

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Protein-Metal-Organic Frameworks and other Hydroxamate-based Crystalline Lattices

Permalink

<https://escholarship.org/uc/item/2dm8x7q4>

Author

Bailey, Jake Brian

Publication Date

2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Protein-Metal-Organic Frameworks and other Hydroxamate-based Crystalline Lattices

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Chemistry

by

Jake Brian Bailey

Committee in charge:

Professor F. Akif Tezcan, Chair
Professor Neal Devaraj
Professor Kamil Godula
Professor Thomas Hermann
Professor Andrea Tao

2020

©

Jake Brian Bailey, 2020

All rights reserved.

The Dissertation of Jake Brian Bailey is approved, and it is acceptable in quality and form for
publication on microfilm and electronically:

Chair

University of California San Diego

2020

Table of Contents

Signature Page	iii
Table of Contents.....	iv
List of Abbreviations.....	vii
List of Figures	x
List of Tables	xvii
Acknowledgments.....	xviii
Vita	xxi
Abstract of the Dissertation	xxiii
Chapter 1. Supramolecular protein self-assembly	1
1.1. Extended 1-, 2-, and 3D protein assemblies found in Nature	2
1.2. Assembly strategies for 1-, 2-, and 3D protein materials.....	6
1.2.1. Charge-based assembly of protein networks.....	7
1.2.2. Protein assembly from genetically fused constructs	10
1.2.3. Assemblies constructed from noncovalent protein-protein interactions.....	12
1.2.4. Disulfide mediated assembly of protein lattices	14
1.2.5. Metal mediated protein self-assembly	17
1.2.6 Protein assemblies based on recognition elements.....	20
1.2.7. Protein-self assembly using multiple interaction motifs.....	23
1.3. Developing a modular methodology for protein self-assembly	26
1.4. References	28
Chapter 2. Design, synthesis, and characterization of the first protein-MOF	34
2.1. Introduction.....	34
2.2. Results and discussion	36
2.2.1. Selection of ferritin as the protein node	36
2.2.2. Design and characterization of a metal-binding site at the C_4 axes of ferritin	38
2.2.3. Design and characterization of a metal-binding site at the C_3 axes of ferritin	40
2.2.4. Selection of a suitable bridging ligand for protein-MOF construction	43
2.2.5. Synthesis and characterization of the first protein-MOF	45
2.3. Methods.....	50
2.3.1. General considerations	50
2.3.2. Site-directed mutagenesis.....	50
2.3.3. Expression, purification, and characterization of ferritin variants	51
2.3.4. Synthesis of N^1, N^4 -dihydroxyterephthalamide (p -H ₂ b ₂ dh)	53
2.3.5. DLS measurements of ferritin assembly conditions	53
2.3.6. Crystallization of the Zn-ferritin node	54

2.3.7. Crystallization of <i>p</i> -bdh-Zn- ⁸⁶ Aferritin and <i>p</i> -bdh-Zn-ferritin	54
2.3.8. sc-XRD data collection and structure determination	55
2.3.9. Iron loading/mineralization in <i>p</i> -bdh-Zn-ferritin crystals	56
2.4. Acknowledgments	57
2.5. References	57
 Chapter 3. Synthetic modularity of ferritin-MOFs	 62
3.1. Introduction.....	62
3.2. Results and discussion	64
3.2.1. Modularity in the ferritin node	64
3.2.2. Modularity in the bridging ligand.....	66
3.2.3. General considerations for the synthesis of ferritin-MOFs	69
3.2.4. Ligands that failed to form ferritin-MOFs	71
3.2.5. Small-angle X-ray scattering as a tool for ferritin-MOF characterization <i>in situ</i>	73
3.2.6. Structural determination of ferritin-MOFs through sc-XRD.....	86
3.3. Methods.....	97
3.3.1. General considerations	97
3.3.2. Structural determination of Ni-ferritin and Co-ferritin nodes through sc-XRD	97
3.3.3. Synthesis of the dihydroxamate bridging ligand library	100
3.3.5. Crystallization of ferritin-MOFs	116
3.3.4. Collection and processing of SAXS profiles.....	117
3.3.6. Structural determination of ferritin-MOFs by sc-XRD.....	118
3.4. Acknowledgments	122
3.5. References	122
 Chapter 4. Thermal behavior of ferritin-MOFs.....	 127
4.1. Introduction.....	127
4.2. Results and discussion	130
4.2.1. Assessing the thermal behavior of ferritin-MOFs by SAXS.....	130
4.2.2. Thermal behavior of six ferritin-MOFs	132
4.2.3. Further characterization of the thermal behavior of 2,5-fdh-Ni-ferritin	136
4.2.4. Characterization of 2,5-fdh-Ni-ferritin in solution with sc-XRD	140
4.3. Methods.....	145
4.3.1. General considerations	145
4.3.2. Calculation of the protein surface area and crystal contact area.....	145
4.3.3. Preparation of ferritin-MOF samples for powder diffraction at SAXS beamlines	145
4.3.4. Temperature-dependent SAXS at synchrotron beamlines.....	146
4.3.5. sc-XRD data collection solution at ambient and elevated temperatures	148
4.4. Acknowledgments	152
4.5. References	152
 Chapter 5. Concluding remarks and future directions.....	 155
5.1. Concluding remarks.....	155
5.2. Future directions.....	157
5.2.1. Design of tetrahedral protein nodes	157

5.2.2. Design of inherently functional bridging ligands.....	164
5.2.3. Development of size-comparable bridging ligands	167
5.3. Methods.....	170
5.3.1. General considerations	170
5.3.2. Design and expression of <i>T. ni</i> variants.....	170
5.3.3. Synthesis of azobenzene precursors.....	173
5.3.4. Expression, purification, and characterization of ³⁰ CROP	174
5.3.5. Expression, purification, and characterization of ⁹⁸ C RhuA.....	175
5.3.6. Synthesis of chelate-based ligands for bioconjugation	176
5.3.7. Bioconjugation of chelate-based ligands to ³⁰ CROP and ⁹⁸ C RhuA.....	183
5.4. Acknowledgments	185
5.5. References	185
Appendix 1. Hydroxamate-based MOFs.....	186
A1.1. Introduction	186
A1.2. Results and discussion.....	188
A1.2.1. Design and synthesis of a chelate-based SBU.....	188
A1.2.2. Synthesis of hydroxamate-based MOFs	191
A1.2.3. Stability of Fe-BPTH	197
A1.2.4. Future research directions and applications of Fe-BPTH and Fe-TPTH	199
A1.3. Methods	200
A1.3.1. General considerations	200
A1.3.2. Synthesis and characterization of the Fe ₄ <i>m</i> -bdh ₆ clusters.....	200
A1.3.3. Synthesis of tetrahydroxamate ligands	201
A1.3.4. Synthesis of hydroxamate-based MOFs	201
A1.3.5. Structural determination of hydroxamate-based MOFs	207
A1.3.6. Assessing the stability of Fe-BPTH in solution	208
A1.4. Acknowledgments	209
A1.5. References.....	209
Appendix 2. NMR Spectra for Synthesized Compounds	212

List of Abbreviations

ALAD	aminolevulinic acid dehydrogenase
AtF	<i>A. fulgidus</i> ferritin
ATP	Adenosine triphosphate
AuNPs	Gold nanoparticles
bcc	body centered cubic
bdc	benzene dicarboxylate
bdh	benzene dihydroxamate
bpdh	biphenyl dihydroxamate
BPTH	biphenyl tetrahydroxamate
CB[8]	cucurbit[8]uril
CCMV	cowpea chlorotic mottle virus
CHES	<i>N</i> -cyclohexyl-2-aminoethanesulfonic acid
ConA	concanavalin A
dgdh	diglyme dihydroxamic acid
DLS	dynamic light scattering
DMP	Dess–Martin periodinane
DNA	deoxyribonucleic acid
DSred	red fluorescent protein
DTT	dithiothreitol
edh	ethene dihydroxamate
fcc	face-centered cubic
fdh	furan dihydroxamate
GST	glutathione S-transferase
HCP1	heme carrier protein 1
HuHF	human heavy-chain ferritin

Hyp	hydroxyproline
IPTG	isopropyl β -D-1-thiogalactopyranoside
kDa	kilodalton
LecA	lectin A
M1	matrix protein 1
mbh	benzene monohydroxamate
MOF	metal-organic framework
MPD	methylpentanediol
MS	mass spectrometry
ndh	napthalene dihydroxamate
NMR	nuclear magnetic resonance
PAMAM	poly(amidoamine)
pdh	pyrrole dihydroxamate
PEG	poly(ethylene glycol)
PEP	pentaerythritol propoxylate 5/4 PO/OH
PMSF	phenylmethylsulfonyl fluoride
podh	pyran-4-one dihydroxamate
pydh	pyridine dihydroxamate
pydh	pyridine dihydroxamate
pzdh	pyrazole dihydroxamate
QPTH	quaterphenyl tetrahydroxamate
RhuA	L-rhamulose 1-phosphate aldolase
RNA	ribonucleic acid
ROP	repressor of primer protein
SAXS	small-angle X-ray scattering
SBA	soybean agglutinin

SBU	secondary building unit
sc-XRD	single-crystal X-ray diffraction
sndh	sulfonated naphthalene dihydroxamate
SP1	Stable Protein 1
tdh	thiophene dihydroxamate
TEM	transmission electron microscopy
TEV	Tobacco Etch Virus nuclear-inclusion-a endopeptidase
TniHF	<i>T. ni</i> light-chain ferritin
TniLF	<i>T. ni</i> heavy-chain ferritin
TPTH	terphenyl tetrahydroxamate
TRAP	trp RNA-binding attenuation protein
Tris	tris(hydroxymethyl)aminomethane
WT	wild-type
xdh	xylene dihydroxamate

List of Figures

Figure 1.1. 1-, 2-, and 3D assemblies in Nature..	5
Figure 1.2. Electrostatic assembly of ferritin and CCMV.....	9
Figure 1.3. 1- and 2D protein assemblies formed through genetic fusions.	11
Figure 1.4. Protein assemblies based on designed noncovalent protein-protein interactions. ...	13
Figure 1.5. Disulfide mediated protein assemblies..	16
Figure 1.6. Metal-mediated protein self-assembly.	19
Figure 1.7. Protein assemblies based on recognition elements.....	22
Figure 1.8. Protein assemblies with multiple interaction motifs.....	25
Figure 2.1. Scheme for metal/ligand-directed self-assembly of ferritin into 3D crystals.....	35
Figure 2.2. Symmetry axes and expected crystal packing of ferritin.	37
Figure 2.3. Analysis of the C_4 axis of ferritin.	39
Figure 2.4. Comparison of the predicted C_3 axis of ferritin with the carbonic anhydrase active site.	41
Figure 2.5. Crystal structure of Zn-ferritin.	41
Figure 2.6. Crystal packing contacts of Zn-ferritin..	42
Figure 2.7. Dynamic light scattering profiles of ferritin-MOF assembly conditions.	44
Figure 2.8. Crystal structure of p -bdh-Zn- ^{86A} ferritin.....	47
Figure 2.9. Crystal packing contacts of p -bdh-Zn- ^{86A} ferritin.	48
Figure 2.10. Light micrograph of p -bdh-Zn-ferritin crystals upon iron loading/mineralization.	49
Figure 2.11. Image of purified ferritin confirmed by SDS-PAGE.	52
Figure 3.1. Ferritin nodes crystallized with Zn^{2+} , Ni^{2+} , and Co^{2+}	65
Figure 3.2. The library of ditopic dihydroxamic acid bridging ligands categorized by the relative geometry of the C-C bond (shown in red).....	68
Figure 3.3. Light micrographs of crystals from fifteen ferritin-MOFs.....	70
Figure 3.4. The impact of DMSO on the tris-His coordination site..	72

Figure 3.5. Comparison of the SAXS profiles of a ligand-mediated body-centered cubic lattice and a Ca-mediated face-centered cubic lattice.	74
Figure 3.6. Diffraction patterns of twenty-one ferritin-MOFs.	75
Figure 3.7. Possible packing geometries for ferritin-MOFs.	78
Figure 3.8. Structural models for the three lattices of observed <i>p</i> -bdh-Ni-ferritin..	78
Figure 3.9. Two possible unit cells, <i>bct</i> (black) and primitive (red), for 2,6-ndh-Zn-ferritin.	80
Figure 3.10. Diffraction patterns of <i>p</i> -H ₂ bdh-based ferritin-MOFs. Crystals formed under precipitant-free conditions are shown in black. Crystals formed with PEP are shown in orange.	81
Figure 3.11. Models depicting the coordination geometry for collinear and bent ligands.	82
Figure 3.12. Diffraction patterns of <i>E</i> -H ₂ edh- and 2,6-H ₂ ndh-based ferritin-MOFs.	83
Figure 3.13. Diffraction patterns of <i>m</i> -H ₂ bdh-, 2,5-H ₂ fdh-, 2,5-H ₂ tdh, and 2,5-H ₂ pdh-based ferritin-MOFs. Crystals formed under precipitant-free conditions are shown in black.	84
Figure 3.14. <i>p</i> -H ₂ xdh-based ferritin-MOF diffraction patterns crystallized under precipitant-free (black) or with PEP (orange).	85
Figure 3.15. sc-XRD structures for a) <i>p</i> -bdh-Zn-ferritin and b) <i>p</i> -xdh-Zn-ferritin.	88
Figure 3.16. A close-up view of the protein-protein interface of <i>p</i> -xdh-Zn-ferritin showing pairs of amino acid residues that may be involved in electrostatic interactions.	89
Figure 3.17. A schematic representation of the <i>p</i> -xdh-Zn-ferritin lattices examined by sc-XRD and SAXS, as viewed along a body diagonal of the unit cell.	89
Figure 3.18. Multiple crystallographic conformations in ferritin-MOFs.	91
Figure 3.19. Crystal structures of <i>m</i> -bdh-(Zn/Ni)- and 2,5-fdh-(Zn/Ni)-ferritin.	94
Figure 3.20. A close-up view of the protein-protein of a single conformation of <i>m</i> -bdh-Zn-ferritin..	95
Figure 4.1. Crystal contact surface area Å ² vs. % of protein surface that forms a crystal contact for >700 proteins.	129
Figure 4.2. Representation of ferritin-MOF decomposition in response to increasing temperature.	131
Figure 4.3. Thermal stability of <i>p</i> -bdh-Zn-ferritin and <i>p</i> -bdh-Ni-ferritin measured by temperature-dependent diffraction studies.	133
Figure 4.4. Thermal stability of six ferritin-MOFs crystallized with 8% PEP measured with temperature-dependent diffraction studies.	135

Figure 4.5. Thermal stability of the Ca ²⁺ -mediated <i>fcc</i> ferritin crystals measured with temperature-dependent diffraction studies.	135
Figure 4.6. Temperature-dependent diffraction patterns of 2,5-fdh-Ni-ferritin with slow stepwise heating (a) or cooling (b)	137
Figure 4.7. The change in unit cell size of the <i>bcc</i> 2,5-fdh-Ni-ferritin is plotted vs. temperature during slow heating (red) and cooling (blue).....	137
Figure 4.8. The low (a) and mid (b) angle temperature-dependent diffraction patterns of 2,5-fdh-Ni-ferritin during cycling are shown.	138
Figure 4.9 Temperature-dependent diffraction patterns of 2,5-fdh-Ni-ferritin crystallized with PEG 300 (a-b) or PEG 4000 (c-d).....	139
Figure 4.10. sc-XRD structure 2,5-fdh-Ni-ferritin collected at 5°C (a), 45 °C (b), and at 5°C (c) after a single heating and cooling cycle.....	142
Figure 4.11. sc-XRD structure 2,5-fdh-Ni-ferritin collected at 45 °C.....	143
Figure 5.1. Possible scaffolds for a tetrahedral protein node. C ₃ axes for the selected 12-mer tetrahedral proteins are shown in blue and red.....	159
Figure 5.2. The hypothetical tris-his binding site on <i>H. halophila</i> dodecin installed through the L9H mutation.	160
Figure 5.3. Structure of <i>T. ni</i> ferritin.....	161
Figure 5.4. Possible lattice symmetries for ferritin-MOFs based on <i>T. Ni</i> variants.	162
Figure 5.5. A potential tris-His metal chelation sites on AfF.....	163
Figure 5.6. The synthetic scheme for the synthesis of a dihydroxamate bridging ligand with a BODIPY spacer.....	164
Figure 5.7. Synthetic schemes for azobenzene-based bridging ligands..	166
Figure 5.8. Cis-trans isomerization leads to a change in lattice spacing for protein-MOFs bridged by azobenzene-based ligands.....	166
Figure 5.9. Protein-based bridging ligands can be created by labeling symmetric proteins with synthetic ligands..	168
Figure 5.10. Possible lattice architectures for heteroprotein-MOFs..	169
Figure A1.1. Hydroxamate-based MOP synthesized from an <i>N</i> -toluyl derivative of <i>m</i> -H ₂ BDH and Ga(III).	189
Figure A1.2. Models of the Δ-Fe ₄ m-bdh ₆ cluster.....	189
Figure A1.3. Crystal structure of Δ-Fe ₄ m-bdh ₆ and Λ-Fe ₄ m-bdh ₆	190

Figure A1.4. a) Scheme for the synthesis of bridging ligands with <i>m</i> -H ₂ bdh head groups. b) Three collinear ligands were synthesized with varying ligand lengths.....	192
Figure A1.5. Powder X-ray diffraction patterns and SEM images for Fe-BPTH (a), Fe-TPTH (b), and Fe-QPTH (c).	193
Figure A1.6. A comparison of powder X-ray diffraction patterns for Fe-BPTH and Fe-TPTH...	194
Figure A1.7. Crystal structure of Fe-BPTH.	195
Figure A1.8. Crystal structure of Fe-TPTH.	196
Figure A1.9. Powder x-ray diffraction of Fe-BPTH after soaking in solutions with different pH (a) or in different solvents (b).....	198
Figure A2.1. ¹ H NMR spectrum of <i>N</i> ¹ , <i>N</i> ⁴ -dihydroxyterephthalamide (<i>p</i> -H ₂ bdh) in DMSO-d ₆	212
Figure A2.2. ¹³ C NMR spectrum of <i>N</i> ¹ , <i>N</i> ⁴ -dihydroxyterephthalamide (<i>p</i> -H ₂ bdh) in DMSO-d ₆	213
Figure A2.3. ¹ H NMR spectrum of <i>N</i> ¹ , <i>N</i> ³ -dihydroxyisophthalamide (<i>m</i> -H ₂ bdh) in DMSO-d ₆	214
Figure A2.4. ¹³ C NMR spectrum of <i>N</i> ¹ , <i>N</i> ³ -dihydroxyisophthalamide (<i>m</i> -H ₂ bdh) in DMSO-d ₆	215
Figure A2.5. ¹ H NMR spectrum of <i>N</i> ² , <i>N</i> ⁶ -dihydroxynaphthalene-2,6-dicarboxamide (H ₂ ndh) in DMSO-d ₆	216
Figure A2.6. ¹³ C NMR spectrum of <i>N</i> ² , <i>N</i> ⁶ -dihydroxynaphthalene-2,6-dicarboxamide (H ₂ ndh) in DMSO-d ₆	217
Figure A2.7. ¹ H NMR spectrum of <i>N</i> ² , <i>N</i> ⁶ -bis(benzyloxy)furan-2,5-dicarboxamide (2,5-Bn ₂ fdh) in DMSO-d ₆	218
Figure A2.8. ¹ H NMR spectrum of <i>N</i> ² , <i>N</i> ⁶ -dihydroxyfuran-2,5-dicarboxamide (2,5-H ₂ fdh) in DMSO-d ₆	219
Figure A2.9. ¹ H NMR spectrum of 2,2'-(1,4-phenylene)bis(<i>N</i> -(benzyloxy)acetamide) (<i>p</i> -Bn ₂ xdh) in DMSO-d ₆	220
Figure A2.10. NMR spectrum of 2,2'-(1,4-phenylene)bis(<i>N</i> -hydroxyacetamide) (<i>p</i> -H ₂ xdh) in DMSO-d ₆	221
Figure A2.11. ¹³ C NMR spectrum of 2,2'-(1,4-phenylene)bis(<i>N</i> -hydroxyacetamide) (<i>p</i> -H ₂ xdh) in DMSO-d ₆	222
Figure A2.12. ¹ H NMR spectrum of <i>N</i> ² , <i>N</i> ⁵ -bis(benzyloxy)furan-2,5-dicarboxamide (2,5-Bn ₂ tdh) in DMSO-d ₆	223
Figure A2.13. ¹ H NMR spectrum of 2,2'-((oxybis(ethane-2,1-diyl))bis(oxy))bis(<i>N</i> -(benzyloxy)acetamide) (Bn ₂ gdh) in DMSO-d ₆	224
Figure A2.14. ¹ H NMR spectrum of 2,2'-((oxybis(ethane-2,1-diyl))bis(oxy))bis(<i>N</i> -hydroxyacetamide) (H ₂ gdh) in DMSO-d ₆	225

Figure A2.15. ^1H NMR spectrum of N^2,N^6 -bis(benzyloxy)pyridine-2,6-dicarboxamide (2,6-Bn ₂ pydh) in DMSO- d_6 .	226
Figure A2.16. ^1H NMR spectrum of N^2,N^6 -dihydroxypyridine-2,6-dicarboxamide (2,6-Bn ₂ pydh) in DMSO- d_6 .	227
Figure A2.17. ^1H NMR spectrum of 4,8-disulfonaphthalene-2,6-dicarboxylic acid (2,6-sndc) in DMSO- d_6 .	228
Figure A2.18. ^1H NMR spectrum of 3,7-bis((benzyloxy)carbamoyl)naphthalene-1,5-disulfonic acid (2,6-Bn ₂ sndh) in DMSO- d_6 .	229
Figure A2.19. ^1H NMR spectrum of 3,7-bis(hydroxycarbamoyl)naphthalene-1,5-disulfonic acid (2,6-H ₂ sndh) in DMSO- d_6 .	230
Figure A2.20. ^1H NMR spectrum of 2-((2,3-dimethylbut-2-en-1-yl)oxy)isoindoline-1,3-dione in DMSO- d_6 .	231
Figure A2.21. ^1H NMR spectrum of <i>O</i> -(2,3-dimethylbut-2-en-1-yl)hydroxylamine hydrochloride in DMSO- d_6 .	232
Figure A2.22. ^1H NMR spectrum of N^1,N^4 -bis((2,3-dimethylbut-2-en-1-yl)oxy)fumaramide in DMSO- d_6 .	233
Figure A2.23. ^1H NMR spectrum of N^1,N^4 -bis((2,3-dimethylbut-2-en-1-yl)oxy)fumaramide (<i>E</i> -H ₂ edh) in DMSO- d_6 .	234
Figure A2.24. ^{13}C NMR spectrum of N^1,N^4 -bis((2,3-dimethylbut-2-en-1-yl)oxy)fumaramide (<i>E</i> -H ₂ edh) in DMSO- d_6 .	235
Figure A2.25. ^1H NMR spectrum of <i>O</i> -tritylhydroxylamine in DMSO- d_6 .	236
Figure A2.26. ^1H NMR spectrum of N^2,N^6 -dihydroxy-1 <i>H</i> -pyrrole-2,5-dicarboxamide (2,5-H ₂ pdh) in DMSO- d_6 .	237
Figure A2.27. ^1H NMR spectrum of N^6,N^5 -dihydroxypyridine-3,5-dicarboxamide (3,5-H ₂ pzd) in DMSO- d_6 .	238
Figure A2.28. ^1H NMR spectrum of N^2,N^6 -dihydroxy-4-oxo-4 <i>H</i> -pyran-2,6-dicarboxamide (2,6-H ₂ podh) in DMSO- d_6 .	239
Figure A2.29. ^1H NMR spectrum of N^2,N^5 -dihydroxypyrazine-2,5-dicarboxamide in DMSO- d_6 .	240
Figure A2.30. ^1H NMR spectrum of [1,1'-biphenyl]-4,4'-dicarboxylic acid in DMSO- d_6 .	241
Figure A2.31. ^1H NMR spectrum of N^2,N^5 -dihydroxythiophene-2,5-dicarboxamide in DMSO- d_6 .	242
Figure A2.32. ^1H NMR spectrum of dimethyl 4,4'-(diazene-1,2-diyl)(<i>E</i>)-dibenzoate in CDCl ₃ .	243

Figure A2.33. ^1H NMR spectrum of dimethyl 2,2'-(diazene-1,2-diylbis(4,1-phenylene))(<i>E</i>)-diacetate in CDCl_3	244
Figure A2.34. ^1H NMR spectrum of 4-(2-chloroacetamido)-2-hydroxybenzoic acid in DMSO-d_6 . Sample contains some starting material.....	245
Figure A2.35. ^1H NMR spectrum of 4-(2-iodoacetamido)-2-hydroxybenzoic acid in DMSO-d_6 .	246
Figure A2.36. ^1H NMR spectrum of 5-(2-chloroacetamido)-2-hydroxybenzoic acid in DMSO-d_6	247
Figure A2.37. ^1H NMR spectrum of 5-(2-iodoacetamido)-2-hydroxybenzoic acid in DMSO-d_6 .	248
Figure A2.38. ^1H NMR spectrum of 2-chloro- <i>N</i> -(8-hydroxyquinolin-5-yl)acetamide in DMSO-d_6	249
Figure A2.39. ^1H NMR spectrum of 2-iodo- <i>N</i> -(8-hydroxyquinolin-5-yl)acetamide in DMSO-d_6 .	250
Figure A2.40. ^1H NMR spectrum of 4-amino- <i>N</i> -hydroxybenzamide in DMSO-d_6	251
Figure A2.41. ^1H NMR spectrum of 4-(2-chloroacetamido)- <i>N</i> -hydroxybenzamide in DMSO-d_6	252
Figure A2.42. ^1H NMR spectrum of 4-(2-iodoacetamido)- <i>N</i> -hydroxybenzamide in DMSO-d_6 . ..	253
Figure A2.43. ^1H NMR spectrum of 4-(tri(1 <i>H</i> -pyrazol-1-yl)methyl)aniline in DMSO-d_6	254
Figure A2.44. ^1H NMR spectrum of 2-chloro- <i>N</i> -(4-(tri(1 <i>H</i> -pyrazol-1-yl)methyl)phenyl)acetamide in DMSO-d_6	255
Figure A2.45. ^1H NMR spectrum of 2-iodo- <i>N</i> -(4-(tri(1 <i>H</i> -pyrazol-1-yl)methyl)phenyl)acetamide in DMSO-d_6	256
Figure A2.46. ^1H NMR spectrum of 2-chloro- <i>N</i> -hydroxyacetamide in DMSO-d_6	257
Figure A2.47. ^{13}C NMR spectrum of 2-chloro- <i>N</i> -hydroxyacetamide in DMSO-d_6	258
Figure A2.48. ^1H NMR spectrum of 2-iodo- <i>N</i> -hydroxyacetamide in DMSO-d_6	259
Figure A2.49. ^{13}C NMR spectrum of 2-iodo- <i>N</i> -hydroxyacetamide in DMSO-d_6	260
Figure A2.50. ^1H NMR spectrum of 3,3',5,5'-tetramethyl-1,1'-biphenyl in CDCl_3	261
Figure A2.51. ^{13}C NMR spectrum of 3,3',5,5'-tetramethyl-1,1'-biphenyl in CDCl_3	262
Figure A2.52. ^1H NMR spectrum of [1,1'-biphenyl]-3,3',5,5'-tetracarboxylic acid in DMSO-d_6 . ..	263
Figure A2.53. ^{13}C NMR spectrum of [1,1'-biphenyl]-3,3',5,5'-tetracarboxylic acid in DMSO-d_6 .	264
Figure A2.54. ^1H NMR spectrum of $N^\beta, N^\beta, N^\delta, N^\delta$ -tetrahydroxy-[1,1'-biphenyl]-3,3',5,5'-tetracarboxamide (H_4BPTH) in DMSO-d_6	265

Figure A2.55. ^{13}C NMR spectrum of $N^\beta, N^\beta, N^\delta, N^\delta$ -tetrahydroxy-[1,1'-biphenyl]-3,3',5,5'-tetracarboxamide (H_4BPTH) in DMSO-d_6 .	266
Figure A2.56. ^1H NMR spectrum of 3,3'',5,5''-tetramethyl-1,1':4',1''-terphenyl in DMSO-d_6 .	267
Figure A2.57. ^1H NMR spectrum of [1,1':4',1''-terphenyl]-3,3'',5,5''-tetracarboxylic acid in DMSO-d_6 .	268
Figure A2.58. ^1H NMR spectrum of $N^\beta, N^\beta, N^\delta, N^\delta$ -tetrahydroxy-[1,1':4',1''-terphenyl]-3,3'',5,5''-tetracarboxamide (H_4TPTH) in DMSO-d_6 .	269
Figure A2.59. ^1H NMR spectrum of 3,3''',5,5'''-tetramethyl-1,1':4',1'':4'',1'''-quaterphenyl in CDCl_3 .	270
Figure A2.60. ^1H NMR spectrum of [1,1':4',1'':4'',1'''-quaterphenyl]-3,3''',5,5'''-tetracarboxylic acid in DMSO-d_6 .	271
Figure A2.61. ^1H NMR spectrum of $N^\beta, N^\beta, N^\delta, N^\delta$ -tetrakis(trityloxy)-[1,1':4',1'':4'',1'''-quaterphenyl]-3,3''',5,5'''-tetracarboxamide (H_4QPTH) in DMSO-d_6 .	272

List of Tables

Table 2.1. Primers for site-directed mutagenesis of human heavy-chain ferritin	50
Table 2.2. Mass spectrometry measurements for all ferritin variants.	52
Table 2.3. Crystallization conditions for ferritin.	55
Table 2.4. X-ray data processing and refinement statistics.....	56
Table 3.1. Unit cell parameters for 21 protein-MOFs without precipitant.....	76
Table 3.2. Crystallization conditions for Ni-ferritin and Co-Ferritin	98
Table 3.3 X-ray data processing and refinement statistics for Ni- and Co-ferritin nodes.....	99
Table 3.4. Crystallization conditions for ferritin-MOFs	119
Table 3.5. X-ray data processing and refinement statistics of <i>bcc</i> and <i>bct</i> ferritin-MOFs	120
Table 3.6. X-ray data processing and refinement statistics of primitive ferritin-MOFs.....	121
Table 4.1. Crystallization conditions for 2,5-fdh-Ni-ferritin	150
Table 4.2. X-ray data processing and refinement statistics of 2,5-fdh-Ni-ferritin	151
Table 5.1. DNA sequences of <i>T. ni</i> ferritin variants	172
Table 5.2. Sequence of ³⁰ CROP.	175
Table A1.1 X-ray data processing and refinement statistics.	208

Acknowledgments

I would first like to thank my advisor, Professor Akif Tezcan, for his support and guidance throughout graduate school. His enthusiasm for science and rigorous work ethic is always inspiring. I would not have been able to accomplish nearly as much without his unwavering support and continuous push for excellence.

I have been extremely fortunate to be a member of the Tezcan Lab and work alongside truly exceptional scientists. This work would not have been possible without input from many members of the Tezcan Lab over the past six years. Starting from the beginning, I would like to thank my original mentors Pam Sontz and Dustin Huard. Their guidance and support helped me quickly acclimate to an entirely new field of research. I would like to thank Lewis Churchfield, both for his research advice and for welcoming me into his group of friends. The members of the Team Ferritin subgroup, especially Ling Zhang and Kenneth Han, always aided me with conducting complicated experiments and helped out during synchrotron trips. I would also like to thank Rob Alberstein, Rohit Subramanian, Julian Esselborn, and Jie Zhu, for answering all of my quick questions even when the answers weren't always quick. I was also fortunate enough to work with two talented undergraduate researchers, Sophia Ahn and Jerika Chiong. It has been great working alongside all of the members of the Tezcan lab over the past six years.

I would like to thank Milan Gembicky for mentoring me as a crystallographer and always helping me tackle intricate problems and Professor Alex Groisman, who designed and constructed microfluidic devices for me, often at a moment's notice. I also wish to thank Steven Weigand (BL 5-ID-D at APS), Thomas Weiss and Ivan Rajakovic (BL 4-2 at SSRL) for their support during my many synchrotron trips. I would also like to thank the many members of the UCSD research community that I was able to work with over the past six years. Everyone I encountered was supportive and willing to lend a bit of their time to help me with my research.

I want to thank all of my friends that have supported me throughout my Ph.D. Living in San Diego would not have been nearly as enjoyable without everyone. I would especially like to thank the PB volleyball crew for providing me with brief escapes from research. Lastly, I would like to thank my family for their unwavering support throughout graduate school.

Chapter 2, in part, is a reprint of the material “A metal-organic framework with spherical protein nodes: rational chemical design of 3D protein crystals.” as it appears in the Journal of the American Chemical Society (Sontz, P. A.*; Bailey*, J. B.; Ahn, S.; Tezcan, F. A. *J. Am. Chem. Soc.* **2015**, *137* (36), 11598-11601). The dissertation author is a co-primary author of this publication. He would like to thank the US Department of Energy (DOE) (Division of Materials Sciences, Office of Basic Energy Sciences, (DE-FG02-10ER46677) for funding this work.

Chapter 3, in part, is a reprint of the material “Synthetic Modularity of Protein-Metal-Organic Frameworks” as it appears in the Journal of the American Chemical Society (Bailey, J. B.; Zhang, L.; Chiong, J. A.; Ahn, S.; Tezcan, F. A. **2017**, *139* (24), 8160-8166). The dissertation author is the primary author of this publication. Chapter 3 also includes work that is currently being prepared for publication with contributions from the following authors: Bailey, J. B.; Tezcan, F. A. The dissertation author will be the primary author of this publication.

Chapter 4, in part, is currently being prepared for publication with contributions from the following authors: Bailey, J. B.; Tezcan, F. A. The dissertation author will be the primary author of this publication.

Chapter 5, in part, is based on the unpublished research efforts of the dissertation author with contributions from Rob Alberstien (⁹⁸C RhuA), Jie Zhu, Yiyang Li (³⁰C ROP), Jerika Chiong (synthesis of organic ligands), and Akif Tezcan.

Appendix 1, in part, is currently being prepared for publication with contributions from the following authors: Chiong, J. A.; Bailey, J. B.; Zhu, J.; Xu, W.; Cohen, S.; Tezcan, F. A. The dissertation author will be a co-primary author of this publication. The author would like to highlight the hard work of Jerika Chiong, who was the primary driving force behind this project, and Prof.

Seth Cohen, who initiated the product and provided valuable guidance. Appendix 1 also contains electron diffraction data that was collected by Zhehao Huang and Laura Sampersky in the lab of Xiaodong Zho at Stockholm University.

The author would like to thank the US Department of Energy (DOE) (Division of Materials Sciences, Office of Basic Energy Sciences, (DE-FG02-10ER46677) and the Lattimer foundation for funding this work.

Vita

Education

The University of California San Diego, CA <i>Ph.D. in Chemistry</i>	March 2020
The University of California San Diego, CA <i>M.S. in Chemistry</i>	May 2015
Northern Arizona University, Flagstaff, AZ <i>B.S. in ACS Extended Chemistry; B.S. in Exercise Science</i>	May 2013

Honors and Awards

Lattimer Award for Excellence in Interdisciplinary Research	Dec 2017
Distinguished Graduate Student Fellowship	Dec 2017
McAlister Scholarship for Undergraduate Research	June 2013
Clinton F. Lane Scholarship for Undergraduate Research	June 2012
Hypercube Scholarship for Undergraduate Research in Computational Chemistry	June 2012

Publications

Golub, E.; Subramanian, R. H.; Esselborn, J.; Alberstein, R. G.; **Bailey, J. B.**; Chiong, J. A.; Yan, X.; Booth, T.; Baker, T. S.; Tezcan, F. A. Constructing Protein Polyhedra from Asymmetric Monomers via Chemically Orthogonal Interactions. *Nature*. **2020**, *578*, 172-176.

Subramanian, R. H.* Smith, S. J.*; Alberstein, R. G.; **Bailey, J. B.**; Cardone, G.; Suominen, L.; Chami, M.; Stahlberg, H.; Baker, T. S.; Tezcan, F. A. Self-Assembly of a Designed Nucleoprotein Architecture through Multimodal Interactions. *ACS Cent. Sci.* **2018**, *4* (11), 1578-1586.

Zhang, L.; **Bailey, J. B.**; Subramanian, R. H.; Groisman, A.; Tezcan, F. A. Hyperexpandable, self-healing macromolecular crystals with integrated polymer networks. *Nature*. **2018**, *557*, 86-91.

De Meulenaere, E.; **Bailey, J. B.** Tezcan, F. A.; Deheyn, D. D. First biochemical and crystallographic characterization of a fast-performing ferritin from a marine invertebrate. *Biochem. J.* **2017**, *474* (24), 4193-4206.

Bailey, J. B.; Zhang, L.; Chiong, J. A.; Ahn, S.; Tezcan, F. A. Synthetic Modularity of Protein-Metal-Organic Frameworks. *J. Am. Chem. Soc.* **2017**, *139* (24), 8160-8166.

Bailey, J. B.*; Subramanian, R. H.*; Churchfield, L. A.*; Tezcan, F. A. Metal-directed design of supramolecular protein assemblies. *Methods. Enzymol.* **2016**, *580*, 223-250.

Sontz, P. A.*; **Bailey, J. B.***; Anh, S.; Tezcan, F. A. A metal-organic framework with spherical protein nodes: rational chemical design of 3D protein crystals. *J. Am. Chem. Soc.* **2015**, *137* (36), 11598-11601.

Sirimulla, S.*; **Bailey, J. B.***; Rahulsimhan, V.; Narayan, M. Halogen interactions in protein-ligand complexes: implications of halogen bonding for rational drug design. *J. Chem. Inf. Model.* **2013**, 53, 2781-2791.

Vadlapudi, A. D.; Vadlapatla, R. V.; Earla, R.; Sirimulla, S.; **Bailey, J. B.**; Pal, D.; Mitra, A. K. Novel biotinylated lipid prodrugs of acyclovir for the treatment of herpetic keratitis (HK): transporter recognition, tissue stability, and antiviral activity. *Pharm. Res.* **2013**, 30 (8), 2063-2076.

Sirimulla, S.; Pal, R.; Raparla, M.; **Bailey, J. B.**; Duran, R.; Altamirano, A. M.; Herndon, W. C.; Narayan, M. Identification of novel nitrosative stress inhibitors through virtual screening and experimental evaluation. *Mol. Inf.* **2012**, 31, 167-172.

Abstract of the Dissertation

Protein-Metal-Organic Frameworks and other Hydroxamate-based Crystalline Lattices

by

Jake Brian Bailey

Doctor of Philosophy in Chemistry

University of California San Diego, 2020

Professor Faik Akif Tezcan, Chair

Proteins are nature's chosen building blocks to create sophisticated and functional 0-, 1-, 2- and 3D assemblies that are necessary for the complexity of life. Accordingly, there is significant interest in the development of highly ordered supramolecular protein arrays that mimic or even surpass the complexity and functional versatility of those assemblies found in nature. However, due to the chemical and structural heterogeneity of proteins, the rational

design of supramolecular protein architectures is challenging. In contrast, coordination polymers such as metal-organic frameworks (MOFs) benefit from a high degree of modularity, where individual components can be interchanged to produce structurally unique architectures with vastly different materials properties. Inspired by the synthetic versatility of MOFs, we developed a method for assembling protein oligomers into highly ordered frameworks, termed protein-MOFs. Akin to more conventional MOFs, the modularity of these lattices stems from the multiple interchangeable components. We engineered a protein node through the installation of outward-facing metal-binding sites at the C_3 vertices of the symmetric cage-like protein, human heavy-chain ferritin. These sites coordinate a variety of transition metals with high fidelity, allowing for the formation of extended crystalline networks when connected by ditopic bridging ligands. The symmetry of the resulting crystalline lattice is dictated by the inherent symmetry of the protein, the coordination preferences of the metal, and the geometry of the bridging ligand. This modularity led to the discovery of stimuli-responsive protein-MOFs capable of dynamic crystal-to-crystal transformations. Protein-MOFs have the potential to be utilized for the construction of functional biomaterials.

Chapter 1. Supramolecular protein self-assembly

Nature's most versatile biomolecules, proteins, are omnipresent across all manner of life. Proteins act as highly evolved molecular machinery and perform vital roles such as transcription, translation, catalysis, molecular transport, signal-recognition, and structural support.¹ Precisely ordered amino acids comprise the primary structure of each protein and dictate the shape of the tertiary fold. The naturally available pool of 20 amino acids provides virtually limitless sequence space for the evolution of diverse structures tailored to perform highly specialized functions. Individual proteins rarely act alone and instead often form synergistic homo- or hetero-oligomeric quaternary complexes. It is estimated that up to 80% of proteins from the most widely studied prokaryote, *Escherichia coli* (*E. coli*), exist primarily as oligomers.² The vast majority of these oligomers form discrete 0D quaternary structures. These 0D complexes range from relatively simple homodimers (such as leucine zipper proteins) to intricate multiprotein oligomers (including ATP-synthases, virus capsids, and the ribosome).³⁻⁶ A much smaller subset of oligomeric proteins that self-assemble into pseudo-infinite 1-, 2-, and 3D functional assemblies. These extended oligomers are implicated throughout biology in cell motility, cell division, intracellular transport, selective sequestration, and structural support, among others.⁷⁻⁹ Thus, there is significant research interest in the development of novel protein materials that mimic or even surpass the functionality of those found in Nature. However, the design of extended protein assemblies is incredibly challenging. The customizable shape and chemical heterogeneity of proteins that make them such attractive, functional building blocks greatly complicate their incorporation into ordered networks. Unlike DNA-based architectures, which benefit from the highly specific Watson-Crick base pairing, there are no generalizable rules for protein self-assembly. Instead, a myriad of design strategies has been employed to direct the formation of self-assembling protein networks. This chapter examines how Nature forms “pseudo-infinite” protein assemblies and describes ongoing efforts to develop methodologies for their construction.

1.1. Extended 1-, 2-, and 3D protein assemblies found in Nature

In most organisms, the proteins with the highest levels of expression are found in self-assembled extended networks. For vertebrates, this protein is collagen. Collagen is an abundant structural protein found in the extracellular matrix that accounts for ca. 33% of the total protein content in humans.⁷ Collagen is comprised of three left-handed polyproline II polypeptide chains that assemble into a right-handed triple helix with a helical pitch of ca. 3.3 residues per turn (Figure 1.1 a). The well-defined helical structure is tightly packed and stabilized by an intrachain hydrogen bond between the amide hydrogen of each glycine (which repeats every third residue) and a backbone carbonyl of a neighboring polypeptide strand (Figure 1.1 a). Only the repeating glycine is conserved; the other positions can accommodate any amino acid, though Pro and Hyp are by far the most common⁷. After assembly of this triple helix (procollagen), the N and C terminal ends are enzymatically truncated to form tropocollagen helices. Multiple tropocollagen helices self-assemble into extended fibrils, which are then covalently crosslinked through exposed lysines creating the collagen fibers found in the extracellular matrix.

While collagen acts as the scaffolding for the extracellular matrix, cytoskeletal proteins provide structural integrity for eukaryotic cells. Actin is the most abundant intracellular protein, accounting for 10% of the total protein content of muscle cells and up to 5% of nonmuscle cells.⁸ This protein is ca. 42 kDa and exists as either a monomeric globular subunit (G-actin) or as homooligomeric 1D filaments (F-actin). F-actin forms the basis for the cytoskeleton and is involved in cell motility, cell division, and many other intracellular processes. F-actin self-assembles into two helical strands that form a left-handed helix with a helical pitch of ca. 2.2 actin subunits per turn.¹⁰ These 1D assemblies are mediated by a series of hydrophobic interactions and salt bridges, leading to stable protein-protein interfaces with large buried surface areas (Figure 1.1 b).¹¹⁻¹² Much like actin, tubulin is another cytoskeletal protein that forms highly-ordered 1D assemblies. Both the α - and β -tubulin subunits are ca. 50 kDa and combine to form hollow

heteropolymeric microtubules with 25 nm diameters that serve as roads for intracellular trafficking.¹³ These microtubules are highly dynamic and continuously reshaped through the addition or subtraction of the tubulin subunits. All of these 1D assemblies are contingent on the formation of multiple, highly specific noncovalent interactions.

2D surface layer (S-layer) proteins are the most common multidimensional assemblies in Nature and are found in the cellular envelopes of nearly all *archaea* and walled *bacteria*.¹⁴ Much like collagen and actin, the structural role of S-layer proteins leads to high expression levels, accounting for ca. 10% of the total protein content within prokaryotic cells.¹⁵ Despite their ubiquitous application across prokaryotic organisms, S-layer proteins share little sequence similarity.¹⁶ The monomeric subunits of S-layer proteins can vary from 40 to 170 kDa. The lattice symmetry of S-layer proteins is also variable; oblique (*P1*, *P2*), square (*P4*), and hexagonal (*P3*, *P6*) symmetries have all been observed (Figure 1.1 c). Depending on both the monomer size and the packing geometry, the pore sizes of S-layers can also vary from 2-8 nm. The lack of sequence similarity and variance in both packing geometry and porosity leads to a variety of proposed functions for S-layer proteins. It is thought that S-layer proteins play a role in pathogenesis, selective permeability, and structural support.¹⁶

Unlike 1- and 2D protein self-assembly, the occurrence of ordered 3D protein crystals in Nature is relatively uncommon. Proteins have evolved to avoid unwanted self-aggregation, and the spontaneous crystallization of intracellular proteins would be detrimental to many cellular functions.¹⁷ The formation of 3D assemblies also provides additional challenges, since the majority of proteins do not have symmetry axes that extend in three dimensions. However, there are select cases where the formation of 3D crystallites is both readily achievable and biologically advantageous. The majority of these cases can be categorized by function, either protein storage, encapsulation, or solid-state catalysis.¹⁸ In *homo sapiens*, pancreatic β -cells store insulin and release in microcrystalline granules.¹⁹ Insulin forms a C_3 symmetric hexameric complex through

the coordination of Zn^{2+} ions. These oligomers further assemble stepwise into ordered crystals.²⁰ It is thought that crystal formation helps prevent the degradation of insulin during storage. Similarly, a subset of insect viruses (such as cytoplasmic polyhedrosis viruses) causes the host to overexpress a specific protein (polyhedrin) that crystallizes around the virus providing a protective shell mediated by extensive noncovalent interactions (Figure 1.1 d-e).²¹ This shell dissolves upon the digestion of the viral crystals, allowing the virus to infect the new host. In eukaryotic cells, peroxisomes often employ porous enzyme crystals as solid-state catalysts.²² These crystals have incredibly high local concentrations of specific enzymes to be packed into a relatively small space. Since the enzyme's substrates are delivered directly to the peroxisome, there is little downside to an immobile 3D enzymatic scaffold. This class of crystalline materials serves as inspiration for the development of ordered heterogenous enzyme-based catalysis to execute complicated stereospecific reactions.

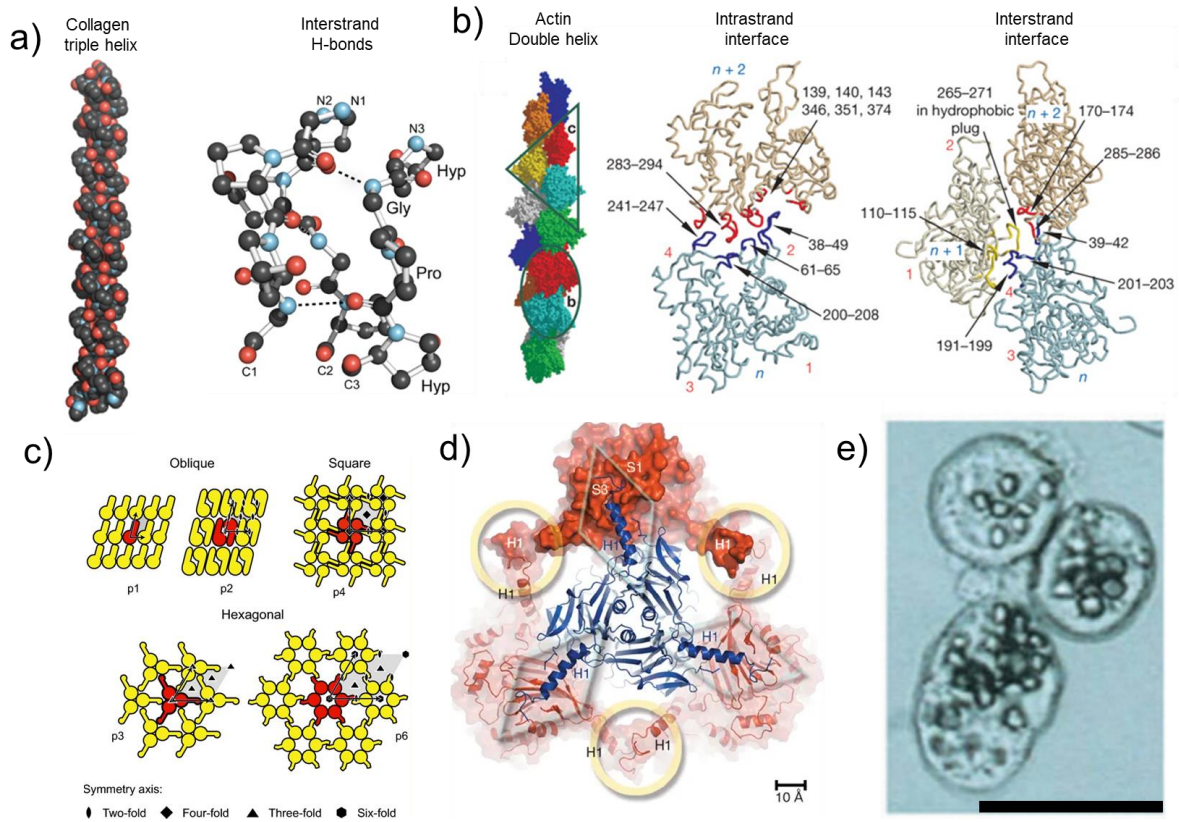


Figure 1.1. 1-, 2-, and 3D assemblies in Nature. a) Collagen triple helix with the highlighted interstrand bonds. b) Actin double helix with the intrastrand and interstrand interactions, respectively. c) Characterized packing symmetries of S-layer proteins. d) Crystal structure of polyhedrin. e) Light microscopy image of CPV infected insect cells. The scale bar is 10 μm. Panel A was adapted from reference 7. Panel B was adapted from reference 12. Panel C was adapted from reference 23. Panels D and E were adapted from reference 21.

1.2. Assembly strategies for 1-, 2-, and 3D protein materials

Designed 1-, 2- and 3D protein assemblies have recently received much attention for potential applications in molecular recognition, drug delivery, biocatalysis, due to their customizable structure, biocompatibility, and potential for large-scale biosynthesis.²⁴⁻²⁶ While there has been considerable progress towards the *de novo* design of pseudo-infinite biomaterials, precisely patterning proteins into multidimensional architectures is still extremely challenging and requires significant trial and error. The diversity in protein size, shape, charge, and displayed functionalities that make proteins such attractive components for biomaterials also complicates their self-assembly. Thus, many different methodologies have been developed for the creation of ordered protein assemblies.²⁶⁻²⁸ The pseudo-infinite arrays found in Nature are primarily stabilized by large interfaces comprised of many complementary, highly specific, and noncovalent protein-protein interactions.¹ Researchers have employed similar methods to develop novel protein architectures through charge-based self-assembly, the genetic fusion of multiple oligomeric domains, or redesign of protein-protein interfaces.²⁹⁻³² Alternatively, highly specific bonding functionalities such as guest-host interactions, metal-mediated contacts, or disulfide bonds offer higher specificity and a smaller design footprint. It is even possible to use multiple complementary design strategies to form even more robust and complex protein assemblies. There have been many reviews written on engineered protein assemblies.^{26-27, 33-36} Only strategies that have been employed to create extended assemblies (1-, 2-, or 3D) are discussed in this chapter.

1.2.1. Charge-based assembly of protein networks

Out of the twenty amino acids that comprise natural proteins, four sidechains are charged at neutral pH. The carboxylate groups on Glu and Asp are negatively charged; the amine and guanidine groups are positively charged in Lys and Arg. These charged residues are often highly solvent-exposed and unevenly distributed across the protein surface. This uneven distribution often leads to local patches of positive and negative charges. The strength of these charged interactions can be modulated by the ionic strength of the solution. In highly symmetric protein, such as the octahedral ferritin and icosahedral cowpea chlorotic mottle virus (CCMV), the patchy regions are symmetrically displayed across the protein surface. Both *P. furiosus* ferritin (diameter = 12 nm) and CCMV (diameter = 28 nm) and have a negative charge at neutral pH.²⁹ Positively charged gold nanoparticles (AuNPs) stabilized with terminal quaternary amines (diameter = 8.5 nm) were synthesized to pair with the negatively-charged proteins. After precisely tuning the interaction strength, *P. furiosus* ferritin and CCMV formed crystals with the positively charged AuNPs.²⁹ The ferritin-AuNP lattice packs into a primitive cubic lattice isostructural with CsCl. The size discrepancy between CCMV and the AuNPs leads to the formation of an *fcc* lattice with eight gold nanoparticles for every CCMV particle and edge lengths of $a = 40.4$ nm (Figure 1.2 a and b). A nearly identical *fcc* lattice was obtained by replacing the AuNPs with positively charged dendrimers (PAMAM-G6).³⁷ Combining much smaller dendrimers (PAMAM-G2) and *P. furiosus* ferritin predictably formed a smaller *fcc* lattice ($a = 14.7$ nm).³⁸ The size of the *P. furiosus* ferritin-AuNP unit cell was systematically controlled by varying the size of the dendrimer (Figure 1.2 c); PAMAM-G4 crystallized into an *fcc* lattice with a 2.2 nm increase in edge length ($a = 15.9$ nm). Further increasing the size of the dendrimer (PAMAM-G5 to -G7) leads to a tightly packed *hcp* lattice. A similar strategy was also employed to form 1D nanowires from the homodecameric, negatively-charged protein stable protein 1 (SP1) and positively charged PAMAM-G5 dendrimers or core linked micelles.³⁹⁻⁴⁰

Unlike ferritin and CCMV, which have local patches of positive and negative charges, the AuNPs, dendrimers, and micelles are amorphous with an isotropic charge distribution. Combining patchy positive and negative components facilitates the formation of different lattices. Avidin, a relatively small (diameter = 4.8 nm), positively-charged tetrameric protein, forms a less densely packed *bcc* lattice with CCMV ($a = 35$ nm) due to the formation of directional charge-based interactions. Disordered linear positively charged peptides have also been used to form crystalline arrays.⁴¹ These peptides interacted with the interior of an engineered soybean seed ferritin through the open 4-fold axis and templated the formation of 2D sheets. Similarly, the positively charged inorganic dye phthalocyanine was small enough to selectively bridge between the negatively charged 2-fold axes of two *P. furiosus* ferritin subunits leading to a tightly-packed *fcc* lattice ($a = 18.3$ nm). In all of these examples, the size mismatch between the components and the flexibility of the positively charged moiety yielded crystals with little short-range order. Instead of relying on natural proteins with patchy areas, the protein subunits can be modified to favor charge-based self-assembly. Human heavy-chain ferritin (HuHF) was reengineered to bear either a highly negative or highly positive charge.⁴² The cocrystallization of the positive and negative variants created a *P4* lattice, where the positively and negatively charged ferritins alternate in a pseudo-*bcc* packing (Figure 1.2 d). Unlike the previous examples, which all had amorphous charge carriers or significant size discrepancies between components, this lattice is highly-ordered and was solved to a resolution of 1.8 Å. This example highlights that weak nondirectional charge-based interactions can template highly ordered assemblies under the right conditions.

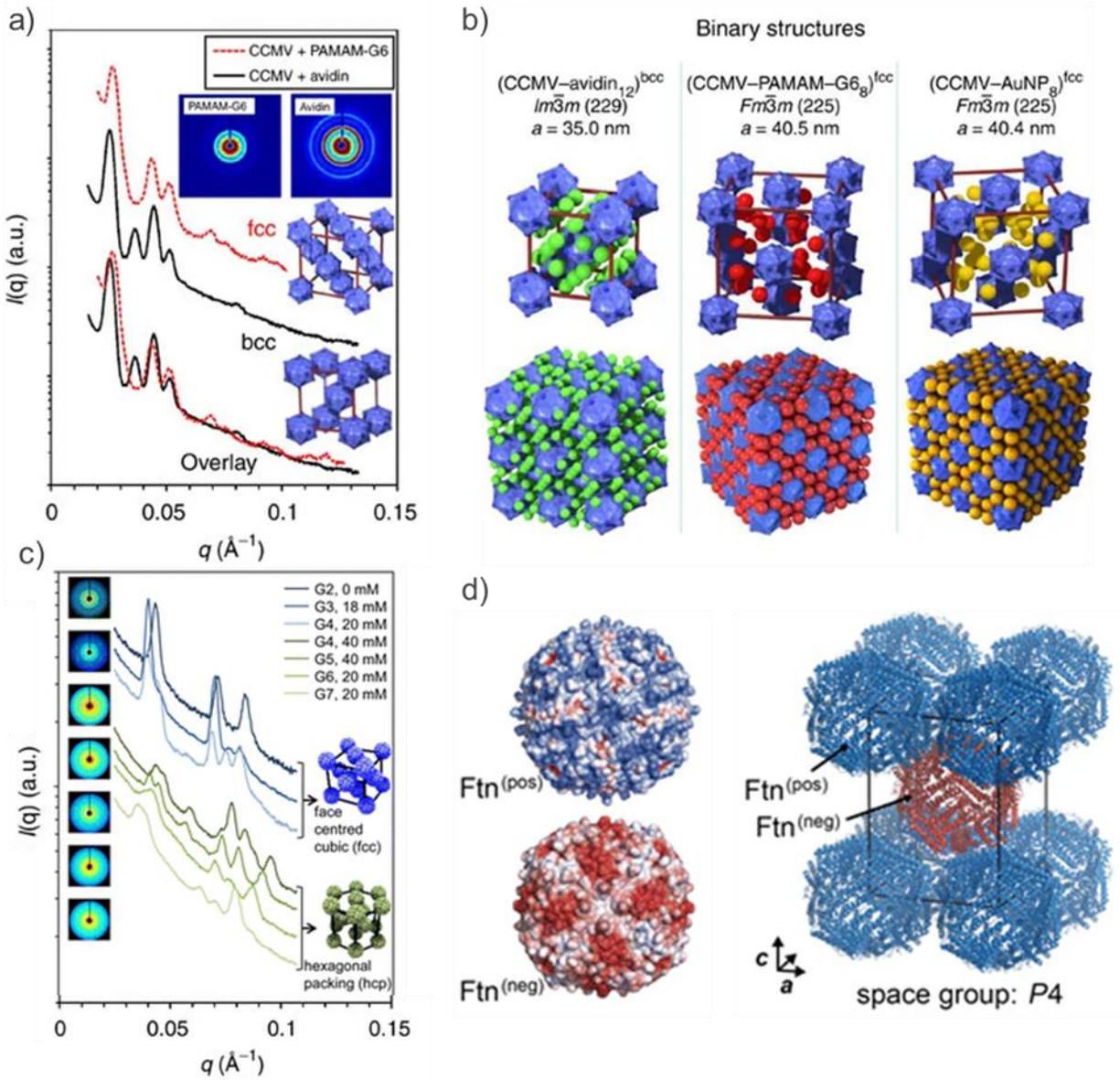


Figure 1.2. Electrostatic assembly of ferritin and CCMV. a) SAXS diffraction of CCMV-avidin (*bcc*) and CCMV-PANAM-G6 (*fcc*) lattices. b) Unit cell dimensions and expected packing of CCMV-avidin, CCMV-PANAM-G6, and CCMV-AuNP. c) Diffraction profiles of ferritin with different dendrimer sizes. d) positively and negatively charged ferritin assembled into a pseudo-*bcc* packing. Panel a and b were adapted from reference 37. Panel c was adapted from reference 38. Panel d was adapted from reference 42.

1.2.2. Protein assembly from genetically fused constructs

The interfaces of native protein oligomers are highly evolved to form with high fidelity. These natural interfaces can be exploited to create extended architectures through the genetic fusion of two (or more) oligomeric proteins. This approach was first employed to form 0D cages and extended 1D filaments.³⁰ The pseudo-infinite assemblies were assembled from a fusion construct containing a single chain from two C_2 dimeric proteins, the viral *influenza* matrix protein M1 and the *P. fluorescens* carboxylesterase, separated by a rigid α -helical linker. Upon expression, the two protein domains each formed the expected C_2 homodimer, leading to the formation of insoluble 1D linear protein chains. Individual chains spontaneously aggregated and formed larger fibrils that were characterized by TEM (Figure 1.3 a). A similar design strategy was used to assemble hexagonal lattices *in vivo* using the octahedral 24-meric HuHF N-terminally fused to a C_2 symmetric citrine subunit.⁴³ This design strategy was improved by utilizing protein oligomers with dihedral symmetries as the building blocks, which removed the need for a rigid linker sequence (Figure 1.3 b).²⁸ Additionally, the updated design used heterologous assembly components to prevent spontaneous aggregation *in vivo*, allowing the components can be expressed separately and later recombined under ideal crystallization conditions. These improvements lead to the formation of both 1D and 2D extended assemblies that could be characterized by TEM (Figure 1.3 c-e). The 1D linear chains were formed from the D_2 red fluorescent protein (DSred) appended with a strep-tag. The D_2 symmetric streptavidin binds the strep tag with high affinity, connecting the DSred tetramers and forming 1D assemblies. Much like the DSred construct, *E. Coli* aminolevulinic acid dehydrogenase (ALAD) was genetically fused to a terminal strep-tag. ALAD, a D_4 symmetric octamer, self-assembled into well ordered 2D sheets with the streptavidin tetramer. Similar lattices were also obtained by replacing the strep-tag and streptavidin with the heterodimeric Lac21E and Lac21K. By fusing proteins with matching symmetries, this method can generate a wide range of different lattices.

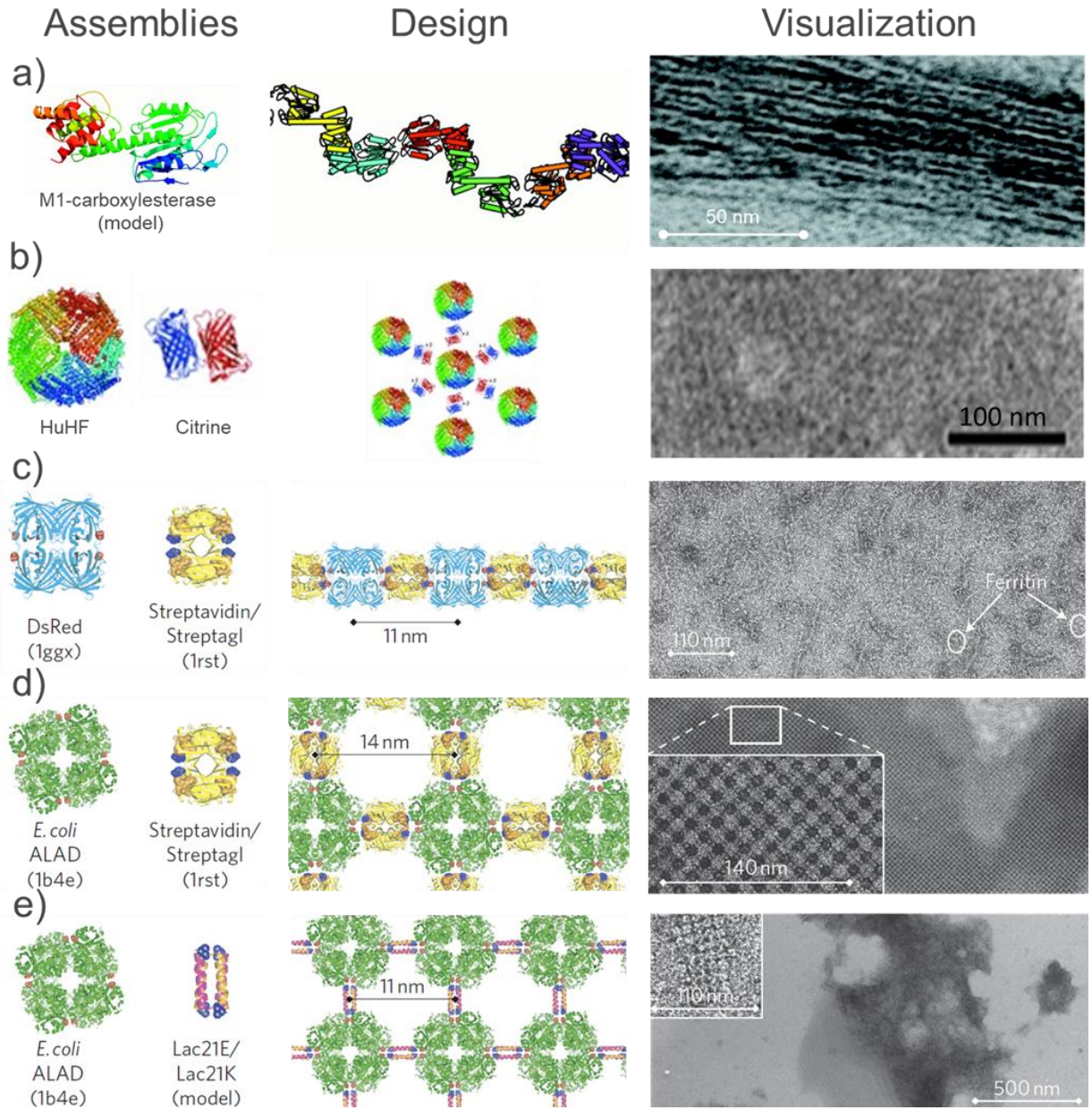


Figure 1.3. 1- and 2D protein assemblies formed through genetic fusions. a) The C_2 matrix protein M1-carboxylesterase fusion forms 1D chains that assemble into larger fibers. b) HuHF-citrine fusion creates hexagonal arrays *in vivo*. c) The combination of the D_2 DsRed-strep and D_2 streptavidin form 1D assemblies. d) D_4 symmetric ALAD-strep and D_2 symmetric streptavidin form ordered 2D sheets. e) ALAD-Lac21E and ALAD-Lac21K form ordered 2D sheets. Panel a was adapted from reference 30. Panel b was adapted from reference 43. Panels c-e were adapted from reference 28.

1.2.3. Assemblies constructed from noncovalent protein-protein interactions

Protein-protein interfaces in Nature are highly complex and comprised of a multitude of weak noncovalent interactions spread over large interfaces. In the previous section, existing protein-protein interfaces were used to create periodic protein arrays. By mimicking nature, novel noncovalent interfaces have been designed to incorporate multiple synergistic hydrophobic or hydrogen-bonding interactions. This strategy repurposes existing protein scaffolds or highly conserved motifs. One such motif, α -helical coiled coils, was utilized to form fibrous 1- and 2D arrays.⁴⁴ Two short complementary polypeptides were synthesized to form dimeric parallel coiled-coil. Positively and negatively charged residues were installed at the N and C terminal halves, respectively. The N-terminal half of one peptide preferentially interacting with the C-terminal half of the other, forming a dimer with two overhanging “sticky ends”. The “sticky ends” promote further linear assembly of the polypeptide chains, followed by aggregation into large fibrils (Figure 1.4 a). Similar coiled-coil motifs were computationally designed to assemble into a 3D protein crystal with $P6$ symmetry.⁴⁵ The outward-facing residues of a tightly packed C_3 symmetric, parallel coiled-coil protein, were screened *in silico* for low energy packing conformations. Out of the five selected sequences, one assembled as predicted into a crystal lattice with $P6$ symmetry (Figure 1.4 b). Improvements in both the accuracy of the force field calculations and the implementation of advanced docking algorithms have enabled the complete redesign of large protein-protein interfaces. Rosetta symmetric docking was employed to design complementary protein-protein interfaces *in silico* for a series of symmetric building blocks.⁴⁶ Out of the designed constructs, three self-assembled into the predicted 2D lattice *in vivo* with $P321$, $P42_12$, and $P6$ symmetry, respectively (Figure 1.4 c). The Rosetta symmetric docking was also used to predict hydrophobic or hydrogen-bonding interfaces that would lead to the 2D assembly for a single dimeric oligomer. In each case, ordered 2D sheets self-assembled.⁴⁷ While the computational accuracy of interface redesign continues to improve, complex interfaces comprised of multiple synergistic mutations are not always needed to form novel noncovalent lattices. A single point mutation (either Phe,

Tyr, or Trp) at a surface-exposed residue on the 4-fold axis was sufficient to induce the formation of 2D or 3D assemblies based exclusively on the engineered noncovalent interactions (Figure 1.4 d).⁴⁸

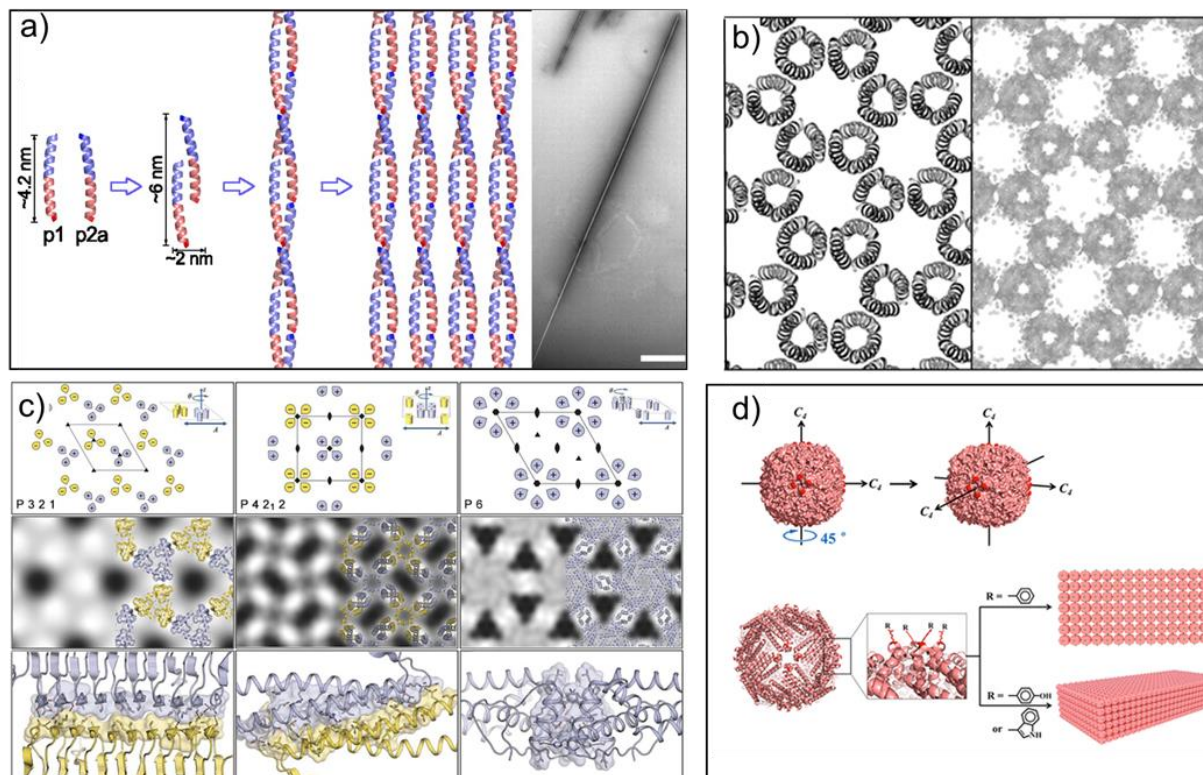


Figure 1.4. Protein assemblies based on designed noncovalent protein-protein interactions. a) Complementary α -helical peptides (p1 and p2a) form a charge-based dimer that templates further assembly into linear chains and extended fibrils. Predominantly positively and negatively charged regions are shown as blue and red, respectively. The scale bar in the negatively-stained TEM image is 2 μ m. b) a single layer of the designed P6 model (left) and the resolvable electron density of a layer of the P6 crystal structure contoured at 2σ . c) Three examples of self-assembled 2D protein crystals mediated by the computationally designed interfaces. d) schematic for the assembly of 2D and 3D arrays of HuHF from a single point mutation. Panel A was adapted from reference 44. Panel B was adapted from reference 45. Panel C originated from reference 26 and was adapted from reference 46. Panel D was adapted from reference 48.

1.2.4. Disulfide mediated assembly of protein lattices

Staying within Nature's existing toolkit, cysteine residues routinely form either inter- or intramolecular disulfide bonds to stabilize folded domains or induce oligomerization. Disulfide bonds are relatively stable, highly specific, and reversible under reducing conditions, making them excellent candidates for directing protein self-assembly. Cysteine residues are also sterically unencumbered and can easily be accommodated into critical positions in most protein scaffolds. Installing cysteine residues at surface-exposed locations on monomeric proteins induce dimerization and help assemble the proteins into an ordered crystal lattice.⁴⁹ Cysteine residues have been strategically placed on opposing faces of symmetric proteins to induce disulfide-bond formation to promote the creation of 1- and 2D extended protein assemblies. One of the first examples utilized Heme carrier protein 1 (HCP1), which exists in nature as a C_6 symmetric hexameric ring (diameter = 9 nm) in solution and packs into highly ordered nanotubes during crystallization. Cysteine residues within disulfide-bonding distance in the native crystal structure were installed on the top (face A) and bottom (face B) of HCP1 (Figure 1.5 a-c) to promote the assembly of the nanotubes in solution.⁵⁰ Under oxidizing conditions, disulfide bond formation templates the construction of short nanotubes. HCP1 variants that contain cysteines on the top or bottom were used to cap these nanotubes and prevent further growth.

A similar design strategy installed cysteine on both the top (face A) and bottom (face B) of a subunit of a similarly sized (diameter = 9 nm) C_{11} symmetric protein trp RNA-binding attenuation protein (TRAP).⁵¹ Under slow oxidizing conditions, the formation of long nanotubes (up to 1 μm) was observed by TEM. Additional characterization of the assembly revealed that the two A faces of TRAP form the expected disulfide-mediated C_2 symmetric dimer.⁵² The cysteine residues on the B faces were unable to form protein-protein disulfide bonds and instead formed heteromeric disulfide bonds with reductant (either DTT or DMP) during nanotube growth. These

two examples highlight the effectiveness of displaying cysteine residues on opposing facets of a subunit to template the formation of 1D networks.

Similarly, 2D materials can be formed by displaying cysteine residues along a plane. To this end, cysteine residues were installed at the corners of the C_4 symmetric tetramer, L-rhamulose 1-phosphate aldolase (RhuA).⁵³ Under slowly oxidizing conditions, the modified RhuA assembled into large (ca. 2 μm) disulfide-mediated sheets that were characterized by TEM (Figure 1.5 d-f). Unlike the previous examples, which all contained multiple disulfide bonds at the protein-protein interface, each RhuA is connected to a neighboring subunit by a single disulfide bond. The adjacent RhuA subunits all alternate, giving rise to a $P4_21_2$ lattice with $a = 114 \text{ \AA}$. Due to the relatively low profile and flexibility of the cysteine residues, these 2D sheets were capable of collective auxetic rearrangements in response to mechanical agitation. Using similar design principles, researchers installed cysteine residues at the C_4 axes of the octahedral HuHF.⁵⁴ The symmetry of ferritin allows for the potential to form a 3D disulfide-mediated lattice. However, only four of the six 4-fold axes formed HuHF-HuHF disulfide bonds, leading to the formation of ordered 2D sheets (Figure 1.5 g-h). While it should be possible to utilize disulfide bonds to construct 3D protein materials, this has yet to be realized.

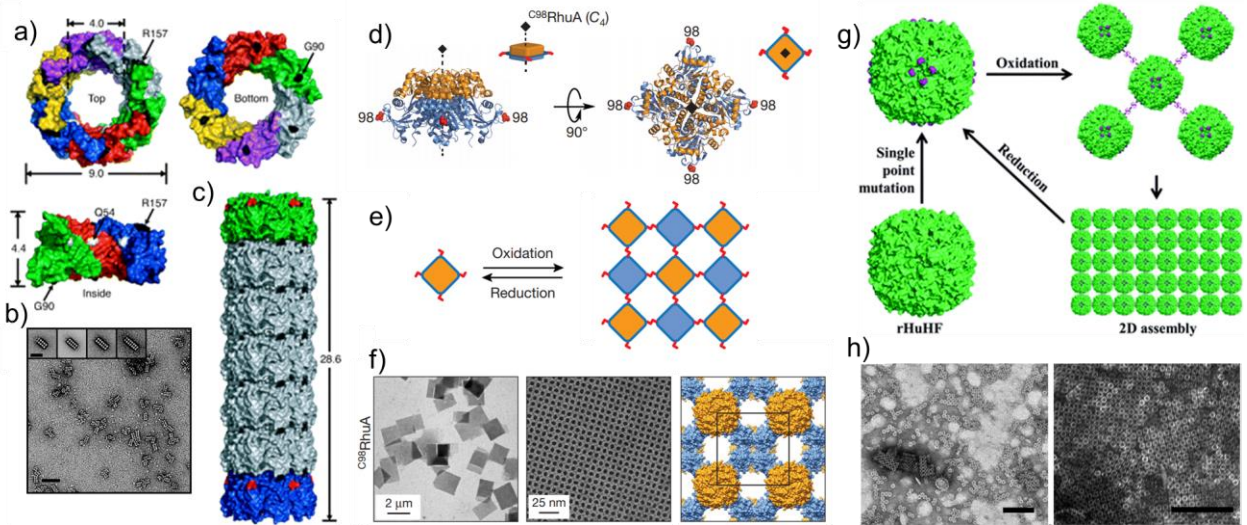


Figure 1.5. Disulfide mediated protein assemblies. a) Hexameric structure of Hcp1 that highlights cysteine sites. b) Negative stain TEM image of Hcp1 nanotubes. The scale bar is 45 nm. c) A model of Hcp1 with two capping subunits. d) tetrameric structure of RhuA that highlights cysteine sites. e) Schematic of RhuA assembly into a 2D protein sheet. f) Negative stain TEM images of RhuA (left and middle) and the proposed model (right). g) Schematic for 2D array formation from ferritin. h) Negative stain TEM images of 2D ferritin sheets after one day (left) and five days (right). The scale bars are 100 nm. Panels a, b, and c were adapted from reference 50. Panels d, e, and f were adapted from reference 53. Panels g and h were adapted from reference 54.

1.2.5. Metal mediated protein self-assembly

Metal ions are found in protein-protein interfaces throughout nature; ca. 5% of structurally characterized oligomers contain a metal ion or metallocofactor at an interfacial site.⁵⁵ Metal ions are most commonly anchored to the protein surface through a combination of Asp, Glu, Cys or His residues. Much like disulfide bonds, metal-mediated interactions are strong, directional, and reversible. Metal coordinating residues have been installed on surface-exposed sites to form crystal-contacts with transition metals in a process termed “metal-mediated synthetic symmetrization” (Figure 1.6 a).⁵⁶ This process led to the discovery of unique metal-mediated packing modes for the heavily studied lysozyme and maltose-binding protein. Metal binding sites have also been engineered to form precise interfacial crystal contacts. The structure of HuHF was first crystallized through the installation of a Ca^{2+} mediated crystal contact inspired by the Cd^{2+} dependent crystal contact in horse spleen ferritin (Figure 1.6 b).⁵⁷ His-tags form strong metal-binding motifs that are often employed during protein purification. An N-terminal His-tag was installed onto the small dimeric enzyme, *S. japonicum* glutathione S transferase, to induce the metal-mediated formation of 1D nanowires upon the addition of Ni^{2+} (Figure 1.6 c).⁵⁸ Similarly, metal-binding motifs can be engineered onto existing protein scaffolds. Highly-ordered 1D nanotubes and 2D sheets were synthesized from an engineered C_2 symmetric cb_{562} dimer coordinated to Zn^{2+} .⁵⁹ Each cb_{562} dimer has a pair of symmetric high-affinity tris-His sites and low-affinity sites containing a carboxylate and the carbonyl and amine of the N-terminus (Figure 1.6 d). At low Zn^{2+} concentrations, metal-induced nucleation occurs slowly, and 2D sheets formed. Increasing the Zn^{2+} concentration increases the number of nuclei and favors 1D nanotube assembly. A similar strategy involving a disulfide mediated D_2 symmetric cb_{562} tetramer was employed in the formation of a series of nanotubes with tailorable diameters.⁶⁰ The Zn^{2+} anchoring sites were comprised of a high-affinity tetracarboxylate site and a similar low-affinity site at the N-terminus (Figure 1.6 e). Non-natural light-responsive metal-coordinating ligands have also been employed in protein self-assembly. Cysteine residues were installed at the top and bottom face

GroEL, a chaperonin protein. These cysteines were site-specifically labeled with a spiropyrans-amended maleimide. Under UV irradiation, the spiropyran ring fragmented, exposing a negatively charged phenol that coordinates to divalent metals such as Mg^{2+} (Figure 1.6 f). These metals provide the connectivity for the GroEL interface, leading to the formation of 1D nanotubes. This process is reversible, and the nanotubes disassemble when exposed to visible light.⁶¹ Metal mediated interactions are incredibly powerful motifs for inducing self-assembly.

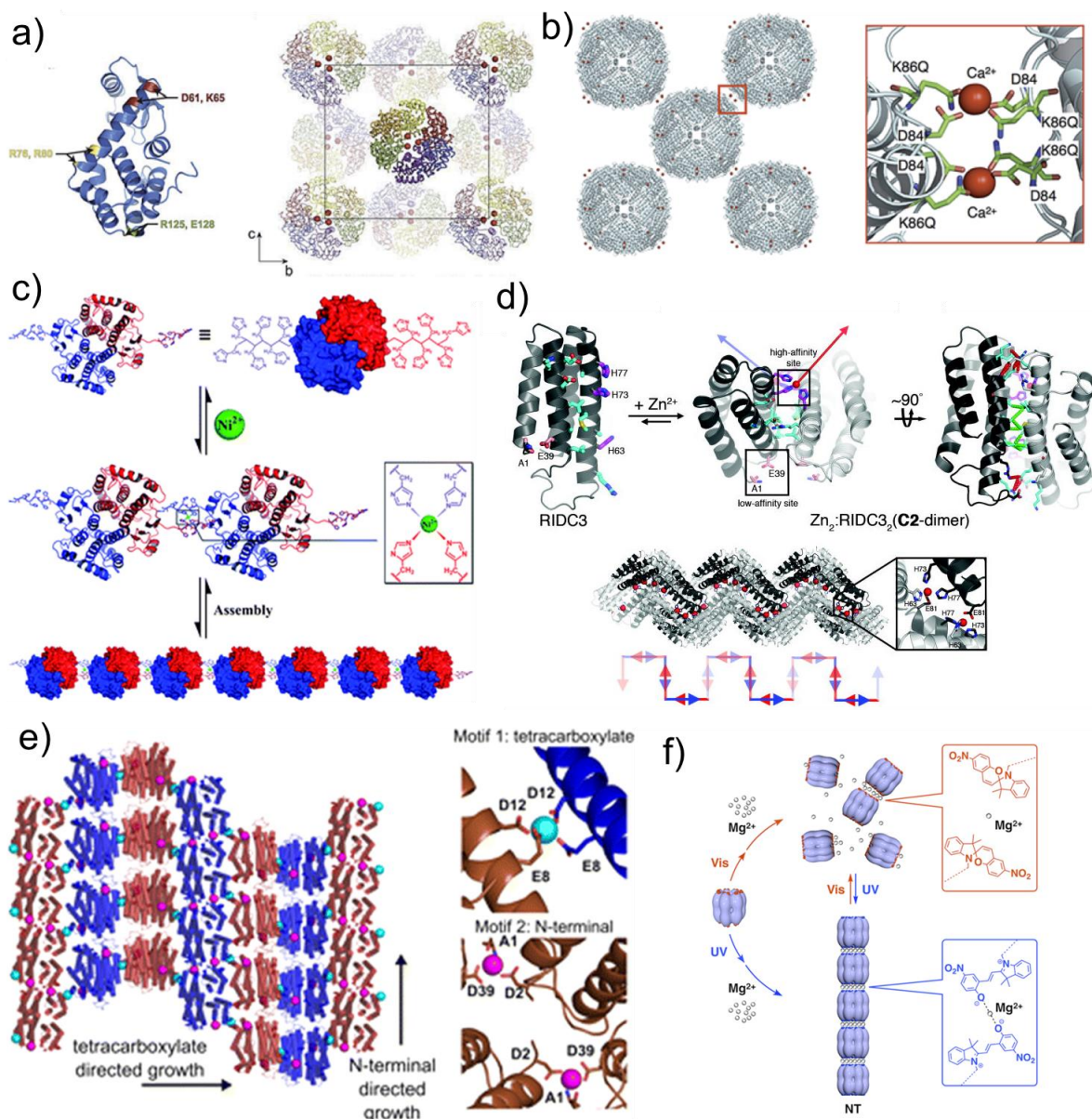


Figure 1.6. Metal-mediated protein self-assembly. a) Zn-mediated 3D crystal packing of a T4L mutant. b) Ca-mediated 3D crystal packing of HuHF. c) Ni-⁶²His assembly of glutathione S transferase nanotubes. d) Zinc mediated assembly of RIDC3 dimers into 2D sheets. e) Zinc mediated assembly of ⁹⁶C-RIDC3 tetramers into 2D sheets. f) UV light triggered the assembly and disassembly of GroEL-spiropyran mediated by Mg ions. Panels a and b originated from references 56 and 57 and were adapted from reference 36. Panel C was adapted from reference 58. Panel D was adapted from reference 59. Panel E was adapted from reference 60. Panel F originated from reference 61 and was adapted from reference 26.

1.2.6 Protein assemblies based on recognition elements

Specific recognition partners, such as evolved protein-cofactor pairs, synthetic guest-host complexes, or even complementary nucleotide sequences, have been used to assemble extended protein architectures. Cofactor-bound proteins in nature have evolved intricate binding pockets that enable the selective recognition and tight binding of a specific compound. For protein self-assembly, the three most common protein-cofactor pairs are lectin-sugar, heme-hemoprotein, and biotin-(strept)avidin. The lectin protein concanavalin A (ConA) is a pseudo-tetrahedral symmetric protein that tightly binds saccharides such as α -D-mannose and α -D-glucose. A synthetic ditopic bismannoside ligand, spaced with a short linker, connects two ConA subunits and templates the assembly of a 3D crystalline material.⁶² Inspired by this work, a polyvalent tetragalactosylated glycocluster was synthesized to promote the assembly of LecA, a similar tetrameric lectin protein.⁶³ Instead of interacting with one lectin binding site on each tetramer (like the bsimannoside ligand), the two sugars on each side of the glycocluster are spaced far enough apart to allow for coordination to two lectin-binding sites on a single LecA tetramer or one lectin-binding site on two LecA tetramers. The polyvalency enables the self-assembly of linear and branched chains mediated by the lectin-sugar binding interaction.

An alternative route to connect cofactor-based proteins involves the bioconjugation of the cofactor to the protein surface. Advances in protein bioconjugation methodology have enabled the attachment of cofactors to specific residues on the protein surface. Like most cofactor-bound proteins, the monomeric *E. coli* hemoprotein b_{562} has a high affinity and specificity for the covalently bound heme-cofactor. A synthetic heme derivative was covalently attached to a cysteine residue on the opposite side of the heme-binding pocket of b_{562} . After the removal of the native heme, the functionalized b_{562} -heme constructs self-assembled into linear head-to-tail chains linked through the synthetic heme (Figure 1.7 a).⁶⁴ Using the same construct, branched 1D chains were formed by introducing a trifunctional ligand displaying three heme cofactors.⁶⁵ A similar approach was employed to assemble linear myoglobin (another monomeric heme protein)

chains. Unlike the previous 1D assemblies, the myoglobin-based linear chains formed 3D aggregates driven through the dimerization of myoglobin by H_2O_2 .⁶⁶ The biotin-(strept)avidin pair, one of Nature's most robust noncovalent complexes, has also been utilized to form synthetic protein assemblies. (Strept)Avidin is a D_2 symmetric tetramer with a biotin binding pocket on each monomer. A symmetric biotin-based building block was constructed through the site-specific labeling of biotin onto a C_4 symmetric RhuA tetramer through thiol-disulfide exchange (Figure 1.7 b).⁶⁷ This building block forms extended 2D assemblies upon the addition of streptavidin.

Instead of relying on Nature for host-guest interactions, synthetic macrocyclic molecules such as crown ethers, cyclodextrins, calixarenes, and cucurbit[n]urils have been widely employed for the selective binding host molecules. Specifically, Cucurbit[8]uril (CB[8]) can accommodate a wide range of guests, including short peptide sequences such as TrpGlyGly or PheGlyGly in a 1:2 stoichiometry. This short PheGlyGly peptide was genetically installed on the N-terminus to the C_2 symmetric dimeric protein glutathione-S transferase (Figure 1.7 c).⁶⁸ Upon the addition of CB[8], 1D nanowires formed through the guest-host mediated interactions. The same strategy was used to construct Ca-responsive contractible nano-springs based on A Recoverin-GST fusion with the N-terminal FG sequence.⁶⁹ Despite the success of the protein-cofactor and guest-host interactions, the inability to modulate the interaction strength is a significant drawback for the formation of highly ordered arrays.

The recognition elements are not limited to guest-host interactions. DNA and RNA are particularly attractive recognition elements due to the highly specific Watson-Crick base pairing of complementary sequences. Previously, DNA-DNA interactions were used to construct ordered inorganic nanoparticles into predesigned 3D lattices using "sticky-end" assembly.⁷⁰ A similar design strategy was used to template multiple distinct 3D protein frameworks.⁷¹ Bovine and *C. gloeosporioides* catalase, two structurally similar C_4 symmetric tetrameric proteins, were functionalized with the lysine-specific NHS attached to DNA. The bovine and *C. gloeosporioides*

catalase had *ca.* 15 and 12 DNA strands per protein, respectively. The addition of complementary DNA strands leads to the formation of either a primitive cubic (CsCl) or body-centered cubic lattices for homo-catalase (bovine-bovine or cg-cg) assemblies, and exclusively primitive cubic lattices for heteromeric (bovine-cg) catalase assemblies (Figure 1.7 d). Unlike other recognition elements, the customizability of DNA allows for the design of incredibly intricate protein-based materials.

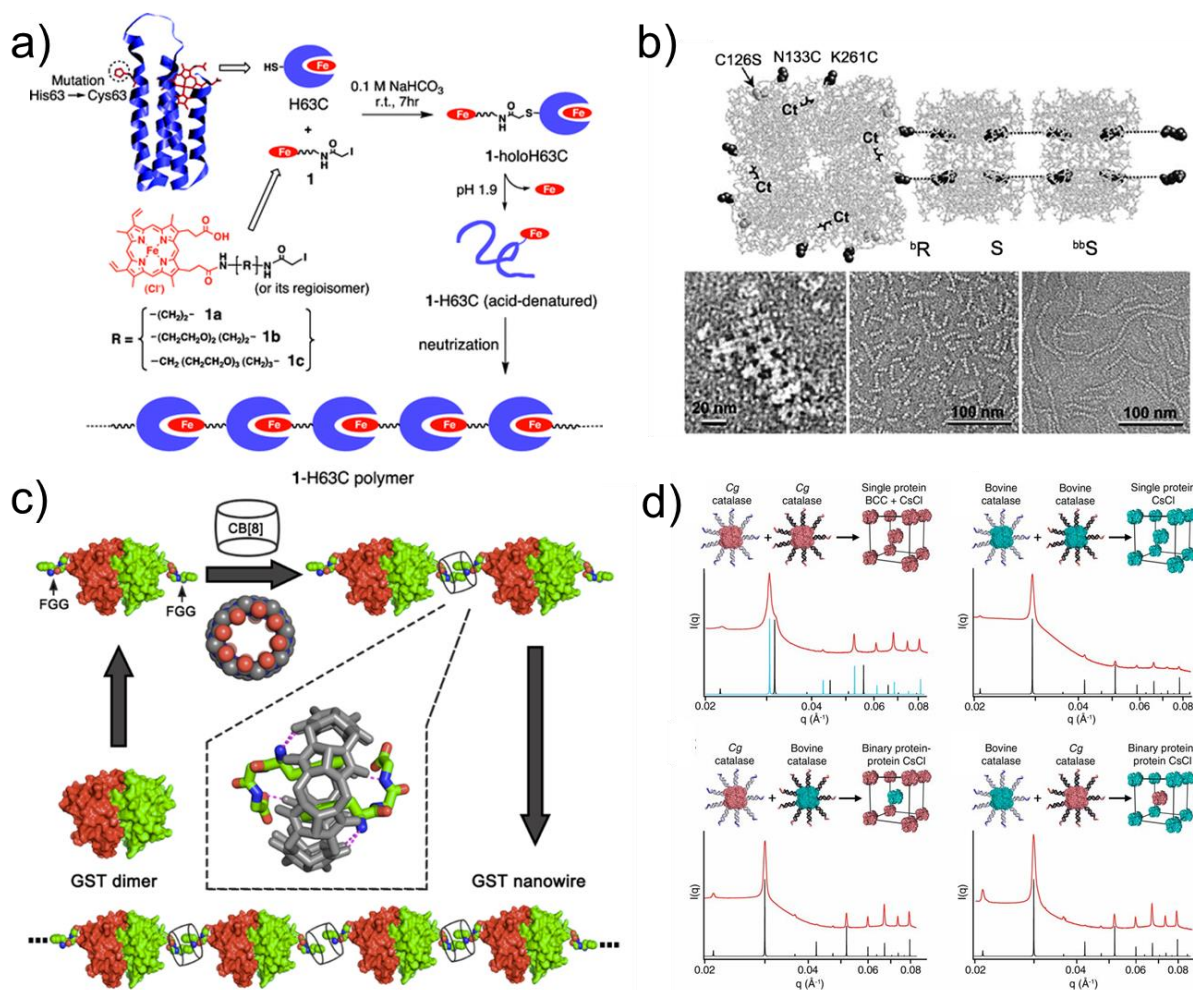


Figure 1.7. Protein assemblies based on recognition elements. a) Schematic for the formation of 1D cytochrome *b*₅₆₂ polymers. b) Representation of the extended chains of RhuA-biotin complexed with streptavidin. Negatively stained TEM images show the formation of short 1D chains. c) schematic for the formation of Glutathione-s-transferase-FGG connected by CB[8]. d) SAXS profiles (red) and predicted patterns (teal and black) of the DNA-mediated lattices formed with *C. gloeosporioides* catalase and bovine catalase. Panel A was adapted from reference 64. Panel B and C originated from reference 67 and 68 and were adapted from reference 26. Panel D was adapted from reference 71.

1.2.7. Protein self-assembly using multiple interaction motifs

The construction of complex pseudo-infinite arrays often requires the design of numerous unique and highly specific interfaces. In these cases, it is often beneficial to incorporate multiple types of interfacial interactions to ensure specificity and provide additional control over self-assembly. The majority of these multi-faceted protein assemblies pair recognition elements with other binding functionalities. As discussed previously, the main drawback to most specific recognition elements is the inability to modulate the strength of the interaction. The self-assembly of highly ordered architectures requires reversible binding functionalities to allow for the correction of defects during nucleation and growth. One example of this is the combination of a bis-biotin group (which coordinates to a single streptavidin) and a tridentate terpyridine (which forms a homodimeric complex with Fe^{2+}). This ligand, when combined with streptavidin and Fe^{2+} , drives the assembly of 1D chains (Figure 1.8 a). A similar strategy can be employed by replacing the terpyridine with a heme.⁷² The hemoprotein (myoglobin) binds the heme, forming a discrete ligand-mediated myoglobin-streptavidin-myoglobin complex. A single Cys mutation was introduced on the back of the myoglobin to promote the assembly of a disulfide-mediated myoglobin-myoglobin C_2 symmetric dimer. When all three components (streptavidin, myoglobin, and heme-bis-biotin) were combined, extended 1D heteroprotein chains self-assembled (Figure 1.8 b).

In addition to metal and disulfide mediated assembly, homodimerizing ligand domains provide additional control of protein self-assembly. Rhodamine, which forms a homodimer based on π - π stacking, was appended site-specifically onto a cysteine that extends above and below the previously described Zn^{2+} -mediated RIDC3 sheets (Figure 1.8 c).⁵⁹ The dimerization of the rhodamine templates the ordered stacking of these 2D sheets in the 3rd dimension. Similar strategies involving rhodamine have been utilized to form lectin-based protein materials. The C_4 symmetric LecA protein has been assembled into a series of 1- and 2D ordered arrays through

the selective coordination of a galactose-rhodamine ligand to the lectin binding pocket, followed by rhodamine dimerization.⁷³ This strategy was again applied with the C_4 symmetric tetrameric soybean agglutinin (SBA) to form highly ordered nanotubes.⁷⁴ A similar rhodamine-mannopyranoside ligand was employed to assemble the pseudo-tetrahedral ConA into 3D lattices.⁷⁵ Unlike the previous bis-mannose ligand, the slower rhodamine dimerization allows for self-correction and the formation of highly ordered 3D crystals (Figure 1.8 d).⁶² The length of the spacer between the rhodamine and mannopyranoside was systematically varied. Short linker lengths favor the formation of a relatively porous $P2_1$ lattice. When the spacer length was increased, an interpenetrated crystal lattice formed with nearly twice the packing density (39% vs. 69%). These crystal lattices are a rare example of modularity in protein biomaterial design.

Short strands of complementary DNA can be used in place of weakly-associating homodimeric ligands to form similar heterodimeric complexes. Short DNA strands were site-specifically labeled onto the cysteine-bearing RIDC3 that assembles into Zn^{2+} -mediated RIDC3 sheets. It was expected that the complementary DNA strands would template stacking of the 2D layers similar to rhodamine. However, a highly sophisticated 3D architecture was induced by the DNA pairing (Figure 1.8 e).⁷⁶ The self-assembly of this crystalline material is predicated on a balance between three types of intramolecular interactions: metal-protein, Watson-Crick base pairing, and nucleic acid-protein interactions. Based on the examples above, combining multiple types of functionalities for interfacial assembly leads to increased complexity and improved multidimensional order of the protein architectures.

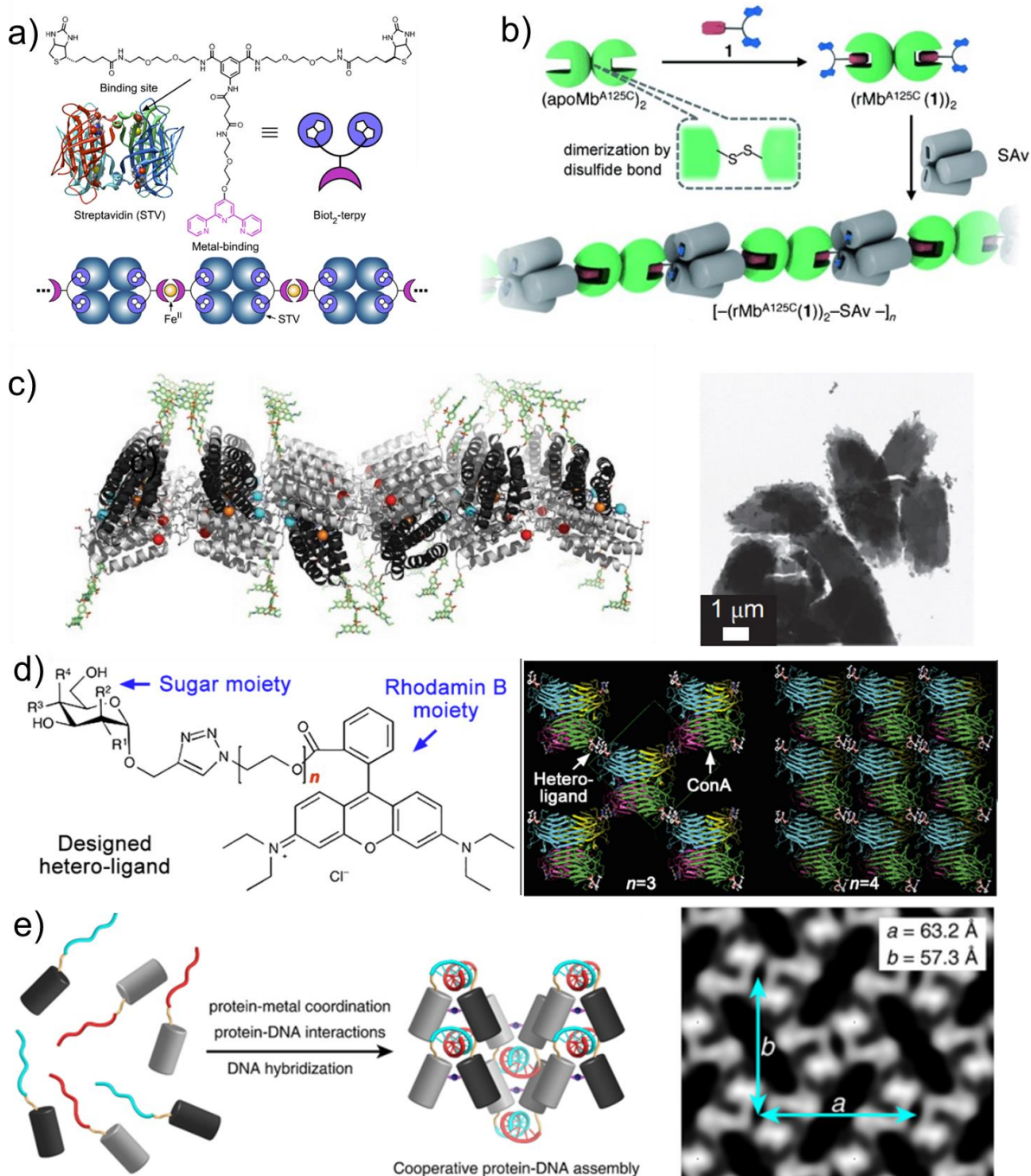


Figure 1.8. Protein assemblies with multiple interaction motifs. a) Schematic for bis-biotin-terphenyl assembly of streptavidin. b) schematic for the assembly of disulfide mediated myoglobin dimers with streptavidin. c) Rhodamine dimerization leads to RIDC3 sheet stacking in the 3rd dimension. d) Rhodamine dimerization leads to the formation of two distinct crystal lattices depending on the linker length. e) DNA-DNA, Protein-DNA and Protein-metal interactions are all necessary for the formation of this 3D assembly. Panel A originated from reference 72 and was adapted from reference 26. Panel B is adapted from reference 72. Panel C is adapted from reference 59. Panel D is adapted from reference 75. Panel E is adapted from reference 76.

1.3. Developing a modular methodology for protein self-assembly

Over the last twenty years, a myriad of strategies have emerged for the construction of highly ordered protein biomaterials. It is necessary to have access to a variety of design strategies since proteins are structurally and chemically diverse biomolecules. The current design strategies do an admirable job for the prediction and construction of 1D extended structures. Linear assemblies are the most straightforward extended architectures and require only two unique interfacial connections across one dimension. The formation of predictable multidimensional periodic arrays is considerably more challenging. Both the genetic fusion symmetrization and noncovalent interface redesign methodologies have significantly high fail rates for multidimensional materials.^{28, 46} The reactive metal-binding residues commonly found on the protein surface can form unexpected chelating motifs, leading to unintended metal-mediated constructs. Lattices based on electrostatic interactions are often poorly ordered since they are connected by a series of weak, nondirectional bonds. Most assemblies based on recognition elements suffer from the opposite problem, where the extremely tight guest-host interactions prevent self-correction during lattice formation. While disulfide-based lattice formation seems to avoid many potential pitfalls due to highly specific and reversible bonds, there are no examples of ordered 3D disulfide-mediated macromolecular assemblies. In addition to a lack of *de novo* predictability, only a few of these strategies are constructed with interchangeable components.

In contrast, coordination polymers such as metal-organic frameworks (MOFs) benefit from a high degree of modularity, where component nodes/struts can be interchanged to produce structurally unique architectures with vastly different materials properties. The wide-ranging applicability directly stems from their modular construction and *de-novo* predictability. In pursuit of a more generalizable assembly platform, we developed a unique method for the formation of crystalline protein arrays based on the construction principles of metal-organic frameworks. These protein-MOFs form highly crystalline structures with a series of interchangeable components. The

symmetry of these protein-MOFs is dependent on both the symmetry of the organic bridging ligand and the coordination preferences of the metal node.⁷⁷⁻⁷⁸ This dissertation describes the design, characterization, and emergent functionality of protein-MOFs as a more modular design strategy for supramolecular protein assembly.

1.4. References

1. Pieters, B. J. G. E.; van Eldijk, M. B.; Nolte, R. J. M.; Mecnovic, J., Natural supramolecular protein assemblies. *Chemical Society Reviews* **2016**, *45* (1), 24-39.
2. Goodsell, D. S.; Olson, A. J., Structural symmetry and protein function. *Annual Review of Biophysics and Biomolecular Structure* **2000**, *29*, 105-153.
3. Hu, J. C.; Sauer, R. T., The Basic-Region Leucine-Zipper Family of DNA Binding Proteins. In *Nucleic Acids and Molecular Biology*, Eckstein, F.; Lilley, D. M. J., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 1992; pp 82-101.
4. Yoshida, M.; Muneyuki, E.; Hisabori, T., ATP synthase — a marvellous rotary engine of the cell. *Nature Reviews Molecular Cell Biology* **2001**, *2* (9), 669-677.
5. Mateu, M. G., Assembly, stability and dynamics of virus capsids. *Archives of Biochemistry and Biophysics* **2013**, *531* (1), 65-79.
6. Steitz, T. A., A structural understanding of the dynamic ribosome machine. *Nature Reviews Molecular Cell Biology* **2008**, *9* (3), 242-253.
7. Shoulders, M. D.; Raines, R. T., Collagen structure and stability. *Annu Rev Biochem* **2009**, *78*, 929-958.
8. Lodish, H.; Berk, A.; Zipursky, L. S.; Matsudaira, P.; Baltimore, D., Darnell, James, *Molecular Cell Biology*, 4th edition. W. H. Freeman: New York, 2000.
9. Fagan, R. P.; Fairweather, N. F., Biogenesis and functions of bacterial S-layers. *Nature Reviews Microbiology* **2014**, *12* (3), 211-222.
10. von der Ecken, J.; Müller, M.; Lehman, W.; Manstein, D. J.; Penczek, P. A.; Raunser, S., Structure of the F-actin–tropomyosin complex. *Nature* **2014**, *519*, 114.
11. Chou, S. Z.; Pollard, T. D., Mechanism of actin polymerization revealed by cryo-EM structures of actin filaments with three different bound nucleotides. *Proceedings of the National Academy of Sciences* **2019**, *116* (10), 4265-4274.
12. Oda, T.; Iwasa, M.; Aihara, T.; Maéda, Y.; Narita, A., The nature of the globular- to fibrous-actin transition. *Nature* **2009**, *457* (7228), 441-445.
13. Brouhard, G. J.; Rice, L. M., Microtubule dynamics: an interplay of biochemistry and mechanics. *Nature Reviews Molecular Cell Biology* **2018**, *19* (7), 451-463.
14. Ilk, N.; Egelseer, E. M.; Sleytr, U. B., S-layer fusion proteins - construction principles and applications. *Current Opinion in Biotechnology* **2011**, *22* (6), 824-831.
15. Pum, D.; Toca-Herrera, J. L.; Sleytr, U. B., S-layer protein self-assembly. *Int J Mol Sci* **2013**, *14* (2), 2484-2501.
16. Gerbino, E.; Carasi, P.; Mobili, P.; Serradell, M. A.; Gómez-Zavaglia, A., Role of S-layer proteins in bacteria. *World Journal of Microbiology and Biotechnology* **2015**, *31* (12), 1877-1887.

17. Ahnert, S. E.; Marsh, J. A.; Hernandez, H.; Robinson, C. V.; Teichmann, S. A., Principles of assembly reveal a periodic table of protein complexes. *Science* **2015**, *350* (6266), 11.
18. Doye, J. P. K.; Poon, W. C. K., Protein crystallization in vivo. *Current Opinion in Colloid & Interface Science* **2006**, *11* (1), 40-46.
19. Dodson, G.; Steiner, D., The role of assembly in insulin's biosynthesis. *Current Opinion in Structural Biology* **1998**, *8* (2), 189-194.
20. Georgiou, D. K.; Vekilov, P. G., A fast response mechanism for insulin storage in crystals may involve kink generation by association of 2D clusters. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103* (6), 1681.
21. Coulibaly, F.; Chiu, E.; Ikeda, K.; Gutmann, S.; Haebel, P. W.; Schulze-Briese, C.; Mori, H.; Metcalf, P., The molecular organization of cypovirus polyhedra. *Nature* **2007**, *446* (7131), 97-101.
22. Vonck, J.; van Bruggen, E. F., Architecture of peroxisomal alcohol oxidase crystals from the methylotrophic yeast *Hansenula polymorpha* as deduced by electron microscopy. *J Bacteriol* **1992**, *174* (16), 5391-5399.
23. Sleytr, U. B.; Schuster, B.; Egelseer, E.-M.; Pum, D., S-layers: principles and applications. *FEMS Microbiol Rev* **2014**, *38* (5), 823-864.
24. Hartje, L. F.; Snow, C. D., Protein crystal based materials for nanoscale applications in medicine and biotechnology. *WIREs Nanomedicine and Nanobiotechnology* **2019**, *11* (4), e1547.
25. Kuan, S. L.; Bergamini, F. R. G.; Weil, T., Functional protein nanostructures: a chemical toolbox. *Chemical Society Reviews* **2018**, *47* (24), 9069-9105.
26. Luo, Q.; Hou, C.; Bai, Y.; Wang, R.; Liu, J., Protein Assembly: Versatile Approaches to Construct Highly Ordered Nanostructures. *Chemical Reviews* **2016**, *116* (22), 13571-13632.
27. Hamley, I. W., Protein Assemblies: Nature-Inspired and Designed Nanostructures. *Biomacromolecules* **2019**, *20* (5), 1829-1848.
28. Sinclair, J. C.; Davies, K. M.; Venien-Bryan, C.; Noble, M. E. M., Generation of protein lattices by fusing proteins with matching rotational symmetry. *Nature Nanotechnology* **2011**, *6* (9), 558-562.
29. Kostianen, M. A.; Hiekkataipale, P.; Laiho, A.; Lemieux, V.; Seitsonen, J.; Ruokolainen, J.; Ceci, P., Electrostatic assembly of binary nanoparticle superlattices using protein cages. *Nature Nanotechnology* **2013**, *8* (1), 52-+.
30. Padilla, J. E.; Colovos, C.; Yeates, T. O., Nanohedra: Using symmetry to design self assembling protein cages, layers, crystals, and filaments. *Proceedings of the National Academy of Sciences of the United States of America* **2001**, *98* (5), 2217-2221.
31. Salgado, E. N.; Ambroggio, X. I.; Brodin, J. D.; Lewis, R. A.; Kuhlman, B.; Tezcan, F. A., Metal templated design of protein interfaces. *Proceedings of the National Academy of Sciences* **2010**, *107* (5), 1827.

32. King, N. P.; Sheffler, W.; Sawaya, M. R.; Vollmar, B. S.; Sumida, J. P.; Andre, I.; Gonen, T.; Yeates, T. O.; Baker, D., Computational Design of Self-Assembling Protein Nanomaterials with Atomic Level Accuracy. *Science* **2012**, 336 (6085), 1171-1174.
33. Bai, Y.; Luo, Q.; Liu, J., Protein self-assembly via supramolecular strategies. *Chemical Society Reviews* **2016**, 45 (10), 2756-2767.
34. Lai, Y.-T.; King, N. P.; Yeates, T. O., Principles for designing ordered protein assemblies. *Trends in Cell Biology* **2012**, 22 (12), 653-661.
35. Kobayashi, N.; Yanase, K.; Sato, T.; Unzai, S.; Hecht, M. H.; Arai, R., Self-Assembling Nano-Architectures Created from a Protein Nano-Building Block Using an Intermolecularly Folded Dimeric de Novo Protein. *Journal of the American Chemical Society* **2015**, 137 (35), 11285-11293.
36. Abe, S.; Ueno, T., Design of protein crystals in the development of solid biomaterials. *Rsc Advances* **2015**, 5 (27), 21366-21375.
37. Liljestrom, V.; Mikkila, J.; Kostianen, M. A., Self-assembly and modular functionalization of three-dimensional crystals from oppositely charged proteins. *Nature Communications* **2014**, 5.
38. Liljestrom, V.; Seitsonen, J.; Kostianen, M. A., Electrostatic Self-Assembly of Soft Matter Nanoparticle Cocrystals with Tunable Lattice Parameters. *Acs Nano* **2015**, 9 (11), 11278-11285.
39. Sun, H.; Miao, L.; Li, J.; Fu, S.; An, G.; Si, C.; Dong, Z.; Luo, Q.; Yu, S.; Xu, J.; Liu, J., Self-Assembly of Cricoid Proteins Induced by "Soft Nanoparticles": An Approach To Design Multienzyme-Cooperative Antioxidative Systems. *Acs Nano* **2015**, 9 (5), 5461-5469.
40. Sun, H.; Zhang, X.; Miao, L.; Zhao, L.; Luo, Q.; Xu, J.; Liu, J., Micelle-Induced Self-Assembling Protein Nanowires: Versatile Supramolecular Scaffolds for Designing the Light-Harvesting System. *Acs Nano* **2016**, 10 (1), 421-428.
41. Yang, R.; Chen, L.; Zhang, T.; Yang, S.; Leng, X.; Zhao, G., Self-assembly of ferritin nanocages into linear chains induced by poly(alpha, L-lysine). *Chemical Communications* **2014**, 50 (4), 481-483.
42. Kunzle, M.; Eckert, T.; Beck, T., Binary Protein Crystals for the Assembly of Inorganic Nanoparticle Superlattices. *Journal of the American Chemical Society* **2016**, 138 (39), 12731-12734.
43. Bellapadrona, G.; Elbaum, M., Supramolecular Protein Assemblies in the Nucleus of Human Cells. *Angewandte Chemie-International Edition* **2014**, 53 (6), 1534-1537.
44. Sharp, T. H.; Bruning, M.; Mantell, J.; Sessions, R. B.; Thomson, A. R.; Zaccai, N. R.; Brady, R. L.; Verkade, P.; Woolfson, D. N., Cryo-transmission electron microscopy structure of a gigadalton peptide fiber of de novo design. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, 109 (33), 13266-13271.
45. Lanci, C. J.; MacDermaid, C. M.; Kang, S. G.; Acharya, R.; North, B.; Yang, X.; Qiu, X. J.; DeGrado, W. F.; Saven, J. G., Computational design of a protein crystal. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, 109 (19), 7304-9.

46. Gonen, S.; DiMaio, F.; Gonen, T.; Baker, D., Design of ordered two-dimensional arrays mediated by noncovalent protein-protein interfaces. *Science* **2015**, *348* (6241), 1365-1368.
47. Chen, Z. B.; Johnson, M. C.; Chen, J. J.; Bick, M. J.; Boyken, S. E.; Lin, B. H.; De Yoreo, J. J.; Kollman, J. M.; Baker, D.; DiMaio, F., Self-Assembling 2D Arrays with de Novo Protein Building Blocks. *Journal of the American Chemical Society* **2019**, *141* (22), 8891-8895.
48. Zhou, K.; Zang, J.; Chen, H.; Wang, W.; Wang, H.; Zhao, G., On-Axis Alignment of Protein Nanocage Assemblies from 2D to 3D through the Aromatic Stacking Interactions of Amino Acid Residues. *ACS Nano* **2018**, *12* (11), 11323-11332.
49. Banatao, D. R.; Cascio, D.; Crowley, C. S.; Fleissner, M. R.; Tienson, H. L.; Yeates, T. O., An approach to crystallizing proteins by synthetic symmetrization. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103* (44), 16230-16235.
50. Ballister, E. R.; Lai, A. H.; Zuckermann, R. N.; Cheng, Y.; Mougous, J. D., In vitro self-assembly from a simple protein of tailorable nanotubes building block. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105* (10), 3733-3738.
51. Miranda, F. F.; Iwasaki, K.; Akashi, S.; Sumitomo, K.; Kobayashi, M.; Yamashita, I.; Tame, J. R. H.; Hedde, J. G., A Self-Assembled Protein Nanotube with High Aspect Ratio. *Small* **2009**, *5* (18), 2077-2084.
52. Nagano, S.; Banwell, E. F.; Iwasaki, K.; Michalak, M.; Pałka, R.; Zhang, K. Y. J.; Voet, A. R. D.; Hedde, J. G., Understanding the Assembly of an Artificial Protein Nanotube. *Advanced Materials Interfaces* **2016**, *3* (24), 1600846.
53. Suzuki, Y.; Cardone, G.; Restrepo, D.; Zavattieri, P. D.; Baker, T. S.; Tezcan, F. A., Self-assembly of coherently dynamic, auxetic, two-dimensional protein crystals. *Nature* **2016**, *533* (7603), 369-+.
54. Zhou, K.; Chen, H.; Zhang, S.; Wang, Y.; Zhao, G., Disulfide-mediated reversible two-dimensional self-assembly of protein nanocages. *Chemical Communications* **2019**, *55* (52), 7510-7513.
55. Song, W. J.; Tezcan, F. A., A designed supramolecular protein assembly with in vivo enzymatic activity. *Science* **2014**, *346* (6216), 1525-1528.
56. Laganowsky, A.; Zhao, M.; Soriaga, A. B.; Sawaya, M. R.; Cascio, D.; Yeates, T. O., An approach to crystallizing proteins by metal-mediated synthetic symmetrization. *Protein Sci* **2011**, *20* (11), 1876-1890.
57. Lawson, D. M.; Artymiuk, P. J.; Yewdall, S. J.; Smith, J. M.; Livingstone, J. C.; Treffry, A.; Luzzago, A.; Levi, S.; Arosio, P.; Cesareni, G.; et al., Solving the structure of human H ferritin by genetically engineering intermolecular crystal contacts. *Nature* **1991**, *349* (6309), 541-4.
58. Zhang, W.; Luo, Q.; Miao, L.; Hou, C.; Bai, Y.; Dong, Z.; Xu, J.; Liu, J., Self-assembly of glutathione S-transferase into nanowires. *Nanoscale* **2012**, *4* (19), 5847-5851.

59. Brodin, J. D.; Ambroggio, X. I.; Tang, C.; Parent, K. N.; Baker, T. S.; Tezcan, F. A., Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. *Nature Chemistry* **2012**, *4* (5), 375-382.
60. Brodin, J. D.; Smith, S. J.; Carr, J. R.; Tezcan, F. A., Designed, Helical Protein Nanotubes with Variable Diameters from a Single Building Block. *Journal of the American Chemical Society* **2015**, *137* (33), 10468-10471.
61. Biswas, S.; Kinbara, K.; Oya, N.; Ishii, N.; Taguchi, H.; Aida, T., A Tubular Biocontainer: Metal Ion-Induced 1D Assembly of a Molecularly Engineered Chaperonin. *Journal of the American Chemical Society* **2009**, *131* (22), 7556-+.
62. Dotan, N.; Arad, D.; Frolow, F.; Freeman, A., Self-assembly of a tetrahedral lectin into predesigned diamondlike protein crystals. *Angewandte Chemie-International Edition* **1999**, *38* (16), 2363-2366.
63. Sicard, D.; Cecioni, S.; Iazykov, M.; Chevolut, Y.; Matthews, S. E.; Praly, J.-P.; Souteyrand, E.; Imbert, A.; Vidal, S.; Phaner-Goutorbe, M., AFM investigation of *Pseudomonas aeruginosa* lectin LecA (PA-IL) filaments induced by multivalent glycoclusters. *Chemical Communications* **2011**, *47* (33), 9483-9485.
64. Kitagishi, H.; Oohora, K.; Yamaguchi, H.; Sato, H.; Matsuo, T.; Harada, A.; Hayashi, T., Supramolecular hemoprotein linear assembly by successive interprotein heme-heme pocket interactions. *Journal of the American Chemical Society* **2007**, *129* (34), 10326-+.
65. Kitagishi, H.; Kakikura, Y.; Yamaguchi, H.; Oohora, K.; Harada, A.; Hayashi, T., Self-Assembly of One- and Two-Dimensional Hemoprotein Systems by Polymerization through Heme-Heme Pocket Interactions. *Angewandte Chemie-International Edition* **2009**, *48* (7), 1271-1274.
66. Oohora, K.; Onoda, A.; Kitagishi, H.; Yamaguchi, H.; Harada, A.; Hayashi, T., A chemically-controlled supramolecular protein polymer formed by a myoglobin-based self-assembly system. *Chemical Science* **2011**, *2* (6), 1033-1038.
67. Ringler, P.; Schulz, G. E., Self-assembly of proteins into designed networks. *Science* **2003**, *302* (5642), 106-109.
68. Hou, C.; Li, J.; Zhao, L.; Zhang, W.; Luo, Q.; Dong, Z.; Xu, J.; Liu, J., Construction of Protein Nanowires through Cucurbit 8 uril-based Highly Specific HostGuest Interactions: An Approach to the Assembly of Functional Proteins. *Angewandte Chemie-International Edition* **2013**, *52* (21), 5590-5593.
69. Si, C.; Li, J.; Luo, Q.; Hou, C.; Pan, T.; Li, H.; Liu, J., An ion signal responsive dynamic protein nano-spring constructed by high ordered host-guest recognition. *Chemical Communications* **2016**, *52* (14), 2924-2927.
70. Mirkin, C. A.; Letsinger, R. L.; Mucic, R. C.; Storhoff, J. J., A DNA-based method for rationally assembling nanoparticles into macroscopic materials. *Nature* **1996**, *382* (6592), 607-609.

71. Brodin, J. D.; Auyeung, E.; Mirkin, C. A., DNA-mediated engineering of multicomponent enzyme crystals. *Proceedings of the National Academy of Sciences of the United States of America* **2015**, *112* (15), 4564-4569.
72. Oohora, K.; Burazerovic, S.; Onoda, A.; Wilson, Y. M.; Ward, T. R.; Hayashi, T., Chemically Programmed Supramolecular Assembly of Hemoprotein and Streptavidin with Alternating Alignment. *Angewandte Chemie-International Edition* **2012**, *51* (16), 3818-3821.
73. Yang, G.; Ding, H.-m.; Kochovski, Z.; Hu, R.; Lu, Y.; Ma, Y.-q.; Chen, G.; Jiang, M., Highly Ordered Self-Assembly of Native Proteins into 1D, 2D, and 3D Structures Modulated by the Tether Length of Assembly-Inducing Ligands. *Angew. Chem. Int. Ed. Engl.* **2017**, *56* (36), 10691-10695.
74. Yang, G.; Zhang, X.; Kochovski, Z.; Zhang, Y.; Dai, B.; Sakai, F.; Jiang, L.; Lu, Y.; Ballauff, M.; Li, X., Precise and Reversible Protein-Microtubule-Like Structure with Helicity Driven by Dual Supramolecular Interactions. *Journal of the American Chemical Society* **2016**, *138* (6), 1932-1937.
75. Sakai, F.; Yang, G.; Weiss, M. S.; Liu, Y.; Chen, G.; Jiang, M., Protein crystalline frameworks with controllable interpenetration directed by dual supramolecular interactions. *Nature Communications* **2014**, *5*.
76. Subramanian, R. H.; Smith, S. J.; Alberstein, R. G.; Bailey, J. B.; Zhang, L.; Cardone, G.; Suominen, L.; Chami, M.; Stahlberg, H.; Baker, T. S.; Tezcan, F. A., Self-Assembly of a Designed Nucleoprotein Architecture through Multimodal Interactions. *ACS Central Sci.* **2018**, *4* (11), 1578-1586.
77. Sontz, P. A.; Bailey, J. B.; Ahn, S.; Tezcan, F. A., A Metal Organic Framework with Spherical Protein Nodes: Rational Chemical Design of 3D Protein Crystals. *Journal of the American Chemical Society* **2015**, *137* (36), 11598-11601.
78. Bailey, J. B.; Zhang, L.; Chiong, J. A.; Ahn, S.; Tezcan, F. A., Synthetic Modularity of Protein-Metal-Organic Frameworks. *Journal of the American Chemical Society* **2017**, *139* (24), 8160-8166.

Chapter 2. Design, synthesis, and characterization of the first protein-MOF

2.1. Introduction

Design of crystalline materials with tunable structural, chemical, and physical properties represents a significant goal common to numerous scientific disciplines.¹⁻⁴ Proteins are especially attractive building blocks for crystalline materials as they provide high chemical and structural diversity and possess inherent functions such as catalysis, electron transfer, and molecular recognition. Three-dimensional (3D) protein lattices have been proposed as stable, porous scaffolds for carrying out catalytic reactions under harsh, non-biological conditions.⁵⁻⁶ Although effective methodologies have been developed for facilitating protein crystallization,⁷⁻⁸ successes in obtaining 3D protein crystals by design have been rare.⁹⁻¹¹ In contrast, the ability to rationally engineer crystalline materials from small organic and inorganic building blocks, though still challenging, is considerably more advanced.^{1, 12-14} A prominent class of these crystalline materials is metal-organic frameworks (MOFs), which have attracted much attention due their applications in separation, storage, and catalysis, etc.^{12, 15-23} The wide-ranging applications of MOFs are primarily fueled by their modular components; MOFs are comprised of metal nodes and bridging organic ligands that form highly ordered, regular network structures. Each node is composed of either a single metal atom or, more commonly, a polynuclear cluster of metal atoms, also known as a secondary building unit (SBU).²⁴⁻²⁵ We hypothesized that a symmetric protein could act as the SBU and form a porous 3D protein framework through both metal- and organic ligand-directed interactions.

The lattice symmetry of MOFs is dictated by the combination of the inherent symmetries of both the metal nodes and the organic bridging ligands.²⁴⁻²⁵ We envisioned a system where the metal node is a symmetrical protein that displays stable, surface-anchored metal ions. This hybrid protein-metal unit could then self-assemble into the desired 3D lattice via the coordination of symmetric ditopic organic ligand (Figure 2.1). When selecting an initial protein scaffold to

appropriate as the node, many criteria must be considered. The structure of the chosen protein should be well characterized and possess higher-order symmetry (i.e., tetragonal or octahedral). The symmetry axes of these scaffolds extend in 3D and are ideal locations for the installation of metal anchoring sites. Utilization of a symmetry axis also satisfies the requirement for multiple coordinating residues, which can lead to high-fidelity metal coordination. A single residue would have significant rotational freedom and would not be able to anchor a metal to the protein surface. The designed coordination environment needs to position the metal to expose an outward-facing open coordination site. This site needs to be near the surface (to allow for ligand coordination) but slightly recessed (to prevent protein-protein dimerization without the bridging ligand). Additionally, ease of protein expression/purification and the stability of the protein node need to be considered.

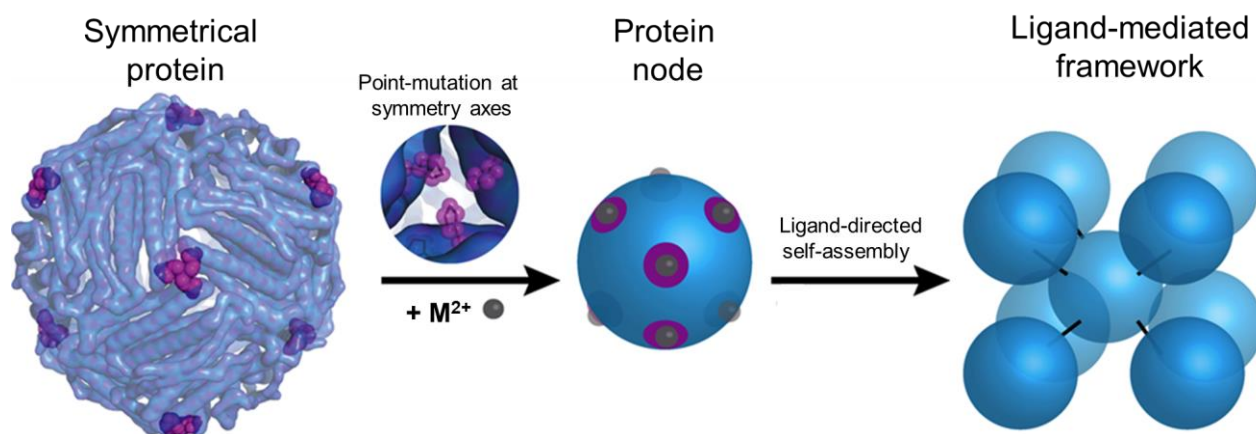


Figure 2.1. Scheme for metal/ligand-directed self-assembly of ferritin into 3D crystals. Surface-exposed binding sites for Zn²⁺ (grey spheres) are engineered at the selected symmetry axes (magenta). In the presence of ditopic organic bridging ligands, the resulting ¹²²Hferritin variant is expected to form a ligand-mediated crystalline framework.

2.2. Results and discussion

2.2.1. Selection of ferritin as the protein node

As an appropriate protein scaffold, we selected the highly stable human heavy-chain ferritin, which is a spherical, cage-like protein with octahedral (432) symmetry.²⁶⁻²⁷ This symmetry, resembling that of the metal nodes in some quintessential MOFs (e.g., MOF-5,²⁸ HKUST-1,¹⁶ MIL-53²⁹), provides several surface sites lying at or near the C_4 , C_3 , and C_2 rotation axes which could potentially be engineered for anchoring metal ions. Lattice connectivity through each of these axes would lead to a primitive, body-centered, or face-centered framework, respectively (Figure 2.2). The selected ΔC^* HuHF variant contains the following mutations, which eliminate all cysteine residues (C90E, C102A, C130A), removing the possibility of intramolecular disulfide bond formation.³⁰ Additionally, the K86Q mutation was included to form a surface-exposed metal anchoring site with 84D that promotes metal-mediated crystallization into a face-centered cubic (*fcc*) lattice.³¹ This crystal packing interaction sits in the middle of a highly ordered loop that spans the entirety of the 2-fold axis. Even though there are a few other locations for the installation of metal anchoring sites along this loop (i.e., positions 84, 87, and 91), it is likely that these selected mutations are too surface exposed and would eschew ligand binding in favor of the formation of metal-mediated protein-protein crystal contacts. The problems associated with surface-exposed metal coordination sites highlights the need for a slightly recessed binding site to necessitate metal-ligand-metal interaction during protein-MOF self-assembly.

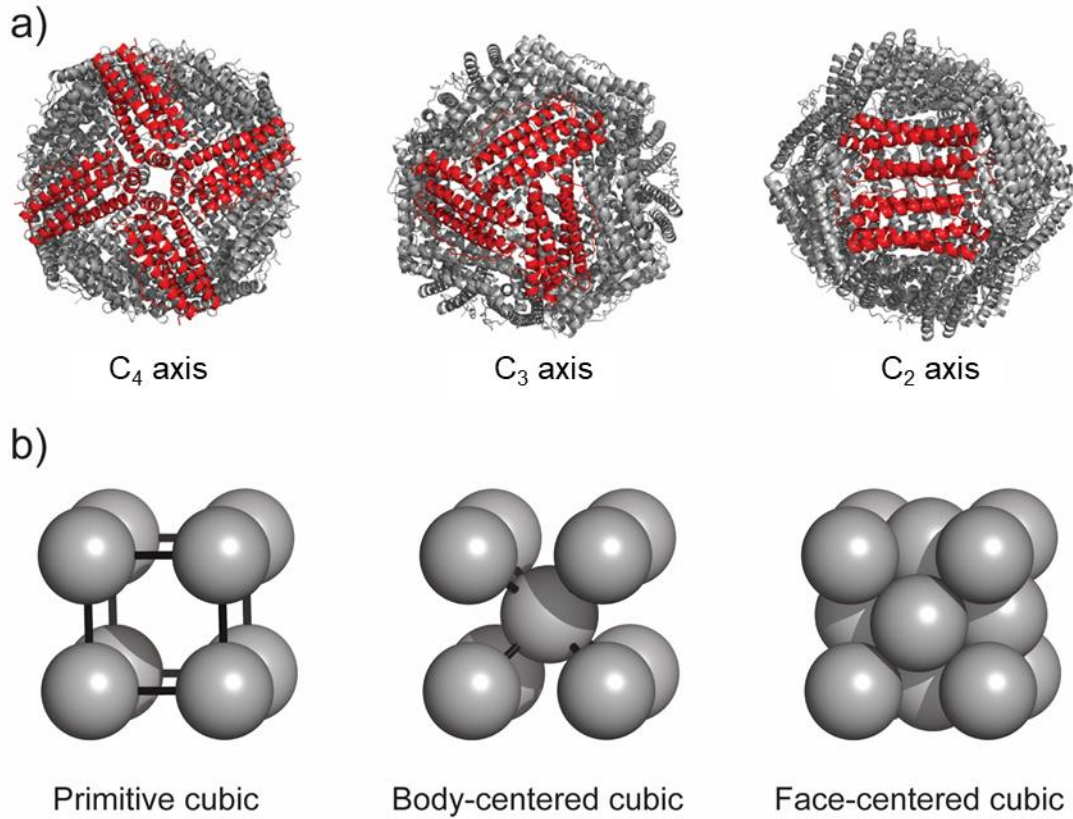


Figure 2.2. Symmetry axes and expected crystal packing of ferritin. a) The C_4 , C_3 , and C_2 symmetry axes of ferritin are highlighted in red. b) Expected lattice symmetry obtained by packing through the C_4 , C_3 , and C_2 axes, respectively.

2.2.2. Design and characterization of a metal-binding site at the C_4 axes of ferritin

Unlike the surface exposed C_2 axes, it is possible to engineer a more recessed metal-anchoring site due to the presence of the small 5th helix, which extends toward the interior of the cage (Figure 2.3 a). Upon further examination of the crystal structure of native ferritin, ^{173}His forms a highly-ordered tetra-His metal-binding site near the C-terminal end of the 5th helix (Figure 2.3 b-c). The metal is positioned to have an outward-facing coordination site that could potentially be accessed by a synthetic ligand. The ^{173}His site has been crystallographically observed to form complexes with late first-row transition metals (Fe^{2+} , Co^{2+} , Ni^{2+} , Cu^{2+} , and Zn^{2+}). However, ^{173}His is significantly buried in the C_4 axis (ca. 16 Å) and would be nearly impossible to access with a bridging ligand. Fortunately, an almost identical binding site could be installed closer to the top of the 5th helix through the single L165H mutation. It is also possible to form a metal anchoring site with the L169H mutation, but this site would be ca. 4 Å deeper in the 4-fold pore. As ^{165}H is the best candidate for the installation of a metal-binding site at the C_4 axis, site-directed mutagenesis was used to install the L165H mutation. To probe the formation of the metal-protein node with the desired coordination motifs, ^{165}H -ferritin (12.5 μM) was crystallized by vapor diffusion through the previously installed 2-fold crystal contact in the presence of excess CuCl_2 (5 mM) at pH 8.5. These conditions yielded crystals with a characteristic octahedron shaped morphology and allowed the determination of the Cu - ^{165}H -ferritin structure at 1.93 Å resolution ($F432$ space group, $a = 180.21$ Å). Despite being nearly identical to the ^{173}His metal-binding site, ^{165}H did not bind Cu^{2+} . Replacing Cu^{2+} with Ni^{2+} during crystallization also resulted in a fully occupied ^{173}His metal-binding site and a non-binding ^{165}H site. The failure of ^{165}H to anchor a metal ion, coupled with the predicted 8 Å distance to the surface, led us to abandon the C_4 axis for the formation of protein-MOFs.

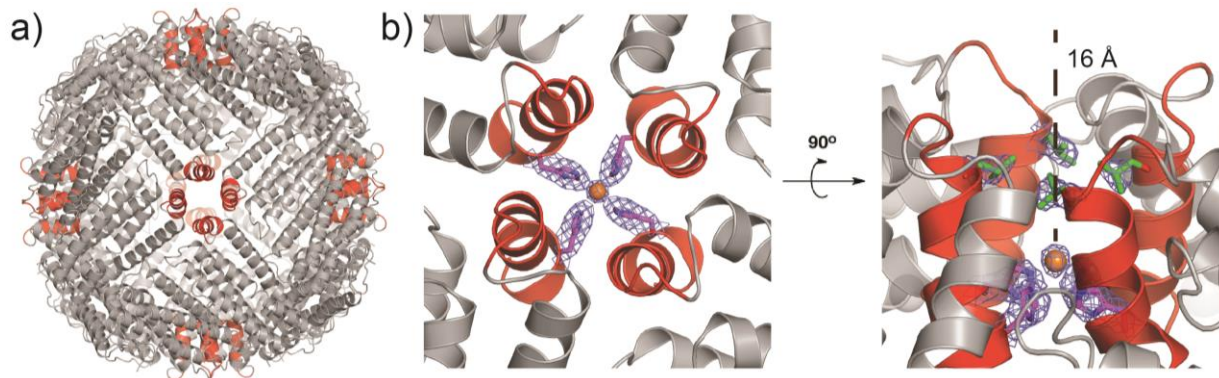


Figure 2.3. Analysis of the C_4 axis of ferritin. a) A single ferritin cage is shown with the 5th helices at the C_4 axis highlighted in red. b) ^{173}His (magenta) forms a tetra-His binding site and coordinates a single Cu atom. The $2F_o - F_c$ map is contoured at 1σ . c) The ^{173}His metal-binding site is buried ca. 16 Å from the protein surface. ^{165}Leu is shown in green near the top of the 5th helix.

2.2.3. Design and characterization of a metal-binding site at the C_3 axes of ferritin

Unlike the C_2 and C_4 axes, the C_3 symmetric pores of ferritin are ideal for the installation of a metal anchoring site. Residue 122 (originally Thr), which lines the C_3 symmetric pores near the exterior surface, was determined to be well-suited for constructing a tripodal metal-binding coordination motif. The side chains at ^{122}Thr positions are oriented toward the center of the pores with $\alpha\text{-C}$ distances of 10.2 Å, which provides sufficient clearance for substitution with histidine residues ($^{122}\text{Hferritin}$) to enable metal coordination. This tripodal motif is reminiscent of the Zn^{2+} active site found in carbonic anhydrase (Figure 2.4).³² This metal-anchoring site is slightly recessed from the outer surface of the protein cage. Residues ^{123}Asp and ^{124}Lys protrude from the 3-fold axis, making metal-mediated dimerization at the tris-His site unlikely. To probe the formation of the metal-protein node with the desired coordination motifs, $^{122}\text{Hferritin}$ (12.5 μM) was crystallized by vapor diffusion in the presence of excess ZnCl_2 (5 mM) at pH 8.5. These conditions yielded crystals with a characteristic octahedron shape (Figure 2.5 a) and allowed the determination of the $\text{Zn-}^{122}\text{Hferritin}$ (referred to as Zn-ferritin) structure at 1.93 Å resolution ($F432$ space group, $a = 180.21$ Å; Table 2.4). This structure revealed that the engineered ^{122}His coordination sites in the C_3 pore coordinate Zn^{2+} ions (8 ions per 24-mer, Figure 2.5 b). Importantly, in the *fcc* arrangement, the $^{122}\text{His-Zn}$ sites are far removed from crystal packing contacts formed near the C_2 symmetry axes of ferritin (Figure 2.6). The engineered Zn coordination sites in the C_3 pores possess a near-ideal tetrahedral geometry created by the $\epsilon\text{-N}$'s of the three ^{122}His side chains and a single water molecule (Figure 2.5 c). The N(His)-Zn-N(His) and N(His)-Zn-OH_2 angles are 109.6° and 109.4° , with Zn-N(His) and Zn-OH_2 bond distances of 2.2 Å and 2.8 Å, respectively. The distances between the $\alpha\text{-C}$'s at position 122 have increased by 0.5 Å to 10.7 Å, indicating that the C_3 pore has slightly expanded to accommodate stable Zn^{2+} coordination. Importantly, the Zn-bound water molecule protrudes out of the exterior ferritin surface, suggesting that it is poised for substitution by the metal-binding moiety of an organic ligand.

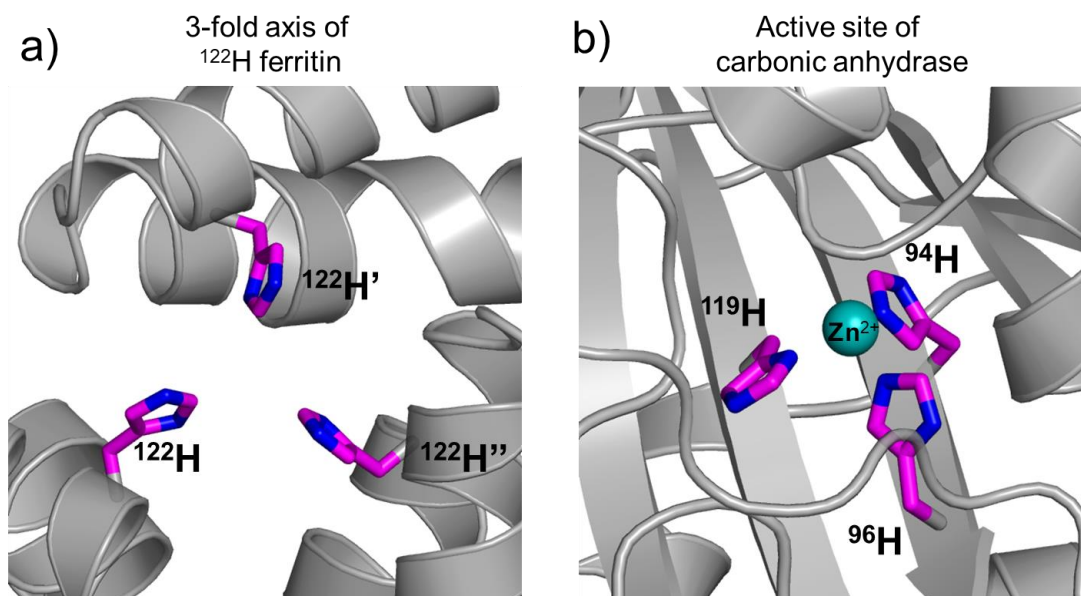


Figure 2.4. Comparison of the predicted C_3 axis of ferritin with the carbonic anhydrase active site. (a) A model of the designed 3-fold axis of ^{122}H ferritin. The three histidines form a tripodal coordination site that is poised for metal coordination. (b) Carbonic anhydrase has a tripodal tris-His active site bound to Zn^{2+} (PDB ID: 3KS3).

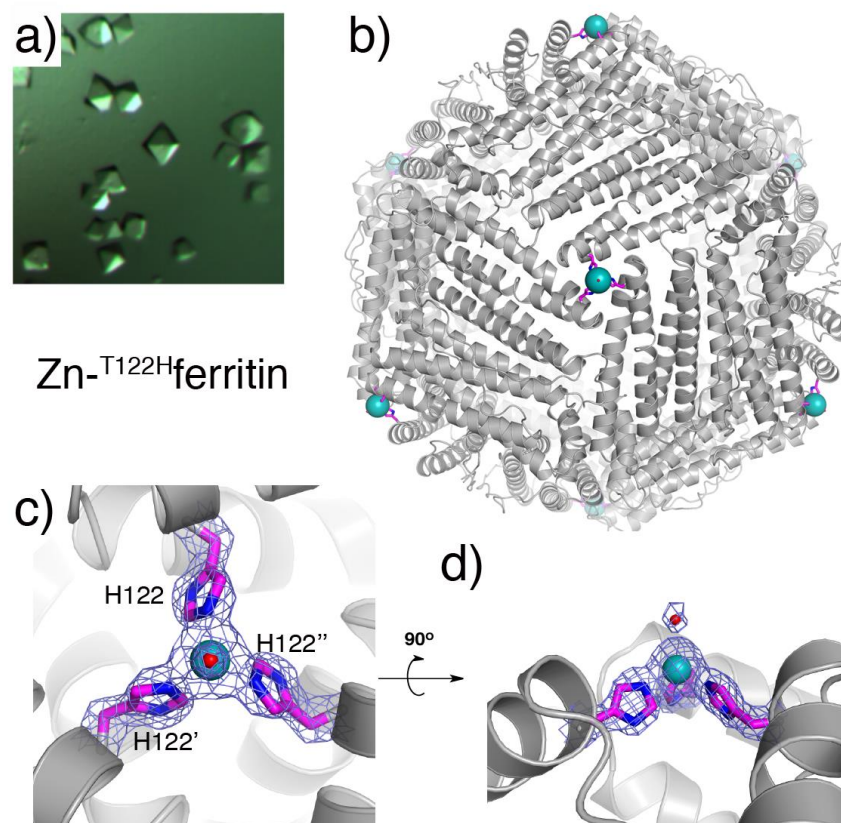


Figure 2.5. Crystal structure of Zn-ferritin. (a) Light micrograph of diamond-like *fcc* crystals of Zn-ferritin. (b) Crystal structure of Zn-ferritin, highlighting the locations of engineered Zn (teal) coordination sites (magenta) in the C_3 pores. (c), (d) Closeup views of the engineered Zn coordination sites (water: red sphere). The $2F_o - F_c$ map is contoured at 1σ .

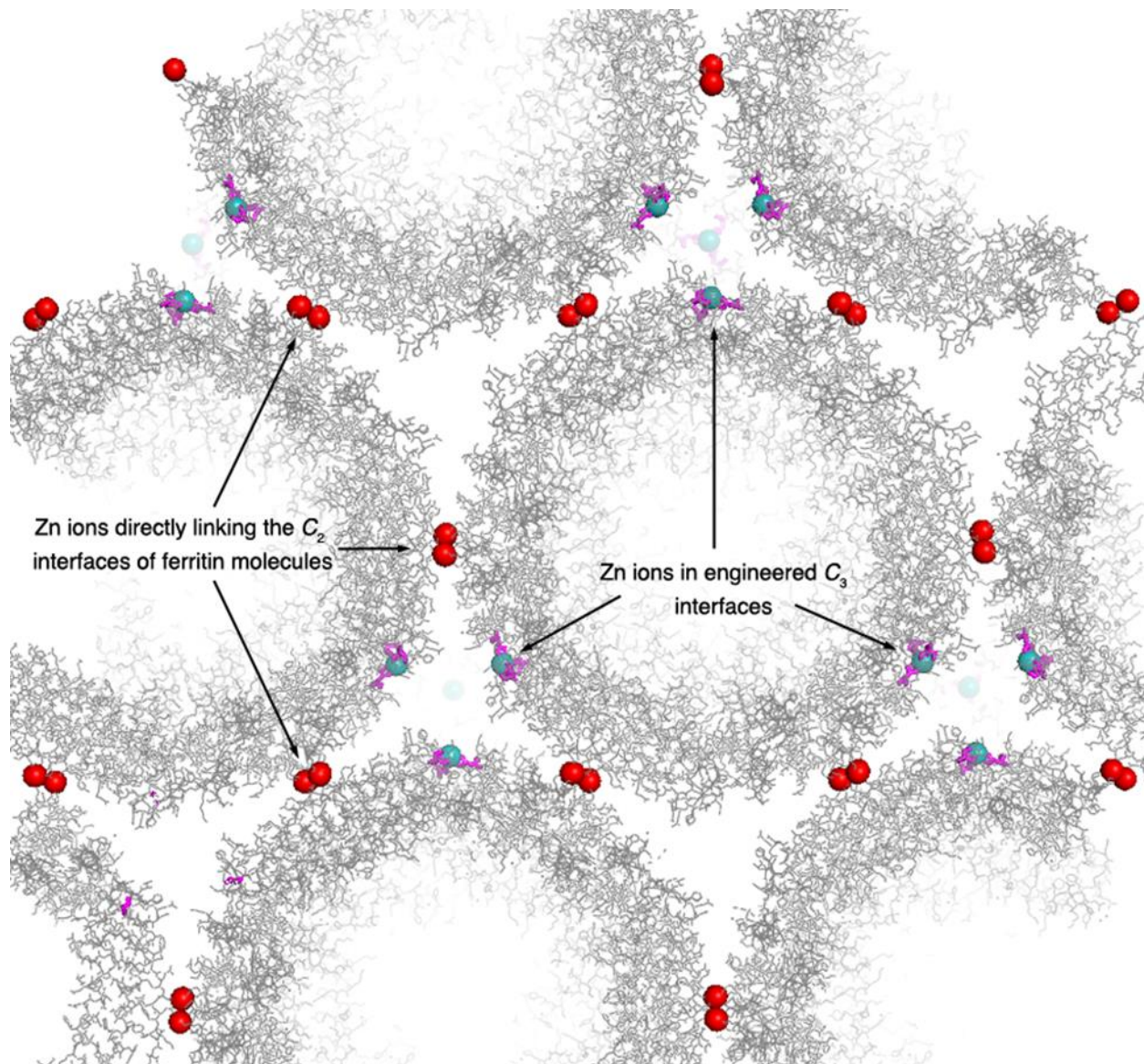


Figure 2.6. Crystal packing contacts of Zn-ferritin. The ferritin molecules are in an *fcc* (F432) arrangement with Zn^{2+} ions (red spheres) at the C_2 interface directly mediating contacts. These Zn^{2+} ions are coordinated by ^{84}Glu and ^{86}Gln sidechains. In this structure, ^{122}His -Zn engineered sites (magenta residues/blue spheres) at the C_3 axis are distant from the crystal packing contacts.

2.2.4. Selection of a suitable bridging ligand for protein-MOF construction

Given the large size of Zn-ferritin nodes (12 nm in diameter) relative to organic linkers used in MOFs, their only feasible linker-mediated supramolecular arrangement would be one directed by a ditopic bridging ligand to yield a body-centered cubic (*bcc*) lattice. We first examined the linker-mediated self-assembly of Zn-ferritin in an aqueous solution using dynamic light scattering (DLS). DLS measurements with ^{122}Hf -ferritin (33 μM monomer, 1.4 μM 24-mer) showed a monodisperse distribution of hydrodynamic diameters (D_h) centered at 13 nm, which is consistent with the expected diameter of the ferritin cage (Figure 2.7 a). A nearly identical distribution was observed when ^{122}Hf -ferritin was combined with an excess of Zn^{2+} (75 μM), indicating that Zn^{2+} alone does not mediate protein self-assembly (Figure 2.7 b). The oft-used *para*-benzene-dicarboxylic acid (*p*- H_2bdc) was chosen as the initial candidate for a linker since it is rigid, linear, and water-soluble.^{28-29, 33} We observed that the addition of excess *p*- H_2bdc (up to 1 mM) with Zn^{2+} (75 μM), also did not result in significant changes in particle size distribution (Figure 2.7 c). This observation is not surprising, as the carboxylate functionalities of *p*- H_2bdc likely do not provide a sufficiently high affinity for Zn^{2+} to hold together Zn-ferritin nodes in an aqueous medium. We surmised that the substitution of carboxylates with bidentate hydroxamate moieties could provide the necessary Zn^{2+} binding affinity while preserving the advantageous properties of *p*- H_2bdc . Hydroxamate functional groups are well known for their ability to form strong bonds with Fe^{3+} as well as a variety of divalent transition metal ions and are found in iron-scavenging siderophores and Zn metalloenzyme inhibitors.³⁴⁻³⁹ The dihydroxamic analog of *p*- H_2bdc , *para*-benzene-dihydroxamic acid (*p*- H_2bdh), was synthesized through a base-mediated amidation reaction from dimethyl terephthalate. Indeed, the addition of *p*- H_2bdh to a solution of ^{122}Hf -ferritin and Zn^{2+} led to the complete depletion of the 13-nm diameter species and the emergence of larger aggregates (>300 nm) (Figure 2.7 d). When the monotopic benzenehydroxamic acid (Hbmh), which cannot act as a linker, was used instead (Figure 2.7 e),

or when Zn^{2+} was omitted (Figure 2.7 f), no high-order oligomeric species were observed. These findings indicate that H_2bdh is capable of inducing the self-assembly of Zn - ^{122}Hf ferritin.

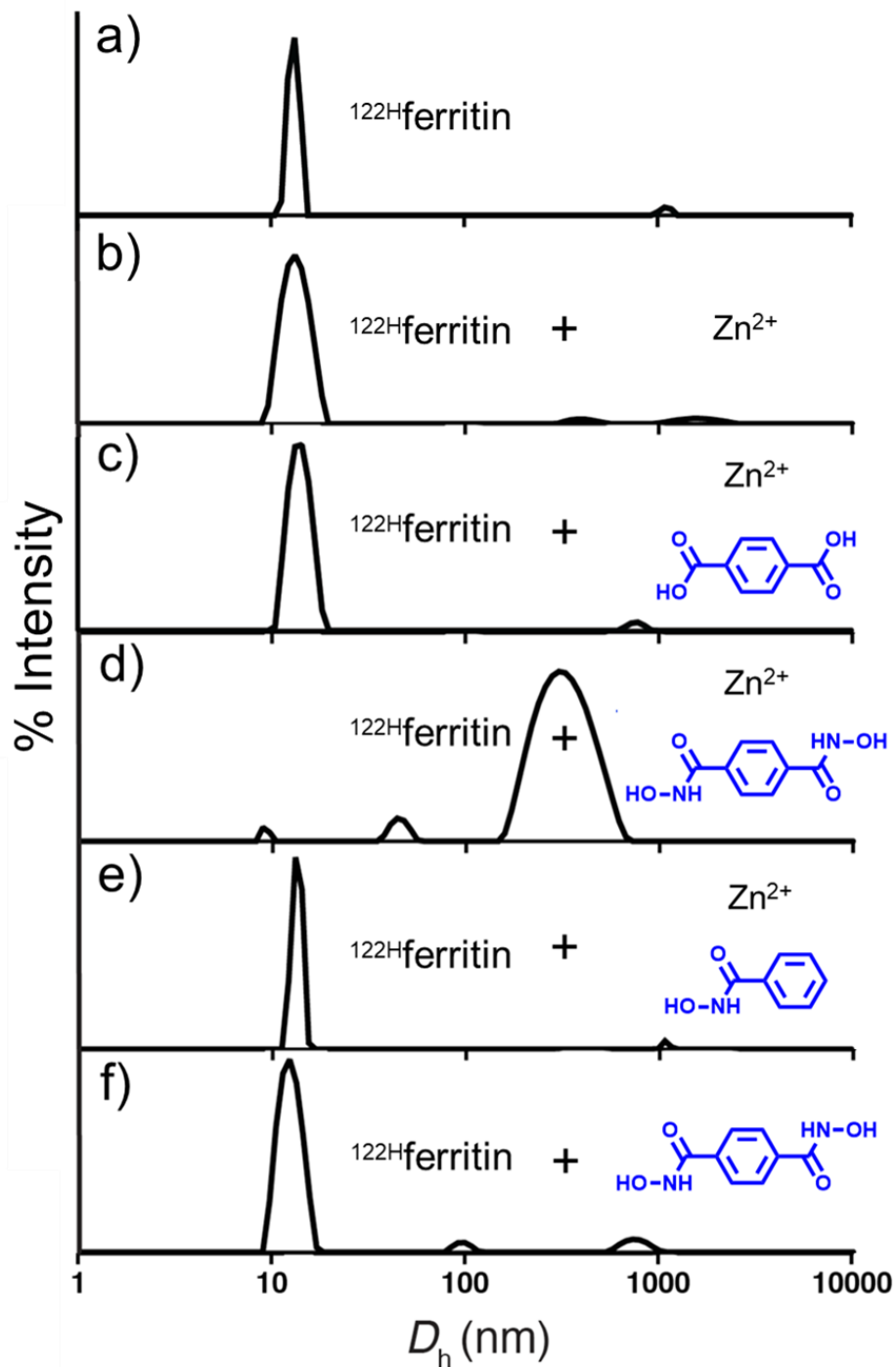


Figure 2.7. Dynamic light scattering profiles of ferritin-MOF assembly conditions. a) ^{122}Hf ferritin, b) ^{122}Hf ferritin + Zn^{2+} , c) ^{122}Hf ferritin + p -H₂mbh, d) ^{122}Hf ferritin + Zn^{2+} + p -H₂bdh, e) ^{122}Hf ferritin + Zn^{2+} + Hbmbh, f) ^{122}Hf ferritin + p -H₂bdh.

2.2.5. Synthesis and characterization of the first protein-MOF

To obtain 3D *p*-bdh-Zn-ferritin crystals, we screened a wide range of solution conditions (pH 5.5-9.0) at room temperature. We consistently observed the formation of distinct, cube-shaped crystals at pH 8.0-9.0 under the following component concentrations: 1-12.5 μM ^{122}Hf ferritin or $^{86}\text{A},^{122}\text{Hf}$ ferritin (where the Q86A mutation is intended to eliminate the possibility of *fcc* packing), 50-250 equivalents of Zn^{2+} (per 24-meric ferritin unit; to saturate all possible Zn sites), 0.5-2 mM *p*-H₂bdh. It was later determined that the Q86A mutation was not necessary and that *fcc* lattice formation rarely observed under conditions that yield ferritin-MOFs. Crystal formation from bulk solution occurred rapidly (≤ 24 h) to form crystals up to 0.5 mm in size (Figure 2.8 a). Consistent with DLS measurements, no crystals formed when *p*-H₂bdh was substituted with *p*-H₂bdc or Hbmh, or when any one of the three components (protein, metal, or ligand) was not included in the solution. When the missing component was introduced into the two-component solutions, which had been incubating overnight, crystals with a rhombic dodecahedral morphology rapidly emerged. Thus, the self-assembly of *p*-bdh-Zn-ferritin crystals is robust and dependent on all three components in solution.

We next turned to single-crystal X-ray diffraction (sc-XRD) to elucidate the structure of the rhombic dodecahedral crystals. Most of the *p*-bdh-Zn-ferritin crystals displayed X-ray diffraction >6 Å. However, after an exhaustive screen, we found a single crystal that provided sufficiently high resolution (3.79 Å) for structure determination. As designed, this crystal has a *bcc* lattice arrangement (*I*432 space group, $a = 155.30$ Å), shown in Figure 2.8 b. An inspection of the initial molecular replacement solution reveals a striking picture where ferritin molecules appear to be suspended in the lattice without any protein-protein contacts (Figure 2.9): the shortest atom-to-atom distance between the neighboring Zn-ferritin nodes near their C_3 symmetry axes is >6 Å (Figure 2.8 c). The engineered Zn^{2+} binding sites, which lie on the crystallographic C_3 axes, are at a separation of 11.2 Å across the interface with some electron density in the intervening region,

consistent with the presence of a single ligand. The rotational averaging of *p*-bdh about the crystallographic C_3 axis, coupled with the low resolution of the diffraction data, gives rise to weak electron density for the linker, particularly around hydroxylamine moieties that are tilted furthest away from the C_3 axis. Nevertheless, a plausible bidentate Zn-*p*-bdh coordination geometry was modeled. In this coordination geometry (shared by three crystallographically related rotamers, Figure 2.8 d), hydroxamate-Zn coordination yields planar five-membered rings with Zn-O(carbonyl) and Zn-O(hydroxylamine) distances of ca. 2.1 Å and 2.3 Å, respectively. The Zn-O(carbonyl) bond is directed nearly parallel to the crystallographic C_3 axis, whereas the Zn-O(hydroxylamine) bond points away from it (at a 60° tilt) and between the ^{122}His side chains. Although other supramolecular protein architectures have been assembled through the use of synthetic organic groups with metal coordinating functionalities⁴⁰⁻⁴⁵ or pre-existing MOFs have been postsynthetically modified with proteins,⁴⁶ the *p*-bdh-Zn- ^{122}Hf ferritin system represents the first example of a ternary protein-metal-organic crystalline framework where each component is an integral part of the construction. The *bcc* arrangement presents a relatively high packing density for a spherical building block (packing factor = 0.68), yet the ferritin hubs themselves are hollow, thus producing a highly porous framework with a solvent content of 67%. As a result, *p*-bdh-Zn-ferritin crystals are highly permeable to solutes, allowing the ferritin nodes to perform their native enzymatic activity within the lattice, which is the oxidation of soluble Fe^{2+} into crystalline Fe^{3+} -oxide mineral (Figure 2.10). The *in crystallo* enzymatic activity creates the future opportunity to template crystalline inorganic materials in 3D with nanoscopic precision. More generally, the *p*-bdh-Zn-ferritin crystals introduce a new class of hybrid materials that combine the chemical and structural versatility of metal ions, organic linkers, and protein building blocks in a modular fashion.

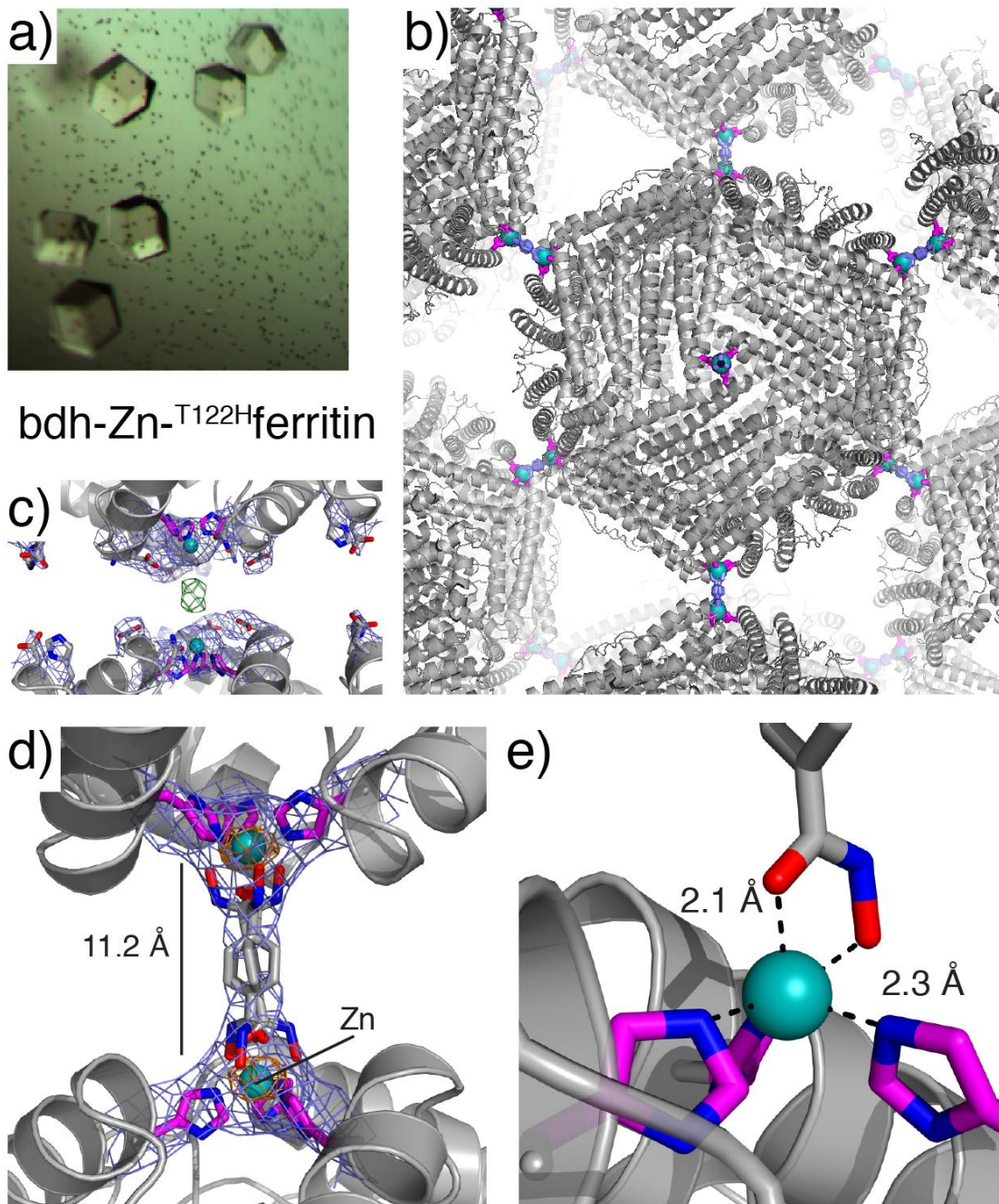


Figure 2.8. Crystal structure of *p*-bdh-Zn-^{86A}ferritin. (a) Light micrograph of cube-shaped *bcc* crystals of bdh-Zn-ferritin. (b) *bcc* packing of the *p*-bdh-Zn-ferritin lattice, mediated by *p*-bdh bridges across the C_3 symmetric ferritin interfaces. (c) View of C_3 symmetric interfaces between neighboring ferritin molecules, showing the lack of direct protein-protein contacts and the presence of electron density between engineered Zn^{2+} ions ($2F_o - F_c$ map: blue- 1σ ; $F_o - F_c$ map: green- 3σ). (d) Closeup view of the three crystallographically related *p*-bdh rotamers that bridge engineered Zn^{2+} ions on neighboring ferritin cages; anomalous difference map (7σ) calculated using Zn K-edge diffraction data is shown in orange. (e) Closeup view of modeled *p*-bdh-Zn coordination.

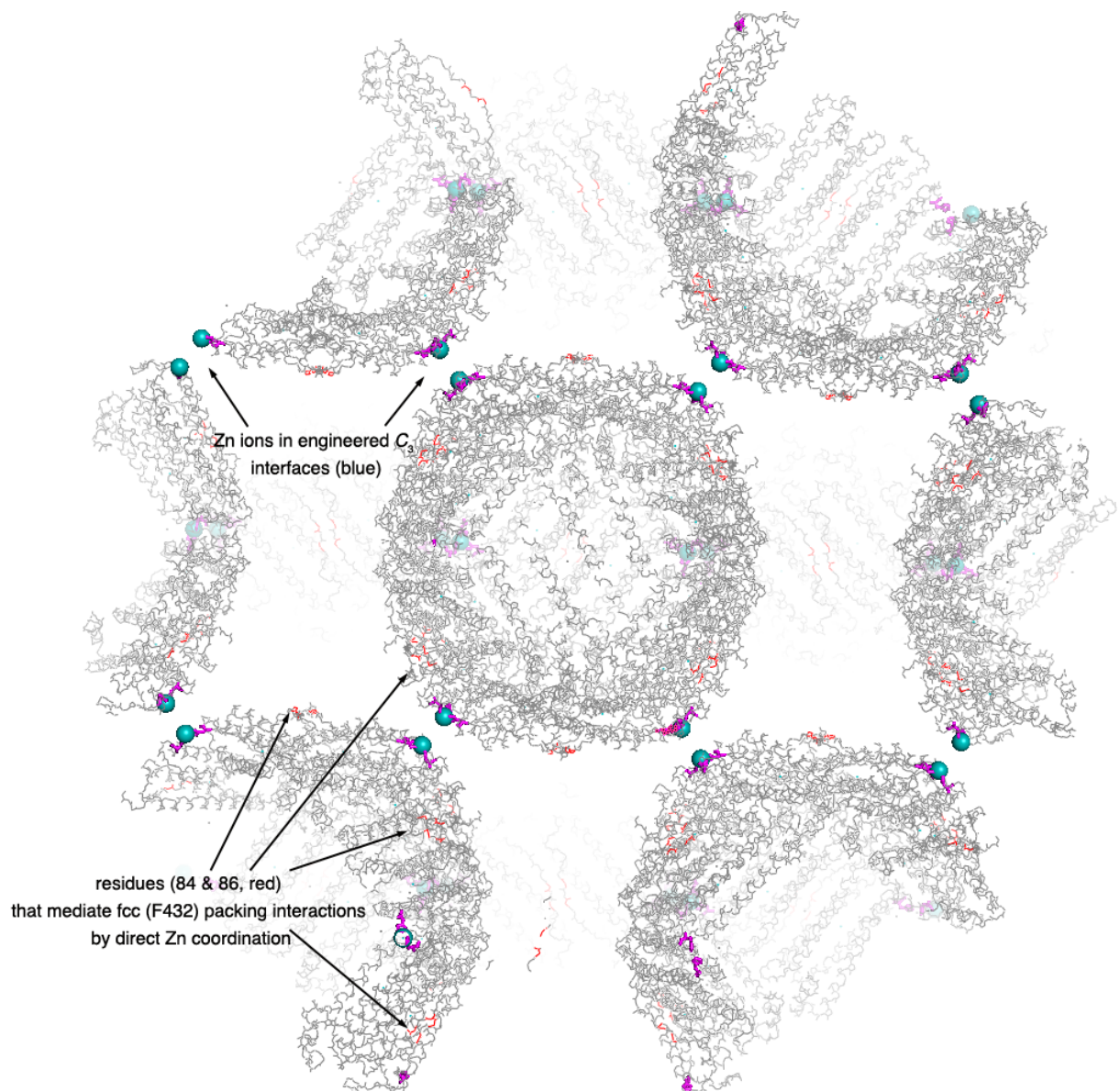


Figure 2.9. Crystal packing contacts of *p*-bdh-Zn^{86A}ferritin. Ferritin molecules are in a *bcc* (*I432*) arrangement and appear suspended in space with no direct protein-protein contacts. Zn²⁺ ions at the engineered sites (blue spheres) are separated by a distance of >11 Å across the C_3 interface. The residues (84 and 86, shown in red) that mediate contacts in the absence of linker are far removed from the contact sites.

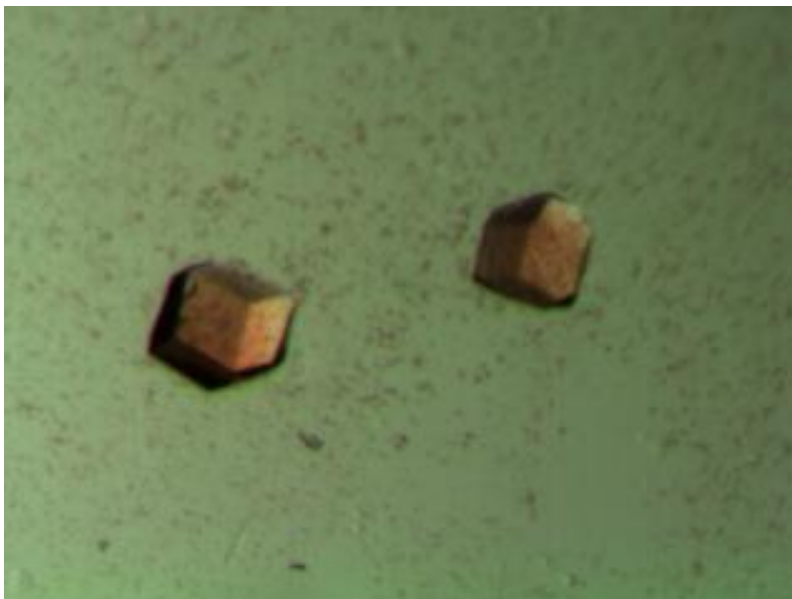


Figure 2.10. Light micrograph of *p*-bdh-Zn-ferritin crystals upon iron loading/mineralization. *p*-bdh-Zn-ferritin crystals become rust-colored after being soaked for 24 hours in a 10 mM (NH₄)₂Fe(SO₄)₂ · 6H₂O solution, indicative of Fe²⁺→Fe³⁺ oxidation.

2.3. Methods

2.3.1. General considerations

All chemical reagents were purchased from Sigma-Aldrich, Fisher Scientific, TCI, or an equivalent supplier. Mass spectrometry (MS) of proteins and small molecules was carried out at the Molecular Mass Spectrometry Facility at UCSD using electrospray ionization (ESI) on a Micromass Quattro Ultima Triple Quadrupole MS. NMR spectra (¹H) were recorded on either a Mercury spectrometer (400 MHz) or a Varian Mercury spectrometer (500 MHz). NMR data are reported as follows: chemical shift, multiplicity (s = singlet, br = broad), and integration. The spectra were internally referenced to the residual solvent signal (DMSO: δ 2.50 ppm).

2.3.2. Site-directed mutagenesis.

Mutations were encoded on the wild-type human H-ferritin pJexpress bacterial vector obtained from DNA 2.0 (Menlo Park, CA) that was optimized for expression in *E. coli*. Variants were prepared using QuikChange mutagenesis (Stratagene) with primers obtained from Integrated DNA Technologies (Table 2.1). Mutant plasmids were transformed into XL-1 blue *E. coli* cells and purified with a QIAprep Spin Miniprep kit (Qiagen). All variants were sequenced (Retrogen) to verify mutagenesis.

Table 2.1. Primers for site-directed mutagenesis of human heavy-chain ferritin

Variant	Mutation (order of addition)	Primer sequence (5'-3': mutation site in red)
ΔC* ferritin ²⁷	K86Q + C90E	GCAGGACATT CAG AAGCCGGAT GAG GACGATTGGG
	C102A	GGCCTGAATGCGATGGAG GCG GCGCTGCATCTGG
	C130A	GAATGATCCGCACCTG GCG GATTTTCATCGAAACGC
L165H ΔC*	L165H	GCACCGGAATCCGGT CAT GCCGAGTACCTGTTT
T122H ΔC*	T122H	GGAGCTGCACAAGCTGGCA CAT GATAAGAATGATCCGCACC
T122H/Q86A ΔC*	Q86A	GCCGCATCTTTCTGCAGGACATT GCG AAGCCGGATGAG

2.3.3. Expression, purification, and characterization of ferritin variants

For ferritin expression, isolation, and purification, we used previously established procedures with minor modifications.²⁷ Mutant plasmids isolated from XL-1 blue cells were transformed into BL21 (DE3) *E. coli* cells and plated on LB agar containing 100 µg/mL ampicillin. A single colony or freezer stock of BL21 containing ferritin variant pJexpress vectors (DNA 2.0) was used to inoculate a starter cell culture (200 mL LB medium, 100 µg/mL ampicillin). The starter culture was incubated for 16 hours at 37 °C while shaking at 250 rpm prior to inoculating 16 x 1 L LB cultures (10 mL per flask) supplemented with 100 mg/L ampicillin. The *E. coli* cells were grown at 37 °C until the OD₆₀₀ was between 0.6 and 0.8. Protein expression was induced by the addition of isopropylβ-D-1-thiogalactopyranoside (IPTG) to a total concentration of 1 mM. Cells were incubated at 37 °C for 3.5 h and harvested by centrifugation (4000 rpm, 10 min). Cell pellets were combined and stored at -80 °C. Frozen cell pellets were thawed and suspended in a buffered solution containing 15 mM Tris (pH 7.4), 150 mM NaCl with 1 mM phenylmethylsulfonyl fluoride (PMSF), 5 mM DTT, and ca. 25 µM lysozyme. Cells were sonicated for 12 min on ice, and the cell lysate was centrifuged at 4 °C (10,000 rpm, 10 min). The supernatant was collected, and the remaining cell pellet was resuspended in a buffered solution containing 15 mM Tris (pH 7.4), 150 mM NaCl with Triton X-100 (1% v/v) and sonicated for 12 min on ice. This lysate was centrifuged at 4 °C (10,000 rpm, 10 min), and the second lysate supernatant was combined with the first. Ferritin was isolated in the intact, 24-meric state from this combined lysate. First, ferritin was enriched to >80% purity by a heat treatment at 65 °C for 10 min and centrifugation (10000 rpm, 10 min). The protein, found in the soluble fraction, was then exchanged into a buffered solution (15 mM Tris, pH 8.0), filtered, and purified using multiple Uno-Q anion exchange column (BioRad) on a DuoFlow chromatography workstation (BioRad) using a linear 0-1 M NaCl gradient. Ferritin elutes at ca. 400 mM NaCl. Fractions containing ferritin were combined and dialyzed into a buffered solution containing 15 mM Tris (pH 7.4), 150 mM NaCl. This protein was then concentrated and loaded onto a Sephacryl S-300 resin size-exclusion column (GE Healthcare)

that was equilibrated with a buffered solution containing 15 mM Tris (pH 7.4), 150 mM NaCl. Protein purity was assessed via SDS-PAGE (Figure 2.11). Electrospray ionization (ESI) mass spectrometry was used to confirm the identity of the isolated protein (Table 2.2). After purification, ferritin was dialyzed into a buffered solution containing 50 mM CHES (pH 8.5) with 150 mM NaCl. 122H ferritin was concentrated to 25 μ M (by 24-meric ferritin cage) and aliquoted. Protein aliquots were stored at 4 °C and warmed to room temperature before use.

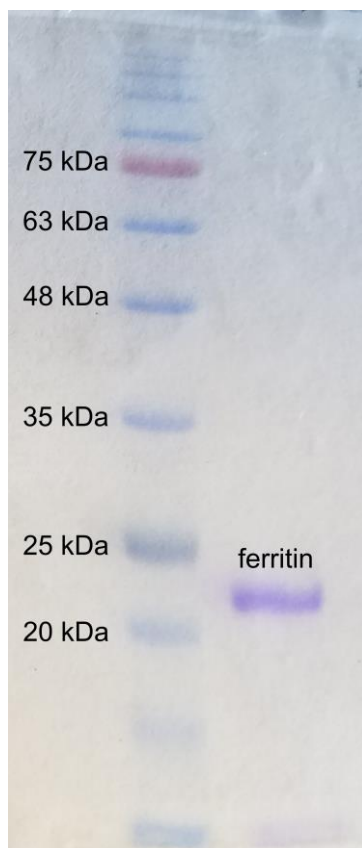
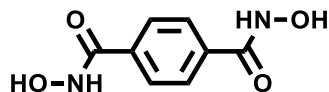


Figure 2.11. Image of purified ferritin confirmed by SDS-PAGE.

Table 2.2. Mass spectrometry measurements for all ferritin variants.

Variant	Expected Mass (Da)	Observed Mass (Da)
ΔC^*	21,056	21,058
T122H ΔC^*	21,093	21,096
T122H,Q86A ΔC^*	21,035	21,036

2.3.4. Synthesis of N^1,N^4 -dihydroxyterephthalamide (p -H₂bdh)



The synthesis for p -H₂bdh was carried out, as previously reported with minor modifications.⁴⁷ Hydroxylamine hydrochloride, 1.06 g (15.45 mmol), and sodium hydroxide, 1.24 g (30.9 mmol) were combined in 20 mL methanol and shaken vigorously. The solution was cooled on ice for 10 min, and NaCl was removed via filtration. The filtrate was combined with 1 g (5.15 mmol) dimethyl terephthalate in 30 mL methanol. The solution was stirred overnight at room temperature. The solvent was removed in vacuo, and the solid material was dissolved in 20 mL H₂O. The solution was acidified to a pH of 5.5 with 5% HCl. The white precipitate was collected by filtration and dried in vacuo. Yield = 65%. ¹H NMR is consistent with previously published values: (400 MHz, DMSO-d₆) δ 11.35 (br s, 2H), δ 9.17 (br s, 2H), δ 7.80 (s, 4H). Measured molecular weight (m/z) = 196.97 (Calc.: 197.05) (M + H⁺). ¹³C NMR (500 MHz, DMSO-d₆): δ 163.42, δ 135.04, δ 126.92.

2.3.5. DLS measurements of ferritin assembly conditions

Dynamic light scattering (DLS) experiments were performed using a Wyatt DynaPro NanoStar instrument. Before the DLS experiments, the protein solution was filtered using a syringe filter (1.0 μm). The temperature was held at 25 °C throughout the experiment. The samples contained 33 μM protein in a solution of 50 mM CHES (pH 8.5), 150 mM NaCl, and 75 μM ZnCl₂ (unless otherwise noted). The stock solutions for the organic linkers contained 5 mM p -H₂bdc, p -H₂bdh, or Hbmbh, dissolved in 50 mM CHES, 150 mM NaCl, pH 9.5. Aliquots from these stock solutions were added to the protein samples to give a final linker concentration of 1 mM. Diameter (D_h) measurements were taken after 1 hour. The experiment runs were performed to collect 10 acquisitions over 10 second time intervals with 657 nm excitation at a power setting of 100%. Peak radius cutoffs were fixed according to default settings (0.5-10000 nm).

Measurements were plotted using a Rayleigh spheres model. Acquisitions from the samples containing protein only or protein with Zn^{2+} exhibiting baselines ≤ 1.026 were used in the average distribution plot. For samples containing organic linkers, all ten acquisitions were included.

2.3.6. Crystallization of the Zn-ferritin node

Crystals of Zn-ferritin were obtained in the *F432* space group through sitting-drop vapor diffusion using solution conditions listed in the table below (Table 2.3). Each sitting drop consisted of 2 μL of protein solution in standard buffer (12.5 μM) with 2 μL of precipitant solution at 25 $^{\circ}\text{C}$. The wells each contained 500 μL of precipitant solution. Crystals appeared after overnight incubation and matured over one week. Before screening, crystals were cryoprotected in perfluoropolyether cryo oil (Hampton) and frozen in liquid nitrogen.

2.3.7. Crystallization of $p\text{-bdh-Zn-}^{86}\text{A}_f$ ferritin and $p\text{-bdh-Zn-ferritin}$

Crystals of $p\text{-bdh-Zn-}^{86}\text{A}_f$ ferritin and $p\text{-bdh-Zn-ferritin}$ were reproducibly self-assembled as described below (Table 2.3). Protein stock solutions (^{122}Hf ferritin and $^{122}\text{H,}^{86}\text{A}_f$ ferritin) were stored in a buffered solution (50 mM CHES, 150 mM NaCl, pH 8.5) before use in all crystallization experiments. ^{122}Hf ferritin or $^{122}\text{H,}^{86}\text{A}_f$ ferritin (1-12.5 μM), 50-250 equivalents of ZnCl_2 (per 24-mer cage), and 0.5-2 mM H_2bdh at 25 $^{\circ}\text{C}$ were mixed in a single drop. Crystal growth was observed over a range of pH values (8.0-9.0). At Zn equivalents greater than 250 per cage or when one or more of the three components ($p\text{-H}_2\text{bdh}$, Zn, or ^{122}Hf ferritin) was omitted, no crystals were observed. In addition, when Hbmh was used in place of H_2bdh , or when ΔC^* ferritin is used in place of ^{122}Hf ferritin, no crystal formation occurs. The $\text{bdh-Zn-}^{86}\text{A}_f$ ferritin crystals selected for screening were soaked in a solution containing ~5 mM H_2bdh (not completely soluble) in 50% xylitol (w/v filtered) for 1 minute and then flash-frozen in liquid nitrogen.⁴⁸

Table 2.3. Crystallization conditions for ferritin.

Protein	Space group	Crystallization/self-assembly conditions
^{122}H ferritin	<i>F</i> 432	25 mM Tris (pH 8), 10 mM CaCl_2 , 10 mM ZnCl_2 , 2% PEG 3350
^{122}H ferritin	<i>I</i> 432	50 μM $^{122}\text{H}\Delta\text{C}^*$, 50 mM CHES (pH 8.5), 150 mM NaCl, 0.1 mM ZnCl_2 , 1 mM H_2bdh
$^{122}\text{H},^{86}\text{A}$ ferritin	<i>I</i> 432	50 μM $^{122}\text{H},^{86}\text{A}\Delta\text{C}^*$, 50 mM CHES (pH 8.5), 150 mM NaCl, 0.3 mM ZnCl_2 , 1 mM H_2bdh

2.3.8. sc-XRD data collection and structure determination

X-ray diffraction data for *p*-bdh- $\text{Zn}^{86\text{A}}$ ferritin and $\text{Zn}^{122\text{H}}$ ferritin were collected at 100 K at Beamline 12-2 of the Stanford Synchrotron Radiation Laboratory (SSRL) using 0.98 Å radiation. Additional data for *p*-bdh- $\text{Zn}^{86\text{A}}$ ferritin were collected at the zinc K-edge (1.27 Å) to identify zinc ions through anomalous scattering. The data were processed using MOSFLM and SCALA.⁴⁹⁻⁵⁰ The structures of $\text{Zn}^{122\text{H}}$ ferritin and *p*-bdh- $\text{Zn}^{86\text{A}}$ ferritin were solved to a resolution of 1.93 Å and 3.79 Å resolution, respectively. Molecular replacement was performed with MOLREP using the ferritin structure (PDB ID: 2CEI) as the search model.⁵¹ For the Zn-ferritin structure, rigid-body, positional, and thermal refinement were carried out using REFMAC.⁶ This refined Zn-ferritin structure was used as a reference model during *p*-bdh- $\text{Zn}^{86\text{A}}$ ferritin refinement using PHENIX.⁵²⁻⁵³ PRODRG was used to generate restraints for the H_2bdh ligand.⁵⁴ COOT was used for iterative manual model building.⁵⁵ Before the last refinement, *p*-bdh was centered across the protein-protein interface, and the hydroxamate functional groups were rotated such that the O(hydroxylamine) was located between the ^{122}His sidechains. All figures were produced with PYMOL.⁵⁶ The refinement statistics are shown in Table 2.4.

Table 2.4. X-ray data processing and refinement statistics.

	Zn- ¹²² Hferritin	bdh-Zn- ¹²² HQ86A ferritin
Data collection		
Space group	<i>F</i> 432	<i>I</i> 432
Cell dimensions		
<i>a, b, c</i> (Å)	180.21	155.32
α, β, γ (°)	90	90
Resolution (Å)	31.86-1.93	54.91-3.79
No. unique reflections	17713	3401
Multiplicity*	13.7(14.2)	11.2 (12.1)
<i>CC</i> 1/2*	0.998 (0.969)	0.997 (0.918)
$\langle I / \sigma I \rangle$ *	19.2 (7.0)	12 (5.0)
Completeness (%)*	91.7 (84.0)	99.8 (100)
<i>R</i> _{sym} *	0.097 (0.328)	0.185 (0.555)
Refinement		
<i>R</i> _{work} / <i>R</i> _{free}	0.14/0.18	0.21/0.26
No. atoms		
Protein	1428	1421
Ligand/ion	9	27
Water	239	0
<i>B</i> -factors (Å ²)		
Protein	15.9	82.0
Ligand/ion	34.6	80.9
Water	28.1	N/A
R.m.s. deviations		
Bond lengths (Å)	0.021	0.005
Bond angles (°)	1.867	1.133
Ramachandran plot (%)		
Most favored	99.0	99.0
Allowed	1.0	1.0

2.3.9. Iron loading/mineralization in *p*-bdh-Zn-ferritin crystals

For iron loading experiments, *p*-bdh-Zn-¹²²Hferritin crystals were transferred to a solution of 25% glutaraldehyde (v/v) (for stabilizing crystals) and incubated overnight. Iron loading was performed by transferring the crystals to 10 μ L of a freshly prepared 10 mM (NH₄)₂Fe(SO₄)₂ · 6H₂O solution.⁵⁷ After soaking for 24 hours under aerobic conditions, the crystals became rust-colored, which is indicative of Fe²⁺ oxidation and mineralization.

2.4. Acknowledgments

Chapter 2, in part, is a reprint of the material “A metal-organic framework with spherical protein nodes: rational chemical design of 3D protein crystals.” as it appears in the Journal of the American Chemical Society (Sontz, P. A.*; Bailey*, J. B.; Ahn, S.; Tezcan, F. A. *J. Am. Chem. Soc.* **2015**, *137* (36), 11598-11601). The dissertation author is a co-primary author of this publication. He would like to thank US Department of Energy (DOE) (Division of Materials Sciences, Office of Basic Energy Sciences, (DE-FG02-10ER46677) for funding this work.

2.5. References

1. Moulton, B.; Zaworotko, M. J., From molecules to crystal engineering: supramolecular isomerism and polymorphism in network solids. *Chemical Reviews* **2001**, *101* (6), 1629-1658.
2. Desiraju, G. R., Crystal engineering: a holistic view. *Angew. Chem. Int. Ed. Engl.* **2007**, *46* (44), 8342-8356.
3. Hollingsworth, M. D., Crystal Engineering: from Structure to Function. *Science* **2002**, *295* (5564), 2410-2413.
4. Dandekar, P.; Kuvadia, Z. B.; Doherty, M. F., Engineering Crystal Morphology. *Annu. Rev. Mater. Res.* **2013**, *43*, 359-386.
5. Margolin, A. L.; Navia, M. A., Protein crystals as novel catalytic materials. *Angewandte Chemie-International Edition* **2001**, *40* (12), 2204-2222.
6. Ueno, T., Porous Protein Crystals as Reaction Vessels. *Chemistry-a European Journal* **2013**, *19* (28), 9096-9102.
7. Cooper, D. R.; Boczek, T.; Grelewska, K.; Pinkowska, M.; Sikorska, M.; Zawadzki, M.; Derewenda, Z., Protein crystallization by surface entropy reduction: optimization of the SER strategy. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2007**, *63* (5), 636-645.
8. Derewenda, Z. S., Application of protein engineering to enhance crystallizability and improve crystal properties. *Acta Crystallogr. Sect. D-Struct. Biol.* **2010**, *66*, 604-615.
9. Lanci, C. J.; MacDermaid, C. M.; Kang, S. G.; Acharya, R.; North, B.; Yang, X.; Qiu, X. J.; DeGrado, W. F.; Saven, J. G., Computational design of a protein crystal. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (19), 7304-9.
10. Sakai, F.; Yang, G.; Weiss, M. S.; Liu, Y.; Chen, G.; Jiang, M., Protein crystalline frameworks with controllable interpenetration directed by dual supramolecular interactions. *Nature Communications* **2014**, *5*.

11. Brodin, J. D.; Ambroggio, X. I.; Tang, C.; Parent, K. N.; Baker, T. S.; Tezcan, F. A., Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. *Nature Chemistry* **2012**, *4* (5), 375-382.
12. Furukawa, H.; Cordova, K. E.; O'Keeffe, M.; Yaghi, O. M., The chemistry and applications of metal-organic frameworks. *Science* **2013**, *341* (6149), 1230444.
13. Jones, J. T. A.; Hasell, T.; Wu, X.; Bacsá, J.; Jelfs, K. E.; Schmidtman, M.; Chong, S. Y.; Adams, D. J.; Trewin, A.; Schiffman, F.; Cora, F.; Slater, B.; Steiner, A.; Day, G. M.; Cooper, A. I., Modular and predictable assembly of porous organic molecular crystals. *Nature* **2011**, *474* (7351), 367-371.
14. Farha, O. K.; Özgür Yazaydın, A.; Eryazici, I.; Malliakas, C. D.; Hauser, B. G.; Kanatzidis, M. G.; Nguyen, S. T.; Snurr, R. Q.; Hupp, J. T., De novo synthesis of a metal-organic framework material featuring ultrahigh surface area and gas storage capacities. *Nat. Chem.* **2010**, *2* (11), 944-948.
15. Peplow, M., Materials science: The hole story. *Nature* **2015**, *520* (7546), 148-50.
16. Chui, S. S.-Y.; Lo, S. M.-F.; Charmant, J. P. H.; Orpen, A. G.; Williams, I. D., A Chemically Functionalizable Nanoporous Material [Cu₃(TMA)₂(H₂O)₃]_n. *Science* **1999**, *283* (5405), 1148-1150.
17. Li, J. R.; Sculley, J.; Zhou, H. C., Metal-organic frameworks for separations. *Chem. Rev.* **2012**, *112* (2), 869-932.
18. Liu, J.; Chen, L.; Cui, H.; Zhang, J.; Zhang, L.; Su, C. Y., Applications of metal-organic frameworks in heterogeneous supramolecular catalysis. *Chem. Soc. Rev.* **2014**, *43* (16), 6011-61.
19. Horike, S.; Shimomura, S.; Kitagawa, S., Soft porous crystals. *Nat. Chem.* **2009**, *1* (9), 695-704.
20. Yoon, M.; Srirambalaji, R.; Kim, K., Homochiral metal-organic frameworks for asymmetric heterogeneous catalysis. *Chemical reviews* **2011**, *112* (2), 1196-1231.
21. Sumida, K.; Rogow, D. L.; Mason, J. A.; McDonald, T. M.; Bloch, E. D.; Herm, Z. R.; Bae, T. H.; Long, J. R., Carbon dioxide capture in metal-organic frameworks. *Chem. Rev.* **2012**, *112* (2), 724-81.
22. Horcajada, P.; Gref, R.; Baati, T.; Allan, P. K.; Maurin, G.; Couvreur, P.; Férey, G.; Morris, R. E.; Serre, C., Metal-organic frameworks in biomedicine. *Chem. Rev.* **2012**, *112* (2), 1232-68.
23. Wang, Z.; Cohen, S. M., Postsynthetic modification of metal-organic frameworks. *Chemical Society Reviews* **2009**, *38* (5), 1315-1329.
24. Eddaoudi, M.; Kim, J.; Rosi, N.; Vodak, D.; Wachter, J.; O'Keeffe, M.; Yaghi, O. M., Systematic design of pore size and functionality in isorecticular MOFs and their application in methane storage. *Science* **2002**, *295* (5554), 469-72.

25. Ockwig, N. W.; Delgado-Friedrichs, O.; O'Keeffe, M.; Yaghi, O. M., Reticular Chemistry: Occurrence and Taxonomy of Nets and Grammar for the Design of Frameworks. *Accounts of Chemical Research* **2005**, *38* (3), 176-182.
26. Toussaint, L.; Bertrand, L.; Hue, L.; Crichton, R. R.; Declercq, J. P., High-resolution X-ray structures of human apoferritin H-chain mutants correlated with their activity and metal-binding sites. *J. Mol. Biol.* **2007**, *365* (2), 440-52.
27. Huard, D. J.; Kane, K. M.; Tezcan, F. A., Re-engineering protein interfaces yields copper-inducible ferritin cage assembly. *Nat. Chem. Biol* **2013**, *9* (3), 169-76.
28. Li, H.; Eddaoudi, M.; O'Keeffe, M.; Yaghi, O. M., Design and synthesis of an exceptionally stable and highly porous metal-organic framework. *Nature* **1999**, *402* (6759), 276-279.
29. Barthelet, K.; Marrot, J.; Riou, D.; Ferey, G., A breathing hybrid organic-inorganic solid with very large pores and high magnetic characteristics. *Angew. Chem. Int. Ed. Engl.* **2002**, *41* (2), 281-4.
30. Santambrogio, P.; Pinto, P.; Levi, S.; Cozzi, A.; Roviola, E.; Albertini, A.; Artymiuk, P.; Harrison, P. M.; Arosio, P., Effects of modifications near the 2-, 3- and 4-fold symmetry axes on human ferritin renaturation. *Biochem. J.* **1997**, *322*, 461-8.
31. Lawson, D. M.; Artymiuk, P. J.; Yewdall, S. J.; Smith, J. M.; Livingstone, J. C.; Treffry, A.; Luzzago, A.; Levi, S.; Arosio, P.; Cesareni, G.; et al., Solving the structure of human H ferritin by genetically engineering intermolecular crystal contacts. *Nature* **1991**, *349* (6309), 541-4.
32. Avvaru, B. S.; Kim, C. U.; Sippel, K. H.; Gruner, S. M.; Agbandje-McKenna, M.; Silverman, D. N.; McKenna, R., A Short, Strong Hydrogen Bond in the Active Site of Human Carbonic Anhydrase II. *Biochemistry* **2010**, *49* (2), 249-251.
33. Liu, L.; Konstas, K.; Hill, M. R.; Telfer, S. G., Programmed pore architectures in modular quaternary metal-organic frameworks. *J. Am. Chem. Soc.* **2013**, *135* (47), 17731-4.
34. Grams, F.; Reinemer, P.; Powers, J. C.; Kleine, T.; Pieper, M.; Tschesche, H.; Huber, R.; Bode, W., X-ray structures of human neutrophil collagenase complexed with peptide hydroxamate and peptide thiol inhibitors. Implications for substrate binding and rational drug design. *Eur. J. Biochem.* **1995**, *228* (3), 830-41.
35. Puerta, D. T.; Cohen, S. M., Elucidating drug-metalloprotein interactions with tris(pyrazolyl)borate model complexes. *Inorg. Chem.* **2002**, *41* (20), 5075-82.
36. Schulze Wischeler, J.; Innocenti, A.; Vullo, D.; Agrawal, A.; Cohen, S. M.; Heine, A.; Supuran, C. T.; Klebe, G., Bidentate Zinc chelators for alpha-carbonic anhydrases that produce a trigonal bipyramidal coordination geometry. *ChemMedChem* **2010**, *5* (9), 1609-15.
37. Jani, M.; Tordai, H.; Trexler, M.; Banyai, L.; Patthy, L., Hydroxamate-based peptide inhibitors of matrix metalloproteinase 2. *Biochimie* **2005**, *87* (3-4), 385-92.
38. Neilands, J. B., Siderophores: Structure and Function of Microbial Iron Transport Compounds. *Journal of Biological Chemistry* **1995**, *270* (45), 26723-26726.

39. Crumbliss, A. L., Iron bioavailability and the coordination chemistry of hydroxamic acids. *Coord. Chem. Rev.* **1990**, *105*, 155-179.
40. Radford, R. J.; Tezcan, F. A., A superprotein triangle driven by nickel (II) coordination: exploiting non-natural metal ligands in protein self-assembly. *J. Am. Chem. Soc.* **2009**, *131* (26), 9136-9137.
41. Radford, R. J.; Nguyen, P. C.; Ditri, T. B.; Figueroa, J. S.; Tezcan, F. A., Controlled protein dimerization through hybrid coordination motifs. *Inorg. Chem.* **2010**, *49* (9), 4362-4369.
42. Radford, R. J.; Nguyen, P. C.; Tezcan, F. A., Modular and Versatile Hybrid Coordination Motifs on α -Helical Protein Surfaces. *Inorg. Chem.* **2010**, *49* (15), 7106-7115.
43. Radford, R. J.; Lawrenz, M.; Nguyen, P. C.; McCammon, J. A.; Tezcan, F. A., Porous protein frameworks with unsaturated metal centers in sterically encumbered coordination sites. *Chem. Commun.* **2011**, *47* (1), 313-315.
44. Burazerovic, S.; Gradinaru, J.; Pierron, J.; Ward, T. R., Hierarchical Self-Assembly of One-Dimensional Streptavidin Bundles as a Collagen Mimetic for the Biomineralization of Calcite. *Angewandte Chemie* **2007**, *119* (29), 5606-5610.
45. Broomell, C. C.; Birkedal, H.; Oliveira, C. L. P.; Pedersen, J. S.; Gertenbach, J.-A.; Young, M.; Douglas, T., Protein cage nanoparticles as secondary building units for the synthesis of 3-dimensional coordination polymers. *Soft Matter* **2010**, *6* (14), 3167-3171.
46. Jung, S.; Kim, Y.; Kim, S.-J.; Kwon, T.-H.; Huh, S.; Park, S., Bio-functionalization of metal-organic frameworks by covalent protein conjugation. *Chem. Commun.* **2011**, *47* (10), 2904-2906.
47. Griffith, D.; Krot, K.; Comiskey, J.; Nolan, K. B.; Marimon, C. J., Monohydroxamic acids and bridging dihydroxamic acids as chelators to ruthenium(III) and as nitric oxide donors: syntheses, speciation studies and nitric oxide releasing investigation. *Dalton Transactions* **2008**, 137-147.
48. Hassell, A. M.; An, G.; Bledsoe, R. K.; Bynum, J. M.; Carter, H. L., 3rd; Deng, S. J.; Gampe, R. T.; Grisard, T. E.; Madauss, K. P.; Nolte, R. T.; Rocque, W. J.; Wang, L.; Weaver, K. L.; Williams, S. P.; Wisely, G. B.; Xu, R.; Shewchuk, L. M., Crystallization of protein-ligand complexes. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2007**, *63* (Pt 1), 72-9.
49. Battye, T. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G., iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67* (Pt 4), 271-81.
50. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., Overview of the CCP4 suite and current developments. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67* (Pt 4), 235-42.
51. Vagin, A.; Teplyakov, A., MOLREP: an Automated Program for Molecular Replacement. *J. Appl. Crystallogr.* **1997**, *30* (6), 1022-1025.

52. Murshudov, G. N.; Vagin, A. A.; Dodson, E. J., Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1997**, *53* (Pt 3), 240-55.
53. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66* (Pt 2), 213-21.
54. Schüttelkopf, A. W.; van Aalten, D. M., PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60* (Pt 8), 1355-63.
55. Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60* (Pt 12 Pt 1), 2126-32.
56. DeLano, W. L., The PYMOL Molecular Graphics System (<http://www.pymol.org>). **2003**.
57. Bertini, I.; Lalli, D.; Mangani, S.; Pozzi, C.; Rosa, C.; Theil, E. C.; Turano, P., Structural Insights into the Ferroxidase Site of Ferritins from Higher Eukaryotes. *J. Am. Chem. Soc.* **2012**, *134* (14), 6169-6176.

Chapter 3. Synthetic modularity of ferritin-MOFs

3.1. Introduction

Periodic protein arrays constitute a significant component of the cellular machinery and are widely utilized as platforms for nano- and biotechnological applications due to their advanced materials properties and precise display of diverse chemical functionalities over the nm- μ m scale.¹⁻³ Accordingly, there has been a growing interest in the construction of artificial protein assemblies.⁴⁻⁷ These efforts have engendered innovative design strategies and chemical approaches, which have led to 0,⁸⁻¹¹ 1,¹²⁻¹⁵ 2,^{12, 16-18} and 3D^{7, 19-22} assemblies with complex structures, and in some cases, sophisticated²³⁻²⁴ and evolvable functions²⁵, as well as emergent properties not yet observed in biology.^{18, 26} Despite these advances, the structural/chemical heterogeneity of proteins still poses a substantial challenge for designing supramolecular protein architectures and arrays. While most design approaches are intended to be generalizable, they still require adjustment of the protein building blocks or the self-assembly procedures on a case-by-case basis. In contrast, the thematically related field of supramolecular chemistry has benefitted from synthetic access to an extensive library of molecular building blocks and bonding strategies,²⁷⁻²⁸ which can be mixed and matched to create structural and functional diversity with relative ease. Such synthetic versatility is aptly highlighted by metal-organic frameworks (MOFs),²⁹⁻³⁹ a vast class of crystalline materials composed of metal-based nodes and organic linkers that can be combined in a modular fashion.

As described in the previous chapter, we adopted the construction strategies of MOFs to create 3D protein lattices in which pseudo-spherical ferritin nodes decorated on their outer surfaces with metal coordination sites were bridged via ditopic metal-chelating ligands (*p*-benzene-1,4-dihydroxamate; *p*-bdh).⁴⁰ Despite the remarkable size mismatch between the *p*-H₂bdh ligands (ca. 9 Å long, 196 g/mol) and ferritin nodes (120 Å across, >505,000 g/mol), their interactions were sufficiently strong to dictate the arrangement of ferritin molecules into the

desired body-centered cubic (*bcc*) lattice. This system introduced a new class of tripartite hybrid materials whose self-assembly is dependent on each of its components: protein, metal, and an organic ligand. In this chapter, we examine if the modularity that is inherent in conventional MOFs and renders them particularly attractive as functional materials also apply to protein-MOFs. Specifically, we asked whether it is possible to *synthetically modulate* the structural parameters of protein-MOFs through discrete molecular interactions.

3.2. Results and discussion

3.2.1. Modularity in the ferritin node

To enable a modular construction approach, we sought to diversify the metal and the organic ligand components of the ferritin-MOFs. As discussed in Chapter 2, we deemed the C_3 symmetric pores of ferritin (essentially the eight vertices of a cube) as the most suitable locations for the installment of metal anchoring sites. These corner locations allow for the installation of surface-exposed tripodal coordination motifs that can bind various metal ions with high affinity, while still presenting outward-facing open coordination sites available for organic linkers. In Chapter 2, we established that ^{122}Hf -ferritin bound Zn^{2+} ions in the expected tetrahedral geometry with a tripodal base of three His side chains and a single, solvent-bound coordination site pointing to the protein exterior (Zn-ferritin).⁴⁰ The Ni- and Co- adducts of ^{122}Hf -ferritin (Ni- and Co-ferritin) were prepared similarly to Zn-ferritin, whereby a significant molar excess of Ni^{2+} or Co^{2+} was added to ferritin solutions to saturate all available metal-binding sites within the protein cage, including the installed tripodal coordination motifs at the C_3 vertices. Ni- and Co-ferritin were crystallized by vapor diffusion in the presence of Ca^{2+} , which selectively induces the formation of face-centered cubic (*fcc*; *F432* space group) lattices by connecting ferritin molecules across their C_2 symmetric interfaces via coordination to pairs of ^{84}Asp and engineered ^{86}Gln residues.⁴¹ The crystal structures of Ni- and Co-ferritin were determined at 1.79 and 1.95 Å resolution, respectively (PDB IDs: 5UP7 and 5VTD). These structures confirm that Ni^{2+} and Co^{2+} ions are anchored by the tripodal ^{122}His coordination motifs at full occupancy, with three aquo ligands completing nearly ideal octahedral coordination spheres (Figure 3.1). Relative to the tetrahedral Zn coordination sites in Zn-ferritin, in Ni- and Co-ferritin, the ^{122}His side chains have moved slightly to accommodate the octahedral geometry, with $\text{N}_{\text{His}}\text{-M-N}_{\text{His}}$ angles of 95° to 96° , $\text{N}_{\text{His}}\text{-M-OH}_2$ angles of 94° (Ni) and 87° (Co), and $\text{OH}_2\text{-M-OH}_2$ angles of 75° (Ni) and 82° (Co). The M-N_{His} distances are 2.1 Å for both nodes, and the M-OH_2 distances are 2.1 Å for Ni-ferritin and 2.2 Å for Co-ferritin.

These observations establish that a series of metals can be displayed at the 3-fold vertices, and like Zn-ferritin, the Ni- and Co-ferritin nodes are poised for coordinating bridging ligands.

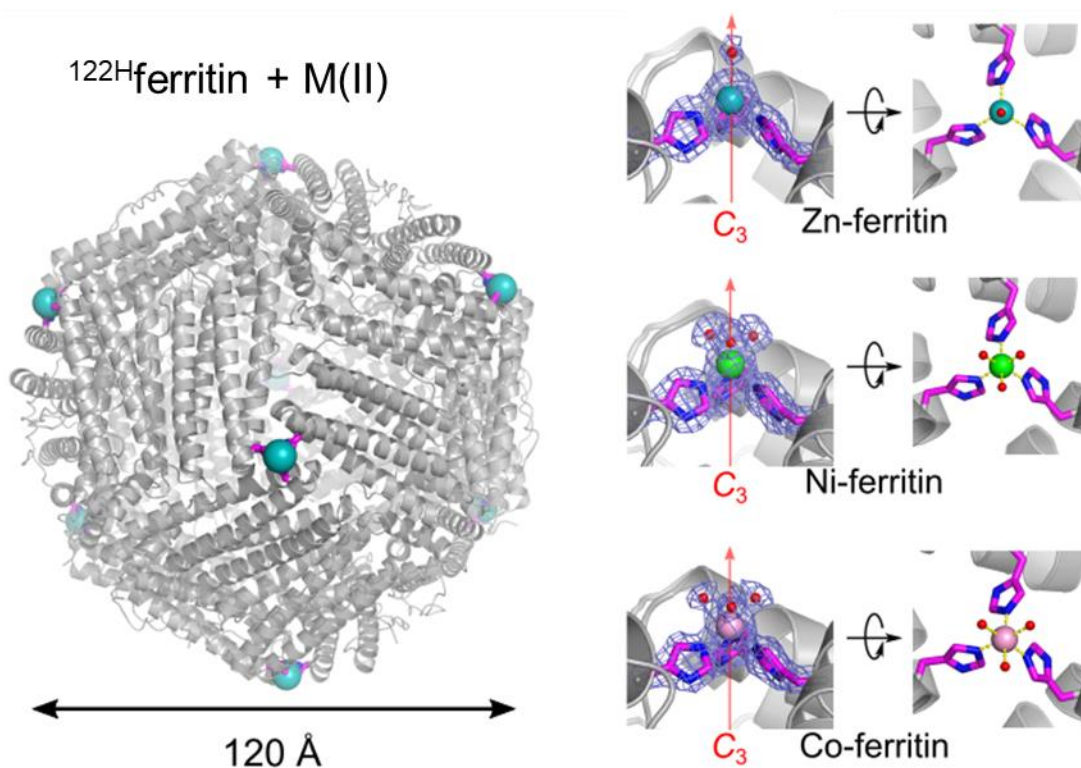


Figure 3.1. Ferritin nodes crystallized with Zn²⁺, Ni²⁺, and Co²⁺. a) The ferritin node (metal ions at the C₃ vertices are shown as teal spheres). Close-up views of the tetragonal Zn²⁺, octahedral Ni²⁺ and Co²⁺ coordination environments. Water molecules are shown as red spheres. 2F_o-F_c maps are contoured to 1σ.

3.2.2. Modularity in the bridging ligand

With the three distinct metallo-ferritin nodes in hand, we next prepared a library of structurally diverse ditopic ligands bearing hydroxamic acid head groups (Figure 3.2). Hydroxamate functionalities were chosen due to their high affinity for Zn^{2+} , Ni^{2+} , and Co^{2+} , as well as their sterically unencumbered nature compared to other commonly used aromatic chelates (e.g., polypyridyl or catecholate-type ligands), allowing for unhindered access to the surface-anchored metal ions without bias from peripheral contacts. The primary differences between these ligands can be described by three parameters: First, the inter-hydroxamate spacing (i.e., ligand length) of these ligands ranges from ca. 7.0 Å to >18 Å. Second, the ligands offer four different geometries in terms of the relative orientations of the two hydroxamate head groups. We define these orientations for the vector along the C-C bond (highlighted in red in Figure 3.2) that appends the hydroxamate moieties to the core of each ligand. Based on this criteria, the ditopic ligands are divided into four categories (collinear, parallel-offset, bent, and flexible). Third, the composition of the spacers can contain heteroatoms (oxygen, nitrogen, or sulfur) to enforce the planarity of the aromatic rings or enhance the solubility of the compounds.

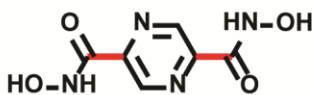
Para-benzene dihydroxamic acid (*p*-H₂bdh) was synthesized through a base-mediated amidation reaction from dimethyl terephthalate, as described previously.^{40, 42} Both *meta*-benzene dihydroxamic acid (*m*-H₂bdh) and 2,6-naphthalene dihydroxamic acid (2,6-H₂ndh) were synthesized using a similar base-mediated amidation with minor modifications. We decided to explore other protocols for the synthesis of hydroxamic acids due to the harsh alkaline conditions, solubility concerns with the longer hydrophobic ligands, and the relative abundance of commercially available ditopic carboxylic acids (when compared to methyl- or ethyl-esters). While there are many established methods for the conversion of carboxylic acids to hydroxamic acids, we selected carbodiimide-based amidation using protected hydroxylamines for the synthesis of the remaining ligands. The commercially available *o*-benzylhydroxylamine hydrochloride was

chosen as the protected hydroxylamine for the synthesis of 2,5-furan dihydroxamic acid (2,5-H₂fhd), *para*-xylene dihydroxamic acid (*p*-H₂xhd), 2,6-pyridine dihydroxamic acid (2,6-H₂pydh), diglyme dihydroxamic acid (H₂gdh), and the sulfonated variant of 2,6-H₂ndh (2,6-H₂sndh). After amidation, the benzyl group was removed by reduction with Pd/C under an H₂ atmosphere. Since this deprotection is not suitable for all of the ligands (i.e., *E*-edh), we employed the acid-labile *O*-2-methylprenyl hydroxylamine that can be removed under acidic conditions.⁴³ This new coupling partner enabled the synthesis of the otherwise inaccessible *E*-ethenedihydroxamate (*E*-edh). However, the synthesis of *O*-2-methylprenyl hydroxylamine required multiple steps and gave relatively poor yields. Finally, we settled on *O*-tritylhydroxylamine as the ideal protected hydroxylamine. *O*-tritylhydroxylamine can be synthesized in two steps in relatively high yields and removed under acidic conditions. Also, the majority of trityl-protected dihydroxamic acids can be purified by selectively precipitating the relatively insoluble *O*-trityl dihydroxamic acids out of solution. *O*-tritylhydroxylamine was used for synthesizing nearly all of the ligands and is highly recommended for future syntheses of hydroxamic acids.

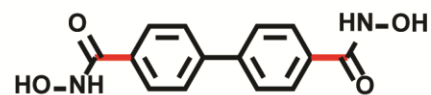
a) Collinear



p-H₂bdh

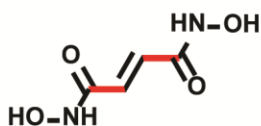


p-H₂pzd

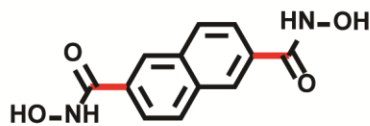


4,4'-H₂bpdh

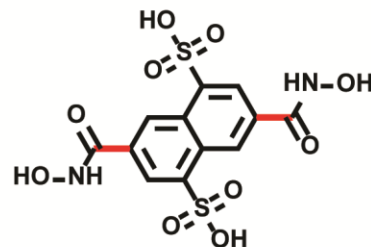
b) Parallel offset



E-H₂edh

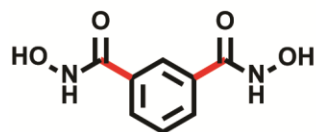


2,6-H₂ndh

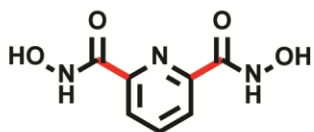


2,6-H₂sndh

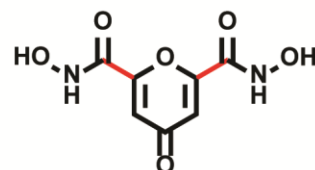
c) Bent



m-H₂bdh



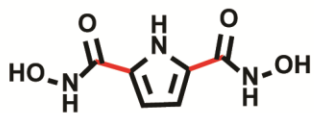
2,6-H₂pydh



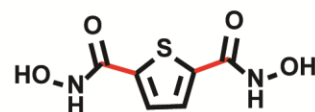
2,6-H₂podh



2,5-H₂fdh

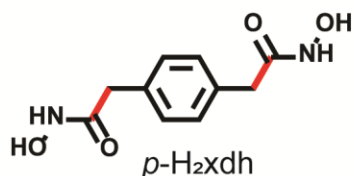


2,5-H₂pdh

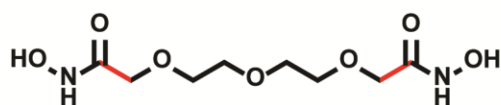


2,5-H₂tdh

d) Flexible



p-H₂xdh



H₂dgdh

Figure 3.2. The library of ditopic dihydroxamic acid bridging ligands categorized by the relative geometry of the C-C bond (shown in red).

3.2.3. General considerations for the synthesis of ferritin-MOFs

The assembly of ferritin-MOFs requires a precise balance of the three components: protein, metal, and ligand. Extensive screening efforts have allowed us to establish a series of guidelines for the assembly of protein-MOFs. Unlike conventional protein crystallization, the formation of ferritin-MOF crystals does not require a reservoir solution; assembly is driven by the metal-ligand-metal coordination, as opposed to the slow concentration and subsequent precipitation of proteins. Therefore, ferritin-MOFs can be synthesized in droplets (2-20 μL) using conventional crystal trays (where a reservoir solution is added to prevent substantial solvent evaporation during crystallization) or in bulk solution (200+ μL) in 24-well culture plates sealed with tape. While it is also possible to synthesize ferritin-MOFs in larger volumes, we had the most success with these two setups. Ferritin has a series of metal-binding residues across the protein surface, buried in the pores, and in the active sites of each monomer, making it is essential to use a moderate excess of metal to ensure all of the 3-fold metal-binding sites are fully occupied. However, if the concentration of either the protein or metal is too high, the weakly coordinating surface residues induce the formation of metal-mediated protein aggregates. Regardless of the ligand used, we had the most success with ferritin-MOF self-assembly with 2-4 μM of ferritin (by 24-mer) and 50-125 equivalents of metal. There seems to be little dependence on the specific identity of the metal (Zn^{2+} , Ni^{2+} , or Co^{2+}). The bridging ligands were always employed at a significant excess of 250-500 equivalents. At these concentrations, the solubility of the ligand must be considered. Deprotonation of the amide hydrogen (pK_a 8.8 - 9.5) substantially increases the solubility of the ligand.⁴⁴ Therefore, the ligand was prepared in a buffered solution at or above pH 9.5. The crystallization conditions varied from pH 8.5-10.0, and most of the ferritin-MOFs were prepared at pH 9.5. *N*-Cyclohexyl-2-aminoethanesulfonic acid (CHES) was used to buffer the solutions since it is a non-chelating ligand with the appropriate pH range. A subset of fifteen ferritin-MOF from the non-heteroatom containing ligand and Zn^{2+} , Ni^{2+} , and Co^{2+} crystallized under the conditions described above (Figure 3.3).

Precipitating agents, such as polyethylene glycol (PEG), are commonly employed during conventional protein crystallization to decrease the solubility of the protein and increase the likelihood of crystal formation. These precipitating agents fill a similar role in the synthesis of Protein-MOFs. A large screen of >15 precipitating agents at varying concentrations, including PEGs with molecular weights between 300-10000, pentaerythritol ethoxylate and propoxylate, glycerol, 2-methyl-2,4-pentanediol (MPD), ethylene glycol, and 1,2-pentanediol was employed. At concentrations < 12% for most of the precipitants, (< 8% for high molecular weight PEGs), crystal growth was improved with no observable change in crystal morphology (rhombic dodecahedrons). The higher precipitant concentrations had the propensity to form octahedron-shaped crystals, which are indicative of the F432 morphology (confirmed by sc-XRD studies). Out of all of the precipitants, pentaerythritol propoxylate 5/4 PO/OH (PEP) was selected as it consistently produced the best crystals.

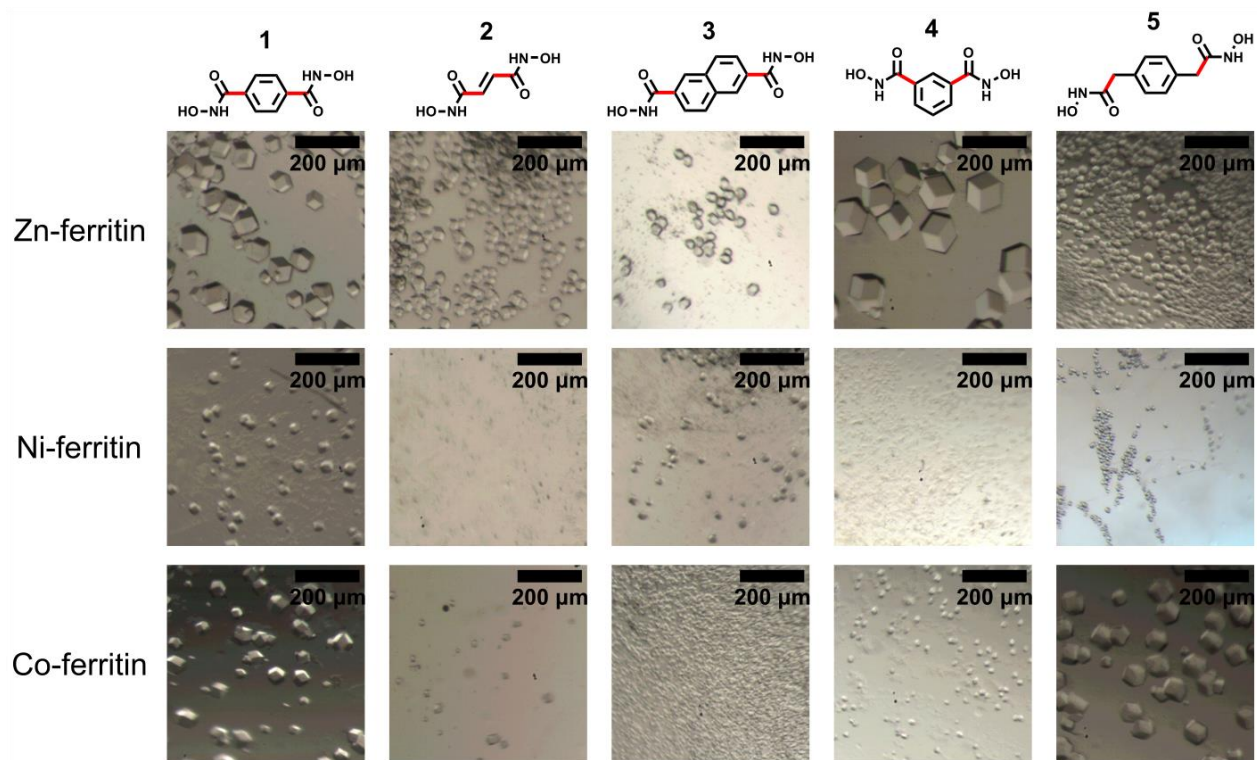


Figure 3.3. Light micrographs of crystals from fifteen ferritin-MOFs.

3.2.4. Ligands that failed to form ferritin-MOFs

The majority of the dihydroxamate ligands produced the expected rhombic dodecahedron-shaped crystals under the conditions described above. Out of the fourteen ligands, only five were unable to reproducibly induce the assembly of the ferritin nodes into observable crystals. Each of these ligands has distinct attributes that likely inhibits protein-MOF synthesis. H₂dgdh is likely too flexible to template an ordered framework and can form dihydroxamate metal chelates to a single metal atom. Both 2,6-pyridine dihydroxamic acid (2,6-H₂pydh) and *para*-pyrazine dihydroxamic acid (*p*-H₂pzdhd) have aromatic nitrogen atoms that are near the hydroxamic acid. In small-molecule complexes, such as Zn-trispyrazolylborate (a tris-His mimic), similar ligands have been observed preferentially coordinating through the aromatic nitrogen and the carbonyl oxygen instead of through the hydroxamate.⁴⁵ This type of coordination may preclude ferritin-MOF formation, as the ligand is no longer able to bridge between two adjacent nodes. The sulfonate groups greatly enhance the solubility of 2,6-H₂ndhd and are strategically positioned far away from the hydroxamates to prevent the formation of an off-pathway chelating motif. The most likely scenario is that the sulfonate groups interact unfavorably with the protein surface (either due to charge or steric clashes) around the 3-fold axis. It is not clear if collinear sulfonated ligands would have similar problems. The 4,4'-biphenyl dihydroxamic acid (4,4'-H₂bpdhd) ligand was not soluble under the conditions that favor ferritin-MOF formation. Attempts to artificially enhance the solubility of the ligands (by adding ca. 10% v/v DMSO or DMF) did not yield any protein-MOF crystals. While the protein is stable under these conditions, the addition of these organic solvents may indirectly impact the geometry of the ¹²²Ni binding site and inhibits crystallization. In an early attempt to probe metal coordination environment, monotopic ligands dissolved in 10% DMSO were soaked into preformed F432 crystals of Ni-ferritin. The crystal structure of Ni-ferritin with 10% (v/v) DMSO was determined to a resolution of 1.70 Å. Unexpectedly, the Ni-¹²²Ni coordination site is inverted in these crystals and faces towards the interior of the ferritin cage (Figure 3.4). The N_{His}-Ni-N_{His} angle is 98° and the OH₂-Ni-OH₂ angle is 82°, which are both very

similar to the previously described outward-facing $N_{\text{His}}\text{-Ni-N}_{\text{His}}$ and $\text{OH}_2\text{-Ni-OH}_2$ angles (95° and 75° respectively). The Ni-N_{His} and Ni-OH_2 distances are both 2.1 \AA , which is identical to the previous structure. Unlike the outward-facing coordination site, the metal in this structure is inaccessible to a bridging ligand and would likely be unable to form ferritin-MOFs. This structure highlights the importance of carefully screening assembly conditions, as the addition of seemingly inert components may lead to unforeseen consequences.

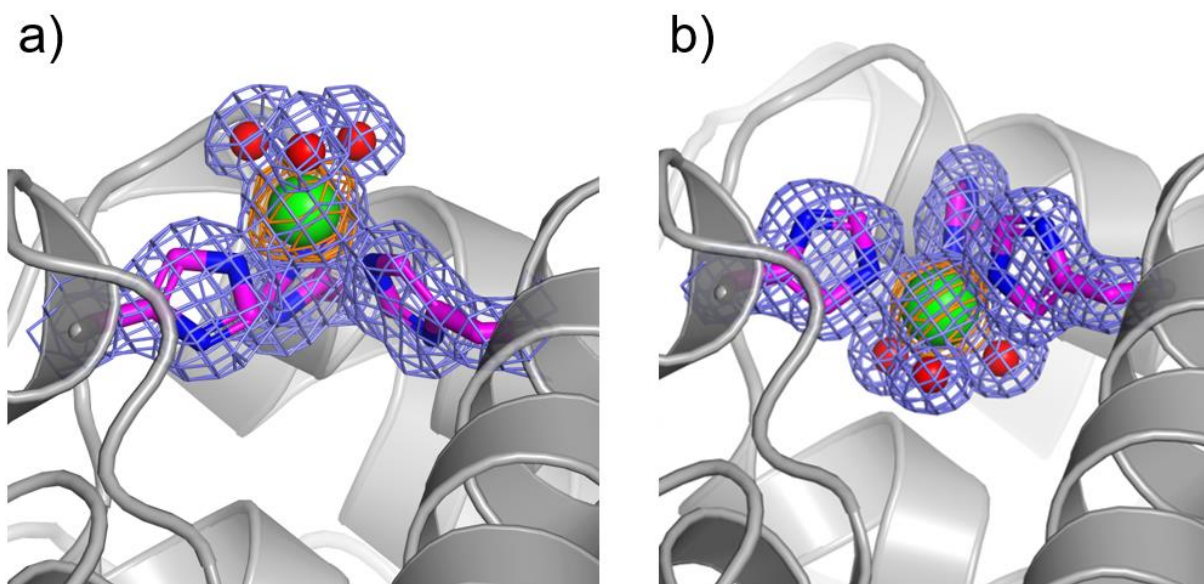


Figure 3.4. The impact of DMSO on the tris-His coordination site. The structures of outward-facing (a) and inward-facing (b) $\text{Ni-}^{122}\text{His}$ coordination sites are shown. The inward-facing coordination site only occurs after the addition of 10% DMSO to the crystallization mixture. The ^{122}His sidechains are shown as magenta sticks, the Ni is shown as a green sphere, and water molecules are shown as red spheres. $2F_o-F_c$ maps are contoured to 1σ (blue) or 3σ (orange).

3.2.5. Small-angle X-ray scattering as a tool for ferritin-MOF characterization *in situ*

Powder X-ray diffraction (PXRD) is a powerful characterization tool for conventional MOFs. Unlike single-crystal X-ray diffraction (sc-XRD), which requires highly-ordered and relatively large crystals, PXRD can often be used to determine the unit cell parameters for microcrystalline samples with less structural order. Additionally, PXRD allows for the analysis of heterogeneous mixtures containing multiple crystallites. Ferritin-MOFs are incompatible with standard PXRD instrumentation due to the large size of the unit cell parameters ($a = 148\text{-}163 \text{ \AA}$). However, small-angle X-ray scattering (SAXS) instrumentation can be utilized to collect powder diffraction data on ferritin-MOFs. The long sample-detector distance (commonly 1-8 m) allows for resolution of the low angle reflections, which can be used for unit cell determination of the ferritin-MOFs. We primarily utilized synchrotron SAXS beamlines (either BL 4-2 at SSRL or BL 5-ID-D at APS) for the characterization of the protein-MOFs presented in this work due to the high precision needed to determine subtle differences in peak splitting between different lattice symmetries.

To assess the accuracy of these powder measurements, we first collected diffraction data on the crystallographically characterized *p*-bdh-Zn-ferritin in solution. The unit cell parameters of *p*-bdh-Zn-ferritin determined by SAXS (*bcc*, $a = 155.97 \text{ \AA}$) are nearly identical to those derived from sc-XRD (*bcc*, $a = 155.81 \text{ \AA}$), shown in Figure 3.5. This diffraction pattern is distinguishable from the *fcc* lattice ($a = 182.13 \text{ \AA}$) that forms when the Zn-ferritin nodes are treated with Ca^{2+} instead of *p*-H₂bdh. These results highlight the effectiveness of employing SAXS beamlines for the unit cell determination of ferritin-MOFs.

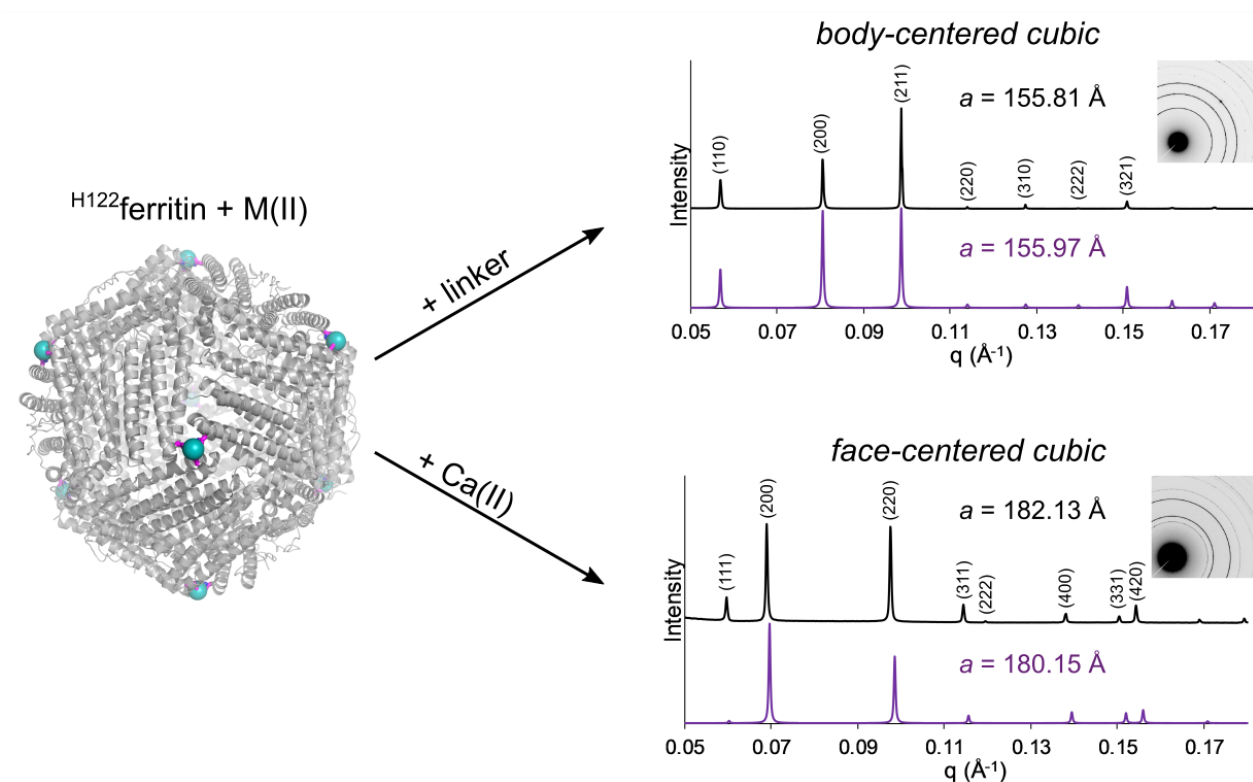


Figure 3.5. Comparison of the SAXS profiles of a ligand-mediated body-centered cubic lattice and a Ca-mediated face-centered cubic lattice. The simulated SAXS profiles (magenta) were generated from the crystal structure coordinates of Ni-ferritin (5UP7) and 1-Zn-ferritin (5UP8) by using the powder diffraction modeling functionality of Mercury.⁴⁶

To avoid potential complications, such as precipitant-mediated crystallization, we first crystallized the ferritin-MOFs under precipitant-free conditions. Out of the nine ligands that readily produce ferritin-MOF, only 2,5-pdh and 2,6-podh failed to form crystals (regardless of metal) under these conditions. Figure 3.6 shows the SAXS profiles and unit cell parameters for the 21 ferritin-MOFs that crystallized under precipitant-free synthesis conditions.

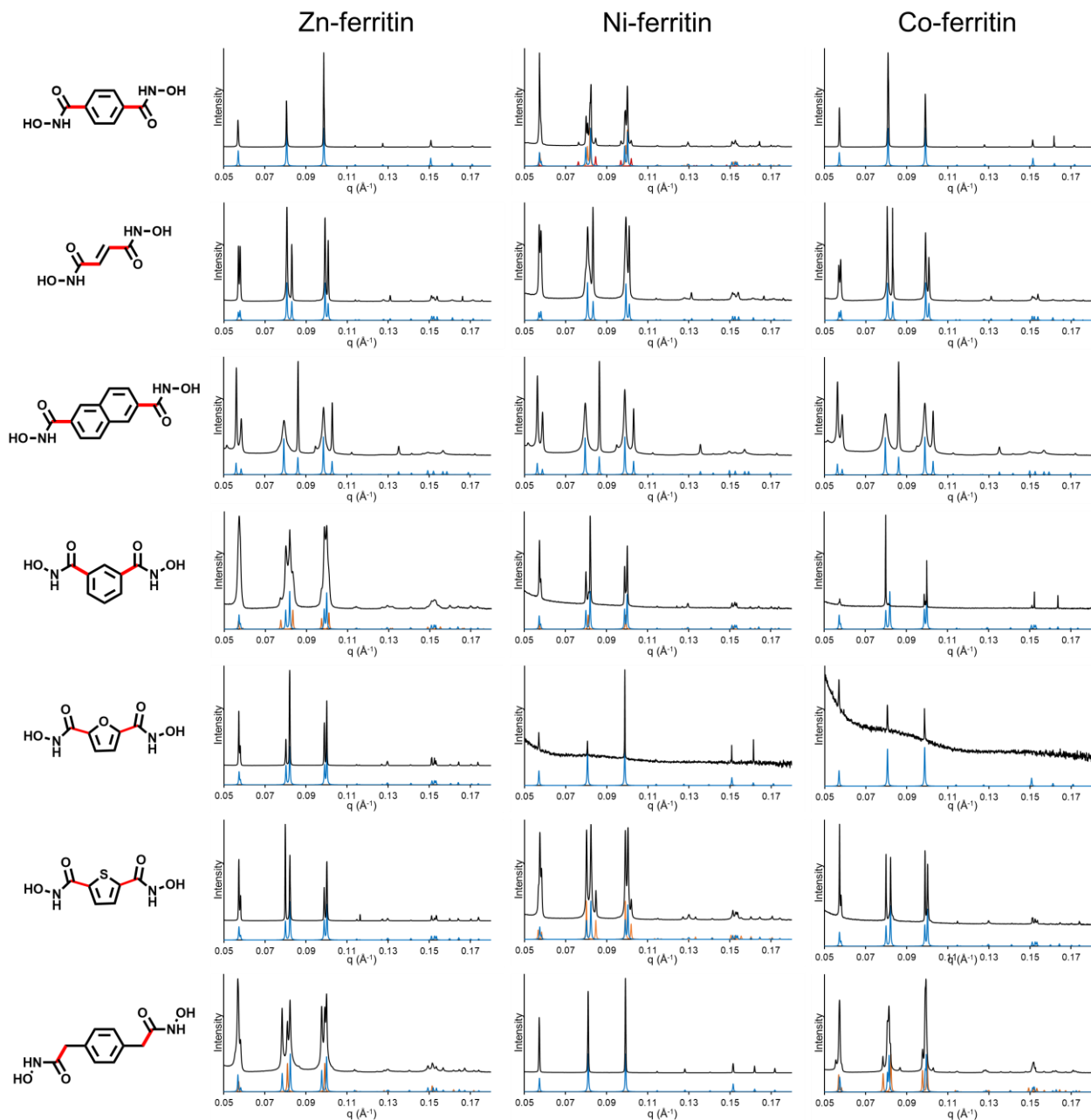
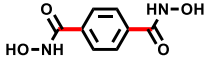
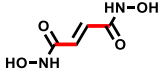
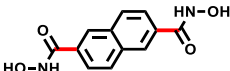
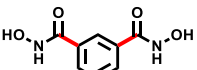
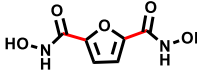
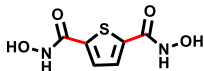
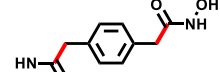


Figure 3.6. Diffraction patterns of twenty-one ferritin-MOFs. The data for each ferritin-MOF is shown in black. The simulated pattern for the primary lattice is shown in blue. Possible indexing solutions for additional lattices are shown in orange and red.

Table 3.1. Unit cell parameters for 21 protein-MOFs without precipitant.

Linker	Zn-ferritin		Ni-ferritin		Co-ferritin	
	Symmetry	Unit Cell	Symmetry	Unit Cell	Symmetry	Unit Cell
	Cubic <i>I</i>	$a = 155.97 \text{ \AA}$	Tetragonal <i>I</i>	$a = 152.89 \text{ \AA}$ $c = 157.48 \text{ \AA}$	Cubic <i>I</i>	$a = 155.23 \text{ \AA}$
			Tetragonal <i>I'</i> *	$a = 153.51 \text{ \AA}$ $c = 155.91 \text{ \AA}$		
			Tetragonal <i>I''</i> **	$a = 148.58 \text{ \AA}$ $c = 164.92 \text{ \AA}$		
	Tetragonal <i>I</i>	$a = 155.90 \text{ \AA}$ $c = 151.36 \text{ \AA}$	Tetragonal <i>I</i>	$a = 155.77 \text{ \AA}$ $c = 150.74 \text{ \AA}$	Tetragonal <i>I</i>	$a = 156.01 \text{ \AA}$ $c = 151.18 \text{ \AA}$
	Tetragonal <i>I</i>	$a = 158.45 \text{ \AA}$ $c = 145.93 \text{ \AA}$	Tetragonal <i>I</i>	$a = 158.05 \text{ \AA}$ $c = 145.46 \text{ \AA}$	Tetragonal <i>I</i>	$a = 157.81 \text{ \AA}$ $c = 146.00 \text{ \AA}$
	Tetragonal <i>I</i>	$a = 153.53 \text{ \AA}$ $c = 157.26 \text{ \AA}$	Tetragonal <i>I</i>	$a = 153.11 \text{ \AA}$ $c = 157.02 \text{ \AA}$	Tetragonal <i>I</i>	$a = 153.50 \text{ \AA}$ $c = 157.28 \text{ \AA}$
	Tetragonal <i>I'</i>	$a = 150.76 \text{ \AA}$ $c = 162.08 \text{ \AA}$	Cubic <i>I'</i>	$a = 154.86 \text{ \AA}$		
	Tetragonal <i>I</i>	$a = 153.11 \text{ \AA}$ $c = 156.63 \text{ \AA}$	Cubic <i>I</i>	$a = 155.91 \text{ \AA}$	Cubic <i>I</i>	$a = 155.96 \text{ \AA}$
	Tetragonal <i>I</i>	$a = 152.89 \text{ \AA}$ $c = 157.17 \text{ \AA}$	Tetragonal <i>I</i>	$a = 152.65 \text{ \AA}$ $c = 156.74 \text{ \AA}$	Tetragonal <i>I</i>	$a = 152.97 \text{ \AA}$ $c = 157.11 \text{ \AA}$
			Tetragonal <i>I'</i>	$a = 156.75 \text{ \AA}$ $c = 148.13 \text{ \AA}$		
	Tetragonal <i>I</i>	$a = 152.67 \text{ \AA}$ $c = 160.27 \text{ \AA}$	Cubic <i>I</i>	$a = 155.16 \text{ \AA}$	Tetragonal <i>I</i>	$a = 154.44 \text{ \AA}$ $c = 156.05 \text{ \AA}$
	Cubic <i>I'</i>	$a = 155.24 \text{ \AA}$			Tetragonal <i>I'</i>	$a = 152.85 \text{ \AA}$ $c = 160.19 \text{ \AA}$

* Simulated SAXS pattern is shown in orange in Figure 3.6.

** Simulated SAXS pattern is shown in red in Figure 3.6.

‡ The additional peaks observed in *p*-bdh-Ni-ferritin, 2,6-ndh-Zn, Ni-, and Co-ferritin, and *m*-bdh-Zn-ferritin samples may be due to a less symmetric primitive lattice.

As expected, these ferritin-MOFs gave well-defined diffraction profiles, which allowed us to determine their lattice symmetries and unit cell parameters. An initial inspection of the diffraction data indicates that bulk samples of most ferritin-MOF variants consist of primarily *bcc* (cubic *I*) and *bct* (tetragonal *I*) lattices. A cartoon explaining the relative ferritin positioning for each lattice symmetry is shown in Figure 3.7. All of the lattices exhibit symmetry and possess unit cell dimensions that fall within $155 \pm 10 \text{ nm}$ (Table 3.1). These observations suggest that each ferritin-MOF forms through the desired, ligand-mediated association of the metal centers at the C_3

vertices of ferritin nodes. While most of the diffraction peaks for the ferritin-MOFs could be assigned to a single lattice, certain metal/ligand combinations had additional unexplained reflections. These peaks were initially assigned to different lattices (shown in different colors in Figure 3.6 and separately listed in Table 3.1), consistent with the expectation that ferritin-MOFs may be inherently flexible and that subtle changes in the metal-ligand coordination and linker geometry can give rise to measurable changes in the lattice parameters. The most prominent example is provided by the simultaneous observation of three distinct *bct* lattices of *p*-bdh-Ni-ferritin whose unit cell parameters vary by ca. 7% (mean dimension: $154.3 \pm 10.0 \text{ \AA}$). This variation points to different extents of tetragonal distortion, i.e., the difference between the *a* and *c* dimensions, which is related to the displacement of the linker-bound metal ions from the body diagonals (see Figure 3.8 for structural models). Both *m*-bdh-Zn and 2,5-tdh-Ni-ferritin samples have reflections that could correspond to significantly different tetragonal lattices. In the bulk samples of *m*-bdh-Ni- and *p*-xdh-Zn-ferritin, both tetragonal and cubic lattices are observed, highlighting that the flexibility in the linker geometry can allow the ferritin C_3 axes to align with or deviate from the body diagonals.

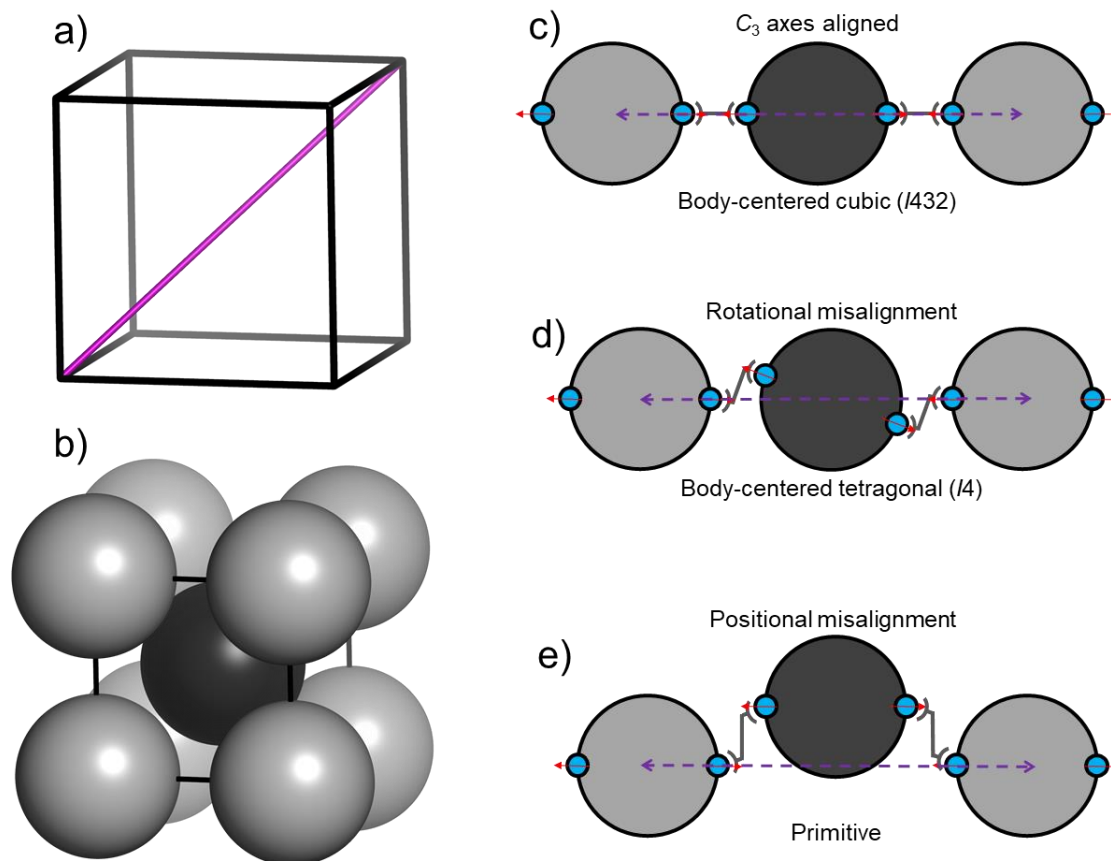


Figure 3.7. Possible packing geometries for ferritin-MOFs. a) The body diagonal of the unit cell is shown in purple. b) ferritin nodes are shown as grey spheres (located on the vertices) and a black sphere (located near the center). The positioning of the ferritin nodes along the body diagonal (purple lines) are shown for *bcc* (c), *bct* (d), and primitive (e) lattice symmetries.

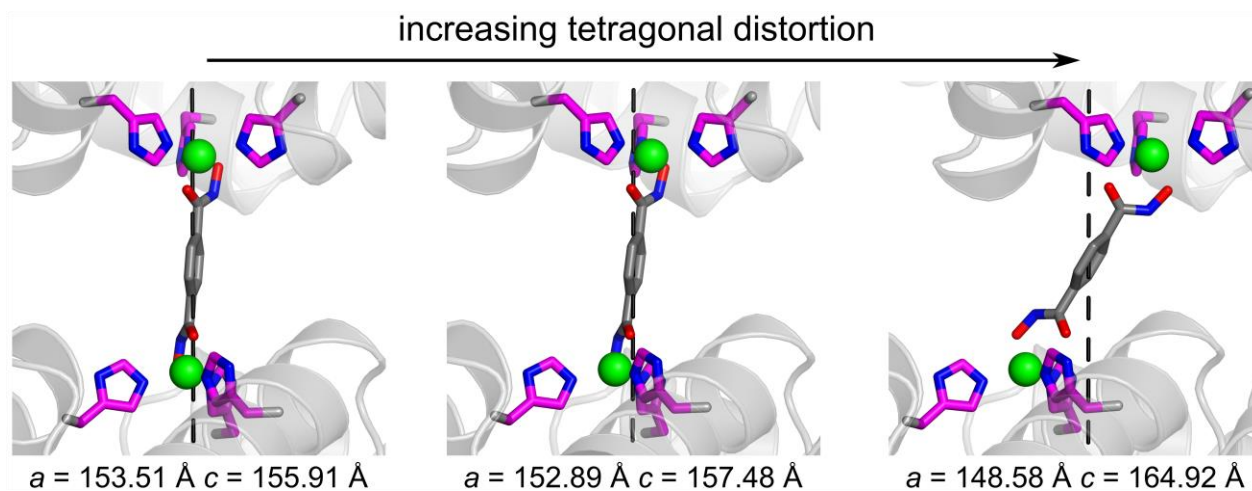


Figure 3.8. Structural models for the three lattices of observed *p*-bdh-Ni-ferritin. As the linker adopts a coordination geometry closer to octahedral, cell parameter c increases, and cell parameter a decreases. The Ni-ferritin positions in these structural models are identical to the coordinates used to simulate the diffraction patterns. The ligand was manually modeled to satisfy a reasonable metal-linker coordination geometry while maintaining the structural equivalence of the metal sites.

Another possible explanation for the appearance of additional reflections is a slight deviation (in either position or rotation) of the central ferritin in the *bct* lattice, which breaks the symmetry and leads to a lower-symmetry primitive lattice with larger cell dimensions. In this case, the $c_{primitive}$ unit cell parameter of the primitive lattice will be nearly identical to c_{bct} , while the $a_{primitive}$ unit cell parameter of the primitive lattice will have approximate dimensions of $\sqrt{2} \times a_{bct}$. A lower symmetry primitive lattice was first identified as a possibility from the 2,6-ndh-(Zn/Ni/Co)-ferritin diffraction patterns and later confirmed through SC-XRD. A model showing these lattices is shown in Figure 3.9. At the time the unit cells corresponding to the additional reflections were determined, little SC-XRD data had been collected on the *bct* protein-MOFs. We now consider it extremely likely that the additional reflections observed in *p*-bdh-Ni-ferritin, *m*-bdh-Zn-ferritin, and 2,6-ndh-(Zn/Ni/Co)-ferritin are indicative of a primitive “pseudo-*bct*” lattice and not due to separate crystalline species. However, the mixed *bcc/bct* lattices observed in *m*-bdh-Ni-ferritin, *p*-xdh-Zn-ferritin, and the *bct/bct* lattices of *p*-xdh-Co-ferritin and 2,5-tdh-Ni are probably separate crystalline species. The presence of highly symmetric peaks in their diffraction patterns cannot be explained by a less symmetric primitive cell.

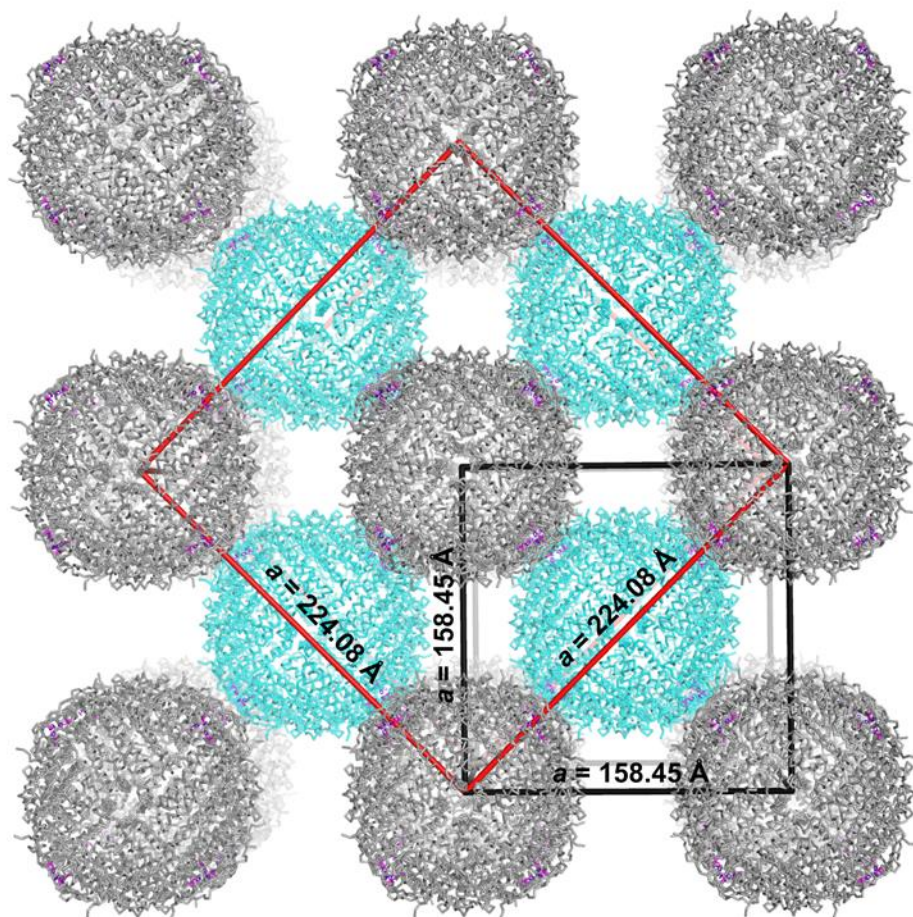


Figure 3.9. Two possible unit cells, *bct* (black) and primitive (red), for 2,6-ndh-Zn-ferritin. Both unit cells have an identical *c* parameter (145.93 Å) but differ with respect to the positioning of the ferritin molecules highlighted in cyan. In the *bct* lattice, the highlighted ferritins would be positioned precisely in the center of eight adjacent ferritins shown in grey, leading to the unit cell outlined in black. In the primitive lattice, the highlighted ferritins would not be positioned in the center of these eight adjacent ferritin molecules due to slightly different ligand-mediated ferritin orientations, leading to a unit cell similar to the one outlined in red.

A closer examination of the SAXS data also reveals some trends related to the metal and ligand components of ferritin-MOFs. The collinear *p*-H₂bdh produces isostructural lattices with both Zn-ferritin and Co-ferritin. Both of these lattices have been consistently observed under a variety of conditions, including when crystallized with precipitants such as PEP (Figure 3.10). As discussed previously, *p*-bdh-Ni-ferritin likely forms a less symmetric primitive cell, indicating that the ligand does not align with *C*₃ axes of neighboring ferritin nodes. The broad peaks observed for *p*-bdh-Ni-ferritin crystallized with PEP suggest that the precise order of this lattice has

decreased. These observations can be ascribed to the similarities between Zn^{2+} and Co^{2+} in terms of their coordination preferences for a tetrahedral geometry (as well as similar ionic radii and reactivity patterns), whereas Ni^{2+} typically tends toward higher coordination numbers and an octahedral geometry.⁴⁷ For $p\text{-H}_2\text{bdh}$ -based ferritin-MOFs, we can determine the coordination geometry between the hydroxamate head groups and the metal based on the lattice symmetry. Indeed, both $p\text{-bdh}(\text{Zn})$ -ferritin and $p\text{-bdh}(\text{Co})$ -ferritin coordinate the hydroxamate head groups in a pseudo-tetrahedral geometry, as this is the only possible coordination that leads to *bcc* symmetry (Figure 3.11 a). Since $p\text{-bdh}(\text{Ni})$ -ferritin crystallizes in a lower symmetry space group, the hydroxamate-Ni coordination must be octahedral (Figure 3.11 b). However, these comparisons only hold for collinear bridging ligands. It is possible for non-collinear ligands to form *bcc* lattices with both tetrahedral and octahedral coordination (Figure 3.11 c-d).

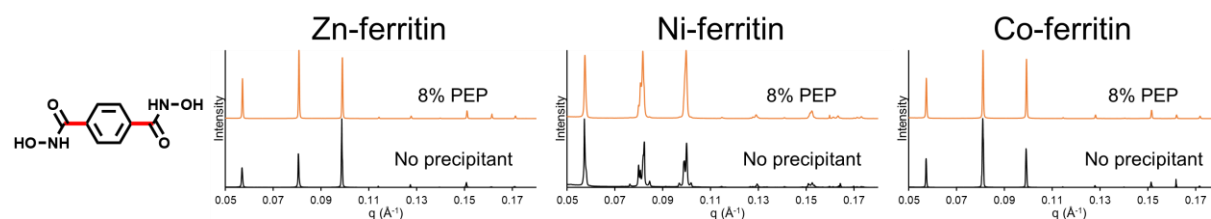


Figure 3.10. Diffraction patterns of $p\text{-H}_2\text{bdh}$ -based ferritin-MOFs. Crystals formed under precipitant-free conditions are shown in black. Crystals formed with PEP are shown in orange.

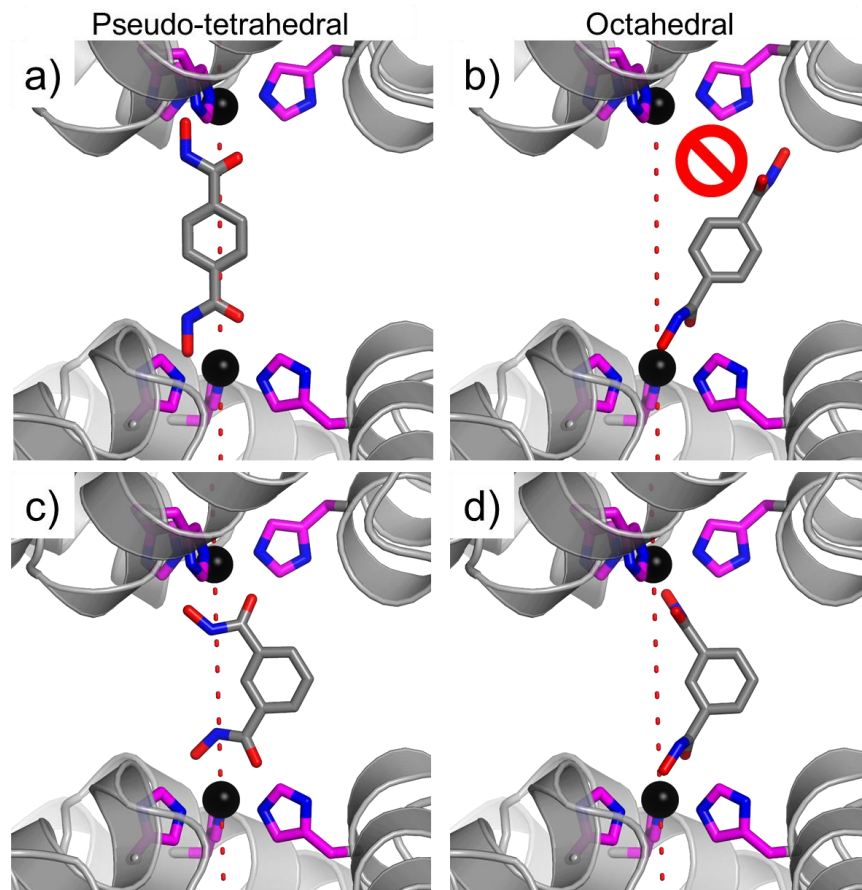


Figure 3.11. Models depicting the coordination geometry for collinear and bent ligands. Lattice symmetry is not correlated with metal coordination geometry for noncollinear ligands. A collinear ligand can form a *bcc* lattice with pseudo-tetrahedral (a), but not octahedral (b), coordination geometry. Both pseudo-tetrahedral (c) and octahedral (d) coordination geometries can form a *bcc* lattice with a noncollinear ligand.

Unlike *p*-H₂bdh, which exhibits a robust geometric preference based on the metal, both parallel-offset ligands (*E*-H₂edh and 2,6-H₂ndh) always yield tetragonal lattices regardless of the metal component. All of these tetragonal lattices have two long sides (*a*) and one short side (*c*). The addition of PEP to increase the yield of these ferritin-MOFs has no impact on the lattice symmetry (Figure 3.12). Since these two ligands display the same relative hydroxamate orientations (parallel-offset), they provide a valid point of comparison in terms of how the ligand dimensions affect lattice metrics. Indeed, we observe that the longer ligand 2,6-H₂ndh consistently produces a more substantial tetragonal distortion ($|a-c| \geq 11.8 \text{ \AA}$) than *E*-H₂edh ($|a-c| \leq 5.0 \text{ \AA}$) for all metallo-ferritin nodes. These observations indicate that synthetic linkers can also modulate the unit cell dimensions of ferritin-MOFs.

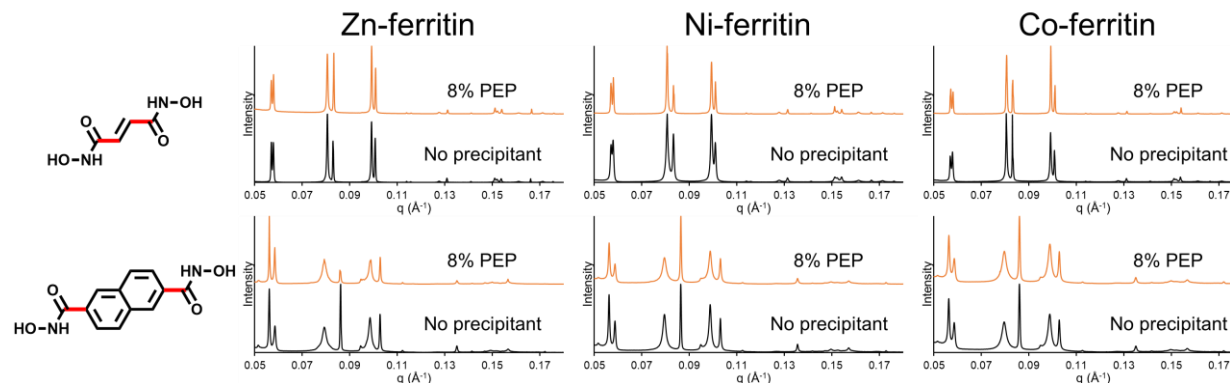


Figure 3.12. Diffraction patterns of *E*-H₂edh- and 2,6-H₂ndh-based ferritin-MOFs. Crystals formed under precipitant-free conditions are shown in black. Crystals formed with PEP are shown in orange. The lattices are all isostructural.

Much like the parallel-offset ligands, the *m*-H₂bdh- and 2,5-H₂tdh-based ferritin-MOFs all form isostructural *bct* lattices regardless of metal (Figure 3.13). The less symmetric primitive cell (shown in *m*-bdh-Zn- and 2,5-fdh-Zn-ferritin) has been observed sporadically throughout all of the *bct* lattices with bent ligands, indicating that there is little energetic penalty for a slight reduction in lattice symmetry. It is not clear if all of the crystals form this larger primitive lattice or if this lattice forms in addition to the *bct* lattice, as all of the reflections in the *bct* lattice are also present in the primitive lattice. In any case, the overall ferritin arrangement of the two is virtually identical. It is difficult to determine if a similar primitive lattice is also present in 2,5-tdh-Ni-ferritin or 2,5-pdh-Zn-ferritin since some of the expected reflections are absent. Since the diffraction pattern for 2,5-pdh-Ni-ferritin can be assigned to the “alternate” *bct* symmetry ($a = 155.89 \text{ \AA}$ $a = 148.73 \text{ \AA}$), it is likely that both 2,5-tdh-Ni-ferritin or 2,5-pdh-Zn-ferritin are comprised of two discrete crystalline species. 2,5-fdh-Zn follows the established trend by crystallizing into either an isostructural *bct* lattice or the larger primitive lattice. Unexpectedly, both 2,5-fdh-Ni- and 2,5-fdh-Co-ferritin form highly symmetric frameworks with *bcc* symmetry, reminiscent of *p*-bdh-Zn and *p*-bdh-Co-ferritin. The unit cell parameters of 2,5-fdh-Ni-ferritin are nearly identical to the previously characterized *p*-bdh-Zn-ferritin lattice ($a = 155.97 \text{ \AA}$) despite the bent geometry and a shorter expected metal-metal distance. The lattice dimensions of 2,5-fdh-Co-ferritin crystallized with PEP ($a = 154.70 \text{ \AA}$) are much closer to the predicted values based on the length of the ligand but deviates from the

2,5-fdh-Co-ferritin in crystallized in precipitant-free conditions ($a = 155.96 \text{ \AA}$). It is currently not known why the furan ligand deviates from the predicted *bct* packing seen in the isostructural thiophene ligand. The unique packing of 2,5-fdh-Ni allows for unusual dynamic behavior, which will be further explored in the next chapter. Crystals for 2,6-podh-based ferritin-MOFs could not be obtained in high enough yields to allow for unit cell determination.

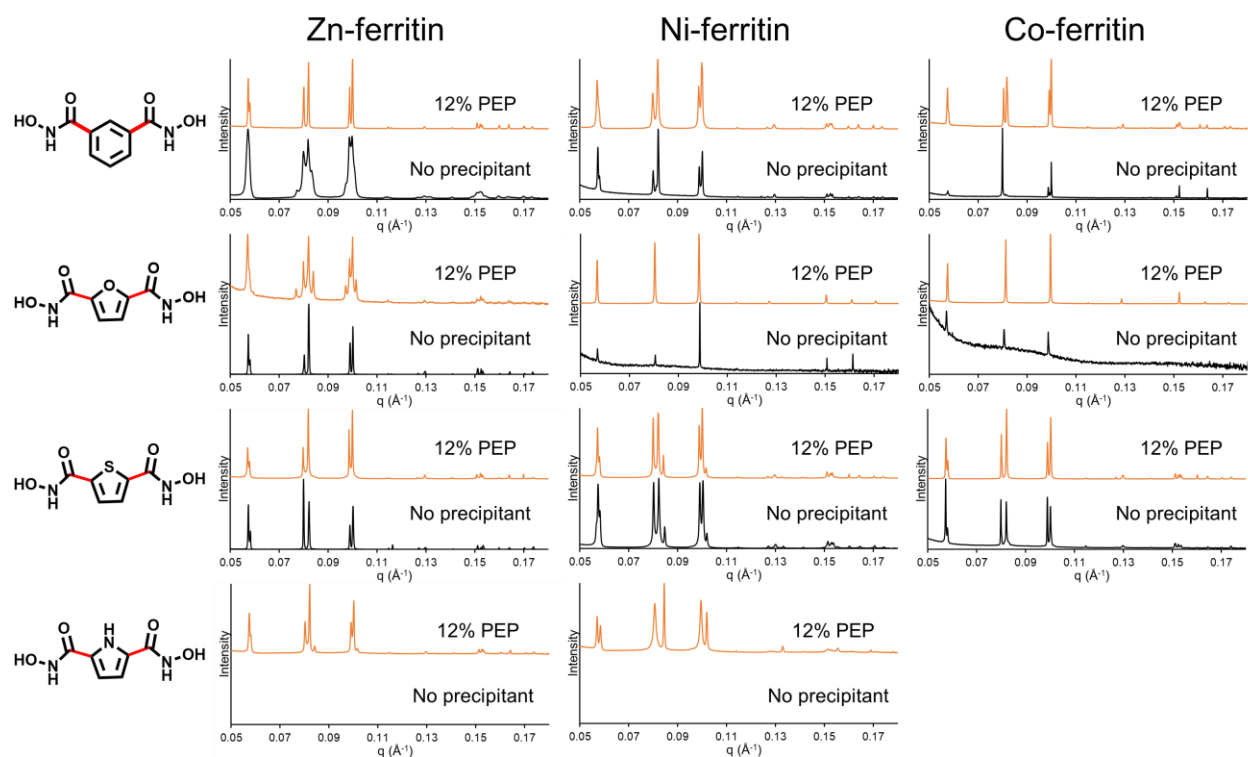


Figure 3.13. Diffraction patterns of *m*-H₂bdh-, 2,5-H₂fdh-, 2,5-H₂tdh-, and 2,5-H₂pdh-based ferritin-MOFs. Crystals formed under precipitant-free conditions are shown in black. Crystals formed with PEP are shown in orange. Crystals of 2,5-H₂pdh-(Zn/Ni)-ferritin could not be obtained in precipitant-free conditions and crystals of 2,5-H₂pdh-Co-ferritin could not be obtained under any conditions.

The *p*-xdh-ferritin adducts can form both *bcc* and *bct* lattices. *p*-xdh-Zn-ferritin is one of the rare cases where both a *bct* and a *bcc* lattice can be observed simultaneously in a near equal ratio. The ability to form both lattices is likely due to the flexibility of the xylene spacer, which has rotatable methylene carbons. The lack of bias for either symmetry indicates that there is very little energetic bias towards either crystal packing arrangement. The *bct* lattice of *p*-xdh-Zn-ferritin has the longest unit cell side ($c = 162.27 \text{ \AA}$) of any of the ferritin-MOFs and has a tetragonal distortion

of ($|a-c| = 7.60 \text{ \AA}$). Only the naphthalene-based 2,6- H_2ndh ligand gives a longer tetragonal distortion. Despite this, the cubic lattice ($a = 155.24 \text{ \AA}$) is smaller than p -bdh-Zn-ferritin. As with many of the examples, p -xdh-Co-ferritin forms very similar lattices to the Zn adduct. The secondary p -xdh-Co-ferritin bct lattice with ($a = 152.85 \text{ \AA}$, $c = 160.19 \text{ \AA}$) is isostructural with p -xdh-Zn-ferritin and has a very similar tetragonal distortion ($|a-c| = 7.34 \text{ \AA}$). The main tetragonal lattice ($a = 154.44 \text{ \AA}$, $c = 156.05 \text{ \AA}$) has the smallest tetragonal distortion ($|a-c| = 1.6 \text{ \AA}$) of any of the lattices, and is very similar to the bcc p -bdh-Zn-ferritin lattice. p -xdh-Ni ferritin forms exclusively a smaller bcc lattice ($a = 155.16 \text{ \AA}$). This bcc lattice becomes the predominant species for all three metal adducts formed in the presence of PEP. Similar results were seen for other precipitants, indicating that specific PEP interactions do not facilitate the increase in lattice symmetry. Ligand-metal combinations that are capable of crystallizing in both bcc and bct symmetries tend to prefer higher-symmetry packing under conditions where the protein is less soluble.

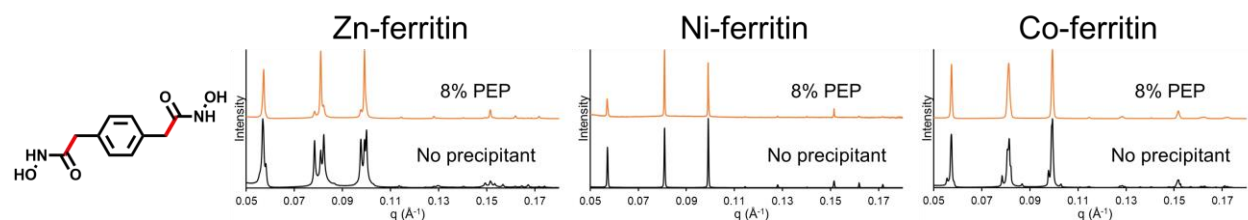


Figure 3.14. p - H_2xdh -based ferritin-MOF diffraction patterns crystallized under precipitant-free (black) or with PEP (orange).

3.2.6. Structural determination of ferritin-MOFs through sc-XRD

To understand the assembly of ferritin-MOFs in detail and to examine the relationship between lattice symmetry and metal/ligand geometry, we pursued single-crystal X-ray diffraction (sc-XRD) experiments. Earlier screening efforts led to single crystals of *p*-bdh-Zn-ferritin, which diffracted to a resolution of 3.8 Å best, with the great majority diffracting to >6.0 Å.⁴⁰ The wide screen described earlier in this chapter enabled us to obtain diffraction-quality single crystals that consistently diffracted to a resolution of <2.75 Å. This dramatic improvement is attributed to three factors. First, increasing the pH from 8.5 to 9.5 increased the solubility of the bridging ligand and allowed for a greater molar excess of ligand to be used. Second, the use of precipitants enabled the growth of large single crystals (ca. 200 μm sides). Finally, improvements in cryoprotection techniques (described in detail in the methods) prevented lattice decomposition during freezing. These improvements directly led to the determination of the structures of six ferritin-MOFs. The structural determination of *p*-bdh-Zn-ferritin (improved from 3.8 Å to 2.63 Å resolution, *I*432, *a* = 155.80 Å; PDB ID: 5UP8), *p*-xdh-Zn-ferritin (2.45 Å resolution, *I*4, *a* = 149.88 Å, *c* = 162.23 Å; PDB ID: 5UP9), and 2,5-fdh-Ni-ferritin (2.28 Å resolution, *I*432, *a* = 154.90 Å) was relatively straightforward, as these unit cell dimensions closely matched the previously determined ones. However, the unit cell determination and structural determination of the other three ferritin-MOFs was challenging. After significant effort, it was determined that these ferritin-MOFs all unexpectedly crystallize into isostructural “pseudo-tetragonal” monoclinic lattices. This allowed for the determination of the structure of *m*-bdh-Zn-ferritin (2.07 Å resolution, *P*₂₁, *a* = 215.34 Å, *b* = 215.61 Å, *c* = 156.35 Å; β = 89.974°), *m*-bdh-Ni-ferritin (2.34 Å resolution, *P*₂₁, *a* = 213.55 Å, *b* = 213.74 Å, *c* = 155.93 Å; β = 90.048°), and 2,5-fdh-Zn-ferritin (2.38 Å resolution, *P*₂₁, *a* = 213.99 Å, *b* = 213.91 Å, *c* = 156.33 Å; β = 89.864°). All of these structures gave insight into the intramolecular metal-ligand interactions that connect the ferritin nodes into ordered lattices.

In the *bcc* crystals of *p*-bdh-Zn-ferritin, the C_3 symmetry axes of the individual ferritin molecules are perfectly aligned with the crystallographic 3-fold symmetry axes that form the body diagonals in the unit cells (Figure 3.15 a). Consequently, the length of each body diagonal (269.9 Å) is the sum of the diameters of two ferritin molecules (i.e., 2×123.4 Å, as measured by the distance between two Zn ions along the C_3 axis within each ferritin molecule) and twice the separation between the Zn ions across the ligand-mediated interface (i.e., 2×11.5 Å). Owing to the cubic symmetry, each *p*-bdh molecule is centered precisely at the intersection of the crystallographic 2- and 3-fold symmetry axes. This positioning, combined with the rotational averaging of *p*-bdh about the 3-fold axis, gives rise to a diffuse electron density for the linker, even at the considerably improved resolution limit (Figure 3.15 a).

These complications are eliminated in the case of the *bct* lattice of *p*-xdh-Zn-ferritin, which lacks crystallographic 3-fold symmetry. The symmetry gives rise to a single unique orientation of *p*-xdh between the two crystallographically non-equivalent Zn centers (Figure 3.15 b). The resultant, well-defined electron density allowed us to build an unambiguous structural model of *p*-xdh and the Zn coordination environments. This model confirms that the hydroxamate head groups bind bidentate to the Zn centers in an apical position. The ligand completes a pseudo-tetrahedral Zn coordination geometry, with both Fe-O coordination distances at 2.2 Å. The xylene moiety of the linker assumes a nearly parallel orientation with respect to the ferritin-ferritin interfacial plane, giving rise to a lateral displacement (ca. 7.5 Å) of the Zn ions from the body diagonal of the *bct* unit cell (Figure 3.15 b). Due to the horizontal orientation of the linker, the interfacial spacing between the ferritin molecules (as defined by the projected Zn-Zn distance) is reduced to 9.5 Å from 11.5 Å observed in the cubic *p*-bdh-Zn-ferritin lattices (Figure 3.15 b). This compaction brings a small number of interfacial side chains into a distance range (3.8-5 Å), where electrostatic interactions between ferritin molecules may be invoked (Figure 3.16). It must be noted that the unit cell parameters for the tetragonal sublattice of *p*-xdh-Zn-ferritin determined by

room-temperature SAXS experiments ($a = 152.67 \text{ \AA}$, $c = 160.27 \text{ \AA}$) indicate a smaller tetragonal distortion compared to the sc-XRD case ($a = 149.88 \text{ \AA}$, $c = 162.23 \text{ \AA}$). This points to a reduced lateral displacement and a ca. 1.1-\AA increase in the interfacial separation (projected Zn-Zn estimated to be ca. 10.5 \AA), which could be readily accommodated by the tilting of the p-bdh ligand with respect to the interfacial plane and bond rotations about the methylene spacers (Figure 3.17). At this interfacial separation, there would be substantially less direct protein-protein contacts, suggesting that the crystallographically observed interprotein contacts are a consequence of the cryogenic temperatures used in sc-XRD experiments.

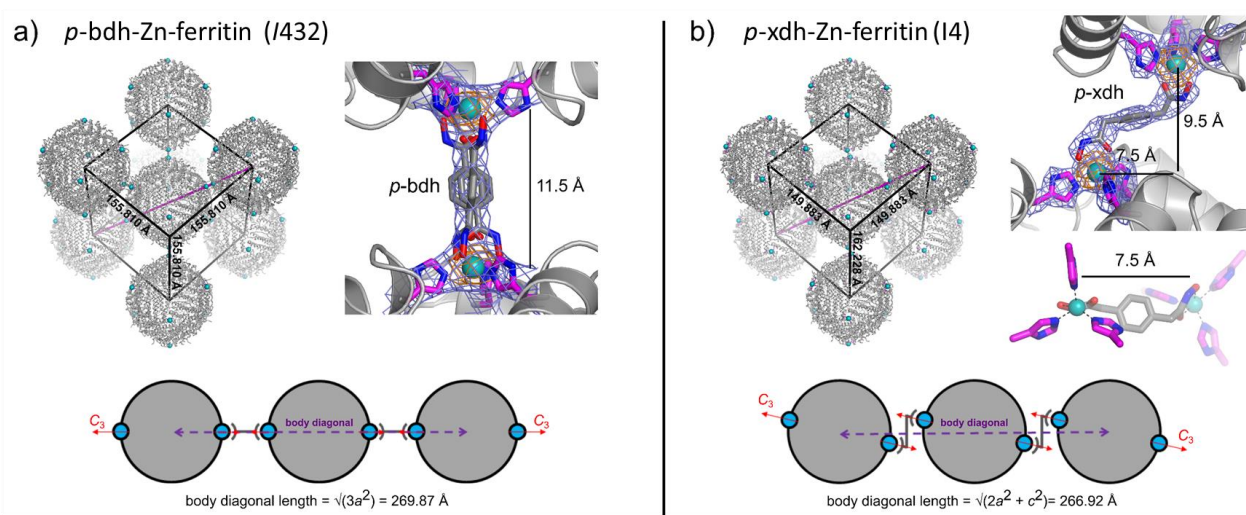


Figure 3.15. sc-XRD structures for a) *p*-bdh-Zn-ferritin and b) *p*-xdh-Zn-ferritin. Zinc atoms are highlighted as teal spheres. $2F_o - F_c$ electron density maps are contoured at 1σ (blue) and 3σ (orange). The body diagonal of the unit cells is highlighted with a purple line.

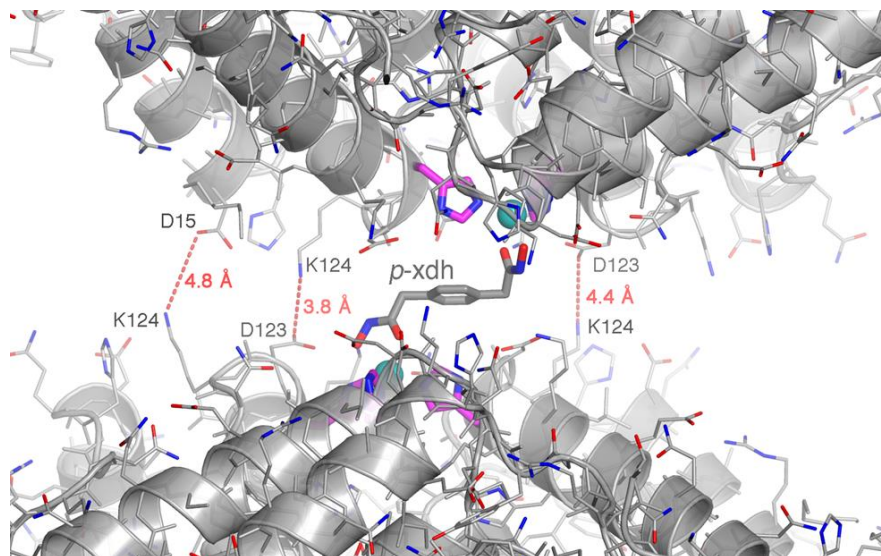


Figure 3.16. A close-up view of the protein-protein interface of *p*-xdh-Zn-ferritin showing pairs of amino acid residues that may be involved in electrostatic interactions.

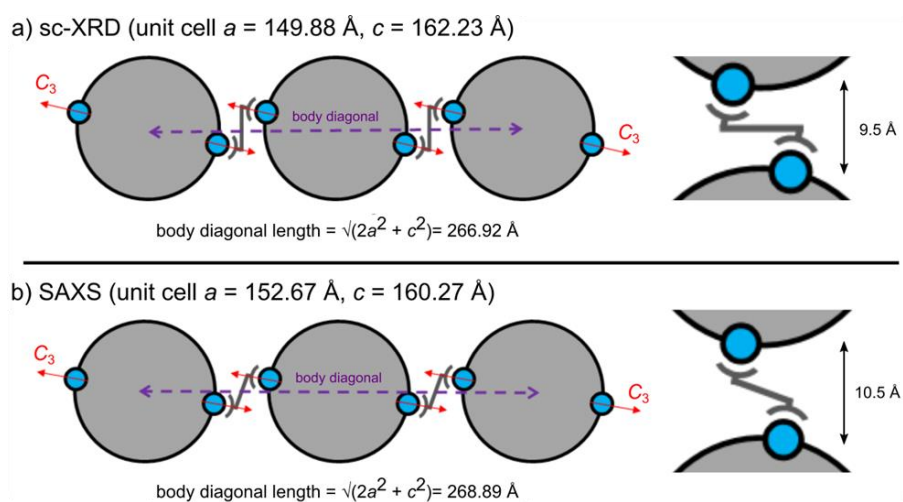


Figure 3.17. A schematic representation of the *p*-xdh-Zn-ferritin lattices examined by sc-XRD and SAXS, as viewed along a body diagonal of the unit cell. Ferritin molecules are shown as gray circles, the Zn ions that lie near the body diagonal are shown as blue circles, and *p*-xdh is depicted as dark gray sticks.

As eluded to earlier, the structural determination of *m*-bdh-Zn-, *m*-bdh-Ni-, and 2,5-fdh-Zn-ferritin was considerably more complicated. Unlike the previously discussed ferritin-MOFs, these structures did not crystallize into the expected body-centered space groups. The apparent symmetry of *m*-bdh-Zn-, *m*-bdh-Ni-, and 2,5-fdh-Zn-ferritin was $P4_212$ with approximate dimensions of $a = 214 \text{ \AA}$ and $c = 156 \text{ \AA}$. The formation of a primitive lattice was previously discussed as a possible explanation for the additional peaks seen in some of the diffraction

patterns of the *bct* ferritin-MOFs. A preliminary model in this space group highlights the structural abnormality behind the apparent $P4_21_2$ symmetry. The apparent primitive lattice is shown as a green square in Figure 3.18 a. The ferritin nodes at the vertices and in the center of the 110 plane (shown in gray) are crystallographically resolved to a single discrete location. Each of these ferritin nodes connects to eight adjacent ferritin nodes through the C_3 axes, as expected for a ligand-mediated ferritin-MOF. However, due to the bent geometry of the ligand, these ferritin nodes (overlaid as red and blue) can occupy multiple equally-favorable conformations. For each of these crystal structures, we were able to resolve two discrete locations of the disordered ferritin node, though there may be alternate conformations with lower occupancies. This primitive unit cell is comprised of two pseudo-*bct* lattices with approximate dimensions $a = 152 \text{ \AA}$, $c = 156 \text{ \AA}$, consistent with the unit cell dimensions that were previously determined. Since the geometry of the disordered ferritin nodes (red and blue) is dependent on the ligand orientation at the interface, there is no long-range patterning of these nodes (Figure 3.18 b). This interface is highlighted in Figure 3.18 c, with both overlaid conformations shown in red and blue.

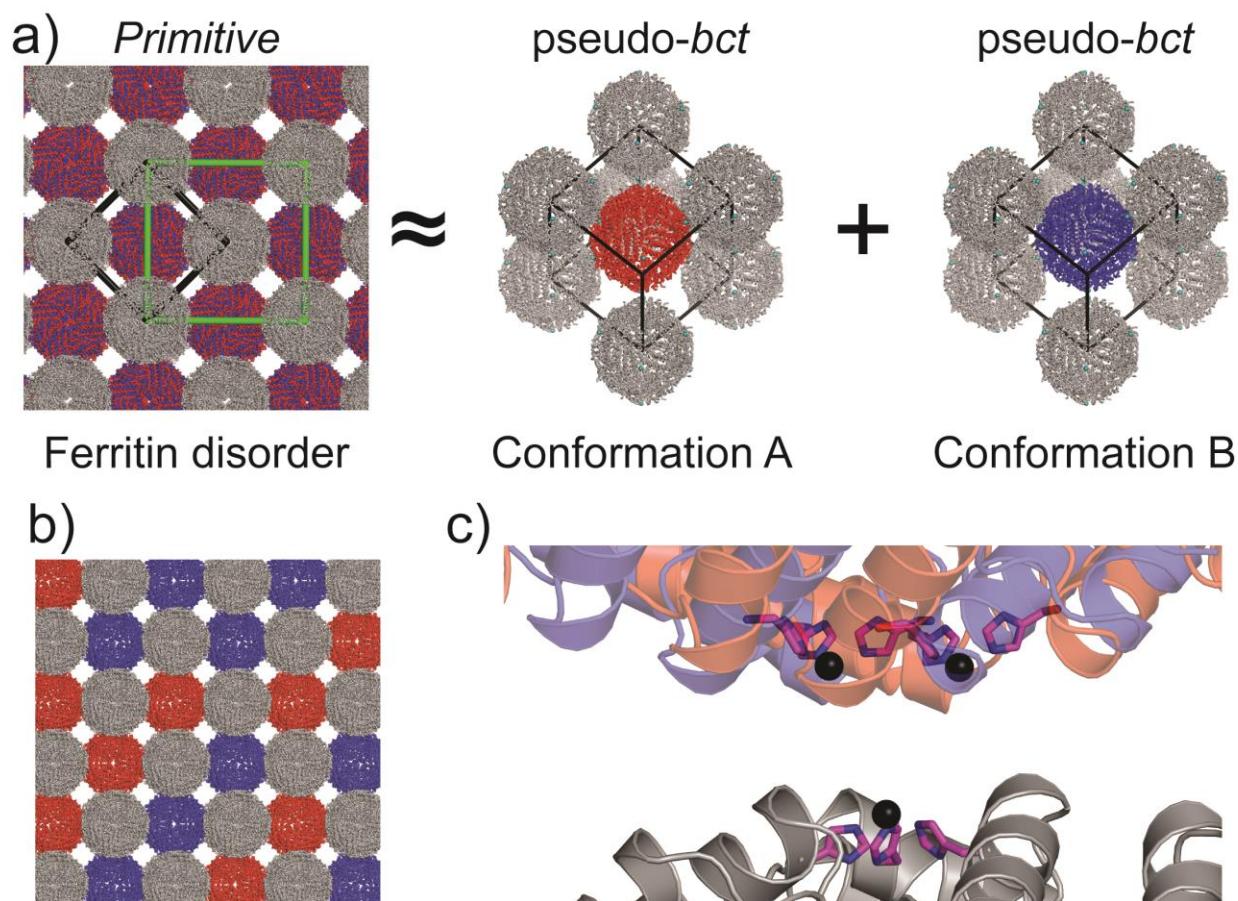


Figure 3.18. Multiple crystallographic conformations in ferritin-MOFs. a) A representation of the crystal lattice for *m*-bdh-Zn-, *m*-bdh-Ni-2,5-fdh-Zn-ferritin is shown. Fully ordered ferritin nodes are shown in gray, and the two alternate conformations of the disordered nodes are shown in red and blue, respectively. The apparent $P4_21_2$ lattice is shown in green, and the underlying pseudo-*bct* lattice is shown in black. b) Hypothetical representation of the patterning shows that there is no long-range repeating unit cell between the red and blue positions. c) The ferritin-ferritin interface is shown with ^{122}His (magenta) and the bound metal ion (black). The alternate ferritin conformations are shown in red and blue.

Though the unit cell parameters are nearly correct, the apparent $P4_21_2$ symmetry does not adequately explain the two discrete ferritin positions. To fully resolve each disordered ferritin position, the symmetry of the lattice needed to be lowered to the monoclinic $P2_1$ space group. Further lowering the symmetry to $P1$ did not improve the structural model of any of the ferritin-MOFs. There is very little deviation between the a and b parameters and β is nearly 90° for *m*-bdh-Zn-ferritin ($a = 215.34 \text{ \AA}$, $b = 215.61 \text{ \AA}$, $c = 156.35 \text{ \AA}$; $\beta = 89.974^\circ$), *m*-bdh-Ni-ferritin ($a = 213.55 \text{ \AA}$, $b = 213.74 \text{ \AA}$, $c = 155.93 \text{ \AA}$; $\beta = 90.048^\circ$), and 2,5-fdh-Zn-ferritin ($a = 213.99 \text{ \AA}$, $b = 213.91 \text{ \AA}$, $c = 156.33 \text{ \AA}$; $\beta = 89.864^\circ$). While we were able to obtain reasonable solutions from the

$P2_1$ space group for all of the ligands, the diffraction pattern might be better explained through a supercell containing more crystallographically unique ferritin nodes. Despite the ferritin node disorder and the extensive non-crystallographic symmetry in the $P2_1$ lattice, we were still able to model the bridging ligands between the ferritin subunits reasonably.

The *m*-bdh-Zn-ferritin has an underlying pseudo-*bct* lattice with $a \approx 152.4$ and $c \approx 156.4$. These values are shorter than the measured values by SAXS ($a = 153.53 \text{ \AA}$ $c = 157.26 \text{ \AA}$). The slightly shorter unit cell dimensions are likely due to the thermal contraction of the cell at 100K, a phenomenon that is observed for most ferritin-MOFs at cryogenic temperatures. The *m*-bdh bridging ligand coordinates bidentate as expected in an apical position, giving a pseudo-tetrahedral binding geometry (Figure 3.19). The Zn-O distances for all of the crystallographically unique sites fall within the expected range (1.9-2.3 \AA). The interfacial spacing between the ferritin nodes is ca. 9.0 \AA , which is less than both p-bdh-Zn-ferritin (11.5 \AA) and p-xdh-Zn-ferritin (9.5 \AA). The lateral displacement of the adjacent nodes is ca. 4.3 \AA from the C_3 axis (shown as a red dashed line). This displacement, combined with the shorter internode spacing, brings ^{123}Asp from one node and ^{124}Lys from the adjacent node close (2.8 \AA). These residues likely form an intermolecular hydrogen bond that helps stabilize the crystal packing (Figure 3.20). This interaction is also observed in *m*-bdh-Ni- and 2,5-fdh-Zn-ferritin. This specific hydrogen-bonding interaction may have biased the disordered ferritin nodes to crystallize into two distinct conformations.

The *m*-bdh-Ni-ferritin pseudo-*bct* lattice is slightly smaller than *m*-bdh-Zn-ferritin, with unit cell parameters $a \approx 151.1$ and $c \approx 155.9$. A close inspection of the 3-fold metal coordination site reveals that the hydroxamate head groups of *m*-bdh coordinate in an octahedral environment. The coordination geometry leads to a slightly shorter interfacial spacing between the ferritin nodes (8.1 \AA), with a very similar lateral displacement (4.2 \AA). The difference in coordination geometry is likely due to the preference for Ni(II) to form octahedral coordination environments. The Ni-O distances for all of the crystallographically unique sites fall within the expected range (2.0-2.2 \AA).

Much like the m-bdh-Zn-ferritin lattice, only a single unique m-bdh position could be resolved in m-bdh-Ni. The 2,5-fdh-Zn lattice closely mirrors the m-bdh-Ni lattice with unit cell parameters $a \approx 151.3$ and $c \approx 155.3$. The ligand density is much more diffuse at the node-node interface and was assigned to four total ligand conformations (two per disordered ferritin). The coordination of the hydroxamate head groups appears to be in the apical position for all of the modeled ligands. However, the diffuse density at the interface makes this difficult to assign definitively. Both the interfacial spacing (8.2 Å) and lateral displacement (4.3 Å) are very similar to the previous two structures.

Out of all of the bent ligands that have been characterized, only 2,5-fdh-(Ni/Co)-ferritin deviate from isostructural *bct* or pseudo-*bct* symmetry. The *bcc* crystals of 2,5-fdh-Ni closely resemble the previously discussed *p*-bdh-Zn crystals; the C_3 symmetry axes of the ferritin nodes are perfectly aligned with the crystallographic 3-fold symmetry axes. Due to the similar lattice dimensions at room temperature determined through powder diffraction, it was unclear if 2,5-fdh-Ni-ferritin would have similar interfacial metal-metal spacing to *p*-bdh-Zn ferritin, even though 2,5-H₂fdh is ca. 1.5 Å shorter than *p*-H₂bdh. As anticipated by the shorter ligand length, the crystal lattice confirms that the interfacial metal-metal distance is shorter in 2,5-fdh-Ni-ferritin (10.1 Å) when compared to *p*-bdh-Zn-ferritin (11.5 Å). The lattice dimensions of 2,5-fdh-Ni-ferritin are also much shorter than those measured at room temperature. The rotational averaging of 2,5-fdh about the 3-fold axis gives rise to a column of density, much like in *p*-bdh-Zn-ferritin. Unfortunately, this complicates the assignment of the hydroxamate-metal geometry, as 2,5-H₂fdh could achieve this coordination with pseudo-tetrahedral or octahedral coordination. Based on the node spacing, and the relatively narrow column of electron density, we believe that it is more likely that the ligand coordinates in a pseudo-tetrahedral geometry, despite the coordination preferences of Ni(II). Additional experiments need to be done to determine if the coordination environment is pseudo-tetrahedral or octahedral.

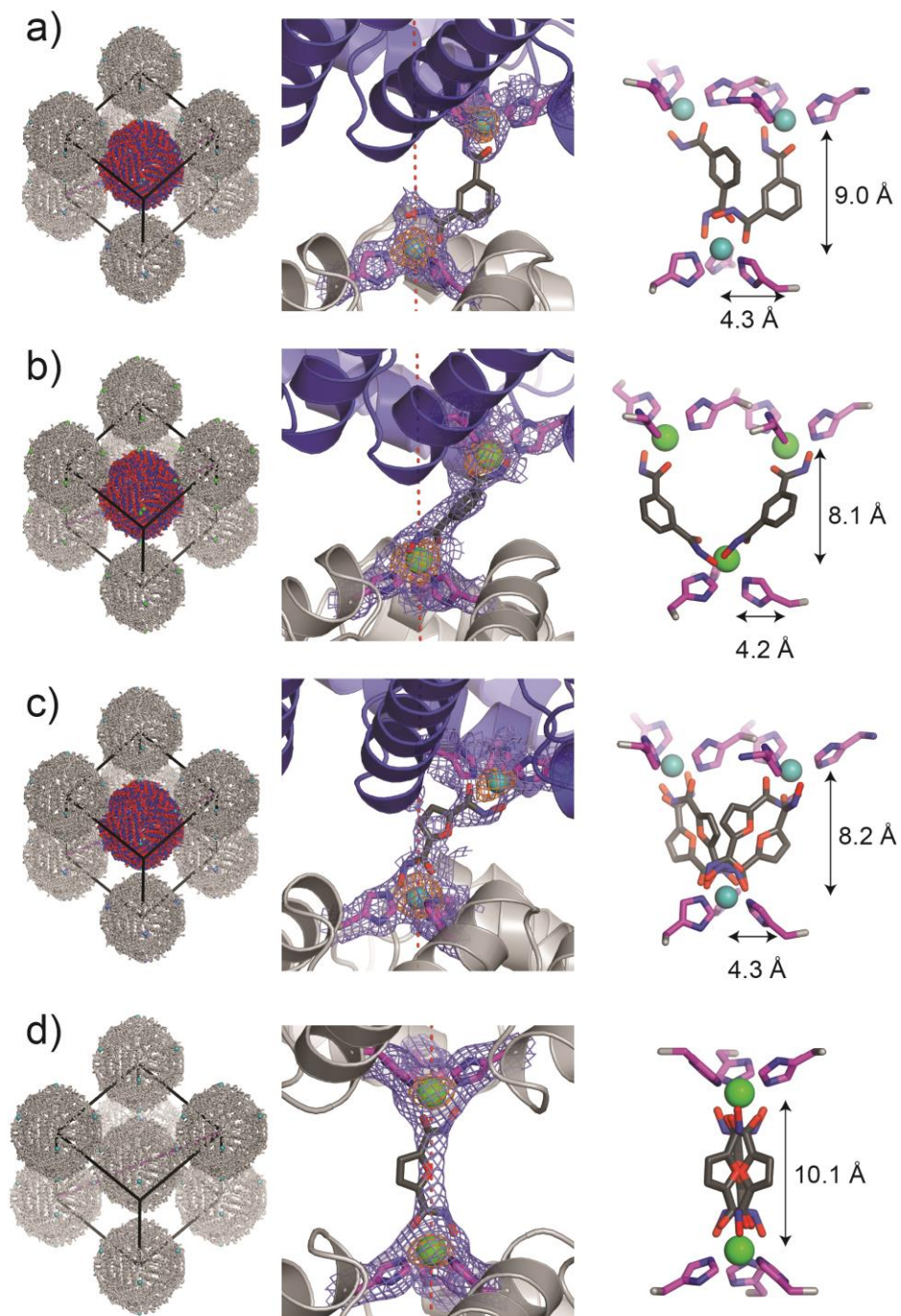


Figure 3.19. Crystal structures of *m*-bdh-(Zn/Ni)- and 2,5-fdh-(Zn/Ni)-ferritin. The pseudo-*bct* or *bcc* unit cells corresponding to *m*-bdh-Zn-ferritin (a), *m*-bdh-Ni-ferritin (b), 2,5-fdh-Zn-ferritin (c) and 2,5-fdh-Ni-ferritin (d) are shown on the right. The node-node interface with a single orientation of the bridging ligand is shown in the middle. $2F_o - F_c$ electron density maps are contoured at 1σ (blue) and 3σ (orange). The plane perpendicular to the C_3 axis is shown as a red dashed line. The alternative conformations of the ligand and node are shown on the right with the internode spacing and lateral displacement. ^{122}His sidechains are shown as magenta sticks, Zn(II), and Ni(II) are shown as teal and green spheres, respectively.

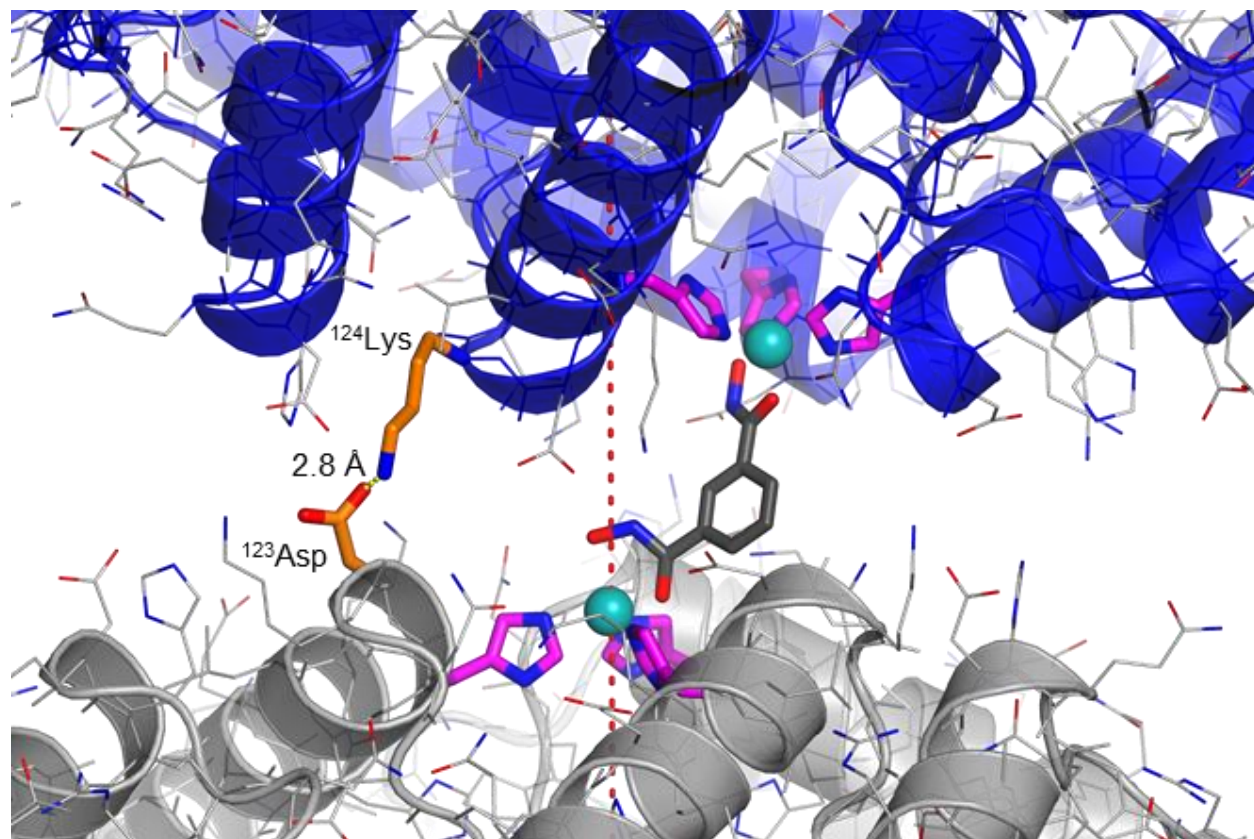


Figure 3.20. A close-up view of the protein-protein of a single conformation of *m*-bdh-Zn-ferritin. ¹²⁴Lys and ¹²³Asp (highlighted as orange sticks) make a highly specific hydrogen bond that helps stabilize this conformation. The plane perpendicular to the C_3 axis is shown as a red dashed line. ¹²²His sidechains are shown as magenta sticks, Zn(II) are shown as teal spheres. A similar hydrogen-bonding interaction is seen in *m*-bdh-Ni-ferritin and 2,5-fdh-Zn-ferritin.

Here we have reported the construction and characterization of an extensive library of 3D, crystalline protein-MOFs by a combination of three different metallo-ferritin nodes and five synthetic linkers bearing hydroxamate head groups. Our results establish that the metal ion and the synthetic linker components can be varied in a modular fashion to influence the structural parameters (i.e., lattice symmetries and unit cell dimensions) of protein-MOFs, akin to conventional MOFs. Despite the remarkable size discrepancy between the ferritin nodes and the organic linkers, the self-assembly of the ferritin-MOFs is highly robust, emphasizing the utility of metal coordination interactions in controlling protein self-assembly. Excitingly, by virtue of the flexibility of the linkers and the fluxionality of metal-linker bonds, several ferritin-MOFs have been observed to adopt multiple conformations, which suggests that they may display dynamic

behavior. The next chapter will investigate the emergent functional/physical properties of protein-MOFs (e.g., their potential dynamic nature) arising from their unique, tripartite composition.

3.3. Methods

3.3.1. General considerations

All chemical reagents were purchased from Sigma-Aldrich, Fisher Scientific, TCI, or an equivalent supplier. Mass spectrometry (MS) of proteins and small molecules was carried out at the Molecular Mass Spectrometry Facility at UCSD using electrospray ionization (ESI) on a Micromass Quattro Ultima Triple Quadrupole MS. NMR spectra (^1H) were recorded on a Varian Mercury spectrometer (400 MHz). NMR data are reported as follows: chemical shift, multiplicity (s = singlet, br = broad), and integration. The spectra were internally referenced to the residual solvent signal (DMSO: δ 2.50 ppm; CDCl_3 : δ 7.26 ppm). ^{122}Hf ferritin was expressed and purified, as described in Chapter 2.

3.3.2. Structural determination of Ni-ferritin and Co-ferritin nodes through sc-XRD

Crystals of Ni-ferritin and Co-ferritin in the *F*432 space group were obtained using sitting-drop vapor diffusion at 25 °C. The crystallization/solution conditions are listed in Table 3.2. Octahedron-shaped crystals appeared overnight and matured over several days. Crystals from the condition labeled Ni-ferritin + DMSO were soaked for 24 h in a solution containing 25 mM Tris (pH 8.0), 150 mM NaCl, 5 mM CaCl_2 , 10 mM catechol, and 10% (v/v) DMSO. Crystals suitable for sc-XRD experiments were harvested and soaked briefly in perfluoropolyether before rapid freezing in liquid N_2 (for Ni-ferritin) or a liquid nitrogen cryostream (Co-ferritin). The sc-XRD data for Ni-ferritin was collected at 100 K at Beamline 12-2 of the Stanford Synchrotron Radiation Laboratory (SSRL) using 0.98 Å radiation. Co-ferritin data were collected on a Bruker Microstar diffractometer equipped with Cu K_α radiation (1.5418 Å) at 100 K. The data were integrated using either iMosflm⁴⁸ (Ni-ferritin) or SAINT (Co-ferritin) and scaled using either Aimless⁴⁹ (Ni-ferritin) or Sadabs⁵⁰ (Co-ferritin). The structures of Ni-ferritin and Co-ferritin and Ni-ferritin + DMSO were solved to resolutions of 1.79 Å, 1.95 Å, and resolution, respectively. Molecular replacement was performed with Phaser⁵¹ using the previously solved ^{122}Hf ferritin structure (PDB ID: 5CMQ) as the

search model. Rigid-body, positional, thermal, and TLS refinements were carried out using Phenix.⁵² The data refinement statistics are shown in Table 3.3.

Table 3.2. Crystallization conditions for Ni-ferritin and Co-Ferritin

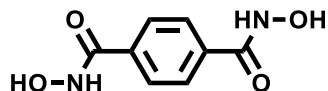
Ni-ferritin	Reservoir	500 μ L total volume: 25 mM Tris (pH 8.0), 10 mM CaCl ₂ , 1 mM NiCl ₂ , 1% PEG 6000
	Sitting Drop	2 μ L reservoir, 2 μ L of 25 μ M ferritin
Co-ferritin	Reservoir	500 μ L total volume: 25 mM Tris (pH 8.0), 12 mM CaCl ₂ , 150 mM NaCl, 0.3 mM CoCl ₂ , 1% PEG 1900 MME
	Sitting Drop	2 μ L reservoir, 2 μ L of 8 μ M ferritin
Ni-ferritin + DMSO	Reservoir	500 μ L total volume: 25 mM Tris (pH 8.0), 10 mM CaCl ₂ , 0.1 mM NiCl ₂ , 1% PEG 1900 MME
	Sitting Drop	2 μ L reservoir, 2 μ L of 25 μ M ferritin

Table 3.3 X-ray data processing and refinement statistics for Ni- and Co-ferritin nodes.

	Ni-ferritin (5UP7)	Co-ferritin (5VTD)	Ni-ferritin + DMSO*
Data collection			
Space group	<i>F</i> 432	<i>F</i> 432	<i>F</i> 432
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	180.15	180.24	178.37
α , β , γ (°)	90	90	90
Resolution (Å)	63.69-1.79	63.76-1.95	50-1.70
No. unique reflections	24176	18852	27311
Multiplicity	10.3 (6.6)	8.1 (4.2)	15.4 (5.6)
<i>CC</i> 1/2	0.996 (0.634)	0.991 (0.797)	1.00 (0.622)
<i>R</i> merge	0.145 (0.791)	0.132 (0.607)	0.056 (0.918)
$\langle I / \sigma \rangle$	12.5 (1.9)	10.9 (2.0)	29.92 (1.58)
Completeness (%)	100 (99.9)	99.5 (98.7)	99.2 (97.0)
Refinement			
<i>R</i> _{work} / <i>R</i> _{free}	0.159/0.189	0.174/0.217	0.151/0.173
No. atoms			
Protein	1436	1415	1465
Ligand/ion	12	9	49
Water	327	268	180
<i>B</i> -factors (Å ²)			
Protein	10.9	14.5	24.57
Ligand/ion	21.3	22.5	36.21
Water	24.4	24.7	42.39
R.m.s. deviations			
Bond lengths (Å)	0.0056	0.0124	0.013
Bond angles (°)	0.765	1.164	1.31
MolProbity ⁵³ score	1.57	1.51	1.69
Clashscore	11.42	9.77	12.07
Ramachandran plot (%)			
Favored	98.24	98.82	97.67
Outliers	0.00	0.00	0.58
Rotamers (%)			
Poor	0.67	0.00	0.63

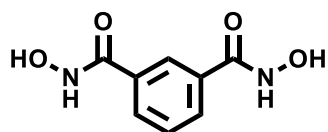
*Crystal structure is preliminary and has not been published. Refinement statistics will improve before deposition.

3.3.3. Synthesis of the dihydroxamate bridging ligand library



Synthesis of *N*¹,*N*⁴-dihydroxyterephthalamide (*p*-H₂bdh)

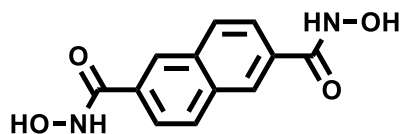
The synthesis for *p*-H₂bdh was carried out as previously reported with minor modifications. Hydroxylamine hydrochloride, 1.06 g (15.45 mmol) and sodium hydroxide, 1.24 g (30.9 mmol) were combined in 20 mL methanol and shaken vigorously. The solution was cooled on ice for 10 min, and NaCl was removed via filtration. The filtrate was combined with 1 g (5.15 mmol) dimethyl terephthalate in 30 mL methanol. The solution was stirred overnight at room temperature. The solvent was removed in vacuo, and the solid material was dissolved in 20 mL H₂O. The solution was acidified to a pH of 5.5 with 5% HCl. The white precipitate was collected by filtration and dried in vacuo. Yield = 65%. ¹H NMR: (400 MHz, DMSO-*d*₆) δ 11.35 (br s, 2H), δ 9.17 (br s, 2H), δ 7.80 (s, 4H). ¹³C NMR: (500 MHz, DMSO-*d*₆): δ 163.42, δ 135.04, δ 126.92. Measured molecular weight (m/z) = 196.97 (Calc.: 197.05) (M + H⁺).



Synthesis of *N*¹,*N*³-dihydroxyisophthalamide (*m*-H₂bdh)

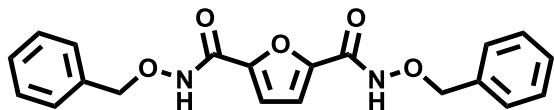
Hydroxylamine hydrochloride, 1.06 g (15.45 mmol) and sodium hydroxide, 1.24 g (30.9 mmol) were combined in 20 mL methanol and shaken vigorously. The solution was cooled on ice for 10 min, and NaCl was removed via filtration. The filtrate was combined with 1 g (5.15 mmol) dimethyl isophthalate in 30 mL methanol. The solution was stirred overnight at room temperature. The solvent was removed in vacuo, and the solid material was dissolved in 20 mL H₂O. The solution was acidified to a pH of 5.5 with 5% HCl. The white precipitate was collected by filtration

and dried in vacuo. Yield = 69%. ^1H NMR: (400 MHz, DMSO-d_6) δ 11.30 (br s, 2H), δ 9.14 (br s, 2H), δ 8.14 (s, 1H), δ 7.85 (dd, 2H), δ 7.53 (t, 1H). ^{13}C NMR (500 MHz, DMSO-d_6): δ 163.65, δ 133.07, δ 129.29, δ 128.54, δ 125.85. Measured molecular weight (m/z) = 197.05 (Calc.: 197.05) ($M + \text{H}^+$).



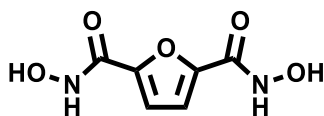
Synthesis of N^2,N^6 -dihydroxynaphthalene-2,6-dicarboxamide (H_2ndh).

Hydroxylamine hydrochloride 2.08 g (29.9 mmol) and sodium hydroxide, 2.4 g (60 mmol) were combined in 30 mL methanol and shaken vigorously. The solution was cooled on ice for 10 min, and NaCl was removed via filtration. The filtrate was combined with 1 g (4.09 mmol) dimethyl naphthalene-2,6-dicarboxylate in 125 mL CHCl_3 . The solution was stirred overnight at room temperature. The precipitate was collected by filtration and dissolved in H_2O . The solution was acidified with 1 M AcOH. The white precipitate was collected by filtration, washed with Et_2O , and dried in vacuo yielding 0.695 g of product (69% yield). ^1H NMR: (400 MHz, DMSO-d_6) δ 11.44 (br s, 2H), δ 9.19 (br s, 2H), δ 8.39 (s, 2H), δ 8.08 (d, 2H), δ 7.88 (d, 2H). ^{13}C NMR: (500 MHz, DMSO-d_6) δ 163.96, 133.27, 131.52, 129.04, 126.92, 124.48. Measured molecular weight (m/z) = 244.95 (Calc.: 245.06) ($M - \text{H}^+$).



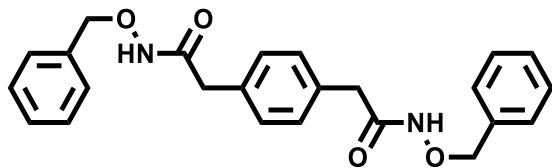
Synthesis of *N*²,*N*⁵-bis(benzyloxy)furan-2,5-dicarboxamide (2,5-Bn₂fdh)

Furan-2,5-dicarboxylic acid (500 mg, 3.24 mmol), HATU (2714 mg, 7.14 mmol), DIPEA (3.40 mL, 19.47 mmol) and *O*-benzylhydroxylamine hydrochloride (1243 mg, 7.79 mmol) were combined in 15 mL of DMF and stirred for 12 h. 50 mL of H₂O was added, and the reaction was extracted with 50 mL of EtOAc three times. The EtOAc layer was washed with 10% w/v citric acid, 10% w/v sodium bicarbonate, and brine. The EtOAc layer was then dried with anhydrous sodium sulfate, and the solvent was removed *in vacuo*. The product was then purified with a silica chromatography column with 0-90% EtOAc in hexanes eluent. Pure fractions were combined, and the solvent was removed *in vacuo* yielding 480 mg of product (41.2% yield). ¹H NMR: (400 MHz, DMSO-*d*₆) δ 11.79 (br s, 2H), δ 7.44 (m, 10H), δ 7.19 (s, 2H), δ 4.93 (s, 4H). Measured molecular weight (*m/z*) = 367.18 (Calc.: 367.12) (*M* + *H*⁺).



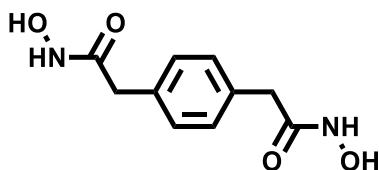
Synthesis of *N*²,*N*⁵-dihydroxyfuran-2,5-dicarboxamide (2,5-H₂fdh)

*N*²,*N*⁵-bis(benzyloxy)furan-2,5-dicarboxamide (100 mg, 0.273 mmol) and Pd/c (100 mg) were suspended in 50 mL of MeOH and placed under 1 atm H₂ while shaking for 2 h. The reaction was filtered through celite, and the solvent was removed *in vacuo*. The powder was resuspended in H₂O, and concentrated HCl was added until the pH was <1. The solution was cooled to 0 °C and the precipitate was collected with vacuum filtration and dried *in vacuo* yielding 50 mg of product (98% yield). ¹H NMR: (400 MHz, DMSO-*d*₆) δ 11.15 (s, 2H), δ 9.34 (s, 2H), δ 7.11 (s, 2H). Measured molecular weight (*m/z*) = 187.05 (Calc.: 187.03) (*M* + *H*⁺).



Synthesis of 2,2'-(1,4-phenylene)bis(*N*-(benzyloxy)acetamide) (*p*-Bn₂xdh)

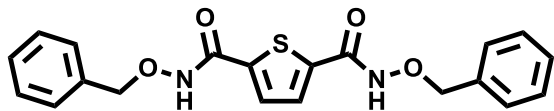
2,2'-(1,4-phenylene)diacetic acid (500 mg, 2.57 mmol), HATU (2154 mg, 5.66 mmol), DIPEA (2.70 mL, 15.45 mmol) and *O*-benzylhydroxylamine hydrochloride (986 mg, 6.18 mmol) were combined in 15 mL of DMF and stirred for 12 h. 50 mL of H₂O was added, and the reaction was extracted with 50 mL of EtOAc three times. The EtOAc layer was washed with 10% w/v citric acid, 10% w/v sodium bicarbonate, and brine. The EtOAc layer was then dried with anhydrous sodium sulfate, and the solvent was reduced *in vacuo*. The precipitate that formed was collected by vacuum filtration and dried *in vacuo* yielding 557 mg of product (54% yield). ¹H NMR: (400 MHz, DMSO-*d*₆) δ 11.30 (s, 2H), δ 7.35 (m, 10H), δ 7.16 (s, 4H), δ 4.77 (s, 4H), δ 3.27 (s, 4H). Measured molecular weight (*m/z*) = 405.24 (Calc.: 405.17) (*M* + *H*⁺).



Synthesis of 2,2'-(1,4-phenylene)bis(*N*-hydroxyacetamide) (*p*-H₂xdh)

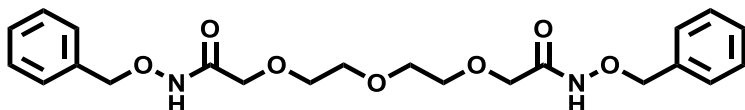
2,2'-(1,4-phenylene)bis(*N*-(benzyloxy)acetamide) (200 mg, 0.494 mmol) and Pd/c (40 mg) were suspended in 50 mL of MeOH and placed under 1 atm H₂ while shaking for 2 h. The reaction was filtered through celite, and the solvent was removed *in vacuo*. The powder was resuspended in H₂O, and concentrated HCl was added until the pH was <1. The solution was cooled to 0 °C and the precipitate was collected with vacuum filtration and dried *in vacuo* yielding 80 mg of

product (72% yield). ¹H NMR: (400 MHz, DMSO-d₆) δ 10.65 (s, 2H), δ 8.79 (s, 2H), δ 7.17 (s, 4H), δ 3.24 (s, 4H). Measured molecular weight (m/z) = 222.98 (Calc.: 223.22) (M - H⁺).



Synthesis of *N*²,*N*⁶-bis(benzyloxy)furan-2,5-dicarboxamide (2,5-Bn₂tdh)

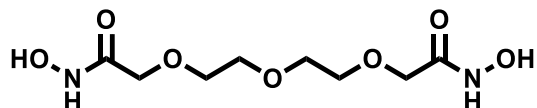
Thiophene-2,5-dicarboxylic acid (500 mg, 2.90 mmol), HATU (2429 mg, 6.36 mmol), DIPEA (3.04 mL, 17.43 mmol) and *O*-benzylhydroxylamine hydrochloride (1113 mg, 6.97 mmol) were combined in 15 mL of DMF and stirred for 12 h. 50 mL of H₂O was added, and the reaction was extracted with 50 mL of EtOAc three times. The EtOAc layer was washed with 10% w/v citric acid, 10% w/v sodium bicarbonate, and brine. The EtOAc layer was then dried with anhydrous sodium sulfate, and the solvent was reduced *in vacuo*. The precipitate that formed was collected by vacuum filtration and dried *in vacuo* yielding 537 mg of product (48% yield). ¹H NMR: (400 MHz, DMSO-d₆) δ 12.03 (br s, 2H), δ 7.60 (s, 2H), δ 7.40 (m, 10H), δ 4.92 (s, 4H). Measured molecular weight (m/z) = 383.08 (Calc.: 383.10) (M + H⁺).



Synthesis of 2,2'-((oxybis(ethane-2,1-diyl))bis(oxy))bis(*N*-(benzyloxy)acetamide) (Bn₂dgdh)

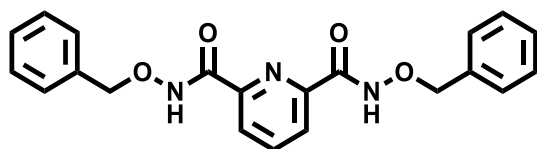
2,2'-((oxybis(ethane-2,1-diyl))bis(oxy))diacetic acid (500 mg, 2.25 mmol), HATU (1882 mg, 4.95 mmol), DIPEA (2.36 mL, 13.50 mmol) and *O*-benzylhydroxylamine hydrochloride (862 mg, 5.40 mmol) were combined in 15 mL of DMF and stirred for 12 h. 50 mL of H₂O was added and the reaction was extracted with 50 mL of EtOAc three times. The EtOAc layer was washed with 10% w/v citric acid, 10% w/v sodium bicarbonate, and brine. The EtOAc layer was then dried

with anhydrous sodium sulfate, and the solvent was removed *in vacuo*. The product was purified with silica column chromatography with 0-15% MeOH in EtOAc as the eluent. Pure fractions were combined, and the solvent was removed *in vacuo* yielding 460 mg of product (47% yield). ¹H NMR: (400 MHz, DMSO-d₆) δ 11.11 (s, 2H), δ 7.39 (m, 10H), δ 4.81 (s, 4H), δ 3.89 (s, 4H), δ 3.54 (s, 8H). Measured molecular weight (m/z) = 433.22 (Calc.: 433.19) (M + H⁺).



Synthesis of 2,2'-((oxybis(ethane-2,1-diyl))bis(oxy))bis(*N*-hydroxyacetamide) (H₂dgdh)

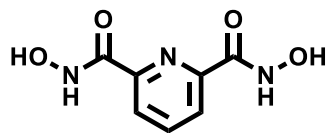
2,2'-((oxybis(ethane-2,1-diyl))bis(oxy))bis(*N*-(benzyloxy)acetamide) (460 mg, 1.064 mmol) and Pd/c (100 mg) were suspended in 50 mL of MeOH and placed under 1 atm H₂ while shaking for 2 h. The reaction was filtered through celite, and the solvent was removed *in vacuo* yielding 172 mg of product (73% yield). ¹H NMR: (400 MHz, DMSO-d₆) δ 10.49 (br s, 2H), δ 8.86 (br s, 2H), δ 3.85 (s, 4H), δ 3.54 (s, 4H), δ 3.54 (s, 8H).



Synthesis of *N*²,*N*⁶-bis(benzyloxy)pyridine-2,6-dicarboxamide (2,6-Bn₂pydh)

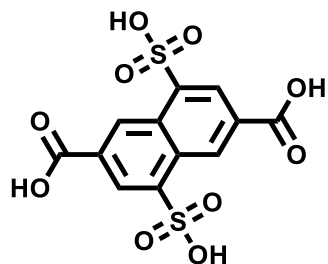
Pyridine-2,6-dicarboxylic acid (500 mg, 2.99 mmol), HATU (2697 mg, 7.09 mmol), DIPEA (3.38 mL, 19.34 mmol) and *O*-benzylhydroxylamine hydrochloride (1235 mg, 7.74 mmol) were combined in 15 mL of DMF and stirred for 12 h. 50 mL of H₂O was added, and the reaction was extracted with 50 mL of EtOAc three times. The EtOAc layer was washed with 10% w/v citric acid, 10% w/v sodium bicarbonate, and brine. The EtOAc layer was then dried with anhydrous sodium sulfate, and the solvent was removed *in vacuo*. The product was then purified with a silica

chromatography column with 0-90% EtOAc in hexanes eluent. Pure fractions were combined, and the solvent was removed *in vacuo* yielding 502 mg of product (41% yield). ¹H NMR: (400 MHz, DMSO-d₆) δ 12.37 (br s, 2H), δ 8.20 (m, 3H), δ 7.41(m, 10H), δ 4.99 (s, 4H).



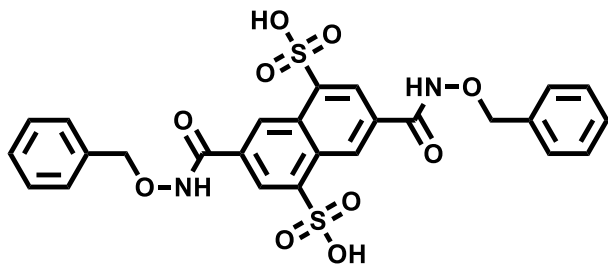
Synthesis of *N,N'*-dihydroxypyridine-2,6-dicarboxamide (2,6-Bn₂pydh)

N,N'-bis(benzyloxy)pyridine-2,6-dicarboxamide (502 mg, 1.33 mmol), and Pd/c (40 mg) were suspended in 50 mL of MeOH and placed under 1 atm H₂ while shaking for 2 h. The reaction was filtered through celite, and the solvent was removed *in vacuo*, yielding 205 mg of product (76% yield). ¹H NMR: (400 MHz, DMSO-d₆) δ 11.83 (br s, 2H), δ 9.29 (br s, 2H), δ 8.13 (m, 3H).



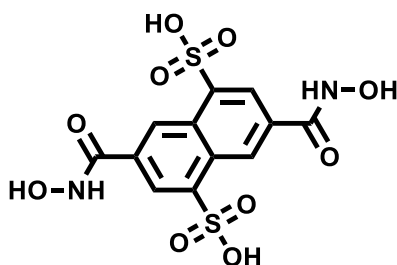
Synthesis of 4,8-disulfonaphthalene-2,6-dicarboxylic acid (2,6-sndc)

Naphthalene-2,6-dicarboxylic acid (1g, 4.63 mmol) was dissolved in 5 mL oleum and heated to 130 °C. The reaction was stirred overnight. 60 mL H₂O was added, and the solution was acidified with 8 mL of 12.1 M HCl. The precipitate that forms was collected by filtration and dried at 50 °C under vacuum yielding 0.915 g of product (52.6% yield). ¹H NMR: (400 MHz, DMSO-d₆) δ 9.60 (s, 2H), δ 8.49 (s, 2H). Measured molecular weight (m/z) = 376.36 (Calc.: 376.96) (M + H⁺).



Synthesis of 3,7-bis((benzyloxy)carbamoyl)naphthalene-1,5-disulfonic acid (2,6-Bn₂sndh)

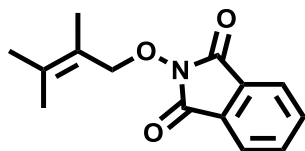
4,8-disulfonaphthalene-2,6-dicarboxylic acid (0.5 g 1.33 mmol) was combined with 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) (0.55 g, 2.925 mmol), Hydroxybenzotriazole (HOBt) (0.395 g, (2.925 mmol), and DIPEA (1.13 mL, 7.95 mmol) in 25 mL of DMF and stirred for 10 min. *O*-benzylhydroxylamine hydrochloride (0.5 g, 3.13 mmol) was added, and the reaction was stirred for 12 h. The solvent was reduced *in vacuo*, and H₂O was used to precipitate out the product, which was collected by filtration and dried *in vacuo* yielding 690 mg (79% yield). ¹H NMR: (400 MHz, DMSO-d₆) δ 11.35 (s, 2H, NH), δ 9.26 (s, 2H), δ 9.06 (s, 2H), δ 8.23 (s, 2H).



Synthesis of 3,7-bis(hydroxycarbamoyl)naphthalene-1,5-disulfonic acid (2,6-H₂sndh)

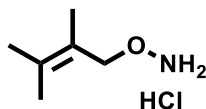
3,7-bis((benzyloxy)carbamoyl)naphthalene-1,5-disulfonic acid (300 mg, 0.511 mmol) was combined with 150 mg of Pd/C in 50 mL of MeOH and hydrogenated under 1 atm H₂ for 2h. The solvent was removed *in vacuo*, and the precipitate was dissolved in 30 mL H₂O. NaOH was added until the pH was ~12. The product was extracted with CHCl₃ (3x 30 mL). The resulting aqueous layer was separated, acidified with concentrated HCl, and the product was collected via filtration and dried *in vacuo* yielding 206 mg of product (89% yield). ¹H NMR: (400 MHz, DMSO-d₆) δ 12.00

(s, 2H), δ 9.33 (s, 2H), δ 8.24 (s, 2H), δ 7.45 (m, 5H), δ 4.96 (s, 4H). Measured molecular weight (m/z) = 404.89 (Calc.: 404.98) (M - H⁺).



Synthesis of 2-((2,3-dimethylbut-2-en-1-yl)oxy)isoindoline-1,3-dione

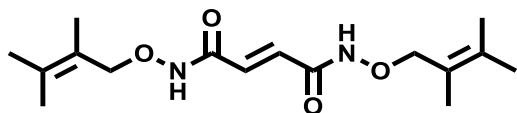
2,3-dimethyl-1,2-butene (9.6 mL, 81 mmol), NBS (14.2 g, 81 mmol), and BPO (250 mg, 1.03 mmol) were dissolved in 50 mL CCl₄ and refluxed for 4 h under N₂. The solvent was removed *in vacuo*, and the precipitate was redissolved in 30 mL of DMSO. A separate solution containing N-hydroxy phthalimide (14.2 g, 82 mmol) and K₂CO₃ (12 g, 87 mmol) in 140 mL DMSO was stirred for 20 min. The two solutions were combined and stirred for 12 h. After 12 h, the reaction was poured into 600 mL of ice water, and a white precipitate was isolated by vacuum filtration. The precipitate was purified by silica column chromatography with 0-30% EtOAc in petroleum ether as the eluent. Pure fractions were combined and concentrated *in vacuo* yielding 5.43 g of product (27.1% yield). ¹H NMR: (400 MHz, CDCl₃) δ 7.76 (m, 4H), δ 4.71 (s, 2H), δ 1.91 (s, 3H), δ 1.76 (s, 3H), δ 1.70 (s, 3H).



Synthesis of O-(2,3-dimethylbut-2-en-1-yl)hydroxylamine hydrochloride

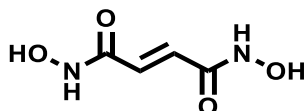
2-((2,3-dimethylbut-2-en-1-yl)oxy)isoindoline-1,3-dione (3.30 g, 13.34 mmol) was stirred in 50 mL of Et₂O with 4 molecular sieves for 12 h. The solution was filtered through a silica plug and washed with Et₂O. The filtrate was acidified with concentrated HCl to pH ~2, and the white

precipitate was collected by filtration yielding 1.21 g of product (60% yield). $^1\text{H NMR}$: (400 MHz, DMSO- d_6) δ 10.56 (s, 2H), δ 7.72 (s, 2H), δ 1.74 (s, 3H), δ 1.68 (s, 6H).



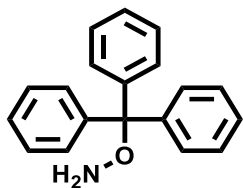
Synthesis of N^1,N^4 -bis((2,3-dimethylbut-2-en-1-yl)oxy)fumaramide

Fumaric Acid (69.6 mg, 0.600 mmol), EDC-HCl (276 mg, 1.44 mmol), HoBT (193 mg, 1.26 mmol), and DIPEA (3.59 mmol 0.627 mL) were combined in 10 mL of DCM and stirred for 10 min. *O*-(2,3-dimethylbut-2-en-1-yl)hydroxylamine hydrochloride (223 mg, 1.44 mmol) was added, and the reaction was stirred for 12 h. The solvent was removed in vacuo, and the product was purified by silica column chromatography with 0-80% EtOAc in hexanes as the eluent. Pure fractions were combined, and the solvent was removed in vacuo yielding 80 mg of product (43% yield). $^1\text{H NMR}$: (400 MHz, DMSO- d_6) δ 11.39 (s, 2H), δ 6.67 (s, 2H), δ 4.33 (s, 4H), δ 1.71 (s, 6H), 1.69 (s, 6H), 1.67 (s, 6H), 3.83 (s, 2H).



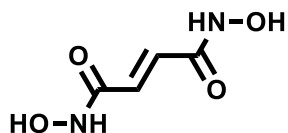
Synthesis of N^1,N^4 -bis((2,3-dimethylbut-2-en-1-yl)oxy)fumaramide (E -H₂edh)

N^1,N^4 -bis((2,3-dimethylbut-2-en-1-yl)oxy)fumaramide (20 mg, 0.258 mmol) was suspended in 10 mL of DCM with 10% v/v TFA. The reaction was stirred for 1 h, and the product was collected by filtration and dried *in vacuo*. This reaction gave >95% yield. $^1\text{H NMR}$: (400 MHz, DMSO- d_6) δ 11.01 (s, 1H), δ 9.24 (s, 2H), δ 6.70 (s, 4H).



Synthesis of O-tritylhydroxylamine

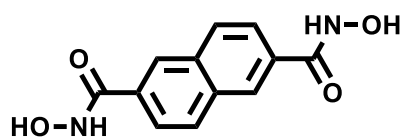
2-hydroxyisoindoline-1,3 dione (20 g, 123 mmol) and DIPEA (23.55 mL, 135 mmol) were combined in 100 mL of DMF and stirred at RT. Trityl Chloride (34.2 g, 123 mmol) was added in portions over 10 min. The reaction was stirred for 2 h. The reaction was poured into 600 mL of H₂O, and the precipitate was collected by filtration. The white precipitate was dissolved in 600 mL CHCl₃, and 30 mL of Hydrazine hydrate in 100 mL MeOH was added. The reaction was stirred at RT for 12 h. The reaction was extracted 3x with 600 mL H₂O, and the CHCl₃ layer was dried *in vacuo*. The product was purified with silica column chromatography using 0-5% EtOAc in Hexanes as the eluent. The pure fractions were combined, and the solvent was slowly evaporated over 24 h. Large crystals of pure product formed and were collected by filtration yielding 18.2 g of product (54%). ¹H NMR: (400 MHz, DMSO-d₆) δ 7.30 (m, 15H), δ 5.64 (s, 2H).



Synthesis of *N*¹,*N*⁴-bis((2,3-dimethylbut-2-en-1-yl)oxy)fumaramide (*E*-H₂edh)

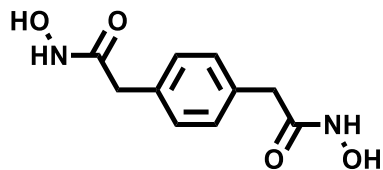
Fumaric acid (116 mg, 1 mmol), HATU (837 mg, 2.2 mmol), DIPEA (1.048 mL, 6 mmol), and O-tritylhydroxylamine (661 mg, 2.4 mmol) was added to 15 mL of DMF and stirred overnight at room temperature. The solution was suspended in 30 mL ethyl acetate and extracted with H₂O (2 x 30 mL) followed by 30 mL brine. The precipitate was collected by filtration, washed with 30 mL ethyl acetate, and dried *in vacuo* yielding 518 mg (82% yield). A portion of the precipitate (158 mg, 0.25 mmol) was suspended in 15 mL dichloromethane with 10% (v/v) trifluoroacetic acid and

stirred for 30 min. The remaining precipitate was filtered, washed with 10 mL dichloromethane, and dried *in vacuo* yielding 35.6 mg of product (97% yield). ¹H NMR: (400 MHz, DMSO-d₆) δ11.02 (br s, 2H), δ9.25 (br s, 2H), δ6.70 (s, 2H) ¹³C NMR: (500 MHz, DMSO-d₆) δ160.74, 129.54. Measured molecular weight (m/z) = 147.07 (Calc.: 147.03) (M + H⁺).



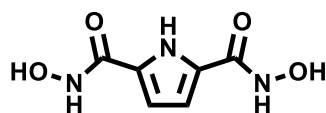
Synthesis of *N*²,*N*⁶-dihydroxynaphthalene-2,6-dicarboxamide (2,6-H₂ndh)

Naphthalene-2,6-dicarboxylic acid (216 mg, 1 mmol), HATU (837 mg, 2.2 mmol), DIPEA (1.048 mL, 6 mmol), and *O*-tritylhydroxylamine (661 mg, 2.4 mmol) were added to 15 mL of DMF and stirred overnight at room temperature. The solution was suspended in 30 mL ethyl acetate and extracted with H₂O (2 x 30 mL) followed by 30 mL brine. The precipitate was collected by filtration, washed with 30 mL ethyl acetate, and dried *in vacuo* yielding 669 mg (92% yield). A portion of the precipitate (183 mg, 0.25 mmol) was suspended in 15 mL dichloromethane with 10% (v/v) trifluoroacetic acid and stirred for 30 min. The remaining precipitate was filtered, washed with 10 mL dichloromethane, and dried *in vacuo* yielding 60.7 mg of product (99% yield). ¹H NMR: (400 MHz, DMSO-d₆) δ11.44 (br s, 2H), δ9.19 (br s, 2H), δ8.39 (s, 2H), δ8.08 (d, 2H), δ 7.88 (d, 2H). ¹³C NMR: (500 MHz, DMSO-d₆) δ163.96, 133.27, 131.52, 129.04, 126.92, 124.48. Measured molecular weight (m/z) = 244.95 (Calc.: 245.06) (M - H⁺).



Synthesis of 2,2'-(1,4-phenylene)bis(*N*-hydroxyacetamide) (*p*-H₂xdh)

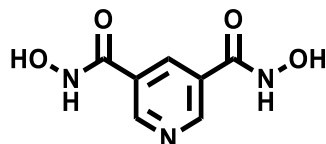
1,4-phenylene diacetic acid (194 mg, 1 mmol), HATU (837 mg, 2.2 mmol), DIPEA (1.048 mL, 6 mmol), and *O*-tritylhydroxylamine (661 mg, 2.4 mmol) were added to 15 mL of DMF and stirred overnight at room temperature. The solution was suspended in 30 mL ethyl acetate and extracted with H₂O (2 x 30 mL) followed by 30 mL brine. The precipitate was collected by filtration, washed with 30 mL ethyl acetate, and dried *in vacuo* yielding 538 mg (76% yield). A portion of the precipitate (177 mg, 0.25 mmol) was suspended in 15 mL dichloromethane with 10% (v/v) trifluoroacetic acid and stirred for 30 min. The remaining precipitate was filtered, washed with 10 mL dichloromethane, and dried *in vacuo* yielding 45.3 mg of product (81% yield). ¹H NMR: (400 MHz, DMSO-*d*₆) δ10.62 (br s, 2H), δ8.80 (br s, 2H), δ7.17 (s, 4H), δ3.23 (s, 4H). ¹³C NMR: (500 MHz, DMSO-*d*₆) δ167.08, 134.23, 128.77 Measured molecular weight (*m/z*) = 223.04 (Calc.: 223.08) (*M* - H⁺).



*N*₂,*N*₅-dihydroxy-1*H*-pyrrole-2,5-dicarboxamide (2,5-H₂pdh)

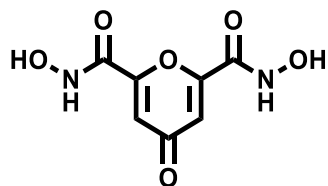
1*H*-pyrrole-2,5-dicarboxylic acid (156 mg, 1 mmol), HATU (837 mg, 2.2 mmol), DIPEA (1.048 mL, 6 mmol), and *O*-tritylhydroxylamine (661 mg, 2.4 mmol) were added to 15 mL of DMF and stirred overnight at room temperature. The solution was suspended in 30 mL ethyl acetate and extracted with H₂O (2 x 30 mL) followed by 30 mL brine. The reaction mixture was purified by silica gel chromatography with 10-50% EtOAc in hexanes as the eluent. Pure fractions were

combined, and the solvent was removed *in vacuo* yielding 286 mg of product (39.8 % yield). The precipitate was suspended in 15 ml dichloromethane with 10% (v/v) trifluoroacetic acid and stirred for 30 min. The remaining precipitate was filtered, washed with 10 mL dichloromethane, and dried *in vacuo* yielding 68 mg of product (86% yield). ¹H NMR: (400 MHz, DMSO-d₆) δ 11.57 (s, 2H), δ 10.93 (s, 1H), δ 9.02 (br s, 2H), δ 6.67 (s, 2H).



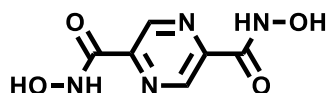
Synthesis of *N*⁸,*N*⁶-dihydroxypyridine-3,5-dicarboxamide (3,5-H₂pzdH)

Pyridine-3,5-dicarboxylic acid (500 mg, 2.99 mmol), HATU (2503 mg, 6.58 mmol), DIPEA (3.14 mL, 17.95 mmol), and *O*-tritylhydroxylamine (1977 mg, 7.18 mmol) were added to 15 mL of DMF and stirred overnight at room temperature. The solution was suspended in 30 mL ethyl acetate and extracted with H₂O (2 x 30 mL) followed by 30 mL brine. The precipitate was collected by filtration, washed with 30 mL ethyl acetate, and dried *in vacuo* yielding 1.0 g (49% yield). A portion of the precipitate (177 mg, 0.25 mmol) was suspended in 15 ml dichloromethane with 10% (v/v) trifluoroacetic acid and stirred for 30 min. The remaining precipitate was filtered, washed with 10 mL dichloromethane, and dried *in vacuo* yielding 200 mg of product (69% yield). ¹H NMR: (400 MHz, DMSO-d₆) δ 11.51 (s, 2H), δ 9.31 (s, 2H), δ 9.00 (s, 2H), δ 8.42 (s, 1H).



Synthesis of *N*²,*N*⁶-dihydroxy-4-oxo-4*H*-pyran-2,6-dicarboxamide (2,6-*H*₂podh)

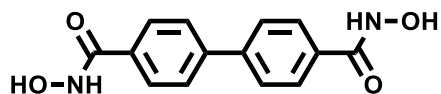
4-oxo-4*H*-pyran-2,6-dicarboxylic acid (184 mg, 1 mmol), HATU (837 mg, 2.2 mmol), DIPEA (1.048 mL, 6 mmol), and *O*-tritylhydroxylamine (661 mg, 2.4 mmol) were added to 15 mL of DMF and stirred overnight at room temperature. The solution was suspended in 30 mL ethyl acetate and extracted with H₂O (2 x 30 mL) followed by 30 mL brine. The precipitate was collected by filtration, washed with 30 mL ethyl acetate, and dried *in vacuo*. A portion of the precipitate (0.25 mmol, 175 mg) was suspended in 15 mL dichloromethane with 10% (v/v) trifluoroacetic acid and stirred for 30 min. The remaining precipitate was filtered, washed with 10 mL dichloromethane, and dried *in vacuo* yielding 24.8 mg of product (46% yield). Additional product could be isolated by removing the filtrate *in vacuo*, resuspending the precipitate with DCM, and filtering the product again. ¹H NMR: (400 MHz, DMSO-*d*₆) δ 11.85 (s, 2H), δ 9.75 (s, 2H), δ 6.85 (s, 2H).



Synthesis of *N*²,*N*⁶-dihydroxypyrazine-2,5-dicarboxamide

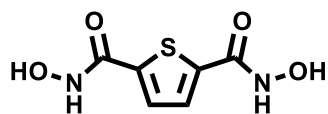
Pyrazine-2,5-dicarboxylic acid dihydrate (204 mg, 1 mmol), HATU (837 mg, 2.2 mmol), DIPEA (1.048 mL, 6 mmol), and *O*-tritylhydroxylamine (661 mg, 2.4 mmol) were added to 15 mL of DMF and stirred overnight at room temperature. The solution was suspended in 30 mL ethyl acetate and extracted with H₂O (2 x 30 mL) followed by 30 mL brine. The precipitate was collected by filtration, washed with 30 mL ethyl acetate, and dried *in vacuo* yielding 560.5 mg (82% yield). A portion of the precipitate (0.25 mmol, 171 mg) was suspended in 15 mL dichloromethane with

10% (v/v) trifluoroacetic acid and stirred for 30 min. The remaining precipitate was filtered, washed with 10 mL dichloromethane, and dried *in vacuo* yielding 42.1 mg (85% yield). ¹H NMR: (400 MHz, DMSO-d₆) δ 11.81 (s, 2H), δ 9.36 (s, 2H), δ 9.09 (s, 2H).



Synthesis of [1,1'-biphenyl]-4,4'-dicarboxylic acid

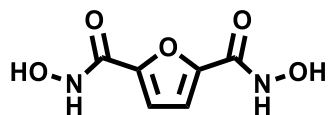
[1,1'-biphenyl]-4,4'-dicarboxylic acid (242 mg, 1 mmol), HATU (837 mg, 2.2 mmol), DIPEA (1.048 mL, 6 mmol), and *O*-tritylhydroxylamine (661 mg, 2.4 mmol) were added to 15 mL of DMF and stirred overnight at room temperature. The solution was suspended in 30 mL ethyl acetate and extracted with H₂O (2 x 30 mL) followed by 30 mL brine. The precipitate was collected by filtration, washed with 30 mL ethyl acetate, and dried *in vacuo* yielding 706.9 mg (93% yield). A portion of the precipitate (0.25 mmol, 189 mg) was suspended in 15 mL dichloromethane with 10% (v/v) trifluoroacetic acid and stirred for 30 min. The remaining precipitate was filtered, washed with 10 mL dichloromethane, and dried *in vacuo* yielding 64.7 mg of product (95% yield). ¹H NMR: (400 MHz, DMSO-d₆) δ 11.31 (s, 2H), δ 9.11 (s, 2H), δ 7.83 (dd, 8H).



Synthesis of *N,N'*-dihydroxythiophene-2,5-dicarboxamide

Thiophene-2,5-dicarboxylic acid (172 mg, 1 mmol), HATU (837 mg, 2.2 mmol), DIPEA (1.048 mL, 6 mmol), and *O*-tritylhydroxylamine (661 mg, 2.4 mmol) was added to 15 mL of DMF and stirred overnight at room temperature. The solution was suspended in 30 mL ethyl acetate and extracted with H₂O (2 x 30 mL) followed by 30 mL brine. The precipitate was collected by filtration, washed with 30 mL ethyl acetate, and dried *in vacuo* yielding 0.488 mg (71% yield). A

portion of the precipitate (0.25 mmol, 172 mg) was suspended in 15 ml dichloromethane with 10% (v/v) trifluoroacetic acid and stirred for 30 min. The remaining precipitate was filtered, washed with 10 mL dichloromethane, and dried *in vacuo* yielding 28.8 mg of product (57% yield). Additional product could be isolated by removing the filtrate *in vacuo*, resuspending the precipitate with DCM, and filtering the product again. ¹H NMR: (400 MHz, DMSO-d₆) δ 11.36 (s, 2H), δ 9.25 (s, 2H), δ 7.54 (s, 2H).



Synthesis of *N*²,*N*⁶-dihydroxyfuran-2,5-dicarboxamide (2,5-H₂fdh)

Furan-2,5-dicarboxylic acid (500 mg, 3.2 mmol), HATU (2680 mg, 7.05 mmol), DIPEA (3.36 mL, 19.22 mmol), and *O*-tritylhydroxylamine (2117 mg, 7.69 mmol) were added to 15 mL of DMF and stirred overnight at room temperature. The solution was suspended in 30 mL ethyl acetate and extracted with H₂O (2 x 30 mL) followed by 30 mL brine. 3 mL of trifluoroacetic acid was added, and the reaction was stirred for 30 min. The remaining precipitate was filtered, washed with 10 mL dichloromethane, and dried *in vacuo* yielding 402 mg of product (67% yield). ¹H NMR: (400 MHz, DMSO-d₆) δ 11.14 (s, 2H), δ 9.34 (s, 2H), δ 7.11 (s, 2H). Measured molecular weight (m/z) = 187.05 (Calc.: 187.03) (M + H⁺).

3.3.5. Crystallization of ferritin-MOFs

Stock solutions of ferritin (25 μM, based on 24-meric state) stored at 4 °C in a buffered solution (50 mM CHES, pH 8.5, 150 mM NaCl) were warmed to room temperature before use in all self-assembly experiments. Formation of protein-MOFs occurs once the three components (ferritin, metal, and ligand) are combined; crystals with rhombic dodecahedral morphology (≥20 μm) appear within 12-24 h. While the ideal conditions for each metal-linker combination vary

slightly, there are common conditions that promote the growth of larger crystals: protein concentrations between 2-4 μM of ferritin, 1-2 mM of bridging linker, and 50-150 equivalents of either ZnCl_2 , NiCl_2 , or CoCl_2 per ferritin cage. We set up ferritin-MOFs in 24-well crystal trays or 24-well culture plates to make crystal harvesting more convenient. A screen of common precipitants was performed with concentrations of 1-16% (w/v) to facilitate the growth of large crystals. Low molecular weight PEGs (PEG 300, PEG 350 MME, and PEG 400) and pentaerythritol propoxylate (5/4 PO/OH) were identified as improving the size and quality of crystal growth. As with their precipitant-free counterparts, no crystals formation was observed when any one of the components (protein, ligand, or metal) was excluded from the sample matrix.

3.3.4. Collection and processing of SAXS profiles

Crystals for SAXS were prepared in 24-well crystallization trays. The reservoir solution contained 500 μL of 50 mM CHES (pH 9.5), 150 mM NaCl, and 474 μM metal (either ZnCl_2 , NiCl_2 , or CoCl_2). The well solution was comprised of 2 μL of a 25 μM ferritin solution in 50 mM CHES pH 8.5, 150 mM NaCl, 2.4 μL of a solution containing 10 mM linker in 50 mM CHES pH 9.5, 150 mM NaCl, and 7.6 μL of the reservoir solution. With these conditions, the upper well contains 4 μM ferritin, 300 μM metal, and 2 mM linker. Each tray contained 24 identical wells with a single metal/linker combination. Crystals formed overnight and were harvested after three days. For each tray, the crystals from all 24 wells were combined in a 1.5 mL Eppendorf tube. After the crystals settled to the bottom of the tube, the crystals were transferred to a 1.5-mm quartz capillary (Hampton Scientific) along with 50 μL of combined reservoir solution. Capillaries were sealed with clay and shipped to beamline 5-ID-D (Argonne National Laboratory-Advanced Photon Source (APS)) for SAXS analysis. Data were collected with collimated X-rays (0.7293 \AA , 17 keV) calibrated with both a glassy carbon standard and a silicon diffraction grating. The X-ray beam had approximate dimensions of 250 μm x 250 μm , and the exposure time used varied between

0.5-2 s. Scattered radiation was detected using a CCD area detector, and 1D scattering data were obtained through the azimuthal averaging of the 2D data to obtain plots of scattering intensity as a function of the scattering vector q :

$$q = 4\pi \sin(\theta) / \lambda$$

Where θ is one-half of the scattering angle, and λ is the wavelength of the X-rays used. Analysis of the 1D data was performed using the powder diffraction processing software JADE (MDI). Minor adjustments (such as increasing q by a factor of 10) are needed for JADE to function properly with such small scattering angles. Under certain conditions, multiple lattices could be indexed. The simulated scattering patterns were obtained by first using the 5UP8 structure to create a protein coordinate file containing three ferritin monomers surrounding the internal C_3 axis of ferritin. The CRYST1 record of the coordinate file was then altered to match the appropriate unit cell geometry for a given lattice. The powder diffraction feature of Mercury⁴⁶ was used to read this coordinate file and simulate the diffraction of each lattice. The peak intensities of the simulated patterns seem to correlate with the observed spectra. This is highlighted in Figure 3.5, where the unaltered crystal structure coordinates (5UP7 and 5UP8) were used to generate the simulated SAXS profiles.

3.3.6. Structural determination of ferritin-MOFs by sc-XRD

Crystals suitable for sc-XRD experiment were harvested and soaked briefly in 50% (w/v) xylitol with 5 mM of *p*-H₂bdh for *p*-bdh-Zn-ferritin; 50% (w/v) xylitol for *p*-xdh-Zn-ferritin; 50% (w/v) PEP, 2 mM *m*-bdh 0.3 mM ZnCl₂, 150 mM NaCl, 50 mM CHES (pH 9.5) for *m*-bdh-Zn-ferritin; 50% (w/v) xylitol, 2 mM *m*-bdh, 0.3 mM NiCl₂, 150 mM NaCl, 50 mM CHES (pH 9.5) for *m*-bdh-Ni-ferritin; 50% (w/v) xylitol for 2,5-fdh-Zn-ferritin; 2 mM 2,5-fdh, 0.3 mM NiCl₂, 150 mM NaCl, 50 mM CHES (pH 9.5) for 2,5-fdh-Ni-ferritin. After cryoprotection, the crystals were quickly frozen in liquid N₂. sc-XRD data were collected at 100 K at BL 12-2 or 9-1 of the Stanford Synchrotron Radiation Laboratory (SSRL) using 0.98 Å radiation. The data were integrated using iMosflm⁴⁸

and scaled using Aimless⁴⁹ (see for statistics). The structures of *p*-bdh-Zn-ferritin, *p*-xdh-Zn-ferritin, *m*-bdh-Zn-ferritin, *m*-bdh-Ni-ferritin, 2,5-fdh-Zn-ferritin, and 2,5-fdh-Ni-ferritin were solved to resolutions of 2.63 Å, 2.45 Å, 2.07 Å, 2.34 Å, 2.38 Å, and 2.28 Å resolution respectively. Molecular replacement was performed with Phaser⁵¹ using the previously solved ¹²²Hf-ferritin structure (PDB ID: 5CMQ) as the search model. Rigid-body, positional, thermal, and TLS refinements were carried out using Phenix.⁵² NCS constraints based on the ordered ferritin node were employed on the alternate conformations for the three P2₁ structures. The eLBOW⁵⁴ module in Phenix was used to generate restraints for all ligands. Coot⁵⁵ was used for iterative manual model building. All figures were produced with Pymol.⁵⁶

Table 3.4. Crystallization conditions for ferritin-MOFs

<i>p</i> -bdh-Zn-ferritin	Reservoir	500 µL total volume: 50 mM CHES (pH 8.5), 150 mM NaCl, 0.140 mM ZnCl ₂
	Sitting Drop	8.6 µL reservoir, 1 µL of 25 µM ferritin, 2.4 µL of 5 mM <i>p</i> -H ₂ bdh in 50 mM CHES (pH 9.5) with 150 mM NaCl
<i>p</i> -xdh-Zn-ferritin	Reservoir	500 µL total volume: 50 mM CHES (pH 8.5), 150 mM NaCl, 0.474 mM ZnCl ₂ , 3% PEG 300
	Sitting Drop	7.6 µL reservoir, 2 µL of 25 µM ferritin, 2.4 µL of 5 mM <i>p</i> -H ₂ xdh in 50 mM CHES (pH 9.5) with 150 mM NaCl
<i>m</i> -bdh-Zn-ferritin	Reservoir	500 µL total volume: 50 mM CHES (pH 9.5), 150 mM NaCl, 0.474 mM ZnCl ₂ , 12.6% PEP (5/4 PO/OH)
	Sitting Drop	7.6 µL reservoir, 2 µL of 25 µM ferritin, 2.4 µL of 5 mM <i>m</i> -H ₂ bdh in 50 mM CHES (pH 9.5) with 150 mM NaCl
<i>m</i> -bdh-Ni-ferritin	Reservoir	500 µL total volume: 50 mM CHES (pH 9.5), 150 mM NaCl, 0.474 mM NiCl ₂ , 10 % (v/v) PEP (5/4 PO/OH)
	Sitting Drop	7.6 µL reservoir, 2 µL of 25 µM ferritin, 2.4 µL of 5 mM <i>m</i> -H ₂ bdh in 50 mM CHES (pH 9.5) with 150 mM NaCl
2,5-fdh-Zn-ferritin	Reservoir	500 µL total volume: 50 mM CHES (pH 8.5), 150 mM NaCl, 0.474 mM ZnCl ₂ , 10% PEP (5/4 PO/OH)
	Sitting Drop	7.6 µL reservoir, 2 µL of 25 µM ferritin, 2.4 µL of 5 mM 2,5-H ₂ fdh in 50 mM CHES (pH 9.5) with 150 mM NaCl
2,5-fdh-Ni-ferritin	Reservoir	500 µL total volume: 50 mM CHES (pH 10.0), 150 mM NaCl, 0.474 mM NiCl ₂ , 19.0 % v/v PEP (5/4 PO/OH)
	Sitting Drop	7.6 µL reservoir, 2 µL of 25 µM ferritin, 2.4 µL of 5 mM 2,5-H ₂ fdh in 50 mM CHES (pH 9.5) with 150 mM NaCl

Table 3.5. X-ray data processing and refinement statistics of *bcc* and *bct* ferritin-MOFs

	<i>p</i> -bdh-Zn-ferritin (5UP8)	<i>p</i> -xdh-Zn-ferritin (5UP9)	2,5-fdh-Ni-ferritin*
Data collection			
Space group	<i>I</i> 432	<i>I</i> 4	<i>I</i> 432
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	155.81	149.88,162.23	154.90
α , β , γ (°)	90	90	90
Resolution (Å)	49.27-2.63	47.76-2.45	63.24-2.28
No. unique reflections	9926	64992	14174
Multiplicity	31.8 (23.0)	5.6 (3.3)	12.0 (8.7)
<i>CC</i> 1/2	0.999 (0.554)	0.993 (0.945)	0.999 (0.762)
<i>R</i> merge	0.189 (1.658)	0.139 (0.314)	0.086 (0.913)
$\langle I / \sigma \rangle$	15.0 (2.3)	8.1 (2.2)	16.3 (1.9)
Completeness (%)	100 (100)	99.1 (95.0)	99.8 (99.3)
Refinement			
<i>R</i> _{work} / <i>R</i> _{free}	0.209/0.262	0.241/0.289	0.211/0.260
No. atoms			
Protein	1403	8547	1427
Ligand/ion	18	218	38
Water	12	1146	88
<i>B</i> -factors (Å ²)			
Protein	84.6	23.8	53.29
Ligand/ion	85.8	38.6	79.81
Water	69.6	31.8	57.08
R.m.s. deviations			
Bond lengths (Å)	0.0029	0.0034	0.118
Bond angles (°)	0.555	0.657	3.51
MolProbity ⁵³ score	2.04	1.67	1.88
Clashscore	15.15	12.26	8.48
Ramachandran plot (%)			
Favored	94.83	97.96	95.93
Outliers	0.57	0.00	0.58
Rotamers (%)			
Poor	0.70	1.22	2.04

*Crystal structure is preliminary and has not been published. Refinement statistics will improve before deposition.

Table 3.6. X-ray data processing and refinement statistics of primitive ferritin-MOFs

	<i>m</i> -bdh-Zn-ferritin*	<i>m</i> -bdh-Ni-ferritin*	2,5-fdh-Zn-ferritin*
Data collection			
Space group	<i>P</i> 2 ₁	<i>P</i> 2 ₁	<i>P</i> 2 ₁
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	215.34, 215.61, 156.35	213.55, 213.74, 155.93	213.99, 213.91, 156.33
α, β, γ (°)	90, 89.974, 90	90, 90.048, 90	90, 89.846, 90
Resolution (Å)	39.29-2.07	39.09-2.34	39.20-2.38
No. unique reflections	847172	575840	538057
Multiplicity	3.5 (3.6)	3.5 (3.6)	3.4 (2.9)
<i>CC</i> 1/2	0.997 (0.501)	0.997 (0.454)	0.997 (0.796)
<i>R</i> merge	0.099 (0.817)	0.098 (0.838)	0.064 (0.259)
<i><I / σI></i>	9.9 (1.6)	11.0 (1.6)	13.6 (3.6)
Completeness (%)	98.2 (98.9)	98.2 (98.4)	96.0 (78.8)
Refinement			
<i>R</i> _{work} / <i>R</i> _{free}	0.182/0.199	0.194/0.255	0.209/0.297
No. atoms			
Protein	103032	103032	103032
Ligand/ion	434	434	626
Water	5722	2393	1953
<i>B</i> -factors (Å ²)			
Protein	31.55	39.94	37.55
Ligand/ion	46.19	61.16	50.32
Water	46.33	50.30	45.50
R.m.s. deviations			
Bond lengths (Å)	0.015	0.017	0.021
Bond angles (°)	1.48	1.66	1.68
MolProbity ⁵³ score	1.59	1.76	1.96
Clashscore	5.36	13.86	13.57
Ramachandran plot (%)			
Favored	97.94	97.43	95.47
Outliers	0	0	0.44
Rotamers (%)			
Poor	0	0	0

*Crystal structures are preliminary and have not been published. Refinement statistics will improve before deposition.

3.4. Acknowledgments

Chapter 3, in part, is a reprint of the material “Synthetic Modularity of Protein-Metal-Organic Frameworks” as it appears in the Journal of the American Chemical Society (Bailey, J. B.; Zhang, L.; Chiong, J. A.; Ahn, S.; Tezcan, F. A. **2017**, *139* (24), 8160-8166). The dissertation author is the primary author of this publication. Chapter 3 also includes work that is currently being prepared for publication with contributions from the following authors: Bailey, J. B.; Tezcan, F. A. The dissertation author will be the primary author of this publication.

3.5. References

1. Mann, S., Life as a nanoscale phenomenon. *Angew. Chem. Int. Ed. Engl.* **2008**, *47* (29), 5306-5320.
2. Yang, L.; Liu, A.; Cao, S.; Putri, R. M.; Jonkheijm, P.; Cornelissen, J. J., Self-Assembly of Proteins: Towards Supramolecular Materials. *Chemistry-A European Journal* **2016**, *22* (44), 15570-15582.
3. Ueno, T., Porous Protein Crystals as Reaction Vessels. *Chemistry-a European Journal* **2013**, *19* (28), 9096-9102.
4. Yeates, T. O.; Liu, Y.; Laniado, J., The design of symmetric protein nanomaterials comes of age in theory and practice. *Current Opinion in Structural Biology.* **2016**, *39*, 134-143.
5. Kobayashi, N.; Arai, R., Design and construction of self-assembling supramolecular protein complexes using artificial and fusion proteins as nanoscale building blocks. *Current Opinion in Biotechnology* **2017**, *46*, 57-65.
6. Salgado, E. N.; Radford, R. J.; Tezcan, F. A., Metal-Directed Protein Self-Assembly. *Accounts of Chemical Research* **2010**, *43* (5), 661-672.
7. Radford, R. J.; Brodin, J. D.; Salgado, E. N.; Tezcan, F. A., Expanding the utility of proteins as platforms for coordination chemistry. *Coord. Chem. Rev.* **2011**, *255* (7-8), 790-803.
8. Lai, Y.-T.; Hura, G. L.; Dyer, K. N.; Tang, H. Y.; Tainer, J. A.; Yeates, T. O., Designing and defining dynamic protein cage nanoassemblies in solution. *Sci. Adv.* **2016**, *2* (12), e1501855.
9. King, N. P.; Bale, J. B.; Sheffler, W.; McNamara, D. E.; Gonen, S.; Gonen, T.; Yeates, T. O.; Baker, D., Accurate design of co-assembling multi-component protein nanomaterials. *Nature* **2014**, *510* (7503), 103-+.
10. Sciore, A.; Su, M.; Koldewey, P.; Eschweiler, J. D.; Diffley, K. A.; Linhares, B. M.; Ruotolo, B. T.; Bardwell, J. C.; Skiniotis, G.; Marsh, E. N. G., Flexible, symmetry-directed approach to

assembling protein cages. *Proceedings of the National Academy of Sciences* **2016**, 113 (31), 8681–8686.

11. Ni, T. W.; Tezcan, F. A., Structural Characterization of a Microperoxidase Inside a Metal-Directed Protein Cage. *Angewandte Chemie-International Edition* **2010**, 49 (39), 7014-7018.

12. Brodin, J. D.; Ambroggio, X. I.; Tang, C. Y.; Parent, K. N.; Baker, T. S.; Tezcan, F. A., Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. *Nature Chemistry* **2012**, 4 (5), 375-382.

13. Brodin, J. D.; Smith, S. J.; Carr, J. R.; Tezcan, F. A., Designed, Helical Protein Nanotubes with Variable Diameters from a Single Building Block. *Journal of the American Chemical Society* **2015**, 137 (33), 10468-10471.

14. Yang, G.; Zhang, X.; Kochovski, Z.; Zhang, Y.; Dai, B.; Sakai, F.; Jiang, L.; Lu, Y.; Ballauff, M.; Li, X., Precise and Reversible Protein-Microtubule-Like Structure with Helicity Driven by Dual Supramolecular Interactions. *Journal of the American Chemical Society* **2016**, 138 (6), 1932-1937.

15. Ballister, E. R.; Lai, A. H.; Zuckermann, R. N.; Cheng, Y.; Mougous, J. D., In vitro self-assembly from a simple protein of tailorable nanotubes building block. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, 105 (10), 3733-3738.

16. Sinclair, J. C.; Davies, K. M.; Venien-Bryan, C.; Noble, M. E. M., Generation of protein lattices by fusing proteins with matching rotational symmetry. *Nature Nanotechnology* **2011**, 6 (9), 558-562.

17. Gonen, S.; DiMaio, F.; Gonen, T.; Baker, D., Design of ordered two-dimensional arrays mediated by noncovalent protein-protein interfaces. *Science* **2015**, 348 (6241), 1365-1368.

18. Suzuki, Y.; Cardone, G.; Restrepo, D.; Zavattieri, P. D.; Baker, T. S.; Tezcan, F. A., Self-assembly of coherently dynamic, auxetic, two-dimensional protein crystals. *Nature* **2016**, 533 (7603), 369-+.

19. Lanci, C. J.; MacDermaid, C. M.; Kang, S. G.; Acharya, R.; North, B.; Yang, X.; Qiu, X. J.; DeGrado, W. F.; Saven, J. G., Computational design of a protein crystal. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, 109 (19), 7304-9.

20. Sakai, F.; Yang, G.; Weiss, M. S.; Liu, Y.; Chen, G.; Jiang, M., Protein crystalline frameworks with controllable interpenetration directed by dual supramolecular interactions. *Nature Communications* **2014**, 5.

21. Künzle, M.; Eckert, T.; Beck, T., Binary Protein Crystals for the Assembly of Inorganic Nanoparticle Superlattices. *Journal of the American Chemical Society* **2016**, 138 (39), 12731-12734.

22. Kostianen, M. A.; Hiekkataipale, P.; Laiho, A.; Lemieux, V.; Seitsonen, J.; Ruokolainen, J.; Ceci, P., Electrostatic assembly of binary nanoparticle superlattices using protein cages. *Nature Nanotechnology* **2013**, 8 (1), 52-+.

23. Joh, N. H.; Wang, T.; Bhate, M. P.; Acharya, R.; Wu, Y.; Grabe, M.; Hong, M.; Grigoryan, G.; DeGrado, W. F., De novo design of a transmembrane Zn²⁺-transporting four-helix bundle. *Science* **2014**, *346* (6216), 1520-1524.
24. Votteler, J.; Ogohara, C.; Yi, S.; Hsia, Y.; Nattermann, U.; Belnap, D. M.; King, N. P.; Sundquist, W. I., Designed proteins induce the formation of nanocage-containing extracellular vesicles. *Nature* **2016**, *540* (7632), 292-+.
25. Song, W. J.; Tezcan, F. A., A designed supramolecular protein assembly with in vivo enzymatic activity. *Science* **2014**, *346* (6216), 1525-1528.
26. Brodin, J. D.; Carr, J. R.; Sontz, P. A.; Tezcan, F. A., Exceptionally stable, redox-active supramolecular protein assemblies with emergent properties. *Proceedings of the National Academy of Sciences of the United States of America* **2014**, *111* (8), 2897-2902.
27. Chakrabarty, R.; Mukherjee, P. S.; Stang, P. J., Supramolecular Coordination: Self-Assembly of Finite Two- and Three-Dimensional Ensembles. *Chemical Reviews* **2011**, *111* (11), 6810-6918.
28. Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J.-L.; Sanders, J. K. M.; Otto, S., Dynamic Combinatorial Chemistry. *Chemical Reviews* **2006**, *106* (9), 3652-3711.
29. Furukawa, H.; Cordova, K. E.; O'Keeffe, M.; Yaghi, O. M., The chemistry and applications of metal-organic frameworks. *Science* **2013**, *341* (6149), 1230444.
30. Zhou, H.-C.; Kitagawa, S., Metal-organic frameworks (MOFs). *Chemical Society Reviews* **2014**, *43* (16), 5415-5418.
31. Eddaoudi, M.; Kim, J.; Rosi, N.; Vodak, D.; Wachter, J.; O'Keeffe, M.; Yaghi, O. M., Systematic design of pore size and functionality in isorecticular MOFs and their application in methane storage. *Science* **2002**, *295* (5554), 469-72.
32. Howarth, A. J.; Peters, A. W.; Vermeulen, N. A.; Wang, T. C.; Hupp, J. T.; Farha, O. K., Best Practices for the Synthesis, Activation, and Characterization of Metal-Organic Frameworks. *Chem. Mater.* **2017**, *29* (1), 26-39.
33. Chui, S. S.-Y.; Lo, S. M.-F.; Charmant, J. P. H.; Orpen, A. G.; Williams, I. D., A Chemically Functionalizable Nanoporous Material [Cu₃(TMA)₂(H₂O)₃]_n. *Science* **1999**, *283* (5405), 1148-1150.
34. Li, J. R.; Sculley, J.; Zhou, H. C., Metal-organic frameworks for separations. *Chem. Rev.* **2012**, *112* (2), 869-932.
35. Horike, S.; Shimomura, S.; Kitagawa, S., Soft porous crystals. *Nat. Chem.* **2009**, *1* (9), 695-704.
36. Yoon, M.; Srirambalaji, R.; Kim, K., Homochiral metal-organic frameworks for asymmetric heterogeneous catalysis. *Chemical reviews* **2011**, *112* (2), 1196-1231.

37. Sumida, K.; Rogow, D. L.; Mason, J. A.; McDonald, T. M.; Bloch, E. D.; Herm, Z. R.; Bae, T. H.; Long, J. R., Carbon dioxide capture in metal-organic frameworks. *Chem. Rev.* **2012**, *112* (2), 724-81.
38. Horcajada, P.; Gref, R.; Baati, T.; Allan, P. K.; Maurin, G.; Couvreur, P.; Ferey, G.; Morris, R. E.; Serre, C., Metal-organic frameworks in biomedicine. *Chem. Rev.* **2012**, *112* (2), 1232-68.
39. Wang, Z.; Cohen, S. M., Postsynthetic modification of metal-organic frameworks. *Chemical Society Reviews* **2009**, *38* (5), 1315-1329.
40. Sontz, P. A.; Bailey, J. B.; Ahn, S.; Tezcan, F. A., A metal organic framework with spherical protein nodes: rational chemical design of 3D protein crystals. *Journal of the American Chemical Society* **2015**, *137* (36), 11598-11601.
41. Lawson, D. M.; Artymiuk, P. J.; Yewdall, S. J.; Smith, J. M.; Livingstone, J. C.; Treffry, A.; Luzzago, A.; Levi, S.; Arosio, P.; Cesareni, G.; Thomas, C. D.; Shaw, W. V.; Harrison, P. M., Solving the structure of human H ferritin by genetically engineering intermolecular crystal contacts. *Nature* **1991**, *349* (6309), 541-4.
42. Griffith, D.; Krot, K.; Comiskey, J.; Nolan, K. B.; Marimon, C. J., Monohydroxamic acids and bridging dihydroxamic acids as chelators to ruthenium(III) and as nitric oxide donors: syntheses, speciation studies and nitric oxide releasing investigation. *Dalton Transactions* **2008**, 137-147.
43. Nikitjuka, A.; Jirgensons, A., Synthesis of Hydroxamic Acids by Using the Acid Labile O-2-Methylprenyl Protecting Group. *Synlett* **2012**, *23* (20), 2972-2974.
44. Dissanayake, D. P.; Senthilnithy, R., Thermodynamic cycle for the calculation of ab initio pKa values for hydroxamic acids. *Journal of Molecular Structure: THEOCHEM* **2009**, *910* (1), 93-98.
45. Dick, B. L.; Cohen, S. M., Metal-Binding Isosteres as New Scaffolds for Metalloenzyme Inhibitors. *Inorganic Chemistry* **2018**, *57* (15), 9538-9543.
46. Macrae, C. F.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Shields, G. P.; Taylor, R.; Towler, M.; Streek, J. v. d., Mercury: visualization and analysis of crystal structures. *Journal of Applied Crystallography* **2006**, *39* (3), 453-457.
47. Bertini, I.; Luchinat, C., The reaction pathways of zinc enzymes and related biological catalysts. In *Bioinorganic Chemistry*, Bertini, I.; Gray, H. B.; Lippard, S. J.; Valentine, J. S., Eds. University Science Books: Mill Valley, 1994; pp 37-106.
48. Battye, T. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G., iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67* (Pt 4), 271-81.
49. Evans, P. R.; Murshudov, G. N., How good are my data and what is the resolution? *Acta Crystallographica Section D: Biological Crystallography* **2013**, *69* (7), 1204-1214.

50. Krause, L.; Herbst-Irmer, R.; Sheldrick, G. M.; Stalke, D., Comparison of silver and molybdenum microfocus X-ray sources for single-crystal structure determination. *Journal of applied crystallography* **2015**, *48* (1), 3-10.
51. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phaser crystallographic software. *Journal of applied crystallography* **2007**, *40* (4), 658-674.
52. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66* (Pt 2), 213-21.
53. Chen, V. B.; Arendall, W. B. I.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C., MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 12-21.
54. Moriarty, N. W.; Grosse-Kunstleve, R. W.; Adams, P. D., electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallographica Section D: Biological Crystallography* **2009**, *65* (10), 1074-1080.
55. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486-501.
56. The PyMOL Molecular Graphics System, V. S., LLC.

Chapter 4. Thermal behavior of ferritin-MOFS

4.1. Introduction

From nm-to- μm -sized microtubules and actin filaments to collagen fibers and spider silk at near-meter length scales, protein-based materials are central to the construction and operation of living systems.¹⁻⁵ These protein assemblies are distinguished by their extended periodic or hierarchical structures as well as their capacity for change in response to various physical, mechanical, and chemical perturbations, which underpins their biological functions as dynamic architectural materials. Ultimately, the emergent properties of protein biomaterials are governed not only by the inherent attributes of each building block, but also the precise nature of the intersubunit connectivity that forms the basis for the assembled architectures. Therefore, the ability to produce protein-based materials with desirable, bulk-scale physical and mechanical properties hinges on our ability to tailor the chemical interactions between protein building blocks in a tunable manner.

Many approaches have been developed over the last decade to design and control protein-protein interactions (discussed in detail in Chapter 1), leading to a variety of protein assemblies with new structures, and occasionally, new functions.⁶⁻¹¹ Despite these advances, there are still relatively few examples of extended protein materials obtained by design that possess novel physical and mechanical properties.^{10, 12} This is partially due to the intersubunit contact areas created through most design strategies, which do not allow for dynamic rearrangements. One of the only designed protein architectures to exhibit coherently dynamic properties, 2D ⁹⁸C RhuA crystals, has minimal protein-protein contact areas (between 361 and 115 \AA^2), which allow for auxetic crystal-to-crystal transformations. However, this behavior was not anticipated *a priori*. It is exceedingly difficult—if not impossible—to project how the design of dynamic intermolecular interactions at the \AA -nm scale would influence the collective properties of extended materials at the bulk scale. Thus, accessing novel materials properties requires a highly

modular design platform that allows for the systematic tuning of the intersubunit connectivity. As described in chapters 2 and 3, we developed a modular methodology for the construction of protein materials and employed it to synthesize an extensive library of ferritin-MOFs.¹³⁻¹⁴ These ferritin-MOFs are ideal for the emergence of novel materials properties due to their modular construction and sparse internode connectivity.

These ferritin-MOFs are highly crystalline, despite the massive size mismatch between the ferritin node and synthetic bridging ligands. In terms of node/linker mass ratios, ferritin-MOFs (ca. 2500/1) are akin to a lattice of soccer balls held together by toothpicks. Remarkably, the total footprint of hydroxamate-metal interactions is ca. 176 Å² (22 Å² × 8) and represents only 0.4% of the outer surface of a ferritin cage (ca. 44,000 Å²). These metal-hydroxamate interactions account for the entirety of the crystal contact area for ferritin-MOFs that form lattices with *bcc* symmetry. The *bct* and pseudo-*bct* ferritin-MOFs also have ultralow internode connectivity, even when accounting for the sparse protein-protein contact area. Figure 4.1 shows the crystal contact surface area (Å²) of >700 protein crystals plotted against the percentage of the protein surface participating in a crystal contact. The six crystallographically characterized protein-MOFs are shown in orange and occupy the lower-left portion of the graph, indicating that both the total contact area and contact area percentage are significantly lower than the average protein crystal, making ferritin-MOFs the most sparsely connected protein-based 3D frameworks. The sparse connectivity of ferritin-MOFs becomes readily apparent when compared with other porous frameworks, such as conventional MOFs. Ferritin-MOFs have a volume of ca. 237,000 Å³ per metal-ligand bond. Even the most porous MOFs, NU110E and DUT60, only have volumes of ca. 845 and 1078 Å³ per metal-ligand bond, respectively.¹⁵⁻¹⁶ The extremely low crystal contact area, coupled with the inherent flexibility of the metal-ligand bonds, allows for the possibility of cooperative dynamic behavior in the ferritin-MOF crystals. Indeed, we first observed dynamic crystal-to-crystal transformations in *bcc* ferritin-MOFs while measuring the thermal stability of the

architectures. This chapter will discuss the thermal stability of ferritin-MOFs as well as the emergence of dynamic temperature-dependent crystal-to-crystal transformations.

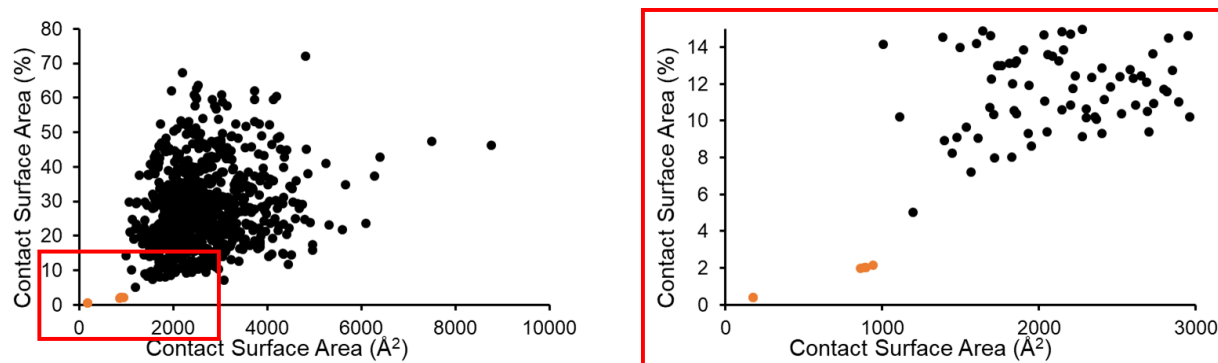


Figure 4.1. Crystal contact surface area \AA^2 vs. % of protein surface that forms a crystal contact for >700 proteins. The ferritin-MOFs characterized with sc-XRD are shown in orange.

4.2. Results and discussion

4.2.1. Assessing the thermal behavior of ferritin-MOFs by SAXS

To determine the thermal behavior of ferritin-MOFs, we turned to time-resolved solution-based PXRD measurements conducted at SAXS beamlines. Both beamline 4-2 (SSRL) and beamline 5-ID-D (APS) have custom-built capillary stages that precisely regulate the capillary temperature. For the determination of ferritin-MOF thermal behavior in solution, there are two different types of thermal gradients (linear or stepwise) that could be employed. A linear gradient allows for a faster heating rate and higher temperature resolution (i.e., increasing the temperature by 2 °C per min and collecting data every 1 °C). However, due to the continually increasing temperature, the ferritin-MOFs are never in equilibrium with the surrounding solution. There may be some lag time between the temperature increase and the onset of thermal decomposition. A stepwise gradient (i.e., increasing the sample by 5 °C and holding the temperature for 10 min before data collection) allows for equilibration of the samples at each temperature. A stepwise gradient comes at the sacrifice of temperature resolution, as the time required for steps smaller than 5 °C over broad temperature ranges would be infeasible. A stepwise gradient also allows for multiple samples to be mounted and measured at once, significantly increasing the efficiency of the experiment. For these reasons, we selected a stepwise gradient for assessing the thermal stability of ferritin-MOFs. Another important consideration is the position of the X-ray beam on the crystalline sample. The beam dimensions range from a height of 0.1 to 0.3 mm and a width of 0.2 to 0.7 mm, which is substantially smaller than the width of the capillary (ca. 1.5 mm) and the height of the crystalline sample (ca. 2 mm). Only a small fraction of the crystalline sample can be observed at once. Figure 4.2 highlights how a difference in beam position could lead to drastically different apparent thermal stability values since only the crystals in the path of the beam contribute to the diffraction pattern. While it may initially seem like the bottom of the capillary (orange) is ideal, the capillary manufacturing process results in significant variability at the terminal end of the capillary. Additionally, the largest single crystals settle to the bottom of the capillary, which

reduces the number of measured orientations causing substantial variability in the relative intensity of specific reflections. The top of the capillary is also not ideal since slight crystal movements (due to either thermal energy or the mechanical motion of the stage) can lead to a decreased sample height and loss of measured diffraction without any thermal decomposition. To address both of these concerns, we selected a beam position near the middle (green), with the beam center 0.3 mm below the top of the sample. In an attempt to control for the amount of crystalline sample, we measured the size of each quartz capillary and only used capillaries with widths from 1.4 to 1.6 mm.

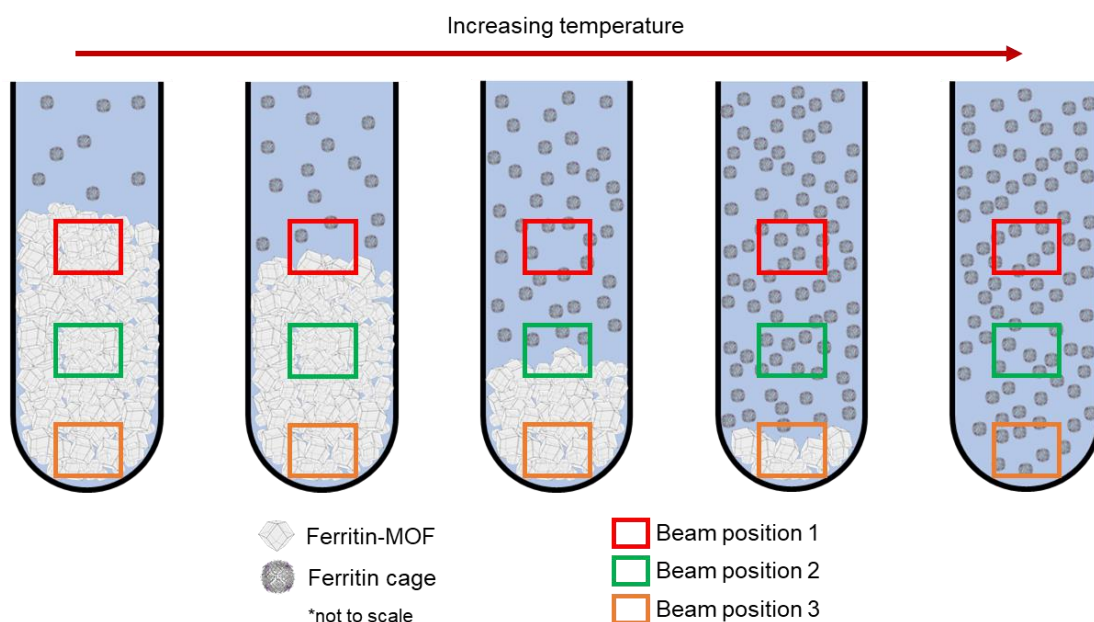


Figure 4.2. Representation of ferritin-MOF decomposition in response to increasing temperature. Three possible beam positions are shown in red, green, and orange.

4.2.2. Thermal behavior of six ferritin-MOFs

The modular construction of ferritin-MOFs afforded us with an extensive library of crystalline structures. We wanted to assess if changes in the lattice symmetry, metal, or ligand significantly impact the materials properties of ferritin-MOFs. Specifically, we choose to investigate the thermal behavior of ferritin-MOFs *in situ*. For this study, we selected the adducts of *p*-H₂bdh, *m*-H₂bdh, and 2,5-H₂fdh with both Zn- and Ni-ferritin. Except for *p*-bdh-Ni, all of the ferritin-MOFs in this series been crystallographically characterized (as described in Chapter 3). First, we examined both *p*-bdh-Zn-ferritin and *p*-bdh-Ni-ferritin at varying concentrations of PEP to assess if the precipitant would affect the lattice integrity. The impact of lattice symmetry on stability was also studied; four of the six ferritin-MOFs have *bct* (or pseudo-*bct*) symmetry, and the other two have *bcc* symmetry. The effects of altering ligand identity were compared through two isostructural pairs (*p*-bdh-Ni-ferritin / *m*-bdh-Ni-ferritin; *m*-bdh-Zn-ferritin / 2,5-fdh-Zn-ferritin). Finally, isostructural *m*-bdh-Ni-ferritin and *m*-bdh-Zn-ferritin structures allow for a direct comparison between metal ions. The ferritin-MOFs based on the other seven bridging ligands and Co-ferritin were excluded primarily due to time considerations, though the thermal behavior of these ferritin-MOFs will be assessed in future experiments.

A quick examination of the temperature-dependent SAXS profiles of *p*-bdh-Ni- and *p*-bdh-Zn-ferritin crystallized at varying levels of PEP (0-12% v/v) reveals two main trends (Figure 4.3). First, *p*-bdh-Ni-ferritin is much more stable than *p*-bdh-Zn-ferritin at all PEP concentrations. None of the *p*-bdh-Ni-ferritin samples show any signs of decomposition up to 75 °C. This upper limit was chosen to prevent the possibility of lattice collapse due to ferritin denaturation, which occurs around 80 °C. In contrast, *p*-bdh-Zn-ferritin has a $T_{\text{disassembly}}$ between 45 and 75 °C, depending on the concentration of PEP. Increasing the PEP concentration increases the apparent thermal stability of the Zn-ferritin-MOFs. This is not surprising, as PEP increases the crystal yield by decreasing the solubility of the ferritin node in solution. It is reasonable to expect that PEP would

decrease the likelihood of ferritin dissociation and solvation by a similar mechanism. A closer examination of *p*-bdh-Zn-ferritin with 0-2% PEP, reveals that the *bcc* lattice undergoes a coherent diffusionless transformation to a *bct* lattice before complete thermal decomposition. The unit cell parameters shift from $a = \text{\AA}$ to $a = \text{\AA}$ and $c = \text{\AA}$, with a tetragonal distortion of $|a-c| = \text{\AA}$. This type of cubic-to-tetragonal transformation has been observed in ferrous alloys such as martensite.¹⁷ Based on the peak intensities, <10% of the *p*-bdh-Zn-ferritin crystals undergo this crystalline transformation. However, it is unlikely that the loss in symmetry is an alternate decomposition pathway since the *bct* lattice persists at 50 °C. The dynamic behavior of *p*-bdh-Zn-ferritin is likely linked to the ultra-low intersubunit contact area of the *bcc* lattice (<0.4% of the surface area). These results also point to a non-innocent role for PEP, which disfavors the crystalline transformation at PEP concentrations $\geq 4\%$. Additional experiments are needed to understand the role of PEP in preventing the diffusionless transformation. While *p*-bdh-Ni shows no signs of decomposition regardless of PEP concentration, the peaks become significantly more broad as the concentration of PEP increases, indicating lattices that are less well ordered.

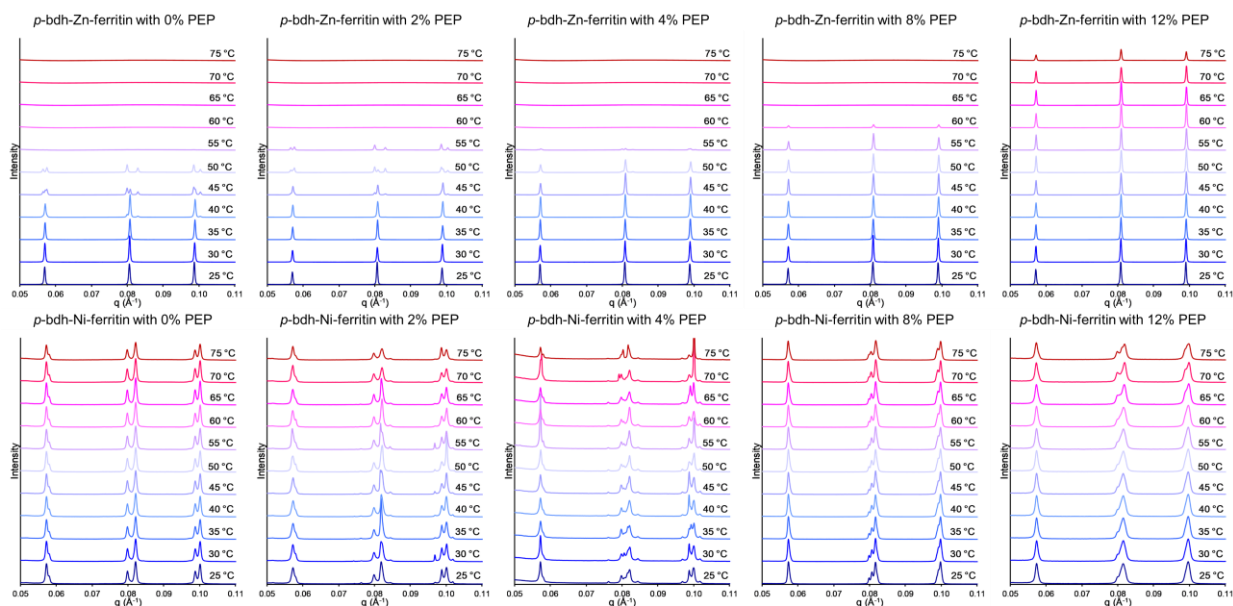


Figure 4.3. Thermal stability of *p*-bdh-Zn-ferritin and *p*-bdh-Ni-ferritin measured by temperature-dependent diffraction studies.

The *m*-bdh- and 2,5-fdh-based ferritin-MOFs exhibit a similar trend; the Zn-ferritin-MOFs begin to disorder between 50 and 55 °C, while the Ni-ferritin lattices show no loss in crystallinity up to 75 °C (Figure 4.4). We believe the inherent stability of the Ni-based lattices (when compared to the Zn-based lattices) is due to the formation of a slightly more stable metal-hydroxamate complex.¹⁸ The Ca²⁺ mediated *fcc* crystals of ferritin are also highly thermostable and show no loss in reflection intensity up to 75 °C (shown in Figure 4.5). This result further confirms that the lower $T_{\text{disassembly}}$ of the Zn-based ferritin MOFs is due to ligand-Zn interactions and not ferritin decomposition. A closer inspection of the *bcc* 2,5-fdh-Ni-ferritin diffraction pattern reveals that this lattice also undergoes a diffusionless transformation. Unlike *p*-bdh-Zn-ferritin, this transition is not suppressed by high PEP concentrations, and the crystal-to-crystal transformation happens in ca. 100% of the 2,5-fdh-Ni-ferritin crystals. The peak position shifts to higher *q* values, which corresponds to a decrease in the unit cell dimensions of the crystals. The cell parameters of the *bcc* lattice isotropically contract from $a = 156.09 \text{ \AA}$ to a smaller *bcc* lattice with cell parameters $a = 154.18 \text{ \AA}$. This change corresponds to a ca. 3.6% decrease in the volume of the unit cell and a 3.3 Å decrease in the body diagonal of the cubic lattice. This is an incredibly sharp phase transition; the lattice does not continue to contract after the initial crystal-to-crystal transformation. This dynamic behavior will be further investigated in the next section.

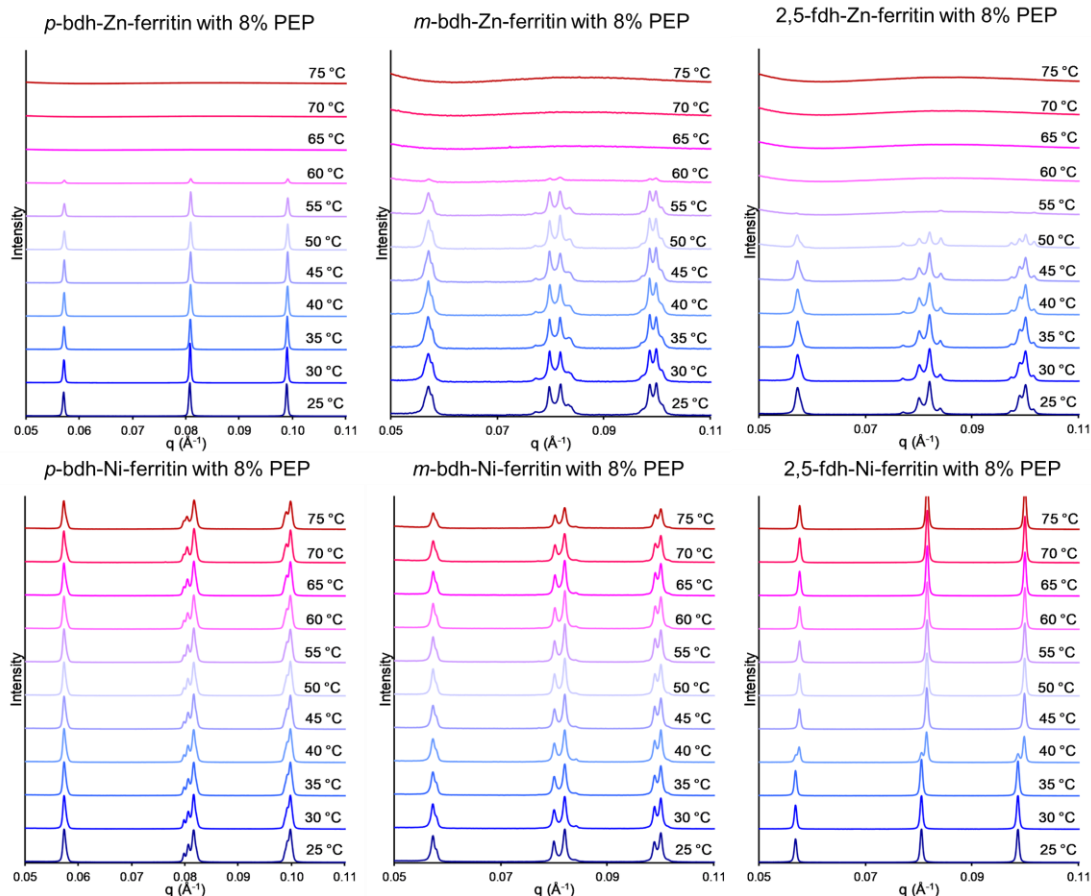


Figure 4.4. Thermal stability of six ferritin-MOFs crystallized with 8% PEP measured with temperature-dependent diffraction studies.

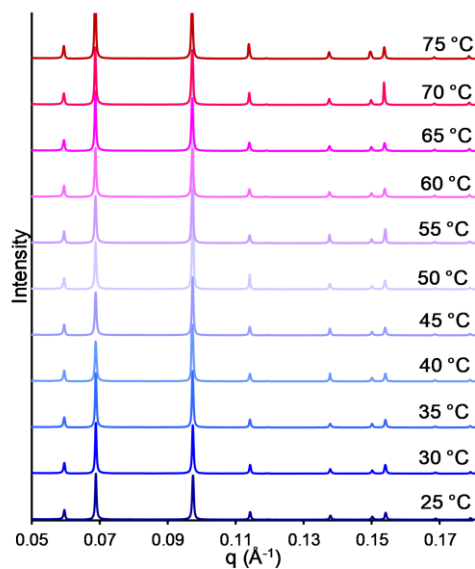


Figure 4.5. Thermal stability of the Ca^{2+} -mediated *fcc* ferritin crystals measured with temperature-dependent diffraction studies.

4.2.3. Further characterization of the thermal behavior of 2,5-fdh-Ni-ferritin

The majority of the ferritin-MOFs exhibit the expected response to thermal stress; the lattice parameters remain relatively unchanged until a critical temperature is reached that leads to liberation of free ferritin and dissolution of the crystal lattice. Both of the *bcc* ferritin-MOFs eschew this behavior and exhibit a temperature-dependent diffusionless transformation. We selected the 2,5-fdh-Ni-ferritin lattice for further study as the transition is observed in ca. 100% of the crystals (as opposed to <10% for *p*-bdh-Zn-ferritin) and the transition temperature is much lower than the $T_{\text{disassembly}}$ of the framework. To more precisely pinpoint the temperature of this phase transition, we slowly heated 2,5-fdh-Ni-ferritin crystals (stepwise, 1° every 15 min) while monitoring the unit cell parameters with X-ray diffraction. The temperature-dependent diffraction patterns are shown in Figure 4.6 (left). The phase transition from the larger *bcc* lattice ($a = 156.10 \text{ \AA}$) to the smaller *bcc* lattice ($a = 154.18 \text{ \AA}$) occurs abruptly between 33 °C and 34 °C. After the transition, the lattice parameters remain constant until at least 75 °C (as seen in Figure 4.4).

Interestingly, the 2,5-fdh-Ni-ferritin does not return to the original lattice ($a = 156.10 \text{ \AA}$) when cooled back to 33 °C. The framework remains in the contracted state ($a = 154.18 \text{ \AA}$) until the temperature is lowered below 24 °C, where it then spontaneously expands back to the initial unit cell dimensions ($a = 156.10 \text{ \AA}$). This 11 °C hysteresis is depicted by the temperature-dependent change in the unit cell parameter a in Figure 4.7. There seems to be no loss in the peak intensity of the low angle reflections after cycling between the expanded and contracted states, indicating that the diffusionless transformation is robust and completely reversible (Figure 4.8 a). While there is little change in the low angle peak intensities between the expanded and contracted state, a significant difference is observed at medium scattering angles ($q = 0.55\text{-}0.85 \text{ \AA}^{-1}$). This loss in intensity at higher temperatures indicates that the crystalline material loses short-range order during contraction (Figure 4.8 b). However, upon subsequent cooling and expansion, 2,5-fdh-Ni-ferritin regains this short-range order and returns to a nearly identical state.

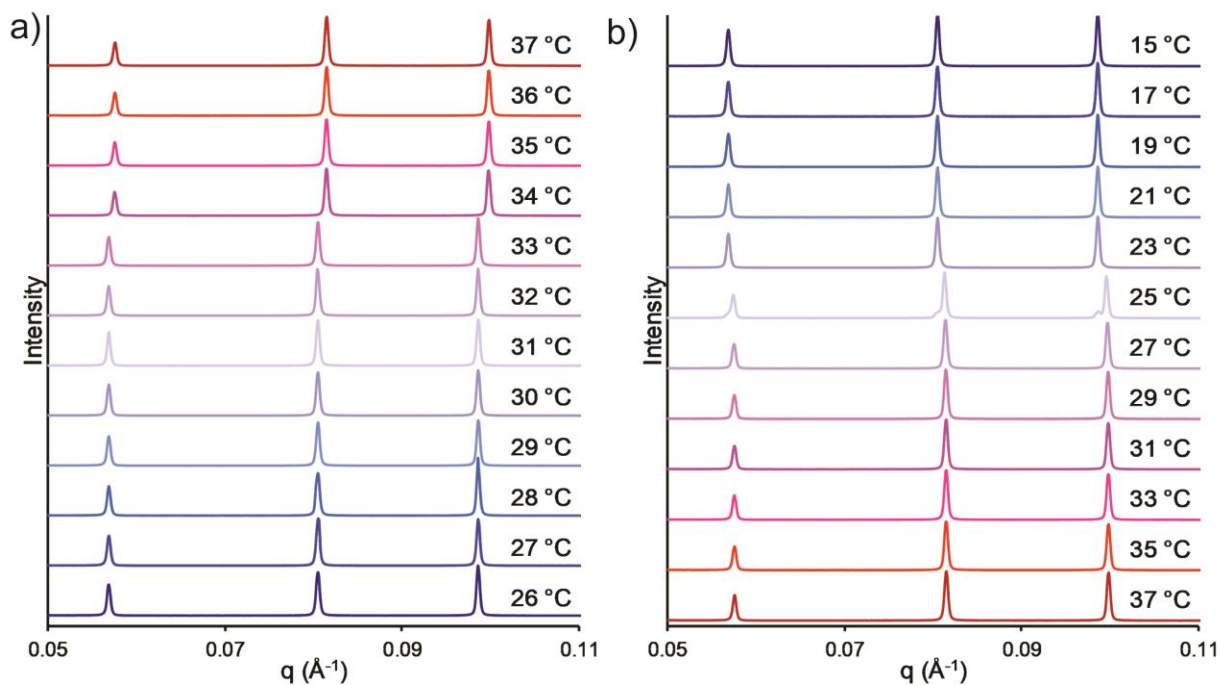


Figure 4.6. Temperature-dependent diffraction patterns of 2,5-fdh-Ni-ferritin with slow stepwise heating (a) or cooling (b). Every other diffraction pattern for b is omitted for clarity.

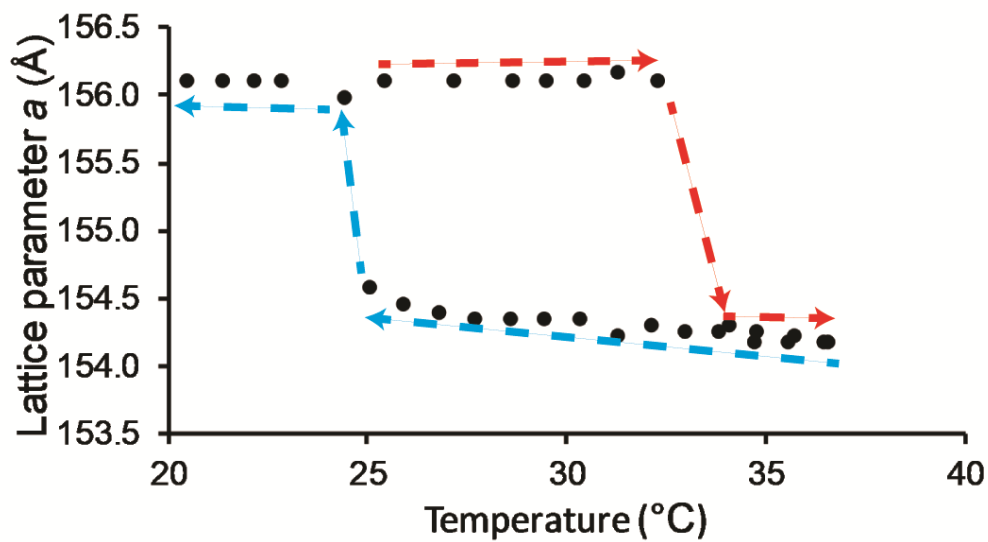


Figure 4.7. The change in unit cell size of the *bcc* 2,5-fdh-Ni-ferritin is plotted vs. temperature during slow heating (red) and cooling (blue).

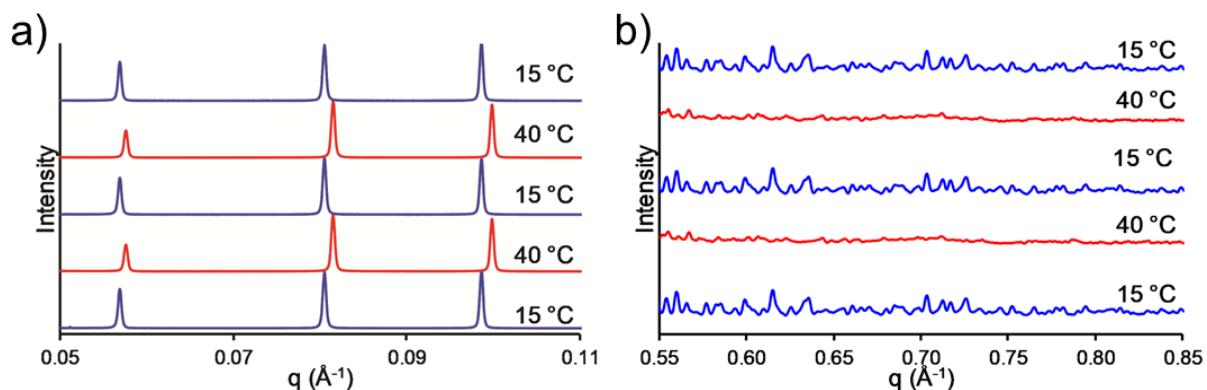


Figure 4.8. The low (a) and mid (b) angle temperature-dependent diffraction patterns of 2,5-fdh-Ni-ferritin during cycling are shown.

Since this phase transition is not observed with any of the other ferritin-MOFs, it is unlikely that the thermal transition is due to specific interactions between the ferritin and the precipitant (PEP). The yields of 2,5-fdh-Ni-ferritin in precipitant-free synthesis conditions are low, and often no crystal formation is observed. To rule out the potential of PEP to participate in the abrupt crystal-to-crystal transformation of 2,5-fdh-Ni-ferritin, we crystallized the ferritin-MOF using similar conditions with PEG 300 and PEG 4000. Using a similar stepwise thermal gradient, we characterized the diffusionless transformations of these ferritin-MOFs (Figure 4.9). The initial unit cell parameters are nearly identical to the previously measured values for both PEG 300 ($a = 156.06 \text{ \AA}$) and PEG 4000 ($a = 156.10 \text{ \AA}$). In both cases, the contracted unit cell is first observed at 29 and 30 °C for the ferritin-MOFs crystallized with PEG 300 and PEG 4000, respectively. After the initial onset, it takes an additional 3-4 °C for all of the crystals to be converted to the contracted *bcc* lattice. The expansion temperature during cooling is also depressed by ca. 3-4 °C relative to the ferritin-MOF crystallized with PEP. This slight depression in both the contraction and expansion temperature yields a nearly identical hysteresis of ca. 11 °C. These experiments confirm that the phase transition is not reliant on specific PEP-protein interactions and robustly occurs regardless of the precipitant used during crystallization.

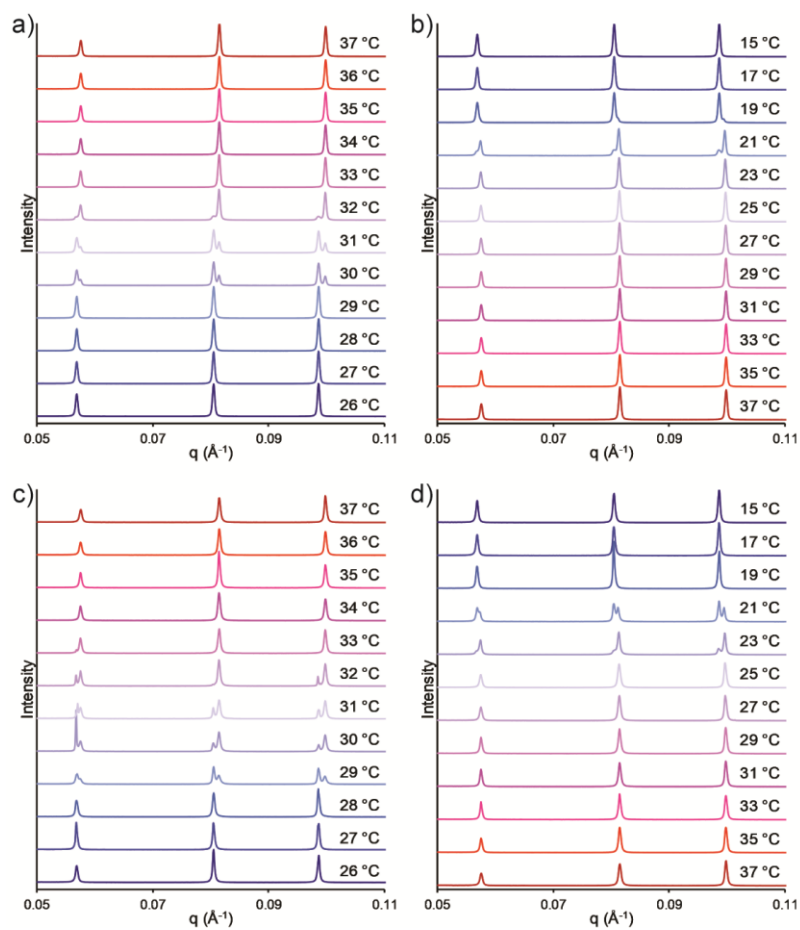


Figure 4.9 Temperature-dependent diffraction patterns of 2,5-fdh-Ni-ferritin crystallized with PEG 300 (a-b) or PEG 4000 (c-d). Every other diffraction pattern for b and d was omitted for clarity.

4.2.4. Characterization of 2,5-fdh-Ni-ferritin in solution with sc-XRD

The solution-based powder diffraction experiments confirmed that the diffusionless transformation is repeatable over many cycles, exhibits significant hysteresis, and is not dependent on the identity of the precipitant. However, these experiments do not provide us with sufficient atomic resolution to determine a possible mechanism for the crystal-to-crystal transformation. For this, we again turned to sc-XRD experiments. We wanted to avoid data collection at cryogenic conditions, as the change from room temperature to 100K has previously shown a 1-2% decrease in the unit cell parameters, which is similar to phase change at elevated temperatures. Utilizing a Bruker Microstar diffractometer equipped with Cu K α radiation (1.5418 Å), we collected three datasets on single 2,5-fdh-Ni-ferritin crystals crystallized with PEP at either 5 °C (below the transition temperature), 45 °C (above the transition temperature), and 5 °C (after a single expansion/contraction cycle). Much like many protein crystals, ferritin-MOFs begin to degrade upon solvent evacuation. Thus, all crystals were measured in solution to prevent crystal disordering due to loss of solvent. A detailed protocol can be found in the methods section.

The structure of 2,5-fdh-Ni-ferritin collected at 5 °C was solved to a resolution of 3.08 Å has unit cell dimensions of $a = 156.08 \pm 0.14$ Å, which is (1.2 Å) larger than the previously described cryogenic structure. These unit cell parameters perfectly match the unit cell dimensions determined by powder diffraction at room temperature (at room temperature), further confirming that the previously observed compressed unit cell was due to the cryogenic temperatures. Despite the difference in unit cell dimensions, the interfacial Ni-Ni distance is 10.1 Å, which is nearly identical to the 2,5-fdh-Ni-ferritin structure collected at 100K. Surprisingly, the b-factors for the protein are relatively similar to the 100K structure as well. The 2,5-fdh bridging ligand is rotationally disordered around the 3-fold crystallographic axis and gives rise to a column of electron density (shown in Figure 4.10 a). Except for the slight difference in the unit cell parameters, this structure is virtually identical to the structure discussed in Chapter 3.

The structure of 2,5-fdh-Ni-ferritin collected at 45 °C could only be solved to a resolution of 6.42 Å. Based on the MAXS data shown in Figure 4.8, we anticipated a similar drop-off in the resolution of the structure. The rate of X-ray damage is also significantly increased at this temperature and may further contribute to the lack of short-range order. Despite the poor resolution, the structure could be indexed, integrated, scaled, and phased through molecular replacement. The unit cell parameter a was 154.10 ± 0.13 Å, which is within error of the unit cell parameters determined through powder diffraction. At this resolution, only broad protein features, such as α -helices, can be resolved (Figure 4.11). The structure suggests that the interfacial distance between adjacent ferritin nodes has decreased by ca. 3 Å (Figure 4.10 b). The ^{122}His sidechains and Ni(II) are not shown in the model, as it is not possible to determine their precise location with reasonable certainty at this resolution. However, the decrease in the interfacial distance necessitates a change in the metal-ligand-protein coordination environment. Additionally, the b-factors are higher than expected, which could be due to both the dynamic crystalline transformation and the elevated temperature of the data collection.

After a single heating and cooling cycle, the structure of 2,5-fdh-Ni-ferritin was collected at 5 °C and solved to a resolution of 3.01 Å. The unit cell dimensions of $a = 156.14 \pm 0.12$ Å are within error of the initial cell parameters before expansion. The substantial increase in the resolution of the expanded *bcc* lattice (when compared to the contracted *bcc* lattice) indicates that the precise atomic order of the structure has been regained during expansion. Indeed, the crystal structure of 2,5-fdh-Ni-ferritin is nearly indistinguishable from the starting lattice. The Ni-Ni distance has slightly decreased from 10.1 Å to 9.8 Å, which is still a large enough gap to accommodate the 2,5-fdh bridging ligand.

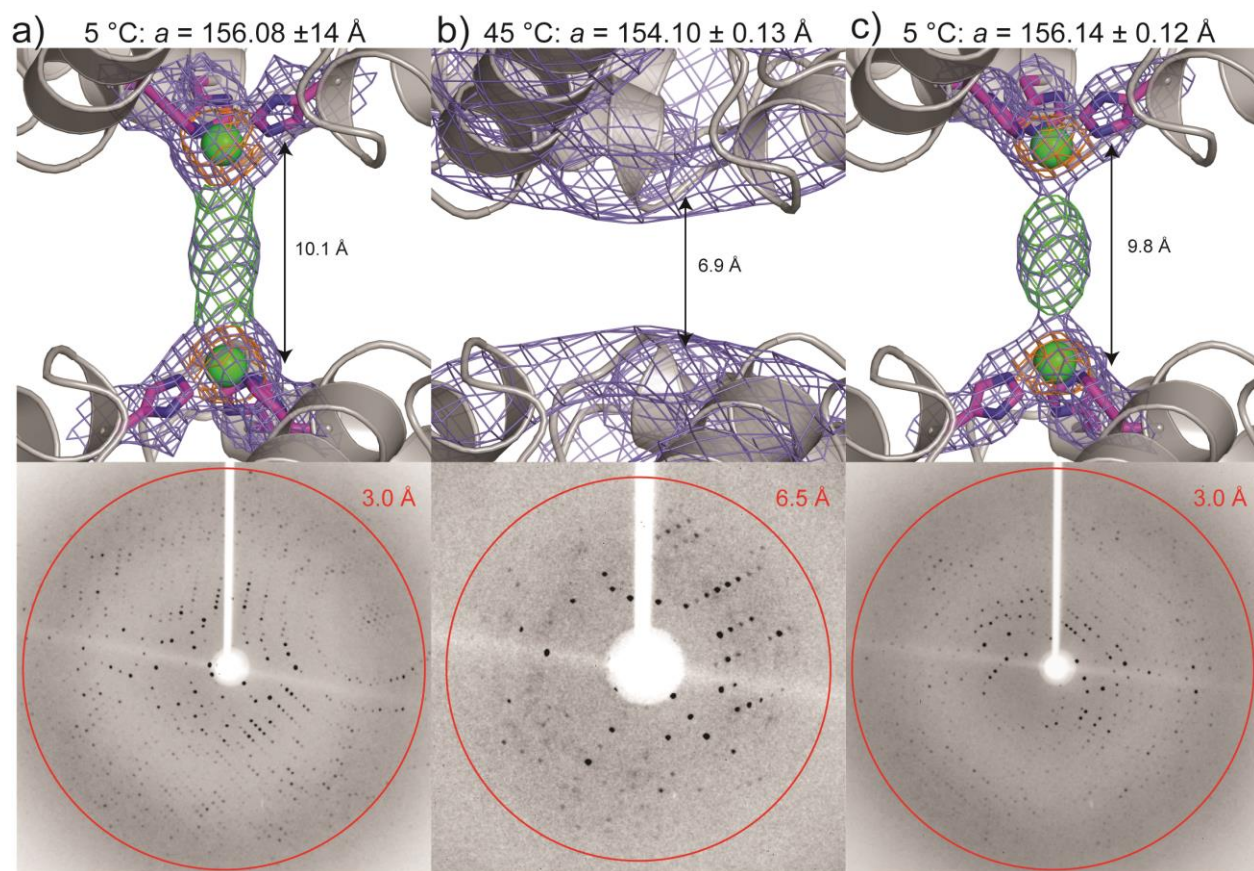


Figure 4.10. sc-XRD structure 2,5-fdh-Ni-ferritin collected at 5°C (a), 45 °C (b), and at 5°C (c) after a single heating and cooling cycle. When modeled, the ^{122}His sites are shown as magenta sticks, and Ni(II) is shown as green spheres. $2F_o - F_c$ electron density maps are contoured at 1σ (blue) and 3σ (orange). Positive $F_o - F_c$ density contoured 3σ are shown as green. A sample diffraction image for each collection is shown below each structure.

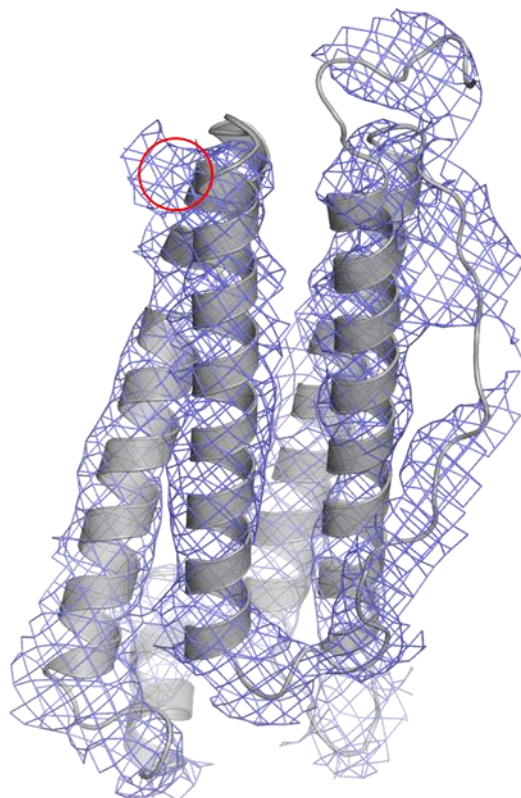


Figure 4.11. sc-XRD structure 2,5-fdh-Ni-ferritin collected at 45 °C. A single ferritin monomer is shown, $2F_o - F_c$ electron density maps are contoured at 1σ (blue). Due to the poor resolution of the structure (6.5 Å), only the rough position of the α -helices can be resolved. The approximate location of ^{122}His is highlighted with a red circle.

Research is still ongoing to uncover the mechanism behind the temperature-dependent rapid phase transition observed in 2,5-fdh-Ni-ferritin. However, the diffraction data (both PXRD and sc-XRD) allows us to draw some reasonable conclusions. First, the phase transition is unlikely to be based on a uniform contraction of the Ni-ferritin node, since this unique behavior is only observed in the *bcc* 2,5-fdh-Ni-lattices. During the isotropic contraction, the 2,5-fhd bridging ligand remains connected to both Ni coordination sites, and both metal sites remain along the crystallographic 3-fold symmetry axis. Deviation by either the ligand in metal would result in a reduction in lattice symmetry (similar to *p*-bdh-Zn-ferritin), or lead to substantial peak broadening. Based on the data collected to date, the most likely hypothesis involves a widening of the C_3 axis of ferritin coupled with the change in metal-ligand coordination geometry from octahedral to a pseudo-trigonal bipyramidal. The change in coordination could allow the Ni^{2+} to move slightly

inward towards the center of the ferritin node. However, this hypothesis only explains what may happen, not why the crystalline transformation occurs. This answer is likely closely related to the specific metal-ligand interactions that promote the assembly of a *bcc* lattice based on Ni- and Co-ferritin with 2,5-H₂fhd, but not the isostructural bridging ligand 2,5-H₂tdh. Assessing the differences in these two bent ligands is paramount to understanding the crystal-to-crystal transformation. Further experiments are still needed to elucidate the concerted protein-metal-ligand rearrangement that contracts the ferritin-MOF crystals.

4.3. Methods

4.3.1. General considerations

All chemical reagents were purchased from Sigma-Aldrich, Fisher Scientific, TCI, or an equivalent supplier. Ferritin was expressed and purified, as described in Chapter 2. The organic bridging ligands and ferritin-MOFs were synthesized, as described in Chapter 3. PXRD data were collected at synchrotron sources, either BL 4-2 (SSRL) or BL-5-ID-D (APS).

4.3.2. Calculation of the protein surface area and crystal contact area

A protein dataset originally compiled by Carugo *et al.* and later curated by Luo *et al.* was used to determine the extent of crystal contact area in protein structures.¹⁹⁻²⁰ The curated dataset contained 773 monomeric proteins with crystallographic resolution better than 2.5 Å and less than 5% heteroatom content, in addition to many other considerations. It is essential to use monomeric proteins, as it is still challenging to discern crystallographic from biological interfaces. All of the protein-protein interfaces will be crystallographic for a monomeric protein. From this dataset, only PDB ID: 1ZLM and 1B02 were excluded from further analysis. The crystal contact area and total surface area for the remaining set of 771 proteins were calculated using PISA.²¹ PISA was also used to calculate the crystal contact area for all six ferritin-MOFs characterized by sc-XRD. Since ferritin is not a monomeric protein, VOIDOO was used to estimate the total surface area on the outside of the ferritin cage.²²

4.3.3. Preparation of ferritin-MOF samples for powder diffraction at SAXS beamlines

Crystals for PXRD were prepared in 24-well culture plates or 24-well crystal trays. For the culture plates, each well contained 33.3 µL of a 25 µM ferritin solution with 50 mM CHES (pH 8.5) and 150 mM NaCl, 40 µL of a 10 mM freshly prepared ligand solution (p-H₂bdh, m-H₂bdh, or 2,5-H₂fdh) with 50 mM CHES (pH 9.5) and 150 mM NaCl, and 126.7 µL of a solution with 0.789 mM of metal (ZnCl₂ or NiCl₂), 12.6% (v/v) precipitant (PEP, PEG 300, or PEG 4000), 50 mM CHES (pH 9.5), and 150 mM NaCl. The total concentration of the solution was 4 µM ferritin, 2 mM ligand,

0.5 mM metal, 8% (v/v) precipitant, 50 mM CHES (pH 9.5), 150 mM NaCl. For the 24-well crystal trays, a 500 μ L of reservoir solution containing 0.473 mM metal (ZnCl_2 or NiCl_2), 12.6% (v/v) precipitant (PEP, PEG 300, or PEG 4000), 50 mM CHES (pH 9.5), and 150 mM NaCl was prepared. Each well contained 2 μ L of a 25 μ M ferritin solution with 50 mM CHES (pH 8.5) and 150 mM NaCl, 2.4 μ L of a 10 mM freshly prepared ligand solution (p- H_2bdh , m- H_2bdh , or 2,5- H_2fdh) and 7.6 μ L of the reservoir solution. The total concentration in the upper well was 4 μ M ferritin, 2 mM ligand, 0.3 mM metal, 8% (v/v) precipitant, 50 mM CHES (pH 9.5), 150 mM NaCl. Samples containing p- H_2bdh were also prepared with a final PEP concentration of 0%, 2%, 4%, and 12%. The wells were all sealed with tape, and crystals formed over 72-120 h before harvesting. Crystals were resuspended in solution and transferred to 1.5 mL Eppendorf tubes. Crystals from either two culture plate wells or twenty-four crystal tray wells were combined for each sample. The crystals settled to the bottom of the tubes over 1-2 hours. The crystals were then loaded into quartz capillaries with ca. 50 μ L of the crystallization solution. The capillary diameter was determined through a custom-built rack containing holes with diameters from 1 to 2 mm, and only capillaries with diameters between 1.4 and 1.6 mm were used. The crystals settled to the bottom of the capillaries overnight and were briefly centrifuged with a hand-crank centrifuge. A thin tube was inserted into each capillary and used to remove crystals until the packed crystal height was ca. 2 mm. The crystal samples were pelleted with a hand-crank centrifuge. The tubes were sealed with clay and shipped at ambient temperature to the synchrotron for analysis.

4.3.4. Temperature-dependent SAXS at synchrotron beamlines

The temperature-dependent PXRD data discussed in Section 4.3.2 were collected at BL 4-2 (SSRL). Capillaries were all loaded into a custom-built multi-capillary temperature-controlled stage. The center of the X-ray beam was positioned ca. 0.3 mm below the top of each crystalline pellet. The X-ray wavelength was 1.12709 \AA (11,000 eV). The beam center and the sample to detector distance was calibrated using an AgBeH standard. Samples were heated to 25 $^\circ\text{C}$ and

incubated at that temperature for 15 min before data collection. X-ray exposures (0.5 s) were collected for each sample position on a Rayonix MX225 CCD area detector. During each collection, the temperature was recorded by a thermocouple placed inside the thermal stage. The temperature was increased to 30 °C over 1 min, and the samples were incubated for 15 min before X-ray data collection. This process was repeated every 5 °C until 75 °C. Data were also collected on a capillary containing DI water in the thermal stage to use as a background. 1D scattering data were obtained through the azimuthal averaging of the 2D data using GSAS-II using a narrow vertical wedge (20 °) to obtain plots of scattering intensity as a function of the scattering vector q :

$$q = 4\pi \sin(\theta) / \lambda$$

where θ is one-half of the scattering angle, and λ is the wavelength of the X-rays used. It is essential to use a narrow vertical wedge at beamline 4-2 since there is significant beam broadening in the horizontal direction. Analysis of the 1D data was performed using the powder diffraction processing software JADE (MDI). Minor adjustments (such as increasing q by a factor of 10) are needed for JADE to function correctly with such small scattering angles. Plotted data are shown in Figure 4.3, Figure 4.4, and Figure 4.5 were all normalized relative to the maximum peak of the initial lattice at 25 °C.

The temperature-dependent PXRD data discussed in Section 4.3.3 were collected at BL 5-ID-D (APS). Capillaries were all loaded into a different custom-built multi-capillary temperature-controlled stage. The center of the X-ray beam was positioned ca. 0.3 mm below the top of each crystalline pellet. The X-ray wavelength was 0.7293 Å (17,000 eV). The beam center and the sample to detector distance was calibrated using a silicon diffraction grating. Samples were heated to 25 °C and incubated at that temperature for 15 min before data collection. X-ray exposures (0.5 s) were collected for each sample position on a Rayonix MX170 CCD area detector (mid-angle scattering was collected simultaneously on a different detector). During each

collection, the temperature was recorded by a thermocouple placed inside the thermal stage. The temperature was increased by 1 °C to 26 °C over 1 min, and the samples were incubated for 15 min before X-ray data collection. This process was repeated every 1 °C until 37 °C. The temperature was then decreased by 1 °C to 36 °C over 1 min, and the samples were incubated for 15 min before X-ray data collection. This process was repeated every 1 °C until 10 °C. The sample was then cycled between 15 °C and 40 °C every 15 min, and X-ray data was collected at each temperature. Data were also collected on a capillary containing DI water in the thermal stage to use as a background. 1D scattering data were obtained through the azimuthal averaging of the 2D data using GSAS-II to obtain plots of scattering intensity as a function of the scattering vector q :

$$q = 4\pi \sin(\theta) / \lambda$$

where θ is one-half of the scattering angle, and λ is the wavelength of the X-rays used. Analysis of the 1D data was performed using the powder diffraction processing software JADE (MDI). Minor adjustments (such as increasing q by a factor of 10) are needed for JADE to function correctly with such small scattering angles. The plotted data, shown in Figure 4.6 and Figure 4.8, were all normalized relative to the maximum peak of the initial lattice at 25 °C.

4.3.5. sc-XRD data collection solution at ambient and elevated temperatures

The single-crystal X-ray diffraction studies were carried out on a Bruker Microstar APEX II CCD diffractometer equipped with Cu K α radiation ($\lambda = 1.54178 \text{ \AA}$). Crystallization conditions for 2,5-fdh-NI-ferritin are described in Table 4.1. Single crystals with diameters $>100 \mu\text{m}$ were individually transferred along with 12 μL of the well solution to 100 μm glass number 50 capillaries (Hampton). The sealed end of the capillary was etched and broken so that the solution could flow through the thin capillary. Mild centrifugal force was used to force the crystal through the capillary. Since the crystal dimensions are larger than the capillary dimensions, the crystals became immobilized as the diameter of the capillary narrows. The crystals were still

surrounded on both sides by the well solution. Each capillary containing a single crystal was mounted on a four-circle goniometer with $\chi = 100^\circ$ to prevent the crystal from moving in the capillary. The temperature was maintained by a cryostream (Oxford cryosystems). X-ray data were collected with a $90^\circ \Phi$ scan with a scan width of 1° to limit X-ray damage while still collecting highly redundant data. The sample to detector distance was 60 mm unless otherwise mentioned. 2,5-fdh-Ni-ferritin #1 was collected at 5°C to limit radiation damage, and a dataset was collected with 60s exposure times. 2,5-fdh-Ni-ferritin #2 was loaded into a capillary in a warm room at 37°C to make sure the crystal remained immobilized at 45°C . The small change in volume (ca. 3.6%) is enough to allow the crystal to move in the capillary. A small set of images were captured at 5°C to confirm the unit cell before 2,5-fdh-Ni-ferritin #2 was heated to 45°C over 5 min and incubated at that temperature for 15 min. A dataset was collected with 60s exposure times and a sample to detector distance of 120 mm (due to the poor resolution). 2,5-fdh-Ni-ferritin #3 was loaded into a capillary in the warm room. The sample was cooled to 5°C by the cryostream, and a small set of images were taken to confirm the starting unit cell. 2,5-fdh-Ni-ferritin #3 was heated to 45°C over 5 min and incubated at that temperature for 15 min, and a small set of images were captured to confirm that the lattice had contracted. The sample was cooled back down to 5°C by the cryostream over 5 min and incubated at the temperature for 15 min. A dataset was collected with 60s exposure times and a sample to detector distance of 60 mm. Data were integrated with SAINT and scaled with Sadabs.²³ The detector parameters (beam center, detector distance, roll, tilt, and yaw) for 2,5-fdh-Ni-ferritin #2 were fixed during integration. These values were determined by collecting data on 2,5-fdh-Ni-ferritin at 5°C with an identical detector distance. The standard deviation of the unit cell parameters was calculated with 20 cycles of Monte Carlo simulations by SAINT. The structures of 2,5-fdh-Ni #1, #2, and #3 were solved to resolutions of 3.08 Å, 6.42 Å, and 3.01 Å, respectively. Molecular replacement was performed with Phaser²⁴ using the previously solved ^{122}Hf ferritin structure (PDB ID: 5CMQ)

as the search model. Rigid-body, positional, thermal, and TLS refinements were carried out using Phenix.²⁵ X-ray processing and refinement statistics are shown in Table 4.2.

Table 4.1. Crystallization conditions for 2,5-fdh-Ni-ferritin

2,5-fdh-Ni-ferritin	Reservoir	500 μ L total volume: 50 mM CHES (pH 10.0), 150 mM NaCl, 0.474 mM NiCl ₂ , 12.6 % v/v PEP (5/4 PO/OH)
	Sitting Drop	12.7 μ L reservoir, 3.3 μ L of 25 μ M ferritin, 4 μ L of 5 mM 2,5-H ₂ fdh in 50 mM CHES (pH 9.5) with 150 mM NaCl

Table 4.2. X-ray data processing and refinement statistics of 2,5-fdh-Ni-ferritin

	2,5-fdh-Ni-ferritin* #1	2,5-fdh-Ni-ferritin* #2	2,5-fdh-Ni-ferritin* #3
Data collection			
Space group	I432	I432	I432
Cell dimensions			
<i>a, b, c</i> (Å)	156.08	154.10	156.08
α, β, γ (°)	90	90	90
Resolution (Å)	49.36-3.08	62.91-6.42	49.38-3.01
No. unique reflections	6301	741	6738
Multiplicity	18.4 (11.3)	17.8 (20.1)	17.7 (9.4)
<i>CC1/2</i>	0.996 (0.899)	0.998 (0.827)	0.973 (0.616)
<i>Rmerge</i>	0.261 (1.186)	0.196 (1.442)	0.346 (0.699)
$\langle I / \sigma I \rangle$	9.8 (2.0)	14.3 (2.1)	11.4 (3.8)
Completeness (%)	100 (100)	99.6 (98.5)	100 (100)
Refinement			
<i>R</i> _{work} / <i>R</i> _{free}	0.203/0.245	0.260/0.313	0.182/0.257
No. atoms			
Protein	1429	1417	1437
Ligand/ion	5	None	30
Water	23	None	31
<i>B</i> -factors (Å ²)			
Protein	37.81	315.06	18.59
Ligand/ion	55.97	None	39.54
Water	33.00	None	14.74
R.m.s. deviations			
Bond lengths (Å)	0.010	0.002	0.009
Bond angles (°)	1.07	0.36	1.05
MolProbity ²⁶ score	1.77	1.60	1.40
Clashscore	3.59	5.06	6.01
Ramachandran plot (%)			
Favored	94.15	95.29	93.57
Outliers	0.59	0.59	0.00
Rotamers (%)			
Outliers	2.00	0.00	0.65

*Crystal structures are preliminary and have not been published. Refinement statistics will improve before deposition.

4.4. Acknowledgments

Chapter 4, in part, is currently being prepared for publication with contributions from the following authors: Bailey, J. B.; Tezcan, F. A. The dissertation author will be the primary author of this publication.

4.5. References

1. Pieters, B. J. G. E.; van Eldijk, M. B.; Nolte, R. J. M.; Mecnovic, J., Natural supramolecular protein assemblies. *Chemical Society Reviews* **2016**, *45* (1), 24-39.
2. Shoulders, M. D.; Raines, R. T., Collagen structure and stability. *Annu Rev Biochem* **2009**, *78*, 929-958.
3. Kluge, J. A.; Rabotyagova, O.; Leisk, G. G.; Kaplan, D. L., Spider silks and their applications. *Trends in Biotechnology* **2008**, *26* (5), 244-251.
4. Pollard, T. D.; Blanchoin, L.; Mullins, R. D., Molecular Mechanisms Controlling Actin Filament Dynamics in Nonmuscle Cells. *Annual Review of Biophysics and Biomolecular Structure* **2000**, *29* (1), 545-576.
5. Brouhard, G. J.; Rice, L. M., Microtubule dynamics: an interplay of biochemistry and mechanics. *Nature Reviews Molecular Cell Biology* **2018**, *19* (7), 451-463.
6. Gonen, S.; DiMaio, F.; Gonen, T.; Baker, D., Design of ordered two-dimensional arrays mediated by noncovalent protein-protein interfaces. *Science* **2015**, *348* (6241), 1365-1368.
7. Sinclair, J. C.; Davies, K. M.; Venien-Bryan, C.; Noble, M. E. M., Generation of protein lattices by fusing proteins with matching rotational symmetry. *Nature Nanotechnology* **2011**, *6* (9), 558-562.
8. Sakai, F.; Yang, G.; Weiss, M. S.; Liu, Y.; Chen, G.; Jiang, M., Protein crystalline frameworks with controllable interpenetration directed by dual supramolecular interactions. *Nature Communications* **2014**, *5*.
9. Brodin, J. D.; Ambroggio, X. I.; Tang, C. Y.; Parent, K. N.; Baker, T. S.; Tezcan, F. A., Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. *Nature Chemistry* **2012**, *4* (5), 375-382.
10. Suzuki, Y.; Cardone, G.; Restrepo, D.; Zavattieri, P. D.; Baker, T. S.; Tezcan, F. A., Self-assembly of coherently dynamic, auxetic, two-dimensional protein crystals. *Nature* **2016**, *533* (7603), 369-+.
11. Liljestrom, V.; Mikkila, J.; Kostianen, M. A., Self-assembly and modular functionalization of three-dimensional crystals from oppositely charged proteins. *Nature Communications* **2014**, *5*.

12. Alberstein, R.; Suzuki, Y.; Paesani, F.; Tezcan, F. A., Engineering the entropy-driven free-energy landscape of a dynamic nanoporous protein assembly. *Nature Chemistry* **2018**, *10* (7), 732-739.
13. Sontz, P. A.; Bailey, J. B.; Ahn, S.; Tezcan, F. A., A metal organic framework with spherical protein nodes: rational chemical design of 3D protein crystals. *Journal of the American Chemical Society* **2015**, *137* (36), 11598-11601.
14. Bailey, J. B.; Zhang, L.; Chiong, J. A.; Ahn, S.; Tezcan, F. A., Synthetic Modularity of Protein-Metal-Organic Frameworks. *Journal of the American Chemical Society* **2017**, *139* (24), 8160-8166.
15. Farha, O. K.; Eryazici, I.; Jeong, N. C.; Hauser, B. G.; Wilmer, C. E.; Sarjeant, A. A.; Snurr, R. Q.; Nguyen, S. T.; Yazaydin, A. Ö.; Hupp, J. T., Metal–Organic Framework Materials with Ultrahigh Surface Areas: Is the Sky the Limit? *Journal of the American Chemical Society* **2012**, *134* (36), 15016-15021.
16. Hönicke, I. M.; Senkovska, I.; Bon, V.; Baburin, I. A.; Bönisch, N.; Raschke, S.; Evans, J. D.; Kaskel, S., Balancing Mechanical Stability and Ultrahigh Porosity in Crystalline Framework Materials. *Angew. Chem. Int. Ed. Engl.* **2018**, *57* (42), 13780-13783.
17. Chen, Y.; Xiao, W.; Jiao, K.; Ping, D.; Xu, H.; Zhao, X.; Wang, Y., Cubic martensite in high carbon steel. *Physical Review Materials* **2018**, *2* (5), 050601.
18. Agrawal, Y. K.; Tandon, S. G., Metal-ligand stability constants of hydroxamic acids. *Journal of Inorganic and Nuclear Chemistry* **1974**, *36* (4), 869-873.
19. Carugo, O.; Djinović-Carugo, K., How many packing contacts are observed in protein crystals? *Journal of Structural Biology* **2012**, *180* (1), 96-100.
20. Luo, J.; Liu, Z.; Guo, Y.; Li, M., A structural dissection of large protein-protein crystal packing contacts. *Scientific Reports* **2015**, *5* (1), 14214.
21. Krissinel, E.; Henrick, K., Inference of Macromolecular Assemblies from Crystalline State. *Journal of Molecular Biology* **2007**, *372* (3), 774-797.
22. Kleywegt, G. J.; Jones, T. A., Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Crystallographica Section D* **1994**, *50* (2), 178-185.
23. Krause, L.; Herbst-Irmer, R.; Sheldrick, G. M.; Stalke, D., Comparison of silver and molybdenum microfocus X-ray sources for single-crystal structure determination. *Journal of applied crystallography* **2015**, *48* (1), 3-10.
24. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phaser crystallographic software. *Journal of applied crystallography* **2007**, *40* (4), 658-674.
25. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66* (Pt 2), 213-21.

26. Chen, V. B.; Arendall, W. B. I.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C., MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 12-21.

Chapter 5. Concluding remarks and future directions

5.1. Concluding remarks

This work has discussed the synthesis, characterization, and unusual dynamic behavior of an extensive library of 3D, crystalline protein-MOFs. All of these protein-MOFs were constructed from a single symmetric protein, ferritin, that was engineered to display outward-facing tris-His metal coordination sites at the C_3 vertices through the T122H mutation. $^{122}\text{Hferritin}$ stably anchors a series of late first-row transiting metals (Zn^{2+} , Ni^{2+} and Co^{2+}). These metal-anchoring sites provide the basis for modular ligand-directed self-assembly with a library of ditopic dihydroxamate bridging ligands. The protein-MOFs rapidly form rhombic dodecahedron-shaped crystals upon the addition of all of the components, often crystallizing in >24h. The symmetry of the ligand and the coordination preferences of the metal all influence the 3D packing of the protein-MOF. Despite the significant mismatch between the size of the protein node (120 Å across, >505,000 g/mol) and bridging ligands (ca. 7-15 Å long, 146-246 g/mol), the ferritin-MOFs are highly crystalline and can be readily characterized by sc-XRD and PXRD (SAXS). These MOF crystals predominantly pack into either *bct* or *bcc* lattices. The *bcc* ferritin-MOFs do not form any protein-protein contacts (determined by sc-XRD); the frameworks are held together solely by the metal-hydroxamate interactions. The *bct* and pseudo-*bct* ferritin-MOFs have a few interfacial side chains in close enough proximity to form electrostatic interactions between ferritin nodes. However, these interactions are unlikely to alter the packing substantially. The lattice symmetry is still dictated by the metal and ligand.

The thermal stability of the ferritin-MOFs is directly linked to the strength of the metal-ligand bonds (i.e., the Ni-ferritin MOFs are more stable than the Zn-ferritin MOFs). Additionally, incredibly small crystal-contact area of the *bcc* ferritin-MOFs enables temperature-dependent dynamic lattice rearrangements. The *bct* and pseudo-*bct* ferritin-MOFs have much larger interfacial contact areas and do not display the same dynamic behavior. Overall, the modular

construction of these protein-MOFs enabled precise control over the lattice symmetry, unit cell dimensions, thermal stability, and allowed for the emergence of dynamic crystal-to-crystal transformations.

Further expanding the versatility of protein-MOFs requires increasing the scope and functionality of both the bridging ligands and the protein scaffold. All of the synthesized protein-MOFs presented in this work have been based on ^{122}Hf ferritin connected by small and relatively inflexible organic ligands. We envision that future studies will include three critical areas. First, the successful engineering of the ^{122}Hf ferritin nodes can serve as a guide to expand the library of protein nodes to include protein scaffolds with different sizes and symmetries. Second, replacing the relatively rigid ligands with stimuli-responsive organic molecules would be an added control over the dynamic responsiveness of the crystalline scaffolds. Third, the size discrepancy of the protein node and bridging ligands needs to be addressed. Expanding the scope of bridging ligands to include size-comparable biomolecules (such as coiled-coil peptides, short DNA sequences, or even symmetric proteins) would dramatically increase the breadth of protein-MOFs. This chapter will discuss each one of these avenues in detail and highlight some early progress towards these goals.

5.2. Future directions

5.2.1. Design of tetrahedral protein nodes

This work explored the synthesis and behavior of protein-MOFs based exclusively on ferritin nodes. While this is a testament to the interchangeability of the other two components (ligand and metal), utilizing protein scaffolds with different symmetries would dramatically increase the accessible lattice arrangements. From a design perspective, the C_3 axis of human heavy-chain ferritin (HuHF) is perfect for node construction. It was incredibly fortuitous that HuHF was selected for the development of the first protein-MOFs. As previously described, the ^{122}His mutation forms a highly robust tris-His coordination motif that anchors many transition metals. Importantly, the site is slightly recessed (to prevent metal-mediated protein-protein dimerization), but close enough to the surface for the coordination of small organic bridging ligands. Tripodal coordination motifs possess a series of advantages over both two- and four-coordinate metal-binding sites. Tripodal sites bias the position of the metal to the centroid of the tripod and only occupy one face of the coordinated metal ions. This leaves multiple open-coordination sites for the binding of chelating ligands and imposes directionality. Symmetric four-coordinate sites tend to coordinate in a planar geometry, leaving only one available open-coordination site per side, making them incompatible with chelating ligands. Most monodentate bridging ligands do not have the binding affinity required to outcompete bulk solvent and drive ordered ligand-based assembly. The metal positioning in two-coordinate metal-chelating motifs is less concrete, and there is no enforced directionality for ligand binding. The binding affinity for a two-coordinate site will also be much lower, increasing the likelihood that a chelating bridging ligand will strip the metal out of the protein scaffold.

The C_3 axes of symmetrical proteins are the most convenient places to install tripodal coordination sites due to the inherent symmetry. It is essential to mention that outward-facing tripodal motifs can also be installed on C_2 axes (which would lead to two tripodal metal-binding

sites per axis) or on nonsymmetric protein faces. The design of either of these would be considerably more complex and require the perfect protein scaffold to build tripodal metal-binding sites amenable for protein-MOF synthesis. Since the C_3 axes of the protein node also need to extend in three dimensions, C_3 and D_3 symmetric proteins can be excluded. This narrows the search to protein oligomers that possess octahedral (O), tetrahedral (T), or icosahedral (I) symmetry. Since ^{122}Hf ferritin has O symmetry, and the formation of protein-MOFs from virus particles is a daunting task, targeting proteins with T symmetry is the most attractive for accessing protein-MOFs with new lattice symmetries. Unlike the O proteins, which are 24-meric in nature and have one distinct C_3 axis, T proteins can be comprised of either twelve or twenty-four subunits and have two different C_3 axes. The protein data bank (rcsb.org) is an excellent resource for identifying appropriate protein scaffolds for node construction. Figure 5.1 shows a series of eight structurally unique dodecameric tetrahedral proteins with their C_3 axes highlighted in red and blue. Out of the sixteen 3-fold axes shown, *S. Aureus* EpiD (red), *C. Albicans* DHQase (red), and *P. aeruginosa* M18 (blue) can all be immediately discarded, since these pores are too open to support a tripodal tris-His site. Unlike ^{122}Hf ferritin, which has a slightly angled alpha helix (ca. 60°), which angles the α -C's to support a tris-His site, many of the other axes are planar complicating incorporation of a tripodal metal-binding site. The *B. Licheniformis* Anti-TRAP (blue), *H. halophila* dodecin (blue), *C. Albicans* DHQase (red), *S. Aureus* EpiD (blue), *B. burgdorferi* NapA (red), *A. tumefaciens* DPS (red) and *Designed* T310 (red and blue) scaffolds can be discarded for this reason. Of the five remaining potential nodes, only *H. halophila* dodecin (red) appears to have the appropriate geometry to satisfy a tris-His metal coordination site (Figure 5.2). This quick exercise highlights the incompatibility of many tetrahedral proteins for node construction.

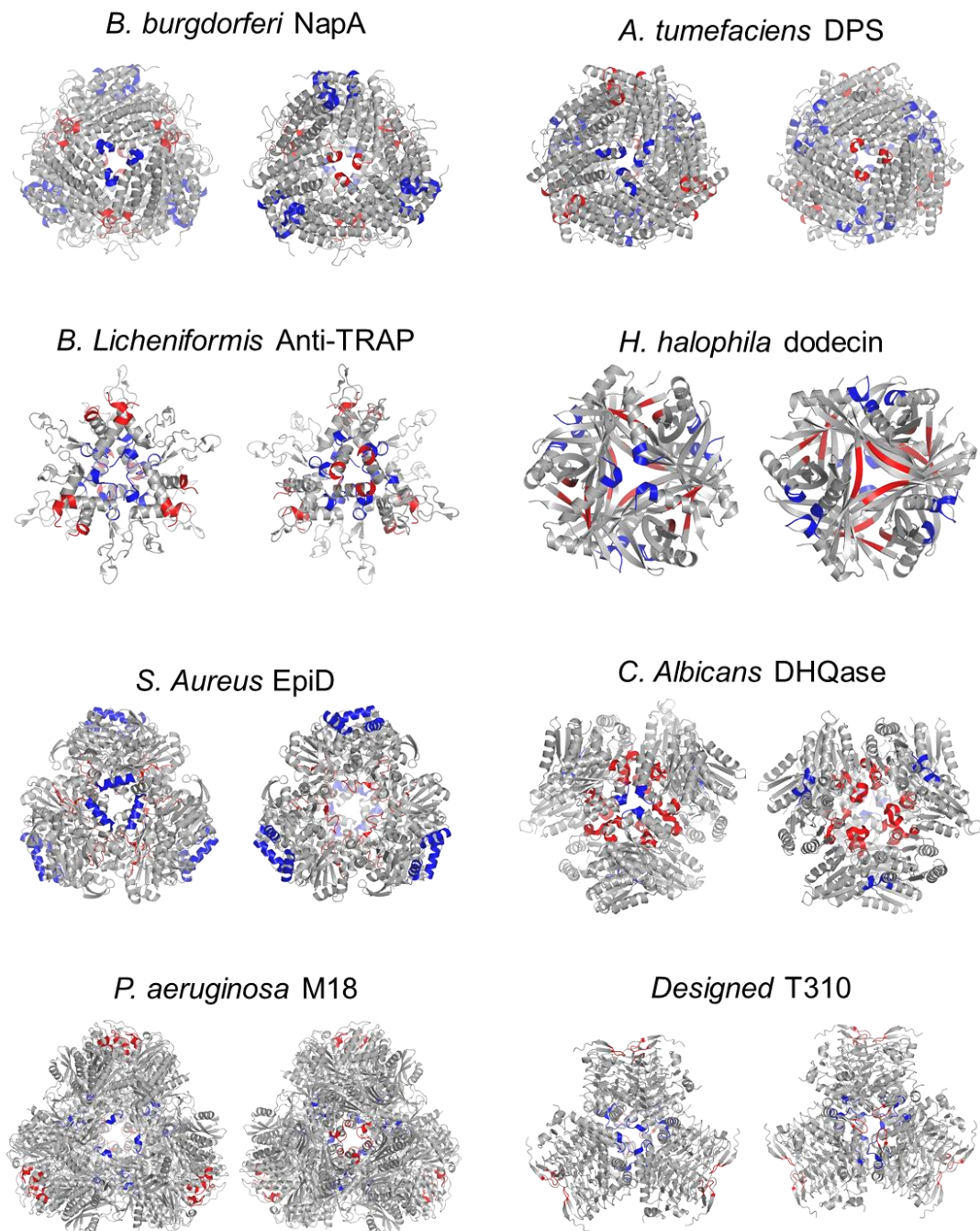


Figure 5.1. Possible scaffolds for a tetrahedral protein node. C_3 axes for the selected 12-mer tetrahedral proteins are shown in blue and red. The selected proteins are *B. burgdorferi* neutrophil-activating protein A (PDB ID: 2PYB); *A. tumefaciens* DNA-binding protein from starved cells (PDB ID: 1O9R); *B. licheniformis* Anti-trp RNA-binding attenuation protein (PDB ID: 3LCZ); *H. halophila* dodecin (PDB ID: 2VXA); *S. Aureus* epidermin biosynthesis protein (PDB ID: 3QJG); *C. albicans* type-II 3-dehydroquinase (PDB ID: 3KIP); *P. aeruginosa* M18 aminopeptidase (PDB ID: 3WT4); Computationally Designed Self-assembling tetrahedron protein T310 (PDB ID: 4EGG).

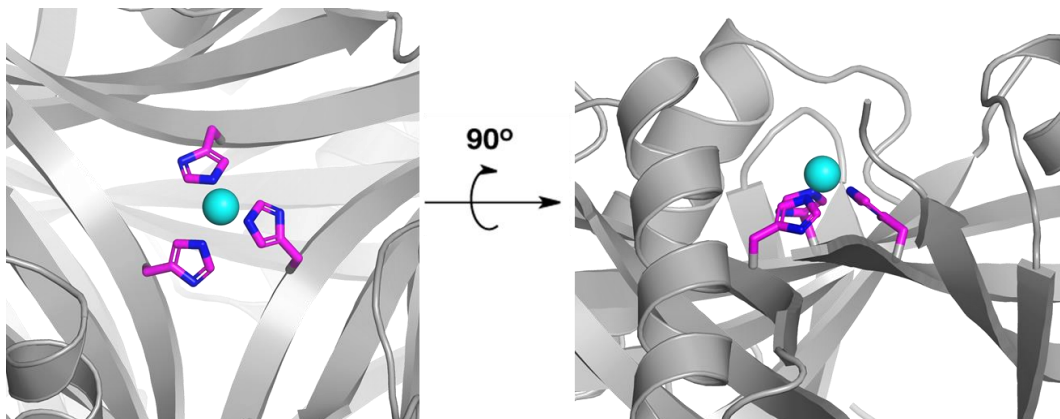


Figure 5.2. The hypothetical tris-his binding site on *H. halophila* dodecin installed through the L9H mutation. The sidechain of the modeled ^9His is shown in magenta, and the metal is shown in teal. This structure designed *in silico* (PDB ID: 2VXA).

Instead of carefully selecting a new protein scaffold that can accommodate a tripodal tris-His metal-binding site, the previously designed ^{122}Hf ferritin site (referred to ^{122}H HuHF in this section to avoid confusion with other ferritins) can be repurposed on tetrahedral ferritin from a different organism. *T. ni* ferritin, which has been crystallographically characterized, is one possible protein.¹ Most organisms that have both light- and heavy-chain ferritin form mixed oligomers based on expression levels. *T. ni* ferritin uniquely forms a stoichiometric 24-meric hetero-oligomer due to a specific interchain disulfide bond between ^3Cys of the heavy-chain and ^{12}Cys of the light-chain (Figure 5.3 a). This leads to two different C_3 axes, comprised entirely of either heavy-chain (TniHF) or light-chain (TniLF). Except for the extended N-terminus, the overall topology of *T. Ni* ferritin is very similar to HuHF, with both chains containing a four-helix bundle and a 5th short helix that extends into the C_4 axis. The C_3 axes of ^{122}H HuHF, wild-type TniHF, and wild-type TniLF are shown in Figure 5.3 b-d. Both TniHF and TniLF have an extended loop that protrudes from the 3-fold axis. This loop could be problematic during protein-MOF formation, as it would require excessively long bridging ligands. The entire C_3 loop was redesigned *in silico* to match the ^{122}H HuHF. A homology model was generated using the I-tasser server for both the remodeled ^{132}H TniHF and ^{149}H TniLF sequences.² The top-scoring result for each variant is shown in Figure 5.4 overlaid with the crystal structure for ^{122}H HuHF. As expected, both of the remodeled ^{132}H TniHF

and ^{149}H TniLF C_3 axes look nearly identical to ^{122}H HuHF. These two *T. ni* variants can be selectively combined to yield three distinct nodes. ^{132}H TniHF + ^{149}H TniLF has eight C_3 axes with tris-His motifs and should assemble into body-centered lattices (similar to ^{122}H HuHF) when combined with ligand and metal. The combination of ^{132}H TniHF + WT-TniLF or ^{149}H TniLF + WT-TniHF will only have four C_3 axes with tris-His motifs arranged in a tetrahedral geometry (Figure 5.4 c). This arrangement would yield highly-porous diamond-like lattices, making the redesigned *T. Ni* ferritin one of the most promising tetrahedral scaffolds.

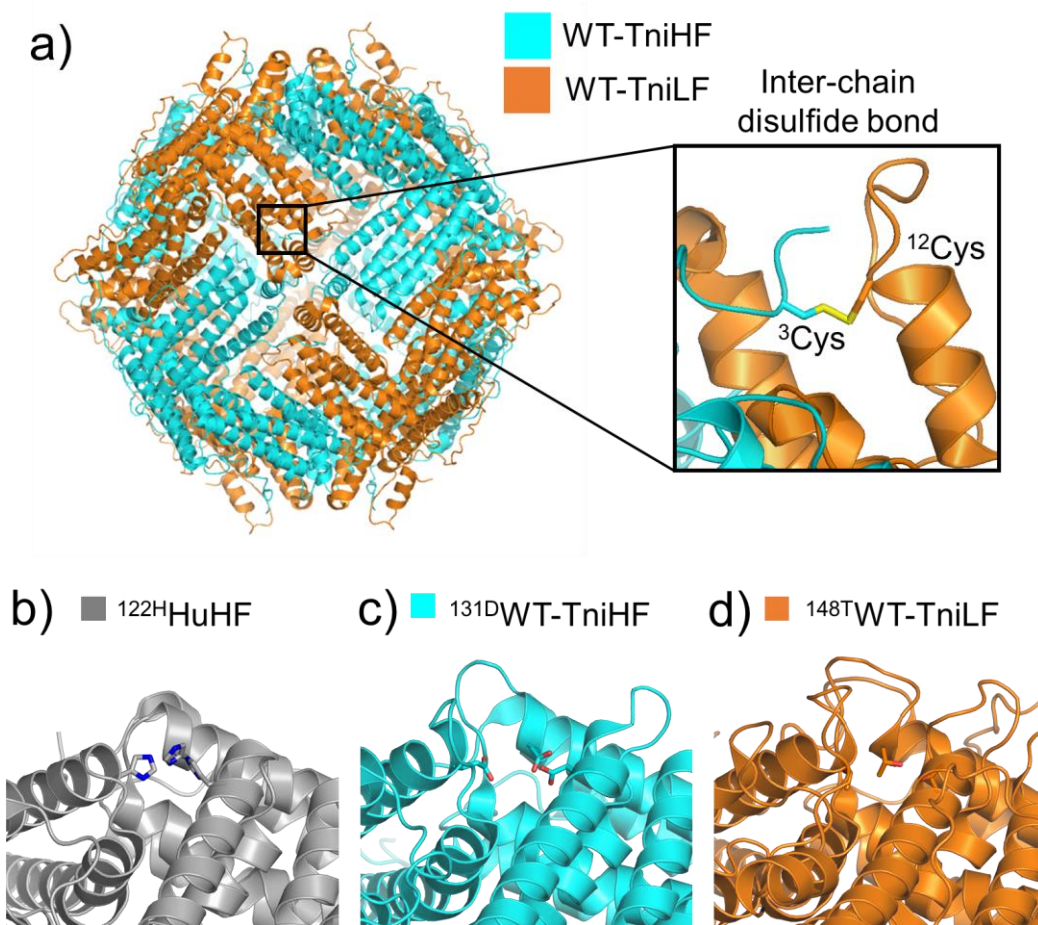


Figure 5.3. Structure of *T. ni* ferritin. a) An interchain disulfide bond leads to the formation of a stoichiometric tetrameric hetero-oligomer for *T. Ni* ferritin. The C_3 axes for ^{122}H HuHF (b) wild-type TniHF (c) and wild-type TniLF (d) are shown.

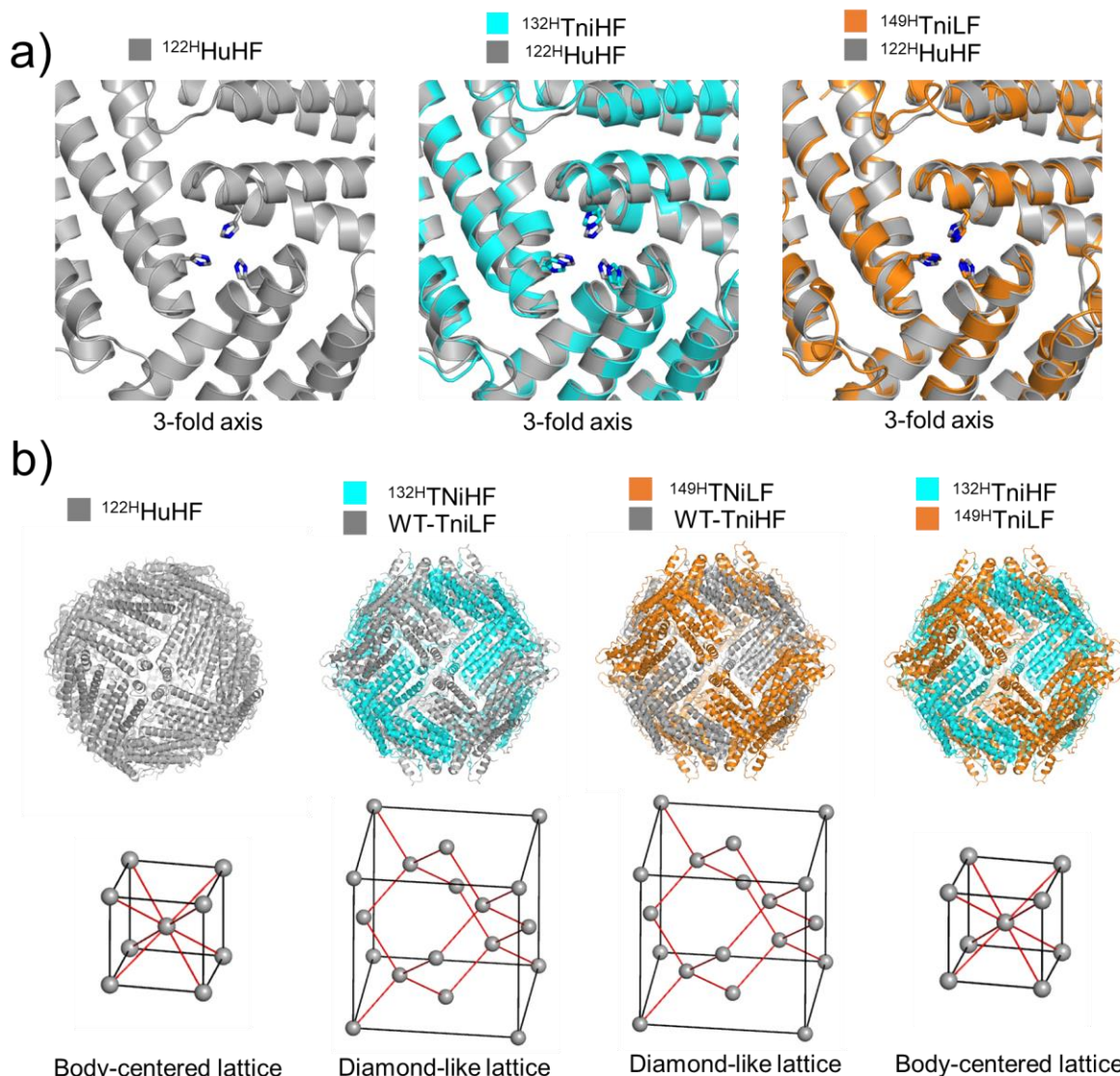


Figure 5.4. Possible lattice symmetries for ferritin-MOFs based on *T. Ni* variants. a) The C_3 axes for ^{122}H -HuHF (left) overlaid with ^{132}H -TniHF (middle) and ^{149}H -TniLF (right) are shown. b) The expected protein-MOF lattice symmetry is shown for all variants.

While the redesigned *T. ni* ferritin offers can form both body-centered and diamond-like lattices, the variants will be challenging to obtain. When separately expressed in *E. coli*, all of the *T. ni* ferritin variants form insoluble aggregates and end up in the inclusion bodies. This behavior is typical of ferritin monomers that do not assemble into a cage *in vivo*.³ The TniHF and TniLF variants need to be reconstituted under mildly reducing conditions to reform the interchain disulfide bond. Refolding and reassembling this tetrahedral ferritin will likely be very challenging

and take substantial screening efforts. Fortunately, there is a more straightforward path towards a tetrahedral ferritin node. Open-pore ferritin variants, such as *A. fulgidus* ferritin (AfF), are 24-meric tetrahedral homo-oligomers with two distinct C_3 axes. One of the 3-fold axes is the “open-pore,” which, as the name suggests, is a large opening that is not suitable for installing a metal-chelating motif (Figure 5.5 a). However, the other C_3 pore of the wild-type AfF is nearly identical to ^{122}H uHF. Figure 5.5 b shows a side-by-side comparison of the C_3 axes of ^{122}H uHF and ^{111}H AfF (the Met111His mutation was modeled *in silico*). The proposed tris-His site of ^{111}H AfF has a slightly longer distance between the α -C's than ^{122}H uHF (11.7 Å vs. 10.5 Å) as well as a longer ϵ -N distance (4.6 Å vs. 3.6 Å). However, this chelating motif should still coordinate transition metal ions. ^{111}H AfF is the most easily accessible protein-node for the construction of diamond-like protein-MOFs with diamond-like lattice arrangements.

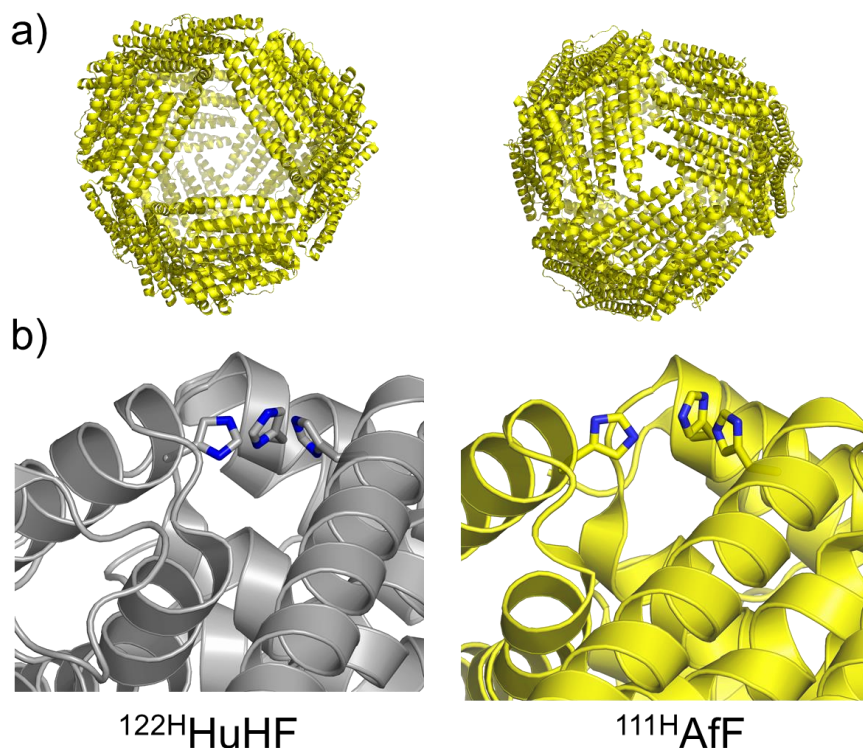


Figure 5.5. A potential tris-His metal chelation sites on AfF. a) Models of the open-pore (left) and closed-pore (right) of AfF are shown (PDB ID: 1SQ3). b) the C_3 axes of ^{122}H uHF (left) and ^{111}H AfF (right) are shown. The ^{111}H AfF mutation was modeled *in silico*.

5.2.2. Design of inherently functional bridging ligands

The series of dihydroxamic bridging ligands used to synthesize the ferritin-MOFs discussed in this work all contained rigid planar spacers. It should be possible to further expand the scope of protein-MOFs by introducing ligands with inherent functionality. To this end, we selected a fluorescent BODIPY dye. A dicarboxylic acid version of this ligand has been previously employed to form traditional MOFs⁴. Figure 5.6 shows the proposed synthetic scheme for BODIPY dihydroxamic acid bridging ligand. The BODIPY dihydroxamate bridging ligand is collinear, with a hydroxamate-hydroxamate spacing of ca. 15 Å, which is a 4 Å increase from *p*-bdh. The BODIPY-based framework would likely be isorecticular with the *p*-bdh series of ferritin-MOFs, forming Cubic lattices with Zn, tetragonal lattices with Ni, and a mixture of both with Co. Unlike 4,4'-bpth, which had poor solubility in aqueous conditions, the charged aromatic scaffold of BODIPY enhances the solubility of the complex. Using BODIPY as a fluorescent handle, it will be possible to monitor ligand loss during ferritin-MOF decomposition, depletion of ligand during crystallization, and explore the possibility of ligand exchange *in crystallo* with other bridging ligands.

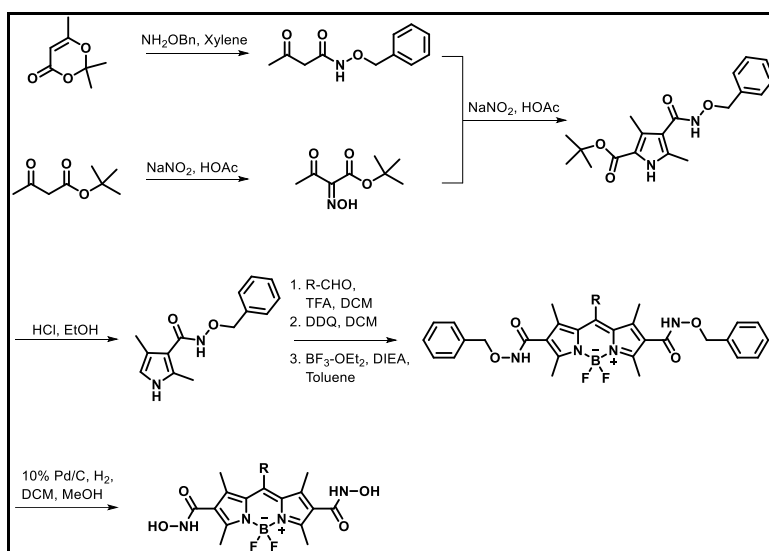


Figure 5.6. The synthetic scheme for the synthesis of a dihydroxamate bridging ligand with a BODIPY spacer.

Utilizing inherently dynamic ligands, such as azobenzenes, should enable crystalline transformations in response to external stimuli. Azobenzene ligands undergo a *trans* to *cis* isomerization upon UV irradiation (320-380 nm).⁵ This process is reversible; the azobenzene will revert to the more stable *trans* conformation upon exposure to visible light (380-450 nm). Figure 5.7 a highlights two possible routes toward the synthesis of the dicarboxylate azobenzene. These carboxylic acids can be converted to the hydroxamic acids through amidation with *O*-tritylhydroxylamine, followed by deprotection with trifluoroacetic acid. This method can be used to synthesize four azobenzene derivatives (Figure 5.7 b). The four proposed photoswitchable azobenzene ligands can be divided into two categories based on connectivity. The *trans*-azobenzene hydroxamates (left), are parallel-offset and should crystallize into a *bct* lattice regardless of metal, similar to the *E*-*edh*- and 2,6-*ndh*-based frameworks. The *trans*-azobenzene with a methylene spacer should behave similarly to *p*-*xdh*-based ferritin MOFs, though the specific lattice orientation is more challenging to predict due to the rotatable bonds. Sulfonated versions of these ligands can also be prepared, since the unsubstituted dihydroxamate azobenzenes may not be soluble enough to interact with the protein nodes in aqueous solutions. The sulfonate is placed far enough away such that it should not interfere with hydroxamate binding. Due to the flexible coordination of the hydroxamates, it should be possible for the azobenzene bridging ligands to remain bound in the *cis* conformation. UV irradiation of azobenzene containing ferritin-MOFs could lead to a contraction of the framework (Figure 5.8), potentially allowing for unprecedented control of a dynamic crystal-to-crystal transformation of a protein-MOF.

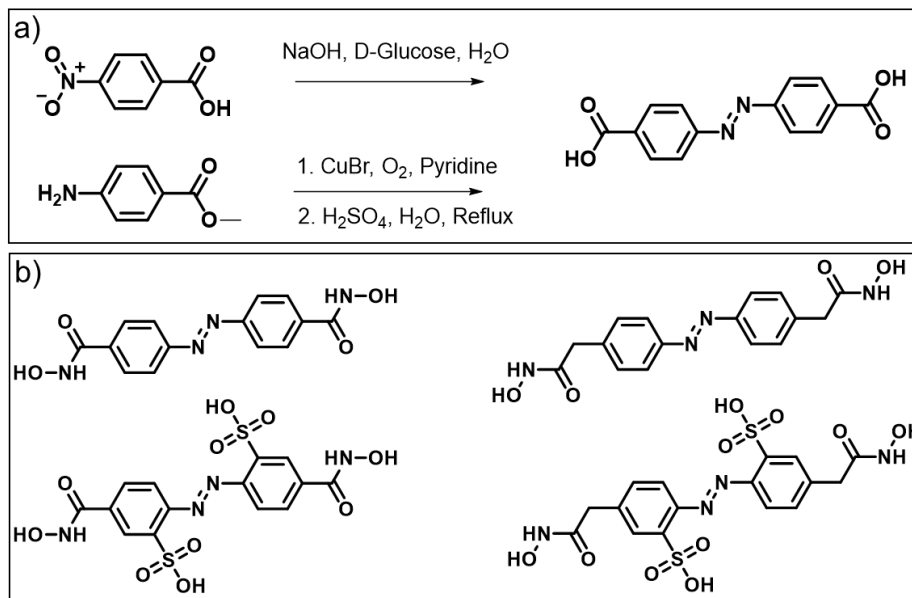


Figure 5.7. Synthetic schemes for azobenzene-based bridging ligands. a) Two proposed synthetic schemes for the synthesis of dicarboxylate azobenzene ligands. b) Four dihydroxamate azobenzene bridging ligands are shown.

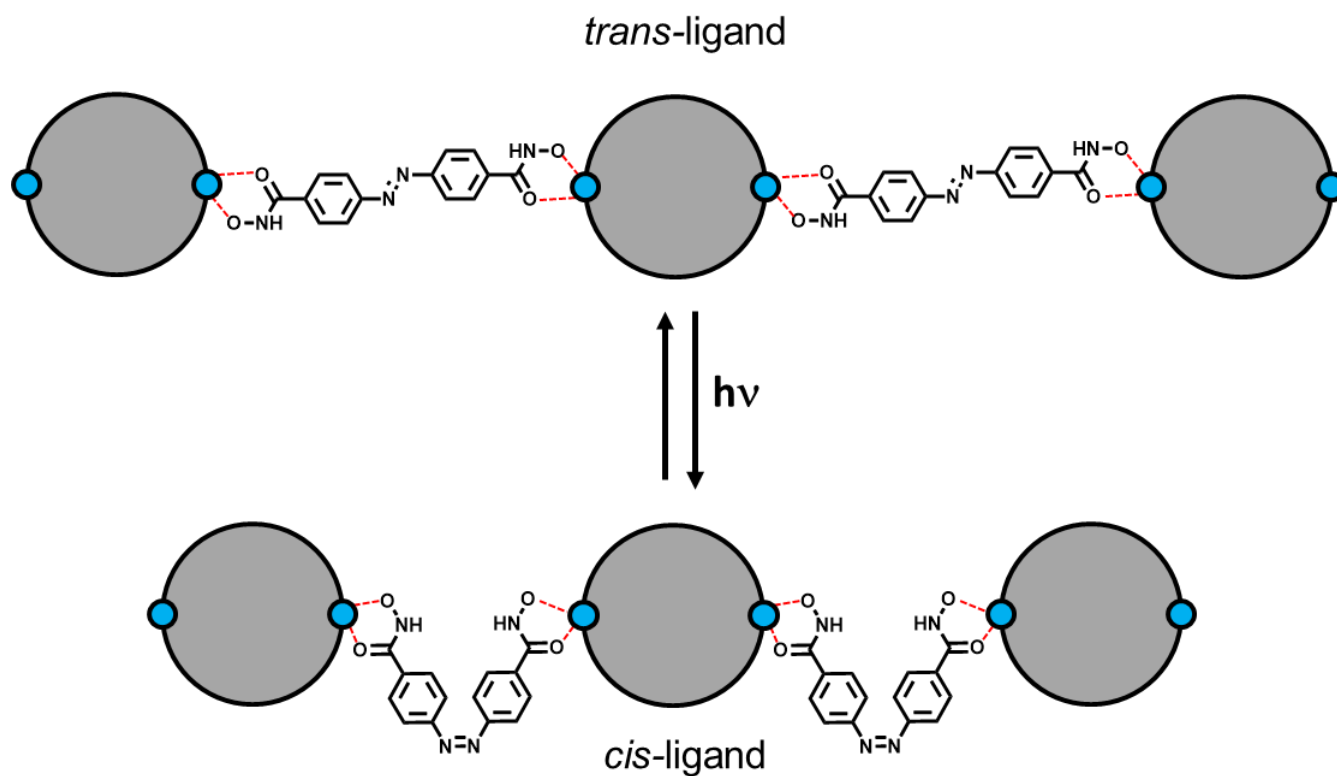


Figure 5.8. Cis-trans isomerization leads to a change in lattice spacing for protein-MOFs bridged by azobenzene-based ligands.

5.2.3. Development of size-comparable bridging ligands

The metal-bound spherical ferritin node, with a diameter of ca. 120 Å and a molecular weight of >505,000 g/mol, is by far the largest MOF node. The robust metal-coordinating motifs allow much smaller bridging ligands to template the assembly of highly-ordered 3D architectures. Much like conventional MOFs, these ditopic ligands can be systematically interchanged, leading to slight differences in both lattice symmetry and unit cell dimensions. These changes are much more drastic in conventional MOFs, where systematic increases in the ligand size significantly impact the unit cell dimensions and pore sizes, and tri- and tetratopic ligands form an entirely different lattice symmetry. Accessing a similar level of structural (and later functional) diversity, size-comparable ligands need to be employed in protein-MOFs. Appropriately sized organic ligands would be incredibly challenging to synthesize and likely lack sufficient solubility in aqueous conditions. Biomolecules, such as coiled-coil peptides, complementary DNA strands, or symmetric oligomeric proteins, are ideal candidates for bridging ligands due to their inherent water solubility and size-similarity to the protein-based nodes. Of the three biopolymers, proteins have the most structural and functional versatility. Additionally, well-established bioconjugation techniques allow for strong metal chelating groups (such as hydroxamic acids) to be appended site specifically to the protein surface. Two protein scaffolds with either C_2 or C_4 symmetry were selected as candidates for bridging ligands in heteroprotein-MOFs. A series of iodoacetamide based metal-chelating groups were synthesized and covalently attached to installed cysteines on the surface of these proteins.

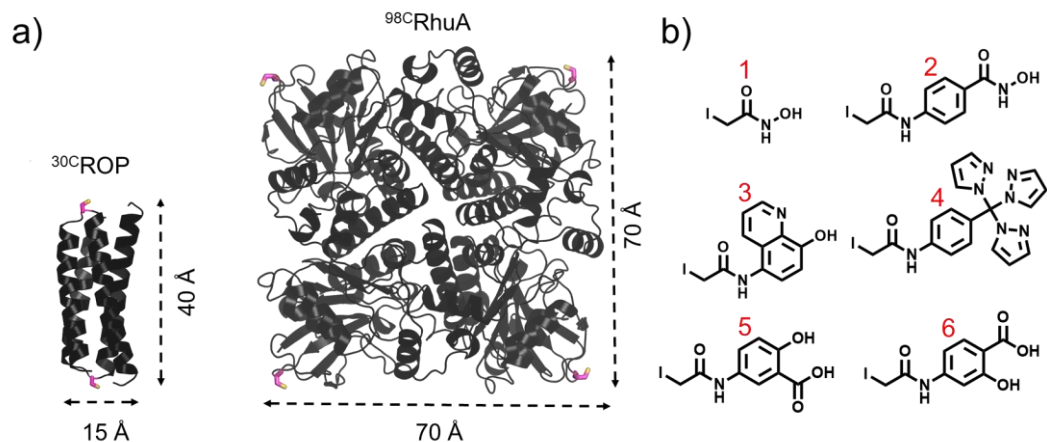


Figure 5.9. Protein-based bridging ligands can be created by labeling symmetric proteins with synthetic ligands. a) The dimensions of the dimeric $^{30}\text{CROP}$ and tetrameric $^{98}\text{CRhuA}$ are shown. Cysteine residues were modeled *in silico* (PDB IDs: 1ROP and 1OJR). b) Six different iodoacetamide-based ligands were synthesized with different metal-chelating groups (1 and 2 are hydroxamic acids, 3 is a hydroxyquinoline, 4 is a tri(pyrazol-1-yl) methane, 5 and 6 are salicylic acids).

The protein ROP (repressor of primer) was selected as the first C_2 symmetric bridging ligand. ROP is a relatively simple dimeric protein comprised of four total α -helices (two per monomer). The protein has a molecular weight of (ca. 7300 g/mol) and a linear length of ca. 40 Å. The loop between the two α -helices is ideal for the installation of a metal-binding group, which transforms ROP into a linear ditopic bridging ligand. To this end, a cysteine was installed at position 30 (S30C mutation), and ^{30}Cys was labeled with the series of metal-binding organic ligands. Due to its symmetry, this ligand would be expected to give body-centered lattices with much larger unit cell parameters. A scheme of ^{122}Hf ferritin assembly with $^{30}\text{CROP}$ hydroxamate is shown in Figure 5.10 a.

Rhamulose-1-phosphate aldolase (RhuA) was selected to act as a C_4 symmetric bridging ligand. RhuA is a tetrameric protein with a molecular weight of ca. 33000 g/mol and an edge length of ca. 70 Å. Previous research on this protein led to the incorporation of ^{98}Cys at the edges of RhuA.⁶ This position was functionalized with the same series of organic ligands, transforming $^{98}\text{CRhuA}$ into a planar tetratopic bridging ligand. The size of RhuA allows each metal-chelating

group to bind to a different ferritin-node. Since there are no comparable tetratopic organic ligands, two possible packings are shown in Figure 5.10 b.

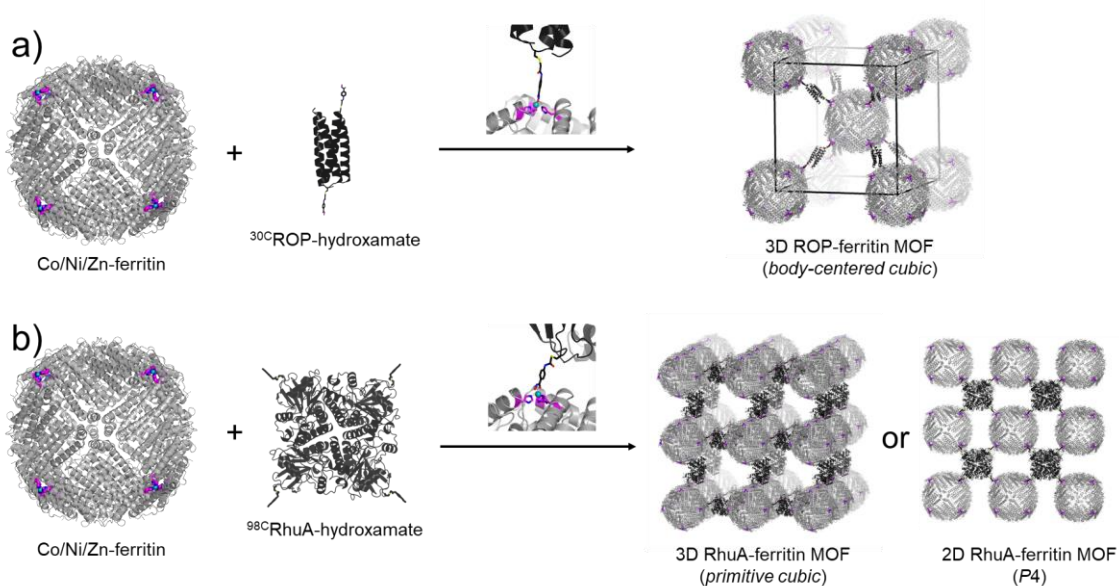


Figure 5.10. Possible lattice architectures for heteroprotein-MOFs. a) Schematic for the assembly of Co/Zn/Ni-ferritin with $^{30}\text{CROP}$ functionalize with a metal-chelating ligand (benzohydroxamic acid). The resulting lattices should be body-centered. b) Schematic for the assembly of Co/Zn/Ni-ferritin with $^{98}\text{CRhuA}$ functionalized with a metal-chelating ligand (benzohydroxamic acid).

5.3. Methods

5.3.1. General considerations

All chemical reagents were purchased from Sigma-Aldrich, Fisher Scientific, TCI, or an equivalent supplier. Mass spectrometry (MS) of proteins and small molecules was carried out at the Molecular Mass Spectrometry Facility at UCSD using electrospray ionization (ESI) on a Micromass Quattro Ultima Triple Quadrupole MS. NMR spectra (^1H) were recorded on a Varian Mercury spectrometer (400 MHz). NMR data are reported as follows: chemical shift, multiplicity (s = singlet, br = broad), and integration. The spectra were internally referenced to the residual solvent signal (DMSO: δ 2.50 ppm; CDCl_3 : δ 7.26 ppm).

5.3.2. Design and expression of *T. ni* variants

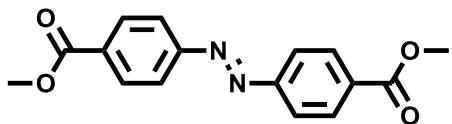
The protein sequences at the C_3 axes for both TniHF and TniLF were modified to mimic the metal coordination site at ^{122}H uHF. For the wild-type TniHF loop, residues 131-136 (GluAspAspSerGluPhe), were replaced with three residues (HisAspLys). For the wild-type TniLF loop, residues 148-155 (ThrArgAsnSerGlnHisLeuHis) were replaced with four residues (HisAspLysAsn). The DNA sequences of ^{132}H TniHF, ^{149}H TniLF, WT-TNiLF, and WT-TniHF (shown in Table 5.1. were all optimized and purchased from Genscript in separate pET-20b(+) vectors. These vectors were transformed into BL21 (DE3) *E. coli* cells and plated on LB agar containing 100 $\mu\text{g}/\text{mL}$ ampicillin. A single colony or freezer stock of BL21 containing ferritin variant was used to inoculate a starter cell culture (200 mL LB medium, 100 $\mu\text{g}/\text{mL}$ ampicillin). The starter culture was incubated for 16 hours at 37 $^\circ\text{C}$ while shaking at 250 rpm before inoculating 16 x 1 L LB cultures (10 mL per flask) supplemented with 100 mg/L ampicillin. The *E. coli* cells were grown at 37 $^\circ\text{C}$ until the OD_{600} was between 0.6 and 0.8. Protein expression was induced by the addition of isopropyl β -D-1-thiogalactopyranoside (IPTG) to a total concentration of 1 mM. Cells were incubated at 37 $^\circ\text{C}$ for 3.5 h and harvested by centrifugation (4000 rpm, 10 min). Cell pellets were combined and stored at -80 $^\circ\text{C}$. Frozen cell pellets were thawed and suspended in a buffered

solution containing 15 mM Tris (pH 7.4), 150 mM NaCl with 1 mM phenylmethylsulfonyl fluoride (PMSF), 5 mM DTT, and ca. 25 μ M lysozyme. Cells were sonicated for 12 min on ice, and the cell lysate was centrifuged at 4 °C (10,000 rpm, 10 min). The supernatant was collected, and the remaining cell pellet was resuspended in a buffered solution containing 15 mM Tris (pH 7.4), 150 mM NaCl with Triton X-100 (1% v/v) and sonicated for 12 min on ice. All of the variants were found in the inclusion bodies. The pellets could be resuspended with 2% SDS, 6 M guanidine-HCl, or 8 M urea in a buffered solution containing 15 mM TRIS (pH 10.0), and 150 mM NaCl.

Table 5.1. DNA sequences of *T. ni* ferritin variants

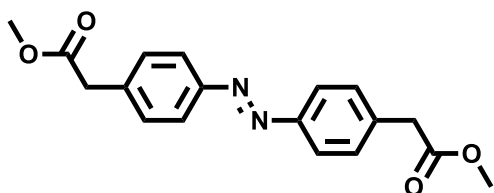
<i>T. Ni</i> ferritin variant	DNA sequence
WT-TniLF	ATGGCGGACACCTGCTACAATGACGTGGCGCTGGACTGCGGTATCACCAGCAA TAGCCTGGCGCTGCCGCGTTGCAATGCGGTTTACGGTGAATACGGCAGCCACG GTAACGTGGCGACCGAGCTGCAGGCGTATGCGAAACTGCACCTGGAACGTAG CTACGATTATCTGCTGAGCGCGGCGTACTTCAACAACCTATCAAACCAACCGTGC GGGCTTCAGCAAACCTGTTTAAGAACTGAGCGACGAGGCGTGGAGCAAGACCA TCGACATCATTAAAGCACGTTACCAAACGTGGTGACAAGATGAACTTCGATCAGC ACAGCACCATGAAAACCGAGCGTAAGAACTACACCGCGGAGAACCACGAACTG GAGGCGCTGGCGAAAGCGCTGGACACCCAAAAGGAACTGGCGGAGCGTGCCT TTTACATCCACCGTGAGGCGACCCGTAACAGCCAGCACCTGCACGACCCGGAA ATTGCGCAATATCTGGAGGAAGAGTTTCATCGAAGATCACGCGGAGAAGATTCTG ACCCTGGCGGGTACACCAGCGACCTGAAGAAATTTATTACCGCGAACAACCG TCACGATCTGAGCCTGGCGCTGTATGTGTTTCGATGAGTATCTGCAAAAAACCGT GTAA
¹⁴⁹ H-TniLF	ATGGCGGACACCTGCTACAATGACGTGGCGCTGGACTGCGGTATCACCAGCAA TAGCCTGGCGCTGCCGCGTTGCAATGCGGTTTACGGTGAATACGGCAGCCACG GTAACGTGGCGACCGAGCTGCAGGCGTATGCGAAACTGCACCTGGAACGTAG CTACGATTATCTGCTGAGCGCGGCGTACTTCAACAACCTATCAAACCAACCGTGC GGGCTTCAGCAAACCTGTTTAAGAACTGAGCGACGAGGCGTGGAGCAAGACCA TCGACATCATTAAAGCACGTTACCAAACGTGGTGACAAGATGAACTTCGATCAGC ACAGCACCATGAAAACCGAGCGTAAGAACTACACCGCGGAGAACCACGAACTG GAGGCGCTGGCGAAAGCGCTGGACACCCAAAAGGAACTGGCGGAGCGTGCCT TTTACATCCACCGTGAGGCGCATGATAAAAACGACCCGGAATTTGCGCAATATC TGGAGGAAGAGTTTCATCGAAGATCACGCGGAGAAGATTTCGTACCCTGGCGGGT CACACCAGCGACCTGAAGAAATTTATTACCGCGAACAACGGTCACGATCTGAGC CTGGCGCTGTATGTGTTTCGATGAGTATCTGCAAAAAACCGTGTAA
WT-TniHF	ATGACCCAATGCAATGTTAATCCGTTCAAATTCGAAAGACTGGATCACCATG CACCGTAGCTGCCGTAATAGCATGCGCCAACAGATCCAGATGGAAGTGGGTGC GAGCCTGCAATACCTGGCGATGGGCGCGCACTTCAGCAAGGATGTGGTTAAC GTCCGGGTTTTGCGCAGCTGTTCTTTGATGCGGCGAGCGAGGAACGTGAGCAT GCGATGAAACTGATTGAATATCTGCTGATGCGTGGCGAGCTGACCAACGATGT GAGCAGCCTGCTGCAAGTTCGTCCGCCACCCGTAGCAGCTGGAAGGGTGGC GTGGAGGCGCTGGAACATGCGCTGAGCATGGAGAGCGACGTGACCAAGAGCA TCCGTAACGTTATTAAGCGTGCGAGGACGATAGCGAGTTCAACGACTACCACC TGTTGATTATCTGACCGGTGACTTTCTGGAGGAACAGTACAAGGGTCAACGTG ATCTGGCGGGCAAAGCGAGCACCCCTGAAGAACTGATGGACCGTCACGAGGC GCTGGGTGAATTTATCTTCGACAAAAAACTGCTGGGCATTGATGTTTAA
¹³² H-TniHF	ATGACCCAATGCAATGTTAATCCGTTCAAATTCGAAAGACTGGATCACCATG CACCGTAGCTGCCGTAATAGCATGCGCCAACAGATCCAGATGGAAGTGGGTGC GAGCCTGCAATACCTGGCGATGGGCGCGCACTTCAGCAAGGATGTGGTTAAC GTCCGGGTTTTGCGCAGCTGTTCTTTGATGCGGCGAGCGAGGAACGTGAGCAT GCGATGAAACTGATTGAATATCTGCTGATGCGTGGCGAGCTGACCAACGATGT GAGCAGCCTGCTGCAAGTTCGTCCGCCACCCGTAGCAGCTGGAAGGGTGGC GTGGAGGCGCTGGAACATGCGCTGAGCATGGAGAGCGACGTGACCAAGAGCA TCCGTAACGTTATTAAGCGTGCCACGACAAGAACGACCCACCTGGTTGATT ATCTGACCGGTGACTTTCTGGAGGAACAGTACAAGGGTCAACGTGATCTGGCG GGCAAAGCGAGCACCCCTGAAGAACTGATGGACCGTCACGAGGCGCTGGGTG AATTTATCTTCGACAAAAAACTGCTGGGCATTGATGTTTAA

5.3.3. Synthesis of azobenzene precursors



Synthesis of dimethyl 4,4'-(diazene-1,2-diyl)(*E*)-dibenzoate

Methyl-4-aminobenzoate (1 g, 6.615 mmol) CuBr (0.949 g, 6.615 mmol) and pyridine (1.6 mL, 19.48 mmol) were combined in 40 mL of toluene in a 1000 mL round-bottom flask. The solution was purged with O₂, and the reaction was heated to 60 C under O₂. The reaction was stirred overnight. The toluene was removed in vacuo, and the crude product was dissolved in CHCl₃ and filtered. The filtrate was dried in vacuo and dry-loaded onto a silica column and purified with a gradient of 0-30% EtOAc in Hexanes. Fractions containing product were combined and dried in vacuo yielding 330 mg of pure product (33.5% yield). ¹H NMR: (400 MHz, CDCl₃) δ 7.86 (d, 4H), δ 6.63 (d, 4H), δ 3.85 (s, 6H).



Synthesis of dimethyl 2,2'-(diazene-1,2-diylbis(4,1-phenylene))(*E*)-diacetate

Methyl-4-aminopheny-acetate (1 mL, 6.9 mmol) CuBr (0.99 g, 6.9 mmol) and pyridine (1.67 mL, 20.7 mmol) were combined in 40 mL of toluene in a 1000 mL round-bottom flask. The solution was purged with O₂, and the reaction was heated to 60 C under O₂. The reaction was stirred overnight. The toluene was removed in vacuo, and the crude product was dissolved in CHCl₃ and filtered. The filtrate was dried in vacuo and dry-loaded onto a silica column and purified with a gradient of 0-30% EtOAc in Hexanes. Fractions containing product were combined and

dried in vacuo yielding 740 mg of pure product (75% yield). ^1H NMR: (400 MHz, CDCl_3) δ 7.88 (d, 4H), δ 7.41 (d, 4H), δ 3.72 (s, 6H), 3.72 (s, 4H).

5.3.4. Expression, purification, and characterization of $^{30}\text{CROP}$

We developed a new method for the expression, isolation, and purification of $^{30}\text{CROP}$. First, we genetically installed an N-terminal His-tag with a *Tobacco Etch Virus* nuclear-inclusion-a endopeptidase (TEV protease) cleavage site. The sequence of $^{30}\text{CROP}$ is shown in Table 5.2. Mutant plasmids isolated from XL-1 blue cells were transformed into BL21 (DE3) *E. coli* cells and plated on LB agar containing 100 $\mu\text{g}/\text{mL}$ ampicillin. A single colony or freezer stock of BL21 containing $^{30}\text{CROP}$ was used to inoculate starter cell cultures (200 mL LB medium, 100 $\mu\text{g}/\text{mL}$ ampicillin). The starter culture was incubated for ca. 16 hours at 37 $^\circ\text{C}$ while shaking at 250 rpm prior to inoculating 16 x 1 L LB cultures (10 mL per flask) supplemented with 100 $\mu\text{g}/\text{mL}$ ampicillin. The *E. coli* cells were grown at 37 $^\circ\text{C}$ until the OD_{600} was between 0.6 and 0.8. Protein expression was induced by the addition of isopropyl β -D-1-thiogalactopyranoside (IPTG) to a total concentration of 1 mM. Cells were incubated at 37 $^\circ\text{C}$ for ca. 4 h and harvested by centrifugation (4000 rpm, 10 min). Cell pellets were combined and stored at -80 $^\circ\text{C}$. Frozen cell pellets were thawed and suspended in a buffered solution containing 50 mM Na_2HPO_4 mM (pH 8.0), 300 mM NaCl, 5 mM β -mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride (PMSF), and ca. 25 μM lysozyme. Cells were sonicated for 30 min on ice, and the cell lysate was centrifuged at 4 $^\circ\text{C}$ (12,000 rpm, 10 min). The supernatant was loaded onto a nickel-NTA column (ThermoFisher) column and eluted using a gradient of 0 – 250 mM imidazole in a buffered solution containing 50 mM Na_2HPO_4 mM (pH 8.0) and 300 mM NaCl. Fractions containing $^{30}\text{CROP}$ eluted at 250 mM imidazole. These fractions were combined and dialyzed against 5 L of a buffered solution containing 50 mM Na_2HPO_4 mM (pH 8.0) and 300 mM NaCl to remove imidazole. The His-tag was removed by incubating 500 μM $^{30}\text{CROP}$ with 5 μM of TEV protease for 12 h. The solution was loaded onto a nickel-NTA column (ThermoFisher) column and eluted using 10 mM imidazole in a

buffered solution containing 50 mM Na₂HPO₄ mM (pH 8.0) and 300 mM NaCl. These fractions were combined and dialyzed against 5L of a buffered solution 50 mM Na₂HPO₄ mM (pH 8.0) and 300 mM NaCl, and 5 mM β-mercaptoethanol (βME), and concentrated in an Amicon stirred cell at 4 °C. The protein was stored at 4 °C until use.

Table 5.2. Sequence of ³⁰CROP.

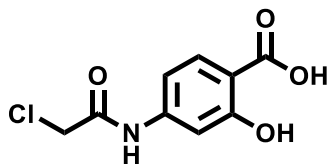
ROP variant	DNA sequence
³⁰ CROP	ATGCACCATCACCATCACCATGGCGGCTCAGAAAATCTTTATTTTCAATCG GGCGGCACTAAACAGGAAAAACCGCCCTTAACATGGCCCGTTTTATCAG AAGCCAGACATTAACGCTGCTGGAGAACTCAACAACTGTGCGCAGATG AACAGGCCGATATTGCT GAATCACTTCACGACCACGCCGACGAGCTTTA CCGCAGCGTCCTCGCACGCTTCGGTGATGACGGTGAAAACCTCTGA

5.3.5. Expression, purification, and characterization of ⁹⁸C RhuA

For ⁹⁸C RhuA expression, isolation, and purification, we used previously established procedures with minor modifications.⁶ Mutant plasmids isolated from XL-1 blue cells were transformed into BL21 (DE3) *E. coli* cells and plated on LB agar containing 100 µg/mL ampicillin. A single colony or freezer stock of BL21 containing ⁹⁸C RhuA was used to inoculate starter cell cultures (200 mL LB medium, 100 µg/mL ampicillin). The starter culture was incubated for ca. 16 hours at 37 °C while shaking at 250 rpm prior to inoculating 16 x 1 L LB cultures (10 mL per flask) supplemented with 100 µg/mL ampicillin. The *E. coli* cells were grown at 37 °C until the OD₆₀₀ was between 0.6 and 0.8. Protein expression was induced by the addition of isopropylβ-D-1-thiogalactopyranoside (IPTG) to a total concentration of 1 mM. Cells were incubated at 37 °C for ca. 12 h and harvested by centrifugation (4000 rpm, 10 min). Cell pellets were combined and stored at -80 °C. Frozen cell pellets were thawed and suspended in a buffered solution containing 10 mM Tris (pH 7.5), 1 mM ZnCl₂, 10 mM β-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride (PMSF), and ca. 25 µM lysozyme. Cells were sonicated for 12 min on ice, and the cell lysate was centrifuged at 4 °C (12,000 rpm, 10 min). The supernatant was collected, and 0.15% (w/v) Polymin-P (Acros) was added. The resulting mixture was stirred for 30 min before centrifugation

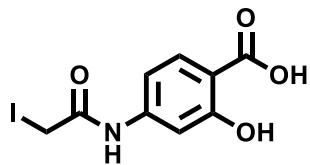
at 4 °C (12,000 rpm, 30 min). The supernatant was loaded onto a DEAE-Sepharose CL-6B (GE Healthcare) column and eluted using a gradient of 0 – 500 mM NaCl in a buffered solution containing 10 mM Tris and 10 mM β -mercaptoethanol at 4 °C. Fractions containing ^{98}C RhuA eluted at ca. 200 mM NaCl. These fractions were combined, and 1.7 M ammonium sulfate was added. The precipitate was pelleted by centrifugation at 4 °C (12,000 rpm, 45 min). The precipitate was dissolved in a buffered solution containing 5 mM sodium phosphate (pH 7.2), 1 mM ZnCl_2 , and 10 mM β ME, and then dialyzed three times against 5 L of the same buffered solution. Further purification was done using a High Q cartridge column (BioRad) using a 0-500 mM NaCl gradient at pH 8.0 on a DuoFlow fast protein chromatography workstation (Bio-Rad). Pure protein fractions eluted at 350 mM NaCl. These fractions were combined, dialyzed three times against a buffered solution containing 10 mM TRIS (pH 7.5), 1 mM ZnCl_2 , and 10 mM β -mercaptoethanol (β ME), and concentrated in an Amicon stirred cell at 4 °C. The protein was stored at 4 °C until use.

5.3.6. Synthesis of chelate-based ligands for bioconjugation



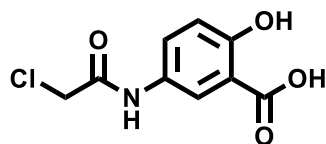
Synthesis of 4-(2-chloroacetamido)-2-hydroxybenzoic acid

4-amino-2-hydroxybenzoic acid (0.5 g, 3.26 mmol) and NaOH (0.39 g, 9.78 mmol) were dissolved in 50 mL H_2O and cooled to 0 °C. Chloroacetyl chloride (0.756 mL, 9.78 mmol) was added dropwise over 10 min. The reaction was gradually warmed to room temperature over 12 h. The product was filtered and dried *in vacuo* yielding 223 mg (26% yield). ^1H NMR: (400 MHz, DMSO-d_6) δ 11.40 (br s, 1H, OH), δ 10.55 (s, 1H), δ 7.74 (d, 1H), 7.33 (s, 1H), 7.06 (dd, 1H), 4.28 (s, 2H).



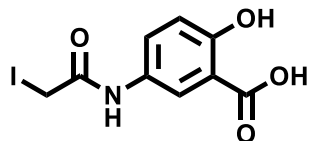
Synthesis of 4-(2-iodoacetamido)-2-hydroxybenzoic acid

4-(2-chloroacetamido)-2-hydroxybenzoic acid (0.1 g, 0.436 mmol) and NaI (0.327 g, 2.18 mmol) were combined in 40 mL of acetone and refluxed for 3 h. The precipitate was removed by filtration, and the filtrate was dried *in vacuo*. The product was washed with 10% (w/v) sodium thiosulfate and water before it was collected by filtration and dried *in vacuo*. The final product is light sensitive and was stored at $-20\text{ }^{\circ}\text{C}$ in the dark. This reaction gave >95% yield. $^1\text{H NMR}$: (400 MHz, DMSO-d_6) δ 11.39 (br s, 1H), δ 10.56 (s, 1H), δ 7.73 (d, 1H), 7.30 (s, 1H), 7.02 (dd, 1H), 3.83 (s, 2H).



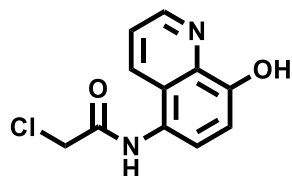
Synthesis of 5-(2-chloroacetamido)-2-hydroxybenzoic acid

5-amino-2-hydroxybenzoic acid (0.5 g, 3.26 mmol) and NaOH (0.39 g, 9.78 mmol) were dissolved in 50 mL H_2O and cooled to $0\text{ }^{\circ}\text{C}$. Chloroacetyl chloride (0.756 mL, 9.78 mmol) was added dropwise over 10 min. The reaction was gradually warmed to room temperature over 12 h. The product was filtered and dried *in vacuo* yielding 290 mg (34% yield). $^1\text{H NMR}$: (400 MHz, DMSO-d_6) δ 11.10 (br s, 1H), δ 10.27 (s, 1H), δ 8.10 (s, 1H), 7.64 (dd, 1H), 6.94 (d, 1H), 4.21 (s, 2H).



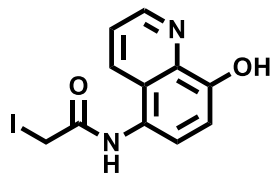
Synthesis of 5-(2-Iodoacetamido)-2-hydroxybenzoic acid

5-(2-chloroacetamido)-2-hydroxybenzoic acid (0.1 g, 0.436 mmol) and NaI (0.327 g, 2.18 mmol) were combined in 40 mL of acetone and refluxed for 3 h. The precipitate was removed by filtration, and the filtrate was dried *in vacuo*. The product was washed with 10% (w/v) sodium thiosulfate and water before it was collected by filtration and dried *in vacuo*. The final product is light sensitive and was stored at $-20\text{ }^{\circ}\text{C}$ in the dark. This reaction gave >95% yield. $^1\text{H NMR}$: (400 MHz, DMSO-d_6) δ 11.10 (br s, 1H, OH), δ 10.30 (s, 1H), δ 8.08 (s, 1H), 7.63 (dd, 1H), 6.92 (d, 1H), 3.79 (s, 2H).



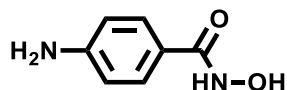
Synthesis of 2-chloro-*N*-(8-hydroxyquinolin-5-yl)acetamide

5-amino-8-hydroxy quinoline (0.5 g, 2.144 mmol) and NaOH (0.343 g, 8.58 mmol) were dissolved in 50 mL H_2O and cooled to $0\text{ }^{\circ}\text{C}$. Chloroacetyl chloride (0.341 mL, 4.29 mmol) was added dropwise over 10 min. The reaction was gradually warmed to room temperature over 12 h. The product was filtered and dried *in vacuo* yielding 92 mg (13% yield). $^1\text{H NMR}$: (400 MHz, DMSO-d_6) δ 10.14 (s, 1H), δ 9.89 (s, 1H), δ 8.88 (dd, 1H), 8.31 (dd, 1H), 7.62 (dd, 1H), 7.45 (d, 1H), 7.07 (d, 1H), 4.38 (s, 2H).



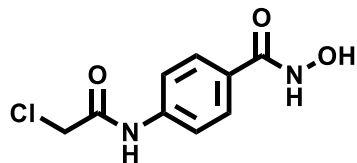
Synthesis of 2-iodo-*N*-(8-hydroxyquinolin-5-yl)acetamide

2-chloro-*N*-(8-hydroxyquinolin-5-yl)acetamide (0.072 g, 0.304 mmol) and NaI (0.228 g, 1.52 mmol) were combined in 40 mL of acetone and refluxed for 3 h. The precipitate was removed by filtration, and the filtrate was dried *in vacuo*. The product was washed with 10% (w/v) sodium thiosulfate and water before it was collected by filtration and dried *in vacuo*. The final product is light sensitive and was stored at $-20\text{ }^{\circ}\text{C}$ in the dark. This reaction gave >95% yield. $^1\text{H NMR}$: (400 MHz, $\text{DMSO-}d_6$) δ 10.18 (s, 1H), δ 9.88 (s, 1H), δ 8.88 (dd, 1H), 8.34 (dd, 1H), 7.62 (dd, 1H), 7.44 (d, 1H), 7.076 (d, 1H), 3.94 (s, 2H).



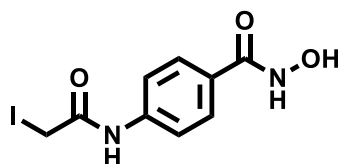
Synthesis of 4-amino-*N*-hydroxybenzamide

Hydroxylamine hydrochloride (1.75 g, 25.44 mmol) and sodium hydroxide (2.04 g, 50.88 mmol) were combined in 20 mL methanol and shaken vigorously. The solution was cooled on ice for 10 min, and NaCl was removed via filtration. The filtrate was combined with methyl 4-aminobenzoate (1 g, 6.36 mmol) dimethyl terephthalate in 40 mL methanol. The solution was stirred overnight at room temperature. The solvent was removed *in vacuo*, and the solid material was dissolved in 20 mL H_2O . The solution was acidified to a pH of 5.5 with 5% HCl. The solvent was removed *in vacuo* and dissolved in MeOH. The precipitate was removed, and the filtrate was concentrated *in vacuo* yielding 350 mg product (35% yield). $^1\text{H NMR}$: (400 MHz, $\text{DMSO-}d_6$) δ 10.75 (br s, 1H), δ 8.67 (br s, 1H), δ 7.45 (d, 2H), 6.51 (d, 2H), 5.58 (s, 2H).



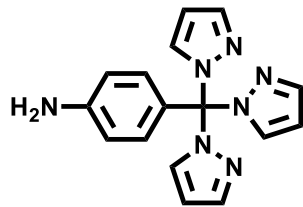
Synthesis of 4-(2-chloroacetamido)-*N*-hydroxybenzamide

4-amino-*N*-hydroxybenzamide (0.3 g, 1.97 mmol) and NaOH (0.236 g, 5.92 mmol) were dissolved in 40 mL H₂O and cooled to 0 °C. Chloroacetyl chloride (0.471 mL, 5.92 mmol) was added dropwise over 10 min. The reaction was gradually warmed to room temperature over 12 h. The product was filtered and dried *in vacuo* yielding 240 mg (53% yield). ¹H NMR: (400 MHz, DMSO-*d*₆) δ 11.13 (br s, 1H, NH), δ 10.50 (s, 1H, NH), δ 8.97 (s, 1H, OH), 7.72 (d, 2H, ArH), 7.65 (d, 2H, ArH), 4.28 (s, 2H, CH₂).



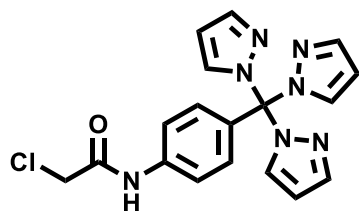
Synthesis of 4-(2-iodoacetamido)-*N*-hydroxybenzamide

4-(2-iodoacetamido)-*N*-hydroxybenzamide (0.1 g, 0.437 mmol) and NaI (0.328 g, 2.19 mmol) were combined in 40 mL of acetone and refluxed for 3 h. The precipitate was removed by filtration, and the filtrate was dried *in vacuo*. The product was washed with 10% (w/v) sodium thiosulfate and water before it was collected by filtration and dried *in vacuo*. The final product is light sensitive and was stored at -20 °C in the dark. This reaction gave >95% yield. ¹H NMR: (400 MHz, DMSO-*d*₆) δ 11.12 (s, 1H), δ 10.53 (s, 1H), δ 8.96 (s, 1H), 7.72 (dd, 1H), 7.62 (dd, 2H), 3.84 (s, 2H).



Synthesis of 4-(tri(1*H*-pyrazol-1-yl)methyl)aniline

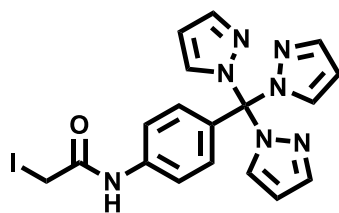
4-(trifluoromethyl)aniline (1.61 g, 10 mmol), 1*H*-pyrazole (2.04 g, 30 mmol) and NaOH (1.2 g, 30 mmol) were all combined in 10 mL DMSO and refluxed for 2 h. H₂O (50 mL) was added, and the reaction was extracted 2x with 50 mL DCM. The organic layer was collected, dried with anhydrous sodium sulfate, and purified by column chromatography with 100% Et₂O as the eluent. Pure fractions were combined, and the solvent was removed in vacuo, yielding 1.4 g of product (46%). ¹H NMR: (400 MHz, DMSO-*d*₆) δ 7.68 (m, 3H), δ 7.32(m, 3H), δ 6.62 (d, 2H), 6.50 (d, 2H), 6.39 (m, 3H), 5.52 (s, 2H).



Synthesis of 2-chloro-*N*-(4-(tri(1*H*-pyrazol-1-yl)methyl)phenyl)acetamide

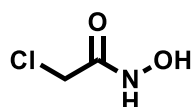
4-(tri(1*H*-pyrazol-1-yl)methyl)aniline (300 mg, 0.98 mmol) and TEA (1.18 mmol, 0.164 mL) were combined in 5 mL of DCM and cooled to -20 °C. 2-chloroacetyl chloride (1.18 mmol, 0.094 mL) was dissolved in 5 mL DCM and cooled to -20 °C. The solution containing 2-chloroacetyl chloride was added to the solution containing 4-(tri(1*H*-pyrazol-1-yl)methyl)aniline and TEA slowly over 2 h at -20 °C. After 2 h, the solution was gradually warmed to RT over 12 h. The solution was removed in vacuo and purified by silica column chromatography 0-5% MeOH in CHCl₃ as the eluent. Pure fractions were combined, and the solvent was removed in vacuo yielding 237 mg

(63% yield). ^1H NMR: (400 MHz, DMSO-d_6) δ 10.56 (s, 1H), δ 7.73 (m, 3H), δ 7.63 (d, 2H), δ 7.33 (m, 3H), 7.03 (d, 2H), 6.45 (m, 3H), 4.29 (s, 2H).



Synthesis of 2-iodo-*N*-(4-(tri(1*H*-pyrazol-1-yl)methyl)phenyl)acetamide

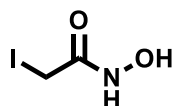
2-chloro-*N*-(4-(tri(1*H*-pyrazol-1-yl)methyl)phenyl)acetamide (0.075 g, 0.196 mmol) and NaI (0.157 g, 0.982 mmol) were combined in 40 mL of acetone and refluxed for 3 h. The precipitate was removed by filtration, and the filtrate was dried *in vacuo*. The product was washed with 10% (w/v) sodium thiosulfate and water before it was collected by filtration and dried *in vacuo*. The final product is light sensitive and was stored at $-20\text{ }^\circ\text{C}$ in the dark. This reaction gave >95% yield. ^1H NMR: (400 MHz, DMSO-d_6) δ 10.69 (s, 1H), δ 7.72 (m, 3H), δ 7.59 (d, 2H), δ 7.33 (m, 3H), 7.00 (d, 2H), 6.44 (m, 3H), 3.83 (s, 2H).



Synthesis of 2-chloro-*N*-hydroxyacetamide

Chloroacetyl chloride (0.58 mL, 7.3 mmol) was dissolved in 2 mL of CH_2Cl_2 and added dropwise to a suspension of *O*-tritylhydroxylamine (2.0 g, 7.3 mmol) and DIPEA (2.5 mL, 14.5 mmol) in 15 mL of CH_2Cl_2 at $0\text{ }^\circ\text{C}$. The reaction mixture was gradually warmed to room temperature and stirred at room temperature for an hour. An additional 15 mL of CH_2Cl_2 was added, and the reaction was extracted with H_2O (3 \times 30 mL). The CH_2Cl_2 solution was collected and evaporated to dryness. A solution containing 15 mL of CH_2Cl_2 with 10% (v/v) trifluoroacetic

acid was added, and the solution was stirred for 30 min. The crude product was purified by silica gel chromatography using a gradient of 0–100% ethyl acetate in hexanes as the eluent. Pure fractions were combined, and the solvent was removed *in vacuo* yielding 437 mg of product (55% yield). ^1H NMR: (300 MHz, DMSO- d_6) δ 10.88 (s, 1H), δ 9.15 (s, 1H), δ 3.93 (s, 2H). ^{13}C NMR: (500 MHz, DMSO- d_6) δ 162.88, δ 40.45. Measured molecular weight (m/z): 108.37 (M – H $^+$); Calculated: 107.99 (M – H $^+$)



Synthesis of 2-iodo-*N*-hydroxyacetamide

2-chloro-*N*-hydroxyacetamide (400 mg, 3.7 mmol) and NaI (2.7 g, 18.3 mmol) were refluxed in 30 mL of acetone for 1 h. The reaction mixture was purified by silica gel chromatography with 100% ethyl acetate as the eluent and dried *in vacuo* yielding 730 mg of product (99% yield). ^1H NMR: (300 MHz, DMSO- d_6) δ 10.81 (s, 1H), δ 9.09 (s, 1H), δ 3.51 (s, 2H). ^{13}C NMR: (500 MHz, DMSO- d_6) δ 164.83, δ –2.01. Measured molecular weight (m/z): 223.85 (M + Na $^+$); Calculated: 223.95 (M + Na $^+$).

5.3.7. Bioconjugation of chelate-based ligands to ^{30}C ROP and ^{98}C RhuA

Purified ^{30}C ROP and ^{98}C RhuA were exchanged into a buffered solution containing 50 mM HEPES (pH 7-8) using Microcon centrifugal filters with a 3 kDa cutoff. Protein concentration was determined by a Bradford assay, and the proteins were diluted to 100 μM (by monomer). A stock solution of 10 mM TCEP was prepared, and the pH was adjusted to ca. 7 through the addition of NaOH. A stock solution of 20 mM of the selected ligand was prepared in DMF. The bioconjugation conditions were 100 μM protein, 1 mM TCEP, and 1-2 mM of the ligand in a buffered solution containing 50 mM HEPES (pH 7-8) and 10% (v/v) DMF. The bioconjugation reactions were stirred at 4 $^\circ\text{C}$ for 12-48 h in the dark. Mass spectrometry was used to confirm labeling. After labeling,

the protein solutions were washed five times with a buffered solution containing 50 mM HEPES (pH 7-8) using Microcon centrifugal filters with a 3 kDa cutoff to remove DMF and excess ligand. The proteins were stored at 4 °C until use.

5.4. Acknowledgments

Chapter 5, in part, is based on the unpublished research efforts of the dissertation author with contributions from Rob Alberstien (⁹⁸C RhuA), Jie Zhu, Yiyang Li (³⁰C ROP), Jerika Chiong (synthesis of organic ligands), and Akif Tezcan.

5.5. References

1. Hamburger, A. E.; West, A. P.; Hamburger, Z. A.; Hamburger, P.; Bjorkman, P. J., Crystal Structure of a Secreted Insect Ferritin Reveals a Symmetrical Arrangement of Heavy and Light Chains. *Journal of Molecular Biology* **2005**, *349* (3), 558-569.
2. Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y., The I-TASSER Suite: protein structure and function prediction. *Nature Methods* **2015**, *12* (1), 7-8.
3. Huard, D. J.; Kane, K. M.; Tezcan, F. A., Re-engineering protein interfaces yields copper-inducible ferritin cage assembly. *Nat. Chem. Biol* **2013**, *9* (3), 169-76.
4. Li, M.; Yao, Y.; Ding, J.; Liu, L.; Qin, J.; Zhao, Y.; Hou, H.; Fan, Y., Spectroscopic and Crystallographic Investigations of Novel BODIPY-Derived Metal–Organic Frameworks. *Inorganic Chemistry* **2015**, *54* (4), 1346-1353.
5. Bandara, H. M. D.; Burdette, S. C., Photoisomerization in different classes of azobenzene. *Chemical Society Reviews* **2012**, *41* (5), 1809-1825.
6. Suzuki, Y.; Cardone, G.; Restrepo, D.; Zavattieri, P. D.; Baker, T. S.; Tezcan, F. A., Self-assembly of coherently dynamic, auxetic, two-dimensional protein crystals. *Nature* **2016**, *533* (7603), 369-+.

Appendix 1. Hydroxamate-based MOFs

A1.1. Introduction

As discussed in previous chapters, metal-organic frameworks (MOFs) are a class of porous, crystalline materials with potential applications in catalysis, gas storage, drug delivery, molecular separations, and molecular sensing.¹⁻⁴ The broad applicability of MOFs stems from their reticular synthesis, wherein interchangeable metal nodes and bridging organic ligands are combined to form highly ordered frameworks. Each node is comprised of either a single metal atom or, more commonly, a polynuclear cluster of metal atoms, also known as a secondary building unit (SBU). The introduction of the SBU allows for the construction of diverse node geometries that form MOFs with greater structural complexity.⁵ Early SBUs were mainly composed of divalent transition metals (Zn^{2+} or Cu^{2+}) and carboxylate groups.⁶⁻⁷ The metal-ligand bonds in these nodes are relatively labile and prone to hydrolysis in aqueous conditions.¹ Infiltration of water molecules into the SBU of these MOFs undermines the structural integrity of the MOFs, eventually leading to collapse of the crystalline framework.

One successful approach to improve the aqueous stability of the framework is to design nodes with stronger metal-ligand bonds. Following Pearson's hard/soft acid/base (HSAB) theory, softer azolate ligands (imidazoles, pyrazoles, etc.) form more stable metal-ligand bonds with softer metal ions such as Zn^{2+} and Cu^{2+} . For example, zeolitic imidazolate frameworks (ZIFs) are based on the strong Zn-azolate bonds.⁸ Carboxylates are much harder ligands and form more stable metal-ligand bonds with high-valent metal ions (Fe^{3+} , Al^{3+} , Cr^{3+} , Zr^{4+} , Ti^{4+}). This has led to a wide variety of stable frameworks, including MIL-53 (Cr^{3+}), MIL-100 (Fe^{3+}), and UiO-66 (Zr^{4+})⁸⁻¹⁰. However, these stable frameworks are still constructed from ligands containing either azolate or carboxylate head groups. The design of protein-MOFs (Chapter 2) highlighted that replacing the weakly coordinating dicarboxylate with a dihydroxamate allows for the formation of ordered

architectures in an aqueous environment. This concept is not unique to protein-MOFs and can be employed to form classical MOFs that are highly stable in aqueous and alkaline conditions.

A survey of the literature reveals that chelating ligands (and especially hydroxamic acids) are underutilized in the construction of MOFs. One of the first examples of hydroxamate implementation into MOFs involved the post-synthetic exchange of *p*-H₂bdt into preformed UiO-66 crystals. The hydroxamate incorporation into the framework resulted in a significant increased alkaline stability of the MOF.¹¹ In 2019, the first example of a traditional MOF synthesized from hydroxamate-based bridging ligands (*p*-H₂bdt) was published.¹² However, the hydroxamate-mediated connectivity of this MOF only extends in two dimensions. Indeed, outside of a small number of catecholate-based frameworks, most coordination solids containing chelating motifs form either 2D sheets connected by noncovalent interactions (either π - π stacking or H-bonding) or employ additional monodentate ligands to facilitate connectivity in the 3rd dimension.¹³⁻¹⁷ This is largely due to the low number of ligands that coordinate to each metal node (usually 2 or 3), which prohibits the formation of 3D frameworks in all but a few select ligand geometries. Much like ferritin-MOFs, where each ferritin node binds eight hydroxamate ligands, a chelate-based SBU would allow for higher ligand coordination numbers and formation of a 3D architecture.¹⁸⁻¹⁹

A1.2. Results and discussion

A1.2.1. Design and synthesis of a chelate-based SBU

Many SBUs have been inspired by discrete, symmetric metal-oxo clusters. Similarly, metal-organic polyhedra (MOPs) have been used as the basis for the construction of novel MOFs²⁰. Unlike extended 3D frameworks, chelate-based ligands are much more prevalent in MOPs²¹⁻²³. In this appendix, a MOP derived from a multitopic, chelating ligand was utilized as the foundation to form a 3D framework. It is the first example of a MOF containing a chelate-based SBU. When selecting a MOP to act as the polynuclear SBU, several criteria need to be met: 1) the MOP needs to form in high fidelity with strong metal-ligand bonds 2) the net of the MOP needs to have strategically placed points-of-attachment to interconnect the discrete clusters through the bridging ligand, and 3) the ligand needs to be synthetically accessible. We took inspiration for the chelate-based node from an M_4L_6 coordination cluster published by Beissel et al., shown in Figure A1.1.²³ This cluster was based on the *N*-toluyl derivative of *meta*-benzene dihydroxamic acid (*m*-H₂bdh). Hydroxamates are an ideal chelating group for the synthesis of extended structures since they are rotationally flexible and sterically unencumbered with a steric footprint similar to the often-used carboxylates. Hydroxamate ligands form neutral tris(chelates) with trivalent metals such as Fe³⁺, Ga³⁺, or Al³⁺. An analogous hydroxamate-based cluster was synthesized by combining *m*-bdh and FeCl₃ in DMF at room temperature (Figure A1.2 a). The *N*-toluyl groups were excluded during ligand synthesis to reduce the steric bulk around the cluster. As expected, each Fe-tris(hydroxamate) forms a C₃ vertex of the discrete cluster. Each cluster has an interior cavity that is ca. 7 Å in diameter (Figure A1.2 a). Since each metal center is chiral, each discrete cluster can have *T* (ΔΔΔΔ or ΛΛΛΛ), C₃ (ΔΔΔΛ or ΛΛΛΔ), or S₄ (ΔΔΛΛ) symmetry in solution. The Fe₄*m*-bdh₆ clusters crystalize readily into *P*2₁3 symmetry. Each unit cell is comprised of eight clusters with *T* symmetry (four ΔΔΔΔ and four ΛΛΛΛ). On average, the Fe-carbonyl and Fe-hydroxyl bond lengths are 2.049 Å and 1.978 Å, respectively. The clusters in this crystal lattice are held together by a series of H-bonding and pi-pi stacking interactions (Figure A1.3). This cluster is ideal for

incorporation as a framework SBU due to the outward-facing connection points that form a pseudo-octahedron shaped net (Figure A1.2 b, highlighted in green). Connecting these vertices should lead to an interconnected 3D framework based on the hydroxamate SBU.

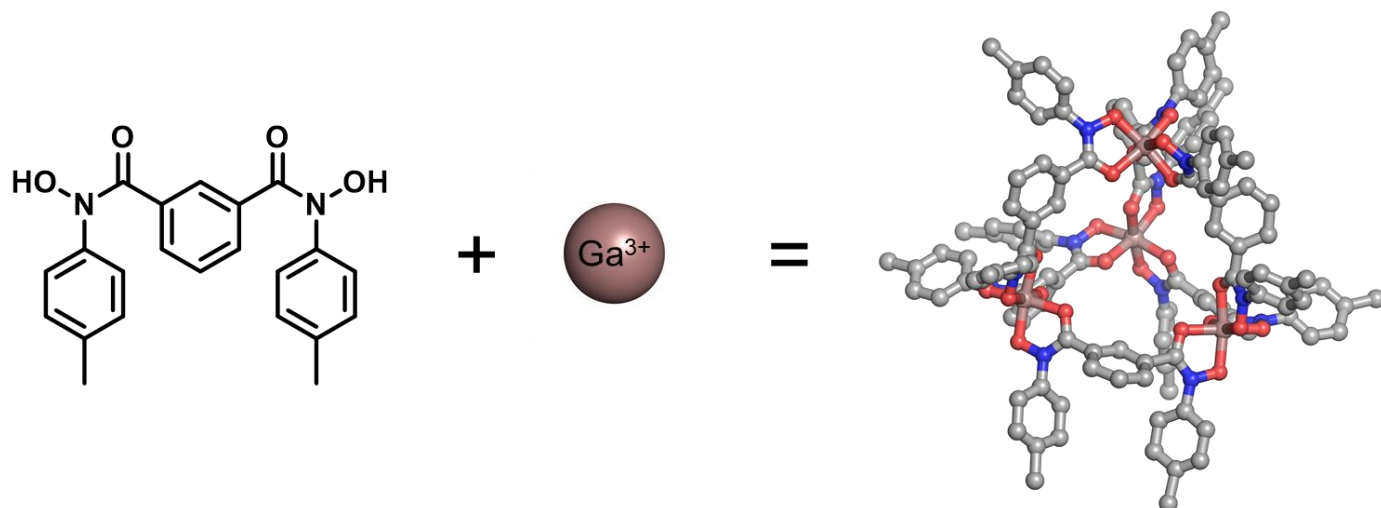


Figure A1.1. Hydroxamate-based MOP synthesized from an *N*-toluyl derivative of *m*-H₂BDH and Ga(III).

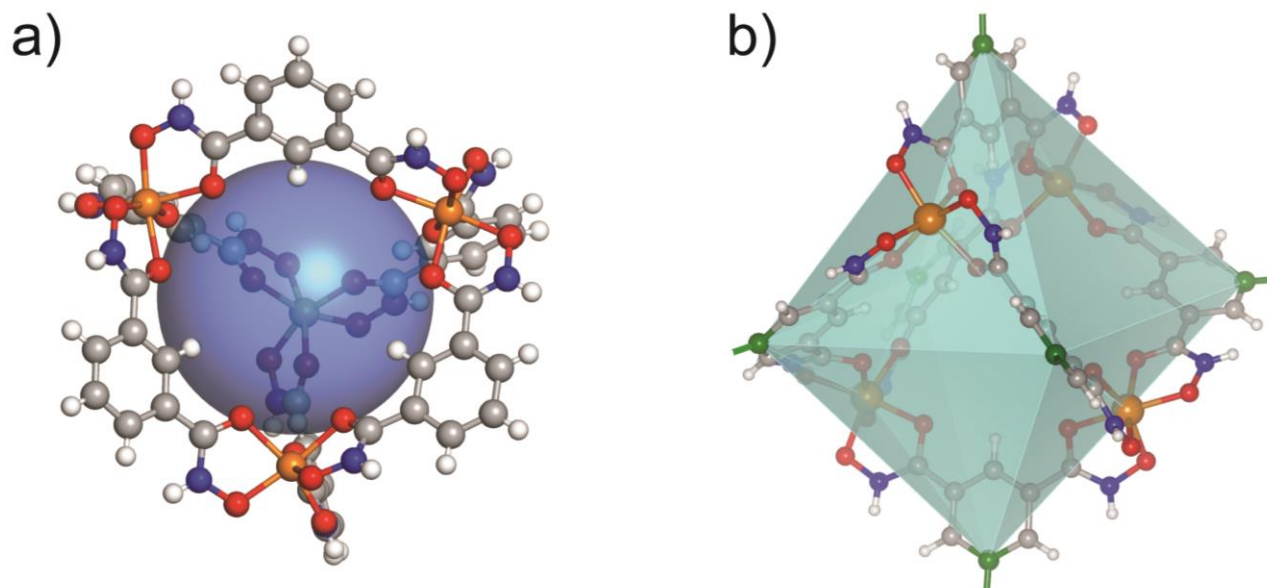


Figure A1.2. Models of the Δ -Fe₄*m*-bdh₆ cluster. a) The blue sphere highlights the pore of the cluster. b) The underlying polyhedron (teal) highlights the vertices of the cluster. The *meta* positions of each *m*-bdh (shown

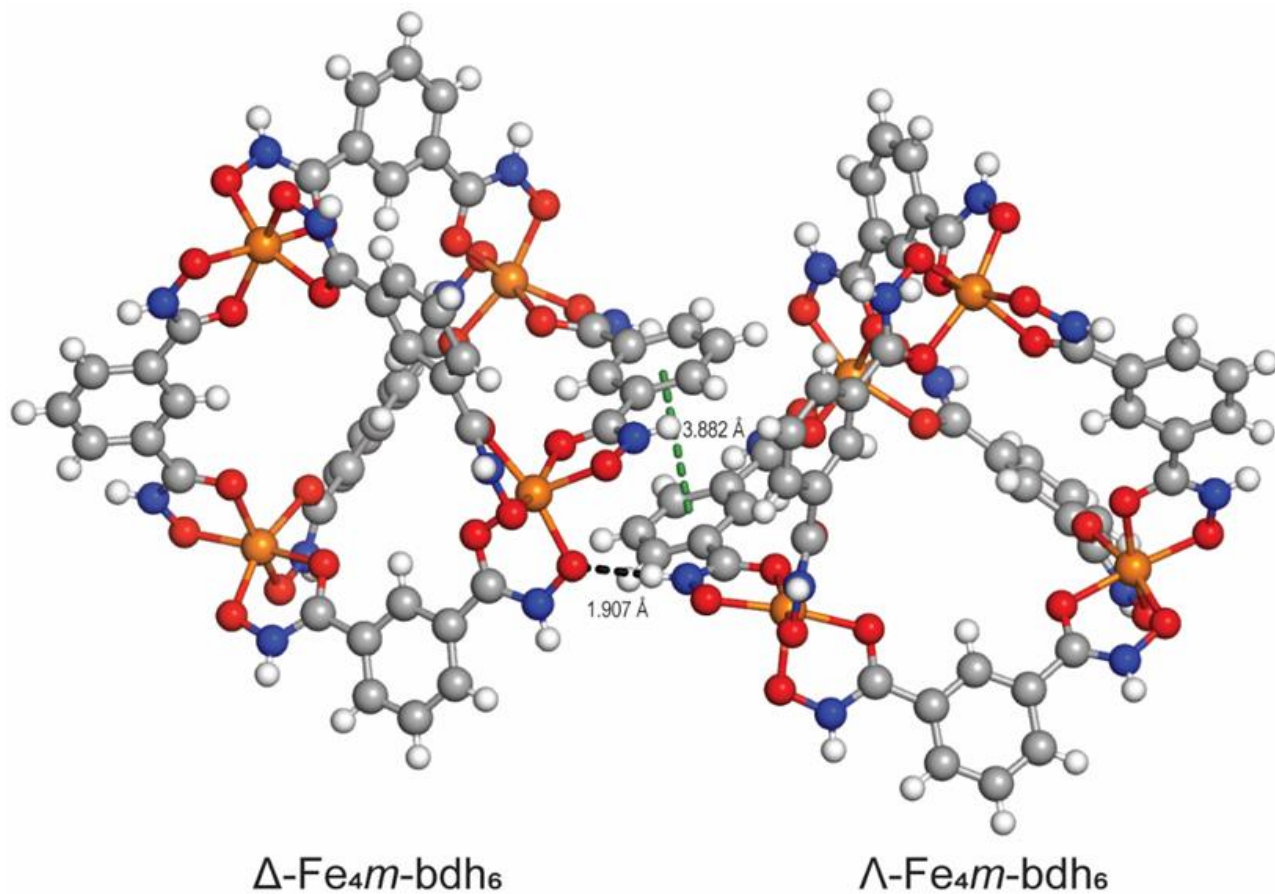


Figure A1.3. Crystal structure of Δ -Fe_{4m}-bdh₆ and Λ -Fe_{4m}-bdh₆. The primary crystal packing interactions are comprised of an intramolecular hydrogen bond between the hydroxyl and amide hydrogen (shown as a black dashed line) and the π - π stacking of the phenyl rings (shown as a green dashed line).

A1.2.2. Synthesis of hydroxamate-based MOFs

Since *m*-H₂bdh facilitates the formation of discrete 0D MOPs, a collinear ligand with *m*-H₂bdh head groups would enable the formation of an interconnected 3D structure. To this end, we devised the synthetic scheme shown in Figure A1.4 a. We synthesized a series of collinear bridging ligands capped with the *m*-H₂bdh head groups (Figure A1.4 b): biphenyl tetrahydroxamic acid (H₄BPTH), terphenyl tetrahydroxamic acid (H₄TPTH), and quaterphenyl (H₄QPTH). These ligands were all used in a broad screen of synthesis conditions, that varied metal identity, metal concentration, ligand concentration, solvent, and assembly temperature. We utilized PXRD to probe these assembly conditions for the formation of ordered frameworks and found a unique set of assembly conditions for each ligand. While the assembly conditions that favored MOF formation were unique for each ligand, we were able to highlight a series of trends. Fe(III) was the only metal that promoted MOF formation. Surprisingly, the oft-interchangeable Ga(III) did not yield ordered frameworks under any conditions. Second, the ideal assembly temperatures for each MOF were 60-80 °C. This temperature is relatively mild for MOF synthesis, but very similar to the temperature used to make the original *N*-toluyl *m*-H₂bdh based cluster. The powder diffraction patterns and SEM images of each MOF are shown in Figure A1.5. The powder diffraction pattern indicates that three MOFs are not isostructural. Therefore, it is not surprising that there is a significant variation in the crystal morphologies of the microcrystalline MOFs. Fe-BPTH forms crystals with a cubic morphology, which likely originate from a highly symmetric structure. Both Fe-TPTH and Fe-QPTH form hexagonal or rectangular rods, respectively, which indicates that these structures are inherently less symmetric.

For the precise determination of lattice symmetry and unit cell dimensions, we collected powder X-ray diffraction data at BL 11-BM (APS). A comparison of the Fe-BPTH and Fe-TPTH diffraction data collected in-house and at APS is shown in Figure A1.6. Unfortunately, the crystalline Fe-QPTH powder did not diffract well enough for unit cell determination. Fe-BPTH

crystallizes into a highly symmetric face-centered cubic (*fcc*) lattice with cell parameters $a = 34.35$ Å. Fe-TPTH crystallizes into a less symmetric trigonal lattice with $a = 18.29$ Å and $c = 16.32$ Å. With help from our collaborators at APS, we were able to determine a preliminary structure for both species. However, we elected not to pursue Rietveld refinement of these structures as we were able to obtain single-crystal diffraction data on both Fe-BPTH and Fe-TPTH.

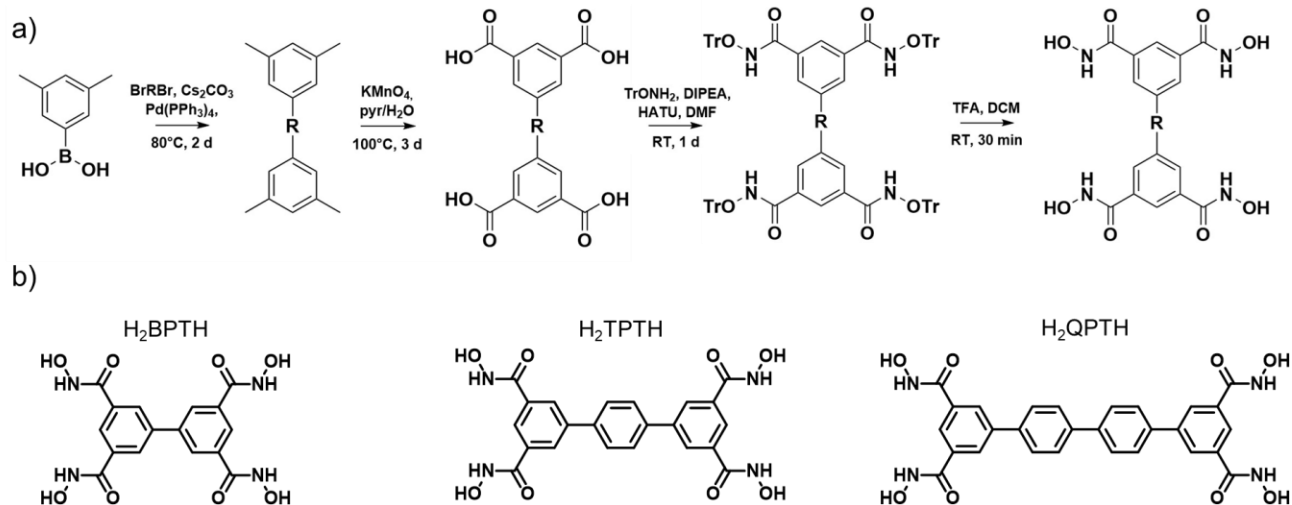


Figure A1.4. a) Scheme for the synthesis of bridging ligands with *m*-H₂bdh head groups. b) Three collinear ligands were synthesized with varying ligand lengths.

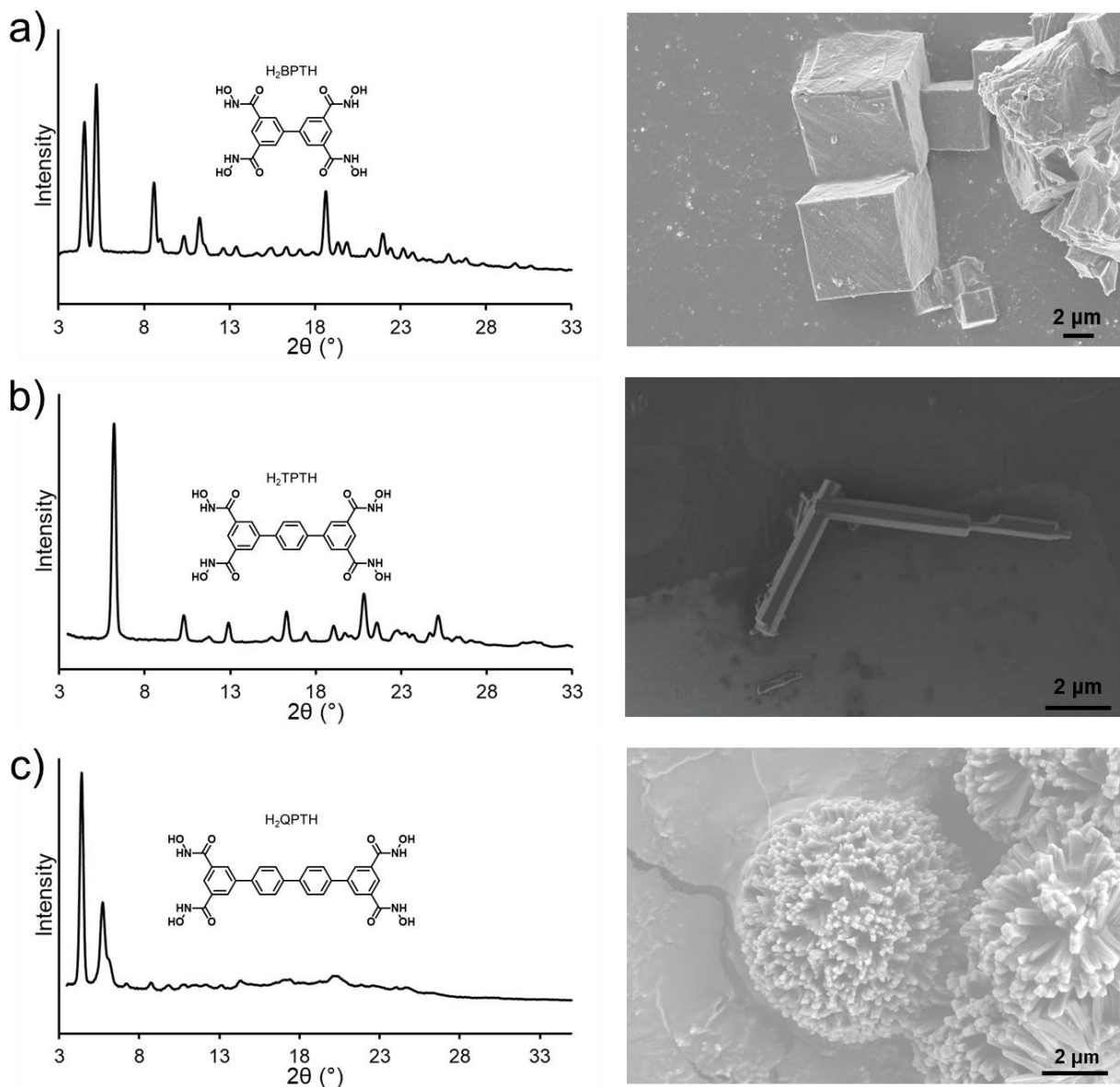


Figure A1.5. Powder X-ray diffraction patterns and SEM images for Fe-BPTH (a), Fe-TPTH (b), and Fe-QPTH (c).

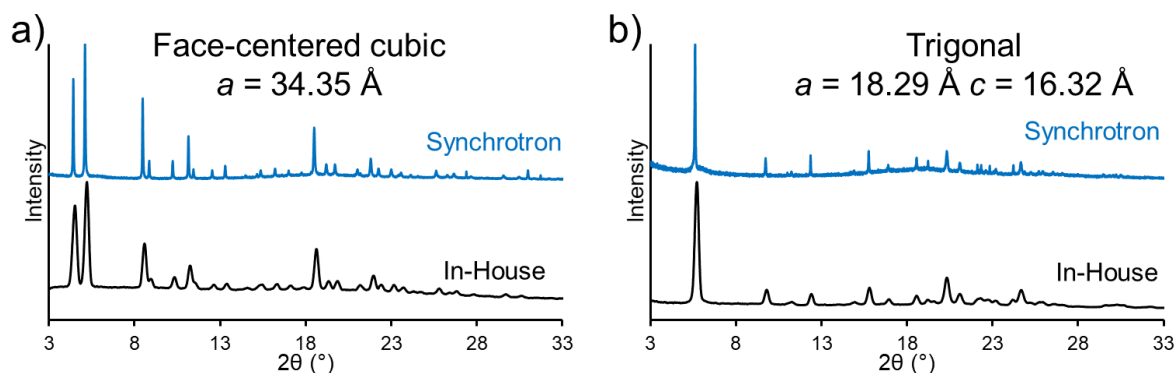


Figure A1.6. A comparison of powder X-ray diffraction patterns for Fe-BPTH and Fe-TPTH. The data was collected either in-house (Bruker SMART Pt135 CCD diffractometer equipped with Cu K α radiation) or at a synchrotron (BL 11-BM at APS) for both Fe-BPTH (a) and Fe-TPTH (b). Synchrotron data was converted to the Cu K α wavelength for comparison.

The size of Fe-BPTH crystals could be significantly increased by adding slow heating and cooling steps to the synthetic protocol. We were able to obtain crystals with a cubic morphology and ca. 120 μm sides, which were suitable for sc-XRD. Two Fe-BPTH crystals were both solved to a resolution of 0.90 \AA using diffraction data obtained at SSRL (Figure A1.7). Similar to the Ca²⁺-dependent ferritin crystals, Fe-BPTH crystals have $F432$ symmetry with ca. $a = 34.42$ \AA . Analysis of these crystal structures reveals that the MOF formed as designed. Each hydroxamate-based SBU is connected to six adjacent SBUs through the bridging ligand. The phenyl-phenyl dihedral angle is 20.4°, which is similar to other MOFs with biphenyl bridging ligands. The Fe-carbonyl and Fe-hydroxyl bond lengths are 2.030 \AA and 1.977 \AA , respectively. Unlike the discrete cluster, which exists as either T , C_3 , or S_4 in solution and a racemic mixture of Δ and Λ enantiomers in the crystal lattice, the MOF connectivity leads to enantiopure (either Δ or Λ) crystals (and the chiral $F432$ space group). Since there is no preference for either the Δ or Λ enantiomer during MOF synthesis, the bulk sample exists as a racemic mixture. The pores formed between individual clusters are approximately 17 \AA in diameter, with 8 \AA apertures. This gives a calculated accessible surface area of ~ 2160 m^2/g . The topology of the MOF is twisted boracite (tbo), with the underlying net consisting of C_3 vertices at each metal and C_4 vertices at each ligand. This is the same topology as the Cu-based MOF HKUST-1.²⁴

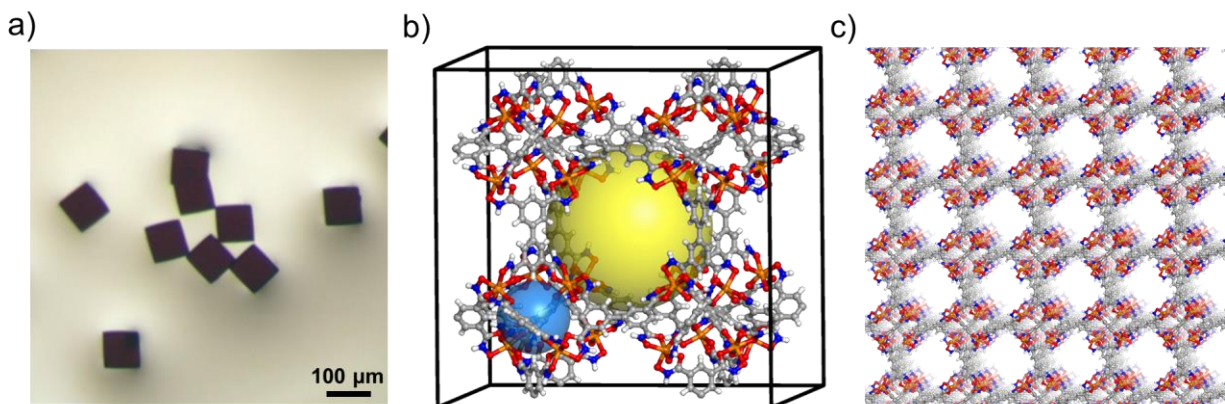


Figure A1.7. Crystal structure of Fe-BPTH. a) Crystals of Fe-BPTH imaged by light microscopy. b) The unit cell of Fe-BPTH. The small pore inside the SBU is shown in blue and the large pore between SBUs is shown in yellow. c) representation of the crystal packing of Fe-BPTH.

Even with significant screening efforts, we were unable to obtain large enough single crystals of Fe-TPTH for sc-XRD. However, our collaborators in the Prof. Xiaodong Zou's lab at Stockholm University were able to determine the structure of the microcrystals through electron diffraction. The crystals have trigonal $P31c$ symmetry with $a = 18.254 \text{ \AA}$ and $c = 16.170 \text{ \AA}$. The crystal structure reveals that the H_4TPTH ligands do not form the expected $Fe_4m\text{-bdh}_6$ clusters and instead pack into interdigitated 2D sheets (Figure 1.8). In this packing, only half of the hydroxamates form the expected tris-chelates with Fe(III). The Fe-carbonyl and Fe-hydroxyl bond lengths are slightly elongated at 2.094 \AA and 2.145 \AA , respectively. The remaining hydroxamates remain uncoordinated and are rotationally disordered in the structure. The 2D sheets are held together primarily through π - π stacking interactions between the phenyl rings of the m -bdh head groups. Multiwalled 1D pores ca. 14 \AA in diameter extend down the c axis of the structure. There is minimal rotational offset between the phenyl rings of each TPDH ligand. Even though this structure was not what we predicted, both the unbound hydroxamates and extensive π - π interactions may allow for unique functionality of Fe-TPTH.

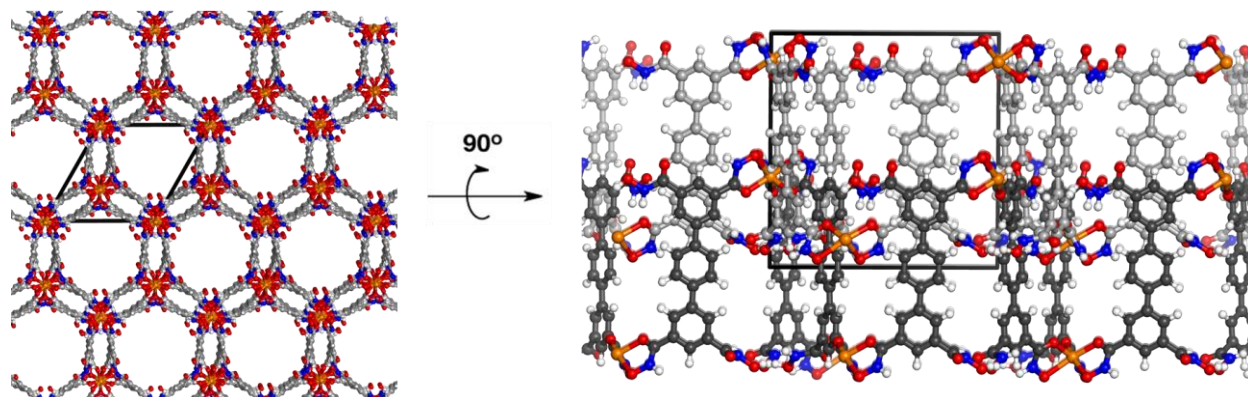


Figure A1.8. Crystal structure of Fe-TPTH. The framework is comprised of multi-walled pores formed through the stacking of interdigitated 2D sheets (shown in light grey or black). Half of the hydroxamates are unbound.

A1.2.3. Stability of Fe-BPTH

The SBUs of Fe-BPTH are constructed entirely from strong metal-hydroxamate interactions, which should render Fe-BPTH stable under harsh conditions. The stability of Fe-BPTH was assessed in water over a wide range of pH values (Figure A1.9 a). As expected, Fe-BPTH retains its crystallinity in all of the tested alkaline conditions, including 5M NaOH. In comparison, carboxylate-based MOFs such as UIO-66 rapidly disorder under similar conditions. Fe-BPTH has similar base stability to the ZIF series, which are among the most stable MOFs under alkaline conditions. Fe-BPTH is also relatively stable in acidic conditions and retains its crystallinity even after being exposed solutions of pH 3.0. At pH 1.0, there is noticeable loss in PXRD peak intensity, which indicates partial decomposition of the framework. The addition of 1M HCl leads to rapid decomposition of the MOF within a few hours. This is not surprising as protonation of the hydroxamate ligand would lead to liberation of the metal ion followed by disassembly. In addition to providing excellent aqueous stability, the strong bonds of the chelate-based node also resist decomposition across a range of coordinating solvents with functional groups that often lead to MOF decomposition. Fe-BPTH is stable in carboxylic acids, primary, secondary, and tertiary amines, as well as all of the organic solvents that were tested (Figure A1.9 b).

Despite the stability of Fe-BPTH in a wide range of solution conditions, activation of this MOF has been extremely challenging. Due to the high boiling point of DEF and the potential for hydrogen bond formation with the amide hydrogen, high temperatures are needed to remove the DEF directly. Although the DEF in the pores can be readily exchanged for a variety of solvents with lower surface tension (i.e., MeOH, DCM, THF, Hexanes), the removal of any of these more volatile solvents leads to complete collapse of the crystalline Fe-BPTH structure. Other activation approaches, such as supercritical CO₂, and lyophilization after solvent exchange into water also led to the collapse of the highly-ordered porous structure (determined by PXRD). DEF molecules

may be needed to stabilize the $\text{Fe}_4m\text{-bdh}_6$ cluster. Further activation experiments are needed to uncover the mechanism behind the structural collapse.

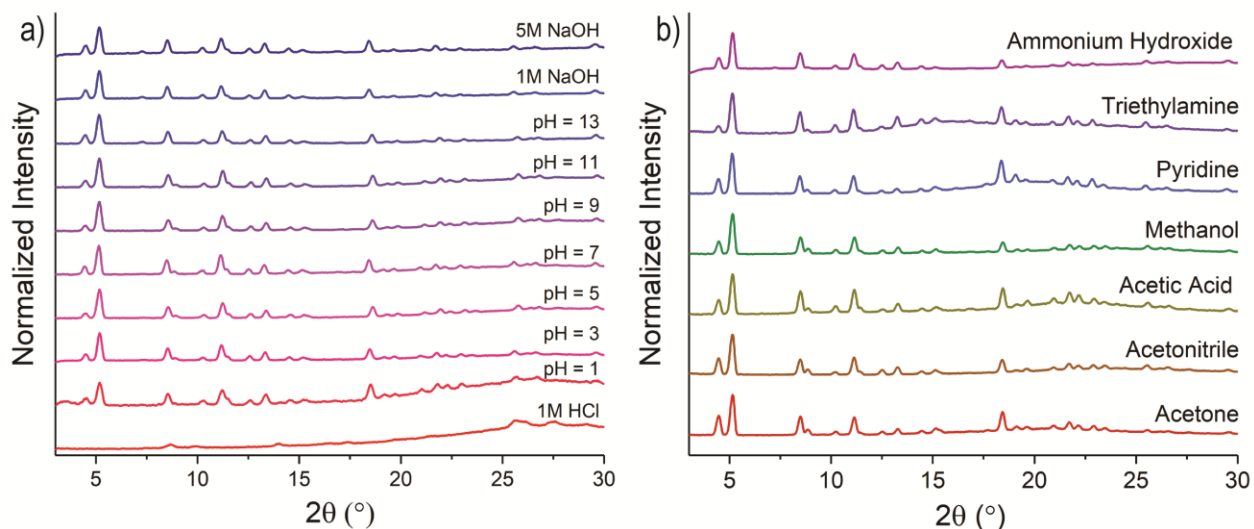


Figure A1.9. Powder x-ray diffraction of Fe-BPTH after soaking in solutions with different pH (a) or in different solvents (b).

A1.2.4. Future research directions and applications of Fe-BPTH and Fe-TPTH

Research efforts into potential applications for both Fe-BPTH and Fe-TPTH are still ongoing. Fe-BPTH has an extremely negative Zeta potential of ca. -44 mV in water (pH 7.0). Preliminary dye uptake studies suggest that Fe-BPTH selectively sequesters small positively charged dyes (methylene blue), while negatively charged dyes (rose bengal) remain in solution. These results, combined with the exceptional solution stability of Fe-BPTH, suggest that the MOF could be utilized to selectively uptake charged molecules in harsh environments. Additionally, the solution stability of Fe-BPTH should allow for post-synthetic modifications based on the deprotonation of the amide hydrogen. Though the structure of Fe-TPTH was not what we had initially envisioned, the MOF has several features that allow for unique functionality. The interdigitated dual-walled tubes that extend throughout the structure may be conducive for electrical conductivity throughout the framework. Currently, collaborators in the Paesani lab at UCSD are investigating if this structure would likely be conducive. The unbound hydroxamates also allow for post-synthetic metalation.

A1.3. Methods

A1.3.1. General considerations

All solvents and reagents were purchased and used without further purification. Mass spectrometry (MS) of small molecules was carried out at the Molecular Mass Spectrometry Facility at UCSD using electrospray ionization (ESI) on a Micromass Quattro Ultima Triple Quadrupole MS. NMR spectra were recorded on either a Varian Mercury spectrometer (500 MHz) or a Bruker AVA spectrometer (300 MHz). NMR data are reported as follows: chemical shift, multiplicity (s = singlet, br = broad), and relative integrated peak area. The spectra were internally referenced to the residual solvent signal (DMSO, δ 2.50 ppm).

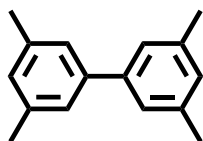
A1.3.2. Synthesis and characterization of the Fe₄*m*-bdh₆ clusters

m-bdh (5.88 mg, 0.03 mmol) and Iron(III) chloride hexahydrate (5.40 mg, 0.02 mmol) were combined in 2 mL of DMF in a 1-dram vial. This vial was placed in a 20 mL scintillation vial, and ca. 10 mL of H₂O was used as the reservoir for crystallization by vapor diffusion. Large cube-shaped crystals formed over three days at room temperature. A 0.24 x 0.24 x 0.24 mm piece of a red crystal was mounted on a Cryoloop with Paratone oil. The single-crystal X-ray diffraction studies were carried out on a Bruker ApexII-Ultra CCD diffractometer equipped with Mo K α radiation ($\lambda = 0.7107 \text{ \AA}$). Data were collected in a nitrogen gas stream at 100(2) K using ω and ϕ scans. Crystal-to-detector distance was 40 mm, and the exposure time was 5 or 60 seconds per frame using a scan width of 0.75°. Data collection was 100 % complete to 25.242° in θ . A total of 94216 reflections were collected covering the indices, $-30 \leq h \leq 22$, $-21 \leq k \leq 32$, $-18 \leq l \leq 34$. 14345 reflections were found to be symmetry independent, with a R_{int} of 0.0624. Indexing and unit cell refinement indicated a primitive cubic lattice. The space group was found to be $P2_13$. 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. All nonhydrogen atoms were

refined anisotropically by full-matrix least-squares (SHELXL-2014). All carbon bonded 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. SQUEEZE was used to remove the poorly ordered solvent (2180 electrons, ca. 30 DMF molecules per unit cell). Crystallographic data are summarized in Table A1..

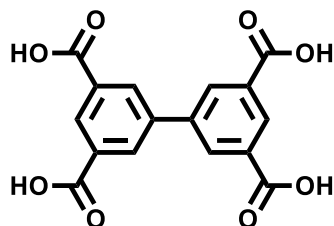
A1.3.3. Synthesis of tetrahydroxamate ligands

The synthesis of the tetracarboxylate precursors was based on published protocols from Lin *et al.* and Blunt *et al.*.²⁵⁻²⁶ The *O*-trityl hydroxylamine was prepared, as described in Chapter 3.



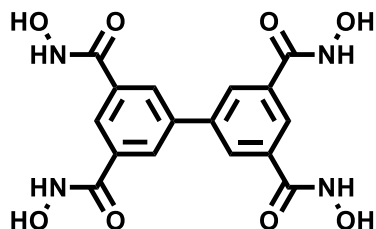
Synthesis of 3,3',5,5'-tetramethyl-1,1'-biphenyl

1-Bromo-3,5-dimethylbenzene (3.67 mL, 27.0 mmol), 3,5-dimethylphenylboronic acid (5.07 g, 33.8 mmol), and Cs_2CO_3 (23.6 g, 72.4 mmol) were mixed in EtOH (175 mL), and the mixture was de-aerated with N_2 . $\text{Pd}(\text{PPh}_3)_4$ (312 mg, 0.27 mmol) was added, and the reaction mixture was refluxed at 80°C for 2 days under N_2 . The reaction mixture was evaporated to dryness, and the residue extracted into CHCl_3 and washed with water. The organic layer was dried over MgSO_4 , and the solvent was removed *in vacuo*. The crude product was purified by silica gel chromatography using 10% DCM in hexanes as the eluent. Pure fractions were combined, and the solvent was removed *in vacuo* yielding 5.57 g of product (98% yield). ^1H NMR: (300 MHz, CDCl_3) δ 7.22 (s, 4H), δ 7.20 (s, 2H), δ 2.40 (s, 12H). ^{13}C NMR: (500 MHz, CDCl_3): δ 141.55, δ 138.23, δ 128.85, δ 125.23, δ 21.56. Measured molecular weight (m/z): 211.07 (Calc.: 211.14) ($\text{M} + \text{H}^+$).



Synthesis of [1,1'-biphenyl]-3,3',5,5'-tetracarboxylic acid

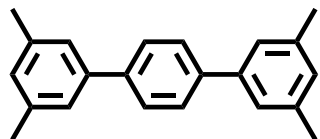
3,3',5,5'-tetramethyl-1,1'-biphenyl (1.68 g, 8.0 mmol) and KMnO_4 (5.1 g, 32.5 mmol) were dissolved in 50 mL of pyridine and 50 mL of H_2O . The reaction was heated to reflux, and 15 portions of KMnO_4 (1.25 g, 7.91 mmol) were added over 2 d. The reaction mixture was filtered over Celite and washed with H_2O . The filtrate was concentrated to 20 mL, and the mixture was filtered again. The filtrate was acidified to pH = 1 using concentrated HCl. The white precipitate was filtered and dried *in vacuo*. Yield: 1.85 g (70%). ^1H NMR: (300 MHz, DMSO-d_6) δ 13.48 (s, 4H), δ 8.52 (s, 2H), δ 8.43 (s, 4H). ^{13}C NMR: (500 MHz, DMSO-d_6): δ 166.38, δ 139.29, δ 132.37, δ 131.45, δ 129.52. Measured molecular weight (m/z): 328.97 (Calc.: 329.25) ($\text{M} - \text{H}^+$).



Synthesis of $N^\beta, N^{\beta'}, N^\delta, N^{\delta'}$ -tetrahydroxy-[1,1'-biphenyl]-3,3',5,5'-tetracarboxamide (H_4BPTH)

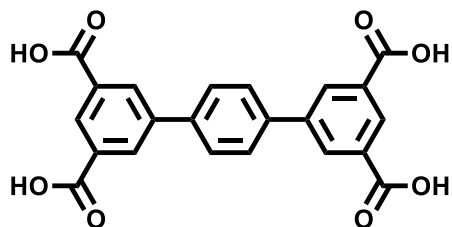
[1,1'-biphenyl]-3,3',5,5'-tetracarboxylic acid (1.85 g, 5.6 mmol), HATU (9.37 g, 24.7 mmol), DIPEA (11.7 mL, 67.2 mmol), and *O*-tritylhydroxylamine (7.4 g, 26.9 mmol) were added to 100 mL of DMF and stirred for 12 h at room temperature. The solution was added to 100 mL of ethyl acetate and extracted with H_2O (3 x 50 mL). The precipitate was collected by filtration and washed with 30 mL ethyl acetate. The precipitate was suspended in 40 mL dichloromethane with 10% (v/v) trifluoroacetic acid and stirred for 30 min at room temperature. The precipitate was filtered, washed with 10 mL dichloromethane, and dried *in vacuo* yielding 1.15 g of product (53% yield). ^1H NMR: (300 MHz, DMSO-d_6) δ 11.47 (s, 4H), δ 9.20 (s, 4H), δ 8.28 (s, 4H), δ 8.18 (s, 4H)

^{13}C NMR: (500 MHz, DMSO-d_6): δ 163.48, δ 139.14, δ 134.05, δ 127.82, δ 125.61. Measured molecular weight (m/z): 389.06 (Calc.: 389.08) ($\text{M} - \text{H}^+$).



Synthesis of 3,3',5,5''-tetramethyl-1,1':4,1''-terphenyl

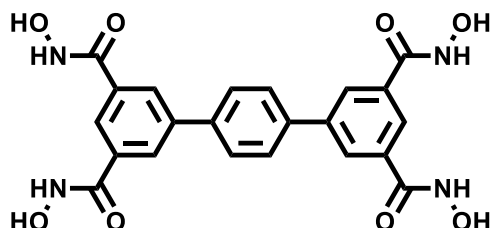
(3,5-dimethylphenyl)boronic acid (4 g, 26.7 mmol), 1,4-dibromobenzene (2.52 g, 10.68 mmol), and Cs_2CO_3 (9.32 g, 28.6 mmol) were combined in 70 mL of EtOH and purged with N_2 for 20 min. $\text{Pd}(\text{Ph}_3\text{P})_4$ (123 mg, 0.107 mmol) was added, and the reaction was refluxed at 80 °C for 2 days under N_2 . The solvent was removed in vacuo, and the product was extracted into CHCl_3 and washed with H_2O . The crude product was purified by silica gel chromatography using 0-5% EtOAc in Hexanes as the eluent. Pure fractions were combined, and the solvent was removed in vacuo yielding 2.47 g of product (81% yield). ^1H NMR: (300 MHz, DMSO-d_6) δ 7.71 (s, 4H), δ 7.32 (s, 4H), δ 7.01 (s, 2H), δ 2.35 (s, 12H).



Synthesis of [1,1':4,1''-terphenyl]-3,3',5,5''-tetracarboxylic acid

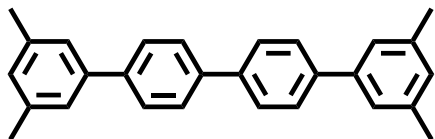
3,3',5,5''-tetramethyl-1,1':4,1''-terphenyl (3 g, 10.47 mmol) and KMnO_4 (6.74 g, 42.6 mmol) were dissolved in 150 mL of pyridine and 150 mL of H_2O . The reaction was heated to reflux, and 15 portions of KMnO_4 (3.38 g, 21.3 mmol) were added over 3 days. The reaction mixture was filtered over Celite and washed with H_2O . The filtrate was dried *in vacuo* and redissolved in H_2O . The filtrate was acidified to pH = 1 using concentrated HCl. The white

precipitate was filtered, washed with H₂O and MeOH, and dried *in vacuo* yielding 4.26 g of product (100% yield). ¹H NMR: (300 MHz, CDCl₃) δ 8.49(s, 4H), δ 8.46 (s, 4H), δ 7.92 (s, 4H).



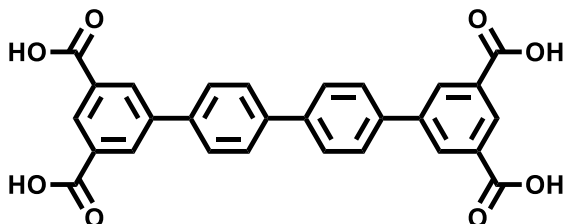
Synthesis of *N*^β,*N*^{β'},*N*^{β''},*N*^{β'''}-tetrahydroxy-[1,1':4',1''-terphenyl]-3,3'',5,5''-tetracarboxamide (H₄TPTH)

[1,1':4',1''-terphenyl]-3,3'',5,5''-tetracarboxylic acid (1.36 g, 3.35 mmol), HATU (5.6 g, 14.73 mmol), DIPEA (7.01 mL, 40.1 mmol), and *O*-tritylhydroxylamine (4.42 g, 16.05 mmol) were added to 100 mL of DMF and stirred for 12 h at room temperature. The solution was added to 100 mL of ethyl acetate and extracted with H₂O (3 x 50 mL). The precipitate was collected by filtration and washed with 30 mL ethyl acetate. The precipitate was suspended in 40 mL dichloromethane with 10% (v/v) trifluoroacetic acid and stirred for 30 min at room temperature. The remaining precipitate was filtered, washed with 10 mL dichloromethane, and dried *in vacuo* yielding 1.08 g of product (69% yield). ¹H NMR: (400 MHz, DMSO-*d*₆) δ 11.47 (s, 4H), δ 9.21 (s, 4H), δ 8.23 (s, 4H), δ 8.19 (s, 2H), δ 7.96 (s, 2H).



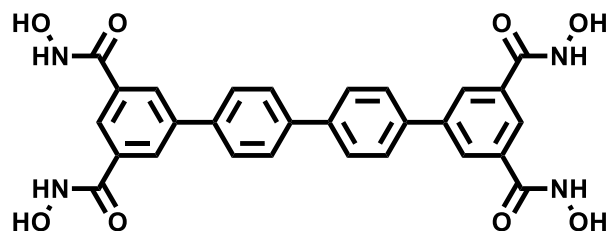
Synthesis of 3,3',5,5'-tetramethyl-1,1':4',1'':4'',1'''-quaterphenyl

(3,5-dimethylphenyl)boronic acid (3.6 g, 24.01 mmol), 4,4'-dibromo-1,1'-biphenyl (3 g, 9.62 mmol), and Cs_2CO_3 (8.4 g, 25.8 mmol) were combined in 150 mL of EtOH and purged with N_2 for 20 min. $\text{Pd}(\text{Ph}_3\text{P})_4$ (111 mg, 0.096 mmol) was added and the reaction was refluxed at 80 °C for 3 d under N_2 . The solvent was removed in vacuo and the product was extracted with CHCl_3 and washed with H_2O . The product was purified by silica gel chromatography using 0-50 % DCM in Hexanes as the eluent. Pure fractions were combined and the solvent was removed in vacuo yielding 3.01 g (86% yield). ^1H NMR: (400 MHz, CDCl_3) δ 8.64 (d, 4H), δ 7.51 (dd, 4H), δ 7.25 (s, 4H), δ 7.09 (s, 2H), δ 2.40 (s, 12H).



Synthesis of [1,1':4',1'':4'',1'''-quaterphenyl]-3,3',5,5'-tetracarboxylic acid

3,3',5,5'-tetramethyl-1,1':4',1'':4'',1'''-quaterphenyl (3 g, 8.28 mmol), KMnO_4 (5.32 g, 33.7 mmol) were dissolved in 120 mL of pyridine and 30 mL of H_2O . The reaction was heated to reflux, and 15 portions of KMnO_4 (2.66 g, 16.9 mmol) were added over 3 days. The reaction mixture was filtered over Celite while hot and washed with H_2O . The filtrate was dried *in vacuo* and redissolved in H_2O and filtered again. The filtrate was acidified to pH = 1 using conc. HCl. The white precipitate was filtered, washed with H_2O and MeOH, and dried *in vacuo* yielding 2.85 g (71% Yield). ^1H NMR: (300 MHz, DMSO-d_6) δ 8.48 (s, 4H), δ 8.45 (s, 2H), δ 7.89 (s, 8H).



Synthesis of $N^6, N^{6''}, N^6, N^{6'''}$ -tetrakis(trityloxy)-[1,1':4',1'':4'', 1'''-quaterphenyl]-3,3''',5,5'''-tetracarboxamide (H_4QPTH)

[1,1':4',1'':4'', 1'''-quaterphenyl]-3,3''',5,5'''-tetracarboxylic acid (2.9 g, 6.01 mmol), HATU (10.06 g, 26.4 mmol), DIPEA (12.60 mL, 72.1 mmol), and *O*-tritylhydroxylamine (7.94 g, 28.9 mmol) were added to 100 mL of DMF and stirred overnight at room temperature. The solution was added to 100 mL of ethyl acetate and extracted with H_2O (3 x 50 mL). The precipitate was collected by filtration and washed with 30 mL ethyl acetate. The precipitate was suspended in 40 mL dichloromethane with 10% (v/v) trifluoroacetic acid and stirred for 30 min at room temperature. The remaining precipitate was filtered, washed with 10 mL dichloromethane, and dried *in vacuo* yielding 3.1 g of product (95% yield). 1H NMR: (400 MHz, $DMSO-d_6$) δ 11.48 (s, 4H), δ 9.21 (s, 4H), δ 8.23 (s, 4H), δ 8.19 (s, 2H), δ 7.94 (s, 8H).

A1.3.4. Synthesis of hydroxamate-based MOFs

Fe-BPTH: $FeCl_3$ (18.0 mg, 0.067 mmol) and H_4BPTH (9.8 mg, 0.025 mmol) were suspended in a mixed solution of containing DEF (1.85 mL) and 1,4-dioxane (0.15 mL) in a 1-dram vial. The sample was slowly heated to 70 °C over 5 h, incubated at 70 °C for 48 h, and slowly cooled to 25 °C over 8 h yielding a red precipitate. The solution was removed, and the crystalline precipitate was washed twice with 2 mL of fresh DEF and stored in DEF at room temperature.

Fe-TPTH: $FeCl_3$ (12.3 mg, 0.033 mmol) and H_4TPTH (11.7 mg, 0.025 mmol) were suspended in a mixed solution of containing DMF (1.80 mL) and acetonitrile (0.20 mL) in a 1-

dram vial. The sample was slowly heated to 70 °C over 5 h, incubated at 70 °C for 48 h, and slowly cooled to 25 °C over 8 h yielding a red precipitate. The solution was removed, and the crystalline precipitate was washed twice with 2 mL of fresh DEF and stored in DEF at room temperature.

Fe-QPTH: FeCl₃ (12.3 mg, 0.067 mmol) and H₄QPTH (13.17 mg, 0.025 mmol) was suspended in a mixed solution containing DMF (2.00 mL) in a 1-dram vial. The sample was slowly heated to 70 °C over 5 h, incubated at 70 °C for 48 h, and slowly cooled to 25 °C over 8 h yielding a red precipitate. The solution was removed, and the crystalline precipitate was washed twice with 2 mL of fresh DEF and stored in DEF at room temperature.

A1.3.5. Structural determination of hydroxamate-based MOFs

Crystals with approximate dimensions 0.12 x 0.12 x 0.12 mm were mounted on Cryoloops with Paratone oil and flash-frozen in liquid N₂. Crystals were shipped to ALS Beamline 8-3-1 for analysis. Data were collected in a nitrogen gas stream at 100(2) K using φ scans. A radiation wavelength of $\lambda = 0.88557 \text{ \AA}$ was used, and data were collected on a Pilatus 3X 6M detector. Crystal-to-detector distance was 165.5 mm, and the exposure time was 0.6 or 1 s per frame using a scan width of 0.15° or 0.5°, respectively. 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. All nonhydrogen atoms were refined anisotropically by full-matrix least-squares (SHELXL-2014). All carbon bonded 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. SQUEEZE was used to remove the disordered solvent. Crystallographic data are summarized in Table A1.1.

Table A1.1 X-ray data processing and refinement statistics.

	Fe ₄ m-bdh ₆	Δ-Fe-BPTH	Λ-Fe-BPTH
Space group	<i>P</i> 2 ₁ 3	<i>F</i> 432	<i>F</i> 432
<i>a</i> , <i>b</i> , <i>c</i> (Å)	27.604(4)	34.4309(7)	34.4159(19)
α, β, γ (°)	90	90	90
Volume (Å ³)	21034(10)	40817(2)	40764(7)
Resolution (Å ⁻¹)	0.80	0.90	0.90
Temperature (K)	100(2)	100(2)	100(2)
Resolution (Å)	0.80	0.90	0.90
ρ Calc. (mg/m ³)	1.094	0.450	.450
μ (mm ⁻¹)	0.609	0.552	.553
Completeness (%)	100	99.6	99.7
Unique Reflections	14345	2475	2474
Restraints	6	0	0
Parameters	694	69	68
R ₁	0.570	0.414	0.0498
wR ₂	0.1368	0.1278	0.1398
GOF	1.053	1.082	1.014
Flack	0.005(4)	0.054(19)	0.44(16)
e ⁻ removed with SQUEEZE	2180	10885	8567

A1.3.6. Assessing the stability of Fe-BPTH in solution

A small amount of Fe-BPTH (< 5 mg) was transferred with minimal DMF to an Eppendorf tube. A series of solutions with different pH values (1, 3, 5, 7, 9, 11, 13) was prepared by adding 1M HCl or NaOH to reach the desired pH. All other solutions were used as purchased. For each solution, Fe-BPTH was resuspended in the solution and exchanged 3x over 1 h. Fe-BPTH was left in the final solution for one week before the decomposition was assessed by powder X-ray diffraction. A small amount of Fe-BPTH was loaded with 50 uL of solution into a 1 mm glass capillary (Hampton Research) and briefly centrifuged to pellet the crystallites. Data were collected in transmission mode on a Bruker SMART Pt135 CCD diffractometer equipped with Cu Kα radiation (λ = 1.54178 Å) calibrated with AgBeH. Images were stitched together and azimuthally averaged using DIFFRAC.EVA.

A1.4. Acknowledgments

Appendix 1, in part, is currently being prepared for publication with contributions from the following authors: Chiong, J. A.; Bailey, J. B.; Zhu, J.; Xu, W.; Cohen, S.; Tezcan, F. A. The dissertation author will be a co-primary author of this publication. The author would like to highlight the hard work of Jerika Chiong, who was the primary driving force behind this project, and Prof. Seth Cohen, who initiated the product and provided valuable guidance. Appendix 1 also contains electron diffraction data that was collected by Zhehao Huang and Laura Sampersky in the lab of Xiaodong Zho at Stockholm University.

A1.5. References

1. Furukawa, H.; Cordova, K. E.; O'Keeffe, M.; Yaghi, O. M., The Chemistry and Applications of Metal-Organic Frameworks. *Science* **2013**, *341* (6149), 1230444.
2. Sumida, K.; Rogow, D. L.; Mason, J. A.; McDonald, T. M.; Bloch, E. D.; Herm, Z. R.; Bae, T. H.; Long, J. R., Carbon dioxide capture in metal-organic frameworks. *Chem. Rev.* **2012**, *112* (2), 724-81.
3. Sun, C.-Y.; Qin, C.; Wang, X.-L.; Su, Z.-M., Metal-organic frameworks as potential drug delivery systems. *Expert Opinion on Drug Delivery* **2013**, *10* (1), 89-101.
4. Yuan, S.; Feng, L.; Wang, K.; Pang, J.; Bosch, M.; Lollar, C.; Sun, Y.; Qin, J.; Yang, X.; Zhang, P.; Wang, Q.; Zou, L.; Zhang, Y.; Zhang, L.; Fang, Y.; Li, J.; Zhou, H.-C., Stable Metal–Organic Frameworks: Design, Synthesis, and Applications. *Advanced Materials* **2018**, *30* (37), 1704303.
5. Kalmutzki, M. J.; Hanikel, N.; Yaghi, O. M., Secondary building units as the turning point in the development of the reticular chemistry of MOFs. *Science Advances* **2018**, *4* (10), eaat9180.
6. Li, H.; Eddaoudi, M.; Groy, T. L.; Yaghi, O. M., Establishing Microporosity in Open Metal–Organic Frameworks: Gas Sorption Isotherms for Zn(BDC) (BDC = 1,4-Benzenedicarboxylate). *Journal of the American Chemical Society* **1998**, *120* (33), 8571-8572.
7. Li, H.; Eddaoudi, M.; O'Keeffe, M.; Yaghi, O. M., Design and synthesis of an exceptionally stable and highly porous metal-organic framework. *Nature* **1999**, *402* (6759), 276-279.
8. Park, K. S.; Ni, Z.; Côté, A. P.; Choi, J. Y.; Huang, R.; Uribe-Romo, F. J.; Chae, H. K.; O'Keeffe, M.; Yaghi, O. M., Exceptional chemical and thermal stability of zeolitic imidazolate frameworks. *Proceedings of the National Academy of Sciences* **2006**, *103* (27), 10186.
9. Serre, C.; Millange, F.; Thouvenot, C.; Noguès, M.; Marsolier, G.; Louër, D.; Férey, G., Very Large Breathing Effect in the First Nanoporous Chromium(III)-Based Solids: MIL-53 or

$\text{CrIII}(\text{OH})\cdot\{\text{O}_2\text{C}-\text{C}_6\text{H}_4-\text{CO}_2\}\cdot\{\text{HO}_2\text{C}-\text{C}_6\text{H}_4-\text{CO}_2\text{H}\}_x\cdot\text{H}_2\text{O}_y$. *Journal of the American Chemical Society* **2002**, *124* (45), 13519-13526.

10. Férey, G.; Serre, C.; Mellot-Draznieks, C.; Millange, F.; Surblé, S.; Dutour, J.; Margiolaki, I., A Hybrid Solid with Giant Pores Prepared by a Combination of Targeted Chemistry, Simulation, and Powder Diffraction. *Angewandte Chemie International Edition* **2004**, *43* (46), 6296-6301.

11. Pereira, C. F.; Howarth, A. J.; Vermeulen, N. A.; Almeida Paz, F. A.; Tomé, J. P. C.; Hupp, J. T.; Farha, O. K., Towards hydroxamic acid linked zirconium metal–organic frameworks. *Materials Chemistry Frontiers* **2017**, *1* (6), 1194-1199.

12. Padial, N. M.; Castells-Gil, J.; Almora-Barrios, N.; Romero-Angel, M.; da Silva, I.; Barawi, M.; García-Sánchez, A.; de la Peña O’Shea, V. A.; Martí-Gastaldo, C., Hydroxamate Titanium–Organic Frameworks and the Effect of Siderophore-Type Linkers over Their Photocatalytic Activity. *Journal of the American Chemical Society* **2019**, *141* (33), 13124-13133.

13. Abrahams, B. F.; Hudson, T. A.; McCormick, L. J.; Robson, R., Coordination Polymers of 2,5-Dihydroxybenzoquinone and Chloranilic Acid with the (10,3)-a Topology. *Crystal Growth & Design* **2011**, *11* (7), 2717-2720.

14. Nakabayashi, K.; Ohkoshi, S.-i., Monometallic Lanthanoid Assembly Showing Ferromagnetism with a Curie Temperature of 11 K. *Inorganic Chemistry* **2009**, *48* (18), 8647-8649.

15. Hmadeh, M.; Lu, Z.; Liu, Z.; Gándara, F.; Furukawa, H.; Wan, S.; Augustyn, V.; Chang, R.; Liao, L.; Zhou, F.; Perre, E.; Ozolins, V.; Suenaga, K.; Duan, X.; Dunn, B.; Yamamoto, Y.; Terasaki, O.; Yaghi, O. M., New Porous Crystals of Extended Metal-Catecholates. *Chemistry of Materials* **2012**, *24* (18), 3511-3513.

16. Padial, N. M.; Castells-Gil, J.; Almora-Barrios, N.; Romero-Angel, M.; da Silva, I.; Barawi, M.; García-Sánchez, A.; de la Peña O’Shea, V. A.; Martí-Gastaldo, C., Hydroxamate Titanium–Organic Frameworks and the Effect of Siderophore-Type Linkers over Their Photocatalytic Activity. *Journal of the American Chemical Society* **2019**.

17. Nguyen, N. T. T.; Furukawa, H.; Gándara, F.; Trickett, C. A.; Jeong, H. M.; Cordova, K. E.; Yaghi, O. M., Three-Dimensional Metal-Catecholate Frameworks and Their Ultrahigh Proton Conductivity. *Journal of the American Chemical Society* **2015**, *137* (49), 15394-15397.

18. Sontz, P. A.; Bailey, J. B.; Ahn, S.; Tezcan, F. A., A metal organic framework with spherical protein nodes: rational chemical design of 3D protein crystals. *Journal of the American Chemical Society* **2015**, *137* (36), 11598-11601.

19. Bailey, J. B.; Zhang, L.; Chiong, J. A.; Ahn, S.; Tezcan, F. A., Synthetic Modularity of Protein-Metal-Organic Frameworks. *Journal of the American Chemical Society* **2017**, *139* (24), 8160-8166.

20. Prakash, M. J.; Lah, M. S., Metal–organic macrocycles, metal–organic polyhedra and metal–organic frameworks. *Chemical Communications* **2009**, (23), 3326-3341.

21. Yeh, R. M.; Xu, J.; Seeber, G.; Raymond, K. N., Large M₄L₄ (M = Al(III), Ga(III), In(III), Ti(IV)) Tetrahedral Coordination Cages: an Extension of Symmetry-Based Design. *Inorganic Chemistry* **2005**, *44* (18), 6228-6239.
22. Abrahams, B. F.; FitzGerald, N. J.; Robson, R., Cages with Tetrahedron-Like Topology Formed from the Combination of Cyclotricatechylene Ligands with Metal Cations. *Angewandte Chemie International Edition* **2010**, *49* (16), 2896-2899.
23. Beissel, T.; Powers, R. E.; Raymond, K. N., Symmetry-Based Metal Complex Cluster Formation. *Angewandte Chemie International Edition in English* **1996**, *35* (10), 1084-1086.
24. Lin, K.-S.; Adhikari, A. K.; Ku, C.-N.; Chiang, C.-L.; Kuo, H., Synthesis and characterization of porous HKUST-1 metal organic frameworks for hydrogen storage. *International Journal of Hydrogen Energy* **2012**, *37* (18), 13865-13871.
25. Lin, X.; Telepeni, I.; Blake, A. J.; Dailly, A.; Brown, C. M.; Simmons, J. M.; Zoppi, M.; Walker, G. S.; Thomas, K. M.; Mays, T. J.; Hubberstey, P.; Champness, N. R.; Schröder, M., High Capacity Hydrogen Adsorption in Cu(II) Tetracarboxylate Framework Materials: The Role of Pore Size, Ligand Functionalization, and Exposed Metal Sites. *Journal of the American Chemical Society* **2009**, *131* (6), 2159-2171.
26. Blunt, M. O.; Russell, J. C.; Gimenez-Lopez, M. d. C.; Taleb, N.; Lin, X.; Schröder, M.; Champness, N. R.; Beton, P. H., Guest-induced growth of a surface-based supramolecular bilayer. *Nature Chemistry* **2011**, *3* (1), 74-78.

Appendix 2. NMR Spectra for Synthesized Compounds

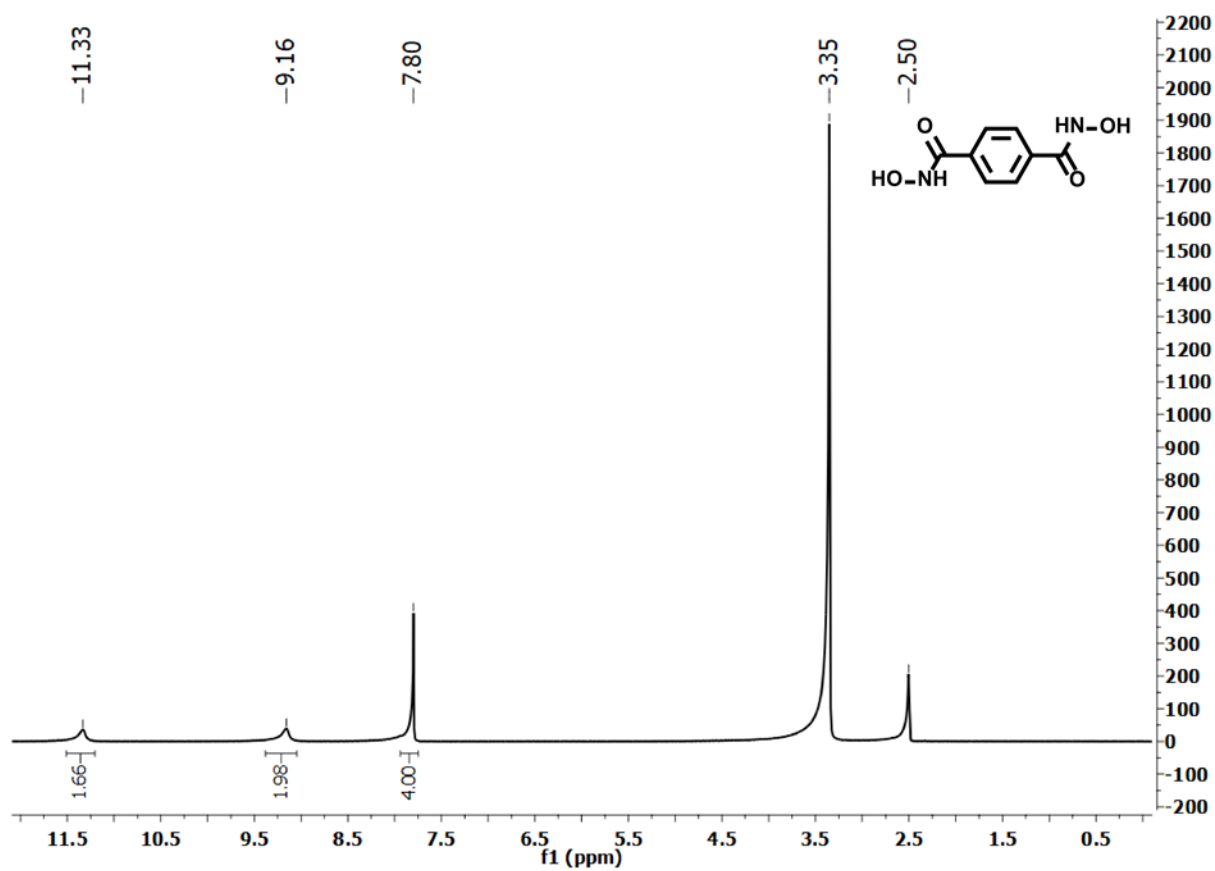


Figure A2.1. ¹H NMR spectrum of *N*¹,*N*⁴-dihydroxyterephthalamide (*p*-H₂bdh) in DMSO-d₆.

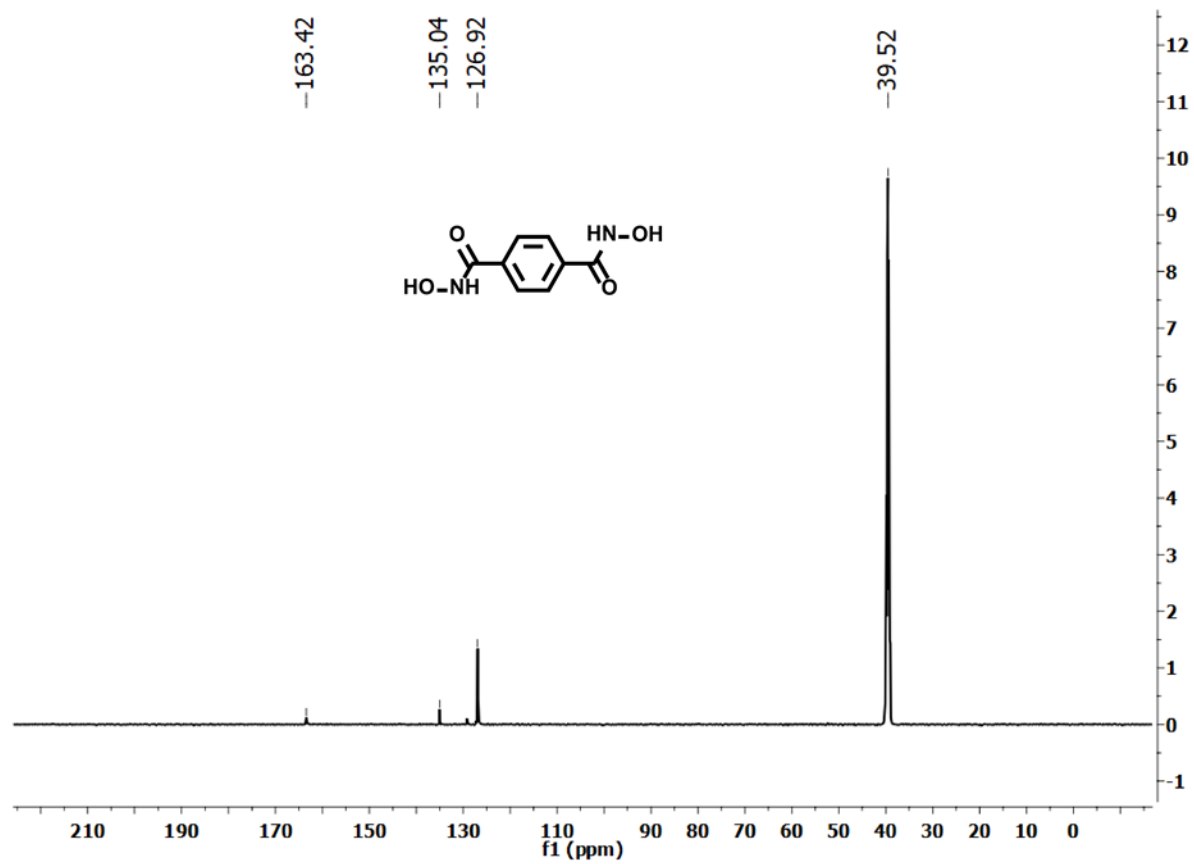


Figure A2.2. ¹³C NMR spectrum of *N*¹,*N*⁴-dihydroxyterephthalamide (*p*-H₂bdh) in DMSO-d₆.

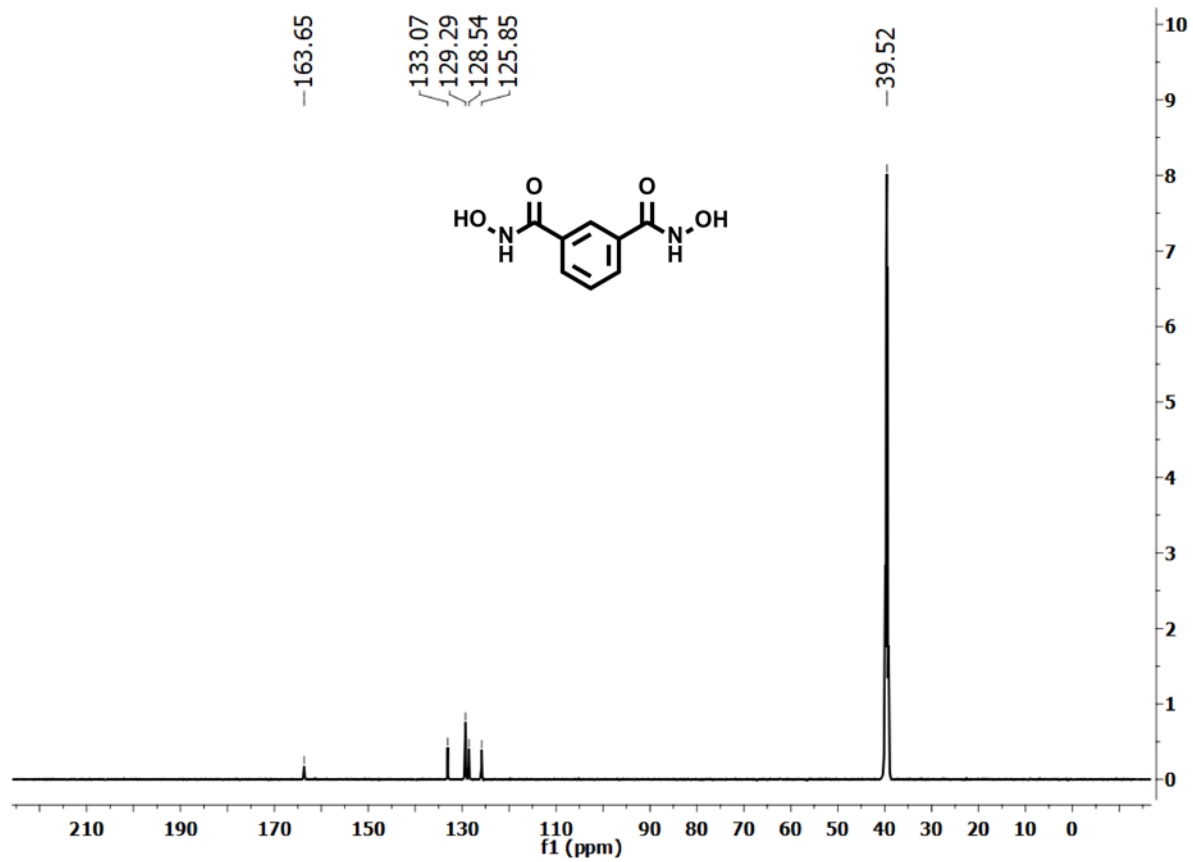


Figure A2.4. ¹³C NMR spectrum of *N*¹,*N*³-dihydroxyisophthalamide (*m*-H₂bdh) in DMSO-d₆.

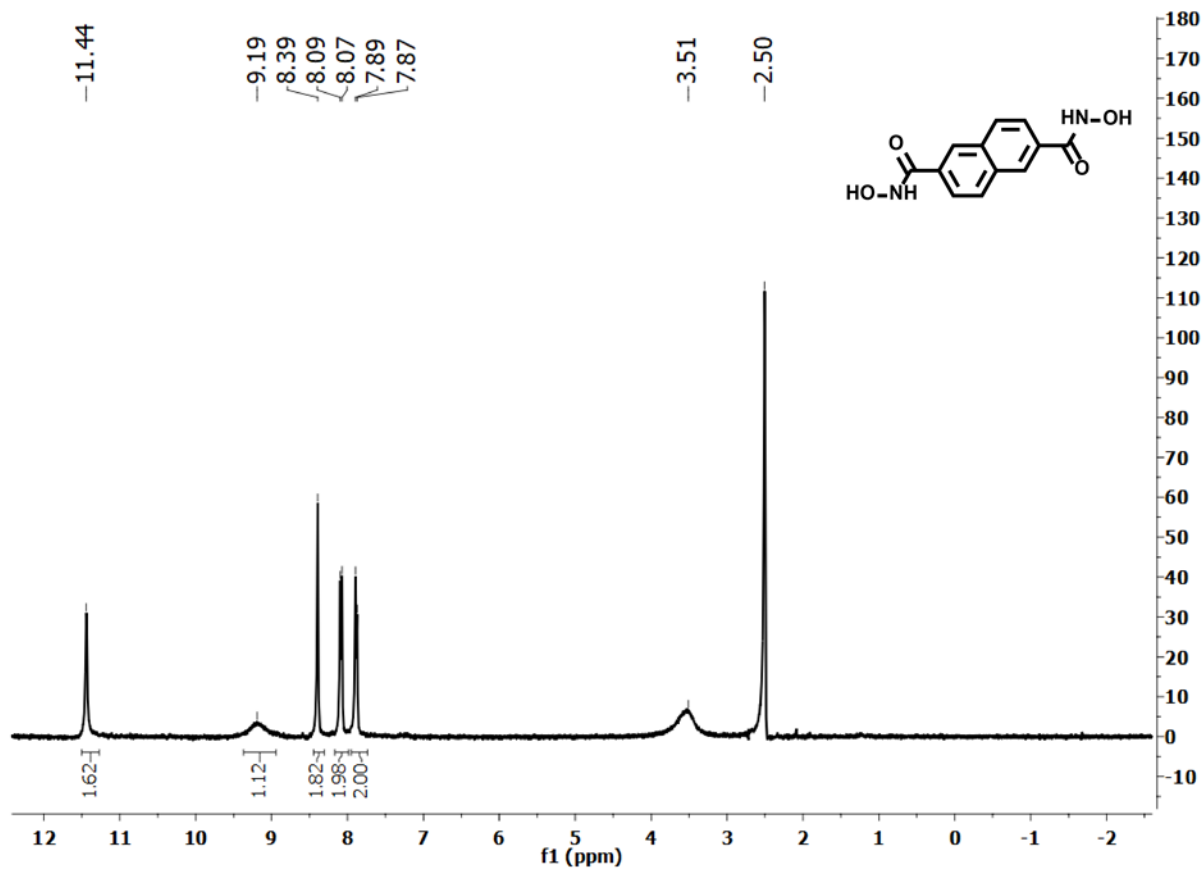


Figure A2.5. ¹H NMR spectrum of *N*²,*N*⁶-dihydroxynaphthalene-2,6-dicarboxamide (H₂ndh) in DMSO-d₆.

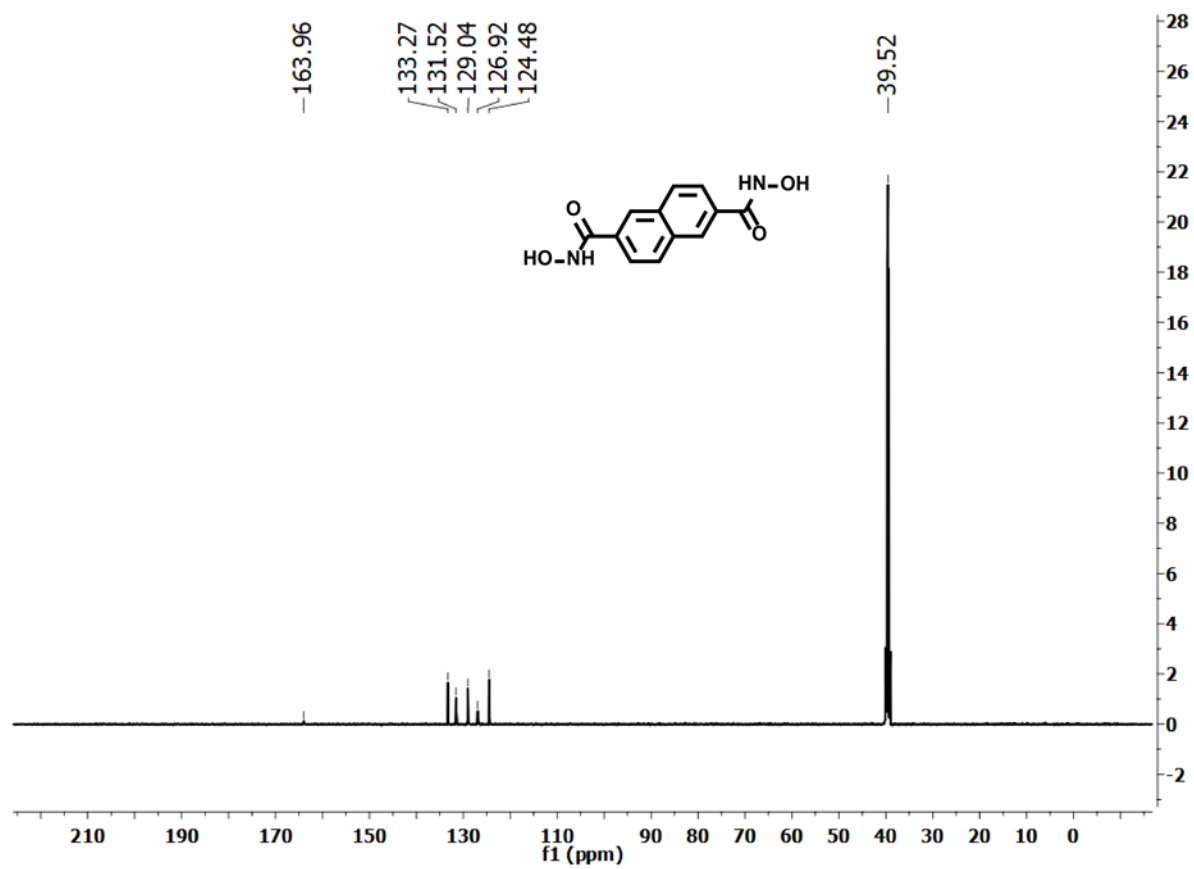


Figure A2.6. ^{13}C NMR spectrum of N^2,N^6 -dihydroxynaphthalene-2,6-dicarboxamide (H₂ndh) in DMSO-d₆.

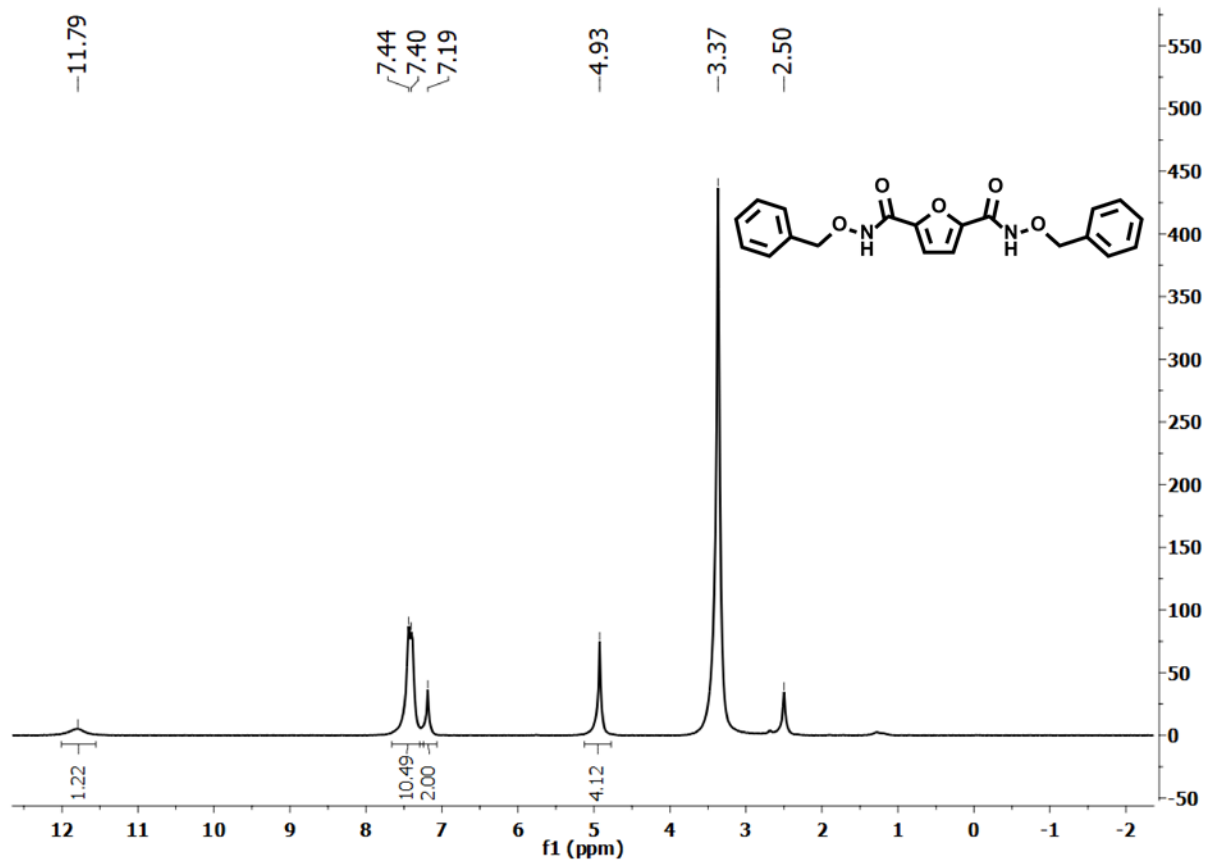


Figure A2.7. ¹H NMR spectrum of *N*²,*N*⁶-bis(benzyloxy)furan-2,5-dicarboxamide (2,5-Bn₂fdh) in DMSO-d₆.

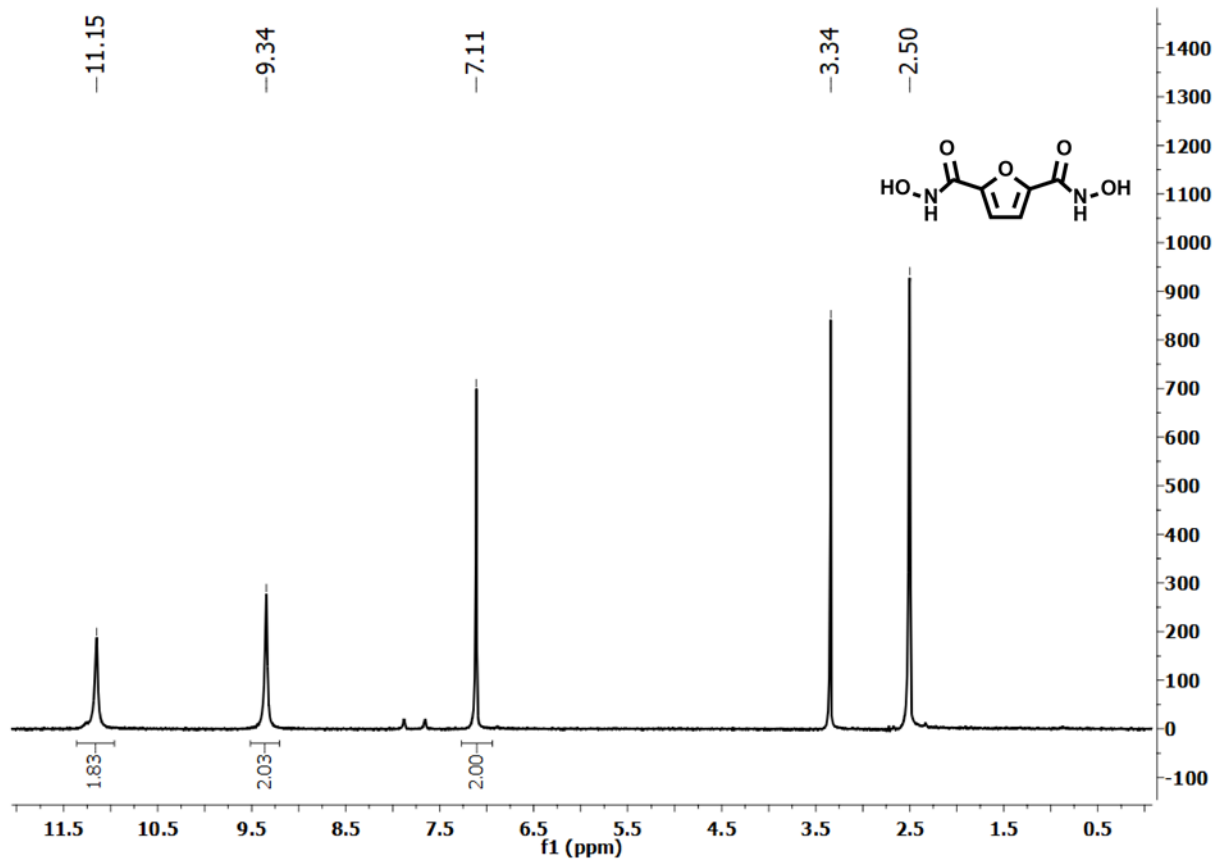


Figure A2.8. ¹H NMR spectrum of *N*²,*N*⁵-dihydroxyfuran-2,5-dicarboxamide (2,5-H₂fdh) in DMSO-d₆.

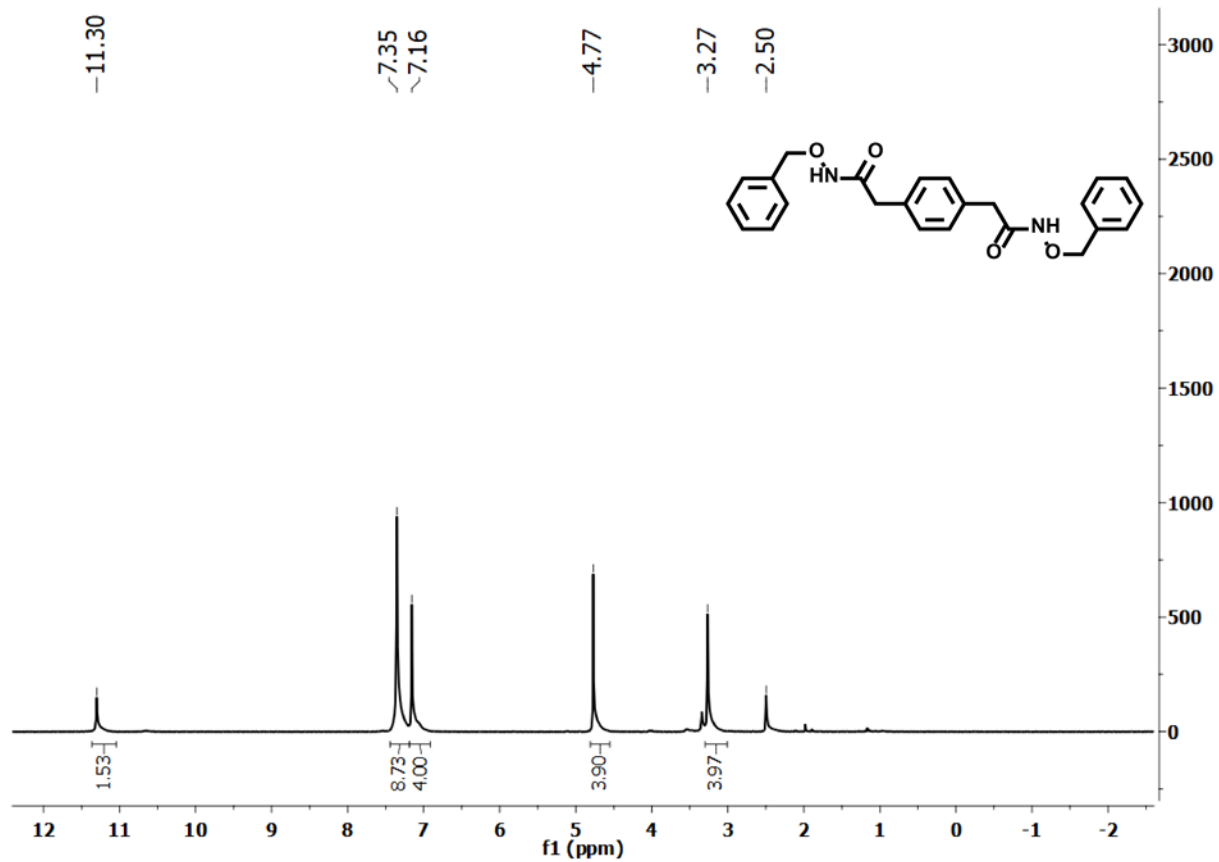


Figure A2.9. ¹H NMR spectrum of 2,2'-(1,4-phenylene)bis(*N*-(benzyloxy)acetamide) (*p*-Bn₂xdh) in DMSO-*d*₆.

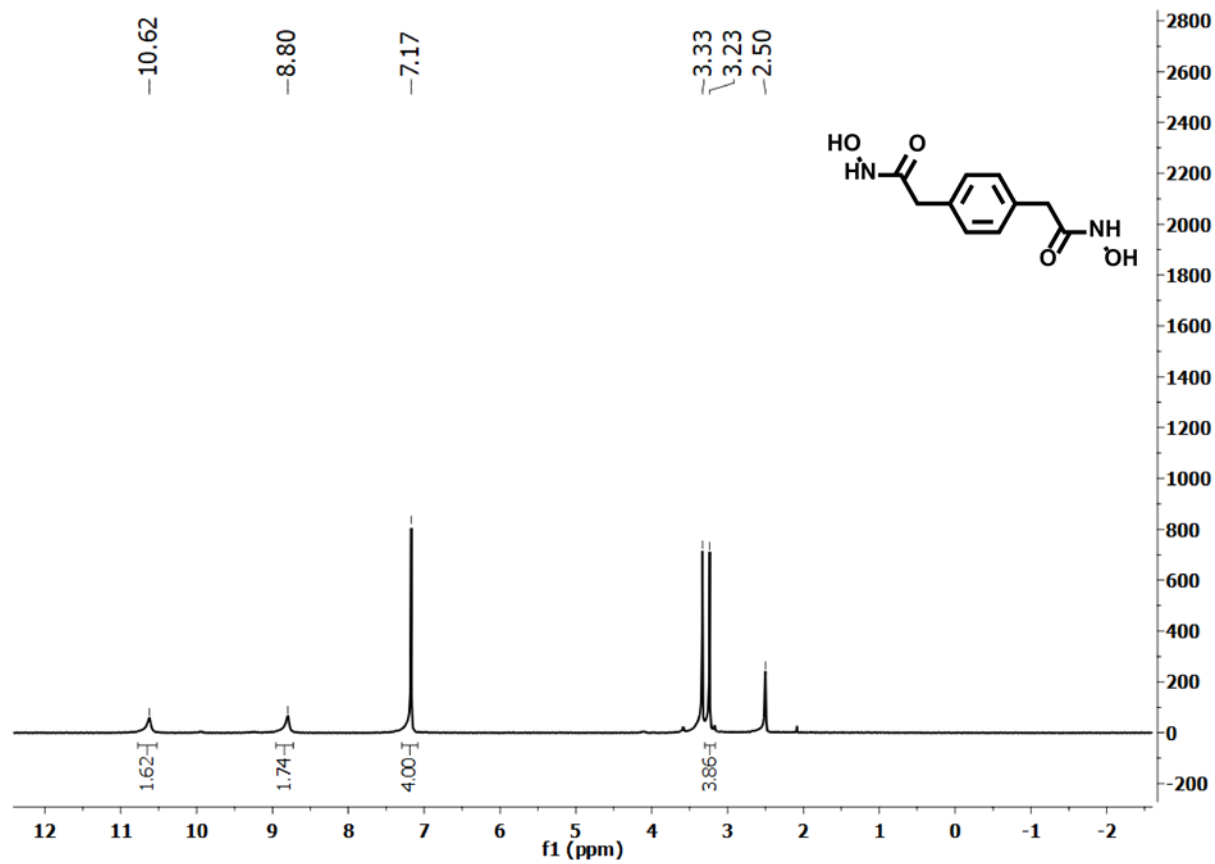


Figure A2.10. NMR spectrum of 2,2'-(1,4-phenylene)bis(*N*-hydroxyacetamide) (*p*-H₂xdh) in DMSO-d₆.

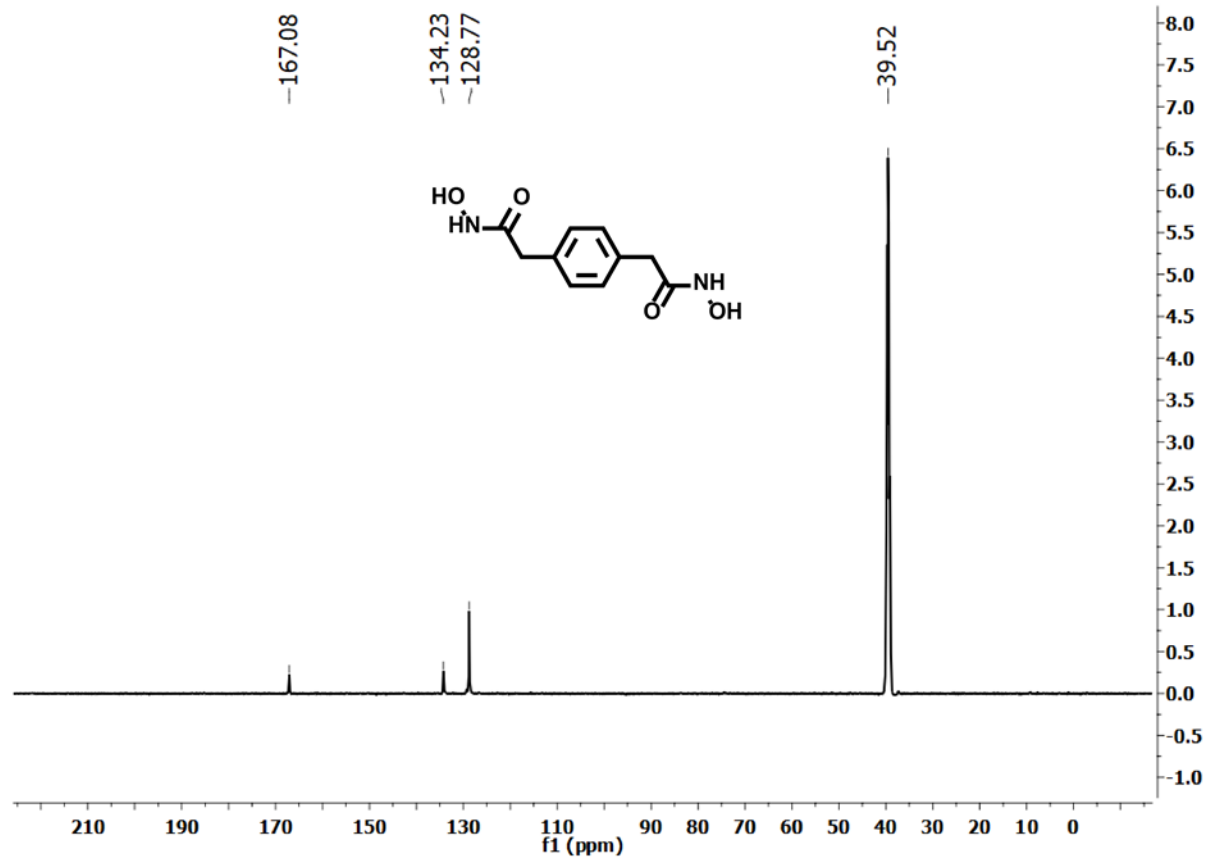


Figure A2.11. ¹³C NMR spectrum of 2,2'-(1,4-phenylene)bis(*N*-hydroxyacetamide) (*p*-H₂xdh) in DMSO-*d*₆.

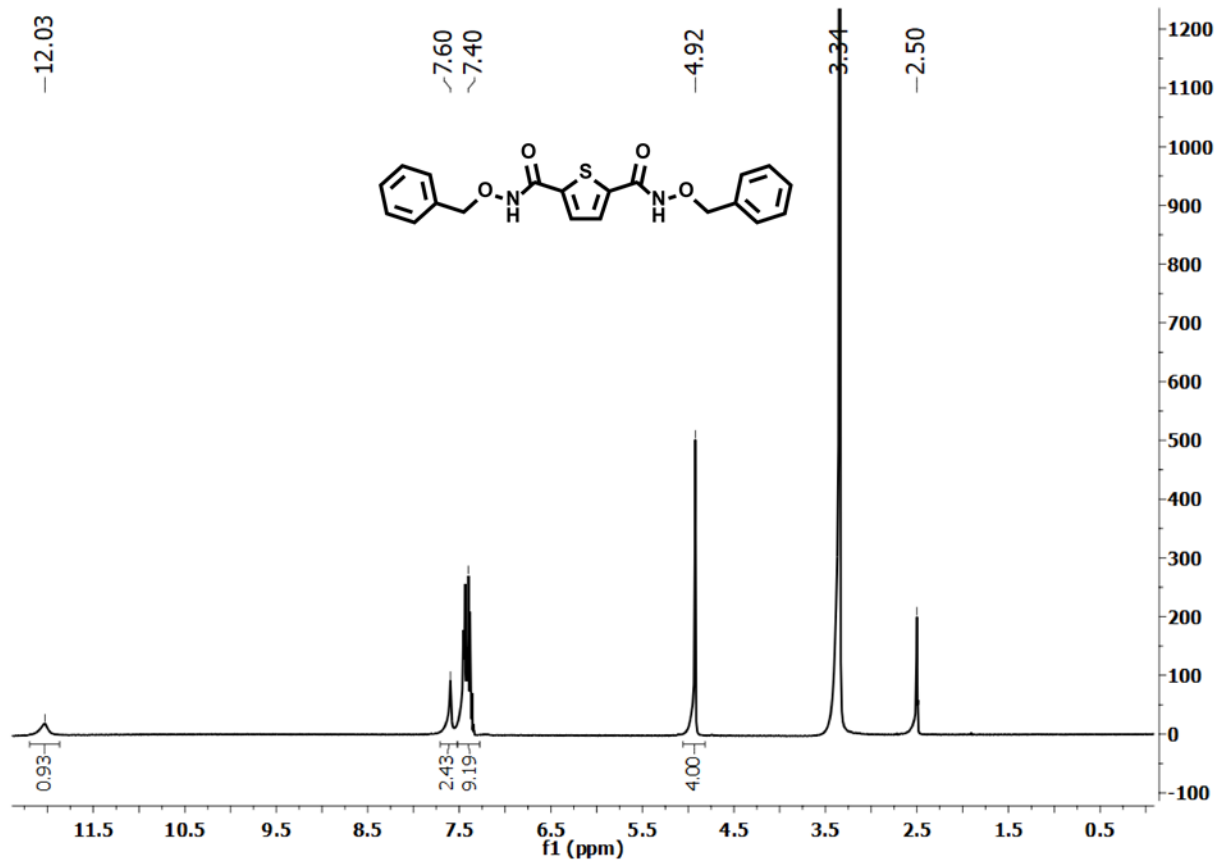


Figure A2.12. ¹H NMR spectrum of *N*²,*N*⁵-bis(benzyloxy)furan-2,5-dicarboxamide (2,5-Bn₂tdh) in DMSO-*d*₆.

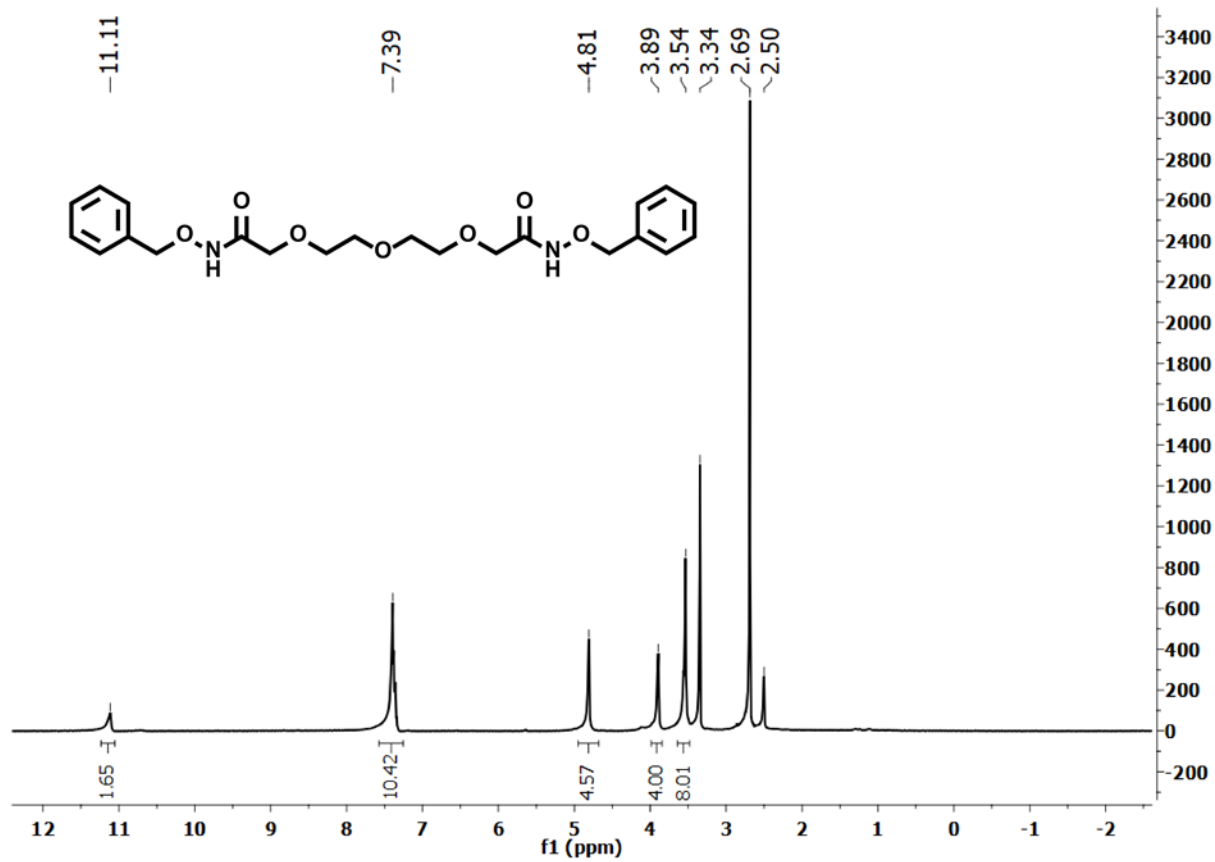


Figure A2.13. ¹H NMR spectrum of 2,2'-((oxybis(ethane-2,1-diyl))bis(oxy))bis(*N*-(benzyloxy)acetamide) (Bh₂dgdh) in DMSO-d₆.

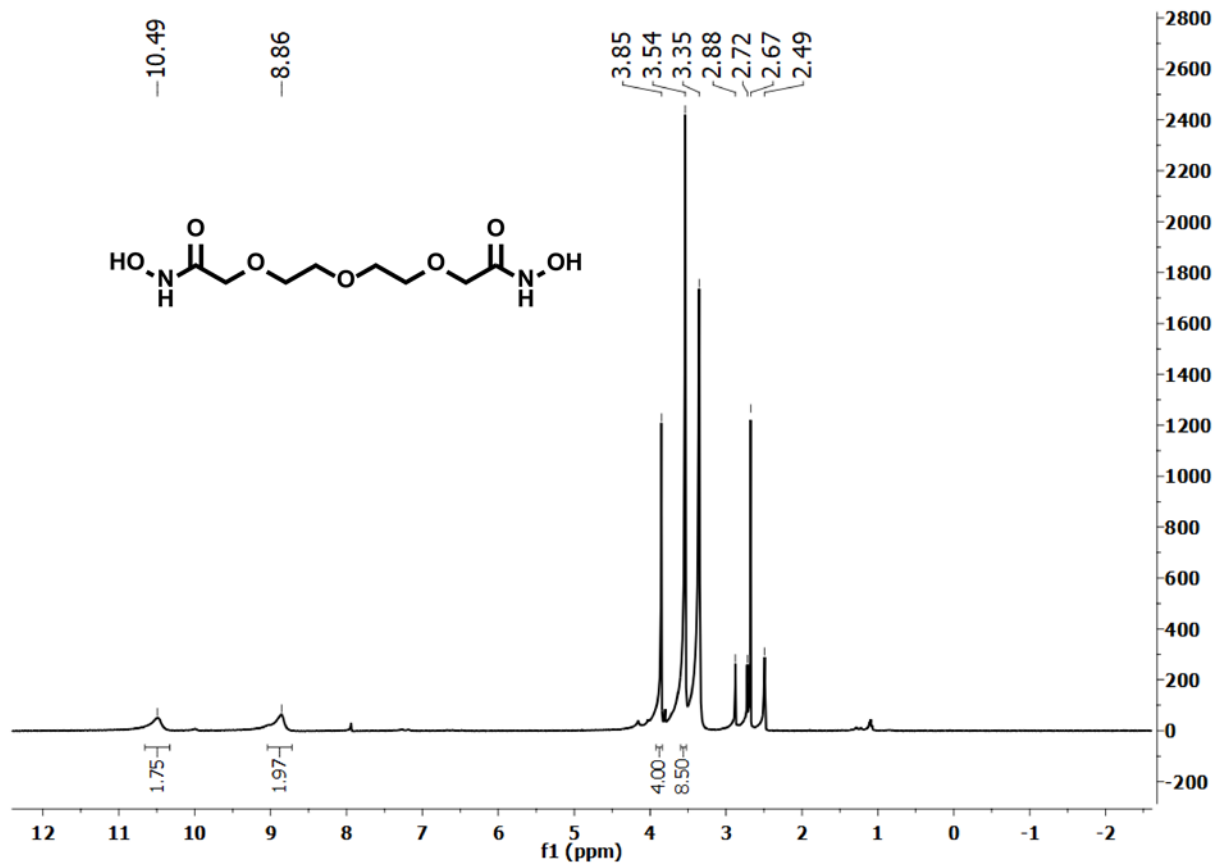


Figure A2.14. ¹H NMR spectrum of 2,2'-((oxybis(ethane-2,1-diyl))bis(oxy))bis(*N*-hydroxyacetamide) (H₂dgdh) in DMSO-d₆.

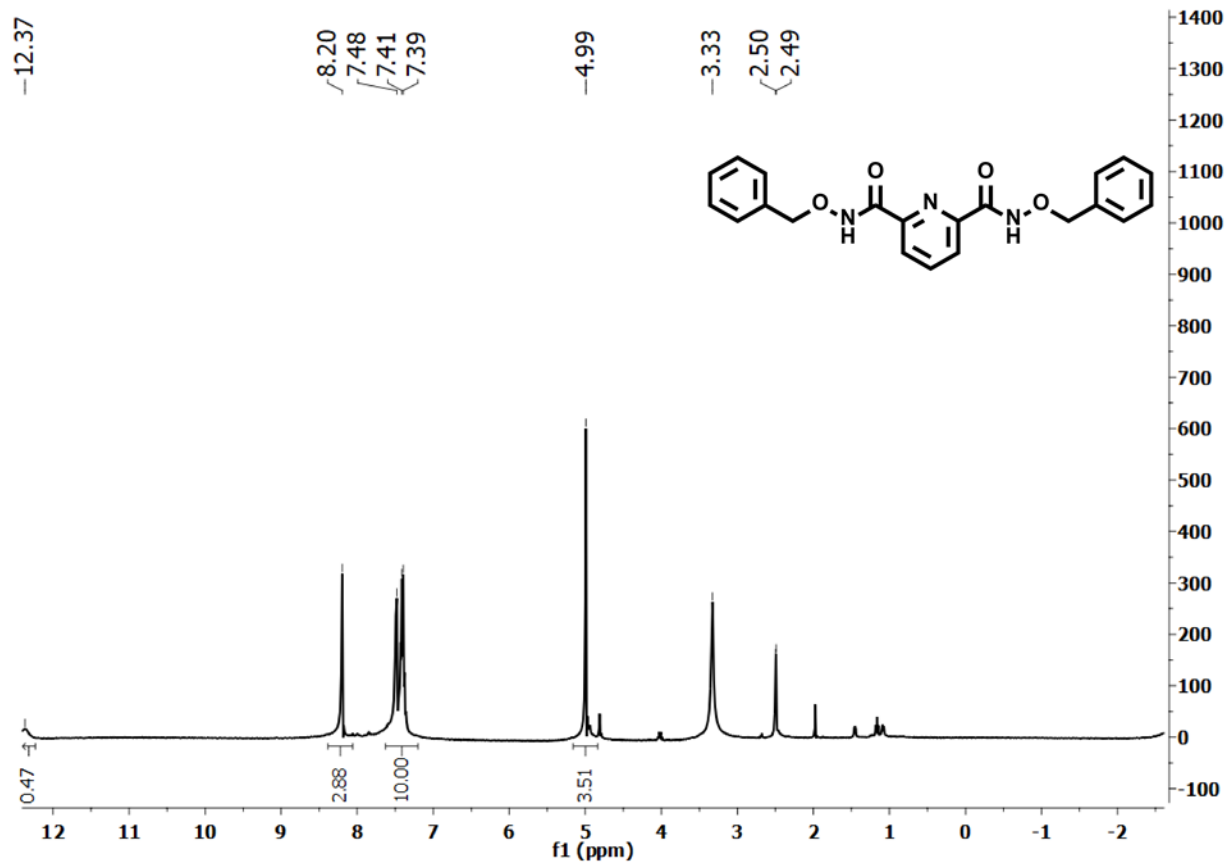


Figure A2.15. ^1H NMR spectrum of N^2,N^6 -bis(benzyloxy)pyridine-2,6-dicarboxamide (2,6-Bn₂pydh) in DMSO-d₆.

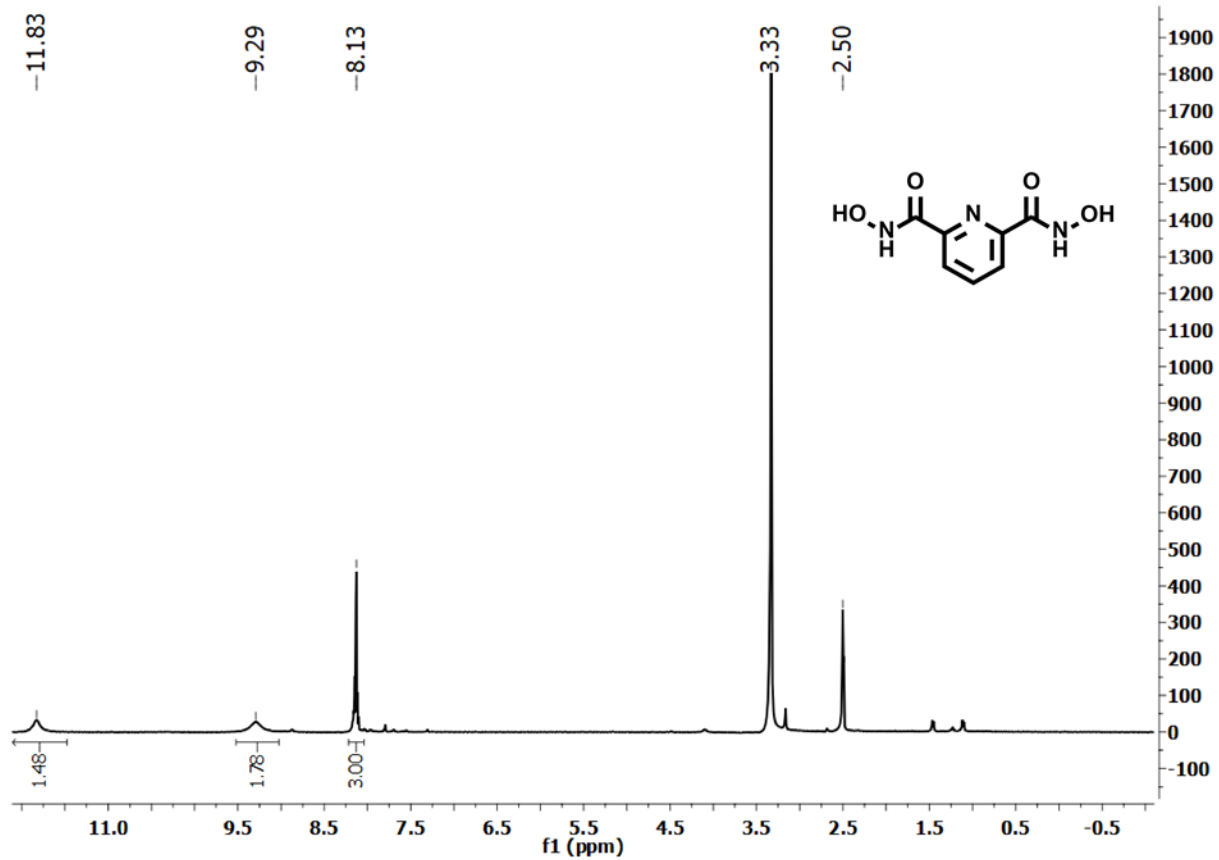


Figure A2.16. ¹H NMR spectrum of *N*²,*N*⁶-dihydroxypyridine-2,6-dicarboxamide (2,6-Bn₂pydh) in DMSO-*d*₆.

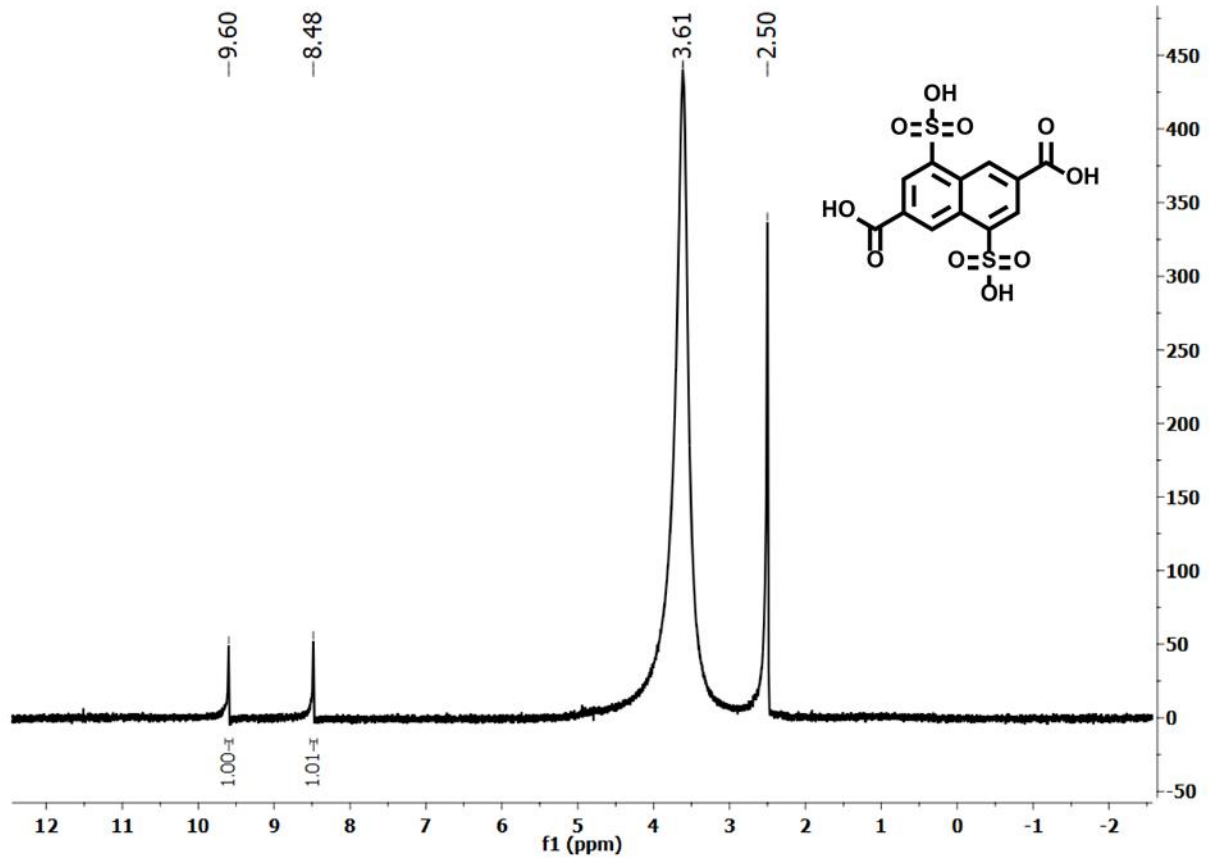


Figure A2.17. ^1H NMR spectrum of 4,8-disulfonaphthalene-2,6-dicarboxylic acid (2,6-sndc) in DMSO-d_6 .

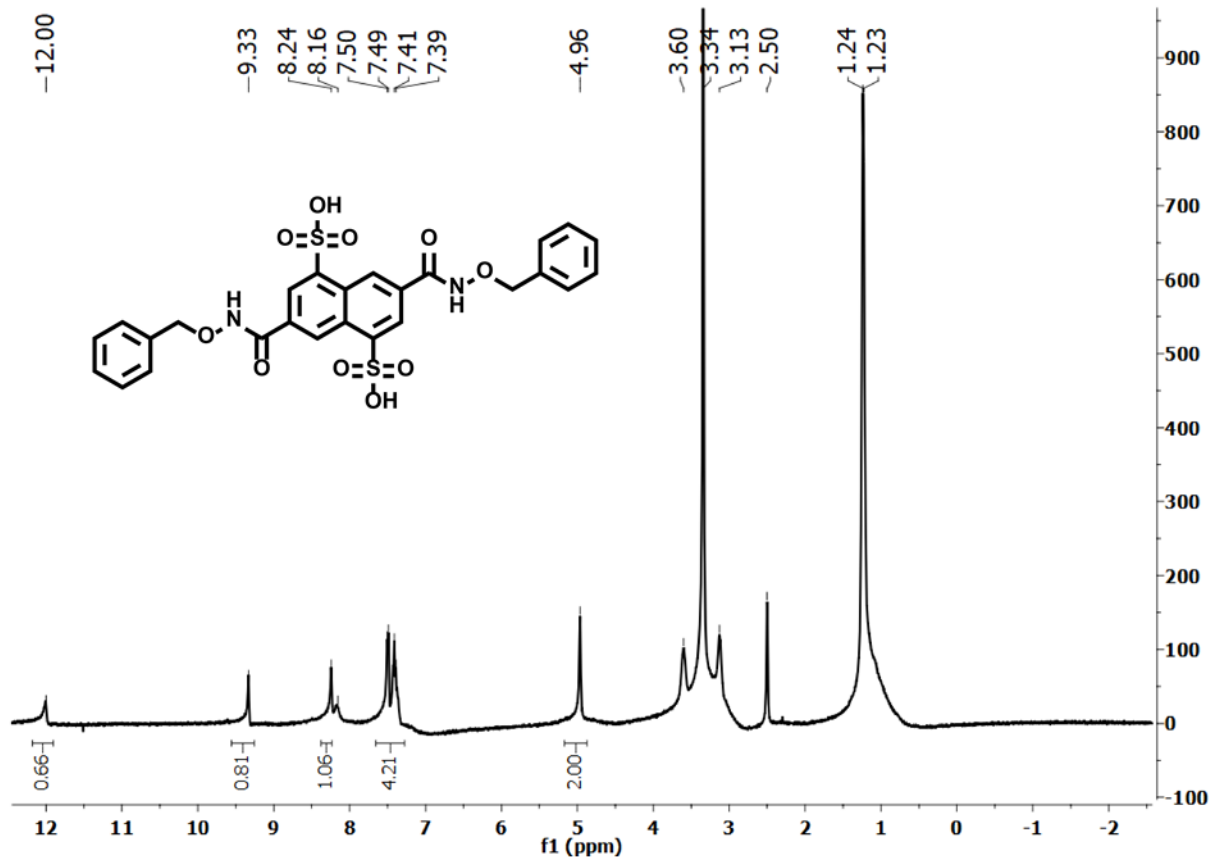


Figure A2.18. ¹H NMR spectrum of 3,7-bis((benzyloxy)carbamoyl)naphthalene-1,5-disulfonic acid (2,6-Bn₂sndh) in DMSO-d₆.

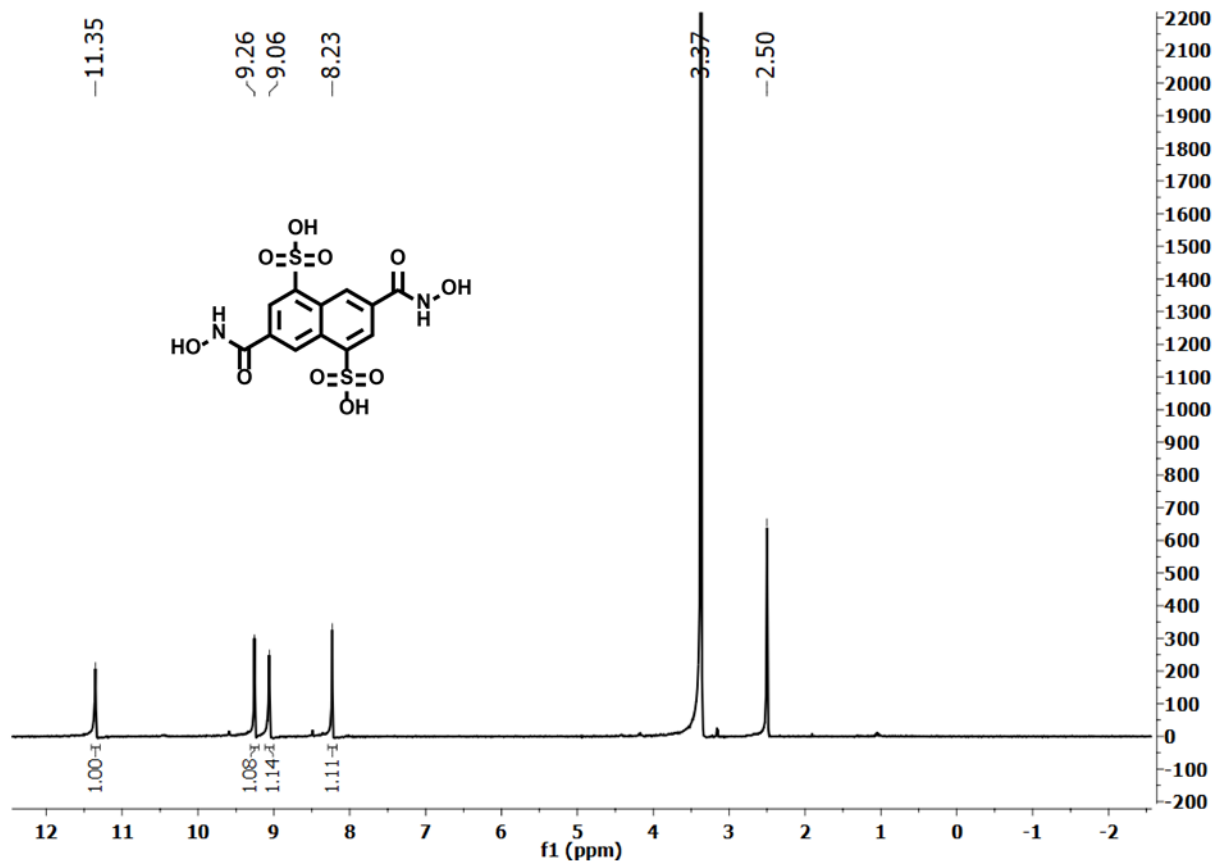


Figure A2.19. ¹H NMR spectrum of 3,7-bis(hydroxycarbamoyl)naphthalene-1,5-disulfonic acid (2,6-H₂sndh) in DMSO-d₆.

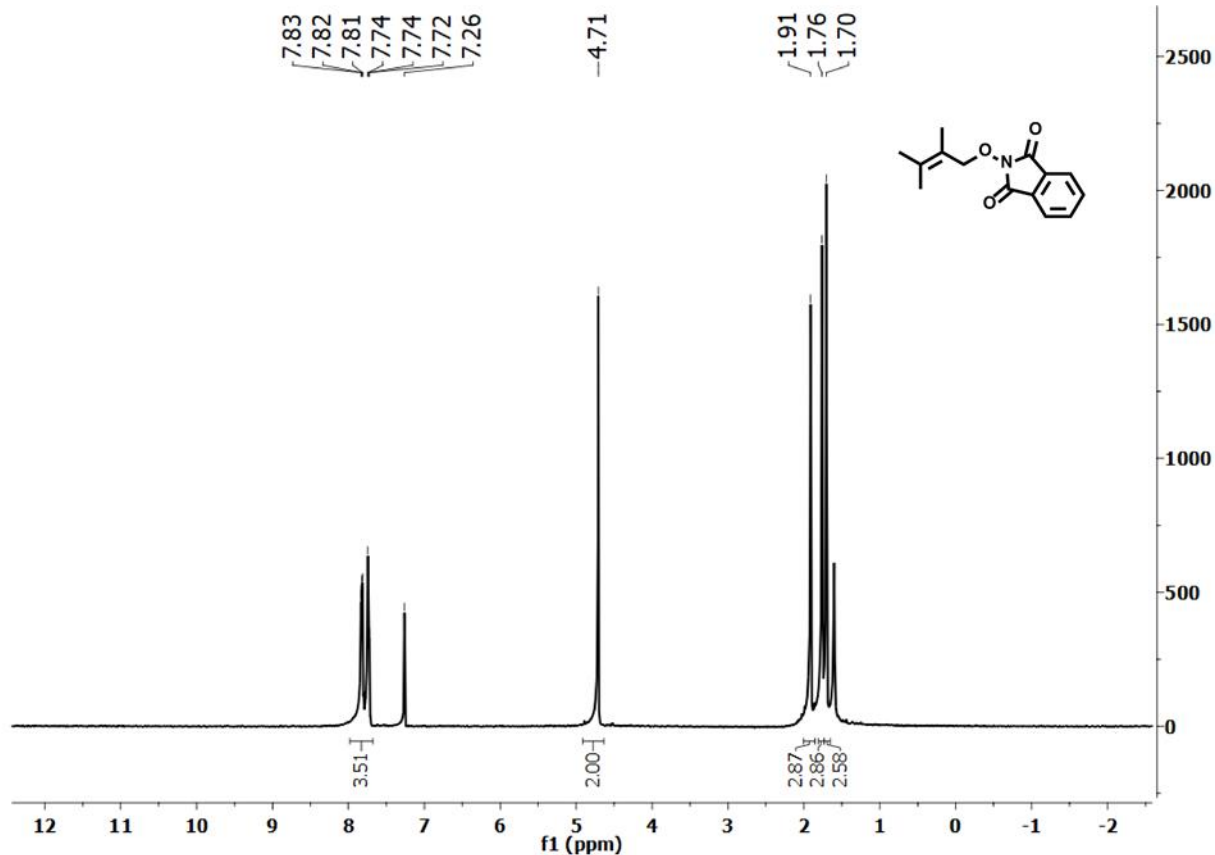


Figure A2.20. ¹H NMR spectrum of 2-((2,3-dimethylbut-2-en-1-yl)oxy)isoindoline-1,3-dione in DMSO-d₆.

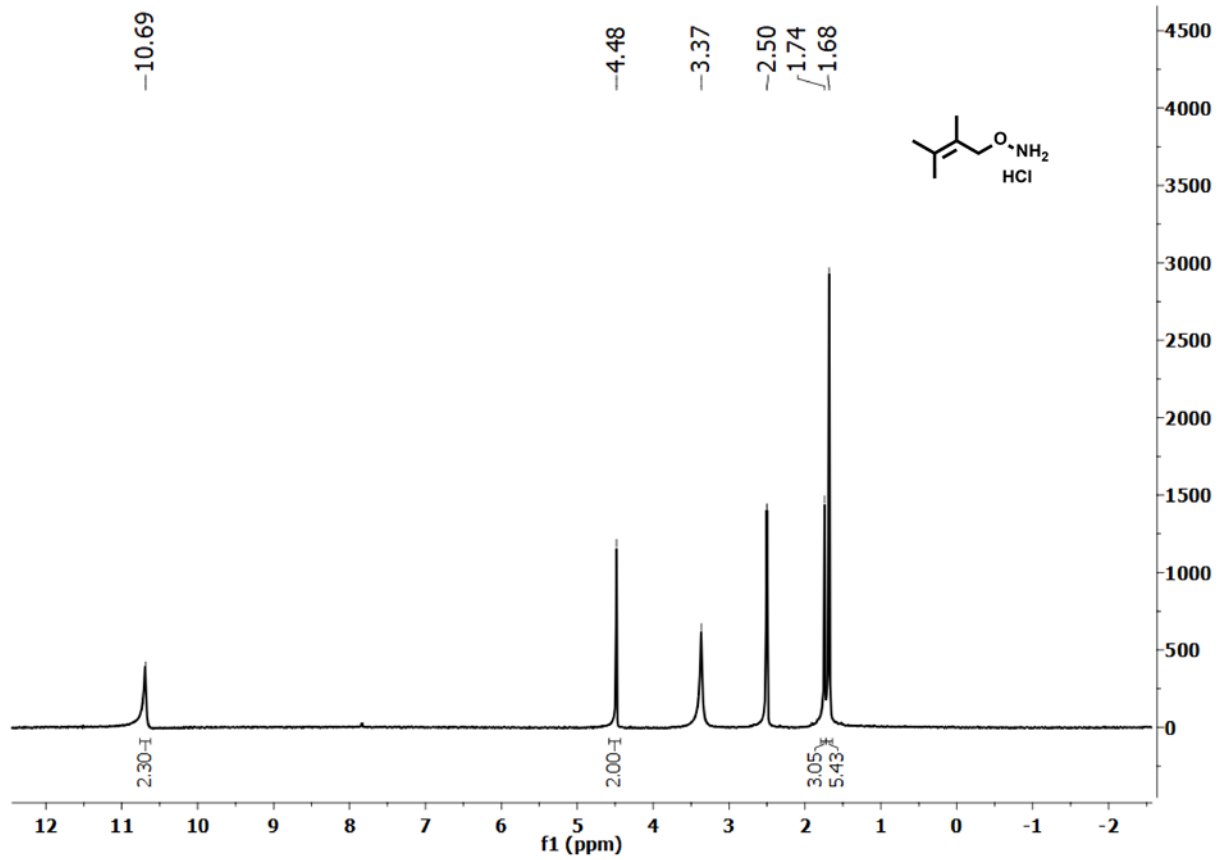


Figure A2.21. ^1H NMR spectrum of O-(2,3-dimethylbut-2-en-1-yl)hydroxylamine hydrochloride in DMSO-d_6 .

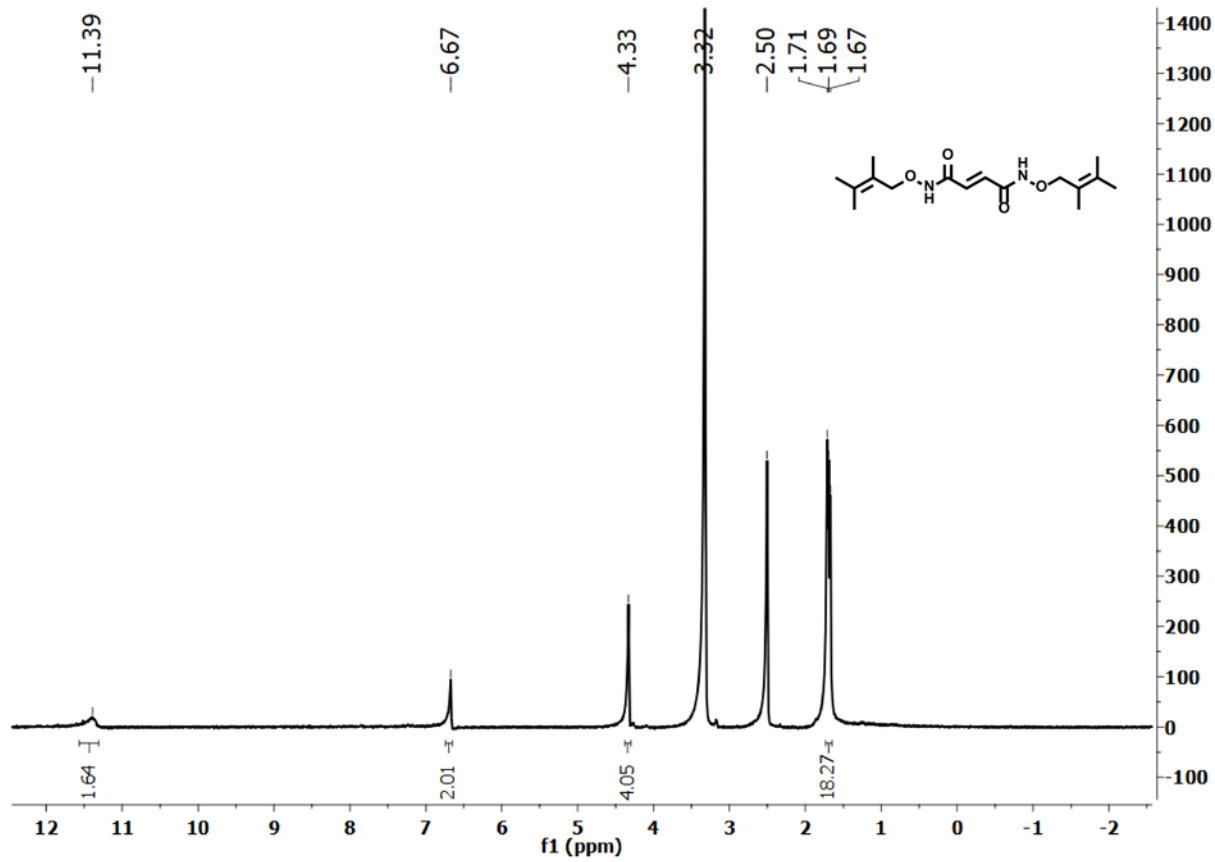


Figure A2.22. ¹H NMR spectrum of *N*¹,*N*⁴-bis((2,3-dimethylbut-2-en-1-yl)oxy)fumaramide in DMSO-*d*₆.

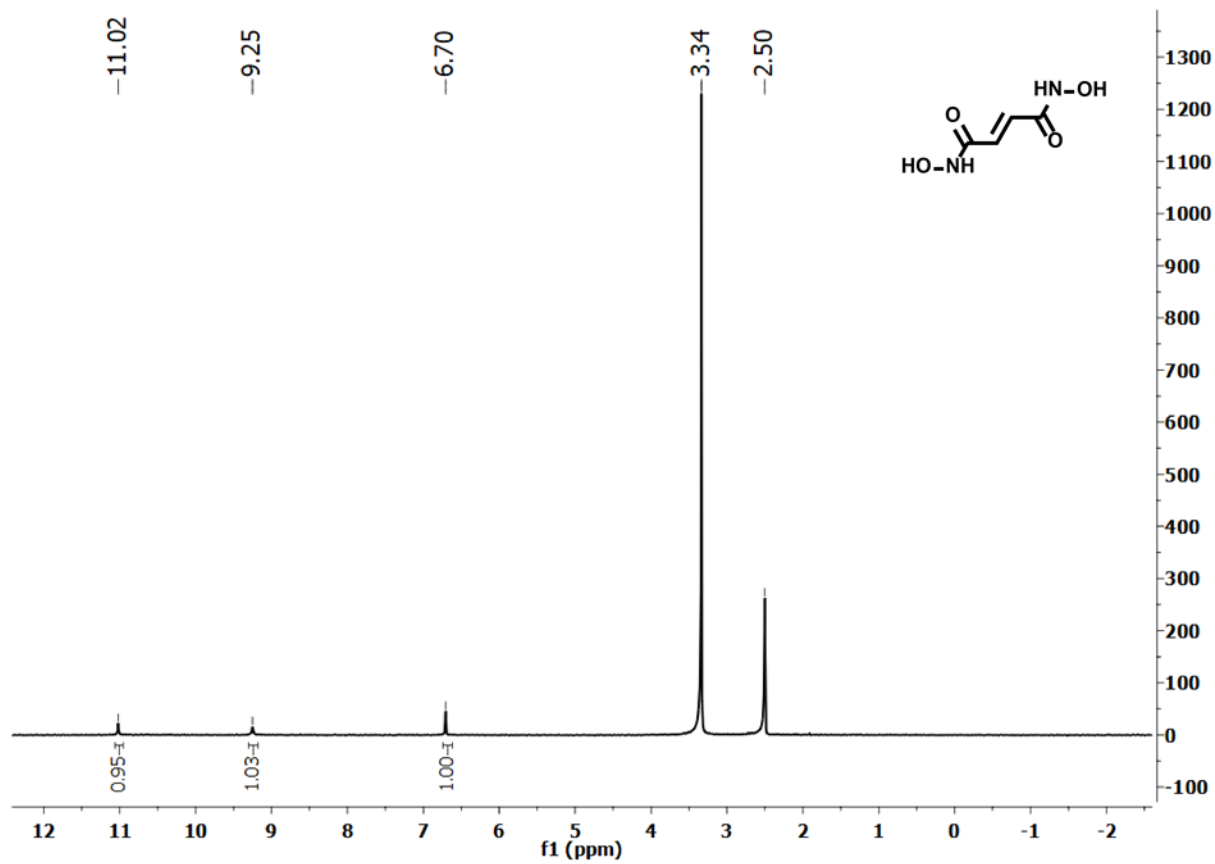


Figure A2.23. ¹H NMR spectrum of *N',N''*-bis((2,3-dimethylbut-2-en-1-yl)oxy)fumaramide (*E*-H₂edh) in DMSO-d₆.

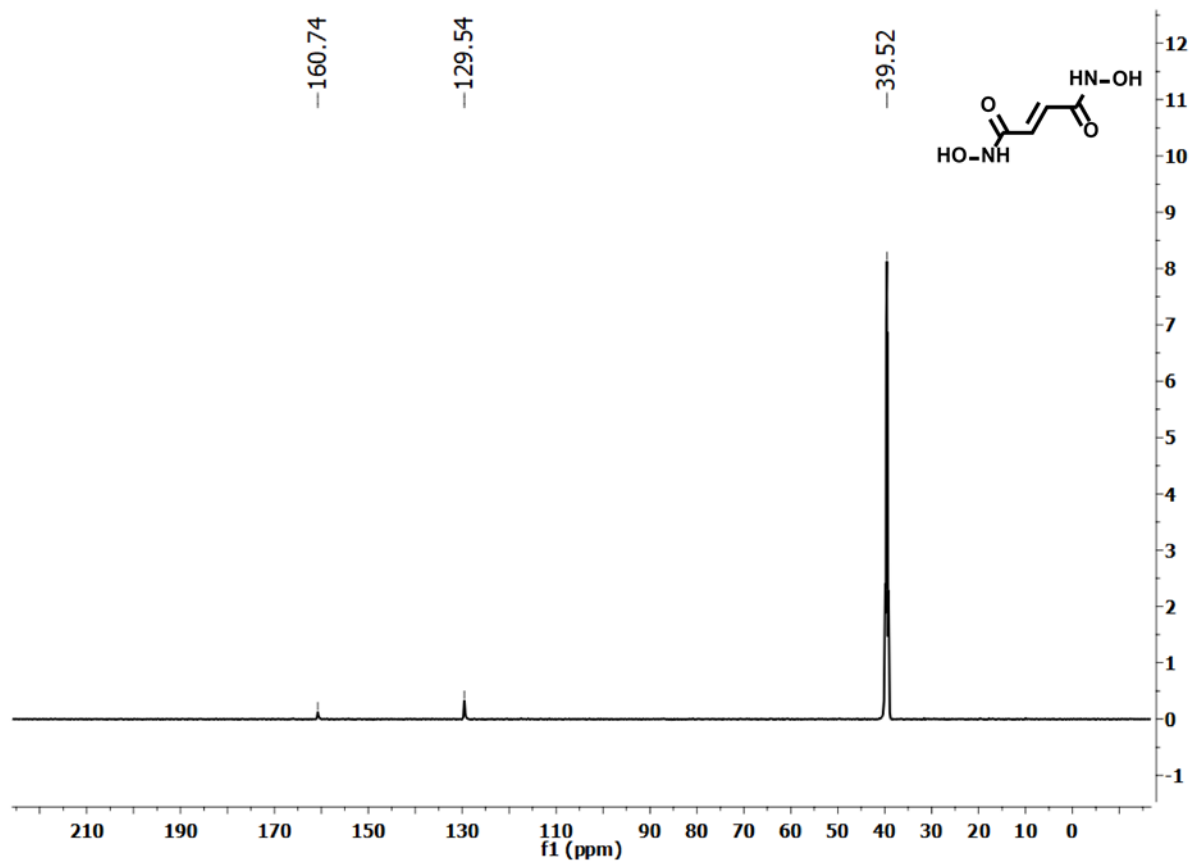


Figure A2.24. ^{13}C NMR spectrum of *N',N''*-bis((2,3-dimethylbut-2-en-1-yl)oxy)fumaramide (*E*-H₂edh) in DMSO- d_6 .

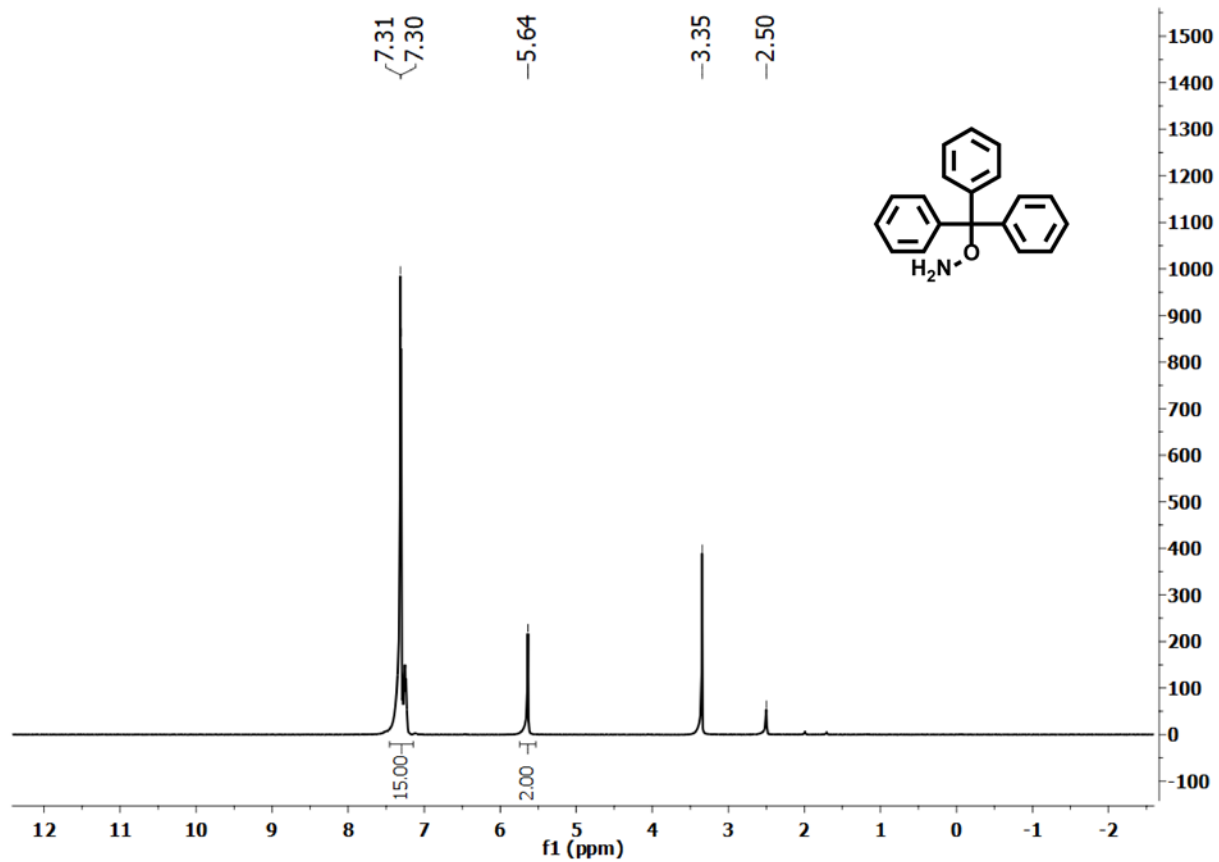


Figure A2.25. ^1H NMR spectrum of O-tritylhydroxylamine in DMSO-d_6 .

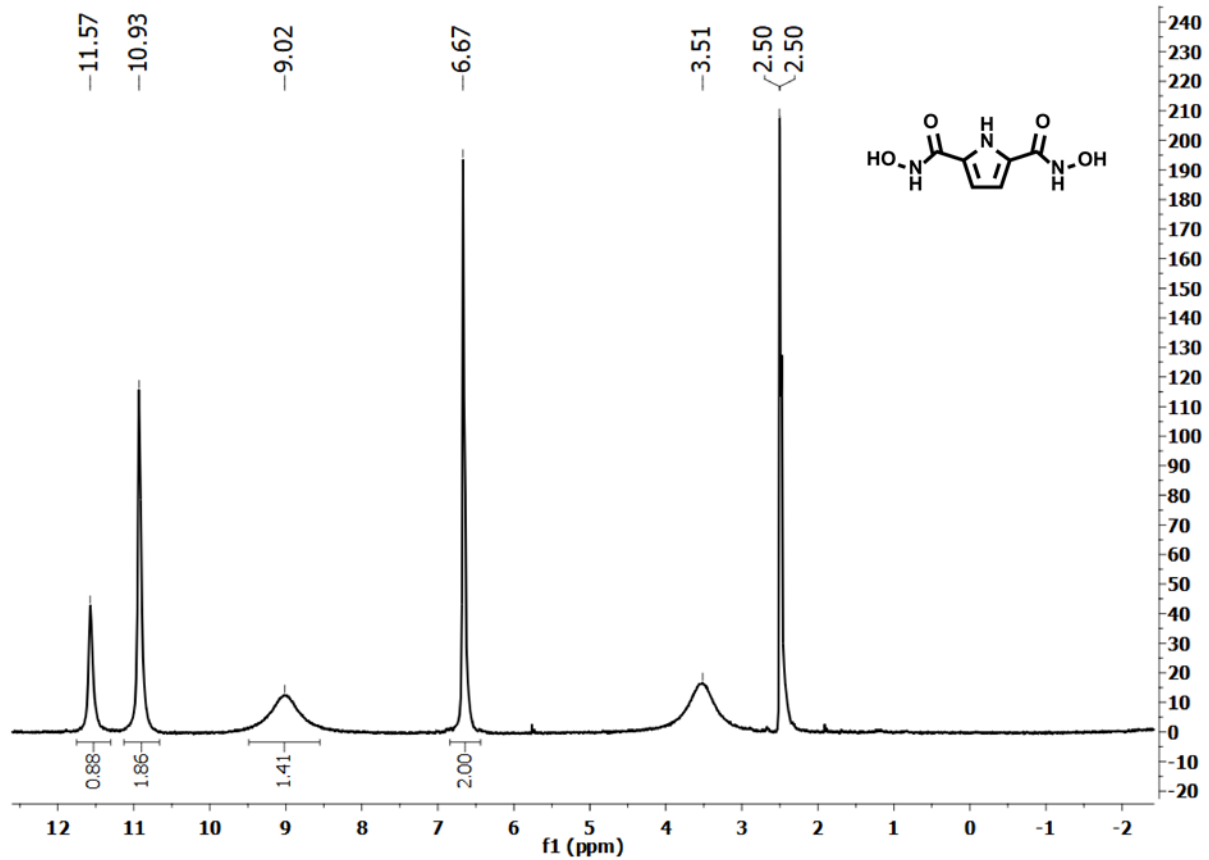


Figure A2.26. ¹H NMR spectrum of *N*²,*N*⁶-dihydroxy-1*H*-pyrrole-2,5-dicarboxamide (2,5-H₂pdh) in DMSO-*d*₆.

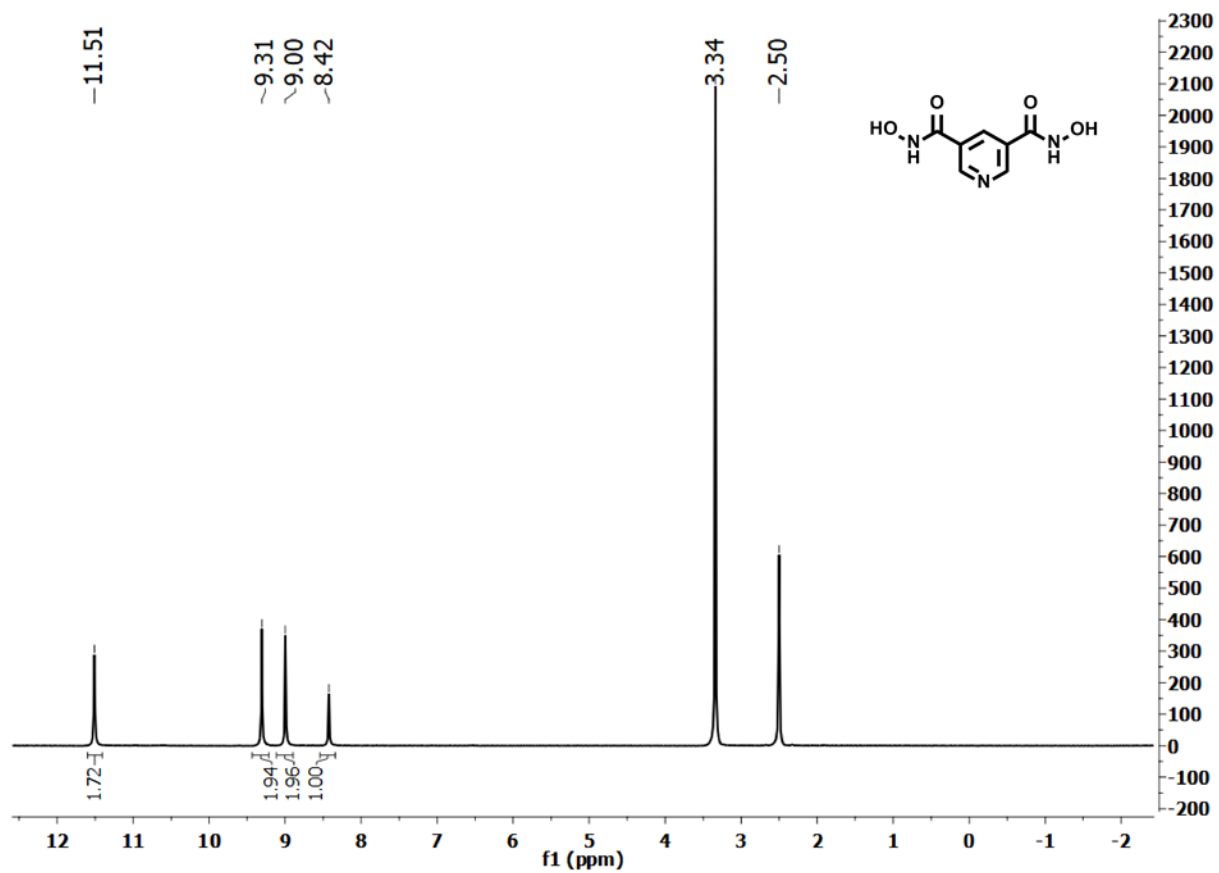


Figure A2.27. ¹H NMR spectrum of *N*³,*N*⁶-dihydroxypyridine-3,5-dicarboxamide (3,5-H₂pzdH) in DMSO-d₆.

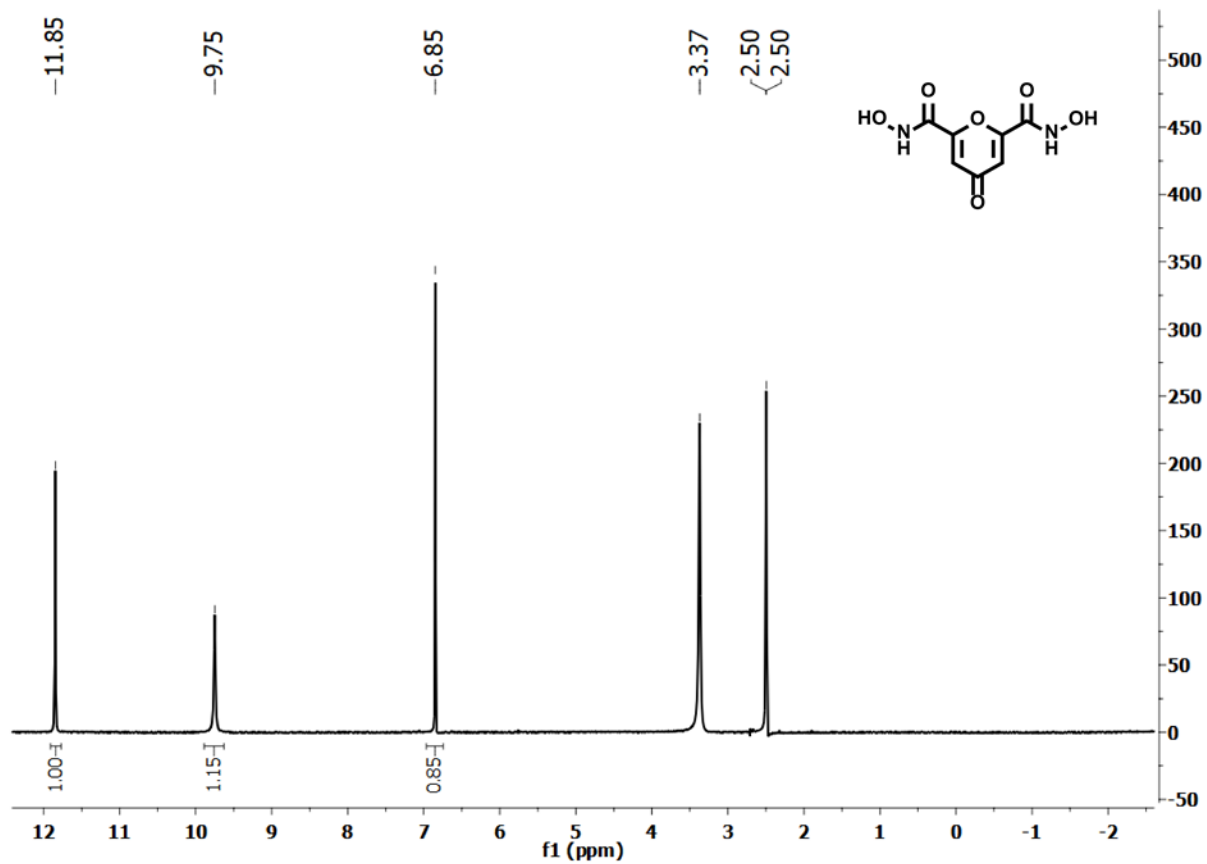


Figure A2.28. ¹H NMR spectrum of *N*²,*N*⁶-dihydroxy-4-oxo-4*H*-pyran-2,6-dicarboxamide (2,6-H₂podh) in DMSO-*d*₆.

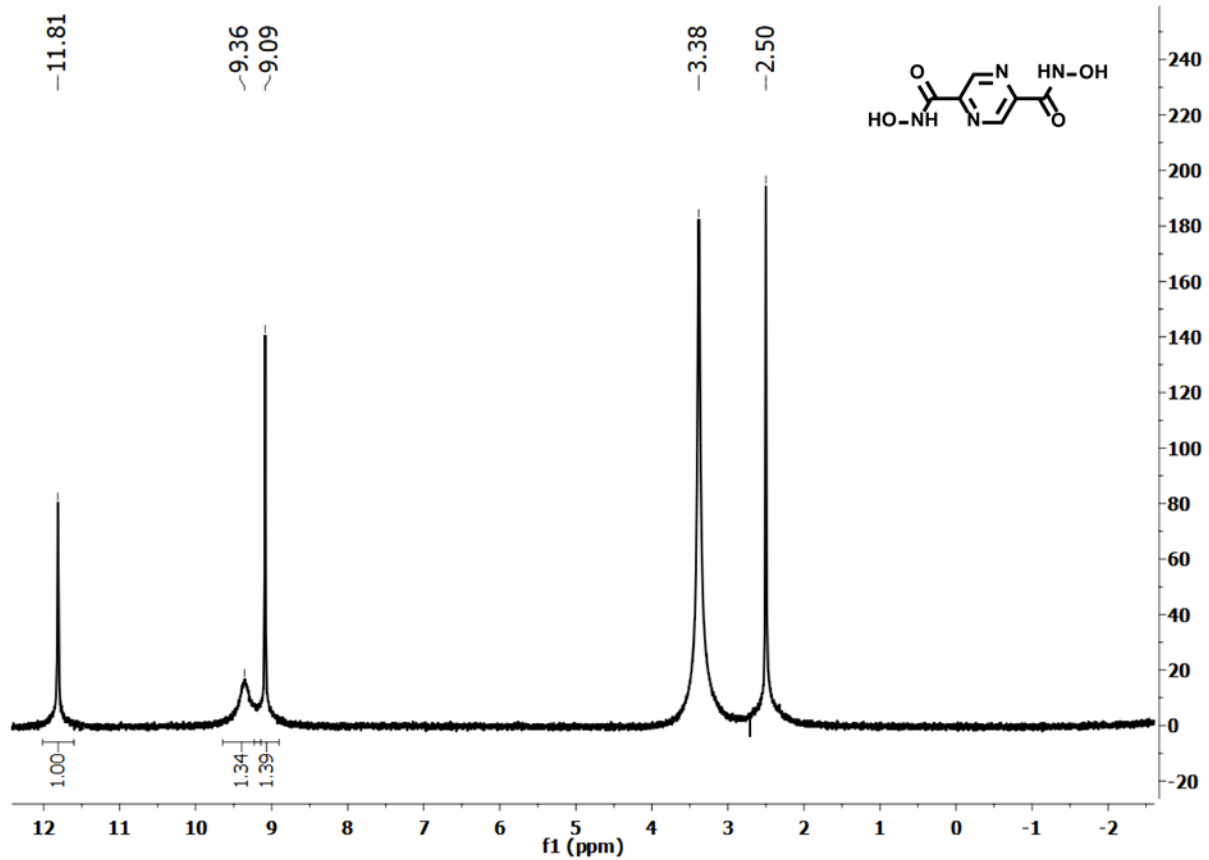


Figure A2.29. ¹H NMR spectrum of *N*²,*N*⁶-dihydroxypyrazine-2,5-dicarboxamide in DMSO-*d*₆.

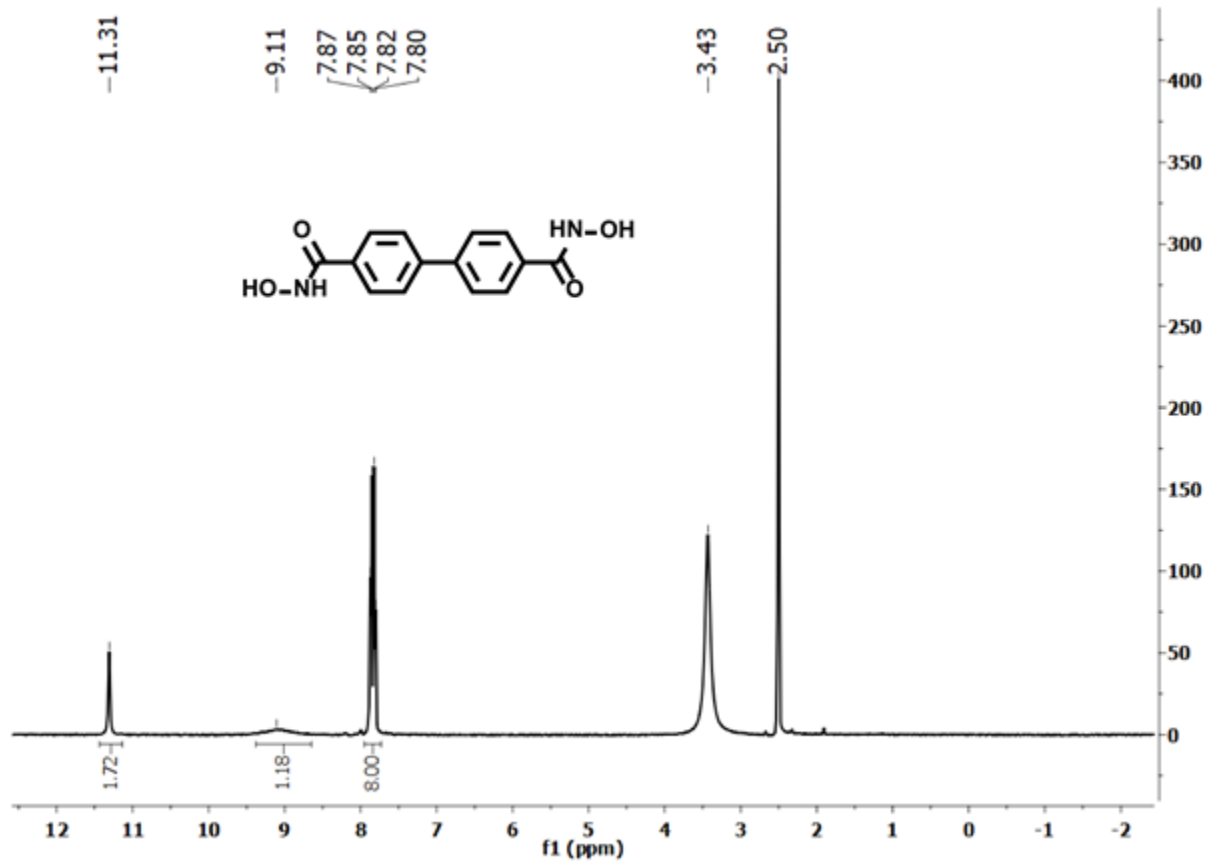


Figure A2.30. ¹H NMR spectrum of [1,1'-biphenyl]-4,4'-dicarboxylic acid in DMSO-d₆.

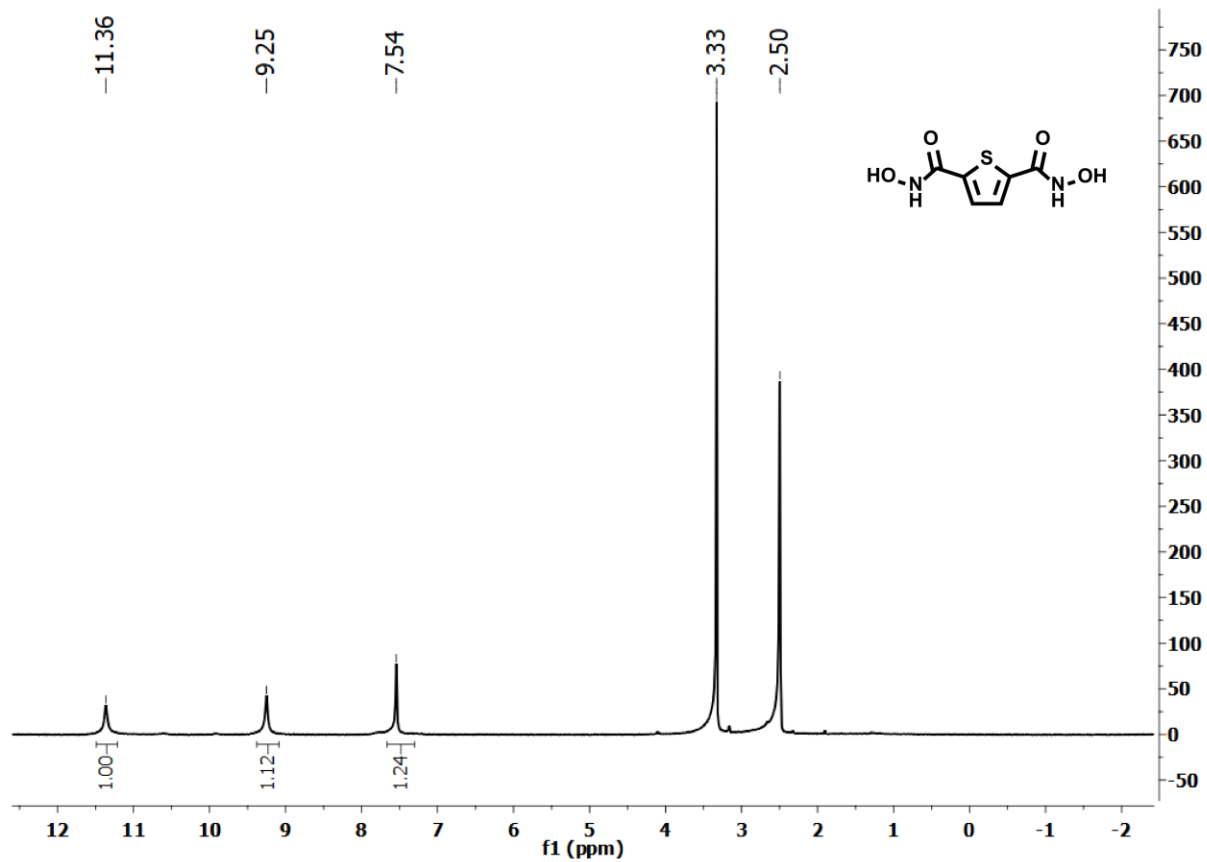


Figure A2.31. ^1H NMR spectrum of N^2,N^6 -dihydroxythiophene-2,5-dicarboxamide in DMSO-d_6 .

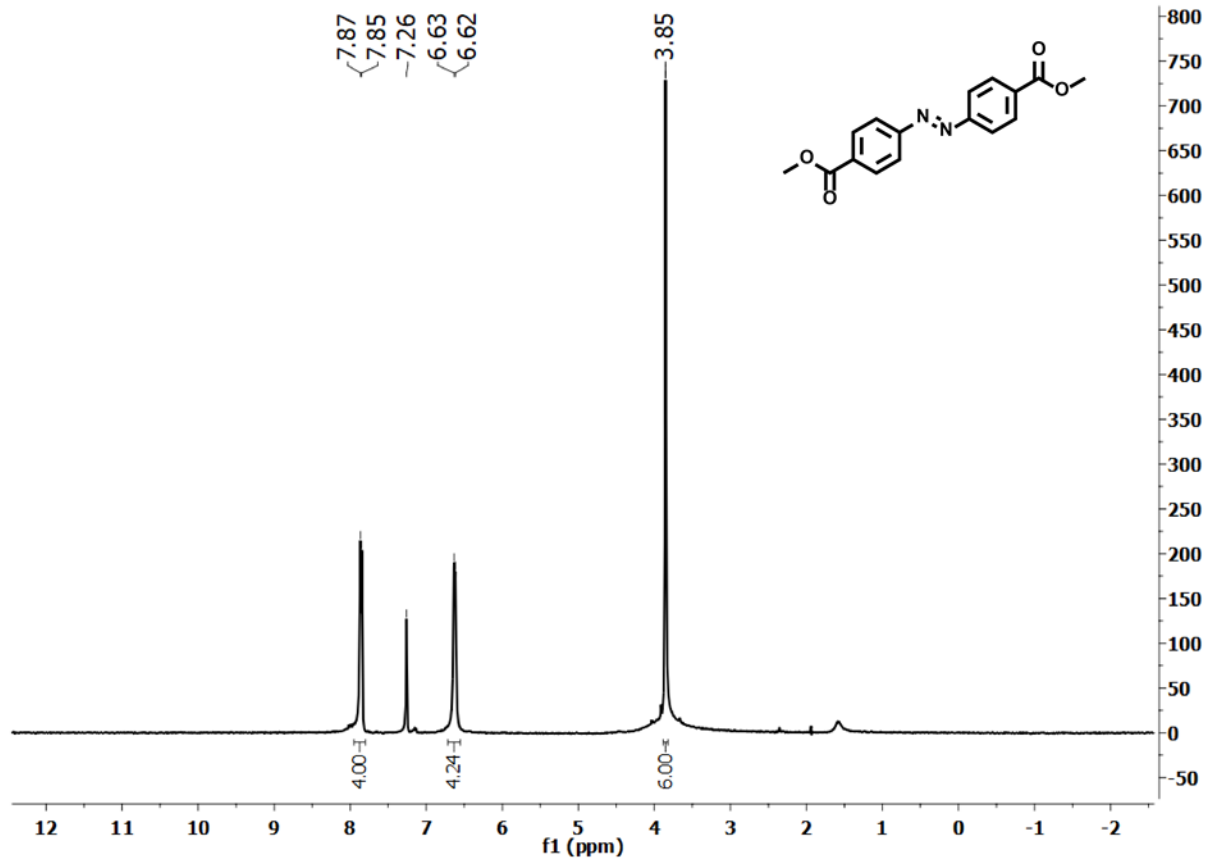


Figure A2.32. ¹H NMR spectrum of dimethyl 4,4'-(diazene-1,2-diyl)(*E*)-dibenzoate in CDCl₃.

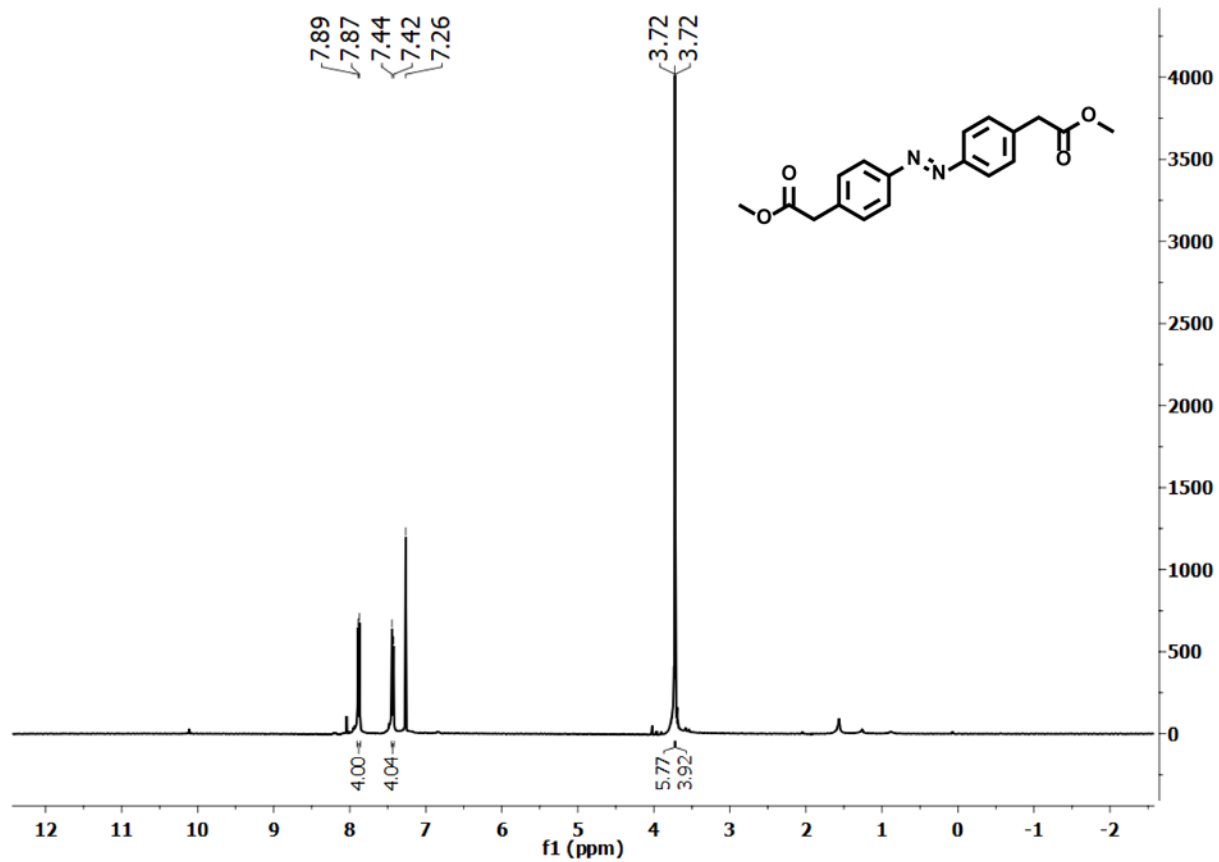


Figure A2.33. ¹H NMR spectrum of dimethyl 2,2'-(diazene-1,2-diylbis(4,1-phenylene))(*E*)-diacetate in CDCl₃.

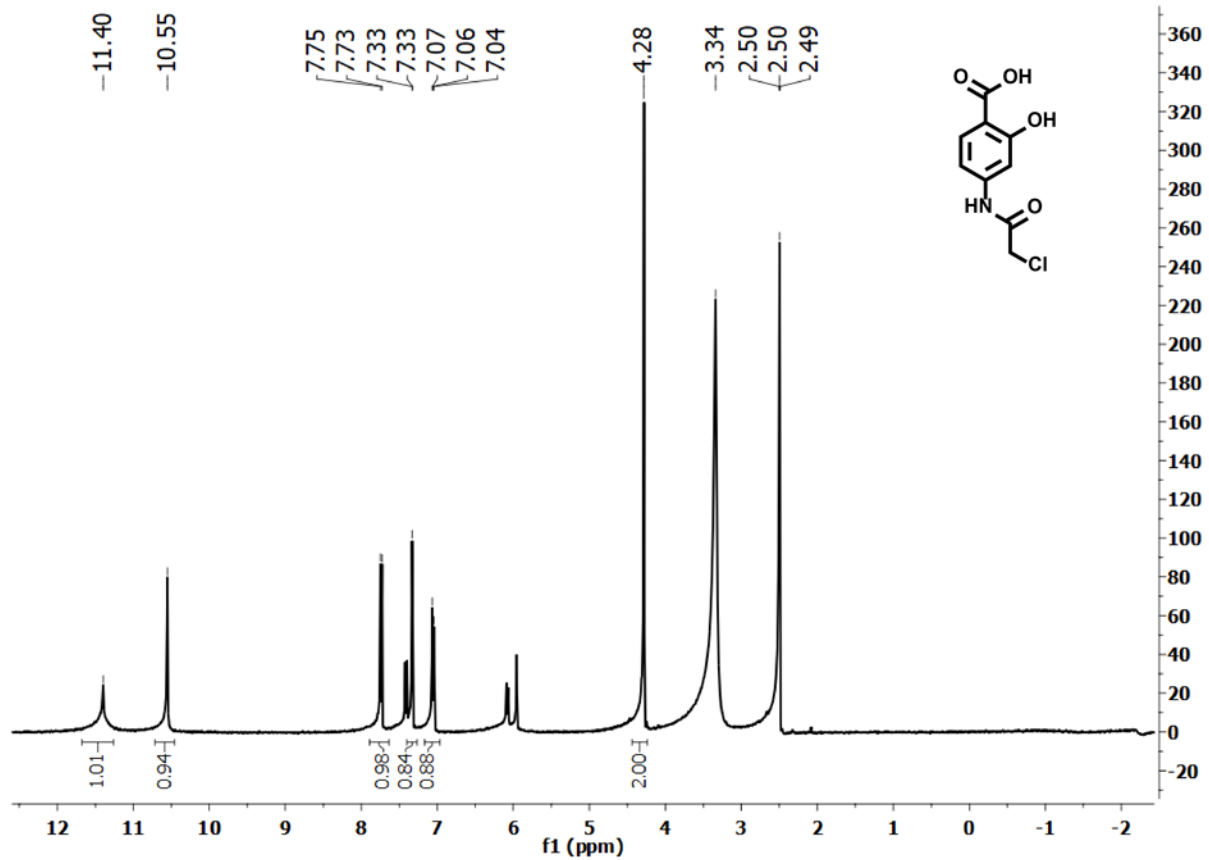


Figure A2.34. ^1H NMR spectrum of 4-(2-chloroacetamido)-2-hydroxybenzoic acid in DMSO-d_6 . Sample contains some starting material.

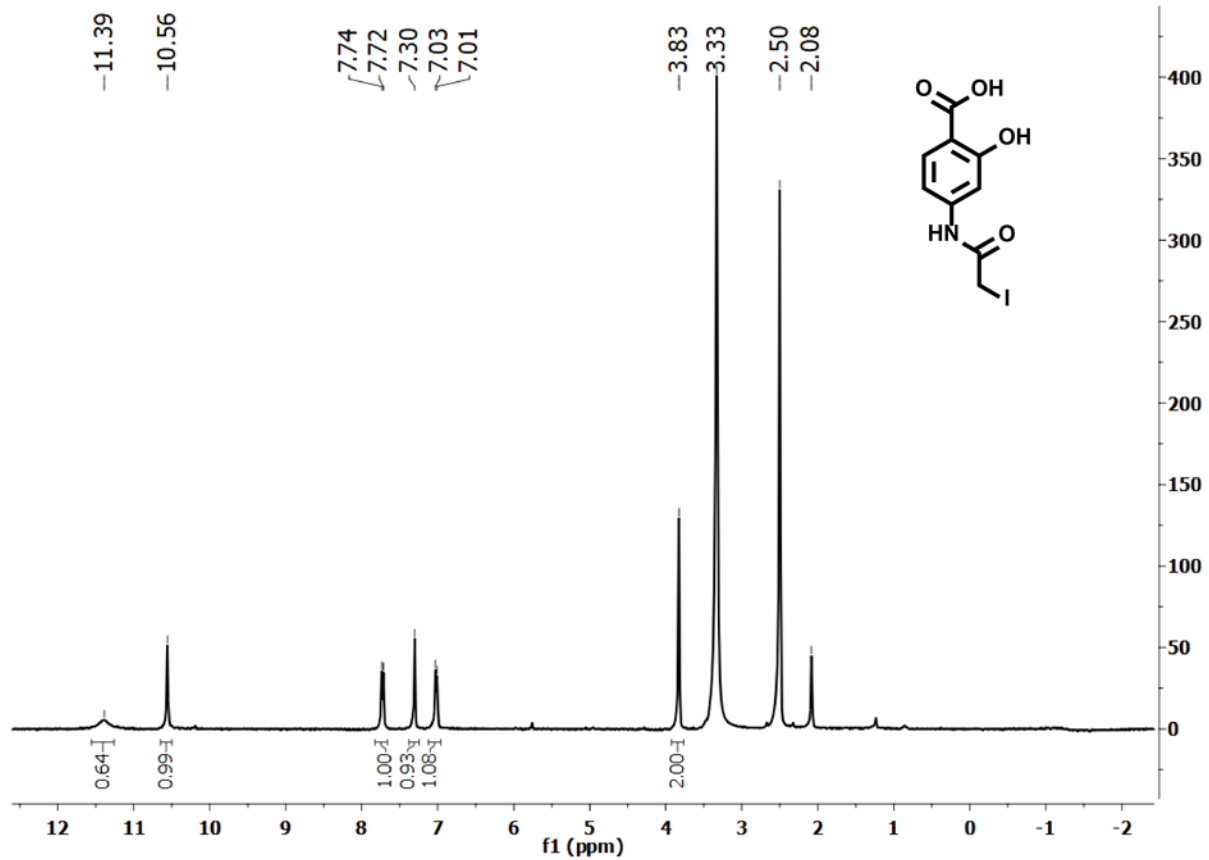


Figure A2.35. ¹H NMR spectrum of 4-(2-iodoacetamido)-2-hydroxybenzoic acid in DMSO-d₆.

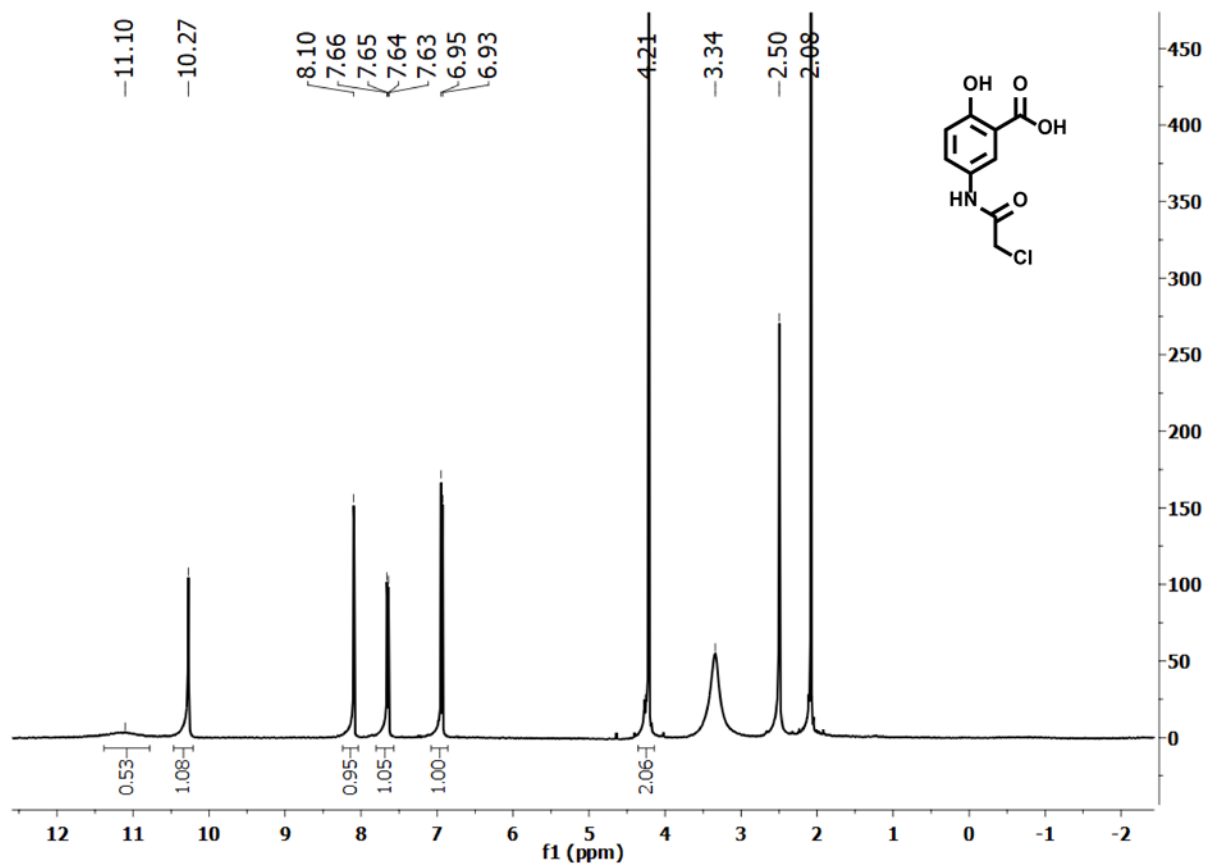


Figure A2.36. ¹H NMR spectrum of 5-(2-chloroacetamido)-2-hydroxybenzoic acid in DMSO-d₆.

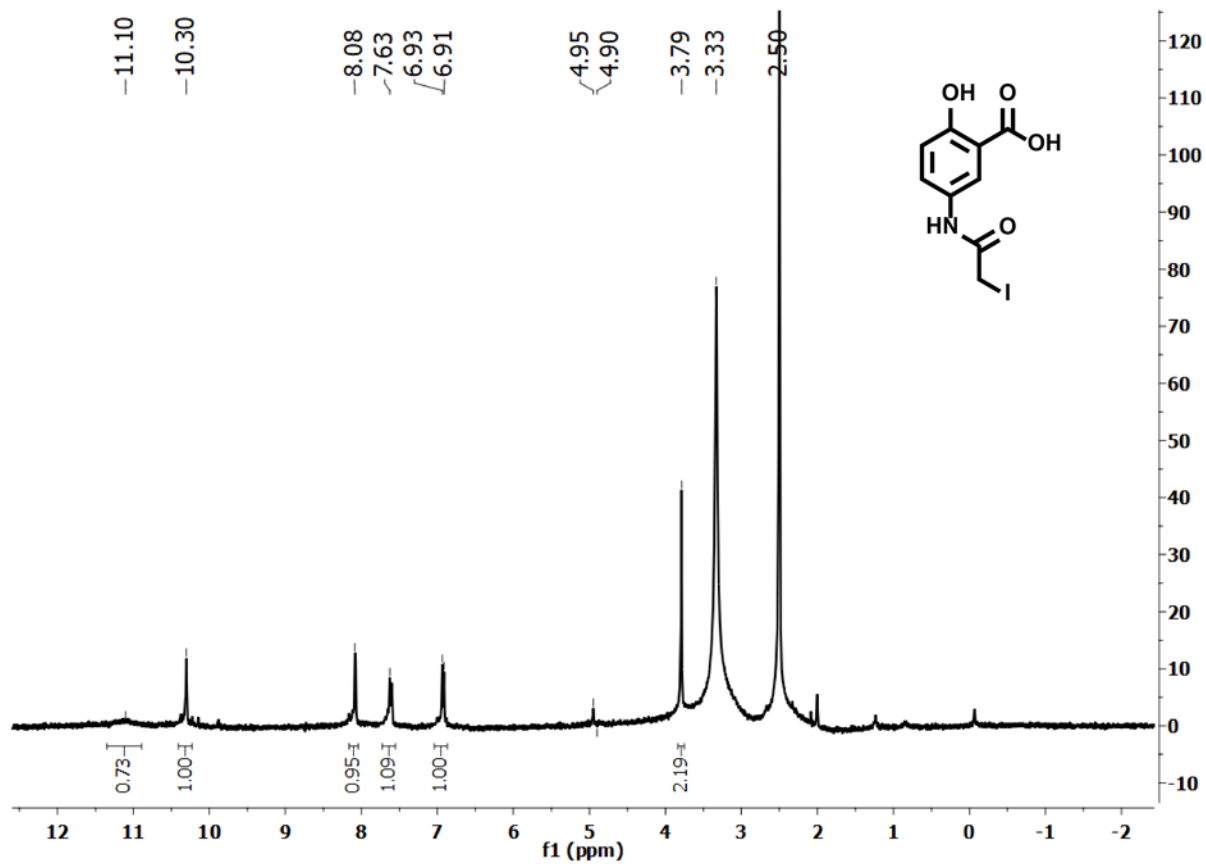


Figure A2.37. ¹H NMR spectrum of 5-(2-iodoacetamido)-2-hydroxybenzoic acid in DMSO-d₆.

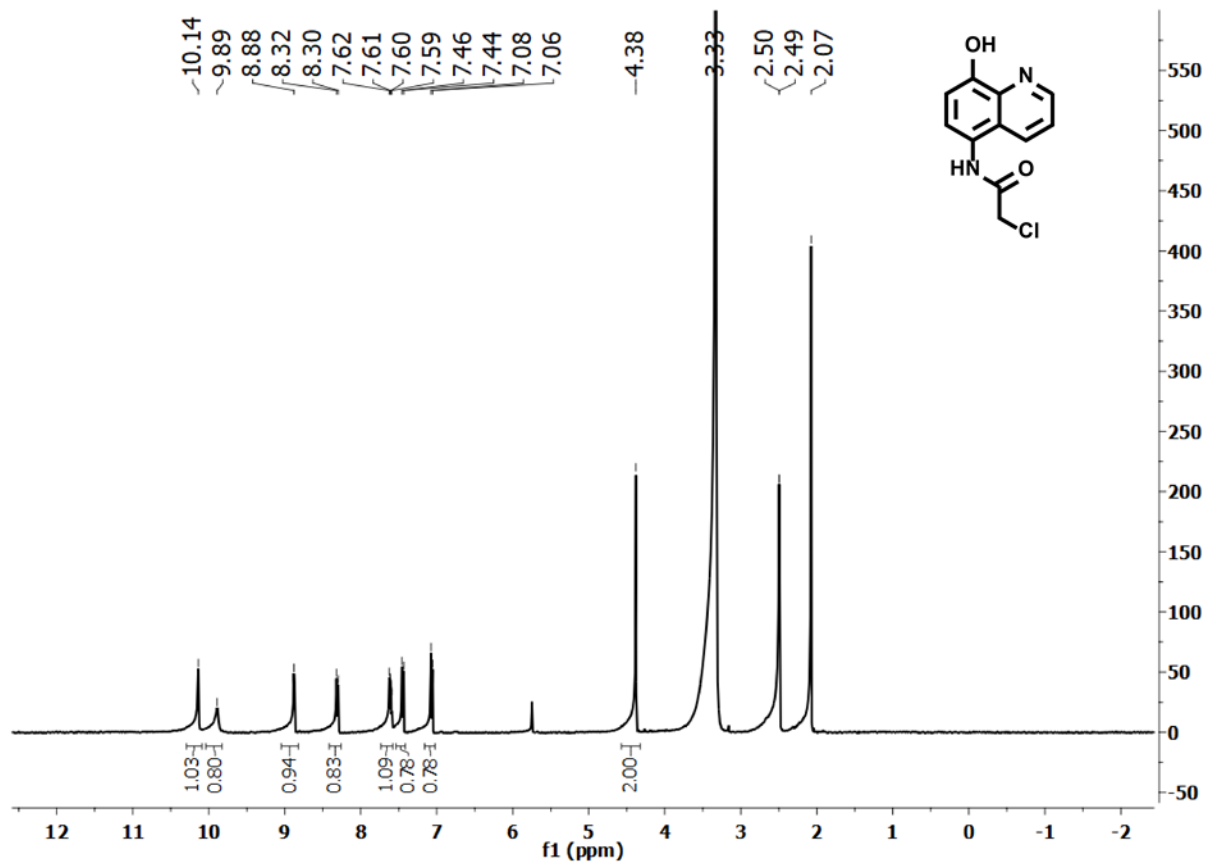


Figure A2.38. ^1H NMR spectrum of 2-chloro-*N*-(8-hydroxyquinolin-5-yl)acetamide in DMSO-d_6 .

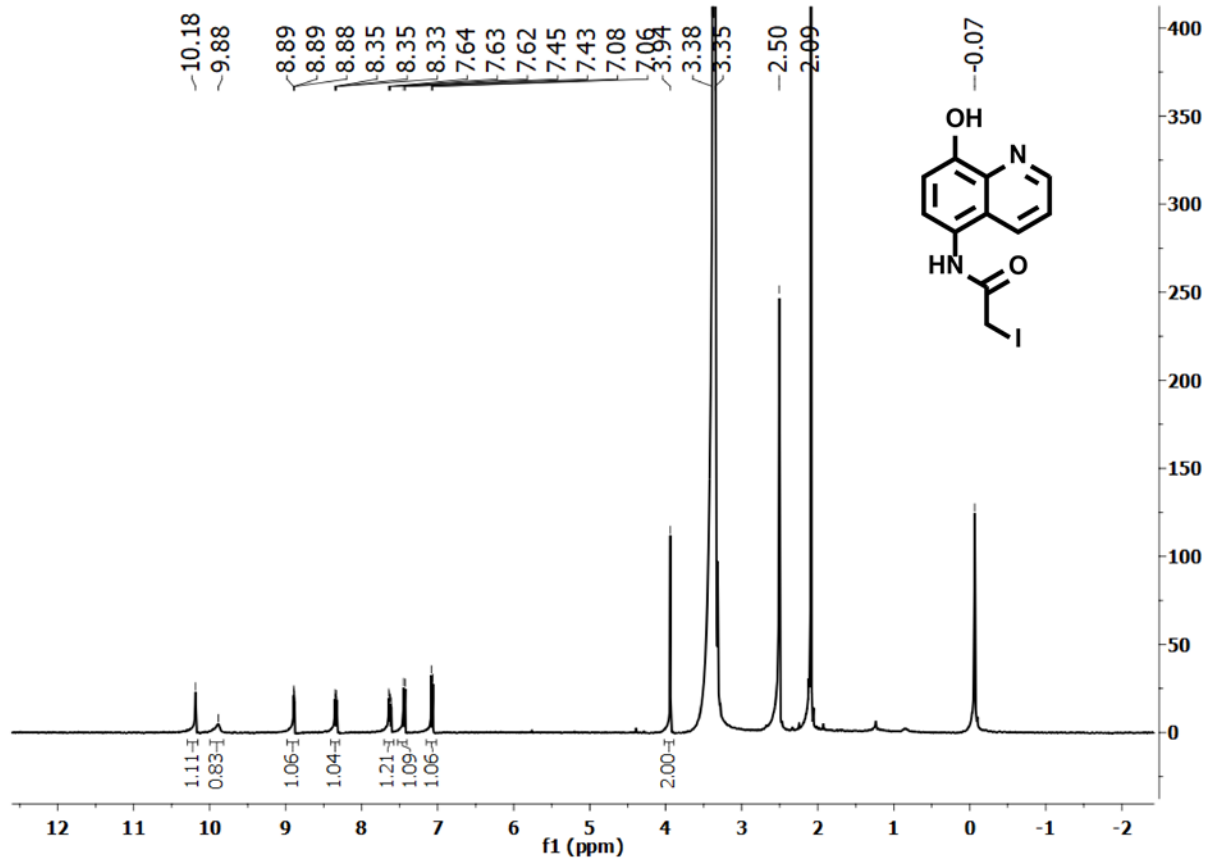


Figure A2.39. ¹H NMR spectrum of 2-iodo-N-(8-hydroxyquinolin-5-yl)acetamide in DMSO-d₆.

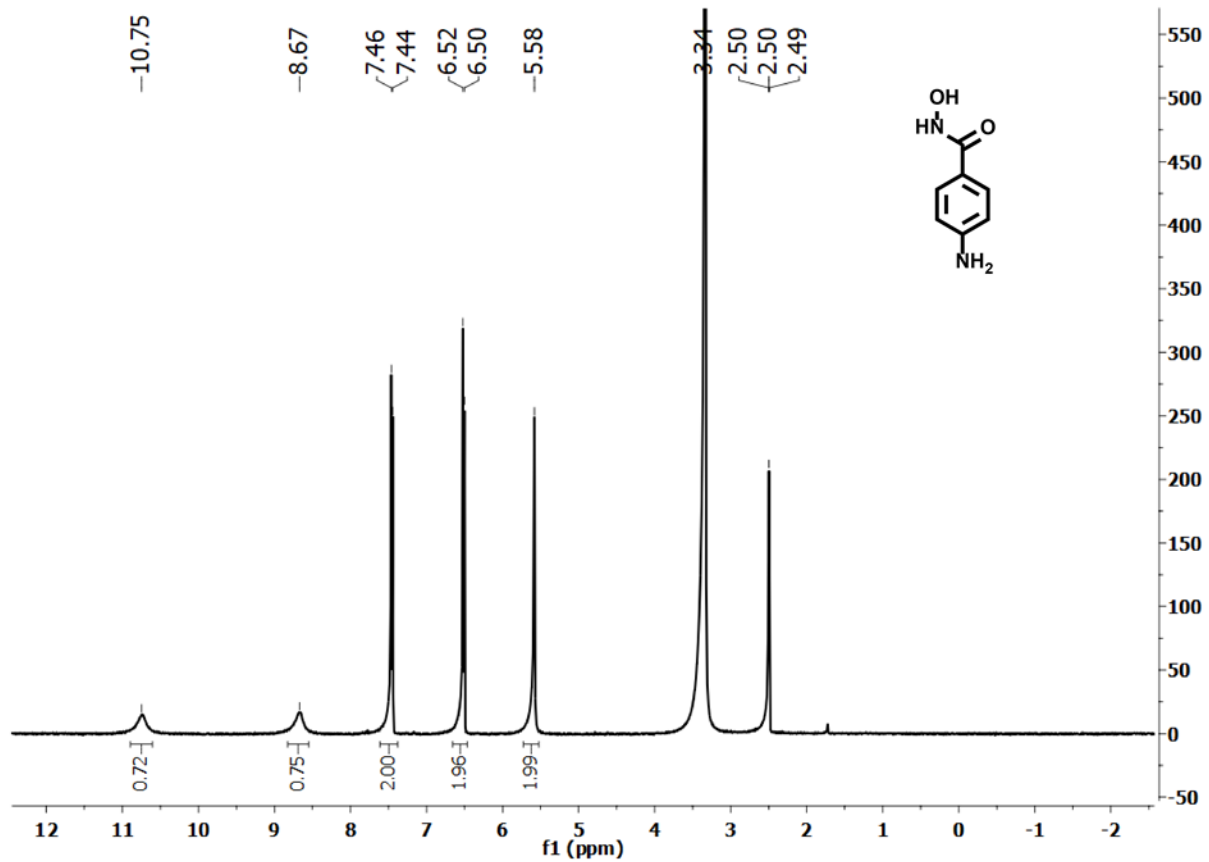


Figure A2.40. ^1H NMR spectrum of 4-amino-*N*-hydroxybenzamide in DMSO-d_6 .

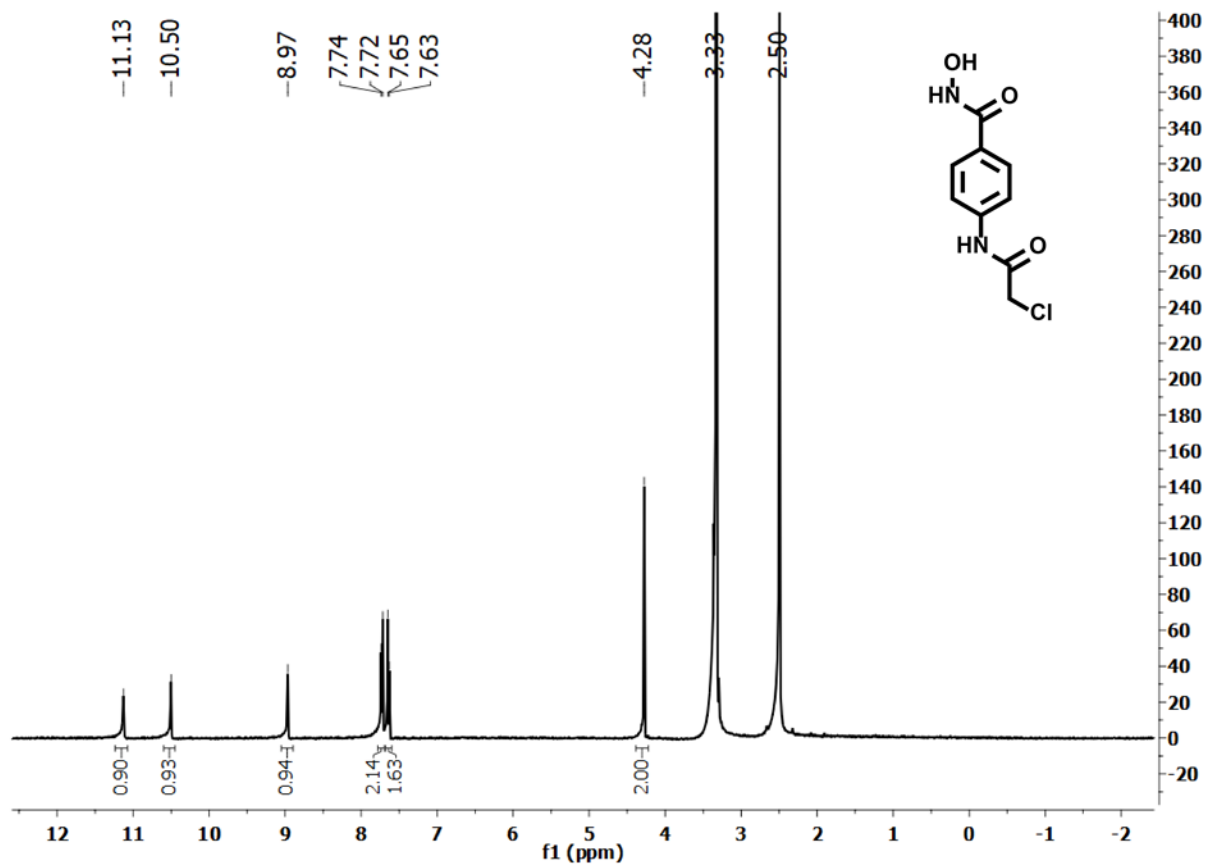


Figure A2.41. ¹H NMR spectrum of 4-(2-chloroacetamido)-N-hydroxybenzamide in DMSO-d₆.

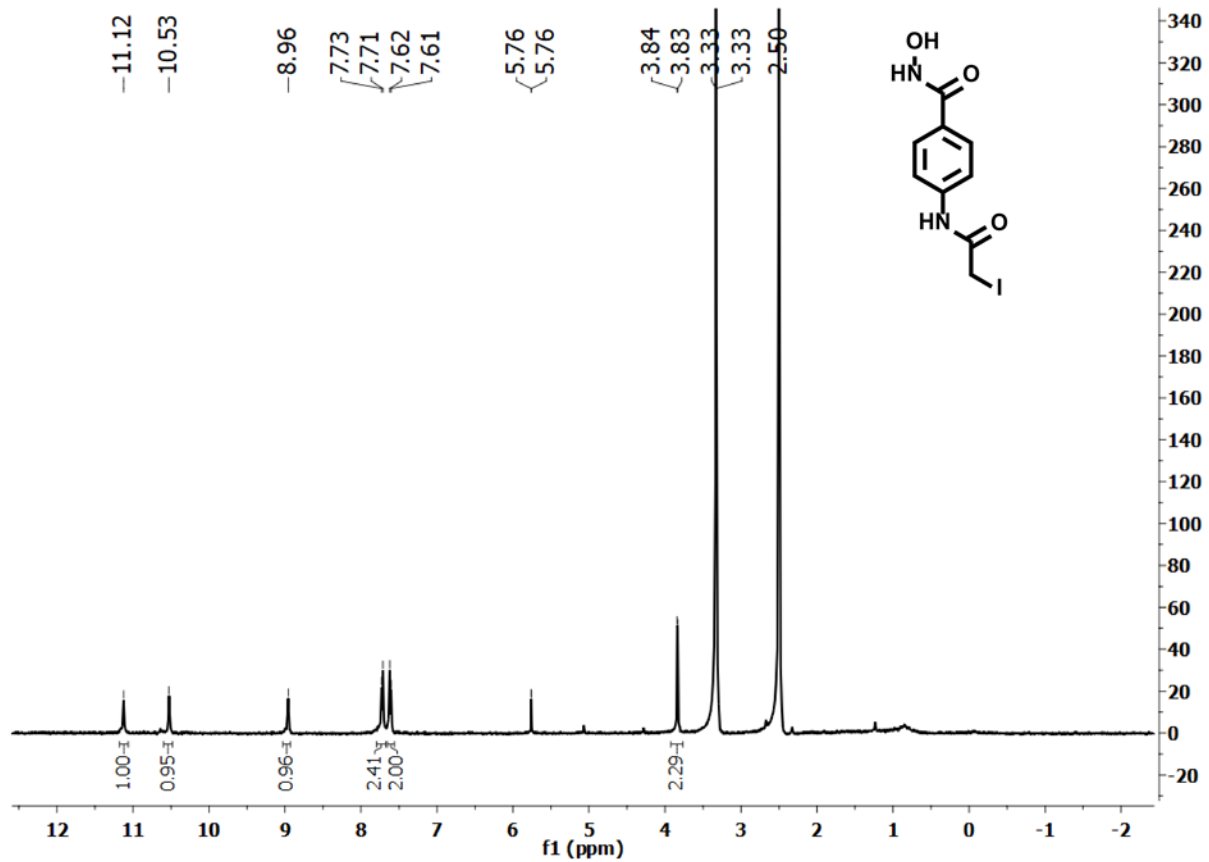


Figure A2.42. ¹H NMR spectrum of 4-(2-iodoacetamido)-*N*-hydroxybenzamide in DMSO-*d*₆.

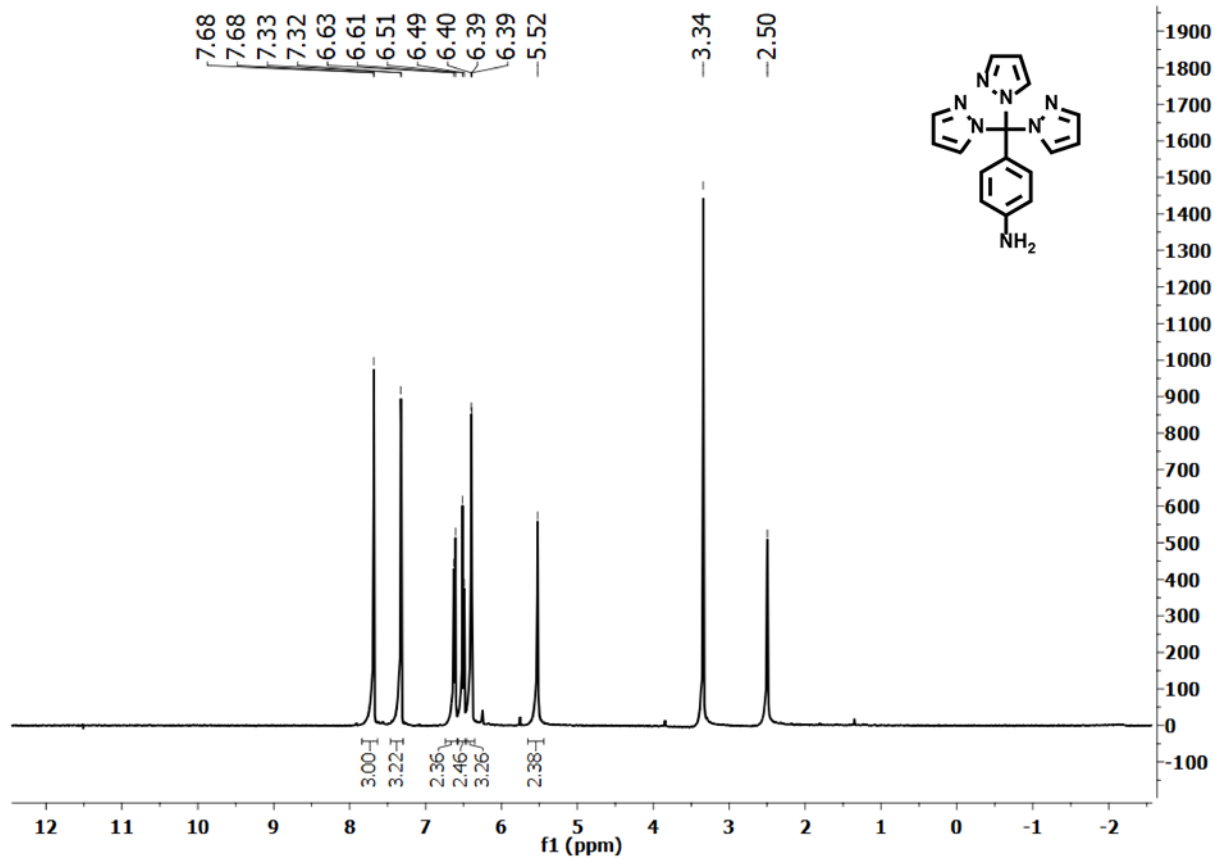


Figure A2.43. ¹H NMR spectrum of 4-(tri(1H-pyrazol-1-yl)methyl)aniline in DMSO-d₆.

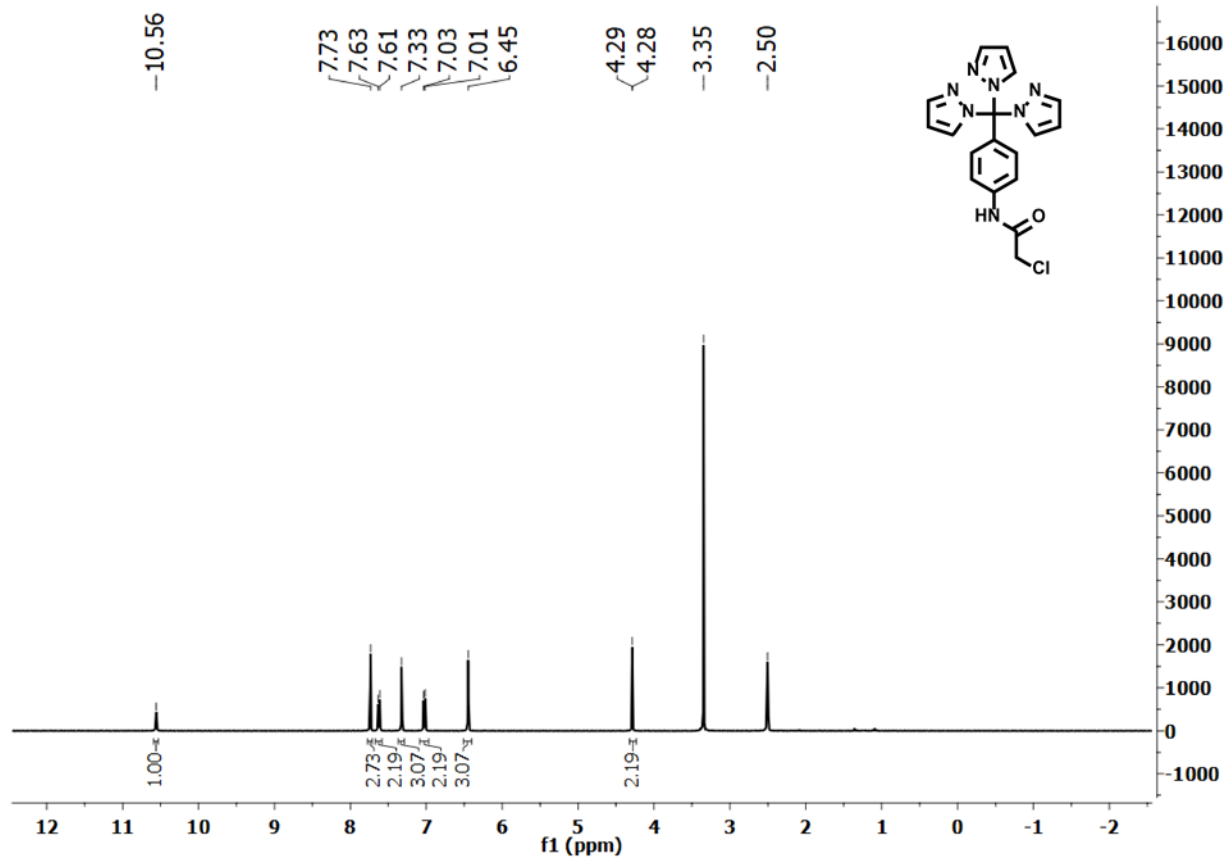


Figure A2.44. ^1H NMR spectrum of 2-chloro-*N*-(4-(tri(1*H*-pyrazol-1-yl)methyl)phenyl)acetamide in DMSO-d_6 .

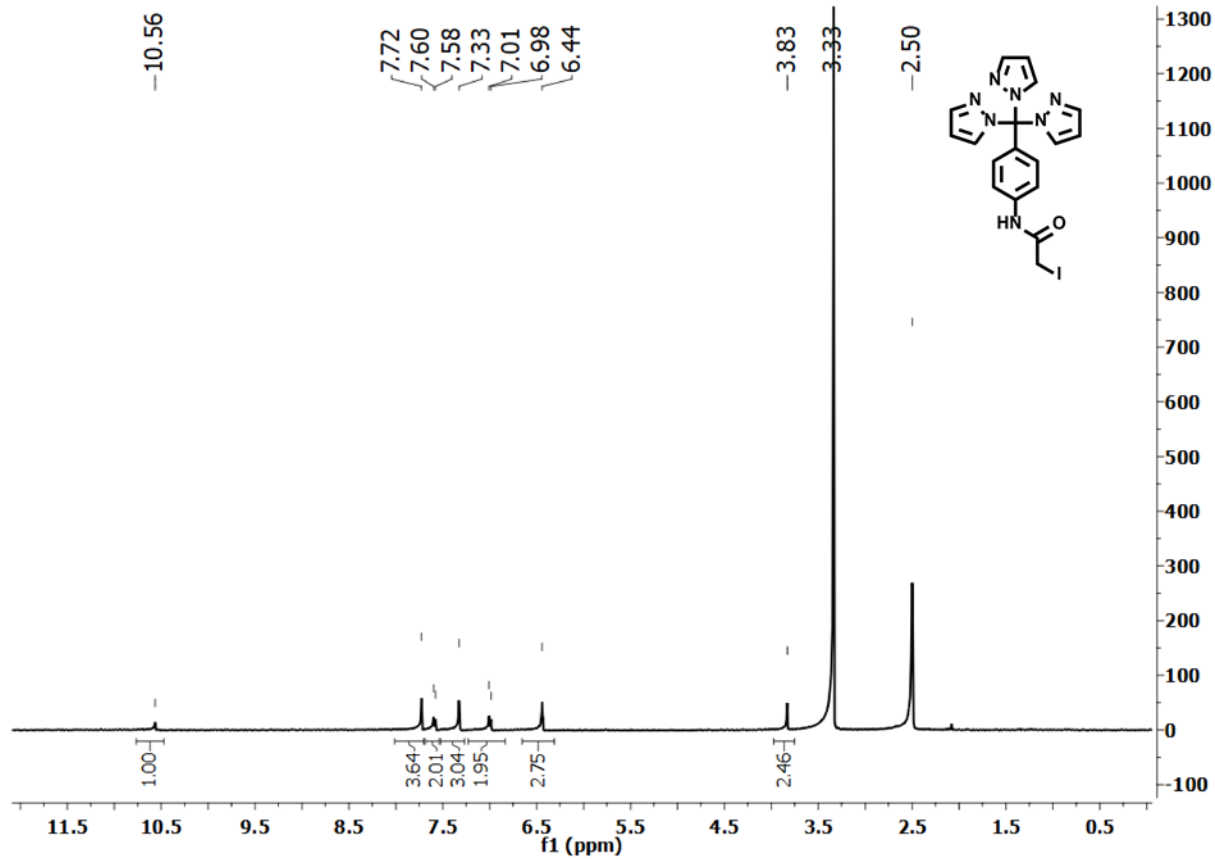


Figure A2.45. ^1H NMR spectrum of 2-iodo-*N*-(4-(tri(1*H*-pyrazol-1-yl)methyl)phenyl)acetamide in DMSO-d_6 .

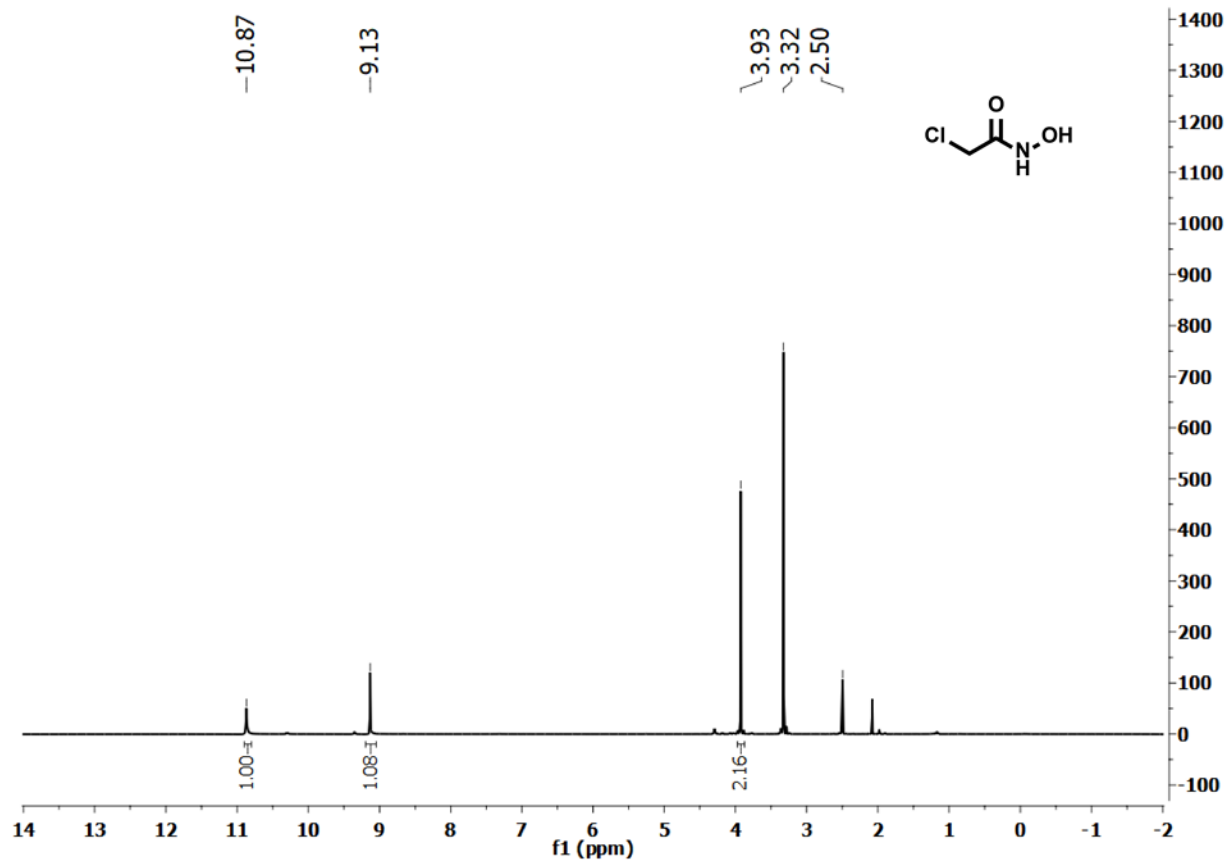


Figure A2.46. ¹H NMR spectrum of 2-chloro-N-hydroxyacetamide in DMSO-d₆.

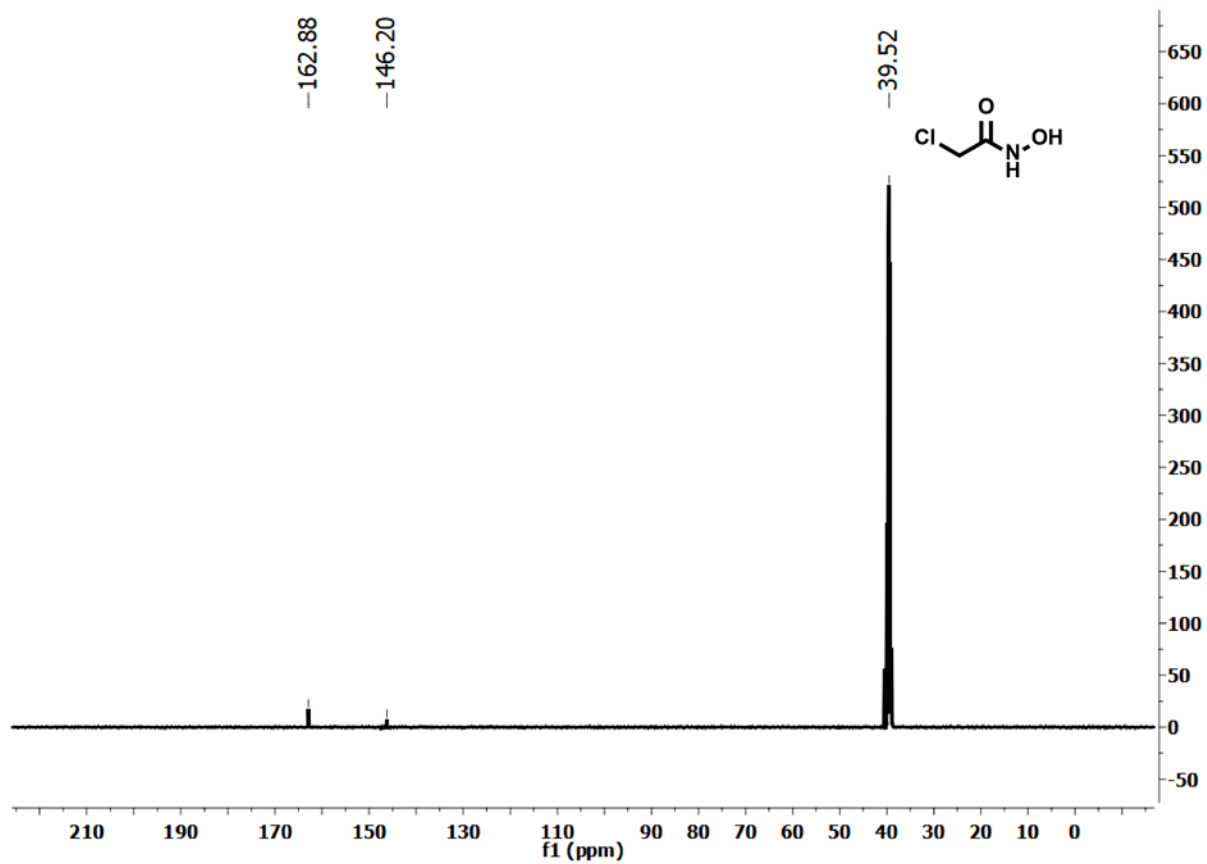


Figure A2.47. ^{13}C NMR spectrum of 2-chloro-*N*-hydroxyacetamide in $\text{DMSO-}d_6$.

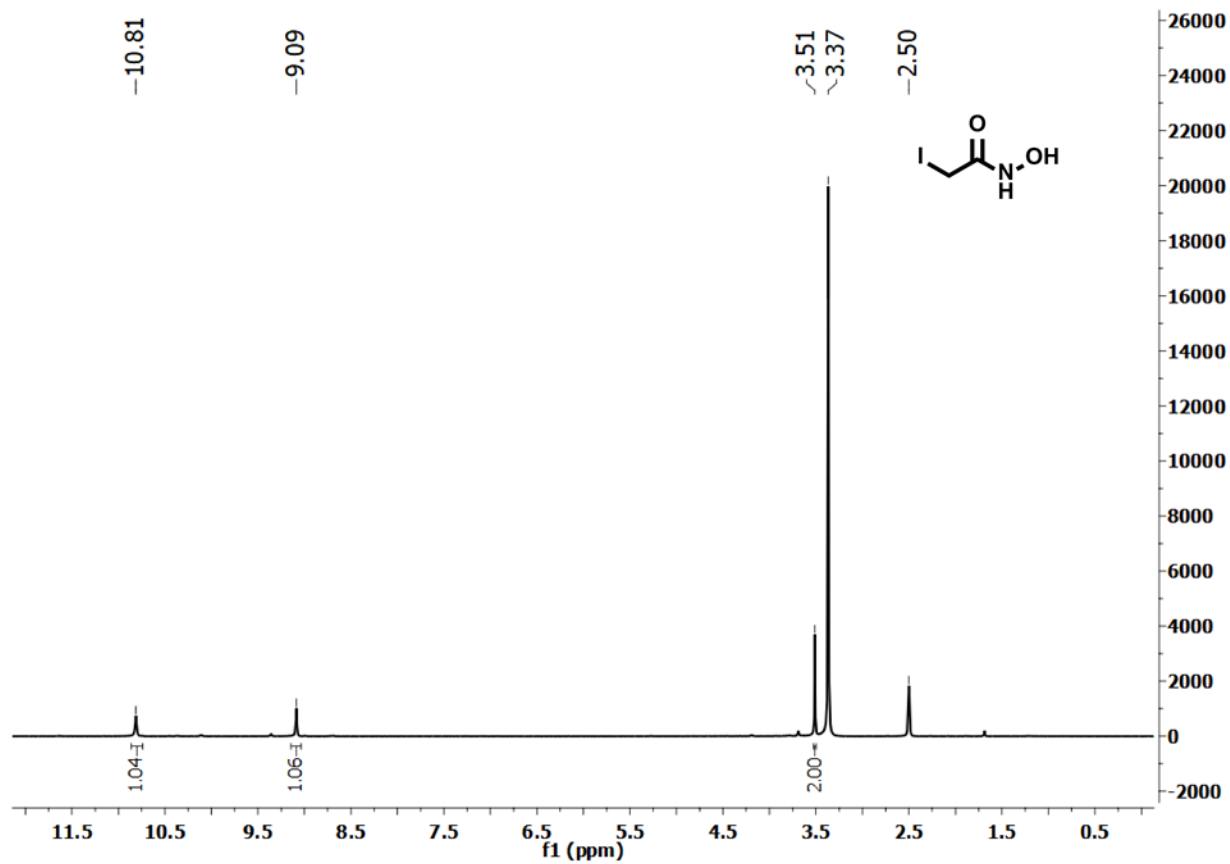


Figure A2.48. ¹H NMR spectrum of 2-iodo-N-hydroxyacetamide in DMSO-d₆.

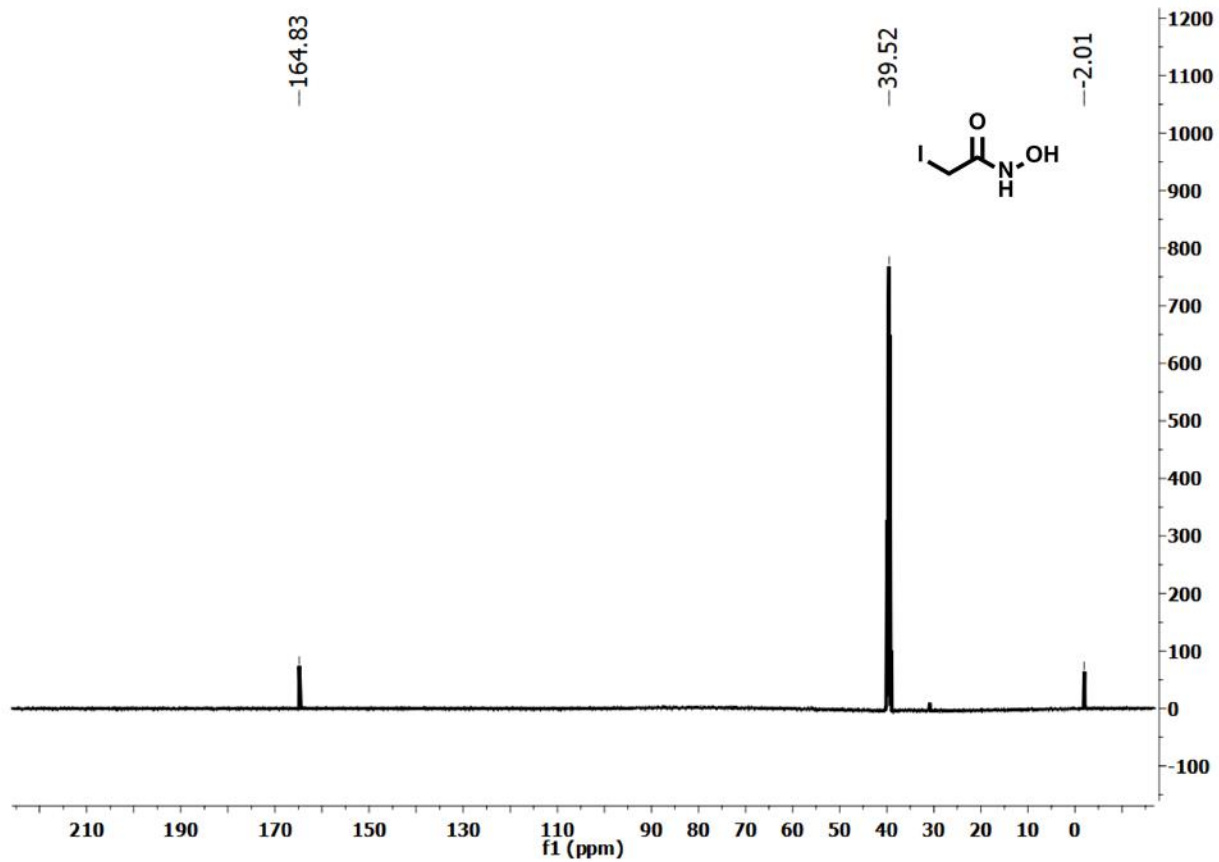


Figure A2.49. ^{13}C NMR spectrum of 2-iodo-*N*-hydroxyacetamide in DMSO-d_6 .

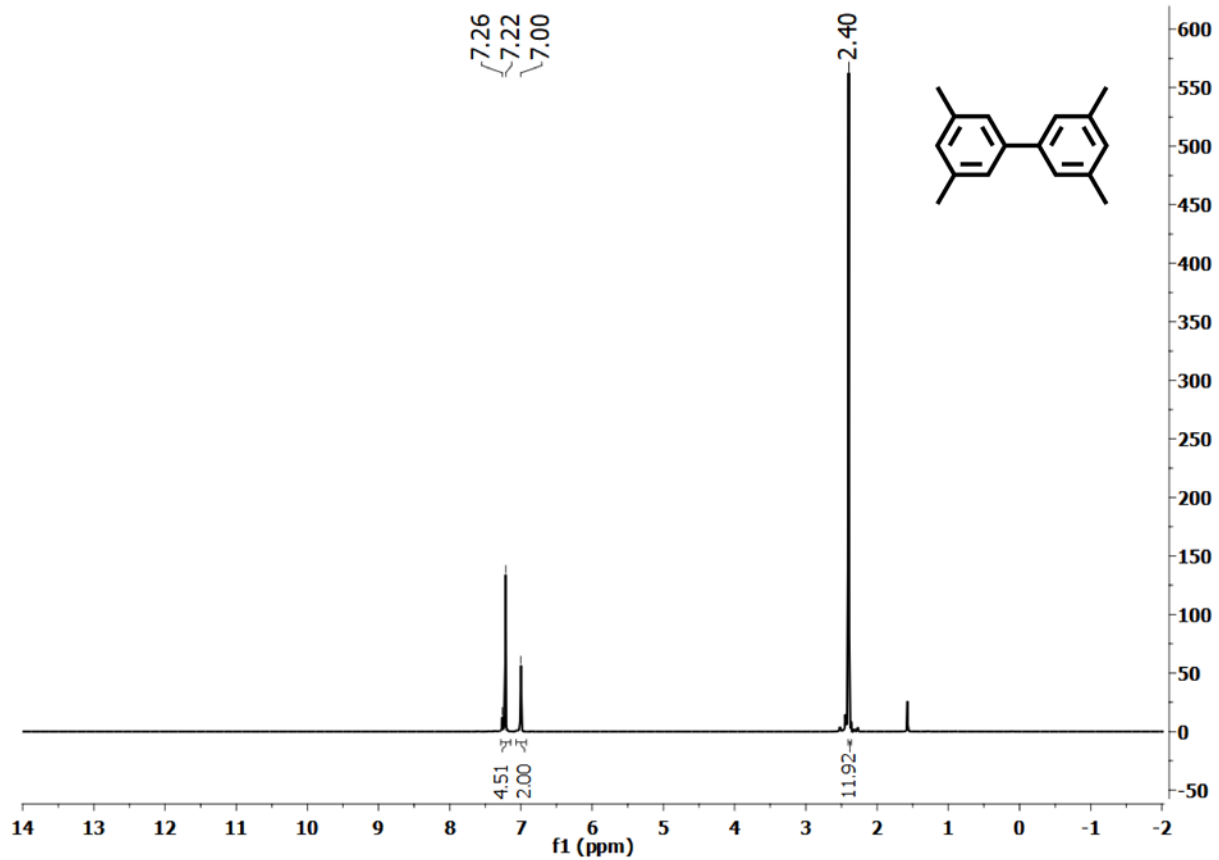


Figure A2.50. ¹H NMR spectrum of 3,3',5,5'-tetramethyl-1,1'-biphenyl in CDCl₃.

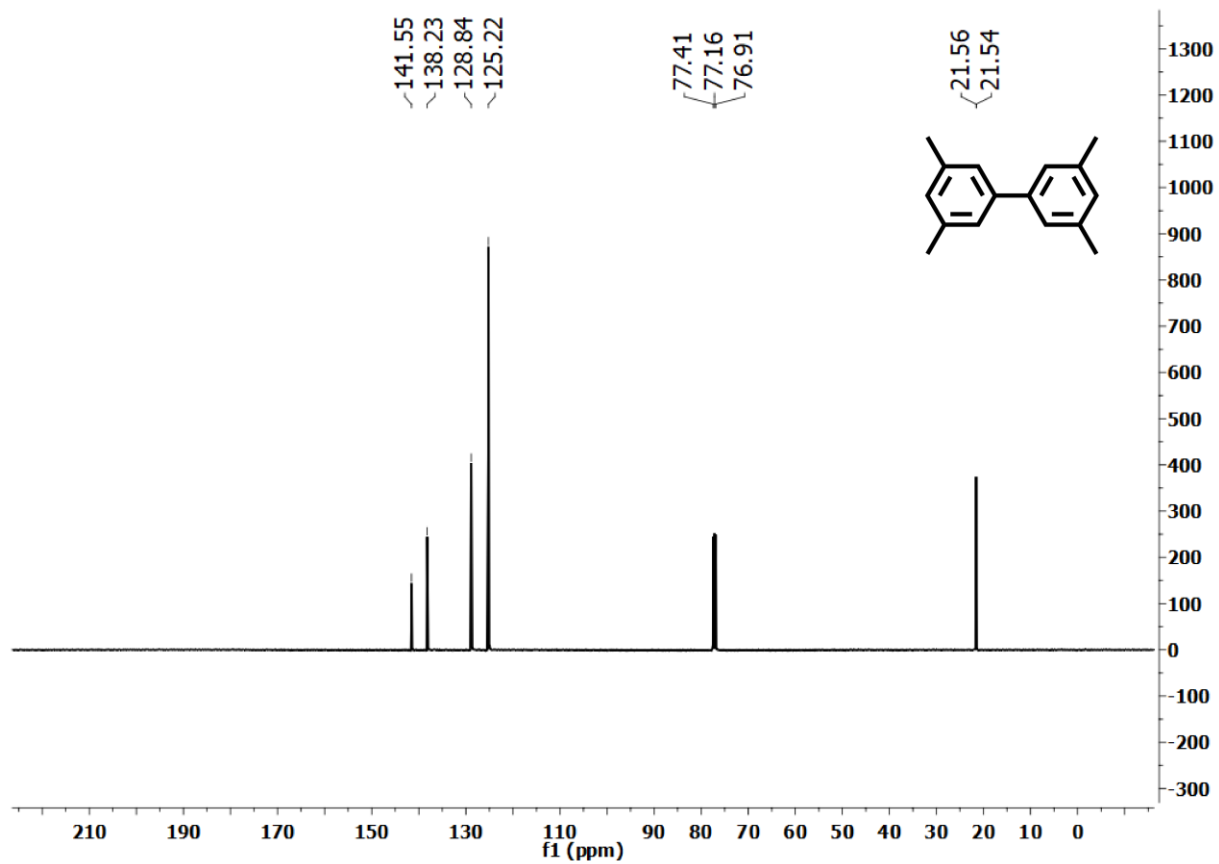


Figure A2.51. ^{13}C NMR spectrum of 3,3',5,5'-tetramethyl-1,1'-biphenyl in CDCl_3 .

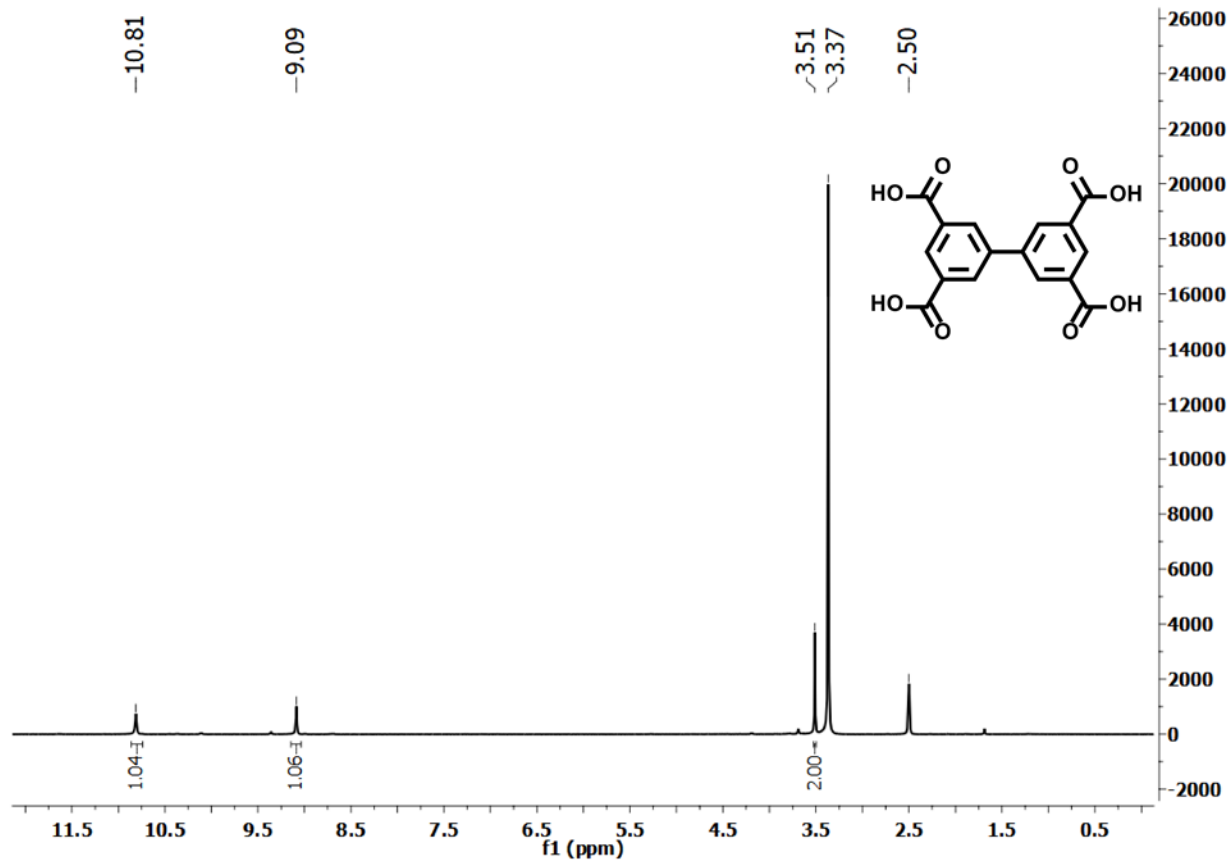


Figure A2.52. ^1H NMR spectrum of [1,1'-biphenyl]-3,3',5,5'-tetracarboxylic acid in DMSO-d_6 .

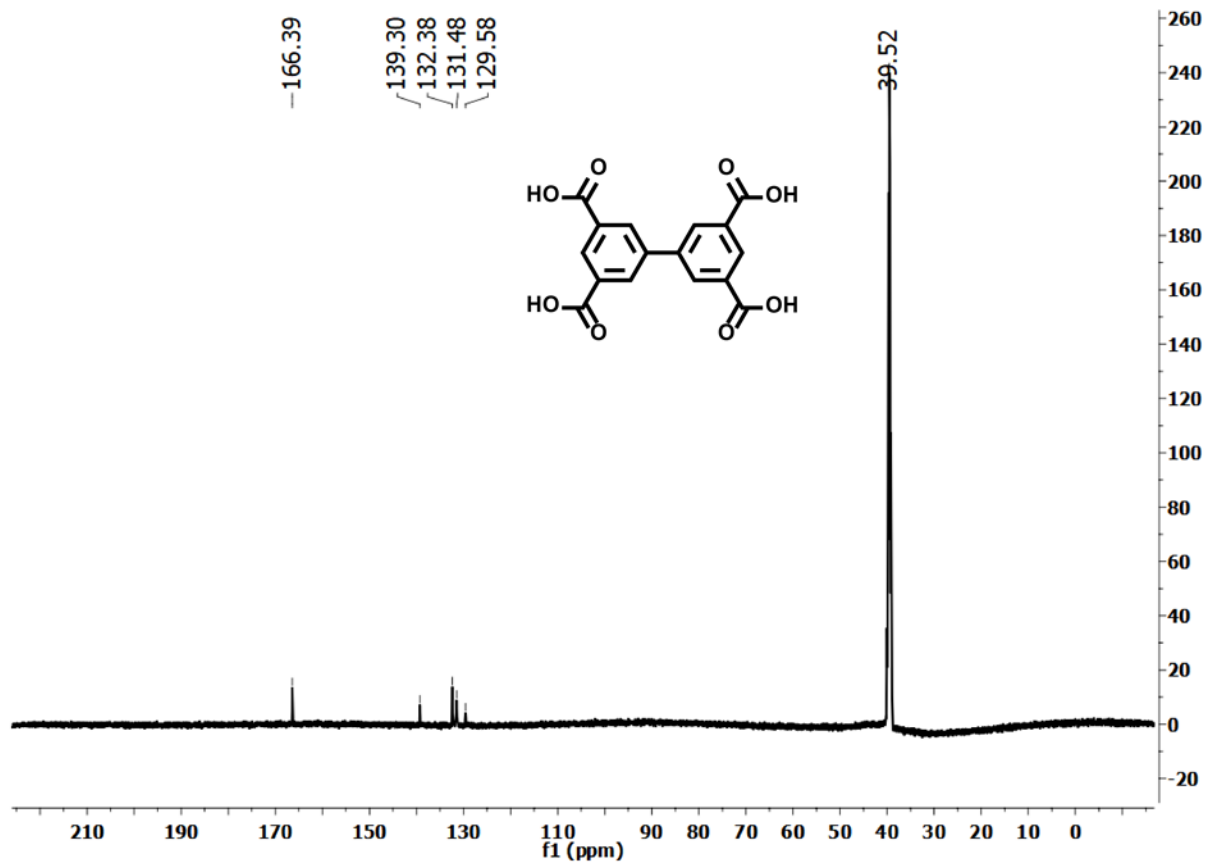


Figure A2.53. ^{13}C NMR spectrum of [1,1'-biphenyl]-3,3',5,5'-tetracarboxylic acid in DMSO-d_6 .

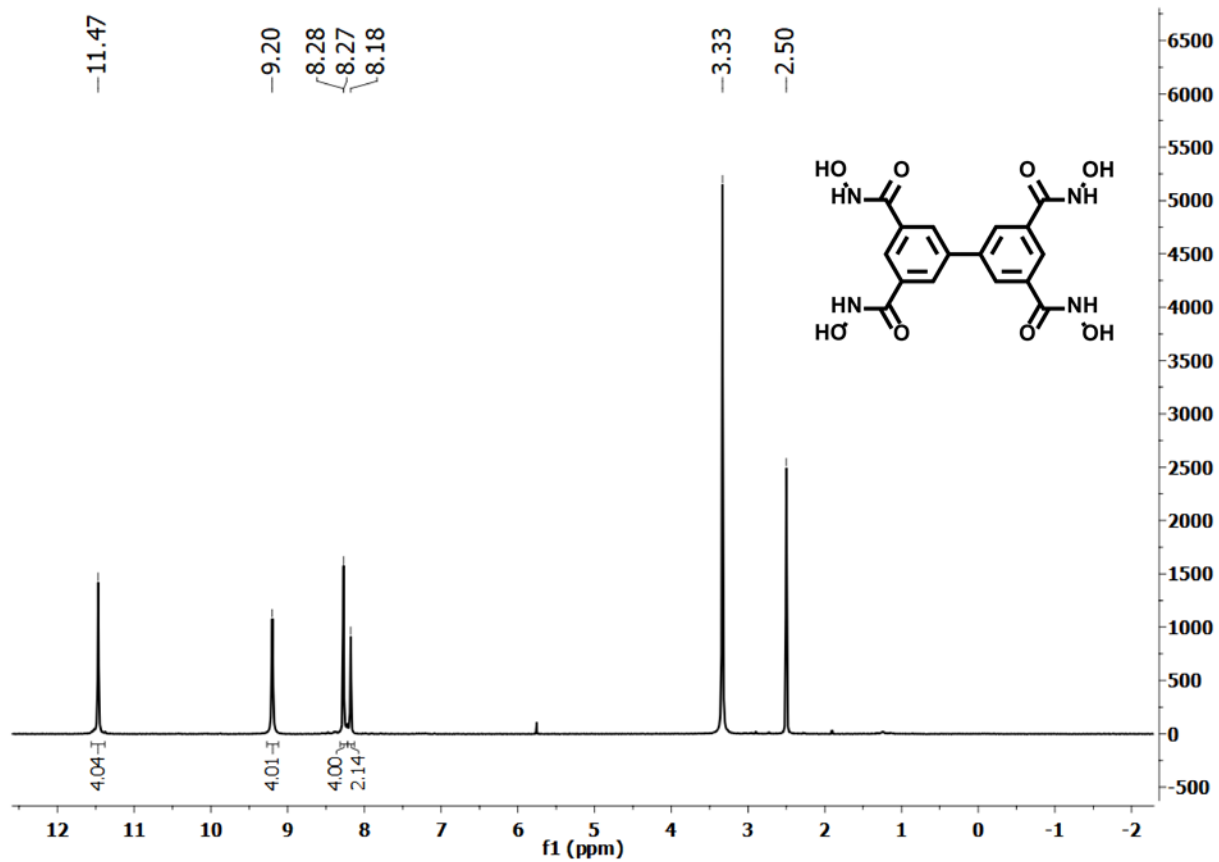


Figure A2.54. ¹H NMR spectrum of *N*^β,*N*^{β'},*N*^δ,*N*^{δ'}-tetrahydroxy-[1,1'-biphenyl]-3,3',5,5'-tetracarboxamide (H₄BPTH) in DMSO-*d*₆.

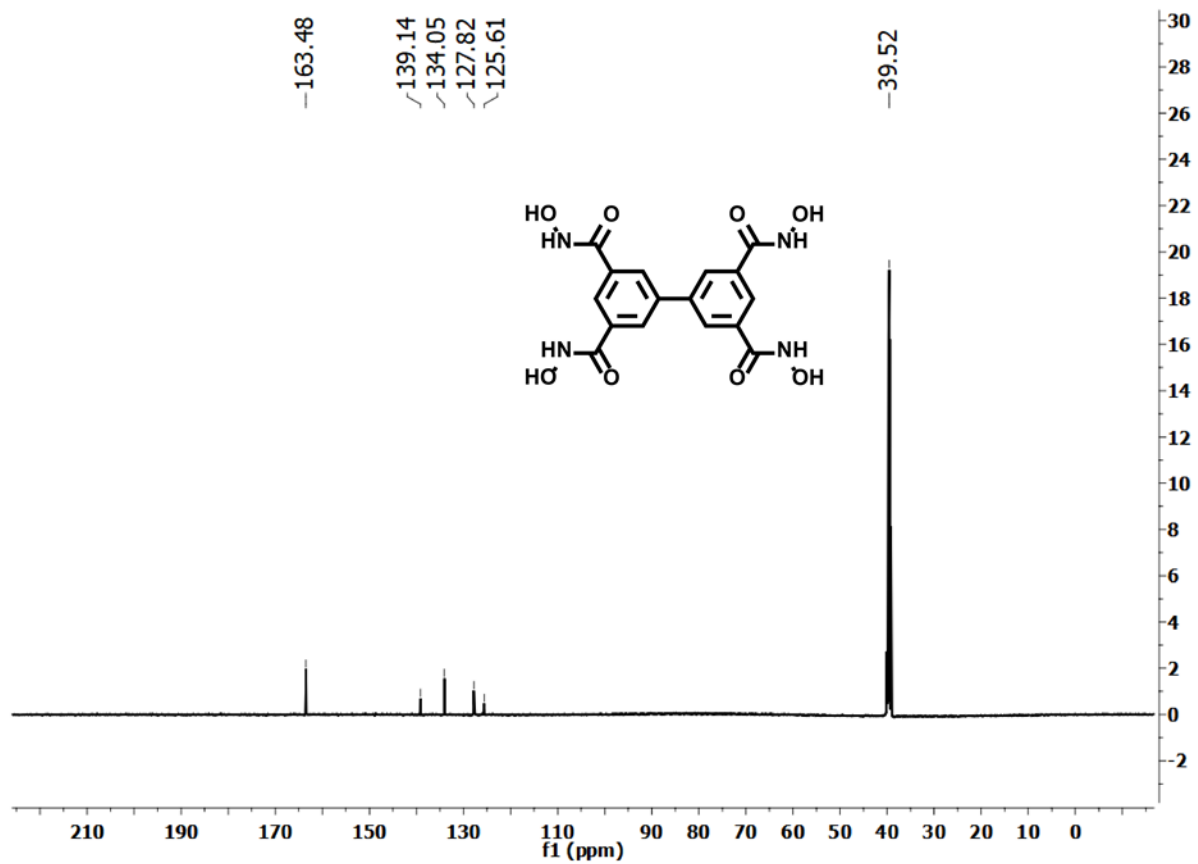


Figure A2.55. ^{13}C NMR spectrum of N^3, N^3', N^5, N^5' -tetrahydroxy-[1,1'-biphenyl]-3,3',5,5'-tetracarboxamide (H₄BPTH) in DMSO- d_6 .

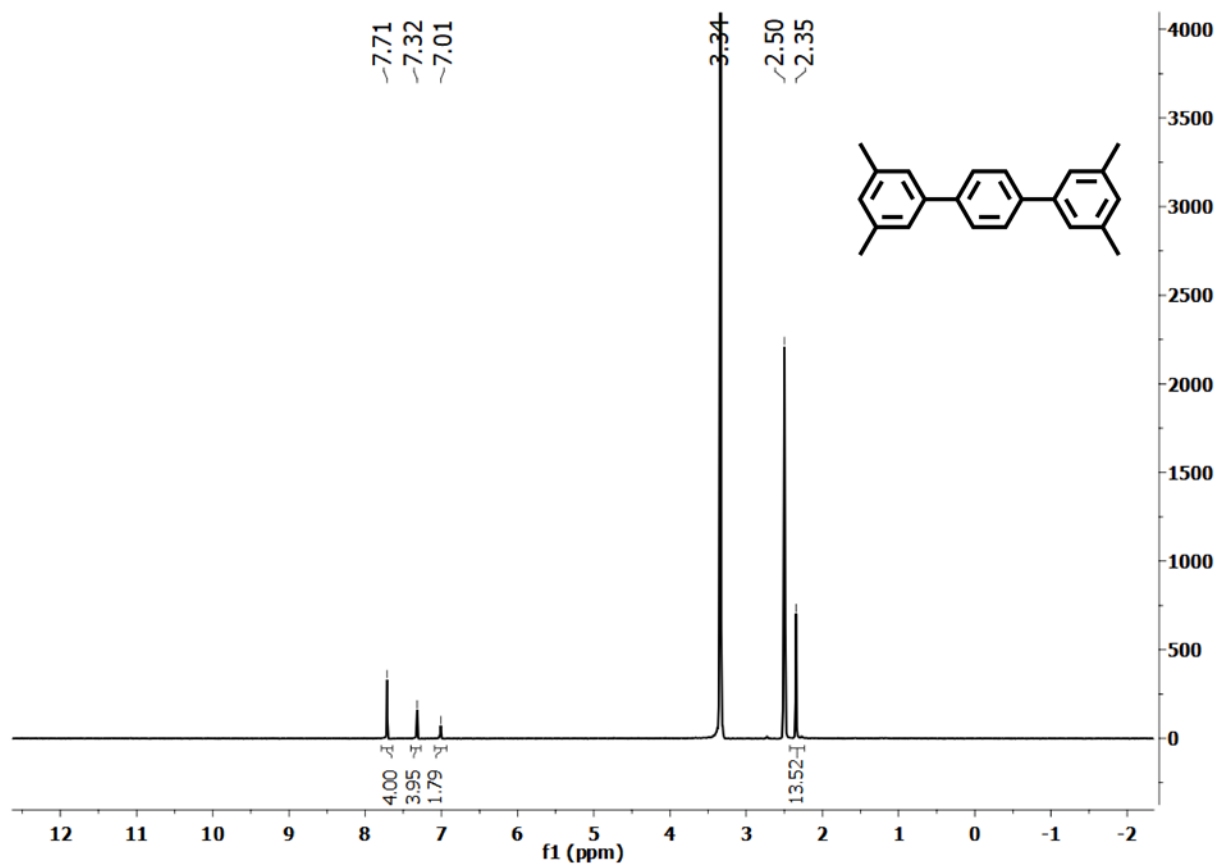


Figure A2.56. ¹H NMR spectrum of 3,3',5,5'-tetramethyl-1,1':4',1''-terphenyl in DMSO-d₆.

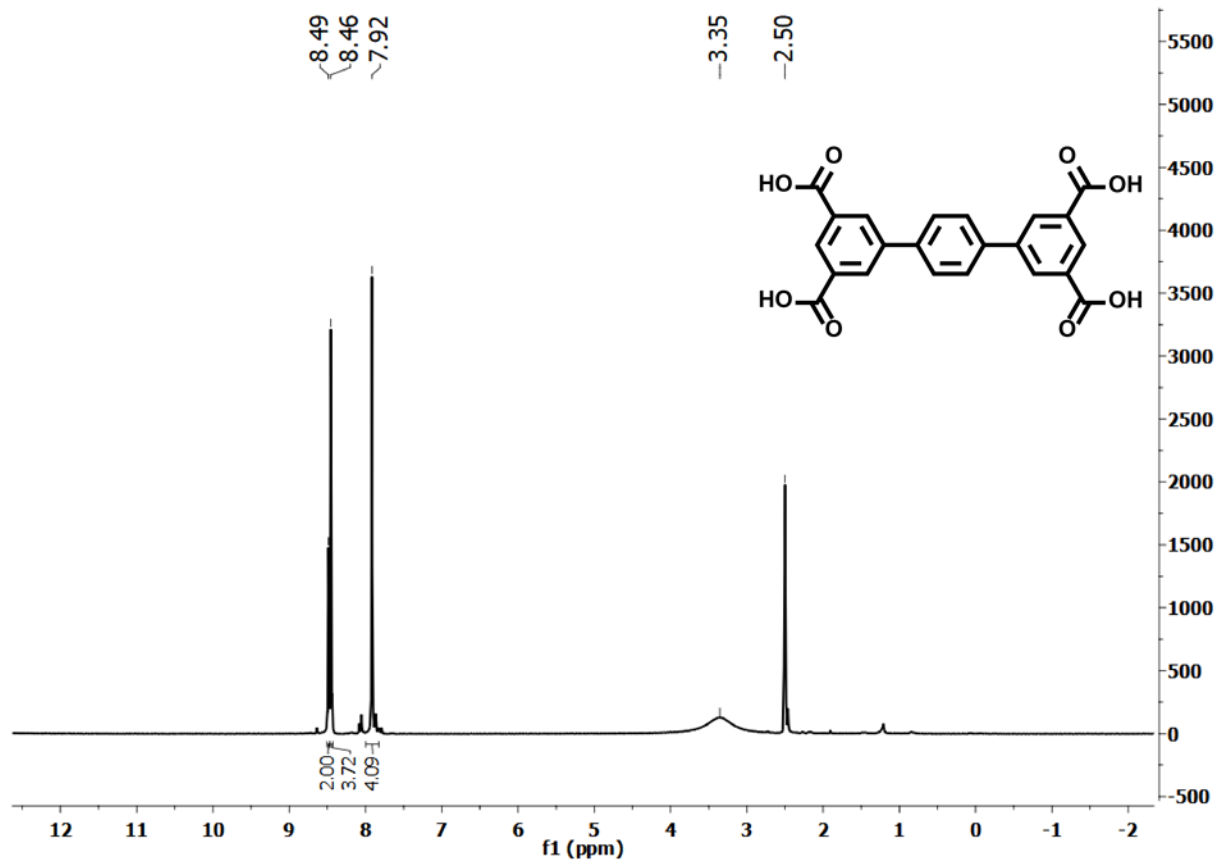


Figure A2.57. ¹H NMR spectrum of [1,1':4,1''-terphenyl]-3,3'',5,5''-tetracarboxylic acid in DMSO-d₆.

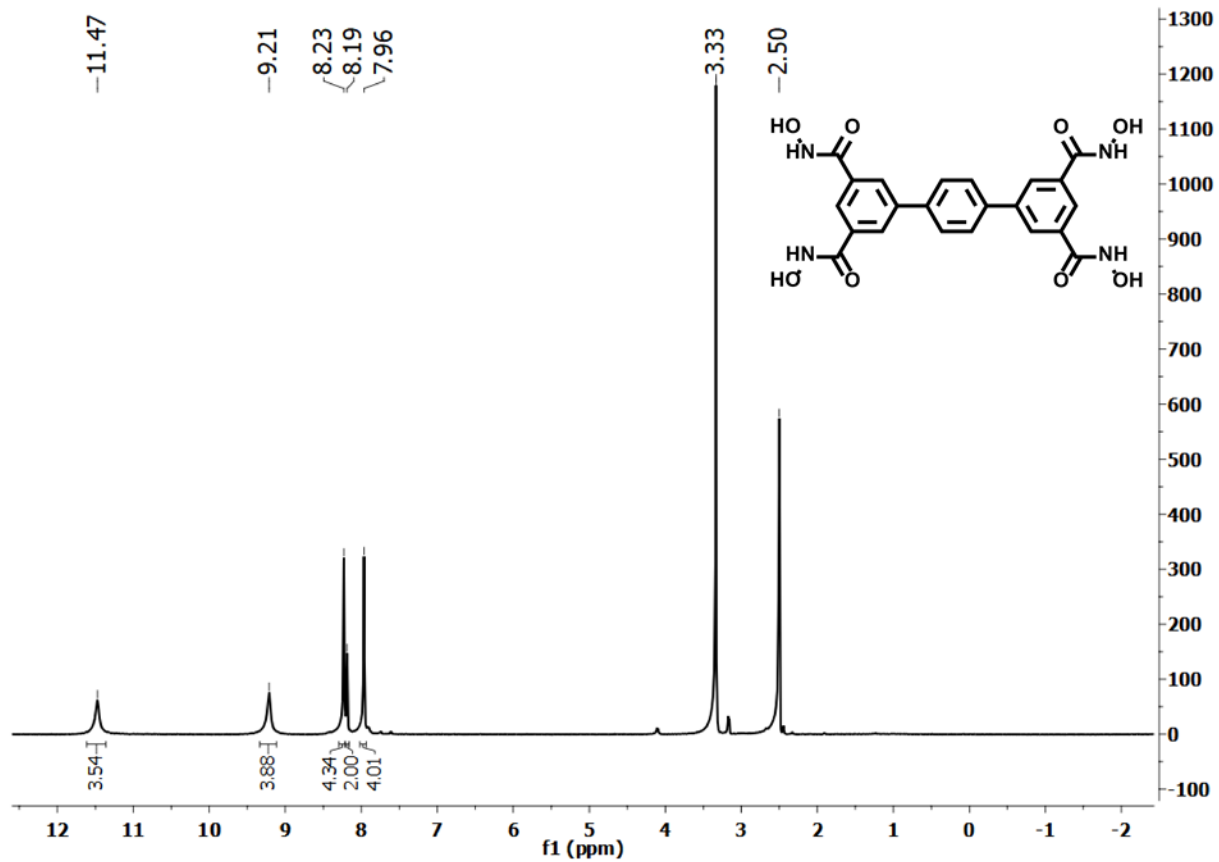


Figure A2.58. ¹H NMR spectrum of *N*^β,*N*^{β'},*N*^{β''},*N*^{β'''}-tetrahydroxy-[1,1':4',1''-terphenyl]-3,3'',5,5''-tetracarboxamide (H₄TPTH) in DMSO-d₆.

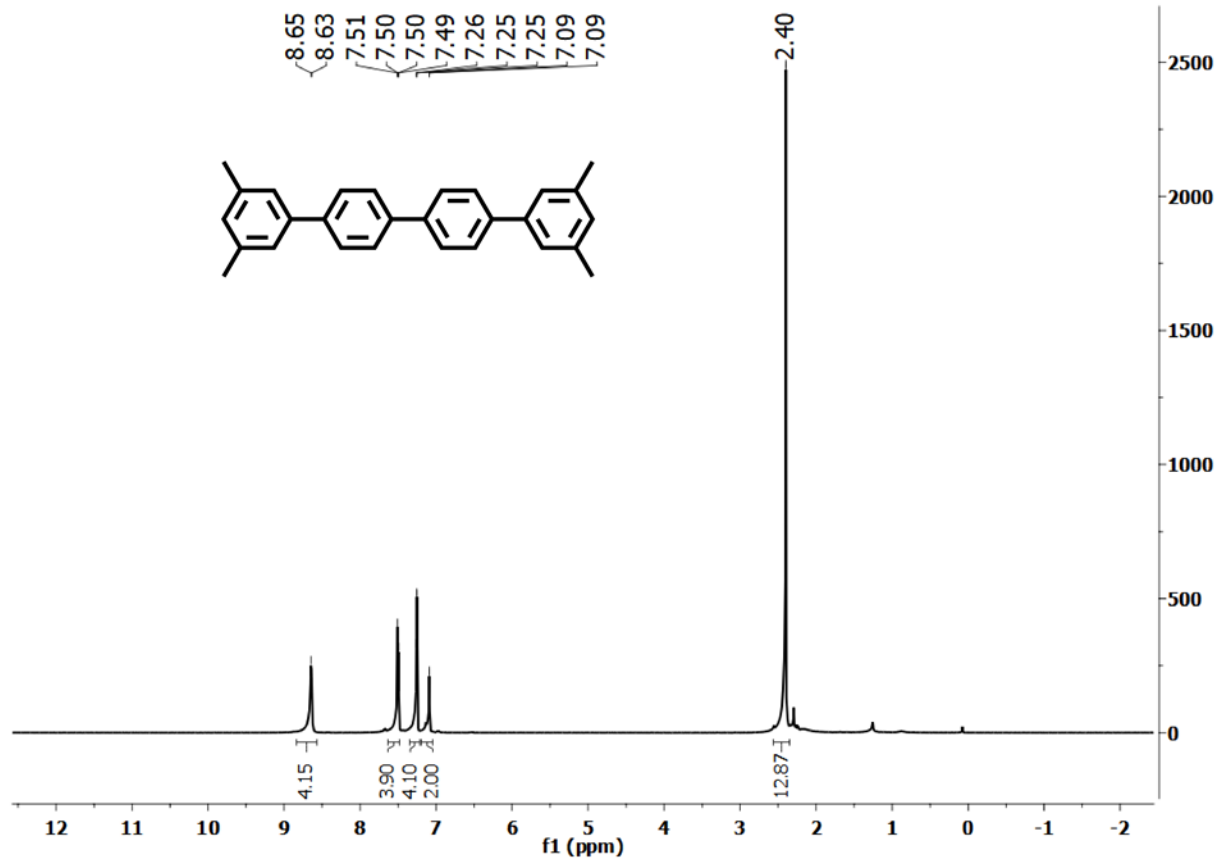


Figure A2.59. ¹H NMR spectrum of 3,3',5,5'-tetramethyl-1,1':4',1'':4'',1'''-quaterphenyl in CDCl₃.

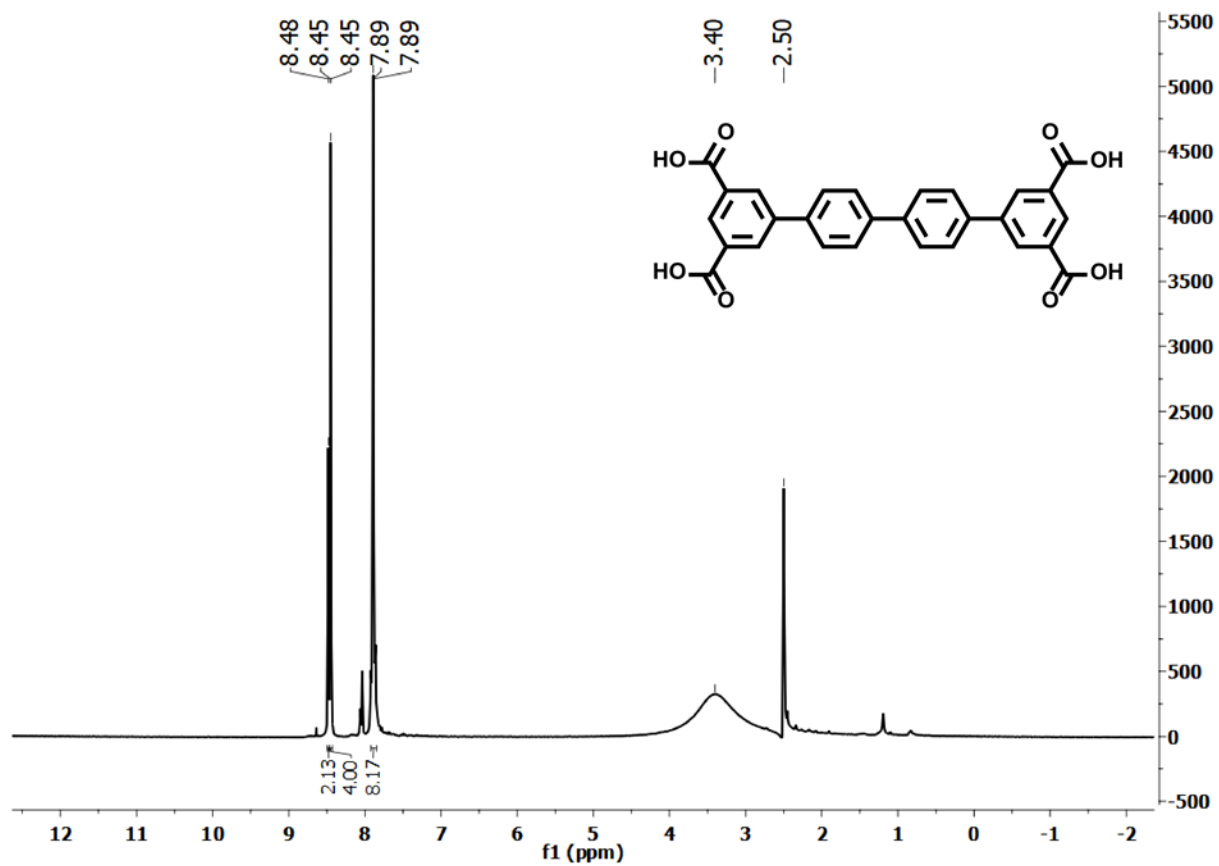


Figure A2.60. ¹H NMR spectrum of [1,1':4',1'':4'',1''':4''',1''''-quaterphenyl]-3,3''',5,5''''-tetracarboxylic acid in DMSO-d₆.

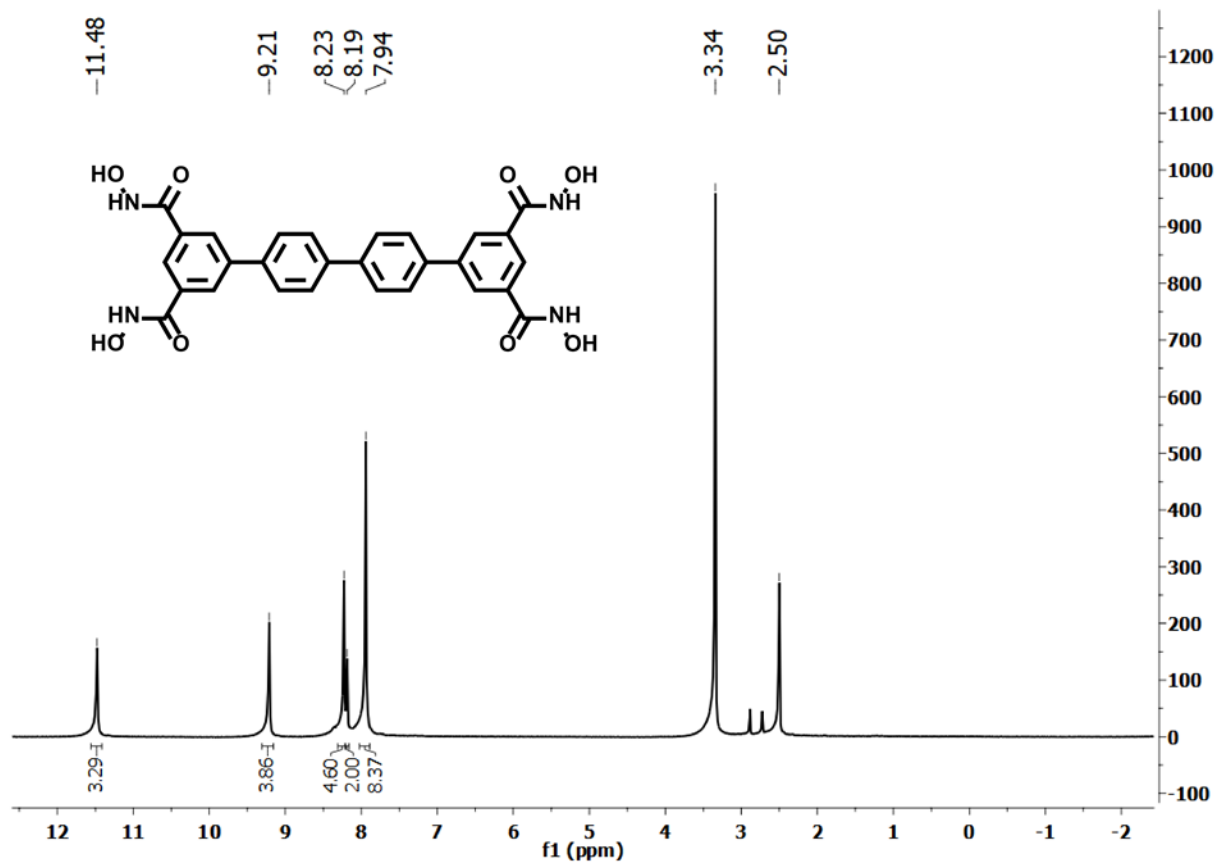


Figure A2.61. ¹H NMR spectrum of *N*³,*N*^{3''},*N*⁶,*N*^{6''}-tetrakis(trityloxy)-[1,1':4',1'':4'',1''':4''',1''''-quaterphenyl]-3,3''',5,5'''-tetracarboxamide (H₄QPTH) in DMSO-*d*₆.