https://doi.org/10.1016/S0006-3495(04)74101-6).
- (25) Ruzza, P.; Biondi, B.; Marchiani, A.; Antolini, N.; Calderan, A. Cell-Penetrating Peptides: A Comparative Study on Lipid Affinity and Cargo Delivery Properties. *Pharmaceuticals* **2010**, *3*, 1045–1062. <https://doi.org/10.3390/ph3041045>.
- (26) Houk, R. J. T.; Tobey, S. L.; Anslyn, E. V. Abiotic Guanidinium Receptors for Anion Molecular Recognition and Sensing. In *Anion Sensing. Topics in Current Chemistry, vol. 255*; Stibor, I., Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2005; pp 199–229. <https://doi.org/10.1007/b101167>.
- (27) Conley, M. P.; Valero, J.; de Mendoza, J. Guanidinium-Based Receptors for Oxoanions. In *Supramol. Chem.*; John Wiley & Sons, Ltd: Chichester, UK, 2012. <https://doi.org/10.1002/9780470661345.smc061>.
- (28) Berger, M.; Schmidtchen, F. P. Zwitterionic Guanidinium Compounds Serve as Electroneutral Anion Hosts. *J. Am. Chem. Soc.* **1999**, *121*, 9986–9993. <https://doi.org/10.1021/ja992028k>.
- (29) Fitch, C. A.; Platzer, G.; Okon, M.; Garcia-Moreno, B. E.; McIntosh, L. P. Arginine: Its PKa Value Revisited. *Protein Sci.* **2015**, *24*, 752–761. <https://doi.org/10.1002/pro.2647>.

- (30) Schmuck, C. Carboxylate Binding by 2-(Guanidiniocarbonyl)Pyrrole Receptors in Aqueous Solvents: Improving the Binding Properties of Guanidinium Cations through Additional Hydrogen Bonds. *Chem. Eur. J.* **2000**, *6*, 709–718. [https://doi.org/10.1002/\(SICI\)1521-3765\(20000218\)6:4<709::AID-CHEM709>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1521-3765(20000218)6:4<709::AID-CHEM709>3.0.CO;2-6).
- (31) Rozas, I.; Kruger, P. E. Theoretical Study of the Interaction between the Guanidinium Cation and Chloride and Sulfate Anions. *J. Chem. Theory Comput.* **2005**, *1*, 1055–1062. <https://doi.org/10.1021/ct050009x>.
- (32) Schmuck, C.; Schwegmann, M. A Molecular Flytrap for the Selective Binding of Citrate and Other Tricarboxylates in Water. *J. Am. Chem. Soc.* **2005**, *127*, 3373–3379. <https://doi.org/10.1021/ja0433469>.
- (33) Dempsey, C. E.; Mason, P. E.; Brady, J. W.; Neilson, G. W. The Reversal by Sulfate of the Denaturant Activity of Guanidinium. *J. Am. Chem. Soc.* **2007**, *129*, 15895–15902. <https://doi.org/10.1021/ja074719j>.
- (34) Echavarren, A.; Galan, A.; Lehn, J. M.; De Mendoza, J. Chiral Recognition of Aromatic Carboxylate Anions by an Optically Active Abiotic Receptor Containing a Rigid Guanidinium Binding Subunit. *J. Am. Chem. Soc.* **1989**, *111*, 4994–4995. <https://doi.org/10.1021/ja00195a071>.
- (35) Anslyn, E. V.; Tobey, S. L. Guanidinium-Based Anion Receptors. In *Encyclopedia of Supramolecular Chemistry*; CRC Press, 2004. <https://doi.org/10.1081/E-ESMC-120012718>.
- (36) Schug, K. A.; Lindner, W. Noncovalent Binding between Guanidinium and Anionic Groups: Focus on Biological- and Synthetic-Based Arginine/Guanidinium Interactions with Phosph[on]Ate and Sulf[on]Ate Residues. *Chem. Rev.* **2005**, *105*, 67–114. <https://doi.org/10.1021/cr040603j>.
- (37) Blondeau, P.; Segura, M.; Pérez-Fernández, R.; de Mendoza, J. Molecular Recognition of Oxoanions Based on Guanidinium Receptors. *Chem. Soc. Rev.* **2007**, *36* (2), 198–210. <https://doi.org/10.1039/B603089K>.
- (38) Jadhav, V. D.; Herdtweck, E.; Schmidtchen, F. P. Addressing Association Entropy by Reconstructing Guanidinium Anchor Groups for Anion Binding: Design, Synthesis, and Host–Guest Binding Studies in Polar and Protic Solutions. *Chem. Eur. J.* **2008**, *14*, 6098–6107. <https://doi.org/10.1002/chem.200702036>.
- (39) Seipp, C. A.; Williams, N. J.; Bryantsev, V. S.; Custelcean, R.; Moyer, B. A. A Conformationally Persistent Pseudo-Bicyclic Guanidinium for Anion Coordination as Stabilized by Dual Intramolecular Hydrogen Bonds. *RSC Adv.* **2015**, *5*, 107266–107269. <https://doi.org/10.1039/C5RA21864K>.
- (40) Seipp, C. A.; Williams, N. J.; Bryantsev, V. S.; Moyer, B. A. Simple Guanidinium Motif for the Selective Binding and Extraction of Sulfate. *Sep. Sci. Technol.* **2018**, *53*, 1864–1873. <https://doi.org/10.1080/01496395.2017.1318922>.
- (41) Rether, C.; Schmuck, C. Carboxylate Binding by Indole-Based Guanidinium Receptors: Acylguanidinium Cations Are Better than Aromatic Guanidinium Cations. *Eur. J. Org. Chem.* **2011**, *2011*, 1459–1466. <https://doi.org/10.1002/ejoc.201001465>.
- (42) Glasovac, Z.; Barešić, L.; Antol, I.; Margetić, D. Benzoylguanidines as Anion-Responsive Systems. *Chempluschem* **2018**, *83*, 845–854. <https://doi.org/10.1002/cplu.201800247>.
- (43) Pérez-Casas, C.; Yatsimirsky, A. K. Detailing Hydrogen Bonding and Deprotonation Equilibria between Anions and Urea/Thiourea Derivatives. *J. Org. Chem.* **2008**, *73*, 2275–2284. <https://doi.org/10.1021/jo702458f>.
- (44) Šekutor, M.; Mlinarić-Majerski, K. Adamantyl Aminoguanidines as Receptors for Oxo-Anions.

*Tetrahedron Lett.* **2014**, *55*, 6665–6670. <https://doi.org/10.1016/j.tetlet.2014.10.062>.

- (45) Koskinen, J. T.; Koskinen, M.; Mutikainen, I.; Mannfors, B.; Elo, H. Experimental and Computational Studies on Aminoguanidine Free Base, Monocation and Dication, Part I: The Crystal and Molecular Structure of Aminoguanidine Monohydrochloride and the Ab Initio Structure of the Endiamine Tautomer of Aminoguanidine Free Base. *Z. Naturforsch. B* **1996**, *51*, 1771–1778. <https://doi.org/doi:10.1515/znb-1996-1215>.
- (46) Koskinen, J. T.; Koskinen, M.; Mutikainen, I.; Tilus, P.; Mannfors, B.; Elo, H. Experimental and Computational Studies on Aminoguanidine Free Base, Monocation and Dication, Part II: Acid-Base Properties, Gas Phase Protonation Energies and Total Energies of Two Tautomers of the Free Base. *Z. Naturforsch. B* **1997**, *52*, 1259–1272. <https://doi.org/doi:10.1515/znb-1997-1018>.
- (47) Pushina, M.; Anzenbacher, P. Biguanides, Anion Receptors and Sensors. *Chem. Commun.* **2017**, *53*, 10074–10077. <https://doi.org/10.1039/C7CC05012G>.
- (48) Dixon, R. P.; Geib, S. J.; Hamilton, A. D. Molecular Recognition: Bis-Acylguanidiniums Provide a Simple Family of Receptors for Phosphodiester. *J. Am. Chem. Soc.* **1992**, *114*, 365–366. <https://doi.org/10.1021/ja00027a059>.
- (49) Gressel, M. C.; Merckel, D. A. S.; Hutchings, M. G. The Effect of Preorganisation on the Solid State Behaviour of Simple ‘Aromatic-Cored’ Bis(Guanidinium) Sulfates. *CrystEngComm* **2003**, *5*, 77–81. <https://doi.org/10.1039/B207803A>.
- (50) Kneeland, D. M.; Ariga, K.; Lynch, V. M.; Huang, C. Y.; Anslyn, E. V. Bis(Alkylguanidinium) Receptors for Phosphodiester: Effect of Counterions, Solvent Mixtures, and Cavity Flexibility on Complexation. *J. Am. Chem. Soc.* **1993**, *115*, 10042–10055. [https://doi.org/10.1021/JA00075A021/SUPPL\\_FILE/JA10042B.PDF](https://doi.org/10.1021/JA00075A021/SUPPL_FILE/JA10042B.PDF).
- (51) Schiessl, P.; Schmidtchen, F. P. Binding of Phosphates to Abiotic Hosts in Water. *J. Org. Chem.* **1994**, *59*, 509–511. <https://doi.org/10.1021/jo00082a001>.
- (52) Linton, B. R.; Goodman, M. S.; Fan, E.; van Arman, S. A.; Hamilton, A. D. Thermodynamic Aspects of Dicarboxylate Recognition by Simple Artificial Receptors. *J. Org. Chem.* **2001**, *66*, 7313–7319. <https://doi.org/10.1021/jo010413y>.
- (53) Dietrich, B.; Fyles, T. M.; Lehn, J.-M.; Pease, L. G.; Fyles, D. L. Anion Receptor Molecules. Synthesis and Some Anion Binding Properties of Macrocyclic Guanidinium Salts. *J. Chem. Soc., Chem. Commun.* **1978**, *21*, 934–936. <https://doi.org/10.1039/C39780000934>.
- (54) Metzger, A.; Lynch, V. M.; Anslyn, E. V. A Synthetic Receptor Selective for Citrate. *Angew. Chem., Int. Ed.* **1997**, *36*, 862–865. <https://doi.org/https://doi.org/10.1002/anie.199708621>.
- (55) Metzger, A.; Anslyn, E. V. A Chemosensor for Citrate in Beverages. *Angew. Chem., Int. Ed.* **1998**, *37*, 649–652. [https://doi.org/10.1002/\(SICI\)1521-3773\(19980316\)37:5<649::AID-ANIE649>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1521-3773(19980316)37:5<649::AID-ANIE649>3.0.CO;2-H).
- (56) Tobey, S. L.; Jones, B. D.; Anslyn, E. V. C<sub>3v</sub> Symmetric Receptors Show High Selectivity and High Affinity for Phosphate. *J. Am. Chem. Soc.* **2003**, *125*, 4026–4027. <https://doi.org/10.1021/ja021390n>.
- (57) Stanzl, E. G.; Trantow, B. M.; Vargas, J. R.; Wender, P. A. Fifteen Years of Cell-Penetrating, Guanidinium-Gich Molecular Transporters: Basic Science, Research Tools, and Clinical Applications. *Acc. Chem. Res.* **2013**, *46*, 2944–2954. <https://doi.org/10.1021/ar4000554>.
- (58) Mitchell, D. J.; Steinman, L.; Kim, D. T.; Fathman, C. G.; Rothbard, J. B. Polyarginine Enters Cells More Efficiently than Other Polycationic Homopolymers. *J. Pept. Res.* **2000**, *56*, 318–325. <https://doi.org/10.1034/j.1399-3011.2000.00723.x>.



- (59) Huang, W.; Seo, J.; Lin, J. S.; Barron, A. E. Peptoid Transporters: Effects of Cationic, Amphipathic Structure on Their Cellular Uptake. *Mol. Biosyst.* **2012**, *8*, 2626–2628. <https://doi.org/10.1039/c2mb25197c>.
- (60) Sarko, D.; Beijer, B.; Garcia Boy, R.; Nothelfer, E.-M.; Leotta, K.; Eisenhut, M.; Altmann, A.; Haberkorn, U.; Mier, W. The Pharmacokinetics of Cell-Penetrating Peptides. *Mol. Pharm.* **2010**, *7*, 2224–2231. <https://doi.org/10.1021/mp100223d>.
- (61) Li, Q.; Xu, M.; Cui, Y.; Huang, C.; Sun, M. Arginine-Rich Membrane-Permeable Peptides Are Seriously Toxic. *Pharmacol. Res. Perspect.* **2017**, *5*, e00334. <https://doi.org/10.1002/prp2.334>.
- (62) Lafarga, V.; Sirozh, O.; Díaz-López, I.; Galarreta, A.; Hisaoka, M.; Zarzuela, E.; Boskovic, J.; Jovanovic, B.; Fernandez-Leiro, R.; Muñoz, J.; Stoecklin, G.; Ventoso, I.; Fernandez-Capetillo, O. Widespread Displacement of DNA- and RNA-Binding Factors Underlies Toxicity of Arginine-Rich Cell-Penetrating Peptides. *EMBO J.* **2021**, *40*, e103311. <https://doi.org/10.15252/emj.2019103311>.
- (63) Maiti, K. K.; Jeon, O.-Y.; Lee, W. S.; Kim, D.-C.; Kim, K.-T.; Takeuchi, T.; Futaki, S.; Chung, S.-K. Design, Synthesis, and Membrane-Translocation Studies of Inositol-Based Transporters. *Angew. Chem., Int. Ed.* **2006**, *45*, 2907–2912. <https://doi.org/10.1002/anie.200600312>.
- (64) Luedtke, N. W.; Carmichael, P.; Tor, Y. Cellular Uptake of Aminoglycosides, Guanidinoglycosides, and Poly-Arginine. *J. Am. Chem. Soc.* **2003**, *125*, 12374–12375. <https://doi.org/10.1021/ja0360135>.
- (65) Dix, A. V; Fischer, L.; Sarrazin, S.; Redgate, C. P. H.; Esko, J. D.; Tor, Y. Cooperative, Heparan Sulfate-Dependent Cellular Uptake of Dimeric Guanidinoglycosides. *Chembiochem* **2010**, *11*, 2302–2310. <https://doi.org/10.1002/cbic.201000399>.
- (66) Inoue, M.; Tong, W.; Esko, J. D.; Tor, Y. Aggregation-Mediated Macromolecular Uptake by a Molecular Transporter. *ACS Chem. Biol.* **2013**, *8*, 1383–1388. <https://doi.org/10.1021/cb400172h>.
- (67) Hamill, K. M.; McCoy, L. S.; Wexselblatt, E.; Esko, J. D.; Tor, Y. Polymyxins Facilitate Entry into Mammalian Cells. *Chem. Sci.* **2016**, *7*, 5059–5068. <https://doi.org/10.1039/C6SC00488A>.
- (68) Bonduelle, C. V; Gillies, E. R. Dendritic Guanidines as Efficient Analogues of Cell Penetrating Peptides. *Pharmaceuticals* **2010**, *3*, 636–666. <https://doi.org/10.3390/ph3030636>.
- (69) Chung, H.-H.; Harms, G.; Min Seong, C.; Choi, B. H.; Min, C.; Taulane, J. P.; Goodman, M. Dendritic Oligoguanidines as Intracellular Translocators. *Pept. Sci.* **2004**, *76*, 83–96. <https://doi.org/https://doi.org/10.1002/bip.10597>.
- (70) Grillaud, M.; Russier, J.; Bianco, A. Polycationic Adamantane-Based Dendrons of Different Generations Display High Cellular Uptake without Triggering Cytotoxicity. *J. Am. Chem. Soc.* **2014**, *136*, 810–819. <https://doi.org/10.1021/ja411987g>.
- (71) Russier, J.; Grillaud, M.; Bianco, A. Elucidation of the Cellular Uptake Mechanisms of Polycationic HYDRAMers. *Bioconjugate Chem.* **2015**, *26*, 1484–1493. <https://doi.org/10.1021/acs.bioconjchem.5b00270>.
- (72) Guo, P.; Gu, W.; Chen, Q.; Lu, H.; Han, X.; Li, W.; Gao, H. Dual Functionalized Amino Poly(Glycerol Methacrylate) with Guanidine and Schiff-Base Linked Imidazole for Enhanced Gene Transfection and Minimized Cytotoxicity. *J. Mater. Chem. B* **2015**, *3*, 6911–6918. <https://doi.org/10.1039/C5TB01291K>.
- (73) Miyazaki, T.; Uchida, S.; Hatano, H.; Miyahara, Y.; Matsumoto, A.; Cabral, H. Guanidine-Phosphate Interactions Stabilize Polyion Complex Micelles Based on Flexible Cationomers to Improve mRNA Delivery. *Eur. Polym. J.* **2020**, *140*, 110028.

<https://doi.org/https://doi.org/10.1016/j.eurpolymj.2020.110028>.

- (74) Pavlovic, I.; Thakor, D. T.; Vargas, J. R.; McKinlay, C. J.; Hauke, S.; Anstaett, P.; Camuña, R. C.; Bigler, L.; Gasser, G.; Schultz, C.; Wender, P. A.; Jessen, H. J. Cellular Delivery and Photochemical Release of a Caged Inositol-Pyrophosphate Induces PH-Domain Translocation in Cellulo. *Nat. Commun.* **2016**, *7*, 10622. <https://doi.org/10.1038/ncomms10622>.
- (75) Kim, Y.; Binauld, S.; Stenzel, M. H. Zwitterionic Guanidine-Based Oligomers Mimicking Cell-Penetrating Peptides as a Nontoxic Alternative to Cationic Polymers to Enhance the Cellular Uptake of Micelles. *Biomacromolecules* **2012**, *13*, 3418–3426. <https://doi.org/10.1021/bm301351e>.
- (76) Fernández-Carneado, J.; Van Gool, M.; Martos, V.; Castel, S.; Prados, P.; de Mendoza, J.; Giralt, E. Highly Efficient, Nonpeptidic Oligoguanidinium Vectors That Selectively Internalize into Mitochondria. *J. Am. Chem. Soc.* **2005**, *127*, 869–874. <https://doi.org/10.1021/ja044006q>.
- (77) Mosquera, J.; Sánchez, M. I.; Valero, J.; Mendoza, J. de; Vázquez, M. E.; Mascareñas, J. L. Sequence-Selective DNA Binding with Cell-Permeable Oligoguanidinium–Peptide Conjugates. *Chem. Commun.* **2015**, *51*, 4811–4814. <https://doi.org/10.1039/C4CC09525A>.
- (78) Okuyama, M.; Laman, H.; Kingsbury, S. R.; Visintin, C.; Leo, E.; Eward, K. L.; Stoeber, K.; Boshoff, C.; Williams, G. H.; Selwood, D. L. Small-Molecule Mimics of an Alpha-Helix for Efficient Transport of Proteins into Cells. *Nat. Methods* **2007**, *4*, 153–159. <https://doi.org/10.1038/nmeth997>.
- (79) Li, M.; Mosel, S.; Knauer, S. K.; Schmuck, C. A Dipeptide with Enhanced Anion Binding Affinity Enables Cell Uptake and Protein Delivery. *Org. Biomol. Chem.* **2018**, *16*, 2312–2317. <https://doi.org/10.1039/C7OB02721D>.
- (80) Tripathi, P. P.; Arami, H.; Banga, I.; Gupta, J.; Gandhi, S. Cell Penetrating Peptides in Preclinical and Clinical Cancer Diagnosis and Therapy. *Oncotarget* **2018**, *9*, 37252–37267. <https://doi.org/10.18632/oncotarget.26442>.
- (81) Matijass, M.; Neundorff, I. Cell-Penetrating Peptides as Part of Therapeutics Used in Cancer Research. *Med. Drug Discov.* **2021**, *10*, 100092. <https://doi.org/https://doi.org/10.1016/j.medidd.2021.100092>.
- (82) Xie, J.; Bi, Y.; Zhang, H.; Dong, S.; Teng, L.; Lee, R. J.; Yang, Z. Cell-Penetrating Peptides in Diagnosis and Treatment of Human Diseases: From Preclinical Research to Clinical Application. *Front. Pharmacol.* **2020**, *11*, 697. <https://doi.org/10.3389/fphar.2020.00697>.
- (83) Deshaies, R. J. Multispecific Drugs Herald a New Era of Biopharmaceutical Innovation. *Nature* **2020**, *580* (7803), 329–338. <https://doi.org/10.1038/s41586-020-2168-1>.
- (84) Lindsley, C. W. Predictions and Statistics for the Best-Selling Drugs Globally and in the United States in 2018 and a Look Forward to 2024 Projections. *ACS Chem. Neurosci.* **2019**, *10* (3), 1115. <https://doi.org/10.1021/acchemneuro.9b00112>.
- (85) Schwarze, S. R.; Hruska, K. A.; Dowdy, S. F. Protein Transduction: Unrestricted Delivery into All Cells? *Trends Cell. Bio.* **2000**, *10*, 290–295. [https://doi.org/10.1016/s0962-8924\(00\)01771-2](https://doi.org/10.1016/s0962-8924(00)01771-2).
- (86) Fuchs, S. M.; Raines, R. T. Internalization of Cationic Peptides: The Road Less (or More?) Traveled. *Cell. Mol. Life Sci.* **2006**, *63*, 1819–1822. <https://doi.org/10.1007/s00018-006-6170-z>.
- (87) Copolovici, D. M.; Langel, K.; Eriste, E.; Langel, Ü. Cell-Penetrating Peptides: Design, Synthesis, and Applications. *ACS Nano* **2014**, *8*, 1972–1994. <https://doi.org/10.1021/nn4057269>.
- (88) Dupont, E.; Prochiantz, A.; Joliot, A. Penetratin Story: An Overview. *Methods Mol. Biol.* **2015**, *1324*, 29–37. [https://doi.org/10.1007/978-1-4939-2806-4\\_2](https://doi.org/10.1007/978-1-4939-2806-4_2).

- (89) Zhu, P.; Jin, L. Cell Penetrating Peptides: A Promising Tool for the Cellular Uptake of Macromolecular Drugs. *Curr. Protein Pept. Sci.* **2018**, *19*, 211–220. <https://doi.org/10.2174/1389203718666170710115240>.
- (90) Derakhshankhah, H.; Jafari, S. Cell Penetrating Peptides: A Concise Review with Emphasis on Biomedical Applications. *Biomed. Pharmacother.* **2018**, *108*, 1090–1096. <https://doi.org/10.1016/j.biopha.2018.09.097>.
- (91) Brock, R. The Uptake of Arginine-Rich Cell-Penetrating Peptides: Putting the Puzzle Together. *Bioconjugate Chem.* **2014**, *25* (5), 863–868. <https://doi.org/10.1021/bc500017t>.
- (92) Gooding, M.; Adigbli, D.; Edith Chan, A. W.; Melander, R. J.; MacRobert, A. J.; Selwood, D. L. A Bifurcated Proteoglycan Binding Small Molecule Carrier for siRNA Delivery. *Chem. Biol. Drug Des.* **2014**, *84* (1), 24–35. <https://doi.org/https://doi.org/10.1111/cbdd.12295>.
- (93) Okuyama, M.; Laman, H.; Kingsbury, S. R.; Visintin, C.; Leo, E.; Eward, K. L.; Stoeber, K.; Boshoff, C.; Williams, G. H.; Selwood, D. L. Small-Molecule Mimics of an  $\alpha$ -Helix for Efficient Transport of Proteins into Cells. *Nat. Methods* **2007**, *4* (2), 153–159. <https://doi.org/10.1038/nmeth997>.
- (94) Schmuck, C. Side Chain Selective Binding of N-Acetyl- $\alpha$ -Amino Acid Carboxylates by a 2-(Guanidiniocarbonyl)Pyrrole Receptor in Aqueous Solvents. *Chem. Commun.* **1999**, No. 9, 843–844. <https://doi.org/10.1039/A901126I>.
- (95) Moiani, D.; Cavallotti, C.; Famulari, A.; Schmuck, C. Oxoanion Binding by Guanidiniocarbonylpyrrole Cations in Water: A Combined DFT and MD Investigation. *Chem. – A Eur. J.* **2008**, *14* (17), 5207–5219. <https://doi.org/10.1002/CHEM.200701745>.
- (96) Kelly, B.; O'Donovan, D. H.; O'Brien, J.; McCabe, T.; Blanco, F.; Rozas, I. Pyridin-2-yl Guanidine Derivatives: Conformational Control Induced by Intramolecular Hydrogen-Bonding Interactions. *J. Org. Chem.* **2011**, *76*, 9216–9227. <https://doi.org/10.1021/jo200954c>.
- (97) Dardonville, C.; Caine, B. A.; de la Fuente, M.; Martín Herranz, G.; Corrales Mariblanca, B.; Popelier, P. L. A. Substituent Effects on the Basicity (PKa) of Aryl Guanidines and 2-(Arylimino)imidazolidines: Correlations of PH-Metric and UV-Metric Values with Predictions from Gas-Phase Ab Initio Bond Lengths. *New J. Chem.* **2017**, *41* (19), 11016–11028. <https://doi.org/10.1039/C7NJ02497E>.
- (98) Stachel, S. J.; Stockwell, S. A.; Van Vranken, D. L. The Fluorescence of Scorpions and Cataractogenesis. *Chem. Biol.* **1999**, *6* (8), 531–539. [https://doi.org/10.1016/S1074-5521\(99\)80085-4](https://doi.org/10.1016/S1074-5521(99)80085-4).
- (99) Szepesi Kovács, D.; Hajdu, I.; Mészáros, G.; Wittner, L.; Meszéna, D.; Tóth, E. Z.; Hegedűs, Z.; Randelović, I.; Tóvári, J.; Szabó, T.; Szilágyi, B.; Milen, M.; Keserű, G. M.; Ábrányi-Balogh, P. Synthesis and Characterization of New Fluorescent Boro- $\beta$ -Carboline Dyes. *RSC Adv.* **2021**, *11* (21), 12802–12807. <https://doi.org/10.1039/D1RA02132J>.
- (100) Huang, R.; Zhang, X.-D.; Wu, X.; Liu, J.-K.  $\beta$ -Carboline-Based PH Fluorescent Probe and Its Application for Monitoring Enzymatic Ester Hydrolysis. *Chem. Biodivers.* **2021**, *18*, e2000829–e2000829. <https://doi.org/https://doi.org/10.1002/cbdv.202000829>.
- (101) Ji, X.; Zhang, D.; Li, L.; Jin, L.; Wu, R. Efficient  $\beta$ -Carboline Alkaloid-Based Probe for Highly Sensitive Imaging of Endogenous Glutathione in Wheat Germ Tissues. *Int. J. Anal. Chem.* **2020**, *2020*, 8675784. <https://doi.org/10.1155/2020/8675784>.
- (102) Ledwon, P. Recent Advances of Donor-Acceptor Type Carbazole-Based Molecules for Light Emitting Applications. *Org. Electron.* Elsevier B.V. December 1, 2019, p 105422. <https://doi.org/10.1016/j.orgel.2019.105422>.

- (103) Wang, B.; Lv, X.; Tan, J.; Zhang, Q.; Huang, Z.; Yi, W.; Wang, L. Bipolar Phenanthroimidazole–Diazacarbazole Hybrids with Appropriate Bandgaps for Highly Efficient and Low Roll-off Red, Green and Blue Electroluminescent Devices. *J. Mater. Chem. C* **2016**, *4*, 8473–8482. <https://doi.org/10.1039/C6TC02683D>.
- (104) Cha, J. R.; Lee, C. W.; Gong, M. S. Bipolar Host Material for Phosphorescent OLEDs Based on 2,7-Diazacarbazole as a New Electron-Transporting Unit. *Bull. Korean Chem. Soc.* **2017**, *38*, 1016–1022. <https://doi.org/10.1002/bkcs.11209>.
- (105) Chichibabin, A. E.; Zeide, O. A. New Reaction for Compounds Containing the Pyridine Nucleus. *J. Russ. Phys. Chem. Soc.* **1914**, *46*, 1216–1236.
- (106) McGill, C. K.; Rappa, A. Advances in the Chichibabin Reaction. In *Advances in Heterocyclic Chemistry*; Katritzky, A. R. B. T.-A. in H. C., Ed.; Academic Press, 1988; Vol. 44, pp 1–79. [https://doi.org/https://doi.org/10.1016/S0065-2725\(08\)60261-5](https://doi.org/https://doi.org/10.1016/S0065-2725(08)60261-5).
- (107) Lindel, T.; Hoffmann, H. Synthesis of Dispacamide from the Marine Sponge *Agelas Dispar*. *Tetrahedron Lett.* **1997**, *38*, 8935–8938. [https://doi.org/10.1016/S0040-4039\(97\)10387-2](https://doi.org/10.1016/S0040-4039(97)10387-2).
- (108) Smith, A. E.; Clapham, K. M.; Batsanov, A. S.; Bryce, M. R.; Tarbit, B. (Dimethoxy- and Dihalopyridyl)Boronic Acids and Highly Functionalized Heteroarylpyridines by Suzuki Cross-Coupling Reactions. *Eur. J. Org. Chem.* **2008**, *8*, 1458–1463. <https://doi.org/10.1002/ejoc.200701156>.
- (109) Dias, A.; Varela, A. P.; Miguel, M. da G.; Macanita, A. L.; Becker, R. S.  $\beta$ -Carboline Photosensitizers. 1. Photophysics, Kinetics and Excited-State Equilibria in Organic Solvents, and Theoretical Calculations. *J. Phys. Chem.* **1992**, *96* (25), 10290–10296. <https://doi.org/10.1021/j100204a036>.
- (110) Albert, A.; Goldacre, R.; Phillips, J. The Strength of Heterocyclic Bases. *J. Chem. Soc.* **1948**, No. 0, 2240–2249. <https://doi.org/10.1039/JR9480002240>.
- (111) <http://supramolecular.org>.
- (112) Thordarson, P. Determining Association Constants from Titration Experiments in Supramolecular Chemistry. *Chem. Soc. Rev.* **2011**, *40* (3), 1305–1323. <https://doi.org/10.1039/C0CS00062K>.
- (113) Brynn Hibbert, D.; Thordarson, P. The Death of the Job Plot, Transparency, Open Science and Online Tools, Uncertainty Estimation Methods and Other Developments in Supramolecular Chemistry Data Analysis. *Chem. Commun.* **2016**, *52* (87), 12792–12805. <https://doi.org/10.1039/C6CC03888C>.
- (114) Tapia, M. J.; Reyman, D.; Viñas, M. H.; Carcedo, C.; Camacho, J. J. Hydrogen-Bonding Interactions of Norharmane in Mixtures of Acetic Acid with Benzene, p-Dioxane and Acetonitrile. *Phys. Chem. Chem. Phys.* **2002**, *4* (15), 3676–3683. <https://doi.org/10.1039/B201526A>.
- (115) *Molinspiration Cheminformatics*, <https://www.molinspiration.com>, *Slovensky Grob, Slovakia*.
- (116) Ahmed Al-Hadhrami, N.; Ladwig, A.; Rahman, A.; Rozas, I.; Paul G. Malthouse, J.; Evans, P. Synthesis of 2-Guanidinyll Pyridines and Their Trypsin Inhibition and Docking. *Bioorg. Med. Chem* **2020**, *28* (16), 115612. <https://doi.org/https://doi.org/10.1016/j.bmc.2020.115612>.
- (117) Ingallina, C.; D’Acquarica, I.; Delle Monache, G.; Ghirga, F.; Quaglio, D.; Ghirga, P.; Berardozi, S.; Markovic, V.; Botta, B. The Pictet-Spengler Reaction Still on Stage. *Curr. Pharm. Des.* **2016**, *22*, 1808–1850. <https://doi.org/10.2174/1381612822666151231100247>.
- (118) Freeman, A. W.; Urvoy, M.; Criswell, M. E. Triphenylphosphine-Mediated Reductive Cyclization of 2-Nitrobiphenyls: A Practical and Convenient Synthesis of Carbazoles. *J. Org. Chem.* **2005**, *70*, 5014–5019. <https://doi.org/10.1021/jo0503299>.

- (119) CN109422743 (A), WUHAN SHANGSAI OPTOELECTRONICS TECH CO LTD, 2019.
- (120) Gribble, G. W. *Indole Ring Synthesis: From Natural Products to Drug Discovery*; 2016.
- (121) Nelson, T. D.; Crouch, R. D. Cu, Ni, and Pd Mediated Homocoupling Reactions in Biaryl Syntheses: The Ullmann Reaction. *Organic Reactions*. October 15, 2004, pp 265–555. <https://doi.org/https://doi.org/10.1002/0471264180.or063.03>.
- (122) Abboud, M.; Aubert, E.; Mamane, V. Double N-Arylation Reaction of Polyhalogenated 4,4'-Bipyridines. Expedient Synthesis of Functionalized 2,7-Diazacarbazoles. *Beilstein J. Org. Chem.* **2012**, *8*, 253–258.
- (123) Awad, H.; Mongin, F.; Trécourt, F.; Quéguiner, G.; Marsais, F. Deprotonation of Chloropyridines Using Lithium Magnesates. *Tetrahedron Lett.* **2004**, *45*, 7873–7877. <https://doi.org/10.1016/j.tetlet.2004.08.151>.
- (124) Bradford, C. N. *The Development of Cell Penetrating Small Molecules, Polymers, and Proteins*, U.W.-Madison, 2014.
- (125) Krężel, A.; Bal, W. A Formula for Correlating PKa Values Determined in D2O and H2O. *J. Inorg. Biochem.* **2004**, *98* (1), 161–166. <https://doi.org/https://doi.org/10.1016/j.jinorgbio.2003.10.001>.
- (126) Suzuki, K.; Kobayashi, A.; Kaneko, S.; Takehira, K.; Yoshihara, T.; Ishida, H.; Shiina, Y.; Oishi, S.; Tobita, S. Reevaluation of Absolute Luminescence Quantum Yields of Standard Solutions Using a Spectrometer with an Integrating Sphere and a Back-Thinned CCD Detector. *Phys. Chem. Chem. Phys.* **2009**, *11* (42), 9850–9860. <https://doi.org/10.1039/B912178A>.
- (127) Cheresh, D. A.; Spiro, R. C. Biosynthetic and Functional Properties of an Arg-Gly-Asp-Directed Receptor Involved in Human Melanoma Cell Attachment to Vitronectin, Fibrinogen, and von Willebrand Factor. *J. Biol. Chem.* **1987**, *262* (36), 17703–17711. [https://doi.org/10.1016/S0021-9258\(18\)45436-1](https://doi.org/10.1016/S0021-9258(18)45436-1).
- (128) Felding-Habermann, B.; Mueller, B. M.; Romerdahl, C. A.; Cheresh, D. A. Involvement of Integrin Alpha V Gene Expression in Human Melanoma Tumorigenicity. *J. Clin. Investig.* **1992**, *89* (6), 2018–2022. <https://doi.org/10.1172/JCI115811>.
- (129) Heitz, F.; Morris, M. C.; Divita, G. Twenty Years of Cell-Penetrating Peptides: From Molecular Mechanisms to Therapeutics. *Br. J. Pharmacol.* **2009**, *157*, 195–206. <https://doi.org/10.1111/J.1476-5381.2009.00057.X>.
- (130) Takeuchi, T.; Futaki, S. Current Understanding of Direct Translocation of Arginine-Rich Cell-Penetrating Peptides and Its Internalization Mechanisms. *Chem. Pharm. Bull. (Tokyo)* **2016**, *64*, 1431–1437. <https://doi.org/10.1248/cpb.c16-00505>.
- (131) Rothbard, J. B.; Jessop, T. C.; Lewis, R. S.; Murray, B. A.; Wender, P. A. Role of Membrane Potential and Hydrogen Bonding in the Mechanism of Translocation of Guanidinium-Rich Peptides into Cells. *J. Am. Chem. Soc.* **2004**, *126* (31), 9506–9507. <https://doi.org/10.1021/ja0482536>.
- (132) Umezawa, N.; Gelman, M. A.; Haigis, M. C.; Raines, R. T.; Gellman, S. H. Translocation of a  $\beta$ -Peptide across Cell Membranes. *J. Am. Chem. Soc.* **2002**, *124*, 368–369. <https://doi.org/10.1021/ja017283v>.
- (133) Nischan, N.; Herce, H. D.; Natale, F.; Bohlke, N.; Budisa, N.; Cardoso, M. C.; Hackenberger, C. P. R. Covalent Attachment of Cyclic TAT Peptides to GFP Results in Protein Delivery into Live Cells with Immediate Bioavailability. *Angew. Chem., Int. Ed.* **2015**, *54*, 1950–1953. <https://doi.org/10.1002/anie.201410006>.
- (134) Nagel, Y. A.; Raschle, P. S.; Wennemers, H. Effect of Preorganized Charge-Display on the Cell-

- Penetrating Properties of Cationic Peptides. *Angew. Chem., Int. Ed.* **2017**, *56*, 122–126. <https://doi.org/10.1002/anie.201607649>.
- (135) Kalafatovic, D.; Giralt, E. Cell-Penetrating Peptides: Design Strategies beyond Primary Structure and Amphipathicity. *Molecules* **2017**, *22*, 1929. <https://doi.org/10.3390/MOLECULES22111929>.
- (136) Pines, M.; Rosenthal, G. A.; Applebaum, S. W. In Vitro Incorporation of L-Canavanine into Vitellogenin of the Fat Body of the Migratory Locust *Locusta Migratoria Migratorioides*. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 5480–5483. <https://doi.org/10.1073/pnas.78.9.5480>.
- (137) Rosenthal, G. A.; Reichhart, J. M.; Hoffmann, J. A. L-Canavanine Incorporation into Vitellogenin and Macromolecular Conformation. *J. Biol. Chem.* **1989**, *264*, 13693–13696. [https://doi.org/https://doi.org/10.1016/S0021-9258\(18\)80053-9](https://doi.org/https://doi.org/10.1016/S0021-9258(18)80053-9).
- (138) Rosenthal, G. A.; Lambert, J.; Hoffmann, D. Canavanine Incorporation into the Antibacterial Proteins of the Fly, *Phormia Terranova* (Diptera), and Its Effect on Biological Activity. *J. Biol. Chem.* **1989**, *264*, 9768–9771.
- (139) Rosenthal, G. A.; Dahlman, D. L. Studies of L-Canavanine Incorporation into Insectan Lysozyme. *J. Biol. Chem.* **1991**, *266*, 15684–15687. [https://doi.org/https://doi.org/10.1016/S0021-9258\(18\)98462-0](https://doi.org/https://doi.org/10.1016/S0021-9258(18)98462-0).
- (140) Boyar, A.; Marsh, R. E. L-Canavanine, a Paradigm for the Structures of Substituted Guanidines. *J. Am. Chem. Soc.* **1982**, *104* (7), 1995–1998. <https://doi.org/10.1021/ja00371a033>.
- (141) Tomiyama, T. The Apparent Dissociation Constants of Canavanine and Canaline. *J. Biol. Chem.* **1935**, *111*, 45–49. [https://doi.org/https://doi.org/10.1016/S0021-9258\(18\)75063-1](https://doi.org/https://doi.org/10.1016/S0021-9258(18)75063-1).
- (142) Radzicka, A.; Wolfenden, R. Comparing the Polarities of the Amino Acids: Side-Chain Distribution Coefficients between the Vapor Phase, Cyclohexane, 1-Octanol, and Neutral Aqueous Solution. *Biochemistry* **1988**, *27*, 1664–1670. <https://doi.org/10.1021/bi00405a042>.
- (143) Gordy, W.; Stanford, S. C. Spectroscopic Evidence for Hydrogen Bonds: Comparison of Proton-attracting Properties of Liquids. II. *J. Chem. Phys.* **1940**, *8*, 170–177. <https://doi.org/10.1063/1.1750625>.
- (144) Hine, J. S. *Physical Organic Chemistry*; McGraw–Hill: New York, 1956.
- (145) Shan, S. O.; Loh, S.; Herschlag, D. The Energetics of Hydrogen Bonds in Model Systems: Implications for Enzymatic Catalysis. *Science (80-. )*. **1996**, *272*, 97–101. <https://doi.org/10.1126/SCIENCE.272.5258.97>.
- (146) Hughes, E.; Burke, R. M.; Doig, A. J. Inhibition of Toxicity in the Beta-Amyloid Peptide Fragment Beta -(25-35) Using N-Methylated Derivatives: A General Strategy to Prevent Amyloid Formation. *J. Biol. Chem.* **2000**, *275*, 25109–25115. <https://doi.org/10.1074/jbc.M003554200>.
- (147) Gordon, D. J.; Sciarretta, K. L.; Meredith, S. C. Inhibition of Beta-Amyloid(40) Fibrillogenesis and Disassembly of  $\beta$ -Amyloid(40) Fibrils by Short  $\beta$ -Amyloid Congeners Containing N-Methyl Amino Acids at Alternate Residues. *Biochemistry* **2001**, *40* (28), 8237–8245. <https://doi.org/10.1021/bi002416v>.
- (148) Gordon, D. J.; Tappe, R.; Meredith, S. C. Design and Characterization of a Membrane Permeable N-Methyl Amino Acid-Containing Peptide That Inhibits Abeta1-40 Fibrillogenesis. *J. Pept. Res.* **2002**, *60*, 37–55. <https://doi.org/10.1034/j.1399-3011.2002.11002.x>.
- (149) Cruz, M.; Tusell, J. M.; Grillo-Bosch, D.; Albericio, F.; Serratos, J.; Rabanal, F.; Giralt, E. Inhibition of Beta-Amyloid Toxicity by Short Peptides Containing N-Methyl Amino Acids. *J. Pept. Res.* **2004**, *63*, 324–328. <https://doi.org/10.1111/j.1399-3011.2004.00156.x>.

- (150) Kokkoni, N.; Stott, K.; Amijee, H.; Mason, J. M.; Doig, A. J. N-Methylated Peptide Inhibitors of Beta-Amyloid Aggregation and Toxicity. Optimization of the Inhibitor Structure. *Biochemistry* **2006**, *45*, 9906–9918. <https://doi.org/10.1021/bi060837s>.
- (151) Pratim Bose, P.; Chatterjee, U.; Nerelius, C.; Govender, T.; Norström, T.; Gogoll, A.; Sandegren, A.; Göthelid, E.; Johansson, J.; Arvidsson, P. I. Poly-N-Methylated Amyloid  $\beta$ -Peptide (A $\beta$ ) C-Terminal Fragments Reduce A $\beta$  Toxicity in Vitro and in *Drosophila Melanogaster*. *J. Med. Chem.* **2009**, *52*, 8002–8009. <https://doi.org/10.1021/jm901092h>.
- (152) Simon, R. J.; Kania, R. S.; Zuckermann, R. N.; Huebner, V. D.; Jewell, D. A.; Banville, S.; Ng, S.; Wang, L.; Rosenberg, S.; Marlowe, C. K. Peptoids: A Modular Approach to Drug Discovery. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9367–9371. <https://doi.org/10.1073/pnas.89.20.9367>.
- (153) Rennert, R.; Wespe, C.; Beck-Sickinger, A. G.; Neundorff, I. Developing Novel HCT Derived Cell-Penetrating Peptides with Improved Metabolic Stability. *Biochim. Biophys. Acta.* **2006**, *1758*, 347–354. <https://doi.org/10.1016/j.bbame.2005.10.006>.
- (154) Chatterjee, J.; Gilon, C.; Hoffman, A.; Kessler, H. N-Methylation of Peptides: A New Perspective in Medicinal Chemistry. *Acc. Chem. Res.* **2008**, *41*, 1331–1342. <https://doi.org/10.1021/ar8000603>.
- (155) Hewitt, W. M.; Leung, S. S. F.; Pye, C. R.; Ponkey, A. R.; Bednarek, M.; Jacobson, M. P.; Lokey, R. S. Cell-Permeable Cyclic Peptides from Synthetic Libraries Inspired by Natural Products. *J. Am. Chem. Soc.* **2015**, *137*, 715–721. <https://doi.org/10.1021/ja508766b>.
- (156) Li, Y.; Li, W.; Xu, Z. Improvement on Permeability of Cyclic Peptide/Peptidomimetic: Backbone N-Methylation as a Useful Tool. *Mar. Drugs* **2021**, *19*, 311. <https://doi.org/10.3390/md19060311>.
- (157) Dougherty, P. G.; Sahni, A.; Pei, D. Understanding Cell Penetration of Cyclic Peptides. *Chem. Rev.* **2019**, *119*, 10241–10287. <https://doi.org/10.1021/acs.chemrev.9b00008>.
- (158) Tan, N. C.; Yu, P.; Kwon, Y.-U.; Kodadek, T. High-Throughput Evaluation of Relative Cell Permeability between Peptoids and Peptides. *Bioorg. Med. Chem.* **2008**, *16*, 5853–5861. <https://doi.org/10.1016/j.bmc.2008.04.074>.
- (159) Grogg, M.; Hilvert, D.; Beck, A.; Seebach, D. Syntheses of Cyanophycin Segments for Investigations of Cell-Penetration. *Synthesis (Stuttg.)*. **2019**, *51*, 31–39. <https://doi.org/10.1055/s-0037-1610202>.
- (160) White, K. N.; Konopelski, J. P. Facile Synthesis of Highly Functionalized N-Methyl Amino Acid Esters without Side-Chain Protection. *Org. Lett.* **2005**, *7*, 4111–4112. <https://doi.org/10.1021/ol051441w>.
- (161) Carpino, L. A.; Shroff, H.; Triolo, S. A.; Mansour, E.-S. M. E.; Wenschuh, H.; Albericio, F. The 2,2,4,6,7-Pentamethyldihydrobenzofuran-5-Sulfonyl Group (Pbf) as Arginine Side Chain Protectant. *Tetrahedron Lett.* **1993**, *34*, 7829–7832. [https://doi.org/https://doi.org/10.1016/S0040-4039\(00\)61487-9](https://doi.org/https://doi.org/10.1016/S0040-4039(00)61487-9).
- (162) Jin, E.; Zhang, B.; Sun, X.; Zhou, Z.; Ma, X.; Sun, Q.; Tang, J.; Shen, Y.; Van Kirk, E.; Murdoch, W. J.; Radosz, M. Acid-Active Cell-Penetrating Peptides for in Vivo Tumor-Targeted Drug Delivery. *J. Am. Chem. Soc.* **2013**, *135*, 933–940. <https://doi.org/10.1021/ja311180x>.
- (163) V. Komarov, I.; Yu. Ishchenko, A.; Hovtvianitsa, A.; Stepanenko, V.; Kharchenko, S.; D. Bond, A.; J. Kirby, A. Fast Amide Bond Cleavage Assisted by a Secondary Amino and a Carboxyl Group—a Model for yet Unknown Peptidases? *Molecules*. 2019, p 572. <https://doi.org/10.3390/molecules24030572>.
- (164) Calabretta, L. O.; Thomas, V. M.; Raines, R. T. Canavanine versus Arginine: Prospects for Cell-Penetrating Peptides. *Tetrahedron Lett.* **2022**, *99*, 153848. <https://doi.org/https://doi.org/10.1016/j.tetlet.2022.153848>.

- (165) Palm, K.; Stenberg, P.; Luthman, K.; Artursson, P. Polar Molecular Surface Properties Predict the Intestinal Absorption of Drugs in Humans. *Pharm. Res.* **1997**, *14*, 568–571. <https://doi.org/10.1023/a:1012188625088>.
- (166) Di, L.; Artursson, P.; Avdeef, A.; Benet, L. Z.; Houston, J. B.; Kansy, M.; Kerns, E. H.; Lennernäs, H.; Smith, D. A.; Sugano, K. The Critical Role of Passive Permeability in Designing Successful Drugs. *ChemMedChem* **2020**, *15*, 1862–1874. <https://doi.org/10.1002/cmde.202000419>.
- (167) Chatterjee, J.; Rechenmacher, F.; Kessler, H. N-Methylation of Peptides and Proteins: An Important Element for Modulating Biological Functions. *Angew. Chem., Int. Ed.* **2013**, *52*, 254–269. <https://doi.org/https://doi.org/10.1002/anie.201205674>.
- (168) Zhang, S.; Prabpai, S.; Kongsaree, P.; Arvidsson, P. I. Poly-N-Methylated  $\alpha$ -Peptides: Synthesis and X-Ray Structure Determination of  $\beta$ -Strand Forming Foldamers. *Chem. Commun.* **2006**, *5*, 497–499. <https://doi.org/10.1039/B513277K>.
- (169) Miyashita, A.; Yasuda, A.; Takaya, H.; Toriumi, K.; Ito, T.; Souchi, T.; Noyori, R. Synthesis of 2,2'-Bis(Diphenylphosphino)-1,1'-Binaphthyl (BINAP), an Atropisomeric Chiral Bis(Triaryl)Phosphine, and Its Use in the Rhodium(I)-Catalyzed Asymmetric Hydrogenation of  $\alpha$ -(Acylamino)Acrylic Acids. *J. Am. Chem. Soc.* **1980**, *102*, 7932–7934. <https://doi.org/10.1021/ja00547a020>.
- (170) Miyashita, A.; Takaya, H.; Souchi, T.; Noyori, R. 2, 2'-Bis(Diphenylphosphino)-1, 1'-Binaphthyl(Binap): A New Atropisomeric Bis(Triaryl)Phosphine. Synthesis and Its Use in the Rh(I)-Catalyzed Asymmetric Hydrogenation of  $\alpha$ -(Acylamino)Acrylic Acids. *Tetrahedron* **1984**, *40*, 1245–1253. [https://doi.org/https://doi.org/10.1016/S0040-4020\(01\)82411-X](https://doi.org/https://doi.org/10.1016/S0040-4020(01)82411-X).
- (171) Noyori, R.; Takaya, H. BINAP: An Efficient Chiral Element for Asymmetric Catalysis. *Acc. Chem. Res.* **1990**, *23*, 345–350. <https://doi.org/10.1021/ar00178a005>.
- (172) Misra, A.; Dwivedi, J.; Kishore, D. Role of the Transition Metal Complexes of 2,2'-Bis(Diphenylphosphino)-1,1'-Binaphthyl (BINAP) in Asymmetric Catalysis. *Synth Commun.* **2017**, *47*, 497–535. <https://doi.org/10.1080/00397911.2016.1267226>.
- (173) Yu, S.; Pu, L. Recent Progress on Using BINOLs in Enantioselective Molecular Recognition. *Tetrahedron* **2015**, *71*, 745–772. <https://doi.org/https://doi.org/10.1016/j.tet.2014.11.007>.
- (174) Telfer, S. G.; Kuroda, R. 1,1'-Binaphthyl-2,2'-Diol and 2,2'-Diamino-1,1'-Binaphthyl: Versatile Frameworks for Chiral Ligands in Coordination and Metallosupramolecular Chemistry. *Coord. Chem. Rev.* **2003**, *242*, 33–46. [https://doi.org/https://doi.org/10.1016/S0010-8545\(03\)00026-2](https://doi.org/https://doi.org/10.1016/S0010-8545(03)00026-2).
- (175) Roquette, P.; Maronna, A.; Peters, A.; Kaifer, E.; Himmel, H.-J.; Hauf, C.; Herz, V.; Scheidt, E.-W.; Scherer, W. On the Electronic Structure of NiII Complexes That Feature Chelating Bisguanidine Ligands. *Chem. Eur. J.* **2010**, *16*, 1336–1350. <https://doi.org/https://doi.org/10.1002/chem.200901479>.
- (176) Maronna, A.; Bindewald, E.; Kaifer, E.; Wadepohl, H.; Himmel, H.-J. Synthesis and Characterization of Novel Guanidine Ligands Featuring Biphenyl or Binaphthyl Backbones. *Eur. J. Inorg. Chem.* **2011**, *2011*, 1302–1314. <https://doi.org/https://doi.org/10.1002/ejic.201000981>.
- (177) Maronna, A.; Hübner, O.; Enders, M.; Kaifer, E.; Himmel, H.-J. Bisguanidines with Biphenyl, Binaphthyl, and Bipyridyl Cores: Proton-Sponge Properties and Coordination Chemistry. *Chem. Eur. J.* **2013**, *19*, 8958–8977. <https://doi.org/https://doi.org/10.1002/chem.201204294>.
- (178) Bailey, P. J.; Pace, S. The Coordination Chemistry of Guanidines and Guanidates. *Coord. Chem. Rev.* **2001**, *214*, 91–141. [https://doi.org/https://doi.org/10.1016/S0010-8545\(00\)00389-1](https://doi.org/https://doi.org/10.1016/S0010-8545(00)00389-1).
- (179) Dinger, M. B.; Henderson, W.; Nicholson, B. K. Organometallic Complexes of Platinum-Group Metals Incorporating Substituted Guanidine Dianion (Triazatrimethylenemethane) Ligands. *J.*



- Organomet. Chem.* **1998**, *556*, 75–88. [https://doi.org/https://doi.org/10.1016/S0022-328X\(97\)00654-2](https://doi.org/https://doi.org/10.1016/S0022-328X(97)00654-2).
- (180) Edelmann, F. T. Lanthanide Amidinates and Guanidinates: From Laboratory Curiosities to Efficient Homogeneous Catalysts and Precursors for Rare-Earth Oxide Thin Films. *Chem. Soc. Rev.* **2009**, *38*, 2253–2268. <https://doi.org/10.1039/B800100F>.
- (181) Edelmann, F. T. Lanthanide Amidinates and Guanidinates in Catalysis and Materials Science: A Continuing Success Story. *Chem. Soc. Rev.* **2012**, *41*, 7657–7672. <https://doi.org/10.1039/C2CS35180C>.
- (182) Kretschmer, R. Ligands with Two Monoanionic N,N-Binding Sites: Synthesis and Coordination Chemistry. *Chem. Eur. J.* **2020**, *26*, 2099–2119. <https://doi.org/https://doi.org/10.1002/chem.201903442>.
- (183) Wild, U.; Hübner, O.; Maronna, A.; Enders, M.; Kaifer, E.; Wadepohl, H.; Himmel, H.-J. The First Metal Complexes of the Proton Sponge 1,8-Bis(N,N,N',N'-Tetramethylguanidino)Naphthalene: Syntheses and Properties. *Eur. J. Inorg. Chem.* **2008**, *2008*, 4440–4447. <https://doi.org/https://doi.org/10.1002/ejic.200800677>.
- (184) Alder, R. W.; Bowman, P. S.; Steele, W. R. S.; Winterman, D. R. The Remarkable Basicity of 1,8-Bis(Dimethylamino)Naphthalene. *Chem. Commun. (London)* **1968**, 723–724. <https://doi.org/10.1039/C19680000723>.
- (185) Raab, V.; Kipke, J.; Gschwind, R. M.; Sundermeyer, J. 1,8-Bis(Tetramethylguanidino)Naphthalene (TMGN): A New, Superbasic and Kinetically Active “Proton Sponge.” *Chem. Eur. J.* **2002**, *8*, 1682–1693. [https://doi.org/https://doi.org/10.1002/1521-3765\(20020402\)8:7<1682::AID-CHEM1682>3.0.CO;2-R](https://doi.org/https://doi.org/10.1002/1521-3765(20020402)8:7<1682::AID-CHEM1682>3.0.CO;2-R).
- (186) Shimizu, S.; Watanabe, N.; Kataoka, T.; Shoji, T.; Abe, N.; Morishita, S.; Ichimura, H. Pyridine and Pyridine Derivatives. In *Ullmann's Encyclopedia of Industrial Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2000. [https://doi.org/10.1002/14356007.a22\\_399](https://doi.org/10.1002/14356007.a22_399).
- (187) Li, Q. Application of Fragment-Based Drug Discovery to Versatile Targets. *Front. Mol. Biosci.* 2020, p 180.
- (188) Kirsch, P.; Hartman, A. M.; Hirsch, A. K. H.; Empting, M. Concepts and Core Principles of Fragment-Based Drug Design. *Molecules* **2019**, *24*, 4309. <https://doi.org/10.3390/molecules24234309>.
- (189) Kim, K.-H.; Song, R.; Kim, K. M. A Metal-Centered Hydrophobic Pocket Recognizing Pyridine over Piperidine. *J. Am. Chem. Soc.* **2003**, *125*, 7170–7171. <https://doi.org/10.1021/ja035126l>.
- (190) Houk, K. N.; Leach, A. G.; Kim, S. P.; Zhang, X. Binding Affinities of Host–Guest, Protein–Ligand, and Protein–Transition-State Complexes. *Angew. Chem., Int. Ed.* **2003**, *42*, 4872–4897. <https://doi.org/https://doi.org/10.1002/anie.200200565>.
- (191) Schulz, N.; Schindler, S.; Huber, S. M.; Erdelyi, M. NMR Determination of the Binding Constant of Ionic Species: A Caveat. *J. Org. Chem.* **2018**, *83*, 10881–10886. <https://doi.org/10.1021/acs.joc.8b01567>.
- (192) Wenzel, M.; Steup, J.; Ohto, K.; Weigand, J. J. Recent Advances in Guanidinium Salt Based Receptors and Functionalized Materials for the Recognition of Anions. *Chem. Lett.* **2021**, *51*, 20–29. <https://doi.org/10.1246/cl.210527>.
- (193) Ellis, G. A.; Palte, M. J.; Raines, R. T. Boronate-Mediated Biologic Delivery. *J. Am. Chem. Soc.* **2012**, *134*, 3631–3634. <https://doi.org/10.1021/ja210719s>.

- (194) Saxon, E.; Bertozzi, C. R. Cell Surface Engineering by a Modified Staudinger Reaction. *Science* (80-). **2000**, *287*, 2007–2010. <https://doi.org/10.1126/science.287.5460.2007>.
- (195) Köhn, M.; Breinbauer, R. The Staudinger Ligation—a Gift to Chemical Biology. *Angew. Chem., Int. Ed.* **2004**, *43*, 3106–3116. <https://doi.org/10.1002/anie.200401744>.
- (196) Saxon, E.; Armstrong, J. I.; Bertozzi, C. R. A “Traceless” Staudinger Ligation for the Chemoselective Synthesis of Amide Bonds. *Org. Lett.* **2000**, *2*, 2141–2143. <https://doi.org/10.1021/ol006054v>.
- (197) Nilsson, B. L.; Kiessling, L. L.; Raines, R. T. Staudinger Ligation: A Peptide from a Thioester and Azide. *Org. Lett.* **2000**, *2*, 1939–1941. <https://doi.org/10.1021/ol0060174>.
- (198) Nilsson, B. L.; Kiessling, L. L.; Raines, R. T. High-Yielding Staudinger Ligation of a Phosphinothioester and Azide to Form a Peptide. *Org. Lett.* **2001**, *3*, 9–12. <https://doi.org/10.1021/ol006739v>.
- (199) Li, W.; Ye, Y. Polyubiquitin Chains: Functions, Structures, and Mechanisms. *Cell. Mol. Life Sci.* **2008**, *65*, 2397–2406. <https://doi.org/10.1007/s00018-008-8090-6>.
- (200) Komander, D. The Emerging Complexity of Protein Ubiquitination. *Biochem. Soc. Trans.* **2009**, *37*, 937–953. <https://doi.org/10.1042/BST0370937>.
- (201) Martin, L. J.; Raines, R. T. Carpe Diubiquitin. *Angew. Chem., Int. Ed.* **2010**, *49*, 9042–9044. <https://doi.org/https://doi.org/10.1002/anie.201005946>.
- (202) Andersen, K. A.; Raines, R. T. Creating Site-Specific Isopeptide Linkages between Proteins with the Traceless Staudinger Ligation. *Methods Mol. Biol.* **2015**, *1248*, 55–65. [https://doi.org/10.1007/978-1-4939-2020-4\\_4](https://doi.org/10.1007/978-1-4939-2020-4_4).
- (203) Aronoff, M. (UW-M. New Tools for Chemical Biology from Main Group Elements., UW-Madison, 2015.
- (204) Shah, L.; Laughlin, S. T.; Carrico, I. S. Light-Activated Staudinger–Bertozzi Ligation within Living Animals. *J. Am. Chem. Soc.* **2016**, *138*, 5186–5189. <https://doi.org/10.1021/jacs.5b13401>.
- (205) Chou, H.-H.; Raines, R. T. Conversion of Azides into Diazo Compounds in Water. *J. Am. Chem. Soc.* **2013**, *135*, 14936–14939. <https://doi.org/10.1021/ja407822b>.
- (206) Van Overschelde, M.; Vervecken, E.; Modha, S. G.; Cogen, S.; Van der Eycken, E.; Van der Eycken, J. Catalyst-Free Alcoholysis of Phosphane-Boranes: A Smooth, Cheap, and Efficient Deprotection Procedure. *Tetrahedron* **2009**, *65*, 6410–6415. <https://doi.org/https://doi.org/10.1016/j.tet.2009.05.063>.
- (207) Huang, Z.; Knaus, E. E. O<sub>2</sub>-(N-Hydroxy(Methoxy)-2-Ethanesulfonamido) Protected Diazen-1-Ium-1,2-Diolates: Nitric Oxide Release via a Base-Induced  $\beta$ -Elimination Cleavage. *Org. Lett.* **2011**, *13*, 1178–1181. <https://doi.org/10.1021/ol200053z>.
- (208) Hu, P.; Feng, T.; Yeung, C.-C.; Koo, C.-K.; Lau, K.-C.; Lam, M. H. W. A Photo-Triggered Traceless Staudinger–Bertozzi Ligation Reaction. *Chem. Eur. J.* **2016**, *22*, 11537–11542. <https://doi.org/https://doi.org/10.1002/chem.201601807>.
- (209) Hu, P.; Berning, K.; Lam, Y.-W.; Ng, I. H.-M.; Yeung, C.-C.; Lam, M. H.-W. Development of a Visible Light Triggerable Traceless Staudinger Ligation Reagent. *J. Org. Chem.* **2018**, *83*, 12998–13010. <https://doi.org/10.1021/acs.joc.8b01370>.
- (210) Klán, P.; Šolomek, T.; Bochet, C. G.; Blanc, A.; Givens, R.; Rubina, M.; Popik, V.; Kostikov, A.; Wirz, J. Photoremovable Protecting Groups in Chemistry and Biology: Reaction Mechanisms and Efficacy. *Chem. Rev.* **2013**, *113*, 119–191. <https://doi.org/10.1021/cr300177k>.

- (211) Antos, J. M.; Francis, M. B. Transition Metal Catalyzed Methods for Site-Selective Protein Modification. *Curr. Opin. Chem. Biol.* **2006**, *10*, 253–262. <https://doi.org/10.1016/j.cbpa.2006.04.009>.
- (212) Caglar, T. I.; Emmanuelle, S.; Yves, M.; A., G. W.; A., T. D. Discovery of Escherichia Coli Methionyl-TRNA Synthetase Mutants for Efficient Labeling of Proteins with Azidonorleucine in Vivo. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 15285–15290. <https://doi.org/10.1073/pnas.0905735106>.
- (213) Li, X.; Zhang, L.; Hall, S. E.; Tam, J. P. A New Ligation Method for N-Terminal Tryptophan-Containing Peptides Using the Pictet–Spengler Reaction. *Tetrahedron Lett.* **2000**, *41*, 4069–4073. [https://doi.org/https://doi.org/10.1016/S0040-4039\(00\)00592-X](https://doi.org/https://doi.org/10.1016/S0040-4039(00)00592-X).
- (214) Tam, J. P.; Yu, Q.; Miao, Z. Orthogonal Ligation Strategies for Peptide and Protein. *Biopolymers* **1999**, *51*, 311–332. [https://doi.org/10.1002/\(SICI\)1097-0282\(1999\)51:5<311::AID-BIP2>3.0.CO;2-A](https://doi.org/10.1002/(SICI)1097-0282(1999)51:5<311::AID-BIP2>3.0.CO;2-A).
- (215) Gilmore, J. M.; Scheck, R. A.; Esser-Kahn, A. P.; Joshi, N. S.; Francis, M. B. N-Terminal Protein Modification through a Biomimetic Transamination Reaction. *Angew. Chem., Int. Ed.* **2006**, *45*, 5307–5311. <https://doi.org/https://doi.org/10.1002/anie.200600368>.
- (216) Myers, E. L.; Raines, R. T. A Phosphine-Mediated Conversion of Azides into Diazo Compounds. *Angew. Chem., Int. Ed.* **2009**, *48*, 2359–2363. <https://doi.org/https://doi.org/10.1002/anie.200804689>.
- (217) Goddard-Borger, E. D.; Stick, R. V. An Efficient, Inexpensive, and Shelf-Stable Diazotransfer Reagent: Imidazole-1-Sulfonyl Azide Hydrochloride. *Org. Lett.* **2007**, *9*, 3797–3800. <https://doi.org/10.1021/ol701581g>.
- (218) van Dongen, S. F. M.; Teeuwen, R. L. M.; Nallani, M.; van Berkel, S. S.; Cornelissen, J. J. L. M.; Nolte, R. J. M.; van Hest, J. C. M. Single-Step Azide Introduction in Proteins via an Aqueous Diazo Transfer. *Bioconjugate Chem.* **2009**, *20*, 20–23. <https://doi.org/10.1021/bc8004304>.
- (219) Carvalho, E.; Iley, J.; de Jesus Perry, M.; Rosa, E. Triazene Drug Metabolites. Part 16.1 Kinetics and Mechanism of the Hydrolysis of Aminoacyltriazenes. *J. Chem. Soc., Perkin Trans. 2* **1998**, 2375–2380. <https://doi.org/10.1039/A805704D>.
- (220) Friscourt, F.; Fahrni, C. J.; Boons, G.-J. Fluorogenic Strain-Promoted Alkyne–Diazo Cycloadditions. *Chem. Eur. J.* **2015**, *21*, 13996–14001. <https://doi.org/https://doi.org/10.1002/chem.201502242>.
- (221) DeAngelis, A.; Dmitrenko, O.; Fox, J. M. Rh-Catalyzed Intermolecular Reactions of Cyclic  $\alpha$ -Diazocarbonyl Compounds with Selectivity over Tertiary C–H Bond Migration. *J. Am. Chem. Soc.* **2012**, *134*, 11035–11043. <https://doi.org/10.1021/ja3046712>.
- (222) Andersen, K. A.; Aronoff, M. R.; McGrath, N. A.; Raines, R. T. Diazo Groups Endure Metabolism and Enable Chemoselectivity in Cellulo. *J. Am. Chem. Soc.* **2015**, *137*, 2412–2415. <https://doi.org/10.1021/ja5095815>.
- (223) Aronoff, M. R.; Gold, B.; Raines, R. T. 1,3-Dipolar Cycloadditions of Diazo Compounds in the Presence of Azides. *Org. Lett.* **2016**, *18*, 1538–1541. <https://doi.org/10.1021/acs.orglett.6b00278>.
- (224) Ravera, M.; Gabano, E.; Sardi, M.; Monti, E.; Gariboldi, M. B.; Osella, D. Antiproliferative Activity of PtII Complexes with Carboxylated Phosphanes in Chelated or Ring-Opened Forms. *Eur. J. Inorg. Chem.* **2012**, *2012* (21), 3441–3448. <https://doi.org/https://doi.org/10.1002/ejic.201200282>.

