Development of Selective Inhibitors and Modulators of Collagen Prolyl 4-Hydroxylase

By

James D. Vasta

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Biochemistry)

at the

UNIVERSITY OF WISCONSIN-MADISON

2015

Date of final oral examination: 12/10/2015

The dissertation is approved by the following members of the Final Oral Committee:

Ronald T. Raines, Professor, Biochemistry
James L. Keck, Professor, Biomolecular Chemistry
Judith N. Burstyn, Professor, Chemistry
Alan D. Attie, Professor, Biochemistry
Douglas B. Weibel, Associate Professor, Biochemistry

Development of Selective Inhibitors and Modulators of Collagen Prolyl 4-Hydroxylase

James David Vasta

Under the Supervision of Ronald T. Raines

At the University of Wisconsin-Madison

Collagens are essential structural proteins in all animals, the distinguishing feature of which is the triple helix, a motif in which three polypeptide chains are interwoven to form a superhelix of great tensile strength. (2*S*, 4*R*)-4-Hydroxyproline is essential for collagen stability, and its presence is due to the posttranslational hydroxylation of prolyl residues in protocollagen by collagen prolyl 4-hydroxylases (CP4Hs). CP4Hs are members of the Fe(II)- and α-ketoglutarate (AKG)-dependent dioxygenase (FAKGD) family. FAKGD enzymes couple the decarboxylation of AKG to the oxidation of an unactivated hydrocarbon substrate.

There are a variety of human diseases associated with defects in collagen or its biosynthesis. Among there are the fibrotic diseases, which are characterized by increased deposition of extracellular matrix, as well the metastasis of some cancers, which requires reorganization of the extracellular matrix and new collagen synthesis. As the catalyst of the rate-limiting step of collagen biosynthesis, CP4Hs are viable therapeutic targets for these diseases.

An effective therapeutic has yet to be developed, and this difficulty can be attributed to the dearth of structural information available for CP4Hs, as well as the difficulty in designing selective inhibitors for members of the FAKGD family. This dissertation focuses on understanding the features of human CP4Hs that could allow for selective inhibition, and the subsequent development of inhibitors and modulators based upon that knowledge.

In CHAPTER 1, we provide a framework for considering the inhibition of CP4Hs as a therapy of the fibrotic diseases and cancer metastasis. We describe the significance of collagen as a biological scaffold and its role in the pathologies above. Then, we review the CP4Hs with a focus on the benefits and drawbacks of previously designed inhibitor scaffolds. Lastly, we discuss the imperative for inhibitors of CP4Hs with therapeutic potential and provide a justification for the work described in the subsequent chapters.

As with all FAKGDs, catalysis by CP4Hs relies on the oxidative decarboxylation of the cosubstrate AKG to generate the highly reactive ferryl ion that is believed to effect hydroxylation. Numerous simple chemical mimics of AKG have been found to inhibit FAKGDs, but most do not offer sufficient selectivity for use as therapeutic agents. CHAPTER TWO describes our investigations of a class of compounds called bipyridinedicarboxylates, which serve as useful chemical probes for prolyl hydroxylases. We find that two of these compounds, namely 2,2′-bipyridine-4,5′-dicarboxylate and 2,2′-bipyridine-5,5′-dicarboxylate inhibit a human CP4H with significantly improved potency and selectivity compared to simple AKG mimics. Our findings clarify a strategy for the development of CP4H inhibitors of clinical utility.

Unfortunately, the bipyridinedicarboxylates described above suffer from a variety of undesirable chemical properties such as poor aqueous solubility and the ability to bind and sequester free iron. CHAPTER 3 describes our efforts to redesign the bipyridine scaffold to improve these undesirable properties. We explore six bipyridine analogues where one of the 6-membered pyridine rings is replaced with an alternative 5- or 6-membered heteroaryl ring containing an alkaloid nitrogen, and find that some of the corresponding dicarboxylate analogues have improved solubility and iron-binding properties while retaining inhibitory potency towards human CP4H. We also find that the diethyl ester prodrugs of two of these compounds, namely ethyl 2-(5-ethoxycarbonylthiazol-2-yl)pyridine-5-carboxylate (diethyl pythiDC) and ethyl 2-(5-ethoxycarbonyl-1-*H*-imidazol-2-yl)pyridine-5-carboxylate (diethyl pyimDC), serve as inhibitors of human CP4H in cultured breast cancer cells at concentrations that do not perturb iron metabolism, with diethyl pythiDC being perhaps the best candidate for further development, both as a therapeutic agent and replacement for the commonly used "P4H" inhibitor ethyl 3,4-dihydroxybenzoate (EDHB).

Most work on CP4Hs has focused on the inhibition of these enzymes for a variety of clinical applications. Still, there might be benefit to *activation* of the enzyme in particular scenarios. The only reported activators of CP4Hs and related FAKGD enzymes are ascorbate and dithiothreitol (DTT), both of which are thought to function in rescuing CP4H from inactive oxidation states at the enzymic iron center, and neither of which are likely to be of clinical utility. In CHAPTER 4, we report the discovery of a series of biheteroarylmonocarboxylates that appear to serve as modest activators of human CP4H1 *in vitro*. We use both chemical and kinetic

approaches to probe the mechanism of these compounds. Specifically, we find that 2,2′-bipyridine-4-carboxylate and 2,2′-bipyridine-5-carboxylate are coactivators that enhance the rate of the ascorbate-dependent re-activation of the enzyme by serving as ligands for the iron center. Moreover we find that this mode of activation is available to many previously reported biheteroaryldicarboxylate inhibitors of the enzyme under particular assay conditions. Lastly, we put forth a general mechanism for the modulation of CP4H activity by bipyridinecarboxylates and related analogues, which provides a framework for further analyses of the kinetic mechanism of CP4Hs and related enzymes.

It has long been noted that CP4H inhibitors based upon the primary protocollagen substrate should offer better selectivity properties compared that of the AKG mimics. Still, our incomplete understanding of the interaction between CP4Hs and their peptide substrates has hindered the development of such inhibitors. CHAPTER 5 describes our efforts to further understand the conformation of peptide substrate for CP4Hs. We employ a chemical approach using alkene isosteres to replace the prolyl peptide bond and restrict the conformation of a model peptide substrate for human CP4Hs. Although neither the *cis* nor the *trans* isosteres served as substrates for human CP4H we find that the *cis* isostere serves as a more potent inhibitor compared to the *trans* isostere, suggesting that the *cis* conformation of the prolyl peptide bond is recognized preferentially. Although these results have not clarified exactly which conformation is hydroxylated by CP4H, this information can inform the design of selective inhibitors of the enzyme.

The complete three-dimensional structure of human CP4H is still unknown, particularly with respect to the catalytic domain and active site. Inspired by affinity- and activity-based proteomic probes used to study other enzyme families, we reasoned that an electrophilic analog of AKG could serve as an irreversible inhibitor of the enzyme and, hence, a useful probe for active-site residues. In CHAPTER 6, we report that 4-oxo-5,6-epoxyhexanoate, an electrophilic epoxide analogue of the AKG co-substrate, irreversibly inactivates human CP4H. The inactivation installs a ketone functionality in CP4H, providing a handle for proteomic experiments. *Caenorhabditis elegans* exposed to the esterified epoxy ketone displays the phenotype of a worm lacking CP4H. Thus, this affinity label can be used to mediate collagen stability in an animal, as is desirable in the treatment of a variety of fibrotic diseases.

Lastly, in CHAPTER 7, we explore the future of CP4H with respect to the development of inhibitors and modulators with therapeutic potential. We discuss the short- and long-term possibilities for the biheteroaryldicarboxylates developed in CHAPTER 2 and CHAPTER 3, both with respect to chemical elaborations of the inhibitors and additional biological model systems in which to explore the benefit of these compounds *in cellulo* and *in vivo*. Moreover, we discuss the immediate future directions for the bipyridinemonocarboxylate activators discovered in CHAPTER 4, which should help to refine the overall mechanism of CP4Hs. Lastly, we discuss the immediate and medium-term needs for further understanding the inhibition observed for the alkene isosteres described in CHAPTER 5, as well as possible starting points for improving the potency of the irreversible inhibitors developed in CHAPTER 6.

Acknowledgments

I cannot thank enough all of the people that have contributed to my development as a scientist. First, I would like to thank my advisor, Professor Ron Raines. Without a doubt, his research and passion for science were major contributors to my decision to pursue a graduate education at UW-Madison. I cannot thank him enough for creating an excellent environment for the development of a scientist, one filled with ample resources and an abundance of fascinating and challenging problems to which those resources may be applied.

I am also grateful to my thesis committee: Professors Jim Keck, Doug Weibel, Judith Burstyn, and Alan Attie. Their guidance, patience, constructive criticism, and excellent advice, have all contributed to my development as a chemical biologist and a scientist in general.

Now, I must acknowledge those individuals that initially inspired me to pursue science as a career. First, I thank my high school chemistry teacher Jacquelyn Anderson and my high school biology teacher Joseph Birzes. Their enthusiasm for their respective fields and passion for teaching were THE reason I chose to study biochemistry in college. Second, I thank my undergraduate mentors Professor Bernard Fried and Professor Joseph Sherma. Despite their acceptance of the *emeritus* title, they are arguably the most productive and tenacious researchers on the Lafayette College campus. Just as my high school teachers inspired me to pursue science in general, Professors Sherma and Fried were a massive driving force in my pursuit of a career in scientific research.

It can often be quite easy to overlook all of the individuals in core facilities that make our research possible. Without them, the scientific endeavor would truly come to a grinding halt. Put simply, I thank them all for their continued service and passion for science. However, I must thank a few individuals in particular for their essential contributions to my work. First, I thank Dr. Martha Vestling, whose leadership of the mass spectrometry facility in the department of chemistry has been impeccable. Her knowledge, guidance, and friendly banter have been invaluable throughout the past 5 years. Second, I thank Dr. Darrel McCaslin for his assistance and insightful discussion in the Biophysics Instrumentation Facility. Lastly, I thank Dr. Mark Anderson for his assistance in NMRFAM and for his undying love of chocolate and baked goods (to which I was an occasional beneficiary).

Over the years, I have had to pleasure of interacting with an abundance of intelligent and passionate people in the Raines laboratory. I must first thank the individuals that I never met, but that paved the way for my work, including Betsy Kersteen, Kelly Gorres, and Josh Higgins. I must also thank the individuals who helped give me a running start in the lab. These include Langdon Martin, Nick McGrath, Katrina Jensen, and Sayani Chattopadhyay, all of which contributed to my familiarization with laboratory practice in synthetic and analytical chemistry, as well as Darryl Wesener, James Thoden, Joelle Lomax and Kevin Desai, who played the same roles in the context of molecular biology and biochemistry. Moreover, I have had the pleasure of working alongside Greg Jakubczak, Amit Choudhary, Ben Caes, Greg Ellis, Mike Palte, Mike Levine, John Lukesh, Nadia Sundlass, Raso Biswas, Ho-Hsuan Chou, Christine Bradford, Brett VanVeller, Caglar Tanrikulu, Chelcie Eller, Robert Presler, Sean Johnston, Kalie Mix, Wen

Chyan, Jesus Dones Monroig, Leland Hyman, Emily Garnett, Brian Graham, Brian Gold, and many others. Mentoring Spencer Peh was always fun and rewarding. I would like to thank my baymates Robert Presler and Val Tripp. It was my absolute pleasure to sit next to both of them and benefit from their insight, passion, and comedic relief. Next, I would like to thank those lab members that I have had the pleasure of calling my friends including Thom Smith, Matt Aronoff, Ian Windsor, Aubrey Ellison, Val Tripp, and Robert Newberry. And last (and most importantly!), I would like to thank two lab members that I am blessed to have in my life and consider more like family: Kristen Andersen and Trish Hoang. I cannot imagine the last five or so years without these two! Kristen and Trish were there through the best and worst of times and supported me when I needed it the most. Moreover, Kristen and I have had a very fruitful collaboration, for which I am especially thankful.

I am grateful to my friends outside of lab, including Doug Begor, Ken Begor, Kyle Begor, Scott Hartkorn, Anthony Ciavardelli, Greg Wells, Paul DeCamp, Kyle Fritz, Michael DiGirolamo, Jorge Sawyer, Pete Coate, Mike Miller, Brian Witkowski (rest in peace), all of which I am blessed to have or have had in my life, and all of which I consider family. They have kept me happy over the years. Moreover, I would like to thank many new friends from Madison including Mike Birschbach, Graham Erwin, Navid Adnani, Brett Kopina, Erin Ronayne, Shruti Waghray, Kate Helmich, Chrissy Petzold, Dan Wolak, the rest of the “Discertators” and the Endurance House Triathlon Team.

I cannot express how grateful I am to my parents, Polly and Jim Vasta, my sister, Lauren Vasta, and the rest of the Vasta’s, Pitonzo’s, Wright’s, and Menefee’s. I have always been able

to count on their love, support, and encouragement, and for that I am truly blessed. I am also grateful for my new family members including Leslie Haber, Ryan Haber, Ted Encelewski, and the rest of the Haber's, Black's, Romig's, Puketapu's and others.

Last, and of course most importantly, I am eternally grateful to my wife and other half, Melissa. Just when you think that there simply isn't a person out there for you, she walks right on into your life! I am blessed to have Melissa's love and support every single day,
and I look forward to spending the rest of my life with her, regardless of the path we take.
Everything is better with someone else by your side…even science!

Table of Contents

Table of Schemes

Table of Figures

Table of Tables

List of Abbreviations

24PDC	pyridine-2,4-dicarboxylic acid
25PDC	pyridine-2,5,-dicarboxylic acid
ε	extinction coefficient
λ_{em}	emission wavelength
λ_{max}	emission wavelength at a local maximum
ACN	acetonitrile
AKG	α-ketoglutarate
Ala, A	alanine
Arg, R	arginine
Asp, D	aspartate, aspartic acid
ATCC	American Type Culture Collection
BCA	bicinchoninic acid
bipy	2,2′-bipyridine
bipy#C	2,2′-bipyridine-#-carboxylic acid
bipy##′DC	2,2′-bipyridine-#,#′-dicarboxylic acid
Boc	*tert*-butyloxycarbonyl
BSA	bovine serum albumin
Bu	butyl

ca.	circa, approximately
CBz	benzyloxycarbonyl
cDNA	complementary deoxyribonucleic acid
CMP	collagen mimetic peptide
CP4H	collagen prolyl 4-hydroxylase
CP4H1#	collagen prolyl 4-hydroxylase holoenzyme containing the α(#) isoform
Cr-P4H	*Chlamydomonas reinhardtii* prolyl 4-hydroxylase
d	day
Da	Dalton
DCC	*N*,*N*′-dicyclohexylcarbodiimide
DCM	dichloromethane or methylene chloride
ddH_2O	distilled, deionized water
DIEA	diisopropylethylamine
DMEM	Dulbecco's modified Eagle's medium
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DPBS	Dulbecco's phosphate-buffered saline
DTT	dithiothreitol
ECM	extracellular matrix
Dans	dansyl

DHB	3,4-dihydroxybenzoic acid
EDHB	ethyl 3,4-dihydroxybenzoate
EDS	Erlers Danlos Syndrome
EDTA	ethylenediaminetetraacetic acid
EMT	epithelial-to-mesenchymal transition
ER	endoplasmic reticulum
Et	ethyl
EtOAc	ethyl acetate
EtOH	ethanol
FAKGD	Fe(II)- and α-ketoglutarate-dependent dioxygenase
FBS	fetal bovine serum
FIH	factor-inhibiting HIF
Flp	4(*R*)-fluoro-L-proline or (2*S*,4*R*)-4-fluoroproline
flp	4(*S*)-fluoro-L-proline or (2*S*,4*S*)-4-fluoroproline
FPLC	fast performance liquid chromatography
Glu, E	glutamate or glutamic acid
Gly, G	glycine
GPC	golgi to plasma membrane carriers
h	hour
HATU	1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate

HCl	hydrochloric acid
HEPES	2[4-(2-hydroxyethyl)-1-piperazinyl]ethanesulfonic acid
HIF	hypoxia-inducible factor
His, H	histidine
HOAT	1-hydroxy-7-azobenzotriazole
HOBT	1-hydroxybenzotriazole hydrate
HPLC	high performance liquid chromatography
HRE	hypoxia responsive element
Hyl	*δ*-hydroxylysine
Hyp	(2*S*,4*R*)-4-hydroxyproline
IC_{50}	half maximal inhibitory complex
IPTG	isopropyl-β-D-1-thiogalactopyranoside
IRP#	iron regulatory protein #
JMJD2E	Jumonji domain-containing 2E
K_a	acid dissociation constant
k_{cat}	first-order enzymatic rate constant
K_d	equilibrium dissociation constant
kDa	kilodalton
Kep	(2*S*)-4-keto-L-proline
K_i	inhibitor dissociation constant
K_M	Michaelis constant

LB	Luria-Bertani medium
Leu, L	leucine
logP	partition coefficient
Lys, K	lysine
MALDI-TOF	matrix-assisted laser desorption ionization-time of flight
MCPBA	meta-chloroperoxybenzoic acid
MEM	minimum essential medium
MeOH	methanol
Met	methyl
Met, M	methionine
MHz	megahertz
min	minute
MS	mass spectrometry
MW	molecular weight
MWCO	molecular weight cut off
NaCl	sodium chloride
NaOH	sodium hydroxide
ND domain	N-terminal dimerization domain
$NHis_6$	*N*-terminal hexahistidine
NMM	*N*-methylmorpholine
NMR	nuclear magnetic resonance

NOA	*N*-oxalylalanine
NOG	*N*-oxalylglycine
OD	optical density
OI	osteogenesis imperfecta
OPr	5-oxaproline
p value	probability value
P3H	prolyl 3-hydroxylase
P4H	prolyl 4-hydroxylase
PBCV-1	*Paramecium bursaria* Chlorella Virus-1
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PDB	protein data bank
PDI	protein disulfide isomerase
PHD	prolyl hydroxylase domain-containing proteins
pK_a	–log(acid dissociation constant)
PP-II, PPII	polyproline II-type
Pro, P	(2*S*)-proline
PSB domain	peptide substrate binding domain
PTM	posttranslational hydroxylation
PVDF	polyvinylidene fluoride
RER	rough endoplasmic reticulum

RNA	ribonucleic acid
RNAi	ribonucleic acid interference
ROS	reactive oxygen species
s	second
SAR	structure–activity relationship
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SHG	second harmonic generation
siRNA	small interfering RNA
T	temperature
TACS-#	tumor associated collagen signature-#
TB	terrific broth
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TIPS	triisopropylsilane
TLC	thin-layer chromatography
T_m	temperature at the midpoint of the denaturation curve
TMS	trimethyl silane
Tris	2-amno-2-(hydroxymethyl)-1,3-propanediol
UPLC	ultra performance liquid chromatography
UPR	unfolded protein response
UV	ultraviolet

UW	University of Wisconsin
Vis	visible
$VitC^{ox}$	dehydroascorbate
$VitC^{red}$	vitamin C, ascorbate, ascorbic acid
V_{max}	maximum velocity
v/v	volume per volume
w/v or wt/vol	weight per volume
Xaa	any amino acid
Yaa	any amino acid

CHAPTER 1

Collagen Prolyl 4-Hydroxylases as Therapeutic Targets in Human Disease

1.1 Abstract

Collagen is an essential structural protein in all animals, the distinguishing feature of which is its triple helix, a motif in which three polypeptide chains are interwoven to form a superhelix of great tensile strength. (2*S*,4*R*)-4-Hydroxyproline is essential for collagen stability, and its presence is due to the post-translational hydroxylation of prolyl residues in protocollagen by collagen prolyl 4-hydroxylases (CP4Hs). CP4Hs are members of the Fe(II)- and α-ketoglutarate (AKG)-dependent dioxygenase family. These enzymes couple the decarboxylation of AKG to the oxidation of an unactivated hydrocarbon substrate. There are a variety of human diseases associated with defects in collagen or its biosynthesis. Among these are the fibrotic diseases, which are characterized by increased deposition of extracellular matrix, as well as cancer metastasis, which is facilitated by an increase in the deposition of collagen in the tumor and a reorganization of the collagen at the tumor boundary. Here, we discuss the significance of collagen as a biological scaffold and its role in the pathology of the fibrotic diseases and cancer metastasis. Moreover, we provide a framework for discussing the disruption of collagen synthesis as a therapy for these diseases, with particular focus on CP4Hs as therapeutic targets. Lastly, we review the types of inhibitors that have been developed for CP4Hs to date, and discuss the imperative for developing inhibitors of CP4Hs with improved therapeutic potential.

1.2 Collagen Structure and Stability

Comprising one-third of the human proteome and three-fourths of the dry weight of human skin, collagen is the most abundant protein in humans and modern animals.[1,2] Although different types play numerous roles in biology, collagen is most importantly a structural protein, serving as the principle component of connective tissue, bone, and the extracellular matrix (ECM) of cells.[1,2] In humans, at least 29 different types of collagens have been identified to date. In contrast to humans and other vertebrates, invertebrates often make use of an even greater variety of collagens. For example, the nematode *Caenorhabditis elegans* uses approximately 170 different collagens to fortify its proteinacious cuticle, the sum of which accounts for 80% of the total protein in this exoskeleton.[3]

Collagen Structure

The distinguishing feature of collagen is its triple helix (Figure 1.1), an elegant structural motif in which three polypeptide chains are interwoven to form a right-handed superhelical structure of great tensile strength (Figure 1.1A).[1] The individual polypeptide chains that form the triple helix are composed of numerous triplet repeats of the sequence Xaa-Yaa-Gly, where Xaa is often (2*S*)-proline (Pro) and Yaa is often (2*S*,4*R*)-4-hydroxyproline (Hyp).[1,2] Due to their high Pro and Hyp content, the individual strands of the triple helix adopt a left-handed polyproline II-type (PP-II) helical conformation.[1] In the triple helix, the three PP-II helices align with a one-residue stagger, which allows for a single N–H$_{(Gly)}$···O=C$_{(Xaa)}$ hydrogen bond per triplet repeat[1] (Figures 1.1B

and 1.1C). The presence of Hyp in collagen is due to the post-translational hydroxylation of Pro residues in the Yaa position by collagen prolyl 4-hydroxylases (CP4Hs, *vide infra*)[4], and is essential for the stability of the triple helix.[5-7] Although a variety of residues are found in the Xaa- and Yaa-positions of the triplet repeat,[8] Gly is invariant as the third residue, as its small size and conformational flexibility are essential for folding of the triple helix.[1]

Effect of Hyp on the Stability of the Triple Helix

It has been well appreciated since the seminal work in the 1970s that the presence of Hyp in collagen is important for the thermostability of the triple helix.[5,7] It was, however, not until more recently that a chemical explanation for that stability was apparent. Prior to 1995, the prevailing theory was that the triple helix was stabilized by an extensive network of water bridges connecting the 4(*R*)-hydroxyl group of Hyp to main-chain oxygen atoms, a theory that was presumably supported by X-ray crystallographic data showing the existence of such water bridges.[9] This paradigm was not, however, beyond criticisms, as triple helices of small collagen mimetic peptides (CMPs) are known to form in solvents such as methanol and propane-1,2-diol, and Hyp is known to enhance triple helix stability even in these anhydrous solvents.[10] Moreover, some researchers were concerned about the large entropic penalty of immobilizing water molecules that is associated with water-bridge formation.[10]

Questioning the water bridge theory, Raines and coworkers sought an alternative mechanism by which Hyp could stabilize the triple helix. Considering the electronic properties of the hydroxyl group, Raines and coworkers proposed that the stability conveyed by Hyp could be

related to the electronic properties of the hydroxyl group in relation to the conformation of the Pro ring. This hypothesis stemmed from investigations of the thermodynamic properties of small-molecule Pro derivatives. While it was known that the peptide bonds of tertiary amides such as Pro have only a slight preference for the trans isomer, [11,12] Raines and coworkers demonstrated that the configuration of the peptide bond in Pro derivatives (Figure 1.2A) can be modulated via electron-withdrawing substituents on the pyrrolidine ring. Specifically, they determined that electron-withdrawing substituents (–OH and –F) at the 4(*R*) position of the pyrrolidine ring exert an inductive effect that increases the preference for the *trans* isomer and enforces a C$^{\gamma}$-*exo* ring pucker (Figure 1.2B) in correlation with the inductive strength of the substituent.[13,14] This finding had implications for the stability of collagen triple helices, as Hyp residues compose approximately 38% of the residues in the Yaa position of the triplet repeat and all peptide bonds in collagen are in the *trans* configuration.[1]

From these data, Raines and coworkers hypothesized that the inductive effect of the 4*R*-hydroxyl group in Hyp is responsible for the enhanced stability of Hyp-containing triple helices and sought to test that hypothesis in a CMP model. Noting that the preference for the *trans* peptide bond and C$^{\gamma}$-*exo* ring pucker increases with the inductive strength of the 4(*R*) substituent, Raines and coworkers reasoned that collagen triple helices containing 4(*R*)-fluoro-l-proline (Flp) residues in the Yaa position of the triple helix would increase triple-helix stability. Moreover, fluoro groups attached to carbon are poor hydrogen-bond acceptors (discussed in detail in ref. 15), thus allowing the researchers to interrogate the necessity of hydrogen bonding for stabilization. On this basis, they synthesized the CMPs of the form $(ProYaaGly)_{10}$ (Yaa = Pro,

Hyp, or Flp) and determined the propensity of these CMPs to form triple helices.[15,16] In agreement with the above hypothesis, these CMPs were capable of forming triple helices as determined by circular dichroism spectroscopy and ultracentrifugation. Moreover, it was found that the thermal stability of the resulting triple helices increases with the inductive strength of the 4*R* substituent, where Pro, Hyp, and Flp containing triple-helices had thermostability (T_m) values of 41, 69, and 91°C, respectively. These results were the first true indication that the strength of the collagen triple helix results from the electronic properties of the 4*R* substituent, rather than the ability of that substituent to participate in hydrogen bonds. The landmark finding led to a host of studies investigating the parameters associated with this new hypothesis for the mechanism of Hyp in stabilizing the triple helix, which have been reviewed recently.[1] In summary, it is now postulated that Hyp stabilizes triple helices through a stereoelectronic effect, which ultimately preorganizes the individual strands of the triple helix into the conformation observed in triple helices. In this way, preorganization reduces the entropic barrier associated with trimer formation and enhances certain interatomic interactions, ultimately leading to increased triple-helical stability.[1]

1.3 Collagen Biosynthesis

Although there are many different types of collagens with various roles in biology (*vide supra*), we focus here on the biosynthesis of the fibrillar collagens (*i.e.*, types I, II, III, V, and XI[1,2,17]) as a model (Figure 1.3). As collagens are secreted proteins, their biosynthesis begins in the rough

endoplasmic reticulum (RER). After correct processing to form a mature mRNA (a splicing process that varies by collagen gene[17]), ribosome-bound mRNA is translated into individual protocollagen polypeptides (also called preprocollagens), which begin post-translational processing in the ER lumen. In addition to cleavage of the signal peptide, these modifications include an extensive list of post-translational modifications (PTMs), many of which are essential for various aspects of triple-helix or supramolecular fibril formation.

Post-translational Modification of Protocollagen

The most abundant modifications of individual protocollagens are the hydroxylation of Pro residues in the Y-position of the triplet repeat to form Hyp,[1,17,18] the implications of which are related primarily to formation and stabilization of the collagen triple helix (*vide supra*), as well as the hydroxylation of Lys residues by lysyl hydroxylase to form *δ*-hydroxylysine (Hyl).[1,17,18] Hyl residues are important in a variety of downstream steps in the synthetic pathway, for example, by serving as sites of *O*-linked glycosylation and as precursors in the downstream cross-linking process (*vide infra*) that increases the tensile strength of mature fibrils.[1,17,18] In addition to the 4-hydroxylation of Pro to form Hyp and hydroxylation of Lys to form Hyl, prolyl 3-hydroxylation has also been observed and is catalyzed by prolyl 3-hydroxylase. As opposed to the prevalence of prolyl 4-hydroxylation, prolyl 3-hydroxylation occurs only at a few specific Xaa position Pro residues, and its exact role among collagens is still under investigation. Still, variation in one of the prolyl 3-hydroxylase (P3H) isoforms is associated with osteogenesis imperfecta (*vide supra*) in both humans and mice.[19,20] Moreover, at least in the case of type VI

collagen, prolyl 3-hydroxylation appears to prevent an aberrant interaction between collagen and the platelet-specific glycoprotein VI, the interaction of which ultimately results in embryonic lethality in mice.[21]

Assembly of Procollagen Triple Helices

After the requisite posttranslational modification, the individual protocollagen strands are assembled into triple helices. The folding of triple helices *in vivo* is facilitated by the globular C- and N-terminal propeptide domains present in each protocollagen strand. In most cases, the C-terminal domains from three protocollagen strands form a globular trimeric complex that is stabilized by essential intramolecular disulfide cross-links between C-terminal domains.[1,2,17] This trimerization sets the register for triple-helix assembly of the downstream collagenous domains, after which the three strands "zip up" from C-terminus to N-terminus to form a procollagen triple helix.[1,2,17] Although the C- and N-terminal propeptides are eventually cleaved from the triple-helical domain, they also serve to enhance the solubility and prevent premature aggregation in the interim.[1,2,17]

As opposed to typical globular proteins, which generally follow the protein folding rules described by Anfinsen,[22] the assembly of collagen triple helices follows an uphill free energy landscape.[23] Hence, specialized chaperone proteins are required to assist in the folding process and enable assembly. In addition to general chaperone proteins available in the ER (*i.e.*, calnexin, calreticulin, PDI, GRP94, and BiP), procollagen triple helix assembly also appears to require additional collagen-associated chaperones such as HSP47, FKBP65, P3H1, CRTAP, and

CYPB.[23] Arguably, CP4H also likely plays a chaperone role, in addition to making assembly more favorable thermodynamically via its enzymatic activity. Still, the majority of known chaperones contribute to folding by binding and stabilizing the triple-helical conformation (HSP47 likely the primary chaperone for this process[23]) or by enhancing the rate of prolyl *cis*/*trans* isomerization,[23,24] which can be rate limiting to protein folding in general.[25] Notably, failures in the process of procollagen folding are associated with a variety of human diseases (*vide infra*).[23]

Collagen Secretion, Fibril Formation, and Cross-linking

Once assembled, procollagen triple helices traverse the secretory pathway and are extruded from the cell.[17] Thoughts about the secretion of collagen and other large cargos are evolving rapidly, and have been the subject of recent review.[26] Hence, the details of this fascinating process will not be discussed at length here. In summary, as procollagen triple helices are substantially longer (~300 nm[1]) than are the typical secreted protein, collagen-producing cells such as fibroblasts appear to use specialized secretory machinery to facilitate the extrusion process.[26,27] After traversing the Golgi apparatus, procollagens are packaged into large internal transport compartments called Golgi to plasma membrane carriers (GPCs). Moreover, as noted above, the N- and C-terminal propeptides of the procollagens must be removed, and there is evidence that the cleavage process (catalyzed by specific procollagen N- and C-proteinases[1,2,26]) begins within these GPCs.[27] After cleavage, the resulting collagen triple helices become insoluble and begin to aggregate and initiate fibril formation, although the extent of such fibril formation is limited

within the GPCs.[26,27] Subsequently, these fibril-containing GPCs are trafficked to extracellular protrusions in the PM called fibripositors, from which collagen fibrils are extruded from the cell and incorporated into growing fibers.[26,27]

Individual triple helices are believed to possess all the requisite information to guide their own self-assembly into larger fibrils. Still, mature collagen fibers observed in biological tissues are composed of a complex hierarchy of fibrillar subunits, and further modifications are required during the supramolecular assembly process that provide additional stability and tailor the mechanical properties of the mature fibers.[1,17,26] Specifically, covalent cross-links are created between the N- and C-terminals of neighboring collagen monomers, the formation of which stabilizes both individual fibrils as well as the supramolecular assembly as a whole.[1,17,26] These cross-links are facilitated by lysyl oxidase, which catalyzes the oxidative deamination of Lys and Hyl residues to form corresponding aldehyde derivatives, after which these aldehydes react with opposing Lys, Hyl, or aldehyde derivatives to generate one of many possible covalent cross-links. The exact chemical nature of such cross-links is complicated[1,17,26] and will not be discussed further here, although we note that the formation of these cross-links (as well as their abundance, localization, and identity) is essential for the overall mechanical properties of the mature collagen fiber.[17,26]

1.4 Collagen-Related Diseases

Abnormalities in the structure and metabolism of collagen are associated with a variety of heritable and nonheritable human diseases. Included among the heritable diseases are the rare genetic disorders Osteogenesis Imperfecta (OI) and Ehrlers Danlos Syndrome (EDS), both of which involve mutations that decrease the production or structural stability of fibrillar collagens in a tissue.[1,2] Diseases within the OI and EDS spectra are most commonly associated with single-base substitutions that lead to the substitution of an obligate glycine with another residue. Such mutations can cause an interruption in the folding of the triple helix at that particular location, or even prevent folding of the triple helix beyond that point.[2] Still, we note that any mutation in a collagen that disrupts triple helix or supramolecular fibril assembly is likely to be associated with a human disease, the severity of which depends on the type of collagen and its required mechanical properties in the tissue of interest.[2]

Of primary interest here are diseases in which normal collagen is overproduced or produced in such a way that it facilitates a pathological process. More specifically, we focus on two such cases: the fibrotic diseases and the metastasis of some cancers. In both such cases, collagen that is produced exacerbates or facilitates the disease, and it is thought that the therapeutic disruption of collagen synthesis during this time would be beneficial.[28-30]

Fibrotic Diseases

Cumulatively, the fibrotic diseases represent a large unmet clinical need, as they contribute to an estimated 45% of deaths in the industrialized word.[31] Fibrosis in general is the formation of

excess fibrous connective tissue in an organ, and likely evolved as a protective response to tissue injury.[32] Under normal circumstances, this excess connective tissue can be processed and removed to restore proper structure and function to the tissue.[32] Nevertheless, a variety of factors, including recurrent or persistent epithelial injury contribute to the development of chronic and pathological fibrosis. Thus, fibrotic diseases are associated with a loss of control of the wound-healing process, the increased deposition of ECM, and ultimately the accumulation of pathogenic levels of collagen-rich scar tissue. Pathological fibrosis can affect many tissues and organs, and common examples of fibrotic diseases include both multisystemic diseases such as systemic sclerosis, as well as those of individual organs including the lungs (cystic fibrosis and idiopathic pulmonary fibrosis), liver (cirrhosis), heart (atrial fibrosis, endomyocardial fibrosis, and scarring from myocardial infarction), intestine (Crohn's disease), kidney (chronic kidney disease) and skin (scleroderma and keloids).[33]

Within the last 5 years, the fibrotic diseases have received considerable attention, and the common underlying molecular and cellular mechanisms (Figure 1.4) that contribute to these chronic states, as well as possible modes of therapeutic intervention, have been reviewed extensively.[32-39] Briefly, many fibrotic diseases are initiated by epithelial injury, which can result from a variety of stimuli including trauma, toxins, inflammatory processes, and infection.[32] Most of these stmuli cause downstream phenotypic changes in the affected tissue including cell death or dysregulated metabolism, production of reactive oxygen species (ROS), ER stress, activation of fibrosis-associated signaling pathways (*i.e.*, TGF-β signaling), recruitment of immune cells, and pathological epithelial-to-mesenchymal transition (EMT).[32] Virtually all fibrosis

accompanies the accumulation of large numbers of mesenchymal cells, the vast majority of which are myofibroblasts.[32] These myofibroblasts are influenced heavily by the phenotypic changes described above (especially activation by TGF-β and immune cell-derived pro-fibrotic cytokines), and are the principle source of the pathologic connective tissue and ECM that accumulates during fibrosis.[32]

The properties of the fibrotic ECM can be vastly different from those of the normal ECM for a particular tissue, which can ultimately disrupt function of the afflicted tissue/organ. Tissues of healthy organs are typically composed of a variety of ECM proteins (*i.e.*, laminins, proteoglycans, and the non-fibrillar type IV collagens), which are normally located in the basement membrane.[32] Still, the location and composition of ECM proteins in fibrotic tissues can be altered markedly. As described above, there is generally an increase in total ECM in fibrotic tissue, and the composition of such pathological ECM includes high levels of the fibrillar types I and III collagens (and other collagen subtypes), as well as increases in fibronectin, osteopontin, hyaluronan, proteoglycans, and laminins.[32] Type I and III collagens are heavily cross-linked and the fibrillar structures they produce can affect a variety of tissue properties including architecture and the behavior of both normal cells and recruited myofibroblasts.[32] As a consequence, fibrotic tissues are generally much more stiff compared to the corresponding healthy tissue, and the changes described above can results in tissue dysfunction and ultimately tissue/organ failure.

Due to the progressive nature of the fibrotic diseases and our limited understanding of the disease etiology, fibrotic diseases have proven difficult to treat. Organ transplantation is still the primary method of treatment, and since there is an inherent limit in the supply of organs, patients

often die waiting for a transplant. At this time, the few antifibrotic drugs available for use in the United States include imatinib mesylate, pirfenidone, and nintedanib. Imatinib mesylate (gleevec®), is a small-molecule inhibitor that selectively inhibits several tyrosine kinases involved in the TGF-β signaling pathway.[33] While it has shown promise in the treatment of a variety of fibrotic diseases, including systemic sclerosis, nephrogenic systemic fibrosis, and scleroderma, this success is tempered significantly by a wide range of side effects (congestive heart failure, edema, muscle cramps, diarrhea, anemia, neutropenia, and thrombocytopenia) as TGF-β signaling is important for normal homeostasis.[33] In addition, pirfenidone is a small-molecule with broad anti-inflammatory activity that disrupts TGF-β and p38 signaling. It is marketed under a variety of trade names (Pirespa® in Japan, Pirfenex® in India, Esbriet® in Europe, Canada, and the United States, Etuary® in China, and KitosCell LP® in Mexico). Lastly, nintedanib (Vargatef®) is a broad-spectrum tyrosine kinase inhibitor that disrupts various signaling pathways involved in fibrosis. Pirfenidone and nintedanib are two of the few treatments available for idiopathic pulmonary fibrosis, though both are not without serious side effects.

Nevertheless, recent advances in our understanding of the molecular mechanisms of induction of tissue fibrosis have been encouraging.[32,33] There are now a variety of appreciated therapeutic targets for fibrosis (Table 1.1), as well as myriad drugs now in various stages of clinical development.[32] Still, we note that many of these drugs target signaling pathways related to the induction of fibrosis and activation of myofibroblasts, and comparatively few target the downstream processes involved in production of pathologic ECM (*i.e.*, collagen, *vide infra*).

Cancer Metastasis

Cancer, a group of diseases characterized by abnormal cell growth with the potential to spread to other parts of the body, is one of the leading causes of morbidity and mortality worldwide.[40] Moreover, invasion and metastasis of individual cancer cells is a fundamental aspect of all cancers. Although most research has focused on understanding and disrupting tumor development (a valid approach to cancer treatment), we note that the majority of cancer-related deaths are caused by cancer metastasis,[41] the understanding of which lags significantly in comparison (*e.g.,* there are no drugs currently available that prevent metastasis[41]). Here, we focus on the metastasis of cancers in the context of collagen and the ECM, and discuss recent findings that highlight unappreciated pathways involved in metastasis and suggest possible therapeutic targets.

There is a growing appreciation for the influence of the ECM on cell and tissue behavior,[42] and this influence is likewise important for cancer cells, a topic that has been reviewed recently.[43] The ECM of cancer cells undergoes considerable changes during tumor progression.[43] In addition to ECM remodeling, the increased deposition of ECM components including collagens type I, III, and IV, fibronectin, and proteoglycans has been observed,[43] as in fibrosis[44] (*vide supra*).

Considering these changes in the ECM, many cancers are known especially for their aberrant ECM phenotypes. The classic example is breast cancer, and breast cancer prognosis has long been correlated with aberrant ECM phenotypes, such as increased tissue density compared to normal breast tissue,[45,46] as well as more recently recognized tumor associated collagen

signatures (TACS) that have been observed using second harmonic generation (SHG) imaging.[46-48] In fact, it is now recognized that the collagenous environment of breast cancer tumors changes considerably during breast tumor progression. More specifically, Keely and coworkers used SHG imaging to investigate mouse models that recapitulate the histological progression of human breast cancer, and found that breast tumors display three distinct TACS throughout tumor progression.[48] The first signature, termed TACS-1, occurs early in tumor formation, and is associated with a localized increased deposition of ECM at the tumor lesion.[48] As the tumor grows and increases in size, TACS-2 is observed, which is characterized by a straightening of collagen fibers that are aligned parallel to the tumor boundary.[48] Lastly, stromal reorganization occurs in the final TACS-3 stage, and collagen fibers are observed to bundle and re-align perpendicular to the tumor boundary.[48] TACS-3 is associated with sites of tumor invasion and the beginnings of metastasis,[46,48] and the presence of TACS-3 is a strongly correlated with tumor prognosis.[47] Moreover, many have observed that individual cancer cells preferentially invade the surrounding tissue by following the perpendicular collagen bundles that form during TACS-3, using these bundles as molecular "highways" to escape the tumor and enter the bloodstream.[46,48,49] In fact, similar perpendicular collagen fibers have been observed in other cancers,[43] suggesting that cancer cells might more generally utilize collagen fibers as a means to escape a tumor during metastasis.

As TACS-3 is correlated with prognosis, it is highly likely that this signature could serve as a useful biomarker in the diagnosis of breast cancer,[47] and possibly other cancers.[43] Additionally, it has been postulated that antimetastatic therapies could involve blocking this

aspect of tumor progression, for example, by preventing induction of TACS-3, by disrupting interactions with the ECM, or by inhibiting collagen deposition itself.[47] Thus, there has been considerable interest in understanding the underlying mechanisms responsible for the increase in collagen deposition observed during tumor progression.

While these mechanisms are under investigation, a few studies have determined that the intratumoral hypoxia that commonly occurs during tumor progression[50,51] is at least partially responsible for the increased collagen synthesis.[28,29,52] More specifically, the intratumural hypoxia causes induction of hypoxia signaling,[53-55] a signal transduction pathway that is mediated by the hypoxia inducible factors HIF-1 and HIF-2. Both HIFs are heterodimeric proteins composed of the HIF-1α or HIF-2α and the constitutively expressed HIF-1β subunit. Under hypoxic conditions, HIF-1α and HIF-2α are stabilized, resulting in their accumulation and trafficking to the nucleus where they ultimately complex with HIF-1β and other transcriptional co-activators to regulate the transcription of target genes.[53] The levels of HIF-1α and HIF-2α are regulated through an oxygen-dependent mechanism requiring the activity of HIF-prolyl-hydroxylases (HIF-PHDs) and the asparaginyl hydroxylase factor inhibiting HIF (FIH), all of which are members of the Fe(II)- and α-ketoglutarate (AKG)-dependent dioxygenase (FAKGD) family of enzymes.[56] Under normoxia, hydroxylation of HIF-1α and HIF-2α by HIF-PHDs on two particular prolyl residues targets them for ubiquitination and subsequent proteasomal degradation.[53] Conversely, hypoxic conditions lower the activity of HIF-PHDs, which results in the significant accumulation of HIF-1α and HIF-2α and initiation of the hypoxia response as described above.

Commonly recognized targets of the HIF transcription factors include proteins involved in ameliorating the effects of hypoxia, at least in the case of certain cancers. Ultimately, the accumulation of HIF-1 increases the expression of various genes of the collagen biosynthetic pathway. These genes include various CP4Hs[29] and lysyl hydroxylase 2,[28] all of which are essential for proper assembly of collagen fibers (*vide supra*) and represent new therapeutic targets for the development of anti-metastatic drugs.

1.5 Targeting Collagen Synthesis as a Therapy for Collagen-Related Diseases

Collagen is essential as a biological scaffold in animals. Nonetheless, disruption of collagen synthesis during particular pathological processes such as the fibrotic diseases and cancer metastasis is a viable therapeutic approach. Here, we focus on the aspects of collagen synthesis that can likely be targeted therapeutically. Then, we focus our attention on a particular target of interest herein, CP4H, and discuss specific reports that have validated this target *in vivo*.

Possible Targets in Collagen Synthesis, Successes, and Consequences

As the fibrilar collagens type I and III are the principle component of the pathological ECM produced in both the fibrotic diseases and in TACS-3 of cancers (*vide supra*), collagen synthesis is an important pathway to consider for therapeutic intervention in these collagen-related diseases. Virtually ever step in collagen synthesis from the upstream signaling pathways and

expression to the downstream fibril formation and cross-linking could be targeted, although the long-term consequences of intervening at any step is unknown.

In the cases of some fibrotic diseases, such as those caused or exacerbated by the action of the patient, removal of the source of the recurring epithelial injury is the most obvious treatment solution. The relevant examples include liver cirrhosis and idiopathic pulmonary fibrosis, which are most commonly induced by the excessive consumption of alcohol[57] and cigarette smoking,[58] respectively. Due to the addictive properties of the agents that induce the fibrosis, it is rare that removal of such practices is successful. As a consequence, most research has focused on disrupting the upstream signaling pathways induced by epithelial injury,[43,59] (*vide supra*), or even the expression of collagen genes themselves.[60,61] These are reasonable strategies and should continue to be explored, but many investigators are concerned with the long-term consequences of inhibiting signaling pathways that are of general homeostatic importance. Hence, targets in the downstream processes in the collagen biosynthetic pathway are also of considerable interest.

Folding of the Triple Helix

As folding of the collagen triple helix is an essential step in the pathway, targets in the early steps of both post-translational modification and triple-helix assembly are particularly attractive. Whereas the misfolding of many proteins in the ER are known to induce the unfolded protein response (UPR),[23] the lack of folding or misfolding of procollagen triple helices appears to affect cells in a different way.[23] Procollagen misfolding (Figure 1.3) appears to induce a response that

is similar to the aggregated protein response, which targets insoluble protein aggregates (such as those of procollagen) for autophagy or lysosomal degradation[62,63] and provides a protective affect against diseases associated with protein aggregation.[64] Thus, this finding provides a tentative validation for the prevention of procollagen folding, although the long-term consequences require further evaluation.

We note that, in particular, the hydroxylation of protocollagen strands by CP4Hs and the activity of various collagen-associated chaperones are arguably the two most important processes that aid in procollagen folding (*vide supra*). Targeting of collagen-associated chaperones has been suggested[61] (*i.e.*, inhibition of HSP47[65]), but the therapeutic inhibition of CP4Hs has received significantly more attention due CP4Hs catalyzing the first committed step in collagen folding. As the therapeutic inhibition of CP4Hs is of principle interest in this thesis, this topic will be explored further below.

Other Post-translational Modifications

PTMs such as prolyl 3-hydroxylation and protocollagen glycosylation have yet to be explored as targets for fibrotic diseases and cancer metastasis. The hydroxylation of Lys to Hyl, however, has been of particular interest. Evaluation of this strategy is still ongoing, but a few studies have suggested that lysyl hydroxylase could be a useful therapeutic target for cancer metastasis. As described above, the intratumoral hypoxia found in many tumors causes increased expression of multiple proteins in the collagen biosynthetic pathway through a HIF-dependent mechanism. One of those proteins is lysyl hydroxylase, and increased expression of multiple isoforms of this

protein has been shown to correlate well with cancer prognosis.[28,66] Moreover, siRNA knockdown experiments in mouse cancer models have shown that lysyl hydroxylase isoform 2 is essential for the metastasis of both breast cancers[66] and sarcomas.[28] Considering these findings, lysyl hydroxylase 2 appears to be an excellent target for the development of anti-metastatic drugs. Still, the long-term consequences of inhibiting the enzyme remain to be evaluated, and the development of selective inhibitors of this enzyme is still in its infancy.[67]

Trafficking, secretion, and propeptide cleavage

Due to their common roles in other pathways, it is not likely that disruption of the intracellular trafficking or secretion of collagens will provide a means to disrupt fibrosis or metastasis in a way that is sufficiently selective for clinical use. In contrast, the steps in which the C- and N-terminal propeptides are cleaved could represent a viable therapeutic target. Although the N-terminal proteinase has received little attention, the C-terminal proteinase has received much because cleavage of the C-terminal propeptide is essential for fibril formation.[59] Some fairly potent peptidomimetic inhibitors of this enzyme have been developed,[68] as well as a series of selective non-peptidic inhibitors.[69,70] Moreover, inhibitors of the enzyme have shown a protective affect from the excessive collagen deposition in rabbit models of cutaneous scarring[71] and are now being tested in clinical studies as dermal anti-scarring agents.[69] Hence, inhibition of the C-terminal proteinase could be fruitful, at least for dermal scarring diseases, though clinical relevance for other collagen-related diseases has yet to be established.

Collagen cross-linking

As the stiffness of the pathological collagens type I and III produced during fibrotic disease and cancer metastasis is highly correlated with their covalent cross-linking during fiber formation, the processes controlling that cross-linking have received considerable attention.[59] The principle enzyme involved in collagen cross-linking is lysyl oxidase (*vide supra*), and many studies have demonstrated that disruption of this enzyme can be beneficial, both for fibrosis[59,72] and cancer metastasis.[73-75] Despite considerable interest in developing inhibitors for this enzyme, very few small-molecule inhibitors with clinical relevance have been developed thus far. However, because collagen cross-linking occurs during fiber formation outside of the cell, antibodies to lysyl oxidase isoforms have proven to be clinically relevant for the treatment of both fibrosis and cancer.[32] For example, the noncompetitive monoclonal antibody to the lysyl oxidase isoform LoxL2 (Table 1.1) is currently in clinical trials for the treatment of multiple fibrotic diseases and cancer, and the results have been promising thus far.

CP4Hs as Therapeutic Targets in Fibrosis and Cancer Metastasis

As catalysts of the first committed step in collagen folding, as well as the rate-limiting step in the overall collagen synthetic pathway, CP4Hs are attractive targets for diseases in which collagens are overproduced. CP4Hs catalyze the hydroxylation of protocollagen strands, which is essential for the folding and thermal stability of the triple helix (*vide supra*).[1,30] Disruption of CP4H activity as a treatment for the fibrotic diseases has been of interest since the discovery of the enzyme itself,[76] and many studies have demonstrated that disrupting this activity could be of

benefit for the fibrotic diseases.[18,59,77-79] Moreover, the development of intratumoral hypoxia during breast cancer has been shown to increase the expression of two CP4H isoforms, and the levels of both correlate well with cancer prognosis.[29] Lastly, these two CP4H isoforms have been shown to be essential for breast cancer metastasis, and disruption of their activity both by genetic knockdown and small-molecule inhibition has been shown to reduce breast cancer metastasis in mice.[29]

The above findings have made CP4Hs even more attractive as therapeutic targets in recent years. Considering the complete absence of anti-metastatic drugs,[41] we note that any clinically useful inhibitor of metastasis would be of clear benefit. Moreover, we note that whereas development of inhibitors for human CP4Hs with sufficient selectivity has proven challenging thus far (*vide infra*), the development of selective small-molecule inhibitors for similar enzymes with anti-metastatic potential (*i.e.,* lysyl oxidase) is still in its infancy. Lastly, the consequences of procollagen misfolding have been discussed above and are most likely tolerable for both the treatment of metastasis, and for the long-term treatments required for most chronic fibrotic diseases. Thus, we further endorse CP4Hs as therapeutic targets in both fibrotic and metastatic diseases. Herein, we review CP4H enzymes as a class. Then, we discuss the progress to date towards the development of a clinically relevant inhibitor of CP4Hs, as well as the shortcomings and hurdles encountered thus far.

1.6 Collagen Prolyl 4-Hydroxylase

A critical process in collagen biosynthesis is the CP4H catalyzed[4] post-translational hydroxylation of Pro residues in the Yaa position to form Hyp.[1,2,4] Considering the estimated abundance of collagen in humans (1/3) and that of Hyp within collagen (~38% $\times {}^{1}/_{3}$)[80], this posttranslational modification is one of the most abundant in humans.[4] Thus, it is not surprising that CP4Hs have been found to be essential in animals,[81-83] as the conformational stability of mature collagen relies on its hydroxylation.[5-7] Collagen is a secreted protein and CP4Hs are located in the lumen of the endoplasmic reticulum where they carry out one of the first steps in the collagen biosynthetic pathway.

Structure of CP4Hs

CP4Hs are members of the FAKGD superfamily, which has myriad roles in biology, ranging from epigenetic regulation and signal transduction to small-molecule biosynthesis.[56] These enzymes couple the oxidative decarboxylation of AKG to the hydroxylation of a hydrocarbon substrate via a radical mechanism involving a highly reactive Fe(IV)=O species (ferryl ion).[56] In mammals, CP4Hs exist as $\alpha_2\beta_2$ tetramers, where the α-subunit is the catalytic subunit and the β-subunit is protein disulfide isomerase, a multifunctional chaperone protein that maintains the α-subunit in a soluble and active conformation.[4,30] Three isoforms of the α-subunit, α(I), α(II), and α(III), have been identified in humans.[4,30] All α-subunit isoforms form tetramers with the β-subunit, which we refer to herein as the CP4H1, CP4H2, and CP4H3 holoenzymes. As the most

prevalent of the isoforms, the α(I)-subunit has been characterized extensively and is the focus of most work described below. Whereas the structure of the tetrameric complex is unknown, those of individual domains of the α(I)-subunit have provided insight into the manner in which CP4Hs interact with the protocollagen substrate, as well as the means by which the α-subunits dimerize to facilitate formation of the tetramer.[84-86]

The α-subunits are multidomain proteins of approximately 59 kDa that contain three domains important for both catalysis and assembly of the tetramer. The N-terminal dimerization domain (ND domain, residues 1–143) is the most recently identified domain and is important for complex assembly.[84] This domain facilitates dimerization of two α-subunits by forming an extended four-helix bundle that includes an antiparallel coiled-coil dimerization motif between the two α-subunits.[84] In contrast to the recently discovered ND domain, the peptide substrate binding (PSB) domain (residues 144–244) and C-terminal catalytic domain (residues 245–517) have long been appreciated.[4,30] The PSB domain is a tetratricopeptide binding domain[85,86] that enhances the affinity of the enzyme for long polymeric substrates generally in the PP-II helical conformation (*i.e.*, protocollagens).[85] Moreover, the PSB domain is believed to possess subsites for the binding of at least three collagen triplet repeats[84] using three tyrosine residues (Tyr233, Tyr230, and Tyr196) spaced similarly to the repeat distance of a PP-II helix to form stacking interactions with prolines in each of the three repeats.[84,86] The ND and PSB domains are, by far, the most structurally characterized elements of the CP4H tetramer, whereas the structures of the C-terminal catalytic domain, human β-subunit, and holoenzyme remain elusive.

The C-terminal domain is homologous to the FAKGD family and contains the enzymic active site. Despite the absence of a crystal structure for this domain, its homology to the FAKGD enzymes has been highly informative, especially with the increasing number of crystal structures of FAKGD enzymes that are now available in the protein data bank (PDB). Like all enzymes of the FAKGD superfamily, catalysis by CP4Hs requires Fe(II), which is bound by a conserved His-X-Asp/Glu…X_n…His motif, as well as the cosubstrates AKG and dioxygen.[56] AKG chelates to enzyme-bound Fe(II) using its C-1 carboxylate and C-2 keto groups, while the C-5 carboxylate group engages in Coulombic interactions with a basic residue (typically arginine or lysine) and additional hydrogen bonds. In the case of the human α(I)-subunit, biochemical data has suggested that the Fe(II) is likely chelated by Asp414, His412, and His483, with Lys493 providing Coulombic stabilization of the C-5 carboxylate of AKG.[87]

The structure of the C-terminal catalytic domain of the human α(I)-subunit remains unknown. The structures of related P4H enzymes from lower organisms have, however, been particularly useful in providing a framework for considering the likely structure of the human catalytic domain (Figure 1.5). The more closely related of these P4H enzymes is from the algae *Chlamydomonas reinhardtii.* Although algae and higher plants do not contain collagen, they do utilize a variety of proline-rich glycoproteins to fortify their cell wall, many of which have high Hyp content.[88,89] The genome of *C. reinhardtii* encodes at least 10 P4H-like polypeptides, though one of the polypeptides termed Cr-P4H-1 (a soluble 29-kDa monomer[90]) has been found to be essential for proper cell assembly *in vivo*.[90] The overall primary sequence identities of Cr-P4H-1 and CP4Hs are relatively low, but the sequence of the processed Cr-P4H-1 polypeptide is 26%

identical to that of the C-terminal catalytic domain of human CP4H1 and all of the essential catalytic residues are conserved.[90] Moreover, Cr-P4H-1 has even been observed to catalyze the hydroxylation of synthetic collagen strands *in vitro*, albeit with less positional specificity (*i.e.,* hydroxylation occurs at both the X- and the Y-positions).[90] Considering the similarities described above between Cr-P4H-1 and human CP4H1, as well as the structural characterization of Cr-P4H-1 (Figure 1.5A),[91,92] this related P4H is often viewed as a reasonable model for the human CP4H1 catalytic domain, although further structural work and additional structure–activity relationships for known inhibitors would be useful for making comparison between the two enzymes.

While Cr-P4H-1 provides arguably the closest model for the human CP4H active site, it should be noted that a crystal structure (Figure 1.5B) of another related P4H (vP4H) from *Paramecium bursaria* Chlorella Virus-1 (PBCV-1) has recently been reported.[93] Eriksson and co-workers first reported vP4H (a soluble, 25-kDa monomer[93]), which is 23% identical to the C-terminal catalytic domain of the human CP4H α(I)-subunit.[94] Moreover, the proposed ferryl ion (*vide infra*) was observed spectroscopically during the reaction cycle of this P4H,[95] providing the first evidence that a P4H actually uses this highly reactive iron species during catalysis. Still, as with Cr-P4H-1, further structural work and comparative structure–activity relationships for known inhibitors would be valuable for making comparisons between vP4H, Cr-P4H-1, and human CP4H1.

Mechanism of CP4Hs

All FAKGDs including CP4Hs are believed to effect catalysis through a similar two-stage mechanism in which AKG is first oxidatively decarboxylated to generate a highly reactive ferryl ion, after which the ferryl ion interacts with a hydrocarbon substrate and effects hydroxylation via a radical rebound process (Figure 1.6 see: Coupled Reaction).[56] Although human CP4Hs are reasonably characterized from an enzymological perspective,[30,78,96] much of what we understand about CP4Hs comes from studies of the corresponding enzyme from chicken (due to the historical ease of isolating the enzyme from chicken tissues[76]), studies of other FAKGD model systems,[56] and theoretical considerations.[97] Classical kinetic studies of chicken CP4H1 have suggested that CP4Hs effect catalysis via an order ter–ter mechanism in which AKG first binds to the CP4H·Fe(II) complex, after which O_2 and the peptide substrate bind in an ordered fashion.[98] After formation of the ferryl ion and subsequent hydroxylation of the peptide substrate, the proposed order of product release is hydroxylated polypeptide product, followed by CO_2 followed by succinate. This order of product release is experimentally supported,[98] but remains under debate.[98]

CP4Hs require the cofactor ascorbate (vitamin C, $VitC^{red}$) for catalytic activity, a property that has classically been linked to scurvy,[99,100] a disease caused by long-term vitamin C deficiency in humans and some higher animals. Still, ascorbate is not required stoichiometrically during normal catalytic cycles, but rather, is thought to rescue the enzyme from unproductive oxidation states at the CP4H iron center.[100] During a typical ter–ter mechanism, there is no requirement for a reducing agent such as vitamin C, as the radical rebound process returns the iron center to the Fe(II) state required for catalysis. This notion is consistent with catalysis by

CP4H, as the CP4H from chicken has been observed to catalyze prolyl hydroxylation at a maximal rate for at least few catalytic cycles in the complete absence of ascorbate.[101] Yet, even at saturating concentrations of the peptide substrate, CP4Hs (and other FAKGDs) have been observed to decarboxylate AKG oxidatively in a manner that does not result in the hydroxylation of the peptide substrate.[101,102] This pathway is termed the "uncoupled reaction" and vitamin C is believed to operate as a nearly specific reducing agent to reduce the iron center from (an) inactive oxidation state(s) (likely Fe(III)[101]) back to the Fe(II) state required for catalysis (Figure 1.6 see: Uncoupled Reaction). Although vitamin C is known to act stoichiometrically during the uncoupled reaction,[102] the more detailed features of the mechanism such as the ligand environment of the iron center during the reduction and putative binding mode of vitamin C remain unknown. Finally, the slow step of the CP4H coupled reaction (Figure 1.6) is not currently known, though the requirement for vitamin C described above suggests that the slow step of the overall reaction, including both the coupled and uncoupled pathways, likely involves a step in the uncoupled pathway.

1.7 Inhibition of Collagen Prolyl 4-Hydroxylases

Given the historical recognition of fibrotic diseases, the inhibition of CP4Hs has been of interest essentially since the enzyme was discovered in the late 1960s and early 1970s. Moreover, the finding that CP4Hs are essential for breast cancer metastasis[29] (and perhaps other cancers) has only increased that interest. As a member of the FAKGD family of enzymes, CP4Hs are

susceptible to inhibition by a variety of mechanisms, many of which have been reviewed recently in the context of the FAKGD family as a whole,[67] as well as with respect to the CP4Hs from chicken.[103] In general, FAKGD enzymes are inhibited by alternative metal ions, metal chelators, mimics of the cosubstrate AKG, mimics of their corresponding primary substrate (*i.e.*, collagen-like peptides for CP4Hs), and various natural products and antioxidants. Here, we discuss the CP4H inhibitors that have been discovered thus far in the context of some of the inhibitor classes described above (and some special cases), focusing more on inhibitor classes of therapeutic value and those that have been used in human cells or other biological contexts. Lastly, we discuss the state of the art for inhibition of CP4Hs *in vivo.*

Alternative Metal Ions

Considering the fairly low affinity of most FAKGD enzymes for ferrous iron, many transition metal ions such as Mn^{2+}, Ni^{2+}, Co^{2+}, Cu^{2+}, and Zn^{2+} have been found to be modest inhibitors of these enzymes.[67] The consequences are perhaps best appreciated when considering the classic hypoxia phenotype observed when treating cells with Co^{2+} (and Ni^{2+} to a lesser extent), which has been found to be a competitive inhibitor of the HIF-PHDs involved in hypoxia signaling.[67,104] CP4Hs have fairly weak affinities for Fe^{2+} (*i.e.*, K_m = 4 μM for human CP4H1 and K_m = 6 μM for human CP4H2[96]) compared to other FAKGD counterparts, and are thus particularly susceptible to inhibition in this manner (*i.e.*, K_i = 0.6 μM for inhibition by the Zn^{2+}

ion[105]). Still, as metal ions are highly unlikely to be useful as CP4H inhibitors in a therapeutic context, this topic with not be reviewed further herein.

Metal Chelators

As with metal ions, virtually all FAKGD enzymes are inhibited by metal chelators (Figure 1.7A). These inhibitors include simple chelators such as ethylenediaminetetraacetic acid (EDTA), deferroxamine, 2,2′-bipyridine (bipy), 1,10-phenanthroline, and catechol, among others. Simple metal chelators are generally thought to inhibit by binding and sequestering free iron (thereby rendering iron unavailable to the enzyme), and generally display variable K_i values in the micromolar range that are dependent upon the amount of free iron added to assays *in vitro*.[106,107] Indeed, bipy itself was classically used in a biological context to sequester free iron and disrupt collagen synthesis,[108,109] providing the initial demonstration that iron was required for the process. Although, metal affinity is an important component of some inhibitor classes (*vide infra*), simple metal chelators are not likely to be of therapeutic value as inhibitors of CP4Hs, as their inhibitory mechanism is not sufficiently selective.[67] Similar molecules are already used clinically to treat conditions such as iron overload[110] or heavy metal toxicity,[111] and thus should be avoided if selective inhibition of an enzyme target is desired. Still, we note that functionalization or modification of many of the scaffolds described above generates more potent inhibitors of the enzyme[67] (*vide infra*). For example, 2,2′-bipyridine-5,5′-dicarboxylic acid (bipy55′DC) is one of the most potent inhibitors of a CP4H recognized to date (K_i = 185 nM[106]) Yet, the residual iron affinity of such compounds remains of concern and complicates the

characterization of the inhibitory mechanism of these and related compounds,[106,112] which is of particular interest herein (see: CHAPTER 2).

AKG Mimics

Prior to the early 1980s, inhibitor design for prolyl hydroxylases was largely empirical, as a chemical mechanism and spacial concept for the active site had yet to be agreed upon. In 1982, Haunauske-Abel and Günzler put forth the first stereochemical concept of the mechanism of prolyl hydroxylases that took into account both theoretical and experimental data.[97] Except for a few modifications, their postulated model is still trusted today, and has been largely confirmed by structural studies of the FAKGD enzymes in general (Figure 1.5).[56] Moreover, their model became the starting point for the most useful developments in the design of inhibitors for P4Hs. This model was particularly relevant for inhibitors that interacted with the AKG binding site, as the model provided a rational explanation for the mode of binding of AKG (*vide supra*), and predicted accurately that analogs of AKG that could not undergo the hypothesized decarboxylation mechanism would serve as potent inhibitors of the enzyme.[97] Moreover, the model also to explained why many of the AKG mimics (*vide infra*) that inhibit CP4Hs do not inhibit the essential enzyme AKG dehydrogenase, which does not require a metal for its catalytic activity.[97,113]

After the model described above was published,[97] a variety of modestly potent inhibitors of the CP4H from chicken were quickly reported (Figure 1.7B) including the pyridinedicarboxylic acids pyridine-2,4-dicarboxylic acid (24PDC) and pyridine 2,5-

dicarboxylic acid (25PDC),[107] 3,4-dihydroxybenzoic acid (DHB), and *N*-oxalylglycine (NOG).[114] Like AKG, all of these compounds posses a metal chelate and a pendant carboxyl group, and thus resemble the structure of AKG in the context of the stereochemical mechanism proposed previously.[97] Moreover, all of these compounds have been found to inhibit the human CP4Hs *in vitro*,[30,115] with similar structure–activity relationships. Whereas 24PDC, 25PDC, and NOG were found to be simple competitive inhibitors with respect to AKG,[107,114] DHB was found to be competitive with both AKG and ascorbate, suggesting that the mechanism of this compound could be more complicated than postulated originally.[116]

Notably, the discovery of each of these simple AKG mimics was a significant advancement in the development of inhibitors of CP4Hs. Still, certain scaffolds appeared more suitable for the design of selective inhibitors of CP4Hs relative to other FAKGDs and were thus investigated further by medicinal chemists. In particular, the 25PDC and NOG scaffolds were of considerable interest, the chemical modification of which will be discussed individually (*vide infra*).

Derivatives of Pyridine-2,5-dicarboxylic Acid

Unlike 24PDC, which was found to inhibit many members of the FAKGD with similar low micromolar potency,[67] 25PDC was found to be highly selective for CP4Hs compared to anti-targets such as the HIF-PHDs. Thus, considerable efforts were made to explore 25PDC derivatives (Figure 1.7B) with the intent of developing a selective antifibrotic drug. In particular, a variety of substitutions were made at both the 2-position[117] and the 5-position[118] of the pyridine

ring. Unfortunately, all of the substitutions for the 2-carboxyl group (*i.e.*, *N*-acyl sulfonamides, tetrazole, hydroxamic acids, 2-imidazolyl, 2-pyrroyl, and others) were made in the absence of the 5-carboxyl group, thereby making it difficult to compare the two series. In general, modification of the 2-position (as described above) in the absence of the 5-position carboxyl group did not yield potent inhibitors.[117] For example, many inhibitors were more potent than the parent pyridine-2-carboxylic acid, but most still inhibited with IC_{50} values in the mid-micromolar range,[117] and have not been developed further.

In contrast to 2-position modifications, modifications of the 5-position carboxyl group were much more fruitful. Substitution of that carboxyl group in many *N*-acyl sulfonamide derivatives yielded 25PDC analogues with similar or improved potency compared to the 25PDC parent.[118] Most of these compounds inhibited with IC_{50} values in the low micromolar range, with phenyl and 4-methoxyphenyl sulfonamides (Figure 1.7B) displaying the highest potency (IC_{50} values of 1.1 μM and 1.0 μM, respectively, compared to 5.5 μM for 25PDC). Yet, considering the binding mode of AKG, it was highly surprising that substitution of the C5-carboxylate with neutral derivatives was tolerated. This finding raised the question of whether or not 25PDC and these analogues bound to the active site in the same manner as did AKG itself. I revisit this idea in CHAPTER 2. Moreover, no clinically-relevant inhibitors based upon these sulfonamide derivatives of 25PDC have been reported to date, suggesting problems in their development.

Derivatives of N-Oxalyl Glycine

While NOG did not display the selectivity properties of 25PDC, and is also known to inhibit AKG dehydrogenase[119] and other enzymes that utilize AKG[120], the NOG scaffold provided a means for researchers to evaluate stereochemistry in the context of the AKG binding pocket. In contrast to the 25PDC scaffold, substitutions to the ω-carboxyl group of the glycine moiety abolished inhibition,[114] suggesting that NOG likely binds in the conventional AKG binding mode. Still, the α carbon of the glycine moiety has a tetrahedral geometry, so substitutions were made to evaluate the nature of the binding pocket on either side of the AKG substrate.[114] Despite substitution of the glycine moiety with a variety of alternative amino acids (*i.e.*, alanine, phenylalanine, leucine, methionine, proline, sarcosine, and others), the only alternative amino-acid substitution that was tolerable was that of alanine, which showed a weaker potency by over an order of magnitude compared to NOG (IC_{50} values of 90.7 μM for racemic *N*-oxalyl alanine (NOA) compared to 2.89 μM for NOG).[114] Although the racemic NOA was of reasonable potency, a striking difference was observed between the *R* and *S* enantiomers, where (*S*)-NOA (Figure 1.7B) was substantially more potent than (*R*)-NOA (IC_{50} values of 38.2 μM and 621 μM, respectively).[114] Whereas the AKG binding pocket of CP4Hs is likely to be constricted, there is a stereochemical preference with respect to out-of-plane substituents.

Combining the Best of Both: Heterocyclic Carbonyl Glycines

Researchers quickly realized that there were advantages to both the 25PDC scaffold and the NOG scaffold that might be combined to yield a more potent and selective inhibitor of CP4Hs.

The heterocyclic carbonyl glycines (Figure 1.7B) were designed to take advantage of the benefits of both of these scaffolds by incorporating both a heterocyclic alkaloid nitrogen in the chelate as well as the pendant glycine for interaction with the active site Lys residue.[77] Heterocyclic carbonyl glycines displayed similar structure–activity relationships to the NOG derivatives, which supported the notion that they still interacted in the AKG-binding pocket.[77] For example, when alanine was substituted for the glycine moiety, there was still a clear preference for the *S*-compared to the *R*-enantiomer.[77] Moreover, it was found that a pendant hydroxyl group attached to the *ortho* position on the heterocycle significantly increased the potency, the mechanism of which was postulated to be a combination of preorganization from an internal hydrogen bond with the adjacent carbonyl group and arguably an additional hydrogen bond within the enzymic active site.[77] Notably, many of these compounds were found to inhibit CP4H *in vitro*, in cultured chicken tendon cells, and even in a rat model of uterine fibrosis, although further development of such compounds were limited due to prohibitive cytotoxicity that was apparently unrelated to inhibition of CP4H activity.[77]

Collagen-Like Peptides

It has long been recognized that P4H inhibitors based upon the primary collagen substrate could be of interest, and might offer improved selectivity compared to that of the AKG mimics (*vide supra*). Whereas many features of the reaction catalyzed by CP4Hs are well characterized, a detailed understanding of the manner in which the protocollagen substrate is recognized by these

enzymes remains incomplete, partially due to the limited structural information available for mammalian CP4Hs.

Although there is still not a clear understanding of the peptide substrate, a few individual examples of peptide-based inhibitors have been reported. For example, CP4Hs are known to be inhibited by polyproline strands in the PP-II conformation, which have been observed to bind in the PSB domain of human CP4H-1.[84] Moreover, some simple analogs of the collagen-repeat sequence in which the composition is modified or rearranged have observed to inhibit the enzyme from chicken. For example, peptides of the sequence $(Gly\text{-}Pro)_n$, $(Gly\text{-}Pro\text{-}Gly)_n$, and $(Pro\text{-}Pro\text{-}Gly\text{-}Gly)_n$ have all been observed to inhibit chicken CP4H, though $(Pro\text{-}Pro\text{-}Gly\text{-}Gly)_n$ has also been observed to be a poor substrate.[103]Moreover, short collagen-like polymers with triplet repeats containing the non-natural Pro derivatives 3- and 4-methyl Pro and 4-fluoro Pro (both as diasteromers and individual enantiomers) serve as inhibitors of chicken CP4H.[121] We also note that a small tetrapeptide containing the sequence Gly-Pro-Flp-Gly was found to be a substrate (and inhibitor based upon its being turned over at a slow rate) for human CP4H-1, furnishing a product containing 4-ketoproline (Kep) in pace of the Flp residue.[122] The Kep was postulated to result from spontaneous collapse of a fluorohydrin that is generated by abstraction of the *pro-R* proton of the 4*S*-Flp residue and replacement with a hydroxyl group,[122] and represents an interesting chemical tool that could possibly be used in a biological context to probe the substrates for CP4Hs in an unbiased manner.

Beyond the peptide-based inhibitors described above, inhibitors based upon the primary collagen substrate are few and far between.[103] We believe that inhibitors based-upon the primary

substrate could be developed, but only after the conformational preferences of the protocollagen substrate are characterized further. We explore this topic in more detail below (see: CHAPTER 5).

Irreversible Inhibitors

Unlike most of the reversible inhibitors discussed above, we note that in a few cases, time-dependent and irreversible inhibition of the enzyme has been observed. These molecules (Figure 1.7C) include certain anthracyclines, peptide substrates containing the Pro derivative 5-oxaproline (Opr), and an electrophilic analogue of the AKG cosubstrate. In the case of the anthracyclines, doxorubicin and daunorubicin were both found to be irreversible inactivators of the CP4H from chicken.[123] They appeared to function via a mechanism that required ferrous iron in solution, but that is competitive with the ascorbate cofactor, with possible alkylation of the enzymic active site via a postulated radical intermediate. Doxorubicin has been used clinically as a chemotherapeutic agent for the treatment of a variety of types of cancer.[124] Moreover, one of the common side effects of doxorubicin treatment is poor wound healing,[125] which is consistent with the notion that one of the many chemotherapeutic effects of doxorubicin is inactivation of CP4Hs (as has been observed in human skin fibroblasts[126]). Unfortunately, doxorubicin in a known intercalator of DNA and has general antiproliferative effects on cells.[124] Thus, despite its clinical utility for some cancers[124] and likely inhibition of collagen synthesis in cells[126] and in *vivo*,[125] replacement of doxorubicin with a more selective inhibitor of CP4Hs would be more a significant improvement for treatment for the fibrotic diseases and metastasis.

Tetrapeptide substrates containing the unnatural Pro derivative Opr (Figure 1.7C) were also found to be irreversible inhibitors of the CP4H from chicken.[127,128] The exact chemical mechanism of this inactivation was never precisely determined, though chemical experiments demonstrated that a peptide containing OPr in the Y-position of the triplet repeat irreversibly inactivated the enzyme in a manner that required ferrous iron and the cosubstrate AKG, but was outcompeted by high concentrations of the peptide substrate $(ProProGly)_{10}$ or ascorbate.[127] Moreover, the tetrapeptide Dansyl-Gly-Phe-OPr-Gly-biotin was found to label the enzyme covalently,[128] although the residue(s) that was modified have not been determined. Peptides containing OPr were even found to inhibit CP4H activity in cultured fibroblasts,[129] though much less effectively compared to that observed *in vitro*,[127] and further investigations into the nature of the inactivation have not been reported.

Finally, an electrophilic analogue of the AKG cosubstrate was also observed to inactivate the CP4H from chicken. This compound was coumalic acid (Figure 1.7C), which resembles AKG in chain length and placement of a pendant carboxyl group, but appears to alkylate the enzyme irreversibly by a postulated nucleophilic attack on the coumarin ring.[130] Moreover, AKG provided a protective affect against the inactivation, suggesting that the alkylation likely occurred in or near the enzymic active site.[130] Unfortunately, coumalic acid itself is unlikely to be of clinical utility, but the observation of this alkylating effect has sparked interest in developing alternative electrophilic molecules as chemical probes of the enzyme (see: CHAPTER 6 and APPENDIX A).

Inhibition of CP4Hs *In Vivo*

A variety of CP4H inhibitors have been tested *in vivo* in the mouse or rat. Few, however, have proven to be sufficiently safe or efficacious to move into clinical studies. As described above, the chemotherapeutic agent doxorubicin[124] (Figure 1.7C) could function, at least in part, by inhibiting CP4Hs,[123] but this compound displays a variety of inhibitory mechanisms[124] and is thus not suitable for most therapeutic applications of interest related to CP4H. A few agents were found to be useful in preclinical studies, the outcomes of which were reviewed recently.[67] Briefly, a prodrug analogue of 24PDC (HOE-077, Figure 1.7D) has been explored in preclinical studies in rats and dogs, and was found to inhibit collagen synthesis in the liver.[67] HOE-077 was, however, found to yield many metabolites, and was found to disrupt collagen synthesis by inactivating cells involved in producing collagen, rather than by direct inhibition of CP4H.[67] Moreover, a 1,10-phenanthroline derivative (FG-0041, Figure 1.7D) was shown to aid in the recovery of left ventricular function in rats that have suffered post myocardial infarction, though this compound probably functions by inhibiting the HIF-PHDs, rather than CP4Hs.[67]

Arguably, the state of the art with respect to inhibiting CP4Hs *in vivo* is ethyl 3,4-dihydroxybenzoate (EDHB, Figure 1.7D). EDHB is the ethyl ester prodrug analogue of DHB, and is commonly used in cell culture as a "P4H inhibitor." Moreover, EDHB has even been shown to be safe enough to use in mouse models of cancer metastasis.[29] A dose of 40 mg/kg per day in mice was found to decrease primary tumor grown, reduce the collagen content of the primary tumor, and significantly reduce metastasis to the lung, all without affecting overall body weight.[29] EDHB, however, is a member of the catechol family, and is also known to perturb iron

metabolism[131] (and likely many other cellular pathways[132]) by serving as an iron chelator[131] and affecting redox homeostasis.[132] Thus, EDHB might not be sufficiently selective for some of the therapeutic applications of interest here.

1.8 Imperative for Inhibitors with Therapeutic Potential

Considering the clear need to treat fibrotic diseases and cancer metastasis, as well as the importance of CP4Hs for both of these pathologies, there is a clear need for inhibitors of CP4Hs with therapeutic potential. Although previously developed inhibitors (*vide supra*) have provided an abundance of information about these enzymes, most fall short of clinical relevance for a variety of reasons, including poor potency in cells, mixed modes of activity, prohibitive toxicity, or other intolerable off-target effects. Unlike the many enzyme activities that can be screened in a high-throughput manner, such screens for prolyl hydroxylation are only in their infancy,[133] and it is not clear how such assays would be adapted to the CP4Hs. We argue that the most fruitful route to CP4H inhibitors with therapeutic potential is likely the re-evaluation of existing inhibitor scaffolds to determine a manner in which these compounds can be made more drug-like (*i.e.*, reduce off-target effects, improve selectivity, maintain, or improve potency). In the chapters that follow, I explore a variety of topics related to developing CP4H inhibitors with therapeutic potential or increasing our understanding of the features of CP4Hs that can be exploited for selective inhibition.

Table 1.1 Examples of antifibrotic drugs and drug candidates[a]

Drug	Target	Type[b]	Pathway / Process	Company (Trade Name)	Development Status
Imatinib Mesylate	Tyrosine kinases	SM	TGF-β signaling	Novartis (Gleevec)	In use for systemic sclerosis
Pirfenidone	Broad anti-inflammatory activity	SM	TGF-β and p38 signaling	Multiple (i.e. Pirespa, Esbriet, Etuary)	In use for idiopathic pulmonary fibrosis
Nintedanib (BIBF 1120)	Tyrosine kinases	SM	Various signaling pathways	Boehringer Ingelheim (Vargatef)	In use for idiopathic pulmonary fibrosis
Fresolimumab (GC1008)	TGF-B	HMA	TGF-β signaling	Sanofi-Aventis	Phase II clinical trials for glomerulosclerosis
LY2382770	TGF-B cytokine	HMA	TGF-β signaling	Lilly	Phase II clinical trials for chronic kidney disease
(STX-100)	$\alpha_v\beta_6$ integrin	HMA	TGF-β signaling	Biogen Idec	Phase II clinical trials for idiopathic pulmonary fibrosis
BMS-986202	LPA receptor-1	SM	Fibroblast migration and apoptosis	Bristol-Myers Squibb	Phase II clinical trials for idiopathic pulmonary fibrosis
Lebrikizumab	Interleukin 13 cytokine	HMA	Innate immune activation	Roche	Phase II clinical trials for asthma
PRM-151	Multiple	Recombinant Pentraxin-2	Immune system modulation	Promedior	Phase II clinical trials for idiopathic pulmonary fibrosis
Simtuzumab (GS-6624)	Lysyl oxidase homolog 2	HMA	Collagen crosslinking and fibrogenic gene expression	Gilead	Multiple clinical trials for fibrotic diseases

[a]Information adapted from refs 32,33.
[b]SM = small-molecule; HMA = human monoclonal antibody.

Figure 1.1 The collagen triple helix

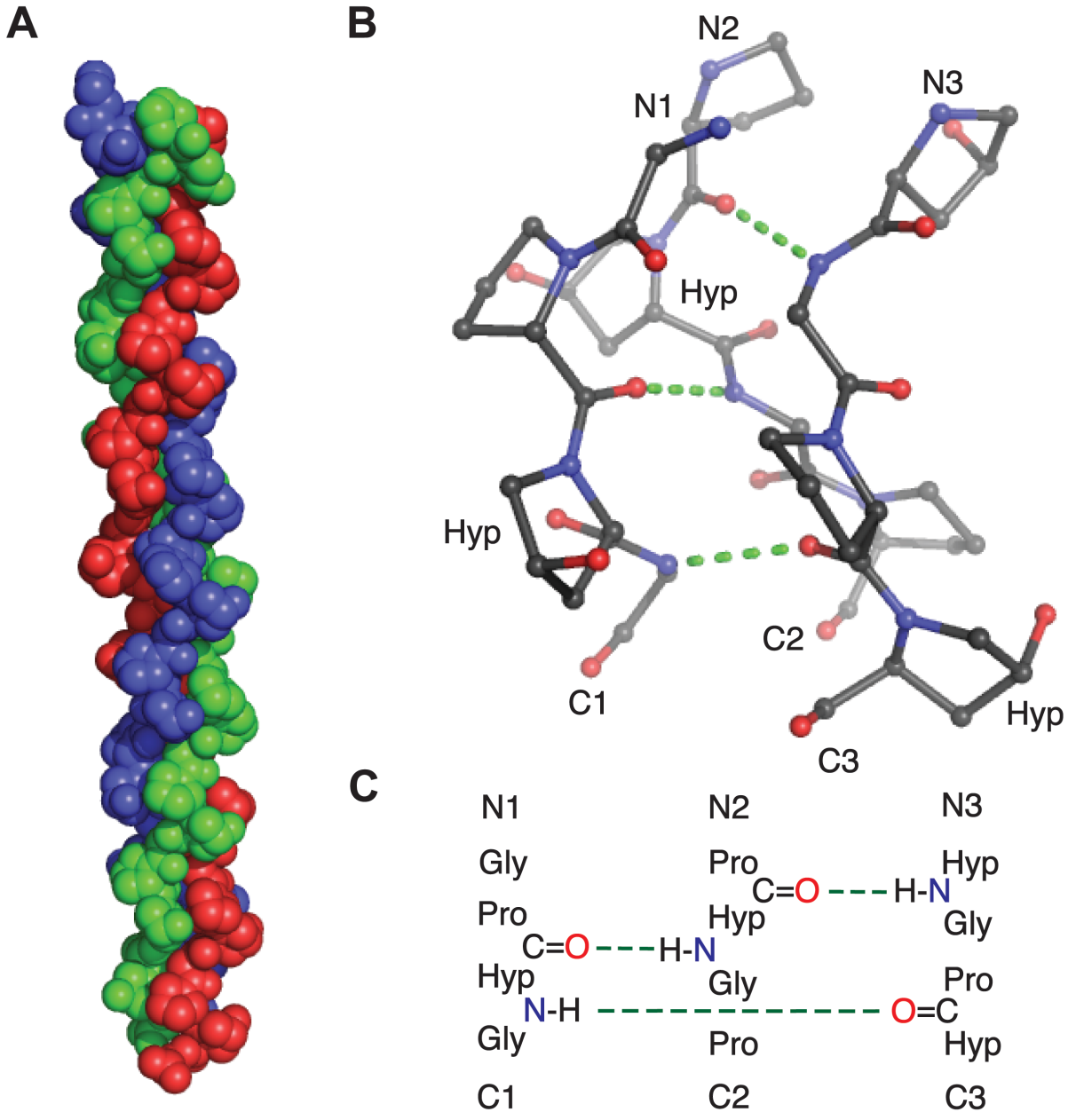


Figure 1.1 The collagen triple helix. (A) Space-filling model $(Pro\text{-}Hyp\text{-}Gly)_n$, a synthetic collagen triple helix. (B) Ball-and-stick model of a short segment from panel A indicating Hyp residues and interstrand hydrogen bonds. (C) Schematic representation showing the register of the three strands in panel B. Coordinates are from PDB entry 1CGD.

Figure 1.2 Structural complexities of proline residues

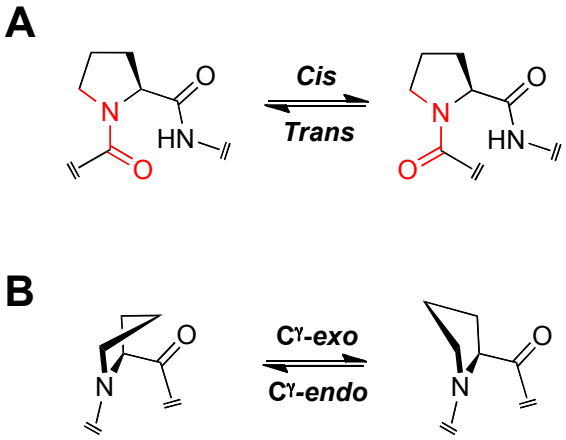


Figure 1.2 Structural complexities of proline residues. (A) Because proline residues are tertiary amides, the *cis* conformation of the prolyl peptide bond occurs more frequently than in non-prolyl residues. $K_{trans/cis} = 4.6$ for Pro residues, while $K_{trans/cis} = 6.1$ for Hyp residues. (B) The pyrrolidine ring of proline residues can exist in one of two ring puckers termed C^{γ}-*exo* and C^{γ}-*endo*. In Pro residues, the ratio of *exo* to *endo* ring puckers is approximately 1:2 in favor of the *endo* pucker.[1] Electron withdrawing substituents at the 4(*R*) position (i.e. –OH and –F) shifts the preference in favor of the *exo* pucker via the gauche effect.[1]

Figure 1.3 The collagen biosynthetic pathway

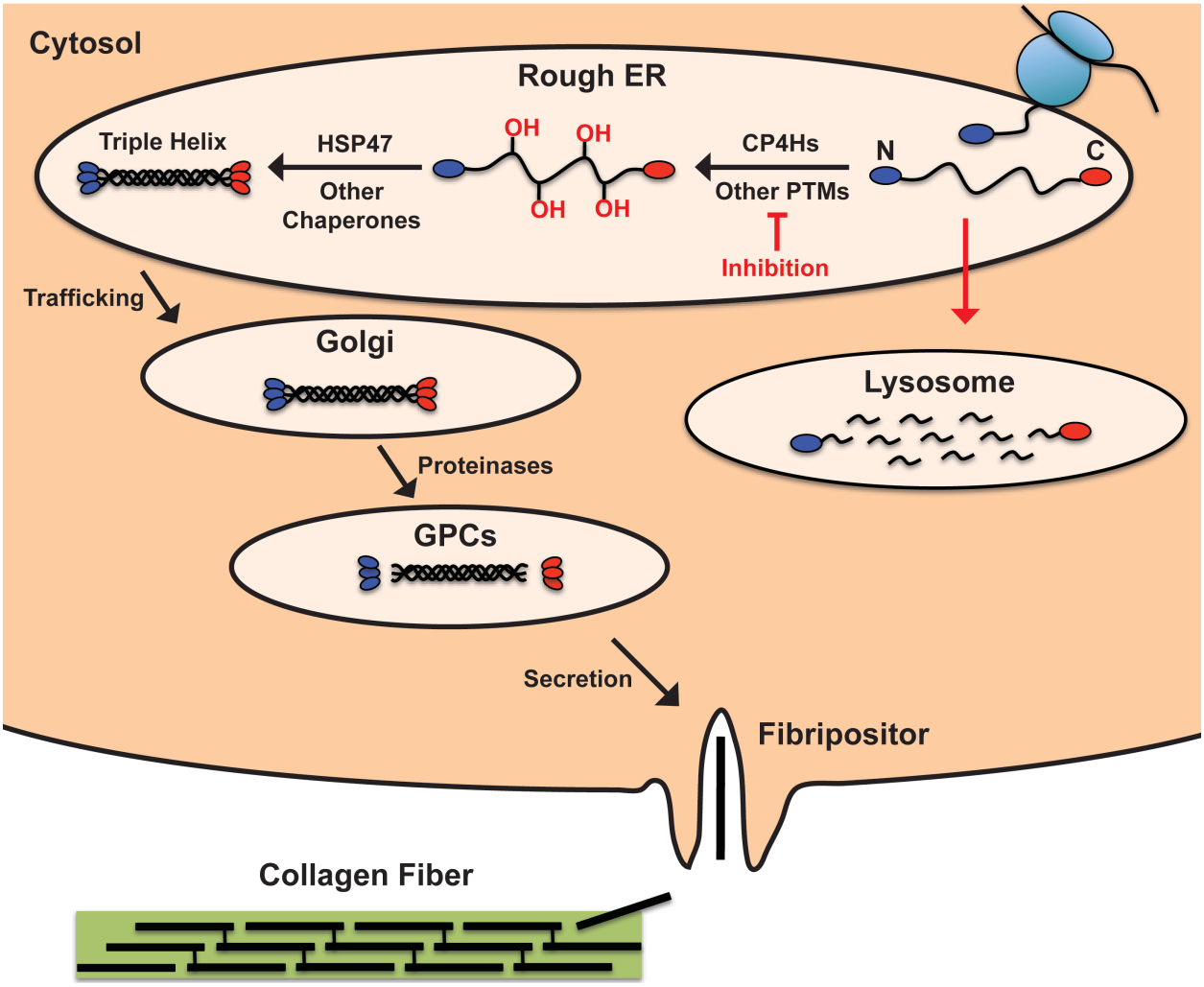


Figure 1.3 The collagen biosynthetic pathway. Collagen synthesis begins in the rough endoplasmic reticulum (rough ER). Protocollagen strands are translated via the ribosome and possess N- and C-terminal propeptides that aid in the folding process. After extensive posttranslational modification including hydroxylation of proline by collagen prolyl 4-hydroxylases (CP4Hs) and prolyl 3-hydroxylases, hydroxylation of lysine residues by lysyl hydroxylases, and *O*-glycosylation, protocollagens assemble into triple helices. Triple helix folding is aided by general ER chaperones as well as collagen-associated chaperones such as HSP47. Procollagen triple helices are trafficked through the golgi and golgi-to-plasma membrane carriers (GPCs), during which time collagen N- and C-terminal proteinases begin cleaving the propeptide domains to allow fibril formation. GPCs are trafficked to plasma membrane protrusions called fibripositors that extrude the fibrils from the cell. Fibrils are incorporated into growing collagen fibers and covalently cross-linked via the action of lysyl oxidases. Disruption of triple helix folding (i.e. via inhibition of CP4Hs) targets protocollagens strands to the lysosome for degradation.

Figure 1.4 Underlying mechanisms for the induction of organ fibrosis

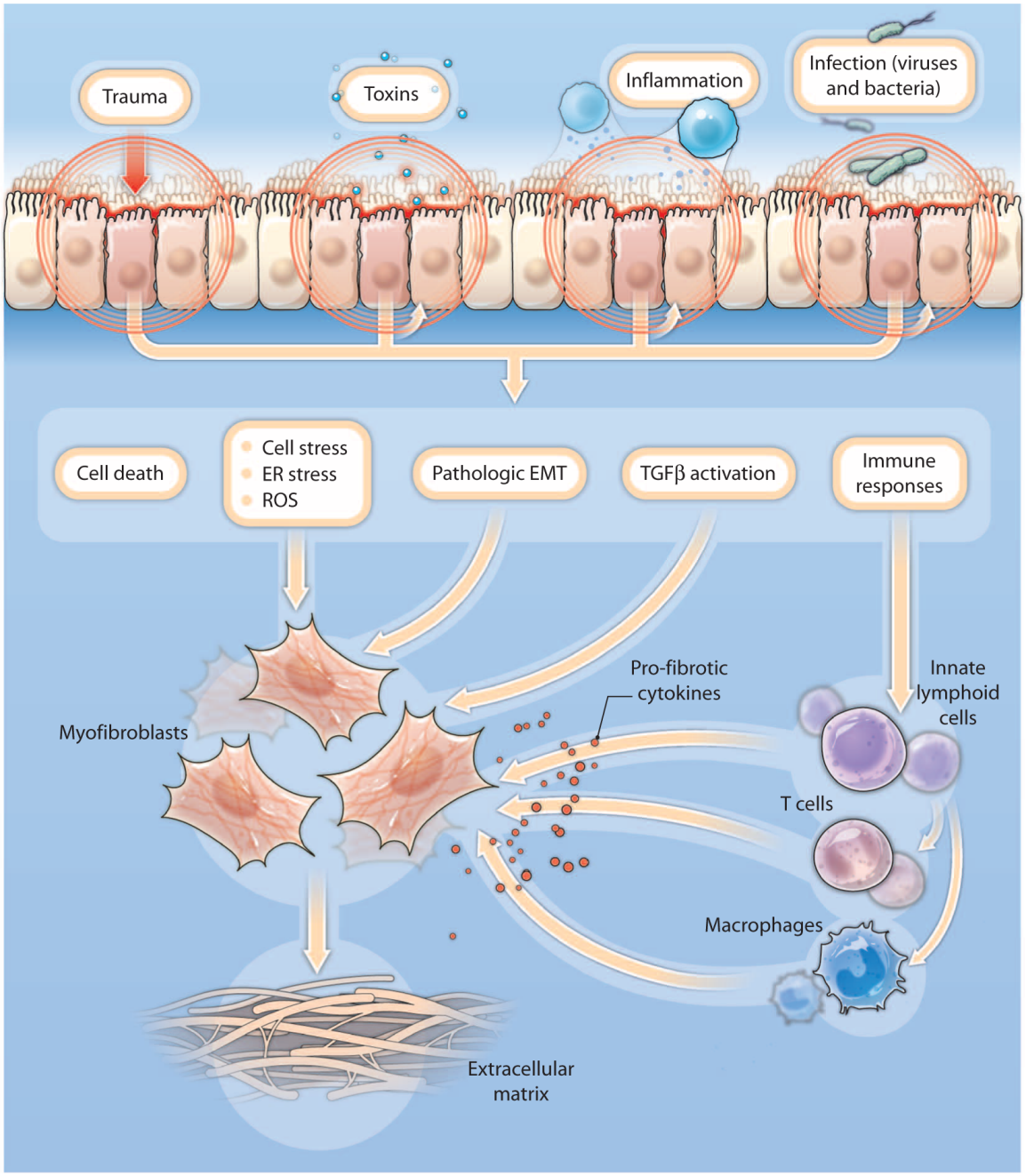


Figure 1.4 Underlying mechanisms for the induction of organ fibrosis (this figure was reproduced from the work of Friedman and co-workers[32]). Many types of injuries to epithelial cells initiate various downstream profibrotic pathways. The metabolism of injured epithelial cells can become dysregulated, leading to ROS production, ER stress, and pathologic EMT. These altered metabolic states activate several downstream pathways such as TGF-β signaling, activation or recruitment of the immune system, and recruitment/stimulation of myofibroblasts, which are the primary source of pathologic ECM produced during fibrosis.

Figure 1.5 Structural homologues of the CP4H C-terminal catalytic domain

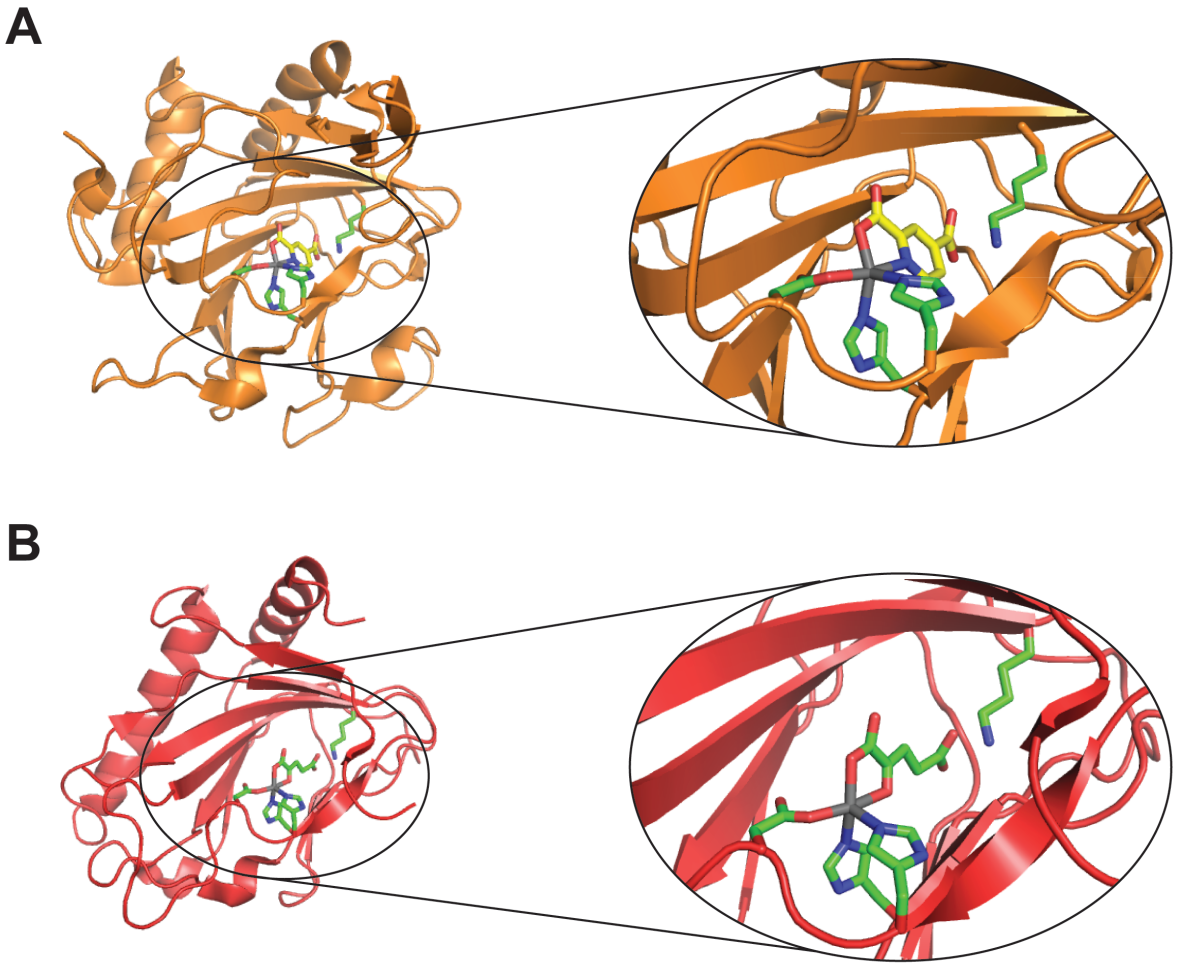


Figure 1.5 Structural homologues of the CP4H C-terminal catalytic domain. (A) Crystal structure (PDB 2JIG) of Cr-P4H-1 in complex with the AKG mimic 2,4-pyridinedicarboxylate (24PDC). A zinc atom (surrogate for iron) is bound to the active site using two His residues and an Asp residue. 24PDC coordinates the zinc atom, using its pyridyl nitrogen and C-2 carboxylate group, with the C-4 carboxylate group engaged in a Coulombic interaction with a Lys residue. (B) Crystal structure (PDB 5C5T) of vP4H in complex with AKG. A zinc atom (surrogate for iron) is bound to the active site using two His residues and an Asp residue. AKG coordinates the zinc atom, using its C-1 carboxylate and C-2 keto groups, with the C-5 carboxylate group engaged in a Coulombic interaction with a Lys residue.

Figure 1.6 The mechanism of the overall reaction catalyzed by CP4Hs

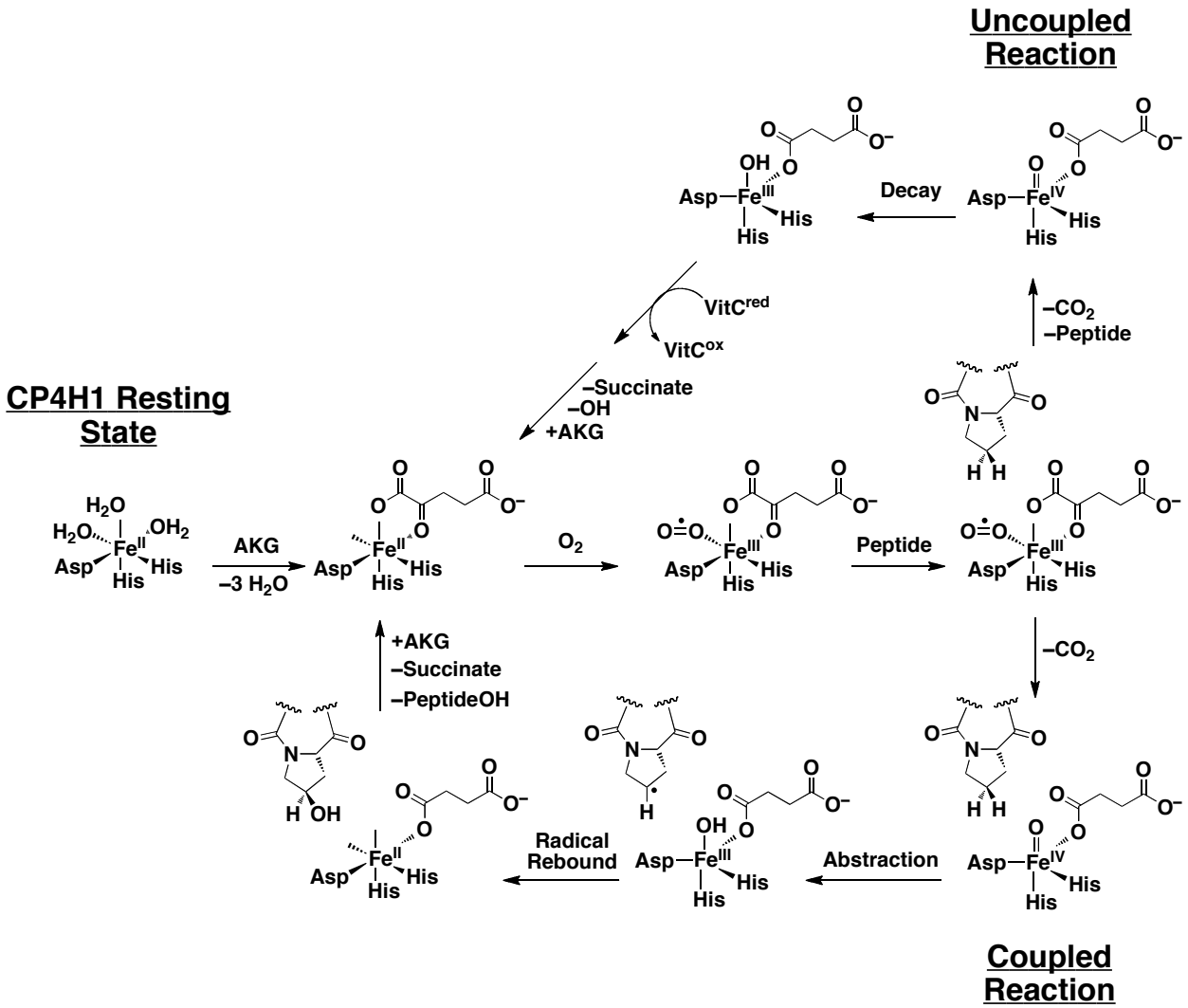


Figure 1.6 The mechanism of the overall reaction catalyzed by CP4Hs. CP4Hs have previously been reported to display an ordered ter-ter mechanism in which AKG first binds to CP4H·Fe(II) complex, after which O_2 and the peptide substrate bind in an ordered fashion.[98] Subsequently, the oxidative decarboxylation of AKG facilitates the formation of a highly reactive Fe(IV)=O species (ferryl ion) that hydroxylates the peptide substrate via a radical rebound process.[56] While all CP4Hs appear to require $VitC^{red}$ for activity, this cofactor is not required during the typical ter-ter mechanism, but rather, has been shown to operate during the "uncoupled reaction" in which CP4H decarboxylates AKG in the absence of the peptide substrate.[102] As suggested by EPR experiments, the ferryl ion appears to decay via an unknown mechanism to an inactive Fe(III) state,[101] and $VitC^{red}$ is required to reduce the iron center back to the Fe(II) state required for catalysis.[101,102] In some cases (particularly for the uncoupled reaction), the ligand environment of the iron center and exact order of binding is not known. Thus, this model is meant to provide a framework for further considerations herein.

Figure 1.7 Examples of known CP4H inhibitors

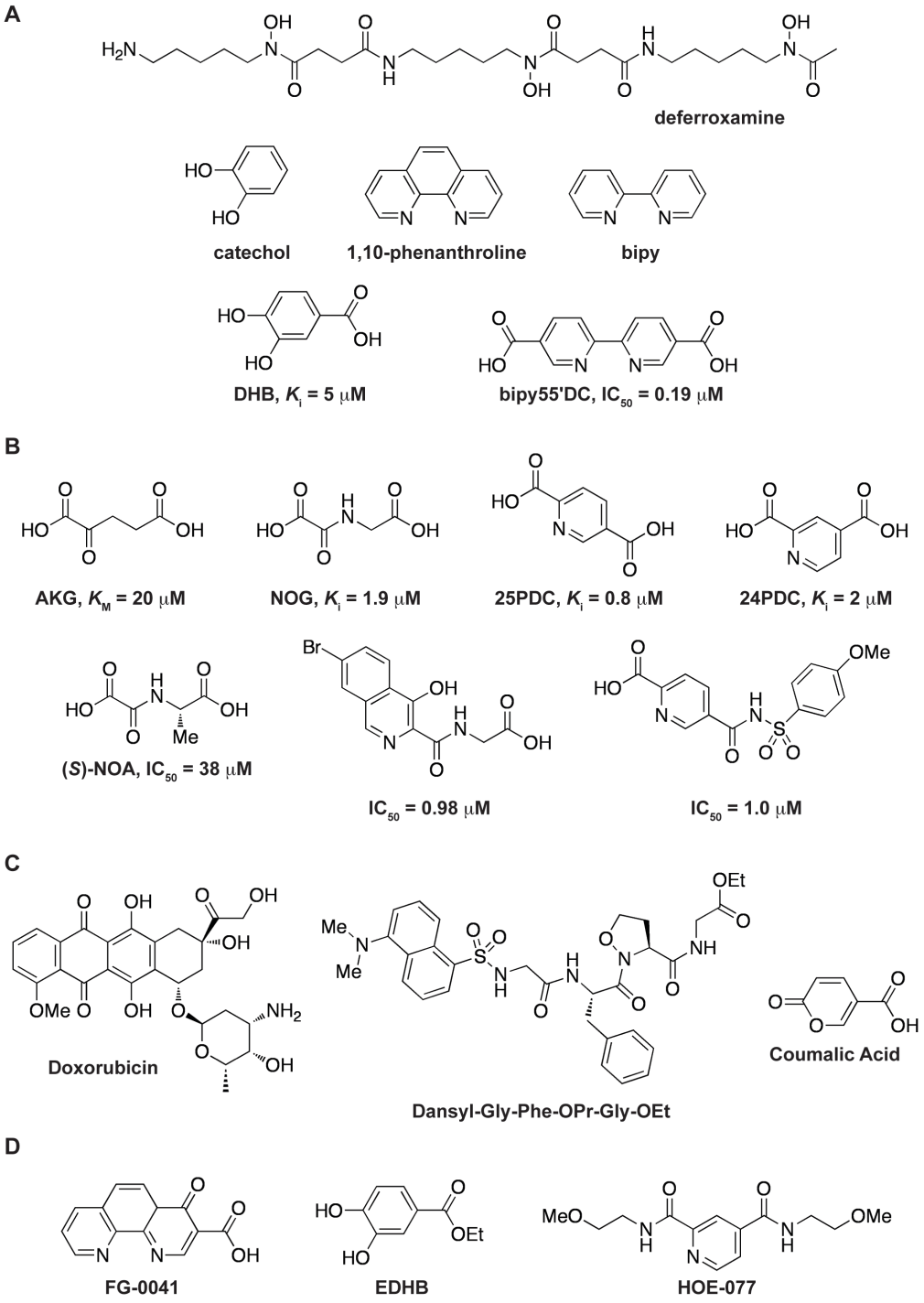
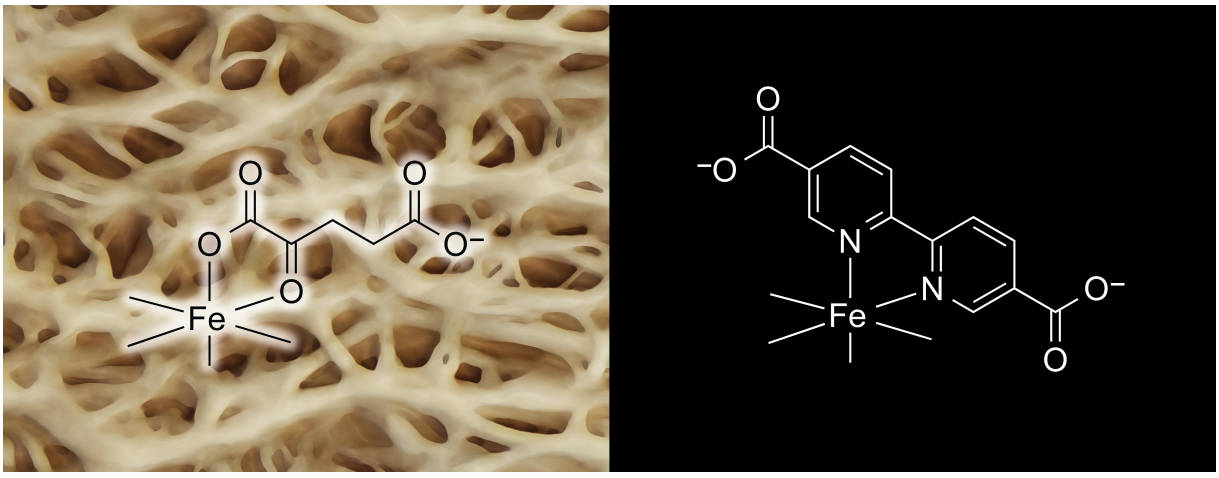


Figure 1.7 Examples of known CP4H inhibitors. In addition to metal ions and some collagen-like peptides, CP4Hs are inhibited by (A) simple metal chelators and their derivatives, (B) simple mimics of the AKG cosubstrate and their derivatives, and (C) a variety of irreversible inhibitors. Only a few compounds (D) have been tested with any success *in vivo*.

CHAPTER 2

Selective Inhibition of Prolyl 4-Hydroxylases by Bipyridinedicarboxylates*



*This chapter has been published in part, under the same title. Reference: Vasta, J.D. & Raines, R.T. Selective inhibition of prolyl 4-hydroylases by bipyridinedicarboxylates. *Biorg. Med. Chem.* All experiments were carried out by J.D.V. Both J.D.V. and R.T.R. contributed to the design of experiments, analysis of data, and preparation of the manuscript and figures.

2.1 Abstract

Collagen is the most abundant protein in animals. A variety of indications are associated with the overproduction of collagen, including fibrotic diseases and cancer metastasis. The stability of collagen relies on the posttranslational modification of proline residues to form (2*S*,4*R*)-4-hydroxyproline. This modification is catalyzed by collagen prolyl 4-hydroxylases (CP4Hs), which are Fe(II)- and α-ketoglutarate (AKG)-dependent dioxygenases located in the lumen of the endoplasmic reticulum. Human CP4Hs are validated targets for treatment of both fibrotic diseases and metastatic breast cancer. Herein, we report on 2,2′-bipyridinedicarboxylates as inhibitors of a human CP4H. Although most 2,2′-bipyridinedicarboxylates are capable of inhibition via iron sequestration, the 4,5′- and 5,5′-dicarboxylates were found to be potent competitive inhibitors of CP4H, and the 5,5′-dicarboxylate was selective in its inhibitory activity. Our findings clarify a strategy for developing CP4H inhibitors of clinical utility.

2.2 Introduction

Collagen is the principle component of bone, connective tissues, and the extracellular matrix in animals.[1] A variety of diseases are associated with the overproduction of collagen, including fibrotic diseases (such as pulmonary fibrosis, renal fibrosis, and hepatic cirrhosis)[32,33] and cancers.[28,29,46,47,73] The stability of collagen relies on posttranslational modifications that occur throughout the secretory pathway.[2] The most prevalent of these modifications is the hydroxylation of collagen strands by collagen prolyl 4-hydroxylases (CP4Hs), which are Fe(II)- and α-ketoglutarate (AKG)-dependent dioxygenases (FAKGDs) located in the lumen of the endoplasmic reticulum.[4] Catalysis by CP4Hs converts (2*S*)-proline (Pro) residues in protocollagen strands into (2*S*,4*R*)-4-hydroxyproline (Hyp) residues (Figure 2.1), which are essential for the conformational stability of mature collagen triple helices.[5] Importantly, CP4Hs are validated targets for treating both fibrotic diseases[30] and metastatic breast cancer,[29] which in aggregate represent a massive unmet clinical need.

Like all enzymes of the FAKGD superfamily, catalysis by CP4Hs requires Fe(II), which is bound by a conserved His-X-Asp/Glu…X_n…His motif, as well as the cosubstrates AKG and dioxygen.[56] AKG chelates to enzyme-bound Fe(II) using its C-1 carboxylate and C-2 keto groups, while the C-5 carboxylate group engages in Coulombic interactions with a basic residue (typically arginine or lysine) and additional hydrogen bonds. All FAKGDs are believed to effect catalysis through a similar two-stage mechanism in which AKG is first oxidatively decarboxylated to generate a highly reactive Fe(IV)=O species (ferryl ion), after which the ferryl

ion interacts with a hydrocarbon substrate and effects hydroxylation via a radical rebound process.[56]

In vertebrates, CP4Hs are known to exist as $\alpha_2\beta_2$ tetramers. In these tetramers, the α-subunit contains the catalytic and substrate-binding domains, and the β-subunit is protein disulfide isomerase, which is a multifunctional protein that is responsible for maintaining the α-subunit in a soluble and active conformation.[4] Three isoforms of the α-subunit, α(I), α(II), and α(III), have been identified in humans.[4] All α-subunit isoforms form tetramers with the β-subunit, which we refer to herein as the CP4H1, CP4H2, and CP4H3 holoenzymes. As the most prevalent of the isoforms, the α(I)-subunit has been characterized extensively. Whereas the structure of the tetrameric complex is unknown, those of individual domains of the α(I)-subunit have provided insight into the manner in which CP4Hs interact with the protocollagen substrate, as well as the means by which the α-subunits dimerize to facilitate formation of the tetramer.[84-86]

The development of CP4H inhibitors has been of keen interest since the mid 1970s. Like many FAKGDs, human CP4Hs are inhibited by simple metal chelators, such as 2,2′-bipyridine (bipy), as well as AKG mimics (Figure 2.2), such as *N*-oxalyl glycine (NOG), pyridine-2,4-dicarboxylic acid (24PDC) and pyridine-2,5-dicarboxylic acid (25PDC).[67] In general, simple metal chelators are thought to inhibit CP4Hs by iron sequestration, as evidenced by IC_{50}-values that are similar to the iron concentration used in assays *in vitro*. In contrast, AKG mimics are thought to bind competitively in the AKG binding site, using a chelating moiety to interact with the enzymic iron and an additional carboxyl group to form favorable Coulombic interactions with Lys493.[67] Compounds of this nature typically display values of IC_{50} or K_i that are less than

the value of K_m for AKG (which is 20 μM for human CP4H1).[98] There are also examples of electrophilic AKG mimics such as coumalic acid[130] and 4-oxo-5,6-epoxyhexanoic acid[134] that appear to inhibit CP4H via covalent modification of the active site. Compounds of this nature could have intriguing chemical and biological utility.

Of the simple AKG mimics, the most potent inhibitor of human CP4Hs reported to date is 25PDC, which was found to have a K_i = 0.8 μM and display modest selectivity for human CP4Hs compared to other FAKGDs, such as the prolyl hydroxylase domain-containing proteins (PHDs), factor-inhibiting HIF (FIH), and Jumonji domain-containing 2E (JMJD2E).[67,107] Thus, much work has focused on modification of the 25PDC scaffold with the intent of generating a cell-active inhibitor that is both potent and selective. Some examples include the 5-position *N*-acylsulfonamide derivatives[118] and 2-heterocyclic glycinamide derivatives[77] of 25PDC (Figure 2.2), though these compounds were not developed further due to insufficient activity in cultured cells or intolerable cytotoxicity.

The most potent inhibitor of human CP4Hs identified to date is 2,2′-bipyridine-5,5′-dicarboxylic acid (bipy55′DC, Figure 2.3).[106] Whereas the parent compound 2,2′-bipyridine-5-carboxylic acid (bipy5C) was found to be a modest inhibitor (IC_{50} = 13 μM), bipy55′DC was reported to be substantially more potent (K_i = 185 nM).[106] The inhibitory mechanism of this compound is, however, not known. Bipy itself is well known to inhibit human CP4Hs via iron sequestration rather than enzymic binding. Although structure–activity relationships (SARs) suggest that enzymic binding contributes to inhibition by bipy55′DC,[106] the role of iron sequestration is unclear.

Herein, we investigate 2,2′-bipyridinedicarboxylates as inhibitors of human FAKGDs, both in terms of the relevant inhibitory mechanism(s) and optimum geometry for inhibition. We access a library of seven 2,2′-bipyridinedicarboxylates and assess the ability of each to sequester iron and bind to a CP4H. We find that both 2,2´-bipyridine-4,5´-dicarboxylate (bipy45′DC) and bipy55′DC are nanomolar inhibitors of a CP4H, though both also bind to free Fe(II). Significantly, we discover that bipy45′DC, but not bipy55′DC, binds to a human PHD. These data provide a roadmap for developing potent, selective inhibitors of human CP4Hs.

2.3 Results and Discussion

2,2′-Bipyridinedicarboxylate library

We accessed a focused library of 2,2′-bipyridinedicarboxylates (bipyDCs) (Figure 2.3). The symmetric members of the library, 2,2′-bipyridine-3,3′-dicarboxylic acid (bipy33′DC), 2,2′-bipyridine-4,4′-dicarboxylic acid (bipy44′DC), bipy55′DC, and 2,2´-bipyridine-6,6′-dicarboxylic acid (bipy66′DC), were obtained from commercial vendors. The asymmetric members were synthesized by a 4-step route in which the key 2,2′-bipyridine intermediates were accessed by the palladium-catalyzed 2-pyridylation of the corresponding pyridine *N*-oxides (Scheme 2.1).[135-137] This route afforded bipy45′DC, 2,2′-bipyridine-4,6′-dicarboxylic acid (bipy46′DC), and 2,2′-bipyridine-5,6′-dicarboxylic acid (bipy56′DC) in overall yields of 53%, 36%, and 15%, respectively.

Iron sequestration

To begin our analysis, we evaluated the bipyDC as iron chelators. We performed titration experiments to determine the half-maximal concentration required to form a complex with 20 μM Fe(II) (Fe_{20}-EC_{50}) at pH 7.0, taking advantage of the strong absorption of complexes of bipy and related analogues with Fe(II) (Table 2.1 and Figure 2.4). Along with an estimation of stoichiometry obtained with Job's method (Table 2.1 and Figure 2.5), the Fe_{20}-EC_{50} value provides a comparative metric for iron affinity. With the exception of bipy33′DC and bipy66′DC, all members of the library were capable of forming distinct red complexes with Fe_{20}-EC_{50} values in the micromolar range (Table 2.1). Because bipy is already known to inhibit CP4Hs, these results suggest that most of these ligands would inhibit CP4H1 by sequestering iron under iron-limiting conditions.

CP4H binding

To identify compounds that inhibit human CP4H1 through enzymic binding rather than iron sequestration, we next sought to determine the extent to which bipy and bipyDCs form a complex with Fe(II) under CP4H assay conditions. Typical *in vitro* assays for human CP4Hs employ a tris(hydroxymethyl)aminomethane (Tris) buffer. Although Tris is known to form complexes with metal ions,[138,139] the use of Tris does not affect the observed hydroxylase activity of purified enzyme.[98] Still, to determine if bipy and bipyDCs can form iron complexes under our assay conditions, we incubated our CP4H assay buffer (50 mM Tris–HCl buffer, pH 7.8, containing 1% w/v bovine serum albumin (BSA), 100 μg/mL catalase, 2 mM sodium ascorbate,

100 μM AKG, 100 μM dithiothreitol (DTT), and 50 μM $FeSO_4$) at 30 °C in the presence or absence of bipy (150 μM), and determined the presence of a $Fe(bipy)_3^{2+}$ complex using spectrophotometry. Under these conditions, we observed the formation of the $Fe(bipy)_3^{2+}$ complex rapidly and in virtually identical abundance to that observed in unbuffered conditions. Moreover, similar results were observed for complexes with bipy44′DC, bipy45′DC, and bipy45′DC ligands. In order of addition experiments, we found that Tris alone inhibits the formation of the $Fe(bipy)_3^{2+}$, but that the subsequent addition of either sodium ascorbate or DTT allowed for complex formation.

Given these results and the requirement of ascorbate for hydroxylase activity, the design of assay conditions that preclude the formation of Fe(II) complexes with bipy and related analogues is highly unlikely. Thus, we next sought to develop *in vitro* screening conditions for human CP4Hs such that the inhibitory effect of iron sequestration would be minimal. We chose an initial screening concentration of 10 μM, which is substantially below the concentration of $FeSO_4$ (50 μM) used in the assay. Importantly, bipy showed virtually no inhibition under these screening conditions whereas 25PDC showed significant inhibition (Figure 2.6A), which validates these conditions for the discovery of compounds where the primary inhibitory mechanism is other than merely iron sequestration. Although most of the compounds screened under these conditions showed little to no inhibition, both bipy55′DC and bipy45′DC were found to be potent inhibitors (more than 90% reduction of CP4H activity) and bipy44′DC was found to be a modest inhibitor that is similar in potency to 25PDC (Figure 2.6A). In subsequent dose–response experiments, the inhibition curves for bipy55′DC and bipy45′DC were found to be

sigmoidal (Figure 2.6B) with IC_{50} values in the low micromolar range (Table 2.2). Yet, the inhibition curve for bipy44′DC was found to be non-sigmoidal (Figure 2.7), which suggests a mixed inhibitory mechanism wherein iron sequestration becomes a contributing factor at higher concentrations (supported by the observation of a red color in the assay solutions). These data suggest that CP4H1 is inhibited strongly by 2,2′-bipyridinedicarboxylates of two different geometries with almost equal potency, and that the inhibition does not rely upon iron sequestration.

To understand further the potent inhibition observed for bipy55′DC and bipy45′DC, we next probed the kinetic mechanism of these compounds. Human CP4Hs have been shown previously to display an ordered ter–ter mechanism in which AKG first binds to the CP4H·Fe(II) complex, after which O_2 and the peptide substrate bind in an ordered fashion.[98] Lineweaver–Burke analyses confirmed that bipy55′DC and bipy45′DC both function as competitive inhibitors with respect to AKG (Figure 2.8), which suggests that they bind to CP4H1 in a manner that requires enzymic iron. Based upon the observed competitive inhibition, we used the Cheng–Prussoff equation to estimate values of K_i for these compounds (Table 2.2). We found that bipy55′DC and bipy45′DC inhibit human CP4H1 with K_i values in the nanomolar range. These K_i values should be taken as upper limits, as it is difficult to take into account the fraction of each that is present in complexes with free iron under assay conditions. Although the competitive inhibition observed here should be interpreted with caution in the absence of structural information, these data are consistent with a mechanism in which bipy55′DC and bipy45′DC inhibit human CP4H1 by binding in the AKG binding pocket. In addition, we tested the biphenyl

analogs of bipy55′DC and bipy45′DC and found that the pyridyl nitrogens were essential for detectable inhibition (data not shown), providing further support for a mechanism in which coordination of the enzymic iron is a requirement.

It is interesting to note the SARs observed above for the 2,2′-bipyridinedicarboxylates parallel that of the pyridinedicarboxylates. For example, human CP4H1 is inhibited by both 24PDC and 25PDC, where the geometry of 25PDC is preferred compared to that of 24PDC.[67] Similarly, both bipy55′DC and bipy45′DC (reminiscent of 25PDC and 24PDC, respectively) are accommodated in the active site, where the geometry of bipy55′DC appears to be preferred compared to that of bipy45′DC. Yet, it is also important to note that bipy55′DC and bipy45′DC are substantially more potent than the corresponding pyridinedicarboxylates. For example, comparing 25PDC and bipy55′DC and assuming energetically similar interactions with the assay buffer, the difference in K_i corresponds to a difference in binding free energy of $\Delta G° = (1.7 \pm 0.1)$ kcal/mol, which is typical for a hydrogen bond. These comparisons suggest that bipy55′DC and bipy45′DC likely bind to human CP4H1 in a manner similar to the corresponding pyridinedicarboxylates, except that their binding is stabilized further by additional enzymic interactions with the second carboxylate group.

Inhibition of PHD2

We reasoned that the second carboxyl group of bipy45′DC and bipy55′DC could allow for the selective inhibition of CP4Hs over other FAKGDs. This issue is, however, difficult to address in the absence of structural information about the catalytic domain of any human or mammalian

CP4H. Although the structure of the N-terminal peptide binding domain of the α-subunit has been characterized with both X-ray crystallography[84,86] and NMR spectroscopy,[85] the structure of the catalytic domain has remained elusive.

Unlike with CP4H, there are numerous crystal structures available for PHD2, including some solved as a complex with AKG or AKG mimics. Moreover, preliminary SARs for PHD2 with respect to simple AKG mimics such as 24PDC, 25PDC, and NOG have already been established, facilitating a comparison between the SARs of CP4H1 and PHD2.[67] For example, NOG and 24PDC inhibit PHD2 at low micromolar concentrations, but 25PDC does not inhibit PHD2 even at high micromolar concentrations. These and other data suggest differences in the active sites of these essential FAKGDs that could be understood further using a combination of both chemical approaches and existing structural information.

Toward this end, we screened our library at 10 μM as inhibitors of PHD2. We found that, whereas bipy45′DC (and to a lesser extent bipy44′DC) inhibited PHD2, most other 2,2′-bipyridinedicarboxylates including bipy55′DC showed virtually no inhibition (Figure 2.9). Moreover, the inhibition observed for bipy45′DC was not greater than that observed for simple AKG mimics, such as NOG and 24PDC. These data are consistent with the hypothesis that bipy45′DC is accommodated in the PHD2 active site, but without the engagement of its 5′-carboxyl group in a favorable enzymic interaction. These data are also in agreement with examinations of available crystal structures of PHD2 (PDB entries 3ouj, 3ouh, 2hqu, and 2g1m), wherein the active site cleft opposite the AKG binding site is exposed to solvent. Using all extant

data, we put forth schematic models of 2,2′-bipyridinedicarboxylates with CP4H1 and PHD2 (Figure 2.10).

Other Isoforms of CP4H

We note that our investigations have focused on the CP4H1 holoenzyme rather than the CP4H2 and CP4H3 isoforms. This choice was largely due to the relative abundance of CP4H1 in most tissues compared to that of the other two isoforms.[4] Nonetheless, BLAST experiments suggest high similarity between the three α-subunits, with the α(I)-subunit 78 and 56% similar to the α(II)- and α(III)-subunits, respectively. Moreover, these similarities improve to >85 and >63% when considering only the catalytic domains. Given these high similarities, the SARs identified for CP4H1 (as well as the model put forth in Figure 2.10A) could also pertain to CP4H2 and CP4H3.

2.4 Conclusions

The need for selective inhibitors of human CP4Hs with therapeutic potential is evident from the growing list of diseases associated with the overproduction of collagen. The lack of structural information available for mammalian CP4Hs has made the development of such inhibitors difficult, and the understanding of known inhibitors such as compounds 25PDC and bipy55′DC likewise evasive. Our investigations of the 2,2′-bipyridinedicarboxylate family reveal new information. For example, 25PDC was known to be a selective inhibitor of human CP4H

compared to other FAKGDs, such as the PHDs, FIH, and JMJD2E.[67] These results, along with additional structure–activity relationships of 25PDC analogues,[118] led researchers to postulate that 25PDC could bind to CP4H in a manner that is distinct from the conventional AKG binding mode, possibly taking advantage of alternative coordination sites on the enzymic iron center. Our finding that bipy45′DC and bipy55′DC are competitive inhibitors of human CP4H1 with almost equal potency contrasts with the previous postulate and suggests that the AKG-binding pocket of this enzyme is likely able to accommodate two different inhibitor geometries (reminiscent of the geometries of either 24PDC or 25PDC) in the conventional AKG binding mode. Still, the exact structural and/or conformational features that allow for this accommodation are not yet understood.

Although achieving selectivity for CP4Hs over other FAKGDs is ideal, it is especially important to achieve selectivity over PHD2. For example, in hypoxic breast cancers, the expression of the CP4H α(I)-and α(II)-subunits is upregulated in a PHD2-dependent manner, which ultimately contributes to the metastatic potential of these tumors.[29] More specifically, the loss in PHD2 activity due to hypoxia allows the accumulation of the PHD2 substrate and transcription factor HIF-1α, which ultimately enhances the expression of these metastatically important CP4H α-subunits.[29] Although treatment with a CP4H inhibitor could be a viable strategy to disrupt breast cancer metastasis, that inhibitor must avoid inhibition of PHD2 so as to prevent a consequent increase in the target CP4Hs. Indeed, the poor performance of many previously characterized CP4H inhibitors in cells could be due, at least in part, to the off-target effects associated with inhibiting PHD2. Our finding that both bipy45′DC and bipy55′DC are

substantially more potent towards human CP4H1 compared to simple AKG mimics, as well as the finding that this increased potency does not occur when inhibiting PHD2, suggests the existence of additional CP4H1 residues that can be used to achieve the desired selectivity. Although the identity of these residues is unclear, previous investigations of the bipy55′DC scaffold showed that additional functional groups such as anilides and sulfonamides are tolerated as replacements of the 5′-carboxylate,[106,118] suggesting that the interaction(s) in this second site might not necessarily be Coulombic in nature.

Lastly, we note that our data reveal that both compound bipy45′DC and compound bipy55′DC are potent inhibitors of CP4H1 and that inhibition arises largely through enzymic binding. Still, bipy45′DC and bipy55′DC do bind to free Fe(II), which could compromise therapeutic efficacy. Our future work will seek to modify these scaffolds to limit Fe(II) binding while maintaining selective affinity for CP4Hs.

2.5 Experimental Section

2.5.1 General Experimental Procedures

Compounds bipy, 24PDC, 25PDC, bipy33′DC, and bipy55′DC were from Sigma–Aldrich (St. Louis, MO). Compound bipy44′DC was from TCI America (Portland, OR). Compound bipy66′DC was from CMS Chemicals (Compton, UK). Biphenyl-4,4′-dicarboxylic acid was from Acros Organics (Geel, Belgium), and biphenyl-3,4′-dicarboxylic acid was from Combi-Blocks (San Diego, CA). Phosphine ligands and phosphonium salts were from either Sigma–

Aldrich or Strem (Newberryport, MA), stored in a dessicator, and used without further purification. $Pd(OAc)_2$ was from Sigma–Aldrich, stored in a dessicator, and used without further purification. HIF-1α $peptide_{556–574}$ was from AnaSpec (Fremont, CA) and used without further purification. All other reagent chemicals were from Sigma–Aldrich, Acros Organics, Combi-Blocks, Oakwood Products (West Columbia, SC), Enamine (Monmouth Junction, NJ), Bachem (Bubendorf, Switzerland), or Calbiochem–Novabiochem (San Diego, CA), and were used without further purification.

All glassware was flame- or oven-dried, and reactions were performed under N_2(g) unless indicated otherwise. Dichloromethane and toluene were dried over a column of alumina. Dimethylformamide was dried over alumina and further purified through an isocyanate scrubbing column. Other anhydrous solvents were obtained in septum-sealed bottles. Flash chromatography was performed with columns of 40–63 Å silica gel, 230–400 mesh from Silicycle (Québec City, Canada). Thin-layer chromatography (TLC) was performed on plates 250-μm silica 60-F_{254} from EMD Merck (Darmstadt, Germany) with visualization by UV light irradiation or staining with $KMnO_4$. The phrase “concentrated under reduced pressure” refers to the removal of solvents and other volatile materials using a rotary evaporator at water aspirator pressure (<20 torr) while maintaining water-bath temperature below 40 °C. Residual solvent was removed from samples at high vacuum (<0.1 torr). The term “high vacuum” refers to vacuum achieved by a mechanical belt-drive oil pump. All reported yields are unoptimized. The purity of all synthesized compounds was assessed by HPLC, and the purity of all final compounds was found to be greater than 95%.

2.5.2 Instrumentation

Melting points were determined with a PMA160 digital melting apparatus from Stanford Research Systems (Sunnyvale, CA). NMR spectra were acquired at ambient temperature with a DMX-400 Avance spectrometer or a Avance 500i spectrometer from Bruker (Billerica, MA) at the National Magnetic Resonance Facility at Madison (NMRFAM) and were referenced to TMS or a residual protic solvent. Some compounds exist as either mixtures of rotamers or tautomers that do not interconvert on the NMR timescale at ambient temperature and therefore exhibit multiple sets of NMR signals (as indicated). Electrospray ionization (ESI) and electron ionization (EI) mass spectrometry were performed with a Micromass LCT® and Micromass AutoSpec® instruments, respectively, from Waters (Milford, MA) at the Mass Spectrometry Facility in the Department of Chemistry at the University of Wisconsin–Madison. The progress of reactions catalyzed by prolyl 4-hydroxylases was determined by analytical HPLC using a system from Waters equipped with a 996 photodiode array detector and Empower 2 software. Compound purity was also determined using the Waters HPLC system above. Preparative HPLC was performed using a Prominence HPLC instrument from Shimadzu (Kyoto, Japan) equipped with two LC-20AP pumps, SPD-M20A photodiode array detector, and CTO-20A column oven. Iron complexes with bipyridine ligands were analyzed by spectrophotometry using a Cary 60 UV–Vis spectrometer from Agilent Technologies (Santa Clara, CA). Protein concentrations were calculated from their absorbance at 280 nm as measured with a NanoVue Plus spectrophotometer from GE Healthcare using an extinction coefficient of 290,000 $M^{-1}cm^{-1}$ for human CP4H[140] and

36,900 $M^{-1}cm^{-1}$ for human PHD2.[141] IC_{50}- and EC_{50}-values were calculated from experimental data with Prism version 6.0 software from GraphPad Software (La Jolla, CA).

2.5.3 Production of Recombinant Human CP4H1

Human CP4H containing the α(I) isoform was produced heterologously in Origami B(DE3) *Escherichia coli* cells and purified as described previously.[140]

2.5.4 Assay of Human CP4H1 Activity in the Presence of Inhibitors

The catalytic activity of human CP4H1 was assayed as described previously.[140] Briefly, activity assays were carried out at 30 °C in 100 μL of Tris–HCl buffer, pH 7.8, containing human CP4H1 (100 nM), inhibitor (0–500 μM), substrate (dansylGlyProProGlyOEt, 500 μM), $FeSO_4$ (50 μM), BSA (1 mg/mL), catalase (0.1 mg/mL), ascorbate (2 mM), DTT (100 μM), and α-ketoglutarate (100 μM). Reaction mixtures were pre-incubated with or without inhibitor for 2 min at 30 °C, after which the reaction was initiated by the addition of α-ketoglutarate. After 15 min, reactions were quenched by boiling for 45 s and centrifuged at 10,000*g*. The supernatant (20–50 μL) was injected into a Nucleodur® C18 Gravity reversed-phase column (4.6 × 250 mm, 5-μm particle size) from Macherey–Nagel (Bethlehem, PA). The column was eluted at 1 mL/min with a gradient of aqueous acetonitrile (20–45% over 20 min) containing 0.1% v/v TFA. The absorbance of the eluent was monitored at 289 nm. All assays were performed in triplicate. Data is reported as activity relative to control reactions lacking inhibitor, where activity is determined

from the percent conversion of substrate to product. Dose-response curves were generated for each inhibitor by plotting the relative activity versus the log of the inhibitor concentration. IC_{50}-values for each inhibitor were interpolated from the dose-response curves by non-linear regression using the sigmoidal dose-response function within the Prism software.

2.5.5 Production of Human PHD2

A cDNA encoding human $PHD2_{181-426}$ possessing an N-terminal hexahistidine (His_6) tag ($NHis_6$-$PHD2_{181-426}$) was cloned using the Gibson strategy,[142] and the encoded protein was produced and purified as described previously.[141]

2.5.6 Assay of Human PHD2 Activity in the Presence of Inhibitors

The catalytic activity of human PHD2 in the absence or presence of putative inhibitors was assayed as described previously.[141] Briefly, activity assays were carried out at 30 °C in 100 μL of Tris–HCl buffer, pH 7.8, containing human $NHis_6$-$PHD2_{181–426}$ (5 μM), inhibitor (0–10 μM), substrate (HIF-1α $peptide_{556–574}$, 50 μM), $FeSO_4$ (50 μM), BSA (1 mg/mL), catalase (0.3 mg/mL), ascorbate (2 mM), DTT (1 mM), and α-ketoglutarate (35 μM). Reactions were pre-incubated with or without inhibitor for 2 min at 30 °C, after which the reaction was initiated by the addition of α-ketoglutarate. After 10 min, reactions were quenched by boiling for 60 s and centrifugation at 10,000*g*. The supernatant (50 μL) was injected into a Nucleodur® C18 Gravity reversed-phase column (4.6 × 250 mm, 5 μm particle size) from Macherey–Nagel. The column

was eluted at 1 mL/min with a gradient of aqueous acetonitrile (5–56% over 34 min) containing 0.1% v/v TFA. The absorbance of the eluent was monitored at 218 nm. All assays were performed in triplicate. Data is reported as activity relative to control reactions lacking inhibitor, where activity is determined from the percent conversion of substrate to product.

2.5.7 Assay of Fe(II)-Affinity of Bipyridine Compounds

The affinity of bipyridine compounds for Fe(II) was determined comparatively by measuring the half maximal concentration (EC_{50}) required for binding 20 μM Fe(II) (Fe_{20}-EC_{50}) in sodium phosphate buffer, pH 7.0. Stock solutions of ligands were prepared in water. Stock solutions of $FeSO_4$ were prepared in H_2O and used within 3 h of preparation. Ligand solutions (3–500 μM depending on affinity) were prepared in 10 mM sodium phosphate buffer, pH 7.0, after which Fe(II) stock solution was added to initiate complex formation. Solutions were allowed to equilibrate for 15 min, after which the absorbance was recorded at the λ_{max} for the complex under study. Absorbance values were corrected by subtracting the absorbance value in the absence of ligand. Dose–response curves were generated for each ligand by plotting the absorbance versus the log of the ligand concentration. Fe_{20}-EC_{50} values for each ligand were interpolated from the dose-response curves by non-linear regression using the sigmoidal dose-response function in Prism. All experiments were performed in triplicate.

2.5.8 Determination of Fe(II) Complex Stoichiometry

The stoichiometry of complexes formed by bipyridine ligands and Fe(II) was estimated with Job's method.[143,144] Briefly, solutions were prepared such that the concentration of ligand and Fe(II) were kept at 0.4 mM, but the mole fraction of the ligand was varied from 0 to 1. Stock solutions of ligands were prepared in water. Stock solutions of $FeSO_4$ were prepared in water and used within 3 h of preparation. Mixtures were prepared in 10 mM sodium phosphate buffer, pH 7.0, and complex formation was initiated by the addition of the Fe(II) solution. Solutions were allowed to equilibrate for 15 min, after which the absorbance was recorded at the appropriate λ_{max} for the complex under study. Absorbance values were corrected by subtracting the absorbance value of a blank solution in the absence of $FeSO_4$, after which the values were normalized relative to the mixture with the highest absorbance value. All experiments were performed at least in duplicate. Job plots were constructed by plotting the normalized absorbance versus the mole fraction of bipyridine ligand, after which the stoichiometry of the complex was estimated from the value on the abscissa of the point of intersection for the two best-fit lines approaching the point of maximum absorbance.

2.5.9 Determination of Compound Purity

All synthesized compounds were analyzed by HPLC to determine their purity. The analyses were performed on the Waters HPLC system using a Nucleodur® C18 Gravity reversed-phase column (4.6 × 250 mm, 5-μm particle size) from Macherey–Nagel. Samples (50 μL) dissolved in H_2O

were injected into the column and eluted at 1 mL/min with a gradient of aqueous acetonitrile (5–95% v/v over 20 min) containing 0.1% v/v TFA. The maximal absorbance in the range of 210–400 nm was used as the detection wavelength.

2.5.10 Synthesis

3-Methoxycarbonylpyridine *N*-oxide

Methyl nicotinate (2.0 g, 14.6 mmoles) and *m*-chloroperoxybenzoic acid (6.5 g, 29.2 mmoles) were dissolved in dry dichloromethane (150 mL) in a flame-dried flask. This reaction mixture was stirred at room temperature for 24 h. The solution was concentrated under reduced pressure, and the resulting residue was purified by chromatography on silica (5–20% MeOH in acetone) to afford the title compound as a pale yellow solid (2.19 g, 98%) with mp 94.6–99.1 °C. **^{1}H NMR (400 MHz, $CDCl_3$, δ):** 8.77 (s, 1 H), 8.34 (d, J = 6.8 Hz, 1 H), 7.86 (d, J = 8.0 Hz, 1 H), 7.38 (dd, J = 6.8, 8.0 Hz, 1 H), 3.98 (s, 3 H); **^{13}C NMR (100 MHz, $CDCl_3$, δ):** 163.2, 142.5, 140.3, 129.9, 126.2, 125.8, 53.1; **HRMS** (ESI) m/z 154.0494 [calc'd for $C_7H_8NO_3$ $(M + H)^+$ 154.0499].

4-Methoxycarbonylpyridine *N*-Oxide

Methyl isonicotinate (1.8 mL, 15.2 mmoles) and *m*-chloroperoxybenzoic acid (6.8 g, 30.4 mmoles) were dissolved in dry dichloromethane (150 mL) in a flame-dried flask. This reaction mixture was stirred at room temperature for 24 h. The solution was concentrated under reduced pressure, and the resulting residue was purified by chromatography on silica (1–4% v/v MeOH

in acetone) to afford the title compound as a white solid (2.26 g, 97%) with mp 118.1–121.3 °C. **^{1}H NMR (400 MHz, $CDCl_3$, δ):** 8.20 (d, J = 4.0 Hz, 2 H), 7.84 (d, J = 4.0 Hz, 2 H), 3.90 (s, 3 H); **^{13}C NMR (100 MHz, $CDCl_3$, δ):** 163.7, 139.4, 126.5, 126.4, 52.8; **HRMS** (ESI) m/z 154.0492 [calc'd for $C_7H_8NO_3$ $(M + H)^+$ 154.0499].

5-Methoxycarbonyl-2-(4-methoxycarbonylpyridin-2-yl)pyridine *N*-oxide

$Pd(OAc)_2$ (36 mg, 0.16 mmoles), $[P(t\text{-}Bu)_3H]BF_4$ (142 mg, 0.49 mmoles), K_2CO_3 (903 mg, 6.5 mmoles), methyl 2-bromoisonicotinate (705 mg, 3.3 mmoles), and 3-methoxycarbonylpyridine *N*-oxide (2000 mg, 13 mmoles) were added to a dried flask. The flask was fitted with a reflux condenser capped with a septum, evacuated, and purged with N_2(g) (~5 times). Dry toluene (22 mL) was added via syringe, and the reaction mixture was stirred at 110 °C for 18 h. The reaction mixture was cooled to room temperature and filtered through Celite®, and the filtrate was concentrated under reduced pressure. The crude product was then purified by chromatography on silica (15% acetone in 1:1 DCM/hexanes) to afford the title compound (537 mg, 57%) as a pale orange solid with mp 170.6–174.1 °C. **^{1}H NMR (400 MHz, $CDCl_3$, δ)**: 9.49 (s, 1 H), 8.89 (s, 1 H), 8.88 (d, J = 5.2 Hz, 1 H), 8.30 (d, J = 8.4 Hz, 1 H), 7.94 (dd, J = 1.2, 5.2 Hz, 1 H), 7.90 (dd, J = 1.2, 8.4 Hz, 1 H), 3.98 (s, 3 H), 3.97 (s, 3 H); **^{13}C NMR (100 MHz, $CDCl_3$, δ):** 165.3, 163.4, 150.2, 149.9, 149.4, 141.9, 138.0, 128.9, 127.7, 125.7, 124.9, 124.1, 53.1, 52.8; **HRMS** (ESI) m/z 289.0826 [calc'd for $C_{14}H_{13}N_2O_5$ $(M + H)^+$ 289.0819].

4-Methoxycarbonyl-2-(6-methoxycarbonylpyridin-2-yl)pyridine *N*-oxide

$Pd(OAc)_2$ (15.6 mg, 0.070 mmoles), $[P(t\text{-}Bu)_3H]BF_4$ (60.4 mg, 0.21 mmoles), K_2CO_3 (384 mg, 2.8 mmoles), methyl 2-bromopicolinate (300 mg, 1.4 mmoles), and 4-methoxycarbonylpyridine *N*-oxide (858 mg, 5.6 mmoles) were added to a dried flask. The flask was fitted with a reflux condenser capped with a septum, evacuated, and purged with N_2(g) (~5 times). Dry toluene (9.0 mL) was added via syringe, and the reaction mixture was stirred at 110 °C for 24 h. The reaction mixture was cooled to room temperature and filtered through Celite®, and the filtrate was concentrated under reduced pressure. The crude product was purified by chromatography on silica (1% v/v MeOH in EtOAc) to afford the title compound (266 mg, 66%) as a pale yellow solid with mp 193.3–197.0 °C. **^{1}H NMR (400 MHz, $CDCl_3$, δ):** 8.88 (d, J = 8.0 Hz, 1 H), 8.74 (d, J = 2.4 Hz, 1 H), 8.31 (d, J = 6.8 Hz, 1 H), 8.18, (d, J = 7.6 Hz, 1 H), 7.99 (t, J = 7.6 Hz, 1 H), 7.87 (dd, J = 2.4, 6.8 Hz, 1 H); **^{13}C NMR (100 MHz, $CDCl_3$, δ):** 165.3, 163.4, 150.2, 149.9, 149.4, 141.9, 138.0, 128.9, 127.7, 125.7, 124.9, 124.1, 53.1, 52.8; **HRMS** (ESI) *m*/*z* 289.0818 [calc'd for $C_{14}H_{13}N_2O_5$ $(M + H)^+$ 289.0819].

5-Methoxycarbonyl-2-(6-methoxycarbonylpyridin-2-yl)pyridine *N*-oxide

$Pd(OAc)_2$ (10.1 mg, 0.045 mmoles), $[P(t\text{-}Bu)_3H]BF_4$ (39.2 mg, 0.14 mmoles), K_2CO_3 (248 mg, 1.8 mmoles), methyl 2-bromopicolinate (195 mg, 0.9 mmoles), and 3-methoxycarbonylpyridine *N*-oxide (553 mg, 3.6 mmoles) were added to a dried flask. The flask was fitted with a reflux condenser capped with a septum, evacuated, and purged with N_2(g) (~5 times). Dry toluene (6.0 mL) was added via syringe, and the reaction mixture was stirred at 110 °C for 24 h. The reaction

mixture was cooled to room temperature and filtered through Celite®, and the filtrate was concentrated under reduced pressure. The crude product was then purified by chromatography on silica (EtOAc) to afford the title compound (135 mg, 52%) as a pale orange solid with mp 159.7–162.0 °C. **^{1}H NMR (400 MHz, $CDCl_3$, δ):** 9.10 (d, J = 8.0 Hz, 1 H), 8.82 (d, J = 0.8 Hz, 1 H), 8.34 (d, J = 8.4 Hz, 1 H), 8.13 (d, J = 7.6 Hz, 1 H), 7.96 (t, J = 8.0 Hz, 1 H), 7.86 (dd, J = 1.2, 8.4 Hz, 1 H), 3.97 (s, 3 H), 3.93 (s, 3 H); **^{13}C NMR (100 MHz, $CDCl_3$, δ):** 165.1, 163.3, 149.2, 149.0, 147.8, 141.7, 137.5, 128.9, 128.6, 128.0, 125.9, 125.9, 53.0, 52.9; **HRMS** (ESI) *m*/*z* 289.0815 [calc'd for $C_{14}H_{13}N_2O_5$ $(M + H)^+$ 289.0819].

Dimethyl 2,2′-bipyridine-4,5′-dicarboxylate

4-Methoxycarbonyl-2-(5-methoxycarbonylpyridin-2-yl)pyridine *N*-oxide (75 mg, 0.26 mmoles) was dissolved in dry $CHCl_3$ (8.7 mL), and PCl_3 (30 μL, 0.31 mmoles) was added to the resulting solution. The ensuing reaction mixture was stirred at 60 °C until the starting material was consumed completely, as judged by TLC (6 h). The reaction mixture was quenched by the dropwise addition of saturated aqueous Na_2CO_3 (10 mL) while stirring on ice. The product was extracted with $CHCl_3$ (4 × 10 mL), and the combined organics were dried over Na_2SO_4(s) and concentrated under reduced pressure to afford the title compound (69 mg, 98%) as a pale yellow solid with mp 163.1–165.2 °C. **^{1}H NMR (400 MHz, $CDCl_3$, δ):** 9.30 (d, J = 1.6 Hz, 1 H), 9.00 (s, 1 H), 8.85 (d, J = 5.2 Hz, 1 H), 8.52 (d, J = 8.0 Hz, 1 H), 8.43 (dd, J = 2.0, 8.4 Hz, 1 H), 7.92 (dd, J = 1.6, 5.2 Hz, 1 H), 4.01 (s, 3 H), 3.99 (s, 1 H); **^{13}C NMR (100 MHz, $CDCl_3$, δ):** 165.7,

165.5, 158.6, 156.2, 150.6, 150.1, 138.6, 138.1, 126.0, 123.5, 121.1, 120.6, 52.8, 52.5; **HRMS** (ESI) *m*/*z* 273.0882 [calc'd for $C_{14}H_{13}N_2O_4$ $(M + H)^+$ 273.0870].

Dimethyl 2,2′-bipyridine-4,6′-dicarboxylate

4-Methoxycarbonyl-2-(6-methoxycarbonylpyridin-2-yl)pyridine *N*-oxide (200 mg, 0.69 mmoles) was dissolved in dry $CHCl_3$ (23 mL). PCl_3 (76 μL, 0.83 mmoles) was added to the resulting solution. The ensuing reaction mixture was stirred at 60 °C until the starting material was consumed completely, as judged by TLC. The reaction mixture was quenched by the dropwise addition of saturated aqueous Na_2CO_3 (25 mL) while stirring on ice. The product was extracted with $CHCl_3$ (4 × 25 mL), and the combined organics were dried over Na_2SO_4(s) and concentrated under reduced pressure to afford the title compound (145 mg, 77%) as a pale yellow solid with mp 182.5–185.0 °C. **^{1}H NMR (400 MHz, $CDCl_3$, δ):** 9.03 (s, 1 H), 8.86 (d, *J* = 5.2 Hz, 1 H), 8.63 (d, *J* = 7.6 Hz, 1 H), 8.18 (d, *J* = 7.6 Hz, 1 H), 8.01 (t, *J* = 8.0 Hz, 1 H), 7.91 (d, *J* = 8.0 Hz, 1 H), 4.05 (s, 3 H), 4.01 (s, 3 H); **^{13}C NMR (100 MHz, $CDCl_3$, δ):** 165.6, 165.3, 155.9, 149.3, 147.9, 139.4, 138.2, 125.6, 124.7, 123.6, 121.3, 52.9, 52.9; **HRMS** (ESI) *m*/*z* 273.0863 [calc'd for $C_{14}H_{13}N_2O_4$ $(M + H)^+$ 273.0870].

Dimethyl 2,2′-bipyridine-5,6′-dicarboxylate

5-Methoxycarbonyl-2-(6-methoxycarbonylpyridin-2-yl)pyridine *N*-oxide (127 mg, 0.44 mmoles) was dissolved in dry $CHCl_3$ (14.7 mL), after which PCl_3 (50 μL, 0.53 mmoles) was added. The

reaction mixture was stirred at 60 °C until the starting material was consumed completely, as judged by TLC. The reaction was quenched by the dropwise addition of saturated aqueous Na_2CO_3 (15 mL) while stirring on ice. The product was extracted with $CHCl_3$ (5 × 15 mL), and the combined organics were dried over Na_2SO_4(s) and concentrated under reduced pressure to afford the title compound (61 mg, 53%) as a pale yellow solid with mp 170.9–172.3 °C. **^{1}H NMR (400 MHz, $CDCl_3$, δ):** 9.25 (d, J = 2.0 Hz, 1 H), 8.66 (d, J = 7.6 Hz, 1 H), 8.62 (d, J = 7.6 Hz, 1 H), 8.41 (dd, J = 2.0, 8.4 Hz, 1 H), 8.16 (d, J = 8.0 Hz, 1 H), 7.98 (t, J = 8.0 Hz, 1 H), 4.03 (s, 3 H), 3.97 (s, 3 H); **^{13}C NMR (100 MHz, $CDCl_3$, δ):** 165.6, 165.6, 158.4, 155.2, 150.4, 147.7, 138.1, 138.1, 126.1, 125.7, 125.0, 121.1, 52.9, 52.4; **HRMS** (ESI) *m/z* 273.0858 [calc'd for $C_{14}H_{13}N_2O_4$ $(M + H)^+$ 273.0870].

2,2′-Bipyridine-4,5′-dicarboxylic Acid (bipy45′DC)

Dimethyl 2,2′-bipyridine-4,5′-dicarboxylate (482 mg, 1.8 mmoles) and KOH (460 mg, 7.1 mmoles) were added to a vial. MeOH (9.0 mL) was added to the vial, and the resulting reaction mixture was heated to 60 °C until the starting material was consumed completely, as judged by TLC. The reaction mixture was cooled to room temperature and concentrated under reduced pressure. The crude product was dissolved in water (20 mL). The aqueous layer was washed with EtOAc (1 × 20 mL), after which the product was precipitated from the aqueous layer by adjusting to pH 3–4 with 1M HCl. After cooling to 4 °C, the product was filtered, washed with water (3 × 5 mL), and dried in vacuo to afford bipy45′DC (412 mg, 95%) as an off-white solid with mp >260 °C. **^{1}H NMR (400 MHz, DMSO-d_6, δ):** 13.67 (bs, 2 H), 9.20 (s, 1 H), 8.92 (d, J =

4.8 Hz, 1 H), 8.88 (s, 1 H), 8.54 (d, J = 8.0 Hz, 1 H), 8.44 (dd, J = 2.0, 8.4 Hz, 1 H), 7.93 (d, J = 4.8 Hz, 1 H); **^{13}C NMR (125 MHz, DMSO-d_6, δ):** 166.5, 166.5, 158.0, 155.8 151.2, 150.8, 140.0, 139.0, 127.3, 124.2, 120.9, 120.7; **HRMS** (ESI) m/z 243.0405 [calc'd for $C_{12}H_7N_2O_4$ (M – H)$^-$ 243.0411].

2,2′-Bipyridine-4,6′-dicarboxylic Acid (bipy46′DC)

Dimethyl 2,2′-bipyridine-4,6′-dicarboxylate (75 mg, 0.28 mmoles) and KOH (80 mg, 1.12 mmoles) were added to a vial. MeOH (8.0 mL) was added to the vial, and the resulting reaction mixture was heated to 60 °C until the starting material was consumed completely, as judged by TLC. The reaction mixture was cooled and concentrated under reduced pressure, after which the crude product was dissolved in water (2 mL). The aqueous layer was washed with EtOAc (1 × 2 mL), after which the product was precipitated from the aqueous layer by adjusting to pH 3–4 with 1 M HCl. After cooling to 4 °C, the product was filtered, washed with water (3 × 1 mL), and dried in vacuo to afford bipy46′DC (49 mg, 73%) as an off-white solid with mp >260 °C. **^{1}H NMR (400 MHz, DMSO-d_6, δ):** 13.66 (bs, 2 H), 8.94 (s, 1 H), 8.90 (d, J = 5.2 Hz, 1 H), 8.62 (dd, J = 1.6, 7.6 Hz, 1 H), 8.17 (t, J = 7.6 Hz, 1 H), 8.13 (dd, J = 1.6, 7.6 Hz, 1 H), 7.92 (dd, J = 1.6, 4.8 Hz, 1 H); **^{13}C NMR (100 MHz, DMSO-d_6, δ):** 166.6, 166.4, 156.1, 155.0, 151.0, 148.7, 140.0, 139.4, 125.8, 124.1, 123.9, 120.4; **HRMS** (ESI) m/z 243.0404 [calc'd for $C_{12}H_7N_2O_4$ (M – H)$^-$ 243.0411].

2,2′-Bipyridine-5,6′-dicarboxylic acid (bipy56′DC)

Dimethyl 2,2′-bipyridine-5,6′-dicarboxylate (39 mg, 0.14 mmoles) and KOH (40 mg, 0.57 mmoles) were added to a vial. MeOH (4.0 mL) was added to the vial, and the resulting reaction mixture was heated to 60 °C until the starting material was consumed completely, as judged by TLC. The reaction mixture was cooled and concentrated under reduced pressure, after which the crude product was dissolved in water (2 mL). The aqueous layer was washed with EtOAc (1 × 2 mL), after which the product was precipitated from the aqueous layer by adjusting to pH 3–4 with 1M HCl. After cooling to 4 °C, the product was filtered, washed with water (3 × 1 mL), and dried in vacuo to afford bipy56′DC (20 mg, 57%) as an off-white solid with mp >260 °C. **^{1}H NMR (400 MHz, DMSO-d_6, δ):** 13.47 (bs, 2 H), 9.18 (d, J = 1.6 Hz, 1 H), 8.68–8.64 (m, 2 H), 8.47 (dd, J = 2.0 , 8.0 Hz, 1 H), 8.20–8.14 (m, 2 H); **^{13}C NMR (100 MHz, DMSO-d_6, δ):** 166.5, 166.3, 158.0, 154.6, 150.6, 148.7, 139.5, 138.8, 127.3, 126.0, 124.7, 121.3; **HRMS** (ESI) *m*/*z* 243.0414 [calc'd for $C_{12}H_7N_2O_4$ $(M - H)^-$ 243.0411].

***N*-(Methoxyoxalyl)glycine Ethyl Ester**

HGlyOEt·HCl (333 mg, 2.4 mmoles) and DMAP (29.1 mg, 0.24 mmoles) were added to a round-bottom flask. The flask was capped with a septum and purged with N_2(g) (~5 times). DCM (6 mL) was added, and the reaction mixture was cooled in an ice bath. DIEA (750 μL, 4.3 mmoles) and monomethyl oxalyl chloride (200 μL, 2.2 mmoles) were added via syringe, and the reaction was stirred and allowed to come to room temperature. After 6 h, the reaction mixture was quenched on ice by the dropwise addition of a saturated aqueous solution of ammonium

chloride (10 mL). The organic layer was collected, and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic extracts were dried over Na_2SO_4(s), and concentrated under reduced pressure to afford a crude yellow oil. The crude product was then purified by chromatography on silica (70% v/v EtOAc in hexanes) to afford the title compound (416 mg, 92%) as a white solid with mp 58.5–62.5 °C. **^{1}H NMR (400 MHz, $CDCl_3$, δ):** 7.65 (bs, 1H), 4.23 (q, J = 7.2 Hz, 2 H), 4.11 (d, J = 5.6 Hz, 2 H), 3.90 (s, 3 H), 1.28 (t, J = 7.2 Hz, 3 H); **^{13}C NMR (100 MHz, $CDCl_3$, δ):** 168.6, 160.4, 156.4, 61.9, 53.7, 41.5, 14.1; **HRMS** (ESI) *m*/*z* 212.0535 [calc'd for $C_7H_{11}NO_5Na$ $(M + Na)^+$ 212.0530].

***N*-Oxalylglycine (NOG)**

N-(Methoxyoxalyl)glycine ethyl ester (100 mg, 0.53 mmoles) and KOH (88 mg, 1.6 mmoles) were added to a vial. MeOH (5.0 mL) was added to the vial, and the resulting reaction mixture was heated to 60 °C until the starting material was consumed completely, as judged by TLC. The reaction mixture was cooled to room temperature and concentrated under reduced pressure. The crude product was dissolved in water (2 mL), and the pH of the resulting solution was adjusted to 3–4 with 1 M HCl. The aqueous solution was concentrated under reduced pressure, and the residue was dissolved in boiling EtOAc. The solution was cooled, filtered, and concentrated under reduced pressure to a viscous oil. The oil was dissolved in a minimal amount of DCM, diluted with hexanes, and concentrated under reduced pressure. This procedure was repeated 3 times, after which the product was dried under high vacuum to afford NOG (77 mg, 99%) as a white solid with mp 104.5–110.5 °C. **^{1}H NMR (500 MHz, DMSO-d_6, δ):** 14.03 (bs, 1 H), 12.76

(bs, 1 H), 9.04 (t, *J* = 6.0 Hz, 1 H), 3.79 (d, *J* = 6.5 Hz, 2 H); **^{13}C NMR (125 MHz, DMSO-d_6, δ):** 170.8, 162.1, 158.9, 41.4; **HRMS** (ESI) *m*/*z* 165.0506 [calc'd for $C_4H_9N_2O_2$ $(M + NH_4)^+$ 165.0506].

***N*-Benzyloxycarbonyl-(2*S*)-prolyl-(2*S*)-prolylglycine Ethyl Ester (CbzProProGlyOEt)**

N-Benzyloxycarbonyl-(2*S*)-prolyl-(2*S*)-proline (CbzProProOH, 967 mg, 2.8 mmoles), HGlyOEt·HCl (428 mg, 3.1 mmoles), EDC·HCl (590 mg, 3.1 mmoles), and HOBT (415 mg, 3.1 mmoles) were added to a round-bottom flask. The flask was capped with a septum and purged with N_2(g) (~5 times). DMF (28 mL) and DIEA (1.9 mL, 11.2 mmoles) were added via syringe, and the reaction mixture was stirred at room temperature until the reaction was judged to be complete by TLC (approximately 48 h). The reaction mixture was concentrated under reduced pressure, and the resulting residue dissolved in EtOAc (50 mL). The organic layer was washed with 5% w/v $KHCO_3$ (2 × 25 mL), 5% w/v $KHSO_4$ (2 × 25 mL), and brine (1 × 25 mL). The combined organics were dried over Na_2SO_4(s) and concentrated under reduced pressure to a crude yellow oil. The crude product was then purified by chromatography on silica (0–2% v/v MeOH in EtOAc) to yield the title compound (755 mg, 63%) as a pale yellow oil. **^{1}H NMR (400 MHz, $CDCl_3$, mixture of 3 or more rotamers, integrations are approximate, δ):** 8.29 (t, *J* = 6.0 Hz, 0.7 H), 7.40–7.20 (m, 15.4 H + $CHCl_3$), 7.15–7.11 (m, 0.4 H), 5.94 (t, *J* = 6.0 Hz, 0.1 H), 5.31 (d, *J* = 12.4 Hz, 0.1 H), 5.18–4.93 (m, 5.0 H), 4.77 (d, *J* = 12.4 Hz, 0.1 H), 4.69 (app dd, *J* = 2.4, 8.0 Hz, 1.0 H), 4.54 (app dd, *J* = 4.0, 4.0 Hz, 1.0 H), 4.49–4.20 (m, 3.4 H), 4.17–3.80 (m, 10.4 H), 3.76–3.39 (m, 10.0 H), 3.31–3.23 (m, 0.9 H), 2.99 (app dd, *J* = 6.0, 17.6 Hz , 0.1 H),

2.53 (app dd, J = 5.6, 12.4 Hz , 0.7 H), 2.35–1.65 (m, 21.4 H), 1.44–1.01 (m, 9.6 H); **^{13}C NMR (100 MHz, $CDCl_3$, mixture of 3 or more rotamers, δ):** 172.9, 172.5, 171.7, 171.6, 171.5, 171.4, 171.3, 169.7, 169.6, 169.3, 168.7, 155.2, 154.9, 154.0, 136.7, 136.5 (multiple signals), 129.0, 128.8, 128.6, 128.4 (multiple signals), 128.1 (2 signals), 127.9 (2 signals), 127.7 (multiple signals), 125.3, 67.3, 67.2, 67.1, 66.9, 61.2 (multiple signals), 61.1, 61.0, 59.7, 59.4, 58.5, 58.3 (multiple signals), 57.4, 47.7, 47.3 (2 signals), 47.1 (multiple signals), 46.9 (multiple signals), 46.7 (multiple signals), 41.5, 41.2, 40.5, 33.9, 31.2, 31.8, 30.4 (multiple signals), 29.4 (2 signals), 26.9, 26.5 (multiple signals), 25.7, 25.2 (multiple signals), 25.0, 24.7, 24.3, 23.7, 23.4, 22.1, 22.0, 14.1 (2 signals); **HRMS** (ESI) *m*/*z* 449.2374 [calc'd for $C_{22}H_{29}N_3O_6Na$ $(M + Na)^+$ 449.2395].

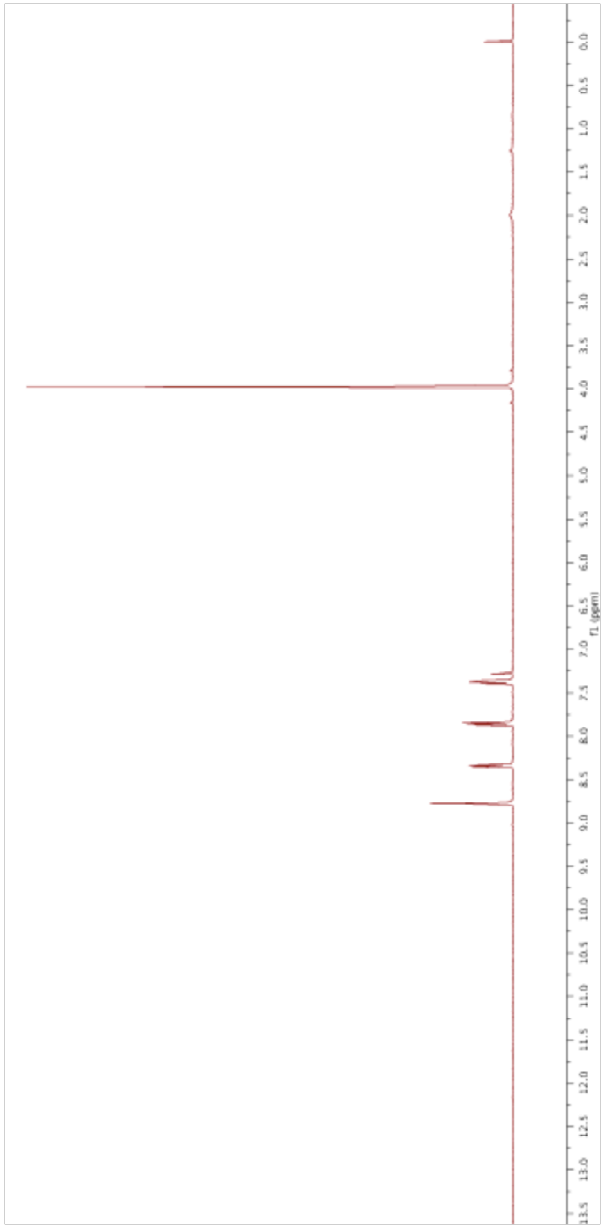
***N*-Dansylglycyl-(2*S*)-prolyl-(2*S*)-prolylglycine ethyl ester (dansylGlyProProGlyOEt)**

A suspension of CbzProProGlyOEt (755 mg, 1.8 mmoles) and Pd/C (76 mg) in MeOH (18 mL) was stirred under an atmosphere of H_2(g) for 3 h. The reaction mixture was filtered through a pad of Celite® and then concentrated under reduced pressure to a pale yellow viscous oil (553 mg). The oil was dissolved in DCM (5.9 mL) and cooled to 0 °C, after which dansylGlyOH (586 mg, 2.0 mmoles) and PyBroP® (973 mg, 2.3 mmoles) were added. The flask was capped with a septum and purged with N_2(g). DIEA (0.63 mL) was added via syringe, and the reaction mixture was stirred at 0 °C for 5 min. The reaction mixture was allowed to come to room temperature and stirred for an additional 6 h. The reaction mixture was concentrated under reduced pressure, and the crude product was purified by chromatography on silica (2% v/v MeOH in EtOAc) and

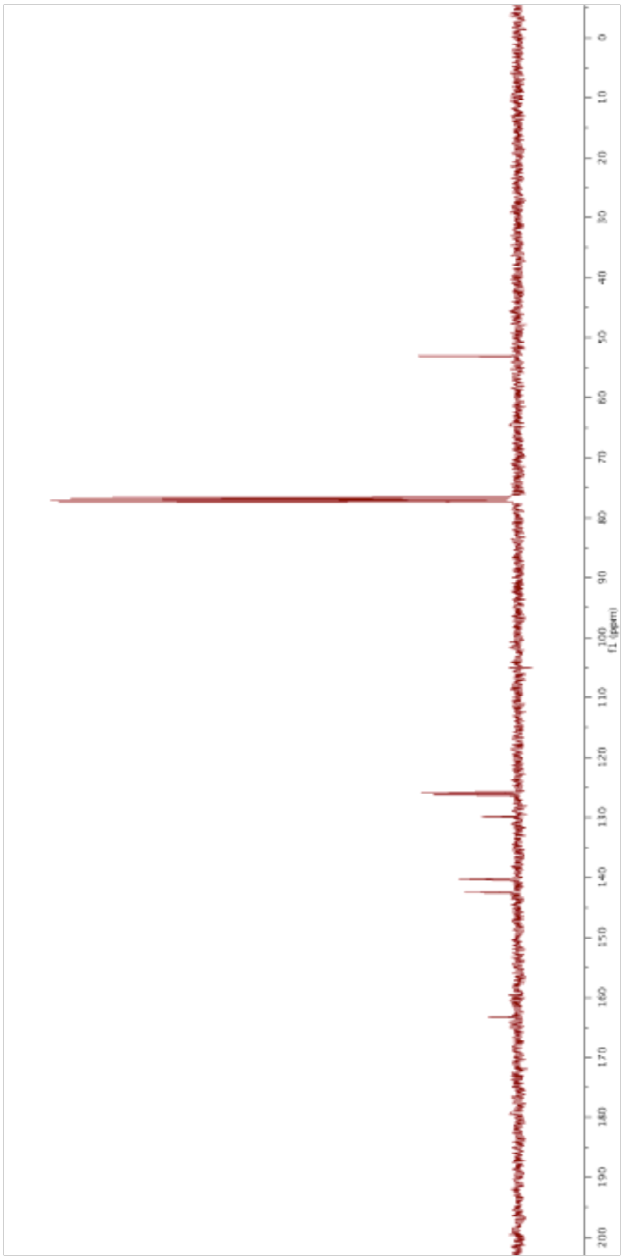
HPLC (25–55% v/v acetonitrile in H_2O with 0.1% v/v TFA) to afford the title compound (950 mg, 92%) as a yellow solid with mp 80.1–85.2 °C. **^{1}H NMR (500 MHz, $CDCl_3$, mixture of 2 or more rotamers, integrations are approximate, δ):** 8.59 (d, J = 8.5 Hz, 1.0 H), 8.37 (dd, J = 8.5, 14.5 Hz, 1.1 H), 8.25 (ddd, J = 1.0, 7.0, 9.5 Hz, 1.0 H), 8.20–8.14 (m, 0.7 H), 7.65–7.55 (m, 2.1 H), 7.26 (d, J = 7.5 Hz, 1.0 H), 7.18 (t, J = 5.0 Hz, 0.7 H), 5.76 (dd, J = 3.5, 5.0 Hz, 0.7 H), 5.67 (dd, J = 3.5, 5.0 Hz, 0.6 H), 4.58 (app dd, J = 2.5, 8.0 Hz, 1.2 H), 4.45 (t, J = 6.5 Hz, 0.7 H), 4.31 (d, J = 8.5 Hz, 0.6 H), 4.24–4.13 (m, 2.1 H), 4.09–4.04 (m, 0.8 H), 3.94–3.80 (m, 2.6 H), 3.74–3.61 (m, 2.3 H), 3.55–3.35 (m, 3.4 H), 2.97–2.95 (m, 4.9 H), 2.56 (app dd, J = 6.0, 12.5 Hz, 0.7 H), 2.36–2.31 (m, 0.8 H), 2.24–1.80 (m, 11.9 H), 1.33–1.26 (m, 2.9 H); **^{13}C NMR (125 MHz, $CDCl_3$, mixture of 2 or more rotamers, δ):** 171.5, 171.4, 170.3, 169.7, 169.4, 165.8, 165.4, 151.3, 150.9, 134.3, 134.2, 130.5, 130.4, 129.8 (2 signals), 129.7 (2 signals), 129.5, 129.4, 128.5, 128.4, 123.4, 123.3, 119.8 (multiple signals), 119.5 (multiple signals), 115.6 (2 signals), 61.4, 61.3, 61.1, 59.7, 59.1, 58.2, 47.2, 47.1, 46.6, 46.2, 45.6, 45.5, 44.6, 44.5, 41.5, 41.3, 31.8, 28.5, 28.4, 27.0, 25.2, 24.9, 24.5, 22.1, 14.2 (2 signals); **HRMS** (ESI) m/z 610.2325 [calc'd for $C_{28}H_{37}N_5O_7SNa$ $(M + Na)^+$ 610.2306].

2.5.11 NMR Spectra

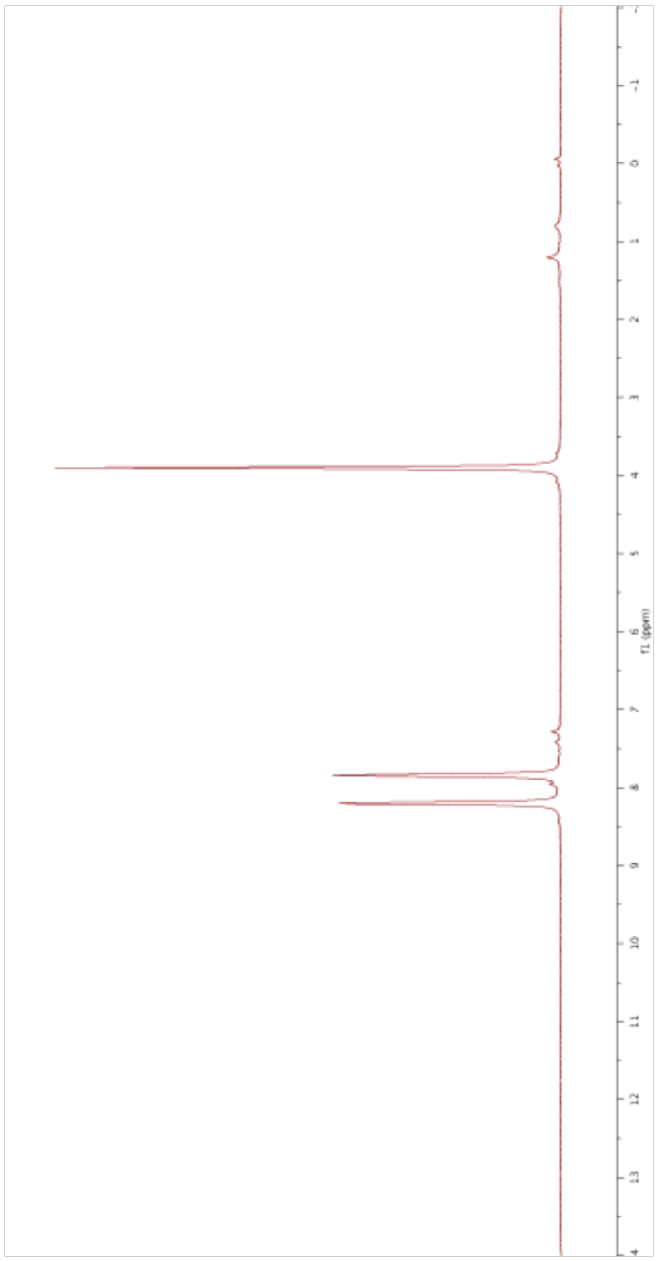
400 MHz ^{1}H NMR spectrum of 3-Methoxycarbonylpyridine *N*-Oxide in $CDCl_3$



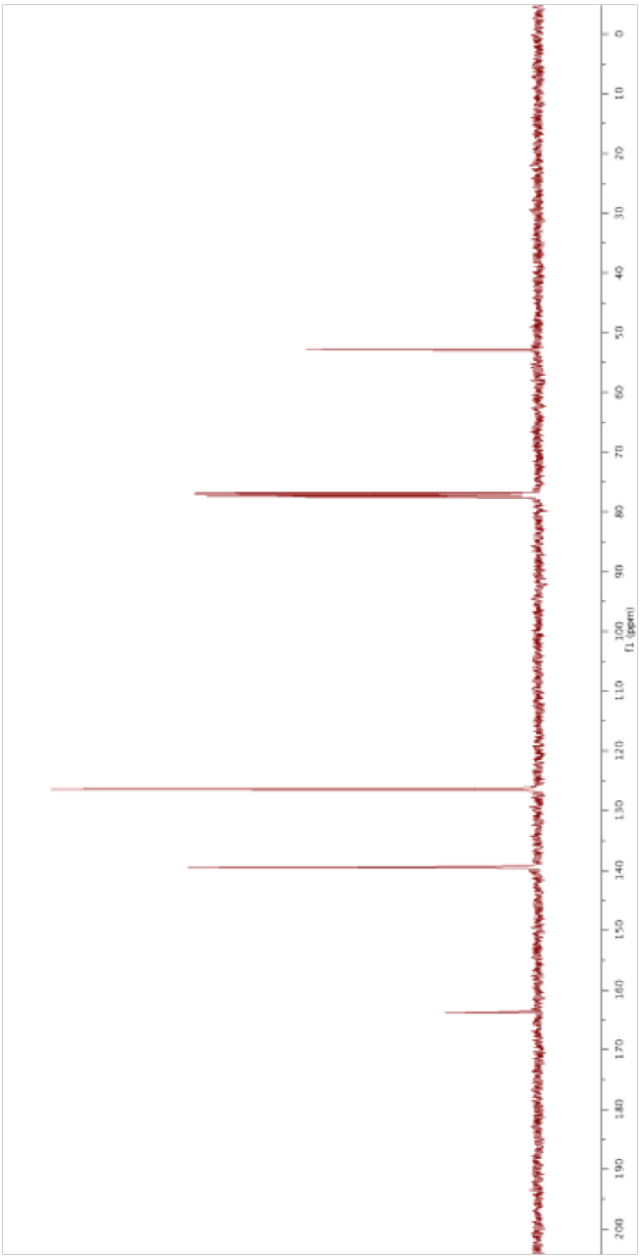
100 MHz ^{13}C NMR spectrum of 3-Methoxycarbonylpyridine *N*-Oxide in $CDCl_3$



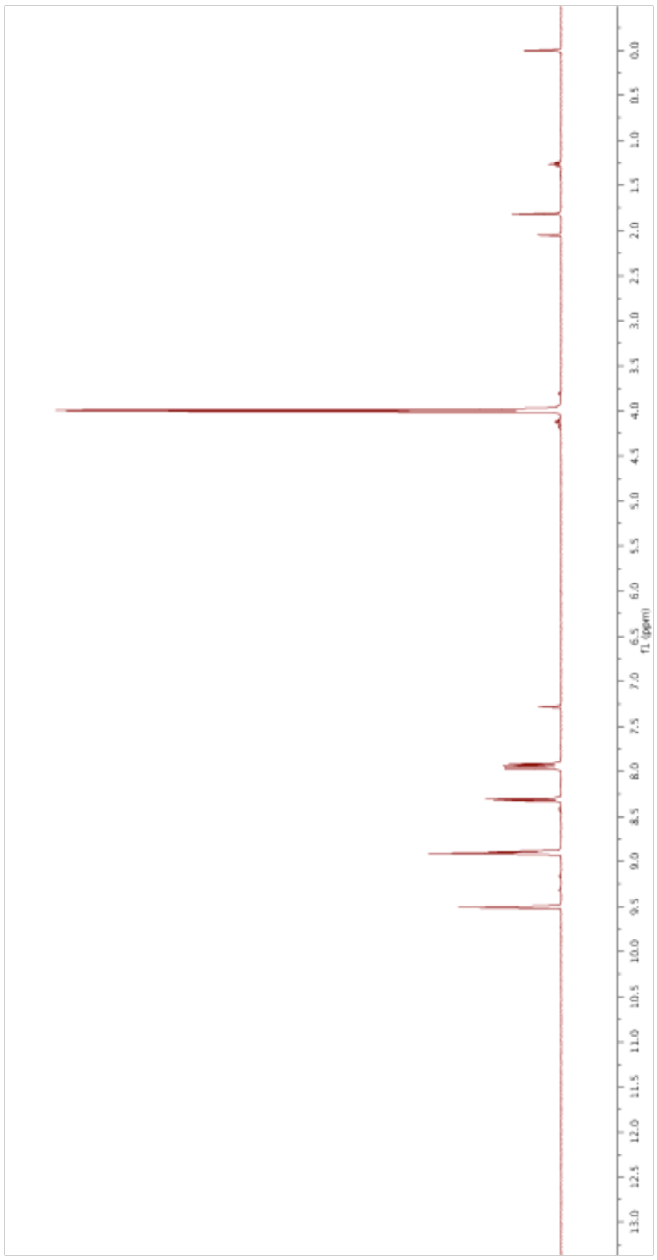
400 MHz ^{1}H NMR spectrum of 4-Methoxycarbonylpyridine *N*-Oxide in $CDCl_3$



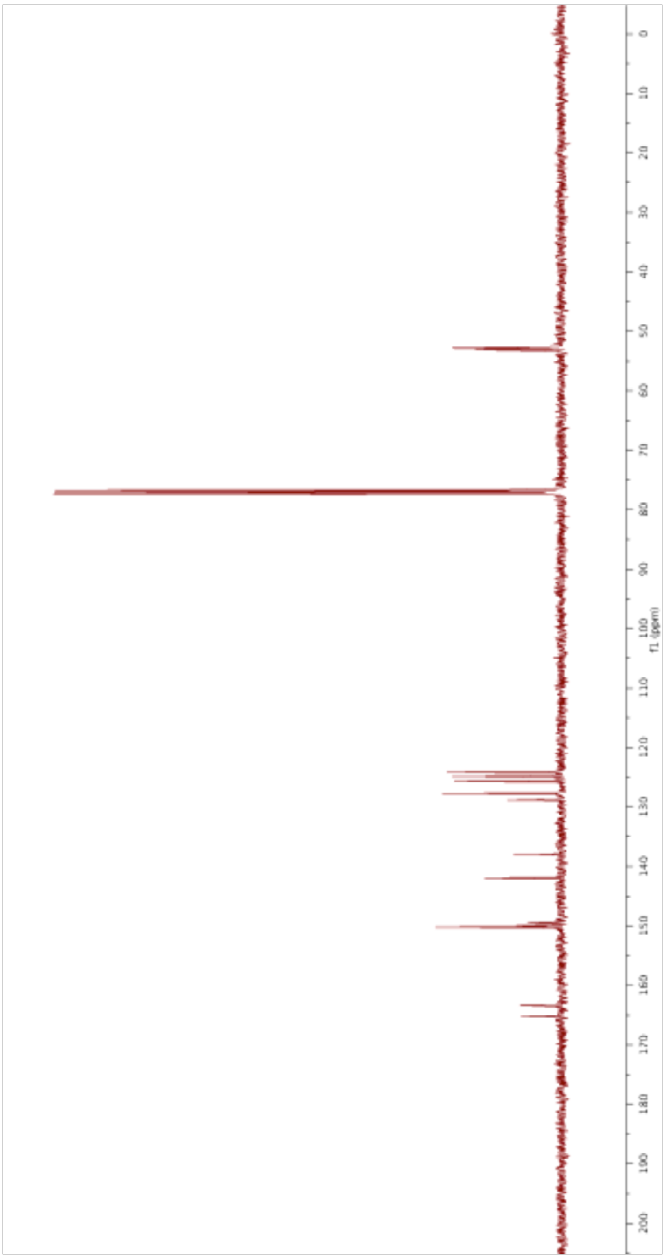
100 MHz ^{13}C NMR spectrum of 4-Methoxycarbonylpyridine *N*-Oxide in $CDCl_3$



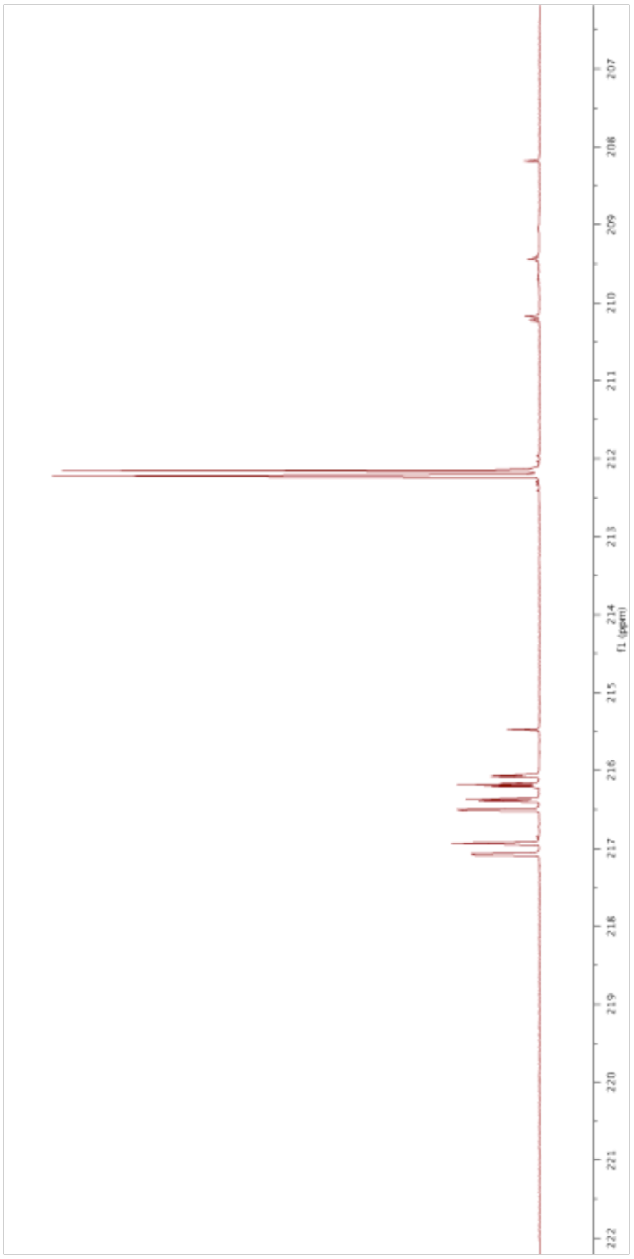
400 MHz ^{1}H NMR spectrum of 5-Methoxycarbonyl-2-(4-methoxycarbonylpyridin-2-yl)pyridine *N*-Oxide in $CDCl_3$



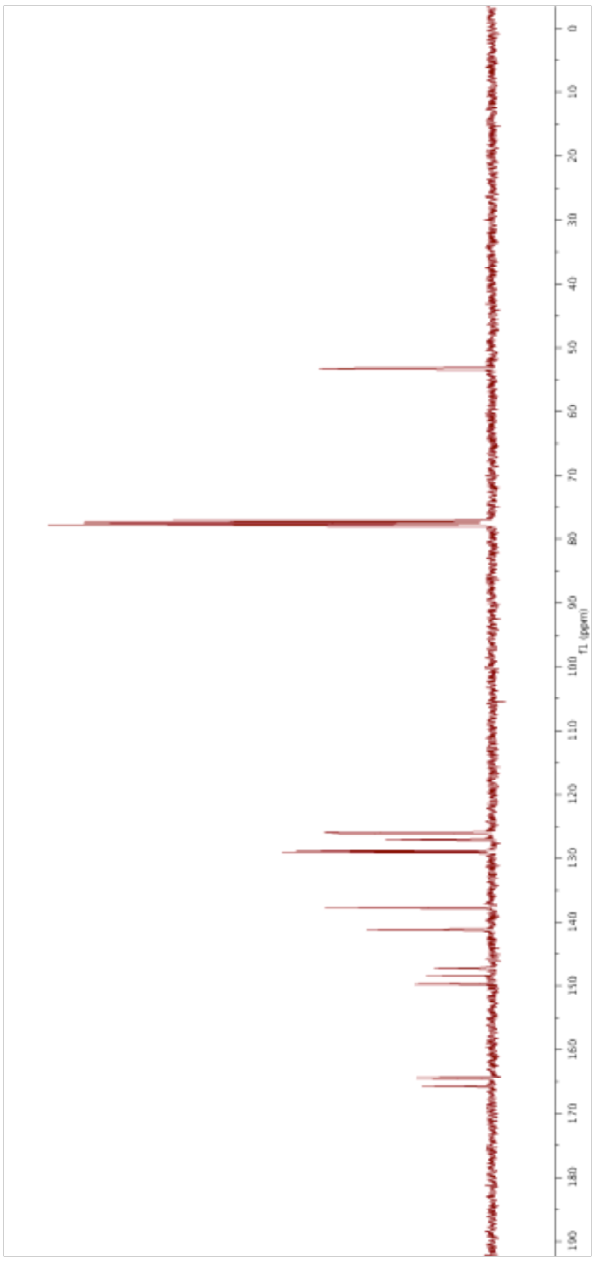
100 MHz ^{13}C NMR spectrum of 5-Methoxycarbonyl-2-(4-methoxycarbonylpyridin-2-yl)pyridine *N*-Oxide in $CDCl_3$



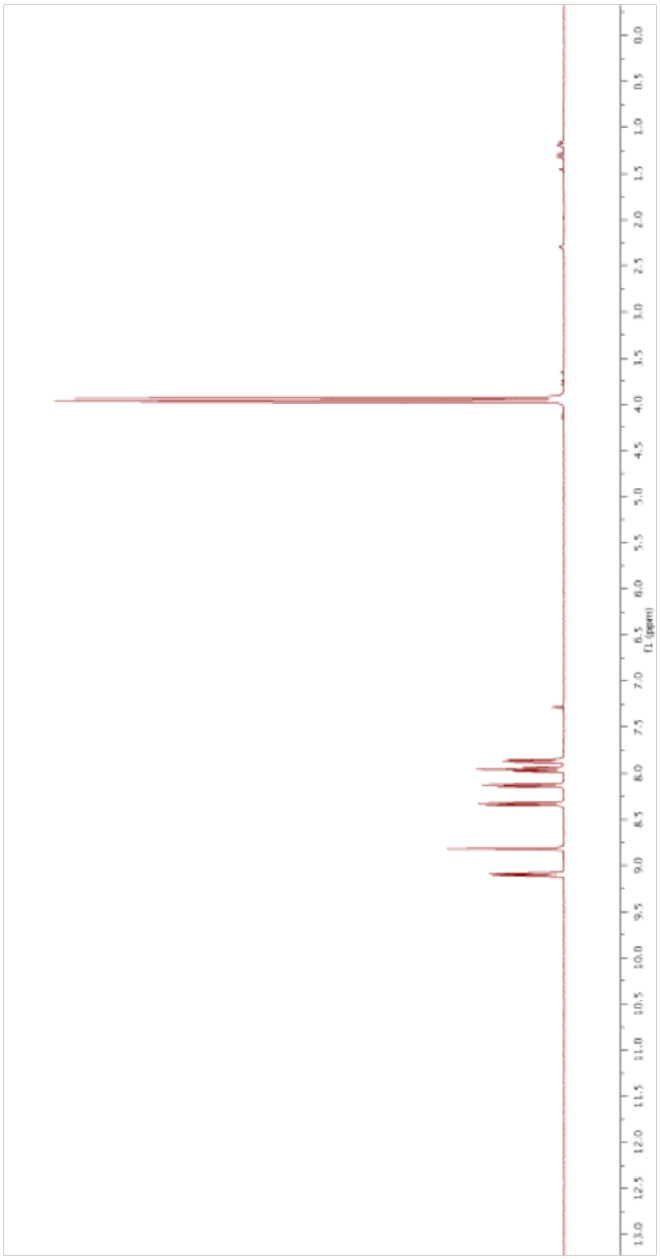
400 MHz ^{1}H NMR spectrum of 4-Methoxycarbonyl-2-(6-methoxycarbonylpyridin-2-yl)pyridine *N*-Oxide in $CDCl_3$



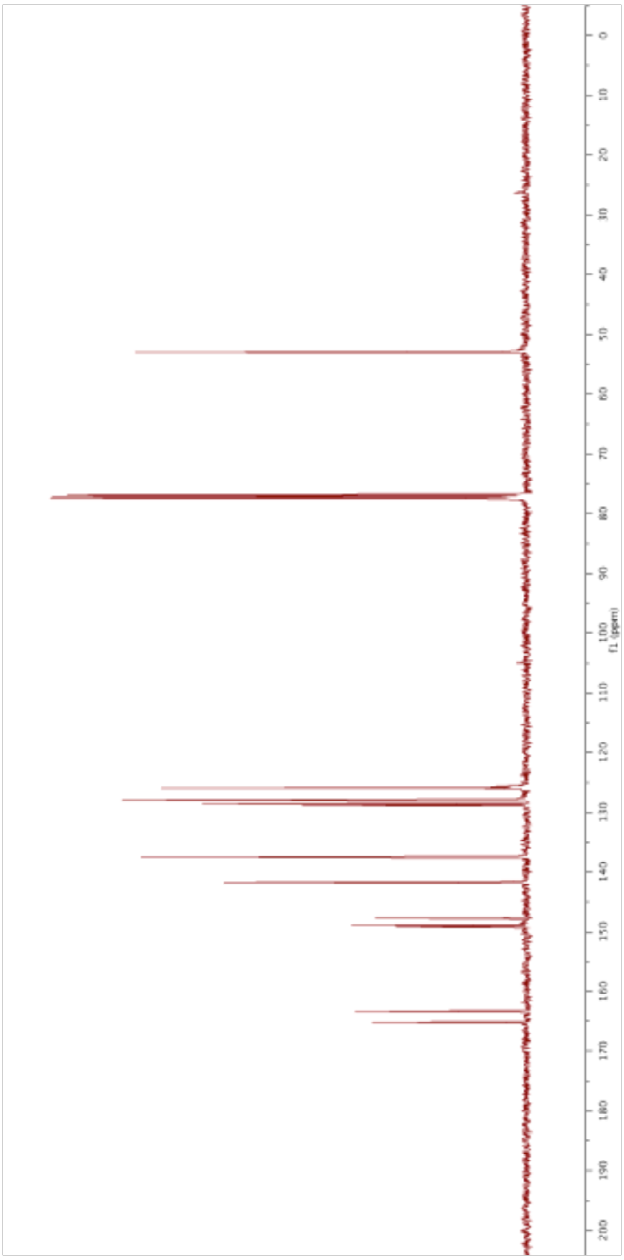
100 MHz ^{13}C NMR spectrum of 4-Methoxycarbonyl-2-(6-methoxycarbonylpyridin-2-yl)pyridine *N*-Oxide in $CDCl_3$



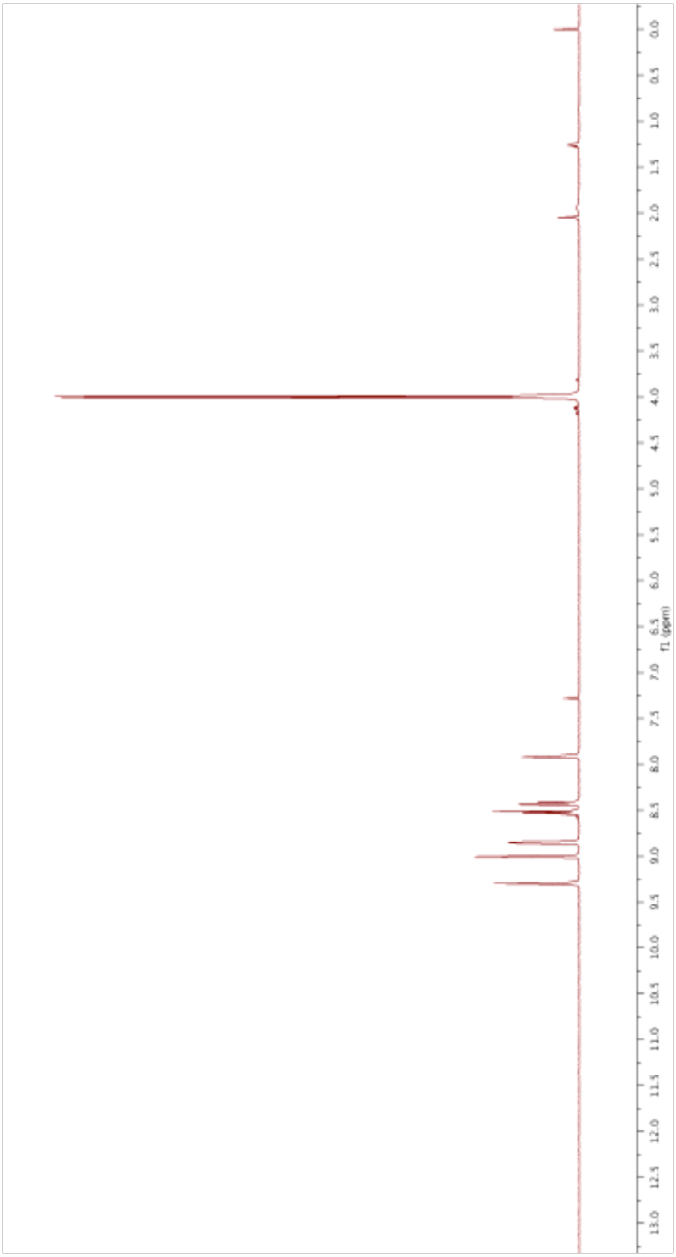
400 MHz ^{1}H NMR spectrum of 5-Methoxycarbonyl-2-(6-methoxycarbonylpyridin-2-yl)pyridine *N*-Oxide in $CDCl_3$



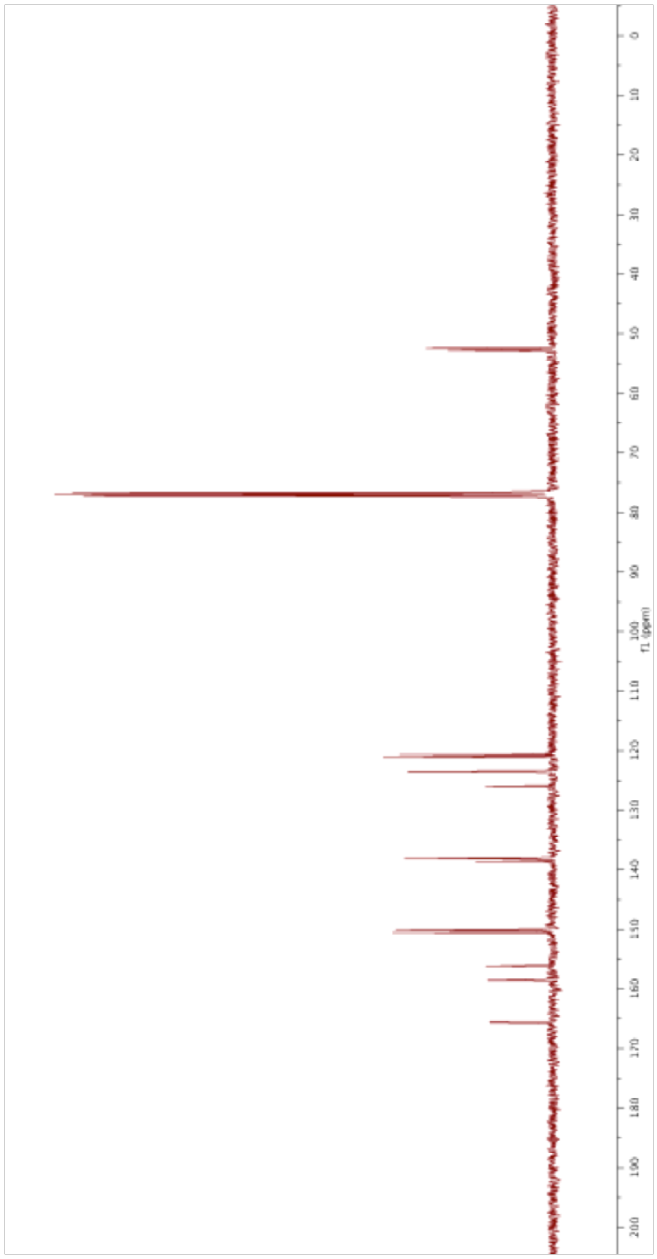
100 MHz ^{13}C NMR spectrum of 5-Methoxycarbonyl-2-(6-methoxycarbonylpyridin-2-yl)pyridine *N*-Oxide in $CDCl_3$



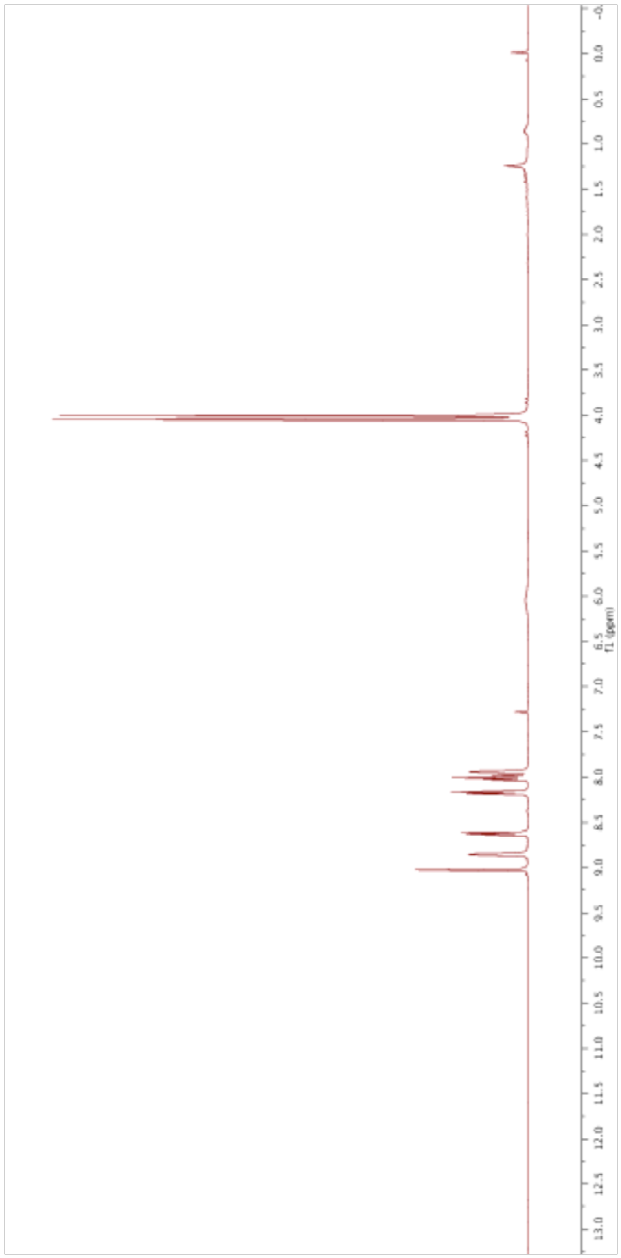
400 MHz ^{1}H NMR spectrum of Dimethyl 2,2′-bipyridine-4,5′-dicarboxylate in $CDCl_3$



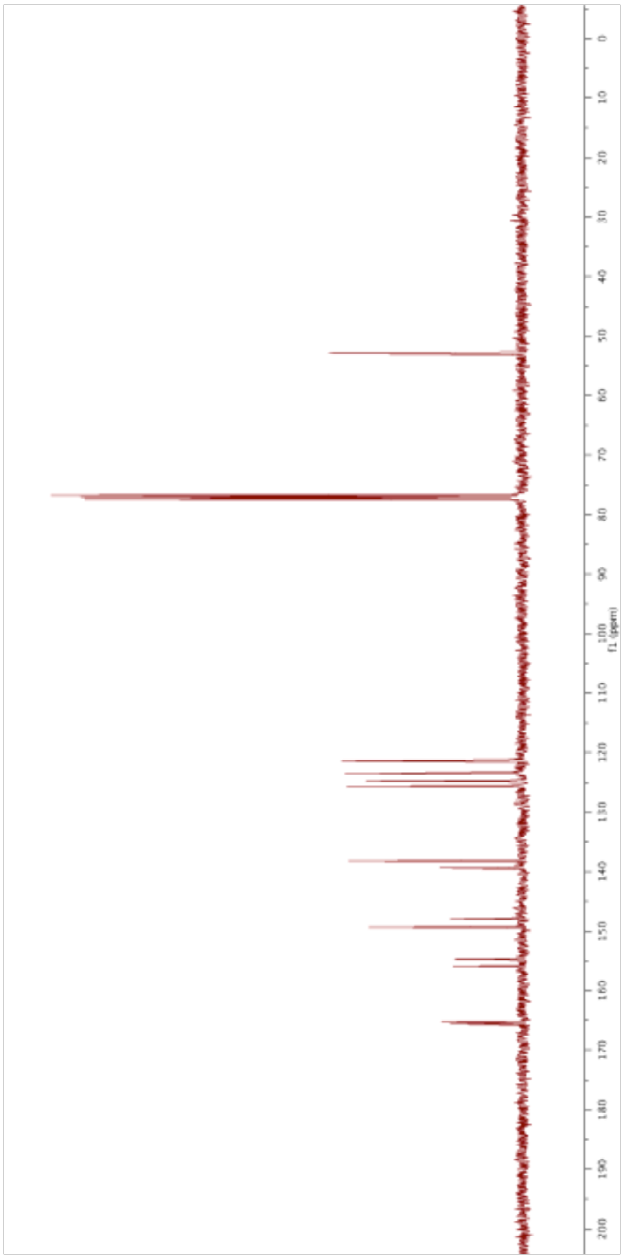
100 MHz ^{13}C NMR spectrum of Dimethyl 2,2′-bipyridine-4,5′-dicarboxylate in $CDCl_3$



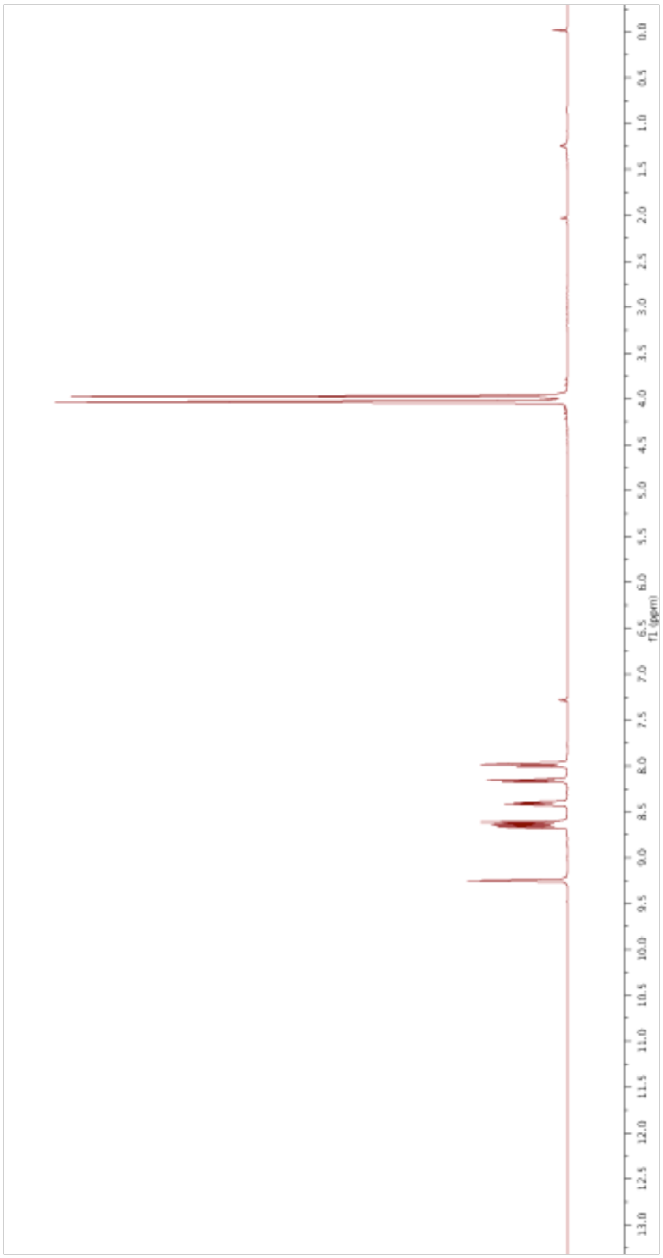
400 MHz ^{1}H NMR spectrum of Dimethyl 2,2′-bipyridine-4,6′-dicarboxylate in $CDCl_3$



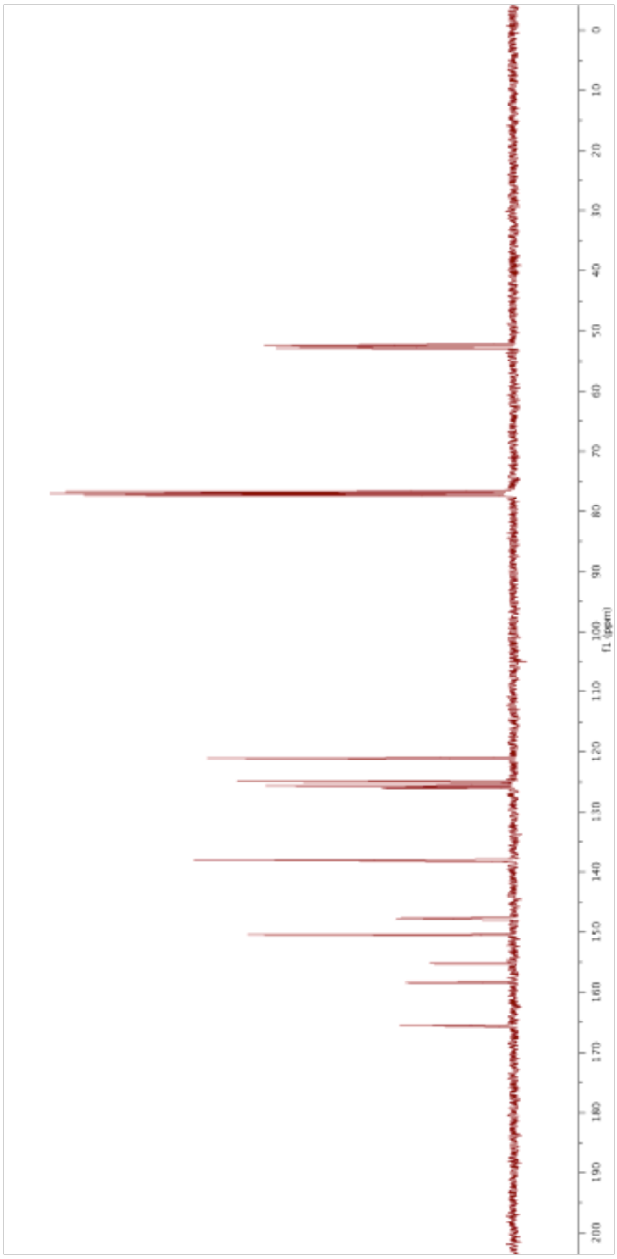
100 MHz ^{13}C NMR spectrum of Dimethyl 2,2′-bipyridine-4,6′-dicarboxylate in $CDCl_3$



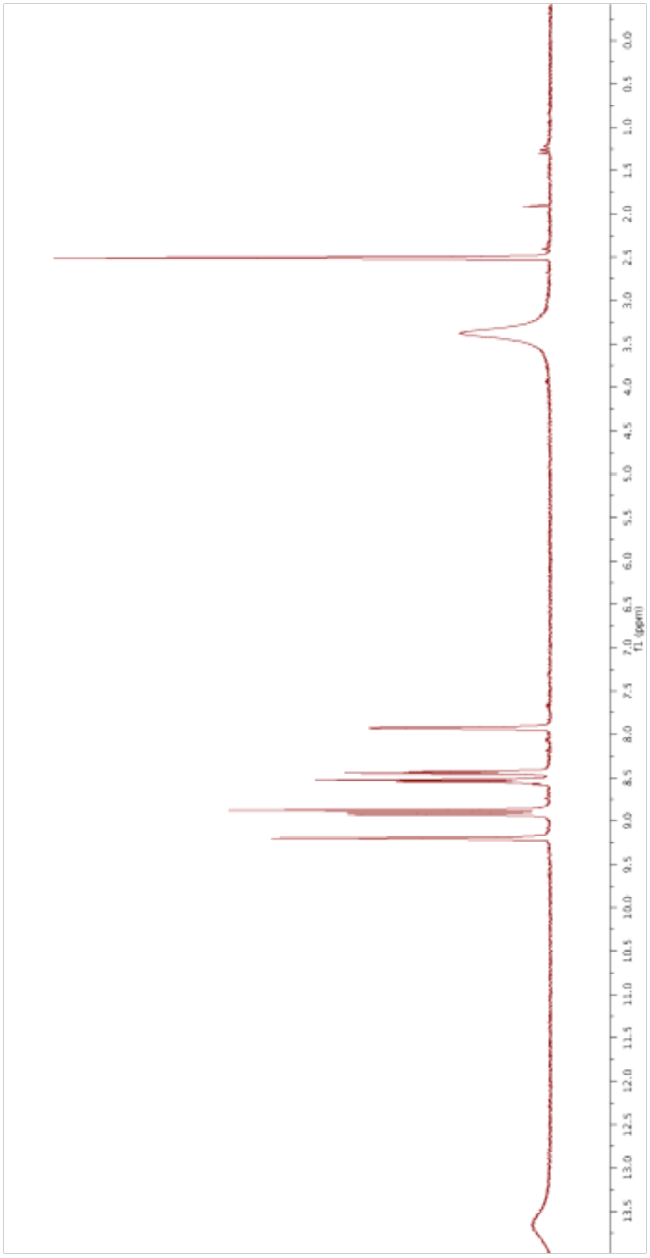
400 MHz ^{1}H NMR spectrum of Dimethyl 2,2′-bipyridine-5,6′-dicarboxylate in $CDCl_3$



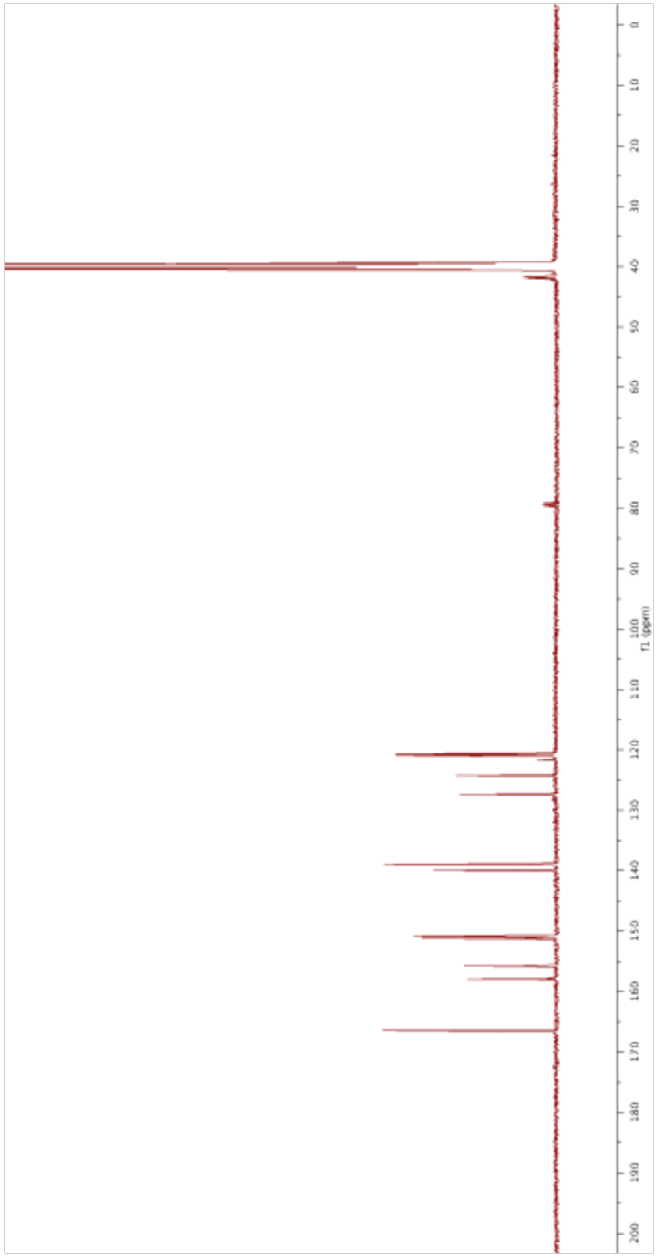
100 MHz ^{13}C NMR spectrum of Dimethyl 2,2′-bipyridine-5,6′-dicarboxylate in $CDCl_3$



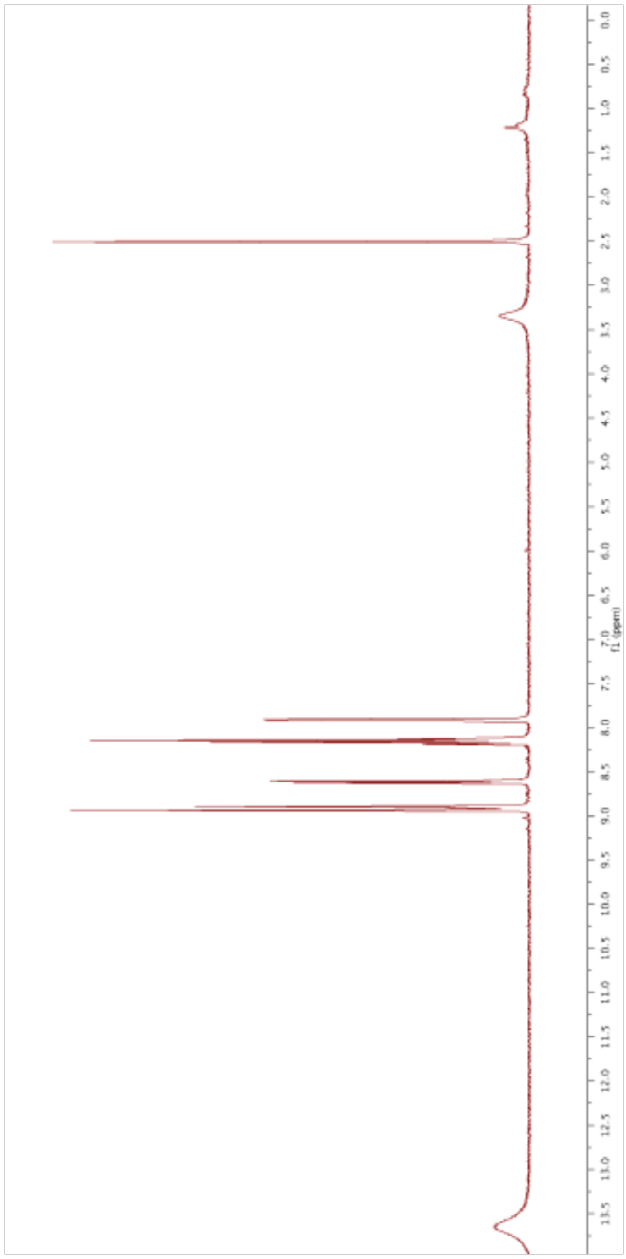
400 MHz ^{1}H NMR spectrum of 2,2′-Bipyridine-4,5′-dicarboxylic acid in DMSO-d_6



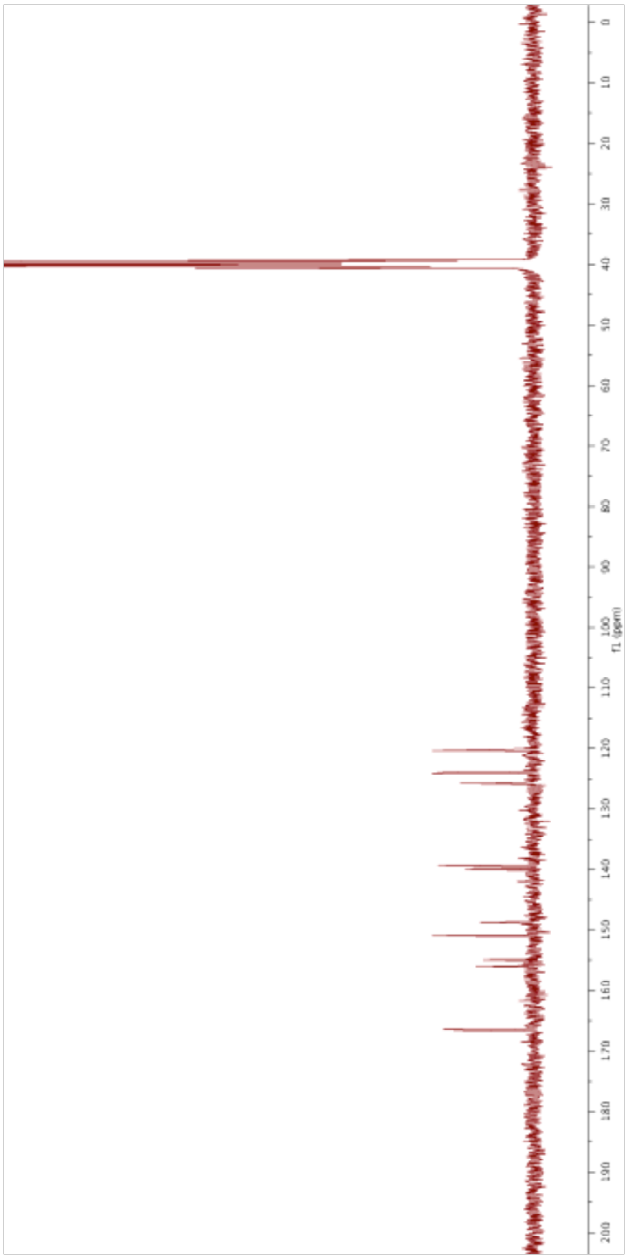
125 MHz ^{13}C NMR spectrum of 2,2′-Bipyridine-4,5′-dicarboxylic acid in DMSO-d_6



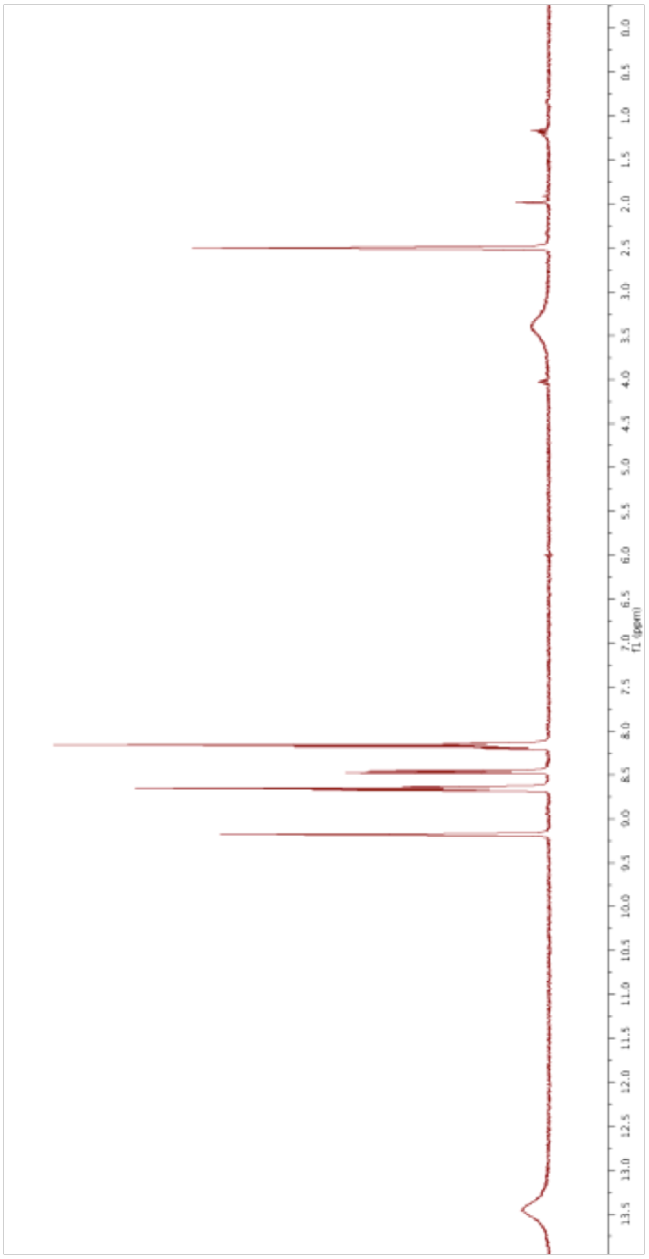
400 MHz ^{1}H NMR spectrum of 2,2′-Bipyridine-4,6′-dicarboxylic acid in DMSO-d_6



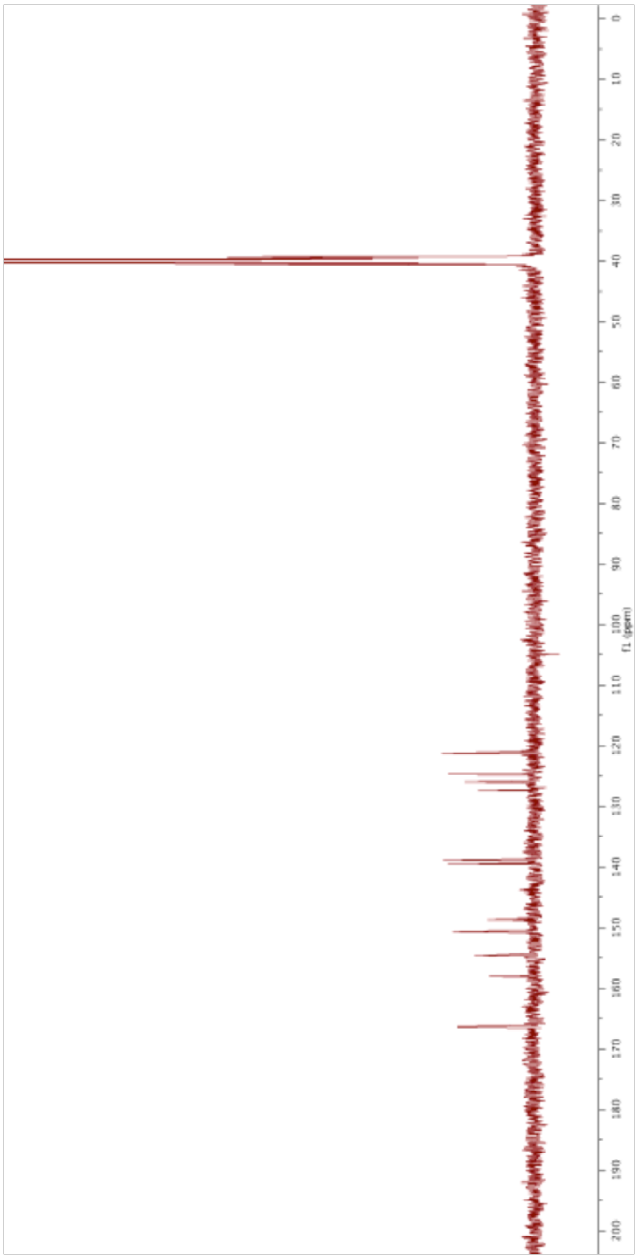
100 MHz ^{13}C NMR spectrum of 2,2′-Bipyridine-4,6′-dicarboxylic acid in DMSO-d_6



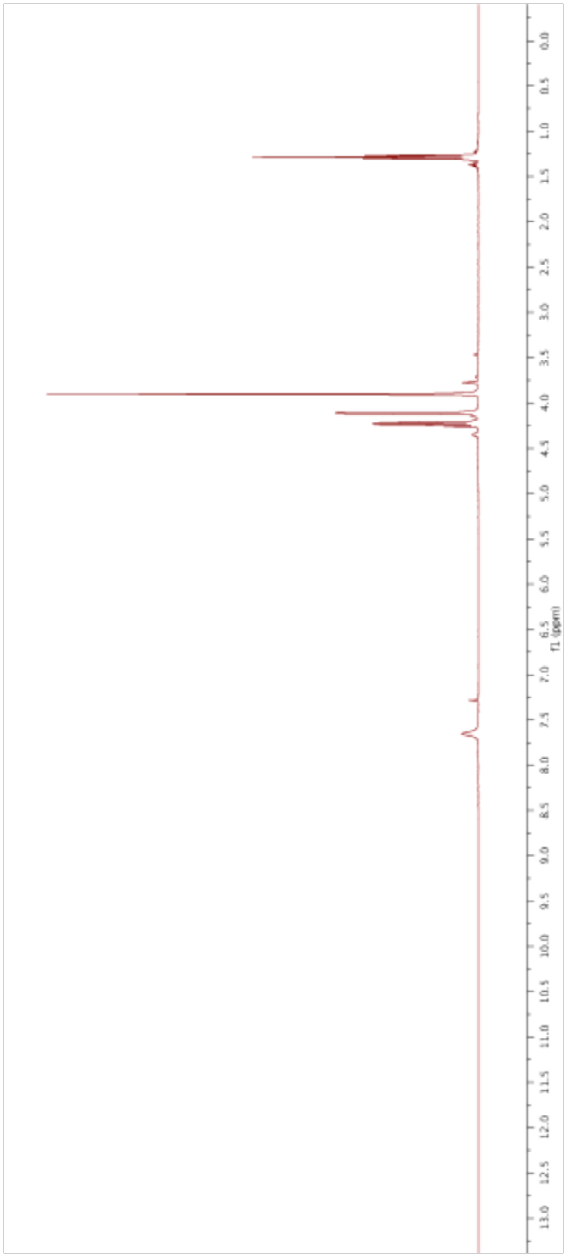
400 MHz ^{1}H NMR spectrum of 2,2′-Bipyridine-5,6′-dicarboxylic acid in DMSO-d_6



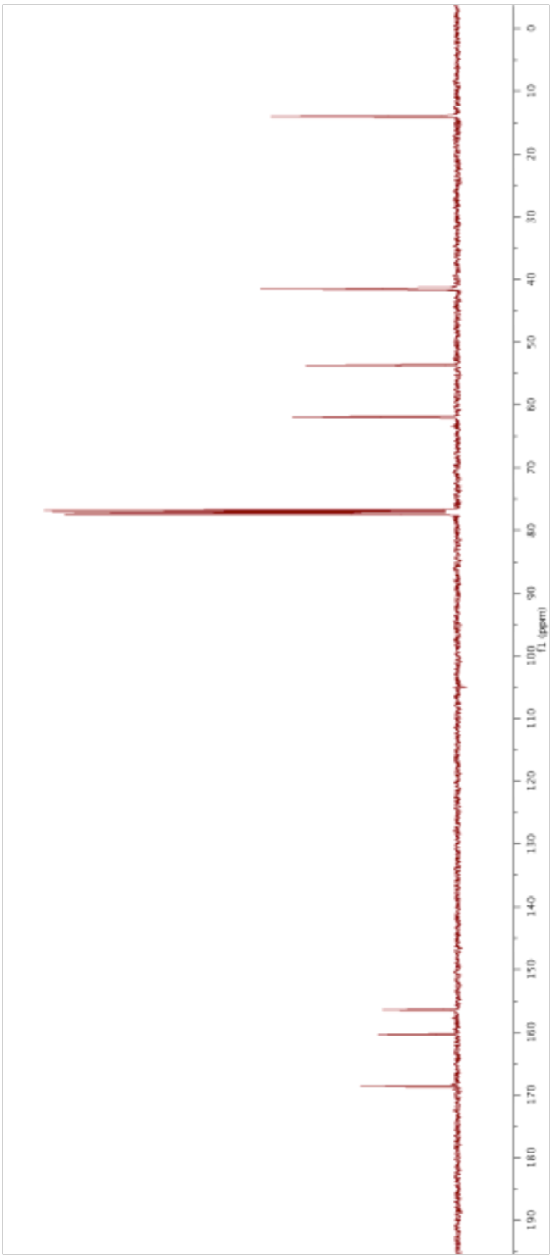
100 MHz ^{13}C NMR spectrum of 2,2′-Bipyridine-5,6′-dicarboxylic acid in DMSO-d_6



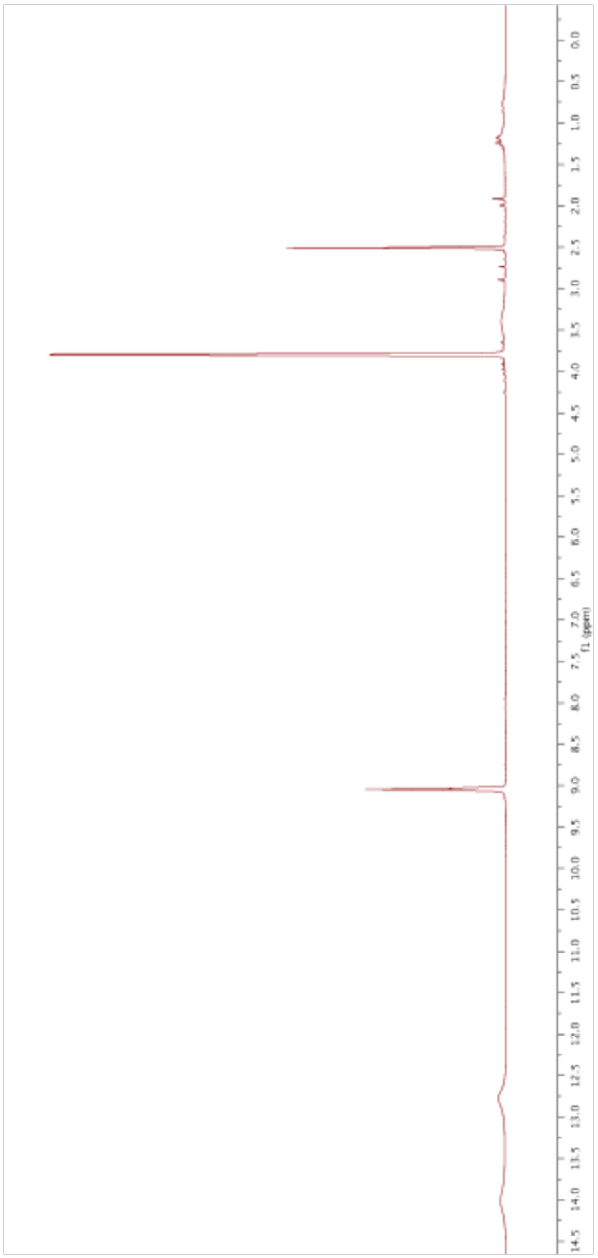
400 MHz ^{1}H NMR spectrum of *N*-(Methoxyoxalyl)glycine ethyl ester in $CDCl_3$



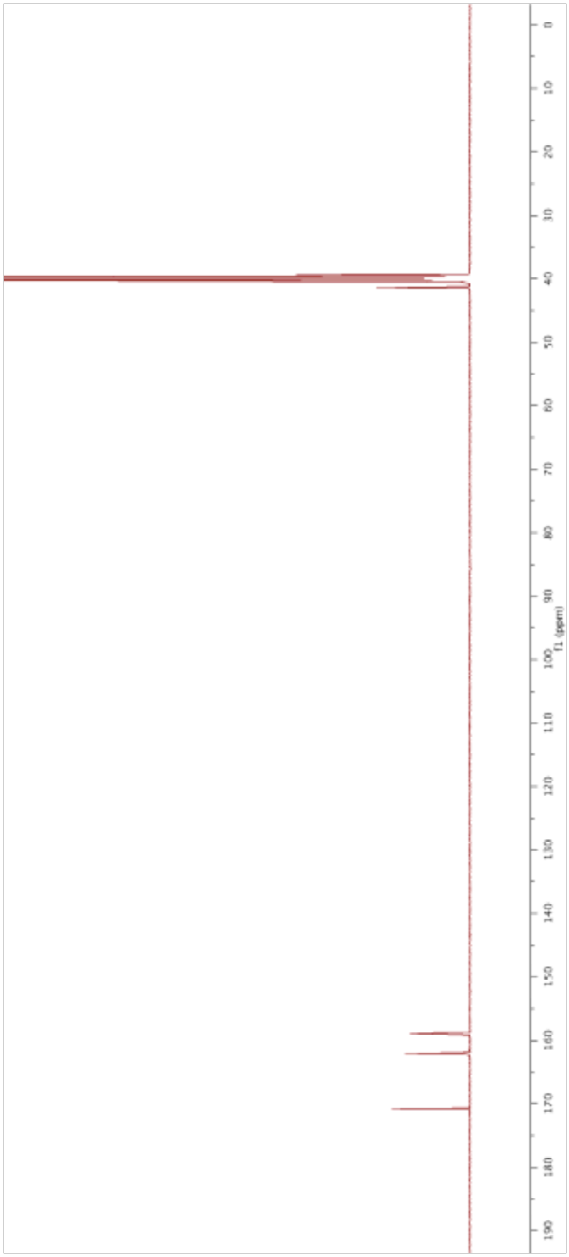
100 MHz ^{13}C NMR spectrum of *N*-(Methoxyoxalyl)glycine ethyl ester in $CDCl_3$



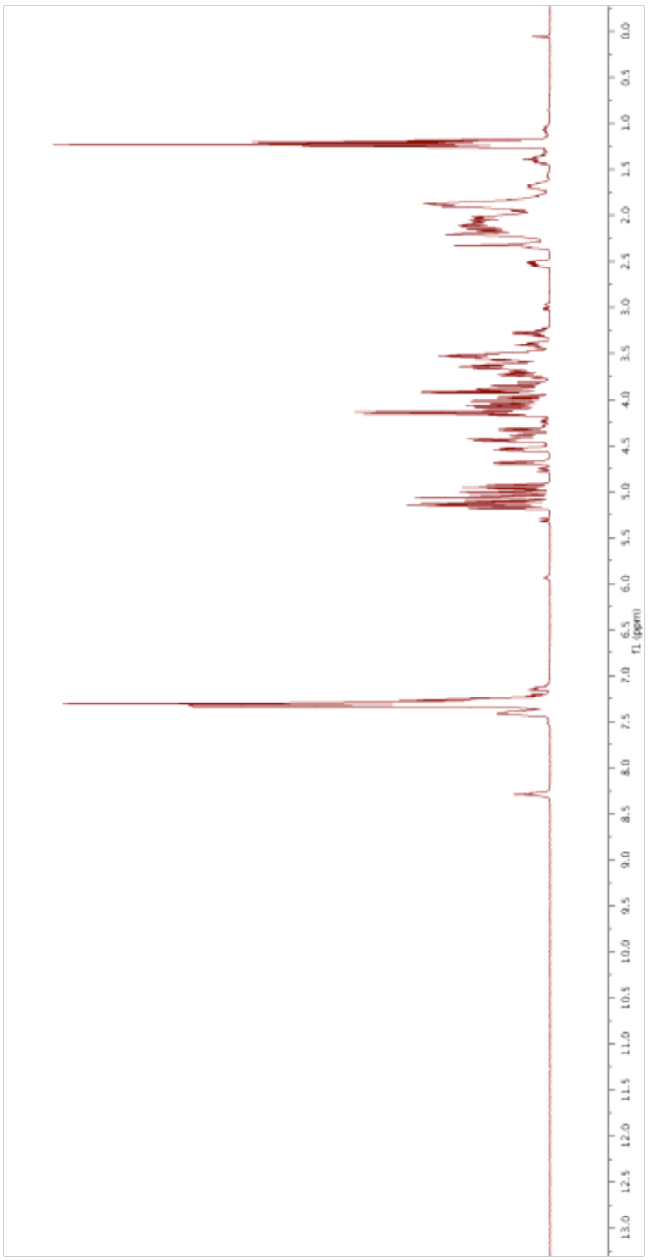
500 MHz ^{1}H NMR spectrum of *N*-Oxalylglycine in DMSO-d_6



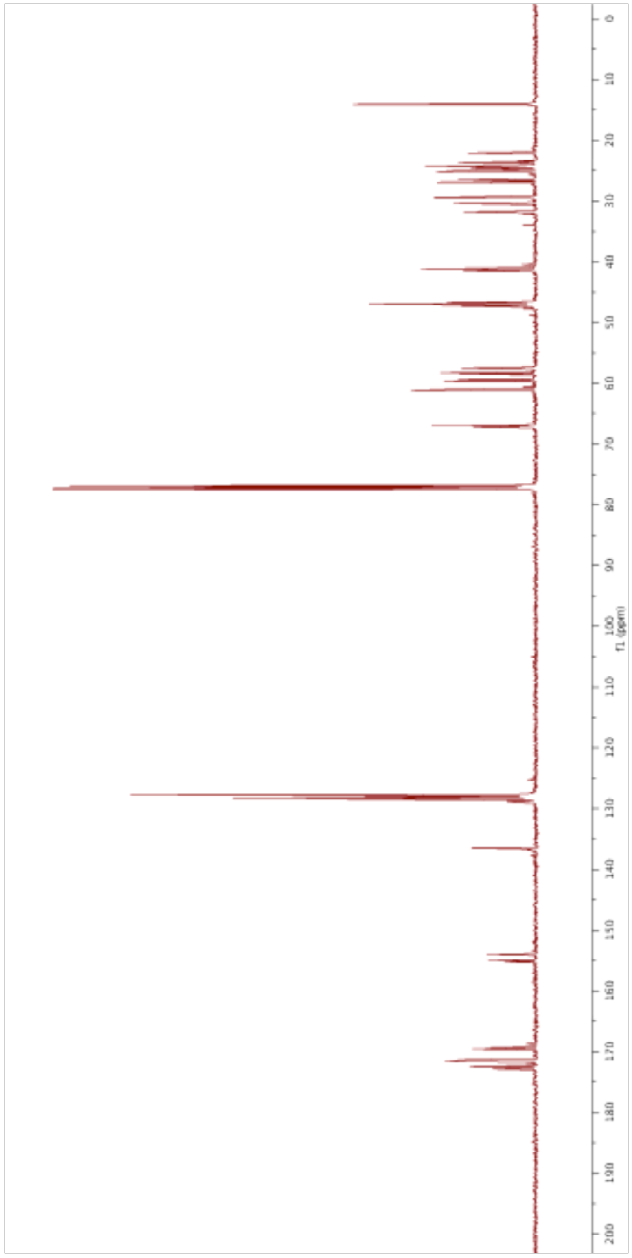
125 MHz ^{13}C NMR spectrum of *N*-Oxalylglycine in DMSO-d_6



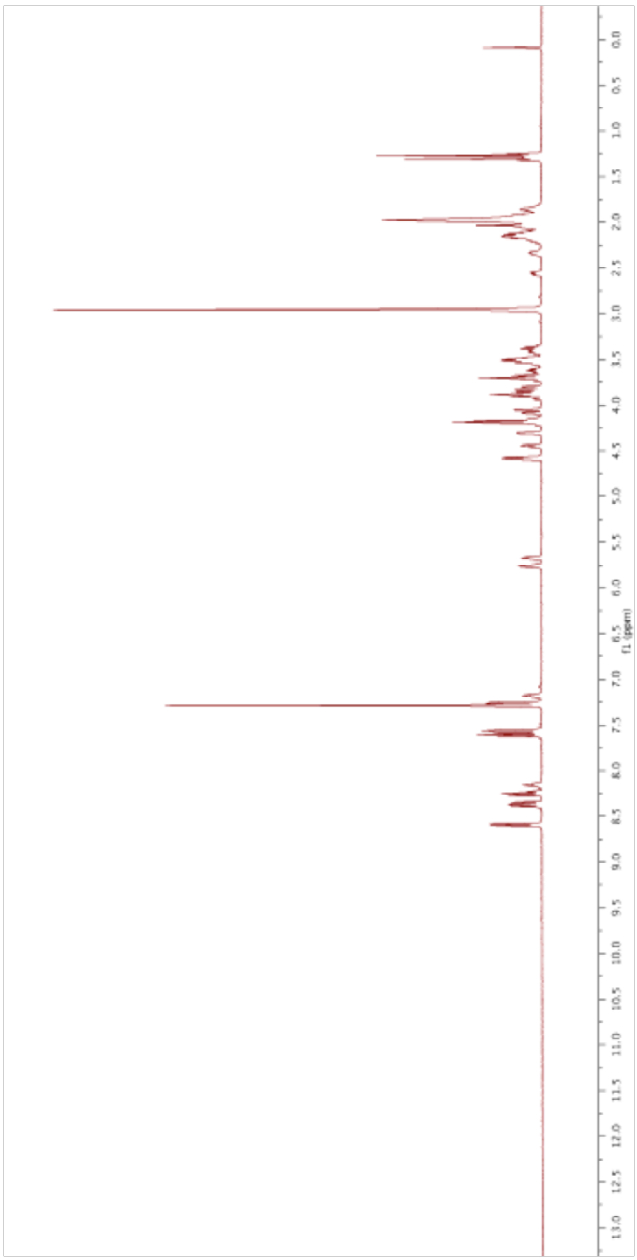
400 MHz ^{1}H NMR spectrum of *N*-Benzyloxycarbonyl-(2*S*)-prolyl-(2*S*)-prolylglycine ethyl ester (CbzProProGlyOEt) in $CDCl_3$



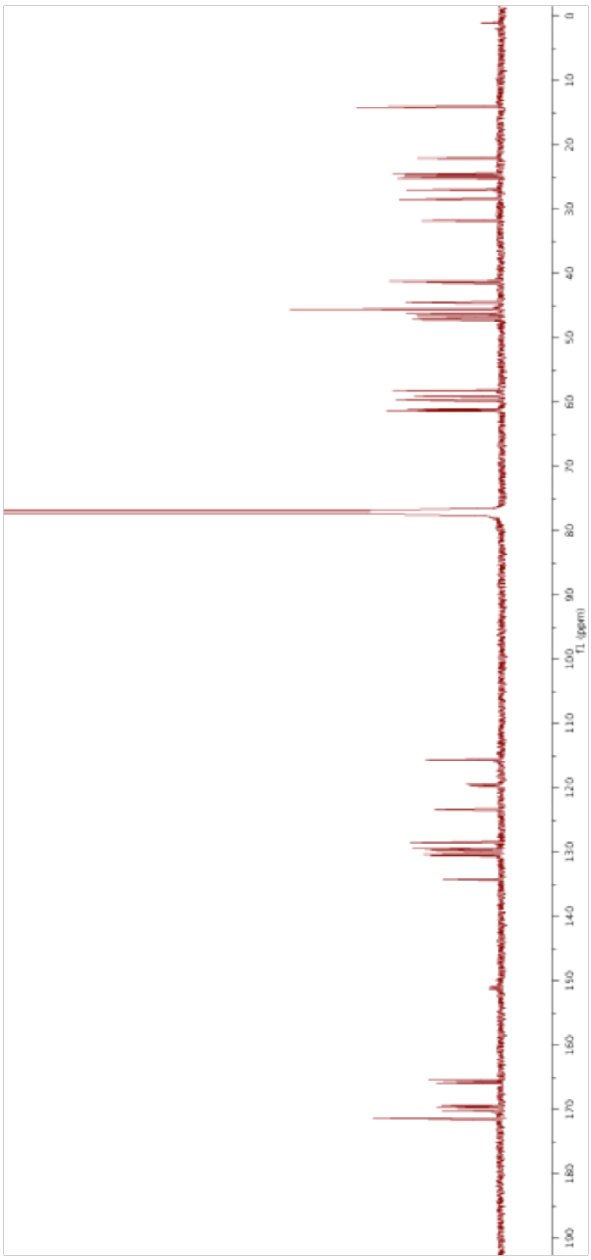
100 MHz ^{13}C NMR spectrum of *N*-Benzyloxycarbonyl-(2*S*)-prolyl-(2*S*)-prolylglycine ethyl ester (CbzProProGlyOEt) in $CDCl_3$



500 MHz ^{1}H NMR spectrum of *N*-Dansylglycyl-(2*S*)-prolyl-(2*S*)-prolylglycine ethyl ester (dansylGlyProProGlyOEt) in $CDCl_3$

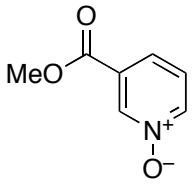


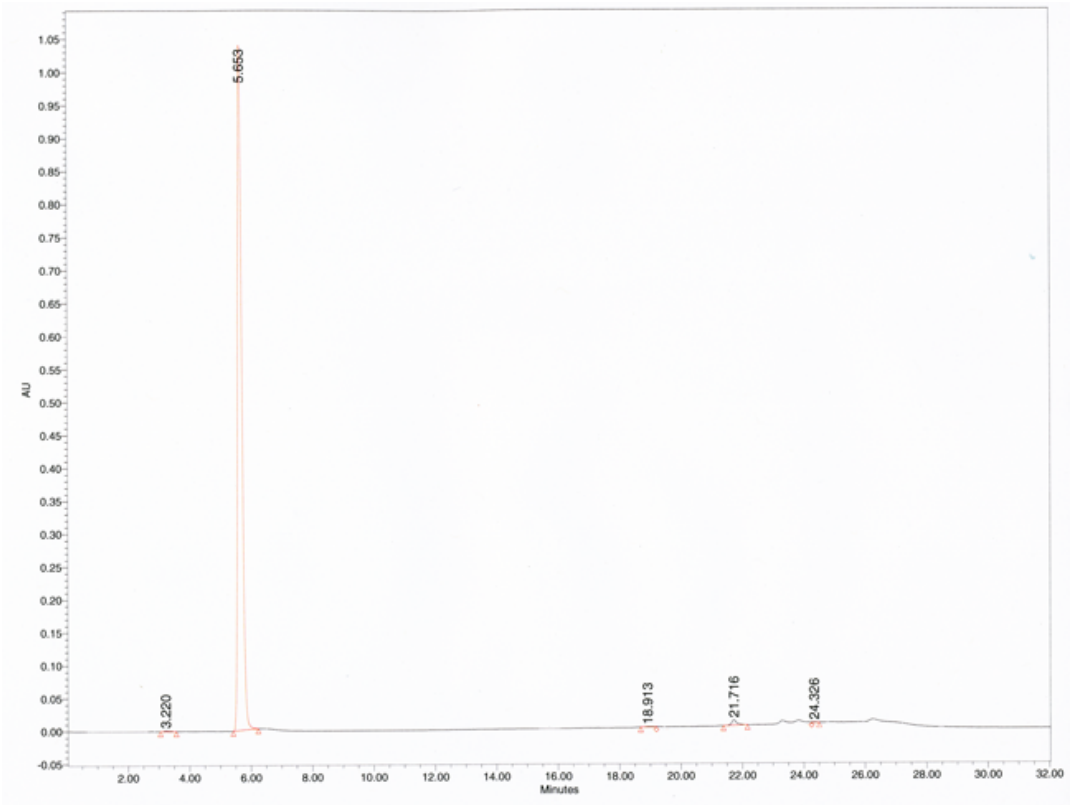
125 MHz ^{13}C NMR spectrum of *N*-Dansylglycyl-(2*S*)-prolyl-(2*S*)-prolylglycine ethyl ester (dansylGlyProProGlyOEt) in $CDCl_3$



2.5.12 HPLC Purity Chromatograms

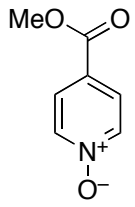
3-Methoxycarbonylpyridine *N*-oxide

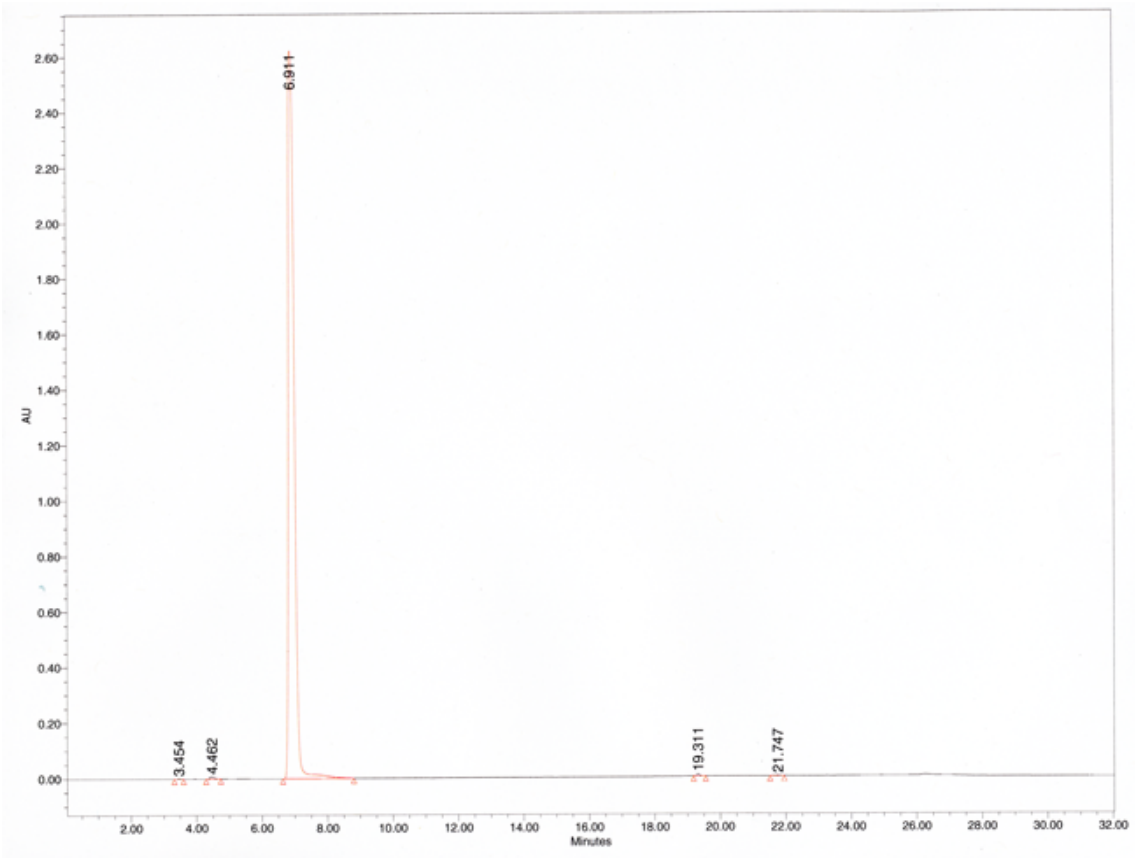




Peak	Retention Time (min)	Area (mV*sec)	% Area	Height (mV)
1	3.22	13456	0.13	1356
2	**5.653**	**10173512**	**98.71**	**1045398**
3	18.913	13293	0.13	623
4	21.716	98669	0.96	8309
5	24.326	7021	0.07	858

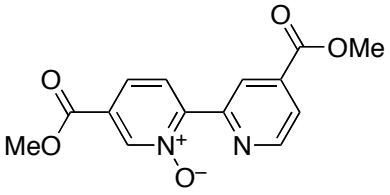
4-Methoxycarbonylpyridine *N*-oxide

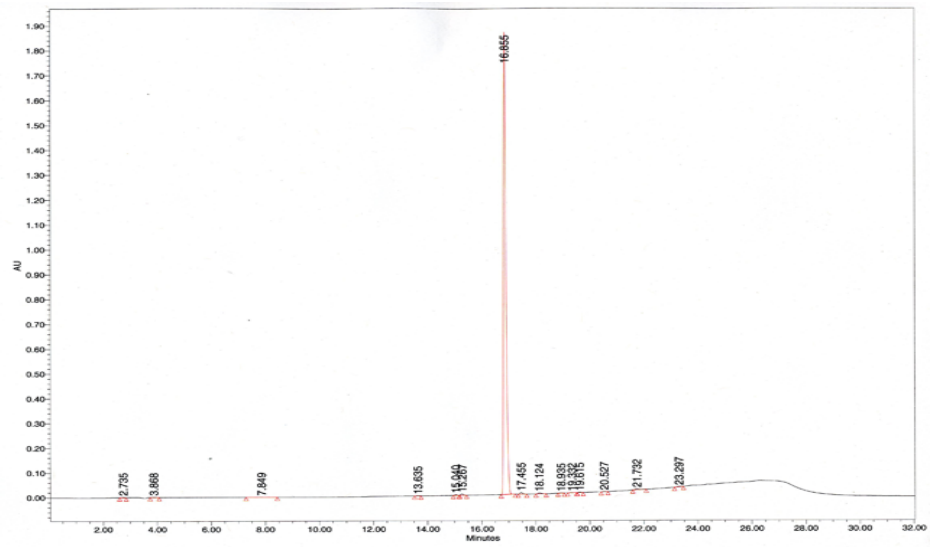




Peak	Retention Time (min)	Area (mV*sec)	% Area	Height (mV)
1	3.454	6669	0.02	1148
2	4.462	68515	0.2	8176
3	**6.911**	**33472343**	**99.59**	**2623673**
4	19.311	48585	0.14	7532
5	21.747	14944	0.04	1754

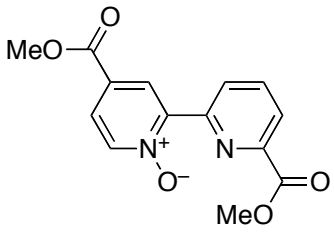
5-Methoxycarbonyl-2-(4-methoxycarbonylpyridin-2-yl)pyridine *N*-Oxide

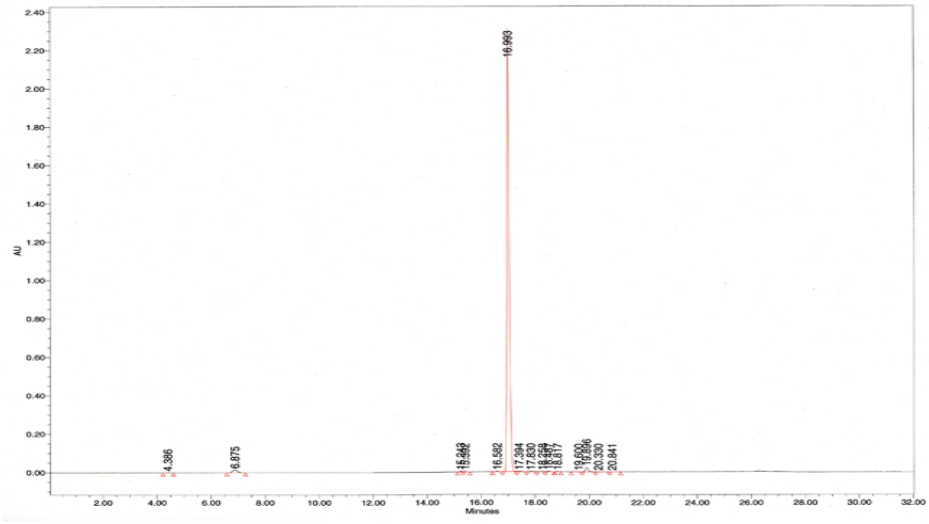




Peak	Retention Time (min)	Area (mV*sec)	% Area	Height (mV)
1	2.735	13037	0.11	2049
2	3.868	10779	0.09	1071
3	7.849	13882	0.11	369
4	13.635	2026	0.02	299
5	15.04	17327	0.14	3104
6	15.267	34877	0.29	6078
7	**16.855**	**11901432**	**98.04**	**1885885**
8	17.455	46443	0.38	7399
9	18.124	30013	0.25	3621
10	18.935	3385	0.03	509
11	19.332	4940	0.04	618
12	19.615	6866	0.06	1127
13	20.527	5029	0.04	707
14	21.732	39489	0.33	3867
15	23.297	9688	0.08	912

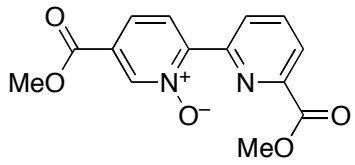
4-Methoxycarbonyl-2-(6-methoxycarbonylpyridin-2-yl)pyridine *N*-Oxide

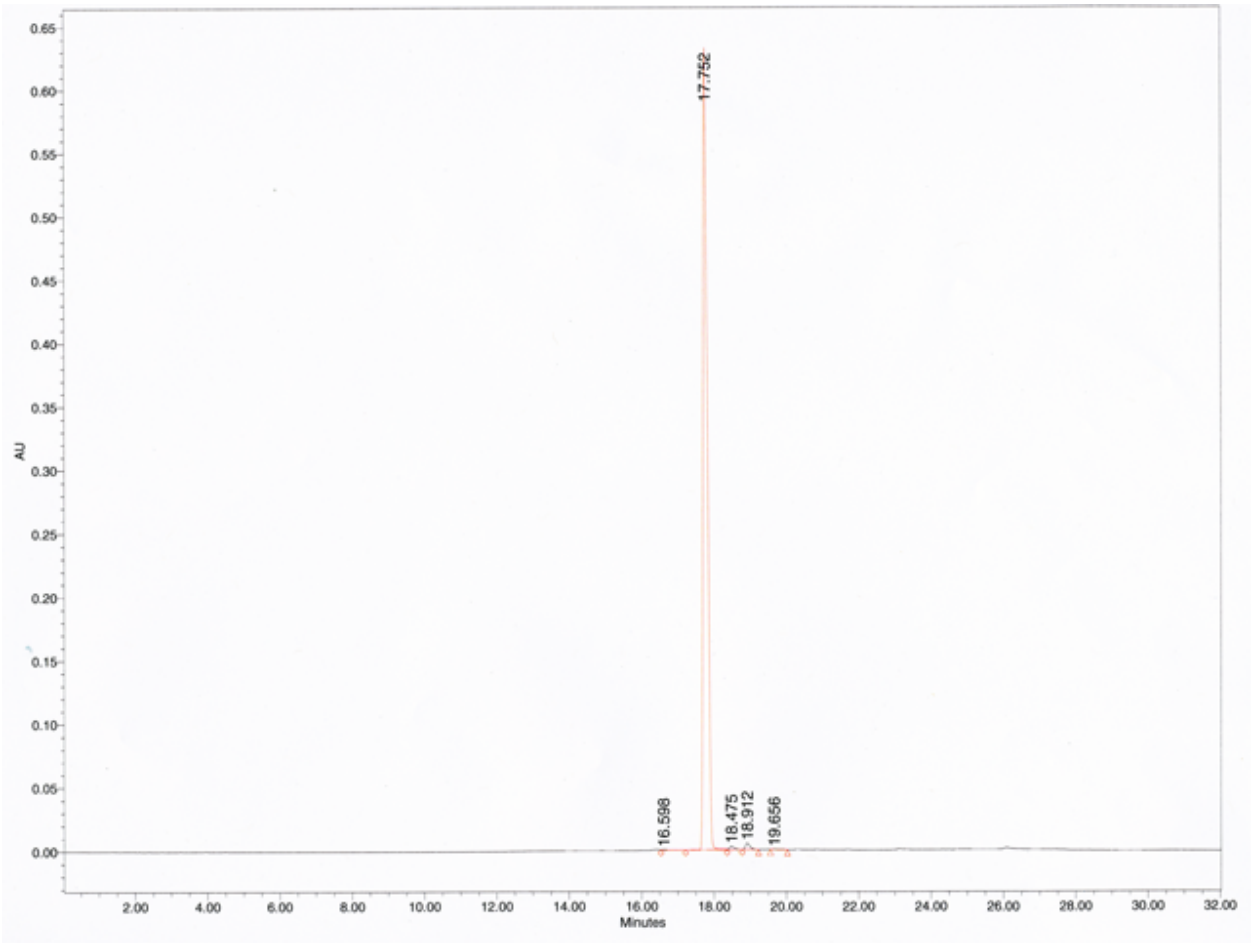




Peak	Retention Time (min)	Area (mV*sec)	% Area	Height (mV)
1	4.386	5642	0.04	563
2	6.875	180530	1.24	14393
3	15.243	21025	0.14	3817
4	15.392	9202	0.06	1524
5	16.582	10787	0.70	1507
6	**16.993**	**14068452**	**96.75**	**2291477**
7	17.394	13684	0.09	1587
8	17.83	11710	0.08	1477
9	18.258	8340	0.06	869
10	18.487	18806	0.13	2734
11	18.817	6780	0.05	1180
12	19.6	3606	0.02	276
13	19.896	176578	1.21	24395
14	20.33	2324	0.02	288
15	20.841	3094	0.02	314

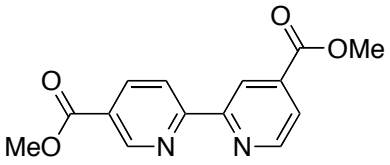
2.5.12.5. 5-Methoxycarbonyl-2-(6-methoxycarbonylpyridin-2-yl)pyridine *N*-Oxide

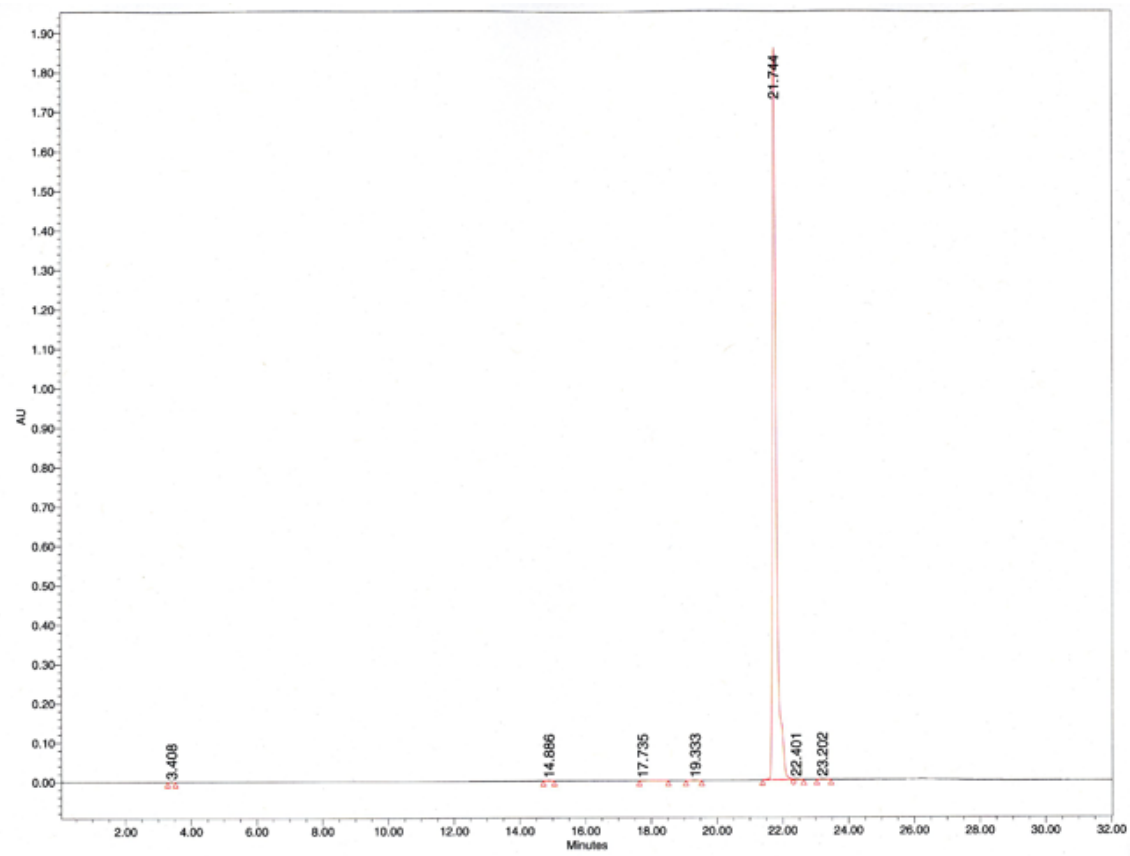




Peak	Retention Time (min)	Area (mV*sec)	% Area	Height (mV)
1	16.598	5066	0.10	390
2	**17.752**	**5074068**	**98.50**	**630451**
3	18.475	24831	0.48	2622
4	18.912	43364	0.84	4994
5	19.656	4232	0.80	320

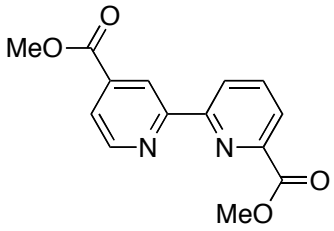
Dimethyl 2,2′-bipyridine-4,5′-dicarboxylate

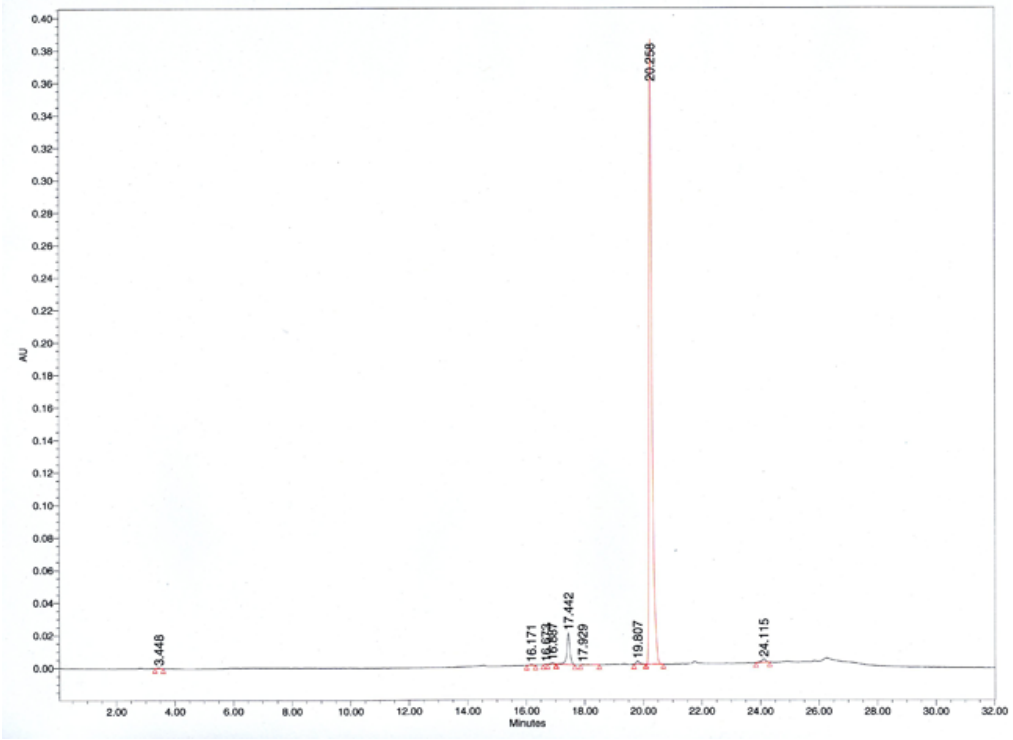




Peak	Retention Time (min)	Area (mV*sec)	% Area	Height (mV)
1	3.408	2855	0.02	480
2	14.886	6487	0.04	971
3	17.735	5271	0.04	342
4	19.333	5392	0.04	454
5	**21.744**	**14975268**	**99.65**	**1855605**
6	22.401	15565	0.10	2001
7	23.202	17522	0.12	1973

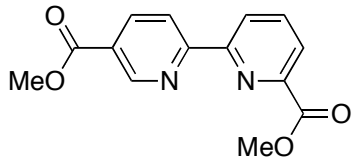
Dimethyl 2,2′-bipyridine-4,6′-dicarboxylate

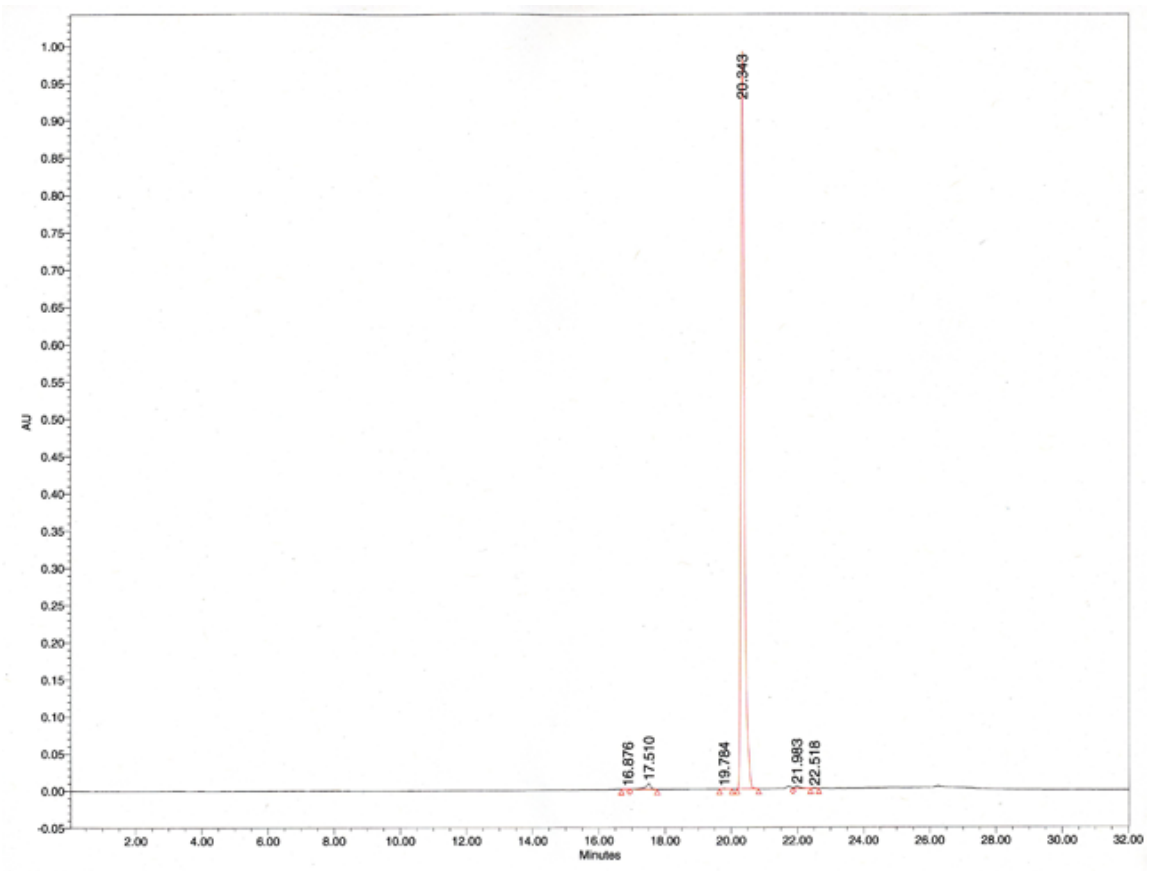




Peak	Retention Time (min)	Area (mV*sec)	% Area	Height (mV)
1	3.448	2487	0.09	430
2	16.171	3810	0.13	502
3	16.673	921	0.03	214
4	16.887	5841	0.20	876
5	17.442	143408	5.01	19470
6	17.929	3204	0.11	246
7	19.807	15865	0.55	2039
8	**20.258**	**2663252**	**93.12**	**382022**
9	24.115	21169	0.74	2136

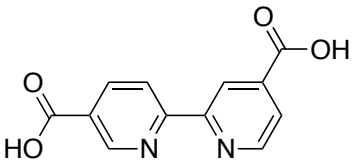
Dimethyl 2,2′-bipyridine-5,6′-dicarboxylate

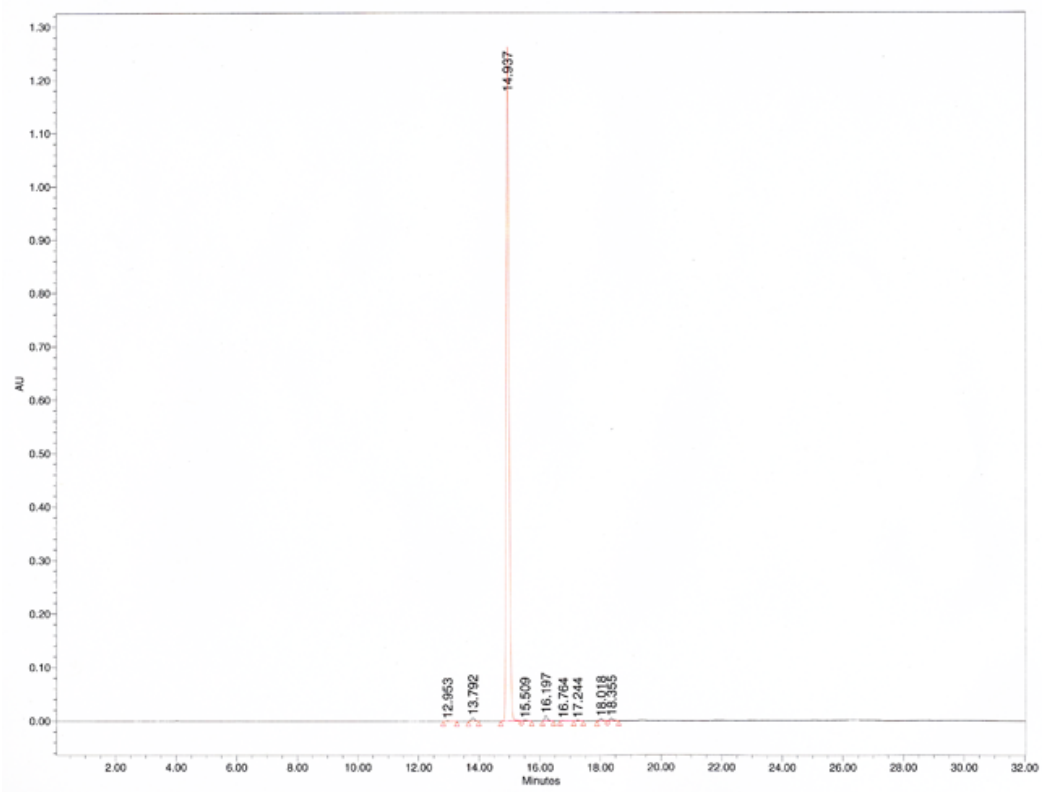




Peak	Retention Time (min)	Area (mV*sec)	% Area	Height (mV)
1	16.876	3500	0.05	327
2	17.51	91422	1.26	7940
3	19.784	3681	0.05	418
4	**20.343**	**7113302**	**98.13**	**985946**
5	21.983	32924	0.45	3450
6	22.518	3734	0.05	574

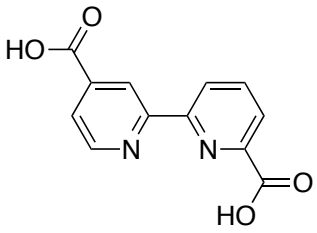
2,2′-Bipyridine-4,5′-dicarboxylic acid

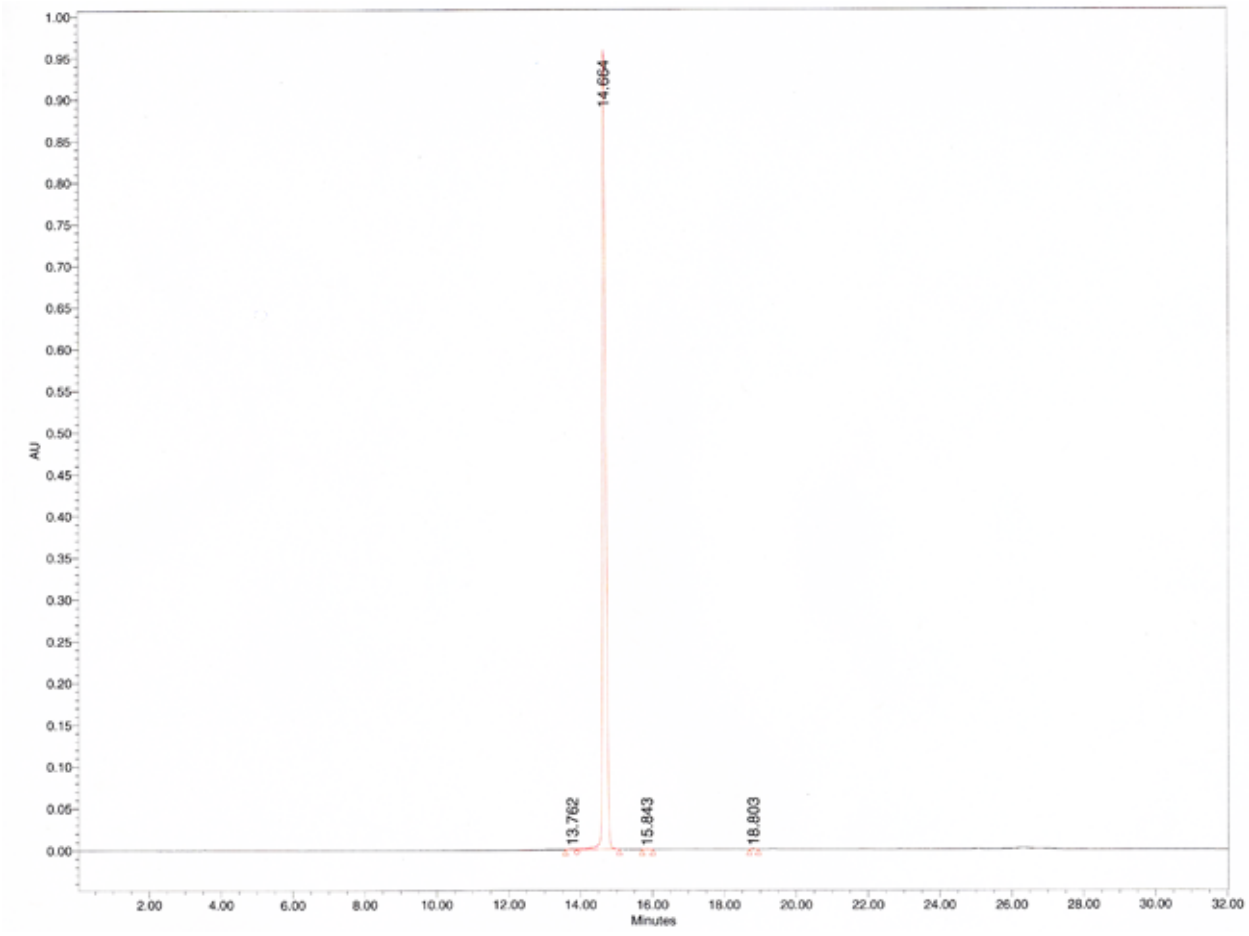




Peak	Retention Time (min)	Area (mV*sec)	% Area	Height (mV)
1	12.953	2349	0.03	342
2	13.792	41201	0.50	6230
3	**14.937**	**8057213**	**97.81**	**1275038**
4	15.509	9592	0.12	1195
5	16.197	59429	0.72	9616
6	16.764	4174	0.05	384
7	17.244	6066	0.07	927
8	18.018	25356	0.31	3755
9	18.355	32573	0.40	4703

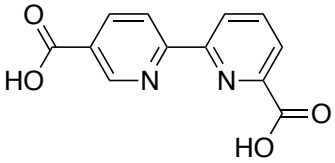
2,2′-Bipyridine-4,6′-dicarboxylic acid

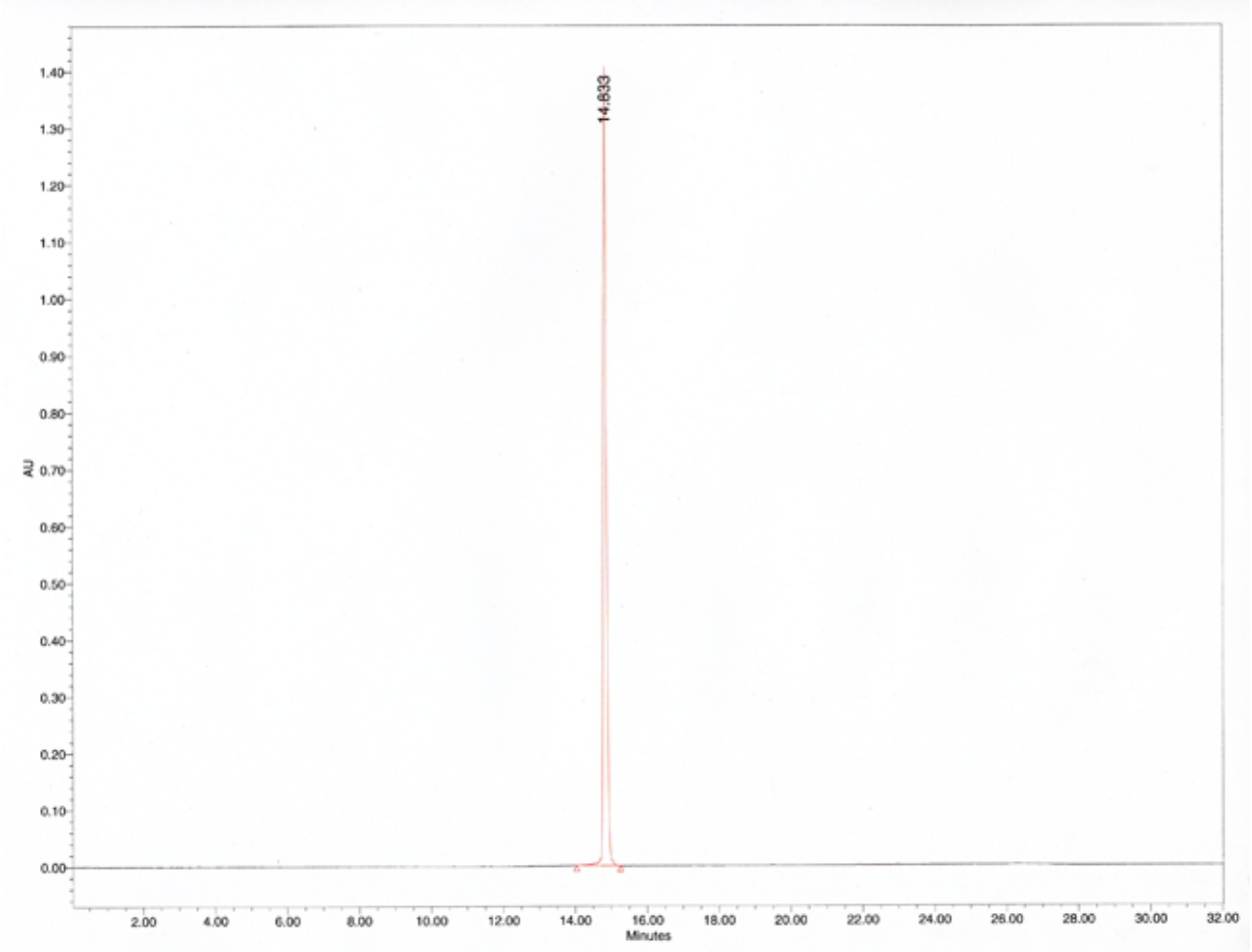




Peak	Retention Time (min)	Area (mV*sec)	% Area	Height (mV)
1	13.762	7264	0.12	768
2	**14.664**	**5934760**	**99.8**	**962317**
3	15.843	1879	0.03	305
4	18.803	2797	0.05	481

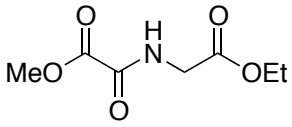
2,2′-Bipyridine-5,6′-dicarboxylic acid

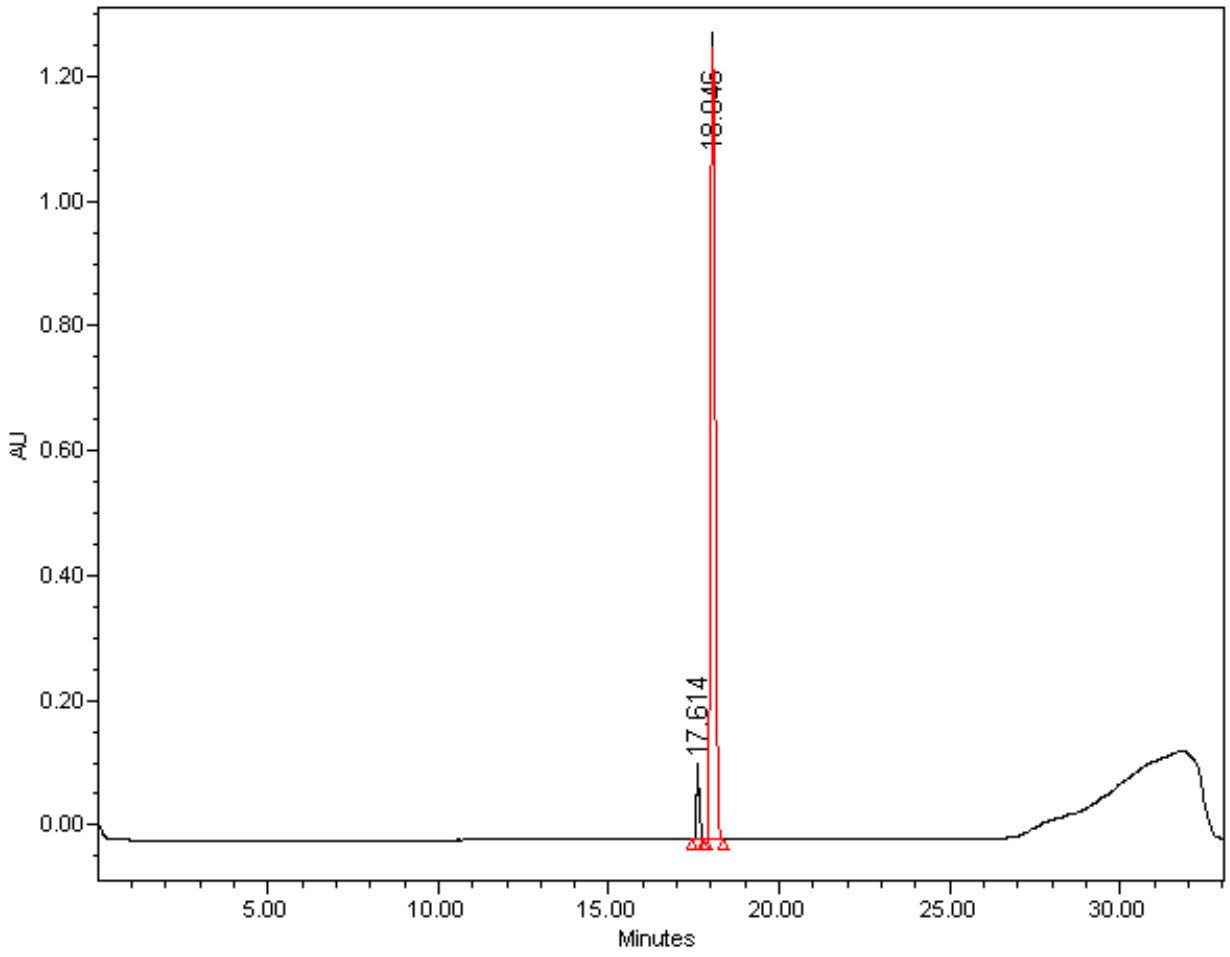




Peak	Retention Time (min)	Area (mV*sec)	% Area	Height (mV)
1	14.833	8630006	100	962317

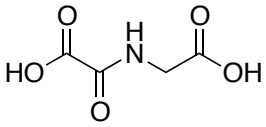
_N_-(Methoxyoxalyl)glycine ethyl ester

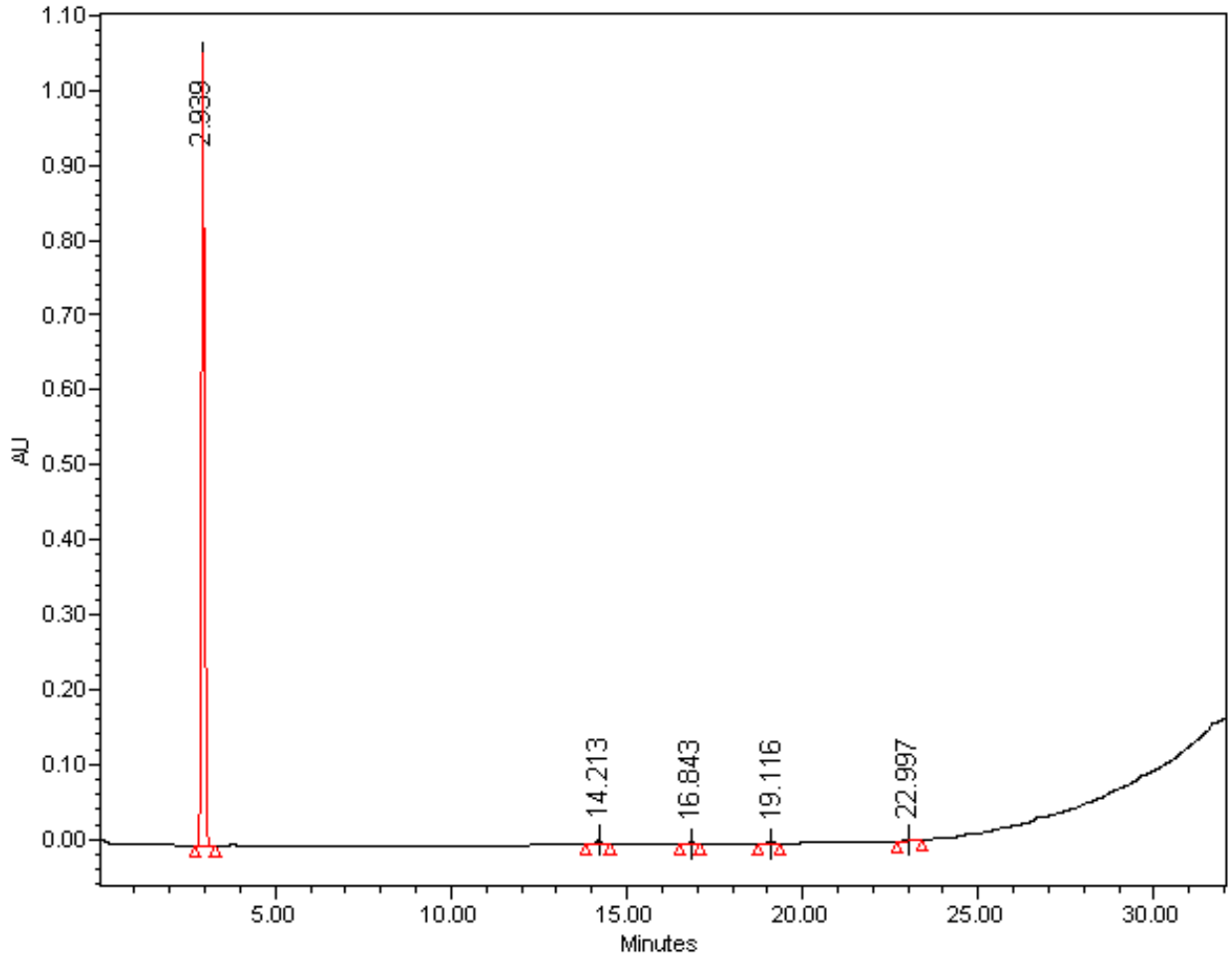




Peak	Retention Time (min)	Area (mV*sec)	% Area	Height (mV)
1	17.614	678345	5.32	97875
2	**18.046**	**12072069**	**94.68**	**1271789**

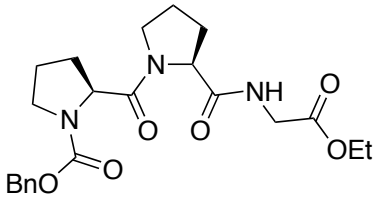
***N*-Oxalylglycine**

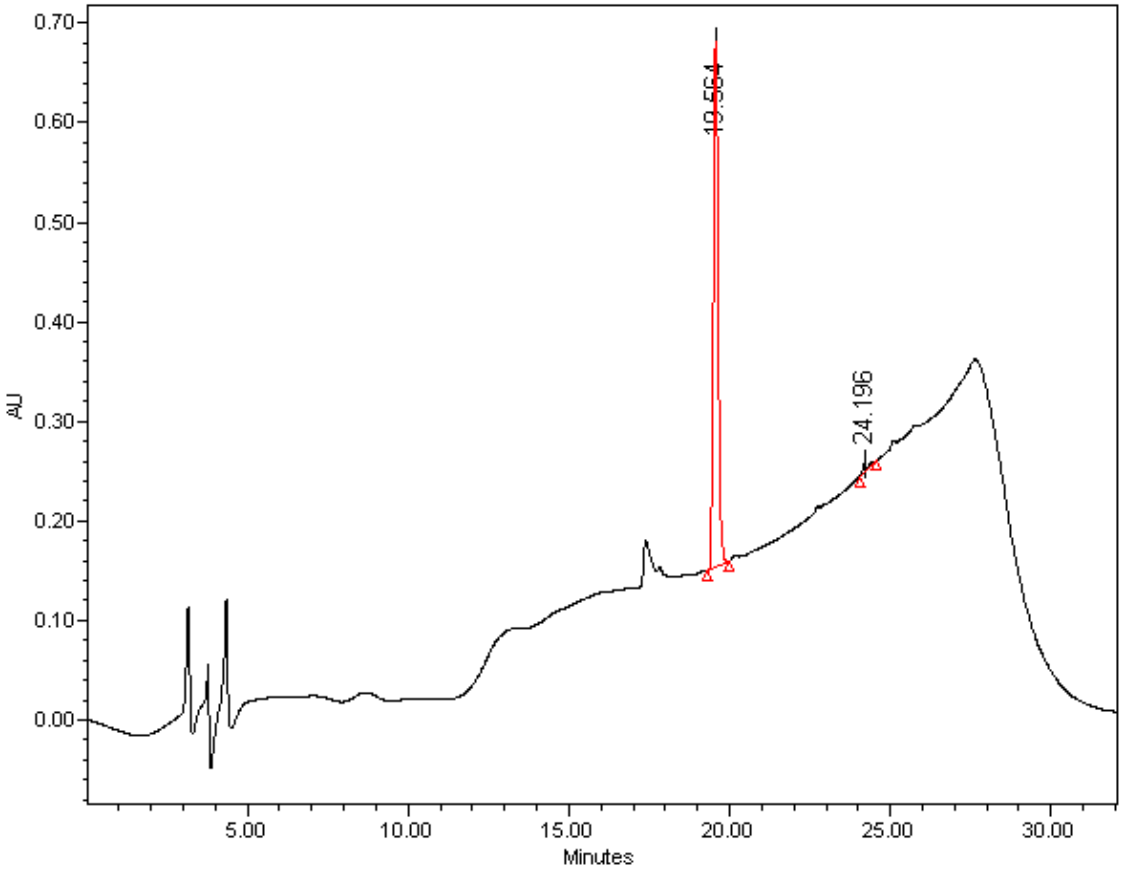




Peak	Retention Time (min)	Area (mV*sec)	% Area	Height (mV)
1	**2.939**	**6423296**	**98.12**	**1052708**
2	14.213	45684	0.7	6442
3	16.843	18876	0.29	1754
4	19.116	14825	0.23	733
5	22.997	43535	0.67	2475

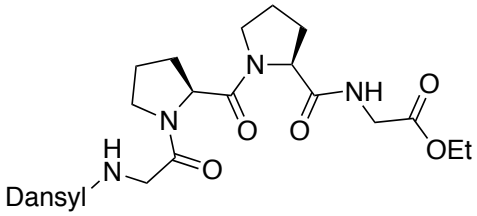
_N_-Benzyloxycarbonyl-(2_S_)-prolyl-(2_S_)-prolylglycine ethyl ester (CbzProProGlyOEt)

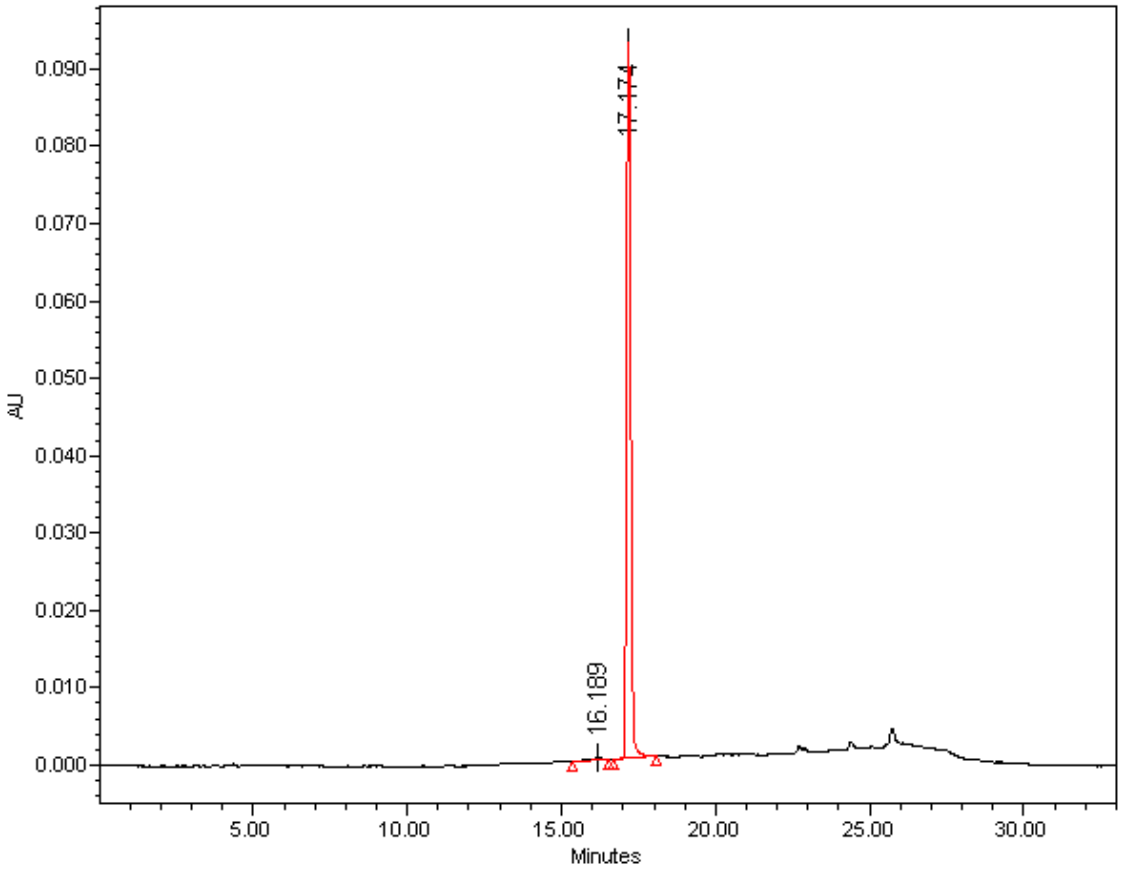




Peak	Retention Time (min)	Area (mV*sec)	% Area	Height (mV)
1	19.564	5386891	98.50	527526
2	**24.196**	**81974**	**1.50**	**8342**

_N_-Dansylglycyl-(2_S_)-prolyl-(2_S_)-prolylglycine ethyl ester (dansylGlyProProGlyOEt)



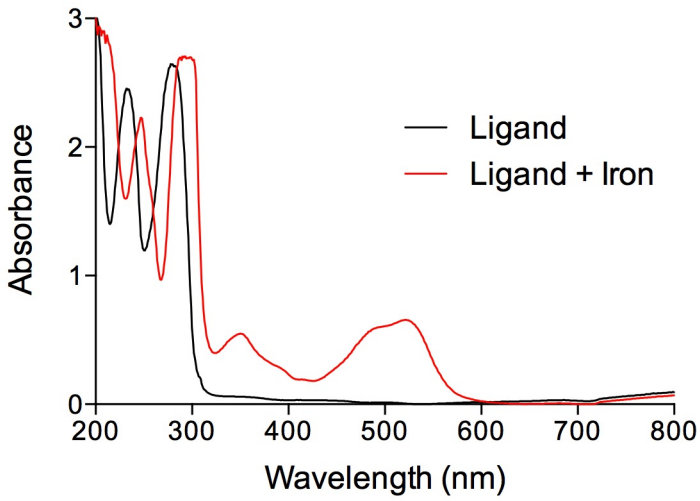


Peak	Retention Time (min)	Area (mV*sec)	% Area	Height (mV)
1	16.189	8436	1.00	460
2	**17.174**	**837020**	**99.00**	**92761**

2.5.13 UV/Vis Spectra of Bipyridine Ligands and Fe(II) Complexes

UV/Vis spectra of bipy in the presence and absence of Fe(II)

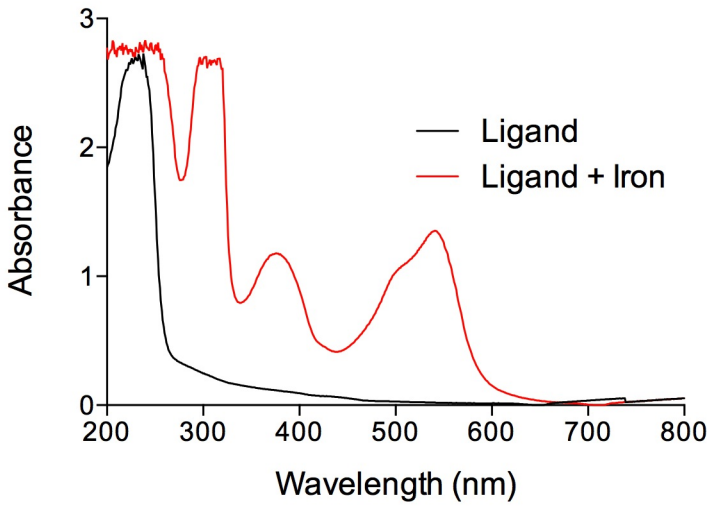
300 μM bipy ± 100 μM Fe(II)SO_4 in 10 mM sodium phosphate, pH 7.0



λ_{max} = 523 nm, ε_{523} = 9,000 ± 100 $M^{-1}cm^{-1}$

UV/Vis spectra of bipy44′DC in the presence and absence of Fe(II)

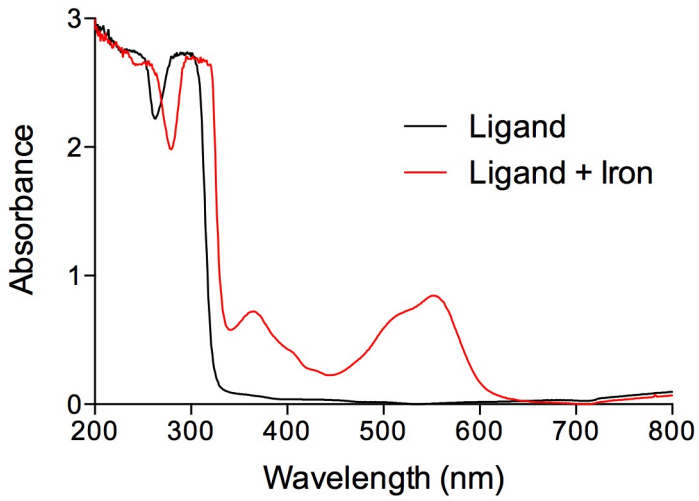
300 μM bipy44′DC ± 100 μM Fe(II)SO_4 in 10 mM sodium phosphate, pH 7.0



λ_{max} = 541 nm, ε_{541} = 15,090 ± 60 $M^{-1}cm^{-1}$

UV/Vis spectra of bipy45′DC in the presence and absence of Fe(II)

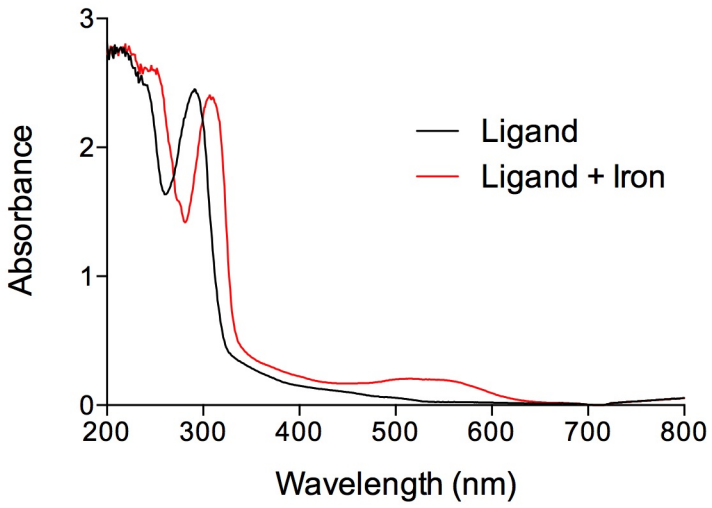
300 μM bipy45′DC ± 100 μM Fe(II)SO_4 in 10 mM sodium phosphate, pH 7.0



λ_{max} = 553 nm, ε_{553} = 11,200 ± 60 $M^{-1}cm^{-1}$

UV/Vis spectra of bipy46′DC in the presence and absence of Fe(II)

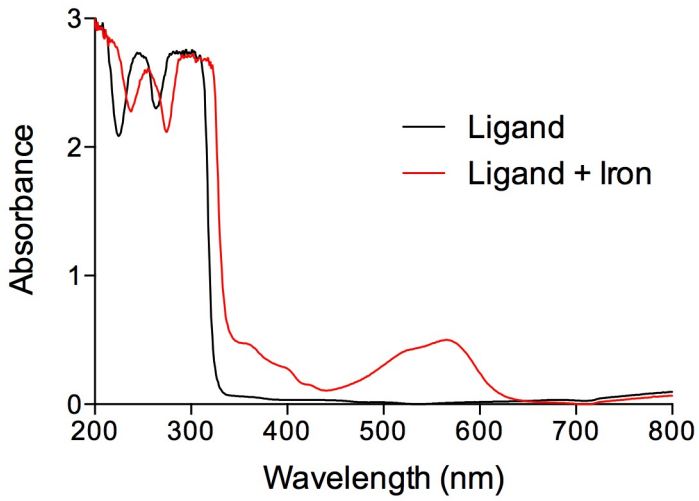
200 μM bipy46′DC ± 100 μM Fe(II)SO_4 in 10 mM sodium phosphate, pH 7.0



λ_{max} = 559 nm, ε_{559} = 1,690 ± 20 $M^{-1}cm^{-1}$

UV/Vis spectra of bipy55′DC in the presence and absence of Fe(II)

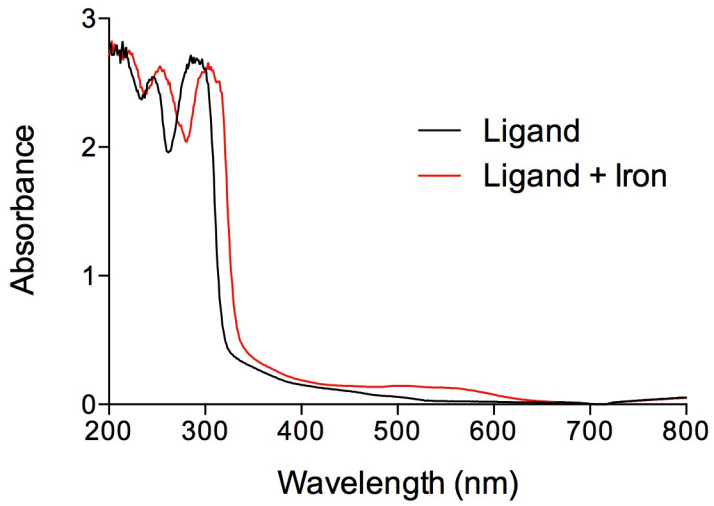
300 μM bipy55′DC ± 100 μM Fe(II)SO_4 in 10 mM sodium phosphate, pH 7.0



λ_{max} = 566 nm, ε_{566} = 7,500 ± 100 $M^{-1}cm^{-1}$

UV/Vis spectra of bipy56′DC in the presence and absence of Fe(II)

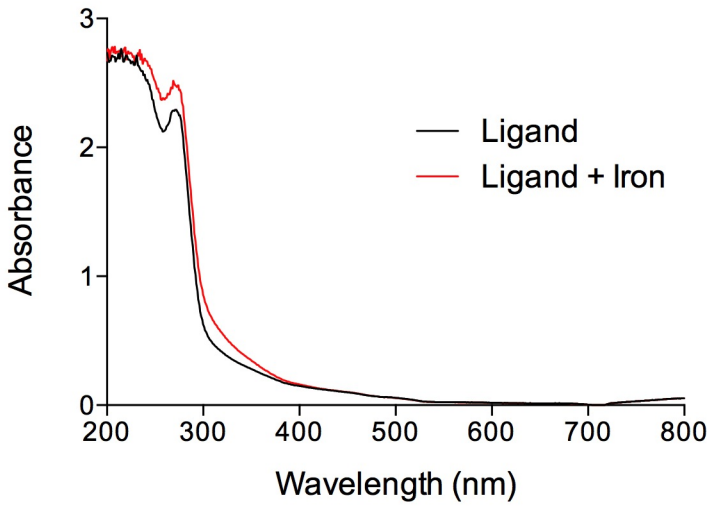
200 μM bipy56′DC ± 100 μM Fe(II)SO_4 in 10 mM sodium phosphate, pH 7.0



λ_{max} = 566 nm, ε_{566} = 1,220 ± 30 $M^{-1}cm^{-1}$

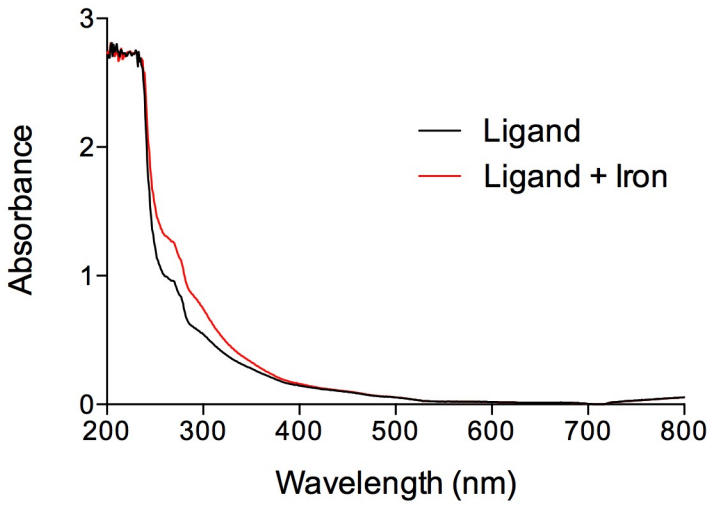
UV/Vis spectra of bipy33′DC in the presence and absence of Fe(II)

300 μM bipy33′DC ± 100 μM Fe(II)SO_4 in 10 mM sodium phosphate, pH 7.0



UV/Vis spectra of bipy66′DC in the presence and absence of Fe(II)

300 μM bipy66′DC ± 100 μM Fe(II)SO_4 in 10 mM sodium phosphate, pH 7.0



2.6. Acknowledgments

This work was supported by Grant R01 AR044276 (NIH). J. D. V. was supported by Molecular Biosciences Training Grant T32 GM007215 (NIH) and a fellowship from the Department of Biochemistry at the University of Wisconsin–Madison. This study made use of the National Magnetic Resonance Facility at Madison, which is supported by Grant P41 GM103399 (NIH). The Micromass LCT® mass spectrometer was obtained with support from Grant CHE-9974839 (NSF).

2.7 Author Contributions

All experiments were carried out by J.D.V. Both J.D.V. and R.T.R. contributed to the design of experiments, analysis of data, and preparation of the manuscript and figures.

Table 2.1 Iron-binding parameters of bipyridine ligands

Compound	λ_{max} of Complex (nm)[a]	Stoichiometry (Ligand:Fe)[b]	Fe_{20}-EC_{50} (μM)[c]
bipy	523	3:1	43 ± 2
bipy33′DC	ND[d]	—	—
bipy44′DC	541	3:1	38 ± 1
bipy45′DC	553	3:1	78 ± 2
bipy46′DC	559	2:1	26 ± 1
bipy55′DC	566	3:1	112 ± 3
bipy56′DC	566	2:1	27 ± 1
bipy66′DC	ND[d]	—	—

[a]Determined by spectrophotometry.
[b]Determined by Job's method.[143,144]
[c]Determined by titration of 20 μM $FeSO_4$. Values are the mean (± SE) of three replicates.
[d]Not detected.

Table 2.2 Inhibition constants for CP4H

Compound	IC_{50} (μM)[a]	K_i (μM)[b]
25PDC	15 ± 1	2.6 ± 0.2
bipy45′DC	1.2 ± 0.1	0.20 ± 0.02
bipy55′DC	0.84 ± 0.08	0.14 ± 0.01

[a]Interpolated from dose–response curves. Values are the mean (± SE) of three replicates.
[b]Estimated with the Cheng–Prusoff equation. Values are the mean (± SE) of three replicates.

Scheme 2.1 Synthesis of bipyridinedicarboxylates examined in this work

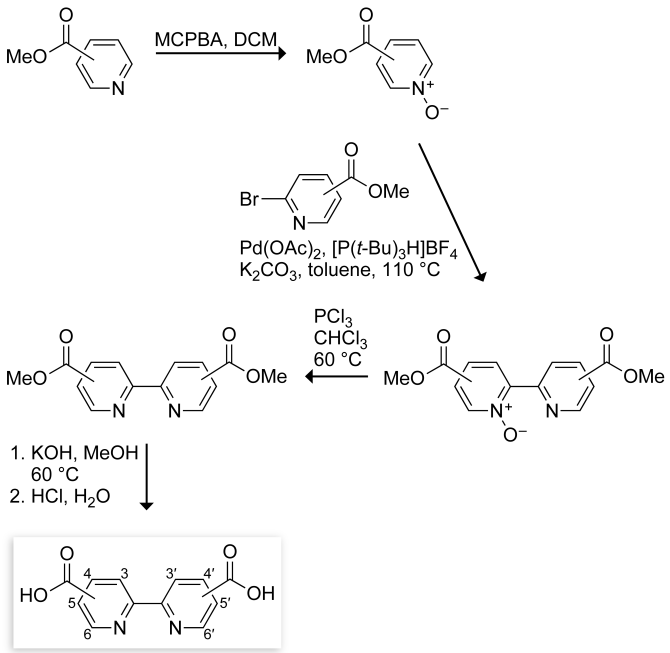


Figure 2.1 The reaction catalyzed by CP4Hs

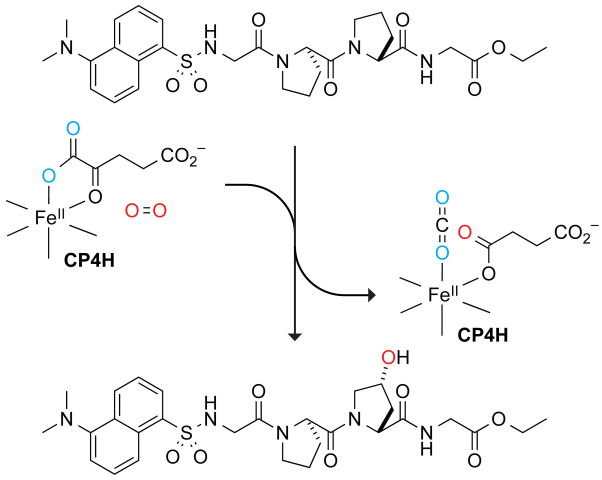


Figure 2.1 The reaction catalyzed by CP4Hs. In an Fe(II)- and AKG-dependent manner, CP4Hs catalyze the hydroxylation of specific Pro residues in collagenous peptides to form Hyp residues. The model peptide substrate used herein is dansylGlyProProGlyOEt, which is hydroxylated to form dansylGlyProHypGlyOEt.

Figure 2.2 Examples of known CP4H inhibitors

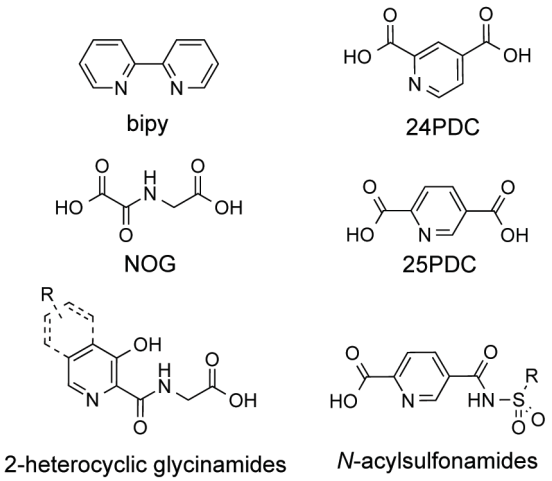


Figure 2.2. Examples of known CP4H inhibitors. Like most FAKGDs, CP4Hs are inhibited by iron chelators (like bipy) and simple AKG mimics (like NOG, 24PDC, and 25PDC).

Figure 2.3 Structure of 2,2′-bipyridinedicarboxylates examined in this work

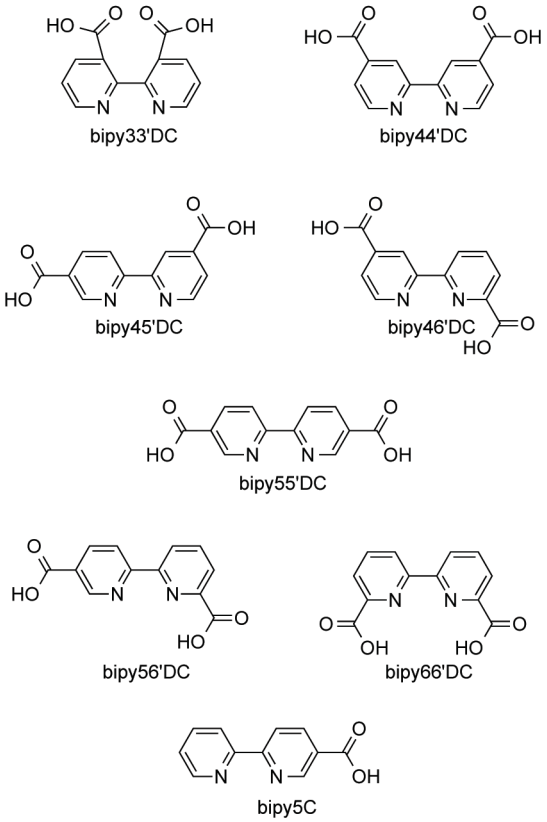


Figure 2.3. Structure of 2,2′-bipyridinedicarboxylates examined in this work and a 2,2′-bipyridinecarboxylate.

Figure 2.4 Ligand titrations for 2,2′-bipyridine complexes

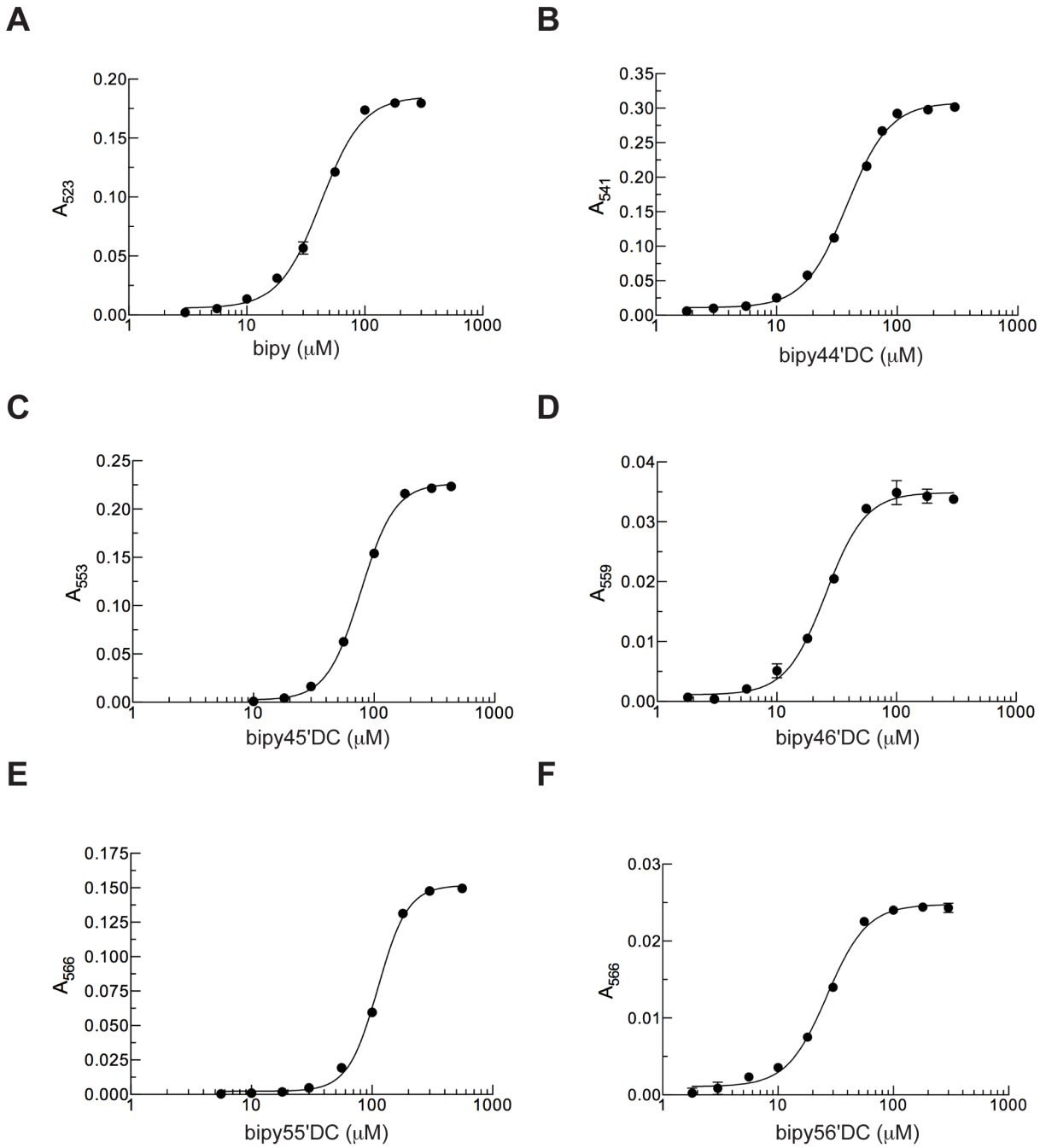


Figure 2.4. Ligand titrations for 2,2′-bipyridine complexes. Experiments were carried out as described in the Experimental Section above, and dose-response plots were prepared for (A) Fe(II)·(bipy)$_3$, (B) Fe(II)·(bipy44′DC)$_3$, (C) Fe(II)·(bipy45′DC)$_3$, (D) Fe(II)·(bipy46′DC)$_2$, (E) Fe(II)·(bipy55′DC)$_3$, and (F) Fe(II)·(bipy56′DC)$_2$. Plots were fit using the dose-response equation, and the associated Fe_{20}-EC_{50} values are reported above in Table 2.1. Correlation coefficient (R^2) values were greater than 0.997 in all cases.

Figure 2.5 Job plots of Fe(II) binding by 2,2′-bipyridine ligands

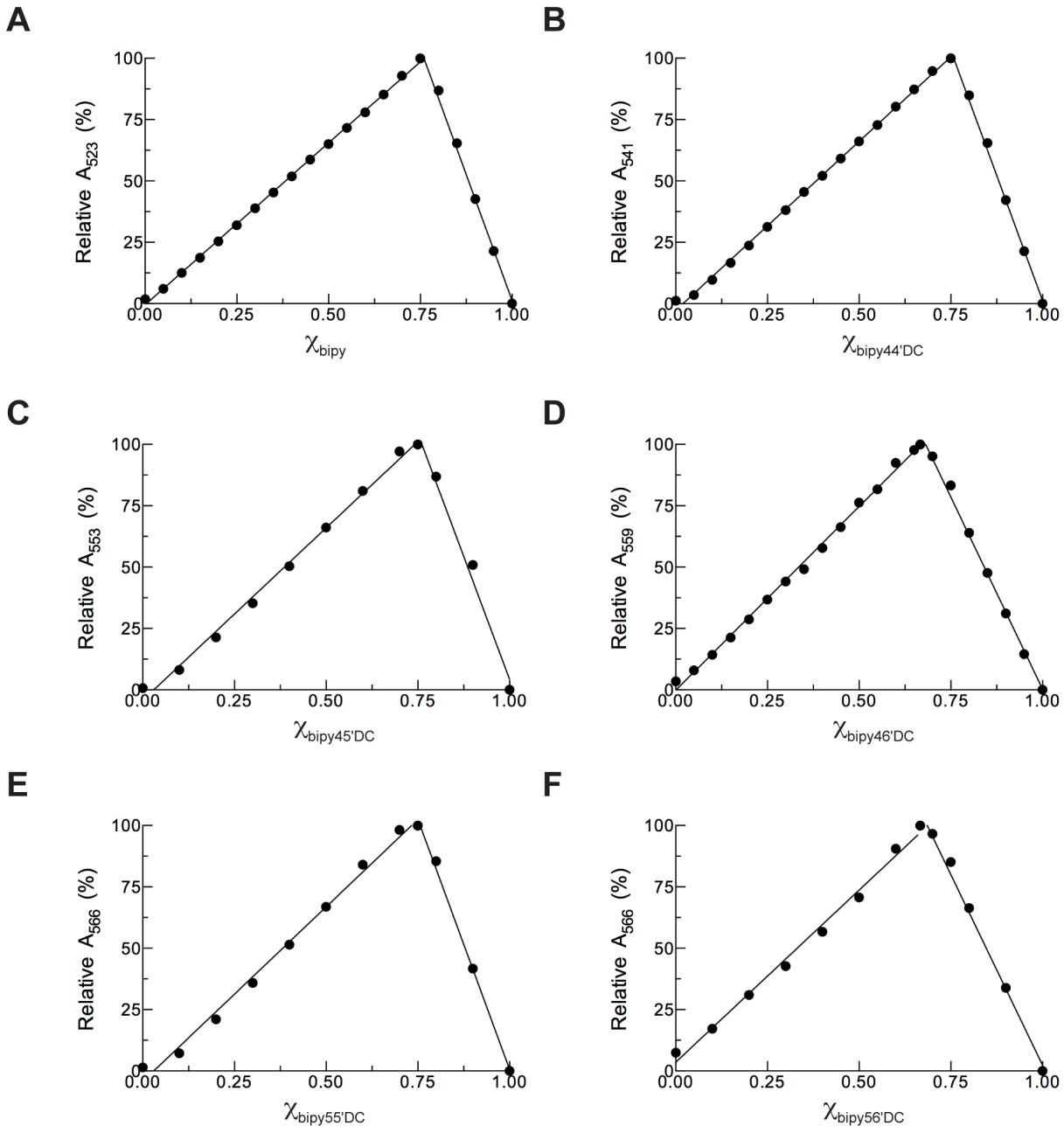


Figure 2.5. Job plots of Fe(II) binding by 2,2′-bipyridine ligands. Experiments were carried out as described in the Experimental Section above, and Job plots were prepared for (A) Fe(II)·$(bipy)_3$, (B) Fe(II)·$(bipy44'DC)_3$, (C) Fe(II)·$(bipy45'DC)_3$, (D) Fe(II)·$(bipy46'DC)_2$, (E) Fe(II)·$(bipy55'DC)_3$, and (F) Fe(II)·$(bipy56'DC)_2$. χ-values at the point intersection of (A) 0.76 ± 0.03, (B) 0.76 ± 0.03, (C) 0.76 ± 0.07, (D) 0.68 ± 0.02, (E) 0.75 ± 0.05, and (F) 0.69 ± 0.04 were used to calculate the stoichiometry for the respective Fe(II) complexes (reported above in Table 2.1).

Figure 2.6 Inhibition of human CP4H1 *in vitro* by 2,2′-bipyridinedicarboxylates

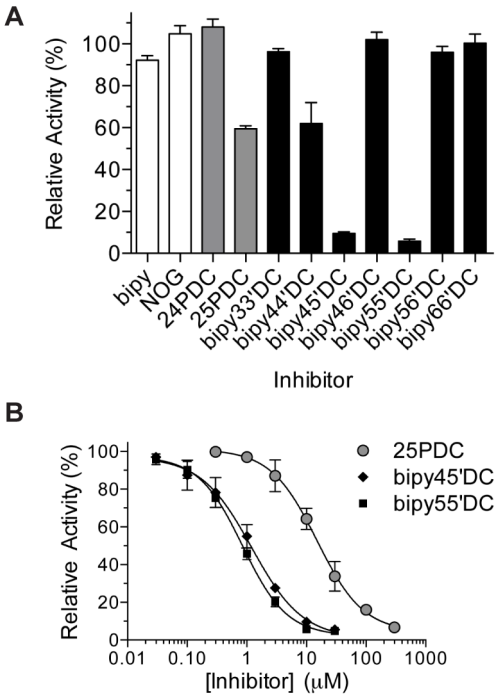


Figure 2.6. Inhibition of human CP4H1 *in vitro* by 2,2′-bipyridinedicarboxylates. (A) To mitigate the effect of iron sequestration, all compounds were initially screened at a concentration of 10 μM in the presence of excess Fe(II) as described in the Experimental Procedures section. Relative activity values are reported as the mean (± SE) of three independent experiments. (B) Dose–response curves for the most potent inhibitors identified in panel A were determined as described in the Experimental Procedures section. Individual points represent the mean (± SE) of three independent experiments. Data were fitted to the dose–response equation to determine IC_{50} values.

Figure 2.7 Inhibition of human CP4H1 by bipy44′DC

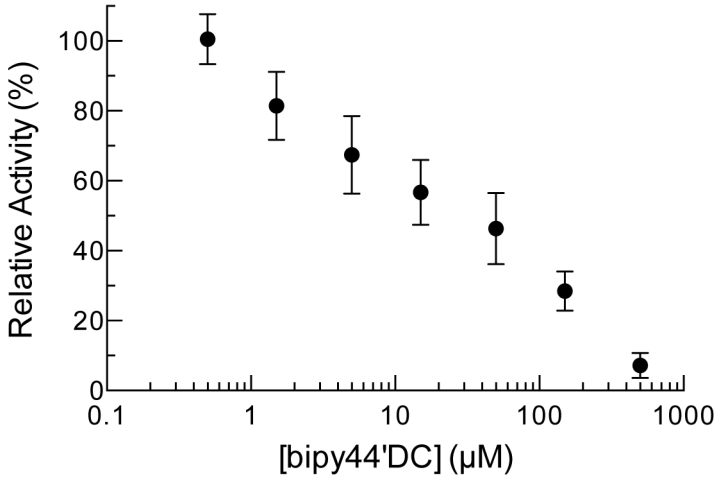


Figure 2.7. Inhibition of human CP4H1 by bipy44′DC. The dose–response curve for bipy44′DC was determined as described in the Experimental Section above. Individual points represent the mean (± SE) of three independent experiments.

Figure 2.8 Lineweaver–Burke analyses for inhibition of human CP4H1 bipy45′DC and bipy55′DC

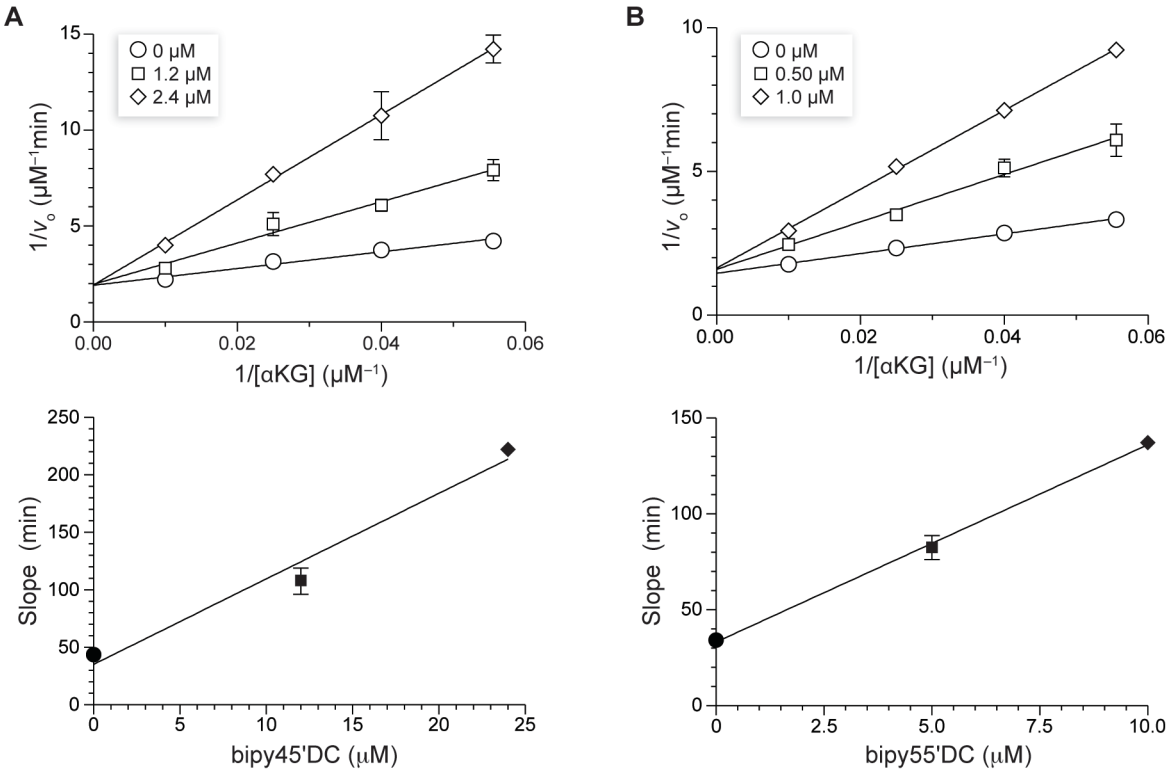


Figure 2.8. Lineweaver–Burke analyses for inhibition of human CP4H1 bipy45′DC and bipy55′DC. Lineweaver-Burke plots for (a) bipy45′DC and (b) bipy55′DC suggest competitive inhibition with respect to the AKG cosubstrate, as evidenced by the intersecting patterns that converge on the ordinate and linearity in the slope replots (shown below the corresponding Lineweaver-Burke plot). The rate of CP4H catalysis with increasing AKG concentration (18–100 μM) was determined in the presence of fixed inhibitor concentrations (specified above) as described in the Experimental Procedures section. Individual points represent the mean (± SE) of three independent experiments.

Figure 2.9 Inhibition of human PHD2 *in vitro* by 2,2′-bipyridinedicarboxylates

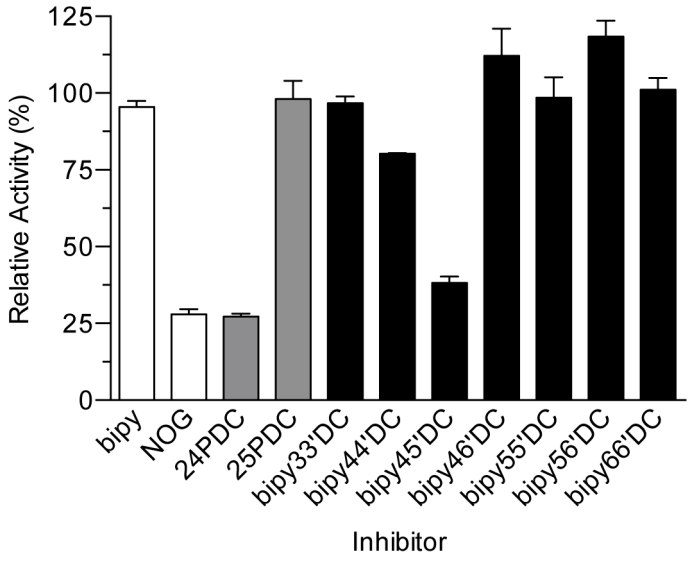


Figure 2.9. Inhibition of human PHD2 *in vitro* by 2,2′-bipyridinedicarboxylates. All compounds were screened at a concentration of 10 μM in the presence of excess Fe(II) as described in the Experimental Procedures section. Relative activity values are reported as the mean (± SE) of three independent experiments.

Figure 2.10 Schematic models of 2,2′-bipyridinedicarboxylate complexes with human CP4H1 and human PHD2

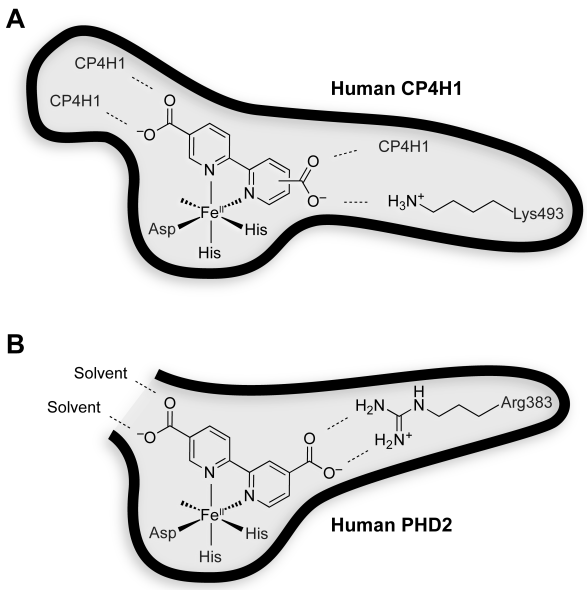
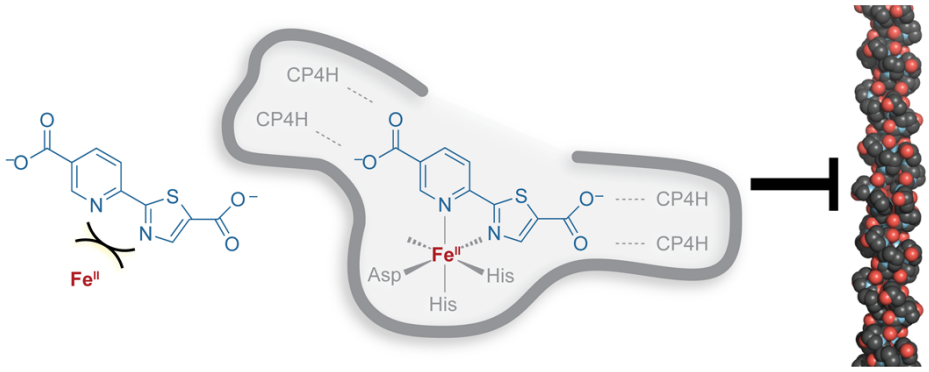


Figure 2.10. Schematic models of 2,2′-bipyridinedicarboxylate complexes with human CP4H1 and human PHD2. (A) Our data suggests that human CP4H1 can bind two different 2,2′-bipyridinedicarboxylate geometries in the conventional AKG binding mode, where the enhanced potency of these inhibitors stems from additional enzymic interactions in the distal active site, which has yet to be characterized. Fe(II) is likely chelated by Asp414, His412, and His483.[87] (B) Unlike CP4H, PHD2 accommodates only bipy45′DC in the AKG binding pocket. That finding and the similar potency of this compound compared to simple AKG mimics (*e.g.*, NOG and 24PDC) are consistent with the formation of hydrogen bonds between the second carboxyl group and solvent. Fe(II) is chelated by Asp315, His313, and His374.[67]

CHAPTER 3

Selective Inhibition of Collagen Prolyl 4-Hydroxylase in Human Cells*



*This chapter has been published in part, under the same title. Reference: Vasta, J.D., Andersen, K.A., Deck, K.M., Nizzi, C.P., Eisenstein, R.S., & Raines, R.T. Selective Inhibition of Collagen Prolyl 4-Hydroxylase in Human Cells. This work was conducted in collaboration with professor R.S. Eisenstein. J.D. Vasta carried out or contributed to carrying out all experiments. K.A. Andersen contributed to carrying out all experiments relating to human cell culture. K.M. Deck, C.P. Nizzi, and R. S. Eisenstein contributed to carrying out experiments relating to the function of IRPs. All authors contributed to the design of experiments, analysis of data, and preparation of the manuscript and figures.

3.1 Abstract

Collagen is the most abundant protein in animals. Its overproduction is associated with fibrosis and cancer metastasis. The stability of collagen relies on post-translational modifications, the most prevalent being the hydroxylation of collagen strands by collagen prolyl 4-hydroxylases (CP4Hs). Catalysis by CP4Hs enlists an iron cofactor to convert proline residues to 4-hydroxyproline residues, which are essential for the conformational stability of mature collagen. Ethyl 3,4-dihydroxybenzoate (EDHB) is commonly used as a "P4H" inhibitor in cells, but suffers from low potency, poor selectivity, and off-target effects that cause iron deficiency. Dicarboxylates of 2,2′-bipyridine are among the most potent known CP4H inhibitors but suffer from a high affinity for free iron. A screen of biheteroaryl compounds revealed that replacing one pyridyl group with a thiazole moiety retains potency and enhances selectivity. A diester of 2-(5-carboxythiazol-2-yl)pyridine-5-carboxylic acid is bioavailable to human cells and inhibits collagen biosynthesis at concentrations that neither cause general toxicity nor disrupt iron homeostasis. These data anoint a potent and selective probe for CP4H and a potential lead for the development of a new class of antifibrotic and antimetastatic agents.

3.2 Introduction

Collagen is the principal component of bone, connective tissues, and the extracellular matrix in animals.[1] The overproduction of collagen is associated with a variety of diseases, including fibrotic diseases[33] and cancers.[28,29,46,47,73] The stability of collagen relies on posttranslational modifications that occur throughout the secretory pathway.[2] By far the most prevalent of these modifications is the hydroxylation of collagen strands by collagen prolyl 4-hydroxylases (CP4Hs), which are Fe(II)- and α-ketoglutarate (AKG)-dependent dioxygenases (FAKGDs) located in the lumen of the endoplasmic reticulum.[4] Catalysis by CP4Hs converts (2*S*)-proline (Pro) residues in protocollagen strands into (2*S*,4*R*)-4-hydroxyproline (Hyp) residues (Figure 3.1A), which are essential for the conformational stability of mature collagen triple helices.[5] Importantly, CP4Hs are validated targets for treating both fibrotic diseases[30] and metastatic breast cancer.[29]

Like all enzymes of the FAKGD superfamily, catalysis by CP4Hs requires Fe(II) and the cosubstrates AKG and dioxygen.[56,97,145] The Fe(II) is bound by a conserved His-X-Asp/Glu…X_n…His motif, and AKG chelates to enzyme-bound Fe(II) using its C-1 carboxylate and C-2 keto groups, while the C-5 carboxylate group forms hydrogen bonds and engages in Coulombic interactions with a basic residue (typically, arginine or lysine). All FAKGDs are believed to effect catalysis through a similar two-stage mechanism in which AKG is first decarboxylated oxidatively to generate a highly reactive Fe(IV)=O species (ferryl ion), which effects the hydroxylation of a hydrocarbon via a radical rebound process.[56,97,145]

In vertebrates, CP4Hs exist as $\alpha_2\beta_2$ tetramers. In these tetramers, the α-subunit contains the catalytic and substrate-binding domains, and the β-subunit is protein disulfide isomerase, which is a multifunctional protein that is responsible for maintaining the α-subunit in a soluble and active conformation.[4] Three isoforms of the α-subunit, α(I), α(II), and α(III), have been identified in humans.[4] All α-subunit isoforms form tetramers with the β-subunit, which we refer to herein as the CP4H1, CP4H2, and CP4H3 holoenzymes. CP4H1 is the most prevalent of the isoforms, and has been characterized most extensively. Whereas the structure of the tetrameric complex is unknown, those of the individual domains of the α(I)-subunit have provided insight into the interaction of CP4Hs with the protocollagen substrate, as well as the dimerization of the α(I)-subunits to facilitate formation of the tetramer.[84-86,146]

The development of CP4H inhibitors has been of interest since the mid-1970s. Like many FAKGDs, human CP4Hs are inhibited by AKG mimics,[98] such as *N*-oxalyl glycine (NOG), pyridine-2,4-dicarboxylic acid (24PDC), pyridine-2,5-dicarboxylic acid (25PDC), and 3,4-dihydroxybenzoic acid (DHB),[67] as well as by simple metal chelators, such as 2,2′-bipyridine (bipy) (Figure 3.1B). These compounds suffer from low potency in cellular assays, insufficient selectivity for CP4H, and intolerable cytotoxicity.[112,118] Still, the ethyl ester of DHB (that is, EDHB) is often used as a cellular "P4H" inhibitor,[131,147] even though DHB is not selective for CP4H compared to other FAKGDs, requires high dosing, and leads to an iron-deficient phenotype.[131]

The most potent inhibitors of human CP4Hs identified to date are bipyridinedicarboxylic acids (bipyDCs) (Figure 3.1B). Two bipyDCs have high potency: 2,2′-bipyridine-4,5′-

dicarboxylic acid (bipy45′DC)[148] and 2,2′-bipyridine-5,5′-dicarboxylic acid (bipy55′DC).[106] Both of these bipyDCs inhibit human CP4H competitively with respect to AKG and bind selectively to human CP4H1 compared to PHD2, another human P4H.[148]

An intrinsic property of bipyDCs limits their utility in a biological context. Like bipy itself, bipyDCs form tight complexes with free iron,[148] which is the dominant metal in life.[149] We hypothesized that replacing the bipy core with another biheteroaryl scaffold could lead to compounds that retain the potency and selectivity of bipy45′DC and bipy55′DC, but not their high affinity for free iron. Accordingly, we synthesized bipy analogs that retain one pyridyl group but replace the second with six distinct 5- or 6-membered nitrogen-containing heterocycles: pyrazine, pyridazine, oxazole, pyrazole, imidazole, pyrrole, and thiazole (Figure 2). We find that an asymmetric scaffold containing a thiazolyl group has especially desirable attributes as an inhibitor of human CP4H.

3.3 Results and Discussion

Biheteroaryl Compounds

To begin our analysis, we prepared a library of six biheteroaryl compounds (Figure 3.2) and evaluated them as iron chelators *in vitro*. The library was assembled either from commercial vendors (pypyr and pyim) or by synthesis using palladium-catalyzed cross-coupling reactions (pypyraz, pypyrid, pyox, and pythi). To evaluate the iron affinity of these compounds *in vitro*, we performed titration experiments to determine the half-maximal concentration required to

form a complex with 20 μM Fe(II) at pH 7.0. The ensuing Fe_{20}-EC_{50} parameter provides a metric for iron affinity.[148] With the exception of pyox, all members of the library were observed to form distinct colored Fe(ligand)$_3^{2+}$ complexes under these conditions (see sections 3.5.21 and 3.5.22).

The Fe_{20}-EC_{50} values of the biheteroaryl compounds varied by more than two orders of magnitude (Figure 3.3; Table 3.1). All, however, were significantly greater than that of bipy or bipyDCs.[148] The Fe_{20}-EC_{50} value appeared to rely on both the ring size and pK_a of the conjugate acid of the alternative heterocycle, as scaffolds with 5-membered rings were substantially weaker chelators than were those with 6-membered rings and pK_a value correlated positively with iron affinity.

Biheteroaryl Dicarboxylates *in Vitro*

Inspired by the low affinity of the biheteroaryl compounds for free iron, we prepared a small library of biheteroaryl dicarboxylates (Figure 3.2) and interrogated those compounds as both iron chelators and inhibitors of human CP4H. With the exceptions of pyimDC (which was synthesized via a classical route to substituted imidazoles[150]) and pypyrDC (which was synthesized by a 1,3-dipolar cycloaddition of a 2-ethynylpyridine with ethyl diazoacetate[151]), all of the biheteroaryl dicarboxylates were synthesized by either palladium-catalyzed direct arylation of an appropriate heterocycle (pypyrazDC, pypyridDC, pyoxDC, pyoxDC*, pythiDC or pythiDC*) with a functionalized 2-bromopyridine[135-137,152,153] or palladium-catalyzed oxidative cross-coupling (pypyrroleDC) between a functionalized pyridine *N*-oxide and an *N*-protected

pyrrole.[154] Typically, direct arylation using methyl- or ethyl-protected carboxylate esters allowed synthesis of the target compounds in 2–4 steps with an acceptable yield. For pyoxDC and pythiDC, cross-coupling yields using the typical inner-sphere base pivalic acid (PivOH) were prohibitively low (<5%, data not shown). We found that the addition of 1-adamantanecarboxylic acid rather than PivOH improved yields markedly and encourage the continued investigation of 1-adamantanecarboxylic acid as an inner-sphere base in palladium-catalyzed direct arylation reactions.

We investigated iron chelation by the biheteroaryl dicarboxylates in a manner similar to that for the parent scaffolds. To our surprise, we were not able to detect complex formation by spectrophotometry for any of the biheteroaryl dicarboxylates at concentrations up to 1 mM, suggesting that the affinity of these compounds for free iron would be negligible in a biological context. Previously, we reported that various bipyDCs have Fe_{20}-EC_{50} values that are similar to that of bipy itself,[148] so our discovery that biheteroaryl dicarboxylates investigated herein have an Fe_{20}-EC_{50} value >1 mM represents an improvement of at least an order of magnitude.

Next, we assessed the biheteroaryl dicarboxylates as inhibitors of human CP4H1. To separate any inhibitory effect that derives from iron sequestration rather than enzymic binding, we employed previously described assay conditions (10 μM compound and 50 μM $FeSO_4$) in which potent chelators like bipy do not cause inhibition.[148] In this initial screen (Figure 3.4), we found that some biheteroaryl dicarboxylates showed little or no inhibition of human CP4H1, consistent with the inability of their heteroatoms to participate in an enzymic interaction. (Both pypyridDC and pypyrDC even showed modest activation under these conditions by a mechanism

that is unclear.) Notably, we found that pyimDC, pyoxDC, and pythiDC were inhibitors of human CP4H1, with pyimDC and pythiDC demonstrating potency only a bit weaker than that of the bipyDCs. Importantly, the regioisomers pythiDC* and pyoxDC* did not show significant inhibition, suggesting that proper regiochemistry is essential for inhibition.

Unlike oxazole or thiazole, imidazole exists as two tautomers, one with a proton on N1 (as in the depiction of pyimDC in Figure 3.2) and another with a proton on N3. Although we did not observe the formation of a complex between pyimDC and free iron by spectrophotometry, we examined this issue more thoroughly. We found that pyimDC was able to deter the formation of the $Fe(bipy)_3^{2+}$ complex in a dose-dependent manner (Figure 3.5). Moreover, competition required a free carboxylate on the imidazole ring. These data are consistent with the formation of a $Fe(pyimDC)_2$ complex with N1 bound to iron. To eliminate this mode of binding, we synthesized NMe-pyimDC (Figure 3.2), which is an analogue of pyimDC that is methylated on N1. We found that NMe-pyimDC was able to deter the formation of the $Fe(bipy)_3^{2+}$ complex, but only at high concentrations (Figure 3.5). We also found that NMe-pyimDC is an inhibitor of human CP4H *in vitro*, but that its potency is less than that of pyimDC (Figure 3.4).

In subsequent dose–response experiments, we found the inhibition curves for pyimDC, pythiDC, and pyoxDC to be sigmoidal, yielding IC_{50} values in the low-micromolar range (Figure 3.6A). The potencies of pyimDC and pythiDC were approximately 10-fold greater than that of pyoxDC. A Lineweaver–Burke analysis of inhibition of CP4H by pythiDC demonstrated competitive inhibition with respect to the AKG cosubstrate (Figure 3.6B).

Like bipy45′DC and bipy55′DC,[106,148] pyimDC, pyoxDC, and pythiDC could bind in the AKG binding pocket and use their second carboxyl group to form additional interactions with residues in the active site of CP4H. If so, then the biheteroaryl dicarboxylates should exhibit structure–activity relationships similar to those of bipy45′DC and bipy55′DC for the inhibition of PHD2, another P4H enzyme that has been characterized extensively. We found that pyimDC, pythiDC, and pyoxDC inhibit human PHD2 only weakly (Figure 3.7), with pythiDC displaying especially low potency.

Biheteroaryl Dicarboxylates *in Cellulo*

Encouraged by the results *in vitro*, we next sought to determine if pyimDC and pythiDC could be suitable for use *in cellulo*. First, we determined the effect of these compounds on iron metabolism in human cells. Because human CP4Hs are validated therapeutic targets for breast cancer, we chose the MDA-MB-231 human breast cancer cell line as a primary model system.[29] We performed additional analyses in human embryonic kidney (HEK) cells. We assessed iron metabolism with immunoblots for ferritin, the transferrin receptor (TfR), and the transcription factor HIF-1α, all of which are known to give distinct phenotypes depending on the status of iron in a human cell.[131] More specifically, levels of ferritin and TfR are regulated by iron regulatory proteins 1 and 2 (IRPs), which have iron-dependent RNA-binding activity that modulates the expression of target genes at the level of translation.[155,156] The stability of HIF-1α is dependent on the prolyl 4-hydroxylase activity of PHD2, which is an activity that is inherently sensitive to

the iron status of the cell. Thus, iron-deficient cells exhibit ferritin levels that are lower and TFR and HIF-1α levels that are higher than those of untreated cells.

The esterification of carboxyl groups is a common strategy to generate effective prodrugs.[157,158] Ethyl dihydroxybenzoate (EDHB), an AKG mimic that is commonly used as a "P4H" inhibitor in cell culture models, served as our benchmark.[131,147] Calculations suggested that the diethyl esters of pyimDC and pythiDC would have log*P* values conducive to cellular uptake (Table 3.2), and we synthesized those two esters. Moreover, the iron affinity of these diethyl esters remained sufficiently low (see section 3.5.21 in the Experimental Procedure section and Figure 3.3), encouraging their use in cellular assays. Cultured MDA-MB-231 cells are known to secrete large amounts of collagen.[29] Due to the importance of CP4H-dependent hydroxylation for collagen stability, the level of collagen secreted by MDA-MB-231 cells is dependent on the activity of this enzyme. Thus, MDA-MB-231 cells are ideal for investigations of both iron deficiency and CP4H activity.

Toward this end, we treated MDA-MB-231 cells with biheteroaryl dicarboxylates, and assayed for cytotoxicity and indicators of iron deficiency. None of the esterified biheteroaryl dicarboxylates exhibited cytotoxic activity at high micromolar concentrations (see section 3.5.11 in the Experimental Procedures section). Cells treated with EDHB demonstrated a strong iron-deficient phenotype, as expected (Figures 3.8A and 3.8B). In contrast, cells treated with diethyl pythiDC appeared to be normal at concentrations as high as 500 μM. Interestingly, diethyl pyimDC showed an intermediate phenotype characterized by a significant decrease in ferritin levels but without an associated accumulation of TfR or HIF-1α. This phenotype was, however,

not observable at diethyl pyimDC concentrations ≤56 μM. The results in MDA-MB-231 cells were replicated in HEK cells. Again, treatment with diethyl pythiDC and low levels of diethyl pyimDC (as well as simply pyim and pythi; Figure 3.9) did not affect the level of TfR, HIF-1α, or ferritin (Figure 3.10).

Next, we assessed the function of IRPs in cells treated with the esterified biheteroaryl compounds. Using a radiolabeled RNA ligand, we performed electrophoretic mobility shift assays on IRPs from treated MDA-MB-231 cells. Whereas DFO, EDHB, and bipy all caused significant increases in RNA-binding by IRPs, diethyl pyimDC and diethyl pythiDC did not (Figure 3.11). This result is again consistent with these compounds having little effect on cellular iron levels.

Lastly, we examined the effect of treating MDA-MB-231 cells with the esterified biheteroaryl compounds on the cellular secretion of type I collagen (Figure 3.8C), which relies on CP4H activity.[29] We found that treatment with either diethyl pyimDC or diethyl pythiDC reduced the level of secreted collagen significantly (Figure 3.8C). Moreover, the efficacy of diethyl pyimDC or diethyl pythiDC was indistinguishable from that of diethyl bipy55′DC. No inhibition of catalysis by CP4H1 or PHD2 was observed *in vitro* with these diethyl esters at a concentration of 100 μM (data not shown), confirming that the inhibitory potential of these compounds is masked by esterification. Lastly, treatment with diethyl pyimDC and diethyl pythiDC did not appear to affect the levels of human CP4H1 itself (Figure 3.12), consistent with the observed reduction in collagen secretion arising from inhibition of CP4H.

3.4 Conclusions

We have identified compounds that can replace EDHB in experiments with cultured cells. Unlike EDHB, we found that diethyl pythiDC, diethyl pyimDC, and even diethyl bipy55′DC inhibit CP4H activity in cultured cells at concentrations that do not cause iron deficiency. Given the subtle ferritin phenotype caused by diethyl pyimDC and the measurable iron affinity of bipy55′DC and related bipyDCs *in vitro*,[148] we put forth diethyl pythiDC as a preferred replacement for EDHB. Moreover, an esterified pythiDC could serve as a lead for clinical development.

3.5 Experimental Procedures

3.5.1 General

2,4-Pyridinedicarboxylic acid (24PDC), 2,5-pyridinedicarboxylic acid (25PDC), 2,2′-bipyridine (bipy), 2-(1*H*-imidazol-2-yl)pyridine (pyim), and 2,2′-bipyridine-5,5′-dicarboxylic acid (bipy55′DC) were from Sigma–Aldrich (St. Louis, MO). 2-(1*H*-Pyrazol-3-yl)pyridine was from Combi-Blocks (San Diego, CA). Phosphine ligands and phosphonium salts were from either Sigma–Aldrich or Strem (Newberryport, MA), stored in a dessicator, and used without further purification. $Pd(OAc)_2$ was from Sigma–Aldrich, stored in a dessicator, and used without further purification. Deferroxamine mesylate was from Santa Cruz Biotechnology (Dallas, TX). Ethyl dihydroxybenzoate (EDHB) was from Combi-Blocks and was recrystallized from EtOAc before use in cell culture experiments. All other reagent chemicals were obtained from commercial

sources (Sigma–Aldrich, Acros, Combi-Blocks, Oakwood Products, Enamine, Bachem, or Novabiochem) and used without further purification. HIF-1α peptide$_{556\text{-}575}$ was from AnaSpec (Fremont, CA) and used without further purification.

All glassware was flame- or oven-dried, and reactions were performed under N_2(g) unless indicated otherwise. DCM and toluene were dried over a column of alumina. Dimethylformamide was dried over alumina and further purified through an isocyanate scrubbing column. Other anhydrous solvents were obtained in septum-sealed bottles. Flash chromatography was performed with columns of 40–63 Å silica gel, 230–400 mesh from Silicycle (Québec City, Canada). Thin-layer chromatography (TLC) was performed on plates of EMD 250 μm silica 60-F_{254} with visualization by UV light or staining with $KMnO_4$.

The phrase "concentrated under reduced pressure" refers to the removal of solvents and other volatile materials using a rotary evaporator at water aspirator pressure (<20 torr) while maintaining water-bath temperature below 40 °C. Residual solvent was removed from samples at high vacuum (<0.1 torr). The term "high vacuum" refers to vacuum achieved by a mechanical belt-drive oil pump. All reported yields are unoptimized. All procedures were performed at ambient temperature unless indicated otherwise.

3.5.2 Instrumentation

NMR spectra were acquired at ambient temperature with a DMX-400 Avance spectrometer or an Avance 500i spectrometer from Bruker (Billerica, MA) at the National Magnetic Resonance

Facility at Madison (NMRFAM) and were referenced to TMS or a residual protic solvent. Some compounds exist as either mixtures of rotomers or tautomers that do not interconvert on the NMR timescale at ambient temperature and therefore exhibit multiple sets of NMR signals (as indicated). Electrospray ionization (ESI) and electron ionization (EI) mass spectrometry were performed with a Micromass LCT® or Micromass AutoSpec® instruments, respectively, from Waters (Milford, MA) at the Mass Spectrometry Facility in the Department of Chemistry at the University of Wisconsin–Madison. The progress of reactions catalyzed by prolyl 4-hydroxylases was determined by analytical HPLC (Waters system equipped with a Waters 996 photodiode array detector, Empower 2 software). Preparative HPLC was performed using a Prominence HPLC instrument from Shimadzu (Kyoto, Japan) equipped with two LC-20AP pumps, an SPD-M20A photodiode array detector, and a CTO-20A column oven. Iron complexes with biheteroaryl ligands were analyzed by spectrophotometry using a Cary 60 UV–Vis spectrophotometer from Agilent Technologies (Santa Clara, CA). Protein concentrations were calculated from their absorbance at 280 nm as measured with a NanoVue Plus spectrophotometer from GE Healthcare using an extinction coefficient[140] of 290,000 $M^{-1}cm^{-1}$ for human CP4H[140] and 36,9000 $M^{-1}cm^{-1}$ for human PHD2.[141] Values of IC_{50}, EC_{50}, and LD_{50} were calculated from experimental data with Prism version 6.0 software from GraphPad Software (La Jolla, CA).

3.5.3 Production of Recombinant Human CP4H1

Human CP4H containing the α(I) isoform was produced heterologously in Origami B(DE3) *Escherichia coli* cells and purified as described previously.[140]

3.5.4 Assay of Human CP4H1 Activity in the Presence of Inhibitors

The catalytic activity of human CP4H1 was assayed as described previously.[140] Briefly, activity assays were carried out at 30 °C in 100 μL Tris–HCl buffer, pH 7.8, containing human CP4H1 (100 nM), inhibitor (0–500 μM), substrate (dansyl-Gly-Pro-Pro-Gly-OEt, 500 μM), $FeSO_4$ (50 μM), BSA (1 mg/mL), catalase (0.1 mg/mL), sodium ascorbate (2 mM), DTT (100 μM), and α-ketoglutarate (100 μM). Reactions were pre-incubated with or without inhibitor for 2 min at 30 °C, after which the reaction was initiated by the addition of α-ketoglutarate. After 15 min, reactions were quenched by boiling for 45 s and centrifuged at 10,000*g*. The supernatant (20–50 μL) was injected into a Nucleodur® C18 Gravity reversed-phase column (4.6 × 250 mm, 5-μm particle size) from Macherey–Nagel (Bethlehem, PA). The column was eluted at 1 mL/min with a linear gradient (20 min) of aqueous acetonitrile (20–45% v/v) containing TFA (0.1% v/v). The absorbance of the eluent was monitored at 289 nm. All assays were performed in triplicate. Data are reported as catalytic activity relative to control reactions lacking inhibitor, where catalytic activity is determined from the percent conversion of substrate to product. Dose–response curves were generated for each inhibitor by plotting the relative activity versus the log

of the inhibitor concentration. IC_{50}-values for each inhibitor were interpolated from the data by non-linear regression to the sigmoidal dose–response equation in Prism software.

3.5.5 Production of Recombinant Human PHD2

A cDNA encoding human $PHD2_{181-426}$ possessing an N-terminal hexahistidine (His_6) tag ($NHis_6$-$PHD2_{181-426}$) was cloned using the Gibson strategy,[142] and the encoded protein was produced and purified as described previously.[141]

3.5.6 Assay of Human PHD2 Activity in the Presence of Inhibitors

The catalytic activity of human PHD2 was assayed as described previously.[141] Briefly, activity assays were carried out at 30 °C in 100 μL Tris–HCl buffer, pH 7.8, containing human $NHis_6$-$PHD2_{181–426}$ (5 μM), inhibitor (0–50 μM), substrate (HIF-1α $peptide_{556–574}$, 50 μM), $FeSO_4$ (50 μM), BSA (1 mg/mL), catalase (0.3 mg/mL), sodium ascorbate (2 mM), DTT (1 mM), and α-ketoglutarate (35 μM). Reaction mixtures were pre-incubated with or without inhibitor for 2 min at 30 °C, after which the reaction was initiated by the addition of α-ketoglutarate. After 10 minutes, reactions were quenched by boiling for 60 s and centrifuged at 10,000*g*. The supernatant (50 μL) was injected into a Nucleodur® C18 Gravity reversed-phase column (4.6 × 250 mm, 5 μm particle size) from Macherey–Nagel. The column was eluted at 1 mL/min with a gradient (34 min) of aqueous acetonitrile (5–56% v/v) containing TFA (0.1% v/v). The absorbance of the eluent was monitored at 218 nm. All assays were performed in triplicate. Data

is reported as activity relative to control reactions lacking inhibitor, where activity is determined from the percent conversion of substrate to product.

3.5.7 Estimation of the pK_a Values of Biheteroaryl Ligands

The pK_a value of biheteroaryl compounds was estimated with potentiometric titration. Compounds (5–10 mM) were dissolved in water and the pH was adjusted to ~12 by adding 10 M NaOH. The pH of the solution was recorded with an Accumet™ XL500 pH meter equipped with an Orion 912600® pH electrode from Thermo Fisher Scientific while titrating the solution with HCl (1 M or 100 μM, as necessary) until a pH value of ~1.75 was achieved. Titration curves were prepared by plotting the solution pH versus the volume of 1 M HCl added, and pK_a values were estimated from derivative plots of the titration curves using Prism software.

3.5.8 Assay of Fe(II)-Affinity for Biheteroaryl Ligands

The affinity of biheteroaryl ligands for Fe(II) was determined comparatively by measuring the half-maximal concentration (EC_{50}) required for binding 20 μM Fe(II) (Fe_{20}-EC_{50}) in sodium phosphate buffer at pH 7.0. Stock solutions of ligands were prepared in either water for high affinity ligands (typically, Fe_{20}-EC_{50} <1 mM) or DMSO for low affinity ligands (typically, Fe_{20}-EC_{50} >1 mM). Stock solutions of $FeSO_4$ were prepared in H_2O and used within 3 h of preparation. Ligand solutions (3 μM–18 mM, depending on affinity) were prepared in 10 mM sodium phosphate buffer, pH 7.0, after which Fe(II) stock solution was added to initiate complex

formation. For high-affinity ligands, solutions were allowed to equilibrate for 15 min, after which the absorbance was recorded at the λ_{max} for the complex under study. For most low affinity ligands, the corresponding Fe(II) complexes were unstable and observed to dissociate over time. Therefore, the absorbance value was determined within 30 s of mixing, where the absorbance was measured at the λ_{max} for the complex under study. Complexes with pyox were not observed under these conditions. Absorbance values were corrected by subtracting the absorbance value in the absence of ligand. Dose–response curves were generated for each ligand by plotting the absorbance *versus* the log of the ligand concentration. Values of Fe_{20}-EC_{50} for each ligand were interpolated from the dose–response curves by non-linear regression using the sigmoidal dose–response function in Prism software. All experiments were performed in triplicate. Experiments to study pyimDC were performed as described above, except that phosphate buffer was excluded from the assay solution, bipy was added last to a final concentration of 300 μM, and the absorbance of the $Fe(bipy)_3^{2+}$ complex was measured at 523 nm.

3.5.9 Determination of Fe(II) Complex Stoichiometry

The stoichiometry of biheteroaryl complexes with Fe(II) was estimated with Job's method. Briefly, reactions were prepared such that the total moles of Fe(II) and ligand were kept constant, but the mole fraction of the ligand was varied from 0 to 1. The total concentration of ligand and Fe(II) used for each individual ligand was based upon the iron affinity and extinction coefficient,

and ranged from 0.4 mM to 2 mM. Stock solutions of ligands were prepared in water. Stock solutions of $FeSO_4$ were prepared in water and used within 3 h of preparation. Reaction mixtures were prepared in 10 mM sodium phosphate buffer, pH 7.0, and complex formation was initiated by the addition of Fe(II) solution. For high-affinity ligands (typically, Fe_{20}-EC_{50} <1 mM), solutions were allowed to equilibrate for 15 min, after which the absorbance was recorded at the λ_{max} for the complex under study. For most low-affinity ligands (typically, Fe_{20}-EC_{50} >1 mM), the corresponding Fe(II) complexes were unstable and observed to dissociate over time. Therefore, the absorbance value was determined within 30 s of mixing, where the absorbance was measured at the λ_{max} for the complex under study. Complexes with pyox were not observed under these conditions. Absorbance values were corrected by subtracting the absorbance value of a blank solution in absence of $FeSO_4$, after which the values were normalized relative to the reaction with the highest absorbance value. All experiments were performed in at least duplicate. Job's plots were constructed by plotting the normalized absorbance versus the mole fraction of biheteroaryl ligand, after which the stoichiometry of the complex was estimated from the mole fraction of the reaction with the highest absorbance value. If necessary, blank titrations using only Fe(II) were used to correct the Job's plots.

3.5.10 General Mammalian Cell Culture

The HEK293T line of human embryonic kidney cells was from American Type Culture Collection (ATCC, Manassas, VA), and the MDA-MB-231 line of human breast cancer cells was

a generous gift from Dr. Beth A. Weaver (University of Wisconsin–Madison). Cell lines were maintained according to the procedures recommended by the ATCC. Cells were grown in a cell culture incubator at 37 °C under CO_2 (5% v/v) in flat-bottomed culture flasks. The culture medium was DMEM supplemented with GIBCO fetal bovine serum (FBS) (10% v/v), penicillin (100 units/mL), streptomycin (100 μg/mL) and L-glutamine (2 mM). Cells were counted by hemocytometry with Trypan Blue prior to use in assays.

3.5.11 Cytotoxicity Assays

The toxicity of esterified biheteroaryl compounds for human breast cancer cells was evaluated with the CellTiter 96® AQueous One Solution Cell Proliferation Assay (MTS) from Promega (Madison, WI). Briefly, MDA-MB-231 cells grown as described in Section X were plated at a concentration of 5,000 cells/well in a clear 96-well plate. The cells were allowed to adhere for 4 h, after which the medium was removed and discarded. Fresh medium was added and the cells were treated with varying concentrations of the test compound at 37 °C for 24 h. The medium was removed, and cells were washed with Dulbecco's PBS. The MTS reagent was added at a ratio of 1:5, and the cells were incubated at 37 °C for 2 h before measuring the absorbance at 490 nm. The average absorbance was measured in triplicate for each concentration tested, and the entire experiment was repeated in duplicate. The percentage of viable cells was determined by normalizing to a PBS control (100% viable), and a H_2O_2 control (0% viable). For all of the compounds tested, the LD_{50}-value could not be measured due to its being higher than the

aqueous solubility limit: diethyl Bipy55′DC, LD_{50} >100 μM; diethyl pyimDC, LD_{50} >1000 μM; diethyl pythiDC, LD_{50} >500 μM.

3.5.12 Effect of CP4H Inhibitors on Iron Metabolism and P4HA1 Levels in Human Cells

MDA-MB-231 or HEK293T cells were plated at 50,000 cells/well into 12-well plates and grown to confluence (~24 h) as described in Section X. The cells were washed, and the medium was replaced with 1.0 mL of serum-free medium. Stock solutions of all test compounds were prepared at 100× in DMSO and added to a final concentration of 1× (0.1–0.5 mM). After addition of the test compound or DMSO vehicle, cells were incubated for 24 h. The cells were then washed with 100 μL MPER solution from Thermo Fisher Scientific (Waltham, MA) and collected. Protein concentrations were determined by BCA assay (Thermo Fisher Scientific), and samples corresponding to 40 μg total protein were analyzed with an immunoblot.

3.5.13 Immunoblotting

For all immunoblot analyses, protein samples where boiled in SDS–PAGE buffer, separated by electrophoresis through a Tris–glycine SDS–PAGE gel from Bio-Rad Laboratories (Hercules, CA), and subsequently transferred to a PVDF membrane for immunoblotting. All primary antibodies were used at the working dilution specified by the manufacturer. Primary antibodies were visualized using a complementary secondary antibody fused to horseradish peroxidase (HRP), with detection of HRP by chemiluminescence, imaging using an ImageQuant LAS 4000

(GE Healthcare), and quantification by densitometry (ImageJ software). The anti-rabbit antibody from Promega and the anti-mouse antibody from Abbiotec (San Diego, CA) were used at the working dilution specified by the manufacturer. Statistical comparisons were performed using Student's *t*-test or one-way ANOVA functions available in Prism.

For investigations of iron metabolism, protein samples (40 μg) were separated by SDS–PAGE through a 12% w/v acrylamide gel, and blots were probed with primary antibodies to human ferritin (rabbit monoclonal) from Abcam (Cambridge, MA), human transferrin receptor (mouse monoclonal) from Invitrogen, human HIF-1α (mouse monoclonal) from BD Biosciences, and human β-actin (rabbit monoclonal) from Cell Signaling Technology (Danvers, MA). Quantifications were normalized to the signal from β-actin, after which treated samples were compared to vehicle-treated controls.

For investigations of P4HA1 levels, protein samples were treated and quantified as described above for iron metabolism except that the blots were probed with primary antibodies to human P4HA1 (rabbit polyclonal) from Thermo Fisher Scientific and human β-actin.

For investigations of collagen secretion, protein samples were separated by SDS–PAGE through a 7.5% w/v acrylamide gel. Blots were first stained for total protein using Ponceau S, after which they were probed with a primary antibody to human type-I collagen (α1) (rabbit polyclonal) from Novus Biologicals (Littleton, CO). Quantifications were normalized to the total protein signal obtained from Ponceau S staining, after which treated samples were compared to vehicle-treated controls.

3.5.14 Effect of CP4H Inhibitors on the Binding of the Iron-Responsive Element by IRPs

MDA-MB-231 cells were plated at 3×10^6 cells/dish into 100 mm × 20 mm dishes and grown to confluence (~24 h) as described in Section X. The cells were washed, and the medium was replaced with 10 mL of serum-free medium. Stock solutions of all test compounds were prepared at 100× in DMSO and added to a final concentration of 1× (as denoted). After addition of the test compound or DMSO vehicle, cells were incubated for 24 h, after which they were harvested by trypsinization and centrifugation. The resultant pellets (~50 μL) were washed with PBS and then lysed using a protocol described previously.[159] Briefly, cell pellets were resuspended in 200 μL of cell lysis buffer, which was 20 mM Hepes–HCl buffer, pH 7.4, containing sodium pyrophosphate (10 mM), sodium fluoride (50 mM), β-glycerophosphate (50 mM), EDTA (5 mM), GTP (1 mM), sodium orthovanadate (1 mM), benzamidine hydrochloride (2 mM), Nonidet NP-40 (0.5% v/v), *p*-nitrophenyl *p′*-guanidinobenzoate·HCl (25 μg/mL), DTT (1 mM), leupeptin (40 μg/mL), pepstatin (4 μg/mL), SBTI (100 μg/mL), MG132 (10 μM), PMSF (200 μM), and BHT (5 μg/mL) by vortexing and lysed on ice by vortexing every few min. After 15 min, the lysates were centrifuged at 14,000 rpm and the supernatants aliquoted and stored at −80 °C until analysis by an electrophoretic mobility shift assay (EMSA).

3.5.15 Electrophoretic Mobility Shift Assay for IRE-Binding by IRPs

EMSAs were performed and quantified essentially as described previously.[160] Briefly [^{32}P]-L-ferritin IRE (1 nM RNA, specific radioactivity ~7,000 dpm/fmol) was incubated with lysate

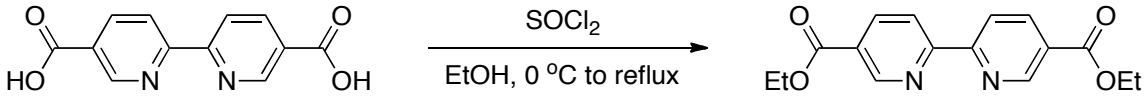
protein (2.5 μg as determined by the Bradford assay) in binding buffer for 10 min on ice. Heparin was added (to 0.5 mg/mL), and the reaction mixture was incubated for an additional 5 min on ice before separating bound and free RNA on a nondenaturing polyacrylamide gel at ambient temperature.

3.5.16 Effect of CP4H Inhibitors on Collagen Production in Mammalian Cells

MDA-MB-231 cells were plated at 50,000 cells/well into 12-well plates and grown to confluence (~24 h) as described in Section X. The cells were washed, and the medium was replaced with 1.5 mL of serum-free medium containing sodium ascorbate (50 μg/mL). Stock solutions of all test compounds were prepared at 100× in DMSO and added to a final concentration of 1× (as denoted). After the addition of a test compound or vehicle (DMSO), cells were incubated for 48 h, after which 1.0 mL of conditioned medium was removed and added to 4.0 mL of chilled acetone (–20 °C). The resultant mixtures were incubated for 3 h at –20 °C, after which the precipitated protein was pelleted by centrifugation at 10,000*g* for 30 min. The supernatants were discarded and the pellets dried briefly in a fume hood. Pellets were stored at –20 °C prior to analysis with an immunoblot.

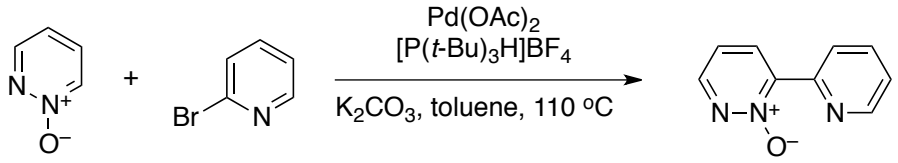
3.5.17 Synthetic Procedures

Diethyl 2,2′-Bipyridine-5,5′-dicarboxylate (diethyl bipy55′DC)



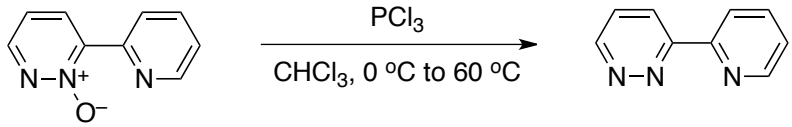
Bipy55′DC (200 mg, 0.82 mmoles) and EtOH (13 mL) were added to a dried flask, and the resulting solution was stirred on ice. Thionyl chloride (1.3 mL) was added dropwise on ice, after which the flask was fitted with a reflux condenser and heated at reflux. After 24 h, the reaction mixture was cooled on ice and quenched by the dropwise addition of saturated aqueous Na_2CO_3 (20 mL). The aqueous layer was extracted with CH_2Cl_2 (4 × 20 mL), and the combined organic extracts were dried over Na_2SO_4(s) and concentrated under reduced pressure. The crude product was purified by chromatography on silica (3% v/v acetone in 1:1 DCM/hexanes) to afford the title compound (190 mg, 77%) as a white solid. **^{1}H NMR** (500 MHz, $CDCl_3$, δ): 9.32 (dd, J = 0.5, 2.0 Hz, 1 H), 8.59 (dd, J = 0.5, 8.5 Hz, 1 H), 8.46 (dd, J = 2.0, 8.5 Hz, 1 H), 4.47 (q, J = 7.5 Hz, 2 H), 1.46 (t, J = 7.5 Hz, 3 H); **^{13}C NMR** (125 MHz, $CDCl_3$, δ): 165.2, 158.3, 150.6, 138.1, 126.6, 121.3, 61.6, 14.3; **HRMS** (ESI) m/z 301.1193 [calc'd for $C_{16}H_{17}N_2O_4$ $(M + H)^+$ 301.1183].

6-(Pyridin-2-yl)pyridazine-1-oxide



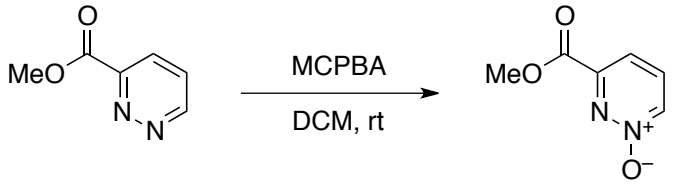
$Pd(OAc)_2$ (29.0 mg, 0.13 mmoles), $[P(t\text{-}Bu)_3H]BF_4$ (113 mg, 0.39 mmoles), pyridazine-1-oxide (500 mg, 5.2 mmoles), and K_2CO_3 (719 mg, 5.2 mmoles) were added to a dried flask. The flask was fitted with a reflux condenser capped with a septum, evacuated, and purged with N_2(g) (~5 times). A degassed solution of 2-bromopyridine (411 mg, 2.6 mmoles) in dry toluene (13 mL) was added via syringe, and the reaction mixture was stirred at 110 °C for 18 h. The reaction mixture was then cooled and filtered through Celite®, and the filtrate was concentrated under reduced pressure. The crude product was purified by chromatography on silica (35% v/v acetone in hexanes) to afford the title compound (206 mg, 46%) as a grey solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 8.90 (d, J = 8.0 Hz, 1H), 8.70 (d, J = 4.0 Hz, 1 H), 8.63 (dd, J = 1.6, 8.0 Hz, 1 H), 8.47 (s, 1 H), 7.84 (t, J = 7.6 Hz, 1 H), 7.36 (dd, J = 5.2, 7.2 Hz, 1 H), 7.19 (dd, J = 5.2, 8.0 Hz, 1 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 149.8, 149.5, 148.2, 142.5, 136.7, 135.6, 125.0, 124.7, 116.5; **HRMS** (ESI) m/z 174.0658 [calc'd for $C_9H_8N_3O$ $(M + H)^+$ 174.0662].

2-(Pyridin-2-yl)pyridazine (pypyrid)



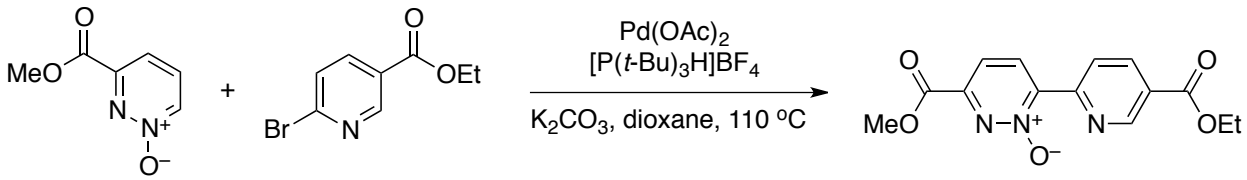
6-(Pyridin-2-yl)pyridazine-1-oxide (150 mg, 0.87 mmoles) was dissolved in dry $CHCl_3$ (4.3 mL) and cooled to 0 °C in an ice bath, after which PCl_3 (230 μL, 2.6 mmoles) was added dropwise with stirring. The reaction mixture was stirred at 60 °C until the starting material was consumed completely, as judged by TLC. The reaction mixture was quenched by the dropwise addition of saturated aqueous Na_2CO_3 (10 mL) while stirring on ice. The product was extracted with $CHCl_3$ (4 × 10 mL), and the combined organic extracts were dried over Na_2SO_4(s) and concentrated under reduced pressure. The crude product was purified by chromatography on silica (40% v/v EtOAc in hexanes) to afford the title compound (50 mg, 37%) as a brown solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 9.19 (dd, J = 1.6, 5.2 Hz, 1H), 8.70–8.67 (m, 2 H), 8.55 (dd, J = 1.6, 8.8 Hz, 1 H), 7.87 (td, J = 1.6, 8.0 Hz, 1 H), 7.59 (dd, J = 4.8, 8.8 Hz, 1 H), 7.38 (ddd, J = 1.2, 4.8, 7.6 Hz, 1 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 158.6, 153.5, 151.2, 149.4, 137.2, 127.0, 124.7, 124.4, 121.6; **HRMS** (ESI) m/z 158.0714 [calc'd for $C_9H_8N_3$ $(M + H)^+$ 158.0713].

Methyl Pyridazine-3-carboxylate-1-oxide



Methyl pyridazine-3-carboxylate (2.0 g, 14.5 mmoles) and *m*-chloroperoxybenzoic acid (3.9 g, 17.4 mmoles) were dissolved in dry DCM (29 mL) in a flame-dried flask. The reaction mixture was stirred at ambient temperature for 5 h. The solvent was evaporated under reduced pressure, and the resulting residue was purified by chromatography on silica (30% v/v acetone in hexanes) to afford the title compound as an off-white solid (1.97 g, 88%). **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 8.28 (dd, J = 2.4, 5.2 Hz, 1 H), 7.79–7.75 (m, 2 H), 4.03 (s, 3 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 162.2, 150.2, 136.9, 134.4, 116.8, 53.6; **HRMS** (ESI) m/z 155.0452 [calc'd for $C_6H_7N_2O_3$ $(M + H)^+$ 155.0452].

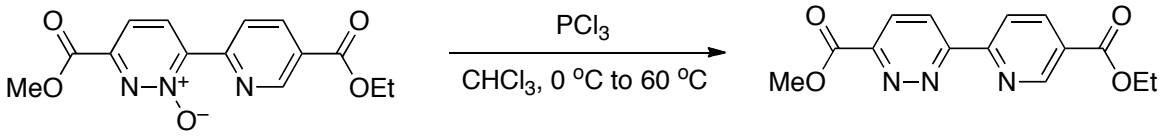
Methyl 6-(5-Methoxycarbonylpyridin-2-yl)pyridazine-3-carboxylate-1-oxide



$Pd(OAc)_2$ (7.3 mg, 0.033 mmoles), $[P(t\text{-}Bu)_3H]BF_4$ (28 mg, 0.098 mmoles), K_2CO_3 (180 mg, 1.3 mmoles), ethyl 6-bromonicotinate (150 mg, 0.65 mmoles), and methyl pyridazine-3-carboxylate-1-oxide (200 mg, 1.3 mmoles) were added to a dried flask. The flask was fitted with a reflux condenser capped with a septum, evacuated, and purged with N_2(g) (~5 times). Dry

dioxane (4 mL) was added via syringe, and the reaction mixture was stirred at 110 °C for 18 h. The reaction mixture was then cooled and filtered through Celite®, and the filtrate was concentrated under reduced pressure. The crude product was purified by chromatography on silica (40% v/v EtOAc in hexanes) to afford the title compound (82 mg) as a pale orange solid. Due to the presence of minor contaminants that were difficult to remove by chromatography or recrystallization, the slightly crude product was used directly in the next reaction before further purification and characterization.**^{1}H NMR** (500 MHz, $CDCl_3$, δ): 9.26 (d, J = 2.0 Hz, 1 H), 9.06 (d, J = 8.5, Hz, 1 H), 8.85 (d, J = 8.5 Hz, 1 H), 8.42 (dd, J = 2.0, 8.5 Hz, 1 H), 7.85 (dd, J = 8.5 Hz, 1 H), 4.41 (q, J = 7.0 Hz, 2 H), 4.00 (s, 3 H), 1.40 (t, J = 7.0 Hz, 3 H); **HRMS** (ESI) m/z 304.0931 [calc'd for $C_{14}H_{14}N_3O_5$ $(M + H)^+$ 304.0928].

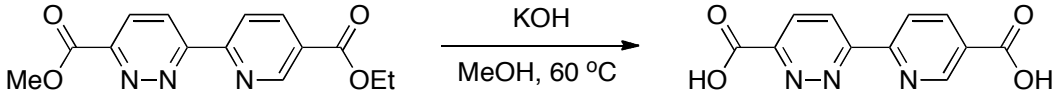
Methyl 6-(5-Ethoxycarbonylpyridin-2-yl)pyridazine-3-carboxylate (ethylmethyl pyrpyridDC)



Methyl 6-(5-methoxycarbonylpyridin-2-yl)pyridazine-3-carboxylate-1-oxide (80 mg, 0.28 mmoles) was dissolved in dry $CHCl_3$ (1.3 mL) and cooled to 0 °C in an ice bath, after which PCl_3 (370 μL, 3.7 mmoles) was added dropwise with stirring. The reaction mixture was stirred at 60 °C until the starting material was consumed completely, as judged by TLC. The reaction was

quenched by the dropwise addition of saturated aqueous Na_2CO_3 (10 mL) while stirring on ice. The product was extracted with $CHCl_3$ (4 × 10 mL), and the combined organic extracts were dried over Na_2SO_4(s) and concentrated under reduced pressure. The crude product was purified by chromatography on silica (35% v/v EtOAc in hexanes) to afford the title compound (37 mg, 20% over two steps) as a pale yellow solid. **^{1}H NMR** (500 MHz, $CDCl_3$, δ): 9.12 (d, J = 1.5 Hz, 1 H), 8.66 (d, J = 8.0 Hz, 1 H), 8.57 (d, J = 8.5 Hz, 1 H), 8.30 (dd, J = 2.0, 8.0 Hz, 1 H), 8.14 (d, J = 8.5 Hz, 1 H), 4.26 (q, J = 7.0 Hz, 2H), 3.92 (s, 3H), 1.25 (t, J = 7.0 Hz, 3H); **^{13}C NMR** (125 MHz, $CDCl_3$, δ): 164.7, 164.3, 158.9, 155.5, 151.1, 150.6, 138.2, 128.2, 127.2, 125.3, 121.6, 61.5, 53.3, 14.1; **HRMS** (ESI) m/z 288.0982 [calc'd for $C_{14}H_{14}N_3O_4$ $(M + H)^+$ 288.0979].

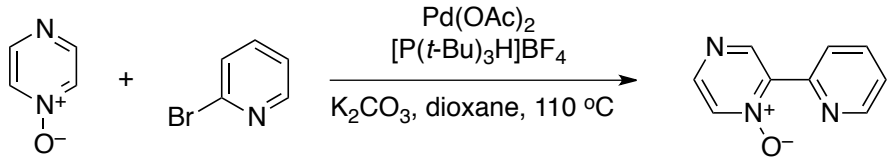
6-(5-Carboxypyridin-2-yl)pyridazine-3-carboxylic Acid (pypyridDC)



Methyl 6-(5-ethoxycarbonylpyridin-2-yl)pyridazine-3-carboxylate (37 mg, 0.13 mmoles) and KOH (34 mg, 0.52 mmoles) were added to a vial. MeOH (1.3 mL) was added to the vial, and the reaction mixture was heated to 60 °C with stirring until the starting material was consumed completely, as judged by TLC. The reaction mixture was cooled and concentrated under reduced pressure. The crude product was dissolved in water (2 mL). The aqueous layer was washed with EtOAc (1 × 2 mL), after which the product was precipitated from the aqueous layer by adjusting to pH 3–4 with 1 M HCl. After cooling to 4 °C, the product was filtered, washed with water (3 ×

1 mL), and dried under high vacuum to afford the title compound (23 mg, 73%) as a pale yellow solid. **^{1}H NMR** (500 MHz, DMSO-d_6, δ): 13.73 (bs, 2H), 9.11 (dd, J = 1.0, 2.0 Hz, 1 H), 8.63 (dd, J = 1.0, 8.5 Hz, 1 H), 8.60 (d, J = 9.0 Hz, 1 H), 8.40 (dd, J = 2.0, 8.5 Hz, 1 H), 8.24 (d, J = 9.0 Hz, 1 H); **^{13}C NMR** (125 MHz, DMSO-d_6, δ): 166.2, 165.3, 158.8, 155.7, 152.5, 150.9, 139.1, 129.2, 128.2, 126.0, 121.9; **HRMS** (ESI) m/z 246.0517 [calc'd for $C_{11}H_8N_3O_4$ $(M + H)^+$ 246.0510].

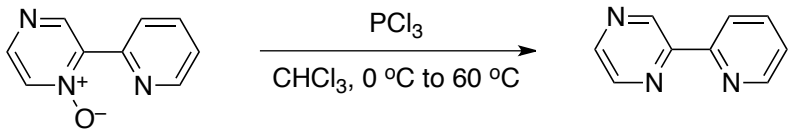
2-(Pyridin-2-yl)pyrazine-1-oxide



$Pd(OAc)_2$ (39.0 mg, 0.17 mmoles), $[P(t\text{-}Bu)_3H]BF_4$ (151 mg, 0.52 mmoles), pyrazine-1-oxide (1.0 g, 10.4 mmoles), and K_2CO_3 (959 mg, 6.9 mmoles) were added to a dried flask. The flask was fitted with a reflux condenser capped with a septum, evacuated, and purged with N_2(g) (~5 times). A degassed solution of 2-bromopyridine (548 mg, 3.5 mmoles) in dry dioxane (17 mL) was added via syringe, and the reaction mixture was stirred at 110 °C for 18 h. The reaction mixture was then cooled and filtered through Celite®, and the filtrate was concentrated under reduced pressure. The crude product was purified by chromatography on silica (30% v/v acetone in hexanes) to afford the title compound (312 mg, 52%) as a pale yellow solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 9.35 (s, 1H), 8.75–8.72 (m, 2 H), 8.38 (d, J = 4.0 Hz, 1 H), 8.15 (dd, J =

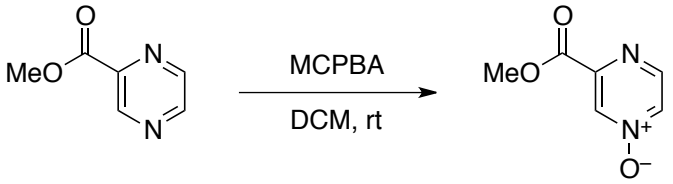
0.4, 4.0 Hz, 1 H), 7.81 (ddd, J = 1.6, 2.0, 8.0 Hz, 1 H), 7.25 (ddd, J = 0.8, 4.8, 8.0 Hz, 1 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 149.9, 149.7, 147.3, 145.8, 142.2, 136.5, 134.5, 125.4, 124.8; **HRMS** (ESI) m/z 174.0662 [calc'd for $C_9H_8N_3O$ $(M + H)^+$ 174.0662].

2-(Pyridin-2-yl)pyrazine (pypyraz)



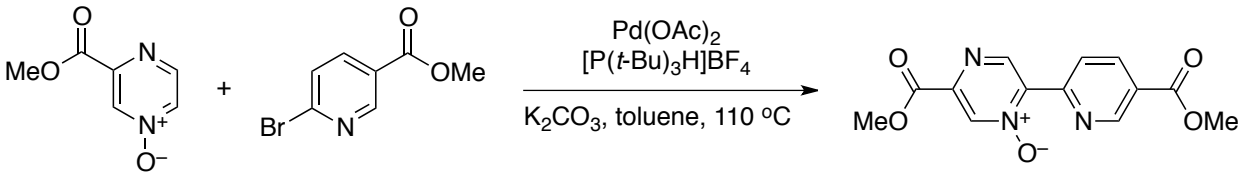
2-(Pyridin-2-yl)pyrazine-1-oxide (200 mg, 1.2 mmoles) was dissolved in dry $CHCl_3$ (23 mL) and cooled to 0 °C in an ice bath, after which PCl_3 (360 μL, 4.2 mmoles) was added dropwise with stirring. The reaction mixture was stirred at 60 °C until the starting material was consumed completely, as judged by TLC. The reaction mixture was quenched by the dropwise addition of saturated aqueous Na_2CO_3 (20 mL) while stirring on ice. The product was extracted with $CHCl_3$ (4 × 20 mL), and the combined organic extracts were dried over Na_2SO_4(s) and concentrated under reduced pressure. The crude product was purified by chromatography on silica (30% v/v EtOAc in hexanes) to afford the title compound (103 mg, 57%) as a pale yellow solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 9.63 (d, J = 1.2 Hz, 1H), 8.71–8.69 (m, 1 H), 8.60–8.58 (m, 2 H), 8.34 (d, J = 8.0 Hz, 1 H), 7.82 (ddd, J = 1.6, 2.0, 8.0 Hz, 1 H), 7.34 (ddd, J = 1.2, 4.8, 7.6 Hz, 1 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 154.2, 151.1, 149.4, 144.4, 143.5, 143.3, 137.0, 124.4, 121.4; **HRMS** (ESI) m/z 158.0708 [calc'd for $C_9H_8N_3$ $(M + H)^+$ 158.0713].

Methyl Pyrazine-3-carboxylate-1-oxide



Methyl pyrazine-2-carboxylate (3.0 g, 21.7 mmoles) and *m*-chloroperoxybenzoic acid (7.3 g, 32.6 mmoles) were dissolved in dry DCM (43 mL) in a flame-dried flask. The reaction mixture was stirred at ambient temperature for 48 h. The solvent was evaporated under reduced pressure, and the resulting residue was purified by chromatography on silica (40–50% v/v acetone in hexanes followed by 80–100% v/v EtOAc in hexanes) to afford the title compound as a white solid (751 mg, 22%). **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 8.76 (dd, J = 0.8, 1.6 Hz, 1 H), 8.57 (dd, J = 0.4, 4.0 Hz, 1 H), 8.22 (dd, J = 2.0, 4.0 Hz, 1 H), 4.05 (s, 3 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 162.6, 147.6, 147.3, 136.0, 135.9, 53.7; **HRMS** (ESI) m/z 155.0451 [calc'd for $C_6H_7N_2O_3$ $(M + H)^+$ 155.0452].

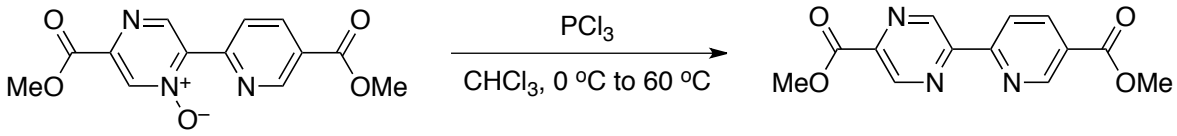
Methyl 2-(5-Methoxycarbonylpyridin-2-yl)pyrazine-5-carboxylate-1-oxide



$Pd(OAc)_2$ (21 mg, 0.09 mmoles), $[P(t\text{-}Bu)_3H]BF_4$ (81 mg, 0.28 mmoles), K_2CO_3 (511 mg, 3.7 mmoles), methyl 6-bromonicotinate (400 mg, 1.9 mmoles), and methyl pyrazine-3-

carboxylate-1-oxide (571 mg, 3.7 mmoles) were added to a dried flask. The flask was fitted with a reflux condenser capped with a septum, evacuated, and purged with N_2(g) (~5 times). Dry toluene (9 mL) was added via syringe, and the reaction mixture was stirred at 110 °C for 24 h. The reaction mixture was then cooled and filtered through Celite®, and the filtrate was concentrated under reduced pressure. The crude product was purified by chromatography on silica (45% v/v EtOAc in hexanes) to afford the title compound (100 mg) as a pale orange solid. Due to the presence of minor contaminants that were difficult to remove by chromatography or recrystallization, the slightly crude product was used directly in the next reaction before further purification and characterization. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 9.63 (s, 1 H), 9.38 (dd, J = 0.4, 2.0, Hz, 1 H), 9.06 (dd, J = 0.8, 8.4 Hz, 1 H), 8.89 (d, J = 0.4 Hz, 1 H), 8.48 (dd, J = 2.0, 8.4 Hz, 1 H), 4.10 (s, 3 H), 4.03 (s, 3 H); **HRMS** (ESI) m/z 290.0770 [calc'd for $C_{13}H_{12}N_3O_5$ (M + H)$^+$ 290.0772].

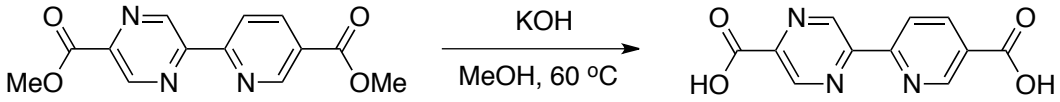
Methyl 2-(5-Methoxycarbonylpyridin-2-yl)pyrazine-5-carboxylate (dimethyl pypyrazDC)



Methyl 2-(5-methoxycarbonylpyridin-2-yl)pyrazine-5-carboxylate-1-oxide (100 mg, 0.35 mmoles) was dissolved in dry $CHCl_3$ (1.5 mL) and cooled to 0 °C in an ice bath, after which PCl_3 (100 μL, 1.0 mmoles) was added dropwise with stirring. The reaction mixture was stirred at 60 °C until the starting material was consumed completely, as judged by TLC. The

reaction was quenched by the dropwise addition of saturated aqueous Na_2CO_3 (10 mL) while stirring on ice. The product was extracted with $CHCl_3$ (3 × 10 mL), and the combined organic extracts were dried over Na_2SO_4(s) and concentrated under reduced pressure. The crude product was recrystallized from $CHCl_3$ and then EtOAc to afford the title compound (28 mg, 6% over two steps) as a pale yellow solid. **^{1}H NMR** (500 MHz, $CDCl_3$, δ): 9.84 (d, J = 1.5 Hz, 1 H), 9.38 (d, J = 1.0 Hz, 1 H), 9.36 (dd, J = 0.5, 2.0 Hz, 1 H), 8.58 (d, J = 0.5, 8.0 Hz, 1 H), 8.50 (dd, J = 2.0, 8.5 Hz, 1 H), 4.11 (s, 3H), 4.03 (s, 3H); **^{13}C NMR** (125 MHz, $CDCl_3$, δ): 165.4, 164.4, 156.4, 152.4, 150.8, 145.2, 143.3, 143.0, 138.4, 126.9, 121.9, 53.3, 52.7; **HRMS** (ESI) *m*/*z* 274.0824 [calc'd for $C_{13}H_{12}N_3O_4$ $(M + H)^+$ 274.0823].

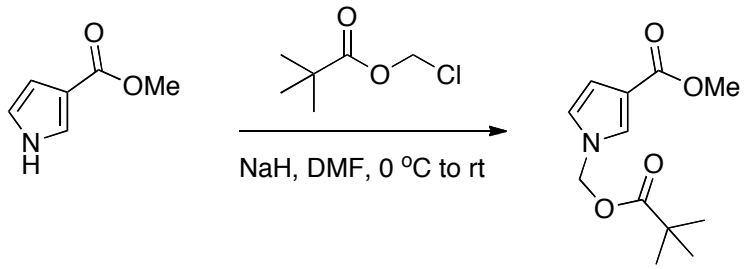
2-(5-Carboxypyridin-2-yl)pyrazine-5-carboxylic Acid (pypyrazDC)



Methyl 2-(5-methoxycarbonylpyridin-2-yl)pyrazine-5-carboxylate (28 mg, 0.10 mmoles) and KOH (26 mg, 0.41 mmoles) were added to a vial. MeOH (3 mL) was added to the vial, and the reaction mixture was heated to 60 °C with stirring until the starting material was consumed completely, as judged by TLC. The reaction mixture was cooled and concentrated under reduced pressure. The crude product was dissolved in water (2 mL). The aqueous layer was washed with EtOAc (1 × 2 mL), after which the product was precipitated from the aqueous layer by adjusting to pH 3–4 with 1 M HCl. After cooling to 4 °C, the product was filtered, washed with water (3 ×

1 mL), and dried under high vacuum to afford the title compound (17 mg, 68%) as a light brown solid. **^{1}H NMR** (500 MHz, DMSO-d_6, δ): 13.84 (bs, 2H), 9.69 (d, J = 1.5 Hz, 1 H), 9.30 (d, J = 1.0 Hz, 1 H), 9.25 (dd, J = 0.5, 2.0 Hz, 1 H), 8.54 (d, J = 7.5 Hz, 1 H), 8.50 (dd, J = 2.0, 8.0 Hz, 1 H); **^{13}C NMR** (125 MHz, DMSO-d_6, δ): 168.8, 165.9, 156.8, 152.1, 151.4, 146.0, 145.3, 143.3, 139.7, 128.7, 122.7; **HRMS** (ESI) m/z 246.0505 [calc'd for $C_{11}H_8N_3O_4$ $(M + H)^+$ 246.0510].

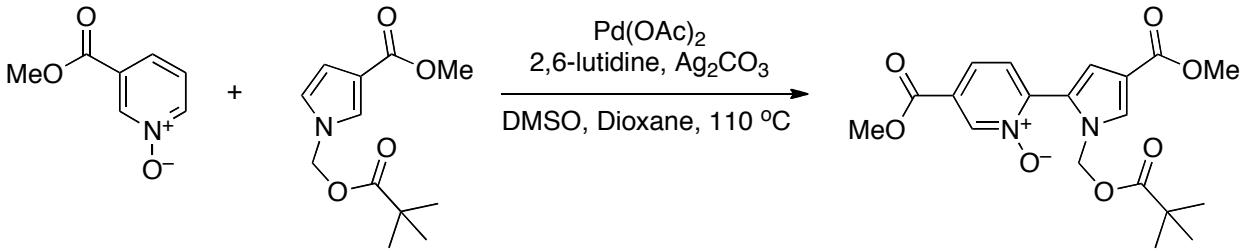
Methyl 1-Pivaloyloxymethyl-1*H*-pyrrole-3-carboxylate



NaH (60% w/v in mineral oil, 267 mg, 6.7 mmoles) was added to a dry flask. The flask was evacuated and purged with N_2(g) (~5 times). Dry DMF (3 mL) was added, and the flask was cooled on ice. A degassed solution of methyl 1*H*-pyrrole-3-carboxylate (836 mg, 6.7 mmoles) in dry DMF (4 mL) was added dropwise over 5 min with stirring. The reaction mixture was stirred on ice for 45 min until gas evolution was complete, after which chloromethyl pivalate (0.97 mL, 6.7 mmoles) was added dropwise. The reaction mixture was allowed to come to ambient temperature overnight, after which the reaction mixture was concentrated under reduced pressure to a crude oil. The oil was taken up in H_2O (25 mL), and the aqueous layer extracted with DCM (5 × 25 mL). The combined organic extracts were dried over Na_2SO_4(s) and concentrated under reduced pressure, after which the crude product was purified by chromatography on silica

(25% v/v EtOAc in hexanes) to afford the title compound (1.20 g, 75%) as a golden oil. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 7.46 (dd, J = 0.4, 2.0 Hz, 1 H), 6.79 (dd, J = 0.4, 2.4 Hz, 1 H), 6.60 (dd, J = 1.2, 1.6 Hz, 1 H), 5.79 (s, 2 H), 3.81, (s, 3 H), 1.18 (s, 9 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 177.7, 164.9, 126.9, 122.6, 117.4, 111.0, 70.8, 51.2, 38.8, 26.8; **HRMS** (ESI) m/z 240.1233 [calc'd for $C_{12}H_{18}NO_4$ $(M + H)^+$ 240.1231].

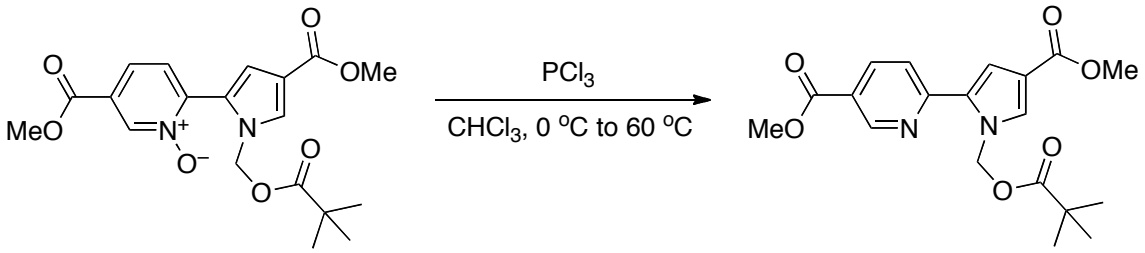
Methyl 2-(4-Methoxycarbonyl-1-pivaloyloxymethyl-1*H*-pyrrol-2-yl)pyridine-5-carboxylate-1-oxide



$Pd(OAc)_2$ (9.2 mg, 0.041 mmoles), Ag_2CO_3 (450 mg, 1.6 mmoles), and 3-methoxycarbonylpyridine-1-oxide (125 mg, 0.82 mmoles) were added to a dried flask. The flask was fitted with a reflux condenser capped with a septum, evacuated, and purged with N_2(g) (~5 times). A degassed solution of methyl 1-pivaloyloxymethyl-1*H*-pyrrole-3-carboxylate (215 mg, 0.90 mmoles), 2,6-lutidine (29 μL, 0.25 mmoles), and DMSO (275 μL, 5% [v/v]) in dry dioxane (5.2 mL) was added via syringe, and the reaction mixture was stirred at 110 °C for 48 h. The reaction mixture was then cooled and filtered through Celite®, and the filtrate was concentrated under reduced pressure. The crude product was purified by chromatography on silica (60–100% v/v EtOAC in hexanes) to afford the title compound (30 mg, 9.4%) as a yellow

solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 8.88 (d, J = 1.2 Hz, 1 H), 7.87 (dd, J = 1.2, 8.0 Hz, 1 H), 7.67 (d, J = 2.0 Hz, 1 H), 7.50 (d, J = 8.0 Hz, 1 H), 6.87 (d, J = 1.6 Hz, 1 H), 6.03 (s, 2 H), 3.99 (s, 3 H), 3.83 (s, 3 H), 1.03 (s, 9 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 177.1, 164.1, 163.2, 145.5, 141.2, 130.6, 128.9, 128.4, 125.9, 125.4, 117.4, 115.8, 71.5, 53.1, 51.4, 38.6, 26.7; **HRMS** (ESI) m/z 391.15041 [calc'd for $C_{19}H_{23}N_2O_7$ $(M + H)^+$ 391.1500].

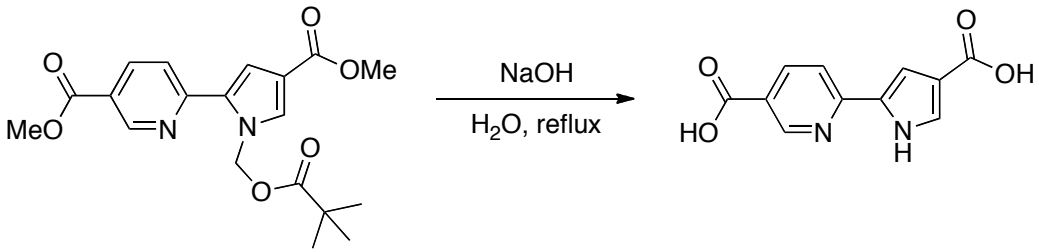
Methyl 2-(4-Methoxycarbonyl-1-pivaloyloxymethyl-1*H*-pyrrol-2-yl)pyridine-5-carboxylate



Methyl 2-(4-methoxycarbonyl-1-pivaloyloxymethyl-1*H*-pyrrol-2-yl)pyridine-5-carboxylate-1-oxide (30 mg, 0.077 mmoles) was dissolved in dry $CHCl_3$ (800 μL) and cooled to 0 °C in an ice bath, after which PCl_3 (8.1 μL, 0.092 mmoles) was added. The reaction mixture was stirred at 60 °C until starting material was consumed completely, as judged by TLC. The reaction was quenched by the dropwise addition of saturated aqueous Na_2CO_3 (3 mL) while stirring on ice. The product was extracted with DCM (4 × 3 mL), and the combined organic extracts were dried over Na_2SO_4(s) and concentrated under reduced pressure to afford the title compound (28 mg, 97%) as a brown solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 9.14 (d, J = 0.8, 2.0 Hz, 1 H), 8.28 (dd, J = 2.4, 8.4 Hz, 1 H), 7.66 (dd, J = 0.8, 8.4 Hz, 1 H), 7.63 (d, J = 2.0 Hz, 1 H), 7.19 (d, J = 1.6 Hz, 1 H), 6.49 (s, 2 H), 3.97 (s, 3 H), 3.86 (s, 3 H), 1.08 (s, 9 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ):

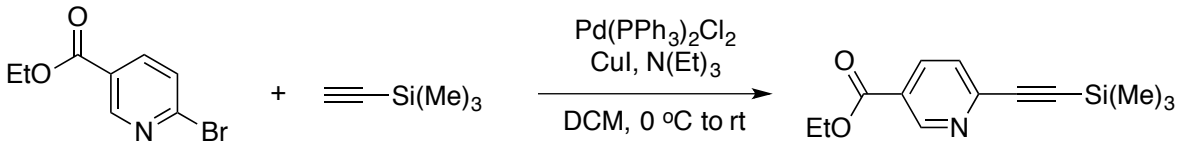
177.5, 165.7, 164.4, 154.2, 150.2, 137.5, 132.1, 131.6, 123.2, 120.3, 116.6, 114.3, 71.5, 52.3, 51.4, 38.7, 26.8; **HRMS** (ESI) *m*/*z* 375.1550 [calc'd for $C_{19}H_{23}N_2O_6$ $(M + H)^+$ 375.1551].

2-(4-Carboxy-1*H*-pyrrol-2-yl)pyridine-5-carboxylic Acid (pypyrroleDC)



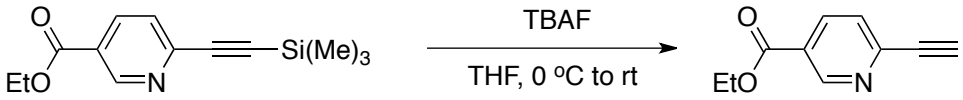
Methyl 2-(4-methoxycarbonyl-1-pivaloyloxymethyl-1*H*-pyrrol-2-yl)pyridine-5-carboxylate (28 mg, 0.075 mmoles) and 2 M NaOH (1.0 mL) were added to a vial. The reaction mixture was heated to reflux with stirring until the starting material was consumed completely, as judged by TLC. The reaction mixture was cooled and concentrated under reduced pressure. The crude product was dissolved in water (2 mL). The aqueous layer was washed with EtOAc (1 × 2 mL), after which the product was precipitated from the aqueous layer by adjusting to pH 3–4 with 1 M HCl. After cooling to 4 °C, the product was filtered, washed with water (3 × 1 mL), and dried under high vacuum to afford the title compound (20 mg, 78%) as a grey solid. **^{1}H NMR** (500 MHz, DMSO-d_6, δ): 13.23 (bs, 1 H), 12.30 (s, 1 H), 12.13, (bs, 1 H), 9.00 (dd, J = 1.0, 2.5 Hz, 1 H), 8.21 (dd, J = 2.5, 8.5 Hz, 1 H), 7.91 (d, J = 8.5 Hz, 1 H), 7.46 (dd, J = 2.5, 3.5 Hz, 1 H), 7.26 (dd, J = 1.5, 2.5 Hz, 1 H); **^{13}C NMR** (125 MHz, DMSO-d_6, δ): 169.3, 168.4, 156.1, 153.3, 140.8, 134.7, 129.9, 126.8, 121.1 (2 signals), 114.0; **HRMS** (EI) *m*/*z* 232.0474 [calc'd for $C_{11}H_8N_2O_4$ $(M)^+$ 232.0479].

Ethyl 2-(2-Trimethylsilylethynyl)pyridine-5-carboxylate



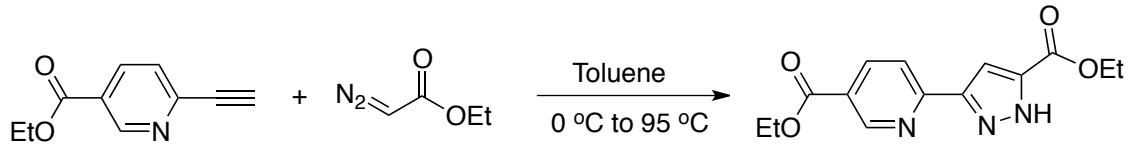
$Pd(PPh_3)_2Cl_2$ (21 mg, 0.030 mmoles), CuI (142 mg, 0.089 mmoles), and ethyl 2-bromonicotinate (532 mg, 2.3 mmoles) were added to a dried flask on ice. The flask was evacuated and purged with N_2(g) (~5 times). A degassed solution of trimethylsilylacetylene (420 μL, 3.0 mmoles) in DCM (5 mL) was added to the flask with stirring on ice. Triethylamine (1.4 mL, 10.0 mmoles) was added to the flask with stirring, after which the flask was allowed to come to ambient temperature overnight. The reaction mixture was diluted with hexanes (5 mL) and filtered through Celite®, and the filtrate was washed with H_2O (50 mL) and brine (50 mL). The organic layer was dried over Na_2SO_4(s) and concentrated under reduced pressure to afford a solid, which was purified by chromatography on silica (5% v/v EtOAc in hexanes) to afford the title compound (544 mg, 95%) as a pale yellow solid. **^{1}H NMR** (500 MHz, $CDCl_3$, δ): 9.17 (d, J = 1.6 Hz, 1 H), 8.26 (dd, J = 2.0, 8.0 Hz, 1 H), 7.54 (d, J = 7.6 Hz, 1 H), 4.43 (q, J = 7.2 Hz, 2 H), 1.43 (t, J = 7.2 Hz, 3 H), 0.30 (s, 9 H); **^{13}C NMR** (125 MHz, $CDCl_3$, δ): 167.5, 153.7, 149.2, 139.8, 129.4, 127.8, 105.8, 101.0, 64.3, 17.0, 2.3; **HRMS** (ESI) m/z 248.1102 [calc'd for $C_{13}H_{18}NO_2Si$ $(M + H)^+$ 248.1102].

Ethyl 2-Ethynylpyridine-5-carboxylate



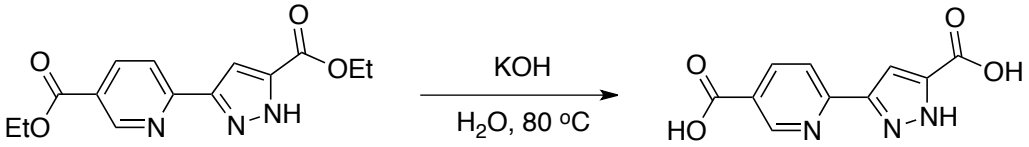
Ethyl 2-(2-trimethylsilylethynyl)pyridine-5-carboxylate (400 mg, 1.6 mmoles) was added to a flame-dried flask. The flask was evacuated and purged with N_2(g), after which THF (1.6 mL) was added on ice. 1.0 M TBAF in THF (1.9 mL) was added dropwise with stirring on ice, after which the flask was allowed to come to ambient temperature with stirring. After 20 min, the reaction mixture was concentrated under reduced pressure, and the crude residue was partitioned between EtOAc (10 mL) and 10% w/v $NaHCO_3$ (10 mL). The organic layer was collected, and the aqueous layer was extracted with EtOAc (10 mL). The combined organic extracts were dried over Na_2SO_4(s) and concentrated under reduced pressure to afford a solid, which was purified by chromatography on silica (20% v/v EtOAc in hexanes) to afford the title compound (144 mg, 51%) as a pale yellow solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 9.19 (d, J = 1.6 Hz, 1 H), 8.29 (dd, J = 2.0, 8.0 Hz, 1 H), 7.56 (d, J = 8.4 Hz, 1 H), 4.43 (q, J = 7.2 Hz, 2 H), 3.33 (s, 1 H), 1.43 (t, J = 7.2 Hz, 3 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 164.6, 151.0, 145.7, 137.2, 126.9, 125.6, 82.3, 79.8, 61.7, 14.2; **HRMS** (ESI) m/z 176.0708 [calc'd for $C_{10}H_{10}NO_2$ $(M + H)^+$ 176.0707].

Ethyl 2-(5-Ethoxycarbonyl-1*H*-pyrazol-3-yl)pyridine-5-carboxylate (diethyl pypyrDC)



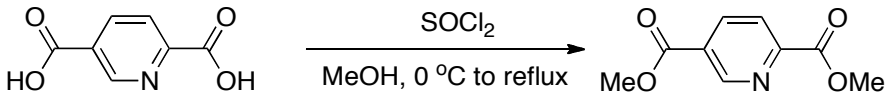
Ethyl 2-ethynylpyridine-5-carboxylate (91 mg, 0.52 mmoles) was dissolved in toluene (5 mL) in a glass vial, and the resulting solution was chilled on ice with stirring. A 73% w/w solution of ethyl diazoacetate in DCM (162 mg, 1.04 mmoles) was added dropwise, after which the vial was purged with N_2(g) and stirred at 95 °C. After 24 h, the reaction mixture was chilled on ice, and a second portion of ethyl diazoacetate (162 mg, 1.04 mmoles) was added. The reaction mixture was stirred at 95 °C for an addition 24 h, after which the reaction was concentrated under reduced pressure to afford a solid. The crude product was purified by chromatography on silica (30% EtOAc in hexanes) to afford the title compound (125 mg, 83%) as a pale yellow solid. **^{1}H NMR** (500 MHz, $CDCl_3$, δ): 12.06 (bs, 1 H), 9.29 (dd, J = 0.5, 2.0 Hz, 1 H), 8.43 (dd, J = 2.0, 8.5 Hz, 1 H), 7.91 (d, J = 8.5 Hz, 1 H), 7.46 (s, 1 H), 4.483 (q, J = 7.0 Hz, 2 H), 4.480 (q, J = 7.0 Hz, 2 H), 1.47 (t, J = 7.0 Hz, 6 H); **^{13}C NMR** (125 MHz, $CDCl_3$, δ): 167.6, 164.0, 154.4, 153.6, 148.4 (br), 144.8 (br), 141.1, 128.3, 122.5, 110.2, 64.3, 64.1, 17.0 (2 signals); **HRMS** (ESI) *m/z* 312.0952 [calc'd for $C_{14}H_{15}N_3O_4Na$ $(M + Na)^+$ 312.0955].

2-(5-Carboxy-1*H*-pyrazol-3-yl)pyridine-5-carboxylic Acid (pypyrDC)



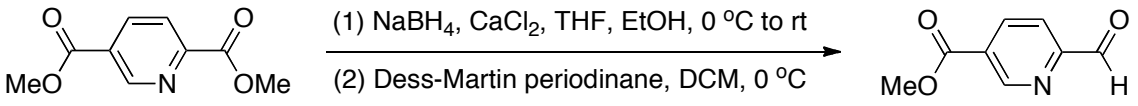
Ethyl 2-(5-ethoxycarbonyl-1*H*-pyrazol-3-yl)pyridine-5-carboxylate (38 mg, 0.13 mmoles) and KOH (34 mg, 0.525 mmoles) were added to a vial. H_2O (1.75 mL) was added to the vial and the reaction mixture was heated to 80 °C and stirred for 2 hr. The reaction mixture was cooled and washed with EtOAc (1 × 3 mL), after which the product was precipitated from the aqueous layer by adjusting to pH 3–4 with 1M HCl. After cooling to 4 °C, the product was filtered, washed with water (3 × 2 mL), and dried in vacuo to afford the title compound (24 mg, 78%) as a white solid. **^{1}H NMR** (500 MHz, DMSO-d_6, δ): 14.16–13.60 (m, 2 H), 9.10 (dd, J = 1.0, 2.5 Hz, 1 H), 8.33 (dd, J = 2.0, 8.0 Hz, 1 H), 8.10 (d, J = 8.5 Hz, 1 H), 7.37 (s, 1 H); **^{13}C NMR** (125 MHz, DMSO-d_6, δ): 169.2, 164.6 (br), 161.4 (q, J = 37 Hz, TFA), 160.0 (br), 153.4, 150.4 (br), 143.0 (br), 141.3, 128.6, 122.7, 118.4 (q, J = 287 Hz, TFA), 110.5; **HRMS** (ESI) m/z 235.0509 [calc'd for $C_{10}H_8N_3O_4$ $(M+H)^+$ 234.0509]. (Note: A small drop of TFA was added to the sample prior to recording the ^{13}C NMR spectrum in order to sharpen the peaks corresponding to the pyrazole C3, C4, C5, and carbonyl carbons, as is common with many free NH pyrazoles.[161])

Dimethyl 2,5-Pyridinedicarboxylate



Thionyl chloride (2.2 mL) was added dropwise to MeOH (14 mL) while stirring on ice. 2,5-Pyridinedicarboxylic acid (1.0 g, 6.0 mmoles) was added, and the reaction mixture was heated at reflux for 3 h. The reaction mixture was cooled, and the solvent removed under reduced pressure. The resulting residue was dissolved in DCM (15 mL), after which saturated aqueous Na_2CO_3 (15 mL) was added while stirring on ice. The aqueous layer was extracted with DCM (3 × 15 mL), and the combined organic extracts were washed with saturated aqueous Na_2CO_3 (2 × 40 mL), dried over Na_2SO_4(s), and concentrated under reduced pressure to afford the title compound (915 mg, 78%) as a pale yellow solid. **^{1}H NMR** (500 MHz, $CDCl_3$, δ): 9.32 (dd, J = 0.5, 2.0 Hz, 1 H), 8.47 (dd, J = 2.0, 8.0 Hz, 1 H), 8.23 (dd, J = 0.5, 8.0 Hz, 1 H), 4.06 (s, 3 H), 4.01 (s, 3 H); **^{13}C NMR** (125 MHz, $CDCl_3$, δ); 165.0, 164.9, 150.8, 150.8, 138.4, 128.6, 124.7, 53.3, 52.8; **HRMS** (ESI) m/z 196.0600 [calc'd for $C_9H_{10}NO_4$ $(M + H)^+$ 196.0605].

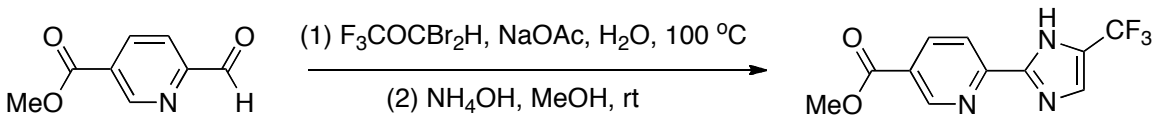
Methyl 6-Formylnicotinate



Dimethyl-2,5-pyridinedicarboxylate (960 mg, 4.9 mmoles) and $CaCl_2$ (2.198 g, 19.7 mmoles) were added to a dried flask. THF (11 mL) and EtOH (12 mL) were then added, and the resulting suspension was stirred on ice for 30 min. $NaBH_4$ (465 mg, 12.3 mmoles) was added portion-wise

with stirring. The reaction mixture was allowed to come to ambient temperature overnight. After 18 h, the reaction was quenched by the dropwise addition of an aqueous solution of NH_4Cl (20 mL of saturated aqueous NH_4Cl plus 40 mL of H_2O) while stirring on ice. The aqueous layer was extracted with DCM (4 × 30 mL), and the combined organic extracts were dried over Na_2SO_4(s) and concentrated under reduced pressure to afford crude product (678 mg) as a pale yellow solid. The crude product was dissolved in dry DCM (45 mL), after which Dess–Martin periodinane (2.6 g, 6.1 mmoles) was added portion-wise with stirring on ice. After 6 h, the reaction was quenched by the dropwise addition of a solution of 5% w/v $Na_2S_2O_3$ in half-saturated $NaHCO_3$ (80 mL). The aqueous layer was extracted with DCM (3 × 40 mL), and the combined organic extracts were dried over Na_2SO_4(s) and concentrated under reduced pressure to afford the title compound (504 mg, 62%) as a pale yellow solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 10.16 (s, 1 H), 9.38 (dd, J = 0.5, 2.0 Hz, 1 H), 8.49 (dd, J = 2.0, 8.0 Hz, 1 H), 8.05 (dd, J = 0.5, 8.0 Hz, 1 H), 4.02 (s, 1 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 192.6, 164.8, 154.9, 151.2, 138.3, 121.1, 52.9; **HRMS** (EI) m/z 165.0415 [calc'd for $C_8H_7NO_3$ (M^+) 165.0421]

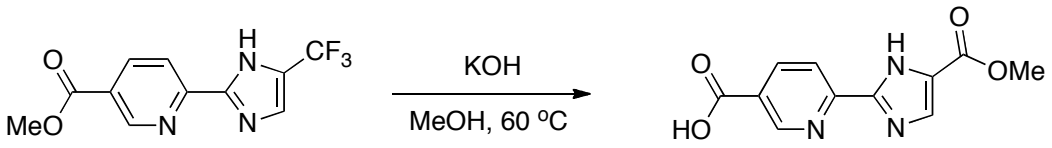
Methyl 2-(5-Trifluoromethyl-1*H*-imidazol-2-yl)pyridine-5-carboxylate



NaOAc (277 mg, 3.4 mmoles) and 2,2-dibromo-1,1,1-trifluoroacetone (454 mg, 1.7 mmoles) were dissolved in H_2O (770 μL) and stirred at 100 °C for 30 min. The reaction mixture was cooled to ambient temperature, after which a solution of methyl 6-formylnicotinate (250 mg, 1.5

mmoles) and concentrated NH_4OH (1.5 mL) in MeOH (4.6 mL), was added dropwise while stirring. The reaction mixture was stirred overnight, after which the solvent was removed under reduced pressure. $CHCl_3$ (25 mL) and 10% v/v $NaHCO_3$ (25 mL) were added to the residue. The aqueous layer was extracted with $CHCl_3$ (2 × 25 mL), and the combined organic extracts were dried over Na_2SO_4(s) and concentrated under reduced pressure. The crude product was purified by chromatography on silica (25% v/v EtOAc in hexanes) and recrystallized from 1:1 hexanes/EtOAc to afford the title compound (128 mg, 47%) as a pale yellow crystalline solid. **^{1}H NMR** (500 MHz, CD_3OD, δ): 9.13 (dd, J = 1.0, 2.0 Hz, 1 H), 8.38 (dd, J = 2.0, 8.5 Hz, 1 H), 8.15 (dd, J = 1.0, 8.5 Hz, 1 H), 7.72 (d, J = 1.0 Hz, 1 H), 3.96 (s, 3 H); **^{13}C NMR** (125 MHz, CD_3OD, δ): 165.2, 150.5, 150.2, 146.7, 137.9, 132.6 (q, J = 39 Hz), 125.7, 121.7 (q, J = 265 Hz), 119.3, 119.1, 51.6; **HRMS** (ESI) m/z 272.0630 [calc'd for $C_{11}H_9F_3N_3O_2$ $(M + H)^+$ 272.0642].

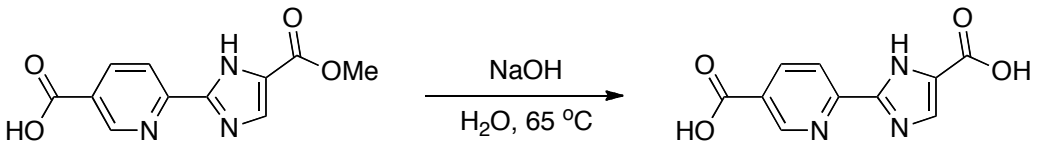
2-(5-Methoxycarbonyl-1-*H*-imidazol-2-yl)pyridine-5-carboxylic Acid



Methyl 2-(5-trifluoromethyl-1*H*-imidazol-2-yl)pyridine-5-carboxylate (103 mg, 0.38 mmoles) and KOH (147 mg, 2.28 mmoles) were added to a vial. MeOH (11 mL) was added to the vial, and the reaction mixture was heated to 60 °C until the starting material was consumed completely, as judged by TLC. The reaction mixture was cooled, and concentrated under reduced pressure. The crude product was dissolved in water (3 mL). The aqueous layer was washed with

EtOAc (1 × 3 mL), after which the product was precipitated from the aqueous layer by adjusting to pH 3–4 with 1 M HCl. After cooling to 4 °C, the product was filtered, washed with water (3 × 1 mL), and dried under high vacuum to afford the title compound (84 mg, 94%) as a white solid. **^{1}H NMR** (500 MHz, DMSO-d_6, δ); 13.74 (s, 1 H), 13.56 (s, 1 H), 9.12 (d, J = 2.0 Hz, 1 H), 8.40 (dd, J = 2.0, 8.0 Hz, 1 H), 8.22 (d, J = 8.0 Hz, 1 H), 7.99 (d, J = 2.5 Hz, 1 H), 3.81 (s, 3 H); **^{13}C NMR** (125 MHz, DMSO-d_6, δ): 166.4, 163.1, 151.4, 150.5, 146.4, 138.9, 134.2, 126.6 (2 signals), 120.3, 51.7; **HRMS** (ESI) m/z 248.0671 [calc'd for $C_{11}H_{10}N_3O_4$ $(M + H)^+$ 248.0666].

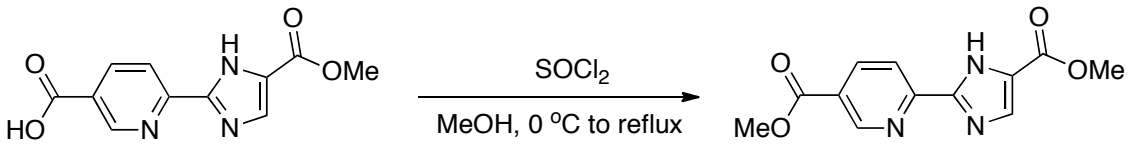
2-(5-Carboxy-1-*H*-imidazol-2-yl)pyridine-5-carboxylic Acid (pyimDC)



2-(5-Methoxycarbonyl-1*H*-imidazol-2-yl)pyridine-5-carboxylic acid (50 mg, 0.20 mmoles), 1 M NaOH (400 μL), and water (3.6 mL) were added to a vial. The reaction mixture was heated to 65 °C with stirring. After 24 h, the reaction mixture was cooled and washed with EtOAc (1 × 3 mL), after which the product was precipitated from the aqueous layer by adjusting to pH 3–4 with 1 M HCl. After cooling to 4 °C, the product was filtered, washed with water (3 × 1 mL), and dried under high vacuum to afford the title compound (44 mg, 93%) as a pale yellow solid. **^{1}H NMR** (500 MHz, DMSO-d_6, δ): 13.62 (s, 1 H), 13.30 (bs, 1 H), 12.74 (bs, 1 H), 9.11 (dd, J = 0.5, 1.0 Hz, 1 H), 8.39 (dd, J = 2.0, 8.5 Hz, 1 H), 8.20 (d, J = 8.0 Hz, 1 H), 7.88 (d, J = 2.5 Hz, 1

H); **^{13}C NMR** (125 MHz, DMSO-d_6, δ): 166.4, 164.0, 151.5, 150.5, 146.1, 138.9, 135.3, 126.5, 126.2, 120.2; **HRMS** (EI) *m*/*z* 233.0432 [calc'd for $C_{10}H_7N_3O_4$ $(M)^+$ 233.18].

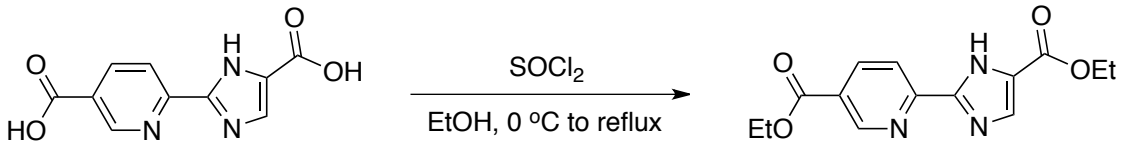
Methyl 2-(5-Methoxycarbonyl-1*H*-imidazol-2-yl)pyridine-5-carboxylate (dimethyl pyimDC)



2-(5-Methoxycarbonyl-1*H*-imidazol-2-yl)pyridine-5-carboxylic acid (100 mg, 0.41 mmoles) and MeOH (7.0 mL) was added to a round-bottom flask and stirred on ice. Thionyl chloride (750 μL) was added dropwise on ice, after which the flask was fitted with a reflux condenser and heated at reflux. After 24 h, the reaction mixture was cooled on ice and quenched by the dropwise addition of saturated aqueous Na_2CO_3 (20 mL). The aqueous layer was extracted with CH_2Cl_2 (5 × 20 mL) and the combined organic extracts were dried over Na_2SO_4(s), and concentrated under reduced pressure. The crude product was purified by chromatography on silica (50% v/v EtOAc in hexanes), to afford a pale yellow solid. The solid was dissolved in minimal DCM, precipitated by the dropwise addition of cold hexanes, and filtered to afford the title compound (50 mg, 48%) as a white solid. **^{1}H NMR** (500 MHz, $CDCl_3$, ~5:2 ratio of 2 tautomers, δ); 11.01 (bs, 1.4 H), 9.18–9.15 (m, 1.4 H), 8.44–8.38 (m, 2.4 H), 8.29 (d, J = 8 Hz, 0.4 H), 7.91 (s, 1 H), 7.84 (s, 0.4 H), 3.99–3.94 (m, 8.4 H); **^{13}C NMR** (125 MHz, $CDCl_3$, ~5:2 ratio of 2 tautomers, δ): 165.3 (2

signals), 163.2, 160.2, 150.7, 150.4 (2 signals), 150.3, 147.8, 146.2, 138.4 (2 signals), 136.7, 135.4, 126.1, 125.9, 124.3 (2 signals), 120.2, 120.0, 52.6 (2 signals), 52.1 (2 signals); **HRMS** (ESI) *m*/*z* 262.0829 [calc'd for $C_{12}H_{12}N_3O_4$ $(M + H)^+$ 262.0823].

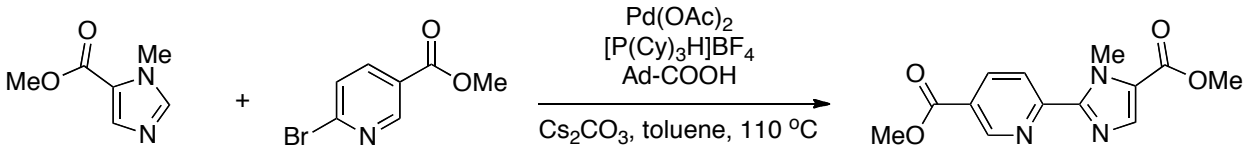
Ethyl 2-(5-Ethoxycarbonyl-1*H*-imidazol-2-yl)pyridine-5-carboxylate (diethyl pyimDC)



2-(5-Carboxy-1*H*-imidazol-2-yl)pyridine-5-carboxylic acid (85 mg, 0.36 mmoles) and EtOH (6.3 mL) were added to a round-bottom flask and stirred on ice. Thionyl chloride (710 μL) was added dropwise on ice, after which the flask was fitted with a reflux condenser and heated at reflux. After 24 h, the reaction mixture was cooled on ice and quenched by the dropwise addition of saturated aqueous Na_2CO_3 (10 mL). The aqueous layer was extracted with CH_2Cl_2 (6 × 10 mL) and the combined organic extracts were dried over Na_2SO_4(s) and concentrated under reduced pressure. The crude product was purified by chromatography on silica (2% v/v MeOH in 1:1 DCM/hexanes) to afford the title compound (95 mg, 90 %) as a pale yellow solid. **^{1}H NMR** (500 MHz, $CDCl_3$, ~5:4 ratio of two tautomers, *δ*): 11.16 (bs, 1 H), 11.02 (bs, 0.8 H), 9.19 (dd, *J* = 1.0, 2.0 Hz, 0.8 H); 9.15–9.14 (m, 1 H), 8.44–8.38 (m, 2.8 H), 8.28 (dd, *J* = 1.0, 8.0 Hz, 0.8 H), 7.90–7.86 (m, 1.8 H), 4.48–4.38 (m, 7.2 H), 1.46–1.39 (m, 10.8 H); **^{13}C NMR** (125 MHz, $CDCl_3$, *δ*): 164.9, 164.8 162.8, 159.8, 150.7, 150.4 (2 signals), 150.3, 147.8, 146.3, 138.4, 138.3,

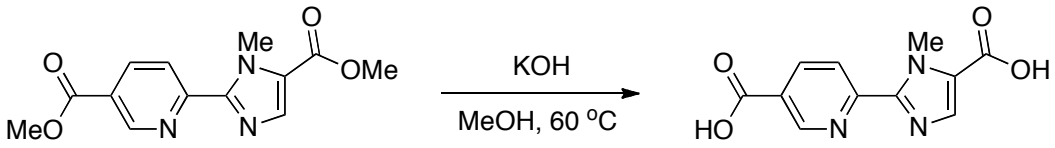
136.5, 135.7, 126.3, 126.1, 124.7, 124.1, 120.1, 120.0, 61.7 (2 signals), 61.3, 61.0, 14.5, 14.4, 14.3 (2 signals); **HRMS** (ESI) *m*/*z* 290.1136 [calc'd for $C_{14}H_{16}N_3O_4$ $(M + H)^+$ 290.1132].

Methyl 2-(5-Methoxycarbonyl-1-methyl-1*H*-imidazol-2-yl)pyridine-5-carboxylate (dimethyl NMe-pyimDC)



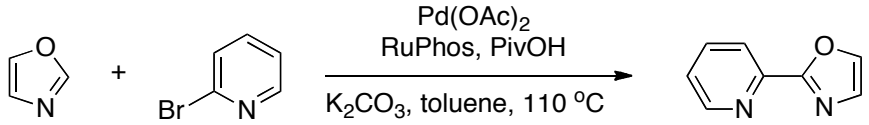
$Pd(OAc)_2$ (32 mg, 0.14 mmoles), $[P(Cy)_3H]BF_4$ (158 mg, 0.43 mmoles), 1-adamantanecarboxylic acid (258 mg, 1.4 mmoles), Cs_2CO_3 (1.9 g, 5.7 mmoles), methyl 6-bromonicotinate (618 mg, 2.9 mmoles) and methyl 1-methyl-1*H*-imidazole-5-carboxylate (400 mg, 2.9 mmoles) were added to a dried flask. The flask was fitted with a reflux condenser capped with a septum, evacuated, and purged with N_2(g) (~5 times). Dry toluene (4.8 mL) was added via syringe, and the reaction mixture was stirred at 110 °C for 48 h. The reaction mixture was then cooled and filtered through Celite®, and the filtrate was concentrated under reduced pressure. The crude product was purified by chromatography on silica (30% v/v EtOAc in hexanes) to afford the title compound (40 mg, 5%) as a pale yellow solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 9.23 (d, J = 1.2 Hz, 1 H), 8.38 (dd, J = 2.0, 8.4 Hz, 1 H), 8.28 (d, J = 8.4 Hz, 1 H), 7.84 (s, 1 H), 4.43 (s, 3 H), 3.98 (s, 3 H), 3.89 (s, 3 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 165.4, 160.8, 153.2, 149.8, 148.2, 137.6, 137.3, 125.6, 125.1, 123.8, 52.5, 51.6, 35.0; **HRMS** (ESI) *m*/*z* 298.0796 [calc'd for $C_{13}H_{13}N_3O_4Na$ $(M + Na)^+$ 298.0799].

2-(5-Carboxy-1-methyl-1*H*-imidazol-2-yl)pyridine-5-carboxylic Acid (NMe-pyimDC)



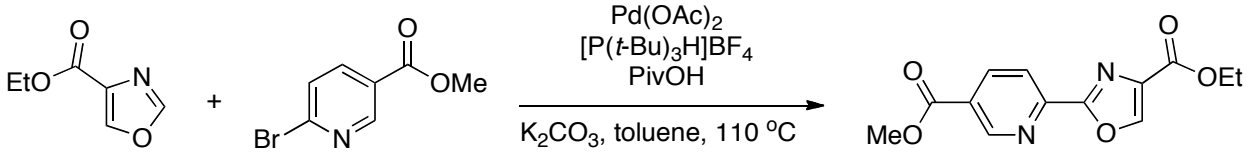
Methyl 2-(5-methoxycarbonyl-1-methyl-1*H*-imidazol-2-yl)pyridine-5-carboxylate (30 mg, 0.11 mmoles) and KOH (28 mg, 0.44 mmoles) were added to a vial. MeOH (3 mL) was added to the vial and the reaction mixture was heated to 60 °C with stirring until the starting material was consumed completely, as judged by TLC. The reaction mixture was cooled and concentrated under reduced pressure, after which the crude product was dissolved in water (2 mL). The aqueous layer was washed with EtOAc (1 × 2 mL), after which the product was precipitated from the aqueous layer by adjusting to pH 3–4 with 1 M HCl. After cooling to 4 °C, the product was filtered, washed with water (3 × 1 mL), and dried under high vacuum to afford the title compound (16 mg, 58%) as a pale yellow solid. **^{1}H NMR** (500 MHz, DMSO-d_6, δ): 13.49 (bs, 1 H), 13.2 (bs, 1 H), 9.16 (d, J = 2.0 Hz, 1 H), 8.41 (dd, J = 2.0, 8.5 Hz, 1 H), 8.25 (d, J = 8.5 Hz, 1 H), 7.79 (s, 1 H), 4.33 (s, 3 H); **^{13}C NMR** (125 MHz, DMSO-d_6, δ): 166.4, 161.7, 153.1, 149.9, 148.0, 138.5, 137.2, 126.7, 126.3, 124.3, 35.1; **HRMS** (EI) m/z 247.0583 [calc'd for $C_{11}H_9N_3O_4$ $(M)^+$ 247.0588].

2-(Oxazol-2-yl)pyridine (pyox)



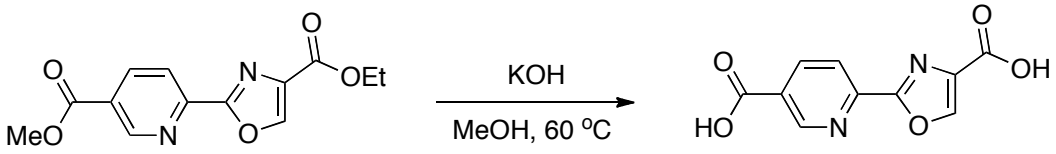
$Pd(OAc)_2$ (33 mg, 0.15 mmoles), RuPhos (135 mg, 0.29 mmoles), pivalic acid (119 mg, 1.2 mmoles) and K_2CO_3 (1.2 g, 8.7 mmoles) were added to a dried flask. The flask was fitted with a reflux condenser capped with a septum, evacuated, and purged with N_2(g) (~5 times). A degassed solution of oxazole (400 mg, 5.8 mmoles) and 2-bromopyridine (457 mg, 2.9 mmoles) in dry toluene (14.5 mL) was added via syringe, and the reaction mixture was stirred at 110 °C for 24 h. The reaction mixture was then cooled and filtered through Celite®, and the filtrate was concentrated under reduced pressure. The crude product was purified by chromatography on silica (20–75% v/v EtOAc in hexanes followed by 10% v/v acetone in 1:1 DCM/hexanes) to afford the title compound (62 mg, 7%) as a golden oil. **^{1}H NMR** (500 MHz, $CDCl_3$, δ): 8.67 (d, J = 5.0 Hz, 1 H), 8.08 (d, J = 8.0 Hz, 1 H), 7.78–7.74 (m, 2 H), 7.30 (ddd, J = 1.0, 5.0, 8.0 Hz, 1 H), 7.25 (s, 1 H); **^{13}C NMR** (125 MHz, $CDCl_3$, δ): 160.7, 149.9, 146.0, 139.8, 137.0, 128.8, 124.7, 122.0; **HRMS** (EI) m/z 146.0472 [calc'd for $C_8H_7N_2O$ $(M)^+$ 146.0475].

Methyl 2-(4-Ethoxycarbonyloxazol-2-yl)pyridine-5-carboxylate (methylethyl pyoxDC*)



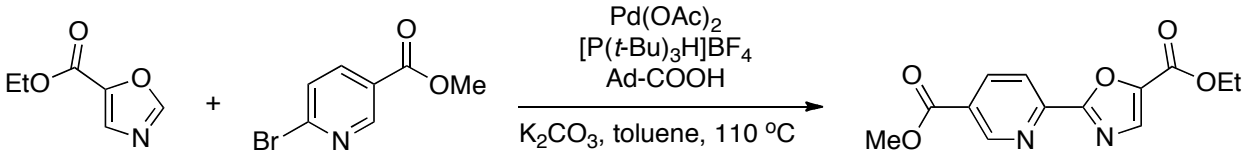
$Pd(OAc)_2$ (15.9 mg, 0.071 mmoles), $[P(t\text{-}Bu)_3H]BF_4$ (61.8 mg, 0.21 mmoles), pivalic acid (29.0 mg, 0.28 mmoles), K_2CO_3 (391 mg, 2.8 mmoles), ethyl oxazole-4-carboxylate (200 mg, 1.4 mmoles), and methyl 6-bromonicotinate (612 mg, 2.8 mmoles) were added to a dried flask. The flask was fitted with a reflux condenser capped with a septum, evacuated, and purged with N_2(g) (~5 times). Dry toluene (7 mL) was added via syringe, and the reaction mixture was stirred at 110 °C for 24 h. The reaction mixture was then cooled and filtered through Celite®, and the filtrate was concentrated under reduced pressure. The crude product was purified via chromatography on silica (2% v/v acetone in DCM followed by 10% v/v acetone in 1:1 DCM/hexanes) to afford the title compound (113 mg, 29%) as a white solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 9.31 (dd, J = 0.8, 2.0 Hz, 1 H), 8.46 (dd, J = 2.0, 8.4 Hz, 1 H), 8.42 (s, 1 H), 8.38 (dd, J = 0.8, 8.4 Hz, 1 H), 4.45 (q, J = 7.2 Hz, 2 H), 4.00 (s, 3 H), 1.42 (t, J = 7.2 Hz, 3 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 164.9, 160.8, 160.2, 151.1, 148.0, 145.2, 138.3, 135.3, 127.1, 122.3, 61.6, 52.7, 14.3; **HRMS** (ESI) m/z 277.0815 [calc'd for $C_{13}H_{13}N_2O_5$ $(M + H)^+$ 277.0823].

2-(4-Carboxyoxazol-2-yl)pyridine-5-carboxylic Acid (pyoxDC*)



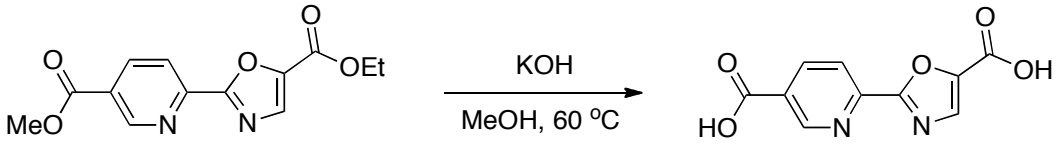
Methyl 2-(4-ethoxycarbonyloxazol-2-yl)-pyridine-5-carboxylate (50 mg, 0.18 mmoles) and KOH (46.7 mg, 0.72 mmoles) were added to a vial. MeOH (5.0 mL) was added to the vial, and the reaction mixture was heated to 60 °C with stirring until the starting material was consumed completely, as judged by TLC. The reaction mixture was cooled and concentrated under reduced pressure, after which the crude product was dissolved in water (2 mL). The aqueous layer was washed with EtOAc (1 × 2 mL), after which the product was precipitated from the aqueous layer by adjusting to pH 3–4 with 1 M HCl. After cooling to 4 °C, the product was filtered, washed with water (3 × 1 mL), and dried under high vacuum to afford the title compound (30 mg, 71%) as a white solid. **^{1}H NMR** (500 MHz, DMSO-d_6, δ): 13.72 (bs, 1 H), 13.40 (bs, 1 H), 9.19 (dd, J = 0.5, 2.0 Hz, 1 H), 9.02 (s, 1 H), 8.47 (dd J = 2.0, 8.5 Hz, 1 H), 8.28 (dd, J = 0.5, 8.5 Hz, 1 H); **^{13}C NMR** (125 MHz, DMSO-d_6, δ): 166.1, 162.2, 160.1, 151.1, 148.0, 147.4, 139.1, 135.4, 128.2, 122.8; **HRMS** (ESI) m/z 233.0195 [calc'd for $C_{10}H_5N_2O_5$ $(M - H)^-$ 233.0203].

Methyl 2-(5-Ethoxycarbonyloxazol-2-yl)pyridine-5-carboxylate (ethylmethyl pyoxDC)



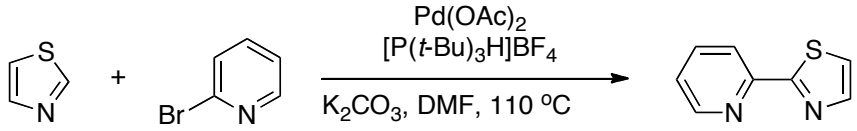
$Pd(OAc)_2$ (31.8 mg, 0.14 mmoles), $[P(t\text{-}Bu)_3H]BF_4$ (123.2 mg, 0.42 mmoles), 1-adamantanecarboxylic acid (153 mg, 0.85 mmoles), K_2CO_3 (782 mg, 5.7 mmoles), and methyl 6-bromonicotinate (918 mg, 4.3 mmoles) were added to a dried flask. The flask was fitted with a reflux condenser capped with a septum, evacuated, and purged with N_2(g) (~5 times). A degassed solution of ethyl oxazole-5-carboxylate (400 mg, 2.8 mmoles) in dry toluene (7 mL) was added via syringe, and the reaction mixture was stirred at 110 °C for 8 h. The reaction mixture was then cooled and filtered through Celite®, and the filtrate was concentrated under reduced pressure. The crude product was purified by chromatography on silica (30% EtOAc in hexanes followed by 2% v/v acetone in 1:1 DCM/hexanes) to afford the title compound (183 mg, 23%) as a white solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 9.37 (dd, J = 0.8, 2.0 Hz, 1 H), 8.47 (dd, J = 2.0, 8.0 Hz, 1 H), 8.30 (dd, J = 0.8, 8 Hz, 1 H), 7.96 (s, 1 H), 4.46 (q, J = 7.2 Hz, 2 H), 4.00 (s, 3 H), 1.43 (t, J = 7.2 Hz, 3 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 164.9, 161.6, 157.5, 151.5, 148.0, 143.8, 138.3, 135.5, 127.3, 122.5, 61.9, 52.7, 14.3; **HRMS** (ESI) m/z 277.0823 [calc'd for $C_{13}H_{13}N_2O_5$ $(M + H)^+$ 279.0819].

2-(5-Carboxyoxazol-2-yl)pyridine-5-carboxylic Acid (pyoxDC)



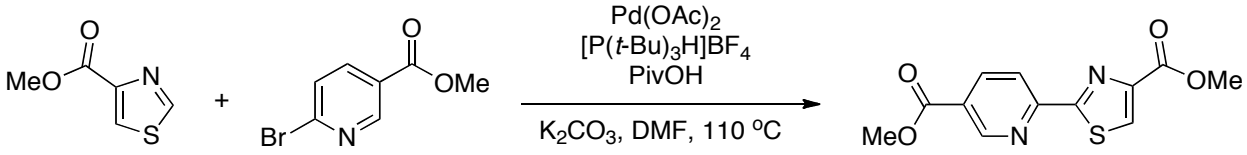
Methyl 2-(5-ethoxycarbonyloxazol-2-yl)-pyridine-5-carboxylate (45 mg, 0.16 mmoles) and KOH (42 mg, 0.65 mmoles). MeOH (4.7 mL) was added to the vial, and the reaction mixture was heated to 60 °C with stirring until the starting material was consumed completely, as judged by TLC. The reaction mixture was cooled and concentrated under reduced pressure. The crude product was dissolved in water (2 mL). The aqueous layer was washed with EtOAc (1 × 2 mL), after which the product was precipitated from the aqueous layer by adjusting to pH 3–4 with 1 M HCl. After cooling to 4 °C, the product was filtered, washed with water (3 × 1 mL), and dried under high vacuum to afford the title compound (24.6 mg, 64%) as a light green solid. **^{1}H NMR** (500 MHz, DMSO-d_6, δ): 13.96 (bs, 1 H), 13.78 (bs, 1 H), 9.22 (dd, J = 1.0, 2.0 Hz, 1 H), 8.48 (dd, J = 2.0, 8.5 Hz, 1 H), 8.23 (dd, J = 1.0, 8.5 Hz, 1 H), 8.16 (s, 1 H); **^{13}C NMR** (125 MHz, DMSO-d_6, δ): 166.1, 161.6, 158.9, 151.2, 147.9, 144.6, 139.1, 135.7, 128.4, 123.3; **HRMS** (ESI) m/z 233.0203 [calc'd for $C_{10}H_5N_2O_5$ $(M - H)^-$ 233.0203].

2-(Thiazol-2-yl)pyridine (pythi)



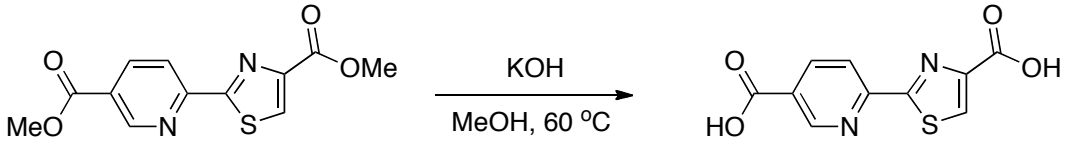
$Pd(OAc)_2$ (35.5 mg, 0.16 mmoles), $[P(t\text{-}Bu)_3H]BF_4$ (137.5 mg, 0.47 mmoles), and K_2CO_3 (874 mg, 6.3 mmoles) were added to a dried flask. The flask was fitted with a reflux condenser capped with a septum, evacuated, and purged with N_2(g) (~5 times). A degassed solution of thiazole (1.08 g, 12.7 mmoles) and 2-bromopyridine (500 mg, 3.2 mmoles) in dry DMF (16 mL) was added via syringe, and the reaction mixture was stirred at 110 °C for 24 h. The reaction mixture was cooled, H_2O (60 mL) was added, and the aqueous layer was extracted with ether (4 × 60 mL). The combined organic extracts were washed with H_2O (1 × 100 mL) and brine (1 × 100 mL), dried over Na_2SO_4(s), and concentrated under reduced pressure. The crude product was purified by silica gel chromatography (20–40% v/v EtOAc in hexanes followed by 2% v/v acetone in 1:1 DCM/hexanes) to afford the title compound (142 mg, 28%) as a white solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 8.58 (d, J = 4.8 Hz, 1 H), 8.16 (d, J = 8.0 Hz, 1 H), 7.89 (d, J = 3.2 Hz, 1 H), 7.75 (td, J = 1.6, 8.0 Hz, 1 H), 7.41 (d, J = 3.2 Hz, 1 H), 7.26 (ddd, J = 1.2, 4.8, 8.0 Hz, 1 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 169.3, 151.3, 149.4, 144.0, 137.0, 124.4, 121.4, 119.6; **HRMS** (EI) m/z 162.0247 [calc'd for $C_8H_7N_2S$ $(M)^+$162.0247].

Methyl 2-(4-Methoxycarbonylthiazol-2-yl)pyridine-5-carboxylate (dimethyl pythiDC*)



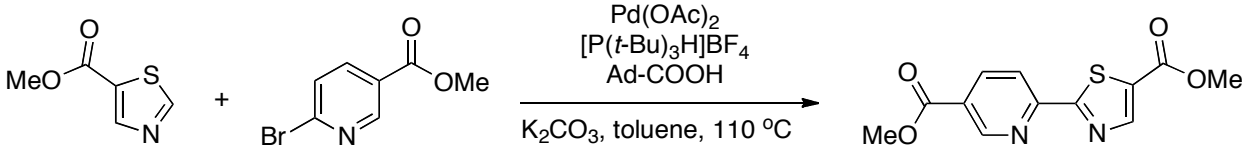
$Pd(OAc)_2$ (15.7 mg, 0.070 mmoles), $[P(t\text{-}Bu)_3H]BF_4$ (60.7 mg, 0.21 mmoles), pivalic acid (28.5 mg, 0.28 mmoles), K_2CO_3 (386 mg, 2.8 mmoles), methyl thiazole-4-carboxylate (200 mg, 1.4 mmoles), and methyl 6-bromonicotinate (603 mg, 2.8 mmoles) were added to a dried flask. The flask was fitted with a reflux condenser capped with a septum, evacuated, and purged with N_2(g) (~5 times). Dry DMF (7 mL) was added via syringe, and the reaction mixture was stirred at 110 °C for 6 h. The reaction mixture was then cooled and filtered through Celite®, and the filtrate was concentrated under reduced pressure. The crude product was purified by chromatography on silica (1% v/v acetone in DCM) to afford the title compound (39 mg, 10%) as a white solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 9.18 (dd, J = 0.8, 2.0 Hz, 1 H), 8.41 (dd, J = 2.0, 8.4 Hz, 1 H), 8.38 (dd, J = 0.8, 8.4 Hz, 1 H), 8.33 (s, 1 H), 3.99 (s, 3 H), 3.98 (s, 3 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 168.7, 165.1, 161.7, 153.3, 150.7, 148.3, 138.3, 130.7, 126.9, 119.6, 52.6, 52.6; **HRMS** (ESI) m/z 279.0423 [calc'd for $C_{12}H_{11}N_2O_4S$ $(M + H)^+$ 279.0435].

2-(4-Carboxythiazol-2-yl)pyridine-5-carboxylic Acid (pythiDC*)



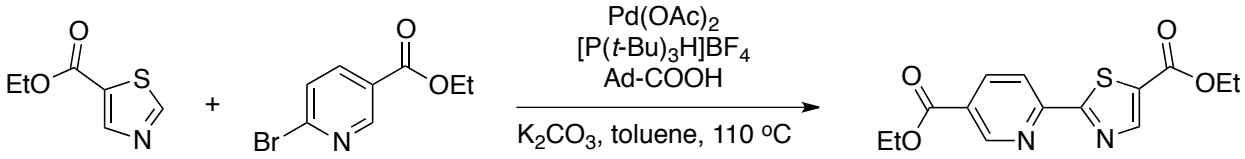
Methyl 2-(4-methoxycarbonylthiazol-2-yl)-pyridine-5-carboxylate (30 mg, 0.11 mmoles) and KOH (27.8 mg, 0.43 mmoles) were added to a vial. MeOH (3.1 mL) was added to the vial, and the reaction mixture was heated to 60 °C with stirring until the starting material was consumed completely, judged by TLC. The reaction mixture was cooled and concentrated under reduced pressure. The crude product was dissolved in water (2 mL). The aqueous layer was washed with EtOAc (1 × 2 mL), after which the product was precipitated from the aqueous layer by adjusting to pH 3–4 with 1 M HCl. After cooling to 4 °C, the product was filtered, washed with water (3 × 1 mL), and dried under high vacuum to afford the title compound (21 mg, 78%) as a white solid. **^{1}H NMR** (500 MHz, DMSO-d_6, δ): 13.51 (bs, 2 H), 9.13 (dd, J = 0.5, 2.0 Hz, 1 H), 8.67 (s, 1 H), 8.47 (dd, J = 2.0, 8.5 Hz, 1 H), 8.28 (dd, J = 0.5, 8.5 Hz, 1 H); **^{13}C NMR** (125 MHz, DMSO-d_6, δ): 168.0, 166.2, 162.5, 153.0, 151.1, 149.4, 139.4, 132.2, 128.6, 119.8; **HRMS** (ESI) m/z 248.9975 [calc'd for $C_{10}H_5N_2O_4S$ $(M - H)^-$ 248.9975].

Methyl 2-(5-Methoxycarbonylthiazol-2-yl)pyridine-5-carboxylate (dimethyl pythiDC)



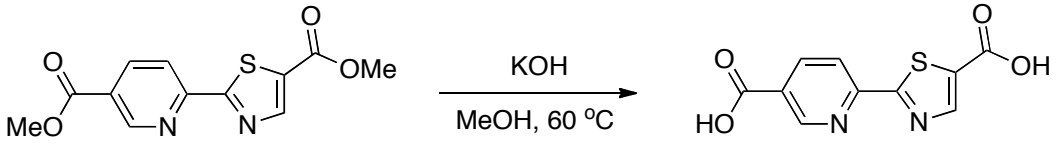
$Pd(OAc)_2$ (3.9 mg, 0.017 mmoles), $[P(t\text{-}Bu)_3H]BF_4$ (10.0 mg, 0.035 mmoles), 1-adamantanecarboxylic acid (18.4 mg, 0.10 mmoles), K_2CO_3 (96.5 mg, 0.70 mmoles), methyl thiazole-5-carboxylate (50 mg, 0.35 mmoles), and methyl 6-bromonicotinate (113 mg, 0.52 mmoles) were added to a dried flask. The flask was fitted with a reflux condenser capped with a septum, evacuated, and purged with N_2(g) (~5 times). Dry toluene (1.7 mL) was added via syringe, and the reaction mixture was stirred at 110 °C for 48 h. The reaction mixture was then cooled and filtered through Celite®, and the filtrate was concentrated under reduced pressure. The crude product was purified by chromatography on silica (3% v/v acetone in 1:1 DCM/hexanes) to afford the title compound (28 mg, 29%) as a white solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 9.22 (s, 1 H), 8.51 (s, 1 H), 8.43 (dd, J = 2.0, 8.0 Hz, 1 H), 8.29 (d, J = 8.0 Hz, 1 H), 4.00 (s, 3 H), 3.96 (s, 3 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 172.7, 165.1, 161.8, 153.6, 150.9, 149.7, 138.4, 131.7, 127.1, 119.6, 52.6, 52.6; **HRMS** (ESI) m/z 279.0439 [calc'd for $C_{12}H_{11}N_2O_4S$ $(M + H)^+$ 279.0435].

Ethyl 2-(5-Ethoxycarbonylthiazol-2-yl)pyridine-5-carboxylate (diethyl pythiDC)



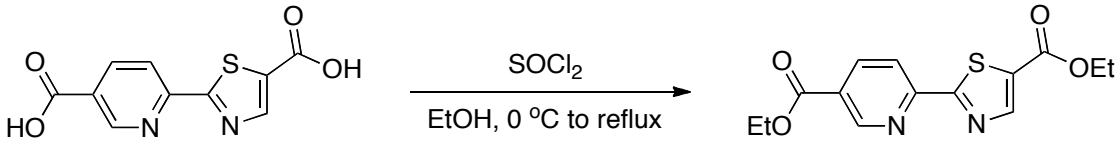
$Pd(OAc)_2$ (150 mg, 0.67 mmoles), $[P(t\text{-}Bu)_3H]BF_4$ (388 mg, 1.34 mmoles), 1-adamantanecarboxylic acid (723 mg, 4.01 mmoles), K_2CO_3 (3.7 g, 27 mmoles), and ethyl 6-bromonicotinate (4.6 g, 20 mmoles) were added to a dried flask. The flask was fitted with a reflux condenser capped with a septum, evacuated, and purged with N_2(g) (~5 times). A solution of ethyl thiazole-5-carboxylate (2.1 g, 13 mmoles) in dry toluene (65 mL) was added, and the resulting solution was stirred at 110 °C for 72 h. The reaction mixture was cooled and filtered through Celite®, and the filtrate was concentrated under reduced pressure. The crude product was purified by chromatography on silica (25% EtOAc/hexanes followed by 3% v/v acetone in 1:1 DCM/hexanes) and recrystallized from 1:1 EtOAc/hexanes to afford the title compound (1.48 g, 40%) as a white crystalline solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 9.22 (s, 1 H), 8.51 (s, 1 H), 8.43 (dd, J = 2.0, 8.0 Hz, 1 H), 8.29 (d, J = 8.0 Hz, 1 H), 4.00 (s, 3 H), 3.96 (s, 3 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 172.7, 165.1, 161.8, 153.6, 150.9, 149.7, 138.4, 131.7, 127.1, 119.6, 52.6, 52.6; **HRMS** (ESI) m/z 279.0439 [calc'd for $C_{12}H_{11}N_2O_4S$ $(M + H)^+$ 279.0435].

2-(5-Carboxythiazol-2-yl)pyridine-5-carboxylic Acid (pythiDC)



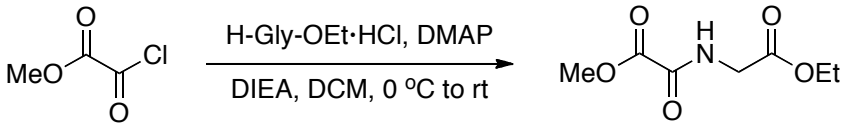
Methyl 2-(5-methoxycarbonylthiazol-2-yl)-pyridine-5-carboxylate (70 mg, 0.25 mmoles) and KOH (63.1 mg, 1.0 mmoles) were added to a vial. MeOH (7.0 mL) was added to the vial, and the reaction mixture was heated to 60 °C with stirring until the starting material was consumed completely, as judged by TLC. The reaction mixture was cooled and concentrated under reduced pressure. The crude product was dissolved in water (4 mL). The aqueous layer was washed with EtOAc (1 × 4 mL), after which the product was precipitated from the aqueous layer by adjusting to pH 2 with 1 M HCl. After cooling to 4 °C, the product was filtered, washed with water (3 × 1 mL), and dried under high vacuum to afford the title compound (38 mg, 61%) as a pale yellow solid. **^{1}H NMR** (500 MHz, DMSO-d_6, δ): 13.80 (bs, 1 H), 13.73 (bs, 1 H), 9.15 (dd, J = 0.5, 2.0 Hz, 1 H), 8.56 (s, 1 H), 8.48 (dd, J = 2.0, 8.0 Hz, 1 H), 8.31 (dd, J = 0.5, 8.0 Hz, 1 H); **^{13}C NMR** (125 MHz, DMSO-d_6, δ): 172.1, 166.1, 162.6, 153.1, 151.2, 149.9, 139.5, 133.6, 128.6, 120.1; **HRMS** (ESI) m/z 248.9977 [calc'd for $C_{10}H_5N_2O_4S$ $(M - H)^-$ 248.9975].

Ethyl 2-(5-Ethoxycarbonylthiazol-2-yl)pyridine-5-carboxylate (diethyl pythiDC)



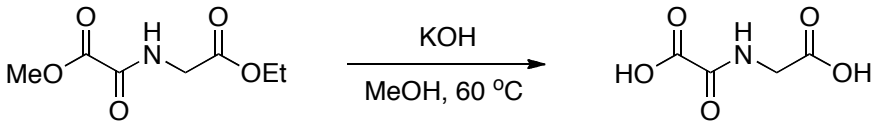
2-(5-Carboxythiazol-2-yl)pyridine-5-carboxylic acid (38 mg, 0.15 mmoles) and EtOH (7 mL) were added to a dried flask and stirred on ice. Thionyl chloride (300 μL) was added dropwise on ice, after which the flask was fitted with a reflux condenser and heated at reflux. After 24 h, the reaction mixture was cooled on ice and quenched by the dropwise addition of saturated aqueous Na_2CO_3 (10 mL). The aqueous layer was extracted with CH_2Cl_2 (3 × 10 mL) and the combined organic extracts were dried over Na_2SO_4(s) and concentrated under reduced pressure. The crude product was purified by chromatography on silica (2% v/v acetone in 1:1 DCM/hexanes) to afford the title compound (40 mg, 86 %) as a pale yellow solid. The spectral data for diethyl pythiDC synthesized by this method matched that reported above for the same compound synthesized by the direct cross coupling described above.

***N*-(Methoxyoxalyl)glycine Ethyl Ester**



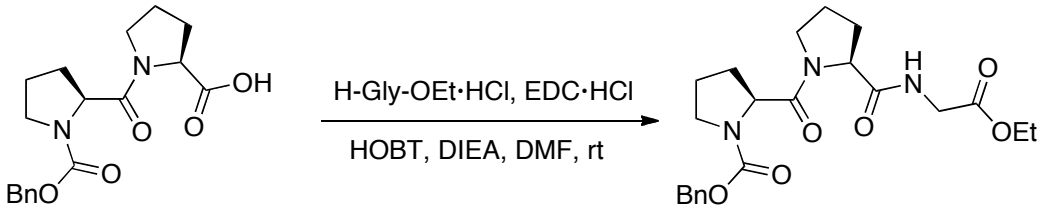
N-(Methoxyoxalyl)glycine ethyl ester was synthesized as reported previously.[148] The spectral data and yield matched that reported previously.

_N_-Oxalylglycine (NOG)



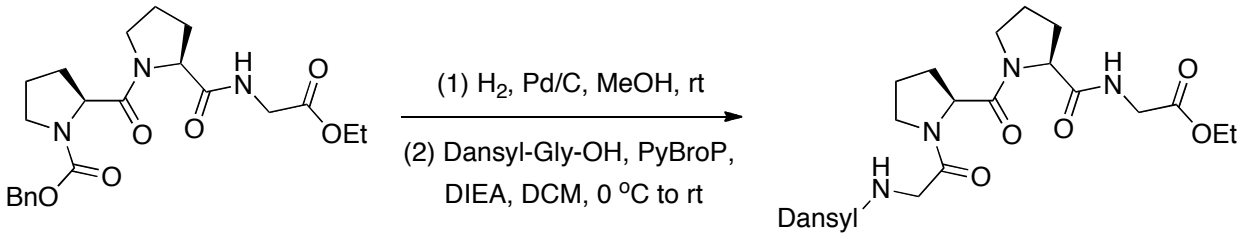
_N_-Oxalylglycine was synthesized as described previously.[148] The spectral data and yield matched that reported previously.

_N_-Benzyloxycarbonyl-(2_S_)-prolyl-(2_S_)-prolylglycine Ethyl Ester (CbzProProGlyOEt)



CBbzProProGlyOEt was synthesized as described previously.[148] The spectral data and yield matched that reported previously.

_N_-Dansylglycyl-(2_S_)-prolyl-(2_S_)-prolylglycine Ethyl Ester (dansylGlyProProGlyOEt):



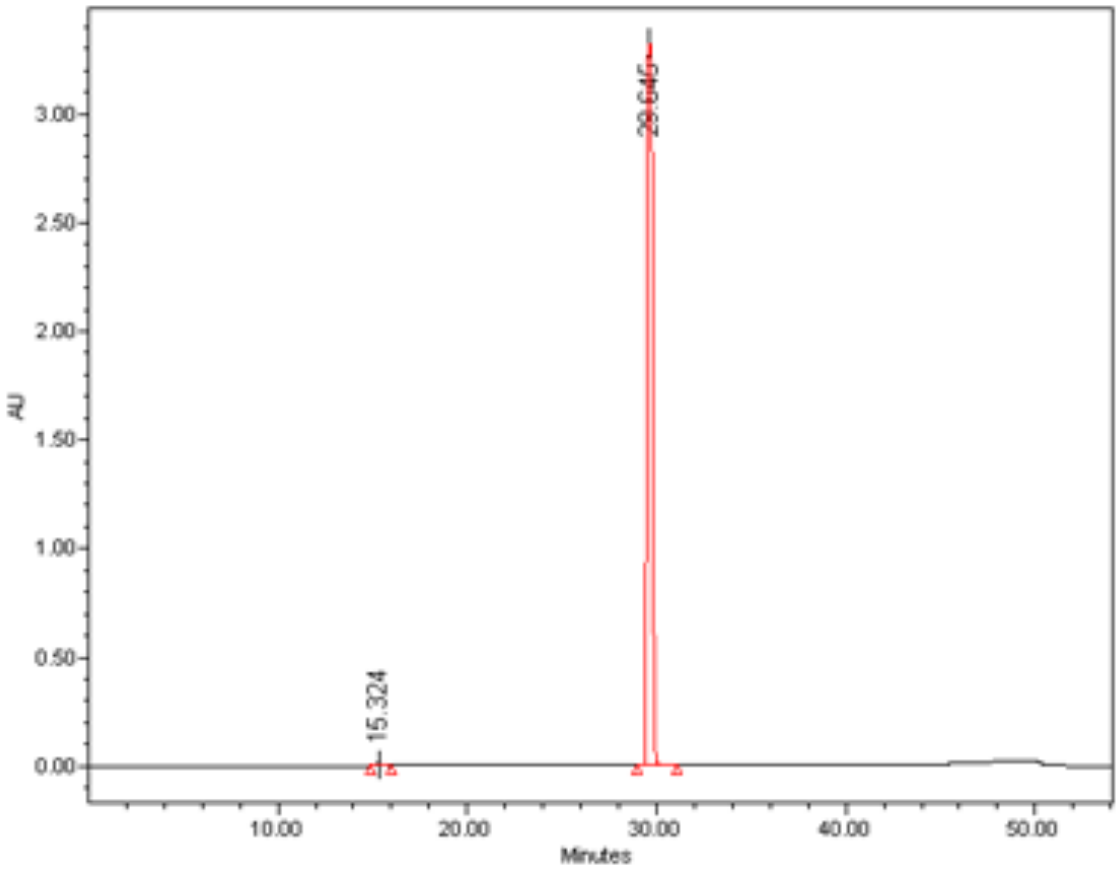
DansylGlyProProGlyOEt was synthesized as described previously.[148] The spectral data and yield matched that reported previously.

3.5.18 Determination of Compound Purity

Compounds were analyzed by HPLC to assess their purity. Analyses were performed with a Waters HPLC system equipped with a Nucleodur® C18 Gravity reversed-phase column (4.6 × 250 mm, 5-μm particle size) from Macherey–Nagel. Compounds dissolved in 50 μL of H_2O were injected onto the column and eluted at 1 mL/min over 32 min with a gradient of aqueous acetonitrile (5–95% v/v containing 0.1% v/v TFA). The maximal absorbance in the range of 210–400 nm was used as the detection wavelength.

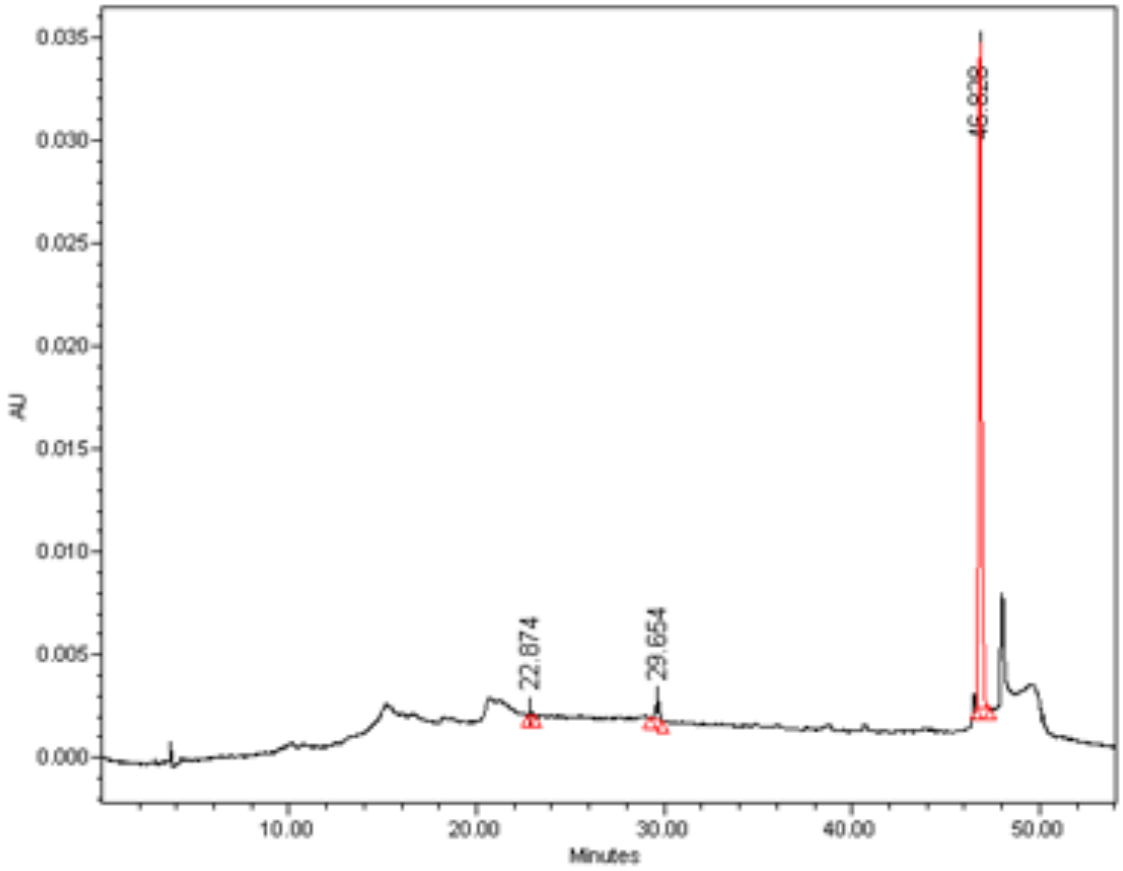
3.5.19 HPLC Chromatograms

EDHB



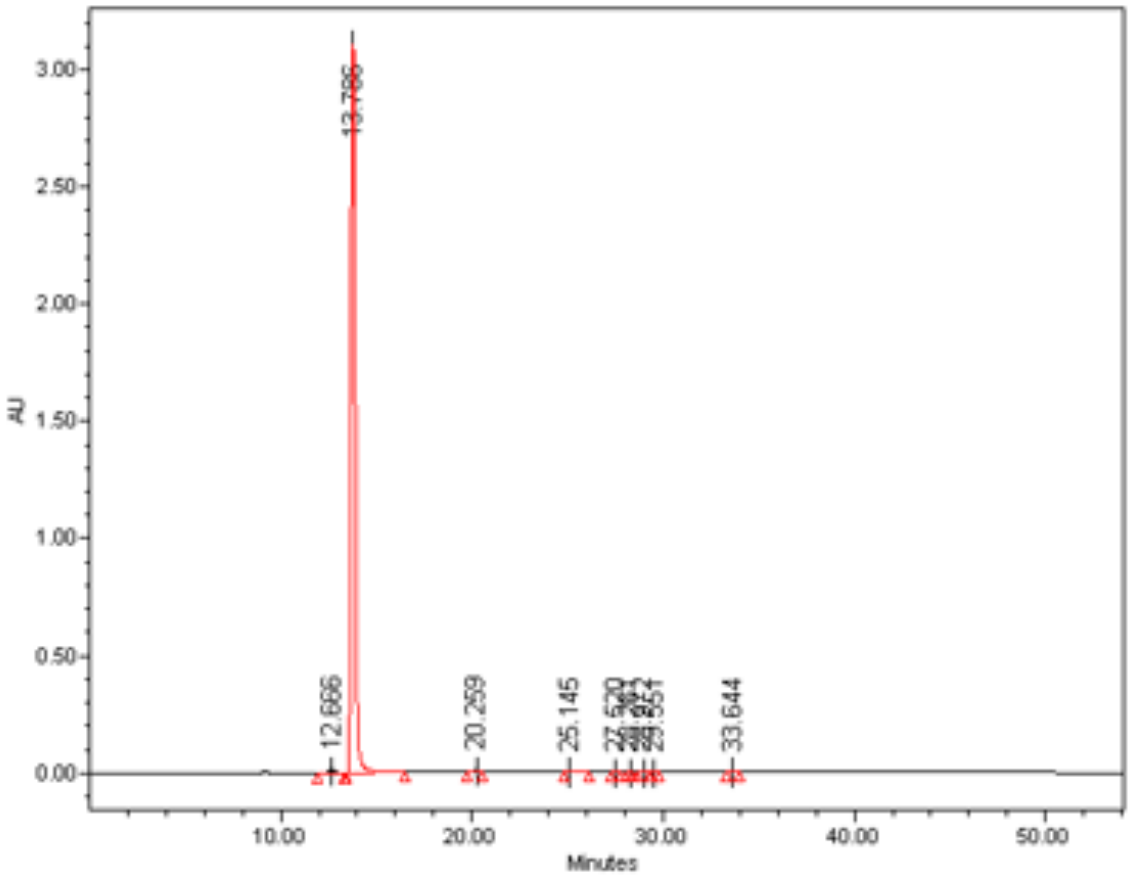
Peak	Retention Time (min)	Area (μV·s)	% Area	Height (μV)
1	15.324	150595	0.22	10840
2	**29.645**	**66918147**	**99.78**	**3314943**

Diethyl Bipy55′DC



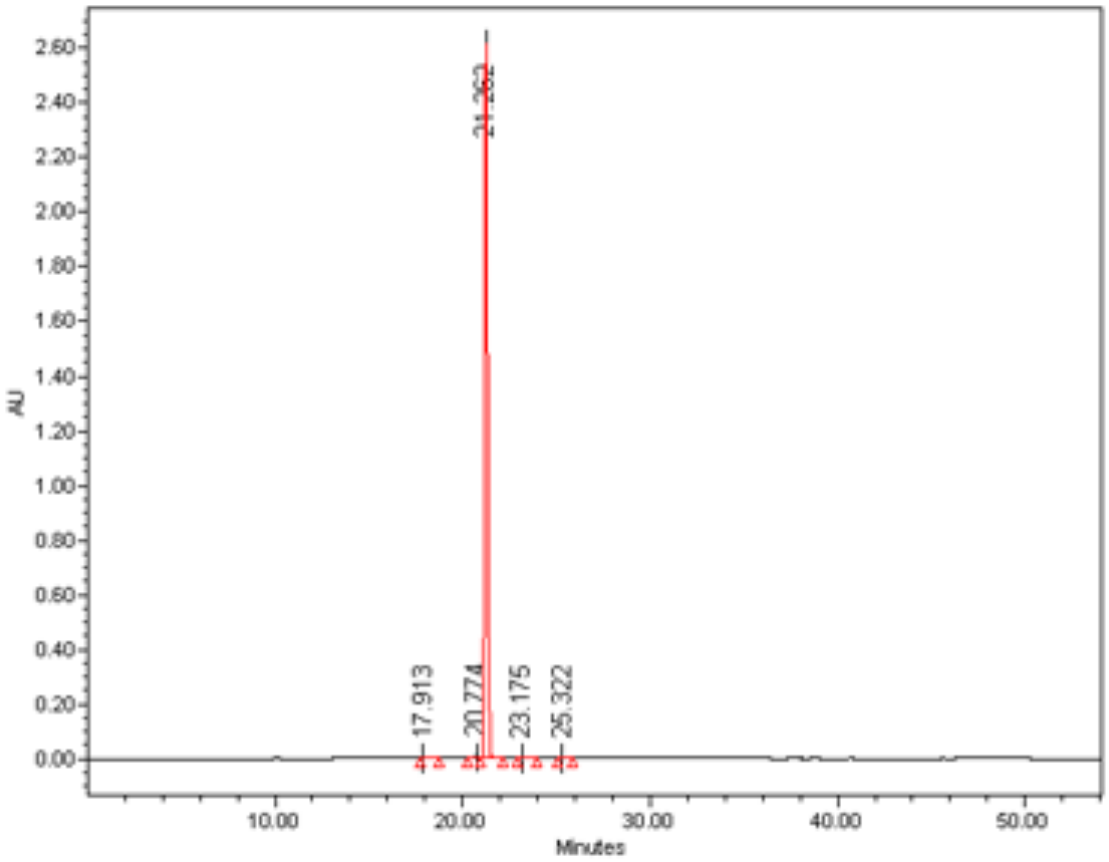
Peak	Retention Time (min)	Area (μV·s)	% Area	Height (μV)
1	22.874	2381	0.71	244
2	29.654	12873	3.84	945
3	**46.828**	**320086**	**95.45**	**32270**

Pypyrid



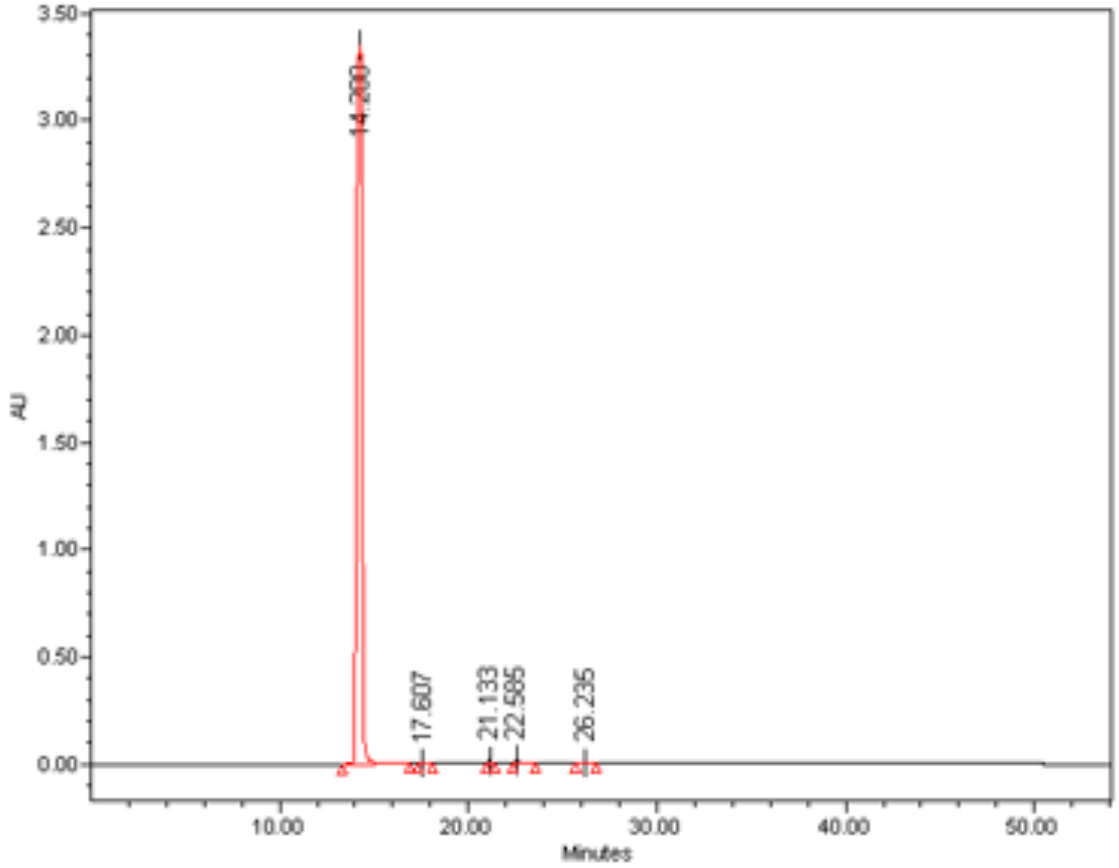
Peak	Retention Time (min)	Area (µV·s)	% Area	Height (µV)
1	12.666	182786	0.38	11247
2	**13.786**	**47247296**	**99.05**	**3107888**
3	20.259	110278	0.23	6186
4	25.145	99102	0.21	3873
5	27.520	6972	0.01	535
6	28.261	10478	0.02	1022
7	28.972	10369	0.02	785
8	29.551	7661	0.02	641
9	33.644	23147	0.05	1836

PypyridDC



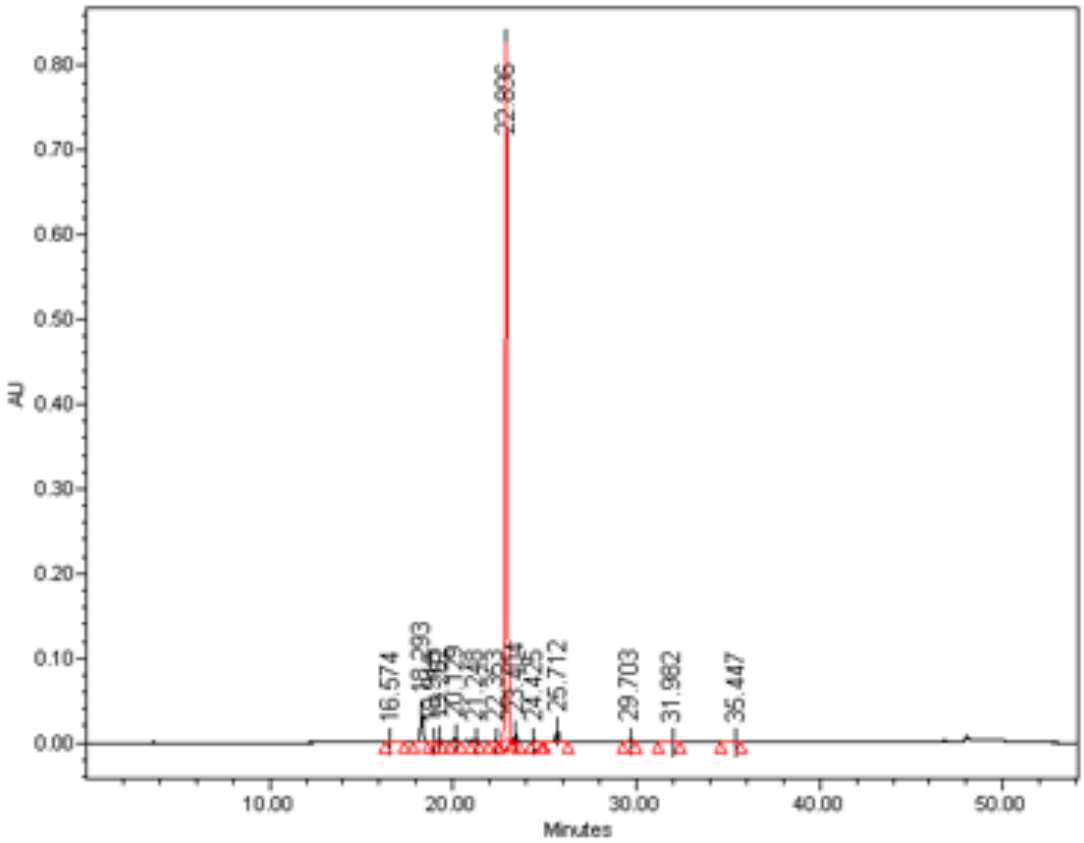
Peak	Retention Time (min)	Area (μV·s)	% Area	Height (μV)
1	17.913	30262	0.11	2810
2	20.774	28341	0.10	2501
3	**21.262**	**27550533**	**99.58**	**2615842**
4	23.175	38610	0.14	1904
5	25.322	19995	0.07	1672

Pypyraz



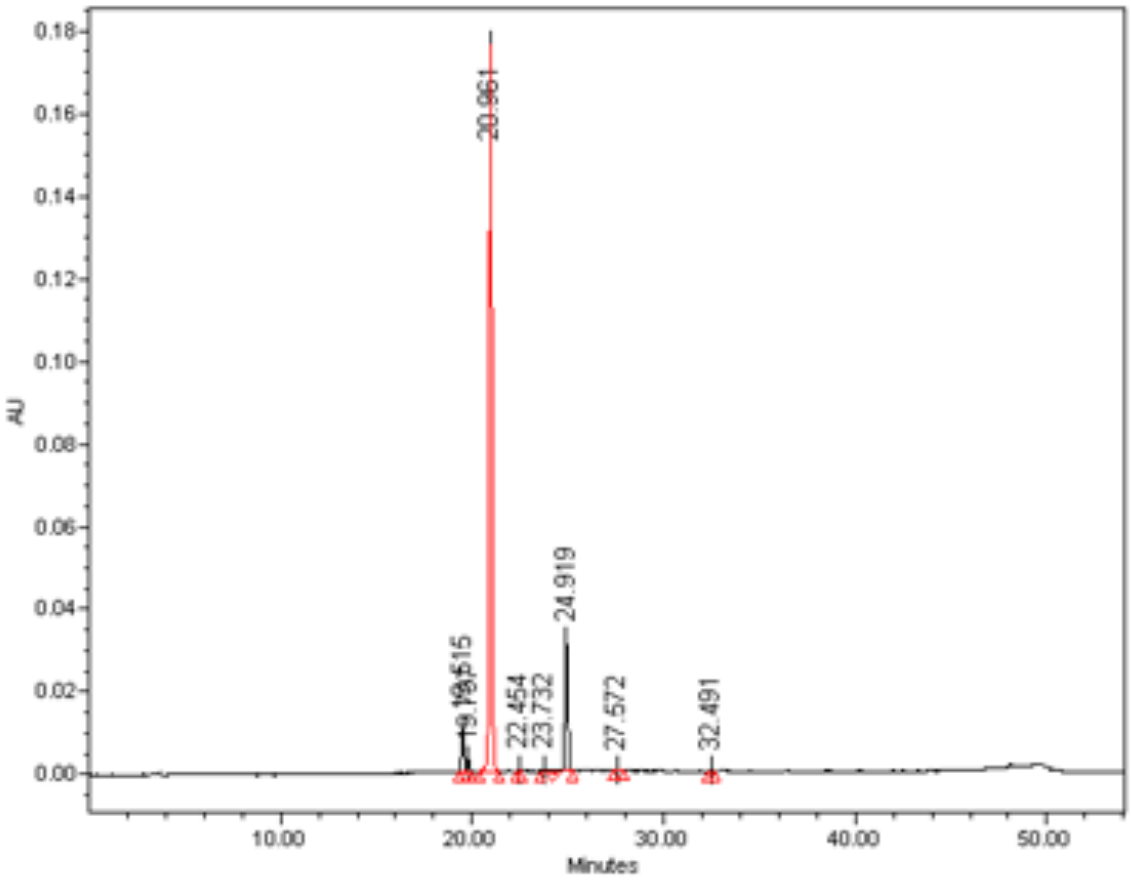
Peak	Retention Time (min)	Area (μV·s)	% Area	Height (μV)
1	**14.280**	**67339739**	**99.13**	**3343388**
2	17.607	72606	0.11	6741
3	21.133	164524	0.24	14808
4	22.585	271119	0.4	19513
5	26.235	79762	0.12	6994

PypyrazDC



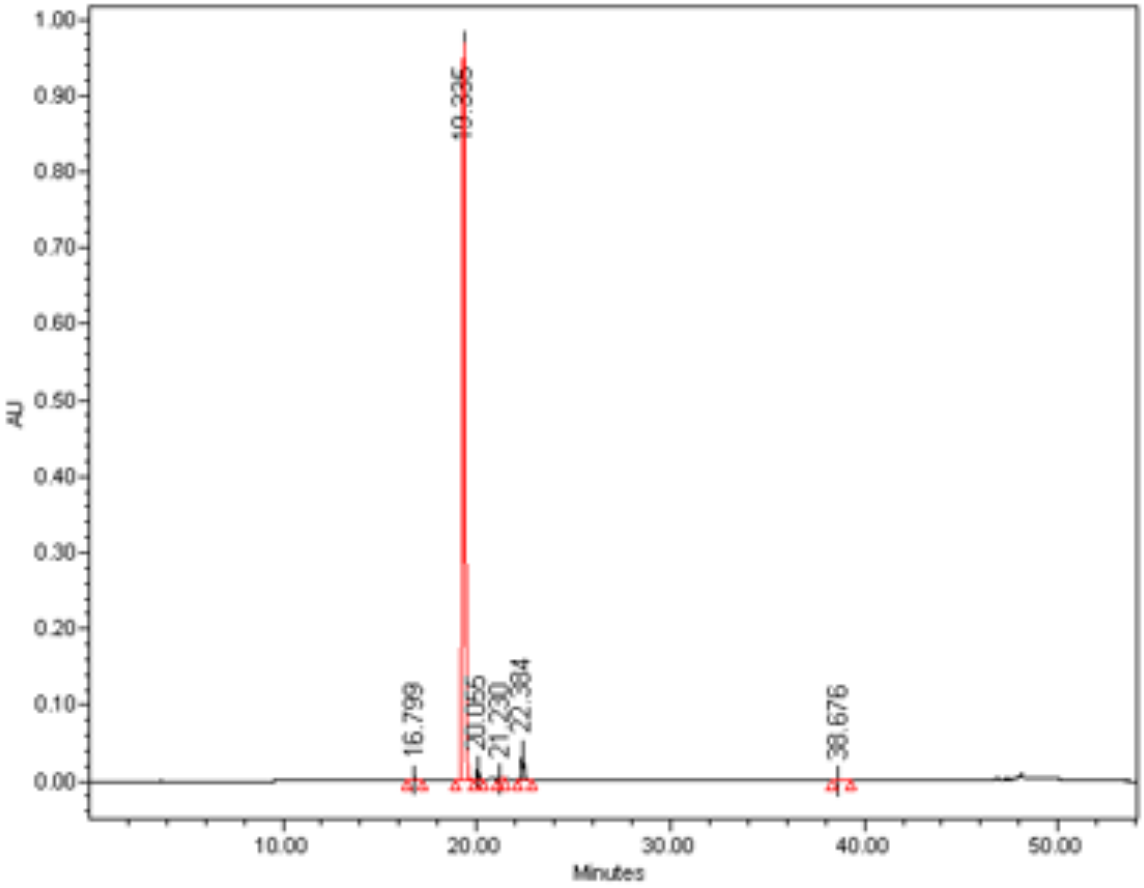
Peak	Retention Time (min)	Area (μV·s)	% Area	Height (μV)
1	16.574	15613	0.18	525
2	18.293	335430	3.85	34147
3	18.915	4923	0.06	446
4	19.282	32690	0.37	3363
5	20.129	57338	0.66	4989
6	21.248	11891	0.14	1009
7	22.353	9429	0.11	1039
8	**22.896**	**7967853**	**91.37**	**828483**
9	23.404	83288	0.96	8722
10	24.425	9131	0.10	858
11	25.712	162359	1.86	13071
12	29.703	14965	0.17	1262
13	31.982	7611	0.09	263
14	35.447	7857	0.09	370

PypyrroleDC



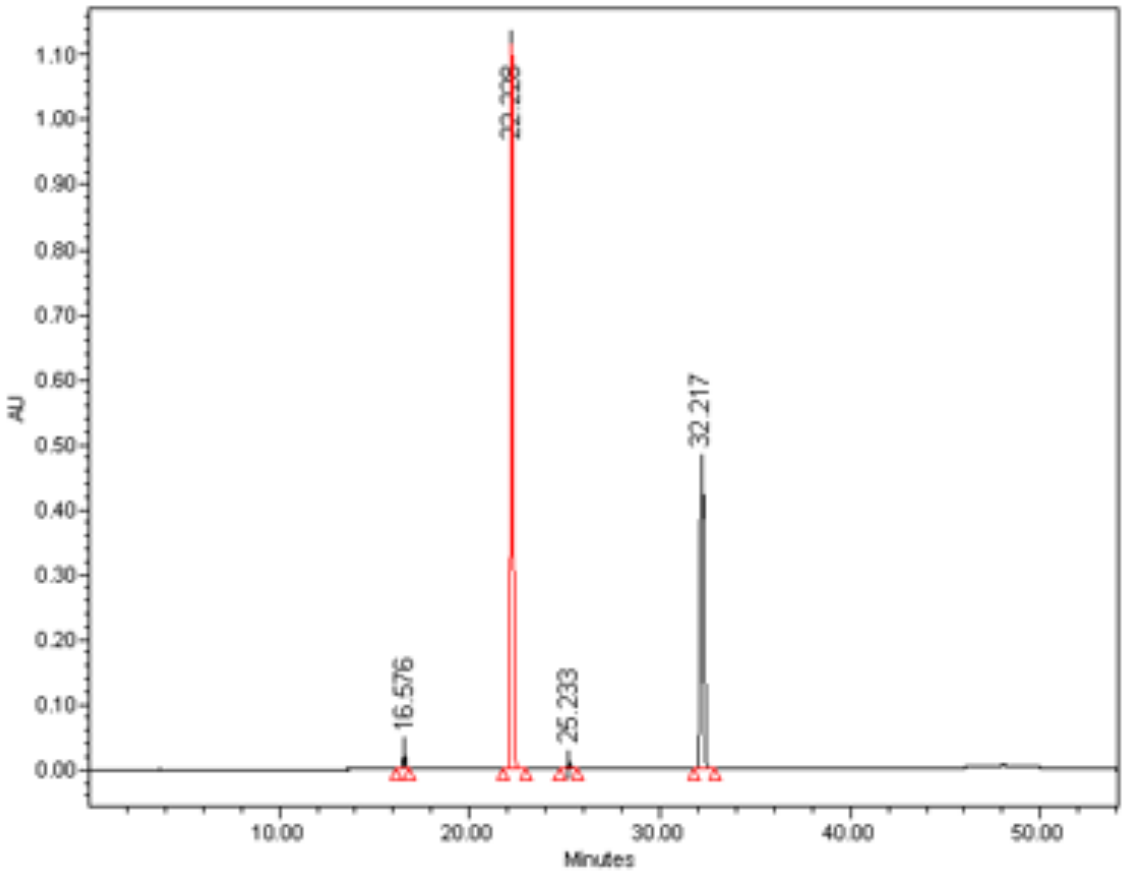
Peak	Retention Time (min)	Area (μV·s)	% Area	Height (μV)
1	**19.515**	99453	4.40	10429
2	19.797	28183	1.25	2870
3	**20.961**	**1789909**	**79.25**	**176621**
4	22.454	4889	0.22	588
5	22.732	5495	0.24	482
6	24.919	323630	14.33	31386
7	27.572	2766	0.12	259
8	32.491	4252	0.19	385

PypyrDC



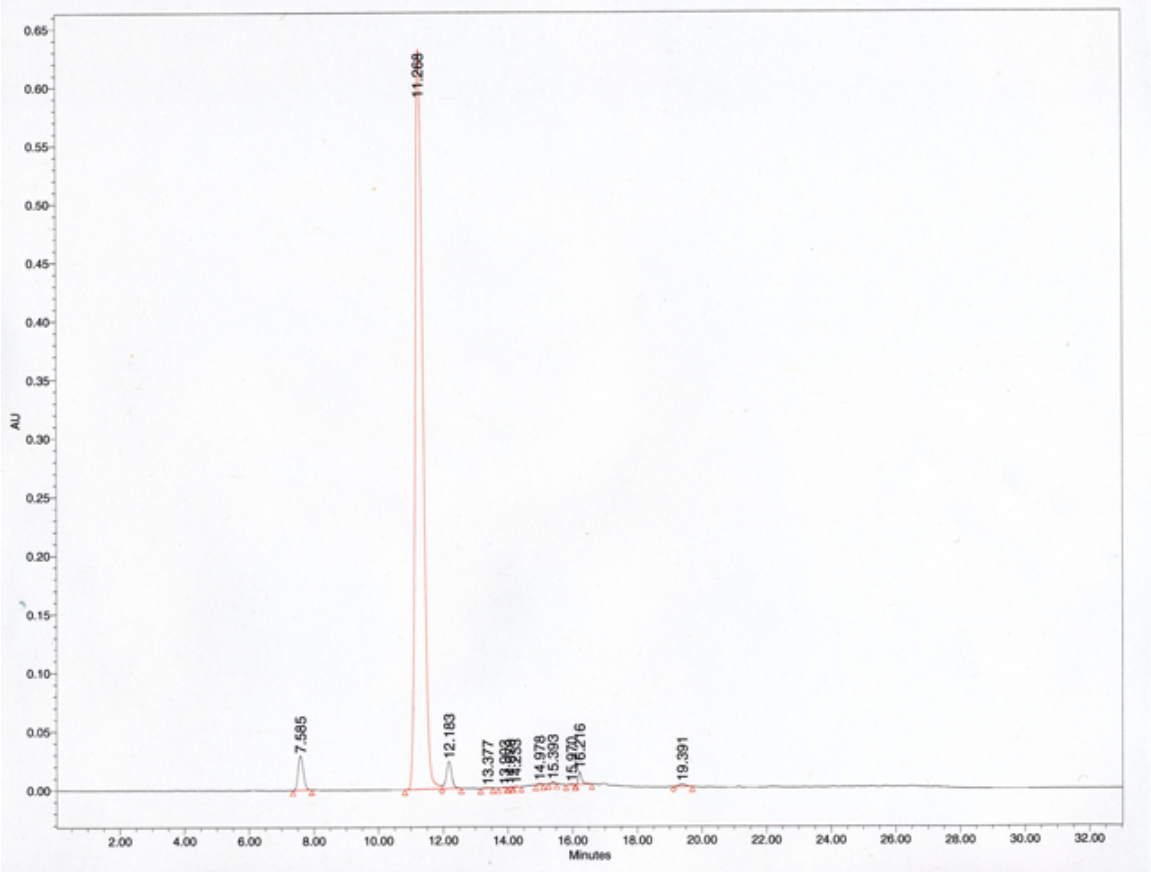
Peak	Retention Time (min)	Area (μV·s)	% Area	Height (μV)
1	16.799	14556	0.15	1315
2	**19.335**	**9322099**	**95.10**	**966494**
3	20.055	123856	1.26	12778
4	21.230	14138	0.14	1516
5	22.384	319544	3.26	34363
6	38.676	8026	0.08	386

2-(5-Methoxycarbonyl-1H-imidazol-2-yl)pyridine-5-carboxylic Acid



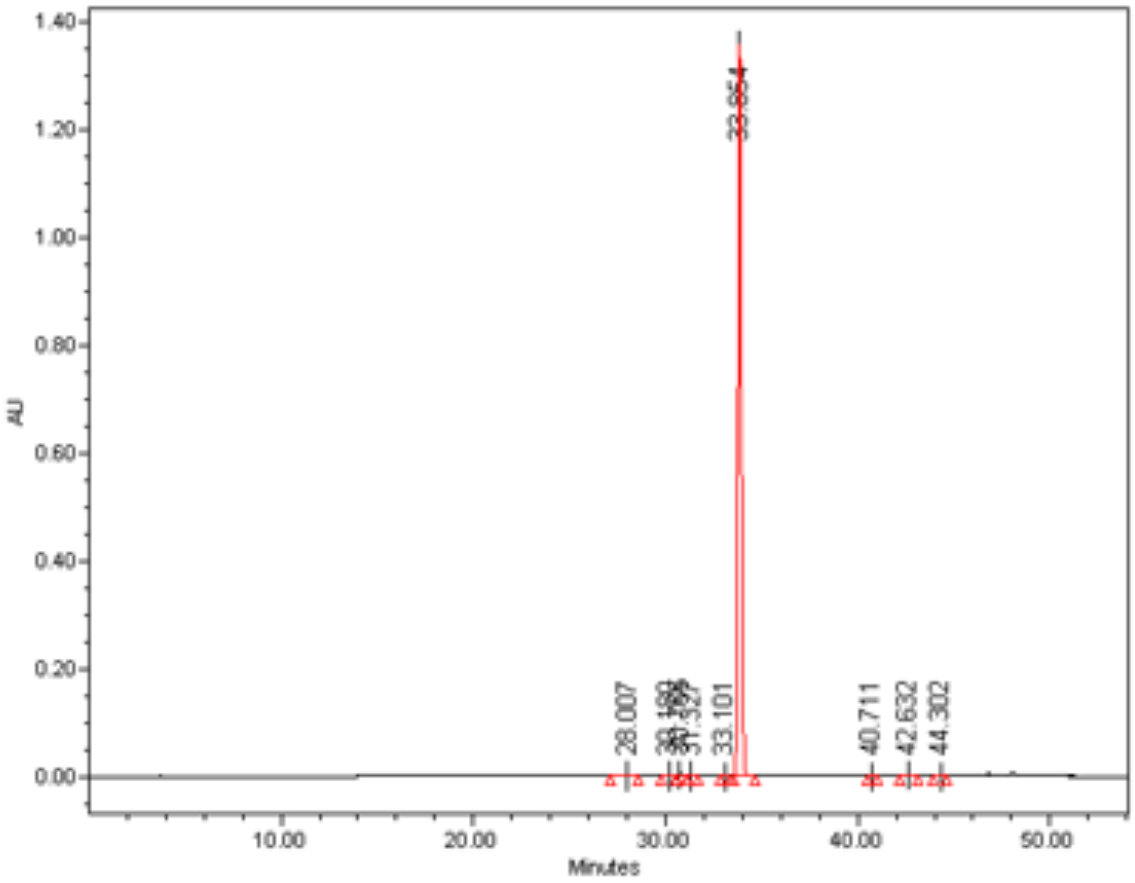
Peak	Retention Time (min)	Area (μV·s)	% Area	Height (μV)
1	16.576	239579	1.47	27336
2	**22.228**	**10474669**	**64.07**	**1113188**
3	25.233	86460	0.53	8268
4	32.217	5549123	33.94	466031

PyimDC



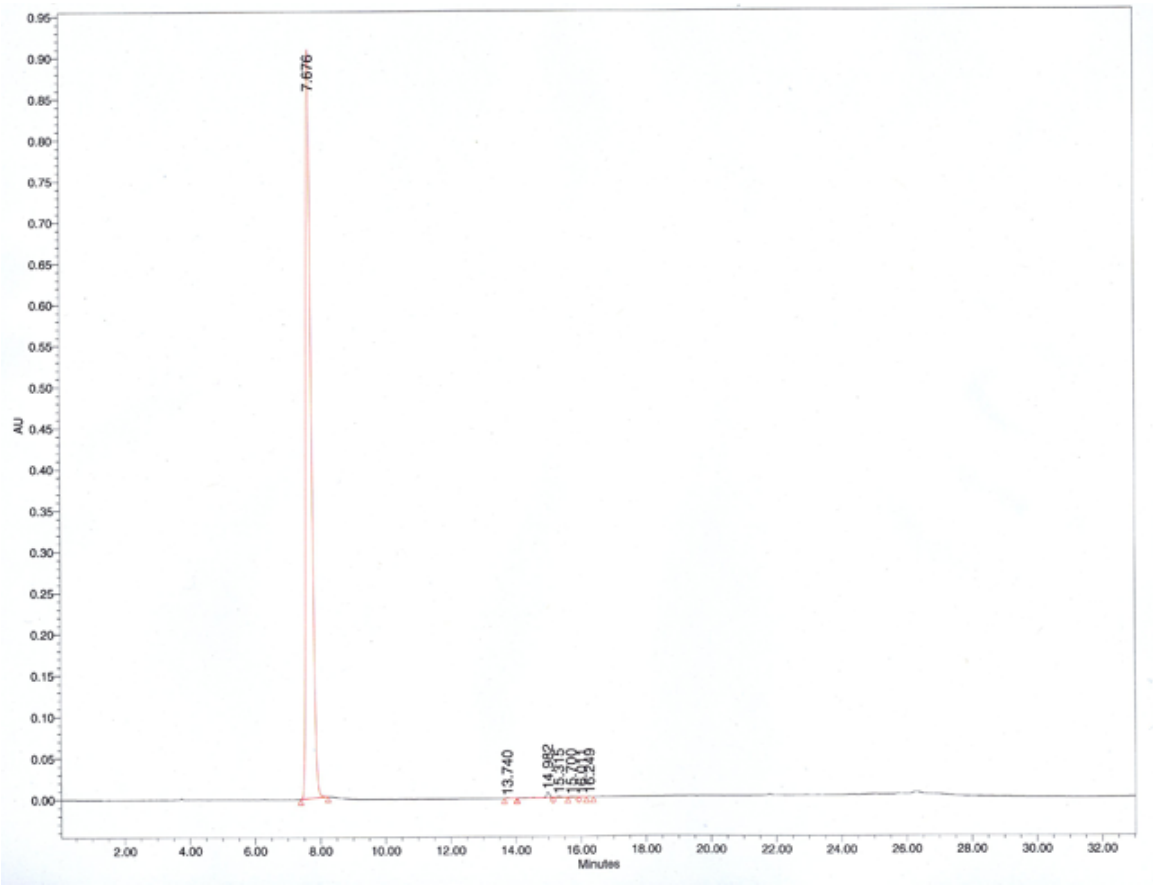
Peak	Retention Time (min)	Area (μV·s)	% Area	Height (μV)
1	7.585	327025	2.95	29352
2	**11.268**	**10327125**	**93.26**	**632576**
3	12.183	271399	2.45	23307
4	13.377	5513	0.05	431
5	13.903	2916	0.03	396
6	14.070	2029	0.02	413
7	14.233	2007	0.02	323
8	14.978	6827	0.06	1163
9	15.393	11308	0.10	1505
10	15.970	2620	0.02	339
11	16.216	82209	0.74	11228
12	19.391	32836	0.30	2188

Diethyl PyimDC



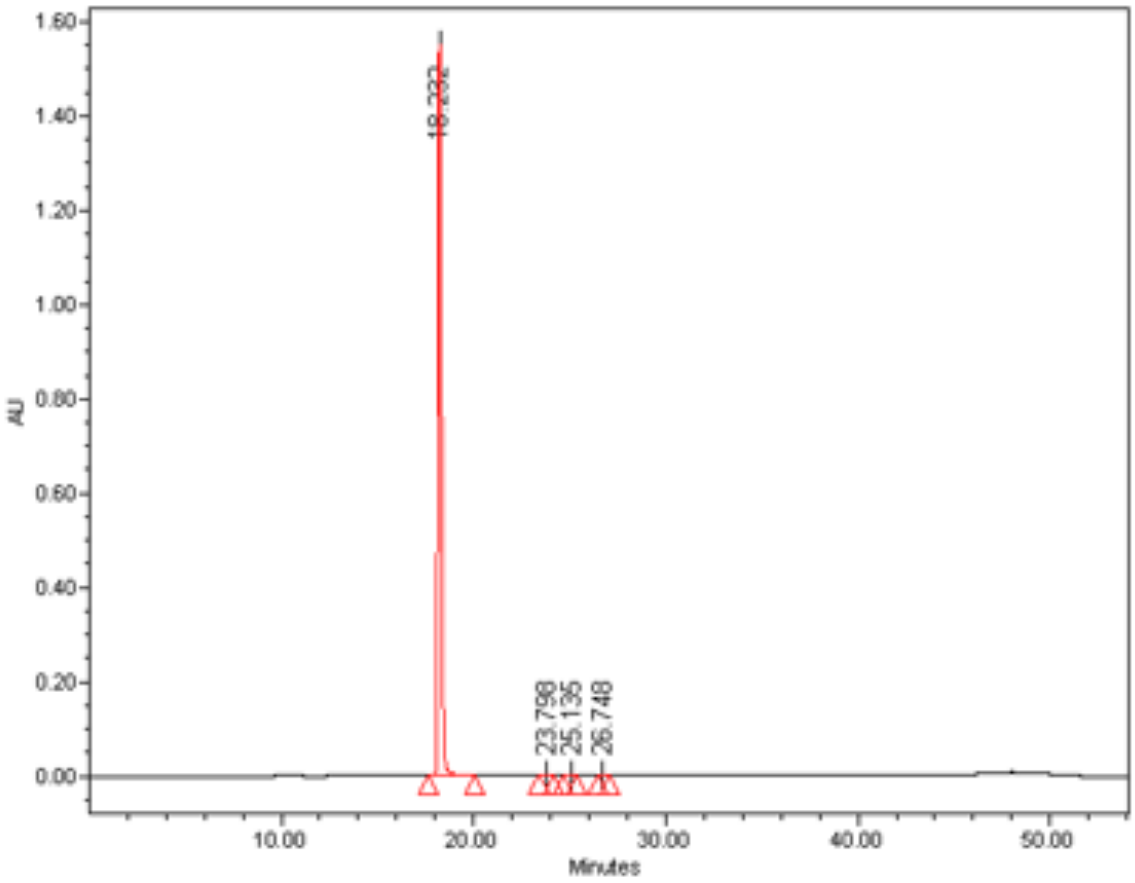
Peak	Retention Time (min)	Area (μV·s)	% Area	Height (μV)
1	28.007	19751	0.12	560
2	30.180	9650	0.06	865
3	30.755	46934	0,28	4158
4	31.327	17379	0.11	1513
5	33.101	5461	0.03	448
6	**33.854**	**16407846**	**99.14**	**1356438**
7	40.711	3385	0.02	240
8	42.632	32365	0.20	2394
9	44.302	7665	0.05	517

NMe-PyimDC



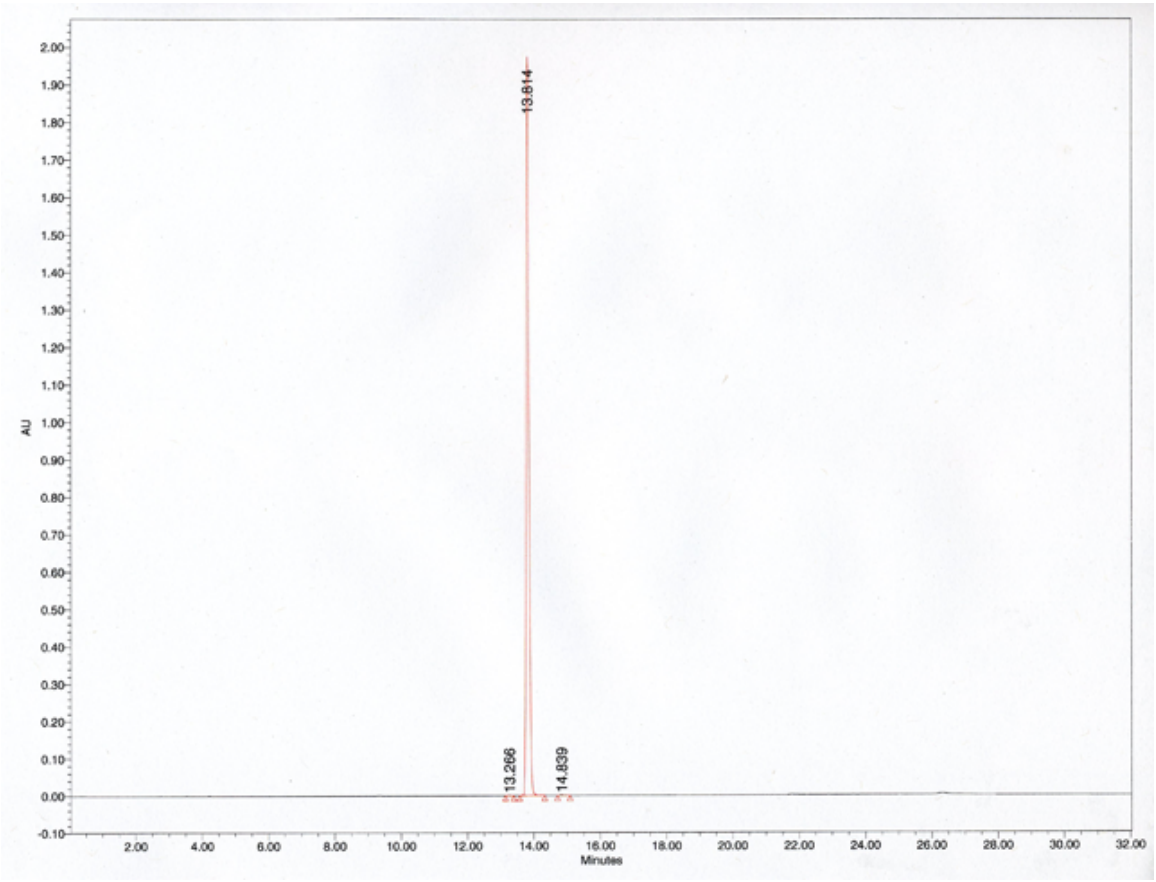
Peak	Retention Time (min)	Area (μV·s)	% Area	Height (μV)
1	**7.676**	**10301024**	**99.36**	**908220**
2	13.74	1890	0.02	232
3	14.982	46920	0.45	6863
4	15.315	8943	0.09	1305
5	15.7	4949	0.05	626
6	16.011	1809	0.02	272
7	18.249	2222	0.02	407

Pyox



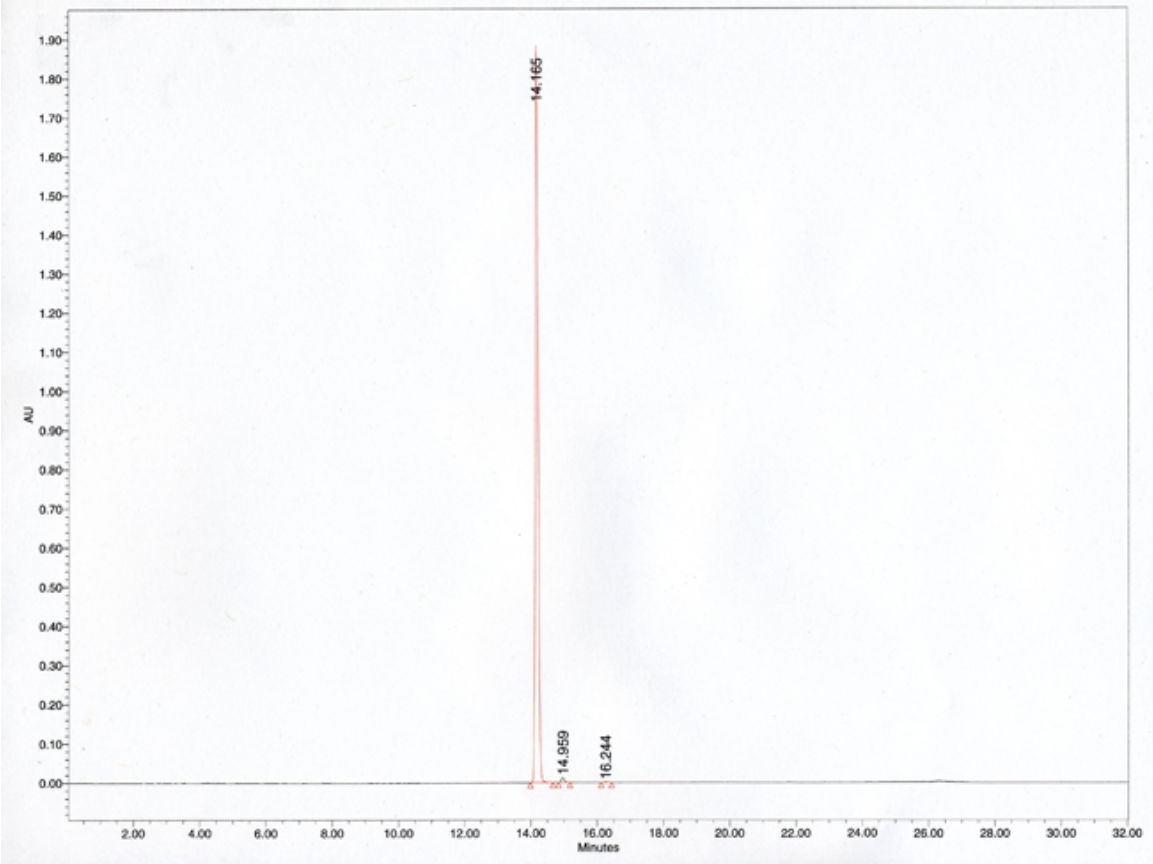
Peak	Retention Time (min)	Area (µV·s)	% Area	Height (µV)
1	18.232	17914865	99.86	1547767
2	23.798	2997	0.02	287
3	25.135	17964	0.10	1625
4	26.748	4629	0.03	396

PyoxDC*



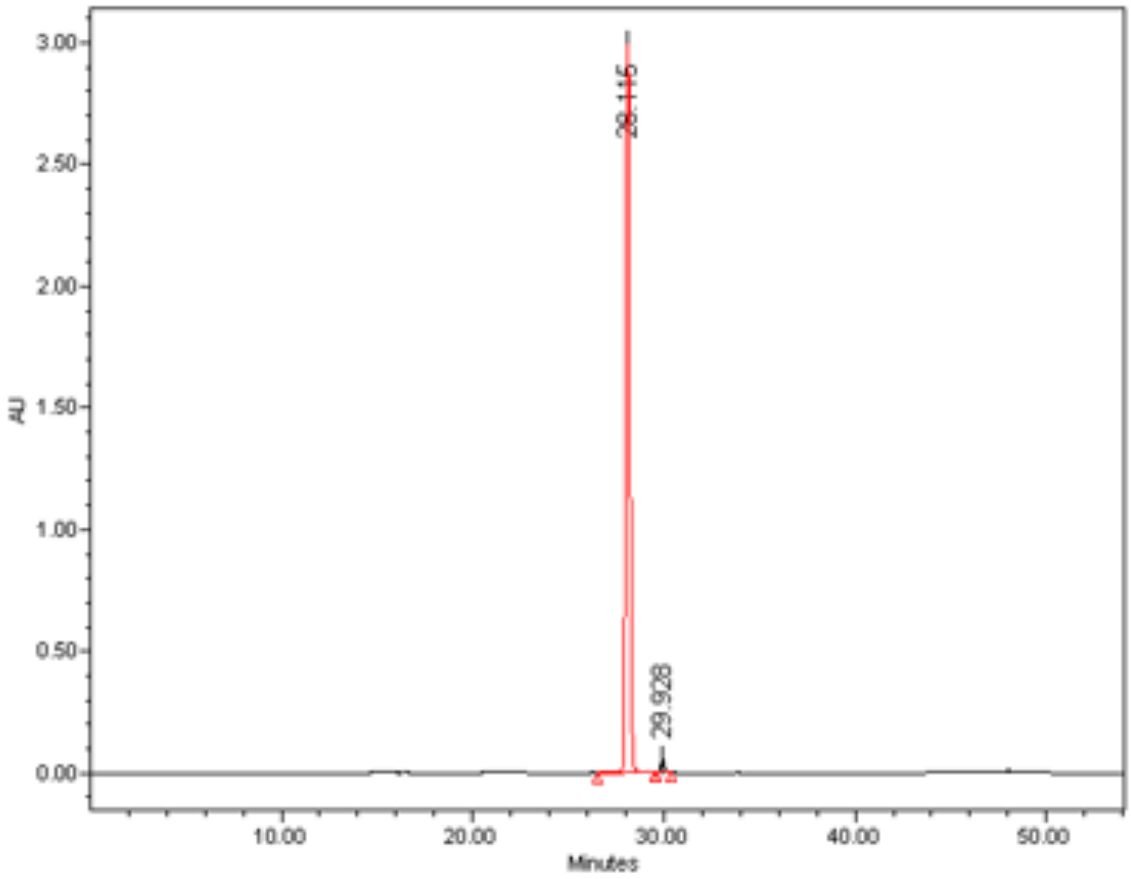
Peak	Retention Time (min)	Area (μV·s)	% Area	Height (μV)
1	13.266	1757	0.01	278
2	**13.814**	**11844742**	**99.89**	**1974317**
3	14.839	10808	0.09	1590

PyoxDC



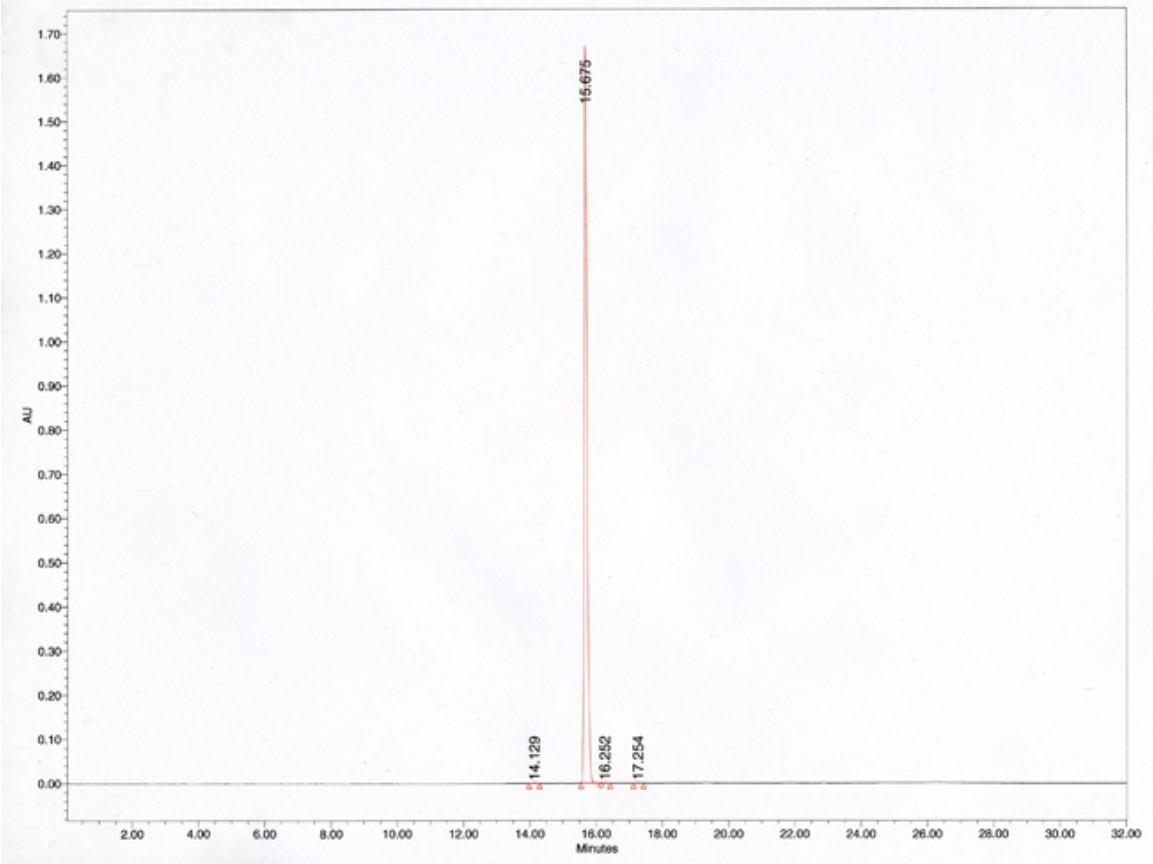
Peak	Retention Time (min)	Area (μV·s)	% Area	Height (μV)
1	**14.165**	**11469567**	**99.25**	**1882775**
2	14.959	79794	0.69	12911
3	16.244	7183	0.06	1120

Pythi



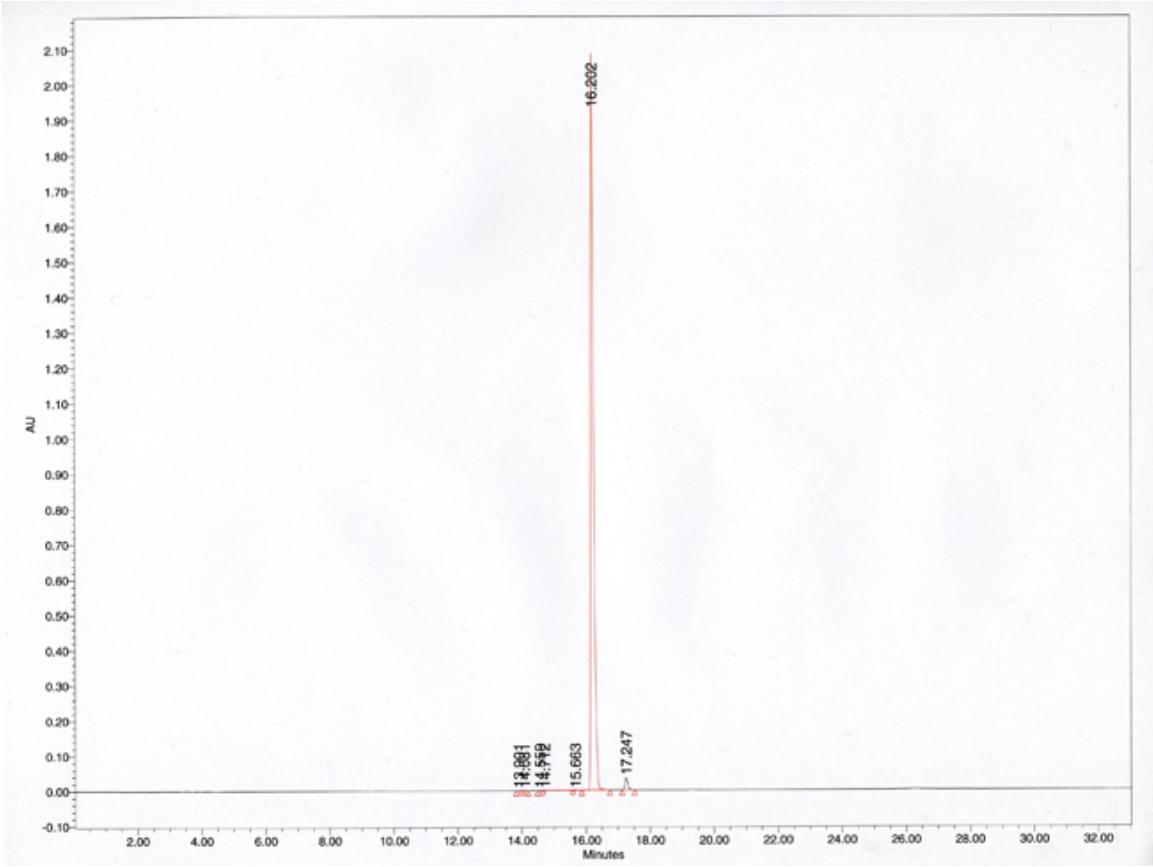
Peak	Retention Time (min)	Area (µV·s)	% Area	Height (µV)
1	**28.115**	**45177691**	**98.36**	**2990871**
2	29.928	753307	1.64	57240

PythiDC*



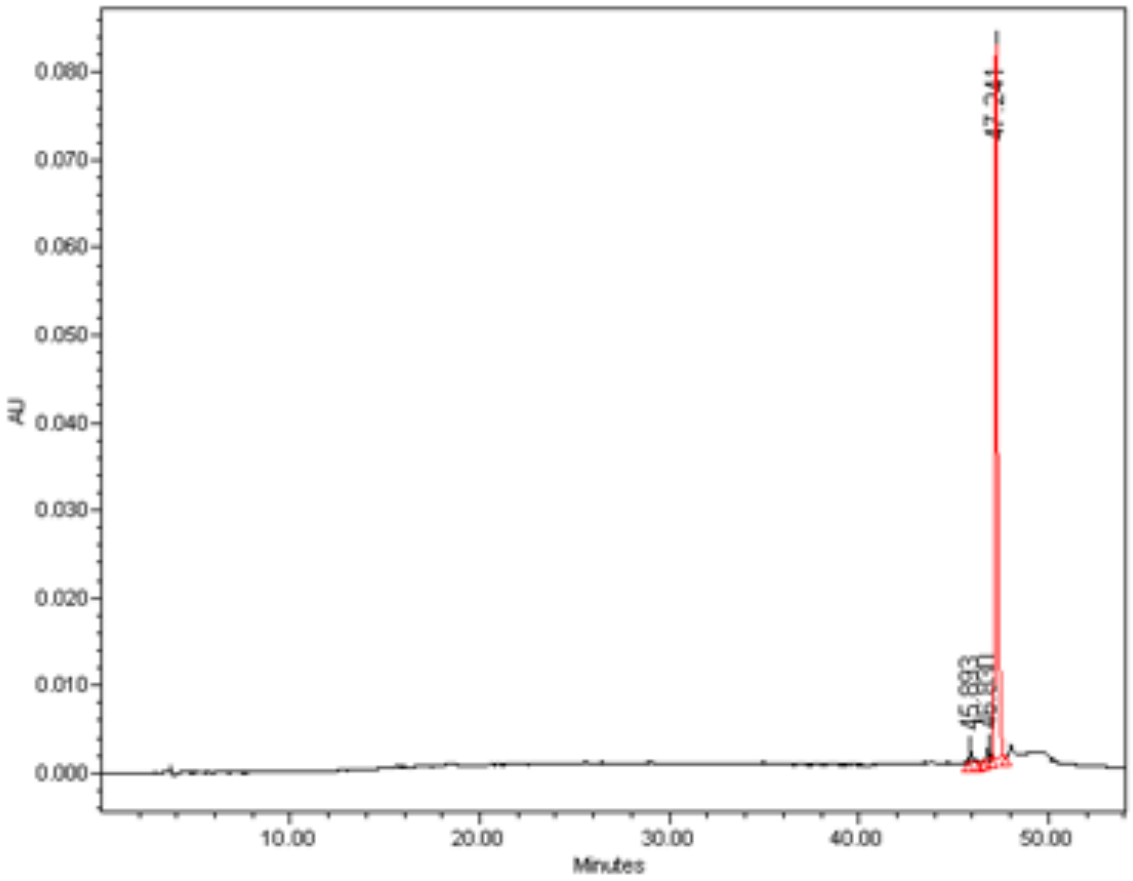
Peak	Retention Time (min)	Area (μV·s)	% Area	Height (μV)
1	14.129	3477	0.03	577
2	**15.675**	**9952801**	**99.87**	**1658904**
3	16.252	5422	0.05	669
4	17.254	4409	0.04	720

PythiDC



Peak	Retention Time (min)	Area (μV·s)	% Area	Height (μV)
1	13.901	2126	0.02	378
2	14.081	1788	0.01	325
3	14.559	1253	0.01	228
4	14.712	3091	0.02	361
5	15.663	1782	0.01	301
6	**16.202**	**13342462**	**98.22**	**2100713**
7	17.247	231648	1.71	35634

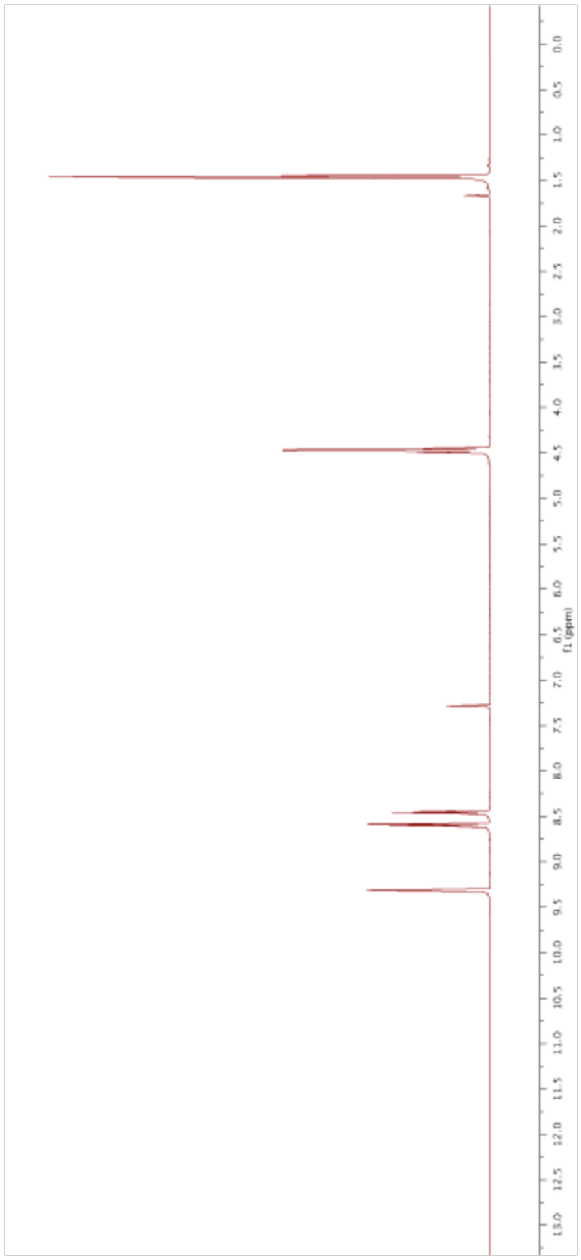
Diethyl PythiDC



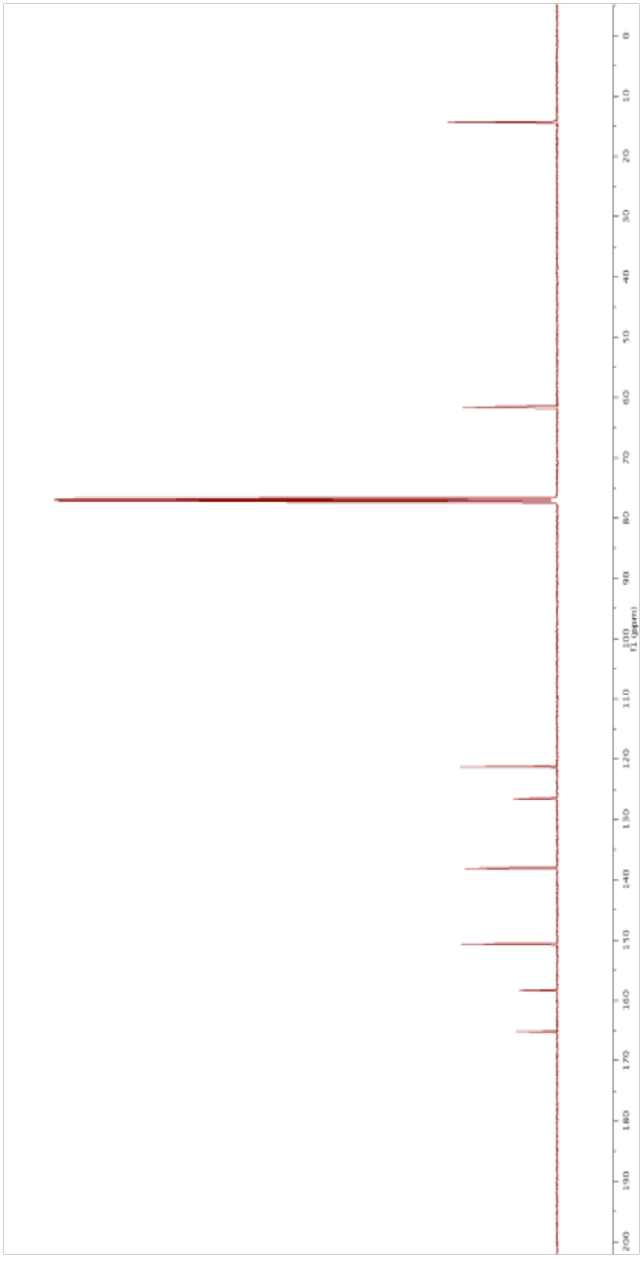
Peak	Retention Time (min)	Area (µV·s)	% Area	Height (µV)
1	45.893	12615	1.5	1276
2	46.830	11096	1.32	1220
3	**47.241**	**815559**	**97.17**	**81319**

3.5.20 NMR Spectra

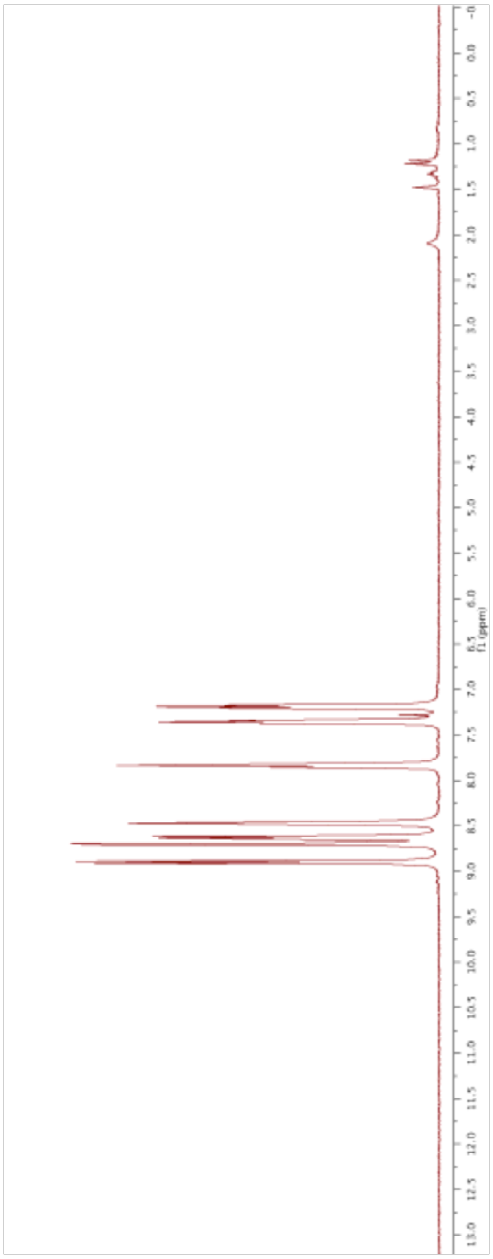
500 MHz ^{1}H NMR spectrum of Diethyl Bipy55′DC in $CDCl_3$



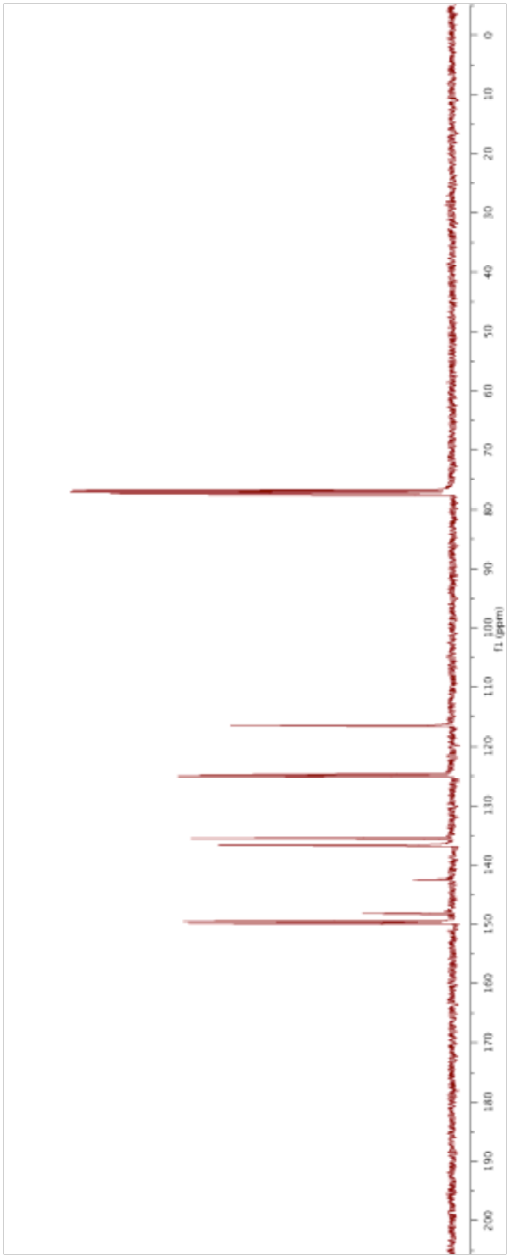
125 MHz ^{13}C NMR spectrum of Diethyl Bipy55′DC in $CDCl_3$



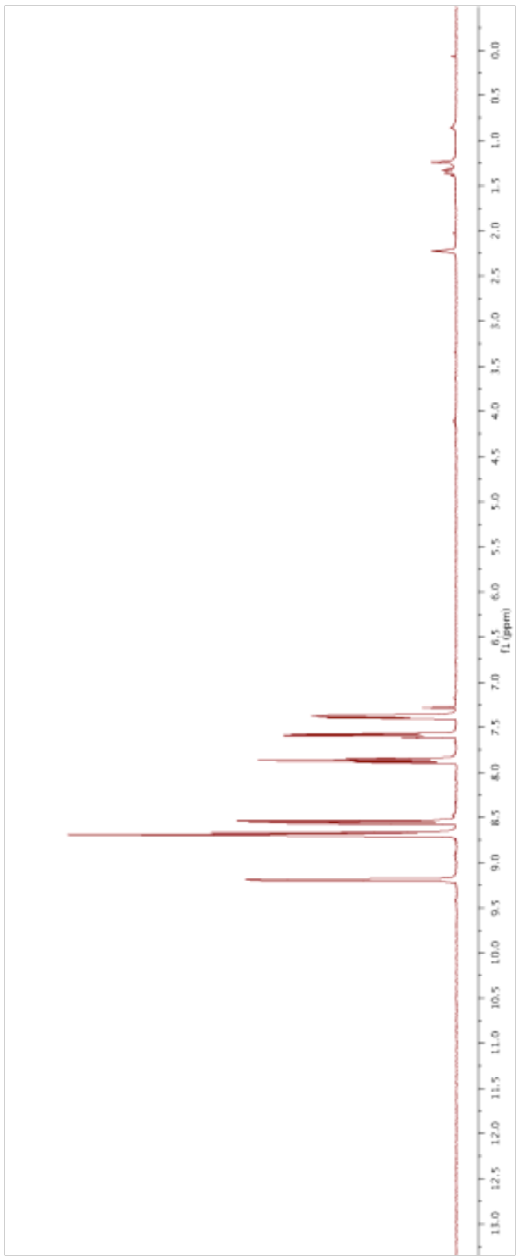
400 MHz ^{1}H NMR spectrum of 6-(Pyridin-2-yl)pyridazine-1-oxide in $CDCl_3$



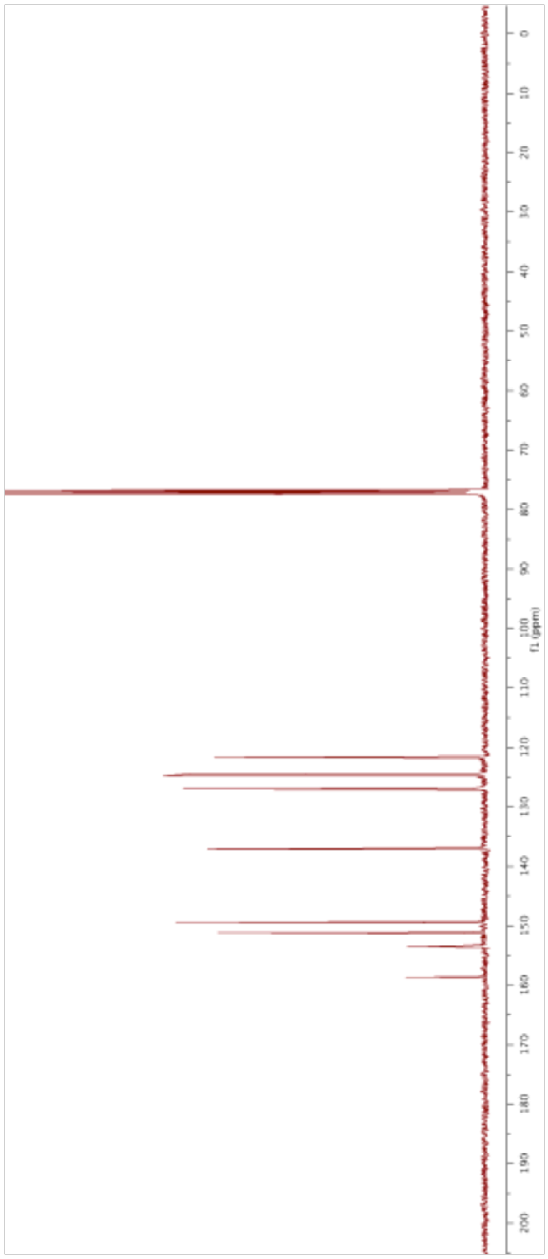
100 MHz ^{13}C NMR spectrum of 6-(Pyridin-2-yl)pyridazine-1-oxide in $CDCl_3$



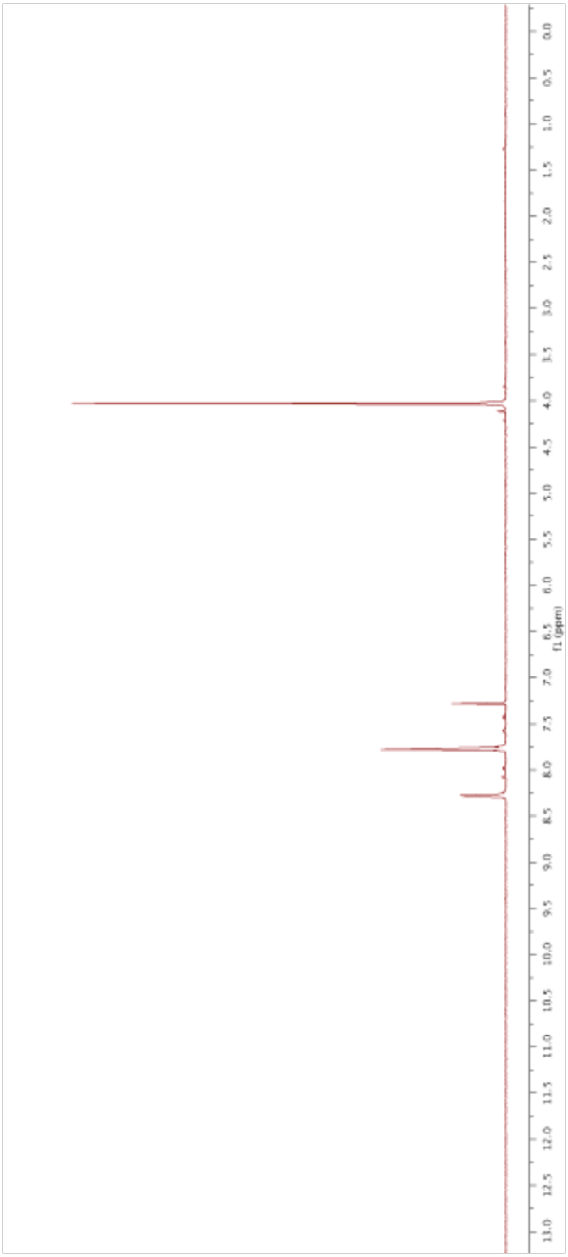
400 MHz ^{1}H NMR spectrum of 3-(Pyridin-2-yl)pyridazine (pypyrid) in $CDCl_3$



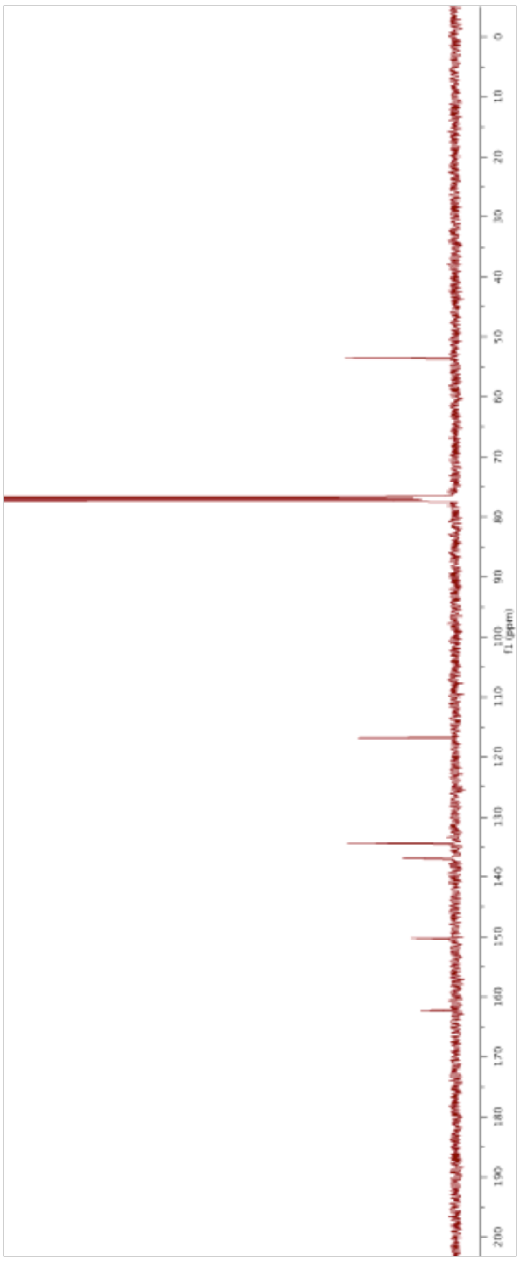
100 MHz ^{13}C NMR spectrum of 3-(Pyridin-2-yl)pyridazine (pypyrid) in $CDCl_3$



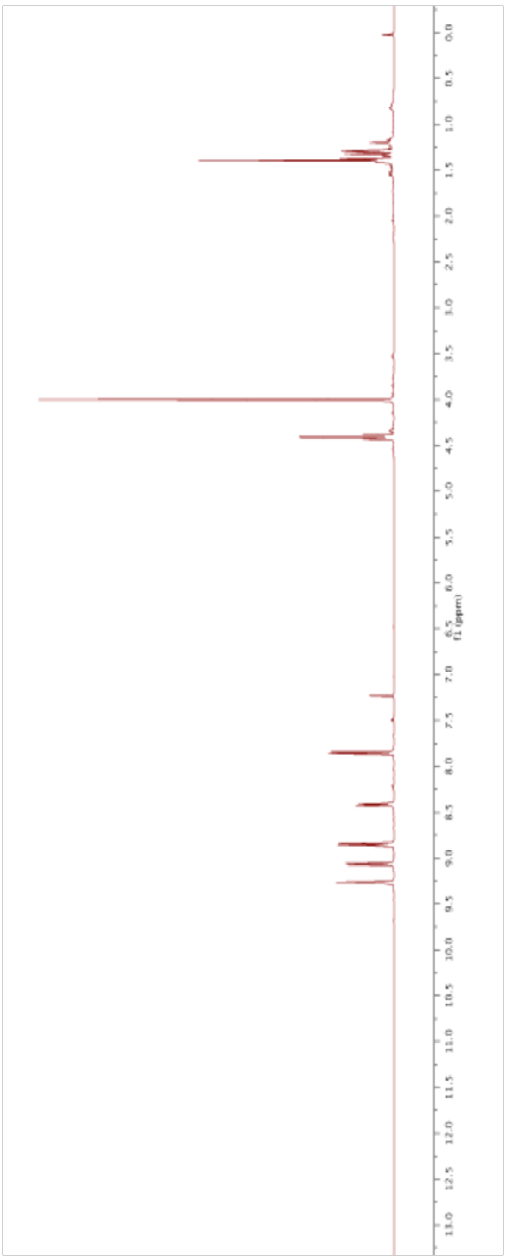
400 MHz ^{1}H NMR spectrum of Methyl Pyridazine-3-carboxylate-1-oxide in $CDCl_3$



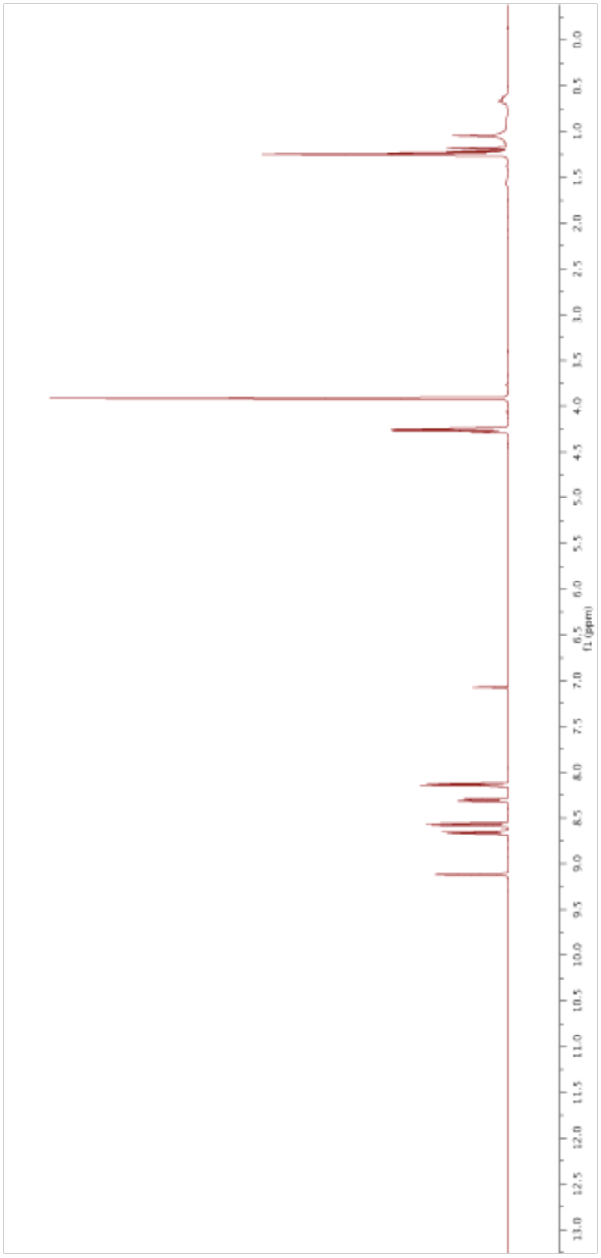
100 MHz ^{13}C NMR spectrum of Methyl Pyridazine-3-carboxylate-1-oxide in $CDCl_3$



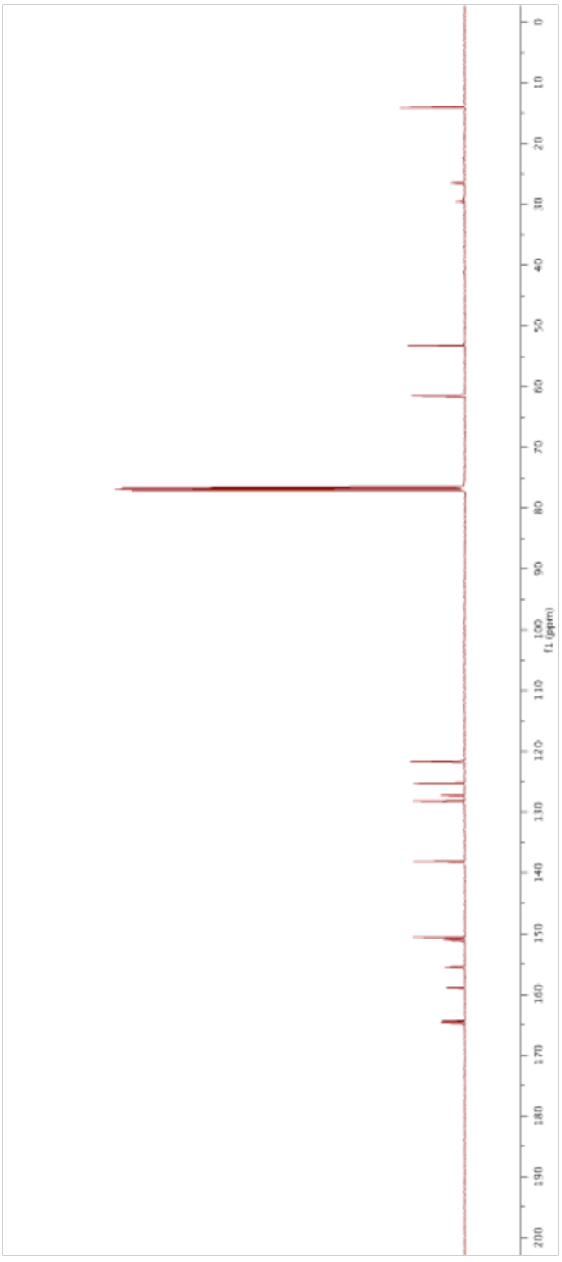
500 MHz ^{1}H NMR spectrum of Crude Methyl 6-(Ethoxycarbonylpyridin-2-yl)pyridazine-3-carboxylate-1-oxide in $CDCl_3$



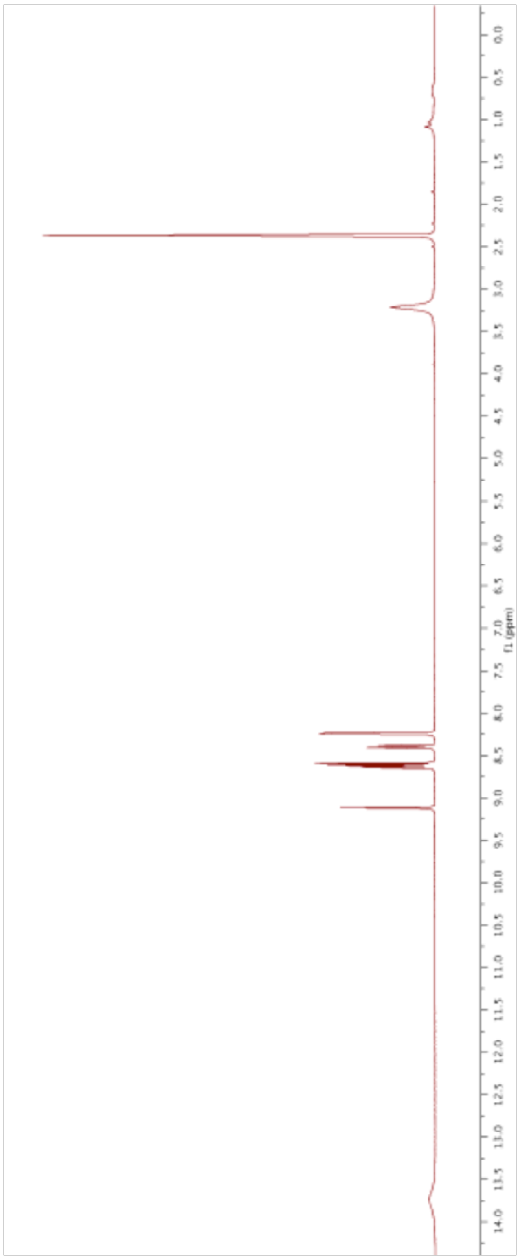
500 MHz ^{1}H NMR spectrum of Methyl 3-(5-Ethoxycarbonylpyridin-2-yl)pyridazine-6-carboxylate (ethylmethyl pypyridDC) in $CDCl_3$



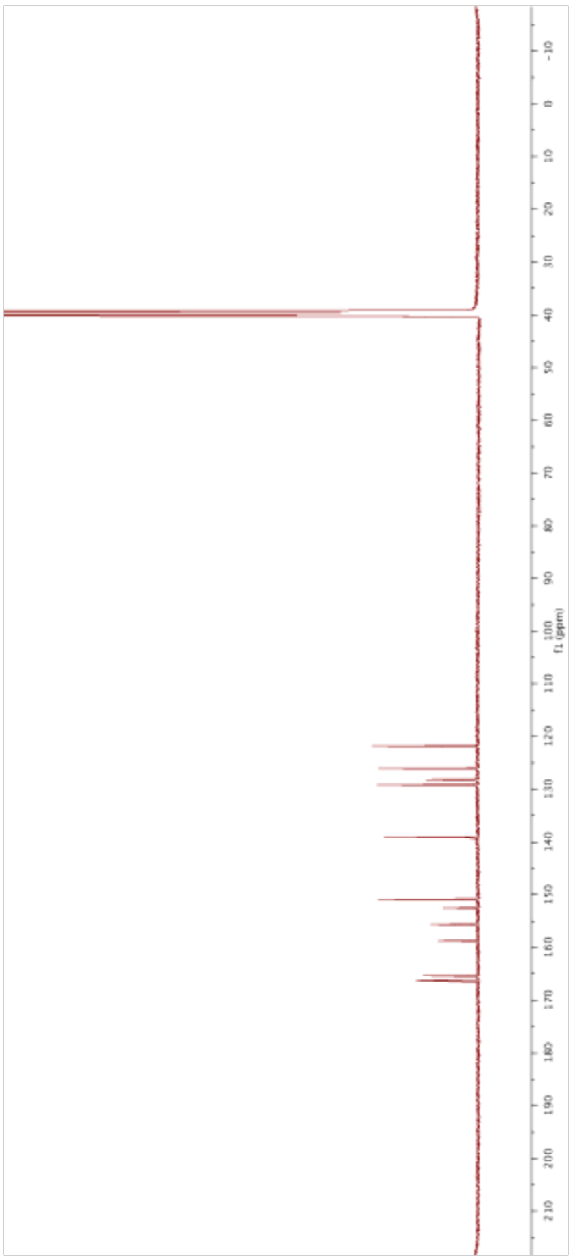
125 MHz ^{13}C NMR spectrum of Methyl 3-(5-Ethoxycarbonylpyridin-2-yl)pyridazine-6-carboxylate (ethylmethyl pypyridDC) in $CDCl_3$



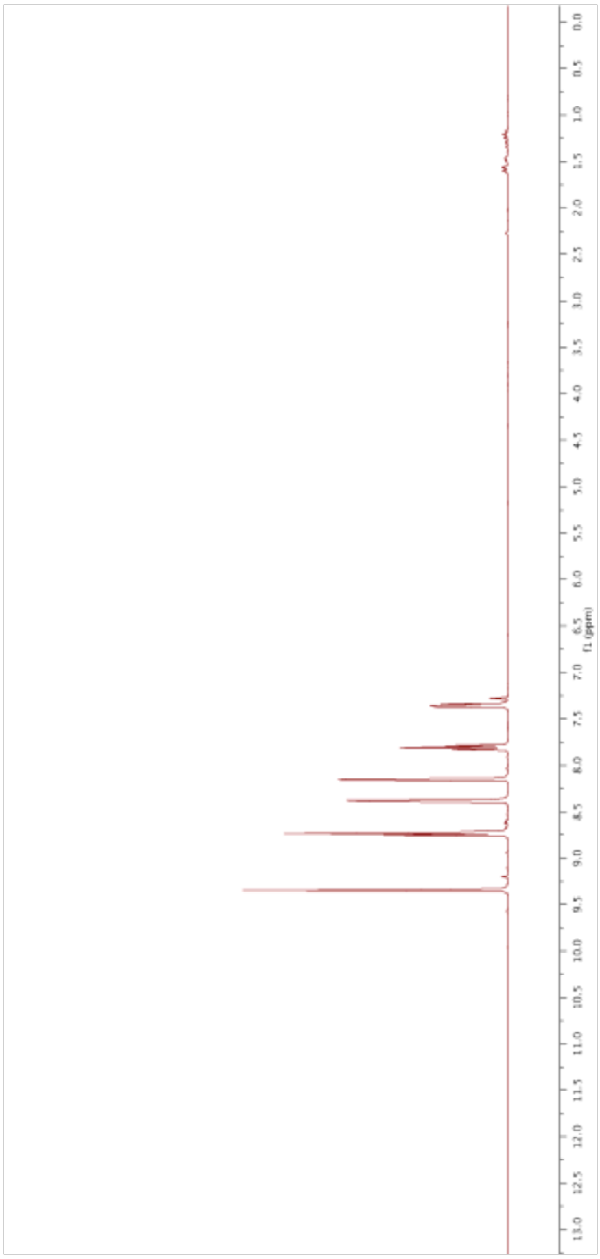
500 MHz ^{1}H NMR spectrum of 3-(5-Carboxypyridin-2-yl)pyridazine-6-carboxylic Acid (pypyridDC) in DMSO-d_6



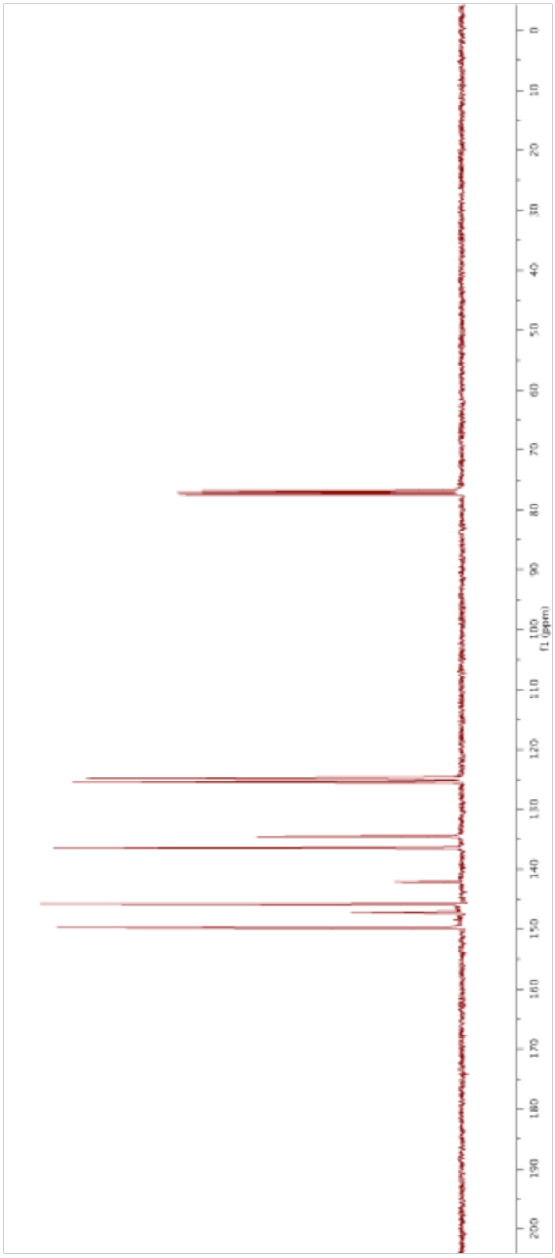
125 MHz ^{13}C NMR spectrum of 3-(5-Carboxypyridin-2-yl)pyridazine-6-carboxylic Acid (pypyridDC) in DMSO-d_6



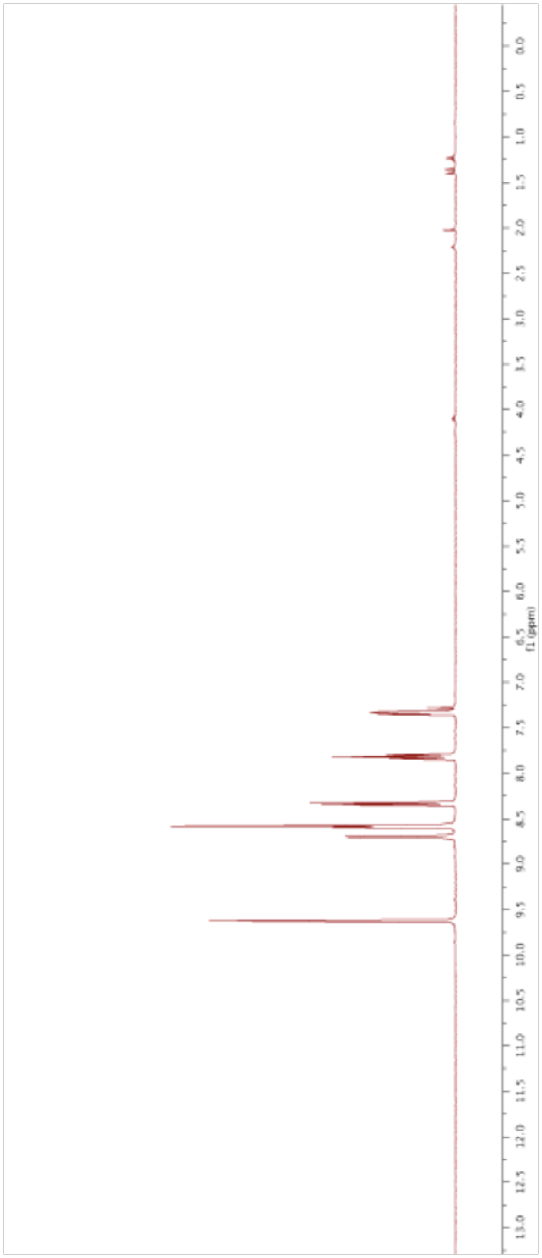
400 MHz ^{1}H NMR spectrum of 2-(Pyridin-2-yl)pyrazine-1-oxide in $CDCl_3$



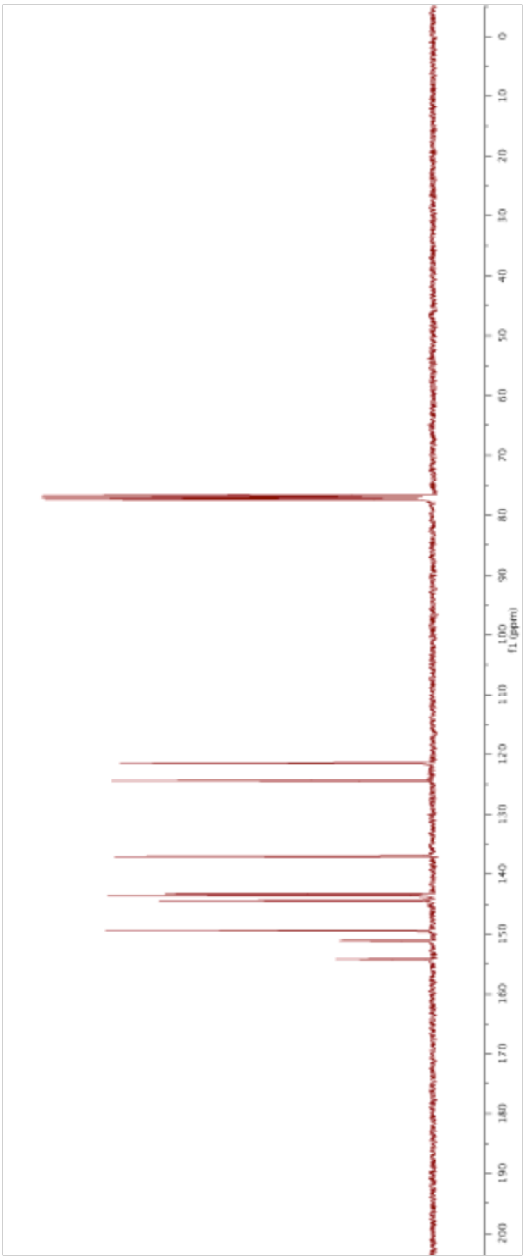
100 MHz ^{13}C NMR spectrum of 2-(Pyridin-2-yl)pyrazine-1-oxide in $CDCl_3$



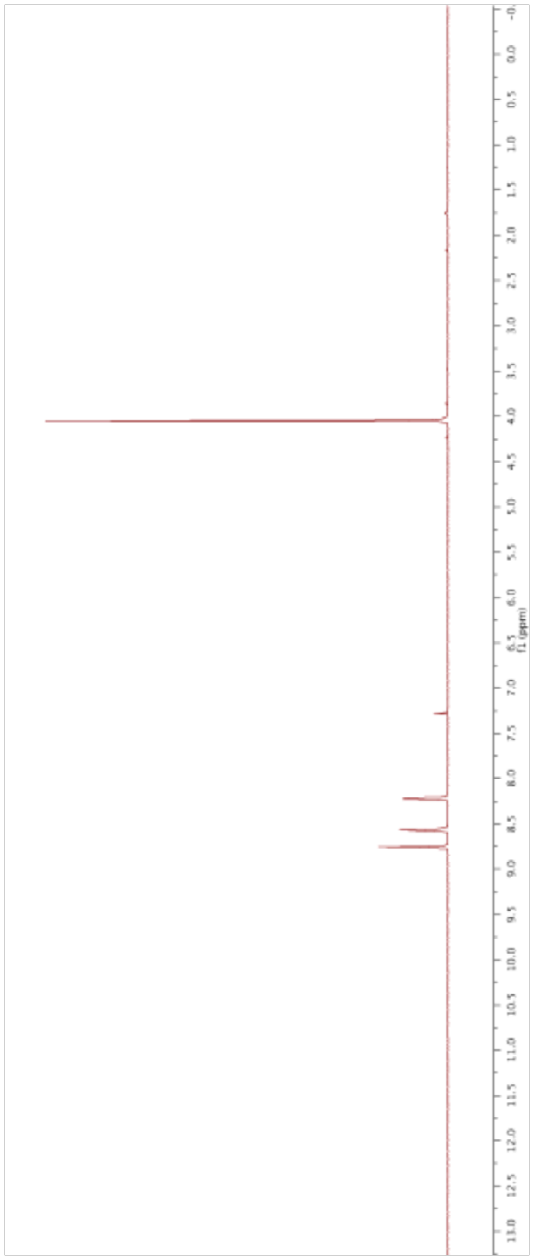
400 MHz ^{1}H NMR spectrum of 2-(Pyridin-2-yl)pyrazine (pypyraz) in $CDCl_3$



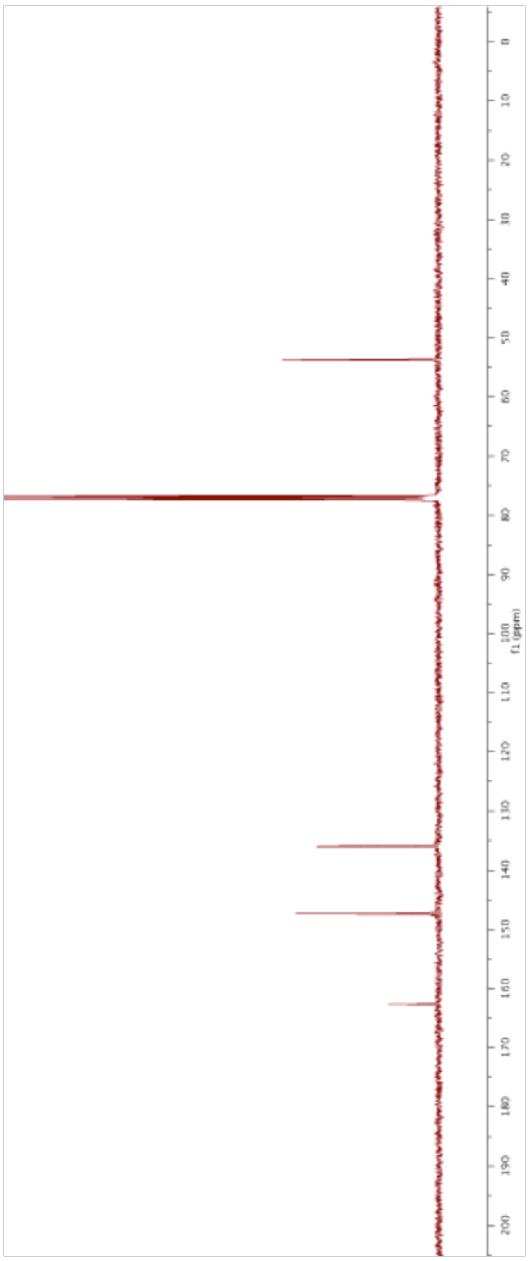
100 MHz ^{13}C NMR spectrum of 2-(Pyridin-2-yl)pyrazine (pypyraz) in $CDCl_3$



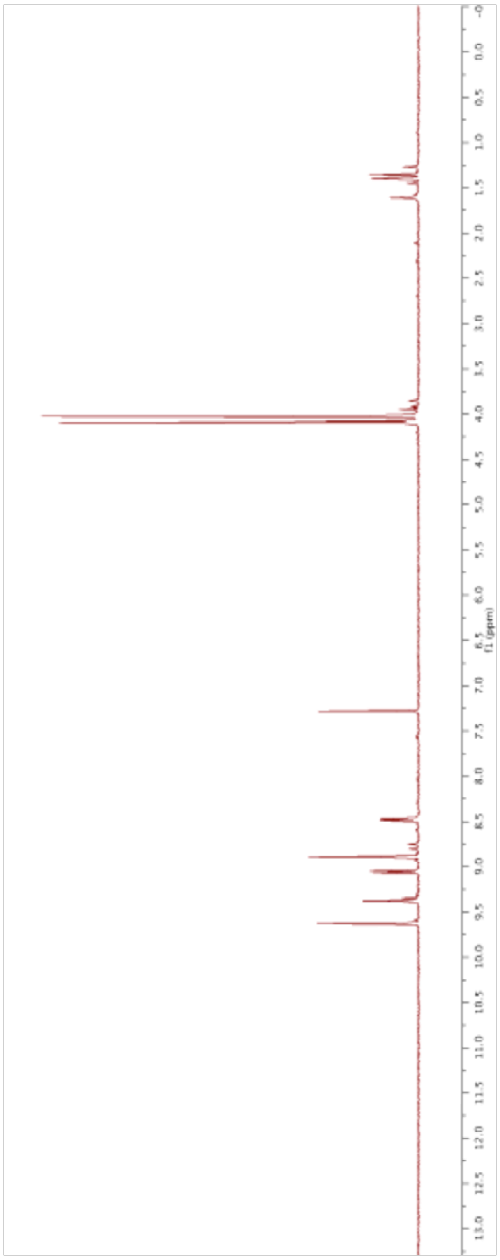
400 MHz ^{1}H NMR spectrum of Methyl Pyrazine-3-carboxylate-1-oxide in $CDCl_3$



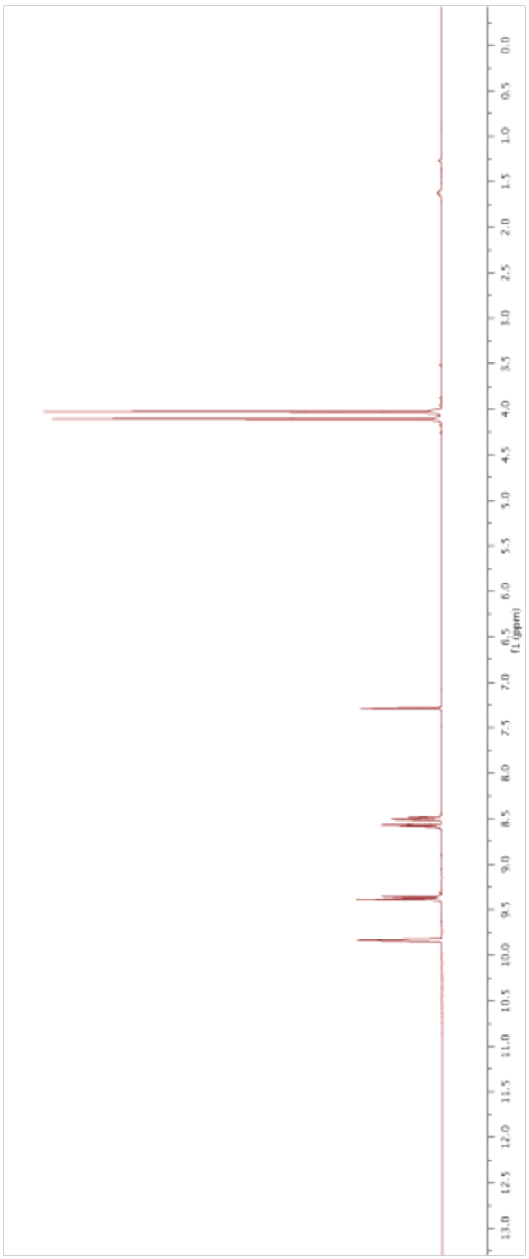
100 MHz ^{13}C NMR spectrum of Methyl Pyrazine-3-carboxylate-1-oxide in $CDCl_3$



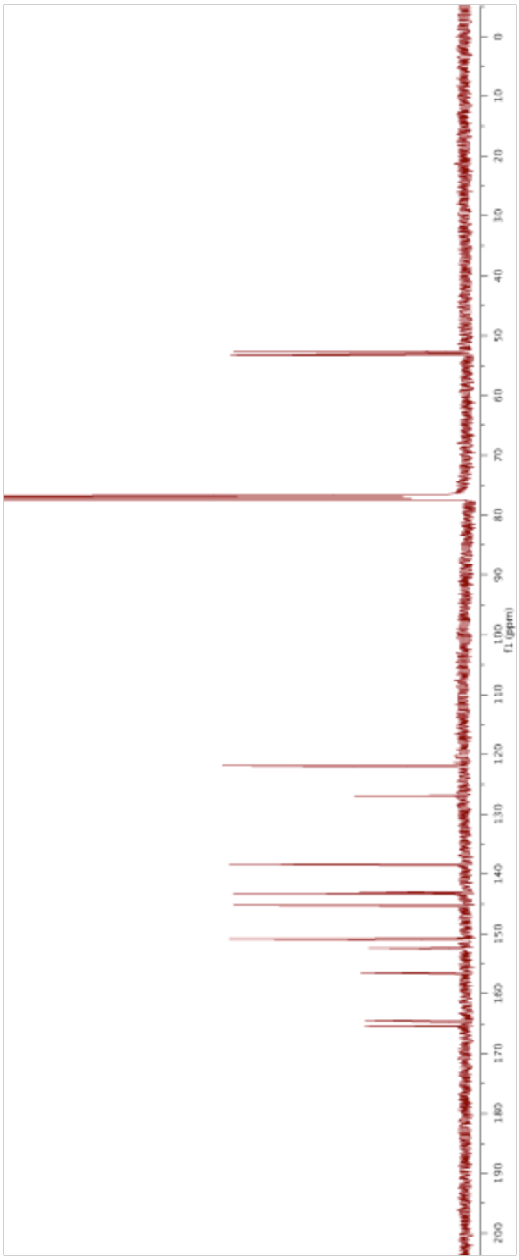
400 MHz ^{1}H NMR spectrum of Crude Methyl 2-(5-Methoxycarbonylpyridin-2-yl)pyrazine-5-carboxylate in $CDCl_3$



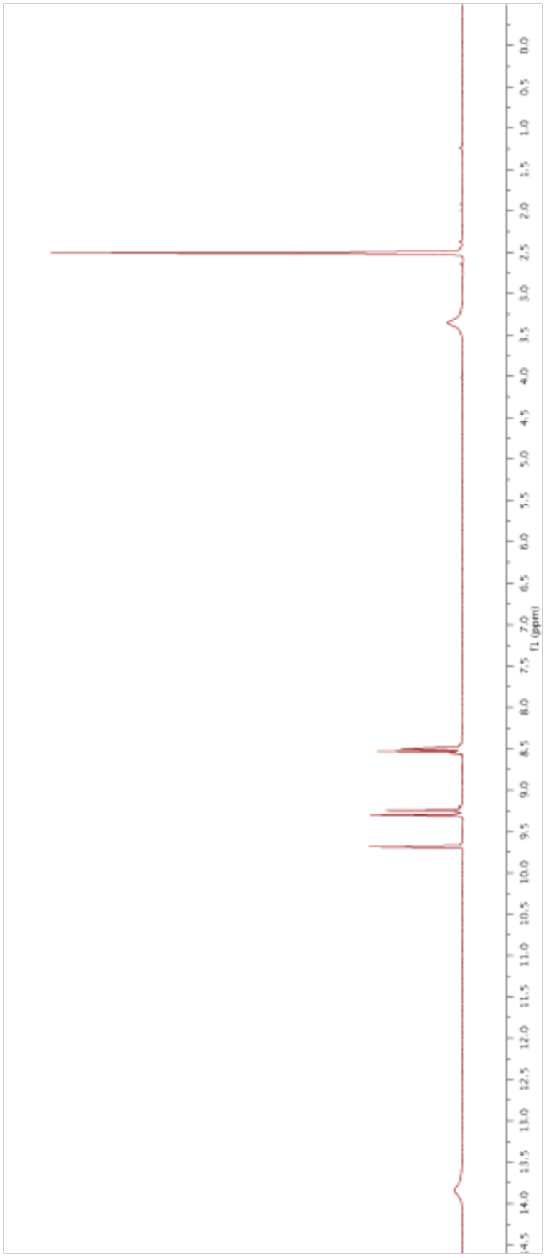
500 MHz ^{1}H NMR spectrum of Methyl 2-(5-Methoxycarbonylpyridin-2-yl)pyrazine-5-carboxylate (dimethyl pypyrazDC) in $CDCl_3$



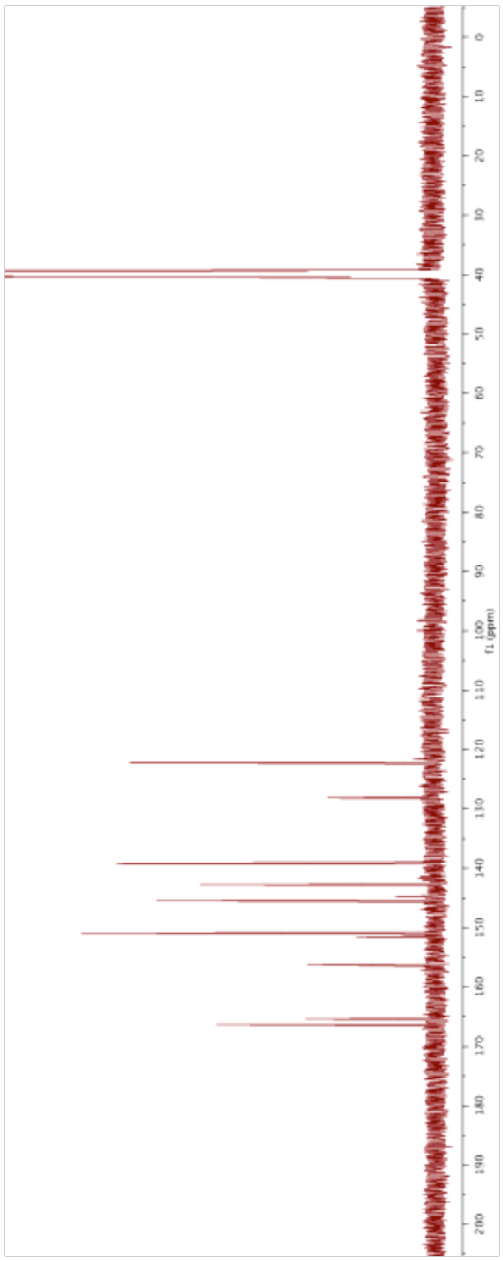
125 MHz ^{13}C NMR spectrum of Methyl 2-(5-Methoxycarbonylpyridin-2-yl)pyrazine-5-carboxylate (dimethyl pypyrazDC) in $CDCl_3$



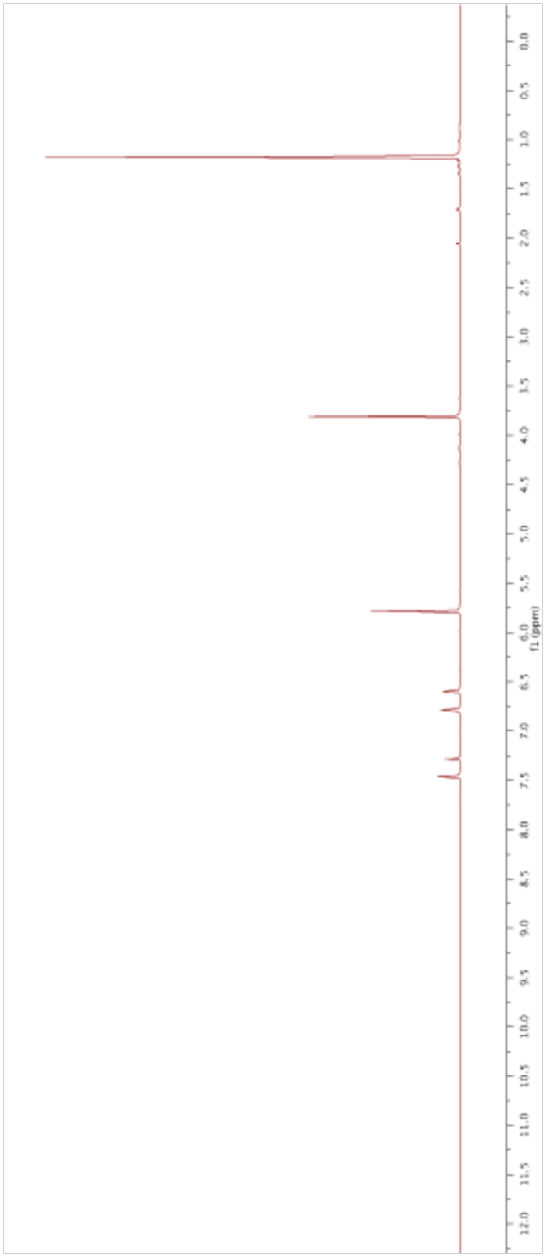
500 MHz ^{1}H NMR spectrum of 2-(5-Carboxypyridin-2-yl)pyrazine-5-carboxylic Acid (pypyrazDC) in DMSO-d_6



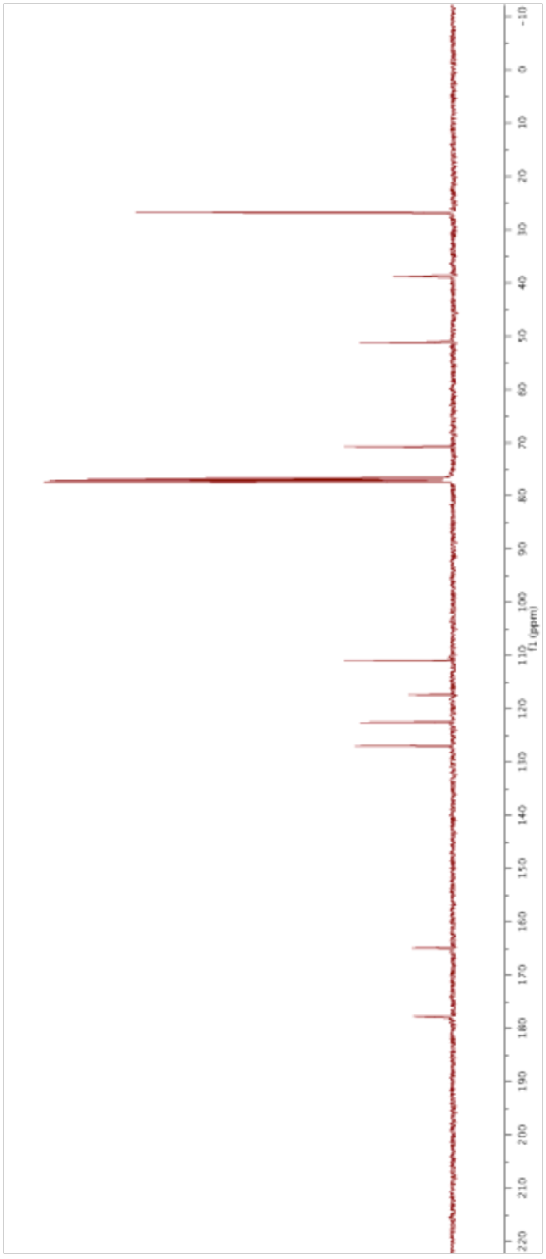
125 MHz ^{13C} NMR spectrum of 2-(5-Carboxypyridin-2-yl)pyrazine-5-carboxylic Acid (pypyrazDC) in DMSO-d_6



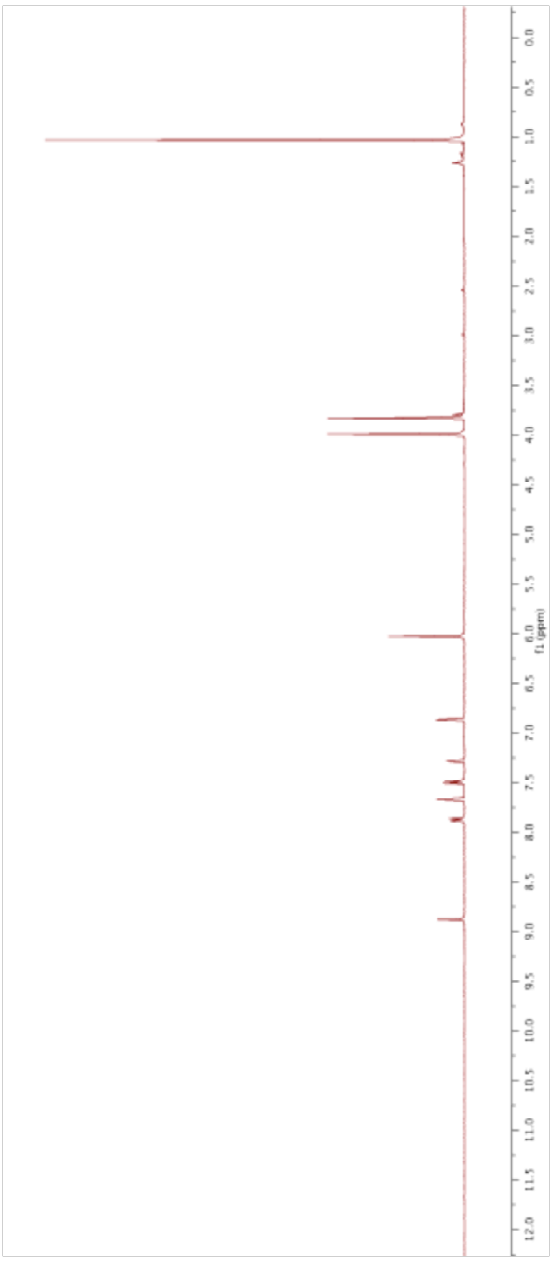
400 MHz ^{1}H NMR spectrum of Methyl 1-Pivaloyloxymethyl-1*H*-pyrrole-3-carboxylate in $CDCl_3$



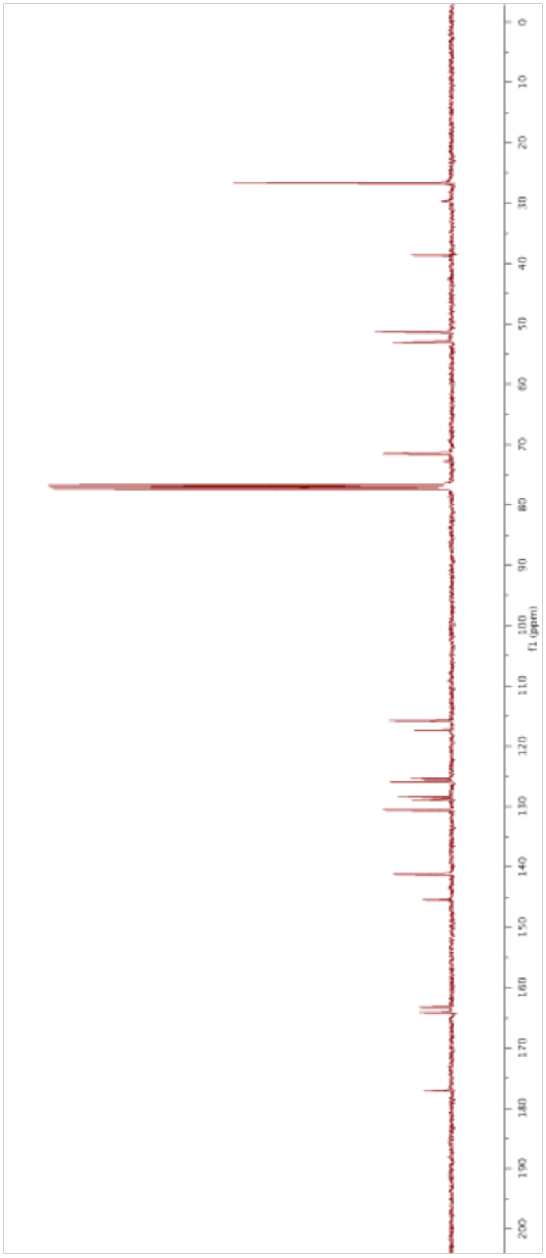
100 MHz ^{13}C NMR spectrum of Methyl 1-Pivaloyloxymethyl-1*H*-pyrrole-3-carboxylate in $CDCl_3$



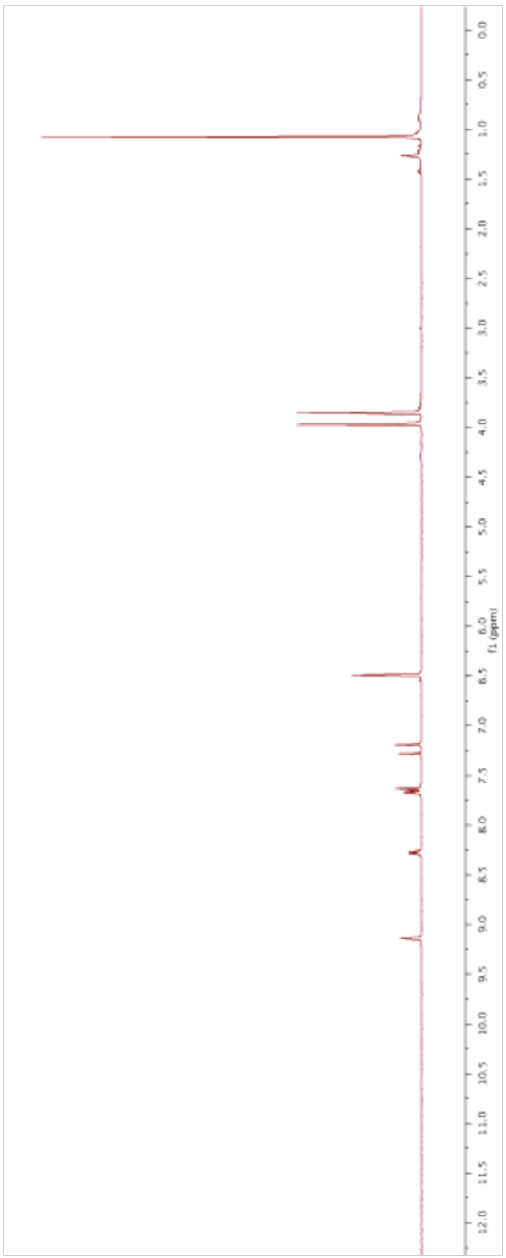
400 MHz ^{1}H NMR spectrum of Methyl 2-(4-Methoxycarbonyl-1-pivaloyloxymethyl-1*H*-pyrrol-2-yl)pyridine-5-carboxylate-1-oxide in $CDCl_3$



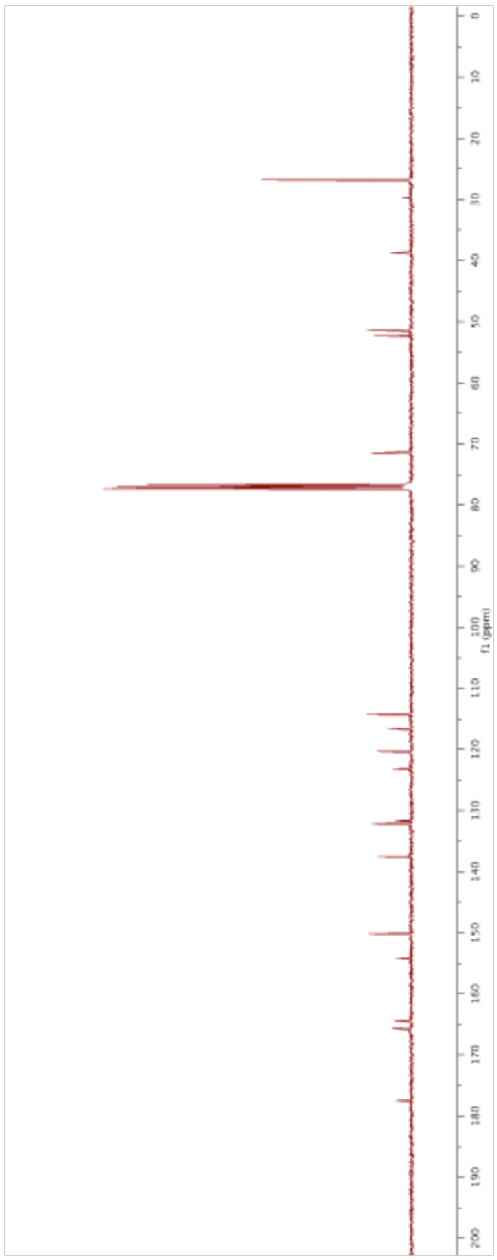
100 MHz ^{13}C NMR spectrum of Methyl 2-(4-Methoxycarbonyl-1-pivaloyloxymethyl-1*H*-pyrrol-2-yl)pyridine-5-carboxylate-1-oxide in $CDCl_3$



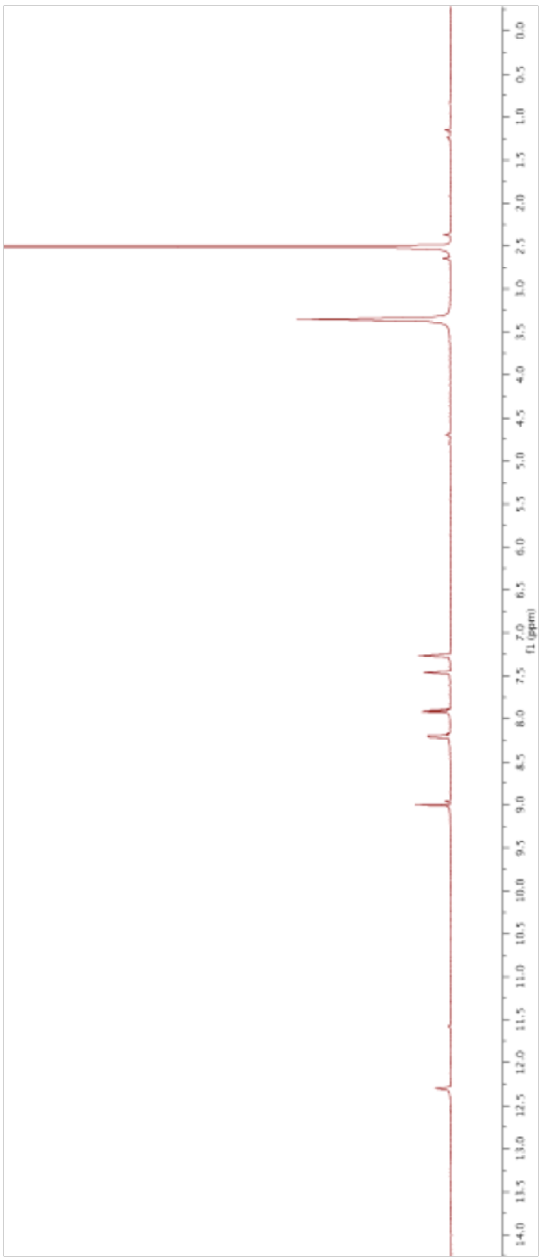
400 MHz ^{1}H NMR spectrum of Methyl 2-(4-Methoxycarbonyl-1-pivaloyloxymethyl-1*H*-pyrrol-2-yl)pyridine-5-carboxylate in $CDCl_3$



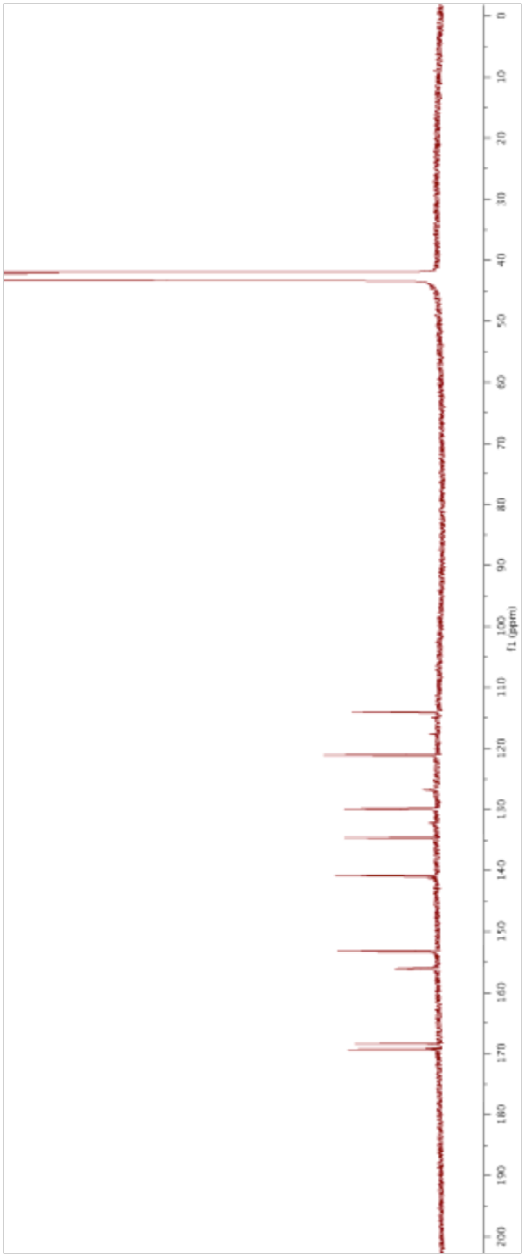
100 MHz ^{13}C NMR spectrum of Methyl 2-(4-Methoxycarbonyl-1-pivaloyloxymethyl-1*H*-pyrrol-2-yl)pyridine-5-carboxylate in $CDCl_3$



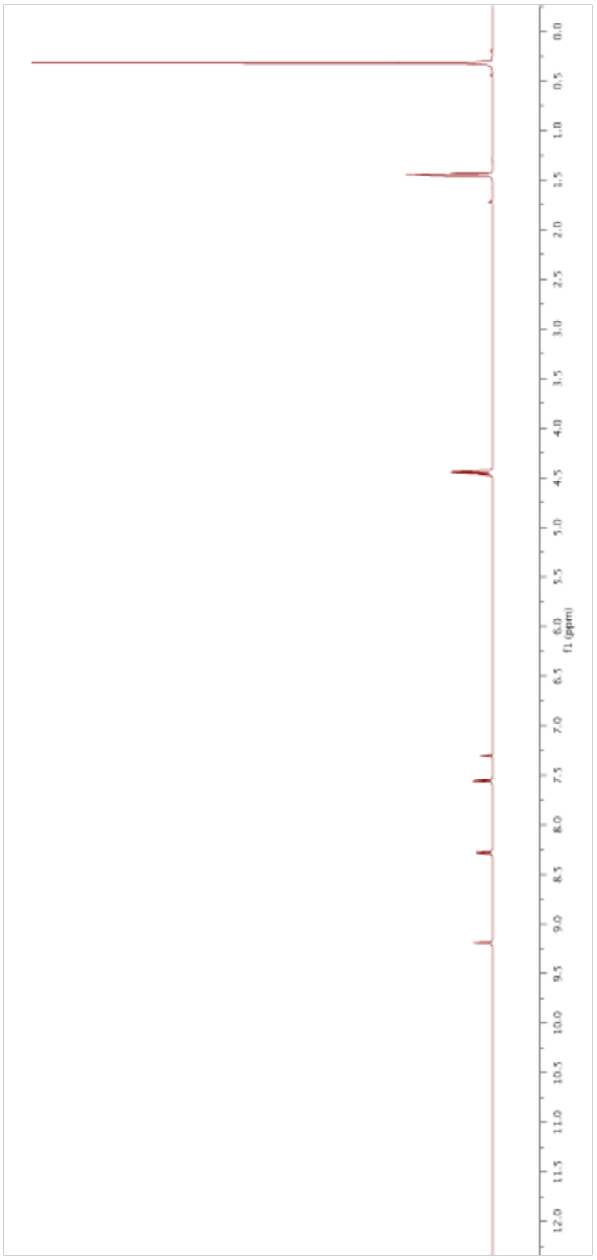
500 MHz ^{1}H NMR spectrum of 2-(4-Carboxy-1*H*-pyrrol-2-yl)pyridine-5-carboxylic Acid (pypyrroleDC) in DMSO-d_6



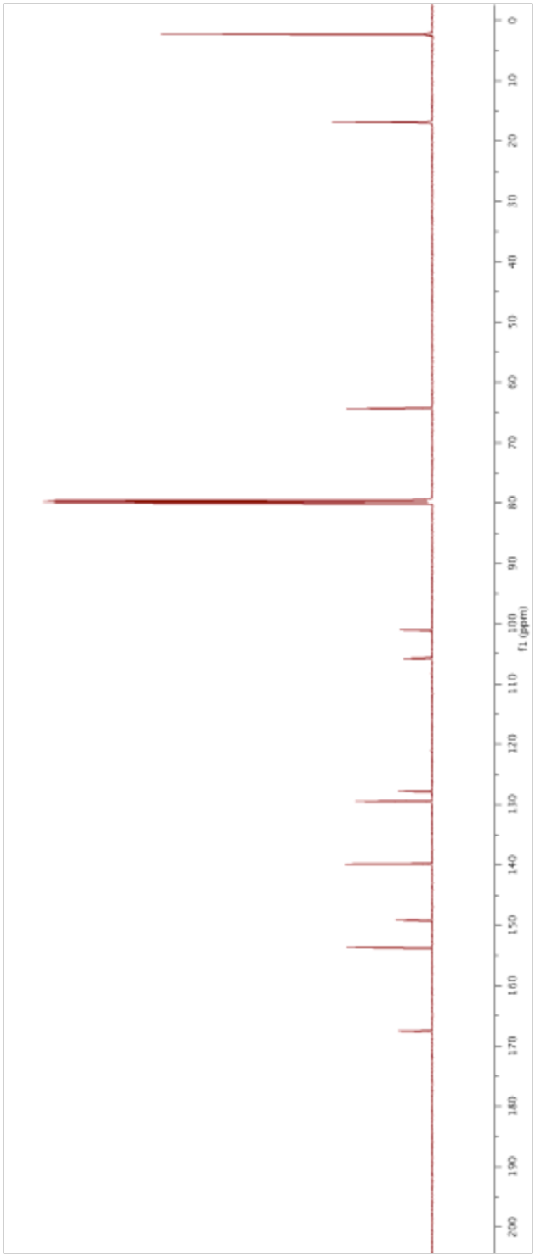
125 MHz ^{13}C NMR spectrum of 2-(4-Carboxy-1*H*-pyrrol-2-yl)pyridine-5-carboxylic Acid (pypyrroleDC) in DMSO-d_6



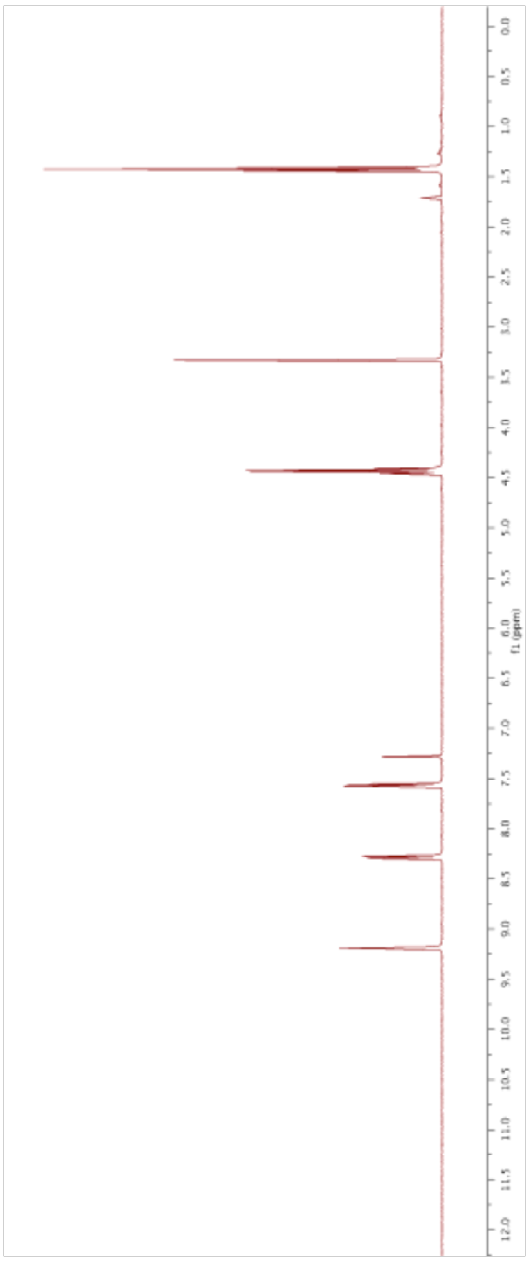
500 MHz ^{1}H NMR spectrum of Ethyl 2-(2-Trimethylsilylethynyl)pyridine-5-carboxylate in $CDCl_3$



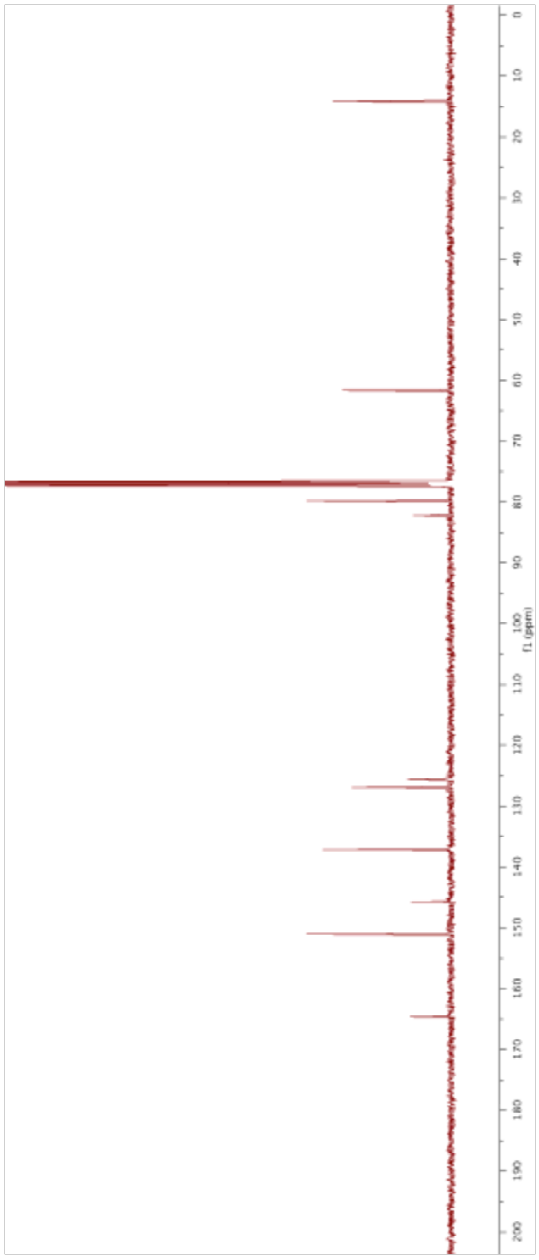
125 MHz ^{13}C NMR spectrum of Ethyl 2-(2-Trimethylsilylethynyl)pyridine-5-carboxylate in $CDCl_3$



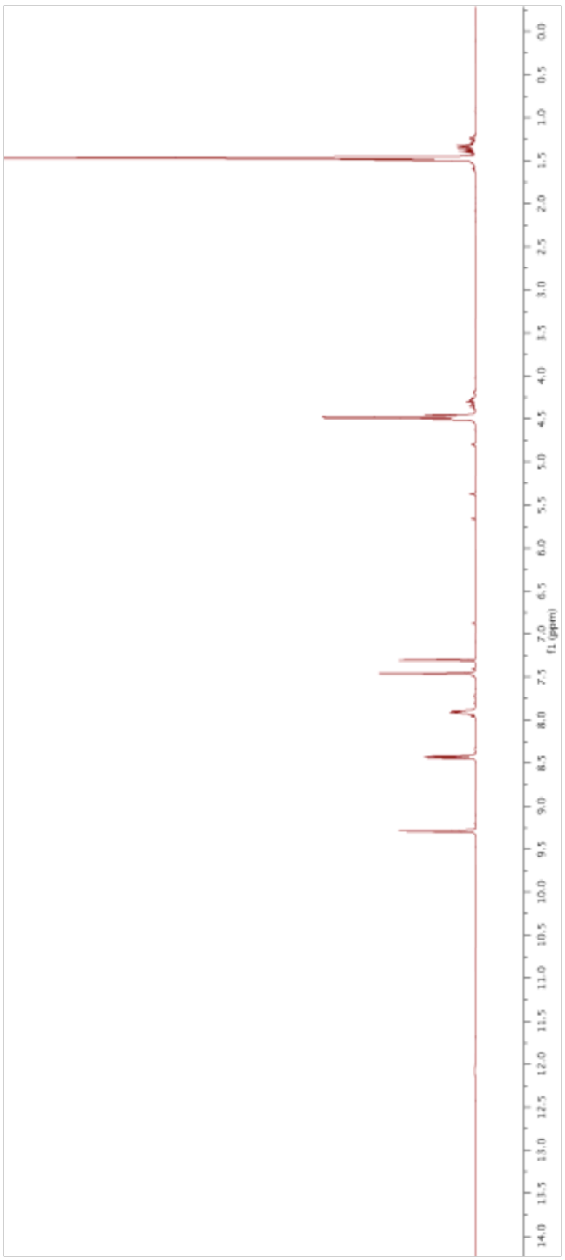
400 MHz ^{1}H NMR spectrum of Ethyl 2-Ethynylpyridine-5-carboxylate in $CDCl_3$



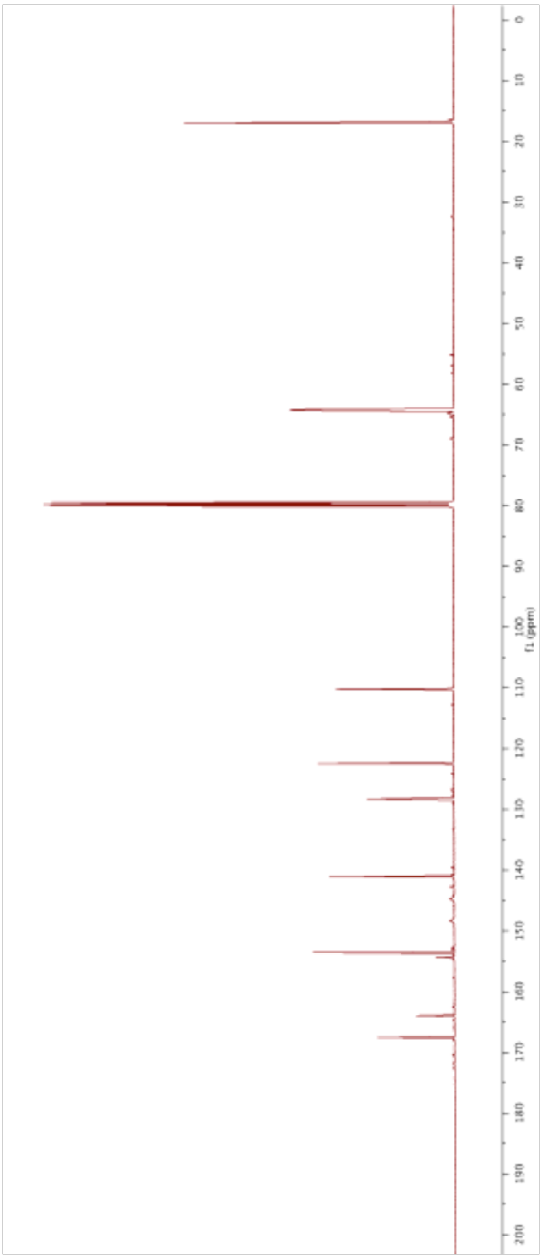
100 MHz ^{13}C NMR spectrum of Ethyl 2-Ethynylpyridine-5-carboxylate in $CDCl_3$



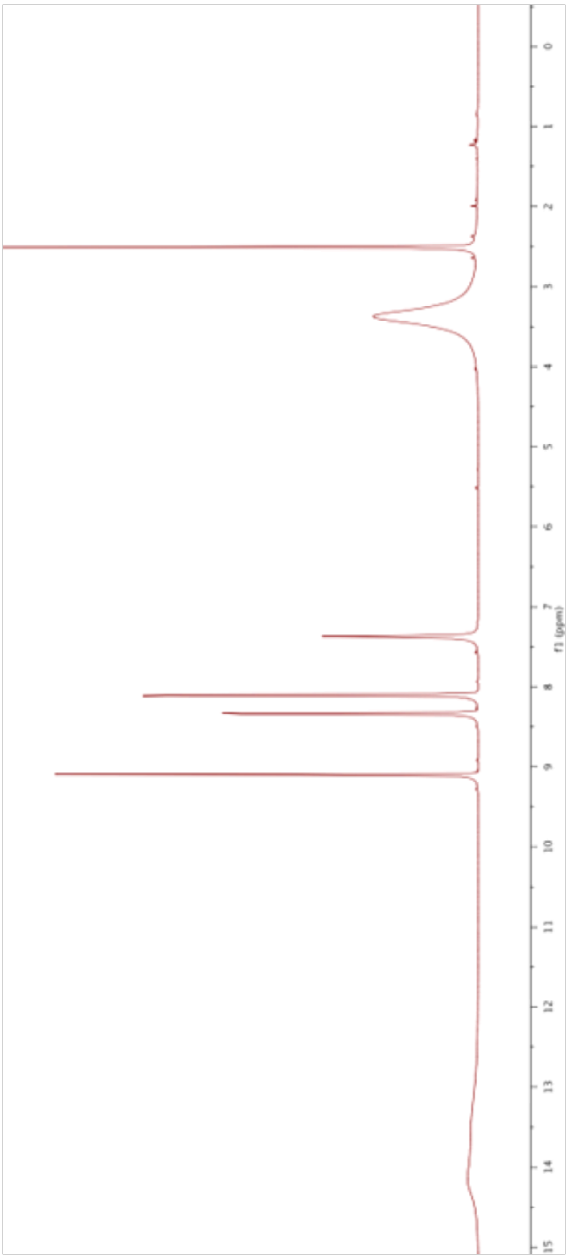
500 MHz ^{1}H NMR spectrum of Ethyl 2-(5-Ethoxycarbonyl-1*H*-pyrazol-3-yl)pyridine-5-carboxylate (diethyl pypyrDC) in $CDCl_3$



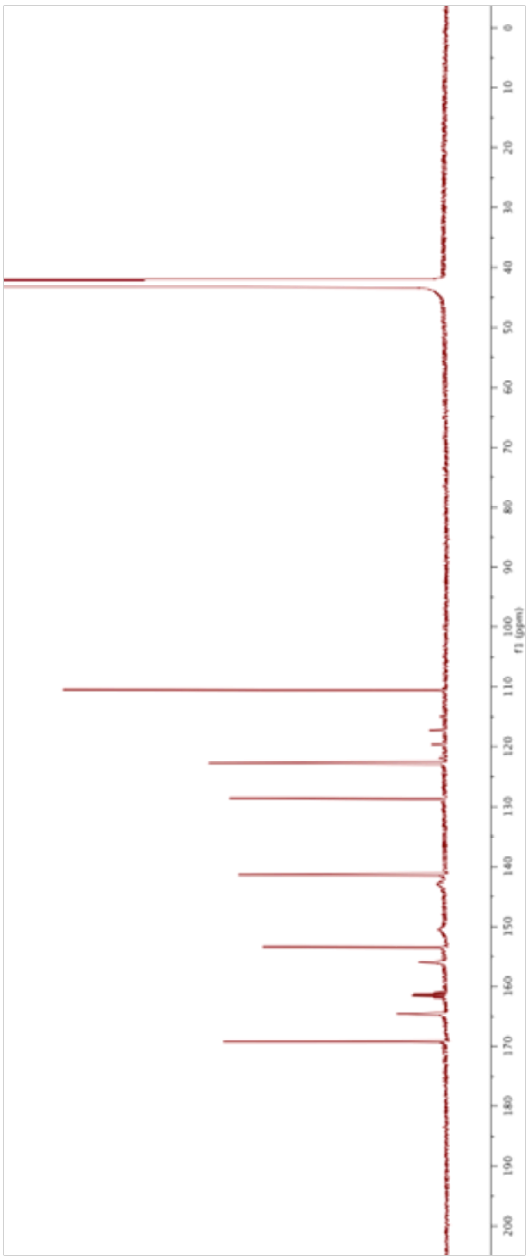
125 MHz ^{13}C NMR spectrum of Ethyl 2-(5-Ethoxycarbonyl-1*H*-pyrazol-3-yl)pyridine-5-carboxylate (diethyl pypyrDC) in $CDCl_3$



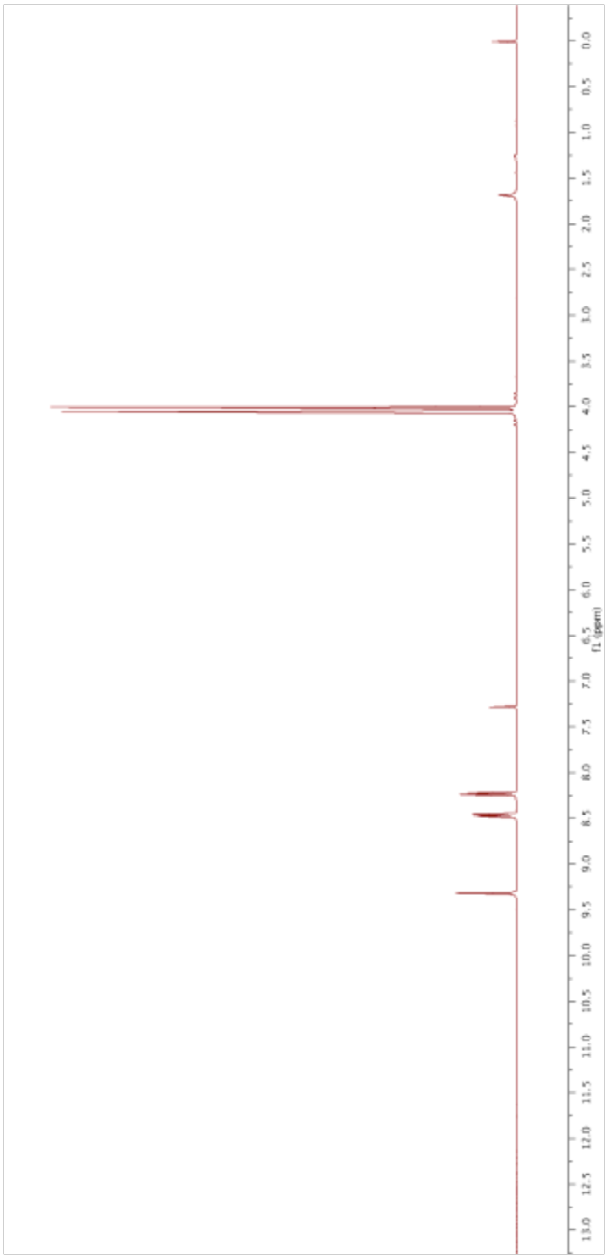
500 MHz ^{1}H NMR spectrum of 2-(3-Carboxy-1*H*-pyrazol-5-yl)pyridine-5-carboxylic Acid (pypyrDC) in DMSO-d_6



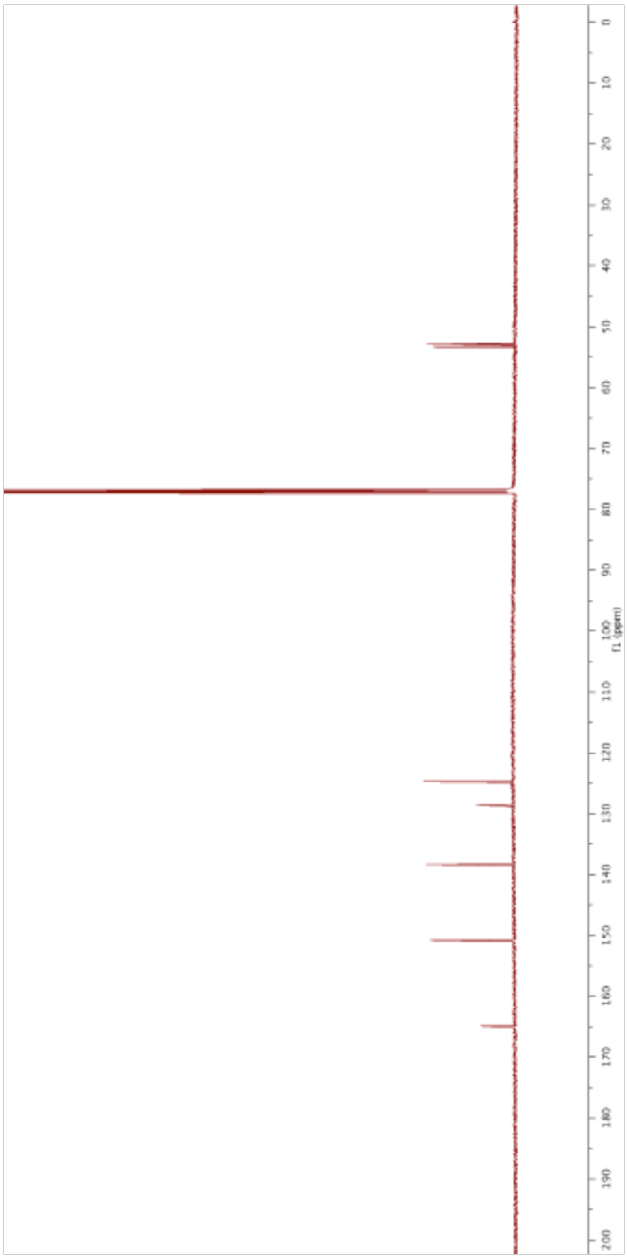
125 MHz ^{13}C NMR spectrum of 2-(3-Carboxy-1*H*-pyrazol-5-yl)pyridine-5-carboxylic Acid (pypyrDC) in DMSO-d_6



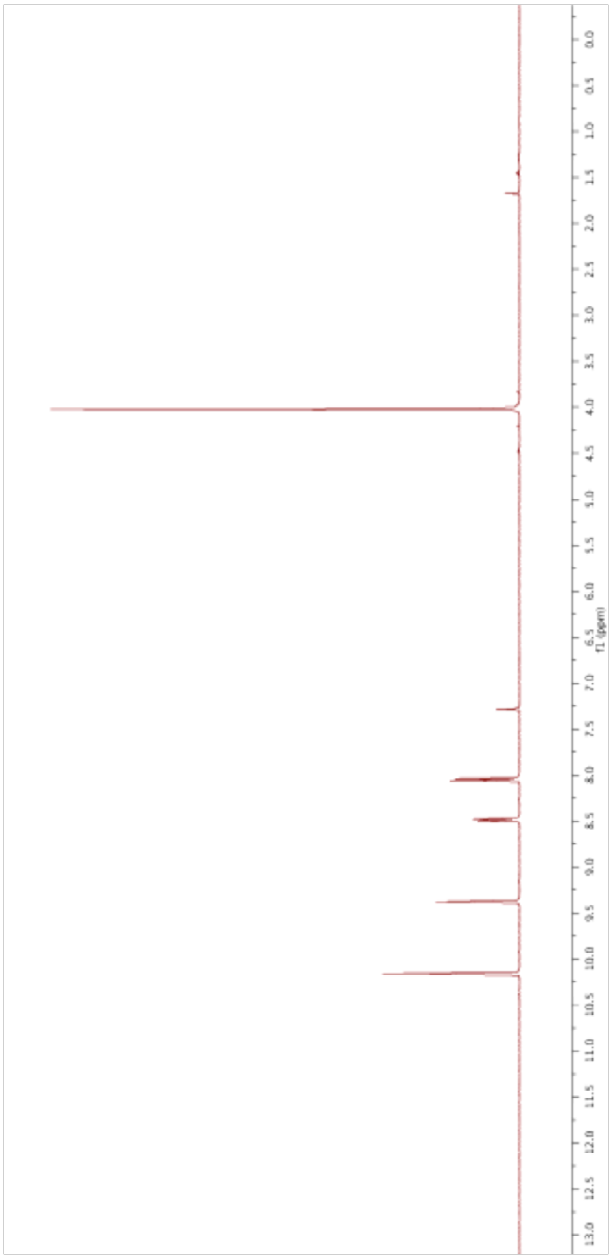
500 MHz ^{1}H NMR spectrum of Dimethyl 25PDC in $CDCl_3$



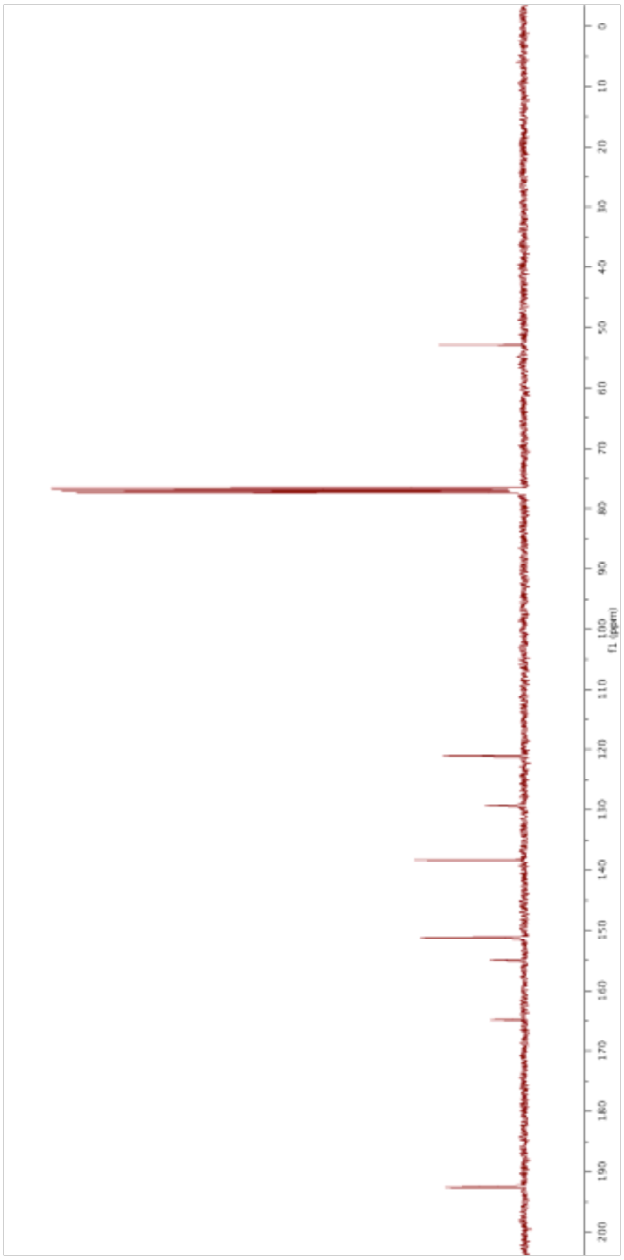
125 MHz ^{13}C NMR spectrum of Dimethyl 25PDC in $CDCl_3$



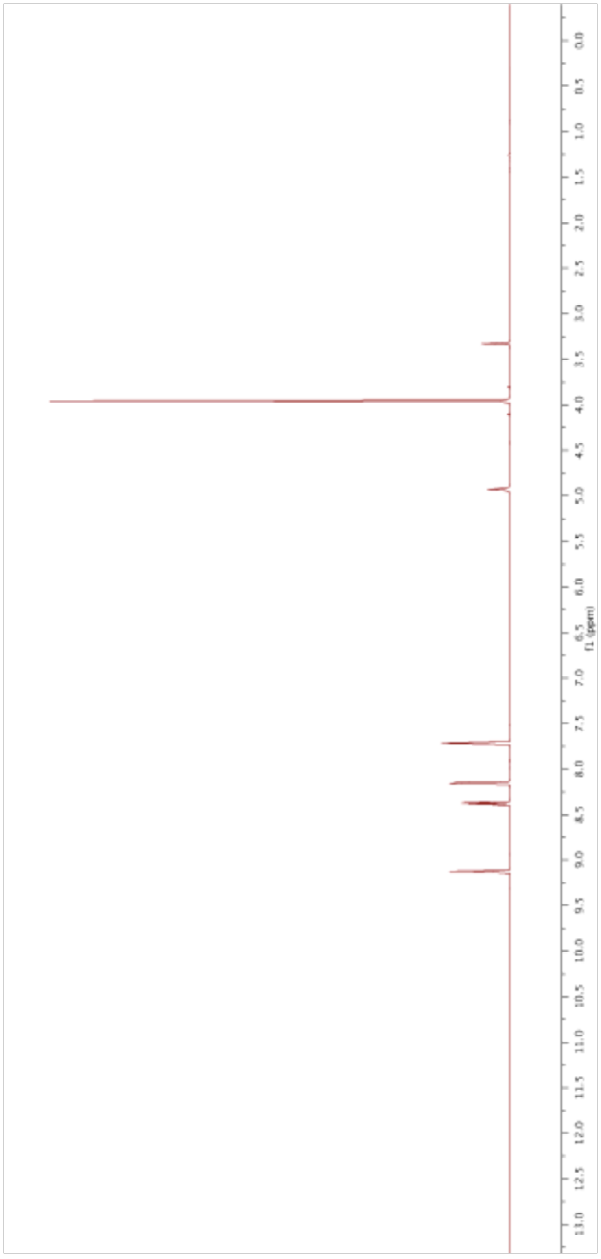
400 MHz ^{1}H NMR spectrum of Methyl 2-Formylnicotinate in $CDCl_3$



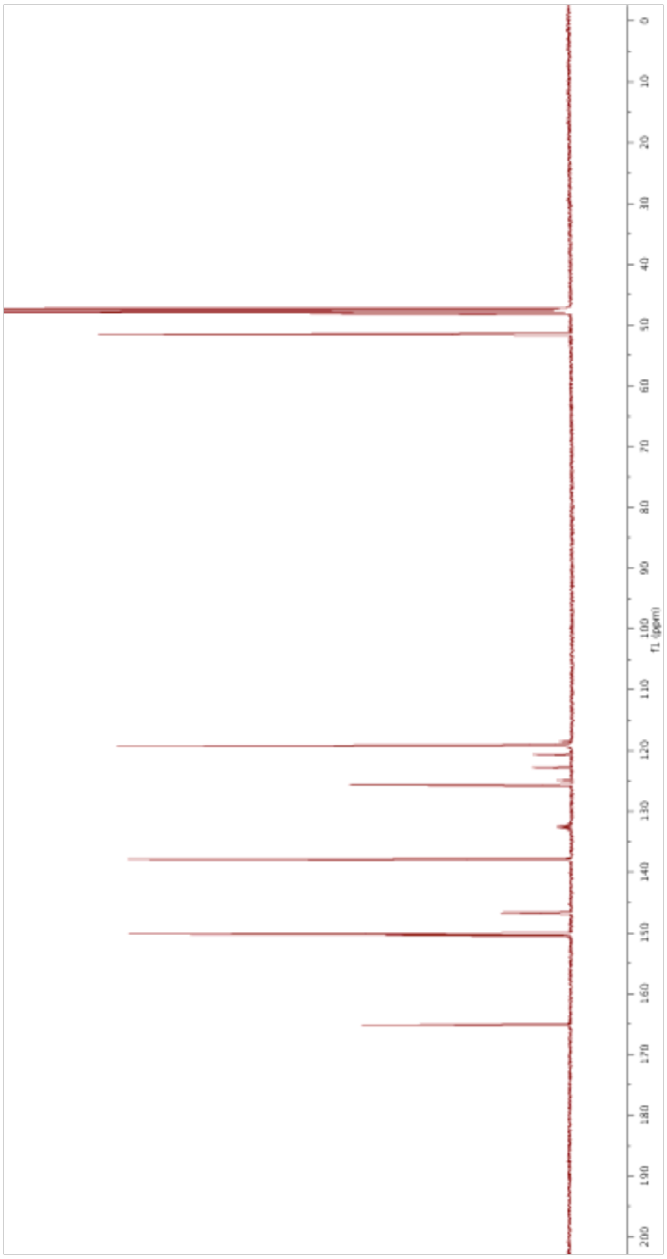
100 MHz ^{13}C NMR spectrum of Methyl 2-Formylnicotinate in $CDCl_3$



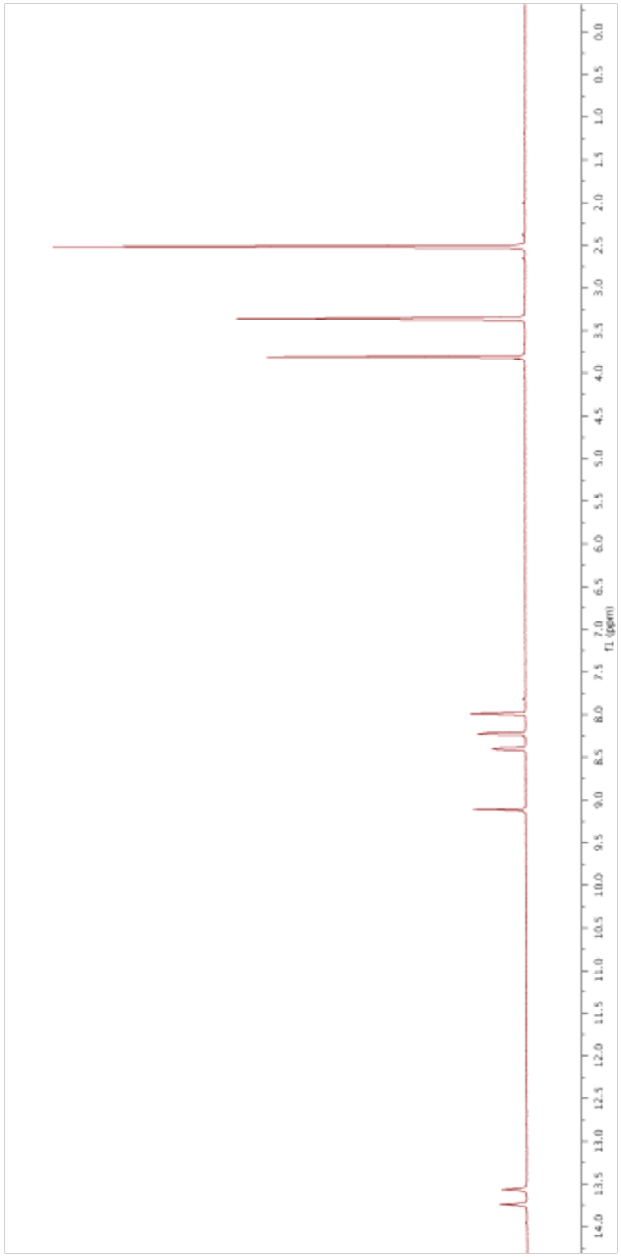
500 MHz ^{1}H NMR spectrum of Methyl 2-(5-Trifluoromethyl-1*H*-imidazol-2-yl)pyridine-5-carboxylate in CD_3OD



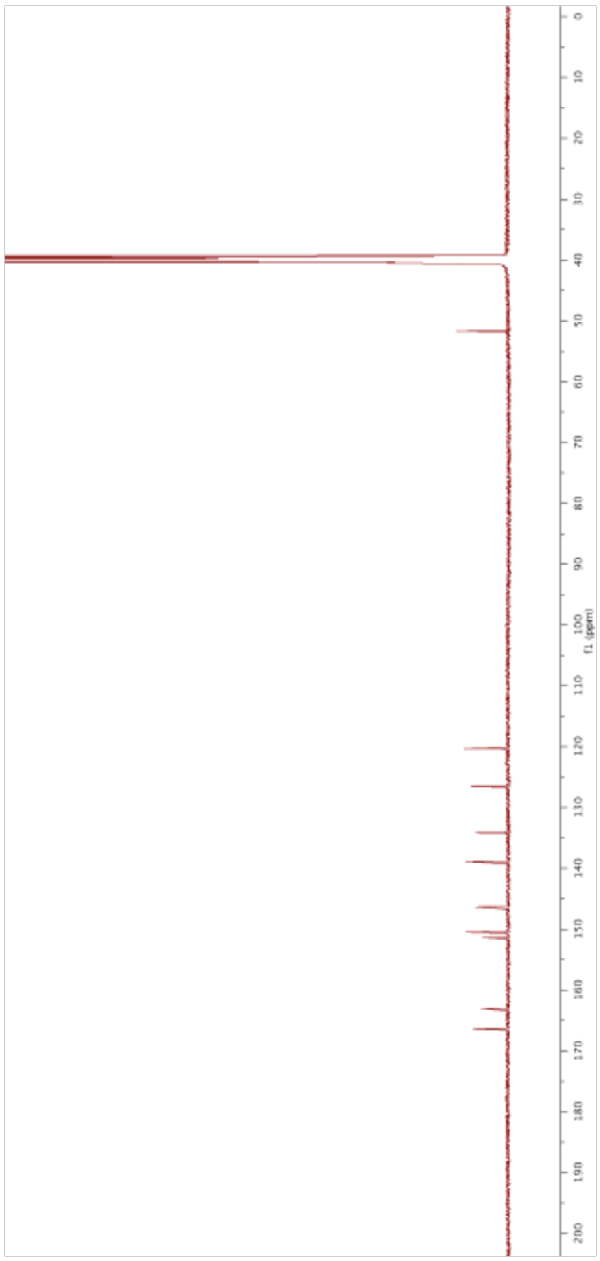
125 MHz ^{13}C NMR spectrum of Methyl 2-(5-Trifluoromethyl-1*H*-imidazol-2-yl)pyridine-5-carboxylate in CD_3OD



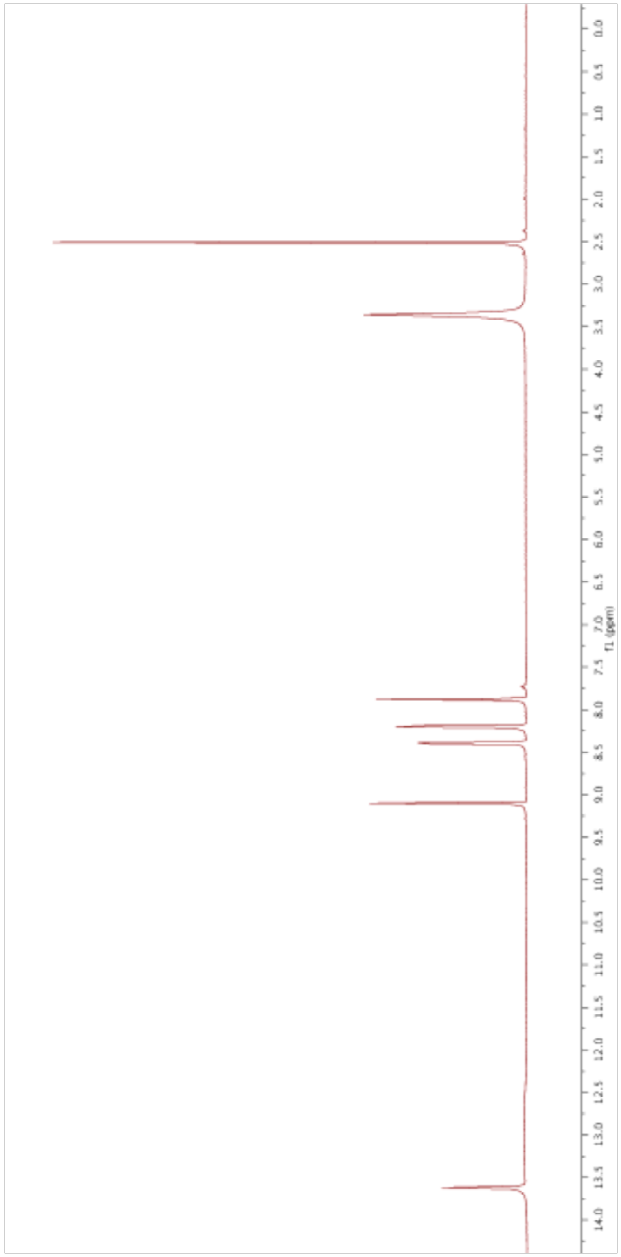
500 MHz ^{1}H NMR spectrum of 2-(5-Methoxycarbonyl-1*H*-imidazol-2-yl)pyridine-5-carboxylic Acid in DMSO-d_6



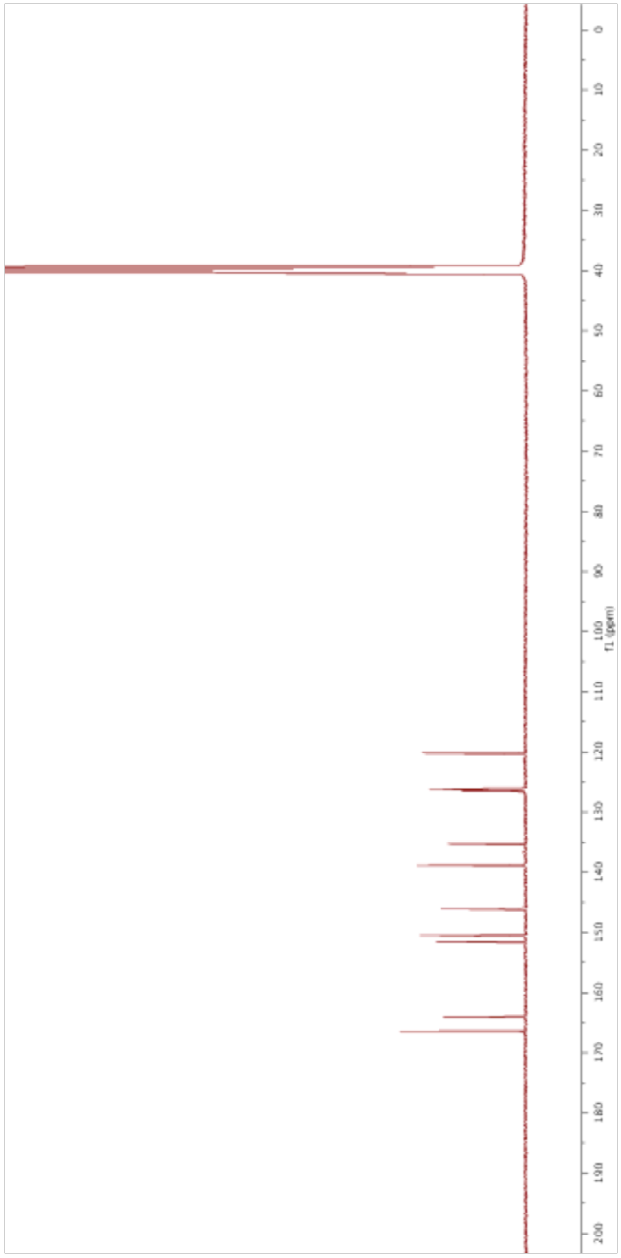
125 MHz ^{13}C NMR spectrum of 2-(5-Methoxycarbonyl-1*H*-imidazol-2-yl)pyridine-5-carboxylic Acid in DMSO-d_6



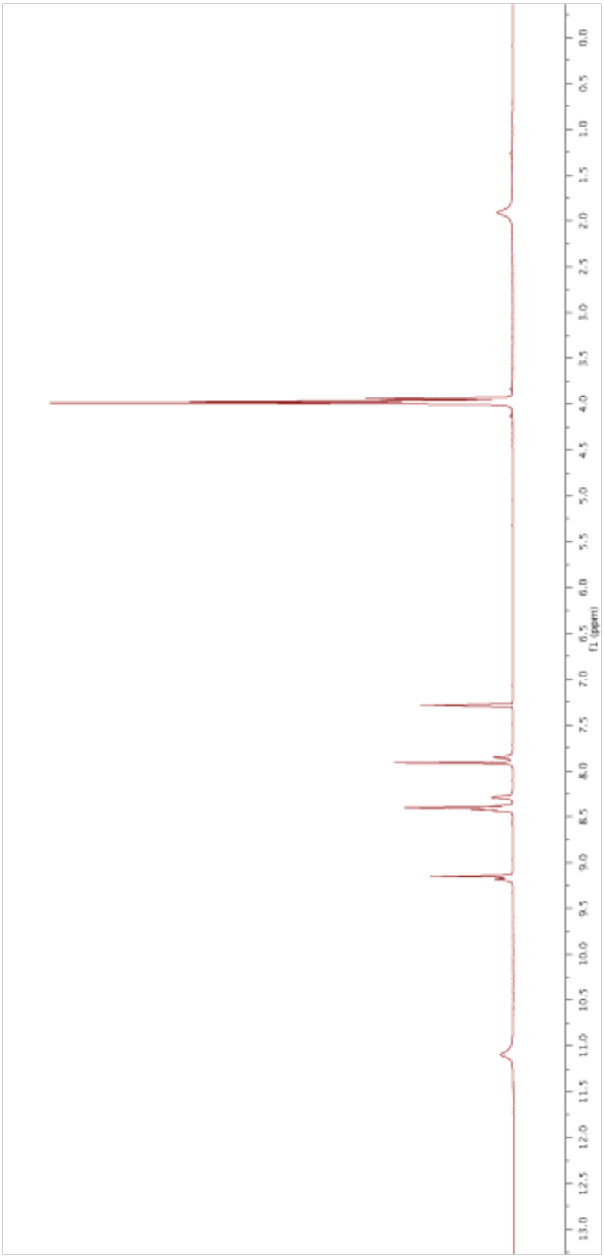
500 MHz ^{1}H NMR spectrum of 2-(5-Carboxy-1*H*-imidazol-2-yl)pyridine-5-carboxylic Acid (pyimDC) in DMSO-d_6



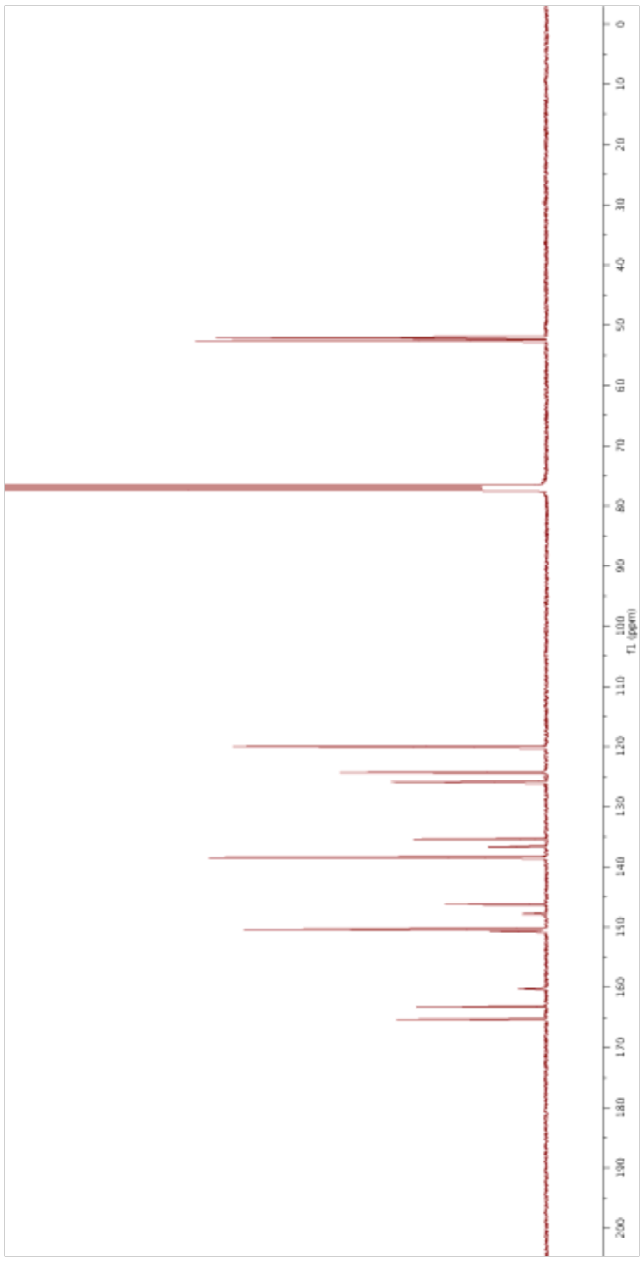
125 MHz ^{13}C NMR spectrum of 2-(5-Carboxy-1*H*-imidazol-2-yl)pyridine-5-carboxylic Acid (pyimDC) in DMSO-d_6



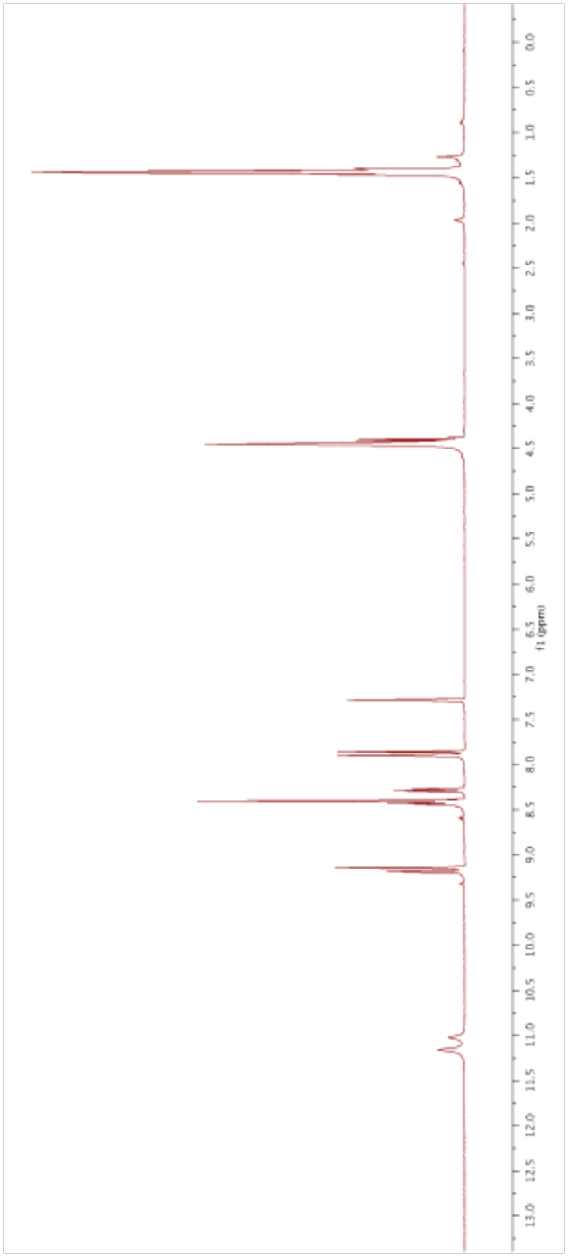
500 MHz ^{1}H NMR spectrum of Methyl 2-(5-Methoxycarbonyl-1*H*-imidazol-2-yl)pyridine-5-carboxylate (dimethyl pyimDC) in $CDCl_3$



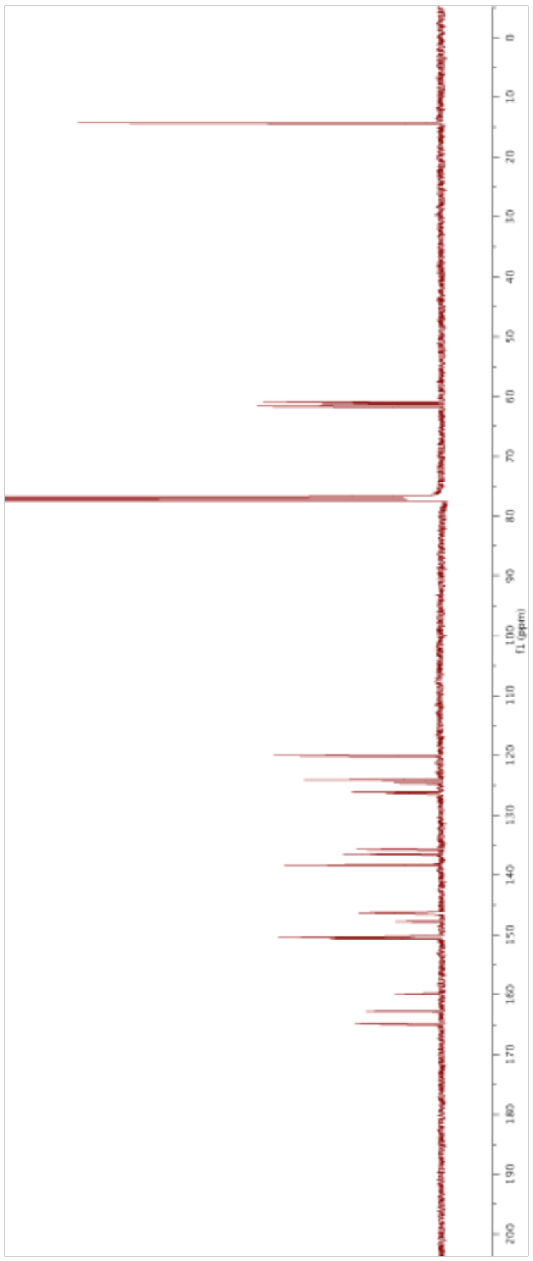
125 MHz ^{13}C NMR spectrum of Methyl 2-(5-Methoxycarbonyl-1*H*-imidazol-2-yl)pyridine-5-carboxylate (dimethyl pyimDC) in $CDCl_3$



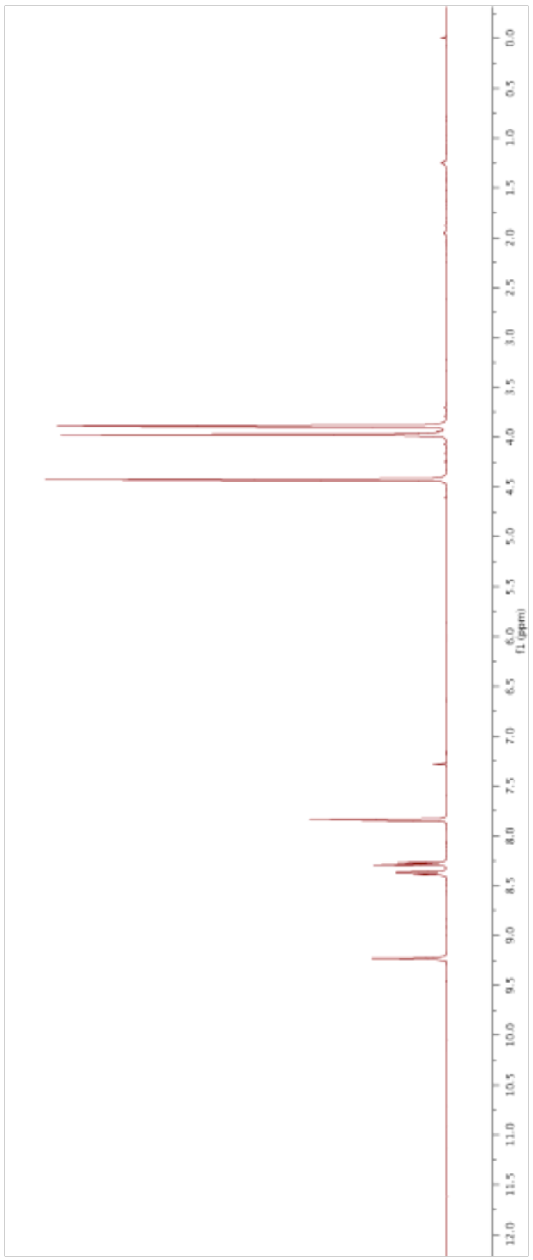
500 MHz ^{1}H NMR spectrum of Ethyl 2-(5-Ethoxycarbonyl-1*H*-imidazol-2-yl)pyridine-5-carboxylate (diethyl pyimDC) in $CDCl_3$



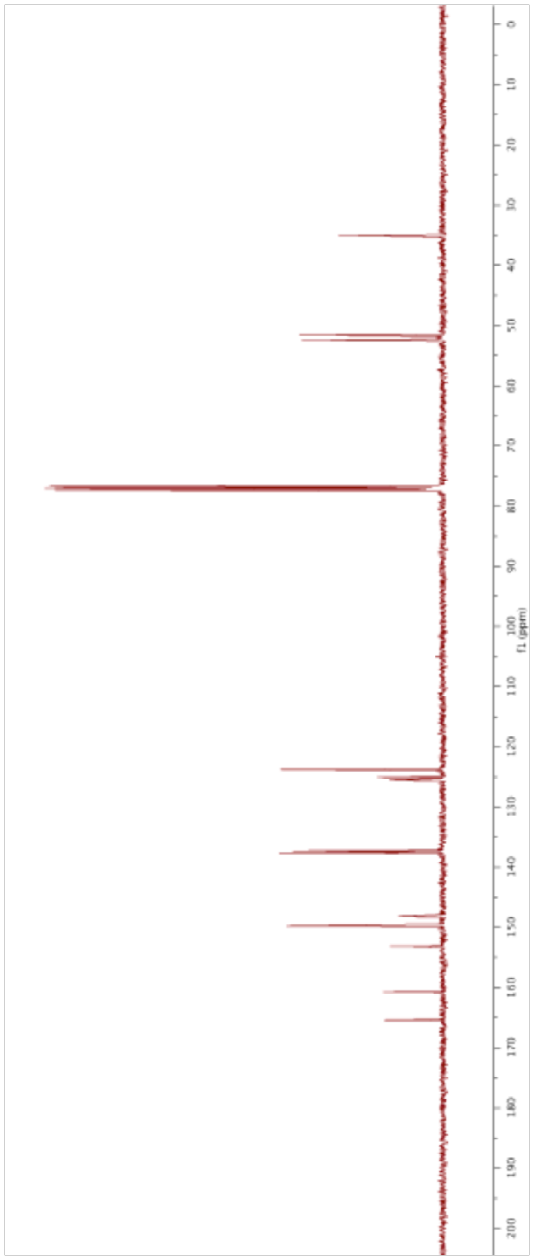
125 MHz ^{13}C NMR spectrum of Ethyl 2-(5-Ethoxycarbonyl-1*H*-imidazol-2-yl)pyridine-5-carboxylate (diethyl pyimDC) in $CDCl_3$



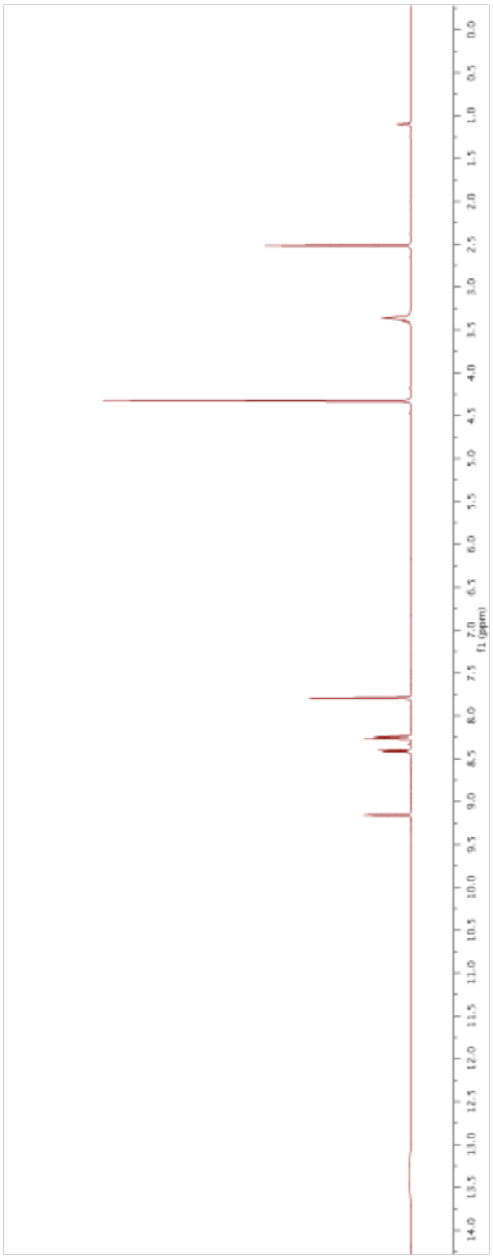
400 MHz ^{1}H NMR spectrum of Methyl 2-(5-Methoxycarbonyl-1-methyl-1*H*-imidazol-2-yl)pyridine-5-carboxylate (dimethyl NMe-pyimDC) in $CDCl_3$



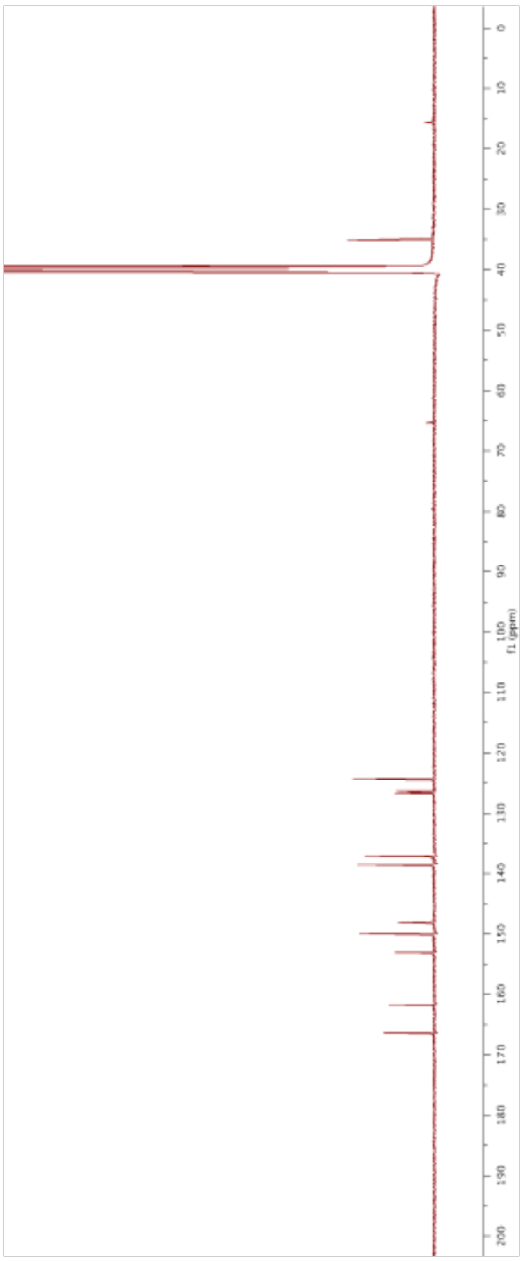
100 MHz ^{13}C NMR spectrum of Methyl 2-(5-Methoxycarbonyl-1-methyl-1*H*-imidazol-2-yl)pyridine-5-carboxylate (dimethyl NMe-pyimDC) in $CDCl_3$



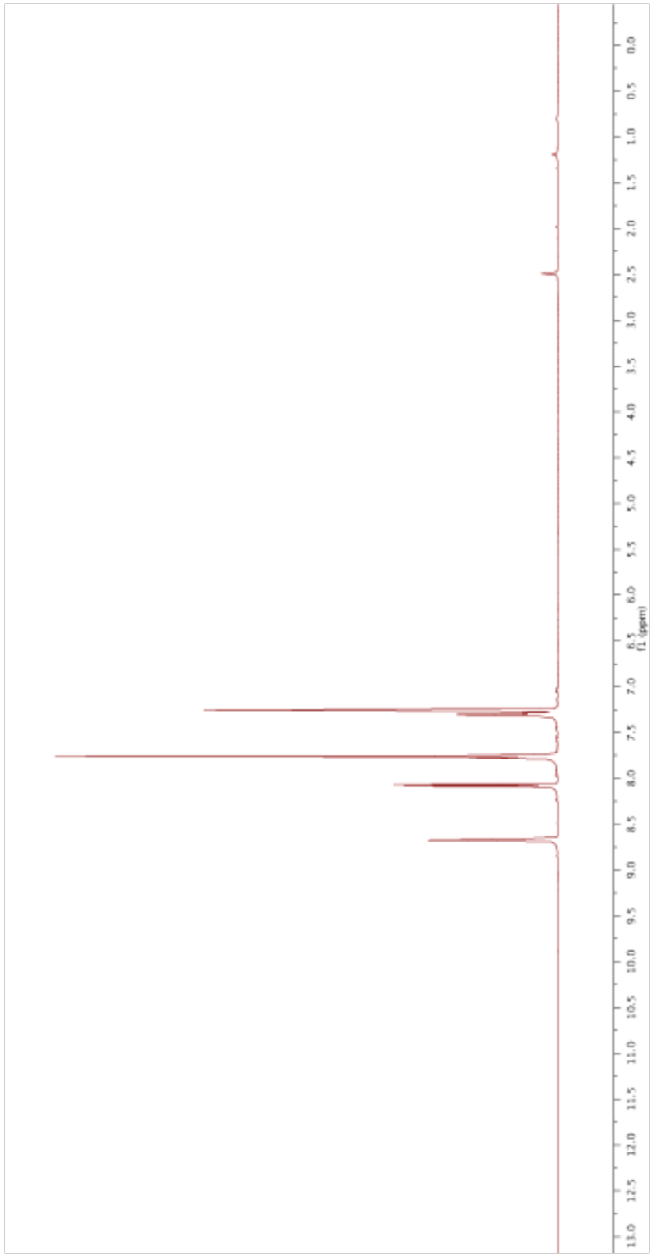
500 MHz ^{1}H NMR spectrum of 2-(5-Carboxy-1-methyl-1*H*-imidazol-2-yl)pyridine-5-carboxylic Acid (NMe-pyimDC) in DMSO-d_6



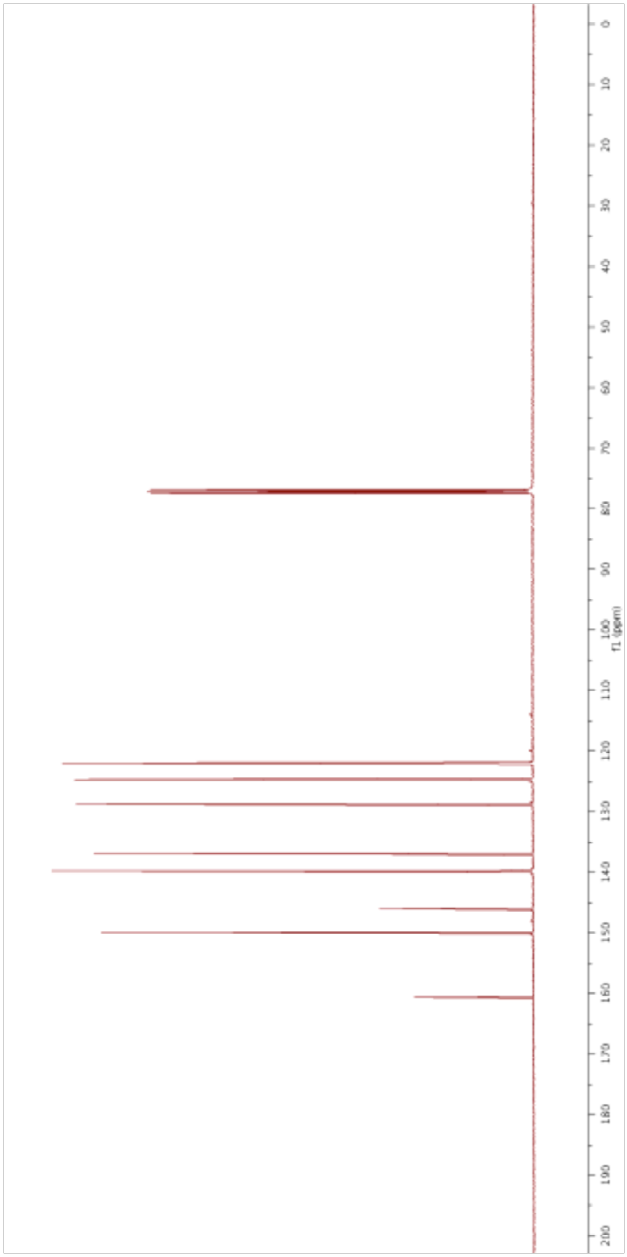
125 MHz ^{13}C NMR spectrum of 2-(5-Carboxy-1-methyl-1*H*-imidazol-2-yl)pyridine-5-carboxylic Acid (NMe-pyimDC) in DMSO-d_6



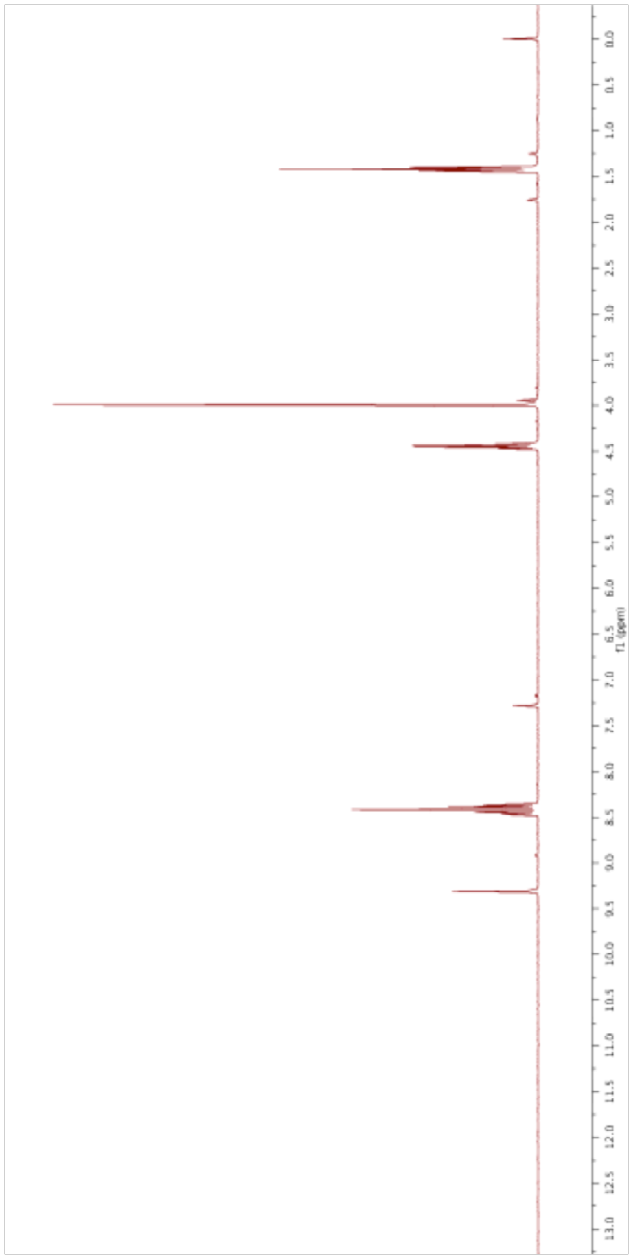
500 MHz ^{1}H NMR spectrum of 2-(Oxazol-2-yl)pyridine (pyox) in $CDCl_3$



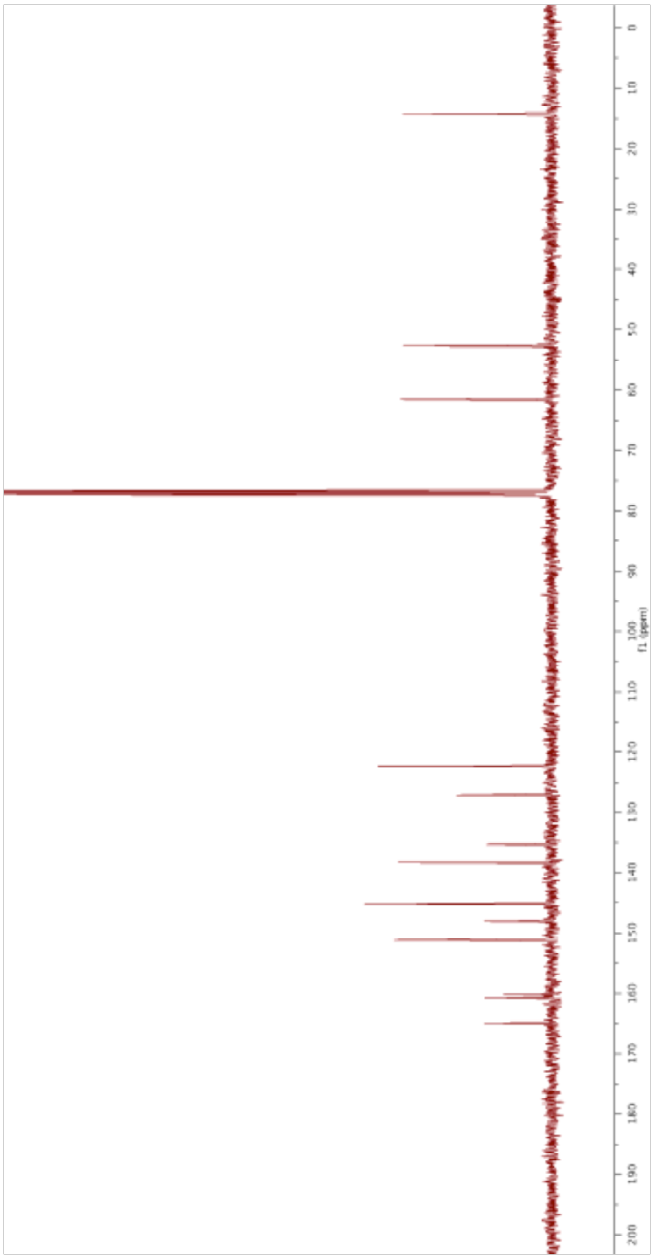
125 MHz ^{13}C NMR spectrum of 2-(Oxazol-2-yl)pyridine (pyox) in $CDCl_3$



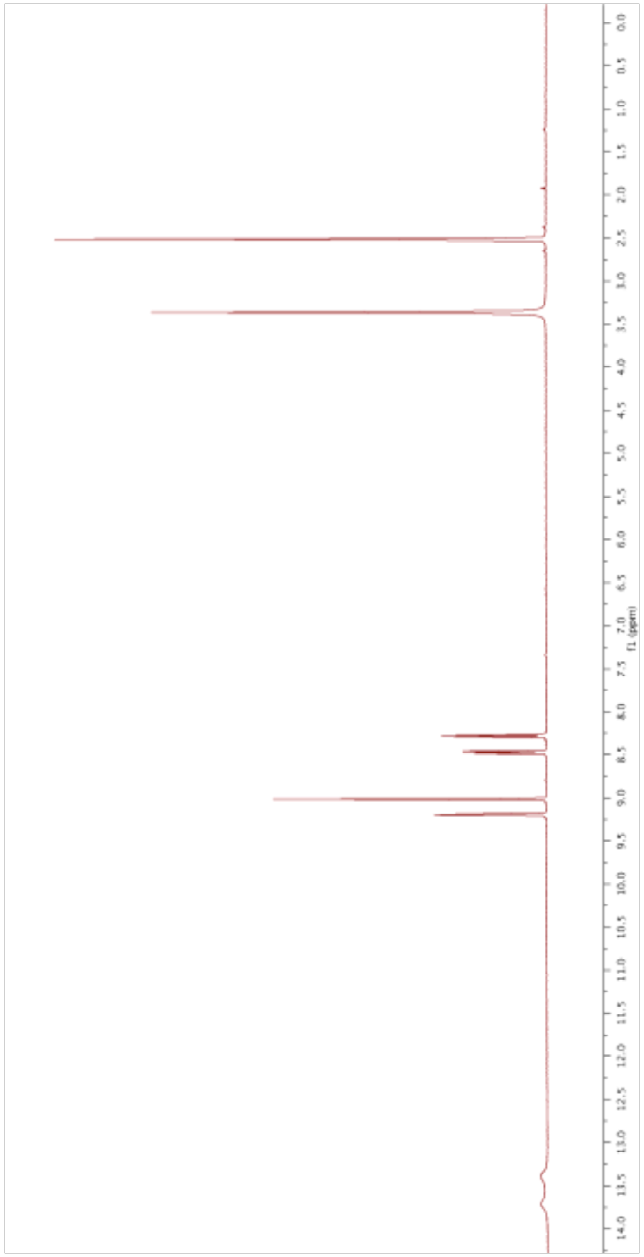
400 MHz ^{1}H NMR spectrum of Methyl 2-(4-Ethoxycarbonyloxazol-2-yl)pyridine-5-carboxylate (ethylmethyl pyoxDC*) in $CDCl_3$



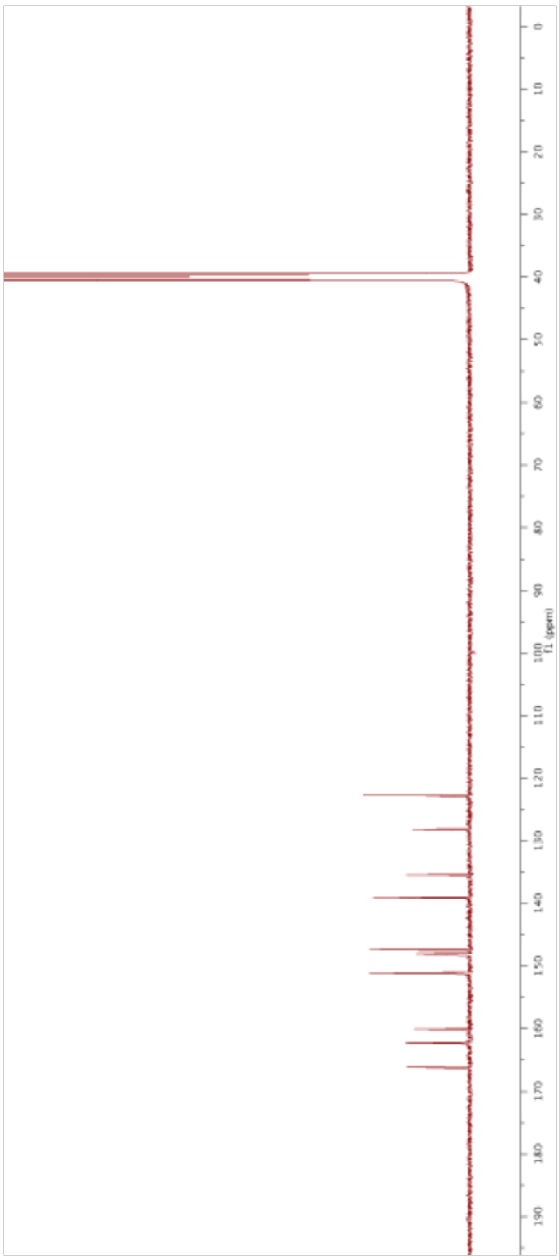
100 MHz ^{13}C NMR spectrum of Methyl 2-(4-Ethoxycarbonyloxazol-2-yl)pyridine-5-carboxylate (ethylmethyl pyoxDC*) in $CDCl_3$



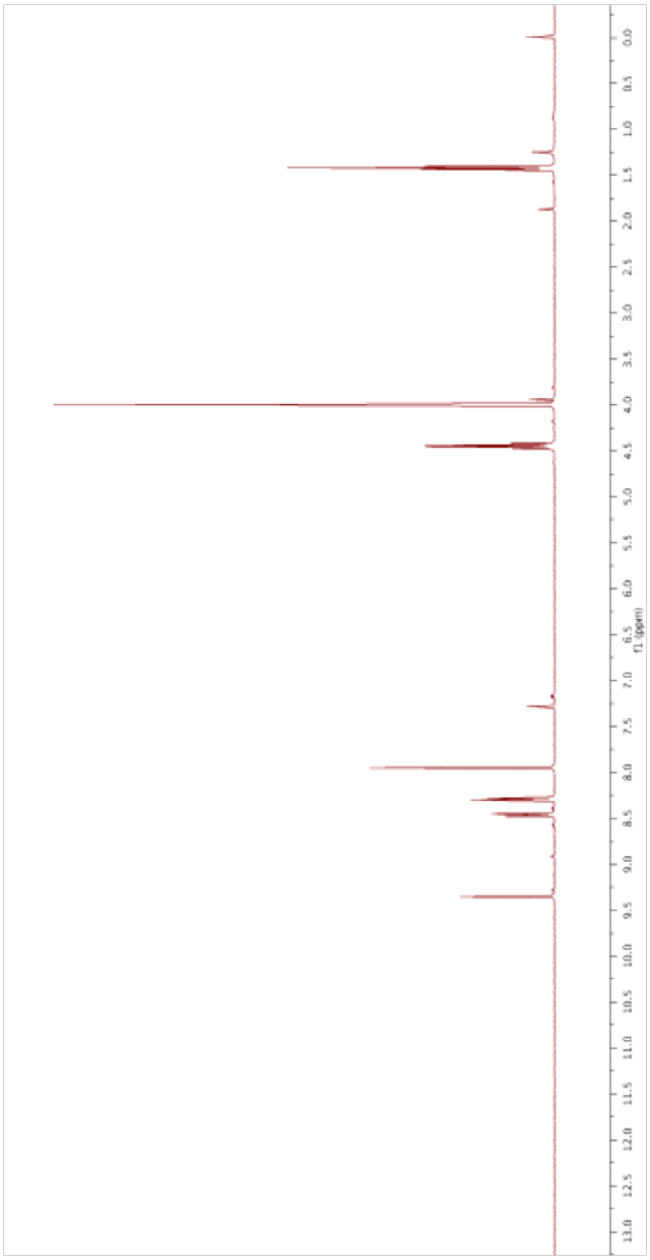
500 MHz ^{1}H NMR spectrum of 2-(4-Carboxyoxazol-2-yl)pyridine-5-carboxylic Acid (pyoxDC*) in DMSO-d_6



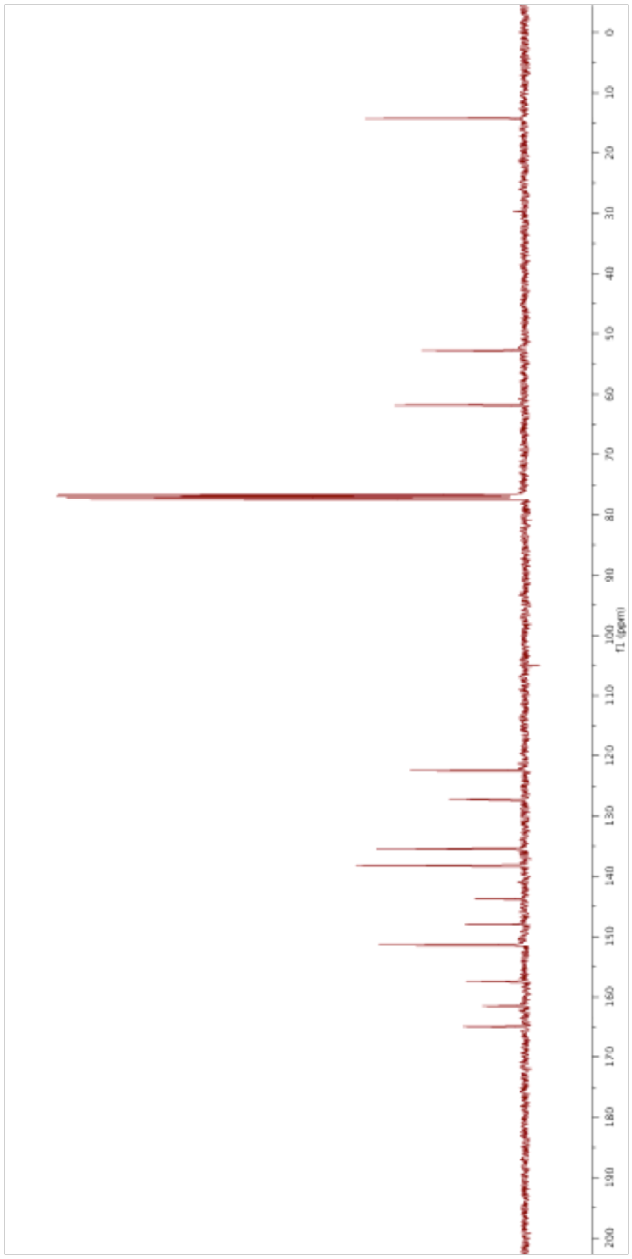
125 MHz ^{13}C NMR spectrum of 2-(4-Carboxyoxazol-2-yl)pyridine-5-carboxylic Acid (pyoxDC*) in DMSO-d_6



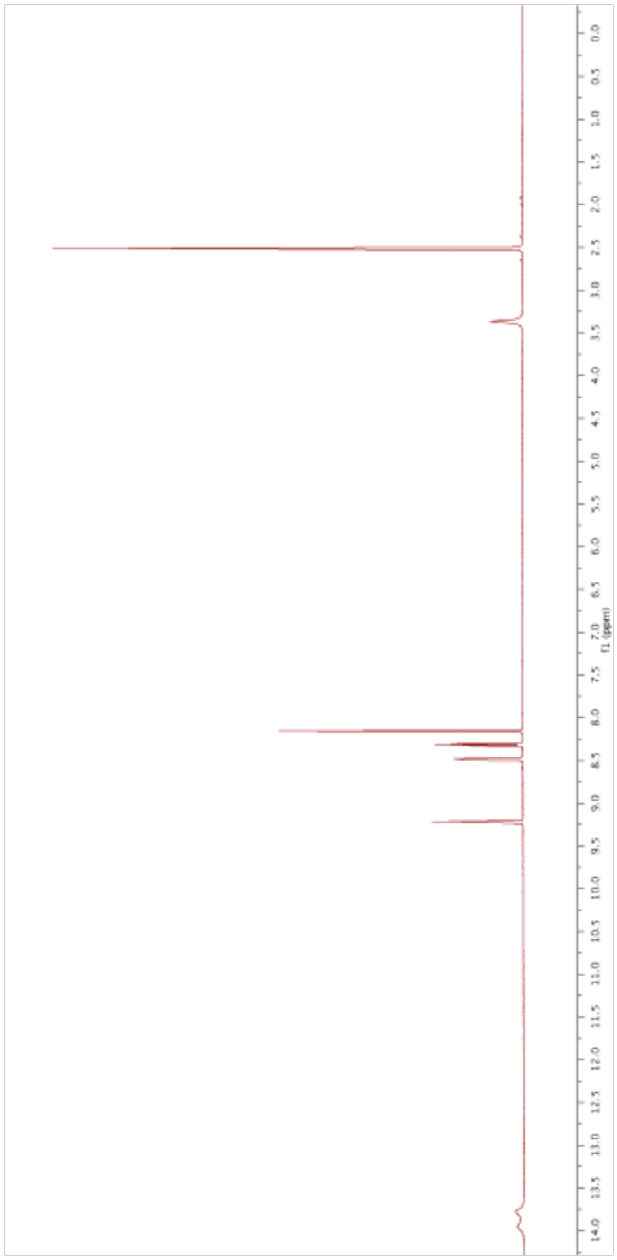
400 MHz ^{1}H NMR spectrum of Methyl 2-(5-Ethoxycarbonyloxazol-2-yl)pyridine-5-carboxylate (ethylmethyl pyoxDC) in $CDCl_3$



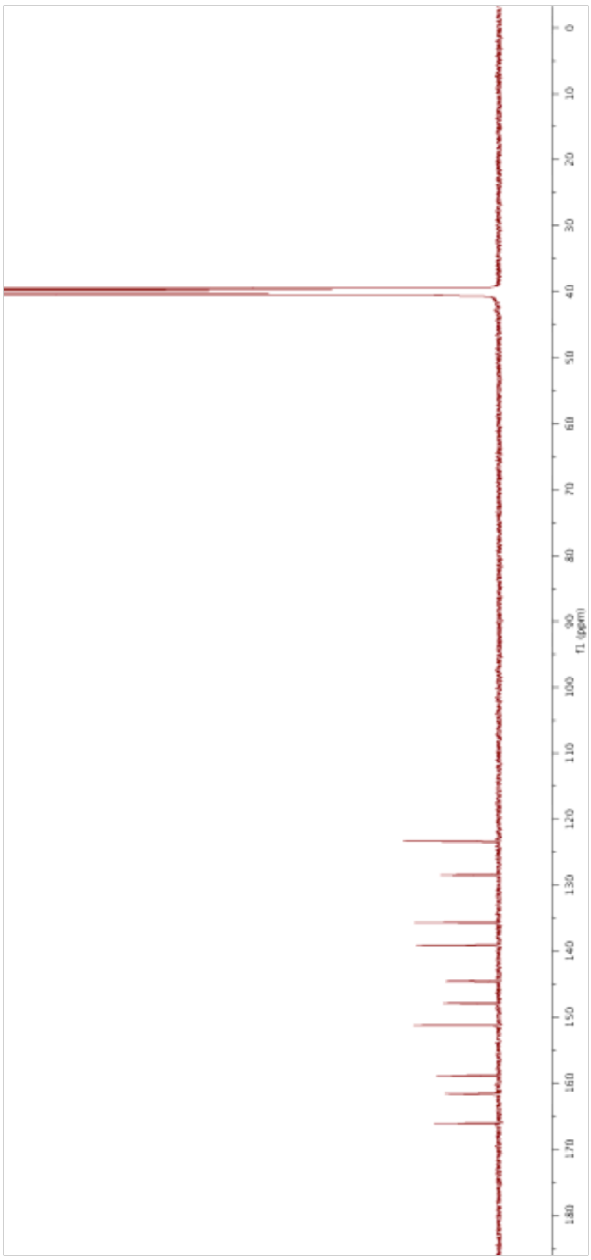
100 MHz ^{13}C NMR spectrum of Methyl 2-(5-Ethoxycarbonyloxazol-2-yl)pyridine-5-carboxylate (ethylmethyl pyoxDC) in $CDCl_3$



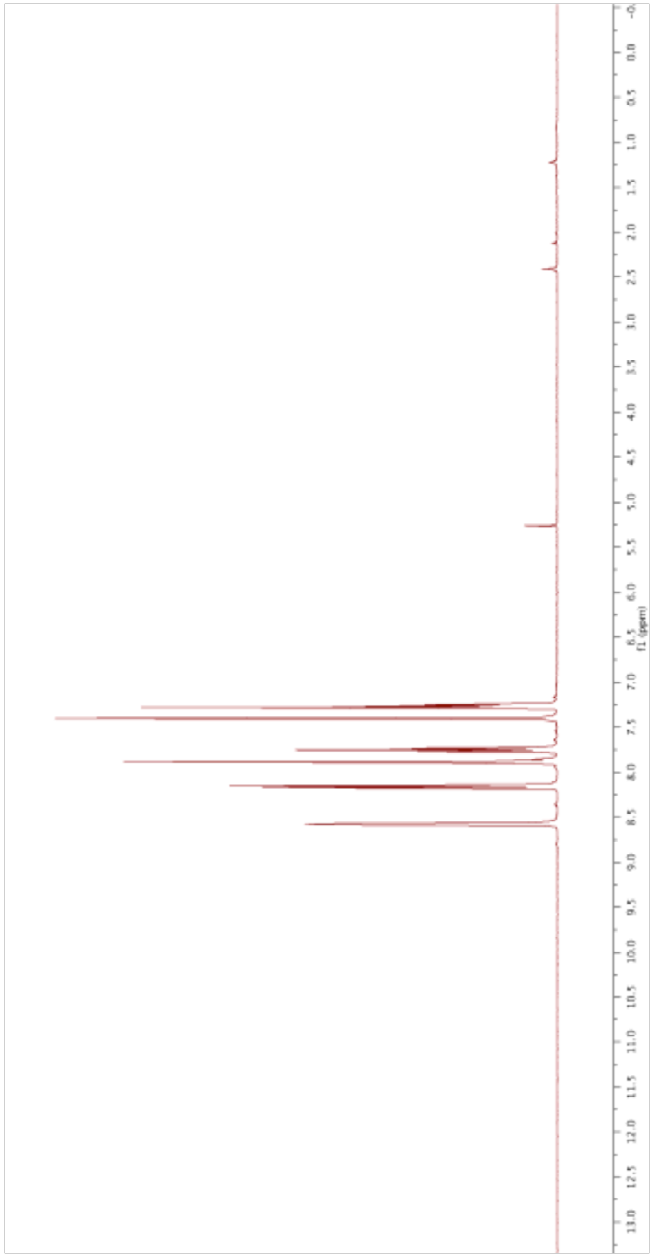
500 MHz ^{1}H NMR spectrum of 2-(5-Carboxyoxazol-2-yl)pyridine-5-carboxylic Acid (pyoxDC) in DMSO-d_6



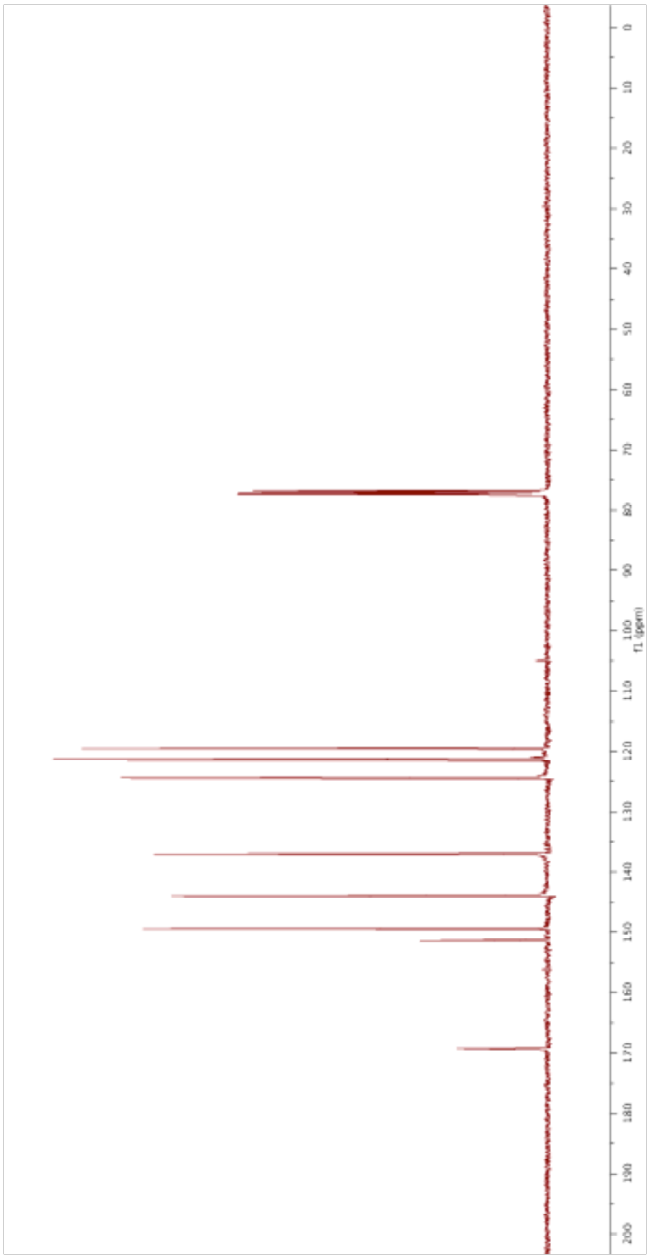
125 MHz ^{13}C NMR spectrum of 2-(5-Carboxyoxazol-2-yl)pyridine-5-carboxylic Acid (pyoxDC) in DMSO-d_6



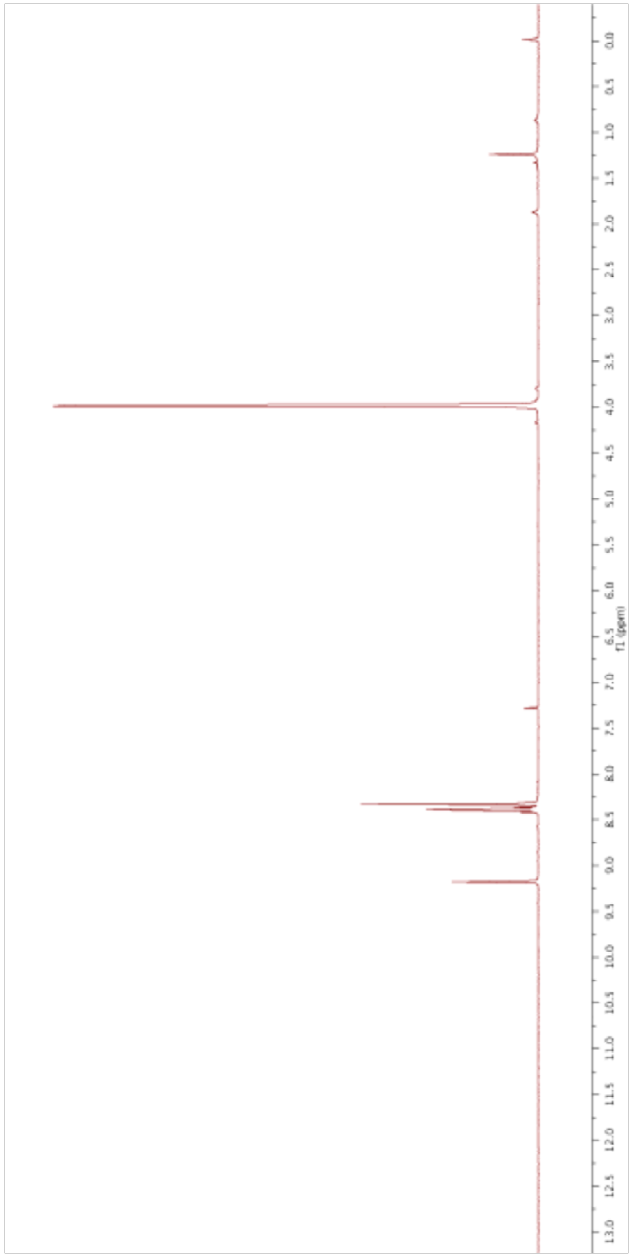
400 MHz ^{1}H NMR spectrum of 2-(Thiazol-2-yl)pyridine (pythi) in $CDCl_3$



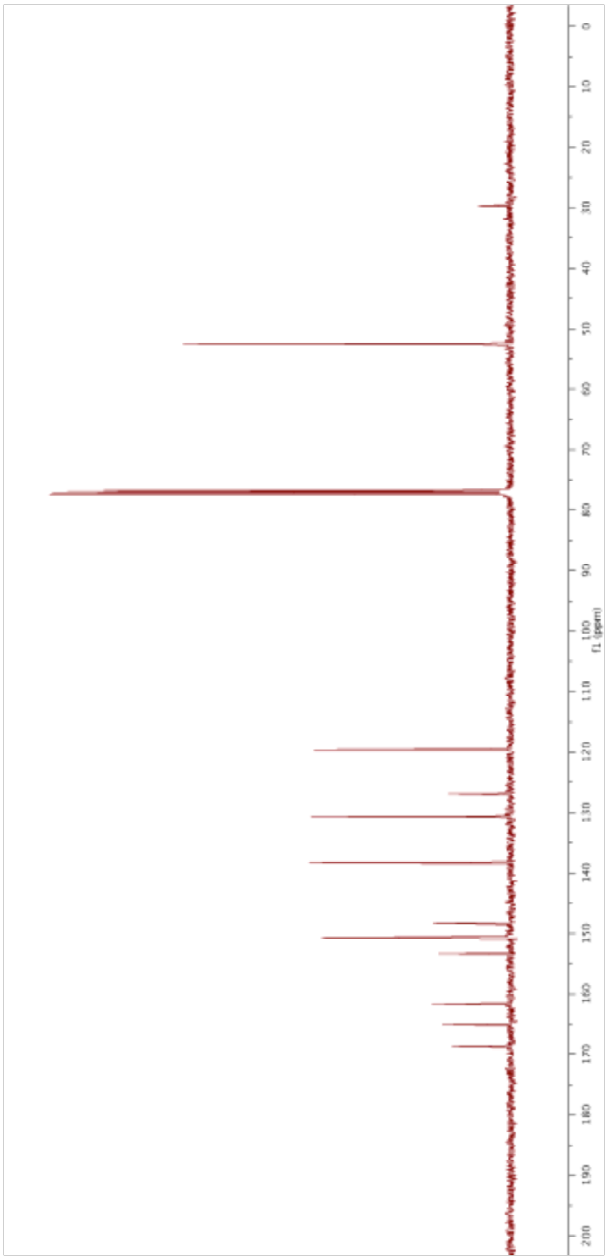
100 MHz ^{13}C NMR spectrum of 2-(Thiazol-2-yl)pyridine (pythi) in $CDCl_3$



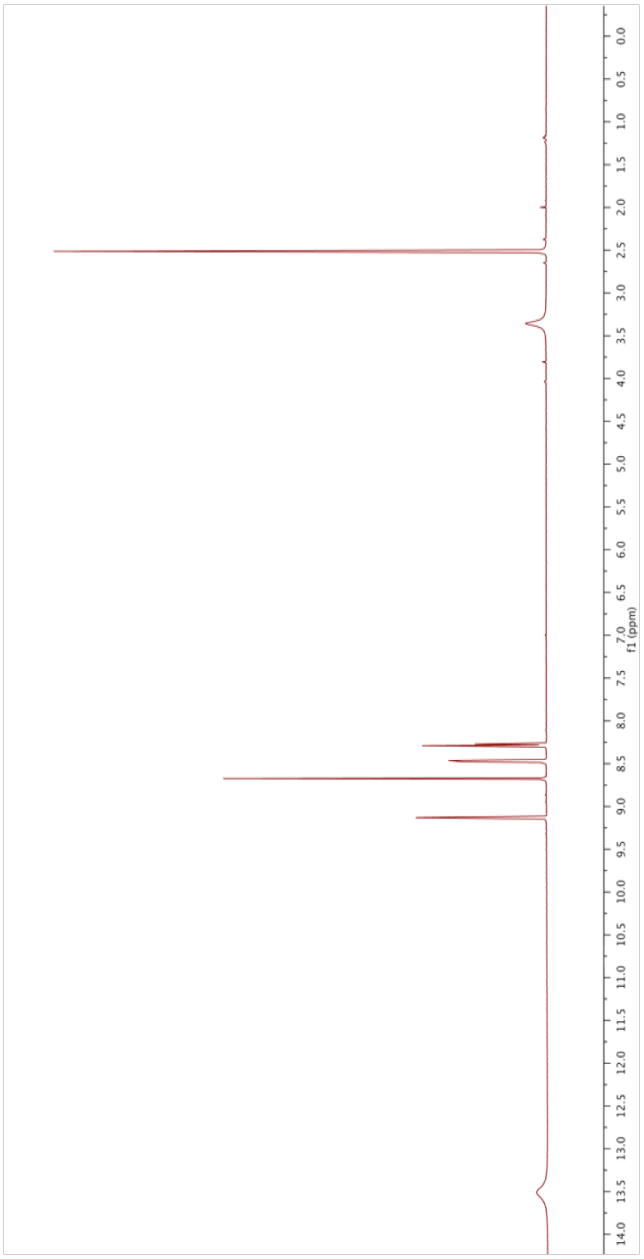
400 MHz ^{1}H NMR spectrum of Methyl 2-(4-Methoxycarbonylthiazol-2-yl)pyridine-5-carboxylate (dimethyl pythiDC*) in $CDCl_3$



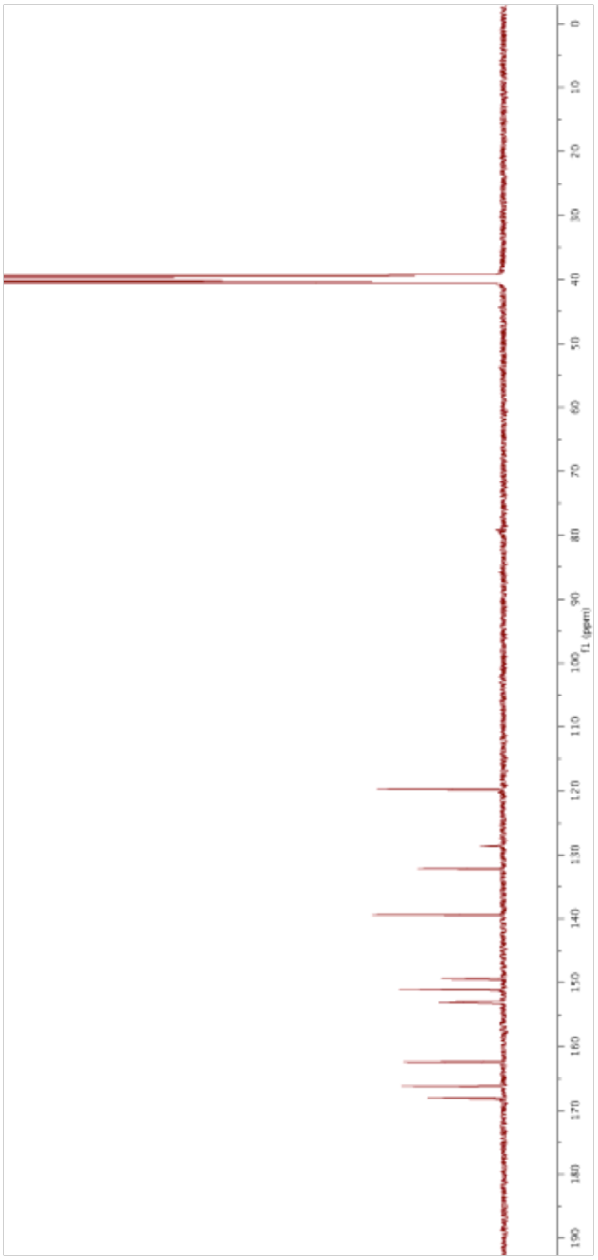
100 MHz ^{13}C NMR spectrum of Methyl 2-(4-Methoxycarbonylthiazol-2-yl)pyridine-5-carboxylate (dimethyl pythiDC*) in $CDCl_3$



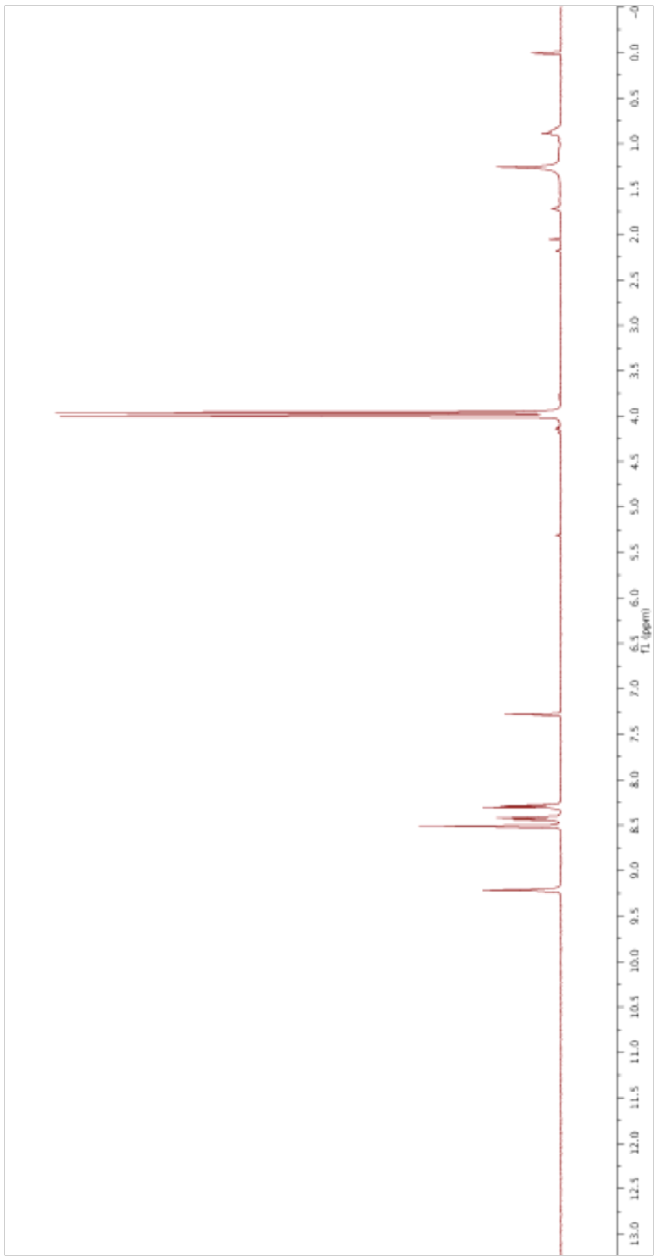
500 MHz ^{1}H NMR spectrum of 2-(4-Carboxythiazol-2-yl)pyridine-5-carboxylic Acid (pythiDC*) in DMSO-d_6



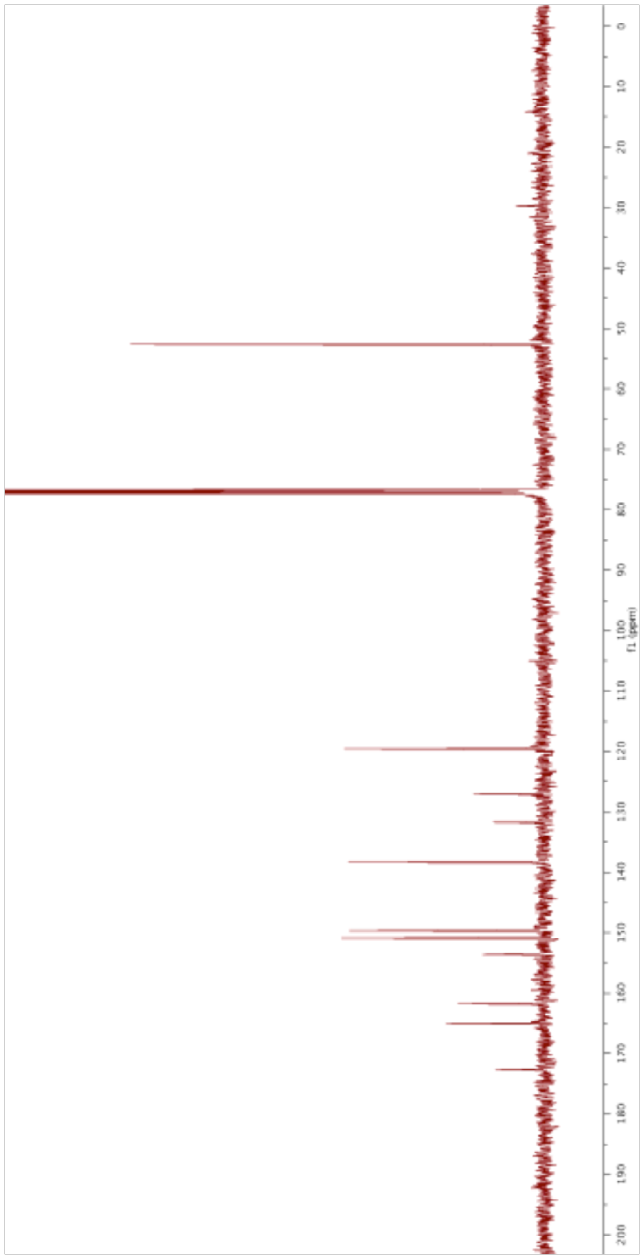
125 MHz ^{13}C NMR spectrum of 2-(4-Carboxythiazol-2-yl)pyridine-5-carboxylic Acid (pythiDC*) in DMSO-d_6



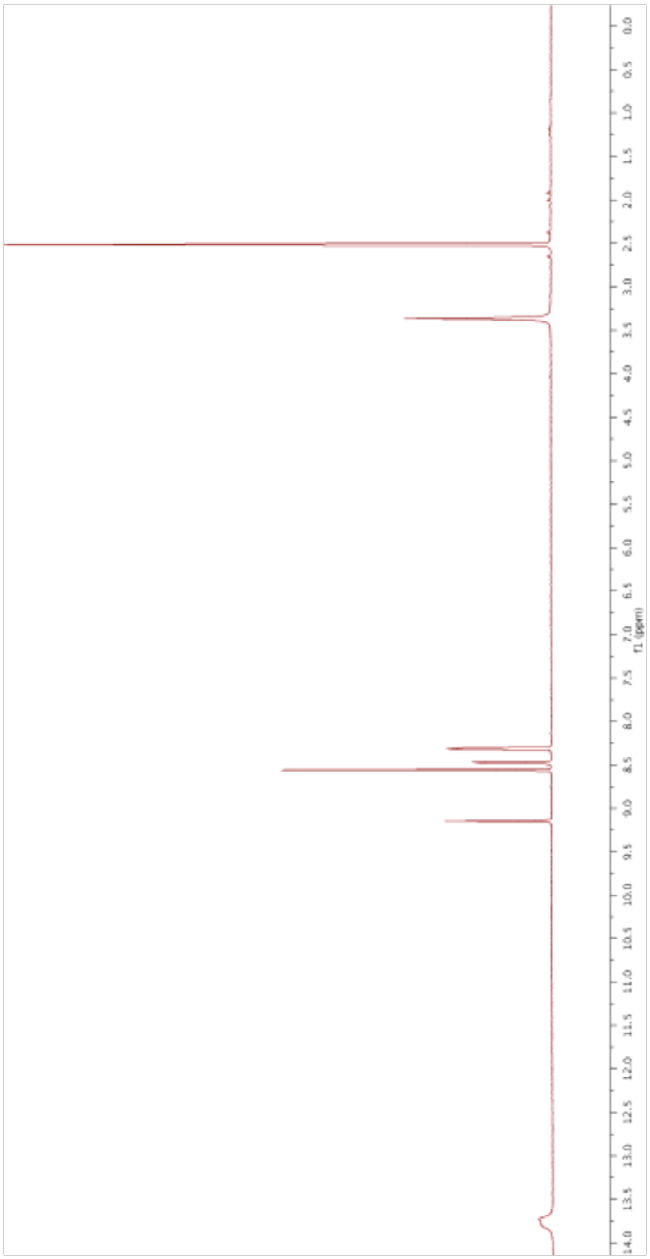
400 MHz ^{1}H NMR spectrum of Methyl 2-(5-Methoxycarbonylthiazol-2-yl)pyridine-5-carboxylate (dimethyl pythiDC) in $CDCl_3$



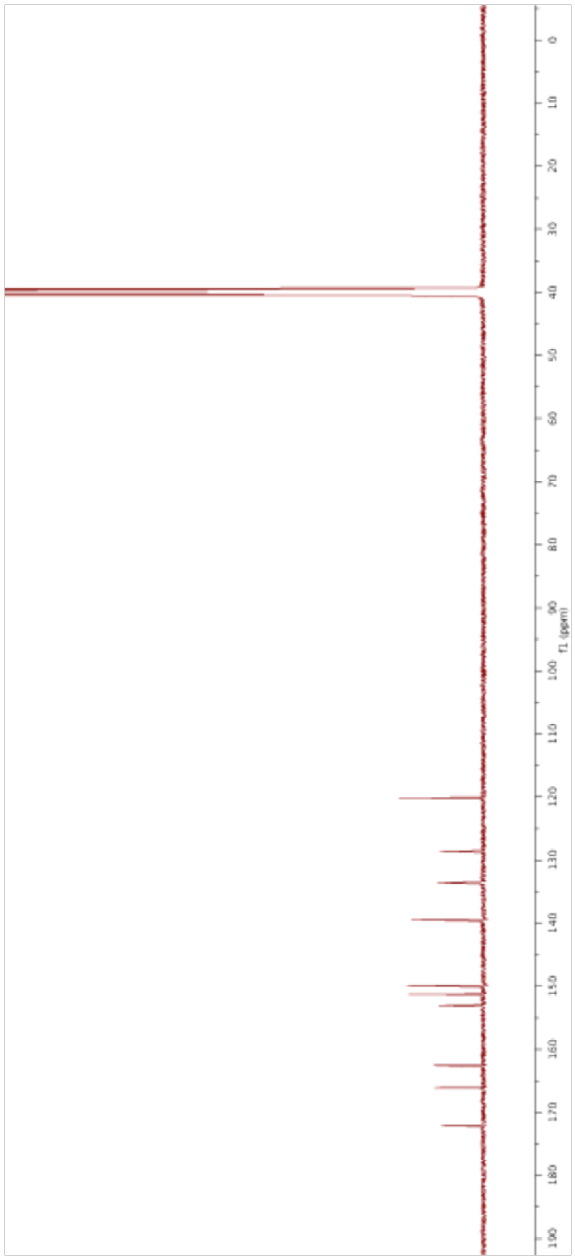
100 MHz ^{13}C NMR spectrum of Methyl 2-(5-Methoxycarbonylthiazol-2-yl)pyridine-5-carboxylate (dimethyl pythiDC) in $CDCl_3$



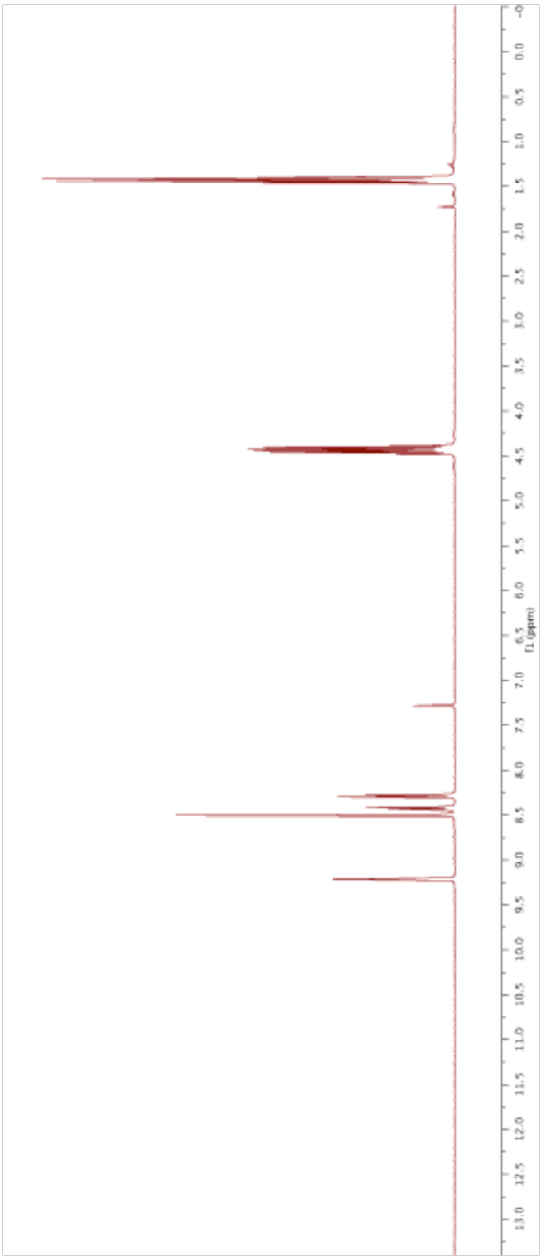
500 MHz ^{1}H NMR spectrum of 2-(5-Carboxythiazol-2-yl)pyridine-5-carboxylic Acid (pythiDC) in DMSO-d_6



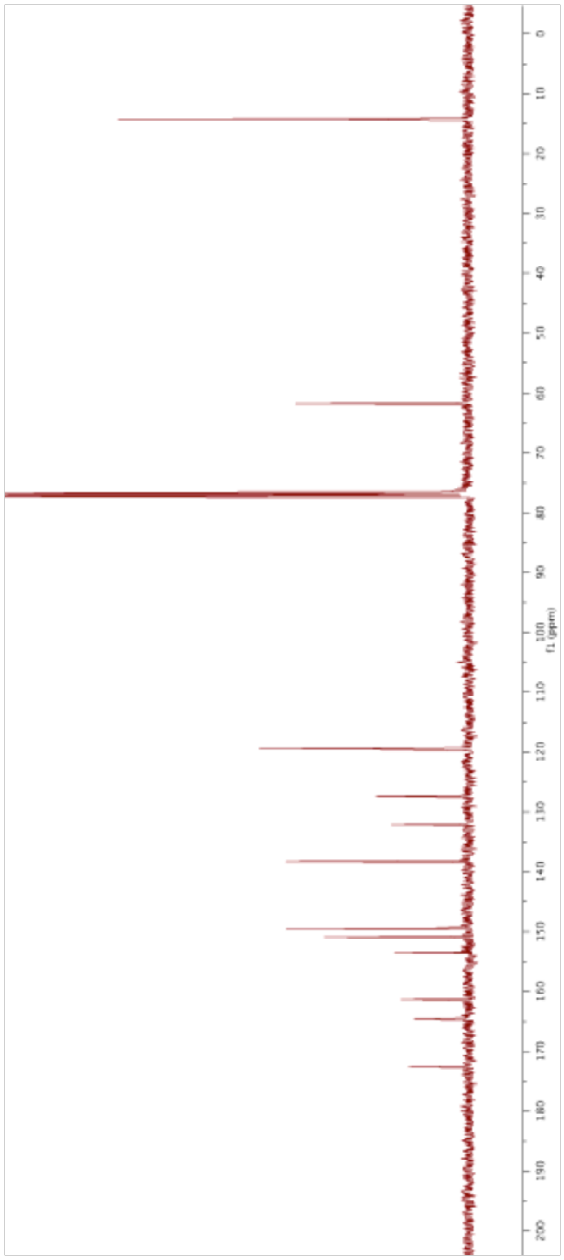
125 MHz ^{13}C NMR spectrum of 2-(5-Carboxythiazol-2-yl)pyridine-5-carboxylic Acid (pythiDC) in DMSO-d_6



400 MHz ^{1}H NMR spectrum of Ethyl 2-(5-Ethoxycarbonylthiazol-2-yl)pyridine-5-carboxylate (diethyl pythiDC) in $CDCl_3$



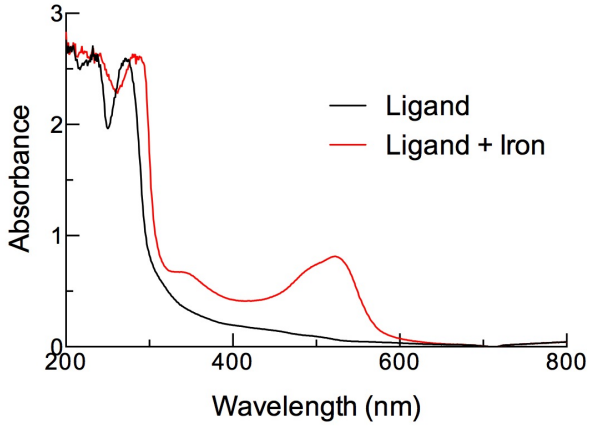
100 MHz ^{13}C NMR spectrum of Ethyl 2-(5-Ethoxycarbonylthiazol-2-yl)pyridine-5-carboxylate (diethyl pythiDC) in $CDCl_3$



3.5.21 Absorption Spectra of Ligands and Complexes

Pypyrid and Fe(pypyrid)$_3$

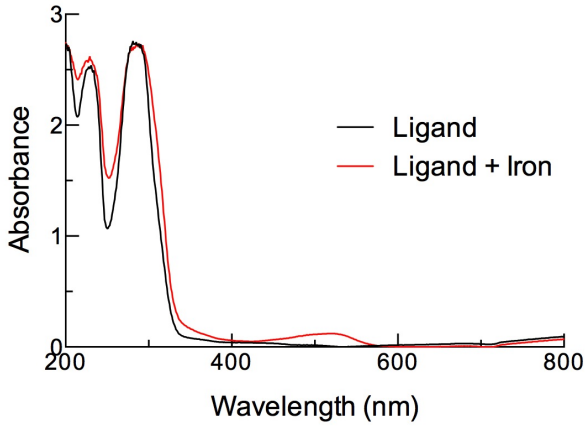
300 μM pypyrid ± 100 μM Fe(II)SO_4 in 10 mM sodium phosphate buffer, pH 7.0



λ_{max} = 522 nm

Pypyraz and Fe(pypyraz)$_3$

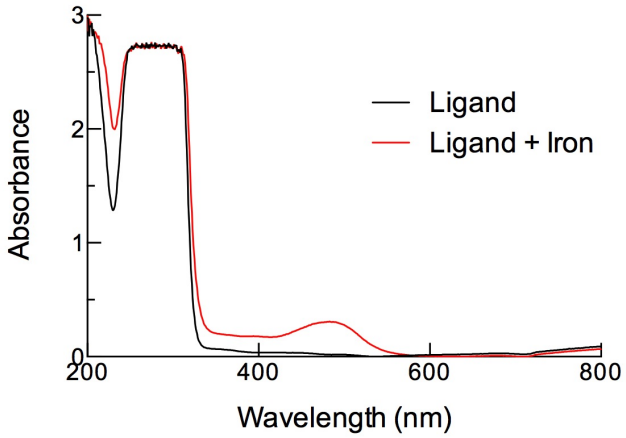
300 μM pypyraz ± 100 μM Fe(II)SO_4 in 10 mM sodium phosphate buffer, pH 7.0



λ_{max} = 523 nm

Pyim and Fe(pyim)$_3$

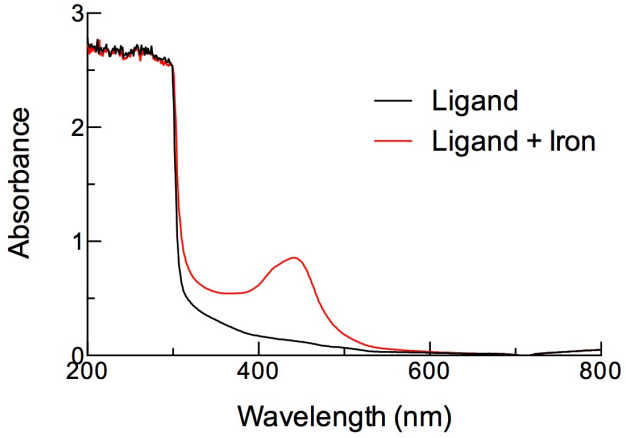
600 μM pyim ± 200 μM Fe(II)SO_4 in 10 mM sodium phosphate buffer, pH 7.0



λ_{max} = 484 nm

Pypyr and Its Complex with Fe(II)

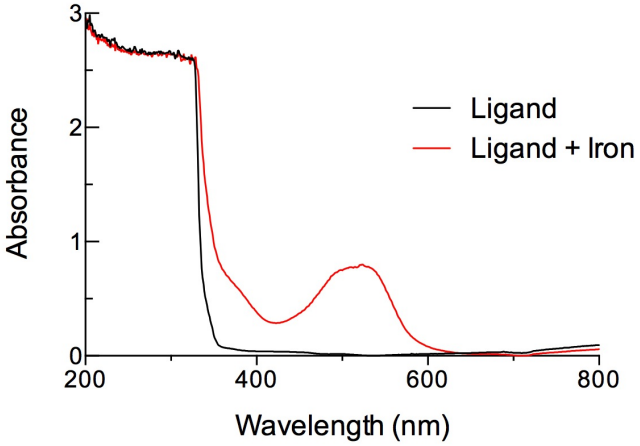
3000 μM pypyr ± 100 μM Fe(II)SO_4 in 10 mM sodium phosphate buffer, pH 7.0



λ_{max} = 441 nm

Pythi and Fe(pythi)$_3$

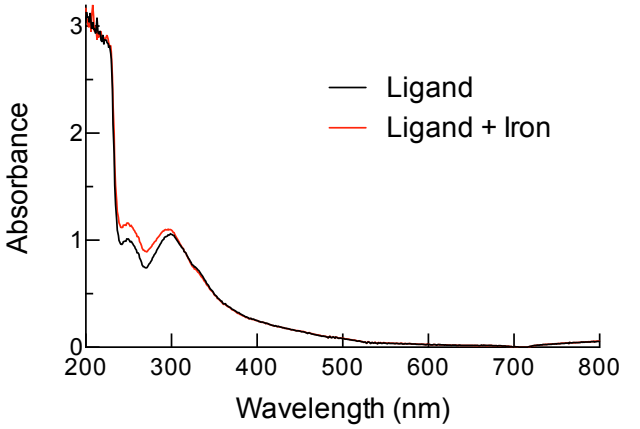
3000 μM pythi ± 1000 μM Fe(II)SO_4 in 10 mM sodium phosphate buffer, pH 7.0



λ_{max} = 524 nm

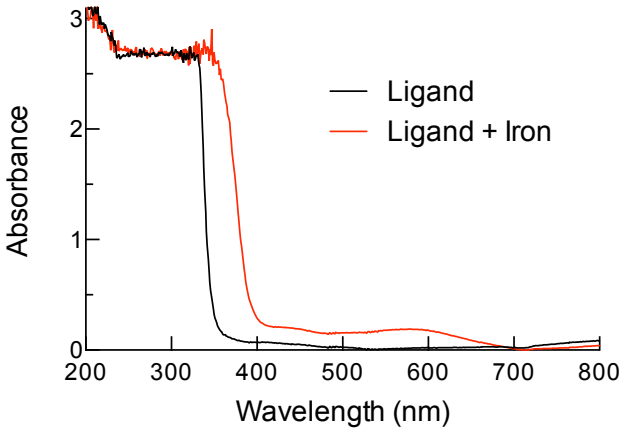
Diethyl Bipy55′DC

100 μM diethyl bipy55′DC ± 20 μM Fe(II)SO_4 in 10 mM sodium phosphate buffer, pH 7.0



Diethyl PyimDC and Its Complex with Fe(II)

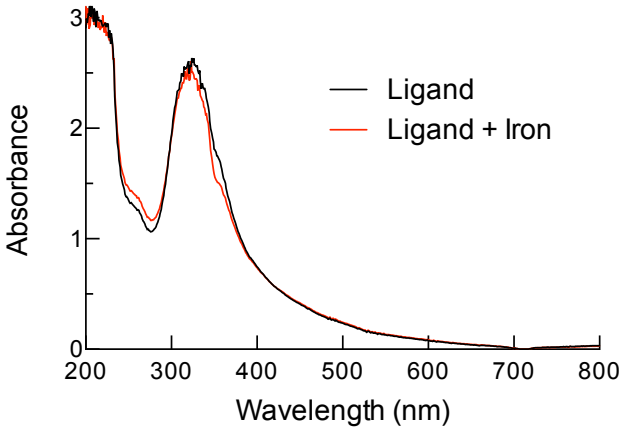
600 μM diethyl pyimDC ± 20 μM Fe(II)SO_4 in 10 mM sodium phosphate buffer, pH 7.0



λ_{max} = 577 nm

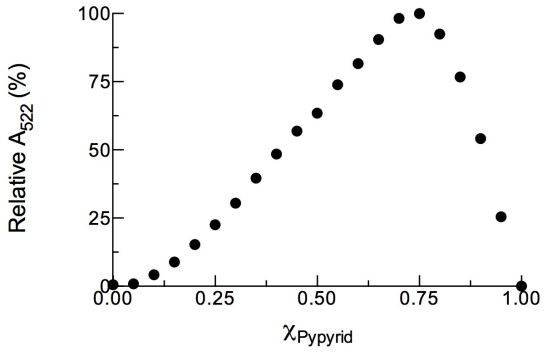
Diethyl PythiDC

200 μM diethyl pythiDC ± 20 μM Fe(II)SO_4 in 10 mM sodium phosphate buffer, pH 7.0

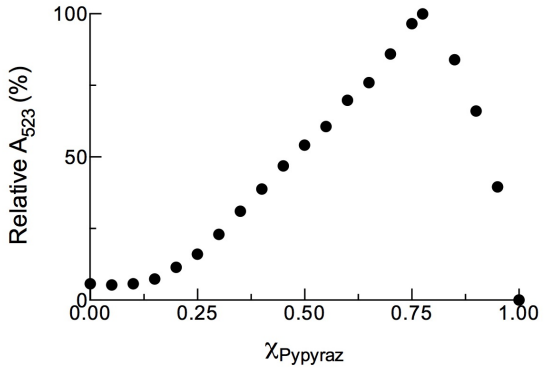


3.5.22 Job's Plots for Fe(II) Complexes

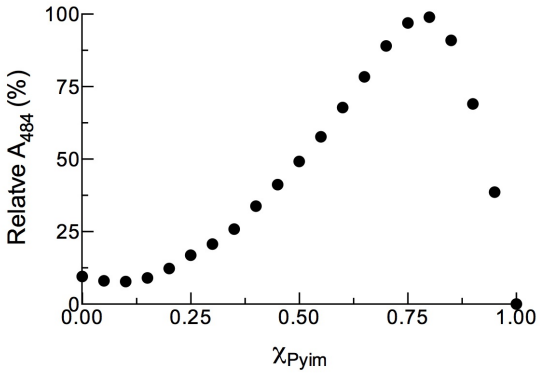
Fe(pypyrid)$_3$



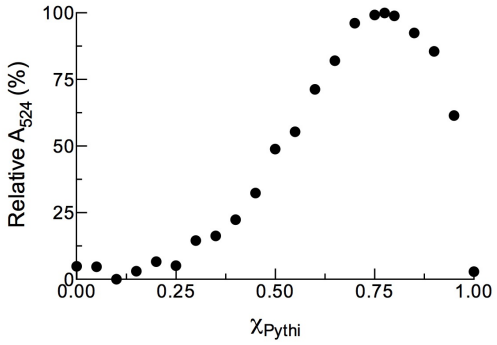
Fe(pypyraz)$_3$



Fe(pyim)$_3$

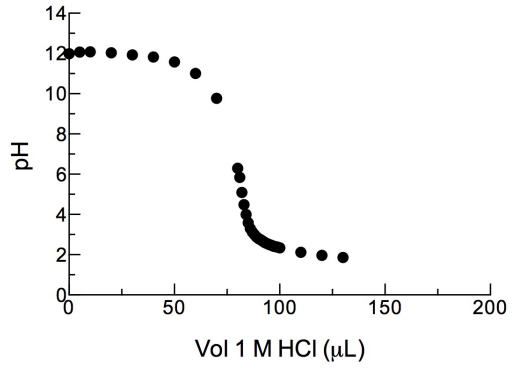


Fe(pythi)$_3$



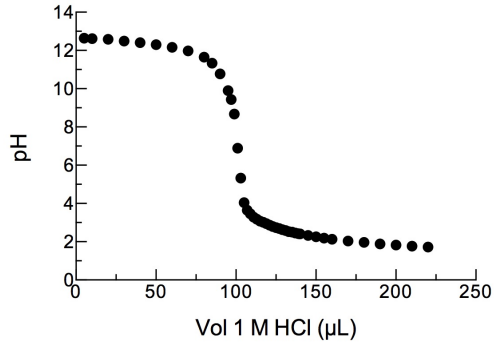
3.5.23 pH-Titration Curves for Biheteroaryl Ligands

Water Blank



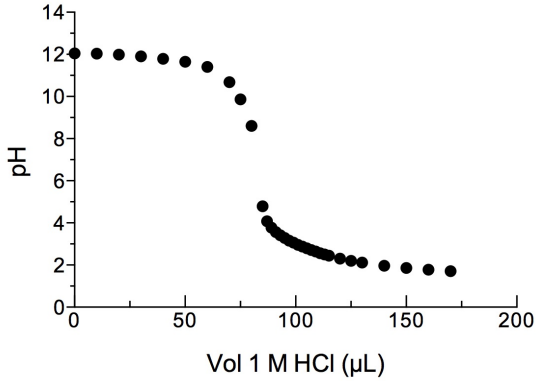
Pypyrid

pK_a <3



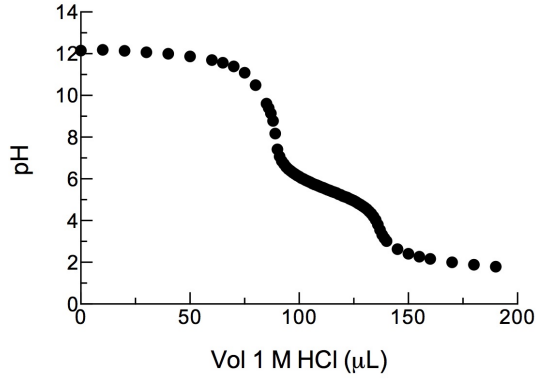
Pypyraz

pK_a <3



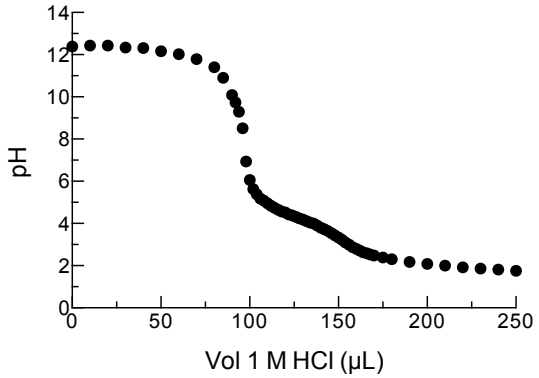
Pyim

pK_a 5.4



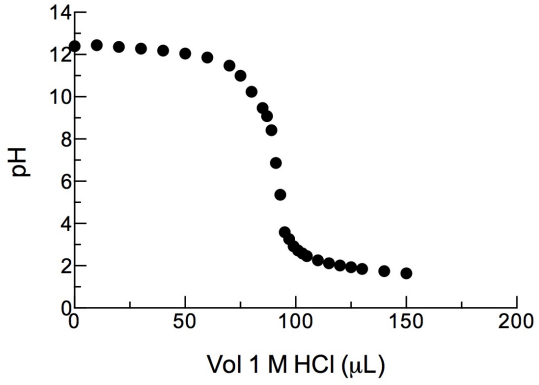
Pypyr

pK_a 4.1



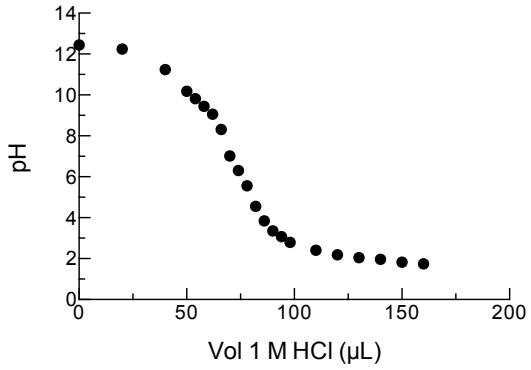
Pythi

pK_a <3



Pyox

pK_a <3



3.6 Acknowledgments

This work was supported by Grant R01 AR044276 to R. T. R. (NIH) and Grant R01 DK066600 to R. S. E. (NIH). J. D. V. was supported by Molecular Biosciences Training Grant T32 GM007215 (NIH). K. A. A was supported by a PhRMA predoctoral fellowship and by Molecular and Cellular Pharmacology Training Grant T32 GM008688 (NIH). The Micromass LCT® mass spectrometer was purchased with funds from Grant CHE-9974839 (NSF). This study made use of the National Magnetic Resonance Facility at Madison, which is supported by Grant P41 GM103399 (NIH).

3.7 Author Contributions

This work was conducted in collaboration with professor R.S. Eisenstein. J.D. Vasta carried out or contributed to carrying out all experiments. K.A. Andersen contributed to carrying out all experiments relating to human cell culture. K.M. Deck, C.P. Nizzi, and R. S. Eisenstein contributed to carrying out experiments relating to the function of IRPs. All authors contributed to the design of experiments, analysis of data, and preparation of the manuscript and figures.

Table 3.1 Iron binding by bipy and biheteroaryl compounds

Compound	Heteroaryl pK_a[a] (ref.)	pK_a[b]	Fe(II) Complex λ_{max}[c]	Stoichiometry (Ligand:Fe)[d]	Fe_{20}-EC_{50}[e]
bipy	5.2 ([162])	4.3	523	3:1	43 ± 2
pypyrid	2.3 ([163])	<3	522	3:1	124 ± 1
pypyraz	0.8 ([164])	<3	523	3:1	400 ± 7
pyim	7.1 ([165])	5.4	484	3:1	900 ± 30
pypyr	2.5 ([166])	4.1	441	ND[f]	2250 ± 100
pythi	2.5 ([167])	<3	524	3:1	5100 ± 100
pyox	0.8 ([168])	<3	NO[g]	NO[g]	>18,000

[a]pK_a value of the conjugate acid of pyridine (bipy) or the non-pyridyl ring.
[b]pK_a value of the conjugate acid as estimated by titration of the entire compound; bipy, ref. [169].
[c]Determined using spectrophotometry.
[d]Determined by Job's method.[143,144]
[e]Determined by titration of 20 μM $FeSO_4$; bipy, ref. [148]. Values are the mean (±SE) of 3 replicates.
[f]Not determined due to prohibitive competing absorbance from iron.
[g]Not observed.

Table 3.2 Estimated log*P* and topological polar surface area (TPSA) values

Compound	log*P*[a]	TPSA (Å^2)[a]
bipy	1.44	25.8
bipy55′DC	0.56	100.4
diethyl bipy55′DC	1.87	78.4
pypyrid	0.74	38.7
pypyraz	0.76	38.7
pyim	0.90	41.6
pypyr	0.85	41.6
pythi	1.64	25.8
pyox	1.00	38.9
pyimDC	0.29	116.2
diethyl pyimDC	1.56	94.2
pythiDC	1.03	100.4
diethyl pythiDC	2.30	78.4

[a]Values were estimated with the molecular properties calculator from Molinspiration (http://www.molinspiration.com).

Figure 3.1 Reaction catalyzed by collagen prolyl 4-hydroxylase and its inhibition

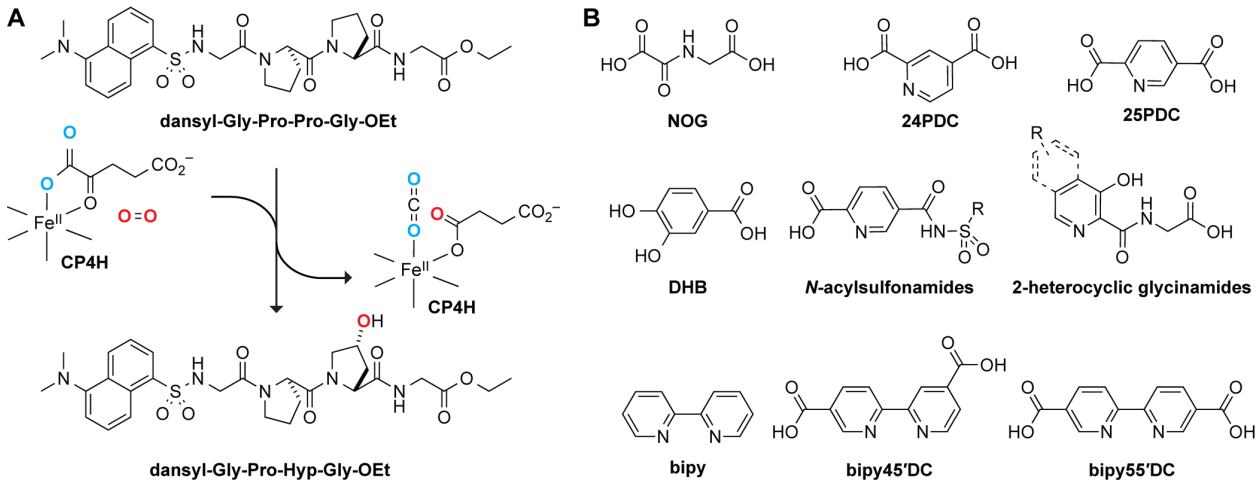


Figure 3.1 Reaction catalyzed by collagen prolyl 4-hydroxylase (CP4H) and its inhibition. (A) In an Fe(II)- and AKG-dependent reaction, CP4Hs catalyze the hydroxylation of Pro residues in collagenous peptides to form Hyp residues. (B) Examples of known human CP4H inhibitors.

Figure 3.2 Biheteroaryl compounds used in this study

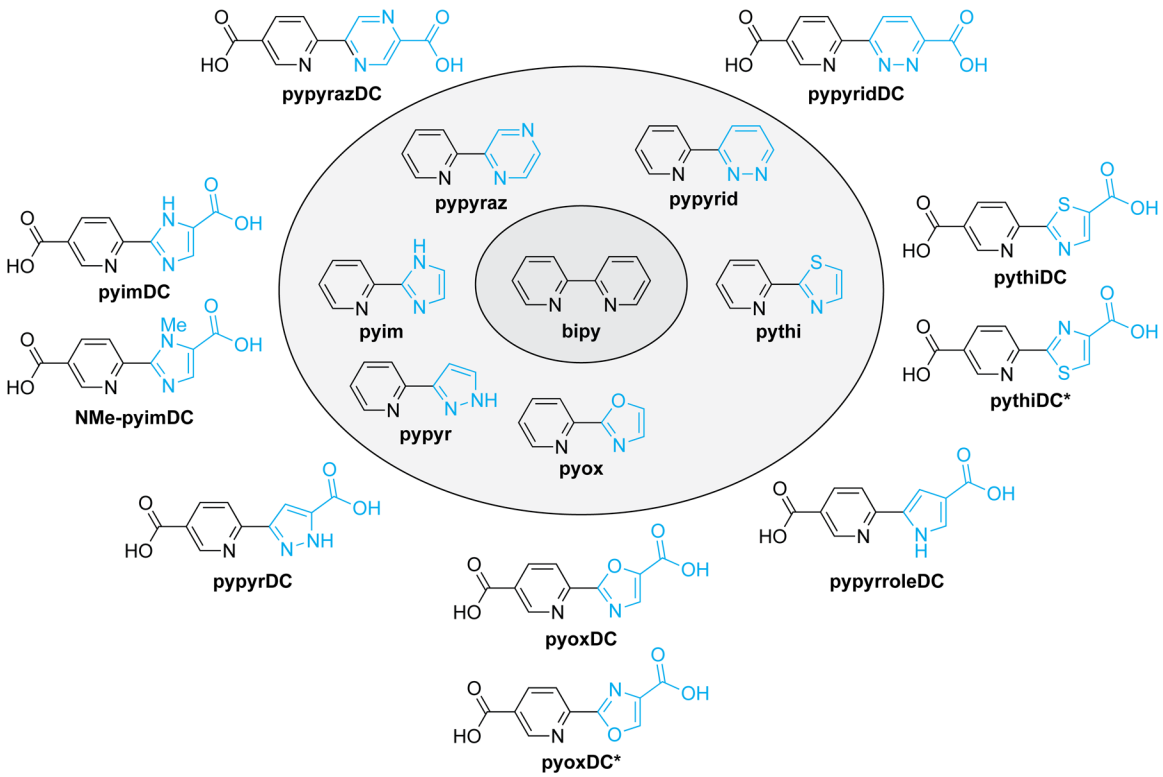


Figure 3.2 Biheteroaryl compounds used in this study. Novel moieties are rendered in cyan, and similar compounds are grouped together.

Figure 3.3 Ligand titrations to form iron complexes

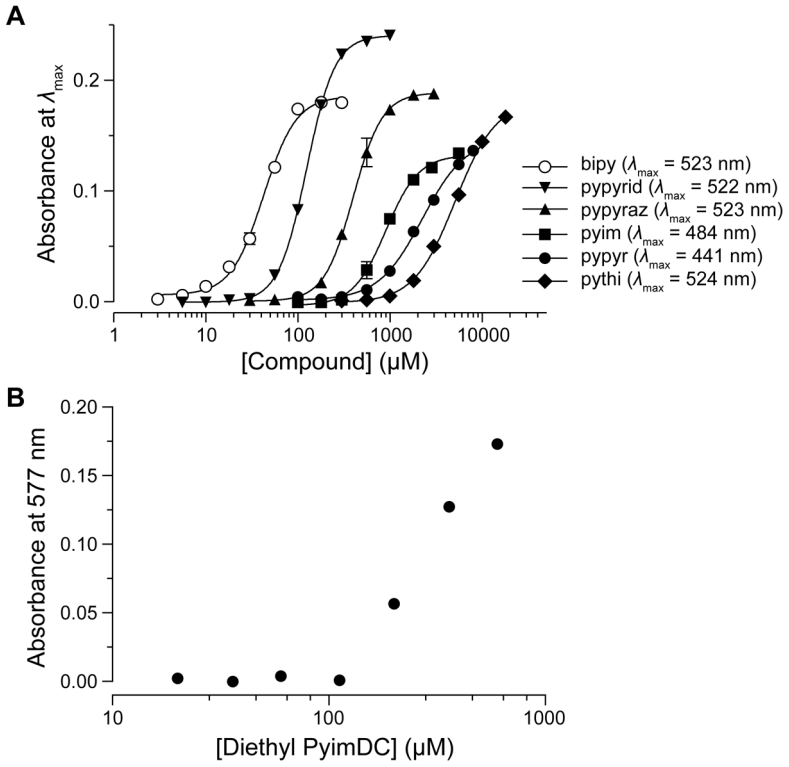


Figure 3.3 Ligand titrations to form iron complexes. Data were acquired as described in the section 3.5.8. (A) Biheteroaryl compounds. $R^2 > 0.99$ for all curve fits. The titration curve for bipy in panel A was reproduced from ref. [148]. (B) Diethyl pyimDC. Fe(II) (20 μM) could not be saturated over the range that diethyl pyimDC remained soluble under the assay conditions. Thus, Fe_{20}-EC_{50} >600 μM, which is the highest tested concentration.

Figure 3.4 Screen for inhibition of human CP4H1

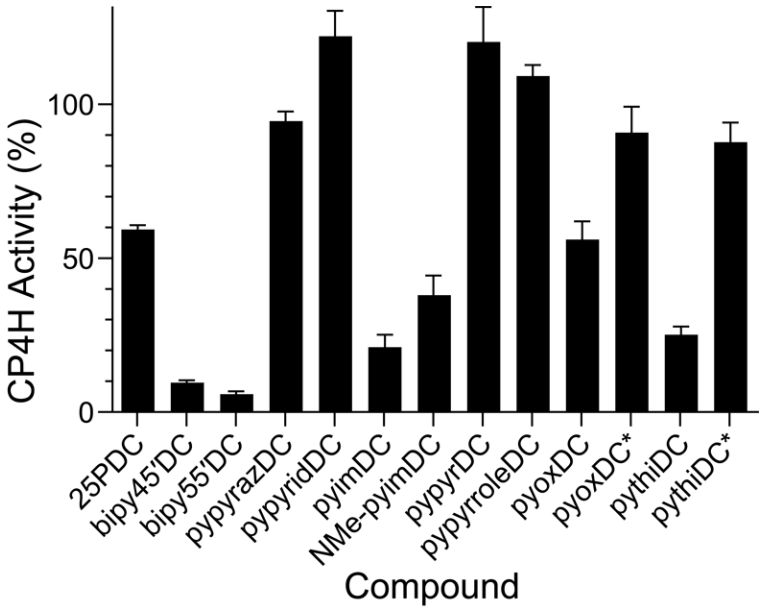
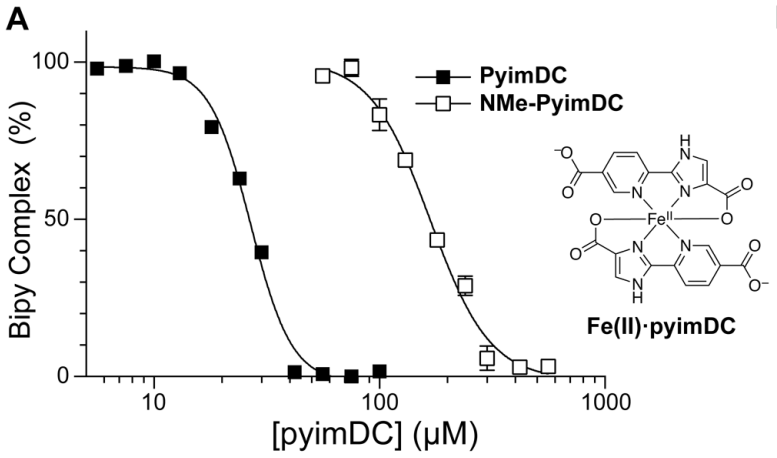
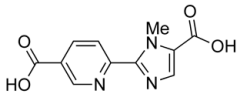
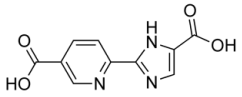
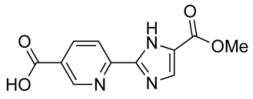


Figure 3.4 Screen for inhibition of human CP4H1. Compounds (10 μM) were screened for inhibition of the catalytic activity of human CP4H1 as described in the section 3.5.4. Relative activity values are the mean (±SD) of three replicates. Data for 25PDC, bipy45′DC, and bipy55′DC are from ref. [148].

Figure 3.5 Competitive iron-binding by pyimDC and ester derivatives



B

Compound	EC_{50} (μM)[a]
 **NMe-PyimDC**	168 ± 7
 **PyimDC**	27 ± 1
	>100

[a]Determined by competitive titration. Values are the mean (±SE) of three replicates.

Figure 3.5 Competitive Iron-Binding by PyimDC and Ester Derivatives. High concentrations of pyimDC can compete with bipy for complexation with Fe(II), as determined with the assay described in section 3.5.8. This phenomenon was dose-dependent (A) and confined to the diacid (B). Together, these data are consistent with pyimDC forming the depicted 2:1 complex with iron.

Figure 3.6 Biheteroaryl dicarboxylates as inhibitors of human CP4H1

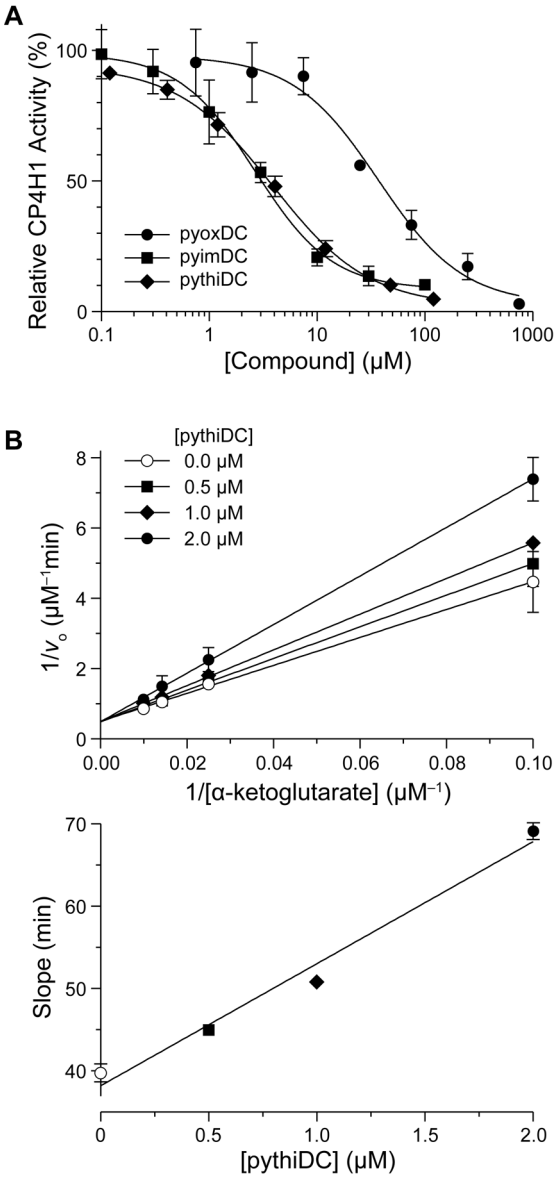


Figure 3.6 Biheteroaryl dicarboxylates as inhibitors of human CP4H1. (A) Individual points represent the mean (±SD) of three independent experiments. Data were fitted to a dose–response equation to determine IC_{50} values: pyimDC, (2.6 ± 0.1) μM; pyoxDC, (33 ± 8) μM; pythiDC, (4.0 ± 0.2) μM. (B) Lineweaver–Burke analysis of inhibition by pythiDC. The rate of the reaction catalyzed by CP4H1 with increasing α-ketoglutarate concentration (10–100 μM) was determined in the presence of a fixed concentration of pythiDC (0.0, 0.5, 1.0, or 2.0 μM). Individual points represent the mean (±SE) of two independent experiments. Data were fitted by linear regression to determine a K_i value of (0.39 ± 0.04) μM for competitive inhibition.

Figure 3.7 Screen for Inhibition of Human PHD2

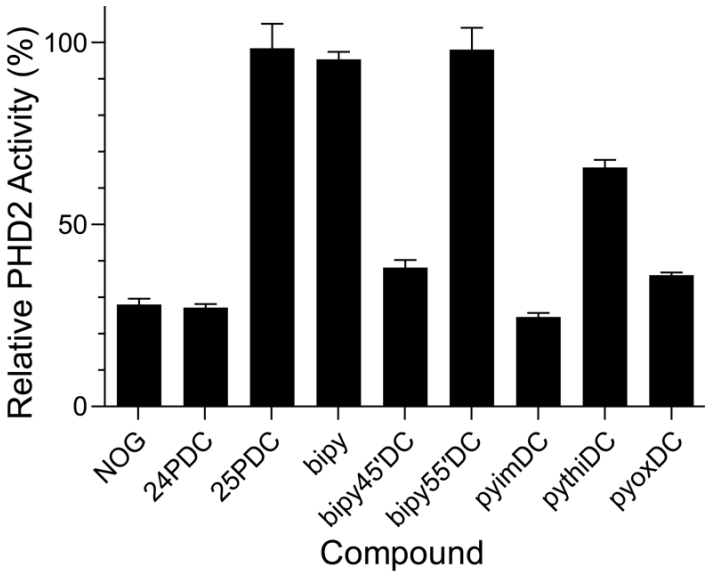


Figure 3.7 Screen for Inhibition of Human PHD2. Compounds (10 μM) were screened for inhibition of the catalytic activity of human PHD2 as described in the section 3.5.6. Relative activity values are the mean (±SD) of three replicates. Data for NOG, 24PDC, 25PDC, bipy, bipy45′DC, and bipy55′DC are from ref. 148.

Figure 3.8 Effect of esterified compounds on iron metabolism and collagen secretion

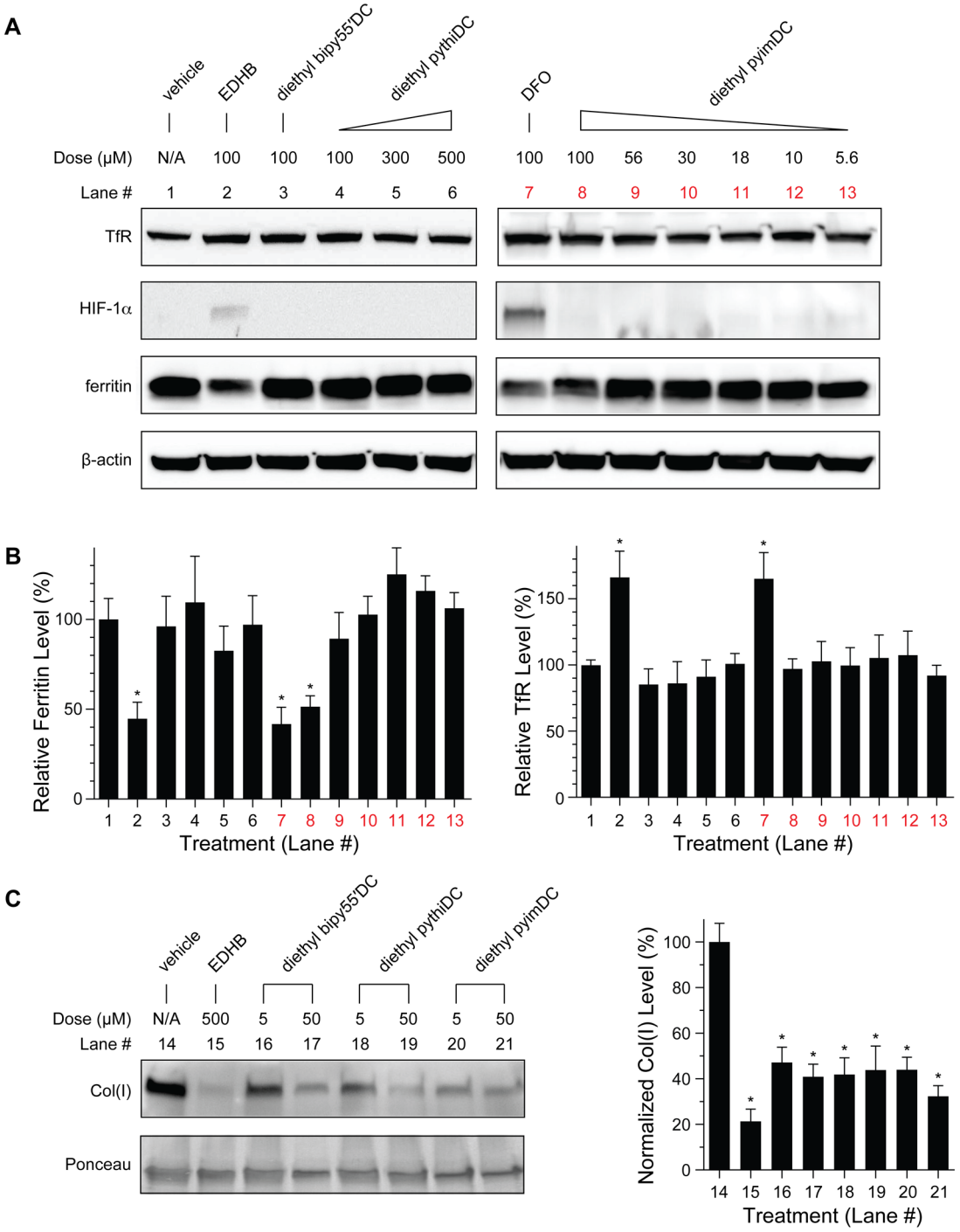


Figure 3.8 Effect of esterified compounds on iron metabolism and collagen secretion. MDA-MB-231 breast cancer cells were treated with deferoxamine (DFO), ethyl dihydroxybenzoate (EDHB), biheteroaryl compounds, or vehicle (DMSO) and then analyzed with an immunoblot (*, $p < 0.05$). (A) Effect of esterified biheteroaryl compounds on iron metabolism. Blots are representative of at least 3 replicates. (B) Densitometric quantitations ($n \geq 3$) corresponding to the blots in panel A and normalized to β-actin. The dose of EDHB (500 μM) is known to diminish collagen secretion significantly.[29] (C) Effect of esterified biheteroaryl compounds on collagen secretion into conditioned media. Blots are representative of at least 5 replicates and quantitations (right) are normalized to total protein using the Ponceau S stained blot.

Figure 3.9 Effect of biheteroaryl compounds on iron metabolism in breast cancer cells

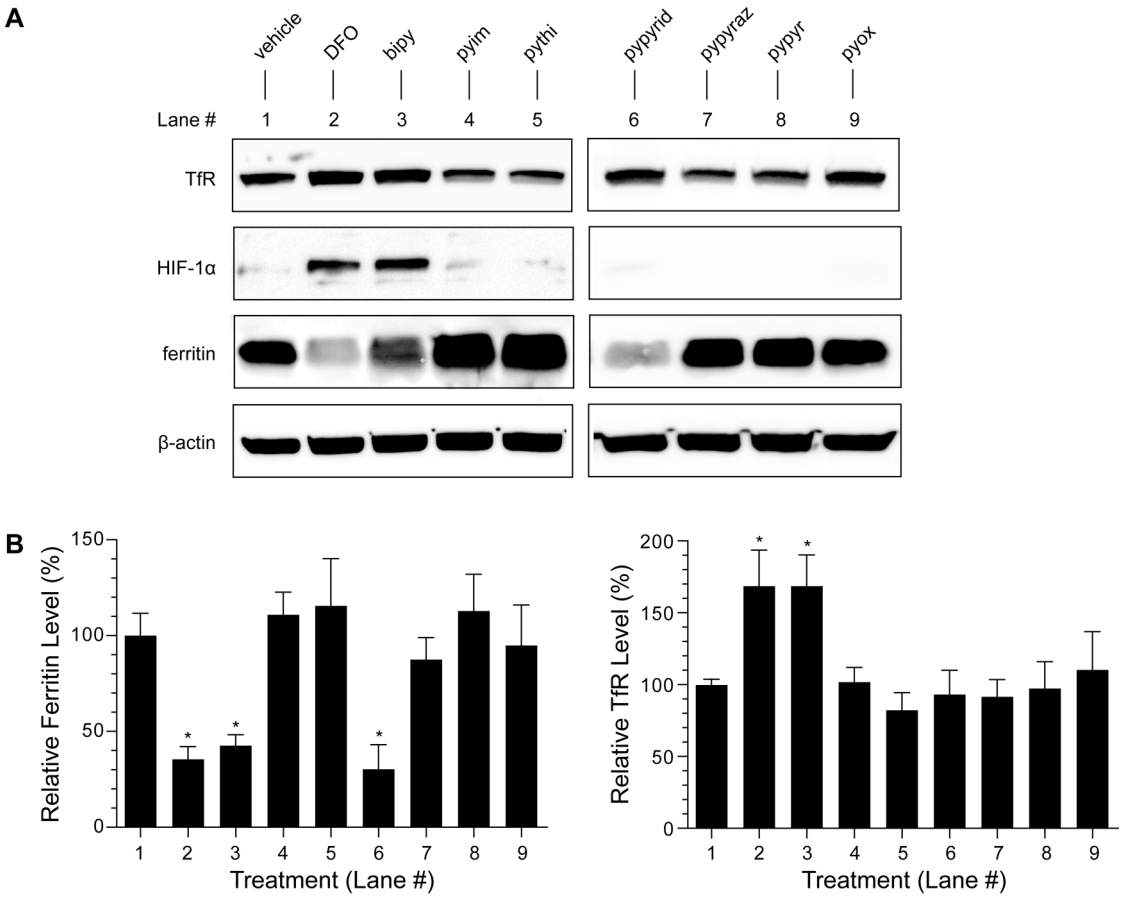


Figure 3.9 Effect of biheteroaryl compounds on iron metabolism in human breast cancer cells. (A) MDA-MB-231 breast cancer cells were treated with deferoxamine (DFO), biheteroaryl compounds, or vehicle (DMSO) as described in section 3.5.12, and analyzed with an immunoblot . (B) Densitometric quantitations ($n \geq 3$) corresponding to the blots in panel A and normalized to β-actin. (*, $p < 0.05$).

Figure 3.10 Effect of biheteroaryl compounds and their diethyl esters on iron metabolism in human embryonic kidney cells

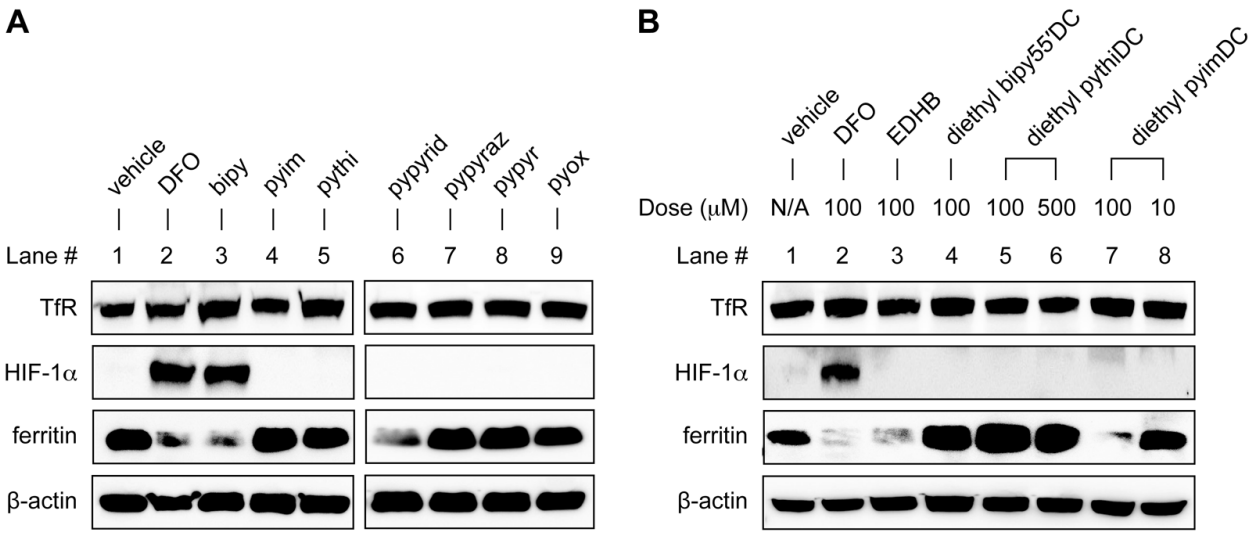


Figure 3.10 Effect of biheteroaryl compounds and their diethyl esters on iron metabolism in human embryonic kidney cells**.** (A) HEK293T cells were treated with biheteroaryl compounds (100 μM), deferoxamine (DFO, 100 μM), or vehicle (DMSO) as described in section 3.5.12, and analyzed with an immunoblot. (B) HEK293T cells were treated with esterified biheteroaryl compounds, deferoxamine (DFO), ethyl dihydroxybenzoate (EDHB), or vehicle (DMSO) as described in the section 3.5.12, and analyzed with an immunoblot.

Figure 3.11 Effect of esterified biheteroaryl compounds on IRE binding by IRPs in human breast cancer cells

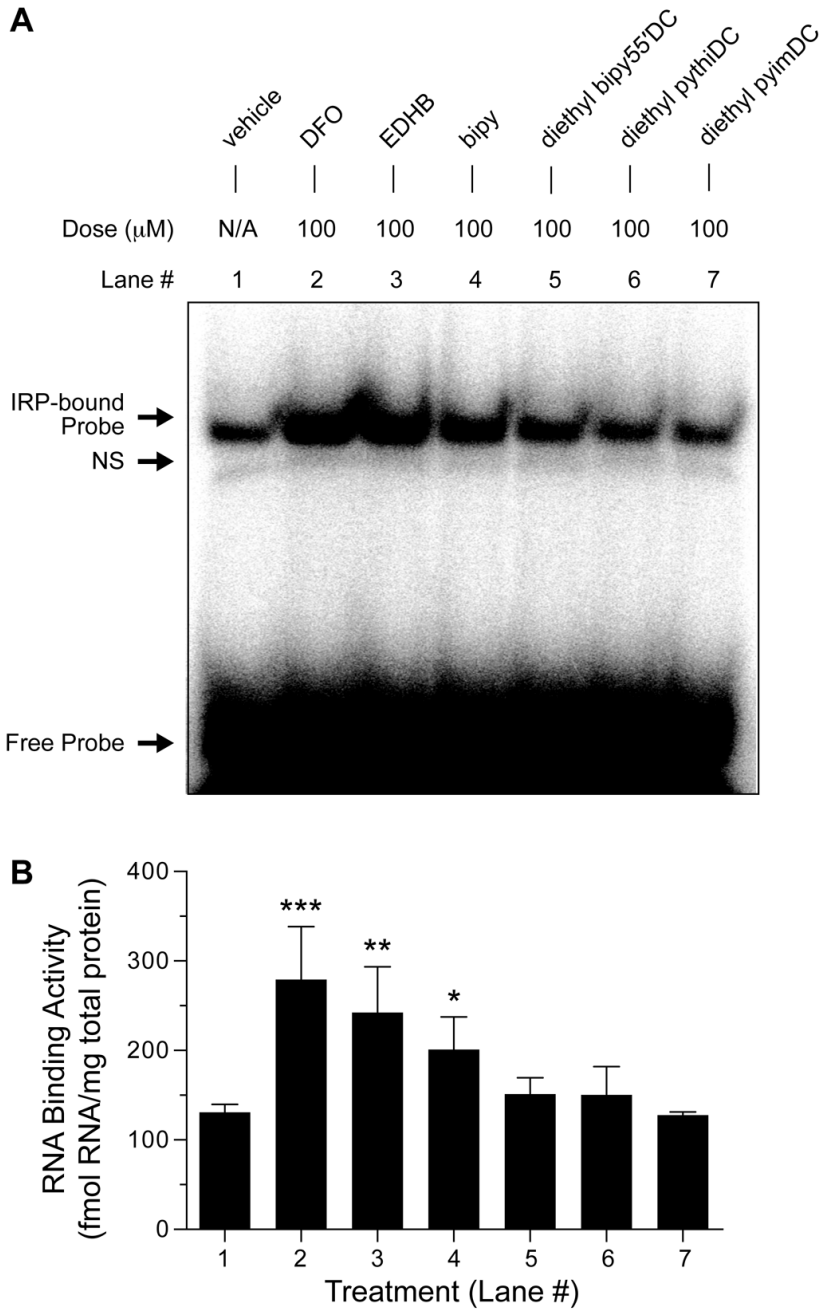


Figure 3.11 Effect of esterified biheteroaryl compounds on IRE binding by IRPs in human breast cancer cells. MDA-MB-231 cells were treated with esterified biheteroaryl compounds, deferoxamine (DFO), ethyl dihydroxybenzoate (EDHB), or vehicle (DMSO) as described in the section 3.5.14, and analyzed by an electrophoretic mobility shift assay (EMSA) using a ^{32}P-labeled RNA ligand for IRPs as described in section 3.5.15. (A) An EMSA representative of 3 replicates. (B) Quantitations of the EMSAs represented in Panel A. Quantitations (n = 3) were obtained by densitometry and derived from the digital light units using a calibration curve obtained by scintillation counting and normalized to total protein and total RNA. Under these conditions, IRP1- and IRP2-bound RNA comigrate in the band labeled "IRP-bound Probe". *p < 0.05, **p < 0.01, ***p < 0.001.

Figure 3.12 Effect of esterified biheteroaryl compounds on P4HA1 levels in human breast cancer cells

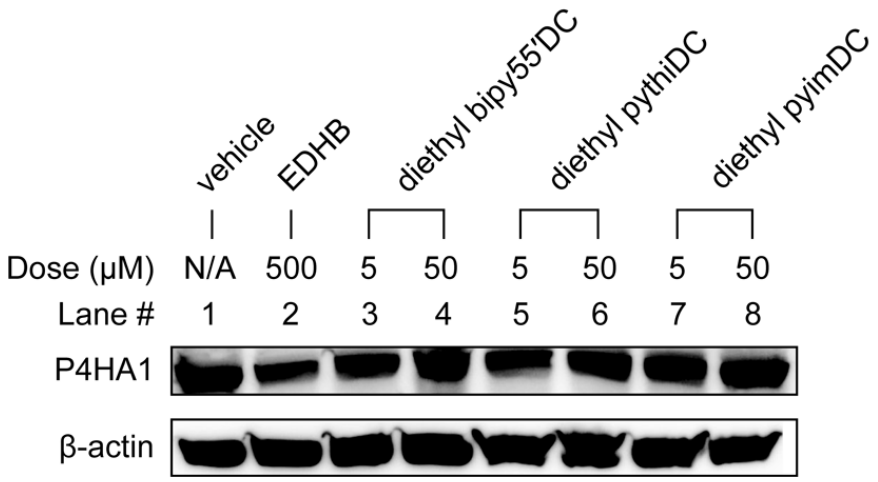
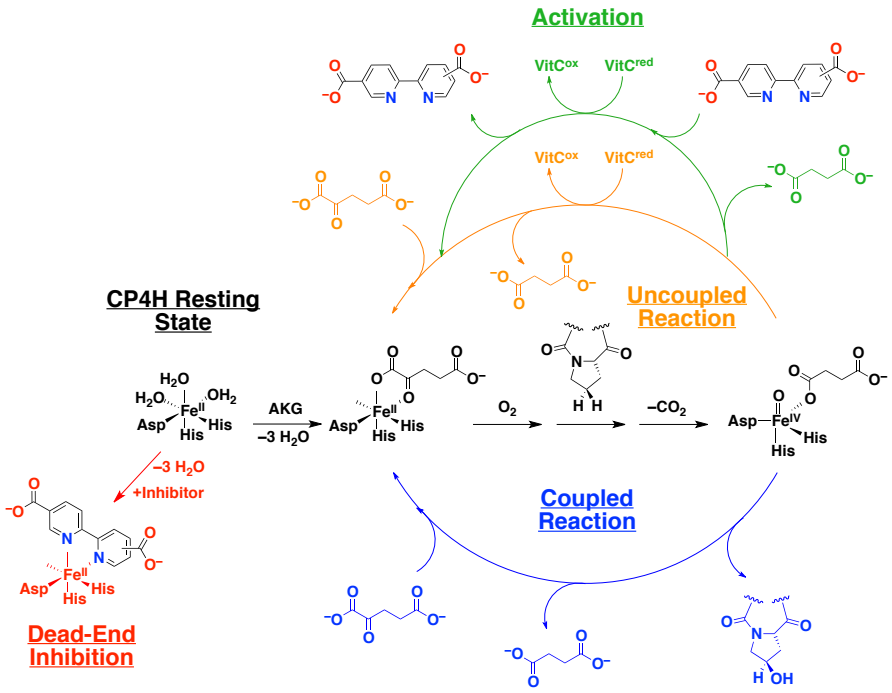


Figure 3.12 Effect of esterified biheteroaryl compounds on P4HA1 levels in human breast cancer cells. MDA-MB-231 cells were treated with esterified biheteroaryl compounds, deferoxamine (DFO), ethyl dihydroxybenzoate (EDHB), or vehicle (DMSO) as described in section 3.5.12, and then analyzed with an immunoblot. Blots are representative of at least 2 replicates.

CHAPTER 4

Bipyridinemonocarboxylates and Related Ligands are Novel Coactivators of Human Collagen Prolyl 4-Hydroxylase*



*This chapter has been prepared for publication in part under the same title. Vasta, J. D. and Raines, R. T. Bipyridinemonocarboxylates and related ligands are novel coactivators of human collagen prolyl 4-hydroxylase. J. D. Vasta designed and carried out all experiments. J. D. Vasta and R. T. Raines contributed to the analysis of data and preparation of the manuscript and figures.

4.1 Abstract

Collagen is one of the most abundant proteins in animals. The posttranslational hydroxylation of proline residues in collagen contributes greatly to its conformational stability, and deficiency in this hydroxylation is associated with a variety of disease states, including scurvy. The hydroxylation of proline residues in collagen is catalyzed by the Fe(II)- and α-ketoglutarate (AKG)-dependent dioxygenase, collagen prolyl 4-hydroxylase (CP4H). Although many inhibitors of CP4H enzymes are known, few compounds besides the peptide substrate and AKG cosubstrate have ever been shown to activate the enzyme. CP4H enzymes have long been known to suffer oxidative inactivation during catalysis, and the cofactor ascorbate (vitamin C) is specifically required to re-activate these enzymes via reduction of the enzymic iron center. Herein, we report the finding that two 2,2′-bipyridinemonocarboxylates serve as novel activators of human CP4H1. Specifically, we find that 2,2′-bipyridine-4-carboxylate (bipy4C) and 2,2′-bipyridine-5-carboxylate (bipy5C) are coactivators that enhance the rate of the ascorbate-dependent re-activation of the enzyme by serving as ligands for the iron center. Moreover, we find that this mode of activation is available to many previously reported biheteroaryldicarboxylate inhibitors of the enzyme under particular assay conditions. Lastly, we put forth a general mechanism for the modulation of CP4H activity by bipyridinecarboxylates and related analogues, providing a framework for further analyses of the kinetic mechanism of CP4Hs and related enzymes.

4.2 Introduction

Much research (see Chapters 1–3 above) has focused on developing selective inhibitors of collagen P4Hs to address the diseases in which collagen is overproduced including the fibrotic diseases and breast cancer metastasis. Still, the loss of CP4H function during ascorbate (vitamin C, $VitC^{red}$) deficiency contributes to scurvy,[99,100] and the complete loss of CP4H activity has been found to be embryonic lethal in the nematode *Caenorhabditis elegans*[81,83] and mice.[82] Moreover, there are a few additional indications, including nonsyndromic high myopia (a leading cause of blindness),[170] and possibly certain neurological process,[171] where the increase in CP4H activity could be of benefit. To our knowledge, the only reported small-molecule activators of CP4Hs and (other FAKGD enzymes) are $VitC^{red}$ and DTT, both of which are thought to assist in the rescue of oxidatively inactivated CP4H molecules and neither of which is likely to be of clinical utility.

Along with our examination of 2,2′-bipyridinedicarboxylates (bipyDCs) as inhibitors of human CP4H1,[148] we also examined a series of bipyridinemonocarboxylates (bipymonoCs). To our surprise, two of these compounds, namely 2,2′-bipyridine-4-carboxylate (bipy4C) and 2,2′-bipyridine-5-carboxylate (bipy5C), caused modest activation of CP4H1 *in vitro*. Activation by these bipymonoCs was unexpected, especially because bipy5C was previously reported as a weak inhibitor of a related CP4H from chicken.[106] Herein, we report our investigations of these bipymonoCs and related analogues. We find that these bipymonoCs serve as novel coactivators of CP4H1 and appear to enhance the rate of the $VitC^{red}$-dependent reduction of the iron center.

Moreover, we find that under the appropriate assay conditions, this mechanism of activation is also available to many known inhibitors of CP4H1, even some of the most potent biheteroaryldicarboxylates reported previously (See: CHAPTER 3).[106,148]

4.3 Results and Discussion

Discovery of BipymonoCs as Activators of CP4H1 and Requirement for Vitamin C

In our previous investigations of bipyridinedicarboxylates as inhibitors of human CP4H1, we also investigated bipy4C and bipy5C (Figure 4.1A), expecting that they would serve as modest inhibitors of the enzyme. In these investigations, we used a standard liquid chromatography-based assay in which we determined the hydroxylation of the model tetrapeptide substrate Dansyl-Gly-Pro-Pro-Gly-OEt by UPLC.[140,148] To our surprise, we found that standard reactions in which bipy5C or bipy4C were added at a concentration of 10 μM produced significantly more hydroxylated product compared to reactions to which vehicle was added. In subsequent investigations, we found that this activation was dose-dependent (Figure 4.1B) and not outcompeted by the cosubstrate α-ketoglutarate (AKG). Eventually, inhibition was observed for both bipy4C and bipy5C at concentrations in the low millimolar range. Moreover, neither bipy4C nor bipy5C appeared to completely replace other required assay components such as AKG or $VitC^{red}$. While the levels of many assay components did not appear to perturb the activating affect caused by bipy4C and bipy5C, we did find that activation required the $VitC^{red}$ cofactor in a dose-dependent manner (Figure 4.1C).

The finding that activation of CP4H1 by bipy4C and bipy5C required $VitC^{red}$ was fascinating, as such a phenomenon has not been reported previously for CP4Hs or for any member of the Fe(II)- and AKG-dependent dioxygenase (FAKGD) family. Thus, we considered the kinetic mechanism of CP4Hs (Figure 4.2) in order to identify a putative mechanism of activation. CP4Hs have previously been reported to display an ordered ter-ter mechanism in which AKG first binds to CP4H·Fe(II) complex, after which O_2 and the peptide substrate bind in an ordered fashion.[98] Subsequently, the oxidative decarboxylation of AKG facilitates the formation of a highly reactive Fe(IV)=O species (ferryl ion) that hydroxylates the peptide substrate via a radical rebound process.[56] Whereas all CP4Hs appear to require $VitC^{red}$ for activity, this cofactor is not required during the typical ter-ter mechanism, but rather, has been shown to operate during the "uncoupled reaction" in which CP4H decarboxylates AKG in the absence of the peptide substrate.[102] As suggested by EPR experiments, the ferryl ion appears to decay via an unknown mechanism to an inactive Fe(III) state,[101] and $VitC^{red}$ is required to reduce the iron center back to the Fe(II) state required for catalysis.[101,102] The finding that $VitC^{red}$ is required for activation by bipy4C and bipy5C suggests that these activators may operate during the uncoupled reaction to assist $VitC^{red}$ in the rescue of oxidatively inactivated CP4H molecules.

A review of the literature suggested that diimine ligands such as 2,2′-bipyridine (bipy) and 1,10-phenanthroline serve as ligand catalysts for the $VitC^{red}$-dependent reduction of Fe(III) to Fe(II) in solution.[172,173] While bipy itself does not serve as an activator of CP4H1,[106,148] we reasoned that bipy4C and bipy5C (resembling the cosubstrate AKG) could be targeted to the CP4H active site via their pendant carboxyl group and serve as ligands to enhance the rate of the

$VitC^{red}$-dependent reduction of iron center. However, as this mechanism has never been reported for a FAKGD enzyme, we first set out to explore the chemical features of these bipymonoCs that are required for activation.

Chemical Determinants for the Activation of CP4H1 by BipymonoCs

According to the above mechanism, bipy4C and bipy5C should serve as ligands for iron and their chelate and carboxylate should be required for activation. First, we chose to investigate the iron-binding properties of bipy4C and bipy5C compared to that of the parent bipy. We performed titration experiments to determine the half-maximal concentration required to form a complex with 20 μM Fe(II) (Fe_{20}-EC_{50}) at pH 7.0, taking advantage of the strong absorption typically observed for complexes of bipy and related analogues with Fe(II). Along with an estimation of stoichiometry obtained with Job's method, the Fe_{20}-EC_{50} value provides a comparative metric for iron affinity.[148] It was previously reported that bipy formed the $Fe(bipy)_3^{2+}$ complex with an Fe_{20}-EC_{50} of 43 ± 2 μM.[148] Like bipy, bipy4C and bipy5C served as ligands for Fe(II), forming distinct red $Fe(ligand)_3$ type complexes (Figure 4.3) with Fe_{20}-EC_{50} values of 33 ± 0.8 μM and 43 ± 2 μM, respectively. Thus, these data demonstrate that bipy4C and bipy5C are capable of serving as ligands for iron in a manner that is similar to their parent bipy.

To determine if their chelate and carboxylate are required for activation, we explored the structure-activity relationships (SARs) of a library of chemical analogues (Figure 4.4A) in which these chemical features were perturbed (Figure 4.4). We found that in either scaffold,

replacement of one or both of the pyridine rings with a phenyl ring (as in 4pyBA, 6Phpy3C, biphenyl4C, 3pyBA, 2Phpy4C, and biphenyl3C in Figure 4.4A) abolished activation. Moreover, repositioning of the pyridyl nitrogen in bipy4C to a non-chelating geometry (see bipy4C* in Figure 4.4A) also abolished activation. Lastly, we found that replacement of the carboxylate of bipy4C with the corresponding methyl ester (see methyl bipy4C in Figure 4.4A) abolished activation. To distinguish between the importance of a carboxylate in the appropriate geometry and the importance of an overall negative charge, we also tested a series of pyridine carboxylates (see py2C, py3C, and py4C in Figure 4.4A) and found that none of these compounds served as activators for CP4H1. Collectively, these SARs confirm the importance of both the bipyridyl chelate and a pendant carboxylate group with appropriate regiochemistry for the activation of CP4H1.

Distinguishing Between Activation by the Free Ligand and the Fe(Ligand)$_3$ Complex

The mechanism proposed above requires that a bipy4C or bipy5C molecule be targeted to the CP4H1 active site to serve as a ligand to enhance the rate of the VitC^{red}-dependent reduction of the iron center. However, the parent bipy and the related 1,10-phenanthroline, have previously been shown to enhance the rate of the VitC^{red}-dependent reduction of Fe(III) to Fe(II) in solution.[172,173] Given the finding that bipy4C and bipy5C could form strong complexes with iron, we next questioned whether it was the free ligand or an Fe(ligand)$_3$ type complex that was involved in the activation. To address this question, we performed preincubation experiments using bipy4C as a model compound. In these experiments, bipy4C (10 μM) was preincubated

with Fe(II) prior to addition to the assay mixture, and the activation was compared to that observed when bipy4C was added last and in the absence of a preincubation step. We found that regardless of the amount of free Fe(II) added to the assay, preincubation of bipy4C with the Fe(II) abolished activation (Figure 4.5). Moreover, at the lowest iron concentration tested (3 μM), we found that bipy4C appeared to inhibit the enzyme, suggesting that the preformed $Fe(ligand)_3$ type complex could sequester iron from CP4H1 over the timecourse of our assay.

These findings suggested that the free ligand is involved in activation of CP4H1, rather than a complex of the ligand with free iron. However, the mechanism proposed above also requires that bipy4C and bipy5C bind in the CP4H1 active site. Thus, we postulated that ligands that activate CP4H1 through this mechanism should also display SARs similar to inhibitors that are though to bind in the active site. The bipyridinedicarboxylates (bipyDCs) bipy-4,5′-dicarboxylate (bipy45′DC) and bipy-5,5′-dicarboxylate (bipy55′DC) are potent inhibitors of human CP4H1 and thought to inhibit the enzyme by binding to the CP4H1 iron center competitively with the AKG cosubstrate. Moreover, we recently reported SARs for the replacement of the bipy core in these molecules with other chelating biheteroaryls (see: CHAPTER 3) and found that some but not all biheteroaryls could replace the bipy core. For example, the dicarboxylate analogue in which the bipy core was replaced with 2-(pyridin-2-yl)thiazole (pythi) served as a potent competitive inhibitor (pythiDC) of CP4H1, while the analogue in which the bipy core was replaced with 3-(pyridine-2-yl)-1*H*-pyrazole (pypyr) showed no inhibition of the enzyme.

Given the SARs for the biheteroaryl dicarboxylates described above, we reasoned that replacement of the bipy core of bipy4C and bipy5C with either pythi or pypyr should also show similar SARs for the activation of CP4H1 if these activators are interacting with the CP4H1 iron center. To address this question, we tested the corresponding monocarboxylates pythi5C and pypyr5C, as well as the regioisomer pythi4C (see Figure 4.4A) as activators of CP4H1 under standard assay conditions (Figure 4.4B). In agreement with our hypothesis, we found that pythi5C served as an activator of CP4H1, while pypyr5C did not. Moreover, the regioisomer pythi4C showed no activation, confirming that correct regiochemistry is required for activation.

A Chemical Mechanism for Activation of CP4H1 by BipymonoCs and Related Ligands

Based upon the cumulative data presented above, we put forth the mechanism described in Figure 4.6 for the activation of CP4H1 by bipymonoCs. During the uncoupled decarboxylation of the cosubstrate AKG, the CP4H1 iron center becomes trapped in an unproductive Fe(III) state,[101] which normally requires reduction by the $VitC^{red}$ cofactor to return it to the catalytically active Fe(II) state.[101,102] We postulate that bipy4C and bipy5C can bind to the unproductive Fe(III) center, displacing any occupying ligands such as water, hydroxide, or the succinate byproduct. Once bound, these ligands serve to enhance the rate of the $VitC^{red}$-dependent reduction of the iron center back to a catalytically active state, after which they dissociate to begin another catalytic cycle. Moreover, this mechanism implies that the $VitC^{red}$-dependent reduction of the CP4H1 iron center is the rate-limiting step in the overall reaction catalyzed by CP4H1.

In previous reports, it was found that CP4H can catalyze its reaction at a high initial rate in the presence of Fe^{2+}, AKG, and the peptide substrate but without the $VitC^{red}$ cofactor.[101] However, that rate gradually decreases within 30s due to the oxidation of the enzyme during the sporadic uncoupled decarboxylation of AKG.[101] Subsequent addition of the $VitC^{red}$ cofactor allowed catalysis to continue at a rate approximately 35% of that observed initially,[101] suggesting that a step in the pathway for the uncoupled reaction (Figure 4.6) becomes rate-limiting for overall catalysis. Our finding that, at best, bipy4C and bipy5C can activate CP4H1 approximately 2.5 to 3-fold under standard conditions is consistent with the report above, and further supports our mechanism and the notion that the $VitC^{red}$-dependent reduction of the CP4H1 iron center is the rate-limiting step in the overall reaction catalyzed by CP4H1.

Based upon this mechanism, it is also expected that activating ligands would be competitive with the succinate byproduct. Thus, as a test of the mechanism, we explored the affect of added succinate on the activation observed for bipy4C and bipy5C. As succinate is also a known product inhibitor that competes with the AKG cosubstrate, we performed these reactions at high and saturating concentrations of AKG (≥ 1 mM) so as to prevent the added succinate from inhibiting the enzyme. Moreover, comparisons were always made to control reactions lacking the activator so as to dissect out any effects associated with adding succinate itself. We found that succinate (1 or 5 mM) was able to outcompete the activating effect of the bipymonoCs (100 μM) in a dose-dependent manner (Figure 4.7), providing further support for the mechanism.

Reconciling Activation and Inhibition of CP4H1 by Biheteroarylcarboxylate Ligands

In considering the mechanism presented in Figure 4.6, we were curious as to whether or not the potent inhibitors bipy45′DC, bipy55′DC and related dicarboxylates could also serve as activators of the enzyme. Interestingly, we found that at high concentrations of AKG, succinate itself appeared to activate the enzyme modestly (Figure 4.8A), a phenomenon that, to our knowledge, had never been reported. This finding was intriguing, and suggested to us that this mode of activation might be available to other inhibitors of CP4H1, so long as their mode of inhibition was prevented by adding high levels of AKG. We probed a variety of compounds (10 μM) including the modest inhibitor 2,5-pyridinedicarboxylate (25PDC), as well as the potent inhibitors bipy45′DC, bipy55′DC, and pythiDC. While the effect of adding bipy itself (Figure 4.8B) was not dependent on the amount of AKG added, we found that 25PDC (Figure 4.8C) displayed modest inhibition at lower concentrations of AKG (100 μM) and apparent activation in the presence of 1 mM AKG. Moreover, succinate was able to disrupt the activation effect (Figure 4.8C), consistent with the mechanism described above (Figure 4.6). In the case of the more potent compounds bipy55′DC (Figure 4.8D), bipy45′DC (Figure 4.8E), and pythiDC (Figure 4.8F), we found that their overall affect on CP4H activity also depended on the amount of AKG added, where they served as potent inhibitors at low AKG (100 μM) concentrations, but apparent activators at very high AKG concentrations up to 10 mM. Again, succinate was able to disrupt the activation effects, consistent with the mechanism proposed in Figure 4.6. Lasty, high concentrations of the AKG cosubstrate did not appear to provide any additional activating effect (Figure 4.8A), suggesting that AKG itself does not bind the Fe(III) species. This finding provides

an explanation as to why the activating effect observed for biheteroarylcarboxylate ligands is not outcompeted by the cosubstrate AKG despite their common interaction in the AKG-binding pocket.

Toward a General Mechanism for the Modulation of CP4H1 by Biheteroarylcarboxylates

Based upon these additional data above, we put forth a general mechanism for the modulation of CP4H1 by biheteroarylcarboxylate ligands and related analogues (Figure 4.9). After the oxidative decarboxylation of AKG and formation of the ferryl ion, CP4H1 catalysis can proceed via two pathways. The primary pathway, or coupled reaction, leads to the productive hydroxylation of the peptide substrate and subsequent return of the enzymic iron to a catalytically active Fe(II) state. The secondary pathway, or uncoupled reaction, occurs in the absence or mispositioning of the peptide substrate, which leads to decay of the ferryl ion to an unproductive Fe(III) species. This Fe(III) species must be reduced by $VitC^{red}$ [101,102] in a step (or series of steps) that likely contains the rate-limiting step for overall catalysis.[101] Biheteroaryl carboxylate ligands and related analogues can modulate CP4H1 activity at multiple points in the catalytic mechanism, that is, by serving as competitive inhibitors with the AKG cosubstrate to form dead-end complexes, or by serving as ligands for the Fe(III) species that enhance the rate of the $VitC^{red}$-dependent reduction of the iron center back to a catalytically active Fe(II) state. Thus, the observed effect of treatment of CP4H1 with a biheteroaryl carboxylate ligand is actually the net effect of both activating and inhibiting processes, where the concentrations of the ligand, the AKG cosubstrate, the succinate byproduct, and the $VitC^{red}$ cofactor all influence that net effect.

Other P4H Enzymes

Lastly, we were curious as to whether this mechanism of activation was available to other P4H enzymes. There are a variety of other P4H enzymes of biological importance in animals, plants, and lower organisms,[4,30] but arguably the most well characterized are the P4Hs involved in hypoxia signaling (PHDs) and the P4H enzymes involved in cell wall biosynthesis in the algae *Chlamydomonas reinhardtii* (CrP4Hs). We chose to investigate these two alternative types of P4H enzymes, using the well characterized PHD isoform 2 (PHD2) and CrP4H isoform 1 (CrP4H-1) as model enzymes. After screening these enzymes under standard assay conditions comparable to those used to assay CP4H1 (Figure 4.10), we found that neither was activated by bipy4C or bipy5C at a modest concentration of 10 μM. Based upon our postulated mechanism for the activation of CP4H1 by these bipymonoCs, this finding suggests that either these activators do not bind appreciably in the active site of these P4Hs, or if they do, that the rate-limiting step in the overall kinetic mechanisms of these other P4Hs is not the $VitC^{red}$-dependent reduction of their enzymic iron center.

4.4 Conclusions

In summary, we have discovered that bipy4C and bipy5C are activators of human CP4H1 via a novel mechanism (Figure 4.6) wherein they serve as ligands that enhance the rate of the $VitC^{red}$-dependent rescue of the enzyme from a catalytically inactive oxidation state at its iron center. To

our knowledge, this report is the first of a small-molecule activator of a CP4H since the discovery that $VitC^{red}$ is a required cofactor for the enzyme. Our postulated mechanism for these activators also suggests that the rate-limiting step of the overall reaction catalyzed by CP4H1 (including both the coupled and uncoupled reaction pathways) is likely the $VitC^{red}$-dependent reduction of the inactive Fe(III) species. Lastly, we found that this mechanism of activation is also available to many potent inhibitors of CP4H1 that are competitive with the cosubstrate AKG. These inhibitors, including 25PDC, bipy45′DC, bipy55′DC, and pythiDC, can also serve as activators of the enzyme, if their inhibitory mechanism is prevented by saturating concentrations of the AKG cosubstrate. The finding that these bipyridinecarboxylate ligands can activate human CP4H1 *in vitro* provides new information about catalysis by this essential enzyme and along with chemical tools to aid in probing the CP4H kinetic mechanism. Further studies are necessary to determine if this mechanism is of significance for the function of CP4H in cells.

4.5 Experimental Procedures

4.5.1 General

Biphenyl-4-carboxylic acid, biphenyl-3-carboxylic acid, 3-(pyridin-2-yl)benzoic acid, 4-(pyridin-2-yl)benzoic acid, 2-phenylisonicotinic acid, 6-phenylnicotinic acid, 5-(pyridin-2-yl)-1*H*-pyrazole-3-carboxylic acid, and 2-(pyridin-2-yl)thiazole-4-carboxylic acid were obtained from Combi-Blocks (San Diego, CA). 2-(Pyridin-2-yl)thiazole-5-carboxylic acid was obtained from Enamine (Monmouth Jct., NJ). 2,2′-Bipyridine-5,5′-dicarboxylate was obtained from Sigma–Aldrich (St. Louis, MO). Phosphine ligands and phosphonium salts were obtained from either Sigma–Aldrich or Strem (Newberryport, MA), stored in a dessicator, and used without further purification. $Pd(OAc)_2$ was obtained from Sigma–Aldrich, stored in a dessicator, and used without further purification. All other reagent chemicals were obtained from commercial sources (Sigma–Aldrich, Acros, Combi-Blocks, Oakwood Products, Enamine, Bachem, or Novabiochem) and used without further purification. The HIF-1α peptide$_{556\text{-}575}$ was from AnaSpec (Fremont, CA) and used without further purification. All glassware was flame- or oven-dried, and reactions were performed under N_2(g) unless indicated otherwise. DCM and toluene were dried over a column of alumina. Dimethylformamide was dried over alumina and further purified through an isocyanate scrubbing column. Other anhydrous solvents were obtained in septum-sealed bottles. Flash chromatography was performed with columns of 40-63 Å silica gel, 230–400 mesh (Silicycle, Québec City, Canada). Thin-layer chromatography (TLC) was performed on plates of EMD 250 μm silica 60-F_{254} with visualization by UV light or

staining with $KMnO_4$. The phrase "concentrated under reduced pressure" refers to the removal of solvents and other volatile materials using a rotary evaporator at water aspirator pressure (<20 torr) while maintaining water-bath temperature below 40 °C. Residual solvent was removed from samples at high vacuum (<0.1 torr). The term "high vacuum" refers to vacuum achieved by a mechanical belt-drive oil pump. All reported yields are unoptimized.

4.5.2 Instrumentation

NMR spectra were acquired at ambient temperature with a Bruker DMX-400 Avance spectrometer or a Bruker Avance 500i spectrometer at the National Magnetic Resonance Facility at Madison (NMRFAM) and were referenced to TMS or a residual protic solvent. Some compounds exist as either mixtures of rotomers or tautomers that do not interconvert on the NMR timescale at ambient temperature and therefore exhibit multiple sets of NMR signals (as indicated). Electrospray ionization (ESI) and electron ionization (EI) mass spectrometry were performed with a Micromass LCT® or Micromass AutoSpec® instruments, respectively, from Waters (Milford, MA) at the Mass Spectrometry Facility in the Department of Chemistry at the University of Wisconsin–Madison. The progress of reactions catalyzed by prolyl 4-hydroxylases was determined by analytical HPLC (Waters system equipped with a Waters 996 photodiode array detector, and Empower 2 software) or analytical UPLC using an Acquity UPLC® H-Class from Waters (equipped with an Acquity photodiode array detector, Acquity quaternary solvent manager, Acquity sample manager with a flow-through needle, and Empower 3 software).

Preparative HPLC was performed using a Prominence HPLC instrument from Shimadzu (Kyoto, Japan) equipped with two LC-20AP pumps, a SPD-M20A photodiode array detector, and a CTO-20A column oven. Iron complexes with biheteroaryl ligands were analyzed by spectrophotometry using a Cary 60 UV–Vis Spectrometer from Agilent Technologies (Santa Clara, CA). Protein concentrations were calculated from their absorbance at 280 nm as measured with a NanoVue Plus spectrophotometer from GE Healthcare using an extinction coefficient of 290,000 $M^{-1}cm^{-1}$ for human CP4H,[140] 36,9000 $M^{-1}cm^{-1}$ for human PHD2,[141] and 44,000 $M^{-1}cm^{-1}$ for *Chlamydomonas reinhardtii* P4H (estimated using EXPASY assuming all cysteines oxidized). EC_{50}-values were calculated from experimental data with Prism version 6.0 from GraphPad Software (La Jolla, CA).

4.5.3 Production of Recombinant Human CP4H1

Human CP4H containing the α(I) isoform was produced heterologously in Origami B(DE3) *Escherichia coli* cells and purified as described previously.[140]

4.5.4 Assay of Human CP4H1 Activity in the Presence of Activators

The catalytic activity of human CP4H1 was assayed as described previously.[140] Briefly, activity assays were carried out at 30 °C in 100 μL Tris–HCl buffer, pH 7.8, containing human CP4H1 (100 nM), activator (0–5 mM), substrate (dansylGlyProProGlyOEt, 500 μM), $FeSO_4$ (50 μM), BSA (1 mg/mL), catalase (0.1 mg/mL), ascorbate (2 mM), DTT (100 μM), and α-ketoglutarate

(0.1-10 mM). Unless otherwise noted, reactions were prepared by adding concentrated stock solutions of each component to concentrated assay buffer in the following order: $FeSO_4$, DTT, ascorbate, BSA, catalase, CP4H1, peptide substrate, and activator (or vehicle). Solutions thus prepared were pre-incubated for 2 min at 30 °C, after which the reaction was initiated by the addition of α-ketoglutarate. After 15 min, reactions were quenched by boiling for 45 s and centrifuged at 10,000*g*. The supernatant (5–10 μL) was injected into an Acquity UPLC® BEH C18 column (2.1 × 50 mm, 1.7 μm particle size) from Waters. The column was eluted at 0.6 mL/min with a gradient (2.9 min) of 20%–68% aqueous acetonitrile containing 0.1% v/v TFA. The absorbance of the eluent was monitored at 289 nm. All assays were performed in triplicate. Data is reported as activity relative to control reactions lacking activator, where activity is determined from the percent conversion of substrate to product. Dose-response curves were generated for chosen activators by plotting the relative activity versus the log of the inhibitor concentration.

4.5.5 Production of Recombinant Human PHD2

A cDNA encoding human $PHD2_{181-426}$ possessing an N-terminal hexahistidine (His_6) tag ($NHis_6$-$PHD2_{181-426}$) was cloned using the Gibson strategy[142] and the encoded protein was produced and purified as described previously.[141]

4.5.6 Assay of Human PHD2 Activity in the Presence of Activators

The catalytic activity of human PHD2 was assayed as described previously.[141] Briefly, activity assays were carried out at 30 °C in 100 μL Tris–HCl buffer, pH 7.8, containing human $NHis_6$-$PHD2_{181–426}$ (5 μM), activator (0–10 μM), substrate (HIF-1α $peptide_{556–574}$, 50 μM), $FeSO_4$ (50 μM), BSA (1 mg/mL), catalase (0.3 mg/mL), ascorbate (2 mM), DTT (1 mM), and α-ketoglutarate (35 μM). Reactions were prepared by adding concentrated stock solutions of each component to concentrated assay buffer in the following order: $FeSO_4$, DTT, ascorbate, BSA, catalase, $NHis_6$-$PHD2_{181–426}$, peptide substrate, and activator (or vehicle). Solutions thus prepared were pre-incubated for 2 min at 30 °C, after which the reaction was initiated by the addition of α-ketoglutarate. After 10 minutes, reactions were quenched by boiling for 60 s and centrifuged at 10,000*g*. The supernatant (50 μL) was injected into a Nucleodur® C18 Gravity reversed-phase column (4.6 × 250 mm, 5 μm particle size) from Macherey-Nagel (Bethlehem, PA). The column was eluted at 1 mL/min with a gradient (34 min) of 5–56% aqueous acetonitrile containing 0.1% v/v TFA. The absorbance of the eluent was monitored at 218 nm. All assays were performed in triplicate. Data is reported as activity relative to control reactions lacking activator, where activity is determined from the percent conversion of substrate to product.

4.5.7 Production of Recombinant *Chlamydomonas reinhardtii* P4H-1

A cDNA encoding *C. reinhardtii* P4H-$1_{30\text{-}245}$ possessing an N-terminal hexahistidine (His_6) tag ($NHis_6$-CrP4H-1) under the control of the T7 promoter was cloned into a pET-22b(+) vector using the Gibson strategy.[142] The expression vector was transformed into Origami B(DE3) cells by electroporation, and transformants were grown on LB agar containing kanamycin (15 μg/mL), ampicillin (100 μg/mL), and tetracycline (12.5 μg/mL). A starter culture of TB medium containing antibiotics as described above was inoculated with a fresh colony after which the cells were grown overnight at 37 °C with shaking at 200 rpm. The starter culture was used to inoculate 4 L of TB medium supplemented with antibiotics as described above to an OD_{600} of 0.02. The culture was incubated at 37 °C with shaking at 200 rpm until an OD_{600} of 0.9 was reached. At this time, protein expression was induced by the addition of isopropyl-1-thio-β-D-galactopyranoside (IPTG, 500 μM), and the culture was grown for 18 h at 21°C with shaking at 200 rpm. Cells were harvested by centrifugation and the cell pellet (6.0 g) was resuspended in 30 mL lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 2.5 mM imidazole, pH 7.4). The cells were lysed at 22 kPSI in a T series Cell Disrupter 2.2 kW from Constant Systems Limited (Northants, UK). Insoluble material was cleared from the lysate by centrifugation at 30,000*g* for 45 min, and $NHis_6$-CrP4H-1 was purified from the soluble fraction by nickel affinity chromatography using a 5 mL HisTrap FF column (GE Healthcare, Piscataway, NJ). The column was equilibrated with lysis buffer, after which the supernant was injected and the A_{280} of the flowthrough was monitored until returning to baseline. $NHis_6$-CrP4H-1 was eluted from the column with an imidazole gradient (2.5–500 mM), with most of the $NHis_6$-CrP4H-1 eluting

between 50–250 mM imidazole. Fractions containing purified $NHis_6$-CrP4H-1 (as determined by 12% SDS-PAGE) were combined, dialyzed against 10 mM Tris–HCl, 100 mM glycine, 100 mM NaCl, pH 7.8, and concentrated to 3 mg/mL using a Vivaspin concentrator with a molecular weight cut off of 8,000 Da. Aliquots were flash frozen in liquid N_2 and stored at −80 °C until use in P4H activity assays. This protocol yielded $NHis_6$-CrP4H-1 at ~24 mg per gram of cell pellet.

4.5.8 Assay of *C. reinhardtii* P4H-1 Activity in the Presence of Activators

The catalytic activity of $NHis_6$-CrP4H-1 was assayed similarly to that described above in section 4.5.4 for human CP4H1 with a few modifications. Briefly, activity assays were carried out at 30 °C in 100 μL HEPES–HCl buffer (50 mM), pH 7.0, containing $NHis_6$-CrP4H-1 (1.2 μM), activator (0–10 μM), substrate (dansylGlyProProGlyOEt, 500 μM), $FeSO_4$ (200 μM), BSA (1 mg/mL), catalase (0.1 mg/mL), ascorbate (2 mM), DTT (100 μM), and α-ketoglutarate (150 μM). Reactions were prepared by adding concentrated stock solutions of each component to concentrated assay buffer in the following order: $FeSO_4$, DTT, ascorbate, BSA, catalase, $NHis_6$-CrP4H-1, peptide substrate, and activator (or vehicle). Solutions thus prepared were pre-incubated for 2 min at 30 °C, after which the reaction was initiated by the addition of α-ketoglutarate. After 1 h, reactions were quenched by the addition of EDTA (0.5 M, 1 μL) and boiled for 45 s. Quenched reactions were subjected to centrifugation at 10,000*g* for 5 min, after which the supernatant (5–10 μL) was injected into an Acquity UPLC® BEH C18 column (2.1 × 50 mm, 1.7 μm particle size) from Waters. The column was eluted at 0.6 mL/min with a gradient

(2.9 min) of 20%–68% aqueous acetonitrile containing 0.1% v/v TFA. The absorbance of the eluent was monitored at 289 nm. All assays were performed in triplicate. Data is reported as activity relative to control reactions lacking activator, where activity is determined from the percent conversion of substrate to product.

4.5.9 Assay of Fe(II)-Affinity for Biheteroaryl Ligands

The affinity of biheteroaryl ligands for Fe(II) was determined comparatively by measuring the half maximal concentration (EC_{50}) required for binding 20 μM Fe(II) (Fe_{20}-EC_{50}) in sodium phosphate buffer, pH 7. Stock solutions of ligands were prepared in water. Stock solutions of $FeSO_4$ were prepared in H_2O and used within 3 hours of preparation. Ligand solutions (3–560 μM depending on affinity) were prepared in 10 mM sodium phosphate buffer, pH 7, after which Fe(II) stock solution was added to initiate complex formation. Solutions were allowed to equilibrate for 15 min, after which the absorbance was recorded at the λ_{max} for the complex under study. Absorbance values were corrected by subtracting the absorbance value in the absence of ligand. Dose-response curves were generated for each ligand by plotting the absorbance versus the log of the ligand concentration. Fe_{20}-EC_{50} values for each ligand were interpolated from the dose-response curves by non-linear regression using the sigmoidal dose-response function in Prism. All experiments were performed in triplicate.

4.5.10 Determination of Fe(II) Complex Stoichiometry

The stoichiometry of biheteroaryl complexes with Fe(II) was estimated via Job's method. Briefly, reactions were prepared such that the total moles of Fe(II) and ligand was kept constant, but the mole fraction of the ligand was varied from 0 to 1. The total concentration of ligand and Fe(II) was 0.4 mM. Stock solutions of ligands were prepared in water. Stock solutions of $FeSO_4$ were prepared in water and used within 3 h of preparation. Reactions were prepared in 10 mM sodium phosphate buffer, pH 7, and complex formation was initiated by the addition of Fe(II) solution. Solutions were allowed to equilibrate for 15 min, after which the absorbance was recorded at the λ_{max} for the complex under study. Absorbance values were corrected by subtracting the absorbance value of a blank solution in absence of $FeSO_4$, after which the values were normalized relative to the reaction with the highest absorbance value. All experiments were performed in at least duplicate. Job's plots were constructed by plotting the normalized absorbance versus the mole fraction of ligand, after which the stoichiometry of the complex was estimated from the value on the abscissa of the point of intersection for the two best-fit lines approaching the point of maximum absorbance.

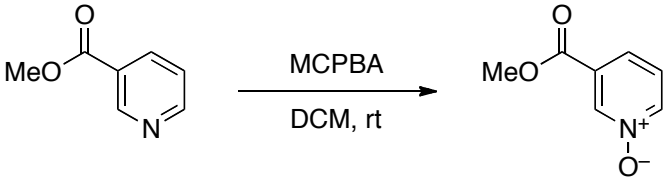
4.5.11 Determination of Compound Purity

All new final compounds were analyzed by HPLC to determine their purity. The analyses were performed on the Waters HPLC system using a Nucleodur® C18 Gravity reversed-phase column (4.6 × 250 mm, 5-μm particle size) from Macherey–Nagel. Samples (50 μL) dissolved in H_2O

were injected into the column and eluted at 1 mL/min with a gradient of aqueous acetonitrile (5–56% v/v over 34 min) containing 0.1% v/v TFA. The maximal absorbance in the range of 210–400 nm was used as the detection wavelength.

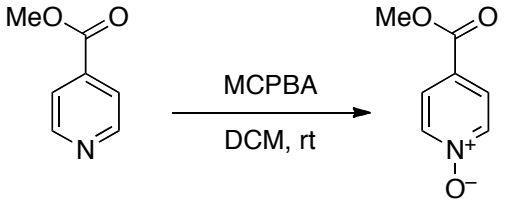
4.5.12 Synthetic Procedures

3-Methoxycarbonylpyridine *N*-oxide



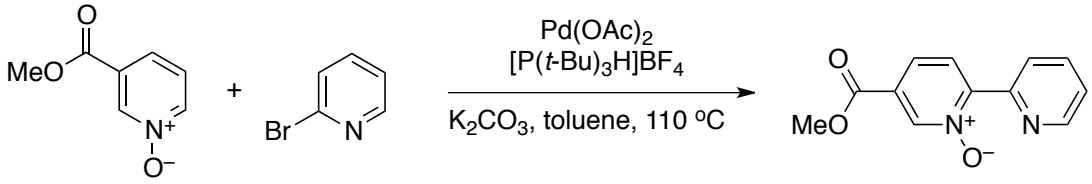
3-Methoxycarbonylpyridine *N*-oxide was prepared by oxidation of methyl nicotinate as described previously.[148] The spectral data and yields matched those reported previously.

4-Methoxycarbonylpyridine *N*-oxide



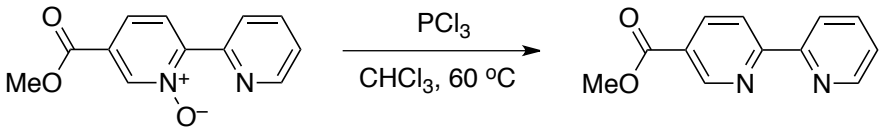
4-Methoxycarbonylpyridine *N*-oxide was prepared by oxidation of methyl isonicotinate as described previously.[148] The spectral data and yields matched those reported previously.

5-Methoxycarbonyl-2-(pyridin-2-yl)pyridine *N*-oxide



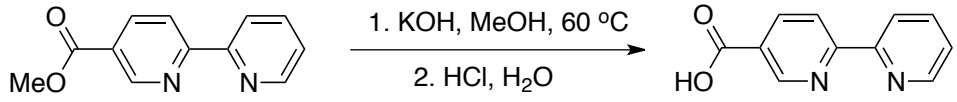
To a dried flask was added $Pd(OAc)_2$ (9 mg, 0.04 mmoles), $[P(t\text{-}Bu)_3H]BF_4$ (36 mg, 0.12 mmoles), K_2CO_3 (226 mg, 1.6 mmoles), and 3-methoxycarbonylpyridine *N*-oxide (500 mg, 3.3 mmoles). The flask was fitted with a reflux condenser capped with a septum, after which the system was evacuated and purged with N_2(g) (~5 times). A degassed solution of 2-bromopyridine (129 mg, 0.82 mmoles) in dry toluene (5 mL) was added via syringe, and the reaction was stirred at 110 °C for 18 hours. The cooled reaction mixture was filtered through Celite®, and the filtrate was concentrated under reduced pressure. The crude product was then purified by chromatography on silica (40% acetone in hexanes) to afford the title compound (100 mg, 53%) as a pale yellow solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 9.00 (dt, J = 1.0, 8.0 Hz, 1 H), 8.89 (dd, J = 0.4, 1.6 Hz, 1 H), 8.75 (ddd, J = 0.8, 1.6, 4.8 Hz, 1 H), 8.32 (d, J = 8.4 Hz, 1 H), 7.90 (dd, J = 1.6, 8.4 Hz, 1 H), 7.85 (td, J = 2.0, 8.0 Hz, 1 H), 7.39 (ddd, J = 0.8, 4.8, 7.6 Hz, 1 H), 3.98 (s, 3 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 163.5, 150.1 (2 signals), 149.6, 148.7, 141.9, 136.4, 128.5, 127.6, 125.7, 124.9, 53.0; **HRMS** (ESI) *m/z* 231.0766 [calc'd for $C_{12}H_{11}N_2O_3$ $(M + H)^+$ 231.0765].

Methyl 2-(Pyridin-2-yl)pyridine-5-carboxylate (methyl bipy5C)



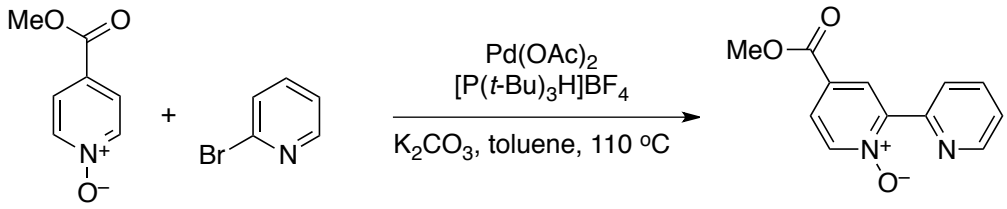
5-Methoxycarbonyl-2-(pyridin-2-yl)pyridine *N*-oxide (75 mg, 0.33 mmoles) was dissolved in dry $CHCl_3$ (3.3 mL), after which PCl_3 (34 μL, 0.39 mmoles) was added. The reaction was stirred at 60 °C until the starting material was consumed completely, as judged by TLC. The reaction was quenched by the dropwise addition of saturated aqueous Na_2CO_3 (5 mL) while stirring on ice. The product was extracted with DCM (4 × 5 mL), and the combined organics were dried over Na_2SO_4(s) and concentrated under reduced pressure to afford the title compound (67 mg, 96%) as a tan solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 9.29 (d, J = 1.2 Hz, 1 H), 8.76 (bs, 1 H), 8.57 (d, J = 8.4 Hz, 1 H), 8.53 (d, J = 8.0 Hz, 1 H), 8.44 (dd, J = 2.0, 8.0 Hz, 1 H), 7.92 (t, J = 7.6 Hz, 1 H), 7.43 (dd, J = 4.8, 6.8 Hz, 1 H), 4.00 (s, 3 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 165.7, 158.6, 154.5, 150.5, 148.8, 138.2, 137.8, 125.9, 124.7, 122.3, 120.8, 52.5; **HRMS** (ESI) *m/z* 215.0814 [calc'd for $C_{12}H_{11}N_2O_2$ $(M + H)^+$ 215.0816].

2-(Pyridin-2-yl)pyridine-5-carboxylic Acid (bipy5C)



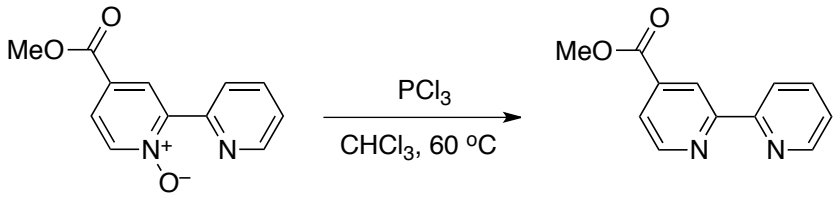
To a vial was added methyl bipy5C (50 mg, 0.23 mmoles) and KOH (60 mg, 0.83 mmoles). MeOH (2.3 mL) was added to the vial and the reaction mixture was heated to 60 °C until complete the starting material was consumed completely, as judged by TLC. The reaction mixture was cooled and concentrated under reduced pressure, after which the crude product was dissolved in water (2 mL). The aqueous layer was washed with EtOAc (1 × 2 mL), after which the product was precipitated from the aqueous layer by adjusting to pH 3–4 with 1M HCl. After cooling to 4°C, the product was filtered, washed with water (2 × 2 mL), and dried in vacuo to afford the title compound (28 mg, 60%) as a white solid. **^{1}H NMR** (400 MHz, DMSO-d_6, δ): 13.39 (bs, 1 H), 9.03 (dd, J = 0.5, 2.0 Hz, 1 H), 8.60 (dq, J = 1.0, 4.5 Hz, 1 H), 8.38 (dd, J = 0.5, 8.0 Hz, 1 H), 8.32 (d, J = 8.0 Hz, 1 H), 8.27 (dd, J = 2.0, 8.0 Hz, 1 H), 7.86 (dt, J = 2.0, 8.0 Hz, 1 H), 7.39 (ddd, J = 1.0, 5.0, 7.5 Hz, 1 H); **^{13}C NMR** (100 MHz, DMSO-d_6, δ): 166.5, 158.7, 154.5, 150.5, 149.9, 138.6, 137.9, 126.9, 125.3, 121.6, 120.6; **HRMS** (EI) *m/z* 200.0577 [calc'd for $C_{11}H_8N_2O_2$ $(M)^+$ 200.0581].

4-Methoxycarbonyl-2-(pyridin-2-yl)pyridine *N*-oxide



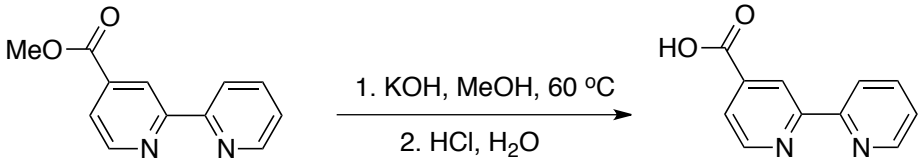
To a dried flask was added $Pd(OAc)_2$ (9 mg, 0.04 mmoles), $[P(t\text{-}Bu)_3H]BF_4$ (36 mg, 0.12 mmoles), K_2CO_3 (226 mg, 1.6 mmoles), and 3-methoxycarbonylpyridine *N*-oxide (500 mg, 3.3 mmoles). The flask was fitted with a reflux condenser capped with a septum, after which the system was evacuated and purged with N_2(g) (~5 times). A degassed solution of 2-bromopyridine (129 mg, 0.82 mmoles) in dry toluene (5 mL) was added via syringe, and the reaction was stirred at 110 °C for 18 hours. The cooled reaction mixture was filtered through Celite®, and the filtrate concentrated under reduced pressure. The crude product was then purified by chromatography on silica (40% acetone in hexanes) to afford the title compound (128 mg, 68%) as a solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 8.94–9.76 (m, 3 H), 8.33 (d, J = 6.8 Hz, 1 H), 7.87–7.83 (m, 2 H), 7.38 (ddd, J = 1.2, 4.8, 7.6 Hz, 1 H), 3.97 (s, 3 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 164.1, 149.6, 148.9, 147.4, 141.0, 136.4, 128.4, 126.4, 125.2, 125.0, 124.6, 52.8; **HRMS** (ESI) *m/z* 231.0766 [calc'd for $C_{12}H_{11}N_2O_3$ $(M + H)^+$ 231.0765].

Methyl 2-(Pyridin-2-yl)pyridine-4-carboxylate (methyl bipy4C)



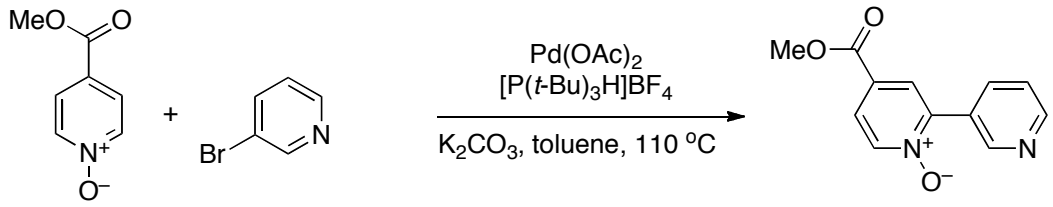
4-Methoxycarbonyl-2-(pyridin-2-yl)pyridine *N*-oxide (100 mg, 0.43 mmoles) was dissolved in dry $CHCl_3$ (4.0 mL), after which PCl_3 (45 μL, 0.52 mmoles) was added. The reaction was stirred at 60 °C until the starting materinal was consumed completely, as judged by TLC. The reaction was quenched by the dropwise addition of saturated aqueous Na_2CO_3 (10 mL) while stirring on ice. The product was extracted with DCM (5 × 10 mL), and the combined organics were dried over Na_2SO_4(s) and concentrated under reduced pressure. The crude product was then purified by chromatography on silica (30% acetone in hexanes) to afford the title compound (83 mg, 89%) as a white solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 8.95 (q, J = 0.8 Hz, 1 H), 8.83 (dd, J = 0.8, 4.8 Hz, 1 H), 8.73 (dq, J = 0.8, 4.8 Hz, 1 H), 8.43 (dt, J = 0.8, 8.0 Hz, 1 H), 7.88–7.82 (m, 2 H), 7.35 (ddd, J = 1.0, 4.8, 7.2 Hz, 1 H), 3.99 (s, 3 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 165.7, 157.3, 155.3, 149.9, 149.3, 138.4, 137.0, 124.1, 122.8, 121.2, 120.4, 52.7; **HRMS** (ESI) *m/z* 205.0818 [calc'd for $C_{12}H_{11}N_2O_2$ $(M + H)^+$ 215.0816].

2-(Pyridin-2-yl)pyridine-4-carboxylic Acid (bipy4C)



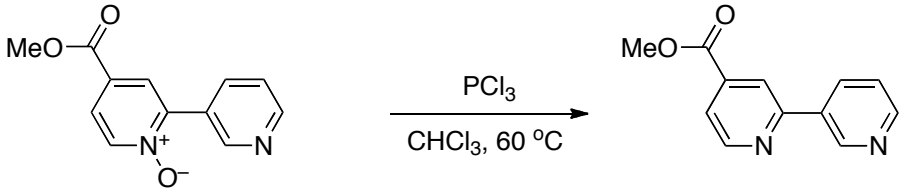
To a vial was added methyl bipy4C (50 mg, 0.23 mmoles) and KOH (60 mg, 0.83 mmoles). MeOH (2.3 mL) was added to the vial and the reaction mixture was heated to 60 °C until the starting material was consumed completely, as judged by TLC. The reaction mixture was cooled and concentrated under reduced pressure, after which the crude product was dissolved in water (2 mL). The aqueous layer was washed with EtOAc (1 × 2 mL), after which the product was precipitated from the aqueous layer by adjusting to pH 3–4 with 1M HCl. After cooling to 4°C, the product was filtered, washed with water (2 × 2 mL), and dried in vacuo to afford the title compound (30 mg, 64%) as a white solid. **^{1}H NMR** (400 MHz, DMSO-d_6, δ): 13.81 (bs, 1 H), 8.89 (dd, J = 0.5, 5.0 Hz, 1 H), 8.84 (dd, J = 0.5, 1.0 Hz, 1 H), 8.74 (ddd, J = 0.5, 1.0, 4.5 Hz, 1 H), 8.43 (dt, J = 1.0, 8.0 Hz, 1 H), 8.00 (td, J = 1.5, 8.0 Hz, 1 H), 7.88 (dd, J = 1.5, 5.0 Hz, 1 H), 7.52 (ddd, J = 1.0, 4.5, 7.5 Hz, 1 H); **^{13}C NMR** (100 MHz, DMSO-d_6, δ): 166.6, 156.8, 154.8, 150.9, 150.0, 139.8, 138.0, 125.2, 123.5, 121.1, 119.9; **HRMS** (EI) *m/z* 200.0583 [calc'd for $C_{11}H_8N_2O_2$ $(M)^+$ 200.0581].

4-Methoxycarbonyl-2-(pyridin-3-yl)pyridine *N*-oxide



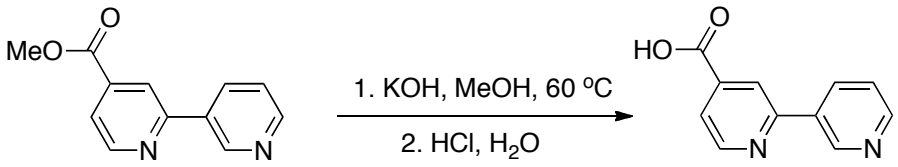
To a dried flask was added $Pd(OAc)_2$ (9 mg, 0.040 mmoles), $[P(t\text{-}Bu)_3H]BF_4$ (36 mg, 0.12 mmoles), K_2CO_3 (226 mg, 1.63 mmoles), and 4-methoxycarbonylpyridine *N*-oxide (500 mg, 3.26 mmoles). The flask was fitted with a reflux condenser capped with a septum, after which the system was evacuated and purged with N_2(g) (~5 times). A degassed solution of 3-bromopyridine (129 mg, 0.82 mmoles) in dry toluene (5 mL) was added via syringe, and the reaction was stirred at 110 °C for 18 hours. The cooled reaction mixture was filtered through Celite®, and the filtrate concentrated under reduced pressure. The crude product was then further purified by chromatography on silica (4% MeOH in EtOAc) to afford the title compound (100 mg) as a white solid. Due to the presence of minor contaminants that were difficult to remove by chromatography or recrystallization, the slightly crude product was used directly in the next reaction before further purification and characterization. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 8.95 (d, J = 1.6 Hz, 1 H), 8.71 (d, J = 0.8, 4.8 Hz, 1 H), 8.35 (d, J = 6.8 Hz, 1 H), 8.30 (dt, J = 1.6, 7.6 Hz, 1 H), 8.10 (d, J = 2.4 Hz, 1 H), 7.87 (dd, J = 2.4, 6.8 Hz, 1 H), 7.44 (dd, J = 5.2, 8.0 Hz, 1 H), 3.97 (s, 3 H); **HRMS** (ESI) *m/z* 231.0761 [calc'd for $C_{12}H_{11}N_2O_3$ $(M + H)^+$ 231.0765].

Methyl 2-(Pyridin-3-yl)pyridine-4-carboxylate



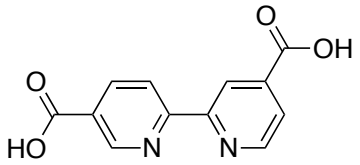
4-Methoxycarbonyl-2-(pyridin-3-yl)pyridine *N*-oxide (75 mg, 0.33 mmoles) was dissolved in dry $CHCl_3$ (3.3 mL), after which PCl_3 (68 μL, 0.78 mmoles) was added. The reaction was stirred at 60 °C until the starting materinal was consumed completely, as judged by TLC. The reaction was quenched by the dropwise addition of saturated aqueous Na_2CO_3 (5 mL) while stirring on ice. The product was extracted with DCM (4 × 5 mL), and the combined organics were dried over Na_2SO_4(s) and concentrated under reduced pressure. The crude product was then purified by chromatography on silica (60% acetone in hexanes) to afford the title compound (40 mg, 31% over 2 steps) as a white solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 9.27 (d, J = 1.6 Hz, 1 H), 8.87 (dd, J = 0.8, 4.8 Hz, 1 H), 8.69 (dd, J = 1.2, 4.8 Hz, 1 H), 8.36 (ddd, J = 2.0, 2.4, 8.0 Hz, 1 H), 8.31 (dd, J = 0.8, 1.2 Hz, 1 H), 7.83 (dd, J = 1.6, 5.2 Hz, 1 H), 7.43 (ddd, J = 0.4, 4.8, 8.0 Hz, 1 H), 4.00 (s, 3 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 165.4, 155.9, 150.8, 150.4, 148.3, 138.4, 134.3, 134.0, 123.6, 121.9, 119.7, 52.8; **HRMS** (ESI) *m/z* 205.0815 [calc'd for $C_{12}H_{11}N_2O_2$ $(M + H)^+$ 215.0816].

2-(Pyridin-3-yl)pyridine-4-carboxylic Acid



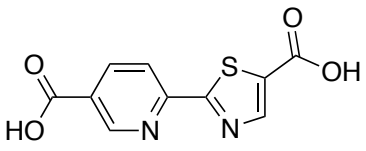
To a vial was added methyl 2-(pyridin-3-yl)pyridine-4-carboxylate (32 mg, 0.15 mmoles) and KOH (38 mg, 0.6 mmoles). MeOH (1.5 mL) was added to the vial and the reaction mixture was heated to 60 °C until the starting material was consumed completely, as judged by TLC. The reaction mixture was cooled and concentrated under reduced pressure, after which the crude product was dissolved in water (2 mL). The aqueous layer was washed with EtOAc (1 × 1 mL), after which the product was precipitated from the aqueous layer by adjusting to pH 3–4 with 1M HCl. After cooling to 4°C, the product was filtered, washed with water (2 × 1 mL), and dried in vacuo to afford the title compound (24 mg, 82%) as a white solid. **^{1}H NMR** (500 MHz, DMSO-d_6, δ): 13.71 (bs, 1 H), 9.17 (d, J = 2.5 Hz, 1 H), 8.77 (dd, J = 0.5, 5.0 Hz, 1 H), 8.54 (dd, J = 1.5, 5.0 Hz, 1 H), 8.36 (ddd, J = 1.5, 3.5, 8.0 Hz, 1 H), 8.24 (s, 1 H), 7.71 (dd, J = 1.5, 5.0 Hz, 1 H), 7.41 (ddd, J = 0.5, 4.5, 8.0 Hz, 1 H); **^{13}C NMR** (125 MHz, DMSO-d_6, δ): 166.5, 155.4, 151.4, 150.7, 148.3, 140.0, 134.7, 133.8, 124.3, 122.6, 119.9; **HRMS** (EI) *m/z* 200.0580 [calc'd for $C_{11}H_8N_2O_2$ $(M)^+$ 200.0581].

Bipy45′DC



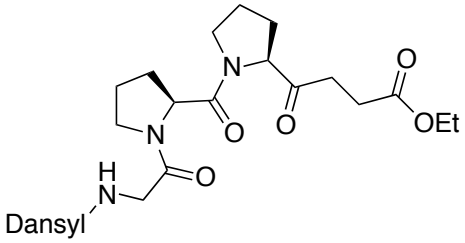
Bipy45′DC was synthesized in 4 steps as described previously,[148] with matching spectral data and yields for all synthetic transformations.

PythiDC



PythiDC was synthesized in 2 steps as described previously (see: CHAPTER 3), with matching spectral data and yields for all synthetic transformations.

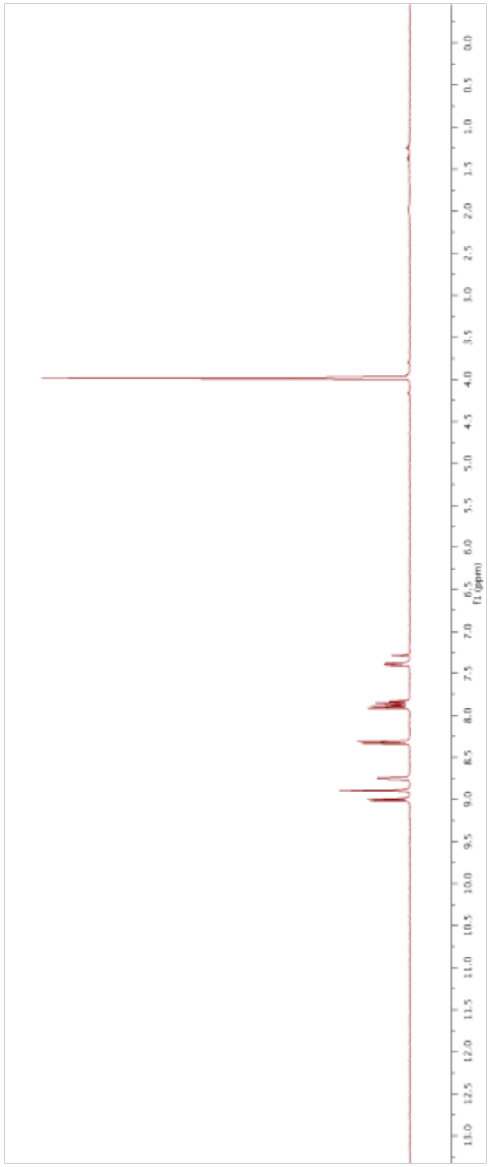
***N*-Dansylglycyl-(2*S*)-prolyl-(2*S*)-prolylglycine Ethyl Ester (DansylGlyProProGlyOEt):**



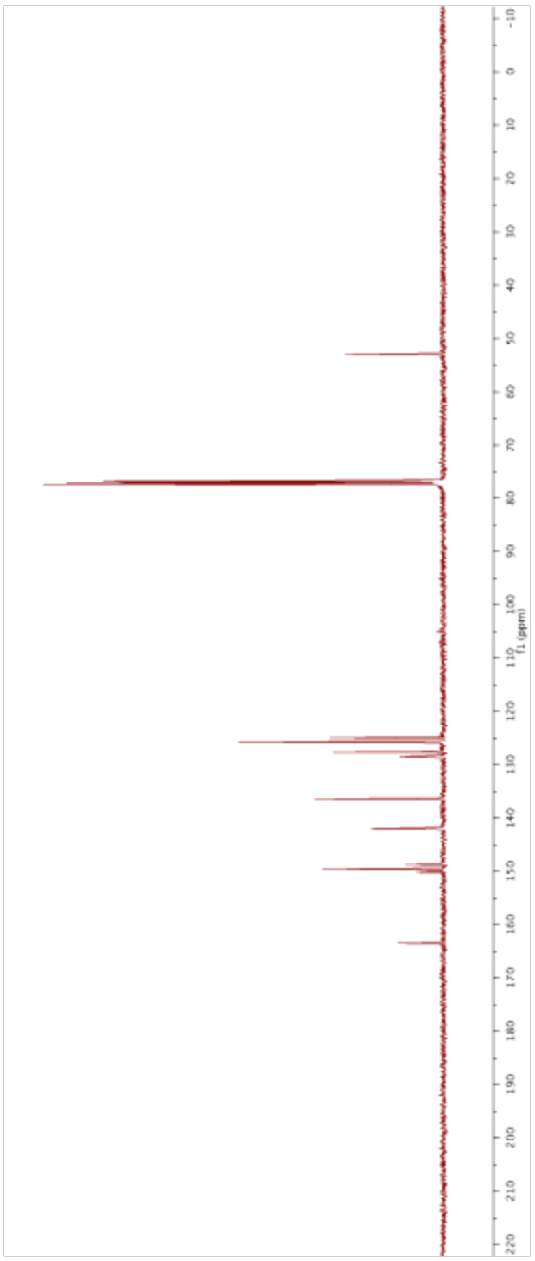
DansylGlyProProGlyOEt was synthesized as described previously.[148] The spectral data and yield matched that reported previously.

4.5.13 NMR Spectra

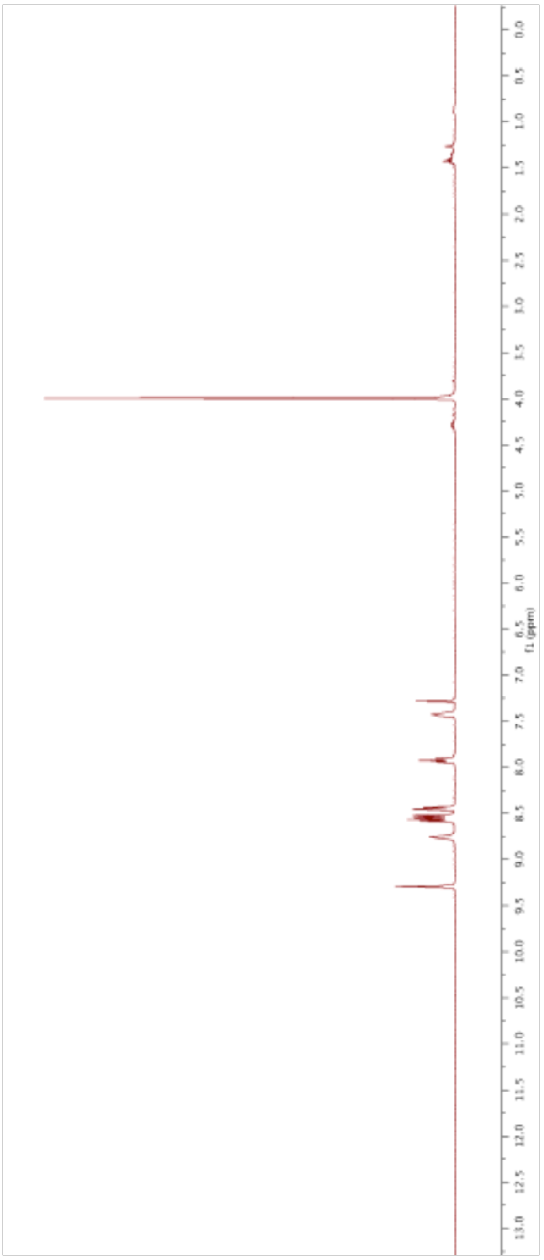
400 MHz ^{1}H NMR Spectrum of 5-Methoxycarbonyl-2-(pyridin-2-yl)pyridine *N*-oxide in $CDCl_3$



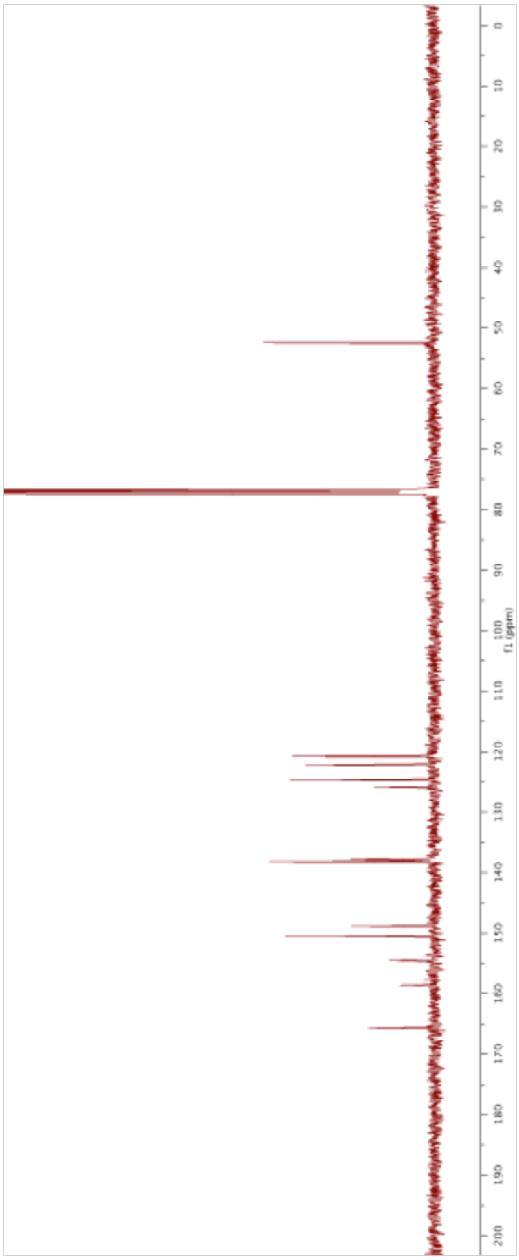
100 MHz ^{13}C NMR Spectrum of 5-Methoxycarbonyl-2-(pyridin-2-yl)pyridine *N*-oxide in $CDCl_3$



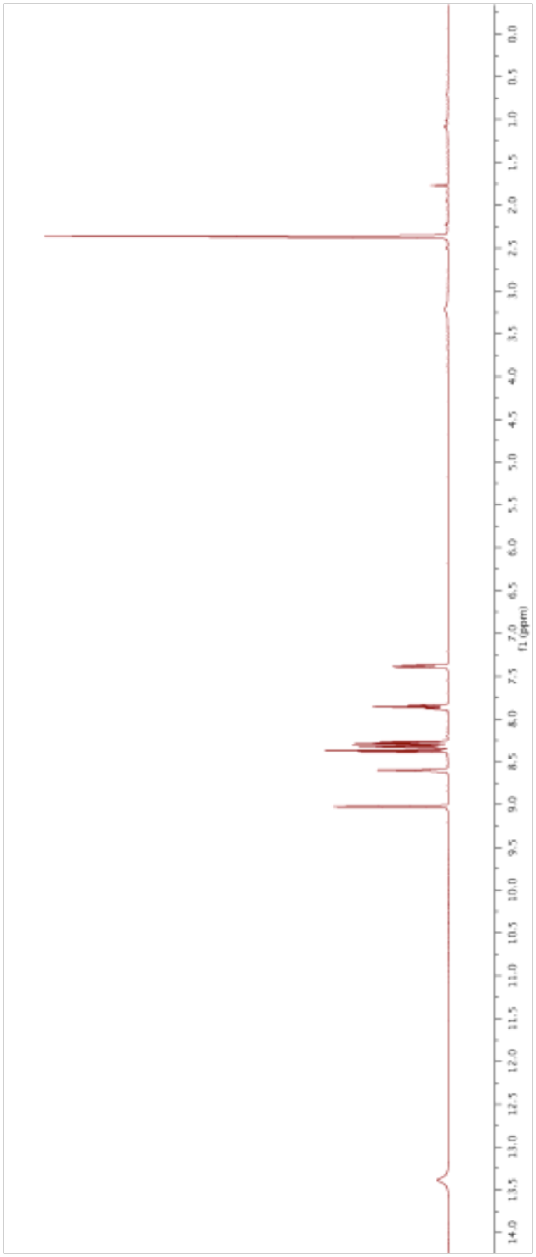
400 MHz ^{1}H NMR Spectrum of Methyl Bipy5C in $CDCl_3$



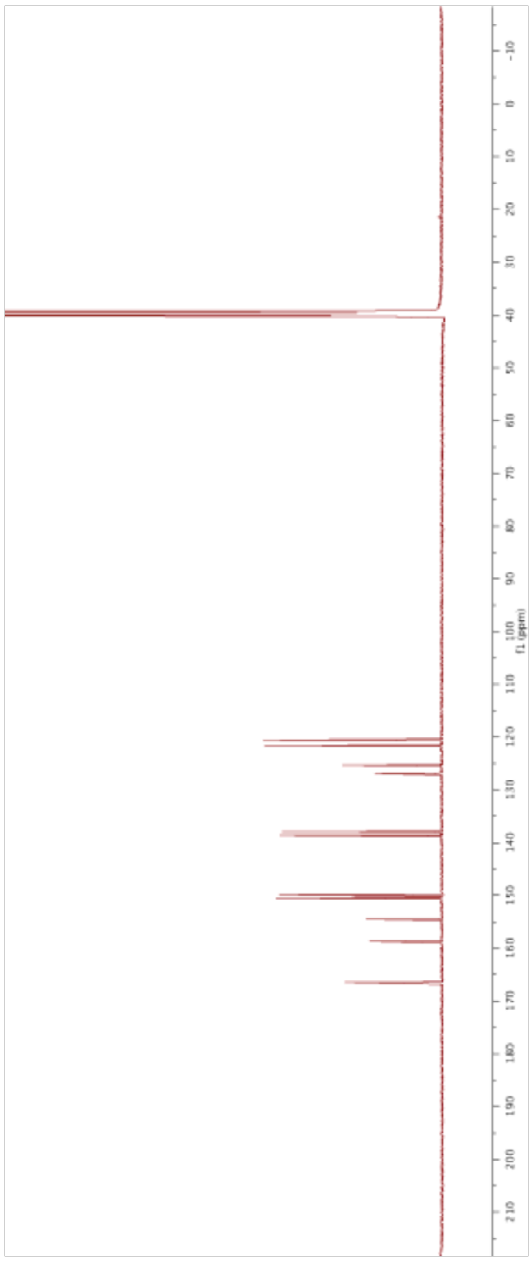
100 MHz ^{13}C NMR Spectrum of Methyl Bipy5C in $CDCl_3$



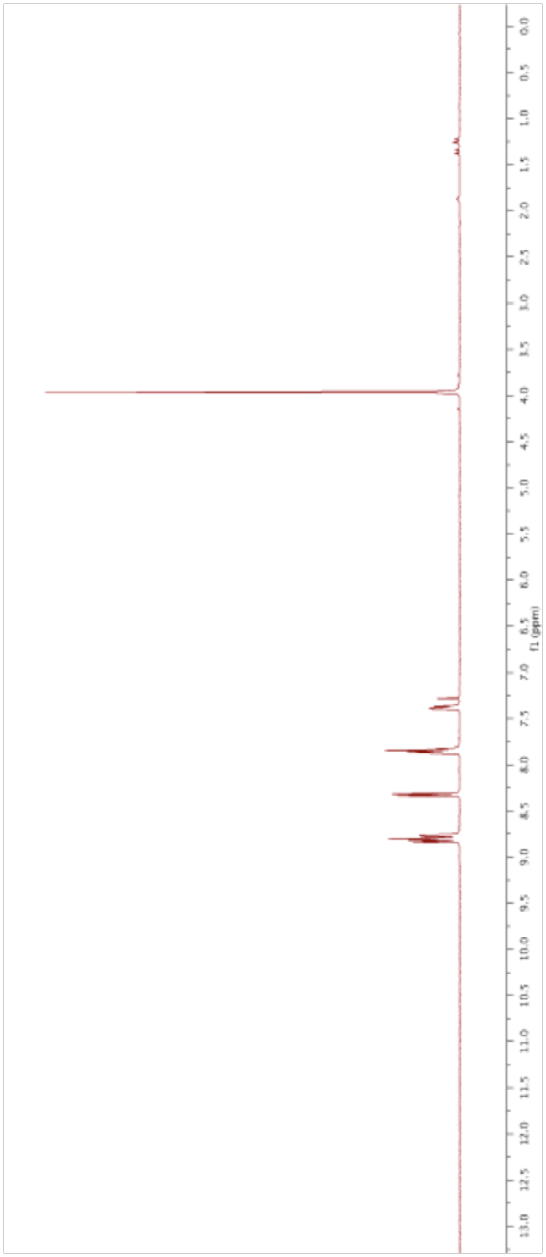
400 MHz ^{1}H NMR Spectrum of Bipy5C in DMSO-d_6



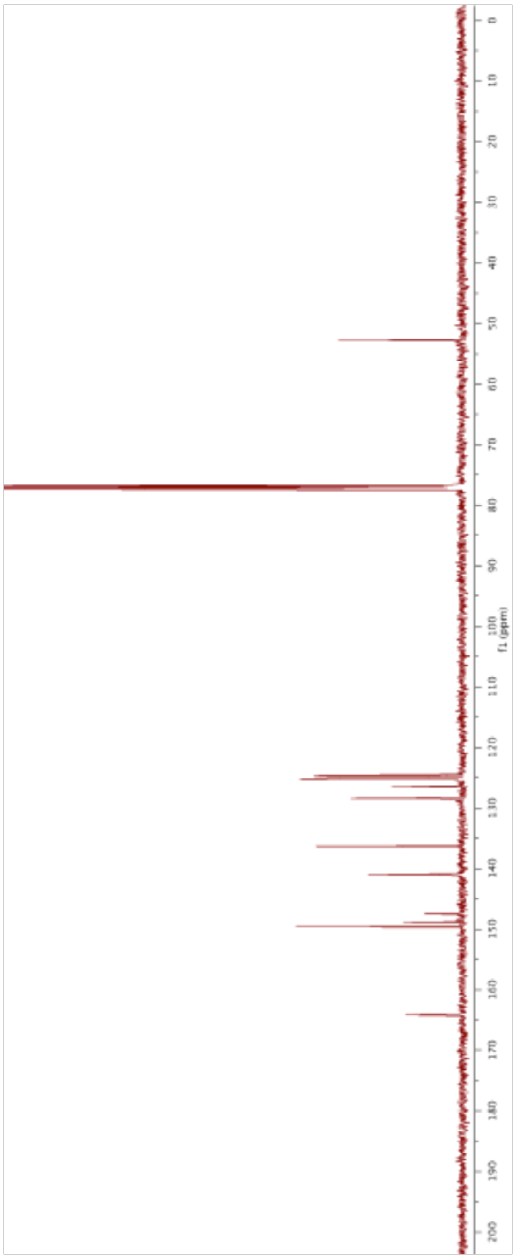
100 MHz ^{13}C NMR Spectrum of Bipy5C in DMSO-d_6



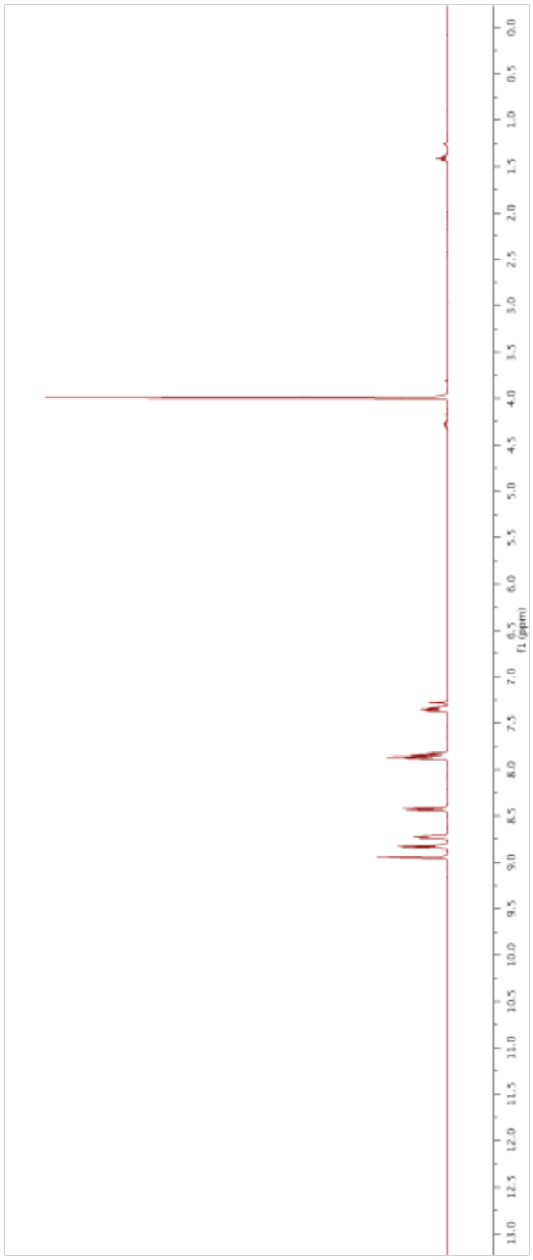
400 MHz ^{1}H NMR Spectrum of 4-Methoxycarbonyl-2-(pyridin-2-yl)pyridine *N*-Oxide in $CDCl_3$



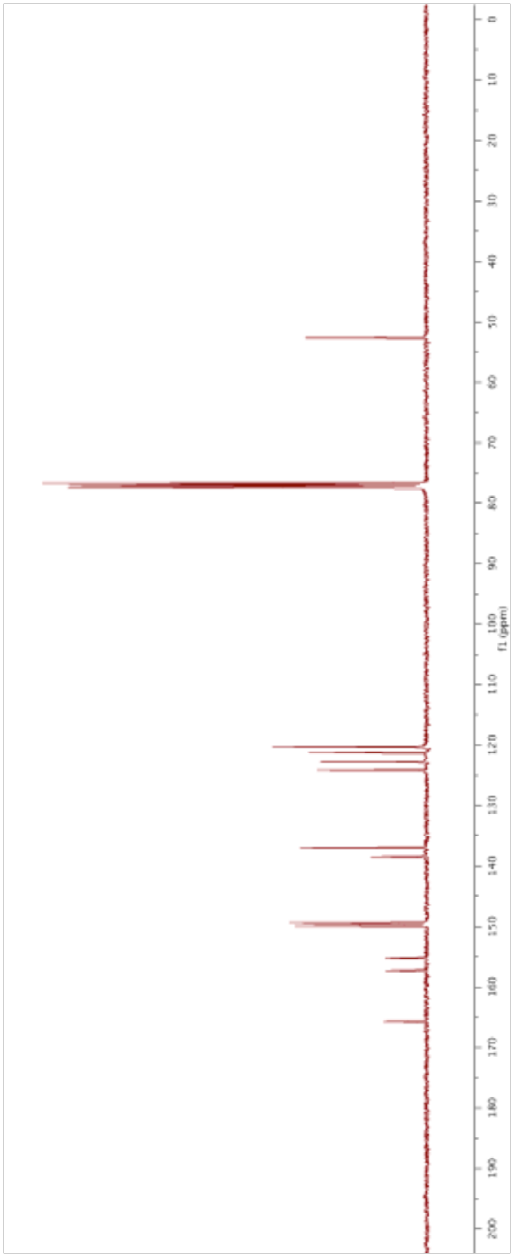
100 MHz ^{13}C NMR Spectrum of 4-Methoxycarbonyl-2-(pyridin-2-yl)pyridine *N*-Oxide in $CDCl_3$



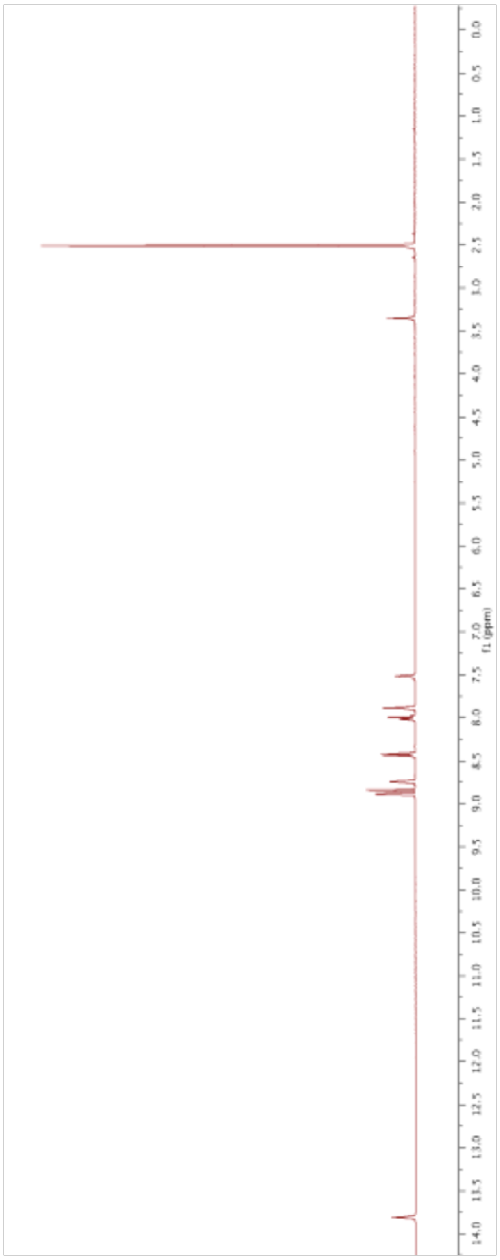
400 MHz ^{1}H NMR Spectrum of Methyl Bipy4C in $CDCl_3$



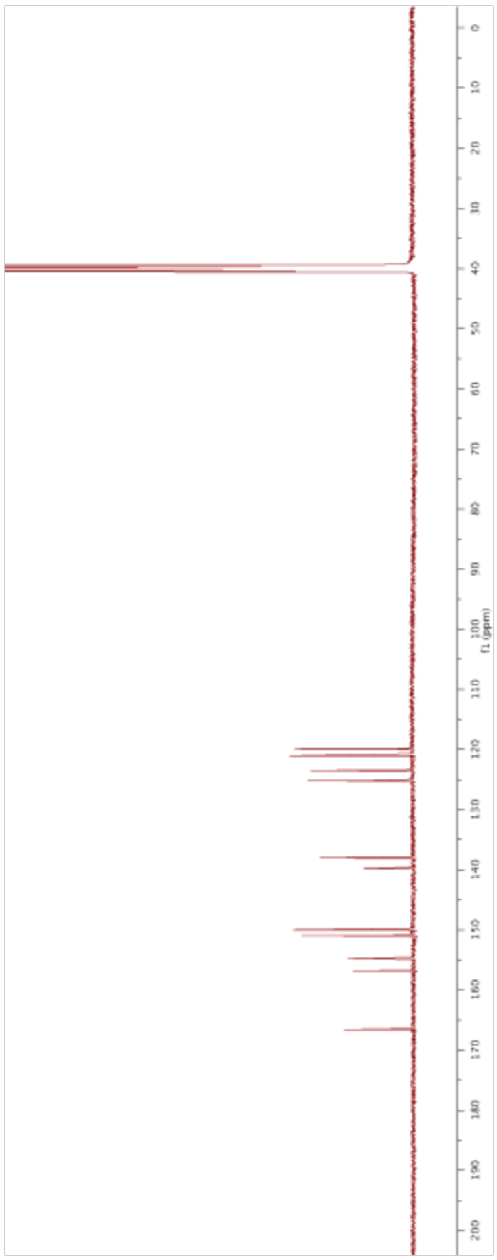
100 MHz ^{13}C NMR Spectrum of Methyl Bipy4C in $CDCl_3$



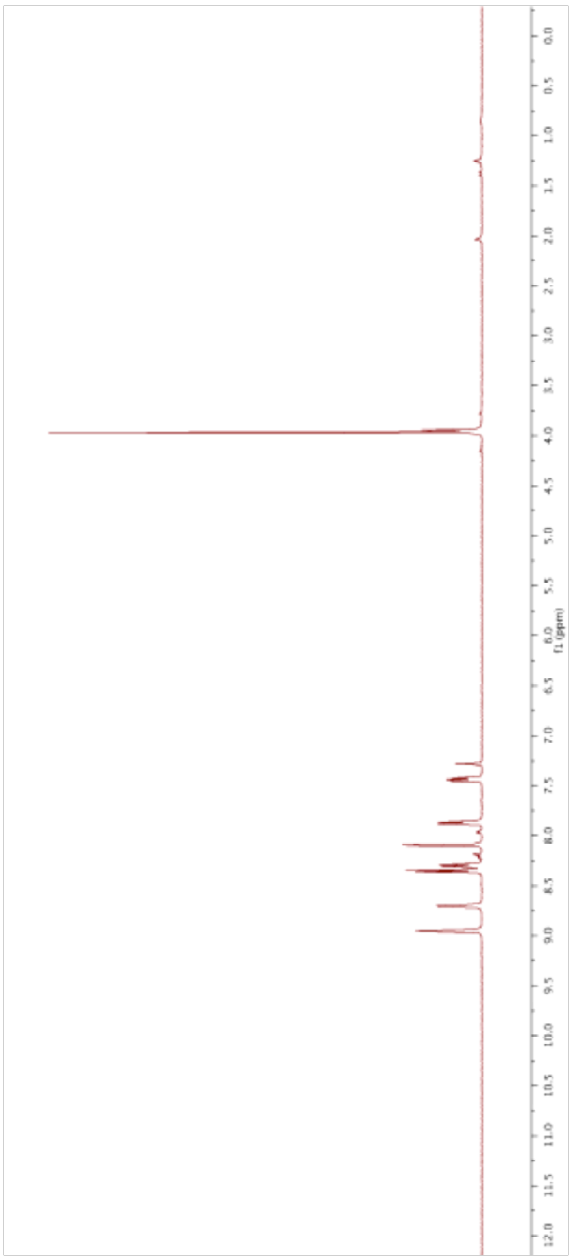
400 MHz ^{1}H NMR Spectrum of Bipy4C in DMSO-d_6



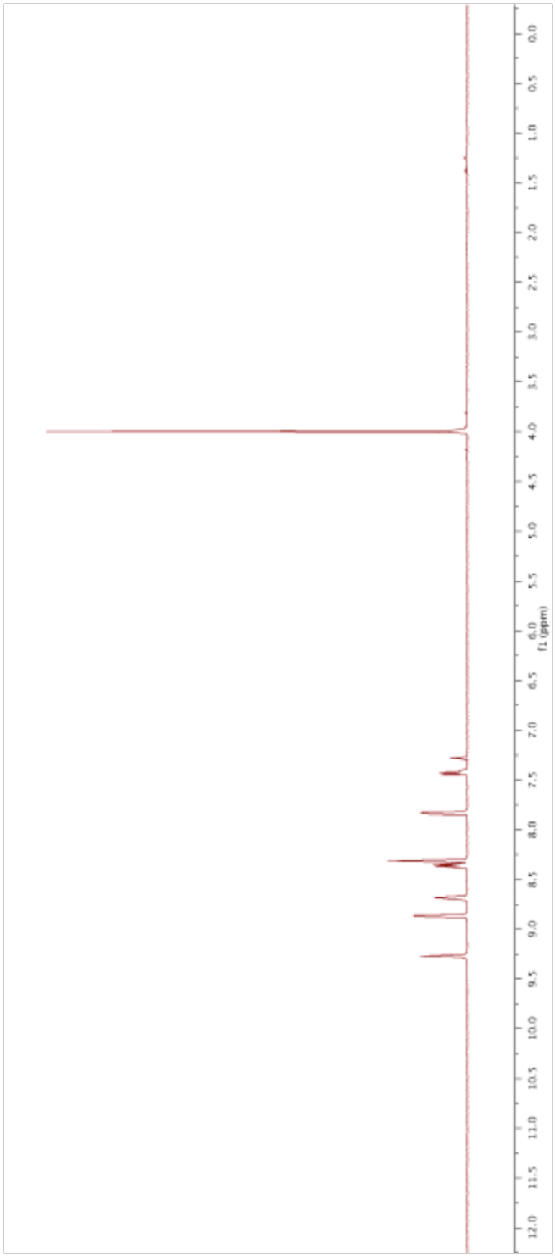
100 MHz ^{13}C NMR Spectrum of Bipy4C in DMSO-d_6



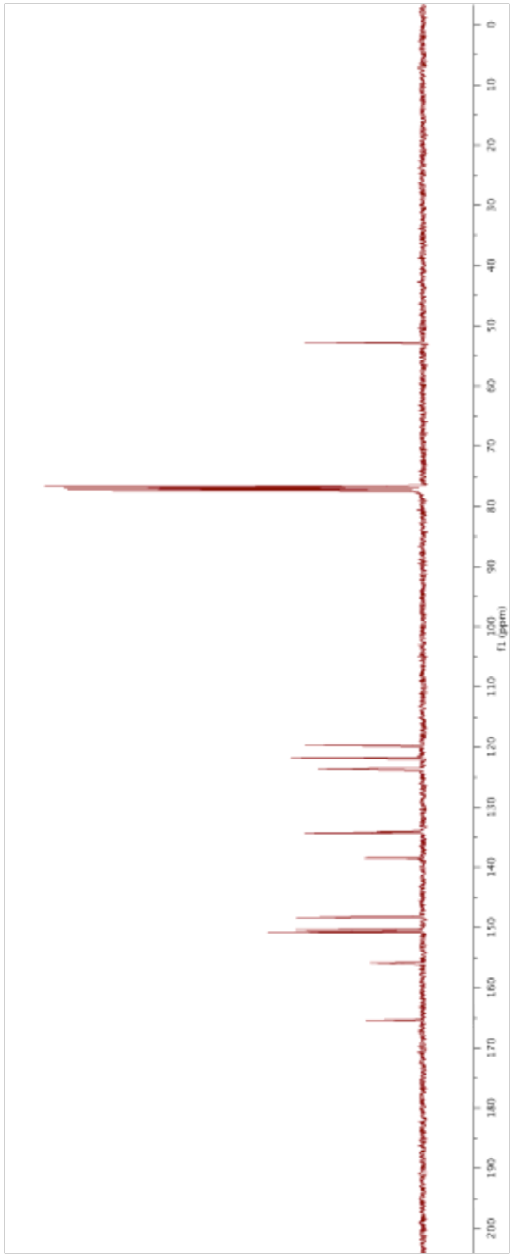
400 MHz ^{1}H NMR Spectrum of Slightly Crude Methyl 2-(pyridin-3-yl)pyridine-4-carboxylate *N*-Oxide in $CDCl_3$



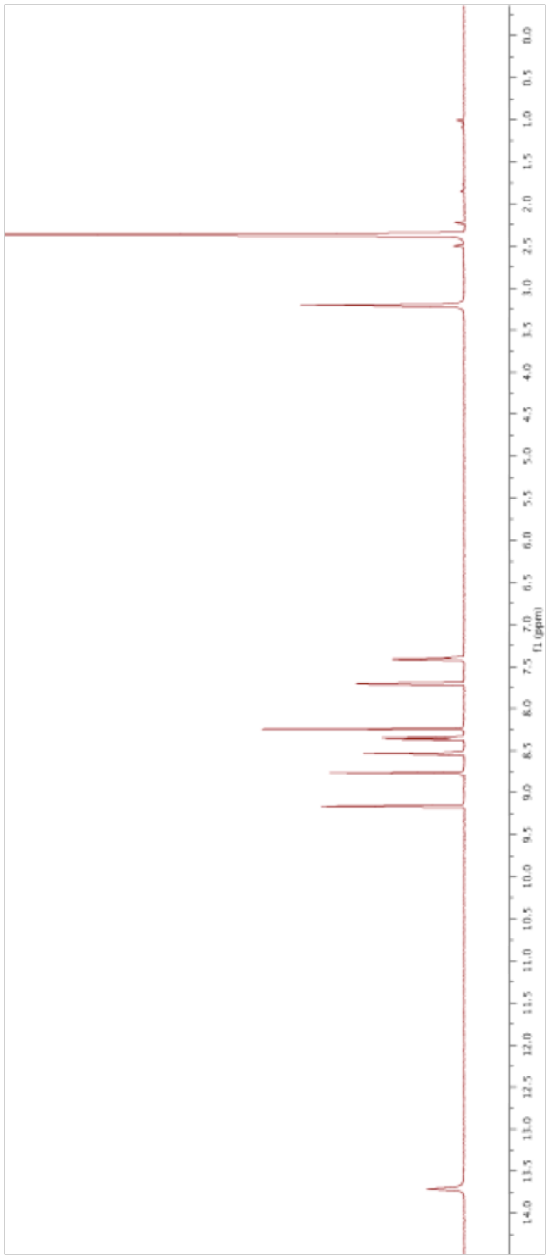
400 MHz ^{1}H NMR Spectrum of Methyl 2-(pyridin-3-yl)pyridine-4-carboxylate in $CDCl_3$



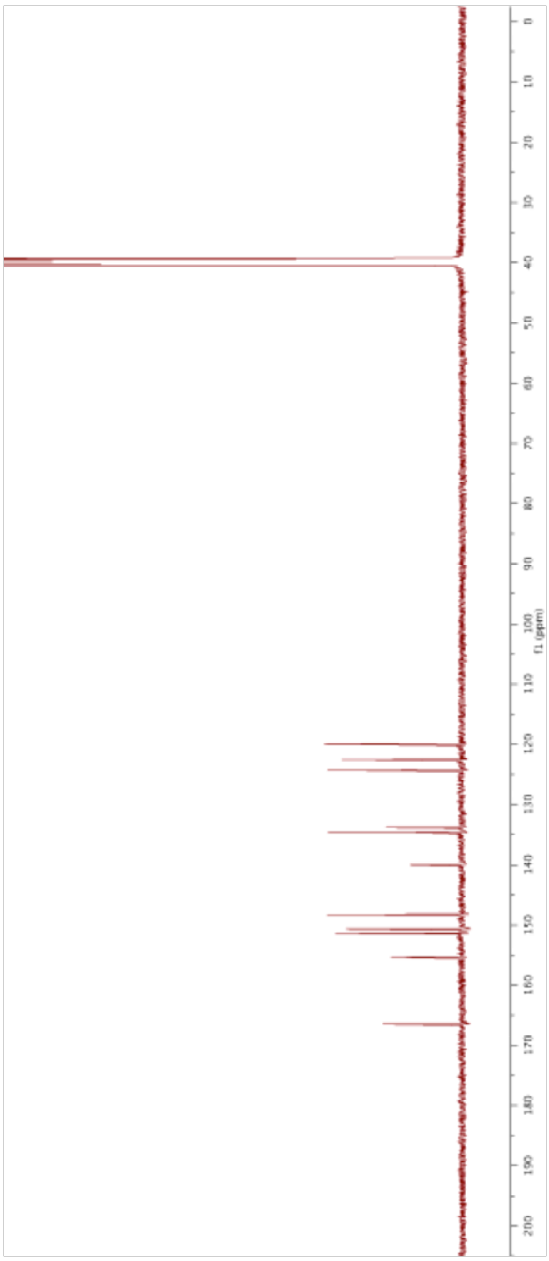
100 MHz ^{13}C NMR Spectrum of Methyl 2-(pyridin-3-yl)pyridine-4-carboxylate in $CDCl_3$



500 MHz ^{1}H NMR Spectrum of 2-(Pyridin-3-yl)pyridine-4-carboxylic Acid in DMSO-d_6

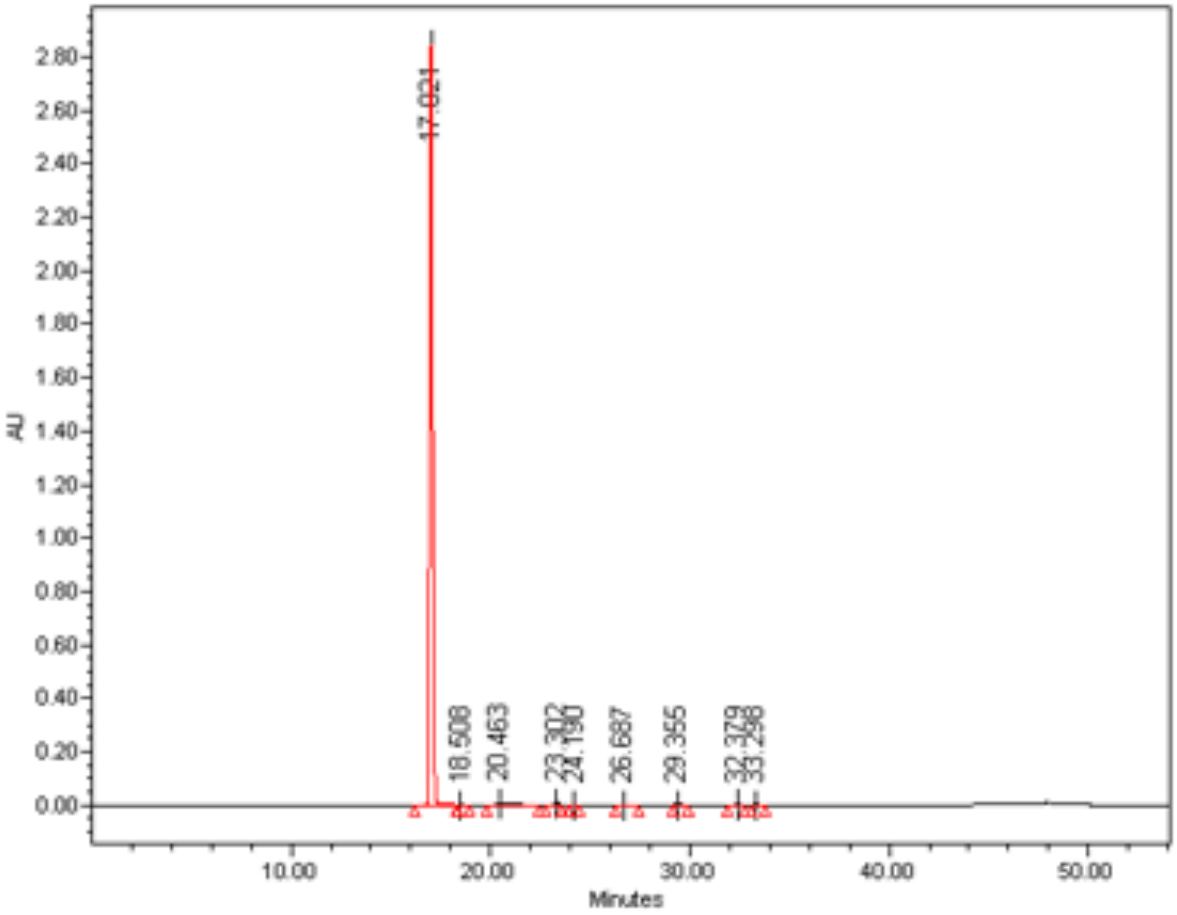


125 MHz ^{13}C NMR Spectrum of 2-(Pyridin-3-yl)pyridine-4-carboxylic Acid in DMSO-d_6



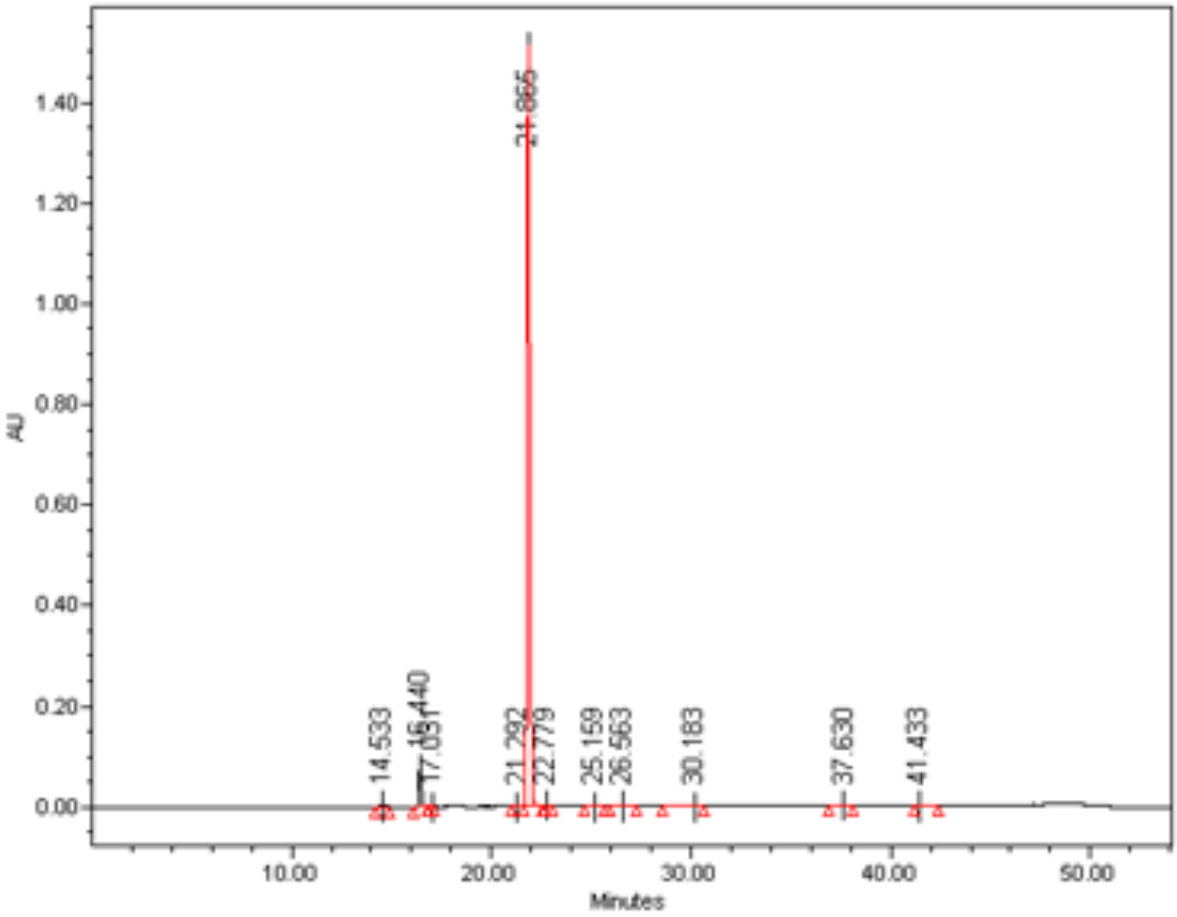
4.5.14 HPLC Purity Chromatograms of Final Compounds

Bipy5C



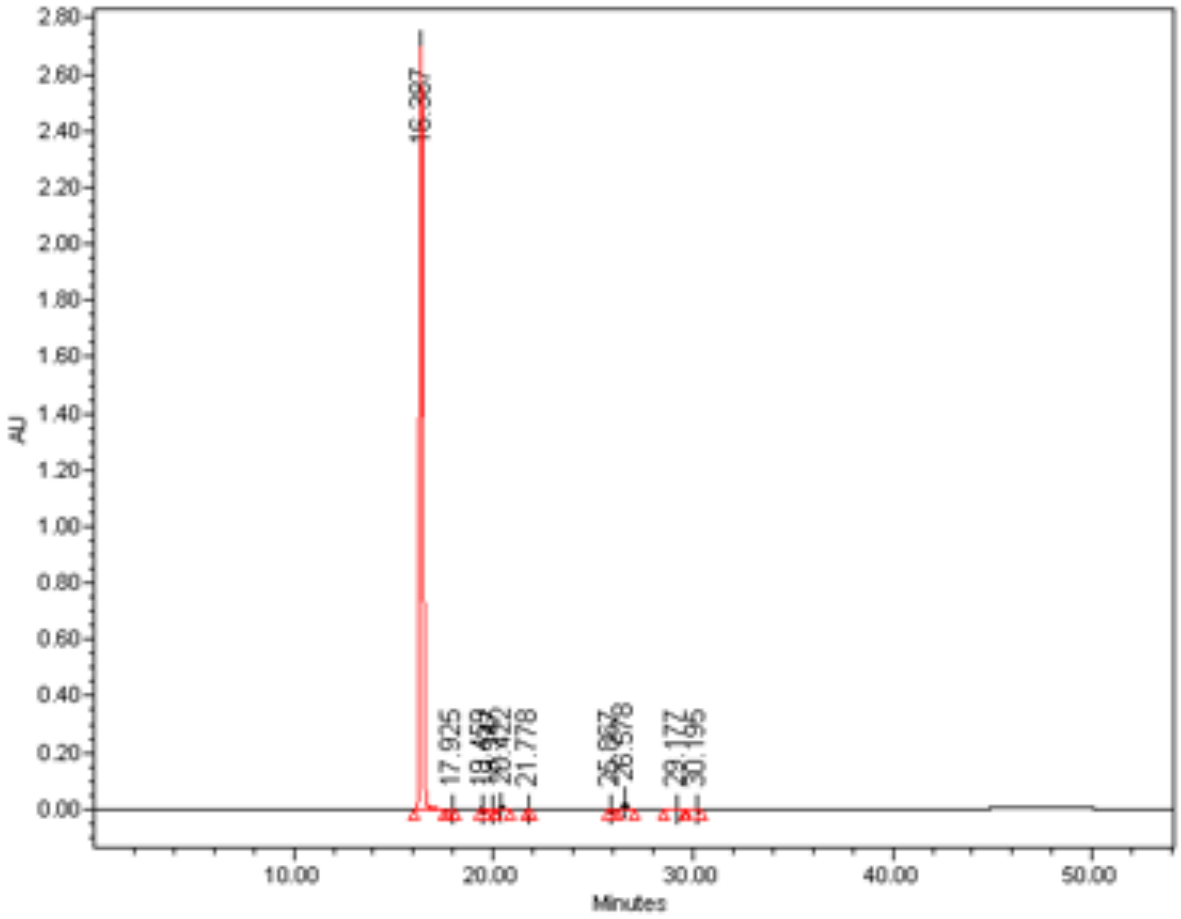
Peak	Retention Time (min)	Area (μV*sec)	% Area	Height (μV)
1	17.021	32559287	98.87	2850687
2	18.508	9619	0.03	1072
3	20.463	151551	0.46	8172
4	23.302	93059	0.28	8529
5	24.190	3841	0.01	326
6	26.687	7768	0.02	411
7	29.355	36489	0.11	2674
8	32.379	36045	0.11	2964
9	33.298	32264	0.1	2651

Methyl Bipy4C



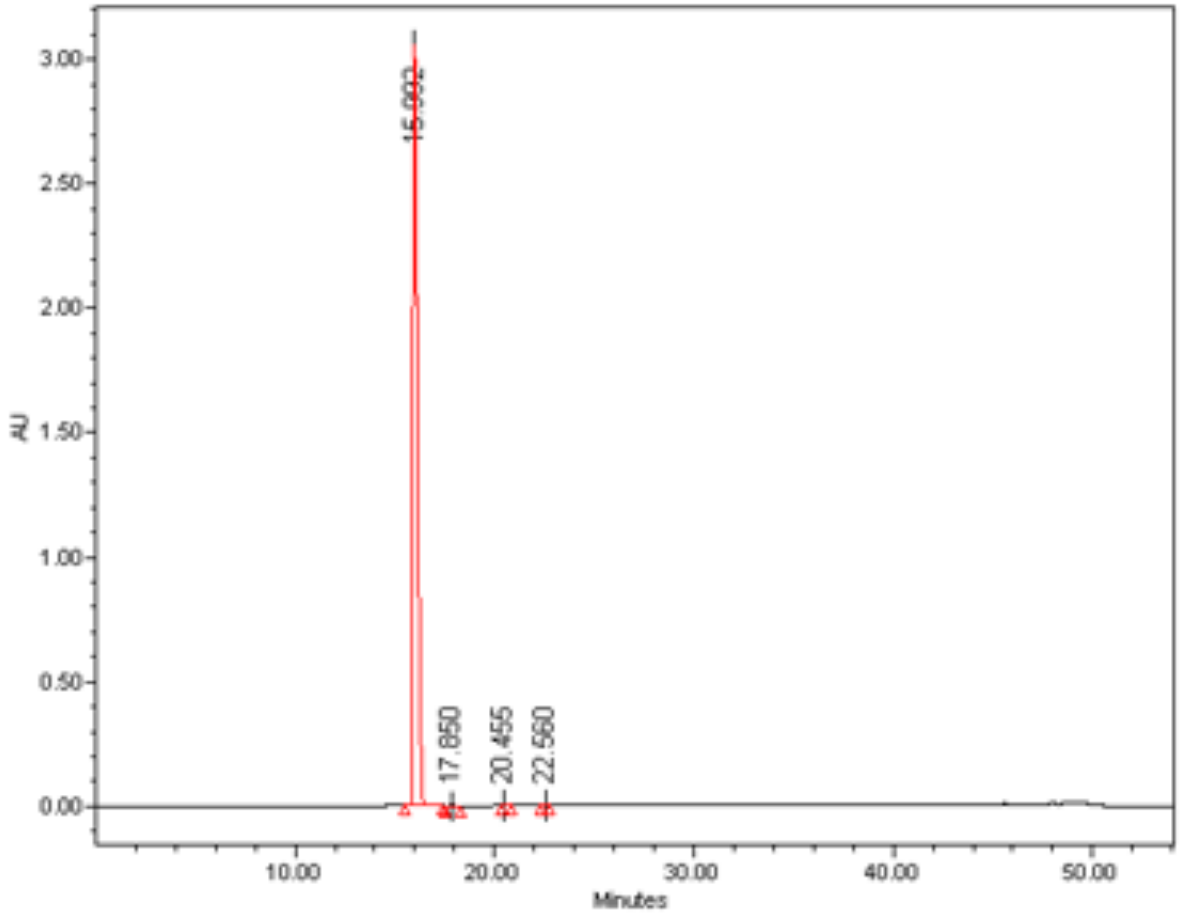
Peak	Retention Time (min)	Area (µV*sec)	% Area	Height (µV)
1	14.533	20406	0.13	1513
2	16.440	669538	4.13	73884
3	17.031	3629	0.02	452
4	21.292	9134	0.06	756
5	**21.865**	**15358256**	**94.83**	**1515017**
6	22.779	19097	0.12	1990
7	25.159	7390	0.05	480
8	26.563	17757	0.11	1287
9	30.183	19086	0.12	645
10	37.630	43906	0.27	2643
11	41.433	28014	0.17	1711

Bipy4C



Peak	Retention Time (min)	Area (μV*sec)	% Area	Height (μV)
1	**16.387**	**31238247**	**98.64**	**2709775**
2	17.925	2791	0.01	291
3	19.459	12298	0.04	1313
4	19.947	3913	0.01	370
5	20.422	88915	0.28	8275
6	21.778	5577	0.02	617
7	25.867	4108	0.01	475
8	26.578	290303	0.92	25471
9	29.177	10734	0.03	644
10	30.195	10885	0.03	838

Bipy4C*



Peak	Retention Time (min)	Area (µV*sec)	% Area	Height (µV)
1	**15.992**	**48280156**	**99.87**	**3059525**
2	17.850	5530	0.01	493
3	20.455	25817	0.05	2505
4	22.560	29168	0.06	2451

4.6 Acknowledgments

This work was supported by Grant R01 AR044276 (NIH). J. D. V. was supported by Molecular Biosciences Training Grant T32 GM007215 (NIH) and a fellowship from the Department of Biochemistry at the University of Wisconsin–Madison. This study made use of the National Magnetic Resonance Facility at Madison, which is supported by Grant P41 GM103399 (NIH). The Micromass LCT® mass spectrometer was obtained with support from Grant CHE-9974839 (NSF).

4.7 Author Contributions

J. D. Vasta designed and carried out all experiments. J. D. Vasta and R. T. Raines contributed to the analysis of data and preparation of the manuscript and figures.

Figure 4.1 Activation of CP4H1 by bipymonoCs is dose-dependent and requires the vitamin C cofactor

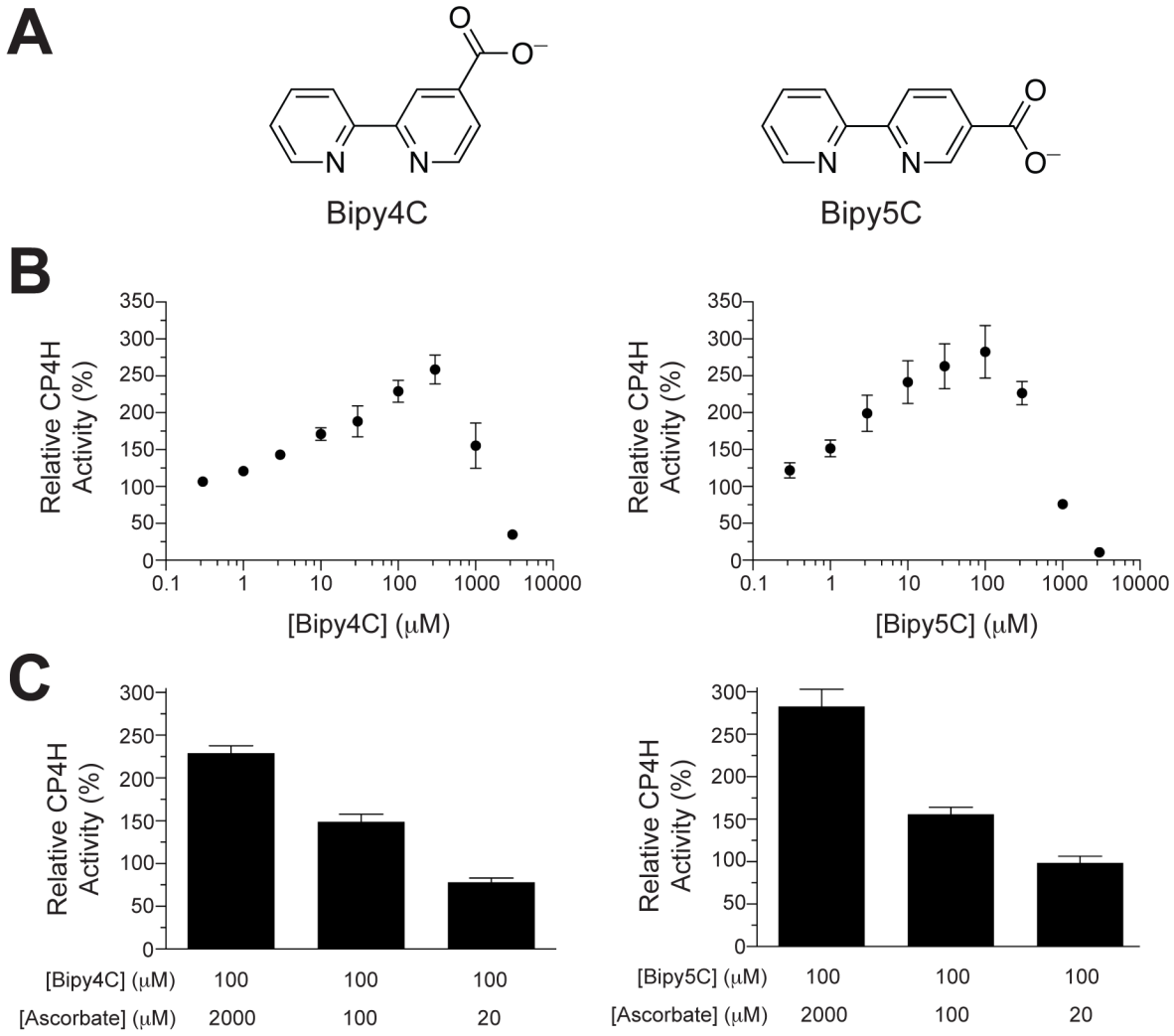


Figure 4.1 Activation of CP4H1 by bipymonoCs is dose-dependent and requires the Vitamin C cofactor. (A) Structures of bipymonoCs used in this study. (B) Dose-response plots for activation (and inhibition) of CP4H1 by bipymonoCs. CP4H assays in the presence and absence of bipymonoC activator (0.3–3000 μM) and ascorbate (2 mM) were conducted as described in the Experimental Procedures section with a saturating concentration of the cosubstrate AKG (1 mM). Relative activity values are the mean (± SE) of three independent experiments) and represent the ratio of CP4H activity observed in the presence and absence of activator. (C) Dependence of the activation observed by bipymonoCs on ascorbate concentration. CP4H assays in the presence and absence of bipymonoC activator (100 μM) and ascorbate (as denoted) were conducted as described in the Experimental Procedures section with a saturating concentration of the cosubstrate AKG (1 mM). Relative activity values are the mean (± SE) of three independent experiments) and represent the ratio of CP4H activity observed in the presence and absence of activator.

Figure 4.2 The mechanism of the overall reaction catalyzed by CP4Hs

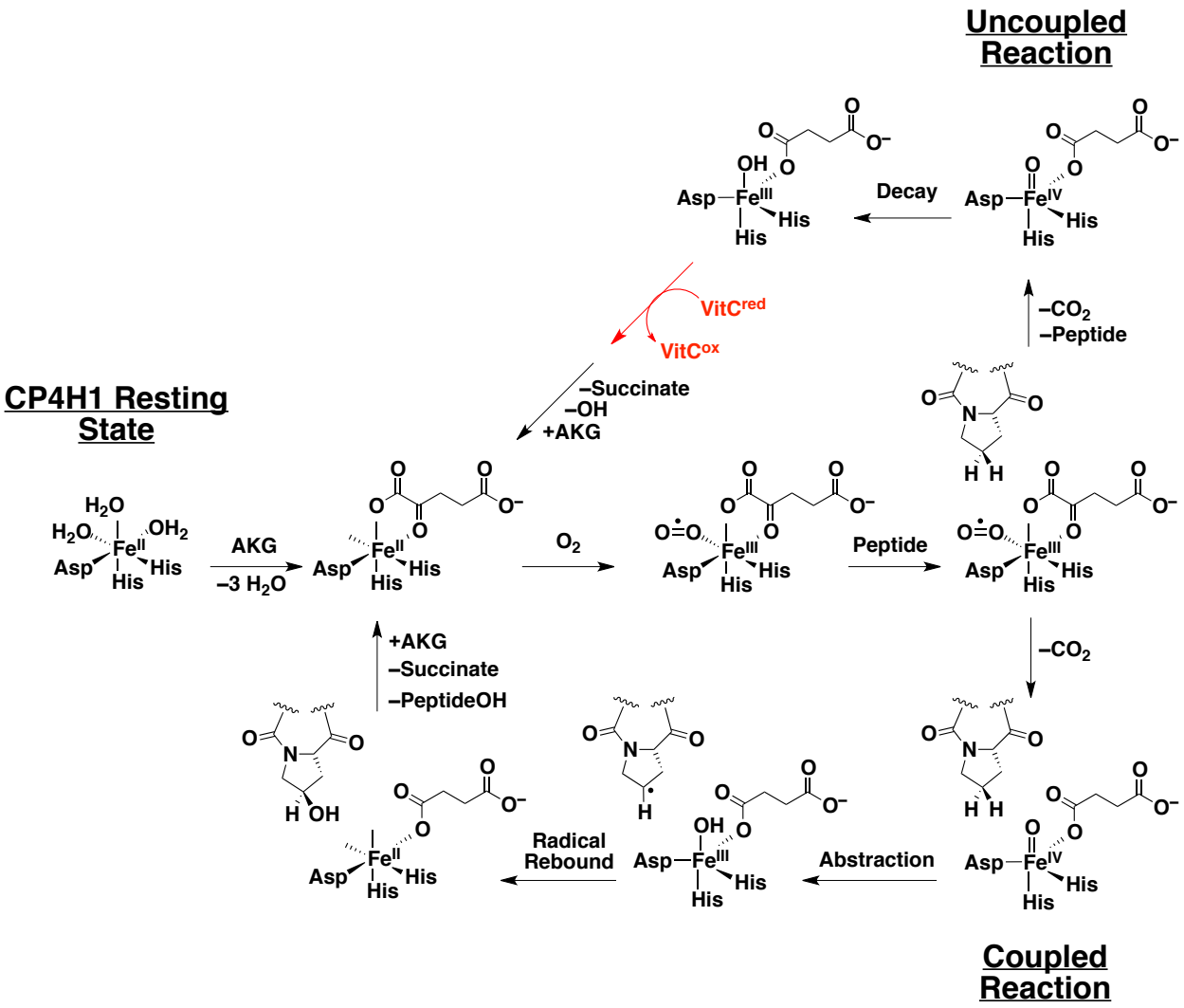


Figure 4.2 The mechanism of the overall reaction catalyzed by CP4Hs. CP4Hs have previously been reported to display an ordered ter-ter mechanism in which AKG first binds to CP4H·Fe(II) complex, after which O_2 and the peptide substrate bind in an ordered fashion.[98] Subsequently, the oxidative decarboxylation of AKG facilitates the formation of a highly reactive Fe(IV)=O species (ferryl ion) that hydroxylates the peptide substrate via a radical rebound process.[56] While all CP4Hs appear to require $VitC^{red}$ for activity, this cofactor is not required during the typical ter-ter mechanism, but rather, has been shown to operate during the "uncoupled reaction" in which CP4H decarboxylates AKG in the absence of the peptide substrate.[102] As suggested by EPR experiments, the ferryl ion appears to decay via an unknown mechanism to an inactive Fe(III) state,[101] and $VitC^{red}$ is required to reduce the iron center back to the Fe(II) state required for catalysis.[101,102] In some cases (particularly for the uncoupled reaction), the ligand environment of the iron center and exact order of binding is not known. Thus, this model is meant to provide a framework for considering the activation of CP4H1 by bipymonoCs.

Figure 4.3 Iron binding properties of bipymonoC activators

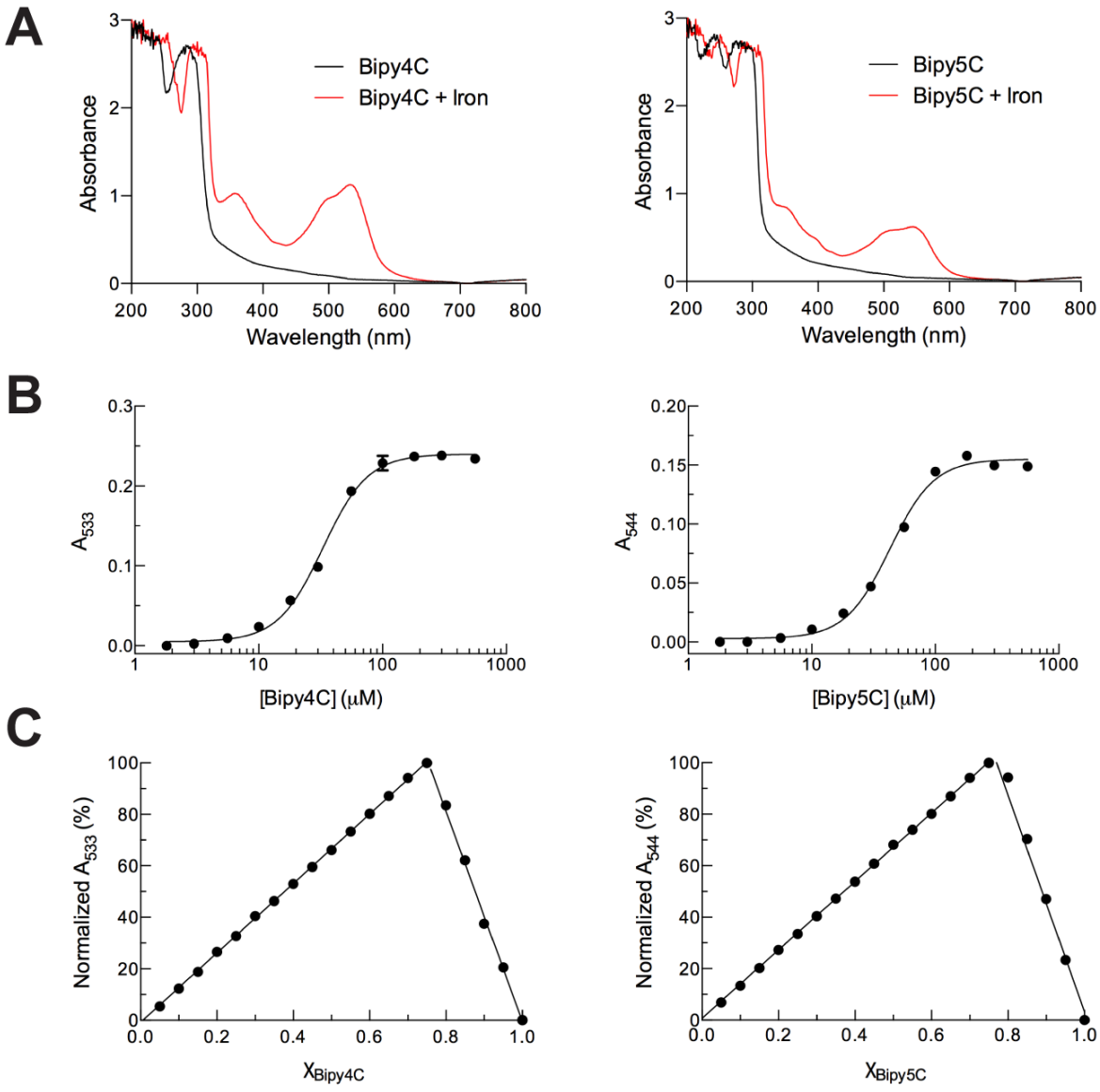


Figure 4.3 Iron binding properties of bipymonoC activators. (A) Absorption spectra of bipymonoCs (300 μM) in the presence and absence of Fe(II)SO_4 (100 μM). Complexes of bipy4C and bipy5C with Fe(II) showed local absorption maxima (λ_{max}) of 533 nm and 544 nm, respectively. (B) Ligand titrations for bipymonoC complexes with Fe(II) were carried out as described in the Experimental Procedures Section. Dose-response plots were fit using the dose-response equation, and the associated Fe_{20}-EC_{50} values for bipy4C and bipy5C complexes were 33 ± 0.8 μM and 43 ± 2 μM, respectively. (C) Job's plots for the bipy4C and bipy5C complexes of Fe(II) were constructed as described in the Experimental Procedures section. χ-values at the point intersection were 0.75 ± 0.02 and 0.76 ± 0.06 for bipy4C and bipy5C, respectively, suggesting the formation of Fe(ligand)$_3$ type complexes.

Figure 4.4 Structure activity relationships of bipymonoCs and related chemical analogues for the activation of human CP4H1

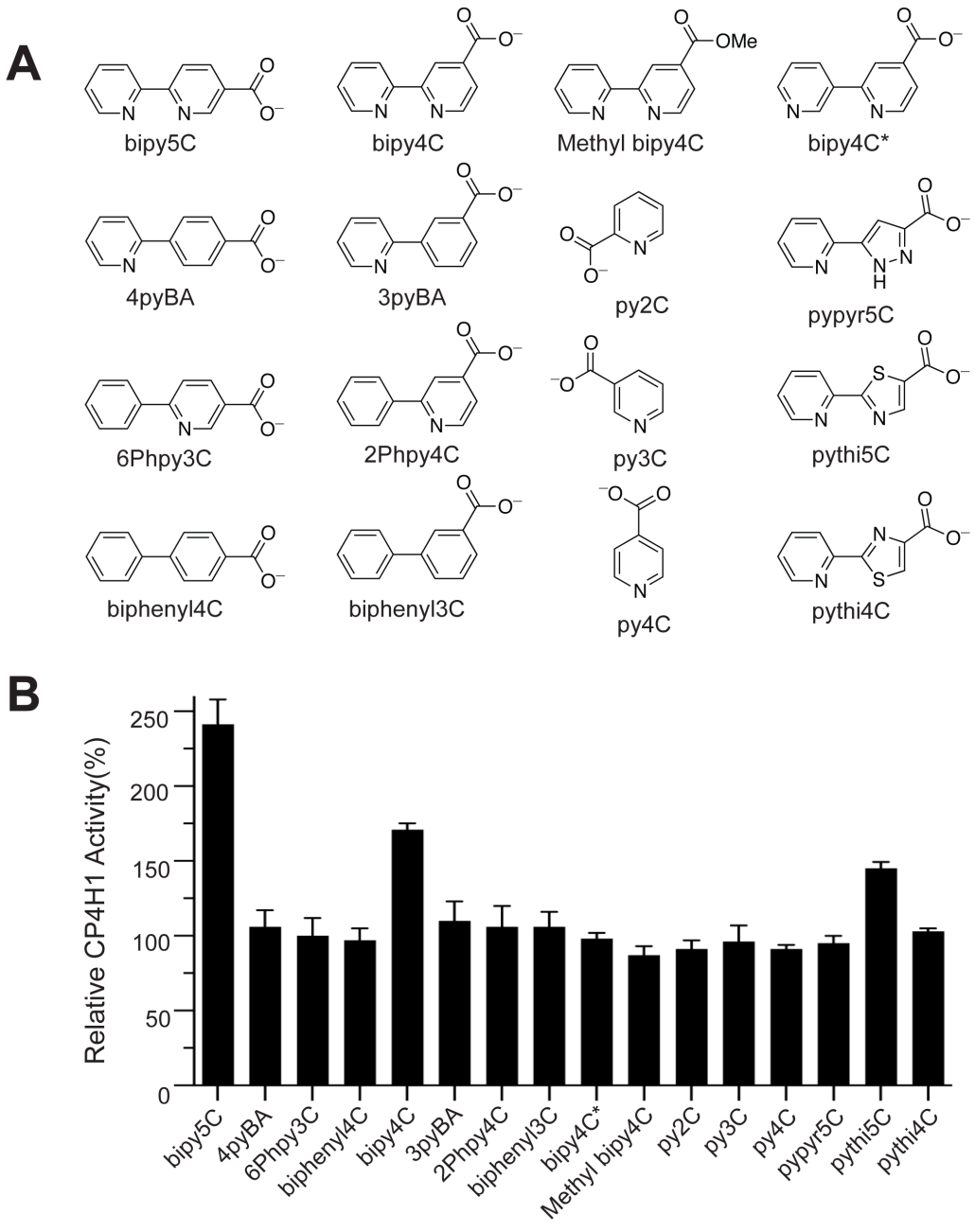


Figure 4.4 Structure activity relationships of bipymonoCs and related chemical analogues for the activation of human CP4H1. (A) Chemical structures and abbreviations. (B) Screen (10 μM) of the compounds depicted in panel A against human CP4H1. Compounds were screened as described in the Experimental Procedures section using 100 μM AKG. Relative activity values are the mean (± SE) of three independent experiments and represent the ratio of CP4H activity observed in the presence and absence of the activator or chemical analogue.

Figure 4.5 Preincubation of bipy4C with iron abolishes activation of CP4H1

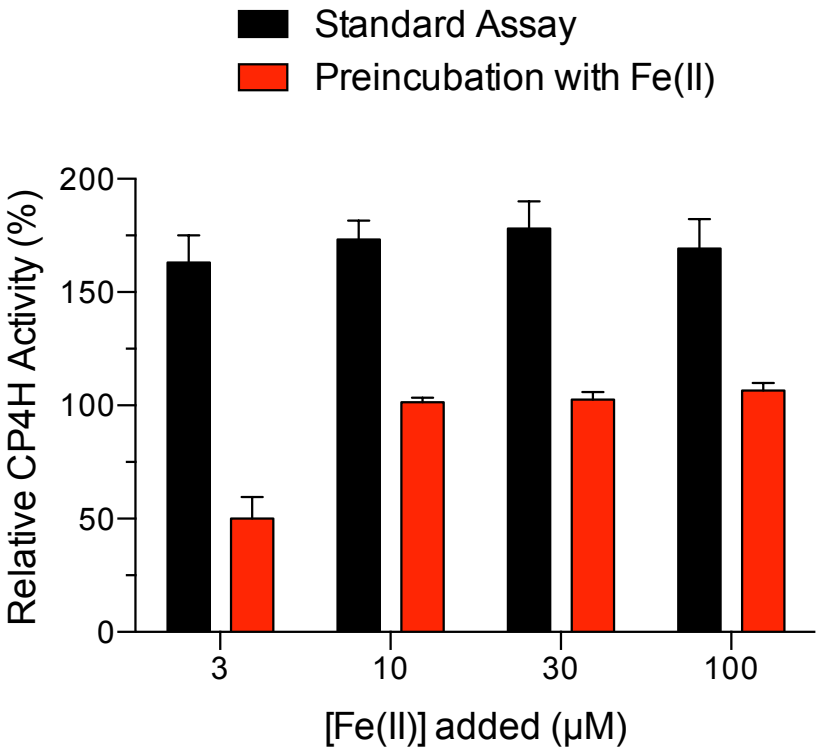


Figure 4.5 Preincubation of bipy4C with iron abolishes activation of CP4H1. CP4H assays were conducted as described in the Experimental Procedures section using 100 μM AKG and 10 μM bipy4C that was either added last (standard) or preincubated with $Fe(II)SO_4$ (as denoted) prior to addition to the assay. Relative activity values are the mean (± SE) of three independent experiments and represent the ratio of CP4H activity observed in the presence and absence of activator.

Figure 4.6 Postulated mechanism for activation of CP4H1 by bipymonoCs

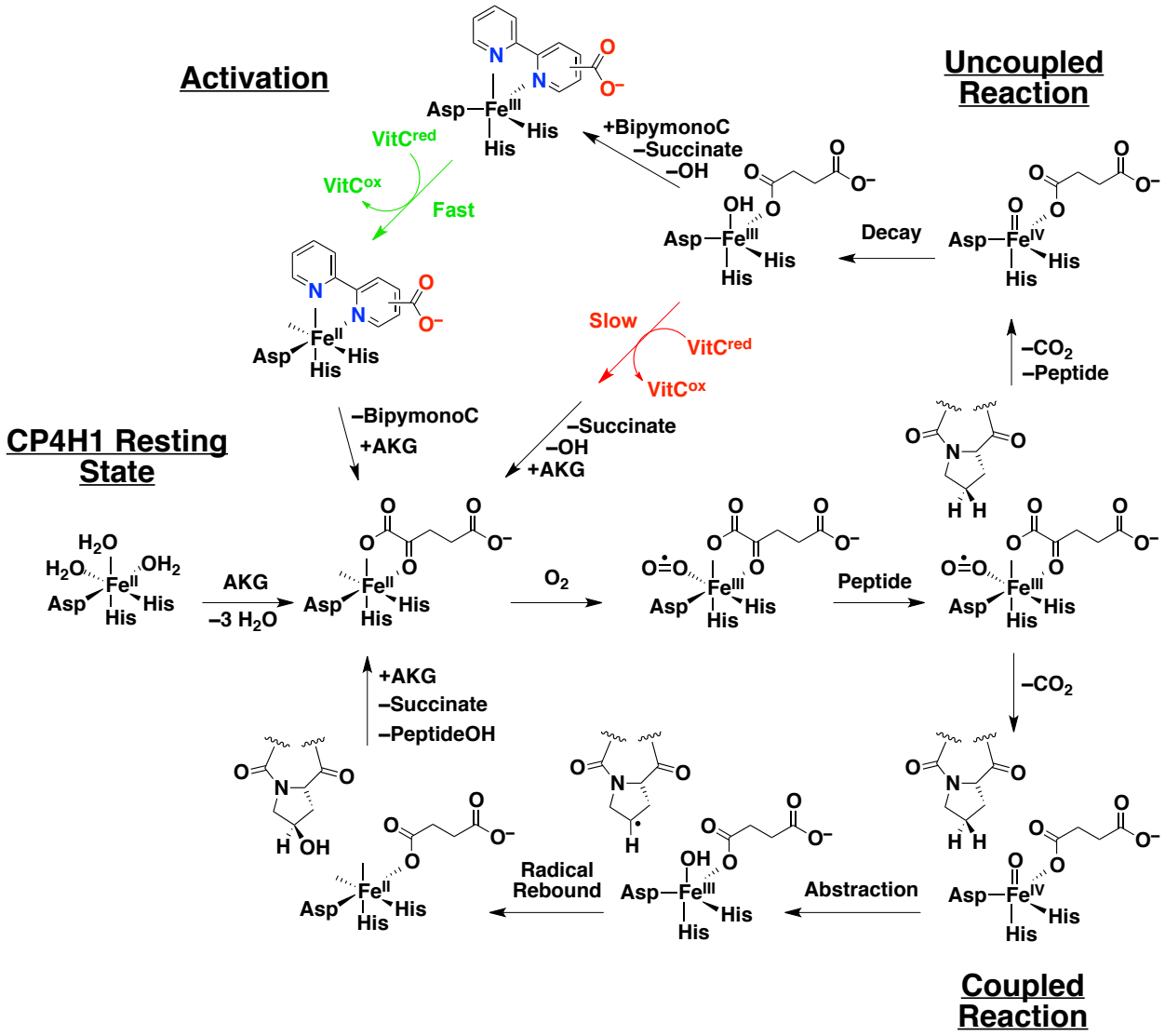


Figure 4.6 Postulated mechanism for activation of CP4H1 by bipymonoCs. Catalysis by CP4H1 and related FAKGD enzymes proceeds through a 2-stage mechanism in which the AKG cosubstrate is first oxidatively decarboxylated to form the Fe(IV)=O species (ferryl ion), after which the ferryl ion normally hydroxylates the peptide substrate via a radical rebound process.[56] The rate-determining step is currently not known. In the absence of the peptide substrate (i.e complete absence, poor positioning, dissociation, etc…) the ferryl ion can decay via the "uncoupled reaction" to an Fe(III) species,[101] which requires the cofactor ascorbate ($VitC^{red}$) for reduction back to the catalytically active Fe(II) state.[101,102] Based upon the information presented above, we postulate that bipymonoCs activate CP4H1 during the uncoupled reaction by serving as ligands that enhances the rate of the $VitC^{red}$-dependent reduction of the iron center. Moreover, this mechanism implies that the catalytic step requiring $VitC^{red}$ is the rate-determining step of the overall reaction catalyzed by CP4H.

Figure 4.7 Activation of CP4H1 by bipymonoCs is competitive with succinate

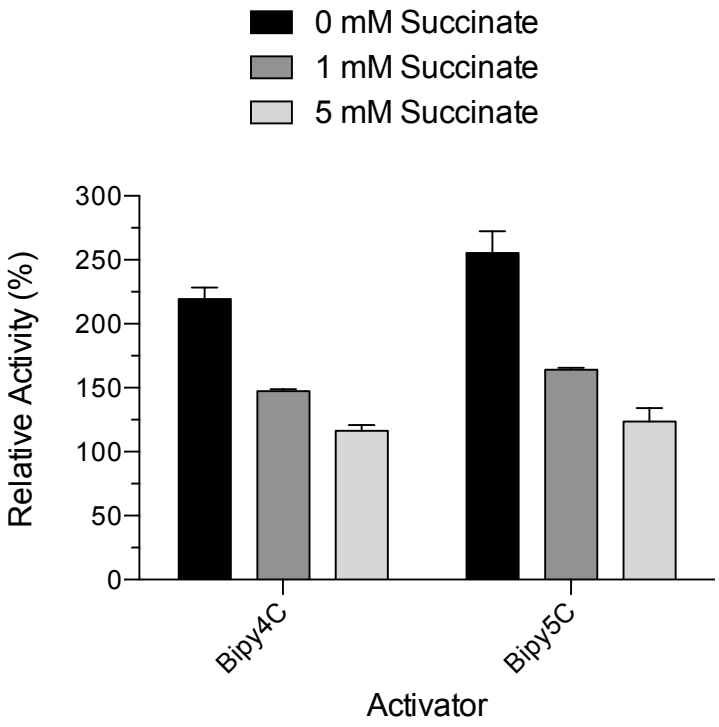


Figure 4.7 Activation of CP4H1 by bipymonoCs is competitive with succinate. CP4H assays were conducted as described in the Experimental Procedures section using 1000 μM AKG and 100 μM bipymonoC activator. Relative activity values are the mean (± SE) of three independent experiments and represent the ratio of CP4H activity observed in the presence and absence of the activator.

Figure 4.8 Inhibition or activation by biheteroarylcarboxylates and related analogues is dependent upon AKG and succinate

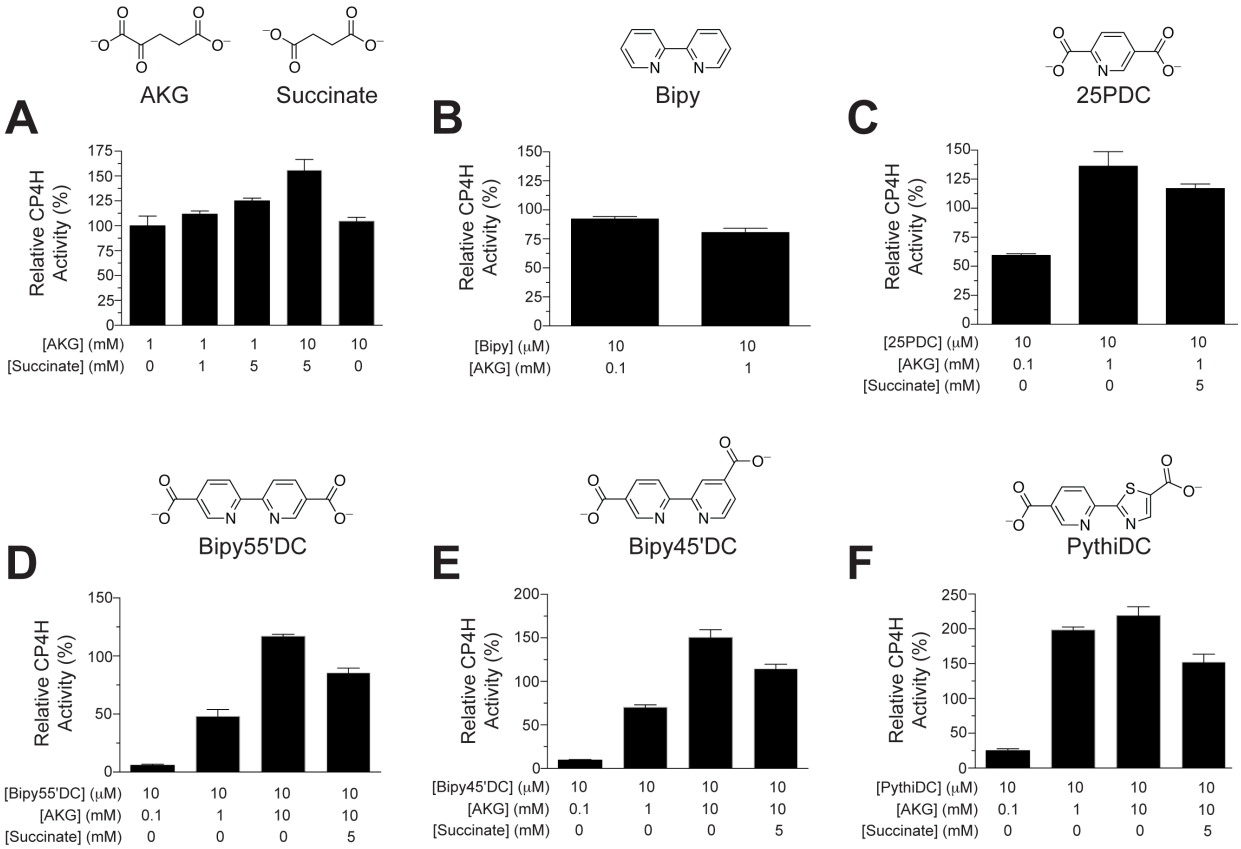


Figure 4.8 Inhibition or activation by biheteroarylcarboxylates and related analogues is dependent upon AKG and succinate. In all cases, reactions catalyzed by CP4H1 were conducted as described in the Experimental Procedures section using $VitC^{red}$ at a concentration of 2 mM. The concentrations of other relevant assay components are as denoted in each panel. Relative activity values are the mean (± SE) of three independent experiments and represent the ratio of CP4H activity observed in the presence and absence of the modulator in question. (A) The effect of added succinate and AKG on CP4H1 activity. (B) The effect of bipy on CP4H1 activity. (C) The effect of 25PDC on CP4H1 activity and dependence upon added AKG and succinate. (D, E, F) The effect of bipy55′DC (D), bipy45′DC (E), and pythiDC (F) on CP4H1 activity and dependence upon added AKG and succinate.

Figure 4.9 A general mechanism for the modulation of human CP4H1 activity by biheteroarylcarboxylate ligands

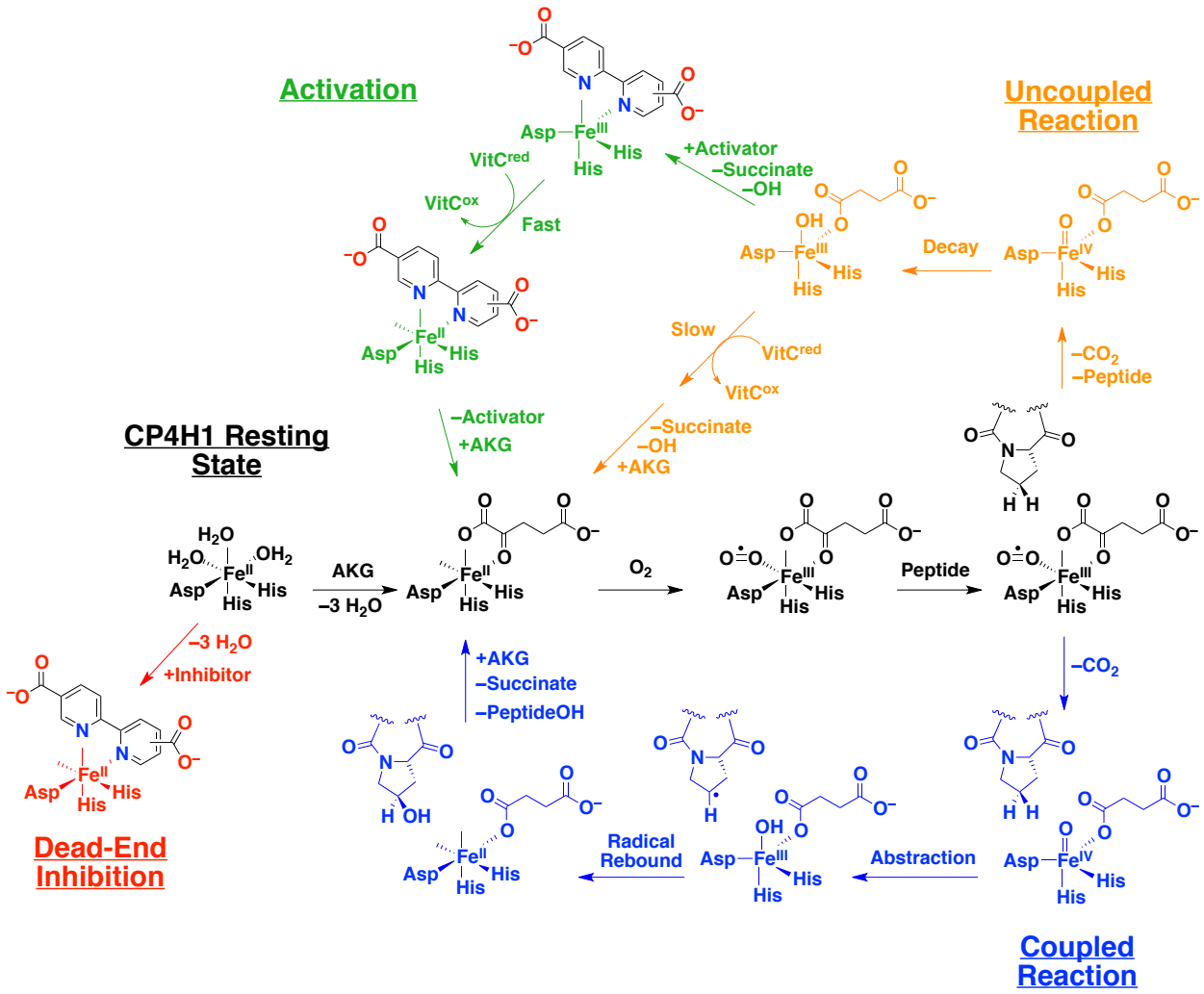


Figure 4.9 A general mechanism for the modulation of human CP4H1 activity by biheteroarylcarboxylate ligands. CP4H1 catalysis proceeds by the 2-stage mechanism presented common to many related P4Hs and FAKGD enzymes.[56] After its generate in stage 1, the ferryl ion can either facilitate the hydroxylation of the peptide substrate in the coupled reaction, or decay to an inactive Fe(III) species in the uncoupled reaction, ultimately requiring rescue by the cofactor $VitC^{red}$ in what is likely the rate determining step of the overall reaction. Common inhibitors that are analogues of the AKG cosubstrate, such as bipyridinedi(mono)carboxylates compete for binding to the Fe(II) species to form dead-end inhibited complexes. However, many of the same compounds can likely serve as ligands to increase the rate of the $VitC^{red}$-dependent reduction of the inactive Fe(III) species. The observed effect of such compounds on catalysis of CP4H1 is thus the net effect of for the activation and inhibition, with the relative contributions of each dependent on a variety of parameters including dose, the concentration of other cosubstrates (AKG), cofactors ($VitC^{red}$), and byproducts (succinate), the relative affinities of the compound of interest to the Fe(II) and Fe(III) species, and likely other factors.

Figure 4.10 Activation of prolyl 4-hydroxylases by bipymonoCs

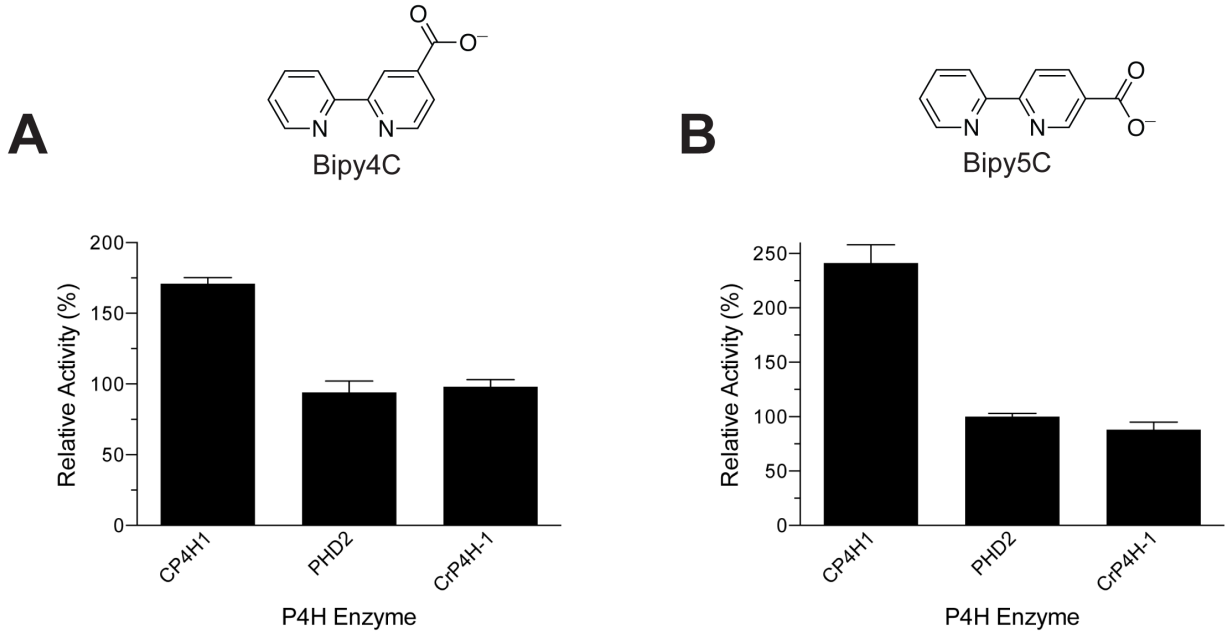


Figure 4.10 Activation of prolyl 4-hydroxylases by bipymonoCs. Human CP4H1. Human PHD2, and *C. reinhardtii* P4H-1 (CrP4H-1) were assayed in the presence or absence of bipy4C (A) or bipy5C (B) at 10 μM under the standard assay conditions described in the Experimental Procedures section. Relative activity values are the mean (± SE) of three independent experiments and represent the ratio of CP4H activity observed in the presence and absence of the activator in question.

CHAPTER 5

Probing the Substrate Conformational Preference of Human Collagen Prolyl 4-Hydroxylase Using Prolyl Peptide Bond Isosteres*

*This chapter has been prepared for publication in part under the same title. Vasta, J. D., Choudhary, A., Jensen, K. H., McGrath, N. A., and Raines, R. T. Probing the substrate conformational preferences of human collagen prolyl 4-hydroxylase using prolyl peptide bond isosteres. J. D. Vasta synthesized model peptide substrates **5.1** and **5.2**. J. D. Vasta, A. Choudhary, K. H. Jensen, and N. A. McGrath all contributed to the synthesis and characterization of corresponding alkene isosteres **5.3**–**5.6**. J. D. Vasta designed and carried out all experiments related to the assay of prolyl 4-hydroxylases. J. D. Vasta, A. Choudhary, K. H. Jensen, and R. T. Raines contributed to the analysis of data and preparation of the manuscript and figures.

5.1 Abstract

Collagen prolyl 4-hydroxylases (CP4Hs) catalyze one of the most abundant posttranslational modifications: the hydroxylation of (2*S*)-proline (Pro) residues in protocollagen strands. The ensuing (2S,4*R*)-4-hydroxyproline is essential for the folding, secretion, and overall stability of the collagen triple helix. Whereas many features of the CP4H catalyzed reaction are well characterized, the precise conformational preferences of CP4Hs with respect to the protocollagen peptide substrate are still unclear. Previous reports have suggested that CP4Hs might recognize a turn conformation in the peptide substrate, and that the ring pucker and/or peptide bond conformation in the prolyl peptide bond of the triplet repeat might be an important determinant. Herein, we report our studies using alkene isosteres of the *cis* and *trans* conformations of the prolyl peptide bond to probe the conformational preferences of human CP4H1. While neither alkene isostere served as a substrate for human CP4H1, the *cis* isostere served as an inhibitor with significantly greater potency that that of the *trans* isostere, suggesting that the *cis* conformation in the prolyl peptide bond is recognized preferentially. Moreover, comparative studies with a related P4H from *Chlamydomonas reinhardtii*, an enzyme with a similar catalytic domain but lacking an N-terminal peptide substrate binding domain, showed a similar preference for the *cis* conformation. Our findings support the hypothesis that CP4Hs preferentially recognize the *cis* conformation in the prolyl peptide bond of collagenous polypeptides in a manner that appears to be localized in the catalytic domain, and inform the design of selective CP4H inhibitors based upon the peptide substrate.

5.2 Introduction

Collagen is the principal component of bone, connective tissues, and the extracellular matrix in animals.[1] The overproduction of collagen is associated with a variety of diseases, including fibrotic diseases[33] and cancers.[28,29,46,47,73] The stability of collagen relies on posttranslational modifications that occur throughout the secretory pathway.[2] By far the most prevalent of these modifications is the hydroxylation of collagen strands by collagen prolyl 4-hydroxylases (CP4Hs), which are Fe(II)- and α-ketoglutarate (AKG)-dependent dioxygenases (FAKGDs) located in the lumen of the endoplasmic reticulum.[4] Catalysis by CP4Hs converts (2*S*)-proline (Pro) residues in protocollagen strands into (2*S*,4*R*)-4-hydroxyproline (Hyp) residues, which are essential for the conformational stability of mature collagen triple helices.[5] Importantly, CP4Hs are validated targets for treating both fibrotic diseases[30] and metastatic breast cancer.[29]

Like all enzymes of the FAKGD superfamily, catalysis by CP4Hs requires Fe(II) and the cosubstrates AKG and dioxygen.[56,97,145] The Fe(II) is bound by a conserved His-X-Asp/Glu…X_n…His motif, and AKG chelates to enzyme-bound Fe(II) using its C-1 carboxylate and C-2 keto groups, while the C-5 carboxylate group forms hydrogen bonds and engages in Coulombic interactions with a basic residue (typically, arginine or lysine). All FAKGDs are believed to effect catalysis through a similar two-stage mechanism in which AKG is first decarboxylated oxidatively to generate a highly reactive Fe(IV)=O species (ferryl ion), which effects the hydroxylation of a hydrocarbon via a radical rebound process.[56,97,145]

In vertebrates, CP4Hs exist as $\alpha_2\beta_2$ tetramers. In these tetramers, the α-subunit contains the C-terminal catalytic domain and N-terminal peptide-substrate-binding domain (PSB domain), and the β-subunit is protein disulfide isomerase, which is a multifunctional protein that is responsible for maintaining the α-subunit in a soluble and active conformation.[4] Three isoforms of the α-subunit, α(I), α(II), and α(III), have been identified in humans.[4] All α-subunit isoforms form tetramers with the β-subunit, which we refer to herein as the CP4H1, CP4H2, and CP4H3 holoenzymes. CP4H1 is the most prevalent of the isoforms, and has been characterized most extensively. Whereas the structure of the tetrameric complex is unknown, those of the individual domains of the α(I)-subunit have provided insight into the interaction of CP4Hs with the protocollagen substrate, as well as the dimerization of the α(I)-subunits to facilitate formation of the tetramer.[84-86,174]

The development of CP4H inhibitors has been of interest since the mid-1970s. Much attention has been paid to developing inhibitors for CP4Hs that mimic the cosubstrate AKG (see Chapters 1–3 above). Still, it has long been noted that inhibitors based upon the primary protocollagen substrate could be of interest, and might offer improved selectivity compared to that of the AKG mimics. Although many features of the reaction catalyzed by CP4Hs are well characterized, a detailed understanding of the manner in which the protocollagen substrate is recognized by these enzymes remains incomplete, due partially to the limited structural information available for mammalian CP4Hs. The structure of the N-terminal PSB domain has been characterized by NMR spectroscopy and X-ray crystallography, and is known to interact with peptides in the polyproline type-II (PPII) helical conformation, using specific tyrosine

residues spaced appropriately to recognize multiple iterations of collagen triplet repeats.[84-86] Moreover, these same structural models have also suggested that Y-position prolyl residues in polymeric peptide substrates are bound in hydrophobic pockets in the PSB, which provides at least a partial explanation for the manner in which CP4Hs avoid inhibition by the corresponding hydroxylated products.[84] Thus, the PSB domain facilitates multivalent interactions with substrates in the PPII conformation, and helps the enzyme avoid product inhibition after hydroxylation.

The conformation of the peptide substrate that is recognized during catalysis is unknown. Other researchers have attempted to address this issue using small collagen mimetic peptides. For example, a variety of biochemical and biophysical investigations have suggested that β-turns in collagen mimetic peptides are sites of prolyl hydroxylation.[175-177] Other conformational preferences are less clear. Because proline is a secondary amine and prolyl peptide bonds are tertiary amides, the *cis* isomer of prolyl peptide bonds (Figure 5.1A) occurs much more frequently than in peptide bonds between non-prolyl residues.[178] Moreover, the conformation of the prolyl peptide bond is correlated with another conformational feature of proline: the pucker of its pyrrolidine ring (Figure 5.1B).[179,180] In a previous report, we used collagen mimetic peptides containing proline analogues that affect both the *cis*/*trans* ratio of the prolyl peptide bond and the ring pucker of the pyrrolidine ring, finding that substrates for human CP4H1 tended to have greater preference for the Cγ-e*ndo* ring pucker and *cis* prolyl peptide bond. While intriguing, these results were only correlative, as none of the model substrates were fixed in a particular conformation.

The above finding with respect to the possible importance of a *cis* prolyl peptide bond in substrates for CP4H is of particular interest herein, as *cis* prolyl peptide bonds are known to be prevalent in procollagens,[181] and *cis*/*trans* isomerization is thought to be the rate-limiting step in the folding of many proteins with high proline content, including collagens.[182-184] Thus, the prevalence of the *cis* prolyl peptide bond in protocollagen substrates for CP4H could be an important determinant for hydroxylation. In this work, we generate small collagen mimetic peptides containing alkene isosteres of the prolyl peptide bond and demonstrate that human CP4H1 and a related P4H from the algae *Chlamydomonas reinhardtii* both recognize the *cis* alkene isostere relative to the *trans*. Moreover, we demonstrate that recognition of the *cis* alkene isostere likely occurs in the catalytic domain. These results provide new insight into how this essential enzyme recognizes its substrate.

5.3 Results and Discussion

Synthesis of Alkene Isosteres

The primary substrates for CP4Hs *in vivo* are protocollagens and other large polymers (i.e. elastin).[4] These large protein substrates are not amenable to most chemical analyses. Thus, we chose to employ a model system based upon the small collagen mimetic tetrapeptide Gly-Pro-Pro-Gly-OEt (GPPGOEt, Figure 5.2A), with corresponding alkene isosteres GP*trans*PGOEt (Figure 5.2B) and GP*cis*PGOEt (Figure 5.2B). Small tetrapeptides of this nature that contain one collagen-like triplet repeat have been found previously to be suitable substrates for

CP4Hs.[122,140,148,175,185] Moreover, due to their small size, they are much more amenable for chemical investigations.[122] Lastly, we chose to utilize model peptides capped (Figure 5.2C) with an N-terminal dansyl sulfonamide, which have been reported previously as substrates for human CP4H,[140,148] with N-terminal CBz-capped peptides[128,186] for supplemental analyses.

The model substrates dansylGPPGOEt (**5.1**) and CBzGPPGOEt (**5.2**) were synthesized by standard solution–phase peptide coupling procedures as described previously for similar peptides.[148] We next synthesized dansylGP*trans*PGOEt (**5.3**) and CBzGP*trans*PGOEt (**5.4**), employing synthetic routes (Scheme 5.1) that were adapted from previous reports.[187-191] The Boc group in aldehyde **5.9** can be in either the *cis* or the *trans* conformation. The facial selectivity in conversion of aldehyde **5.9** to alcohol **5.10** is particularly striking. In the *cis* conformation of aldehyde **5.9**, the bulky *tert*-butyl group blocks the non-reactive face. This face is blocked by the donor carbonyl oxygen in the *trans* conformation. The key step in the synthesis involved a stereoselective Ireland–Claisen rearrangement of **5.11** to **5.13**. The coupling of H_2N-Gly-OEt to **5.17** employed a hindered base, collidine. Lastly, deprotection of the N-terminal Boc group of **5.18** followed by subsequent coupling of dansyl-Gly-OH or CBz-Gly-OH yielded the final products **5.3** or **5.4**, respectively.

The corresponding *cis* alkene isosteres dansylGP*cis*PGOEt (**5.5**) and CBzGP*cis*PGOEt (**5.6**) proved to be more challenging synthetic targets. Inspired by the synthetic routes developed by Etzkorn and co-workers for similar systems,[192,193] we developed the route described in Scheme 5.2 to prepare these *cis* alkene isosteres. A variety of reducing agents were evaluated for the reduction of ketone **5.21** to alcohol **5.22** including L-Selectride, K-Selectride, $NaBH_4$,

DIBALH, and $LiAlH_4$. Unfortunately, the best reducing agent was $LiAlH_4$, which afforded a nearly equal mixture of the diastereometric alcohols. The absolute stereochemistry of alcohol **5.22** was established using Mosher's acid. Although it is possible to isolate the desired diastereomer of alcohol **5.22**, separation of diastereomers following alkylation to **5.23** is more facile. The key step in the synthesis involved a Still–Wittig rearrangement of **5.23** to **5.24** with transfer of chirality. As for the synthesis of the *tran*s isosteres, the coupling of H_2N-Gly-OEt to **5.26** again employed collidine. Lastly, deprotection of the N-terminal Boc group of **5.27** followed by subsequent coupling of dansyl-Gly-OH or CBz-Gly-OH yielded the final products **5.5** or **5.6**, respectively.

Interaction of Model Peptides and Alkene Isososteres with Human CP4H

With the model peptides and corresponding alkene isosteres in hand, we next set out to probe the interaction of these peptides with human CP4H, either as substrates or inhibitors of the enzyme. First, we tested the model peptides (**5.1** and **5.2**) as substrates for human CP4H, employing a standard UPLC assay adapted from a similar HPLC assay described previously.[122,140,148] Initial tests confirmed that both dansylGPPGOEt and CBzGPPGOEt were substrates for human CP4H1, as expected from previous reports.[140,148] Surprisingly, neither the *cis* nor the *trans* alkene isosteres (**5.3**–**5.6**) served as substrates for the enzyme (data not shown), regardless of the N-terminal cap used or concentration range tested (0.1–3 mM for dansyl isosteres and 0.1–10 mM for CBz isosteres).

Although isosteres **5.3**–**5.6** did not serve as substrates for human CP4H, we reasoned that they could still serve as inhibitors of human CP4H. Using the dansyl model peptides, we found that both dansylGP*trans*PGOEt and dansylGP*cis*PGOEt served as inhibitors of human CP4H (Figure 5.3A), with the *cis* isostere being significantly more potent than the *trans*. In subsequent dose–response experiments, the inhibition curves for both dansyl isosteres were found to be sigmoidal (Figure 5.3B), with IC_{50} values of (75 ± 6) μM and (410 ± 30) μM for the *cis* and *trans* isosteres, respectively. Moreover, this same trend where the *cis* isostere inhibits with greater potency than the *trans* isostere was also observed with the CBz-capped isosteres **5.4** and **5.6** (Figure 5.4), suggesting that the trend observed is independent of the N-terminal group. Due to more reliable quantitation and sensitivity compared to the CBz-peptides, we proceeded with the dansyl isosteres **5.3** and **5.5** in subsequent analyses.

With an interest in characterizing further the inhibition observed for the dansyl isosteres, we next considered their kinetic mechanism. CP4Hs have been shown previously to display an ordered ter–ter mechanism in which AKG first binds to the CP4H·Fe(II) complex, after which O_2 and the peptide substrate bind in an ordered fashion.[98] If the isosteres inhibited by binding in a manner that is similar to a substrate, they would be expected to display competitive inhibition with respect to that peptide substrate. Thus, we attempted to perform Lineweaver–Burke analyses to characterize the inhibition by the dansyl isosteres versus the dansylGPPGOEt model peptide. We found that dansylGPPGOEt did not display Michaelis–Menten kinetics, but rather a pattern recognizable as substrate inhibition (Figure 5.5A), which was even more apparent in the corresponding Lineweaver–Burke plot (Figure 5.5B). Substrate inhibition had been observed

previously for a CP4H when using the longer collagen mimetic peptide substrate $(ProProGly)_{10}$.[98] Moreover, substrate inhibition by $(ProProGly)_{10}$ was reported to be competitive with the cosubstrate AKG.[98] Substrate inhibition by dansylGPPGOEt in our assays was, however, observed even at saturating concentrations of AKG, suggesting that the inhibitory binding mode of dansylGPPGOEt could not be eliminated in this manner. Unfortunately, the limited solubility of these tetrapeptides prevented us from fully characterizing the observed substrate inhibition mechanism for dansylGPPGOEt, as well as any affect that the dansyl isosteres may have had on that substrate inhibition.

Given the complexities in the kinetics with respect to the peptide substrate, we reasoned that if ordered binding still occurred in the kinetic mechanism, it would be expected that the dansyl isosteres would display uncompetitive inhibition versus the AKG cosubstrate. We performed a Lineweaver–Burke analysis using *cis* isostere **5.5** as a model inhibitor, but found an inhibition pattern consistent with noncompetitive inhibition versus the AKG cosubstrate (Figure 5.6A). The slope (5.6B) and intercept (5.6C) replots for the noncompetitive inhibition were linear, and K_{is} and K_{ii} values of (56 ± 13) μM and (18 ± 8) μM, respectively, were extracted from the slope (K_{is}) and intercept (K_{ii}) replots after linear regression analysis. A value for K_{is} of (56 ± 13) μM is in reasonable agreement with the IC_{50} value of (76 ± 13) μM reported above for *cis* isostere **5.5**. The observation of noncompetitive inhibition for *cis* isostere **5.5** versus AKG could suggest that either this isostere binds in a CP4H subsite that is independent of the site for hydroxylation, or that there is not a strict ordered binding of the small collagenous tetrapeptides

compared to other substrates such as AKG and dioxygen. Kinetics alone cannot distinguish between these two possibilities.

Preliminarily, the inhibition pattern reported above is consistent with the hypothesis that human CP4H preferentially binds to the *cis* conformation compared to the *trans* conformation in collagen-like peptides. The CP4H α-subunit is composed of multiple domains including a C-terminal catalytic domain as well as an N-terminal PSB domain,[4,30] and both the C-terminal catalytic domain and PSB domain likely interact with the typical polymeric substrate. Hence, the manner in which CP4H distinguishes these conformations and the putative location in which the alkene isosteres interact is still unknown.

As described above, the structure of the PSB domain has been characterized, and is known to interact with peptides in the PPII helical conformation with trans peptide bonds.[84-86] Given this information, it is not surprising that human CP4H has some affinity for *trans* isostere **5.3**, as this peptide isostere is likely to adopt an extended conformation similar to a PPII helix. Given its locked conformation, however, *cis* isostere **5.5** is less likely to interact favorably in the PSB domain, which raises the question of whether or not the C-terminal catalytic domain contains the subsite that preferentially recognizes the *cis* conformation.

To address this question, we sought out a more simple P4H model system. We chose to investigate a related P4H from the algae *Chlamydomonas reinhardtii* (CrP4H-1).[90] CrP4H-1 is a soluble, structurally characterized,[91,92] 29-kDa monomer and is 26% identical in amino acid sequence to the catalytic domain of human CP4H1.[90] As FAKGD enzymes are generally conserved at the level of fold,[56] a sequence identity of 29% is significant, coupled with the

observation that all essential active site residues are conserved in the two enzymes.[90] Moreover, although substrates for algal P4Hs such as CrP4H-1 are typically proline- and serine- rich polymers of the algal cell wall *in vivo*, CrP4H-1 has been observed to hydroxylate collagen-like peptides, albeit with less positional specificity in its prolyl hydroxylation (*i.e.*, CrP4H-1 hydroxylates proline residues in both the X- and Y- position of the collagen tripeptide repeat).[90]

Considering the similarities to human CP4H1 described above, CrP4H-1 can be viewed as a reasonable model system for the catalytic domain of the human CP4H1 alpha subunit, without the complexities of the N-terminal PSB domain or the CP4H β-subunit. Thus, we next sought to investigate the interaction of dansyl peptides **5.1**, **5.3**, and **5.5** with CrP4H-1. Under assay conditions similar to that used for that used to assay human CP4H1, we found that model peptide **5.1** served as a substrate for CrP4H-1, albeit with a substrate inhibition phenomenon similar to that observed above for human CP4H1 (Figure 5.7). Again, we did not observe hydroxylation of either of the dansyl isosteres **5.3** or **5.5** by CrP4H-1 (data not shown). W did find, however, that both dansyl isosteres **5.3** and **5.5** served as inhibitors of CrP4H-1. We observed an inhibition pattern virtually identical to that observed for human CP4H1 with *cis* isostere **5.5** displaying significantly greater potency than that of *trans* isostere **5.3** (Figure 5.8). Similar to human CP4H1, these data suggest that CrP4H-1 preferentially recognizes the *cis* conformation in peptide substrate analogues.

5.4 Conclusions and Caveats

Given the similarities between human CP4H1 and the algal CrP4H-1 described above, as well as CrP4H-1 being essentially a monomeric P4H catalytic domain, our data cumulatively suggest that both P4Hs preferentially recognize the *cis* conformation in peptide substrate analogues in a manner that only requires the catalytic domain. This finding is consistent with a hypothesis that CP4Hs preferentially recognize the *cis* conformation in their peptide substrates, which was put forth previously.[122]

Still, neither the *cis* nor *trans* isostere served as a substrate for either P4H. Hence, it is not clear if catalysis requires a *cis* conformation. For example, it could be that CP4Hs and related enzymes preferentially recognize the *cis* conformation in their polymeric peptide substrates using a distinct subsite that is located in the catalytic domain, but not necessarily located in the CP4H active site. In this manner, the conformational preference could be used for binding affinity, rather than recognition of the specific proline residue to be hydroxylated. Another possibility is that CP4Hs hydroxylate prolyl residues adjacent to *cis* prolyl peptide bonds, but a conformational change in the substrate is required for hydroxylation. In this case, perhaps neither the *cis* nor the *trans* alkene isostere would be expected to be a substrate because these peptides are locked in a particular subset of conformations and unable to undergo the conformational change required for catalysis. A third possibility is that the alkene isosteres do not serve as substrates because of inherent chemical differences between the isosteres and their corresponding model peptides. Further studies using a combination of computational modeling and/or structural approaches should look to determine more precisely the conformation of the

alkene isosteres in terms of the relevant ring puckers and other parameters, which would facilitate a more meaningful comparison between the isosteres and their corresponding model peptides.

Although neither the *cis* nor the *trans* isostere served as a substrate for human CP4H, the *cis* isostere served as a more potent inhibitor, suggesting that the *cis* conformation of the prolyl peptide bond is recognized preferentially. While these results have not clarified exactly which conformation is hydroxylated by CP4H, this data still provides valuable knowledge about how this essential enzyme recognizes collagenous peptides as well as information in support of the end goal of developing selective inhibitors for CP4Hs.

5.5 Experimental Procedures

5.5.1 General

All reagent chemicals were obtained from commercial sources (Sigma–Aldrich, Acros, Combi-Blocks, Oakwood Products, Enamine, Bachem, or Novabiochem) and used without further purification. All glassware was flame- or oven-dried, and reactions were performed under N_2(g) unless indicated otherwise. DCM and toluene were dried over a column of alumina. Dimethylformamide was dried over alumina and further purified through an isocyanate scrubbing column. Other anhydrous solvents were obtained in septum-sealed bottles. Flash chromatography was performed with columns of 40–63 Å silica gel, 230–400 mesh (Silicycle, Québec City, Canada). Thin-layer chromatography (TLC) was performed on plates of EMD 250 μm silica 60-F_{254} with visualization by UV light or staining with $KMnO_4$. The phrase "concentrated under reduced pressure" refers to the removal of solvents and other volatile materials using a rotary evaporator at water aspirator pressure (<20 torr) while maintaining water-bath temperature below 40 °C. Residual solvent was removed from samples at high vacuum (<0.1 torr). The term "high vacuum" refers to vacuum achieved by a mechanical belt-drive oil pump. All reported yields are unoptimized.

5.5.2 Instrumentation

NMR spectra were acquired at ambient temperature with a Bruker DMX-400 Avance spectrometer or a Bruker Avance 500i spectrometer at the National Magnetic Resonance Facility

at Madison (NMRFAM) and were referenced to TMS or a residual protic solvent. Some compounds exist as either mixtures of rotomers or tautomers that do not interconvert on the NMR timescale at ambient temperature and therefore exhibit multiple sets of NMR signals (as indicated). Electrospray ionization (ESI) and electron ionization (EI) mass spectrometry were performed with a Micromass LCT® or Micromass AutoSpec® instruments, respectively, from Waters (Milford, MA) at the Mass Spectrometry Facility in the Department of Chemistry at the University of Wisconsin–Madison. The progress of reactions catalyzed by prolyl 4-hydroxylases was determined by analytical UPLC using an Acquity UPLC® H-Class from Waters (equipped with an Acquity photodiode array detector, Acquity quaternary solvent manager, Acquity sample manager with a flow-through needle, and Empower 3 software). Preparative HPLC was performed using a Prominence HPLC instrument from Shimadzu (Kyoto, Japan) equipped with two LC-20AP pumps, a SPD-M20A photodiode array detector, and a CTO-20A column oven. The concentration of dansyl peptides was determined using a Cary 60 UV–Vis Spectrometer from Agilent Technologies (Santa Clara, CA) using an extinction coefficient of 3,900 $M^{-1}cm^{-1}$. Protein concentrations were calculated from their absorbance at 280 nm as measured with a NanoVue Plus spectrophotometer from GE Healthcare using an extinction coefficient of 290,000 $M^{-1}cm^{-1}$ for human CP4H,[140] and 44,000 $M^{-1}cm^{-1}$ for *Chlamydomonas reinhardtii* P4H (estimated using EXPASY assuming all cysteines oxidized). IC_{50}-values were calculated from experimental data with Prism version 6.0 from GraphPad Software (La Jolla, CA).

5.5.3 Production of Recombinant Human CP4H1

Human CP4H containing the α(I) isoform was produced heterologously in Origami B(DE3) *Escherichia coli* cells and purified as described previously.[140]

5.5.4 Assay of Human CP4H1 Activity in the Presence of Inhibitors

The catalytic activity of human CP4H1 was assayed as described previously.[140] Briefly, activity assays were carried out at 30 °C in 100 μL Tris–HCl buffer (50 mM), pH 7.8, containing human CP4H1 (100 nM), inhibitor (0–3000 μM), substrate (dansylGlyProProGlyOEt, 0–1860 μM), $FeSO_4$ (50 μM), BSA (1 mg/mL), catalase (0.1 mg/mL), ascorbate (2 mM), DTT (100 μM), and α-ketoglutarate (13–1000 μM). Unless otherwise noted, reactions were prepared by adding concentrated stock solutions of each component to concentrated assay buffer in the following order: $FeSO_4$, DTT, ascorbate, BSA, catalase, CP4H1, peptide substrate, and inhibitor (or vehicle). Solutions thus prepared were pre-incubated for 2 min at 30 °C, after which the reaction was initiated by the addition of α-ketoglutarate. After 15 min, reactions were quenched by boiling for 45 s and centrifuged at 10,000*g*. The supernatant (5–10 μL) was injected into an Acquity UPLC® BEH C18 column (2.1 × 50 mm, 1.7 μm particle size) from Waters. The column was eluted at 0.6 mL/min with a gradient (2.9 min) of 20%–68% aqueous acetonitrile containing 0.1% v/v TFA. The absorbance of the eluent was monitored at 289 nm. All assays were performed in triplicate. Data is reported as activity relative to control reactions lacking inhibitor, where activity is determined from the percent conversion of substrate to product. Dose-response

curves were generated for each inhibitor by plotting the relative activity versus the log of the inhibitor concentration. IC_{50}-values for each inhibitor were interpolated from the dose-response curves by non-linear regression using the sigmoidal dose-response function available in Prism. For subsrate inhibition by dansylGlyProProGlyOEt, data was fit to the substrate inhibition function available in Prism.

5.5.5 Production of Recombinant *Chlamydomonas reinhardtii* P4H-1

A cDNA encoding *C. reinhardtii* P4H-$1_{30\text{-}245}$ possessing an N-terminal hexahistidine (His_6) tag ($NHis_6$-CrP4H-1) under the control of the T7 promoter was cloned into a pET-22b(+) vector using the Gibson strategy.[142] The expression vector was transformed into Origami B(DE3) cells by electroporation, and transformants were grown on LB agar containing kanamycin (15 μg/mL), ampicillin (100 μg/mL), and tetracycline (12.5 μg/mL). A starter culture of TB medium containing antibiotics as described above was inoculated with a fresh colony, after which the cells were grown overnight at 37 °C with shaking at 200 rpm. The starter culture was used to inoculate 4 L of TB medium supplemented with antibiotics as described above to an OD_{600} of 0.02. The culture was incubated at 37 °C with shaking at 200 rpm until an OD_{600} of 0.9 was reached. At this time, protein expression was induced by the addition of isopropyl-1-thio-β-D-galactopyranoside (IPTG, 500 μM), and the culture was grown for 18 h at 21°C with shaking at 200 rpm. Cells were harvested by centrifugation and the cell pellet (6.0 g) was resuspended in 30 mL lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 2.5 mM imidazole, pH 7.4). The cells

were lysed at 22 kPSI in a T series Cell Disrupter 2.2 kW from Constant Systems Limited (Northants, UK). Insoluble material was cleared from the lysate by centrifugation at 30,000*g* for 45 min, and $NHis_6$-CrP4H-1 was purified from the soluble fraction by nickel affinity chromatography using a 5 mL HisTrap FF column (GE Healthcare, Piscataway, NJ). The column was equilibrated with lysis buffer, after which the supernant was injected and the A_{280} of the flowthrough was monitored until returning to baseline. $NHis_6$-CrP4H-1 was eluted from the column with an imidazole gradient (2.5–500 mM), with most of the $NHis_6$-CrP4H-1 eluting between 50–250 mM imidazole. Fractions containing purified $NHis_6$-CrP4H-1 (as determined by 12% SDS-PAGE) were combined, dialyzed against 10 mM Tris–HCl, 100 mM glycine, 100 mM NaCl, pH 7.8, and concentrated to 3 mg/mL using a Vivaspin concentrator with a molecular weight cut off of 8,000 Da. Aliquots were flash frozen in liquid N_2 and stored at −80 °C until use in P4H activity assays. This protocol yielded $NHis_6$-CrP4H-1 at ~24 mg per gram of cell pellet.

5.5.6 Assay of *C. reinhardtii* P4H-1 Activity in the Presence of Inhibitors

The catalytic activity of $NHis_6$-CrP4H-1was assayed similarly to that described above in section 5.5.4 for human CP4H1 with a few modifications. Briefly, activity assays were carried out at 30 °C in 100 μL HEPES–HCl buffer (50 mM), pH 6.8, containing $NHis_6$-CrP4H-1 (1.2 μM), inhibitor (0–100 μM), substrate (dansylGlyProProGlyOEt, 500 μM), $FeSO_4$ (200 μM), BSA (1 mg/mL), catalase (0.1 mg/mL), ascorbate (2 mM), DTT (100 μM), and α-ketoglutarate (1 mM). Reactions were prepared by adding concentrated stock solutions of each component to

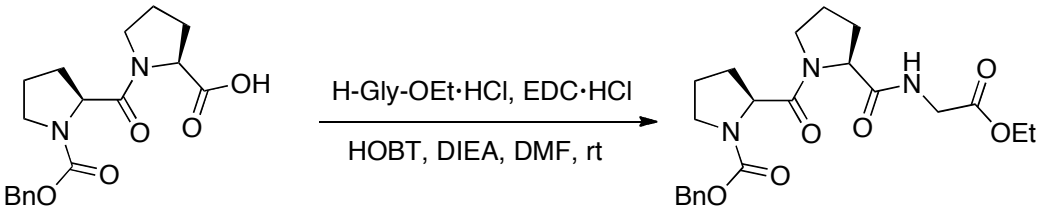
concentrated assay buffer in the following order: $FeSO_4$, DTT, ascorbate, BSA, catalase, $NHis_6$-CrP4H-1, peptide substrate, and inhibitor (or vehicle). Solutions thus prepared were pre-incubated for 2 min at 30 °C, after which the reaction was initiated by the addition of α-ketoglutarate. After 1 h, reactions were quenched by the addition of EDTA (0.5 M, 1 μL) and boiled for 45 s. Quenched reactions were centrifuged at 10,000*g* for 5 min, after which the supernatant (5–10 μL) was injected into an Acquity UPLC® BEH C18 column (2.1 × 50 mm, 1.7 μm particle size) from Waters. The column was eluted at 0.6 mL/min with a gradient (2.9 min) of 20%–68% aqueous acetonitrile containing 0.1% v/v TFA. The absorbance of the eluent was monitored at 289 nm. All assays were performed in triplicate. Data is reported as activity relative to control reactions lacking inhibitor, where activity is determined from the percent conversion of substrate to product.

5.5.7 Synthetic Procedures

Synthesis of Model Tetrapeptide Substrates

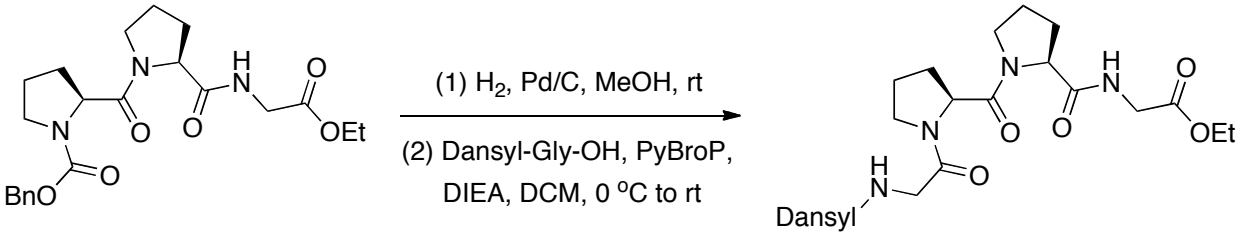
The model tetrapeptide substrates dansylGPPGOEt (**5.1**) and CBzGPPGOEt (**5.2**) were synthesized by standard solution phase peptide coupling procedures as described previously for similar peptides. To avoid redundancy, only synthetic procedures and spectral data for new compounds are provided below.

_N_-Benzyloxycarbonyl-(2_S_)-prolyl-(2_S_)-prolylglycine Ethyl Ester (CbzProProGlyOEt)



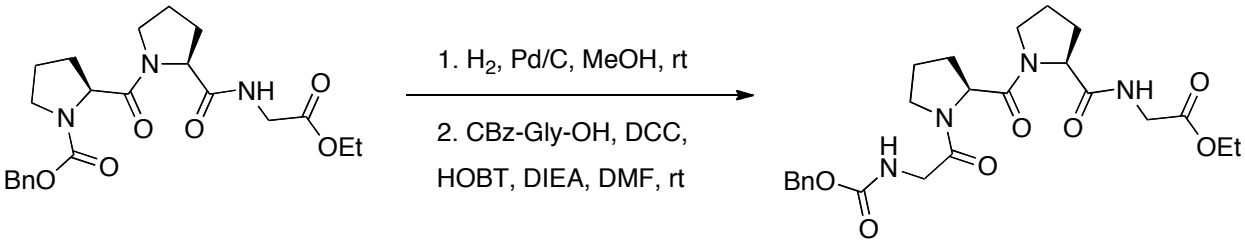
CbzProProGlyOEt was synthesized as described previously.[148] The spectral data and yield are in agreement with that reported previously.

_N_-Dansylglycyl-(2_S_)-prolyl-(2_S_)-prolylglycine Ethyl Ester (DansylGlyProProGlyOEt, 5.1)



DansylGlyProProGlyOEt was synthesized as described previously.[148] The spectral data and yield are in agreement with that reported previously

_N_-Carboxybenzyloxyglycyl-(2_S_)-prolyl-(2_S_)-prolylglycine Ethyl Ester (CbzGlyProProGlyOEt, 5.2)



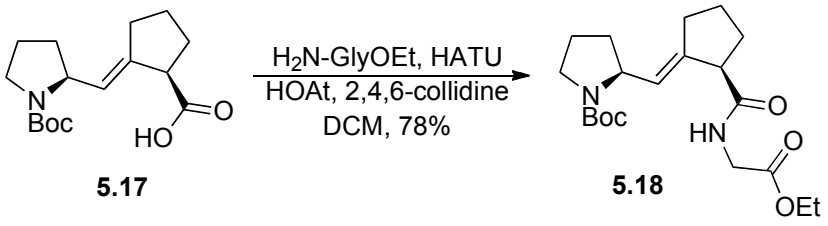
A suspension of CbzProProGlyOEt (1.08 g, 2.5 mmoles) and Pd/C (114 mg) in MeOH (25 mL) was stirred under an atmosphere of H_2(g) for 3 h. The mixture was filtered through a pad of Celite®, and concentrated under reduced pressure to a pale yellow viscous oil (835 mg). A portion of the oil was carried on to the subsequent coupling step. The oil (682 mg, 2.29 mmoles) was dissolved in anhydrous DMF (20 mL) to which Cbz-Gly-OH (436 mg, 2.09 mmoles), 1-hydroxybenzotriazole hydrate (288 mg, 2.09 mmoles), and *N*,*N*′-dicyclohexylcarbodiimide (430 mg, 2.09 mmoles) were added and the suspension stirred at room temperature under an atmosphere of N_2(g). DIEA (1.2 mL) was added and the reaction was stirred at room temperature for another 34 hr. The reaction was filtered through filter paper and concentrated under reduced pressure, after which the residue was dissolved in EtOAc (50 mL) and washed with 5% $KHCO_3$ (2 x 50 mL), 5% $KHSO_4$ (2 x 50 mL) and brine (1 x 50 ml). The organics were dried over Na_2SO_4(s) and concentrated under reduced pressure to afford a crude yellow oil. The crude product was purified by chromatography on silica (20% MeOH in EtOAc) and HPLC (30–90% v/v acetonitrile in H_2O with 0.1% v/v TFA) to afford the title compound (282 mg, 28%) as a yellow oil. **^{1}H NMR** (400 MHz, $CDCl_3$, mixture of 3 or more rotomers, integrations are approximate, *δ*) 10.3 (s, 4.4 H), 8.38 (t, *J* = 5.6 Hz, 1.0 H), 7.53–7.28 (m, 17.3 H), 7.13 (m, 0.2 H), 6.12–5.95 (m, 0.6 H), 5.87 (m, 1.5 H), 5.71 (m, 0.7 H), 5.13–5.05 (m, 6.0 H), 4.66–4.42 (m, 5.3 H), 4.20–3.89 (m, 16.2 H), 3.79–3.39 (m, 11.9 H), 2.97 (s, 0.2 H), 2.53–2.44 (m, 1.0 H), 2.28–1.78 (m, 21.1 H), 1.26–1.20 (m, 7.5 H); **^{13}C NMR** (100 MHz, $CDCl_3$, mixture of 3 or more rotomers, *δ*) 172.3, 172.0, 171.8, 171.7, 171.6, 171.3, 171.2, 171.0, 169.6, 169.1, 168.0, 167.9, 159.4, 159.0, 158.6, 158.2, 156.7, 156.4, 136.2, 128.5 (2 signals), 128.2, 128.1, 128.0, 127.9,

116.3, 113.5, 67.3, 67.2, 67.0, 61.6, 61.4, 61.3, 61.1, 61.0, 60.1, 60.0 59.2, 58.4, 47.5, 47.3 (2 signals), 46.7, 46.5, 43.3, 43.2, 42.6, 42.5, 41.5, 41.3, 31.8, 31.7, 28.4, 28.3, 28.1, 27.7, 25.1, 24.7, 22.1, 22.0, 14.0; **HRMS** (ESI) *m/z* 511.2155 [calc'd for $C_{24}H_{32}N_4O_7Na$ $(M + Na)^+$ 511.2164].

Synthesis of *Trans* Alkene Isosteres

The syntheses of dansylGlyPro*trans*ProGlyOEt (**5.3**) and CBzGlyPro*trans*ProGlyOEt (**5.4**) were adapted from previous reports[187-191] using the route described in Scheme 5.1. To avoid redundancy, only synthetic procedures and spectral data for new intermediates and the final targets are provided below.

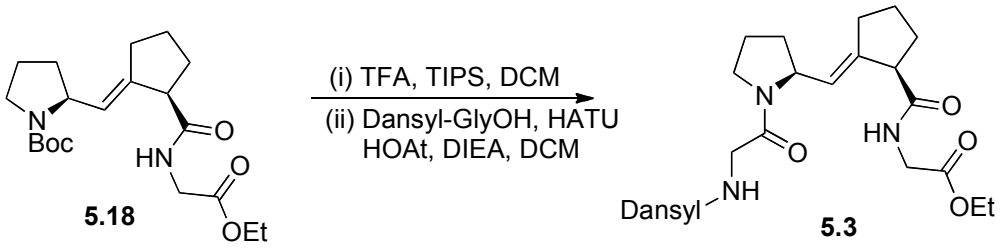
(E)-(*S*,*S*)-*N*-(2-(*N*-Benzylpyrrolidin-2-ylmethylene)-cyclopentyl)carbonyl]glycine Ethyl Ester (5.18)



Acid **5.17** (100 mg, 0.34 mmoles) was dissolved in anhydrous DCM (6.8 mL) after which glycine ethyl ester hydrochloride (52 mg, 0.373 mmoles), 1-hydroxy-7-azobenzotriazole (138 mg, 1.01 mmoles), and 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate (386 mg, 1.01 mmoles) were added and the suspension was stirred at

room temperature under an atmosphere of N_2(g). 2,4,6-Collidine (0.22 mL, 1.70 mmoles) was added and the reaction was stirred at room temperature for 30 min. Brine was added to the reaction mixture, after which it was extracted with EtOAc. The combined organics were dried over Na_2SO_4(s) and concentrated under reduced pressure to afford a crude oil. The crude product was further purified by chromatography on silica (30–100% EtOAc in hexanes) to afford the title compound (60 mg, 47%) as a pale yellow oil, which was taken on slightly crude to the next step. **^{1}H NMR** (400 MHz, $CDCl_3$, δ) 5.50 (bs, 1 H), 4.38 (bs, 1 H), 4.21 (app q, J = 7.6 Hz, 2 H), 4.02–4.00 (m, 2 H), 3.44–3.39 (m, 2 H), 3.26 (bs, 1 H), 2.57–2.48 (m, 1 H), 2.32–2.24 (m, 1 H), 2.14–2.04 (m, 2 H), 1.95–1.67 (m, 6 H), 1.60 (s, 3 H), 1.42 (s, 9 H), 1.28 (t, *J* = 7.2 Hz, 3 H).

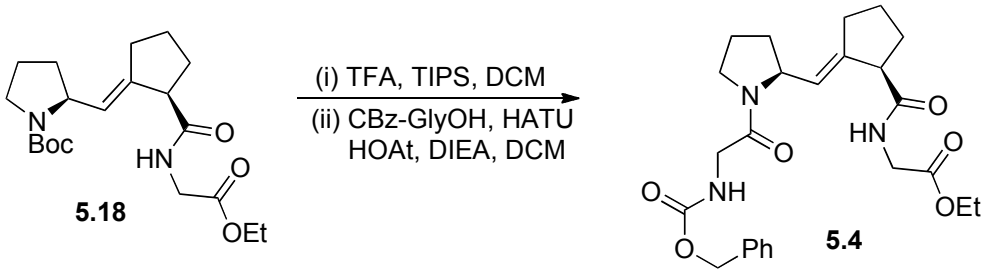
DansylGlyPro*trans*ProGlyOEt (5.3)



Alkene **5.18** (60 mg, 0.16 mmoles) was dissolved in dry DCM (0.88 mL) in a vial, after which triisopropylsilane (0.08 mL, 0.16 mmoles) was added. TFA (0.88 mL) was added and the reaction was stirred at room temperature for 1 h. The reaction was concentrated under reduced pressure, and the residue resuspended in dry DCM (3.2 mL). Dansyl-Gly-OH (49 mg, 0.16 mmoles), 1-hydroxy-7-azobenzotriazole (65 mg, 0.16 mmoles), and 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate

(180 mg, 0.16 mmoles) were added and the reaction mixture was stirred at room temperature. DIEA (0.14 mL) was added and the reaction was stirred at room temperature for 1 h. The reaction was concentrated under reduced pressure, after which the crude product was purified by chromatography on silica (70% EtOAc in hexanes) and HPLC 25–55% v/v acetonitrile in H_2O with 0.1% v/v TFA) to afford the title compound (32 mg, 35%) as a pale yellow solid. **^{1}H NMR** (500 MHz, $CDCl_3$, mixture of 2 or more rotomers, integrations are approximate, δ) 8.55 (d, J = 9.0 Hz, 1.0 H), 8.33 (dd, J = 2.0, 9.0 Hz, 1.0 H), 8.22 (d, J = 7.0 Hz, 0.5 H), 8.16 (d, J = 6.5 Hz, 0.5 H), 7.58 (q, J = 8.0 Hz, 1.0 H), 7.53–7.48 (m, 1.0 H), 7.20 (dd, J = 5.0, 7.5 Hz, 0.9 H), 6.98 (s, 0.1 H), 6.47 (t, J = 5.0 Hz, 0.5 H), 6.18 (t, J = 5.0 H, 0.5 H), 5.87–5.84 (m, 0.9 H), 5.32–5.27 (m, 1.0 H), 4.50–4.46 (m, 0.5 H), 4.22–4.12 (m, 2.4 H), 4.06–3.94 (m, 2.0 H), 3.66 (d, J = 4.5 Hz, 1.1 H), 3.62 (d, J = 4.5 Hz, 0.4 H), 3.57 (d, J = 4.0 Hz, 0.3 H), 3.54 (d, J = 4.0 Hz, 0.2 H), 3.49–3.38 (m, 1.0 H), 3.34–3.30 (m, 0.5 H), 3.25–3.15 (m, 0.9 H), 3.09–3.06 (m, 0.5 H), 2.89 (s, 6.0 H), 2.57–2.50 (m, 0.5 H), 2.36–1.96 (m, 4.2 H), 1.93–1.59 (m, 13.0 H), 1.43 (s, 0.6 H), 1.29–1.25 (m, 5.0 H). 0.93–0.82 (m, 0.5 H), 0.07 (s, 0.6 H); **^{13}C NMR** (125 MHz, $CDCl_3$, mixture of 2 or more rotomers, δ) 174.2, 173.5 170.5, 170.3, 165.9, 165.2, 152.0, 151.9, 144.8, 144.2, 134.6, 134.5, 130.9, 130.8, 130.2 (2 signals), 130.1, 129.7, 129.5, 129.0, 128.9, 124.4 (2 signals), 123.4, 123.2, 119.6, 119.5, 115.9, 115.7, 61.9 (2 signals), 57.5, 57.0, 52.3, 52.0, 47.0, 46.0, 45.8, 45.0, 44.4, 41.8, 33.8, 31.5, 31.0, 30.7 (2 signals), 30.0, 29.4, 29.2, 25.3, 25.2, 24.7, 23.0, 14.5, 1.4; **HRMS** (ESI) m/z 571.2580 [calc'd for $C_{29}H_{39}N_4O_6S$ $(M+H)^+$ 571.2585].

CBzGlyPro*trans*ProGlyOEt (5.4)



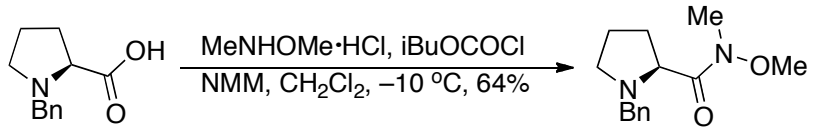
Alkene **5.18** (87 mg, 0.23 mmoles) was dissolved in dry DCM (1.3 mL) in a vial, after which triisopropylsilane (0.094 mL, 0.46 mmoles) was added. TFA (1.3 mL) was added and the reaction was stirred at room temperature for 1 h. The reaction was concentrated under reduced pressure, and the residue resuspended in dry DCM (4.6 mL). CbzGlyOH (48 mg, 0.23 mmoles), 1-hydroxy-7-azobenzotriazole (94 mg, 0.69 mmoles), and 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate (262 mg, 0.69 mmoles) were added and the reaction mixture was stirred at room temperature. DIEA (0.6 mL) was added and the reaction was stirred at room temperature for 1 h. The reaction was concentrated under reduced pressure, after which the crude product was purified by HPLC (30–90% v/v acetonitrile in H_2O with 0.1% v/v TFA) to afford the title compound (45 mg, 42%) as a viscous yellow oil. **^{1}H NMR** (500 MHz, $CDCl_3$, mixture of 2 or more rotomers, integrations are approximate, δ) 7.38–7.32 (m, 7.5 H), 6.65 (t, J = 4.5 Hz, 1.0 H), 6.44 (t, J = 4.5 Hz, 0.7 H), 5.82–5.80 (m, 1.5 H), 5.53–5.47 (m, 1.6 H), 5.18–5.10 (m, 3.3 H), 4.66–4.62 (m, 1.0 H), 4.43–4.39 (m, 0.8 H), 4.26–4.18 (m, 3.3 H), 4.12 (d, J = 5.5 Hz, 0.4 H), 4.09–3.88 (m, 6.4 H), 3.65–3.43 (m, 4.6 H), 3.33–3.28 (m, 2.1 H), 2.69–2.62 (m, 1.0 H), 2.55–2.48 (m, 0.8 H), 2.40–2.20 (m, 2.8 H), 2.17–1.70 (m, 12.4 H), 1.33–1.27 (m, 4.7 H); **^{13}C NMR** (100 MHz, $CDCl_3$, mixture of 2 or more

rotomers, *δ*) 174.7, 174.1 170.0, 169.9, 167.8, 166.9, 156.5, 144.4, 143.5, 136.4, 129.9, 128.9, 128.7, 128.5, 128.1, 127.9 (2 signals), 124.5, 124.2, 68.3, 66.9, 61.5, 57.2, 56.9, 51.9, 51.7, 46.8, 45.9, 43.5, 42.8, 41.5, 33.6, 31.1, 30.7, 30.5, 29.2, 28.9, 25.0, 24.9, 24.5, 22.6, 14.1; **HRMS** (ESI) m/z 472.2442 [calc'd for $C_{25}H_{34}N_3O_6$ $(M+H)^+$ 472.2442].

Synthesis of *Cis* Alkene Isosteres

The alkene isosteres DansylGlyPro*cis*ProGlyOEt (**5.5**) and CBzGlyPro*cis*ProGlyOEt (**5.6**) were synthesized using the route described in Scheme 5.2. To avoid redundancy, only synthetic procedures and spectral for new intermediates and the final targets are provided below.

***N*-methoxy-*N*-methyl-1-(phenylmethyl)-(2*S*)-2-pyrrolidinecarboxamide (5.20)**

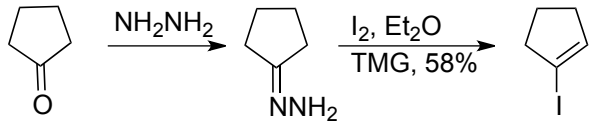


5.19 **5.20**

To a round bottom flask was added acid **5.19** (8.1 g, 39 mmol). The flask was evacuated and purged with Ar(g), after which DCM (100 mL) was added and the reaction mixture cooled to −10°C in an ice/NaCl bath with stirring. NMM (5.2 mL, 47 mmol) was added after which isobutylchoroformate (7.7 mL, 59 mmol) was added dropwise. The reaction mixture was stirred for 30 min, after which N,O-Dimethylhydroxylamine hydrochloride (5.8 mL, 59 mmol) was added. A second portion of NMM (9.4 mL, 86 mmol) was added and the reaction stirred at −10 °C for 20 min. The reaction mixture was allowed to warm to ambient temperature and stirred for

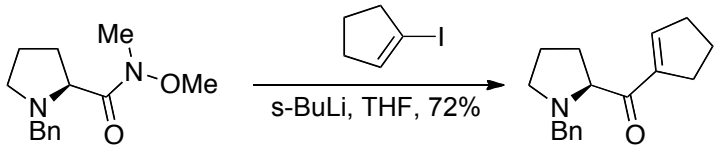
another 2 h, after which the mixture was diluted with DCM and washed with H_2O and brine. The crude organics were dried over Na_2SO_4(s), filtered, and concentrated under reduced pressure. The crude product was purified by chromatography on silica (1% MeOH + 0.25% NEt_3 in DCM, followed by 2% MeOH + 0.5% NEt_3 in DCM) to afford the title compound (6.2 g, 64%) as a yellow oil. **^{1}H NMR** (500 MHz, $CDCl_3$, δ) 7.35 (d, J = 7.3 Hz, 2 H), 7.30 (dd, J = 7.3, 7.3 Hz, 2 H), 7.24 (d, J = 7.1 Hz, 1 H), 3.94 (d, J = 12.8 Hz, 1 H), 3.57 (s, 3 H), 3.55 (d, J = 12.8 Hz, 1 H), 3.17 (s, 3 H), 3.10 (ddd, J = 8.5, 8.5, 2.8 Hz, 1 H), 2.44 (apparent q, J = 7.9 Hz, 1 H), 2.19–2.11 (m, 1 H), 1.97–1.75 (m, 4 H); **^{13}C NMR** (125 MHz, $CDCl_3$, δ) 175.1, 138.7, 129.5, 128.2, 127.1, 62.1, 61.3, 58.1, 53.0, 32.5, 29.2, 23.1; **HRMS** (ESI) m/z 249.1594 [calc'd for $C_{14}H_{20}N_2O_2$ $(M+H)^+$ 249.1598].

1-Iodocyclopentene



1-Iodocyclopentene was prepared via literature procedures[191] and resulted in identical spectra and comparable yields in each synthetic transformation.

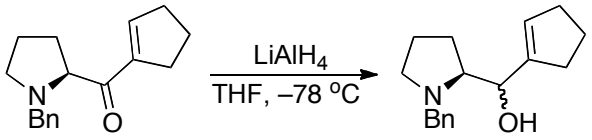
(*S*)-(*N*-Benzylpyrrolidin-2-yl)-cyclopent-1-enyl-methanone (5.21)



5.20 **5.21**

To a flame dried flask was added 1-Iodopentene (6.1 g, 31.3 mmol), after which the flask was evacuated and purged with Ar(g). Anhydrous THF (60 mL) was added to the flask, and the reaction mixture was cooled to −40°C in a dry ice/MeCN bath with stirring. A freshly titrated solution of sec-BuLi (1.02 M in cyclohexane, 30.7 mL, 31.3 mmol) was added dropwise, and the reaction mixture was stirred at −40°C for 90 min. to prepare a solution of cyclopentenyllithium. In a separate flask, weinreb amide **5.20** (4.6 g, 18.5 mmol) was dissolved in anhydrous THF (50 mL) and the reaction was cooled to −78 °C. The solution of cyclopentenyllithium was added dropwise, and the reaction mixture was allowed to warm to ambient temperature and stirred for 1 h. The reaction was quenched by the addition of saturated NH_4Cl. The quenched reaction was diluted with EtOAc. The combined organics were washed with NH_4Cl and brine, after which the organics were dried over Na_2SO_4(s), filtered, and concentrated under reduced pressure. The crude product was dissolved in benzene and frozen overnight, after which the product was purified by chromatography on silica (25% EtOAc + 0.5% NEt_3 in hexanes) to afford the title compound (1.65 g, 72%) as a brown oil. **^{1}H NMR** (400 MHz, $CDCl_3$, δ) 7.30−7.20 (m, 5 H), 6.88 (bs, 1 H), 3.86 (d, J = 12.7 Hz, 1 H), 3.66 (dd, J = 9.1, 7.4 Hz, 1 H), 3.39 (d, J = 12.7 Hz, 1 H), 3.09−3.03 (m, 1 H), 2.57−2.51 (m, 4 H), 2.33 (apparent q, J = 8.2 Hz, 1 H), 2.16−2.07 (m, 1 H), 1.98−1.74 (m, 5 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ) 199.7, 144.4, 144.0, 138.7, 129.5, 128.3, 127.2, 69.6, 58.8, 53.4, 34.4, 31.3, 30.3, 23.1, 22.5; **HRMS** (ESI) m/z 256.1703 [calc'd for $C_{17}H_{21}NO$ $(M+H)^+$ 256.1696].

(*S*,*S*)-(*N*-Benzylpyrrolidin-2-yl)-cyclopent-1-enyl-methanol (5.22)

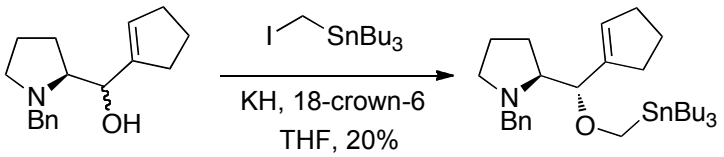


5.21 **5.22**

To a 250 mL flame-dried round bottom flask was added ketone **5.21** (3.036 g, 11.89 mmol) and a magnetic stir bar. The flask was purged with argon and anhydrous THF (50 mL) was added. The reaction mixture was cooled to −78 °C in a dry ice/acetone bath. To the cooled reaction mixture was added $LiAlH_4$ (3.69 g, 97.2 mmol) in three portions. The reaction mixture was stirred for 1 h and then warmed to ambient temperature. The reaction mixture was cooled to 0 °C in an ice bath, and quenched by the dropwise addition of MeOH followed by the addition of saturated NH_4Cl. The mixture was diluted with EtOAc and washed with saturated NH_4Cl and saturated sodium potassium tartrate. The combined aqueous phase was extracted thrice with CH_2Cl_2 (emulsions developed). The combined organic phase was washed with brine, dried over Na_2SO_4(s), filtered, and concentrated under reduced pressure to afford the title compound as a crude yellow oil (~1:1 mixture of diasteromers), which was used without further purification. Note: The desired diastereomer (*S*,*S*) can be isolated, but separation of the diasteromers following alcohol alkylation is more facile. (*S*,*S*)-diastereomer **^{1}H NMR** (500 MHz, $CDCl_3$, δ) 7.34–7.29 (m, 4 H), 7.27–7.24 (m, 1 H), 5.73 (bd, J = 1.4 Hz, 1 H), 4.01 (d, J = 12.1 Hz, 1 H), 3.94 (bs, 1 H), 3.44 (d, J = 12.9 Hz, 1 H), 3.02–2.98 (m, 1 H), 2.94–2.89 (m, 1 H), 2.42–2.35 (m, 2 H), 2.34–2.21 (m, 3 H), 2.01–1.93 (m, 1 H), 1.92–1.85 (m, 2 H), 1.76–1.66 (m, 4 H); **^{13}C NMR** (125 MHz, $CDCl_3$, δ) 146.1, 139.7, 128.9, 128.5, 127.3, 126.2, 73.1, 66.0, 61.3, 54.4,

32.3, 32.2, 29.6, 24.4, 23.7; **HRMS** (ESI) m/z 258.1847 [calc'd for $C_{17}H_{23}NO$ $(M+H)^+$ 258.1853].

(*S*,*S*)-*N*-Benzyl-2-(cyclopent-1-enyl-tributylstannanylmethoxy-methyl)-pyrrolidine (5.23)

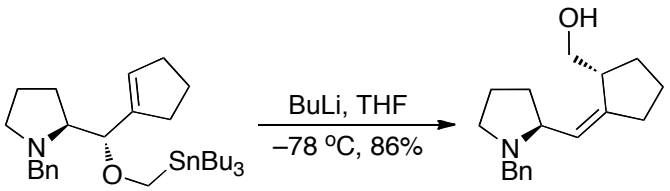


5.22 **5.23**

To a 100 mL flame-dried round bottom flask was added 230 mg KH (230 mg, 30% by wt in oil, 1.7 mmol) and a magnetic stir bar. The flask was purged with argon, and dry THF (2 mL) and 18-crown-6 (710 mg, 2.69 mmol) were added. In a separate dry flask, alcohol **5.22** (257 mg, mixture of diastereomers, *S*,*S*:*S*,*R* = 1:2, 1.00 mmol) was dissolved in dry THF (2 mL), and the resultant solution was transferred to the reaction mixture *via* syringe. A further 2 mL of THF was used for rinsing. In a separate dry flask, Bu_3SnCH_2I (647 mg, 1.7 mmol) was dissolved in THF (2 mL), and the resultant solution was transferred to the reaction mixture via syringe. A further 2 mL of THF was used for rinsing. After stirring at ambient temperature for 1 h, the reaction mixture was quenched by the dropwise addition of saturated NH_4Cl. The mixture was diluted with EtOAc and washed with saturated NH_4Cl and brine. The organic phase was dried over Na_2SO_4(s), filtered, and concentrated under reduced pressure. The desired diastereomer (*S*,*S*) was purified by chromatography on silica (0.25% NEt_3 + 2.5% EtOAc/hexanes to 5% EtOAc/hexanes) to yield the title compound (149 mg, 20% yield, 80% of theoretical yield) as a

yellow oil. **^{1}H NMR** (500 MHz, $CDCl_3$, δ) 7.35 (d, J = 7.1 Hz, 2 H), 7.29 (dd, J = 7.2, 7.2 Hz, 2 H), 7.21 (dd, J = 7.2, 7.2 Hz, 1 H), 5.64 (bs, 1 H), 4.49 (d, J = 13.2 Hz, 1 H), 3.75 (d, J = 7.7 Hz, 1 H), 3.69 (d, J = 9.8 Hz, 1 H), 3.47 (d, J = 10.1 Hz, 1 H), 3.38 (d, J = 13.2 Hz, 1 H), 2.90–2.86 (m, 1 H), 2.71 (apparent q, J = 7.8 Hz, 1 H), 2.39–2.30 (m, 3 H), 2.19–2.13 (m, 2 H), 1.91–1.83 (m, 2 H), 1.70–1.54 (m, 3 H), 1.53–1.39 (m, 6 H), 0.85 (t, J = 7.2 Hz, 9 H); **^{13}C NMR** (125 MHz, $CDCl_3$, δ) 143.3, 140.5, 129.7, 129.3, 128.1, 126.6, 89.4, 65.0, 60.7, 58.5, 55.0, 32.1, 31.2, 29.4, 28.1, 27.6, 23.7, 22.9, 13.9, 9.0; **HRMS** (ESI) m/z 558.3041 [calc'd for $C_{30}H_{51}NOSn$ $(M+H)^+$ 558.3061].

(Z)-(*S*,*S*)-[2-(*N*-Benzylpyrrolidin-2-ylmethylene)-cyclopentyl]-methanol (5.24)

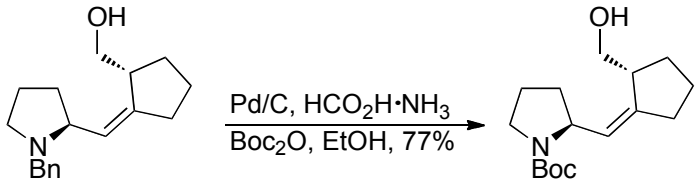


5.23 **5.24**

To a 100 mL flame-dried flask was added **5.23** (2.44 g, 4.36 mmol) and a magnetic stir bar. The flask was purged with argon and anhydrous THF (27 mL) was added. The reaction mixture was cooled to –78 °C in a dry ice/acetone bath. To the cooled reaction mixture was added dropwise a freshly titrated solution of *n*BuLi (1.85 mL, 2.35 M in hexanes, 4.36 mmol). The reaction mixture was stirred for 1 h and then warmed to ambient temperature. The reaction mixture was quenched by the addition of saturated NH_4Cl. The mixture was diluted with Et_2O and washed with saturated NH_4Cl and brine. The organic phase was dried over Na_2SO_4(s), filtered, and

concentrated under reduced pressure. The product was purified by chromatography on silica (1% NEt_3 in 1:1 EtOAc/hexanes) to yield the title compound (1.062 g, 86.3%) as a white solid. **^{1}H NMR** (500 MHz, $CDCl_3$, δ) 7.35–7.27 (m, 4 H), 7.25–7.20 (m, 1 H), 5.40 (dd, J = 8.7, 1.9 Hz, 1 H), 3.90 (d, J = 12.6 Hz, 1 H), 3.61 (dd, J = 10.6, 6.8 Hz, 1 H), 3.48 (dd, J = 10.6, 7.5 Hz, 1 H), 3.19 (d, J = 12.5 H, 1 H), 3.22–3.15 (m, 1 H), 2.92–2.82 (m, 2 H), 2.39–2.23 (m, 2 H), 2.19 (apparent q, J = 8.8 Hz, 1 H), 2.02–1.92 (m, 1 H), 1.87–1.52 (m, 7 H); **^{13}C NMR** (125 MHz, $CDCl_3$, δ) 146.7, 139.8, 129.0, 128.4, 127.0, 124.9, 65.2, 64.6, 57.9, 52.7, 43.5, 34.2, 31.1, 29.9, 23.8, 22.0; **HRMS** (ESI) m/z 272.2010 [calc'd for $C_{18}H_{25}NO$ $(M+H)^+$ 272.2009].

(Z)-(*S*,*S*)-[2-(*N*-*tert*-butyloxycarbonylpyrrolidin-2-ylmethylene)-cyclopentyl]-methanol (5.25)

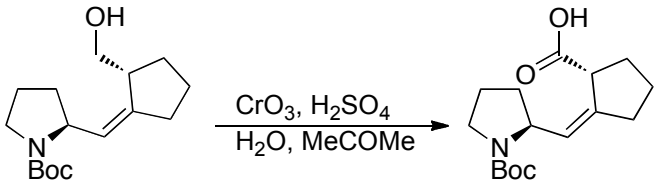


5.24 **5.25**

To a flame dried flask was added alcohol **5.24** (700 mg, 2.57 mmol), Boc_2O (1.14 g, 5.24 mmol), ammonium formate (229 mg, 3.63 mmol), and a magnetic stir bar. The flask was purged with argon and ethanol (10 mL) was added. A portion of Pd/C (500 mg, 10% by mass) was added, and the reaction mixture was stirred at ambient temperature and monitored by TLC. After approximately 30 min, the reaction mixture was filtered through a plug of Celite,® rinsed with methanol, and concentrated under reduced pressure. The product was purified by

chromatography on silica (1:3 EtOAc/hexanes) to yield the title compound (559 mg, 77%) as a colorless oil. **^{1}H NMR** (500 MHz, $CDCl_3$, δ) 5.34 (bs, 1 H), 5.27 (d, J = 9.5 Hz, 1 H), 4.66 (ddd, J = 7.4, 7.4, 7.4 Hz, 1 H), 3.63–3.52 (m, 2 H), 3.44–3.32 (m, 2 H), 2.93 (bs, 1 H), 2.45 (apparent quintet, J = 8.0 Hz, 1 H), 2.20 (apparent quintet, J = 7.1 Hz, 1 H), 2.08 (ddd, J = 13.2, 7.0, 7.0 Hz, 1 H), 1.90–1.74 (m, 4 H), 1.68–1.63 (m, 1 H), 1.60–1.51 (m, 2 H), 1.46 (s, 9 H); **^{13}C NMR** (125 MHz, $CDCl_3$, δ) 155.2, 144.8, 124.6, 79.7, 66.7, 57.2, 47.3, 43.2, 33.7, 33.4, 30.0, 28.8, 24.2, 23.9; **HRMS** (ESI) m/z 282.2061[calc'd for $C_{16}H_{27}NO_3$ $(M+H)^+$ 282.2064].

(Z)-(*S*,*S*)-[2-(*N*-Benzylpyrrolidin-2-ylmethylene)-cyclopentyl]carboxylic Acid (5.26)

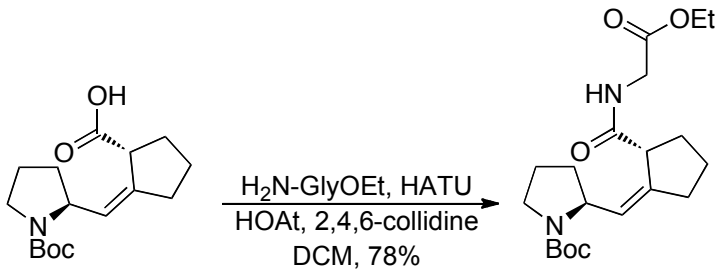


5.25 **5.26**

The Jones Reagent was prepared by addition of 1.0 mL concentrated H_2SO_4 to 1.0 g of CrO_3. The solution was cooled to 0 °C and 3 mL of cold water was added slowly with rapid stirring. In a 50 mL round bottom flask, alcohol **5.25** (40 mg, 0.175) was dissolved in acetone (2 mL) under an atmosphere of argon. The reaction flask was cooled to 0 °C and Jones Reagent (140 μL, 0.35 mmol) was added dropwise. After 10 min, the reaction mixture was quenched by the dropwise addition of *iso*-propanol until the reaction mixture turned from orange to green. Brine (10 mL) was added, and the mixture was extracted with EtOAc (3 x 10 mL). The combined organic phase was washed with brine, dried over Na_2SO_4(s), filtered, and concentrated under reduced pressure

to afford the title compound (15 mg) which was taken on crude to the next step. **HRMS** (ESI) m/z 294.1711 [calc'd for $C_{16}H_{24}NO_4$ $(M–H)^-$ 294.1710].

(Z)-(*S*,*S*)-*N*-(2-(*N*-Benzylpyrrolidin-2-ylmethylene)-cyclopentyl)carbonyl]glycine Ethyl Ester (5.27)

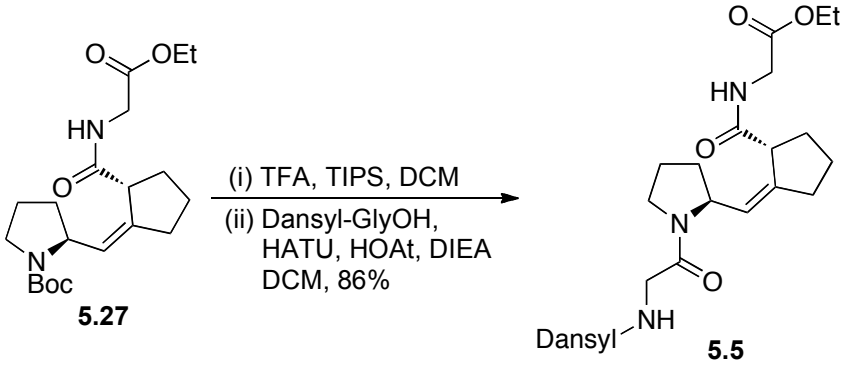


5.26 **5.27**

Acid **5.26** (15 mg, 0.051 mmoles) was dissolved in anhydrous DCM (1 mL) after which glycine ethyl ester hydrochloride (8 mg, 0.056 mmoles), 1-hydroxy-7-azobenzotriazole (21 mg, 0.153 mmoles), 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate (58 mg, 0.153 mmoles) and were added and the suspension stirred at ambient temperature under an atmosphere of N_2(g). 2,4,6-Collidine (0.034 mL, 0.255 mmoles) was added and the reaction was stirred at ambient temperature for 30 min. Brine was added to the reaction mixture, after which it was extracted with EtOAc. The combined organics were dried over Na_2SO_4(s) and concentrated under reduced pressure to afford a crude oil. The crude product was purified by chromatography on silica (30–100% EtOAc in hexanes) to afford the title compound (15 mg, 78%) as a pale yellow oil, which was taken on slightly crude to the next step. **^{1}H NMR** (400 MHz, $CDCl_3$, δ) 8.71 (bs, 1 H), 5.42 (d, J = 9.2 Hz, 1 H), 4.36 (d, J = 7.2,

16.4 Hz, 1 H), 4.20–4.07 (m, 2 H), 3.99 (dd, J = 6.4, 17.2 Hz, 1 H), 3.88 (dd, J = 5.6, 17.2 Hz, 1 H), 3.44–3.29 (m, 3 H), 2.53–2.44 (m, 1 H), 2.39–2.32 (m, 1 H), 2.22 (apparent quintet, J = 7.6 Hz, 1 H), 2.07 (apparent sextet, J = 6.0 Hz, 1 H), 1.89–1.73 (m, 3 H), 1.64–1.54 (m, 5 H), 1.4 (s, 9 H), 1.23 (t, J = 6.8 Hz, 3 H); **^{13}C NMR** (125 MHz, $CDCl_3$, δ) 173.0, 170.2, 154.9, 141.1, 127.3, 79.4, 60.7, 57.2, 47.4, 47.1, 41.6, 33.2, 33.1, 31.2, 28.5, 24.2, 23.8, 14.2.

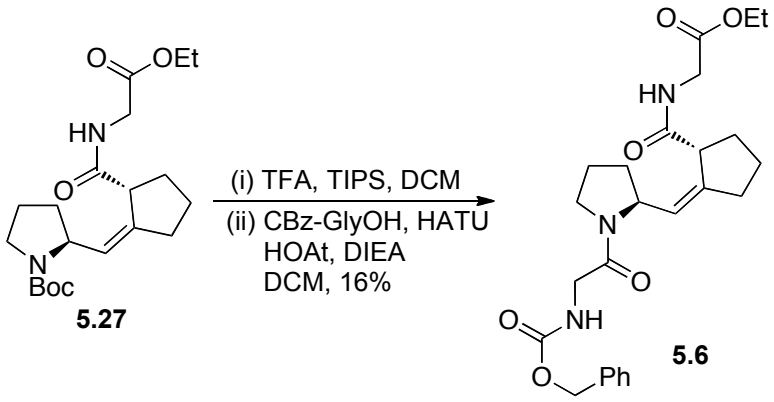
DansylGlyPro*cis*ProGlyOEt (5.5)



Alkene **5.27** (38 mg, 0.1 mmoles) was dissolved in dry DCM (0.56 mL) in a vial, after which triisopropylsilane (0.04 mL, 0.2 mmoles) was added. TFA (0.56 mL) was added and the reaction was stirred at ambient temperature for 1 h. The reaction was concentrated under reduced pressure, and the residue resuspended in dry DCM (2 mL). Dansyl-Gly-OH (30 mg, 0.1 mmoles), 1-hydroxy-7-azobenzotriazole (41 mg, 0.3 mmoles), and 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate (114 mg, 0.3 mmoles) were added and the reaction mixture was stirred at ambient temperature. DIEA (0.08 mL) was added and the reaction was stirred at room temperature for 1 h. The

reaction was concentrated under reduced pressure, after which the crude product was purified by chromatography on silica (70% EtOAc in hexanes) and HPLC 25–65% v/v acetonitrile in H_2O with 0.1% v/v TFA) to afford the title compound (49 mg, 86%) as a pale yellow solid. **^{1}H NMR** (500 MHz, $CDCl_3$, mixture of 2 or more rotamers, integrations are approximate, δ) 8.62 (t, J = 9.0 Hz, 1.0 H), 8.45 (t, J = 6.0 Hz, 0.5 H), 8.32 (d, J = 7.5 Hz, 0.5 H), 7.71 (q, J = 8.0 Hz, 1.0 H), 7.53 (d, J = 7.5 Hz, 0.5 H), 5.84 (t, J = 4.0 Hz, 0.5 H), 5.26 (d, J = 9.0 Hz, 0.5 H), 4.36 (q, J = 8.0 Hz, 0.6 H), 4.20–4.10 (m, 1.2 H), 4.05–3.98 (m, 0.2 H), 3.87–3.50 (m, 7.8 H), 3.41–3.04 (m, 5.1 H), 2.44–2.34 (m, 0.6 H), 2.27–2.18 (m, 1.1 H), 2.12–2.03 (m, 0.7 H), 2.00–1.81 (m, 1.7 H), 1.67–1.52 (m, 1.8 H), 1.43 (s, 0.2 H), 1.28–1.25 (m, 4.2 H), 0.89–0.82 (m, 0.5 H), 0.07 (m, 0.1 H); **^{13}C NMR** (125 MHz, $CDCl_3$, mixture of 2 or more rotomers, δ) 177.0, 172.3, 168.2, 163.0 (q, J = 39.4 Hz, TFA) 147.7, 144.4, 137.7, 132.9, 132.4, 131.2, 130.7, 130.6, 128.7, 128.0, 126.4, 119.8, 118.1 (q, J = 286 Hz, TFA) 64.0, 61.2, 49.8, 49.0, 48.9, 47.4, 44.4, 36.1 (2 signals), 34.8, 34.7, 34.4, 33.0, 32.4 (2 signals), 32.3, 32.2, 32.1, 32.0, 31.8, 27.4, 27.1, 26.6, 25.4, 16.9, 16.8; **HRMS** (ESI) m/z 571.2579 [calc'd for $C_{29}H_{39}N_4O_6S$ $(M+H)^+$ 571.2585].

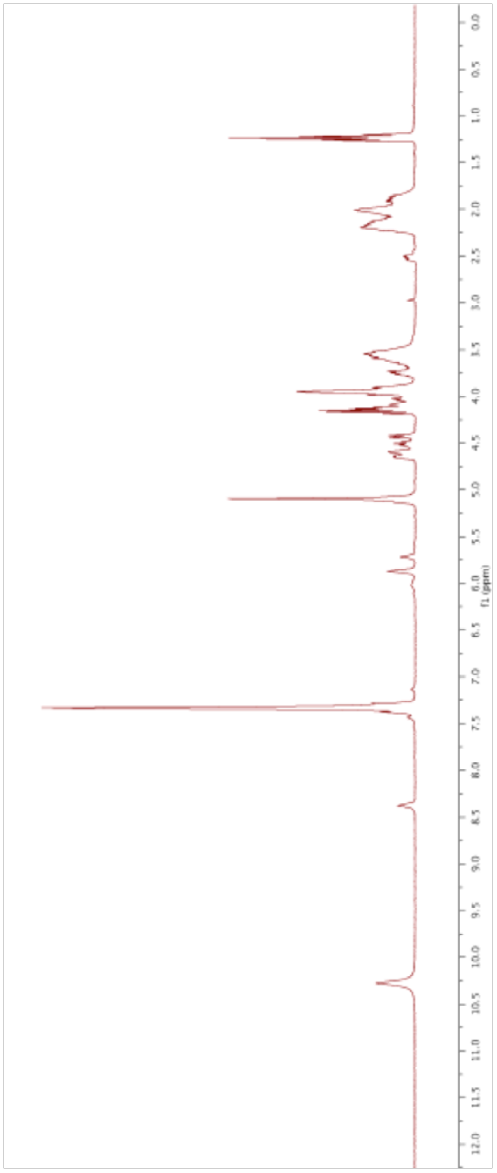
CBzGlyPro*cis*ProGlyOEt (5.6)



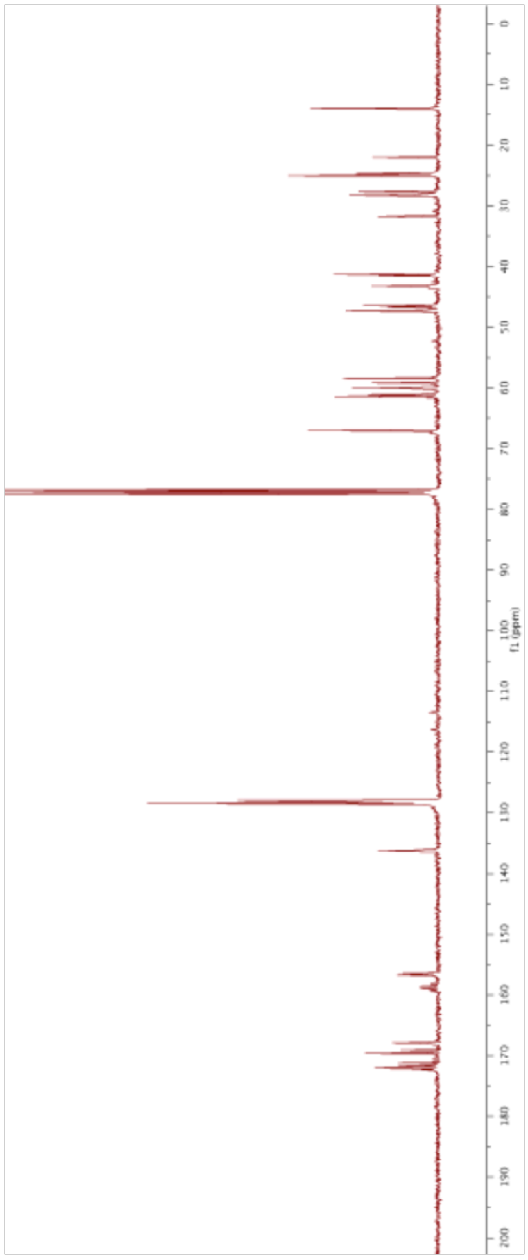
Alkene **5.27** (15 mg, 0.039 mmoles) was dissolved in dry DCM (0.22 mL) in a vial, after which triisopropylsilane (0.016 mL, 0.078 mmoles) was added. TFA (0.22 mL) was added and the reaction was stirred at ambient temperature for 1 h. The reaction was concentrated under reduced pressure, and the residue resuspended in dry DCM (0.8 mL). Cbz-Gly-OH (8.2 mg, 0.039 mmoles), 1-hydroxy-7-azobenzotriazole (16 mg, 0.12 mmoles), and 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate (45 mg, 0.17 mmoles) were added and the reaction mixture was stirred at ambient temperature. DIEA (0.034 mL) was added and the reaction was stirred at room temperature for 1 h. The reaction was concentrated under reduced pressure, after which the crude product was purified by HPLC (30–90% v/v acetonitrile in H_2O with 0.1% v/v TFA) to afford the title compound (3 mg, 16%) as a yellow oil. **^{1}H NMR** (500 MHz, $CDCl_3$, mixture of 2 or more rotamers, integrations are approximate, δ) 8.61 (t, J = 6.0 Hz, 1.0 H), 7.39–7.28 (m, 5.5 H plus $CHCl_3$), 5.66 (s, 1.0 H), 5.40 (d, J = 9.0 Hz, 1.0 H), 5.14–5.09 (m, 2.1 H), 4.57 (app q, J = 8.0 Hz, 1.1 H), 4.20–4.09 (m, 2.1 H), 4.02–3.90 (m, 2.9 H), 3.49–3.43 (m, 2.2 H), 2.51–2.45 (m, 1.9 H), 2.36–2.33 (m, 1.4 H), 2.28–2.14 (m, 2.8 H), 2.08–1.88 (m, 6.9 H), 1.69–1.61 (m, 2.2 H), 1.28–1.22 (m, 3.5 H); 0.91–0.74 (m, 2.6 H); **^{13}C NMR** (125 MHz, $CDCl_3$, δ) 173.4, 170.0, 166.9, 156.1, 141.9, 136.3, 128.6, 128.3, 128.2, 126.1, 67.1, 61.0, 58.3, 47.4, 46.4, 43.7, 41.7, 33.2, 32.3, 31.5, 24.6, 23.9, 14.1; **HRMS** (ESI) m/z 472.2442 [calc'd for $C_{25}H_{34}N_3O_6$ $(M+H)^+$ 472.2442].

5.5.8 NMR Spectra

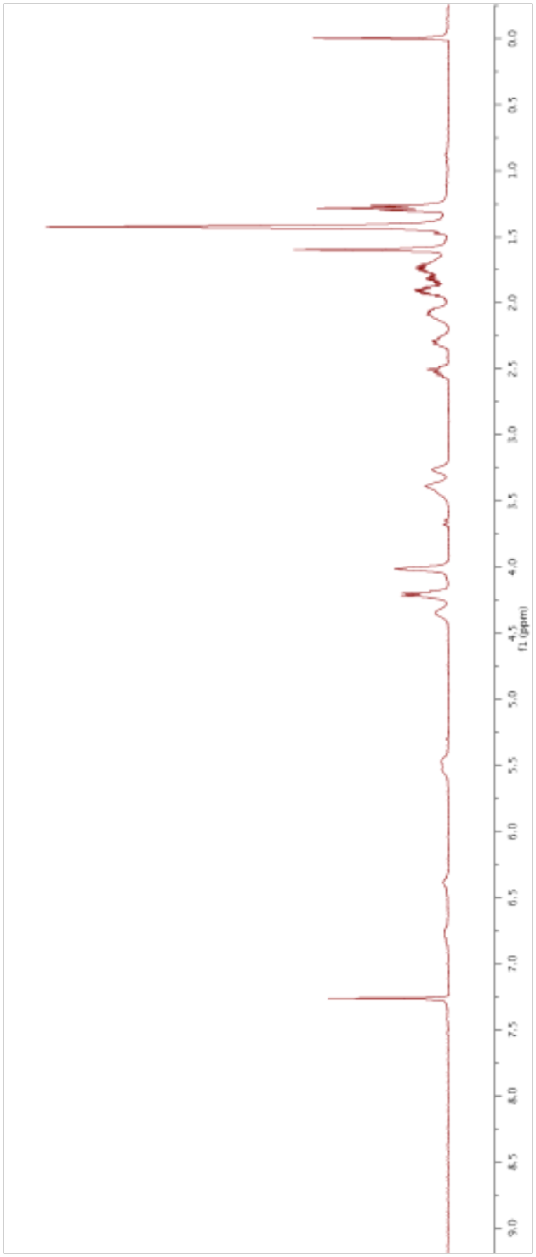
400 MHz ^{1}H NMR of CBzGPPGOEt (5.2) in $CDCl_3$



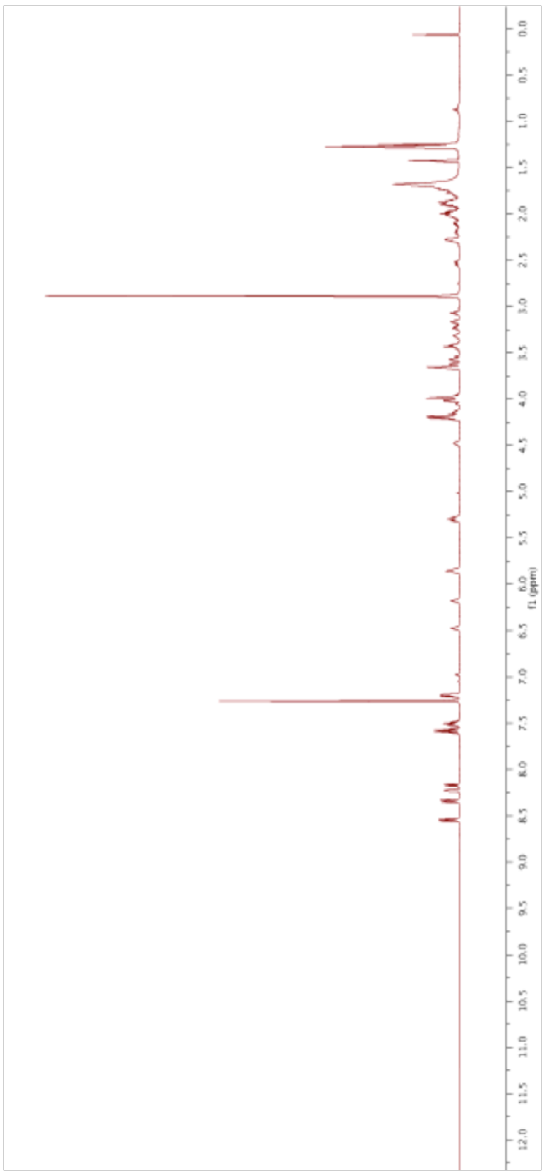
100 MHz ^{13}C NMR of CBzGPPGOEt (5.2) in $CDCl_3$



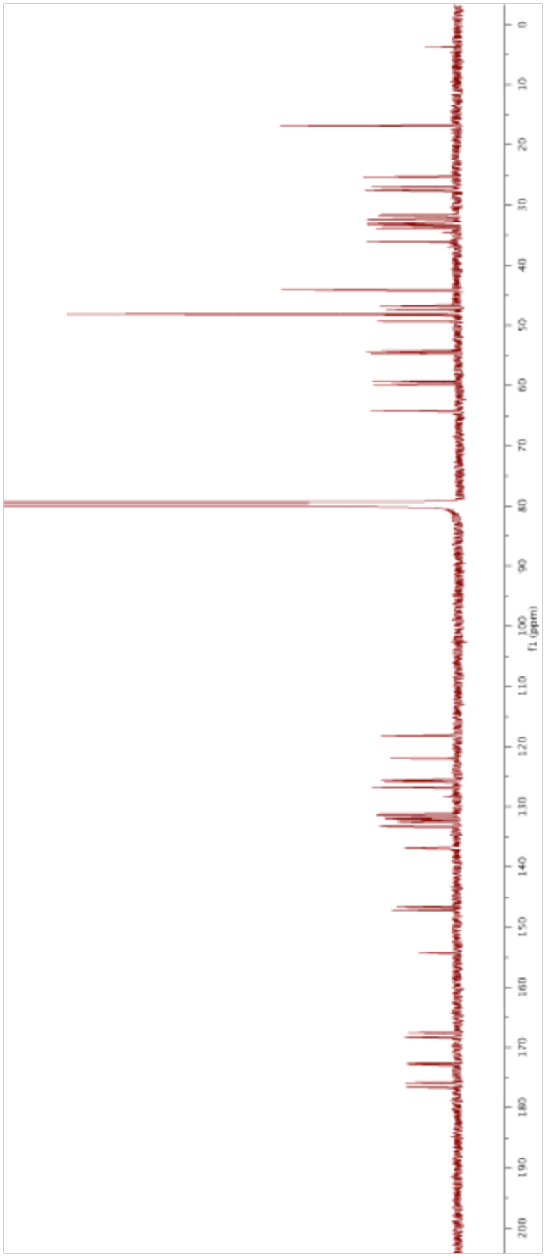
400 MHz ^{1}H NMR of (E)-(*S*,*S*)-*N*-(2-(*N*-Benzylpyrrolidin-2-ylmethylene)-cyclopentyl)carbonyl]glycine Ethyl Ester (5.18) in $CDCl_3$



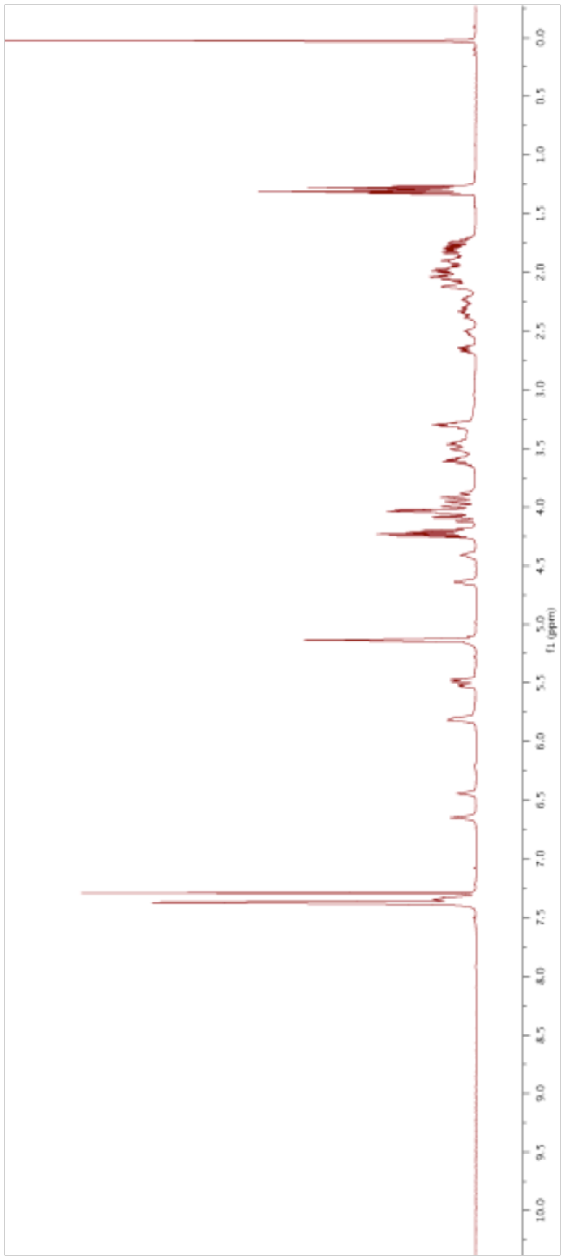
500 MHz ^{1}H NMR Spectrum of DansylGlyPro*trans*ProGlyOEt (5.3) in $CDCl_3$



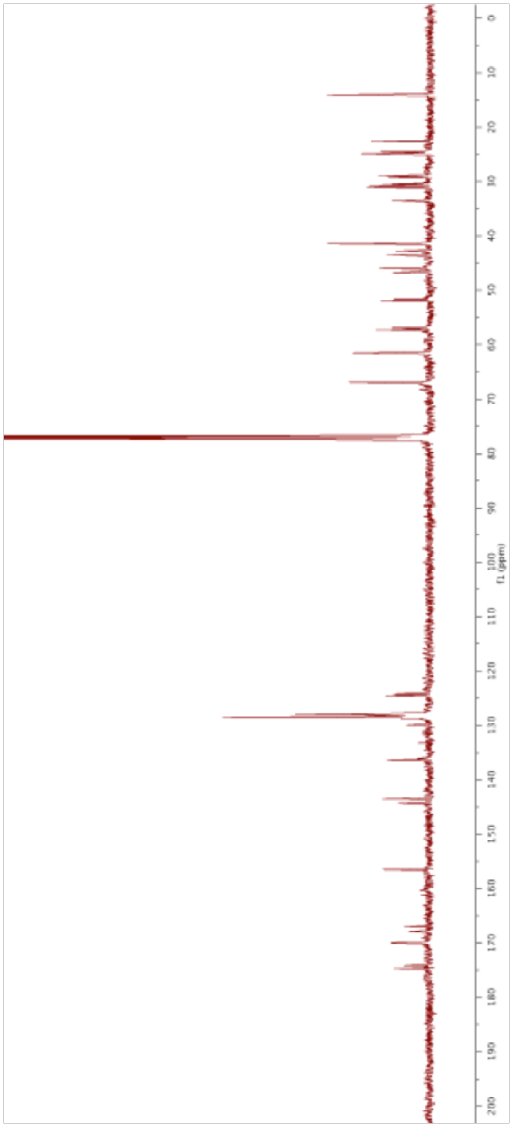
125 MHz ^{13}C NMR Spectrum of DansylGlyPro*trans*ProGlyOEt (5.3) in $CDCl_3$



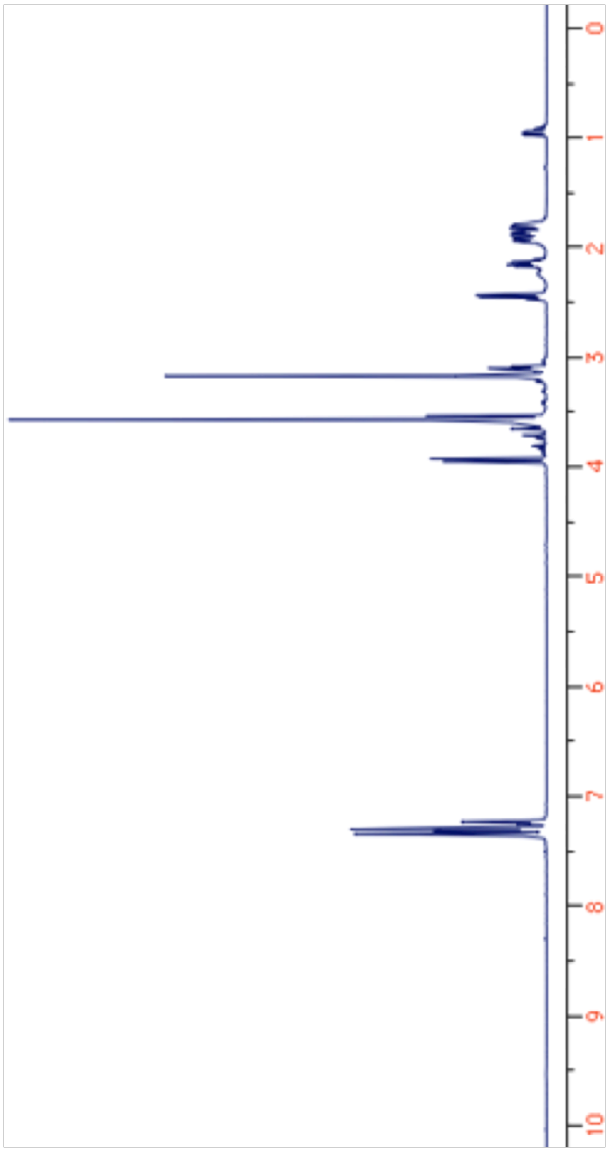
500 MHz ^{1}H NMR Spectrum of CBzGlyPro*trans*ProGlyOEt (5.4) in $CDCl_3$



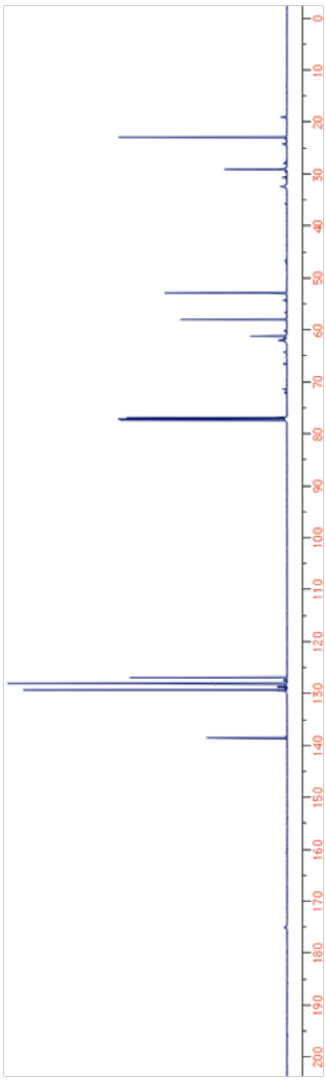
125 MHz ^{13}C Spectrum of CBzGlyPro*trans*ProGlyOEt (5.4) in $CDCl_3$



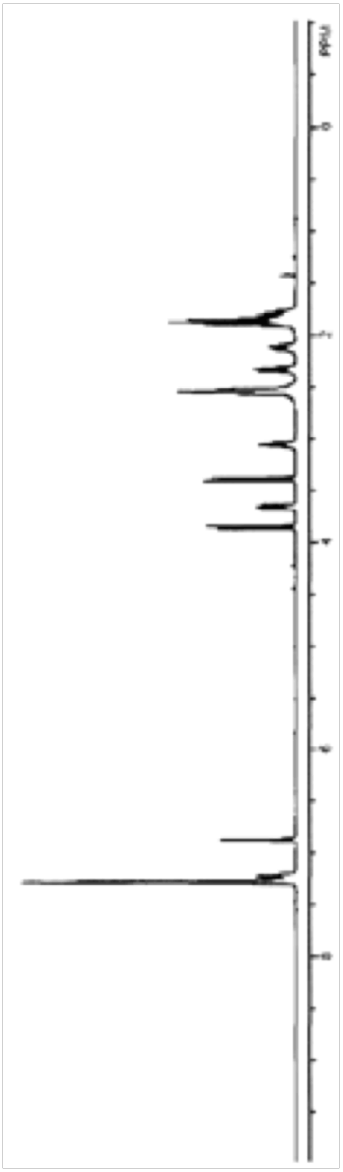
500 MHz ^{1}H NMR Spectrum of *N*-Methoxy-*N*-methyl-1-(phenylmethyl)-(2*S*)-2-pyrrolidinecarboxamide (5.20) in $CDCl_3$



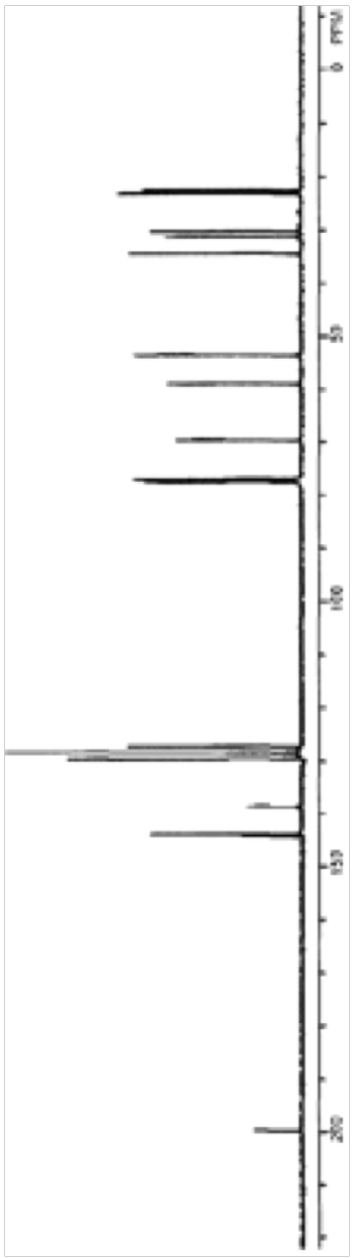
125 MHz ^{13}C NMR Spectrum of *N*-Methoxy-*N*-methyl-1-(phenylmethyl)-(2*S*)-2-pyrrolidinecarboxamide (5.20) in $CDCl_3$



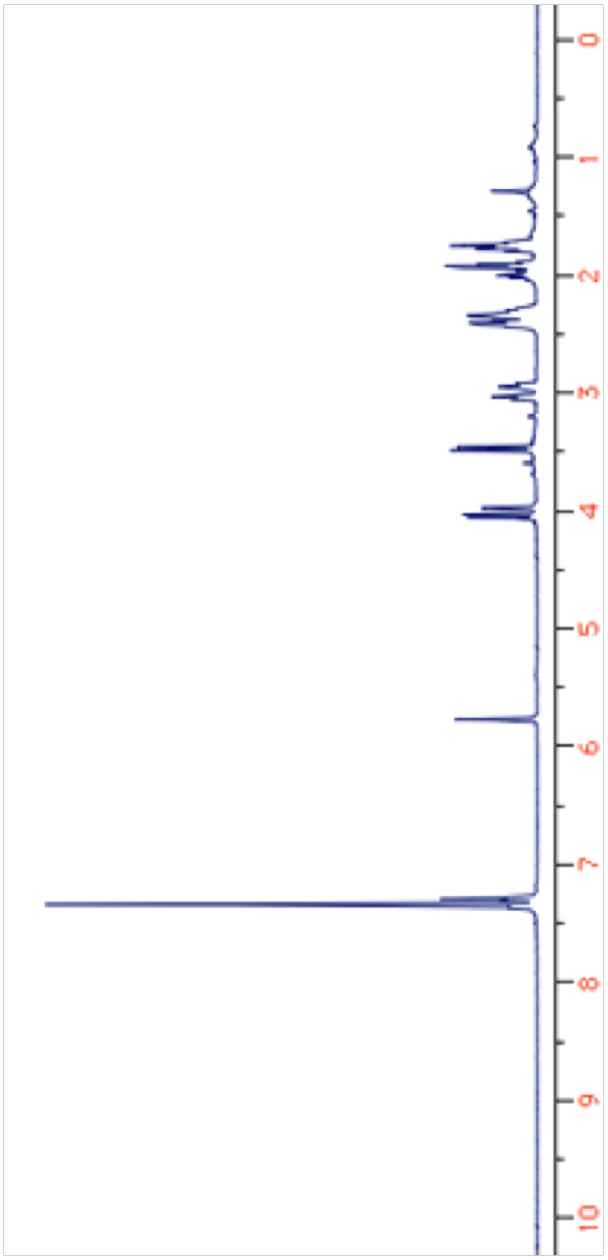
400 MHz ^{1}H NMR Spectrum of (*S*)-(*N*-Benzylpyrrolidin-2-yl)-cyclopent-1-enyl-methanone (5.21) in $CDCl_3$



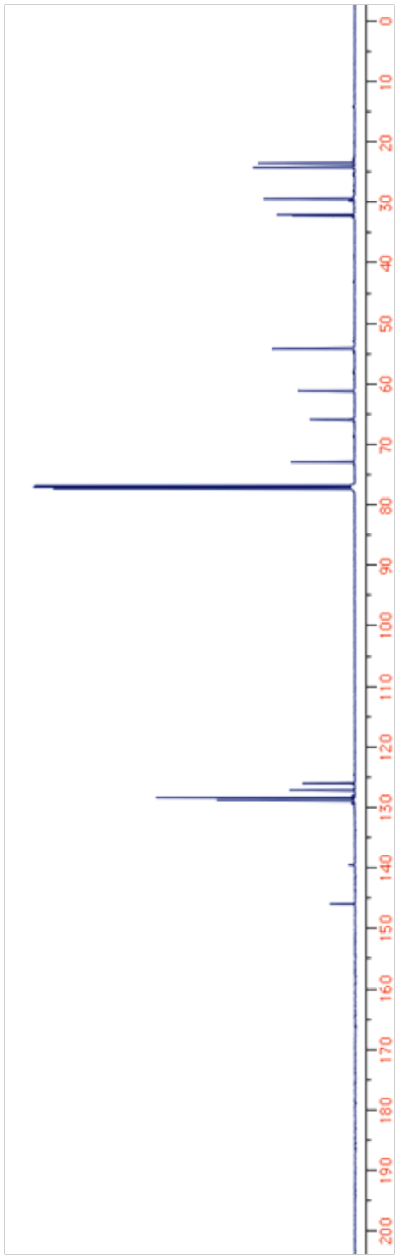
100 MHz ^{13}C NMR Spectrum of (*S*)-(*N*-Benzylpyrrolidin-2-yl)-cyclopent-1-enyl-methanone (5.21) in $CDCl_3$



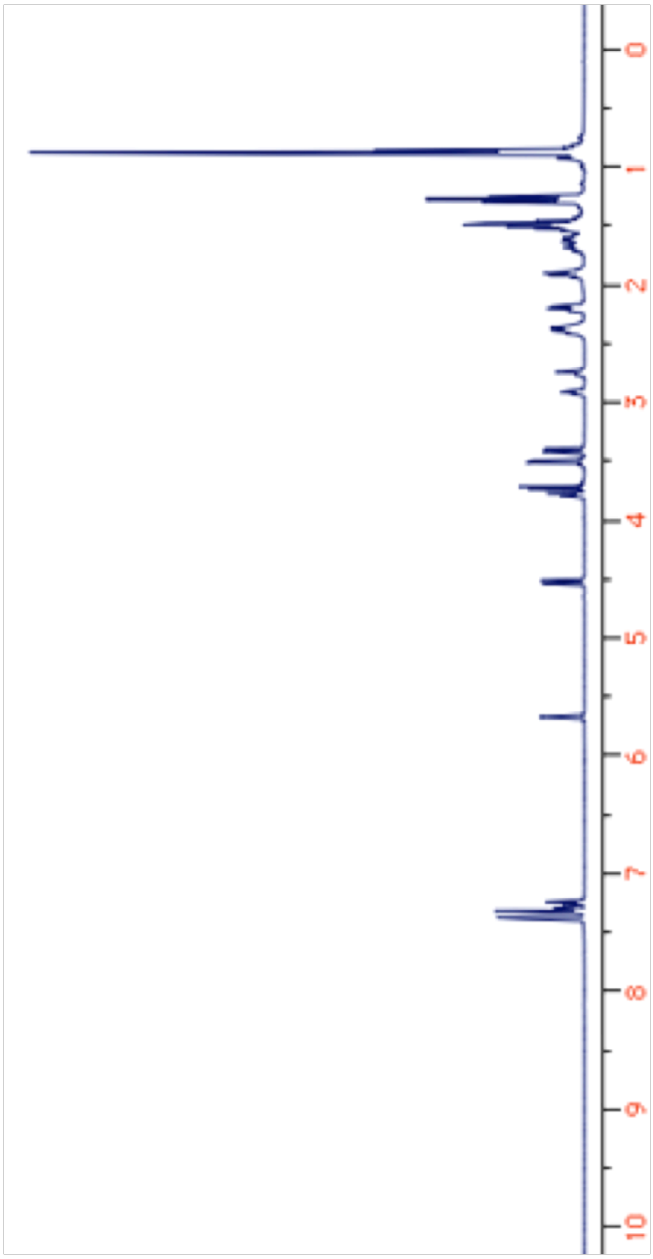
500 MHz ^{1}H NMR Spectrum of (*S*,*S*)-(*N*-Benzylpyrrolidin-2-yl)-cyclopent-1-enyl-methanol (5.22) in $CDCl_3$



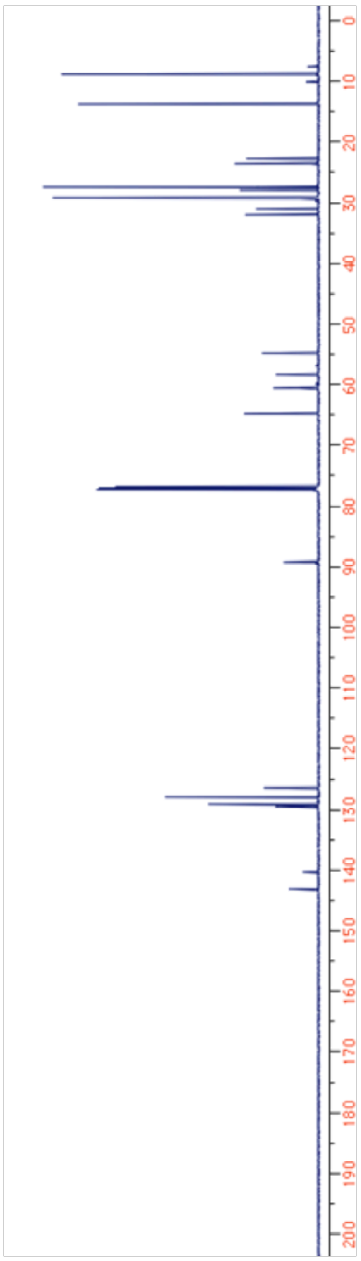
125 MHz ^{13}C NMR Spectrum of (*S*,*S*)-(*N*-Benzylpyrrolidin-2-yl)-cyclopent-1-enyl-methanol (5.22) in $CDCl_3$



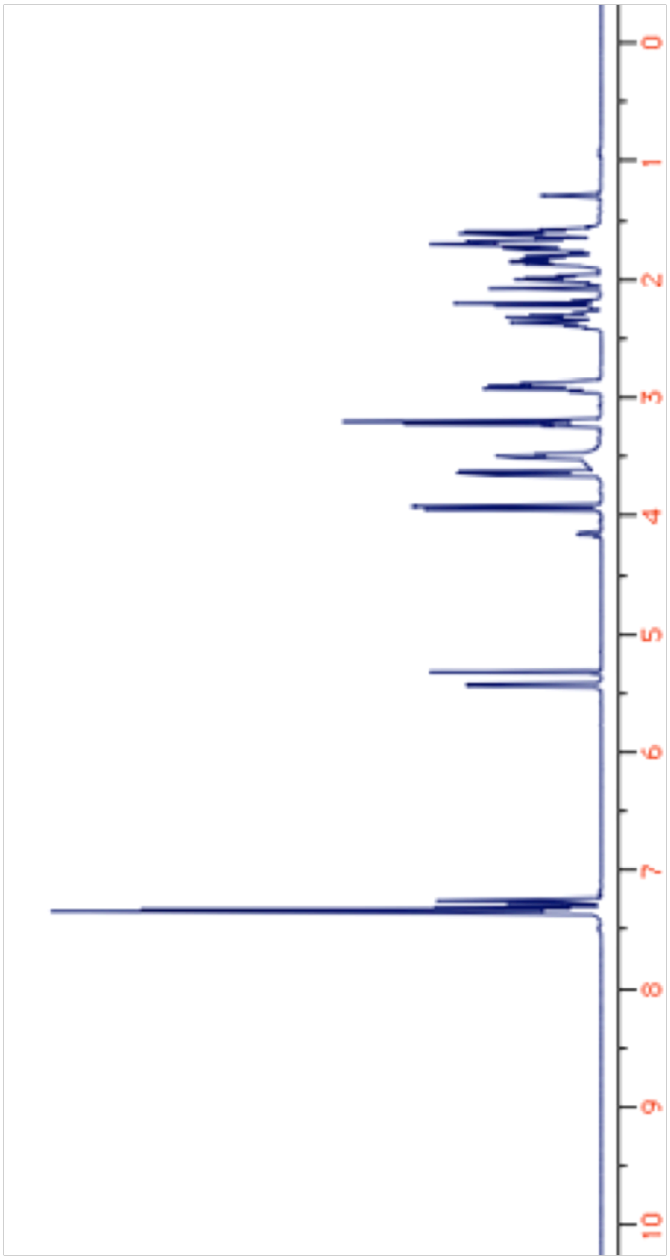
500 MHz ^{1}H NMR Spectrum of (*S*,*S*)-*N*-Benzyl-2-(cyclopent-1-enyl-tributylstannanylmethoxy-methyl)-pyrrolidine (5.23) in $CDCl_3$



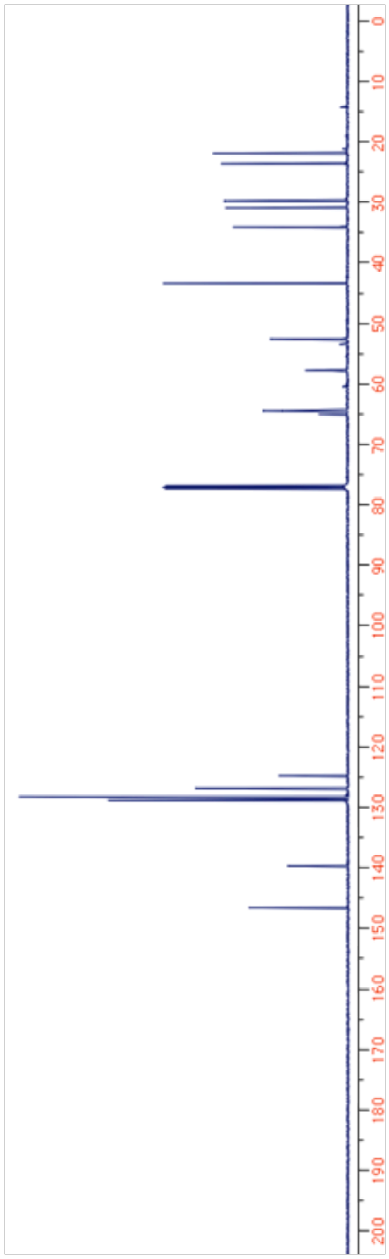
125 MHz ^{13}C NMR Spectrum of (*S*,*S*)-*N*-Benzyl-2-(cyclopent-1-enyl-tributylstannanylmethoxy-methyl)-pyrrolidine (5.23) in $CDCl_3$



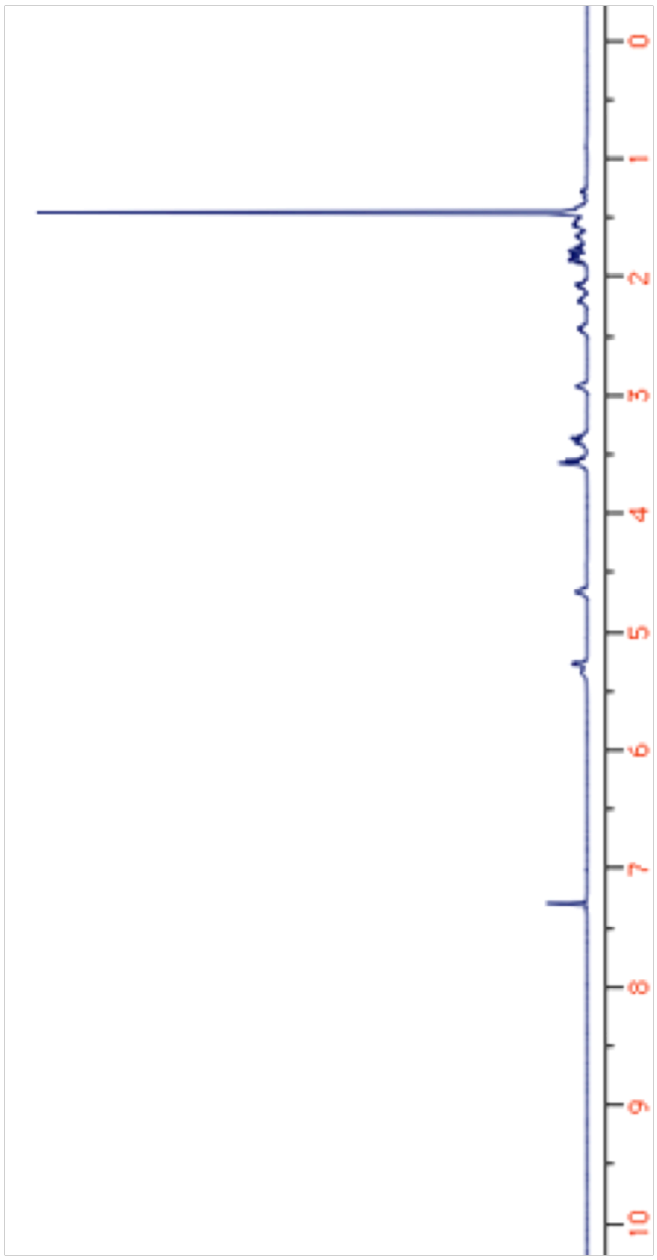
500 MHz ^{1}H NMR Spectrum of (Z)-(*S*,*S*)-[2-(*N*-Benzylpyrrolidin-2-ylmethylene)-cyclopentyl]-methanol (5.24) in $CDCl_3$



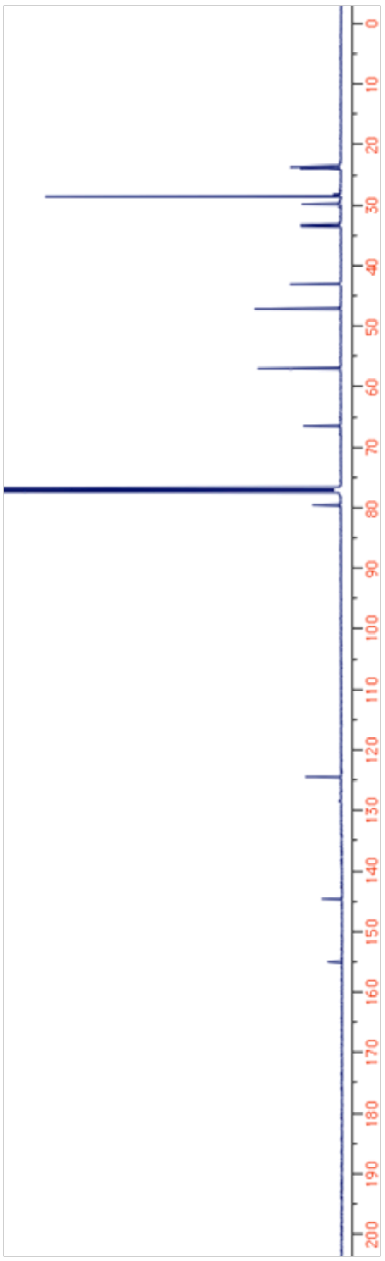
500 MHz ^{13}C NMR Spectrum of (Z)-(*S*,*S*)-[2-(*N*-Benzylpyrrolidin-2-ylmethylene)-cyclopentyl]-methanol (5.24) in $CDCl_3$



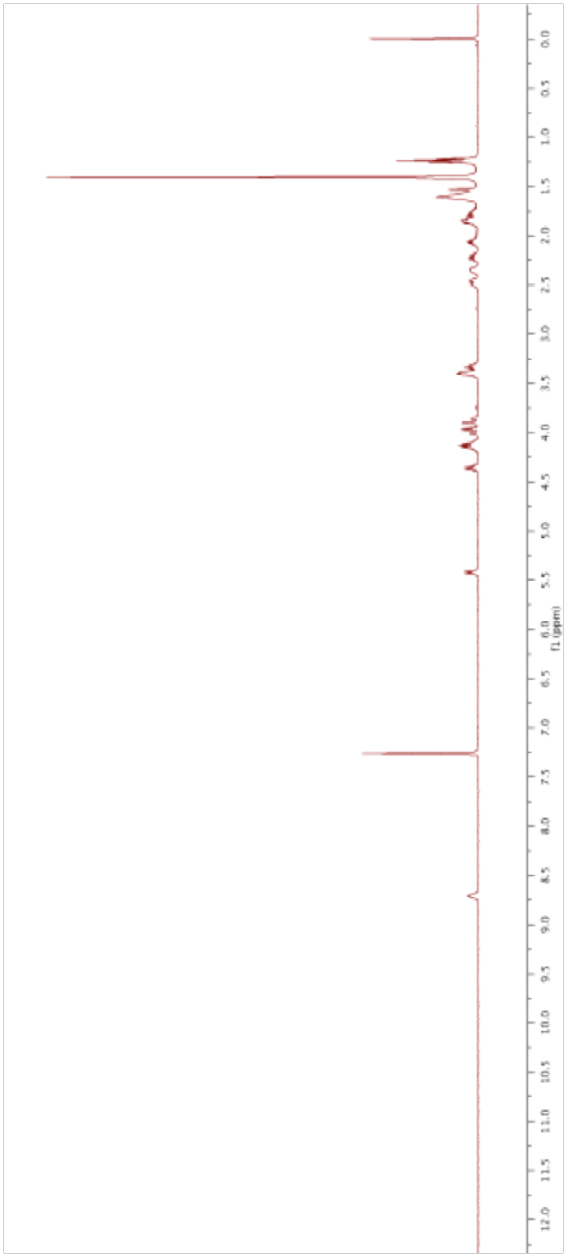
500 MHz ^{1}H NMR Spectrum of (Z)-(*S*,*S*)-[2-(*N*-*tert*-butyloxycarbonylpyrrolidin-2-ylmethylene)-cyclopentyl]-methanol (5.25) in $CDCl_3$



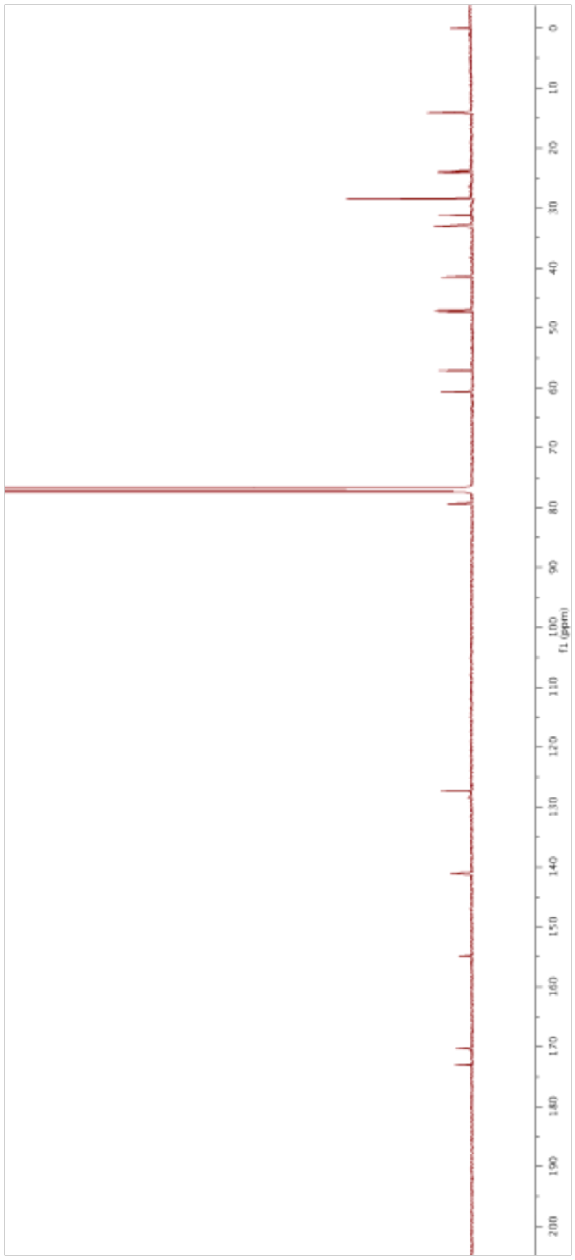
125 MHz ^{13}C NMR Spectrum of (Z)-(*S*,*S*)-[2-(*N*-*tert*-butyloxycarbonylpyrrolidin-2-ylmethylene)-cyclopentyl]-methanol (5.25) in $CDCl_3$



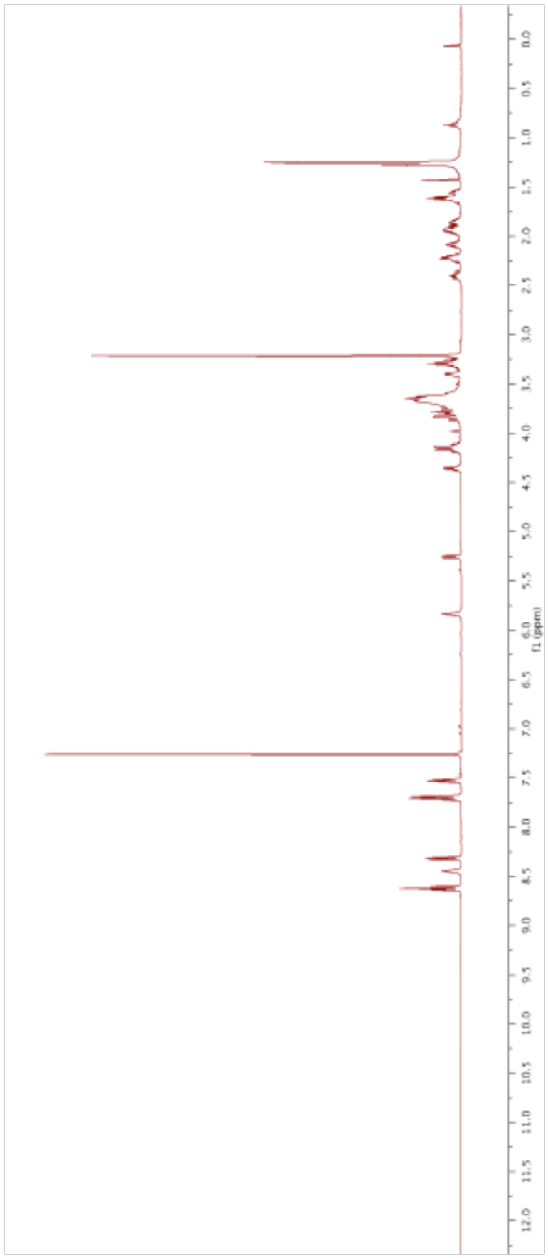
400 MHz ^{1}H NMR Spectrum of (Z)-(*S*,*S*)-*N*-(2-(*N*-Benzylpyrrolidin-2-ylmethylene)-cyclopentyl)carbonyl]glycine Ethyl Ester (5.27) in $CDCl_3$



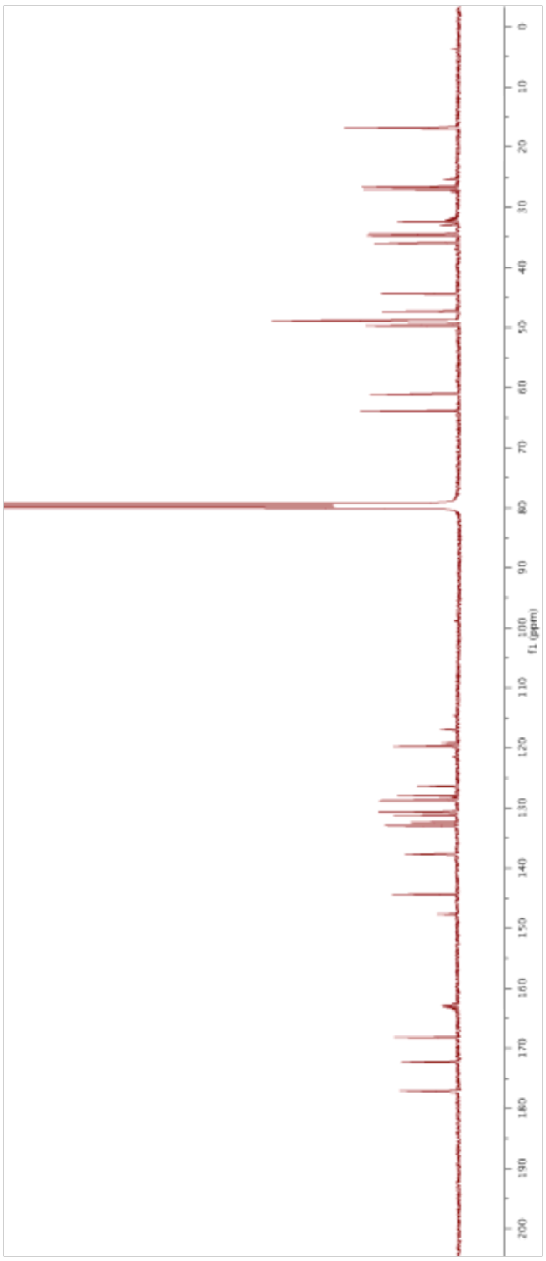
125 MHz ^{13}C NMR Spectrum of (Z)-(*S*,*S*)-*N*-(2-(*N*-Benzylpyrrolidin-2-ylmethylene)-cyclopentyl)carbonyl]glycine Ethyl Ester (5.27) in $CDCl_3$



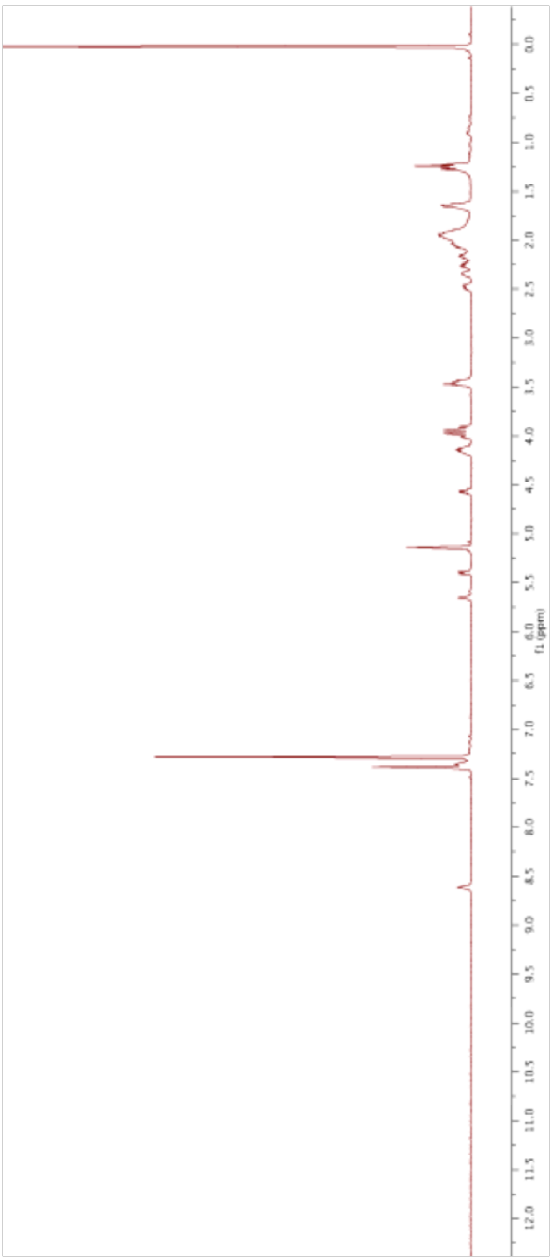
500 MHz ^{1}H NMR Spectrum of DansylGlyPro*cis*ProGlyOEt (5.5) in $CDCl_3$



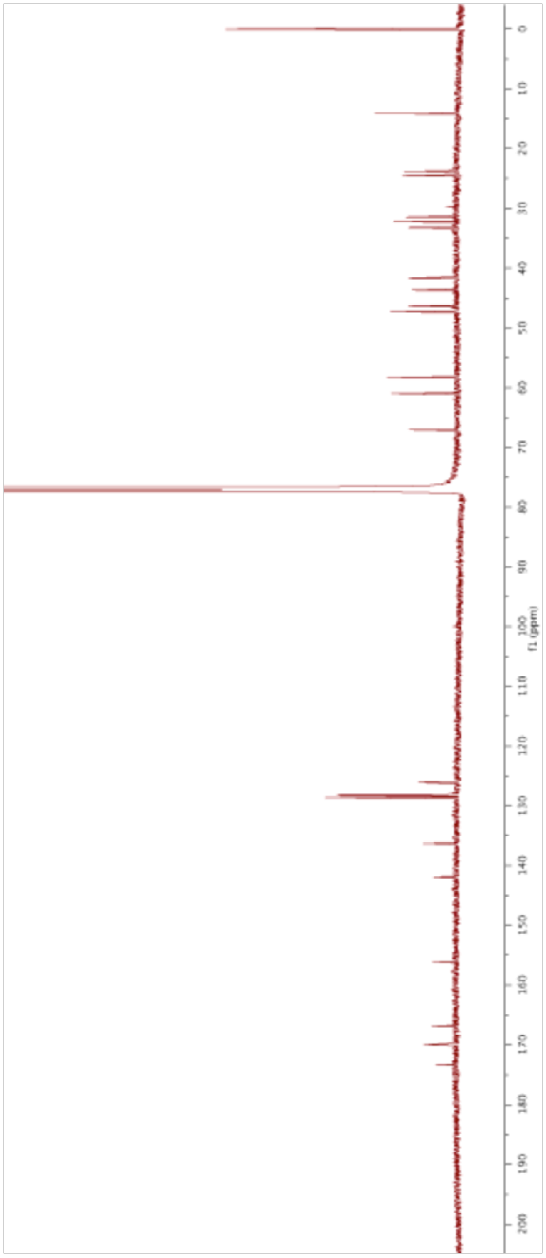
125 MHz ^{13}C NMR Spectrum of DansylGlyPro*cis*ProGlyOEt (5.5) in $CDCl_3$



500 MHz ^{1}H NMR Spectrum of CBzGlyPro*cis*ProGlyOEt (5.6) in $CDCl_3$



125 MHz ^{13}C NMR Spectrum of CBzGlyPro*cis*ProGlyOEt (5.6) in $CDCl_3$

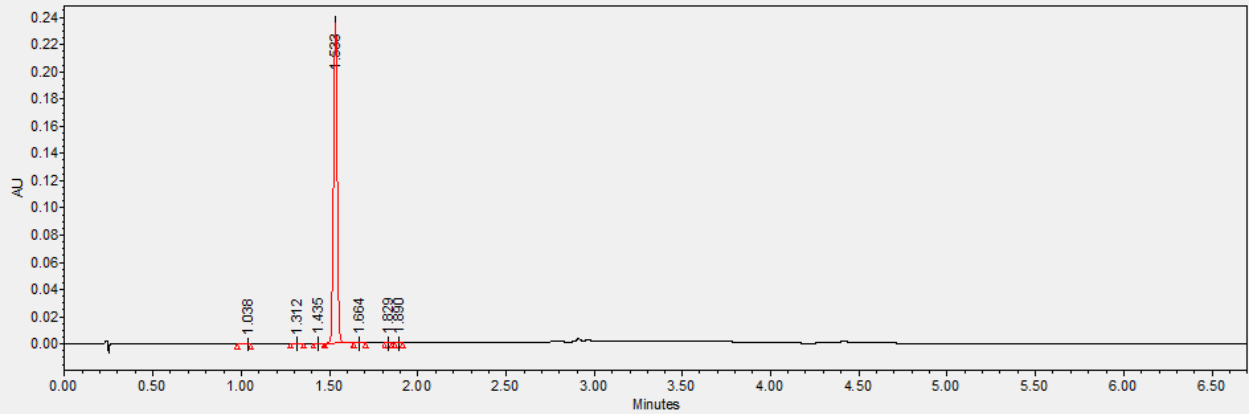


5.5.9 Determination of Compound Purity

All new final compounds were analyzed by UPLC to determine their purity. The analyses were performed on the Waters UPLC system using an Acquity UPLC® BEH C18 column (2.1 × 50 mm, 1.7 μm particle size) from Waters. Samples (5 μL) dissolved in H_2O were injected into the column and eluted at 0.6 mL/min with a gradient of aqueous acetonitrile (20–68% v/v over 2.9 min for dansyl peptides and 10–90% v/v over 2.9 min for CBz peptides) containing 0.1% v/v TFA. For dansyl peptides, the absorbance at 289 nm was used as the detection wavelength. For CBz peptides, the absorbance at 218 nm was used as the detection wavelength.

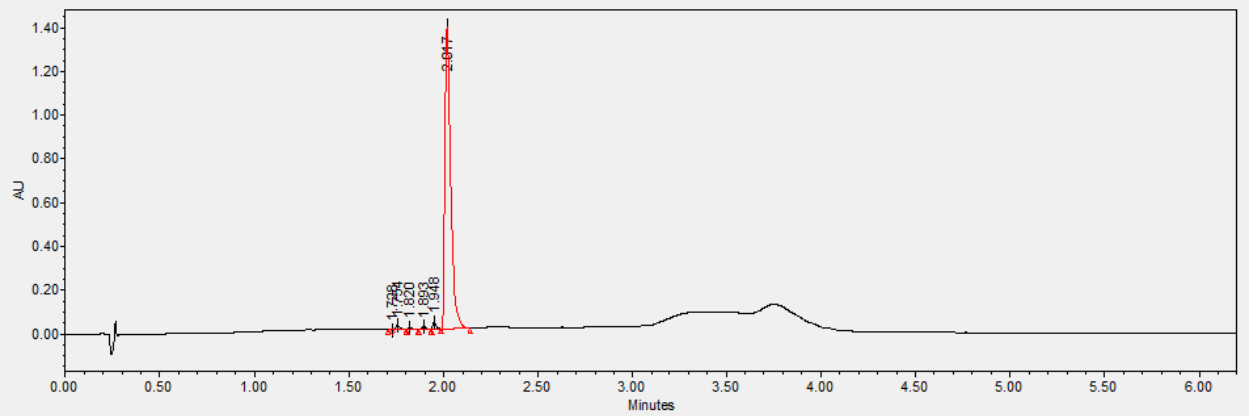
5.5.10 Chromatograms of Final Compounds

DansylGlyProProGlyOEt (5.1)



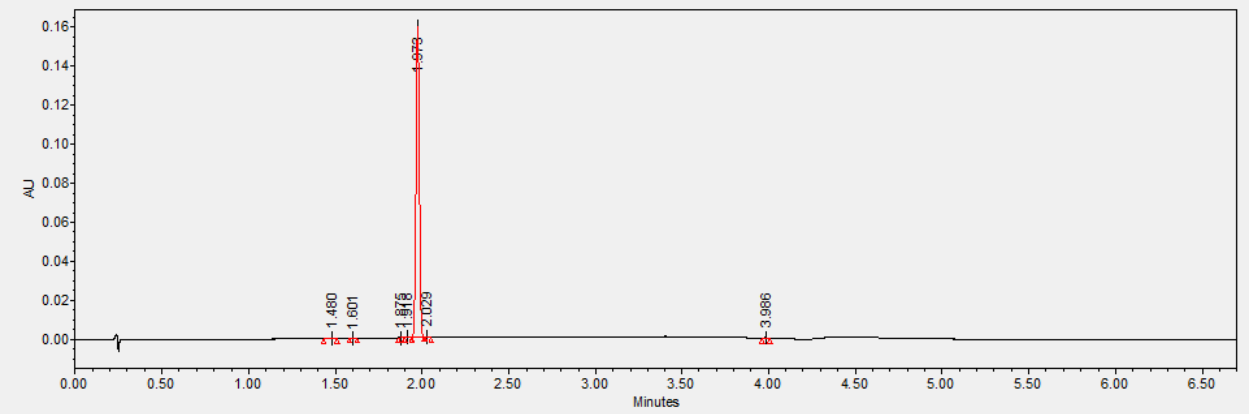
Peak	Retention Time (min)	Area (µV*sec)	Area (%)	Height (µV)
1	1.038	87	0.03	44
2	1.312	41	0.01	28
3	1.435	47	0.01	35
4	**1.533**	**335060**	**99.90**	**235752**
5	1.664	71	0.02	32
6	1.829	51	0.02	42
7	1.890	47	0.01	34

CBzGlyProProGlyOEt (5.2)



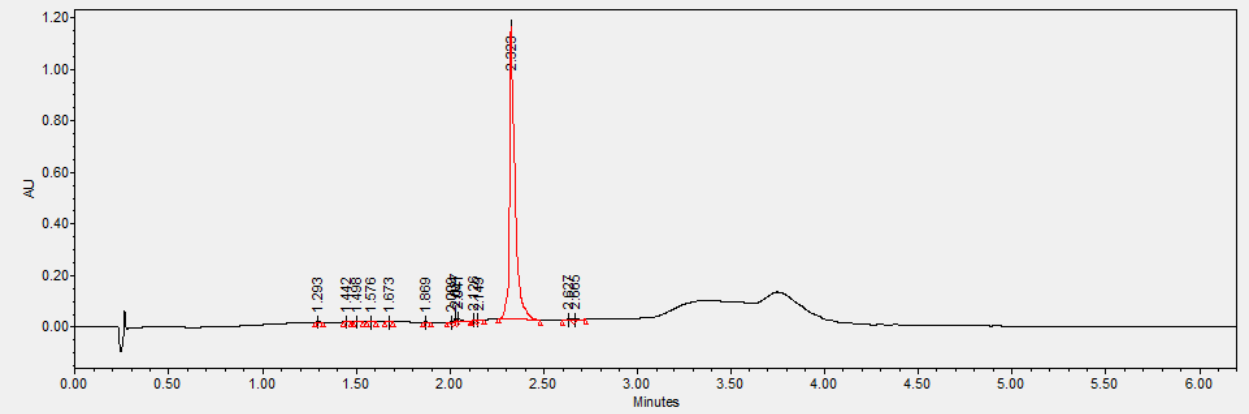
Peak	Retention Time (min)	Area (µV*sec)	Area (%)	Height (µV)
1	1.728	439	0.01	744
2	1.754	29212	0.96	20800
3	1.82	13229	0.44	11425
4	1.893	23989	0.79	15200
5	1.948	38104	1.26	32938
6	**2.017**	**2930062**	**96.54**	**1382473**

DansylGlyPro*trans*ProGlyOEt (5.3)



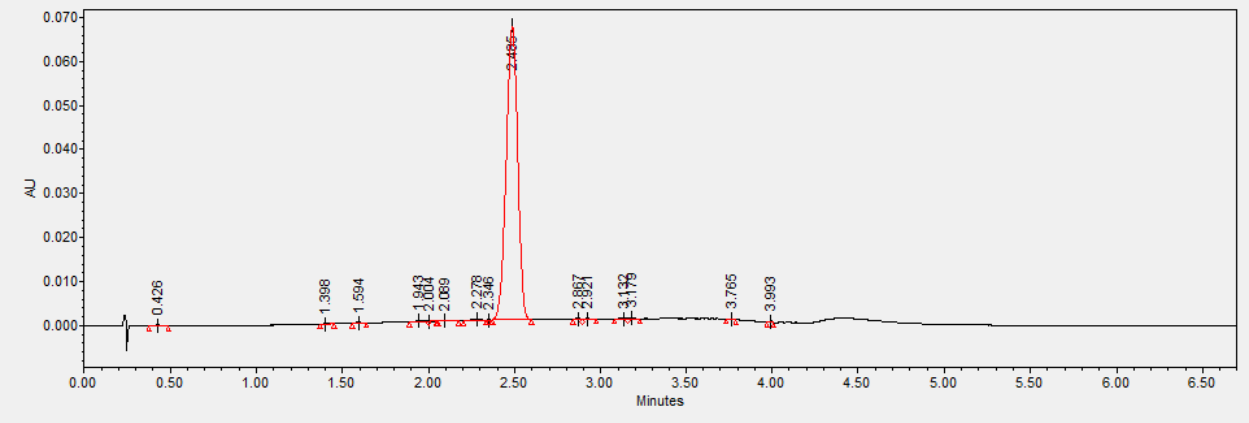
Peak	Retention Time (min)	Area (µV*sec)	Area (%)	Height (µV)
1	1.480	448	0.21	284
2	1.601	103	0.05	71
3	1.875	96	0.05	93
4	1.918	333	0.16	332
5	**1.973**	**210229**	**99.18**	**159413**
6	2.029	211	0.10	209
7	3.986	553	0.26	646

CBzGlyPro*trans*ProGlyOEt (5.4)



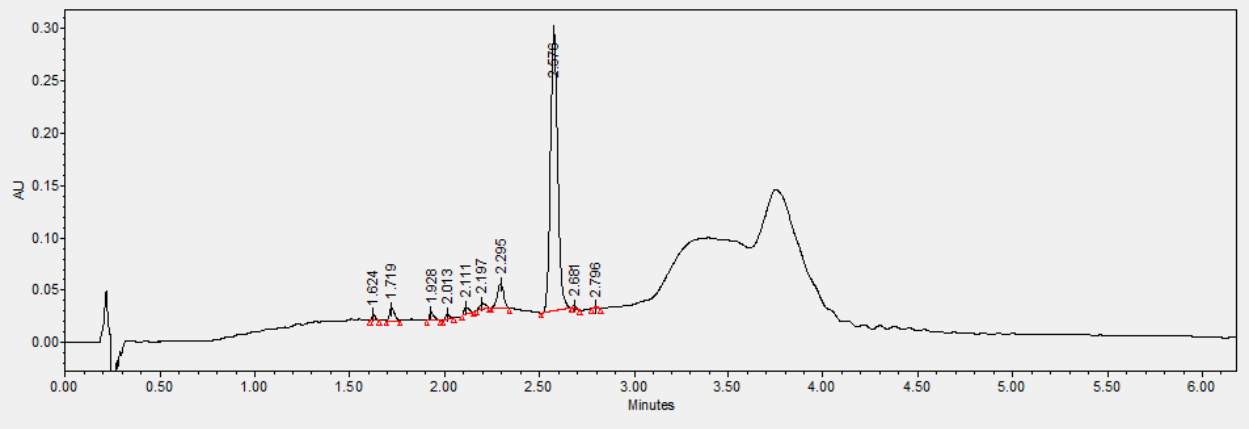
Peak	Retention Time (min)	Area (µV*sec)	Area (%)	Height (µV)
1	1.293	2810	0.12	3408
2	1.442	2064	0.09	2183
3	1.498	2953	0.13	2639
4	1.576	1344	0.06	1150
5	1.673	1670	0.07	1225
6	1.869	2355	0.1	2076
7	2.009	1691	0.07	1976
8	2.027	15496	0.66	16469
9	2.041	17644	0.75	12622
10	2.126	3890	0.17	3219
11	2.149	2466	0.1	1974
12	**2.323**	**2295591**	**97.41**	**1136083**
13	2.627	4690	0.2	3286
14	2.665	2056	0.09	1320

DansylGlyPro*cis*ProGlyOEt (5.5)



Peak	Retention Time (min)	Area (µV*sec)	Area (%)	Height (µV)
1	0.426	369	0.12	241
2	1.398	111	0.04	73
3	1.594	318	0.10	222
4	1.943	622	0.20	385
5	2.004	61	0.02	41
6	2.089	403	0.13	157
7	2.278	1333	0.43	393
8	2.346	113	0.04	127
9	**2.485**	**308032**	**98.34**	**66713**
10	2.867	367	0.12	238
11	2.921	301	0.10	132
12	3.132	302	0.10	135
13	3.179	481	0.15	306
14	3.765	135	0.04	123
15	3.993	298	0.10	352

CBzGlyPro*cis*ProGlyOEt (5.6)



Peak	Retention Time (min)	Area (μV*sec)	Area (%)	Height (μV)
1	1.624	6708	0.79	5748
2	1.719	21346	2.53	11731
3	1.928	13352	1.58	8013
4	2.013	7427	0.88	5055
5	2.111	12284	1.45	6750
6	2.197	12457	1.47	5434
7	2.295	55179	6.53	23049
8	**2.576**	**711750**	**84.24**	**266043**
9	2.681	2446	0.29	2473
10	2.796	2007	0.24	1519

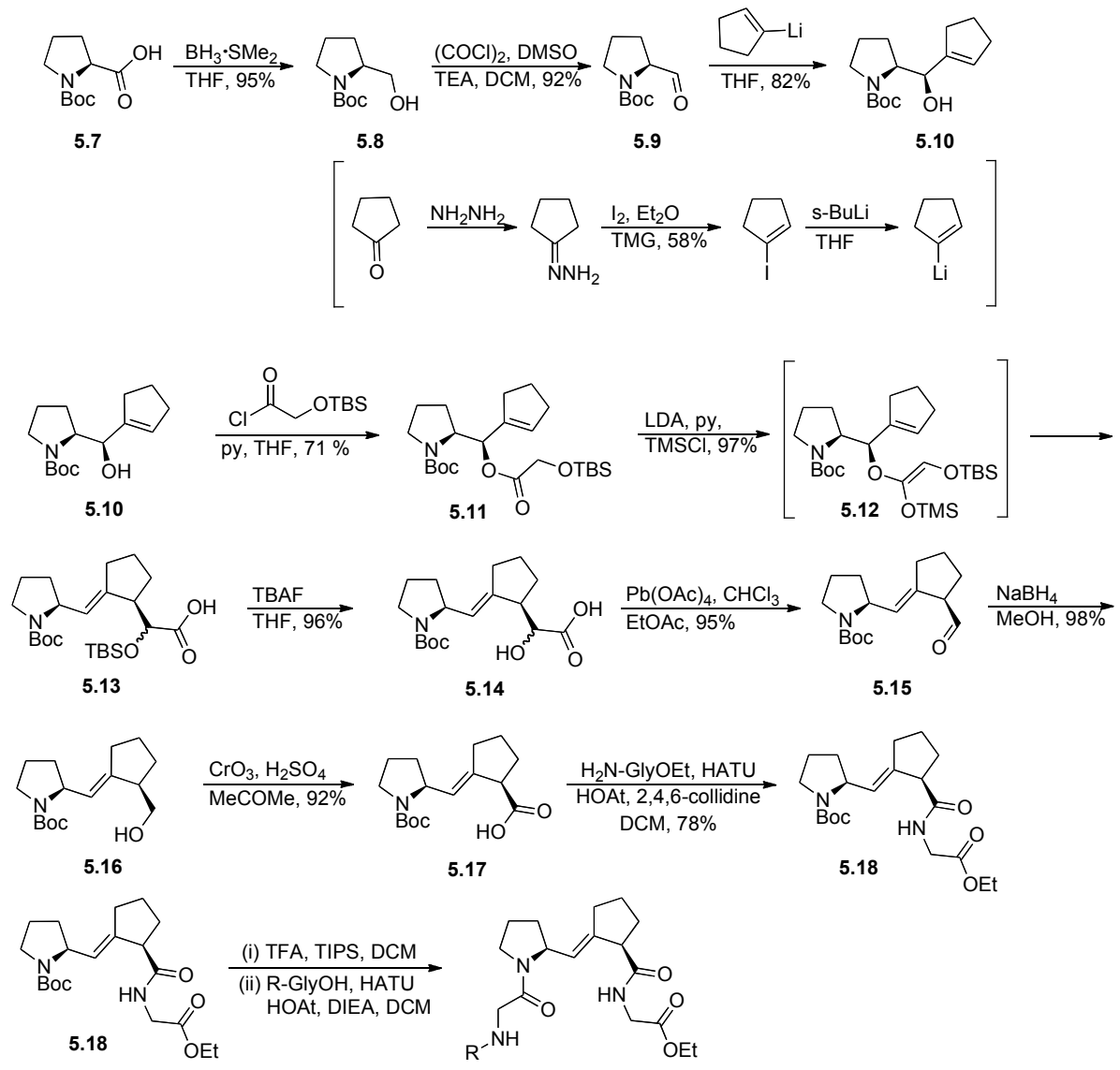
5.6 Acknowledgments

This work was supported by Grant R01 AR044276 (NIH). J. D. V. was supported by Molecular Biosciences Training Grant T32 GM007215 (NIH) and a fellowship from the Department of Biochemistry at the University of Wisconsin–Madison. N. A. M. was supported by postdoctoral fellowship F32 GM096712 (NIH). This study made use of the National Magnetic Resonance Facility at Madison, which is supported by Grant P41 GM103399 (NIH). The Micromass LCT® mass spectrometer was obtained with support from Grant CHE-9974839 (NSF).

5.7 Author Contributions

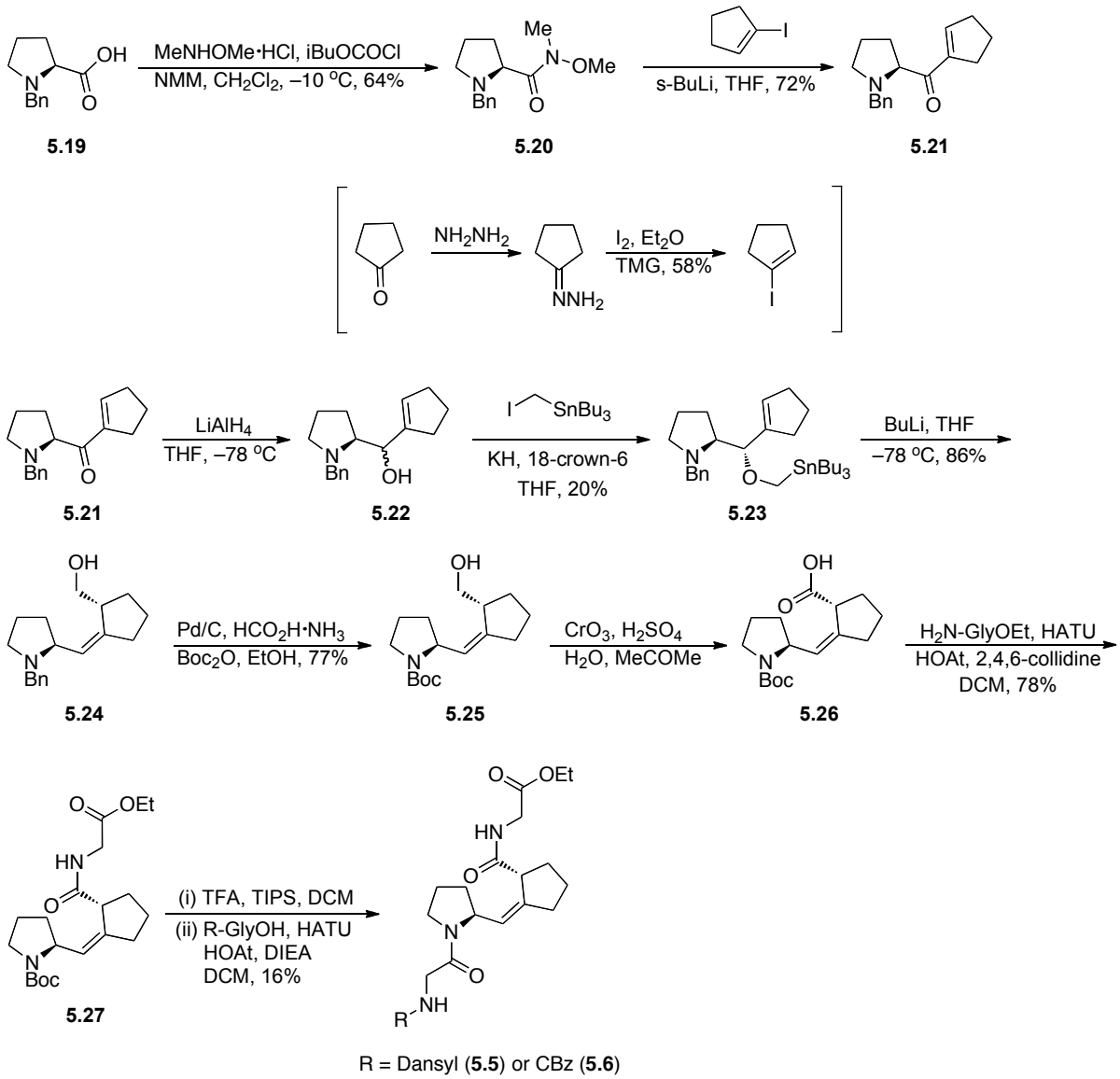
J. D. Vasta, A. Choudhary, K. H. Jensen, and N. A. McGrath all contributed to the synthesis and characterization of peptides and corresponding alkene isosteres. J. D. Vasta designed and carried out all experiments related to the assay of prolyl 4-hydroxylases. J. D. Vasta, A. Choudhary, K. H. Jensen, and R. T. Raines contributed to the analysis of data and preparation of the manuscript and figures.

Scheme 5.1 Synthetic route to GP*trans*PGOEt alkene isosteres 5.3 and 5.4



R = Dansyl (**5.3**) or CBz (**5.4**)

Scheme 5.2 Synthetic route to GP*cis*PGOEt alkene isosteres 5.5 and 5.6



R = Dansyl (**5.5**) or CBz (**5.6**)

Figure 5.1 Structural complexities of proline residues

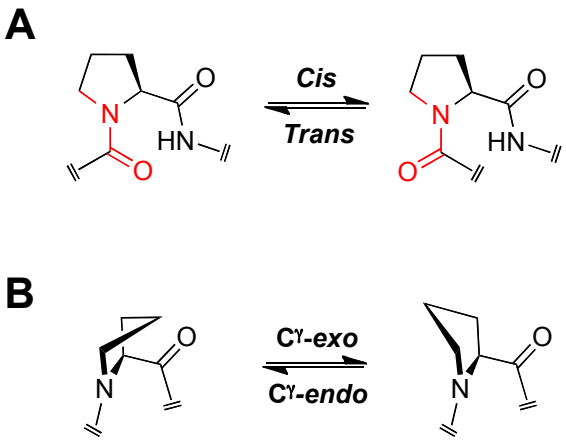


Figure 5.1 Structural complexities of proline residues. (A) Because proline residues are tertiary amides, the *cis*-conformation of the prolyl peptide bond occurs more frequently than in non-prolyl residues. (B) The pyrrolidine ring of proline residues can exist in one of two ring puckers.

Figure 5.2 Model peptide substrates and locked alkene isosteres

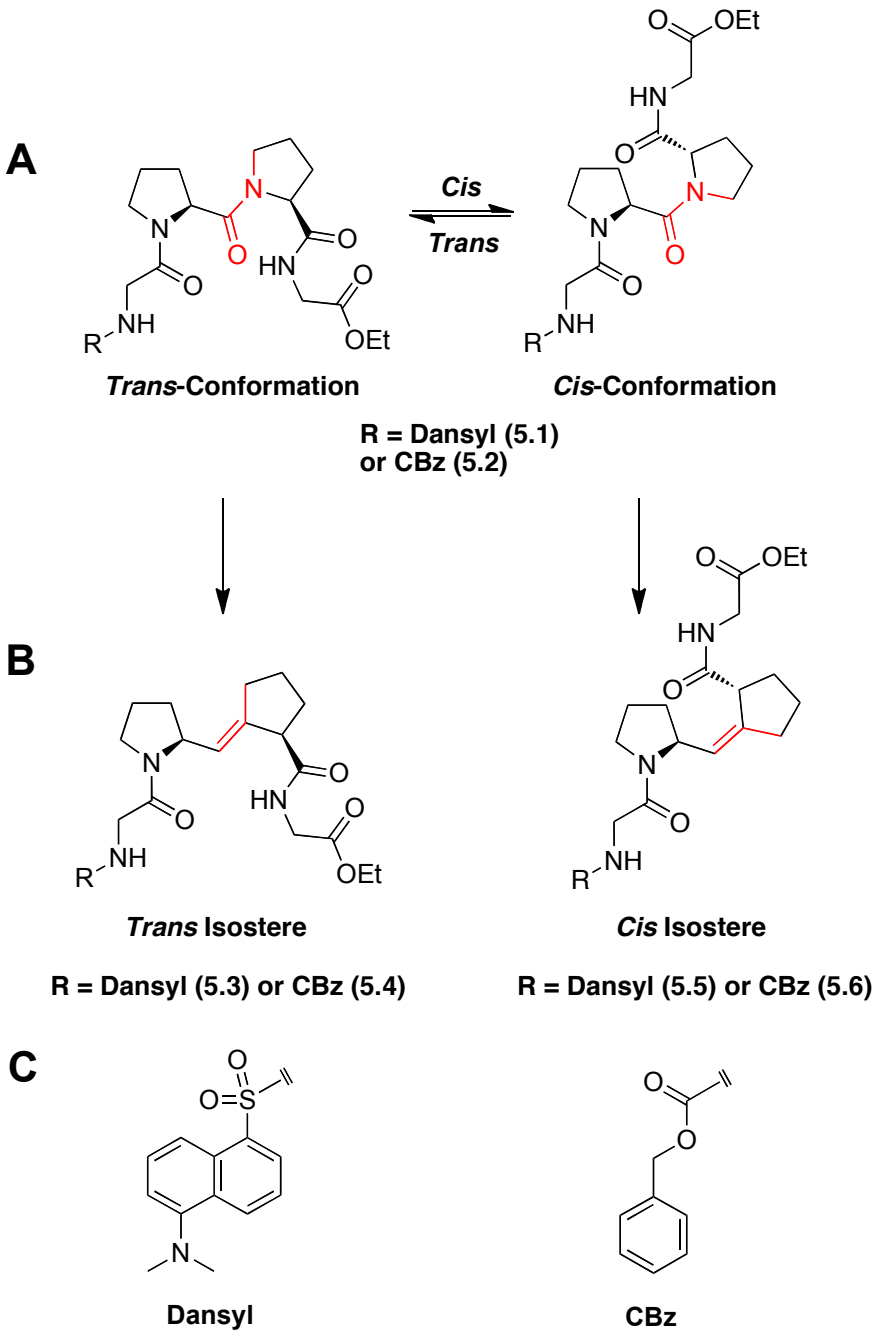


Figure 5.2 Model peptide substrates and locked alkene isosteres. (A) The model tetrapeptides R-GlyProProGly-OEt **5.1** and **5.2** can adopt either the *cis*- or *trans*- conformation in solution. (B) Corresponding alkene isosteres of the *trans*- (**5.3** and **5.4)** and *cis*- (**5.5** and **5.6**) conformations. (C) Renditions of the dansyl and CBz N-terminal caps used for the model peptides and alkene isosteres.

Figure 5.3 Inhibition of human CP4H1 by dansyl isosteres 5.3 and 5.5

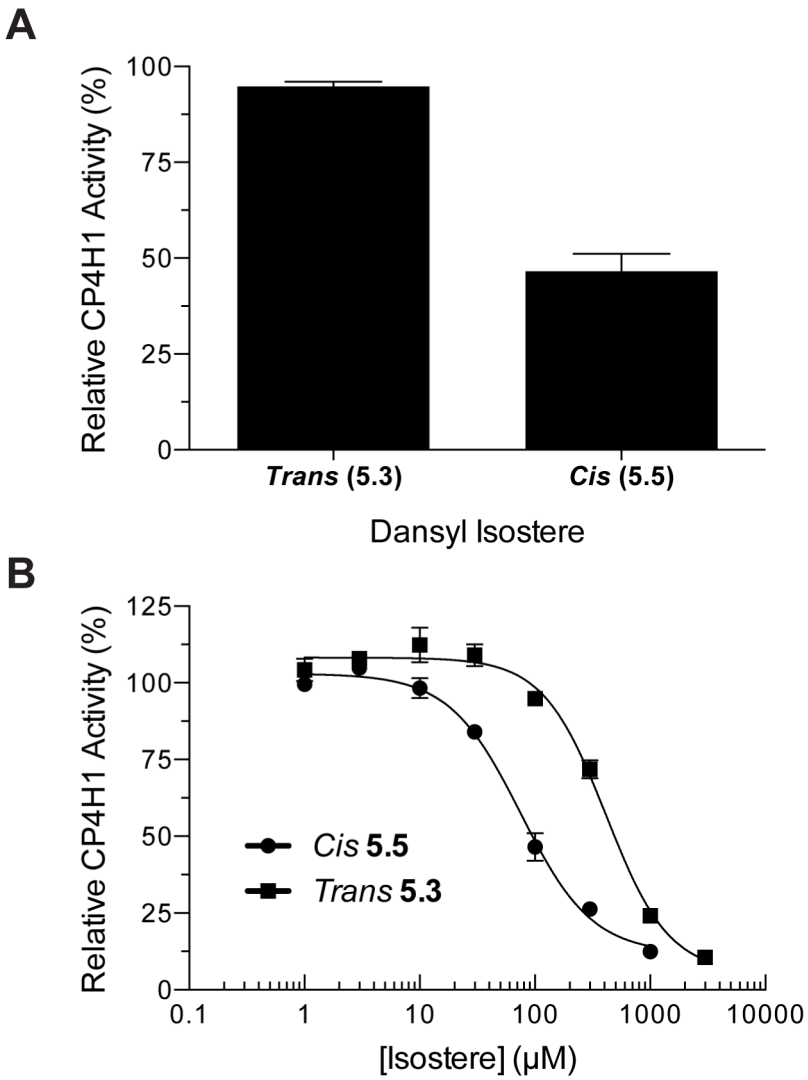


Figure 5.3 Inhibition of human CP4H1 by dansyl isosteres **5.3** and **5.5**. CP4H assays were conducted as described in the Experimental Procedures section using 1000 μM AKG, 500 μM dansylGPPGOEt substrate, and 100 μM dansyl isostere inhibitor. Relative activity values are the mean (± SD) of three independent experiments and represent the ratio of CP4H activity observed in the presence and absence of the isostere.

Figure 5.4 Inhibition of human CP4H1 by CBz isosteres 5.4 and 5.6

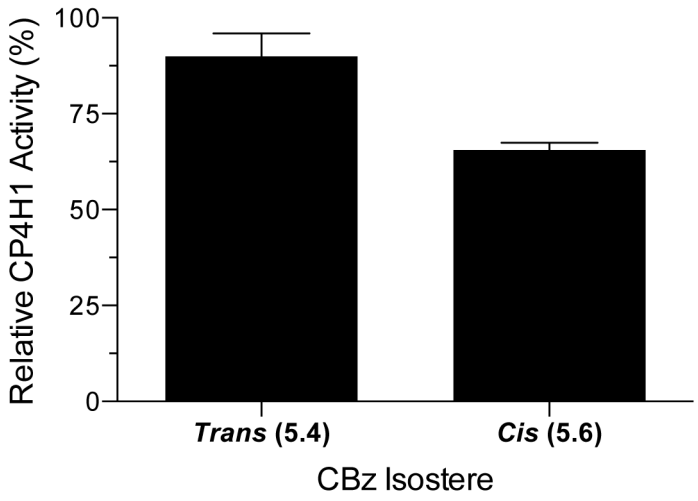
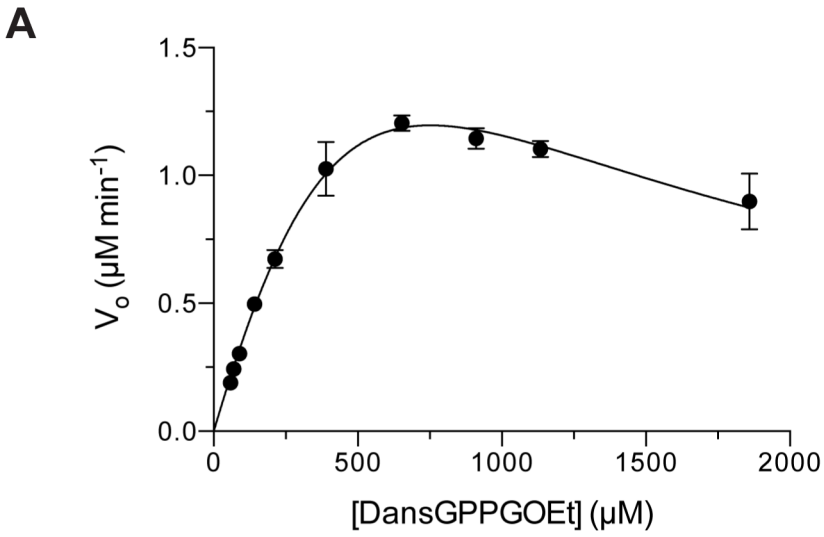
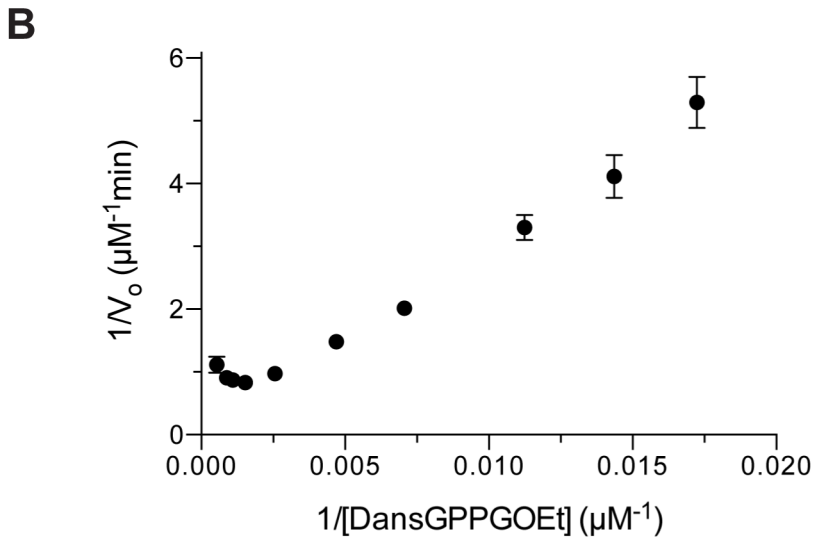


Figure 5.4 Inhibition of human CP4H1 by CBz isosteres **5.4** and **5.6**. CP4H assays were conducted as described in the Experimental Procedures section using 1000 μM AKG, 1 mM CBzGPPGOEt substrate, and 1 mM CBz isostere inhibitor. Relative activity values are the mean (± SD) of three independent experiments and represent the ratio of CP4H activity observed in the presence and absence of the isostere.

Figure 5.5 Substrate inhibition of human CP4H1 by dansylGPPGOEt (5.1)





C

$$v = \frac{V_{max} * [S]}{K_m + [S] + \frac{[S]^2}{K_i}}$$

Figure 5.5 Substrate inhibition of human CP4H1 by dansylGPPGOEt (**5.1**). (A) Velocity plot for human CP4H1 with **5.1** as the substrate. Assays were conducted as described in the Experimental Procedures section using 1000 μM AKG and 58–1860 μM **5.1**. (B) A Lineweaver-Burke plot of the data in panel A, further demonstrating substrate inhibition at high concentrations of **5.1**. (C) The equation for substrate inhibition, where K_i is the dissociation constant for substrate binding in an inhibitory manner, was fit to the data in panel A. While the fit was in reasonable agreement with the data ($R^2 = 0.98$), interpolated values of V_{max}, K_m, and K_i were 7.5 ± 2.9 μM min^{-1}, 2000 ± 850 μM, and 290 ± 130 μM, respectively, with standard errors that are prohibitively large. The limited solubility of **5.1** prohibited meaningful improvement of these interpolated parameters.

Figure 5.6 Lineweaver–Burke analysis of *cis* isostere 5.5

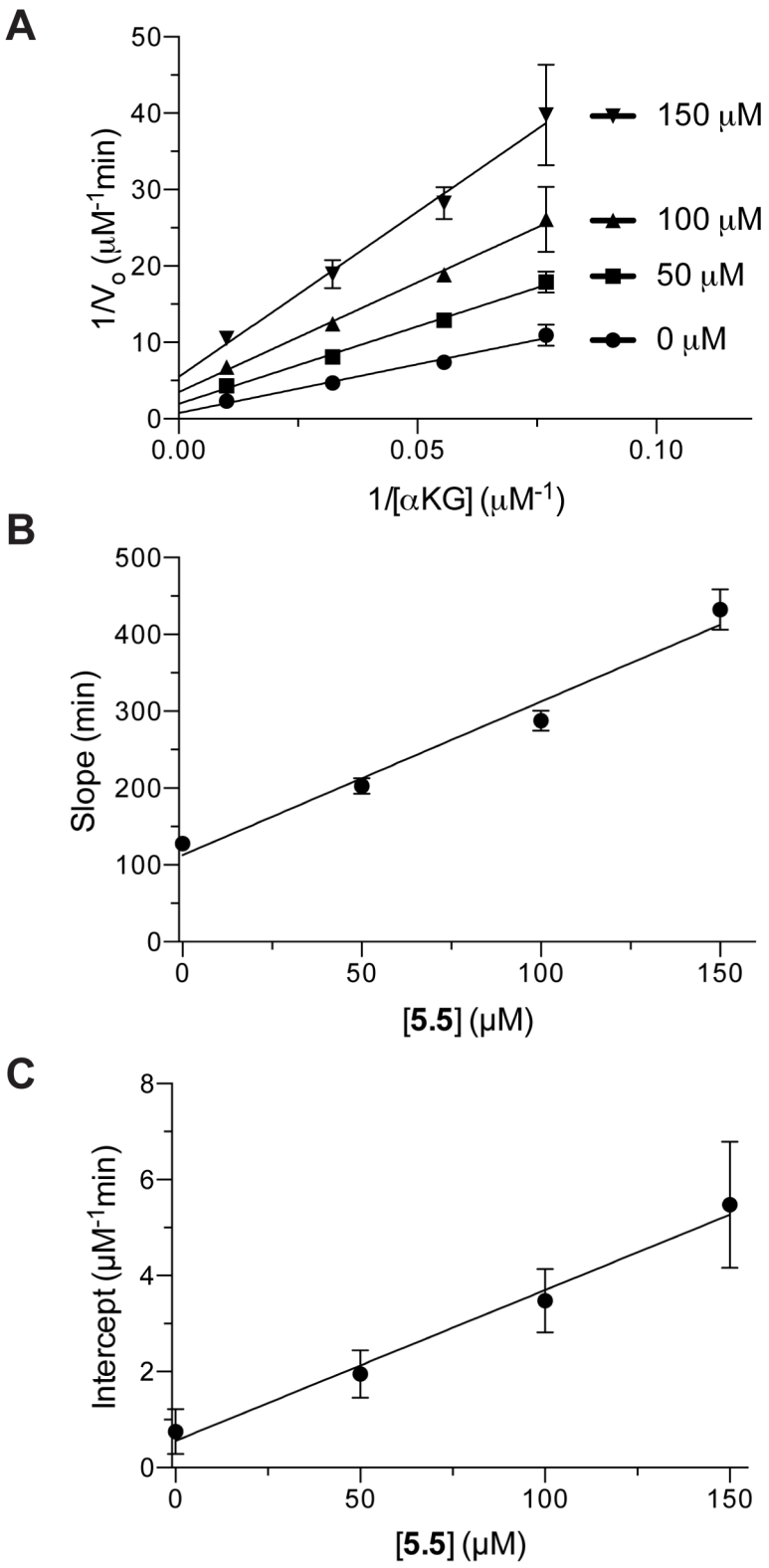


Figure 5.6 Lineweaver–Burke analysis of *cis* isostere **5.5**. (A) A Lineweaver-Burke plot for **5.5** suggests noncompetitive inhibition with respect to AKG, as evidenced by the intersecting pattern that does not converge on the ordinate and linearity in the slope (B) and intercept (C) replots. The rate of CP4H catalysis with increasing AKG concentration (13–100 μM) was determined in the presence of fixed inhibitor concentrations (specified above) as described in the Experimental Procedures section. Individual points represent the mean (± SE) of three independent experiments.

Figure 5.7 Substrate inhibition of CrP4H-1 by dansylGPPGOEt (5.1)

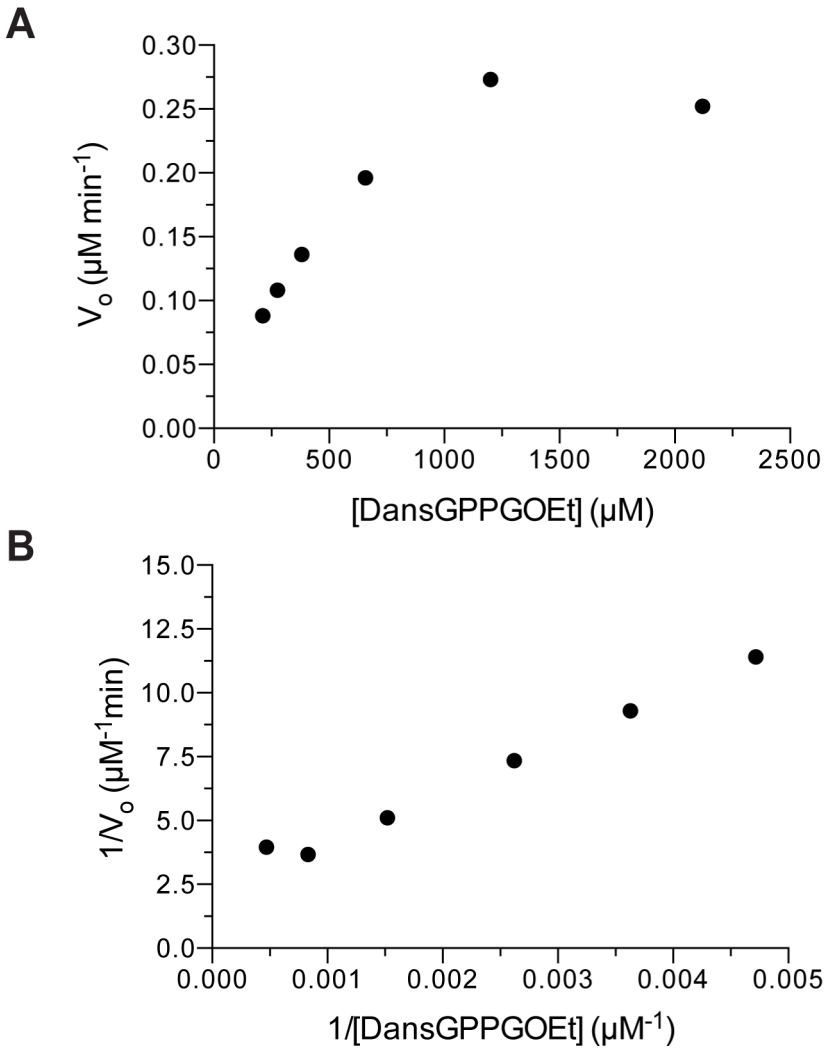


Figure 5.7 Substrate inhibition of CrP4H-1 by dansylGPPGOEt (**5.1**). (A) A Velocity plot for CrP4H-1 with **5.1** as the substrate. Assays were conducted as described in the Experimental Procedures section using 1000 μM AKG and 212–2120 μM **5.1**. (B) A Lineweaver-Burke plot of the data in panel A, further demonstrating substrate inhibition at high concentrations of **5.1**.

Figure 5.8 Inhibition of CrP4H-1 by dansyl isosteres 5.3 and 5.5

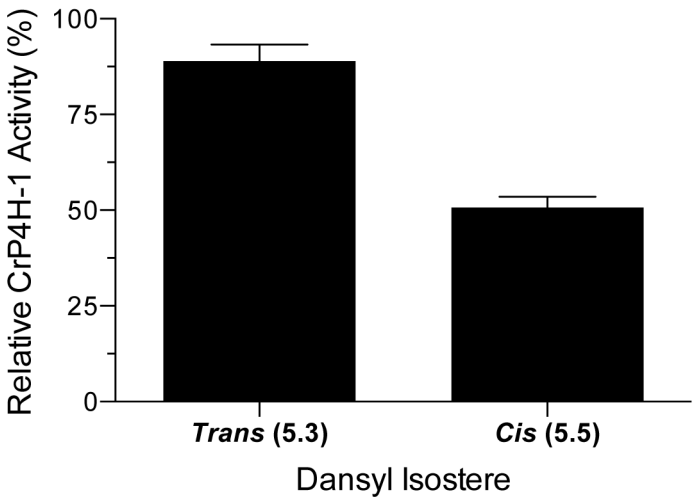
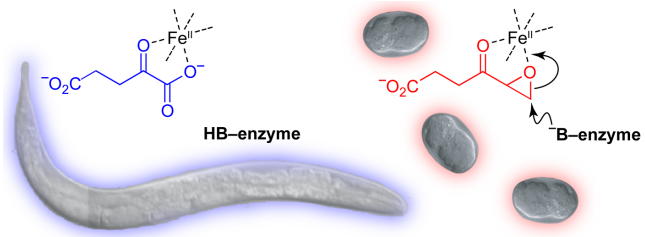


Figure 5.8 Inhibition of CrP4H-1 by dansyl isosteres **5.3** and **5.5**. Assays were conducted as described in the Experimental Procedures section using 1000 μM AKG, 500 μM dansylGPPGOEt substrate, and 100 μM dansyl isostere inhibitor. Relative activity values are the mean (± SD) of three independent experiments and represent the ratio of CrP4H-1 activity observed in the presence and absence of the isostere.

CHAPTER 6

Bioavailable Affinity Label for Collagen Prolyl 4-Hydroxylase*



*This chapter has been published in part under the same title. Vasta, J.D., Higgin, J.J., Kersteen, E.A., and Raines, R.T. Bioavailable affinity label for collagen prolyl 4-hydroxylase. J.D.V., E.A.K, and J.J.H. contributed to the design and execution of experiments. All authors contributed to the analysis of data and preparation of the manuscript and figures.

6.1 Abstract

Collagen is the most abundant protein in animals. Its prevalent 4-hydroxyproline residues contribute greatly to its conformational stability. The hydroxyl groups arise from a post-translational modification catalyzed by the non-heme iron-dependent enzyme, collagen prolyl 4-hydroxylase (P4H). Here, we report that 4-oxo-5,6-epoxyhexanoate, a mimic of the α-ketoglutarate co-substrate, inactivates human P4H. The inactivation installs a ketone functionality in P4H, providing a handle for proteomic experiments. *Caenorhabditis elegans* exposed to the esterified epoxy ketone displays the phenotype of a worm lacking P4H. Thus, this affinity label can be used to mediate collagen stability in an animal, as is desirable in the treatment of a variety of fibrotic diseases.

6.2 Introduction

Collagen is the most abundant protein in animals, and maintains the structural integrity of various connective tissues.[1] The overproduction of collagen is associated with fibrotic diseases, including myocardial infarction, stroke, peripheral vascular disease, diabetes, and severe anemias.[30] The stability of collagen relies on catalysis by collagen prolyl 4-hydroxylase (P4H).[4] This enzyme is located in the lumen of the endoplasmic reticulum, where it catalyzes a remarkable post-translational modification—the conversion of (2S)-proline residues (Pro) to (2S,4R)-4-hydroxyproline residues (Hyp) in protocollagen strands. The human enzyme consists of an $\alpha_2\beta_2$ tetramer in which each α subunit contains an active site. P4H is essential for animals,[81-83] as the conformational stability of mature collagen relies on its hydroxylation.[5-7] Moreover, P4H is a validated target for the treatment of fibrotic diseases.[30]

P4H is a non-heme iron(II) dioxygenase that uses O_2 and α-ketoglutarate as co-substrates (Figure 6.1A). The three-dimensional structure of mammalian P4H is unknown. We reasoned that an electrophilic analog of α-ketoglutarate could serve as an irreversible inhibitor of the enzyme and, hence, a useful probe for active-site residues. Inspired by natural products such as trapoxin and epolactaene as well as artificial affinity labels,[194] we designed 4-oxo-5,6-epoxyhexanoate (**6.1**) for this purpose. Our thought was that the epoxide oxygen would chelate to the active-site iron and thereby be activated for nucleophilic attack (Figure 6.1B). We also noted that α-ketoglutarate and epoxy ketone **6.1** have the same number and a similar arrangement of non-hydrogen atoms. This attribute is important because of the inability of larger α-

ketoglutarate analogs to inhibit catalysis by human P4H.[195] Epoxy ketone **6.1**, which also resembles 5-aminolevulinic acid, is a known competitive inhibitor of 5-aminolevulinic acid dehydratase but does not alkylate that enzyme.[196]

6.3 Results and Discussion

Effect of Epoxy Ketone 6.1 on Catalysis by P4H

We produced human P4H by heterologous expression in Escherichia coli,[122,140,185,197,198] and we synthesized racemic **6.1**. We assayed the effect of this putative affinity label on the ability of P4H to catalyze the hydroxylation of a fluorescent peptidic substrate, dansyl-Gly–Phe–Pro–Gly-OEt.[122,140,185,198] The data revealed that epoxy ketone **6.1** elicits irreversible inactivation of P4H, and does so in a concentration-dependent manner (Figure 6.2A). α-Ketoglutarate protected P4H against inactivation (Figure 6.2B), suggesting that epoxy ketone **6.1** competes with this co-substrate for binding to the active site.

Utility of Epoxy Ketone 6.1 as a Proteomic Probe

Noting that a ketone has chemical reactivity orthogonal to that in natural amino acids,[199] we examined the utility of epoxy ketone **6.1** in a model proteomic experiment. Epoxy ketone **6.1** was incubated with P4H in the presence of iron(II). The mixture was treated with biocytin hydrazide and then $NaBH_4$ to reduce the incipient *N*-acyl hydrazone (which is unstable[200]). Small molecules were removed by gel-filtration chromatography. The P4H was added to wells of a

microplate containing immobilized streptavidin. The wells were washed and exposed to a monoclonal antibody directed against P4H. Visualization with a secondary antibody conjugated to horseradish peroxidase revealed that P4H is associated with streptavidin only in the presence of epoxy ketone **6.1** (Figure 6.2C). The conjugation was confirmed by SDS–PAGE (data not shown). These data indicate that epoxy ketone **6.1** could serve as a chemoselective probe for P4H. Thus, epoxy ketone **6.1** complements a (2*S*,4*S*)-4-fluoroproline substrate that spawns a ketone upon P4H catalysis.[185]

Effect of epoxy ketone 6.1 on an animal

Finally, we determined the effect of epoxy ketone **6.1** on a live animal. Like all animals (including dinosaurs[201]), *Caenorhabditis elegans* has an abundance of collagen. In wild-type worms, P4H exists largely as a tetramer of two α subunits (DPY-18 and PHY-2, which each contain an active site) and two β subunits.[202] The *dpy-18* gene is essential for body morphology, whereas the function of the *phy-2* gene is not apparent.[81,83] The deletion of both genes results in embryos that grow to the twofold stage, but are unable to maintain their shape and consequently explode.[81] These two forms of the α subunit have 56.5% amino-acid sequence identity with each other and ~50% identity with their human homologues. α-Ketoglutarate analogs can eliminate the remaining P4H activity in *dpy-18 C. elegans* at concentrations that do not affect the wild-type worm.[81] This synthetic lethality provided the basis for our assay.

Dpy-18 worms were placed onto solid medium containing known concentrations of esterified epoxy ketone **6.1**. The ester group, which can be hydrolyzed by cellular esterases,

served to mask the anionic carboxylate and thus enhance uptake.[81,158] Concentration-dependent embryonic lethality of *dpy-18 C. elegans* was observed with LD_{50} = 0.11 mM (Figure 6.3). These data are consistent with lethality being due to P4H inactivation.

6.4 Conclusions

The need for small-molecule probes and inhibitors of P4H is evident from the lack of structural information about this essential enzyme as well as the list of fibrotic diseases associated with collagen overproduction.[30] The newly discovered role of P4Hs in hypoxia[53,54,203] and cancer,[204] and the cardioprotection conferred by P4H inhibitors[205-207] increase the imperative.[67] Our identification of 4-oxo-5,6-epoxyhexanoate as an affinity label for P4H that is bioavailable upon esterification provides a means to address this need. On-going efforts in our laboratory are directed at identifying the alkylated enzymic residue, improving the potency of the epoxy ketone as an affinity label, and evaluating its selectivity towards other P4Hs.

6.5 Experimental Procedures

6.5.1 General

BakerDry™ solvents in cyclotainers™ (DMF, ≤20 ppm water; CH_2Cl_2, ≤30 ppm water; THF, ≤10 ppm water) were from J. T. Baker (Phillipsburg, NJ). Unless noted otherwise, all other chemicals were from Aldrich Chemical (Milwaukee, WI), Fisher Scientific (Pittsburgh, PA), or NovaBiochem (San Diego, CA), and were used without further purification.

Human collagen prolyl 4-hydroxylase (P4H) was produced in *E. coli* and purified as described previously.[122,140,185,197,198] Dansyl-Gly–Phe–Pro–Gly-OEt, a fluorescent P4H substrate, was synthesized as described previously.[122,140,185,198]

In the synthetic procedures, the term "concentrated under reduced pressure" refers to the removal of solvents and other volatile materials using a rotary evaporator at water aspirator pressure (<20 torr) while maintaining the water-bath temperature below 50 °C.

6.5.2 Instrumentation

NMR spectra were obtained with a Bruker AC+ 300 MHz, Bruker DMX 400 WB, or Bruker Avance III 500i spectrometer. Mass spectra were obtained with a Micromass LCT® electrospray ionization, time-of-flight analyzer from Waters (Milford, MA). HPLC analyses were preformed with an instrument from Waters controlled with the Millenium (version 3.20) software package and equipped with two 515 pumps, a 717plus autosampler, and a 996 photodiode array detector.

6.5.3 Synthesis of 4-Oxo-5,6-epoxyhexanoate (6.1)

Epoxy ketone **6.1** was synthesized in 5 steps with an overall yield of 38% by the route shown in Scheme 6.1 and described below.

Benzyl Succinic Acid (6.2)

Benzyl succinic acid was synthesized by following a known procedure.[208] Here, succinic anhydride (30 g, 300 mmol) was dissolved in benzyl alcohol (31.6 mL, 33 g, 300 mmol), and the resulting solution was heated at reflux for 4 h. The reaction mixture was dissolved in ether (100 mL), and the insoluble succinic acid was removed by filtration. The filtrate was extracted with saturated aqueous Na_2CO_3 (3 × 100 mL), and the combined aqueous extracts were acidified with 2 M HCl (2.0 L). The precipitate was collected by filtration and dried under vacuum to give **6.2** (30 g, 48%). **^{1}H NMR** ($CDCl_3$, 300 MHz) δ 7.35 (m, 5H), 5.15 (s, 2H), 2.70 (m, 4H). **^{13}C NMR** ($CDCl_3$, 75 MHz) δ 177.7, 172.1, 135.7, 128.8, 128.5, 128.4, 66.9, 29.1, 29.0.

Benzyl 4-Chlorooxobutyrate (6.3)

Benzyl 4-chlorooxobutyrate was synthesized by following a known procedure.[209] Here, benzyl succinic acid (4.14 g, 20 mmol) was dissolved in CH_2Cl_2 (150 mL), and the resulting solution was cooled to 0 °C. Oxalyl chloride (13.8 g, 109 mmol) was added slowly, and the reaction mixture was stirred at 0 ℃ for 5 min. The reaction mixture was allowed to warm to 15 ℃ and stirred for an additional 30 min. The reaction mixture was then concentrated under reduced pressure, and the residue was dissolved in benzene. The residue was concentrated under reduced

pressure to give **6.3** (4.5 g, 99%) as a clear viscous liquid. **^{1}H NMR** ($CDCl_3$, 300 MHz) δ 7.34 (m, 5H), 5.13 (s, 2H), 3.19 (t, J = 7 Hz, 2H), 2.69 (t, J = 7 Hz, 2H). **^{13}C NMR** ($CDCl_3$, 75 MHz) δ 173.2, 170.9, 135.6, 128.8, 128.6, 128.4, 67.1, 29.5, 28.5.

Benzyl 4-Oxo-5-hexenoate (6.4)

Benzyl 4-oxo-5-hexenoate (**6.4**) was synthesized in 82% yield from benzyl 4-chlorooxobutyrate in the same manner as was ethyl 4-oxo-5-hexenoate (6) from ethyl 4-chlorooxobutyrate, *vide infra*. **^{1}H NMR** ($CDCl_3$, 300 MHz) δ 7.35 (m, 5H), 6.38 (dd, J = 10 and 18 Hz, 1H), 6.26 (dd, J = 1 and 18 Hz, 1H), 5.86 (dd, J = 1 and 10 Hz, 1H), 5.13 (s, 2H), 2.94 (t, J = 7 Hz, 2H), 2.70 (t, J = 7 Hz, 2H).

Benzyl 4-Oxo-5,6-epoxyhexanoate (6.5)

Benzyl 4-oxo-5,6-epoxyhexanoate (**6.5**) was synthesized in 98% yield from benzyl 4-oxo-5-hexenoate in the same manner as was ethyl 4-oxo-5,6-epoxyhexanoate (**6.6**) from ethyl 4-oxo-5-hexenoate (**6.7**), *vide infra*. **^{1}H NMR** ($CDCl_3$, 300 MHz) δ 7.34 (m, 5H), 5.10 (s, 2H), 3.46 (dd, J = 2 and 5 Hz, 1H), 2.99 (dd, J = 5 and 6 Hz, 1H) 2.94 (dd, J = 2 and 6 Hz, 1H), 2.64 (m, 4H). **^{13}C NMR** ($CDCl_3$, 75 MHz) δ 206.1, 172.4, 135.9, 128.7, 128.5, 128.4, 66.6, 53.7, 46.3, 31.3, 27.5.

4-Oxo-5,6-epoxyhexanoic Acid (1, free acid)

Benzyl 4-oxo-5,6-epoxyhexanoate (340 mg, 1.45 mmol) was dissolved in EtOAc (20 mL). Pd/C (10% w/w, 40 mg) was added, and the reaction mixture was stirred under H_2(g) for 4 h. The reaction mixture was filtered through Celite®, and concentrated under reduced pressure to give the free acid of epoxy ketone **6.1** (183 mg, 99%). **ESI MS EMM** *m*/*z*: $[M - H]^-$ 143.0340, calcd. 143.0344. **^{1}H NMR** ($CDCl_3$, 300 MHz) δ 10.74 (br s, 1H), 3.48 (dd, J = 2 and 5 Hz, 1H). 3.03 (dd, J = 5 and 5 Hz, 1H), 2.96 (dd, J = 2 and 5 Hz, 1H), 2.64 (m, 4H). **^{13}C NMR** ($CDCl_3$ 75 MHz) δ 206.1, 178.3, 53.7, 46.4, 31.1, 27.2.

Sodium 4-Oxo-5,6-epoxyhexanoate (1, sodium salt)

4-Oxo-5,6-epoxyhexanoic acid (206 mg, 1.4 mmol) was dissolved in aqueous $NaHCO_3$ (3.6 mL, 400 mM), and the resulting solution was stirred for 1 h. The reaction mixture was concentrated to dryness by lyophilization to give the water-soluble sodium salt of epoxy ketone **6.1** (237 mg, quant), which was used for assays *in vitro*. **^{1}H NMR** (D_2O, 500 MHz) δ 3.69 (dd, J = 3 and 5 Hz, 1H), 3.00 (t, J = 5 and 6 Hz, 1H), 2.80 (dd, J = 3 and 6 Hz, 1H), 2.70 (m, 2H), 2.34 (t, J = 7 Hz, 2H). **^{13}C NMR** (D_2O, 125 MHz) δ 211.7, 182.1, 54.2, 48.7, 36.1, 31.5.

6.5.4 Synthesis of Ethyl 4-Oxo-5,6-epoxyhexanoate (6.6, esterified 6.1)

Esterified **6.1** was synthesized in 2 steps with an overall yield of 70% by the route shown in Scheme 6.2 and described below.

Ethyl 4-Oxo-5-hexenoate (6.7)

Ethyl 4-oxo-5-hexenoate was synthesized by following a known procedure.[210] Here, ethyl 4-chlorooxobutyrate (2.5 g, 15.2 mmol) was dissolved in HMPA (10 mL). Tributylvinyltin (5.0 g, 15.2 mmol) and benzylchlorotriphenylphosphine palladium(II) (52 mg) were added, and the reaction mixture was heated open to air to 65 °C for 5 min. At that time, the reaction mixture turned black. The reaction mixture was allowed to cool to room temperature. H_2O (50 mL) and NaCl (1.0 g) were added, and the resulting solution was stirred for 5 min. The solution was extracted with diethyl ether (3 × 60 mL). The combined organic extracts were washed with H_2O (4 × 60 mL), and the organic layer was dried over $MgSO_4$(s) and concentrated under reduced pressure to ~20 mL. A half-saturated solution of KF in MeOH (25 mL) was added to precipitate the $SnBu_3F$, which was removed by filtration through Celite®. The filtrate was concentrated under reduced pressure and purified by chromatography on silica gel, eluting with ethyl acetate (10% v/v) in CH_2Cl_2 to give **6.7** (1.944 g, 82%). **ESI MS** *m*/*z*: $[M + Na]^+$ 179.0, calcd 179.1. **^{1}H NMR** ($CDCl_3$, 400 MHz) δ 6.35 (dd, J = 10 and 18 Hz, 1H), 6.23 (dd, J = 1 and 18 Hz, 1H), 5.84 (dd, J = 1 and 10 Hz, 1H), 4.10 (q, J = 7 Hz, 2H), 2.89 (t, J = 7 Hz, 2H), 2.59 (t, J = 7 Hz, 2H), 1.21 (t, J = 7, 3H). **^{13}C NMR** ($CDCl_3$, 100 MHz) δ 198.5, 172.7, 136.2, 128.4, 60.6, 34.0, 27.9, 14.1.

Ethyl 4-Oxo-5,6-epoxy-hexanoate (6.6, esterified 6.1)

Ethyl 4-oxo-5,6-epoxyhexanoate was synthesized by following a known procedure.[196] Here, ethyl 4-oxo-5-hexenoate (1.107 g, 7.09 mmol) and H_2O_2 (30% v/v, 1.75 mL) were dissolved in *t*BuOH (11.81 mL) and H_2O (5.9 mL) at 0 °C. Saturated aqueous K_2CO_3 (606 μL) at 0 °C was added slowly. The reaction mixture was stirred for 1 h at 0 °C, allowed to warm to room temperature, and then stirred for an additional 3 h. The reaction mixture was diluted with H_2O (100 mL) and extracted with diethyl ether (3 × 30 mL). The combined organic extracts were dried over $MgSO_4$(s) and concentrated under reduced pressure. The product was purified by chromatography on silica gel, eluting with ethyl acetate (10% v/v) in CH_2Cl_2 to give **6.6** (1.04 g, 85%). **ESI MS** *m*/*z*: $[M + Na]^+$ 195.0, calcd 195.1. **^{1}H NMR** ($CDCl_3$, 300 MHz) δ 4.11 (q, J = 7 Hz, 2H), 3.48 (dd, J = 3 and 4 Hz, 1H), 3.03 (dd, J = 4 and 6 Hz, 1H), 3.01 (dd, J = 3 and 6 Hz, 1H), 2.66 (m, 4H), 1.25 (t, J = 7 Hz, 3H). **^{13}C NMR** ($CDCl_3$, 75 MHz) δ 206.3, 172.6, 60.9, 53.7, 46.3, 31.3, 27.5, 14.3.

6.5.5 *In Vitro* Assay of P4H Activity

The enzymatic activity of P4H was assayed as described previously.[140] Briefly, substrate (dansyl-Gly–Phe–Pro–Gly-OEt) was added to 1 mL of 50 mM Tris–HCl buffer, pH 7.8, containing P4H (0.43 μM), $FeSO_4$ (0.1 mM), BSA (1 mg/mL), catalase (0.1 mg/mL), ascorbate (2 mM), DTT (0.1 mM), and α-ketoglutarate (0.5 mM). Aliquots (30 μL) were removed at 60, 180, and 300 s, quenched by boiling for 30 s, and injected onto a Microsorb-MV C18 reversed-phase HPLC

column (4.6 × 250 mm, 100-Å pore size) from Varian (Palo Alto, CA). The column was eluted at 1 mL/min with aqueous acetonitrile (50% v/v), and the absorbance of the eluent was monitored at 337.5 nm. All assays were preformed in triplicate.

6.5.6 *In Vitro* Assay for Irreversible Inhibition of P4H

P4H (43 μM) was incubated for 2.5, 5, 10, or 15 min with $FeSO_4$ (2 mM) and epoxy ketone **6.1** in 50 mM Tris–HCl buffer, pH 7.8. BSA (1 mg/mL), catalase (0.1 mg/mL), ascorbate (2 mM), DTT (0.1 mM), and α-ketoglutarate (0.5 mM) were added to this reaction mixture. Substrate (dansyl-Gly–Phe–Pro–Gly-OEt) was added to initiate the reaction, and assays for P4H activity were performed as described above. All assays were preformed in triplicate. Data are reported as activity relative to control reaction mixtures that lacked epoxy ketone **6.1**. Half-times for inactivation at various inhibitor concentrations are determined from linear least-squares regression analysis of the log of the relative rates. Half-times for inactivation were plotted against the reciprocal of inhibitor concentration in a Kitz–Wilson plot[211] to determine the value of k_{inact}/K_i.

6.5.7 Protection of P4H by α-Ketoglutarate

Epoxy ketone **6.1** (356 mM), P4H (43 μM), and $FeSO_4$ (2 mM) were incubated for 5 min with α-ketoglutarate (0, 15, 30, or 50 mM) in 50 mM Tris–HCl buffer, pH 7.8. BSA (1 mg/mL), catalase (0.1 mg/mL), ascorbate (2 mM), DTT (0.1 mM), and substrate (dansyl-Gly–Phe–Pro–Gly-OEt)

(2.5 mM) were added to initiate the reaction. Aliquots were removed at 15, 60, 120, 180, 240, and 300 s, quenched by boiling for 30 s, and injected onto the C18 reversed-phase HPLC column described above. The column was eluted at 1 mL/min with aqueous acetonitrile (50% v/v), and the absorbance of the eluent was monitored at 337.5 nm. Reaction rates were compared to the reactions without inhibitor.

6.5.8 Immobilization of P4H

P4H (50 μg) was incubated for 30 min at 37 °C with $FeSO_4$ (0.1 mM), ascorbate (5 mM), and epoxy ketone **6.1** (13 mM) in 1 mL of PBS. Biocytin hydrazide was added (to 11 mM), and the reaction mixture was incubated for an additional 30 min at 37 °C. $NaBH_4$ (30 mg/mL in 10 mM NaOH) was added to the reaction mixture to reduce the acylhydrazone. The P4H complex was desalted using a NICK™ column from Amersham Pharmacia (Uppsala, Sweden) and incubated on Reacti-Bind™ streptavidin coated microtiter plates from Pierce Chemical (Rockford, IL) for 1 h at 37 °C. The wells were washed (4×) with blocking buffer (0.1% w/v BSA and 0.05% v/v Tween 20® in PBS). A mouse monoclonal antibody specific to the α-subunit of human P4H (ICN Biomedicals, Costa Mesa, CA) was added to the wells, and incubated for 1 h at 37 °C. The wells were washed again (4×) with blocking buffer. Anti-mouse IgG–peroxidase conjugate from Sigma Chemical (St. Louis, MO) was added to the wells, and the plate was incubated for 1 h at 37 °C. The wells were washed again (4×) with blocking buffer. *O*-Phenylenediamine solution (0.05 M citric acid/0.05 M sodium phosphate buffer, pH 5.0, containing 1 mg/mL *O*-

phenylenediamine and 0.003% v/v H_2O_2) was added to the wells, and catalysis of *O*-phenylenediamine oxidation by peroxidase was monitored at 490 nm after quenching the reaction with H_2SO_4 (2.5 M).

6.5.9 *In Vivo* Assay of Epoxy Ketone 6.1

The effect of epoxy ketone **6.1** on *dpy-18 Caenorhabditis elegans* was assayed by modification of a known procedure.[81] Here, **6.6** (esterified **6.1**) in DMSO was added to a plate containing solid growth medium (4 mL), which consisted of agar (17 g/L), Bacto-Peptone (2.5 g/L), NaCl (50 mM), cholesterol (5 mg/L), $CaCl_2$ (1 mM), and $MgSO_4$ (1 mM) in 1 mM potassium phosphate buffer, pH 6.[212] The inhibitor was allowed to diffuse throughout the agar for 24 h. Four L4 worms were then added to each plate. After sufficient time for the L4 worms to grow to adult stage and lay eggs (24–48 h), the adult worms were removed to simplify counting. After sufficient time for all live eggs to hatch into worms (24 h), live worms and dead embryos (typically, 40–100) were counted by using a light microscope.

6.6 Acknowledgments

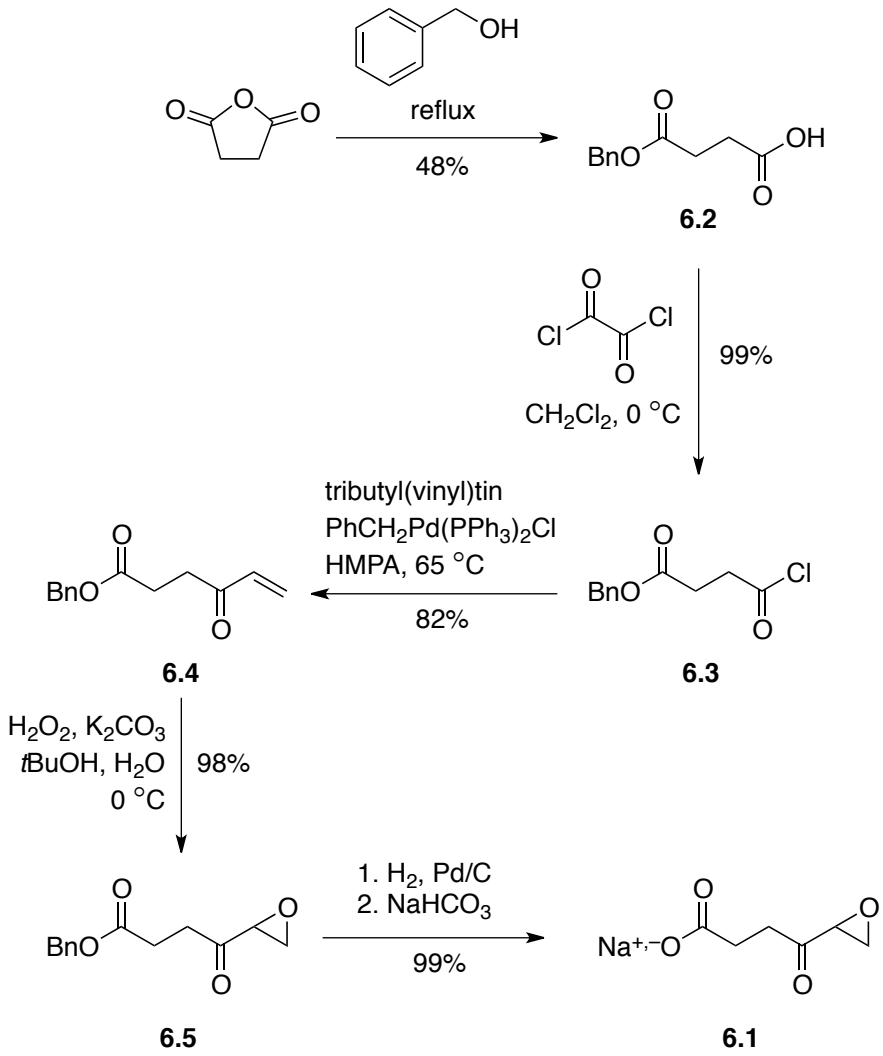
This paper is dedicated to Professor Ashraf Brik on the occasion of his winning the 2013 Tetrahedron Young Investigator Award in Bioorganic & Medicinal Chemistry. We are grateful to Dr. Lisa Friedman for her guidance in acquiring the data shown in Figure 6.3. This work was supported by Grant R01 AR044276 (NIH). J.D.V. was supported by Molecular Biosciences

Training Grant T32 GM007215 (NIH). E.A.K. was supported by Biotechnology Training Grant T32 GM083149 (NIH). The Micromass LCT® mass spectrometer was purchased with funds from Grant CHE-9974839 (NSF). This study made use of the National Magnetic Resonance Facility at Madison, which was supported by Grants P41 RR002301 and P41 GM066326 (NIH). Additional NMR equipment was purchased with funds from the University of Wisconsin, the NIH (RR 002781, RR 008438), the NSF (DMB-8415048, OIA-9977486, BIR-9214394), and the USDA.

6.7 Author Contributions

J.D.V., E.A.K, and J.J.H. contributed to the design and execution of experiments. All authors contributed to the analysis of data and preparation of the manuscript and figures.

Scheme 6.1 Route for the synthesis of epoxy ketone 6.1



Scheme 6.2 Route for the synthesis of 6.6 (esterified 6.1)

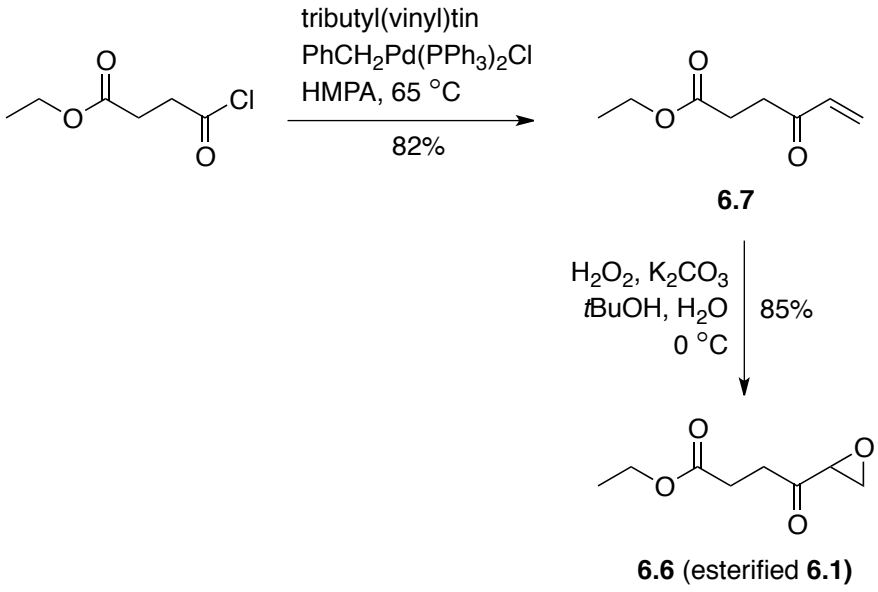


Figure 6.1 Prolyl hydroxylase reaction and active site binding modes

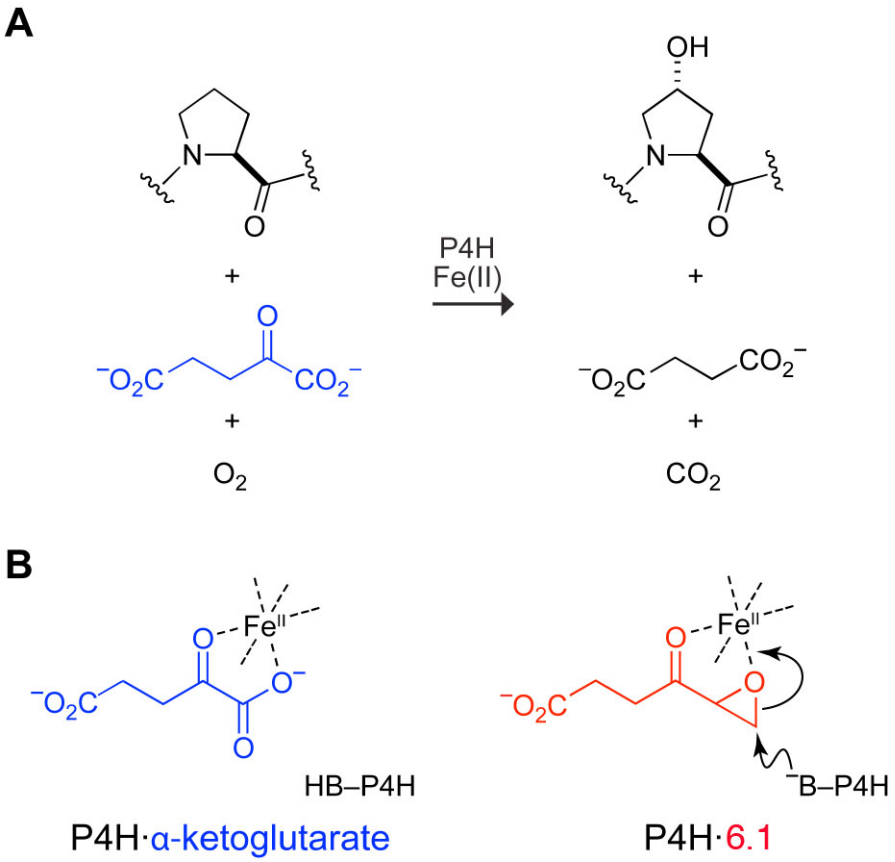


Figure 6.1 Prolyl hydroxylase reaction and active site binding modes (A) Reaction catalyzed by prolyl 4-hydroxylase. (B) Notional image of α-ketoglutarate (left) or epoxy ketone **6.1** (right) bound in the active site of P4H and interacting with an enzymic base, B. Alkylation is more likely to occur by nucleophilic attack at C6 than C5.[213]

Figure 6.2 *In vitro* assays of epoxy ketone 6.1

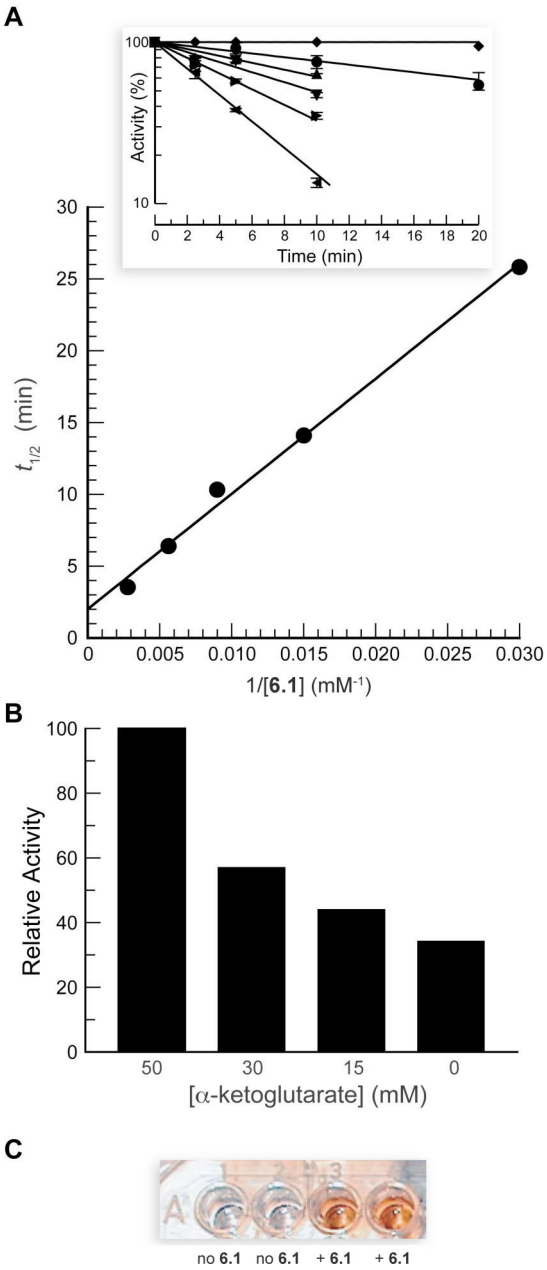


Figure 6.2. *In vitro* assays of epoxy ketone **6.1**. (A) Kitz–Wilson plot[211] showing the effect of different concentrations of epoxy ketone **6.1** on the catalytic activity of P4H. k_{inact}/K_i = 0.09 $M^{-1}s^{-1}$. Inset: Raw data. (B) Effect of α-ketoglutarate concentration on the inactivation of P4H by epoxy ketone **6.1**. (C) Immobilization of P4H after its reaction with epoxy ketone **6.1**. Wells 1 and 2, A_{490} = 0.14; wells 3 and 4, A_{490} = 0.98.

Figure 6.3 *In vivo* assay of epoxy ketone 6.1

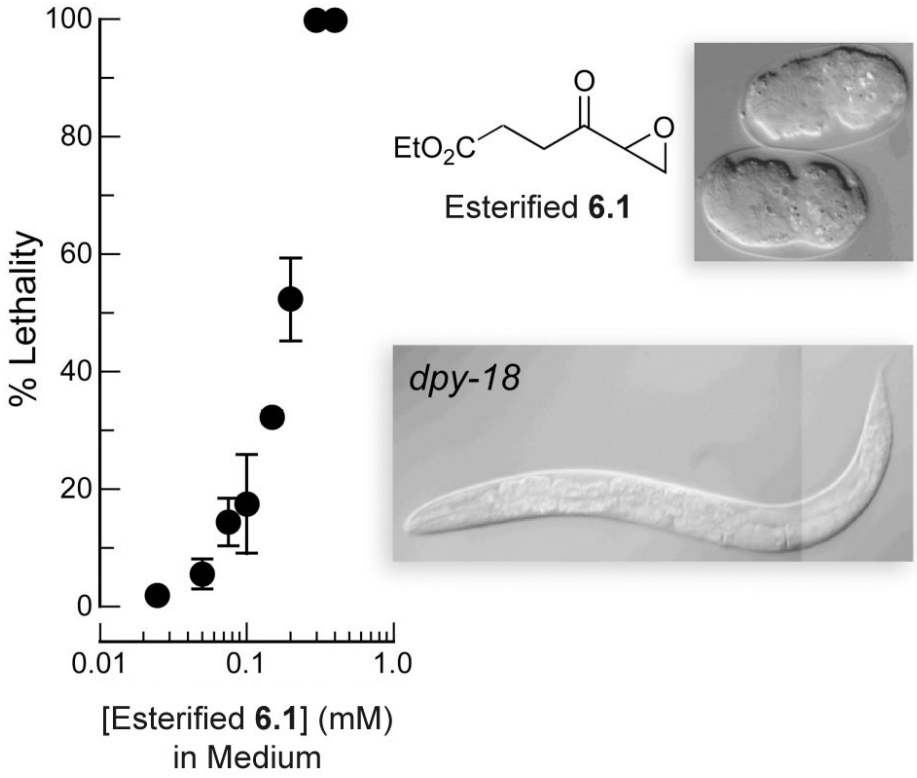


Figure 6.3 *In vivo* assay of epoxy ketone **6.1**. *Dpy-18 C. elegans* worms (four in stage L4) were exposed to known concentrations of esterified epoxy ketone **6.1** in the growth medium. Lethality refers to eggs that do not hatch. Insets: Image of lethality observed in the embryos of either a *dpy-18*/*phy-2* animal or a *dpy-18* animal exposed to esterified epoxy ketone **6.1** (top); image of an adult *dpy-18* animal not exposed to esterified epoxy ketone **6.1** (bottom).

CHAPTER 7

Future Directions

7.1 Biheteroaryldicarboxylate inhibitors of CP4Hs

In CHAPTERS 2 and 3, I presented data in support of biheteroaryldicarboxylates as potent and selective inhibitors of human CP4Hs. These compounds, including pythiDC, pyimDC, and even bipy55′DC, can be made cell-permeable via esterification (a common prodrug strategy), and appear to inhibit CP4H activity in human breast cancer cells. Moreover, they do so without disrupting iron metabolism or inhibiting the HIF-PHDs, which are significant shortcomings for previously used cell-active P4H inhibitors like EDHB. Thus, I advocate for the replacement of EDHB with these biheteroaryldicarboxylates and corresponding esters, with pythiDC arguably the best candidate for further development.

Investigations of biheteroaryl dicarboxylates *in vitro*

There are still many open questions regarding the inhibition of human CP4Hs by biheteroaryldicarboxylates. My data suggest that these inhibitors bind competitively in the AKG binding pocket and take advantage of interactions with the CP4H iron center. Still, the exact CP4H residues that interact with the inhibitors requires characterization, particularly with respect to the residues that interact with the second carboxyl group. Structural studies would be ideal for probing these active site residues, though the human CP4H holoenzymes have proven to be difficult to study by X-ray crystallography. As discussed in Chapter 5, the P4H from *C. reinhardtii* (Cr-P4H-1) is arguably the most closely related FAKGD enzyme available, and represents a reasonable model for the human CP4H1 catalytic domain. In addition, vP4H

(discussed in CHAPTER 1) has recently been characterized by X-ray crystallography,[93] and represents another model protein that could be used as a surrogate for the human CP4H1 holoenzyme in structural studies. Both of these P4H enzymes are significantly more soluble than is human CP4H and crystallize readily under a variety of conditions.[91-93] Moving forward, the biheteroaryldicarboxylate libraries developed in CHAPTERS 2 and 3 could be investigated as inhibitors of Cr-P4H-1 and vP4H. If the biheteroaryldicarboxylates serve as inhibitors of Cr-P4H-1 and/or vP4H with similar structure–activity relationships (SARs) compared to human CP4H1, structural studies could be initiated using these more simple P4Hs to investigate the binding modes of the biheteroaryldicarboxylates. If, however, the SARs are not comparable between the simple P4Hs and human CP4H1, such a finding would still provide information regarding the suitability of these simple P4Hs as models of the human enzyme.

In addition to structural studies of the interaction of the biheteroaryldicarboxylates with P4H enzymes, we also advocate for the continued exploration of these scaffolds using medicinal chemistry approaches (Figure 7.1). For example, various functional groups could be appended to many of the positions on both of the heteroaryl rings. These functional groups could likely include halogens and small alkyl groups initially, followed by polar substituents such as hydroxyl or amino groups, and then more complicated functional groups depending on the initial results. Due to the comparative ease of synthesis, it might be useful to probe the SARs of bipy45′DC and bipy55′DC analogues first, and then adapt any information that is learned from these scaffolds to the more complicated synthetic targets (*i.e.*, pythiDC and pyimDC analogues). Such medicinal chemistry studies could provide useful information about the binding of

biheteroaryldicarboxylates to human CP4H1, especially if used in conjunction with structural studies of the simple P4Hs above. Moreover, these studies could also provide additional analogues for further investigation in cells and preclinical models.

Investigations of biheteroaryldicarboxylates *in cellulo* and beyond

As noted above, I put forth pythiDC as a selective chemical probe for human CP4Hs, with diethyl pythiDC as a preferred cell-permeable analogue. Nonetheless, I suggest a continued investigation of this compound and related analogues in a cellular context. For example, investigations of diethyl pythiDC (as well as additional ester analogues) should be pursued both in other cancerous cell lines as well as in other collagen-producing cells such as fibroblasts. Such investigations would provide useful information about the therapeutic potential for pythiDC and related compounds with respect to other types of cancers and the fibrotic diseases. Then, once a more elaborate library of cell-active biheteroaryldicarboxylate compounds is assembled, investigations of these compounds in preclinical models of fibrosis and cancer metastasis should be pursued, with mouse models of breast cancer arguably the best models for initial investigations.

7.2 Bipyridinemonocarboxylate activators of human CP4H1

In CHAPTER 4, I presented data suggesting that bipyridinemonocarboxylates such as bipy4C and bipy5C serve as novel activators of human CP4H1. Moreover, I put forth a putative

mechanism of activation wherein bipy4C and bipy5C serve as ligands for the CP4H iron center and assist in the ascorbate-dependent reduction from an inactive iron oxidation state to a catalytically active Fe(II) state. I note that such a mechanism of activation has not been reported for any FAKGD enzyme. Still, the model I put forth needs to be tested further, especially with respect to the ligand environment and oxidation state of the iron center during the activation step. From EPR studies of the CP4H from chicken, it is thought that the uncoupled oxidation of AKG leads to a catalytically inactive Fe(III) species, and ascorbate is required as an almost specific reducing agent to return the iron center to the Fe(II) state.[101] Thus, based upon my postulated mechanism for bipy4C and bipy5C, these bipyridinemonocarboxylates should significantly enhance the rate of reduction of the CP4H metal center, and that rate enhancement should be observable.

Although such studies might be challenging, I argue that EPR spectroscopy would be an ideal method to study the kinetics of the ascorbate-dependent reduction of the CP4H iron center. Moreover, EPR spectroscopy has already been used to study the same process for the chicken CP4H,[101] further supporting the feasibility of such experiments. Beyond obtaining the necessary quantities of protein and establishing the method in-house or with a collaborator, the kinetic measurements of the reduction reaction will likely present the most significant challenge. As noted previously, ligands such as bipy and 1,10-phenanthroline are known to catalyze the ascorbate-dependent reduction of Fe(III) to Fe(II) in solution. My data suggests that bipy4C and bipy5C do not activate CP4H in this manner, but I do note that bipy and 1,10-phenanthroline can provide rate enhancements of at least an order of magnitude for reduction of free Fe(III).[101] Thus,

the study of the effect of bipy4C and bipy5C on the reduction of the CP4H iron center might require a stopped-flow set up in order to measure the rate enhancement accurately (if, in fact, the model is correct).

7.3 Conformational preferences of CP4Hs for the peptide substrate

In CHAPTER 5, I explored the interaction of human CP4H1 and Cr-P4H-1 with *cis* and *trans* alkene isosteres of the peptide substrate R-Gly-Pro-Pro-Gly-OEt (where R = Dansyl or Cbz). The isosteres mimicked the two possible conformations of the prolyl peptide bond preceding the Y-position of the collagen triplet repeat, and were designed originally to probe the conformational preferences with respect to hydroxylation of the Y-position Pro residue. Yet, neither the *cis* nor *trans* isostere served as a substrate for human CP4H1 or Cr-P4H-1. In contrast, I found that both the *cis* and *trans* isosteres served as inhibitors of these enzymes, with the *cis* isostere substantially more potent than the *trans* isostere, Thus, my data suggest that CP4H preferentially interacts with the *cis* conformation over the *trans* conformation in collagen-like polypeptides, though the exact conformation that is hydroxylated remains to be determined.

In CHAPTER 5, I also put forth a variety of possible explanations that could explain why the alkene isosteres are not substrates, but rather, serve as inhibitors of these P4Hs. These explanations include the possibility that a conformational change in the peptide substrate is required for catalysis, that the conformational preference could be recognized in a subsite that is not proximal to the active-site iron center, or that the chemical properties of the alkene isosteres

are not suitable for hydroxylation. With respect to the latter explanation, I suggest the use of computational and/or structural studies to investigate the detailed chemical nature of the alkene isosteres in order to determine whether or not they are comparable to the model substrate R-Gly-Pro-Pro-Gly-OEt. Special attention should be paid to the ring pucker of the Y-position Pro analogue, as it has been postulated that a C^{γ}-endo ring pucker in the Y-position Pro residue could be an important determinant of catalysis.[122] If it is determined that the alkene isosteres are sufficiently comparable to the model peptide substrate, such a result would suggest that either a conformational change in the substrate is required for catalysis, or that the conformational preference is recognized using an alternative (and, as of yet, unappreciated) subsite. Structural investigations aimed at determining a structure of CP4H in complex with the *cis* isostere would help to distinguish between these two possibilities. Whereas the structure of the catalytic domain of human CP4H1 has yet to be determined by any method, I note again that Cr-P4H-1 is likely a reasonable model for the human catalytic domain. Thus, as the alkene isosteres also served as inhibitors of this P4H, a crystal structure of Cr-P4H-1 (or perhaps that of vP4H) in complex with the *cis* isostere would likely provide a wealth of information with respect to the manner in which CP4Hs interact with this collagen mimetic peptide. Still, additional attention should be paid to the N- and C-terminal capping groups for the *cis* isostere, especially with respect to groups that could improve the solubility of these peptides for crystallographic purposes or more rigorous kinetic analyses.

7.4 Irreversible inhibitors of CP4Hs

In CHAPTER 6, I developed an electrophilic epoxy ketone analogue of AKG as an irreversible inhibitor of human CP4H. This compound replaced the C-1 carboxylate of AKG with an epoxide moiety, and inhibited human CP4H in a manner that was irreversible, time-dependent, and competitive with AKG. Moreover, this compound could be made bioavailable by esterification, and served as an inhibitor of CP4Hs *in vivo* in the nematode *Caenorhabditis elegans*. However, the potency of this epoxide analog of AKG was poor, likely because the epoxide oxygen is a poor ligand for the CP4H iron center compared a carboxylate oxygen. As noted above, the pyridine dicarboxylates 24PDC and 25PDC were both found to be modestly potent inhibitors of CP4Hs that bind to CP4H with a higher affinity compared to that of AKG. Thus, it is possible that 2-oxiranyl analogues of 24PDC and 25PDC (Figure 7.2B) will serve as more potent irreversible inhibitors than the epoxy ketone reported in CHAPTER 6. I begin to explore the preparation of such compounds in APPENDIX A.

Considering the finding that biheteroaryldicarboxylates are highly potent inhibitors of human CP4H1, I postulate that a second carboxylate could be incorporated into the epoxy ketone in order to improve binding affinity. I put forth the epoxy ketones and 2-oxiranylpyridines depicted in Figure 7.2C and Figure 7.2D as possible starting points for further investigations. All of these compounds could likely be synthesized via Heck coupling of an acrylate ester (Figure 7.2C) or 3-butenoate ester (Figure 7.2D) to an acyl chloride or 2-halopyridine, followed by conversion to the corresponding epoxide directly via oxidation, or in the case of the pyridine

analogs, via conversion to a vicinal bromohydrin and collapse to the epoxide under mild alkaline conditions, as discussed further in APPENDIX A.

7.5 Conclusions

In this thesis, I have attempted to explore various aspects of human CP4Hs that can be exploited for the selective inhibition of these enzymes, both *in vitro* and beyond. Many of the chemical tools developed here have provided useful information regarding that end goal, and could serve as starting points for the development of CP4H inhibitors with therapeutic potential for the fibrotic diseases and cancer metastasis. Even after more than 50 years of rigorous research, there are still many fascinating and unanswered questions regarding the structure and function of these essential enzymes. Certainly, the future of CP4H remains bright, perhaps even more so now than ever!

Figure 7.1 Elaboration of the biheteroaryldicarboxylate scaffolds

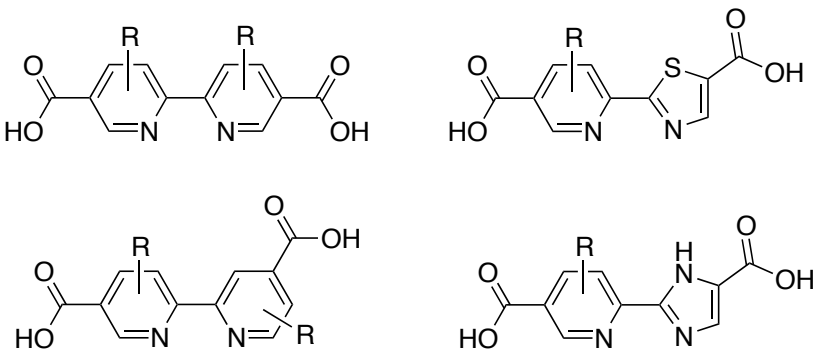


Figure 7.1 Elaboration of the biheteroaryldicarboxylate scaffolds. A medicinal chemistry approach can be pursued to elaborate the bipyDC, pythiDC, and pyimDC scaffolds in hopes of improving the potency and/or selectivity of these CP4H inhibitors. The R group could include halogens or small alkyl groups initially, or possibly polar substituents like hydroxyl or amino groups. Depending on the initial findings, more complicated functional groups could be tested.

Figure 7.2 Future starting points for the design of irreversible CP4H inhibitors

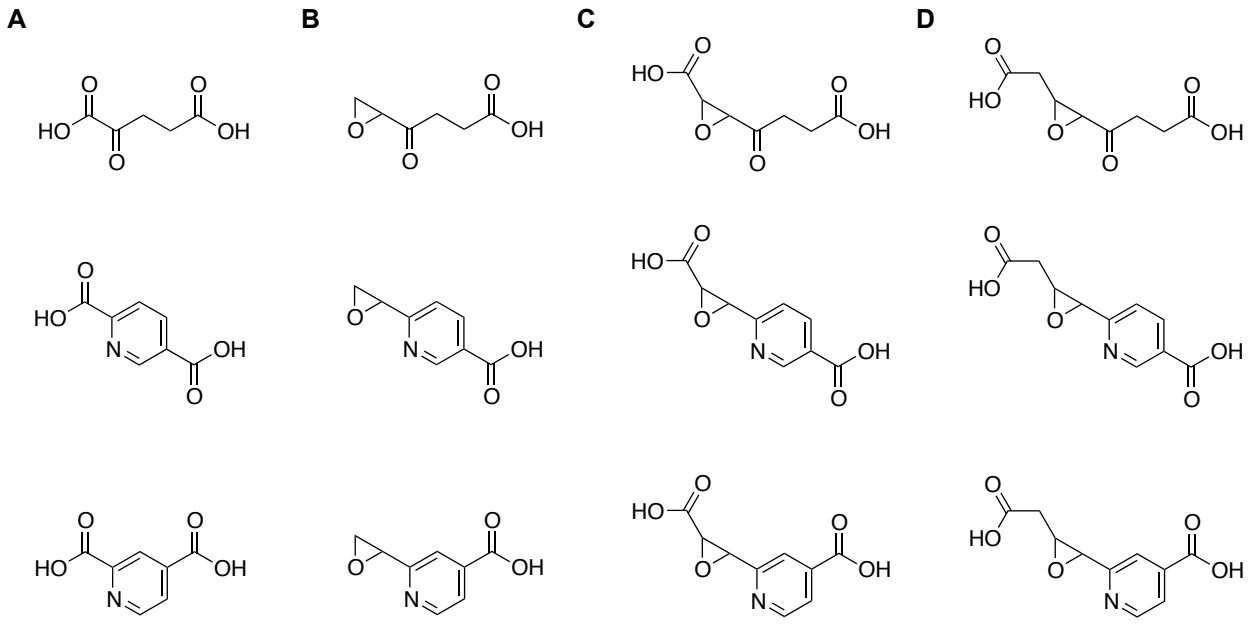


Figure 7.2 Future starting points for the design of irreversible CP4H inhibitors. (A) Renditions of the cosubstrate AKG and pyridine dicarboxylate analogues. (B) First generation epoxide inhibitors investigated in CHAPTER 6 and APPENDIX A. (C, D) Possible starting points for the design of more potent irreversible inhibitors that could take advantage of additional interactions with the second carboxyl group.

APPENDIX A

Synthesis of 2-Pyridylepoxides as Inhibitors of Collagen Prolyl 4-Hydroxylase*

*This chapter has been adapted from the senior honors thesis of Keng Hee (Spencer) Peh in pursuit of a bachelor of science degree in Biochemistry. K.H. Peh conducted the majority of experiments under the guidance of J.D. Vasta. Both K.H. Peh and J.D. Vasta contributed to the design of experiments and analysis of data. K.H. Peh prepared the manuscript and contributed with J.D. Vasta and R.T. Raines to the final editing of the manuscript.

A.1 Abstract

Collagen is an essential extracellular protein in animals. Its prevalent 4-hydroxyproline (Hyp) residues contribute significantly to its conformational stability. These Hyp residues arise from a post-translational modification catalyzed by collagen prolyl 4-hydroxylases (CP4Hs), which are members of the Fe(II)- and α-ketoglutarate (AKG)-dependent dioxygenase (FAKGD) family of enzymes. CP4Hs represent important therapeutic targets for the treatment of fibrotic and metastatic diseases. Recently, 4-oxo-5,6-epoxyhexanoate, an epoxide mimic of the cosubstrate AKG, was reported as an irreversible inhibitor and affinity label of a human CP4H. Despite its poor inhibitory potency, this epoxy ketone showed promise in proteomic experiments and a bioavailable ester analogue could be used to mediate collagen stability in the nematode *Caenorhabditis elegans*. 2,5-Pyridinedicarboxylate (25PDC) is competitive inhibitor with respect to the AKG cosubstate, and binds to CP4H with improved affinity compared to that of AKG. The goal of this work was to develop a synthetic route to the corresponding 2-epoxide analogue of 25PDC, which could serve as an irreversible inhibitor of CP4H with improved potency compared to 4-oxo-5,6-epoxyhexanoate.

A.2 Introduction

Collagen is an essential extracellular protein in mammals comprised of three α-chains intertwined to form a triple helix. Each chain is composed of triplet repeats of amino acid sequence Xaa-Yaa-Gly, where Xaa and Yaa are most often Proline (Pro) and (2*S*, 4*R*)-4-hydroxyproline (Hyp), respectively. The production of fully functional collagen is facilitated by the enzyme, collagen prolyl 4-hydroxylase (CP4H). CP4H catalyzes the post-translational hydroxylation of Pro residues on the Yaa position to form Hyp, which is necessary for the stability of the collagen helix.[1] Collagen plays a structural role in forming connective tissues such as tendons and ligaments, as well as bone and extracellular matrix.[1,2] It is also involved in the wound healing process in the body.[214]

Despite the importance of collagen, overproduction of this protein poses a potential threat to human health. Tissue scarring after surgery is a life-threatening situation caused by the excess accumulation of collagen at the wound site.[214,215] Also, the overproduction of collagen has been shown to contribute to the fibrotic diseases including fibrosarcomas, arteriosclerosis, scleroderma,[215,216] as well as cancer metastasis.[28,29,46,47,73] Studies in the medical field have shown that the therapeutic inhibition of CP4Hs is a potential strategy for the treatment of fibrotic diseases and scarring,[79,214] as well as metastatic disease,[29] although no viable therapeutics have been developed to date based upon this strategy.[67]

In mammals, CP4Hs exist as $\alpha_2\beta_2$ tetramers.[2,4] The α-subunit contains the catalytic domain for the hydroxylation of collagen strands and is a member of the Fe(II)- and α-

ketoglutarate (AKG)-dependent dioxygenase (FAKGD) family, a family with a variety of roles in biology.[56] In the active site, the co-substrate AKG chelates the Fe(II) using the C-1 carboxyl and C-2 keto groups, and a basic residue (usually Arg or Lys) is used to stabilize the C-5 carboxyl group. The β-subunit is protein disulfide isomerase (PDI), a multi-functional protein that acts as a chaperone to maintain the α-subunit in a soluble and active conformation. There are many different prolyl hydroxylases within the FAKGD family including the HIF prolyl hydroxylases, prolyl 3-hydroxylase, and transmembrane prolyl 4-hydroxylase, but CP4H is specific to its function.[4]

Previously, we reported a novel compound 4-oxo-5, 6-epoxyhexanoic acid (Epox456), and determined that this compound acted as an irreversible inhibitor of CP4H *in vitro*.[134] Epox456 was hypothesized to be analogous to the chemical structure of AKG (Figure A.1A) and thought to inhibit CP4H via the mechanism shown in Figure A.1B. Despite preliminary success, Epox456 was not found to be very potent, suggesting that improvements could be made to this inhibitor.[134] Studies of more potent CP4H inhibitors such as 2,4-pyridinedicarboxylate (24PDC) and 2,5-pyridinedicarboxylate (25PDC) suggest that they inhibit CP4H by mimicking AKG and binding in the CP4H active site (Figure A.1C), due to similar carbon chain length and positioning of a chelating moiety.[67] As 25PDC is known to bind to CP4H more tightly than AKG, we reasoned that an epoxide analogue of 25PDC would have improved potency towards CP4H compared to Epox456. Additionally, 25PDC has been shown to selectively inhibit human CP4H compared to other FAKGD enzymes, which is a desirable property of a potential therapeutic.[67] In this work, we sought to develop a synthetic route to 6-oxiranylnicotinic acid

(epoxide **A.1**), an electrophilic 25PDC analogue depicted in Figure A.1D.

A.3 Results and Discussion

Initially, we developed the synthetic route depicted in Scheme A.1 for the synthesis of epoxide **A.1**. A Steglich esterification of the commercially available 6-bromonicotinic acid with benzyl alcohol provided ester **A.2** in good yield (60%). Next, we used a Stille coupling of ester **A.2** with tributyl(vinyl)tin to form ester **A.3**. With further optimization of the workup to efficiently remove the organotin byproducts via precipitation, this reaction provided ester **A.3** in good yield (66%). Subsequent oxidation of ester **A.3** with MCPBA proceeded efficiently, providing *N*-oxide **A.4** in good yield (60%), but left us with the synthetic challenge of reducing the pyridine *N*-oxide and deprotecting the benzyl ester without perturbing the desired epoxide moiety. Initially, we envisioned a simultaneous deprotection of the benzyl ester and pyridine *N*-oxide using a palladium reduction with sodium borohydride as the reducing reagent. However, we found that under these conditions, the epoxide moiety was easily opened to form the primary alcohol, as evidenced by the characteristic triplets of the $–CH_2–CH_2–$ moiety in the ^1H NMR spectrum of the crude product (Figure A.2). Review of the synthetic literature found a previous report suggesting that aryl epoxides can be opened under palladium reduction conditions[217] (analogous to the reduction of benzyl esters), which confirmed our findings. As an alternative, we attempted reduce the pyridine *N*-oxide individually using PCl_3, but found that these conditions were also too harsh to preserve the target epoxide moiety (Figure A.3).

With the previously described difficulties in synthesizing epoxide **A.1** via Scheme A.1, we next sought to prepare epoxide **A.1** via an alternative synthetic route that did not require an oxidation step. We developed the route depicted in Scheme 2, which would collapse a vicinal bromohydrin under alkaline conditions to form the epoxide moiety, with simultaneous deprotection of the ethyl ester to afford epoxide **A.1**. Initially, we used a Stille coupling of the commercially available ethyl 6-chloronicotinate with tributyl(vinyl)tin to provide ethyl ester **A.5** in good yield (66%). Then, we reacted ethyl ester **A.5** with *N*,*N*-dibromo-*p*-toluenesulfonamide ($TsNBr_2$),[218] which afforded vicinal bromohydrin **A.6** in reasonable yield (40%). In the final step, we made use of mild alkaline conditions to both saponify the ethyl ester and collapse the bromohydrine to an epoxide. A ^1H NMR spectrum of the crude product suggested the presence of the target epoxide **A.1** (Figure A.4), although an isolation of epoxide **A.1** has not yet been optimized.

A.4 Conclusions and Future Directions

Consistent with our original goal, we have developed a synthetic route to epoxide **A.1**, an electrophilic analogue of 25PDC. In the short term, an isolation of epoxide **A.1** must be optimized such that the epoxide moiety is preserved. This isolation is not trivial, but could include a combination of recrystallization and/or reversed phase high performance liquid chromatography (HPLC). Once epoxide **A.1** can be isolated efficiently, the compound can be tested as an inhibitor of human CP4H using a previously established HPLC assay. If the overall

hypothesis is correct, epoxide **A.1** should serve as an irreversible inhibitor of human CP4H that is more potent than the previously reported Epox456. Moreover, epoxide **A.1** should be more selective for human CP4H compared to other members of the FAKGD family, as has been reported for the parent compound 25PDC. In the long term, the CP4H residue(s) that is(are) alkylated by these electrophilic AKG analogues should be identified, which will provide useful information about the active site residues of human CP4H.

A.5 Experimental Procedures Section

A.5.1 General Experimental Procedures

6-Bromonicotinic acid was from Combi-Blocks (San Diego, CA). Ethyl 6-chloronicotinate was from Sigma–Aldrich (St. Louis, MO). Tributyl(vinyl)tin was from Gelest, Inc. (Morrisville, PA). Palladium precatalysts and phosphine ligands were from either Sigma–Aldrich or Strem (Newberryport, MA), storred in a dessicator, and used without further purification. All other reagent chemicals were obtained from commercial sources (Sigma–Aldrich, Acros, Combi-Blocks, Oakwood Products, Enamine, etc.) and used without further purification. All glassware was flame- or oven-dried, and reactions were performed under N_2(g) unless indicated otherwise. DCM and toluene were dried over a column of alumina. Dimethylformamide was dried over alumina and further purified through an isocyanate scrubbing column. Other anhydrous solvents were obtained in septum-sealed bottles. Flash chromatography was performed with columns of 40-63 Å silica gel, 230–400 mesh (Silicycle, Québec City, Canada). Thin-layer chromatography (TLC) was performed on plates of EMD 250 μm silica 60-F_{254} with visualization by UV light or staining with $KMnO_4$. The phrase “concentrated under reduced pressure” refers to the removal of solvents and other volatile materials using a rotary evaporator at water aspirator pressure (<20 torr) while maintaining water-bath temperature below 40 °C. Residual solvent was removed from samples at high vacuum (<0.1 torr). The term “high vacuum” refers to vacuum achieved by a mechanical belt-drive oil pump. All reported yields are unoptimized.

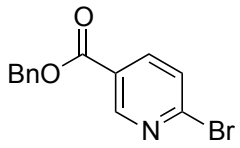
A.5.2 Instrumentation

NMR spectra were acquired at ambient temperature with a Bruker DMX-400 Avance spectrometer at the National Magnetic Resonance Facility at Madison (NMRFAM) and were referenced to TMS or a residual protic solvent. Electrospray ionization (ESI) mass spectrometry was performed with a Micromass LCT® instrument from Waters (Milford, MA) at the Mass Spectrometry Facility in the Department of Chemistry at the University of Wisconsin–Madison.

A.5.3 Synthetic Procedures

Two synthetic routes were attempted to prepare epoxide **A.1**. The first route (albeit unsuccessful) is depicted in Scheme A.1. The second and more successful synthetic route is depicted in Scheme A.2.

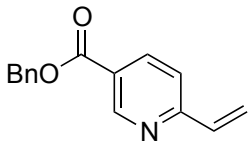
Benzyl 6-Bromonicotinate (A.2)



6-Bromonicotinic acid (5.0 g, 25 mmoles), 4-dimethylaminopyridine (300 mg, 2.5 mmoles), and *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (5.2 g, 27 mmoles) were added to a dried reaction flask. The flask was evacuated and purged with N_2(g), after which anhydrous DMF (25 mL) and benzyl alcohol (2.6 mL, 25 mmoles) were added via syringe, and the reaction was stirred at ambient temperature. The reaction was completed after 4 h as judged by thin layer

chromatography. Saturated aqueous ammonium chloride (20 mL) was added to the reaction and the resultant mixture was stirred for 15 min. The reaction mixture was extracted with ether (3 × 20 mL) and the combined organic extracts were washed with H_2O (2 × 20 mL) and by brine (20 mL). The organic layer was dried over Na_2SO_4(s) and concentrated under reduced pressure. The crude product was purified by chromatography on silica (30% EtOAc in hexanes) to afford the title compound (4.4 g, 60%) as a pale yellow crystalline solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ) 9.01 (d, J = 2.4 Hz, 1 H), 8.16 (dd, J = 2.4, 8.4 Hz, 1 H), 7.60, (d, J = 8.4 Hz , 1 H), 7.47–7.37 (m, 5 H), 5.41 (s, 2 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ) 164.4, 151.5, 146.9, 139.2, 135.2, 128.7, 128.4, 128.1, 125.5, 125.3, 67.4; **HRMS** (ESI) *m/z* 291.9962 [calc'd for $C_{13}H_{10}BrNO_2$ $(M + H)^+$ 291.9968].

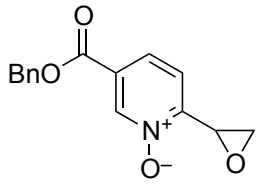
Benzyl 6-Vinylnicotinate (A.3)



Benzyl 6-bromonicotinate (**A.2**) (3 g, 10.3 mmoles) and *trans*-dichlorobis(triphenylphosphine) palladium (II) (360 mg, 0.48 mmoles) were added to a dried reaction flask. The flask was evacuated and purged with N_2(g), after which anhydrous DMF (18 mL) and tributyl(vinyl)tin (3.0 mL, 10.3 mmoles) were added and the reaction was stirred at 80 °C. The reaction was completed after 8 h as judged by thin layer chromatography. NaCl (500 mg) in H_2O (10 mL) was added to the reaction and the resulting mixture was stirred for 15 min. The reaction mixture was

extracted with ether (3 × 20 mL) and the combined organic extracts were washed with H_2O (2 × 20 mL) and brine (20 mL) and then concentration under reduced pressure. To the residue was added chilled saturated potassium fluoride in H_2O (20 mL) and chilled ether (15 mL) and the layers were allowed to partition. The organic layer was collected and the aqueous layer extracted with ether (2 × 15 mL). The combined organic extracts were concentrated under reduced pressure and the crude product was purified by chromatography on silica (30% EtOAc in hexanes) to afford the title compound (1.6 g, 66%) as yellow oil. **^{1}H NMR** (400 MHz, $CDCl_3$, δ) 9.23 (d, J = 1.6 Hz, 1 H), 8.29 (dd, J = 2.4, 8.4 Hz, 1 H), 7.48–7.36 (m, 6 H), 6.88 (dd, J = 10.8, 17.6 Hz, 1 H), 6.37 (d, J = 17.2 Hz, 1 H), 5.64 (d, J = 10.8 Hz, 1 H), 5.41 (s, 2 H); **HRMS** (ESI) *m/z* 239.0937 [calc'd for $C_{15}H_{13}NO_2$ $(M + H)^+$ 239.0941].

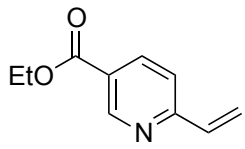
Benzyl 6-Oxiranylnicotinate-1-oxide (A.4)



Benzyl 6-vinylnicotinate (**A.3**) (1.0 g, 4.2 mmole) and *m*-chloroperoxybenzoic acid (2.8 g, 12.5 mmoles) were added to a dried reaction flask. The reaction flask was placed in an ice bath, evacuated, and purged with N_2(g). Anhydrous DCM (10 mL) was added slowly to the reaction mixture, after which it was stirred and allowed to warm to ambient temperature. The reaction was completed after 72 h as judged by thin layer chromatography. The solvent was removed under reduced pressure and the crude product was purified by chromatography on silica (50%

acetone in hexanes) to afford the title compound (610 mg, 60%) as yellow oil. **^{1}H NMR** (400 MHz, $CDCl_3$, δ) 8.86 (d, J = 1.2 Hz, 1 H), 7.88 (dd, J = 0.8, 8.0 Hz, 1 H), 7.44–7.34 (m, 6 H), 5.34 (s, 1 H), 4.56 (dd, J = 2.4, 4.0 Hz, 1 H), 3.35 (dd, J = 4.4, 5.6 Hz, 1 H), 2.74 (dd, J = 2.0, 5.6 Hz, 1 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ) 162.6, 152.5, 140.4, 134.7, 128.8 (2 signals), 128.5 (2 signals), 126.7, 121.8, 67.9, 50.4, 47.7; **HRMS** (ESI) *m/z* 272.0912 [calc'd for $C_{15}H_{13}NO_4$ $(M + H)^+$ 272.0918].

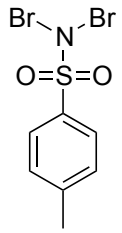
Ethyl 6-Vinylnicotinate (A.5)



Ethyl 6-chloronicotinate (5 g, 27 mmoles) and *trans*-dichlorobis(triphenylphosphine) palladium (II) (430 mg, 1.35 mmoles) were added to a dried reaction flask. The flask was evacuated and purged with N_2(g), after which anhydrous DMF (45 mL) and tributyl(vinyl)tin (7.9 mL, 27 mmoles) were added and the reaction was stirred at 80 °C. The reaction was completed after 5 h as judged by thin layer chromatography. NaCl (500 mg) in H_2O (50 mL) was added to reaction and the resulting mixture was stirred for 15 min. The reaction mixture was extracted with ether (3 × 25 mL) and the combined organic extracts were washed with H_2O (2 × 25 mL) and brine (25 mL) and concentrated under reduced pressure. Saturated potassium fluoride in MeOH (10 mL) and cold ether (15 mL) were added to the concentrated crude product, after which a white precipitate formed. The white precipitate was removed by filtration and the filtrate was

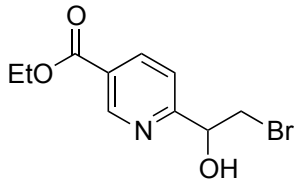
concentrated under reduced pressure. The precipitation procedure was repeated two more times, after which the combined organic extracts were concentrated under reduced pressure. The crude product was purified by chromatography on silica (30% EtOAc in hexanes) to afford the title compound (3.2 g, 66%) as an orange oil. **^{1}H NMR** (400 MHz, $CDCl_3$, δ) 9.18 (d, J = 1.6 Hz, 1 H), 8.26 (dd, J = 2.0, 8.0 Hz, 1 H), 7.41 (dd, J = 0.4, 8.0 Hz, 1 H), 6.88 (dd, J = 10.8, 17.6 Hz, 1 H), 6.35 (dd, J = 0.8, 17.2 Hz, 1 H), 5.63 (dd, J = 1.2, 10.8 Hz, 1 H), 4.42 (q, J = 6.8 Hz, 2 H), 1.42 (t, J = 7.2 Hz, 3 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ) 165.3, 159.1, 150.8, 137.6, 136.2, 124.8, 120.9, 120.6, 61.3, 12.3; **HRMS** (ESI) *m/z* 178.0863 [calc'd for $C_{10}H_{11}NO_2$ $(M + H)^+$ 178.0863].

***N,N*-Dibromo-*p*-toluenesulfonamide ($TsNBr_2$)**



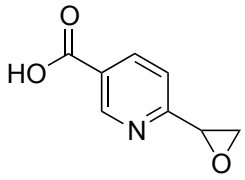
$TsNBr_2$ was prepared as described previously.[219] Chloramine-T trihydrate (5 g, 17.8 mmoles) and H_2O (100 mL) were added to a dried reaction flask. Bromine (1 mL, 19.5 mmoles) was added dropwise over two min while stirring. The reaction mixture was filtered under vacuum and the solid washed with deionized H_2O (3 × 10 mL). The resulting pale orange solid was dried under high vacuum and stored in a desiccator until use.

Ethyl 6-(2-Bromo-1-hydroxyethyl)nicotinate (A.6)



Ethyl 6-vinylnicotinate (**A.5**) (100 mg, 0.55 mmoles) was added to a dried reaction flask and dissolved in acetonitrile:H_2O (4:1, 3.5 mL). $TsNBr_2$, (210 mg, 0.65 mmoles) was added slowly, after which the reaction was stirred for 10 min. Aqueous sodium thiosulfate (10%, 2.5 mL) was added to the reaction and the resulting mixture was stirred for 15 min at ambient temperature. The mixture was extracted with ether (3 × 5 mL) and the combined organic extracts were washed with H_2O (2 × 5 mL) and brine (5 mL) and dried over Na_2SO_4(s). The crude product was concentrated under reduced pressure and purified by chromatography on silica (30% EtOAc in hexanes) to afford title compound (59 mg, 40%) as an orange solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ) 9.19 (d, J = 1.2 Hz, 1 H), 8.39 (dd, J = 2.0, 8.0 Hz, 1 H), 7.59 (d, J = 8.0 Hz, 1 H), 5.11 (t, J = 4.8 Hz, 1 H), 4.44 (q, J = 7.2 Hz, 2 H), 4.11 (bs, 1 H), 3.82 (dd, J = 4.8, 10.4 Hz, 1 H), 3.74 (dd, J = 5.6, 10.4 Hz, 1 H), 1.43 (t, J = 7.2 Hz, 3 H); **HRMS** (ESI) *m/z* 274.0073 [calc'd for $C_{10}H_{12}BrNO_3$ $(M + H)^+$ 274.0074]

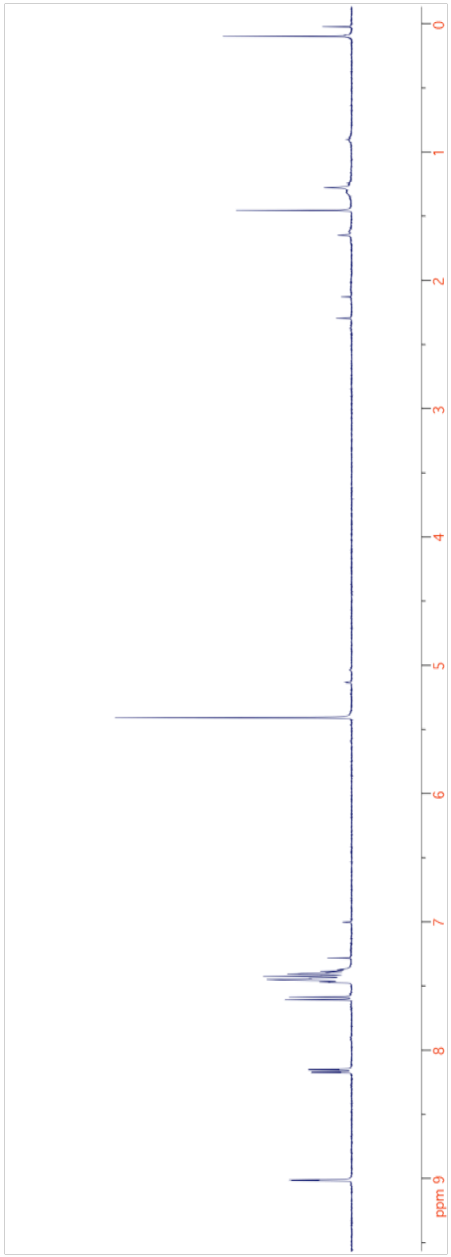
6-Oxiranylnicotinic Acid (A.1)



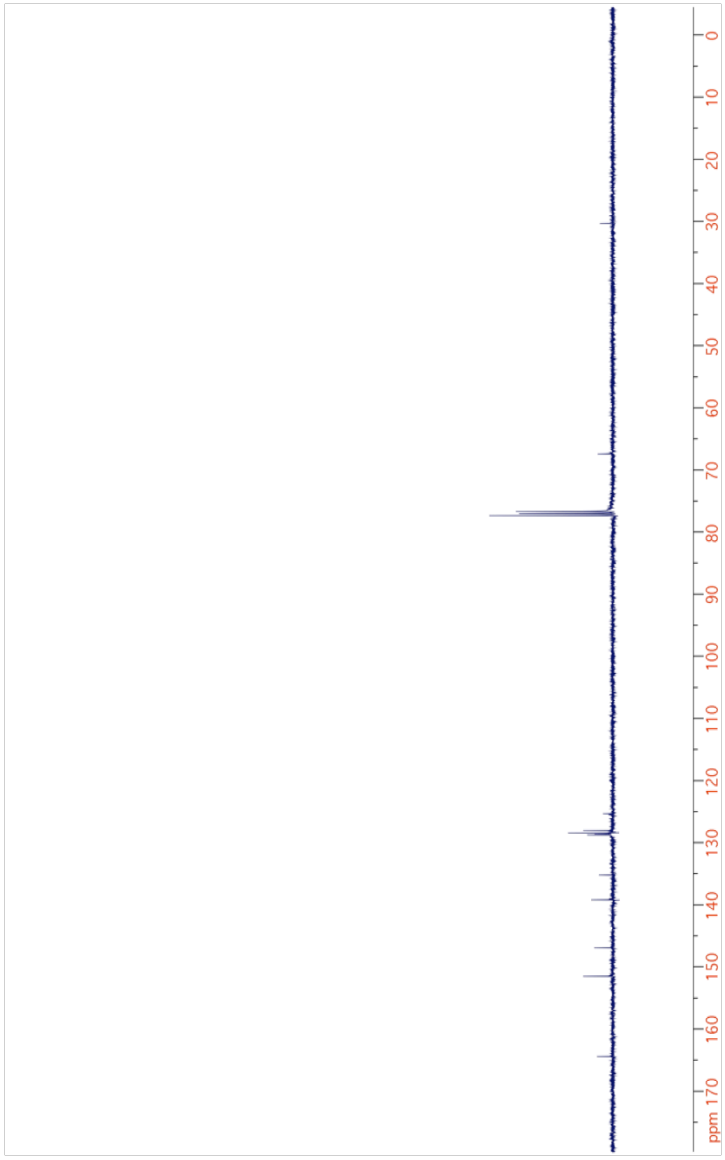
Ethyl 6-(2-bromo-1-hydroxyethyl)nicotinate (**A.6**) (50 mg, 0.2 mmoles) and potassium bicarbonate (76 mg, 0.6 mmoles) were added to dried reaction flask. The reaction flask was extracted and purged with N_2(g), after which MeOH (700 μL) was added and the reaction was stirred at room temperature. The reaction was completed after 3 h as judged by thin layer chromatography. The solvent was evaporated under reduced pressure to afford the crude product. A ^{1}H NMR (Figure A.4) of the crude product suggested the presence of the title compound, although the isolation has yet to be optimized.

A.5.4 NMR Spectra

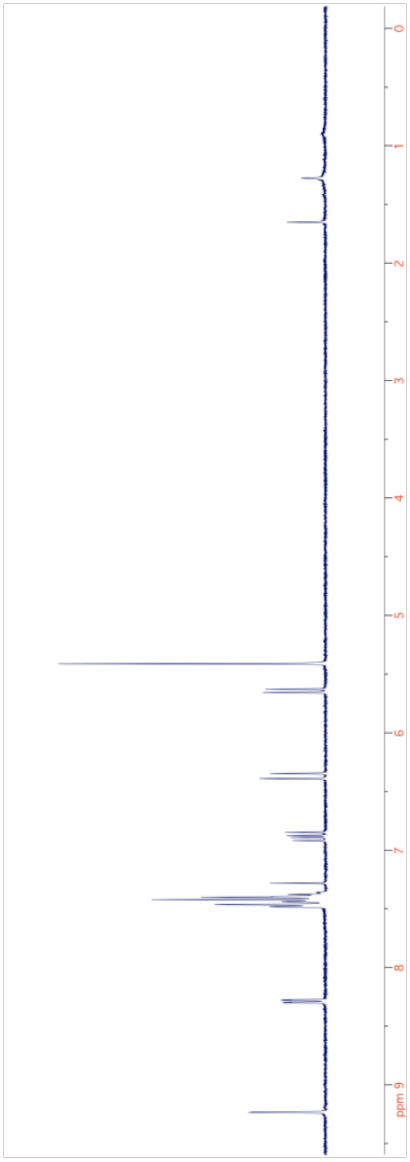
400 MHz ^{1}H NMR spectrum of Benzyl 6-Bromonicotinate in $CDCl_3$



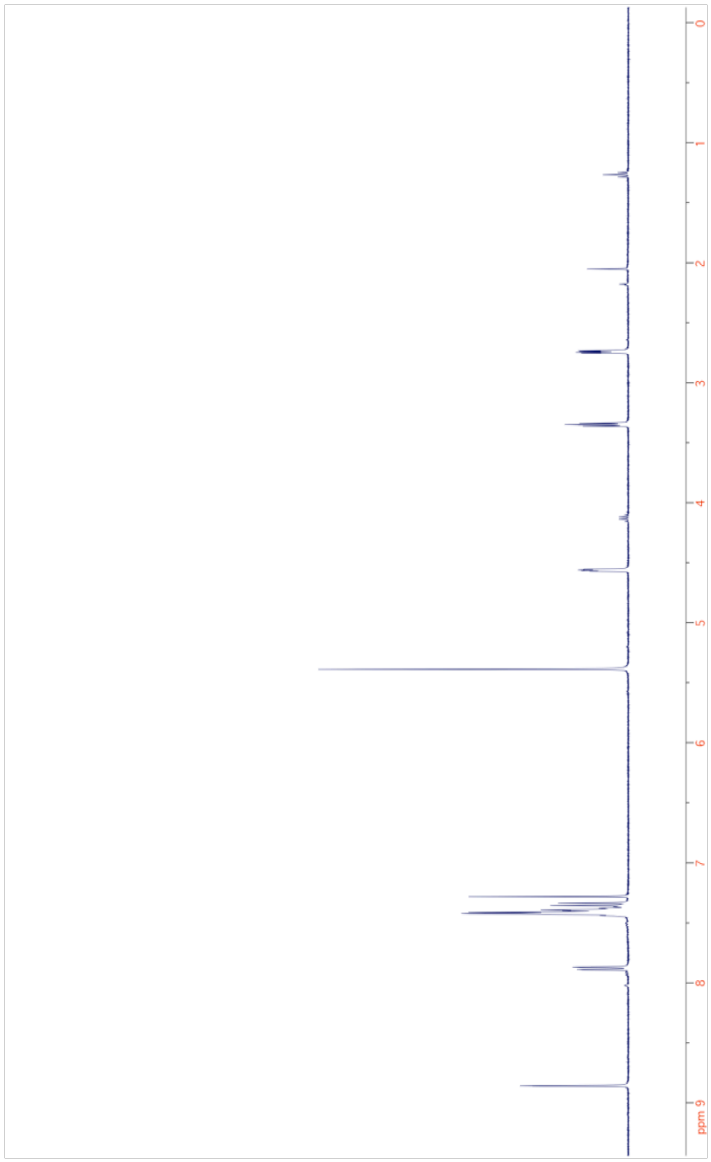
100 MHz ^{13}C NMR spectrum of Benzyl 6-Bromonicotinate in $CDCl_3$



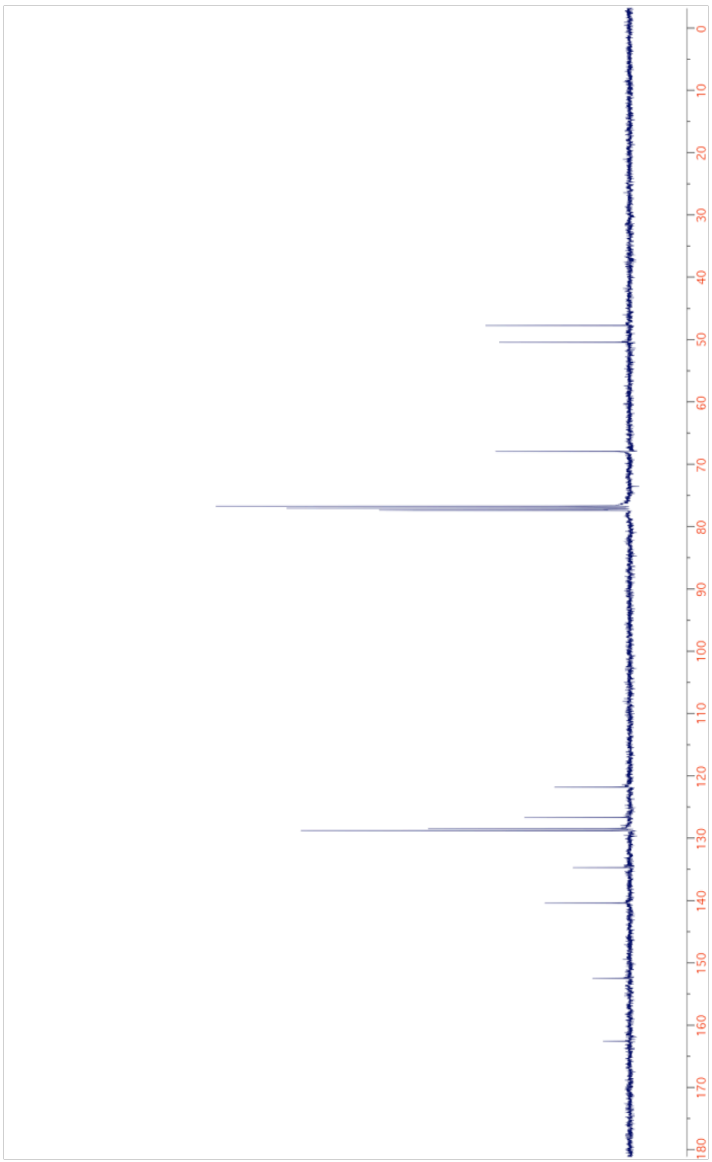
400 MHz ^{1}H NMR spectrum of Benzyl 6-Vinylnicotinate in $CDCl_3$



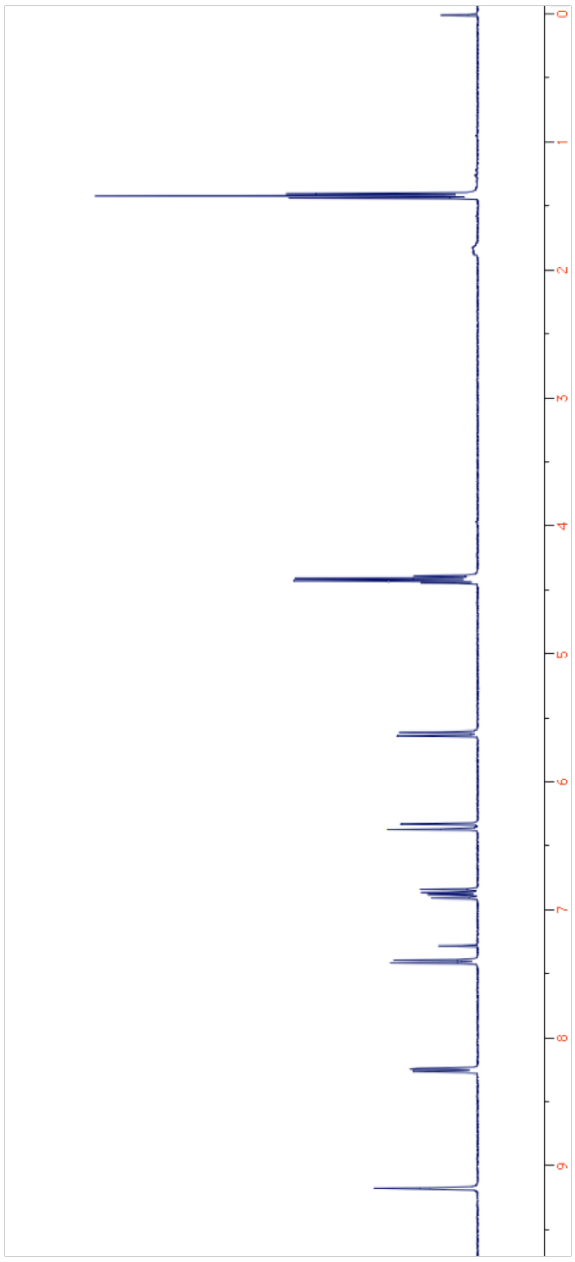
400 MHz ^{1}H NMR spectrum of Benzyl 6-Oxiranylnicotinate-1-oxide in $CDCl_3$



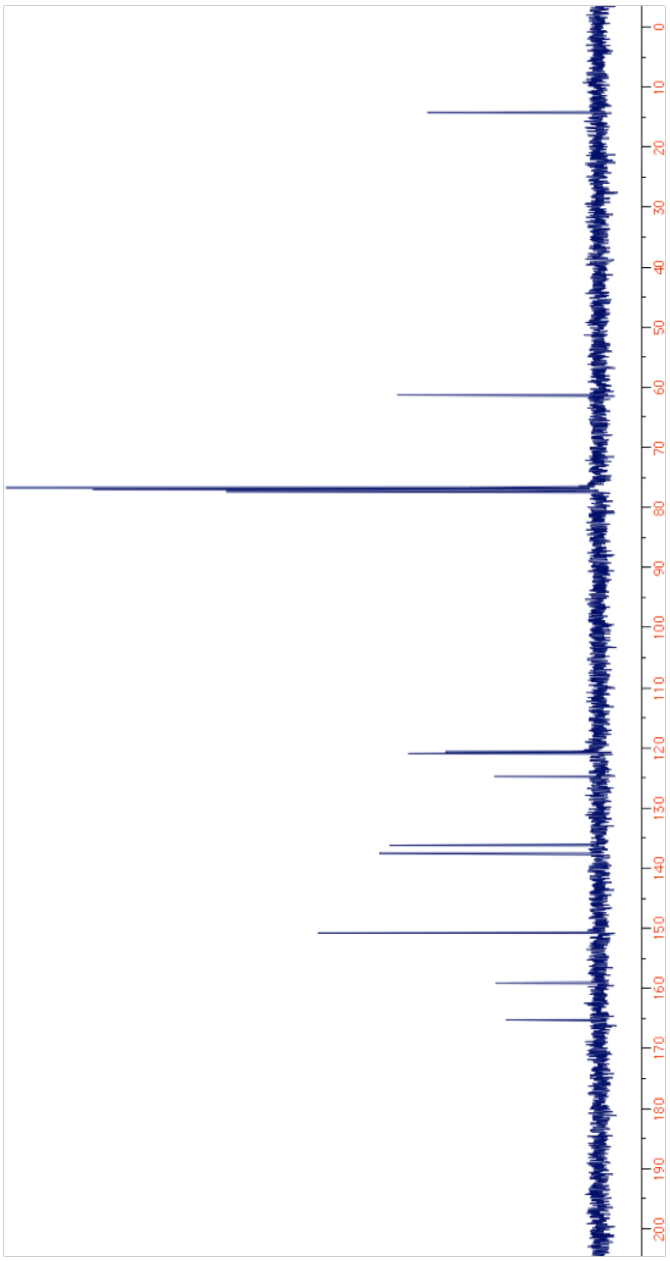
100 MHz ^{13}C NMR spectrum of Benzyl 6-Oxiranylnicotinate-1-oxide in $CDCl_3$



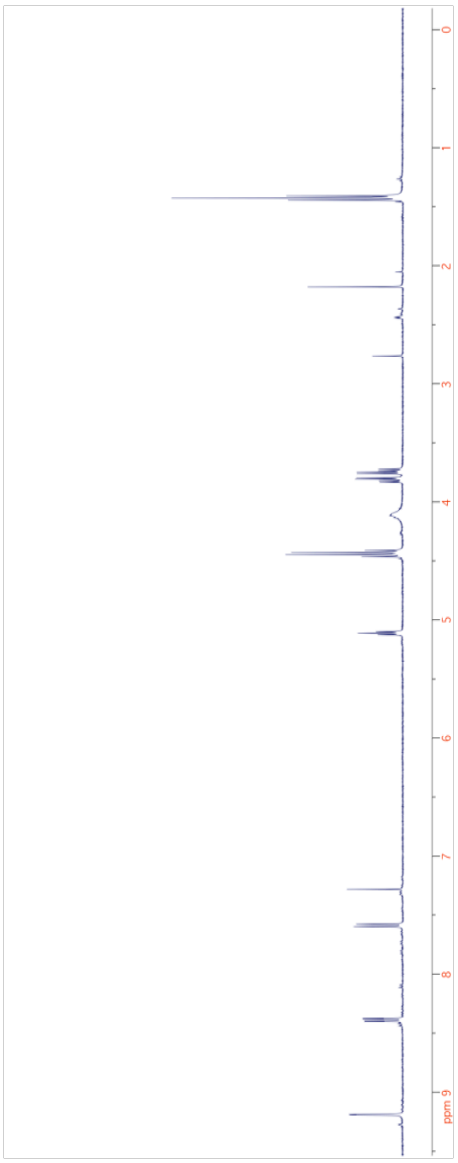
400 MHz ^{1}H NMR spectrum of Ethyl 6-Vinylnicotinate in $CDCl_3$



100 MHz ^{13}C NMR spectrum of Ethyl 6-Vinylnicotinate in $CDCl_3$



400 MHz ^{1}H NMR spectrum of Ethyl 6-(2-Bromo-1-hydroxyethyl)nicotinate in $CDCl_3$



A.6 Acknowledgments

This work was supported by Grant R01 AR044276 (NIH). J. D. V. was supported by Molecular Biosciences Training Grant T32 GM007215 (NIH) and a fellowship from the Department of Biochemistry at the University of Wisconsin–Madison. This study made use of the National Magnetic Resonance Facility at Madison, which is supported by Grant P41 GM103399 (NIH). The Micromass LCT® mass spectrometer was obtained with support from Grant CHE-9974839 (NSF).

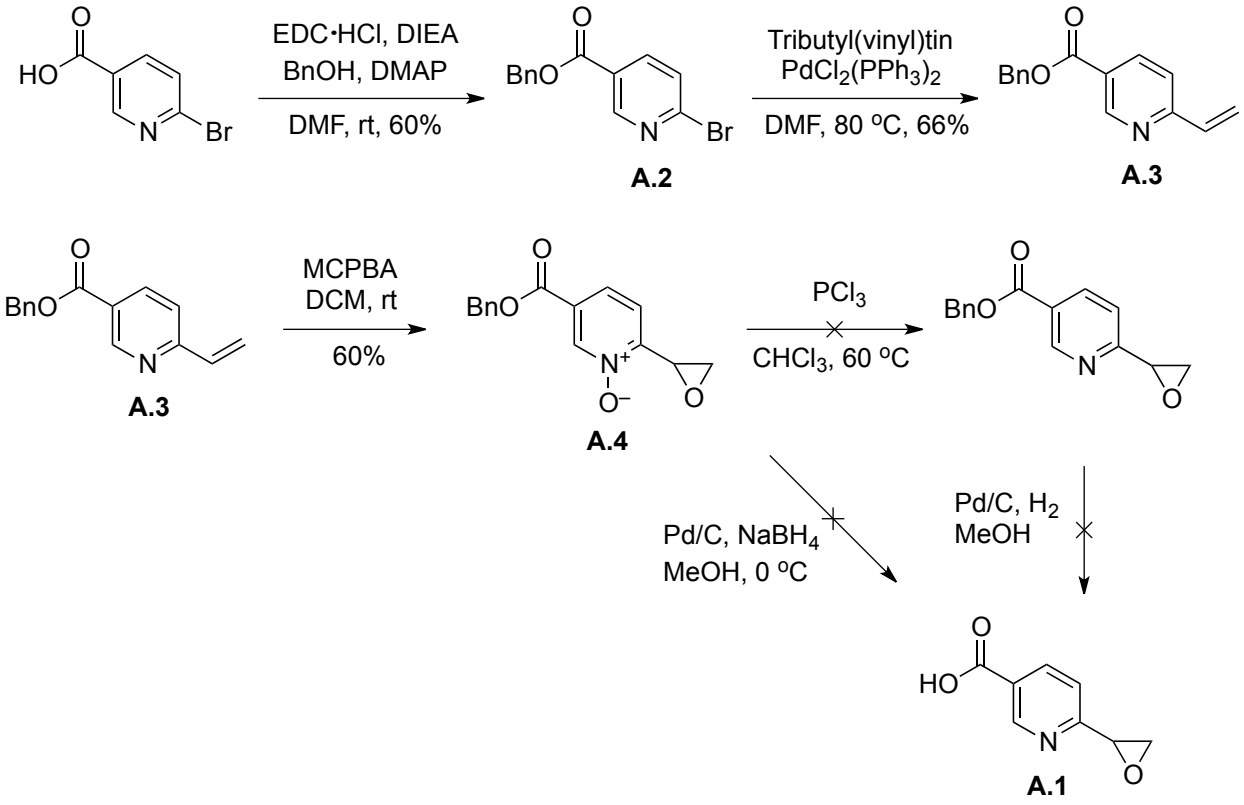
As Written by K. H. Peh

I would like to thank my mentor James Vasta for constantly giving advice and guiding me through my project. I would also like to thank other colleagues including Dr. K. Desai, T. Hoang, R. Presler, R. Windsor, K. Anderson, C. Eller, Dr. J. Lomax, and Professor R. T. Raines.

A.7 Author Contributions

K. H. Peh carried out the majority of synthetic procedures under the guidance of J. D. Vasta. Both K. H. Peh and J. D. Vasta contributed to the design of experiments. The manuscript was adapted from the senior honors thesis of K. H. Peh and edited here by J. D. Vasta and R. T. Raines.

Scheme A.1 Initial, albeit unsuccessful, synthetic route to epoxide A.1



Scheme A.2 Final synthetic route to epoxide A.1 via bromohydrin collapse

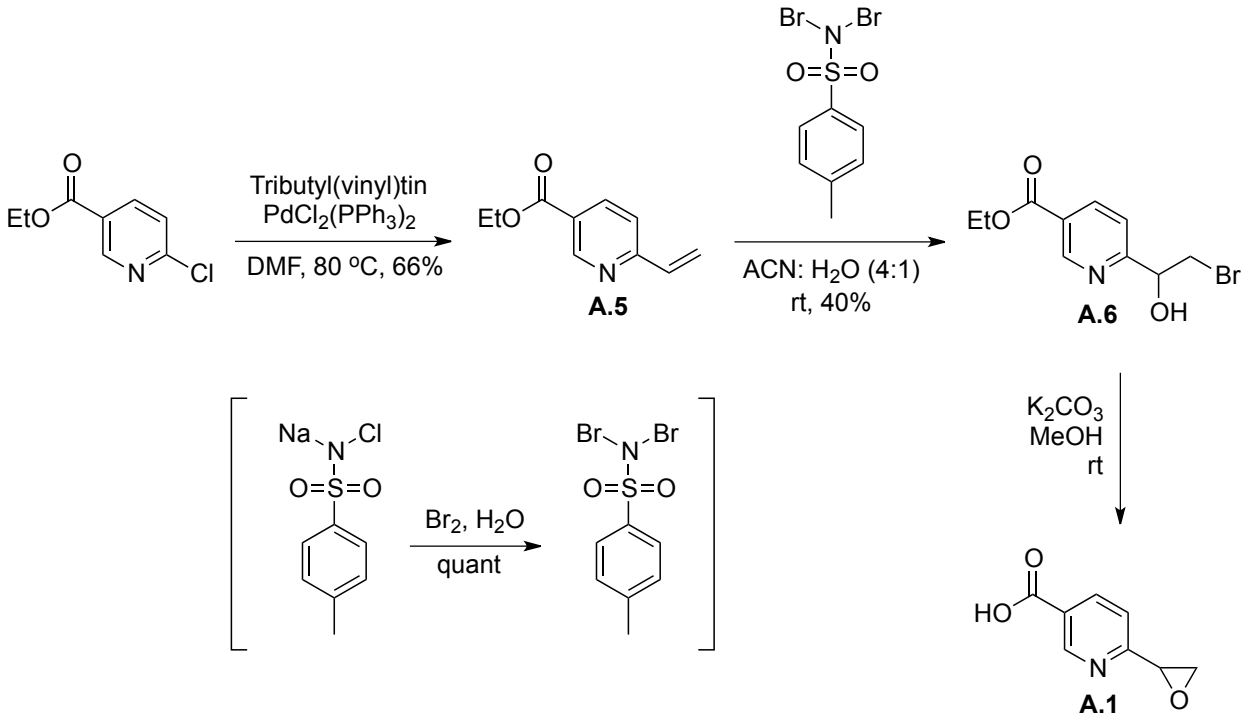


Figure A.1 Putative binding modes of AKG and analogs to the CP4H active site

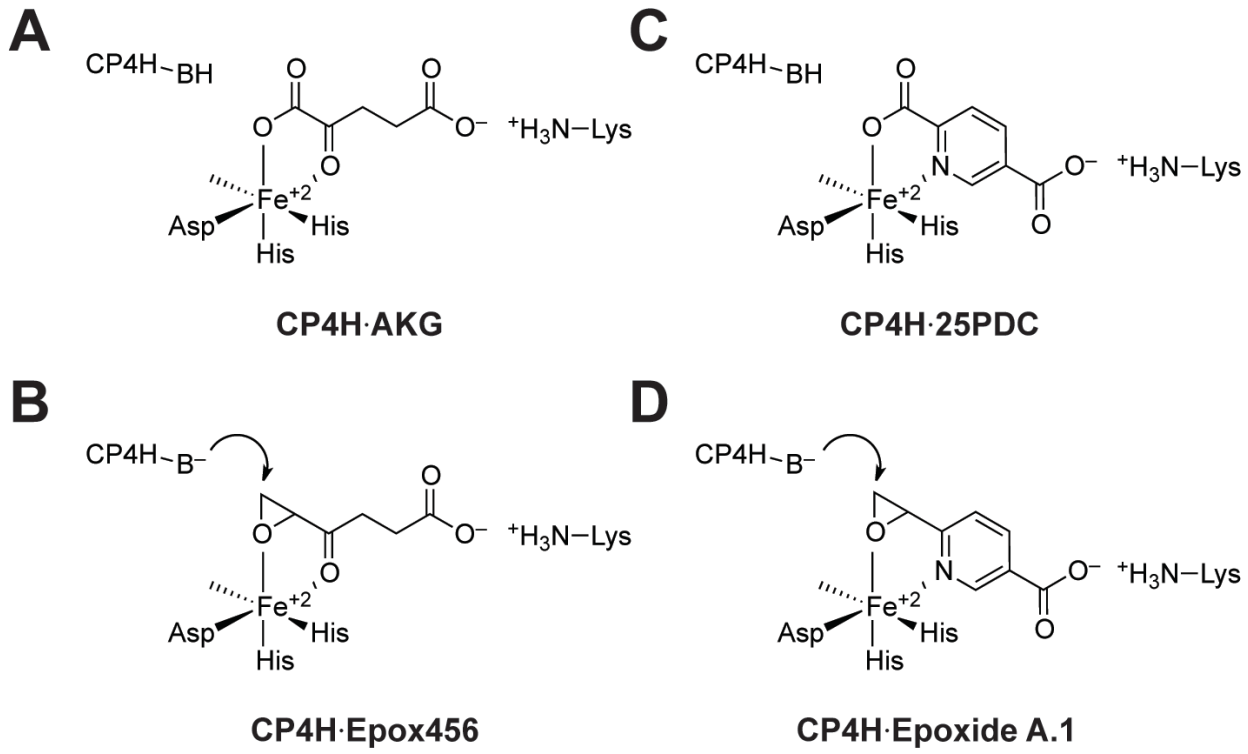


Figure A.1. Putative binding modes of AKG and analogs to the human CP4H activate site. (A) Schematic binding mode of AKG, the natural cosubstrate for the reaction catalyzed by human CP4Hs and related FAKGD enzymes. (B) Putative mechanism of the electrophilic AKG analog Epox456 as reported previously.[134] (C) Schematic binding mode of 25PDC, a competitive inhibitor of human CP4H. (D) Depiction and hypothesized inhibitory mechanism of epoxide **A.1**, an electrophilic analog of 25PDC.

Figure A.2 Crude ^{1}H NMR spectrum of the product obtained from the palladium reduction of *N*-Oxide A.4 using $NaBH_4$ as the reducing agent

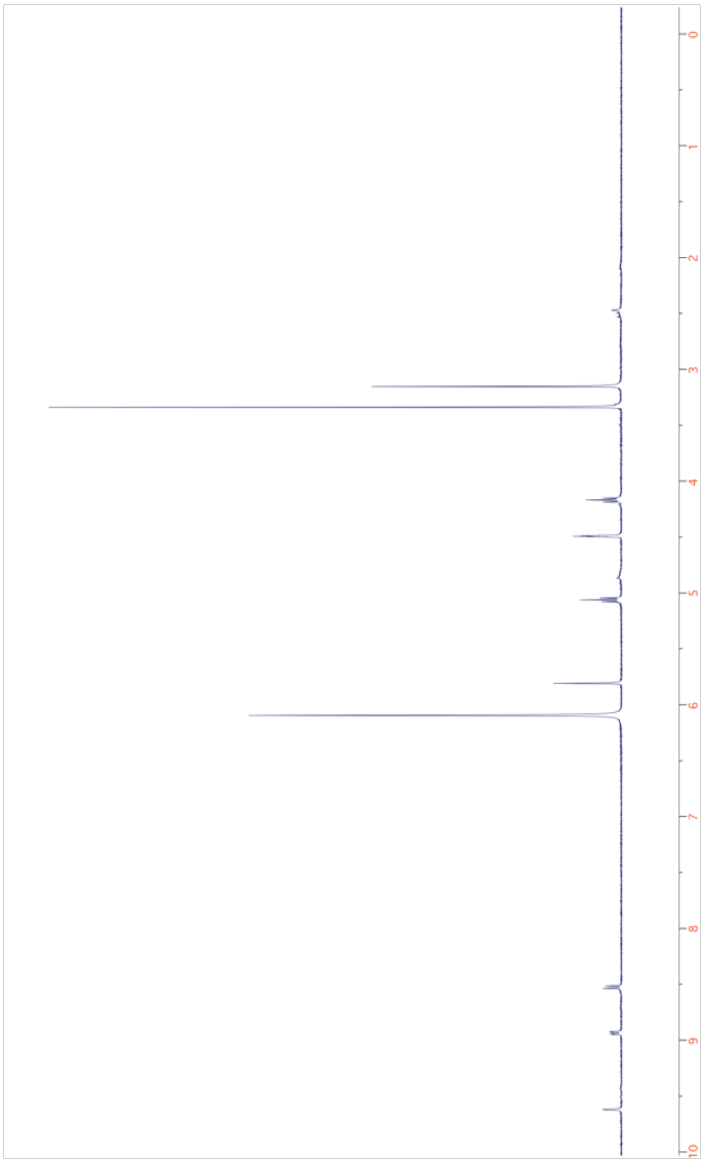


Figure A.2. Crude ^{1}H NMR spectrum of the product obtained from the palladium reduction of *N*-Oxide **A.4** using $NaBH_4$ as the reducing agent. The spectrum supports the opening of the epoxide ring to form the primary alcohol, as evidenced by the two triplets that are characteristic of the $–CH_2–CH_2–$ moiety.

Figure A.3 Crude ^{1}H NMR spectrum of the product obtained after the PCl_3 reduction of *N*-Oxide A.4

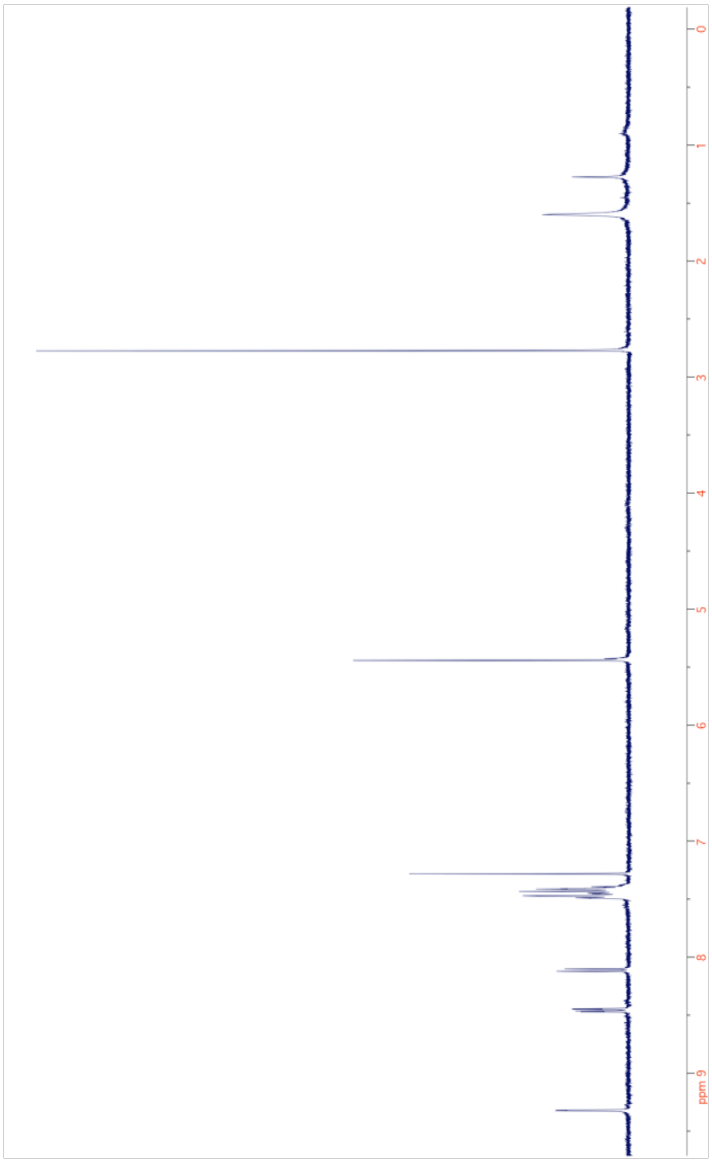


Figure A.3. Crude ^{1}H NMR spectrum of the product obtained after the PCl_3 reduction of *N*-Oxide **A.4**.

Figure A.4 Crude ^{1}H NMR of the product obtained after the alkaline hydrolysis of vicinal bromohydrin A.6

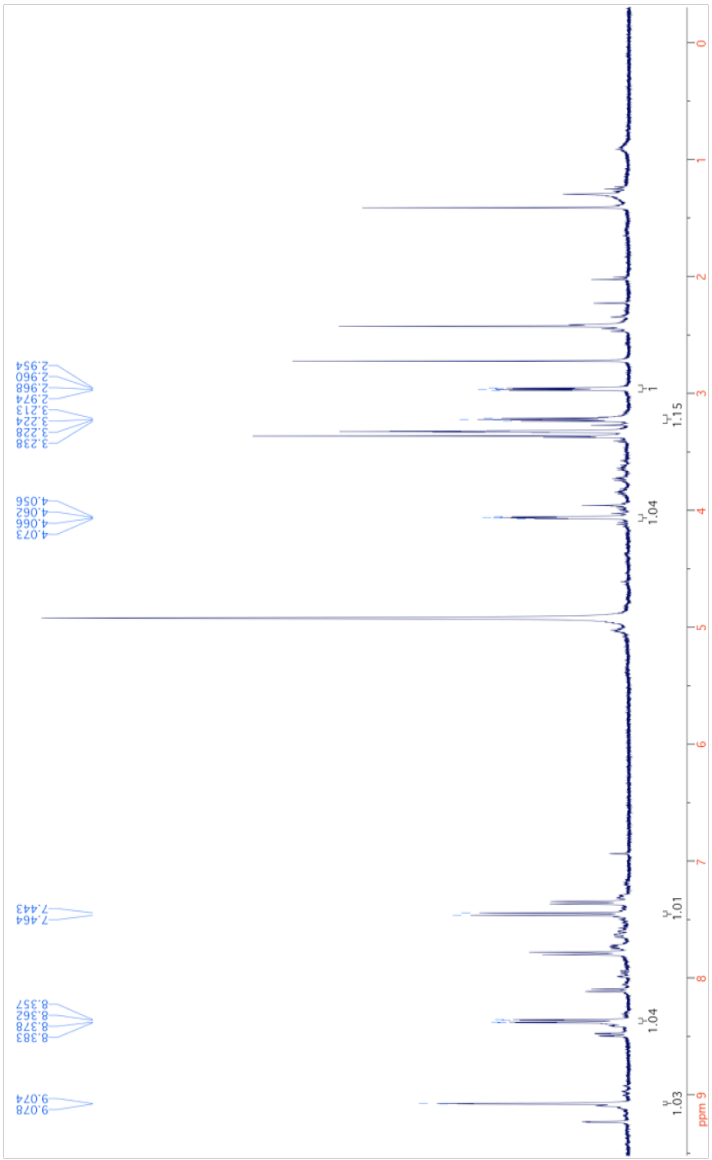


Figure A.4. Crude ^{1}H NMR of the product obtained after alkaline hydrolysis of vicinal bromohydrin **A.6**. Annotations correspond to the putative resonances of epoxide **A.1** and suggest the successful formation of this target compound.

APPENDIX B

IOX2 is Not a Potent Inhibitor of Human Collagen Prolyl 4-Hydroxylase 1*

*This chapter has been prepared for publication in part under the same title. Vasta, J. D., Andersen, K. A., and Raines, R. T. IOX2 is not a potent inhibitor of human collagen prolyl 4-hydroxylase 1. J. D. Vasta carried out or contributed to carrying out all experiments. K. A. Andersen contributed to carrying out all experiments relating to human cell culture. All authors contributed to the design of experiments, analysis of data, and preparation of the manuscript and figures.

B.1 Abstract

Hypoxia-Inducible Factor (HIF)-1 is an important, αβ dimeric transcription factor in animal biology, particularly for its role in the hypoxia signaling pathway. The level of its alpha subunit (HIF-1α), is regulated through an oxygen-dependent mechanism requiring the activity of HIF-prolyl-hydroxylases (HIF-PHDs), which are members of the Fe(II)- and α-ketoglutarate (AKG)-dependent dioxygenase (FAKGD) family of enzymes. Under normoxia, hydroxylation of HIF-1α by HIF-PHDs on two particular prolyl residues targets HIF-1α for ubiquitination and subsequent proteasomal degradation. Conversely, hypoxic conditions lower the activity of HIF-PHDs, which results in the significant accumulation of HIF-1α and initiation of the hypoxia signaling pathway. Upregulation of HIF-1 has been of increasing therapeutic interest, especially with respect to conditions such as anemia and ischemic/vascular disease, and inhibition of HIF-PHDs is of particular interest for these purposes. Recently, a series of 4-hydroxy-2-oxo-1,2-dihydroquinoline glycinamide analogues were reported as selective inhibitors of HIF-PHDs, with 2-(4-hydroxy-1-benzyl-2-oxo-1,2-dihydroquinoline-3-carboxamido)acetic acid (IOX2) anointed as a preferred functional probe. However, it was noted that the selectivity panel chosen in that report was incomplete, due to the large number of FAKGDs (~60) that are present throughout biology. Herein, we provide support confirming the function of IOX2 as a cell-active chemical probe for the HIF-PHDs, and demonstrate that this compound is not an inhibitor of human collagen prolyl 4-hydroxylase 1, another biologically important FAKGD.

B.2 Introduction

Hypoxia-Inducible Factor (HIF)-1 is an important, αβ dimeric transcription factor in animal biology, particularly for its role in the hypoxia signaling pathway.[53-55] Proper function of HIF-1 depends on both the beta subunit (HIF-1β), which is a ubiquitous nuclear protein involved in the transcription of target genes in a variety of signaling pathways, as well as the alpha subunit (HIF-1α), which provides the specificity for the hypoxic response. Under hypoxic conditions, HIF-1α is stabilized, resulting its accumulation and trafficking to the nucleus where it ultimately complexes with HIF-1β and other transcriptional co-activators to regulate the transcription of target genes involved in ameliorating the effects of hypoxia including erythropoietin (EPO), vascular endothelial growth factor (VEGF), and glucose transporter 1 (Glut1).[53]

The level of HIF-1α is regulated through an oxygen dependent mechanism requiring the activity of HIF-prolyl-hydroxylases (HIF-PHDs) and Factor Inhibiting HIF (FIH), all of which are members of the Fe(II)- and α-ketoglutarate (AKG)-dependent dioxygenase (FAKGD) family of enzymes.[56] FAKGDs couple the oxidative decarboxylation of the cosubstrate AKG to the hydroxylation of a hydrocarbon substrate via a radical rebound process.[56] Under normoxia, hydroxylation of HIF-1α by HIF-PHDs on two particular prolyl residues targets HIF-1α for ubiquitination and subsequent proteasomal degradation.[53] Conversely, hypoxic conditions lower the activity of HIF-PHDs, which results in the significant accumulation of HIF-1α and initiation of the hypoxia response as described above.

Upregulation of HIF-1 has been of increasing therapeutic interest, especially with respect to conditions such as anemia and ischemic/vascular disease,[30,67] and inhibition of HIF-PHDs are of particular interest for these purposes.[67,133] The development of clinically useful inhibitors for members of the FAKGD family of enzymes has proven difficult historically due to the challenge in achieving adequate selectivity, although significant progress has been made recently with the increasing number of crystal structures of FAKGDs that are now becoming available.[67] In fact, there are now PHD inhibitors in both early and late stage clinical trials (i.e. roxadustat, GSK-1278863, and AKB-6548) for the treatment of anemia, providing further evidence of this progress. Still, the selectivity profiles of these compounds are poorly characterized, which could be of significance with respect to their possible side effects.

Inhibition of the FAKGDs as a whole has been reviewed recently.[67] However, the vast majority of inhibitors with therapeutic potential are mimics of AKG, which are thought to inhibit by binding competitively in the AKG pocket and coordinating the enzymic iron center. While most simple AKG mimics such as *N*-oxalyl glycine and the pyridinedicarboxylates are of insufficient selectivity to be clinically useful, a variety of inhibitor scaffolds have begun to emerge with improved selectivity properties. For example 2,2′-bipyridine-4,5′-dicarboxylate and 2,2′-bipyridine-5,5′-dicarboxylate were identified as inhibitors of human collagen prolyl 4-hydroxylase (CP4H),[106,148] another biologically important FAKGD enzyme. The geometry of these compounds and the presence of a second carboxylate is believed to impart both enhanced potency compared to the simple AKG mimics, and also enzymic selectivity for CP4H compared to HIF-PHD isoform 2 (PHD2).[148] Unfortunately, these 2,2′-bipyridines (bipys) suffered from

selectivity issues related to the binding of free iron, but replacement of the bipy core with the weaker chelating biheteroaryl 2-(thiazol-2-yl)-pyridine (pythi) generated a potent and selective CP4H inhibitor that is active in human cells without perturbing iron metabolism (See CHAPTER 3). Moreover, Wang and coworkers recently reported the development of methylstat, a cell-active inhibitor and probe that is selective for the Jumonji C domain-containing proteins (JMJDs).[220] Methylstat was formed by fusing a primary substrate (methyllysine) mimic to an AKG mimic via a flexible linker,[220] which is another viable strategy that may be applicable to other FAKGDs of therapeutic interest.

Recently, a series of 4-hydroxy-2-oxo-1,2-dihydroquinoline glycinamide analogues were reported as selective inhibitors of the HIF-PHDs, with 2-(4-hydroxy-1-benzyl-2-oxo-1,2-dihydroquinoline-3-carboxamido)acetic acid (IOX2) anointed as a preferred functional probe.[133] IOX2 and related analogues were thought to serve as AKG mimics, using an intriguing 6-membered ring chelate to bind the active site iron center.[133] However, the authors noted that the selectivity panel chosen in that report was incomplete, due to the large number of FAKGDs (~60) that are present throughout biology. Herein, we report our investigations of IOX2. While IOX2 is a highly potent inhibitor of the HIF-PHDs, we find that it is not an inhibitor of a human CP4H *in vitro*. Moreover, we confirm that IOX2 inhibits the HIF-PHDs in cultured cells, and demonstrate that it does so without perturbing iron metabolism.

B.3 Results and Discussion

We were inspired by the potency and selectivity of IOX2 reported previously.[133] However, we noted that while the selectivity panel studied in the first report was thorough and included a variety of FAKGD enzymes such as the HIF-PHDs, FIH, and various histone demethylases (the JMJDs, Jumonji/ARID domain-containg protein 1C [JARID1C], and F-box and leucine-rich repeat protein 11 [FBKL11]), it lacked a few elements of interest. For example, the collagen hydroxylases, particularly CP4H isoform 1 (CP4H1), were not investigated. These hydroxylases represent important anti targets in the use of HIF-PHD inhibitors due to their contribution to a variety of diseases including scurvy.[30] Moreover, many inhibitors of FAKGDs rely on coordination to the active site iron center, and compounds possessing a high affinity for iron can often inhibit by sequestering free iron and disrupting iron metabolism.[131] Thus, we aimed to assess the inhibitory properties of IOX2 with respect to CP4H1 and to determine whether this probe could inhibit the HIF-PHDs in cells without disrupting iron metabolism.

First, we sought to determine if IOX2 was an inhibitor of the collagen prolyl hydroxylases, using CP4H1 (the most abundant isoform[30]) as a model protein. We synthesized IOX2 using a synthetic scheme that was adapted from the original report,[133] which provided IOX2 in 4 steps with an overall yield of 25%. To separate the possible inhibitory effect of iron sequestration rather than enzymic binding, we next screened IOX2 as an inhibitor of CP4H1 under assay conditions (10 μM compound and 50 μM $FeSO_4$) in which potent chelators do not cause inhibition.[148] Compared to the simple AKG mimic 2,5-pyridinedicarboxylate (25PDC), which inhibits CP4H1 with moderate potency, IOX2 did not cause any inhibition of CP4H1

under these assay conditions (Figure B.1). Compared to its reported IC_{50} for inhibition of PHD2 (0.022 μM)[133] this finding suggests that IOX2 is at least 450-fold selective for PHD2 compared to CP4H1, which is in general agreement for the selectivity of IOX2 for PHD2 compared to a panel of other FAKGDs (2–5,000 fold selective).[133]

Next, we sought to determine whether IOX2 can inhibit the HIF-PHDs in human cells without disrupting iron metabolism. To assess both iron metabolism and inhibition of HIF-PHDs, we performed immunoblots for ferritin, the transferrin receptor (TfR), and HIF-1α, all of which are known to give distinct phenotypes depending on the status of iron in a human cell.[131] More specifically, levels of ferritin and TfR are regulated by iron regulatory proteins 1 and 2 (IRPs), which have iron-dependent RNA-binding activity that modulates the expression of target genes at the level of translation.[155,156] Moreover, the stability of HIF-1α, which is dependent on the prolyl 4-hydroxylase activity of the HIF-PHDs, is inherently sensitive to the iron status of the cell. Thus, iron-deficient cells exhibit ferritin levels that are lower and TfR and HIF-1α levels that are higher than those of untreated cells, and inhibitors that are selective for the iron center of the HIF-PHDs compared to free iron should display only increases in HIF-1α levels.

Toward this end, we treated MDA-MB-231 breast cancer cells (a readily available human cell line in which the iron phenotypes are established for various compounds of interest) with IOX2 and a panel of controls including both iron chelators and P4H inhibitors (Figure B.2). Cells treated with the iron chelators deferroxamine (DFO) and bipy, as well as the "P4H" inhibitor ethyl 3,4-dihydroxybenzoate (EDHB), demonstrated a strong iron-deficient phenotype, whereas cells treated with a cell permeable diester prodrug of *N*-oxalyl glycine demonstrated only

increases in HIF-1α levels, consistent with the anticipated phenotypes described above (Figure B.2). Cells treated with IOX2 demonstrated significant increases in HIF-1α levels without changes in the level of ferritin or TfR (Figure B.2), which provides support that IOX2 inhibits the HIF-PHDs in cells without perturbing iron metabolism.

B.4 Conclusions

With its high potency and selectivity, IOX2 represents an important functional probe in the study of the HIF-PHDs and the hypoxia signaling pathway both *in vitro* and in cells.[133] This study has provided further support for the selectivity of IOX2 for the HIF-PHDs. The finding that IOX2 does not inhibit CP4H1 is significant considering the essentiality of CP4Hs in animal biology[4,81-83,202] and their importance in human disease.[30] Also, as iron is one of the most important metals in biology,[149] the finding that IOX2 does not perturb iron metabolism in cultured cells at concentrations required to inhibit the HIF-PHDs is notable and suggests that the 4-hydroxy-2-oxo-1,2-dihydroquinoline scaffold may be more generally useful as an inhibitor template for other enzymes of the FAKGD family. While evaluations of the overall selectivity of IOX2 are still incomplete, this study provides valuable selectivity information and further endorses IOX2 as a preferred functional probe for the HIF-PHDs.

B.5 Experimental Procedures Section

B.5.1 General Experimental Procedures

Isatoic anhydride was from VWR Scientific (Radnor, PA). 2,5-pyridinedicarboxylate was from Sigma-Aldrich (St. Louis, MO). All other reagent chemicals were obtained from commercial sources (Sigma–Aldrich, Acros, Combi-Blocks, Oakwood Products, Enamine, Bachem, or Novabiochem) and used without further purification. All glassware was flame- or oven-dried, and reactions were performed under N_2(g) unless indicated otherwise. DCM and toluene were dried over a column of alumina. Dimethylformamide was dried over alumina and further purified through an isocyanate scrubbing column. Other anhydrous solvents were obtained in septum-sealed bottles. Flash chromatography was performed with columns of 40-63 Å silica gel, 230–400 mesh (Silicycle, Québec City, Canada). Thin-layer chromatography (TLC) was performed on plates of EMD 250 μm silica 60-F_{254} with visualization by UV light or staining with $KMnO_4$. The phrase "concentrated under reduced pressure" refers to the removal of solvents and other volatile materials using a rotary evaporator at water aspirator pressure (<20 torr) while maintaining water-bath temperature below 40°C. Residual solvent was removed from samples at high vacuum (<0.1 torr). The term "high vacuum" refers to vacuum achieved by a mechanical belt-drive oil pump. All reported yields are unoptimized.

B.5.2 Instrumentation

NMR spectra were acquired at ambient temperature with a Bruker DMX-400 Avance

spectrometer or a Bruker Avance 500i spectrometer at the National Magnetic Resonance Facility at Madison (NMRFAM) and were referenced to TMS or a residual protic solvent. Electrospray ionization (ESI) mass spectrometry was performed with a Micromass LCT® instrument from Waters (Milford, MA) at the Mass Spectrometry Facility in the Department of Chemistry at the University of Wisconsin–Madison. The progress of reactions catalyzed by prolyl 4-hydroxylases was determined by analytical HPLC (Waters system equipped with a Waters 996 photodiode array detector, and Empower 2 software). Preparative HPLC was performed using a Prominence HPLC instrument from Shimadzu (Kyoto, Japan) equipped with two LC-20AP pumps, a SPD-M20A photodiode array detector, and a CTO-20A column oven. Protein concentrations were calculated from their absorbance at 280 nm as measured with a NanoVue Plus spectrophotometer from GE Healthcare using an extinction coefficient of 290,000 $M^{-1}cm^{-1}$ for human CP4H.[140]

B.5.3 Production of Recombinant Human CP4H1

Human CP4H containing the α(I) isoform was produced heterologously in Origami B(DE3) *Escherichia coli* cells and purified as described previously.[140]

B.5.4 Assay of Human CP4H1 Activity in the Presence of Inhibitors

The catalytic activity of human CP4H1 was assayed as described previously.[140] Briefly, activity assays were carried out at 30 °C in 100 μL Tris–HCl buffer (50 mM), pH 7.8, containing human CP4H1 (100 nM), inhibitor (0–10 μM), substrate (dansylGlyProProGlyOEt, 500 μM), $FeSO_4$ (50

μM), BSA (1 mg/mL), catalase (0.1 mg/mL), ascorbate (2 mM), DTT (100 μM), and α-ketoglutarate (100 μM). Reactions were pre-incubated with or without activator for 2 min at 30 °C, after which the reaction was initiated by the addition of α-ketoglutarate. After 15 min, reactions were quenched by boiling for 45 s and centrifuged at 10,000*g*. The supernatant (50 μL) was injected into a Nucleodur® C18 gravity reversed-phase column (4.6 × 250 mm, 5-μm particle size) from Macherey–Nagel (Bethlehem, PA). The column was eluted at 1 mL/min with a gradient (20 min) of 20%–45% aqueous acetonitrile containing 0.1% v/v TFA. The absorbance of the eluent was monitored at 289 nm. All assays were performed in triplicate. Data is reported as activity relative to control reactions lacking activator, where activity is determined from the percent conversion of substrate to product.

B.5.5 General Mammalian Cell Culture

The MDA-MB-231 cell line was a generous gift from Dr. Beth A. Weaver (University of Wisconsin–Madison). Cell lines were maintained according to the procedures recommended by the ATCC. Cells were grown in a cell culture incubator at 37 °C under CO_2 (5% v/v) in flat-bottomed culture flasks. The culture medium was DMEM supplemented with GIBCO fetal bovine serum (FBS) (10% v/v), penicillin (100 units/mL), streptomycin (100 μg/mL) and L-glutamine (2 mM). Cells were counted by hemocytometry with Trypan Blue prior to use in assays.

B.5.6 Effect of Iron Chelators or P4H Inhibitors on Iron Metabolism and Inhibition of HIF-PHDs in Human Cells

MDA-MB-231 cells were plated at 50,000 cells/well into 12-well plates and grown to confluence (~24 h) as described above. The cells were washed, and the medium was replaced with 1.0 mL of serum-free medium. Stock solutions of all test compounds were prepared at 100× in DMSO and added to a final concentration of 1× (0–1 mM). After addition of the test compound or DMSO vehicle, cells were incubated for 24 h. The cells were then washed with 100 μL MPER solution from Thermo Fisher Scientific (Waltham, MA) and collected. Protein concentrations were determined by BCA assay (Thermo Fisher Scientific), and samples corresponding to 40 μg total protein were analyzed with an immunoblot.

B.5.7 Immunoblotting

For all immunoblot analyses, protein samples where boiled in SDS–PAGE buffer, separated by electrophoresis through a Tris–glycine SDS–PAGE gel from Bio-Rad Laboratories (Hercules, CA), and subsequently transferred to a PVDF membrane for immunoblotting. All primary antibodies were used at the working dilution specified by the manufacturer. Primary antibodies were visualized using a complementary secondary antibody fused to horseradish peroxidase (HRP), with detection of HRP by chemiluminescence, imaging using an ImageQuant LAS 4000 (GE Healthcare), and quantitation by densitometry (ImageJ software). The anti-rabbit antibody from Promega and the anti-mouse antibody from Abbiotec (San Diego, CA) were used at the

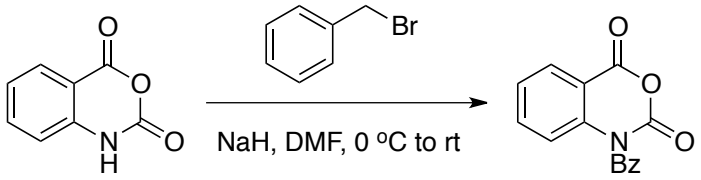
working dilution specified by the manufacturer. Statistical comparisons were performed using the one-way ANOVA function available in Prism.

For investigations of iron metabolism and HIF-PHD inhibition, protein samples (40 μg) were separated by SDS–PAGE through a 12% w/v acrylamide gel, and blots were probed with primary antibodies to human ferritin (rabbit monoclonal) from Abcam (Cambridge, MA), human transferrin receptor (mouse monoclonal) from Invitrogen, human HIF-1α (mouse monoclonal) from BD Biosciences, and human β-actin (rabbit monoclonal) from Cell Signaling Technology (Danvers, MA).

B.5.8 Synthetic Procedures

The synthetic route (Scheme B.1) to IOX2 and was adapted from the initial report of this compound as a selective inhibitor of the HIF-PHDs.[133]

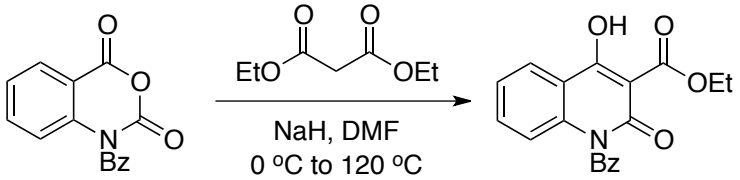
N-Benzylisatoic Anhydride



To an oven-dried flask was added NaH (60% w/w in mineral oil, 2.45 g, 61.3 mmoles), after which the flask was evacuated, purged with N_2(g), and cooled in an ice bath. Anhydrous DMF (20 mL) was added with continuous stirring, after which a solution of isatoic anhydride (10.0 g, 61.3 mmoles) in anhydrous DMF (20 mL) was added dropwise over 5 min with continued

stirring. The resultant mixture was allowed to stir on ice for 30 min, after which bubbling ceased. Benzyl bromide (7.28 mL, 61.3 mmol) was added dropwise with continuous stirring, after which the reaction was allowed to come to ambient temperature overnight. The reaction was quenched by pouring into ice cold H_2O (100 mL) with constant stirring, and the resulting mixture was extracted with DCM (4 × 100 mL). The combined organics were dried over Na_2SO_4(s) and concentrated under reduced pressure to afford a crude red solid. The solid was purified by recrystallization from toluene to afford the title compound (10.3 g, 67%) as a beige crystalline solid. **^{1}H NMR** (400 MHz, $CDCl_3$, δ): 8.19 (dd, J = 1.2, 8.0 Hz, 1 H), 7.66 (ddd, J = 1.2, 7.6, 8.4 Hz, 1 H), 7.40–7.27 (m, 6 H), 7.14 (d, J = 8.4 Hz, 1 H), 5.33 (s, 2 H); **^{13}C NMR** (100 MHz, $CDCl_3$, δ): 158.3, 148.5, 141.4, 137.2, 134.4, 130.9, 129.1, 128.1, 126.6, 124.2, 114.7, 111.8, 48.5; **HRMS** (ESI) *m/z* 271.1077 [calc'd for $C_{15}H_{15}N_2O_3$ $(M + NH_4)^+$ 271.1078].

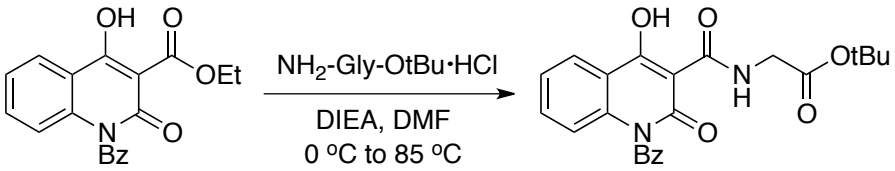
Ethyl 4-Hydroxy-1-benzyl-2-oxo-1,2-dihydroquinoline-3-carboxylate



To an oven-dried flask was added NaH (60% w/w in mineral oil, 1.58 g, 39.5 mmoles), after which the flask was evacuated, purged with N_2(g), and cooled in an ice bath. Anhydrous DMF (40 mL) was added with continuous stirring, after which diethyl malonate (6.00 mL, 39.5 mmoles) was added dropwise over 5 min. The resulting mixture was allowed to stir on ice for 15 min, after which bubbling ceased. *N*-Benzylisatoic anhydride (10.0 g, 39.5 mmoles) was added

in one portion and the temperature of the reaction mixture was raised to 120 °C slowly over 10 min. The reaction mixture was stirred at 120 °C overnight, after which resulting mixture was cooled to ambient temperature and concentrated under reduced pressure to a crude oil. To the oil was added H_2O (100 mL) and glacial acetic acid (2.25 mL), after which a precipitate was observed. The resulting mixture was extracted with EtOAc (3 × 100 mL), and the combined organic extracts were concentrated under reduced pressure to afford a crude solid. The solid was purified by recrystallization from MeOH to afford the title compound (7.85 g, 61%) as a pale orange crystalline solid. **^{1}H NMR** (500 MHz, $CDCl_3$, δ): 14.40 (s, 1 H), 8.22 (dd, J = 1.5, 8.0 Hz, 1 H), 7.56 (ddd J = 1.5, 7.5, 9.0 Hz, 1 H), 7.32–7.21 (m, 7 H), 5.54 (bs, 2 H), 4.98 (q, J = 7.0 Hz, 2 H), 1.51 (t, J = 7.0 Hz, 3 H); **^{13}C NMR** (125 MHz, $CDCl_3$, δ): 172.8, 172.1, 159.8, 140.8, 136.4, 134.4, 128.8, 127.2, 126.5, 125.8, 122.1, 115.2, 115.1, 97.8, 62.5, 45.7, 12.3; **HRMS** (ESI) *m/z* 321.1235 [calc'd for $C_{19}H_{18}NO_4$ $(M + H)^+$ 324.1231].

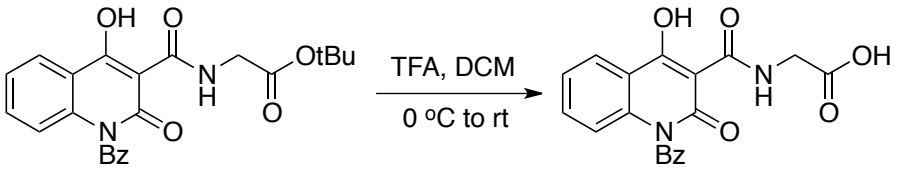
tert-Butyl 2-(4-Hydroxy-1-benzyl-2-oxo-1,2-dihydroquinoline-3-carboxamido)acetate (tBu IOX2)



To an oven dried flask was added ethyl 4-hydroxy-1-benzyl-2-oxo-1,2-dihydroquinoline-3-carboxylate (4.0 g, 12.4 mmoles) and glycine tert-butyl ester hydrochloride (2.49 g, 14.8 mmoles), after which the flask was evacuated and purged with N_2(g). DMF (50 mL) was added

to the flask, after which DIEA (6.46 mL, 37.1 mmoles) was added and the reaction was stirred at 85 °C. After 5 h, the reaction was cooled and concentrated under reduced pressure. Saturated aqueous NH_4Cl (100 mL) was added to the residue and the resulting mixture was extracted with EtOAc (3 × 100 mL). The combined organics were dried over Na_2SO_4(s) and concentrated under reduced pressure to a white solid. The solid was purified by chromatography on silica (20% EtOAc in hexanes) and recrystallized from EtOAc to afford the title compound (3.77 g, 75%) as a white crystalline solid. **^{1}H NMR** (500 MHz, $CDCl_3$, δ): 10.77 (t, J = 5.0 Hz, 1 H), 8.23 (dd, J = 1.5, 8.0 Hz, 1 H), 7.56 (ddd, J = 1.5, 7.0, 8.5 Hz, 1 H), 7.33–7.20 (m, 7 H), 5.56 (bs, 2 H), 4.16 (d, J = 5.5 Hz, 2 H), 1.52 (s, 9 H); **^{13}C NMR** (125 MHz, $CDCl_3$, δ): 172.1, 171.1, 168.3, 162.8, 139.6, 136.2, 133.9, 128.9, 127.4, 126.4, 125.6, 122.5, 116.3, 115.1, 96.7, 82.4, 45.7, 41.9, 28.1; **HRMS** (ESI) *m/z* 409.1752 [calc'd for $C_{23}H_{25}N_2O_5$ $(M + H)^+$ 409.1758].

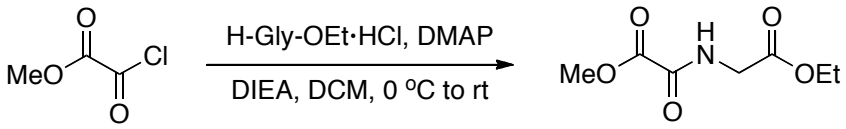
2-(4-Hydroxy-1-benzyl-2-oxo-1,2-dihydroquinoline-3-carboxamido)acetic Acid (IOX2)



To an oven dried flask was added tBu IOX2 (3.0 g, 7.3 mmoles), after which the flask was evacuated and purged with N_2(g). DCM (17 mL) was added and the flask was cooled in an ice bath with stirring. TFA (11.2 mL, 147 mmoles) was added dropwise on ice with stirring, after which the reaction was allowed to come to ambient temperature. The reaction was stirred for 3 h,

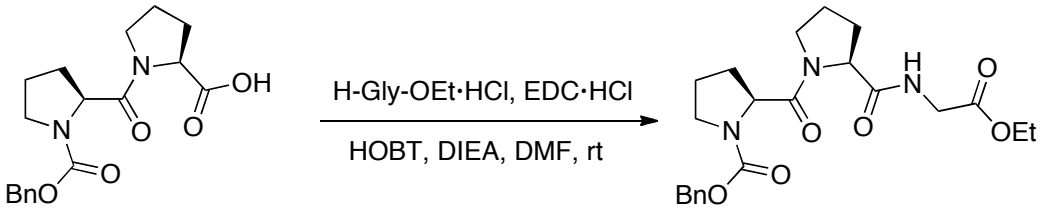
after which a white precipitate was observed. The resulting mixture was concentrated under reduced pressure to afford a white solid. The solid was recrystallized from EtOH to afford the title compound (2.1 g, 82%) as a fluffy white solid. **^{1}H NMR** (500 MHz, $CDCl_3$, δ): 12.99 (bs, 1 H), 10.54 (t, J = 5.5 Hz, 1 H), 8.14 (d, J = 9.0 Hz, 1 H), 7.72 (t, J = 7.5 Hz, 1 H), 7.48 (d, J = 8.5 Hz, 1 H), 7.37–7.22 (m, 6 H), 5.57 (bs, 2 H), 4.16 (d, J = 5.5 Hz, 2 H); **^{13}C NMR** (125 MHz, $CDCl_3$, δ): 172.5, 171.6, 171.4, 162.7, 140.2, 137.6, 135.4, 129.7, 128.1, 127.4, 125.8, 123.7, 116.8, 116.3, 96.9, 45.6, 42.0; **HRMS** (ESI) *m/z* 353.1139 [calc'd for $C_{19}H_{17}N_2O_5$ $(M + H)^+$ 353.1132].

***N*-(Methoxyoxalyl)glycine Ethyl Ester**



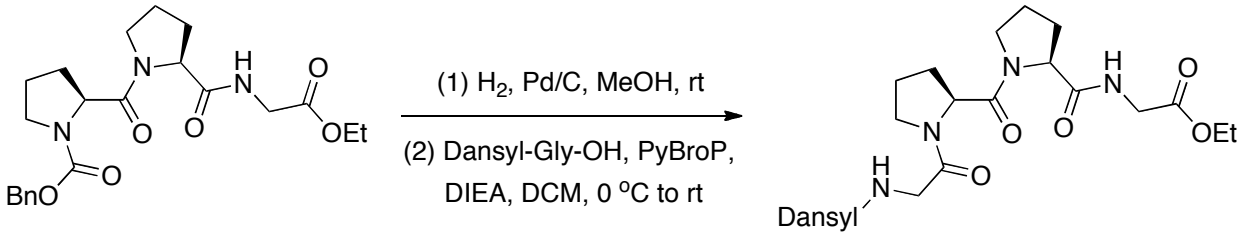
N-(Methoxyoxalyl)glycine ethyl ester was synthesized as reported previously.[148] The spectral data and yield matched that reported previously.

***N*-Benzyloxycarbonyl-(2*S*)-prolyl-(2*S*)-prolylglycine Ethyl Ester (CbzProProGlyOEt)**



CBbzProProGlyOEt was synthesized as described previously.[148] The spectral data and yield matched that reported previously.

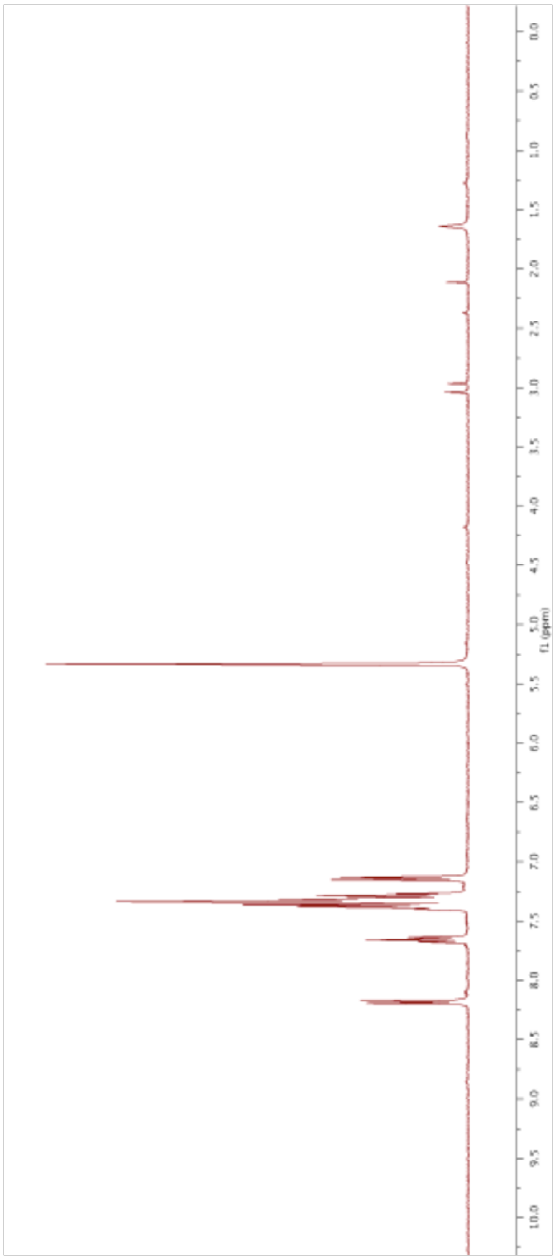
***N*-Dansylglycyl-(2*S*)-prolyl-(2*S*)-prolylglycine Ethyl Ester (DansylGlyProProGlyOEt):**



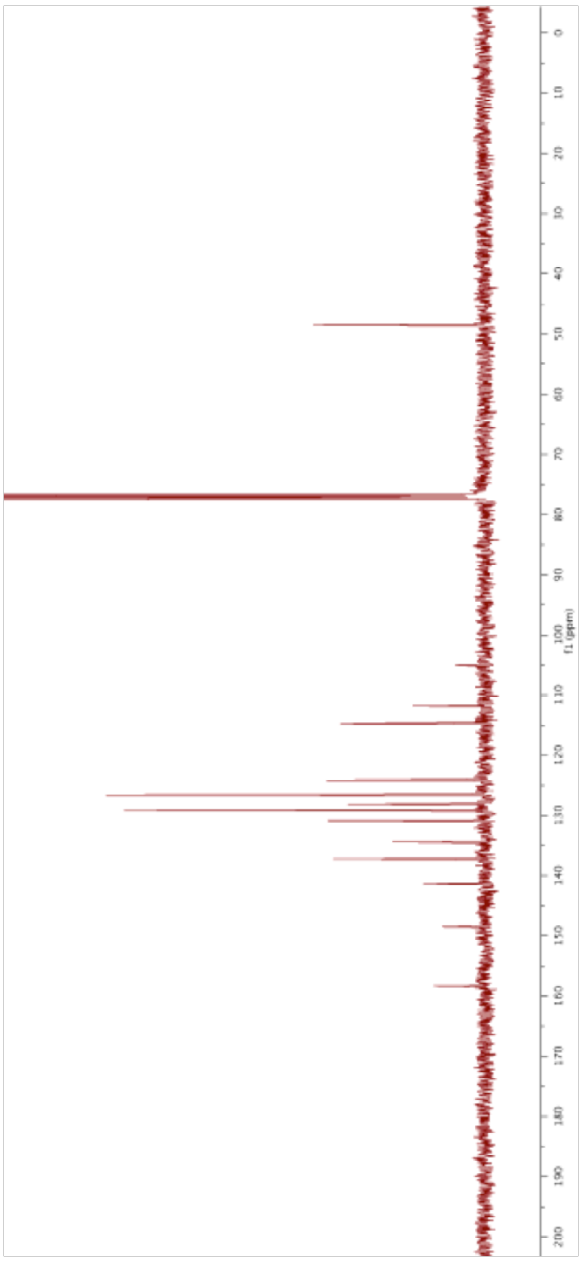
DansylGlyProProGlyOEt was synthesized as described previously.[148] The spectral data and yield matched that reported previously.

B.5.9 NMR Spectra

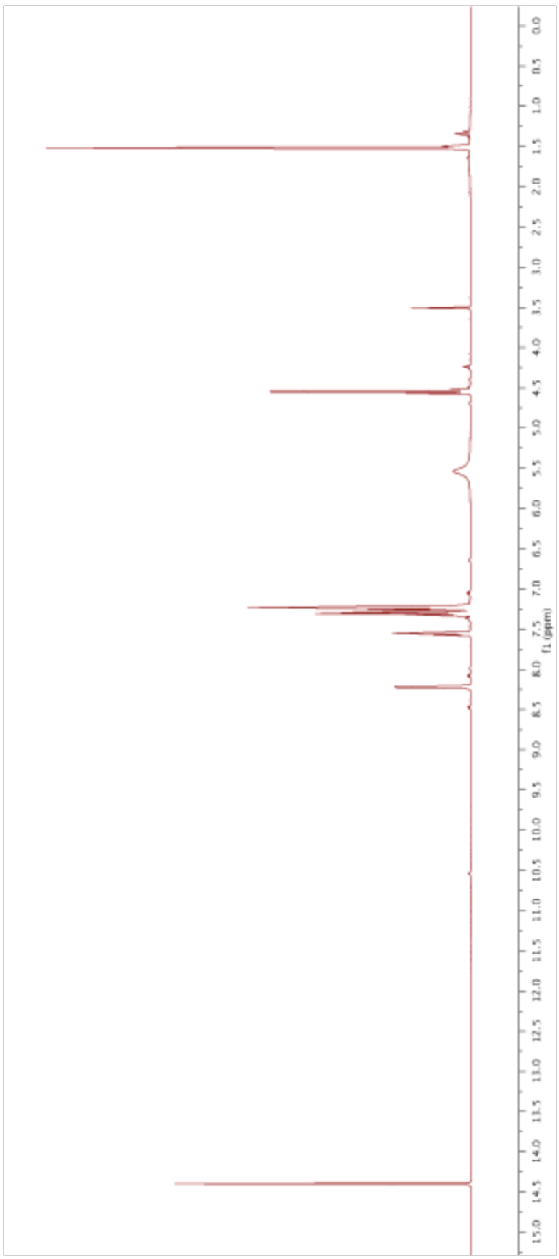
400 MHz ^{1}H NMR Spectrum of *N*-Benzylisatoic Anhydride in $CDCl_3$



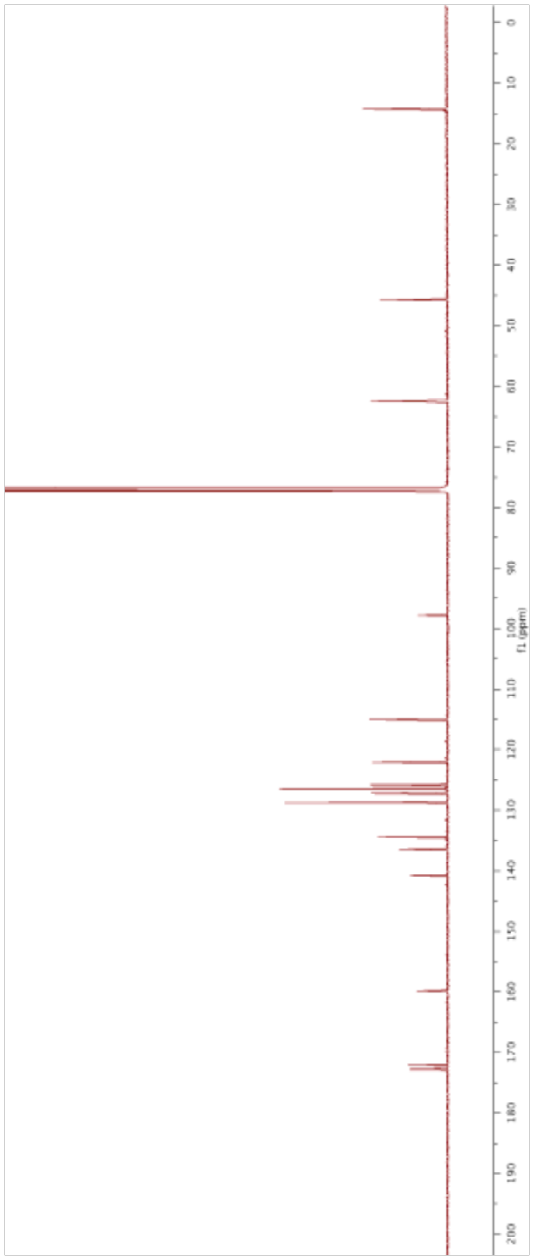
100 MHz ^{13}C NMR Spectrum of *N*-Benzylisatoic Anhydride in $CDCl_3$



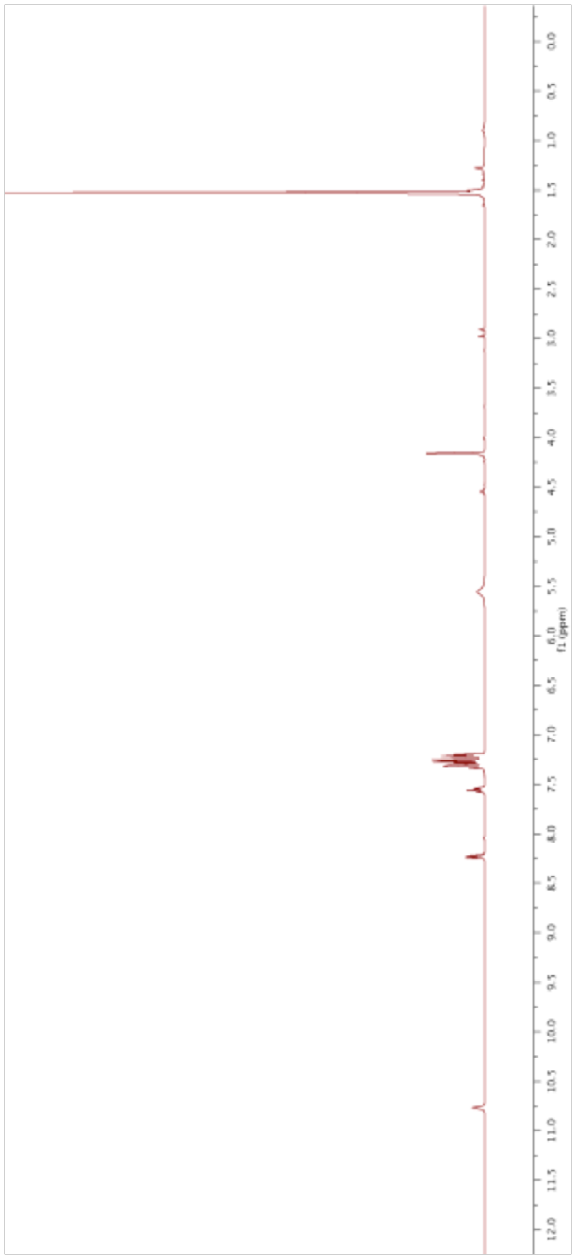
500 MHz ^{1}H NMR Spectrum of Ethyl 4-Hydroxy-1-benzyl-2-oxo-1,2-dihydroquinoline-3-carboxylate in $CDCl_3$



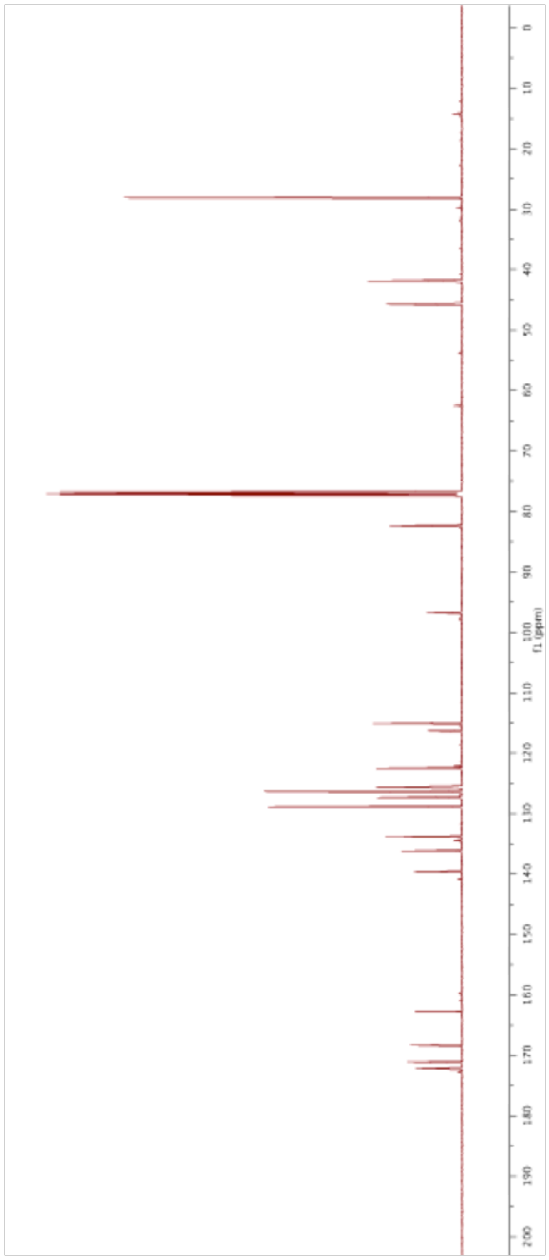
125 MHz ^{13}C NMR Spectrum of Ethyl 4-Hydroxy-1-benzyl-2-oxo-1,2-dihydroquinoline-3-carboxylate in $CDCl_3$



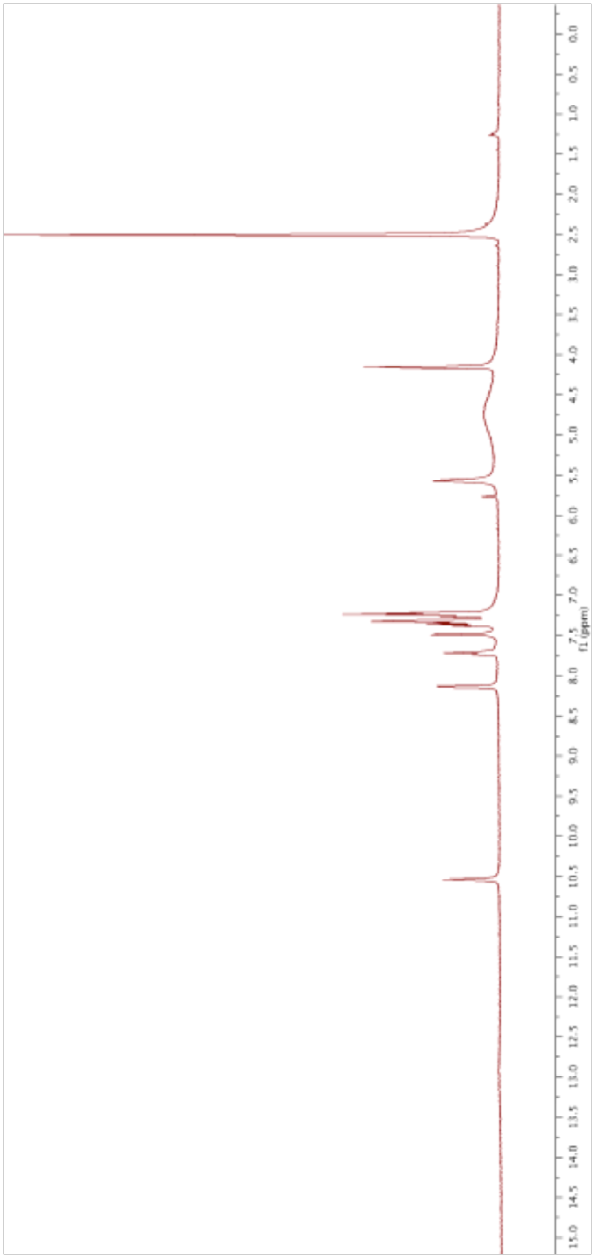
500 MHz ^{1}H NMR Spectrum of tBu IOX2 in $CDCl_3$



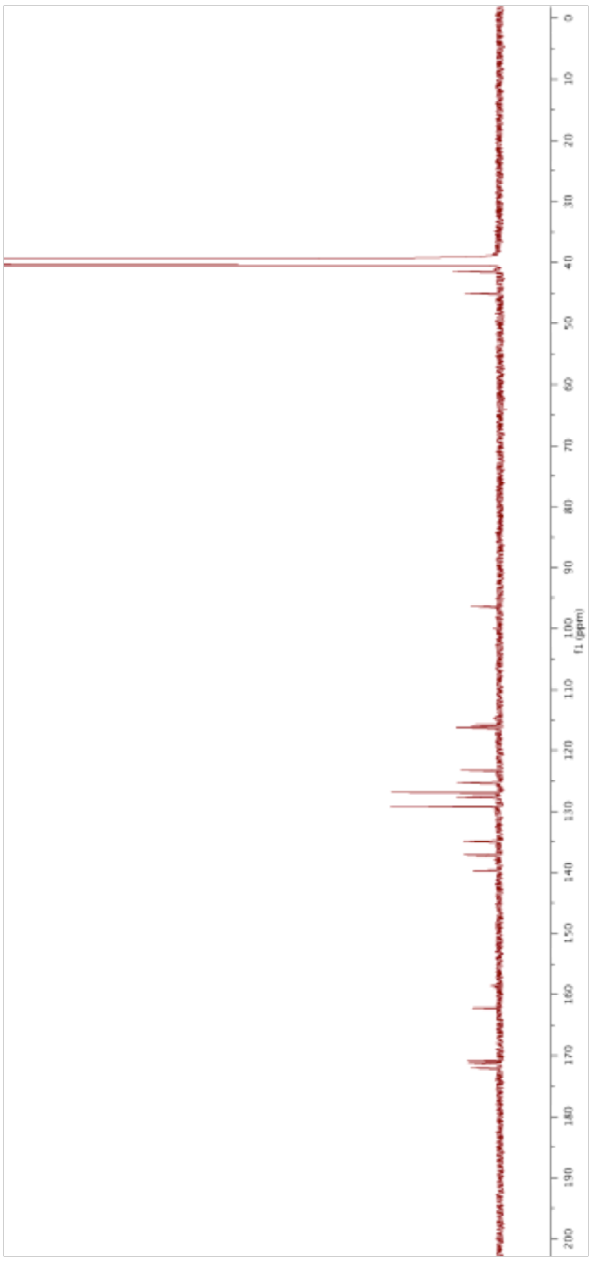
125 MHz ^{13}C NMR Spectrum of tBu IOX2 in $CDCl_3$



500 MHz ^{1}H NMR Spectrum of IOX2 in DMSO-d_6



125 MHz ^{13}C NMR Spectrum of IOX2 in DMSO-d_6



B.6 Acknowledgments

This work was supported by Grant R01 AR044276 (NIH). J. D. V. was supported by Molecular Biosciences Training Grant T32 GM007215 (NIH) and a fellowship from the Department of Biochemistry at the University of Wisconsin–Madison. K. A. A was supported by a PhRMA predoctoral fellowship and by Molecular and Cellular Pharmacology Training Grant T32 GM008688 (NIH). This study made use of the National Magnetic Resonance Facility at Madison, which is supported by Grant P41 GM103399 (NIH). The Micromass LCT® mass spectrometer was obtained with support from Grant CHE-9974839 (NSF).

B.7 Author Contributions

J. D. V. carried out or contributed to carrying out all experiments. K. A. A. contributed to carrying out all experiments relating to human cell culture. J. D. V. and R. T. R. contributed to the design of experiments, analysis of data, and preparation of the manuscript and figures.

Scheme B.1 Synthesis of IOX2

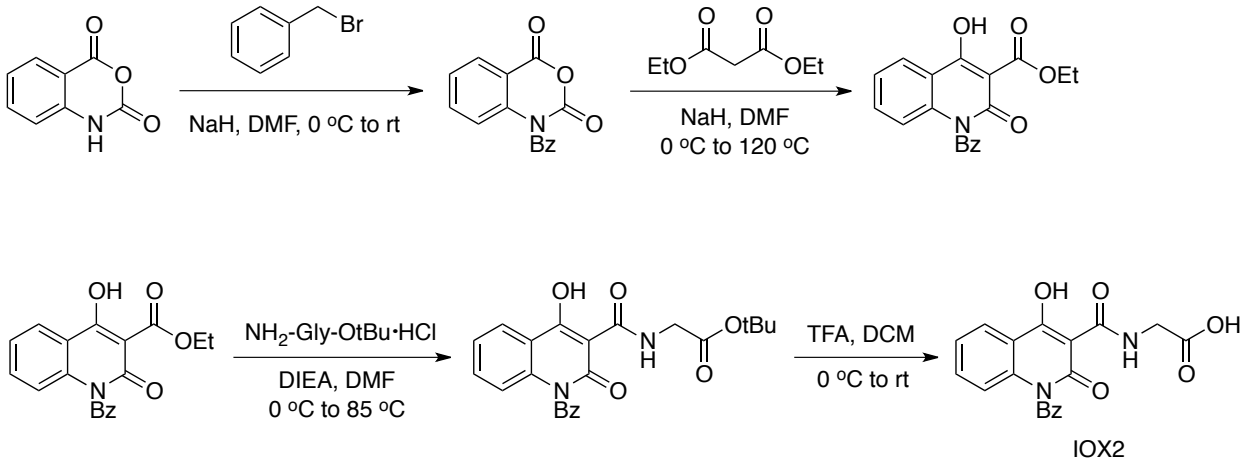


Figure B.1 Inhibition of human CP4H1 *in vitro*

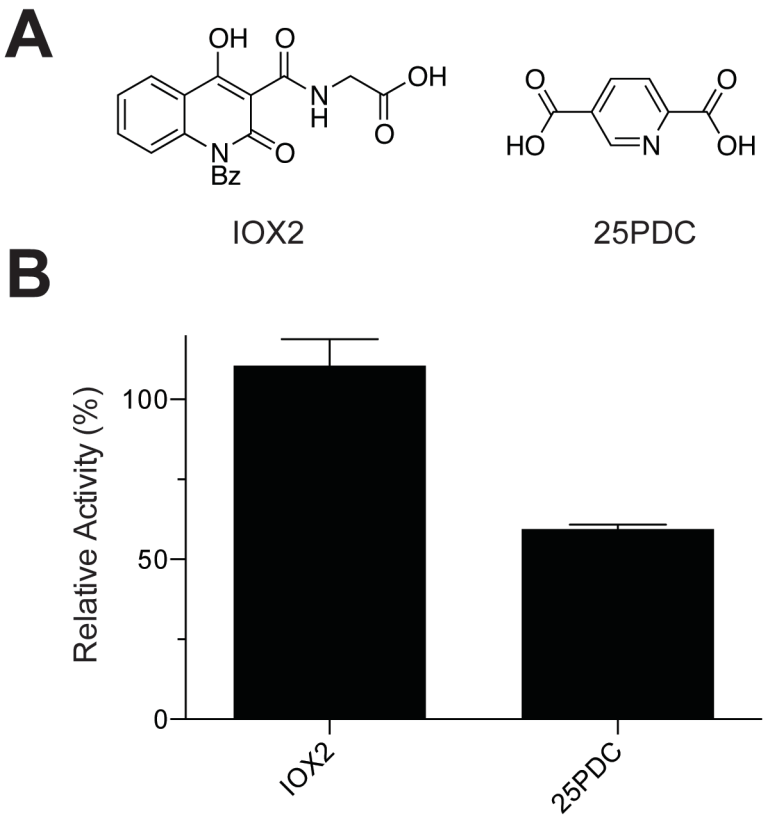


Figure B.1 Inhibition of human CP4H1 *in vitro*. Compounds (A) were tested at a concentration of 10 μM as described in the Experimental Procedures section. Relative activity values (B) are reported as the mean (± SE) of three independent experiments.

Figure B.2 Effect of chelators or P4H Inhibitors on iron metabolism and inhibition of HIF-PHDs in human breast cancer cells

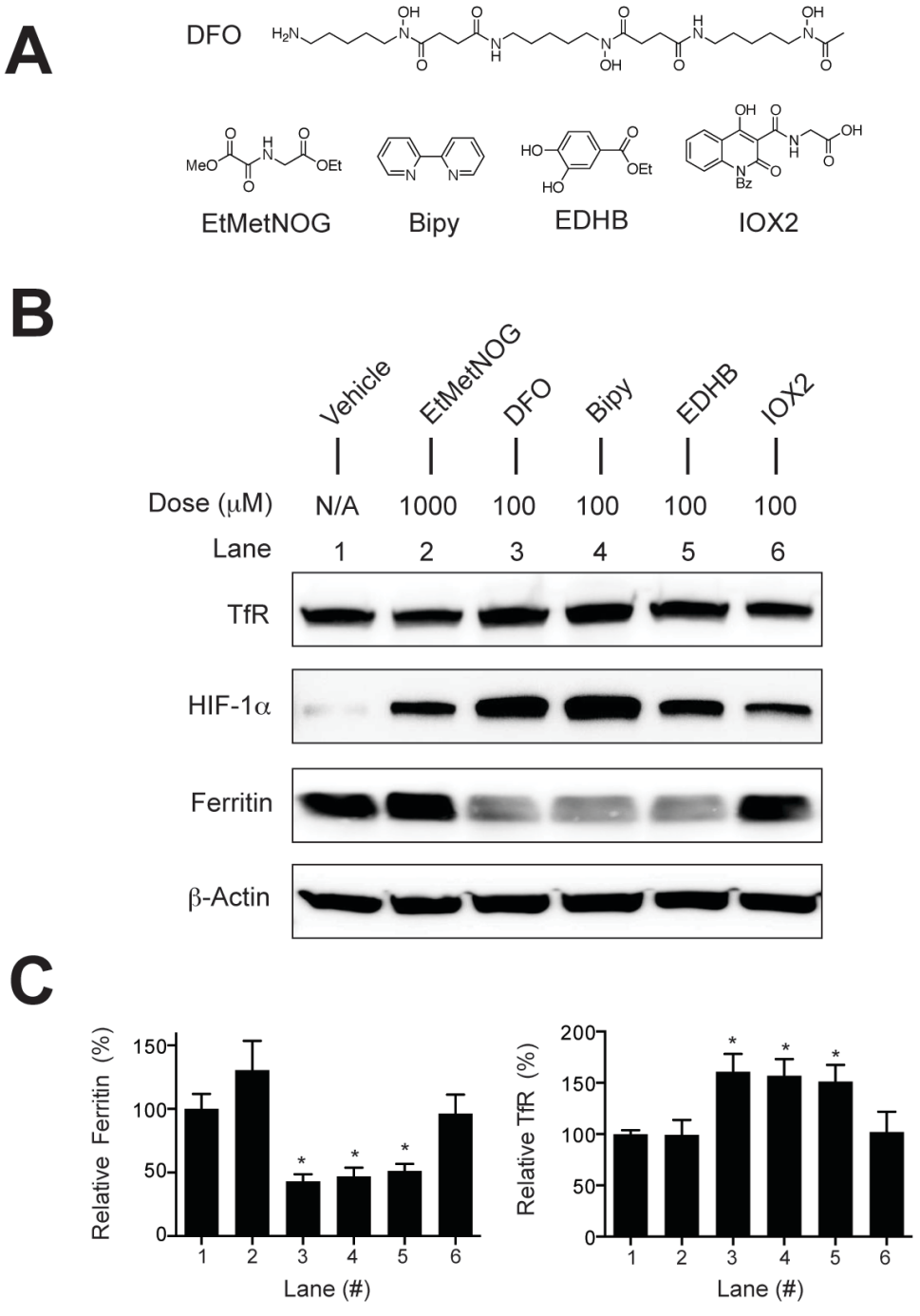


Figure B.2 Effect of chelators or P4H Inhibitors on iron metabolism and inhibition of HIF-PHDs in human breast cancer cells. MDA-MB-231 breast cancer cells were treated with compounds of interest (A) at the specified dosage (B) or DMSO vehicle as described in the Experimental Procedures section, and analyzed with an immunoblot (B). Blots are representative of at least 4 replicates. (C) Densitometric quantitations ($n \geq 4$) corresponding to the blots in panel B and normalized to β-actin.

REFERENCES

1. Shoulders, M. D., and Raines, R. T. (2009) Collagen structure and stability, *Annu. Rev. Biochem. 78*, 929-958.

2. Myllyharju, J., and Kivirikko, K. I. (2004) Collagens, modifying enzymes and their mutations in humans, flies and worms, *Trends Genet. 20*, 33-43.

3. Page, A. P., and Johnstone, I. L. (2007) The cuticle, In *WormBook*, ed. The *C. elegans* Research Community.

4. Gorres, K. L., and Raines, R. T. (2010) Prolyl-4-hydroxylase, *Crit. Rev. Biochem. Mol. Biol. 45*, 106-124.

5. Berg, R. A., and Prockop, D. J. (1973) The thermal transition of a non-hydroxylated form of collagen. Evidence for a role for hydroxyproline in stabilizing the triple helix of collagen, *Biochem. Biophys. Res. Comm. 52*, 115-120.

6. Burjanadze, T. V. (2000) New analysis of the phylogenetic change of collagen thermostability, *Biopolymers 53*, 523-528.

7. Rosenbloom, J., Harsch, M., and Jimenez, S. (1973) Hydroxyproline content determines the denaturation temperature of chick tendon collagen, *Arch. Biochem. Biophys. 158*, 478-484.

8. Ramachandran, G. N., and Reddi, A. H., (Eds.) (1976) *Biochemistry of Collagen*, Plenum Press, New York.

9. Bella, J., Brodsky, B., and Berman, H. M. (1995) Hydration structure of a collagen peptide, *Structure 3*, 893-906.

10. Engel, J., Chen, H.-T., Prockop, D. J., and Klump, H. (1977) The triple helix–coil conversion of collagen-like polytripeptides in aqueous and nonaqueous solvents. Comparison of the thermodynamic parameters and the binding of water to (L-Pro–L-Pro–Gly)$_n$ and (L-Pro–L-Hyp–Gly)$_n$, *Biopolymers 16*, 601-622.

11. Evans, C. A., and Rabenstein, D. L. (1974) Nuclear magnetic resonance studies of the acid-base chemistry of amino acids and peptides. II. Dependence of the acidity of the C-terminal carboxyl group on the conformation of the C-terminal peptide bond, *J. Am. Chem. Soc. 96*, 7312-7317.

12. Madison, V., and Schellman, J. (1970) Location of proline derivatives in conformational space. I. Conformational calculations; optical activity and NMR experiments, *Biopolymers 9*, 511-567.

13. Eberhardt, E. S., Panasik, N., Jr., and Raines, R. T. (1996) Inductive effects on the energetics of prolyl peptide bond isomerization: Implications for collagen folding and stability, *J. Am. Chem. Soc. 118*, 12261-12266.

14. Panasik, N., Jr., Eberhardt, E. S., Edison, A. S., Powell, D. R., and Raines, R. T. (1994) Inductive effects on the structure of proline residues, *Int. J. Pept. Protein Res. 44*, 262-269.

15. Holmgren, S. K., Bretscher, L. E., Taylor, K. M., and Raines, R. T. (1999) A hyperstable collagen mimic, *Chem. Biol. 6*, 63-70.

16. Holmgren, S. K., Taylor, K. M., Bretscher, L. E., and Raines, R. T. (1998) Code for collagen's stability deciphered, *Nature 392*, 666-667.

17. Gelse, K., Pöschl, E., and Aigner, T. (2003) Collagens—structure, function, and biosynthesis, *Adv. Drug Deliv. Rev. 55*, 1531-1546.

18. Myllyharju, J., and Kivirikko, K. I. (2001) Collagens and collagen-related diseases, *Ann. Med. 33*, 7-21.

19. Vranka, J. A., Pokidysheva, E., Hayashi, L., Zientek, K., Mizuno, K., Ishikawa, Y., Maddox, K., Tufa, S., Keene, D. R., Klein, R., and Bächinger, H. P. (2010) Prolyl 3-hydroxylase 1 null mice display abnormalities in fibrillar collagen-rich tissues such as tendons, skin, and bones, *J. Biol. Chem. 285*, 17253-17262.

20. Marina, J. C., Cabral, W. A., and Barnes, A. M. (2010) Null mutations in LEPRE1 and CRTAP cause severe recessive osteogenesis imperfecta, *Cell Tissue Res. 339*, 59-70.

21. Pokidysheva, E., Boudko, S., Vranka, J. A., Zientek, K., Maddox, K., Moser, M., Fässler, R., Ware, J., and Bächinger, H. P. (2014) Biological role of prolyl 3-hydroxylation in type IV collagen, *Proc. Natl. Acad. Sci. U.S.A. 111*, 161-166.

22. Anfinsen, C. B. (1973) Principles that govern the folding of protein chains, *Science 181*, 223-230.

23. Makareeva, E., Nydea, A. A., and Leikin, S. (2011) Chaperoning osteogenesis: New protein-folding disease paradigms, *Trends Cell Biol. 21*, 168-176.

24. Engel, J., and Bächinger, H. P. (2005) Structure, stability and folding of the collagen triple helix, *Top. Curr. Chem. 247*, 7-33.

25. Wedemeyer, W. J., Welker, E., and Scheraga, H. A. (2002) Proline cis-trans isomerization and protein folding, *Biochemistry 41*, 14637-14644.

26. Mouw, J. K., Ou, G., and Weaver, V. M. (2014) Extracellular matrix assembly: A multiscale deconstruction, *Nat. Rev. Mol. Cell Biol. 15*, 771-785.

27. Canty, E. G., Lu, Y., Meadows, R. S., Shaw, M. K., Holmes, D. F., and Kadler, K. E. (2004) Coalignment of plasma membrane channels and protrusions (fibripositors) specifies the parallelism of tendon, *J. Cell Biol. 165*, 553-563.

28. Eisinger-Mathason, T. S., Zhang, M., Qiu, Q., Skuli, N., Nakazawa, M. S., Karakasheva, T., Mucaj, V., Shay, J. E. S., Stangenberg, L., Sadri, N., Puré, E., Yoon, S. S., Kirsch, D. G., and Simon, M. C. (2013) Hypoxia-dependent modification of collagen networks promotes sarcoma metastasis, *Cancer Discov. 3*, 1190-1205.

29. Gilkes, D. M., Chaturvedi, P., Bajpai, S., Wong, C. C., Wei, H., Pitcairn, S., Hubbi, M. E., Wirtz, D., and Semenza, G. L. (2013) Collagen prolyl hydroxylases are essential for breast cancer metastasis, *Cancer Res. 73*, 3285-3296.

30. Myllyharju, J. (2008) Prolyl 4-hydroxylases, key enzymes in the synthesis of collagens and regulation of the response to hypoxia, and their roles as treatment targets, *Ann. Med. 40*, 402-417.

31. Wynn, T. A. (2008) Cellular and molecular mechanisms of fibrosis, *J. Pathol. 214*, 199-210.

32. Friedman, S. L., Sheppard, D., Duffield, J. S., and Violette, S. (2013) Therapy for Fibrotic Diseases: Nearing the Starting Line, *Sci. Transl. Med. 5*, 1-17.

33. Rosenbloom, J., Castro, S. V., and Jimenez, S. A. (2010) Narrative review: Fibrotic diseases: Cellular and molecular mechanisms and novel therapies, *Ann. Intern. Med. 152*, 159-166.

34. Wynn, T. A., and Ramalingam, T. R. (2012) Mechanisms of fibrosis: therapeutic translation for fibrotic disease, *Nat. Med. 18*, 1028-1040.

35. Datta, A., Scotton, C. J., and Chambers, R. C. (2011) Novel therapeutic approaches for pulmonary fibrosis, *Br. J. Pharmacol. 163*, 141-172.

36. O'Connell, O. J., Kennedy, M. P., and Henry, M. T. (2011) Idiopathic pulmonary fibrosis: Treatment update, *Adv. Ther. 28*, 986-999.

37. Cohen-Naftaly, M., and Friedman, S. L. (2011) Current status of novel antifibrotic therapies in patients with chronic liver disease, *Therap. Adv. Gastroenterol. 4*, 391-417.

38. Kisseleva, T., and Brenner, D. A. (2011) Anti-fibrogenic strategies and the regression of fibrosis, *Best Pract. Res. Clin. Gastroentrol. 25*, 305-317.

39. Mathew, A., Cunard, R., and Sharma, K. (2011) Antifibrotic treatment and other new strategies for improving renal outcomes, *Contrib. Nephrol. 170*, 217-227.

40. Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., and Forman, D. (2011) Global cancer statistics, *CA Cancer J. Clin. 61*, 69-90.

41. Stock, A. M., Troost, G., Niggermann, B., Zänker, K. S., and Entschladen, F. (2013) Targets for anti-metastatic drug development, *Curr. Pharm. Des. 19*, 5127-5134.

42. Kim, S. H., Turnbull, J., and Guimond, S. (2011) Extracellular matrix and cell signaling: The dynamic cooperation of integrin, proteoglycan, and growth factor receptor., *J. Endocrinol. 209*, 139-151.

43. Fang, M., Yuan, J., Peng, C., and Li, Y. (2014) Collagen as a double-edged sword in tumor progression, *Tumor Biol. 35*, 2871-2882.

44. O'Connor, J. W., and Gomez, E. W. (2014) Biomechanics of TGFβ-induced epithelial-mesenchymal transition: Implications for fibrosis and cancer, *Clin. Trans. Med. 3*.

45. Wolfe, J. N. (1976) Risk for breast cancer development determined by mammographic parenchymal pattern, *Cancer 37*, 2486-2492.

46. Provenzano, P. P., Inman, D. R., Eliceiri, K. W., Knittel, J. G., Yan, L., Rueden, C. T., White, J. G., and Keely, P. J. (2008) Collagen density promotes mammary tumor initiation and progression, *BMC Med. 6*, 11.

47. Conklin, M. W., Eickhoff, J. C., Riching, K. M., Pehlke, C. A., Eliceiri, K. W., Provenzano, P. P., Friedi, A., and Keely, P. J. (2011) Aligned collagen is a prognostic signature for survival in human breast carcinoma, *Am. J. Pathol. 178*, 1221-1232.

48. Provenzano, P. P., Eliceiri, K. W., Campbell, J. M., Inman, D. R., White, J. G., and Keely, P. J. (2006) Collagen reorganization at the tumer-stromal interface facilitates local invasion, *BMC Med. 4*, 38.

49. Wyckoff, J. B., Wang, Y., Lin, E. Y., Li, J. F., Goswami, S., Stanley, E. R., Segall, J. E., Pollard, J. W., and Condeelis, J. (2007) Direction visualization of macrophage-assisted tumor cell intravasation in mammary tumors, *Cancer Res. 67*, 2649-2656.

50. Dewhirst, M. W. Intermittent hypoxia furthers the rationale for hypoxia-inducible factor-1 targeting, *Cancer Res. 67*, 854-855.

51. Brahimi-Horn, M. C., Chiche, J., and Pouyssegur, J. (2007) Hypoxia and cancer, *J. Mol. Med. 85*, 1301-1307.

52. Bozec, L., van der Heijden, G., and Horton, M. (2007) Collagen fibrils: Nanoscale ropes, *Biophys. J. 92*, 70-75.

53. Chowdhury, R., Hardy, A., and Schofield, C. J. (2008) The human oxygen sensing machinery and its manipulation, *Chem. Soc. Rev. 37*, 1308-1319.

54. Kaelin Jr., W. G., and Ratcliffe, P. J. (2008) Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway, *Mol. Cell 30*, 393-402.

55. Schofield, C. J., and Ratcliffe, P. J. (2004) Oxygen sensing by HIF hydroxylases, *Nat. Rev. Mol. Cell. Biol. 5*, 343-354.

56. Hausinger, R. P. (2004) Fe(II)/α-ketoglutarate-dependent hydroxylases and related enzymes, *Crit. Rev. Biochem. Mol. 39*, 21-68.

57. Rehm, J., Taylor, B., Mohapatra, S., Irving, H., Baliunas, D., Patra, J., and Roerecke, M. (2010) Alcohol as a risk factor for liver cirrhosis: A systematic review and meta-analysis, *Drug Alcohol Rev. 29*, 437-445.

58. Oh, C. K., Murray, L. A., and Molfino, N. A. (2012) Smoking and idiopathic pulmonary fibrosis, *Pulm. Med. 2012*.

59. Frankin, T. J. (1997) Therapeutic approaches to organ fibrosis, *Int. J. Biochem. Cell Biol. 29*, 79-89.

60. Fritz, D., Stefanovic, L., and Stefanovic, B. (2011) Progress toward discovery of antifibrotic drugs targeting synthesis of type I collagen, *Curr. Med. Chem. 18*, 3410-3416.

61. Thompson, K. J., McKillop, I. H., and Schrum, L. W. (2011) Targeting collagen expression in alcoholic liver disease, *World J. Gastroenterol. 17*, 2473-2481.

62. Gotkin, M. G., Ripley, C. R., R., L. S., Bateman, J. F., and Bienkowski, R. S. (2004) Intracellular trafficking and degradation of unassociated proalpha2 chains of collagen type I, *Exp. Cell Res. 296*, 307-316.

63. Ishida, Y., Yamamoto, M., Kitamura, A., Lamandé, S. R., Yoshimori, T., Bateman, J. F., Kubota, H., and Nagata, K. (2009) Autophagic elimination of misfolded procollagen aggregates in the endoplasmic reticulum as a means of cell protection, *Mol. Biol. Cell 20*.

64. Madeo, F., Eisenberg, T., and Kroemer, G. (2009) Autophagy for the avoidance of neurodegeneration, *Genes Dev. 23*, 2253-2259.

65. Thomson, C. A., Atkinson, H. M., and Anantharayanan, V. S. (2005) Identification of small molecule chemical inhibitors of the collagen-specific chaperone Hsp47, *J. Med. Chem. 48*, 1680-1684.

66. Gilkes, D., Bajpai, S., Wong, C. C., Chaturvedi, P., Hubbi, M. E., Wirtz, D., and Semenza, G. L. (2013) Procollagen lysyl hydroxylase 2 is essential for hypoxia-induced breast cancer metastasis, *Mol. Cancer Res. 11*, 456-466.

67. Rose, N. R., Mcdonough, M. A., King, O. N. F., Kawamura, A., and Schofield, C. J. (2011) Inhibition of 2-oxoglutarate dependent oxygenases, *Chem. Soc. Rev. 40*, 4364-4397.

68. Delaet, N. G. J., Robinston, L. A., Wilson, D. M., Sullivan, R. W., Bradley, E. K., Dankwardt, S. M., Martin, R. L., Van Wart, H. E., and Walker, K. A. M. (2003) Novel inhibitors of procollagen C-terminal proteinase. Part 1: Diamino acid hydroxamates.

69. Bailey, S., Fish, P. V., Billotte, S., Bordner, J., Greiling, D., James, K., McElroy, A., Mills, J. E., Reed, C., and Webster, R. (2008) Succinyl hydroxamates as potent and selective non-peptidic inhibitors of procollagen C-proteinase: Design, synthesis and evaluation as topically applied, dermal anti-scarring agents, *Bioorg. Med. Chem. Lett. 18*, 6562-6567.

70. Fish, P. V., Allan, G. A., Bailey, S., Blagg, J., Butt, R., Collis, M. G., Greiling, D., James, K., Kendall, J., McElroy, A., McCleverty, D., Reed, C., Webster, R., and Whitlock, G. A. (2007) Potent and selective nonpeptidic inhibitors of procollagen C-proteinase, *J. Med. Chem. 50*, 3442-3456.

71. Reid, R. R., Mogford, J. E., deGiorgio-Miller, A., and Mustoe, T. A. (2006) Inhibition of procollagen C-proteinase reduces scar hypertrophy in a rabbit model of cutaneous scarring, *Wound Repair Regen. 14*, 138-141.

72. Chien, J. W., Richards, T. J., Gibson, K. F., Zhang, Y., Lindell, K. O., Shao, L., Lyman, S. K., Adamkewicz, J. I., Smith, V., Kaminski, N., and O'Riordan, T. (2014) Serum lysyl

oxidase-like 2 levels and idiopathic pulmonary fibrosis disease progression, *Eur. Respir. J. 43*, 1430-1438.

73. Cox, T. R., Bird, D., Baker, A.-M., Barker, H. E., Ho, M. W.-Y., Lang, G., and Erler, J. T. (2014) Lox-mediated collagen crosslinking is responsible for fibrosis-enhanced metastasis, *Cancer Res. 73*, 1721-1732.

74. Nishioka, T., Eustace, A., and West, C. (2012) Lysyl oxidase: From basic science to future cancer treatment, *Cell Struct. Funct. 37*, 75-80.

75. Xiao, Q., and Ge, G. (2012) Lysyl oxidase, extracellular matrix remodeling and cancer metastasis, *Cancer Microenviron. 5*, 261-273.

76. Kivirikko, K. I., and Prockop, D. J. (1967) Hydroxylation of proline in synthetic polypeptides with purified protocollagen hydroxylase, *J. Biol. Chem. 242*, 4007-4012.

77. Franklin, T. J., Hales, N. J., Johnstone, D., Morris, W. B., Cunliffe, C. J., Millest, A. J., and Hill, G. B. (1991) Approaches to the design of anti-fibrotic drugs, *Biochem. Soc. Trans. 19*, 812-815.

78. Kivirikko, K. I., and Myllyharju, J. (1998) Prolyl 4-hydroxylases and their protein disulfide isomerase subunit, *Matrix Biol. 16*, 357-368.

79. Nwogu, J. I., Geenen, D., Bean, M., Brenner, M. C., Huang, X., and Buttrick, P. M. (2001) Inhibition of collagen synthesis with a prolyl 4-hydroxylase inhibitor improves left ventricular function and alters the pattern of left ventricular dilatation after myocardial infarction, *Circulation 104*, 2216-2221.

80. Ramshaw, J. A. M., Shah, N. K., and Brodsky, B. (1998) Gly–X–Y tripeptide frequencies in collagen: A context for host–guest triple-helical peptides, *J. Struct. Biol. 122*, 86-91.

81. Friedman, L., Higgin, J. J., Moulder, G., Barstead, R., Raines, R. T., and Kimble, J. (2000) Prolyl 4-hydroxylase is required for viability and morphogenesis in *Caenorhabditis elegans*, *Proc. Natl. Acad. Sci. USA 97*, 4736-4741.

82. Holster, T., Pakkanen, O., Soininen, R., Sormunen, R., Nokelainen, M., Kivirikko, K. I., and Myllyharju, J. (2007) Loss of assembly of the main basement membrane collagen, Type IV, but not fibril-forming collagens and embryonic death in collagen prolyl 4-hydroxylase I null mice, *J. Biol. Chem. 282*, 2512-2519.

83. Winter, A. D., and Page, A. P. (2000) Prolyl 4-hydroxylase is an essential procollagen-modifying enzyme required for exoskeleton formation and the maintenance of body shape in the nematode *Caenorhabditis elegans*, *Mol. Cell Biol. 20*, 4084-4093.

84. Anantharajan, J., Koski, M. K., Kursula, P., Hieta, R., Bergmann, U., Myllyharju, J., and Wierenga, R. K. (2013) The structural motifs for substrate binding and dimerization of the α subunit of collagen prolyl 4-hydroxylase, *Structure 21*, 2107-2118.

85. Hieta, R., Kukkola, L., Permi, P., Pirilä, P., Kivirikko, K. I., Kilpeläinen, I., and Myllyharju, J. (2003) The peptide-substrate binding domain of human collagen prolyl 4-hydroxylases. Backbone assignments, secondary structure, and binding of proline-rich peptides, *J. Biol. Chem. 278*, 34966-34974.

86. Pekkala, M., Hieta, R., Bergmann, U., Kivirikko, K. I., Wierenga, R. K., and Myllyharju, J. (2004) The peptide-substrate-binding domain of collagen prolyl 4-hydroxylase is a

tetratricopeptide repeat domain with functional aromatic residues, *J. Biol. Chem. 279*, 52255-52261.

87. Myllyharju, J., and Kivirikko, K. I. (1997) Characterization of the iron- and 2-oxoglutarate-binding sites of human prolyl 4-hydroxylase, *EMBO J. 16*, 1173-1180.

88. Cassab, G. I. (1998) Plant cell wall proteins, *Annu. Rev. Plant Physiol. Plant Mol. Biol. 49*, 281-309.

89. Kieliszewski, M. J., and Shpak, E. (2001) Synthetic genes for the elucidation of glycosylation codes for arabinogalactan-proteins and other hydroxyproline-rich glycoproteins, *Cell. Mol. Life Sci. 58*, 1386-1398.

90. Keskiaho, K., Hieta, R., Sormunen, R., and Myllyharju, J. (2007) *Chlamydomonas reinhardtii* has multiple prolyl 4-hydroxylases, one of which is essential for proper cell wall assembly, *Plan Cell 19*, 256-269.

91. Koski, M. K., Hieta, R., Böllner, C., Kivirikko, K. I., Myllyharju, J., and Wierenga, R. K. (2007) The active site of an algal prolyl 4-hydroxylase has large structural plasticity, *J. Biol. Chem. 282*, 37112-37123.

92. Koski, M. K., Hieta, R., Hirsilä, M., Rönkä, A., Myllyharju, J., and Wierenga, R. K. (2009) The crystal structure of an algal prolyl 4-hydroxylase complexed with a proline-rich peptide reveals a novel buried tripeptide binding motif, *J. Biol. Chem. 284*, 25290-25301.

93. Longbotham, J. E., Levy, C., Johannissen, L. O., Tarhonskaya, H., Jiang, S., Loenarz, C., Flashman, E., Hay, S., Schofield, C. J., and Scrutton, N. S. (2015) Structure and mechanism of a viral collagen prolyl hydroxylase, *Biochemistry 54*, 6093-6105.

94. Eriksson, M., Myllyharju, J., Tu, H., Hellman, M., and Kivirikko, K. I. (1999) Evidence for 4-hydroxyproline in viral proteins. Characterization of a viral prolyl 4-hydroxylase and its peptide substrates, *J. Biol. Chem. 274*, 22131-22134.

95. Hoffart, L. M., Barr, E. W., Guyer, R. B., Bollinger, J. M., and Krebs, C. (2006) Direct spetroscopic detection of a C-H-cleaving high-spin Fe(IV) complex in a prolyl 4-hydroxylase, *Proc. Natl. Acad. Sci. U.S.A. 103*, 14738-14743.

96. Myllyharju, J. (2003) Prolyl 4-hydroxylases, the key enzymes of collagen biosynthesis, *Matrix Biol. 22*, 15-24.

97. Hanauske-Abel, H. M., and Günzler, V. (1982) A stereochemical concept for the catalytic mechanism of prolylhydroxylase: Applicability to classification and design of inhibitors, *J. Theor. Biol. 94*, 421-455.

98. Myllylä, R., Tuderman, L., and Kivirikko, K. I. (1977) Mechanism of the prolyl hydroxylase reaction. 2. Kinetic analysis of the reaction sequence, *Eur. J. Biochem. 80*, 349-357.

99. Kivirikko, K. I., and Risteli, L. (1976) Biosynthesis of collagen and its alterations in pathological states, *Med. Biol. 54*, 159-186.

100. Myllylä, R., Kuutti-Savolainen, E. R., and Kivirikko, K. I. (1978) The role of ascorbate in the prolyl hydroxylase reaction, *Biochem. Biophys. Res. Commun. 83*, 441-448.

101. De Jong, L., Albracht, S. P. J., and A., K. (1982) Prolyl 4-hydroxylase activity in relation to the oxidation state of enzyme-bound iron: The role of ascorbate in peptidyl proline hydroxylation, *Biochim. Biophys. Acta 704*, 326-332.

102. Myllylä, R., Majamaa, K., Gunzler, V., Hanauske-Abel, H. M., and Kivirikko, K. I. (1984) Ascorbate is consumed stoichiometrically in the uncoupled reactions catalyzed by prolyl 4-hydroxylase and lysyl hydroxylase, *J. Biol. Chem. 259*, 5403-5405.

103. Guzman, N. A., (Ed.) (1998) *Prolyl Hydroxylase, Protein Disulfide Isomerase, and Other Structurally Related Proteins*, Marcel Dekker, New York.

104. Hirsilä, M., Koivunen, P., Xu, L., Seeley, T., Kivirikko, K. I., and Myllyharju, J. (2005) Effect of desferrioxamine and metals on the hydroxylases in the oxygen sensing pathway, *FASEB J. 19*, 1308-1310.

105. Kivirikko, K. I., and Myllylä, R. (1987) Recent developments in postranslational modification: Intracellular processing, *Methods Enzymol. 144*, 96-114.

106. Hales, N. J., and Beattie, J. F. (1993) Novel inhibitors of prolyl 4-hydroxylase. 5. The intriguing structure–activity relationships seen with 2,2'-bipyridine and its 5,5'-dicarboxylic acid derivatives, *J. Med. Chem. 36*, 3853-3858.

107. Majamaa, K., Hanauske-Abel, H. M., Günzler, V., and Kivirikko, K. I. (1984) The 2-oxoglutarate binding site of prolyl 4-hydroxylase. Identification of distinct subsites and evidence for 2-oxoglutarate decarboxylation in a ligand reaction at the enzyme-bound ferrous ion, *Eur. J. Biochem. 138*, 239-245.

108. Switzer, B. R., and Summer, G. K. (1973) Inhibition of collagen synthesis by α,α′-dipyridyl in human skin fibroblasts in culture, *In Vitro 9*, 160-166.

109. Whitson, T. C., and Peacock, E. E. J. (1969) Effect of alpha alpha dipyridyl on collagen synthesis in healing wounds, *Surg. Gynecol. Obstet. 128*, 1061-1064.

110. Sheth, S. (2014) Iron chelation: An update, *Curr. Opin. Hematol. 21*, 179-185.

111. Flora, S. J. S., and Pachauri, V. (2010) Chelation in metal intoxication, *Int. J. Environ. Res. Public Health 7*, 2745-2788.

112. Franklin, T. J., Morris, W. P., Edwards, P. N., Large, M. S., and Stephenson, R. (2001) Inhibition of prolyl 4-hydroxylase *in vitro* and *in vivo* by members of a novel series of phenanthrolinones, *Biochem. J. 353*, 333-338.

113. Majamaa, K., Turpeenniemi-Hujanen, T. M., Latipää, P., Günzler, V., Hanauske-Abel, H. M., Hassinen, I. E., and Kivirikko, K. I. (1985) Differences between collagen hydroxylases and 2-oxoglutarate dehydrogenase in their inhibition by structural analogues of 2-oxoglutarate, *Biochem. J. 229*, 127-133.

114. Cunliffe, C. J., Franklin, T. J., Hales, N. J., and Hill, G. B. (1992) Novel inhibitors of prolyl 4-hydroxylase. 3. Inhibition by the substrate analogue N-oxaloglycine and its derivatives, *J. Med. Chem. 35*, 2652-2658.

115. Kukkola, L., Hieta, R., Kivirikko, K. I., and Myllyharju, J. (2003) Identification and characterization of a third human, rat, and mouse collagen prolyl 4-hydroxylase isoenzyme, *J. Biol. Chem. 278*, 47685-47693.

116. Majamaa, K., Günzler, V., Hanauske-Abel, H. M., Myllylä, R., and Kivirikko, K. I. (1986) Partial identity of the 2-oxoglutarate and ascorbate binding sites of prolyl 4-hydroxylase, *J. Biol. Chem. 261*, 7819-7823.

117. Dowell, R. I., Hales, N. H., and Tucker, H. (1993) Novel inhibitors of prolyl 4-hydroxylase. Part 4. Pyridine-2-carboxylic acid analogues with alternative 2-substituents, *Eur. J. Med. Chem. 28*, 513-516.

118. Dowell, R. I., and Hadley, E. M. (1992) Novel inhibitors of prolyl 4-hydroxylase, *J. Med. Chem. 35*, 800-804.

119. Rife, J. E., and Cleland, W. W. (1980) Determination of the chemical mechanism of glutamate dehydrogenase for pH studies, *Biochemistry 19*, 2328-2333.

120. Northrop, D. B., and Cleland, W. W. (1974) The kinetics of pig heart triphosphopyridine nucleotide-isoscitrate dehydrogenase, *J. Biol. Chem. 249*, 2928-2931.

121. Hutton, J. J., Marglin, A., Witkop, B., Kurtz, J., Berger, A., and Udenfriend, S. (1968) Synthetic polypeptides as substrates and inhibitors of collagen proline hydroxylase, *Arch. Biochem. Biophys. 125*, 779-785.

122. Gorres, K. L., Edupuganti, R., Krow, G. R., and Raines, R. T. (2008) Conformational preferences of substrates for human prolyl 4-hydroxylase, *Biochemistry 47*, 9447-9455.

123. Günzler, V., Hanauske-Abel, H. M., Myllyla, R., Kaska, D. D., Hanauske, A., and Kivirikko, K. I. (1988) Syncatalytic inactivation of prolyl 4-hydroxylase by anthracyclines, *Biochem. J. 251*, 365-372.

124. Calabresi, P., and Parks Jr., R. E. (1980) Chemotherapy of neoplastic diseases, In *The Pharmacological Basis of Therapeutics* (Gilman, A. G., Goodman, L. S., and Gilman, A., Eds.), pp 1249-1313, Macmillan Publishing Co., New York.

125. Bland, K. I., Palin, W. E., Fraunhofer, J. A., Morris, R. R., Adcock, R. A., and Tobin II, G. R. (1984) Experimental and clinical observations of the effects of cytotoxic chemotherapeutic drugs on wound healing., *Ann. Surg. 199*, 782-790.

126. Sasaki, T., Holeyfield, K. C., and Uitto, J. (1987) Doxorubicin-induced inhibition of prolyl hydroxylation during collagen biosynthesis in human skin fibroblast cultures, *J. Clin. Invest. 80*, 1735-1741.

127. Günzler, V., Brocks, D., Henke, S., Myllyla, R., Geiger, R., and Kivirikko, K. I. (1988) Syncatalytic inactivation of prolyl 4-hydroxylase by peptides containing the unphysiologic amino acid 5-oxaproline, *J. Biol. Chem. 263*, 19498-19504.

128. Wu, M., Moon, H.-S., and Begley, T. P. (1999) Mechanism-based inactivation of the human prolyl-4-hydroxylase by 5-oxaproline-containing peptides: Evidence for a prolyl radical intermediate, *J. Am. Chem. Soc. 121*, 587-588.

129. Karvonen, K., Ala-Kokko, L., Pihlajaniemi, T., Helaakoski, T., Henke, S., Günzler, V., Kivirikko, K. I., and Savolainen, E. R. (1990) Specific inactivation of prolyl 4-hydroxylase and inhibition of collagen synthesis by oxaproline-containing peptide in cultured human skin fibroblasts, *J. Biol. Chem. 265*, 8415-8419.

130. Günzler, V., Hanauske-Abel, H. M., Myllylä, R., Mohr, J., and Kivirikko, K. I. (1987) Time-dependent inactivation of chick-embryo prolyl 4-hydroxylase by coumalic acid, *Biochem. J. 242*, 163-169.

131. Wang, J., Buss, J. L., Chen, G., Ponka, P., and Pantopoulos, K. (2002) The prolyl 4-hydroxylase inhibitor ethyl-3,4-dihydroxybenzoate generates effective iron deficiency in cultured cells, *FEBS Lett. 529*, 309-312.

132. Schweigert, N., Alexander, J., Zehnder, B., and Eggen, R. I. L. (2001) Chemical properties of catechols and their molecular modes of toxic action in cells, from microorganisms to mammals, *Environ. Microbiol. 3*, 81-91.

133. Chowdhury, R., Candela-Lena, J. I., Chan, M. C., Greenald, D. J., Yeoh, K. K., Tian, Y., McDonough, M. A., Tumber, A., Rose, N. R., Conejo-Garcia, A., Demetriades, M., Mathavan, S., Kawamura, A., Lee, M. K., van Eeden, F., Pugh, C. W., Ratcliffe, P. J., and Schofield, C. J. (2013) Selective small molecule probes for the hypoxia inducible factor (HIF) prolyl hydroxylases, *ACS Chem. Biol. 8*, 1488-1496.

134. Vasta, J. D., Higgin, J. J., Kersteen, E. A., and Raines, R. T. (2013) Bioavailable affinity label for collagen prolyl 4-hydroxylase, *Bioorg. Med. Chem. 21*, 3597-3601.

135. Campeau, L.-C., Rosseaux, S., and Fagnou, K. (2005) A solution to the 2-pyridyl organometallic cross-coupling problem: Regioselective catalytic direct arylation of pyridine *N*-oxides, *J. Am. Chem. Soc. 127*, 18020-18021.

136. Campeau, L.-C., Stuart, D. R., Leclerc, J.-P., Bertrand-Laperle, M., Villemure, E., Sun, H.-Y., Lasserre, S., Guimond, N., Lecavallier, M., and Fagnou, K. (2009) Palladium-

catalyzed direct arylation of azine and azole *N*-oxides: Reaction development, scope, and applications in synthesis, *J. Am. Chem. Soc. 131*, 3291-3306.

137. Duric, S., and Tzschucke, C. C. (2011) Synthesis of unsymmetrically substituted bipyridines by palladium-catalyzed direct C–H arylation of pyridine-*N*-oxides, *Org. Lett. 13*, 2310-2313.

138. Fischer, B. E., Häring, U. K., Tribolet, R., and Sigel, H. (1979) Metal Ion/Buffer Interactions: Stability of Binary and Ternary Complexes Containing 2-Amino-2(hydroxymethyl)-1,3-propanediol (Tris) and Adenosine 5'-Triphosphate (ATP), *Eur. J. Biochem. 94*, 523-530.

139. Hanlon, D. P., Watt, D. S., and Westhead, E. W. (1966) The Interaction of Divalent Metal Ions with Tris Buffer in Dilute Solution, *Anal. Biochem. 16*, 225-233.

140. Kersteen, E. A., Higgin, J. J., and Raines, R. T. (2004) Production of human prolyl 4-hydroxylase in *Escherichia coli*, *Protein Expr. Purif. 38*, 279-291.

141. Hewitson, K. S., Schofield, C. J., and Ratcliffe, P. J. (2007) Hypoxia-inducible factor prolyl-hydroxylase: Purification and assays of PHD2, *Methods Enzymol. 435*, 25-42.

142. Gibson, D. G. (2011) Enzymatic assembly of overlapping DNA fragments, *Methods Enzymol. 498*, 349-361.

143. Huang, C. Y. (1982) Determination of binding stoichiometry by the continuous variation method: The Job plot, *Methods Enzymol. 87*, 509-525.

144. Job, P. (1928) Formation and stability of inorganic complexes in solution, *Annali di Chimica Applicata 9*, 113-203.

145. Costas, M., Mehn, M. P., Jensen, M. P., and Que, L. J. (2004) Dioxygen activation at mononuclear nonheme iron active sites: Enzymes, models, and intermediates, *Chem. Rev. 104*, 939-986.

146. Abraham, R. J., and McLauchlan, K. A. (1962) The proton resonance spectra and conformations of the prolines. Part II., *Mol. Phys. 5*, 513-523.

147. Sasaki, T., Majamaa, K., and Uitto, J. (1987) Reduction of collagen production in keloid fibroblast cultures by ethyl-3,4-dihydroxybenzoate, *J. Biol. Chem. 262*, 9397-9403.

148. Vasta, J. D., and Raines, R. T. (2015) Selective inhibition of prolyl 4-hydroxylases by bipyridinedicarboxylates, *Bioorg. Med. Chem. 23*, 3081-3090.

149. Frey, P. A., and Reed, G. H. (2012) The ubiquity of iron, *ACS Chem. Biol. 7*, 1477-1481.

150. Baldwin, J. J., Kasinger, P. A., Novello, F. C., Sprague, J. M., and Duggan, D. E. (1975) 4-Trifluoromethylimidazoles and 5-(4-pyridyl)-1,2,4-triazoles, new classes of xanthine oxidase inhibitors, *J. Med. Chem. 18*, 895-900.

151. Vuluga, D., Legros, J., Crousse, B., and Donnet-Delpon, D. (2009) Synthesis of pyrazoles through catalyst-free cycloaddition of diazo compounds to alkynes, *Green Chem. 11*, 156-159.

152. Martin, T., Verrier, C., Hoarau, C., and Marsais, F. (2008) Direct C-2 arylation of alkyl 4-thiazolecarboxylates: New insights in synthesis of heterocyclic core of thiopeptide antibiotics, *Org. Lett. 10*, 2909-2912.

153. Strotman, N. A., Chobanian, H. R., Guo, Y., He, J., and Wilson, J. E. (2010) Highly regioselective palladium-ctalyzed direct arylation of oxazole at C-2 or C-5 with aryl bromides, chlorides, and triflates, *Org Lett 12*, 3578-3581.

154. Liu, W., Yu, X., Li, Y., and Kuang, C. (2014) Palladium-catalyzed oxidative CH/CH cross-coupling of pyridine *N*-oxides with five-membered heterocycles, *Chem. Commun. 50*, 9291-9294.

155. Anderson, C. P., Shen, L., Eisenstein, R. S., and Leibold, E. A. (2012) Mammalian iron metabolism and its control by iron regulatory proteins, *Biochim. Biophys. Acta 1823*, 1468-1483.

156. Muckenthaler, M. U., Bruno, G., and Hentze, M. W. (2008) Systemic iron homeostasis and the iron-responsive element/iron-regulatory protein (IRE/IRP) regulatory network, *Annu. Rev. Nutr. 28*, 197-213.

157. Rautio, J., Kumpulainen, H., Heimbach, T., Oliyai, R., Oh, D., Järvinen, T., and Savolainen, J. (2008) Prodrugs: design and clinical applications, *Nat Rev Drug Discov 7*, 255-270.

158. Testa, B., and Mayer, J. M. (2003) *Hydrolysis in Drug and Prodrug Metabolism: Chemistry, Biochemistry, and Enzymology*, Academic Press, New York, NY.

159. Eisenstein, R. S., Tuazon, P. T., Schalinske, K. L., Anderson, S. A., and Traugh, J. A. (1993) Iron-responsive element-binding protein. Phosphorylation by protein kinase C., *J. Biol. Chem. 268*, 27363-27370.

160. Goforth, J. B., Anderson, S. A., Nizzi, C. P., and Eisenstein, R. S. (2010) Multiple determinants within iron-responsive elements dictate iron regulatory protein binding and regulatory hierarchy, *RNA 16*, 154-169.

161. Begtrup, M., Boyer, G., Cabildo, P., Cativiela, C., Claramunt, R. M., Elguero, J., García, J. I., Toiron, C., and Vedsø, P. (1993) ^{13}C NMR of pyrazoles, *Magn. Reson. Chem. 31*, 107-168.

162. Pearson, R. G., and Williams, F. V. (1953) Rates of ionization of pseudo acids. V. Steric effects in the base-catalyzed ionization of nitroethane, *J. Am. Chem. Soc. 75*, 3073-3075.

163. Albert, A., Goldacre, R., and Phillips, J. (1948) The strength of heterocyclic bases, *J. Chem. Soc.*, 2240-2249.

164. Ford, P. C., DeForest, P. R., Gaunder, R., and Taube, H. (1968) Synthesis and properties of pentaamminepyridineruthenium(II) and related pentaammineruthenium complexes of aromatic nitrogen heterocycles, *J. Am. Chem. Soc. 90*, 1187-1194.

165. Walba, H., and Isensee, R. W. (1961) Acidity constants of some arylimidazoles and their cations, *J. Org. Chem. 26*, 2789-2791.

166. Catalan, J., and Elguero, J. (1987) Basicity and acidity of azoles, *Adv. Heterocycl. Chem. 41*, 187-274.

167. Zoltewicz, J. A., and Deady, L. W. (1978) Quaternization of heteroaromatic compounds: Quantitative aspects, *Adv. Heterocycl. Chem. 22*, 71-121.

168. Brown, D. J., and Ghosh, P. B. (1969) The spectra, ionization, and deuteriation of oxazoles and related compounds, *J. Chem. Soc. B*, 270-276.

169. Krumholz, P. (1949) Ferrous mono-α-α′-dipyridyl, *J. Am. Chem. Soc. 71*, 3654-3656.

170. Guo, H., Tong, P., Liu, Y., Xia, L., Wang, T., Tian, O., Li, Y., Hu, Y., Zheng, Y., Jin, X., Li, Y., Xiong, W., Tang, B., Feng, Y., Li, J., Pan, O., Hu, Z., and Xia, Z. (2015) Mutations of P4HA2 encoding prolyl 4-hydroxylase 2 are associated with nonsyndromic high myopia, *Genet. Med. 17*, 300-306.

171. Torpe, N., and Pocock, R. (2014) Regulation of axonal midline guidance by prolyl 4-hydroxylation in *Caenorhabditis elegans*, *J. Neurosci. 34*, 16348-16357.

172. Kimura, M., Yamamoto, M., and Yamaba, S. (1982) Kinetics and mechanism of the oxidation of L-ascorbic acid by tris(oxalato)cobaltate(III) and tris(1,10-phenanthroline)iron(III) complexes in aqueous solution, *J. Chem. Soc. Dalton Trans.*, 423-427.

173. Younathan, E. S., and Frieden, E. (1961) Catalysis of ascorbate oxidation by 1,10-phenanthroline and related ligands, *Biochim. Biophys. Acta 46*, 51-58.

174. Abraham, R. J., and McLauchlan, K. A. (1962) The proton resonance spectra and conformations of the prolines. Part I., *Mol. Phys. 5*, 195-203.

175. Atreya, P. L., and Ananthanarayanan, V. S. (1991) Interaction of prolyl 4-hydroxylase with synthetic peptide substrates, *J. Biol. Chem. 266*, 2852-2858.

176. Brahmachari, S. K., and Anantharayanan, V. S. (1979) *β*-Turns in nascent procollagen are sites of posttranslational enzymatic hydroxylation of proline., *Proc. Natl. Acad. Sci. U.S.A. 76*, 5119-5123.

177. Rapaka, R. S., Renugopalakrishnan, V., Urry, D. W., and Bhatnagar, R. S. (1978) Hydroxylation of proline in polytripeptide models of collagen: Stereochemistry of polytripeptide-prolyl hydroxylase interaction., *Biochemistry 17*, 2892-2989.

178. Fischer, G. (2000) Chemical aspects of peptide bond isomerisation, *Chem. Soc. Rev. 29*, 119-127.

179. Bretscher, L. E., Jenkins, C. L., Taylor, K. M., DeRider, M. L., and Raines, R. T. (2001) Conformational stability of collagen relies on a stereoelectronic effect, *J. Am. Chem. Soc. 123*, 777-778.

180. DeRider, M. L., Wilkens, S. J., Waddell, M. J., Bretscher, L. E., Weinhold, F., Raines, R. T., and Markley, J. L. (2002) Collagen stability: Insights from NMR spectroscopic and hybrid density functional computational investigations of the effect of electronegative substituents on prolyl ring conformations, *J. Am. Chem. Soc. 124*, 2497-2505.

181. Sarkar, S. K., Young, P. E., Sullivan, C. E., and Torchia, D. A. (1984) Detection of *cis* and *trans* C-Pro peptide bonds in proteins by ^{13}C NMR: Application to collagen, *Proc. Natl. Acad. Sci. U.S.A. 81*, 4800-4803.

182. Bächinger, H. P., Bruckner, P., Timpl, R., Prockop, D. J., and Engel, J. (1980) Folding mechanism of the triple helix in type-III collagen and type-III pN-collagen, *Eur. J. Biochem. 106*, 619-632.

183. Bruckner, P., Bächinger, H. P., Timpl, R., and Engel, J. (1978) Three conformationally distinct domains in the amino-terminal segment of type III procollagen and its rapid triple helix – coil transition, *Eur. J. Biochem. 90*, 593-603.

184. Bruckner, P., Eikenberry, E. F., and Prockop, D. J. (1981) Formation of the Triple Helix of Type I Procollagen in cellulo. A Kinetic Model based on cis-trans Isomerization of Peptide Bonds, *Eur. J. Biochem. 118*, 607-613.

185. Gorres, K. L., and Raines, R. T. (2009) Direct and continuous assay for prolyl 4-hydroxylase, *Anal. Biochem. 386*, 181-185.

186. Tandon, M., Wu, M., and Begley, T. P. (1998) Substrate specificity of human prolyl-4-hydroxylase, *Bioorg. Med. Chem. Lett. 8*, 1139-1144.

187. Bandur, N. G., Harms, K., and Koert, U. (2005) First stereoselective synthesis of a Pro–Pro *E*-alkene dipeptide isosteres, *Synlett*, 773-776.

188. Bandur, N. G., Harms, K., and Koert, U. (2007) Synthesis of two dipeptide isosteres containing di- and trisubstituted E-configured double bonds, *Synthesis*, 2720-2730.

189. Dai, N., and Etzkorn, F. A. (2009) Cis-trans proline isomerization effects on collagen triple-helix stability are limited, *J. Am. Chem. Soc. 131*, 13728-13732.

190. Dai, N., Wang, X. J., and Etzkorn, F. A. (2008) The effect of a *trans*-locked Gly-Pro alkene isostere on collagen triple helix stability, *J. Am. Chem. Soc. 130*, 5396-5397.

191. Wang, X. J., Hart, S. A., Xu, B., Mason, M. D., Goodell, J. R., and Etzkorn, F. A. (2003) Serine-*cis*-proline and serine-*trans*-proline isosteres: Stereoselective synthesis of (*Z*)- and (*E*)-alkene mimics by Still–Wittig and Ireland–Claisen rearrangements, *J. Org. Chem. 68*, 2343-2349.

192. Hart, S. A., Sabat, M., and Etzkorn, F. A. (1998) Enantio- and regioselective synthesis of a (Z)-alkene cis-proline mimic, *J. Org. Chem. 63*, 7580-7581.

193. Still, W. C., and Mitra, A. (1978) A highly stereoselective synthesis of z-trisubstituted olefins via [2,3]-sigmatropic rearrangement. Preference for a psuedoaxially substituted transition state., *J. Am. Chem. Soc. 100*, 1927-1928.

194. Cravatt, B. F., Wright, T. W., and Koszarich, J. W. (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry, *Annu. Rev. Biochem. 77*, 383-414.

195. Higgin, J. J. (2004) Small-molecule enzyme inhibitors utilizing active-site metal chelation: Prolyl 4-hydroxylase and microbial ribonuclease, In *Chemistry*, University of Wisconsin–Madison, Madison.

196. Appleton, D., Duguid, A. B., Lee, S.-K., Ha, Y.-J., Ha, H.-J., and Leeper, F. J. (1998) Synthesis of analogues of 5-aminolaevulinic acid and inhibition of 5-aminolaevulinic acid dehydratase, *J. Chem. Soc. Perkin Trans. 1*, 89-101.

197. Pinkas, D. M., Ding, S., Raines, R. T., and Barron, A. E. (2011) Tunable, post-translational hydroxylation of collagen domains in Escherichia coli, *ACS Chem. Biol 6*, 320-324.

198. Gorres, K. L., Pua, K. H., and Raines, R. T. (2009) Stringency of the 2-His-1-Asp active-site motif in prolyl 4-hydroxylase, *PLoS One 4*, e7635.

199. Zhang, Z., Smith, B. A., Wang, L., Brock, A., Cho, C., and Schultz, P. G. (2003) A new strategy for the site-specific modification of protein in vivo, *Biochemistry 42*, 6735-6746.

200. Kalia, J., and Raines, R. T. (2008) Hydrolytic stability of hydrazones and oximes, *Angew. Chem. Int. Ed. 47*, 7253-7256.

201. Schweitzer, M. H., Zheng, W., Organ, C. L., Avci, R., Suo, Z., Freimark, L. M., Lebleu, V. S., Duncan, M. B., Vander Heiden, M. G., Neveu, J. M., Lane, W. S., Cottrell, J. S., Horner, J. R., Cantley, L. C., Kalluri, R., and Asara, J. M. (2009) Biomolecular characterization and protein sequences of the Campanian hadrosaur *B. canadensis*, *Science 324*, 626-631.

202. Myllyharju, J., Kukkola, L., Winter, A. D., and Page, A. P. (2002) The exoskeleton collagens in *Caenorhabditis elegans* are modified by prolyl 4-hydroxylases with unique combinations of subunits, *J. Biol. Chem. 277*, 29187-29196.

203. Fraisl, P., Aragonés, J., and Carmeliet, P. (2009) Inhibition of oxygen sensors as a therapeutic strategy for ischaemic and inflammatory disease, *Nat. Rev. Drug Discov. 8*, 139-152.

204. Kaelin Jr., W. G. (2011) Cancer and altered metabolism: Potential importance of hypoxia-inducible factor and 2-oxoglutarate-dependent dioxygenases, *Cold Spring Harb. Symp. Quant. Biol. 76*, 335-345.

205. Wright, G., Higgin, J. J., Raines, R. T., Steenbergen, C., and Murphy, W. (2003) Activation of the prolyl hydroxylase oxygen-sensor results in induction of GLUT1, heme oxygenase-1, and nitric-oxide synthase proteins and confers protection from metabolic inhibition to cardiomyocytes, *J. Biol. Chem. 278*, 20235-20239.

206. Ong, S. G., and Hausenloy, D. J. (2012) Hypoxia-inducible factor as a therapeutic target for cardioprotection, *Pharmacol. Ther. 136*, 69-81.

207. Sen Banerjee, S., Thirunavukkarasu, M., Tipu Rishi, M., Sanchez, J. A., Maulik, N., and Maulik, G. (2012) HIF-prolyl hydroxylases and cardiovascular diseases, *Toxicol. Mech. Methods 22*, 347-358.

208. Linstead, R. P., Weedon, B. C. L., and Wladislaw, B. (1955) Anodic synthesis. Part XIII. Chain extension of fatty acids by electrolysis with benzyl half esters, *J. Chem. Soc.*, 1097-1100.

209. Abell, A. D., Morris, K. B., and Litten, J. C. (1990) Synthesis and deprotection of [1-(ethoxycarbonyl)-4-[(diphenylmethoxy)carbonyl]-1-methyl-2-oxobutyl]triphenylphosphonium chloride: A key intermediate in the Wittig reaction between a cyclic anhydride and a stabilized ylide, *J. Org. Chem. 55*, 5217-5221.

210. Milstein, D., and Stille, J. K. (1979) Mild, selective, general method of ketone synthesis from acid chlorides and organotin compounds catalyzed by palladium, *J. Org. Chem. 44*, 1613-1618.

211. Kitz, R., and Wilson, I. B. (1962) Esters of methanesulfonic acid as irreversible inhibitors of acetylcholinesterase, *J. Biol. Chem. 237*, 3245-3249.

212. Hope, I. A. (1999) *C. elegans: A Practical Approach*, Oxford University Press, New York.

213. Lauret, C. (2001) Epoxy ketones as versatile building blocks in organic synthesis, *Tetrahedron: Asymmetry 12*, 2359-2383.

214. Mutsaers, S. E., Bishop, J. E., McGrouther, G., and Laurent, G. J. (1997) Mechanisms of tissue repair: From wound healing to fibrosis, *Int. J. Biochem. Cell Biol. 29*, 5-17.

215. Rieder, F., Brenmoehl, J., Leeb, S., Schölmerich, J., and Rogler, G. (2007) Wound healing and fibrosis in intestinal disease, *Gut 56*, 130-139.

216. Gabrielli, A., Avvedimento, E. V., and Krieg, T. (2009) Scleroderma, *N. Engl. J. Med. 360*, 1989-2003.

217. Ley, S. V., Stewart-Liddon, A. J. P., Pears, D., Perni, R. H., and Treacher, K. (2006) Hydrogenation of aromatic ketones, aldehydes, and epoxides with hydrogen and Pd(0)Encat^{TM} 30NP, *Bellstein J. Org. Chem. 2*.

218. Phukan, P., Chakraborty, P., and Katak, D. (2006) A simple and efficient method for regioselective and stereoselective synthesis of vicinal bromohydrins and alkoxybromides from an olefin, *J. Org. Chem. 71*, 7533-7537.

219. Nair, C. G. R., and Indrasenan, P. (1976) New redox titrants in non-aqueous or partially aqueous media-VI: Potentiometric determinations using dibromamine-T and some further applications of dichloramine-T, *Talanta 23*, 239-241.

220. Luo, X., Liu, Y., Kubicek, S., Myllyharju, J., Tumber, A., Ng, S., Che, K. H., Podoll, J., Heightman, T. D., Opperman, U., Schreiber, S. L., and Wang, X. (2011) A selective inhibitor and probe of the cellular functions of Jumonji C domain-containing histone demethylases, *J. Am. Chem. Soc. 133*, 9451-9456.