UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Innovative Approaches to Library Design for Fragment-based Drug Discovery

Permalink

Author

Stokes, Ryjul Wynn

Publication Date

2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Innovative Approaches to Library Design for Fragment-based Drug Discovery

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Chemistry

by

Ryjul Wynn Stokes

Committee in charge:

    Professor Seth M. Cohen, Chair
    Professor Michael Gilson
    Professor Darren Lipomi
    Professor Arnold Rheingold
    Professor Jerry Yang

2022

The Dissertation of Ryjul Wynn Stokes is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2022

DEDICATION

For Sydnee

Home is wherever I'm with you

EPIGRAPH

There is not a "fragment" in all nature, for every relative fragment of one thing is a full harmonious unit in itself.

-John Muir

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

2D	Two-Dimensional
3D	Three-Dimensional
Å	Ångström
Ac_2O	Acetic anhydride
AcOH	Acetic acid
ACS	American Chemical Society
ADMET	Absorption, Distribution, Metabolism, Excretion, and Toxicity
$AlCl_3$	Aluminum Chloride
ALS	Advanced Light Source
ATP	Adenosine Triphosphate
B3LYP	Becke 3-parameter Lee–Yang–Parr
Bn	Benzyl
BnOH	Benzyl alcohol
Br_2	Bromine
bs	broad singlet
C	Celsius
CADD	Computer-Aided Drug Design
CCD	Charge-Coupled Device
CCDC	Cambridge Crystallographic Data Centre
$(CD_3)_2CO$	Deuterated acetone
CD_3OD	Deuterated methanol
CDI	Carbonyldiimidazole

clogP	Calculated log P
CHAPS	Cholamidopropyl)dimethylammonio]-1-propanesulfonate.
CI	Confidence Interval
CO	Carbon Monoxide
CP	Cyclopentadienyl
CPCM	Conductor-like Polarizable Continuum Model
CSD	Cambridge Structural Database
CTAB	Cetrimonium Bromide
CuI	Copper Iodide
d	Doublet
Da	Daltons
DBU	1,8-Diazabicyclo(5.4.0)undec-7-ene
DCM	Dichloromethane
dd	doublet of doublets
dt	doublet of triplets
DFT	Density Functional Theory
DNA	Deoxyribonucleic acid
DNAse	Deoxyribonuclease
DMF	Dimethylformamide
DMSO	Dimethyl Sulfoxide
DTT	Dithiothreitol
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid

ESI-MS	Electrospray Ionization Mass Spectrometry
Et$_2$O	Diethyl ether
EtOAc	Ethyl Acetate
EtOH	Ethanol
FAM	6-carboxyfluorescein
FBDD	Fragment-Based Drug Discovery
FDA	Food and Drug Administration
FITC	Fluorescein Isothiocyanate
FRET	Förster resonance energy transfer
g	Times Gravity
G	Gibbs Free Energy
GSK3β	β-isoform of glycogen synthase kinase 3
h	hours
H	Hydrogen
H$_2$O	Water
HAC	Heavy Atom Count
hCAII	human Carbonic Anhydrase II
HCl	Hydrochloric acid
Hex	Hexanes
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	Human Immunodeficiency Virus
HNO$_2$	Nitrous acid
HOBt	Hydroxybenzotriazole

HONH$_2$·HCl	Hydroxylamine hydrochloride
HPLC	High-Performance Liquid Chromatography
HR-ESI-MS	High-Resolution Electrospray Ionization Mass Spectrometry
HRMS	High Resolution Mass Spectrometry
Hsp90	N-terminal domain of heat shock protein 90-α
HTL	Hit-To-Lead
HTS	High-Throughput Screening
IC$_{50}$	Half maximal Inhibitor Concentration
IPTG	β-D-1-thiogalactopyranoside
J	Coupling constant
K	Kelvin
KCl	Potassium Chloride
K$_2$CO$_3$	Potassium Carbonate
KDM	Jumonji-domain of Histone Lysine Demethylase
KOH	Potassium hydroxide
L	Liter
LB	Luria broth
LC-TOFMS	Liquid Chromatography Time-of-Flight Mass Spectrometry
LE	Ligand Efficiency
LiAlH$_4$	Lithium Aluminum Hydride
LiCl	Lithium Chloride
LogP	Octanol-water partition coefficient
m	multiplet

M	Molar
MBI	Metal-Binding Isostere
MBP	Metal-Binding Pharmacophore
MD	Molecular Dynamic
MeCN	Acetonitrile
MeOH	Methanol
mF	metallofragment
min	minutes
MM	Molecular Mechanical
MOE	Molecular Operating Environment
mg	Milligram
$MgCl_2$	Magnesium Chloride
$MgCl_2 \cdot 6H_2O$	Magnesium Chloride Hexahydrate
$MgSO_4$	Magnesium Sulfate
MHz	Megahertz
mL	milliliters
mM	millimolar
mmol	millimoles
$MnCl_2$	Manganese Chloride
MW	Molecular Weight
MWCO	Molecular Weight Cutoff
MV	Molecular Volume
m/z	mass-to-charge ratio

n-BuLi	n-Butyllithium
$NaBH_4$	Sodium Borohydride
NaCl	Sodium Chloride
NaH	Sodium Hydride
NaN_3	Sodium azide
$NaNO_2$	Sodium Nitrite
NaOAc	Sodium Acetate
NaOH	Sodium Hydroxide
Na_2PO_4	Sodium Phosphate
NDM-1	New Delhi metallo-β-lactamase-1
NEt_3	Triethylamine
NH	Ammonia
NH_4Cl	Ammonium Chloride
nm	nanometer
nM	nanomolar
NMR	Nuclear Magnetic Resonance
NNRTI	Non-Nucleoside Reverse Transcriptase Inhibitor
NSAID	Nonsteroidal Anti-Inflammatory Drug
OS	Octasporines
OD_{600}	Optical density at 600 nm
PA_N	N-terminal domain of the Polymerase Acidic subunit of the RNA-dependent RNA polymerase of the influenza virus
PB1	Polymerase Basic protein 1

PB2	Polymerase Basic protein 2
PC	Physicochemical
PK	Pharmacokinetic
pK_a	Acid dissociation constant
PDB	Protein Data Bank
PE	Petroleum Ether
PEG	Polyethylene Glycol
PMI	Principal Moment of Inertia
PPM	Parts Per Million
q	quartet
QM	Quantum Mechanical
RMSD	Root-Mean-Square Deviation
RNA	Ribonucleic acid
RPM	Revolutions Per Minute
s	singlet
SAR	Structure-Activity Relationship
SBVS	Structure-Based Virtual Screening
SDS-PAGE	Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis
sec	seconds
SMILES	Simplified Molecular Input Line Entry System
S_NAr	Nucleophilic Aromatic Substitution
SOC	Super Optimal Concentrated
$SOCl_2$	Thionyl Chloride

t	triplet
TAM	5-carboxytetramethylrhodamine
TCEP	Tris(2-carboxyethyl)phosphine
THF	Tetrahydrofuran
TEV	Tobacco Etch Virus protease
TFA	Trifluoro acetic acid
TLC	Thin layer chromatography
TMS	Trimethylsilyl
Tris	Tris(hydroxymethyl)aminomethane
UV	Ultraviolet
VS	Virtual Screening
$ZnCl_2$	Zinc Chloride
δ	Chemical shift; ppm
Δ	Difference
ζ	Average of the sum of the deviation of 6 unique metal–ligand bond lengths around the central metal atom
λ	Wavelength
μL	Microliter
μM	Micromolar
Σ	Sum of the deviation of 12 unique cis ligand–metal–ligand angles from 90°

ACKNOWLEDGEMENTS

Many people have played a part in the journey that has led to this point in my life. Consequently, I have many people to thank and acknowledge. Unfortunately, there is not room here to name all those who have played an important role in my life; however, I would like to thank and acknowledge all of those who are not explicitly listed here.

I can trace my interest in science back to early science questions from my father, Richard Stokes, who first taught me to love learning. His influence on my life and my career cannot be understated. I must also thank my mother, Keli Stokes, for her continual support and enthusiasm towards my studies. She has been a constant source of encouragement and has always demonstrated unmatched interest in and attention toward my research progress. I would also like to thank my siblings Jazmyn Murphy, Madyson Cohen, and Cyrus Stokes for the important influences they have been in my life. I would also like to thank my mother- and father-in-law, Melanie and Byde Willis, for their constant help and support.

I've had the opportunity to work with many amazing people throughout the course of my studies. I'd like to recognize Eric Anderson for his early and important influence on me as an aspiring adolescent scientist. I would also like to thank my excellent undergraduate mentor Whitney Walker, who patiently taught me how to work in a lab. I must also acknowledge my undergraduate advisor David Michaelis, who first taught me organic chemistry and exposed me to research as an undergraduate student. My experience working with him was transformative and had the single biggest influence on the trajectory of my career. I would like to thank him for the many important experiences I had in his lab and for laying the foundation of my research career. I would also like to thank Michael Schmidt, my mentor during my time at Bristol-Myers Squibb, for his patience and for the important experiences he provided me with. Furthermore,

I would also like to thank the excellent mentors I had in graduate school. Cy Credille, Christine Morrison, Kathleen Prosser, and Johannes Karges were all important mentor figures to me. I would like to thank each of them for their constant patience and guidance throughout my studies. Additionally, I would like to thank Merritt Andrus, Kyle Barcus, Scott Burt, Steven Castle, Jeffrey Macedone, Anthony Mrse, Conor O'herin, Yongxuan Su, Michael Talley, and all Cohen Lab members past and present for their help and support. Lastly, I would like to thank my Ph.D. advisor Seth Cohen, from whom I have learned many valuable lessons. Seth has pushed me to be a better scientist and communicator and has helped me to address many difficult challenges. I am grateful to him for his guidance, and for his advocacy, both of which have been critical during my time in graduate school.

I would also like to acknowledge my children Wren and Elliot Stokes. They have both provided me with joy and motivation, and I appreciate the important things they have taught me. The wonder and creativity they consistently display has been a source of inspiration for me as I seek to learn and innovate.

Last, and most importantly, I would like to thank my partner Sydnee Stokes, who has been the most important person in my life for the past 10 years. She has been my constant support throughout my entire education. I have watched her struggle and sacrifice as she enabled me to pursue my goals, for which I will always be grateful. She provided me with the support and advice that I needed during the most difficult times, and I will always appreciate her steadfast companionship. Of the many people acknowledged here, I have her to thank the most.

Chapter 1, in part, is a reprint of the material as it appears in "Metal complexes for therapeutic applications." *Trends Chem.,* **2021,** *3*, 523. The dissertation author was a primary

author of this paper and gratefully acknowledges the contributions of coauthors Johannes Karges and Seth M. Cohen.

Chapter 2, in large part, is a reprint of the material as it appears in "Computational Prediction of the Binding Pose of Metal-Binding Pharmacophores." *ACS Med. Chem. Lett.,* **2022,** 1*3*, 3, 428. The dissertation author was a primary author of this paper and gratefully acknowledges the contributions of coauthors Johannes Karges and Seth M. Cohen. Reprinted with permission from *ACS Med. Chem. Lett.* 2022, *13*, 3, 428–435. Copyright 2022 American Chemical Society.

Chapter 3, in large part, is a reprint of the material as it appears in "Carboxylic Acid Isostere Derivatives of Hydroxypyridinones as Core Scaffolds for Influenza Endonuclease Inhibition," which has been accepted for publication. The dissertation author was a primary author of this paper and gratefully acknowledges the contributions of coauthors Alysia Kohlbrand, Hyeonglim Seo, Banumathi Sankaran, Johannes Karges, and Seth M. Cohen. Reprinted with permission from *ACS Med. Chem. Lett*. Copyright 2022 American Chemical Society.

Chapter 4, in large part, is a reprint of the material as it appears in "Evaluation of 3-Dimensionality in Approved and Experimental Drug Space." *ACS Med. Chem. Lett.,* **2020,** *11*, 6, 1292-1298. The dissertation author was a primary author of this paper and gratefully acknowledges the contributions of coauthors Kathleen E. Prosser and Seth M. Cohen. Reprinted with permission from *ACS Med. Chem. Lett.* **2020**, *11*, 6, 1292-1298. Copyright 2020 American Chemical Society.

Chapter 5, in large part, is a reprint of the material as it appears in "Expanding medicinal chemistry into 3D space: metallofragments as 3D scaffolds for fragment-based drug discovery."

VITA

2017   Bachelor of Arts in Chemistry, Brigham Young University

2019   Master of Science in Chemistry, University of California San Diego

2022   Doctor of Philosophy in Chemistry, University of California San Diego

PUBLICATIONS

14. **Stokes, R. W.**[†]; Kohlbrand, A. J.[†]; Karges, J.; Seo, H.; Sankaran, B.; Cohen S. M. Carboxylic Acid Isostere Derivatives of Hydroxypyridinones as Core Scaffolds for Influenza Endonuclease Inhibition. *ACS Med. Chem Lett. In Press.*

13. Karges, J. K**.; Stokes, R. W**.; Cohen S. M. Computational Prediction of the Binding Pose of Metal-Binding Pharmacophores for Fragment-Based Drug Discovery. *ACS Med. Chem. Lett. 2022, 13,* 428–435.

12. Rogolino, D.[†]; Naesens, L.[†]; Bartolia, J.; Carcelli M.; De Luca, L.; Pelosi, G.; Steavert, A.; **Stokes, R. W.**; Van Berwaer, R.; Vittorio, S.; Cohen S. M. Exploration of the 2,3-Dihydroisoindole Pharmacophore for Inhibition of the Influenza Virus PA Endonuclease. *Bioorg. Chem.* **2021**, *116*, 105388.

11. **Stokes, R. W.**[†]; Karges, J.[†]; Cohen, S. M. Metal Complexes for Therapeutic Applications. *Trends in Chemistry.* **2021**, *3*, 523 – 534.

10. Adamek, R. N.[†]; Suire, C. N.[†]; **Stokes, R. W.**; Brizuela, M. K.; Cohen, S. M.; Leissring, M. A. Hydroxypyridinethione Inhibitors of Human Insulin-Degrading Enzyme. *ChemMedChem.* **2021**, *16*, 1776 – 1788.

9. Karges, J.; **Stokes, R. W.**; Cohen, S. M. Rationally Designed Photorelease of a Metal-Binding Pharmacophore from a Ru(II) Polypyridine Complex. *Dalton Trans.* **2021**. *50*, 2757 – 2765.

8. Prosser, K. E.; **Stokes, R. W.;** Cohen, S. M. An Evaluation of 3-Dimensionality in Existing Approved and Experimental Drug Space. *ACS Med. Chem. Lett.* **2020**, 11, 1292 – 1298.

7. **Stokes, R. W.**[†]; Morrison, C. N.[†]; Prosser, K. E[†].; Cordes, A.; Metzler-Nolte, N.; Cohen, S. M. Expanding Medicinal Chemistry into 3D Space: Metallofragments as 3D Scaffolds for Fragment-Based Drug Discovery. *Chem. Sci.* **2020**, *11*, 1216 – 1225.

6. **Stokes, R. W.**[†]; Credille, C. V.[†]; Morrison, C. N.[†]; Dick, B. L. [‡]; Yifan, F. [‡]; Sun, J.; Chen, Y.; Cohen, S. M. SAR Exploration of Tight-Binding Inhibitors of Influenza Virus PA Endonuclease. *J. Med. Chem.* **2019***, 62,* 9438 – 9449.

5. Ence, C. C.[†]; Walker, W. K.[†]; **Stokes, R. W.**; Martinez, E. E.; Sarager, S. M.; Smith, S. J.; Michaelis, D. J. Synthesis of chiral titanium-containing phosphinoamide ligands for enantioselective heterobimetallic catalysis. *Tetrahedron.* **2019***, 75,* 3341 – 3347.

4. Credille, C. V.; Dick, B. L.; Morrison, C. N.; **Stokes, R. W.**; Adamek, R. N.; Wu, N. C.; Wilson, I. A.; Cohen, S. M. Structure-Activity Relationships in Metal-Binding Pharmacophores for Influenza Endonuclease. *J. Med. Chem.* **2018***, 61,* 10206 – 10217.

3. Schmidt, M. A.; **Stokes, R. W**.; Davies, M. L.; Roberts, F. 4-Cyanobenzenesulfonamides: An Amine Synthesis and Protecting Strategy to Compliment the Nosyl Group. *J. Org. Chem.*, **2017**, *82*, 4550–4560.

2. **Stokes, R, W.** [†]; Talley M. R. [†]; Walker, W. K.; Michaelis, D. J. Electrophilic activation of alkynes for enyne cycloisomerization reactions with *in situ* generated early/late heterobimetallic Pt–Ti catalysts. *Dalton Trans.* **2016**, *45*, 9770–9773

1. Walker, W. K.; Anderson, D. L.; **Stokes, R. W**.; Smith, S. L.; Michaelis, D. J. Allylic Aminations with Hindered Secondary Amine Nucleophiles Catalyzed by Heterobimetallic Ti–Pd Complexes. *Org. Lett.* **2015**, *17* (3), 752–755.

[†] Equal contribution

ABSTRACT OF THE DISSERTATION

Innovative Approaches to Library Design for Fragment-based Drug Discovery

by

Ryjul Wynn Stokes

Doctor of Philosophy in Chemistry

University of California San Diego, 2022

Professor Seth M. Cohen, Chair

Fragment-based drug discovery (FBDD) is an important technology in drug discovery that seeks to identify molecules that can inhibit targets of therapeutic relevance. Once a fragment 'hit' is identified, it can be elaborated into a molecule with improved inhibition and favorable pharmaceutical properties. In this thesis, several strategies to augment existing FBDD techniques are discussed. While these approaches are quite different from each other,

each is rooted in fundamental inorganic chemistry and utilizes principles of coordination chemistry to address problems in FBDD. First, computational screening techniques are discussed which may help accelerate FBDD against metalloenzyme targets. This strategy aims to predict inhibitor binding to the metal-containing active site of a metalloenzyme and establish structure-activity relationships. Next, the uses of isosteres, which are molecules that seek to mimic certain molecular functionalities while modulating other properties that may be advantageous are discussed in the context of metalloenzyme FBDD. Here, a subset of carboxylic acid isosteres are evaluated for their ability to inhibit influenza endonuclease and modulate the pharmacological properties of the parent carboxylic acid-containing molecule. Next, the role of three-dimensionality in FBDD is discussed, and a library of metal complexes that is highly three-dimensional is evaluated against several protein targets. Finally, recent approaches to develop a thoroughly representative 3D library from existing compounds are discussed. Overall, this thesis aims to show how concepts in inorganic coordination chemistry can augment FBDD.

Chapter 1: Library Design in Fragment-based Drug Discovery

1.1 Introduction to Fragment-based Drug Discovery

The pharmaceutical industry has largely approached drug discovery via high-throughput screening (HTS). HTS employs large libraries of molecules (up to 10^6 or more) that are screened against biomolecular targets of interest. Once a 'hit' is identified, further development is required to improve potency, selectivity, physicochemical (PC), and pharmacokinetic (PK) properties. However, due to the types of molecules that are typically screened in HTS campaigns, the compounds that are identified do not always represent ideal starting points for drug development.[1] As an example, hits are often highly lipophilic, which impedes the drug development process due to poor aqueous solubility. Similarly, the often large, complex molecules that are identified via HTS can be synthetically difficult to manipulate and do not always lend themselves easily to late-stage functionalization, resulting in longer, more expensive hit-to-lead (HTL) campaigns. In an attempt to address some of the challenges associated with traditional HTS, several specialized drug discovery techniques have been developed, such as NMR-based screening, structure-based screening (i.e., macromolecular X-ray crystallography), and proteome-wide drug screening, to name a few.[2-4] Each of these techniques has its own strengths and drawbacks (e.g., throughput, information content, etc.) that is governed by many factors. Among these approaches, fragment-based drug discovery (FBDD) has emerged as a powerful and efficient small molecule drug discovery alternative to HTS.[5] Unlike HTS, FBDD relies on libraries that are much smaller in size – hundreds to thousands of molecules – rather than the hundreds of thousands or millions typical of HTS. FBDD libraries can be screened against targets of interest to identify low-affinity fragment hits. These hit molecules may be grown, merged, or linked together to form lead compounds with higher affinity and selectivity (Figure 1.1). Additionally, just as many drug-like molecules are often described as obeying the "rule of five,"[6] which predisposes molecules to

possess favorable qualities for oral bioavailability, successful fragment molecules often obey an

analogous "rule of three." The rule of three states that useful fragment molecules should have less

than three hydrogen bond donors, a molecular mass of less than 300 daltons, and an octanol-water

partition coefficient (LogP) of less than three.[7] The rule of three provides guidelines for library

design that will result in fragments that are able to make good binding interactions with the protein,

are small enough to not make unfavorable interactions (steric or electronic), and can be elaborated

into more drug-like molecules.[4]
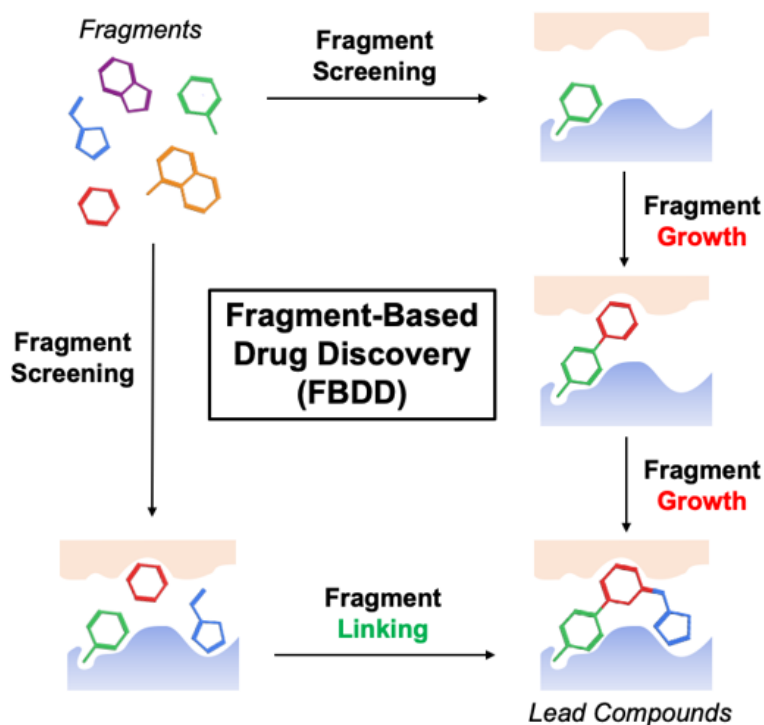
Figure 1.1. Fragment-based drug discovery. Typically, a fragment library (*top left*) is screened against a target of interest. Once a hit is identified, the fragment can be grown (*right*) into a lead compound by introducing substituents that are able to form specific interactions with the target (blue and beige regions). Sometimes, multiple hits in close proximity are identified and fragment linking approaches (*bottom*) can be used to combine these fragments together to create a lead compound with higher target affinity.

One distinct advantage unique to FBDD is its ability to cover a greater representation of chemical space, due the smaller size restriction placed upon fragment molecules. It is estimated that there are 10^{60-200} druglike molecules within the acceptable size limits of drug-like compounds, an amount that is impossible to thoroughly sample with even the largest HTS libraries.[8-9] However, due to the smaller nature of fragments, it is estimated that the total number of possibilities that obey the rule of three is $\sim 10^7$ total molecules, which can be more systematically sampled.[10] Using a combination of techniques, including X-ray crystallography, molecular modeling, and biochemical/biophysical assays to quantify inhibition, fragments can be elaborated into full-length inhibitors through iterative cycles of refinement.

1.2 Library Design in Fragment-based Drug Discovery

Library design is of utmost importance when assembling a fragment library, and the success of FBDD campaigns depends on the careful selection of the library constituents.[11] Additionally, library design has evolved as a reflection of considerations relevant to both screening conditions and HTL development. It has been noted that despite some of the advantages offered by FBDD, it is not necessarily faster than HTS, and consequently, many efforts have been devoted towards optimizing elements of library design, including relevant screening technologies, available fragment collections, and the accessible chemistry of the fragment molecules.[11-14] Although too numerous to cover comprehensively here, a myriad of different strategies have been applied toward library design in recent years in an effort to select for the best fragments that can be grown and optimized into lead-like inhibitors. Some of these considerations include the previously mentioned rule of three,[7] appropriate size distribution and complexity,[15] synthetically accessible growth vectors for development,[16] and avoiding compounds that result in aggregation,

reactivity, or false positive screening results,[17] among others. It has been demonstrated that augmented libraries can drastically improve early phase drug discovery,[11] and the development of novel libraries can address specific gaps in available technologies.

1.3 Using Inorganic Chemistry to Accelerate FBDD

The goal of this thesis is to investigate novel methodologies that can address limitations associated with FBDD. Here, two distinct applications are discussed in detail. While unique in their approach, both methods are grounded in fundamental principles of inorganic chemistry to augment library design for FBDD.

First, Section 1.4 of this chapter introduces metalloenzyme FBDD, which has been pioneered in the Cohen group, and discusses the unique considerations that are relevant when designing fragment libraries to target metalloenzymes. Section 1.5 discusses the use of isosteres in medicinal chemistry, and how their incorporation into metalloenzyme FBDD libraries may overcome specific challenges pertinent to inhibitor development.

Next, fragment shape, specifically three-dimensionality, is introduced as an important consideration in FBDD (Section 1.6). While this approach is unrelated to metalloenzyme inhibition, it is another important consideration in FBDD that can be directly addressed by inorganic chemistry. Section 1.7 expounds upon this idea by introducing metal complexes as highly three-dimensional (3D) molecules that are uniquely suited to create topologically rich fragment libraries.

Taken together, these two approaches, both stemming from the utilization of principles of inorganic chemistry or inherit features of inorganic coordination complexes, create novel avenues of study that can address important problems in FBDD library design.

1.4 Library Design for Metalloenzyme FBDD

It has been estimated that nearly 50% of the proteome is composed of metal-dependent proteins (i.e., metalloproteins).[18] These proteins carry out a wide variety of functions across many different protein classes. The role of the metal ion in metalloproteins can be broadly divided between two main categories: structure and function.[19] Structural metal ion cofactors aid in protein folding, while functional metal ions aid in performing a range of functions, including catalysis and electron transfer. In addition to performing many critical roles in normal physiology, metalloenzymes are also implicated in a range of different pathologies. Due to the many processes controlled and regulated by metalloenzymes, they are attractive targets in drug discovery campaigns. However, despite comprising a large percentage of the proteome, metalloenzymes are underrepresented as targets of drugs approved by the Food and Drug Administration (FDA).[20] Many metalloenzymes present one or more metal ions at the enzyme active site, which are central for catalytic function. One strategy to inhibit these metalloenzymes relies on the use of inhibitors possessing Lewis basic atoms to form a coordinate covalent bond to the metal ions in the active site. If the inhibitor molecule possesses sufficient affinity for the active site/metal, it can block the enzyme from performing its desired function. This approach is one of the central hypotheses this thesis. For many years, hydroxamic acids were used almost exclusively as the metal-binding group for many therapeutics that targeted a range of metalloenzymes, despite the pharmacological disadvantages they present.[20] In recent years, the Cohen lab has addressed the overreliance on hydroxamic acids as by developing novel metal binding groups, also known as metal binding pharmacophores (MBPs). In addition to introducing novel MBPs, which are capable of binding to a variety of different metal ions, the Cohen group has also developed MBPs that demonstrate

some selectivity for therapeutically relevant metalloenzymes,[21] display improved pharmacokinetic properties,[22] and move beyond the narrow scope of hydroxamic acid inhibitors (Figure 1.2).
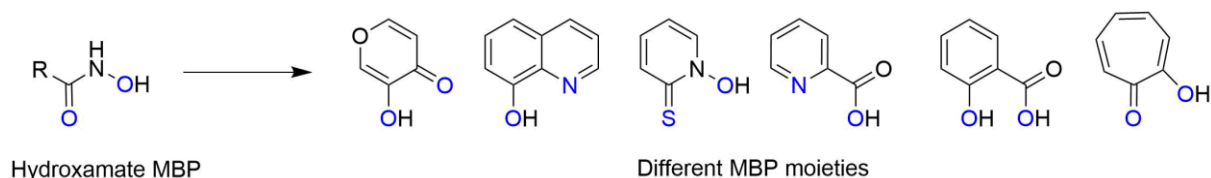
Figure 1.2. Several MBPs used by the Cohen group for FBDD that move beyond the scope of hydroxamic acid-based inhibitors. Donor atoms are highlighted in blue.

Work in the Cohen group has introduced MBP libraries that can produce hits against therapeutically relevant metalloenzyme targets.[23-24] The small size (<300 Daltons) and high affinity of MBPs for cationic metal ions in the active site of metalloenzymes makes them ideal starting points for FBDD. During the HTL process in FBDD an MBP library can be biochemically screened against a protein target utilizing techniques such as ^{19}F NMR, colorimetric, fluorescence, and fluorescence polarization-based assays.[21, 25-26] Due to the Lewis basic properties of the MBP, there are strong interactions between the cationic active site metal(s) and the MBP fragment hit. This allows the binding site of the fragment to be readily assumed at the metal-containing active site. In conventional FBDD, which does not always rely on a singular dominant interaction such as this, the binding location is not as easily predicted. However, once an MBP fragment hit is identified, additional efforts can be made to elucidate and confirm the binding mode of the fragment hit. Some strategies to more precisely determine and validate the binding mode include protein X-ray crystallography, computational modeling, and the use of inorganic model complexes, which can used to recapitulate metalloenzyme active sites.[27-29] Once the binding mode is understood, it can be optimized to coordinate the metal ion(s) and facilitate growth vectors to explore structure-activity-relationships (SAR) between the inhibitor and the protein. These

relationships can be further developed through iterative synthesis combined with computational and biophysical techniques. Through iterative cycles of refinement, highly potent and selective metalloenzyme inhibitors can be developed.[27, 30-31] Once good enzyme inhibition activity is achieved, lead compounds which meet additional downstream development criteria, will be selected. Some of these criteria include stability, selectivity, and good absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties (Figure 1.3).
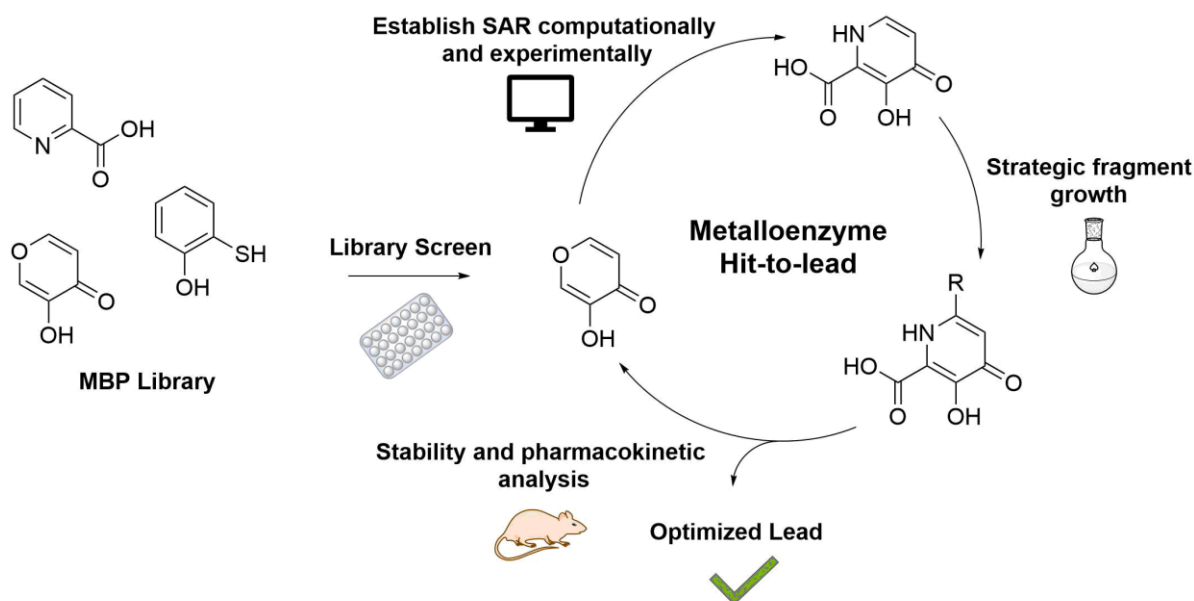
Figure 1.3. Scheme demonstrating HTL against metalloenzyme targets. Once a hit is identified, biophysical and computational methods can be employed to determine the binding pose of the fragment. Later, using the identified SAR, strategic additions can be employed to make additional protein-ligand contacts and improve inhibition. Once an inhibitor with high potency is identified, pharmacokinetic analysis can inform the suitability of the drug molecule in biological systems. The process can iterate many times until an optimized lead molecule with sufficient inhibitory activity and a satisfactory pharmacokinetic profile is identified. Metalloenzyme HTL follows the same principles applied in conventional HTL development, except that the library is specifically designed to provide coordination to the active site metal(s).

1.5 Isosteric Replacement

Isosteric replacement is a common strategy in medicinal chemistry; however, the development of libraries and sub-libraries with isosteric functionalities targeting metalloenzyme inhibition is a recent development.[32] The classical definition of an isostere dates back to 1919, when Irving Langmuir described isosteres as an atom (or group of atoms) that shares the same number and arrangement of electrons with another atom or group of atoms (for example, N_2 and CO).[33-34] The more modern definition of isosteres (sometimes referred to as bioisosteres) traces its roots back to work in the 1930s by Hans Erlenmeyer, who demonstrated that antibodies were unable to differentiate between similar groups of atoms in artificial antigens.[35] This latter example better exemplifies the modern definition of (bio)isosteres, which focuses on identifying groups that can elicit a similar biological response, whose physical and electronic structure may or may not resemble that of the original group. Isosteres are commonly used as a method to mitigate pharmacological liabilities while preserving the relevant biological interaction/function (Figure 1.4) and are typically strategically introduced later in the drug development process as PK liabilities are identified. Because isosteres strive to mimic biological function rather than exact structural features, the use and efficacy of a particular isostere will be context dependent.[35-36]
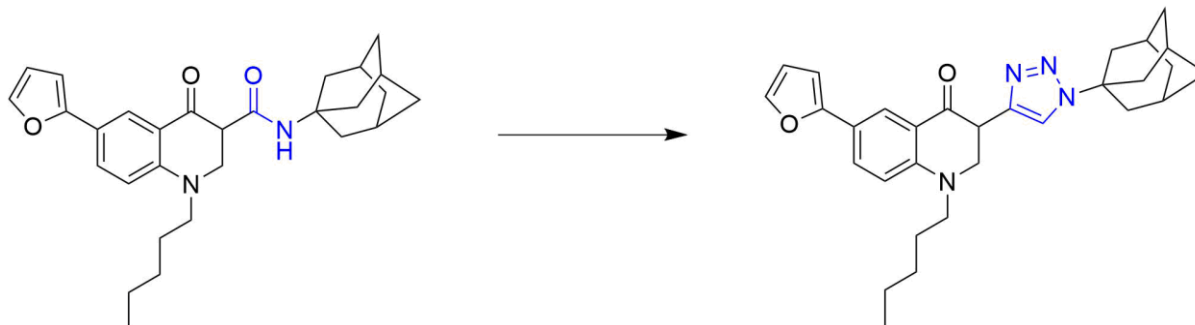
Figure 1.4. Representation of isosteric replacement. In this example, an amide group was replaced with a triazole (both highlighted in blue). This isosteric replacement resulted in comparable inhibitory activity with the parent molecule against the cannabinoid-2 receptor, which is implicated in cancer, inflammation, and neuropathy, and avoids the psychoactive effects of activation of the cannabinoid-1 receptor.[37] Here, the triazole-containing molecule demonstrated improved selectivity and solubility.[38-39]

Recent work in our own laboratory has sought to introduce the concept of isosteric replacement to metalloenzyme FBDD (Figure 1.5).[22, 32, 40] In the case of metalloenzyme library development, the Lewis basic atoms that contribute to metal-binding do not always present ideal PK properties. Isosteric replacement of MBPs is an emerging strategy that can aid in the development of new MBPs that are capable of making strong interactions with metal ions while distinguishing themselves from traditional MBPs with their improved PK properties. These metal-binding isosteres (MBIs) have the potential to address problems that may arise during metalloenzyme FBDD while maintaining the critical coordination necessary for inhibition. The screening of libraries of isosteres based on MBPs with known pharmacological liabilities will accelerate the development process of metalloenzyme inhibitors. Chapter 3 details the synthesis, biological, and crystallographic analysis of isosteres of the potent hydroxypyridinone fragment against a viral metalloenzyme target.
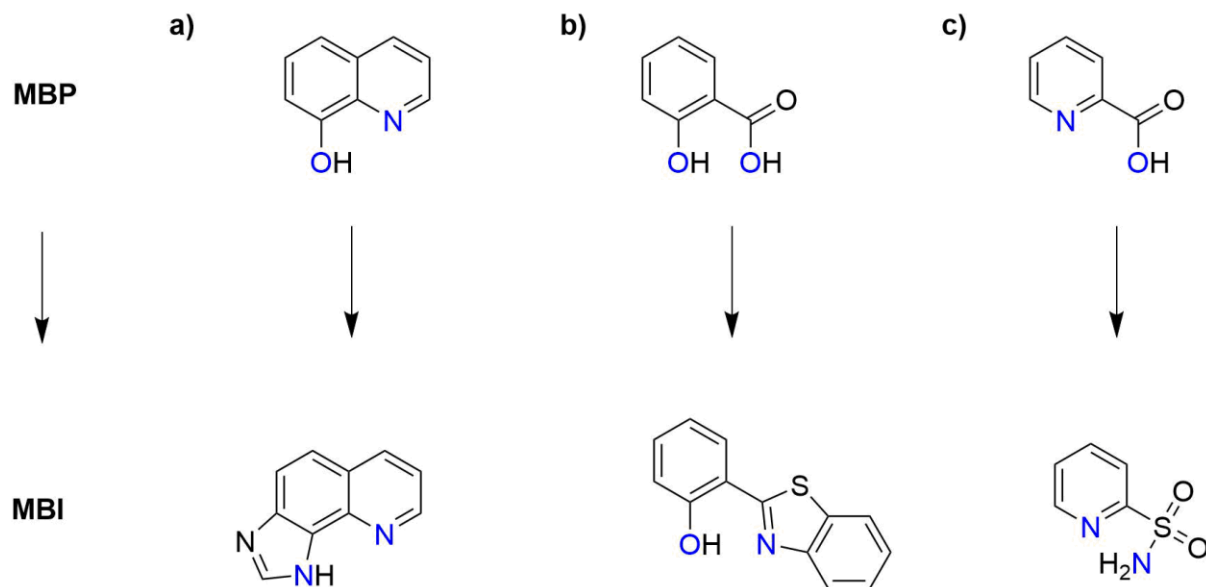
Figure 1.5. Recent MBI examples from the Cohen Lab. (a) 8-hydroxy quinoline,[22] (b) salicylic acid,[40] and (c) picolinic acid[32] have all been successfully developed into MBIs. Despite modifying the coordinating group, each of these MBIs demonstrated good inhibitory activity in enzymatic assays. Donor atoms are highlighted in blue.

1.6 3D Library Design

Though FBDD has helped to address some of the challenges associated with HTS, it is not without its shortcomings. One important limitation that hinders FBDD is a bias towards flat, small fragment molecules. The overwhelming planarity of the fragments typically used in FBDD was highlighted by a study that analyzed 18,534 conventional fragments according to their shape. The study revealed that the overwhelming majority of these fragments (~75%) showed mostly linear or planar geometries.[41] This predisposition towards linear and planar geometries, especially in the context of FBDD, limits the final geometries of the final, full-length inhibitor.

It has been shown that broad shape diversity in a library correlates with broad biological activity.[42] Consequently, a host of libraries that have attempted to increase the 3D quality of fragments have emerged in recent years. While not an exhaustive list, some of these recent approaches to improve three-dimensionality include diversity-oriented synthesis,[41] the

11

introduction of chiral carbon centers,[43] the incorporation of cubanes,[44] and the use of combinatorial libraries.[42] While many of these methods do increase three-dimensionality, they often require difficult synthetic routes, which limits their use as library components (Figure 1.6). Due to the potential benefits of 3D library design, and the lack of scalable, robust methods to achieve broad 3D representation, increased research into 3D fragment libraries will aid in the development of novel inhibitors.
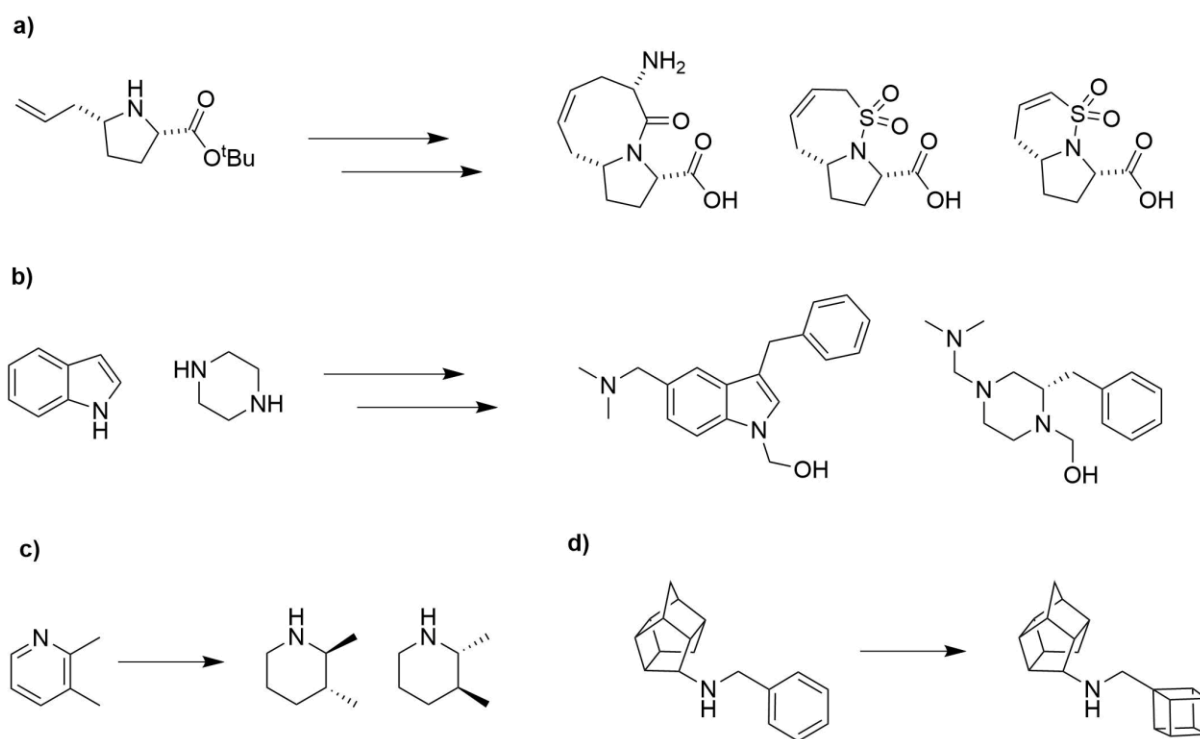
Figure 1.6. Improving 3-dimensionality in fragment libraries. (a) Approach demonstrating diversity-oriented synthesis.[41] From simple proline-based derivatives, fragment libraries with diverse stereochemistry and ring sizes/configurations can be generated. (b) Approach demonstrating combinatorial libraries.[42] By combining a library of scaffolds (*left*) with another library of building blocks (-Ph, -OH, -N(CH$_3$)$_2$) a large library of diverse molecules can be created. (c) Saturation has been used as a metric of molecular complexity, which facilitates higher 3-dimensionality without a significant increase in molecular weight.[43]   (d) The incorporation of cubanes onto inhibitor molecules has also been shown to improve 3-dimensionality. In the selected example, replacing the phenyl functional group with cubane resulted in an improvement in inhibitory activity against σ$_1$ and σ$_2$ receptors due to the increased steric bulk it afforded.[44-45]

1.7 Metal Complexes as 3D Fragments

As described above, molecular shape is an important consideration in library development. Because biomolecular recognition is often driven by protein–ligand interactions, biological macromolecules will bind to molecules (including their native binding partners) with complementary 3D binding surfaces.[46] Thus, many research efforts have been devoted toward molecules with high structural and topological diversity.[41, 46] It has been demonstrated that increasing the three-dimensionality of molecules can lead to an increased likelihood of clinical success.[15] For example, 3D molecules exhibit higher solubility due to increased solvation and diminished solid state crystal lattice packing.[47] Additionally, other ADMET properties have been shown to be influenced by molecular topology.[48] Other metrics of three-dimensionality, such as the fraction of sp^3 hybridized carbons, have also been studied as methods of improving drug-like properties.[43, 49] Despite the advantages of highly 3D molecules, modern drug discovery relies heavily on 2D and planar structures for both fragment and lead-like molecules.[47, 49] Although there have been many recent efforts to design libraries comprised of highly 3D molecules for drug discovery purposes,[41, 43, 50-51] most drug-like molecules trend toward planarity.[49]

Organic molecules are typically constructed from linear, planar, or tetrahedral building blocks, primarily driven by the hybridization, valency, and geometry of carbon atoms (Figure 1.7A). However, metal centers offer the opportunity to utilize building blocks with increased valency and more varied geometries (Figure 1.7). For example, an octahedral metal center with six different substituents is capable of forming 30 stereoisomers (compared to the two possible isomers formed by chiral carbon centers).[52] It has been demonstrated that binding selectivity toward biomolecules correlates with shape and stereochemical complexity[53] and the simple but

well-defined, rigid shapes that can be found in some metal complexes may possess the complexity necessary to achieve high protein-binding specificity.[54]
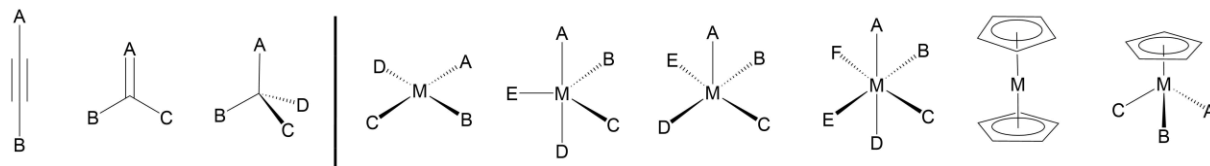
Figure 1.7. Geometries of carbon- and metal-containing complexes. (A) Linear, planar, and tetrahedral geometry of carbon centers. (B) Square planar, trigonal bipyramidal, square pyramidal, octahedral, sandwich, and half-sandwich geometries of metal complexes

In a series of seminal studies, Meggers and coworkers demonstrated that organometallic complexes are capable of acting as highly selective protein kinase inhibitors.[55-58] Ru(II) and Ir(III) pyridocarbazole metal complexes were derived from the natural product staurosporine, which is a relatively nonspecific protein kinase inhibitor.[59-60] The resulting metal complexes, termed octasporines (**OS**), contained an organic component related to staurosporine that could also function as a bidentate ligand to a metal center. The staurosporine derivatives were bound to different metal complexes, which allowed for the introduction of new ligands and molecular geometries/shapes. These inert metal complexes presented well-defined molecular surfaces that aided in the identification of new interactions that could contribute to binding strength and selectivity toward kinases. As shown in Figure 1.8a, **Λ-OS1** demonstrates a highly globular and complimentary molecular surface to the active site of the β-isoform of glycogen synthase kinase 3 (GSK3β). **Λ-OS1** occupies the ATP-binding site, similar to staurosporine (Figure 1.8b) and other organic inhibitors.[61-62] In addition to a range of well-defined hydrogen bonds, hydrophobic interactions, and edge-to-face stacking, the CO ligand of the metal complex interacts with a flexible glycine-rich loop caused by a conformational change, which is not possible with the

natural product staurosporine. This example highlights how metal complexes with diverse geometries can occupy regions of chemical and topological space that cannot be readily accessed by conventional organic compounds,[63-64] and demonstrates the potential of octahedral metal complexes to act as enzyme inhibitors with highly complementary binding surfaces.
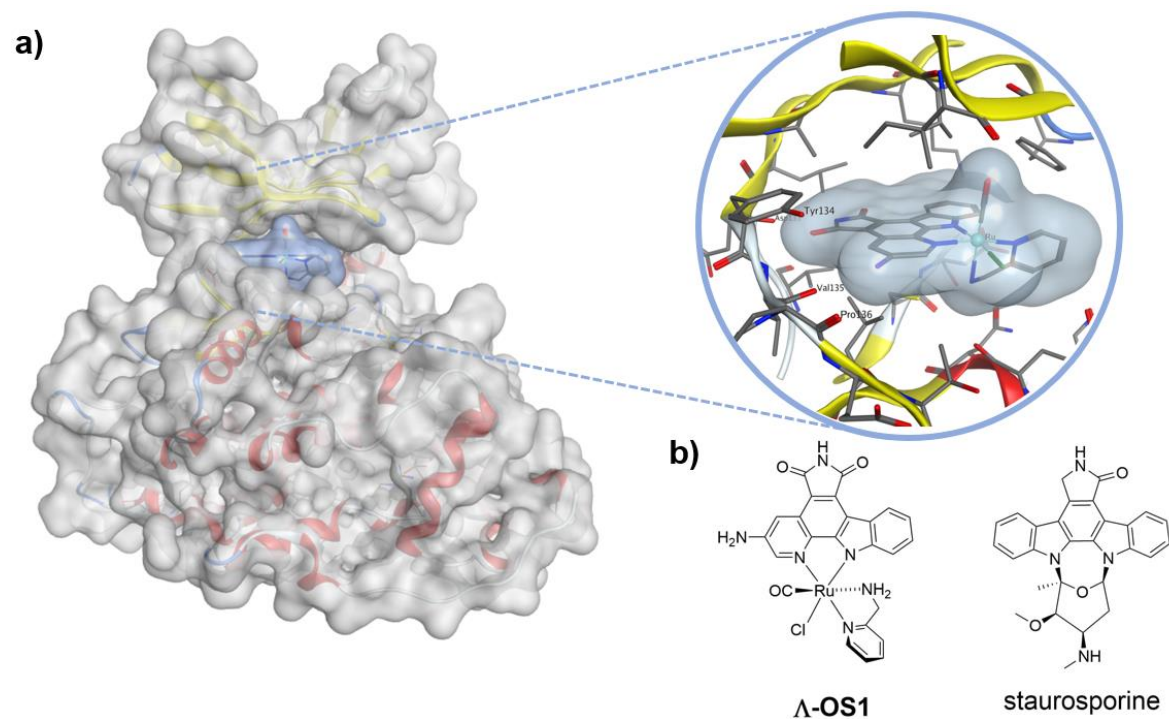
Figure 1.8. (a) X-ray crystal structure of **Λ-OS1** bound to GSK3β. **Λ-OS1** shows a highly complementary molecular surface that is able to form a novel interaction with the glycine-rich loop via an induced-fit binding mode. (b) Structures of **Λ-OS1** and staurosporine. PDB accession codes are 3PUP and 1Q3D for protein-bound **Λ-OS1** and staurosporine, respectively.[58]

This work by Meggers and coworkers demonstrates the capability of metal complexes to serve as highly 3D inhibitors. These effective 'metalloisosteres' served as the proof of concept for the work that will be discussed in Chapters 5 and 6, which explores the various geometries and high three-dimensionality of metal complexes in FBDD. These properties make metal complexes attractive candidates for a library of topologically diverse fragments, which can be composed of

trigonal bipyramidal, square pyramidal, octahedral, sandwich, and half-sandwich geometries (Figure 1.7).[63, 65] These coordination compounds can serve as useful scaffolds in the creation of unique shapes and geometries that can selectively engage enzyme targets.

1.8 Scope of this Dissertation

This dissertation focuses on establishing new approaches for library design in FBDD. Each method outlined below addresses a particular challenge associated with FBDD. These approaches facilitate the creation of libraries that can be utilized for specific purposes to aid in the identification of hit molecules that may be useful in drug discovery campaigns.

Chapter 2 describes a novel workflow to computationally predict the binding pose of MBPs and their elaborated inhibitors bound to metalloenzymes. This methodology enables the creation and use of virtual MBP libraries that can be screened, elaborated, and experimentally evaluated.

Chapter 3 details the creation of a small sub-library of MBIs and demonstrates the influence of the isosteric group on the PK properties of each molecule. Additionally, the binding mode of select MBIs was evaluated when bound to the protein target, revealing new information about the coordination modes of these MBIs.

Chapter 4 describes an analysis of the topologies of molecules in the Protein Data Bank (PDB) and the DrugBank. This analysis highlights the lack of 3D diversity in these relevant databases, and explores the advantages of libraries with high topological diversity.

Chapter 5 details the creation of a library of metal complexes that can be used in FBDD. Due to the geometries of these 'metallofragments,' they can occupy 3D regions of chemical space that have been difficult to achieve using organic molecules alone. Here, their utility as fragments is demonstrated in a screen against severally therapeutically relevant targets.

Chapter 6 describes progress towards the creation of a highly 3D virtual library of metallofragments whose components will consist of structures mined from the Cambridge Structural Database (CSD). Filters will be applied to select for metal complexes that are stable in biological systems and which demonstrate variable molecular topologies. This library will provide comprehensive coverage of chemical space and will be virtually (and eventually experimentally) screened against a wide range of therapeutically relevant protein targets.

1.9 Acknowledgments

Chapter 1, in part, is a reprint of the material as it appears in "Metal complexes for therapeutic applications." *Trends Chem.,* **2021,** *3*, 523. The dissertation author was a primary author of this paper and gratefully acknowledges the contributions of coauthors Johannes Karges and Seth M. Cohen.

1.10 References

1.   Scott, D. E.; Coyne, A. G.; Hudson, S. A.; Abell, C., Fragment-Based Approaches In Drug Discovery And Chemical Biology. *Biochemistry* **2012,** *51* (25), 4990-5003.

2.   Pellecchia, M.; Sem, D. S.; Wüthrich, K., Nmr in drug discovery. *Nat. Rev. Drug Discov.* **2002,** *1* (3), 211-219.

3.   Maveyraud, L.; Mourey. L. Protein X-ray Crystallography and Drug Discovery *Molecules*, **2020**, 25 (5), 1030.

4.   Zhou, Y.; Liu, Z.; Rothschild, K. J.; Lim, M. J., Proteome-wide drug screening using mass spectrometric imaging of bead-arrays. *Sci. Rep.* **2016,** *6* (1), 26125.

5. Jencks William, P., On the attribution and additivity of binding energies. *Proc. Natl. Acad. Sci. U. S. A.* **1981,** *78* (7), 4046-4050.

6. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental And computational Approaches To Estimate Solubility And Permeability In Drug Discovery And Development Settings. *Adv. Drug Deliv. Rev.* **2001,** *46* (1), 3-26.

7. Congreve, M.; Carr, R.; Murray, C.; Jhoti, H., A 'Rule of Three' For Fragment-Based Lead Discovery? *Drug Discov. Today* **2003,** *8* (19), 876-877.

8. Bohacek, R. S.; McMartin, C.; Guida, W. C., The art and practice of structure-based drug design: a molecular modeling perspective. *Med. Res. Rev.* **1996,** *16* (1), 3-50.

9. Ertl, P., Cheminformatics Analysis of Organic Substituents: Identification of the Most Common Substituents, Calculation of Substituent Properties, and Automatic Identification of Drug-like Bioisosteric Groups. *J. Chem. Inf. and Comput. Sci.* **2003,** *43* (2), 374-380.

10. Fink, T.; Reymond, J.-L., Virtual Exploration of the Chemical Universe up to 11 Atoms of C, N, O, F: Assembly of 26.4 Million Structures (110.9 Million Stereoisomers) and Analysis for New Ring Systems, Stereochemistry, Physicochemical Properties, Compound Classes, and Drug Discovery. *J. Chem. Inf. Model.* **2007,** *47* (2), 342-353.

11. Troelsen, N. S. C., M. H., Library Design Strategies To Accelerate Fragment-Based Drug Discovery. *Chem. Eur. J.* **2020,** *26*, 11391-11403.

12. Erlanson, D. A.; Fesik, S. W.; Hubbard, R. E.; Jahnke, W.; Jhoti, H., Twenty years on: the impact of fragments on drug discovery. *Nat. Rev. Drug Discov.* **2016,** *15* (9), 605-619.

13. Murray, C. W.; Rees, D. C., Opportunity Knocks: Organic Chemistry for Fragment-Based Drug Discovery (FBDD). *Angew. Chem. Int. Ed.* **2016,** *55* (2), 488-492.

14.   Keserű, G. M.; Erlanson, D. A.; Ferenczy, G. G.; Hann, M. M.; Murray, C. W.; Pickett, S. D., Design Principles for Fragment Libraries: Maximizing the Value of Learnings from Pharma Fragment-Based Drug Discovery (FBDD) Programs for Use in Academia. *J. Med. Chem.* **2016,** *59* (18), 8189-8206.

15.   Kombo, D. C.; Tallapragada, K.; Jain, R.; Chewning, J.; Mazurov, A. A.; Speake, J. D.; Hauser, T. A.; Toler, S., 3D Molecular Descriptors Important For Clinical Success. *J. Chem. Inf. Model.* **2013,** *53* (2), 327-342.

16.   Palmer, N.; Peakman, T. M.; Norton, D.; Rees, D. C., Design and synthesis of dihydroisoquinolones for fragment-based drug discovery (FBDD). *Org. & Biomol. Chem.* **2016,** *14* (5), 1599-1610.

17.   Davis, B. J.; Erlanson, D. A., Learning from our mistakes: The 'unknown knowns' in fragment screening. *Bioorg. Med. Chem. Lett.* **2013,** *23* (10), 2844-2852.

18.   Waldron, K. J. R., J. C.; Ford, D.; Robinson, N. J., Metalloproteins and Metal Sensing. *Nature* **2009,** *460*, 823-830.

19.   Chen, A. Y.; Adamek, R. N.; Dick, B. L.; Credille, C. V.; Morrison, C. N.; Cohen, S. M., Targeting metalloenzymes for therapeutic intervention. *Chem. Rev.* **2019,** *119* (2), 1323-1455.

20.   Cohen, S. M., A Bioinorganic Approach to Fragment-Based Drug Discovery Targeting Metalloenzymes. *Acc. Chem. Res.* **2017,** *50* (8), 2007-2016.

21.   Adamek, R. N.; Suire, C. N.; Stokes, R. W.; Brizuela, M. K.; Cohen, S. M.; Leissring, M. A., Hydroxypyridinethione Inhibitors of Human Insulin-Degrading Enzyme. *ChemMedChem* **2021,** *16* (11), 1776-1788.

22.   Seo, H.; Jackl, M. K.; Kalaj, M.; Cohen, S. M., Developing Metal-Binding Isosteres of 8-Hydroxyquinoline as Metalloenzyme Inhibitor Scaffolds. *Inorg. Chem.* **2022,** *61* (19), 7631-7641.

23.   Agrawal, A.; Johnson, S. L.; Jacobsen, J. A.; Miller, M. T.; Chen, L.-H.; Pellecchia, M.; Cohen, S. M., Chelator Fragment Libraries for Targeting Metalloproteinases. *ChemMedChem* **2010,** *5* (2), 195-199.

24.   Jacobsen, J. A.; Fullagar, J. L.; Miller, M. T.; Cohen, S. M., Identifying Chelators for Metalloprotein Inhibitors Using a Fragment-Based Approach. *J. Med. Chem.* **2011,** *54* (2), 591-602.

25.   Prosser, K. E.; Kohlbrand, A. J.; Seo, H.; Kalaj, M.; Cohen, S. M., 19F-Tagged metal binding pharmacophores for NMR screening of metalloenzymes. *Chem. Comm.* **2021,** *57* (40), 4934-4937.

26.   Morrison, C. N.; Prosser, K. E.; Stokes, R. W.; Cordes, A.; Metzler-Nolte, N.; Cohen, S. M., Expanding Medicinal Chemistry Into 3D space: Metallofragments As 3D Scaffolds For Fragment-Based Drug Discovery. *Chem. Sci.* **2020,** *11* (5), 1216-1225.

27.   Credille, C. V.; Dick, B. L.; Morrison, C. N.; Stokes, R. W.; Adamek, R. N.; Wu, N. C.; Wilson, I. A.; Cohen, S. M., Structure–Activity Relationships in Metal-Binding Pharmacophores for Influenza Endonuclease. *J. Med. Chem.* **2018,** *61* (22), 10206-10217.

28.   Karges, J.; Stokes, R. W.; Cohen, S. M., Computational Prediction of the Binding Pose of Metal-Binding Pharmacophores. *ACS Med. Chem. Lett.* **2022,** *13* (3), 428-435.

29.   Puerta, D. T.; Cohen, S. M., Elucidating Drug-Metalloprotein Interactions with Tris(pyrazolyl)borate Model Complexes. *Inorg. Chem.* **2002,** *41* (20), 5075-5082.

30.   Credille, C. V.; Chen, Y.; Cohen, S. M., Fragment-Based Identification of Influenza Endonuclease Inhibitors. *J. Med. Chem.* **2016,** *59* (13), 6444-6454.

31.   Credille, C. V.; Morrison, C. N.; Stokes, R. W.; Dick, B. L.; Feng, Y.; Sun, J.; Chen, Y.; Cohen, S. M., SAR Exploration of Tight-Binding Inhibitors of Influenza Virus PA Endonuclease. *J. Med. Chem.* **2019,** *62* (21), 9438-9449.

32.   Dick, B. L.; Cohen, S. M., Metal-Binding Isosteres as New Scaffolds for Metalloenzyme Inhibitors. *Inorg. Chem.* **2018,** *57* (15), 9538-9543.

33.   Langmuir, I., ISOMORPHISM, ISOSTERISM AND COVALENCE. *J. Am. Chem. Soc.* **1919,** *41* (10), 1543-1559.

34.   Patani, G. A.; LaVoie, E. J., Bioisosterism: A Rational Approach in Drug Design. *Chem. Rev.* **1996,** *96* (8), 3147-3176.

35.   Meanwell, N. A., Synopsis of Some Recent Tactical Application of Bioisosteres in Drug Design. *J. Med. Chem.* **2011,** *54* (8), 2529-2591.

36.   Lassalas, P.; Gay, B.; Lasfargeas, C.; James, M. J.; Tran, V.; Vijayendran, K. G.; Brunden, K. R.; Kozlowski, M. C.; Thomas, C. J.; Smith, A. B.; Huryn, D. M.; Ballatore, C., Structure Property Relationships of Carboxylic Acid Isosteres. *J. Med. Chem.* **2016,** *59* (7), 3183-3203.

37.   Cabral, G. A.; Griffin-Thomas, L., Emerging role of the cannabinoid receptor CB2 in immune regulation: therapeutic prospects for neuroinflammation. *Expert Reviews in Molecular Medicine* **2009,** *11*, e3.

38.   Mugnaini, C.; Nocerino, S.; Pedani, V.; Pasquini, S.; Tafi, A.; De Chiaro, M.; Bellucci, L.; Valoti, M.; Guida, F.; Luongo, L.; Dragoni, S.; Ligresti, A.; Rosenberg, A.; Bolognini, D.; Cascio, M. G.; Pertwee, R. G.; Moaddel, R.; Maione, S.; Di Marzo, V.; Corelli, F., Investigations on the 4-Quinolone-3-Carboxylic Acid Motif Part 5: Modulation of the Physicochemical Profile of a Set of Potent and Selective Cannabinoid-2 Receptor Ligands through a Bioisosteric Approach. *ChemMedChem* **2012,** *7* (5), 920-934.

39.   Kumari, S.; Carmona, A. V.; Tiwari, A. K.; Trippier, P. C., Amide Bond Bioisosteres: Strategies, Synthesis, and Successes. *J. Med. Chem.* **2020,** *63* (21), 12290-12358.

40.   Jackl, M. K.; Seo, H.; Karges, J.; Kalaj, M.; Cohen, S. M., Salicylate metal-binding isosteres as fragments for metalloenzyme inhibition. *Chem. Sci.* **2022,** *13* (7), 2128-2136.

41.   Hung, A. W.; Ramek, A.; Wang, Y.; Kaya, T.; Wilson, J. A.; Clemons, P. A.; Young, D. W., Route To Three-Dimensional Fragments Using Diversity-Oriented Synthesis. *Proc. Natl. Acad. Sci.* **2011,** *108* (17), 6799-6804.

42.   Sauer, W. H. B.; Schwarz, M. K., Molecular Shape Diversity Of Combinatorial Libraries: A Prerequisite For Broad Bioactivity. *J. Chem. Inf. Model.* **2003,** *43* (3), 987-1003.

43.   Lovering, F.; Bikker, J.; Humblet, C., Escape From Flatland: Increasing Saturation As An Approach To Improving Clinical Success. *J. Med. Chem.* **2009,** *52* (21), 6752-6756.

44.   Reekie, T. A.; Williams, C. M.; Rendina, L. M.; Kassiou, M., Cubanes in medicinal chemistry. *J. Med. Chem.* **2019,** *62* (3), 1078-1095.

45.   Banister, S. D.; Manoli, M.; Barron, M. L.; Werry, E. L.; Kassiou, M., N-substituted 8-aminopentacyclo[5.4.0.02,6.03,10.05,9]undecanes as σ receptor ligands with potential neuroprotective effects. *Bioorg. Med. Chem. Lett.* **2013,** *21* (19), 6038-6052.

46.   Galloway, W. R. J. D.; Isidro-Llobet, A.; Spring, D. R., Diversity-Oriented Synthesis As A Tool For The Discovery Of Novel Biologically Active Small Molecules. *Nat. Commun.* **2010,** *1* (1), 80.

47.   Firth, N. C.; Brown, N.; Blagg, J., Plane Of Best Fit: A Novel Method To Characterize The Three-Dimensionality Of Molecules. *J. Chem. Inf. Model.* **2012,** *52* (10), 2516-2525.

48.   Yang, Y.; Engkvist, O.; Llinàs, A.; Chen, H., Beyond Size, Ionization State, And Lipophilicity: Influence Of Molecular Topology On Absorption, Distribution, Metabolism, Excretion, And Toxicity For Druglike Compounds. *J. Med. Chem.* **2012,** *55* (8), 3667-3677.

49.   Meyers, J.; Carter, M.; Mok, N. Y.; Brown, N., On The Origins Of Three-Dimensionality In Drug-Like Molecules. *Future Med. Chem.* **2016,** *8* (14), 1753-1767.

50.   Leach, A. R.; Gillet, V. J.; Lewis, R. A.; Taylor, R., Three-Dimensional Pharmacophore Methods In Drug Discovery. *J. Med. Chem.* **2010,** *53* (2), 539-558.

51.   Twigg, D. G.; Kondo, N.; Mitchell, S. L.; Galloway, W. R. J. D.; Sore, H. F.; Madin, A.; Spring, D. R., Partially Saturated Bicyclic Heteroaromatics As An Sp3-Enriched Fragment Collection. *Angew. Chem. Int. Ed.* **2016,** *55* (40), 12479-12483.

52.   Meggers, E., Targeting Proteins With Metal Complexes. *Chem. Commun.* **2009,** (9), 1001-1010.

53.   Clemons, P. A.; Bodycombe, N. E.; Carrinski, H. A.; Wilson, J. A.; Shamji, A. F.; Wagner, B. K.; Koehler, A. N.; Schreiber, S. L., Small Molecules Of Different Origins Have Distinct Distributions Of Structural Complexity That Correlate With Protein-Binding Profiles. *Proc. Natl. Acad. Sci.* **2010,** *107* (44), 18787-18792.

54.   Zhang, L.; Carroll, P.; Meggers, E., Ruthenium Complexes As Protein Kinase Inhibitors. *Org. Lett.* **2004,** *6* (4), 521-523.

55.   Bregman, H.; Williams, D. S.; Atilla, G. E.; Carroll, P. J.; Meggers, E., An Organometallic Inhibitor For Glycogen Synthase Kinase 3. *J. Am. Chem. Soc.* **2004,** *126* (42), 13594-13595.

56.   Bregman, H.; Carroll, P. J.; Meggers, E., Rapid Access To Unexplored Chemical Space By Ligand Scanning Around A Ruthenium Center: Discovery Of Potent And Selective Protein Kinase Inhibitors. *J. Am. Chem. Soc.* **2006,** *128* (3), 877-884.

57.   Atilla-Gokcumen, G. E.; Williams, D. S.; Bregman, H.; Pagano, N.; Meggers, E., Organometallic Compounds With Biological Activity: A Very Selective And Highly Potent Cellular Inhibitor For Glycogen Synthase Kinase 3. *ChemBioChem* **2006,** *7* (9), 1443-1450.

58.   Feng, L.; Geisselbrecht, Y.; Blanck, S.; Wilbuer, A.; Atilla-Gokcumen, G. E.; Filippakopoulos, P.; Kräling, K.; Celik, M. A.; Harms, K.; Maksimoska, J.; Marmorstein, R.; Frenking, G.; Knapp, S.; Essen, L.-O.; Meggers, E., Structurally Sophisticated Octahedral Metal Complexes As Highly Selective Protein Kinase Inhibitors. *J. Am. Chem. Soc.* **2011,** *133* (15), 5976-5986.

59.   Tamaoki, T.; Nomoto, H.; Takahashi, I.; Kato, Y.; Morimoto, M.; Tomita, F., Staurosporine, A Potent Inhibitor Of PhospholipidCa^{++} Dependent Protein Kinase. *Biochem. Biophys. Res. Commun.* **1986,** *135* (2), 397-402.

60.   Caravatti, G.; Meyer, T.; Fredenhagen, A.; Trinks, U.; Mett, H.; Fabbro, D., Inhibitory Activity And Selectivity Of Staurosporine Derivatives Towards Protein Kinase C. *Bioorg. Med. Chem. Lett.* **1994,** *4* (3), 399-404.

61.   Bertrand, J. A.; Thieffine, S.; Vulpetti, A.; Cristiani, C.; Valsasina, B.; Knapp, S.; Kalisz, H. M.; Flocco, M., Structural Characterization Of The GSK-3β Active Site Using Selective And Non-Selective ATP-Mimetic Inhibitors. *J. Mol. Biol.* **2003,** *333* (2), 393-407.

62.   Meijer, L.; Skaltsounis, A.-L.; Magiatis, P.; Polychronopoulos, P.; Knockaert, M.; Leost, M.; Ryan, X. P.; Vonica, C. A.; Brivanlou, A.; Dajani, R.; Crovace, C.; Tarricone, C.; Musacchio, A.; Roe, S. M.; Pearl, L.; Greengard, P., GSK-3-Selective Inhibitors Derived From Tyrian Purple Indirubins. *Chem. Biol.* **2003,** *10* (12), 1255-1266.

63.   Meggers, E., Exploring Biologically Relevant Chemical Space With Metal Complexes. *Curr. Opin. Chem. Biol.* **2007,** *11* (3), 287-292.

64.   Davies, C. L.; Dux, E. L.; Duhme-Klair, A.-K., Supramolecular Interactions Between Functional Metal Complexes And Proteins. *Dalton Trans.* **2009,** (46), 10141-10154.

65.   Mjos, K. D.; Orvig, C., Metallodrugs In Medicinal Inorganic Chemistry. *Chem. Rev.* **2014,** *114* (8), 4540-4563.

Chapter 2: Computational Prediction of the Binding Pose of Metal-Binding Pharmacophores

2.1 Introduction

During the course of (metalloenzyme) inhibitor development, atomic resolution structural information is not always available. Fortunately, in the absence of detailed structural information, computational methods may be used to model and predict protein-ligand interactions. Advances in computer software and hardware have made computer-aided drug design (CADD) ubiquitous in drug discovery and development efforts. In addition to providing or supplementing structural information, CADD approaches increase the efficacy of inhibitor development by predicting factors such as toxicity, activity, and efficacy, in many cases, before experiments have been performed in vitro.[1] Among the techniques most relevant to FBDD are virtual fragment screening and fragment growth.

Virtual screening (VS) is the process of using virtual libraries to identify compounds that possess bioactive properties when evaluated against a target of interest. VS can be described as the application of filters that select for molecules which demonstrate predicted biological activity against a target.[2] One of the techniques used in VS methodologies is structure-based virtual screening (SBVS), which uses the experimentally-determined structural information of the target to dock ligands and rank them based on predicted affinity and binding site recognition.[2] When SBVS is considered in the context of FBDD, it may be used to predict relevant initial fragment placement in the active site, and can be used to rationalize fragment hits. Once a fragment hit has been confirmed virtually or experimentally, virtual fragment growth can proceed. This important step seeks to define growth vectors on the fragment, and grow the molecule into a more potent inhibitor while considering the inherent steric factors of the protein (Figure 2.1).[3] This step is often performed iteratively and in tandem with experimental chemical synthesis and biological evaluation. Combining experimental and computational results greatly accelerates the

development process by decreasing the number of experiments that must be performed and by providing additional insight to observed results.
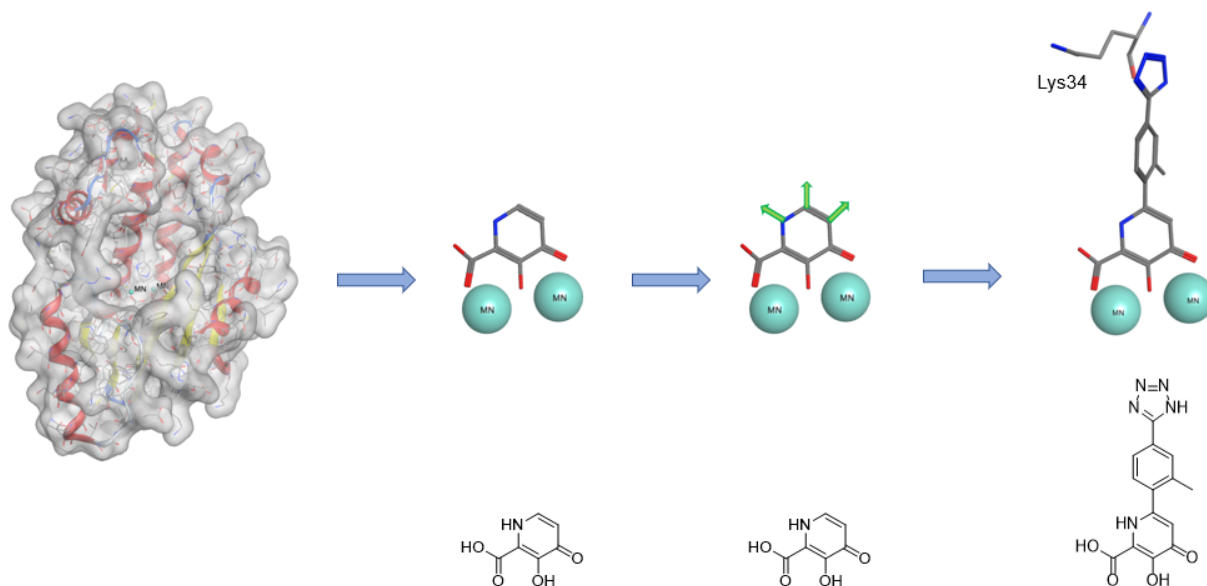
Figure 2.1. Example workflow of computational metalloenzyme FBDD (structure is PDB ID 6E6W).[4] In this chapter, this workflow was performed on a previously published inhibitor to retroactively predict and confirm the mode of MBP binding to a metalloenzyme. First, given structural information (*left*) a known MBP can be docked to the active site metal ions. Next, based on the steric factors of the protein, growth vectors can be defined, and substituent groups can be added to the MBP to make additional interactions with the protein active site. In the example shown, extension of the MBP with a tetrazole group makes a contact with Lys34 (*right*), which results in improved inhibitory activity.

2.2 VS in Metalloenzyme FBDD

As noted in Chapter 1, approximately 40-50% of all enzymes are metal ion dependent,[5-6] with the metal ion(s) in these metalloenzymes playing either a structural or a catalytic/functional role. In drug discovery efforts targeting metalloenzymes, the use of computational methods to aid in inhibitor development can be advantageous. Substantial progress has been made to improve the computational algorithms and scoring functions used to find and identify the energetically lowest conformation of an inhibitor-protein complex.[7-10] Despite advancements in computational

28

techniques, predicting accurate binding affinities and geometries of metalloenzyme-inhibitor interactions remains challenging.[11-13] A recent computational study compared the performance of several commonly used docking programs (AutoDock4, AutoDock4Zn, AutoDock Vina, Quick Vina 2, LeDock, PLANTS, and UCSF DOCK6) with metalloenzymes. While some of these programs could predict the correct binding geometry of a ligand, none of them were able to successfully score docking poses.[14] Some attention has also been devoted towards molecular dynamic (MD) strategies. One study demonstrated the docking of inhibitors with a hydroxamic acid MBP to Zn^{2+}-dependent matrix metalloproteinase 9 upon dividing the computational process into several tasks: (i) docking with the requirement to generate a metal-MBP interaction of the hydroxamic acid using the FlexX program, (ii) optimization of the geometry of the MBP-metalloenzyme complex by quantum mechanical and molecular mechanical (QM/MM) calculations, (iii) conformational optimization of the MBP-metalloenzyme complex without changes in the metal-MBP interaction by MD calculations, and (iv) a single point QM/MM energy calculation. Using this complex workflow, the authors were able to correlate the calculated and experimentally determined binding affinity.[15] However, the aforementioned methods are complex and require significant computational resources and expertise, making them unsuitable for routine virtual screening, docking, and modeling. More user-friendly and accessible methods for predicting metalloenzyme-inhibitor/MBP interactions is needed.

Herein, a readily accessible method for the prediction of MBP-metalloenzyme interactions is presented. The binding pose of the MBP fragment was predicted using a combination of density functional theory (DFT) calculations and a genetic algorithm with Genetic Optimization for Ligand Docking (GOLD) program. The obtained scaffold pose was then extended with MOE and energetically minimized. The method was evaluated against several structurally characterized

metalloenzymes from the Protein Data Bank (PDB) with different metal ion active sites and a variety of bound inhibitors. The predicted binding poses were found to be in good agreement with the crystallographically determined binding modes, providing an accessible tool for computational drug discovery campaigns against metalloenzymes.

2.3 Target Selection

To ensure the generality of the approach, metalloenzymes were examined that contain different metal ions, different numbers of metal ions in the active site, and varying coordination geometries. The following enzymes were selected:

1) Human Carbonic Anhydrase II (hCAII): Human carbonic anhydrase II is a Zn^{2+}-dependent metalloenzyme. The active site consists of a tetrahedral Zn^{2+} ion, bound by three His residues and a capping water molecule. The enzyme catalyzes the hydration of carbon dioxide to a bicarbonate anion that modulates carbon dioxide levels, as well as the pH in the blood stream. hCAII inhibitors are used as a treatment for glaucoma.[16-18]

2) Jumonji-Domain of Histone Lysine Demethylase (KDM): Histone lysine demethylases with a conserved jumonji-domain are Fe^{2+}-dependent enzymes (crystal structures often possess Ni^{2+} or Mn^{2+} ions as a surrogate for Fe^{2+}). The enzyme active site consists of an octahedral Fe^{2+} ion bound by two His and one Glu residue and three capping water molecules. The enzyme catalyzes the demethylation of lysine residues using α-ketoglutaric acid and oxygen. Certain KDMs have been associated with cancer, as well as various mental disorders.[19-21]

3) N-terminal Domain of the Polymerase Acidic subunit of the RNA-dependent RNA Polymerase of the Influenza Virus (PA_N): The N-terminal domain of the polymerase acidic

subunit is a Mn^{2+} or Mg^{2+} dependent metalloenzyme. The enzyme active site consists of two Mn^{2+} or Mg^{2+} ions, bound by one His, one Ile, one Asp, and two Glu residues and five capping water molecules. The enzyme cleaves a 5′-mRNA cap which is necessary for eukaryotic translation. PA_N is considered an important pharmacological target against the influenza virus with one compound approved for human use.[22-24]

2.4 Methodology

The modeling experiment was divided into several tasks for which different docking programs were used based on their individual strengths (Figure 2.2). The three-dimensional structure of a MBP fragment was generated and energetically optimized using DFT calculations with Gaussian. Other molecule editors (i.e., ORCA, Spartan) were also found to be suitable to generate an appropriate, three-dimensional structure of the MBP. Next, the metalloenzyme-inhibitor structure was obtained from the PDB. Using the Molecular Operating Environment (MOE) software suite, water molecules, and other small molecules (e.g., buffer components, cryoprotectant, inhibitors, counter ions) were removed, hydrogen atoms were added, and side chains protonated at physiological pH.

The optimized MBP fragments were docked with a genetic algorithm using GOLD. The structure of the protein was considered rigid and the metal ions in the active site were set with a predefined coordination geometry (e.g., tetrahedral, octahedral). The resulting binding poses were evaluated during docking using the ChemPLP scoring function. Following this, the binding poses from the docking experiment were re-scored using the GoldScore scoring function. Previous studies have shown that rescoring improves the identification of fragment hits than use of a single scoring function.[25-26] Because docking with GOLD requires high computational resources that

would limit virtual screening of large libraries, the docking process was divided into two tasks. The MBP was docked using GOLD, followed by virtual MBP 'fragment growth' into a lead-like molecule using MOE. To elaborate the MBP fragment into the complete inhibitor, the MBP model obtained from GOLD was loaded into MOE and the MBP was manually elaborated to generate the complete inhibitor. The inhibitor was then energetically minimized, while keeping the pose of the MBP generated by GOLD fixed.

To ensure that the metalloenzyme structures studied here are highly homologous, the amino acid sequences and three-dimensional structures were compared. As expected, each target was found with a sequence homology of at least ~93% and a sequence coverage of at least ~91% (Appendix, Table 2.S1). Notably, portions of the structures that did not align were the termini, which were not resolved in all structures. To evaluate the accuracy of the inhibitor docking, the computational and crystallographic structures were aligned using MOE using a protein alignment algorithm. The sequences of the structures were aligned and then the three-dimensional structures superimposed. While the aligned proteins showed generally a high congruency, some portions of the structures showed significant discrepancies. To improve the superposition, the structures were further aligned with a focus on the active site, specifically on the metal ion(s) and adjacent amino acid atoms as a reference point (see Appendix: Supporting Information for details). As a means to quantify the difference between the computational and crystallographic binding geometries, the root-mean-square deviation (RMSD) values were calculated using LigRMSD 1.0.[27]
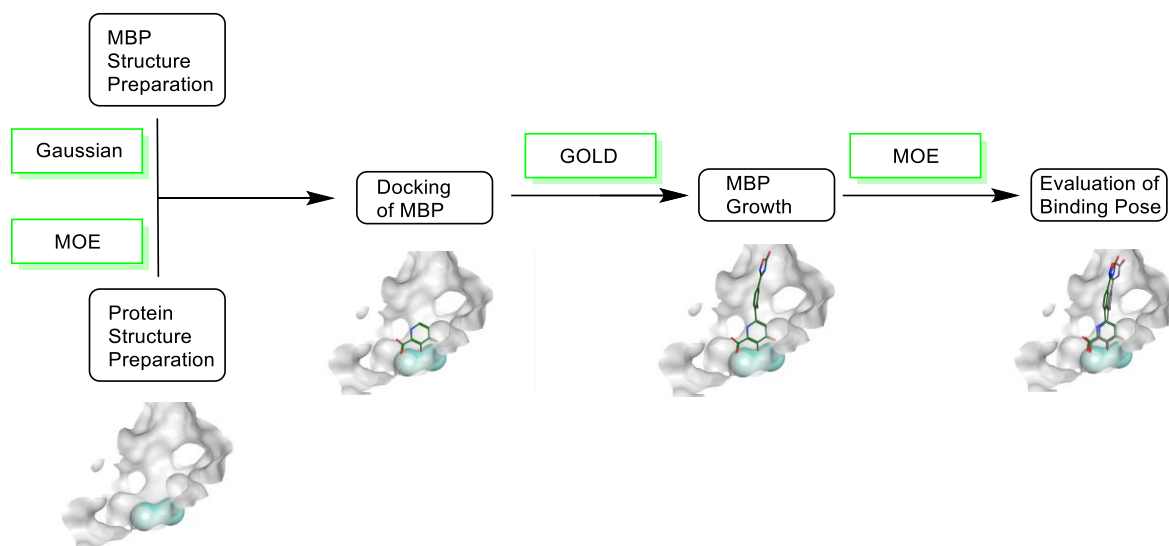
Figure 2.2. Workflow for modeling metalloenzyme-inhibitor interactions. Protein surface (PA$_N$) shown in gray, active site metals shown in cyan, and MBPs and inhibitors shown as sticks colored by atom type.

2.5 MBP Fragment Modeling

To validate the docking methodology, four structurally diverse and crystallographically characterized MBPs for each enzyme target were examined.[4, 28-37] Using DFT calculations and docking with GOLD, the best scoring binding poses of all MBPs with hCAII and PA$_N$, and the majority of MBP with KDM were found coordinated to the active site metal ion(s). In the case of 2,4-pyridinedicarboxylic acid with KDM (PDB 2VD7), the MBP did not dock to the active site, but rather at a distal site (Appendix, Figure 2.S1). To address this outlier, docking was repeated with a box of 20 Å centered around the KDM active site metal ion resulting in 2,4-pyridinedicarboxylic acid docked to the active site metal ion in KDM.

The computational binding poses were compared to the crystallographic structures. A visual comparison shows good alignment for the vast majority of the modelled and experimentally determined metalloenzyme-MBP complexes (Figure 2.3). The only fragment that showed a significant discrepancy was 3-(dimethylamino)benzohydrazide (PDB 6RMP) bound to hCAII.

33

This MBP was computationally predicted to bind the metal with the keto hydrazide moiety in an orientation that was reversed from the crystallographic structure. The quality of the docking was quantified by determining the RMSD values between the computational and crystallographic configurations. As expected, the predicted binding poses of the MBP fragment were found to be in good agreement with an average RMSD value of 0.87 Å (Table 2.1).
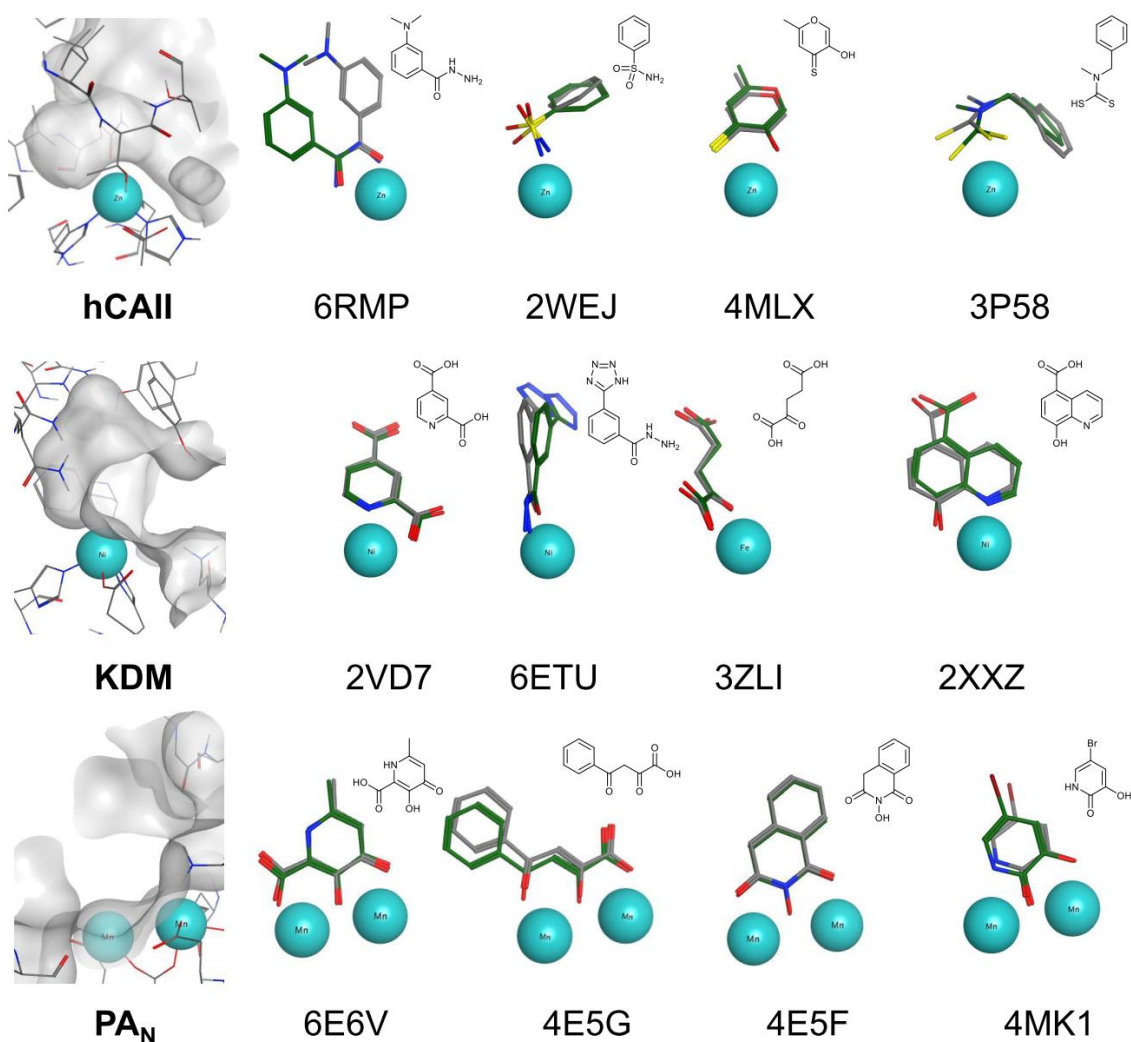
Figure 2.3. Comparison of MBP binding poses from crystallographic data (gray carbons) and computational docking (green carbons) using GOLD. Examples from hCAII (*top row*) demonstrate inhibitor binding to zinc, while the middle row (KDM) demonstrates binding to iron or nickel. Examples binding the dinuclear manganese active site of PA$_N$ (*bottom row*) are also shown. PDB accession codes are shown beneath each example, while chemical structures are shown above. The active site amino acid residues are also shown to the left of each row.

34

Table 2.1. RMSD values of computationally and crystallographically determined metalloenzyme-MBP complexes.

Enzyme	Entry	RMSD / Å	Reference
hCAII	2WEJ	0.49	[28]
	3P58	0.86	[29]
	4MLX	0.32	[30]
	6RMP	3.75	[31]
KDM	2VD7	0.22	[32]
	2XXZ	0.68	[33]
	3ZLI	0.34	[34]
	6ETU	0.94	[35]
PA$_N$	4E5F	0.23	[36]
	4E5G	0.63	[36]
	4MK1	1.67	[37]
	6E6V	0.34	[4]

2.6 Inhibitor Modeling

Following successful MBP geometry prediction, the modeling of lead-like inhibitors based on these MBPs was performed. A total of 47 crystal structures of metalloenzyme-inhibitor complexes from the PDB were examined. Specifically, benzenesulfonamide inhibitors with hCAII,[38-55] 4-pyridinecarboxylic acid inhibitors with KDM,[32, 56-63] and 3-hydroxy-4-oxo-1,4-dihydropyridine-2-carboxylic acid inhibitors with PA$_N$[4, 36, 64] were modeled. The bound inhibitors were selected to have broad chemical composition to test the generality of the methodology (Figures 2.S2-2.S4). To verify the binding pose of the MBP fragment, the computational poses from GOLD were compared to the crystal structures of bound lead-like inhibitors, comparing only on the MBP moiety (Appendix, Figure 2.S5). Overall, the fragments showed good agreement to the experimentally binding geometries with RMSD values of 0.90 Å for hCAII, 0.52 Å for KDM,

and 0.78 Å for PA_N (for a complete list of RMSD values see Appendix, Table 2.3). That stated, while the crystallographically characterized binding poses for the fragments bound to hCAII and KDM were very similar, the MBP for PA_N was found in more varied orientations across the crystal structures (Figures 2.S5c-d). While the bond distances for MBPs in PA_N were all similar, the MBPs are bound with orientational differences relative to the active site. A binding angle was arbitrarily defined as the carbon atom of the carboxyl group of Asp108, a centroid between both metal centers, and a centroid of the aromatic MBP moiety. This angle ranged from 175° (PDB 4E5I, Appendix, Figure 2.S5) to 145° (PDB 6E3N, Figure 2.S5) across structures and likely originates from the influence of other interactions these full-length inhibitors form with subpockets in the relatively large, open active site of PA_N.

The benzenesulfonamide fragment in hCAII showed the largest RMSD discrepancy between the docked and experimental structures. These differences could negatively influence the next stage in the docking the full-length inhibitor, where upon ligand extension the predicted binding pose for the MBP could lead to inaccuracies in the full ligand pose. Therefore, a modified MBP docking protocol was used for the benzenesulfonamide fragment in hCAII. The binding pose of the MBP is primarily driven by metal coordination and hydrogen bonding interactions of the sulfonamide moiety with hCAII. Therefore, the docking of this fragment was repeated with a bias that added additional metal binding terms to the GoldScore scoring function, as previously described.[65] With these adjustments to the scoring function, the binding pose of benzenesulfonamide was better aligned with the experimentally determined structures (Appendix, Figure 2.S6), and produced a low RMSD value of 0.47 Å (Table 2.4). This experiment shows that additional information about expected MBP-active site interactions can be used to better predict the binding pose. To avoid adding a user bias in applying our method (caused by modifying the

36

applied scoring function), the initial binding pose (Figure 2.S5) was used for the further ligand extensions.

As described above, the three-dimensional protein structures of the computationally and experimentally determined enzyme-inhibitor complexes were superimposed using an automatic algorithm and then further refined with focus on the active site with an intentional bias towards the active site metal ion(s) as well as the coordinating amino acids (Figure 2.S7). As the inhibitors are expected to coordinate to the metal ion as part of their mechanism of action, these elements should be aligned with the highest possible accuracy. This procedure resulted in average RMSD values of 0.23 Å for hCAII, 0.13 Å for KDM, and 0.14 Å for PA_N for the respective active sites alignments.

Following the protein structure alignment, the computationally and experimentally determined enzyme-inhibitor complexes were compared (Figure 2.4, Figures 2.S8-2.S10). Overall, a high congruency between the docked models and the experimentally determined binding poses is observed with average RMSD values of 2.44 Å for hCAII, 1.57 Å for KDM, and 1.60 Å for PA_N (for a full list of RMSD values see Table 2.S4, Appendix). It is noteworthy that the discrepancies of the computationally modelled and experimentally determined binding poses are within the average crystal structure resolution in the PDB, with the most entries possessing a resolution of 1.8-2.0 Å.[66]
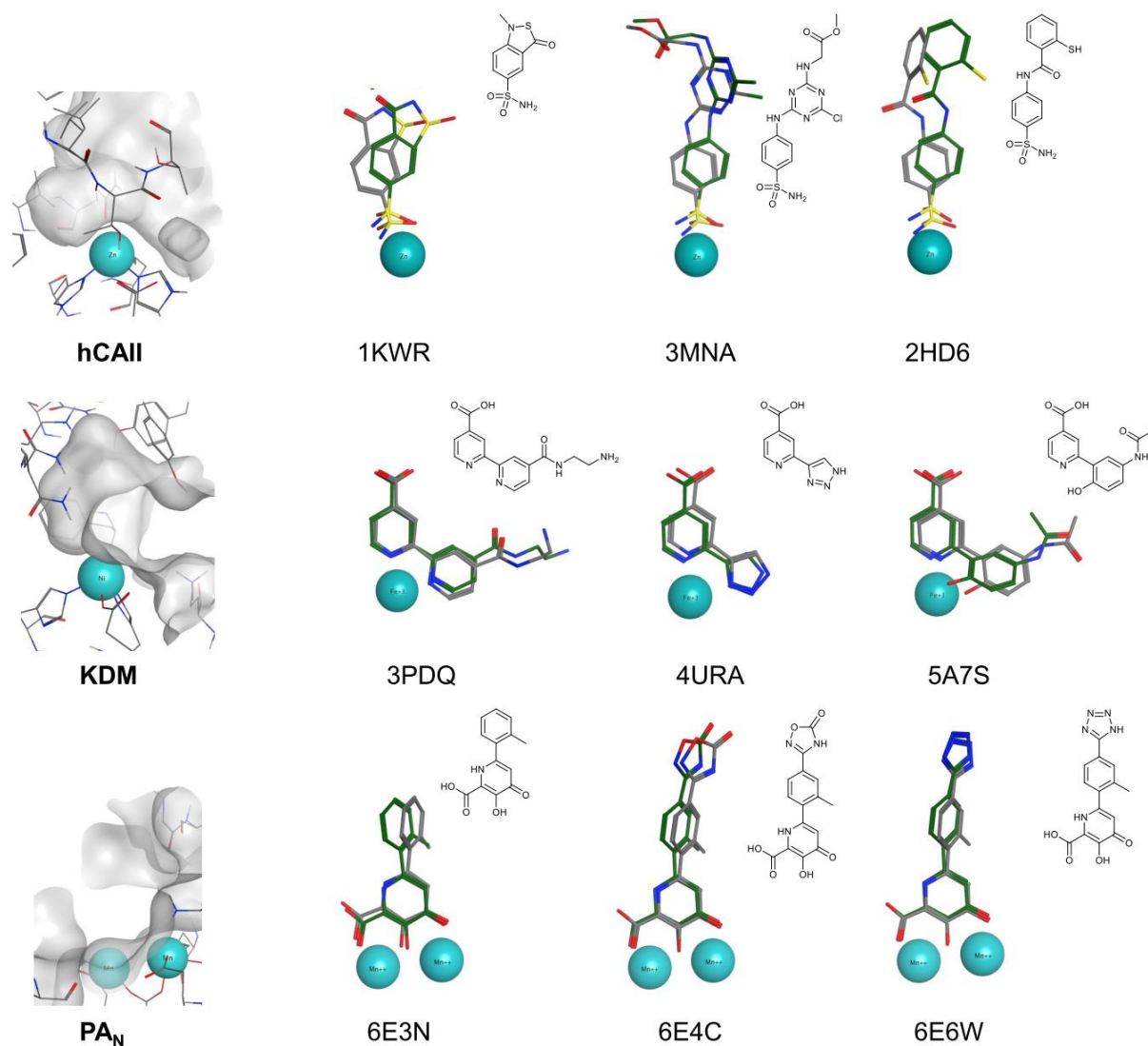
Figure 2.4. Comparison of binding modes of crystallographically determined structures (gray carbons) and computationally derived inhibitor poses (green carbons) from several PDB entries (PDB entry codes shown beneath each example). Examples from hCAII (*top row*), KDM (*middle row*), and PA$_N$ (*bottom row*) are shown. demonstrates binding to iron or nickel. Examples binding the dinuclear manganese active site of PA$_N$ (*bottom row*) are also shown. For reference, the active site amino acid residues are also shown to the left of each row.

The RMSD distribution of the structure alignments was analyzed (Figure 2.5). While a few of the docked compounds showed large discrepancies to the experimentally determined structures (RMSD >4 Å), the majority of docked inhibitors were in good agreement with the experimental structures, with RMSD values well within the resolution range of typical protein-

inhibitor co-crystal structures (RMSD <2 Å). These results suggest that the computational approach described here could be useful for the developing MBP fragments into lead-like compounds via fragment growth.
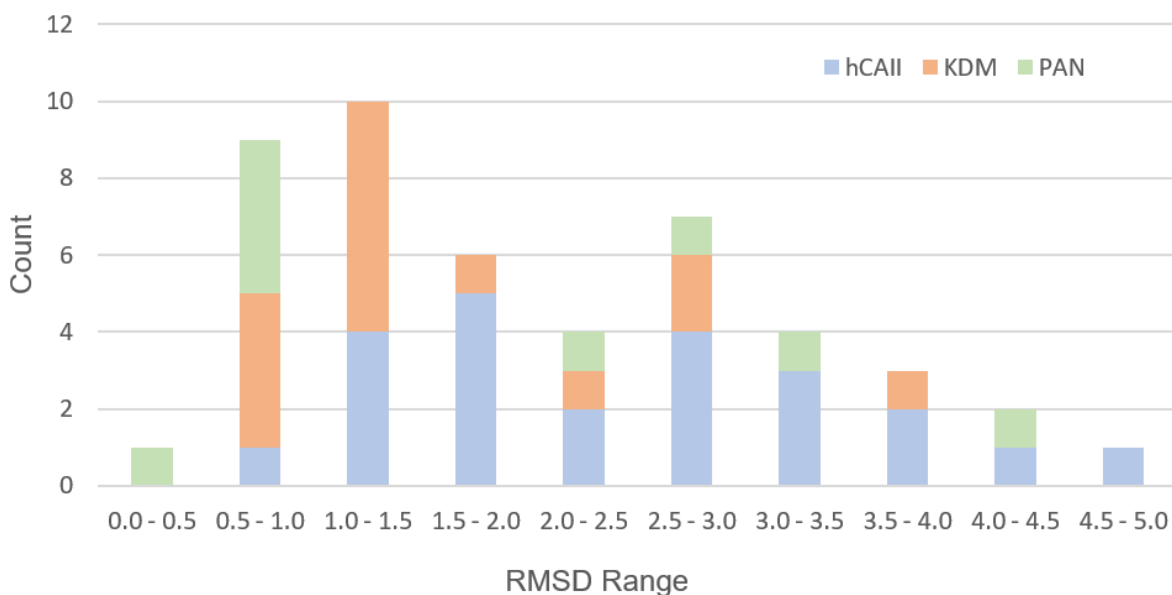
Figure 2.5. Distribution of RMSD values of computational versus experimental binding poses for metalloenzyme inhibitors. Color-coded by metalloenzyme: hCAII (blue), KDM (orange), and PA_N (green).

Finally, the docking method was compared with commonly used docking programs AutoDock Vina (Figure 2.S11) and SwissDock (Figure 2.S12) under standard docking procedures using hCAII and PA_N as model systems. Upon applying these methods, the MBP was not found in the active site and were not consistent with the crystallographically characterized structures. These findings further highlight the value of the methods described here for the prediction of binding poses of inhibitors in metalloenzymes.

2.7 Conclusion

A new computational strategy for modeling the binding of MBPs in metalloenzymes is presented. The approach uses a combination of DFT calculations and docking with a genetic algorithm (GOLD). The methodology was evaluated by comparing computational models with experimentally determined metalloenzyme-inhibitor complexes. Metalloenzymes with different metal ions, different numbers of metal ions in the active site, and different metal coordination geometries were examined to determine the scope and robustness of the approach. Good agreement was found between the computational models and experimentally determined structures, well within the resolution of most crystal structures in the PDB. The workflow described here is a robust method for predicting MBP binding poses and further computational methods that address energetic optimization of complete, lead-like inhibitors would be useful next steps in improved modeling of metalloenzyme inhibitors. Overall, the presented methodology may provide a more accessible approach for the design of novel metalloenzyme inhibitors, especially in the context of FBDD. Work in Chapter 3 will use this methodology to predict MBP binding of novel MBPs against PA$_N$ endonuclease. These predicted poses will then be examined by comparing them to newly obtained protein co-crystal structures.

2.8 Acknowledgements

2.9 Appendix: Supporting Information

Structure Preparation

The geometry of the fragment was determined using density-functional theory (DFT) calculations with the Gaussian software package (Gaussian, Inc., Wallingford CT, 2016). The calculations were performed using the Pople double-zeta basis set with a single set of polarization functions on non-hydrogen atoms (6-31G(d)). Solvent effects were included using a polarizable continuum model (PCM). The structures of all calculated molecules correspond to minima on the ground state potential energy surfaces with no imaginary frequencies present. The metalloprotein structures human carbonic anhydrase II (PDB 6GOT),[67] jumonji-domain of histone lysine demethylase (PDB 5ANQ),[60] and N-terminal domain of the polymerase acidic subunit of the RNA-dependent RNA polymerase of the influenza virus (PDB 6E6W)[4] were downloaded from the Protein Data Bank (https://www.rcsb.org/) and prepared for the docking experiment using the Molecular Operating Environment (MOE, Chemical Computing Group ULC, Montreal, QC, Canada, 2019) software package. Water molecules, other small molecules, inhibitors as well as ions were removed. Hydrogen atoms were added and the side chains protonated at physiological pH.

Protein Sequence Homology

Percent sequence identity and percent sequence coverage was determined using BLAST.[68] hCAII sequences were compared to the sequence of PDB accession code 6RKN. KDM sequences were compared to the sequence of PDB accession code 2VD7. PA$_N$ sequences were compared to the sequence of PDB accession code 6DCZ.

Fragment Docking

The structurally optimized fragments were docked using the Genetic Optimization for Ligand Docking (GOLD, Cambridge Crystallographic Data Centre, Cambridge, United Kingdom, 2019) software package. The protein structure was considered rigid. The binding pose of the compounds was predicted using a genetic algorithm with a population size of 200, selection pressure of 1.2, number of operations of 500000, number of islands of 5, niche size of 2, crossover frequency of 95, mutation frequency of 95, and a migration frequency of 10. The genetic algorithm was set to 100 runs. The metal ions within the active site were predefined to have a tetrahedral geometry for Zn and an octahedral geometry for Fe or Mn. During the docking procedure the binding poses were evaluated using the ChemPLP scoring function. After docking, the obtained solutions were re-scored using the GoldScore fitness function.

Protein Alignment

PDB files and the docking pose from GOLD were loaded into MOE and identical subunits were removed. The sequences were aligned, and the pocket residues were superposed. For PA_N, endonuclease, residues 109-119 were superposed with an RMSD of 0.17 Å. The active site alignment was further refined by superposing the proteins using two manganese atoms and the amine in His41 as alignment points, with an RMSD of 0.14 Å (Figure 2.S7a.). For KDM, the sequences were aligned, and the pocket residues were superposed with an RMSD of 0.28 Å. The active site alignment was further refined using the Fe ion, the amine from His188 and the amine in His276 as alignment points, with an RMS distance of 0.13 Å (Figure 2.S7b.). For hCAII, residues the pocket residues were superposed with an RMSD of 0.23 Å, which provided sufficient alignment and did not require further refinement (Figure 2.S7c.).

Fragment Growth

Fragment growth was performed in MOE. The model structure from GOLD was loaded into MOE. The fragment model for each protein was manually extended, and the added group was energy minimized using the Amber10:EHT forcefield with a gradient of 0.1 RMS kcal/mol/ Å^2, while the MBP was not altered. For the PA_N model, the atoms in the core of the MBP were manually changed to match the corresponding PDB structure.

Evaluation of binding pose prediction

The binding pose was evaluated by comparing the computationally docked model with the corresponding PDB structure. After the proteins were aligned, poses of the PDB ligand and the computationally-derived ligand were uploaded to the LigRMSD 1.0 server[27] and an RMSD value was calculated.

Docking with AutoDock Vina

The structurally optimized fragments for 2WEJ, 1I9L, 6E6V, and 6E6W were docked using the AutoDock Vina software (version 1.2.0, The Scripps Research Institute, La Jolla CA, 2021). The protein structure was considered rigid. The cubic grid for the docking experiment was chosen so that the whole protein was covered. As default parameters, the energy range was set at 4 and the exhaustiveness at 8. The best scored binding pose was compared with the crystallographically determined binding pose.

Docking with SwissDock

      The structurally optimized fragments 2WEJ, 1I9L, 6E6V, and 6E6W were docked using SwissDock (Swiss Institute of Bioinformatics, Lausanne, 2021). The protein structure was considered rigid. No information about the binding pocket was provided, allowing docking towards the whole protein. The best scored binding pose was compared with the crystallographically determined binding pose.

SUPPORTING Figures AND TABLES

Table 2.S1. Percent sequence identity and percent sequence coverage between proteins.

Protein	PDB	Percent Identity	Percent Coverage	Protein	PDB	Percent Identity	Percent Coverage	Protein	PDB	Percent Identity	Percent Coverage
hCAII	4MTY	Parent	Parent	KDM	2VD7	Parent	Parent	PA$_N$	6DCZ	Parent	Parent
	1IF8	100	99		3PDQ	100	100		4E5I	92.97	96
	1I9L	99	99		4GD4	100	100		4E5J	92.97	96
	2HD6	100	99		4URA	100	94		6E3M	100	100
	1KWR	100	100		5A7N	100	100		6E3N	100	100
	3K34	100	100		5A7O	99.74	100		6E3O	100	100
	3MNA	100	100		5A7P	100	100		6E3P	100	100
	3MMF	100	100		5A7Q	100	100		6E4C	100	100
	3MZC	100	100		5A7S	100	100		6E6V	100	100
	3NON	100	100		5A7W	99.74 - 100*	100		6E6W	100	100
	3N4B	100	100		5A80	100	100				
	4ILX	100	98		5ANQ	100	94				
	4ITP	100	98		5F5I	100	95				
	4XE1	100	100		5VMP	97.33	98				
	5BYI	100	100		6CG2	100	91				
	5E2R	100	100								
	5LJK	100	99								
	5N1S	100	100								
	5N1R	100	100								
	5N24	100	100								
	6EQU	100	100								
	6IOW	100	100								
	6SDS	100	100								

* Two different sequences reported in PDB accession code 5A7W.
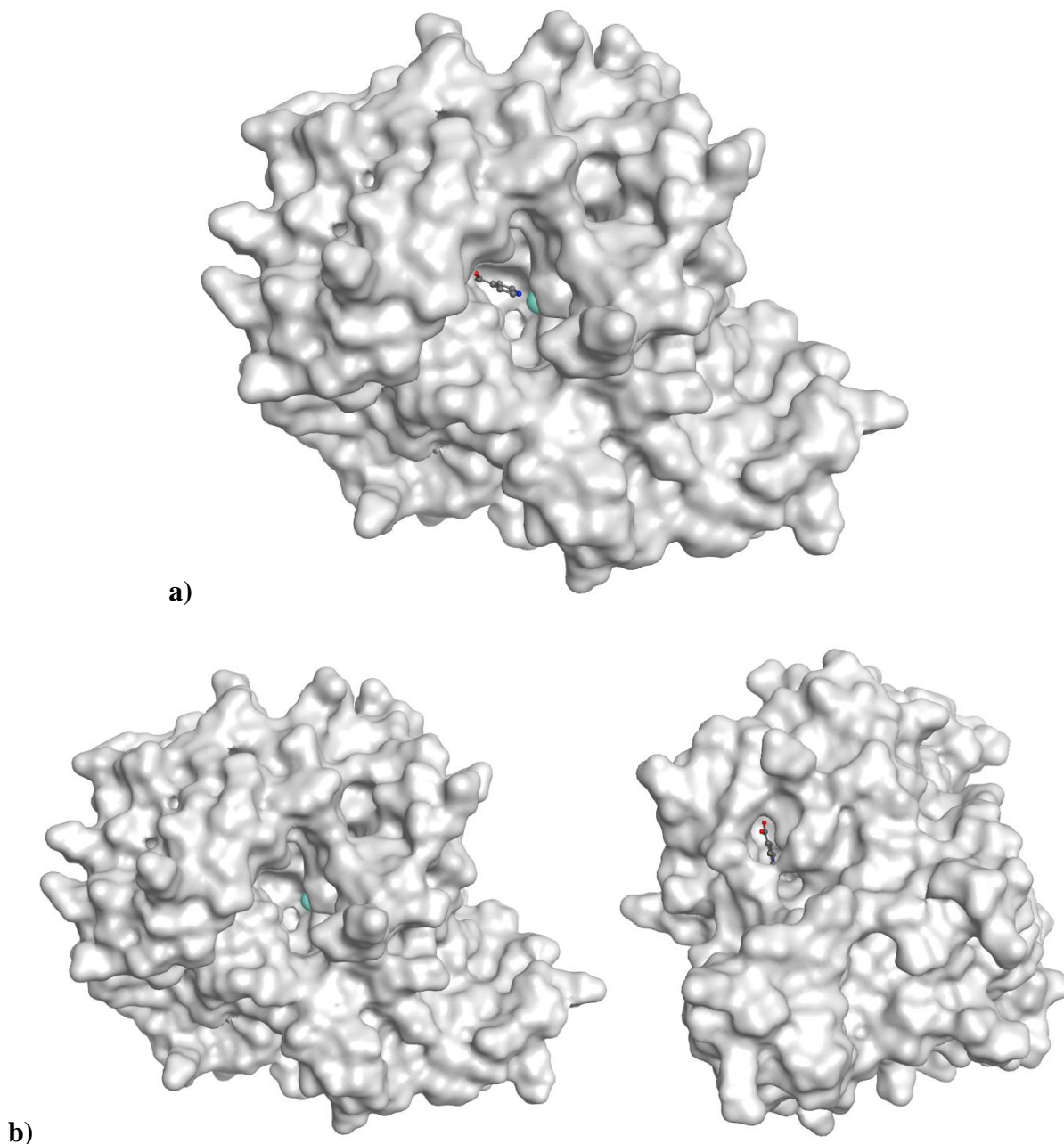
Figure 2.S1. Binding pose of MBP 2,4-pyridinedicarboxylic acid: a) restricted to a box of 20 Å centered around the KDM active site metal ion (where the MBP ultimately binds), or b) without any restrictions. Without any restrictions, the MBP does not bind at the active site (left), but rather at a distal site (right, protein image rotated ~120° to show docked site) in KDM. The distal site is Protein (PA$_N$) shown as gray surface, active site metals shown as cyan surface, and MBP shown as sticks colored by atom type.

MBP Fragment

Full Length Inhibitors

1I9L	**1IF8**	**1KWR**	**2HD6**	**3K34**	**3MMF**	**3MNA**	**3N0N**	**3N4B**

3MZC	**4ILX**	**4ITP**	**4MTY**	**4XE1**	**5BYI**	**5E2R**	**5LJQ** **5N1R**
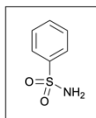
5N1S	**5N24**	**6EQU**	**6I0W**	**6SDS**
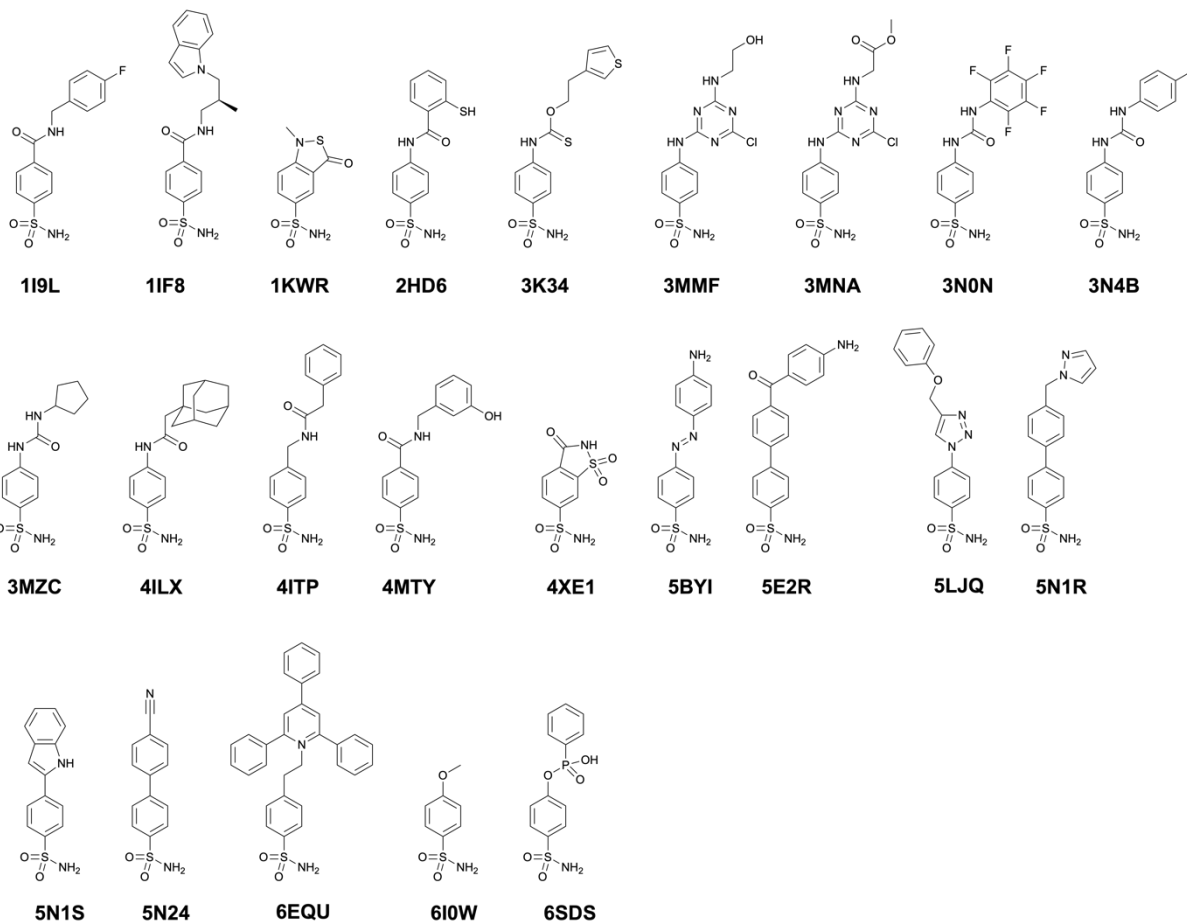
Figure 2.S2. Chemical Structures of the docked fragment as well as full length inhibitors for hCAII. Name of the inhibitors belongs to their corresponding PDB code.[38-55]

47

MBP Fragment

Full Length Inhibitors

| 2VD7 | 3PDQ | 4GD4 | 4URA | 5A7N | 5A7O | 5A7P | 5A7Q |

| 5A7S | 5A7W | 5A80 | 5ANQ | 5F5I | 5VGI | 6CG2 |
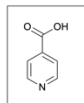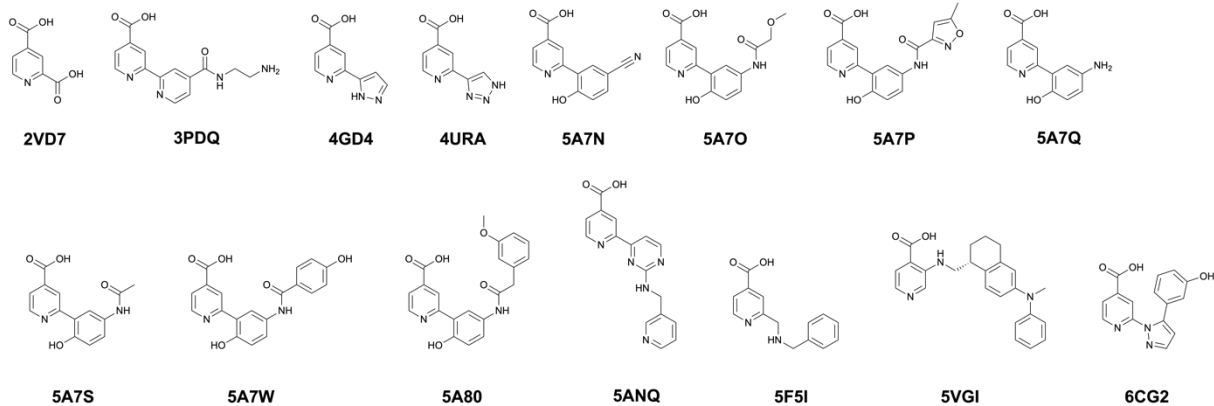
Figure 2.S3. Chemical Structures of the docked fragment as well as full length inhibitors for KDM. Name of the inhibitors belongs to their corresponding PDB code.[32, 56-63]

MBP Fragment

Full Length Inhibitors

| 4E5I | 4E5J | 6DCZ | 6E3M | 6E3N | 6E3O | 6E3P | 6E4C | 6E6W |
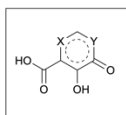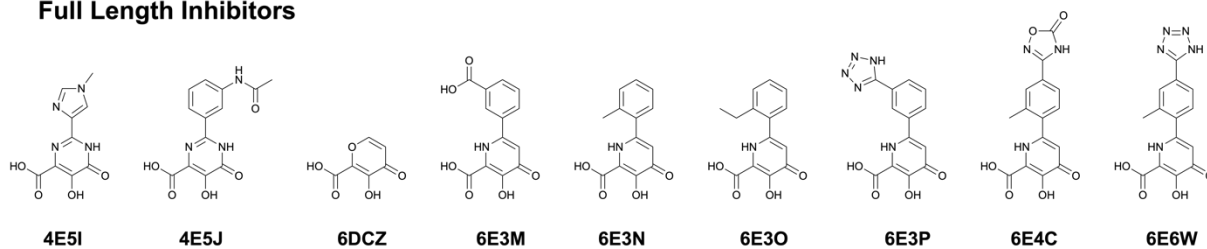
Figure 2.S4. Chemical Structures of the docked fragment as well as full length inhibitors for PA$_N$. Name of the inhibitors belongs to their corresponding PDB code.[4, 36, 64]

Table 2.S2. Root-mean-square deviation of atomic positions (RMSD) between the computational model and only the MBP of the evaluated PDB structures.

Protein	PDB ID	RMSD	Protein	PDB ID	RMSD	Protein	PDB ID	RMSD
hCAII	Model		KDM	Model		PA$_N$	Model	
	1I9L	0.71		2VD7	0.5		4E5J	1.14
	1IF8	0.85		3PDQ	0.22		4E5I	1.92
	1KWR	0.87		4GD4	0.69		6E4C	0.36
	2HD6	0.91		4URA	0.69		6E3N	0.42
	3K34	0.86		5A7N	0.23		6DCZ	0.69
	3MMF	0.76		5A7O	0.47		6E3M	1.07
	3MNA	0.77		5A7P	0.44		6E6W	0.30
	3MZC	1.01		5A7Q	0.93		6E3P	0.76
	3N0N	0.91		5A7S	0.76		6E3O	0.33
	3N4B	0.93		5A7W	0.45		Average	0.78
	4ILX	0.97		5A80	0.47			
	4ITP	0.79		5ANQ	0.24			
	4MTY	0.84		5F5I	0.56			
	4XE1	1.08		5VGI	0.71			
	5BYI	0.86		6CG2	0.46			
	5E2R	1.02		Average	0.52			
	5N1R	1.03						
	5N24	1.10						
	6EQU	0.96						
	6I0W	0.77						
	6SDS	0.79						
	Average	0.90						
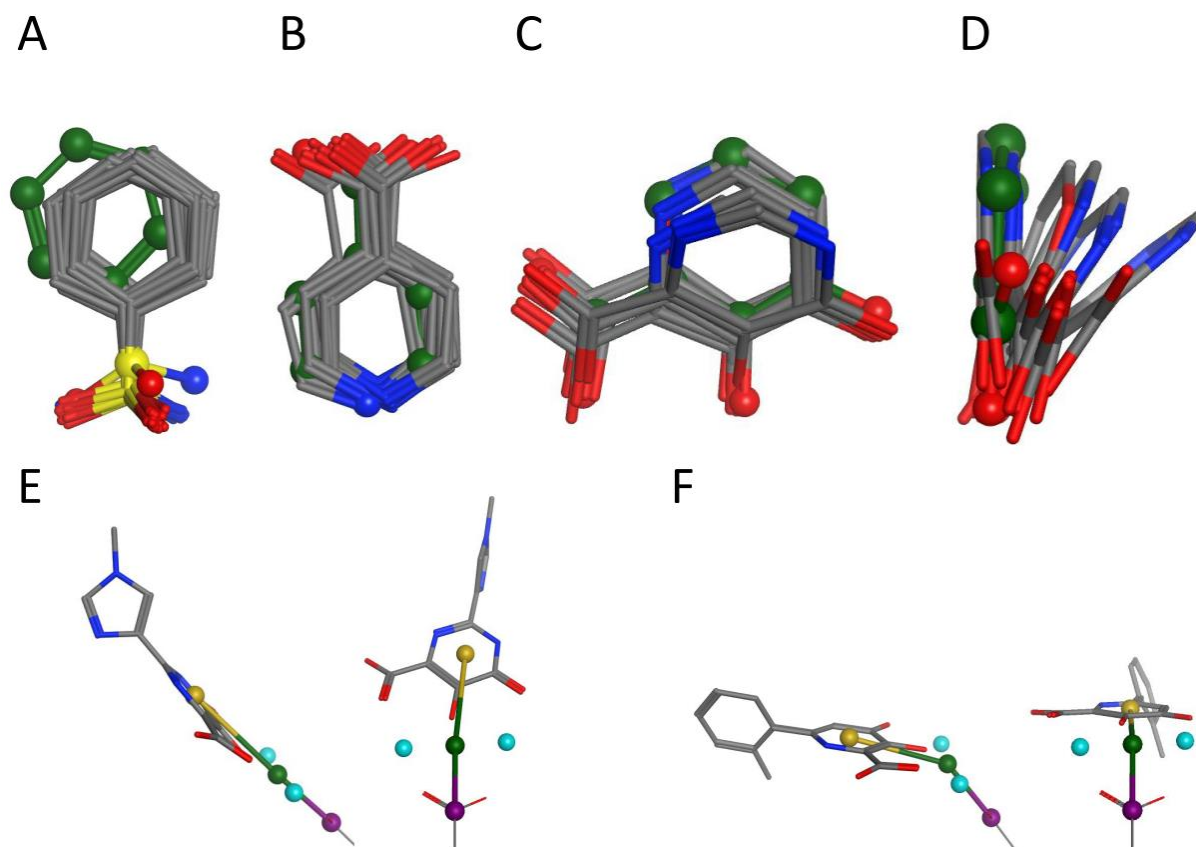
Figure 2.S5. Overlay between the docked MBP using GOLD (green) and the MBP fragment from the corresponding PDB structures for: (A) hCAII, (B) KDM, (C) PA$_N$, and (D) PA$_N$ (perspective of this pose is rotated by 90° compared to the other images to highlight the 'tilt' of the MBP poses). Angle between the carbon atom of the carboxyl group of Asp108 (purple), a centroid between both metal centers (green), and a centroid of the aromatic MBP moiety (yellow) for (E) 4E5I (175° binding angle) and (F) 6E3N (145° binding angle).
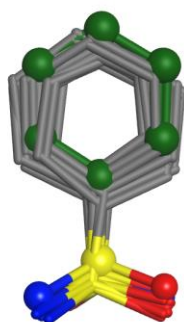
Figure 2.S6. Overlay between the computationally docked MBP (green) and corresponding MBPs from PDB structures for hCAII (gray). Using a modified scoring function with a metal-binding bias the benzenesulfonamide MBP is much better aligned with the experimentally determined structures (compare with Figure 2.S5a).

Table 2.S3. Root-mean-square deviation of atomic positions (RMSD) between the computational model using a scoring function with a metal-binding bias and only the MBP of the evaluated PDB structures.

Protein	PDB ID	RMSD
hCAII	Model	
	1I9L	0.41
	1IF8	0.60
	1KWR	0.34
	2HD6	0.46
	3K34	0.42
	3MMF	0.30
	3MNA	0.27
	3MZC	0.42
	3N0N	0.34
	3N4B	0.38
	4ILX	0.34
	4ITP	0.29
	4MTY	0.46
	4XE1	0.51
	5BYI	0.51
	5E2R	0.53
	5LJQ	0.38
	5N1R	1.19
	5N1S	0.52
	5N24	0.53
	6EQU	0.48
	6I0W	0.56
	6SDS	0.53
	Average	0.47
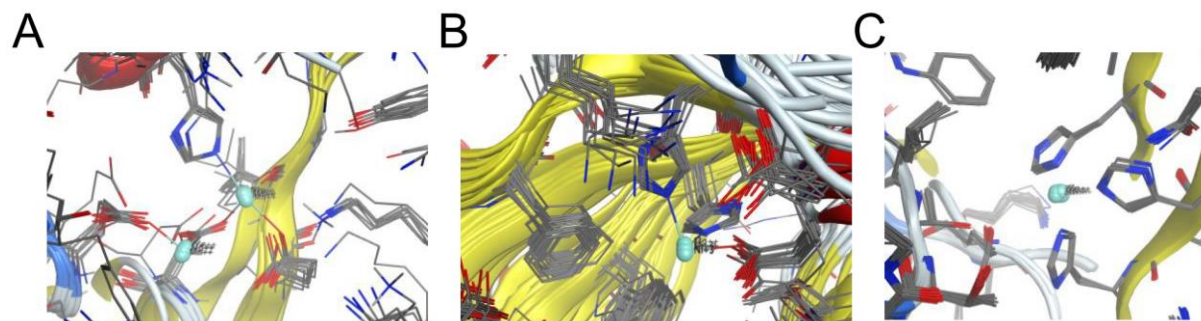
Figure 2.S7. Active site alignment for: (A) PA$_N$, (B) KDM, and (C) hCAII.
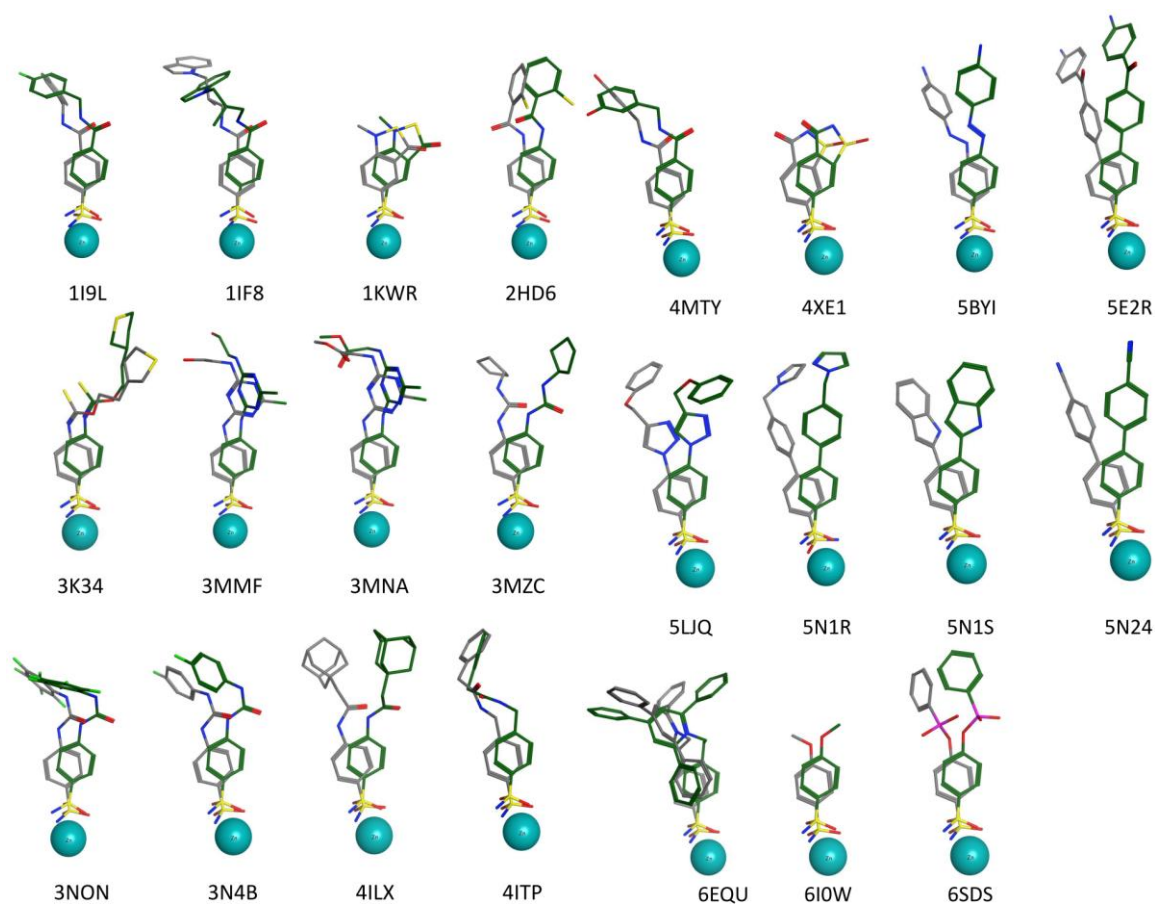
Figure 2.S8. Comparison of binding modes of computationally derived hCAII inhibitor poses (green) and crystallographically determined structures (gray) from PDB entries (PDB entry codes shown).
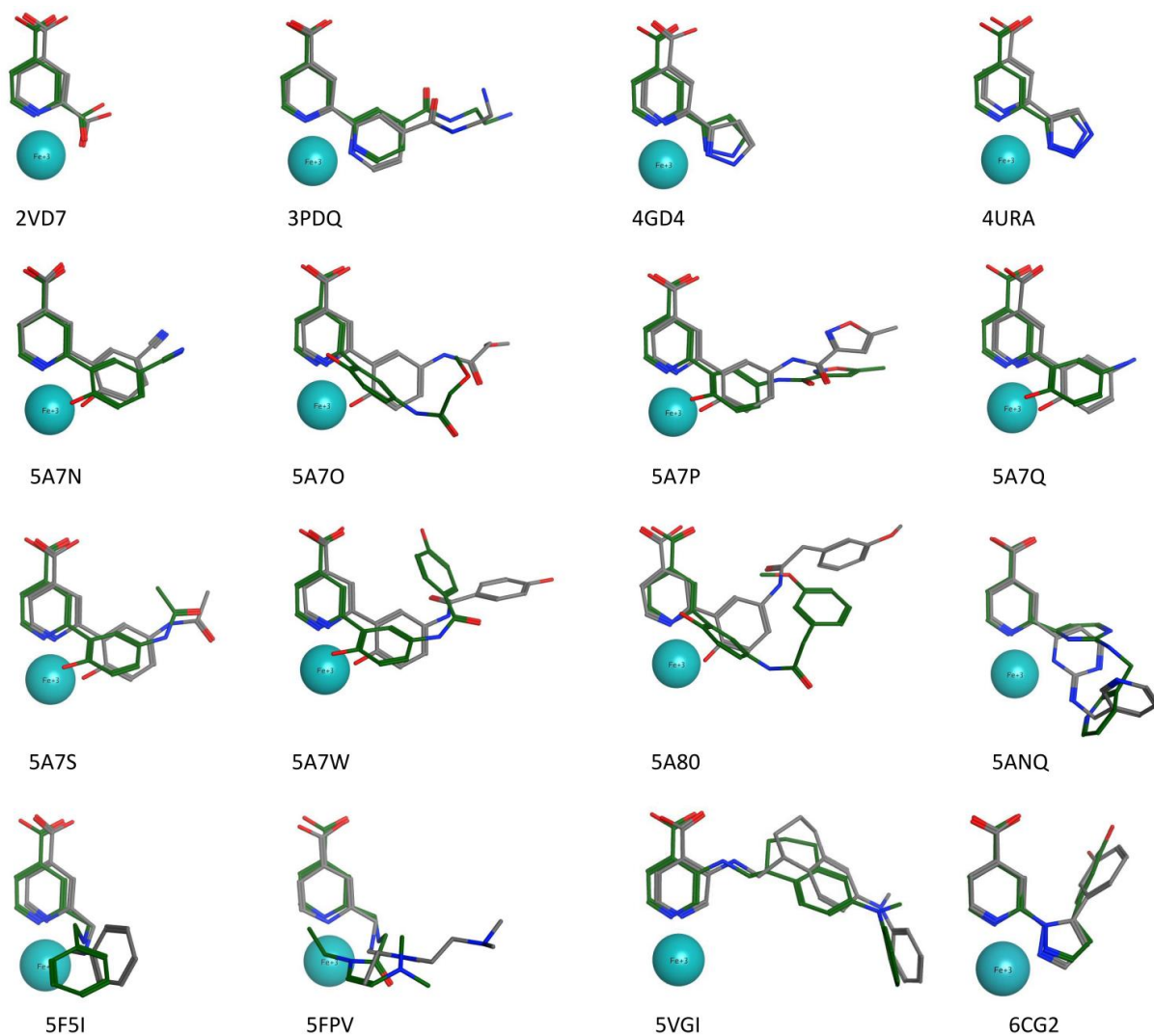
Figure 2.S9. Comparison of binding modes of computationally derived KDM inhibitor poses (green) and crystallographically determined structures (gray) from PDB entries (PDB entry codes shown).
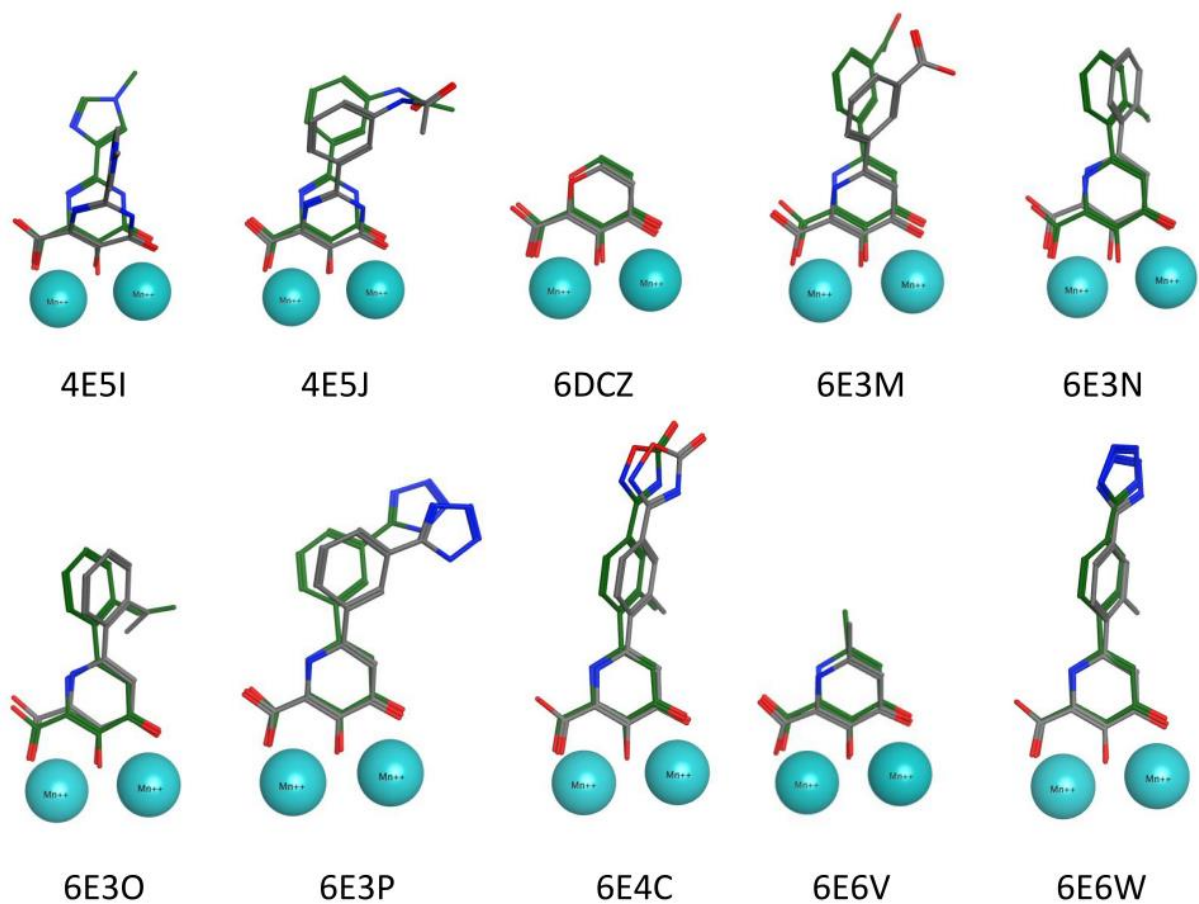
Figure 2.S10. Comparison of binding modes of computationally derived PA$_N$ inhibitor poses (green) and crystallographically determined structures (gray) from PDB entries (PDB entry codes shown).

Table 2.S4. RMSD values of computationally and crystallographically determined full-length inhibitors of the enzyme-inhibitor complexes.

Enzyme	Entry	RMSD	Reference
hCAII	1I9L	1.69	6
	1IF8	3.18	7
	1KWR	1.07	8
	2HD6	1.74	9
	3K34	1.91	10
	3MMF	1.67	11
	3MNA	1.20	11
	3N0N	3.79	12
	3N4B	3.03	12
	3MZC	3.73	12
	4ILX	4.11	13
	4ITP	0.95	14
	4MTY	1.86	15
	4XE1	1.33	16
	5BYI	2.79	17
	5E2R	3.07	18
	5LJQ	4.58	19
	5N1R	2.80	20
	5N1S	2.87	20
	5N24	2.61	20
	6EQU	2.48	21
	6I0W	1.24	22
	6SDS	2.35	23
	Average	2.44	
KDM	2VD7	0.52	24
	3PDQ	0.85	25
	4GD4	0.69	26
	4URA	0.67	27
	5A7N	1.04	28
	5A7O	2.77	28
	5A7P	1.30	28
	5A7Q	1.02	28
	5A7S	1.32	28
	5A7W	2.70	28
	5A80	3.92	28
	5ANQ	2.10	2
	5F5I	1.42	29
	5VGI	1.79	31
	6CG2	1.46	30-31
	Average	1.57	
PA$_N$	4E5I	4.22	32
	4E5J	3.05	32
	6DCZ	0.69	33
	6E3M	2.61	3
	6E3N	0.65	3
	6E3O	0.83	3
	6E3P	2.09	3
	6E4C	0.57	3
	6E6W	0.39	3
	Average	1.60	
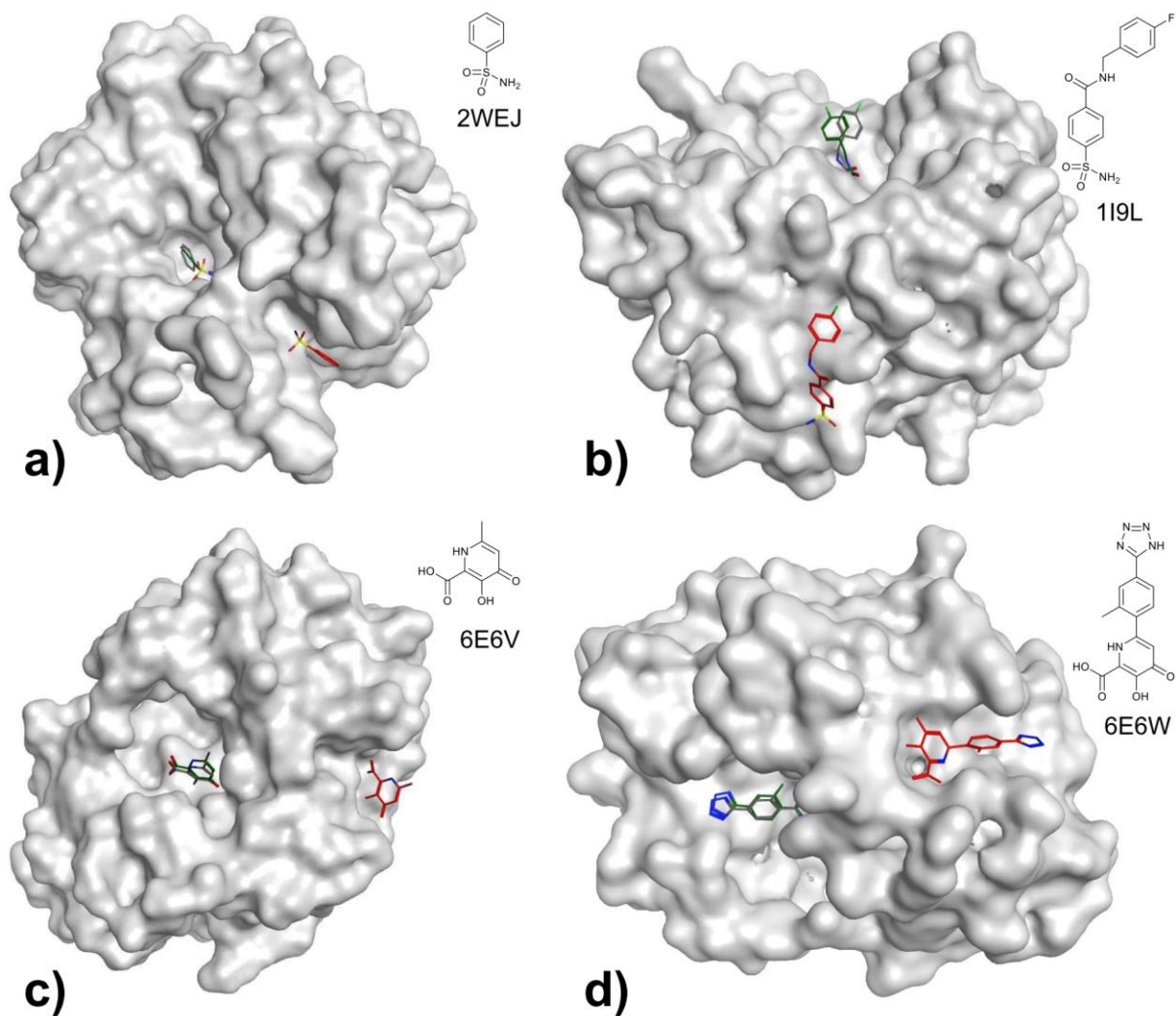
Figure 2.S11. Comparison of binding modes of computationally derived hCAII inhibitor poses upon docking with AutoDock Vina (red), our reported docking procedure (green), and crystallographically determined structures (gray) from PDB entries: a) 2WEJ, b) 1I9L, c) 6E6V, d) 6E6W.

56
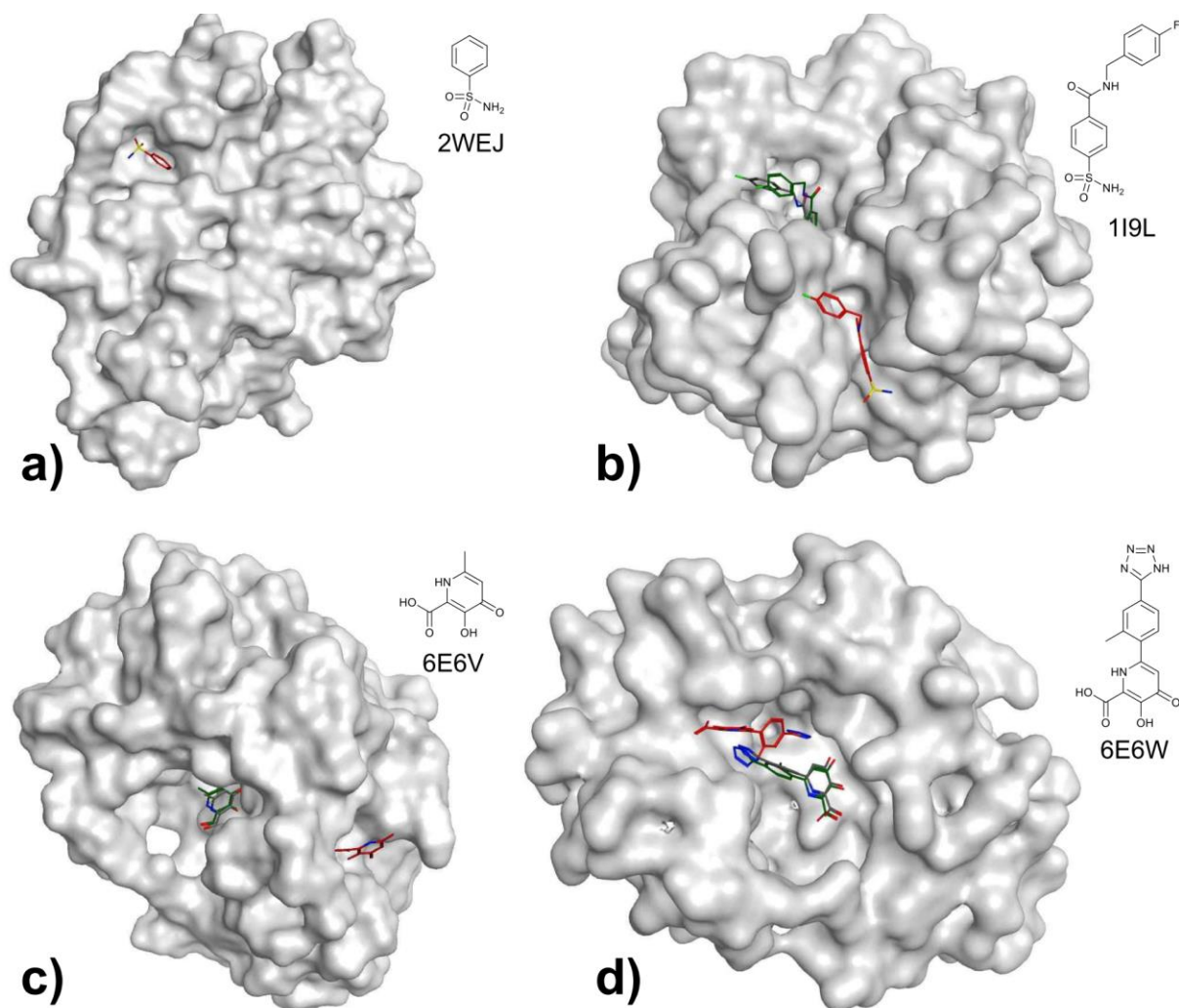
Figure 2.S12. Comparison of binding modes of computationally derived PA$_N$ inhibitor poses upon docking with SwissDock (red), our reported docking procedure (green), and crystallographically determined structures (gray) from PDB entries: a) 2WEJ (the active site is found on the other side of this protein and therefore only the with SwissDock determined binding pose can be seen), b) 1I9L, c) 6E6V, d) 6E6W.

2.10 References

1. Maia, E. H. B.; Assis, L. C.; de Oliveira, T. A.; da Silva, A. M.; Taranto, A. G., Structure-Based Virtual Screening: From Classical to Artificial Intelligence. *Front. Chem.* **2020,** *8*.

2. Lavecchia, A.; Giovanni, D. C., Virtual Screening Strategies in Drug Discovery: A Critical Review. *Curr. Med. Chem.* **2013,** *20* (23), 2839-2860.

3. de Souza Neto, L. R.; Moreira-Filho, J. T.; Neves, B. J.; Maidana, R. L. B. R.; Guimarães, A. C. R.; Furnham, N.; Andrade, C. H.; Silva, F. P., In silico Strategies to Support Fragment-to-Lead Optimization in Drug Discovery. *Front. Chem.* **2020,** *8*.

4. Credille, C. V.; Morrison, C. N.; Stokes, R. W.; Dick, B. L.; Feng, Y.; Sun, J.; Chen, Y.; Cohen, S. M., SAR Exploration Of Tight-Binding Inhibitors Of Influenza Virus PA Endonuclease. *J. Med. Chem.* **2019,** *62* (21), 9438-9449.

5. Andreini, C.; Bertini, I.; Cavallaro, G.; Holliday, G. L.; Thornton, J. M., Metal Ions In Biological Catalysis: From Enzyme Databases To General Principles. *J. Biol. Inorg. Chem.* **2008,** *13* (8), 1205-1218.

6. Waldron, K. J.; Rutherford, J. C.; Ford, D.; Robinson, N. J., Metalloproteins And Metal Sensing. *Nature* **2009,** *460* (7257), 823-830.

7. Pagadala, N. S.; Syed, K.; Tuszynski, J., Software For Molecular Docking: A Review. *Biophys. Rev.* **2017,** *9* (2), 91-102.

8. Sulimov, V. B.; Kutov, D. C.; Sulimov, A. V., Advances In Docking. *Curr. Med.* **2019,** *26* (42), 7555-7580.

9. Morris, G. M.; Lim-Wilby, M., Molecular Docking. In *Molecular Modeling of Proteins*, Kukol, A., Ed. Humana Press: Totowa, NJ, 2008; pp 365-382.

10.   Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J., Docking And Scoring In Virtual Screening For Drug Discovery: Methods And Applications. *Nat. Rev. Drug Discov.* **2004,** *3* (11), 935-949.

11.   Sciortino, G.; Rodríguez-Guerra Pedregal, J.; Lledós, A.; Garribba, E.; Maréchal, J.-D., Prediction Of The Interaction Of Metallic Moieties With Proteins: An Update For Protein-Ligand Docking Techniques. *J. Comput. Chem.* **2018,** *39* (1), 42-51.

12.   Chen, E.; Swift, R. V.; Alderson, N.; Feher, V. A.; Feng, G.-S.; Amaro, R. E., Computation-Guided Discovery Of Influenza Endonuclease Inhibitors. *ACS Med. Chem. Lett.* **2014,** *5* (1), 61-64.

13.   Englebienne, P.; Fiaux, H.; Kuntz, D. A.; Corbeil, C. R.; Gerber-Lemaire, S.; Rose, D. R.; Moitessier, N., Evaluation Of Docking Programs For Predicting Binding Of Golgi α-Mannosidase II Inhibitors: A Comparison With Crystallography. *Proteins* **2007,** *69* (1), 160-176.

14.   Çınaroğlu, S. S.; Timuçin, E., Comparative Assessment Of Seven Docking Programs On A Nonredundant Metalloprotein Subset Of The PDBbind Refined. *J. Chem. Inf. Model.* **2019,** *59* (9), 3846-3859.

15.   Khandelwal, A.; Lukacova, V.; Comez, D.; Kroll, D. M.; Raha, S.; Balaz, S., A Combination Of Docking, QM/MM Methods, And MD Simulation For Binding Affinity Estimation Of Metalloprotein Ligands. *J. Med. Chem.* **2005,** *48* (17), 5437-5447.

16.   Eriksson, A. E.; Jones, T. A.; Liljas, A., Refined Structure Of Human Carbonic Anhydrase II At 2.0 Å Resolution. *Proteins* **1988,** *4* (4), 274-282.

17.   De Simone, G.; Alterio, V.; Supuran, C. T., Exploiting The Hydrophobic And Hydrophilic Binding Sites For Designing Carbonic Anhydrase Inhibitors. *Expert Opin. Drug Discov.* **2013,** *8* (7), 793-810.

18.   McKenna, R.; Supuran, C. T., Carbonic Anhydrase Inhibitors Drug Design. In *Carbonic Anhydrase: Mechanism, Regulation, Links to Disease, and Industrial Applications*, Frost, S. C.; McKenna, R., Eds. Springer Netherlands: Dordrecht, 2014; pp 291-323.

19.   Fueyo, R.; García, M. A.; Martínez-Balbás, M. A., Jumonji Family Histone Demethylases In Neural Development. *Cell Tissue Res.* **2015,** *359* (1), 87-98.

20.   Shi, Y., Histone Lysine Demethylases: Emerging Roles In Development, Physiology And Disease. *Nat. Rev. Genet.* **2007,** *8* (11), 829-833.

21.   Pedersen, M. T.; Helin, K., Histone Demethylases In Development And Disease. *Trends Cell Biol.* **2010,** *20* (11), 662-671.

22.   Reich, S.; Guilligay, D.; Pflug, A.; Malet, H.; Berger, I.; Crépin, T.; Hart, D.; Lunardi, T.; Nanao, M.; Ruigrok, R. W. H.; Cusack, S., Structural Insight Into Cap-Snatching And RNA Synthesis By Influenza Polymerase. *Nature* **2014,** *516* (7531), 361-366.

23.   Ju, H.; Zhang, J.; Huang, B.; Kang, D.; Huang, B.; Liu, X.; Zhan, P., Inhibitors Of Influenza Virus Polymerase Acidic (PA) Endonuclease: Contemporary Developments And Perspectives. *J. Med. Chem.* **2017,** *60* (9), 3533-3551.

24.   Karges, J.; Stokes, R. W.; Cohen, S. M., Photorelease Of A Metal-Binding Pharmacophore From A Ru(ii) Polypyridine Complex. *Dalton Trans.* **2021,** *50* (8), 2757-2765.

25.   Oda, A.; Tsuchida, K.; Takakura, T.; Yamaotsu, N.; Hirono, S., Comparison of Consensus Scoring Strategies for Evaluating Computational Models of Protein−Ligand Complexes. *J. Chem. Inf. Model.* **2006,** *46* (1), 380-391.

26.   Cheng, T.; Li, X.; Li, Y.; Liu, Z.; Wang, R., Comparative Assessment of Scoring Functions on a Diverse Test Set. *J. Chem. Inf. Model.* **2009,** *49* (4), 1079-1093.

27.   Velázquez-Libera, J. L.; Durán-Verdugo, F.; Valdés-Jiménez, A.; Núñez-Vivanco, G.; Caballero, J., LigRMSD: A Web Server For Automatic Structure Matching And RMSD Calculations Among Identical And Similar Compounds In Protein-Ligand Docking. *Bioinformatics* **2020,** *36* (9), 2912-2914.

28.   Scott, A. D.; Phillips, C.; Alex, A.; Flocco, M.; Bent, A.; Randall, A.; O'Brien, R.; Damian, L.; Jones, L. H., Thermodynamic Optimisation In Drug Discovery: A Case Study Using Carbonic Anhydrase Inhibitors. *ChemMedChem* **2009,** *4* (12), 1985-1989.

29.   Carta, F.; Aggarwal, M.; Maresca, A.; Scozzafava, A.; McKenna, R.; Supuran, C. T., Dithiocarbamates: A New Class Of Carbonic Anhydrase Inhibitors. Crystallographic And kinetic Investigations. *Chem. Commun.* **2012,** *48* (13), 1868-1870.

30.   Martin, D. P.; Blachly, P. G.; Marts, A. R.; Woodruff, T. M.; de Oliveira, C. A. F.; McCammon, J. A.; Tierney, D. L.; Cohen, S. M., 'Unconventional' Coordination Chemistry By Metal Chelating Fragments In A Metalloprotein Active Site. *J. Am. Chem. Soc.* **2014,** *136* (14), 5400-5406.

31.   Glöckner, S.; Heine, A.; Klebe, G., A Proof-of-Concept Fragment Screening Of A Hit-Validated 96-Compounds Library Against Human Carbonic Anhydrase II. *Biomolecules* **2020,** *10* (4), 518.

32.   Rose, N. R.; Ng, S. S.; Mecinović, J.; Liénard, B. M. R.; Bello, S. H.; Sun, Z.; McDonough, M. A.; Oppermann, U.; Schofield, C. J., Inhibitor Scaffolds For 2-Oxoglutarate-Dependent Histone Lysine Demethylases. *J. Med. Chem.* **2008,** *51* (22), 7053-7056.

33.   Che, K. H.; Yue, W. W.; Krojer, T.; Muniz, J. R. C.; Ng, S. S.; Tumber, A.; Daniel, M.; Burgess-Brown, N.; Savitsky, P.; Ugochukwu, E.; Filippakopoulos, P.; Arrowsmith, C.; Weigelt,

J.; Edwards, A.; Bountra, C.; Oppermann, U., *Protein Data Bank* **2010**, DOI: 10.2210/pdb2XXZ/pdb.

34. Walport, L. J.; Hopkinson, R. J.; Vollmar, M.; Madden, S. K.; Gileadi, C.; Oppermann, U.; Schofield, C. J.; Johansson, C., Human UTY(KDM6C) Is A Male-Specific N^ϵ-Methyl Lysyl Demethylase. *J. Biol. Chem.* **2014,** *289* (26), 18302-18313.

35. Małecki, P. H.; Rüger, N.; Roatsch, M.; Krylova, O.; Link, A.; Jung, M.; Heinemann, U.; Weiss, M. S., Structure-Based Screening Of Tetrazolylhydrazide Inhibitors Versus KDM4 Histone Demethylases. *ChemMedChem* **2019,** *14* (21), 1828-1839.

36. DuBois, R. M.; Slavish, P. J.; Baughman, B. M.; Yun, M.-K.; Bao, J.; Webby, R. J.; Webb, T. R.; White, S. W., Structural And Biochemical Basis For Development Of Influenza Virus Inhibitors Targeting The PA Endonuclease. *PLoS Pathog.* **2012,** *8* (8), e1002830.

37. Bauman, J. D.; Patel, D.; Baker, S. F.; Vijayan, R. S. K.; Xiang, A.; Parhi, A. K.; Martínez-Sobrido, L.; LaVoie, E. J.; Das, K.; Arnold, E., Crystallographic Fragment Screening And Structure-Based Optimization Yields A New Class Of Influenza Endonuclease Inhibitors. *ACS Chem. Bio.* **2013,** *8* (11), 2501-2508.

38. Kim, C.-Y.; Chandra, P. P.; Jain, A.; Christianson, D. W., Fluoroaromatic−Fluoroaromatic Interactions Between Inhibitors Bound In The Crystal Lattice Of Human Carbonic Anhydrase II. *J. Am. Chem. Soc.* **2001,** *123* (39), 9620-9627.

39. Grzybowski, B. A.; Ishchenko, A. V.; Kim, C.-Y.; Topalov, G.; Chapman, R.; Christianson, D. W.; Whitesides, G. M.; Shakhnovich, E. I., Combinatorial Computational Method Gives New Picomolar Ligands For A Known Enzyme. *Proc. Natl. Acad. Sci.* **2002,** *99* (3), 1270-1273.

40.   Grüneberg, S.; Stubbs, M. T.; Klebe, G., Successful Virtual Screening For Novel Inhibitors Of Human Carbonic Anhydrase: Strategy And Experimental Confirmation. *J. Med. Chem.* **2002,** *45* (17), 3588-3602.

41.   De Simone, G.; Vitale, R. M.; Di Fiore, A.; Pedone, C.; Scozzafava, A.; Montero, J.-L.; Winum, J.-Y.; Supuran, C. T., Carbonic Anhydrase Inhibitors: Hypoxia-Activatable Sulfonamides Incorporating Disulfide Bonds That Target the Tumor-Associated Isoform IX. *J. Med. Chem.* **2006,** *49* (18), 5544-5551.

42.   Behnke, C. A.; Le Trong, I.; Godden, J. W.; Merritt, E. A.; Teller, D. C.; Bajorath, J.; Stenkamp, R. E., Atomic Resolution Studies Of Carbonic Anhydrase II. *Acta Crystallogr. D* **2010,** *66* (5), 616-627.

43.   Carta, F.; Garaj, V.; Maresca, A.; Wagner, J.; Avvaru, B. S.; Robbins, A. H.; Scozzafava, A.; McKenna, R.; Supuran, C. T., Sulfonamides Incorporating 1,3,5-Triazine Moieties Selectively And Potently Inhibit Carbonic Anhydrase Transmembrane Isoforms IX, XII And XIV Over Cytosolic Isoforms I And II: Solution And X-ray Crystallographic Studies. *Bioorg. Med. Chem.* **2011,** *19* (10), 3105-3119.

44.   Pacchiano, F.; Aggarwal, M.; Avvaru, B. S.; Robbins, A. H.; Scozzafava, A.; McKenna, R.; Supuran, C. T., Selective Hydrophobic Pocket Binding Observed Within The Carbonic Anhydrase II Active Site Accommodate Different 4-Substituted-Ureido-Benzenesulfonamides And Correlate To Inhibitor Potency. *Chem. Commun.* **2010,** *46* (44), 8371-8373.

45.   Biswas, S.; Carta, F.; Scozzafava, A.; McKenna, R.; Supuran, C. T., Structural Effect Of Phenyl Ring Compared To Thiadiazole Based Adamantyl-Sulfonamides On Carbonic Anhydrase Inhibition. *Bioorg. Med. Chem.* **2013,** *21* (8), 2314-2318.

46.   Güzel-Akdemir, Ö.; Biswas, S.; Lastra, K.; McKenna, R.; Supuran, C. T., Structural Study Of The Location Of The Phenyl Tail Of Benzene Sulfonamides And The Effect On Human Carbonic Anhydrase Inhibition. *Bioorg. Med. Chem.* **2013,** *21* (21), 6674-6680.

47.   Buratto, J.; Colombo, C.; Stupfel, M.; Dawson, S. J.; Dolain, C.; Langlois d'Estaintot, B.; Fischer, L.; Granier, T.; Laguerre, M.; Gallois, B.; Huc, I., Structure Of A Complex Formed By A Protein And A Helical Aromatic Oligoamide Foldamer At 2.1 Å Resolution. *Angew. Chem. Int. Ed.* **2014,** *53* (3), 883-887.

48.   Alterio, V.; Tanc, M.; Ivanova, J.; Zalubovskis, R.; Vozny, I.; Monti, S. M.; Di Fiore, A.; De Simone, G.; Supuran, C. T., X-ray Crystallographic And Kinetic Investigations Of 6-Sulfamoyl-Saccharin As A Carbonic Anhydrase Inhibitor. *Org. Biomol. Chem.* **2015,** *13* (13), 4064-4069.

49.   Runtsch, L. S.; Barber, D. M.; Mayer, P.; Groll, M.; Trauner, D.; Broichhagen, J., Azobenzene-Based Inhibitors Of Human Carbonic Anhydrase II. *Beilstein J. Org. Chem.* **2015,** *11*, 1129-1135.

50.   La Regina, G.; Coluccia, A.; Famiglini, V.; Pelliccia, S.; Monti, L.; Vullo, D.; Nuti, E.; Alterio, V.; De Simone, G.; Monti, S. M.; Pan, P.; Parkkila, S.; Supuran, C. T.; Rossello, A.; Silvestri, R., Discovery Of 1,1′-Biphenyl-4-Sulfonamides As A New Class Of Potent And Selective Carbonic Anhydrase XIV Inhibitors. *J. Med. Chem.* **2015,** *58* (21), 8564-8572.

51.   Nocentini, A.; Ferraroni, M.; Carta, F.; Ceruso, M.; Gratteri, P.; Lanzi, C.; Masini, E.; Supuran, C. T., Benzenesulfonamides Incorporating Flexible Triazole Moieties Are Highly Effective Carbonic Anhydrase Inhibitors: Synthesis And Kinetic, Crystallographic, Computational, And Intraocular Pressure Lowering Investigations. *J. Med. Chem.* **2016,** *59* (23), 10692-10704.

52. Ferraroni, M.; Cornelio, B.; Sapi, J.; Supuran, C. T.; Scozzafava, A., Sulfonamide Carbonic Anhydrase Inhibitors: Zinc Coordination And Tail Effects Influence Inhibitory Efficacy And Selectivity For Different Isoforms. *Inorg. Chim. Acta* **2018,** *470*, 128-132.

53. Alterio, V.; Esposito, D.; Monti, S. M.; Supuran, C. T.; De Simone, G., Crystal Structure Of The Human Carbonic Anhydrase II Adduct With 1-(4-sulfamoylphenyl-ethyl)-2,4,6-triphenylpyridinium Perchlorate, A Membrane-Impermeant, Isoform Selective Inhibitor. *J. Enzyme Inhib. Med. Chem.* **2018,** *33* (1), 151-157.

54. Glöckner, S.; Ngo, K.; Sager, C. P.; Hüfner-Wulsdorf, T.; Heine, A.; Klebe, G., Conformational Changes In Alkyl Chains Determine The Thermodynamic And Kinetic Binding Profiles Of Carbonic Anhydrase Inhibitors. *ACS Chem. Biol.* **2020,** *15* (3), 675-685.

55. Nocentini, A.; Alterio, V.; Bua, S.; Micheli, L.; Esposito, D.; Buonanno, M.; Bartolucci, G.; Osman, S. M.; Alothman, Z. A.; Cirilli, R.; Pierini, M.; Monti, S. M.; Di Cesare Mannelli, L.; Gratteri, P.; Ghelardini, C.; De Simone, G.; Supuran, C. T., Phenyl(thio)phosphon(amid)ate Benzenesulfonamides As Potent And Selective Inhibitors Of Human Carbonic Anhydrases II And VII Counteract Allodynia In A Mouse Model Of Oxaliplatin-Induced Neuropathy. *J. Med. Chem.* **2020,** *63* (10), 5185-5200.

56. Chang, K.-H.; King, O. N. F.; Tumber, A.; Woon, E. C. Y.; Heightman, T. D.; McDonough, M. A.; Schofield, C. J.; Rose, N. R., Inhibition Of Histone Demethylases By 4-Carboxy-2,2′-Bipyridyl Compounds. *ChemMedChem* **2011,** *6* (5), 759-764.

57. King, O. N. F.; Krojer, T.; Arrowsmith, C. H.; Edwards, A.; Bountra, C.; McDonough, M. A.; Schofield, C. J., Crystal Structure Of JMJD2A Complexed With Inhibitor. to be published, PDB Code: 4GD4.

58.   England, K. S.; Tumber, A.; Krojer, T.; Scozzafava, G.; Ng, S. S.; Daniel, M.; Szykowska, A.; Che, K.; von Delft, F.; Burgess-Brown, N. A.; Kawamura, A.; Schofield, C. J.; Brennan, P. E., Optimisation Of A Triazolopyridine Based Histone Demethylase Inhibitor Yields A Potent And Selective KDM2A (FBXL11) Inhibitor. *Med. Chem. Commun.* **2014,** *5* (12), 1879-1886.

59.   Korczynska, M.; Le, D. D.; Younger, N.; Gregori-Puigjané, E.; Tumber, A.; Krojer, T.; Velupillai, S.; Gileadi, C.; Nowak, R. P.; Iwasa, E.; Pollock, S. B.; Ortiz Torres, I.; Oppermann, U.; Shoichet, B. K.; Fujimori, D. G., Docking And Linking Of Fragments To Discover Jumonji Histone Demethylase Inhibitors. *J. Med. Chem.* **2016,** *59* (4), 1580-1598.

60.   Roatsch, M.; Robaa, D.; Pippel, M.; Nettleship, J. E.; Reddivari, Y.; Bird, L. E.; Hoffmann, I.; Franz, H.; Owens, R. J.; Schüle, R.; Flaig, R.; Sippl, W.; Jung, M., Substituted 2-(2-aminopyrimidin-4-yl)pyridine-4-carboxylates As Potent Inhibitors Of JumonjiC Domain-Containing Histone Demethylases. *Future Med. Chem.* **2016,** *8* (13), 1553-1571.

61.   Bavetsias, V.; Lanigan, R. M.; Ruda, G. F.; Atrash, B.; McLaughlin, M. G.; Tumber, A.; Mok, N. Y.; Le Bihan, Y.-V.; Dempster, S.; Boxall, K. J.; Jeganathan, F.; Hatch, S. B.; Savitsky, P.; Velupillai, S.; Krojer, T.; England, K. S.; Sejberg, J.; Thai, C.; Donovan, A.; Pal, A.; Scozzafava, G.; Bennett, J. M.; Kawamura, A.; Johansson, C.; Szykowska, A.; Gileadi, C.; Burgess-Brown, N. A.; von Delft, F.; Oppermann, U.; Walters, Z.; Shipley, J.; Raynaud, F. I.; Westaway, S. M.; Prinjha, R. K.; Fedorov, O.; Burke, R.; Schofield, C. J.; Westwood, I. M.; Bountra, C.; Müller, S.; van Montfort, R. L. M.; Brennan, P. E.; Blagg, J., 8-Substituted Pyrido[3,4-d]pyrimidin-4(3H)-one Derivatives As Potent, Cell Permeable, KDM4 (JMJD2) And KDM5 (JARID1) Histone Lysine Demethylase Inhibitors. *J. Med. Chem.* **2016,** *59* (4), 1388-1409.

62.   Nie, Z.; Shi, L.; Lai, C.; O'Connell, S. M.; Xu, J.; Stansfield, R. K.; Hosfield, D. J.; Veal, J. M.; Stafford, J. A., Structure-Based Design And Discovery Of Potent And Selective KDM5 Inhibitors. *Bioorg. Med. Chem. Lett.* **2018,** *28* (9), 1490-1494.

63.   Chen, Y. K.; Bonaldi, T.; Cuomo, A.; Del Rosario, J. R.; Hosfield, D. J.; Kanouni, T.; Kao, S.-c.; Lai, C.; Lobo, N. A.; Matuszkiewicz, J.; McGeehan, A.; O'Connell, S. M.; Shi, L.; Stafford, J. A.; Stansfield, R. K.; Veal, J. M.; Weiss, M. S.; Yuen, N. Y.; Wallace, M. B., Design Of KDM4 Inhibitors With Antiproliferative Effects In Cancer Models. *ACS Med. Chem. Lett.* **2017,** *8* (8), 869-874.

64.   Credille, C. V.; Dick, B. L.; Morrison, C. N.; Stokes, R. W.; Adamek, R. N.; Wu, N. C.; Wilson, I. A.; Cohen, S. M., Structure–Activity Relationships In Metal-Binding Pharmacophores For Influenza Endonuclease. *J. Med. Chem.* **2018,** *61* (22), 10206-10217.

65.   Kirton, S. B.; Murray, C. W.; Verdonk, M. L.; Taylor, R. D., Prediction Of Binding Modes For Ligands In The Cytochromes P450 And Other Heme-Containing Proteins. *Proteins* **2005,** *58* (4), 836-844.

66.   Brzezinski, D.; Dauter, Z.; Minor, W.; Jaskolski, M., On The Evolution Of The Quality Of Macromolecular Models In The PDB. *FEBS J.* **2020,** *287* (13), 2685-2698.

67.   Angeli, A.; Tanini, D.; Capperucci, A.; Malevolti, G.; Turco, F.; Ferraroni, M.; Supuran, C. T., Synthesis Of Different Thio-Scaffolds Bearing Sulfonamide With Subnanomolar Carbonic Anhydrase II And IX Inhibitory Properties And X-ray Investigations For Their Inhibitory Mechanism. *Bioorg. Chem.* **2018,** *81*, 642-648.

68.   Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J., Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990,** *215* (3), 403-410.

Chapter 3: Carboxylic Acid Isostere Derivatives of Hydroxypyridinones as Core Scaffolds for Influenza Endonuclease Inhibition

3.1 Introduction

In addition to the computational prediction of MBP binding, which was discussed in Chapter 2, and will be further discussed in this chapter, another goal of this thesis is to investigate MBP binding modes experimentally, which provides important structural information in early FBDD campaigns. Chapter 1.5 discussed the application of isosteres to metalloenzyme FBDD, and here the binding of several MBP isosteres is evaluated experimentally. As a model system to study the binding of isosteres, the RNA-dependent RNA polymerase acidic N-terminal (PA$_N$) endonuclease[1] of the influenza virus was chosen, as it is both an important therapeutic target and would benefit from the development of isosteres that can address some of the PK liabilities associated with recently developed inhibitors.[2] PA$_N$ represents an especially attractive target due to its critical role in viral replication,[3] its high conservation, its lack of a human analog,[4] and the high mortality of seasonal influenza epidemics.[5] PA$_N$ contains a dinuclear Mg^{2+} or Mn^{2+} active site that is highly conserved and has been the target of several recent drug discovery efforts,[6-8] including the approval of baloxavir marboxil by the United States Food and Drug Administration (FDA).[9] Some efforts have used metalloenzyme-focused FBDD to identify potent inhibitors of PA$_N$.[2, 10-12]

Several of these FBDD campaigns have identified MBPs that utilize a triad of oxygen donors, including a free carboxylic acid, that binds the metal ions, making these fragments very polar and non-ideal starting points for the development of novel therapeutics. Among some previously characterized MBPs, a carboxylic acid modified hydroxypyridinone was identified, achieving IC$_{50}$ values as low as ~17 nM against PA$_N$.[2] Despite this impressive activity, this MBP (and elaborated inhibitors derived from it) were found to be only moderately active in influenza infected cells (Figure 3.1). This disparity between antiviral activity and enzymatic inhibition has

also been observed among structurally related compounds, likely due to low membrane permeability originating from the high polarity of these molecules.[8] It is well-known that carboxylic acid moieties are associated with pharmacological liabilities, which can limit their use in drug development.[13-15] The carboxylic acid functional group can be replaced by a different moiety with similar biochemical interactions, but different physiochemical properties via isosteric replacement to try to address possible shortcomings of MBP **1**.[16, 17]
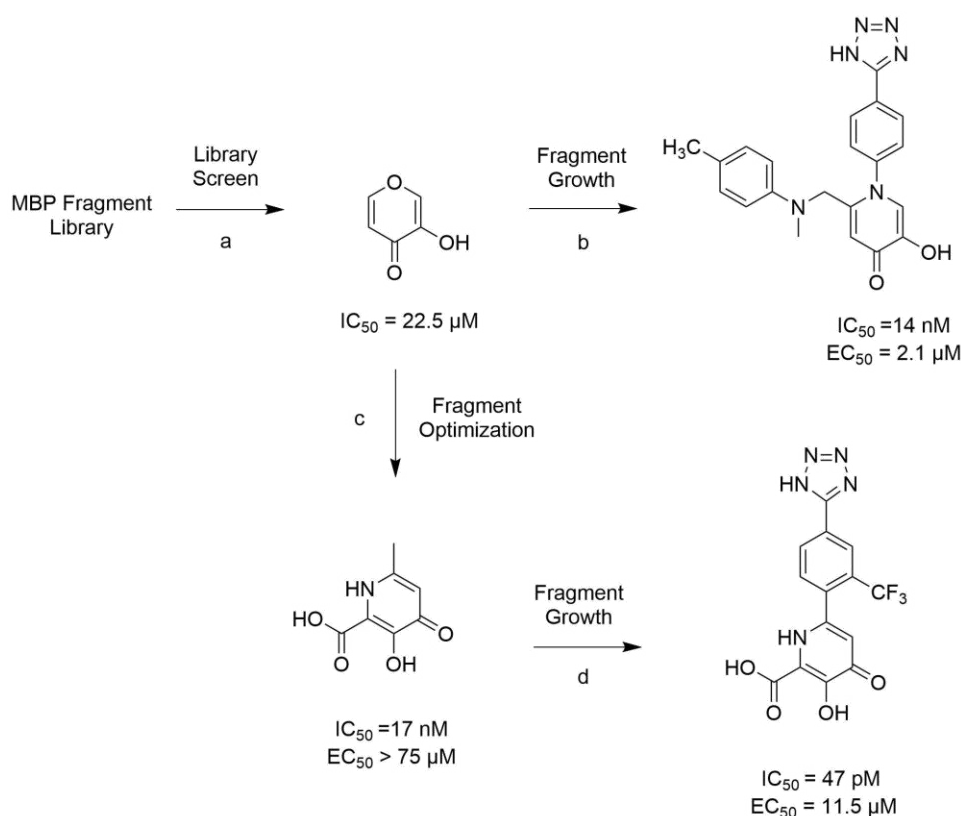
Figure 3.1. Previous studies in our group have explored the development of PA_N endonuclease inhibitors. (A) From an MBP library screen, pyromeconic acid was identified as having low micromolar activity against PA_N. (B) Further SAR investigations resulted in a functionalized inhibitor which demonstrated low nanomolar efficacy in enzymatic studies.[10] (C) Optimization of the MBP resulted in a hydroxypyridinone fragment which exhibited high inhibitory activity in an enzymatic assay, but which demonstrated poor efficacy in cellular studies.[11]   Finally, functionalization of the hydroxypyridinone MBP resulted in a highly potent inhibitor, which exhibited low picomolar activity in enzymatic studies.[2] Unfortunately, despite improved activity, equivalent potency was not observed in cellular models.

To address these limitations, the design and experimental evaluation of novel metal-binding isosteres (MBIs) is reported (Figure 3.2). Nine MBIs (**3-11**) of hydroxypyridinone MBP **1** were synthesized, which installed isosteric replacements for the carboxylic acid moiety (Figure 3.2). These MBIs also possess a bromine atom in the 6-position of the ring to aid in crystallography (i.e., via heavy atom substitution and anomalous scattering) and to serve as a synthetic handle for future compound elaboration.[2] To account for the electronic and lipophilic effects of adding this bromine substituent, a direct analog of **1** was synthesized that contains a bromine atom (**2**, Figure 3.2). Table 3.S1 lists the calculated cLogP values of these fragments compared to analogues substituted with -H, -CH$_3$, or -Ph groups. The MBIs were found to possess good inhibition of PA$_N$ endonuclease, with many displaying half maximal inhibitory concentration (IC$_{50}$) values in the low nanomolar range, comparable to the parent carboxylic acid MBP (compound **1**, Figure 3.2). Using X-ray crystallography, the binding of these compounds in the active site of PA$_N$ was examined and compared with the predicted docking results. While many compounds bound as expected (based on prior findings),[2, 11] some isosteres exhibited somewhat unexpected binding interactions. A preliminary analysis of the physical properties, such as clogP and calculated pK_a, revealed that the isosteres demonstrate tunability across these properties, potentially circumventing the limitations of the previously identified carboxylic acid scaffold.
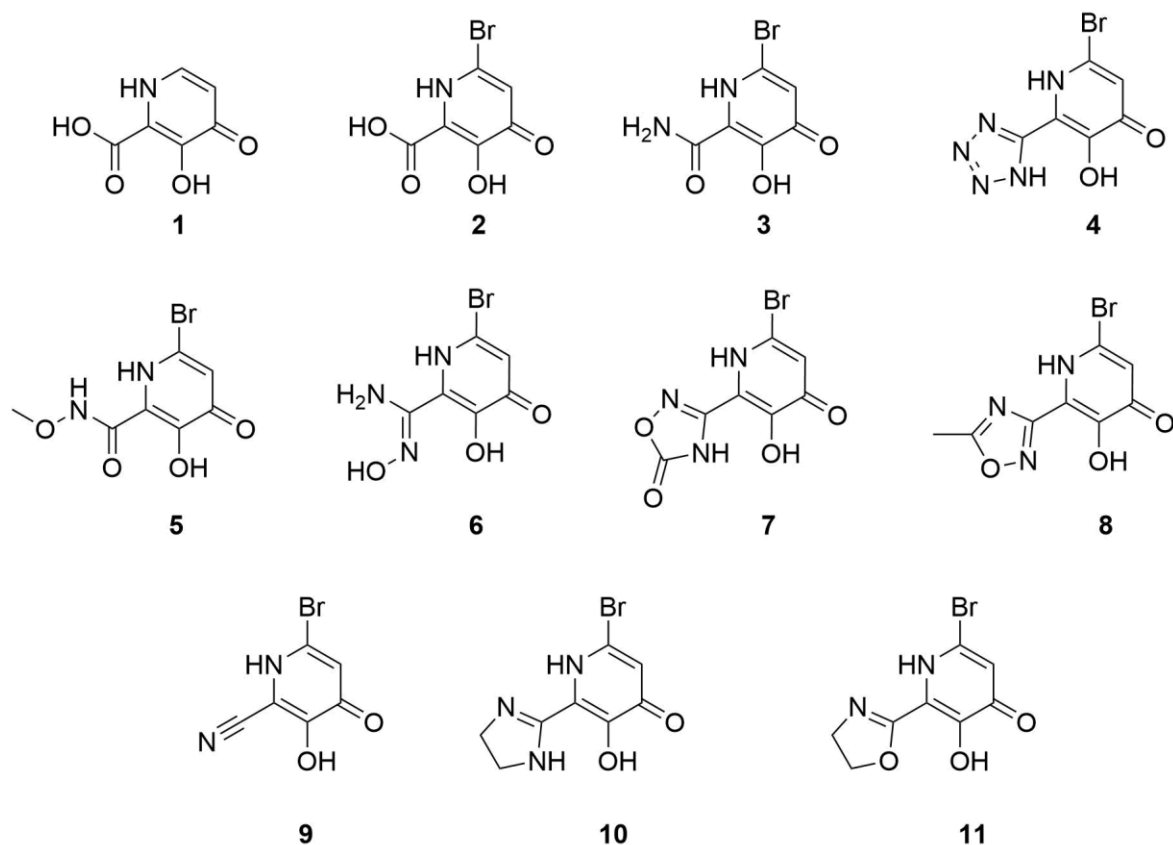
Figure 3.2. Chemical structures of metal-binding isosteres (MBIs) investigated in this study.

3.2 Compound Synthesis

All bromine-containing compounds (**2-11**) were synthesized from precursor **12**, 3-hydroxypicolinonitrile (Scheme 3.1). Commercially available **12** can be dibrominated selectively in the 4- and 6-positions upon treatment with elemental bromine to afford **13** in high yield. An oxygen atom that will later participate as a Lewis base to bind the Mn^{2+} center can be introduced via S_NAr chemistry. This nucleophilic addition results in the selective introduction of benzyl alcohol in the 4-position to afford intermediate **14** in good yield. Compound **14** serves as a crucial intermediate for the isosteres evaluated in this study. The versatility of the nitrile functional group allows for facile conversion to several functional groups that have metal-binding capabilities, but

with electronic properties that may mitigate some of the difficulties associated with carboxylic acids. Hydrolysis of **14** under basic conditions yields the benzyl-protected compounds **2a** and **3a**, which can be deprotected under acidic conditions to afford compounds **2** and **3**, the acid- and amide-functionalized molecules, respectively. Compound **2** can be further functionalized to incorporate an *N*-methoxy carboxamide group (**5**) by conversion to the acid chloride and subsequent treatment with methoxyamine hydrochloride. A click reaction between compound **14** and sodium azide results in the formation of a tetrazole, a commonly employed carboxylic acid isostere. This compound, (**4a**) can be deprotected to afford the tridentate MBI compound **4**. Treatment of compound **14** with catalytic zinc chloride and ethanolamine or ethylenediamine results in the partially saturated heterocycles **11a** and **12a**. Saturated heterocycles can be advantageous as they tend to have higher lipophilicity, solubility, and three-dimensionality than their unsaturated counterparts.[18] Deprotection of these compounds with HCl results in the formation of the partially saturated imidazoline (**10**) and oxazoline (**11**). The nitrile group in compound **14** can also be transformed into an *N*-hydroxy amidine functional group when treated with hydroxylamine hydrochloride and triethylamine, followed by debenzylation to afford compound **6**. The structure of compound 6 was unambiguously confirmed by obtaining a crystal structure of compound **6** bound to hydrotris(5,3-methylphenylpyrazolyl)borate (Tp[Ph,Me]) (Figure 3.S1). Finally, simple deprotection of compound **14** under acidic conditions results in the nitrile-containing MBI **9**. In some cases, protection of the phenol in compound **14** was beneficial. Treatment of **14** with benzyl bromide and potassium carbonate resulted in the formation of **15** in good yield. The protected *N*-hydroxy amide intermediate was synthesized using similar conditions to **6a**, though longer reaction times were employed. Compound **16** can be acylated upon the addition of acetic anhydride to afford compound **17**, which when exposed to heat and a mixture of

hydrochloric acid, acetic acid, and trifluoroacetic acid, undergoes both debenzylation and an acid-promoted cyclization to afford oxadiazole **8** in a single step. The structure of compound **8** was also unambiguously confirmed by obtaining a crystal structure of **8** in complex with hydrotris(5,3-methylphenylpyrazolyl)borate (Figure 3.S2). Additionally, treatment of compound **16** with carbonyldiimidazole (CDI) results in the formation of the oxadiazolone-functionalized molecule **7a**, which can be deprotected to afford compound **7**.
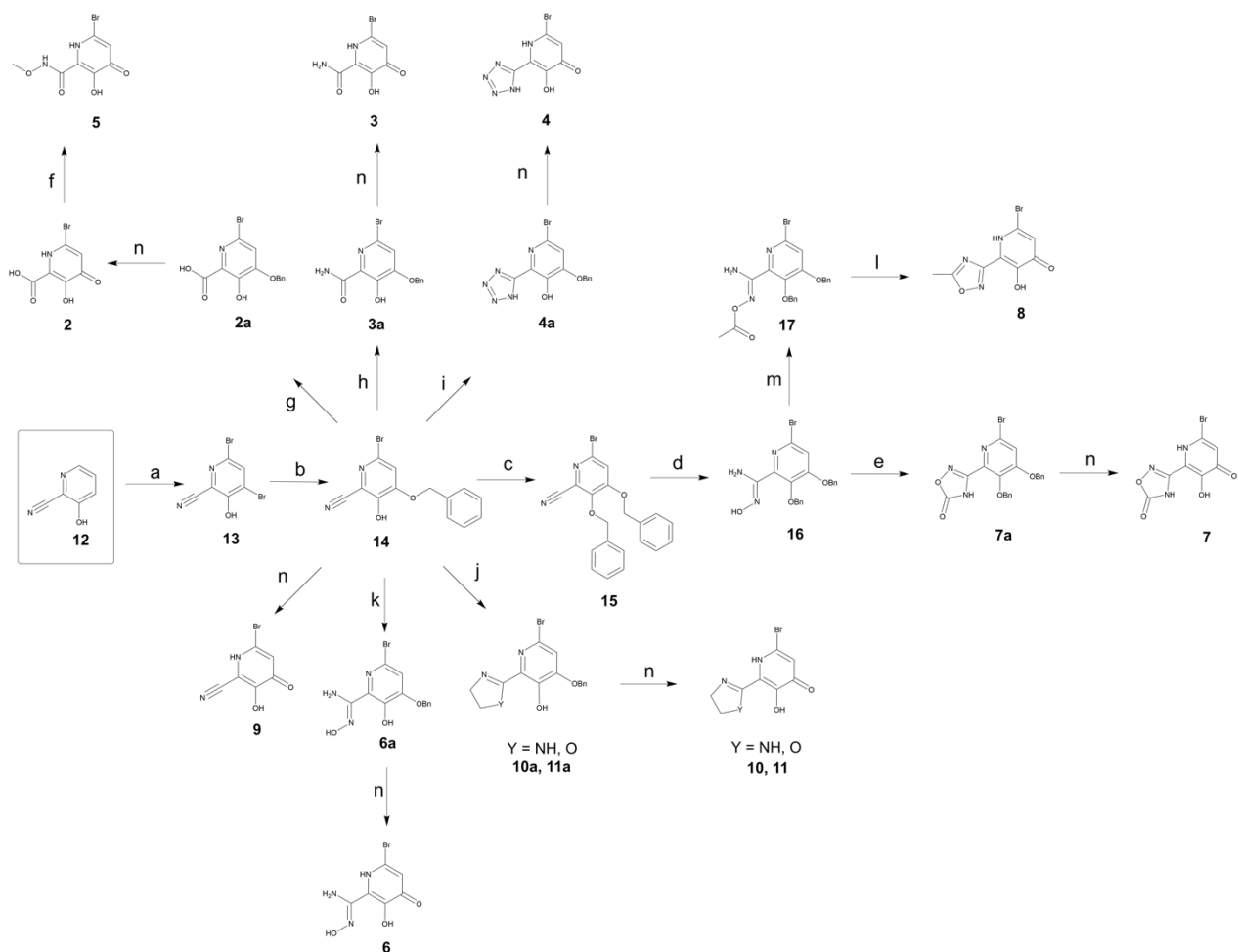
Scheme 3.1. Synthesis of MBIs. Reagents and Conditions: (a) NaOAc, H$_2$O:MeOH, Br$_2$ 0-25 °C; (b) NaH, BnOH, DMSO, 60 °C, 18 h; (c) BnBr, K$_2$CO$_3$, DMF, 45 °C, 19 h; (d) HONH$_2$·HCl, NEt$_3$, EtOH, 50 °C, 15 h; (e) CDI, NEt$_3$, DMF, 80 °C, 8 h; (f) i. SOCl$_2$, DCM, reflux, 12 h. ii. Methoxyamine hydrochloride, NEt$_3$, DCM, -78 °C – 40 °C, 48 h; (g) NaOH, EtOH, 100 °C, 20 h; (h) KOH, EtOH:H2O, 100 °C, 16 h; (i) NaN$_3$, NH$_4$Cl, DMF, 115 °C, 4 h; (j) Ligand, ZnCl$_2$, Toluene, 90 °C – 130 °C, 16-20 h; (k) HONH$_2$·HCl, NEt$_3$, EtOH, 50 °C 1 h; (l) HCl:AcOH:TFA (5:5:1), 50 °C, 48 h; (m) Ac$_2$O, DBU, THF, 0 °C – 80 °C, 8 h; (n) HCl, 25 – 100 °C, 12 – 48 h.

3.3 Density Functional Theory Calculations and Molecular Modeling

Previous studies have correlated the acidity of the phenolic proton (pK_a) of the hydroxypyridinone fragment with the ability to inhibit the activity of PA$_N$.[11] In addition, changes in pK_a can be telling indicators of changes in the electronic structure of the heterocycle.

Accordingly, pK_a values of the MBIs were calculated to probe the electronic differences introduced by the isostere groups. The pK_a value of the phenolic proton was estimated using density functional theory (DFT) calculations. More precisely, the energetic states within a thermodynamic cycle of the proton transfer reaction were calculated and the resulting free energy correlated to the pK_a values as has been previously described (see Supporting Information for details).[19] Compound **2** was predicted to have a pK_a value of 9.4. Using a pH-metric titration, the measured pK_a value of compound **2** was measured at 9.79 ± 0.01, showing reasonable agreement with the DFT calculation. The isosteric replacement of this functional group was identified as an influencing factor of the acidity of the phenolic proton (a full list of predicted and several additional measured pK_a values, all of which show close agreement, is provided in Table 3.S2. The isostere modification of the carbonyl group gave calculated pK_a values ranging from 8.1 to 12.6 (Table 3.2) indicating that changes in the 2-position of the hydroxypyridinone ring can have large effects on the electronics of the ring system and is a useful means to tune the electronic properties of the ring.

In addition to the acidity of the phenolic proton, the coordination and active site geometries can also influence metal ion coordination and affect inhibitory activity. Accordingly, the ability of the MBIs to coordinate the dinuclear Mn^{2+} active site of PA_N, was predicted using the docking methodology from Chapter 2.[20] The MBIs were docked using a genetic algorithm in the Genetic Optimization for Ligand Docking (GOLD) software package (Figure 3.3). As expected, the hydroxypyridinone core, which is found in all tested compounds, was found to bind with a coordination geometry similar to that which has been previously observed.[2] The phenolic moiety was coordinated to both metal ions and the carbonyl group bound axially to the octahedral Mn_1. Select poses are shown in Figure 3.3, and additional poses are shown in Figure 3.S4). Carbonyl-

containing compounds **3** and **5** did not influence the binding pose of the MBP, with the oxygen atom of the carbonyl moiety coordinated to the active site metal ion. In comparison, the isosteres with heterocycles (**4, 8,** and **10**) indicated a metal interaction with a basic nitrogen atom in the heterocycle, generating a tridentate inhibitor. Curiously, the docking algorithm suggested a twist in the oxadiazolone and oxadiazole rings (**7** and **8**) relative to the hydroxypyridinone core. Otherwise, the isosteric replacement did not drastically alter the predicted binding pose of the MBIs (when compared to MBP **1**), suggesting these molecules may be useful in the development of PA_N inhibitors.
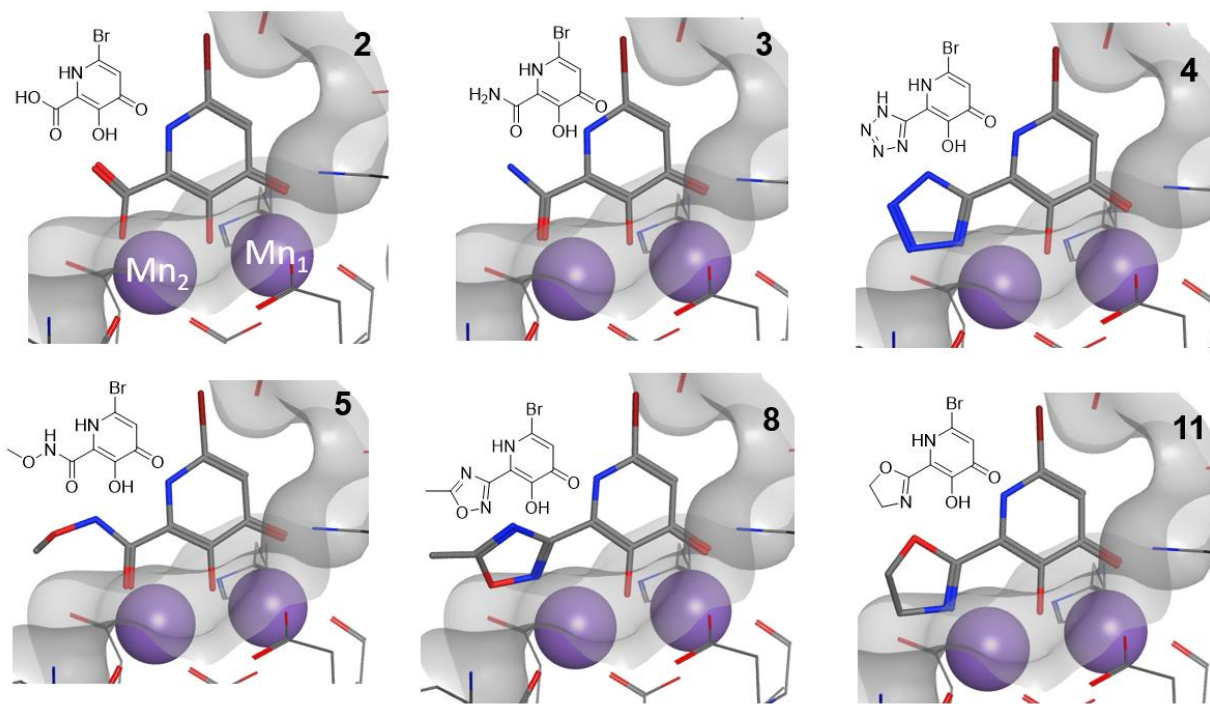
Figure 3.3. Computationally predicted binding poses of representative MBIs **2-5**, **8**, and **11**. All MBIs demonstrated octahedral coordination to the dinuclear Mn^{2+} active site.

77

3.4 X-ray Crystallography

To determine the binding mode of the MBIs in the PA_N active site, several compounds were crystallized with the enzyme and the co-crystal structures determined. Tables of collection and refinement statistics for X-ray structures of PA_N bound with compounds are provided in the Supporting Information (Tables 3.S3, 3.S4). In the ligand-free structure of PA_N, the Mn^{2+} cations are coordinated by His41, Glu80, Asp108, Glu 119, and Ile120 and five water molecules, producing an octahedral coordination geometry at each metal center. As anticipated, based on the structure of compound **2** to PA_N, the addition of a bromine atom to the MBIs produced no apparent change in the binding to the active site metal ions (compare structures of **1** and **2**, Figure 3.4) and aided in ligand placement within the electron density map. Both compounds **1** and **2** coordinated by bridging the metal ions through the phenolate donor. This provides a triad of oxygen donor atoms (including oxygen donor atoms from the ketone and carboxylic acid groups), creating a 5-membered chelate ring at Mn_1 and a 6-membered chelate ring at Mn_2. Compounds **1** and **2** also displayed hydrogen bonding between the ketone donor from the hydroxypyridinone ring and Lys134.[21]

Previous crystallographic studies have focused on compounds containing a triad of chelating oxygen atoms that bind in a tridentate fashion, displacing three of the coordinating water molecules.[2, 11] Compounds **1, 2, 3,** and **5** present the same oxygen triad as previously reported compounds and displayed similar binding poses, coordinating Mn_1 through the ketone and Mn_2 through the carboxylate, with the hydroxyl group bridging Mn_1 and Mn_2 (Figure 3.4). [21] The binding and inhibitory activity (vide infra) of these compounds are comparable to the parent MBP **1**. The average distance from Mn_2 to the carboxylic acid isosteric group in each of these structures is 2.2 Å. A list of metal-donor atom distances is provided in Table S3.5. Interestingly, several

nitrogen-containing isosteres were found to have a bridging water molecule between the nitrogen of the hydroxypyridinone core and a nitrogen from the isosteric groups (see compounds **3**, **4**, **5**, and **10**, Figure 3.4). This interaction is reminiscent of the carbamoyl pyridine bicycle in Baloxavir,[8] but creates a cyclized, rigid structure by hydrogen bonding with water, rather than a synthetic, covalent ring structure. This pseudo five-membered ring may be a useful in future inhibitor design, as it provides conformational stabilization without the necessity to form a bicycle, affording additional synthetic opportunities to develop interactions in the pocket away from the metals.

Among measured structures, compounds **4, 7, 8, 9,** and **10** break with the chelating oxygen triad motif by replacing the carboxylic acid oxygen with a nitrogen atom. As expected, Mn_1 is bound between the phenol and the ketone of the hydroxypyridinone core, but unlike previous MBIs with a third oxygen donor, rotations of varying degrees were observed with the heterocyclic MBIs when bound to Mn_2. Among the most interesting observations, the oxadiazolone ring of compound **8** is noticeably rotated out of the plane of the hydroxypyridinone core. A possible explanation for the rotation is the driving force to maintain an octahedral coordination geometry at both metal centers. To preserve an acceptable angle/distance between the metal and the donor atom at both metal sites, a rotation of the heterocyclic isostere may occur, the metal-ligand distances may be adjusted, or a combination thereof.

To further examine this hypothesis, octahedral distortion parameters ζ and Σ were calculated for compound **8** (relative to Mn2) and compared with the distortion parameters of compound **3** and the ligand free structure, which presented the most ideal octahedral geometries. To visualize the changes in octahedral character, Octadist,[22] a free octahedral distortion analysis software, was used (Figure 3.S3). The ζ and Σ values are general octahedral distortion parameters

used to characterize deviations from ideal octahedral geometry and are used here to evaluate the isostere binding geometry to Mn_2.[22] The value of ζ is the average of the sum of the deviation of 6 unique metal–ligand bond lengths around the central metal atom (d_i) from the average value (d_{mean}).[23] The value of Σ is the sum of the deviation of 12 unique *cis* ligand–metal–ligand angles (ϕ_i) from 90°.[24, 25] These distortion values were calculated for compound **8**, although it should be noted that these values do not take into account the resolution of the crystal structure, and should be weighed accordingly. The value for ζ is very small, 0.574 Å, suggesting only very small changes to ligand-metal distances for each of the six donor atoms. The Σ value is also rather small, 90.81°, with the largest deviations from 90° resultant from contributions of the more rigid angles associated with Glu80 and Asp108. When these distortion parameters are calculated for a structure of compound **8** that is forced into an idealized co-planar ligand geometry, the ζ value increases to 0.606 Å and Σ increases to 98.347°, suggesting that the planar MBI structure enforces an inferior octahedral geometry at the metal centers compared to the experimental structure. Comparing both the rotated and planar structures of compound **8** to compound **3** (0.357 Å, 46.669°) and the ligand free structure (0.685 Å, 44.571°), both conformations of compound **8** exhibit poorer octahedral character.

The tetrazole in compound **4** displays a slight twist, perhaps due to the same reason as compound **8**, though the rotation appears less pronounced. Interestingly, instead of incorporating a more distinct twist, this structure adopts different metal-ligand binding distances. The distance between Mn_1 and the ketone functional group is 2.3 Å, while the distance between the tetrazole and Mn_2 is 1.8 Å, both the longest and shortest metal-donor distances observed among these inhibitors, respectively. The co-crystal structure of **6** with PA_N gave ambiguous ligand density and could not be satisfactorily refined. To further probe the metal-binding abilities of the

molecule, $[(Tp^{Ph,Me})Zn(\mathbf{6})]$ was synthesized and examined by small molecule X-ray crystallography. The structure revealed multiple metal coordination modes for **6** (Figure 3.S1, Table 3.S7), which may explain the inability to resolve the structure of this compound in PA_N. The average distance from the metal cation in each of these nitrogen donating structures is 2.2 Å, matching the donor-metal distance in oxygen triad MBIs, furthering the argument that the rotations of compounds **4** and **8** result from attempting to optimize metal binding interactions.

Overall, MBIs following the same donor oxygen triad displayed similar metal binding activity and binding poses to previously reported structures. MBIs presenting a nitrogen donor atom instead of the carboxylate oxygen have a greater variability in metal binding behavior and pose. Heterocyclic nitrogen MBIs were found to have varying degrees of internal ligand rotation, most notably with compound **8**, which is most likely explained by the need to optimize the octahedral coordination geometry at the Mn_2 metal center. Gratifyingly, the experimental structures showed close agreement to the predicted structures calculated in Section 3.3, with all overlayed structures showing RMSD values of less than 1 Å (Figure 3.S5). This finding is an additional validation of the use of computational metalloenzyme FBDD to predict MBP binding to active site metals.
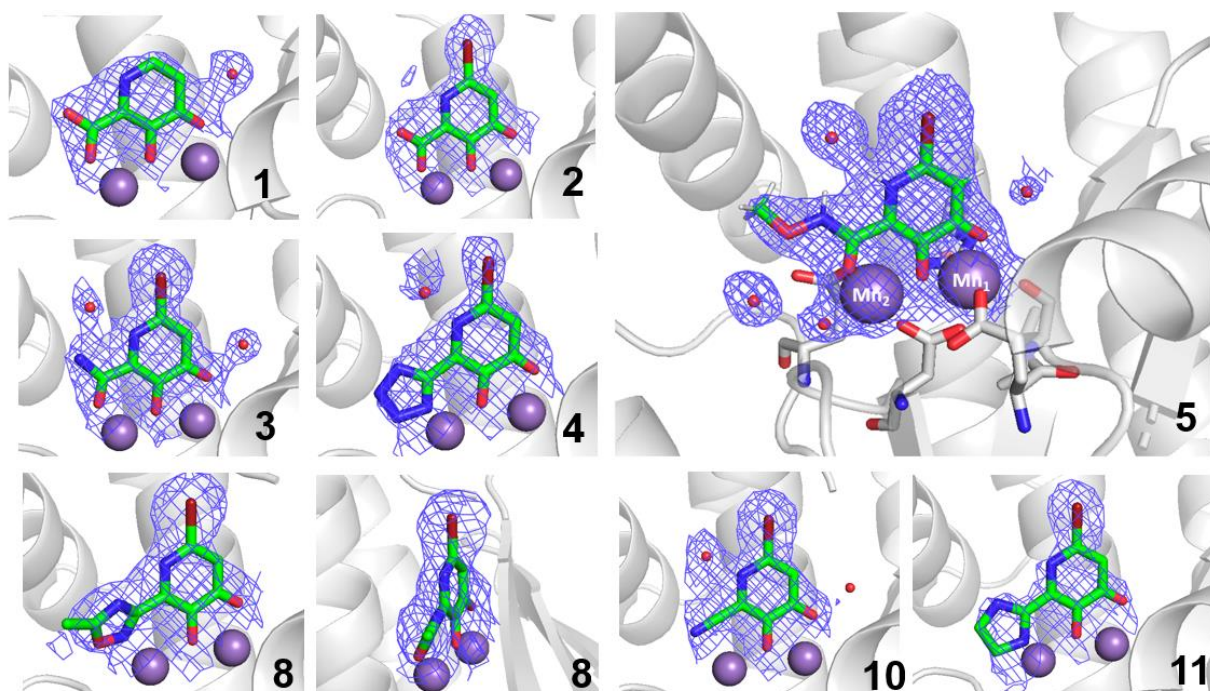
Figure 3.4. Co-crystal structures of PA_N endonuclease with compounds **1-5**, **8**, **9**, and **10**. The large panel depicts MBI **5** coordinating Mn_1 through the exocyclic carbonyl oxygen atom, Mn_2 through the carboxylate group, and the hydroxyl donor bridging between Mn_1 and Mn_2. The protein backbone is shown as a gray cartoon. Mn^{2+} ions and water molecules are shown as purple and red spheres, respectively. Residues coordinating the Mn^{2+} ions are shown in sticks. Atom colors are: carbon (green MBI, gray protein), oxygen (red), nitrogen (blue), bromine (dark red). The electron density of each MBI and the Mn^{2+} ions are displayed as a blue mesh. Mesh is $2F_o$-F_c contoured at 2σ.

3.5 Biochemical Evaluation
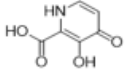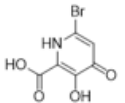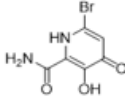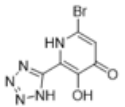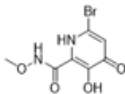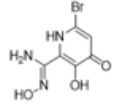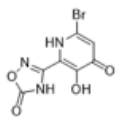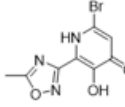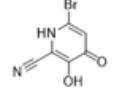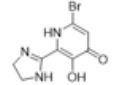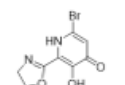
To determine the efficacy of these molecules against PA_N, the MBIs were evaluated using a fluorescence-quenching-based nuclease activity assay.[26] In this assay, a single-stranded, 17-mer DNA oligonucleotide was labeled with a fluorescence emitter and a fluorescence quencher (see Experimental section for details). When PA_N is active, the nucleotide is cleaved, resulting in a fluorescent signal, whereas when PA_N is inhibited, the fluorescence signal is quenched. Using this assay, dose response curves were obtained to determine the half-maximal inhibitory concentration (IC_{50}) values of these compounds. Importantly, IC_{50} values are strongly dependent on the buffer, pH, enzyme concentration, and substrate used, and as such may vary somewhat with assay

conditions.[27] Here, a protein concentration of 25 nM was employed.[6] As previously identified hydroxypyridinones have been shown to be tight binders of PA_N,[2] it is expected that molecules with good metal-binding motifs will demonstrate low IC_{50} values under these assay conditions.[28] The results of the dose response experiment are shown in Table 3.1 as IC_{50} values, with the error displayed as a 95% confidence interval (CI).

Compound **2** was used as a control to determine whether the addition of a bromine atom negatively affected the metal-binding ability of the MBP. Fortunately, compounds **1** and **2** showed similar inhibition activity (Table 3.1), indicating that addition of bromide did not substantially impact metal-binding relative to the parent compound. Due to their small nature and high inhibitory activity, both compounds display extremely high ligand efficiencies (LE) (Table 3.1). Despite conversion from a potentially anionic carboxylate to a neutral carboxyamide, compound **3** showed little difference in activity when compared with compounds **1** or **2**. Compound **4** showed comparable activity to the parent structure, likely due to the shorter metal-ligand bond distances (Table 3.S5). This is perhaps unsurprising as tetrazoles have been widely investigated as isosteres for carboxylic acids due to their resistance to biological degradation, lack of biological activity, and metabolic resistance.[29] Compound **5** demonstrated moderate inhibition, suggesting somewhat poorer metal-binding ability, perhaps due to the electron withdrawing nature of the methyoxy group. Compound **7** demonstrated the most potent IC_{50} value of the evaluated isosteres; however, a crystal structure was not obtained. Compound **8**, functionalized with a methyl-oxadiazole, showed good inhibition, indicating that the rotation of the heterocycle may contribute to improved metal binding via optimizing the metal-ligand geometry. While compound **8** does bind favorably, it demonstrates higher distortion parameters than compound **3**, possibly accounting for the decreased activity of **8** compared to **3**. Compound **9** showed relatively poor inhibition, due to the

loss of a strong interaction with Mn_2, which cannot be provided by the nitrile isostere in this context. However, the inhibition observed for **9** is higher than what might be expected for a compound only interacting with one of the active site metal ions. Previous structures have shown bidentate hydroxypyridinones coordinating to Mn_1, with the ketone bridging the Mn_1 and Mn_2 metals with the hydroxyl group bound to only Mn_1.[2, 6] Consistent with every structure reported in this study, compound **9** bridges the metal ions through the anionic phenolic group, while the ketone group interacts with Mn_2 and Lys134. This result indicates that even with poorer-binding isosteres like **9**, the introduction of a substituent in the 2-position of the hydroxypyridine ring drives the MBI to coordinate in such a way that bridges the metals with the phenolic group, which is a stronger donor than the ketone and should give better overall affinity. Finally, the saturated heterocycle isosteres **10** and **11** showed markedly poorer inhibition than the parent MBI, possibly due to the softer nature of the heterocyclic donor nitrogen atoms, resulting in poorer electron donation to Mn_2.

Table 3.1. Half maximal inhibitory concentration (IC_{50}), 95% confidence interval (95% CI), $logIC_{50}$ (pIC_{50}), and ligand efficiency (LE) of compounds **1-11** against PA_N endonuclease.

Compound	Structure	IC_{50} (μM)	95% CI	pIC_{50}	LE
1		0.05	±0.02	7.30	0.93
2		0.07	±0.04	7.15	0.83
3		0.09	±0.05	7.05	0.82
4		0.09	±0.11	7.05	0.70
5		0.32	±0.02	6.49	0.65
6		0.50	±0.16	6.30	0.68
7		0.04	±0.02	7.40	0.69
8		0.12	±0.06	6.91	0.65
9		0.40	±0.24	6.40	0.81
10		>0.5	-	<6.30	<0.63
11		>0.5	-	<6.30	<0.63

3.6 Conclusions

In summary, this chapter details the development of several MBIs of the highly active hydroxypyridinone scaffold for the inhibition of PA_N endonuclease. These MBIs were designed to incorporate isosteric replacement of the nonideal carboxylic acid group and retain metal-binding capability for use in metalloenzyme FBDD. A variety of fragments with isosteric functional groups were synthesized via a common intermediate and evaluated in a FRET-based enzymatic assay. Many of these molecules retain comparable inhibitory activity to the parent carboxylic acid fragment, confirming successful isosteric replacement. Promisingly, the MBIs possess a range of physicochemical properties, which may help to overcome pharmacologic liabilities of the parent carboxylic acid functionalized fragment. These molecules were further evaluated crystallographically to improve understanding of the structural and metal-binding properties affected by replacement of the carboxylic acid group. The findings from this chapter demonstrate that when evaluated against PA_N, isosteric groups can replace a problematic carboxylic acid, demonstrate tunable PK properties, and maintain potent inhibition. The use of these compounds and others like them in metalloenzyme FBDD will help identify potent binding molecules that are less susceptible to some of the problems associated with highly polar MBPs.

3.7 Acknowledgements

Chapter 3, in large part, is a reprint of the material as it appears in "Carboxylic Acid Isostere Derivatives of Hydroxypyridinones as Core Scaffolds for Influenza Endonuclease Inhibition," which has been accepted for publication in *ACS Med. Chem. Lett.* The dissertation author was a primary author of this paper and gratefully acknowledges the contributions of coauthors Alysia Kohlbrand, Hyeonglim Seo, Banumathi Sankaran, Johannes Karges, and Seth M. Cohen. Reprinted with permission from *ACS Med. Chem. Lett.* Copyright 2022 American Chemical Society.

3.8 Appendix: Supporting Information

General Experimental Details

All solvents and reagents, unless otherwise noted, were obtained from commercial sources and used without further purification. All reactions, unless otherwise stated, were performed under a nitrogen atmosphere. Silica chromatography was performed using a CombiFlash Rf Teledyne ISCO system using hexane (Hex), ethyl acetate (EtOAc), methylene chloride (DCM), or methanol (MeOH) as eluents. C_{18} reverse phase chromatography was performed using the same instrument using 0.1% formic acid in methanol, acetonitrile, or water as eluent. Several final compounds were further purified using a Gilson preparatory HPLC equipped with a UV detector, using Waters SunFire preparative C_{18} OBD columns (5 μM, 19×50 mm or 19×100 mm). A linear gradient from 10 to 90% MeCN in H_2O was used for 15 min with a flow rate of 20 mL/min. ^{1}H and ^{13}C NMR spectra were obtained on a Varian (400 MHz) spectrometer, Jeol (500 MHz) spectrometer or a VX (500 MHz) equipped with XSens cold probe (Varian) spectrometer in the Department of Chemistry and Biochemistry at U.C. San Diego. The purity of all compounds used in assays was determined

to be ≥95% by HPLC analysis. Mass spectrometry was performed at the U.C. San Diego Molecular Mass Spectrometry Facility. HRMS analysis was performed using an Agilent 6230 Accurate-Mass LC-TOFMS located at the U.C. San Diego Molecular Mass Spectrometry Facility. pK_a values were determined using a Sirius T3 instrument. All titrations were performed in 0.15 M KCl with 0.5 M HCl and KOH. The pK_a values were determined by analyzing the MBI sample in triplicate using potentiometric titration. Experiments were performed over a pH range of 2.0–12.0.

Compound Synthesis

3-Hydroxy-4-oxo-1,4-dihydropyridine-2-carboxylic   acid   (1).   4-(Benzyloxy)-3-hydroxypicolinonitrile (200 mg, 0.88 mmol) was dissolved in 2.1 mL of a 5:5:1 mixture of concentrated hydrochloric acid, acetic acid, and trifluoro acetic acid. The reaction mixture was heated to 110 °C for 24 h. The solvent was removed in vacuo and the remaining residue was purified by C_{18} chromatography eluting in an H_2O/MeOH system to afford the product (12 mg, 9%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 7.81 (d, J = 6.3 Hz, 1H), 7.02 (d, J = 6.4 Hz, 1H). ^{13}C NMR (126 MHz, DMSO-d_6): δ 164.6, 163.0, 152.0, 132.3, 126.5, 112.4. HR-ESI-MS experimental: 156.0292. Calculated for $[C_6H_6NO_4]^+$: 156.0291.

6-Bromo-3-hydroxy-4-oxo-1,4-dihydropyridine-2-carboxylic acid (2). 4-(Benzyloxy)-6-bromo-3-hydroxypicolinic acid (**2a**, 50 mg, 0.15 mmol) was dissolved in 5 mL of concentrated hydrochloric acid and stirred at room temperature for 30 h. The solvent was removed in vacuo and the remaining residue was taken up in water and extracted with ethyl acetate. The organic solvent was removed in vacuo and purified over C_{18} using H_2O/MeOH to afford the product (14 mg, 41 %). ^{1}H NMR

(500 MHz, DMSO-d_6): δ 7.05 (s, 1H). ^{13}C NMR (126 MHz, DMSO-d_6): δ 170.7, 156.3, 149.5, 131.3, 129.9, 117.6. HR-ESI-MS experimental: 231.9252. Calculated for [C$_6$H$_3$BrNO$_4$]$^-$: 231.9251.

6-Bromo-3-hydroxy-4-oxo-1,4-dihydropyridine-2-carboxamide (**3**). KOH (184 mg, 3.28 mmol) and 4-(benzyloxy)-6-bromo-3-hydroxypicolinonitrile (**14**, 200 mg, 0.66 mmol) were combined in EtOH:H$_2$O (10 mL) and heated to 100 °C for 16 h. The reaction was allowed to cool to room temperature, and the solvent was removed in vacuo. The resulting residue was taken up in water and acidified with 4M HCl, resulting in the formation of a white-yellow precipitate (**3a**). This solid was then dissolved in HCl and stirred for 48 h at room temperature. The reaction was dried in vacuo and purified over C$_{18}$ using H$_2$O/MeOH to afford the product (36 mg, 24 %). ^{1}H NMR (500 MHz, (CD$_3$)$_2$CO): δ 12.83 (s, 1H), 9.94 (s, 1H), 8.18 (s, 1H), 7.52 (s, 1H), 7.52, (s, 1H). ^{13}C (126 MHz, (CD$_3$)$_2$CO): δ 171.2, 155.5, 146.7, 130.6, 129.1, 117.4. HR-ESI-MS Experimental: 230.9411. Calculated for [C$_6$H$_4$BrN$_2$O$_3$]$^+$: 230.9411.

6-Bromo-3-hydroxy-2-(1H-tetrazol-5-yl)89yridine-4(1H)-one (**4**). 4-(Benzyloxy)-6-bromo-2-(1H-tetrazol-5-yl)pyridine-3-ol (**4a**) was taken directly from the previous step and dissolved in 5 mL of concentrated hydrochloric acid. The reaction mixture was heated to 100 °C for 8 h. Upon completion, the solvent was removed in vacuo and the resulting residue was purified over C$_{18}$ using H$_2$O/MeOH to afford the product (33.4 mg, 40 %). ^{1}H NMR (500 MHz, CD$_3$OD): δ 7.07 (s, 1H). ^{13}C NMR (126 MHz, CD$_3$OD): δ 155.6, 149.9, 143.6, 131.5, 128.0, 115.8. HR-ESI-MS Experimental: 255.9475. Calculated for [C$_6$H$_3$BrN$_5$O$_2$]$^-$: 255.9476.

6-Bromo-3-hydroxy-N-methoxy-4-oxo-1,4-dihydropyridine-2-carboxamide (**5**). 6-Bromo-3-hydroxy-4-oxo-1,4-dihydropyridine-2-carboxylic acid (**2**, 42 mg, 0.18 mmol) was dissolved in 2 mL of thionyl chloride and heated to reflux for 12 h. The solvent was removed in vacuo and the resulting acid chloride was dissolved in 2 mL of DCM and cooled to -78 °C. Methoxyamine hydrochloride (29 mg, 0.35 mmol) and triethylamine (73 µL, 0.52 mmol) were added, and the reaction was warmed to r.t. and then heated to 40 °C for 48 h. The reaction mixture was filtered, and the filtrate was concentrated under reduced pressure. The remaining residue was purified over C_{18} using H_2O/MeOH to afford the product (15.7 mg, 34 %). ^{1}H NMR (500 MHz, CD_3OD): δ 7.04 (s, 1H), 3.82 (s, 3H). ^{13}C NMR (126 MHz, CD_3OD): δ 155.8, 149.9, 148.0, 129.9, 129.3, 117.1, 63.2. HR-ESI-MS Experimental: 260.9518. Calculated for $[C_7H_6BrN_2O_4]^-$: 260.9516.

6-Bromo-N',3-dihydroxy-4-oxo-1,4-dihydropyridine-2-carboximidamide (**6**). N'-acetoxy-3,4-bis(benzyloxy)-6-bromopicolinimidamide (**6a**, 100 mg, 0.21 mmol) was dissolved in 4 mL of concentrated hydrochloric acid. The solution was heated to 50 °C and stirred for 16 h. The reaction was cooled to room temperature, and the solvent was removed in vacuo. The resulting residue was taken up in water and extracted with ethyl acetate, which was removed in vacuo to afford the product (41 mg, 78%). ^{1}H NMR (500 MHz, CD_3OD): δ 6.91 (s, 1H). ^{13}C NMR (126 MHz, CD_3OD): δ 154.9, 153.1, 143.9, 131.3, 129.1, 114.5. HR-ESI-MS Experimental: 247.9666. Calculated for $[C_6H_7BrN_3O_3]^+$: 247.9665.

3-(6-Bromo-3-hydroxy-4-oxo-1,4-dihydropyridin-2-yl)-1,2,4-oxadiazol-5(4H)-one (**7**). 3-(3,4-Bis(benzyloxy)-6-bromopyridin-2-yl)-1,2,4-oxadiazol-5(2H)-one (**7a**, 51 mg, 0.11 mmol) was dissolved in 4 mL of concentrated hydrochloric acid. The solution was heated to 50 °C and stirred

for 16 h. The reaction was cooled to room temperature, and the solvent was removed in vacuo. The resulting residue was taken up in water and extracted with ethyl acetate. The organic layer was concentrated in vacuo and purified over C_{18}, using H_2O/MeOH as a gradient to afford the product (17 mg, 55%). ^{1}H NMR (500 MHz, CD$_3$OD): δ 7.03 (s, 1H). ^{13}C NMR (126 MHz, CD$_3$OD): δ 159.8, 157.0 155.5, 144.1, 131.0, 127.4, 115.8. HR-ESI-MS Experimental: 271.9312. Calculated for [C$_7$H$_3$BrN$_3$O$_4$]$^-$: 271.9312.

6-Bromo-3-hydroxy-2-(5-methyl-1,2,4-oxadiazol-3-yl)9lyridine-4(1H)-one (**8**). N'-acetoxy-3,4-bis(benzyloxy)-6-bromopicolinimidamide (**8a**, 90 mg, 0.19 mmol) was dissolved in 11 mL of a 5:5:1 mixture of concentrated hydrochloric acid, acetic acid, and trifluoroacetic acid, and stirred for 48 h at 50 °C. The solvent was removed in vacuo and the product was purified over C_{18} using a H_2O/MeOH gradient (20 mg, 39%). ^{1}H NMR (500 MHz, CD$_3$OD): δ 7.06 (s, 1H), 2.74 (s, 3H). ^{13}C (126 MHz, CD$_3$OD): δ 177.5, 165.8, 155.9, 144.3, 131.4, 129.6, 115.9, 10.8). HR-ESI-MS Experimental: 271.9664. Calculated for [C$_8$H$_7$BrN$_3$O$_3$]$^+$: 271.9665

6-Bromo-3-hydroxy-4-oxo-1,4-dihydropyridine-2-carbonitrile (**9**). 4-(Benzyloxy)-6-bromo-3-hydroxypicolinonitrile (**14**, 100 mg, 0.33 mmol) was dissolved in 4 mL of concentrated hydrochloric acid and heated to 50 °C overnight. The solvent was removed in vacuo, and the resulting solids were dissolved in water and extracted with DCM. The organic layer was dried in vacuo to afford the product (21 mg, 30 %). ^{1}H NMR (500 MHz, CD$_3$OD): δ 7.07 (s, 1H). ^{13}C NMR (126 MHz, CD$_3$OD): δ 155.5, 149.2, 130.9, 118.9, 116.5, 114.2. HR-ESI-MS Experimental: 212.9306. Calculated for [C$_6$H$_2$BrN$_2$O$_2$]$^-$: 212.9305.

6-Bromo-2-(4,5-dihydro-1H-imidazol-2-yl)-3-hydroxypyridin-4(1H)-one (**10**). 4-(Benzyloxy)-6-bromo-2-(4,5-dihydro-1H-imidazol-2-yl)92yridine-3-ol (**10a**, 77 mg, 0.22 mmol) was dissolved in 3 mL of concentrated hydrochloric acid and stirred overnight at room temperature. Upon completion, the solvent was removed, and the resulting residue was purified over C_{18} using H_2O/MeOH to afford the product (12 mg, 22 %). ^{1}H NMR (500 MHz, CD_3OD): δ 7.14 (s, 1H), 4.06 (s, 4H). ^{13}C NMR (126 MHz, CD_3OD): δ 162.0, 155.8, 147.3, 131.1, 125.4, 117.0, 44.2. HR-ESI-MS Experimental: 257.9871. Calculated for $[C_8H_9BrN_3O_2]^+$: 257.9873.

6-Bromo-2-(4,5-dihydrooxazol-2-yl)-3-hydroxypyridin-4(1H)-one (**11**). 4-(Benzyloxy)-6-bromo-2-(4,5-dihydrooxazol-2-yl)pyridin-3-ol (**11a**, 90 mg, 0.26 mmol) was dissolved in 4 mL of concentrated hydrochloric acid. The solution was heated to 50 °C and stirred for 16 h. The reaction was cooled to room temperature, and the solvent was removed in vacuo. The resulting residue was taken up in water and extracted with ethyl acetate. The organic layer was concentrated in vacuo and the resulting residue purified over C_{18}, using H_2O/MeOH as a gradient to afford the product (22 mg, 33 %). ^{1}H NMR (500 MHz, CD_3OD): δ 6.78 (s, 1H), 3.81 (t, 2H, J = 5.1 Hz), 3.52 (t, 2H, J = 5.1). ^{13}C NMR (126 MHz, CD_3OD): δ 161.2, 159.8, 159.7, 123.9, 122.9, 112.9, 59.4, 43.8. HR-ESI-MS experimental: 258.9718. Calculated for $[C_8H_8BrN_2O_3]^+$: 258.9713.

4-(Benzyloxy)-6-bromo-3-hydroxypicolinic   acid   (2a).   4-(Benzyloxy)-6-bromo-3-hydroxypicolinonitrile (**14**, 1.0 g, 3.3 mmol) was heated to 100 °C in Ethanol (20 mL) and 6.0 M NaOH (10 mL) for 20 h, during which, a white precipitate formed. The reaction was diluted with 25 mL of water and 25 mL EtOAc were added. The solution was mixed and filtered. To the white precipitate was added 50 mL water, and the solution was acidified to pH ~2. The mixture was then

filtered, affording 4-(benzyloxy)-6-bromo-3-hydroxypicolinic acid as a white solid in 83% yield.
^{1}H NMR (500 MHz, DMSO-d_6): δ 7.50 (s, 1H), 7.47 – 7.31 (m, 5H), 5.26 (s, 2H). ^{13}C NMR (126 MHz, DMSO-d_6): δ 170.7, 156.0, 149.8, 135.8, 130.7, 130.3, 129.0, 128.9, 128.7, 115.9, 71.2. HR-ESI-MS Experimental: 321.9717. Calculated for $[C_{13}H_9BrNO_4]^-$: 321.9720.

4-(Benzyloxy)-6-bromo-2-(1H-tetrazol-5-yl)pyridin-3-ol (**4a**). 4-(Benzyloxy)-6-bromo-3-hydroxypicolinonitrile (**14**, 100 mg, 328 µmol), sodium azide (32 mg, 492 µmol), and ammonium chloride (53 mg, 0.99 mmol) were combined in DMF (5 mL) and heated to 115 °C for 4 h. After the reaction mixture had cooled, it was poured into cool water (note: maintain the concentration of azide in solution <5%) and was acidified to pH 2 with 4M HCl. A white precipitate formed, and the solution was stirred in an ice bath for 30 min, after which time the solids were collected by filtration and dried to afford the product. The resulting solids were dissolved in EtOAc and washed several times with a 5% LiCl solution. Residual DMF was difficult to remove and did not affect the next step, which was undertaken immediately. ^{1}H NMR (500 MHz, DMSO-d_6): δ 7.54 – 7.50 (m, 3H), 7.45 – 7.41 (m, 2H), 7.39 – 7.35 (m, 2H), 5.35 (s, 2H). ^{13}C NMR (126 MHz, DMSO-d_6): δ 156.1, 144.2, 135.8, 131.4, 129.7, 129.0, 128.9, 128.7, 128.6, 114.4, 71.3. HR-ESI-MS Experimental: 345.9943. Calculated for $[C_{13}H_9BrN_5O_2]^-$: 345.9945.

4-(Benzyloxy)-6-bromo-N',3-dihydroxypicolinimidamide (**6a**). 4-(Benzyloxy)-6-bromo-3-hydroxypicolinonitrile (**14**, 1.0 g, 3.3 mmol), hydroxylamine hydrochloride (0.5 g, 6.6 mmol), and triethylamine (0.9 mL, 6.6 mmol) were combined in ethanol (10 mL) in a synth vial. The reaction mixture was heated to 50 °C for 1 h and cooled to room temperature, resulting in the formation of an off-white precipitate. The reaction mixture was diluted with water and extracted with EtOAc

two times. The addition of brine aided with emulsions. The organic layers were combined and concentrated under vacuum to afford the product (0.9 g, 90%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 12.35 (s, 1H), 10.38 (s, 1H), 7.48 – 7.36 (m, 5H), 7.35 (s, 1H), 6.30 (s, 2H), 5.22 (s, 2H). ^{13}C NMR (126 MHz, DMSO-d_6): δ 155.5, 153.1, 145.1, 136.1, 131.7, 129.3, 129.0, 128.8, 128.7, 113.1, 70.9. HR-ESI-MS Experimental: 335.9989. Calculated for $[C_{13}H_{11}BrN_3O_3]^-$: 335.9989.

3-(3,4-Bis(benzyloxy)-6-bromopyridin-2-yl)-1,2,4-oxadiazol-5(4H)-one   (**7a**).   3,4-Bis(benzyloxy)-6-bromo-N'-hydroxypicolinimidamide (**16**, 200 mg, 0.47 mmol), CDI (151 mg, 0.93 mmol), and triethylamine (78.1 µL, 0.56 mmol) were dissolved in 5 mL of DMF. The reaction mixture was heated to 80 °C for 8 h. The solvent was removed in vacuo and ethyl acetate was added. The solution was rinsed with water and the organic solvent removed in vacuo to afford the product (142 mg, 67 %). ^{1}H NMR (500 MHz, CDCl$_3$): δ 7.45 (s, 5H), 7.38 – 7.36 (m, 2H), 7.33 – 7.30 (m, 3H), 7.21 (s, 1H), 5.21 (s, 2H), 5.16 (s, 2H).   HR-ESI-MS Experimental: 452.0258. Calculated for $[C_{21}H_{15}BrN_3O_4]^-$: 452.0251

4-(Benzyloxy)-6-bromo-2-(4,5-dihydro-1H-imidazol-2-yl)pyridin-3-ol (**10a**).   4-(Benzyloxy)-6-bromo-3-hydroxypicolinonitrile (**14**, 100 mg, 0.33 mmol) was dissolved in toluene (3 mL), followed by the addition of ethane-1,2-diamine (29.5 mg, 0.30 µL, 0.49 mmol) and zinc chloride (0.6 mg). The reaction was heated to 130 °C for 20 h. Upon completion, the reaction was allowed to cool to room temperature and the solvent was removed in vacuo. Ethyl acetate was added, and solution was washed with water. The organic solvent was removed in vacuo and the resulting residue purified over silica using a 0-10% (DCM/MeOH) gradient to afford the product (77 mg, 68%). ^{1}H NMR (500 MHz, CDCl$_3$): δ 7.43 (d, J = 7.3 Hz, 2H), 7.36 (t, J = 7.3 Hz, 2H), 7.34 –

7.31 (m, 1H), 7.27 (s, 1H), 5.09 (s, 2H), 3.92 – 3.82 (m, 4H). HR-ESI-MS experimental: 348.0342.

Calculated for $[C_{15}H_{15}BrN_3O_2]^+$: 348.0342.

4-(Benzyloxy)-6-bromo-2-(4,5-dihydrooxazol-2-yl)pyridin-3-ol (**11a**). 4-(Benzyloxy)-6-bromo-3-hydroxypicolinonitrile (**14**, 100 mg, 0.33 mmol) was dissolved in toluene (2 mL), followed by the addition of 2-aminoethan-1-ol (29.7 µL, 0.49 mmol) and zinc chloride (0.6 mg). The reaction mixture was heated to 90 °C for 16 h. Upon completion, the reaction was allowed to return to room temperature and the solvent was removed in vacuo. The remaining residue was taken up in ethyl acetate and washed with water. The organic layer was concentrated to afford the product (96 mg, 84%). ^{1}H NMR (500 MHz, CD$_3$OD): δ 7.51 – 7.47 (m, 2H), 7.40 – 7.34 (m, 2H), 7.34 – 7.29 (m, 1H), 6.92 (s, 1H), 5.13 (s, 2H), 3.79 (t, J = 5.1 Hz, 2H), 3.50 (t, J = 5.1 Hz, 2H). HR-ESI-MS experimental: 349.0183. Calculated for $[C_{15}H_{14}BrN_2O_3]^+$: 349.0182.

4,6-Dibromo-3-hydroxypicolinonitrile (**13**).   Compound **13** was prepared as previously described.[30] To a mechanically stirred solution of 3-hydroxypicolinoitrile (**12**, 10.0 g, 83.3 mmol) and sodium acetate (13.7 g, 167 mmol) in water (150 mL) and methanol (50 mL) at 0 °C was slowly added Br$_2$ (13 mL, 250 mmol) from a pressure-equalizing dropping funnel while maintaining the temperature at <20 °C. The reaction mixture was then stirred overnight at room temperature. After the reaction was complete, as indicated by TLC analysis, the reaction mixture was cooled to 5-10 °C, and slowly charged with an aqueous solution of 25% NaHSO$_3$ (100 mL) while keeping the temperature at <20 °C. The resulting suspension was stirred for 30 min and filtered. The filter cake was washed with water, dried in air for several hours and then under vacuum to afford the product as a light-yellow solid (21.5 g, 93 %). ^{1}H NMR (500 MHz, DMSO-

d_6): δ 8.27 (s, 1H). ^{13}C NMR (126 MHz, DMSO-d_6): δ 156.3, 136.2, 130.0, 126.6, 122.0, 115.2. HR-ESI-MS experimental: 274.8462. Calculated for [$C_6HBr_2N_2O$]$^-$: 274.8461.

4-(Benzyloxy)-6-bromo-3-hydroxypicolinonitrile (**14**). Compound **14** was prepared as previously described.[2] In a flame-dried flask, NaH (3.6 g, 90 mmol) was stirred in dry DMSO (20 mL, dried on molecular sieves). Benzyl alcohol (9.4 mL, 90 mmol) was added dropwise to the mixture, pausing when excessive foam formed. When the reaction was complete, as evidenced by the cessation of hydrogen evolution, 4,6-dibromo-3-hydroxypicolinonitrile (**13**, 5.0 g, 18 mmol) in dry DMSO (10 mL) was added dropwise to the stirring solution, and the mixture was heated to 60 °C for 18 h, monitoring conversion by thin-layer chromatography. When the reaction was complete, the mixture was placed in an ice bath, and 150 mL of water were added. The mixture was stirred for 10 min, during which time the solution clouded and then clarified. To remove residual benzyl alcohol, the aqueous solution was extracted twice with 100 mL portions of diethyl ether. Residual ether was then removed from the aqueous portion under vacuum, and the solution was cooled to 5 °C on ice. The addition of 4M HCl (until pH ~2) resulted in a solid precipitate. It is imperative to maintain an ice bath temperature during acidification; otherwise, excess heat or residual ether present in the solution will result in the formation of a gummy, orange oil in lieu of off-white crystallites. If an oil forms, it can be slowly dissolved in basic (pH > 10) water with sonication to repeat the precipitation. The acidified mixture was stirred for 30 min on ice, and the resultant solids were filtered, rinsed with H_2O, and dried. Purification via silica gel chromatography in an EtOAc/Hex system afforded 4-(benzyloxy)-6-bromo-3-hydroxypicolinonitrile (4.3 g, 78%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 11.62 (s, 1H), 7.57 (s, 1H), 7.53 – 7.48 (m, 2H), 7.44 – 7.34 (m,

3H), 5.33 (m, 2H). HR-ESI-MS experimental: 302.9777. Calculated for $[C_{13}H_8BrN_2O_2]^-$: 302.9775.

3,4-Bis(benzyloxy)-6-bromopicolinonitrile (**15**). 4-(Benzyloxy)-6-bromo-3-hydroxypicolinonitrile (**14**, 1.0 g, 3.3 mmol) and potassium carbonate (0.9 g, 6.6 mmol) were taken up in DMF (30 mL), and (bromomethyl)benzene (0.8 mL, 6.6 mmol) was added slowly. The reaction was stirred for 19 h at 40 °C. After cooling, 30 mL of ethyl acetate was added to the reaction mixture, which was then washed with 30 mL of water. The organic layer was collected, and the aqueous layer was extracted once more with 30 mL of ethyl acetate. The organic layers were combined and concentrated in vacuo. The product was separated over silica using a 0-50% Hex/EtOAc gradient. (1.2 g, 91%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 7.83 (s, 1H), 7.55 – 7.50 (m, 2H), 7.48 – 7.41 (m, 3H), 7.35 – 7.32 (m, 5H), 5.38 (s, 2H), 5.22 (s, 2H). ^{13}C NMR (126 MHz, DMSO-d_6): δ 159.8, 148.8, 136.9, 135.8, 135.3, 129.3, 129.2, 128.9, 128.8, 126.8, 117.6, 114.8, 75.9, 72.1. HR-ESI-MS experimental: 395.0387. Calculated for $[C_{20}H_{16}BrN_2O_2]^+$: 395.0390.

3,4-Bis(benzyloxy)-6-bromo-N'-hydroxypicolinimidamide (**16**). 3,4-Bis(benzyloxy)-6-bromopicolinonitrile (**15**, 2.1 g, 5.0 mmol), hydroxylamine hydrochloride (0.9 g, 13.2 mmol), and triethylamine (1.8 mL, 18.8 mmol) were combined in Ethanol (200 mL) in a round-bottom flask. The reaction mixture was heated to 50 °C for 15 h. The solvent was removed in vacuo and the residue was taken up in ethyl acetate and rinsed with water. The organic layer was collected and removed in vacuo to afford the product (2.1 g, 97 %). ^{1}H NMR (500 MHz, (CD$_3$)$_2$CO: δ 9.34 (s, 1H), 7.08 – 7.01 (m, 3H), 7.01 – 6.89 (m, 5H), 6.87 – 6.81 (m, 3H), 5.31 (s,

2H), 4.85 (s, 2H), 4.51 (s, 2H).   HR-ESI-MS experimental: 428.0599.   Calculated for [C$_{20}$H$_{19}$BrN$_3$O$_3$]$^+$: 428.0604.

N'-acetoxy-3,4-bis(benzyloxy)-6-bromopicolinimidamide (**17**).   To a 5 °C solution of 4-(benzyloxy)-N',3-dihydroxypicolinimidamide (**16**, 100 mg, 0.23 mmol) and DBU (77.4 µL, 0.51 mmol) in dry THF stirring on ice was added acetic anhydride (44.1 µL, 0.47 mmol) dropwise. The solution was stirred 30 min on ice and was then heated to 80 °C for 8 h. The solvent was then evaporated in vacuo. The resulting mixture was dissolved in ethyl acetate and rinsed with water. The organic layer was concentrated in vacuo and the resulting residue was separated over silica using Hex/EtOAc as a gradient to afford the product (110 mg, 80%). ^{1}H NMR (500 MHz, CDCl$_3$): δ 7.46 – 7.43 (m, 2H), 7.41 (s, 5H), 7.29 – 7.25 (m, 3H), 7.13 (s, 1H), 5.64 – 5.54 (bs, 2H), 5.15 (s, 2H), 5.14 (s, 2H), 2.22 (s, 3H).   HR-ESI-MS experimental: 470.0707.   Calculated for [C$_{22}$H21BrN$_3$O$_4$]$^-$: 470.0710.

PA$_N$ Endonuclease Assay

PA$_N$ endonuclease activity assays were carried out as previously reported.[10] Assays were performed using Black Costar 96-well plates. Each well contained a total volume of 100 µL comprised of buffer (20 mM Tris, 150 mM NaCl, 2 mM MnCl$_2$, 0.2% TritonX100, pH 8.0, 0.2% 2-mercaptoethanol), influenza PA$_N$ endonuclease (25 nM), inhibitor (various concentrations), and fluorescent ssDNA-oligo substrate (250 nM). A single-stranded, 17-mer DNA substrate labeled with a 5′-FAM fluorophore and a 3′-TAM quencher ([6-FAM]AATCGCAGGCAGCACTC[TAM], Sigma-Aldrich) was employed as the substrate. All assay components were pipetted into the plate, and ultimately, the substrate was added using a

multi-channel pipette, and the assay was immediately started. Samples were prepared in triplicates. Background wells consisting of all assay components except enzyme were prepared for each sample. Positive and negative controls were prepared on each plate to gauge the fluorescence signal of fully active protein and the absence of protein. Change in fluorescence of each well was measured by a Synergy H4 Hybrid Multi-Mode Microplate Reader (BioTek) at 39 second intervals over 45 min at 37 °C (λ_{ex} = 485 nm; λ_{em} = 528 nm). The gain was set to 100. Typically, data collected between 20 and 35 min was used in the activity calculations, as this data range had a linear slope. The slope of the fluorescence signal for each sample was background corrected and percent inhibition was determined by normalizing the slope of the sample to that of the positive and negative controls. Dose response curves were generated for inhibitors by plotting percent inhibition versus log of the concentration for each inhibitor. The data were fit with a sigmoidal curve to determine the IC_{50} value using a four parameter MATLAB script. Ligand efficiency (LE) values were calculated using the following equation: LE = ΔG/heavy atom count (HAC). This equation can be calculated for IC_{50} values using the equation: $-RT\ln(IC_{50})$/HAC, which can be approximated as $1.4(pIC_{50})$/HAC.[31]

Computational Prediction of pK_a Values

   The predictions were performed slightly modified from a published procedure.[19] Density functional theory (DFT) calculations were performed with the Gaussian software package (Gaussian, Inc., Wallingford CT, 2016). The calculations were done with the B3LYP hybrid functional and the 6-31++ g(d,p) basis set. Solvent effects were included using CPCM model. The concentration of bulk water was considered to be 55.49 M. The structures of all calculated molecules correspond to ground state minima on the ground state potential energy surfaces with no imaginary frequencies present. The following thermodynamic cycle and equations were used to determine the pK_a value:

$$\text{HA(g)} \quad + \quad \text{H}_2\text{O(g)} \quad \xrightarrow{\Delta G_g} \quad \text{A}^-\text{(g)} \quad + \quad \text{H}_3\text{O}^+\text{(g)}$$

$$\uparrow \text{-}\Delta G_s(\text{HA}) \qquad \uparrow \text{-}\Delta G_s(\text{H}_2\text{O}) \qquad \downarrow \Delta G_s(\text{A}^-) \qquad \downarrow \Delta G_s(\text{H}_3\text{O}^+)$$

$$\text{HA(s)} \quad + \quad \text{H}_2\text{O(s)} \quad \xrightarrow{\Delta G_s} \quad \text{A}^-\text{(s)} \quad + \quad \text{H}_3\text{O}^+\text{(s)}$$
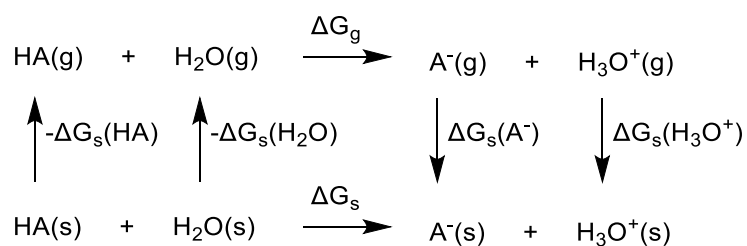
s=aqueous phase, g=gas phase

$$\Delta G_s = \Delta G_g + \Delta G_s(\text{A}^-) + \Delta G_s(\text{H}_3\text{O}^+) - \Delta G_s(\text{HA}) - \Delta G_s(\text{H}_2\text{O})$$

$$pK_a = (\Delta G_s / 1.364) - \log(\text{H}_2\text{O})$$

Computational Prediction of Binding Poses

The geometry of the fragment was determined using density-functional theory (DFT) calculations with the Gaussian software package (Gaussian, Inc., Wallingford CT, 2016). The calculations were performed using the Pople double-zeta basis set with a single set of polarization functions on non-hydrogen atoms (6-31G(d)). Solvent effects were included using a polarizable continuum model (PCM). The structures of all calculated molecules correspond to minima on the ground state potential energy surfaces with no imaginary frequencies present. The metalloprotein structure (PDB Code: 6E6V)[2] was downloaded from the Protein Data Bank (https://www.rcsb.org/) and prepared for the docking experiment using the Molecular Operating Environment (MOE, Chemical Computing Group ULC, Montreal, QC, Canada, 2019) software package. Water molecules, other small molecules, inhibitors as well as ions were removed. Hydrogen atoms were added, and the side chains protonated at physiological pH. The structurally optimized molecules were docked using the Genetic Optimization for Ligand Docking (GOLD, Cambridge Crystallographic Data Centre, Cambridge, United Kingdom, 2019) software package. The protein structure was considered rigid. The binding pose of the compounds was predicted using a genetic algorithm with a population size of 200, selection pressure of 1.2, number of operations of 500000, number of islands of 5, niche size of 2, crossover frequency of 95, mutation frequency of 95, and a migration frequency of 10. The genetic algorithm was set to 100 runs. The metal ions within the active site were predefined to have an octahedral geometry. During the docking procedure the binding poses were evaluated using the ChemPLP scoring function. After docking, the obtained solutions were re-scored using the GoldScore fitness function.

Table 3.S1. Comparative predictions of cLogP.[*]

Compound	R = H	R = Br	R = Me	R = Ph
2R	-0.35	0.47	-0.14	0.95
3R	-1.15	-0.33	-0.95	0.15
4R	-1.04	-0.22	-0.84	0.26
5R	-0.78	0.04	-0.58	0.52
6R	-1.09	-0.27	-0.88	0.21
7R	-0.31	0.5	-0.11	0.98
8R	-0.10	0.72	0.05	0.88
9R	-0.12	0.70	0.08	1.18
10R	-0.75	0.70	-0.55	0.55
11R	0.07	0.88	0.27	1.36

[*]Values calculated in MarvinSketch 19.6.

Table 3.S2. Predicted and measured phenolic pK_a values.

Compound	Calculated pK_a	Measured pK_a
1	8.5	-
2	9.4	9.79 ± 0.01
3	9.2	9.64 ± 0.02
4	12.6	>12
5	8.4	-
6	8.9	9.06 ± 0.02
7	9.8	-
8	10.7	-
9	8.1	-
10	9.7	-
11	9.5	-

Protein Expression and Purification

Expression and purification of PA_N endonuclease was performed as reported previously. The pandemic H1N1 N-terminal PA (PA_N) endonuclease Δ52-64:Gly truncated construct was expressed from a pET-28a parent vector containing a kanamycin-resistance reporter gene with expression inducible by the Lac1 operon. PA_N endonuclease was expressed as an 8-histidine tagged fusion protein cleavable by TEV protease. The transformation protocol was adapted from pET system manual (Novagen) using single competent BL21 cells. Briefly, 1 µL of 25 ng/µL recombinant plasmid was used for transformation. Cells were mixed with the plasmid and were heat shocked at 42 °C for 30 sec followed by incubation on ice for 3 min. Outgrowth was plated on LB agarose plates containing 50 µg/mL kanamycin and was incubated overnight at 37 °C. One colony was scraped from the LB plate and added to 5 mL of SOC broth containing 50 µg/mL kanamycin and was incubated overnight at 37 °C with shaking at 125 rpm. SOC media (100 mL) containing 50 µg/mL kanamycin was combined with the 5 mL overnight growth and was incubated with shaking at 200 rpm at 37 °C until the OD_{600} of this starter culture reached >2 (3-4 hours). The culture was equally divided into 6, 2L flasks containing 1L of expression media (TB media with added 0.2% dextrose, 0.1 mM $MnCl_2$, and 0.1 mM $MgSO_4$, 50 µg/mL kanamycin). Cells were grown to and OD_{600} between 0.4-0.6 at room temperature with shaking at 200 rpm (3-4 hours). Expression was then induced by addition of IPTG to a final concentration of 0.1 mM. The cultures were grown with vigorous shaking (250 rpm) overnight at room temperature. After ~18 h the cells were harvested by centrifuging at 2000g for 30 min at 4 °C. The resulting paste was stored at -80 °C prior to lysis.

Cell paste was thawed in batches on ice for 2 h and resuspended in 25-35 mL of lysis buffer (1% Triton-X, 1mM $MgCl_2 \cdot 6H_2O$, 2mM DTT, 10-100 µg/mL DNAse-1, 1mg/ml lysozyme, and

1% glycerol) and EDTA free protease inhibitor (Roche). Cells were lysed using a probe sonicator (Fisherbrand model 120) with cycles of 25 second pulses and 59 second rest at 60% amplitude. Cell debris was then pelleted by centrifugation at 10000 rpm for 45 min at 4 °C. The supernatant was decanted from the pellet, and a HisTrap FF (Cytiva) column was utilized to isolate His-tagged fusion protein from the cell lysates according to the manufacturer's recommendations at 4 °C. Briefly, cell-free lysates from 6 L growth were loaded on 5 mL column that had previously been charged with Ni ions. The column was then washed with binding buffer (20 mM Na_2PO_4, 500 mM NaCl, 25 mM imidazole, pH 7.4) until fraction absorbance reached a steady baseline. The protein was then eluted over a gradient from 0-100% elution buffer (20 mM Na_2PO_4, 500 mM NaCl, 500 mM imidazole, pH 7.4) at a flow rate of 4 mL/min. PA_N endonuclease eluted between 40-60% elution buffer. SDS-PAGE analysis showed a band corresponding to PA_N endonuclease running at ~23 kDa with several small impurities.

Fractions containing PA_N endonuclease were combined in a 10K MWCO dialysis bag with 1000 units of TEV protease and were dialyzed against dialysis buffer (100 mM NaCl, 1 mM dithiothreitol, 1 mM $MnCl_2$, 20 mM Tris, 5% glycerol, pH 8.0) overnight with three buffer exchanges. The proteolytic cleavage of the fusion protein is slow and greatly benefits from the addition of excess TEV protease. A white precipitate forms over time. After buffer exchange, the solution was filtered through a 0.45 μm filter. The solution was run through the Histrap FF column equilibrated with the same binding buffer as before. The resulting flow through contained His-cleaved PA_N endonuclease, which was then concentrated to 5-10 mg/mL using a pressurized Amicon and/or spin Amicon concentrator. The concentrated protein was then purified on a gel-permeation size exclusion column (GE Superdex 75, 10/300 GL) according to manufacturer recommendations in buffer (150 mM NaCl, 2 mM $MgCl_2$, 2 mM $MnCl_2$, 20 mM HEPES, pH 7.5).

A large peak corresponding to the cleaved PA$_N$ endonuclease eluted at ~12 mL eluent. A small shoulder before the main peak was occasionally observed, which contained primarily uncleaved and/or unfolded PA$_N$ endonuclease construct. Fractions containing pure cleaved PA$_N$ endonuclease were combined and concentrated to 2-5 mg/mL. Stored protein was flash-frozen in liquid nitrogen and was kept at -80 °C. This protein was suitable for use in enzyme or thermal shift assays or for protein crystallography.

Protein Crystallography

Purified protein for crystallization was stored at 2.2-4.3 mg/mL at -80 °C after flash freezing in buffer consisting of 150 mM sodium chloride, 20 mM HEPES (pH 7.5), 2 mM MgCl$_2$, and 2 mM MnCl$_2$. Co-crystallization and crystal soaking methods were used to obtain co-crystal structures of inhibitors bound to PA$_N$ endonuclease. For co-crystallization, protein was incubated with 0.5 mM inhibitor for 1 hour on ice prior to setting the crystallization drops. For crystal soaking, fully formed holo crystals were transferred to a new drop containing 5 µL of reservoir solution and 1 µL of 50 mM DMSO inhibitor stock solution (final concentration 8.3 mM). Crystals were left undisturbed overnight and either stored in liquid nitrogen, or collected on an in-house X-ray diffractometer the following day. In both crystallization methods, crystals were grown using hanging drop and set in 24-well pre-greased plates (Hampton HR3-171) with siliconized glass slides (Hampton HR3-231). A 5:1 ratio of purified protein to reservoir solution at room temperature was found to be the optimal ratio for the largest crystal formation. Reservoir solution consisted of 22-34% PEG (MW 4000 g/mol), 100 mM Tris (pH 8.35), and 220 mM sodium acetate. Colorless crystals with hexagonal bipyramidal morphology appeared within 2 days and reached full size after 1-2 weeks. Crystals were typically 50 to 200 microns in diameter. Crystals were

cryoprotected with perfluoroether (Hampton HR2-814) prior to flash freezing in liquid nitrogen. Crystals were stored in liquid nitrogen until data collection.

Datasets for protein co-crystallized with compounds **5, 9,** and **10** were collected on an in-house X-ray diffractometer. For these experiments, diffraction data was collected at 100 K on a Bruker X8 Proteum diffractometer using a Bruker Microfocus Rotating Anode (MicroStar FR-592) X-ray generator with a Bruker APEX II CCD detector at wavelength 1.54178 Å. Data was integrated, scaled, and merged using the Bruker APEX3 software package (Bruker, 2017). Datasets for protein co-crystallized with compounds **1, 2, 3, 4,** and **8** were collected at the Advanced Light Source (ALS), Lawrence Berkeley National Laboratory, in collaboration with Dr. Banu Sankaran through Collaborative Crystallography program on beamline 8.2.1. For all datasets (including datasets collected on the in-house X-ray generator), phasing was determined by molecular replacement against a previously published PA_N endonuclease structure (PDB 4AWM) using PHASER. All structures were refined with Phenix version 1.19.2.

Table 3.S3. X-ray crystallographic data collection and refinement statistics.

Compound	1	2	3	4
PDB	8CTF	8DDB	7V04	8DDE
Data collection statistics				
Resolution range	57.45 - 2.14 (2.217 - 2.14)	58.02 - 2.15 (2.227 - 2.15)	65.32 - 1.91 (1.978 - 1.91)	57.81 - 2.22 (2.299 - 2.22)
Space group	$P6_222$	$P6_222$	$P6_222$	$P6_222$
Cell dimensions a, b, c, (Å) α, β, γ (°)	75.66 75.66 119.451 90 90 120	76.228 76.228 121.608 90 90 120	75.421 75.421 119.317 90 90 120	75.979 75.979 121.005 90 90 120
Unique reflections	11738 (1134)	11965 (1158)	16216 (1580)	10783 (1053)
Completeness (%)	99.65 (99.74)	99.91 (99.83)	99.86 (100.00)	99.81 (99.15)
Mean I/sigma(I)	23.37 (2.50)	25.56 (2.58)	29.60 (2.59)	24.87 (2.54)
R-merge	0.07424 (1.636)	0.0699 (2.003)	0.05714 (1.42)	0.08323 (2.05)
R-measured	0.07587 (1.669)	0.07145 (2.043)	0.05841 (1.455)	0.08492 (2.092)
R-work	0.2030 (0.3203)	0.1824 (0.2868)	0.1967 (0.3112)	0.1876 (0.3225)
R-free	0.2419 (0.3374)	0.2385 (0.3637)	0.2347 (0.3702)	0.2442 (0.4187)
RMS(bonds)	0.003	0.011	0.008	0.01
RMS(angles)	0.44	1.1	0.98	1.02
Ramachandran favored (%)	98.29	94.86	98.86	94.89
Ramachandran outliers (%)	0	0.57	0	0.57
Average B-factor	68.07	68.88	51.24	64.51
Redundancy	25	25	24.4	26.2
CC1/2	1.000	0.999	0.999	1.000

* Metrics for highest resolution shell given in parentheses

Table 3.S4. X-ray crystallographic data collection and refinement statistics.

Compound	**5**	**8**	**9**	**11**
PDB	8DJV	8DHN	8DAL	8DJY
Data collection statistics				
Resolution range	37.24-2.08 (2.15-2.08)	64.98-2.40 (2.48-2.40)	40.15-2.2 (2.28-2.2)	44.29-2.5 (2.5-2.59)
Space group	P6$_2$22	P6$_2$22	P6$_2$22	P6$_2$22
Cell dimensions a, b, c, (Å) α, β, γ (°)	74.465 74.465 120.259 90 90 120	75.034 75.034 121.24 90 90 120	75.403 75.403 120.436 90 90 120	75.807 75.807 119.978 90 90 120
Unique reflections	12322 (1077)	15004 (1221)	11981 (1049)	13372 (732)
Completeness (%)	99.1 (90.8)	99.74 (100.0)	100.0 (100.0)	100.0 (100.0)
Mean I/sigma(I)	13 (2.3)	20 (2.0)	17.2 (1.5)	20 (2.0)
R-merge	0.1 (0.775)	0.1 (1.439)	0.09 (1.368)	0.091 (1.243)
R-measured	0.103 (0.815)	0.102 (1.478)	0.092 (1.432)	0.093 (1.292)
R-work	0.2626 (0.3450)	0.2006(0.2792)	0.1998 (0.3086)	0.2116 (0.3358)
R-free	0.291 (0.354)	0.2563 (0.3205)	0.2576 (0.3866)	0.2860 (0.3964)
RMS(bonds)	0.01	0.004	0.004	0.008
RMS(angles)	1.21	0.62	0.57	0.86
Ramachandran favored (%)	94.32	96.57	98.86	96.02
Ramachandran outliers (%)	0	0	0	0
Average B-factor	48.6	54.1	54.5	57.6
Redundancy	17.7	29	19.6	19.1
CC1/2	0.999	1.000	1.000	0.999

* Metrics for highest resolution shell given in parentheses

Table 3.S5. Metal-binding distances between coordinating atoms and metals. IC_{50} values are shown for context.

Compound	Isostere-Mn_2	OH-Mn_2	OH-Mn_1	Carbonyl-Mn_1	Mn-Mn	IC_{50}
1	2.2 Å	1.9 Å	2.3 Å	1.7 Å	3.9 Å	0.05 μM
2	2.1 Å	2.1 Å	2.2 Å	2.2 Å	3.7 Å	0.07 μM
3	2.1 Å	2 Å	2.2 Å	2.2 Å	3.8 Å	0.09 μM
4	1.8 Å	2.3 Å	1.9 Å	2.3 Å	3.7 Å	0.09 μM
5	2.0 Å	2.2 Å	2.1 Å	2 Å	3.8 Å	0.32 μM
8	2.1 Å	2.3 Å	2.1 Å	2.1 Å	3.9 Å	0.12 μM
9	2.8 Å	2.2 Å	2.2 Å	2.3 Å	3.9 Å	0.40 μM
10	2.0 Å	2.3 Å	2 Å	2.3 Å	3.8 Å	>0.5 μM

Table 3.S6. Octahedral Distortion Parameters.

Compound	D_{mean}	ζ	Σ
Ligand free (water)	2.1914	0.685765	44.5718
3	2.1662	0.357853	46.6692
4	2.1378	1.151002	64.0144
8	2.2385	0.574277	90.8137
8 (artificially forced planar)	2.2330	0.606847	98.3470
9	2.3006	1.042725	92.6387

Synthesis and Crystallization of Tris(pyrazolyl)borate Model Complexes

[(Tp^{Ph,Me})ZnOH] was synthesized as reported using [Tp^{Ph,Me}K], which was prepared as previously reported.[32] [(Tp^{Ph,Me})ZnOH] (10 mg, 0.02 mmol) and 6-bromo-3-hydroxy-2-(5-methyl-1,2,4-oxadiazol-3-yl)pyridin-4(1H)-one (**8**, 5 mg, 0.02 mmol) were dissolved in 1 mL DCM and 0.5 mL of MeOH. The reaction mixture was stirred overnight at room temperature. The resulting mixture was evaporated to dryness via rotary evaporation and subsequently dissolved in a minimal amount of benzene. The solution was filtered to remove any undissolved solids. The resulting complex in benzene was recrystallized using vapor diffusion with pentane to afford crystals. The TpZn structure with compound **6** was prepared similarly, though additional [(Tp^{Ph,Me})ZnOH] was used. [(Tp^{Ph,Me})ZnOH] (60 mg, 0.11 mmol) was dissolved in 2 mL of DCM. 6-bromo-N',3-dihydroxy-4-oxo-1,4-dihydropyridine-2-carboximidamide (**6**, 15 mg, 0.06 mmol) was dissolved in 2 mL of DCM was added, and the reaction mixture was stirred overnight under a nitrogen atmosphere. The resulting mixture was evaporated to dryness via rotary evaporation and subsequently dissolved in a minimal amount of benzene. The solution was filtered to remove any

undissolved solids. The resulting complex in benzene was recrystallized using vapor diffusion with pentane.

Single Crystal X-ray Diffraction

Suitable crystals of $[(Tp^{Ph,Me}Zn)_2(\mathbf{6})]$ and $[(Tp^{Ph,Me})Zn(\mathbf{8})]$ ($Tp^{Ph,Me}$ = hydrotris(3,5-phenylmethylpyrazolyl)borate) were selected and data was collected at 100 K on a Bruker APEX-II Ultra diffractometer with a Mo-Kα Microfocus Rotating Anode and a APEX-II CCD area detector or a Bruker Kappa diffractometer equipped with a Bruker X8 APEX II Mo sealed tube and a Bruker APEX-II CCD. The data were integrated using the Bruker SAINT Software program and scaled using the SADABS software program. The structure was solved with the ShelXT[33] structure solution program using direct methods and refined with the XL refinement package using least squares minimization using Olex2.[34] The crystal data files were deposited into the Cambridge Crystallographic Data Centre (CCDC).

111
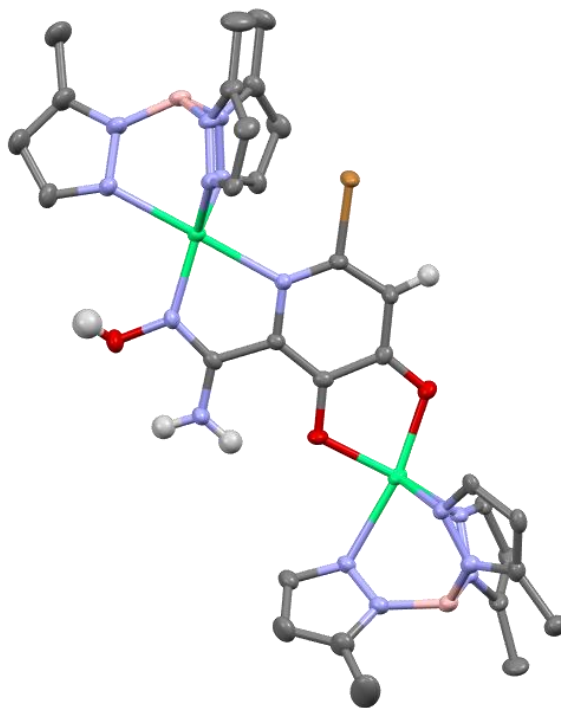
Figure 3.S1. Crystal structure of $[(Tp^{Ph,Me})_2Zn_2(\mathbf{6})]$ (ORTEP, 50% probability ellipsoids) highlighting the ability of the compound to form multiple metal chelates. Phenyl rings are omitted for clarity.
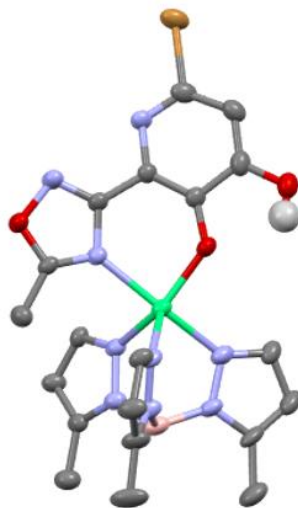
Figure 3.S2. Crystal structure of $[(Tp^{Ph,Me})Zn(\mathbf{8})]$ (ORTEP, 50% probability ellipsoids). Phenyl rings are omitted for clarity.

Table 3.S7. Crystal data and structure refinement for $[(Tp^{Ph,Me})_2Zn_2(\mathbf{6})]$ and $[(Tp^{Ph,Me})Zn(\mathbf{8})]$ complexes.

Compound	$[(Tp^{Ph,Me})_2Zn_2(\mathbf{6})]$	$[(Tp^{Ph,Me})Zn(\mathbf{8})]$
Identification code	2163562	2163563
Empirical formula	$C_{33}H_{30}BBr_{0.5}N_{7.5}O_{1.5}Zn$	$C_{41}H_{36}BBrN_9O_3Zn$
Formula weight	671.78	858.88
Temperature/K	100	100
Crystal system	triclinic	monoclinic
Space group	P-1	P21/c
a/Å	13.0275(5)	24.986(2)
b/Å	13.1334(6)	13.5052(12)
c/Å	19.2346(8)	23.071(2)
α/°	79.7020(10)	90
β/°	78.8350(10)	94.770(3)
γ/°	89.1920(10)	90
Volume/Å^3	3175.9(2)	7758.1(12)
Z	4	8
ρ_{calc}g/cm^3	1.405	1.471
μ/mm^{-1}	1.443	1.715
F(000)	1380	3512
Crystal size/mm^3	$0.1 \times 0.1 \times 0.1$	$0.1 \times 0.1 \times 0.1$
Radiation	MoKα (λ = 0.71073)	MoKα (λ = 0.71073)
2Θ range for data collection/°	2.194 to 52.244	3.776 to 48.342
Index ranges	-16 ≤ h ≤ 16, -16 ≤ k ≤ 15, -23 ≤ l ≤ 23	-28 ≤ h ≤ 27, -15 ≤ k ≤ 15, -26 ≤ l ≤ 26
Reflections collected	33400	167523
Independent reflections	12591 [R_{int} = 0.0394, R_{sigma} = 0.0472]	12421 [Rint = 0.2372, Rsigma = 0.0814]
Data/restraints/parameters	12591/0/812	12421/0/1022
Goodness-of-fit on F^2	1.055	1.001
Final R indexes [I>=2σ (I)]	R_1 = 0.0517, wR$_2$ = 0.1673	R1 = 0.0584, wR2 = 0.1338
Final R indexes [all data]	R_1 = 0.0673, wR$_2$ = 0.1787	R1 = 0.1141, wR2 = 0.1689
Largest diff. peak/hole / e Å^{-3}	3.56/-0.39	2.02/-1.72

Octahedral Distortion
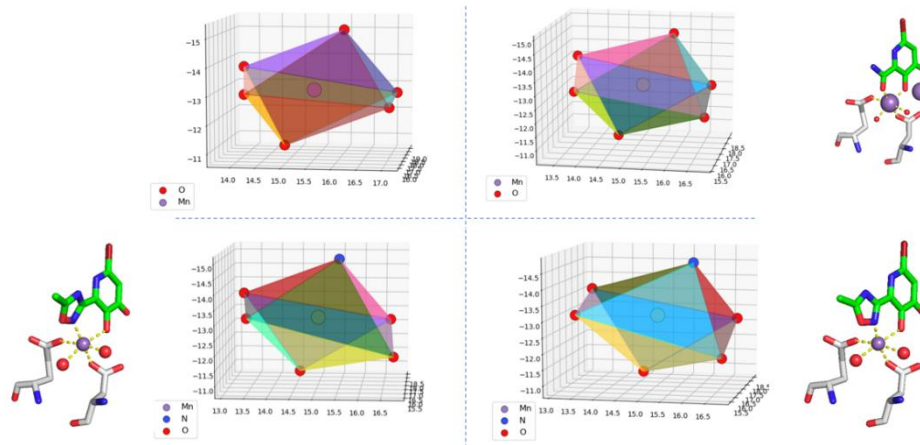
Figure 3.S3. Visualization of octahedral distortion around Mn_2 in PA_N endonuclease. The ligand coordination sphere is shown with the central purple sphere representing Mn_2, red spheres representing oxygen donor atoms, and blue spheres representing nitrogen donor atoms. *Top left*: Octahedron of Mn_2 with no inhibitor (bound water molecules), *Top right*: Octahedron of experimentally measured structure of compound **3**. *Bottom left*: Octahedron of experimentally measured structure of compound **8**. *Bottom right*: An 'artificial structure' of compound **8** forced into a planar conformation. Octahedral distortion parameters are shown in Table 3.S6.
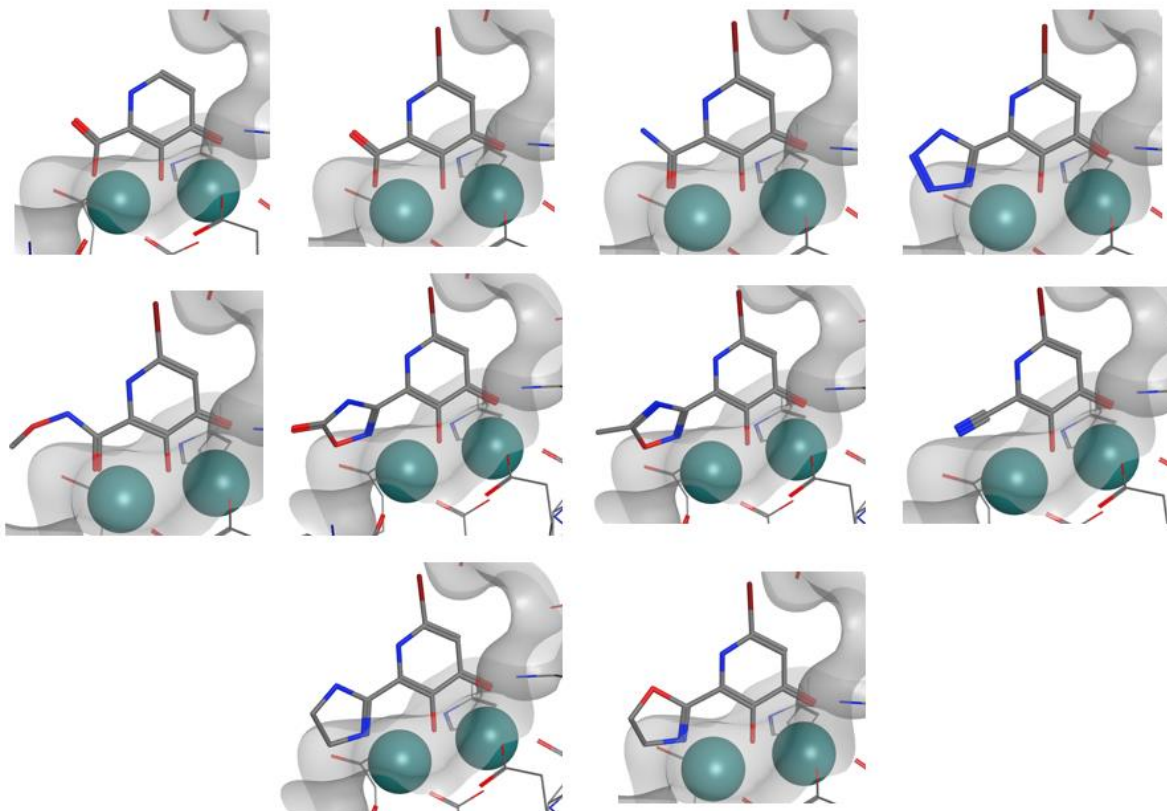
Figure 3.S4. Docking poses of MBI's predicted by GOLD.
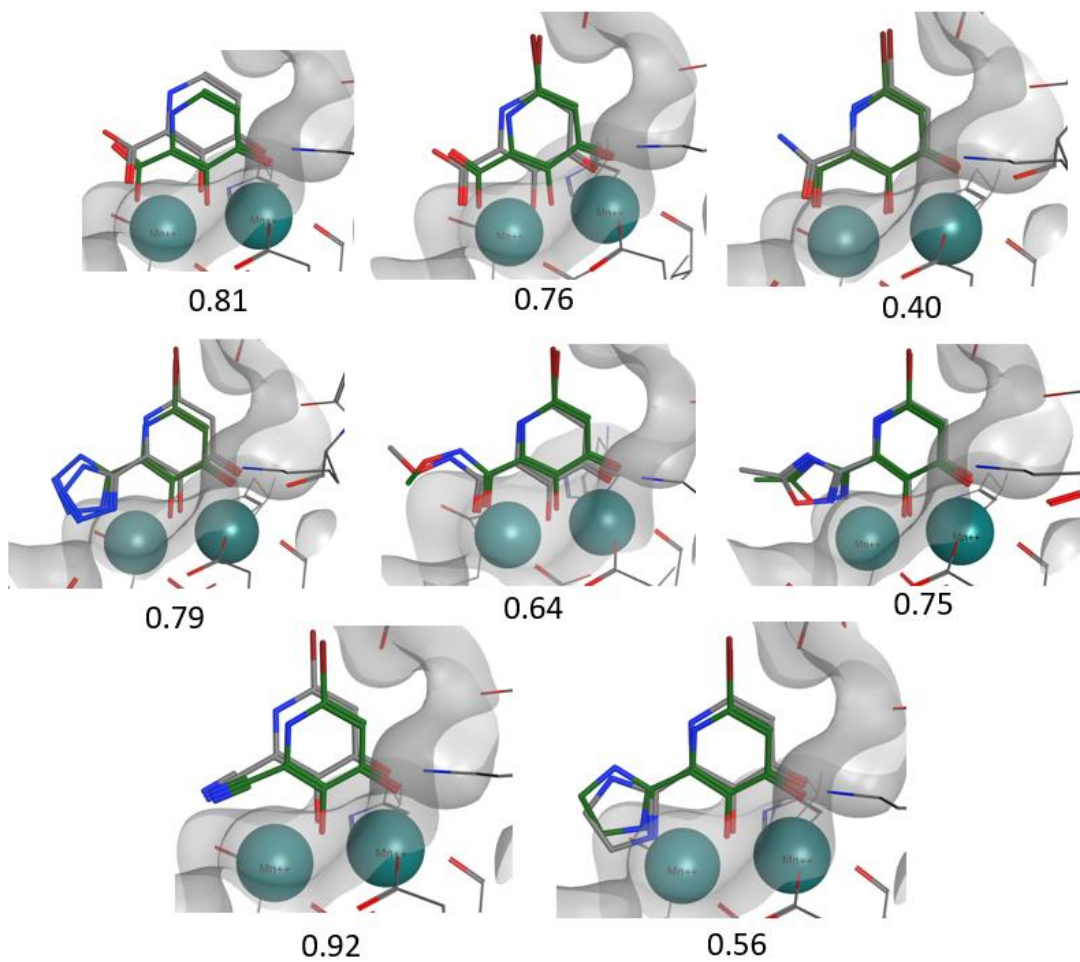
Figure 3.S5. RMSD values between computationally predicted structures (green) and experimentally obtained co-crystal structures (grey). Calculated using LigRMSD.[35]

3.9 References

1.   Yuan, P.; Bartlam, M.; Lou, Z.; Chen, S.; Zhou, J.; He, X.; Lv, Z.; Ge, R.; Li, X.; Deng, T.; Fodor, E.; Rao, Z.; Liu, Y., Crystal structure of an avian influenza polymerase PAN reveals an endonuclease active site. *Nature* **2009,** *458* (7240), 909-913.

2.   Credille, C. V.; Morrison, C. N.; Stokes, R. W.; Dick, B. L.; Feng, Y.; Sun, J.; Chen, Y.; Cohen, S. M., SAR Exploration Of Tight-Binding Inhibitors Of Influenza Virus PA Endonuclease. *J. Med. Chem.* **2019,** *62* (21), 9438-9449.

3.   Hutchinson, E. C.; Fodor, E., Transport of the influenza virus genome from nucleus to nucleus. *Viruses* **2013,** *5* (10), 2424-2446.

4.   Monod, A.; Swale, C.; Tarus, B.; Tissot, A.; Delmas, B.; Ruigrok, R. W.; Crépin, T.; Slama-Schwok, A., Learning from structure-based drug design and new antivirals targeting the ribonucleoprotein complex for the treatment of influenza. *Expert Opin. Drug Discovery* **2015,** *10* (4), 345-371.

5.   Iuliano, A. D.; Roguski, K. M.; Chang, H. H.; Muscatello, D. J.; Palekar, R.; Tempia, S.; Cohen, C.; Gran, J. M.; Schanzer, D.; Cowling, B. J.; Wu, P.; Kyncl, J.; Ang, L. W.; Park, M.; Redlberger-Fritz, M.; Yu, H.; Espenhain, L.; Krishnan, A.; Emukule, G.; van Asten, L.; Pereira da Silva, S.; Aungkulanon, S.; Buchholz, U.; Widdowson, M.-A.; Bresee, J. S.; Azziz-Baumgartner, E.; Cheng, P.-Y.; Dawood, F.; Foppa, I.; Olsen, S.; Haber, M.; Jeffers, C.; MacIntyre, C. R.; Newall, A. T.; Wood, J. G.; Kundi, M.; Popow-Kraupp, T.; Ahmed, M.; Rahman, M.; Marinho, F.; Sotomayor Proschle, C. V.; Vergara Mallegas, N.; Luzhao, F.; Sa, L.; Barbosa-Ramírez, J.; Sanchez, D. M.; Gomez, L. A.; Vargas, X. B.; Acosta Herrera, a.; Llanés, M. J.; Fischer, T. K.; Krause, T. G.; Mølbak, K.; Nielsen, J.; Trebbien, R.; Bruno, A.; Ojeda, J.; Ramos, H.; an der Heiden, M.; del Carmen Castillo Signor, L.; Serrano, C. E.;

Bhardwaj, R.; Chadha, M.; Narayan, V.; Kosen, S.; Bromberg, M.; Glatman-Freedman, A.; Kaufman, Z.; Arima, Y.; Oishi, K.; Chaves, S.; Nyawanda, B.; Al-Jarallah, R. A.; Kuri-Morales, P. A.; Matus, C. R.; Corona, M. E. J.; Burmaa, A.; Darmaa, O.; Obtel, M.; Cherkaoui, I.; van den Wijngaard, C. C.; van der Hoek, W.; Baker, M.; Bandaranayake, D.; Bissielo, A.; Huang, S.; Lopez, L.; Newbern, C.; Flem, E.; Grøneng, G. M.; Hauge, S.; de Cosío, F. G.; de Moltó, Y.; Castillo, L. M.; Cabello, M. A.; von Horoch, M.; Medina Osis, J.; Machado, A.; Nunes, B.; Rodrigues, A. P.; Rodrigues, E.; Calomfirescu, C.; Lupulescu, E.; Popescu, R.; Popovici, O.; Bogdanovic, D.; Kostic, M.; Lazarevic, K.; Milosevic, Z.; Tiodorovic, B.; Chen, M.; Cutter, J.; Lee, V.; Lin, R.; Ma, S.; Cohen, A. L.; Treurnicht, F.; Kim, W. J.; Delgado-Sanz, C.; de mateo Ontañón, S.; Larrauri, A.; León, I. L.; Vallejo, F.; Born, R.; Junker, C.; Koch, D.; Chuang, J.-H.; Huang, W.-T.; Kuo, H.-W.; Tsai, Y.-C.; Bundhamcharoen, K.; Chittaganpitch, M.; Green, H. K.; Pebody, R.; Goñi, N.; Chiparelli, H.; Brammer, L.; Mustaquim, D., Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *Lancet* **2018,** *391* (10127), 1285-1300.

6. Bauman, J. D.; Patel, D.; Baker, S. F.; Vijayan, R. S. K.; Xiang, A.; Parhi, A. K.; Martínez-Sobrido, L.; LaVoie, E. J.; Das, K.; Arnold, E., Crystallographic Fragment Screening and Structure-Based Optimization Yields a New Class of Influenza Endonuclease Inhibitors. *ACS Chem. Biol.* **2013,** *8* (11), 2501-2508.

7. Jones, J. C.; Marathe, B. M.; Lerner, C.; Kreis, L.; Gasser, R.; Pascua, P. N. Q.; Najera, I.; Govorkova, E. A., A Novel Endonuclease Inhibitor Exhibits Broad-Spectrum Anti-Influenza Virus Activity In Vitro. *Antimicrob. Agents Chemother.* **2016,** *60* (9), 5504-5514.

8. Miyagawa, M.; Akiyama, T.; Taoda, Y.; Takaya, K.; Takahashi-Kageyama, C.; Tomita, K.; Yasuo, K.; Hattori, K.; Shano, S.; Yoshida, R.; Shishido, T.; Yoshinaga, T.; Sato, A.;

Kawai, M., Synthesis and SAR Study of Carbamoyl Pyridone Bicycle Derivatives as Potent Inhibitors of Influenza Cap-dependent Endonuclease. *J. Med. Chem.* **2019,** *62* (17), 8101-8114.

9.   Ng, K. E., Xofluza (Baloxavir Marboxil) for the Treatment Of Acute Uncomplicated Influenza. *P. T.* **2019,** *44* (1), 9-11.

10.   Credille, C. V.; Chen, Y.; Cohen, S. M., Fragment-Based Identification of Influenza Endonuclease Inhibitors. *J. Med. Chem.* **2016,** *59* (13), 6444-6454.

11.   Credille, C. V.; Dick, B. L.; Morrison, C. N.; Stokes, R. W.; Adamek, R. N.; Wu, N. C.; Wilson, I. A.; Cohen, S. M., Structure–Activity Relationships In Metal-Binding Pharmacophores For Influenza Endonuclease. *J. Med. Chem.* **2018,** *61* (22), 10206-10217.

12.   Karges, J.; Stokes, R. W.; Cohen, S. M., Photorelease Of A Metal-Binding Pharmacophore From A Ru(ii) Polypyridine Complex. *Dalton Trans.* **2021,** *50* (8), 2757-2765.

13.   Lassila, T.; Hokkanen, J.; Aatsinki, S.-M.; Mattila, S.; Turpeinen, M.; Tolonen, A., Toxicity of Carboxylic Acid-Containing Drugs: The Role of Acyl Migration and CoA Conjugation Investigated. *Chem. Res. Toxicol.* **2015,** *28* (12), 2292-2303.

14.   Pajouhesh, H.; Lenz, G. R., Medicinal chemical properties of successful central nervous system drugs. *NeuroRx* **2005,** *2* (4), 541-553.

15.   Dick, B. L.; Cohen, S. M., Metal-Binding Isosteres as New Scaffolds for Metalloenzyme Inhibitors. *Inorg. Chem.* **2018,** *57* (15), 9538-9543.

16.   Patani, G. A.; LaVoie, E. J., Bioisosterism: A Rational Approach in Drug Design. *Chem. Rev.* **1996,** *96* (8), 3147-3176.

17.   Lassalas, P.; Gay, B.; Lasfargeas, C.; James, M. J.; Tran, V.; Vijayendran, K. G.; Brunden, K. R.; Kozlowski, M. C.; Thomas, C. J.; Smith, A. B.; Huryn, D. M.; Ballatore, C.,

Structure Property Relationships of Carboxylic Acid Isosteres. *J. Med. Chem.* **2016,** *59* (7), 3183-3203.

18.   Marson, C. M., Chapter Two - Saturated Heterocycles with Applications in Medicinal Chemistry. In *Advances in Heterocyclic Chemistry*, Scriven, E. F. V.; Ramsden, C. A., Eds. Academic Press: 2017; Vol. 121, pp 13-33.

19.   Pliego, J. R., Thermodynamic cycles and the calculation of pKa. *Chem. Phys. Lett.* **2003,** *367* (1), 145-149.

20.   Karges, J.; Stokes, R. W.; Cohen, S. M., Computational Prediction of the Binding Pose of Metal-Binding Pharmacophores. *ACS Med. Chem. Lett.* **2022,** *13* (3), 428-435.

21.   Zhang, C.; Xiang, J.; Xie, Q.; Zhao, J.; Zhang, H.; Huang, E.; Shaw, P.; Liu, X.; Hu, C., Identification of Influenza PAN Endonuclease Inhibitors via 3D-QSAR Modeling and Docking-Based Virtual Screening. *Molecules* **2021,** *26* (23).

22.   Ketkaew, R.; Tantirungrotechai, Y.; Harding, P.; Chastanet, G.; Guionneau, P.; Marchivie, M.; Harding, D. J., OctaDist: a tool for calculating distortion parameters in spin crossover and coordination complexes. *Dalton Trans.* **2021,** *50* (3), 1086-1096.

23.   Buron-Le Cointe, M.; Hébert, J.; Baldé, C.; Moisan, N.; Toupet, L.; Guionneau, P.; Létard, J. F.; Freysz, E.; Cailleau, H.; Collet, E., Intermolecular control of thermoswitching and photoswitching phenomena in two spin-crossover polymorphs. *Phys. Rev. B* **2012,** *85* (6), 064114.

24.   Drew, M. G. B.; Harding, C. J.; McKee, V.; Morgan, G. G.; Nelson, J., Geometric control of manganese redox state. *J. Chem. Soc., Chem. Commun.* **1995,** (10), 1035-1038.

25.   Guionneau, P.; Brigouleix, C.; Barrans, Y.; Goeta, A. E.; Létard, J.-F.; Howard, J. A. K.; Gaultier, J.; Chasseau, D., High pressure and very low temperature effects on the crystal structures of some iron(II) complexes. *C. R. Acad. Sci., Ser. IIc: Chim.* **2001,** *4* (2), 161-171.

26.   Baughman, B. M.; Jake Slavish, P.; DuBois, R. M.; Boyd, V. A.; White, S. W.; Webb, T. R., Identification of Influenza Endonuclease Inhibitors Using a Novel Fluorescence Polarization Assay. *ACS Chem. Biol.* **2012,** *7* (3), 526-534.

27.   Kalliokoski, T.; Kramer, C.; Vulpetti, A.; Gedeck, P., Comparability of Mixed IC50 Data – A Statistical Analysis. *PLoS One* **2013,** *8* (4), e61007.

28.   Todd, B.; Tchesnokov, E. P.; Götte, M., The active form of the influenza cap-snatching endonuclease inhibitor baloxavir marboxil is a tight binding inhibitor. *J. Biol. Chem.* **2021,** *296*, 100486.

29.   Malik, M. A.; Wani, M. Y.; Al-Thabaiti, S. A.; Shiekh, R. A., Tetrazoles as carboxylic acid isosteres: chemistry and biology. *J. Inclusion Phenom. Macrocyclic Chem.* **2014,** *78* (1), 15-37.

30.   Renga, J. M. Z., Y.; Whiteker, G. T.; Choy, N.; Stockman, K. E, Process for the preparation of 4-alkoxy-3-hydroxypicolinic acids. *Patent WO2016007634* **2016**

31.   Carr, R. A. E.; Congreve, M.; Murray, C. W.; Rees, D. C., Fragment-based lead discovery: leads by design. *Drug Discovery Today* **2005,** *10* (14), 987-992.

32.   Puerta, D. T.; Cohen, S. M., Elucidating Drug-Metalloprotein Interactions with Tris(pyrazolyl)borate Model Complexes. *Inorg. Chem.* **2002,** *41* (20), 5075-5082.

33.   Sheldrick, G., SHELXT - Integrated space-group and crystal-structure determination. *Acta Cryst.* **2015,** *A71* (1), 3-8.

34.   Dolomanov, O. V.; Bourhis, L. J.; Gildea, R. J.; Howard, J. A. K.; Puschmann, H., OLEX2: a complete structure solution, refinement and analysis program. *J. Appl. Cryst.* **2009,** *42* (2), 339-341.

35.	Velázquez-Libera, J. L.; Durán-Verdugo, F.; Valdés-Jiménez, A.; Núñez-Vivanco, G.; Caballero, J., LigRMSD: A Web Server For Automatic Structure Matching And RMSD Calculations Among Identical And Similar Compounds In Protein-Ligand Docking. *Bioinformatics* **2020,** *36* (9), 2912-2914.

Chapter 4: Evaluation of 3-dimensionality in Approved and Experimental Drug Space

4.1 Introduction

As discussed in Chapter 1.6, three-dimensionality is an important consideration in fragment library design because the diversity of druggable protein targets requires ligands to possess both structural and conformational variety to generate effective pharmaceuticals.[1] Many compounds used in FBDD have the propensity to trend towards planarity, as was highlighted in Chapter 1.6. This chapter seeks to analyze large libraries of compounds that can be used in (fragment-based) drug discovery, quantify their three-dimensionality, and explore new approaches to address the lack of highly 3D fragment libraries.

The 3D shape of a molecule can be described by the number of sp^3 carbons,[2] the plane of best fit calculation,[3] or normalized principal moment of inertia (PMI) values,[4] among other parameters.[5, 6] These descriptors have facilitated the topological analysis of large fragment[7] and high-throughput screening (HTS) libraries;[8] to the best of our knowledge, only limited analyses of approved and late-stage drugs have been undertaken.[9, 10, 11] The normalization that is common practice for PMI analysis allows for the comparison of compounds with diverse structural features and molecular weights (MW) and is a straightforward calculation that can be readily carried out by many commercial software packages. The PMI parameter was demonstrated to be a robust method for evaluating 3D topology when compared to the other methods.[3]

Using PMI analyses, this chapter aims to evaluate the 3D conformations of structures reported in the DrugBank, including approved and experimental therapeutics. Of particular interest is the relationship between the energy minimized/optimized conformations of approved therapeutics and the conformation of those molecules when bound to proteins in the PDB. Previous efforts have focused on the exploration of flexible conformations upon protein binding as they relate to in silico screening and predicting binding conformations rather than the inherent

change in 3D topology as described herein.[1, 12, 13] Lastly, this chapter consider the role that inorganic chemistry could play in increasing the prevalence of more 3D molecules in the drug discovery pipeline.

4.2 3D Diversity in the DrugBank

The DrugBank is an online cheminformatics and bioinformatics database that provides extensive information on both drugs and drug targets.[14] The DrugBank repository contains over 13,000 chemical entities, with over 4000 unique targets. Of the 13,000+ chemical entities, 8696 can be downloaded as 3D structures (version 5.1.4), with 8532 having MW >100 Da. These structures include approved, experimental, investigational, illicit, and withdrawn drugs, as well as nutraceuticals. Molecular Operating Environment (MOE) is a drug discovery software platform that can be used for the curation and evaluation of these structures. MOE was used to determine the moments of inertia for each compound ($I_1 \leq I_2 \leq I_3$), which describe the torque needed for a given angular acceleration. These values were used to determine the normalized PMI values, I_1/I_3 and I_2/I_3, for structures with MW >100 Da. The most common presentation of these data is graphically (Figure 4.1a), where the [0 1] apex is a linear molecule (dimethylbutadiyne), the [0.5 0.5] corner is a planar molecule (benzene), and the [1 1] apex is a spherical structure (adamantane). Where a conformation falls within these parameters provides an indication of its topology, with molecules having larger scores ($I_1/I_3 + I_2/I_3$) being considered more 3D in character.
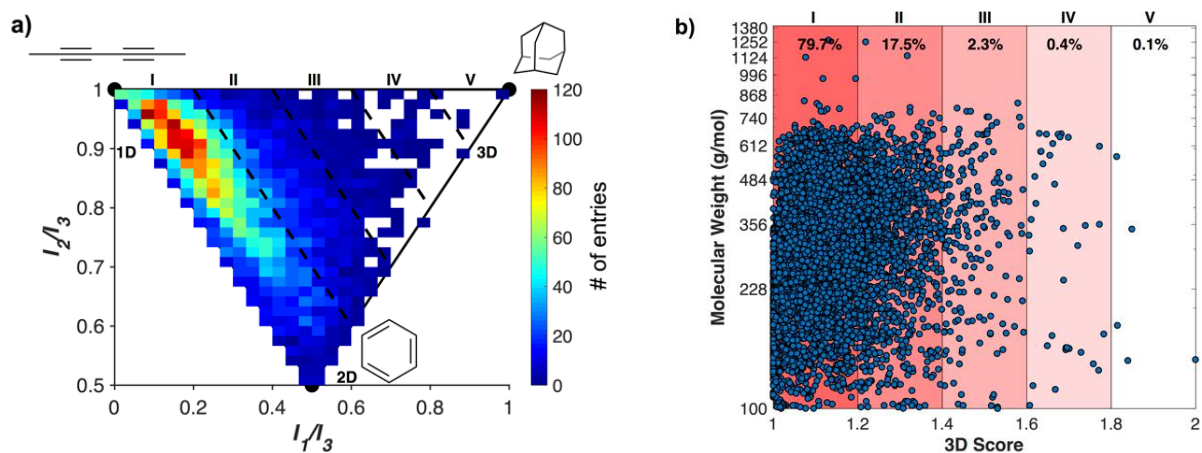
Figure 4.1. PMI analysis of (a) 8532 entries from the DrugBank, plotted on a typical PMI triangle plot and (b) as 3D Score ($I_1/I_3 + I_2/I_3$) vs. MW.

The DrugBank structures, when analyzed by normalized PMIs, are found to have largely linear and planar topologies, with just under 80% having a 3D score <1.2 (Figure 4.1). Structures having 3D scores >1.6 (Figure 4.1b, **IV** and **V**) can be considered 'highly' 3D, and only 0.5% of the DrugBank falls in these regions. This is in agreement with a previous report that found molecules in the ChEMBL Database (a database of bioactive compounds with drug-like properties) tend towards planarity.[15] The topological distribution of the DrugBank structures is also typical of organic libraries used in drug development, such as the ZINC database of virtual fragments, where ~75% of the fragments have a 3D score of <1.2.[7] This is not to imply that contemporary drug discovery does not include concerted design and synthetic efforts at achieving the best 3D molecular shape for a target, only that when libraries of currently and previously used compounds and fragments are retroactively analyzed they are often observed to lack decidedly 3D topologies. The traditional chemical feedstocks used in drug discovery campaigns are dominated by sp^2-rich molecules and retrosynthetic analyses of large drug-like libraries have confirmed this, demonstrating the bias toward using planar building blocks in medicinal chemistry.[15] Though it has been demonstrated that relatively planar scaffolds may be combined in a way that provides

access to a reasonable sampling of chemical and topological space, it has also been shown that constraints in synthetic accessibility limit a more unbiased and diverse sampling of topological space.[7] Additionally, adherence to the rule of five[16] (Ro5) in drug design, which helps select for oral bioavailability, may predispose molecules to flatter geometries, for example, by limiting the number of rotatable bonds.

The 3D scores of the DrugBank structures were evaluated with respect to their MWs, ranging from 100 to 1268 Da. There is no obvious relationship between the MW of a compound and its 3D shape (Figure 1b). Interestingly, in each 3D score range the MWs are broadly distributed between 100 and ~700 Da. The molecules with the greatest 3D scores (**IV**) range from 140 to 568 Da. This demonstrates that what might perhaps be the more intuitive notion, that larger molecules have greater 3-dimensionality, is not the case, as has been discussed recently.[17]

4.3 3D Diversity in the PDB

The PDB is the premier repository for biological macromolecular structures, containing over 150,000 total entries and over 11,000 depositions in 2019 alone.[18] Within the PDB, ligands bound to protein structures having DrugBank IDs have been identified, and their corresponding ligand coordinates can be downloaded. Limiting these structures to compounds with a MW >100 Da yields 7411 structures that comprise ~500 unique chemical entities. The 3D analysis of all the protein-bound structures generates a PMI plot whose general features are similar to those of the DrugBank molecules, with the highest concentration of compounds in the upper left corner (Appendix, Figure 4.S2). However, this data set is biased by molecules that appear with greater frequency. For example, sapropterin, a dietary supplement administered to reduce blood phenylalanine levels appears in this list 1036 times. To better illustrate the conformational

diversity of the PDB, the PMI values for each duplicate molecular entry were averaged over all conformations to yield 502 unique PMI values and 3D scores (Figure 4.2). As was the case for the DrugBank and individual PBD entries, the averaged PMI values from the different PDB conformations localize primarily in the upper left linear region. Despite the large differences in protein targets and MW values of the structures analyzed here, both the energy minimized conformations (DrugBank) and protein bound conformations remain largely linear and/or planar. There have been in-depth discussions regarding the extensive molecular rearrangements drug molecules are capable of undergoing in order to facilitate protein binding.[12, 19, 20] It is interesting to note that these rearrangements do not translate to an overall change in 3D topology, at least for the molecular structures explored here.

The 3D score vs. MW of the averaged PDB ligand conformations was plotted (Figure 4.2b) and found to be similar to the outcome for the minimized DrugBank 3D structures (Figure 4.1b). The 3D scores represent molecules of broad MW distributions and there is no obvious relationship between MW and 3D topology. As demonstrated in the PMI plot (Figure 4.2a), the majority (83.5%) of compounds fall into bin **I** (Figure 4.2b), with decreasing prevalence in each bin of increasing 3-dimensionality. This follows the same general trend as the geometry minimized DrugBank structures and further highlights the absence of highly 3D topologies in currently used and proposed drug molecules. As described earlier, it is possible that these linear/planar features are the result of the feedstocks and libraries used for drug discovery campaigns.

The energy of protein bound ligand structures in the context of conformational change tolerance has been explored elsewhere.[19-22] However, the 3D topologies of these different conformations has not been evaluated through the PMI analysis. Many of the DrugBank structures appear in the PDB with high frequency, though many do not undergo significant changes (<0.1) in 3D score.

Diclofenac (Figure 4.3), an orally administered nonsteroidal anti-inflammatory drug (NSAID), was chosen as a case study to explore both energy and topology. Diclofenac has 51 protein-bound structures in the PDB with 3D scores ranging from 1.03 (linear, **I**) to 1.52 (moderately 3D, **III**).
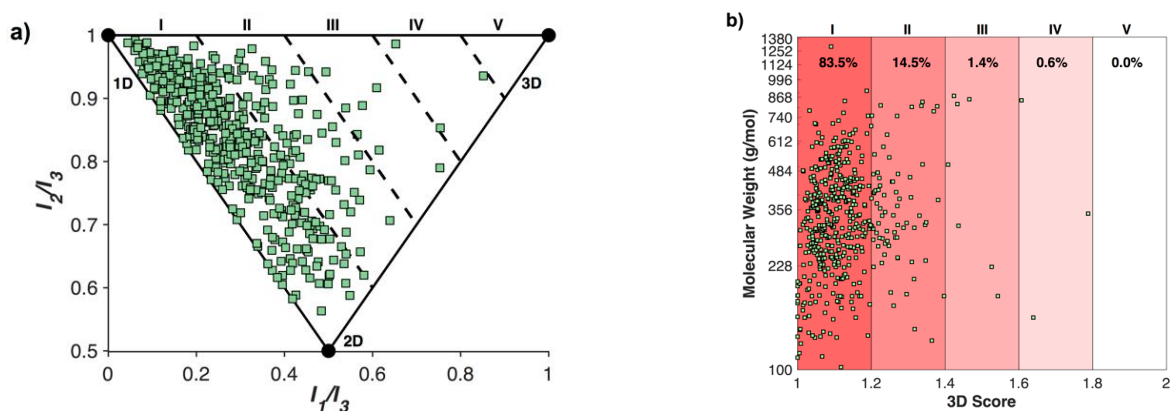
**Figure 4.2.** Averaged PMI analysis of (a) 502 entries from the PDB, plotted on a typical PMI triangle plot and (b) as 3D Score ($I_1/I_3 + I_2/I_3$) vs. MW.

Three representative conformations of diclofenac, the least and most 3D, as well as the energy minimized structures, are shown in Figure 4.3a. These structures have 3D scores of 1.03, 1.52, and 1.22, respectively, and can be visually identified as significant and obvious changes in their 3D conformations. The least 3D structure has the two aromatic rings nearly in plane with each other, while in the most 3D topology the rings are much closer to perpendicular, and the energy minimized conformation falls in between. When the potential energy surfaces of the different conformations are determined, there is no clear relationship to their diverse 3D topologies (Figure 4.3b). The higher energy conformations (white/orange) include some of the most and least 3D topologies. The lower energy conformations (black) span 3D scores between 1.1 and 1.4, representing significant topological translations. These findings are in agreement with previous reports stating that a wide range of strain energies can be tolerated by therapeutics when bound to target sites.[1, 12, 19, 21, 22]
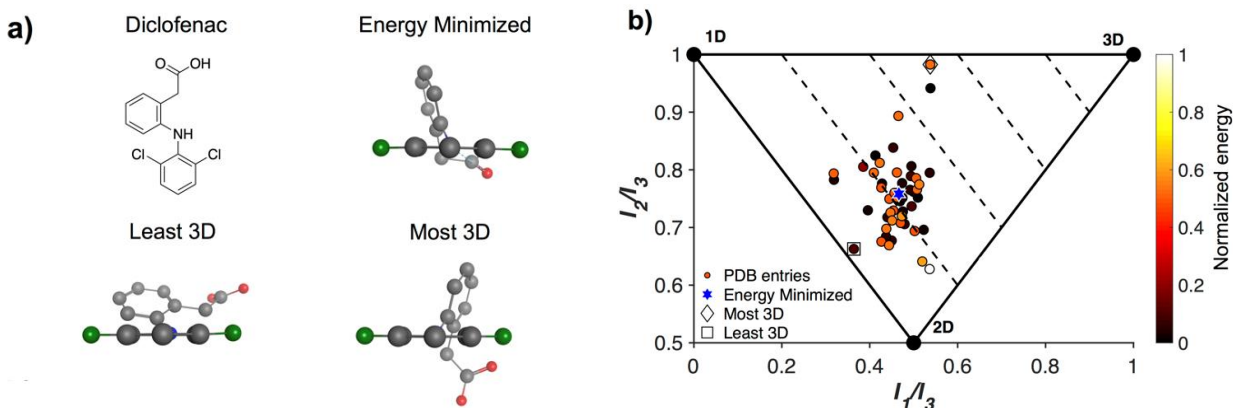
Figure 4.3. Diclofenac is an NSAID with 51 PDB entries. a) The most and least 3D conformations of diclofenac, as well as an energy-minimized conformation. b) An analysis of the relative energies of the various conformations of the diclofenac PDB entries.

To more closely examine the effects of a protein active site on a molecule, a PMI analysis of several FDA-approved human immunodeficiency virus (HIV) drugs from two distinct target classes was conducted. Nevirapine and efavirenz are non-nucleoside reverse transcriptase inhibitors (NNRTIs), while nelfinavir, lopinavir, and saquinavir are protease inhibitors.[23] Relevant ligand conformations were downloaded from the PDB, and a PMI analysis was performed for each entry (Figure 4.4).

The aforementioned inhibitors were bound to HIV reverse transcriptases and proteases that contained a range of sequence mutations. In this analysis, sequence identity between reverse transcriptases was conserved to at least 98% when compared to a parent sequence (Table 4.S1). Similarly, homology between proteases was maintained to a minimum of 78% in order to ensure comparison of inhibitor structures bound to similar proteins. For the purposes of this analysis, it was assumed that tertiary structure remained largely conserved across the mutations. For each protein-bound inhibitor, the PMI remained mostly static, indicating a high conservation of binding motif, as expected for compounds bound to conserved targets. Nelfinavir and nevirapine showed

almost no variation in conformation across 11 and 17 structures, respectively, while saquinavir remained mostly consistent, with two notable exceptions. Both exceptions, at [0.40 0.97] and [0.56 0.61] from PDB entries 3S56 and 3UFN, respectively, are due to a second saquinavir molecule that is bound to an extension of the active site cavity near the protein surface.[24, 25] These second drug molecules exhibit a more 3D conformation than the saquinavir molecule found in the active site, which in the 3S56 PDB structure has a PMI of [0.31 0.86]. While this more 3D conformation is likely due to both crystal packing and mutational differences between the proteases, it is an interesting example of one molecule displaying significantly different topologies even when interacting with the same protein. In particular, this example demonstrates that an active site can dictate the topology of the bound ligand quite substantially, and any straying from that binding pocket can result in notable changes in conformation. It is interesting to observe that with a single exception across 52 structures, each protease inhibitor adopts a conformation with a 3D score <1.2. These topologies are likely due to both the shape of the active site of HIV proteases, as well as their chemical structural similarities. Both NNRTIs (efavirenz and nevirapine) also show closely related 3D scores, though slightly higher than those of the protease inhibitors.
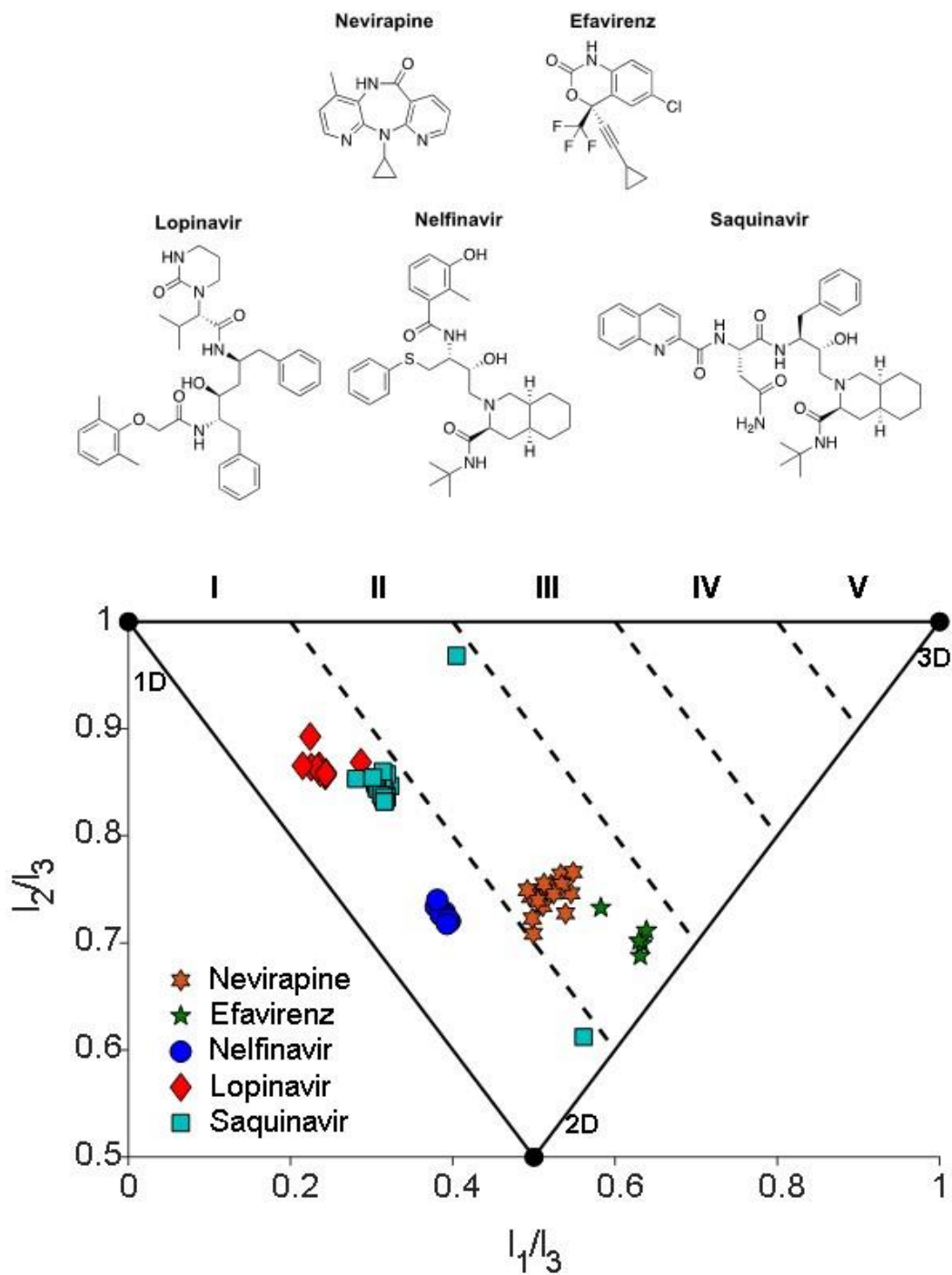
Figure 4.4. PMI analysis of FDA-approved antiviral therapies, nevirapine, efavirenz (reverse transcriptase inhibitors), nelfinavir, lopinavir, and saquinavir (protease inhibitors).

These similar PMI values across several drugs designed for the same target highlight the importance of the protein-binding pocket shape in drug discovery. While the need for 3D scaffolds to explore many current and desired target sites exists, it is also true that in some cases a linear/planar molecule will be best suited. It is for this reason that topological shape diversity is desired in library development and is predicated on the ability to generate molecules of high 3-dimensionality, as low 1- and 2-dimensionality is already broadly covered. Attempts to address the size and shape of protein active sites has resulted in a suite of tools that can be used to aid in drug discovery efforts. [26-29]

4.4 Achieving 3-dimensionality in Medicinal Chemistry

To address the lack of 3-dimensionality in pharmaceuticals there have been efforts in academic and industry labs to design and synthesize libraries of 3D molecules to be used in drug discovery campaigns.[5, 7, 30-33] These structures include compounds prepared through diversity-oriented syntheses with a higher than typical number of sp^3 carbons, as well as chiral and spiro centers. Such strategies are often regarded as possible solutions to the over representation of flat molecules in fragment libraries and drug discovery feedstocks.[2, 4, 7, 17, 31-33] To evaluate the success of such approaches, a PMI analysis (Figure 4.5) was carried out on 15 unique fragments from three different reports (Figure 4.S4).[7, 32, 33] These studies each sought to generate organic fragments with high degrees of 3-dimensionality. Of the compounds evaluated, 6/15 (40%) effectively escape the linear/planar region (3D Score >**I**). These results represent a significant enhancement in 3D character compared to the typical fragments libraries that have 75% of molecules in "flatland."[7] While these results are promising, all organic fragments presented here fail to achieve 3D Scores >**II,** indicating their relatively limited topological diversity. This is not

133

a comprehensive survey of current efforts in this area, but it does suggest that there is room for further improvement.

The medicinal inorganic community has considered the inclusion of metal centers into existing drugs as an intriguing method for augmenting existing therapeutics.[34] Tamoxifen is a chemotherapeutic agent that has been modified to include a ferrocene group (ferrocifen), and the anticancer activity against estrogen-receptor positive cancers increases upon the inclusion of the iron moiety. An unanticipated consequence of this modification is that the 3-dimensionality of the molecule increases slightly from tamoxifen to ferrocifen (Figure 4.5). With an understanding of the inherent 3-dimensionality of metal-complexes, Meggers and co-workers set out to incorporate organometallic groups into kinase inhibitors as deliberately 3D moieties.[34-36] When a representative compound from Meggers' work is analyzed using the PMI method, the inclusion of a Ru(II) center increases the 3-dimensionality of staurosporine quite significantly (Figure 4.5). In one representative example the authors crystalized a DAPK1 (PDB 2YAK) selective staurosporine derivative OS4 with the protein kinase.[37] When the molecular structure is compared to the PDB structure, there is no significant change in the 3D topology, suggesting the shape is relevant to the selectivity of the inhibitor. In addition to these and prior efforts which aim to augment existing organic therapeutics, the work in Chapter 5 introduces the creation of a 3D metallofragment (mF) library (Figure 4.S5).[30] A PMI analysis of 13 compounds from this library, including metallocenes, piano-stool, and octahedral complexes, found that 100% of the compounds fall outside of flatland (Figure 4.5). This mF library broadly covers 3D chemical space with a relatively small number of molecules and starts to address the need for topologically diverse feedstocks for drug discovery campaigns.
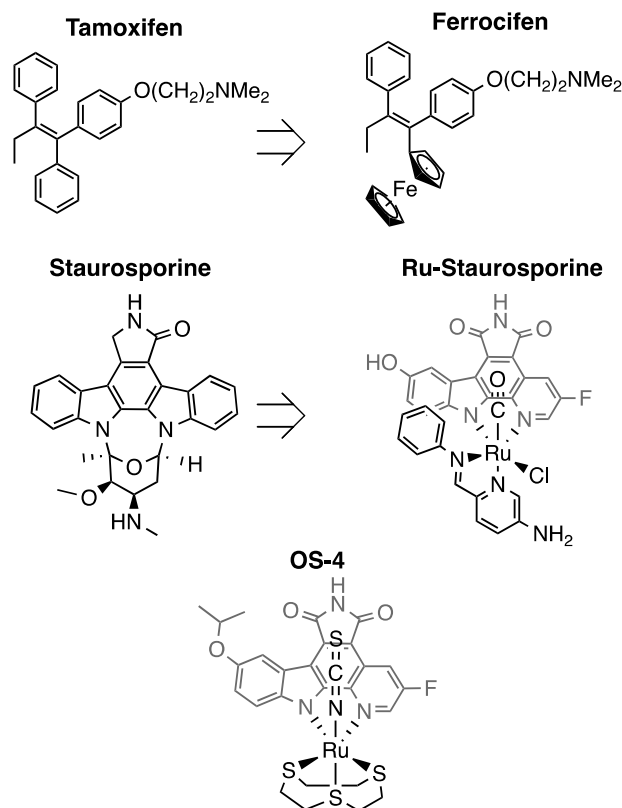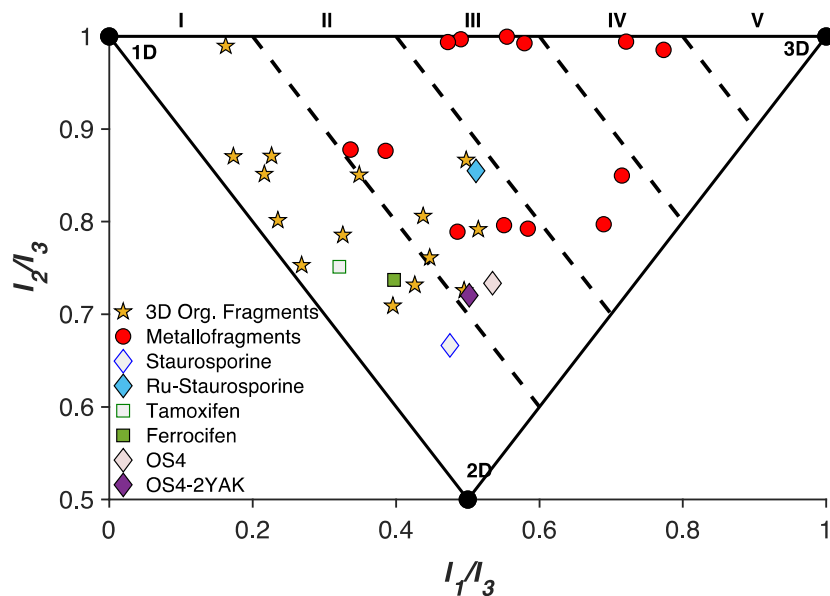
Figure 4.5. PMI analysis of 3D organic fragments (☆) and mFs (o). PMI analysis and structures of tamoxifen and staurosporine, along with their metal containing analogues/derivatives ferrocifen, Ru-staurosporine, and OS-4.

These efforts to prepare both organic and inorganic libraries with high 3-dimensionality will allow for exploration of biological space that has previously been inaccessible with the limited 1D/2D molecular entities.

4.5 Conclusion

This chapter explores the three-dimensionality of large molecule libraries that are relevant in drug discovery efforts. It was found that both energy-minimized (DrugBank) and protein-bound conformations (PDB) of small molecule drugs are overwhelmingly linear and planar. This holds true across a wide range of MWs. Less than 1% of any of the molecular conformations surveyed fall in the highest 3D-area (**V**, 3D Score >1.8). While the 3D score of any particular molecule is partly due to the intrinsic ability of the molecule to adopt a 3D conformation, the capacity of the target protein to accept a 3D conformation also plays a critical role in determining the final bound shape of the molecule. The features of an active site, such as nearby or reactive residues, are already taken into consideration in drug design and in the modeling of bound inhibitors. This analysis demonstrates the relevance of the overall shape of the binding site to the final protein-drug complex.

The true need for topological diversity in feedstocks and final drug molecules remains unclear given the overwhelming number of linear and planar drugs. The question remains as to whether more 3D compounds represent attractive and untapped therapeutic space, or if more linear/planar molecules are indeed the best topologies for bioactive molecules. Similarly, is the linearity and planarity of current ligands because they produce the most bioactive compounds, or are they simply a consequence of the scarcity of highly 3D molecules in the pipeline? Given the range of proteins currently considered undruggable,[38, 39] or even those targets for which no therapy

has been successfully designed, it is anticipated that a broader survey of protein target sites will yield greater shape diversity. Furthermore, the development of topologically rich and varied feedstocks will subsequently yield elaborated bioactive molecules with high 3-dimensionality that may allow for the development of molecules with greater potency and specificity than their planar predecessors. Chapter 5 will introduce metal complexes as highly 3D fragments to help address the lack of three-dimensionality in molecular libraries demonstrated here.

4.6 Acknowledgements
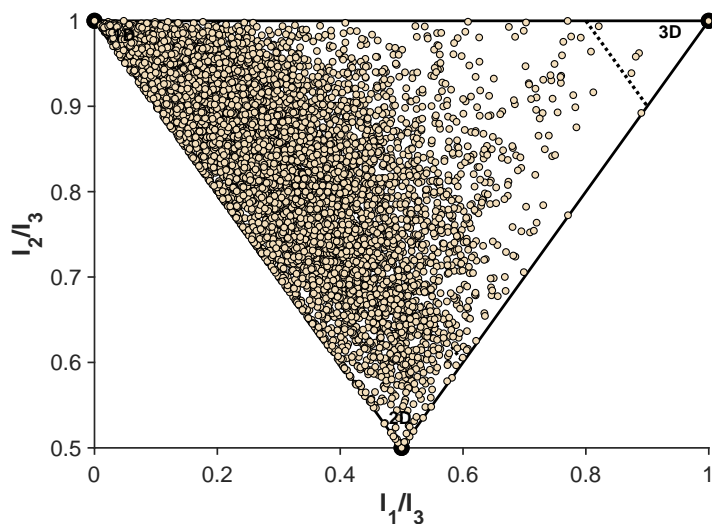
Figure 4.S1. PMI plot for all 8532 structures in the DrugBank downloaded as 3D minimized structures.

Figure 4.S2. PMI values for PDB ligands that are DrugBank entries, having MW >100 g/mol. N = 7411.
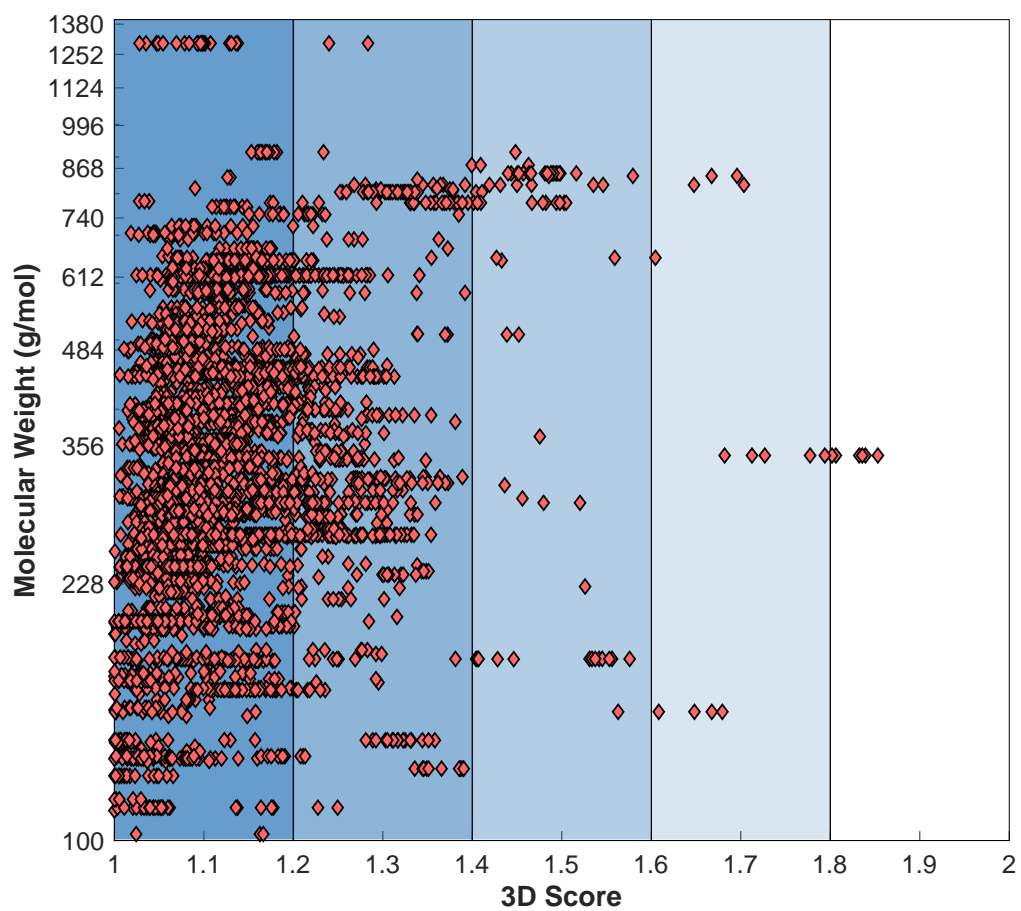
Figure 4.S3. The 7411 entries for ligands in the PDB that have DrugBank identities evaluated by MW and their 3D score.

3D Fragment library chemical structures
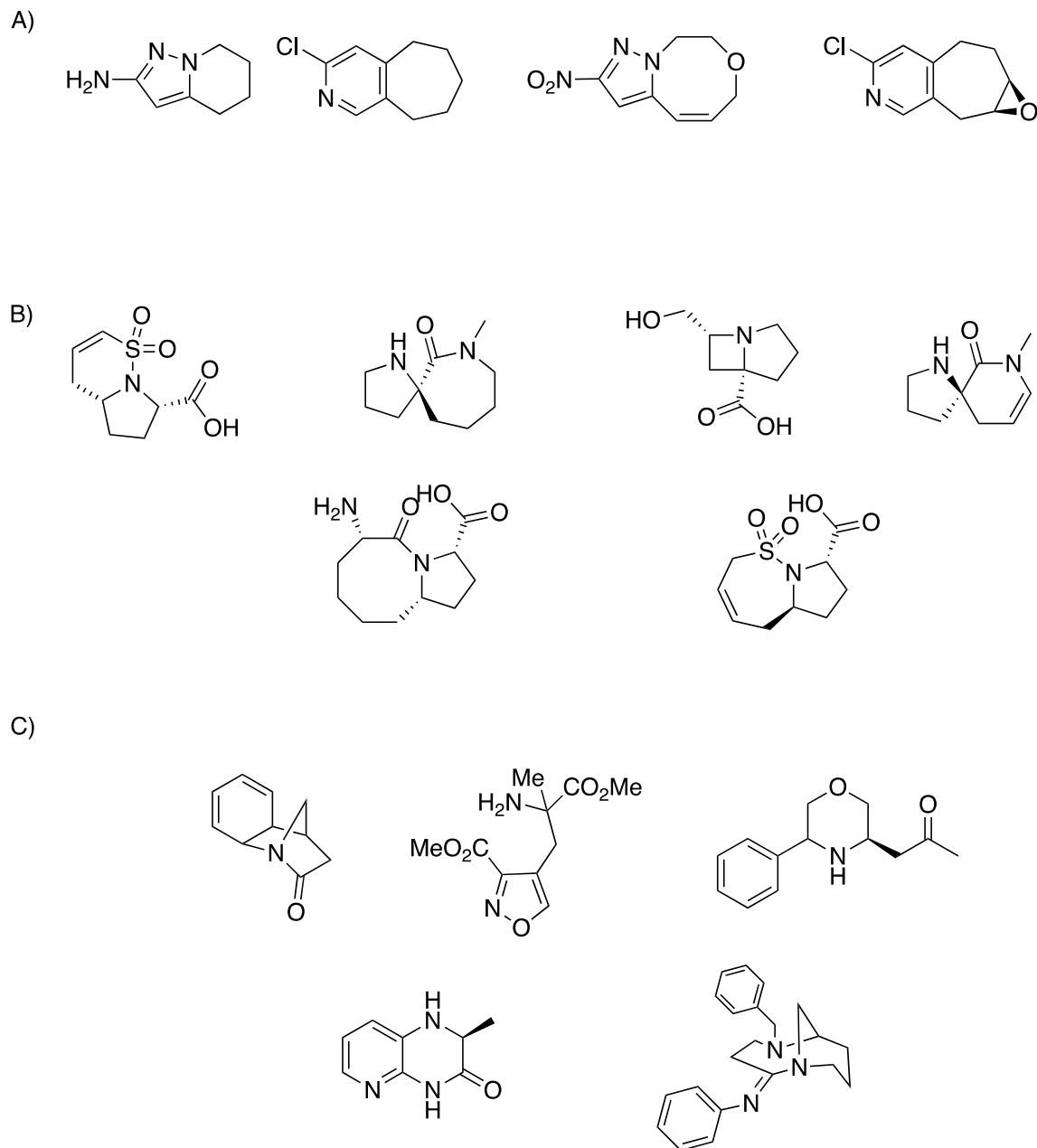
Figure 4.S4. Set of 15 organic fragments from three literature reports (A[32] B[7] C[33]) aimed at preparing 3D organic fragments.
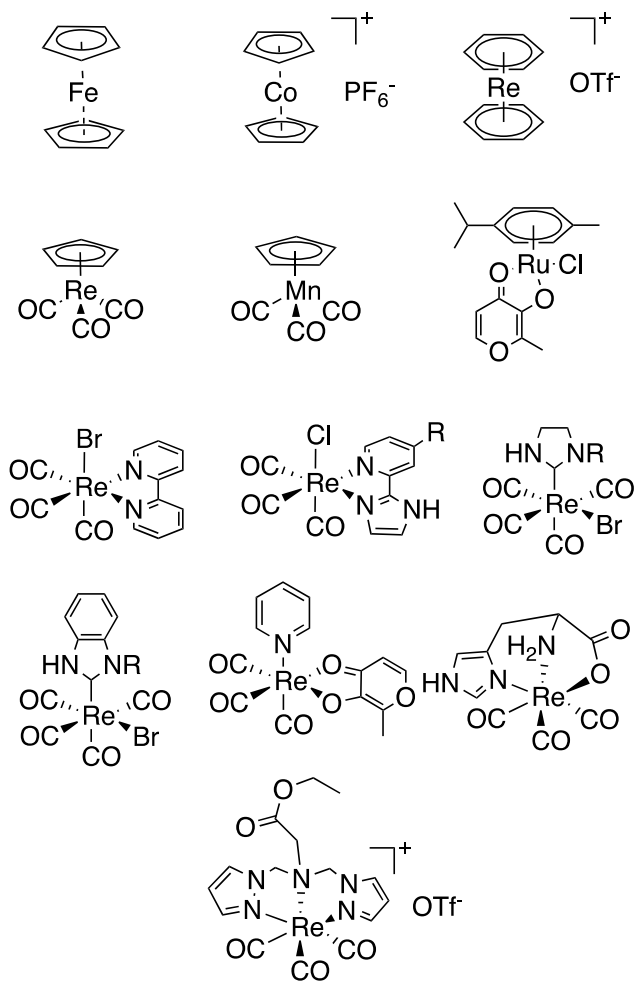
Figure 4.S5. Inorganic mFs discussed in this work as having greater 3D topological diversity.[30]

PDB ID Values for HIV Drug Conformations

Table 4.S1. The HIV therapies and the associated PDB ID values used to evaluate therapeutic conformations in the active site. Percent sequence identity for nevirapine and efavirenz-bound structures was calculated against 3LP1 for HIV reverse transcriptases. Percent sequence identity for nelfinavir, lopinavir, and saquinavir-bound structures was calculated against 3EKX for HIV proteases.

Compound	PDB IDs
Nevirapine	1FKP, 1JLB, 1JLF, 1LW0, 1LWC, 1LWE, 1LWF, 1S1U, 1S1X, 1VRT, 2HND, 2HNY, 3HVT, 3LP0, 3LP1, 3QIP, 5HBM.
Efavirenz	1FK9, 1FKO, 1IKV, 1IKW, 1JKH.
Nelfinavir	1OHR, 2PYM, 2PYN, 2Q63, 2Q64, 2QAK, 2R5Q, 3EKX, 3EL0, 3EL5.
Lopinavir	1MUI, 1RV7, 2O4S, 2Q5K, 2QHC, 2RKF, 2RKG, 2Z54, 4L1A, 6DJ1, 6DJ2.
Saquinavir	1C6Z, 1HXB, 2NMY, 2NMZ, 2NNK, 2NNP, 3CYX, 3D1X, 3D1Y, 3EKQ, 3EL4, 3K4V, 3NDT, 3NDU, 3OXC, 3PWR, 3S56, 3TKG, 3TL9, 3UFN, 4QGI, 5KQX, 5KR2.

4.8 References

1.   Friedrich, N.-O.;   Simsir, M.; Kirchmair, J., How Diverse Are the Protein-Bound Conformations of Small-Molecule Drugs and Cofactors? *Front. Chem.* **2018,** *6*, 68.

2.   Lovering, F.; Bikker, J.; Humblet, C., Escape from Flatland: Increasing Saturation as an Approach to Improving Clinical Success. *J. Med. Chem.* **2009,** *52* (21), 6752-6756.

3.   Firth, N. C.; Brown, N.; Blagg, J., Plane of Best Fit: A Novel Method to Characterize the Three-Dimensionality of Molecules. *J. Chem. Inf. Model.* **2012,** *52* (10), 2516-2525.

4.   Sauer, W. H. B.; Schwarz, M. K., Molecular Shape Diversity of Combinatorial Libraries: A Prerequisite for Broad Bioactivity. *J. Chem. Inf. Model.* **2003,** *43* (3), 987-1003.

5.   Leach, A. R.; Gillet, V. J.; Lewis, R. A.; Taylor, R., Three-Dimensional Pharmacophore Methods in Drug Discovery. *J. Med. Chem.* **2010,** *53* (2), 539-558.

6.   Kortagere, S.;   Krasowski, M. D.; Ekins, S., The importance of discerning shape in molecular pharmacology. *Trends Pharmacol. Sci.* **2009,** *30* (3), 138-147.

7.   Hung, A. W.; Ramek, A.; Wang, Y.; Kaya, T.; Wilson, J. A.; Clemons, P. A.; Young, D. W., Route to three-dimensional fragments using diversity-oriented synthesis. *PNAS* **2011,** *108* (17), 6799.

8.   Kalászi, A.; Szisz, D.; Imre, G.; Polgár, T., Screen3D: A Novel Fully Flexible High-Throughput Shape-Similarity Search Method. *J. Chem. Inf. Model.* **2014,** *54* (4), 1036-1049.

9.   Vasudevan, S. R.; Moore, J. B.; Schymura, Y.; Churchill, G. C., Shape-Based Reprofiling of FDA-Approved Drugs for the H1 Histamine Receptor. *J. Med. Chem.* **2012,** *55* (16), 7054-7060.

10.   Gadhe, C. G.; Lee, E.; Kim, M.-h., Finding new scaffolds of JAK3 inhibitors in public database: 3D-QSAR models & shape-based screening. *Arch. Pharmacal Res.* **2015,** *38* (11), 2008-2019.

11.   Reymond, J.-L., The Chemical Space Project. *Accounts of Chemical Research* **2015,** *48* (3), 722-730.

12.   Perola, E.; Charifson, P. S., Conformational Analysis of Drug-Like Molecules Bound to Proteins: An Extensive Study of Ligand Reorganization upon Binding. *J. Med. Chem.* **2004,** *47* (10), 2499-2510.

13.   Friedrich, N.-O.; Meyder, A.; de Bruyn Kops, C.; Sommer, K.; Flachsenberg, F.; Rarey, M.; Kirchmair, J., High-Quality Dataset of Protein-Bound Ligand Conformations and Its Application to Benchmarking Conformer Ensemble Generators. *J. Chem. Inf. Model.* **2017,** *57* (3), 529-539.

14.   Wishart, D. S.; Feunang, Y. D.; Guo, A. C.; Lo, E. J.; Marcu, A.; Grant, J. R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z.; Assempour, N.; Iynkkaran, I.; Liu, Y.; Maciejewski, A.; Gale, N.; Wilson, A.; Chin, L.; Cummings, R.; Le, D.; Pon, A.; Knox, C.; Wilson, M., DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* **2018,** *46* (D1), D1074-d1082.

15.   Meyers, J.; Carter, M.; Mok, N. Y.; Brown, N., On the origins of three-dimensionality in drug-like molecules. *Future Med. Chem.* **2016,** *8* (14), 1753-1767.

16.   Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **2001,** *46* (1-3), 3-26.

17.   Lipkus, A. H.; Watkins, S. P.; Gengras, K.; McBride, M. J.; Wills, T. J., Recent Changes in the Scaffold Diversity of Organic Chemistry As Seen in the CAS Registry. *J. Org. Chem.* **2019,** *84* (21), 13948-13956.

18.   Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E., The Protein Data Bank. *Nucleic Acids Res.* **2000,** *28* (1), 235-242.

19.   Boström, J.; Norrby, P.-O.; Liljefors, T., Conformational energy penalties of protein-bound ligands. *J. Comput.-Aided Mol. Des.* **1998,** *12* (4), 383-383.

20.   Nicklaus, M. C.; Wang, S.; Driscoll, J. S.; Milne, G. W. A., Conformational changes of small molecules binding to proteins. *Bioorg. Med. Chem.* **1995,** *3* (4), 411-428.

21.   Günther, S.; Senger, C.; Michalsky, E.; Goede, A.; Preissner, R., Representation of target-bound drugs by computed conformers: implications for conformational libraries. *BMC Bioinf.* **2006,** *7*, 293-293.

22.   Vieth, M.; Hirst, J. D.; Brooks, C. L., Do active site conformations of small ligands correspond to low free-energy solution structures? *J. Comput.-Aided Mol. Des.* **1998,** *12* (6), 563-572.

23.   Margolis, A. M.; Heverling, H.; Pham, P. A.; Stolbach, A., A review of the toxicity of HIV medications. *J. Med. Toxicol.* **2014,** *10* (1), 26-39.

24.   Tie, Y.; Wang, Y. F.; Boross, P. I.; Chiu, T. Y.; Ghosh, A. K.; Tozser, J.; Louis, J. M.; Harrison, R. W.; Weber, I. T., Critical differences in HIV-1 and HIV-2 protease specificity for clinical inhibitors. *Protein Science* **2012,** *21* (3), 339-350.

25.   Agniswamy, J.; Shen, C.-H.; Aniana, A.; Sayer, J. M.; Louis, J. M.; Weber, I. T., HIV-1 Protease with 20 Mutations Exhibits Extreme Resistance to Clinical Inhibitors through Coordinated Structural Rearrangements. *Biochemistry* **2012,** *51* (13), 2819-2828.

26.   Liang, J.; Edelsbrunner, H.; Woodward, C., Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. *Protein science : a publication of the Protein Society* **1998,** *7* (9), 1884-1897.

27.   Le Guilloux, V.; Schmidtke, P.; Tuffery, P., Fpocket: An open source platform for ligand pocket detection. *BMC Bioinformatics* **2009,** *10* (1), 168.

28.   Weisel, M.; Proschak, E.; Schneider, G., PocketPicker: analysis of ligand binding-sites with shape descriptors. *Chem Cent J* **2007,** *1*, 7-7.

29.   Pu, L.; Govindaraj, R. G.; Lemoine, J. M.; Wu, H.-C.; Brylinski, M., DeepDrug3D: classification of ligand-binding pockets in proteins with a convolutional neural network. *PLoS Comput. Biol.* **2019,** *15* (2), e1006718.

30.   Morrison, C. N.; Prosser, K. E.; Stokes, R. W.; Cordes, A. L.; Metzler-Nolte, N.; Cohen, S. M., Expanding Medicinal Chemistry into 3D Space: Metallofragments as 3D Scaffolds for Fragment-Based Drug Discovery. *Chem. Sci.* **2020,** *11*, 1216-1225.

31.   Morton, D.; Leach, S.; Cordier, C.; Warriner, S.; Nelson, A., Synthesis of Natural-Product-Like Molecules with Over Eighty Distinct Scaffolds. *Angew. Chem., Int. Ed.* **2009,** *48* (1), 104-109.

32.   Twigg, D. G.; Kondo, N.; Mitchell, S. L.; Galloway, W. R. J. D.; Sore, H. F.; Madin, A.; Spring, D. R., Partially Saturated Bicyclic Heteroaromatics as an sp3-Enriched Fragment Collection. *Angew. Chem., Int. Ed.* **2016,** *55* (40), 12479-12483.

33.   Kidd, S. L.; Osberger, T. J.; Mateu, N.; Sore, H. F.; Spring, D. R., Recent Applications of Diversity-Oriented Synthesis Toward Novel, 3-Dimensional Fragment Collections. *Front. Chem.* **2018,** *6*, 460.

34.   Meggers, E., Targeting proteins with metal complexes. *Chem. Commun.* **2009,** (9), 1001-1010.

35.   Maksimoska, J.; Feng, L.; Harms, K.; Yi, C.; Kissil, J.; Marmorstein, R.; Meggers, E., Targeting Large Kinase Active Site with Rigid, Bulky Octahedral Ruthenium Complexes. *J. Am. Chem. Soc.* **2008,** *130* (47), 15764-15765.

36.   Meggers, E.; Atilla-Gokcumen, G. E.; Bregman, H.; Maksimoska, J.; Mulcahy, S. P.; Pagano, N.; Williams, D. S., Exploring chemical space with organometallics: ruthenium complexes as protein kinase inhibitors. *Synlett* **2007,** *2007* (08), 1177-1189.

37.   Feng, L.; Geisselbrecht, Y.; Blanck, S.; Wilbuer, A.; Atilla-Gokcumen, G. E.; Filippakopoulos, P.; Kräling, K.; Celik, M. A.; Harms, K.; Maksimoska, J.; Marmorstein, R.; Frenking, G.; Knapp, S.; Essen, L.-O.; Meggers, E., Structurally Sophisticated Octahedral Metal Complexes as Highly Selective Protein Kinase Inhibitors. *Journal of the American Chemical Society* **2011,** *133* (15), 5976-5986.

38.   Dang, C. V.; Reddy, E. P.; Shokat, K. M.; Soucek, L., Drugging the 'undruggable' cancer targets. *Nat. Rev. Cancer* **2017,** *17* (8), 502-508.

39.   Russo, A.; Aiello, C.; Grieco, P.; Marasco, D., Targeting "Undruggable" Proteins: Design of Synthetic Cyclopeptides. *Curr. Med. Chem.* **2016,** *23* (8), 748-62.

Chapter 5: Expanding Medicinal Chemistry into 3D Space: Metallofragments as 3D Scaffolds for Fragment-based Drug Discovery

5.1 Introduction

As noted in Chapter 4, it remains a challenge to create a high degree of structural diversity in fragment libraries.[1-5] However, structural diversity is highly desirable because molecular shape is among the most important factors dictating biological effects of molecules, such as recognition by a biological target.[1-2, 6-7] Additionally, increased 3D shape can lead to greater aqueous solubility due to greater solvation and poorer solid-state crystal lattice packing, as well as improved ADMET properties.[6] As a result, increasing the 3D shape of molecules has been correlated to broader biological activity.[1, 7] It has also been shown that molecular shape is more strongly dictated by the core compound scaffold rather than the shape or positioning of substituents decorating the core scaffold,[1, 7] further incentivizing the exploration of fragment libraries consisting of a variety of 3D scaffolds, which are expected to display a wider range of biological activities compared to single scaffold libraries.[1-2, 5-7]

As an example of the dominance of 2D structures in FBDD, a previously reported analysis of 18,534 organic fragments from the ZINC database (a collection of commercially available chemicals used for virtual screening) showed that the majority (~75%) of conventional fragments have a linear (1D) or planar (2D) shape (Figure 5.1).[1] The ZINC database was analyzed using the method of Sauer and Schwartz,[2] as detailed in Chapter 4.
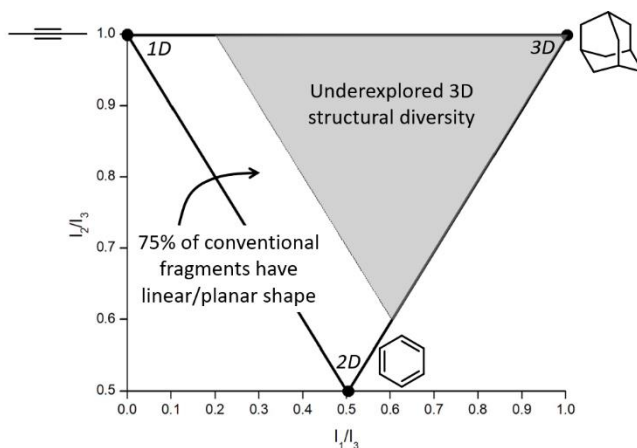
Figure 5.1. Normalized PMI values of a molecule can be plotted to assess molecular topology. Analysis of the ZINC database showed that ~75% of conventional fragments have a linear/planar shape (fall in the white region of the plot), indicating that fragments with 3D topology (gray region of the plot) are vastly underexplored in FBDD.

The lack of structural diversity in fragment libraries is in large part due to the challenge of producing small, organic molecules with inherent 3D shape.[1] Efforts to create 3D organic fragments have included diversity-oriented synthesis,[1, 4, 7] combinatorial libraries,[2] incorporation of cubanes,[8] and incorporation of chiral carbon atoms,[3] all of which pose significant synthetic challenges. This chapter discusses how the highly flat FBDD landscape may be addressed by introducing the first mF library composed entirely of small inorganic complexes with inherent 3D topologies. This small, proof-of-concept library consists of 71 compounds divided into 13 different classes based on metal center and structural homology. The prospective value of this mF library is demonstrated by screening the library against three therapeutic targets, including an antiviral target, an antibacterial target, and an anticancer target. The specific proteins screened in this study are PA_N endonuclease, New Delhi metallo-β-lactamase-1 (NDM-1), and the N-terminal domain of heat shock protein 90-α (Hsp90), respectively. The findings of this chapter show that 3D mFs are an innovative and potentially useful new tool for FBDD that are capable of targeting topological space not readily accessible by conventional organic fragment libraries.

5.2 Metallofragment Library Design

The bioinorganic community has explored the use of coordination and organometallic compounds as inhibitors or as auxiliary groups to augment existing organic moieties,[9-14] and some of these metal-containing inhibitors having entered clinical trials, such as ferroquine.[15-16] A few uses of organometallic groups to augment existing organic inhibitors have produced spectacular results, including highly selective and active kinase inhibitors.[10, 17] However, none of these efforts have approached the use of coordination compounds as fragments for FBDD. This is an important distinction from prior studies: rather than using the metal complex alone or to augment an existing molecule, the approach presented here starts with the coordination compound as a core structural scaffold. Additionally, Dyson and coworker have recently presented a new type of fragment-based approach using metal complexes.[18] Their approach involves merging known, bioactive metal compounds to create bimetallic compounds with the potential to form short- or long-range crosslinks in DNA and protein targets.[18] This is substantially different from our approach, which focuses on 3D mononuclear complexes as structural scaffolds that can be elaborated into more drug-like molecules through modification/elaboration of the ligand components of the fragments.

The heart of this work is a novel mF library, which consists of 13 classes of various sandwich, half-sandwich, and octahedral metal complexes (Figure 5.2). Members within each class share the same metal and core geometry, but feature ligands with different functional groups and/or heterocycles. Approximately 15% of the library was purchased from commercial sources and used without further modification, while the remaining majority of complexes were prepared according to literature procedures (see Appendix for details). Of the prepared metal compounds, ~30% (19/71) represent previously unreported chemical entities. The majority of these novel complexes are ruthenium arene derivatives and rhenium tricarbonyl complexes (Classes **J** and **K**).

In addition to their described application in this work as mFs, these new compounds will have potential applications in the development of new anticancer agents (Ru(II) arenes)[19-22] and model agents for imaging applications (Re(I) tricarbonyls).[23] The majority of the ligands in the complexes in Classes **J** and **K** are derived from previously reported metal-binding pharmacophores.[24] The other synthesized compounds include metallocene and piano-stool derivatives with substituent modifications carried out on their respective aromatic rings (Classes **A**, **B**, **D**, **E**),[25] and a variety of carbene and diimine Re(I) complexes that have been reported extensively elsewhere.[26-29] One particular advantageous aspect of this library is that most complexes, both reported and novel, were prepared in one or two steps from commercially available starting material. This presents the opportunity in this and future studies to rapidly expand the contents of the mF library, in part due to the intrinsically modular nature of ligand and metal complex syntheses. In addition to library expansion, the modularity of ligand synthesis will allow for fragment growth or linking, as is the practice in traditional FBDD campaigns. Indeed, the mFs could potentially be screened in tandem with organic fragment libraries, providing insight to the appropriate ligand substituents needed to achieve potent target binding. Every class of mFs contain positions for modification, via ligand modification, that will facilitate future hit-to-lead or fragment building efforts.

To explore extensive chemical space, features such as charge (e.g., neutral Class **A** versus cationic Class **B**), hydrophobicity, synthetic accessibility, structural diversity, aqueous stability, and rule-of-three compliance (see below) were used to guide this initial library design.[30] Fragments were selected with the aim of having a kinetically and thermodynamically stable core, ideally posing no significant pharmacokinetic challenge beyond that found for conventional organic fragments. Each compound in the library presented herein consists of only one metal ion.

Other studies have examined the utility of metal cluster compounds in drug discovery;[18, 31] however, these tend to be much larger compounds, which makes them less suitable for FBDD.

Within the mF library there are three sub-groups described by their general structure (Figure 5.2): sandwich or metallocene complexes (Classes **A**-**C**), half-sandwhich or piano-stool complexes (Classes **D**, **E** and **J**), and six-coordinate octahedral complexes (Classes **F**-**I** and **K**-**M**). Class **A** contains ferrocene derivatives, which are one of the most common metal-containing scaffolds explored in medicinal bioinorganic chemistry due to their ease of functionalization, stability, and low cost.[11, 32] Ferrocene was first introduced as a bioisostere for aryl/heteroaryl rings, and ferrocene has been utilized to improve anticancer, antimalarial, and antibacterial properties of organic therapies.[11, 21, 32] Classes **B** and **C** are comprised of cobaltocenes and bis(arene)rhenium scaffolds, which are structurally similar to Class **A** but possess a positive charge.

The half-sandwich compounds included in our library are Re(I) compounds (Class **D**) that have been used for biomedical imaging applications,[33] Mn(I) complexes (Class **E**) that have been used as CO releasing agents, and Ru(II) agents (Class **J**) that have been extensively studied as potential therapeutics in their own right.[11, 16, 21 20] Despite the large number of reports on the biological activity of the Ru(II) arene complexes, this work represents, to the best of our knowledge, the first attempt to use Ru(II) arene complexes as core structural scaffolds.

Sandwich complexes:
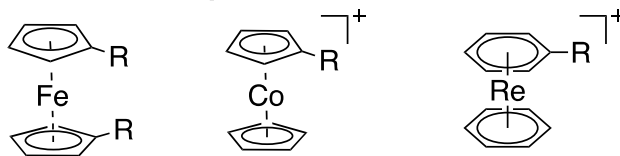
Class A   Class B   Class C

Half-sandwich/piano stool complexes:
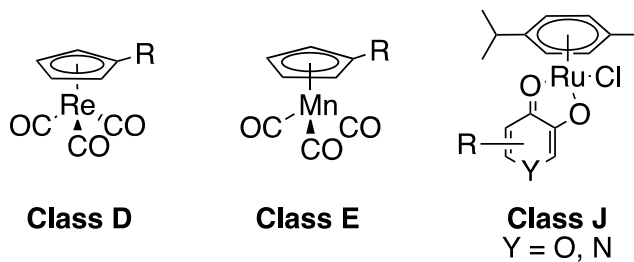
Class D   Class E   Class J
                                  Y = O, N

Octahedral complexes:
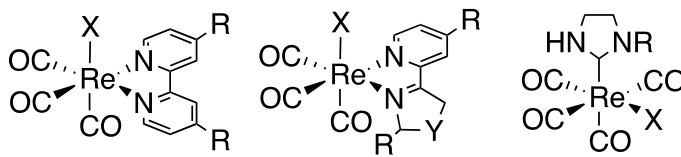
Class F   Class G   Class H
                 Y = O, S
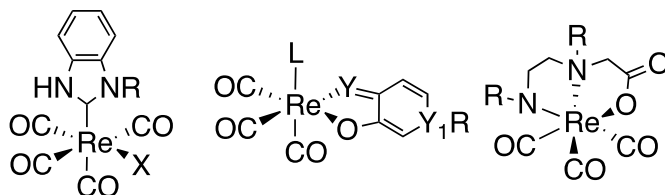
Class I   Class K   Class L
                 Y = O, S
                 Y1 = O, N
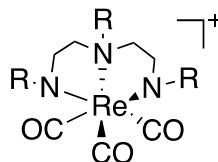
Class M

Figure 5.2. Classes of compounds in the mF library, separated into sub-groups defined by their overall geometry.

All the octahedral complexes presented in our mF library contain Re(I) with 3-4 carbonyl ligands (Classes **F-I** and **K-M**, Figure 5.2). Complexes similar to those in Classes **F** and **G** with bidentate *N,N* donors have been investigated for their anticancer properties.[14, 28] Classes **H** and **I** are carbene complexes; most reported biologically active metal-carbene complexes are prepared with Ag(I) and Au(I) and have shown anticancer and antimicrobial properties. [11, 21] Class **K** molecules consist of O,O and S,O heterocyclic bidentate ligands, while Classes **L** and **M** are prepared with N,N,O (Class **L**) and N,N,N (Class **M**) tridentate donor ligands. Again, to the best of our knowledge, this work is the first time that the Re(I)-based compounds in Classes **H**, **I**, **K**, **L**, and **M** are being utilized in FBDD. It is worth noting that octahedral complexes in the library containing asymmetric bidentate ligands (Classes **G** and **K**) form enantiomeric mixtures due to different binding orientations of the bidentate ligand. Small molecule crystal structures of some Class **K** mFs (Figure 5.S14) show that both enantiomers are present in the product. No attempt to separate the enantiomers was made as it is beyond the scope of this work. Use of enantiomeric mixtures in early-stage drug discovery is commonplace and as such does not represent a significant shortcoming of the mF library. Future work dedicated to hit-to-lead development of Class **G** and **K** molecules will address purification of the enantiomers.

5.3 Redefining the 'rule of three' for mFs

The concept of 'drug-like properties' is constantly evolving. For example, it was recently shown that the average molecular weight of drug molecules has increased substantially in the last 20 years, and its validity as an indicator of drug-likeness has been called into question.[34] As stated earlier, fragments for FBDD are generally designed to be 'rule-of-three' compliant, which includes MW $\leq$300 Da, cLogP $\leq$3, number of hydrogen bond donors $\leq$3, number of hydrogen bond

acceptors ≤3, and ≤3 rotatable bonds.[30] mFs generally satisfy all of these rules except MW ≤300 Da. This rule – much like the Lipinski rule stating a 500 Da cutoff for drug-like molecules[35] – can be considered a proxy to account for molecular size, which can impact permeability and uptake, rather than a strict restriction on MW. Although transition metal ions have a much higher atomic weight than a carbon atom, the actual molecular volume (MV, Å^3) of transition metal ions is not proportionally larger than a carbon atom.

mF K5   X-ray structure   Surface lipophilicity
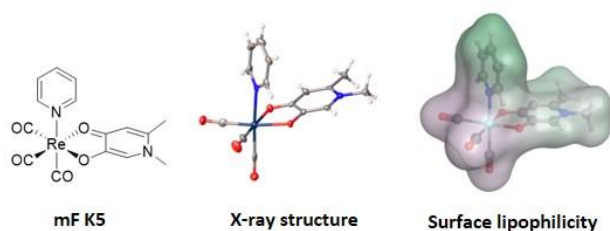
Figure 5.3. Representation of mF **K5** (from left-to-right): chemical structure, X-ray structure, and molecular surface colored by lipophilicity. Hydrophilic and lipophilic regions are represented by pink and green surfaces, respectively. The molecular volume of **K5** was determined to be 292 Å^3 and only one enantiomer is shown from the X-ray structure (see ESI for details).

With this in mind, an analysis of representative mFs was performed to redefine the rule-of-three parameter for MW in terms of molecular size. This redefinition was validated by comparing the heavy atom count (HAC; the number of non-hydrogen atoms) and 'apparent MW' of mFs (where the atomic weight of the metal ion is substituted for a carbon atom) to that of conventional organic fragments. The MV of representative mFs was evaluated against their apparent MW and their HAC. The result of this analysis (Figure 5.S15) shows that the MV of mFs varies in a manner that is indistinguishable from the MV of conventional organic fragments based on HAC and apparent MW. Thus, although the mFs have a greater MW compared to organic fragments, they are not proportionally greater in size, and hence should operate as suitable scaffolds for FBDD. Based on this analysis, in lieu of MW ≤300 Da, we propose a new rule-of-

three for mFs: MV ≤ 300 Å^3. As a representative example, the chemical structure, X-ray structure,

and molecular surface of mF **K5** is shown in Figure 5.3; the molecular volume of **K5** is 292 Å^3

(calculated by Molecular Operating Environment, v. 2019.0101). [36]

5.4 3-Dimensional Analysis of Metallofragments

To confirm that the initial fragment selection encompassed the desired 3D molecular space,

a normalized PMI analysis of each mF was performed as described using the Molecular Operating

Environment program (version 2019.0101, see Appendix for details).[2] The normalized PMI ratios

were benchmarked using the same standards as previously reported (Figure 5.4): 2-butyene (1D),

benzene (2D), and adamantane (3D). As shown in Figure 5.4, the mF library broadly covers the

3D section of the normalized PMI plot. When compared to existing fragment libraries, such as the

ZINC library (Figure 5.1), the mF library covers a much broader 3D topological space using far

fewer compounds. This was quantified by determining the percentage of mFs above the left-hand

boundary given by the equation $x + y = 1.2$ or $(I_1/I_3) + (I_2/I_3) = 1.2$. In our analysis, which is based

on the ZINC library analysis performed by Hung and coworkers,[1] fragments that satisfy the

equation $(I_1/I_3) + (I_2/I_3) > 1.2$ are considered to have a 3D shape. Of the 71 mFs in the library, 55

(77 %) satisfy $(I_1/I_3) + (I_2/I_3) > 1.2$ and can be considered to have a 3D shape. Comparatively,

Hung's analysis of the ZINC library showed that only ~25% of conventional fragments have 3D

shape.[1] Thus, by the normalized PMI metric to analyze molecular shape, the mF library clearly

achieves the goal of providing greater access to 3D scaffolds.

The topology of the mFs was also analyzed by complex type (sandwich, half-sandwich,

and octahedral complexes). The normalized PMI plots (Figure 5.4) show that mFs belonging to

the same complex type tend to have similar topology. The sandwich mFs cluster near the linear

region of the plot, half-sandwich mFs occupy the top of the plot in the linear to spherical region, and octahedral mFs are concentrated between the planar and spherical region. This analysis validates the notion that the core scaffold of a molecule contributes more to its overall shape than does its substituents.[1,7] Knowledge of the general topology of various scaffolds may be useful in targeted FBDD.
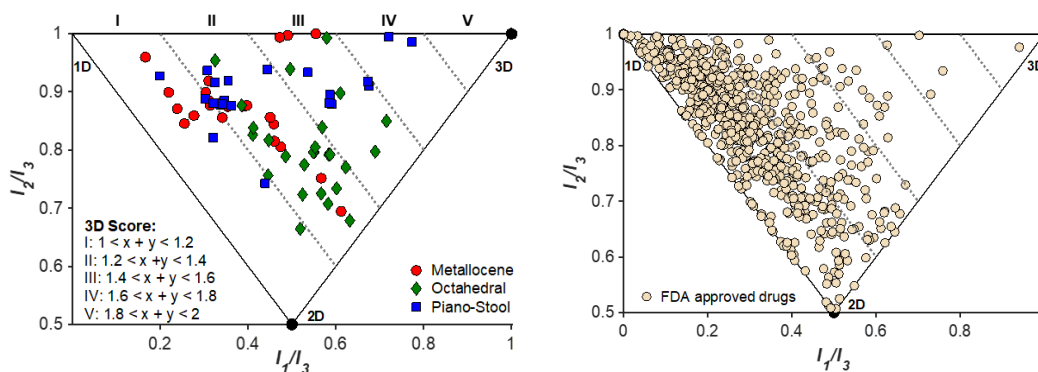
Figure 5.4. *Top*: Normalized PMI analysis of the entire mF library shows that the mF library broadly populates 3D topological space. mFs of each complex type tend to have a related topology. Within the 3D region, metallocene complexes (*red circles*) are more linear/planar, piano-stool complexes (*blue squares*) are more linear/spherical, and octahedral complexes (*green diamonds*) are more planar/spherical. *Bottom*: Normalized PMI analysis of approved drug molecules in DrugBank (v. 5.1.3). In both plots, the degree of increasing 3D topologies, or the 3D scores, is delineated by lines with $x + y \leq 1.2$ (**I** linear/flat), 1.4 (**II**), 1.6 (**III**), 1.8 (**IV**), and $x + y > 1.8$ (**V**).

For comparison to the mF library, the molecular topology of FDA-approved drugs was assessed (Figure 5.4), using structures in the DrugBank (version 5.1.3 from 4 April 2019). Structures were downloaded as '3D', meaning that the downloaded structure represents the lowest energy conformer of the free drug molecule. Normalized PMI calculations of the lowest energy conformer were performed on all structures, and the results are shown in Figure 5.4. To the best of our knowledge, this 3D analysis of approved drugs has not previously been performed and it serves as an interesting insight to the current scope of therapeutic structural diversity. Of the

approved drugs with structures in the DrugBank database 23% (161/712) have 3D scores falling above 1.2 ((I_1/I_3) + (I_2/I_3) > 1.2). The largely 2D character of drug molecules has been previously described,[6] but it is important to note that these energy-minimized structures may not reflect the protein-bound or solution confirmations of the drugs.[6, 37] These energy-minimized structures may also demonstrate some bias towards the left side of the PMI plot due to the steric interactions of side-chains, producing some artificial linearity/planarity. Even concerted efforts in FBDD campaigns to drive towards 3D diversity fail to achieve high 3-dimensionality (3D score > **III**).[1] Of the 712 approved therapies examined here, only 5 (0.7%) compounds are considered highly 3D (3D score > **III**), while with only 71 entries, the mF platform places 2 complexes, ~3% of the library, in this space. Given the 3D nature of these core fragment scaffolds, the mF library offers more direct access to molecules that occupy the previously unexplored 3D space at both the fragment and drug level.

5.5 Metallofragment Library Evaluation and Screening

As described above, many of the sandwich, half-sandwich, and octahedral complexes that comprise the mF library have been broadly examined for their biological activity. In this work the mFs are intended to serve as inert scaffolds upon which fragment growth can be carried out. To determine the general stability of each fragment class, ^{1}H NMR analysis was carried out in deuterated DMSO prior to screening against potential protein targets. Spectra were collected of the first entry in each Class and for each complex **A-I1** and **L-M1** only a single species was observed (Figure 5.S19). Complexes in Class **K** undergo partial solvation through the loss of the monodentate heterocycle to produce a second species with a coordinated DMSO (Figure 5.S18). Such complexes have been reported,[38] and it is anticipated that all components of these solutions

will be aquated once dissolved in aqueous media. Similarly, the Ru(arene) scaffolds in Class **J** exhibit some ligand exchange in DMSO; the speciation of such complexes has been studied extensively in both organic and aqueous media.[39] While these particular scaffolds may not be an ideal mF motif, the relevance of these compounds to the bioinorganic literature and their 3D topological diversity prompted the inclusion of Class **J** as a starting point for these studies.

To assess the utility of the mF library for FBDD, the library was screened against three therapeutically relevant targets (Figure 5.5). The selected targets were: the polymerase acidic N-terminal (PA_N) endonuclease domain from the H1N1 influenza A virus (antiviral target), New Delhi metallo-β-lactamase-1 (NDM-1; antibacterial target), and the N-terminal domain of heat shock protein 90-α (Hsp90; anticancer target).[40-41] The role of these enzymes in their respective diseases and inhibitor development for each are presented elsewhere.[40-42] Briefly, PA_N endonuclease is one of three proteins in the RNA-dependent RNA polymerase complex of the influenza A virus, along with the polymerase basic protein 1 (PB1) and the polymerase basic protein 2 (PB2).[43] PA_N endonuclease contains a dinuclear metal active site, in which two Mn^{2+} or Mg^{2+} cations catalyze endonuclease activity.[44] A functional RNA polymerase complex is essential to viral replication,[43] and the first treatment targeting this protein, Baloxavir marboxil, has now gained FDA approval.[45] Baloxavir, along with other leading drug discovery efforts, have targeted the metal centers nestled in the large active site of PA_N as a means of enzyme inhibition. NDM-1 is a protein found in both Gram-negative and Gram-positive bacteria that has been shown to hydrolyze clinically relevant β-lactam antibiotics.[46] The active site of NDM-1 is largely hydrophobic, with fluctional conformations and two Zn(II) ions that participate in antibiotic hydrolysis.[47] There has been substantial effort to develop NDM-1 inhibitors and the target is still considered promising,[40] though NDM-1 inhibitors have yet to gain FDA approval. The final target,

Hsp90, is a ubiquitous molecular chaperone with many diverse functions including the folding, stability, and activity of many proteins ('clients').[48-49] Several Hsp90 clients have been identified as oncoproteins that are associated with cancer hallmarks.[41, 49] The binding site of the reported inhibitors is a 15 Å deep pocket capable of binding to polypeptide chains.[50] More than a dozen Hsp90 inhibitors with a variety of mechanisms of action have entered clinical trials, but none have received FDA approval.[41, 49]

The mF library was screened at a fragment concentration of 200 µM against all three targets using established, high-throughput screening assays for each protein (Figure 5.5).[44, 51] Considering the library as an entirety, a hit rate (percent inhibition > 50%) of approximately ~28 % (20/71) was achieved against PA_N. Classic high-throughput screening (HTS) of drug-like molecules have reported hit rates between 0.001% and 0.2 %, while organic fragment libraries are reported to have hit rates ranging from 3% to 30%.[52-53] Thus, at least against PA_N, the mF library performs as well as or better than traditional screening libraries, speaking to the promise of this screening platform.

Class **A** ferrocene derivatives were observed to exhibit photosensitivity.[54-55] DMSO stock solutions of certain Class **A** mFs exposed to ambient light resulted in photodecomposition. Many of the derivatives in class A contain a ferrocenyl carbonyl motif, which has been previously reported to undergo photoaquation (λ >280 nm) in wet DMSO to produce a monocyclopentadienyliron cation, the anionic ligand, and free cyclopentadiene.[55] Suspecting issues with photostability, compound **A7** was dissolved in DMSO-d_6, exposed to ambient room light (fluorescent light bulb), and monitored for stability by NMR. Indeed, photoinstability was confirmed by the observance of free cyclopentadienyl peaks appearing in the ^{1}H NMR spectrum. The instability of these compounds highlights the importance of rigorously evaluating fragment stability for each fragment class. Accordingly, screening data for Class **A** compounds is only

presented against PA$_N$, and compound stock solutions were prepared fresh and kept isolated from light prior to evaluation. While these compounds may present promising scaffolds for future work on fragment growth and elaboration, it is important to recognize their general lack of specificity. Excessive lipophilicity, along with properties such as redox activity and metal chelation are considered to be hallmarks of potential ligand promiscuity.[56] The lipophilicity of ferrocene and ferrocene derivatives, discussed in terms of the partition coefficient log P, are typically between 2-5.[57] These values fall on the higher end of what is generally considered acceptable for drug-like molecules (log $P < 5$), and it is likely that their lipophilicity would only increase with fragment growth. Given this, should a Class **A** mF be selected as a hit compound in future efforts, precautions would be advised to ensure that the complex binds to the desired binding pocket rather than non-specific hydrophobic interactions leading to inhibitory responses. Additional care should be taken to ensure that photodecomposition does not occur. The non-specificity and photosensitivity demonstrated by these compounds represent some of the challenges inherent to using metal complexes in FBDD. Accordingly, appropriate precautions should be taken to characterize the stability, speciation, and shelf-life of mFs used in future efforts.
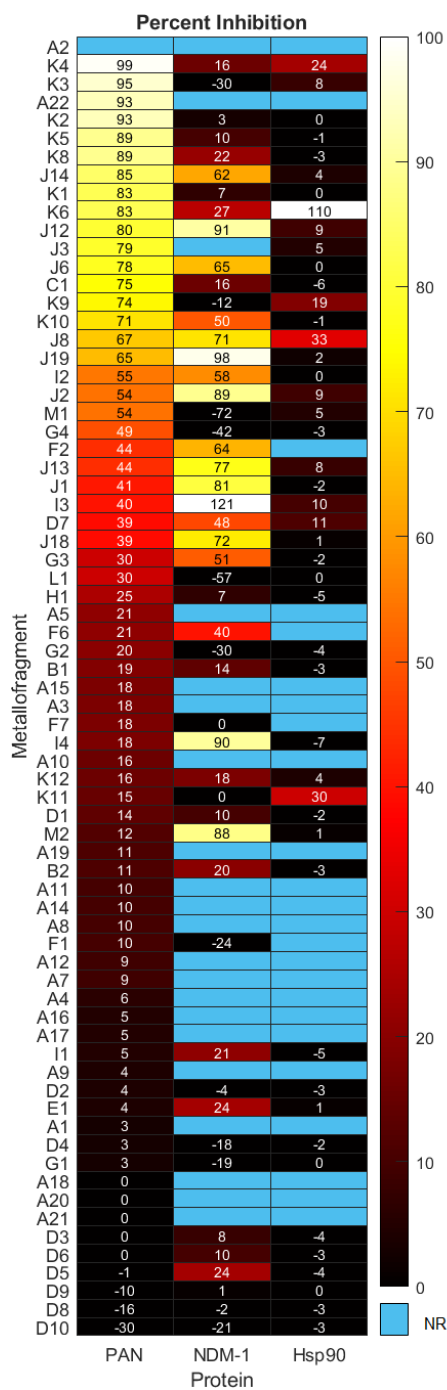
Figure 5.5. Screening results, presented as percent inhibition, for the mF library tested at 200 μM mF concentration against the viral target PA$_N$, the bacterial target NDM-1, and the human cancer target Hsp90 (NR = Not Reported).

In addition to validating the suitability of the mF platform for future FBDD campaigns, these screening results allow for the examination of the relationship between mF 3D topologies

and their biological activity. Based on the PMI analysis presented above, the large majority of mFs have 3D scores of **II** ($1.2 \leq (I_1/I_3) + (I_2/I_3) < 1.4$) compared to the ZINC fragment library where ~75% of fragments have $(I_1/I_3) + (I_2/I_3) < 1.2$, a 3D score of **I**.[1] Within **II** there are mFs from each subgroup, and they all exhibit a broad range of inhibitory effects against the three targets. Of the most 3D fragments, those that fall into **IV**, none of them achieved percent inhibition values above 50%. There are only 2 mFs with this score, representing ~3% of the library, and both are from the half-sandwich complex subgroup. It is challenging to draw any significant conclusions regarding their moderate biological activity given the limited scope. Nonetheless, these highly 3D fragments demonstrated weak inhibition at 200 µM against at least one of the targets, and as such could potentially be pursued for fragment growth and lead development. In future studies it will be important to populate these higher 3D scores so that more informative analyses of the relationship between 3-dimensionality and biological activity can be undertaken. The ability to measure and compare these inhibitors will allow future FBDD campaigns to take full advantage of the topological diversity afforded by the mF library.

5.6 Structure-activity relationship of mFs

In an effort to elucidate the possible mechanism of inhibition, preliminary docking exercises were carried out on a representative mF, **K6**. Fragment **K6** showed good inhibitory activity against PA_N endonuclease, and also stood out as a potent inhibitor of Hsp90, providing a model mF for this docking study. The coordinates of an aquated version of **K6** were determined from the X-ray crystal structure of **K5** (see Appendix Figure 5.S14 and Table 5.S1) and the complex was docked against a reported inhibitor-bound PA_N endonuclease structure (6E3M) and an inhibitor-bound crystal structure of Hsp90 (1YET) using the Molecular Operating Environment

program (version 2019.0101).[58] The best scoring pose against each protein is shown in Figure 5.6. While these poses may not indicate the true binding mode of **K6** against these two targets, they serve as an initial starting point in the rational development of new inhibitors. In both docking studies the proteins and mF are mapped from pink to green based on their lipophilicities (Figure 5.6). Interestingly, many of the mF library entries do not have hydrogen bond donating or accepting entities, as is the case for **K6**. As such, an analysis of the molecular interactions shows that the docking of the mF in the pockets of PA$_N$ endonuclease and Hsp90 is driven primarily by the steric interactions, directly related to the fragment 3-dimensionality.
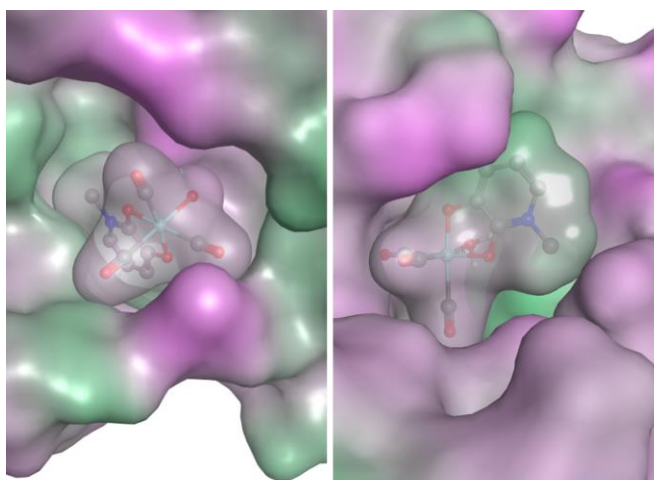
Figure 5.6. Aquated fragment **K6** docked against PA$_N$ endonuclease (left) and Hsp90 (right). Fragments and protein are shown with molecular surface maps colored to indicate lipophilicity. Hydrophilic and lipophilic regions are represented by pink and green surfaces, respectively.

5.7 Conclusions

The use of metal complexes to complement organic fragment libraries for drug discovery applications has not been considered elsewhere. The benefit this approach aims to impart on the FBDD methodology is the ability to access underexplored 3D chemical topologies. To evaluate the feasibility and the validity of such an approach, this chapter details the design, synthesized and

characterization of a modest library of coordination and organometallic complexes with diverse 3D topologies. A comparative shape analysis by the PMI method impressively demonstrates the validity of our approach, with 77% populating the 3D region, as opposed to only 25% from the much larger ZINC library of purely organic compounds. As a proof-of-concept, the mF library was then screened against three different, relevant biological enzyme targets, i.e. PA_N endonuclease, NDM-1, and Hsp90 at 200 µM fragment concentration. These assays generated a range of inhibitory responses, in which some mF classes performed well, while others achieved only moderate to poor inhibition. Through the combination of fragment screening and molecular docking, this chapter demonstrates that these mFs are capable of the same types of analyses undergone by traditional organic fragments in medicinal chemistry campaigns. Such analyses will be critical for future works carrying out HTL development utilizing the mF platform.

An analysis of the 3D topologies of ~700 approved therapeutics demonstrates that a large majority fall under a linear/flat regime. Because of the limitations in the synthesis of 3D rich drug/fragment libraries, it is difficult to know if these flat/linear structures represent ideal geometries, or if they are simply a consequence of the tools used to prepare drug discovery libraries. The mFs presented here have comparatively highly 3D topologies, but this space has been so challenging to access that the potential of truly 3D scaffolds has yet to be determined. Overall, this work showcases the utility of a novel mF library of modest size with 3D diversity that exhibits a broad range of biological responses. Though the methodology applied in this chapter represents a novel approach to FBDD, it must be carefully applied, as metal complexes may be susceptible to challenges not common in traditional FBDD.[54] Future efforts will build on these exciting proof-of-concept studies by expanding the mF library to include more highly 3D fragments with careful consideration of their kinetic and thermodynamic stability. Specifically,

the analysis of a large database of metal complexes will help to identify metal complexes that provide thorough and systematic coverage of chemical space. With a second-generation library in development, these mFs can be further developed into lead-like molecules and previously inaccessible or challenging targets will be examined.

5.8 Acknowledgements

5.9 Appendix: Supporting Information

All reagents and solvents were obtained from commercial sources and used without further purification. Microwave reactions were performed in 10 mL or 35 mL microwave vials using a CEM Discover S reactor. Silica gel column chromatography was performed on a CombiFlash Rf Teledyne ISCO system with prepacked silica cartridges or High-Performance Gold C18 columns. Reverse phase column chromatography (C18 column) was performed on the same instrument using 0.1% formic acid in methanol, acetonitrile, or water as eluent. Separations were monitored by mass spectrometry via a Teledyne ISCO RF[+] PurIon ESI-MS or APCI-MS detector with 1 Da resolution. ^{1}H NMR spectra were recorded at ambient temperature on a 400 or 500 MHz Varian FT-NMR instrument located in the Department of Chemistry and Biochemistry at the University

of California, San Diego. Some ^{1}H NMR spectra were recorded at ambient temperature on a 200, 250, 300, or 400 Bruker FT-NMR instrument located in the Department of Chemistry at the Ruhr University in Bochum, Germany. ^{1}H NMR data is expressed in parts per million (ppm) relative to the residual non-deuterated solvent signals, and spin multiplicities are given as s (singlet), d (doublet), dd (doublet of doublets), t (triplet), dt (doublet of triplets), q (quartet), m (multiplet), and bs (broad singlet). Available coupling constants (J) are reported in hertz (Hz). Mass spectra were obtained at the Molecular Mass Spectrometry Facility (MMSF) in the Department of Chemistry and Biochemistry at the University of California, San Diego. Some mass spectra were obtained in the Department of Chemistry and Biochemistry at the Ruhr University in Bochum, Germany. Further details on synthesis may be found in the following sections of the Supporting Information. The purity of all compounds used in assays was determined to be ≥95% by ^{1}H NMR spectroscopy and confirmed by mass spectrometry.

Synthesis and characterization of mF library

Class A compounds
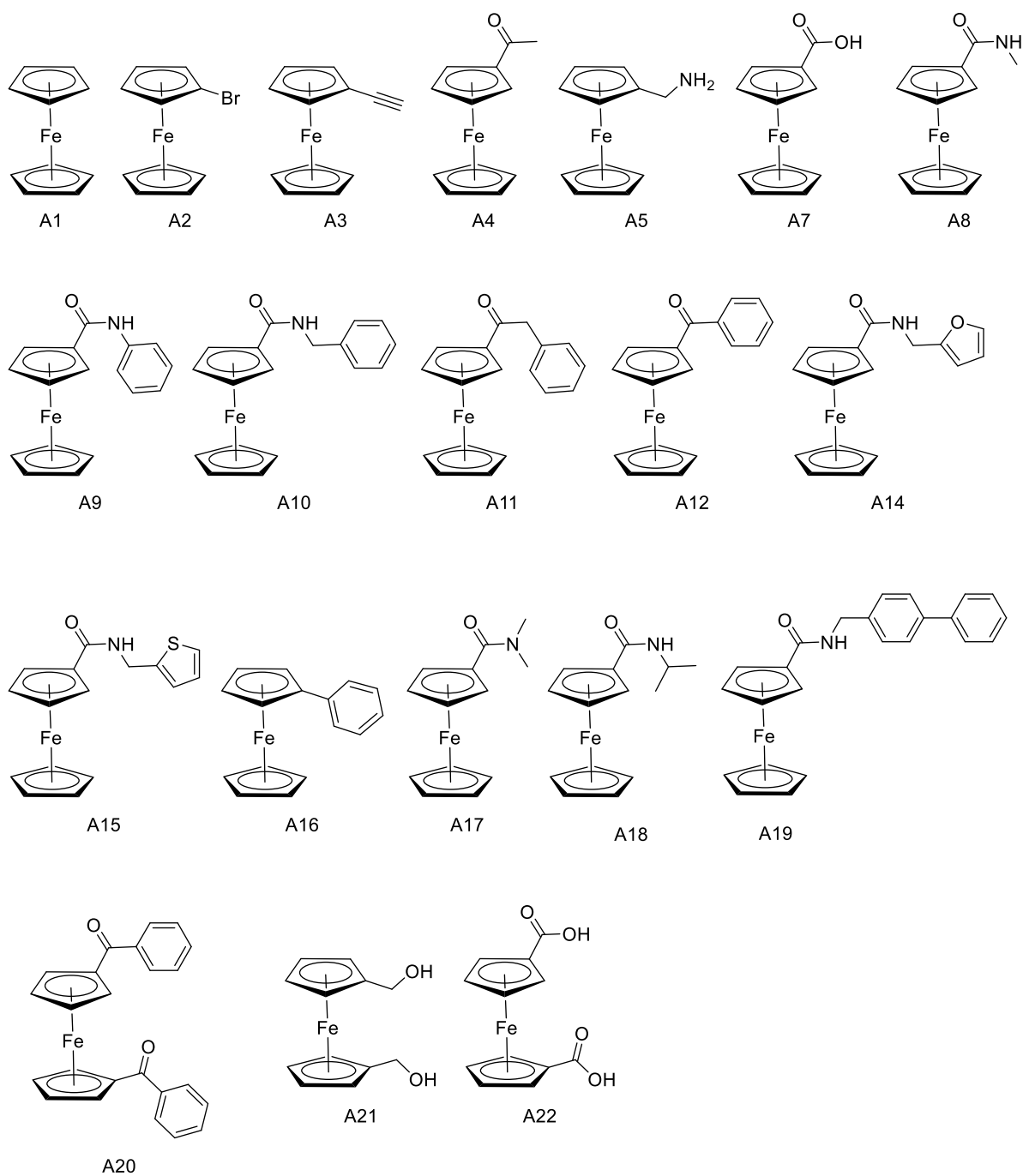
Figure 5.S1. Ferrocene derivatives that comprise the Class **A** mFs.

Fe(η^5-C$_5$H$_5$)$_2$] (ferrocene; A1). Compound A1 was purchased from Sigma Aldrich and used without further purification.

[Fe(η^5-C$_5$H$_5$)(η^5-C$_5$H$_4$Br)] (bromoferrocene; A2). Compound A2 was purchased from Sigma Aldrich and used without further purification.

[Fe(η^5-C$_5$H$_5$)(η^5-C$_5$H$_4$CCH)] (ethynylferrocene; A3). Compound A3 was purchased from Sigma Aldrich and used without further purification.

[Fe(η^5-C$_5$H$_5$)(η^5-C$_5$H$_4$COCH$_3$)] (acetylferrocene; A4). Compound A4 was purchased from Sigma Aldrich and used without further purification.

[Fe(η^5-C$_5$H$_5$)(η^5-C$_5$H$_4$CH$_3$NH$_2$)] (ferrocenemethylamine; A5). Ferroceneoxime was made from ferrocenecarboxaldehyde following a previously reported procedure.[25] To a solution of ferroceneoxime (214 mg, 0.93 mmol) in 12 mL dry THF in a N$_2$-flooded Schlenk flask, LiAlH$_4$ (2.4 M in THF, 2.06 mL) was added slowly. The reaction mixture was stirred for 72 h at room temperature, then poured onto ice. Et$_2$O was added and the resulting gray solid was collected via filtration. The aqueous layer was washed with Et$_2$O, and the combined organic layers were dried over MgSO$_4$ and the solvent was evaporated under reduced pressure to yield A5. Yield: 174 mg (87%). ^{1}H NMR (200 MHz, Chloroform-d_1): δ 4.24 – 4.06 (m, 9H), 3.55 (s, 2H). ESI-MS(+): *m/z* 224.7 [M+H]$^+$.

[Fe(η^5-C$_5$H$_5$)(η^5-C$_5$H$_4$COOH)] (ferrocenecarboxylic acid; A7). Compound A7 was purchased from Sigma Aldrich and used without further purification.

[Fe(η^5-C$_5$H$_5$)(η^5-C$_5$H$_4$CONHCH$_3$)] (A8). 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) and hydroxybenzotriazole (HOBt) were added to a suspension of ferrocenecarboxylic acid in DCM (5 mL). After stirring for 1 h at room temperature, the mixture was cooled to 0 °C and methylamine was added. The reaction mixture was allowed to warm and stir at room temperature for 1 h before the solvent was removed. The product was purified with silica-based column chromatography and then recrystallized from DCM/EtOAc (5:1), yielding orange crystals. Yield: 55%. ^{1}H NMR (500 MHz, DMSO-d_6): δ 7.74 (bs, 1H), 4.74 (t, J = 1.96 Hz, 2H), 4.32 (t, J = 1.83 Hz, 2H), 4.14 (s, 5H), 2.69 (d, J = 4.65 Hz, 3H). ESI-MS(+): m/z 244.22 [M+H]$^+$.

[Fe(η^5-C$_5$H$_5$)(η^5-C$_5$H$_4$CONHC$_6$H$_5$)] (A9). EDC and HOBt were added to a suspension of ferrocenecarboxylic acid in DCM (5 mL). After stirring for 1 h at room temperature, the mixture was cooled to 0 °C and aniline was added. The reaction mixture was allowed to warm and stir at room temperature for 1 h before the solvent was removed. The product was purified with silica-based column chromatography and then recrystallized from DCM/EtOAc (5:1), yielding orange crystals. Yield: 15%. ^{1}H NMR (500 MHz, DMSO-d_6): δ 9.44 (s, 1H), 7.68-7.72 (m, 2H), 7.29 - 7.36 (m, 2H), 7.06 (m, 1H), 5.01 (t, J = 1.9 Hz, 2H), 4.45 (t, J = 1.9 Hz, 2H), 4.22 (s, 5H). ESI-MS(+): m/z 306.43 [M+H]$^+$.

[Fe(η^5-C$_5$H$_5$)(η^5-C$_5$H$_4$CONHCH$_2$C$_6$H$_5$)] (A10). EDC and HOBt were added to a suspension of ferrocenecarboxylic acid in DCM (5 mL). After stirring for 1 h at room temperature, the mixture

was cooled to 0 °C and benzylamine was added. The reaction mixture was allowed to warm and stir at room temperature for 1 h before the solvent was removed. The product was purified with silica-based column chromatography and then recrystallized from DCM/EtOAc (5:1), yielding orange crystals. Yield: 65%. ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.37 (t, J = 6.05 Hz, 1H), 7.30 - 7.36 (m, 4H), 7.20 - 7.26 (m, 1H), 4.82 (t, J = 1.9 Hz, 2H), 4.34 - 4.39 (m, 4H), 4.12 (s, 5H). ESI-MS(+): m/z 320.29 [M+H]$^+$.

[Fe(η^5-C$_5$H$_5$)(η^5-C$_5$H$_4$COCH$_2$C$_6$H$_5$)] (A11). To a stirred solution of ferrocene in DCM (5 mL) at 0 °C was added AlCl$_3$ in portions. A solution of 2-phenylacetyl chloride in DCM (5 mL) was added dropwise over 1 min while keeping the temperature at 0 °C. After addition, the reaction mixture was warmed to room temperature and stirred for 1 h. The reaction mixture was concentrated under reduced pressure and purified with silica-based column chromatography using a gradient of 0-100% Hex/EtOAc. Yield: 90%. ^{1}H NMR (500 MHz, DMSO-d_6): δ 7.28 - 7.38 (m, 4H), 7.20 - 7.27 (m, 1H), 4.88 (t, J = 1.8 Hz, 2H), 4.60 (t, J = 1.8 Hz, 2H), 4.19 (s, 5H), 4.04 (s, 2H). ESI-TOFMS(+): m/z 305.0622 [M+H]$^+$.

[Fe(η^5-C$_5$H$_5$)(η^5-C$_5$H$_4$COC$_6$H$_5$)] (A12). To a stirred solution of ferrocene in DCM (5 mL) at 0 °C was added AlCl$_3$ in portions. A solution of benzoyl chloride in dichloromethane (5 mL) was added dropwise over 1 min while maintaining the temperature at 0 °C. After addition, the reaction mixture was warmed to room temperature and stirred for 1 h. The reaction mixture was concentrated under reduced pressure and purified with silica-based column chromatography using a gradient of 0-100% Hex/EtOAc. Yield: 58%. ^{1}H NMR (500 MHz, DMSO-d_6): δ 7.87 (m, 2H),

7.51-7.64 (m, 3H), 4.83 (m, 2H), 4.70 (m, 2H), 4.25 (s, 5H). ESI-TOFMS(+): *m/z* 291.0467 [M+H]$^+$.

[Fe(η^5-C$_5$H$_5$)(η^5-C$_5$H$_4$CONHCH$_2$C$_4$OH$_3$)] (A14). EDC and HOBt were added to a suspension of ferrocenecarboxylic acid in DCM (5 mL). After stirring for 1 h at room temperature, the mixture was cooled to 0 °C and furanylmethylamine was added. The mixture was stirred for 1 h at room temperature before the solvent was removed under reduced pressure. The product was purified with silica-based column chromatography and then recrystallized from DCM/EtOAc (5:1), yielding orange crystals. Yield: 65%. ^{1}H NMR (500 MHz, DMSO-d_6): δ 8.24 (t, J = 5.87 Hz, 1H), 7.58 (dd, J = 0.86, 1.83 Hz, 1H), 6.41 (dd, J = 1.71, 3.18 Hz, 1H), 6.26 (dd, J = 0.73, 3.18 Hz, 1H), 4.81 (t, J = 1.96 Hz, 2H), 4.37 (d, J = 5.87 Hz, 2H), 4.34 (t, J = 1.96 Hz, 2H), 4.12 (s, 5H). ESI-TOFMS(+): *m/z* 310.0523 [M+H]$^+$. To the best of our knowledge, compound **A14** has not been reported previously in the literature.

[Fe(η^5-C$_5$H$_5$)(η^5-C$_5$H$_4$CONHCH$_2$C$_4$SH$_3$)] (A15). EDC and HOBt were added to a suspension of ferrocenecarboxylic acid in DCM (5 mL). After stirring for 1 h at room temperature, the mixture was cooled to 0 °C and thiophenylmethylamine was added. The mixture was stirred for 1 h at room temperature before the solvent was removed under reduced pressure. The product was purified with silica-based column chromatography and then recrystallized from DCM/EtOAc (5:1), yielding orange crystals. Yield: 71%. ^{1}H NMR (500 MHz, DMSO-d_6): δ 8.44 (t, J = 5.99 Hz, 1H), 7.40 (dd, J = 1.34, 5.01 Hz, 1H), 7.02 (dd, J = 1.22, 3.42 Hz, 1H), 6.97 (dd, J = 3.55, 5.01 Hz, 1H), 4.81 (t, J = 1.96 Hz, 2H), 4.53 (d, J = 5.62 Hz, 2H), 4.35 (t, J = 1.96 Hz, 2H), 4.12 (s, 5H). ESI-TOFMS(+): *m/z* 326.0295 [M+H]$^+$.

[Fe(η^5-C$_5$H$_5$)(η^5-C$_5$H$_4$C$_6$H$_5$)] (phenylferrocene; A16). Aniline (1.2 mmol), H$_2$O (2 ml) and concentrated HCl (2 ml) were combined in a round-bottom flask and cooled to 0 – 5 °C. While maintaining the temperature of the solution below 5 °C, 0.1 g NaNO$_2$ in 2 ml H$_2$O was added dropwise. After the addition, the mixture was stirred for 1–1.5 h at 5 °C. Adequate urea was added to decompose the surplus HNO$_2$. Then, again keeping the temperature below 5 °C, a solution of ferrocene (in 5 ml ether) and 0.1 g cetrimonium bromide (CTAB, C$_{16}$H$_{33}$N(CH$_3$)$_3$Br) was added dropwise over 5 mins with stirring. The solution was stirred for 2 h at room temperature. The solvent was evaporated under reduced pressure, and the product was purified with silica-based column chromatography using EtOAc/Hex (gradient 0-10%). Yield: 18%. ^{1}H NMR (500 MHz, DMSO-d_6): δ 7.53 (m, 2H), 7.29 (m, 2H), 7.18 (m, 1H), 4.78 (m, 2H), 4.34 (m, 2H), 4.01 (s, 5H). ESI-MS(+): m/z 262.46 [M+H]$^+$.

[Fe(η^5-C$_5$H$_5$)(η^5-C$_5$H$_4$CON(CH$_3$)$_2$)] (A17). EDC and HOBt were added to a suspension of ferrocenecarboxylic acid in DCM (5 mL). After stirring for 1 h at room temperature, the mixture was cooled to 0 °C and dimethylamine was added. The mixture was stirred for 1 h at room temperature before the solvent was removed under reduced pressure. The product was purified with silica-based column chromatography and then recrystallized from DCM/EtOAc (5:1), yielding orange crystals. Yield: 76%. ^{1}H NMR (500 MHz, DMSO-d_6): δ 4.60 (t, J = 1.83 Hz, 2H), 4.36 (t, J = 1.96 Hz, 2H), 4.22 (s, 5H), 2.7-3.25 (bd, 6H). ESI-MS(+): m/z 258.20 [M+H]$^+$.

[Fe(η^5-C$_5$H$_5$)(η^5-C$_5$H$_4$CONCH(CH$_3$)$_2$)] (A18). EDC and HOBt were added to a suspension of ferrocenecarboxylic acid in DCM (5 mL). After stirring for 1 h at room temperature, the mixture

was cooled to 0 °C and isopropylamine was added. The mixture was stirred for 1 h at room temperature before the solvent was removed under reduced pressure. The product was purified with silica-based column chromatography and then recrystallized from DCM/EtOAc (5:1), yielding orange crystals. Yield: 72%. ^{1}H NMR (500 MHz, DMSO-d_6): δ 7.47 (d, J = 7.83 Hz, 1H), 4.80 (m, 2H), 4.31 (m, 2H), 4.13 (s, 5H), 3.98 - 4.08 (m, 1H), 1.13 (d, J = 6.60 Hz, 6H). ESI-MS(+): m/z 272.17 [M+H]$^+$.

[Fe(η^5-C$_5$H$_5$)(η^5-C$_5$H$_4$CONHCH$_2$C$_6$H$_4$C$_6$H$_5$)] (A19). EDC and HOBt were added to a suspension of ferrocenecarboxylic acid in DCM (5 mL). After stirring for 1 h at room temperature, the mixture was cooled to 0 °C and (biphenyl-4-yl)methylamine was added. The mixture was stirred for 72 h at room temperature before the solvent was removed under reduced pressure. The product was purified with silica-based column chromatography using a gradient of 0-100% Hex/EtOAc. Yield: 40 %. ^{1}H NMR (500 MHz, DMSO-d_6): δ 8.37 (t, J = 6.0 Hz, 1H), 7.62 (d, J = 8.1 Hz, 4H), 7.41 (t, J = 7.8 Hz, 4H), 7.31 (t, J = 7.4 Hz, 1H), 4.81 (t, J = 1.8 Hz, 2H), 4.39 (d, J = 6.0 Hz, 2H), 4.34 – 4.32 (m, 2H), 4.11 (s, 5H). m/z 396.24 [M+H]$^+$.

[Fe(η^5-C$_5$H$_4$COC$_6$H$_5$)$_2$] (A20). To a stirred solution of ferrocene in DCM (5 mL) at 0 °C was added AlCl$_3$ in portions. A solution of benzoyl chloride in dichloromethane (5 mL) was added dropwise over 1 min while keeping the temperature at 0 °C. After addition, the reaction mixture was warmed to room temperature and stirred for 1 h. The reaction mixture was concentrated under reduced pressure and purified with silica-based column chromatography using a gradient of 0-100% Hex/EtOAc. Yield: 26%. ^{1}H NMR (500 MHz, DMSO-d_6): δ 7.69 - 7.81 (m, 4H), 7.56 -

7.64 (m, 2H), 7.42 - 7.54 (m, 4H), 4.78 - 4.94 (m, 4H), 4.62 - 4.77 (m, 4H). ESI-TOFMS(+): *m/z* 395.0727 [M+H]$^+$.

[Fe(η^5-C$_5$H$_4$CH$_2$OH)$_2$] (1,1'-ferrocenedimethanol; A21). Compound A21 was purchased from Sigma Aldrich and used without further purification.

[Fe(η^5-C$_5$H$_4$COOH)$_2$] (1,1'-ferrocenedicarboxylic acid; A22). Compound A22 was purchased from Sigma Aldrich and used without further purification.

Class B compounds
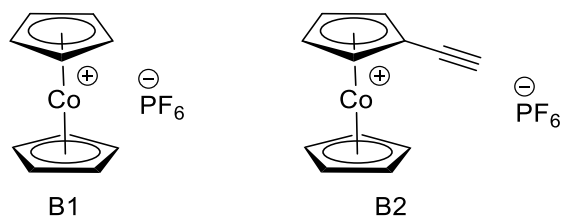
Figure 5.S2. Cobaltocene scaffolds that comprise the Class **B** mFs.

[Co(η⁵-C₅H₅)₂]⁺[PF₆]⁻ (cobaltocenium hexafluorophosphate; B1). Compound B1 was purchased from Sigma Aldrich and used without further purification.

[Co(η⁵-C₅H₅)(η⁵-C₅H₄CCH)]⁺[PF₆]⁻ (B2). Compound B2 was synthesized following a previously reported procedure.[60] ^{1}H NMR (500 MHz, DMSO-d_6): δ 6.19 (t, J = 2.1 Hz, 2H), 5.91 (t, J = 2.0 Hz, 2H), 5.88 (s, 5H), 4.71 (s, 1H). ESI-MS(+): m/z 212.7 [M+H]⁺.

Class C compounds
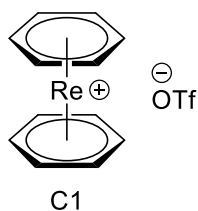
Figure 5.S3. The rhenium sandwhich complex that comprises Class **C**.

[Re(η⁶-C₆H₆)₂]⁺[OTf]⁻ (Bis(arene)rhenium triflate; C1). Compound C1 was synthesized following a previously reported procedure.[19] ^{1}H NMR (250 MHz, DMSO-d_6): δ 6.04 (s, 12H). ESI-MS(+): m/z 342.5 [M+H]$^+$.

Class D compounds
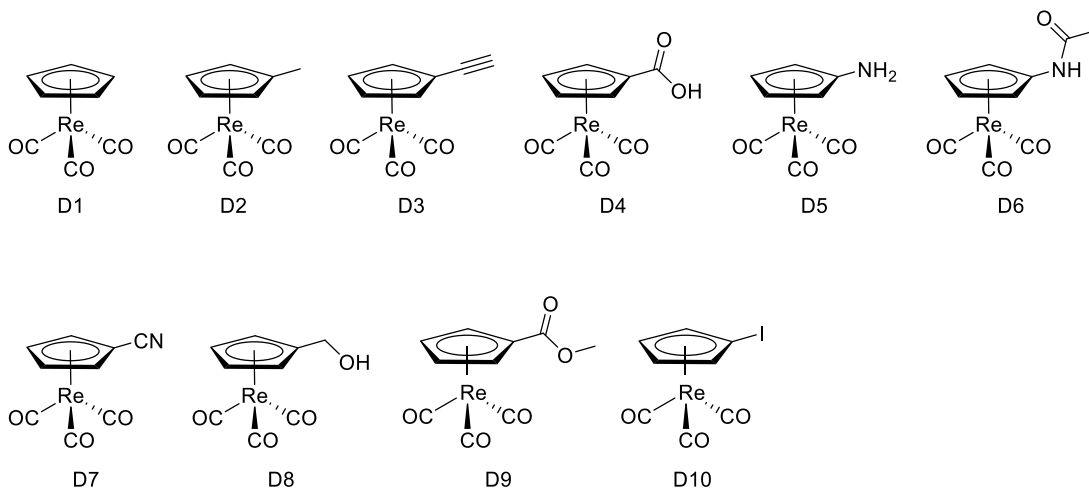
Figure 5.S4. Rhenium piano-stool scaffolds that comprise the Class **D** mFs.

Note: Mass spectra could not be obtained for most Class D compounds.

[Re(η⁵-C₅H₅)(CO)₃] (D1). Compound D1 was purchased from commercial sources and used without further purification.

[Re(η⁵-C₅H₄CH₃)(CO)₃] (D2). Compound D2 was synthesized following a previously reported procedure.[61] ^{1}H NMR (200 MHz, Chloroform-d_1): δ 5.23 (s, 4H), 2.23 (s, 3H).

[Re(η⁵-C₅H₄CCH)(CO)₃] (D3). To make the Re(Cp-acetylene-TMS)(CO)₃ precursor, compound D10 (121 mg, 0.26 mmol) and TMS-acetylene (0.05 mL, 0.38 mmol) were dissolved in 3.5 mL of a mixture of NEt₃/DMF/THF (0.5:1:2). The mixture was degassed by three freeze-pump-thaw cycles before Pd(MeCN)₂Cl₂ (5 mol%) and CuI (10 mol%) were added. The reaction mixture stirred overnight at room temperature. The solvent was evaporated under reduced pressure, and Re(Cp-acetylene-TMS)(CO)₃ was purified with silica-based column chromatography using PE/EtOAc (20:1) as eluent. Yield: 84 mg (76%). To make compound D3, Re(Cp-acetylene-TMS)(CO)₃ (81 mg, 0.19 mmol) and K_2CO_3 (29 mg, 0.21 mmol) were diluted with 2 mL MeOH, and the white suspension stirred for 3 h at room temperature. After the addition of 4 mL H_2O, the product was extracted with Et_2O. Yield: 61 mg (90%). ¹H NMR (300 MHz, Chloroform-d_1): δ 5.66 (t, J = 2.2 Hz, 2H), 5.28 (t, J = 2.2 Hz, 2H), 2.82 (s, 1H).

[Re(η⁵-C₅H₄COOH)(CO)₃] (D4). Compound D1 (110 mg, 0.33 mmol) was dissolved in 6 mL dry THF in a N₂-flooded Schlenk flask, and the solution was cooled to -78 °C. After the slow addition of n-BuLi (1.6 M in n-hexane; 0.31 mL), the reaction mixture was kept cooled and stirred for 1 h. CO_2 was bubbled through the mixture while it was allowed to warm up very slowly over several hours. The reaction was quenched with water, and the aqueous layer was washed with Et_2O. The pH of the aqueous layer was adjusted to 1 with 1 M HCl, and the product was extracted with Et_2O. Yield: 41 mg (33%). ¹H NMR (500 MHz, DMSO-d_6): δ 6.22 (t, J = 1.9 Hz, 2H), 5.75 (t, J = 1.9 Hz, 2H). EI-MS(+): *m/z* 379.9 [M].

[Re(η⁵-C₅H₄NH₂)(CO)₃] (D5). Compound D5 was synthesized following a previously reported procedure.[61] After extraction, the product was purified with silica-based column chromatography

using PE/EtOAc (10:1) as the eluent. Solvent was evaporated from the product under reduced pressure, and the product was recrystallized by slow evaporation of a mixture of DCM/n-hexane. ^{1}H NMR (200 MHz, Chloroform-d_1): δ 5.05 (t, J = 2.2 Hz, 2H), 4.92 (t, J = 2.2 Hz, 2H), 3.28 (s, 2H). EI-MS(+): m/z 350.9 [M].

Re(η^5-C$_5$H$_4$NCOCH$_3$)(CO)$_3$ (D6). Compound D6 was synthesized following a previously reported procedure.[62] ^{1}H NMR (200 MHz, Chloroform-d_1): δ 6.90 (s, 1H), 5.69 (q, J = 2.0 Hz, 2H), 5.19 (q, J = 2.0, 2H), 2.07 (s, 3H). ESI-TOFMS(+): m/z 394.0081 [M+H]$^+$.

[Re(η^5-C$_5$H$_4$CN)(CO)$_3$] (D7). Re(η^5-C$_5$H$_4$COH)(CO)$_3$ was synthesized following a previously reported procedure with some modification.[27] To make the Re(η^5-C$_5$H$_4$COH)(CO)$_3$ precursor, compound D1 (200 mg, 0.60 mmol) was dissolved in 12 mL dry THF in a N$_2$-flooded Schlenk flask, and the solution was cooled to -78 °C. After the slow addition of n-BuLi (1.6 M in n-hexane, 0.56 mL), the reaction mixture was kept cool and stirred for 35 min. Dry DMF (0.11 mL, 1.49 mmol) was added, and the mixture was stirred first for 1.5 h at -78 °C and then for 1 h at -5 °C. The reaction was quenched with saturated aqueous NH$_4$Cl (3 mL). The aqueous phase was extracted with Et$_2$O, and the combined organic layers were dried over MgSO$_4$ and filtered. The solvent was evaporated under reduced pressure. Purification of Re(η^5-C$_5$H$_4$COH)(CO)$_3$ was performed with silica-based column chromatography using DCM/PE (1:2) as eluent. Yield: 169 mg (78%). Compound D7 was synthesized following a previously reported procedure with some modification.[63] To Re(η^5-C$_5$H$_4$COH)(CO)$_3$ (169 mg, 0.47 mmol) and NH$_2$OH·HCl (48 mg, 0.70 mmol) was added 6 mL of formic acid. The mixture was stirred for 30 min at 80 °C. After cooling to ambient temperature, it was poured onto ice. The aqueous layer was extracted with

Et$_2$O, and the combined organic layers were washed with H$_2$O and brine, dried over MgSO$_4$, and filtered. The solvent was evaporated under reduced pressure. The compound was dissolved in DCM, and the organic layer was washed with saturated aqueous NaHCO$_3$, dried over MgSO$_4$, and filtered. The solvent was evaporated under reduced pressure, and the product was purified with silica-based column chromatography using DCM/n-hexane (1:1) as the eluent. Yield: 72 mg (41%). ^{1}H NMR (200 MHz, Chloroform-d_1): δ 5.91 (t, J = 2.3 Hz, 2H), 5.42 (t, J = 2.3 Hz, 2H).

[Re(η^5-C$_5$H$_4$CH$_2$OH)(CO)$_3$] (D8). Re(η^5-C$_5$H$_4$COH)(CO)$_3$ was synthesized following a previously reported procedure with some modification, as described for D7.[27] To make compound D8, Re(η^5-C$_5$H$_4$COH)(CO)$_3$ (130 mg, 0.36 mmol) was dissolved in 6 mL dry Et$_2$O in a N$_2$-flooded Schlenk flask at 0 °C, and LiAlH$_4$ (7 mg, 0.18 mmol) was added. The reaction mixture stirred for 1 h at 0 °C and 1.5 h at room temperature. The reaction was quenched with saturated aqueous NH$_4$Cl (3 mL), and the aqueous layer was extracted with Et$_2$O. The combined organic layers were washed with H$_2$O and brine, dried over MgSO$_4$, and filtered. The solvent was evaporated under reduced pressure, and the product was purified with silica-based column chromatography using DCM as the eluent. Yield: 86 mg (66%). ^{1}H NMR (200 MHz, Chloroform-d_1): δ 5.31 (t, J = 2.2 Hz, 2H), 5.45 (t, J = 2.2 Hz, 2H), 4.44 (d, J = 5.2 Hz, 2H).

[Re(η^5-C$_5$H$_4$COOCH$_3$)(CO)$_3$] (D9). Compound D1 (100 mg, 0.30 mmol) was dissolved in 12 mL dry THF in a N$_2$-flooded Schlenk flask, and the solution was cooled to -78 °C. After the slow addition of n-BuLi (1.6 M in n-hexane, 0.22 mL), the reaction mixture was kept cooled and stirred for 30 min. Methyl chloroformate (0.03 mL, 0.43 mmol) was added, and the mixture stirred first for 1.5 h at -78 °C, then for 1 h at -5 °C. The reaction was quenched with saturated aqueous

NH$_4$Cl (3 mL). The aqueous phase was extracted with Et$_2$O, and the combined organic layers were dried over MgSO$_4$ and filtered. The solvent was evaporated under reduced pressure and purification of the product was done with silica-based column chromatography using DCM/n-hexane (1:3) as the eluent. Yield: 114 mg (98%). ^{1}H NMR (200 MHz, Chloroform-d_1): δ 6.10 – 5.91 (m, 2H), 5.43 – 5.30 (m, 2H), 3.81 (s, 3H).

[Re(η^5-C$_5$H$_4$I)(CO)$_3$] (D10). ReCp(CO)$_3$ (125 mg, 0.37 mmol) was dissolved in 10 mL dry THF in a N$_2$-flooded Schlenk flask, and the solution was cooled to -78 °C. After the addition of n-BuLi (1.6 M in n-hexane; 0.35 mL), the reaction mixture was kept cooled and stirred for 1 h. I$_2$ (142 mg, 0.56 mmol) was dissolved in 3 mL dry THF and slowly added to the reaction mixture. The mixture was allowed to warm gradually to reach room temperature and then continued to stir overnight. The reaction was quenched with 3 mL of a saturated aqueous NH$_4$Cl solution, extracted with Et$_2$O, and washed with brine. The combined organic layers were dried over MgSO$_4$ and filtered, and the solvent was evaporated under reduced pressure. Purification was performed with silica-based column chromatography using n-hexane/EtOAc (20:1) as the eluent. Yield: 121 mg (70%). ^{1}H NMR (200 MHz, Chloroform-d_1): δ 5.61 (t, J = 2.2 Hz, 2H), 5.27 (t, J = 2.2 Hz, 2H).

Class E compounds

E1
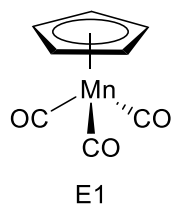
Figure 5.S5. The manganese tricarbonyl complex is the Class **E** mF.

[Mn(η⁵-C₅H₅)(CO)₃] (E1). Compound E1 was purchased from commercial sources and used without further purification

Class F compounds

F1   F2   F6   F7
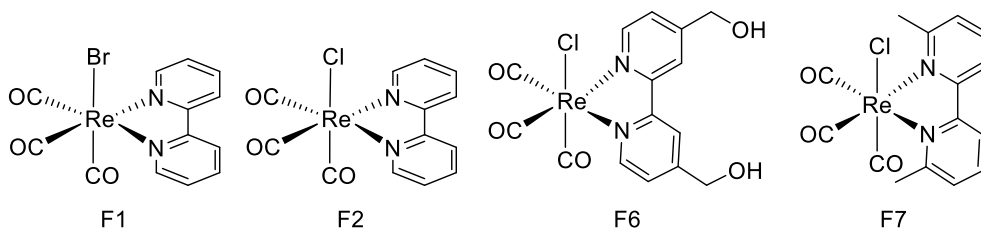
Figure 5.S6. Rhenium tricarbonyl scaffolds that comprise the Class **F** mFs.

[Re(2,2'-bipyridine)Br(CO)₃] (F1). Re(CO)₅Br (50 mg, 0.14 mmol) and 2,2'-bipyridine (27 mg, 0.15 mmol) were suspended in 5 mL toluene (stored over molecular sieves). The suspension was saturated with N_2 and stirred overnight at 60 °C before the solvent was evaporated under reduced pressure. The residue was dissolved in DCM, and MeOH was added to the solution. The DCM

was slowly evaporated under reduced pressure until the product started to precipitate. The yellow crystalline product was collected via filtration. Yield: 66 mg (97%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 9.04 (d, J = 5.3 Hz, 2H), 8.77 (d, J = 8.1 Hz, 2H), 8.33 (td, J = 8.0 Hz, 1.6 Hz, 2H), 7.83 – 7.70 (m, 2H). ESI-TOFMS(+): m/z 523.9589 [M+NH$_4$]$^+$, 528.9142 [M+Na]$^+$.

[Re(2,2'-bipyridine)Cl(CO)$_3$] (F2). Re(CO)$_5$Cl (50 mg, 0.14 mmol) and 2,2'-bipyridine (27 mg, 0.15 mmol) were suspended in 5 mL toluene (stored over molecular sieves). The suspension was saturated with N$_2$ and stirred overnight at 60 °C before the solvent was evaporated under reduced pressure. The residue was dissolved in DCM, and MeOH was added to the solution. The DCM was slowly evaporated under reduced pressure until the product started to precipitate. The yellow crystalline product was collected via filtration. Yield: 66 mg (97%). ^{1}H NMR (500 MHz, Chloroform-d_1): δ 9.02 (d, J = 5.2 Hz, 2H), 8.77 (d, J = 8.1 Hz, 2H), 8.34 (td, J = 7.9 Hz, 1.3 Hz, 2H), 7.76 (m, 2H). ESI-TOFMS(+): m/z 427.0084 [M-Cl]$^+$.

[Re([2,2'-bipyridine]-4,4'-diyldimethanol)Cl(CO)$_3$] (F6). Re(CO)$_5$Cl (50 mg, 0.14 mmol) and [2,2'-bipyridine]-4,4'-diyldimethanol (31 mg, 0.15 mmol) were suspended in 5 mL toluene (stored over molecular sieves). The suspension was saturated with N$_2$ and stirred overnight at 60 °C before the solvent was evaporated under reduced pressure. The resulting residue was suspended in a mixture of DCM and MeOH and was sonicated for 0.5 h. The yellow powdery product was collected via filtration. Yield: 48 mg (67%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 8.94 (d, J = 5.7 Hz, 2H), 8.61 (d, J = 1.6 Hz, 2H), 7.69 (dd, J = 5.7, 1.6 Hz, 2H), 5.81 (t, J = 5.7 Hz, 2H), 4.77 (d, J = 5.7 Hz, 4H). ESI-MS(+): m/z 544.93 [M+Na]$^+$. ESI-MS(-): m/z 520.95 [M-H]$^-$.

[Re(6,6'-dimethyl-2,2'-bipyridine)Cl(CO)₃] (F7). Re(CO)₅Cl (50 mg, 0.14 mmol) and 6,6'-dimethyl-2,2'-bipyridine (27 mg, 0.15 mmol) were suspended in 5 mL toluene (stored over molecular sieves). The suspension was saturated with N_2 and stirred overnight at 60 °C before the solvent was evaporated under reduced pressure. The residue was dissolved in DCM, and MeOH was added to the solution. The DCM was slowly evaporated under reduced pressure until the product started to precipitate. The yellow crystalline product was collected via filtration. Yield: 66 mg (97%). ^{1}H NMR (500 MHz, Chloroform-d_1): δ 7.99 (dd, $J = 7.7$, 1.0 Hz, 2H), 7.89 (t, $J = 7.9$ Hz, 2H), 7.45 (dd, $J = 7.7$, 1.2 Hz, 2H), 3.13 (s, 6H). ESI-MS(+): m/z 513.00 [M+Na]$^+$, 455.17 [M-Cl]$^+$.

Class G compounds
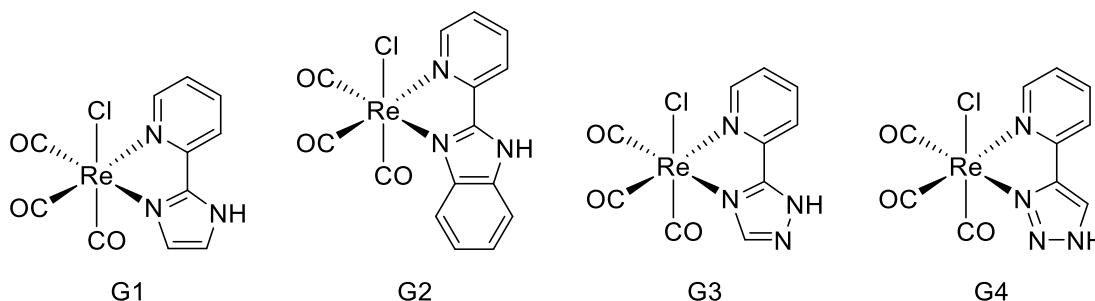
Figure 5.S7. Rhenium tricarbonyl scaffolds that comprise the Class **G** mFs.

[Re(2-(1*H*-imidazol-2-yl)pyridine)Cl(CO)₃] (G1). Re(CO)₅Cl (50 mg, 0.14 mmol) and 2-(1*H*-imidazol-2-yl)pyridine (21 mg, 0.15 mmol) were suspended in 5 mL toluene (stored over molecular sieves). The suspension was saturated with N_2 and stirred overnight at 60 °C. The solvent was evaporated under reduced pressure, and the yellow residue was washed with pentane and dried in vacuo. Recrystallization from hot acetone gave the product as yellow crystals. Yield: 61 mg (98%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 8.90 (d, $J = 5.4$ Hz, 1H), 8.34 – 8.21 (m, 2H),

7.69 (d, $J = 1.3$ Hz, 1H), 7.62 (ddd, $J = 7.3, 5.4, 1.8$ Hz, 1H), 7.50 (d, $J = 1.4$ Hz, 1H). ESI-MS(+): m/z 416.04 [M-Cl]$^-$, 433.66 [M-Cl+NH$_3$].

[Re(2-(1*H*-benzimidazol-2-yl)pyridine)Cl(CO)$_3$] (G2). Re(CO)$_5$Cl (50 mg, 0.14 mmol) and 2-(1*H*-benzimidazol-2-yl)pyridine (28 mg, 0.15 mmol) were suspended in 5 mL toluene (stored over molecular sieves). The suspension was saturated with N$_2$ and stirred overnight at 60 °C. The solvent was evaporated under reduced pressure, and the yellow residue was washed with pentane and dried in vacuo. Recrystallization from hot acetone gave the product as yellow crystals. Yield: 62 mg (90%). ^{1}H NMR (500 MHz, Methanol-d_4): δ 9.10 (dt, $J = 5.5, 1.2$ Hz, 1H), 8.41 – 8.34 (m, 1H), 8.30 (td, $J = 7.8, 1.5$ Hz, 1H), 7.97 – 7.92 (m, 1H), 7.76 – 7.68 (m, 2H), 7.57 – 7.50 (m, 2H). ESI-MS(+): m/z 466.04 [M-Cl]$^-$, 433.66 [M-Cl+NH$_3$].

[Re(2-(1*H*-1,2,4-triazol-5-yl)pyridine)Cl(CO)$_3$] (G3). Re(CO)$_5$Cl (50 mg, 0.14 mmol) and 2-(1*H*-1,2,4-triazol-5-yl)pyridine (21 mg, 0.15 mmol) were suspended in 5 mL toluene (stored over molecular sieves). The suspension was saturated with N$_2$ and stirred overnight at 60 °C. The solvent was evaporated under reduced pressure, and the residue was suspended in acetone. The suspension was filtered, and hexane was added slowly to the filtrate. The resulting precipitate was collected via filtration and was re-dissolved in a mixture of acetone and H$_2$O. The acetone was slowly evaporated under reduced pressure until a white solid started to precipitate. The precipitate was collected via filtration and evaporated to dryness. The product was recrystallized from acetone. Yield: 15 mg (24%). ^{1}H NMR (500 MHz, Acetone-d_6): δ 9.47 (s, 1H), 9.07 (ddd, $J = 5.5, 1.5, 0.9$ Hz, 1H), 8.39 (dt, $J = 8.0, 1.2$ Hz, 1H), 8.34 (td, $J = 7.7, 1.5$ Hz, 1H), 7.84 – 7.76 (m,

1H). ESI-MS(+): *m/z* 417.14 [M-Cl]$^+$ 434.70 [M-Cl+H$_2$O]$^+$. To the best of our knowledge, compound **G3** has not been reported previously in the literature

[Re(2-(1*H*-1,2,3-triazol-5-yl)pyridine)Cl(CO)$_3$] (G4). Re(CO)$_5$Cl (50 mg, 0.14 mmol) and 2-(1*H*-1,2,3-triazol-5-yl)pyridine (21 mg, 0.15 mmol) were suspended in 3 mL acetone. The suspension was heated to 60 °C overnight. The solvent was evaporated under reduced pressure, and the residue was purified with silica-based column chromatography using DCM/MeOH as eluent (3 min 0% MeOH, 22 min 0-10.1% MeOH, 3.4 min 10.1% MeOH). Yield: 18 mg (29%). ^{1}H NMR (500 MHz, Acetone-*d$_6$*): δ 9.13 (s, 1H), 9.06 (dt, *J* = 5.6, 1.1 Hz, 1H), 8.33 (dt, *J* = 7.9, 1.2 Hz, 1H), 8.27 (td, *J* = 7.7, 1.4 Hz, 1H), 7.67 (ddd, *J* = 7.4, 5.6, 1.4 Hz, 1H). ESI-MS(+): *m/z* 417.02 [M-Cl]$^+$ 434.69 [M-Cl+H$_2$O]$^+$. ESI-MS(-): *m/z* 451.03 [M-H]$^-$. To the best of our knowledge, compound **G4** has not been reported previously in the literature.

Class H compounds
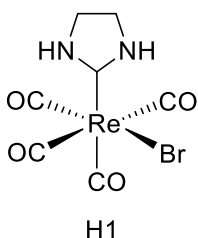
Figure 5.S8. Rhenium tricarbonyl scaffold that is the Class **H** mF.

[Re(imidazolidine)Br(CO)$_3$] (H1). Compound H1 was synthesized following a previously reported procedure.[64] ^{1}H NMR (200 MHz, Chloroform-*d$_1$*): δ 6.87 (bs, 2H), 3.78 (s, 4H). ESI-TOFMS(+): *m/z* 465.9398 [M+NH$_4$]$^+$, 470.8937 [M+Na]$^+$.

Class I compounds
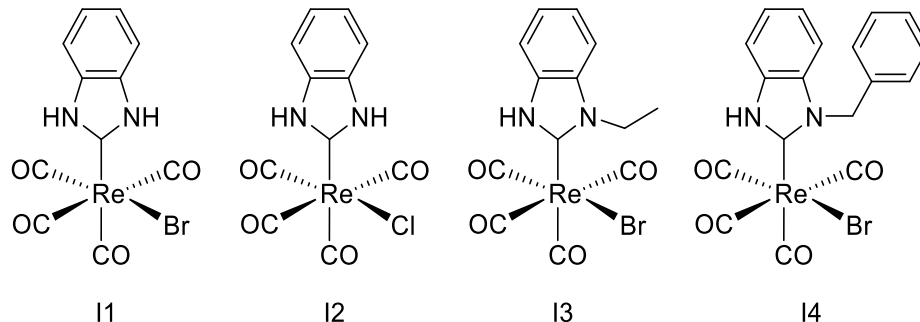
Figure 5.S9. Rhenium tricarbonyl scaffolds that comprise the Class **I** mFs.

[Re(2,3-dihydro-1*H*-benzo[d]imidazole)Br(CO)₃] (I1). To make the 2-((triphenyl-λ^5-phosphaneylidene)amino)aniline ligand, 2-azidoaniline (224 mg, 1.67 mmol, 1 eq.) was dissolved in toluene in a nitrogen atmosphere. After 10 minutes the reaction mixture was treated with a solution of triphenylphosphine (440 mg, 1.7 mmol, 1.02 eq.) in toluene. The solution was stirred for 20 h at room temperature before the solvent was removed under reduced pressure, and the crude foamy product was purified by silica-based column chromatography using hexanes/EtOAc (4:1 to 0:1) as the eluent. The ligand was obtained as an orange crystalline solid and dried under reduced pressure. Yield: 550 mg (89%). ^{1}H NMR (250 MHz, Chloroform-d_1): δ 7.82 – 7.72 (m, 6H), 7.56 – 7.42 (m, 9H), 6.75 – 6.70 (m, 1H), 6.60 – 6.52 (m, 1H), 6.41 – 6.31 (m, 2H), 4.36 (bs, 2H). ESI-MS(+): *m/z* 369.0 [M+H]$^+$.

The following synthetic procedure for compound I1 was originally published elsewhere.[26] Under a nitrogen atmosphere, bromopentacarbonylrhenium(I) (220 mg, 0.54 mmol, 1 eq.) and 2-((triphenyl-λ^5-phosphaneylidene)amino)aniline (200 mg, 0.54 mmol, 1 eq.) were dissolved in toluene and stirred for 24 h at ambient temperature. The solvent was removed under reduced

188

pressure, and the crude product was purified by silica-based column chromatography using hexanes/EtOAc (9:1) as the eluent. The product was obtained as an off-white solid and dried under reduced pressure. Yield: 230 mg (86%). ^{1}H NMR (200 MHz, DCM-d_2): δ 10.54 (bs, 2H), 7.59 – 7.53 (m, 2H), 7.42 – 7.35 (m, 2H). EI-MS(+): m/z 495.8 [M]$^+$.

[Re(2,3-dihydro-1*H*-benzo[d]imidazole)Cl(CO)₃] (I2). Under a nitrogen atmosphere, pentacarbonylchlororhenium(I) (300 mg, 0.83 mmol, 1 eq.) and 2-((triphenyl-λ^5-phosphaneylidene)amino)aniline (305 mg, 0.83 mmol, 1 eq.) were dissolved in toluene and stirred for 18 h at ambient temperature. The solvent was removed under reduced pressure, and the crude product was purified by silica-based column chromatography using PE:EtOAc (4:1) as the eluent. The desired product was obtained as a beige solid and dried under reduced pressure. Yield: 282 mg (75%). ^{1}H NMR (200 MHz, Chloroform-d_1): δ 10.58 (bs, 2H), 7.58 – 7.50 (m, 2H), 7.42 – 7.34 (m, 2H). ESI-TOFMS(+): m/z 469.9899 [M+NH₄]$^+$, 474.9452 [M+Na]$^+$.

[Re(1-ethyl-2,3-dihydro-1*H*-benzo[d]imidazole)Br(CO)₃] (I3). To make the *N*-ethyl-2-((triphenyl-λ^5-phosphaneylidene)amino)aniline ligand, 2-azido-N-ethylaniline (166 mg, 1 mmol, 1 eq.) was dissolved in toluene in a nitrogen atmosphere. The reaction mixture was treated with a solution of triphenylphosphine (268 mg, 1 mmol, 1 eq.) in toluene. The solution was stirred for 20 h at room temperature. Subsequently, the solvent was removed under reduced pressure, and the crude product was purified by silica-based column chromatography using DCM as the eluent. The product was obtained as an orange crystalline solid. Yield: 190 mg (48%). ^{1}H NMR (200 MHz, Chloroform-d_1): δ 7.81-7.72 (m, 6H), 7.56-7.41 (m, 9H), 6.70-6.64 (m, 1H), 6.59-6.54 (m,

1H), 6.42-6.38 (m, 1H), 6.33-6.27 (m, 1H), 4.99 (bs, 1H), 3.24 (q, J = 7.1 Hz, 2H), 1.34 (t, J = 7.1 Hz, 3H). ESI-MS(+): m/z 397.0 [M+H]$^+$.

The following synthetic procedure for compound I3 was originally published elsewhere.[26] Under a nitrogen atmosphere, N-ethyl-2-((triphenyl-λ^5-phosphaneylidene)amino)aniline (160 mg, 0.4 mmol, 1 eq.) was dissolved in dry toluene (8 ml) and treated with rheniumpentacarbonylbromide (166 mg, 0.4 mmol, 1 eq.). The orange solution was stirred for 18 h at room temperature. Subsequently, the solvent was removed under reduced pressure, and the beige crude product was purified by silica-based column chromatography using hexanes:EtOAc (4:1) as the eluent. The product was obtained as a beige crystalline solid. Yield: 127 mg (60%). ^{1}H NMR (200 MHz, Chloroform-d_1): δ 11.19 (bs, 1H), 7.55 − 7.47 (m, 2H), 7.41 − 7.35 (m, 2H), 4.52 (q, J = 7.2 Hz, 2H), 1.56 (t, J = 7.2 Hz, 3H). ESI-TOFMS(+): m/z 445.0199 [M-HBr]$^+$, 541.9690 [M+NH$_4$]$^+$.

[Re(1-benzyl-2,3-dihydro-1H-benzo[d]imidazole)Br(CO)$_3$] (I4). To make the N-benzyl-2-((triphenyl-λ^5-phosphaneylidene)amino)aniline ligand, 2-azido-N-benzylaniline (200 mg, 0.9 mmol, 1 eq.) was dissolved in toluene under a nitrogen atmosphere and treated with a solution of triphenylphosphine (239 mg, 0.91 mmol, 1.05 eq.) in toluene. The solution was stirred for 24 h at room temperature. Subsequently, the solvent was removed under reduced pressure, and the crude product was purified by silica-based column chromatography using hexanes/EtOAc (4:1) as the eluent. The product was obtained as an orange solid and dried under reduced pressure. Yield: 310 mg (75%). ^{1}H NMR (250 MHz, Chloroform-d_1): δ 7.91-7.80 (m, 6H), 7.65-7.48 (m, 11H), 7.45-7.28 (m, 3H), 6.67-6.57 (m, 2H), 6.50-6.44 (m, 1H), 6.37-6.29 (m, 1H), 5.50 (bs, 1H), 4.54 (s, 2H). ESI-MS(+): m/z 458.9 [M+H]$^+$.

The following synthetic procedure for compound I4 was originally published elsewhere.[26] Under a nitrogen atmosphere, bromopentacarbonylrhenium(I) (112 mg, 0.28 mmol, 1 eq.) and *N*-benzyl-2-((triphenyl-λ^5-phosphaneylidene)amino)aniline (126 mg, 0.28 mmol, 1 eq.) were dissolved in THF and stirred for 24 h at room temperature. The solvent was removed in vacuo and the crude product was purified by silica-based column chromatography using hexanes/EtOAc (4:1) as the eluent. The product was obtained as a grey solid and dried under reduced pressure. Yield: 137 mg (85%). ^{1}H NMR (200 MHz, Chloroform-d_1): δ 11.35 (s, 1H), 7.54 – 7.50 (m, 1H), 7.36 – 7.21 (m, 6H), 7.04 – 7.00 (m, 2H), 5.67 (s, 2H).

Class J compounds
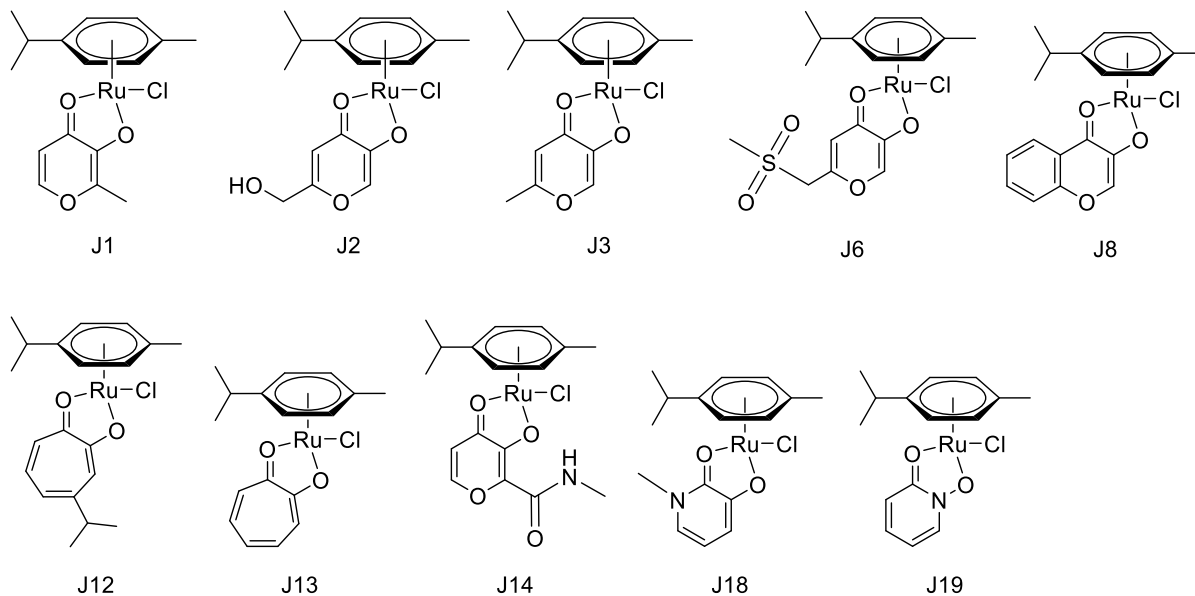
Figure 5.S10. Ruthenium piano-stool scaffolds that comprise the Class **J** mFs.

[Ru(η^6-*p*-cymene)Cl(maltol)] (J1). A solution of 0.2 M sodium methoxide was made by dissolving sodium (230 mg, 10.0 mmol) into methanol (50 mL). The reaction was kept as dry as possible, and the sodium methoxide solution was kept under a nitrogen atmosphere.

Bis[dichlorido(η^6-p-cymene)Ru(II)] (0.9 eq.) was added to a solution of maltol (1 eq.) and sodium methoxide (1.1 eq.) in methanol (80 mL). The reaction mixture was stirred at room temperature under a nitrogen atmosphere for 4–8 h. Afterwards the solvent was evaporated under reduced pressure, and the residue was dissolved in dichloromethane and filtered to remove NaCl. Solvent was evaporated from the filtrate under reduced pressure, and the resulting solid was purified with silica-based column chromatography using DCM/MeOH as eluent (gradient of 0-10% MeOH). Yield: 56.0 mg (78%). ^{1}H NMR (500 MHz, Chloroform-d_1): δ 7.55 (d, J = 5.1 Hz, 1H), 6.50 (d, J = 5.1 Hz), 5.53 – 5.49 (m, 2H), 5.32 – 5.27 (m, 2H), 2.91 – 2.88 (m, 1H), 2.41 (s, 3H), 2.32 (s, 3H), 1.34 – 1.29 (m, 6H). ESI-MS(+): m/z = 361.23 [M-Cl]$^+$.

[Ru(η^6-p-cymene)Cl(6-(hydroxymethyl)-4-oxo-4H-pyran-3-olate)] (J2). A solution of 0.2 M sodium methoxide was made by dissolving sodium (230 mg, 10.0 mmol) into methanol (50 mL). The reaction was kept as dry as possible, and the sodium methoxide solution was kept under a nitrogen atmosphere. Bis[dichlorido(η^6-p-cymene)Ru(II)] (0.9 eq.) was added to a solution of 5-hydroxy-2-(hydroxymethyl)-4H-pyran-4-one (1 eq.) and sodium methoxide (1.1 eq.) in methanol (80 mL). The reaction mixture was stirred at room temperature under a nitrogen atmosphere for 4–8 h. Afterwards the solvent was evaporated under reduced pressure, and the residue was dissolved in dichloromethane and filtered to remove NaCl. Solvent was evaporated from the filtrate under reduced pressure, and the resulting solid was purified with silica-based column chromatography using DCM/MeOH as eluent (gradient of 0-10% MeOH). Yield: 47.8 mg (64%). ^{1}H NMR (500 MHz, Chloroform-d_1): δ 7.68 (s, 1H), 6.63 (s, 1H), 5.54 – 5.51 (m, 2H), 5.31 (s, 2H), 4.44 (s, 2H), 2.92 – 2.90 (m, 1H), 2.70 (s, 1H), 2.30 (s, 3H), 1.33 – 1.29 (m, 6H). ESI-MS(+): m/z = 377.23 [M-Cl]$^+$.

[Ru(η⁶-*p*-cymene)Cl(allomaltol)] (J3). A solution of 0.2 M sodium methoxide was made by dissolving sodium (230 mg, 10.0 mmol) into methanol (50 mL). The reaction was kept as dry as possible, and the sodium methoxide solution was kept under a nitrogen atmosphere. Bis[dichlorido(η⁶-*p*-cymene)Ru(II)] (0.9 eq.) was added to a solution of allomaltol (1 eq.) and sodium methoxide (1.1 eq.) in methanol (80 mL). The reaction mixture was stirred at room temperature under a nitrogen atmosphere for 4–8 h. Afterwards the solvent was evaporated under reduced pressure, and the residue was dissolved in dichloromethane and filtered to remove NaCl. Solvent was evaporated from the filtrate under reduced pressure, and the resulting solid was purified with silica-based column chromatography using DCM/MeOH as eluent (gradient of 0-10% MeOH). Yield: 44.3 mg (62%). ^{1}H NMR (500 MHz, Chloroform-d_1): δ 7.62 (s, 1H), 6.30 (s, 1H), 5.49 (t, J = 6.6 Hz, 2H), 5.28 (d, J = 5.9 Hz, 2H), 2.91 – 2.85 (m, 1H), 2.27 (s, 3H), 2.21 (s, 3H), 1.28 (dd, J = 6.6, 4.5 Hz, 6H). ESI-MS(+): m/z = 361.25 [M-Cl]$^+$.

[Ru(η⁶-*p*-cymene)Cl(6-(methylsulfonyl)-4-oxo-4H-pyran-3-olate)] (J6). A solution of 0.2 M sodium methoxide was made by dissolving sodium (230 mg, 10.0 mmol) into methanol (50 mL). The reaction was kept as dry as possible, and the sodium methoxide solution was kept under a nitrogen atmosphere. Bis[dichlorido(η⁶-*p*-cymene)Ru(II)] (0.9 eq.) was added to a solution of 5-hydroxy-2-(methylsulfonyl)-4*H*-pyran-4-one (1 eq.) and sodium methoxide (1.1 eq.) in methanol (80 mL). The reaction mixture was stirred at room temperature under a nitrogen atmosphere for 4–8 h. Afterwards the solvent was evaporated under reduced pressure, and the residue was dissolved in dichloromethane and filtered to remove NaCl. Solvent was evaporated from the filtrate under reduced pressure, and the resulting solid was purified with silica-based column chromatography using DCM/MeOH as eluent (gradient of 0-10% MeOH). Yield: 18.4 mg (25%).

¹H NMR (500 MHz, Chloroform-d_1): δ 7.76 (s, 1H), 6.67 (s, 1H), 5.56 (t, J = 4.6 Hz, 2H), 5.34 (dd, J = 5.1, 3.9 Hz, 2H), 4.18 (d, J = 2.9 Hz, 2H), 2.96 (s, 2H), 2.31 (s, 3H), 1.35 – 1.33 (dd, J = 6.9, 0.8 Hz, 6H). ESI-MS(+): m/z = 439.00 [M-Cl]⁺. To the best of our knowledge, compound **J6** has not been reported previously in the literature

[Ru(η⁶-*p*-cymene)Cl(4-oxo-4*H*-chromen-3-olate)] (J8). A solution of 0.2 M sodium methoxide was made by dissolving sodium (230 mg, 10.0 mmol) into methanol (50 mL). The reaction was kept as dry as possible, and the sodium methoxide solution was kept under a nitrogen atmosphere. Bis[dichlorido(η⁶-*p*-cymene)Ru(II)] (0.9 eq.) was added to a solution of 3-hydroxy-4*H*-chromen-4-one (1 eq.) and sodium methoxide (1.1 eq.) in methanol (80 mL). The reaction mixture was stirred at room temperature under a nitrogen atmosphere for 4–8 h. Afterwards the solvent was evaporated under reduced pressure, and the residue was dissolved in dichloromethane and filtered to remove NaCl. Solvent was evaporated from the filtrate under reduced pressure, and the resulting solid was purified with silica-based column chromatography using DCM/MeOH as eluent (gradient of 0-10% MeOH). Yield: 18.8 mg (24%). ¹H NMR (500 MHz, Chloroform-d_1): δ 8.82 (dd, J = 8.2, 1.5 Hz, 1H), 7.93 (s, 1H), 7.60 (ddd, J = 8.6, 7.0, 1.6 Hz, 1H), 7.49 – 7.44 (m, 1H), 7.34 (ddd, J = 8.1, 7.1, 0.9 Hz, 1H), 5.65 – 5.59 (m, 2H), 5.40 (t, J = 4.8 Hz, 2H), 2.99 – 2.94 (m, 1H), 2.36 (s, 3H), 1.37 – 1.33 (m, 6H). ESI-MS(+): m/z = 397.21 [M-Cl]⁺. To the best of our knowledge, compound **J8** has not been reported previously in the literature

[Ru(η⁶-*p*-cymene)Cl(3-isopropyl-7-oxocyclohepta-1,3,5-trien-1-olate)] (J12). A solution of 0.2 M sodium methoxide was made by dissolving sodium (230 mg, 10.0 mmol) into methanol (50 mL). The reaction was kept as dry as possible, and the sodium methoxide solution was kept under

a nitrogen atmosphere. Bis[dichlorido(η^6-*p*-cymene)Ru(II)] (0.9 eq.) was added to a solution of 2-hydroxy-4-isopropylcyclohepta-2,4,6-trien-1-one (1 eq.) and sodium methoxide (1.1 eq.) in methanol (80 mL). The reaction mixture was stirred at room temperature under a nitrogen atmosphere for 4–8 h. Afterwards the solvent was evaporated under reduced pressure, and the residue was dissolved in dichloromethane and filtered to remove NaCl. Solvent was evaporated from the filtrate under reduced pressure, and the resulting solid was purified with silica-based column chromatography using DCM/MeOH as eluent (gradient of 0-10% MeOH). Yield: 66.4 mg (84%). ^{1}H NMR (500 MHz, Chloroform-d_1): δ 7.21 (d, J = 1.5 Hz, 1H), 7.11 – 7.09 (m, 2H), 6.73 – 6.66 (m, 1H), 5.53 (t, J = 6.7 Hz, 2H), 5.32 (t, J = 5.9 Hz, 2H), 2.90 (dt, J = 13.8, 6.9 Hz, 1H), 2.73 (dt, J = 13.7, 6.9 Hz, 1H), 2.32 (d, J = 2.7 Hz, 3H), 1.32 – 1.30 (m, 6H), 1.19 – 1.17 (m, 6H). ESI-MS(+): m/z = 399.19 [M-Cl]$^+$. To the best of our knowledge, compound **J12** has not been reported previously in the literature.

[Ru(η^6-*p*-cymene)Cl(7-oxocyclohepta-1,3,5-trien-1-olate)] (J13). A solution of 0.2 M sodium methoxide was made by dissolving sodium (230 mg, 10.0 mmol) into methanol (50 mL). The reaction was kept as dry as possible, and the sodium methoxide solution was kept under a nitrogen atmosphere. Bis[dichlorido(η^6-*p*-cymene)Ru(II)] (0.9 eq.) was added to a solution of 2-hydroxycyclohepta-2,4,6-trien-1-one (1 eq.) and sodium methoxide (1.1 eq.) in methanol (80 mL). The reaction mixture was stirred at room temperature under a nitrogen atmosphere for 4–8 h. Afterwards the solvent was evaporated under reduced pressure, and the residue was dissolved in dichloromethane and filtered to remove NaCl. Solvent was evaporated from the filtrate under reduced pressure, and the resulting solid was purified with silica-based column chromatography using DCM/MeOH as eluent (gradient of 0-10% MeOH). Yield: 53.0 mg (75%). ^{1}H NMR (500

MHz, Chloroform-d_1): δ 7.24 – 7.14 (m, 4H), 6.81 – 6.73 (m, 1H), 5.54 (d, J = 6.1 Hz, 2H), 5.33 (d, J = 6.1 Hz, 2H), 2.90 (dt, J = 13.9, 6.9 Hz, 1H), 2.32 (s, 3H), 1.31 (d, J =7.0 Hz, 6H). ESI-MS(+): m/z = 357.13 [M-Cl]$^+$.

[Ru(η^6-*p*-cymene)Cl(2-(methylcarbamoyl)-4-oxo-4*H*-pyran-3-olate)] (J14). A solution of 0.2 M sodium methoxide was made by dissolving sodium (230 mg, 10.0 mmol) into methanol (50 mL). The reaction was kept as dry as possible, and the sodium methoxide solution was kept under a nitrogen atmosphere. Bis[dichlorido(η^6-*p*-cymene)Ru(II)] (0.9 eq.) was added to a solution of 3-hydroxy-N-methyl-4-oxo-4*H*-pyran-2-carboxamide (1 eq.) and sodium methoxide (1.1 eq.) in methanol (80 mL). The reaction mixture was stirred at room temperature under a nitrogen atmosphere for 4–8 h. Afterwards the solvent was evaporated under reduced pressure, and the residue was dissolved in dichloromethane and filtered to remove NaCl. Solvent was evaporated from the filtrate under reduced pressure, and the resulting solid was purified with silica-based column chromatography using DCM/MeOH as eluent (gradient of 0-10% MeOH). Yield: 31.3 mg (39%). ^{1}H NMR (500 MHz, Chloroform-d_1): δ 9.08 (d, J = 4.4 Hz, 1H), 7.83 (d, J = 5.1 Hz, 1H), 6.59 (d, J = 5.1 Hz, 1H), 5.66 – 5.55 (m, 2H), 5.37 (dd, J = 5.2, 3.7 Hz, 2H), 2.99 (d, J = 4.9 Hz, 3H), 2.96 – 2.84 (m, 1H), 2.33 (s, 3H), 1.36 (t, J = 6.7 Hz, 6H). ESI-MS(+): m/z = 404.07 [M-Cl]$^+$. To the best of our knowledge, compound **J14** has not been reported previously in the literature

[Ru(η^6-*p*-cymene)Cl(methyl-2-oxo-1,2-dihydropyridin-3-olate)] (J18). A solution of 0.2 M sodium methoxide was made by dissolving sodium (230 mg, 10.0 mmol) into methanol (50 mL). The reaction was kept as dry as possible, and the sodium methoxide solution was kept under a

nitrogen atmosphere. Bis[dichlorido(η^6-p-cymene)Ru(II)] (0.9 eq.) was added to a solution of 3-hydroxymethylpyridin-2(1H)-one (1 eq.) and sodium methoxide (1.1 eq.) in methanol (80 mL). The reaction mixture was stirred at room temperature under a nitrogen atmosphere for 4–8 h. Afterwards the solvent was evaporated under reduced pressure, and the residue was dissolved in dichloromethane and filtered to remove NaCl. Solvent was evaporated from the filtrate under reduced pressure, and the resulting solid was purified with silica-based column chromatography using DCM/MeOH as eluent (gradient of 0-10% MeOH). Yield: 40.2 mg (56%). ^{1}H NMR (500 MHz, Chloroform-d_1): δ 6.63 (dd, J = 7.7, 1.4 Hz, 1H), 6.43 (dd, J = 6.5, 1.4 Hz, 1H), 6.20 (dd, J = 7.7, 6.5 Hz, 1H), 5.55 – 5.51 (m, 2H), 5.32 (t, J = 5.5 Hz, 2H), 3.62 (s, 3H), 2.96 – 2.89 (m, 1H), 2.32 (s, 3H), 1.36 – 1.31 (m, 6H). ESI-MS(+): m/z = 360.24 [M-Cl]$^+$. To the best of our knowledge, compound **J18** has not been reported previously in the literature

[Ru(η^6-p-cymene)Cl(2-oxopyridin-1(2H)-olate)] (J19). A solution of 0.2 M sodium methoxide was made by dissolving sodium (230 mg, 10.0 mmol) into methanol (50 mL). The reaction was kept as dry as possible, and the sodium methoxide solution was kept under a nitrogen atmosphere. Bis[dichlorido(η^6-p-cymene)Ru(II)] (0.9 eq.) was added to a solution of hydroxypyridin-2(1H)-one (1 eq.) and sodium methoxide (1.1 eq.) in methanol (80 mL). The reaction mixture was stirred at room temperature under a nitrogen atmosphere for 4–8 h. Afterwards the solvent was evaporated under reduced pressure, and the residue was dissolved in dichloromethane and filtered to remove NaCl. Solvent was evaporated from the filtrate under reduced pressure, and the resulting solid was purified with silica-based column chromatography using DCM/MeOH as eluent (gradient of 0-10% MeOH). Yield: 31.4 mg (45%). ^{1}H NMR (500 MHz, Chloroform-d_1): δ 7.82 (dd, J = 6.7, 1.7 Hz, 1H), 7.14 (ddd, J = 8.7, 7.1, 1.7 Hz, 1H), 6.75 (dd, J = 8.7, 1.6 Hz, 1H), 6.36

(td, *J* = 6.9, 1.6 Hz, 1H), 5.53 (dd, *J* = 9.7, 5.8 Hz, 2H), 5.30 (dd, *J* = 8.0, 5.9 Hz, 2H), 2.97 – 2.91

(m, 1H), 2.34 (s, 3H), 1.34 (dd, *J* = 6.9, 2.9 Hz, 6H). ESI-MS(+): *m/z* = 346.11 [M-Cl]⁺. To the

best of our knowledge, compound **J19** has not been reported previously in the literature.

Class K compounds
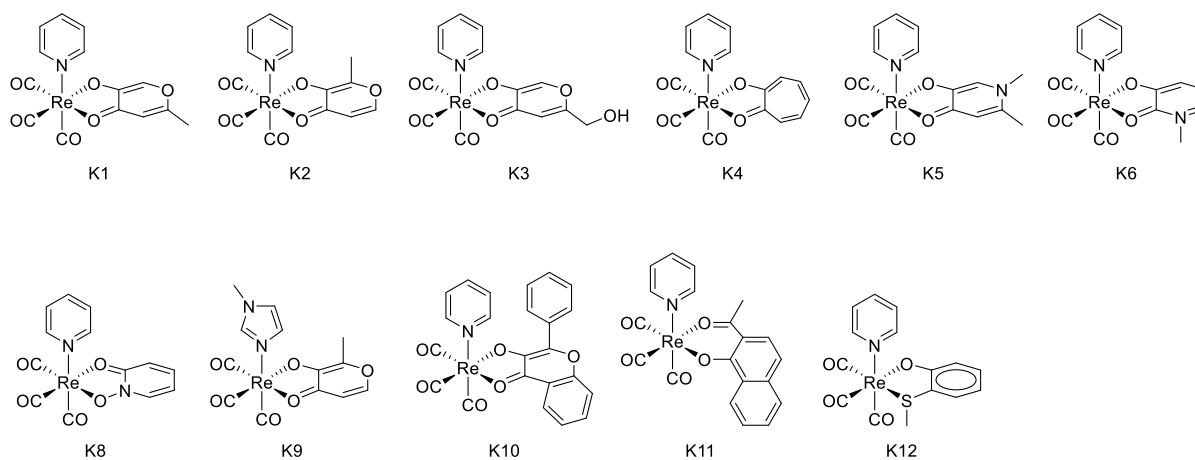
Figure 5.S11. Rhenium tricarbonyl scaffolds that comprise the Class **K** mFs.

[Re(pyridine)(allomaltol)(CO)₃] (K1).   Re₂(CO)₁₀ (50 mg, 0.08 mmol), allomaltol (22 mg,

0.18 mmol) and pyridine (12 µL, 0.15 mmol) were suspended in 3 mL toluene (stored over

molecular sieves) in a Teflon-lined steel vessel. The steel vessel was placed in an oven maintained

at 140 °C for 48 h. The mixture was transferred to a round bottom flask, and the solvent was

evaporated under reduced pressure.   The residue was purified by silica-based column

chromatography using n-hexane/EtOAc as eluent (3 min 0% EtOAc, 22.7 min 0-100% EtOAc, 4.3

min 100% EtOAc). Yield: 41 mg (57%). ¹H NMR (500 MHz, Acetone-*d₆*): *δ* 8.60 (tq, *J* = 3.2,

1.7 Hz, 2H), 8.08 – 8.01 (m, 1H), 7.87 (t, *J* = 2.4 Hz, 1H), 7.58 (ddd, *J* = 7.1, 4.7, 2.4 Hz, 2H),

6.53 (t, *J* = 2.4 Hz, 1H), 2.34 (q, *J* = 2.0 Hz, 3H). ESI-MS(+): *m/z* = 497.80 [M+Na]⁺; 446.71

[M-CO] $^+$; 350.08 [M-(O,O-Lig)]$^+$. To the best of our knowledge, compound **K1** has not been reported previously in the literature.

[Re(pyridine)(maltol)(CO)₃] (K2). $Re_2(CO)_{10}$ (75 mg, 0.11 mmol), maltol (33 mg, 0.26 mmol) and pyridine (19 µL, 0.23 mmol) were suspended in 3 mL toluene (stored over molecular sieves) in a Teflon-lined steel vessel. The steel vessel was placed in an oven maintained at 140 °C for 48 h. The mixture was transferred to a round bottom flask, and the solvent was evaporated under reduced pressure. The residue was purified by silica-based column chromatography using n-hexane/EtOAc as eluent (3 min 0% EtOAc, 22.7 min 0-100% EtOAc, 4.3 min 100% EtOAc). Yield: 64 mg (59%). ^{1}H NMR (400 MHz, Acetone-d_6): δ 8.70 – 8.57 (m, 2H), 8.02 (dd, $J = 9.5$, 6.5 Hz, 2H), 7.57 (t, $J = 6.9$ Hz, 2H), 6.63 (d, $J = 5.0$ Hz, 1H), 2.39 (s, 3H). ESI-MS(+): $m/z =$ 446.98 [M-CO] $^+$; 350.15 [M-(O,O-Lig)]$^+$. To the best of our knowledge, compound **K2** has not been reported previously in the literature.

[Re(pyridine)(5-hydroxy-2-(hydroxymethyl)-4*H*-pyran-4-one)(CO)₃] (K3). $Re_2(CO)_{10}$ (80 mg, 0.12 mmol), 5-hydroxy-2-(hydroxymethyl)-4*H*-pyran-4-one (40 mg, 0.28 mmol) and pyridine (20 µL, 0.25 mmol) were suspended in 3 mL toluene (stored over molecular sieves) in a Teflon-lined steel vessel. The steel vessel was placed in an oven maintained at 140 °C for 48 h. The mixture was transferred to a round bottom flask, and the solvent was evaporated under reduced pressure. The residue was purified by silica-based column chromatography using n-hexane/EtOAc as eluent (1.8 min 0% EtOAc, 2 min 0-40% EtOAc, 18 min 40-100% EtOAc, 3.7 min 100% EtOAc). Yield: 70 mg (58%). ^{1}H NMR (400 MHz, Acetone-d_6): δ 8.61 (d, $J = 5.6$ Hz, 2H), 8.05 (t, $J = 7.9$ Hz, 1H), 7.92 (s, 1H), 7.59 (t, $J = 6.9$ Hz, 2H), 6.71 (s, 1H), 4.84 (d, $J = 6.9$ Hz, 1H),

4.49 (t, J = 5.8 Hz, 2H). ESI-MS(+): m/z = 529.67 [M+K]$^+$; 513.78 [M+Na]$^+$; 350.10 [M-(O,O-Lig)]$^+$. To the best of our knowledge, compound **K3** has not been reported previously in the literature.

[Re(pyridine)(tropolone)(CO)₃] (K4). Re$_2$(CO)$_{10}$ (50 mg, 0.08 mmol), tropolone (22 mg, 0.18 mmol) and pyridine (12 µL, 0.15 mmol) were suspended in 3 mL toluene (stored over molecular sieves) in a Teflon-lined steel vessel. The steel vessel was placed in an oven maintained at 140 °C for 48 h. The mixture was transferred to a round bottom flask, and the solvent was evaporated under reduced pressure. The residue was purified by silica-based column chromatography using n-hexane/EtOAc as eluent (1.8 min 0% EtOAc, 20 min 0-100% EtOAc, 4 min 100% EtOAc). Yield: 24 mg (33%). ^{1}H NMR (400 MHz, Acetone-d_6): δ 8.67 – 8.51 (m, 2H), 8.02 (t, J = 8.6, 6.7 Hz, 1H), 7.58 (t, J = 5.9 Hz, 2H), 7.54 (t, J = 9.2 Hz, 2H), 7.35 (d, J = 11.0 Hz, 2H), 7.13 (t, J = 9.6 Hz, 1H). APCI-MS: m/z = 392.96 [M-pyridine+H]$^+$; 382.87 [M-cycloheptatriene+H]$^+$. ESI-MS(+): m/z = 443.3 [M-CO]$^+$; 431.5 [M-pyridine+K]$^+$; 414.5 [M-pyridine+Na]$^+$.

[Re(pyridine)(5-hydroxy-1,2-dimethylpyridin-4(1H)-one)(CO)₃] (K5). Re$_2$(CO)$_{10}$ (50 mg, 0.08 mmol), 5-hydroxy-1,2-dimethylpyridin-4(1H)-one (25 mg, 0.18 mmol) and pyridine (12 µL, 0.15 mmol) were suspended in 3 mL toluene (stored over molecular sieves) in a Teflon-lined steel vessel. The steel vessel was placed in an oven maintained at 140 °C for 48 h. The mixture was transferred to a round bottom flask, and the solvent was evaporated under reduced pressure. The residue was purified by silica-based column chromatography using n-hexane/EtOAc as eluent (1 min 0% EtOAc, 5 min 0-100% EtOAc, 15 min 100% EtOAc). Yield: 32 mg (43%). ^{1}H NMR

(400 MHz, Acetone-d_6): δ 8.66 – 8.56 (m, 2H), 7.99 (t, J = 7.9 Hz, 1H), 7.53 (t, J = 6.9 Hz, 2H), 7.16 (s, 1H), 6.33 (s, 1H), 3.71 (s, 3H), 2.31 (s, 3H). ESI-MS(+): m/z = 487.9 [M]$^+$; 460.0 [M-CO]$^+$. To the best of our knowledge, compound **K5** has not been reported previously in the literature.

[Re(pyridine)(3-hydroxy-1-methylpyridin-2(1H)-one)(CO)$_3$] **(K6).**   Re$_2$(CO)$_{10}$ (50 mg, 0.08 mmol), 3-hydroxy-1-methylpyridin-2(1H)-one (22 mg, 0.18 mmol) and pyridine (12 µL, 0.15 mmol) were suspended in 3 mL toluene (stored over molecular sieves) in a Teflon-lined steel vessel. The steel vessel was placed in an oven maintained at 140 °C for 48 h. The mixture was transferred to a round bottom flask, and the solvent was evaporated under reduced pressure. The residue was purified by silica-based column chromatography using n-hexane/EtOAc as eluent (1 min 0% EtOAc, 20 min 0-100% EtOAc, 5 min 100% EtOAc). Yield: 40 mg (55%). ^{1}H NMR (400 MHz, Acetone-d_6): δ 8.66 (dt, J = 5.1, 1.6 Hz, 2H), 8.01 (tt, J = 7.7, 1.6 Hz, 1H), 7.61 – 7.50 (m, 2H), 6.83 (dd, J = 6.5, 1.4 Hz, 1H), 6.59 (dd, J = 7.7, 1.4 Hz, 1H), 6.35 (dd, J = 7.7, 6.5 Hz, 1H), 3.71 (s, 3H). ESI-TOFMS(+): m/z 395.9890 [M-pyridine]$^+$. To the best of our knowledge, compound **K6** has not been reported previously in the literature.

[Re(pyridine)(1-hydroxypyridin-2(1H)-one)(CO)$_3$] (K8). Re$_2$(CO)$_{10}$ (75 mg, 0.11 mmol), 1-hydroxypyridin-2(1H)-one (29 mg, 0.26 mmol) and pyridine (19 µL, 0.23 mmol) were suspended in 3 mL toluene (stored over molecular sieves) in a Teflon-lined steel vessel. The steel vessel was placed in an oven maintained at 140 °C for 48 h. The mixture was transferred to a round bottom flask, and the solvent was evaporated under reduced pressure. The residue was purified by silica-based column chromatography using n-hexane/EtOAc as eluent (3.0 min 0% EtOAc, 22.7 min 0-

201

100% EtOAc, 9.3 min 100% EtOAc). Yield: 41 mg (39%). Single crystals were obtained by layering a DCM solution of the compound with n-hexanes. ^{1}H NMR (400 MHz, Acetone-d_6): δ 8.68 (d, J = 6.1 Hz, 2H), 8.04 (d, J = 8.0 Hz, 2H), 7.59 (d, J = 6.7 Hz, 2H), 7.43 – 7.34 (m, 1H), 6.80 (d, J = 9.0 Hz, 1H), 6.63 (d, J = 7.4 Hz, 1H). ESI-MS(+): m/z = 444.7 [M-O]$^+$; 483.7 [M-O-pyridine+NH$_4$]$^+$; 365.8 [M-O-pyridine]$^+$; 337.8 [M-O-pyridine-CO]$^+$. To the best of our knowledge, compound **K8** has not been reported previously in the literature.

[Re(1-methyl-1H-imidazole)(3-hydroxy-2-methyl-4H-pyran-4-one)(CO)₃] (K9). Re$_2$(CO)$_{10}$ (50 mg, 0.08 mmol), 3-hydroxy-2-methyl-4H-pyran-4-one (22 mg, 0.18 mmol) and 1-methyl-1H-imidazole (12 µL, 0.15 mmol) were suspended in 3 mL toluene (stored over molecular sieves) in a Teflon-lined steel vessel. The steel vessel was placed in an oven maintained at 140 °C for 48 h. The mixture was transferred to a round bottom flask, and the solvent was evaporated under reduced pressure. The residue was purified by silica-based column chromatography using n-hexane/EtOAc as eluent (5 min 0% EtOAc, 25 min 0-100% EtOAc, 3 min 100% EtOAc). Yield: 50 mg (68%). Single crystals were obtained by layering a DCM solution of the compound with n-hexanes. ^{1}H NMR (400 MHz, Acetone-d_6): δ 8.03 (d, J = 5.1 Hz, 1H), 7.83 (s, 1H), 7.13 (s, 1H), 6.98 (s, 1H), 6.59 (d, J = 5.1 Hz, 1H), 3.78 (s, 3H), 2.40 (s, 3H). ESI-MS(+): m/z = 516.8 [M+K]$^+$; 500.8 [M+Na]$^+$; 478.0 [M]$^+$; 450.3 [M-CO]$^+$. To the best of our knowledge, compound **K9** has not been reported previously in the literature.

[Re(pyridine)(3-hydroxyflavone)(CO)₃] (K10). Re$_2$(CO)$_{10}$ (50 mg, 0.08 mmol), 3-hydroxyflavone (42 mg, 0.18 mmol) and pyridine (12 µL, 0.15 mmol) were suspended in 3 mL toluene (stored over molecular sieves) in a Teflon-lined steel vessel. The steel vessel was placed

in an oven maintained at 140 °C for 48 h. Orange powder and dark colored crystals were obtained, filtered, washed with hexane, and air-dried. Yield: 67 mg (74%). ^{1}H NMR (400 MHz, Acetone-d_6): δ 8.78 – 8.70 (m, 4H), 8.23 (d, J = 8.1 Hz, 1H), 7.99 – 7.91 (m, 1H), 7.88 – 7.80 (m, 2H), 7.63 – 7.47 (m, 6H). ESI-MS(+): m/z = 587.8 [M]$^+$; 559.2 [M-CO]$^+$; 531.7 [M-pyridine+Na]$^+$.

[Re(pyridine)(2-acetyl-1-naphthol)(CO)₃] (K11). Re$_2$(CO)$_{10}$ (50 mg, 0.08 mmol), 2-acetyl-1-naphthol (33 mg, 0.18 mmol) and pyridine (12 µL, 0.15 mmol) were suspended in 3 mL toluene (stored over molecular sieves) in a Teflon-lined steel vessel. The steel vessel was placed in an oven maintained at 140 °C for 48 h. The mixture was transferred to a round bottom flask, and the solvent was evaporated under reduced pressure. The residue was purified by silica-based column chromatography using n-hexane/EtOAc as eluent (5 min 0% EtOAc, 16 min 0-25%, 6 min 25-100% EtOAc, 3 min 100% EtOAc). Yield: 31 mg (38%). Single crystals were obtained by layering a DCM solution of the compound with n-hexanes. ^{1}H NMR (400 MHz, Acetone-d_6): δ 8.75 (dt, J = 4.9, 1.6 Hz, 2H), 8.57 (d, J = 8.3 Hz, 1H), 8.07 – 7.98 (m, 1H), 7.69 – 7.60 (m, 2H), 7.57 (ddd, J = 7.7, 5.0, 1.5 Hz, 2H), 7.53 – 7.45 (m, 2H), 6.88 (d, J = 9.2 Hz, 1H), 2.63 (s, 3H). ESI-MS(+): m/z = 534.8 [M]$^+$; 506.9 [M-CO]$^+$. To the best of our knowledge, compound **K11** has not been reported previously in the literature.

[Re(pyridine)(2-(methylthio)phenol)(CO)₃] (K12). Re$_2$(CO)$_{10}$ (50 mg, 0.08 mmol), 2-(methylthio)phenol (21 µL, 0.18 mmol) and pyridine (12 µL, 0.15 mmol) were suspended in 3 mL toluene (stored over molecular sieves) in a Teflon-lined steel vessel. The steel vessel was placed in an oven maintained at 140 °C for 48 h. The mixture was transferred to a round bottom flask, and the solvent was evaporated under reduced pressure. The residue was purified by silica-based

column chromatography using n-hexane/EtOAc as eluent (5 min 0% EtOAc, 20 min 0-50% EtOAc, 5 min 50% EtOAc). Yield: 34 mg (45%). Single crystals were obtained by layering a DCM solution of the compound with n-hexanes. ^{1}H NMR (400 MHz, Acetone-d_6): δ 8.79 (s, 2H), 8.02 (s, 1H), 7.56 (s, 2H), 7.33 (s, 1H), 7.02 (s, 1H), 6.78 (s, 1H), 6.46 (s, 1H), 3.08 (s, 3H). ESI-MS(+): m/z = 512.5 [M+Na]$^+$. To the best of our knowledge, compound **K12** has not been reported previously in the literature.

Class L compounds

L1
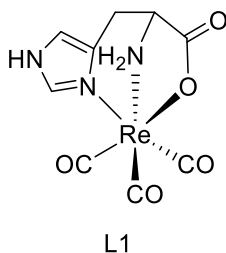
Figure 5.S12. Rhenium tricarbonyl scaffold that is the Class **L** mF.

[Re(histidine)(CO)₃] (L1). Re(CO)₅Cl (50 mg, 0.14 mmol), histidine (21 mg, 0.14 mmol), and NaOH (6 mg, 0.14 mmol) were dissolved in 3 mL H$_2$O. The reaction mixture stirred overnight at 60 °C. The mixture was concentrated under reduced pressure, and the white precipitate was collected via filtration, washed with H$_2$O, and dried in vacuo. The filtrate and washing solution were combined, concentrated, and stored at -18 °C overnight. The precipitate was collected via filtration, and both precipitation fractions were combined. The product was recrystallized from a MeOH/H$_2$O mixture. Yield: 17 mg (29%). ^{1}H NMR (500 MHz, Methanol-d_4): δ 8.01 (d, J = 1.4

Hz, 1H), 7.02 (s, 1H), 5.74 (d, J = 6.7 Hz, 1H), 5.19 – 5.01 (m, 1H), 4.03 (ddd, J = 6.6, 4.3, 2.5

Hz, 2H), 3.18 (ddd, J = 17.1, 4.3, 1.5 Hz, 2H). ESI-MS(+): m/z 448.14 [M+Na]$^+$.

Class M compounds
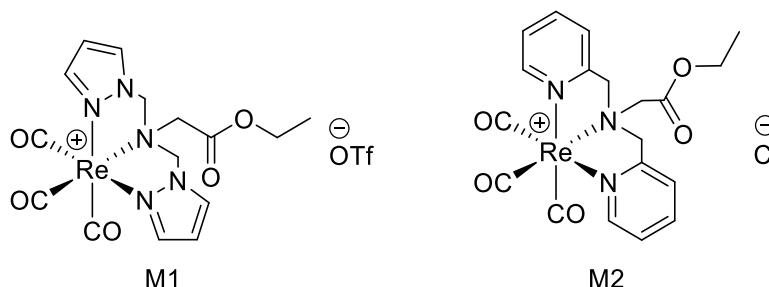
Figure 5.S13. Rhenium tricarbonyl scaffolds that comprise the Class **M** mFs.

[Re(ethyl-bis((1*H*-pyrazol-1-yl)methyl)glycinate)(CO)₃]⁺[OTf]⁻ (M1).

[Re(SO(CH₃)₂)₃(CO)₃]$^+$ triflate (75 mg, 0.11 mmol) was mixed with ethyl bis((1*H*-pyrazol-1-

yl)methyl)glycinate (33 mg, 0.13 mmol), and the mixture was stirred for 1 h at 40 °C and for

additional 0.5 h at 50 °C. MeOH was added, and the mixture was filtered. The filtrate was

concentrated under reduced pressure. The residue was dissolved with MeCN, poured into cold

Et₂O, and stored at -20 °C overnight. The obtained white solid was separated from the solution by

decantation and air dried. Yield: 35 mg (45%). ^{1}H NMR (400 MHz, Acetone-d_6): δ 8.19 (d, J =

2.8 Hz, 2H), 8.05 (d, J = 2.3 Hz, 2H), 6.54 (d, J = 12.1 Hz, 2H), 6.44 (t, J = 2.5 Hz, 2H), 5.79 (d,

J = 12.1 Hz, 2H), 4.33 (q, J = 7.1 Hz, 2H), 1.31 (t, J = 7.1 Hz, 3H). ESI-MS(+): m/z 534.17 [M$^+$].

To the best of our knowledge, compound **M1** has not been reported previously in the literature.

To make the ligand ethyl bis((1*H*-pyrazol-1-yl)methyl)glycinate, glycine ethyl ester hydrochloride (203 mg, 1.46 mmol) was suspended in 20 mL DCM. Triethylamine (203 µL, 1.46 mmol) was added slowly, and the mixture was stirred for 10 min until most of the starting material was dissolved. (1*H*-pyrazol-1-yl)methanol (300 mg, 3.06 mmol) was added, and the mixture was stirred for 5 days at room temperature. H_2O and brine were used to wash the organic layer, which was then dried over $MgSO_4$ and evaporated to dryness. Residual (1*H*-pyrazol-1-yl)methanol was sublimed from the product in vacuo at 50 °C. The product was obtained as a colorless oil. Yield: 310 mg (81%). ^{1}H NMR (400 MHz, Chloroform-d_1): δ 7.56 (dd, J = 19.2, 2.1 Hz, 4H), 6.30 (t, J = 2.1 Hz, 2H), 5.11 (s, 4H), 4.11 (q, J = 7.1 Hz, 2H), 3.61 (s, 2H), 1.24 (t, J = 7.1 Hz, 3H). ESI-MS(+): *m/z* 286.16 [M+Na]$^+$.

To make the precursor $[Re(SO(CH_3)_2)_3(CO)_3]^+$ triflate, $Re(CO)_5Cl$ (100 mg, 0.28 mmol) in DMSO (60 µL, 0.84 mmol) was suspended in 8 mL acetone, and the flask was covered with aluminum foil before $AgCF_3SO_3$ (71 mg, 0.28 mmol) was added. The mixture was refluxed overnight under exclusion from light. The precipitation was collected via filtration, and the filtrate was concentrated under reduced pressure before it was poured into Et_2O and stored at -20 °C overnight. The obtained colorless crystals were separated from the solvent by decantation and air dried. Yield: 143 mg (87%).

[Re(ethyl-bis((1*H*-pyrazol-1-yl)methyl)glycinate)(CO)₃]⁺[Cl]⁻ (M2). $Re(CO)_5Cl$ (52 mg, 0.14 mmol) and ethyl bis(pyridin-2-ylmethyl)glycinate (49 mg, 0.17 mmol) were suspended in 1 mL MeOH, and the mixture was irradiated in a microwave at 110 °C for 10 min. The resulting solution was concentrated under reduced pressure and poured into cold Et_2O. The white precipitation was

separated from the solution through centrifugation and dried in vacuo. Yield: 78 mg (86%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 8.82 (d, J = 5.5 Hz, 2H), 8.01 (td, J = 7.8, 1.5 Hz, 2H), 7.66 (d, J = 7.9 Hz, 2H), 7.43 (t, J = 6.8 Hz, 2H), 5.20 (d, J = 17.2 Hz, 2H), 5.10 (d, J = 17.1 Hz, 2H), 4.78 (s, 2H), 4.24 (q, J = 7.1 Hz, 2H), 1.28 (t, J = 7.2 Hz, 3H). ESI-MS(+): m/z 556.10 [M$^+$].

To make ethyl (pyridin-2-ylmethyl)glycinate, glycine ethyl ester hydrochloride (474 mg, 3.39 mmol) was dissolved in 5 mL MeOH (stored over molecular sieves) under a N$_2$ atmosphere, and the solution was cooled to 0 °C. Picolinaldehyde (708 µL, 7.47 mmol) was added, followed by addition of NaBH$_4$ (206 mg, 5.43 mmol) in small portions. The reaction mixture was stirred overnight coming to room temperature during this time. H$_2$O (1 mL) was added slowly to quench the reaction, and the mixture was stirred for an additional 5 h. DCM was used to extract the product. The combined organic layers were washed with H$_2$O and brine, dried over MgSO$_4$, and evaporated to dryness. The product was purified by silica-based column chromatography using DCM/MeOH as eluent (5 min 0% MeOH, 32 min 0-8% MeOH, 10 min 8% MeOH). Yield: 522 mg (79%). ^{1}H NMR (400 MHz, Chloroform-d_1): δ 8.57 (s, 1H), 7.66 (d, J = 6.4 Hz, 1H), 7.35 (d, J = 6.4 Hz, 1H), 7.27 (d, J = 4.9 Hz, 1H), 7.18 (q, J = 6.2 Hz, 1H), 4.77 (d, J = 5.2 Hz, 1H), 4.20 (p, J = 6.9 Hz, 2H), 3.97 (d, J = 5.1 Hz, 2H), 3.49 (d, J = 5.6 Hz, 2H), 1.28 (q, J = 6.7 Hz, 3H).

To make ethyl bis(pyridin-2-ylmethyl)glycinate, ethyl (pyridin-2-ylmethyl)glycinate (522 mg, 2.69 mmol) and 2-(bromomethyl)pyridine hydrobromide (680 mg, 2.69 mmol) were dissolved in DMF (stored over molecular sieves), followed by addition of triethylamine (783 µL, 5.65 mmol). More DMF was added until most of the precipitation was gone. The reaction mixture stirred for 36 h at room temperature. H$_2$O was added, and the product was extracted with EtOAc. The

combined organic layers were washed with H_2O and brine and evaporated to dryness. Yield: 582 mg (76%). ^{1}H NMR (400 MHz, Acetone-d_6): δ 8.48 (d, J = 4.9 Hz, 2H), 7.74 (dt, J = 7.6, 3.8 Hz, 2H), 7.61 (d, J = 7.9 Hz, 2H), 7.22 (dd, J = 7.5, 4.9 Hz, 2H), 4.13 (q, J = 7.1 Hz, 2H), 3.98 (s, 4H), 3.46 (s, 2H), 1.22 (t, J = 7.1 Hz, 3H).

X-ray Crystallography

Class K compounds were crystallized as described in the synthesis and characterization section. All compounds crystallized as light-yellow blocks. The single crystal X-ray diffraction studies were carried out on a Bruker Kappa APEX-II CCD diffractometer equipped with Mo K$_\alpha$ radiation (λ = 0.71073 Å). Suitable crystals were mounted on a cryoloop with Paratone oil. Data were collected in a nitrogen gas stream at 100(2) K using Φ and ω scans. Crystal-to-detector distance was 40 mm and exposure time was 5 seconds per frame using a scan width of 0.75-2.0°. Data collection was 100% complete to 25.00° in θ. The data were integrated using the Bruker SAINT software program and scaled using the SADABS software program. Solution by direct methods (SHELXT) produced a complete phasing model consistent with the proposed structure.[65] All nonhydrogen atoms were refined anisotropically by full-matrix least-squares (SHELXL-2014). All hydrogen atoms were placed using a riding model. Their positions were constrained relative to their parent atom using the appropriate HFIX command in SHELXL-2014. For compound K5, the absolute stereochemistry of the molecule was established by anomalous dispersion using the Parson's method with a Flack parameter of 0.025(14). The crystal data file of all complexes was deposited into the Cambridge Crystallographic Data Centre (CCDC) with deposition numbers 1962326-1962330. Crystal structures are shown in Figure 5.S14, and crystallographic data are summarized in Table 5.S1.
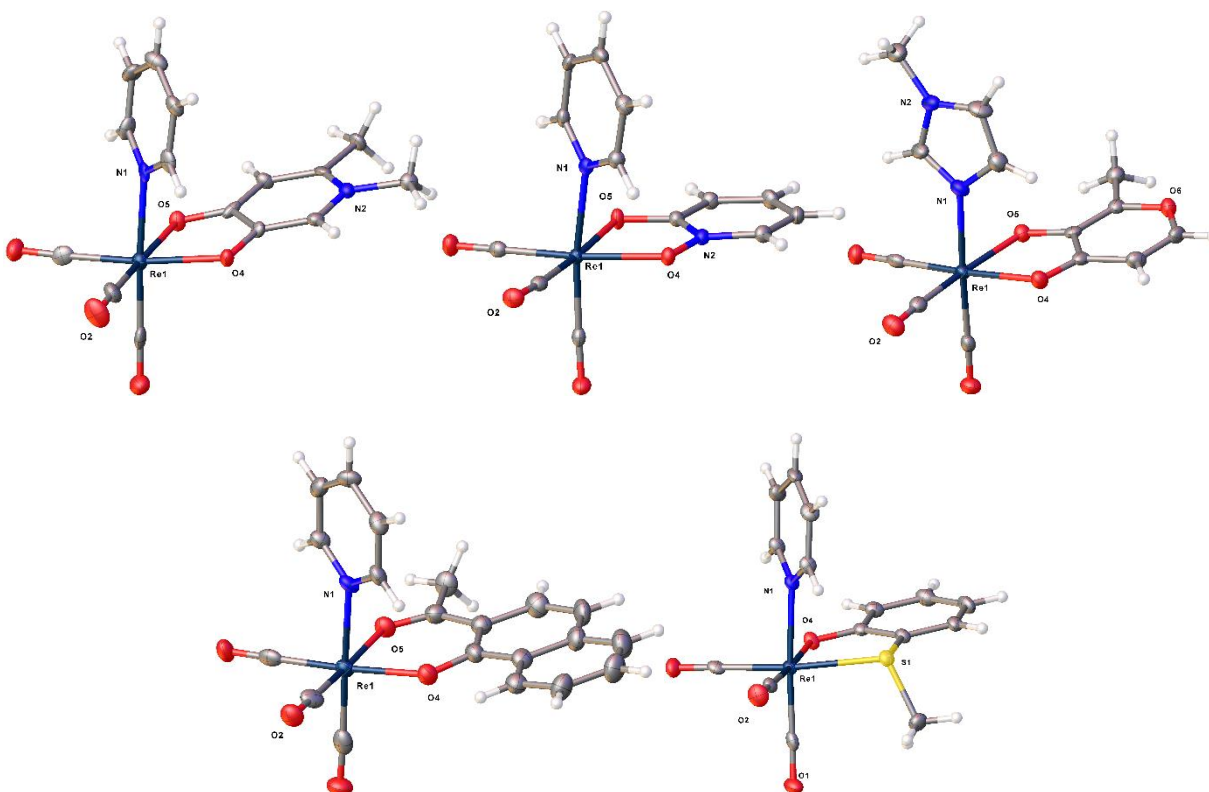
Figure 5.S14. Structures of Class K compounds rendered as an ORTEP with atoms at 50% thermal probability ellipsoids. *Top row*: K5, K8, and K9; *bottom row*: K11 and K12. Only one molecule is shown for each mF; however, all these compounds form in enantiomeric mixtures due to different binding orientations of the bidentate ligand, and both enantiomers are present in the unit cell. Color scheme: carbon = gray, hydrogen = white, oxygen = red, nitrogen = blue, sulfur = yellow, rhenium = navy blue.

Table 5.S1. Crystal data and structure refinement for Class K compounds

Compound	K5	K8	K9
Identification code	1962326	1962327	1962328
Empirical formula	$C_{15}H_{13}N_2O_5Re$	$C_{13}H_9N_2O_5Re$	$C_{13}H_{11}N_2O_6Re$
Molecular formula	$C_{15}H_{13}N_2O_5Re$	$C_{13}H_9N_2O_5Re$	$C_{13}H_{11}N_2O_6Re$
Formula weight (g/mol)	487.47	459.42	477.44
Temperature (K)	100.0	100.0	100.0
Crystal system	Orthorhombic	Monoclinic	Triclinic
Space group	$P2_12_12_1$	$P\,1\,2_1/c\,1$	P-1
a (Å)	6.3450(2)	9.1929(5)	9.1080(5)
b (Å)	14.5492(5)	13.3161(7)	9.5099(5)
c (Å)	16.7909(7)	11.3352(6)	9.7673(5)
α (°)	90	90	70.669(2)
β (°)	90	106.800(2)	80.012(2)
γ (°)	90	90	61.781(2)
Volume (Å^3)	1550.05(10)	1328.36(12)	703.28(7)
Z	4	4	2
ρ_{calc} (g/cm^3)	2.089	2.297	2.255
Absorption coefficient (mm^{-1})	7.866	9.171	8.671
F(000)	928	864	452
Crystal size (mm^3)	0.213 × 0.145 × 0.117	0.131 × 0.106 × 0.094	0.127 × 0.115 ×0.086
Theta range for data collection (°)	1.852 to 26.396	2.314 to 26.404	2.210 to 26.368
Index ranges	-7 ≤ h ≤ 7, -18 ≤ k ≤ 18, -20 ≤ l ≤ 20	-11 ≤ h ≤ 11, -16 ≤ k ≤ 16, -14 ≤ l ≤ 13	-11 ≤ h ≤ 11, -11 ≤ k ≤ 8, -12 ≤ l ≤ 12
Reflections collected	17683	25383	10244
Independent reflections	3161 [R_{int} = 0.0536, R_{sigma} = 0.0444]	2729 [R_{int} = 0.0373, R_{sigma} = 0.0197]	2866 [R_{int} = 0.0300, R_{sigma} = 0.0326]
Data / restraints / parameters	3161 / 0 / 211	2729 / 0 / 191	2866 / 37 / 245
Goodness-of-fit on F^2	1	1.068	1.047
Final R indices [$I{\geq}2\sigma(I)$]	R_1 = 0.0239, wR_2 = 0.0424	R_1 = 0.0145, wR_2 = 0.0302	R_1 = 0.0200, wR_2 = 0.0378
R indices [all data]	R_1 = 0.0517, wR_2 = 0.0501	R_1 = 0.0175, wR_2 = 0.0309	R_1 = 0.0233, wR_2 = 0.0387
Absolute structure parameter	0.025(14)	n/a	n/a
Largest diff. peak and hole (e/Å^3)	0.494 and -0.533	0.404 and -0.587	0.563 and -0.513

Table 5.S1. Crystal data and structure refinement for Class K compounds (continued).

Compound	K11	K12
Identification code	1962329	1962330
Empirical formula	$C_{20.50}H_{15}ClNO_5Re$	$C_{15}H_{12}NO_4ReS$
Molecular formula	$C_{20}H_{14}NO_5Re$, $0.5(CH_2Cl_2)$	$C_{15}H_{12}NO_4ReS$
Formula weight (g/mol)	576.98	488.52
Temperature (K)	100.0	100.0
Crystal system	Monoclinic	Monoclinic
Space group	$P\,1\,2_1/c\,1$	$P\,1\,2_1/n\,1$
a (Å)	20.753(2)	8.5304(5)
b (Å)	7.0352(8)	8.1560(5)
c (Å)	13.7941(18)	22.7941(15)
α (°)	90	90
β (°)	105.388(7)	96.365(2)
γ (°)	90	90
Volume ($Å^3$)	1941.8(4)	1576.10(17)
Z	4	4
ρ_{calc} (g/cm^3)	1.974	2.059
Absorption coefficient (mm^{-1})	6.428	7.858
F(000)	1108	928
Crystal size (mm^3)	$0.153 \times 0.147 \times 0.136$	$0.087 \times 0.065 \times 0.063$
Theta range for data collection (°)	2.036 to 25.381	2.470 to 26.379
Index ranges	$-24 \le h \le 24$, $-8 \le k \le 6$, $-16 \le l \le 16$	$-9 \le h \le 10$, $-10 \le k \le 10$, $-28 \le l \le 28$
Reflections collected	26253	21549
Independent reflections	3563 [R_{int} = 0.0959, R_{sigma} = 0.0639]	3228 [R_{int} = 0.0421, R_{sigma} = 0.0297]
Data / restraints / parameters	3563 / 1 / 267	3228 / 0 / 200
Goodness-of-fit on F^2	1.049	1.051
Final R indices [$I \ge 2\sigma(I)$]	R_1 = 0.0376, wR_2 = 0.0759	R_1 = 0.0207, wR_2 = 0.0355
R indices [all data]	R_1 = 0.0535, wR_2 = 0.0815	R_1 = 0.0285, wR_2 = 0.0372
Largest diff. peak and hole (e/Å^3)	0.923 and -1.342	0.465 and -0.821

Molecular volume analysis of mF library
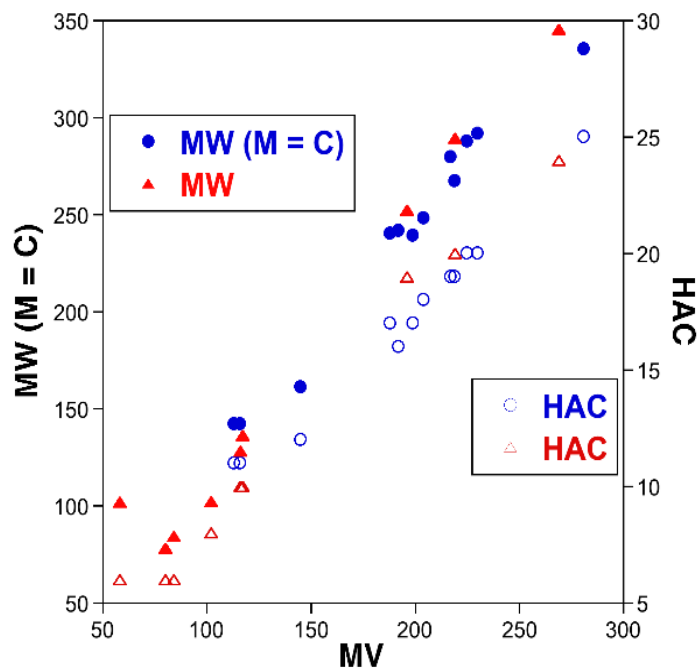
Figure 5.S15. The size (molecular volume, MV in Å^3) of organic fragments (red triangles) versus mF (blue circles) are indistinguishable when these molecules are compared using an apparent MW (filled symbols, where the metal ion is assigned the mass of a carbon atom, M = C) or when comparing heavy atom count (open symbols, HAC, e.g. the number of non-hydrogen atoms) in the fragment. In addition, the correlation of either MW or HAC vs. MV of organic and mFs is effectively the same; therefore, MV $\leq 300 \ \text{Å}^3$ is proposed as a suitable 'rule' for mF rule-of-3 compliance rather than MW ≤ 300 Da.

PA$_N$ endonuclease protein expression and purification

Expression and purification of PA$_N$ endonuclease was performed as reported previously.[66] Pandemic H1N1 N-terminal PA (PA$_N$) endonuclease Δ52-64:Gly truncated construct was expressed from a pET-28a parent vector containing a kanamycin-resistance reporter gene with expression inducible by the Lac 1 operon. PA$_N$ endonuclease was expressed as an 8-histidine tagged fusion protein with cleavable by TEV protease. The Transformation protocol was adapted from pET system manual (Novagen) using single competent BL21 cells. Briefly, 1 µL of 25 ng/µL recombinant plasmid was used for transformation. Cells were mixed by flicking with plasmid and were heat shocked at 42 ℃ for 30 sec followed by incubation on ice for 5 minutes. Outgrowth was plated on LB agarose plates contain 50 µg/mL kanamycin and was incubated overnight at 37 ℃. One colony was scraped from the LB plate and added to 50 mL of SOC broth containing 50 µg/mL kanamycin and was incubated for 5 h at 37 ℃ with shaking at 125 rpm. Glycerol stocks of this culture were prepared (0.9 mL cultured media + 0.1 mL 80% glycerol) and column frozen for future expressions. SOC media (100 mL) containing 50 µg/mL kanamycin was combined with 1 mL frozen cell glycerol stock and was incubated with shaking at 200 rpm at 37 ℃. When the OD$_{600}$ of this starter culture reached >2 (~5-6 h), the culture was equally divided into 6 × 1 L batches of expression media (TB media with added 0.2% dextrose, 0.1 mM MnCl$_2$, and 0.1 mM MgSO$_4$) containing 50 µg/mL kanamycin. Cells were grown to the beginning of log phase (OD$_{600}$ = between 0.4-0.6) at 37 ℃ with shaking at 200 rpm. When the OD$_{600}$ = 0.4-0.6, the cultures were cooled to room temperature over ice. Expression of PA$_N$ endonuclease was then induced by the addition of IPTG to a final concentration of 0.1 mM. The cultures were grown with vigorous shaking (250 rpm) overnight at room temperature. The caps of the flasks were completely removed

to increase aeration. After ~18 h, the cells were harvested by centrifuging at 2000g for 30 min at 4 ℃. The resulting cell paste was stored at -80 ℃ prior to lysis.

The cell paste was thawed on ice for 2 h and re-suspended 1:1 (v/v) with lysis buffer (20 mM Na_2PO_4, 500 mM NaCl, 25 mM imidazole, 1 mM $MgCl_2$, 2 mM dithiothreitol, 0.2% Triton-X, pH=7.4) plus EDTA free protease inhibitor (1 pellet per ~50 mL lysis buffer), lysozyme (1 mg/mL), and DNAse-1 (10-100 µg/mL). Further lysis was performed using a probe sonicator over 10 min with cycles of 10 sec sonication and 20 sec rest. The cell suspension was kept in a water/ice bath during sonication. The lysates were free-flowing and homogenous. The lysates were shaken at 125 rpm for 30 min on ice. Cell debris was then pelleted by centrifugation at 10000 rpm for 45 min at 4 ℃. The supernatant was decanted from the pellet and was filtered through 0.45 µm syringe filters. The resulting lysate was loaded onto 2×5 mL HisTrap HP (Pharmacia) columns that had been previously charged with Ni ions. The columns were washed with binding buffer (20 mM Na_2PO_4, 500 mM NaCl, 25 mM imidazole, pH 7.4) until fraction absorbance reached a steady baseline. The protein was then eluted over a 45 min gradient at a flow rate of 4 mL/min, from 0-100% elution buffer (20 mM Na_2PO_4, 500 mM NaCl, 500 mM imidazole, pH 7.4). PA_N endonuclease eluted between 40-60% elution buffer. SDS-PAGE analysis showed a band corresponding to PA_N endonuclease running at ~23 kDa.

Fractions containing endonuclease protein were combined in a 10K MWCO dialysis bag with 2000 units of TEV protease and dithiothreitol final concentration of 1 mM. The solutions were dialyzed against dialysis buffer (100 mM NaCl, 1 mM dithiothreitol, 1 mM $MnCl_2$, 20 mM Tris, 5% glycerol, pH 8.0) overnight at 4 ℃ with two buffer exchanges. A white precipitate forms over time along the inside and outside walls of the dialysis tubing. After buffer exchange, the solution was filtered through a 0.45 µm filter and was concentrated to 5-10 mg/ml using a

pressurized Amicon system. This protein was suitable for use in fluorescence quenching-based nuclease assays.

PA$_N$ endonuclease fluorescence quenching-based nuclease activity assay

PA$_N$ endonuclease activity assays were carried out as previously reported, with slight modification.[66] The only alteration made was to remove 2-mercaptoethanol from the assay buffer. Assays were performed using Black Costar 96-well plates. Each well contained a total volume of 100 μL comprised of buffer (20 mM Tris, 150 mM NaCl, 2 mM MnCl$_2$, 0.2% Triton-X100, pH 8.0), influenza PA$_N$ endonuclease (25 nM), inhibitor (various concentrations), and fluorescent ssDNA-oligo substrate (200 nM). DMSO was present at a final concentration of 5% in each well. The presence of this concentration of DMSO was found to have a negligible effect on the assay. A single-stranded, 17-mer DNA substrate labeled with a 5′-FAM fluorophore and a 3′-TAMRA quencher ([6-FAM]AATCGCAGGCAGCACTC[TAM], Sigma-Aldrich) was employed as the substrate. All assay components were pipetted into the plate, and ultimately, the substrate was added using a multi-channel pipette, and the assay was immediately started. Samples were prepared in triplicates. Background wells consisting of all assay components except enzyme were prepared for each sample. Positive and negative controls were prepared on each plate to gauge the fluorescence signal of fully active protein and the absence of protein. Change in fluorescence of each well was measured by a Synergy H4 Hybrid Multi-Mode Microplate Reader (BioTek) at 39 second intervals over 45 min at 37 ºC (λ_{ex} = 485 nm; λ_{em} = 528 nm). The gain was set to 100. Typically, data collected between 20 and 35 minutes was used in the activity calculations, as this data range yielded the steadiest slope. The slope of the fluorescence signal for each sample was background corrected, and percent inhibition was determined by normalizing the slope of the

215

sample to that of the positive and negative controls. For general library screening to identify hits, mFs were screened at a concentration of 200 µM.

NDM-1 meropenem activity assay

NDM-1 protein was expressed in purified in the laboratory of Dr. Michael W. Crowder according to previously published methods.[51] The NDM-1 activity assay was performed following a modified published procedure using meropenem as the substrate.[51] Briefly, the decrease in absorption of meropenem at 300 nm was monitored in UV-transparent 96-well plates (Corning product #3635). The buffer used was 50 mM HEPES and 2 mM CHAPS at pH 7. To each well, 1 µL of each compound (various concentrations) and 69 µL of NDM-1 (2.50 nM final concentration) was added. After incubating the plate at 25 ℃ for 20 min, 30 µL of meropenem (180 µM final concentration) was added to each well to initiate the reaction. Positive control wells consisted of NDM-1 and meropenem (no inhibitor; fully active), and negative control wells consisted of meropenem (no enzyme; no reaction). Absorbance of each plate was measured using a Synergy H4 Hybrid Multi-Mode Microplate Reader (BioTek) at 300 nm over 5 min with 15 sec intervals.

Hsp90 fluorescence polarization activity assay

Hsp90α N-terminal domain (Hsp90) assay kits were purchased from BPS Bioscience (catalog #50293). The fluorescence polarization assays were carried out as described in the kit instruction with slight modification. The only alteration made was to replace the 2 mM dithiothreitol (final concentration) with 200 µM tris(2-carboxyethyl)phosphine (TCEP), which was validated using positive and negative controls, as well as geldanamycin dose response evaluation. Briefly, 5 µL FITC-labeled geldanamycin (100 nM final concentration) and 10 µL

inhibitor (various concentrations) were added to each well of a 96-well, black, low binding, microtiter plate. Each well also contained 15 µL assay buffer, 5 µL TCEP (at 4 mM), 5 µL bovine serum albumin (at 2 mg/mL), and 40 µL H$_2$O. The reaction was initiated by adding 20 µL Hsp90 (at 17 ng/µL final concentration) to each well. Positive control wells consisted of FITC-labeled geldanamycin and Hsp90 (no inhibitor), and negative control wells consisted of FITC-labeled geldanamycin (no inhibitor or protein). Plates were incubated for 2 hours with slow shaking at room temperature prior to reading. After incubation, using a Synergy H4 Hybrid Multi-Mode Microplate Reader (BioTek), excitation was performed at 485 nm, and emission was measured at 530 nm. Fluorescence polarization was calculated according to the following equation: $mP = \left(\frac{I_{parallel} - I_{perpendicular}}{I_{parallel} + I_{perpendicular}} \right) x1000$. A G-factor correction was applied. Inhibition values were calculated by comparing the fluorescence polarization of sample wells versus control wells. Z-scores ranged from 0.82 – 0.93.

For general library screening to identify hits, mFs were screened at a concentration of 200 µM. Runs were performed in duplicate or triplicate.
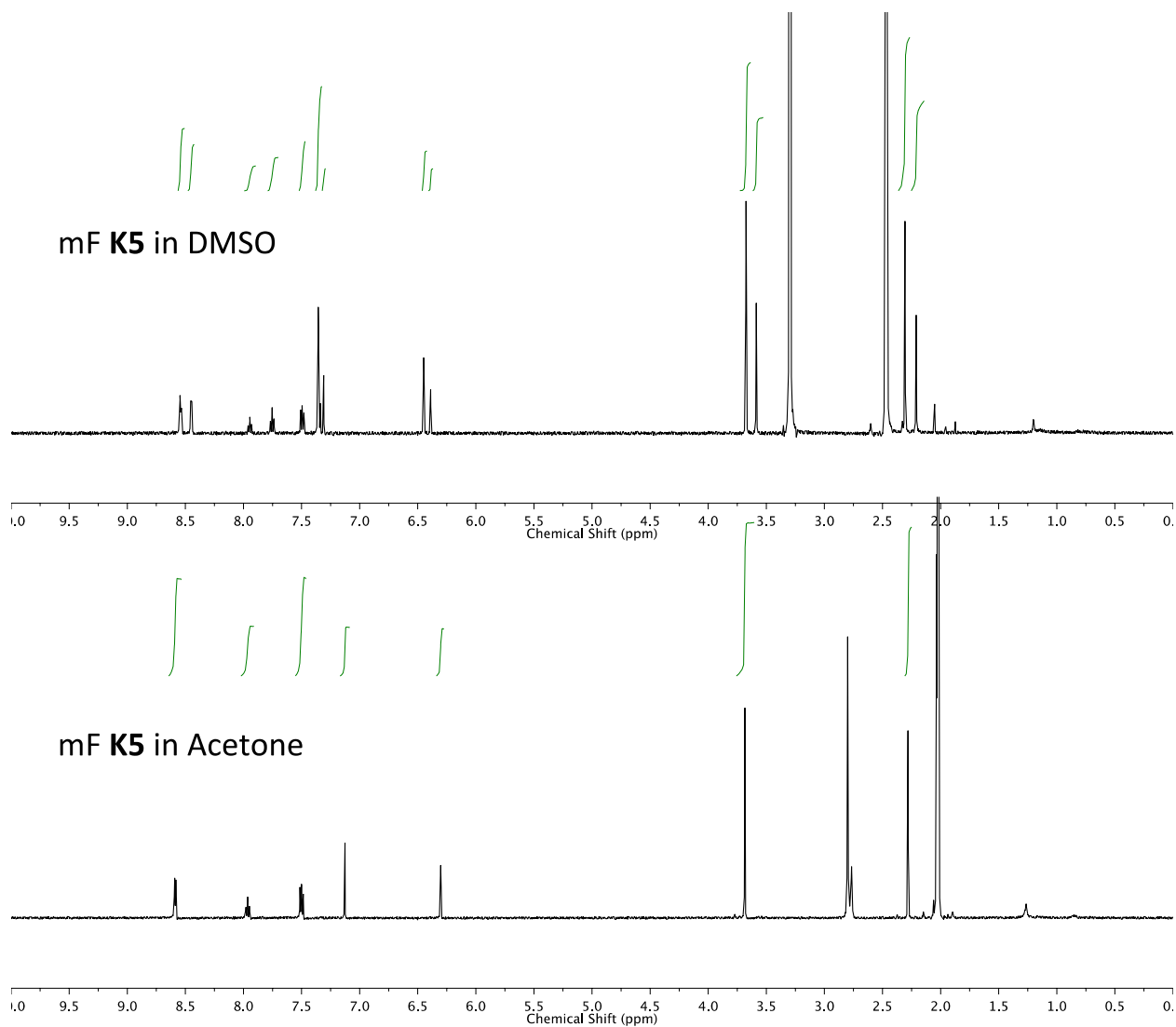
Figure 5.S16. ^{1}H NMR analysis of a representative Class **K** mF in acetone (no loss of heterocycle) and in DMSO, in which the pyridine is lost to generate a second DMSO coordinated complex.
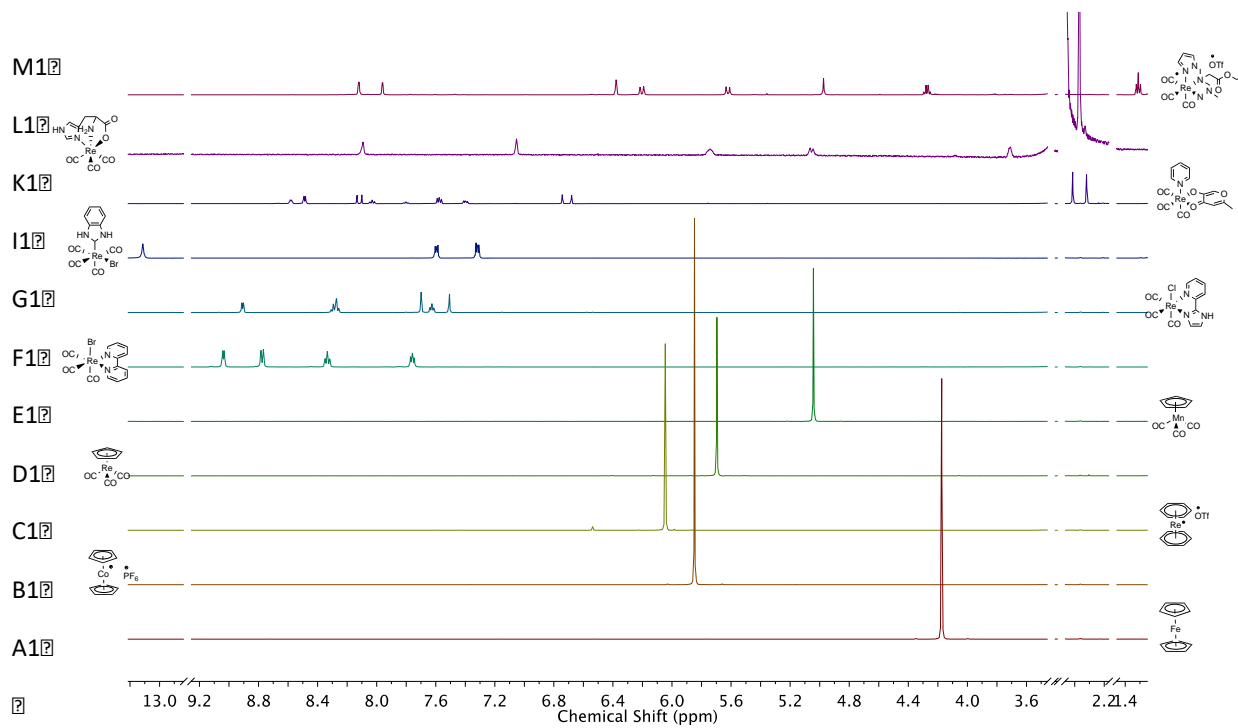
Figure 5.S17. ^{1}H NMR analysis of the stability of a representative entry from each fragment class in d6-DMSO.

PMI and 3D Structure Analysis

For each mF, a search was carried out in the Cambridge Structural Database (CSD) system (version 5.40, November 2018)[67] using the ConQuest platform (2.0.0)[68] for identical or the most analogous structure. The CIF file for each of the reported database identifiers in Table 5.S2 was downloaded and opened in MOE, and the PMI was calculated from the identical or minimally modified structure.

Table 5.S2. The mFs and the corresponding CSD database identifiers for the X-ray crystallographic structures used for the PMI calculations, the two normalized PMI values, and their 3D score.

mF	CSD Identifier	I_1/I_3	I_2/I_3	3D Score
A1	FEHYAS	0.49	1.00	1.49
A2	DUKQOQ	0.34	0.86	1.20
A3	SURFIU	0.47	0.81	1.28
A4	BIWKEX	0.46	0.84	1.30
A5	DEQCAD	0.39	0.88	1.27
A7	CIZFAS	0.45	0.86	1.31
A8	DEXHOE	0.40	0.88	1.27
A9	UZOMEC	0.22	0.90	1.12
A10	BATLEO	0.28	0.86	1.14
A11	PACFER	0.24	0.87	1.11
A12	ZZZHCY01	0.31	0.88	1.20
A14	FAJVOE	0.31	0.88	1.19
A15	FAJVOE	0.30	0.90	1.20
A16	DICBIA	0.31	0.92	1.23

Table 5.S2. The mFs and the corresponding CSD database identifiers for the X-ray crystallographic structures used for the PMI calculations, the two normalized PMI values, and their 3D score (continued).

mF	CSD Identifier	I_1/I_3	I_2/I_3	3D Score
A17	XUXKOQ	0.34	0.88	1.22
A18	GAQQAQ	0.35	0.87	1.23
A19	BATLEO	0.26	0.85	1.10
A20	DBEFER01	0.17	0.96	1.13
A21	ZZDKO01	0.57	0.75	1.32
A22	FEROCA	0.61	0.69	1.31
B1	BEWBEM	0.47	0.99	1.47
B2	BOVQIM	0.46	0.82	1.28
C1	PEVDEB	0.55	1.00	1.55
D1	COCPRE	0.72	0.99	1.72
D2	KITCUL	0.68	0.91	1.58
D3	COXCOI	0.59	0.88	1.47
D4	KIXZOH	0.54	0.93	1.47
D5	UFAZUY	0.67	0.92	1.59
D6	QAYCIF	0.35	0.92	1.27
D7	COXCOI	0.58	0.88	1.46
D8	KIXZOH	0.59	0.89	1.48
D9	KIXZOH	0.44	0.94	1.38
D10	DUKREH	0.31	0.94	1.24
E1	CPMNCO	0.77	0.99	1.76
F1	ETUQAN	0.72	0.85	1.56
F2	XELXAO	0.60	0.73	1.34
F3	MUHTEQ	0.52	0.66	1.18

Table 5.S2. The mFs and the corresponding CSD database identifiers for the X-ray crystallographic structures used for the PMI calculations, the two normalized PMI values, and their 3D score (continued).

mF	CSD Identifier	I_1/I_3	I_2/I_3	3D Score
F4	SEPPAH	0.63	0.68	1.31
G1	DIYJAY	0.58	0.79	1.38
G2	DIYJAY	0.52	0.72	1.25
G3	DIYJAY	0.58	0.79	1.38
G4	DIYJAY	0.59	0.79	1.38
H1	WAPJOO	0.69	0.80	1.49
I1	PEHJEU	0.39	0.88	1.26
I2	PEHJEU	0.33	0.95	1.28
I3	PEHJEU	0.41	0.84	1.25
I4	PEHJEU	0.45	0.76	1.20
J1	YULJAR	0.34	0.88	1.21
J2	QUHNOX	0.32	0.82	1.14
J3	YULJIZ	0.35	0.89	1.23
J6	YULJIZ	0.20	0.93	1.13
J8	ODELAN	0.30	0.89	1.19
J12	OTIMOV	0.32	0.88	1.20
J13	OTIMOV	0.34	0.88	1.22
J14	YULJAR	0.32	0.92	1.24
J18	IJELAL	0.36	0.88	1.24
J19	KARYAF	0.44	0.74	1.18
K1	1962326	0.49	0.79	1.27
K2	1962326	0.55	0.80	1.34
K3	1962326	0.41	0.83	1.24

Table 5.S2. The mFs and the corresponding CSD database identifiers for the X-ray crystallographic structures used for the PMI calculations, the two normalized PMI values, and their 3D score (continued).

mF	CSD Identifier	I_1/I_3	I_2/I_3	3D Score
K4	1962326	0.53	0.77	1.30
K5	1962326	0.45	0.82	1.26
K6	1962326	0.55	0.80	1.36
K8	1962327	0.58	0.71	1.29
K9	1962328	0.57	0.73	1.29
K10	1962326	0.57	0.84	1.41
K11	1962329	0.50	0.94	1.43
K12	1962330	0.62	0.77	1.39
L1	EZATAA	0.55	0.80	1.35
M1	DISDEP	0.58	0.99	1.57
M2	BETDIO	0.61	0.90	1.51

5.10 References

1.   Hung, A.; Ramek, A.; Wang, Y.; Kaya, T.; Wilson, J.; Clemons, P.; Young, D., Route to three-dimensional fragments using diversity-oriented synthesis. *Proc. Natl. Acad. Sci. USA* **2011,** *108*, 6799-6804.

2.   Sauer, W. H. B.; Schwarz, M. K., Molecular shape diversity of combinatorial libraries: A prerequisite for broad bioactivity. *J. Chem. Inf. Comp. Sci.* **2003,** *43*, 987-1003.

3.   Lovering, F.; Bikker, J.; Humblet, C., Escape from flatland: Increasing saturation as an approach to improving clinical success. *J. Med. Chem.* **2009,** *52*, 6752-6756.

4.   Kidd, S. L.; Osberger, T. J.; Mateu, N.; Sore, H. F.; Spring, D. R., Recent spplications of diversity-oriented synthesis toward novel, 3-dimensional fragment collections. *Front. Chem.* **2018,** *6*, 460.

5.   Aldeghi, M.; Malhotra, S.; Selwood, D. L.; Chan, A. W., Two- and three-dimensional rings in drugs. *Chem. Biol. Drug Des.* **2014,** *83* (4), 450-61.

6.   Firth, N. C.; Brown, N.; Blagg, J., Plane of best fit: A novel method to characterize the three-dimensionality of molecules. *J. Chem. Inf. Model.* **2012,** *52* (10), 2516-2525.

7.   Galloway, W. R. J. D.; Isidro-Llobet, A.; Spring, D. R., Diversity-oriented synthesis as a tool for the discovery of novel biologically active small molecules. *Nature Commun.* **2010,** *1*, 80.

8.   Reekie, T. A.; Williams, C. M.; Rendina, L. M.; Kassiou, M., Cubanes in medicinal chemistry. *J. Med. Chem.* **2019,** *62* (3), 1078-1095.

9.   Can, D.; Spingler, B.; Schmutz, P.; Mendes, F.; Raposinho, P.; Fernandes, C.; Carta, F.; Innocenti, A.; Santos, I.; Supuran, C. T.; Alberto, R., [(Cp-R)M(CO)$_3$] (M = Re or ^{99m}Tc) Arylsulfonamide, arylsulfamide, and arylsulfamate conjugates for selective targeting of human carbonic anhydrase IX. *Angew. Chem. Int. Ed.* **2012,** *57*, 3354-3357.

10.	Dorr, M.; Meggers, E., Metal complexes as structural templates for targeting proteins. *Curr. Opin. Chem. Biol.* **2014,** *19*, 76-81.

11.	Gasser, G.; Ott, I.; Metzler-Nolte, N., Organometallic anticancer compounds. *J. Med. Chem.* **2011,** *54* (1), 3-25.

12.	Huisman, M.; Kodanko, J. P.; Arora, K.; Herroon, M.; Alnaed, M.; Endicott, J.; Podgorski, I.; Kodanko, J. J., Affinity-enhanced luminescent Re(I)- and Ru(II)-based inhibitors of the cysteine protease cathepsin L. *Inorg. Chem.* **2018,** *57* (13), 7881-7891.

13.	Vaden, R. M.; Guillen, K. P.; Salvant, J. M.; Santiago, C. B.; Gibbons, J. B.; Pathi, S. S.; Arunachalam, S.; Sigman, M. S.; Looper, R. E.; Welm, B. E., A cancer-selective zinc ionophore inspired by the natural product naamidine A. *ACS Chem. Biol.* **2019,** *14* (1), 106-117.

14.	Konkankit, C. C.; Vaughn, B. A.; MacMillan, S. N.; Boros, E.; Wilson, J. J., Combinatorial synthesis to identify a potent, necrosis-inducing rhenium anticancer agent. *Inorg. Chem.* **2019,** *58* (6), 3895-3909.

15.	M. Galanski, V. B. A., M. A. Jakupec and B. K. Keppler, Recent developments in the field of tumor-inhibiting metal complexes. *Curr. Pharma. Design* **2003,** *9*, 2078-2089.

16.	Parveen, S.; Arjmand, F.; Tabassum, S., Development and future prospects of selective organometallic compounds as anticancer drug candidates exhibiting novel modes of action. *Eur. J. Med. Chem.* **2019,** *175*, 269-286.

17.	Blanck, S.; Geisselbrecht, Y.; Kräling, K.; Middel, S.; Mietke, T.; Harms, K.; Essen, L.-O.; Meggers, E., Bioactive cyclometalated phthalimides: Design, synthesis and kinase inhibition. *Dalton Trans.* **2012,** *41*, 9337-9348.

18.	Batchelor, L. K.; Dyson, P. J., Extrapolating the fragment-based approach to inorganic drug discovery. *Trends in Chemistry*.

19.   Meola, G.; Braband, H.; Schmutz, P.; Benz, M.; Spingler, B.; Alberto, R., Bis-arene complexes [Re(η^6-arene)$_2$]$^+$ as highly stable bioorganometallic scaffolds. *Inorg. Chem.* **2016,** *55* (21), 11131-11139.

20.   Peacock, A. F. A.; Habtemariam, A.; Fernández, R.; Walland, V.; Fabbiani, F. P. A.; Parsons, S.; Aird, R. E.; Jodrell, D. I.; Sadler, P. J., Tuning the reactivity of osmium(II) and ruthenium(II) arene complexes under physiological conditions. *J. Am. Chem. Soc.* **2006,** *128* (5), 1739-1748.

21.   Zhang, P.; Sadler, P. J., Advances in the design of organometallic anticancer complexes. *J. Organomet. Chem.* **2017,** *839*, 5-14.

22.   Trifonova, E.; Perekalin, D.; Lyssenko, K.; R. Kudinov, A., Synthesis and structures of cationic bis(arene)rhenium complexes. *J. Organomet. Chem.* **2013,** *727*, 60–63.

23.   Klenc, J.; Lipowska, M.; Abhayawardhana, P. L.; Taylor, A. T.; Marzilli, L. G., Structure and Properties of fac-[ReI(CO)3(NTA)]2– (NTA3– = Trianion of Nitrilotriacetic Acid) and fac-[ReI(CO)3(L)]n– Analogues Useful for Assessing the Excellent Renal Clearance of the fac-[99mTcI(CO)3(NTA)]2– Diagnostic Renal Agent. *Inorg. Chem.* **2015,** *54* (13), 6281-6290.

24.   Cohen, S. M., A Bioinorganic Approach to Fragment-Based Drug Discovery Targeting Metalloenzymes. *Acc Chem Res* **2017,** *50* (8), 2007-2016.

25.   Baramee, A.; Coppin, A.; Mortuaire, M.; Pelinski, L.; Tomavo, S.; Brocard, J., Synthesis and in vitro activities of ferrocenic aminohydroxynaphthoquinones against *Toxoplasma gondii* and *Plasmodium falciparum. Bioorg. Med. Chem.* **2006,** *14* (5), 1294-1302.

26.   Siegmund, D. Rhenium(I) N-heterocyclic carbene complexes as novel organometallic antibacterial agents. Ruhr University Bochum, 2018.

27.   Heldt, J.-M.; Fischer-Durand, N.; Salmain, M.; Vessières, A.; Jaouen, G., Preparation and characterization of poly(amidoamine) dendrimers functionalized with a rhenium carbonyl complex and PEG as new IR probes for carbonyl metallo immunoassay. *J. Organomet. Chem.* **2004,** *689* (25), 4775-4782.

28.   Knopf, K. M.; Murphy, B. L.; MacMillan, S. N.; Baskin, J. M.; Barr, M. P.; Boros, E.; Wilson, J. J., In vitro anticancer activity and in vivo biodistribution of rhenium(I) tricarbonyl aqua complexes. *J. Am. Chem. Soc.* **2017,** *139* (40), 14302-14314.

29.   Kolobova, N. E.; Valueva, Z. P.; Solodova, M. Y., Synthesis of formylcyclopentadienyltricarbonylrhenium and some of its properties. *Bull. Acad. Sci. USSR, Div. Chem. Sci.* **1980,** *29* (10), 1701-1705.

30.   Congreve, M.; Carr, R. A. E.; Murray, C. W.; Jhoti, H., A 'rule of three' for fragment-based lead discovery? *Drug Discov. Today* **2003,** *8*, 876-877.

31.   Wong, E. L.-M.; Sun, R. W.-Y.; Chung, N. P. Y.; Lin, C.-L. S.; Zhu, N.; Che, C.-M., A mixed-valent ruthenium−oxo oxalato cluster $Na_7[Ru_4(\mu^3\text{-}O)_4(C_2O_4)_6]$ with potent anti-HIV activities. *J. Am. Chem. Soc.* **2006,** *128* (15), 4938-4939.

32.   Singh, A.; Lumb, I.; Mehra, V.; Kumar, V., Ferrocene-appended pharmacophores: An exciting approach for modulating the biological potential of organic scaffolds. *Dalton Trans.* **2019,** *48* (9), 2840-2860.

33.   Policar, C.; Waern, J. B.; Plamont, M.-A.; Clède, S.; Mayet, C.; Prazeres, R.; Ortega, J.-M.; Vessières, A.; Dazzi, A., Subcellular IR imaging of a metal–carbonyl moiety using photothermally induced resonance. *Angew. Chem. Int. Ed.* **2011,** *50* (4), 860-864.

34.   Shultz, M. D., Two decades under the influence of the rule of five and the changing properties of approved oral drugs. *J. Med. Chem.* **2019,** *62* (4), 1701-1714.

35.   Lipinski, C. A., Lead- and drug-like compounds: The rule-of-five revolution. *Drug Discov. Today: Tech.* **2004,** *1*, 337-341.

36.   *Molecular Operating Environment (MOE)*, Chemical Computing Group ULC: Montreal QC, Canada, 2019.0101.

37.   Morley, A. D.; Pugliese, A.; Birchall, K.; Bower, J.; Brennan, P.; Brown, N.; Chapman, T.; Drysdale, M.; Gilbert, I. H.; Hoelder, S.; Jordan, A.; Ley, S. V.; Merritt, A.; Miller, D.; Swarbrick, M. E.; Wyatt, P. G., Fragment-based hit identification: Thinking in 3D. *Drug Discov. Today* **2013,** *18* (23), 1221-1227.

38.   Grzegorczyk, M.; Kapturkiewicz, A.; Nowacki, J.; Trojanowska, A., Center-symmetric dimeric Re(CO)3+ complexes with Schiff base derivatives of salicylic aldehyde. *Inorg. Chem. Commun.* **2011,** *14* (11), 1773-1776.

39.   Patra, M.; Joshi, T.; Pierroz, V.; Ingram, K.; Kaiser, M.; Ferrari, S.; Spingler, B.; Keiser, J.; Gasser, G., DMSO-Mediated Ligand Dissociation: Renaissance for Biological Activity of N-Heterocyclic-[Ru(η6-arene)Cl2] Drug Candidates. *Chem. - Eur. J.* **2013,** *19* (44), 14768-14772.

40.   Chen, A. Y.; Adamek, R. N.; Dick, B. L.; Credille, C. V.; Morrison, C. N.; Cohen, S. M., Targeting metalloenzymes for therapeutic intervention. *Chem. Rev.* **2019,** *119* (2), 1323-1455.

41.   Sidera, K.; Patsavoudi, E., HSP90 inhibitors: current development and potential in cancer therapy. *Recent Pat. Anticancer Drug. Discov.* **2014,** *9* (1), 1-20.

42.   Credille, C. V.; Morrison, C. N.; Stokes, R. W.; Dick, B. L.; Feng, Y.; Sun, J.; Chen, Y.; Cohen, S. M., SAR Exploration of Tight-Binding Inhibitors of Influenza Virus PA Endonuclease. *J. Med. Chem.* **2019,** *62*, 9438-9449.

43.   Dias, A.; Bouvier, D.; Crépin, T.; McCarthy, A. A.; Hart, D. J.; Baudin, F.; Cusack, S.; Ruigrok, R. W. H., The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* **2009,** *458*, 914-918.

44.   Credille, C. V.; Dick, B. L.; Morrison, C. N.; Stokes, R. W.; Adamek, R. N.; Wu, N. C.; Wilson, I. A.; Cohen, S. M., Structure-activity relationships in metal-binding pharmacophores for influenza endonuclease. *J. Med. Chem.* **2018,** *61* (22), 10206-10217.

45.   Heo, Y.-A., Baloxavir: First global approval. *Drugs* **2018,** *78* (6), 693-697.

46.   Nordmann, P.; Poirel, L.; Walsh, T. R.; Livermore, D. M., The emerging NDM carbapenemases. *Trends Microbiol.* **2011,** *19* (12), 588-95.

47.   Sun, Z.; Hu, L.; Sankaran, B.; Prasad, B. V. V.; Palzkill, T., Differential active site requirements for NDM-1 β-lactamase hydrolysis of carbapenem versus penicillin and cephalosporin antibiotics. *Nature Communications* **2018,** *9* (1), 4524.

48.   Wegele, H.; Muller, L.; Buchner, J., Hsp70 and Hsp90: A relay team for protein folding. *Rev. Physiol. Biochem. Pharmacol.* **2004,** *151*, 1-44.

49.   Butler, L. M.; Ferraldeschi, R.; Armstrong, H. K.; Centenera, M. M.; Workman, P., Maximizing the therapeutic potential of HSP90 inhibitors. *Mol. Cancer Res.* **2015,** *13* (11), 1445-1451.

50.   Stebbins, C. E.; Russo, A. A.; Schneider, C.; Rosen, N.; Hartl, F. U.; Pavletich, N. P., Crystal Structure of an Hsp90–Geldanamycin Complex: Targeting of a Protein Chaperone by an Antitumor Agent. *Cell* **1997,** *89* (2), 239-250.

51.   Chen, A. Y.; Thomas, P. W.; Cheng, Z.; Xu, N. Y.; Tierney, D. L.; Crowder, M. W.; Fast, W.; Cohen, S. M., Investigation of dipicolinic acid isosteres for the inhibition of metallo-β-lactamases. *ChemMedChem* **2019,** *14* (0).

52.   Hajduk, P. J.; Huth, J. R.; Fesik, S. W., Druggability Indices for Protein Targets Derived from NMR-Based Screening Data. *J. Med. Chem.* **2005,** *48* (7), 2518-2525.

53.   Ansgar, S.; Simon, R.; Andreas, M.; Wolfgang, J.; Paul, S.; Edgar, J., Library Design for Fragment Based Screening. *Curr. Top. Med. Chem.* **2005,** *5* (8), 751-762.

54.   Morrison, C. N.; Prosser, K. E.; Stokes, R. W.; Cordes, A.; Metzler-Nolte, N.; Cohen, S. M., Correction: Expanding medicinal chemistry into 3D space: metallofragments as 3D scaffolds for fragment-based drug discovery. *Chem. Sci.* **2022,** *13* (32), 9450-9452.

55.   Ali, L. H.; Cox, A.; Kemp, T. J., Photochemistry of ferrocenyl ketones and acids in dimethyl sulphoxide and related solvents. *J. Chem. Soc., Dalton Trans.* **1973,** (14), 1468-1475.

56.   Blagg, J.; Workman, P., Choose and Use Your Chemical Probe Wisely to Explore Cancer Biology. *Cancer cell* **2017,** *32* (1), 9-25.

57.   Maschke, M.; Alborzinia, H.; Lieb, M.; Wölfl, S.; Metzler-Nolte, N., Structure–Activity Relationship of Trifluoromethyl-Containing Metallocenes: Electrochemistry, Lipophilicity, Cytotoxicity, and ROS Production. *ChemMedChem* **2014,** *9* (6), 1188-1194.

58.   Vilar, S.; Cozza, G.; Moro, S., Medicinal chemistry and the molecular operating environment (MOE): application of QSAR and molecular docking to drug discovery. *Curr Top Med Chem* **2008,** *8* (18), 1555-72.

59.   Morrison, C. N.; Prosser, K. E.; Stokes, R. W.; Cordes, A. L.; Metzler-Nolte, N.; Cohen, S. M., Expanding Medicinal Chemistry into 3D Space: Metallofragments as 3D Scaffolds for Fragment-Based Drug Discovery. *Chem. Sci.* **2020,** *11*, 1216-1225.

60.   Vanicek, S.; Kopacka, H.; Wurst, K.; Müller, T.; Schottenberger, H.; Bildstein, B., Chemoselective, practical synthesis of cobaltocenium carboxylic acid hexafluorophosphate. *Organometallics* **2014,** *33* (5), 1152-1156.

61.   Chong, D.; Laws, D. R.; Nafady, A.; Costa, P. J.; Rheingold, A. L.; Calhorda, M. J.; Geiger, W. E., [Re(η^5-C$_5$H$_5$)(CO)$_3$]$^+$ Family of 17-electron compounds: Monomer/dimer equilibria and other reactions. *J. Am. Chem. Soc.* **2008,** *130* (8), 2692-2703.

62.   Cais, M.; Narkis, N., Organometallic studies VIII. Aminocyclopentadienylmanganese tricarbonyl and some related compounds. Application of the schmidt reaction to metallocenyl ketones. *J. Organomet. Chem.* **1965,** *3* (3), 188-199.

63.   Kolobova, N. E.; Valueva, Z. P.; Solodova, M. Y., Synthesis of formylcyclopentadienyltricarbonylrhenium and some of its properties. *Bulletin of the Academy of Sciences of the USSR, Division of chemical science* **1980,** *29* (10), 1701-1705.

64.   Liu, C.-Y.; Chen, D.-Y.; Lee, G.-H.; Peng, S.-M.; Liu, S.-T., Synthesis of cyclic diamino-substituted metal carbene complexes. *Organometallics* **1996,** *15* (3), 1055-1061.

65.   Sheldrick, G., Crystal structure refinement with SHELXL. *Acta Cryst. C* **2015,** *71* (1), 3-8.

66.   Credille, C. V.; Chen, Y.; Cohen, S. M., Fragment-based identification of influenza endonuclease inhibitors. *J. Med. Chem.* **2016,** *59* (13), 6444-6454.

67.   Allen, F. H., The Cambridge Structural Database: A quarter of a million crystal structures and rising. *Acta Cryst.* **2002,** *B58*, 380-388.

68.   Bruno, I. J. C., J. C.; Edgington, P. R.; Kessler, M.; Macrae, C. F.; McCabe, P.; Pearson, J.; Taylor, R. , New software for searching the Cambridge Structural Database and visualizing crystal structures. *Acta Cryst.* **2002,** *B58*, 389-397.

Chapter 6: Broadly Accessing 3D Chemical Space with Metal Complexes

6.1 Introduction

As described in Chapters 1, 4, and 5, 3-dimensionality is an important consideration in library design, as highly 3D molecules have favorable properties in drug discovery efforts.[1-2] As previously noted, various challenges associated with synthesizing topologically diverse molecules has hindered progress towards the development of highly 3D libraries that thoroughly represent chemical space. While many research groups (including our own) have attempted to assemble highly 3-dimensional libraries, these efforts have not resulted in a library that completely represents 3D space.[3-6] As demonstrated in Chapters 4 and 5, metal complexes display topological diversity and have been successfully used in FBDD applications. While the exploration of the topologies of additional metal complexes that have been evaluated in a drug discovery context would be beneficial, no database of such metal complexes exists.[7] Additionally, though many metal complexes have been evaluated for their bioactivity, many more metal complexes with diverse properties and topologies remain unexplored.

Here, we propose that data from the Cambridge Structural Database (CSD, v5.45), provided by the Cambridge Crystallographic Data Centre (CCDC), can be utilized to develop a library of topologically diverse metal complexes suitable for FBDD. The CSD represents a suitable starting point for library development for several reasons. First, every entry in the CSD has been crystallographically characterized, enabling access to structural data for topological analysis and VS. Second, the large number of compounds (>1.18 million structures) in the CSD provides complete and thorough topological coverage of 3D space (Figure 6.1). Additionally, each CSD entry is associated with a literature reference, which can facilitate text mining to deduce the potential of the molecule to be included in a fragment library. This text mining will strive to identify molecules that have been experimentally evaluated as inhibitors and will also attempt to

discover molecules that may be useful in this context, but which have not been previously considered for such an application. By first manually applying selective drug-like filters to CSD entries and afterwards refining the results with this text mining approach, a library of stable, topologically diverse mFs can be assembled. This library will consist of molecules that provide a thorough representation of topological space and are stable under biological conditions. Additionally, molecules with well-defined exchangeable ligands may also be identified, which can be assembled into an additional 3D library of reactive mFs,[8] in which the fragments can form coordinative covalent bonds with the target upon ligand exchange.
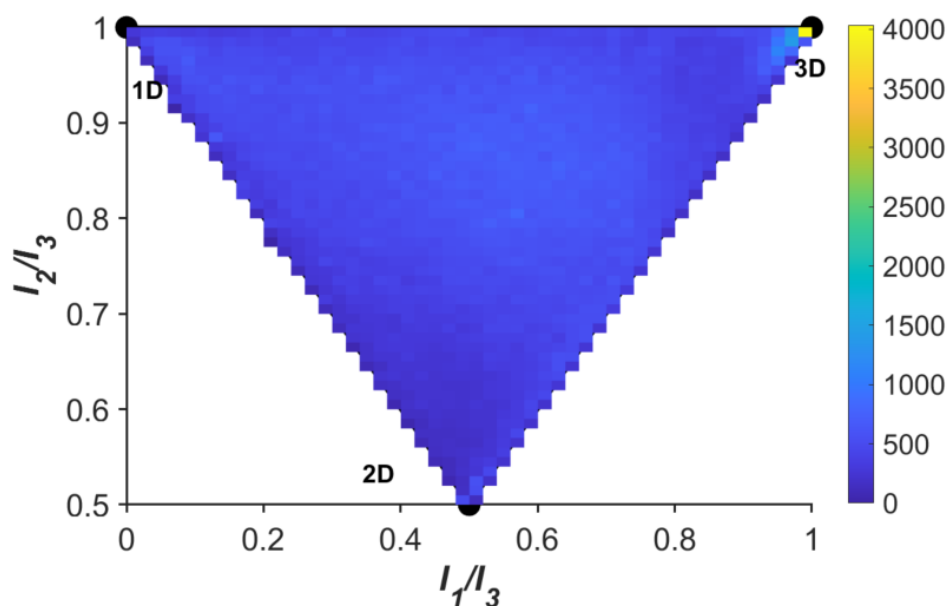

Figure 6.1. PMI representation of >600,000 metal complexes in the CCDC, demonstrating broad topological coverage of 3D space.

6.2 Data Reduction of Structures in the CSD

While the CSD has a large amount of structural data, it contains many compounds that are reactive, unstable, or are otherwise unsuitable for FBDD applications. Therefore, several filters were applied to CSD entries in order to broadly exclude molecules that would likely be

problematic. Using version 5.45 of the CCDC, including the March 2022 update, CCDC entries were sorted using Conquest (2022.01).[9] All metal-containing structures that had 3D coordinates determined (621,322 entries) were selected. Because there are typically solvents, salts, or other molecules present in single crystal structures, the largest molecule from each entry was exported in .SDfile format and loaded into MOE (2020.901). Once in MOE, each entry was prepared, and parameters such as PMI, the number of atoms, the number of carbon atoms, molecular volume, the number of violations to Lipinski's rules, and Simplified Molecular Input Line System (SMILES) strings were calculated. These values were used as parameters to further filter and reduce the dataset. Organic molecules that were exported as the largest molecules and structures weighing less than 200 Da, or which had three atoms or less were removed, resulting in 609,860 structures. Next, in order to filter out molecules too large to be useful in general drug discovery, 364,799 molecules with molecular volumes greater than 500 Å^3 were removed, reducing the data set to 242,285 molecules. Non-unique structures were removed according to CCDC identifiers, resulting in the removal of an additional 23,748 structures. Next, compounds with no carbon atoms were removed, further reducing the dataset to 189,232 structures. To help select for molecules amenable for drug discovery, molecules which had more than one violation to Lipinski's rules were removed, resulting in 172,875 entries. To remove entries with potentially problematic or toxic atoms, structures containing mercury, cadmium, tin, lead, arsenic, antimony, lanthanum, actinium, polonium, astatine, cerium, thorium, neodymium, uranium, promethium, neptunium, samarium, europium, americium, terbium, berkelium, dysprosium, californium, holmium, einsteinium, erbium, fermium, thulium, mendelevium, ytterbium, nobelium, and lutetium were removed. This resulted in the removal of 23,688 entries, corresponding to a working dataset of 149,187 compounds. Next, the few entries that had keywords indicating stability from the CCDC

remarks were analyzed. Entries with words or phrases including, 'air sensitive', 'light sensitive', 'moist', or 'unstable' were removed, resulting in the removal of 3,458 entries.

Next, to prepare for text mining, the dataset of 145,619 entries was matched with its corresponding literature citation. After doing so, it became apparent that a large percentage of the entries came from a relatively small number of journals. Several journals from the American Chemical Society (ACS) were among the journals with the most entries. Initially, data was only solicited from the ACS to expedite text analysis, as only one publisher would need to grant permission to mine the data, and because the data would be provided in a uniform format. For text analysis, 7,839 articles from *Organometallics*, 7,923 articles from *Inorganic Chemistry*, 3,460 articles from *Journal of the American Chemical Society*, 274 articles from *Journal of Organic Chemistry*, 236 articles from *Organic Letters*, and 111 articles from *Journal of Medicinal Chemistry* were selected, representing a total dataset of 33,685 compounds from 19,843 articles. Permission was requested and received from the American Chemical Society to text mine the data, which is ongoing. This data reduction workflow is outlined in Figure 6.2.
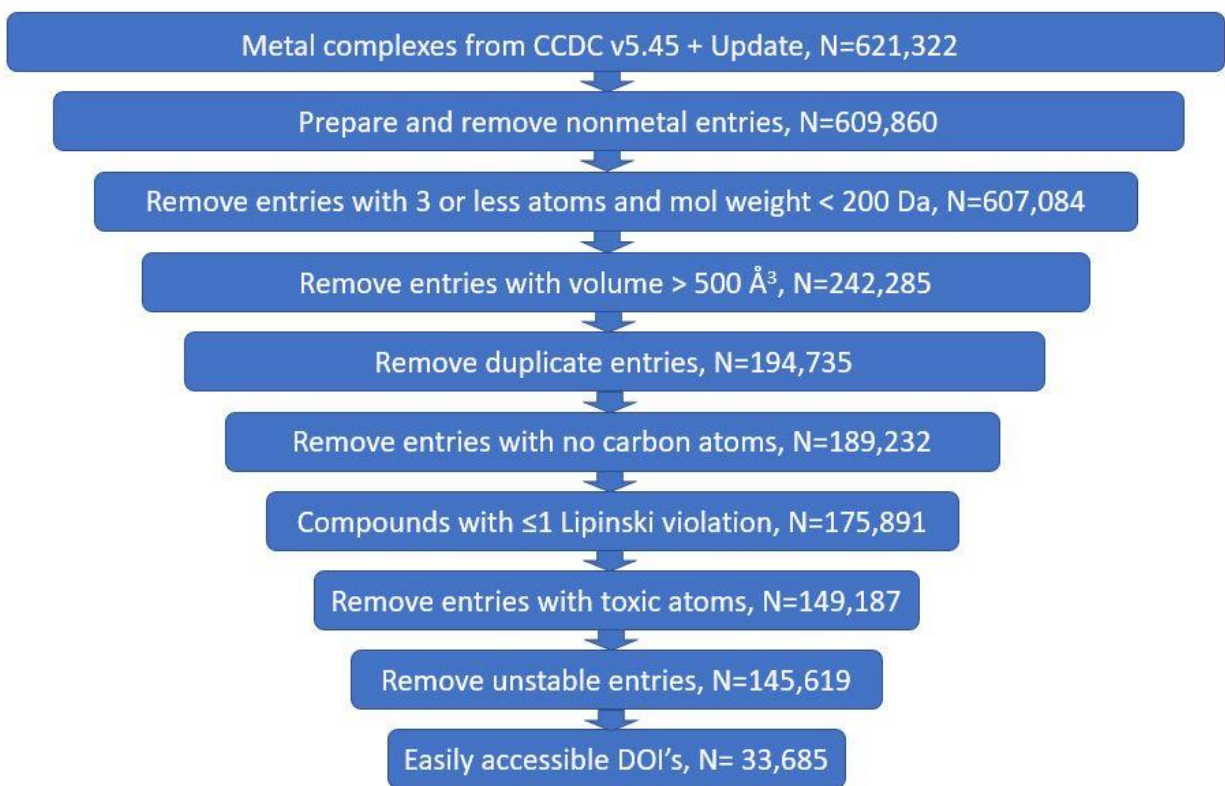
Figure 6.2. Workflow demonstrating the reduction of >1.8 million compounds in the CSD to 33,685 organometallic structures that may be useful in a 3D screening library.

6.3 3D Analysis of Reduced Dataset

As a proof of concept prior to text mining (which will aim to identify experimentally useful molecules) an analysis of different subsets of the reduced data set was performed. First, as a point of reference, the DrugBank[10] (which was analyzed in depth in Chapter 4) was analyzed. This dataset comprises of 8,532 entries, with an average molecular weight of 336.0 g/mol. Additionally, it possesses an average $I_1/I_3 = 0.29$, and an average $I_2/I_3 = 0.84$, corresponding to an average overall 3D score of 1.14. By contrast, all metal complexes in the CSD (621,322 entries) have an average molecular volume of 843.0 Å^3 and an average weight of 1228.5 g/mol. The average $I_1/I_3 = 0.51$, and the average $I_2/I_3 = 0.85$, with an overall average 3D score of 1.36. While this 3D score is much higher than that of the DrugBank, the average molecular volume and molecular weight of the

compounds in this dataset are too high to be broadly useful in fragment screening. However, upon

filtering the dataset to include molecules with more desirable properties (145,619 compounds), the

average volume was reduced to 364 Å^3, with the average weight reduced to 540.0 g/mol, in good

agreement with the metrics discussed in Chapter 5. Interestingly, despite reducing the average

molecular volume and molecular weight, the overall 3D score was unaffected, with an average

$I_1/I_3 = 0.48$, and an average $I_2/I_3 = 0.84$, corresponding to an average overall 3D score of 1.32.

Finally, once the data was filtered to select for the referenced journals above, the resulting dataset

of 33,685 compounds had a combined average 3D score of 1.36. This score of 1.36 is equivalent

to the average 3D score of all metal complexes in the CSD prior to filtering. Additionally, this 3D

score demonstrates an impressive improvement in three-dimensionality when compared to the

average 3D score of the DrugBank (3D score = 1.14, Figure 6.3). Several representative examples

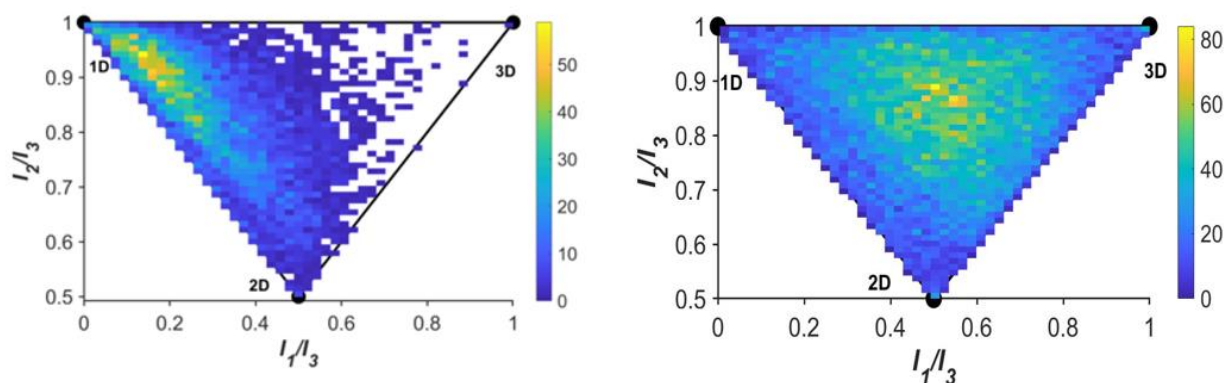of molecules identified using the approach are shown in Figure 6.4.

Figure 6.3. PMI plots of the DrugBank (n = 8,532, *left*) and the reduced dataset (n = 33,685, *right*). The reduced dataset demonstrates a thorough and complete coverage of topological space, while the DrugBank is unable to populate highly 3D regions of the PMI plot.
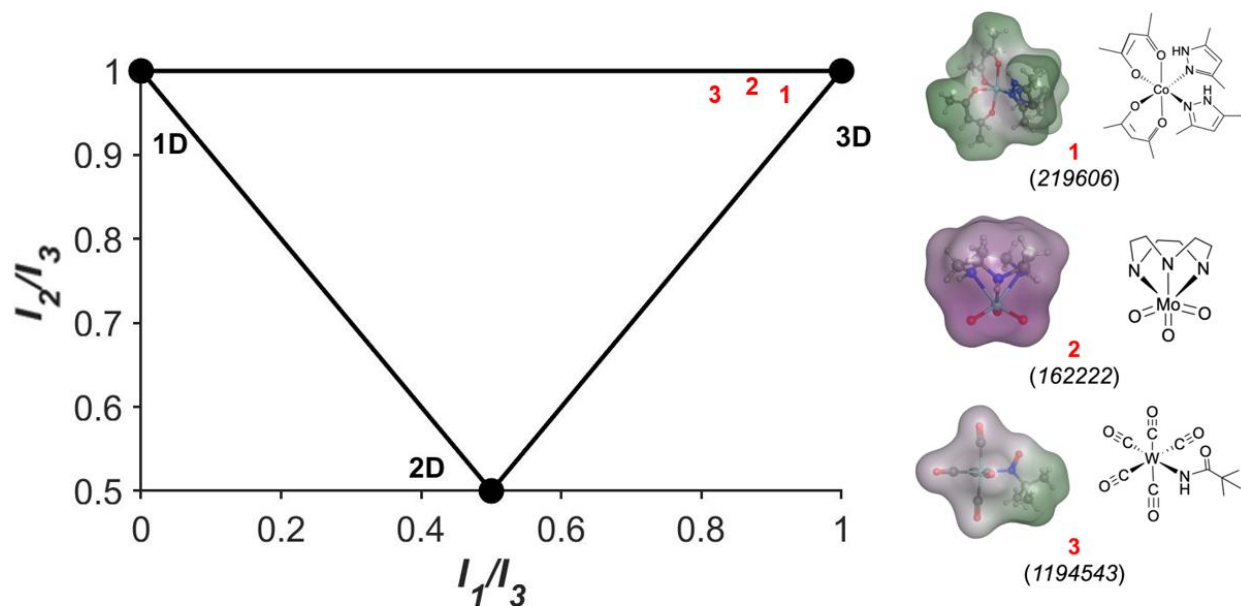
Figure 6.4. Chemical and single crystal X-ray structure (ball-and-stick) of three, select mF candidates showing the molecular surface colored by lipophilicity. Hydrophilic and lipophilic regions are shown in pink and green, respectively. The location of these mF candidates in the PMI plot is highlighted by red numbers. CSD deposition numbers are provided for each mF.

6.4 Future Work Towards the Assembly of a Representative 3D Library

Moving forward in this effort, the 33,685 compounds identified will be further refined to select for molecules that can be included in a topologically diverse fragment library that systematically represents all areas of chemical space. In collaboration with Prof. Jihan Kim, we plan to text mine the corresponding publication of each CSD entry for keywords relevant to the ability of the molecule to be stable in an aqueous biological system. Because many of the molecules in this dataset were selected from publications and applications that were not considering fragment screening, it is likely that their stability in aqueous media may not be explicitly stated or evaluated. We recently developed a prototype that was able to successfully identify both unstable molecules that will be excluded from, and potentially useful molecules that may be included in the final mF library from a subset of CSD entries. A more robust text mining program will be applied on a large scale. This approach will aim to exclude obviously unstable

molecules based on their descriptions in the text and will seek to identify potentially useful compounds and score them according to their inferred stabilities in aqueous media. Next, topological diversity will be used as a criterion for library inclusion to ensure that the final libraries provide broad and thorough coverage of topological space (Figure 6.5). Using this method, virtual libraries will be created that can be screened against protein targets of interest. Additional filtering will be applied to account for synthetic accessibility and suitability for derivatization, which are important requirements for library assembly and HTL development. Accordingly, fragments that can be readily synthesized in 1-3 synthetic steps will be prioritized. Ultimately, this final data reduction step will result in reasonably sized libraries for chemical synthesis and experimental screening. As the mFs are synthesized to assemble the libraries, hits from VS that occupy compelling regions of chemical space will be prioritized.
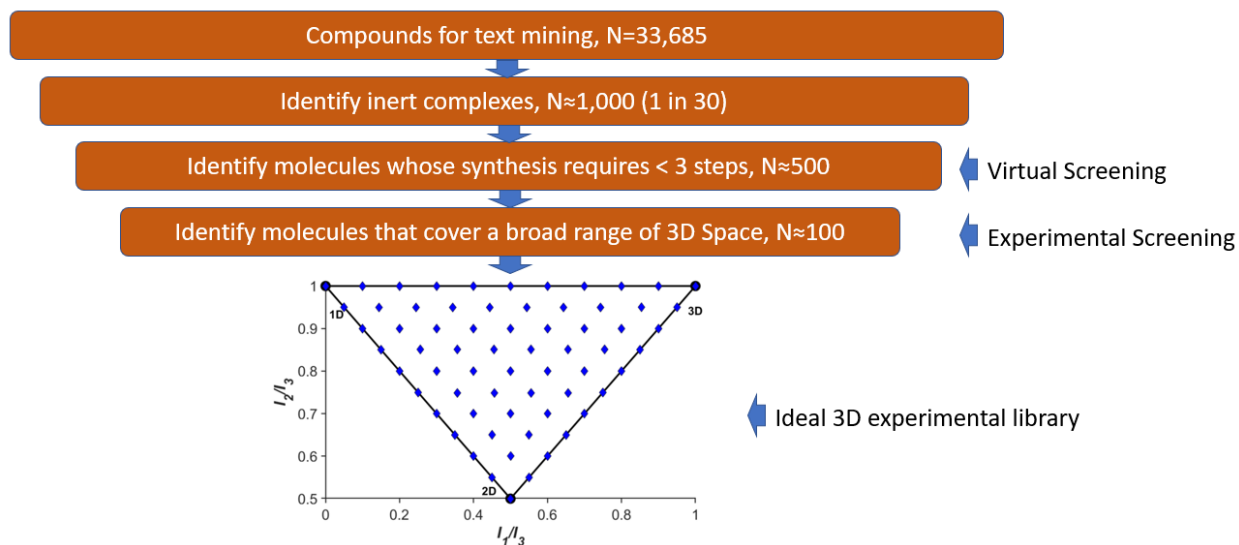

Figure 6.5. Workflow demonstrating the remaining steps to produce a highly topological 3D library of mFs.

6.5 Conclusions

Highly 3D compounds are desirable in drug discovery efforts, though their implementation in fragment libraries has proven challenging.[11] While metal complexes have been shown to be highly three dimensional,[12] and have been applied in screening applications,[6] no easily accessible, comprehensive 3D fragment library exists. Here, data from the CSD was collected and reduced in an attempt to identify topologically diverse molecules that may be suitable for drug discovery applications. From more than 1.8 million structures in the CSD, selective filters were applied to identify 145,619 potentially useful compounds. Filters included molecular size and molecular weight, violations to the rule of five, and the removal of potentially toxic atoms. This reduced set of molecules showed improved three-dimensionality when compared with the DrugBank. Next, text mining techniques are being performed to exclude molecules with obvious instability in aqueous media and infer stability where it is not explicitly stated. These efforts are ongoing, and it is hoped that they will enable the creation of virtual mF libraries with diverse topologies. Separate libraries of inert and reactive mFs will be assembled, and VS will be performed. Based on the results of the results of the VS, mFs will be synthesized into libraries that can be evaluated experimentally. These libraries will thoroughly represent 3D space and will lead to the identification of useful inhibitors of a wide range of therapeutically relevant targets.

6.6 Acknowledgements

Chapter 6 contains unpublished material that was written with Jihan Kim and Seth M. Cohen. The dissertation author was the primary author of this chapter. We acknowledge the American Chemical Society for providing articles for text mining.

6.7 References

1.   Galloway, W. R. J. D.; Isidro-Llobet, A.; Spring, D. R., Diversity-Oriented Synthesis As A Tool For The Discovery Of Novel Biologically Active Small Molecules. *Nat. Commun.* **2010,** *1* (1), 80.

2.   Firth, N. C.; Brown, N.; Blagg, J., Plane Of Best Fit: A Novel Method To Characterize The Three-Dimensionality Of Molecules. *J. Chem. Inf. Model.* **2012,** *52* (10), 2516-2525.

3.   Meggers, E., Exploring Biologically Relevant Chemical Space With Metal Complexes. *Curr. Opin. Chem. Biol.* **2007,** *11* (3), 287-292.

4.   Zhang, L.; Carroll, P.; Meggers, E., Ruthenium Complexes As Protein Kinase Inhibitors. *Org. Lett.* **2004,** *6* (4), 521-523.

5.   Feng, L.; Geisselbrecht, Y.; Blanck, S.; Wilbuer, A.; Atilla-Gokcumen, G. E.; Filippakopoulos, P.; Kräling, K.; Celik, M. A.; Harms, K.; Maksimoska, J.; Marmorstein, R.; Frenking, G.; Knapp, S.; Essen, L.-O.; Meggers, E., Structurally Sophisticated Octahedral Metal Complexes As Highly Selective Protein Kinase Inhibitors. *J. Am. Chem. Soc.* **2011,** *133* (15), 5976-5986.

6.   Morrison, C. N.; Prosser, K. E.; Stokes, R. W.; Cordes, A.; Metzler-Nolte, N.; Cohen, S. M., Expanding Medicinal Chemistry Into 3D space: Metallofragments As 3D Scaffolds For Fragment-Based Drug Discovery. *Chem. Sci.* **2020,** *11* (5), 1216-1225.

7.   Medina-Franco, J. L.; López-López, E.; Andrade, E.; Ruiz-Azuara, L.; Frei, A.; Guan, D.; Zuegg, J.; Blaskovich, M. A. T., Bridging informatics and medicinal inorganic chemistry: Toward a database of metallodrugs and metallodrug candidates. *Drug Discovery Today* **2022,** *27* (5), 1420-1430.

8.   Karges, J.; Kalaj, M.; Gembicky, M.; Cohen, S. M., ReI Tricarbonyl Complexes as Coordinate Covalent Inhibitors for the SARS-CoV-2 Main Cysteine Protease. *Angew. Chem., Int. Ed.* **2021,** *60* (19), 10716-10723.

9.   Bruno, I. J.; Cole, J. C.; Edgington, P. R.; Kessler, M.; Macrae, C. F.; McCabe, P.; Pearson, J.; Taylor, R., New software for searching the Cambridge Structural Database and visualizing crystal structures. *Acta Crystallogr.* **2002,** *B58* (3 Part 1), 389-397.

10.   Wishart, D. S.; Knox, C.; Guo, A. C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J., DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* **2006,** *34* (suppl_1), D668-D672.

11.   Prosser, K. E.; Stokes, R. W.; Cohen, S. M., Evaluation Of 3-Dimensionality In Approved And Experimental Drug Space. *ACS Med. Chem. Lett.* **2020,** *11* (6), 1292-1298.

12.   Meggers, E., Targeting Proteins With Metal Complexes. *Chem. Commun.* **2009,** (9), 1001-1010.