STRUCTURAL AND MECHANISTIC STUDIES ON PROLYL HYDROXYLASES

A thesis submitted to the Board of the Faculty of Physical Sciences at the University of Oxford in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Rasheduzzaman Chowdhury
Hertford College, Oxford
Trinity Term 2008

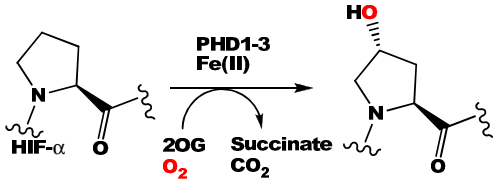
Rasheduzzaman Chowdhury
Hertford College

D. Phil.
Trinity Term 2008

STRUCTURAL AND MECHANISTIC STUDIES ON PROLYL HYDROXYLASES

Abstract

Oxygen dependent prolyl-4-hydroxylation of the α-subunit of the hypoxia inducible transcription factor (HIF-α) plays an essential role in the hypoxic response.[1,2] Hydroxylation of proline residues in the N- or C-terminal oxygen dependent degradation domains (NODD or CODD) increases the affinity of HIF-α to the von Hippel-Lindau protein (pVHL) by ~1000 fold so signalling for HIF-α degradation.[3] With limiting oxygen, HIF-α hydroxylation slows, it dimerises with HIF-β and activates the transcription of a gene array.[4] Prolyl-4-hydroxylation also stabilises the triple helix structure of collagen, the most abundant human protein.[5] Both the collagen and the HIF prolyl hydroxylases (PHDs) are Fe(II) and 2-oxoglutarate (2OG) dependent oxygenases.[6]



Crystal structures of PHD2 in complex with CODD were determined in the current study. Together with biochemical analyses, the results demonstrate that catalysis involves a mobile region of PHD2 that encloses the hydroxylation site and stabilises the PHD2.Fe(II).2OG complex. When bound to PHD2 the pyrrolidine ring of the non-hydroxylated proline-residue adopts a C^4-endo conformation. Evidence is provided that 4*R*-hydroxylation enables a stereoelectronic effect that changes the proline conformation to the C^4-exo state, as observed when hydroxylated HIF-α is bound to pVHL and in collagen. The results help to rationalise NODD/CODD selectivity data for PHD isoforms and the effects of clinically observed mutations on PHD2 catalysis. Analyses on the interaction of nitric oxide with PHD2 are described and discussed with respect to regulation of the hypoxic response by nitric oxide.

1. Epstein, A.C., Gleadle, J.M., McNeill, L.A., Hewitson, K.S., O'Rourke, J., Mole, D.R., Mukherji, M., Metzen, E., Wilson, M.I., Dhanda, A., Tian, Y.M., Masson, N., Hamilton, D.L., Jaakkola, P., Barstead, R., Hodgkin, J., Maxwell, P.H., Pugh, C.W., Schofield, C.J. & Ratcliffe, P.J. C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell* **107**, 43-54 (2001).
2. Schofield, C.J. & Ratcliffe, P.J. Oxygen sensing by HIF hydroxylases. *Nat Rev Mol Cell Biol* **5**, 343-54 (2004).
3. Jaakkola, P., Mole, D.R., Tian, Y.M., Wilson, M.I., Gielbert, J., Gaskell, S.J., Kriegsheim, A., Hebestreit, H.F., Mukherji, M., Schofield, C.J., Maxwell, P.H., Pugh, C.W. & Ratcliffe, P.J. Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. *Science* **292**, 468-72 (2001).
4. Semenza, G.L. Hydroxylation of HIF-1: oxygen sensing at the molecular level. *Physiology (Bethesda)* **19**, 176-82 (2004).
5. Jenkins, C.L. & Raines, R.T. Insights on the conformational stability of collagen. *Nat Prod Rep* **19**, 49-59 (2002).
6. Clifton, I.J., McDonough, M.A., Ehrismann, D., Kershaw, N.J., Granatino, N. & Schofield, C.J. Structural studies on 2-oxoglutarate oxygenases and related double-stranded beta-helix fold proteins. *J Inorg Biochem* **100**, 644-69 (2006).

Acknowledgements

I'm indebted to my supervisor, Prof. Chris Schofield for giving me the opportunity to work on an exciting project, for almost day-to-day supervision, encouragement, expert analyses and for all the support not only confined to academic but also to the personal level.

My greatest thanks to Dr Emily Flashman and Dr Mike McDonough for being my biochemistry and crystallography mentors. Without their adroit supervision, endless support and constructive criticism, I could not make this far. Indeed, "In the world of crystallography (almost) everything seems to be model biased". Special thanks to Dr Ian Clifton and Adam for all (apparent) software related problems.

I thank Emma Longman (Biochemistry, Oxford) for assistance with the sedimentation velocity experiments, Tom Walter (OPPF, Oxford) and Oleg Fedorov (SGC, Oxford) for assistance with crystallisation robotics and thermofluor assays; E.D. Lowe (LMB, Oxford) for assistance in data collection; and the scientists of beamlines ID23-2 (ESRF), 10.1 (SRS) and I04 (Diamond), the Oxford Supercomputer Centre and the UK National HPC facilities for technical assistance.

I appreciate the last minute assistance of my proof-reader squad: Celia, Emily and Mike who had much better things to do than going through my thesis over and over again. Special thanks go again to Celia, Mike, Wei Ge and Maggie for offering me their help to childmind my kids, to Emily and Aman for the support in my hardest times.

Needless to say that I am grateful to all of my workmates and friends within and outside CJS group for their support (and at times tolerance to my loud music). I came across a number of wonderful peoples over the past four years of D. Phil. whom I owe big thanks (in no particular order): Ed, Jasmin (for the criticism, long talks on craving for papers), Tom B, Steve, Nadia, Danica, D. Lancaster (for the scientific debates on FIH vs. PHD and 'cryptic' digestion), Ben, Anna, David Genn, Kirsty (my sole source of plasmids), Dom, Leanne, Ran, Christoph, Nathan, Eleanor, Ming, Mukram, Wendy, Val, Manso and Zhihong. And of course, special thanks to my rock buddies: Aman (for all times), TG, Rich and Luke. Thanks to many great friends both in Oxford and further a field who were always encouraging and good company. The following peoples deserve a special mention, Amit, Azad bhai, Abdullah bhai, Liza apa and Mithudi plus many more (my apologies for not mentioning all the names), and my Grad Centre buddies, Mike (Lokale) and Mairin.

I acknowledge the financial support from Commonwealth Scholarship 2004 and all the help from my academic administrators at the CSC, Anna O'Flynn and at the British Council, Mike Bray for their support.

And last but never the least, I appreciate the support from my wife for being the best friend/ companion ever, my kids for being my biggest source of inspiration and my parents/ in-laws for their constant blessings.

Contents

Contents i-x

Abbreviations xi-xii

List of contributors xiii

Chapter 1 – Introduction

1.1 Physiological background 1

1.2 Hypoxia-Inducible Factor (HIF): Discovery of a master regulator of oxygen homoeostasis 1

1.3 HIF target genes and the hypoxic response 2

1.4 Domain architecture and isoforms of HIF 3
- 1.4.1 HIF-α isoforms 3
- 1.4.2 Domain architecture of HIF subunits 4

1.5 DNA binding and dimerisation of HIF 4

1.6 Regulation of HIF activity by oxygen – basic components of hypoxic system 7
- 1.6.1 Prolyl hydroxylation of the central oxygen dependent degradation domains (ODDs) 7
- 1.6.2 Structure of the HIF targeting E3 ubiquitin ligase and implications 8
- 1.6.3 Asparaginyl hydroxylation of transactivation domains (CAD) 9

1.7 Regulation of HIF stability via the oxygen independent pathway 11

1.8 Factors affecting HIF levels and activity 11

1.9 The HIF hydroxylases 14

1.10 Functional differences between PHD isoforms 17
- 1.10.1 Variable expression of PHDs 17
- 1.10.2 Differential subcellular localisation of PHDs 17
- 1.10.3 PHD2 is the key regulator in normoxia 18

1.11 Other potential substrates for the HIF hydroxylases 18

1.12 Genetic diseases associated with HIF 19

1.13 Possibilities for therapeutic intervention 20

1.14 Objectives of the current research 21

1.15 References 22

Chapter 2 – Crystallisation and structure solution of PHD2 complexes

2.1	Selection of PHD2 construct for structural work		28
2.2	Crystallisation attempts with tPHD2 that were successful for other related human 2OG oxygenases		29
2.3	Systematic approach to optimise protein sample composition		31
	2.3.1	Pre-crystallisation testing	33
	2.3.2	Thermal shift measurements using thermofluor assays	33
	2.3.3	Sedimentation velocity analysis using analytical ultracentrifugation (AUC)	37
	2.3.4	Verification of complex formation by MS analyses	38
	2.3.5	Inhibition studies	39
2.4	Truncations and related modifications of tPHD2 for crystallisation		40
	2.4.1	Activity and MS analyses of tPHD2Δloop variant	40
2.5	Soaking experiments and structure solution		41
	2.5.1	Soaking with CODD/NODD	41
	2.5.2	Soaking with 2OG/NOG	43
2.6	Cross linking between CODD and tPHD2 in an attempt to form a stable complex		45
	2.6.1	Activity analyses of tPHD2 I280C using CODD/CODD L574C as substrates	46
	2.6.2	MS analyses of tPHD2 I280C–CODD L574C complexes under denaturing conditions	47
	2.6.3	Optimisation of complex formation	48
2.7	Mutations aimed to block homotrimeric crystal packing		50
	2.7.1	Activity and MS analyses of the variants	51
	2.7.2	Crystallisation of variants blocking homotrimer formation	54
2.8	Data collection and structure solution		55
2.9	Optimisation of initial conditions for tPHD2.CODD crystallisation		56
	2.9.1	Thermal stability assays for stabilising the tPHD2 R398A.CODD/NODD complexes	57
2.10	Summary		58
2.11	References		59

Chapter 3 – Structural analyses of tPHD2.CODD complexes

3.1	B-Factor analyses of tPHD2 complexes suggest flexible regions	61
3.2	General structural features and comparison of different tPHD2 complexes	62
3.3	Overall mode of CODD binding	64
3.4	Binding of catalytic $Pro564_{CODD}$ at the tPHD2 active site	65
3.5	Hydrophobic binding sites in tPHD2, including those involving binding of LXXLAP motif	65
3.6	Evidence that a flexible 'β2β3/loop' enables 'open' and 'closed' tPHD2 conformations	66
3.7	Lid closure stabilises the tPHD2.CODD complex for efficient catalysis	68
3.8	Further strategies to obtain a tPHD2 structure with the 'closed' form of the 'β2β3/loop'	71
3.9	The role of tPHD2 C-terminal helix α4 in CODD binding	73
3.10	Additional CODD C-terminal interactions to tPHD2	74
3.11	Summary	75
3.12	References	75

Chapter 4 – Biochemical investigations of tPHD2 for substrate selectivity

4.1	Substrate selectivities amongst the PHDs	76
4.2	Role of 'β2β3/loop' in NODD/CODD selectivity	76
4.3	MDS studies predict a similar binding mode for NODD LXXLAP	78
4.4	Selectivity determinants from the conserved LXXLAP binding site	79
4.5	Modelling of tPHD2.NODD complexes	82
	4.5.1 Variant analyses towards validating tPHD2.NODD model 1	84
	4.5.2 Analyses of alternative tPHD2.NODD models 2 and 3	86
4.6	Selectivity determinants involving the tPHD2 C-terminal substrate binding site	89
4.7	Clinically observed mutations	92
4.8	Modelling of PHD3.CODD/NODD complexes	94
4.9	Summary	96
4.10	References	96

Chapter 5 – Studies with HIF substrate analogues to investigate differential regulation of NODD and CODD

5.1	Preferential hydroxylation of CODD over NODD	98
5.2	Significance of the presence of two ODDs in HIF	99
5.3	Evidence for coordinated regulation of CODD and NODD hydroxylation	101
5.3.1	Hydroxylation of 'NODD-CODD' hybrid peptide	102
5.3.2	Hydroxylation of 'CODD-NODD' hybrid peptide	102
5.4	Catalytic differences observed using CODD analogues and tPHD2 variants	104
5.5	HIF-2α clinical mutations	106
5.6	Summary	108
5.7	References	108

Chapter 6 – PHD2 active site chemistry and further studies towards investigating the oxygen sensing properties of PHD2

6.1	Active site, 2OG and metal binding	110
6.2	Water molecule at the active site iron – implications for O_2 binding	111
6.2.1	Iron-bound oxygen molecule in JMJD2A indicates an O_2 delivery system	111
6.2.2	Difficulty in replacing the metal bound water	112
6.2.3	Rearrangement possibility of 2OG 1-carboxylate or ferryl intermediate	116
6.3	Variable CODD conformations with different binding partners	117
6.4	Stereochemical endo-exo conversion of $Pro564_{CODD}$/$Pro402_{NODD}$	118
6.4.1	Stereospecificity of the PHD reaction	119
6.4.2	Studies with peptide analogues (endo/exo)	119
6.4.3	4*R*-$Hyp564_{CODD}$ is required for pVHL recognition	121
6.5	Implications for inhibitor design	123
6.6	Summary	124
6.7	References	124

Chapter 7 – Crystallisation and structure solutions of tPHD2 in complex with a new series of 2OG analogues

7.1	Inhibition data of the BB-287 series compounds		127
7.2	Crystallisation with the BB-287 series of compounds and structure solution		128
	7.2.1	Crystallisation of tPHD2 with BB-287 compounds	128
	7.2.2	Structure solution	129
7.3	Comparison of tPHD2.Fe(II).BB-287 with tPHD2.Fe(II).A/B crystal structures		131
	7.3.1	Fe(II) coordination	131
	7.3.2	2OG C5 carboxylate binding	132
	7.3.3	Hydrophobic interactions	132
	7.3.4	Interactions with symmetry related molecules	133
7.4	Attempts to crystallise tPHD2 in the presence of BB-287 and CODD substrate		134
7.5	Methylation of tPHD2 in attempts to improve crystal contacts		135
	7.5.1	Characterisation of the methylation products	135
	7.5.2	Crystallisation of methylated tPHD2	136
	7.5.3	Structure solution and model building	137
	7.5.4	Comparison with tPHD2 structures	139
7.6	Activity analyses of methylated tPHD2		139
7.7	Crystallisation of PHD2-FL (Full-length)		141
	7.7.1	Further attempts for PHD2-FL crystallisation	141
7.8	Summary		143
7.9	References		143

Chapter 8 – Investigations into the role of nitric oxide in the hypoxic response

8.1 Endogenous regulation and chemistry of NO/NO donors 144
8.1.1 Sources of NO 144
8.1.2 Reactions of NO 144

8.2 General role of S-nitrosylation 145

8.3 Direct effect of NO on PHD2 mediated HIF hydroxylation 145
8.3.1 NO diminishes PHD2 activity 146
8.3.2 NO alters Fe(II) EPR signals in PHD2 146

8.4 Crystallisation attempts – evidence for S-nitrosylation 148

8.5 Characterisation of S-nitrosylated tPHD2 in solution 151
8.5.1 Preparation of S-nitrosylated Apo-tPHD2 and native MS analyses 151
8.5.2 LC/MS analyses 152

8.6 Effects of S-nitrosylation on purified PHD2 155
8.6.1 Role of reducing agents on denitrosylation 155
8.6.2 Time dependence 156
8.6.3 Pre-incubation 157
8.6.4 Oxygen dependence 157
8.6.5 Effects of transition metals 158

8.7 Variation of conserved cysteines has no significant effect on PHD2 activity 160

8.8 Effect of Cu(I) and GSH on the cysteine variants 162

8.9 Treatment of HeLa cells with GSNO upregulates HIF 163
8.9.1 Design of experiments 163
8.9.2 Immunoblot analyses of cell inputs 163
8.9.3 Immunoprecipitation assays 164
8.9.4 Western blot analyses using recombinant tPHD2 164

8.10 Discussion and future directions 165

8.11 References 169

Chapter 9 – Cloning and characterisation of human PHD3

9.1	Difficulty in expression of PHD3		171
9.2	Optimisation of lysis and expression conditions		171
9.3	Construct design		172
	9.3.1	Truncation and deletion variants	172
	9.3.2	C-Terminal His_6-tagged PHD3	173
	9.3.3	Surface cysteine variants	173
	9.3.4	PHD3 purification and optimisation of buffer for protein storage	173
	9.3.5	Improvement of internal salt-bridges	175
9.4	Activity analyses of PHD3 and the variants		175
9.5	Co-expression trials		177
	9.5.1	Co-expression with HIF-1α	177
	9.5.2	Co-expression with chaperones	177
9.6	Crystallography		178
9.6.1	9.6.1	Crystallisation optimum solubility screening	178
9.6.2	9.6.2	Reductive lysine methylation	179
9.6.3	9.6.3	Thermal stability scanning	179
9.6.4	9.6.4	Constructing further variants by introducing a homotrimeric interface	179
9.7	Summary and future direction		180
9.8	References		181

Chapter 10 – Materials and Methods

10.1	General experimental methods		182
	10.1.1	Enzymes and chemicals	182
	10.1.2	Centrifugation	182
	10.1.3	pH measurement	183
10.2	Molecular biology		183
	10.2.1	Preparation of oligonucleotides	183
	10.2.2	Plasmid constructs	183
	10.2.3	Amplification of DNA sequence using PCR (sub-cloning)	183
	10.2.4	Site-directed mutagenesis	184
	10.2.5	Multi site-directed mutagenesis	185
	10.2.6	Dpn I digestion of the PCR products (mutagenesis)	186
	10.2.7	DNA restriction digest and clean-up of DNA (sub-cloning)	186
	10.2.8	Agarose gel electrophoresis	186
	10.2.9	DNA ligation	187
	10.2.10	Competent cells	188
	10.2.11	Transformation	188
	10.2.12	Plasmid purification	188
	10.2.13	DNA concentration determination	189
	10.2.14	DNA sequence verification	189
10.3	Sterile techniques		189
	10.3.1	Growth media	190
	10.3.2	Preparation of competent cells	190
10.4	Protein expression and purification		191
	10.4.1	Protein expression trial	191
	10.4.2	Scale-up of protein expression	191
	10.4.3	Preparation of cell extract	191
	10.4.4	Protein purification	192
	10.4.4.1	Purification of His_6-tagged proteins	192
	10.4.4.2	Ion exchange chromatography	193
	10.4.4.3	Cleavage of His_6-tag	193
	10.4.4.4	Size exclusion chromatography (SEC)	194
	10.4.4.5	Protein concentration determination	194
10.5	Protein characterisation		195
	10.5.1	SDS polyacrylamide gel electrophoresis (SDS-PAGE)	195
	10.5.2	Native PAGE	196
	10.5.3	Mass spectrometry	196
	10.5.3.1	Denaturing ESI-MS	196
	10.5.3.2	Non-denaturing ESI-MS	197
	10.5.3.3	MALDI-MS	197
	10.5.4	Elemental analysis by atomic absorption spectroscopy	197
	10.5.5	Circular dichroism (CD) spectroscopy	197
10.6	Biochemical assays		198
	10.6.1	2OG turnover assay	198
	10.6.2	Hydroxylation assay	199
10.7	Protein crystallisation		199
	10.7.1	Crystallisation trials	199
	10.7.2	In-house data collection and processing	200

10.8	General cloning and purification of tPHD2 variants (Chapters 2 to 8)		200
	10.8.1	Cloning of tPHD2 variants	200
	10.8.2	Purification	201
10.9	Experimental details for Chapter 2		201
	10.9.1	Analysis of protein stability by ThermoFluor	201
	10.9.2	Sedimentation velocity analytical ultracentrifugation	201
	10.9.3	Crystallisation details	201
	10.9.4	Data collection and processing	202
	10.9.5	Structure solution by molecular replacement and refinement	202
10.10	Experimental details for Chapters 3 and 4		202
	10.10.1	Crystallisation details	202
	10.10.2	Data collection and processing	203
	10.10.3	Structure solution by molecular replacement and refinement	203
	10.10.4	Modelling of tPHD2 and PHD3 complexes	203
	10.10.5	Molecular dynamics simulations studies	203
10.11	Experimental details for Chapter 5		204
	10.11.1	Synthesis of Peptides	204
10.12	Experimental details for Chapter 6		205
	10.12.1	Crystallisation details	205
	10.12.2	Data collection and processing	205
	10.12.3	Structure solution by molecular replacement and refinement	205
	10.12.4	[^{35}S]-pVHL capture assays	205
	10.12.5	Time-resolved fluorescence resonance energy transfer (TR-FRET)	206
	10.12.6	FT-ICR MS experiments	207
	10.12.7	Synthesis of labelled amino acids	207
10.13	Experimental details for Chapter 7		207
	10.13.1	TR-FRET assay	207
	10.13.2	Reductive lysine methylation	207
	10.13.3	Crystallisation details	207
	10.13.4	Data collection and processing	208
	10.13.5	Structure solution by molecular replacement and refinement	208
10.14	Experimental details for Chapter 8		208
	10.14.1	Nitric oxide solution preparation	208
	10.14.2	Preparation of GSNO and SNAP solution	208
	10.14.3	Preparation of S-nitrosylated tPHD2	208
	10.14.4	S-nitrosylation assay by the modified ‘Biotin-switch method’	209
	10.14.5	Tryptic digestion and LC-MS analyses	209
	10.14.6	MS/MS analyses	210
	10.14.7	Electron paramagnetic resonance (EPR) experimental conditions	210
	10.14.8	Crystallisation details	210
	10.14.9	Data collection and processing	211
	10.14.10	Structure solution by molecular replacement and refinement	211
	10.14.11	In vivo assays	211
	10.14.11.1	Antibodies	211
	10.14.11.2	Transfection	212
	10.14.11.3	Immunoprecipitation/purification of endogenous PHD2	212
	10.14.11.4	Western blotting/immunoblotting	212

10.15	Experimental details for Chapter 9		213
	10.15.1	Construction of plasmids	213
	10.15.2	Co-transformation for co-expression	213
	10.15.3	Purification of PHD3 wt and variants	214
	10.15.4	Non-denaturing MS and activity analyses	214
	10.15.5	CD measurements	214
	10.15.6	Thermal stability assay	214
	10.15.7	Crystallisation	215
10.16	References		215

Appendices I-XXXV

Appendix 1	Summary of the tPHD2 constructs/ proteins produced in the current study	I
Appendix 2	Compounds used in the thermal stability (Thermofluor) assays and/or MS studies	III
Appendix 3	Hydrogen bonds and hydrophobic interactions between tPHD2 and CODD as apparently observed in the tPHD2.CODD complex structure	XIV
Appendix 4	Summary of results for 2OG turnover and hydroxylation assays with NODD and CODD with tPHD2 and tPHD2 variants	XVI
Appendix 5	Hydrogen bonds and hydrophobic interactions as apparently observed in tPHD2.NODD models	XIX
Appendix 6	MS/MS data on tPHD2Me-I tryptic fragments showing peptide mass fingerprints, sequence coverage and MS/MS ion scores.	XX
Appendix 7	MS/MS data on tPHD2Me-II tryptic fragments showing peptide mass fingerprints, sequence coverage and MS/MS ion scores.	XXIII
Appendix 8	MS/MS data on tPHD2-SNO tryptic fragments showing peptide mass fingerprints, sequence coverage and MS/MS ion scores.	XXVII
Appendix 9	List of lysis buffer conditions used for PHD3 expression trials	XXXI
Appendix 10	Summary of the PHD3 constructs/ proteins produced in the current study	XXXII
Appendix 11	List of primers used for generation of tPHD2 variants	XXXIII
Appendix 12	List of primers used for generation of PHD3 variants	XXXV

Abbreviations

(v/v)	Volume to volume ratio
(w/v)	Weight to volume ratio
2OG	2-Oxoglutarate
2× TY	2× Tryptone/Yeast extract
3D	3-Dimensional
aa	Amino acids
A	Absorbance
ANS	Anthocyanidin synthase
APS	Ammonium persulfate
ATP	Adenosine triphosfate
ARD	Ankyrin repeat domain
ARNT	Aryl hydrocarbon receptor nuclear translocator
bp	Base pair
BSA	Bovine serum albumin
bHLH	Basic helix-loop-helix
BLAST	Basic local alignment search tool
C	Celsius
CAD	C-Terminal transactivation domain
CAS	Clavaminic acid synthase
CBP	CREB (cyclic AMP-response element binding protein) binding protein
CD	Circular dichroism
cDNA	Complementary DNA
CH-1	Cysteine/histidine rich
CITED 2	CBP/p300 interacting transactivator with ED-rich tail 2
CODD	C-Terminal oxygen dependent degradation domain
CPH	Collagen prolyl-4-hydroxylase
CoA	Coenzyme A
CoIP	Co-immunoprecipitation
Cul2	Cullin 2
CV	Column volume
DAOCS	Deacetoxycephalosporin C synthase
ddH_2O	Doubly deionised water
DMSO	Dimethylsulfoxide
dNTP	Deoxyribonucleotide triphosfate
DSBH	Double-stranded β helix
dsDNA	Double stranded DNA
E/S	Enzyme to substrate ratio
EDTA	Ethylenediaminetetraacetic acid
EPO	Erythropoietin
ESI-MS	Electrospray-ionisation mass spectrometry
FIH	Factor inhibiting HIF
FPLC	Fast protein liquid chromatography
GPCR	G-Protein coupled receptor
GST	Glutathione-S-transferase
HEPES	(*N*-[2-Hydroxyethyl]piperazine-*N*′-[2-ethanesulfonic acid])
HIF	Hypoxia-inducible factor
HPLC	High performance liquid chromatography
HRE	Hypoxia-response elements
Hsp90	Heat-shock protein 90
hTRF	Homogeneous time-resolved fluorescence
Hyp	4*R*-hydroxyproline
IKK	IκB kinase
IκB	Inhibitor of NF-κB
IP	Immunoprecipitation
IPAS	Inhibitory Per-ARNT-Sim
IPNS	Isopenicillin-N-synthase
IPTG	Isopropyl-β-*D*-thiogalactoside
IVTT	In vitro transcription/translation
kb	Kilo base
kDa	Kilo Dalton
K_M	Michaelis-Menten constant
LB	Luria broth
LC-MS	Liquid chromatography coupled with mass spectrometry
MALDI	Matrix-assisted laser desorption/ ionisation
MDS	Molecular dynamics simulations
MS	Mass spectrometry
MW	Molecular weight
MWCO	Molecular weight cut-off
MYND	Myeloid translocation protein 8, Nervy, and DEAF1
m/z	Mass to charge ratio
NAD	N-Terminal transactivation domain
NADPH	Nicotinamide adenine dinucleotide phosfate

NF-κB	Nuclear factor enhancing kappa light-chain gene expression in B cells
NLS	Nuclear localisation signal
NOG	*N*-oxalylglycine
OD_{600}	Optical density at 600 nm
ODD	Oxygen-dependent degradation domain
PAGE	Polyacrylamide gel electrophoresis
PAS	Per-ARNT-Sim
PER	Period circadian protein
PCR	Polymerase chain reaction
PDB	Protein data bank
PHD	Prolyl hydroxylase domain containing protein
tPHD2	PHD2 (181-426)
PHD3-CH	PHD3 with a C-terminal hexa-His tag
PHD3-NH	PHD3 with an N-terminal hexa-His tag
Pfu	*Pyrococcus furiosus*
ppm	Parts per million
PTM	Post-translational modifications
pVHL	von Hippel-Lindau tumour suppressor protein
QMM	Quantum Molecular Mechanics
Q-TOF	Quadrupole time-of-flight
RACK1	Receptor of activated protein kinase C
Rbx1	Ring-box 1
RCC	Renal cell carcinoma
ROS	Reactive oxygen species
rpm	Revolutions per minute
RNAi	RNA interference
s	Seconds
SCF	Skp-1-Cdc53/Cullin-F-box
SDS	Sodium dodecylsulfate
SIM	Single-minded protein
siRNA	Small interfering RNA
SPR	Surface plasmon resonance
ssDNA	Single stranded DNA
TAD	Transactivation domain
TAE	Tris.HCl/ Acetic acid (glacial)/ EDTA
TEMED	*N*,*N*,*N*′,*N*′-Tetramethylethylene-diamine
TF	Transcription factor
T_m	Melting temperature
Tris	Tris(hydroxymethylethyl)amino-methane
TauD	Taurine/α-ketoglutarate dioxygenase
UV	Ultraviolet
V	Volts
VCB	pVHL-elongin B-elongin C
VCBCR	pVHL-elongin B-elongin C-Cul2-Rbx1
VEGF	Vascular endothelial growth factor
V_{max}	Maximal velocity
wt	Wildtype
z	Charge state

Standard abbreviations for nucleotides, amino acids and SI-units/prefixes are used throughout.

List of contributors

The following experiments were carried out in collaboration with, or with assistance from, the listed personnel from the Chemistry Research Laboratory, Department of Chemistry, University of Oxford, except where stated.

Experiments	**Contributor**
Non-denaturing ESI-MS (Chapters 2-9)	J. Mecinović and Dr B. M. R. Lienard
Data collection of tPHD2.Fe(II).B.CODD$_{Hyp564}$ and tPHD2.Fe(II).B (NO exposed) crystals (Chapters 2 and 8)	Dr M. A. McDonough
Molecular dynamics simulations (MDS) studies (Chapters 3-6)	Dr. C. Domene[1]
TR-FRET assay (Chapters 6 and 7)	Dr L. A. McNeill
Experiments with human cells (Chapter 8)	Dr M. E. Cockman[2]

[1]Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford; [2]Centre for Cellular and Molecular Physiology, University of Oxford, Oxford.

The following reagents were prepared or assays carried out by the listed personnel from the Chemistry Research Laboratory, Department of Chemistry, University of Oxford.

Experiments	**Contributor**
Cloning of PHD2$_{181-426}$ in pET-28a(+) (Chapters 2-8) and PHD3 in pET-28a(+) (Chapter 9)	Dr K. S. Hewitson
Cloning of PHD2$_{181-426}$/pET-24a(+) and PHD2$_{181-406}$/pET-28a(+) (Chapters 2-8)	Dr E. G. Flashman
Cloning of PHD2$_{181-402}$, tPHD2-PHD3 and tPHD2-PHD1 (Chapter 4)	Dr E. G. Flashman
Kinetic assays with tPHD2, PHD2$_{181-402}$, tPHD2-PHD3 and tPHD2-PHD1 (Chapter 4)	E. A. L. Bagg and Dr E. G. Flashman
Peptide synthesis (Chapter 5)	C. Loenarz
Syntheses of peptides and amino acids with stereospecific (endo/exo) proline analogues (Chapter 6)	C. Loenarz
Studies with substrate analogues – deuterium exchange (FT-ICR MS analyses), hydroxylation assays and pVHL capture (Chapter 6)	C. Loenarz
VCB production and Eu^{3+} labelling (Chapter 6)	C. Loenarz
EPR experiments (Chapter 8)	Dr E. G. Flashman

Chapter 1 – Introduction*

1.1 Physiological background

The mechanism by which animals adapt to conditions of limiting oxygen (hypoxia) has been a long-standing physiological problem.[2] More than a century ago physiologists observed that limited oxygen availability, such as occurring at high altitude, resulted in an increase in red blood cells.[2] Evidence that the effect was mediated via an effector molecule, erythropoietin (EPO), came from work on animals in which serum from anaemic rabbits was shown to elicit erythropoiesis after injection into normal rabbits.[3] These fundamental observations led ultimately to the purification of EPO and cloning of the gene encoding for EPO. Recombinant EPO is now very widely used for the treatment of anaemia.[4] Although EPO is normally produced continuously by the body in order to ensure renewal of red blood cells, its rate of production is increased when oxygen becomes limiting, including at altitude, in tumours, and with CO poisoning.[5]

1.2 Hypoxia-Inducible Factor (HIF): Discovery of a master regulator of oxygen homoeostasis

The expression of eukaryotic genes including EPO is regulated by multiple transcription factors (TFs) some of which may control protein synthesis in response to external stimuli. All (TFs) are modular in structure and contain a DNA binding domain and a trans-activating domain (*trans*-acting: involving a different molecule; *cis*-acting: involving the same molecule).[6] The trans-activating domains enable binding of the TF to other proteins involved in transcription or its regulation. Some TFs, or TF complexes, contain sensing domains that regulate the activity of the TF, sometimes in response to binding of a small molecule.

Studies with transgenic mice demonstrated that there were DNA sequences on both the 3' and 5' sides of the EPO gene that regulated its expression, some in a hypoxia dependent manner.[7] Liver cell lines that expressed EPO in an oxygen dependent manner were then identified and used to narrow down sequences that were responsible for the hypoxic regulation of EPO.[8,9] A specific sequence (to the 3'-side) that enhanced EPO expression under hypoxic conditions (up to a 1000 fold) termed the hypoxia response element (HRE), was identified along with other sequences that are important for hypoxic regulation.[10]

*In a modified form, part of this introduction has been published as a review.[1] Some of the results described in this introduction section were reported during the course of my work.

The core conserved region within the HRE is a pentanucleotide sequence 5'-A/(G)CGTG-3'; this sequence is repeated for the regulation of some hypoxically regulated genes. DNA affinity chromatography led to the characterisation of hypoxia-inducible factor (HIF), as an α,β-heterodimeric TF that binds to the core DNA sequence 5'-A/(G)CGTG-3' within the cis-acting HRE.[11] The oxygen dependency of HIF activity relies on post-translational modifications to the HIF-α subunit that occur in a central oxygen dependent degradation domain (ODD)[12] and a C-terminal transactivation domain (CAD).[13] HIF-β, which is constitutively expressed in nuclei, is not regulated by oxygen.[11] An important observation was that hypoxic regulation by the HIF system is not limited to EPO but involves an array of human genes implying a complex and coordinated response to chronic hypoxia.[14-17]

1.3 HIF target genes and the hypoxic response

The transcriptional control over the hypoxic genes elicited by HIF is possibly the most well-defined molecular mechanism for multicellular O_2 homeostasis. Cells employ a considerable amount of energy for continuous synthesis and, interestingly at the same time, degradation of HIF under conditions of normal oxygen supply. HIF has been termed a master regulator for the hypoxic response.[17] More than seventy HIF regulated genes have been directly identified by demonstration that HIF binds to an associated HRE (Fig. 1.1). Microarray analyses imply that the number will be substantially larger, possibly up to 200 or more genes are up or down regulated by hypoxia or hypoxia mimics.[18] Many of them have been identified on the basis of loss of hypoxia-driven gene expression in HIF-α null cells or siRNA treated cells. It is uncertain how many of these genes have functional HREs. The large number of genes that are HIF regulated is consistent with the complex nature of the hypoxic response.

Functions associated with maximising oxygen delivery, preventing damage or optimising metabolism for hypoxia can be assigned to some of the HIF target genes.[17] Significantly, expression of genes encoding for glycolytic enzymes are HIF regulated as well as those involved in increasing red blood cell production (EPO) and angiogenesis (blood vessel formation from a pre-existing bed of vessels). There is also evidence that the tricarboxylic acid (TCA) cycle is actively downregulated in hypoxic conditions.[17] A number of HIF related genes encode for Fe(II) related proteins including some of the HIF regulating hydroxylases (discussed later), transferrin, transferrin receptor (both involved in iron transport), haem oxygenase, inducible NO synthase, tyrosine hydroxylase, and procollagen

prolyl-hydroxylase (CPH) α-subunit (like the HIF hydroxylases CPH is a 2-oxoglutarate (2OG) oxygenase).[16,17]

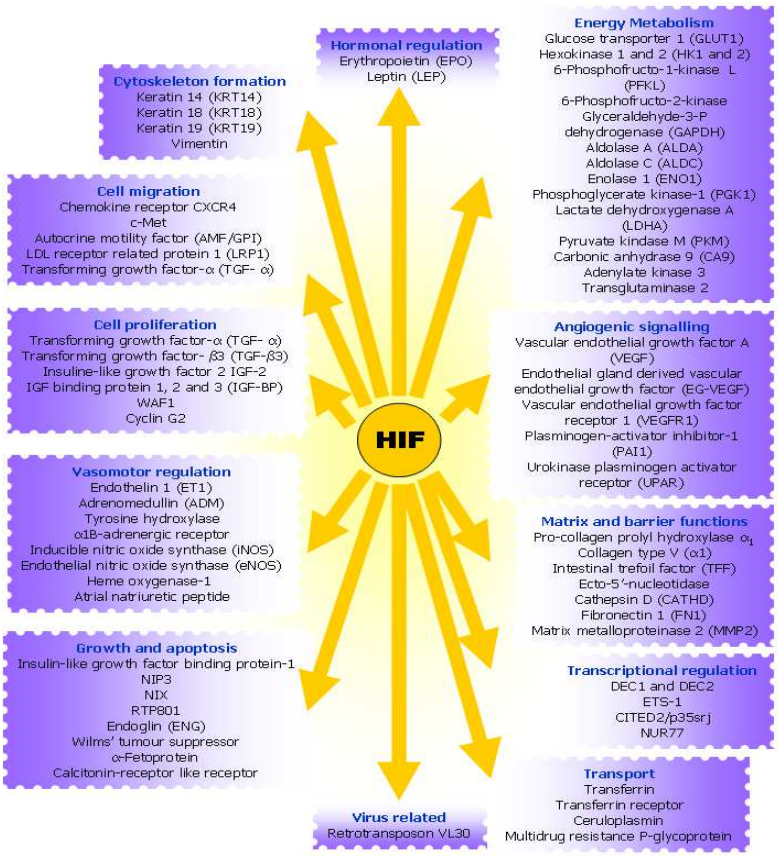


Fig. 1.1 HIF target genes. The complete number of genes transcriptionally activated by HIF may exceed 200 as demonstrated by microarray analyses.[18] Abbreviations: c-Met, tyrosine-kinase product of the met proto-oncogene; WAF1, protein implicated in p53 transcriptional regulation; Nip3, novel immunogenic protein 3 and NIX, Bnip3-like protein X are members of the Bcl2 family of cell-death factors; DEC1 and DEC2, differentiated embryo chondrocyte 1 and 2, ETS is a DNA-binding domain that defines a family of transcription factors; p35srj, 35-kDa serine/glycine-rich junction; NUR77, orphan nuclear receptor; CITED2, CBP/p300-interacting transactivator with ED-rich C-terminal domain 2. Figure updated from reference 16,17.

1.4 Domain architecture and isoforms of HIF

1.4.1 HIF-α isoforms

Three isoforms of human HIF-α have been reported to date: HIF-1α, 2α and 3α.[16] In other organisms the number of HIF-α isoforms may vary. Whilst HIF-1α and HIF-2α, upon dimerisation with HIF-β, apparently have overlapping but distinct specificities regarding target gene activation and physiological induction, HIF-3α is proposed to act as a negative regulator of the hypoxic response.[19] An alternative splice variant of the HIF-3α gene was found to dimerise with HIF-β and inhibit transcriptional activity by sequestering HIF-β from a functional HIF complex.[20] HIF-3α is hypoxically inducible[21] and may therefore provide a potential negative feedback mechanism for HIF-α regulation.[22]

1.4.2 Domain architecture of HIF subunits

Both the HIF-α and -β subunits are basic helix-loop-helix (bHLH) and PAS domain

containing TF proteins (Fig. 1.2).[11] PAS is an acronym referring to the eukaryotic proteins PER (period circadian protein), ARNT (aryl-hydrocarbon receptor nuclear translocator), and SIM (single-minded protein). HIF-1α and -2α share a common domain architecture: the N-terminal bHLH-PAS domains for DNA binding/dimerisation, the central oxygen dependent degradation domains for stability and the C-terminal transactivation domains for transcriptional activation. HIF-3α, having bHLH and PAS domains, can bind to DNA and/or dimerise with HIF-β; yet it does not contain a transactivation domain and therefore, plays a negative role in the hypoxic response system (Fig. 1.2).

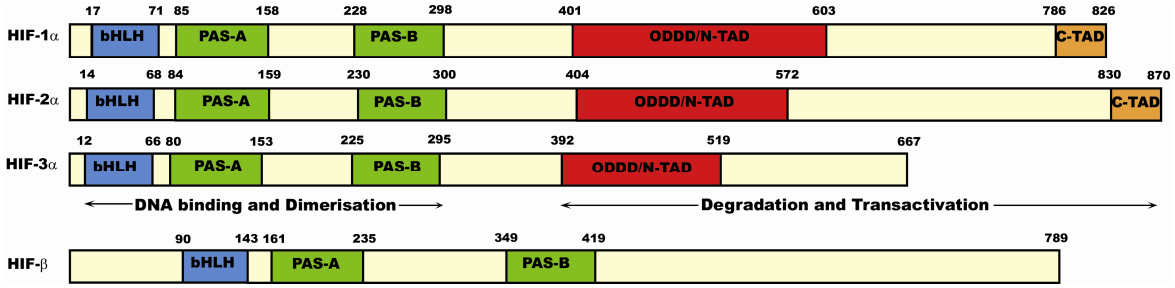


Fig. 1.2 Domain architectures of human HIF-1α, -2α, -3α and HIF-β. The bHLH domain followed by two internal repeats of the PAS domains (A and B) are involved in DNA binding and dimerisation. The central portion of HIF-α consists of the oxygen-dependent degradation domain (ODD). The C-terminal part comprises the transactivation domains (N/C-TAD) that are separated by an inhibitory domain and the nuclear localisation signal.[23] The oxygen-sensitive ODD partially overlaps with N-TAD and is involved in stabilisation/degradation process. The transactivation domains recruit co-activators such as CBP/p300 to initiate transcription. Nuclear localisation sequences (not shown) target HIF-α to the nucleus.

1.5 DNA binding and dimerisation of HIF

All bHLH transcriptional regulators must dimerise to form functional DNA-protein complexes.[24] The basic sequences in bHLH directly bind to DNA and the helix-loop-helix domains enable dimerisation. The bHLH proteins are broadly classified into three subfamilies: those having only the bHLH domain and those with an additional dimerisation domain, either a leucine zipper (bHLH-LZ, e.g. MAX, USF)[25-28] or a PAS domain (e.g. HIF-α/β). Both the bHLH and bHLH-LZ proteins recognise the classic E-box core enhancer sequence (CANNTG), whilst bHLH-PAS proteins often tend to recognise slight variation in the consensus sequence.[24]

Detailed information on the interaction of the HIF heterodimer and the cognate (A/(G)CGTG) sequence is not yet available; studies on the interactions between other bHLH proteins and response elements reveal that a conserved His-Glu-Arg triad binds to the major groove of the DNA (Fig. 1.3).[29] HIF-β likely binds in a similar way though the

triad appears not to be conserved in HIF-α. Methylation at the cytosine-5-position of the conserved (A/G)CGTG sequence reduces HIF binding to the HRE providing one possible mechanism for cell-type specificity.[30] Co-immunoprecipitation (CoIP) of HIF-1α and HIF-β from nuclear extracts indicated that these subunits can dimerise even in the absence of DNA. Deletion of N-terminal basic domain abrogated DNA binding without disrupting the dimeric form.[10]

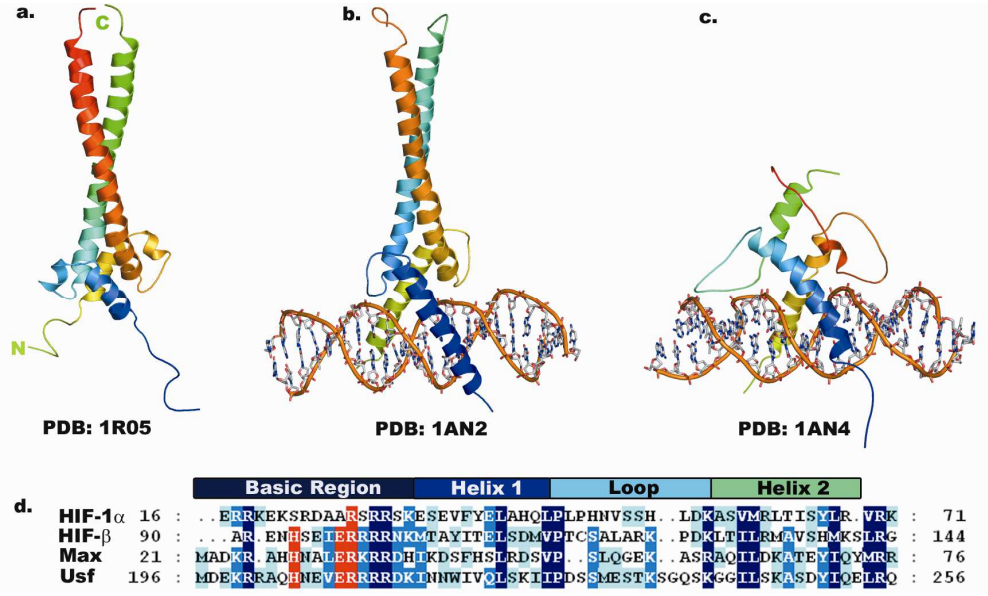


Fig. 1.3 Cartoon showing binding of the homodimeric Max (MYC-associated protein X) and USF (upstream stimulatory factor) transcription factors (bHLH-LZ type) to the recognition sequence (CACGTG) suggesting how a functional HIF complex could bind to a hypoxic response element. (a) View from an NMR structure of the bHLH of Max (PDB: 1RO5),[27] (b) view from a crystal structure of Max in complex with target gene promoter E-Box (5'-CACGTG-3') DNA sequences (PDB: 1AN2),[27] (c) view from a crystal structure of USF bHLH in complex with DNA (PDB: 1AN4).[26]

PAS domains in bHLH-PAS proteins contain two internal repeats, each of ~130 amino acids (PAS-A and PAS-B) (Fig. 1.2). They act as a secondary dimerisation interface (bHLH being the primary interface) that recognises the binding partners. Partner selection is highly specific for PAS proteins and determines target gene recognition.[24] An NMR structure of the HIF-2α PAS-B domain (residues 240-350) reveals a central β-sheet region flanked by α-helices that are mainly involved in HIF-α/β dimerisation (Fig. 1.4a).[25] Current evidence suggests that the HIF-2α PAS-B interacts specifically with the HIF-β PAS domain and not PAS domains from other proteins. HIF-1α/2α PAS-B have sequence identity of ~75% (Fig. 1.4); homology modelling and mutational studies predict similar folding of PAS-B in the two HIF-α isoforms.[25]

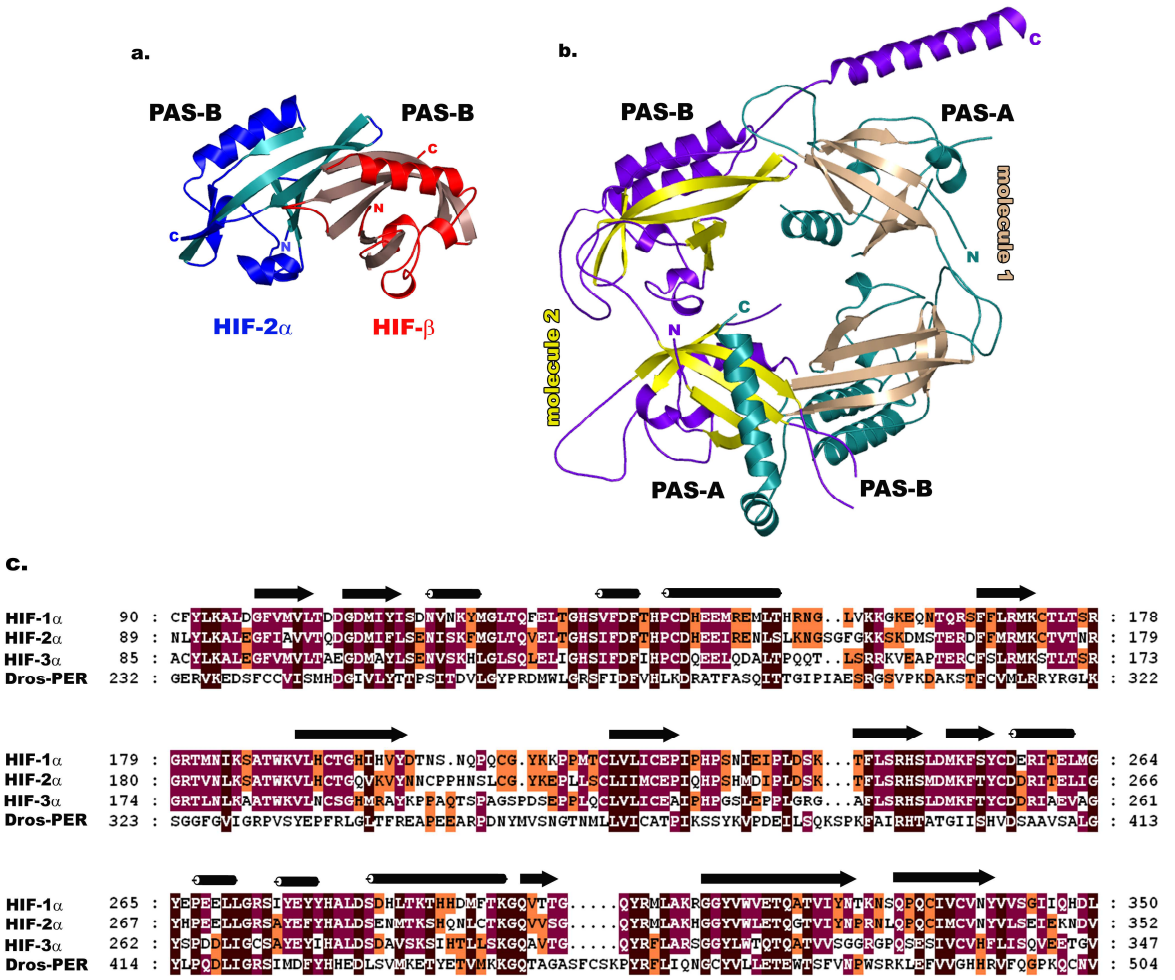


Fig. 1.4 PAS domains of HIF-α and HIF-β. (a) NMR structure of the heterodimeric complex of HIF-2α PAS-B (residues 240–350) and HIF-β PAS-B (residues 356–470) (PDB: 2A24),[25] (b) Views from crystal structure of the PAS domain of *Drosophila* clock protein PERIOD (PDB: 1WA9),[28] (c) Sequence alignment of PAS domains of HIF-α isoforms and *Drosophila* PERIOD circadian protein. Secondary structures for PAS-A were assigned on the basis of PAS-A of the *Drosophila* PERIOD and those of PAS-B were assigned on the basis of PAS-B of *Drosophila* PERIOD/HIF-2α structures.

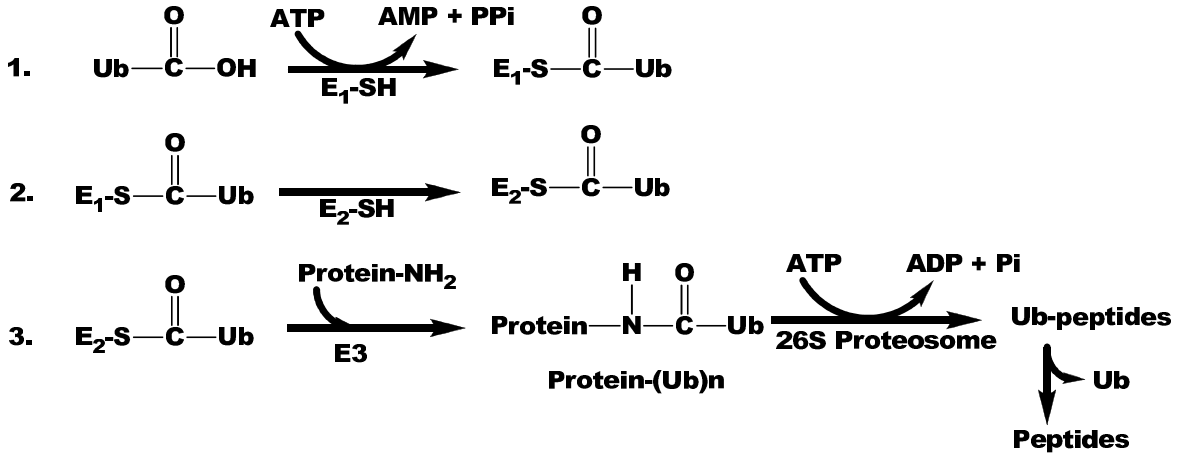
HIF-β is also known as the aryl hydrocarbon receptor nuclear translocator (ARNT) and is a ubiquitous nuclear protein that is involved in the transcription of genes other than those involved in the hypoxic response.[11] To date ARNT has been reported to form both homo/heterodimers with the aryl hydrocarbon receptor (AhR), HIF-α and SIM. ARNT/AhR senses the dioxin pollutants whilst ARNT/SIM mediates control over neural development processes.[24]

1.6 Regulation of HIF activity by oxygen – basic components of hypoxic system

1.6.1 Prolyl hydroxylation of the central oxygen dependent degradation domains (ODDs)

In normoxic cells, only very low levels of HIF-α, if any, can be detected by Western blotting (antibody staining). Human HIF-1α and -2α (~100 kDa) both contain a large central regulatory region of ~200 residues that regulates their oxygen dependent degradation. Within this ODD there are two sites for oxygen dependent prolyl hydroxylation (a Pro402 in the N-terminal ODD, NODD; Pro564 in the C-terminal ODD of human HIF-1α, CODD).[31,32] Both of these sites contain an LXXLAP motif that is highly conserved in other metazoans. Hydroxylation at the *trans*-4-position of either $Pro564_{CODD}$ or $Pro402_{NODD}$ is sufficient to target HIF-α to an ubiquitin ligase (see below) that results in attachment of ubiquitin protein molecules.[33] The oxygen in the hydroxyproline was shown to arise from molecular oxygen providing a direct link between oxygen availability and HIF-1α stability.[34] As for other proteins labelled with ubiquitin, ubiquitinylation targets HIF-α to the proteasome where it is hydrolysed into short peptides. Under hypoxic conditions, the rate of prolyl hydroxylation of the NODD and CODD is slowed, most of the HIF-α avoids hydroxylation and can therefore translocate to the nucleus (if not already there) and dimerise with HIF-β to enable the hypoxic response.

The conjugation of a ubiquitin moiety (76 amino acid protein, MW 8.5 kDa) to a target protein substrate involves three steps[35]: (1) ATP-dependent activation of ubiquitin that results in a thio-ester bond formation between the C-terminal carboxylate of ubiquitin and the -SH (cysteine) of a ubiquitin activating enzyme (E1), followed by (2) transfer of ubiquitin from the E1 ubiquitin complex to the active site cysteine of a ubiquitin-conjugating enzyme (E2) via a trans(thio)esterification reaction (Scheme 1.1). Finally, in step (3), ubiquitin is transferred from the E2 enzyme to a target lysine in the target protein by a ubiquitin-protein ligase (E3). E3 ligase complexes act as the recognition component of the system and are capable of interacting with both E2 and the target proteins. The ligation reaction is generally catalysed by HECT (Homologous to the E6-AP Carboxyl Terminus) or RING (Really Interesting New Gene) domains of the E3 ligase complexes.[35]



Scheme 1.1 Enzymatic reactions involved in the ubiquitination system. Ub = Ubiquitin, PPi = inorganic diphosphate, Pi = inorganic phosphate. See text for discussion.

1.6.2 Structure of the HIF targeting E3 ubiquitin ligase and implications

The recognition component of the HIF-1α ubiquitin ligase is the von Hippel-Lindau protein (pVHL).[33] pVHL forms a complex with elongins B and C (VCB)[36], Cul2[37] and the Ring-H2 finger protein Rbx1 (VCBCR complex).[38] The VCBCR complex shares sequence and structural similarities with the well-characterised SCF (Skp-1-Cdc53/Cullin-F-box)[39] class of E3 ubiquitin ligase complexes. A 2.7 Å resolution crystal structure of the ternary complex of VCB (PDB ID: 1VCB) shows two interfaces, one (mostly helical and hence the α-domain) between pVHL and elongin C and the other between elongins C and B.[36] Evidence that pVHL binds to Cul2 via elongin C in native cells comes from biochemical data where pVHL-Cul2 complexes were confirmed by CoIP and copurifications.[37]

Prior to the observation of HIF-α prolyl hydroxylation, the oxygen dependent nature of complex formation between HIF-α and VCB had been an important step in identifying the oxygen sensing mechanism.[33] In vitro studies have shown that prolyl-4-hydroxylated CODD peptides bind ~1000 fold more tightly to the VCB complex than non-hydroxylated CODD.[40] X-ray diffraction analyses have revealed that the post-translationally introduced alcohol is positioned to bind via hydrogen-bonds to the side chains of Ser111 and His115 of the von Hippel-Lindau protein in VCB (Fig. 1.5a).[40,41] In the absence of the VCB template, NMR studies reveal this region of HIF-α as disordered. The hydroxylated CODD proline residue has an exo-conformation as observed in the collagen triple-helix fold due to the operation of the gauche stereoelectronic effect.[42]

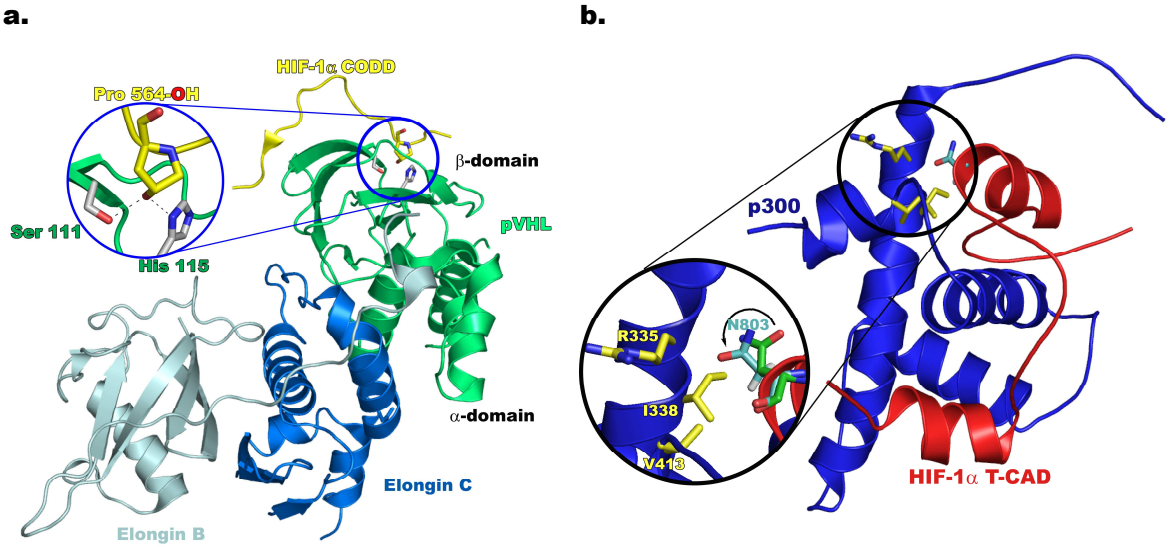


Fig. 1.5 Views from structures of key oxygen dependent components of the HIF system. Inset diagrams provide closer views. (a) The HIF-VCB interaction;[40,41] VCB shows exquisite discrimination between hydroxylated and non-hydroxylated proline containing ODDs. This is mediated by two hydrogen bonds formed between the alcohol of the hydroxylated proline and two residues of pVHL (Ser111 and His115). The α-domain of pVHL binds to elongin C and Cul2 whereas the β-domain interacts with hydroxylated HIF-α (PDB: 1LQB) (b) The HIF-p300 interaction;[43,44] upon binding to p300, HIF-1α CAD forms two α-helices which are connected by an intervening loop (PDB: 1L3E). Hydroxylation at the pro-*S* β-carbon of Asn803 may provide a steric clash with Ile-338 of p300.[45] Note the conformational changes of Asn803 (inset) when bound to p300 (cyan sticks) and FIH (green sticks).

1.6.3 Asparaginyl hydroxylation of transactivation domains (CAD)

An oxygen dependent post-translational hydroxylation also occurs within the CAD region of HIF-α.[46] The oxygen-dependent hydroxylation of a single asparagine residue within the CAD was first identified in mouse HIF-2α at position 851, corresponding to Asn803 in human HIF-1α.[46] Later hydroxylation at the pro-*S* β-position of Asn803 was observed in human HIF-1α CAD.[45] Hydroxylation at this single asparagine residue was found to significantly decrease binding of HIF-α to the cysteine/histidine rich (CH-1) domain of the transcriptional co-activator complex CBP/p300 disabling HIF mediated transcription. In the HIF-α-p300/CBP complex,[43,44] HIF-1α Asn803 contributes to an α-helix; hydroxylation is proposed to disrupt the hydrophobic binding interactions between this α-helix of HIF-1α and CH-1 (Fig. 1.5b).

In summary, two types of oxygen dependent post-translational hydroxylation of HIF-α have been identified. Prolyl hydroxylation establishes a protein-protein interaction that signals for HIF degradation. Asparaginyl hydroxylation disrupts a protein-protein interaction disabling HIF mediated transcription. Both the HIF-α prolyl and asparaginyl hydroxylations are catalysed by oxygenases that employ Fe(II) as a cofactor and 2OG and

oxygen as cosubstrates (discussed further below).

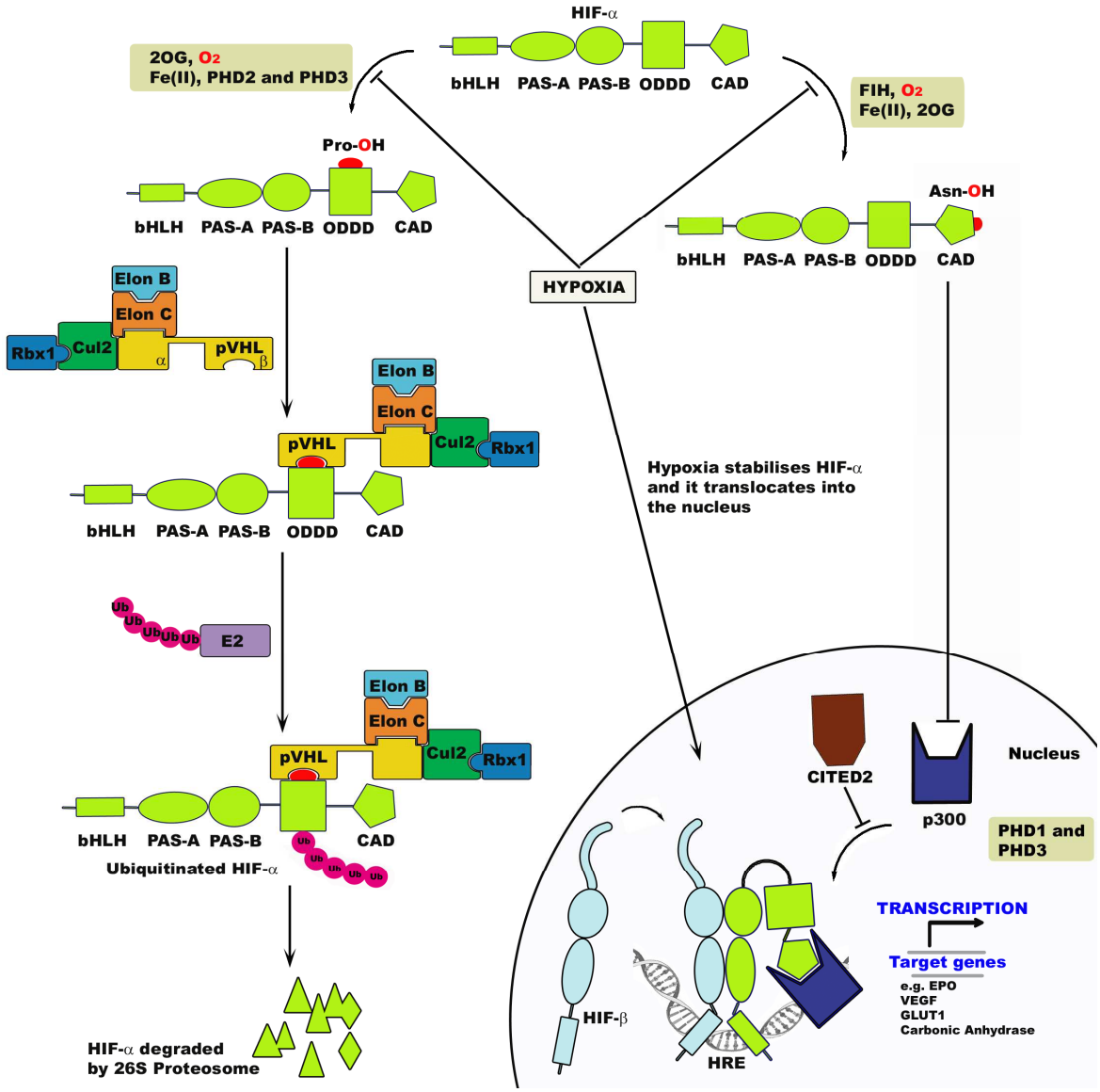


Fig. 1.6 Simplified overview of the HIF system including reactions of the HIF hydroxylases. Under normoxic conditions, the PHDs and FIH downregulate and inactivate HIF-α subunits. PHD1 is exclusively localised in the nucleus, PHD2 and FIH are mainly found in the cytoplasm whereas PHD3 is rather homogenously distributed in both the nucleus and cytoplasm.[47] In hypoxia, the activity of the HIF hydroxylases is reduced enabling HIF-α to escape PHD and FIH mediated control. HIF-α translocates to the nucleus, dimerises with HIF-β via bHLH-PAS domains leading to a transcriptionally active complex. The functional heterodimeric HIF then binds to hypoxia response elements (HREs) in the regulatory regions of target genes using bHLH of both subunits and also recruits transcriptional machinery (p300/CBP) using its transactivation domain. CITED2 which binds p300 with higher affinity than HIF-1α, negatively regulates HIF-1 transcriptional activity[48]. Abbreviations: PAS = PER ARNT SIM (see text), bHLH = basic-helix-1-loop-helix-2, ODD = oxygen dependent degradation domain, CAD = transactivation domain, Elon B = elongin B, Elon C = elongin C, Rbx1 = ring-box 1, Cul2 = cullin 2, Ub = ubiquitin, E2 = E2 ligase, CITED2 = CBP/p300-interacting transactivator with ED-rich carboxy-terminal domain 2.

1.7 Regulation of HIF stability via the oxygen independent pathway

Regulation of HIF-α stability is also mediated by at least one oxygen-independent pathway. Recent studies have revealed that HIF-1α is degraded in a PHD/VHL-independent mechanism via interaction of the receptor of activated protein kinase C (RACK1) with the PAS-A domain of HIF-1α.[49] It is proposed that RACK1 forms a ubiquitin ligase complex similar to that formed by pVHL; RACK1 is able to recruit elongin C plus other subunits which can target HIF-α for proteasomal degradation via a parallel O_2-independent pathway.[50] RACK1 mediated control over HIF-1α is limited by the chaperone, heat-shock protein 90 (Hsp90). In the cytoplasm, HIF-1α can bind to Hsp90, an association that stabilises HIF-α. However, upon displacement of HIF-α from Hsp90 by Hsp90 inhibitors, RACK1 is able to bind to HIF-α and recruit ubiquitin ligase machinery and potentiate the degradation of HIF-α.[51,52]

1.8 Factors affecting HIF levels and activity

Cells become adapted to hypoxia over extended periods and lose the ability to stabilise HIF-α, only inducing HIF-α when made severely hypoxic. Two of the prolyl hydroxylases (PHDs) are themselves hypoxically regulated providing a negative feedback mechanism that may enable the more rapid degradation of HIF-α upon reoxygenation of hypoxic cells.[53,54] Regulation of PHD levels, and as a consequence HIF-α levels, provides a mechanism for adjusting the set-point of the HIF system. This is essential if the hypoxic response is to operate in the very different cellular environments encountered in humans.

Various other feedback and modulating processes are likely to regulate HIF levels via regulation of the activity of the HIF hydroxylases.[16,55] Factors to be considered, for which there is already evidence, include: (i) the intrinsic catalytic activity of the HIF hydroxylases under different conditions, (ii) the forms of HIF and the HIF hydroxylases that are present in a given cell type; modifications may occur via the production of different transcripts or post-translational modifications[56] such as phosphorylation, acetylation, *S*-nitrosylation, (iii) the rates of production and degradation of HIF, the HIF hydroxylases and associated proteins involved in transcription and transport, (iv) the availability of the Fe(II) cofactor for the HIF hydroxylases, (v) the redox state of the cell; in addition to exogenous iron supply,[57] Fe(II) availability is regulated by the redox state of the cell which affects the Fe(II)/Fe(III) ratio; as for CPH activity, PHD and (with some

substrates) FIH activity is stimulated by ascorbate, (vi) the number and availability of HRE sequences for HIF binding; availability of HREs may be regulated by changes to chromatin structure (e.g. cytosine methylation) or protein (e.g. histone modification), and (vii) small molecules (e.g. succinate, fumarate, NO) (Fig. 1.7).

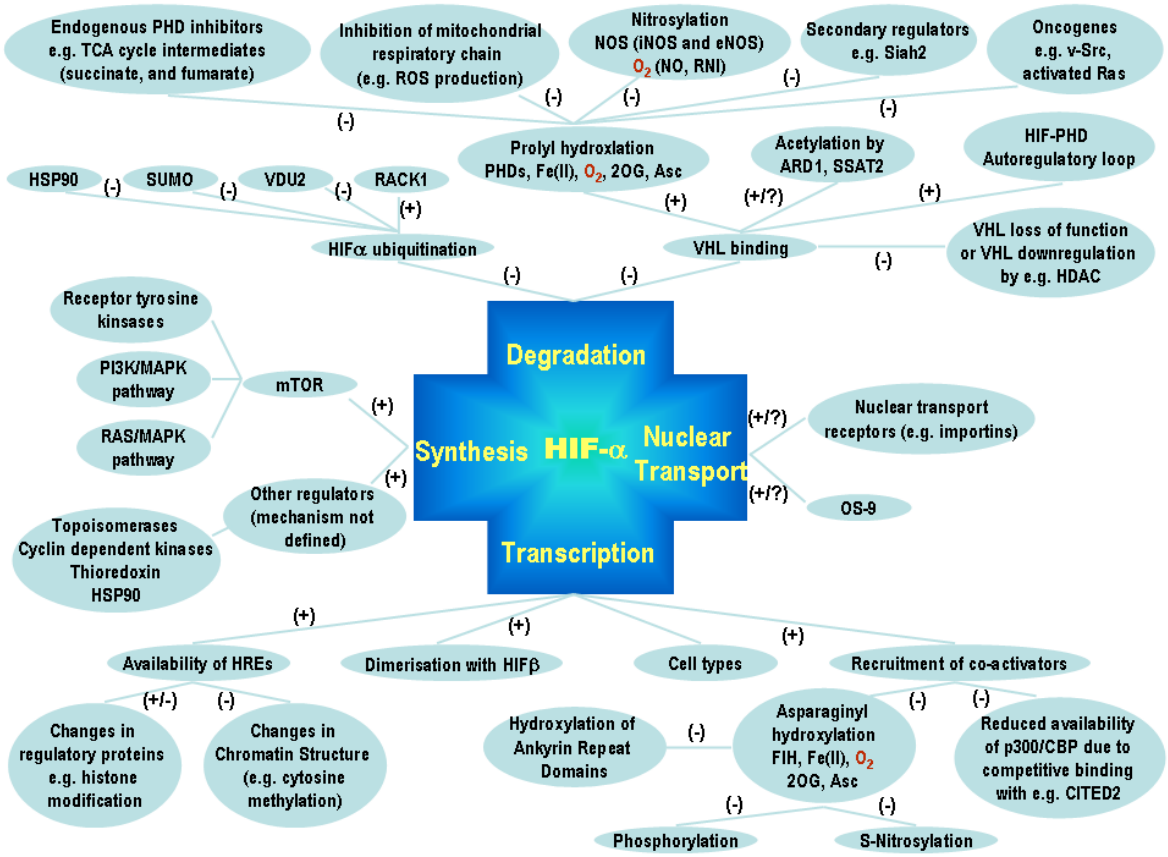


Fig. 1.7 Major factors that (may) affect the HIF system. Abbreviations: ROS, reactive oxygen species; iNOS an eNOS, inducible and endothelial nitric oxide synthase; RNI, reactive nitrogen intermediates; Siah2, seven in absentia homolog 2; ARD1, arrest-defective protein 1; SSAT2, spermidine/spermine N1-acetyltransferase 2; HSP90, heat shock protein 90; RACK1, receptor of activated protein kinase C 1; pVHL, von Hippel-Lindau tumour suppressor protein, VDU2, pVHL-interacting deubiquitinating enzyme 2; HDAC, histone deacetylases; PI3K, phosphoinositide-3-kinase; MAPK, mitogen-activated protein kinase; CITED2, CBP/p300-interacting transactivator with ED-rich carboxy-terminal domain 2; SUMO, Small ubiquitin-related modifier.

From a medicinal perspective the possibility of endogenous and exogenous regulation of HIF activity by small molecules is attractive. Succinate dehydrogenase[58] and fumarate hydratase[59] mutants causing increases in fumarate and succinate levels are characteristic of some tumour cells. Because succinate and fumarate are inhibitors of PHDs, these mutations may contribute to activation of the hypoxic response through stabilising HIF. The interfaces between NO and the HIF system appear complex. Under normoxic conditions NO upregulates HIF (possibly via hydroxylase inhibition involving direct interaction) whereas under hypoxic conditions it downregulates HIF.[60]

The links between the HIF system and reactive species (ROS) are also a subject of ongoing discussion.[61,62] Mitochondrial electron transport chain inhibitors inhibit hypoxically mediated HIF stabilisation. It has been proposed that mitochondria produce a HIF hydroxylase inhibitor;[60] alternatively it could be that the set-point for oxygen dependency of the HIF hydroxylases has changed, e.g. altered local oxygen availability or by one or more potential regulatory factors described above. Mitochondrial respiration accounts for >90% of oxygen consumption in humans. It is proposed that under hypoxia NO can reduce mitochondrial O_2 consumption via inhibition of cytochrome C oxidase (redox centre IV in the mitochondrial electron-transport chain) making more O_2 available for oxygenases such as the PHDs.[63] The transcription factor JunD enables upregulation of antioxidants. The deletion/loss of JunD leads to HIF stabilisation via a process proposed to be mediated either directly or indirectly by an increase in ROS.[64] One possibility is that the ROS inhibit the HIF hydroxylases by irreversible active site oxidation.

Ongoing studies are identifying other factors that have an impact on the HIF system. HIF could be stabilised by promoting degradation of two of the three PHD isoforms (PHD1/PHD3) via the ubiquitin ligase Siah1a/Siah2.[65] Transforming growth factor (TGF) β has been reported to downregulate PHD2 mRNA and protein levels stabilising HIF levels via Smad signalling pathway.[66] G-Protein-coupled receptor (GPCR) agonists, including angiotensin II and thrombin, have been shown to induce HIF-1 in vascular smooth muscle cells.[67] Inhibition of PHDs by NADPH oxidase mediated ROS production and subsequent decrease in ascorbate levels has been recently proposed to be a regulatory mechanism of HIF-1 induction under angiotensin II treatment.[68] Other studies showed that angiotensin II might suppress SM-20 (rat homologue of PHD3) mRNA expression in a rat pheochromocytoma cell line and thereby promote HIF stabilisation.

Although the available evidence indicates that the regulation of HIF levels is an important and probably dominant mechanism in the adaptation of cells to hypoxia, it is also important to appreciate that there are other transcription factors regulated by oxygen, including the NFκb TF (NF = nuclear factor) inflammatory response system.[69] One connection between the HIF and NFκβ systems is provided by the observation that PHD1 interacts with IκB kinases that regulate IκBα levels that in turn inhibits NFκB.[70]

1.9 The HIF hydroxylases

Three human prolyl hydroxylases (PHD 1-3/ EGLN 1-3) catalyse hydroxylation of specific proline(s) within the ODDs.[71,72] The K_M (Michaelis-Menten constant) value of PHD2 for molecular oxygen (see below) supports the role of PHDs as cellular oxygen sensors. This provides a direct link between the regulation of HIF and molecular oxygen.[73] Although they share high sequence conservation within their C-terminal catalytic domains, each PHD isoform shows different kinetic properties and primary cellular localisation.[47] The N-terminal domains are more divergent and have poorly characterised functions. Sequence analyses indicate that the non-conserved N-terminus of PHD2 contains a Myeloid translocation protein 8, Nervy, and DEAF1 (MYND)-type zinc finger domain; deletion of MYND finger was reported to increase the activity of both recombinant and in vitro translated PHD2.[74] The three PHDs are selective for the HIF-α CODD over the NODD; PHD3 is the most selective, accepting only CODD in in vitro studies.[75,76] PHD2 is thought to be the most important of the PHDs in catalysing HIF-α hydroxylation under normoxic conditions in healthy tissue.[77] A recent report has demonstrated that another human prolyl hydroxylase can carry out HIF-α hydroxylation, but its relevance to the hypoxic response is not yet established.[78]

As for other 2OG oxygenases the available data implies that catalysis by FIH and PHD2 proceeds via a general mechanism involving the sequential binding of 2OG (cosubstrate), substrate and finally, oxygen.[79] Kinetic and biophysical analyses have revealed that PHD2 has unusually tight binding constants for Fe(II) and 2OG, compared to FIH and most other 2OG oxygenases studied.[80] The reported K_M of HIF-α fragment substrates for PHD2 and FIH are in the range of <0.1 – 10 μM.[75,79,81] The apparent K_M values of FIH and the PHDs for molecular oxygen were ~90 and 250 μM, respectively, using 19-residue peptides.[73,79,82] Studies using longer HIF-α substrates (35-mer peptide/ recombinant proteins of M.W. >20 kDa) lowered the K_M (oxygen) for PHD to ~100 μM.[79,81] Given that physiological oxygen concentrations are reported to be around 4–40 μM[83] (below the estimated K_M) and the fact that the PHDs are present within cells in a femtomolar concentration range[84], this implies that the hydroxylases are poised to respond to physiologically relevant changes in intracellular oxygen concentrations, with any decrease in oxygen tension likely to reduce the catalytic rate of the enzymes. Overall these properties appear to make the HIF hydroxylases suitable for a role as oxygen sensors.

The relatively higher affinity of FIH for molecular O_2 suggests that PHDs would be inactivated before FIH as a function of severity of hypoxia.[82] This means that even if a small amount of HIF escapes from PHD mediated degradation, FIH would still prevent transcriptional activation of HIF via hydroxylation of the CAD region. In cell-based assays using Hep3B and U-2OS cells, FIH activity was found to be detectable even at severe hypoxia (0.2% O_2 concentration) whereas PHDs were found essentially inactive under the same conditions.[85] Other groups have shown that FIH (but not the PHDs) sensitivity towards decreasing oxygen levels can vary between different cell types.[86] However, given that the enzyme and substrate concentrations are orders of magnitude lower and that there are pericellular/intracellular oxygen gradients in vivo, the results from the in vitro studies may not always correlate to in vivo assessments of enzyme activity.[85]

Crystal structures for the catalytic domain of PHD2 (tPHD2)[87] and FIH[88-90] have revealed that their overall folds follow that of other 2OG oxygenases (Fig. 1.8). Both contain the canonical double-stranded β-helix (DSBH) fold that has been observed in all 2OG oxygenases identified to date. They also both contain the highly, but not fully, conserved 2His-1carboxylate HXD...H facial triad of iron binding residues (Fig. 1.9). The structure of tPHD2 revealed a narrow entrance to the active site consistent with the high affinity for Fe(II) and 2OG. Kinetic and structural analyses suggest that the NODD/CODD selectivity of the three PHDs is partly determined by two non-conserved regions at their C-termini or a mobile loop region that may fold to enclose the active site on substrate binding.[75,76]

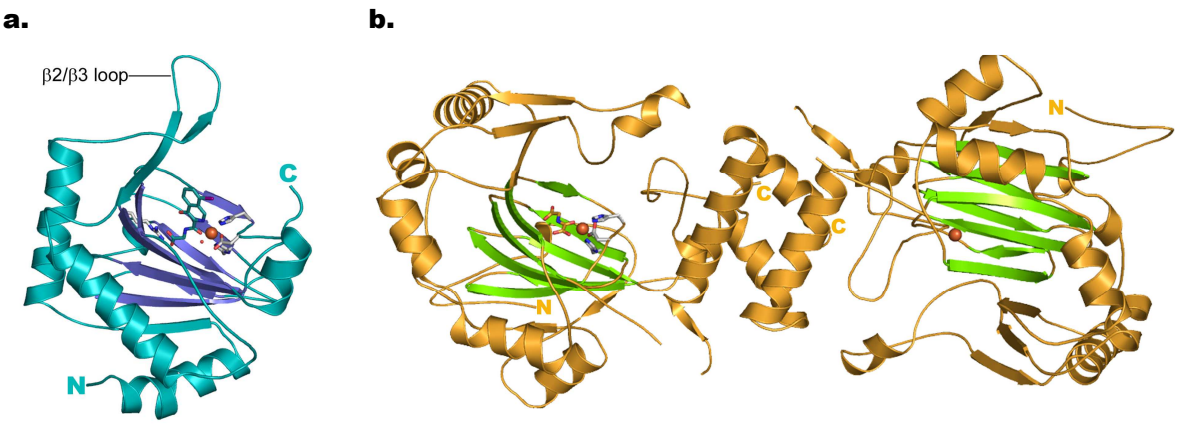
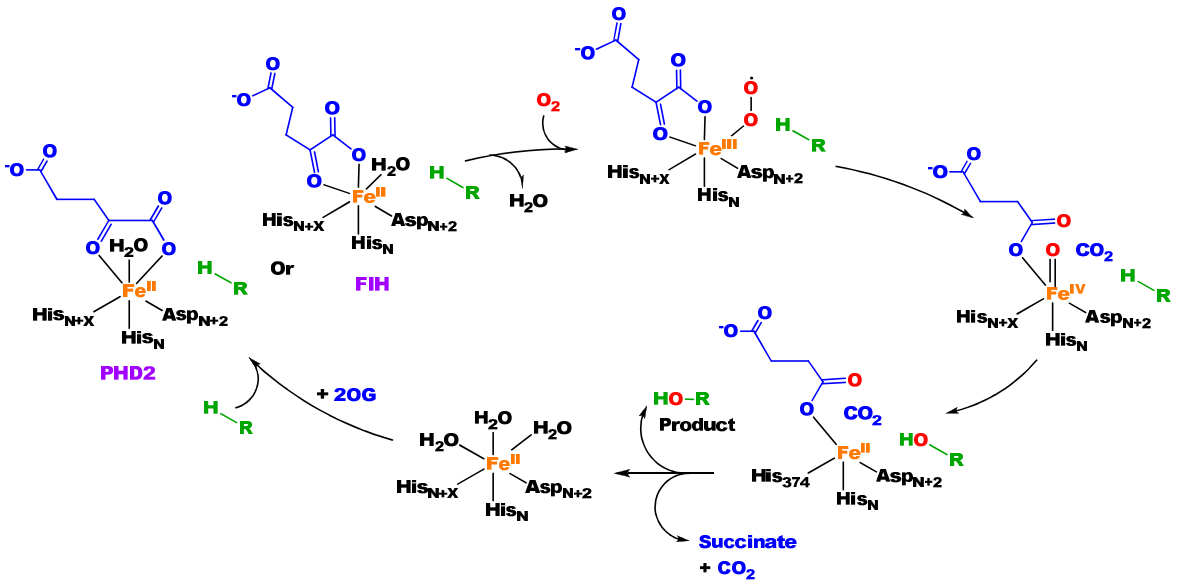


Fig 1.8. Ribbon representations of the crystal structures of HIF hydroxylases. Overall structures of (a) monomeric tPHD2 (cyan, PDB: 2G19); (b) dimeric FIH (bright orange, PDB: 1H2N) showing the conserved DSBH fold in dark blue (tPHD2) and lemon green (FIH), the iron and 2OG binding sites (residues highlighted in white sticks) and orange spheres for iron centre.

FIH-CAD complex structures have been reported and reveal that residues 795–803 of CAD bind in groove at the FIH active site and adopt an extended conformation interacting with FIH through hydrogen bonds and hydrophobic interactions (Fig. 1.9).[89] CAD Asn803 and Ala804 form an inverse γ-turn, stabilised by a hydrogen bond between the backbone CO of Val802 and NH of Ala804. Asn803 of the CAD is enclosed and its 3-pro-*S* Asn hydrogen that must be oxidised projects directly towards the iron.[45]

Structural and functional studies on HIF hydroxylases coupled with Quantum Molecular Mechanics (QMM)[91] analyses on relevant non-haem iron-containing oxygenases[92-94] including 2OG oxygenases has allowed mechanistic predictions to be made. In outline 2OG oxygenases, including the HIF hydroxylases, couple the two electron oxidation of 2OG into succinate and carbon dioxide to the two electron oxidation of a substrate. The reaction is proposed to proceed via formation of a highly reactive Fe(IV)=O intermediate (Fig. 1.9). Evidence for this high-spin Fe(IV)=O intermediate comes from Mössbauer and electron paramagnetic resonance spectroscopy carried out on 2OG oxygenases, *viz.*, taurine/α-ketoglutarate dioxygenase (TauD)[95] and a viral collagen prolyl-4-hydroxylase (CPH).[96] To date, in humans the only type of substrate oxidation identified has been hydroxylation. In microorganisms and plants they catalyse a range of oxidative transformations including desaturations and oxidative ring closures in secondary metabolism of, e.g. antibiotics and flavonoids.[97,98]

a.



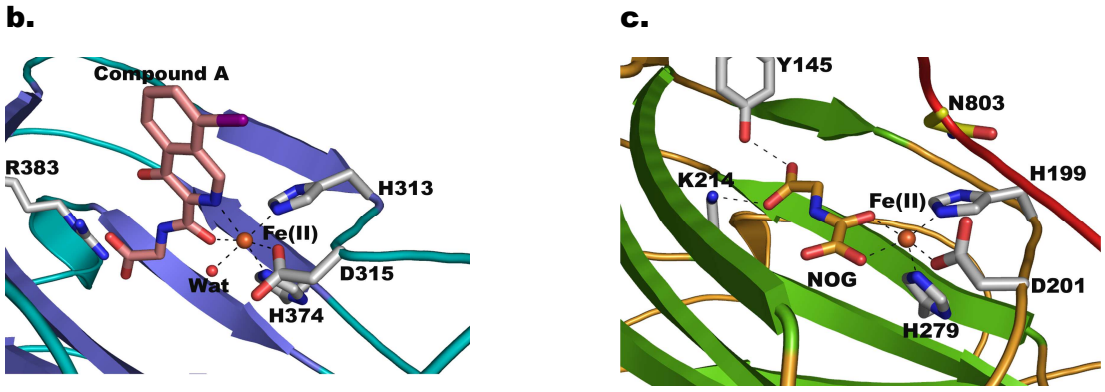


Fig. 1.9 Outline mechanism and active sites of the HIF hydroxylases. (a) Outline mechanism for the HIF hydroxylases. Fe(II) is octahedrally coordinated at the enzyme active site by the conserved HXD..H triad and the remaining coordination sites are occupied by water molecules. Crystal structures of FIH and other 2OG-dependent oxygenases in presence/absence of substrate molecule suggest that the cosubstrate 2OG and then the substrate (HIF-α) bind sequentially to the active site. This is followed by binding of molecular oxygen which is proposed to replace the remaining water molecule from the iron centre, leading to decarboxylation of 2OG and a highly reactive Fe(IV)=O intermediate.[95,96] The latter is responsible for hydroxylating the peptide substrate. In case of FIH, binding of substrate displaces the water molecule from Fe(II). Views from (c) the PHD and (d) the FIH active sites.

1.10 Functional differences between PHD isoforms

1.10.1 Variable expression of PHDs

The mRNAs for the three human PHDs are widely expressed in various tissues and cell lines albeit at different levels.[53,99-101] Of the three isoforms, PHD2 is the most abundant across all tissues and studied cell lines, whilst PHD3 is detected only in cardiac and neural tissue under normoxia and PHD1 has the highest expression in the placenta and is the only PHD found in testis.[99-101] PHD2 and PHD3 are themselves hypoxically regulated suggesting an auto-regulatory mechanism, with PHD3 showing the most striking induction under HIF control.[53,102-105] The role of PHD3 may therefore be more important in hypoxia.[53] Recently, an oxygen-sensing role for PHD1 was reported in hypoxia tolerance; conditional knockdown of PHD1 was shown to induce hypoxic tolerance rapidly.[106] Loss of PHD1 was associated with the oxygen consumption in skeletal muscles via upregulation of a Pparα pathway.[106]

1.10.2 Differential subcellular localisation of PHDs

Whilst PHD1 has been exclusively detected in the nucleus, PHD2 is mainly cytoplasmic, although during tumour progression, it can translocate into the nucleus.[47,107] PHD3 is more homogeneously distributed in both cytoplasm and nucleus.[47] The reason(s) for such restricted distribution or the mechanism behind the nuclear transportation of PHDs are poorly understood although a consensus sequence KK(R)XK(R) for monopartite nuclear

localisation signals (NLS)[108] is present in variable non-conserved regions of all three PHDs.

1.10.3 PHD2 is the key regulator in normoxia

The PHDs showed specific activities in an order of PHD2 = PHD3 > PHD1 using 19-mer peptides from the three HIF isoforms.[84] A separate study using the recombinant PHDs and GST-HIF-1$\alpha_{(531-575)}$ protein substrate suggested that PHD2 has the highest specific activity towards the primary hydroxylation site (Pro564).[77] To address the precise role of different PHDs in vivo, short interfering RNAs were used to specifically silence each of the PHDs.[103] Loss of PHD2 was found to be sufficient to stabilise and activate HIF-1α in normoxic cells whereas loss of PHD1 and PHD3 had no significant effect under these conditions. PHD2 was therefore proposed to be the dominant human PHD in normoxia.[103] The siRNA data also revealed that the PHDs manifest a further degree of selectivity between HIF-1α and HIF-2α isoforms, with PHD2 having more effect on HIF-1α and PHD3 on HIF-2α.[53]

1.11 Other potential substrates for the HIF hydroxylases

One important question is whether the HIF hydroxylases have alternative substrates. Some of the PHDs have been reported to interact with other proteins including OS-9,[109,110] ING4,[111] RNA polymerase II[112] and IκB kinase-β,[70] an enzyme involved in regulating the inflammatory response. As yet, prolyl hydroxylation of these potential targets has not been demonstrated by mass spectrometry. OS-9, has been identified using yeast two-hybrid assays and immunoprecipitation analyses and is suggested to have a role in trafficking of HIF proteins between the nucleus and the ER.[109,110] Protein interaction screens led to the identification of ankyrin repeat domain (ARD) proteins as FIH substrates.[113] The ARD is a 33-residue loop-helix-turn-helix motif that is amongst the most common of structural motifs in human cells.[114] To date, FIH has been reported to catalyse the hydroxylation of ARD proteins from the NFκB and Notch ARD protein families (intracellular Notch domain and some NFκB proteins are TFs, involved in the inflammatory response and development).[115] Recently, the ARD of a SOCS (suppressor of cytokine signalling) box protein (ASB4) has also been reported as an FIH substrate.[116] The link with Notch proteins is notable because HIF-1α is recruited to Notch responsive promoters under hypoxic conditions (independently of HIF-β). Collectively, these results suggest that ARD

hydroxylation by FIH may be common; they imply that hydroxylation of cytoplasmic proteins may be much more ubiquitous than previously believed.

1.12 Genetic diseases associated with HIF

HIF regulated genes such as EPO (for the treatment of anaemia) and VEGF (targeted for tumour treatment) are already the subject of therapies.[117,118] Because of the pivotal role of HIF in oxygen homeostasis, it is attracting interest from a therapeutic perspective. Ischemic diseases, e.g. coronary heart disease, result from a restricted blood supply leading to a lack of oxygen. For ischemic disease management, upregulation of HIF activity is desirable. In contrast, tumours are normally hypoxic leading to upregulation of HIF; thus, for tumour therapy it is desirable to starve the tumours of oxygen/other nutrients by down-regulating HIF activity. There is evidence that renal cell carcinomas (RCCs) selectively overexpress the gene encoding for HIF-2α rather than that for HIF-1α; inhibition of HIF-2α was shown to be sufficient to suppress tumour growth.[119] Mice in which HIF-2α was deleted were reported to display anaemia suggesting that HIF-2α is the predominant HIF isoform regulating EPO in adults. Other examples of HIF-α isoform selective gene regulation have also been reported including HIF-1α selective regulation of carbonic anhydrase IX.[120] However, further work is required to achieve a molecular understanding of how selectivity is achieved. Beyond the scope of this introduction, there are also links between the HIF system and other signalling systems, such as the inflammatory response, that may influence selectivity.

Two inherited mutations to PHD2 have been linked with familial erythrocytosis (an abnormally high red blood count).[121,122] One of these, $Pro317_{PHD2}$, is located two residues away from one of the Fe(II) binding residues ($Asp315_{PHD2}$) and is located close to the apparently unusually narrow entrance to the PHD2 active site. The other, $Arg371_{PHD2}$, is also close to another of the triad of metal-binding residues, $His374_{PHD2}$. Both mutations have been shown to result in a decrease in catalytic activity, but they may also affect the in vivo stability of PHD2.

von Hippel-Lindau (VHL) disease is an autosomal dominant inherited cancer syndrome that is characterised by RCCs and tumours. VHL syndrome occurs at a frequency of ~1 in 35,000 humans and is caused by mutations in the VHL gene. The lack of wildtype VHL results in accumulation of HIF in cells and subsequent overexpression of target genes.

Abnormal increases in red blood cell production, or polycythaemia, endemic in the Chuvash population (in Russia) also arise from VHL mutations. However, these VHL mutations do not result in tumours but cause venous abnormalities and a tendency towards thrombosis. Why different VHL mutations lead to different phenotypes is unknown.

1.13 Possibilities for therapeutic intervention

Basic research on oxygen sensing and the HIF system has uncovered opportunities for therapeutic intervention. Therapeutic possibilities involving the HIF system may be divided into those that upregulate or those that downregulate HIF target genes/proteins. Knowledge of the HRE sequences can be used to selectively target the expression of exogenous genes to hypoxic tissues, either for imaging or for delivering therapeutic genes. However, the effects of HIF-silencing are not straightforward, and do not necessarily correspond with a decrease in tumour size or vasculature. The potential for use of elements of the HIF system for employing either the HRE or ODD components has been elegantly demonstrated by work on mice.[123] Use of polyamides that bind to specific regulatory sequences is a possible route to achieving regulation of a subset of HIF target genes.[123]

Screens to identify small molecules that down-regulate HIF-α have been carried out (for reviews see references [117,118]). These have led to molecules that target HIF associated molecules, though how selective these targets are to HIF and the hypoxia response is unclear. Inhibitors of the mammalian target of rapamycin (mTOR) result in lowered HIF levels, and are in clinical development as anti-cancer agents.[117,118] Curcumin, which is present in some spices, is reported to promote HIF-α degradation.[124] Inhibitors of heat shock protein 90 (HSP90) that stabilises HIF-1α in PHD/pVHL independent manner, also reduce HIF-1α levels.[118] Promoting the activity of the HIF hydroxylases in cancer cells may reduce the availability of HIF; treatment with 2OG has shown decreased VEGF production in Hep3B cells;[125] it is unclear if 2OG will be limiting in vivo though there is a possibility that in some tumours succinate and fumarate compete with 2OG binding to the PHDs.[58] Ensuring PHD activity is not limited by ascorbate, or the actual in vivo reducing agent - if there is one, may also help to suppress HIF levels.

Other potential small molecule targets to block HIF function include blocking HIF-α/β dimerisation and preventing HIF from binding to its cotranscriptional activators, specifically p300. The natural product chetomin, a disulphide containing diketopiperazine

(epidithiodiketopiperazine), is a potent blocker of binding of HIF-α to p300, both with isolated proteins and in cells, though the mechanism of action remains unknown.[126]

In the case of upregulating HIF for the treatment of ischemic disease or anaemia (by upregulating EPO), two clear possibilities are blocking either the HIF-VHL interaction or the inhibiting activity of the HIF hydroxylases. Studies with peptides and the 2OG analogue, *N*-oxalylglycine (first used as a procollagen prolyl hydroxylase inhibitor),[127] verified these approaches.[72] Further evidence for targeting the HIF hydroxylases as a therapeutic approach comes from the use of Co(II) as a hypoxia mimic - at least in part the ability of Co(II) to act in this way results from inhibition of the HIF hydroxylases. Indeed Co(II) has been used as a clinical treatment for anaemia prior to the advent of EPO, but has toxic effects probably unrelated to its HIF-based mechanism of action.[128] Recent efforts have focused on identifying other organic inhibitors of the HIF hydroxylases.[129,130] Most, if not all, of those reported act as 2OG competitors and chelate to the active site Fe(II) in a bidentate manner. It will be desirable to develop inhibitors that are selective for the HIF hydroxylases over other 2OG oxygenases. To what extent selectivity in terms of HIF target genes is desirable is uncertain and may depend on the therapeutic goal. Indeed for some applications induction of the hypoxic response in a manner as close as possible to the natural response may be desirable; if so the point of oxygen sensing (i.e. HIF hydroxylases) may be a preferred target.

1.14 Objectives of the current research

During the course of my work, crystal structures had been reported for the catalytic domain of PHD2 (tPHD2) in complex with Fe(II) and two 2OG competitive bicyclic inhibitors, A (PDB ID 2G19, 2G1M) and B (PDB ID 2HBT, 2HBU).[87] The structures revealed the stereotypical iron binding motif and DSBH fold of the 2OG oxygenases but did not address the central question of how PHDs bind HIF-α. The structures did not provide insights into the reasons for the unusual properties of PHD2, nor into the factors governing PHD selectivity, and in particular why PHD3 is highly selective for CODD over NODD. To gain a molecular understanding of how the PHDs act as oxygen sensors in normal and disease states, information on how they bind HIF-α is required. This thesis describes work aimed at providing a basis for understanding the molecular features leading to PHD selectivity.

1.15 References

1. Chowdhury, R., Hardy, A. & Schofield, C.J. The human oxygen sensing machinery and its manipulation. *Chem Soc Rev* **37**, 1308-19 (2008).
2. Ratcliffe, P.J. From erythropoietin to oxygen: hypoxia-inducible factor hydroxylases and the hypoxia signal pathway. *Blood Purif* **20**, 445-50 (2002).
3. Erslev, A. Humoral regulation of red cell production. *Blood* **8**, 349-57 (1953).
4. Arcasoy, M.O. The non-haematopoietic biological effects of erythropoietin. *Br J Haematol* **141**, 14-31 (2008).
5. Jelkmann, W. Erythropoietin after a century of research: younger than ever. *Eur J Haematol* **78**, 183-205 (2007).
6. Lee, T.I. & Young, R.A. Transcription of eukaryotic protein-coding genes. *Annu Rev Genet* **34**, 77-137 (2000).
7. Semenza, G.L., Nejfelt, M.K., Chi, S.M. & Antonarakis, S.E. Hypoxia-inducible nuclear factors bind to an enhancer element located 3' to the human erythropoietin gene. *Proc Natl Acad Sci U S A* **88**, 5680-4 (1991).
8. Maxwell, P.H., Pugh, C.W. & Ratcliffe, P.J. Inducible operation of the erythropoietin 3' enhancer in multiple cell lines: evidence for a widespread oxygen-sensing mechanism. *Proc Natl Acad Sci U S A* **90**, 2423-7 (1993).
9. Semenza, G.L. & Wang, G.L. A nuclear factor induced by hypoxia via de novo protein synthesis binds to the human erythropoietin gene enhancer at a site required for transcriptional activation. *Mol Cell Biol* **12**, 5447-54 (1992).
10. Jiang, B.H., Rue, E., Wang, G.L., Roe, R. & Semenza, G.L. Dimerization, DNA binding, and transactivation properties of hypoxia-inducible factor 1. *J Biol Chem* **271**, 17771-8 (1996).
11. Wang, G.L., Jiang, B.H., Rue, E.A. & Semenza, G.L. Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. *Proc Natl Acad Sci U S A* **92**, 5510-4 (1995).
12. Huang, L.E., Gu, J., Schau, M. & Bunn, H.F. Regulation of hypoxia-inducible factor 1alpha is mediated by an O2-dependent degradation domain via the ubiquitin-proteasome pathway. *Proc Natl Acad Sci U S A* **95**, 7987-92 (1998).
13. Jiang, B.H., Zheng, J.Z., Leung, S.W., Roe, R. & Semenza, G.L. Transactivation and inhibitory domains of hypoxia-inducible factor 1alpha. Modulation of transcriptional activity by oxygen tension. *J Biol Chem* **272**, 19253-60 (1997).
14. Kaelin, W.G. Proline hydroxylation and gene expression. *Annu Rev Biochem* **74**, 115-28 (2005).
15. Ozer, A. & Bruick, R.K. Non-heme dioxygenases: cellular sensors and regulators jelly rolled into one? *Nat Chem Biol* **3**, 144-53 (2007).
16. Schofield, C.J. & Ratcliffe, P.J. Oxygen sensing by HIF hydroxylases. *Nat Rev Mol Cell Biol* **5**, 343-54 (2004).
17. Semenza, G.L. Oxygen-dependent regulation of mitochondrial respiration by hypoxia-inducible factor 1. *Biochem J* **405**, 1-9 (2007).
18. Elvidge, G.P., Glenny, L., Appelhoff, R.J., Ratcliffe, P.J., Ragoussis, J. & Gleadle, J.M. Concordant regulation of gene expression by hypoxia and 2-oxoglutarate-dependent dioxygenase inhibition: the role of HIF-1alpha, HIF-2alpha, and other pathways. *J Biol Chem* **281**, 15215-26 (2006).
19. Gu, Y.Z., Moran, S.M., Hogenesch, J.B., Wartman, L. & Bradfield, C.A. Molecular characterization and chromosomal localization of a third alpha-class hypoxia inducible factor subunit, HIF3alpha. *Gene Expr* **7**, 205-13 (1998).
20. Makino, Y., Cao, R., Svensson, K., Bertilsson, G., Asman, M., Tanaka, H., Cao, Y., Berkenstam, A. & Poellinger, L. Inhibitory PAS domain protein is a negative regulator of hypoxia-inducible gene expression. *Nature* **414**, 550-4 (2001).
21. Makino, Y., Kanopka, A., Wilson, W.J., Tanaka, H. & Poellinger, L. Inhibitory PAS domain protein (IPAS) is a hypoxia-inducible splicing variant of the hypoxia-inducible factor-3alpha locus. *J Biol Chem* **277**, 32405-8 (2002).
22. Lisy, K. & Peet, D.J. Turn me on: regulating HIF transcriptional activity. *Cell Death Differ* (2008).
23. Hara, S., Kobayashi, C. & Imura, N. Nuclear localization of hypoxia-inducible factor-2alpha in bovine arterial endothelial cells. *Mol Cell Biol Res Commun* **2**, 119-23 (1999).
24. Kewley, R.J., Whitelaw, M.L. & Chapman-Smith, A. The mammalian basic helix-loop-helix/PAS family of transcriptional regulators. *Int J Biochem Cell Biol* **36**, 189-204 (2004).
25. Card, P.B., Erbel, P.J. & Gardner, K.H. Structural basis of ARNT PAS-B dimerization: use of a common beta-sheet interface for hetero- and homodimerization. *J Mol Biol* **353**, 664-77 (2005).
26. Ferre-D'Amare, A.R., Pognonec, P., Roeder, R.G. & Burley, S.K. Structure and function of the b/HLH/Z domain of USF. *Embo J* **13**, 180-9 (1994).

27. Sauve, S., Tremblay, L. & Lavigne, P. The NMR solution structure of a mutant of the Max b/HLH/LZ free of DNA: insights into the specific and reversible DNA binding mechanism of dimeric transcription factors. *J Mol Biol* **342**, 813-32 (2004).
28. Yildiz, O., Doi, M., Yujnovsky, I., Cardone, L., Berndt, A., Hennig, S., Schulze, S., Urbanke, C., Sassone-Corsi, P. & Wolf, E. Crystal structure and interactions of the PAS repeat region of the Drosophila clock protein PERIOD. *Mol Cell* **17**, 69-82 (2005).
29. Michel, G., Minet, E., Mottet, D., Remacle, J. & Michiels, C. Site-directed mutagenesis studies of the hypoxia-inducible factor-1alpha DNA-binding domain. *Biochim Biophys Acta* **1578**, 73-83 (2002).
30. Wenger, R.H., Kvietikova, I., Rolfs, A., Camenisch, G. & Gassmann, M. Oxygen-regulated erythropoietin gene expression is dependent on a CpG methylation-free hypoxia-inducible factor-1 DNA-binding site. *Eur J Biochem* **253**, 771-7 (1998).
31. Ivan, M., Kondo, K., Yang, H., Kim, W., Valiando, J., Ohh, M., Salic, A., Asara, J.M., Lane, W.S. & Kaelin, W.G., Jr. HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. *Science* **292**, 464-8 (2001).
32. Jaakkola, P., Mole, D.R., Tian, Y.M., Wilson, M.I., Gielbert, J., Gaskell, S.J., Kriegsheim, A., Hebestreit, H.F., Mukherji, M., Schofield, C.J., Maxwell, P.H., Pugh, C.W. & Ratcliffe, P.J. Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. *Science* **292**, 468-72 (2001).
33. Maxwell, P.H., Wiesener, M.S., Chang, G.W., Clifford, S.C., Vaux, E.C., Cockman, M.E., Wykoff, C.C., Pugh, C.W., Maher, E.R. & Ratcliffe, P.J. The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature* **399**, 271-275 (1999).
34. McNeill, L.A., Hewitson, K.S., Gleadle, J.M., Horsfall, L.E., Oldham, N.J., Maxwell, P.H., Pugh, C.W., Ratcliffe, P.J. & Schofield, C.J. The use of dioxygen by HIF prolyl hydroxylase (PHD1). *Bioorg Med Chem Lett* **12**, 1547-50 (2002).
35. Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu Rev Biochem* **67**, 425-79 (1998).
36. Stebbins, C.E., Kaelin, W.G., Jr. & Pavletich, N.P. Structure of the VHL-ElonginC-ElonginB complex: implications for VHL tumor suppressor function. *Science* **284**, 455-61 (1999).
37. Lonergan, K.M., Iliopoulos, O., Ohh, M., Kamura, T., Conaway, R.C., Conaway, J.W. & Kaelin, W.G., Jr. Regulation of hypoxia-inducible mRNAs by the von Hippel-Lindau tumor suppressor protein requires binding to complexes containing elongins B/C and Cul2. *Mol Cell Biol* **18**, 732-41 (1998).
38. Kamura, T., Koepp, D.M., Conrad, M.N., Skowyra, D., Moreland, R.J., Iliopoulos, O., Lane, W.S., Kaelin, W.G., Jr., Elledge, S.J., Conaway, R.C., Harper, J.W. & Conaway, J.W. Rbx1, a component of the VHL tumor suppressor complex and SCF ubiquitin ligase. *Science* **284**, 657-61 (1999).
39. Zheng, N., Schulman, B.A., Song, L., Miller, J.J., Jeffrey, P.D., Wang, P., Chu, C., Koepp, D.M., Elledge, S.J., Pagano, M., Conaway, R.C., Conaway, J.W., Harper, J.W. & Pavletich, N.P. Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. *Nature* **416**, 703-9 (2002).
40. Hon, W.C., Wilson, M.I., Harlos, K., Claridge, T.D., Schofield, C.J., Pugh, C.W., Maxwell, P.H., Ratcliffe, P.J., Stuart, D.I. & Jones, E.Y. Structural basis for the recognition of hydroxyproline in HIF-1 alpha by pVHL. *Nature* **417**, 975-8 (2002).
41. Min, J.H., Yang, H., Ivan, M., Gertler, F., Kaelin, W.G., Jr. & Pavletich, N.P. Structure of an HIF-1alpha -pVHL complex: hydroxyproline recognition in signaling. *Science* **296**, 1886-9 (2002).
42. DeRider, M.L., Wilkens, S.J., Waddell, M.J., Bretscher, L.E., Weinhold, F., Raines, R.T. & Markley, J.L. Collagen stability: insights from NMR spectroscopic and hybrid density functional computational investigations of the effect of electronegative substituents on prolyl ring conformations. *J Am Chem Soc* **124**, 2497-505 (2002).
43. Dames, S.A., Martinez-Yamout, M., De Guzman, R.N., Dyson, H.J. & Wright, P.E. Structural basis for Hif-1 alpha /CBP recognition in the cellular hypoxic response. *Proc Natl Acad Sci U S A* **99**, 5271-6 (2002).
44. Freedman, S.J., Sun, Z.Y., Poy, F., Kung, A.L., Livingston, D.M., Wagner, G. & Eck, M.J. Structural basis for recruitment of CBP/p300 by hypoxia-inducible factor-1 alpha. *Proc Natl Acad Sci U S A* **99**, 5367-72 (2002).
45. McNeill, L.A., Hewitson, K.S., Claridge, T.D., Seibel, J.F., Horsfall, L.E. & Schofield, C.J. Hypoxia-inducible factor asparaginyl hydroxylase (FIH-1) catalyses hydroxylation at the beta-carbon of asparagine-803. *Biochem J* **367**, 571-5 (2002).
46. Lando, D., Peet, D.J., Whelan, D.A., Gorman, J.J. & Whitelaw, M.L. Asparagine hydroxylation of the HIF transactivation domain a hypoxic switch. *Science* **295**, 858-61 (2002).
47. Metzen, E., Berchner-Pfannschmidt, U., Stengel, P., Marxsen, J.H., Stolze, I., Klinger, M., Huang, W.Q., Wotzlaw, C., Hellwig-Burgel, T., Jelkmann, W., Acker, H. & Fandrey, J. Intracellular localisation of human HIF-1 alpha hydroxylases: implications for oxygen sensing. *J Cell Sci* **116**, 1319-26 (2003).

48. Bhattacharya, S., Michels, C.L., Leung, M.K., Arany, Z.P., Kung, A.L. & Livingston, D.M. Functional role of p35srj, a novel p300/CBP binding protein, during transactivation by HIF-1. *Genes Dev* **13**, 64-75 (1999).
49. Semenza, G.L. Hypoxia-inducible factor 1 (HIF-1) pathway. *Sci STKE* **2007**, cm8 (2007).
50. Baek, J.H., Liu, Y.V., McDonald, K.R., Wesley, J.B., Zhang, H. & Semenza, G.L. Spermidine/spermine N(1)-acetyltransferase-1 binds to hypoxia-inducible factor-1alpha (HIF-1alpha) and RACK1 and promotes ubiquitination and degradation of HIF-1alpha. *J Biol Chem* **282**, 33358-66 (2007).
51. Liu, Y.V., Baek, J.H., Zhang, H., Diez, R., Cole, R.N. & Semenza, G.L. RACK1 competes with HSP90 for binding to HIF-1alpha and is required for O(2)-independent and HSP90 inhibitor-induced degradation of HIF-1alpha. *Mol Cell* **25**, 207-17 (2007).
52. Liu, Y.V. & Semenza, G.L. RACK1 vs. HSP90: competition for HIF-1 alpha degradation vs. stabilization. *Cell Cycle* **6**, 656-9 (2007).
53. Appelhoff, R.J., Tian, Y.M., Raval, R.R., Turley, H., Harris, A.L., Pugh, C.W., Ratcliffe, P.J. & Gleadle, J.M. Differential function of the prolyl hydroxylases PHD1, PHD2, and PHD3 in the regulation of hypoxia-inducible factor. *J Biol Chem* **279**, 38458-65 (2004).
54. Stiehl, D.P., Wirthner, R., Koditz, J., Spielmann, P., Camenisch, G. & Wenger, R.H. Increased prolyl 4-hydroxylase domain proteins compensate for decreased oxygen levels. Evidence for an autoregulatory oxygen-sensing system. *J Biol Chem* **281**, 23482-91 (2006).
55. Schofield, C.J. & Ratcliffe, P.J. Signalling hypoxia by HIF hydroxylases. *Biochem Biophys Res Commun* **338**, 617-26 (2005).
56. Wei, W. & Yu, X.D. Hypoxia-inducible factors: crosstalk between their protein stability and protein degradation. *Cancer Lett* **257**, 145-56 (2007).
57. Pan, Y., Mansfield, K.D., Bertozzi, C.C., Rudenko, V., Chan, D.A., Giaccia, A.J. & Simon, M.C. Multiple factors affecting cellular redox status and energy metabolism modulate hypoxia-inducible factor prolyl hydroxylase activity in vivo and in vitro. *Mol Cell Biol* **27**, 912-25 (2007).
58. Selak, M.A., Armour, S.M., MacKenzie, E.D., Boulahbel, H., Watson, D.G., Mansfield, K.D., Pan, Y., Simon, M.C., Thompson, C.B. & Gottlieb, E. Succinate links TCA cycle dysfunction to oncogenesis by inhibiting HIF-alpha prolyl hydroxylase. *Cancer Cell* **7**, 77-85 (2005).
59. Isaacs, J.S., Jung, Y.J., Mole, D.R., Lee, S., Torres-Cabala, C., Chung, Y.L., Merino, M., Trepel, J., Zbar, B., Toro, J., Ratcliffe, P.J., Linehan, W.M. & Neckers, L. HIF overexpression correlates with biallelic loss of fumarate hydratase in renal cancer: novel role of fumarate in regulation of HIF stability. *Cancer Cell* **8**, 143-53 (2005).
60. Brune, B. & Zhou, J. Nitric oxide and superoxide: interference with hypoxic signaling. *Cardiovasc Res* **75**, 275-82 (2007).
61. Kaelin, W.G., Jr. ROS: really involved in oxygen sensing. *Cell Metab* **1**, 357-8 (2005).
62. Klimova, T. & Chandel, N.S. Mitochondrial complex III regulates hypoxic activation of HIF. *Cell Death Differ* **15**, 660-6 (2008).
63. Hagen, T., Taylor, C.T., Lam, F. & Moncada, S. Redistribution of intracellular oxygen in hypoxia by nitric oxide: effect on HIF1alpha. *Science* **302**, 1975-8 (2003).
64. Gerald, D., Berra, E., Frapart, Y.M., Chan, D.A., Giaccia, A.J., Mansuy, D., Pouyssegur, J., Yaniv, M. & Mechta-Grigoriou, F. JunD reduces tumor angiogenesis by protecting cells from oxidative stress. *Cell* **118**, 781-94 (2004).
65. Nakayama, K., Frew, I.J., Hagensen, M., Skals, M., Habelhah, H., Bhoumik, A., Kadoya, T., Erdjument-Bromage, H., Tempst, P., Frappell, P.B., Bowtell, D.D. & Ronai, Z. Siah2 regulates stability of prolyl-hydroxylases, controls HIF1alpha abundance, and modulates physiological responses to hypoxia. *Cell* **117**, 941-52 (2004).
66. McMahon, S., Charbonneau, M., Grandmont, S., Richard, D.E. & Dubois, C.M. Transforming growth factor beta1 induces hypoxia-inducible factor-1 stabilization through selective inhibition of PHD2 expression. *J Biol Chem* **281**, 24171-81 (2006).
67. Wolf, G. Role of reactive oxygen species in angiotensin II-mediated renal growth, differentiation, and apoptosis. *Antioxid Redox Signal* **7**, 1337-45 (2005).
68. Page, E.L., Chan, D.A., Giaccia, A.J., Levine, M. & Richard, D.E. Hypoxia-inducible Factor-1{alpha} Stabilization in Nonhypoxic Conditions: Role of Oxidation and Intracellular Ascorbate Depletion. *Mol Biol Cell* **19**, 86-94 (2008).
69. Cummins, E.P. & Taylor, C.T. Hypoxia-responsive transcription factors. *Pflugers Arch* **450**, 363-71 (2005).
70. Cummins, E.P., Berra, E., Comerford, K.M., Ginouves, A., Fitzgerald, K.T., Seeballuck, F., Godson, C., Nielsen, J.E., Moynagh, P., Pouyssegur, J. & Taylor, C.T. Prolyl hydroxylase-1 negatively regulates IkappaB kinase-beta, giving insight into hypoxia-induced NFkappaB activity. *Proc Natl Acad Sci U S A* **103**, 18154-9 (2006).

71. Bruick, R.K. & McKnight, S.L. A conserved family of prolyl-4-hydroxylases that modify HIF. *Science* **294**, 1337-40 (2001).
72. Epstein, A.C., Gleadle, J.M., McNeill, L.A., Hewitson, K.S., O'Rourke, J., Mole, D.R., Mukherji, M., Metzen, E., Wilson, M.I., Dhanda, A., Tian, Y.M., Masson, N., Hamilton, D.L., Jaakkola, P., Barstead, R., Hodgkin, J., Maxwell, P.H., Pugh, C.W., Schofield, C.J. & Ratcliffe, P.J. C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell* **107**, 43-54 (2001).
73. Hirsila, M., Koivunen, P., Gunzler, V., Kivirikko, K.I. & Myllyharju, J. Characterization of the human prolyl 4-hydroxylases that modify the hypoxia-inducible factor. *J Biol Chem* **278**, 30772-80 (2003).
74. Choi, K.O., Lee, T., Lee, N., Kim, J.H., Yang, E.G., Yoon, J.M., Kim, J.H., Lee, T.G. & Park, H. Inhibition of the catalytic activity of hypoxia-inducible factor-1alpha-prolyl-hydroxylase 2 by a MYND-type zinc finger. *Mol Pharmacol* **68**, 1803-9 (2005).
75. Flashman, E., Bagg, E.A., Chowdhury, R., Mecinovic, J., Loenarz, C., McDonough, M.A., Hewitson, K.S. & Schofield, C.J. Kinetic rationale for selectivity toward N- and C-terminal oxygen-dependent degradation domain substrates mediated by a loop region of hypoxia-inducible factor prolyl hydroxylases. *J Biol Chem* **283**, 3808-15 (2008).
76. Villar, D., Vara-Vega, A., Landazuri, M.O. & Del Peso, L. Identification of a region on hypoxia-inducible-factor prolyl 4-hydroxylases that determines their specificity for the oxygen degradation domains. *Biochem J* **408**, 231-40 (2007).
77. Huang, J., Zhao, Q., Mooney, S.M. & Lee, F.S. Sequence determinants in hypoxia-inducible factor-1alpha for hydroxylation by the prolyl hydroxylases PHD1, PHD2, and PHD3. *J Biol Chem* **277**, 39792-800 (2002).
78. Koivunen, P., Tiainen, P., Hyvarinen, J., Williams, K.E., Sormunen, R., Klaus, S.J., Kivirikko, K.I. & Myllyharju, J. An endoplasmic reticulum transmembrane prolyl 4-hydroxylase is induced by hypoxia and acts on hypoxia-inducible factor alpha. *J Biol Chem* **282**, 30544-52 (2007).
79. Ehrismann, D., Flashman, E., Genn, D.N., Mathioudakis, N., Hewitson, K.S., Ratcliffe, P.J. & Schofield, C.J. Studies on the activity of the hypoxia-inducible-factor hydroxylases using an oxygen consumption assay. *Biochem J* **401**, 227-34 (2007).
80. McNeill, L.A., Flashman, E., Buck, M.R., Hewitson, K.S., Clifton, I.J., Jeschke, G., Claridge, T.D., Ehrismann, D., Oldham, N.J. & Schofield, C.J. Hypoxia-inducible factor prolyl hydroxylase 2 has a high affinity for ferrous iron and 2-oxoglutarate. *Mol Biosyst* **1**, 321-4 (2005).
81. Koivunen, P., Hirsila, M., Kivirikko, K.I. & Myllyharju, J. The Length of Peptide Substrates Has a Marked Effect on Hydroxylation by the Hypoxia-inducible Factor Prolyl 4-Hydroxylases. *J Biol Chem* **281**, 28712-20 (2006).
82. Koivunen, P., Hirsila, M., Gunzler, V., Kivirikko, K.I. & Myllyharju, J. Catalytic properties of the asparaginyl hydroxylase (FIH) in the oxygen sensing pathway are distinct from those of its prolyl 4-hydroxylases. *J Biol Chem* **279**, 9899-904 (2004).
83. Kaelin, W.G., Jr. & Ratcliffe, P.J. Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway. *Mol Cell* **30**, 393-402 (2008).
84. Tuckerman, J.R., Zhao, Y., Hewitson, K.S., Tian, Y.M., Pugh, C.W., Ratcliffe, P.J. & Mole, D.R. Determination and comparison of specific activity of the HIF-prolyl hydroxylases. *FEBS Lett* **576**, 145-50 (2004).
85. Stolze, I.P., Tian, Y.M., Appelhoff, R.J., Turley, H., Wykoff, C.C., Gleadle, J.M. & Ratcliffe, P.J. Genetic analysis of the role of the asparaginyl hydroxylase factor inhibiting hypoxia-inducible factor (HIF) in regulating HIF transcriptional target genes. *J Biol Chem* **279**, 42719-25 (2004).
86. Bracken, C.P., Fedele, A.O., Linke, S., Balrak, W., Lisy, K., Whitelaw, M.L. & Peet, D.J. Cell-specific regulation of hypoxia-inducible factor (HIF)-1alpha and HIF-2alpha stabilization and transactivation in a graded oxygen environment. *J Biol Chem* **281**, 22575-85 (2006).
87. McDonough, M.A., Li, V., Flashman, E., Chowdhury, R., Mohr, C., Lienard, B.M., Zondlo, J., Oldham, N.J., Clifton, I.J., Lewis, J., McNeill, L.A., Kurzeja, R.J., Hewitson, K.S., Yang, E., Jordan, S., Syed, R.S. & Schofield, C.J. Cellular oxygen sensing: Crystal structure of hypoxia-inducible factor prolyl hydroxylase (PHD2). *Proc Natl Acad Sci U S A* **103**, 9814-9 (2006).
88. Dann, C.E., 3rd, Bruick, R.K. & Deisenhofer, J. Structure of factor-inhibiting hypoxia-inducible factor 1: An asparaginyl hydroxylase involved in the hypoxic response pathway. *Proc Natl Acad Sci U S A* **99**, 15351-6 (2002).
89. Elkins, J.M., Hewitson, K.S., McNeill, L.A., Seibel, J.F., Schlemminger, I., Pugh, C.W., Ratcliffe, P.J. & Schofield, C.J. Structure of factor-inhibiting hypoxia-inducible factor (HIF) reveals mechanism of oxidative modification of HIF-1 alpha. *J Biol Chem* **278**, 1802-6 (2003).
90. Lee, C., Kim, S.J., Jeong, D.G., Lee, S.M. & Ryu, S.E. Structure of human FIH-1 reveals a unique active site pocket and interaction sites for HIF-1 and von Hippel-Lindau. *J Biol Chem* **278**, 7558-63 (2003).

91. Borowski, T., Bassan, A. & Siegbahn, P.E. Mechanism of dioxygen activation in 2-oxoglutarate-dependent enzymes: a hybrid DFT study. *Chemistry* **10**, 1031-41 (2004).
92. Bugg, T.D. Oxygenases: mechanisms and structural motifs for O(2) activation. *Curr Opin Chem Biol* **5**, 550-5 (2001).
93. Bugg, T.D. & Ramaswamy, S. Non-heme iron-dependent dioxygenases: unravelling catalytic mechanisms for complex enzymatic oxidations. *Curr Opin Chem Biol* **12**, 134-40 (2008).
94. Xin, M. & Bugg, T.D. Evidence from mechanistic probes for distinct hydroperoxide rearrangement mechanisms in the intradiol and extradiol catechol dioxygenases. *J Am Chem Soc* **130**, 10422-30 (2008).
95. Price, J.C., Barr, E.W., Tirupati, B., Bollinger, J.M., Jr. & Krebs, C. The first direct characterization of a high-valent iron intermediate in the reaction of an alpha-ketoglutarate-dependent dioxygenase: a high-spin FeIV complex in taurine/alpha-ketoglutarate dioxygenase (TauD) from Escherichia coli. *Biochemistry* **42**, 7497-508 (2003).
96. Hoffart, L.M., Barr, E.W., Guyer, R.B., Bollinger, J.M., Jr. & Krebs, C. Direct spectroscopic detection of a C-H-cleaving high-spin Fe(IV) complex in a prolyl-4-hydroxylase. *Proc Natl Acad Sci U S A* **103**, 14738-43 (2006).
97. Clifton, I.J., McDonough, M.A., Ehrismann, D., Kershaw, N.J., Granatino, N. & Schofield, C.J. Structural studies on 2-oxoglutarate oxygenases and related double-stranded beta-helix fold proteins. *J Inorg Biochem* **100**, 644-69 (2006).
98. Hausinger, R.P. FeII/alpha-ketoglutarate-dependent hydroxylases and related enzymes. *Crit Rev Biochem Mol Biol* **39**, 21-68 (2004).
99. Cioffi, C.L., Liu, X.Q., Kosinski, P.A., Garay, M. & Bowen, B.R. Differential regulation of HIF-1 alpha prolyl-4-hydroxylase genes by hypoxia in human cardiovascular cells. *Biochem Biophys Res Commun* **303**, 947-53 (2003).
100. Hirsila, M., Koivunen, P., Gunzler, V., Kivirikko, K.I. & Myllyharju, J. Characterization of the human prolyl 4-hydroxylases that modify the hypoxia-inducible factor. *J. Biol. Chem.* **278**, 30772-30780 (2003).
101. Lieb, M.E., Menzies, K., Moschella, M.C., Ni, R. & Taubman, M.B. Mammalian EGLN genes have distinct patterns of mRNA expression and regulation. *Biochem Cell Biol* **80**, 421-6 (2002).
102. Aprelikova, O., Chandramouli, G.V., Wood, M., Vasselli, J.R., Riss, J., Maranchie, J.K., Linehan, W.M. & Barrett, J.C. Regulation of HIF prolyl hydroxylases by hypoxia-inducible factors. *J Cell Biochem* **92**, 491-501 (2004).
103. Berra, E., Benizri, E., Ginouves, A., Volmat, V., Roux, D. & Pouyssegur, J. HIF prolyl-hydroxylase 2 is the key oxygen sensor setting low steady-state levels of HIF-1alpha in normoxia. *Embo J* **22**, 4082-90 (2003).
104. D'Angelo, G., Duplan, E., Boyer, N., Vigne, P. & Frelin, C. Hypoxia up-regulates prolyl hydroxylase activity: a feedback mechanism that limits HIF-1 responses during reoxygenation. *J Biol Chem* **278**, 38183-7 (2003).
105. Marxsen, J.H., Stengel, P., Doege, K., Heikkinen, P., Jokilehto, T., Wagner, T., Jelkmann, W., Jaakkola, P. & Metzen, E. Hypoxia-inducible factor-1 (HIF-1) promotes its degradation by induction of HIF-alpha-prolyl-4-hydroxylases. *Biochem J* **381**, 761-7 (2004).
106. Aragones, J., Schneider, M., Van Geyte, K., Fraisl, P., Dresselaers, T., Mazzone, M., Dirkx, R., Zacchigna, S., Lemieux, H., Jeoung, N.H., Lambrechts, D., Bishop, T., Lafuste, P., Diez-Juan, A., Harten, S.K., Van Noten, P., De Bock, K., Willam, C., Tjwa, M., Grosfeld, A., Navet, R., Moons, L., Vandendriessche, T., Deroose, C., Wijeyekoon, B., Nuyts, J., Jordan, B., Silasi-Mansat, R., Lupu, F., Dewerchin, M., Pugh, C., Salmon, P., Mortelmans, L., Gallez, B., Gorus, F., Buyse, J., Sluse, F., Harris, R.A., Gnaiger, E., Hespel, P., Van Hecke, P., Schuit, F., Van Veldhoven, P., Ratcliffe, P., Baes, M., Maxwell, P. & Carmeliet, P. Deficiency or inhibition of oxygen sensor Phd1 induces hypoxia tolerance by reprogramming basal metabolism. *Nat Genet* **40**, 170-80 (2008).
107. Jokilehto, T., Rantanen, K., Luukkaa, M., Heikkinen, P., Grenman, R., Minn, H., Kronqvist, P. & Jaakkola, P.M. Overexpression and nuclear translocation of hypoxia-inducible factor prolyl hydroxylase PHD2 in head and neck squamous cell carcinoma is associated with tumor aggressiveness. *Clin Cancer Res* **12**, 1080-7 (2006).
108. Chelsky, D., Ralph, R. & Jonak, G. Sequence requirements for synthetic peptide-mediated translocation to the nucleus. *Mol Cell Biol* **9**, 2487-92 (1989).
109. Baek, J.H., Mahon, P.C., Oh, J., Kelly, B., Krishnamachary, B., Pearson, M., Chan, D.A., Giaccia, A.J. & Semenza, G.L. OS-9 interacts with hypoxia-inducible factor 1alpha and prolyl hydroxylases to promote oxygen-dependent degradation of HIF-1alpha. *Mol Cell* **17**, 503-12 (2005).
110. Flashman, E., McDonough, M.A. & Schofield, C.J. OS-9: another piece in the HIF complex story. *Mol Cell* **17**, 472-3 (2005).
111. Ozer, A., Wu, L.C. & Bruick, R.K. The candidate tumor suppressor ING4 represses activation of the hypoxia inducible factor (HIF). *Proc Natl Acad Sci U S A* **102**, 7481-6 (2005).

112. Kuznetsova, A.V., Meller, J., Schnell, P.O., Nash, J.A., Ignacak, M.L., Sanchez, Y., Conaway, J.W., Conaway, R.C. & Czyzyk-Krzeska, M.F. von Hippel-Lindau protein binds hyperphosphorylated large subunit of RNA polymerase II through a proline hydroxylation motif and targets it for ubiquitination. *Proc Natl Acad Sci U S A* **100**, 2706-11 (2003).
113. Cockman, M.E., Lancaster, D.E., Stolze, I.P., Hewitson, K.S., McDonough, M.A., Coleman, M.L., Coles, C.H., Yu, X., Hay, R.T., Ley, S.C., Pugh, C.W., Oldham, N.J., Masson, N., Schofield, C.J. & Ratcliffe, P.J. Posttranslational hydroxylation of ankyrin repeats in IkappaB proteins by the hypoxia-inducible factor (HIF) asparaginyl hydroxylase, factor inhibiting HIF (FIH). *Proc Natl Acad Sci U S A* **103**, 14767-72 (2006).
114. Barrick, D., Ferreiro, D.U. & Komives, E.A. Folding landscapes of ankyrin repeat proteins: experiments meet theory. *Curr Opin Struct Biol* **18**, 27-34 (2008).
115. Coleman, M.L., McDonough, M.A., Hewitson, K.S., Coles, C., Mecinovic, J., Edelmann, M., Cook, K.M., Cockman, M.E., Lancaster, D.E., Kessler, B.M., Oldham, N.J., Ratcliffe, P.J. & Schofield, C.J. Asparaginyl hydroxylation of the Notch ankyrin repeat domain by factor inhibiting hypoxia-inducible factor. *J Biol Chem* **282**, 24027-38 (2007).
116. Ferguson, J.E., 3rd, Wu, Y., Smith, K., Charles, P., Powers, K., Wang, H. & Patterson, C. ASB4 is a hydroxylation substrate of FIH and promotes vascular differentiation via an oxygen-dependent mechanism. *Mol Cell Biol* **27**, 6407-19 (2007).
117. Semenza, G.L. Targeting HIF-1 for cancer therapy. *Nat Rev Cancer* **3**, 721-32 (2003).
118. Semenza, G.L. Evaluation of HIF-1 inhibitors as anticancer agents. *Drug Discov Today* **12**, 853-9 (2007).
119. Raval, R.R., Lau, K.W., Tran, M.G., Sowter, H.M., Mandriota, S.J., Li, J.L., Pugh, C.W., Maxwell, P.H., Harris, A.L. & Ratcliffe, P.J. Contrasting properties of hypoxia-inducible factor 1 (HIF-1) and HIF-2 in von Hippel-Lindau-associated renal cell carcinoma. *Mol Cell Biol* **25**, 5675-86 (2005).
120. Wykoff, C.C., Beasley, N.J., Watson, P.H., Turner, K.J., Pastorek, J., Sibtain, A., Wilson, G.D., Turley, H., Talks, K.L., Maxwell, P.H., Pugh, C.W., Ratcliffe, P.J. & Harris, A.L. Hypoxia-inducible expression of tumor-associated carbonic anhydrases. *Cancer Res* **60**, 7075-83 (2000).
121. Percy, M.J., Furlow, P.W., Beer, P.A., Lappin, T.R., McMullin, M.F. & Lee, F.S. A novel erythrocytosis-associated PHD2 mutation suggests the location of a HIF binding groove. *Blood* **110**, 2193-6 (2007).
122. Percy, M.J., Zhao, Q., Flores, A., Harrison, C., Lappin, T.R., Maxwell, P.H., McMullin, M.F. & Lee, F.S. A family with erythrocytosis establishes a role for prolyl hydroxylase domain protein 2 in oxygen homeostasis. *Proc Natl Acad Sci U S A* **103**, 654-9 (2006).
123. Olenyuk, B.Z., Zhang, G.J., Klco, J.M., Nickols, N.G., Kaelin, W.G., Jr. & Dervan, P.B. Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc Natl Acad Sci U S A* **101**, 16768-73 (2004).
124. Choi, H., Chun, Y.S., Kim, S.W., Kim, M.S. & Park, J.W. Curcumin inhibits hypoxia-inducible factor-1 by degrading aryl hydrocarbon receptor nuclear translocator: a mechanism of tumor growth inhibition. *Mol Pharmacol* **70**, 1664-71 (2006).
125. Matsumoto, K., Imagawa, S., Obara, N., Suzuki, N., Takahashi, S., Nagasawa, T. & Yamamoto, M. 2-Oxoglutarate downregulates expression of vascular endothelial growth factor and erythropoietin through decreasing hypoxia-inducible factor-1alpha and inhibits angiogenesis. *J Cell Physiol* **209**, 333-40 (2006).
126. Kung, A.L., Zabludoff, S.D., France, D.S., Freedman, S.J., Tanner, E.A., Vieira, A., Cornell-Kennon, S., Lee, J., Wang, B., Wang, J., Memmert, K., Naegeli, H.U., Petersen, F., Eck, M.J., Bair, K.W., Wood, A.W. & Livingston, D.M. Small molecule blockade of transcriptional coactivation of the hypoxia-inducible factor pathway. *Cancer Cell* **6**, 33-43 (2004).
127. Cunliffe, C.J., Franklin, T.J., Hales, N.J. & Hill, G.B. Novel inhibitors of prolyl 4-hydroxylase. 3. Inhibition by the substrate analogue N-oxaloglycine and its derivatives. *J Med Chem* **35**, 2652-8 (1992).
128. Goldberg, M.A., Dunning, S.P. & Bunn, H.F. Regulation of the erythropoietin gene: evidence that the oxygen sensor is a heme protein. *Science* **242**, 1412-5 (1988).
129. McDonough, M.A., McNeill, L.A., Tilliet, M., Papamicael, C.A., Chen, Q.Y., Banerji, B., Hewitson, K.S. & Schofield, C.J. Selective inhibition of factor inhibiting hypoxia-inducible factor. *J Am Chem Soc* **127**, 7680-1 (2005).
130. Safran, M., Kim, W.Y., O'Connell, F., Flippin, L., Gunzler, V., Horner, J.W., Depinho, R.A. & Kaelin, W.G., Jr. Mouse model for noninvasive imaging of HIF prolyl hydroxylase activity: assessment of an oral agent that stimulates erythropoietin production. *Proc Natl Acad Sci U S A* **103**, 105-10 (2006).

Chapter 2 – Crystallisation and structure solution of PHD2 complexes

This chapter details various approaches used to obtain a PHD2-substrate complex structure.

2.1 Selection of PHD2 construct for structural work

Sequence analyses of PHD2 reveal that its N-terminal 180 amino acids constitute an intrinsically disordered region and a Myeloid Translocation Protein 8, Nervy and DEAF1 (MYND) zinc finger domain (Fig. 2.1). Primary sequence comparisons suggest that the zinc finger domain is not apparently present in other members of the 2OG oxygenase superfamily.

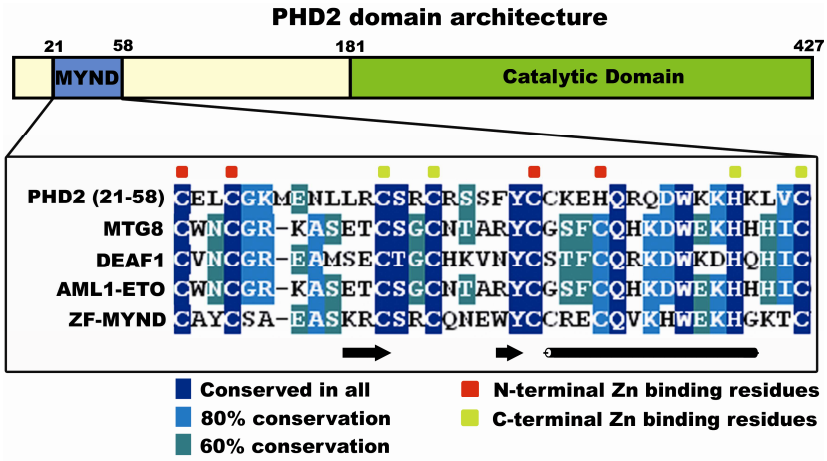


Fig. 2.1 PHD2 domain architecture showing the N-terminal MYND finger and the C-terminal catalytic domains.[1] Sequence alignment of PHD2 MYND zinc finger with other MYND domain containing proteins is presented in the box. To assign the secondary structures and Zn-binding sites, the following PDB coordinates were used: 2DJ8 (Myeloid Translocation Gene 8, MTG8), 2JW6 (Deformed Epidermal Autoregulatory Factor 1, DEAF1)[2], 2OD1 (Acute Myelogenous Leukemia 1-Eight Twenty One fusion protein, AML1-ETO)[3], 2D8Q (Zinc finger MYND domain-containing protein 10).

Because, at least some, of the N-terminal 180 residues of PHD2 were predicted to be largely disordered, it was considered that they may hinder crystallisation of full length PHD2. Even if a crystal could be obtained with the disordered region, structural disorder within the crystal lattice may impede data collection at higher resolution. Therefore, an N-terminally truncated PHD2 was considered suitable for structural work. A plasmid construct encoding N-terminal His_6 tagged $PHD2_{181-426}$ (tPHD2) was cloned into a pET-28a(+) vector (Novagen) by Dr K. S. Hewitson for crystallisation purposes. During the course of this work, N-terminally truncated forms of PHD2, i.e. without the MYND finger, were reported to catalyse hydroxylation of various HIF-1α CODD substrates (residues 401-603, GST-tagged and residues 556-575, biotinylated) more efficiently than the full length PHD2.[1] However, the presence of the MYND finger domain was not found to

substantially affect the activity of PHD2 in the current study, causing <20% loss of activity using a 19-residue HIF-1α $CODD_{556-574}$ substrate (CODD hereafter) (Fig. 2.2).

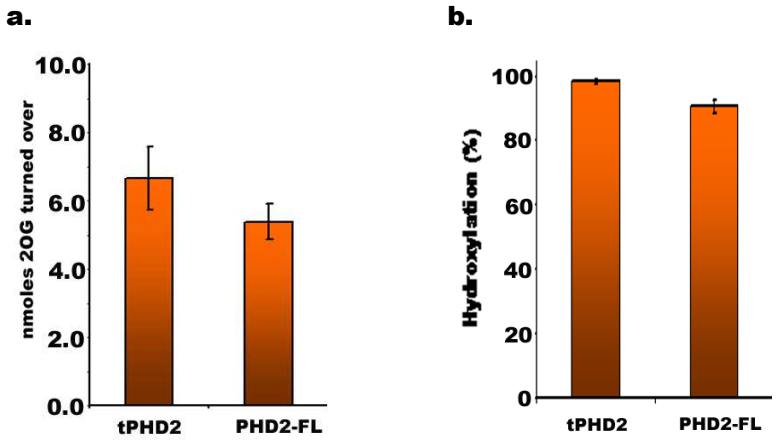


Fig. 2.2 2OG turnover (a) and hydroxylation assay (b) results of tPHD2 and the full-length PHD2 (PHD2-FL) using a CODD peptide substrate. See Appendix 1 for purification and characterisation of tPHD2.

2.2 Crystallisation attempts with tPHD2 that were successful for other related human 2OG oxygenases

Initial crystallisation attempts with tPHD2 utilised similar protein sample compositions that were previously successful in obtaining structures of some 2OG oxygenases in complex with peptide/nucleotide substrates (Table 2.1). PHD2 is highly unusual, if not unique, amongst studied 2OG oxygenases in that it forms a relatively stable complex with Fe(II) and 2OG.[4] tPHD2 produced in bacterial cells is reported to copurify with endogenous levels of Fe(II) cofactor and 2OG cosubstrate.[4] A heterogeneous mixture of the tPHD2.Fe(II) and tPHD2.Fe(II).2OG complexes thus obtained can be observed by spectroscopic techniques and by non-denaturing electrospray ionisation mass spectroscopy (ESI-MS).[4]

Some 2OG oxygenases are reported to undergo self inactivation via oxidation in the presence of Fe(II) and 2OG; in some cases this can lead to fragmentation (e.g. for phytanoyl-CoA 2-hydroxylase, PAHX).[5] These observations raise the possibility of generating reactive oxygen species even in the absence of substrate leading to a heterogeneous population, which is undesirable for crystallisation. Therefore, instead of using 2OG, its near isosteric analogue, *N*-oxalylglycine (NOG) that has been found to inhibit 2OG oxygenases, including the HIF hydroxylases, was used for crystallisation.[6]

With the aim of obtaining crystals of tPHD2.Fe(II).2OG/NOG complexes both with or without CODD, varying concentrations of 2OG/NOG and/or CODD were added to tPHD2 (10 mg.ml^{-1}). All crystallisation trials using Fe(II) were performed under near anaerobic conditions (pO_2 <0.1 ppm). Since no obvious correlations between crystallisation conditions and protein structure/family have been established so far, sparse matrix crystallisation screens were performed. Anaerobic screening conditions included pre-equilibriated Crystal Screen and Crystal Screen II (Hampton Research). In most cases these screens resulted in clear sitting/hanging drops with no apparent crystallisation hits.

Table 2.1. Protein sample compositions used in previously reported crystallisation of substrate complexes of human 2OG oxygenases. Substrate complex structures for FIH.2OG/NOG.CAD,[7] JMJD2A.2OG/NOG/succinate.H3K9Me(n)/H3K36Me[8-10] and ABH2.2OG.dsDNA[11] have been reported during the course of the current study.

PDB	Enzyme		Metal ions* (mM)	Ligand/ Additive (mM)	Substrate (mM)
	Construct	Conc. (mM)			
Factor Inhibiting HIF (FIH)					
1H2K[7]	wt (1-349)/ pET28a	~0.3	Fe^{2+} (1)	NOG (2.5)	$CAD_{786-826}$ (2.5)
1H2L[7]	wt (1-349)/ pET28a	~0.3	Fe^{2+} (1)	2OG (2.5)	$CAD_{786-826}$ (2.5)
1H2M[7]	wt (1-349)/ pET28a	~0.3	Zn^{2+} (1)	NOG (2.5)	$CAD_{775-826}$ (2.5)
xxxx[12]	wt (1-349)/ pET28a	~0.3	Fe^{2+} (1)	2OG (2.5)	mNotch-$1_{1930-1949}$ (2.5)
xxxx[12]	wt (1-349)/ pET28a	~0.3	Fe^{2+} (1)	2OG (2.5)	mNotch-$1_{1997-2016}$ (2.5)
JmjC domain-containing histone demethylation protein 3A (JMJD2A)					
2OQ6[10]	ΔC (1-359)/ pNIC28-Bsa4	~0.3	Ni^{2+} (4)	NOG (10)	H3K9Me3 (5)
2OS2[10]	ΔC (1-359)/ pNIC28-Bsa4	~0.3	Ni^{2+} (4)	NOG (10)	H3K36Me3 (5)
2OT7[10]	ΔC (1-359)/ pNIC28-Bsa4	~0.3	Ni^{2+} (4)	NOG (10)	H3K9Me1 (5)
2OX0[10]	ΔC (1-359)/ pNIC28-Bsa4	~0.3	Ni^{2+} (4)	NOG (10)	H3K9Me2 (5)
2Q8C[9]	ΔC (1-350)/ pET15b	n.r.	Ni^{2+} (1)	2OG (2x)[f]	H3K9Me3 (n.r.)
2Q8D[9]	ΔC (1-350)/ pET15b	n.r.	Ni^{2+} (1)	Succinate (2x)[f]	H3K36Me2 (n.r.)
2Q8E[9]	ΔC (1-350)/ pET15b	n.r.	Ni^{2+} (1)	NOG (2x)[f]	H3K36Me3 (n.r.)
2P5B[8]	ΔC (1-350)/ pGEX-4T-2	~0.4	Fe^{2+} (n.r.)	NOG (n.r.)	H3k36me3 (n.r.)
2PXJ[8]	ΔC (1-350)/ pGEX-4T-2	~0.4	Fe^{2+} (n.r.)	NOG (n.r.)	H3k36me1 (n.r.)
Alkylated DNA repair protein homologue 2 (ABH2)[§]					
3BTX[11]	ΔN(55-258) C67S/C165S/C192S/E175C		Mn^{2+} (1)	2OG (2)	dsDNA2
3BTY[11]	ΔN(55-258) C67S/C165S/C192S/169C		-	-	dsDNA6
3BTZ[11]	ΔN(55-258) C67S/C165S/C192S/E175C		-	-	dsDNA3
3BU0[11]	ΔN(55-258) C67S/C165S/C192S/E175C		Mn^{2+} (1)	2OG (2)	dsDNA2
3BUC[11]	ΔN(55-258) C67S/C165S/C192S/169C		Mn^{2+} (1)	2OG (2)	dsDNA6

* Protein samples containing Fe^{2+} were incubated anaerobically for crystallisation trials.

n.r. = Not reported; xxxx = coordinates not deposited; wt = wildtype protein.

ΔC, ΔN = truncation of C- or N-terminus, respectively.

[f] A two-fold molar excess of ligand (2OG/NOG/Succinate) was added.[9]

[§] The covalently linked complexes between ABH2 and dsDNA were purified and concentrated to 5 $mg.ml^{-1}$ before application of a hanging drop crystallisation screen. Crystals formed in this way were finally soaked in 2.5 mM Mn^{2+} and 2OG.[11]

Substitution of the iron by different transition metal ions has been used as a potential strategy to crystallise catalytically inert complexes of 2OG oxygenases, including a HIF hydroxylase (FIH) (Table 2.1). This strategy can be particularly useful for high-throughput screens using robots because these were not available under anaerobic atmosphere. Mass spectrometric analyses revealed that, except for Mg(II) and Ca(II), all other tested metal ions formed complexes with tPHD2.[13] Therefore, attempts were made to generate stable inactive complexes of tPHD2 using Co(II), Ni(II), Zn(II) and Mn(II) for crystallisation. Screens used included Crystal Screen, Crystal Screen 2, Index, Grid Screen PEG 6000, Grid Screen Ammonium Sulfate, Salt-Rx and PEG/Ion (Hampton Research), and the Wizard I and II (Emerald BioSystems). No suitable crystal hits were obtained using these screens, suggesting that optimisation of the protein sample itself was required.

2.3 Systematic approach to optimise protein sample composition

A more systematic approach was then followed to optimise the protein sample composition, as summarised in Fig. 2.3. In addition to optimising the protein buffer, variation of ligands and metal ions was carried out in the hope of obtaining high quality crystals of tPHD2.HIF-1α complexes. HIF-1α $CODD_{556-574}$ (CODD) and HIF-2α $CODD_{522-542}$ were chosen for the initial crystallisation experiments because they form more stable complexes with tPHD2 than HIF-1α $NODD_{395-413}$ (NODD hereafter), as determined by mass spectrometric analyses under non-denaturing conditions (discussed later). A variety of crystallisation screens (Section 2.2) were set up with the aim of crystallising tPHD2.metal ion.2OG/2OG analogue complexes, with or without HIF-1α/-2α CODD peptides, in 96-well Greiner crystallisation plates (using 100 nl protein + 100 nl buffer per sitting drop). Protein sample compositions for high-throughput crystallisation screening were optimised based on the results of fluorescence based thermal scans (Section 2.3.2), sedimentation velocity analyses for monodispersity (Section 2.3.3) and soft-ionisation mass spectrometric analyses (Section 2.3.4). The number of ligands was narrowed down on

the basis of the inhibition assay results (Section 2.3.5) or their ability to stabilise tPHD2 (Section 2.3.4). Crystals/crystallites, crystalline precipitates or phase separation were generally considered 'leads' or 'hits' for further optimisation.

The initial crystallisation hits thus obtained were fine-tuned by setting up further experiments with variations around the hit conditions. 24-well VDX (Hampton Research) plates were used for the optimisation; these employed larger quantities of protein samples (1 µl protein + 1 µl buffer per sitting/hanging drop). Many of the initial hit conditions could not be reproduced on scale-up, presumably because of the modified kinetics of evaporation/mixing, the surface-to-volume ratio, the size of the crystallisation drop and mass transport effects in the drops used by the nanolitre scale.[14]

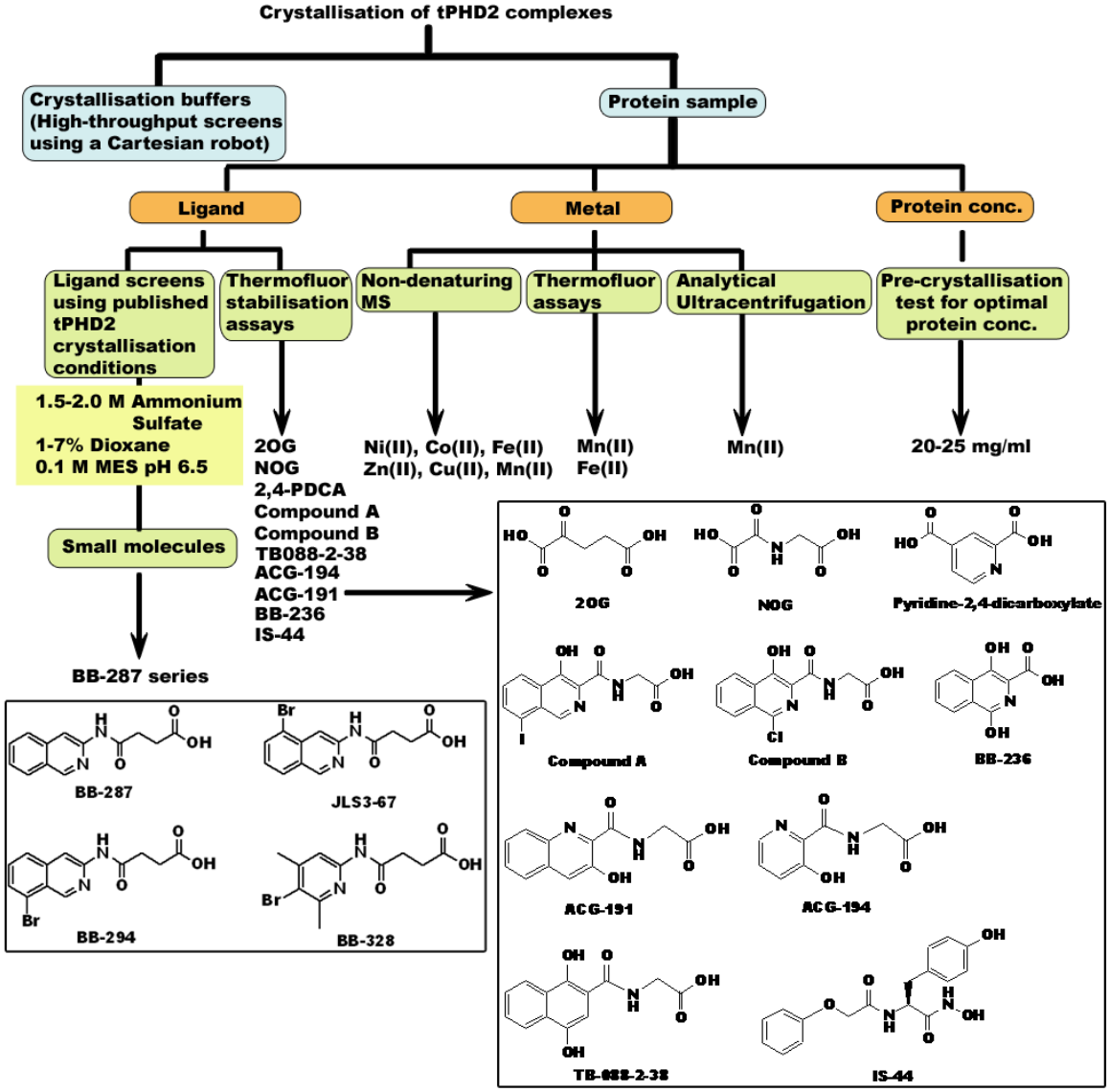


Fig. 2.3 A summary of different approaches and small molecules employed in the crystallisation attempts with tPHD2 complexes.

2.3.1 Pre-crystallisation testing

Crystallisation can be regarded as a process involving phase transition between liquid solutes and crystalline solid that requires the formation of effective nuclei followed by crystal growth (Fig. 2.4).[14] An appropriate protein concentration is a pre-requisite for crystallisation; too high a concentration can result in an amorphous precipitate whilst too low a concentration can result in clear drops. For further crystallisation experiments, an optimum tPHD2 concentration was determined as 20-25 $mg \cdot ml^{-1}$ using a pre-crystallisation test (PCT, Hampton Research).

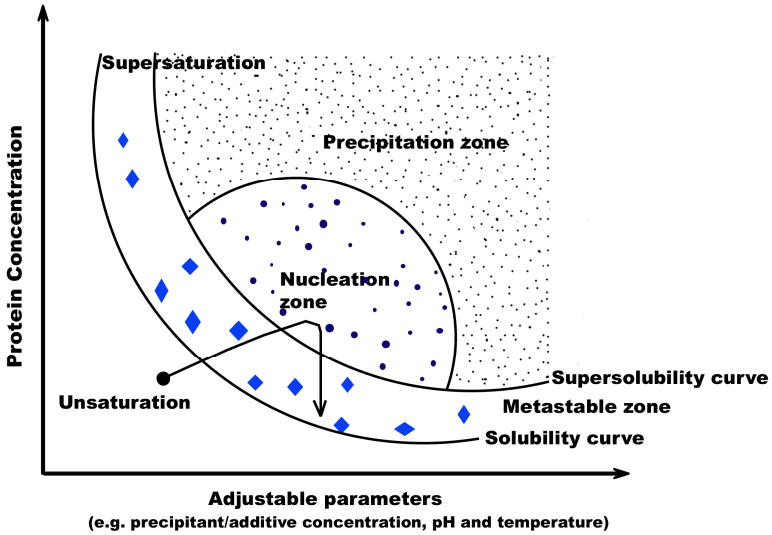


Fig. 2.4 Schematic representation of a phase diagram for protein crystallisation.[14] The solid black circle in the unsaturation phase represents a starting point. The solubility is defined as a protein concentration that is in equilibrium with crystals. The supersolubility curve is defined as the line separating conditions under which spontaneous nucleation/ phase separation/ precipitation occurs from those under which the crystallisation solution remains clear. The figure was modified from reference 14.

2.3.2 Thermal shift measurements using thermofluor assays

A fluorescence-based thermal shift assay was used for the high-throughput screening of protein stability as a function of protein buffers, pH and/or additives (see Chapter 10 for experimental procedures).[15,16] From a melting curve, the midpoint of the unfolding transition is taken as an approximation of the melting temperature (T_m) of a given protein/complex.[16] A positive shift in T_m indicates stabilisation of the protein, presumably by an increase in structural order and a reduction in conformational flexibility, while a negative shift in T_m indicates destabilisation. Comparison of T_m in the presence or absence of a particular ligand gives an indication of its binding affinity/magnitude of stabilisation.

Initially, attempts were made to optimise the buffer pH for tPHD2 over a test range of 4.0-9.4 using the T_m shift assay. The results indicated that tPHD2 appeared to be more stable at neutral to alkaline pH (7.0 to 9.4) and less stable at an acidic pH (Fig. 2.5). However, there

was no sign of precipitation of the protein within the experimental pH range, as there would have been no fluorescence signal if the protein had precipitated (Fig. 2.5).[16]

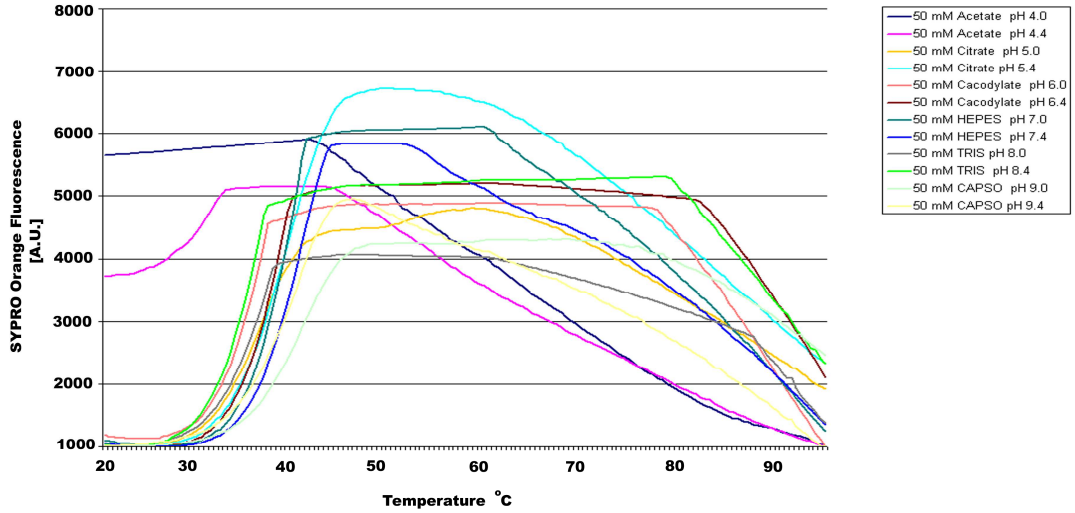


Fig. 2.5 Results of a thermal stability assay on tPHD2; T_m shifts were measured in different buffers with varying pH values. Of the conditions analysed, the protein appears to be most stable in 50 mM HEPES pH 7.0/7.4, 50 mM citrate pH 5.0/5.4 and 50 mM CAPSO pH 9.0/9.4 buffers.

The effect of different metal ions on tPHD2 stability was then investigated using the T_m shift assay (Fig. 2.6). Of all the transition metal ions tested (Fig. 2.6), Mn(II) showed the highest positive T_m shift. Addition of Ni(II) and in some cases, Zn(II) resulted in unusual unfolding (Fig. 2.6), which may reflect binding of more than one metal ion. Evidence for the binding of more than one Zn(II) ion to tPHD2 came from non-denaturing mass spectrometry.[13] With Fe(III), the transition was of low cooperativity and low fluorescence intensity was observed, both suggesting that the protein was most likely precipitated (Fig. 2.6). Consistent with this result, MS data under non-denaturing condition indicated that Fe(III) does not bind to tPHD2.

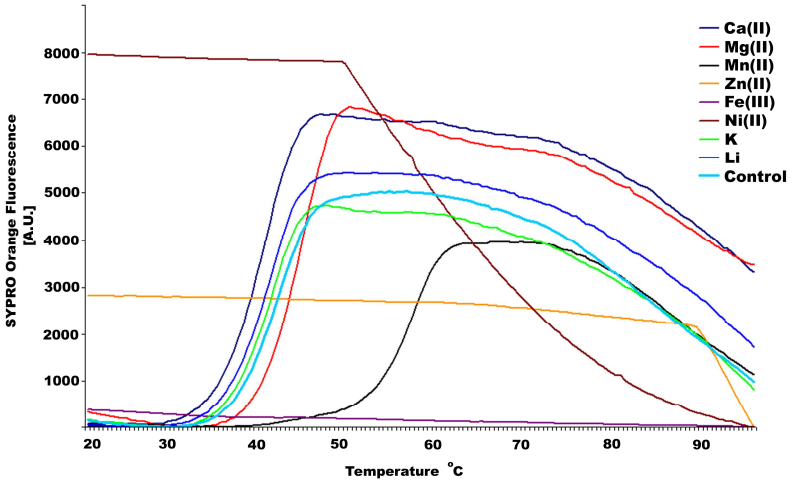
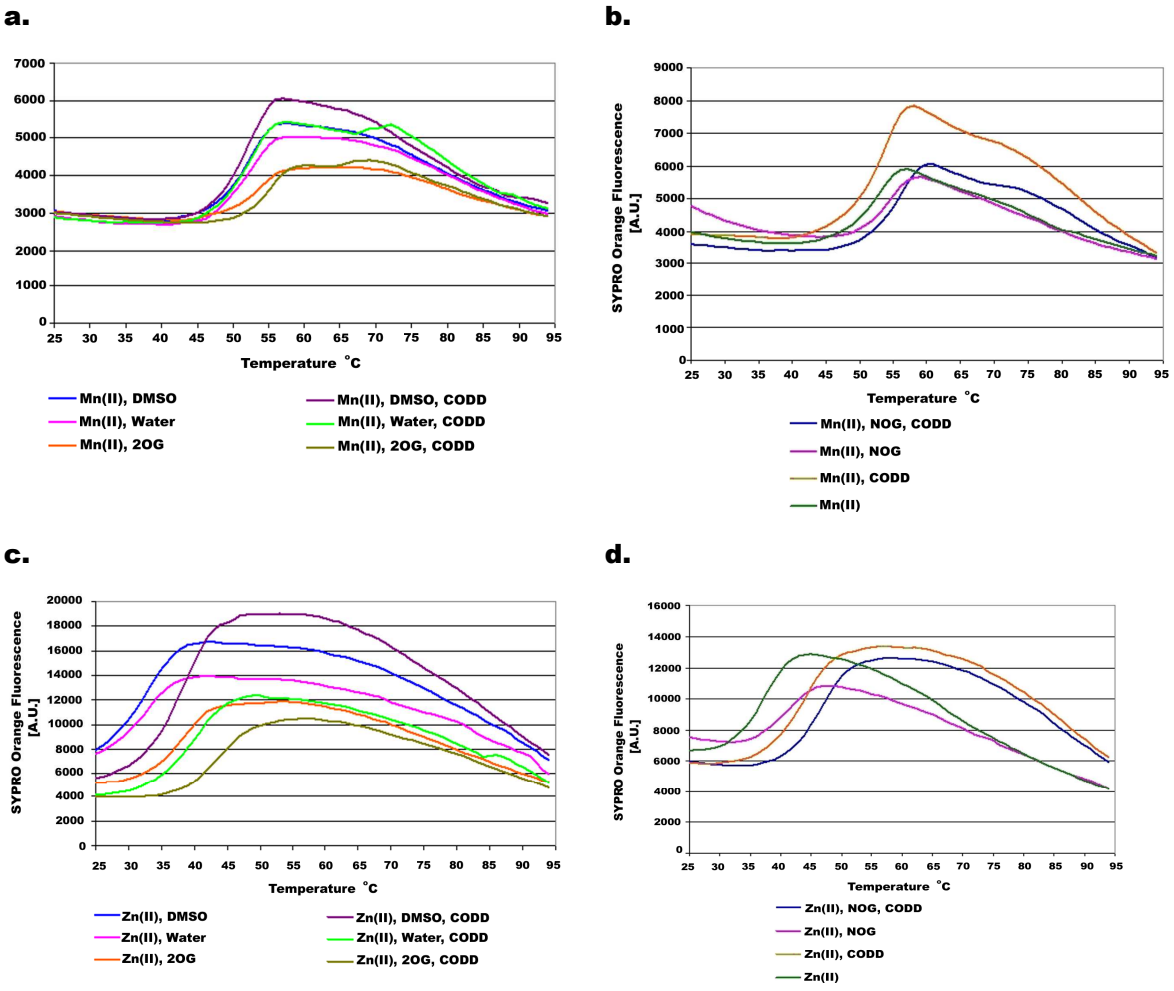


Fig. 2.6 The effect of transition metal ions on the stability of tPHD2 as measured by a fluorescence-coupled thermal shift assay.

The effects of 2OG/NOG in the presence/absence of CODD using Mn(II), Zn(II), Fe(II) and Ni(II), were then examined. Fe(II) was kept (as far as possible) in a reduced environment, provided by the addition of sodium dithionate. In all cases, the addition of either 2OG or NOG, showed a marked stabilising T_m shift (Fig. 2.7). Further stabilisation could be apparently attained by the addition of CODD to tPHD2.2OG/NOG (Fig. 2.7). However, CODD alone (without ligand) did not show significant stabilisation. Moreover, the addition of CODD to tPHD2.Fe(II)/Mn(II) resulted in apparent cooperative unfolding with or without any ligand (2OG/ NOG).

In summary, the largest stabilising effect was observed only in the presence of both a ligand (2OG and NOG) and CODD. These observations also suggest ligand and peptide bound to tPHD2 in the presence of different transition metal ions (Mn(II), Zn(II) and Fe(II)). The unusual unfolding with Ni(II), as observed with the metal ion alone (Fig. 2.6), was not altered by the addition of ligand (e.g. NOG) or CODD peptide (Fig. 2.7f).



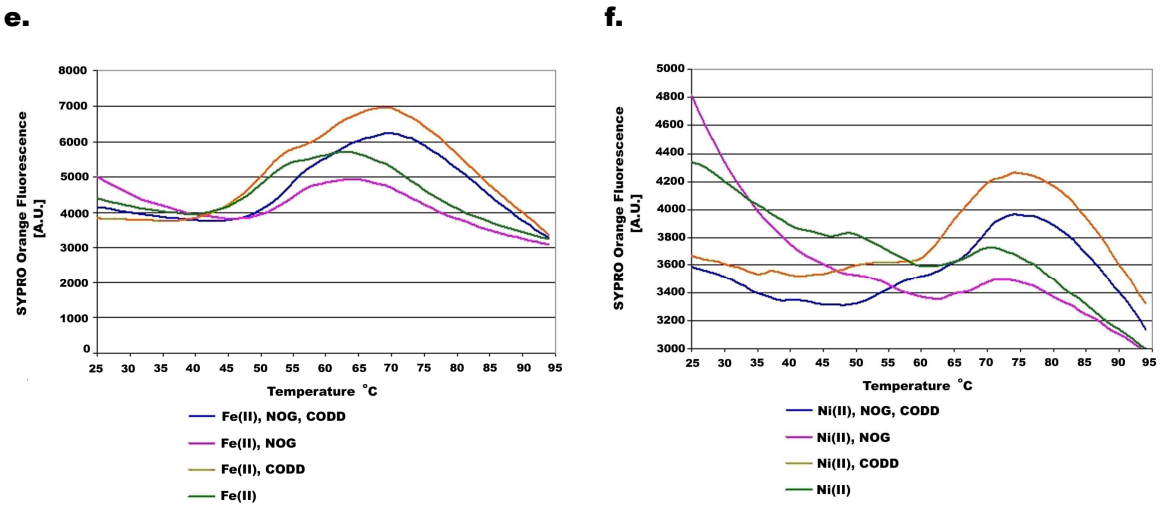


Fig. 2.7 Effect of Mn(II) (a and b), Zn(II) (c and d), Fe(II) (e) and Ni(II) (f), 2OG, NOG and a short CODD (19-mer) peptide substrate in stabilisation of tPHD2.

The thermal stability assay data indicated that Mn(II), along with 2OG or NOG, stabilised the tPHD2.CODD complex to the highest degree. Further attempts were made to identify other ligands that could impart stability to the same extent as 2OG, or to a higher degree. A total of ~200 compounds (mostly 2OG analogues) were included in the thermal stabilisation screen irrespective of their potency as tPHD2 inhibitors (Section 2.3.5). Because recombinant tPHD2 copurifies with Fe(II) and 2OG,[4] it was necessary to remove these cofactors in order to examine the absolute effect of other potential ligands on T_m shift.

An apo form of tPHD2 was produced by treating tPHD2 (as purified) with a 5000 fold excess of EDTA for 12 hours, to enable (almost) complete removal of Fe(II) and 2OG followed by a single step purification using gel filtration chromatography. A tPHD2.Mn(II) complex (produced by adding 1 equivalent of Mn(II) to apo-tPHD2 and verified by MS analyses under non-denaturing conditions) was then screened against a library of small molecules in an attempt to find an alternative stabilising agent (Fig. 2.8). The results identified a number of 2OG analogues that significantly stabilised the putative tPHD2.Mn(II).CODD complex (Fig. 2.8)

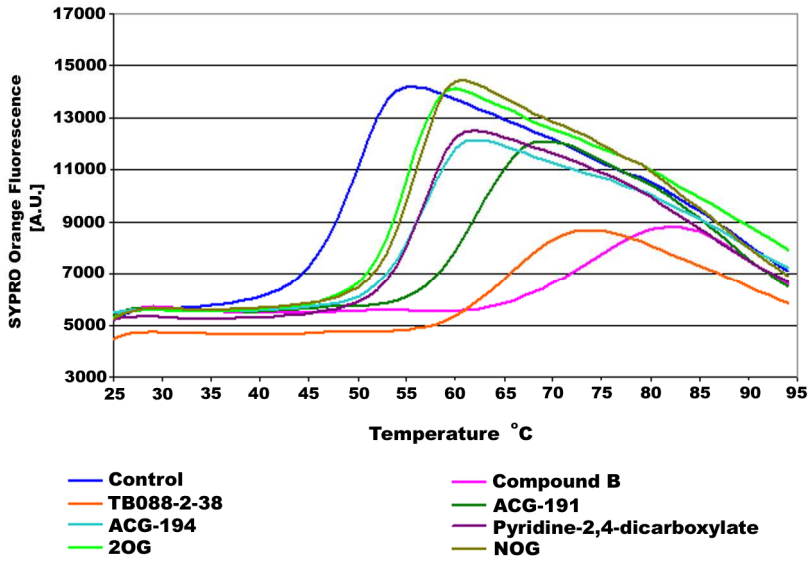
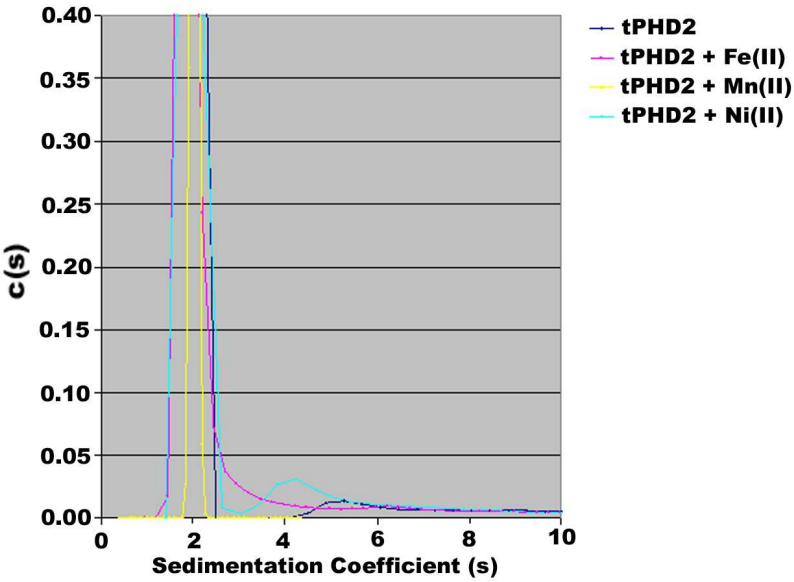


Fig. 2.8 Fluorescence-based T_m shift assay using the tPHD2.Mn(II).CODD complex. Results show that potential 2OG analogues, including acyclic (NOG), monocyclic (ACG-194, pyridine-2,4-dicarboxylate or 2,4-PDCA) and bicyclic (TB088-2-38, Compound B, ACG-191) compounds stabilise the complex. tPHD2.Mn(II).CODD without any ligand was used as the control. See Fig. 2.3 for chemical structures.

2.3.3 Sedimentation velocity analysis using analytical ultracentrifugation (AUC)

Consistent with the thermal stabilisation assays, analytical ultracentrifugation studies in velocity mode, demonstrated (almost) complete monodisperse behaviour of tPHD2 in solution; amongst the transition metal ions tested, Mn(II) showed the maximum monodispersity (Fig. 2.9).



	s	% as main peak
tPHD2	1.95	85
tPHD2 + Fe(II)	1.90	76
tPHD2 + Mn(II)	2.02	100
tPHD2 + Ni(II)	1.97	83

Fig. 2.9 Sedimentation coefficient distributions (c(s)) for tPHD2. The table shows rate of monodispersity as a function of the metal ion added. Amongst the tested metal ions, only Mn(II) caused monodisperse behaviour which was consistent with the results of the thermal scanning assays.

2.3.4 Verification of complex formation by MS analyses

MS under non-denaturing conditions revealed that, except for Mg(II) and Ca(II), the tested metal ions, including Co(II), Ni(II), Cu(II), Mn(II) and Zn(II) all bound to tPHD2,[13] as did the 2OG analogues, NOG, 2,4-PDCA, compound A/B, ACG-191, ACG-194 and ACG-198 (Fig. 2.10). Interestingly, the bicyclic analogues (Fig. 2.3) formed weaker ternary complexes with tPHD2.CODD as compared to monocyclic/acyclic compounds (Fig. 2.10). Addition of CODD to the tPHD2.Fe(II).inhibitor complexes yielded heterogeneous mixtures of tPHD2.Fe(II).CODD (major complex in case of bicyclic compounds), tPHD2.Fe(II).inhibitor and tPHD2.Fe(II).inhibitor.CODD (the major complex in case of acyclic compounds) complexes (Fig. 2.10). In addition, BB-287 (Chapter 7) formed only a weak complex with tPHD2; addition of CODD to tPHD2.Fe(II).BB-287 yielded a sample consisting of almost complete tPHD2.Fe(II).CODD complex (Fig. 2.10).

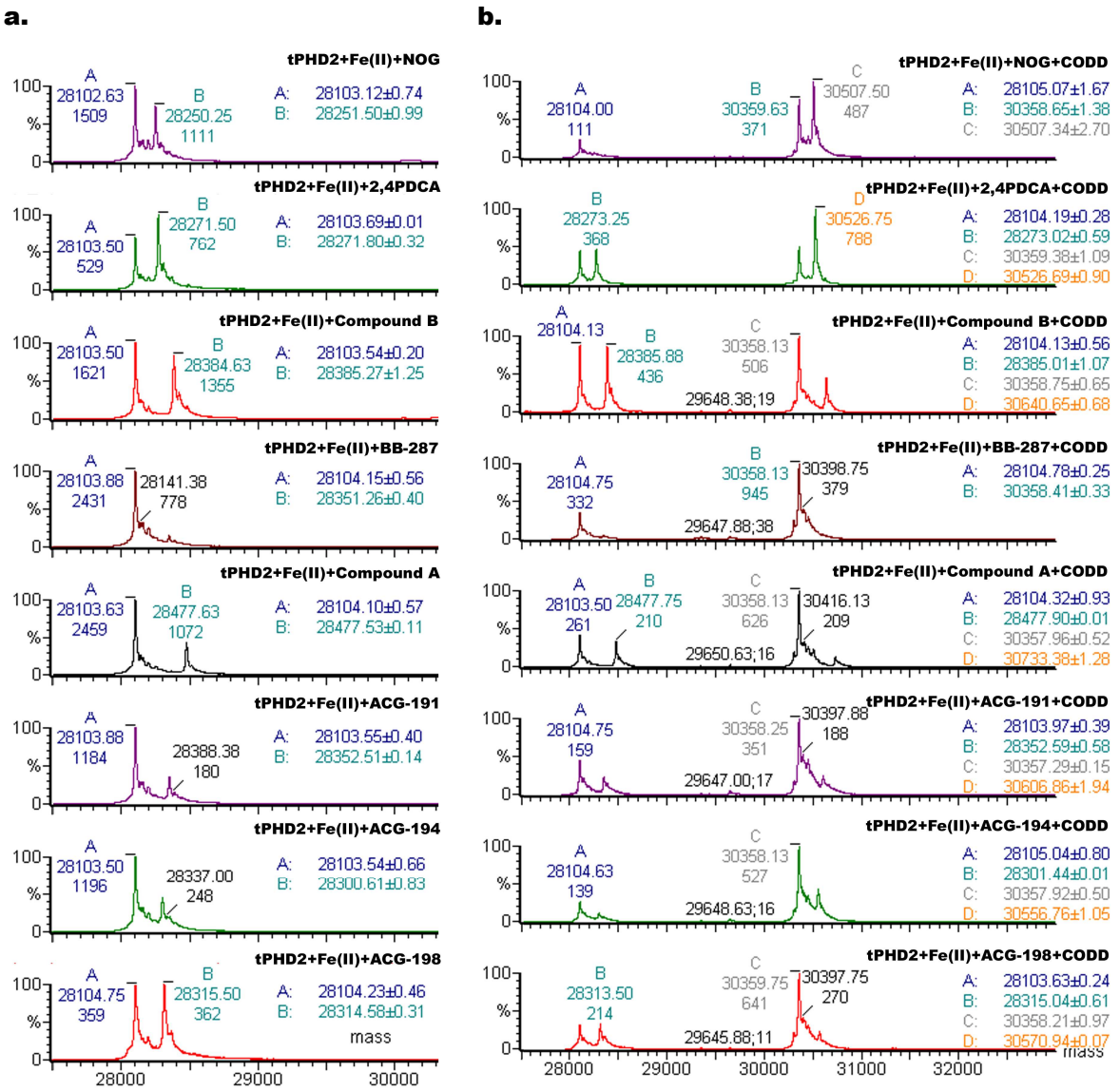


Fig. 2.10 MS analyses under non-denaturing conditions showing 2OG analogue binding in the absence (a) or presence of CODD (b).

2.3.5 Inhibition studies

A range of 2OG analogues have been reported to inhibit prolyl hydroxylase activity including small 'peptide-like' molecules, with IC_{50} values of ~100 μM using crude cell extracts or purified enzymes. Examples include:

a) *N*-oxalyl amino acid derivatives, e.g. NOG and its dimethyl ester, dimethyloxalyl glycine (DMOG)[6];
b) thiol derivatives of 2OG, e.g. 2-mercaptoglutarate, thioester of NOG[17];
c) 2-oxoethyl oxalates with oxygen instead of amide nitrogen in DMOG[17];
d) naturally occurring dealanylalahopcin and synthetically derived cyclic hydroxamates[18];
e) phosphonic acid derivatives of acetic, propionic and butyric acid[17];
f) glutamate derivatives, e.g. *N*-acetyl-(*S*)-glutamate, *N*-benzoyl-glutamate[17];
g) aspartate derivatives, e.g. *N*-acetyl-(*S*)-aspartate[17];
h) dipeptides, e.g. (*S*)-Glu-Gly, *Z*-(*S*)-Glu-Gly, (*S*)-Asp-Gly and β-(*S*)-Asp-Gly[17];
i) TCA cycle intermediates, e.g. fumarate, succinate and isocitrate.[19]

Heterocyclic compounds with two or more fused rings (e.g. quinoline, isoquinoline, phenanthroline) containing carboxylic acid and other lipophilic functionalities were also found to inhibit PHD2.[20-22] These compounds act as 2OG mimetics and form stronger complexes as suggested by MS analyses (Fig. 2.10). Some of these compounds have IC_{50} values as low as <5 μM (using purified PHD2). In the recently reported crystal structures of tPHD2, the enzyme was crystallised with isoquinoline type 2OG analogues (compound A and B). Therefore, after examining their ability to form complexes by mass spectrometry, attempts were made to crystallise tPHD2 (with/without CODD) with potent inhibitors (including unpublished compounds, Appendix 2), using either standard crystallisation conditions (1.5-2.0 M ammonium sulfate, 1-7 % dioxane, 0.1 M MES pH 6.5) or Crystal Screen and Crystal Screen 2 (Hampton Research).

Despite numerous hits, no diffraction-quality crystals were obtained for tPHD2.substrate complexes in these attempts (outlined in Fig. 2.3). Nonetheless, parallel screens using standard crystallisation conditions for tPHD2 enabled crystallisation of the enzyme with a new series of compounds (Chapter 7).

2.4 Truncations and related modifications of tPHD2 for crystallisation

Protein engineering, including truncations/deletions of the flexible N- and/or C-terminal regions, removal of flexible loops or extraction of individual domains from a multidomain protein has become an increasingly popular technique amongst protein crystallographers as a means of obtaining well diffracting protein crystals.[23] Examples of its applications include the HIV-1 gp120 glycoprotein,[24] and the regulatory subunit of cAMP-dependent protein kinase A.[25] B-factor analyses of the tPHD2.Fe(II).A/B structures revealed three major flexible regions, including one internal loop (residues 237-254), a C-terminal region involving residues after helix α4 (residues 404-426) and a short 'hairpin' (residues 345-354) connecting strand IV and V (DSBH numbering) that is seen in almost all 2OG oxygenases.[26] No electron density was observed for the C-terminal part of tPHD2 (residues 404-426) due to disorder; hence these residues were not modelled in the final structure.[27]

Except for the hairpin, the other two regions (C-terminus and β2/β3 loop) were targeted for deletion (Chapter 1, Fig. 1.8) as they might interfere with the crystallisation in the presence of a substrate, and/or worsen the diffraction data quality and refinement statistics due to their flexible nature. With this aim, residues 238–250 of tPHD2 were replaced by an Ala-Ala sequence, i.e. ^{237}DGQLVSQKSDSSKDI^{251} was replaced with ^{237}DAAI^{251}. The Ala-Ala insert was intended to provide a short linker between β2 and β3 while still maintaining the overall fold. A C-terminal truncation (181-406) was also cloned into a pET-28a(+) vector. After an extensive crystallisation trials using 2OG with/without CODD substrate, these variants did not yield suitable crystals. Interestingly, unlike the C-terminal truncation, the β2/β3 loop deletion mutant (tPHD2Δloop) appeared to lose the ability to crystallise even with the bicyclic inhibitors that normally enable crystallisation of tPHD2 in the $P6_3$ crystal form.

2.4.1 Activity and MS analyses of tPHD2Δloop variant

Although deleting the thirteen loop residues (238-250) was found to increase the turnover of 2OG in the presence of substrate peptides (Fig. 2.11), hydroxylation assays demonstrated that the tPHD2Δloop variant catalysed <5% substrate (CODD/NODD) hydroxylation under the standard conditions of 15 minutes at 37°C (Fig. 2.11). Non-denaturing protein MS assays indicated that the tPHD2Δloop variant formed a less stable complex with Fe(II) and 2OG than tPHD2, but that it was capable of binding CODD and

NODD even though hydroxylation was low (5%) (Fig 2.11). Together, these observations indicate that the β2/β3 loop is involved in the coupling of 2OG decarboxylation to substrate hydroxylation.

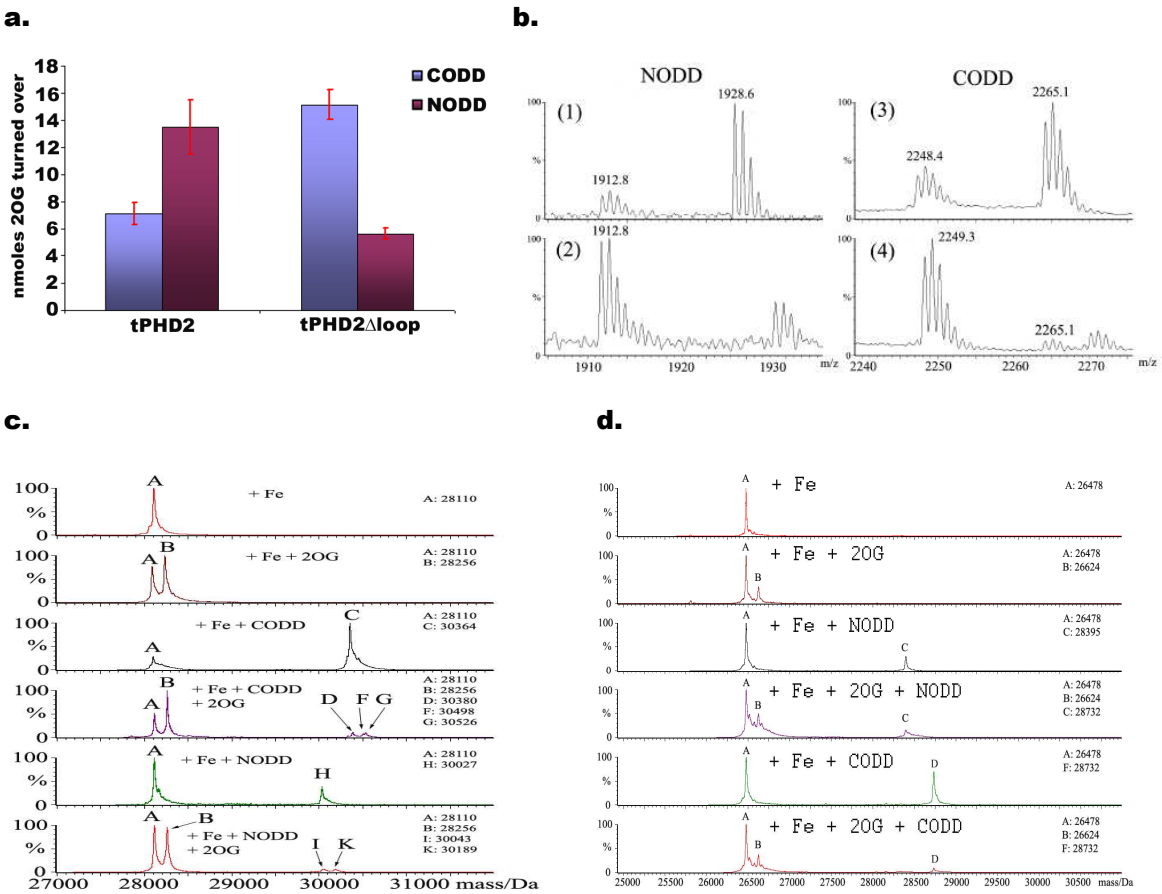


Fig. 2.11 Comparative activity and MS analyses of the tPHD2Δloop mutant with CODD or NODD substrate. (a) 2OG turnover assay results showing significantly higher activity of tPHD2Δloop as compared to the wt tPHD2 (n≥3). (b) MALDI MS data comparing tPHD2 (1 and 3) vs. tPHD2Δloop (2 and 4) mediated NODD and CODD hydroxylation.[28] (c) Non-denaturing mass spectrometry of tPHD2 binding to NODD and CODD peptide substrates. Peaks represent (A) tPHD2.Fe; (B) tPHD2.Fe.2OG; (C) tPHD2.Fe.CODD; (D) tPHD2.Fe.CODD$_{Hyp564}$; (F) tPHD2.Fe.CODD$_{Hyp564}$.succinate; (G) tPHD2.Fe.CODD$_{Hyp564}$.2OG; (H) tPHD2.Fe.NODD; (I) tPHD2.Fe.NODD$_{Hyp402}$; (J) tPHD2.Fe.NODD.2OG and (K) tPHD2.Fe.NODD$_{Hyp402}$.2OG. (d) Non-denaturing mass spectrometry of tPHD2Δloop binding to NODD and CODD peptide substrates. Peaks represent mass shifts corresponding to A) tPHD2.Fe; B) tPHD2.Fe.2OG; C) tPHD2.Fe.NODD (or tPHD2.Fe.2OG.NODD) and D) tPHD2.Fe.CODD (or tPHD2.Fe.2OG.CODD).

2.5 Soaking experiments and structure solution

2.5.1 Soaking with CODD/NODD

Although monomeric in solution, tPHD2 crystallises in a 'head to tail' homotrimeric form (space group *P*6$_3$) in the presence of bicyclic 2OG analogue inhibitors (compounds A and B). Attempts were made to obtain tPHD2.HIFαODD crystal structures by soaking preformed crystals (*P*6$_3$) of tPHD2.Fe(II).A/B complexes with HIFα peptides (CODD and

NODD). Soaking tPHD2.Fe(II).B crystals with hydroxylated Pro564 HIF-1$\alpha_{556\text{-}574(Hyp564)}$ ($CODD_{Hyp564}$) yielded a 2.3Å resolution structure (space group $P6_3$) with C-terminal CODD residues 568-574, ^{568}MDDDFQL^{574} apparent in the electron density maps. The phases were solved by molecular replacement using a search model PDB 2G19; data processing and refinement statistics are presented in Table 2.2. No density was observed for the hydroxylated Pro564 and adjacent residues, because in this crystal form the C-terminal helix (α4) of a symmetry related molecule blocks the active site and prevents the rest of the CODD residues from binding (Figs 2.12 and 2.13). The tPHD2.Fe(II).B.$CODD_{Hyp564}$ (tPHD2.$CODD_{Hyp}$ hereafter) structure with observed density for seven CODD residues is consistent with the previous proposal that PHDs bind first to the C-terminal site of $Pro564_{CODD}$ and the interaction may stimulate subsequent binding to the N-terminal site, possibly depending on the availability of these N-terminal residues.[29]

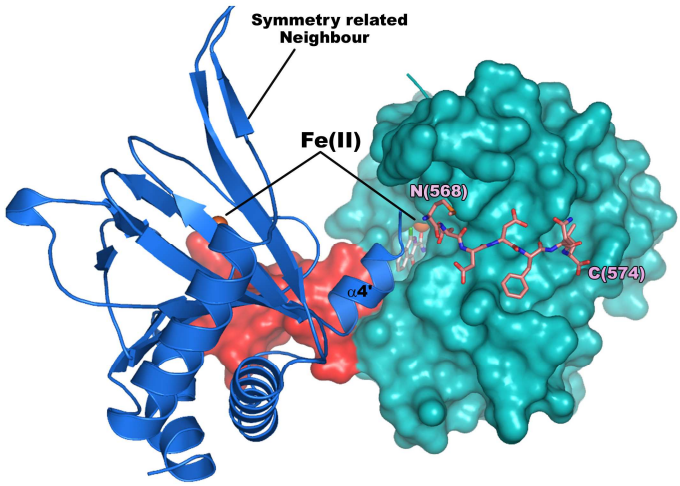


Fig. 2.12 Surface representation of tPHD2.CODDHyp564 complex (cyan); CODD residues 568-574 were shown as salmon sticks. The C-terminal α4 helix of a symmetry related tPHD2 molecule (blue) apparently blocks the entrance to the active site.

Table 2.2 Crystallographic data processing and refinement statistics of the tPHD.Fe(II).B.$CODD_{556\text{-}574(Hyp564)}$ crystal.

Measurement	**tPHD2.Fe(II).B.$CODD_{556\text{-}574(Hyp564)}$**
Data Collection	
Space Group	$P6_3$
Cell dimensions a,b,c (Å)	111.82 111.82 39.62
Resolution (Å)	30.55-2.30 (2.38-2.30)*
No. of unique reflections	12581
Completeness (%)	99.1 (96.2)*
Redundancy	3.2
R_{sym}**	0.068 (0.312)*
Mean I/σ(I)	16.7 (3.3)*

Refinement	
R_{factor}	0.188
R_{free}	0.222
R.m.s. deviation	
Bond length, Å	0.007
Bond angle, °	1.5
No. of atoms	
Protein	1704
Peptide	61
Ligand	19
Water	126
$<B_{factor}>Å^2$	
Protein	32.02
Peptide	54.21
Ligand	13.86 (B)
Metal ion	16.10 (Fe)
Water	39.37

*Highest resolution shell shown in parentheses.

**$R_{sym} = \sum|I-<I>|/\sum I$, where I is the intensity of an individual measurement and $<I>$ is the average intensity from multiple observations.

2.5.2 Soaking with 2OG/NOG

Binding of bicyclic compounds A/B at the active site of tPHD2 involves hydrophobic contacts with Val401 (helix α4) of a neighbouring symmetry related molecule (Fig. 2.13). This interaction, amongst others between the tPHD2 and a symmetry related neighbour, appeared to contribute to the stabilisation of crystallographic trimer formation (Fig. 2.13). Consistent with the structural analyses, MS studies showed that the bicyclic compounds formed weak ternary (tPHD2.metal.inhibitor.CODD) complexes (Section 2.3.5). These observations led to a hypothesis that the presence of these analogues may block binding of CODD at the enzyme active site.

Attempts were therefore made to soak out these analogues from the preformed hexagonal crystals using 10 mM 2OG/NOG in the crystallisation buffer (the mother liquor). After 48 hours at room temperature, there was no visible morphological change in the crystal shape. The tPHD2.Fe(II).B crystals (soaked with 2OG/NOG) were screened using in-house X-ray radiation (Bruker-AXS). Collection of a few frames enabled indexing that revealed the same space group ($P6_3$) and similar cell constants suggesting unsuccessful replacement of the bicyclic compounds. A structure obtained by molecular replacement using PDB ID: 2G19 as the search model and the difference electron density map (F_o-F_c) calculated thereafter revealed the presence of bicyclic compound (B) in the active site.

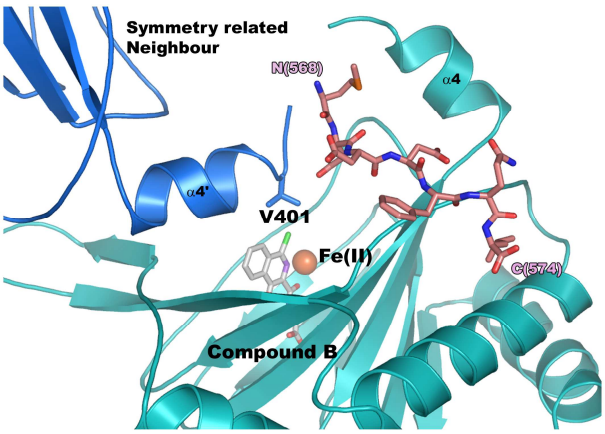


Fig. 2.13 Ribbon representation of tPHD2 (cyan) in complex with CODD (modelled residues 568-574 in salmon sticks) showing hydrophobic contacts between Val401 of a symmetry related tPHD2 (blue) and the inhibitor (compound B in white sticks). Fe(II) is shown as orange sphere.

Apart from Val401 residue of a neighbouring molecule, the compounds A/B, and probably 2OG (based on docking studies) make hydrophobic contacts with residues derived from the core DSBH, notably Met299, Tyr303, Tyr310, Ile327, Leu343, Val376, Ala385 and Trp389 (Fig. 2.14).[27] The binding of 2OG/2OG analogues also involves electrostatic interactions including one with Arg383 of strand VIII (DSBH numbering), which has been observed in some other DSBH containing 2OG oxygenases. Substituting Arg383 with Ala ablates 2OG binding and reduces activity to below the level of detection (<5%).[27]

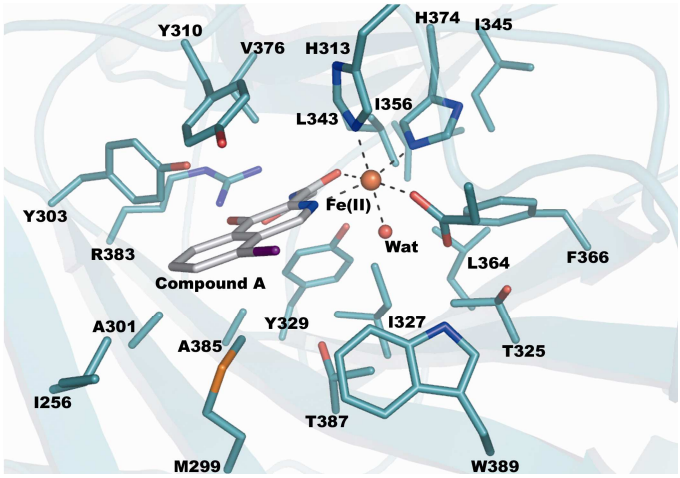


Fig. 2.14 Active site of tPHD2 as observed in the tPHD2.Fe(II).A crystal structure.[27] Compound A maintains hydrophobic contacts with residues derived from the core DSBH.

For competitive soaking experiments, two variants (M299V and V401A) were produced with the aim of loosening but not completely abrogating 2OG/2OG analogue binding (Fig. 2.14). M299V crystallised with compound B and formed the hexagonal crystal form (Fig. 2.15); soaking experiments were conducted using the same procedure as above. Data processing using multiple datasets collected on crystals soaked with 2OG/NOG converged into the crystal structure for this variant revealing the bicyclic analogue at the active site.

After repeatedly unsuccessful soaking attempts, co-crystallisation experiments were then pursued, following strategies aimed at obtaining more stable tPHD2.substrate complexes or hindering formation of the trimeric form.

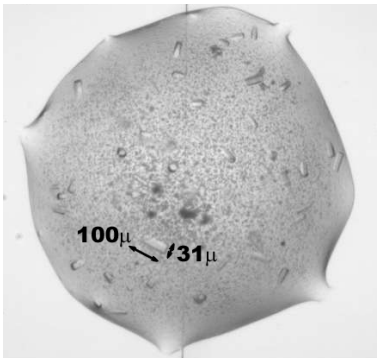
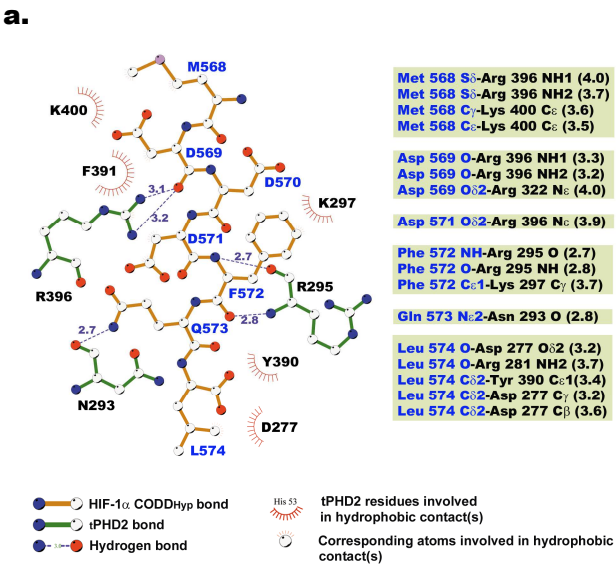


Fig. 2.15 Hexagonal crystals of tPHD2 M299V variant as a complex with Fe(II) and compound B. Concentrations used for M299V variant and compound B were 18.8 mg.ml^{-1} and 1 mM, respectively. Crystallisation conditions (Hampton Research Crystal Screen 2) included polyethylene glycol monomethyl ether 5000 30 % w-v, ammonium sulfate 0.2 M and MES pH 6.5 0.1 M.

2.6 Cross linking between CODD and tPHD2 in an attempt to form a stable complex

In the tPHD2.CODD$_{Hyp}$ structure (Section 2.5.1), several interactions including hydrophobic, ionic and hydrogen bonding occur between tPHD2 and CODD residues (Fig. 2.16a). Leu574$_{CODD}$ is bound in a hydrophobic pocket in tPHD2, formed by the side chains of Ile280, Ile292, Tyr390 and the side chain methylenes of Arg281 and Asp277 (Fig. 2.16b). Although the tPHD2.CODD$_{Hyp}$ structure did not represent a catalytically productive complex, it provided a possible strategy to cross-link the substrate and tPHD2 by introducing cysteine mutations in both within a close proximity. A disulfide bond between tPHD2 and substrate may enable stronger and hopefully irreversible complex formation, suitable for crystallisation. With this aim in mind, a 19-mer (residues 556-574) CODD L574C mutant peptide and recombinant tPHD2 I280C mutant were produced.



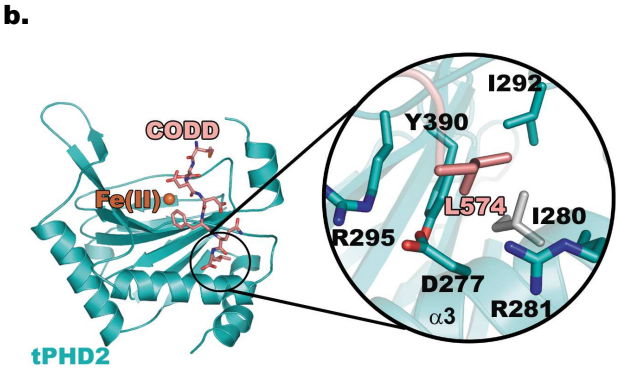
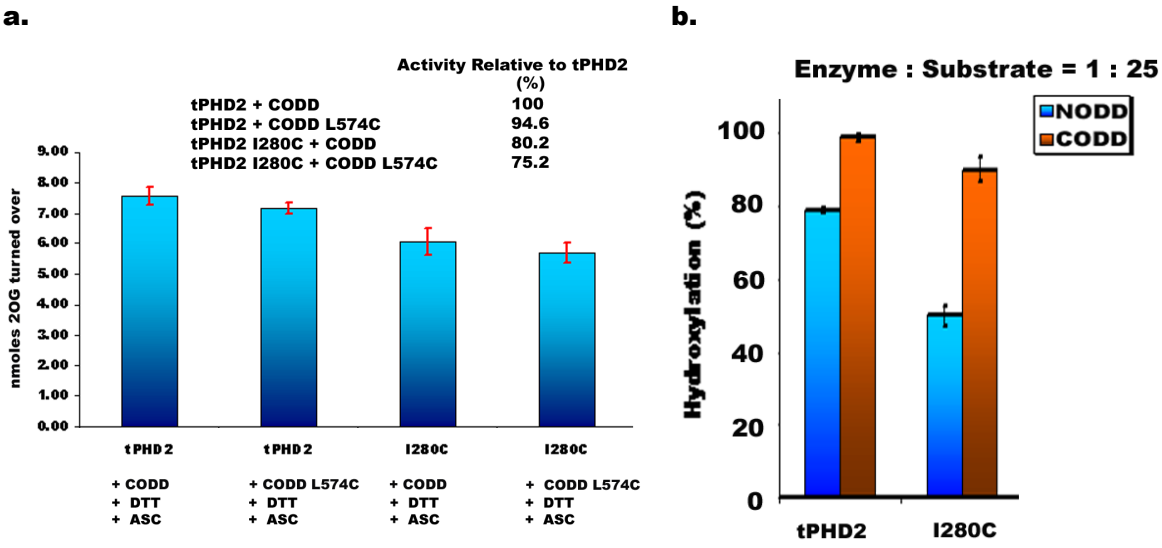


Fig. 2.16 Views from the crystal structure of tPHD2.CODD$_{Hyp}$ showing interactions between CODD$_{Hyp564}$ and tPHD2. (a) Ligplot[30] representations of tPHD2.CODDHyp564 structure showing electrostatic interactions between tPHD2 and CODD residues. For clarity, the observed hydrophobic as well as electrostatic interactions are also presented in the table. (b) highlights mutation site(s) for cross-linking.

2.6.1 Activity analyses of tPHD2 I280C using CODD/CODD L574C as substrates

tPHD2 I280C was found to be less active than the wt tPHD2 using both CODD and NODD substrates (Fig. 2.17). Consistent with the important role for Leu574$_{CODD}$ in prolyl hydroxylation in vivo,[31] CODD L574C caused reduced 2OG turnover when incubated with either wt tPHD2 or the I280C variant. The effect was most significant in case of the tPHD2 I280C and CODD L574C combination suggesting reduced complex formation. However, the possibility remained that once the hydroxylated CODD L574C bound covalently to tPHD2 I280C via disulfide formation, it could block the binding of another CODD L574C molecule and thereby reduce activity. However, SDS-PAGE analysis under non-reducing conditions was not indicative of a covalent complex formation since the complex would be expected to show a band ~ +2 kDa relative to the non-complexed form (Fig. 2.17d).



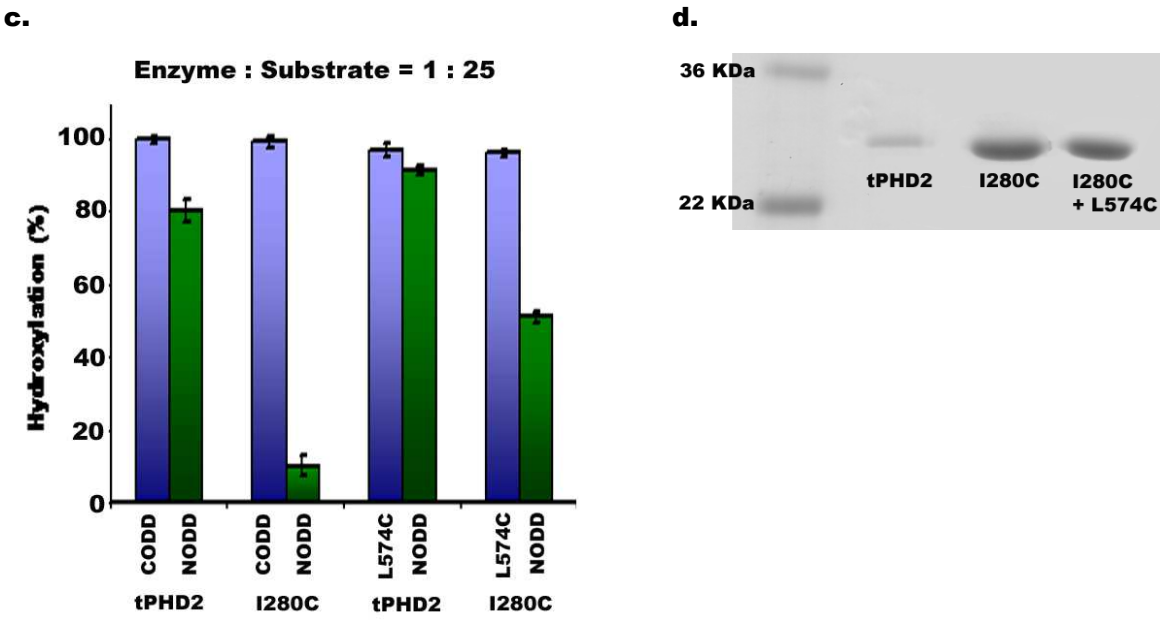


Fig. 2.17 2OG turnover and hydroxylation assay of tPHD2 and I280C tPHD2 variant. (a) 2OG turnover assay results showing tPHD2 I280C variant is less active than wt tPHD2 when incubated with CODD and CODD L574C peptide substrates. Substrate hydroxylation was measured by MALDI-MS using individual substrates (b) and in competition to each other (c). (d) an SDS-PAGE gel showing bands corresponding to tPHD2 (~28 kDa as a control), the tPHD2 I280C variant and a mixture of the tPHD2 I280C variant and CODD L574C.

2.6.2 MS analyses of tPHD2 I280C–CODD L574C complexes under denaturing conditions

Since the objective was to observe formation of a covalent complex between tPHD2 I280C and CODD L574C, MS analyses of complexes were carried out under denaturing conditions using 1% trifluoroacetic acid (TFA). The MS analyses revealed that incubation of tPHD2 I280C and CODD L574C (1 hour, 4°C) did not yield any detectable covalent complex under the analytical conditions. As an alternative, tPHD2 I280C and CODD L574C were also incubated with disulfiram, a compound reported to inhibit aldehyde dehydrogenase by inducing an intramolecular disulfide bond formation.[32,33] Disulfiram interacts with cysteines to form an unstable mixed disulfide that in turn reacts with a neighbouring cysteine residue located within 6–8 Å, to form a stable inter/intramolecular disulfide bond.[32] This protocol yielded ~10% covalent complex between tPHD2 I280C and CODD L574C, as evidenced by the MS analyses under denaturing conditions (Fig. 2.18). Further optimisation would therefore be required to improve the product yield, if this approach were to be viable.

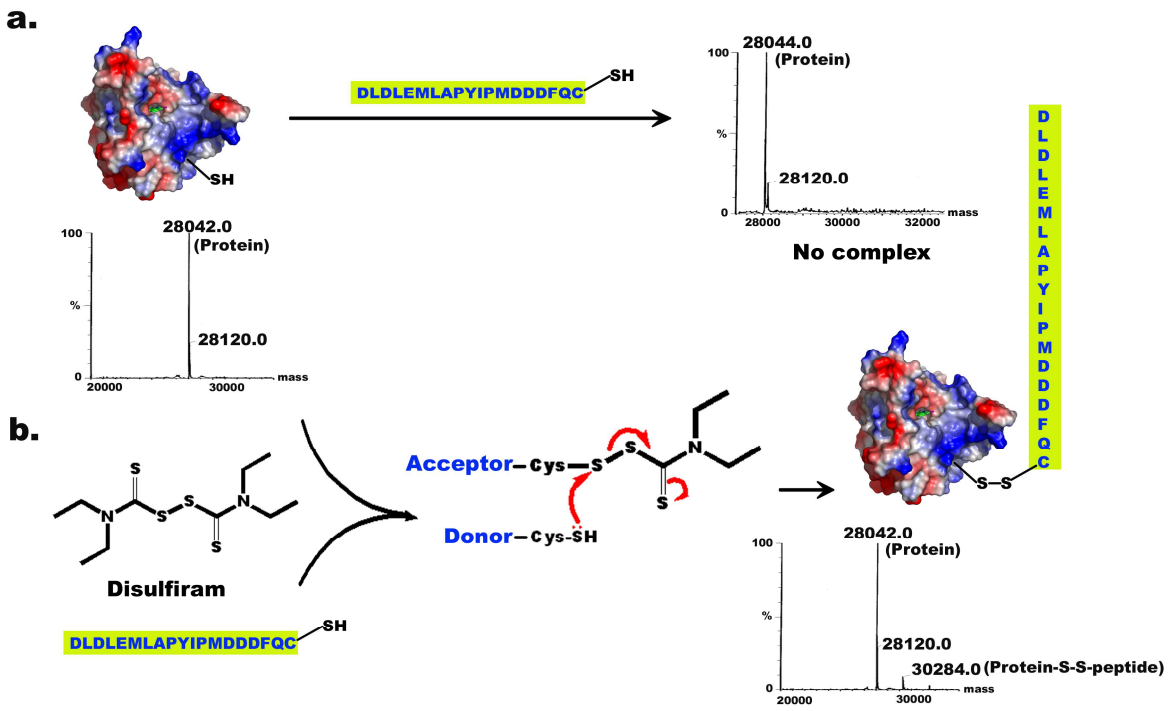


Fig. 2.18 Strategies to obtain a tPHD2 I280C-CODD L574C covalent complex. (a) No covalent complex was formed after incubation of CODD L574C peptide with tPHD2 I280C. (b) A small amount (10%) of a complex was obtained by incubating tPHD2 I280C and CODD L574C with disulfiram. Both tPHD2 I280C and the peptide CODD L574C can act as a donor/acceptor depending on the relative nucleophilicity (pK_a) of the participating cysteines.

2.6.3 Optimisation of complex formation

In a new approach, a selenenylsulfide-mediated conjugation strategy that had been successfully used in a protein glycoconjugation was investigated.[34,35] A cysteine residue from either the enzyme (tPHD2 I280C), or the substrate (CODD L574C), was targeted for conversion into the corresponding (phenylselenenyl)sulfide. The electrophilic nature of the sulfur atom in an S-Se bond can allow efficient nucleophilic substitution (Fig. 2.19). Although the tPHD2 gene encodes seven cysteine residues (201, 208, 266, 283, 302, 323 and 326), it was hoped that the binding of substrate to the specific recognition sites on tPHD2 may minimise non-specific complex formation.

Two parallel strategies were explored in which the cysteine(s) of tPHD2 I280C acted as either electrophile or nucleophile (Fig. 2.19). In the first approach (i.e. activation of tPHD2 I280C), derivatisation of the protein cysteine(s) into an electrophilic (phenylselenenyl)sulfide (PhSe-S) was attempted by treatment with phenylselenenyl bromide (PhSeBr).[35] However, MS analyses revealed a major peak of +624 Da (corresponding to the addition of 4 PhSe) relative to tPHD2 I280C mass (28044.0 Da) along with three smaller peaks representing further complexes (Fig. 2.19). Theoretical

solvent accessibility calculations (Chapter 8) indicated that the apparent addition of four PhSe groups may involve modification of three surface-exposed cysteines in tPHD2 (201, 208 and 266) in addition to Cys280 of the I280C variant. Addition of CODD L574C to this heterogeneous mixture would likely result in the formation of multiple non-specific complexes. Hence, further attempts were not pursued.

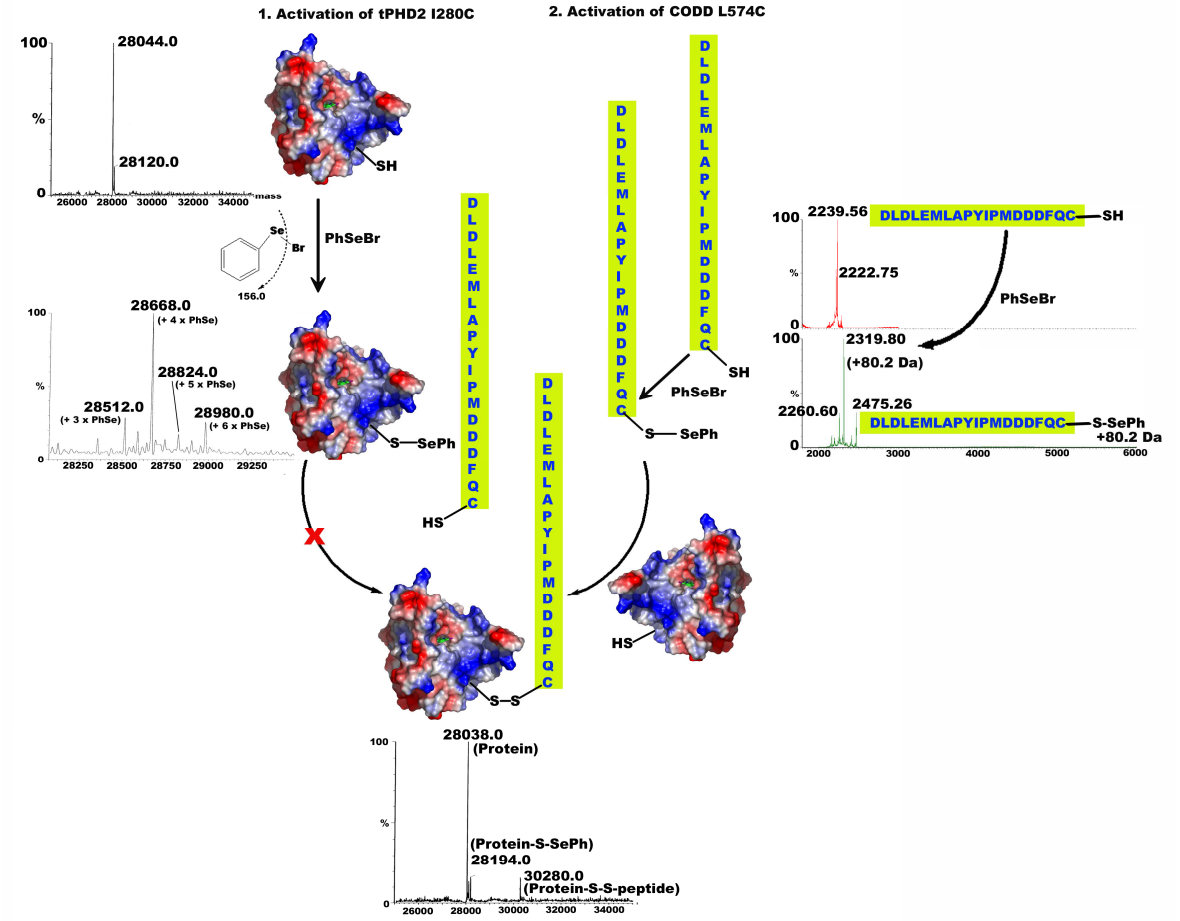


Fig. 2.19 Two potential cysteine conjugation strategies in which the tPHD2 cysteine residues play either an electrophilic role (1) or a nucleophilic role (2). The first approach failed due to non-specific activation of multiple cysteine residues in the protein. The alternative approach, however, resulted in ~20% complex formation. Evidence for PhSe transfer to a free -SH of tPHD2 in the final step came from the observation of an additional peak at 28194.0 Da (+156.0 Da) with an equal intensity of the complex.

In the second approach, the substrate peptide (CODD L574C) thiol was first converted into a selenenylsulfide derivative using PhSeBr in order to couple it to a protein cysteine acting as a nucleophile. MALDI-MS and UV traces after HPLC purification were not suggestive of any peptide dimer formation, as would be expected in case of incomplete derivatisation (Fig. 2.19). Addition of tPHD2 I280C to the activated CODD L574C (CODD L574C-S-SePh) afforded at least 20% complex plus an additional complex with equal intensity corresponding to mass 28194.0 Da (+156.0 Da) (Fig. 2.19). This additional peak most likely represented formation of a (phenylselenenyl)sulfide derivatised tPHD2, which

indicated that the reaction of CODD L574C-S-SePh with the tPHD2 I280C variant may proceed via a mixed disulfide interchange step (Fig. 2.19).

The aforementioned experiments indicated difficulty in forming a stoichiometric covalent complex between tPHD2 I280C and CODD L574C. However, because of the specific recognition of PHD2 for CODD, the tPHD2 I280C and the CODD L574C variants presumably formed a non-covalent complex consistent with the activity studies (Fig. 2.16). Crystallisation screens (Crystal Screen, Crystal Screen 2, PEG-Ion, Ammonium Sulfate Grid Screen, Hampton Research) were setup with the following combinations: tPHD2 I280C, CODD L574C and NOG (1:1:1); tPHD2 I280C, CODD L574C, NOG and disulfiram (1:1:1:1); tPHD2 I280C, CODD L574C-SePh (plus other derivatives, Fig. 2.19) and NOG (1:1:1). However, the experiments did not lead to crystallisation of a tPHD2.CODD complex, possibly due to lack of homogeneity.

During the writing of this thesis, the 2OG dependent DNA repair enzymes AlkB and ABH2 were reported to be crystallised with dsDNA following a chemical cross-linking approach[11] similar to that investigated here. Several cysteine residues were engineered into the AlkB/ABH2 active sites or regions relatively remote from the active site.[11] The mutation sites in both enzymes were chosen based on the previously published AlkB-ssDNA (trinucleotide) complexes.[11,36] In order to avoid any possible non-specific disulfide cross-linking, all the native cysteines in the protein(s) were mutated to serines. The yield of cross-linked complexes were analysed by SDS-PAGE under non-reducing conditions, which identified the optimum site(s) for modification in the protein(s).[11]

2.7 Mutations aimed to block homotrimeric crystal packing

Analyses of crystal packing in the tPHD2.bicyclic inhibitor structure ($P6_3$) indicated that the crystalline homotrimer is stabilised by electrostatic and hydrogen bond interactions including those between residues from C-terminal helix α4 and β2/β3 loop of one monomer and the active site region of a neighbouring subunit (Table 2.3 and Fig. 2.20).[27] As a result, the C-terminal helix of a neighbouring trimer subunit blocks the opening of the active site and, together with the bicyclic inhibitors, prevents the substrate from approaching the metal ion in this crystal form, rationalising the inability to obtain a structure representative of a catalytically productive complex with this crystal form.

Table 2.3 Intermolecular interactions apparently stabilising the homotrimeric crystal form ($P6_3$) in the tPHD2.Fe(II).A crystal structure.

Molecule I residue atom	Residue location	Molecule II (symmetry related) Residue atom/ compound A	Residue location	Distance (Å)
Ser289 Oγ	Short loop after helix α3	Arg252 NH2	β2/β3 loop	2.90
		Arg252 Nε	β2/β3 loop	3.22
		Asp250 Oδ2	β2/β3 loop	2.63
Lys291 Nζ	Short loop after helix α3	Asp237 Oδ1	β2/β3 loop	2.73
Asp394 Oδ2	C-terminal helix α4	Thr236 Oγ1	β2/β3 loop (base)	2.66
Arg398 NH1	C-terminal helix α4	Asp254 Oδ1	β2/β3 loop (base)	2.79
Arg398 NH2	C-terminal helix α4	Asp254 Oδ1	β2/β3 loop (base)	3.06
Val401 Cγ1	C-terminal helix α4	Compound A	Active site	4.21
Val401 Cγ2	C-terminal helix α4	Compound A	Active site	4.23

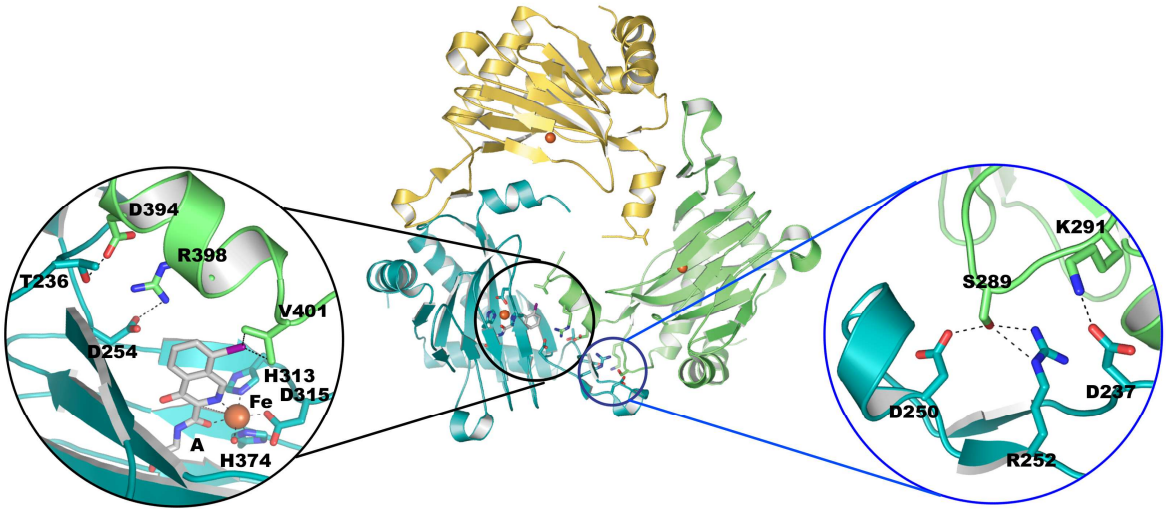


Fig. 2.20 Homotrimeric crystal packing as observed in the tPHD2.Fe(II).A crystal structure. Detailed interactions are given in Table 2.3.

Because the C-terminal α4-helix apparently blocks the active site of a neighbouring symmetry related molecule in the $P6_3$ crystal form, tPHD2 variants (T236A, R252A, D254A, D394A, R398A and S289K) aimed at hindering homotrimer formation were produced, with the aim of obtaining a tPHD2.CODD structure more representative of a catalytically active complex.

2.7.1 Activity and MS analyses of the variants

Both 2OG turnover and hydroxylation assays revealed that, with the exception of R252A and D254A, the above variants did not significantly alter the catalytic activity under standard conditions (Fig. 2.21). Explanations for the catalytic differences of R252A and D254A variants, when compared to wt tPHD2 are given in Chapter 3, Section 3.6.

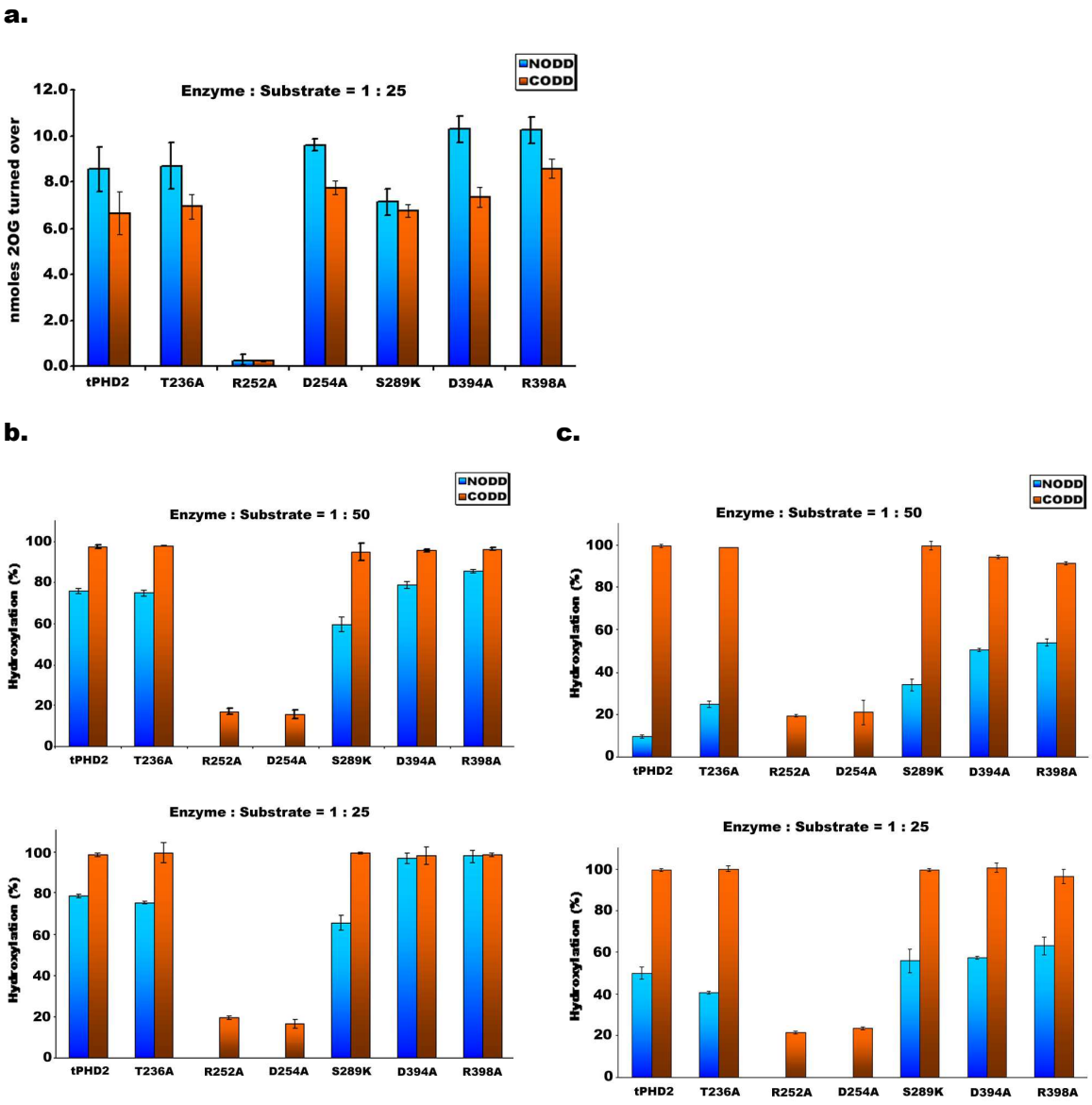


Fig. 2.21 Activity analyses of tPHD2 variants aimed at preventing homotrimeric crystal form. (a) 2OG turnover (b) hydroxylation assay results using NODD and CODD separately (c) competition experiments where both substrates were incubated with the enzyme. In hydroxylation assays, a peak corresponding to minus $Leu574_{CODD}$ was observed occasionally possibly due to peptide instability in solution. This fragment peptide was however, hydroxylated with almost similar efficiency to the 19-mer CODD peptide.

MS analyses showed that the R252A and D254A variants (c and d in Fig. 2.22) have a lower affinity for 2OG than wt tPHD2 and the other tPHD2 variants that were analysed. R252A also has a much lower affinity for CODD compared to wt tPHD2. In contrast, D254A was observed to bind both NODD and CODD equally as wt tPHD2. In all cases, NODD binding was observed to be significantly weaker than CODD binding.

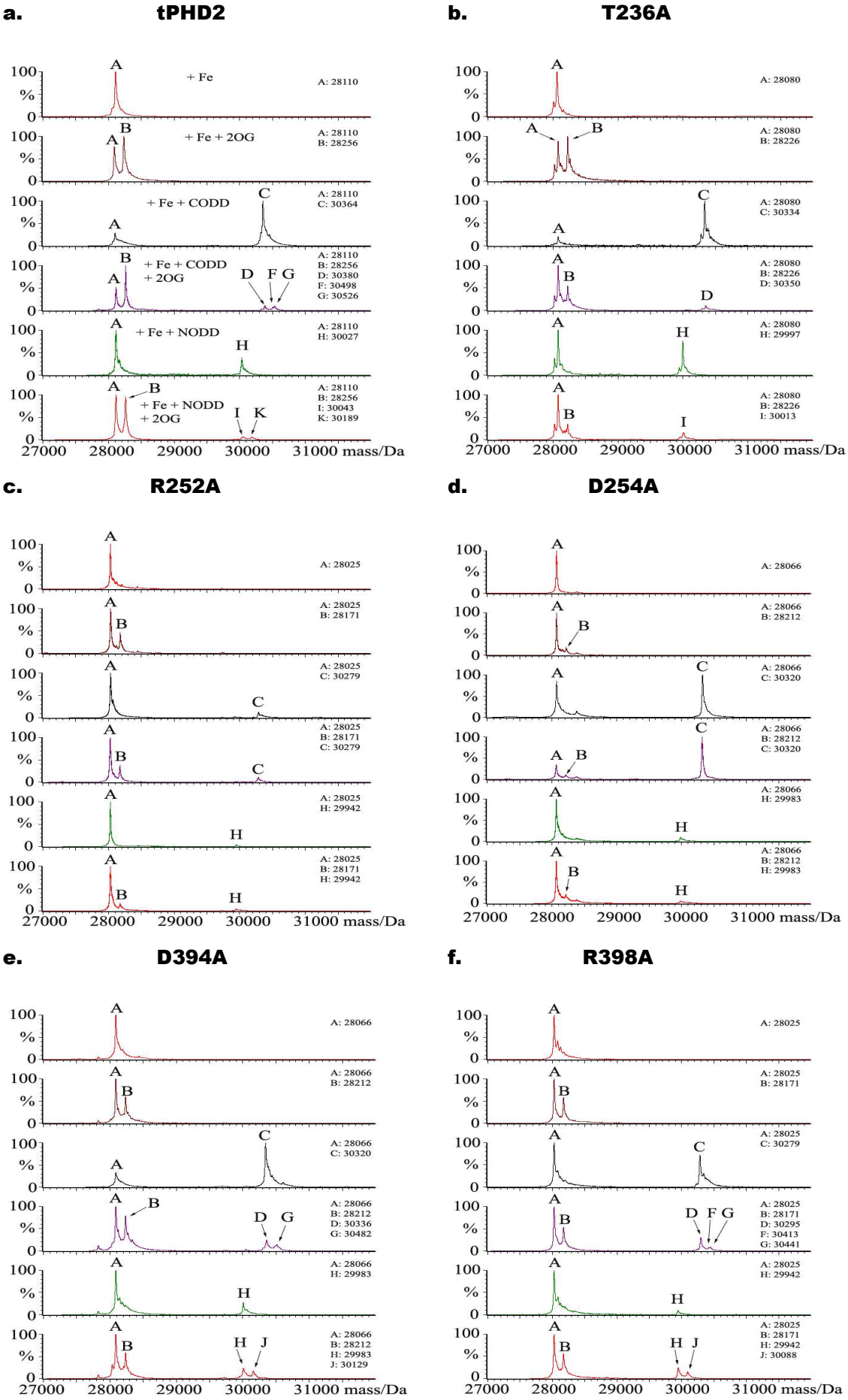

a.
tPHD2
+ Fe
A: 28110
+ Fe + 2OG
A: 28110
B: 28256
+ Fe + CODD
A: 28110
C: 30364
+ Fe + CODD
+ 2OG
A: 28110
B: 28256
D: 30380
F: 30498
G: 30526
+ Fe + NODD
A: 28110
H: 30027
+ Fe + NODD
+ 2OG
A: 28110
B: 28256
I: 30043
K: 30189
b.
T236A
A: 28080
A: 28080
B: 28226
A: 28080
C: 30334
A: 28080
B: 28226
D: 30350
A: 28080
H: 29997
A: 28080
B: 28226
I: 30013
c.
R252A
A: 28025
A: 28025
B: 28171
A: 28025
C: 30279
A: 28025
B: 28171
C: 30279
A: 28025
H: 29942
A: 28025
B: 28171
H: 29942
d.
D254A
A: 28066
A: 28066
B: 28212
A: 28066
C: 30320
A: 28066
B: 28212
C: 30320
A: 28066
H: 29983
A: 28066
B: 28212
H: 29983
e.
D394A
A: 28066
A: 28066
B: 28212
A: 28066
C: 30320
A: 28066
B: 28212
D: 30336
G: 30482
A: 28066
H: 29983
A: 28066
B: 28212
H: 29983
J: 30129
f.
R398A
A: 28025
A: 28025
B: 28171
A: 28025
C: 30279
A: 28025
B: 28171
D: 30295
F: 30413
G: 30441
A: 28025
H: 29942
A: 28025
B: 28171
H: 29942
J: 30088
100
%
0
27000
28000
29000
30000
31000 mass/Da

Fig. 2.22 Comparison of the binding of Fe(II), 2OG and the substrates (CODD and NODD) to wt tPHD2 and variants by non-denaturing MS analyses. (a) wt tPHD2, (b) T236A, (c) R252A, (d) D254A, (e) D394A and (f) R398A. Panel 1 shows each enzyme in the presence of an equimolar amount of Fe(II). Panel 2 shows each enzyme in the presence of equimolar amounts of Fe(II) and 2OG. Interactions with each PHD protein were analysed in the presence of equimolar amounts of Fe(II) and CODD (panel 3) or NODD (panel 5). Panels 4 and 6 show analyses for each enzyme in the presence of equimolar amounts of Fe(II), 2OG and CODD and Fe(II), 2OG and NODD, respectively. Peaks represent (A) tPHD2.Fe; (B) tPHD2.Fe.2OG; (C) tPHD2.Fe.CODD; (D) tPHD2.Fe.$CODD_{Hyp564}$; (E) tPHD2.Fe.CODD.2OG; (F) tPHD2.Fe.$CODD_{Hyp564}$.succinate; (G) tPHD2.Fe.$CODD_{Hyp564}$.2OG; (H) tPHD2.Fe.NODD; (I) tPHD2.Fe.$NODD_{Hyp402}$; (J) tPHD2.Fe.NODD.2OG and (K) tPHD2.Fe.$NODD_{Hyp402}$.2OG.

2.7.2 Crystallisation of variants blocking homotrimer formation

Following the activity and MS analyses, the tPHD2 variants (T236A, R252A, D254A, D394A, R398A and S289K) were set up for crystallisation trials with bicyclic 2OG analogues using the standard tPHD2 crystallisation conditions, as well as with Crystal Screen, Crystal Screen II, PEG/Ion and Ammonium Sulfate Grid Screen (Hampton Research). The lack of crystal hits suggested that all of the tested variants apparently lost the ability to crystallise with bicyclic inhibitors as seen for tPHD2Δloop variant (Section 2.4), indicating inability to produce homotrimers. These variants in complex with CODD/NODD, HIF-2α CODD/NODD were then subjected to high-throughput crystallisation using ligand hits obtained from the thermal scans, notably, 2OG, NOG, succinate and pyridine-2,4-dicarboxylate (Section 2.3.2).

Following extensive crystallisation trials involving various combinations of Mn(II), 2OG/2OG analogues (as above) and tPHD2 variants (T236A, R252A, D254A, D294A, R398A and S289K), crystals of an R398A tPHD2 variant in complex with CODD were obtained under aerobic conditions. The complex (40 $mg \cdot ml^{-1}$ protein /1 mM Mn(II)/1 mM NOG/0.1 mM HIF-1α$CODD_{556-574}$) crystallised in a buffer containing 20% polyethylene glycol 3350 (w/v) and 0.2 M $MgCl_2.6H_2O$ (pH 5.9). Crystallisation plates were monitored hourly for the first day and then 4 hourly for the first week after plate set up (using an automated imaging system). The frequency of monitoring dropped after 3 months to only once a month. As Fig. 2.23 shows, no crystals appeared within 141 days after setting the crystallisation plates, except for possible globular nuclei (Fig. 2.23). After a long incubation of over five months (168 days), crystals with a rhombohedral morphology formed.

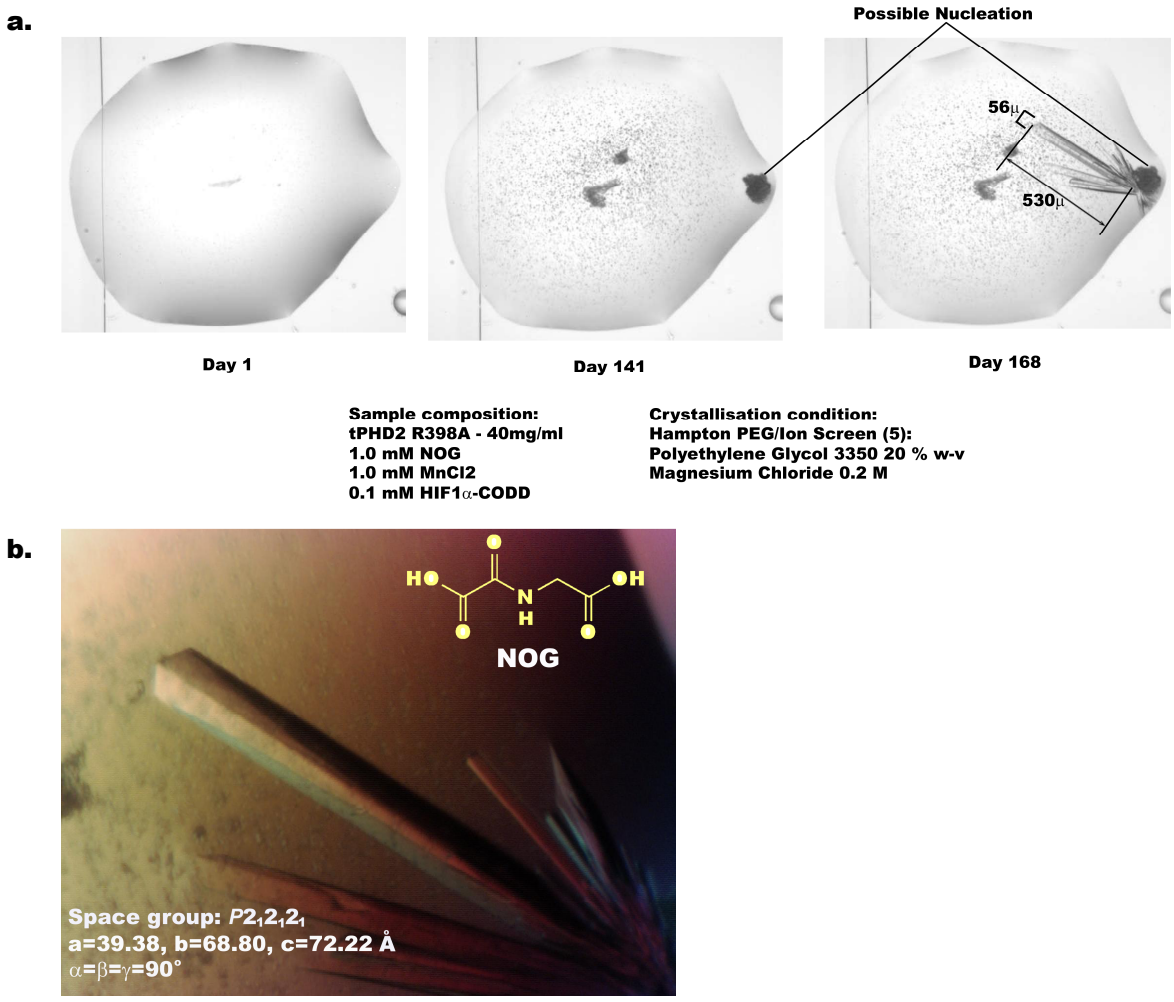


Fig. 2.23 Crystals of the tPHD2R398A.Mn(II).NOG.CODD complex. (a) Monitoring of crystallisation sitting drop containing tPHD2, Mn(II), NOG and CODD after 1, 141 and 168 days. It is possible that nucleation was the rate-limiting step in the crystallisation process. Once formed the effective nuclei apparently led to a more rapid crystal growth. (b) Close-up view of the crystals showing the rhombohedral crystal morphology.

2.8 Data collection and structure solution

With NOG and Mn(II) substituting for 2OG and Fe(II) respectively, the tPHD2 R398A.Mn(II).NOG.HIF-1αCODD$_{556\text{-}574}$ (hereafter tPHD2.CODD) complex crystallised in space group $P2_12_12_1$ (Fig. 2.23 and Table 2.4). Details for freezing of the crystals and data collection are given in Chapter 10. The structure was solved by molecular replacement (using search model: PDB 2G19) and refined to 2.0 Å resolution (R_{factor} = 0.234 and R_{free} = 0.248). A total of 17 CODD residues were refined in the final structure including catalytic Pro564$_{CODD}$ and the conserved LXXLAP motif residues. All the peptide bonds in the structure were in the *trans*-configuration except for that between Lys186$_{PHD2}$ and Pro187$_{PHD2}$; the presence of this *cis*-peptide may be biologically significant given that FKBP38 peptidyl-prolyl *cis*/*trans* isomerase has been identified recently as a protein that interacts with PHD2.[37]

Table 2.4. Crystallographic data processing and refinement statistics of tPHD2.CODD complex structure.

Measurement	**tPHD2R398A.Mn(II).NOG.CODD$_{556-574}$**
Data Collection	
Space Group	$P2_12_12_1$
Cell dimensions a,b,c (Å)	39.38
	68.80
	72.22
Resolution (Å)	34.18-2.00 (2.07-2.00)*
No. of unique reflections	15231
Completeness (%)	99.2 (98.0)*
Redundancy	4.5
R_{sym}**	0.157 (0.419)*
Mean I/σ(I)	7.9 (1.9)*
Refinement	
R_{factor}	0.234
R_{free}	0.248
R.m.s. deviation	
Bond length, Å	0.008
Bond angle, °	1.6
No. of atoms	
Protein	1770
Peptide	145
Ligand	10
Water	155
$<B_{factor}>$Å^2	
Protein	23.81
Peptide	33.74
Ligand	19.59
Metal ion	12.85
Water	34.25

*Highest resolution shell shown in parentheses.
**$R_{sym} = \sum|I-<I>|/\sum I$, where I is the intensity of an individual measurement and $<I>$ is the average intensity from multiple observations.

2.9 Optimisation of initial conditions for tPHD2.CODD crystallisation

After the successful crystallisation of the tPHD2.CODD complex, attempts were made to produce crystals on a larger scale in order to optimise cryo-conditions and to screen crystals for the best possible diffraction quality. Optimisation screens were also set up using Fe(II), NOG and varied concentrations of CODD under anaerobic conditions. This reproduced the same orthorhombic crystal form ($P2_12_12_1$) after a relatively short incubation time of two and half months (Fig. 2.24). However, it was not possible to produce a completely homogeneous sample, most likely due to peptide degradation. The sitting drops with crystals often had 'skins' and precipitation that made harvesting these crystals rather difficult (Fig. 2.24). None of the crystals obtained in this way diffracted to beyond 2.5 Å.

In addition, the mosaicity of these crystals was higher (~1.5) than that for the initial tPHD2.CODD complex.

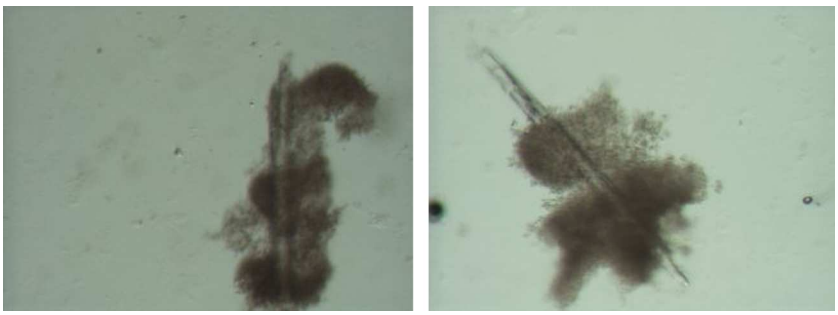
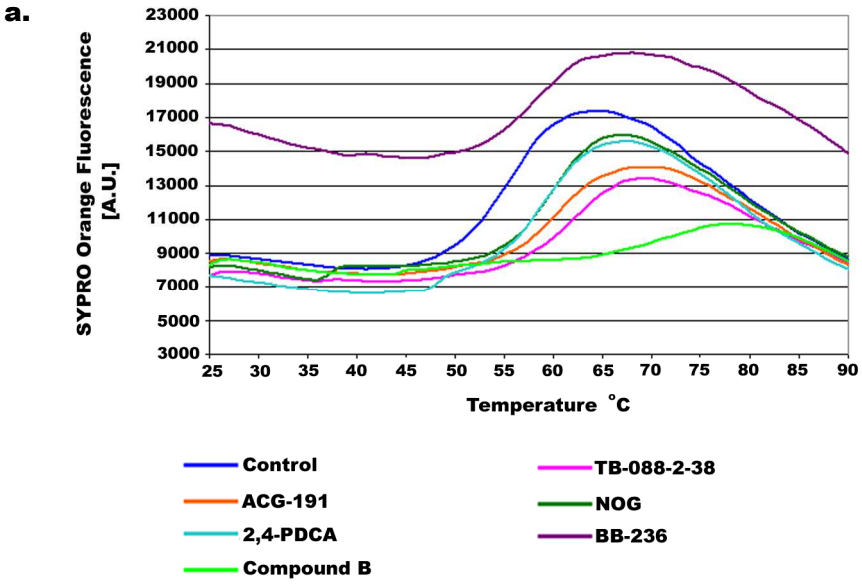


Fig. 2.24 Crystals of tPHD2 R398A.Fe(II).NOG.CODD soaked in a cryo-solution containing crystallisation buffer (20% PEG 3350, 0.2 M $MgCl_2$) and 30% glycerol (v/v). Notably, globular precipitates surrounding the crystals made harvesting extremely difficult.

2.9.1 Thermal stability assays for stabilising the tPHD2 R398A.CODD/NODD complexes

Initial thermal stability assays had been performed on wt tPHD2 with the aim of identifying stabilising agents (Section 2.3.2). However, the subsequently obtained tPHD2.CODD crystals contained a tPHD2 R398A variant. Therefore, further thermal stability assays were performed to screen for ligands that would stabilise tPHD2 R398A.Mn(II).CODD/ NODD complexes with the aim of obtaining higher resolution data. The assay results identified at least two new ligands, namely BB-236 and ACG-194 that had not been observed to stabilise wt tPHD2 significantly (Fig. 2.25).



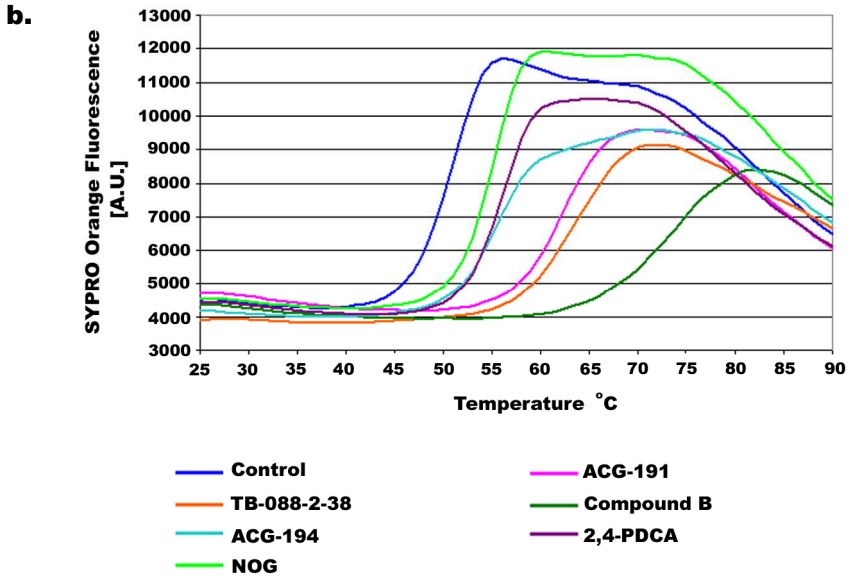


Fig. 2.25 Results of the thermal stability assay using tPHD2 R398A.Mn(II).CODD (a) and tPHD2 R398A.Mn(II).NODD (b) complexes reveal potential stabilising agents. See Fig. 2.3 for chemical structures.

Crystallisation attempts using ACG-194, 2,4-PDCA and NOG (replacing 2OG) and Mn(II) (replacing Fe(II)) were carried out with the aim of obtaining crystals of tPHD2.NODD as well as tPHD2.CODD$_{(Hyp564)}$ (product) complexes. However, these attempts did not lead to diffraction-quality crystals. Further crystallisation attempts using a longer CODD protein (residues 530-698) were also unsuccessful.

2.10 Summary

Considerable efforts were expended on obtaining a crystallisation strategy for a tPHD2-HIF-α complex. Whilst not all of these were successful, some of the strategies employed may be useful for other members of the 2OG oxygenase family, including the mutation approach aimed at blocking 'unwanted' oligomerisation. Although unsuccessful in the current study, the cross-linking strategy to trap labile protein-substrate complexes could be useful to obtain crystal structures of complexes. Future experiments for crystallisation of tPHD2.NODD/HIF-2αCODD complexes could involve variation of native cysteines to serines (to prevent non-specific cross-linking) followed by introducing cysteine variants at different possible interaction points for stable cross-links. In this regard, knowledge from the tPHD2.CODD complex structure (Chapter 3) may be useful.

2.11 References

1. Choi, K.O., Lee, T., Lee, N., Kim, J.H., Yang, E.G., Yoon, J.M., Kim, J.H., Lee, T.G. & Park, H. Inhibition of the catalytic activity of hypoxia-inducible factor-1alpha-prolyl-hydroxylase 2 by a MYND-type zinc finger. *Mol Pharmacol* **68**, 1803-9 (2005).
2. Spadaccini, R., Perrin, H., Bottomley, M.J., Ansieau, S. & Sattler, M. Structure and functional analysis of the MYND domain. *J Mol Biol* **358**, 498-508 (2006).
3. Liu, Y., Chen, W., Gaudet, J., Cheney, M.D., Roudaia, L., Cierpicki, T., Klet, R.C., Hartman, K., Laue, T.M., Speck, N.A. & Bushweller, J.H. Structural basis for recognition of SMRT/N-CoR by the MYND domain and its contribution to AML1/ETO's activity. *Cancer Cell* **11**, 483-97 (2007).
4. McNeill, L.A., Flashman, E., Buck, M.R., Hewitson, K.S., Clifton, I.J., Jeschke, G., Claridge, T.D., Ehrismann, D., Oldham, N.J. & Schofield, C.J. Hypoxia-inducible factor prolyl hydroxylase 2 has a high affinity for ferrous iron and 2-oxoglutarate. *Mol Biosyst* **1**, 321-4 (2005).
5. Schofield, C.J. & Ratcliffe, P.J. Oxygen sensing by HIF hydroxylases. *Nat Rev Mol Cell Biol* **5**, 343-54 (2004).
6. Epstein, A.C., Gleadle, J.M., McNeill, L.A., Hewitson, K.S., O'Rourke, J., Mole, D.R., Mukherji, M., Metzen, E., Wilson, M.I., Dhanda, A., Tian, Y.M., Masson, N., Hamilton, D.L., Jaakkola, P., Barstead, R., Hodgkin, J., Maxwell, P.H., Pugh, C.W., Schofield, C.J. & Ratcliffe, P.J. C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell* **107**, 43-54 (2001).
7. Elkins, J.M., Hewitson, K.S., McNeill, L.A., Seibel, J.F., Schlemminger, I., Pugh, C.W., Ratcliffe, P.J. & Schofield, C.J. Structure of factor-inhibiting hypoxia-inducible factor (HIF) reveals mechanism of oxidative modification of HIF-1 alpha. *J Biol Chem* **278**, 1802-6 (2003).
8. Chen, Z., Zang, J., Kappler, J., Hong, X., Crawford, F., Wang, Q., Lan, F., Jiang, C., Whetstine, J., Dai, S., Hansen, K., Shi, Y. & Zhang, G. Structural basis of the recognition of a methylated histone tail by JMJD2A. *Proc Natl Acad Sci U S A* **104**, 10818-23 (2007).
9. Couture, J.F., Collazo, E., Ortiz-Tello, P.A., Brunzelle, J.S. & Trievel, R.C. Specificity and mechanism of JMJD2A, a trimethyllysine-specific histone demethylase. *Nat Struct Mol Biol* **14**, 689-95 (2007).
10. Ng, S.S., Kavanagh, K.L., McDonough, M.A., Butler, D., Pilka, E.S., Lienard, B.M., Bray, J.E., Savitsky, P., Gileadi, O., von Delft, F., Rose, N.R., Offer, J., Scheinost, J.C., Borowski, T., Sundstrom, M., Schofield, C.J. & Oppermann, U. Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. *Nature* **448**, 87-91 (2007).
11. Yang, C.G., Yi, C., Duguid, E.M., Sullivan, C.T., Jian, X., Rice, P.A. & He, C. Crystal structures of DNA/RNA repair enzymes AlkB and ABH2 bound to dsDNA. *Nature* **452**, 961-5 (2008).
12. Coleman, M.L., McDonough, M.A., Hewitson, K.S., Coles, C., Mecinovic, J., Edelmann, M., Cook, K.M., Cockman, M.E., Lancaster, D.E., Kessler, B.M., Oldham, N.J., Ratcliffe, P.J. & Schofield, C.J. Asparaginyl hydroxylation of the Notch ankyrin repeat domain by factor inhibiting hypoxia-inducible factor. *J Biol Chem* **282**, 24027-38 (2007).
13. Mecinovic, J., Chowdhury, R., Lienard, B.M., Flashman, E., Buck, M.R., Oldham, N.J. & Schofield, C.J. ESI-MS studies on prolyl hydroxylase domain 2 reveal a new metal binding site. *ChemMedChem* **3**, 569-72 (2008).
14. Chayen, N.E. & Saridakis, E. Protein crystallization: from purified protein to diffraction-quality crystal. *Nat Methods* **5**, 147-53 (2008).
15. Ericsson, U.B., Hallberg, B.M., Detitta, G.T., Dekker, N. & Nordlund, P. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal Biochem* **357**, 289-98 (2006).
16. Geerlof, A., Brown, J., Coutard, B., Egloff, M.P., Enguita, F.J., Fogg, M.J., Gilbert, R.J., Groves, M.R., Haouz, A., Nettleship, J.E., Nordlund, P., Owens, R.J., Ruff, M., Sainsbury, S., Svergun, D.I. & Wilmanns, M. The impact of protein characterization in structural proteomics. *Acta Crystallogr D Biol Crystallogr* **62**, 1125-36 (2006).
17. Mole, D.R., Schlemminger, I., McNeill, L.A., Hewitson, K.S., Pugh, C.W., Ratcliffe, P.J. & Schofield, C.J. 2-oxoglutarate analogue inhibitors of HIF prolyl hydroxylase. *Bioorg Med Chem Lett* **13**, 2677-80 (2003).
18. Schlemminger, I., Mole, D.R., McNeill, L.A., Dhanda, A., Hewitson, K.S., Tian, Y.M., Ratcliffe, P.J., Pugh, C.W. & Schofield, C.J. Analogues of dealanylalahopcin are inhibitors of human HIF prolyl hydroxylases. *Bioorg Med Chem Lett* **13**, 1451-4 (2003).
19. Hewitson, K.S., Lienard, B.M., McDonough, M.A., Clifton, I.J., Butler, D., Soares, A.S., Oldham, N.J., McNeill, L.A. & Schofield, C.J. Structural and mechanistic studies on the inhibition of the hypoxia-inducible transcription factor hydroxylases by tricarboxylic acid cycle intermediates. *J Biol Chem* **282**, 3293-301 (2007).

20. Warshakoon, N.C., Wu, S., Boyer, A., Kawamoto, R., Renock, S., Xu, K., Pokross, M., Evdokimov, A.G., Zhou, S., Winter, C., Walter, R. & Mekel, M. Design and synthesis of a series of novel pyrazolopyridines as HIF-1alpha prolyl hydroxylase inhibitors. *Bioorg Med Chem Lett* **16**, 5687-90 (2006).
21. Warshakoon, N.C., Wu, S., Boyer, A., Kawamoto, R., Sheville, J., Bhatt, R.T., Renock, S., Xu, K., Pokross, M., Zhou, S., Walter, R., Mekel, M., Evdokimov, A.G. & East, S. Design and synthesis of substituted pyridine derivatives as HIF-1alpha prolyl hydroxylase inhibitors. *Bioorg Med Chem Lett* **16**, 5616-20 (2006).
22. Warshakoon, N.C., Wu, S., Boyer, A., Kawamoto, R., Sheville, J., Renock, S., Xu, K., Pokross, M., Zhou, S., Winter, C., Walter, R., Mekel, M. & Evdokimov, A.G. Structure-based design, synthesis, and SAR evaluation of a new series of 8-hydroxyquinolines as HIF-1alpha prolyl hydroxylase inhibitors. *Bioorg Med Chem Lett* **16**, 5517-22 (2006).
23. Derewenda, Z.S. The use of recombinant methods and molecular engineering in protein crystallization. *Methods* **34**, 354-63 (2004).
24. Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J. & Hendrickson, W.A. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393**, 648-59 (1998).
25. Su, Y., Dostmann, W.R., Herberg, F.W., Durick, K., Xuong, N.H., Ten Eyck, L., Taylor, S.S. & Varughese, K.I. Regulatory subunit of protein kinase A: structure of deletion mutant with cAMP binding domains. *Science* **269**, 807-13 (1995).
26. Clifton, I.J., McDonough, M.A., Ehrismann, D., Kershaw, N.J., Granatino, N. & Schofield, C.J. Structural studies on 2-oxoglutarate oxygenases and related double-stranded beta-helix fold proteins. *J Inorg Biochem* **100**, 644-69 (2006).
27. McDonough, M.A., Li, V., Flashman, E., Chowdhury, R., Mohr, C., Lienard, B.M., Zondlo, J., Oldham, N.J., Clifton, I.J., Lewis, J., McNeill, L.A., Kurzeja, R.J., Hewitson, K.S., Yang, E., Jordan, S., Syed, R.S. & Schofield, C.J. Cellular oxygen sensing: Crystal structure of hypoxia-inducible factor prolyl hydroxylase (PHD2). *Proc Natl Acad Sci U S A* **103**, 9814-9 (2006).
28. Flashman, E., Bagg, E.A., Chowdhury, R., Mecinovic, J., Loenarz, C., McDonough, M.A., Hewitson, K.S. & Schofield, C.J. Kinetic rationale for selectivity toward N- and C-terminal oxygen-dependent degradation domain substrates mediated by a loop region of hypoxia-inducible factor prolyl hydroxylases. *J Biol Chem* **283**, 3808-15 (2008).
29. Koivunen, P., Hirsila, M., Kivirikko, K.I. & Myllyharju, J. The Length of Peptide Substrates Has a Marked Effect on Hydroxylation by the Hypoxia-inducible Factor Prolyl 4-Hydroxylases. *J Biol Chem* **281**, 28712-20 (2006).
30. Wallace, A.C., Laskowski, R.A. & Thornton, J.M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* **8**, 127-34 (1995).
31. Kageyama, Y., Koshiji, M., To, K.K., Tian, Y.M., Ratcliffe, P.J. & Huang, L.E. Leu-574 of human HIF-1alpha is a molecular determinant of prolyl hydroxylation. *Faseb J* **18**, 1028-30 (2004).
32. Sauna, Z.E., Shukla, S. & Ambudkar, S.V. Disulfiram, an old drug with new potential therapeutic uses for human cancers and fungal infections. *Mol Biosyst* **1**, 127-34 (2005).
33. Shen, M.L., Lipsky, J.J. & Naylor, S. Role of disulfiram in the in vitro inhibition of rat liver mitochondrial aldehyde dehydrogenase. *Biochem Pharmacol* **60**, 947-53 (2000).
34. Bernardes, G.J., Chalker, J.M., Errey, J.C. & Davis, B.G. Facile conversion of cysteine and alkyl cysteines to dehydroalanine on protein surfaces: versatile and switchable access to functionalized proteins. *J Am Chem Soc* **130**, 5052-3 (2008).
35. Gamblin, D.P., Garnier, P., van Kasteren, S., Oldham, N.J., Fairbanks, A.J. & Davis, B.G. Glyco-SeS: selenenylsulfide-mediated protein glycoconjugation--a new strategy in post-translational modification. *Angew Chem Int Ed Engl* **43**, 828-33 (2004).
36. Yu, B., Edstrom, W.C., Benach, J., Hamuro, Y., Weber, P.C., Gibney, B.R. & Hunt, J.F. Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB. *Nature* **439**, 879-84 (2006).
37. Barth, S., Nesper, J., Hasgall, P.A., Wirthner, R., Nytko, K.J., Edlich, F., Katschinski, D.M., Stiehl, D.P., Wenger, R.H. & Camenisch, G. The peptidyl prolyl cis/trans isomerase FKBP38 determines hypoxia-inducible transcription factor prolyl-4-hydroxylase PHD2 protein stability. *Mol Cell Biol* **27**, 3758-68 (2007).

Chapter 3 – Structural analyses of tPHD2.CODD complexes

This chapter describes how PHD2 binds to a HIF-1α CODD substrate in the light of the tPHD2.CODD structure (Chapter 2). Variation of residues that are involved in CODD binding or stabilisation of the tPHD2.CODD complex coupled with activity studies were carried out with the aim of providing insights into substrate recognition by the different human PHD isoforms. For clarity, the secondary structure nomenclature of various crystal forms ($P6_3$ and $P2_12_12_1$) discussed here uses that of the tPHD2.Fe(II).A/B structures ($P6_3$).[1]

3.1 B-Factor analyses of tPHD2 complexes suggest flexible regions

In both the tPHD2 complex structures ($P6_3$ and $P2_12_12_1$), the B-factors for residues corresponding to the β2/β3 loop (residues 237-254, 29.1 Å^2, $P2_12_12_1$) and the hairpin (residues 345-354, 27.1 Å^2, $P2_12_12_1$) regions were observed to be significantly higher than the average protein B-factor (23.8 Å^2, $P2_12_12_1$), consistent with the proposed mobile nature of these residues (Fig. 3.1, discussed below). It is possible that substrate binding stabilises these flexible regions as reflected both by lower B-factors for the residues in the β2/β3 loop and hairpin regions as well as by a lower mean B-factor of tPHD2.CODD (23.8 Å^2, $P2_12_12_1$) compared to that of the tPHD2.Fe(II).A structure (26.7 Å^2, $P6_3$) (Fig. 3.1).

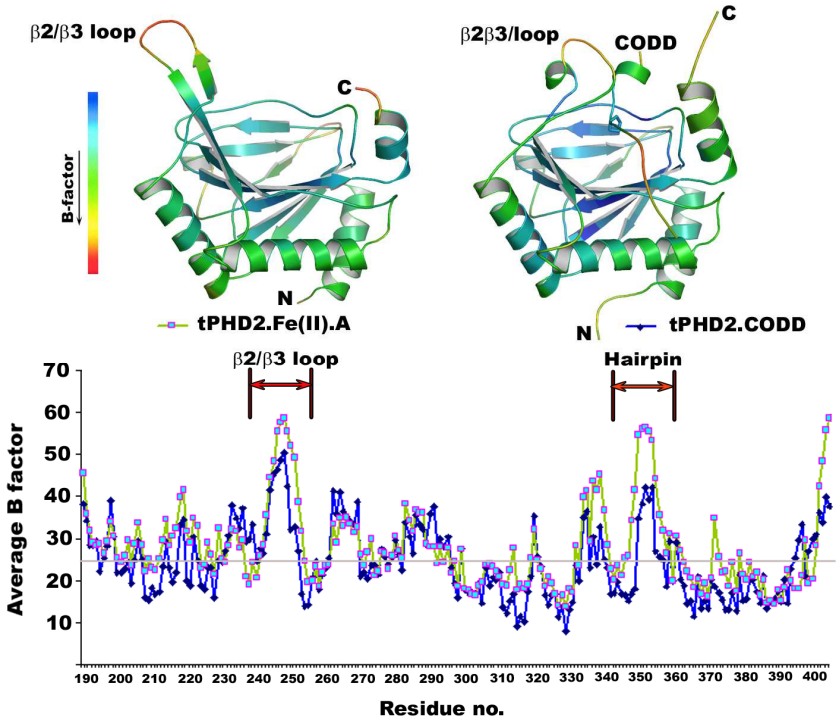


Fig. 3.1 Comparative B-factor (atomic temperature factor) analyses of tPHD2.Fe(II).A (PDB ID: 2G19) and tPHD2.CODD structures.

3.2 General structural features and comparison of different tPHD2 complexes

An overlay of the unrefined model of the tPHD2.Fe(II).NOG.CODD complex obtained after molecular replacement (data not shown) and the tPHD2.CODD structure suggested that the overall structures were very similar. Furthermore, molecular dynamics simulations (MDS) predict that, in solution, the overall structure of the tPHD2.Fe(II).2OG.CODD complex is very similar to that observed for the analogous crystal structure with Mn(II) and NOG (Fig. 3.2).

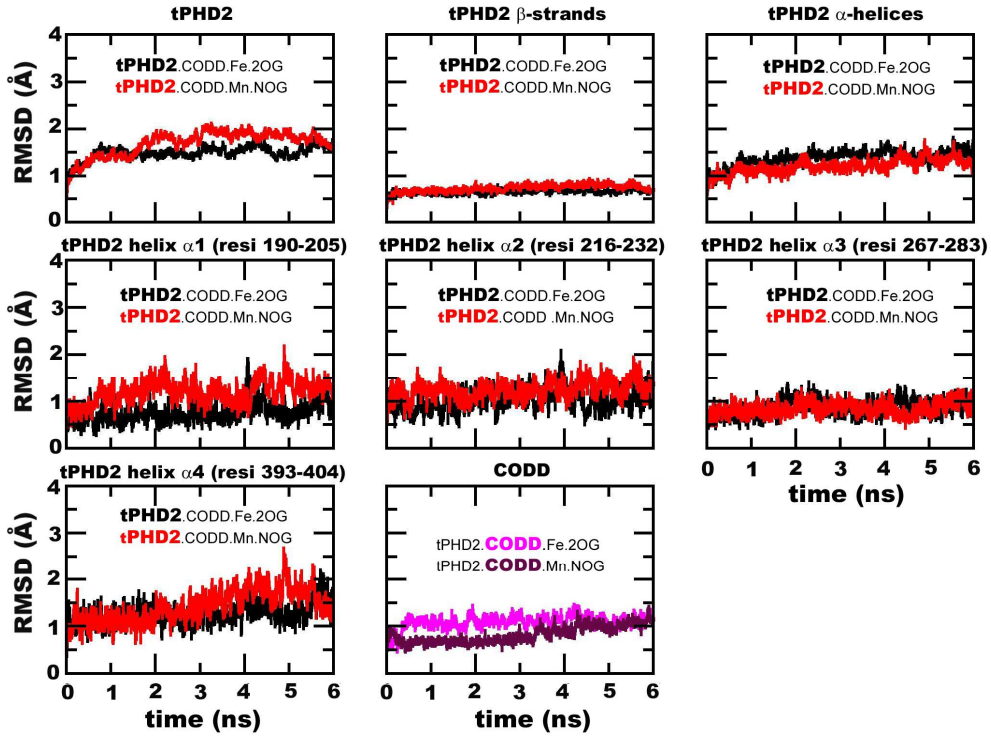


Fig. 3.2 Molecular Dynamics Simulations (MDS) studies on tPHD2.CODD and tPHD2.Fe(II).2OG.CODD modelled complex. The figure shows the Cα root-mean-square distances (RMSD) from the template tPHD2.CODD structure as a function of simulation time for different structural elements, indicated in the figure titles with different coloured labels.

Superimposition of tPHD2.Fe(II).A/B ($P6_3$) and tPHD2.CODD ($P2_12_12_1$) structures revealed similar backbone conformations for the major (β1, β8, β5, β10, β4) and minor (β7, β6, β9, β-II) β-sheets of the DSBH, and three (α1-3) of the four α-helices (Fig. 3.2). Evidence for two mobile regions was apparent; one involving a β2/β3 loop in the $P6_3$ form (residues 237-254; 'β2β3/loop' hereafter) that is folded around the substrate in tPHD2.CODD (Fig. 3.3). RMSD for the 202 equivalent Cα atoms excluding residues 237-254 was 0.710 Å (Fig. 3.3c). The other potentially mobile region involved the C-terminal helix α4 that, compared to the tPHD2.Fe(II).A, rotates relative to the core DSBH by ~15°. The C-terminus of other 2OG oxygenases belonging to the same structural sub-family as the PHDs has been implicated in catalyses.[2,3]

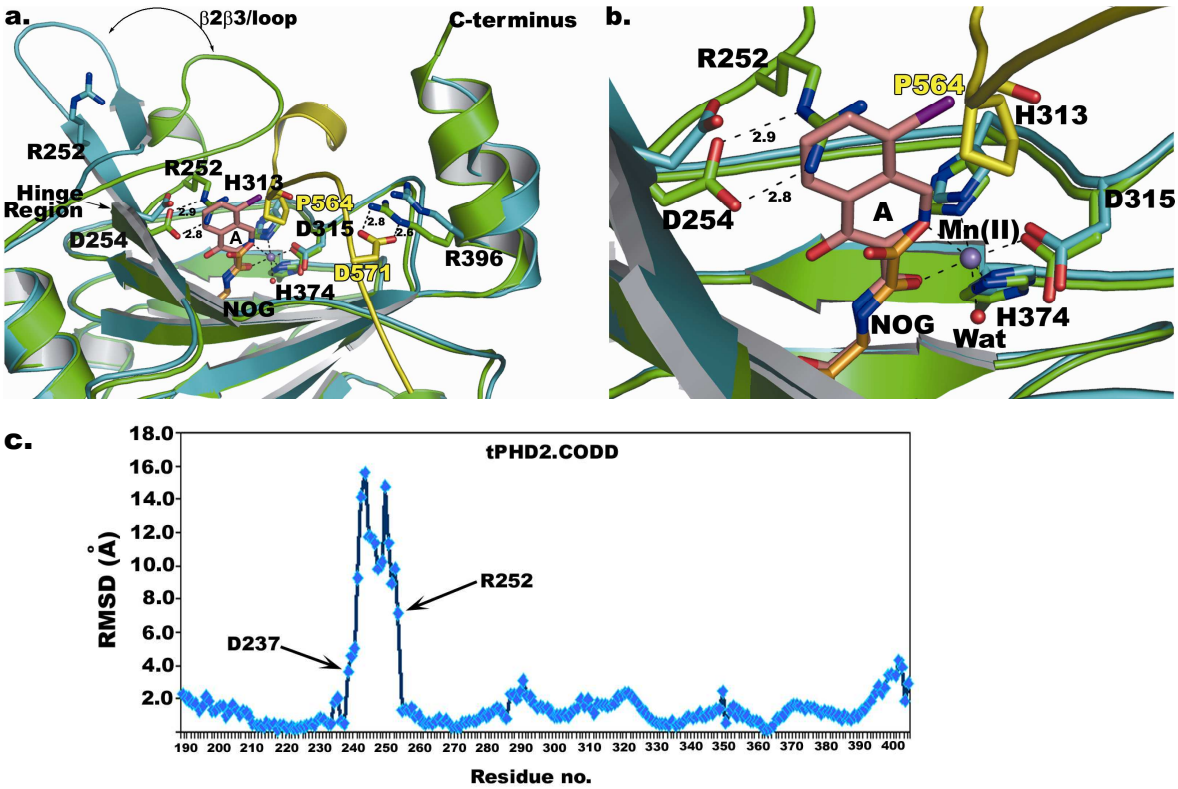


Fig. 3.3 Conformational changes in PHD2 structures. (a) Ribbons representation of the tPHD2.CODD complex (green) superimposed with tPHD2.Fe(II).A (cyan) showing structural changes in the β2β3/loop (PHD2$_{237-254}$) and C-terminal α4-helix. (b) Close-up view of R252:D254 salt-bridge interaction that stabilises 'closed' 'β2β3/loop' conformation in tPHD2.CODD structure (see the text). Bicyclic inhibitors, such as A (salmon) will disrupt the R252:D254 salt-bridge observed in the 'closed' substrate binding conformation ($P2_12_12_1$) and rotate Tyr310 C_β-C_γ (by ~45°) relative to that observed in the $P6_3$ form (Tyr310 side chains are not shown in the figure for clarity). (c) Plot of RMSD for tPHD2 residues in tPHD2.CODD complex from tPHD2.Fe(II).A complex suggesting significant structural changes in the β2/β3 loop (residues 237-254).

The overall tPHD2 folds in both the tPHD2.CODD$_{Hyp}$ and tPHD2.Fe(II).A/B structures are very similar. Comparing the tPHD2.CODD$_{Hyp}$ ($P6_3$) and the tPHD2.CODD ($P2_12_12_1$) structures, the backbone conformations of CODD residues Met568 to Leu574 are similar, but not identical (RMSD = 1.872 Å for residues 568-574), supporting the overall proposed binding mode. There are differences in some side chain orientations of CODD within this sequence, but these may, in part, be due to packing constraints particularly in the $P6_3$ trimeric form (Fig. 3.6). Most strikingly in the tPHD2.CODD structure, electron density for the target proline, Pro564$_{CODD}$, was observed in the $2F_o$-F_c OMIT map (Fig. 3.4) revealing the position of Pro564$_{CODD}$ with respect to the iron centre. No electron density for any of the residues of the conserved LXXLAP motif, including the hydroxylated proline (Pro564$_{CODD}$) was observed in the $P6_3$ form. Together with the results of biochemical analyses (see below) these structural observations reveal that tPHD2.CODD better represents the structure of a catalytically productive complex over the $P6_3$ crystal form.

3.3 Overall mode of CODD binding

Seventeen CODD residues (558-574) were refined in the final tPHD2.CODD structural model, including for the conserved LXXLAP motif (Fig. 3.4). CODD binding apparently involves a combination of main-chain to main-chain/side-chain and side-chain to side-chain interactions that accounts for both substrate recognition and selectivity (Appendix 3). CODD binds to tPHD2 in a well-defined cleft and adopts a loosely extended form including several bends and an N-terminal 3_{10}-helix (residues 559-562) (Fig. 3.3a). From the N-terminus, CODD interacts with tPHD2 β-II, the 'β2β3/loop', β10, the C-terminal helix α4, the loop prior to β4, and α3 (Fig. 3.4a; discussed below).

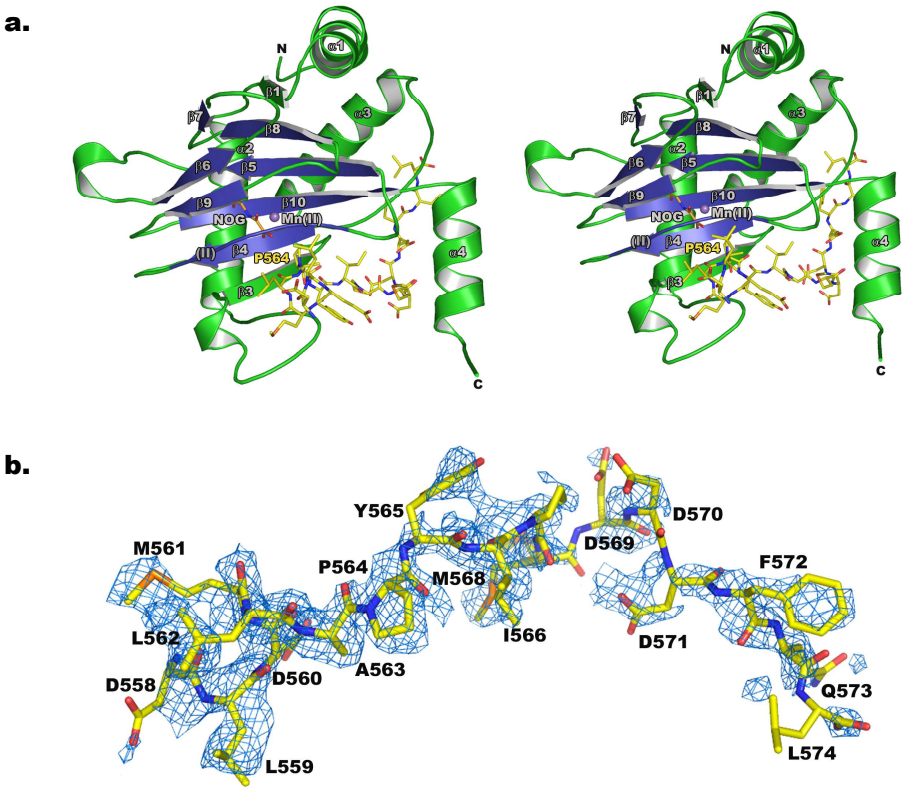


Fig. 3.4 The overall fold of tPHD2 in the tPHD2.CODD complex structure ($P2_12_12_1$ crystal form). (a) Stereoview ribbons representation. tPHD2 fold comprises 4 α-helices and 10 β-strands of which 8 form a DSBH (dark blue)[1]. Three of the four α-helices (α1, α2 and α3) pack along the major β-sheet and stabilise the DSBH. (b) Electron density map (simulated annealing $2F_o$-F_c OMIT contoured to 1σ) for CODD residues 558-574 in the tPHD2.CODD structure. CODD residues were left out for the unweighted OMIT map calculation. The peptide is shown as a stick model where the carbon, nitrogen, oxygen and sulfur atoms are in yellow, blue, red and orange, respectively. (II) indicates the β-strand II of the DSBH.[1]

3.4 Binding of Pro564$_{CODD}$ at the tPHD2 active site

In the tPHD2.CODD structure, Pro564$_{CODD}$ is located at the apex of a loose bend and, like the metal (Mn(II) substituting for Fe(II)) and 2OG is apparently embedded in a deep active site pocket. Hydrogen bonds between the Ala563$_{CODD}$ main chain carbonyl O and Val241$_{PHD2}$ backbone amide NH of the 'β2β3/loop', and the Pro564$_{CODD}$ backbone carbonyl O and Arg322$_{PHD2}$ side chain NH1/NH2 (Fig. 3.11, Appendix 3) appear to position Pro564$_{CODD}$ adjacent to the metal.

3.5 Hydrophobic binding sites in tPHD2, including those involving binding of LXXLAP motif

The hydroxylation sites of both NODD and CODD occur in a highly conserved LXXLAP motif;[4] Leu559$_{CODD}$, Leu562$_{CODD}$, and Ala563$_{CODD}$ of this motif form part of a 3_{10}-helix and make hydrophobic interactions with tPHD2 (Appendix 3 and Fig. 3.5). Overall, the tPHD2.CODD structure reveals five major hydrophobic binding sites for CODD binding (sites I-V in Fig. 3.5). Site I involves Met561$_{CODD}$ that interacts with a non-conserved 'β2β3/loop' residue Lys244$_{PHD2}$ (Arg in PHD1/3) (Fig. 3.5b). The Leu 559$_{CODD}$ (of LXXLAP) binding site (site II) is formed by Val314$_{PHD2}$ and Pro317$_{PHD2}$ side-chains (Fig. 3.5c); and that of Leu562$_{CODD}$ by Ile251$_{PHD2}$ (Ile in PHD1 and Leu in PHD3), Val311$_{PHD2}$ side-chains and the Cβ of the metal coordinating His313$_{PHD2}$ (site III) (Fig. 3.5d).

Away from the LXXLAP motif, the Ile566$_{CODD}$ binding involves hydrophobic contacts with Phe391$_{PHD2}$ and Trp389$_{PHD2}$ (site IV) (Fig. 3.5e). Hydrophobic interactions at the peptide C-terminus occur between the side chain of Leu574$_{CODD}$, which is important for HIF-1α hydroxylation in cells,[5] and a hydrophobic pocket in tPHD2, formed by the side chains of Ile280, Ile292, Tyr390 and Arg281/Arg295 (via methylenes) (site V, Fig. 3.5f). Interactions with both the LXXLAP motif and Leu574$_{CODD}$ are involved in determining NODD/CODD selectivity (see Chapter 4).

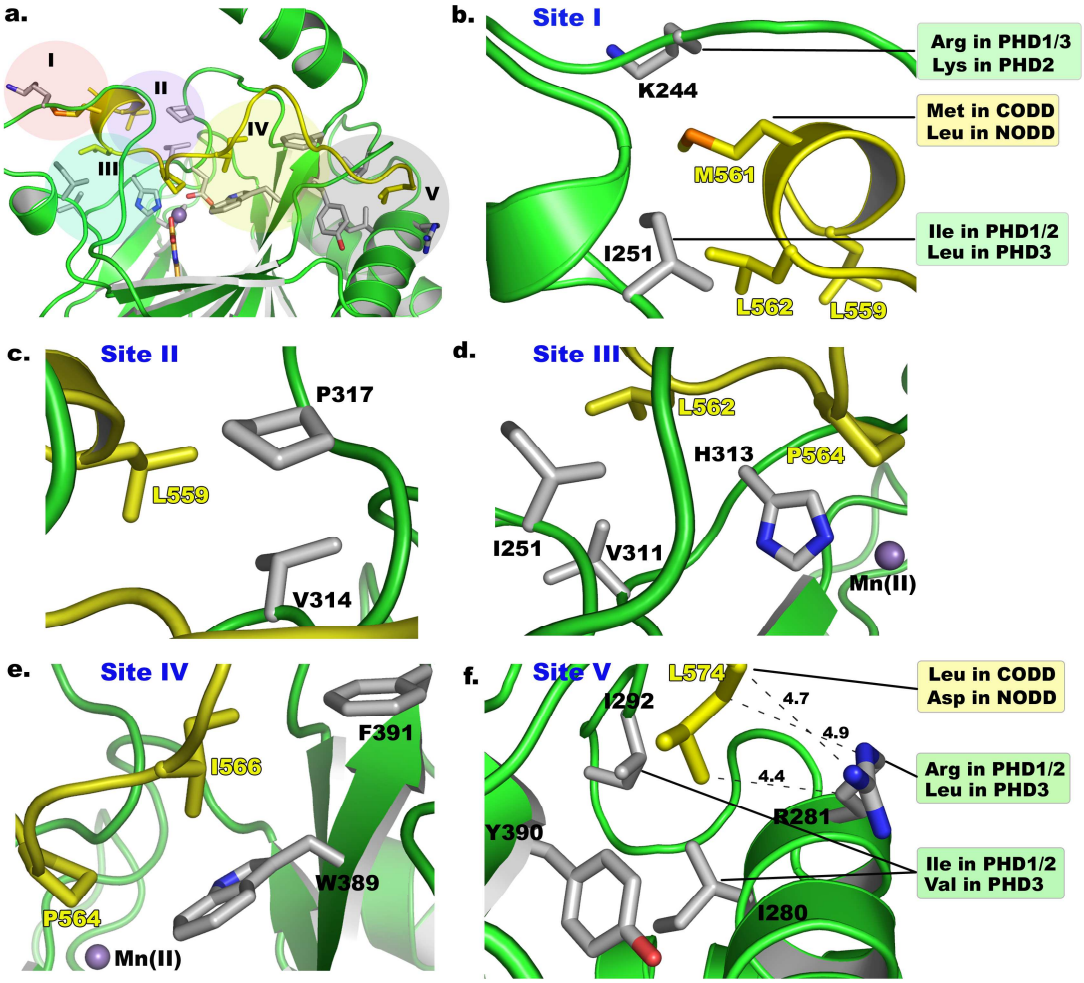


Fig 3.5 Hydrophobic sites on tPHD2 that are involved in CODD binding. (a) View from the tPHD2.CODD complex structure showing the proposed hydrophobic CODD binding sites (I-V); (b) proposed differences between PHD1-3 in binding to $Met561_{CODD}$ and $Leu562_{CODD}$ (site I); (c) the $Leu559_{CODD}$ binding site (site II), is formed by the side-chains of $Val314_{PHD2}$ and $Pro317_{PHD2}$; (d) the $Leu562_{CODD}$ binding site (site III), by the side-chains of $Ile251_{PHD2}$, $Val311_{PHD2}$ and Cα of the metal coordinating $His313_{PHD2}$; (e) the $Ile566_{CODD}$ binding site (site IV) by the side chains of $Trp389_{PHD2}$ and $Phe391_{PHD2}$; (f) the $Leu574_{CODD}$ is bound in a hydrophobic pocket (site V) formed by the tPHD2 Ile280, Arg281, Ile292, and Tyr390 side chains. Mn(II) is represented as a violet sphere.

3.6 Evidence that a flexible 'β2β3/loop' enables 'open' and 'closed' tPHD2 conformations

In order for CODD to bind to tPHD2 the 'β2β3/loop' residues must almost entirely move from the conformation observed in the tPHD2.CODD structure (Fig. 3.6). In the $P6_3$ form the β2/β3 loop residues adopt a finger-like conformation that may enable an 'open' active site, capable of binding CODD. In tPHD2.CODD, the same residues (residues 237-254) form an extended lid and a 3_{10}-helix (residues 248-251) that entirely enclose the conserved LXXLAP motif of CODD (Fig 3.6). This change in conformation suggests a dynamic nature for this loop in substrate-binding. Interactions between the 'β2β3/loop' residues and

CODD include two backbone-to-backbone hydrogen bonds and hydrophobic contacts (Appendix 3). The base of the 'β2β3/loop' appears to be anchored at the active site by a salt-bridge between $Arg252_{PHD2}$ and $Asp254_{PHD2}$ (Fig. 3.3). Binding of bicyclic aromatic inhibitors will disrupt this salt-bridge, consistent with the 'open' state observed in the $P6_3$ crystal form (Fig. 3.3).

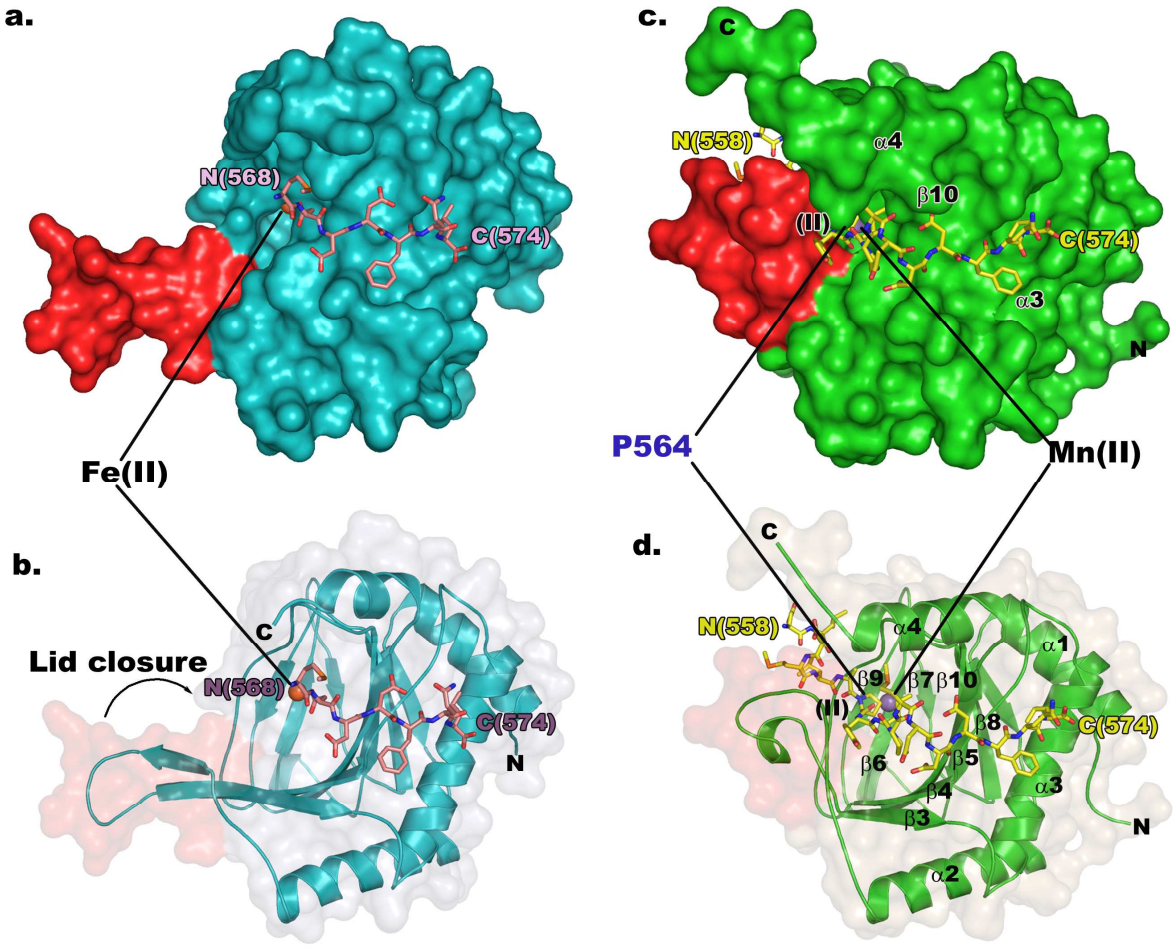


Fig. 3.6 The PHD2 HIFα binding site and comparison of the tPHD2.Mn(II).NOG.CODD (tPHD2.CODD) with tPHD2.Fe(II).B.$CODD_{Hyp564}$ (tPHD2.$CODD_{Hyp564}$) structures. (a) Surface representation of tPHD2.$CODD_{Hyp564}$ complex (cyan); CODD residues 568-574 (salmon sticks) were modelled. The C-terminal α4 helix of a symmetry related tPHD2 molecule (blue) blocks the active site entrance. The 'β2β3/loop' is highlighted in red. (b) Ribbons representation of tPHD2.$CODD_{Hyp564}$ complex (cyan) in the 'open' conformation. (c) Surface representation of tPHD2.CODD complex showing the channel where CODD (yellow sticks) binds. Note the 'β2β3/loop' (red) moves to enclose the substrate. The active site is almost completely buried upon lid-closure that rationalises the high affinity of tPHD2 for Fe(II) and 2OG. (d) Ribbons representation of the tPHD2.CODD complex (green) in 'closed' conformation.

MDS predict that, in solution, the presence of CODD stabilises the 'β2β3/loop' as well as the C-terminal helix α4 (Fig. 3.7).

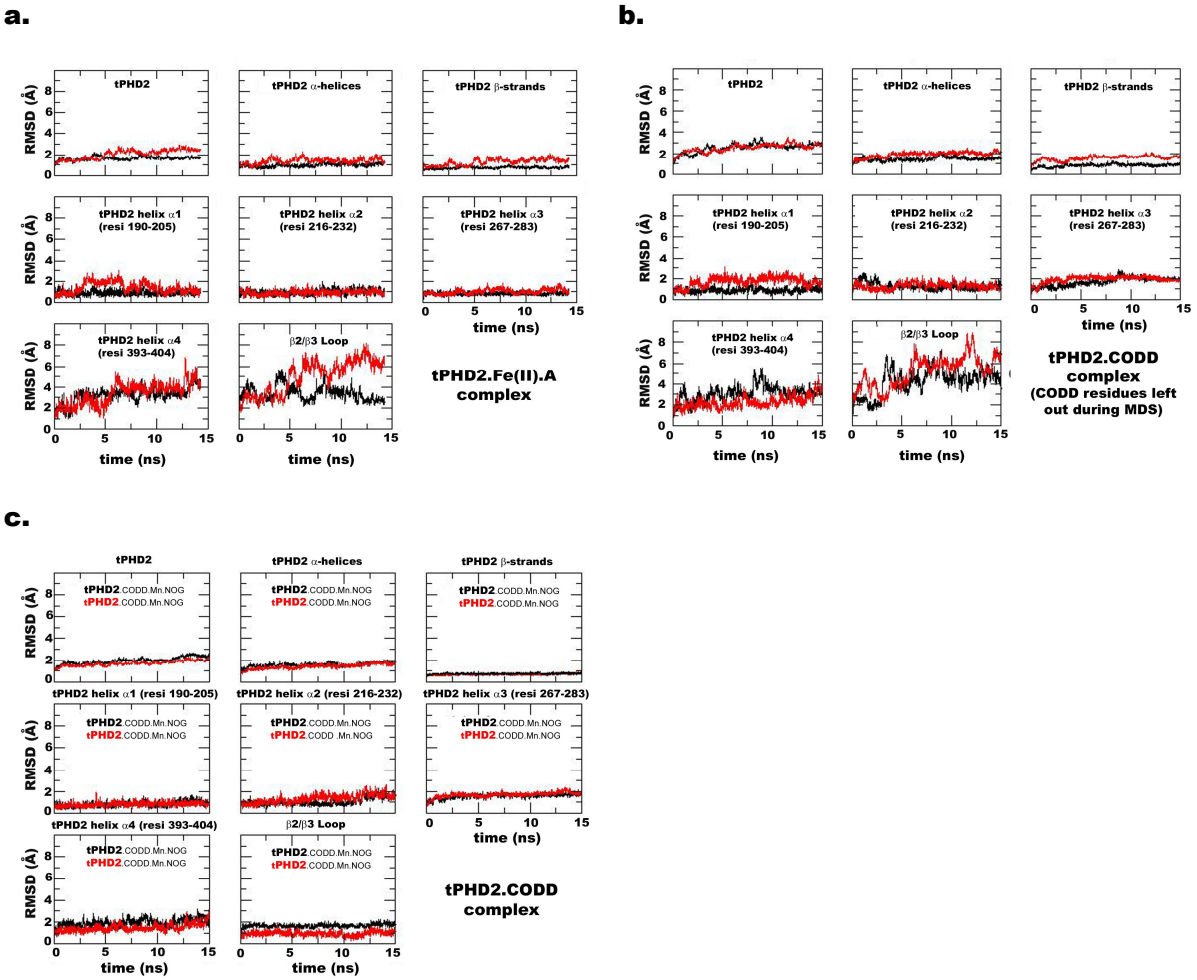


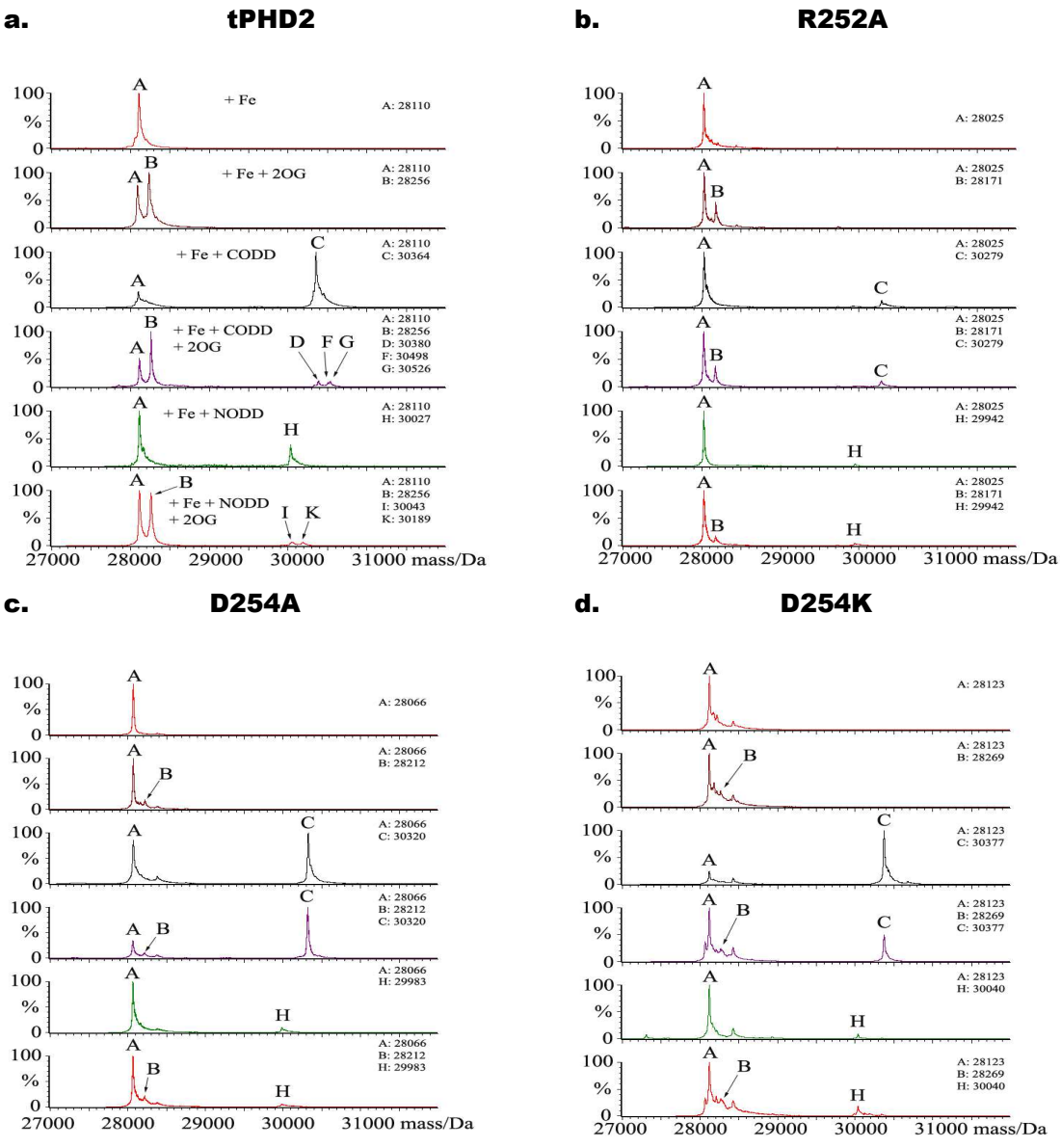
Fig. 3.7 MDS studies on tPHD2.Fe.A (PDB 2G19) and tPHD2.CODD complexes. The figure shows the Cα root-mean-square distances (RMSD) from the template tPHD2.CODD structure as a function of simulation time for different structural elements, indicated in the figure titles.

3.7 Lid closure stabilises the tPHD2.CODD complex for efficient catalysis

tPHD2 variants in which either the 'β2β3/loop' was deleted, or residues apparently important in stabilising the 'enclosed' conformation were substituted (tPHD2 R252A, D254K, D254A), demonstrated a significant reduction in CODD hydroxylation (<25%) compared to wt tPHD2 (Fig 3.10 and Appendix 4). Notably, tPHD2 R252A formed a weaker complex with 2OG or the substrates (CODD/NODD) than wt tPHD2 as determined by MS studies under non-denaturing conditions (Fig. 3.8); the variant was almost unable to turnover 2OG (<5%) or catalyse hydroxylation of CODD (<20%) and NODD (not detected) (Fig. 3.10).

The tPHD2.CODD structure shows that $Arg252_{PHD2}$ is positioned to form a hydrogen bond network involving a water molecule and $Tyr565_{CODD}$ main chain carbonyl O (Tyr565 O-H_2O (2.3 Å); H_2O-Arg252 NH1 (2.7 Å)) (Appendix 3). This water molecule has a very low

B factor (7.2 $Å^2$) indicating that it may not be easily displaced in solution. Replacing Arg252 with alanine will presumably disrupt this network and weaken substrate binding, explaining why the R252A variant forms a significantly less stable complex with CODD/NODD (Fig. 3.8). In contrast, D254K and D254A and the β2/β3 loop deletion variants formed less stable 2OG complexes than tPHD2, yet they displayed increased 2OG turnover, both in the absence and presence of CODD (Figs 3.8 and 3.10). Collectively, these results reveal that the 'β2β3/loop' is important in enabling PHD2 to form stable complexes with Fe(II) and 2OG, as well as coordinating the oxidation of 2OG with prolyl hydroxylation. Both of these properties are probably required for PHD2 to act as an efficient oxygen sensor.



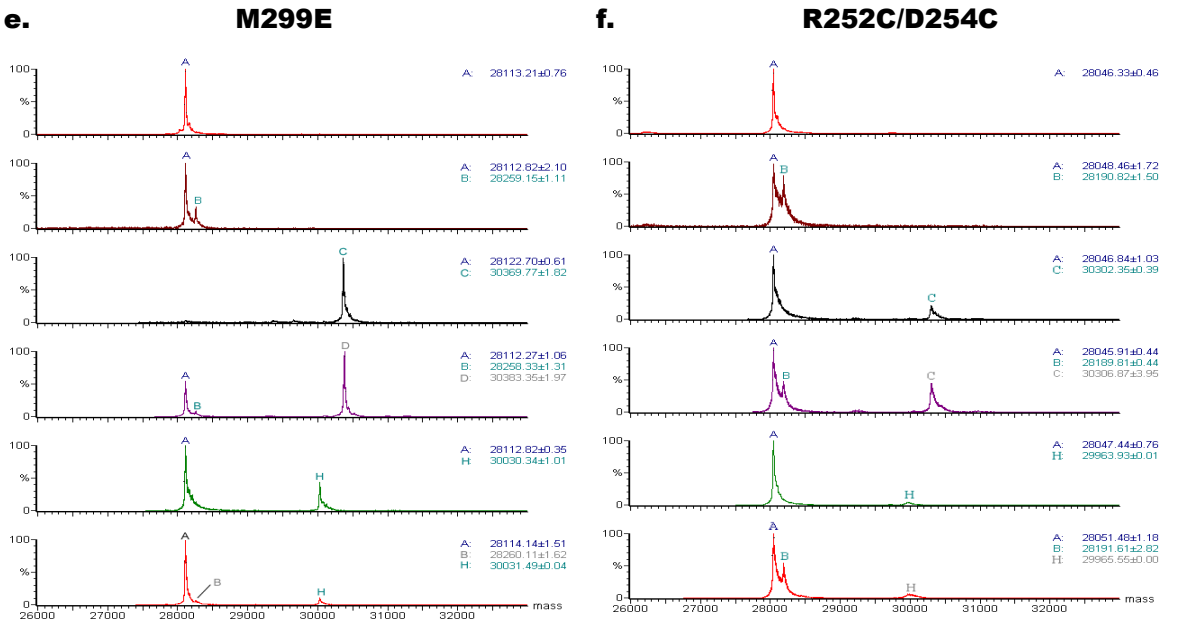


Fig. 3.8 Non-denaturing MS analyses of wt and point mutant tPHD2 binding to NODD and CODD peptide substrates. (a) wt tPHD2, (b) R252A, (c) D254A, (d) D254K, (e) M299E and (f) R252C/D254C. Peaks represent (A) tPHD2.Fe; (B) tPHD2.Fe.2OG; (C) tPHD2.Fe.CODD; (D) tPHD2.Fe.CODD$_{Hyp564}$; (E) tPHD2.Fe.CODD.2OG; (F) tPHD2.Fe.CODD$_{Hyp564}$.succinate; (G) tPHD2.Fe.CODD$_{Hyp564}$.2OG; (H) tPHD2.Fe.NODD; (I) tPHD2.Fe.NODD$_{Hyp402}$; (J) tPHD2.Fe.NODD.2OG and (K) tPHD2.Fe.NODD$_{Hyp402}$.2OG. Notably, R252A (d), D254A (e) and D254K (f) have a lower affinity for 2OG than the wt tPHD2 and the other tPHD2 mutants. The data further show that R252A has a much lower affinity for CODD than wt tPHD2 and other mutants, whereas D254A/K binds both NODD and CODD equally as wt tPHD2. The greater ability to turnover 2OG and a weaker binding of D254A and D254K possibly reflects an 'open' loop conformation.

Consistent with the MS data, Arg252$_{PHD2}$ is positioned to interact, although possibly weakly, with the non-coordinating carboxyl O (Arg252 NH2 - O1, 3.7 Å) of NOG and possibly 2OG. A similar interaction is also observed in almost all structurally related 2OG oxygenases usually involving an arginine/lysine or an asparagine residue. Support for this observation came from the MDS data (Fig. 3.9); during the simulations, Arg252$_{PHD2}$ was often observed to move towards the 2OG to form more stable ionic interactions (data not shown).

In addition to substitution of Arg252$_{PHD2}$, variants of nearby residues were also produced with the aim of disrupting the interaction between Arg252$_{PHD2}$ and 2OG. Met299$_{PHD2}$ was targeted for this purpose because it is in close proximity to Arg252 (Fig. 3.9) and the side chain of Met299$_{PHD2}$ is in a 'head-to-head' orientation to that of Arg252$_{PHD2}$, i.e. the Met299 Sδ and Arg252 NH2/NH1 are facing each other. Non-denaturing MS analyses show that the M299E variant weakly binds to 2OG although CODD binding was comparable to that of wt tPHD2 (Fig. 3.8). Consistent with these observations, the M299E variant gave significantly reduced activity in both 2OG turnover (~10% of wt tPHD2) and

hydroxylation (< 50% of wt tPHD2) assays (Fig. 3.10). $Met299_{PHD2}$ is located at the active site entrance; because of the head-to-head side chain orientation with $Arg252_{PHD2}$, the Glu299 of M299E can form an additional ion-pairing with $Arg252_{PHD2}$. This may hinder 2OG binding by narrowing the active site entrance for 2OG (Fig. 3.9); alternatively reactions of 2OG may be slowed.

MDS of a tPHD2 M299E variant model, based on the tPHD2.CODD structure, predicted formation of a 'complex salt-bridge'[6] between Glu199 Oε1 of M299E variant and Arg252 NH2 and between Asp254 and Arg252 (Fig. 3.9). The energetic contribution of such complex salt bridges has been reported to play a significant role in protein stabilisation.[7]

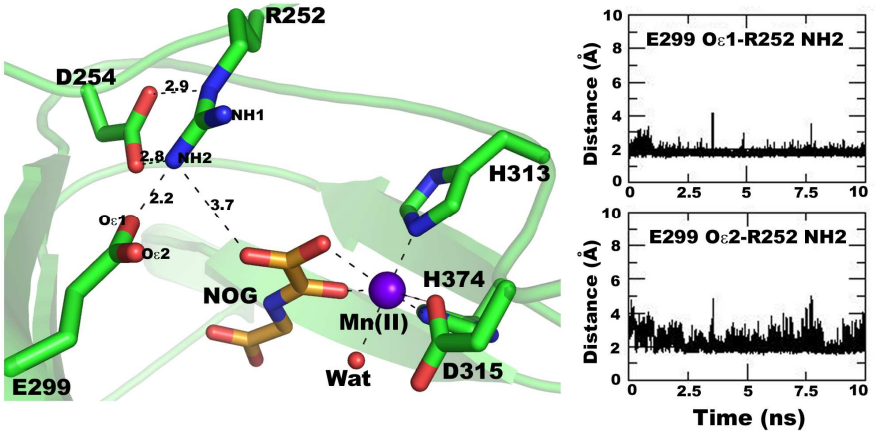


Fig. 3.9 Active site of the tPHD2 M299E model showing surrounding residues that are proposed to stabilise an 'open' 'β2β3/loop' conformation. MDS predicted a hydrogen bond between Glu199 Oε1 of M299E variant and Arg252 NH2. See text for discussion.

3.8 Further strategies to obtain a tPHD2 structure with the 'closed' form of the 'β2β3/loop'

Although the β2/β3 loop conformation observed in the $P6_3$ crystal form ('β2β3/loop' in tPHD2.CODD) may enable an 'open' active site that reflects a CODD binding conformation, it may also be an artifact of the homotrimeric crystal packing. This possibility raises the question of whether or not the exact lid conformation(s) of the resting tPHD2, i.e. in the CODD unbound states, are represented by the $P6_3$ crystal form structures. With the M299E variant, the combined binding energy gained via two salt-bridges with Arg252 (to Asp254 and to Glu299 in the variant) may enable this variant to preferentially stabilise the 'closed' lid conformation. This suggested a possible strategy for obtaining a resting tPHD2 structure, i.e. a substrate unbound form without using the bicyclic inhibitors; this had not been possible with wt tPHD2 despite numerous attempts.

Another possible way of stabilising the 'closed' lid conformation would be to crosslink the residues at the 'β2β3/loop' hinge section (Arg252 and Asp254) via a disulfide bond. With this aim, a R252C/D254C variant was produced.

Non-denaturing MS studies on R252C/D254C (Fig. 3.8) revealed that this variant displayed a significantly reduced ability to bind the substrate (CODD/NODD); the observation was reflected in the activity studies, where CODD hydroxylation was <20%, and NODD hydroxylation was below the limits of detection. However, attempts to crystallise M299E and R252C/D254C variants using variable amounts of 2OG/NOG, with or without CODD substrate were unsuccessful.

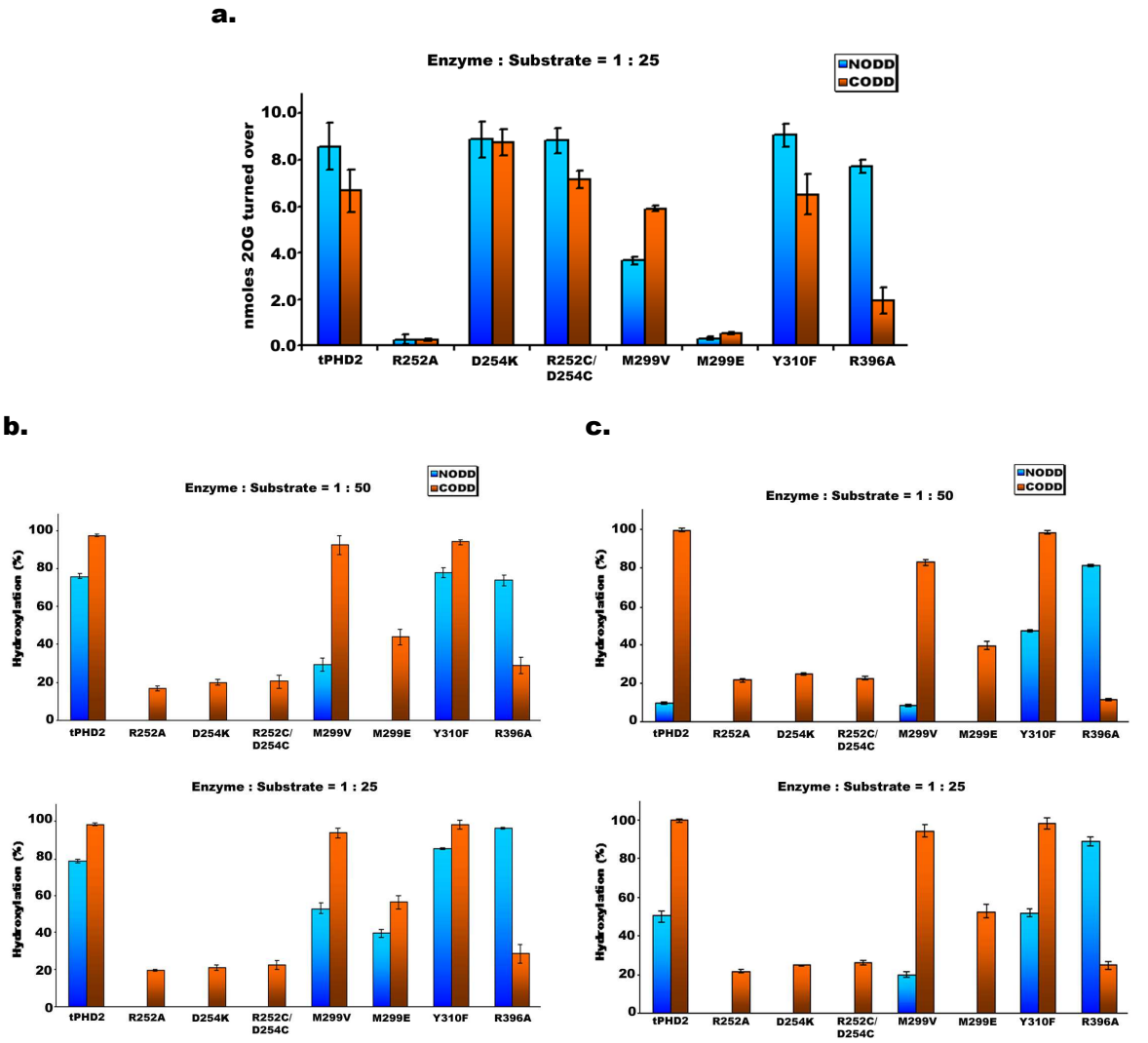


Fig. 3.10 Activity analyses of tPHD2 variants. (a) 2OG turnover (b) substrate hydroxylation as determined by MALDI-TOF mass spectrometry with two enzyme concentrations (2 μM and 4 μM with 100 μM substrate) (c) substrate hydroxylation in NODD/CODD competition experiments with two enzyme concentrations.

3.9 The role of tPHD2 C-terminal helix α4 in CODD binding

Compared to the $P6_3$ crystal form, the C-terminal helix α4 in the tPHD2.CODD complex moves towards the active site (Fig. 3.3). The side facing towards the protein core of helix α4 interacts with CODD, including a salt bridge between $Asp571_{CODD}$ and $Arg396_{PHD2}$ (which is conserved in PHD1-3) (Figs 3.3 and 3.11). These observations are consistent with studies on the role of the C-terminus in PHD2 catalysis; truncations after tPHD2 residue 402 cause loss of hydroxylation activity suggesting a role for helix α4 (residues 393-404) in catalysis.[8]

$Arg396_{PHD2}$ is conserved in PHD1 and PHD3, in fact, in all prolyl hydroxylases amongst human and other higher animals, although not in bacteria or fungi, which is indicative of the fact that HIF is probably not a prime substrate for them.

R396A tPHD2 displayed significantly reduced CODD (<30%), but increased NODD (>120%) hydroxylation both in assays with individual NODD/CODD substrates (Fig. 3.10) and in competition assays (Fig. 3.10). This result is consistent with the lack of an acidic residue in an analogous position to $Asp571_{CODD}$ in NODD (Chapter 4, Fig. 4.8) and reveals a more important role for the $Asp571_{CODD}$-$Arg396_{PHD2}$ salt-bridge in CODD compared to NODD catalysis. The observation that a single point variation can selectively ablate activity with one of two substrates is of relevance to searches for alternative PHD substrates, because it indicates that the selectivity of the PHD isoforms for different non-HIF substrates could be near absolute.

In the tPHD2.CODD complex, CODD residues 569 and 570 have high B-factors indicating possible plasticity within this part of the structure. The cluster of these three sequential aspartyl residues (569-571) in CODD binds to a region of positive charge on the surface of tPHD2 formed by Lys262, Lys291, Lys297, Arg322, Arg325, Arg396 and Lys400. Although $Arg396_{PHD2}$ (which forms a salt-bridge with $Asp571_{CODD}$) plays an important role in CODD binding/catalysis (Fig. 3.3), mutation of the $Asp571_{CODD}$ to Ala retains >80 % activity with wt tPHD2 under standard conditions (Chapter 5).[9] Consistent with the observed role of $Arg396_{PHD2}$ in CODD hydroxylation, double (D569A/D571A or D570A/D569A) and triple mutations (D569A/D570A/D571A) of CODD were shown to have 50-80 % and <10 % activity, respectively, compared to wt tPHD2.[9]

3.10 Additional CODD C-terminal interactions with tPHD2

In the tPHD2.CODD structure, four antiparallel hydrogen bonds were observed between the main chains of CODD residues (Leu574, Phe572) and tPHD2 residues (Asn293, Arg295) from the α3-β4 loop (residues 283-298) (Fig. 3.11). The aromatic side chain of $Phe572_{CODD}$ is positioned to interact with the cationic side-chain of adjacent $Arg295_{PHD2}$. In the $P2_12_12_1$ crystal form, two refined side chain conformations of $Met568_{CODD}$ are positioned to interact with $Tyr403_{PHD2}$. One of the $Met568_{CODD}$ conformations is also positioned to make hydrophobic contact with the methylenes of $Lys400_{PHD2}$. The higher K_m and k_{cat} values observed for the C-terminal truncation of tPHD2 (181-402)[8] compared to the wt tPHD2 might in part be related to the observed hydrophobic interactions between $Met568_{CODD}$ and $Tyr403_{PHD2}$ in the tPHD2.CODD crystal structure.

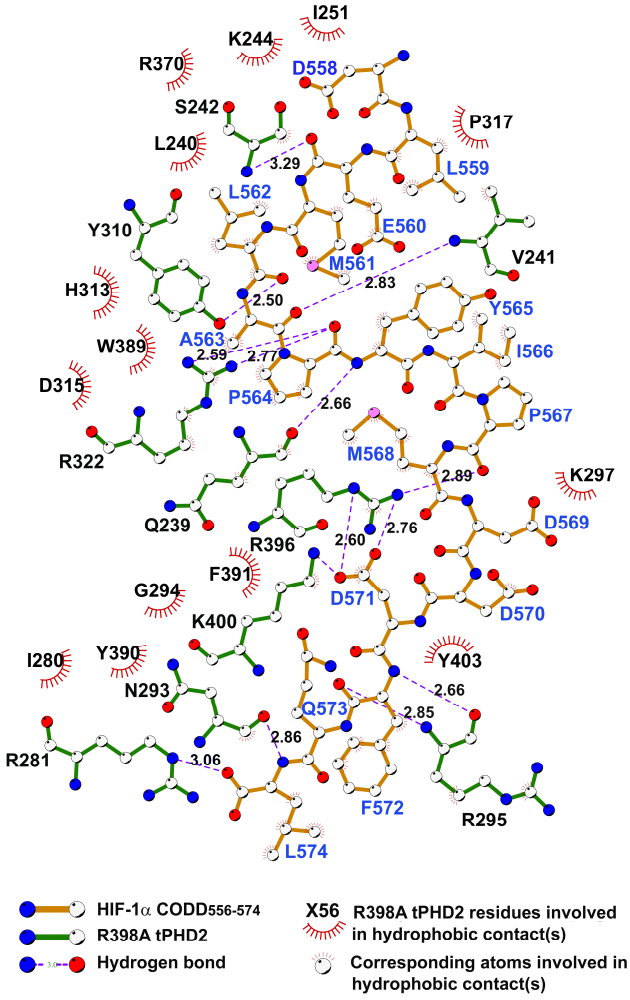


Fig. 3.11 Ligplot[10] representations of tPHD2.CODD complex showing electrostatic interactions of CODD to tPHD2.

3.11 Summary

The structural studies presented here have enabled an understanding of how CODD binds tPHD2 involving a flexible 'β2β3/loop' that encloses the active site, positioning the hydroxylated $Pro564_{CODD}$ in the deep-seated active site. The combined biochemical data demonstrate that variation of some of the CODD binding residues reduces or ablates catalytic ability while other variants, such as Y310F do not interfere with activity. Although the possibility remains that these variants may show differences in kinetic properties when using continuous assays,[8] it seems likely that their corresponding binding partners in CODD may tolerate some degree of plasticity in terms of catalytically productive binding. Future work could be directed toward determining the detailed kinetics of 2OG and CODD for interesting tPHD2 variants including Y310F and R396A.

3.12 References

1. McDonough, M.A., Li, V., Flashman, E., Chowdhury, R., Mohr, C., Lienard, B.M., Zondlo, J., Oldham, N.J., Clifton, I.J., Lewis, J., McNeill, L.A., Kurzeja, R.J., Hewitson, K.S., Yang, E., Jordan, S., Syed, R.S. & Schofield, C.J. Cellular oxygen sensing: Crystal structure of hypoxia-inducible factor prolyl hydroxylase (PHD2). *Proc Natl Acad Sci U S A* **103**, 9814-9 (2006).
2. Elkins, J.M., Hewitson, K.S., McNeill, L.A., Seibel, J.F., Schlemminger, I., Pugh, C.W., Ratcliffe, P.J. & Schofield, C.J. Structure of factor-inhibiting hypoxia-inducible factor (HIF) reveals mechanism of oxidative modification of HIF-1 alpha. *J Biol Chem* **278**, 1802-6 (2003).
3. Oster, L.M., van Scheltinga, A.C., Valegard, K., Hose, A.M., Dubus, A., Hajdu, J. & Andersson, I. Conformational flexibility of the C terminus with implications for substrate binding and catalysis revealed in a new crystal form of deacetoxycephalosporin C synthase. *J Mol Biol* **343**, 157-71 (2004).
4. Masson, N., Willam, C., Maxwell, P.H., Pugh, C.W. & Ratcliffe, P.J. Independent function of two destruction domains in hypoxia-inducible factor-alpha chains activated by prolyl hydroxylation. *Embo J* **20**, 5197-206 (2001).
5. Kageyama, Y., Koshiji, M., To, K.K., Tian, Y.M., Ratcliffe, P.J. & Huang, L.E. Leu-574 of human HIF-1alpha is a molecular determinant of prolyl hydroxylation. *Faseb J* **18**, 1028-30 (2004).
6. Sarakatsannis, J.N. & Duan, Y. Statistical characterization of salt bridges in proteins. *Proteins* **60**, 732-9 (2005).
7. Horovitz, A., Serrano, L., Avron, B., Bycroft, M. & Fersht, A.R. Strength and co-operativity of contributions of surface salt bridges to protein stability. *J Mol Biol* **216**, 1031-44 (1990).
8. Flashman, E., Bagg, E.A., Chowdhury, R., Mecinovic, J., Loenarz, C., McDonough, M.A., Hewitson, K.S. & Schofield, C.J. Kinetic rationale for selectivity toward N- and C-terminal oxygen-dependent degradation domain substrates mediated by a loop region of hypoxia-inducible factor prolyl hydroxylases. *J Biol Chem* **283**, 3808-15 (2008).
9. Li, D., Hirsila, M., Koivunen, P., Brenner, M.C., Xu, L., Yang, C., Kivirikko, K.I. & Myllyharju, J. Many amino acid substitutions in a hypoxia-inducible transcription factor (HIF)-1alpha-like peptide cause only minor changes in its hydroxylation by the HIF prolyl 4-hydroxylases: substitution of 3,4-dehydroproline or azetidine-2-carboxylic acid for the proline leads to a high rate of uncoupled 2-oxoglutarate decarboxylation. *J Biol Chem* **279**, 55051-9 (2004).
10. Wallace, A.C., Laskowski, R.A. & Thornton, J.M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* **8**, 127-34 (1995).

Chapter 4 – Biochemical investigations into the substrate selectivity of tPHD2

Despite having the same cofactor and cosubstrate requirements, the three human PHDs have different physiological roles that are proposed to be related to their catalytic activities.[1,2] The multiple interactions observed between catalytic tPHD2 and CODD in the tPHD2.CODD structure raise the question as to how the different substrate selectivities of the PHD isoforms is achieved. This chapter describes biochemical and modelling studies aimed at addressing this question. In addition, crystal structures of two clinically observed PHD2 mutations causing familial erythrocytosis were obtained. These variants were found to be significantly selective for CODD over NODD hydroxylation. Residues that are not conserved in all three PHDs were identified as important determinants of substrate selectivity.

4.1 Substrate selectivities amongst the PHDs

Despite the highly conserved nature of the PHD catalytic domains (Fig. 4.1), the three PHD isoforms exhibit different CODD/NODD selectivities, with PHD3 almost exclusively catalysing CODD hydroxylation[3,4] and PHD1 being less active toward NODD than CODD relative to PHD2.[4] Apart from their very different N-terminal domains (residues 1-164 in PHD1 and 1-180 in PHD2), sequence conservation is substantially lost in two other regions (Fig. 4.1): the C-terminus and the flexible β2/β3 'finger-like' loop, suggesting a potential role for these regions in determining such selectivity.[5] Experiments were therefore conducted in an attempt to understand whether these sequence variations between the PHDs enable their selectivity for CODD and NODD.

4.2 Role of 'β2β3/loop' in NODD/CODD selectivity

The observation that deletion of 'β2β3/loop' in tPHD2 significantly reduced hydroxylation of HIF-1α substrates to below the level of detection (Chapter 2), suggested an important role for the 'β2β3/loop' in coupling 2OG decarboxylation to substrate hydroxylation.[5] 'Chimeric' PHDs where the tPHD2 'β2β3/loop' residues were exchanged with those from PHD1/PHD3 possessed altered NODD and CODD selectivities; notably a tPHD2-PHD3 chimera (i.e. tPHD2 with the PHD3 loop) displayed significantly higher affinity for CODD (K_M ~6.4 μM) than NODD (K_M ~56.0 μM).[5]

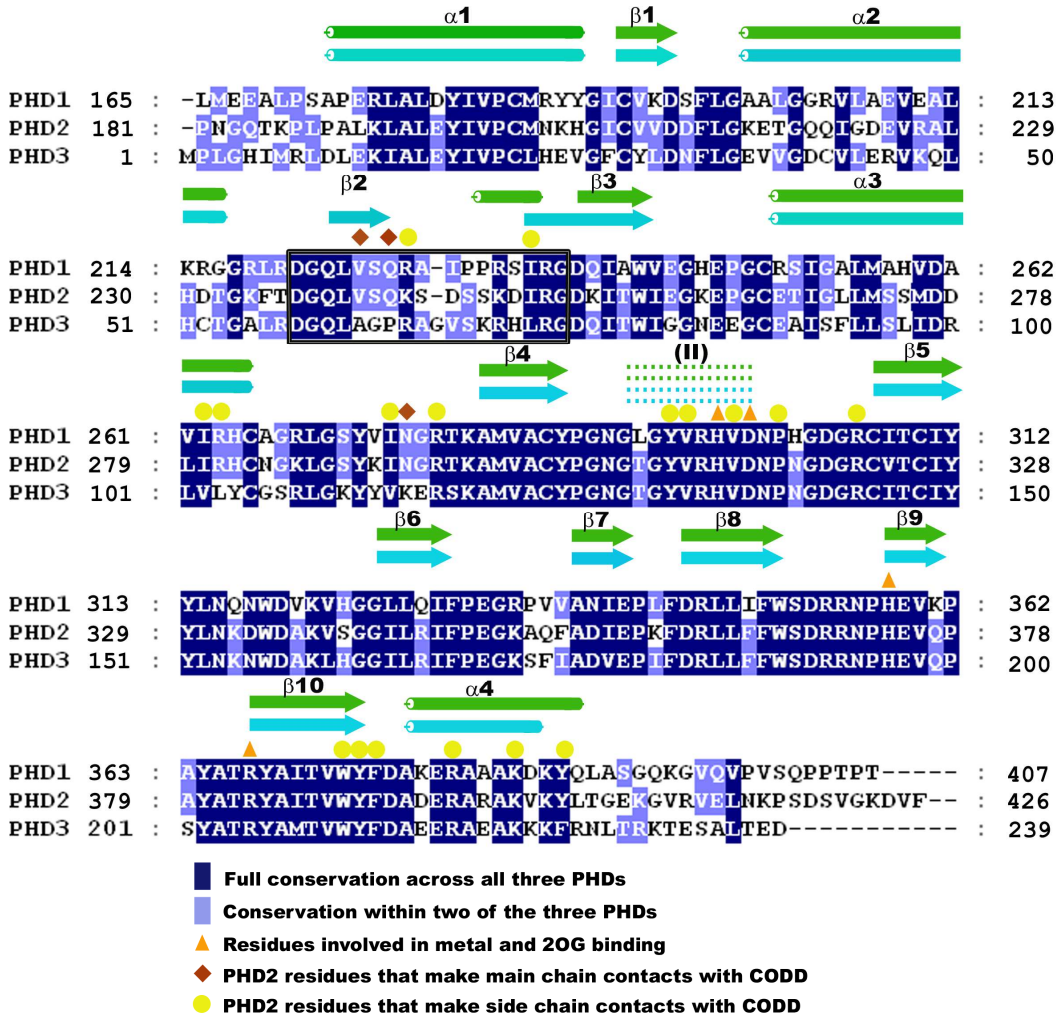


Fig. 4.1 Sequence alignments of human PHDs produced using ClustalW.[6] Amino acid sequence comparison of the human PHD isoforms. Secondary structures in the $P6_3$ (tPHD2.Fe(II).A/B) and $P2_12_12_1$ (tPHD2.CODD) crystal forms are in cyan and green, respectively. Sequences corresponding to the β2β3/loop regions are boxed in black. (II) indicates the β-strand II of the DSBH.[5]

Further, competition experiments where equal amounts of both NODD and CODD were incubated with the enzyme revealed that a tPHD2-PHD3 chimera catalysed CODD to a similar extent as observed with wt tPHD2, but NODD hydroxylation was almost completely ablated (Fig. 4.3).[5] The combined data from kinetic measurements and competition experiments suggest that the β2/β3 loop residues likely enables PHD3 to bind CODD more efficiently although these residues do not completely block NODD binding under the assay conditions.[5] Consistent with the in vitro data, cell-based experiments using yeast two hybrid and luciferase reporter assays with PHD and HIF variants support the proposal that the β2/β3 loop residues are important in activating PHD3 for CODD.[7] However, the observations both in vitro and in cells that PHD3 can catalyse NODD hydroxylation at levels ranging from zero to a very low level,[9,10] implies at least some of the selectivity determinants seem to be from other regions of PHD. In addition, deleting

β2/β3 loop residues did not ablate binding of either CODD or NODD even though hydroxylation was barely detectable[5] indicated that hydroxylation requires binding of a substrate in a productive orientation and that, residues other than those from the β2/β3 loop are also involved in correctly orienting both the tPHD2.CODD and tPHD2.NODD complexes for catalysis.

The combined kinetic analyses on the PHD isoforms[4] and the 'chimeric' PHDs[5] suggest that differences in rates of CODD/NODD hydroxylation (and therefore selectivity) likely arise from differences in substrate binding as well as in the rate of catalysis of the hydroxylation step once the enzyme-substrate complexes have been formed. An understanding of the precise molecular interactions between the three PHDs and both NODD and CODD is desirable in order to identify the factors contributing to substrate selectivity. The crystal structure of PHD2.CODD complex reveals that tPHD2, from the N-terminus, employs multiple residues (both conserved and non-conserved in the PHD isoforms) from strand (βII), the β2β3/loop, β10, the *C*-terminal helix α4 and the loop prior to β4, and α3 to interact with CODD (Chapter 3). Whilst numerous attempts to obtain a crystal structure of tPHD2 complexed with NODD were unsuccessful, site-directed mutagenesis was used to evaluate specific interactions predicted by the computational models of tPHD2.NODD complex. Biochemical studies were particularly aimed at investigating whether a series of interactions involving the side chains of non-conserved PHD residues enable their selectivity for CODD and NODD hydroxylations.

4.3 MDS studies predict a similar binding mode for NODD LXXLAP

In order to investigate the specific determinants of sequence selectivity for the PHDs, it is imperative to understand the exact nature of their multiple interactions with the CODD/NODD substrates. Because molecular dynamic simulations (MDS) can be useful in predicting molecular interactions provided that structures are available for use as controls,[8] MDS were carried out on tPHD2, which accepts both CODD and NODD, and PHD3 which is highly selective for CODD. To investigate the possible CODD/NODD interaction modes, MDS were first carried out on the tPHD2 modelled complexes with the conserved LXXLAP motifs of CODD and NODD. MDS analyses on tPHD2.Fe(II).2OG.LXXLAP-CODD/LXXLAP-NODD complexes using the tPHD2.CODD structure as a template predicted that the residues of the NODD and CODD LXXLAP motifs[9] will bind in near identical conformations to PHD2 (Fig. 4.2).

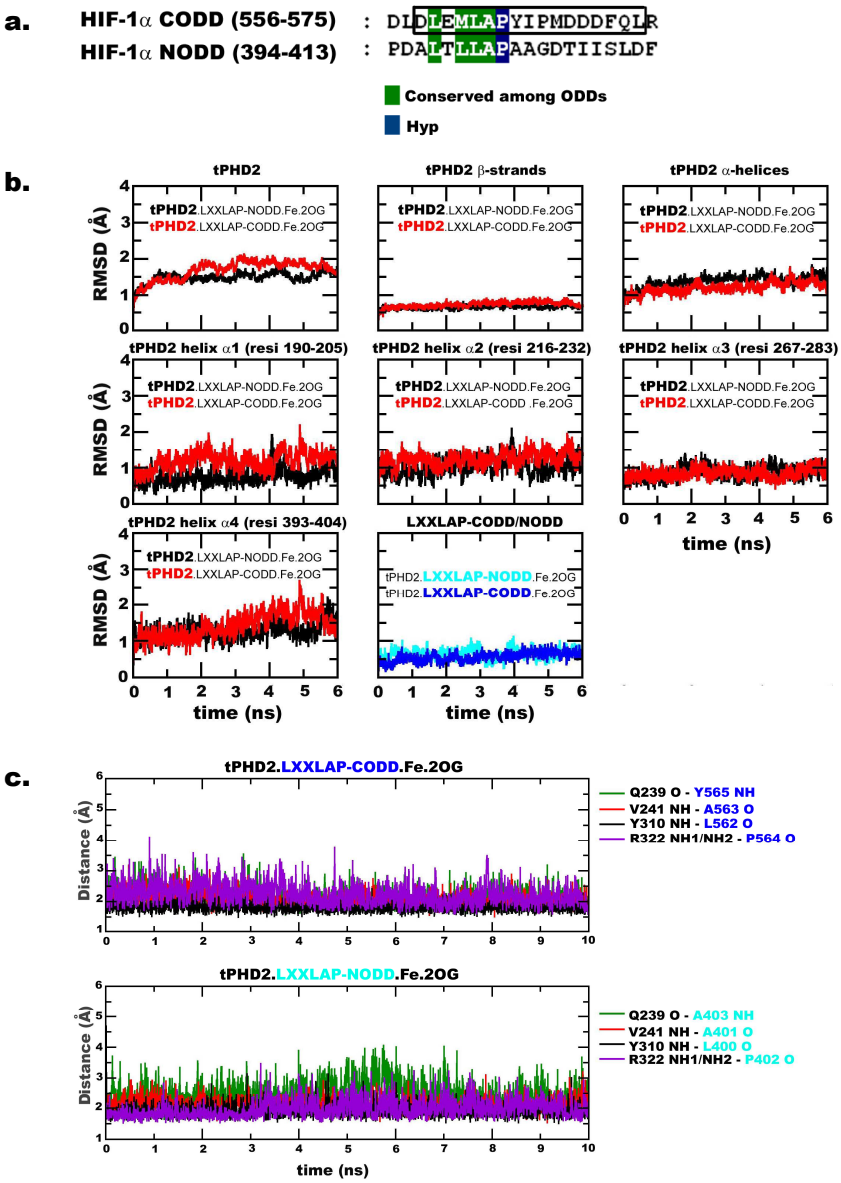


Fig. 4.2 Sequence alignment of CODD and NODD along with MDS studies on tPHD2.Fe.2OG.CODD-LXXLAP and tPHD2.Fe.2OG.LXXLAP-NODD modelled complexes. (a) Sequence alignment of HIF-1α $CODD_{(556\text{-}575)}$ and $NODD_{(394\text{-}413)}$. The hyroxylated proline (Hyp) is highlighted in blue. CODD residues modelled in the tPHD2.CODD structure are highlighted in a box. The panels in (b) show the Cα root-mean-square distances (RMSD) from the initial model as a function of simulation time for different structural elements as indicated in the figure titles by the different coloured labels. (c) shows hydrogen bond distances between the tPHD2 and LXXLAP-CODD/NODD residues during simulations of tPHD2.LXXLAP-CODD and tPHD2.LXXLAP-NODD.

4.4 Selectivity determinants from the conserved LXXLAP binding site

Binding of the LXXLAP motif of both CODD and NODD involves PHD2 residues from the β2/β3 loop, β-strand II and a loop-like structure at the end of β-strand II (Chapter 3). Interestingly, two non-conserved residues of the LXXLAP motifs ($Glu560_{CODD}$ and $Met561_{CODD}$, analogous to $Thr398_{NODD}$ and $Leu399_{NODD}$, respectively) are positioned to make hydrophobic contacts with three β2/β3 loop residues of tPHD2 ($Ser242_{PHD2}$, $Lys244_{PHD2}$ and $Ile251_{PHD2}$), which are not conserved in PHD3. Because the loop appeared

to play an important role in substrate selectivity[6,8] and the MDS data suggested a similar binding mode for CODD-/NODD-LXXLAP motif with the tPHD2 loop residues, it was considered possible that these three residues ($Ser242_{PHD2}$, $Lys244_{PHD2}$ and $Ile251_{PHD2}$) could be involved in determining NODD/CODD selectivity. The other two interactions mediated by the PHD2 loop residues ($Val241_{PHD2}$ and $Gln243_{PHD2}$ analogous to $Ala62_{PHD3}$ and $Pro64_{PHD3}$) only involve backbone interactions and hence are considered less likely to enable substrate selectivity. Point variants of tPHD2 substituting $Ser242_{PHD2}$, $Lys244_{PHD2}$ and $Ile251_{PHD2}$, to glycine, arginine and leucine residues respectively, as in PHD3 ($Gly63_{PHD3}$, $Arg65_{PHD3}$, $Leu73_{PHD3}$) were produced for activity analyses.

When tested for substrate hydroxylation, the S242G variant displayed CODD hydroxylation similar to tPHD2 (Fig. 4.3). However, the NODD hydroxylation appeared to be significantly reduced both in single peptide and in competition experiments with CODD after a standard 15-min incubation. The tPHD2.CODD structure suggested an interaction between side chains of $Ser242_{PHD2}$ of the β2β3/loop and $Glu560_{CODD}$ ($Thr398_{NODD}$) of the LXXLAP motif via a potential hydrogen bond (Ser242 Oγ - Glu560 Oε1, 3.6Å). MDS studies indicated that $Ser242_{PHD2}$ could also maintain similar electrostatic interaction with $Thr398_{NODD}$ and that Oε1 of $Glu560_{CODD}$ could move away from $Ser242_{PHD2}$ Oγ, destabilising the interaction. Therefore, substitution of Ser242 by Gly (as in PHD3) can possibly alter binding of PHDs to both CODD and NODD to variable degrees.

Substitution of $Lys244_{PHD2}$ to Arg led to reduction in activity (relative to wt) with both CODD (<30% reduction) and NODD (>60% reduction) substrates in 15-min standard incubation, further supporting a role for the mobile loop in substrate selectivity (Fig 4.3). In the tPHD2.CODD crystal structure, $Lys244_{PHD2}$ maintains hydrophobic interactions with $Met561_{CODD}$ ($Leu399_{NODD}$ in the analogous position) of LXXLAP motif. Although the effect of lysine-to-arginine substitutions ($Lys244_{PHD2}$/$Arg228_{PHD1}$ and $Arg65_{PHD3}$) may appear subtle from a structural viewpoint, in the folded state, lysines may have more freedom than the bulkier arginines with a guanidinium group that does not allow as much internal rotational freedom due to resonance stabilisation.[10] Thus, substituting $Lys244_{PHD2}$ to Arg might cause suboptimal binding with the substrate(s). Interestingly, sequence analyses of PHDs (Fig. 4.1) reveal $Arg228_{PHD1}$ (in the place of $Lys244_{PHD2}$) as the only difference amongst the CODD binding residues of PHD1 and PHD2. Thus, the data for this variant supported previous observation that PHD1 and a tPHD2-PHD1 chimera (tPHD2

with PHD1 loop residues) were less efficient in catalysing NODD hydroxylation as compared to PHD2.[4,5]

The I251L variant displayed a slight reduction in CODD (~10%) but a more pronounced decrease in NODD (~50%) hydroxylation after a 15 min incubation suggesting that the variant was selective for CODD (Fig. 4.3). To investigate the relative importance of a hydrophobic residue (that binds the first leucine of LXXLAP motif) at this position, an I251G variant was produced and tested for activity. The results show that there is a substantial loss of both CODD (>40% loss) and NODD (>80% loss) in the single substrate experiments with almost none to very low levels of NODD hydroxylation in the competition experiments (Fig. 4.3).

Multiple variants combining the above mutations were then produced with the hope that multiple substitutions might be even less functionally tolerated, and thereby might show more distinct preferences for CODD over NODD hydroxylation. Substitution of $Lys244_{PHD2}$ and $Ile251_{PHD2}$, to arginine and leucine residues respectively, as in PHD3 ($Arg65_{PHD3}$, $Leu73_{PHD3}$) led to a slight reduction in activity with CODD, but a significant (>40%) reduction with NODD, further supporting a role for the mobile loop in determining substrate selectivity (Fig 4.3). In NODD/CODD competition experiments, the K244R/I251L variant catalysed <40% NODD hydroxylation at 1:25 enzyme-substrate ratio (E/S) and almost none when performed at a 1:50 ratio (Fig. 4.3); CODD was fully hydroxylated in these competition experiments. A further S242G/K244R/I251L triple variant had similar activity as tPHD2 in the single substrate experiments (Fig. 4.3). Interestingly, like the tPHD2-PHD3 chimera[5], the triple variant catalysed complete CODD hydroxylation, but with no detectable NODD hydroxylation in substrate competition experiments at the same enzyme-substrate ratios (1:25 and 1:50) (Fig. 4.3).

Collectively, these biochemical data suggest that the three non-conserved residues (Ser242, Lys244 and Ile251) from the β2/β3 loop might have different roles in productive binding of the substrate LXXLAP motif and hence regulating NODD/CODD selectivity. From the structural point of view, whilst a serine-to-glycine substitution (S242G) can enhance binding possibly by increasing substrate accessibility, a lysine-to-arginine and an isoleucine-to-leucine replacements can cause mild steric clash or loss of hydrophobic interaction with the LXXLAP motif (respectively) resulting in an overall reduced binding

affinity. The isoleucine-to-leucine substitution appeared to destabilise more NODD than CODD binding.

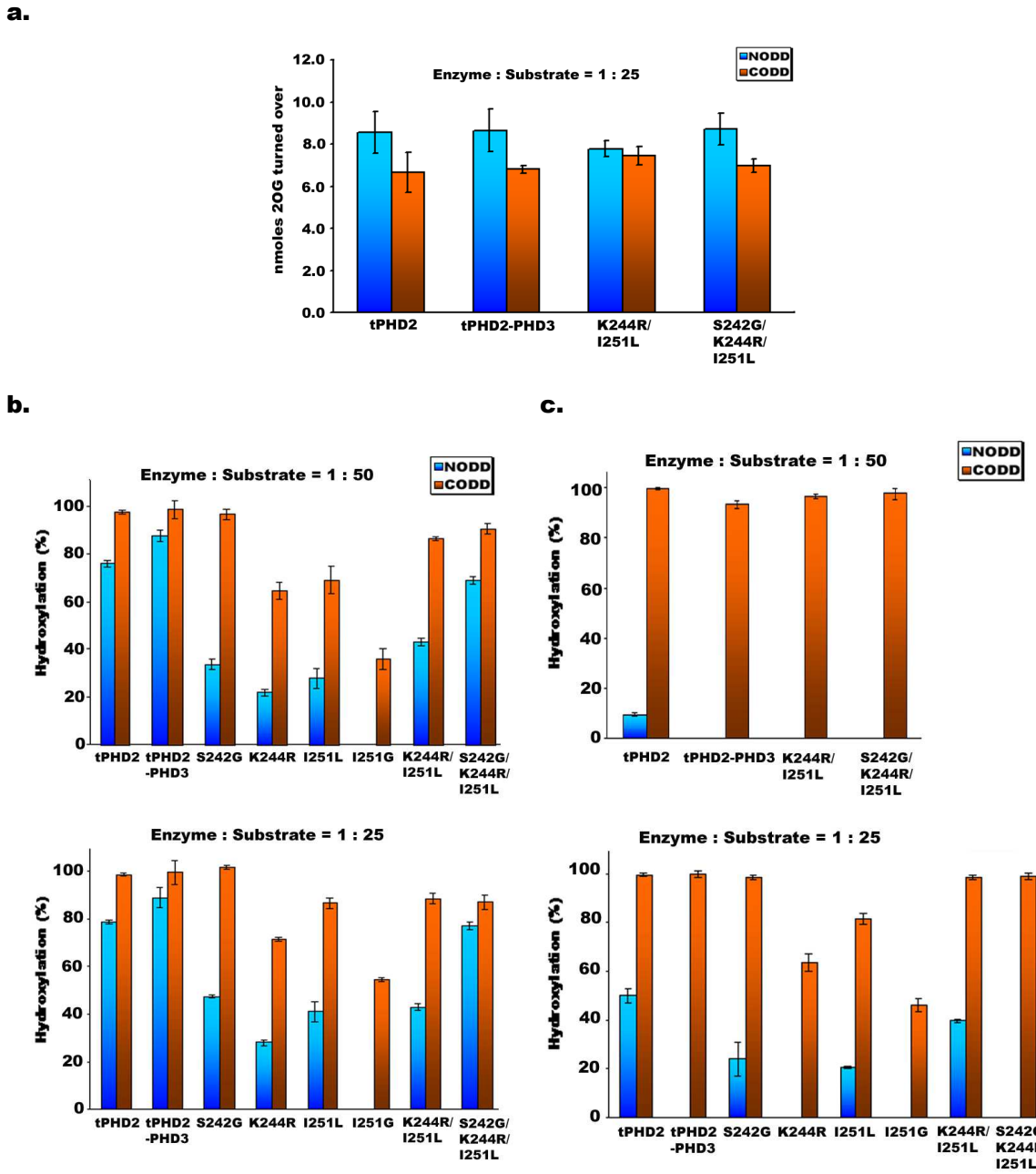


Fig. 4.3 Activity analyses of tPHD2 variants where the CODD residues proposed to interact with the β2/β3 loop were substituted by analogous PHD3 residues. (a) 2OG turnover (b) substrate hydroxylation as determined by MALDI-TOF MS using variable enzyme concentration (c) substrate hydroxylation in competition experiments with variable enzyme concentration.

4.5 Modelling of tPHD2.NODD complexes

Because it was not possible to obtain a tPHD2.NODD crystal structure, models for tPHD2.NODD complex were generated using the tPHD2.CODD crystal structure as a template taking three important considerations into account: (a) the structural studies combined with biochemical data presented here, and those published elsewhere[6,8] revealed

involvement of the β2/β3 loop in binding the LXXLAP motif that is conserved in all HIF[9]; (b) NODD has more hydrophobic residues than CODD; in particular the three sequential aspartyl residues in CODD are substituted by two hydrophobic ($Ile408_{NODD}$, $Ile409_{NODD}$) and a neutral residue ($Thr407_{NODD}$). Therefore, it is likely that NODD may employ more hydrophobic interactions in binding to tPHD2 than CODD; (c) the MDS studies predicted that the conserved LXXLAP motifs of NODD and CODD would bind in a near identical manner (Section 4.3).

The first step was to identify surface clefts in tPHD2 using the SURFNET cleft location algorithm[11] that determined cleft volume and accessibility. Because the tPHD2.Fe(II).A ($P6_3$) structure may reflect an 'open' form of the β2/β3 loop and represent a possible substrate binding conformation, this structure was used to identify hydrophobic patches as potential NODD binding sites. Analyses of the tPHD2 surface (PDB ID: 2G19) revealed ten hydrophobic clefts (Fig. 4.4) that could potentially serve as NODD binding regions. Interestingly, the C-terminus of CODD was observed to bind to one of these computationally derived clefts (cleft 9) in the tPHD2.CODD structure thus partially validating such cleft analyses. All the surface clefts identified by PDBSum were considered for model building. In addition, because the quality of such model complexes depends on the alignment used, multiple sequence alignments of CODD and NODD were employed.

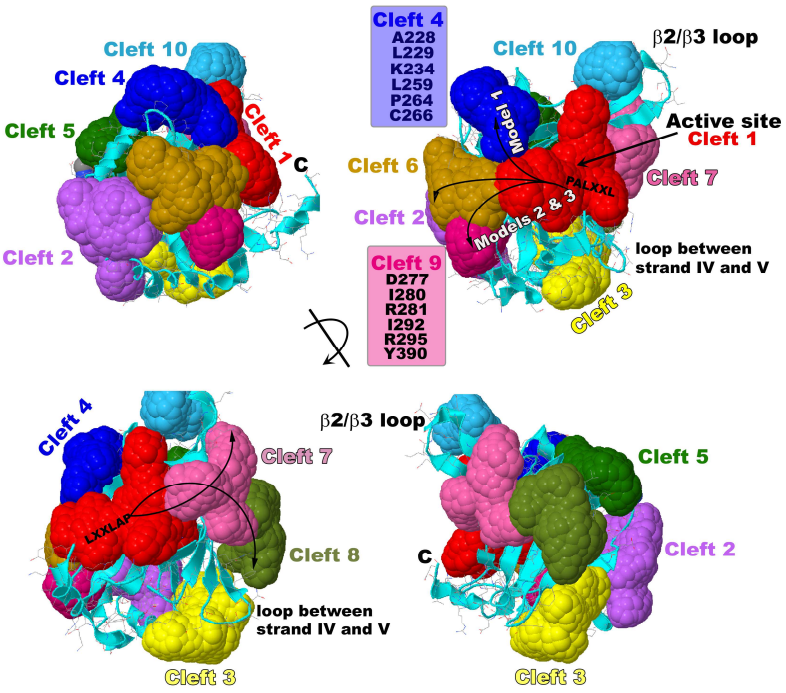


Fig. 4.4 Views from surface cleft analyses using the tPHD2.Fe(II).A structure ($P6_3$). Multiple hydrophobic clefts were generated by PDBSum server.[11]

To include a range of possible modes of interaction between tPHD2 and NODD, an ensemble of 20 models was generated. The five top tPHD2.NODD models were selected on the basis of best fit to the geometric restraints and stereo-chemical quality.[12] Two of these models did not reveal significant interactions with the flexible tPHD2 β2/β3 loop and were subsequently ruled out on the basis of biochemical data on the loop variants (Section 4.4). The remaining three models were subjected to MDS analyses; the resulting PDB coordinates that reached the plateau levels during MDS were used for structural analyses and further biochemical studies.

In the final three models (Fig. 4.5), the C-terminus of NODD was observed to involve interactions with either cleft 4 (model 1) or cleft 9 (models 2 and 3). All three models satisfied the spatial restraints for the LXXLAP motif. Variants of residues that composed surface patches of clefts 4 and 9 were then produced with the aim of validating the modelling studies.

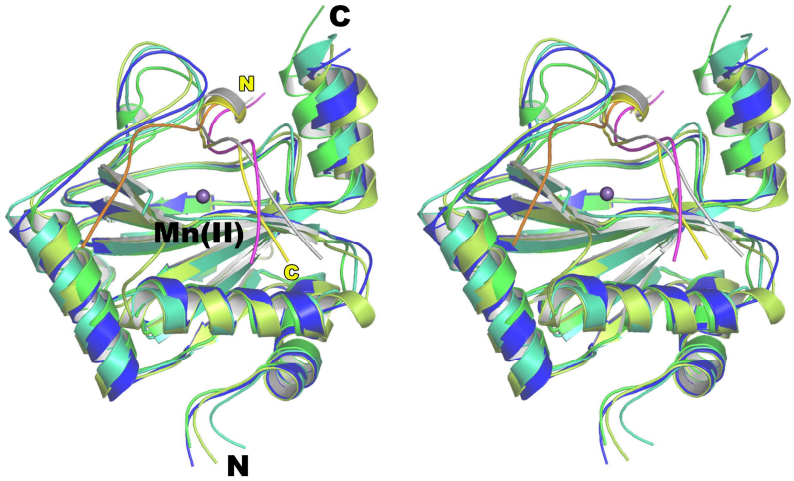


Fig. 4.5 Stereoviews of superimposition of tPHD2.CODD crystal structure with three tPHD2.NODD models derived after MDS. Cartoons represent tPHD2 (green).CODD (yellow) crystal structure and tPHD2 (lemon green).NODD (orange) model 1, tPHD2 (cyan).NODD (purple) model 2 and tPHD2 (blue).NODD (grey) model 3 modelled complexes. Note the apparent structural flexibility in the β2β3/loop region and the α helices, in particular the C-terminal helix α4.

4.5.1 Variant analyses towards validating tPHD2.NODD model 1

The tPHD2.NODD model 1 predicted that tPHD2 residues (cleft 4 in Figs 4.4 and 4.6) Ala228, Leu229, Lys234, Leu259, Pro264 and Cys266 make hydrophobic contacts with $Ile408_{NODD}$, $Ile409_{NODD}$ and $Leu411_{NODD}$ residues C-terminal to the LXXLAP motif. Of note, the tPHD2 residues, Ala228, Lys234, Thr236 and Pro264 are not conserved in PHD3. The following variants were produced: S242G/K244R/I251L/P264E and S242G/K244R/I251L/A228Q/K234A/T236R/P264E where these predicted NODD binding

residues in the tPHD2 were replaced by the equivalent PHD3 residues. The aim was to investigate whether these substitutions caused any reduction in NODD hydroxylation. The results of activity assays show that these two variants displayed similar activity when compared to the S242G/K244R/I251L variant using both NODD and CODD substrates. In fact, the S242G/K244R/I251L/A228Q/K234A/T236R/P264E variant showed slightly more NODD hydroxylation in a competition experiment with CODD (Fig. 4.7).

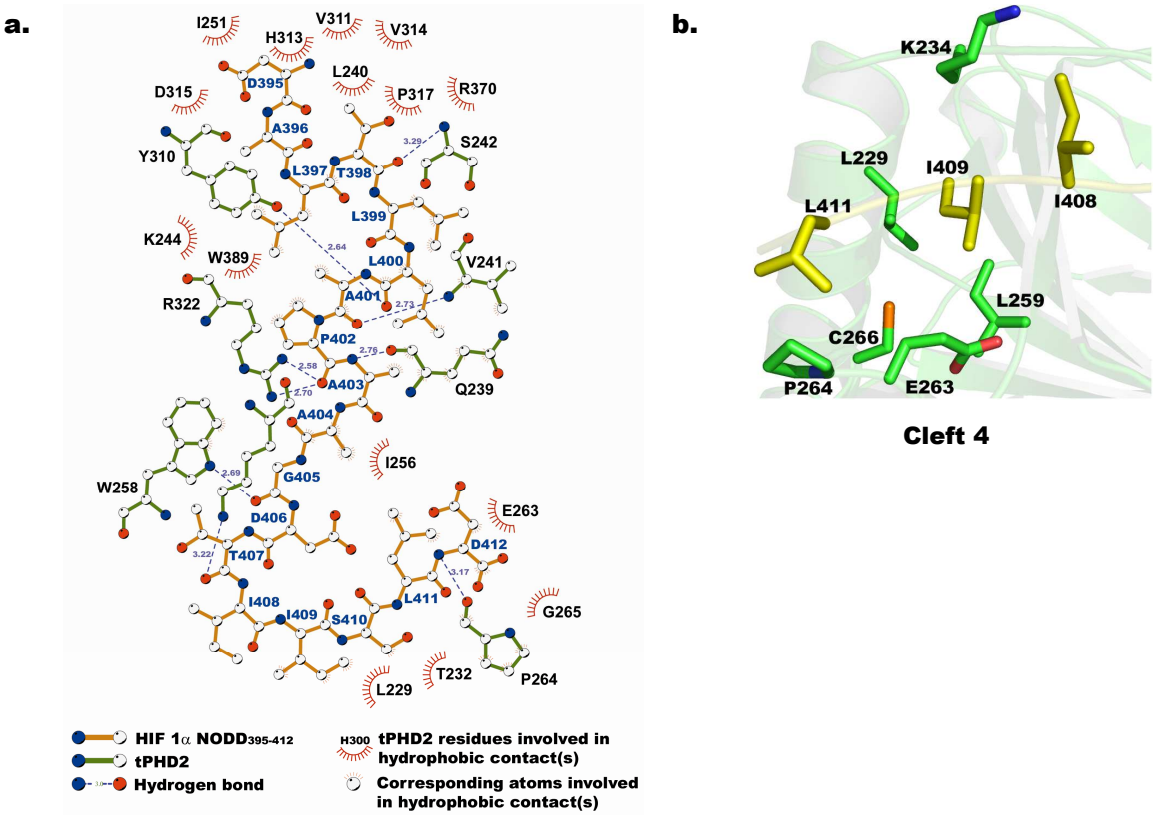


Fig. 4.6 Interactions observed in tPHD2.NODD model 1. (a) Ligplot representations of tPHD2.NODD model 1. Binding of the C-terminal residues of NODD to tPHD2 was predicted to rely more on hydrophobic interactions than that of CODD due to the presence of three sequential neutral/hydrophobic residues, Thr407 Ile408 and Ile409 in the NODD sequence as opposed to the three sequential aspartyl residues in CODD. (b) Hydrophobic interactions of the C-terminal NODD residues (Ile408, Ile409 and Leu411 in yellow) with tPHD2 cleft 4 (Fig. 4.4) in green. The hydrophobic cleft is formed by $Ala228_{PHD2}$, $Leu229_{PHD2}$, $Lys234_{PHD2}$, $Leu259_{PHD2}$, $Pro264_{PHD2}$ and $Cys266_{PHD2}$ of which Ala228, Lys234, Thr236 and Pro264 are not conserved in PHD3.

In addition, a variant of $Cys266_{PHD2}$ that formed part of the predicted $Ile409_{NODD}$ binding site in tPHD2.NODD model 1 was substituted for alanine (Fig. 4.6b). The C266A variant did not show significant difference in activity using CODD (Fig. 4.7), although slightly reduced activity (>10%) in 2OG turnover was observed with NODD.

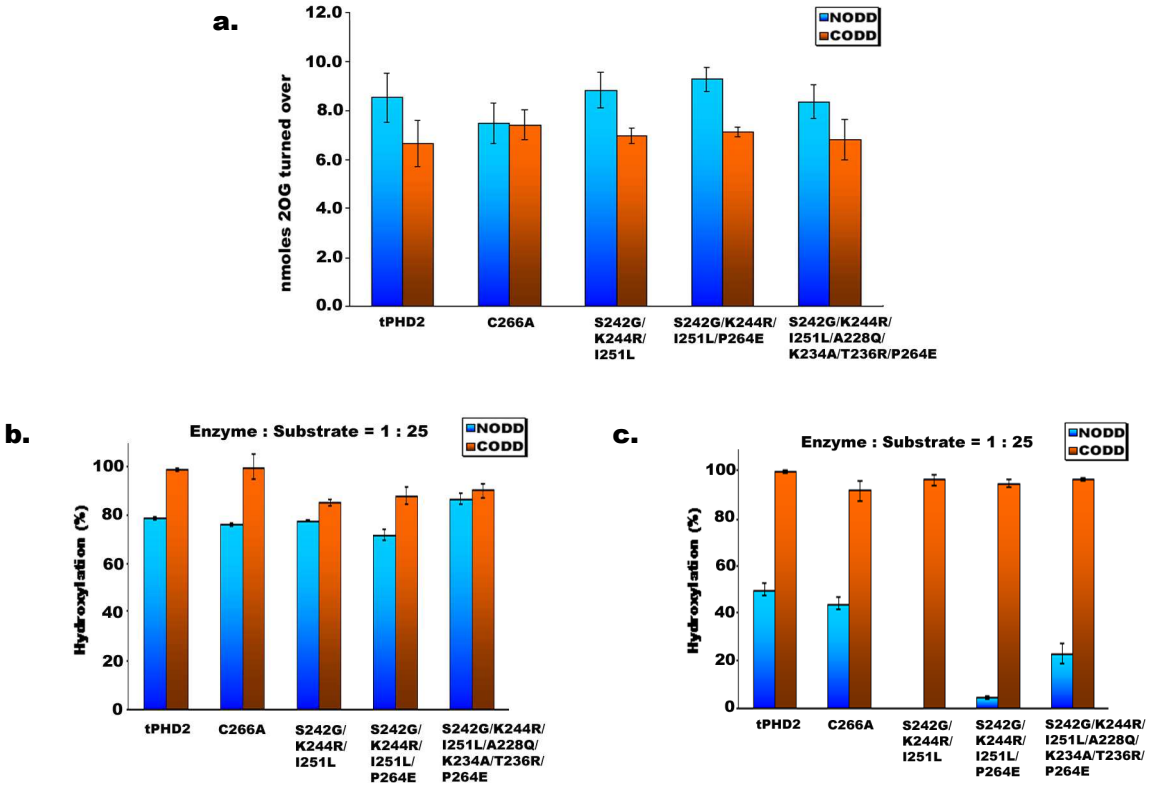


Fig. 4.7 Activity analyses of tPHD2 variants where tPHD2 residues were substituted by analogous PHD3 residues based on the tPHD2.NODD model 1. a) 2OG turnover b) substrate hydroxylation as determined by MALDI-TOF mass spectrometry at 1:25 E/S ratio c) substrate hydroxylation in competition experiments using the same E/S ratio.

Furthermore, during the MD simulations of the tPHD2.NODD model 1, the C-terminal part of NODD was found to be rather loosely bound to tPHD2. Collectively the data from the biochemical and simulations studies were not supportive of the tPHD2.NODD model 1.

4.5.2 Analyses of alternative tPHD2.NODD models 2 and 3

The remaining two tPHD2.NODD models utilised slightly different sequence alignments of NODD/CODD for modelling the C-terminal regions of NODD relative to the conserved LXXLAP motif (Fig. 4.8a). For model 3, a gap was inserted after LXXLAP in the NODD sequence to align the C-terminal $Leu574_{CODD}$ and $Leu411_{NODD}$ (Fig. 4.8a). Consistent with the previous simulations studies using NODD/CODD LXXLAP (Section 4.3), these two models displayed similar modes of interaction with the LXXLAP motif (Figs 4.5 and 4.8). Away from the LXXLAP, MDS predicted that the backbones of NODD in these two models would adopt similar but not identical conformations to CODD (Fig. 4.8).

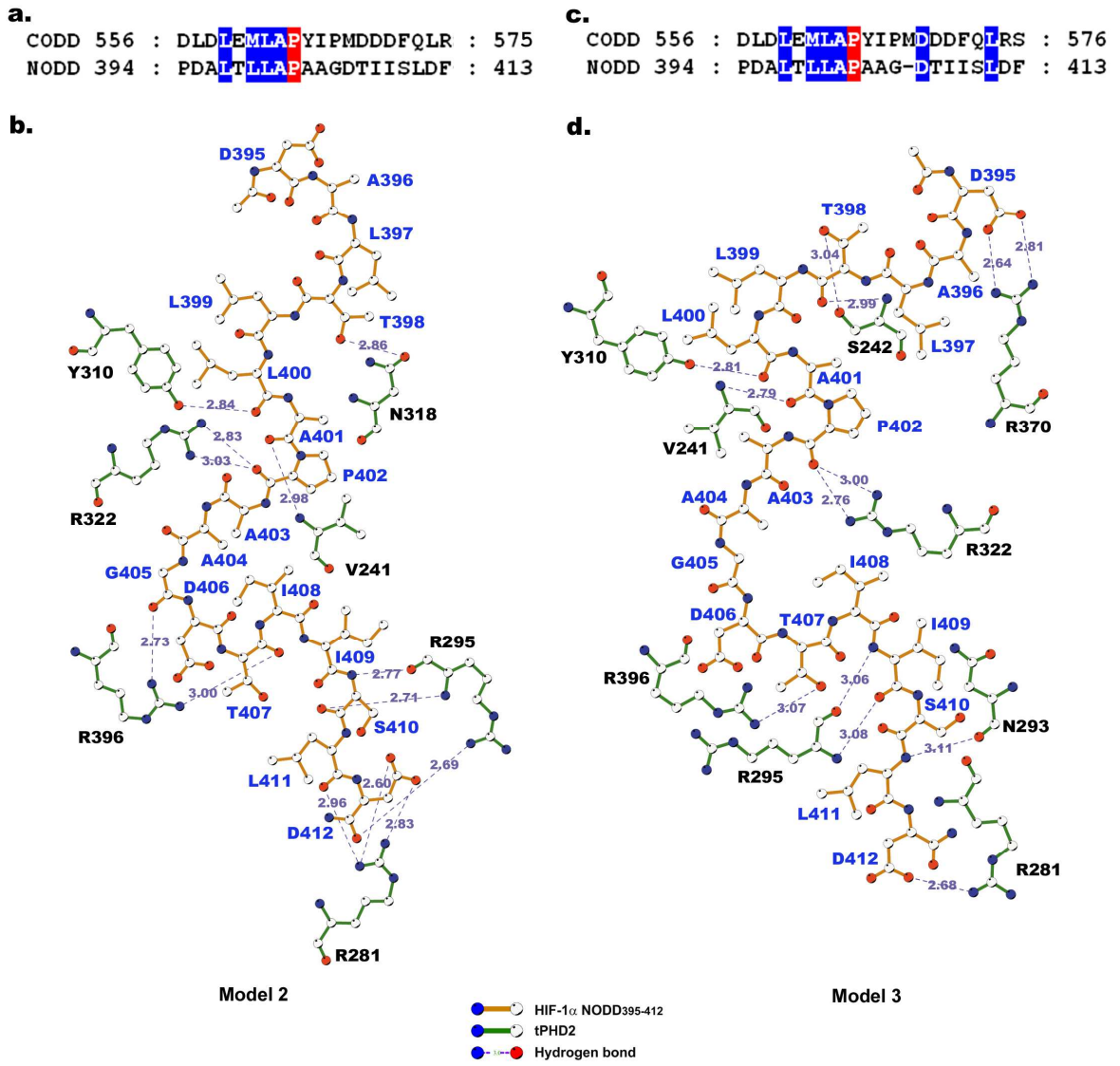
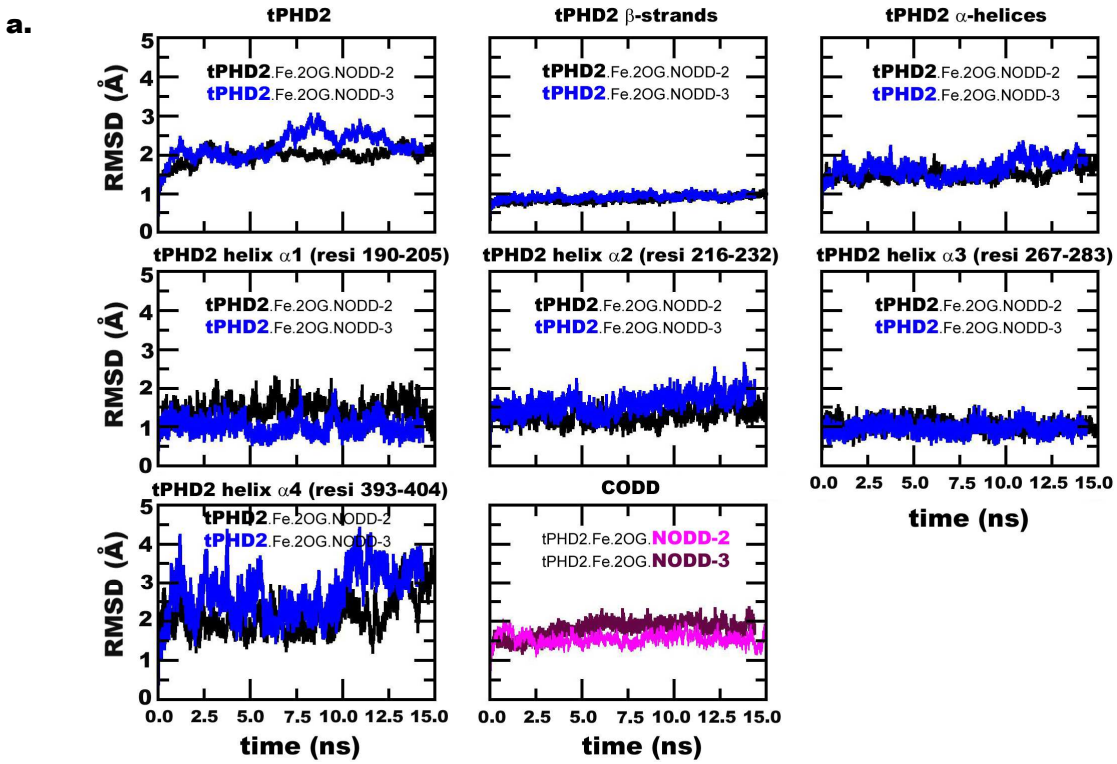


Fig. 4.8 Ligplot representations of tPHD2.Fe(II).2OG. NODD models (2 and 3) with the sequence alignments for CODD and NODD. Sequence alignments of CODD and NODD used to generate models 2 and 3 are presented in (a) and (c), respectively. Ligplot representations of tPHD2.NODD models (2 and 3) predicting H-bonds/electrostatic interactions between tPHD2 and NODD are in (b) and (d).

MDS predicted that electrostatic interactions at the N-terminus of NODD, containing LXXLAP motif, would involve the side chains of the non-conserved $Asn318_{PHD2}$ ($His302_{PHD1}$, $Asn140_{PHD3}$) and two conserved residues, $Tyr310_{PHD2}$ and $Arg322_{PHD2}$ in model 2. Analogous interactions involving these two conserved residues were also observed in model 3, in addition to a salt-bridge between $Arg370_{PHD2}$ and $Asp395_{NODD}$. Interestingly, variation of the residue following $Arg370_{PHD2}$, R371H has been identified as a clinical mutation causing familial erythrocytosis (discussed in Section 4.7).[13]

Residues C-terminal to the LXXLAP motif in NODD had multiple interactions in both models predominantly between the side chain atoms of tPHD2 and NODD main chain

atoms, as predicted by MDS (Figs 4.8 and 4.9). Due to a gap inserted after the LXXLAP sequence in the alignment used for model 3 (Fig. 4.8a), many of the interactions were observed between the same tPHD2 residues and those +1 shifted in NODD; for example, $Arg395_{PHD2}$ interacted with $Ser410_{NODD}$ main chain NH/O in model 2 and $Ile409_{NODD}$ main chain NH/O in model 3. $Arg396_{PHD2}$, that formed a salt-bridge interaction with $Asp571_{CODD}$ in the tPHD2.CODD crystal structure, was predicted to form hydrogen bonds with $Gly405_{NODD}$, $Asp406_{NODD}$ and $Thr407_{NODD}$ main chain atoms in model 2. In model 3, the interactions of $Arg396_{PHD2}$ with NODD residues were comparatively weaker as suggested by predicted higher bond distances (Appendix 5). Thus, the two modelled tPHD2.NODD complexes reveal slightly different binding modes for NODD at the C-terminal region. Nevertheless and strikingly, both models coupled with MDS analyses suggested that $Asp412_{NODD}$ ($Leu574_{CODD}$) can make a stable electrostatic interaction with $Arg281_{PHD2}$ ($Leu103_{PHD3}$) in solution (Fig. 4.10). The hydrophobic interactions observed in both simulated models are summarised in Appendix 5.



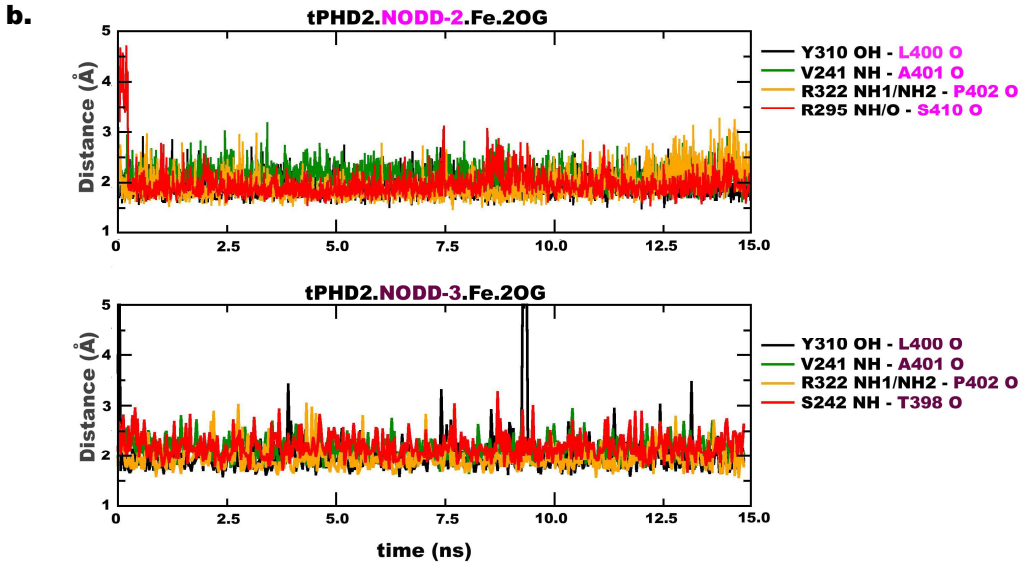


Fig. 4.9 MDS studies on tPHD2.NODD modelled complexes (model 2 and model 3). The panels in (a) show the Cα RMSD from the initial model as a function of simulation time for different structural elements, indicated in the figure titles by different labels. (b) shows hydrogen bond distances between the tPHD2 and NODD residues during simulations of tPHD2.NODD model 2 and 3 as labelled.

4.6 Selectivity determinants involving the tPHD2 C-terminal substrate binding site

The predicted $Leu574_{CODD}$ binding site, which is important for HIF-1α hydroxylation in cells,[14] is notably different in PHD3 to that of PHD1/2 (Chapter 3, Fig 3.5f). This is important because NODD has an acidic residue ($Asp412_{NODD}$) at the analogous position to $Leu574_{CODD}$ (Fig. 4.8a). The predicted $Leu574_{CODD}$ binding site in PHD3 possesses a leucine residue ($Leu103_{PHD3}$), compared to the presence of arginine residues in PHD1 ($Arg263_{PHD1}$) and PHD2 ($Arg281_{PHD2}$) at the analogous position (Fig. 4.10). Substitution of $Arg281_{PHD2}$ with a leucine residue led to little or no reduction in activity with CODD, but significant reduction (>40%) in activity with NODD. The effect was more pronounced in a NODD/CODD competition experiment, where NODD was hydroxylated to about <10% when using a 1:25 E/S ratio and remained almost completely unmodified using a 1:50 E/S ratio (Fig. 4.11).

A further triple tPHD2 variant, I280V/I292V/R281L was produced with the aim of making the R281L tPHD2 variant even more 'PHD3 like'. Substitution of $Ile280_{PHD2}$, $Arg281_{PHD2}$ and $Ile292_{PHD2}$ with the corresponding PHD3 residues, $Val102_{PHD3}$, $Leu103_{PHD3}$ and $Val114_{PHD3}$ was carried out. The triple variant hydroxylated CODD to a similar level of tPHD2 R281L variant; however, NODD hydroxylation was unexpectedly increased to

>70% (for triple variant) compared to ~30% by the R281L variant (Fig. 4.11). When in competition with CODD, NODD was hydroxylated <20% using a 1:50 E/S ratio compared to almost complete lack of NODD hydroxylation by the R281L variant. Using more enzyme (E/S ratio, 1:25), the I280V/I292V/R281L and R281L variants catalysed >30% and <10% NODD hydroxylation, respectively in the competition experiments (Fig. 4.11). The results clearly suggested that replacing two isoleucines (or at least one) by valine (1 methylene less) residues might improve CODD hydroxylation and even rescue NODD hydroxylation to some extent. Modelling studies suggested that such a substitution could reduce the size of the hydrophobic pocket in the tPHD2 I280V/I292V/R281L variant (and therefore possibly in PHD3) that might serve as a 'snug fit' for $Leu412_{NODD}$ or the following residue, $Phe413_{NODD}$. This is consistent with yeast-two hybrid interaction data where the $Leu574_{CODD}$ binding site in PHD3 was shown to have a strong preference for both leucine and phenylalanine, whereas PHD2 had a slightly higher preference for valine (shorter) over leucine in this position.[15]

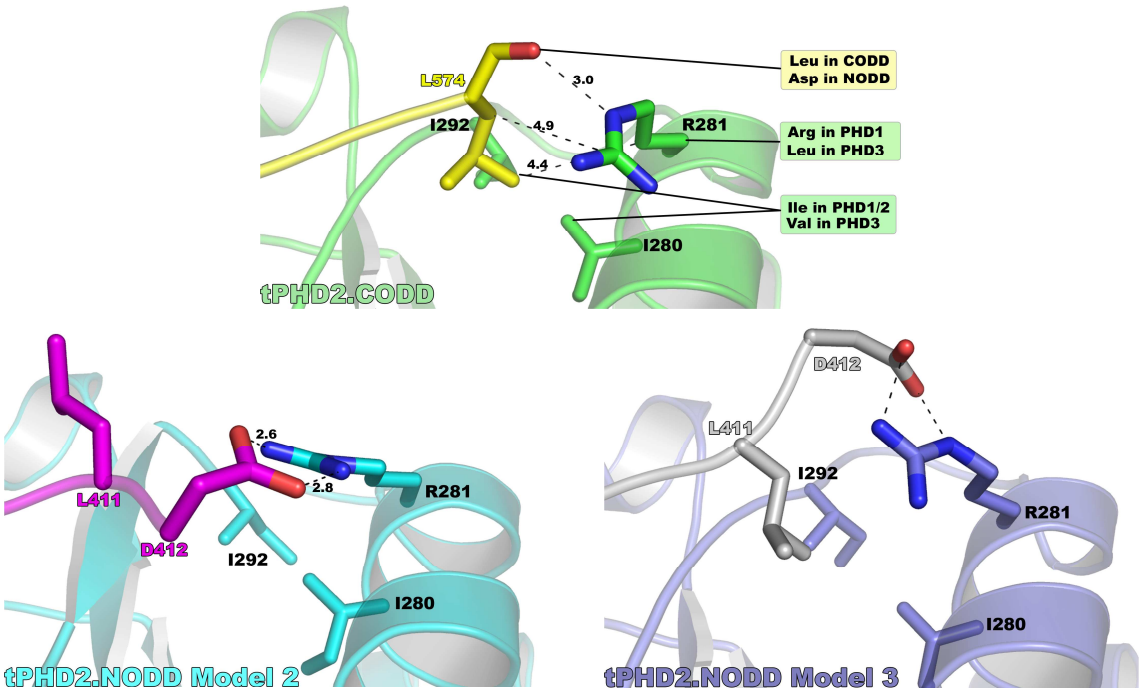


Fig. 4.10 $Leu574_{CODD}$ binding site showing the position of $Arg281_{PHD2}$ relative to $Leu574_{CODD}$. Note that the residues 293 – 295 form a short β-strand in the MDS derived tPHD2.NODD model 2, which was not observed either in the crystal structure of tPHD2.CODD or in the tPHD2.NODD model 3.

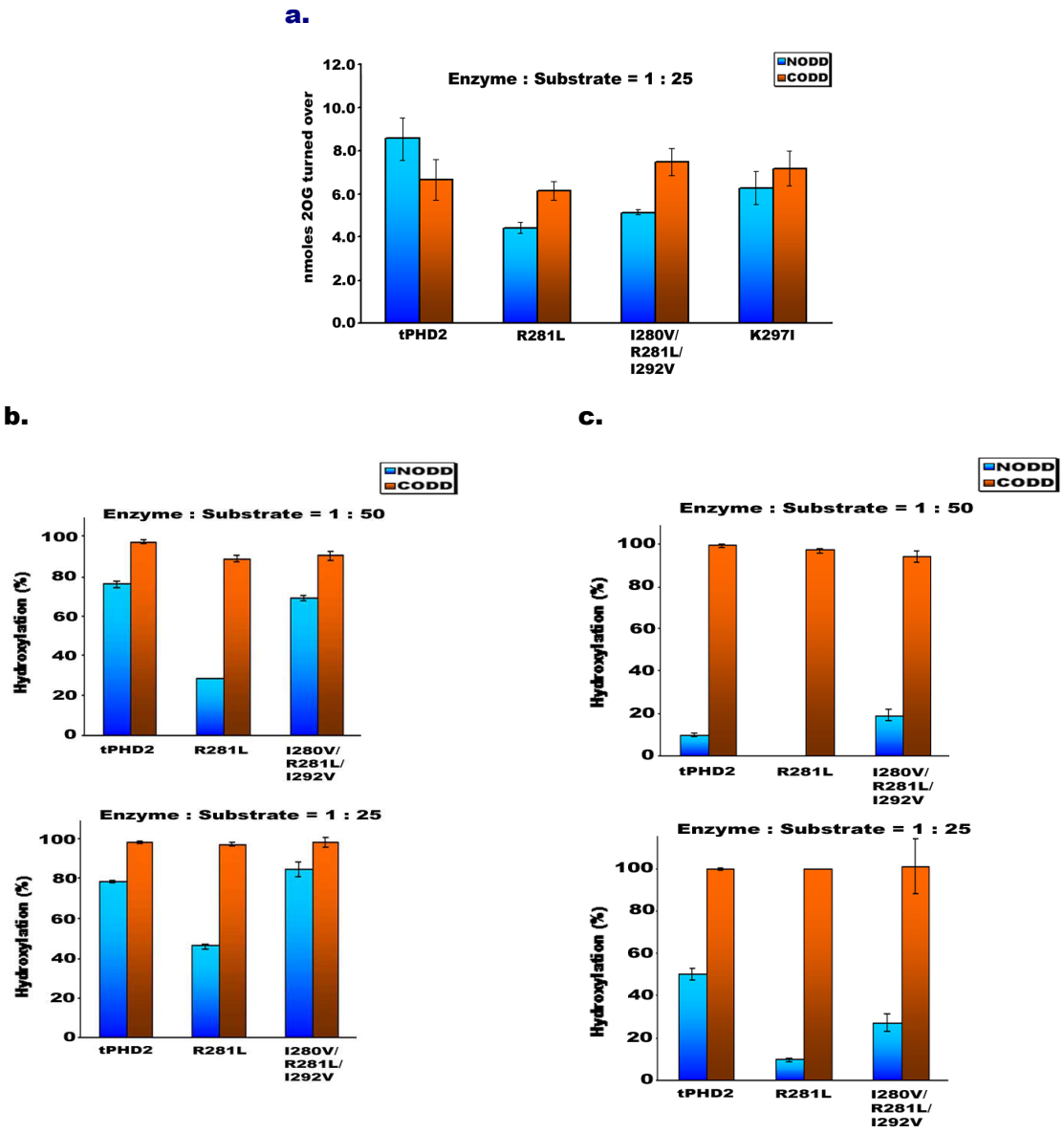


Fig. 4.11 Activity analyses of tPHD2 variants where tPHD2 residues were substituted by analogous PHD3 residues based on the tPHD2.NODD model 2 and 3. a) 2OG turnover b) substrate hydroxylation as determined by MALDI-TOF mass spectrometry at variable enzyme concentrations c) substrate hydroxylation in competition experiments using the same enzyme-to-substrate ratios.

Thus it is proposed that in PHD1/2 the Arg 263_{PHD1}/ Arg 281_{PHD2} residue can enable binding (via its methylenes) of both the hydrophobic side chain of Leu 574_{CODD} and the acidic side chain of Asp 412_{NODD} (via its basic guanidine group) (Fig. 4.10). A positively charged surface likely favours binding of the negatively charged side-chains of Asp 412_{NODD}. In PHD3, the predicted pocket is predominantly hydrophobic so disfavours binding of the acidic Asp412_{NODD} side chain. Support for this proposal comes from yeast-two hybrid experiments where except for phenylalanine, substitution of Leu574 with all other residues was disfavoured for a stable interaction with PHD3.[15] The observation that the triple variant (I280V/I292V/R281L) was slightly better than R281L tPHD2 variant in

catalysing both CODD and NODD hydroxylation suggests an important role for at least one of the two valines (in places of $Ile280_{PHD2}$ and $Ile292_{PHD2}$). Therefore, in addition to the surface charge, the size of the Leu574 binding site is also important in maintaining productive substrate binding.

4.7 Clinically observed mutations

Two clinically observed heterozygous mutations in PHD2, P317R and R371H, are linked with familial erythrocytosis.[13,16] $Pro317_{PHD2}$ is proximal to the iron coordinating HXD motif ($His313_{PHD2}$ and $Asp315_{PHD2}$), whilst $Arg371_{PHD2}$ is close to the other metal coordinating residue, $His374_{PHD2}$ (β9).[5] Crystal structures of tPHD2 P317R.Mn(II).B (1.9 Å) and tPHD2 R371H.Mn(II).B (1.7 Å) were solved by molecular replacement using PDB ID 2G19 as a search model. Data collection and refinement statistics are given in Table 4.1. The structures revealed no discernible perturbation of the metal binding site (Fig. 4.12). In the tPHD2.CODD structure $Pro317_{PHD2}$ is part of a hydrophobic patch that binds the 3_{10}-helix of the LXXLAP motif (Fig. 4.12) whilst $Arg371_{PHD2}$ is relatively remote from the substrate binding region (Fig. 4.12). Activity analyses showed that tPHD2 R371H retained >60% activity with NODD but was equally active as wildtype with CODD under standard conditions (Fig. 4.13).

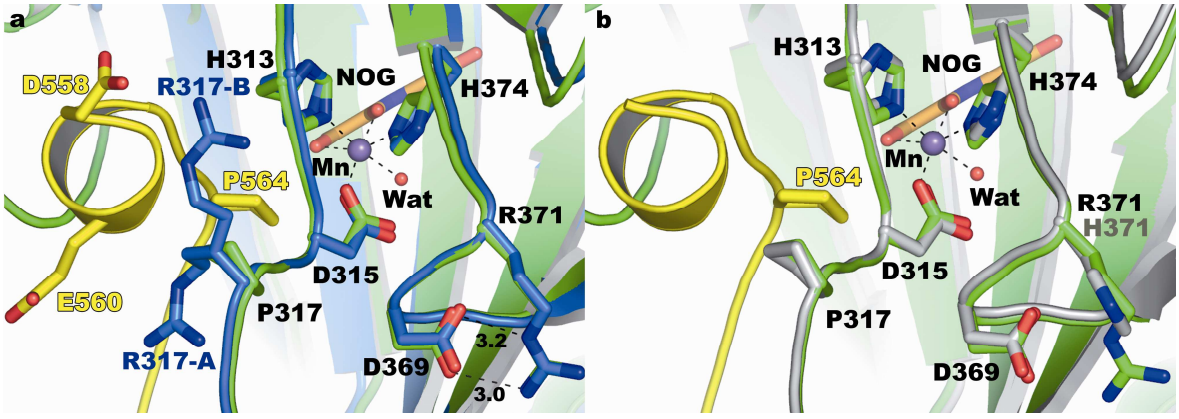


Fig. 4.12 Superimposition (separately) of the crystal structures for two clinically observed PHD2 mutations (P317R and R371H) in complex with Mn(II).B with the tPHD2.CODD structure (green). (a) Structure of the P317R mutant (dark blue) showing the two alternate conformations of Arg317; (b) Structure of the R371H mutation (gray) showing how this mutant does not directly affect CODD binding or Fe(II) coordination. In wt tPHD2, $Arg371_{PHD2}$ is involved in a salt-bridge with $Asp369_{PHD2}$.

P317R PHD2 (full-length) has been reported to cause reduction in HIF-$1\alpha_{549\text{-}575}$ hydroxylation (<10% of wt PHD2) in pull down assays.[16] In contrast, it was found in the current study that under standard conditions, tPHD2 P317R possessed ~80% activity with

CODD compared to wt tPHD2, but did not hydroxylate NODD (Fig. 4.13). In the tPHD2 P317R.Mn(II).B crystal structure, $Arg317_{PHD2}$ was modelled in two alternate conformations with equal occupancy (Fig. 4.12). In one of these conformations the $Arg317_{PHD2}$ side chain points towards the core DSBH, such that the methylenes occupy equivalent positions to C^3 and C^4 of Pro317 in wt tPHD2. This conformation may enable the P317R mutant to interact with the 3_{10} helix in CODD in a similar manner to wt tPHD2. The other modelled conformation for $Arg317_{PHD2}$ appears to clash with $Leu559_{CODD}$ of the LXXLAP motif. The lack of activity of P317R with NODD is consistent with a relatively important role for binding of the LXXLAP residues for NODD compared to CODD catalysis (NODD cannot make the $Asp571_{CODD}$-$Arg\ 396_{PHD2}$ salt-bridge, Fig. 4.8). This suggests that the phenotype of the P317R mutant may partly result from decreased NODD versus CODD hydroxylation.

Table 4.2. Crystallographic data processing and refinement statistics.

Measurement	**tPHD2P317R.Mn(II).B**	**tPHD2R371H.Mn(II).B**
Data Collection		
Space Group	$P6_3$	$P6_3$
Cell dimensions a,b,c (Å)	111.18	110.67
	111.18	110.67
	40.03	40.23
Resolution (Å)	27.80-1.90 (2.00-1.90)*	32.05-1.70 (1.79-1.70)*
No. of unique reflections	20673	30735
Completeness (%)	99.1 (98.6)*	100 (100)*
Redundancy	2.8	10.5
R_{sym}**	0.088 (0.354)*	0.087 (0.500)*
Mean I/σ(I)	9.1 (2.9)*	18.0 (3.8)*
Refinement		
R_{factor} ‡	0.214	0.212
R_{free}	0.226	0.230
R.m.s. deviation		
Bond length, Å	0.007	0.005
Bond angle, °	1.3	1.3
No. of atoms		
Protein	1721	1716
Ligand	19 (B)	19 (B)
Water	138	159
$<B_{factor}>Å^2$		
Protein	26.75	26.18
Ligand	12.54 (B)	12.81 (B)
Metal	11.20 (Mn)	12.27 (Mn)
Water	33.46	35.51

*Highest resolution shell shown in parentheses.

**Rsym = $\sum|I-<I>|/\sum I$, where I is the intensity of an individual measurement and <I> is the average intensity from multiple observations.

‡ $R_{factor} = \sigma hkl||Fobs\ (hkl)| - k\ |Fcalc\ (hkl)||/\sigma hkl|Fobs\ (hkl)|$ for the working set of reflections; R_{free} is the R-value for ~6% of the reflections excluded from refinement.

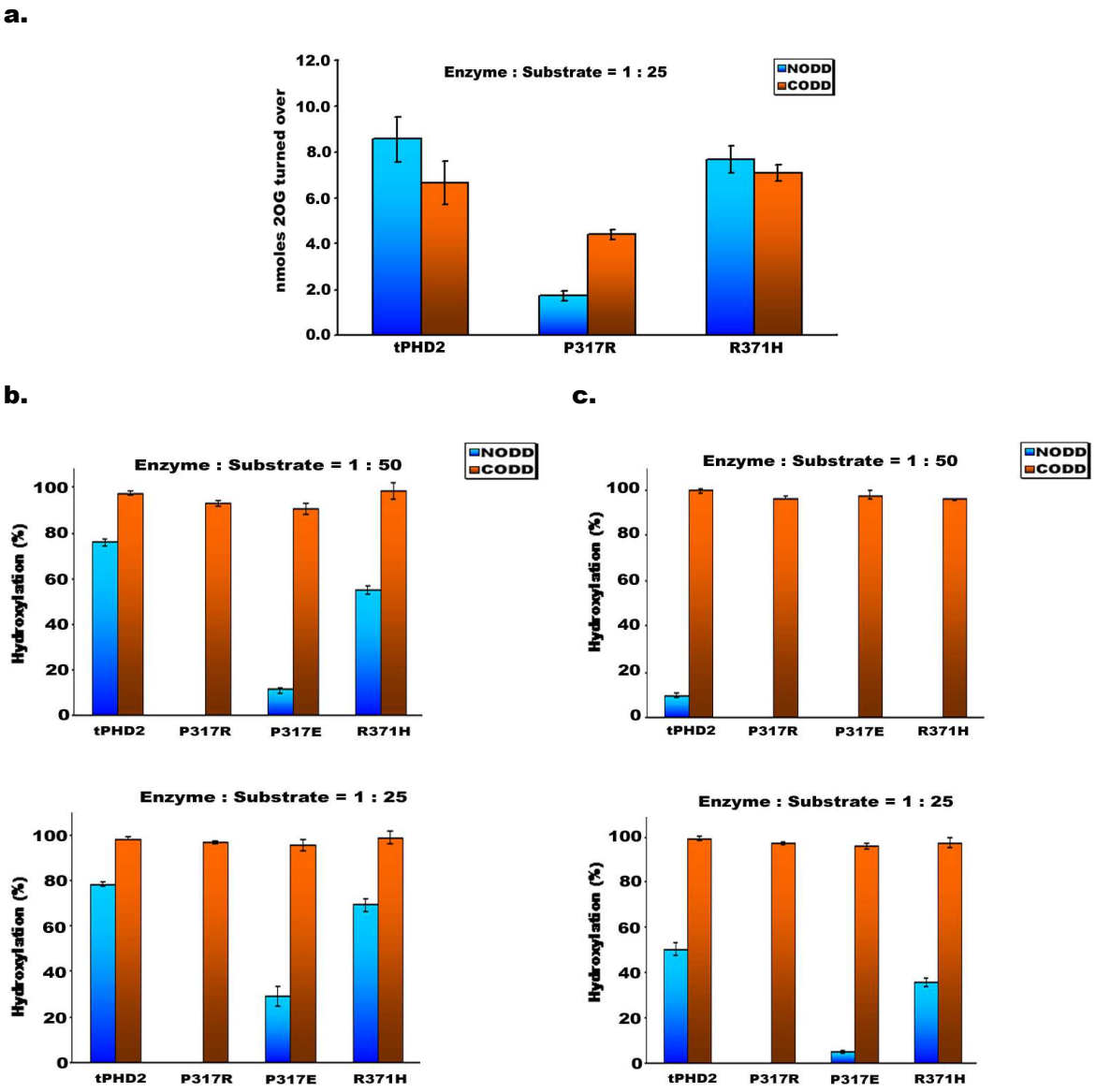


Fig. 4.13 Comparative activity analyses of the clinically observed tPHD2 variants, P317R and R371H along with tPHD2 variant P317E. (a) 2OG turnover activities of P317R and R371H compared with wt tPHD2; (b) substrate hydroxylation as determined by MALDI-TOF mass spectrometry at variable enzyme concentrations (2 and 4 μM) c) substrate hydroxylation in competition experiments using the same enzyme-to-substrate ratios.

4.8 Modelling of PHD3.CODD/NODD complexes

Models for PHD3.Fe(II).2OG.CODD and PHD3.Fe(II).2OG.NODD were generated as described (Section 4.3). The results of MDS studies on tPHD2.Fe(II).2OG.CODD and PHD3.Fe(II).2OG.CODD modelled complexes suggested that tPHD2 and PHD3 may adopt similar folds in solution with significant differences arising in the C-terminal helix α4 reflecting the sequence differences within this region (Fig. 4.14). MDS also predict that the CODD conformations, and hence presumably the binding patterns, will be similar when in complex with tPHD2 and PHD3. However, in contrast to the tPHD2.CODD/NODD and PHD3.CODD complexes, and consistent with the biochemical

and cell based selectivity assay data,[3,4] the PHD3.NODD complex was predicted to be unstable in the MDS analyses (Fig. 4.14).

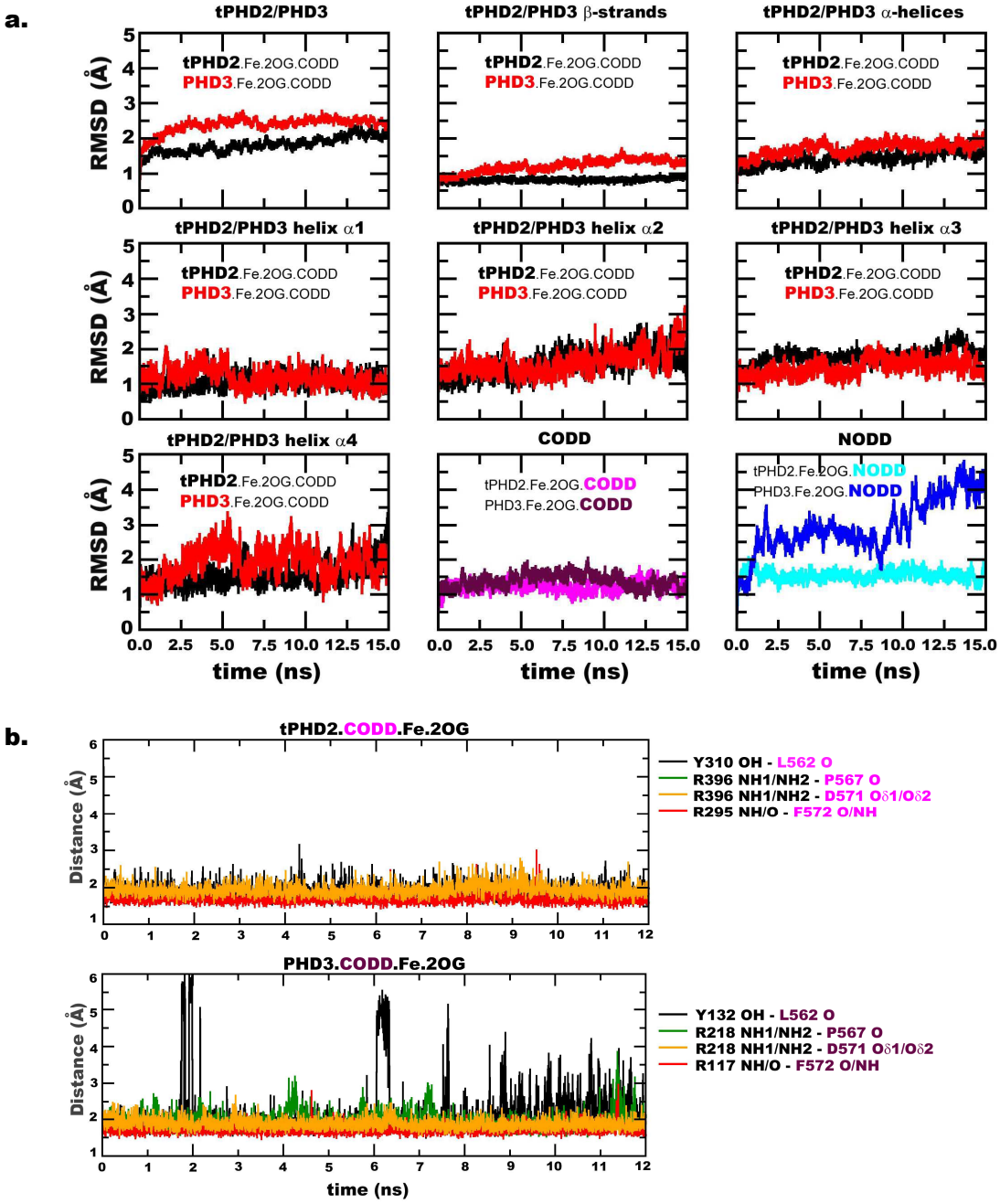


Fig. 4.14 MDS analyses on tPHD2.CODD and PHD3.CODD modelled complexes. The panels in figure (a) show the Cα RMSD from the initial model as a function of simulation time for different structural elements, indicated in the figure titles by the different coloured labels. Figure (b) shows hydrogen bond distances between the tPHD2/PHD3 and CODD residues during the simulations.

4.9 Summary

Analyses of the crystal structures and models, coupled to biochemical assay data, lead to the proposal that interactions involving the 'β2β3/loop' residues, the LXXLAP motif, and $Leu574_{CODD}$ are all involved in determining NODD/CODD selectivity. Together with the biochemical analyses presented, the sequence differences in these regions account for the exclusive selectivity of PHD3 for CODD. Substrate selectivity appears to involve two major determinants: (a) multiple interactions/residues that enable PHD3 to have higher affinity for CODD over other PHDs (e.g. $Gly63_{PHD3}$) and (b) residues in PHD1/PHD3 that directly interfere with NODD binding (e.g. $Leu103_{PHD3}$). Future experiments can be directed towards measuring detailed kinetic parameters for single amino acid variants, notably, tPHD2 S242G, K244R, I251L and R281L, to quantitatively establish their contributions to NODD/CODD selectivity.

4.10 References

1. Bishop, T., Gallagher, D., Pascual, A., Lygate, C.A., de Bono, J.P., Nicholls, L.G., Ortega-Saenz, P., Oster, H., Wijeyekoon, B., Sutherland, A.I., Grosfeld, A., Aragones, J., Schneider, M., van Geyte, K., Teixeira, D., Diez-Juan, A., Lopez-Barneo, J., Channon, K.M., Maxwell, P.H., Pugh, C.W., Davies, A.M., Carmeliet, P. & Ratcliffe, P.J. Abnormal sympathoadrenal development and systemic hypotension in PHD3-/- mice. *Mol. Cell. Biol.* **28**, 3386-3400 (2008).
2. Takeda, K., Ho, V.C., Takeda, H., Duan, L.J., Nagy, A. & Fong, G.H. Placental but not heart defects are associated with elevated hypoxia-inducible factor alpha levels in mice lacking prolyl hydroxylase domain protein 2. *Mol. Cell. Biol.* **26**, 8336-8346 (2006).
3. Epstein, A.C., Gleadle, J.M., McNeill, L.A., Hewitson, K.S., O'Rourke, J., Mole, D.R., Mukherji, M., Metzen, E., Wilson, M.I., Dhanda, A., Tian, Y.M., Masson, N., Hamilton, D.L., Jaakkola, P., Barstead, R., Hodgkin, J., Maxwell, P.H., Pugh, C.W., Schofield, C.J. & Ratcliffe, P.J. C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell* **107**, 43-54 (2001).
4. Hirsila, M., Koivunen, P., Gunzler, V., Kivirikko, K.I. & Myllyharju, J. Characterization of the human prolyl 4-hydroxylases that modify the hypoxia-inducible factor. *J Biol Chem* **278**, 30772-80 (2003).
5. Flashman, E., Bagg, E.A., Chowdhury, R., Mecinovic, J., Loenarz, C., McDonough, M.A., Hewitson, K.S. & Schofield, C.J. Kinetic rationale for selectivity toward N- and C-terminal oxygen-dependent degradation domain substrates mediated by a loop region of hypoxia-inducible factor prolyl hydroxylases. *J Biol Chem* **283**, 3808-15 (2008).
6. Thompson, J.D., Higgins, D.G. & Gibson, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680 (1994).
7. Villar, D., Vara-Vega, A., Landazuri, M.O. & Del Peso, L. Identification of a region on hypoxia-inducible-factor prolyl 4-hydroxylases that determines their specificity for the oxygen degradation domains. *Biochem J* **408**, 231-40 (2007).
8. Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964-72 (2007).
9. Masson, N., Willam, C., Maxwell, P.H., Pugh, C.W. & Ratcliffe, P.J. Independent function of two destruction domains in hypoxia-inducible factor-alpha chains activated by prolyl hydroxylation. *Embo J* **20**, 5197-206 (2001).
10. Berezovsky, I.N., Chen, W.W., Choi, P.J. & Shakhnovich, E.I. Entropic stabilization of proteins and its proteomic consequences. *PLoS Comput Biol* **1**, e47 (2005).
11. Laskowski, R.A. SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J Mol Graph* **13**, 323-30, 307-8 (1995).

12. Morris, A.L., MacArthur, M.W., Hutchinson, E.G. & Thornton, J.M. Stereochemical quality of protein structure coordinates. *Proteins* **12**, 345-364 (1992).
13. Percy, M.J., Furlow, P.W., Beer, P.A., Lappin, T.R., McMullin, M.F. & Lee, F.S. A novel erythrocytosis-associated PHD2 mutation suggests the location of a HIF binding groove. *Blood* **110**, 2193-6 (2007).
14. Kageyama, Y., Koshiji, M., To, K.K., Tian, Y.M., Ratcliffe, P.J. & Huang, L.E. Leu-574 of human HIF-1alpha is a molecular determinant of prolyl hydroxylation. *Faseb J* **18**, 1028-30 (2004).
15. Landazuri, M.O., Vara-Vega, A., Viton, M., Cuevas, Y. & del Peso, L. Analysis of HIF-prolyl hydroxylases binding to substrates. *Biochem Biophys Res Commun* **351**, 313-20 (2006).
16. Percy, M.J., Zhao, Q., Flores, A., Harrison, C., Lappin, T.R., Maxwell, P.H., McMullin, M.F. & Lee, F.S. A family with erythrocytosis establishes a role for prolyl hydroxylase domain protein 2 in oxygen homeostasis. *Proc Natl Acad Sci U S A* **103**, 654-9 (2006).

Chapter 5 – Studies with HIF substrate analogues to investigate differential regulation of NODD and CODD

In this chapter, work using various substrate (CODD and NODD) analogues to further investigate the interactions of NODD and CODD with PHDs is described. Previous studies identified a conserved LXXLAP motif necessary for recognition of HIF-α by pVHL.[1] The current study aims to investigate the role of the LXXLAP motif and the non-conserved C-terminal regions of the substrates.

5.1 Preferential hydroxylation of CODD over NODD

Substrate hydroxylation assays revealed that tPHD2 catalyses almost complete hydroxylation of CODD but ~80% hydroxylation of NODD within a 15 minutes incubation under standard conditions. The catalytic differences were more obvious in substrate competition experiments, where tPHD2 was incubated with equimolar amounts of CODD and NODD (Fig. 5.1). The results from competition experiments revealed that CODD was almost fully hydroxylated with varying extents of NODD hydroxylation observed depending on the tPHD2 concentration. In such competition experiments, tPHD2 catalysed <10% and ~50% NODD hydroxylation at tPHD2 concentrations of 2 and 4 μM, respectively (Fig. 5.1).

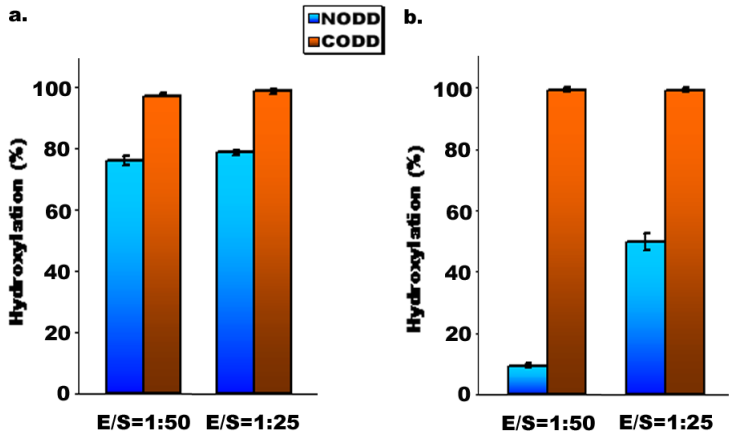


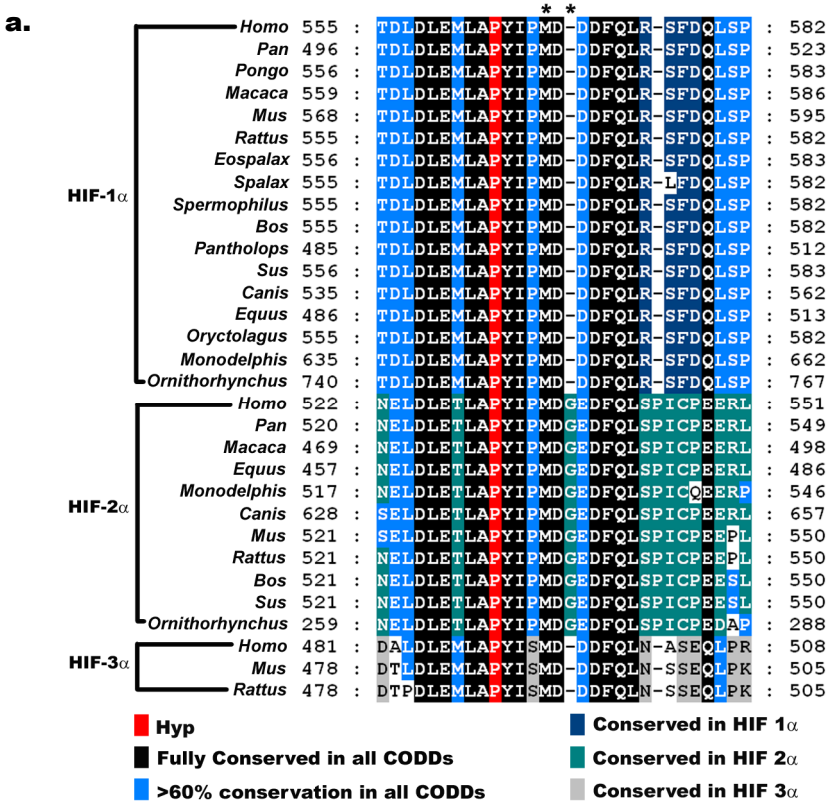
Fig. 5.1 Hydroxylation assay results using tPHD2. (a) tPHD2 (2 and 4 μM) was incubated with NODD and CODD separately. (b) Competition experiment results where tPHD2 was incubated with both CODD and NODD. E/S = enzyme-substrate ratio.

Consistent with the activity assay data, non-denaturing MS analyses revealed stronger complex formation between tPHD2 and CODD than tPHD2 and NODD. In fact, the amount of tPHD2.Fe(II).NODD complex observed after stoichiometric addition of NODD to tPHD2 was comparable (if not less) to that observed for the complex between

tPHD2.Fe(II) and Pro564 hydroxylated CODD ($CODD_{Hyp564}$). Collectively the results of the competition assays and MS studies (Chapter 3, Fig. 3.8a), suggested that because of its weaker binding, NODD hydroxylation may be more sensitive towards the endogenous PHD levels and physiological changes in oxygen levels or the availability of other cofactors such as 2OG and Fe(II). The data support the previous cell-based assays where the hydroxylation of $Pro564_{CODD}$ was found to be the primary target of all three prolyl hydroxylases within the full-length HIF-1α protein.[2] Furthermore, the K_M values of PHD1 and PHD2 for NODD were reported to be significantly higher than those for the CODD.[3,4]

5.2 Significance of the presence of two ODDs in HIF

Although hydroxylation at one of the two HIF-α prolines is sufficient for pVHL-mediated recognition and degradation,[1,5] the presence of two proline hydroxylation sites in the ODD of HIF-1α/-2α suggests a differential role for each site. Mutation of either $Pro402_{NODD}$ or $Pro564_{CODD}$ is sufficient to partially stabilise HIF-1α under normoxic conditions.[1] The two hydroxylation sites are highly evolutionarily conserved across higher organisms (Fig. 5.2), raising the question of why there are two hydroxylation sites and how each site contributes to the overall HIF-1α degradation.



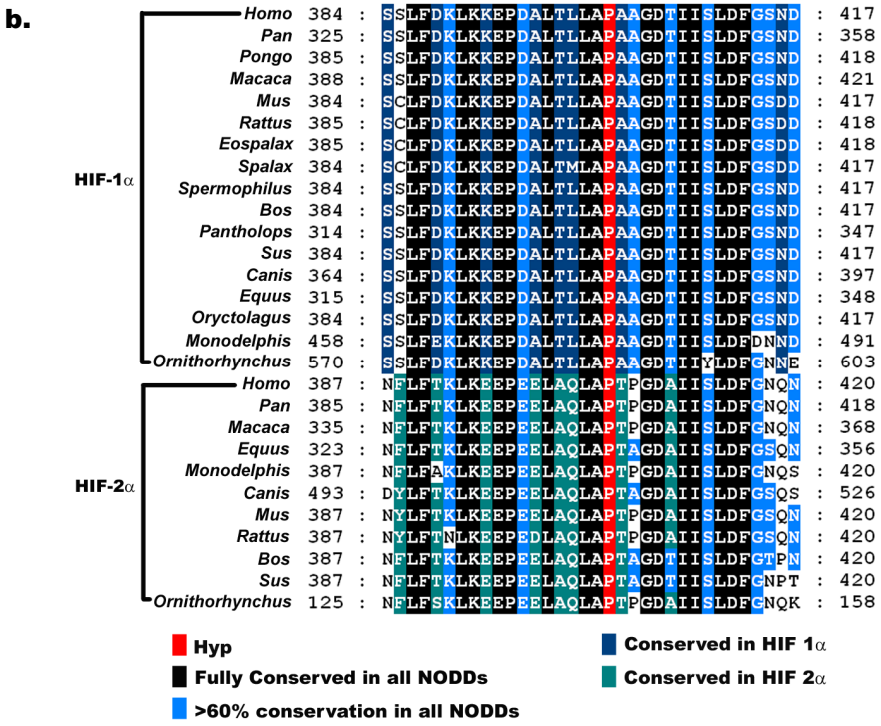


Fig. 5.2 Sequence alignments of ODDs. (a) Sequence alignment of CODD regions of all annotated HIFs across the mammalian species. '*' indicates hotspots for observed clinical mutations in the *HIF2A* gene (Section 5.5); (b) Sequence alignment of NODD regions of mammalian species. Notably, HIF-3α lacks NODD.

Given that CODD is preferentially hydroxylated over NODD and VCB has a remarkably higher (~1000-fold) affinity for hydroxylated CODD than the PHDs (Fig. 5.2), what is the reason for NODD hydroxylation? Further, given that CODD, but not NODD is conserved in many early multicellular organisms, including *Trichoplax* (a proposed ancestor of metazoans) why did NODD hydroxylation evolve at all?

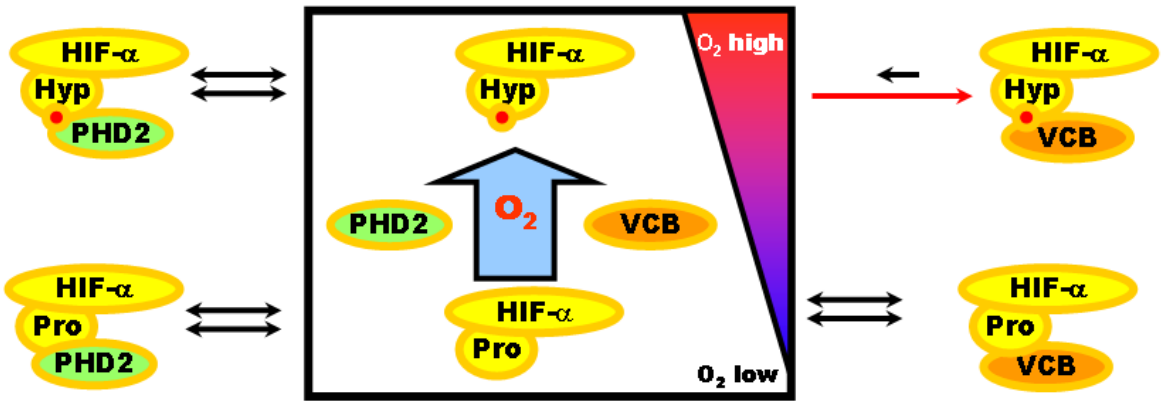


Fig. 5.3 The HIF signalling system showing the role of PHD catalysed prolyl-4-hydroxylation. The ability of VCB to discriminate between hydroxylated and non-hydroxylated HIFα is remarkable, with the difference in K_D for hydroxylated versus non-hydroxylated CODD peptides being ~1000-fold (33 nM versus 34 μM, respectively).[6] PHD2 ≡ PHD2.Fe(II).2OG complex (both Fe(II) and 2OG are tightly bound to PHD2).

5.3 Evidence for coordinated regulation of CODD and NODD hydroxylation

Several groups have attempted to link the independent and non-redundant operation of each hydroxylation event to the regulation of HIF-1α.[1,5] Previous observations, based on immunoblotting studies, have suggested that $Pro564_{CODD}$ hydroxylation affects $Pro402_{NODD}$ hydroxylation.[5] In these studies, the $CODD_{558-565}$ domain was exchanged with the $NODD_{398-407}$ domain in HIF-1α protein to identify the sequences that might cause a preferential hydroxylation of $Pro564_{CODD}$ over $Pro402_{NODD}$.

The domains that Chan et al.[5] swapped (as above) were short and did not include the originally identified sequence for prolyl hydroxylase recognition.[2,4,7] In particular, the three sequential aspartyl residues in CODD, including $Asp571_{CODD}$ (that forms an ionic interaction with tPHD2 in the tPHD2.CODD crystal structure), were not included in the domain swapping experiments. Because the salt bridge interaction imparted significant stability to the overall tPHD2.CODD complex (Chapter 3), a series of hybrid peptides was produced that swapped the LXXLAP motif and the N-terminal residues between NODD and CODD (Fig. 5.4) with the aim of investigating the relative contributions from the LXXLAP motif and the C-terminal non-conserved region.

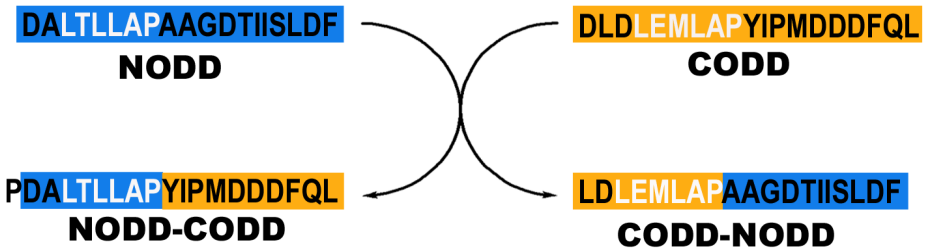


Fig. 5.4 Sequences of NODD, CODD and the hybrid peptides.

MDS studies on tPHD2 complexed with the LXXLAP motifs of CODD and NODD suggested a near identical binding mode (Chapter 4). However, it was considered that switching the C/N-terminal domains in the NODD and CODD hybrid peptides may shed some light on the specific contributions of these domains to the differential affinity for wt tPHD2, as well as the tPHD2 variants that demonstrated PHD3-like selectivity for CODD/NODD hydroxylation.

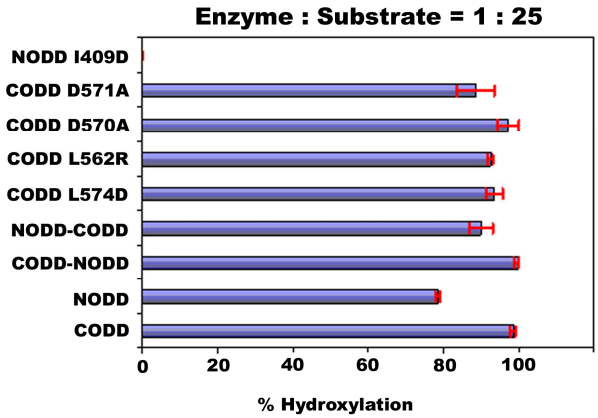


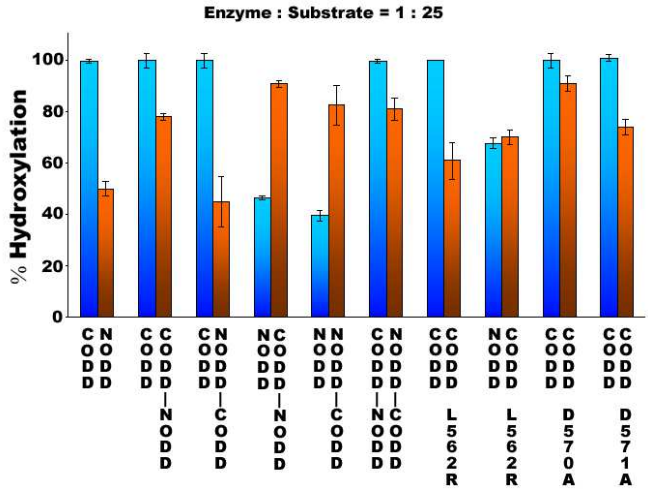
Fig. 5.5 The extent of hydroxylation of variant substrates by tPHD2 as compared to NODD and CODD hydroxylation. Order of hydroxylation: CODD = CODD-NODD > NODD-CODD > NODD.

5.3.1 Hydroxylation of the 'NODD-CODD' hybrid peptide

Substituting residues C-terminus to the NODD LXXLAP with those of CODD to give 'NODD-CODD' (Fig. 5.4) appeared to slightly increase hydroxylation of the hybrid 'NODD-CODD' peptide, compared to wt NODD. This may be due to the introduction of a stabilising $D571_{CODD}$-$R396_{PHD2}$ salt-bridge as observed in the tPHD2.CODD structure in the analogous spatial position (Fig. 5.5).

5.3.2 Hydroxylation of the 'CODD-NODD' hybrid peptide

Unlike NODD, the 'CODD-NODD' hybrid peptide was completely hydroxylated (Fig. 5.5). This is interesting because the wt NODD was only <80% hydroxylated under standard assay conditions. These observations suggest that the non-conserved residues in the CODD-LXXLAP sequence ($Glu560_{CODD}$ and $Met561_{CODD}$) might increase binding of CODD with tPHD2 compared to NODD. Consistent with this data, substitution of $Glu560_{CODD}$ and $Met561_{CODD}$ by almost all amino acids including $Thr398_{NODD}$ and $Leu399_{NODD}$ that are in the analogous positions in NODD, was previously shown to cause reduced 2OG turnover in radioactive 2OG decarboxylation assays.[8]



% Hydroxylation in competition assays		Binding affinity	Relatively more stabilising interactions
DLDLEMLAPYIPMDDDFQL CODD (100%) >	LDLEMLAPAAGDTIISLDF CODD-NODD (78%)	CODD > CODD-NODD	$Asp571_{CODD}$-$Arg396_{PHD2}$ salt-bridge
DLDLEMLAPYIPMDDDFQL CODD (100%) >	PDALTLLAPYIPMDDDFQL NODD-CODD (45%)	CODD > NODD-CODD	CODD-LXXLAP hydrophobic
LDLEMLAPAAGDTIISLDF CODD-NODD (100%) >	PDALTLLAPYIPMDDDFQL NODD-CODD (81%)	NODD-CODD > CODD-NODD	CODD-LXXLAP hydrophobic

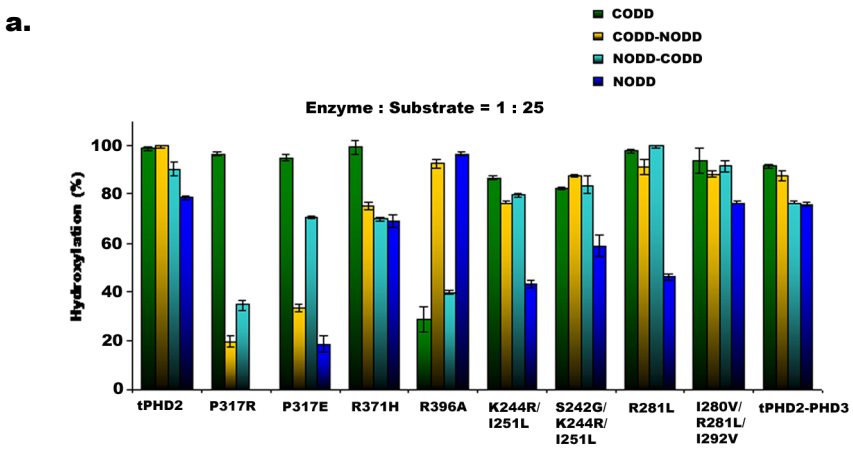
Fig. 5.6 Competition experiments where pairs of substrates were incubated with tPHD2. Notably, despite having an aspartate in the analogous position to $Asp571_{CODD}$ (that forms salt-bridge with $Arg396_{PHD2}$), NODD-CODD was hydroxylated to a lesser extent when in competition with CODD-NODD, thus indicating a more significant role of CODD-LXXLAP over the salt-bridge interaction. Data rationalise at least in part, why a subtle mutation in the LXXLAP interacting PHD2 residue (such as Ile251 to Leu as in PHD3) could make a significant difference in activity (Chapter 4, Fig. 4.3).

In the competition experiments (Fig. 5.6), it was evident that CODD was preferentially hydroxylated over both the hybrid peptides (CODD-NODD and NODD-CODD) and the single amino acid variants of CODD (see below). The preferential hydroxylation of CODD over CODD-NODD hybrid (lacking $Asp571_{CODD}$) once again supports a stabilising role of the $R396_{PHD2}$-$D571_{CODD}$ interaction. Although NODD could make an ion-pair interaction ($Asp412_{NODD}$-$Arg281_{PHD2}$) as predicted by the MDS analyses on tPHD2.NODD modelled complexes and mutagenesis studies (Chapter 4), non-denaturing MS showed CODD binds >10 fold stronger to tPHD2 than CODD-NODD. However, CODD-NODD was more efficiently hydroxylated when in competition with NODD-CODD despite having an $Asp571_{CODD}$ that can form salt-bridge as in the case of CODD (Fig. 5.6). The superior hydroxylations of CODD and CODD-NODD over NODD-CODD led to the proposal that the NODD-LXXLAP is not suitable for optimum interaction with tPHD2. These findings collectively suggest that the LXXLAP itself, as well as the C-terminal part of the two ODDs, is responsible for differential affinity between NODD and CODD.

Several point variations on the CODD peptide were also produced to investigate their effects on tPHD2 mediated hydroxylation. Although all the CODD variants were hydroxylated almost completely after a standard 15 minutes incubation (Fig. 5.5), consistent with the previous observations,[8] variable degrees of hydroxylation were observed in the competition assays with wt CODD (Fig. 5.6). While substitution of $Asp570_{CODD}$ (part of the DDD motif) resulted in almost complete hydroxylation when in competition with wt CODD, the D571A CODD variant ($Asp571_{CODD}$ forms a salt-bridge) was <80% hydroxylated. A L562R (part of the LXXLAP) variant was proposed to increase CODD binding in cases of both PHD2 and to a lesser extent, PHD3, based on a yeast-two hybrid assay.[9] However, this variant was hydroxylated with lower efficiency than CODD and with almost equal efficiency as NODD in competition assays (Fig. 5.6).

5.4 Catalytic differences observed using CODD analogues and tPHD2 variants

tPHD2 variants that were observed to show variable CODD and NODD hydroxylation (Chapter 4) were then tested with the variant peptide substrates (Fig. 5.7). The order of hydroxylation for these substrate analogues against tPHD2 variants is summarised in Table 5.1. Collectively, these activity results support the observation that the peptides containing CODD-LXXLAP (CODD and CODD-NODD) resulted in preferred binding and/or hydroxylation when compared to those containing NODD-LXXLAP.



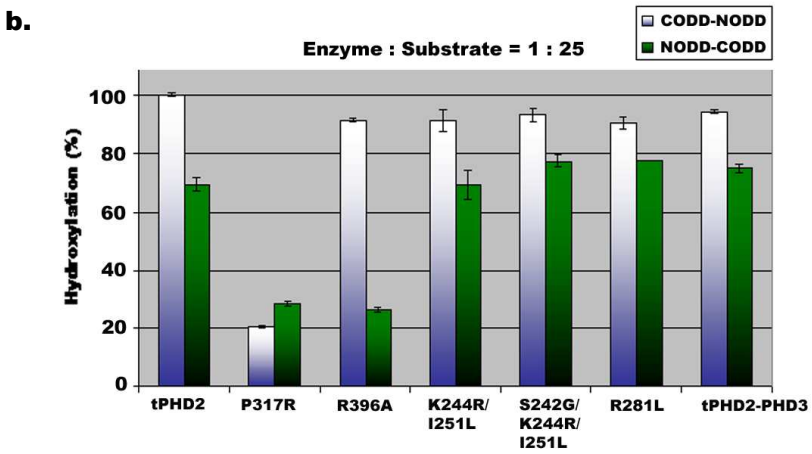


Fig. 5.7 Hydroxylation of CODD/NODD and the hybrid peptide substrates using tPHD2 wt and variants. (a) Individual substrate experiment and (b) competition experiments when both CODD-NODD and NODD-CODD were incubated with wt and variant tPHD2.

tPHD2 variants were also investigated for differential hydroxylation in competition assays with the hybrid peptides CODD-NODD and NODD-CODD. The results showed that, as with wt tPHD2, all the variants except for R396A demonstrated preferred hydroxylation/binding for CODD-NODD relative to NODD-CODD. The R396A variant caused ~92% and ~26% hydroxylation of CODD-NODD and NODD-CODD, respectively. This compares with 100% and ~70% with wt tPHD2. Due to lack of an $Arg396_{PHD2}$, the R396A variant is probably unable to form a salt-bridge interaction with one of the three sequential aspartates at the C-terminal part of CODD, explaining the significantly reduced hydroxylation of NODD-CODD.

Table 5.1 **Summary of hydroxylation assay results showing that the order of hydroxylation supports the proposal that CODD-LXXLAP is the preferred part of the substrate by tPHD2 and the variants.**

Enzyme	Hydroxylation order	Inferences
tPHD2	CODD = CODD-NODD > NODD-CODD > NODD	Consistent with the proposal that CODD-LXXLAP was preferred over NODD-LXXLAP
P317R	CODD >> NODD-CODD > CODD-NODD > NODD	Glu317 (in P317E), but not Arg317 (in P317R) could form a hydrogen bond with $Thr398_{NODD}$ and result in a small amount of NODD hydroxylation.
P317E	CODD > NODD-CODD > CODD-NODD > NODD	
R371H	CODD > CODD-NODD > NODD-CODD > NODD	CODD-LXXLAP was preferred over NODD-LXXLAP
R396A	NODD = CODD-NODD > NODD-CODD > CODD	Lost salt-bridge with $Asp571_{CODD}$ supporting the important role of this interaction.
K244R/I251L	CODD > NODD-CODD > CODD-NODD > NODD	Lost binding to both NODD and CODD LXXLAP; CODD LXXLAP was better accepted.
S242G/K244R/I251L	CODD > NODD-CODD > CODD-NODD > NODD	Similar to K244R/I251L
R281L	CODD > NODD-CODD > CODD-NODD > NODD	$Arg281_{PHD2}$ (Leu in PHD3) interacts with NODD C-terminal Asp412, as predicted by modelling studies.

I280V/R281L/I292V	CODD > NODD-CODD > CODD-NODD > NODD	Hydroxylation was less affected compared to single mutation R281L possibly due to an altered size of the hydrophobic pocket (Chapter 4).
PHD2-3 chimera	CODD = CODD-NODD > NODD-CODD = NODD	CODD-LXXLAP was preferred over NODD-LXXLAP

5.5 HIF-2α clinical mutations

Four heterozygous mutations were recently identified in exon 12 of the *HIF2A* gene leading to translation of HIF-2α M535V, M535I, G537R and G537W variants.[10-12] These variants have been linked to idiopathic erythrocytosis and in all cases, the patients were diagnosed with elevated EPO levels. Mutations at Gly537 were more frequent than Met535, occurring in ~80% of patients with erythrocytosis.[11] pVHL capture assays, hydroxylation assays coupled to MALDI-MS as well as immunoblotting revealed that the G537W mutation reduces hydroxylation of HIF-2α$_{527-542}$ by PHD2 (full-length), as well as the subsequent recognition of HIF-2α by pVHL.[12]

To further investigate whether any of these clinical mutations of HIF-2α had an effect the extent of hydroxylation, M535V, G537R and G537W variants of the HIF-2α CODD$_{523-543}$ were produced. In contrast to previous studies, none of the tested variants of HIF-2α (M535V, G537R and G537W) appeared to display altered tPHD2 mediated hydroxylation. Care must be taken to compare results from different types of experiments, especially when the G537W HIF-2α variant was tested against a full-length PHD2 in the previous assays.[12] In contrast, in the current study, the variant peptide was tested against a truncated form of PHD2 that lacks the N-terminal MYND finger; the truncated form was found to be more active than the full-length PHD2 using CODD substrates (Chapter 2).

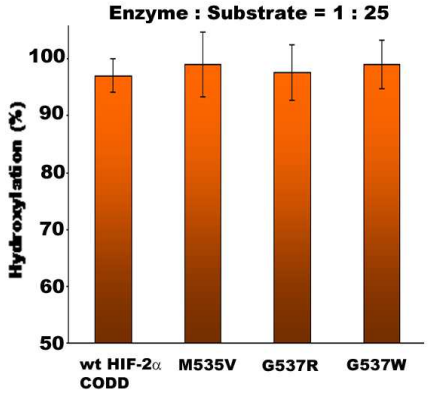


Fig. 5.8 Hydroxylation assays using tPHD2 with wt HIF-2α and the variants M535V, G537R and G537W showed that the HIF-2α variants were able to retain full hydroxylation activity.

Sequence analyses revealed that Met535 in HIF-2α is conserved across all the mammalian species and in HIF-1α isoforms ($Met568_{CODD}$ in human HIF-1α) (Fig. 5.2a). In the tPHD2.CODD crystal structure, $Met568_{CODD}$ was modelled in two alternate conformations that made hydrophobic contacts with $Tyr403_{PHD2}$. The HIF-2α M535V variant was hydroxylated similarly to the wt HIF-2α (Fig. 5.8). Therefore, one may speculate that substituting a methionine with a hydrophobic valine or isoleucine may have relatively little effect on catalytic activity.

In contrast, Gly537 is conserved only in HIF-2α amongst across many species (Fig. 5.8); sequence analyses further reveal that the residue appears to be an insertion between the two C-terminal sequential aspartates in HIF-1α/-3α (Fig. 5.8). In the tPHD2.CODD crystal structure, $Asp569_{CODD}$ and $Asp570_{CODD}$ that would surround the 'extra' glycine have relatively high B-factors indicating conformational flexibility in this part of the structure. Furthermore, the MDS analyses indicated that this glycine could be accommodated whilst maintaining interactions between PHD2 and the ODD (Fig. 5.9). Therefore, this part of the structure could be sufficiently flexible to accommodate such bulky amino acids as tryptophan or arginine. These observations were consistent with the hydroxylation assay results that showed no significant variation in hydroxylation of G537R and G537W HIF-2α variants using tPHD2 (Fig. 5.9).

Thus, at this time there is a discrepancy between the results with the G537W HIF-2α peptide presented here and those previously reported with relatively shorter HIF fragments and a full-length PHD2.[12] However, more detailed kinetic analyses may reveal differences not apparent in the end-point assays described here.

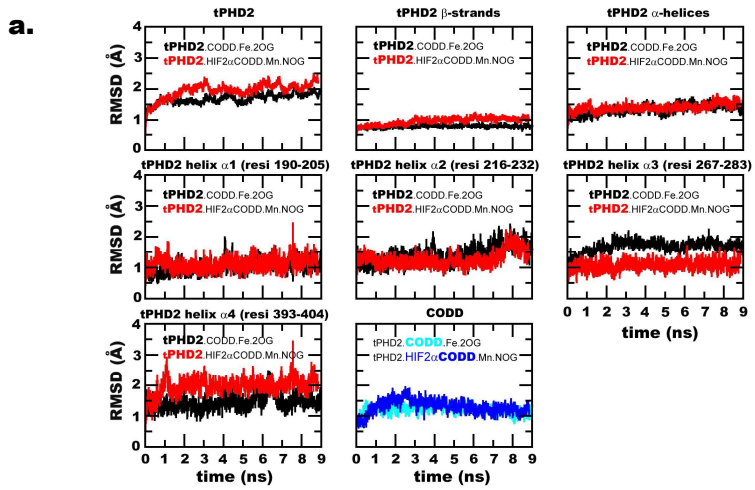
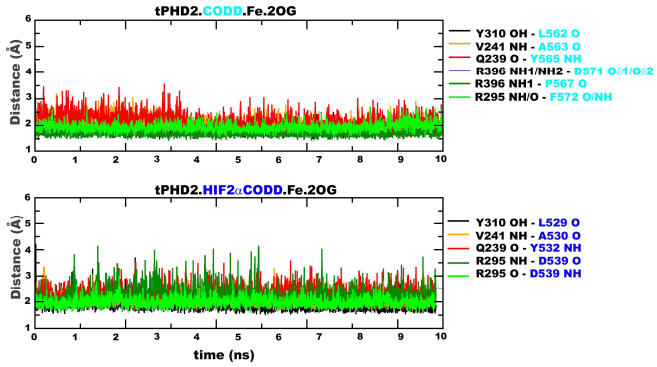


Fig. 5.9 MDS studies on tPHD2.Fe.2OG.CODD and tPHD2.Fe.2OG.HIF-2α CODD modelled complexes. The panels in (a) show the RMSD from the initial model as a function of simulation time for different structural elements as indicated in the figure titles by the different coloured labels. (b) shows hydrogen bond distances between tPHD2 and HIF-1α/-2αCODD residues during simulations.

b.



5.6 Summary

Structural and biochemical studies described in Chapter 3 identified a salt-bridge interaction between $Asp571_{CODD}$ and $Arg396_{PHD2}$ that plays an important role in substrate binding and hydroxylation. In addition, substrate analogue studies presented here enabled identification of LXXLAP of CODD as a region that is optimal for binding with PHD2.

5.7 References

1. Masson, N., Willam, C., Maxwell, P.H., Pugh, C.W. & Ratcliffe, P.J. Independent function of two destruction domains in hypoxia-inducible factor-alpha chains activated by prolyl hydroxylation. *Embo J* **20**, 5197-206 (2001).
2. Epstein, A.C., Gleadle, J.M., McNeill, L.A., Hewitson, K.S., O'Rourke, J., Mole, D.R., Mukherji, M., Metzen, E., Wilson, M.I., Dhanda, A., Tian, Y.M., Masson, N., Hamilton, D.L., Jaakkola, P., Barstead, R., Hodgkin, J., Maxwell, P.H., Pugh, C.W., Schofield, C.J. & Ratcliffe, P.J. C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell* **107**, 43-54 (2001).
3. Flashman, E., Bagg, E.A., Chowdhury, R., Mecinovic, J., Loenarz, C., McDonough, M.A., Hewitson, K.S. & Schofield, C.J. Kinetic rationale for selectivity toward N- and C-terminal oxygen-dependent degradation domain substrates mediated by a loop region of hypoxia-inducible factor prolyl hydroxylases. *J Biol Chem* **283**, 3808-15 (2008).
4. Hirsila, M., Koivunen, P., Gunzler, V., Kivirikko, K.I. & Myllyharju, J. Characterization of the human prolyl 4-hydroxylases that modify the hypoxia-inducible factor. *J Biol Chem* **278**, 30772-80 (2003).
5. Chan, D.A., Sutphin, P.D., Yen, S.E. & Giaccia, A.J. Coordinate regulation of the oxygen-dependent degradation domains of hypoxia-inducible factor 1 alpha. *Mol Cell Biol* **25**, 6415-26 (2005).
6. Hon, W.C., Wilson, M.I., Harlos, K., Claridge, T.D., Schofield, C.J., Pugh, C.W., Maxwell, P.H., Ratcliffe, P.J., Stuart, D.I. & Jones, E.Y. Structural basis for the recognition of hydroxyproline in HIF-1 alpha by pVHL. *Nature* **417**, 975-8 (2002).
7. Huang, J., Zhao, Q., Mooney, S.M. & Lee, F.S. Sequence determinants in hypoxia-inducible factor-1alpha for hydroxylation by the prolyl hydroxylases PHD1, PHD2, and PHD3. *J Biol Chem* **277**, 39792-800 (2002).
8. Li, D., Hirsila, M., Koivunen, P., Brenner, M.C., Xu, L., Yang, C., Kivirikko, K.I. & Myllyharju, J. Many amino acid substitutions in a hypoxia-inducible transcription factor (HIF)-1alpha-like peptide cause only minor changes in its hydroxylation by the HIF prolyl 4-hydroxylases: substitution of 3,4-dehydroproline or azetidine-2-carboxylic acid for the proline leads to a high rate of uncoupled 2-oxoglutarate decarboxylation. *J Biol Chem* **279**, 55051-9 (2004).
9. Landazuri, M.O., Vara-Vega, A., Viton, M., Cuevas, Y. & del Peso, L. Analysis of HIF-prolyl hydroxylases binding to substrates. *Biochem Biophys Res Commun* **351**, 313-20 (2006).
10. Martini, M., Teofili, L., Cenci, T., Giona, F., Torti, L., Rea, M., Foa, R., Leone, G. & Larocca, L.M. A novel heterozygous HIF2AM535I mutation reinforces the role of oxygen sensing pathway disturbances in the pathogenesis of familial erythrocytosis. *Haematologica* **93**, 1068-71 (2008).

11. Percy, M.J., Beer, P.A., Campbell, G., Dekker, A.W., Green, A.R., Oscier, D., Rainey, M.G., van Wijk, R., Wood, M., Lappin, T.R., McMullin, M.F. & Lee, F.S. Novel exon 12 mutations in the HIF2A gene associated with erythrocytosis. *Blood* **111**, 5400-2 (2008).
12. Percy, M.J., Furlow, P.W., Lucas, G.S., Li, X., Lappin, T.R., McMullin, M.F. & Lee, F.S. A gain-of-function mutation in the HIF2A gene in familial erythrocytosis. *N Engl J Med* **358**, 162-8 (2008).

Chapter 6 – PHD2 active site chemistry and further studies towards investigating the oxygen sensing properties of PHD2

The K_M value for oxygen (250 μM) for PHD2 is significantly higher than the binding constants for 2OG (60 μM) and long HIF-α fragments (<5 μM), a property proposed to be important in the role of PHD2 as an oxygen sensor.[1-3] PHD2 is also highly unusual amongst studied 2OG oxygenases in that it forms a long-lived complex with Fe(II) and 2OG under aerobic conditions.[4] Biochemical assays and the structural work described here identify features that enable this unusual behaviour, including the mobile loop that encloses the active site and limits access (Chapter 3), the position of the 2OG 1-carboxylate, and a water molecule positioned to be tightly bound to the metal. Further, the studies imply that a precise stereoelectronic arrangement, involving the C^4-endo $Pro564_{CODD}$ conformation is required for productive reaction between the proposed ferryl oxo (Fe^{IV}=O) intermediate and HIF-1α substrate.

6.1 Active site, 2OG and metal binding

Consistent with sequence and mutation analyses,[5-7] the tPHD2.CODD crystal structure revealed that the metal ion of PHD2 is coordinated by a facial triad of residues (His313, Asp315, His374); the NOG binds to the metal in a bidentate manner with a water molecule completing an octahedral coordination geometry (Fig. 6.1). The NOG coordinates the metal via its oxalate oxygen (O-Mn(II); 2.4 Å; *trans* to His374 Nε2 (Nε2-Mn(II); 1.9 Å)) and the amide α-carbonyl oxygen (Oα-Mn(II); 2.1 Å) forming a five membered chelate ring. MDS studies predict that, in solution, Fe(II) and 2OG will bind tPHD2 in a similar manner as observed for Mn(II) and NOG in the tPHD2.CODD structure (Chapter 3).

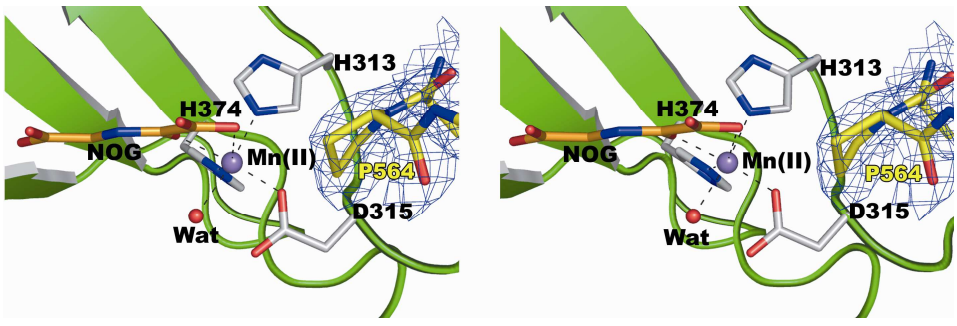


Fig. 6.1 Stereoview of CODD (yellow sticks) binding at the tPHD2 active site showing the experimentally observed electron density around $Pro564_{CODD}$ in a simulated annealing $2F_o$-F_c OMIT map contoured to 1σ; CODD atoms were left out of the map calculation.

The carboxylate of NOG is bound to $Arg383_{PHD2}$ and $Tyr329_{PHD2}$ side chains through electrostatic and hydrogen bond interactions (Arg383 NH1-O3, 3.0 Å; Arg383 NH2-O1, 2.5 Å; and Tyr329 OH-O1, 2.4 Å). The R383A tPHD2 variant did not catalyse 2OG turnover/hydroxylation (<5%),[8] whereas Y329F tPHD2 retained almost complete hydroxylation activity (Fig. 6.2) indicating that the ion-pair between $Arg383_{PHD2}$ and 2OG carboxylate was sufficient to sustain catalysis.

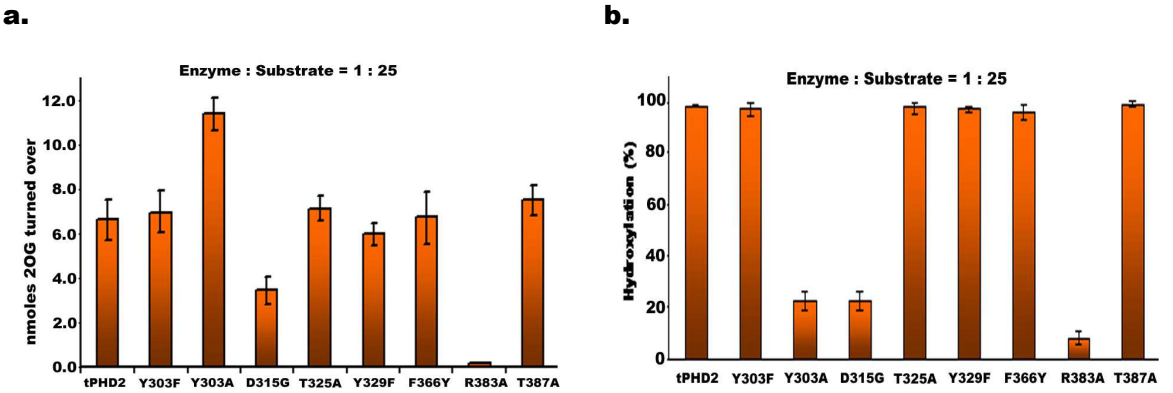


Fig. 6.2. Activity analyses of tPHD2 variants where residues around the active site including those involved in Fe(II) and 2OG binding were modified. (a) 2OG turnover and (b) substrate hydroxylation as determined by MALDI-TOF mass spectrometry using 4 μM enzyme and 100 μM CODD.

6.2 Water molecule at the active site iron – implications for O_2 binding

The observation of a metal bound water molecule *trans* to $His313_{PHD2}$ (Nε2-Mn(II); 2.0 Å; H_2O-Mn(II), 2.2 Å), in both ($P6_3$) and ($P2_12_12_1$) crystal forms suggested that oxygen binding in this position might occur with subsequent loss of CO_2 and formation of a ferryl intermediate (for which there is evidence with other 2OG oxygenases[7,9,10]) to enable $Pro564_{CODD}$ hydroxylation. Support for the formation of a PHD2.Fe(II).2OG.HIF-1α complex in catalysis, analogous to that observed crystallographically, came from non-denaturing MS studies where a tPHD2.Fe(II).2OG.CODD complex had been observed (Chapter 3, Fig. 3.8a).

6.2.1 Iron-bound oxygen molecule in JMJD2A indicates an O_2 delivery system

A recently reported structure of JMJD2A complexed with $H3K36Me_3$ (PDB ID 2P5B) provided evidence of a dioxygen coordinating the active site iron.[11] To date, this is the only 2OG oxygenase structure with an oxygen molecule bound supporting previous proposals of replacing a metal coordinated water during catalysis.[7] It was proposed that the O_2 molecule

could be recruited by $Lys241_{JMJD2A}$ and delivered to the iron to complete a catalytic cycle. $Lys241_{JMJD2A}$ is positioned on a long hairpin-like structure (containing an α helix and 3 short α helices) that connects strand IV and V of the DSBH, and is conserved in JMJDs. A similar role for a lysine was observed in another non-heme iron enzyme, superoxide reductase (SOR).[12] The evolutionarily conserved $Lys48_{SOR}$ is part of a flexible-loop; when the loop is in the 'closed' conformation, the residue forms a hydrogen bond network involving oxygen and two water molecules. The 'open' conformation of the loop prevents $Lys48_{SOR}$ from interacting with the oxygen atom.[12]

These observations raised the possibility that oxygen may similarly require binding to a protein residue in tPHD2 prior to activating the iron centre. Neither of the tPHD2 structures with substrate bound or unbound, provided evidence of direct attachment of molecular oxygen to the iron or a protein residue. The flexible β2β3/loop has two lysines, $Lys244_{PHD2}$ and $Lys249_{PHD2}$; $Lys244_{PHD2}$ makes hydrophobic contact with $Met561_{CODD}$ (Chapter 3). None of the lysines in tPHD2 are close to the active site; moreover, methylation of lysines did not ablate catalytic activity (Chapter 7) indicating the lysines are probably not involved in binding oxygen as proposed for JMJD2A and SOR.

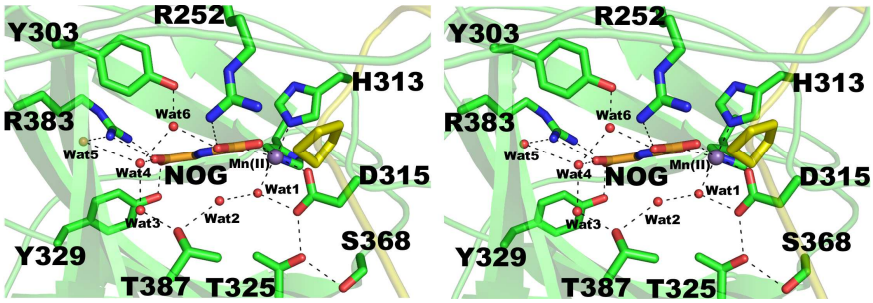


Fig. 6.3 Stereoview ribbons representation of the tPHD2 active site (green) in tPHD2.CODD complex showing a water network and associated tPHD2 residues. CODD, NOG and Mn(II) are shown as yellow and orange sticks and violet sphere, respectively.

Other residues near the tPHD2 active site that could possibly form hydrogen bond to oxygen were Tyr303, Thr387, Thr325 and Arg252. These residues are conserved across all the species. Using site-directed mutagenesis, R252A, Y303F, Y303A, T387A and T325A variants were produced and tested for their ability to turnover 2OG. Since 2OG turnover has been shown in this work not always to correlate with substrate hydroxylation, the variants were also tested for their ability to hydroxylate CODD substrate. The results revealed that, except for R252A and Y303A, all other variants were equally active as wt tPHD2 in both assays under standard conditions (Fig. 6.2).

In both substrate bound/unbound structures of tPHD2, $Thr387_{PHD2}$ connects a chain of five hydrogen-bonded water molecules that extends from the iron towards the 2OG C5-carboxylate binding site (in order from Fe–Wat1–Wat2–Thr387–Wat3–Wat4–Wat5, Fig. 6.3). $Thr325_{PHD2}$ forms part of a hydrogen bond network between Wat1-Asp315 (active site)-Thr325 that is observed in other 2OG oxygenases (DAOCS[13,14], ANS[15] and an algal collagen prolyl hydroxylase[16]).

$Tyr303_{PHD2}$ was reported to hydrogen bond to a hydroxyl group of compounds A/B in the tPHD2.Fe(II).A/B crystal structures.[8] In the tPHD2.CODD structure, the residue forms a hydrogen bond network with a water molecule (Wat6 in Fig. 6.3) and NOG. The Y303F variant retained full enzymatic activity and was inhibited by compound A, indicating that the hydroxyl of $Tyr303_{PHD2}$ is not catalytically essential.[8] In contrast, the Y303A variant was found to increase 2OG turnover mostly in an uncoupled manner, as the variant catalysed only <20% substrate hydroxylation. The increased 2OG turnover of Y303A variant may be due to increased volume around the active site iron (Fig. 6.2) as observed for other variants such as tPHD2Δloop (Chapter 2).

The proposed multiple roles of $Arg252_{PHD2}$ in substrate/2OG binding, including stabilising a closed lid conformation, are described in Chapter 3. The R252A variant showed a marked reduction in activity in both 2OG turnover and hydroxylation assays (Chapter 2). Thus, in addition to its role in stabilising the finger-lid conformation and 2OG binding, $Arg252_{PHD2}$ may be involved in oxygen binding. However, the structural and activity studies using R252A variants coupled to MS analyses described here are not sufficient to conclusively establish its role as an oxygen carrier.

6.2.2 Difficulty in replacing the metal bound water

In the tPHD2.CODD structure, the metal-bound water (*trans* to His313, Wat1), has a 2-fold lower B-factor (17.2 $Å^2$) than the average for all the crystallographic waters (34.3 $Å^2$) thus suggesting less plasticity (Fig. 6.4). Wat1 is positioned to hydrogen-bond to the non-coordinating oxygen of $Asp315_{PHD2}$, apparently securing the Fe(II) in a 6-coordinate state (6C) that is proposed to prevent substrate uncoupled oxidation (Fig. 6.5).[17,18] In the asparaginyl hydroxylase, factor inhibiting HIF (FIH), Fe(II) is in a 5-coordinate state (5C) upon substrate binding as observed in the crystal structures,[19] possibly due to the lack of an internal hydrogen bond from the monodentate $Asp201_{FIH}$ to a water ligand (Fig. 6.5). The

non-coordinating oxygen of $Asp201_{FIH}$ binds to a backbone amide of its substrate (HIF-1α CAD) instead.[19]

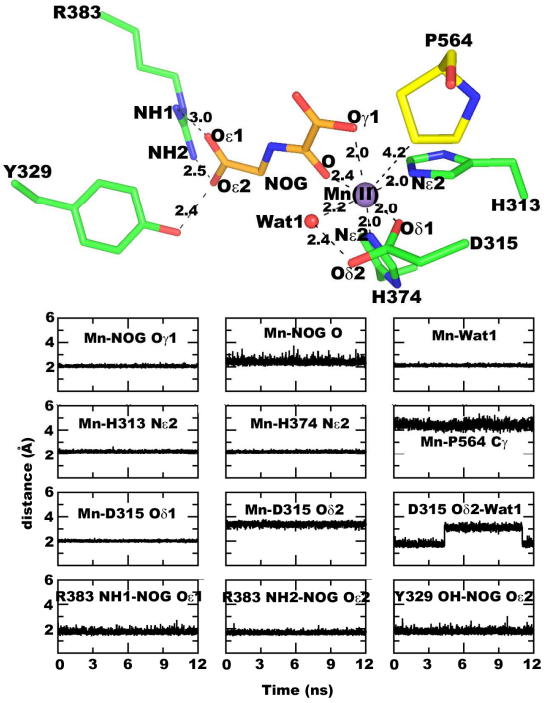


Fig. 6.4 MDS analyses on the tPHD2.CODD crystal structure showing the bond/atom distances as a function of simulation time. Notably, the water molecule (Wat1) at the active site was predicted to be tightly coordinated with the metal ion whilst the hydrogen with the non-coordinating $Asp315_{PHD2}$ Oδ2 could be flexible.

The effect of variation of $Asp315_{PHD2}$ was then investigated. The results showed that a D315G tPHD2 variant still could catalyse ~50% 2OG turnover compared to the wt tPHD2 in the presence of CODD; however, the hydroxylation of CODD was significantly reduced (~20%) after an incubation of 15 minutes (Fig. 6.2). In contrast, it was recently reported that whilst FIH D201A and D201E variants did not catalyse HIF-1α hydroxylation, though displayed an enhanced 2OG turnover compared to the wt FIH, the FIH D201G variant catalysed both 2OG turnover and asparaginyl hydroxylation at comparable levels to wt FIH.[20]

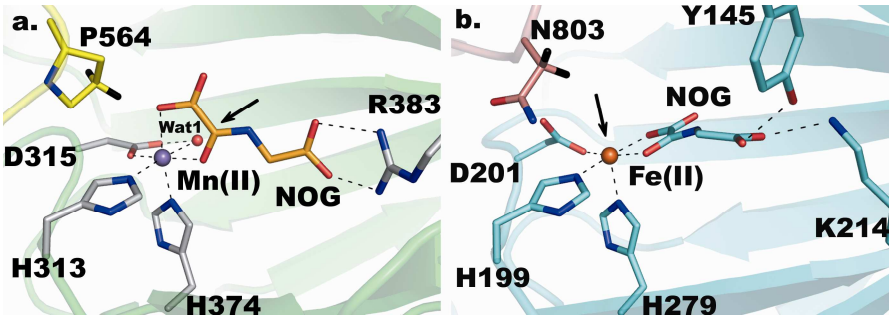


Fig. 6.5 The single Mn(II) ion (violet sphere) is octahedrally coordinated by a facial triad of tPHD2 residues, (His313, Asp315, His374), the 2OG analogue NOG (orange sticks) in a bidentate manner and a water molecule. The metal bound water (Wat1) is positioned *trans* to His313 (Nε2-Mn(II); 2.0 Å; H_2O-Mn(II), 2.2 Å) and hydrogen bonded with Asp315 Oδ2 (Oδ2-Wat1; 2.4 Å). Asp315 Oδ1 coordinates the metal (2.0 Å) approximately *trans* to the amide carbonyl (Oα) of NOG. Arrows indicate proposed oxygen binding site.

Spectroscopic and electronic structural studies on the non-haem Fe(II) containing extradiol dioxygenase 1,2-dihydroxybiphenyl dioxygenase (1,2-DHBD) revealed that 1,2-DHBD is five-coordinate (5C) due to the steric interaction of a tyrosine residue ($Tyr250_{DHBD}$) near the iron site (Fig. 6.6).[18,21] By introducing a similar steric clash with the Fe(II)-bound water in tPHD2, it could be possible to weaken the water binding. This may facilitate a 6C→5C conversion and subsequent oxygen binding to enable tPHD2 to catalyse a higher substrate hydroxylation.

With this aim of weakening water binding at the active site, $Phe366_{PHD2}$ was mutated to tyrosine. The conserved $Phe366_{PHD2}$ residue was targeted because the phenyl ring projects towards the metal bound water (Phe366 Cζ-Mn(II), 4.6 Å; Phe366 Cζ-Wat1, 3.3 Å) in tPHD2 crystal structures and therefore, an introduction of a hydroxyl group (C-O, 1.3 Å) might be sufficient to cause a steric clash with Wat1.

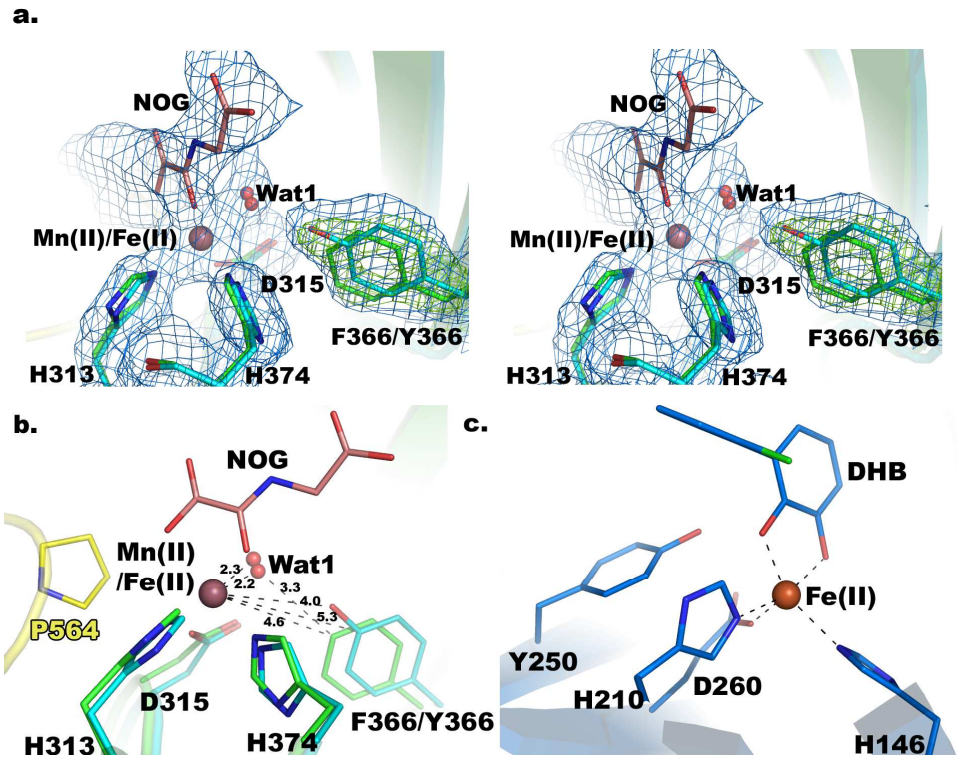


Fig. 6.6 6C and 5C states of tPHD2 and 1,2-DHBD, as revealed by the crystal structures. (a) Stereoview of the superimposed active sites of tPHD2 in tPHD2.CODD (green) and tPHD2 F366Y variant (gray) showing the experimentally observed electron density for the variant structure. The blue $2F_o$-F_c map was contoured to 1σ; the green F_o-F_c OMIT map was calculated by removing Tyr366 of the variant and was contoured to 3σ. The bicyclic analogue in F366Y.Fe(II) complex is not shown for clarity. (b) Views from superimposed tPHD2.CODD and F366Y variant active sites showing the displacement of the metal coordinated water and Tyr366 side chain. $Pro564_{CODD}$ is in yellow sticks. (c) Active site of 1,2-DHBD in the same orientation as tPHD2 (in b) showing a proposed role for $Tyr250_{DHBD}$ to enable the enzyme to render the enzyme in a 5C state. DHB is the substrate, 2'-chlorodihydroxybiphenyl.

The F366Y variant was crystallised with a bicyclic 2OG analogue in the homotrimeric $P6_3$ crystal form; the structure was solved by molecular replacement to 2.0 Å resolution using PDB ID 2G19 as the search model (Fig. 6.6, Table 6.1). Comparison of the tPHD2.CODD and F366Y tPHD2 structures reveals that, in the variant structure, Tyr366 is moved away from the active site approximately by the length of a hydroxyl group (Tyr366 Cζ-Fe(II), 5.3Å; Phe366 Cζ-Mn(II), 4.6 Å). The variant appeared to be active as tPHD2 (Fig. 6.2) at an E/S ratio of 1:25. In future work, it may be of interest to look at its activity with variation in substrate and/or oxygen concentration.

Table 6.1 Crystallographic data processing and refinement statistics of tPHD2.CODD complex structure.

Measurement	**tPHD2 F366Y.Fe(II).BB-287**
Data Collection	
Space Group	$P6_3$
Cell dimensions a,b,c (Å)	111.12
	111.12
	40.06
Resolution (Å)	55.6-2.2 (2.3-2.2)*
No. of unique reflections	14463
Completeness (%)	98.8 (91.4)*
Redundancy	9.63 (4.73)*
R_{sym}**	0.067 (0.360)*
Mean I/σ(I)	26.52 (4.92)*
Refinement	
R_{factor}	0.2008
R_{free}	0.2337
R.m.s. deviation	
Bond length, Å	0.006
Bond angle, °	1.41
No. of atoms	
Protein	1693
Ligand	18 (BB-287)
Water	96
$<B_{factor}>$ Å^2	
Protein	25.88
Ligand	17.20
Metal ion	15.02 (Fe)
Water	33.16

*Highest resolution shell shown in parentheses.
**$R_{sym} = \sum|I-<I>|/\sum I$, where I is the intensity of an individual measurement and $<I>$ is the average intensity from multiple observations.

6.2.3 Rearrangement possibility of 2OG 1-carboxylate or ferryl intermediate

MDS studies predict that NOG (a near isosteric analogue of 2OG) binding may reflect the way 2OG would bind in the active site (Chapter 3, Fig. 3.2). In the tPHD2.CODD crystal

structure, the NOG 1-carboxylate is positioned in the coordination site adjacent to Pro 564_{CODD}; thus the oxygen must either bind to Fe(II) at the less accessible metal coordination site located at the interior of the active site, or the NOG/2OG 1-carboxylate must rearrange to enable oxygen binding *trans* to $His374_{PHD2}$ (Figs. 6.5 and 6.15). Evidence for catalysis proceeding via a mechanism involving such a rearrangement has been proposed for the related 2OG oxygenase, CAS (clavulanic acid synthase).[22]

Coupled to the limited access to the active site iron, in which the β2β3/loop plays an important role (Chapter 3), this arrangement may provide an explanation as to why oxygen has a significantly higher K_M value for PHD2 compared to the other cofactor/cosubstrate. The role of PHDs in the human hypoxic response appears to be different to that of FIH. Unlike PHD2, the crystal structures of FIH.2OG.Fe(II) in substrate bound/unbound forms reveal that FIH does not possess a metal coordinated water[19] (hence Fe(II) is in a 5C state, Fig. 6.5) and the reported K_M of oxygen for FIH is much lower (~90μM) than that for PHDs (Chapter 1, Section 1.9).[1] Taken together these results suggest that PHD2 plays a more important role than FIH as an oxygen sensor.

6.3 Variable CODD conformations with different binding partners

Sequence analyses of HIF-α isoforms predict that both CODD/NODD and CAD regions are largely unstructured. However, they can become more ordered when influenced by the cognate binding partner(s) as observed, for example, in the case of p300.CAD complex structure. Conformational analyses of CODD peptides bound to tPHD2 and VCB complex indicated that the CODD peptides can exhibit multiple conformations (Fig. 6.7). Interestingly, the hydroxylated $Pro564_{CODD}$ adopts two different conformations - C^4-endo and C^4-exo when bound to PHD2 and VCB, respectively. The endo-conformation as observed in tPHD2.CODD structure allows the proline to have a bend at the apex of the peptide. These observations provide evidence that suggests prolyl hydroxylation could be stereo-specific.

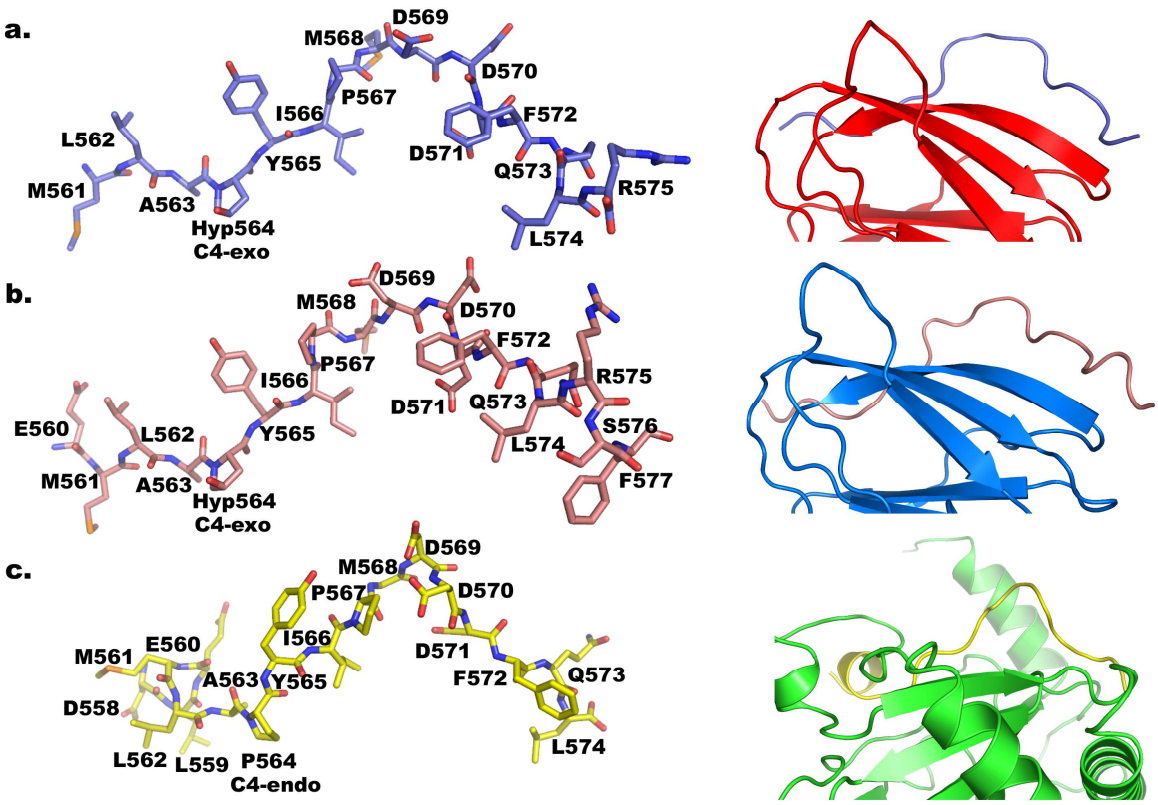


Fig. 6.7 Multiple conformations of the HIF-α CODD peptide residues when in complex with VCB (a and b) and tPHD2 (c). (a) Hydroxylated HIF-1α $CODD_{Hyp564}$ fragment (residues 561-575 modelled) from VCB complex using PDB: 1LM8 (b) Hydroxylated HIF-1α $CODD_{Hyp564}$ fragment (residues 560-577 modelled) from VCB complex using PDB: 1LQB (c) HIF-1α CODD (residues 558-574 modelled) from tPHD2.CODD complex. RMSDs of the peptides in (a) and (b) are 0.47Å, (a) and (c) are 2.8 Å and (b) and (c) are 2.6 Å for 14 residues.

6.4 Stereochemical endo-exo conversion of $Pro564_{CODD}$/ $Pro402_{NODD}$

In order to investigate whether the observed ring puckering of hydroxylated $Pro564_{CODD}$ was not due to packing constraints under crystalline conditions, additional experiments using analogues of $Pro564_{CODD}$ were performed by C. Loenarz (Chemistry Research Laboratory, Oxford).

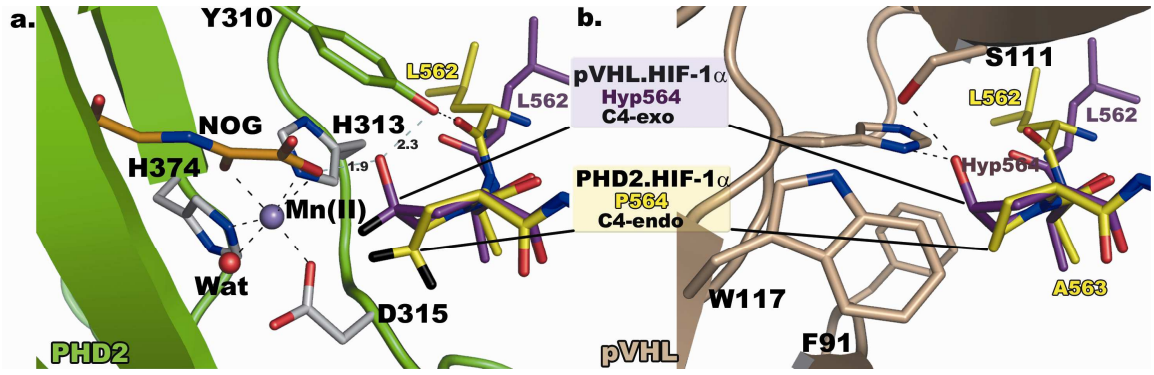


Fig. 6.8 Comparisons of the $Pro564_{CODD}$ conformations when bound to tPHD2 (a, green) and VCB (b, wheat). Superimposition of CODD (yellow)/ CODDHyp564 (purple) in complex with tPHD2 /VCB yielded RMSD of 2.6 Å for 14 residues (561-574, Cα atoms). The $Pro564_{CODD}$ C^4-methylene is in the endo-conformation when bound to tPHD2, whilst in the exo-conformation when $Hyp564_{CODD}$ is bound to the VCB complex. The C^4 Pro564 hydrogen(s) are shown in black.

6.4.1 Stereospecificity of the PHD reaction

At the active site, the pyrrolidine ring of $Pro564_{CODD}$ adopts the C^4-endo conformation (the 4-methylene and 1-carboxamide groups are on the same face, Fig. 6.8). This arrangement projects the 4-*pro*-*R* hydrogen towards the metal and led to the proposal that the hydroxylation step of catalysis, involving reaction of the ferryl intermediate with a C-H bond, requires a specific prolyl ring conformation, and occurs with retention of stereochemistry. Indeed incubation of HIF-1α CODD containing a labelled Pro564 4-*pro*-*R* ^2H with tPHD2 gave hydroxylated CODD with >95% loss of the 4-*pro*-*R* ^2H (C. Loenarz, unpublished data).

6.4.2 Studies with peptide analogues (endo/exo)

Proline residues can exist as an equilibrium of conformers including C^4-endo and C^4-exo forms. Due to the gauche effect operating within the vicinal N-C^5-C^4-X arrangement, stereoelectronic analysis predicts that *trans* prolyl-4-hydroxylation will bias the C^4-endo/ C^4-exo equilibrium to the latter.[23] Indeed the C^4-exo prolyl conformation present in *trans* hydroxyproline residues stabilises the collagen helix[23]. To test the possibility that product release in PHD catalysis is aided by a C^4-endo to C^4-exo switch (Fig. 6.8), C. Loenarz synthesised and assayed analogues of Pro564, predicted to preferentially adopt C^4-endo or C^4-exo conformations, for binding to and as substrates of tPHD2 (Figs 6.9 and 6.10).

In support of a bias to the C^4-exo conformation upon hydroxylation, modelling studies indicated that the side chain hydroxyl group of $Tyr310_{PHD2}$ is positioned to hydrogen bond with the 4*R*-alcohol of hydroxyprolyl564 when in the C^4-exo but not in C^4-endo conformation (Fig. 6.8); thus a C^4-endo to C^4-exo switch may assist in product release.

Binding of $Pro564_{CODD}$ analogues with a preference for the C^4-endo prolyl conformation (i.e. with an electronegative fluoro or hydroxyl group at the 4*S*-position) to tPHD2 was ~five times stronger than for analogues that stabilise the C^4-exo conformation (i.e. with a 4*R*-electronegative group) as analysed by non-denaturing MS, with the 4*R*-fluoroprolyl564 binding the weakest (Fig. 6.9).

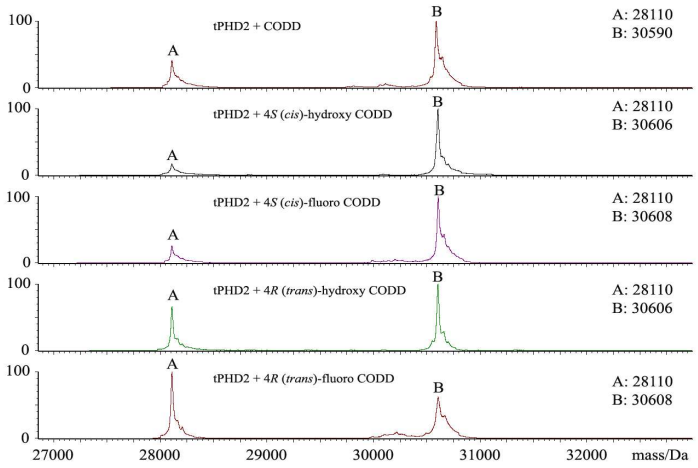
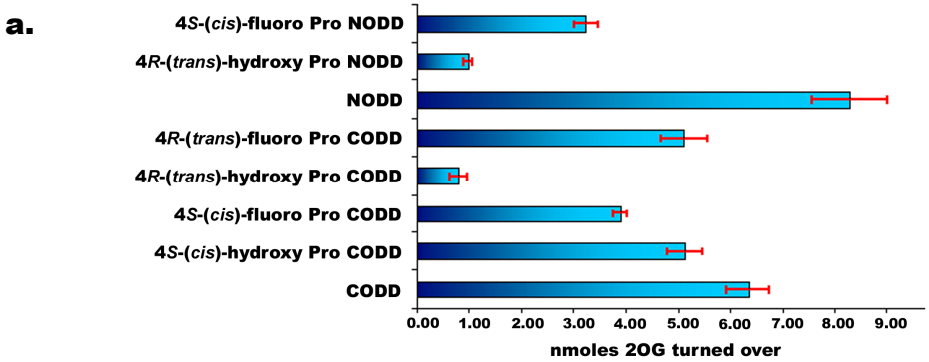


Fig. 6.9 ESI-MS of tPHD2 in the presence of an equimolar amount of Fe(II) and CODD (panel 1); 4*S* (*cis*)-hydroxy Pro 564 CODD (panel 2); 4*S* (*cis*)-fluoro Pro 564 CODD (panel 3); 4*R* (*trans*)-hydroxy Pro 564 CODD (panel 4); 4*R* (*trans*)-fluoro Pro 564 CODD D (panel 5). Peaks represent A) tPHD2.Fe; B) tPHD2.Fe.CODD/Pro 564_{CODD} analogue complexes. Notably, CODD substrates predicted to preferentially adopt a C^4-endo prolyl 564 conformation (4*S*-hydroxyproline and 4*S*-fluoroproline) were observed to bind tPHD2 more strongly than those predicted to have a C^4-exo conformation (4*R*-hydroxyproline and 4*R*-fluoroproline).

On incubation with tPHD2 under catalytic conditions, the 4*S*-hydroxy/fluoroprolyl 564 analogues gave the same 4-oxoprolyl product, supporting the proposal that an electronegative group at the *4S* position stabilises the catalytically productive C^4-endo conformation (Fig. 6.10). In contrast, the 4*R*-fluoroprolyl 564_{CODD} analogue was not hydroxylated, but displayed increased 2OG turnover (Fig. 6.10), suggesting that it can enable reaction with oxygen and 2OG but not hydroxylation, supporting the proposal that it occupies a C^4-exo conformation (Fig. 6.8).



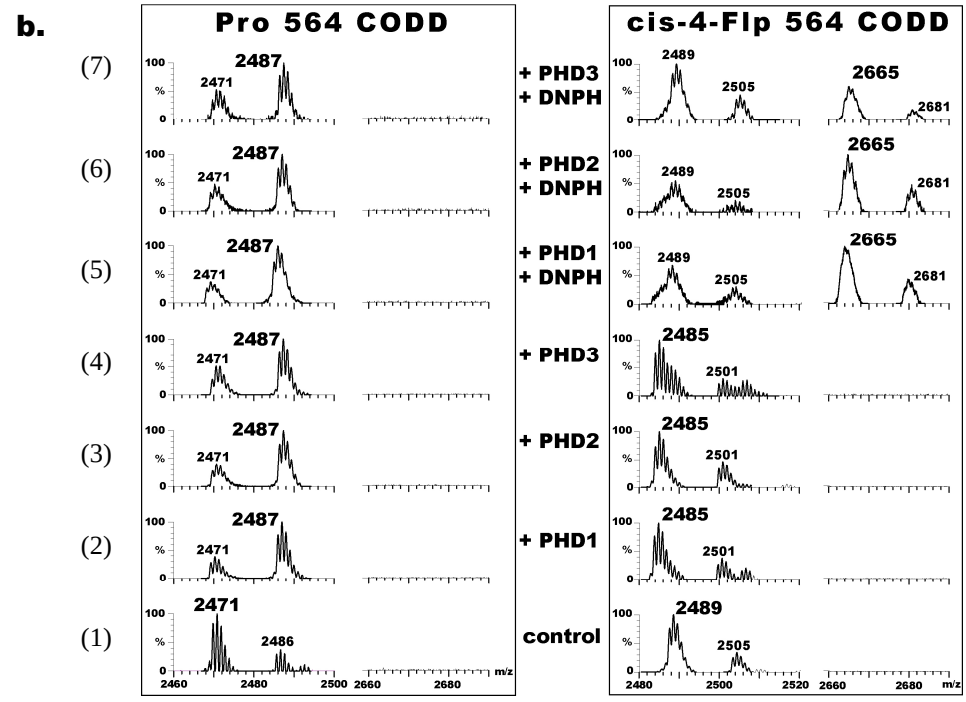


Fig. 6.10 2OG turnover assay and mass spectra showing 4-oxoprolyl formation after incubation of tPHD2 with Pro564$_{CODD}$ and Pro402$_{NODD}$ analogues. (a) 2OG assay of Pro 564$_{CODD}$ analogues with 4*S*-Hyp, 4*S*-fluoroproline (Flp), 4*R*-Hyp and 4*R*-Flp. Pro 564$_{CODD}$ analogues predicted to preferentially adopt a C^4-endo conformation showed significantly higher 2OG turnover than for those predicted to adopt a C^4-exo conformation. The 4*R*-fluoroprolyl 564 analogue was not hydroxylated after incubation with tPHD2 (data not shown) yet stimulated 2OG turnover (>80 % of tPHD2). (b) MALDI-TOF mass spectra obtained after incubation of CODD analogues with PHDs. Each box shows spectra of the peptide analogues (treated with 2,4-dinitrophenyl hydrazine (DNPH), panel 1), after treatment with PHD1 (panels 2 and 5), tPHD2 (panels 3 and 6) and PHD3 (panels 4 and 7). Mass shifts correspond to hydroxylation (+16 Da) of Pro 564$_{CODD}$ and formation of the 4-oxoprolyl 564 product using 4*S*-Hyp and 4*S*-Flp CODD analogues (-2 and -4 Da for 4*S*-Hyp and 4*S*-Flp, respectively). Ketone formation was confirmed by DNPH derivatisation (+180 Da). Formation of a +16 Da peak relative to starting material was observed in controls (i.e. without tPHD2 incubation, panel 1) due to partial methionine (Met560/568$_{CODD}$) oxidation as observed previously.[24] Fig. 6.10b was prepared by C. Loenarz.

6.4.3 4R-Hyp 564$_{CODD}$ is required for pVHL recognition

Given the apparently important role of the Pro564 conformation in PHD catalysis, its role was then investigated in pVHL binding (Figs 6.8 and 6.11). Hydroxylation results in a 1000-fold decrease in the binding constant of HIF-α to the pVHL.Elongin C/B (VCB) complex.[25] In the VCB.CODD$_{Hyp564}$ complex structures, the hydroxyproline (Hyp) 564$_{CODD}$ residue adopts the C^4-exo conformation (Fig. 6.8).[25,26] The Hyp564$_{CODD}$ hydroxyl is positioned to make two hydrogen bonds with pVHL (to the side chains of Ser111 and His115). Pro 564$_{CODD}$ and Pro 402$_{NODD}$ analogues containing C^4 substituents known to bias towards C^4-endo or C^4-exo conformations as well as the 4-oxo derivative were then screened for pVHL binding. None of the 4*S*-fluoro/hydroxy or 4-oxo Pro-analogues bound to pVHL within the detection limits of a pull-down assay (< 10% of CODD$_{Hyp564}$, Fig. 6.11), as neither did the 4*R*-fluoro-analogue.

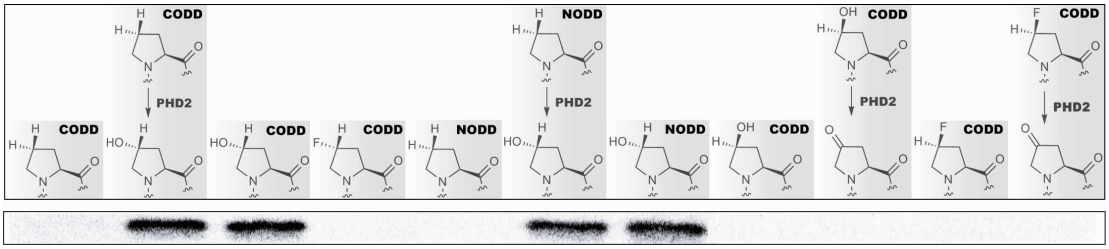


Fig. 6.11 [^{35}S]-pVHL capture assays of non-hydroxylated and 4*R*-/4*S*- Pro564$_{CODD(556-574)}$/Pro 402$_{NODD(395-413)}$ analogues and a 4-oxo-Pro 564CODD as determined by an in vitro transcription/translation (IVTT) coupled pull-down assay (bands in the lower panel show [^{35}S]-pVHL binding).

Non-denaturing MS analyses on the binding of the analogues to the intact VCB complex also revealed that binding of the 4-position analogues to VCB was < 10% of the level of binding observed for 4*R*-Hyp 564$_{CODD}$ (Fig. 6.12).

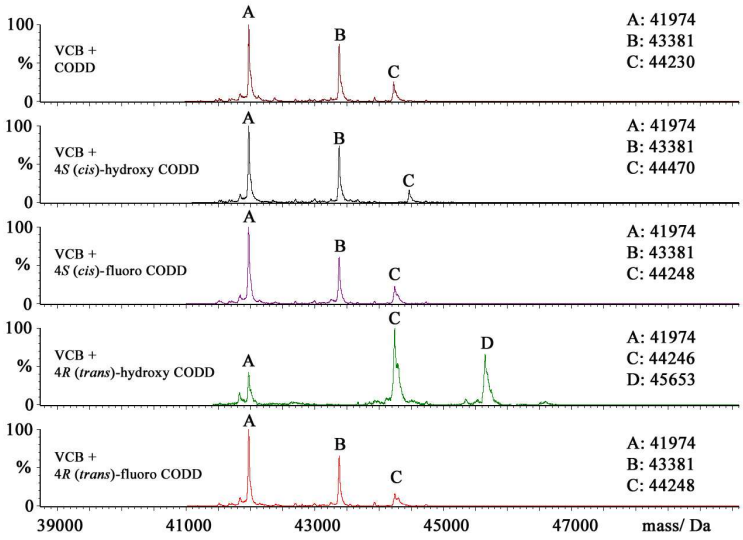


Fig. 6.12 Non-denaturing ESI-MS of VCB in the presence of an equimolar amount of CODD (panel 1); biotinylated 4*S* (*cis*)-hydroxy Pro 564 CODD (panel 2); 4*S* (*cis*)-fluoro Pro 564 CODD (panel 3); 4*R* (*trans*)-hydroxy Pro 564 CODD (panel 4); 4*R* (*trans*)-fluoro Pro 564 CODD (panel 5). Peaks represent A) VCB; B) VCB with uncleaved linker (GSYVQGHLPLVPR); C) the VCB.CODD peptide complex; D) VCB-linker.CODD peptide complex (See Chapter 10 for experimental details).

To more accurately quantify the relative binding of the analogues to VCB, Dr L.A. McNeill (Chemistry Research Laboratory, Oxford) then carried out time-resolved fluorescence (TR-FRET) assays (Fig. 6.13). The following order of binding was observed for CODD and the Pro-564 analogues: 4*R*-Hyp$_{CODD}$ (100 ± 5.3 %), 4*S*-Hyp$_{CODD}$ (5.75 ± 0.1 %), 4*R*-fluoro Pro$_{CODD}$ (1.25 ± 0.1 %), 4*S*-fluoro Pro$_{CODD}$ (1.38 ± 0.1 %) and Pro$_{CODD}$ (0.99 ± 0.1 %). These results reveal the remarkable selectivity of the pVHL pocket for 4*R*-Hyp and demonstrate that both hydrogen bonds from pVHL to the alcohol of Hyp564$_{CODD}$ must contribute significantly to the binding of hydroxylated HIF-α to VCB.

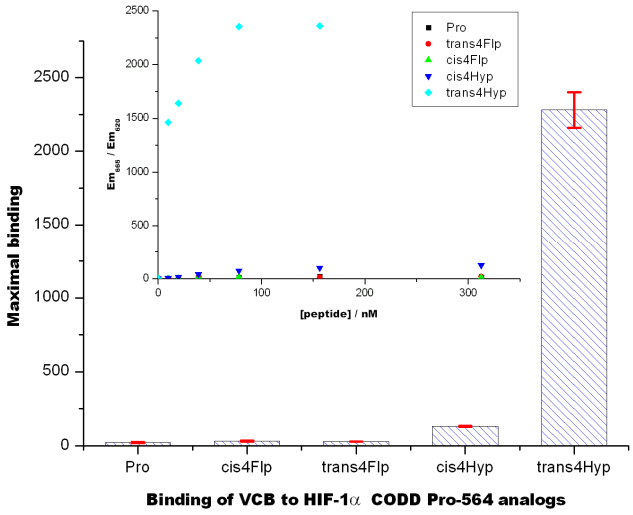


Fig. 6.13 TR-FRET assay of Pro 564$_{CODD}$ peptide analogues revealing differential binding to VCB. Binding of Pro 564$_{CODD}$ peptide analogues to VCB as determined by TR-FRET assay. Maximal binding values were obtained by increasing peptide concentration and measuring the TR-FRET Em_{655}/Em_{615} signal (encapsulated figure). Analysis of the binding curves (see Chapter 10 for experimental detail) for individual peptides result in the following maximal binding values Pro, 22.5 ± 3.0; *cis*-4-Flp, 31.4 ± 2.6; *trans*-4-Flp, 28.4 ± 0.9; *cis*-4-Hyp 131.0 ± 2.7; *trans*-4-Hyp, 2280.0 ± 121.6; Abbreviations: Pro, proline; Hyp, hydroxyproline; Flp, fluoroproline. Figure was prepared by C. Loenarz.

6.5 Implications for inhibitor design

The combined structures enable categorisation of metal chelating PHD inhibitors into those, such as NOG, that coordinate the metal and allow substrate binding with the mobile finger-lid in the closed position (as in crystal form $P2_12_12_1$), and those, such as A and B, that chelate the metal but block both active site substrate and the finger-lid binding in the closed position (notably by disrupting the enzyme R252:D254 salt-bridge), forcing the 'β2β3/loop' to(wards) the conformation observed in the $P6_3$ crystal form (Fig. 6.14).

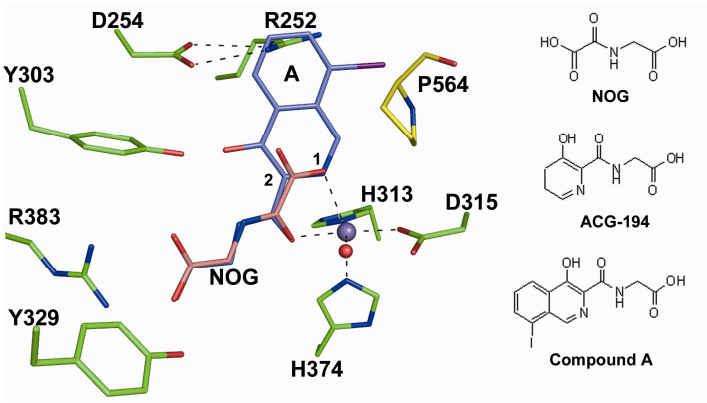


Fig. 6.14 PHD2 active site showing how different classes of PHD2 inhibitors can result in different modes of inhibition. See the text.

6.6 Summary

Overall the results reveal insights into how the structural features of PHD2 enable it to act as an oxygen sensing enzyme. PHD catalysed HIF-α proline hydroxylation appears to induce a conformational bias at Pro564/Pro402 (HIF-1α, C^4 endo/exo) that is part of the hypoxic response mechanism (Fig. 6.15). The conformational bias produced by the gauche effect has not been widely considered in proteins outside the context of collagen stabilisation; coupled with recent reports suggesting that post-translational hydroxylation of intracellular proteins may be more widespread than previously perceived, the results suggest that this effect may be involved in modulating other protein-protein interactions including in signalling pathways.

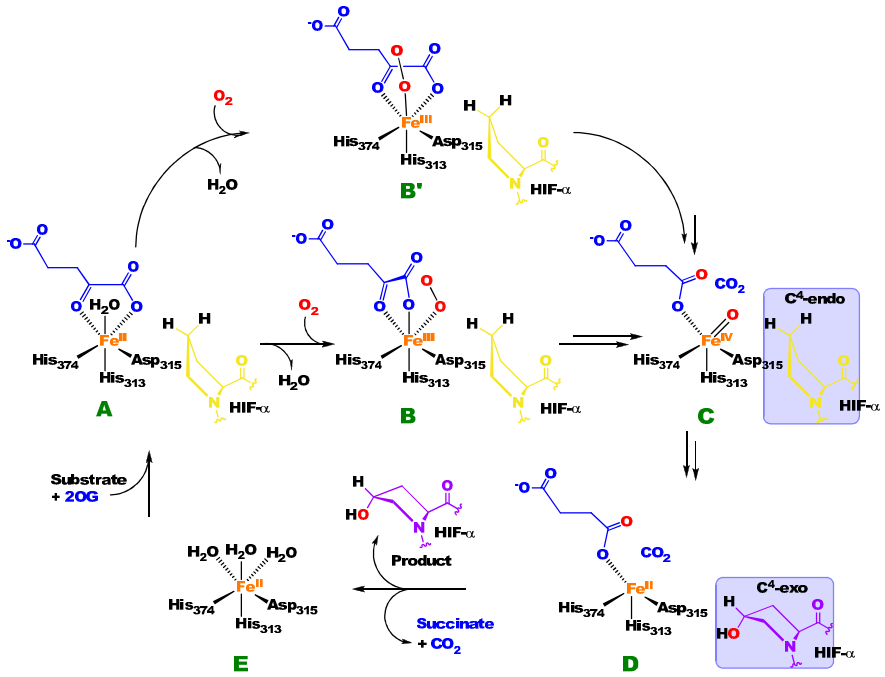


Fig. 6.15 Outline PHD2 catalytic cycle showing the proposed ferryl and other intermediates, highlighting (boxed) the proposed switch in the Pro564$_{CODD}$ ring conformation. The position of oxygen binding is either *trans* to His313$_{PHD2}$ or His374$_{PHD2}$ (Section 6.2.3).

6.7 References

1. Ehrismann, D., Flashman, E., Genn, D.N., Mathioudakis, N., Hewitson, K.S., Ratcliffe, P.J. & Schofield, C.J. Studies on the activity of the hypoxia-inducible-factor hydroxylases using an oxygen consumption assay. *Biochem J* **401**, 227-34 (2007).
2. Hirsila, M., Koivunen, P., Gunzler, V., Kivirikko, K.I. & Myllyharju, J. Characterization of the human prolyl 4-hydroxylases that modify the hypoxia-inducible factor. *J Biol Chem* **278**, 30772-80 (2003).
3. Koivunen, P., Hirsila, M., Kivirikko, K.I. & Myllyharju, J. The Length of Peptide Substrates Has a Marked Effect on Hydroxylation by the Hypoxia-inducible Factor Prolyl 4-Hydroxylases. *J Biol Chem* **281**, 28712-20 (2006).
4. McNeill, L.A., Flashman, E., Buck, M.R., Hewitson, K.S., Clifton, I.J., Jeschke, G., Claridge, T.D., Ehrismann, D., Oldham, N.J. & Schofield, C.J. Hypoxia-inducible factor prolyl hydroxylase 2 has a high affinity for ferrous iron and 2-oxoglutarate. *Mol Biosyst* **1**, 321-4 (2005).
5. Epstein, A.C., Gleadle, J.M., McNeill, L.A., Hewitson, K.S., O'Rourke, J., Mole, D.R., Mukherji, M., Metzen, E., Wilson, M.I., Dhanda, A., Tian, Y.M., Masson, N., Hamilton, D.L., Jaakkola, P.,

Barstead, R., Hodgkin, J., Maxwell, P.H., Pugh, C.W., Schofield, C.J. & Ratcliffe, P.J. C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell* **107**, 43-54 (2001).

6. Kaelin, W.G. Proline hydroxylation and gene expression. *Annu Rev Biochem* **74**, 115-28 (2005).
7. Schofield, C.J. & Ratcliffe, P.J. Oxygen sensing by HIF hydroxylases. *Nat Rev Mol Cell Biol* **5**, 343-54 (2004).
8. McDonough, M.A., Li, V., Flashman, E., Chowdhury, R., Mohr, C., Lienard, B.M., Zondlo, J., Oldham, N.J., Clifton, I.J., Lewis, J., McNeill, L.A., Kurzeja, R.J., Hewitson, K.S., Yang, E., Jordan, S., Syed, R.S. & Schofield, C.J. Cellular oxygen sensing: Crystal structure of hypoxia-inducible factor prolyl hydroxylase (PHD2). *Proc Natl Acad Sci U S A* **103**, 9814-9 (2006).
9. Hoffart, L.M., Barr, E.W., Guyer, R.B., Bollinger, J.M., Jr. & Krebs, C. Direct spectroscopic detection of a C-H-cleaving high-spin Fe(IV) complex in a prolyl-4-hydroxylase. *Proc Natl Acad Sci U S A* **103**, 14738-43 (2006).
10. Price, J.C., Barr, E.W., Tirupati, B., Bollinger, J.M., Jr. & Krebs, C. The first direct characterization of a high-valent iron intermediate in the reaction of an alpha-ketoglutarate-dependent dioxygenase: a high-spin FeIV complex in taurine/alpha-ketoglutarate dioxygenase (TauD) from Escherichia coli. *Biochemistry* **42**, 7497-508 (2003).
11. Chen, Z., Zang, J., Kappler, J., Hong, X., Crawford, F., Wang, Q., Lan, F., Jiang, C., Whetstine, J., Dai, S., Hansen, K., Shi, Y. & Zhang, G. Structural basis of the recognition of a methylated histone tail by JMJD2A. *Proc Natl Acad Sci U S A* **104**, 10818-23 (2007).
12. Katona, G., Carpentier, P., Niviere, V., Amara, P., Adam, V., Ohana, J., Tsanov, N. & Bourgeois, D. Raman-assisted crystallography reveals end-on peroxide intermediates in a nonheme iron enzyme. *Science* **316**, 449-53 (2007).
13. Oster, L.M., van Scheltinga, A.C., Valegard, K., Hose, A.M., Dubus, A., Hajdu, J. & Andersson, I. Conformational flexibility of the C terminus with implications for substrate binding and catalysis revealed in a new crystal form of deacetoxycephalosporin C synthase. *J Mol Biol* **343**, 157-71 (2004).
14. Valegard, K., van Scheltinga, A.C., Lloyd, M.D., Hara, T., Ramaswamy, S., Perrakis, A., Thompson, A., Lee, H.J., Baldwin, J.E., Schofield, C.J., Hajdu, J. & Andersson, I. Structure of a cephalosporin synthase. *Nature* **394**, 805-9 (1998).
15. Wilmouth, R.C., Turnbull, J.J., Welford, R.W., Clifton, I.J., Prescott, A.G. & Schofield, C.J. Structure and mechanism of anthocyanidin synthase from Arabidopsis thaliana. *Structure* **10**, 93-103 (2002).
16. Koski, M.K., Hieta, R., Bollner, C., Kivirikko, K.I., Myllyharju, J. & Wierenga, R.K. The active site of an algal prolyl 4-hydroxylase has a large structural plasticity. *J Biol Chem* **282**, 37112-23 (2007).
17. Neidig, M.L., Brown, C.D., Light, K.M., Fujimori, D.G., Nolan, E.M., Price, J.C., Barr, E.W., Bollinger, J.M., Jr., Krebs, C., Walsh, C.T. & Solomon, E.I. CD and MCD of CytC3 and taurine dioxygenase: role of the facial triad in alpha-KG-dependent oxygenases. *J Am Chem Soc* **129**, 14224-31 (2007).
18. Neidig, M.L., Kavana, M., Moran, G.R. & Solomon, E.I. CD and MCD studies of the non-heme ferrous active site in (4-hydroxyphenyl)pyruvate dioxygenase: correlation between oxygen activation in the extradiol and alpha-KG-dependent dioxygenases. *J Am Chem Soc* **126**, 4486-7 (2004).
19. Elkins, J.M., Hewitson, K.S., McNeill, L.A., Seibel, J.F., Schlemminger, I., Pugh, C.W., Ratcliffe, P.J. & Schofield, C.J. Structure of factor-inhibiting hypoxia-inducible factor (HIF) reveals mechanism of oxidative modification of HIF-1 alpha. *J Biol Chem* **278**, 1802-6 (2003).
20. Hewitson, K.S., Holmes, S.L., Ehrismann, D., Hardy, A.P., Chowdhury, R., Schofield, C.J. & McDonough, M.A. Evidence that two enzyme derived histidine ligands are sufficient for iron binding and catalysis by factor inhibiting HIF (FIH). *J Biol Chem* (2008).
21. Davis, M.I., Wasinger, E.C., Decker, A., Pau, M.Y., Vaillancourt, F.H., Bolin, J.T., Eltis, L.D., Hedman, B., Hodgson, K.O. & Solomon, E.I. Spectroscopic and electronic structure studies of 2,3-dihydroxybiphenyl 1,2-dioxygenase: O2 reactivity of the non-heme ferrous site in extradiol dioxygenases. *J Am Chem Soc* **125**, 11214-27 (2003).
22. Zhang, Z., Ren, J., Harlos, K., McKinnon, C.H., Clifton, I.J. & Schofield, C.J. Crystal structure of a clavaminate synthase-Fe(II)-2-oxoglutarate-substrate-NO complex: evidence for metal centered rearrangements. *FEBS Lett* **517**, 7-12 (2002).
23. Jenkins, C.L. & Raines, R.T. Insights on the conformational stability of collagen. *Nat Prod Rep* **19**, 49-59 (2002).
24. Jaakkola, P., Mole, D.R., Tian, Y.M., Wilson, M.I., Gielbert, J., Gaskell, S.J., Kriegsheim, A., Hebestreit, H.F., Mukherji, M., Schofield, C.J., Maxwell, P.H., Pugh, C.W. & Ratcliffe, P.J. Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. *Science* **292**, 468-72 (2001).

25. Hon, W.C., Wilson, M.I., Harlos, K., Claridge, T.D., Schofield, C.J., Pugh, C.W., Maxwell, P.H., Ratcliffe, P.J., Stuart, D.I. & Jones, E.Y. Structural basis for the recognition of hydroxyproline in HIF-1 alpha by pVHL. *Nature* **417**, 975-8 (2002).
26. Min, J.H., Yang, H., Ivan, M., Gertler, F., Kaelin, W.G., Jr. & Pavletich, N.P. Structure of an HIF-1alpha -pVHL complex: hydroxyproline recognition in signaling. *Science* **296**, 1886-9 (2002).

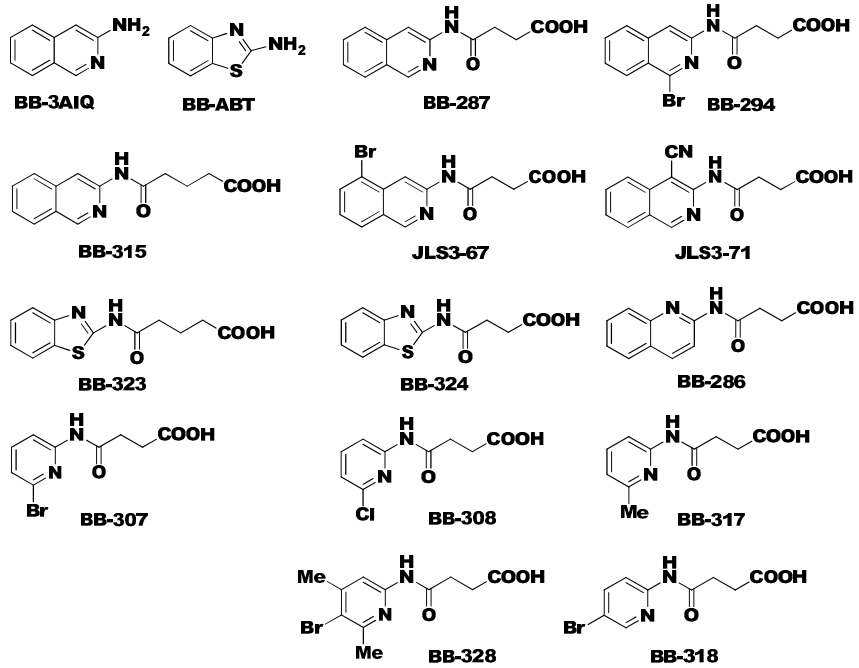
Chapter 7 – Crystallisation and structure solutions of tPHD2 in complex with a new series of 2OG analogues

As described in Chapter 2, a number of 2OG analogues were screened with the aim of crystallising tPHD2 using the reported tPHD2 crystallisation conditions.[1] This enabled crystallisation of tPHD2 in complex with a new series of compounds synthesised by Dr B. Banerji and Dr J. L. Sorensen (Chemistry Research Laboratory, Oxford). The inhibition profiles of these 2OG analogues against tPHD2 as well as binding assay results suggested that compounds of this series did not bind tightly. Nonetheless, four compounds of this series were able to be crystallised with tPHD2. Further attempts were made to enhance the diffraction quality of the crystals by lysine methylation.[2] The impact of lysine methylation on the tPHD2 structure as well as on catalysis is also described in this Chapter.

7.1 Inhibition data of the BB-287 compounds

The compounds in the BB-287 series (Fig. 7.1a) were evaluated for their inhibition profiles against tPHD2 using a time-resolved fluorescence resonance energy transfer (TR-FRET) assay[3] (Dr L.A. McNeill, unpublished data). The compounds in this series showed only weak inhibition even when they were pre-incubated with tPHD2 before assay (Fig. 7.1b). They were categorised as weak inhibitors with an apparent IC_{50} of ~500 μM for BB-287, the most potent of these compounds.

a.



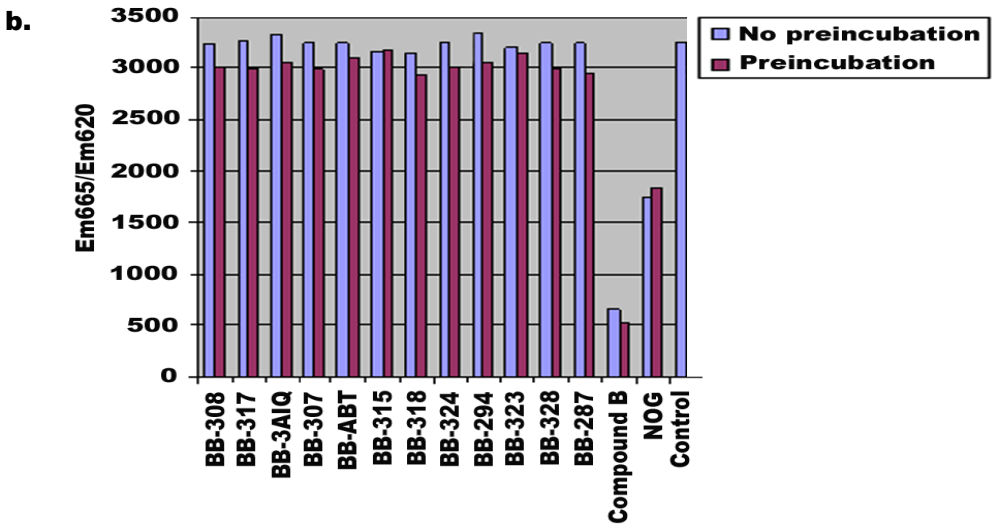


Fig. 7.1 (a) Structures of BB-287 compounds. (b) Activity of tPHD2 in the presence of the BB-287 compounds at 400 μM. Compound B and NOG (400 μM) were used as control inhibitors for comparison. A negative control utilised the same experimental conditions without addition of any inhibitor. The ratio of the two emission signals (Em665 and Em620) is a measure of FRET/ relative fluorescence that indicates the capture of hydroxylated product by Eu^{3+} labelled VCB. For experimental details, see Chapter 10, Section 10.11.5.

7.2 Crystallisation with the BB-287 series of compounds and structure solution

Parallel ligand screens using standard tPHD2 crystallisation conditions (1.5-2.0 M ammonium sulphate, 1-7% dioxane, 0.1M MES pH 6.5)[1] demonstrated that tPHD2 can crystallise with the BB-287 compounds. This was interesting because none of the compounds of this series were identified in the systematic search for ligands using thermal stabilisation or MS analyses (Chapter 2).

7.2.1 Crystallisation of tPHD2 with BB-287 compounds

For crystallisation with tPHD2, compounds at a stock concentration of 0.1 M (in DMSO) were added to purified tPHD2 (20 $mg.ml^{-1}$ in 50 mM Tris.HCl pH 7.5) in a twofold molar excess and incubated for 3-4 hours on ice. Of all the tested compounds of the BB-287 series (Fig. 7.1a), four compounds crystallised with tPHD2 (Fig. 7.2). Crystals were grown by the vapour-diffusion method under near anaerobic (PO_2 <0.1 ppm, with 1 mM $FeSO_4$) or aerobic (with 1 mM $MnCl_2$) conditions. Hexagonal crystals for tPHD2.Fe(II).BB-287/BB-328 and tPHD2.Mn(II).JLS3-67 complexes appeared several hours after the plates were set up and grew to full size within 48 hours. In contrast, tPHD2.Fe(II).BB-294 and tPHD2.Fe(II).BB-327 crystals appeared in the hanging drops after ~3 months of plate setup, in a prismatic shape, indicating the possibility of different crystal packing than the hexagonal $P6_3$ form.

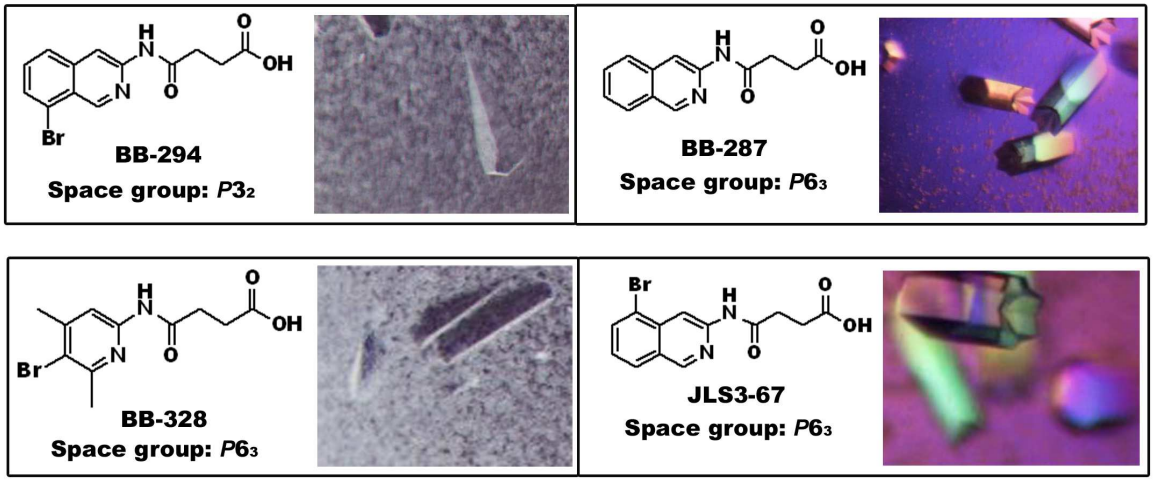


Fig. 7.2 tPHD2 crystals formed in the presence of Fe(II) or Mn(II) and BB-287 compounds.

7.2.2 Structure solution

The crystals of tPHD2.BB-287/BB-294/BB-328/JLS3-67 complexes were soaked in a cryoprotectant comprised of reservoir solution plus 30% glycerol (v/v) and then flash-frozen in liquid N_2 using Hampton Research nylon loops. After being mounted on a goniometer fitted with N_2 cryo-stream, tPHD2.BB-287/BB-328/JLS3-67 complex crystals diffracted in a range of 2.0-2.5 Å with a hexagonal $P6_3$ space group as seen for compounds A/B. The data collection and processing details are summarised in Table 7.1. Phases for tPHD2.BB-287/BB-328/JLS3-67 crystals were determined by molecular replacement using PDB 2G19. Iterative cycles of model building and refinement led to the final structures (Fig. 7.3).

Unfortunately, none of the crystals with BB-294 diffracted below 3.5 Å. Data collection of 0.5° oscillation images over 90° range enabled determination of space group of $P3_2$ for tPHD2.Fe(II).BB-294 crystals to 3.8 Å resolution with unit cell parameters a = b = 154.49, c = 84.63 Å (Table 7.1). The large unit cell appeared to contain 6 molecules per asymmetric unit as suggested by Mathews' coefficient analyses[4] and the structure was solved by Phaser[5] using PDB 2G19 as the search model.

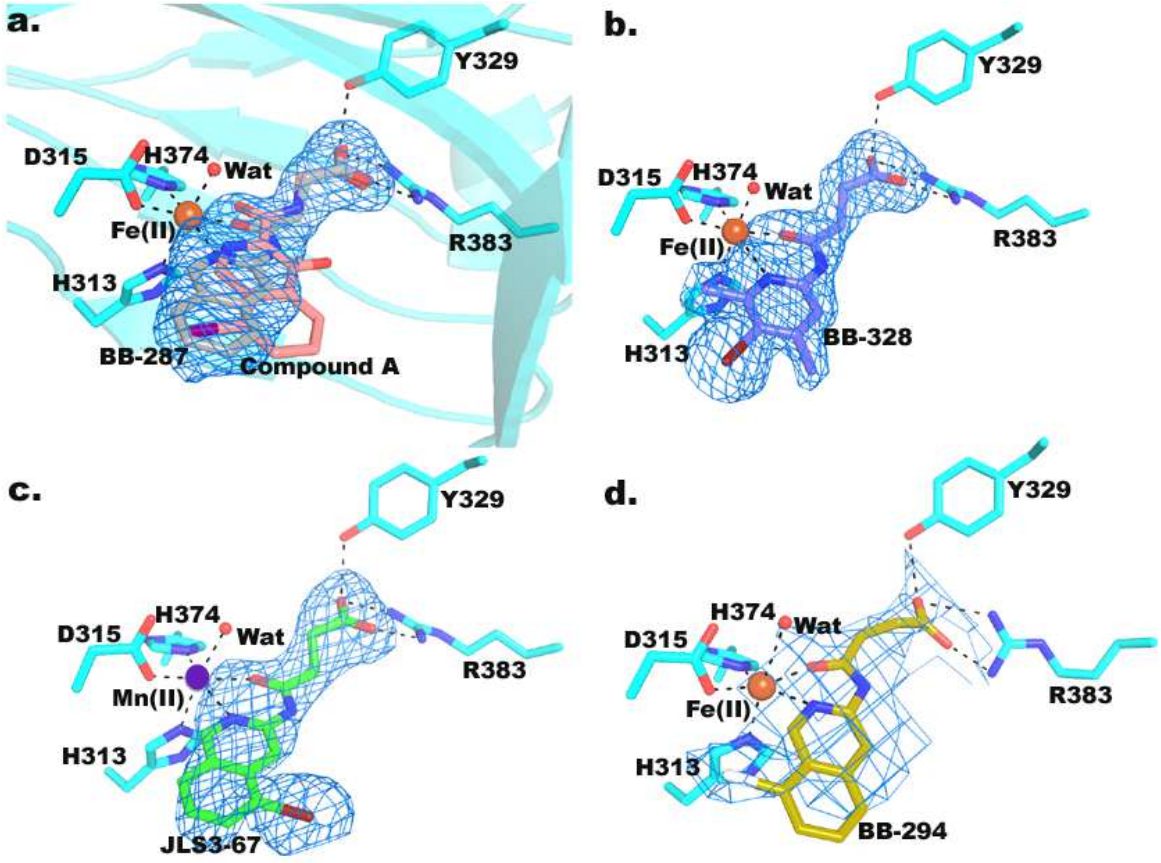


Fig. 7.3 Crystal structures of tPHD2 in complexes with BB-287 series compounds showing the *Fo-Fc* OMIT map contoured to 3σ around the ligands. Dotted lines indicate apparent hydrogen-bonds/ polar interactions. Overall tPHD2 folds were similar to those reported (rmsd for Cα ≤ 0.3 Å).[1]

Table 7.1 Data statistics for the tPHD2.BB-287/BB-328/JLS3-67/BB-294 complex crystals and refinement statistics.

Measurement	**tPHD2.Fe.BB-287**	**tPHD2.Fe.BB-328**	**tPHD2.Fe.JLS-367**	**tPHD2.Fe.BB-294**
Data Collection				
Space Group	$P6_3$	$P6_3$	$P6_3$	$P3_2$
Cell dimensions a,b,c (Å)	109.993	110.807	109.459	154.485
	109.993	110.807	109.459	154.485
	39.368	39.889	39.077	84.628
No. of molecules per a.s.u.	1	1	1	6
Resolution (Å)	50.00-2.00	50.00-1.99	50.00-2.37	50.00-3.80
	(2.03-2.00)*	(2.03-1.99)*	(2.48-2.37)*	
No. of unique reflections	18381	19823	12315	19597
Completeness (%)	98.3 (86.0)*	97.1 (96.2)*	88.9 (95.2)*	87.9 (51.3)*
Redundancy	6.7 (2.3)*	5.4 (4.8)*	3.2 (3.0)*	-
R_{sym}**	0.085 (0.549)*	0.098 (0.757)*	0.108 (0.455)*	-
Mean I/σ(I)	16.46 (1.8)*	13.75 (1.98)*	7.96 (2.07)*	(2.4)*

Refinement				
R_{factor}	0.217	0.228	0.218	0.211
R_{free}	0.251	0.254	0.266	0.234
R.m.s. deviation				
Bond length, Å	0.008	0.007	0.008	0.014
Bond angle, °	1.17	1.26	1.59	1.49
No. of atoms				
Protein	1696	1696	1723	9995
Ligand	18 (BB-287)	17 (BB-328)	19 (JLS-367)	19×6 (BB-294)
Water	110	127	111	6
$<B_{factor}>Å^2$				
Protein	39.55	32.53	44.38	61.05
Ligand	31.19	23.01	46.33	89.21
Metal	27.22 (Fe)	20.26 (Fe)	35.19 (Mn)	16.68
Water	43.65	40.34	47.52	40.08

*Highest resolution shell shown in parenthesis.

**$R_{sym} = \sum|I-<I>|/\sum I$, where I is the intensity of an individual measurement and $<I>$ is the average intensity from multiple observations.

As shown in Fig. 7.3, the compounds in the BB-287 series (BB-287, BB-328, BB-294 and JLS3-67) appeared to have similar interaction to tPHD2 although there were some differences in binding as compared to Compounds A/B that are outlined in the next Section.

7.3 Comparison of tPHD2.Fe(II).BB-287 with tPHD2.Fe(II).A/B crystal structures

7.3.1 Fe(II) coordination

All three compounds, A/B/BB-287 were observed to coordinate Fe(II) via the isoquinoyl nitrogen and an oxygen of the amide carbonyl (O11 in compounds A and B, O12 in BB-287) (Table 7.2). Interestingly, the bidentate coordination of Fe(II) by compounds A/B forms an approximately planar five membered chelate ring whereas BB-287 compounds form a six membered ring (Figs 7.4 and 7.5). The chelate ring size may affect the stability of the complex.[6]

The amide carbonyl oxygen (O11/O12) of compounds A/B/BB-287 is positioned to coordinate the iron approximately *trans* to the $Asp315_{PHD2}$ side chain and their isoquinoyl nitrogen is *trans* to $His374_{PHD2}$ Nε2 (Fig. 7.4). This binding mode allows compounds A/B to form a hydrogen bond between their phenolic oxygen and the side-chain hydroxyl of the conserved $Tyr303_{PHD2}$ residue. BB-287 lacks a phenolic oxygen at the analogous position (C4) to A and B and therefore cannot interact in this way (Fig. 7.4).

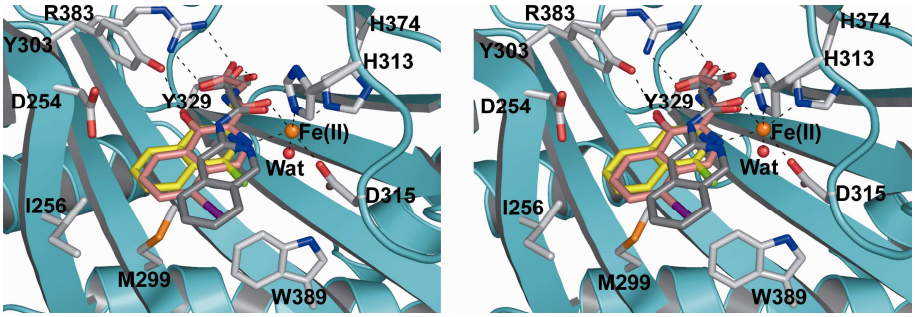


Fig. 7.4 Stereoviews showing comparison of $P6_3$ crystal forms with compounds A, B and BB-287 revealing similar binding modes. Compounds A, B and BB-287 are in salmon, yellow and grey sticks.

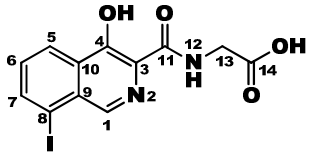
7.3.2 2OG C5 carboxylate binding

The 2OG C5 equivalent carboxylate of the three compounds (C14 of compounds A/B and C15 of BB-287) is positioned to form an ion-pair with the $Arg383_{PHD2}$ guanidino group and to hydrogen bond with the phenolic OH of $Tyr329_{PHD2}$ (Fig. 7.4).

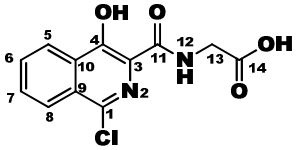
7.3.3 Hydrophobic interactions

Except for the Fe(II) and 2OG binding residues (His313, Asp315, His374, Arg383 and Tyr329), the tPHD2 active site is predominantly lined by hydrophobic residues derived from the DSBH (Table 7.2). The bicyclic aromatic rings of the compounds project through the active site opening and are sandwiched between the side chains of Tyr310, Met299, and Trp389 (Fig. 7.4 and Table 7.2). Besides these DSBH residues, an additional hydrophobic shelf formed by the side chains of Ile256 and Trp258 leads to the active site opening. Interestingly, C13 and C14 atoms of BB-287 undergo a slight twisting to accommodate the additional side chain carbon (Figs 7.4 and 7.5). Thus, BB-287 has four bonds capable of facile rotation as compared to three such bonds for compounds A/B (Table 7.2). Because the overall binding energy correlates with the degrees of freedom within the molecule, more rigid compounds will bind tighter at the active site.

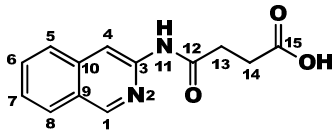
Table 7.2 The observed hydrophobic interactions of 2OG bicyclic analogues and NOG to the tPHD2 active site in the tPHD2 crystal structures. (-) Indicates no apparent interaction was observed.



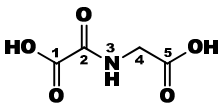
Compound A



Compound B



BB-287



NOG

Residue	DSBH Strand	tPHD2.Fe(II).A; PDB ID: 2G19 (distance, Å)	tPHD2.Fe(II).B; PDB ID: 2HBT (distance, Å)	tPHD2.Fe(II).BB-287 (distance, Å)	tPHD2.Mn(II).NOG. CODD (distance, Å)
Met299	I	M299Cε - C5 (3.5)	M299Cε - C5 (3.6)	M299Cε – C4 (3.7)	M299Cε – C1 (4.2)
Ala301	I	A301Cβ - C5 (4.7)	A301Cβ - C5 (4.7)	-	-
Tyr303	I	Y303Cζ - C4 (4.7)	Y303Cζ - C4 (4.7)	Y303Cζ – C13 (5.0)	-
Tyr310	II	Y310Cζ - C8 (4.1) Y310Cε2 - C10 (3.7)	Y310 Cζ – C9 (3.7) Y310 Cε2 - C10 (3.6)	Y310Cζ – C4 (3.8) Y310Cε1 – C4 (3.8)	Y310Cδ2 – C1 (3.5) Y310Cε2 - C10 (3.4)
Ile327	III	I327Cδ1 - C13 (3.7)	I327Cδ1 - C13 (3.9)	I327Cδ1 - C14 (3.5) I327Cδ1 – C15 (4.0)	I327Cδ1 – C4 (3.7) I327Cδ1 – C5 (3.8)
Leu343	IV	L343Cδ1 - C13 (4.0)	L343Cδ1 - C13 (3.9)	L343Cδ1 - C13 (4.2) L343Cδ1 - C14 (4.1)	L343Cδ1 - C4 (3.6)
Val376	VII	V376Cγ2 - C14 (3.7) V376Cγ2 - C13 (3.6)	V376Cγ2 - C14 (3.6) V376Cγ2 - C13 (3.4)	V376Cγ2 - C12 (3.3) V376Cγ2 – C15 (4.1)	V376Cγ2 – C4 (3.3) V376Cγ2 – C5 (3.5)
Ala385	VIII	A385Cβ – C14 (4.9)	A385Cβ – C14 (5.0)	A385Cβ – C15 (4.6)	A385Cβ – C5 (4.8)
Trp389	VIII	W389Cζ2 – C1 (4.3) W389Cη2 – C1 (4.7)	W389Cζ2 – CL (4.0) W389Cη2 – C1 (4.7)	W389Cζ2 – C1 (4.4) W389Cη2 – C9 (4.5)	W389Cζ2 – C1 (4.5) W389Cη2 – C1 (4.8)

7.3.4 Interactions with symmetry related molecules

The aromatic bicycles of compounds A/B/BB-287 and the iodine atom of compound A, make hydrophobic interactions with Val401 of C-terminal helix α4 of a neighbouring molecule in the $P6_3$ crystal form (Fig. 7.5). The shorter distances observed between the

aromatic ring of BB-287 and $Val401_{PHD2}$ of a symmetry related molecule suggest that BB-287 may have stronger hydrophobic interactions with a symmetry related neighbour and possibly play a relatively more important role in stabilising the homotrimeric form than compounds A/B.

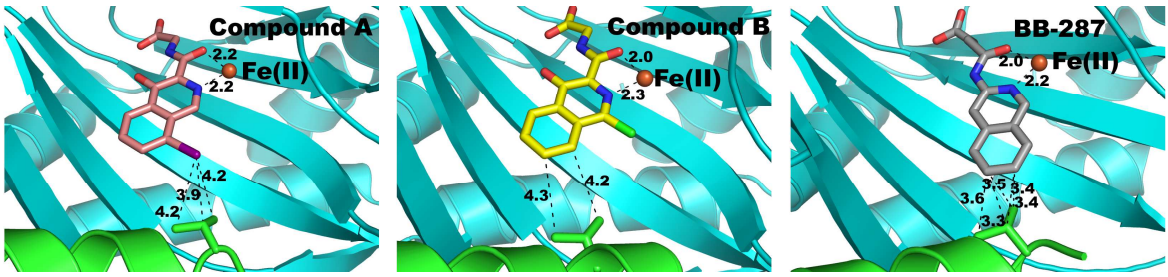


Fig. 7.5 **Hydrophobic interactions observed in the *P*6_3 crystal form between 2OG bicyclic inhibitors and the $Val401_{PHD2}$ of a symmetry related molecule**.

In conclusion, the structural observations suggest that compared to compounds A/B, BB-287 binds less tightly. This is consistent with the non-denaturing MS studies (Chapter 2, Fig. 2.10) and a significantly higher IC_{50} value for BB-287. With the hope that a weaker binding could be beneficial for competitive soaking experiments, 2OG analogues were used in attempts to replace BB-287 in the crystalline state. However, as for compounds A/B, these attempts were unsuccessful.

7.4 Attempts to crystallise tPHD2 in the presence of BB-287 and CODD substrate

Attempts to co-crystallise tPHD2 with CODD and BB-287 resulted in several hits (Fig. 7.6). However, these conditions could not be optimised; the available crystals also did not diffract well.



Fig. 7.6 tPHD2 crystals in presence of BB-287 and CODD

7.5 Methylation of tPHD2 in attempts to improve crystal contacts

Surface engineering of proteins by chemical modification such as reductive lysine methylation or microengineering of flexible surface residues, e.g. lysine or glutamate, has been successfully used to change crystal contacts within a lattice.[2,5] tPHD2 has 21 lysines, some of which are more surface exposed than others. Therefore, instead of mutating all 21 lysine residues, which would be a time- and labour-intensive procedure, reductive lysine methylation was performed with the hope of obtaining better diffraction-quality crystals of tPHD2. This could also lead to a new crystal form with a different conformation, particularly of the finger-lid, other than that observed in the $P6_3$ and $P2_12_12_1$ crystal forms.

Briefly, the protein in a low concentration (0.5-1 mg.ml^{-1}), was reductively methylated by the sequential addition of 1M dimethylamine-borane complex and 1M formaldehyde with gentle mixing.[2] The process was repeated twice before finally incubating for 24 hours at 4°C. Interestingly, the methylation reaction product was resolved in two separate peaks by size exclusion chromatography (SEC) (Fig. 7.7).

7.5.1 Characterisation of the methylation products

Appropriate peak fractions for the methylated tPHD2 products (tPHD2Me-I and tPHD2Me-II, hereafter) were pooled and concentrated (Fig. 7.7). Although the SEC profile of tPHD2Me-I (broader peak) indicated a heterogeneous mixture of multiple components, SDS-PAGE analyses of both products did not reveal significant differences in homogeneity (Fig. 7.7 inset)

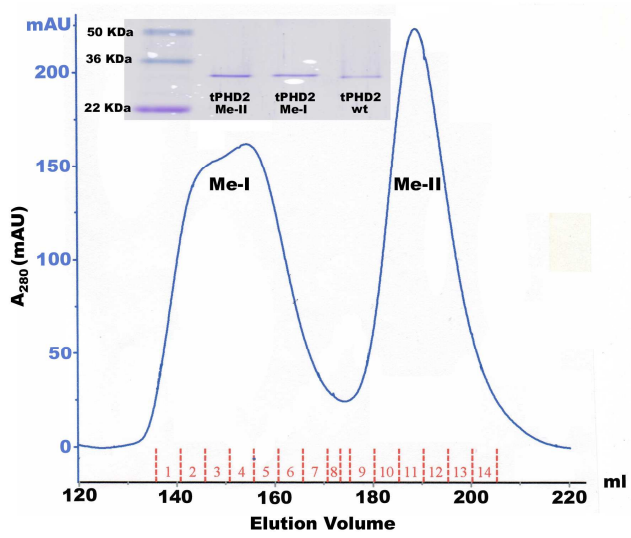


Fig. 7.7 Size exclusion chromatography (SEC) profile and SDS-PAGE analyses of methylated tPHD2 products.

Consistent with SDS-PAGE analyses, MS (under denatured conditions) of both tPHD2Me-I and -II revealed two major peaks at 28674.0 (+610 Da) and 28664.0 (+600 Da), respectively. The additional masses corresponded to ~22/21 dimethylation events (21 lysines plus the N-terminus) (Fig. 7.8).

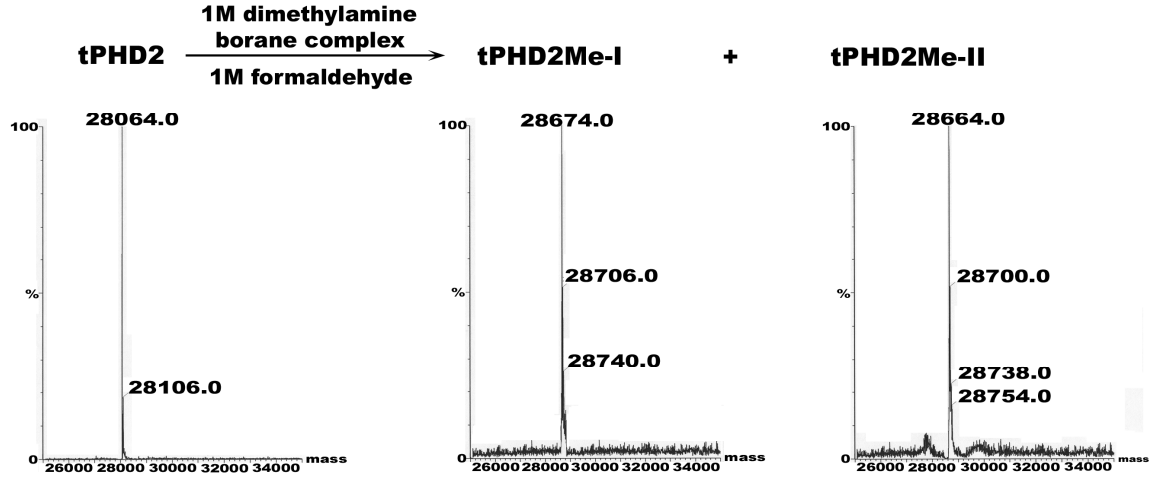


Fig. 7.8 Denatured MS of the unmodified tPHD2 and products after treatment with methylation reagents.

The observation of two distinct peaks (tPHD2Me-I and tPHD2Me-II) in SEC, despite having similar molecular weights (Fig. 7.7), suggested that one of the methylation products might be aggregated. SEC separates proteins on the basis of gyration radius or the hydrodynamic volume of the components. Since the tPHD2Me-II protein eluted in a similar elution volume to that of wt tPHD2 (~180 ml in a 300 ml Superdex S75, Pharmacia column), it was considered more likely that tPHD2Me-II was folded similarly to wt tPHD2 in solution.

To investigate the methylation status of lysine residues, the two products (tPHD2Me-I and tPHD2Me-II) were treated with proteolytic trypsin. The fragmentation patterns thus obtained on the LC-MS spectra appeared to be different relative to that of unmodified tPHD2 due to a number of missed cleavages (Appendices 6 and 7). LC-MS/MS analyses on the fragments of tPHD2Me-I and tPHD2Me-II provided MS/MS evidence for di-methylation of most the lysine residues except where stated otherwise (Table 7.4).

7.5.2 Crystallisation of methylated tPHD2

Attempts were made to co-crystallise both tPHD2Me-I (18 $mg.ml^{-1}$) and tPHD2Me-II (18 $mg.ml^{-1}$) with and without 2OG analogues as described before (Chapter 2). The methylation products crystallised with BB-287 using different crystallisation conditions in

typical hexagonal rod morphology. tPHD2Me-II could be crystallised in a number of conditions including one that was further optimised (Fig. 7.9). In contrast, tPHD2Me-I also crystallised but with significant amounts of precipitation in all initial hit conditions plus in all optimised conditions (Fig. 7.9).

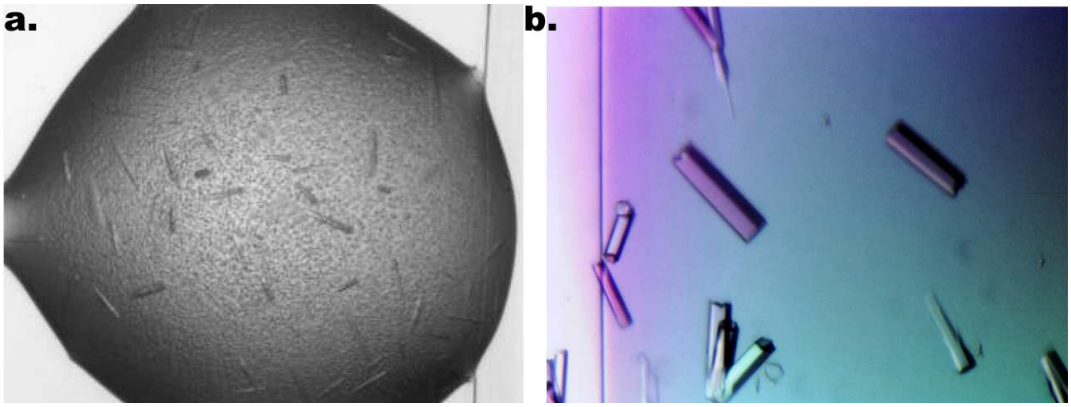


Fig. 7.9 Crystals of tPHD2Me-I and -II using the sample composition of ~18 mg.ml^{-1} protein, 1 mM Fe(II) and 1 mM BB-287. (a) tPHD2Me-I crystallised with heavy precipitation in well conditions containing 1.7 % polyethylene glycol 400 (v/v), 15 % glycerol (v/v), 1.7 M ammonium sulphate and 0.085 M HEPES-Na pH 7.5. (b) tPHD2Me-II crystallised in 1.1 M di-ammonium tartrate pH 7.0 in a hexagonal shape.

7.5.3 Structure solution and model building

The crystals of both tPHD2Me-I and tPHD2Me-II diffracted to ~2.4 Å in a typical $P6_3$ space group. The data collection and processing details are summarised in Table 7.3. The structures were determined by molecular replacement using Phaser[5] to calculate the phases (search model: PDB ID 2G19). The initial models used for tPHD2Me-I and tPHD2Me-II did not include any methylation of lysine residues. Iterative model building and rounds of refinements of these models (without methylation) were performed in CNS until the decreasing R_{cryst}/R_{free} with acceptable geometry no longer converged (R_{cryst} = 0.218 and R_{free} = 0.256).

Table 7.3 Data statistics for the tPHD2Me-I/II.Fe.BB-287 crystals and refinement statistics.

Measurement	**tPHD2Me-I.Fe.BB-287**	**tPHD2Me-II.Fe.BB-287**
Data Collection		
Space Group	$P6_3$	$P6_3$
Cell dimensions a,b,c (Å)	110.93 110.93 39.66	111.86 111.86 40.04
Resolution (Å)	36.3-2.4 (2.40-2.32)*	50-2.40 (2.49-2.40)*
No. of unique reflections	14397	11311
Completeness (%)	99.9 (99.9)*	99.0 (98.3)*
Redundancy	5.5 (5.5)*	5.5 (4.8)*
R_{sym}**	0.136 (0.784)*	0.181 (0.738)*
Mean I/σ(I)	8.4 (2.01)*	8.8 (2.20)*

*Highest resolution shell shown in parenthesis.

**R_{sym} = $\sum|I-\langle I\rangle|/\sum I$, where I is the intensity of an individual measurement and $\langle I\rangle$ is the average intensity from multiple observations.

In order to avoid model bias that may reinforce the refinement process towards a partially incorrect model, simulated annealed (SA) OMIT maps were calculated excluding all the lysines. However, electron densities for most of the lysines were not obvious enough to determine the methylation states accurately. Therefore, the structures were SA refined using a slow-cooling protocol for mono, di and tri-methylation states of each lysine residue (Table 7.4). For the final structures, the methylation state was assigned based on the lowest R_{free} and the results of MS/MS analyses (Tables 7.4 and 7.5).

Table 7.4 Monitoring R_{free} during the refinement of individual lysines in tPHD2Me-I and tPHD2Me-II with all three methylation states. Note that the apparently observed tri-methylation state (*) could also be an average of multiple alternate conformations of di-methylated lysines. (-) indicates residues that were not modelled.

	tPHD2Me-I					**tPHD2Me-II**				
Lys	R_{free}			Final	MS-	R_{free}			Final	MS-
No.	Mono	Di	Tri	structure	MS	Mono	Di	Tri	structure	MS
192	0.2563	0.2558	0.2572	2	n.d.	0.2549	0.2553	0.2547	3*	n.d.
204	0.2555	0.2551	0.2566	2	2	0.2572	0.2586	0.2587	-	n.d.
216	0.2571	0.2576	0.2567	2	2	0.2565	0.2563	0.2572	2	n.d.
234	0.2571	0.2566	0.2557	2	2/-	0.2575	0.2569	0.2585	2	2
244	0.2574	0.2567	0.2561	2	2	-	-	-	-	2/-
249	0.2575	0.2574	0.2571	-	2	0.2570	0.2560	0.2562	2	2
255	0.2554	0.2576	0.2566	2	2	0.2572	0.2572	0.2578	2	2/1/-
262	0.2553	0.2562	0.2574	1	2/1	0.2575	0.2574	0.2574	1	2/1/-
286	0.2565	0.2557	0.2582	2	2	0.2567	0.2566	0.2561	2	2
291	0.2580	0.2564	0.2571	2	2	0.2573	0.2559	0.2566	2	2
297	0.2559	0.2563	0.2583	1	2/1	0.2573	0.2571	0.2581	2	2
332	0.2563	0.2561	0.2549	3*	2/3/-	0.2558	0.2578	0.2577	1	2/-
337	0.2571	0.2560	0.2570	-	2	0.2575	0.2578	0.2578	-	2/-
350	0.2552	0.2567	0.2564	1	2	0.2572	0.2572	0.2578	-	2
359	0.2553	0.2560	0.2556	-	2	0.2574	0.2572	0.2563	-	2
400	0.2562	0.2576	0.2577	1	2/-	0.2553	0.2553	0.2630	2	2
402	0.2558	0.2560	0.2570	2	2	0.2565	0.2562	0.2569	2	2

MS/MS data interpretation:
2/- = mixture of di-methylated/unmodified lysines, 2/1 = mixture of di-methylated/mono-methylated lysines, 2/3/- = mixture of di-methylated/tri-methyalted/unmodified lysines; 2/1/- = mixture of di-methylated/mono-methylated/unmodified lysines.

Table 7.5 Final refinement statistics for the tPHD2Me-I/II.Fe.BB-287 structures with appropriately methylated lysines.

Measurement	tPHD2Me-I.Fe.BB-287	tPHD2Me-II.Fe.BB-287
Refinement		
R_{factor}	0.224	0.208
R_{free}	0.253	0.252
R.m.s. deviation		
Bond length, Å	0.007	0.006
Bond angle, °	1.37	1.26
No. of atoms		
Protein	1726	1711
Ligand	18 (BB-287)	18 (BB-287)
Water	45	60
$<B_{factor}>Å^2$		
Protein	46.11	43.79
Ligand	29.07 (BB-287)	33.42 (BB-287)
Metal	26.9 (Fe)	21.41 (Fe)
Water	43.98	44.77

7.5.4 Comparison with tPHD2 structures

The overall folds of tPHD2Me-I and tPHD2Me-II were similar to the previously solved tPHD2.Fe(II).BB-287 complex, with the β2/β3 loop in an 'open' conformation. There was no significant local change in the finger-lid or in the C-terminal α4 helix regions. Interestingly and consistent with the MS analyses (Fig. 7.8 and Table 7.4), tPHD2Me-I and tPHD2Me-II revealed multiple methylation states, unlike previous reports on other proteins (Table 7.4).[2] This could be because of the relative accessibility of lysines to the methylation reagents; some of the lysines in tPHD2 are protected in the DSBH core or buried in the secondary structure interfaces. Although lysine methylation did not alter the space group, many of the lysine residues, particularly Lys216, Lys262, Lys291, Lys350, Lys400 and Lys402, became more ordered upon methylation as evident in the $2F_o$-F_c electron density maps.

7.6 Activity analyses of methylated tPHD2

In order to investigate whether the lysine modification or more precisely, the variable lysine methylation states in tPHD2Me-I and tPHD2Me-II affected catalysis, both the products were tested for hydroxylation and 2OG turnover activity (Fig. 7.10). The results show that tPHD2Me-I was significantly less active than wt tPHD2 against both NODD and CODD (Fig. 7.10). It is important to appreciate that tPHD2Me-I is likely to be

heterogeneous population because of non-specific protein aggregation, as suggested. However, it may be that methylation affects specific interactions.

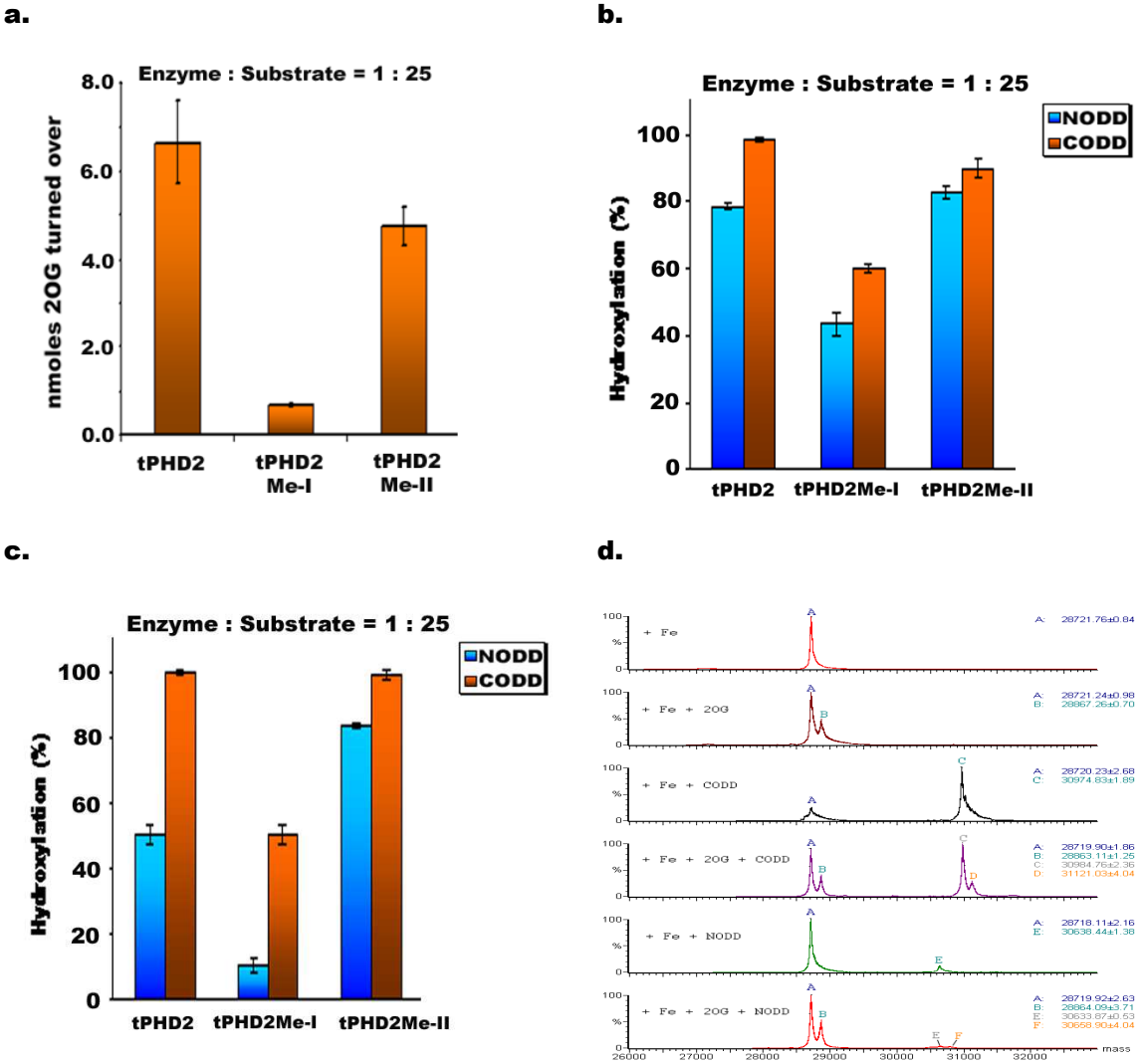


Fig. 7.10 Activity analyses of wt tPHD2 and the methylation products along with non-denaturing MS analyses on tPHD2Me-II. (a) 2OG turnover using CODD as a substrate; (b) individual substrate experiments where CODD and NODD were incubated separately; (c) competition experiments where both the substrates were incubated with the enzyme; (d) MS analyses of tPHD2Me-II under non-denaturing conditions. Peaks represent (A) tPHD2Me-II.Fe; (B) tPHD2Me-II.Fe.2OG; (C) tPHD2Me-II.Fe.CODD; (D) tPHD2Me-II.Fe.CODD.2OG; (E) tPHD2Me-II.Fe.NODD; (F) tPHD2Me-II.Fe.NODD$_{Hyp402}$. For comparison purposes, MS analyses on tPHD2 are given in Fig. 3.8, Chapter 3.

tPHD2Me-II was also less active than the wt tPHD2 in a 2OG turnover assay (Fig. 7.10a). In the tPHD2.CODD structure, Lys244$_{PHD2}$ and Lys400$_{PHD2}$ appeared to make hydrophobic contacts with CODD (Appendix 3). Lys400$_{PHD2}$ also forms electrostatic interaction with Asp571$_{CODD}$ (Asp571 Oδ1- Lys400 Nζ, 3.0 Å). Methylation of lysines could therefore alter binding including that involving Lys400$_{PHD2}$ and cause the activity differences.

In contrast, NODD hydroxylation appeared to be enhanced with tPHD2Me-II (Fig. 7.10). In addition to $Lys244_{PHD2}$, $Lys297_{PHD2}$ was predicted to have hydrophobic contacts with $Ile408_{NODD}$ in the tPHD2.NODD model 2 (Chapter 4 and Appendix 5). Therefore, it is speculated that methylation of $Lys244_{PHD2}$ and $Lys297_{PHD2}$ might alter the hydrophobic interactions with NODD and cause an increase in the hydroxylation with tPHD2Me-II.

7.7 Crystallisation of PHD2-FL (Full-length)

Attempts were also made to crystallise full-length PHD2 (PHD2-FL) with 2OG/various 2OG analogues at a protein concentration of ~15 mg/ml. PHD2-FL crystals were grown in complex with compound JLS3-67 (2 mM) using sitting drops containing 0.2 M Na-acetate, 30% PEG 8000 (w/v) and 0.1 M Na-cacodylate pH 6.5. The complex crystallised in a fully grown hexagonal crystals (300µ x 25µ x 25µ) within 30 days (Fig. 7.11). Crystals were frozen using 15% glycerol in the crystallisation buffer.

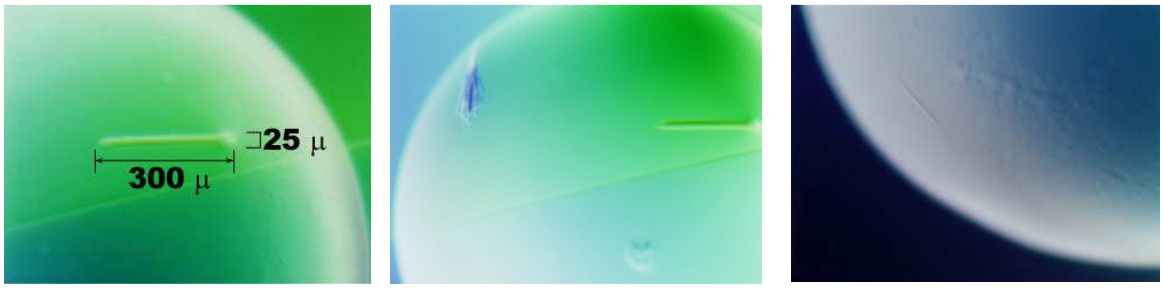


Fig. 7.11 Crystals of PHD2-FL.Fe(II).JLS3-67

Unfortunately, the PHD2-FL.Fe(II).JLS3-67 crystals did not diffract well. Data collection using in-house X-ray radiation (CuKα 1.542 Å) enabled space group determination. PHD2-FL crystallised in the $P6_3$ space group; the very similar cell constants of these crystals to those of tPHD2 indicated that the protein might have undergone proteolysis during crystallisation to yield an N-terminally truncated PHD2 that crystallised possibly in a homotrimeric form like tPHD2. However, the structure could not be solved due to insufficient diffraction data.

7.7.1 Further attempts for PHD2-FL crystallisation

A structure of full-length PHD2 (PHD2-FL) would be interesting since it could reveal the proposed inhibitory role of its N-terminal MYND domain in catalysis (Chapter 2). In the tPHD2.CODD crystal structure, the C-terminal $Leu574_{CODD}$ is close to the N-terminal $Gln184_{PHD2}$ of a symmetry related molecule (Gln184 N- Leu574 O, 5.2 Å). This positioning could be a result of crystal packing but it was also possible that the N-terminus

of PHD2 might alter substrate binding (Figure 7.12). However, more experimental data will be required to support this proposal.

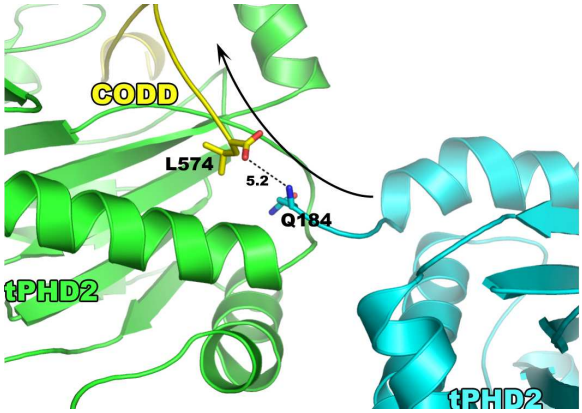


Fig. 7.12 View from the tPHD2.CODD crystal structure showing the N-terminus of a symmetry related tPHD2 molecule is close to the C-terminus of CODD.

MS analyses showed that recombinant PHD2-FL, produced in baculovirus and purified by Ni(II) affinity chromatography, was a heterogeneous mixture giving rise to at least five different peaks. These corresponded to ApoPHD2-FL, +Fe, +Fe+Zn and salt adducts (Fig. 7.13). With the aim of producing a homogeneous sample for crystallisation with ligands other than JLS3-67, apo-PHD2-FL was produced using a similar procedure as described elsewhere (Chapter 2, Section 2.3.2). The protein was then treated with one-fold excess of Zn(II). No significant crystal hits were obtained with Zn(II) containing PHD2-FL despite numerous attempts.

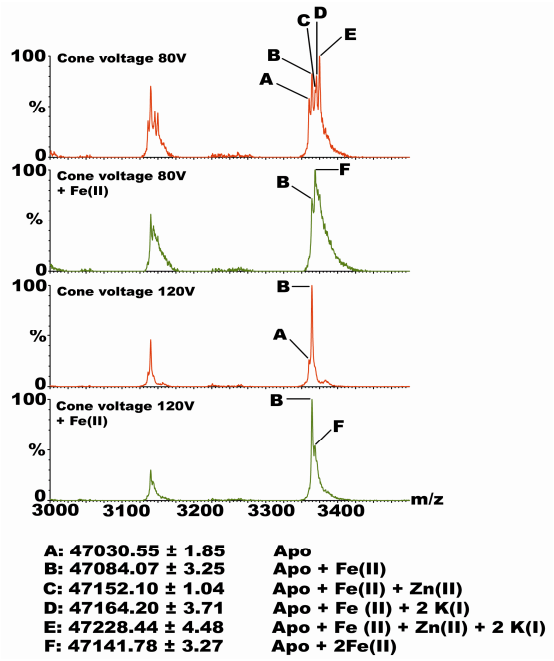


Fig. 7.13 Non-denaturing MS studies on PHD2-FL using two cone voltages, 80 V and 200 V. In addition to weaker Zn(II) binding, the protein revealed some salt adduct peaks at 80 V cone voltage, which could be removed at 200V.

7.8 Summary

The ability to induce formation of the homotrimeric crystal form did not correlate with in vitro inhibitory potency, e.g. all of Compounds A, B, and BB-287 enabled crystallisation with equal efficiency, yet A and B were significantly more potent inhibitors than BB-287 (most potent of this series) under standard in vitro assay conditions using CODD as substrate. The structure of tPHD2 complexed with BB-287/BB-328/JLS3-67 revealed a similar inhibitor binding mode to that observed in the reported tPHD2.Fe.Compound A/B structures.[1] In order to improve diffraction quality of tPHD2.Fe(II).BB-287 crystals, reductive lysine methylation was performed which yielded two products separable by size exclusion chromatography. Both of these products were crystallised under different crystallisation conditions. Variable lysine methylation states as observed in the crystal structures of the two products and/or multiple aggregation states of tPHD2Me-I may account for their activity differences. In addition, methylation appeared to improve the electron density for some of the lysine side chains.

7.9 References

1. McDonough, M.A., Li, V., Flashman, E., Chowdhury, R., Mohr, C., Lienard, B.M., Zondlo, J., Oldham, N.J., Clifton, I.J., Lewis, J., McNeill, L.A., Kurzeja, R.J., Hewitson, K.S., Yang, E., Jordan, S., Syed, R.S. & Schofield, C.J. Cellular oxygen sensing: Crystal structure of hypoxia-inducible factor prolyl hydroxylase (PHD2). *Proc Natl Acad Sci U S A* **103**, 9814-9 (2006).
2. Walter, T.S., Meier, C., Assenberg, R., Au, K.F., Ren, J., Verma, A., Nettleship, J.E., Owens, R.J., Stuart, D.I. & Grimes, J.M. Lysine methylation as a routine rescue strategy for protein crystallization. *Structure* **14**, 1617-22 (2006).
3. Tegley, C.M., Viswanadhan, V.N., Biswas, K., Frohn, M.J., Peterkin, T.A., Chang, C., Burli, R.W., Dao, J.H., Veith, H., Rogers, N., Yoder, S.C., Biddlecome, G., Tagari, P., Allen, J.R. & Hungate, R.W. Discovery of novel hydroxy-thiazoles as HIF-alpha prolyl hydroxylase inhibitors: SAR, synthesis, and modeling evaluation. *Bioorg Med Chem Lett* **18**, 3925-8 (2008).
4. Matthews, B.W. Solvent content of protein crystals. *J Mol Biol* **33**, 491-7 (1968).
5. McCoy, A.J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D Biol Crystallogr* **63**, 32-41 (2007).
6. Merkel, M., Pascaly, M., Krebs, B., Astner, J., Foxon, S.P. & Schindler, S. Chelate ring size variations and their effects on coordination chemistry and catechol dioxygenase reactivity of iron(III) complexes. *Inorg Chem* **44**, 7582-9 (2005).

Chapter 8 – Investigations into the role of nitric oxide in the hypoxic response

Following binding of HIF-α substrate, the next step in the catalytic cycle of the PHDs is proposed to be oxygen binding. Attempts to obtain an oxygen bound tPHD2 structure using varying amounts of H_2O_2 or O_2 were unsuccessful. Therefore nitric oxide (NO) was used as a substitute for molecular oxygen. Because NO has been used as an unproductive O_2 analogue in studies with other 2OG dependent oxygenases,[2,3] it is possible that NO competes with O_2 for iron binding. Indeed, there is growing evidence for regulation of the hypoxic response by NO. In normoxia, NO is proposed to stabilise HIF-α possibly via PHD inhibition.[1] The current study aimed to provide direct evidence of how NO can modulate HIF prolyl hydroxylation by crystallographic and EPR analyses in conjunction with biochemical investigations.

8.1 Endogenous regulation and chemistry of NO/ NO donors

8.1.1 Sources of NO

Endogenous NO is produced from the oxidation of L-arginine by constitutive (endothelial and neuronal) and inducible isoforms of nitric oxide synthase (NOS).[2] The inducible form of NOS is regulated by the HIF system, suggesting that increased NO production following HIF activation may establish a feedback loop. NO can also be generated from non-enzymatic sources including S-nitrosothiols, organic nitrates, iron-nitrosyl complexes, sydnonimines, C-nitroso compounds and secondary amine/ NO complex ions (reviewed in reference 3).

8.1.2 Reactions of NO

NO reacts in biological systems with oxygen (O_2), superoxide anion (O_2^-), and transition metals to give higher oxides (NO_x), peroxynitrite ($OONO^-$), and metal-NO adducts.[4] NO can also react with endogenous free thiols (a reaction known as S-nitrosylation) possibly in an oxygen dependent manner.[5] Various studies support the involvement of higher oxides such as N_2O_3 in S-nitrosothiol production.[8,9] The available evidence suggests that NO exists in cells as S-nitroso adducts of low molecular weight thiols such as glutathione and L-cysteine or abundant thiol containing proteins like albumin, haemoglobin and calbindin.[6]

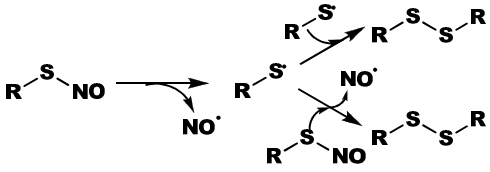
$$2NO + O_2 \rightarrow 2NO_2 \quad (1)$$

$$NO_2 + NO \rightarrow N_2O_3 \quad (2)$$

$$N_2O_3 + H_2O \rightarrow 2H^+ + 2NO_2^- \quad (3)$$

$$N_2O_3 + GSH \rightarrow H^+ + NO_2^- + GSNO \quad (4)$$

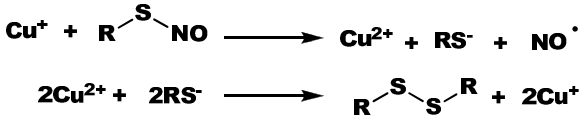
$$2GSNO \rightarrow GSSG + 2NO \quad (5)$$



Thus, S-nitrosothiols can serve as non-enzymatic sources of NO (reaction 5) via transnitrosation (thiol/S-nitrosothiol exchange) that can lead to regeneration of NO at the site of action via nucleophilic attack of a thiolate anion (RS^-) on the S-nitrosothiol sulphur.[10,11] Alternately, the reaction may proceed via formation of N-hydroxysulfenamides by the nucleophilic attack of thiol on the RSNO nitrogen.[7]

8.2 General role of S-nitrosylation

S-Nitrosylation/denitrosylation represents the most convenient way to store, transport and finally release NO.[13,14] In fact, the reactivity of RSNO species is proposed to be dependent on their ability to regenerate NO.[8] S-Nitrosothiols can decompose into NO and corresponding disulfides by the following routes: (a) spontaneously at a slow rate, (b) catalytically in the presence of trace amount of transition metals (e.g. Cu(I) and to a lesser extent, Cu(II)), or (c) photolytically by irradiation via homolytic cleavage of the S-NO bond to generate NO and RS·.[11,15,16]



The decomposition of S-nitrosothiols also appears to be accelerated by the presence of reducing agents such as ascorbate.[11,17]

8.3 Direct effect of NO on PHD2 mediated HIF hydroxylation

Studies on the effect of NO (or NO donors) on the HIF system have suggested that at least in some cell types, it has a bimodal profile with respect to oxygen availability.[1] Under normoxia, NO donors stabilise HIF-1α and promote HIF-mediated transcriptional activity in cultured cells[9] and in macrophages during bacterial infection.[10] In contrast, under hypoxic conditions, NO and reactive nitrogen species (RNS) antagonise HIF activity.[11,12]

There is evidence that the effect of NO on the hypoxic response system is related to HIF-α prolyl hydroxylation.[11,12] Despite numerous studies, very little is known about how NO interacts with PHDs at a molecular level and how it attenuates the hydroxylation of prolines in NODD and CODD. In the current study, experiments were performed with the aim of investigating the possibility of direct reaction of NO with tPHD2 and its effect(s) on PHD catalysis.

8.3.1 NO diminishes PHD2 activity

The effect of 'free' NO on the activity of the catalytic domain of PHD2 (tPHD2) was initially investigated. Freshly prepared NO solutions (ranging from 3 to 50 μM) were added to tPHD2 (4 μM) under near anaerobic (PO2 <0.1 ppm) conditions for an hour prior to activity assays. The NO-treated tPHD2 was then assayed by using a 2OG decarboxylation assay under standard aerobic conditions. The data was analysed by Graphpad Prism software to calculate the inhibition profile (Fig. 8.1). The results reveal that, under standard incubation conditions, NO inhibits tPHD2 in a dose dependent manner with an IC_{50} of 10 μM.

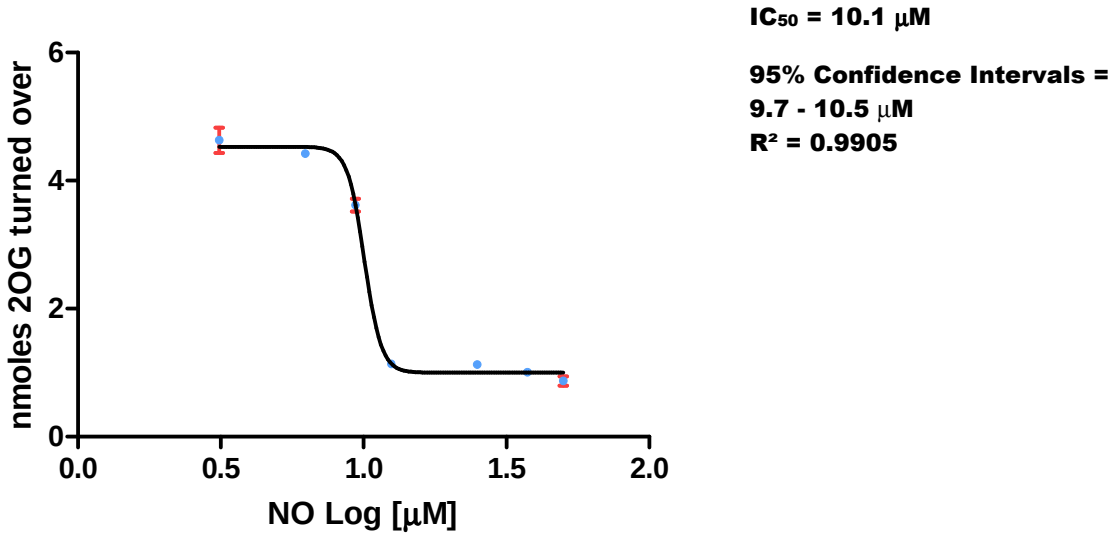


Fig. 8.1 PHD2 inhibition by NO. NO (~3-50 μM) was added to tPHD2 (4 μM) and incubated for an hour prior to activity analyses.

8.3.2 NO alters Fe(II) EPR signals in PHD2

To test whether NO binds to Fe(II) at the PHD2 active site, as has been observed for some other 2OG oxygenases and related enzymes,[2,3] electron paramagnetic resonance (EPR) experiments were then performed in collaboration with Dr E.G. Flashman (Chemistry Research Laboratory, Oxford) (Fig. 8.2). The spectrum of tPHD2 alone revealed a small

feature at g=4.3, characteristic of a small amount of ferric iron present in the sample, as observed previously.[13] Anaerobic addition of dissolved NO resulted in the formation of an EPR signal in the g=4 region, typical of the [FeNO], S=3/2 signal previously observed for NO binding to the ferrous active site of other non-haem iron oxygenases.[19,20] Compared to the tPHD2.Fe(II) complex, the tPHD2.Fe(II).2OG complex was apparently unable to/has a very much reduced ability to bind NO at its active site Fe(II) with only a small feature appearing in the g= 4 region (at least within the timescale of analyses).

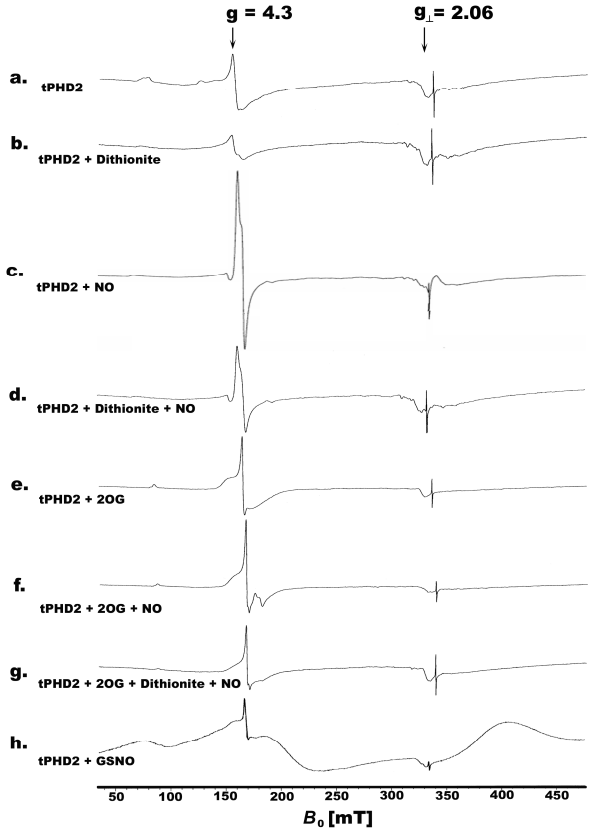


Fig. 8.2 EPR spectra of tPHD2 complexes under different conditions. GSNO = S-nitrosoglutathione.

In the tPHD2.Fe(II) complex, up to three unoccupied coordination sites are likely to be hydrated as observed for other 2OG oxygenases.[14] In principle, NO could replace any one, or more, of these water molecules producing a paramagnetic Fe(III) signal in EPR. In the tPHD2.Fe(II).2OG complex, NO may have to replace a tightly metal coordinated water molecule hydrogen-bonded to $Asp315_{PHD2}$, rationalising the reduced level of NO binding to the tPHD2.Fe(II).2OG complex. No signal for [FeNO] was detectable with a substantial intensity when GSNO (250 equivalents) was added under near anaerobic conditions (Fig. 8.2h), raising the possibility that NO itself and NO donors may react differently with PHD2 under the experimental conditions.

8.4 Crystallisation attempts – evidence for S-nitrosylation

Attempts were then made to obtain a crystal structure of a tPHD2.Fe.NO complex. Because NO and NO donors may react differently, initial attempts were made to co-crystallise tPHD2 with both NO and a NO donor, GSNO. However, these attempts were not successful. Next, pre-grown tPHD2.Fe(II).compound B crystals were exposed to NO gas under anaerobic conditions, whilst tPHD2.Zn(II).B crystals were soaked with 50 mM GSNO prior to cryo-cooling in liquid nitrogen. Structural models were generated by molecular replacement using the tPHD2.Fe(II).A structure (PDB ID: 2G19). Data are given in Table 8.1. NO binding to the Fe(II) was not observed in either of the structures obtained. Given that 2OG appeared to block NO binding to Fe(II) as observed by EPR (Fig. 8.2), it is possible that the bicyclic inhibitor (compound B), used to induce tPHD2 crystallisation, also blocks NO binding. However, both datasets had positive difference density extending from the side chain of $Cys302_{PHD2}$, suggesting the addition of an NO group to this residue.

Table 8.1 Crystallographic data processing and refinement statistics.

Measurement	**tPHD2-SNO.Fe(II).B (NO exposed)**	**tPHD2-SNO.Fe(II).B (GSNO soaked)**	**tPHD2 C302A.Fe(II).C**
Data Collection			
Space Group	$P6_3$	$P6_3$	$P6_3$
Cell dimensions a,b,c (Å)	110.30	111.05	110.31
	110. 30	111.05	110.31
	40.38	40.33	39.99
Resolution (Å)	50-2.10 (2.18-2.10)*	37.19-2.0 (2.11-2.0)*	50.0-2.60 (2.69-2.60)*
No. of unique reflections	16616	19446	15639
Completeness (%)	99.7 (100.0)*	99.7 (98.1)*	93.3 (79.6)*
Redundancy	3.0 (3.0)*	5.1 (3.6)*	3.6 (2.7)*
R_{sym}**	0.069 (0.387)*	0.102 (0.343)*	0.059 (0.436)*
Mean I/σ(I)	15.02 (2.78)*	14.2 (3.2)*	15.47 (1.66)*
Refinement			
R_{factor}	0.237	0.213	0.222
R_{free}	0.263	0.245	0.270
R.m.s. deviation			
Bond length, Å	0.01	0.01	0.01
Bond angle, º	1.24	1.30	1.24
No. of atoms			
Protein	1678	1710	1686
Ligand	19 (B)	19 (B)	20 (C)
Water	81	143	32
$<B_{factor}>Å^2$			
Protein	38.57	31.38	55.64
Ligand	17.17 (B)	18.65 (B)	33.72 (C)
Metal	25.58 (Fe)	17.43 (Zn)	35.54
Water	39.66	41.32	42.25

*Highest resolution shell shown in parenthesis.

**$R_{sym} = \sum|I-<I>|/\sum I$, where I is the intensity of an individual measurement and $<I>$ is the average intensity from multiple observations.

The tPHD2SNO.Fe(II)/Zn(II).B structures were initially refined using the stereochemical restraints for an S-nitrosocysteine (SNC) residue as described for the crystal structure of blackfin tuna myoglobin (PDB ID: 2NRM, 1.09 Å resolution[15]) with the C–S–N–O dihedral angle un-restrained. Attempts to refine the structural models imposing planar geometric restraints for the cis/trans-SNO dihedral angles were not productive because the O of the SNO group could not be fitted within the experimental electron density. The exact chemical nature of S-nitroso tPHD2 adducts is therefore unclear. Considering the apparent dihedral angle (~90°) observed for the C–S–N–O group in tPHD2SNO.Fe(II)/Zn(II). inhibitor complexes, it is possible that the Cys-302 bound nitrogen oxide is at least in part, in the R–S–N•–OH form, a proposed reaction intermediate of NO with thiols[16] or the O-centered radical thionitroxide radical (R–S–NH–O•) form;[17] other potential forms such as R-S-NH-OH may also result from partial reduction during data collection using synchrotron radiation.

So far structural insights into S-nitrosylation have lagged behind biological work. To date, there are only a few protein structures with S-nitrosothiols (nitrophorin[18], haemoglobin[19], myoglobin[15] and thioredoxin-1[20]). S-Nitrosylation has not been shown to cause significant structural changes. Superimposition of the tPHD2SNO.Fe(II).B complexes ($P6_3$) with tPHD2.Fe(II).A complex ($P6_3$) structure revealed no significant structural change, although superimposition with the tPHD2.CODD complex ($P2_12_12_1$) suggests a few local changes (Fig. 8.3).

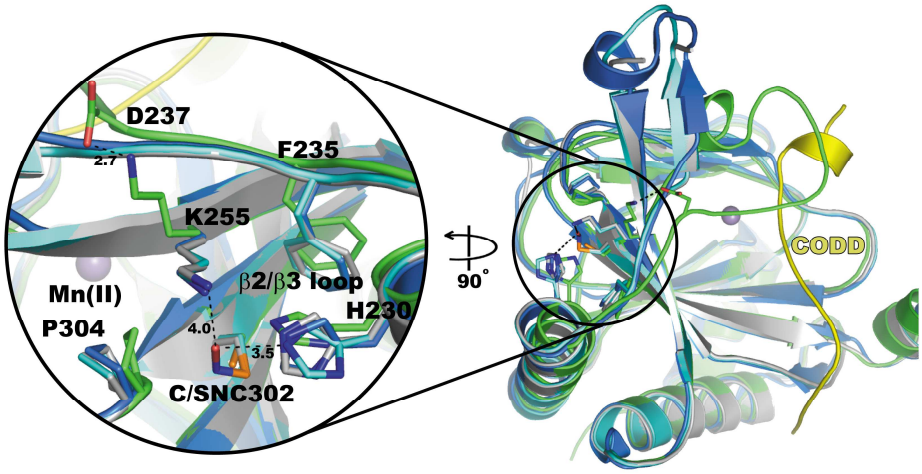


Fig. 8.3 Superimposition of tPHD2.CODD, tPHD2.Fe(II).A, tPHD2.Fe(II).B(NO) and tPHD2.Fe(II).B(GSNO) structures showing local changes around the S-nitrosylated Cys302$_{PHD2}$.

Nitrosylation increased the length of Cys302$_{PHD2}$ side chain by ~2.5 Å; the SNO group of Cys302$_{PHD2}$ is apparently positioned to make weak interaction with the neighbouring

$Lys255_{PHD2}$ (Fig. 8.3). In contrast, in the tPHD2.CODD structure, $Lys255_{PHD2}$ projects towards the protein interior and forms an electrostatic interaction with $Asp237_{PHD2}$ (K255 Nζ-D237 Oδ1, 2.7Å) (Fig. 8.3). However, the conformational changes of $Lys255_{PHD2}$, $Pro304_{PHD2}$, $His230_{PHD2}$ and $Phe235_{PHD2}$ side chains were apparently due to substrate binding and/or β2β3/loop closure (Chapter 2) and therefore are less likely to correlate with the S-nitrosylated $Cys302_{PHD2}$ side chain orientation.

$Cys302_{PHD2}$ is conserved in all three human PHDs and is located in a conserved hydrophobic region formed by the side chains of $Lys255_{PHD2}$, $Pro304_{PHD2}$, $Thr382_{PHD2}$, $Tyr384_{PHD2}$, $His230_{PHD2}$ and $Phe235_{PHD2}$. Thus, access to $Cys302_{PHD2}$ is largely hindered and the residue lacks solvent accessibility as calculated by AREAIMOL[21] and NACCESS[22] (Table 8.2). The buried nature of $Cys302_{PHD2}$ is consistent with those of the Cys residues that were reported to be nitrosylated when thioredoxin was treated with GSNO.[20]

Table 8.2 Calculated solvent accessibility and pKa for seven cysteines in tPHD2

Residue	Solvent accessible Surface area (Å^2)[a]	Solvent accessible Surface area (Å^2)[b]	Calculated pK_a [c]
Cys201	12.80	14.19	8.77
Cys208	0.00	0.00	12.71
Cys266	4.30	4.02	9.00
Cys283	0.70	0.40	9.94
Cys302	0.00	0.00	9.51
Cys323	0.00	0.00	9.99
Cys326	0.00	0.00	12.06

[a] The accessible surface area (ASA) of Cys Sγ atoms was calculated using AREAIMOL[21] tool in the CCP4 Programme Suite with a 1.4 Å probe solvent radius.

[b] The accessible surface area (ASA) of Cys Sγ atoms was calculated using NACCESS[22] with a 1.4 Å probe solvent radius.

[c] Cysteine thiol pK_a values were calculated using tPHD2 structure (PDB 2G19) with the protein pKa prediction software, PROPKA[30,31]

Although the $Cys302_{PHD2}$ side chain projects away from the active site iron, $Cys302_{PHD2}$ is located on a region that forms part of the active site. The residue is located adjacent to $Tyr303_{PHD2}$, a residue proposed to be involved in catalysis.[23] Variation analyses revealed that Y303F retained catalytic activity whilst a Y303A variant lost hydroxylation but gained 2OG turnover activity mostly because of the uncoupled turnover (Chapter 6). To investigate whether modification of $Cys302_{PHD2}$ had any effect on PHD2 catalysis, a C302A variant was produced and its activity was measured by 2OG turnover and

hydroxylation assays. The results showed that the variant can catalyse 2OG turnover and substrate hydroxylation at levels comparable to wt tPHD2 (data shown later). To investigate whether substitution of $Cys302_{PHD2}$ had any effect on structure, the variant was crystallised with compound B. The structure of tPHD2 C302A variant in complex with Fe(II) and a 2OG analogue, compound C[32,33] was solved by molecular replacement using PDB ID 2G19 (Space group $P6_3$). The redundant structure did not reveal any perturbation or local changes when compared to the wt tPHD2.Fe(II).A/B complexes. Data statistics are in Table 8.1.

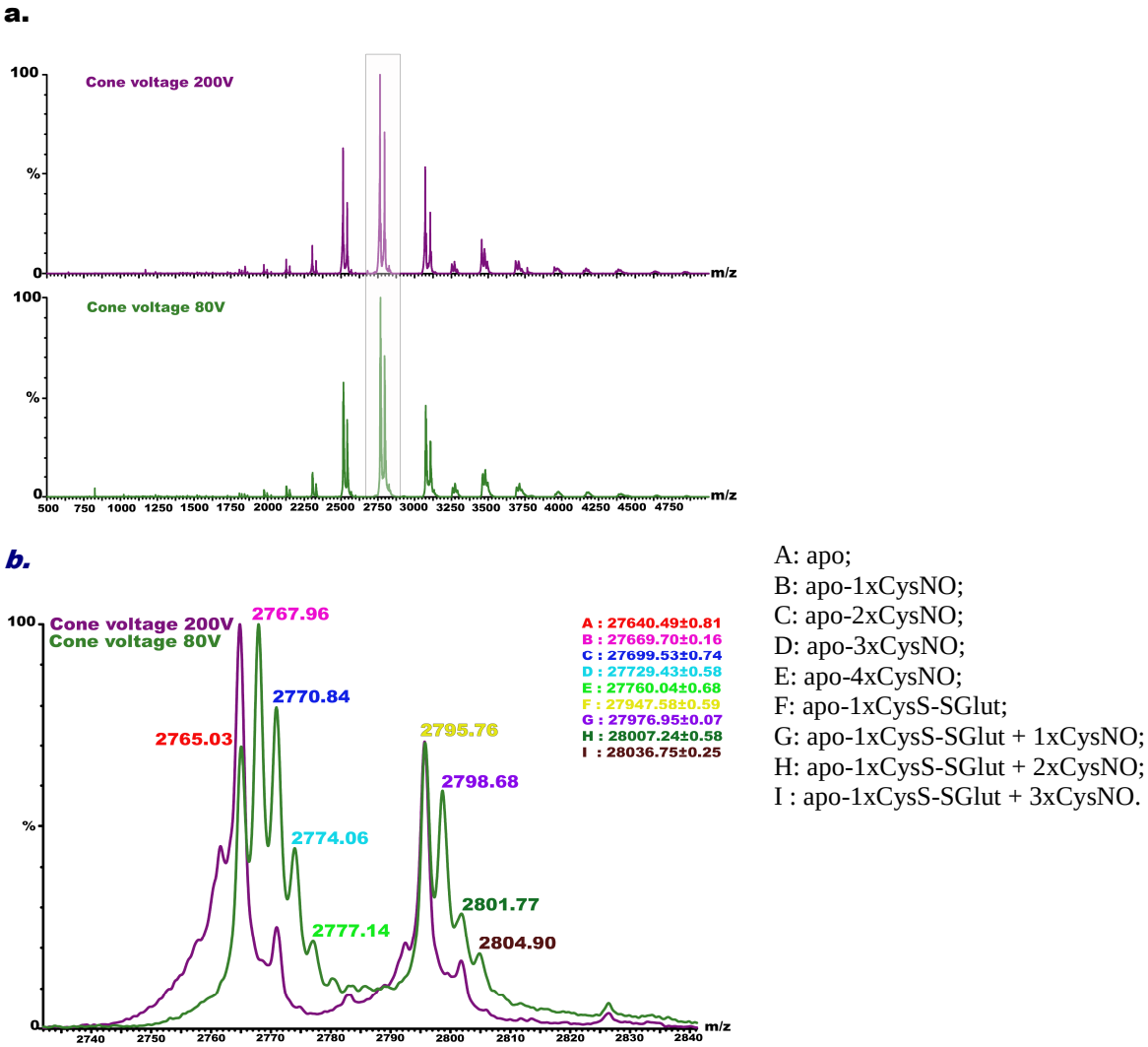
8.5 Characterisation of S-nitrosylated tPHD2 in solution

To investigate whether S-nitrosylation of tPHD2 occurs in solution, tPHD2.Fe(II) and apo-tPHD2 were treated with GSNO (250 equivalents) followed by removal of excess GSNO using size exclusion chromatography (SEC).

8.5.1 Preparation of S-nitrosylated apo-tPHD2 and native MS analyses

Mass spectrometry (MS) under non-denaturing conditions was then used to test for modification of the GSNO treated tPHD2 that was successfully used for other proteins.[34,35] The MS analyses indicated that GSNO was able to nitrosylate both tPHD2.Fe(II) and apo-tPHD2 to similar levels (Fig. 8.4; data for S-nitrosylated tPHD2.Fe(II) not shown). With this technique it is not possible to distinguish readily between different types of modifications, e.g. between reaction at Fe(II) and reaction with cysteinyl thiols. However, treatment of apo-tPHD2 with an excess of GSNO resulted in a heterogeneous mixture containing multiple nitrosylated products (Fig. 8.4). Regardless of the applied cone voltage (80 V or 200 V), the spectra showed two series of peaks: At a cone voltage of 80 V, the lower mass series revealed mass increases of ~30 Da, corresponding to the addition of NO, in addition to the unmodified protein, suggesting substantial covalent binding of between one and four molecules of NO (Fig. 8.4). The higher mass series suggested formation of a disulfide adduct between apo-tPHD2 and glutathione (+307.1 Da relative to the apo-tPHD2). In addition, multiple signals were observed indicating that the apo-tPHD2-glutathione product bound to between one and three NO molecules in this higher mass series (Fig. 8.4b). At a cone voltage of 200 V, the apo-tPHD2-glutathione adduct was stable (Fig. 8.4b), but most of the observed peaks corresponding to nitrosylation were lost. Interestingly, two NO molecules apparently still remained bound to the enzyme at a cone

voltage of 200 V, although their signal intensities were significantly reduced compared to those at 80 V.



A: apo;
B: apo-1xCysNO;
C: apo-2xCysNO;
D: apo-3xCysNO;
E: apo-4xCysNO;
F: apo-1xCysS-SGlut;
G: apo-1xCysS-SGlut + 1xCysNO;
H: apo-1xCysS-SGlut + 2xCysNO;
I : apo-1xCysS-SGlut + 3xCysNO.

Fig. 8.4 MS spectra showing the products of tPHD2 after treatment with GSNO at cone voltages of 80V and 200V. (a) Overall spectra showing two series of peaks (b) Analysis of the most stable complexes at the +10 charge state reveals apparent S-nitroso as well as mixed disulfide adducts of glutathione.

8.5.2 LC-MS analyses

To investigate the sites of S-nitrosylation on tPHD2 in solution, tPHD2 was treated with a GSNO/NO solution, followed by in-solution tryptic digestion and analysis of the resultant fragments by LC-MS. Although it has been questioned whether it would be possible to detect labile S-nitrosylation[36] by MS, recently, 'direct' measurement of protein S-nitrosylation by LC-MS analyses have been reported.[35,37,38] In the present study, apparent S-nitrosylation of four cysteines, Cys201 (>30%), Cys208 (<5%), Cys302 (>50%, charge state z = 2; >25%, z = 1), Cys323/326 (>35%) were detected (Fig. 8.5); S-nitrosylation at Cys283 cannot be ruled out because no peak was observed for the fragment ^{282}H**C**NGK^{286}

in either the untreated or GSNO/NO treated tPHD2 samples. Of the four S-nitrosylation sites, one site could not be unambiguously assigned because SNO-Cys peptide contained two cysteine residues, $Cys323_{PHD2}$ and $Cys326_{PHD2}$. Further MS/MS analyses of the tryptic fragments support S-nitrosylation of Cys-201, Cys-208 and Cys-302 (Fig 8.5).

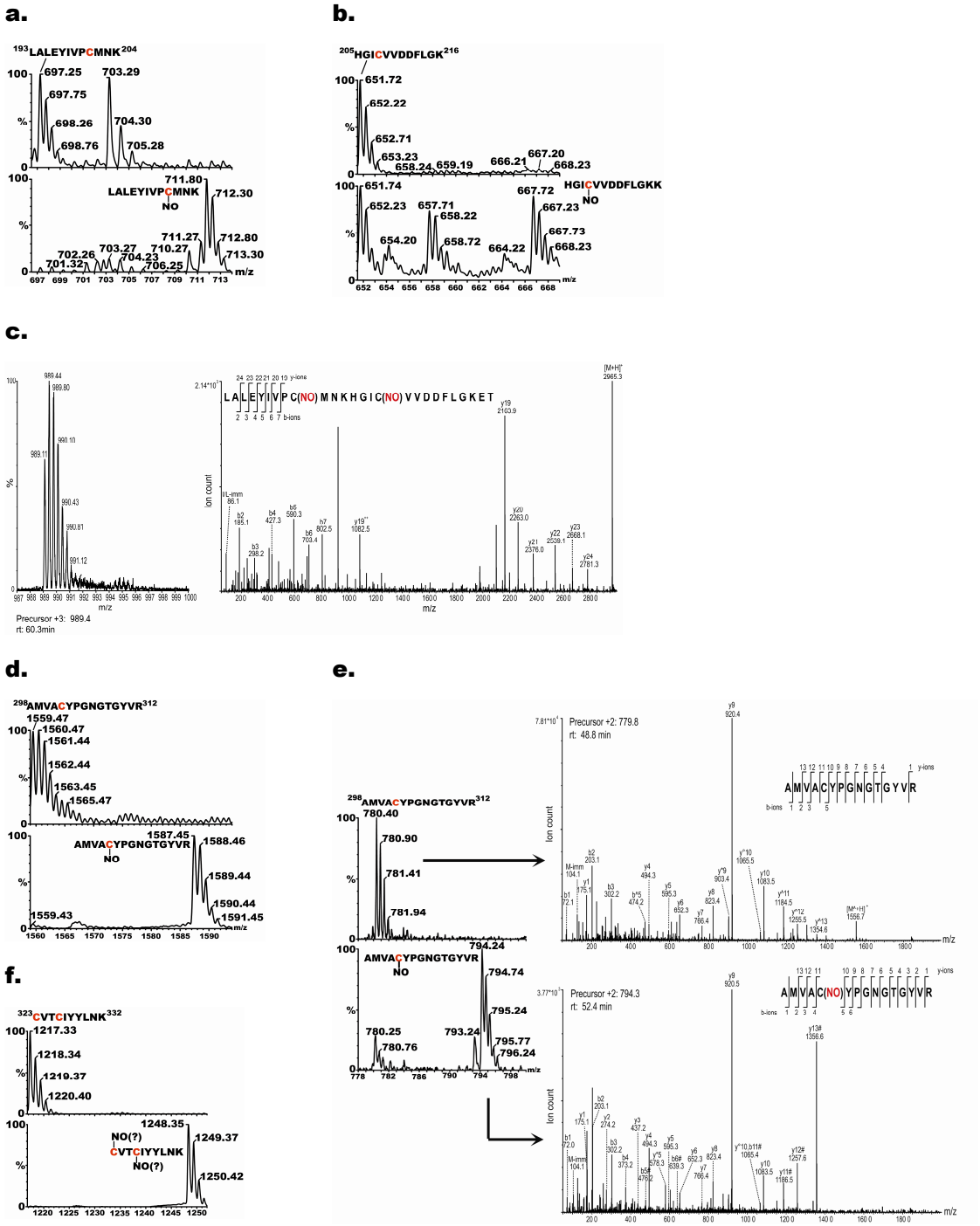


Fig. 8.5 MS analyses demonstrating S-nitrosylation of Cys201, Cys208, Cys302 and Cys323/326 after treatment of tPHD2 with GSNO. (a) and (b) LC-MS spectra showing tryptic fragments containing Cys-201 (a), Cys-208 (b) before (*upper*) and after (*lower*) treatment with GSNO. (c) MS/MS analyses of a doubly S-nitrosylated precursor ion containing Cys-201 and Cys-208. (d) and (e) LC-MS resolved fragments containing Cys-302 prior to (*upper*) and after (*lower*) GSNO treatment at +1 (d) and

+2 (e, *left*) charge states and the MS/MS spectra of the corresponding precursor ions (e, *right*); (f) LC-MS resolved mono-S-nitrosylated fragment containing Cys-323 and Cys-326. MS/MS analyses show y- and b- fragment ions as well as immonium ions where observed. The following processes, that can occur during peptide fragmentation in tandem mass spectrometry experiments, are denoted as follows: ° Loss of water/ dehydration (-18 Da); * Loss of ammonia/deamination (-17 Da); # Loss of nitric oxide from *S*-nitrosocysteine/ denitrosylation with protonation (-29 Da); ^ Formal loss of 2H from free cysteine (-2 Da).

A modification of a previously described 'biotin-switch' method[36] was then used to further investigate the selectivity of S-nitrosylation of Cys residues on tPHD2. After treatment of tPHD2 with GSNO/NO solution, the unmodified cysteines in the protein were capped via methylthiolation by reaction with methyl methanethiosulfonate (MMTS). Under the reaction conditions, methyl methanethiosulfonate is reported not to react with nitrosothiols or disulfide bonds (Fig. 8.6).[36] The nitrosothiols were then reduced with ascorbate to produce free thiols. The proteins were then digested with trypsin and the fragments separated and analysed by LC-MS.

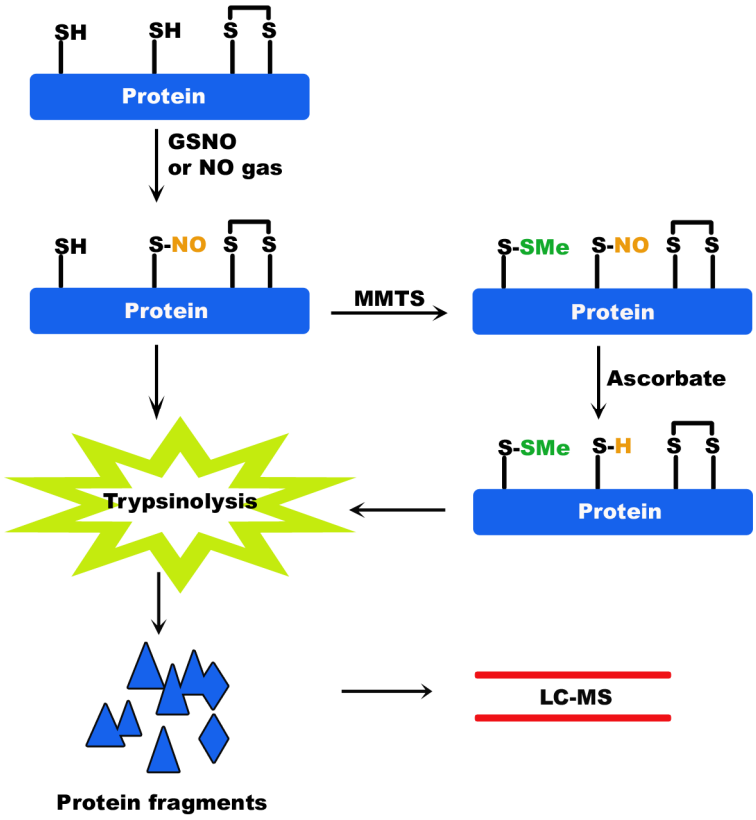


Fig. 8.6 Schematic representation of direct and indirect approaches for measuring S-nitrosylation of PHD2. Treatment of S-nitrosylated protein with MMTS followed by reduction with ascorbate was performed. S-nitrosylated and S-methylthiolated proteins were subjected to trypsin mediated proteolysis and the derived fragments were analysed by LC-MS/MS.

Consistent with the non-denaturing MS data and the results of 'direct' measurement of the S-nitrosothiols, S-methylthiolation (indicative of non-nitrosylation) was observed >60%

for Cys201, >70% and >80% for Cys208 (charge states, z = 1 and 2, respectively), >90% for Cys266 (z = 2), >70% and >50% for Cys302 (z = 1 and 2, respectively) fragments. Interestingly, two additional sets of signals were observed corresponding to S-Me capped ^{323}**C**VT**C**IYYLNK^{332} fragments at a charge state, z = 1: ~10% for a mono and >55% as a di S-Me capped fragments. The results suggest that one of the two cysteines within this fragment is S-nitrosylated (<45%); the other cysteine is <5% (if not at all) S-nitrosylated. This result is consistent with the results from direct measurement. Additional support for mono-S-nitrosylation of this fragment came from the observation of a ~90% S-Me capped cysteine.

8.6 Effects of S-nitrosylation on purified PHD2

Further experiments were carried out to investigate the role of PHD2 S-nitrosylation on catalysis.

8.6.1 Role of reducing agents on denitrosylation

It has been well studied with numerous proteins that S-nitrosylation and corresponding denitrosylation reactions are redox dependent reversible modifications. The effects of ascorbate and thiol-based reducing agents, such as dithiothreitol (DTT), on these reactions are well characterised.[24] However PHD2, like some other 2OG oxygenases (e.g. AlkB)[25] requires L-ascorbate for activity.[26] This essentially creates difficulty in determining the effect of S-nitrosylation on PHD2 activity. Initial attempts were made to reduce the amount of ascorbate in assays or to replace ascorbate with alternative reducing agents that might not interfere with S-nitroso adduct formation.

The results (Fig. 8.7) showed that a combination of L-ascorbate (4 mM) and DTT (1 mM) led to optimal 2OG turnover by tPHD2. None of the reducing agents tested were individually capable of maintaining full activity. Use of either DTT or β-mercaptoethanol retained >40% 2OG turnover. TCEP and Na-dithionite resulted in a significant loss of activity (<20%) relative to the optimal conditions. The effect of a NO-donor, GSNO, on PHD activity was then tested in the presence of the above reducing agents/ combinations. Addition of GSNO at 100 μM final concentration did not appear to have a significant effect on 2OG turnover (Fig. 8.7).

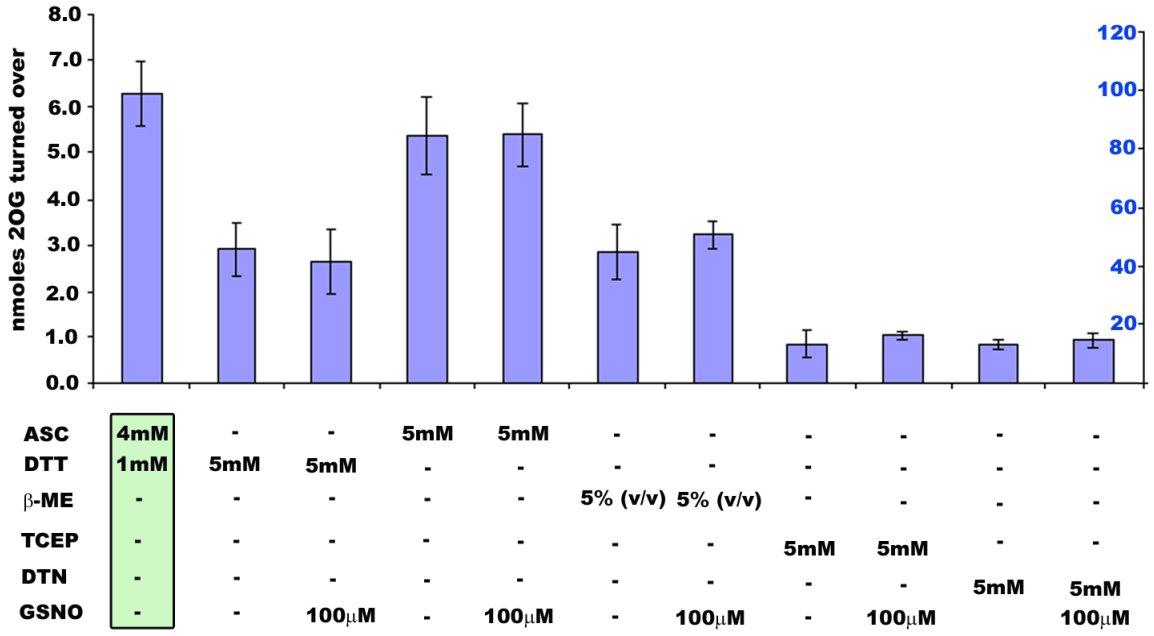


Fig. 8.7 Effect of different reducing agents on 2OG turnover by tPHD2. tPHD2 and CODD concentrations used in these experiments were 4 and 100 μM, respectively. The first column is the standard assay mixture containing 4mM L-ascorbate and 1mM dithiothreitol. Effects of different ascorbate analogues have been reported elsewhere.[26] ASC = L-ascorbate, DTT = dithiothreitol, β-ME = β-mercaptoethanol, TCEP = Tris(2-carboxyethyl)phosphine, DTN = Na-dithionite.

8.6.2 Time dependence

To investigate whether a longer reaction time can increase PHD2 S-nitrosylation and alter catalytic activity, additional experiments were performed over longer periods (up to 1 hour). Because ascorbate was necessary for activity and was able to retain >80% 2OG turnover as the sole reducing agent (Fig. 8.7), further experiments were performed using 4 mM ascorbate without addition of DTT. Purified tPHD2 (4 μM) was treated with three different concentrations of GSNO (100 μM, 400 μM and 1 mM) for four time points (5 min, 15 min, 30 min and 1 hour). As shown in Fig. 8.8, the nmoles of 2OG turned over increased as a function of time and was almost unaffected by the concentration of GSNO.

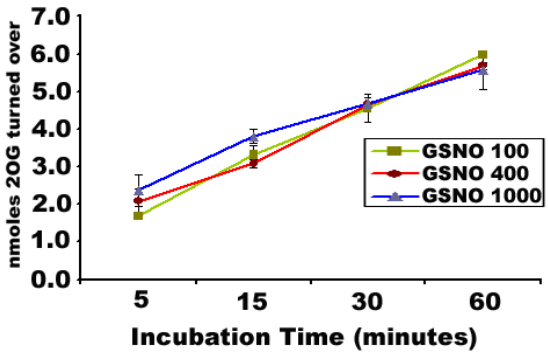


Fig. 8.8 Time course experiment showing addition of GSNO to the reaction mixture did not have significant effect on 2OG turnover activity. In the control experiment, tPHD2 in the absence of GSNO catalysed 2.96 ± 0.28 nmoles of 2OG turnover after 15 minutes of incubation.

8.6.3 Pre-incubation

The studies described in the previous sections revealed that GSNO, when added to the reaction mixture, does not appear to have a significant effect on PHD2 activity. This may due to the destabilising effect of reducing agent(s) on S-nitroso adducts. Further experiments were performed where tPHD2.Fe(II) was pre-incubated with variable amounts of GSNO (before adding other co-substrates) and 2OG decarboxylation assays were carried out under standard (aerobic) incubation conditions.

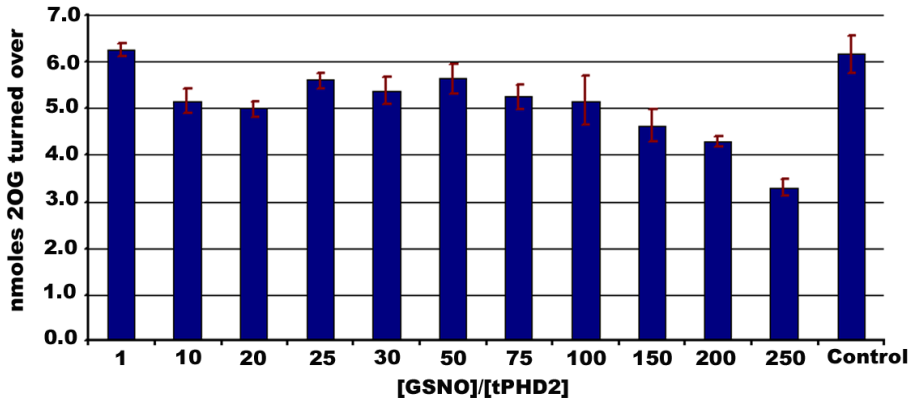


Fig. 8.9 2OG turnover activities of tPHD2-SNO when tPHD2 (4μM) was pre-incubated with appropriate amounts of GSNO under oxygenated conditions.

The results show that at 250 equivalents, GSNO appeared to reduce tPHD2 activity to below 60% of the level of untreated tPHD2 (Fig. 8.9). In contrast, when apo-tPHD2 was pre-incubated with GSNO, no inhibition was observed under aerobic conditions, indicating that NO mediated inhibition likely occurs via active site Fe(II) binding. These observations suggested the possibility of a redox dependent release of NO from S-nitrosylated tPHD2/ thionitrite.

8.6.4 Oxygen dependence

Previous reports proposed that NO release from thionitrites could be accelerated using molecular O_2.[11,12] An additional set of experiments was therefore conducted to investigate the possible role of O_2 in 'NO release' from S-nitrosylated tPHD2. The 2OG turnover activity was measured for GSNO treated tPHD2 using oxygenated or reduced buffers prepared in an anaerobic glove box. For activity analyses using reduced buffers, the pre-incubation step with GSNO (250 equivalents) was performed under anaerobic conditions to minimise O_2 contamination.

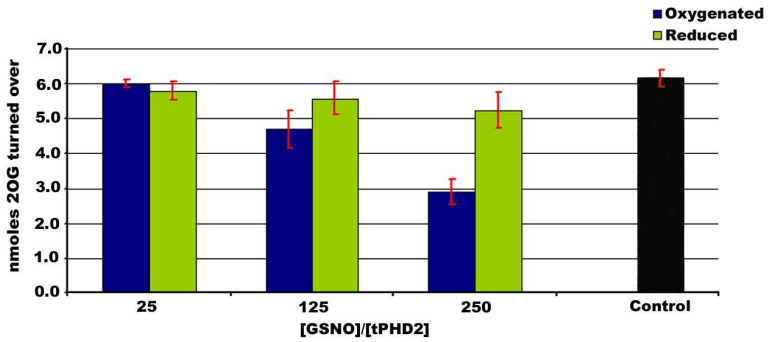


Fig. 8.10 2OG turnover activities of tPHD2-SNO when tPHD2 (4μM) was pre-incubated with appropriate amounts of GSNO under reducing (green) and oxygenated (dark blue) conditions.

The results indicate that GSNO, in the presence of oxygenated buffers, significantly reduces (~50%) the ability of PHD2 to turnover 2OG (Fig. 8.10). A dose-dependent reduction was also observed under reducing conditions but the extent of activity reduction was much less than under oxygenated conditions.

In addition to GSNO, another NO-donor, S-nitroso-N-acetyl-DL-penicillamine (SNAP) was used to assess the effects of S-nitrosylation on tPHD2 catalysis under oxygenated conditions. Similar to the results with GSNO, SNAP reduced the 2OG turnover activity of tPHD2 in a dose-dependent manner under oxygenated conditions (Fig. 8.11).

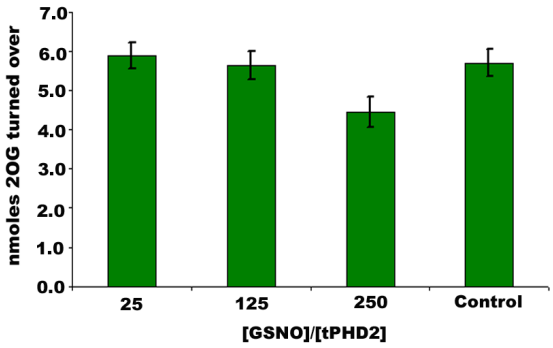


Fig. 8.11 2OG turnover activities of tPHD2-SNO when tPHD2 (4 μM) was pre-incubated with variable amounts of SNAP under non-reducing condition.

8.6.5 Effects of transition metals

NO release via de-nitrosylation of RSNO can be stimulated by traces of transition metals such as Cu(I)[27] or Cu(II)[28], reducing agents including thiols and ascorbate.[24] Several other metal ions have been examined as possible reagents for RSNO decomposition, including Zn(II), Ca(II), Mg(II), Ni(II), Co(II), Mn(II) and Cr(III); none catalyse thionitrite decomposition.[29] Some reports suggested weak catalytic decomposition by Fe(II), whilst Fe(III) was inactive.[24] In order to investigate such a role in denitrosylation, further

experiments were performed using tPHD2apo-SNO (SEC purified, Section 8.5.1) with varying amounts of Cu(I)/Cu(II). Because Cu(II) has the ability to replace Fe(II) at the PHD2 active site,[30] initial experiments were performed to see whether Cu(I) and Cu(II) have inhibitory roles in 2OG turnover.

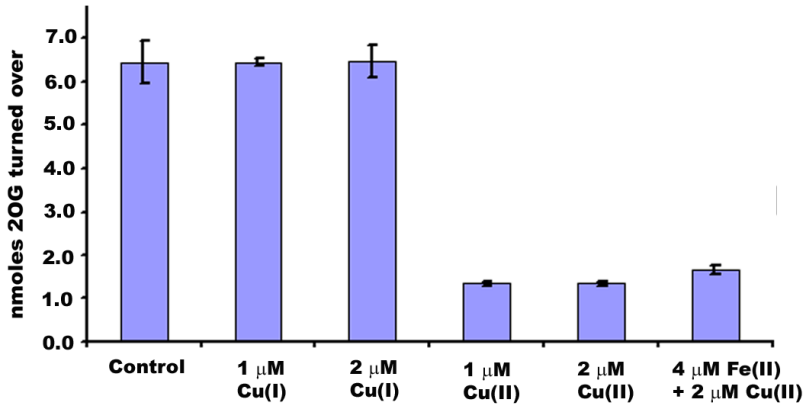


Fig. 8.12 Effect of Cu(I) and Cu(II) on 2OG turnover of tPHD2.

The results revealed that Cu(II) (but not Cu(I)) was found to inhibit tPHD2.Fe(II) at sub-stoichiometric concentrations (≤ 2 μM, 0.5 equivalent) under the standard conditions (Fig. 8.12), consistent with a report that Cu(II) inhibits PHD2 (IC_{50} 2.3 μM), independent of the iron concentration[45] (Fig. 8.12). Addition of an excess of Fe(II) did not appear to rescue Cu(II) mediated loss of activity. Because Cu(I) did not appear to have a significant effect on 2OG turnover by tPHD2 at the concentrations tested, further experiments on tPHD2apo-SNO were performed in the presence of Cu(I).

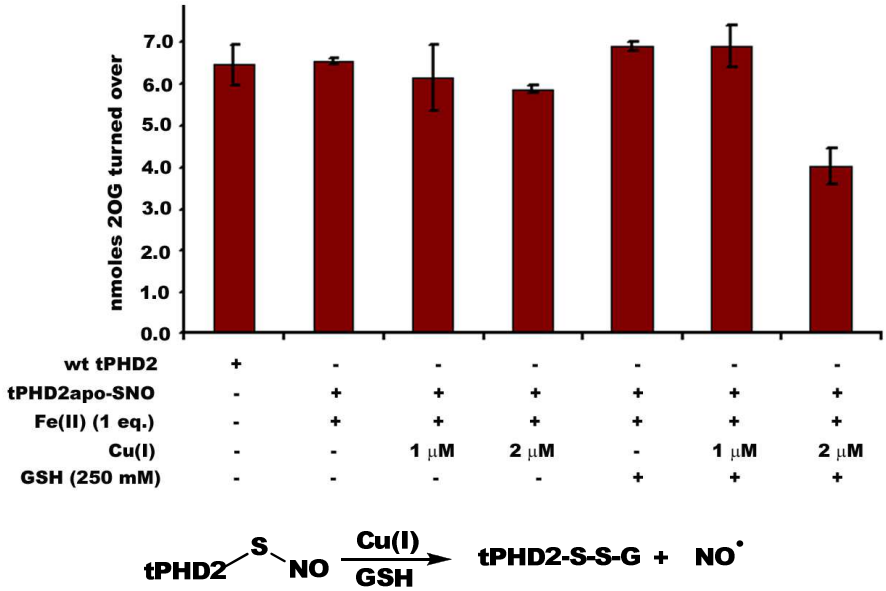


Fig. 8.13 2OG turnover by tPHD2apo-SNO (4μM) after addition of Fe(II), Cu(I) and glutathione (GSH).

However, treatment of an S-nitrosylated tPHD2.Fe(II) complex (prepared by anaerobic treatment of a purified apo-tPHD2-SNO with Fe(II)) with 0.5 equivalent of Cu(I) ions and an excess of reduced glutathione demonstrated ~40% inhibition (Fig. 8.13). The reaction was carried out under conditions aimed at minimising the amount of Cu(II) ions (1000 and 250 fold excesses of ascorbate and DTT, respectively, relative to tPHD2). No inhibition was observed in controls with Cu(I)/reduced glutathione and in the absence of copper ions glutathione slightly stimulated PHD activity (Fig. 8.13). Thus, it is reasonable to propose that inhibition arises by release of NO from R-SNO species.

8.7 Variation of conserved cysteines has no significant effect on PHD2 activity

There are seven cysteine residues in tPHD2, none of which appear to be directly involved in substrate binding in the tPHD2.CODD structure (Fig. 8.14). MS analyses provide evidence for S-nitrosylation of Cys201, Cys208, Cys302 and Cys323/Cys326 to variable extents. In addition, in all known tPHD2 structures, including with bound substrate, one of the alternate conformations of the side chains of Cys201 and Cys208 was located close enough to form a disulfide bond (Fig. 8.14). To further investigate the possible effect of cysteinyl-modification on PHD2 catalysis, variants of all the cysteinyl residues (C201A, C208A, C266A, C283A, C302A, C323A and C326A) were prepared.

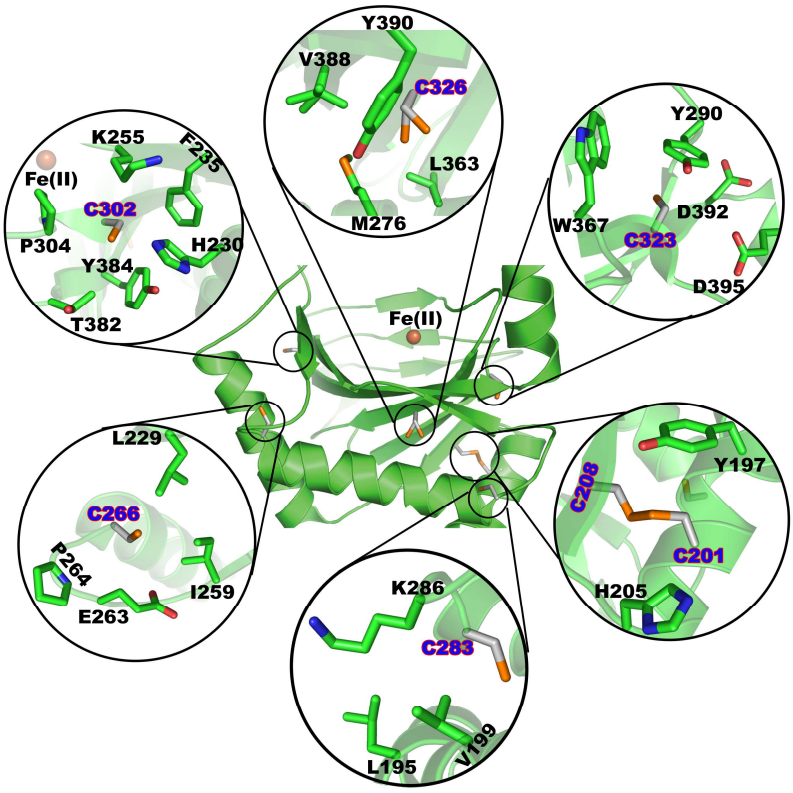


Fig. 8.14 Cysteine residues in tPHD2 including those involving S-nitrosylation.

To determine whether variations of the cysteinyl residues caused any conformational change, circular dichroism (CD) was employed to monitor the secondary structural contents. The individual spectra of all Cys-variants were typical of wildtype tPHD2 with combination of α-helical and β-sheet/β-turn content consistent with the available crystal structures of tPHD2 (Fig. 8.15).

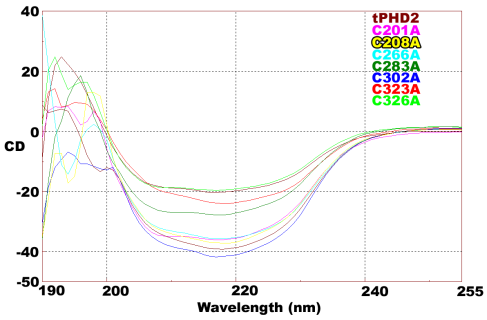


Fig. 8.15 CD spectra of wt tPHD2 and the cysteine variants.

Activity analyses monitoring 2OG decarboxylation and substrate hydroxylation showed that most of the variants appear to possess similar activity to that of tPHD2 within the range of experimental error, although slightly higher rates of 2OG turnover were observed for the C302A and C326A variants (Fig. 8.16). Hydroxylation assays revealed only small differences in activity between the Cys variants and wildtype tPHD2 including in NODD/CODD competition assays where both NODD (HIF-1$\alpha_{395-413}$) and CODD (HIF-1$\alpha_{556-574}$) peptides were simultaneously incubated with the enzyme (Fig. 8.16).

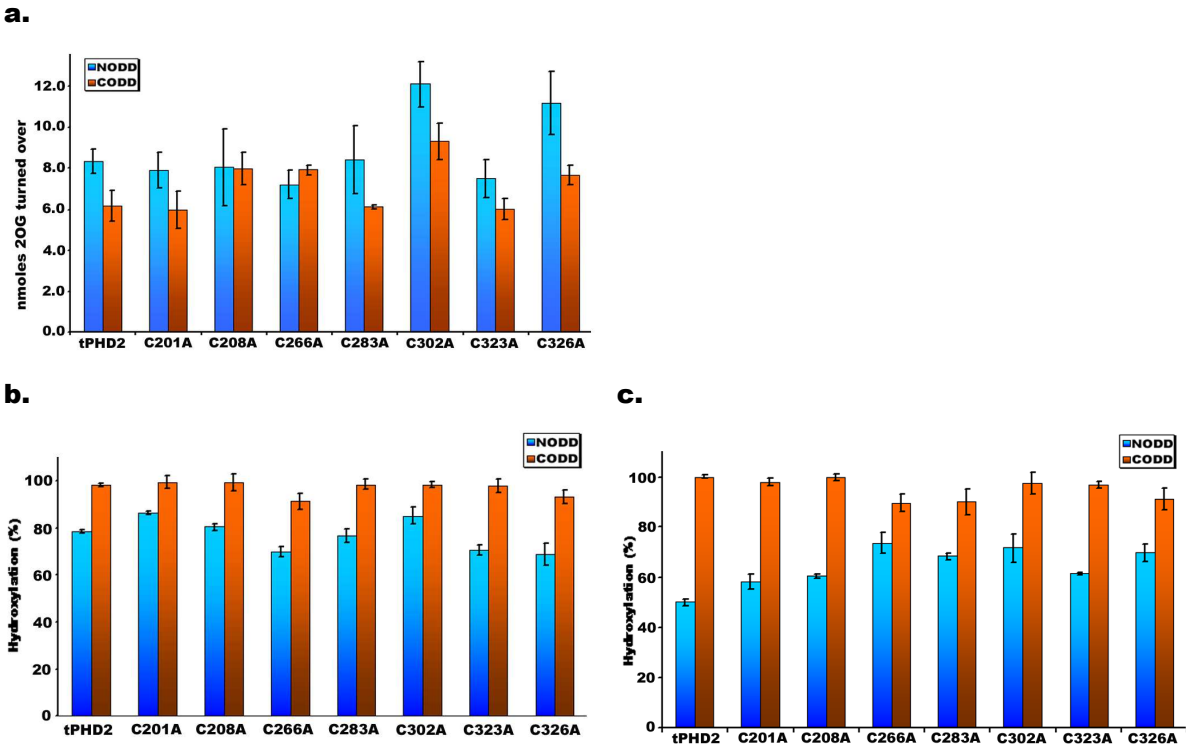


Fig. 8.16 Activity analyses of tPHD2 cys variants (a) 2OG turnover (b) individual substrate hydroxylation and (c) competition experiments were both NODD and CODD were incubated with the wt and variants.

The cysteine variants were then subjected to S-nitrosylation with 250 equivalents of GSNO, as described in Section 8.6.5 for wt tPHD2. The results showed that 2OG turnover activity dropped significantly with both the wt and the cysteine variants upon treatment with GSNO (Fig. 8.17).

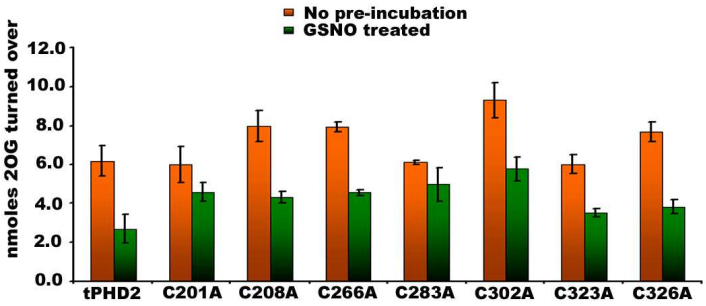


Fig. 8.17 2OG turnover activity of wt and tPHD2 variants before and after addition of 250 equivalents of GSNO. CODD (19-mer) peptide was used as substrate.

8.8 Effect of Cu(I) and GSH on the cysteine variants

Additional experiments were then performed to investigate whether the nitrosylated Cys variants show any difference in activity in the presence of Cu(I) and GSH. As shown before, the activity of tPHD2apo-SNO could be reduced with Cu(I) and GSH. The apo forms of all cysteine variants were produced and treated with 250 equivalents of GSNO. Activity analyses showed that except for C283A and C323A, all other variants displayed some reduced activity after incubating the apo variant-SNO adduct with Cu(I) and GSH (Fig. 8.18). Consistent with the MS data (section 8.5), Cys-201, Cys-208, Cys-266, Cys-302 and Cys-326 appeared to be the key residues that reacted with NO and underwent through a transnitrosylation process with Cu(I) and GSH. The results also indicated that the loss of PHD activity upon treatment with Cu(I) and GSH could result from a concerted 'release of NO' from all the above residues rather than from a single cysteine.

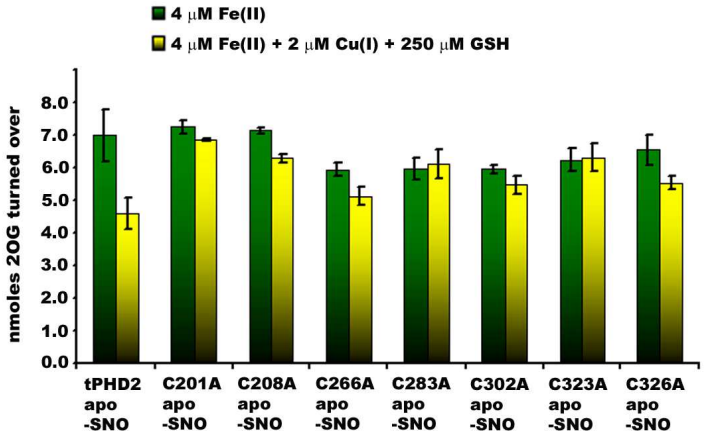


Fig. 8.18 2OG turnover activity of S-nitrosylated wt and tPHD2 variants after addition of Fe(II), Cu(I) and reduced GSH. The CODD (19-mer) peptide was used as a substrate.

8.9 Treatment of HeLa cells with GSNO upregulates HIF

Attempts were made to investigate the role of S-nitrosylation on PHD catalysis in cells. Immunoblot (or Western blot) analyses have been used previously to demonstrate dose- and time-dependent upregulation of normoxic HIF-1α by GSNO in human embryonic kidney (HEK293) cells.[9] In addition, more recent studies show that GSNO can induce expression of PHD2 and PHD3 mRNA under HIF control, reinforcing the existence of a HIF-PHD autoregulatory feedback loop.[46,47]

8.9.1 Design of experiments

The objective of this cell-based work was to investigate whether PHD2 S-nitrosylation took place under physiological conditions after treating cells with GSNO. GSNO has been used extensively as an S-nitrosylating agent in intact cells. To investigate endogenous S-nitrosylation of PHD2, Dr M.E. Cockman (CCMP, Oxford) produced a stable HeLa cell line that could overexpress FLAG-tagged PHD2. In this study, the cells were divided into six treatment groups where three groups were transfected with Flag-tagged PHD2 and the remaining three were non-transfected. Then two groups of cells (transfected/ non-transfected) were treated for an hour with GSNO (1 mM) or L-cysteine (1 mM) as control.

8.9.2 Immunoblot analysis of cell inputs

Before analyses of protein S-nitrosylation in cells, the effects of GSNO on endogenous levels of PHD2 and normoxic HIF-1α were studied. Consistent with previous findings,[9] the immunoblot analyses showed that GSNO (1 mM) stabilised normoxic HIF-1α both in the PHD2 transfected and non-transfected cells (Fig. 8.19). Endogenous levels of HIF-1α were detectable even under normoxic conditions in all non-transfected groups; the corresponding PHD2 transfected cells showed lower levels of HIF-1α (Fig. 8.19).

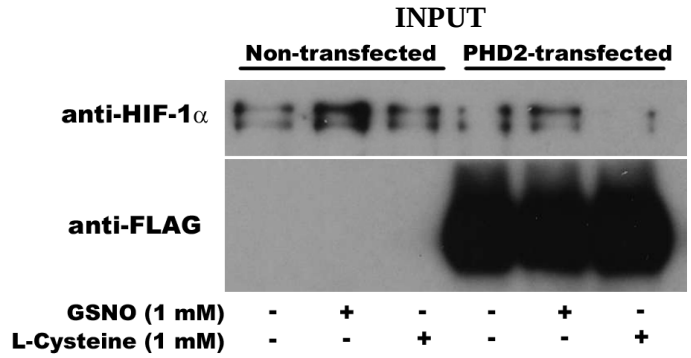


Fig. 8.19 HIF-1α accumulation after GSNO treatment. HeLa cells were stably transfected with functional FLAG-PHD2. Cells in the treatment groups were either treated with 1 mM GSNO or 1 mM L-

cysteine for an hour before lysis for immunoblot analysis. Immunoblots for FLAG-tag and HIF-1α indicated the presence of PHD2 and HIF-1α, respectively. Consistent with the previously published results, HIF-1α protein levels were increased significantly by GSNO treatment, although the protein level was apparently lower in cells transfected with FLAG-PHD2 compared to non-transfected.

8.9.3 Immunoprecipitation assays

Immunoprecipitation (IP) of transfected PHD2 using anti-FLAG coated sepharose beads was then carried out to amplify PHD2 signals in the treatment group.

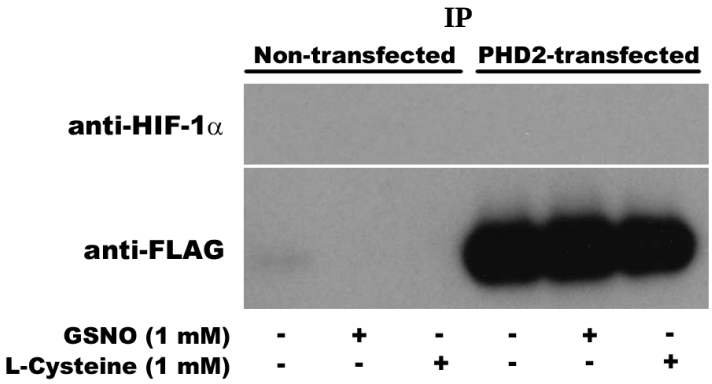


Fig. 8.20 Immunoprecipitation (IP) of PHD2 followed by FLAG and HIF-1α immunoblot analyses. The results show that cells produce sufficient PHD2 for analyses of S-nitrosylation.

The overexpression of transfected PHD2 was detected in all groups (GSNO, L-cysteine and untreated) using anti-PHD2 as well as anti-FLAG antibodies (Fig. 8.20). The intensities of the bands suggested that there may be sufficient PHD2 immunoprecipitates to enable analysis for S-nitrosylation using anti-SNO antibodies.

8.9.4 Immunoblot analysis using recombinant tPHD2

Recombinant tPHD2 treated with GSNO was used as a control to investigate whether the anti-SNO antibody used in the cell-based studies could discriminate between the S-nitrosylated form and unmodified tPHD2. Prior to immunoblot analysis, GSNO-treated tPHD2 was tested for S-nitrosylation by MS measurements under non-denaturing conditions (Section 8.5).

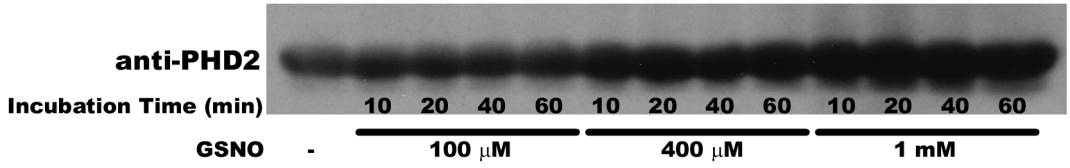


Fig. 8.21 Immunoblot analyses of recombinant S-nitrosylated tPHD2 as opposed to the unmodified wt tPHD2. The anti-SNO antibody could not detect any tPHD2-SNO (as verified by MS) at a recommended ratio (1:500) (blot not shown).

The immunoblot show that the anti-PHD2 was able to detect tPHD2 (Fig. 8.21). Although anti-SNO was unable to bind S-nitrosylated protein samples at a dilution of 1:500 in blocking buffer, the antibody at a higher ratio (1:50) bound non-specifically to tPHD2-SNO as well as the unmodified tPHD2 (data not shown). Thus, the sporadic immunoreactivities of anti-SNO antibody did not allow unambiguous identification of possible S-nitrosylated PHD2 in cells and hence, further experiments could not be carried out to investigate the effects of PHD2 S-nitrosylation on endogenous HIF levels.

8.10 Discussion and future directions

Despite the in vivo evidence that GSNO and other NO donors stabilise normoxic HIF levels,[1] little or no information is known about the details of the underlying molecular mechanisms involved. The current study demonstrates that NO can react with tPHD2 both by binding to the active site Fe(II), as shown by EPR analyses and/or via S-nitrosylation of cysteine residues as shown via MS and crystallographic studies (Fig. 8.22). Accurate quantification of the precise in vivo NO concentrations, especially at subcellular levels, is challenging due to heterogenous tissue distribution, limitations in the microelectrodes currently used for detection and the physicochemical nature of NO itself. Electrochemical detection methods suggest that NO concentrations may range from subnanomolar to micromolar values; however other studies, including modelling analyses, indicate NO concentrations are in the picomolar to nanomolar range, with probable increased levels in disease states.[31] Endogenous concentrations of PHD2 (the most abundant of the PHDs) are estimated to be approximately 4 nM in hypoxic MDA-MB-435 cells[49] and because both PHD2 and PHD3 are themselves induced by hypoxia,[50,51] the normoxic levels of PHDs may be even significantly lower. Given the existing cellular data on the effect of NO on HIF and the ratio of NO to PHD2 where NO mediated inhibition was observed in the current study, it seems reasonable that at least in some pathophysiological conditions, the complex interactions between NO and endogenous PHD2 may have physiological relevance.

There is good precedent for the reaction of NO with non-haem Fe(II) centres.[2,3] NO-mediated PHD2 inhibition therefore likely arises from reaction of NO at the Fe(II) centre in the absence of 2OG (Fig. 8.22a), either resulting in blocked O_2 binding, blocked 2OG binding or oxidation to the catalytically inert Fe(III) form. The lack of observed binding of NO to the Fe(II) of PHD2 in the presence of 2OG is notable (Fig. 8.22a); it contrasts with

the 2OG oxygenase clavaminate synthase,[32] and may reflect the apparently unusual stability of the PHD2.Fe(II).2OG complex[13] (note that NO can possibly still react with cysteinyl thiols in this complex). The tPHD2.Fe(II).2OG complex may form a substantial fraction of the total PHD2 concentration present in cells, with consequent limitations on NO binding to the active site iron. Inhibition of the PHDs by NO may be of more relevance when 2OG is limiting. Thus, the lack of binding of NO to the metal centre in tPHD2SNO.Fe(II)/Zn(II).B crystal structure is likely due to the presence of the iron binding inhibitor required for crystallisation. In order to bind to a tPHD2.inhibitor complex NO, like O_2, may have to replace a water molecule that is tightly bound to the active site Fe(II) and is positioned to hydrogen-bond with Asp315.

In a second mechanism of reaction with NO donors, and under appropriate conditions possibly NO itself, PHD2 (in apo form and as Fe(II) or Fe(II).2OG complexes) can undergo reaction on its cysteinyl thiols (Fig. 8.22a). Notably, the EPR analyses imply that NO itself but not GSNO reacts with the Fe(II) in PHD2 at least under the conditions tested. Further, 'release' of NO from S-nitrosylated tPHD2 can result in inhibition via reaction at the active site Fe(II) (Fig. 8.22a). Thus, the manner in which NO is delivered to the PHDs should be taken into account in cellular and physiological studies on the effect of NO/ NO donors on the HIF system.

Substitution studies on the seven cysteinyl residues present in the catalytic domain of PHD2 reveal that individual modifications at these residues have little effect under standard incubation conditions. However, it cannot be ruled out that nitrosylation at one or more sites affect catalytic activity directly or affect other aspects of PHD biochemistry, e.g. protein-protein interaction. The MS results suggest S-nitrosylation of at least four tPHD2 residues (Cys201, Cys208, Cys302 and Cys323/Cys326). It should also be noted that full length PHD2 has a conserved zinc finger the cysteinyl-residues of which have the potential to react with NO.

In contrast to the solution analyses, the crystallisation studies provided evidence for only one such modification at Cys302. In a recent study on the nucleophilicity of the tPHD2 cysteinyl thiols, Cys201 was found to be predominantly modified by thiols or N-ethylmaleimide (J Mecinovic, unpublished data). The apparently selective modification of Cys302 in the crystal structure may reflect kinetic (accessibility) and/or thermodynamic

(stability) factors for the various cysteine residues. The stability of the adduct is likely to be a factor, as proposed for related studies on thioredoxin-1[20] and protein-tyrosine phosphatase 1B.[33] Cys302 is conserved in all three human PHDs and is located in a conserved hydrophobic pocket formed by the hydrophobic side chains of Lys255, Pro304, Thr382, Tyr384, His230 and Phe235 (Fig. 8.3). The buried nature of Cys302 may result in a stable adduct formation.

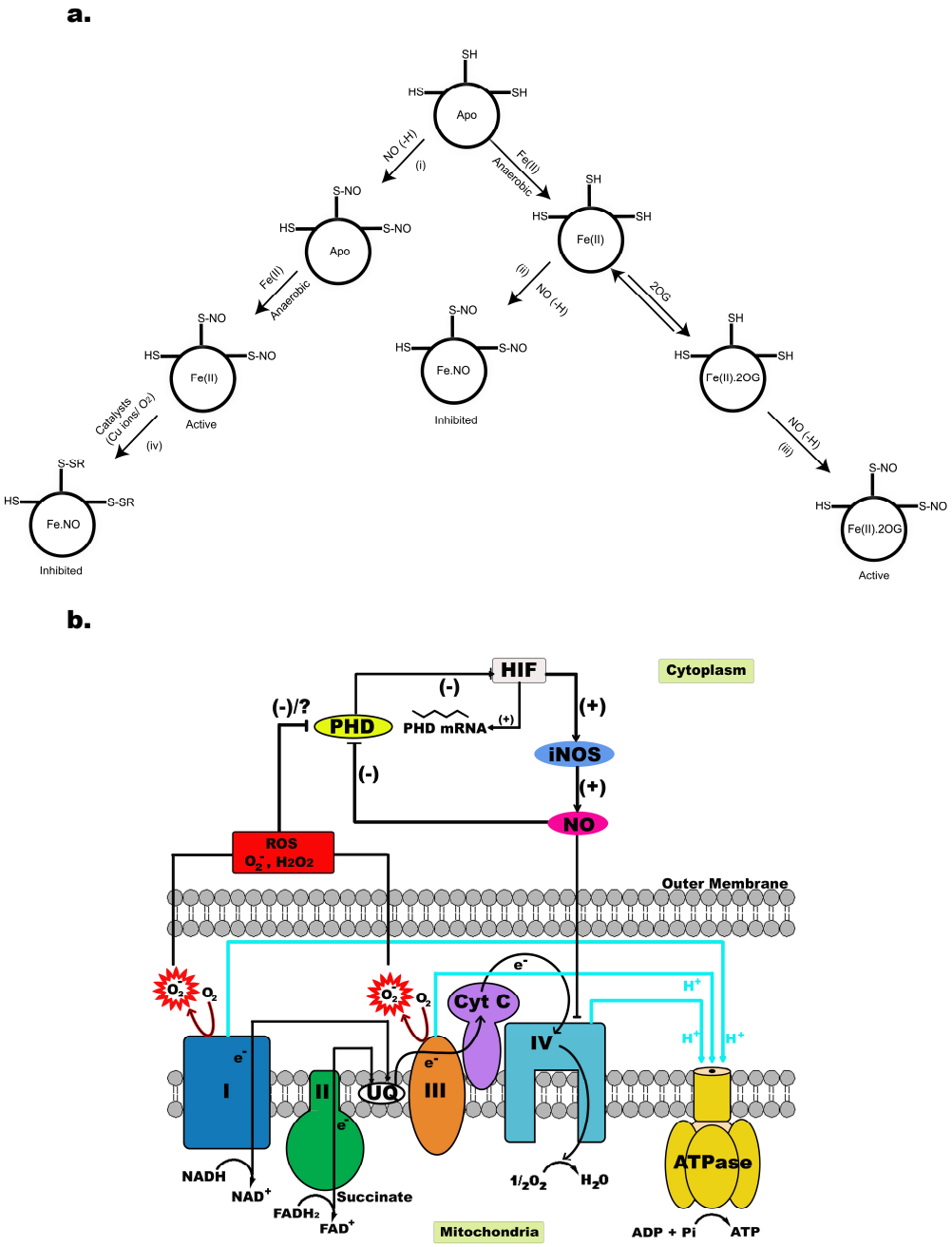


Fig. 8.22 Role of NO in hypoxic system. (a) The potentially complex reaction of NO with PHD2. The figure summarises some of the likely types of reaction of NO with apo-PHD2 (i) or the Fe(II) (ii)/ Fe(II).2OG (iii) complexes. Note that the reactions have the potential to be more complex because there are seven cysteinyl residues in tPHD2 which likely react differently and transnitrosylation / mixed disulfide formation occurs including when using NO transfer reagents such as GSNO in the presence of catalysts (e.g. O_2, copper ions). S-Nitrosylated cysteines may also undergo redox mediated release of NO that may inhibit PHD activity via Fe(II) binding (iv). Mixed disulfides are unlikely to, at least potently,

cause PHD2 inhibition. (b) Model for the targets of NO and associated oxides in hypoxic system and mitochondrial respiration. Whereas S-nitrosylation of Cys800 of HIF-1α can stimulate the transcriptional activity of HIF-1 via increased binding to CREB/p300[34], S-nitrosylation of PHD2 may serve as reserve that may increase local NO level via de-nitrosylation, depending on factors including O_2 and free copper concentration. NO released in this way or produced enzymatically via arginine metabolism can directly inhibit PHDs by coordinating Fe(II). Alternatively, NO has been reported to produce reactive oxygen species (ROS, such as O_2^- and H_2O_2) via inhibition of a mitochondrial respiratory centre, cytochrome c oxidase.[35] Increased ROS production can alter Fenton chemistry leading to increased Fe(III).[36] The role of ROS in modulating the protein amount of HIF-1α is controversially discussed.[1]

Further evidence for the importance of reagents/ conditions on the effects of NO/NO donors on PHD2, came from studies on S-nitrosylation. Under the standard conditions, the NO-donor, GSNO only reduces tPHD2 activity with an aerobic pre-incubation period when using tPHD2-Fe(II), but not apo-tPHD2. Because O_2 can accelerate the release of NO from nitrosothiols,[11,12] it is possible that NO gas may be produced from nitrosylated cysteine residues under aerobic conditions, and that the released NO inhibits PHD2 via binding at the Fe(II) centre (Fig. 8.22a). Copper ions are reported to catalyse release of NO from RSNO species in a process accelerated by reducing agents.[41,42] Further support for the potential of inhibition by 'NO release' comes from the observation of a clear loss of tPHD2 activity after pre-incubation of an S-nitrosylated tPHD2.Fe(II) complex (prepared by anaerobic treatment of a purified apo-tPHD2-SNO with Fe(II)) with Cu(I) and reduced glutathione (Fig. 8.22). The potential role of reduced glutathione in this process is notable because intracellular levels of glutathione are high[37] and therefore, may have physiological relevance.

Overall, the results demonstrate the potential for multiple interactions between NO/ NO donors and the HIF hydroxylases at both active site Fe(II) and cysteinyl residues (Fig. 8.22). This dual reactivity coupled to the seven cysteinyl residues in the catalytic domain (and others in the N-terminal Zn(II)-binding domain) of PHD2 suggests that the PHD activity may be fine-tuned by reaction with NO and/or small molecules/ ions via multiple reactions (Fig. 8.22). However, the available evidence suggests that the most direct mechanism for modulation of hydroxylase activity by NO is via competition with oxygen for binding to the active site Fe(II) (Fig. 8.22a). Although NO is reported to occur in the nanomolar range and below in cells, and the PHDs are also reported to occur at nanomolar levels, raised NO levels in disease states mean it is possible that this mechanism has physiological relevance by inhibiting PHD2, and possibly other human 2OG oxygenases under pathophysiological conditions.

8.11 References

1. Brune, B. & Zhou, J. Nitric oxide and superoxide: interference with hypoxic signaling. *Cardiovasc Res* **75**, 275-82 (2007).
2. Nathan, C. & Xie, Q.W. Nitric oxide synthases: roles, tolls, and controls. *Cell* **78**, 915-8 (1994).
3. Al-Sa'doni, H. & Ferro, A. S-Nitrosothiols: a class of nitric oxide-donor drugs. *Clin Sci (Lond)* **98**, 507-20 (2000).
4. Stamler, J.S. Redox signaling: nitrosylation and related target interactions of nitric oxide. *Cell* **78**, 931-6 (1994).
5. Stamler, J.S., Simon, D.I., Osborne, J.A., Mullins, M.E., Jaraki, O., Michel, T., Singel, D.J. & Loscalzo, J. S-nitrosylation of proteins with nitric oxide: synthesis and characterization of biologically active compounds. *Proc Natl Acad Sci U S A* **89**, 444-8 (1992).
6. Hess, D.T., Matsumoto, A., Kim, S.O., Marshall, H.E. & Stamler, J.S. Protein S-nitrosylation: purview and parameters. *Nat Rev Mol Cell Biol* **6**, 150-66 (2005).
7. Singh, S.P., Wishnok, J.S., Keshive, M., Deen, W.M. & Tannenbaum, S.R. The chemistry of the S-nitrosoglutathione/glutathione system. *Proc Natl Acad Sci U S A* **93**, 14428-33 (1996).
8. Girard, P. & Potier, P. NO, thiols and disulfides. *FEBS Lett* **320**, 7-8 (1993).
9. Metzen, E., Zhou, J., Jelkmann, W., Fandrey, J. & Brune, B. Nitric oxide impairs normoxic degradation of HIF-1alpha by inhibition of prolyl hydroxylases. *Mol Biol Cell* **14**, 3470-81 (2003).
10. Peyssonnaux, C., Datta, V., Cramer, T., Doedens, A., Theodorakis, E.A., Gallo, R.L., Hurtado-Ziola, N., Nizet, V. & Johnson, R.S. HIF-1alpha expression regulates the bactericidal capacity of phagocytes. *J Clin Invest* **115**, 1806-15 (2005).
11. Hagen, T., Taylor, C.T., Lam, F. & Moncada, S. Redistribution of intracellular oxygen in hypoxia by nitric oxide: effect on HIF1α. *Science* **302**, 1975-8 (2003).
12. Huang, L.E., Willmore, W.G., Gu, J., Goldberg, M.A. & Bunn, H.F. Inhibition of hypoxia-inducible factor 1 activation by carbon monoxide and nitric oxide. Implications for oxygen sensing and signaling. *J. Biol. Chem.* **274**, 9038-44 (1999).
13. McNeill, L.A., Flashman, E., Buck, M.R., Hewitson, K.S., Clifton, I.J., Jeschke, G., Claridge, T.D., Ehrismann, D., Oldham, N.J. & Schofield, C.J. Hypoxia-inducible factor prolyl hydroxylase 2 has a high affinity for ferrous iron and 2-oxoglutarate. *Mol Biosyst* **1**, 321-4 (2005).
14. Schofield, C.J. & Ratcliffe, P.J. Oxygen sensing by HIF hydroxylases. *Nat Rev Mol Cell Biol* **5**, 343-54 (2004).
15. Schreiter, E.R., Rodriguez, M.M., Weichsel, A., Montfort, W.R. & Bonaventura, J. S-nitrosylation-induced conformational change in blackfin tuna myoglobin. *J Biol Chem* **282**, 19773-80 (2007).
16. Chan, N.L., Kavanaugh, J.S., Rogers, P.H. & Arnone, A. Crystallographic analysis of the interaction of nitric oxide with quaternary-T human hemoglobin. *Biochemistry* **43**, 118-32 (2004).
17. Zhao, Y.L. & Houk, K.N. Thionitroxides, RSNHO*: the structure of the SNO moiety in "S-Nitrosohemoglobin", a possible NO reservoir and transporter. *J. Am. Chem. Soc.* **128**, 1422-3 (2006).
18. Weichsel, A., Maes, E.M., Andersen, J.F., Valenzuela, J.G., Shokhireva, T., Walker, F.A. & Montfort, W.R. Heme-assisted S-nitrosation of a proximal thiolate in a nitric oxide transport protein. *Proc Natl Acad Sci U S A* **102**, 594-9 (2005).
19. Chan, N.L., Rogers, P.H. & Arnone, A. Crystal structure of the S-nitroso form of liganded human hemoglobin. *Biochemistry* **37**, 16459-64 (1998).

20. Weichsel, A., Brailey, J.L. & Montfort, W.R. Buried S-nitrosocysteine revealed in crystal structures of human thioredoxin. *Biochemistry* **46**, 1219-27 (2007).
21. Lee, B. & Richards, F.M. The interpretation of protein structures: estimation of static accessibility. *J Mol Biol* **55**, 379-400 (1971).
22. Hubbard, S.J. & Thornton, J.M. NACCESS, computer program. *Department of Biochemistry and Molecular Biology, University College, London* (1993).
23. McDonough, M.A., Li, V., Flashman, E., Chowdhury, R., Mohr, C., Lienard, B.M., Zondlo, J., Oldham, N.J., Clifton, I.J., Lewis, J., McNeill, L.A., Kurzeja, R.J., Hewitson, K.S., Yang, E., Jordan, S., Syed, R.S. & Schofield, C.J. Cellular oxygen sensing: Crystal structure of hypoxia-inducible factor prolyl hydroxylase (PHD2). *Proc Natl Acad Sci U S A* **103**, 9814-9 (2006).
24. Stamler, J.S. & Toone, E.J. The decomposition of thionitrites. *Curr Opin Chem Biol* **6**, 779-85 (2002).
25. Welford, R.W., Schlemminger, I., McNeill, L.A., Hewitson, K.S. & Schofield, C.J. The selectivity and inhibition of AlkB. *J Biol Chem* **278**, 10157-61 (2003).
26. Webb, J.D. Studies on the selectivity of oxygen-sensing hydroxylases. *University of Oxford, Part II Thesis* (2004).
27. Burg, A., Cohen, H. & Meyerstein, D. The reaction mechanism of nitrosothiols with copper(I). *J Biol Inorg Chem* **5**, 213-7 (2000).
28. Noble, D.R. & Williams, D.L. Structure-reactivity studies of the Cu(2+)-catalyzed decomposition of four S-nitrosothiols based around the S-Nitrosocysteine/S-nitrosoglutathione structures. *Nitric Oxide* **4**, 392-8 (2000).
29. McAninly, J., Williams, D.L.H., Askew, S.C., Butler, A.R. & Russell, C. Metal ion catalysis in nitrosothiol (RSNO) decomposition. *J. Chem. Soc., Chem. Commun.*, 1758-1759 (1993).
30. Mecinovic, J., Chowdhury, R., Lienard, B.M., Flashman, E., Buck, M.R., Oldham, N.J. & Schofield, C.J. ESI-MS studies on prolyl hydroxylase domain 2 reveal a new metal binding site. *ChemMedChem* **3**, 569-72 (2008).
31. Thomas, D.D., Ridnour, L.A., Isenberg, J.S., Flores-Santana, W., Switzer, C.H., Donzelli, S., Hussain, P., Vecoli, C., Paolocci, N., Ambs, S., Colton, C.A., Harris, C.C., Roberts, D.D. & Wink, D.A. The chemical biology of nitric oxide: implications in cellular signaling. *Free Radic Biol Med* **45**, 18-31 (2008).
32. Zhang, Z., Ren, J., Harlos, K., McKinnon, C.H., Clifton, I.J. & Schofield, C.J. Crystal structure of a clavaminate synthase-Fe(II)-2-oxoglutarate-substrate-NO complex: evidence for metal centered rearrangements. *FEBS Lett* **517**, 7-12 (2002).
33. Chen, Y.Y., Chu, H.M., Pan, K.T., Teng, C.H., Wang, D.L., Wang, A.H., Khoo, K.H. & Meng, T.C. Cysteine S-nitrosylation protects protein-tyrosine phosphatase 1B against oxidation-induced permanent inactivation. *J. Biol. Chem.* **283**, 35265-72 (2008).
34. Yasinska, I.M. & Sumbayev, V.V. S-nitrosation of Cys-800 of HIF-1alpha protein activates its interaction with p300 and stimulates its transcriptional activity. *FEBS Lett* **549**, 105-9 (2003).
35. Palacios-Callender, M., Quintero, M., Hollis, V.S., Springett, R.J. & Moncada, S. Endogenous NO regulates superoxide production at low oxygen concentrations by modifying the redox state of cytochrome c oxidase. *Proc Natl Acad Sci U S A* **101**, 7630-5 (2004).
36. Cash, T.P., Pan, Y. & Simon, M.C. Reactive oxygen species and cellular oxygen sensing. *Free Radic Biol Med* **43**, 1219-25 (2007).
37. Meister, A. & Anderson, M.E. Glutathione. *Annu Rev Biochem* **52**, 711-60 (1983).

Chapter 9 – Cloning and characterisation of human PHD3

This chapter describes attempts to improve the heterologous expression of human PHD3 in *E. coli* for crystallographic and biochemical studies. Different approaches were followed, including cloning different constructs, optimising expression/lysis conditions and co-expression with HIF-1α fragments and molecular chaperones. Despite significant improvement in the soluble expression level, the activity results show that the wt and variants of PHD3 were significantly less active than wt tPHD2.

9.1 Difficulty in expression of PHD3

The expression of recombinant human PHD3 in bacterial cells has proved difficult since its discovery in 2001.[1] Several groups have reported that the activity of PHD3 is reduced upon concentration, probably due to aggregation.[2] Indeed, PHD3 was recently reported to form subcellular aggregates when overexpressed in HeLa cells under normoxia.[3] In the current study, an N-terminally His_6-tagged PHD3 (PHD3-NH hereafter) was found to be only <0.5% soluble in bacterial cells with the majority of expressed proteins remaining as inclusion bodies.

9.2 Optimisation of lysis and expression conditions

Initial attempts were made to produce an optimal amount of soluble PHD3 protein for subsequent purification by using a sparse matrix approach that allows a rapid sampling of a large 'solubility space' (Appendix 9),[4] by optimising expression conditions (induction temperature, IPTG concentration and incubation time) or by expressing the constructs in different *E. coli* cell lines including the 'rare codon' enriched Rosetta DE3 strain. However the expression levels of both wt PHD3 and a C-terminally truncated variant (Section 9.3.1) appeared to be unaltered in most of the lysis buffer conditions although the protein expression levels seemed to slightly increase using *E. coli* BL21 DE3 strains (data not shown). The relative expression levels were assessed by SDS-PAGE.

Previously it was observed in the cases of other proteins that heterologous gene expression in *E. coli* can result in insoluble/nonfunctional target protein because of the lack of adequate cofactors. Therefore, variable concentrations of iron and/or 2OG were added during induction. Whilst the addition of 2OG made no significant difference, the addition

of 1-5 µM iron did appear to slightly improve the protein expression. Further efforts were then focused on cloning optimised expression construct.

9.3 Construct design

Sequence analyses of PHD2 and PHD3 predicted some disordered regions in the proteins (Fig. 9.1). Several programmes were used that predicted structural disorder within regions of PHD2 that were agreeable with the tPHD2 crystal structure. Of note, the N-terminal domain of PHD2, except for residues 21-58 of the Zn-MYND finger, was predicted to be unstructured. Apart from that, B-factor analyses of all tPHD2 structures including with substrate revealed two major flexible regions: a β2/β3 loop (residues 237-254) and the C-terminus (residues 400-403). C-Terminal residues 404-426 were not modelled in structures because no electron density could be observed. The results of disorder prediction supported the structural finding; residues 185-406 (excluding β2/β3 loop residues 237-254) were predicted to be more structured than the rest, which validated the believability of such predictions. Analyses of the PHD3 sequence suggested that residues after $Arg218_{PHD3}$ were likely to be unstructured.

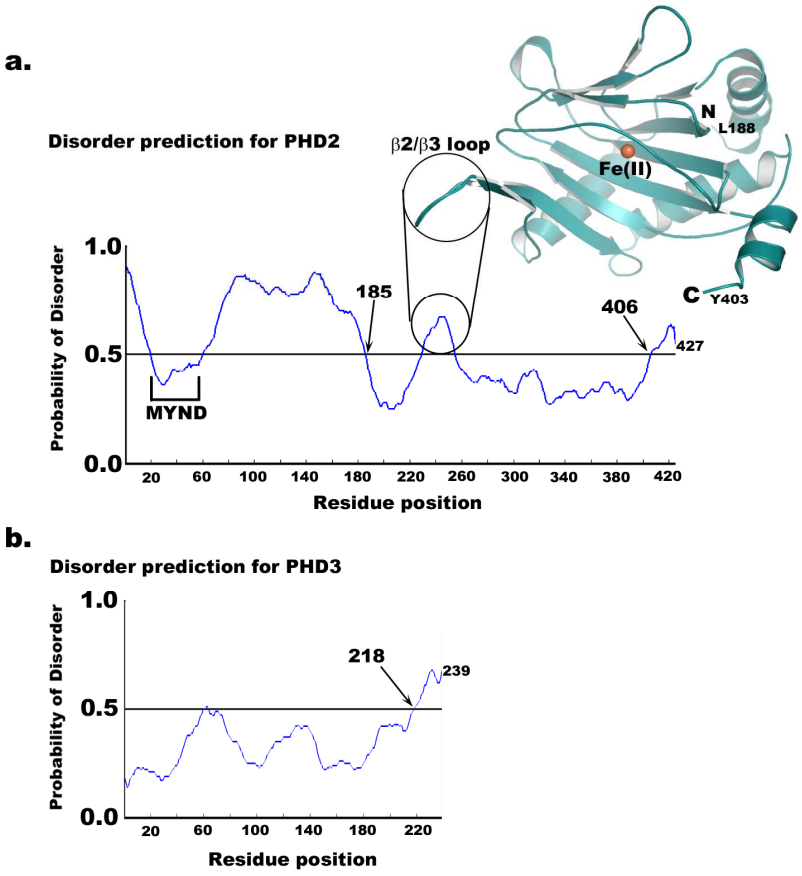


Fig. 9.1 Prediction of disordered regions in PHD2 (a) and PHD3 (b). The tPHD2 structure supports the disorder predictions.

9.3.1 Truncation and deletion variants

Attempts were made to improve soluble expression of PHD3, notably by constructing a series of C- and N-terminal truncations (Appendix 10), and by deleting an 'extra' residue in the predicted β2/β3 loop in PHD3 (ΔR71, ΔH72) (Chapter 4, Fig. 4.1). However, these variants did not appear to improve protein solubility.

9.3.2 C-Terminal His_6-tagged PHD3

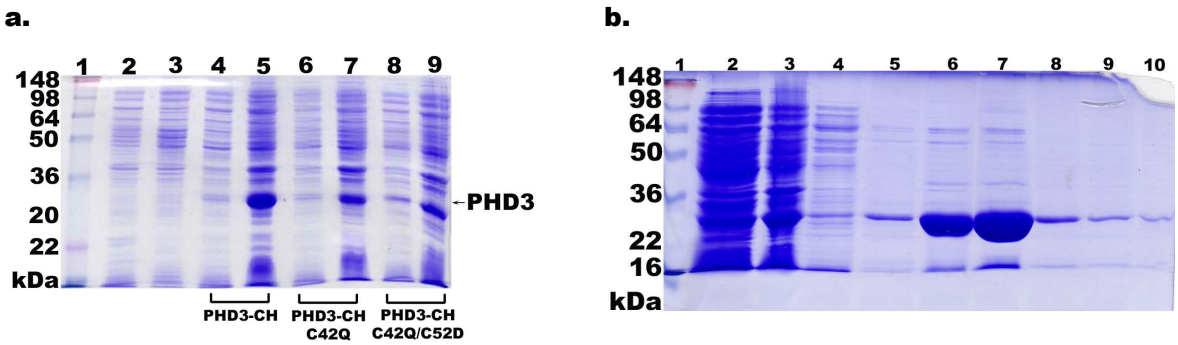
A C-terminal His_6-tagged *phd3* construct was cloned into the pET-29b(+) vector with a thrombin cleavage site (LVPRGS) engineered at the C-terminus of the PHD3 sequence. The C-terminally His_6-tagged PHD3 (PHD3-CH hereafter) was found to be more soluble (~1%) in *E. coli* BL21 (DE3) than PHD3-NH (Appendix 10). However, the soluble protein had a tendency to heavily precipitate when concentrated to more than 1.5 $mg.ml^{-1}$.

9.3.3 Surface cysteine variants

Modelling studies of PHD3 using tPHD2 crystal structure as a template suggested that the presence of two cysteine residues ($Cys42_{PHD3}$ and $Cys52_{PHD3}$) located on the predicted protein surface could cause aggregation. In the analogous position to these two cysteines, PHD2 has $Gln221_{PHD2}$ and $Asp231_{PHD2}$. Two variants of PHD3 were made (C42Q and C42Q/C52D) whereby the cysteines were replaced by the PHD2 equivalent residues. Soluble expression levels of C-terminally His_6-tagged PHD3 C42Q and C42Q/C52D double mutant improved to ~3-4% when grown at ~22°C for 5-6 hours, with induction by 0.5 mM IPTG (Fig. 9.3a).

9.3.4 PHD3 purification and optimisation of buffer for protein storage

PHD3-NH C42Q, PHD3-CH, PHD3-CH C42Q and PHD3-CH C42Q/C52D were purified using Ni^{2+} affinity chromatography followed by size exclusion chromatography in 50 mM Tris.HCl pH 7.5, 0.3 M NaCl, 10% glycerol and 0.1% β- mercaptoethanol (Fig. 9.2).



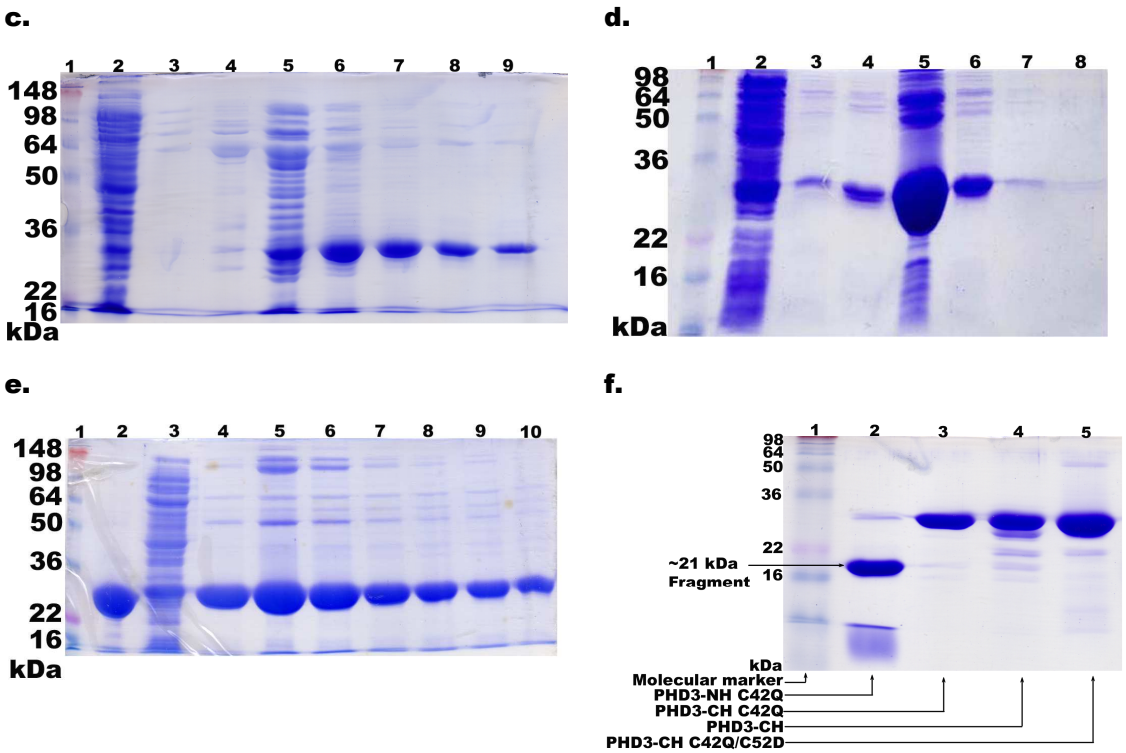


Fig. 9.2 Analysis of expression levels and purification by SDS-PAGE. (a) Expression levels of PHD3-CH, PHD3-CH C42Q and PHD3-CH C42Q/C52D in *E. coli* BL21 (DE3) (0.5 mM IPTG/ 25°C). Lane 1 = molecular marker 2, 3 = soluble and insoluble un-induced cell expression, 4, 5 = soluble and insoluble expression of PHD3-CH, 6, 7 = soluble and insoluble expression of PHD3-CH C42Q, 8, 9 = soluble and insoluble expression of PHD3-CH C42Q/C52D. Figures (b-e) show purification of PHD3-NH C42Q, PHD3-CH, PHD3-CH C42Q and PHD3-CH C42Q/C52D. (b) Coomassie blue stained 12.5% SDS-PAGE gel of PHD3-NH C42Q fractions from nickel affinity column: 1 = molecular weight markers, 2 = load (cell extract), 3 = insoluble cell debris 4 = flow through from column, 5-10 = elution samples following a His-column; (c) coomassie blue stained 12.5% SDS-PAGE gel of PHD3-CH fractions: 1 = molecular weight markers, 2 = load, 3-9 = elution samples from column; (d) SDS-PAGE gel of PHD3-CH C42Q fractions: 1 = molecular weight markers, 2 = load, 3-8 = elution samples from column; (e) SDS-PAGE gel of PHD3-CH C42Q/C52D fractions along with a tPHD2 control: 1 = molecular weight markers, 2 = tPHD2 (control) 3 = load, 4-10 = elution samples from Ni^{2+} affinity column; (f) analysis of protein purity by SDS-PAGE. Lane 2 shows a thrombin cleaved PHD3-NH C42Q product (~21 kDa).

Attempts were made to formulate an optimal buffer system for further activity analysis; with this aim the double cysteine variant PHD3-CH C42Q/C52D was dialysed into buffers with lower salt and/or different buffering compounds (HEPES, MES and phosphate). Addition of 10-20% glycerol or 5-10% PEG 400 generally had a stabilising effect on protein stability (judged by less tendency to precipitate), although different buffers did not appear to make an observable difference. The PHD3-CH C42Q/C52D variant could be concentrated to 4-6 $mg.ml^{-1}$ using a buffer system comprising 50 mM Tris.HCl pH 7.5, 0.2 M NaCl, 10 % glycerol and 2 mM TCEP. If high amounts of reducing agent were included (~10 mM TCEP or ~30 mM DTT or ~40 mM β-mercaptoethanol), then the protein could be concentrated up to 12-14 $mg.ml^{-1}$.

9.3.5 Improvement of internal salt-bridges

The observations that thermophilic proteins contain more salt bridges than their mesophilic counterparts[5-7] and substitution of residues involved in buried salt bridges decreases stability by 3-5 kcal mol^{-1} as measured by pK_a shifts[8,9] suggest protein salt-bridge interactions have a significant role in overall protein stability. Such favourable electrostatic interactions also depend on the ion-pair geometries.[10] Other studies have shown that most globular proteins have evolved a `surface-entropy shield' created mostly by lysine and glutamate residues, which prevents non-specific aggregation and precipitation.[11]

Modelling studies showed that PHD3 lacks at least 3 such salt bridges/ ionic interactions that were observed in the crystal structures of tPHD2, viz., $Asp224_{PHD2}$-$Arg227_{PHD2}$, $Asp278_{PHD2}$-$Arg281_{PHD2}$ and $Asp394_{PHD2}$-$Arg398_{PHD2}$ in the crystal structure of tPHD2. Using PHD3-CH C42Q/C52D as a template, a C42Q/E45D/K48R/C52D/R65K/R100D/ L103R/Y104H/E216D/E220R variant (PHD3-CH MM-3, hereafter) was therefore cloned where the PHD3 residues were replaced by analogous PHD2 residues so that the variant can form additional salt-bridges. Treatment of the PHD3-NH C42Q variant with thrombin yielded an ~21 kDa fragment (Fig. 9.2f) due to the presence of an internal thrombin cleavage site ($Arg65_{PHD3}$). The cleavage site was also modified in PHD3-CH MM-3 by substitution with a lysine, present in PHD2 ($Lys244_{PHD2}$) in the analogous position. The soluble expression level for the PHD3-CH MM-3 variant significantly improved (>5%) because of these modifications (data not shown).

9.4 Activity analyses of PHD3 and variants

PHD3-CH, PHD3-CH C42Q and PHD3-CH C42Q/C52Q variants were found to be only slightly active in both 2OG turnover and hydroxylation assays as compared to tPHD2. A more sensitive TR-FRET assay was employed for activity analyses since PHD3 was reported to aggregate and display diminished activity at higher concentrations. The TR-FRET assay enabled measurement of activity using nanomolar amounts of PHD3. The results indicated that the PHD3-CH could be active at a lower concentration range, although still in a significantly lower level when compared to tPHD2 (Fig. 9.3).

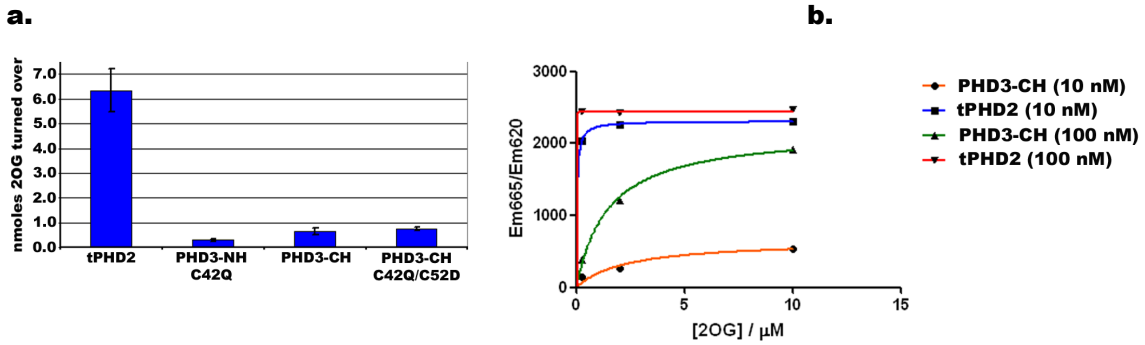


Fig. 9.3 Activity analysis of wt and variants of PHD3. (a) 2OG turnover data (b) TR-FRET assay results analysed by GraphPad Prism 5 software. CODD (19-mer) was used as substrate.

To investigate whether the proteins were correctly folded, CD measurements were carried out. CD spectra of the tested proteins indicated that both the wt PHD3-CH and PHD3-CH C42Q variant could be folded with reasonably similar amounts of secondary structure contents to tPHD2 (Fig. 9.4a). Consistent with the CD measurements, non-denaturing MS experiments revealed that the PHD3-CH C42Q point mutation had the ability to bind a 2-OG analogue inhibitor, compound C[12] (Fig. 9.4b). Although these data were suggestive of correctly folded protein, it was not possible to rationalise why all the PHD3 proteins were found to be essentially inactive in the 2OG turnover/ hydroxylation assays.

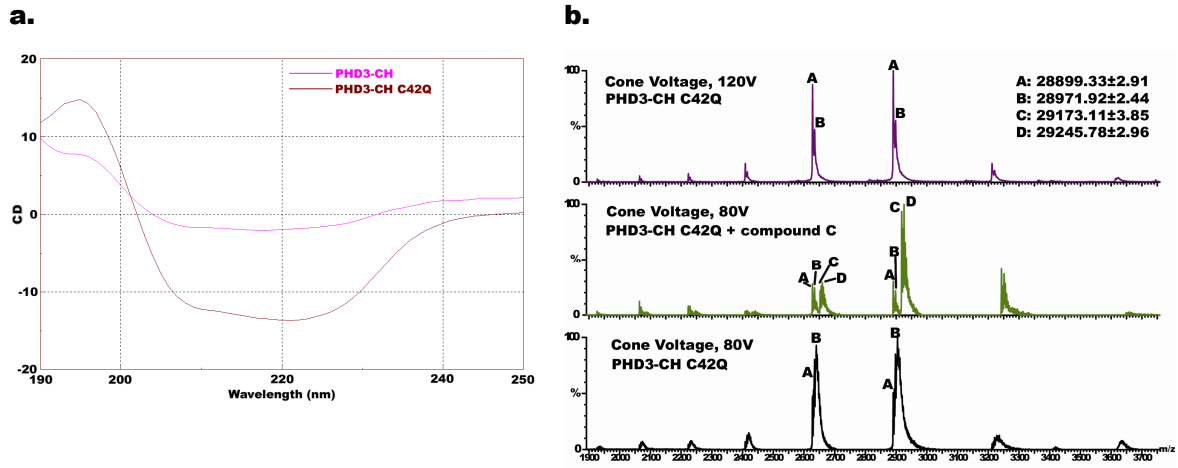


Fig. 9.4 Characterisation of wt and variants of PHD3. (a) CD measurements of PHD3-CH (10 µM) and PHD3-CH C42Q (25 µM) variant revealing similar secondary structure content (for comparison with tPHD2, see Chapter 8, Fig. 8.17) (b) MS analysis of PHD3-CH C42Q variant under non-denaturing condition. Peaks indicate, A = PHD3-CH C42Q, B = PHD3-CH C42Q salt adducts; C = PHD3-CH C42Q.compound C complex and D, corresponding salt adducts of the complex.

9.5 Co-expression trials

In some cases, co-expression of multiple genes in *E. coli*, such as the members of a stable multiprotein complex[13] or a protein with a chaperone[14] has been demonstrated to significantly alleviate solubility problems and/or protein misfolding.

9.5.1 Co-expression with HIF-1α

Attempts were therefore made to co-express PHD3-CH and two variants (C42Q and C42Q/C52D) with different GST-tagged HIF-1α CODD fragments (530-652 and 530-826) in the hope that a stabilising interaction would promote production of correctly folded protein. However, none of the HIF-1α CODD fragments appeared to increase soluble expression of PHD3 (Fig. 9.5).

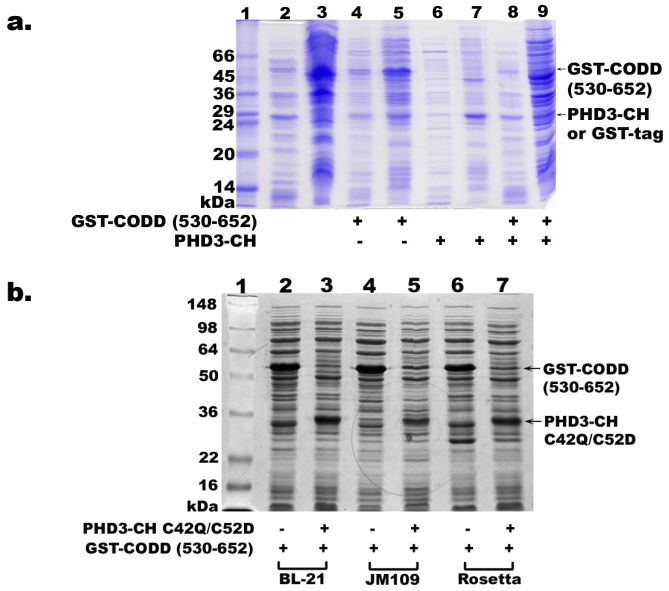


Fig. 9.5 SDS-PAGE analyses of PHD3 levels when co-expressed with GST-CODD.
(a) Co-expression of PHD3-CH with GST-$CODD_{530-652}$ fragment in *E. coli* BL21 DE3. Lane 1 = molecular marker 2, 3 = soluble and insoluble un-induced cell expression, 4, 5 = soluble and insoluble expression of GST-$CODD_{530-652}$, 6, 7 = soluble and insoluble expression of PHD3-CH, 8, 9 = soluble and insoluble expression of PHD3-CH plus GST-$CODD_{530-652}$.
(b) Co-expression of PHD3-CH C42Q/ C52D with GST-$CODD_{530-652}$ Lane 1 = molecular marker, 2 = GST-$CODD_{530-652}$ in BL21 (DE3), 3 = PHD3-CH C42Q/C52D plus GST-$CODD_{530-652}$ in BL21 (DE3), 4 = GST-$CODD_{530-652}$ in JM109 (DE3), 5 = PHD3-CH C42Q/C52D plus GST-$CODD_{530-652}$ in JM109 (DE3), 6 = GST $CODD_{530-652}$ in Rosetta (DE3), 7 = PHD3-CH C42Q/C52D plus GST-$CODD_{530-652}$ in Rosetta (DE3).

9.5.2 Co-expression with chaperones

In a previous study, PHD3 had been reported to co-purify with the cytosolic chaperonin TRiC/CTT from several mammalian cell extracts.[15] The authors reasoned that PHD3 could be a TRiC substrate, providing a regulatory mechanism by which PHD3 might be activated.[15] Because TRiC is a homologue of the bacterial chaperonin GroEL,[16] co-expression with GroEL-GroES system may stimulate PHD3 folding in *E. coli*. The PHD3-CH MM-3 that demonstrated highest expression so far, was co-expressed with chaperones belonging to the DnaK–DnaJ–GrpE (Takara Bio INC., Japan) and GroEL–GroES (Takara

Bio INC., Japan) systems.[14] The expression level of PHD3-CH MM-3 variant as analysed by SDS-PAGE appeared to remain unaltered (data not shown). Studies are ongoing to purify the variant.

9.6 Crystallography

Initial crystallisation attempts using PHD3-CH (1 $mg.ml^{-1}$), PHD3-CH C42Q (1.5 $mg.ml^{-1}$) and PHD3-CH C42Q/C52D (5 $mg.ml^{-1}$) variants in complex with 2OG analogues (2 eq.) yielded several potential needle-like hits, with many of the conditions containing 2.0 M ammonium sulphate. Screens employed included Crystal Screen, Crystal Screen II and PEG/Ion (Hampton Research). Attempts to optimise these initial crystallisation hits were performed by varying the ammonium sulphate concentration (1.5 - 2.2 M) and buffer pH (6.0-8.5). Additionally, additives (Hampton Research) were also screened in the presence of initial hit conditions. However, none of the crystallisation hits could be optimised in this way. Notwithstanding, it is noteworthy that the highest number of hits was obtained with BB-287 (Chapter 7) amongst all the 2OG analogues tested.

9.6.1 Crystallisation optimum solubility screening

Because the initial crystallisation attempts led to mostly clear drops, presumably due to low protein concentration, further attempts were made to crystallise and/or identify a buffer system that could stabilise PHD3-CH C42Q/C52D at a higher concentration. The variant was concentrated to ~26 $mg.ml^{-1}$ using a buffer system containing 50 mM Tris.HCl pH 7.5, 0.2 M NaCl, 10% glycerol and 5 mM TCEP prior to setting up the crystal trays. Examination of crystallisation drops containing PHD3-CH C42Q/C52D (~13/ ~26 $mg.ml^{-1}$ variant + 1mM Fe + 2mM 2OG analogue) revealed some form of precipitation in nearly 98% of the drops (most after one day) irrespective of the ligand used.

Upon examination of the conditions of the clear drops, it was observed that they all contained either PEG 1500 (30%, w/v) or PEG 3350 (20%, w/v). Therefore, buffers containing low molecular weight PEG were considered likely to improve the solubility of this variant. In addition some of the clear drop conditions also contained 0.2 M $CaCl_2/MgCl_2$. Notably, Mg(II) or Ca(II) amongst other tested metals were not observed to bind tPHD2 in MS experiments under non-denaturing conditions (Chapter 2).[17] Thus a solubility-enhancing buffer system for PHD3 was formulated containing Tris.HCl or HEPES-Na pH 7-8.5 with 0.1-0.2 M salt (e.g $MgCl_2/CaCl_2/NaCl/(NH4)_2SO_4$), 10-20%

glycerol/PEG 1500/PEG 3350 and a reducing agent (preferably 2 mM TCEP). The result was consistent with the initially performed searches for lysis buffer using a sparse matrix method (Section 9.2).

9.6.2 Reductive lysine methylation

Reductive lysine methylation as described in Chapter 7 was carried out for PHD3 with the hope of obtaining suitable crystals for diffraction. Crystal screens were set up using ~10 mg.ml^{-1} of the methylated PHD3-CH C42Q/C52D variant. Various attempts to crystallise pre-formed complexes of the variant in the presence of a few 2OG analogues were unsuccessful.

9.6.3 Thermal stability scanning

Thermal stability assays using the PHD3-CH C42Q/C52D variant revealed that the protein was essentially unfolded (high fluorescence) in the absence or presence of 2OG analogues (Fig. 9.6). The variant seemed to be somewhat stabilised in the presence of flavonoid-type compounds, such as luteolin and quercetin (Fig. 9.6). Quercetin was previously reported to inhibit PHD2 and activate the hypoxic system in colon epithelial cell lines.[18]

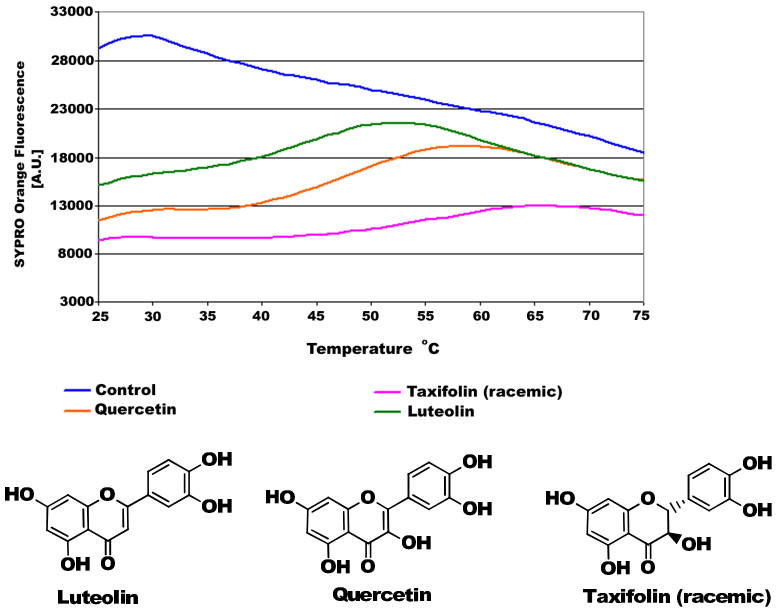


Fig. 9.6 Thermal stability assay showing predominantly unfolded nature of PHD3-CH C42Q/C52D. See the text for description.

9.6.4 Constructing further variants by introducing a homotrimeric interface

In the presence of appropriate inhibitors (compounds A/B/C or BB-287) tPHD2 crystallises as a homotrimer involving residues from the β2/β3 loop (Asp237, Asp250, Arg252, Asp254), a short loop after helix α3 (Ser289 and Lys291) and the C-terminus (Asp394, Arg398, Val401).[19] Except for the catalytically important $Arg252_{PHD2}$ and $Asp254_{PHD2}$, none of these residues are conserved in PHD3. Previous attempts to crystallise PHD3 C42Q and PHD3 C42Q/C52D were not successful. Hence, two multiple variants of PHD3 were produced, R57T/R65K/H72D/K111S/Y113K/E220R/K223V (PHD3-CH MM-1) R57T/R65K/H72D/K111S/Y113K/E216D/E220R/K223V (PHD3-CH MM-2) where the homotrimer forming residues of tPHD2 were introduced by substitution of the analogous PHD3 residues. Studies are ongoing towards purification and further analysis of PHD3-CH MM-1 and MM-2.

9.7 Summary and future direction

C-Terminally His_6-tagged PHD3 and variants were produced that significantly improve PHD3 soluble expression in bacterial cells. Various attempts to optimise buffer and expression conditions as well as a formulation of solubility-enhancing buffer system are presented here (summarised in Fig. 9.7) Future work could be directed towards purification of the multiple PHD3 variants PHD3-CH MM-1, MM-2 and MM-3 for crystallisation and/or biochemical studies.

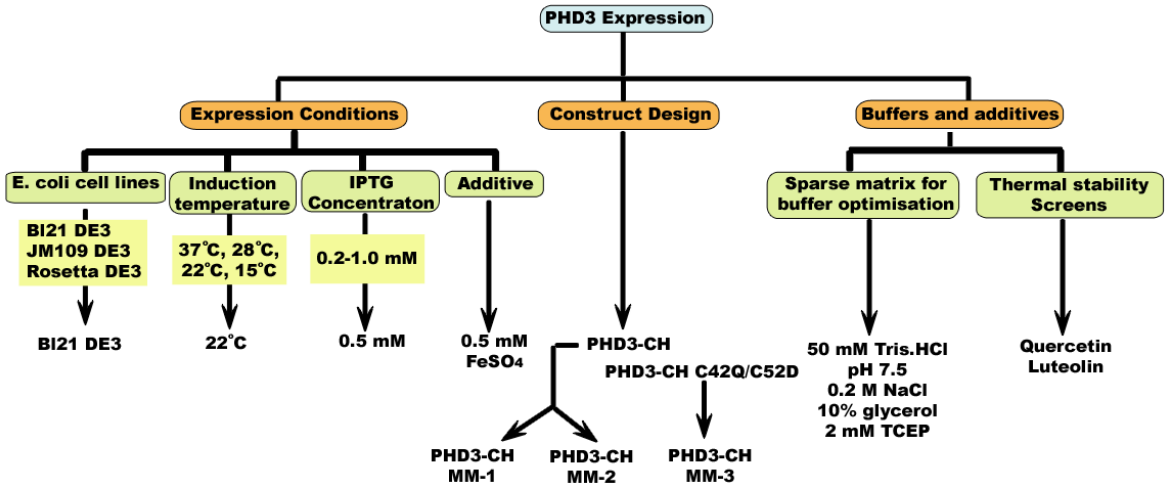


Fig. 9.7 Summary of different approaches followed to improve PHD3 solubility and/or stability.

9.8 References

1. Epstein, A.C.R., Gleadle, J.M., McNeill, L.A., Hewitson, K.S., O'Rourke, J., Mole, D.R., Mukherji, M., Metzen, E., Wilson, M.I., Dhanda, A., Tian, Y.M., Masson, N., Hamilton, D.L., Jaakkola, P., Barstead, R., Hodgkin, J., Maxwell, P.H., Pugh, C.W., Schofield, C.J. & Ratcliffe, P.J. C-elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell* **107**, 43-54 (2001).
2. Oehme, F., Jonghaus, W., Narouz-Ott, L., Huetter, J. & Flamme, I. A nonradioactive 96-well plate assay for the detection of hypoxia-inducible factor prolyl hydroxylase activity. *Anal Biochem* **330**, 74-80 (2004).
3. Rantanen, K., Pursiheimo, J., Hogel, H., Himanen, V., Metzen, E. & Jaakkola, P.M. Prolyl Hydroxylase PHD3 Activates Oxygen-dependent Protein Aggregation. *Mol Biol Cell* **19**, 2231-40 (2008).
4. Lindwall, G., Chau, M., Gardner, S.R. & Kohlstaedt, L.A. A sparse matrix approach to the solubilization of overexpressed proteins. *Protein Eng* **13**, 67-71 (2000).
5. Elcock, A.H. & McCammon, J.A. Electrostatic contributions to the stability of halophilic proteins. *J Mol Biol* **280**, 731-48 (1998).
6. Xiao, L. & Honig, B. Electrostatic contributions to the stability of hyperthermophilic proteins. *J Mol Biol* **289**, 1435-44 (1999).
7. Kumar, S., Ma, B., Tsai, C.J. & Nussinov, R. Electrostatic strengths of salt bridges in thermophilic and mesophilic glutamate dehydrogenase monomers. *Proteins* **38**, 368-83 (2000).
8. Fersht, A.R. Conformational equilibria in -and -chymotrypsin. The energetics and importance of the salt bridge. *J Mol Biol* **64**, 497-509 (1972).
9. Anderson, D.E., Becktel, W.J. & Dahlquist, F.W. pH-induced denaturation of proteins: a single salt bridge contributes 3-5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry* **29**, 2403-8 (1990).
10. Kumar, S. & Nussinov, R. Relationship between ion pair geometries and electrostatic strengths in proteins. *Biophys J* **83**, 1595-612 (2002).
11. Doye, J.P., Louis, A.A. & Vendruscolo, M. Inhibition of protein crystallization by evolutionary negative design. *Phys Biol* **1**, P9-13 (2004).
12. Allen, J.R., Tegley, C.M., Biswas, K., Burli, R., Muller, K.M., Frohn, M.J., Golden, J.E., Mercede, S.J., Neira, S.C., Peterkin, T.A.N., Hungate, R.W., Kurzeja, R., Yu, V. & Dao, J. Quinolone based compounds exhibiting prolyl hydroxylase inhibitory activity, and compositions, and uses thereof. . *PCT Int. Appl.* **WO 2007070359 A2 20070621**(2007).
13. Li, C., Schwabe, J.W., Banayo, E. & Evans, R.M. Coexpression of nuclear receptor partners increases their solubility and biological activities. *Proc Natl Acad Sci U S A* **94**, 2278-83 (1997).
14. Thomas, J.G., Ayling, A. & Baneyx, F. Molecular chaperones, folding catalysts, and the recovery of active recombinant proteins from E. coli. To fold or to refold. *Appl Biochem Biotechnol* **66**, 197-238 (1997).
15. Masson, N., Appelhoff, R.J., Tuckerman, J.R., Tian, Y.M., Demol, H., Puype, M., Vandekerckhove, J., Ratcliffe, P.J. & Pugh, C.W. The HIF prolyl hydroxylase PHD3 is a potential substrate of the TRiC chaperonin. *FEBS Lett* **570**, 166-70 (2004).
16. Leroux, M.R. & Hartl, F.U. Protein folding: versatility of the cytosolic chaperonin TRiC/CCT. *Curr Biol* **10**, R260-4 (2000).
17. Mecinovic, J., Chowdhury, R., Lienard, B.M., Flashman, E., Buck, M.R., Oldham, N.J. & Schofield, C.J. ESI-MS studies on prolyl hydroxylase domain 2 reveal a new metal binding site. *ChemMedChem* **3**, 569-72 (2008).
18. Jeon, H., Kim, H., Choi, D., Kim, D., Park, S.Y., Kim, Y.J., Kim, Y.M. & Jung, Y. Quercetin activates an angiogenic pathway, hypoxia inducible factor (HIF)-1-vascular endothelial growth factor, by inhibiting HIF-prolyl hydroxylase: a structural analysis of quercetin for inhibiting HIF-prolyl hydroxylase. *Mol Pharmacol* **71**, 1676-84 (2007).
19. McDonough, M.A., Li, V., Flashman, E., Chowdhury, R., Mohr, C., Lienard, B.M., Zondlo, J., Oldham, N.J., Clifton, I.J., Lewis, J., McNeill, L.A., Kurzeja, R.J., Hewitson, K.S., Yang, E., Jordan, S., Syed, R.S. & Schofield, C.J. Cellular oxygen sensing: Crystal structure of hypoxia-inducible factor prolyl hydroxylase (PHD2). *Proc Natl Acad Sci U S A* **103**, 9814-9 (2006).

Chapter 10 – Materials and Methods

10.1 General experimental methods

10.1.1 Enzymes and chemicals

The following enzymes and chemicals were used in the work:

Enzymes and chemicals	Source
Tryptone, Yeast Extract and Agar (bacteriological)	Oxoid or Difco
Antibiotics (Ampicillin, Kanamycin and Chloramphenicol)	Melford Laboratories Ltd.
DTT and IPTG	Apollo Scientific Ltd.
Electrophoresis grade agarose	Bioline
Plasmid vectors (pET-28a(+), pET-24a(+), pET-29b(+))	Novagen
Restriction endonucleases (*Nde I, Nhe I, EcoR I, Xho I*)	Promega or New England Biolabs Ltd.
Molecular weight markers for DNA electrophoresis (2-log DNA Ladder 0.1-10 kb)	New England Biolabs Ltd.
T4 DNA Ligase, Pfu Turbo® polymerase, dNTPs	Promega or Stratagene
Tris-carboxymethyl ethylene diamine (TED) resin (immobilised Ni^{2+} affinity chromatography)	Macherey-Nagel
GST purification kit and resin	GE Amersham biosciences
Molecular weight markers for SDS-PAGE (SeeBlue® Plus2)	Invitrogen
Radioactive labelled α-[1-^{14}C]-ketoglutaric acid	Perkin Elmer Life Sciences
4*S*-hydroxyproline, 4*R*-fluoroproline, 4*S*-fluoroproline and Fmoc-Hyp(tBu)-OH	Bachem Ltd.
$^{2}H_2$ gas (99.7 %) and Nitric oxide gas (NO)	Argo International Ltd.

Water was doubly purified using a Millipore Milli-RO system followed by a Milli-Q-Plus system equipped with a 0.22 μm filter to 18.2 MΩ. All other chemicals were obtained in the highest available quality from Sigma-Aldrich, Fisher Scientific Ltd., or Fluka.

10.1.2 Centrifugation

Samples of volumes <1.5 ml were centrifuged either at 4°C using a Beckman Coulter Microfuge® 22R Centrifuge (14,000 rpm), or at room temperature using a Fisher Scientific accuSpin™ Micro (13,000 rpm). All other centrifugation procedures were carried out at 4°C using a Beckman GS-6R centrifuge (4,000 rpm). Cell cultures were centrifuged in a Beckman J2-21 (9,000 rpm). Cell lysates were centrifuged in a Beckman Avanti J-25 Centrifuge (25,000 rpm). Protein concentration was carried out using a Beckman Allegra 21R centrifuge (4,000 rpm).

10.1.3 pH measurement

The pH of solutions was determined using a Mettler Delta 340 pH meter, with a Mettler Toledo Inlab 410 electrode. The pH meter was calibrated to pH 7.0 and pH 4.0 or pH 10.0 before use and the electrode was stored in KCl (3 M).

10.2 Molecular biology

10.2.1 Preparation of oligonucleotides

Oligonucleotides were obtained from Sigma-Genosys Ltd. as lyophilised pellets on a 0.025 μmol scale and purified by desalting. Before use, the oligonucleotides were resuspended in sterile TE (Tris·HCl-EDTA) buffer to give a 100 μM stock solution, which was diluted to 10 μM with sterile deionised water for use in PCR experiments.

10.2.2 Plasmid constructs

cDNA sequences encoding tPHD2 have previously been inserted into the pET-28a(+)/pET-24a(+) vector (Novagen), by Dr K.S. Hewitson and Dr E.G. Flashman,[1] to yield tPHD2 protein with/without an N-terminal His_6-tag. cDNA sequences encoding PHD3 were inserted into the pET-28a(+) vector by Dr K.S. Hewitson. These constructs were used as templates for further sub-cloning or mutagenesis experiments.

10.2.3 Amplification of DNA sequence using PCR (sub-cloning)

Primer oligonucleotides flanking the desired sequence were designed to incorporate the restriction enzyme sites at 5' and 3' ends. Polymerase chain reactions (PCR) were conducted to amplify the gene of interest using a Genius Thermal Cycler from Techne. In all the PCR experiments, Stratagene's high fidelity *Pfu* Turbo DNA polymerase was used, which allows for proofreading of misincorporated nucleotides. A typical PCR reaction mixture contained the following:

dsDNA template (0.1-0.5μg/μl)	1 μl
Forward primer (10 μM)	5 μl
Reverse primer (10 μM)	5 μl
10× reaction buffer	5 μl
dNTP mixture (2 mM)	5 μl
Sterile deionised H_2O	28 μl
*Pfu*Turbo DNA polymerase (2.5 U/μl)	1 μl

$MgCl_2$ (1.5 µl, 0.1 M) and/or DMSO (5 µl) were added to the PCR mixture wherever deemed necessary. The final volume was maintained at 50 µl. The thermal cycler was programmed for an initial hold of 1 minute at 95°C that allowed denaturing of the dsDNA template. Further thermal cycling was followed for 30-35 cycles, where one cycle was typically as below:

Temperature	**Time**	**Step**
94°C	30 seconds	Denaturation
55°C	40 seconds	Annealing
72°C	2 minutes per kb DNA	Extension

A final extension was then carried out at 72°C for 10 minutes and the mixture was then cooled to 4°C.

10.2.4 Site-directed mutagenesis

Stratagene's QuikChange site-directed mutagenesis kit (Stratagene Instruction Manual #200518) was used to make point mutations or enable the deletion of single amino acids. The forward and reverse primers, each incorporating the mutation and complementary to opposite strands of the vector were used to amplify the sequence in opposite directions by PCR. A typical PCR was set up according to Stratagene's QuikChange® Site-Directed Mutagenesis protocol:

dsDNA template (0.1-0.5 ng/µl)	1 µl
Forward primer (10 µM)	2 µl
Reverse primer (10 µM)	2 µl
10× reaction buffer	5 µl
dNTP mixture (2 mM)	5 µl
Sterile deionised H_2O	34 µl
*Pfu*Turbo DNA polymerase (2.5 U/µl)	1 µl

$MgCl_2$ (1.5 µl, 0.1 M) was added to the PCR mixture, if necessary, whilst maintaining the final volume at 50 µl. As before, the PCR mixture was heated to 95°C for 1 minute to denature the DNA. Further thermal cycling (12-18 cycles) with the mutagenic primers yielded an amplified and mutated plasmid. Typical cycling parameters for the QuikChange Site-Directed Mutagenesis were as follows:

Temperature	Time	Step
95°C	30 seconds	Denaturation
55°C*	1 minute	Annealing
68°C	1 minute/kb of plasmid length	Extension

* The annealing temperature was varied from 45 to 60°C depending on the binding of mutagenic primers to the template as verified later by DNA sequencing.

10.2.5 Multi site-directed mutagenesis

Stratagene's QuikChange multi site-directed mutagenesis kit (Stratagene Instruction Manual #200514) was used to produce variants of multiple residues. The mutagenic primers, each incorporating the mutation at different loci were designed to bind the same strand of the template DNA. A high fidelity enzyme blend containing *Pfu* Turbo DNA polymerase was then used to generate dsDNA with one strand bearing multiple mutations and nicks that were sealed by non-specified components in the enzyme blend. A typical PCR reaction mixture was set up according to manufacturer's protocol:

	1-3 mutations	4-5 mutations
dsDNA template (0.1-0.5 µg/µl)	0.5 µl	0.5 µl
Mutagenic primers (10 µM)	1 µl (each)	0.5 µl (each)
10× reaction buffer	2.5 µl	2.5 µl
dNTP mixture (as supplied)	1 µl	1 µl
Pfu Turbo DNA polymerase (2.5 U/µl)	1 µl	1 µl
Sterile deionised H_2O	q.s.	q.s.
Final volume	25 µl	25 µl.

A typical thermal cycling protocol for the QuikChange Multi Site-Directed Mutagenesis used the following parameters:

Temperature	Time	Step
95°C	1 minute	Denaturation
55°C*	1 minute	Annealing
65°C	2 minutes/kb of plasmid length	Extension

* The annealing temperature was varied from 45 to 60°C depending on the extent of binding of primers to the template.

10.2.6 Dpn I digestion of the PCR products (mutagenesis)

The PCR reactions yielded product mixtures containing mutated plasmids and parental (nonmutated) dsDNA templates. It was necessary to remove the non-modified templates before transformation since they both contain the same antibiotic resistance gene. This was achieved by digesting the PCR products with the Dpn I restriction enzyme (Stratagene Instruction Manual #200518). Since Dpn I works on methylated/ hemimethylated DNAs, the plasmid DNA templates were isolated from a dam^+ *E. coli* strain so that the parent DNA could be digested. Dpn I (1 µl) was added to the PCR products, gently mixed and centrifuged for 1 minute before being digested for 3 hours at 37°C.

10.2.7 DNA restriction digest and clean-up of DNA (sub-cloning)

In order to incorporate a DNA insert (gene of interest) into an expression vector, the inserts as well as expression vectors were doubly digested with restriction enzymes for 3 hours at 37°C, and then run on an agarose gel. Gel slices containing the target DNA fragments were excised and the DNA extracted from the gel using the QIAquick gel extraction kit 50 (Qiagen), following the manufacturer's protocol (QIAquick Spin Handbook, catalogue # 28704). A typical restriction digestion mixture contained the following:

	PCR product (µl)	**Purifed plasmid** (µl)	**pET Vector (0.5µg/µl)** (µl)
DNA	30.0	10.0	6.00
Restriction enzyme I (20 U/µl)	1.00	0.05	1.00
Restriction enzyme II (20 U/µl)	1.00	0.05	1.00
10 × NEB double digestion buffer	5.00	2.00	3.00
BSA (1 µg/µl)	5.00	2.00	3.00
Sterile ddH_2O	8.00	5.00	16.0
Final volume	50.0	20.0	30.0.

10.2.8 Agarose gel electrophoresis

Agarose gel electrophoresis was employed to analyse the progression of DNA restriction digestion, to determine the yield and purity of PCR or gene cleaned products, alongside a molecular weight marker (2-log DNA Ladder). A 1% agarose gel was prepared by boiling

0.4 g agarose in 40 ml TAE (Tris·HCl-Acetate-EDTA) buffer and then allowed to cool. Gels were run on a Bio-Rad microprocessor (Model 1000/500 or Model 200/2.0) system at a constant potential of 100 V using 1× TE (Tris·HCl-EDTA) as running buffer, stained with ethidium bromide (0.5 $\mu g \cdot ml^{-1}$) and rinsed with 1x TAE.

Buffers for agarose gel electrophoresis

The following buffers were used for electrophoresis:

1× TE buffer

Tris·HCl	10 mM
EDTA	1 mM

The pH was adjusted to 7.4 with HCl and the buffer solution was autoclaved.

50× Running buffer solution (50× TAE) (per litre)

Tris·HCl	242 g
Glacial acetic acid	57.1 ml
EDTA	14.6 g

Adjusted to pH 8 with HCl.

5× Sample loading buffer

Glycerol	30% (v/v)
Xylene cyanol FF	0.25% (v/v)
Bromophenol blue	0.25% (w/v)
TAE buffer	2% (v/v)

10.2.9 DNA ligation

For ligation reactions, the DNA insert and vector (3:1 w/w) were incubated in a reaction mixture (10 µl) containing 1 µl of T4 ligase (3 U/µl) at either 4°C for 24 hours, or 15°C for 16 hours. A typical ligation mixture contained the following reagents:

DNA PCR insert	6.5 µl
Vector	1 µl
10× ligase buffer	1 µl
DNA ligase (T4)	1 µl

10.2.10 Competent cells

The following *E. coli* strains were used with the genotype (Stratagene, Competent Cells Brochure):

Strain	Application	Genotype
XL1-Blue	Plasmid preparation	endA1 gyrA96(nal^R) thi-1 recA1 relA1 lac glnV44 F'[::Tn10 proAB$^+$ lacI^q Δ(lacZ)M15] hsdR17(r_K^- m_K^+)
XL10-Gold	Transformation of ligation product DNA	endA1 glnV44 recA1 thi-1 gyrA96 relA1 lac Hte Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 tet^R F'[proAB lacI^{q}ZΔM15 Tn10(Tet^R Amy Cm^R)]
BL21 (DE3)	Protein expression	F$^-$ ompT gal dcm lon hsdS_B(r_B^- m_B^-) λ(DE3)*
JM109 (DE3)	Protein expression	endA1 glnV44 thi-1 relA1 gyrA96 recA1 mcrB$^+$ Δ(lac-proAB) e14- [F' traD36 proAB$^+$ lacI^q lacZΔM15] hsdR17($r_K^-m_K^+$) λ(DE3)*
Rosetta (DE3)	Protein expression	F$^-$ ompT hsdS_B(R_B^- m_B^-) gal dcm λ(DE3)* lacY1 pRARE (Cam^R).

* DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5] denotes a λ prophage carrying a chromosomal copy of the T7 RNA polymerase gene and lacI^q
For DNA amplification, a dam$^+$ (DNA adenine methylation) *E. coli* strain (Section 10.2.6) was used to ensure that the isolated plasmids were suitable for Dpn I digestion.

10.2.11 Transformation

For transformation, an aliquot of competent cells (50 µl) was transferred to a 1.5 ml sterile Eppendorf vial and the plasmid/ PCR product (1 µl) or ligation mixture (2 µl) was added with gentle mixing. In the case of transforming ligation mixtures or PCR products, cells were treated with β-mercaptoethanol (2 µl) to enhance transformation efficiency. The transformation mixture was chilled on ice for 30 minutes, followed by a heat shock for 30 seconds at 42°C to induce DNA uptake. In order to allow the bacterial cells recover from the heat shock, the cells were incubated at 37°C for an hour with 200 µl 2× TY medium (or SOC medium in case of *E. coli* Rosetta (DE3)) and plated on LB-agar plates containing the appropriate antibiotics (kanamycin at 30 $\mu g \cdot ml^{-1}$, ampicillin at 100 $\mu g \cdot ml^{-1}$, chloramphenicol 35 $\mu g \cdot ml^{-1}$). Finally the plates were incubated at 37°C overnight.

10.2.12 Plasmid purification

A single colony resulting from transformation was inoculated into 10 ml 2× TY medium containing the appropriate antibiotic(s). The starter culture was grown overnight at 37°C with shaking at 220 rpm. Plasmid DNA was purified from the culture using the Qiagen

Miniprep Kit following the manufacturer's protocol. To verify whether or not the insert had been properly cloned, 10 μl of miniprep product was doubly digested with relevant restriction enzymes and subjected to analysis by agarose gel electrophoresis.

10.2.13 DNA concentration determination

DNA concentration was determined by measuring the absorbance at 260 nm (A_{260}) sample (1.5 μl) using a Nanodrop ND-100 spectrophotometer and the millimolar extinction coefficient (ε_0), defined: A_{260}/ε_0 = [DNA] mM. The value of ε_0 was determined by the base composition of the DNA and was obtained by summing the contributions of each base.

Base	*Contribution to ε_0*
T	$8.8 mM^{-1} \cdot cm^{-1}$
C	$7.3 mM^{-1} \cdot cm^{-1}$
G	$11.7 mM^{-1} \cdot cm^{-1}$
A	$15.4 mM^{-1} \cdot cm^{-1}$

10.2.14 DNA sequence verification

All plasmids were verified by DNA sequencing either at the Geneservice Ltd., Department of Biochemistry, University of Oxford, or at the William Dunn School of Pathology, University of Oxford. The results were compared to wt sequence to ensure the amplified insert or the mutated plasmid sequence was as desired and in the correct reading frame.

10.3 Sterile techniques

Standard sterile techniques were used throughout the current study. Plates were poured and streaked in a laminar flow hood (Envair Ltd. Microbiological Safety Cabinet Class II) using media and equipment that had been sterilised in a Thermo Life Sciences BA 852 autoclave at 121°C for 20 minutes. Solutions of IPTG, antibiotic and other heat labile components were filter-sterilised using 0.2 μm filters. Bacterial plate cultures were incubated in a Heraeus Type B 6030 Incubator. Liquid starter cultures were grown overnight at 37°C with shaking (250 rpm), in a New Brunswick Scientific G24 Environmental orbital shaker.

10.3.1 Growth media

The following growth media were used:

Luria Broth (LB) medium (per litre)

Tryptone	10 g
Yeast extract	5 g
NaCl	10 g

15g of agar per litre was added prior to autoclaving for solid media.

2× Tryptone/Yeast extract (2× TY) (per litre)

Tryptone	16 g
Yeast extract	10 g
NaCl	5 g

S.O.C. medium (per litre)

Tryptone	20 g
Yeast extract	5 g
NaCl	58.44 mg
KCl	18.64 mg
$MgCl_2.6H_2O$	2.03 g
$MgSO_4.7H_2O$	2.47 g
Glucose monohydrate	3.96 g.

10.3.2 Preparation of competent cells

Bacterial strains were made 'competent' by treatment with $CaCl_2$ in the early log phase of growth. Glycerol stocks of *E. coli* strains BL21 (DE3), JM109 (DE3) and Rosetta (DE3) were from Promega. Bacterial strains were streaked onto LB-agar plates with appropriate antibiotics and colonies grown overnight at 37°C. One colony was used to inoculate 100 ml sterile 2× TY medium and was grown on a shaking platform (250 rpm) at 37°C overnight. This starter culture (800 µl) was used to inoculate 400 ml LB/ SOC medium, which was grown with shaking at 37°C until the optical density at 600 nm (A_{600}) reached 0.6-0.8. Cells were centrifuged at 750 ×g, 4°C for 5 minutes, and resuspended in 200 ml sterile 0.1 M $CaCl_2$/ 20% glycerol. After incubation on ice for 20 minutes, the cells were centrifuged as before and resuspended in 8 ml 0.1 M $CaCl_2$/ 20% glycerol. Aliquots (200 µl) of cells were stored in sterile Eppendorf vials at -80°C.

10.4 Protein expression and purification

10.4.1 Protein expression trial

Protein expression trials were performed to determine suitable IPTG concentrations, induction temperatures and incubation periods for optimum protein production. Following transformation into the bacterial cells (Section 10.2.10), single colonies were inoculated into 10 ml 2× TY media (overnight, 37ºC, shaken at 200-250 rpm). Starter culture (1 ml) thus obtained was used to inoculate 100 ml of 2× TY media which was then grown at 37ºC with shaking until the cells reached the mid-log phase of growth where A_{600} = 0.6-0.8. The flasks were then transferred to a pre-equilibrated incubator at the desired temperature (15°C, 22°C, 28°C or 37°C) for 15 minutes. Variable amounts of IPTG (0.1-1.0 mM) were added to induce protein expression and the flasks were incubated for different lengths of period (usually overnight at 15°C, 6 hours at 22°C and 3-4 hours at 28/37°C). The cells were then harvested at 9,000 rpm (4°C, 10 minutes). Protein expression levels were assessed by SDS-PAGE analyses (Section 10.5.1) using the crude lysates (Section 10.4.3) for each expression condition.

10.4.2 Scale-up of protein expression

On the basis of optimum expression condition, 5-10 ml of the overnight starter culture was used to inoculate 2 l flasks containing 500-600 ml of 2× TY growth media with appropriate antibiotic added. The flasks were incubated at 37°C until the A_{600} reached in the range of 0.6-0.8, after which IPTG was added to the most favourable concentration for the production of soluble protein and growth was continued for an optimum incubation period. Cells were harvested by centrifugation at 9,000 rpm, 4ºC for 15 minutes and the wet cell pellets were stored at -80ºC for protein purification.

10.4.3 Preparation of cell extract

Frozen cell pellets were resuspended in 5 ml lysis buffer (50 mM Tris·HCl pH 7.5 or column binding buffer as below) per gram of wet cell pellet. The cell suspension was then lysed by sonication on ice (4 × 45 sec sonication, 45 sec rest in between), using a SANYO Soniprep 150 sonicator, 10 μm microns amplitude (for small scale) or SONICS VibraCell sonicator, with a microtip at 60% amplitude (for large scale). DNAse I (0.02 mg/ml) and $MgCl_2$ (5 mM) were added to the lysate to digest chromosomal DNA, which was left on ice for 5 minutes, and then centrifuged at 9,000 rpm (20 minutes at 4°C) to separate the

soluble proteins from insoluble cellular contents. The supernatant was retained for the assessment of protein expression levels or for the purification of soluble proteins.

10.4.4 Protein purification

Enzymes were purified at 4°C using an AKTA Fast Protein Liquid Chromatography (FPLC) System, Amersham Pharmacia Biotech, (Liquid Chromatography Controller UPC-900 equipped with a high precision P-920 pump). The FPLC system was controlled from a PC running UNICORN v.4.00. All buffers were freshly prepared with ddH_2O and filtered through 0.2 µm Millipore filter before use.

10.4.4.1 Purification of His_6-tagged proteins

Immobilised Ni^{2+} ion affinity chromatography (IMAC) was employed for rapid purification of His_6-tagged proteins. A strong metal chelator, pentadentate Protino Ni-TED (Tris-carboxymethyl ethylene diamine) resin was used for this purpose (Macherey-Nagel User Manual, Nov. 2005/Rev. 02). TED occupies five of the six binding sites in the coordination sphere of the Ni^{2+} ion, leaving the remaining coordination site for protein binding. Compared to TED, other chelating agents such as NTA (nitrilotriacetic acid) have four binding sites available for the Ni^{2+} ion with two remaining sites for protein binding. The additional chelation site of TED with Ni^{2+} minimises metal contamination during purification and increases specificity for His_6-tagged proteins.

The cell extracts were prepared in Protino Ni-TED lysis-equilibriation-wash (LEW) buffer as described before (Section 10.4.3). The crude cell lysate was filtered through a Millipore 0.45 µm filter before application on a TED column. The cleared lysate was then loaded onto a column pre-equilibriated with 10× column volume LEW buffer (1 CV = ~10ml). Followed by another 10 CV column wash, the protein was finally eluted with ~2 CV of elution buffer. The procedure yielded His_6-tagged proteins of >90% purity by SDS-PAGE (see section 10.5.1). The protein was concentrated by ultrafiltration (10 kDa MW cut-off, Amicon ultra-15, Millipore) and subjected to thrombin cleavage when required. The buffers used were as follows:

Buffer	NaCl	Imadazole	Tris·HCl	pH
1× Lysis-equilibriation-wash (LEW) buffer	0.5 M	-	20 mM	7-8
1× Elution Buffer	0.5 M	0.5 M	20 mM	7-8.

10.4.4.2 Ion exchange chromatography

Ion-exchange chromatography involves the separation of proteins based on the reversible interaction between a charged protein molecule and an oppositely charged chromatography medium. All proteins in the current study had theoretical isoelectric points (pI) greater than 7.0. A cation exchanger (S, SP) was preferred for purification because the proteins with higher pI values would have a net positive charge at an acidic pH of the mobile phase (0.1 M MES pH 5.8) used for the cation exchanger. The cell extracts were prepared in this mobile phase (buffer A) and passed over a 50 ml SP Sepharose column (Amersham Biosciences) attached to the FPLC. The protein was eluted with an increasing gradient of buffer B containing 0.1 M MES, 1 M NaCl, pH 5.8. All the eluted fractions and the discarded flow-through (during the wash step) were analysed by SDS-PAGE. The fractions containing protein of highest purity were pooled and subjected to further purification by size exclusion chromatography (SEC), if necessary.

The cation exchange (SP Sepharose) buffers were as follows:

Buffer	MES hydrate	NaCl	pH
1 × Buffer A (binding)	0.1 M	-	5.8
1 × Buffer B (elution)	0.1 M	1 M	5.8

10.4.4.3 Cleavage of His_6-tag

High-grade thrombin (Novagen) was used to remove the His_6-tag from the N- or C-terminus of a fusion protein according to standard protocol: an accurate amount of thrombin (1 U per mg protein) plus 10× thrombin cleavage buffer were added to the protein solution as per manufacturer's instructions. The thrombin-protein mixture was incubated at 4°C overnight and then resolved by SEC (see below).

10.4.4.4 Size exclusion chromatography (SEC)

SEC (gel-filtration or gel-permeation chromatography, GPC), uses porous particles to separate molecules of different sizes. A Superdex S75, 300 ml (Pharmacia) gel filtration column was used throughout the study for final step purification. Typically, the column was fixed to the Akta FPLC and equilibrated with 1CV of water and 2CV of buffer (0.1 M Tris·HCl pH 7.5, 0.1 M NaCl) at a flow rate of 2-3 ml·minute^{-1}. The proteins were concentrated to around 1-5 ml using an Amicon Ultra-15, Millipore centrifugal filter membrane and loaded onto the column using 2 ml/ 10 ml loop at 0.5-1 ml·minute^{-1}. After loading the proteins, the flow rate was increased to 2-3 ml·minute^{-1} with the UV baseline corrected to zero. The protein fractions were collected by an automatic fraction collector by setting the UV peak level at 25 mAU and a typical fraction size of 5 ml. The eluted fractions were analysed for purity by SDS-PAGE. Those fractions containing pure protein were pooled together and buffer exchanged to 50 mM Tris·HCl pH 7.5 (unless specified otherwise) using a PD-10 desalting column (Amersham Biosciences) and stored at -80°C.

10.4.4.5 Protein concentration determination

Protein concentrations were determined using the Bradford method. Bradford reagent (2 ml) was added to the protein sample (5 µl) and incubated at room temperature for 1~3 minutes. Absorbance at 595 nm was recorded with reference to a blank sample containing Bradford reagent and 5 µl of buffer. If absorbance exceeded 1.0, the protein sample was diluted 10-fold. Protein concentrations were then calculated by reference to a calibration curve of bovine serum albumin at known concentrations. Variations in Bradford reagent were apparent between preparations, so calibration curves were produced for each reagent preparation. The protein was finally concentrated to <40 mg·ml^{-1} (typically 25 mg·ml^{-1}).

Bradford reagent contains the following:

Coomassie brilliant blue G250	0.01% (w/v)
Phosphoric acid	8.80% (v/v)
Ethanol	4.75% (v/v)

10.5 Protein characterisation

10.5.1 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

Sodium dodecyl sulphate (SDS) is an anionic detergent, which binds in a non-sequence specific manner to denatured proteins in the presence of a reducing agent, e.g. β-mercaptoethanol. The resultant SDS-polypeptides are separated on a polyacrylamide gel according to their molecular weights. The molecular weights of the protein samples can therefore be approximated when compared with standard markers. Briefly, bilayered gels are run in a reservoir containing running buffer of a different pH and ionic strength so that the negatively charged SDS-polypeptide migrates through a porous stacking gel, followed by a resolving gel with higher acrylamide content to separate the samples. The polyacrylamide gels were made using the Bio-Rad Mini Protean II system. A typical protein sample contained 15 µl protein plus 5 µl 4× SDS loading buffer. The protein samples were boiled at 95°C for 3-5 minutes, and 5µl (typically 5-10 µg) loaded onto the gel, which was electrophoresed at 200 V for 55 minutes. The gels were then stained with Coomassie Blue for 15 minutes before being rinsed in a mixture of 10% methanol and 10% acetic acid. When completely destained, gels were dried using gel drying films (from Promega).

Buffers for SDS-PAGE

10× Running buffer solution (per litre)

Tris·HCl	30 g
Glycine	144 g
SDS	10 g

4× SDS-PAGE sample loading buffer (per 100ml)

Tris·HCl, 1.0 M, pH 6.8	5 ml	(50 mM)
Serva blue 10% (w/v)	0.04 ml	(0.04%, w/v)
SDS	2 g	(2%, w/v)
Glycerol	10 ml	(10% v/v)
ddH_2O	q.s. to 95 ml.	

95 ml of this mixture was heated to ~50°C to dissolve the SDS and then 5 ml of β-mercaptoethanol was added.

SDS-PAGE gels

Reagent	Separating gel (12.5%)	Stacking gel (3%)
0.7 M Tris·HCl pH 8.8	2.50 ml	-
1.0 M Tris·HCl pH 6.8	-	1.25 ml
ddH_2O	3.14 ml	3.14 ml
30% Acrylamide	4.16 ml	0.5 ml
10% SDS	100 µl	50 µl
10% APS	100 µl	50 µl
TEMED	8 µl	5 µl

The freshly prepared APS and the TEMED were added immediately before pouring the gels.

SDS-PAGE gel stain

Acetic acid (glacial)	10% (v/v)
Methanol	50% (v/v)
Coomassie brilliant blue R250	0.25% (w/v)

SDS-PAGE gel destain

Acetic acid (glacial)	10% (v/v)
Methanol	10% (v/v)

10.5.2 Native PAGE

Polyacrylamide gel electrophoresis (PAGE) under non-denaturing conditions was performed on samples to investigate the presence of possible aggregates by following the above-mentioned procedure for SDS PAGE except without using the detergent (SDS). The samples were run at a lower voltage of 80-100 for a longer period of time (usually, 2-3 hours) at 4°C.

10.5.3 Mass spectrometry

10.5.3.1 Denaturing ESI-MS

Denaturing mass spectrometry of protein samples at 10 µM was performed using a Fisons Platform Quadupole mass spectrometer equipped with an ESI source fitted to an Agilent

1100 autosampler inlet. The sample (10 µl) was quenched in a mixture of 50% acetonitrile/ 0.1% formic acid in water with a cone voltage of 30 V.

10.5.3.2 Non-denaturing ESI-MS

For non-denaturing MS measurements, proteins were buffer exchanged into 15 mM ammonium acetate buffer (pH 7.5). Spectra were obtained on a Waters Q-TOF mass spectrometer (Q-TOF micro, Micromass) with a chip voltage of 1.70 kV ± 0.2 kV and a delivery pressure of 0.25 psi (1 psi = 6.89 kPa) using a NanoMate HD Robot chip-based nano-electrospray device (Advion Biosciences). Calibration and sample acquisitions were performed in the positive ion mode in the range of 500–5,000 *m/z*. Typically, 15 µM protein samples were sprayed at a cone voltage of 80 V (alternately, 50 V and 200 V wherever mentioned) with acquisition/scan times of 10 s/1 s. The pressure at the interface between the atmospheric source and the high vacuum region was 6.60 mbar. Data were processed with MASSLYNX 4.0 (Waters).

10.5.3.3 MALDI-MS

Hydroxylation and/or other modification of peptides were analysed using MALDI-TOF microMXmass spectrometer (Waters Micromass) in the negative ion reflectron mode. All samples for MS analyses were mixed with α-cyano-4-hydroxy-cinnamic acid (CHCA) as MALDI matrix (LaserBio Labs) (1:1) and spotted onto the target plate. Unless otherwise stated, the following conditions were used: source voltage = 12,000 V, reflectron voltage = 5,200 V, pulse voltage = 1,950 V, micro channel plate (MCP) detector voltage = 2,350 V, laser energy = 100-150 %, matrix suppression = 800 amu and TLF delay = 500 ns.

10.5.4 Elemental analysis by atomic absorption spectroscopy

For elemental analysis by atomic absorption spectroscopy, the protein samples were diluted with 50 mM Tris·HCl pH 7.5 to a final concentration of ~7 µM (10 ml). The samples were analysed for Fe^{3+}, Co^{2+}, Ni^{2+} and Zn^{2+} using an inductively coupled plasma optical emission spectrometer (ICP-OES, Thermo Jarrell Ash AtomScan 16, Thermo Electron Corporation) at the Inorganic Chemistry Laboratory (ICL), Oxford.

10.5.5 Circular dichroism (CD) spectroscopy

Circular dichroism (CD) spectroscopy is frequently used in proteomics to estimate the proportion of secondary structures, i.e. α-helices, β-strands or random coils in a protein molecule depending on the absorption characteristics of its component L-amino acids. In the current study, all CD spectra were recorded on a Jasco 720 spectropolarimeter at 20ºC in a 500 μm slit width over the range of 190–255 nm at a protein concentration of ~25 μM in 10 mM phosphate buffer, pH 7.2. Scanning took place at 50-100 nm·minute^{-1}, with readings taken every nm. Each spectrum was repeated three times to obtain an average result.

10.6 Biochemical assays

10.6.1 2OG turnover assay

Purified recombinant enzymes and synthetic peptides were tested for their ability to stimulate PHD-dependent decarboxylation of 1-[^{14}C]-labelled 2OG, as described.[2,3] Standard assay conditions comprised a total volume of 100 μl: 50 mM Tris·HCl pH 7.5, 4 mM ascorbate, 288 μM 2OG, 3.7 μM 1-[^{14}C]-2OG (specific activity 56.8 μCi·nmol^{-1}, stock concentration 1.83 mM), 100 μM $FeSO_4 \cdot 7H_2O$, 0.66 mg·ml^{-1} catalase (H_2O_2 scavenger), 4 μM enzyme and 100 μM substrate. Briefly, the assay was set up in three drops, one containing enzyme (10 μl 40 μM), another containing substrate (10 μl 1 mM) and the third (80 μl) containing all other reagents in a 5 ml tube. A 0.5 ml Eppendorf vial containing 200 μl hyamine hydroxide was added and the tube was then sealed with a rubber septum. The reaction mixtures were incubated with shaking at 37°C for 15 minutes and then quenched with methanol (200 μl). Reaction tubes were then kept on ice for 20 minutes, before the hyamine hydroxide was removed and treated with scintillant liquid for radioactive count measurement (Beckman, LS6500). Conversion of 2OG was calculated from the nmoles of 1-[^{14}C]-2OG that had been converted to $^{14}CO_2$ gas using the equation,

$$\frac{C_{Exp}}{C_{Total}} \times 29.17 = \text{nmoles of 2OG turned over},$$

where C_{Exp} is the experimental count in dpm, C_{Total} is the total count (dpm) for 29.17 nmoles of 2OG used per reaction. The activity (frequency or turnover number) is therefore reported as nmoles 2OG turned over/0.4 nmole enzyme/15 minutes. Assays were performed in triplicate and were repeated for consistency.

10.6.2 Hydroxylation assay

Assay conditions utilised similar 2OG turnover assay conditions except without the radioactive 2OG. Reactions were quenched with an equal volume of 0.1 % formic acid/ MeOH at 4°C. 4-Oxoprolyl products were derivatised as 2,4-dinitrophenylhydrazone adducts using 2,4-dinitrophenylhydrazine (1:50) prior to MS analyses (Section 10.5.3.3). Control experiments using equimolar amounts of non-hydroxylated and proline-4-hydroxylated peptides (10 µl 50 µM each) suggested that an addition of 16 Da corresponding to hydroxyl group might not significantly interfere with the ionisation/desorption of the peptides. Percentage of hydroxylation was calculated as illustrated in Fig. 10.1.

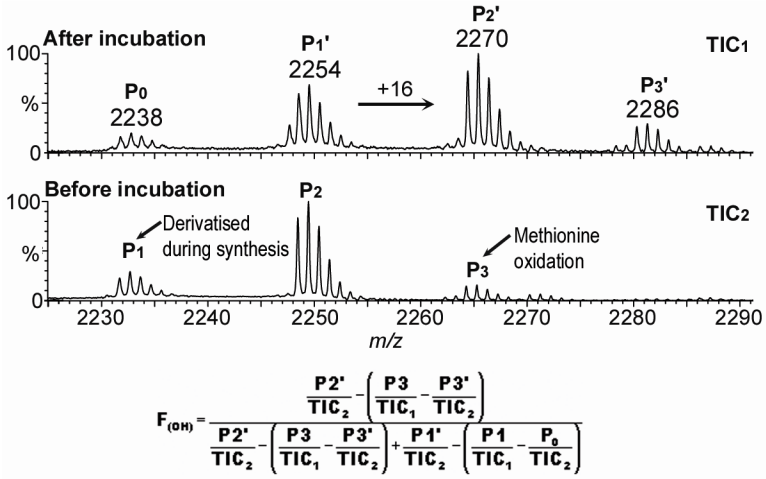


Fig. 10.1 **Calculation of percentage of hydroxylation from MALDI-MS data.** Percentage of hydroxylation was calculated from peak intensity to corresponding total ion count (TIC) ratios that minimise any possible change(s) in experimental conditions (e.g. concentration of peptide or the matrix, spot size, pipette errors etc.). In the control experiment, MS spectrum of a typical CODD peptide without being incubated with the enzyme contained P1: a possible derivatised (e.g. cyclised) peptide corresponding to -16 Da of the calculated mass, P2: the non-hydroxylated peptide, P3: a partial methionine (Met560/568_{CODD}) oxidation product that was observed previously.[4] All these can be hydroxylated after incubation with the enzyme.

10.7 Protein crystallisation

10.7.1 Crystallisation trials

Protein samples of high purity (≥95%) were concentrated to ~20mg·ml^{-1} in 50 mM Tris·HCl pH 7.5 for initial crystal trials on the basis of precipitation results using the Hampton Pre-Crystallisation Test (Hampton Research). Initial crystallisation screens were performed using the sitting-drop vapour-diffusion method at the Oxford Protein Production Facility (OPPF), Oxford. A Robbins Hydra (Robbins Scientific) was used to dispense solutions promoting crystal growth into 96-well plates (Greiner). Screens included Crystal Screen, Crystal Screen 2, Index, Grid Screen PEG 6000, Grid Screen PEG/LiCl, Grid

Screen MPD, Grid Screen Ammonium Sulfate and PEG/Ion (Hampton Research) and Wizard I and II (Emerald BioSystems). Crystallisation drops contained 100 nl protein and 100 nl reservoir solution, pipetted using a Cartesian Dispensing Robot and a Cartesian Microsys dispenser (Genomic Solutions), respectively. Crystallisation hits were scaled up using VDX plates by sitting/hanging drop vapour diffusion method. Plates were stored in a temperature-controlled TAP Homebase storage vault (The Automation Partnership, Royston, England) at 21°C/4°C. Multiple optimisation screens of initial crystal hits were set up around all crystal-forming conditions, both manually and using a RoboGo dispenser (MWG Biotech AG). Optimisation parameters that were varied included buffer composition, molarity and pH, drop size, protein: reservoir solution ratio and concentration.

10.7.2 In-house data collection and processing

For diffraction experiments crystals were transferred directly to appropriate cryo solution (usually crystallisation buffer containing 20-30% glycerol) and frozen in liquid N_2. In order to screen diffraction quality and/or collect datasets, crystals were mounted in an Oxford Cryosystems Cobra cryostream at 100 K by using a X8 Proteum (Bruker-AXS) equipped with a Microstar microfocus rotating anode X-ray generator producing CuK_α (1.542 Å) radiation, four circle KAPPA goniometer, and SMART 6000 CCD detector. Data collection strategy and indexing were performed by using PROTEUMPLUS aiming for a target redundancy (usually 20) in fine-slice mode (0.25°). Data were integrated using SAINT and scaled by using SADABS followed by analysis and manipulation in XPREP (Bruker-AXS, 2005). In the cases where better resolution data were desired, crystals were preserved in liquid N_2 for further diffraction experiments using synchrotron radiation (see methods for chapters below).

10.8 General cloning and purification of tPHD2 variants (Chapters 2 to 8)

10.8.1 Cloning of tPHD2 variants

Stratagene's QuickChange site-directed mutagenesis kit was used to make point variants and C-terminal truncations using the *phd2*$_{(181\text{-}426)}$/pET28a(+)/pET24a(+) plasmid construct(s) (Section 10.2.2) as template (for primers see Appendix 11). Stratagene's QuikChange multi site-directed mutagenesis kit was used to produce variants of multiple residues using the aforementioned constructs. A deletion variant (PHD2Δloop) was produced whereby residues 238–250 (the β2/β3 loop) were replaced by an Ala-Ala linker

using multi site-directed mutagenesis. All constructs were verified by DNA sequencing at the Sir William Dunn School of Pathology, University of Oxford or at the Geneservice Ltd., Department of Biochemistry, University of Oxford.

10.8.2 Purification

Variant or wt tPHD2 was purified by immobilised Ni^{2+} affinity chromatography followed by cleavage of the His_6-tag by thrombin and a final purification by gel filtration (Superdex S75, 300 ml, Pharmacia). Alternatively the proteins were purified using cation exchange chromatography (SP Sepharose 50, Amersham Biosciences) followed by SEC. Finally they were buffer exchanged into 50 mM Tris·HCl pH 7.5 and concentrated to ≤40 mg/ml. Proteins were of >95 % homogeneity, as determined by SDS-PAGE; and were characterised by MS analyses, typically under non-denaturing conditions (see Section 10.5.3.2).

10.9 Experimental details for Chapter 2

10.9.1 Analysis of protein stability by ThermoFluor

A fluorescence-based thermal shift (ThermoFluor) assay[5] was used to monitor protein unfolding with respect to temperature using SYPRO orange as a fluorophore at the OPPF, Oxford or at the Structural Genomics Consortium (SGC), Oxford. Typically, the reaction mixture contained 5 µl (10 µg) protein sample, 5 µl dye (10×) and 40 µl experimental buffer per experiment and was cycled/fed into an Opticon thermal cycler (MJ Research Inc.) or Mx3005p (Stratagene). The midpoints of the unfolding transition were calculated as an approximation of melting point (T_m)[6] from the melting curve using OpticonMONITOR Analysis Software version 3.1 supplied by the manufacturer.

10.9.2 Sedimentation velocity analytical ultracentrifugation

Sedimentation velocity experiments were conducted with an Optima XLI (Beckman Coulter) using an An50 Ti eight-hole rotor and with 400-µl of tPHD2 sample (~1 $mg \cdot ml^{-1}$) in standard double-sector Epon centrepieces equipped with sapphire windows.

10.9.3 Crystallisation details

Crystallisation screens were carried out in 96-well Greiner crystallisation plates using a Cartesian Dispensing Robot (Genomic Solutions Ltd.) at the OPPF.[7] Crystals of

tPHD2.CODD complex were obtained at 20°C using 200 nl drops (100 nl protein (40 $mg \cdot ml^{-1}$) in 50 mM Tris·HCl pH 7.5/1 mM $MnCl_2$/1 mM NOG/0.1 mM HIF-1α $CODD_{556-574}$ plus 100 nl well solution). The well solution consisted of 0.2 M $MgCl_2$ and 20% polyethylene glycol 3350. The crystals were transferred directly to a 30% glycerol/70% well cryo-solution and frozen in liquid N_2. A tPHD2.Fe(II).B.$CODD_{568-574}$ crystal was obtained by soaking pre-formed $PHD2_{181-426}$.Fe(II).B crystals in 50% sodium malonate pH 7.5 solution containing 10 mM 4*R*-hydroxylated Pro 564 HIF-1α $CODD_{556-574(Hyp564)}$ peptide under anaerobic conditions for 72 hrs. tPHD2.Fe(II).A/B crystals were obtained as described[8].

10.9.4 Data collection and processing

All datasets were collected on a single crystal at 100 K. Data set for the tPHD2.CODD complex was collected at the ESRF microfocus beamline ID23-2 equipped with a MarMOSAIC 225 CCD detector. tPHD2.Fe(II).B.$CODD_{556-574(Hyp\ 564)}$ dataset was obtained at Daresbury beamline SRS-10.1 equipped with a MarMOSAIC 225 CCD detector. Data were integrated and scaled using HKL2000[9].

10.9.5 Structure solution by molecular replacement and refinement

Structures were solved by molecular replacement using PHASER[10] (search model PDB ID 2G19). Parameter and topology files for compounds B were generated using PRODRG[11] for refinement in CNS[12]. Iterative cycles of model building in COOT[13] and slowcool-simulated annealing refinement by using the maximum-likelihood function and bulk-solvent modelling in CNS proceeded until the decreasing R_{cryst}/R_{free} no longer converged. PROCHECK[14] was used to monitor the geometric quality of the model between refinement cycles and identify poorly modelled areas needing attention. Water molecules were added to peaks >1.8σ in $2F_o - F_c$ electron density maps that were within hydrogen bonding distance to protein with reasonable hydrogen bonding geometry.

10.10 Experimental details for Chapters 3 and 4

10.10.1 Crystallisation details

tPHD2 P317R.Mn(II).B and R371H.Mn(II).B crystals were grown by vapour diffusion at 20°C with 0.1 M tri-sodium citrate pH 5.6, 20% PEG 4000 plus 20% 2-propanol and 0.1 M MES-Na pH 6.5, 30% PEG monomethyl ether 5000 plus 0.2 M ammonium sulphate,

respectively. Stock solutions of all 2OG analogues were prepared using 100% DMSO to a final concentration of 0.1 M.

10.10.2 Data collection and processing

Data sets for tPHD2 P317R/R371H.Mn(II).B were collected at the Diamond beamline I04 with an ADSC Q315 3x3 CCD detector and were processed with MOSFLM and SCALA (CCP4 suite)[15].

10.10.3 Structure solution by molecular replacement and refinement

Structure solution and refinement was as described in Section 10.9.5.

10.10.4 Modelling of tPHD2 and PHD3 complexes

Models for tPHD2.HIF2αCODD$_{525\text{-}542}$, tPHD2.HIF1αNODD$_{395\text{-}412}$, PHD3.HIF1α CODD$_{558\text{-}574}$ and PHD3.HIF1αNODD$_{395\text{-}412}$ were generated using the crystal structure of tPHD2.CODD as a template in the automodel feature of MODELLERv8.1.[16]

10.10.5 Molecular dynamics simulations studies

MDS experiments described throughout this thesis were performed in collaboration with Dr. C. Domene, Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford. The tPHD2.CODD crystal structure was used as a template for MDS studies. The following systems were analysed: tPHD2.Mn.NOG.HIF1αCODD$_{558\text{-}574}$, tPHD2.Fe(II).2OG. HIF1αCODD$_{558\text{-}574}$, tPHD2.Fe(II).2OG.HIF2αCODD$_{525\text{-}542}$, tPHD2.Fe(II).2OG.LEMLAPY (LXXLAP of CODD), tPHD2.Fe(II).2OG.LTLLAPA (LXXLAP of NODD), tPHD2.Fe(II).2OG.HIF1αNODD$_{395\text{-}412}$ (three models), PHD3.Fe(II).2OG.HIF1α CODD$_{558\text{-}574}$, PHD3.Fe(II).2OG.HIF1αNODD$_{395\text{-}412}$, (three models). An acetyl group was added to the N-terminal (Gln184$_{PHD2}$), and the C-terminal was amidated (Lys408$_{PHD2}$). The NODD (Asp395-Asp412) and CODD (Asp558-Leu574) peptides, and the different LXXLAP motifs were capped similarly. All the residues were in their default protonation state and H atoms were added using the software psfgen of VMD2.6.[17] Histidines were protonated at the delta position. A total of 155 crystallographic water molecules were included in the models. Systems were solvated with a box padding of 18×18×18Å^3 dimension. Counterions were added to achieve neutrality of the simulation systems. In the current study, the drift from an initial model was used as a simple measure

of the relative 'stability' of a given structure in a simulation. The drift was measured as the time-dependent Cα atom root mean square deviation (RMSD) from the initial model. The initial rise in Cα RMSD in the first 0.2-0.3 ns is common in such simulations and may be attributed to relaxation of the protein upon its transfer to the solution environment and/or inaccuracies in the potential function.

Simulations were performed with NAMD2.6[17], and the CHARMM-27 all atom force field[18] with the TIP3 model for water molecules.[19] A set of force field parameters for the 2OG and NOG were developed. Standardised protocols based on the original development of the force field were followed to ensure the transferability of the new chemical compounds. A 1 fs time step was used to integrate the equation of motion. Coordinates were saved every 1 ps. Long-range electrostatic interactions were treated by the Particle Mesh Ewald algorithm,[20] and a 10 Å cut-off was used for the van der Waals interactions. Langevin dynamics controlled the temperature at 25°C, with a damping factor of 5 ps^{-1}. The Nosé-Hoover Langevin pressure control[21,22] was used to keep a pressure of 1 bar (piston period was 200 fs and the damping time scale was 100 fs). Energy was first minimised by a steepest descent procedure (10,000 steps), followed by production runs of 6 to 15 ns of unconstrained dynamics in the NpT ensemble. Overall, this corresponds to a total simulation time of 122 ns. Each system is composed of ~60,000 atoms.

10.11 Experimental details for Chapter 5

10.11.1 Synthesis of peptides

HIF-1α NODD (395-413) and CODD (556-574) peptides used for all Crystallisation trials and activity assays were from Peptide Protein Research Ltd., Fareham, UK. Except for the wt NODD and CODD, all variant peptides were produced by C. Loenarz (Chemistry Research Laboratory, Oxford). Peptides were prepared using a CS Bio CS336 solid phase peptide synthesiser using Rink amide linker, PL-AMS resin (Polymer Laboratories) and standard Fmoc/DIC/HOBt strategy. Peptides were N-terminally biotinylated using biotin p-nitrophenyl ester (Novabiochem), final cleavage (TFA:phenol:H_2O:triisopropylsilane 88:5:5:2) gave peptides as C-terminal amides which were purified by preparative reversed-phase HPLC using Vydac 218TP C18 column (Grace Davison Discovery Sciences), lyophilised and the mass of the peptide product was measured using a Waters Micromass MALDI-TOF mass spectrometer.

10.12 Experimental details for Chapter 6

10.12.1 Crystallisation details

Crystals of tPHD2 F366Y (~25 mg·ml^{-1}) were obtained in complex with Fe(II) (equimolar) and BB-287 (2 equivalents) by sitting drop vapour diffusion at 20°C. The crystallisation conditions employed 2% PEG 400 (v/v), 0.1 M HEPES-Na pH 7.5 and 2 M ammonium sulphate. A stock solution of BB-287 was prepared using 100% DMSO to a final concentration of 0.1 M. Crystals were frozen in a cryo-solution containing crystallisation buffers and 30% glycerol.

10.12.2 Data collection and processing

A dataset was collected for tPHD2 F366Y crystal using an X8 Proteum (Bruker-AXS) equipped with a Microstar microfocus rotating anode X-ray generator producing CuKα radiation (see Section 10.7.2 for further details). Data were integrated using SAINT and scaled by using SADABS followed by analysis and manipulation in XPREP (Bruker-AXS, 2005).

10.12.3 Structure solution by molecular replacement and refinement

Structure solution and refinement was as described in Section 10.9.5.

10.12.4 [^{35}S]-pVHL capture assays

[^{35}S]-Methionine labeled proteins were synthesised by in vitro transcription/translation (IVTT) reactions using TNT T7 Quick Coupled Rabbit Reticulocyte Lysate (Promega), as described previously[23]. N-Terminal biotinylated peptides were used, corresponding to HIF-1α residues 395-413 (NODD) or 556-574 (CODD) with Pro 564 substituted by the indicated proline analogue. Biotinylated peptides (100 µM) were incubated at 37 °C, in a reaction mix containing 4 µM tPHD2, 50 mM Tris·HCl pH 7.5, 300 µM 2OG, 2.5 mM ascorbate, and 50 µM $FeSO_4$. Reactions were commenced by addition of the tPHD2 and stopped after 1.5 hours by dilution in ice-cold 'stop buffer' (20 mM Tris·HCl pH 7.5, 150 mM NaCl, 0.5% octylphenyl-polyethylene glycol and 300 µM desferrioxamine mesylate). The biotinylated peptide was captured using streptavidin-conjugated magnetic beads in 'capture buffer' (20 mM Tris·HCl pH 7.5, 150 mM NaCl), 0.5 % octylphenyl-polyethylene glycol and 100 µM desferrioxamine mesylate) followed by washing. The peptide-coated beads were then further incubated at 4°C with excess [^{35}S]-pVHL produced

by IVTT reaction, diluted in 'capture buffer'. Following two further washes, captured [^{35}S]-pVHL was quantified by SDS-PAGE and autoradiography using a phosphorimager.

10.12.5 Time-resolved fluorescence resonance energy transfer (TR-FRET)

To analyse binding of biotinylated peptides to VCB, a homogeneous TR-FRET assay was carried out by Dr L.A. McNeill (Chemistry Research Laboratory, Oxford).[24] The ternary GST-VCB complex was produced as described[25], purified by successive GST affinity and SEC, thrombin cleavage, and further GST affinity and SEC. VCB was labelled with a Eu^{3+}-cryptate linker (LANCE Eu-W1024-ITC, PerkinElmer; ~6 Eu per VCB). Biotinylated peptide substrate and analogues were made as described in Section 10.11.1. The assay was performed in costar 96 well black clear bottom microassay plates at a final volume of 50 µl in reaction buffer containing 30 mM MES-Na pH 6.0, 10 mM NaCl, 10 mM $CaCl_2$, 0.25% Brij-35. The solution was allowed to reach binding equilibrium by shaking for an hour before detection on a RUBY Star plate reader (BMG Labtech). The reaction was stopped with 50 mM succinic acid in detection buffer (50 mM Tris·HCl, pH 8.0, 0.1 M NaCl, 0.05% Tween 20, 0.5% NaN_3) containing a final concentration of 25 nM streptavidin-APC and 2.5 nM VCB-Eu. The data output is the ratio of the 665 nm and 620 nm emission signals resulting from the 620 nm excitation of streptavidin-allophycocyanin and the 320 nm excitation of Eu^{3+}, respectively. Analyses of the binding curves were performed using curve-fitting procedures (equation $y = y_0 + A \cdot e^{-x/t}$; y, amount of binding; y_0, maximal binding; x, peptide concentration in nM). A comparison of assay mixes for 2OG decarboxylation assay and TR-FRET are given in Table 10.1.

Table 10.1. Comparison of the experimental conditions used for 2OG radioactive assay vs. TR-FRET assay

	2-OG radioactive assay	TR-FRET Assay
Enzyme	4 µM	10 nM, 100 nM
Substrate (CODD)	100 µM	100 nM
2-OG	291.6 µM*	0.25, 2, 10 100 µM
Fe(II)	100 µM	100 µM
Ascorbate	4 mM	2 mM
Dithiothreitol (DTT)	1 mM	-
Catalase	0.3 mg/ml	-
Total reaction mixture	100 µl	50 µl

* includes 3.6 µM of [^{14}C]-2OG

10.12.6 FT-ICR MS experiments

LC FT-ICR MS and MS-MS analyses of peptides was performed at > 100,000 FWHM resolution on a 9.4 T Bruker APEX-Qe hybrid FT-ICR MS coupled to a Dionex Ultimat 3000 LC system.

10.12.7 Synthesis of labelled amino acids

Labelled amino acids and associated peptides were synthesised by C. Loenarz (Chemistry Research Laboratory, Oxford).

10.13 Experimental details for Chapter 7

10.13.1 TR-FRET assay

For inhibition studies TR-FRET assay was employed using the previously described protocol (Section 10.12.5).

10.13.2 Reductive lysine methylation

Protein lysine methylation reaction was performed as previously described[26] using tPHD2 (0.5 mg/ml) in 50 mM HEPES-Na pH 7.5 buffer. Briefly, 20 µl of freshly prepared 1 M dimethylamine-borane complex (ABC) and 40 µl of 1 M formaldehyde were added per ml of tPHD2 solution and were gently mixed and incubated at 4°C. After 2 hours, a further 20 µl ABC and 40 µl formaldehyde were added and the incubation continued for another 2 hour. Following a final addition of 10 µl ABC, the reaction was incubated overnight at 4°C. The reaction mixture was then centrifuged to remove any unwanted precipitation before being purified by SEC. Appropriate peak fractions were pooled and concentrated for further studies.

10.13.3 Crystallisation details

Crystals of tPHD2.Fe(II).BB-287 compounds (~20 $mg \cdot ml^{-1}$ protein + 1mM Fe(II)/ 1mM Mn(II) + 2 mM compound) were grown by vapour diffusion at 20°C under near anaerobic (with Fe) or aerobic (with Mn) conditions. The crystallisation buffer used contained 1.6–2.0 M $(NH_4)_2SO_4$, 2–8% dioxane, 0.1 M MES-Na pH 6.5 and 1 mM $FeSO_4$. Stock solutions of BB-287 compounds were prepared using 100 % DMSO to a final concentration of 0.1 M. tPHD2Me-I.Fe(II).BB-287 (~18 $mg \cdot ml^{-1}$ protein + 1mM Fe(II) + 2 mM BB-287) crystallised in well conditions containing 1.7 % polyethylene glycol 400

(v/v), 15 % glycerol (v/v), 1.7 M ammonium sulphate and 0.085 M HEPES-Na pH 7.5. tPHD2Me-II.Fe(II).BB-287 (~18 $mg \cdot ml^{-1}$ protein + 1mM Fe(II) + 2 mM BB-287) crystallised in 1.1 M di-ammonium tartrate pH 7.0 in a hexagonal shape. Crystals were transferred directly into cryo-solution containing crystallisation buffers and 30 % glycerol and frozen in liquid N_2.

10.13.4 Data collection and processing

Data sets for crystals of tPHD2 and the methylation products (tPHD2Me-I and tPHD2Me-II) were collected at the Diamond beamline I04 with an ADSC Q315 3x3 CCD detector and were processed with MOSFLM and SCALA (CCP4 suite)[15].

10.13.5 Structure solution by molecular replacement and refinement

Structure solution and refinement was as described in Section 10.9.5.

10.14 Experimental details for Chapter 8

10.14.1 Nitric oxide solution preparation

NO solutions were prepared as previously described.[27] Briefly, a solution of 50 mM ammonium bicarbonate solution, pH 8.0, in a rubber-stoppered tube was sparged for 15 minutes with N_2 and then 15 minutes with NO gas (35 bar, density 1.34 $g \cdot L^{-1}$). This resulted in a saturated solution of NO (1.25 mM), which might contain higher oxides of NO that were not quantified.[27] Accurate measurement of NO was not possible since most of the measurement procedures in the literature involve conversion of NO to higher oxide (e.g. Saville-Griess assay[28]).

10.14.2 Preparation of GSNO and SNAP solution

GSNO and SNAP solutions (0.1 M stock) were prepared in 50 mM Tris·HCl pH 7.5 with the final pH adjusted to 7.0.

10.14.3 Preparation of S-nitrosylated tPHD2

Preparative quantities of tPHD2-SNO were produced in 50 mM Tris·HCl pH 7.5 by mixing tPHD2apo (1 $mg \cdot ml^{-1}$) with 250 equivalents of GSNO. After incubation on ice for 2 hours, the resultant tPHD2-SNO was separated from unreacted reagent/ possible by-products by gel filtration column (Superdex S75, 300 ml, Pharmacia). S-Nitrosylation was

verified by mass spectrometry under non-denaturing conditions before further experiments. Prior to activity analyses, 1 eq. of Fe(II) was added under anaerobic conditions. For LC-MS analyses (see below), a 25 fold excess of GSNO (or NO solution) was used. The NO donor or any unwanted product was removed by centrifuging through Bio-Rad spin columns pre-equilibriated with 50 mM Tris·HCl pH 7.5, 200 mM EDTA.

10.14.4 S-Nitrosylation assay by the modified 'Biotin-switch method'

A modified version of the biotin-switch method[29] was also used as an 'indirect' method to detect protein S-nitrosylation. To a 100 µl S-nitrosylated tPHD2 (~40µM, prepared as above) was added 4 volumes of 'blocking buffer' (50 mM Tris·HCl pH 7.5, 2.5% SDS, 20 mM methylmethanethiosulfonate, MMTS) at 50 ºC for 20 minutes with frequent vortexing. The MMTS was then removed by buffer exchanging with Bio-Rad spin columns pre-equilibriated with 50 mM Tris·HCl pH 7.5. L-ascorbate was added to a final concentration of 1 mM and incubated for 1 hour at 25 ºC. The protein sample was then subjected to tryptic digestion and fragments were analysed for thiomethylation by MS (see below).

10.14.5 Tryptic digestion and LC-MS analyses

The S-nitrosylated/ S-thiomethylated tPHD2 samples was buffer exchanged into 0.1 M ammonium bicarbonate pH 7.8 using a NICK column (Amersham Biosciences) and the final concentration adjusted to 25 µM. The modified protein sample was then digested 'in solution' with porcine trypsin (Promega) in 20 mM ammonium bicarbonate at 1:10 protease-to-protein ratio; incubations were carried out for ~16 hours at 37ºC. An aliquot (10 µl) of the digest was then desalted and concentrated using in-tip reverse-phase resins (C-18 ZipTips, Millipore) as described, by the manufacturer prior to MS analyses.

An Agilent 1100 capillary LC system (Palo Alto, CA) was used to inject sample volumes of typically 8 µl onto a 5 µm Phenomenex Jupiter capillary LC column (150 × 0.5 mm). The fragment peptides were resolved with a linear gradient of 5-95% solvent B for 50 minutes, followed by a second gradient of 95-5% for 40 minutes at a flow rate of 15 µl/minute. Mobile phase solvent A was 0.1% aqueous formic acid and solvent B comprised 0.1% formic acid in 90% acetonitrile. The LC column eluate was subsequently analysed from 0 to 35 minute using positive ion electrospray with a Waters (Milford, MA) QTof micromass spectrometer. When needed for calibration, synthetic [Glu^1]-Fibrinopeptide B (EGVNDNEEGFFSAR, $[M+2H]^{2+}$ charge state, *m/z* = 785) (Sigma-Aldrich) peptide was

also run on this system. The following conditions were used: capillary voltage = 3000 V, cone voltage = 35 V, extraction cone voltage = 4 V, desolvation and nebulising nitrogen gas flow = 300 L per hour, desolvation temperature = 300°C, source temperature = 80°C, MCP detector voltage = 2400 V and Tof Flight tube voltage = 5630 V.

10.14.6 MS-MS analyses

For MS/MS analyses, the fragment peptides (prepared as above) were resolved by nano-ultraperformance liquid chromatography tandem mass spectrometry (nano-UPLC MS/MS) using a 75 µm inner diameter 25 cm length C18 nano-AcquityTM UPLCTM column (1.7 µm particle size; Waters) and a 90 min gradient of 2–45% solvent B (solvent A: 99.9% H2O, 0.1% formic acid; solvent B: 99.9% acetonitrile, 0.1% formic acid) on a Waters nanoAcquity UPLC system (final flow rate 250 nl·min^{-1}; 6000-7000 p.s.i.) coupled to a Q-TOF Premier tandem mass spectrometer (Waters). Data acquisition was performed in data-directed analysis mode (DDA; MS to MS/MS switching at precursor ion counts greater than 10 and MS/MS collision energy dependent on precursor ion mass and charge state). The mass accuracy of the raw data was calibrated using [Glu1]-Fibrinopeptide peptide (200 fmol·µl^{1} in 50% actenotrile, 0.1% formic acid; 700 nl·min^{-1} flow rate) which was infused into the mass spectrometer during sample analysis. MS and MS/MS data were calibrated at intervals of 30 s. MS/MS database searches were carried out using the MASCOT algorithm on an in-house server (CBRG, University of Oxford) using a UniProtKB/Swiss-Prot database allowing for S-nitrosylation (Cys) and methionine oxidation as modifications. MS/MS data were also inspected and assigned manually.

10.14.7 Electron paramagnetic resonance (EPR) experimental conditions

EPR experiments were performed by Dr E.G. Flashman (Chemistry Research Laboratory, Oxford). Continuous-wave EPR spectra of tPHD2 complexes were obtained at a temperature of 10 K and a frequency of 9.757 GHz with a modulation amplitude of 0.5 mT and a microwave power of 0.6 mW using a Bruker EMX spectrometer.

10.14.8 Crystallisation details

tPHD2.Fe(II).B crystals were obtained as described.[8] Crystals of tPHD2.Zn(II).B (tPHD2 (~20mg·ml^{-1} tPHD2apo + 1mM Zn-acetate + 1mM compound B) grew in sitting drops with well solution containing 30% PEG 4000 (w-v), 0.2 M sodium acetate, 0.1 M Tris·HCl pH 8.5. tPHD2.Fe(II).B crystals were exposed to NO gas (1 ml, density ~1.3 g/L) for an

hour whilst tPHD2.Zn(II).B crystals were soaked with 50 mM GSNO for 12 hours before being frozen in liquid nitrogen. tPHD2 C302A.Fe(II).compound C crystals were obtained under similar conditions as for the tPHD2.Fe(II).B complex.

10.14.9 Data collection and processing

A dataset for a tPHD2.Fe(II).B (NO exposed) crystal was collected at SRS Daresbury Station 14.1 on an ADSC Quantum 4 detector (Area Detector Systems Corp.). Data were integrated and scaled using HKL2000.[9] Datasets for the crystals of tPHD2.Zn(II).B (GSNO soaked) were collected at the Diamond beamline I04 with an ADSC Q315 3x3 CCD detector and were processed with MOSFLM and SCALA (CCP4 suite).[15] A dataset for a tPHD2 C302A.Fe(II).compound C crystal was obtained using an X8 Proteum (Bruker-AXS) equipped with a Microstar microfocus rotating anode X-ray generator and processed as described in Section 10.8.2.

10.14.10 Structure solution by molecular replacement and refinement

Structure solution and refinement was as described in Section 10.9.5.

10.14.11 In vivo assays

For the detection of endogenous S-nitrosylation of PHD2, a HeLa cell line was obtained from Dr M. E. Cockman (Centre for Cellular and Molecular Physiology, Oxford) that could stably overexpress FLAG-tagged PHD2.

10.14.11.1 Antibodies

The following primary antibiotics were used:

Primary Antibody	Conjugation	Source/ catalogue no.	Host	Immunogen	Working dilution
Anti-HIF-1α	horseradish peroxidise (HRP)	Novus Biologicals (NB100-123H)	Mouse	HIF-1α (432-528)	1:500
Anti-PHD2	-	Novus Biologicals (NB100-138)	Rabbit	PHD2 (350-426)	1:500
Anti-FLAG	horseradish peroxidise (HRP)	Sigma-Aldrich (A8592)	Mouse	-	1:1000
Anti-S-nitrosocysteine	-	Sigma-Aldrich (N5411)	Rabbit	S-nitrosylated cysteine-Keyhole Limpet Hemocyanin (KLH)	1:500 (or 1:50)

10.14.11.2 Transfection

Plasmid DNA was transfected into HeLa cells by using FuGENE 6 (Roche).[30]

10.14.11.3 Immunoprecipitation/purification of endogenous PHD2

To investigate the effect of S-nitrosylation, cells were treated with 1 mM GSNO or 1 mM L-cysteine as control; after an hour of incubation at 37°C, cells were washed with phosphate buffered saline (PBS). Cell extracts were prepared in IP+ buffer [20 mM Tris·HCl pH 7.4/0.1 M NaCl/5 mM $MgCl_2$/0.5% (v/v) Igepal CA-630/ 1× Complete protease inhibitor mixture (Roche)][30], and aliquots were used as INPUTs. For purification, FLAG-tagged PHD2 was immunoprecipitated with anti-FLAG conjugated agarose beads (Sigma-Aldrich). Beads were then washed repeatedly and finally eluted in 50 mM NH_4OH pH ~10, 0.5 mM EDTA at 4°C. The samples were then desalted by MeOH/$CHCl_3$ precipitation protocol followed by resuspension in 50 mM Tris·HCl pH 7.5.

10.14.11.4 Western blotting/immunoblotting

Proteins resolved on SDS-PAGE under non-reducing condition were electroblotted onto Immobilon-P™ polyvinylidene fluoride (PVDF, Millipore) membrane using a Mini Trans-Blot apparatus (Bio-Rad) according to the manufacturer's protocol. Briefly, a transfer cassette comprising the pre-equilibrated gel and PVDF membrane, sandwiched between Whatman 3MM filter paper and fibre pads (Scotch-Brite™) was assembled in transfer buffer (10 mM Tris·HCl pH 7.4, 0.1 M glycine, 10% methanol v/v) to minimise the trapping of air bubbles. The cassette was then placed in the transfer tank, immersed in pre-chilled transfer buffer and electrophoresed at 100 V for 1 hour at 4°C. Following electrophoresis, membranes were blocked (PBS containing 5% w/v dried skimmed milk) to prevent non-specific antibody binding and incubated for an hour at room temperature with rocking. Primary antibodies were appropriately diluted in blocking solution (Section 10.14.10.1) and applied to the membrane for overnight at room temperature. Following incubation with primary antibody, membranes were washed three times with PBS containing 0.1% Tween-20. Anti-mouse/ Anti-rabbit horse-radish peroxidise (HRP) secondary antibody (Dako) was suspended in blocking solution at a dilution of 1:2000 and incubated for an hour. Following washing of the excess secondary antibody, a chemiluminescent peroxidise substrate (Pierce) was added and bound HRP activity was visualised by chemiluminescence (ECL+Plus™, Amersham) using X-OMAT LS film.

Protein loading was verified after immunoblotting by staining the PVDF membrane with SDS-PAGE stain.

10.15 Experimental details for Chapter 9

10.15.1 Construction of plasmids

Using *phd3*/pET-28a(+) DNA construct as a template, a C-terminally His-tagged *phd3* was cloned into a pET-29b(+) vector (for primers see Appendix 12). In addition, sequence encoding for a thrombin cleavage site (LVPRGS) was engineered at the end of *phd3* gene within the reading frame. Primer oligonucleotides (Sigma-Genosys Ltd.) flanking the *phd3* sequence were designed to incorporate *Nde* I and *Xho I* restriction enzyme sites at 5' and 3' ends including an upstream thrombin cleavage site immediately before the *Xho I*.

An N-terminally truncated PHD3 variant (13 - 239) was sub-cloned into a pET-28a(+) vector using *Nde* I and *Xho I* restriction sites. C-Terminal truncations were amplified using either *phd3*/pET-28a(+) (*Nde I*/*EcoR I* digested) or $phd3_{13\text{-}239}$/pET-28a(+) (*Nde I*/*Xho I* digested). The two deletion variants (Δ71R or Δ72H) and a PHD3-NH C42Q were generated using *phd3*/pET-28a(+) as the template. The remainder of the point variants with C-terminal His_6-tag were constructed by site-directed mutagenesis using *Nde I*/*Xho I* digested *phd3*/pET-29b(+) template. The multi-variants, PHD3-CH MM-1, MM-2, MM-3 were produced by Stratagene's QuikChange multi site-directed mutagenesis protocol using PHD3-CH C42Q/C52D construct with multiple rounds of PCR. Mutagenic PCR products were digested with Dpn I before being transformed into competent cells (Section 10.2.10). All constructs were verified by DNA sequencing (Section 10.2.14).

10.15.2 Co-transformation for co-expression

Transformations for co-expressions were performed in a similar way to those for single plasmid transformations (Section 10.2.12) with the exception that 1 µl of each of the plasmids with different antibiotic resistance markers were mixed with the competent cells before heat shock treatment. The transformed competent cells were streaked onto agar media containing appropriate antibiotics. Plasmids for N-terminal GST fusion proteins (HIF-$1\alpha_{530\text{-}652}$ and HIF-$1\alpha_{530\text{-}826}$) were obtained from Dr K.S. Hewitson (Chemistry Research Laboratory, Oxford) and those for chaperones from Takara Bio Inc.

Plasmids used for the co-transformation are as follows:

Construct	Vector	Promoter	Inducer	Resistant Marker
His-PHD3	pET-28a(+) or pET-29b(+)	*T7 lac*	IPTG	Kanamycin
GST-HIF-1$\alpha_{530-652}$ or GST-HIF-1$\alpha_{530-826}$	pGEX-6p-1	*tac*	IPTG	Ampicillin
DnaK–DnaJ–GrpE	pKJE7	*araB*	L-Arabinose	Chloramphenicol
GroEL–GroES	pGro7	*araB*	L-Arabinose	Chloramphenicol

10.15.3 Purification of PHD3 wt and variants

Following transformation into *E. coli* BL21 (DE3) cells, the variant/ wildtype PHD3 protein expression was induced with 0.5 mM IPTG (6 hours at 22°C). In addition, cells were also supplemented with 5 μM $FeSO_4$ during induction. Cells were harvested and lysed by sonication in lysis-equilibriation-wash (LEW) buffer containing 20 mM Tris·HCl pH 7.0, 0.5 M NaCl, 10% glycerol and 0.1% β-mercaptoethanol. In some cases, the cell extract was supplemented with EDTA-free protease inhibitors cocktail, Roche. The target protein (< 5% total soluble extract) was purified by immobilised Ni^{2+} affinity chromatography in elution buffer (LEW plus 0.5 M imidazole). The final purification was carried out using a gel filtration column equilibriated with elution buffer (50 mM Tris·HCl pH 7.5, 0.2 M NaCl, 10% glycerol and 2 mM TCEP).

10.15.4 Non-denaturing MS and activity analyses

MS analysis under non-denaturing conditions and activity analysis were carried out as described in Section 10.7.

10.15.5 CD measurements

CD measurements were carried out as described in Section 10.5.5.

10.15.6 Thermal stability assay

Thermal stability assays on PHD3 were performed using SYPRO orange (Invitrogen) as a fluorophore at the Structural Genomics Consortium (SGC), Oxford. A real-time PCR device (Mx3005p from Stratagene) was used to monitor protein unfolding (i.e. increase in fluorescence). PHD3-CH C42Q/C52D (10 μM or 25–100 $\mu g \cdot ml^{-1}$) in 50 mM Tris·HCl pH 7.5, 0.2 M NaCl, 10% glycerol and 2 mM TCEP was incubated with the ligands (10 μM –

1 mM in DMSO) in a reaction volume of 20–25 µl in 96-well PCR microplates (ABGene) Before initiating a full screen, an initial scan was performed using variable concentration of PHD3-CH C42Q/C52D to determine an optimal concentration for detectable fluorescent signals. The samples were heated at 1°C per minute, from 25°C to 75°C with fluorescence intensity measured every 1–3°C. Fluorescence intensities were plotted as a function of temperature by using an SGC developed software package.[31]

10.15.7 Crystallisation

Crystallisation (96-well, Greiner) plates for PHD3-CH (1 $mg \cdot ml^{-1}$), PHD3-CH C42Q (1.5 $mg \cdot ml^{-1}$) and PHD3-CH C42Q/C52D (5 $mg \cdot ml^{-1}$/ ~13 $mg \cdot ml^{-1}$/ ~26 $mg \cdot ml^{-1}$) variants were set up using the sitting-drop vapour-diffusion at the OPPF, Oxford (Section 10.8.1). Screens included Crystal Screen, Crystal Screen 2 and PEG/Ion (Hampton Research). Crystallisation drops for PHD3-CH C42Q/C52D (~13 $mg \cdot ml^{-1}$/ ~26 $mg \cdot ml^{-1}$) contained 100 nl protein and 100 nl reservoir solution. For samples with a low concentration, the enzyme to reservoir ratio was set as 3:1. Attempts were made to reproduce/ optimise the initial crystallisation hits using VDX plates by sitting/hanging drop vapour diffusion at either room temperature or at 4°C.

10.16 References

1. McNeill, L.A., Flashman, E., Buck, M.R., Hewitson, K.S., Clifton, I.J., Jeschke, G., Claridge, T.D., Ehrismann, D., Oldham, N.J. & Schofield, C.J. Hypoxia-inducible factor prolyl hydroxylase 2 has a high affinity for ferrous iron and 2-oxoglutarate. *Mol Biosyst* **1**, 321-4 (2005).
2. Hewitson, K.S., McNeill, L.A., Riordan, M.V., Tian, Y.M., Bullock, A.N., Welford, R.W., Elkins, J.M., Oldham, N.J., Bhattacharya, S., Gleadle, J.M., Ratcliffe, P.J., Pugh, C.W. & Schofield, C.J. Hypoxia-inducible factor (HIF) asparagine hydroxylase is identical to factor inhibiting HIF (FIH) and is related to the cupin structural family. *J Biol Chem* **277**, 26351-5 (2002).
3. Sabourin, P.J. & Bieber, L.L. Purification and characterization of an α-ketoisocaproate oxygenase of rat liver. *J. Biol. Chem.* **257**, 7460-7467 (1982).
4. Jaakkola, P., Mole, D.R., Tian, Y.M., Wilson, M.I., Gielbert, J., Gaskell, S.J., Kriegsheim, A., Hebestreit, H.F., Mukherji, M., Schofield, C.J., Maxwell, P.H., Pugh, C.W. & Ratcliffe, P.J. Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. *Science* **292**, 468-72 (2001).
5. Geerlof, A., Brown, J., Coutard, B., Egloff, M.P., Enguita, F.J., Fogg, M.J., Gilbert, R.J., Groves, M.R., Haouz, A., Nettleship, J.E., Nordlund, P., Owens, R.J., Ruff, M., Sainsbury, S., Svergun, D.I. & Wilmanns, M. The impact of protein characterization in structural proteomics. *Acta Crystallogr D Biol Crystallogr* **62**, 1125-36 (2006).
6. Pantoliano, M.W., Petrella, E.C., Kwasnoski, J.D., Lobanov, V.S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, B.A., Lane, P. & Salemme, F.R. High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J Biomol Screen* **6**, 429-40 (2001).
7. Walter, T.S., Diprose, J.M., Mayo, C.J., Siebold, C., Pickford, M.G., Carter, L., Sutton, G.C., Berrow, N.S., Brown, J., Berry, I.M., Stewart-Jones, G.B., Grimes, J.M., Stammers, D.K., Esnouf, R.M., Jones, E.Y., Owens, R.J., Stuart, D.I. & Harlos, K. A procedure for setting up high-throughput nanolitre crystallization experiments. Crystallization workflow for initial screening, automated storage, imaging and optimization. *Acta Crystallogr. D Biol. Crystallogr.* **61**, 651-657 (2005).
8. McDonough, M.A., Li, V., Flashman, E., Chowdhury, R., Mohr, C., Lienard, B.M., Zondlo, J., Oldham, N.J., Clifton, I.J., Lewis, J., McNeill, L.A., Kurzeja, R.J., Hewitson, K.S., Yang, E.,

Jordan, S., Syed, R.S. & Schofield, C.J. Cellular oxygen sensing: Crystal structure of hypoxia-inducible factor prolyl hydroxylase (PHD2). *Proc Natl Acad Sci U S A* **103**, 9814-9 (2006).

9. Otwinowski, Z., Minor, W. & Jr., C.W.C. Processing of X-ray diffraction data collected in oscillation mode. in *Methods in Enzymology*, Vol. **276**, 307-326 (Academic Press, 1997).
10. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. & Read, R.J. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674 (2007).
11. Schuttelkopf, A.W. & van Aalten, D.M. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 1355-1363 (2004).
12. Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. & Warren, G.L. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**, 905-21 (1998).
13. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32 (2004).
14. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283-291 (1993).
15. Bailey, S. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760-763 (1994).
16. Sali, A. & Blundell, T.L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779-815 (1993).
17. Kale, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N., Philips, J., Shinozaki, A., Varadarajan, K. & Schulten, K. Molecular dynamics programs design. NAMD2: greater suitability for parallel molecular dynamics. *J. Comput. Phys.* **151**, 283-312 (1999).
18. MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher, W.E., Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D. & Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586-3616 (1998).
19. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. & Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926-935 (1983).
20. Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H. & Pedersen, L.G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577-8593 (1995).
21. Feller, S.E., Zhang, Y., Pastor, R.W. & Brooks, B.R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **103**, 4613-4621 (1995).
22. Martyna, G.J., Tobias, D.J. & Klein, M.L. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **101**, 4177-4189 (1994).
23. Tuckerman, J.R., Zhao, Y., Hewitson, K.S., Tian, Y.M., Pugh, C.W., Ratcliffe, P.J. & Mole, D.R. Determination and comparison of specific activity of the HIF-prolyl hydroxylases. *FEBS Lett.* **576**, 145-150 (2004).
24. Frohn, M., Viswanadhan, V., Pickrell, A.J., Golden, J.E., Muller, K.M., Burli, R.W., Biddlecome, G., Yoder, S.C., Rogers, N., Dao, J.H., Hungate, R. & Allen, J.R. Structure-guided design of substituted aza-benzimidazoles as potent hypoxia inducible factor-1alpha prolyl hydroxylase-2 inhibitors. *Bioorg Med Chem Lett* (2008).
25. Hon, W.C., Wilson, M.I., Harlos, K., Claridge, T.D., Schofield, C.J., Pugh, C.W., Maxwell, P.H., Ratcliffe, P.J., Stuart, D.I. & Jones, E.Y. Structural basis for the recognition of hydroxyproline in HIF-1 alpha by pVHL. *Nature* **417**, 975-8 (2002).
26. Walter, T.S., Meier, C., Assenberg, R., Au, K.F., Ren, J., Verma, A., Nettleship, J.E., Owens, R.J., Stuart, D.I. & Grimes, J.M. Lysine methylation as a routine rescue strategy for protein crystallization. *Structure* **14**, 1617-22 (2006).
27. Mirza, U.A., Chait, B.T. & Lander, H.M. Monitoring reactions of nitric oxide with peptides and proteins by electrospray ionization-mass spectrometry. *J Biol Chem* **270**, 17185-8 (1995).
28. Archer, S. Measurement of nitric oxide in biological models. *Faseb J* **7**, 349-60 (1993).
29. Jaffrey, S.R., Erdjument-Bromage, H., Ferris, C.D., Tempst, P. & Snyder, S.H. Protein S-nitrosylation: a physiological signal for neuronal nitric oxide. *Nat Cell Biol* **3**, 193-7 (2001).
30. Cockman, M.E., Lancaster, D.E., Stolze, I.P., Hewitson, K.S., McDonough, M.A., Coleman, M.L., Coles, C.H., Yu, X., Hay, R.T., Ley, S.C., Pugh, C.W., Oldham, N.J., Masson, N., Schofield, C.J. & Ratcliffe, P.J. Posttranslational hydroxylation of ankyrin repeats in IkappaB proteins by the hypoxia-inducible factor (HIF) asparaginyl hydroxylase, factor inhibiting HIF (FIH). *Proc Natl Acad Sci U S A* **103**, 14767-72 (2006).
31. Vedadi, M., Niesen, F.H., Allali-Hassani, A., Fedorov, O.Y., Finerty, P.J., Jr., Wasney, G.A., Yeung, R., Arrowsmith, C., Ball, L.J., Berglund, H., Hui, R., Marsden, B.D., Nordlund, P.,

Sundstrom, M., Weigelt, J. & Edwards, A.M. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc Natl Acad Sci U S A* **103**, 15835-40 (2006).

Appendix 1

Summary of the PHD2 constructs/ proteins produced in the current study.

Construct/Protein	Expression Vector	Purification method	Observed mass (Da)	Calculated mass (Da)
$PHD2_{181-426}$ or tPHD2	pET24a	CA	27642	27643
	pET28a	IMAC	28054	28055
Variants relating to homotrimeric crystal packing				
T236A	pET28a	IMAC	28024	28025
D394A	pET28a	IMAC	28010	28011
S289K	pET28a	IMAC	28095	28096
R398A	pET28a	IMAC	27969	27970
Variants relating to β2/β3 loop				
tPHD2Δloop	pET24a	CA	26422	26425
tPHD2-PHD3 (cloned and purified by Dr E. Flashman)	pET24a	CA	27700	27701
R252A	pET28a	IMAC	27969	27970
D254A	pET28a	IMAC	28010	28011
D254K	pET28a	IMAC	28067	28068
R252C/D254C	pET28a	IMAC	27990	27990
Clinically observed and related mutations				
P317R	pET28a	IMAC	28113	28114
P317E	pET28a	IMAC	28086	28087
R371H	pET24a	CA	27623	27624
Variants relating to 2OG and Fe(II) interacting residues				
M299V	pET28a	IMAC	28023	28023
M299E	pET28a	IMAC	28057	28053
Y303F	pET28a	IMAC	28038	28039
Y303A	pET28a	IMAC	27962	27963
D315G (cloned by E. Bagg)	pET28a	IMAC	27996	27997
T325A	pET28a	IMAC	28024	28025
Y329F	pET28a	IMAC	28038	28039
F366Y	pET28a	IMAC	28070	28071
R383A	pET28a	IMAC	27969	27970
T387A	pET28a	IMAC	28024	28025
Variants relating to LXXLAP binding residues				
S242G	pET28a	IMAC	28025	28022
K244R	pET28a	IMAC	28083	28081
I251L	pET28a	IMAC	28054	28055
I251G	pET28a	IMAC	27999	28000
K244R/I251L	pET24a	CA	27668	27671
S242G/K244R/I251L	pET24a	CA	27640	27641
S242G/K244R/I251L/P264E	pET24a	CA	27670	27673
S242G/K244R/I251L/A228Q/K234A/T236R /P264E	pET24a	CA	27727	27728

Variants of other residues involved in substrate binding/ selectivity

R281L	pET28a	IMAC	28011	28012
I280V/I292V/R281L	pET28a	IMAC	27983	27984
I280C	pET28a	IMAC	28044	28045
K297I (cloned and purified by J. Mecinovic)	pET28a	IMAC	28040	28040
Y310F	pET28a	IMAC	28038	28039
R396A	pET28a	IMAC	27969	27970

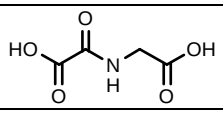
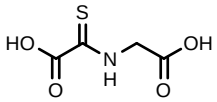
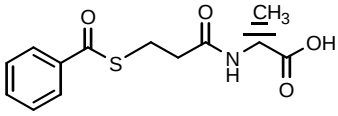
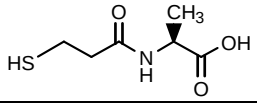
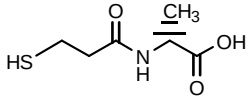
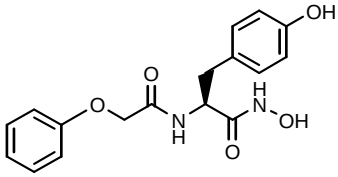
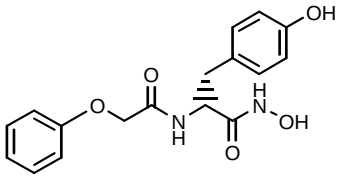
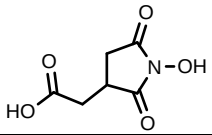
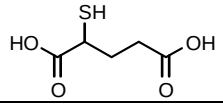
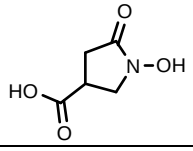
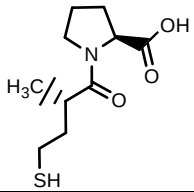
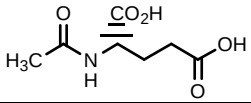
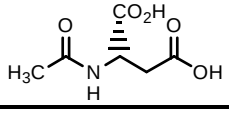
Cysteine variants

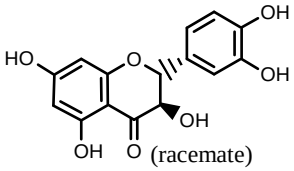
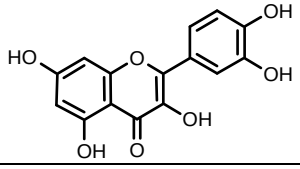
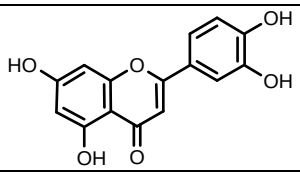
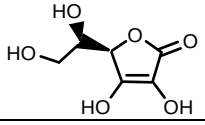
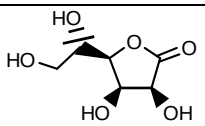
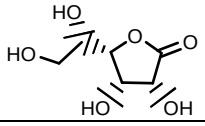
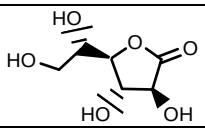
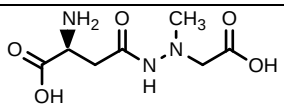
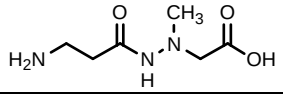
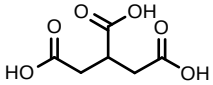
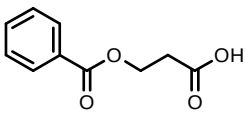
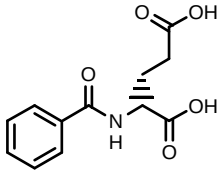
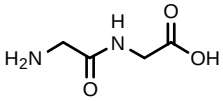
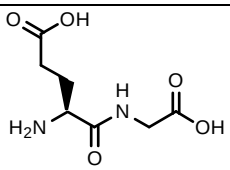
C201A	pET28a	IMAC	28022	28023
C208A	pET28a	IMAC	28022	28023
C266A	pET28a	IMAC	28022	28023
C283A	pET28a	IMAC	28022	28023
C302A	pET28a	IMAC	28022	28023
C323A	pET28a	IMAC	28022	28023
C326A	pET28a	IMAC	28022	28023

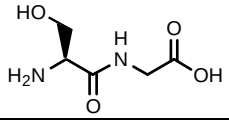
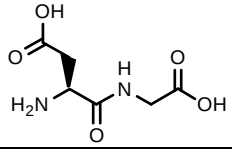
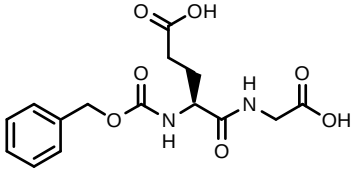
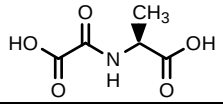
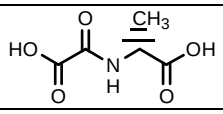
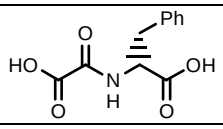
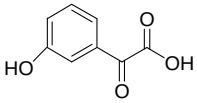
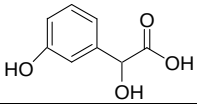
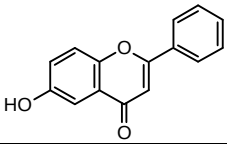
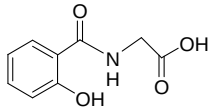
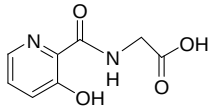
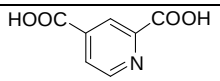
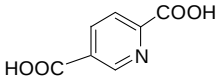
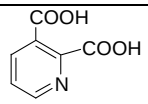
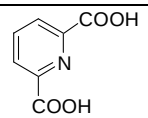
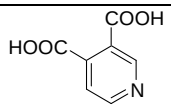
* Average protein yield was 1-1.5 mg of protein per g of wet cell pellet when purified by Ni^{2+} immobilized metal affinity chromatography (IMAC, with ≥95% homogeneity as determined by SDS-PAGE) and 2-2.5 mg of protein per g of wet cell pellet when purified by cation exchange (CA, with ≥90% homogeneity as determined by SDS-PAGE). See Chapter 10 for experimental details.

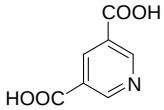
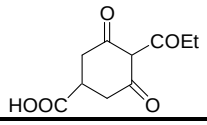
Appendix 2

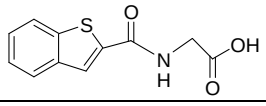
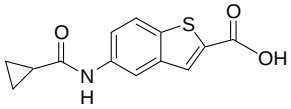
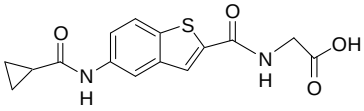
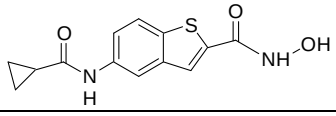
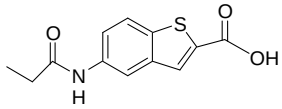
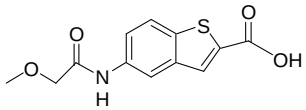
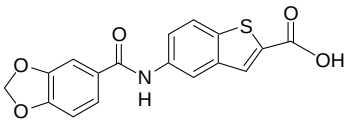
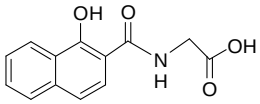
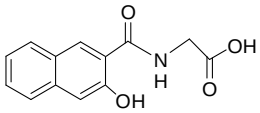
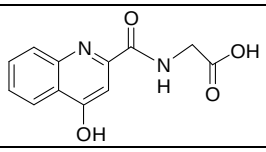
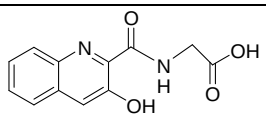
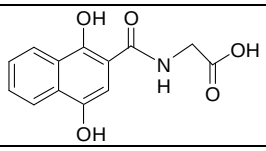
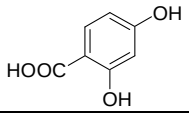
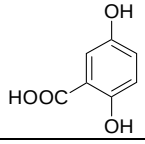
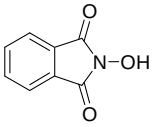
Compounds used in the thermal stability (Thermofluor) assays and/or MS studies.

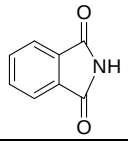
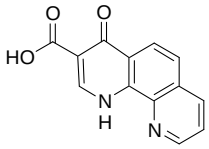
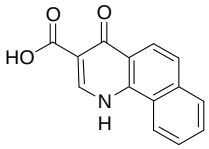
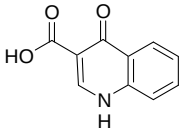
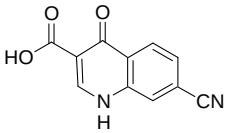
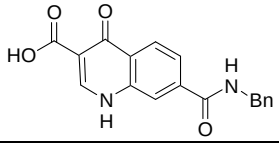
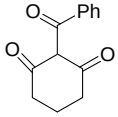
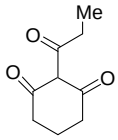
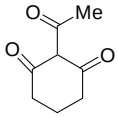
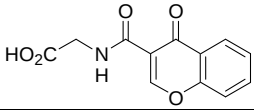
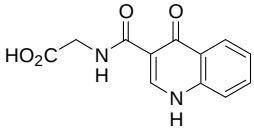
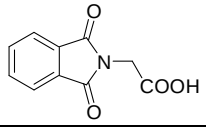
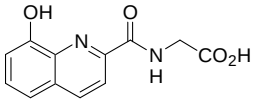
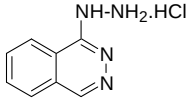
Compound	Structure	Molecular Formula	Molecular Weight
NOG		$C_4H_5NO_5$	147.09
IS3		$C_4H_5NO_4S$	163.15
IS21		$C_{13}H_{15}NO_4S$	281.23
IS37		$C_6H_{11}NO_3S$	177.22
IS38		$C_6H_{11}NO_3S$	177.22
IS44		$C_{17}H_{18}N_2O_5$	330.34
IS45		$C_{17}H_{18}N_2O_5$	330.34
IS52		$C_6H_7NO_5$	173.12
IS54		$C_5H_8O_4S$	164.18
ANU2		$C_5H_7NO_4$	145.11
C2		$C_{10}H_{17}NO_3S$	231.31
C6		$C_7H_{11}NO_5$	189.17
C7		$C_6H_9NO_5$	175.14

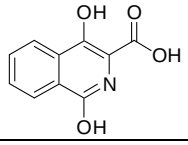
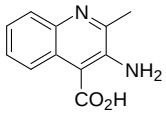
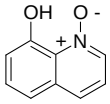
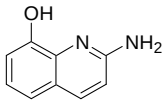
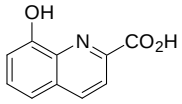
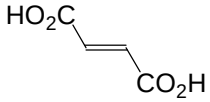
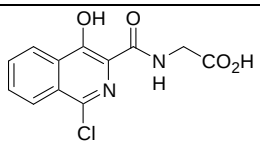
R1	 (racemate)	$C_{15}H_{12}O_7$	304.25
R2		$C_{15}H_{10}O_7$	302.24
R3		$C_{15}H_{12}O_7$	286.24
R35		$C_6H_8O_6$	176.12
R32		$C_6H_{10}O_6$	178.14
R34		$C_6H_{10}O_6$	178.14
R33		$C_6H_{10}O_6$	178.14
HKS20		$C_7H_{13}N_3O_5$	219.20
HKS71		$C_6H_{13}N_3O_3$	175.19
C13		$C_6H_8O_6$	176.12
IS84		$C_{10}H_{10}O_4$	251.24
IS90		$C_{12}H_{13}NO_5$	251.24
C17		$C_4H_8N_2O_3$	132.12
C21		$C_7H_{12}N_2O_5$	204.18

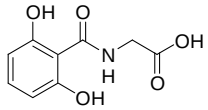
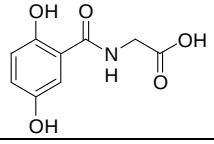
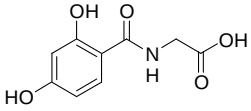
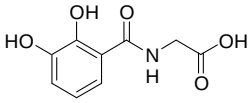
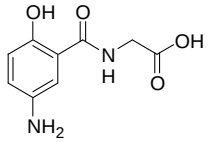
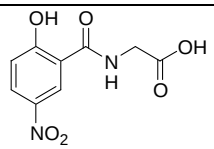
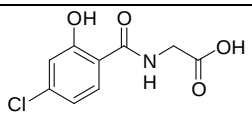
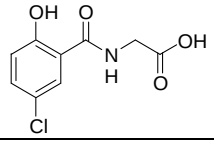
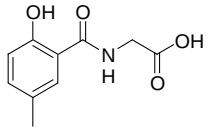
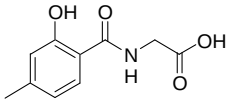
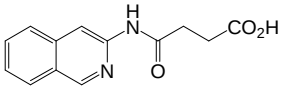
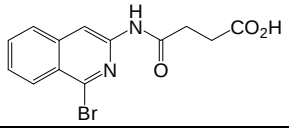
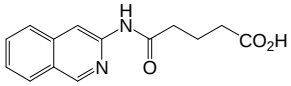
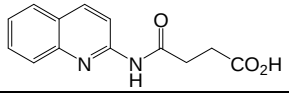
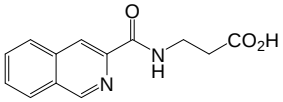
C23		$C_5H_{10}N_2O_4$	162.14
C18		$C_6H_{10}N_2O_5$	190.15
C22		$C_{15}H_{18}N_2O_7$	338.31
IS70		$C_5H_7NO_5$	161.11
IS71		$C_5H_7NO_5$	161.11
NOFD		$C_{11}H_{11}NO_5$	237.21
ACG-29		$C_8H_6O_4$	166.13
3-Hydroxy mandelate		$C_8H_8O_4$	168.15
6-hydroxy flavone		$C_{15}H_{10}O_3$	238.24
ACG-179a		$C_9H_9NO_4$	195.17
ACG-194b		$C_8H_8N_2O_4$	196.16
2,4-PDCA		$C_7H_5NO_4$	167.12
2,5-PDCA		$C_7H_5NO_4$	167.12
2,3-PDCA		$C_7H_5NO_4$	167.12
2,6-PDCA		$C_7H_5NO_4$	167.12
3,4-PDCA		$C_7H_5NO_4$	167.12

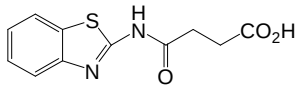
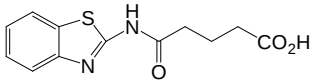
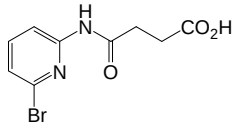
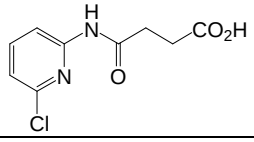
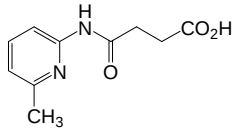
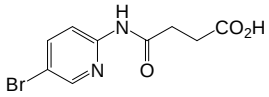
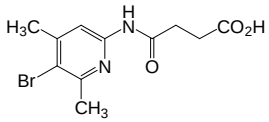
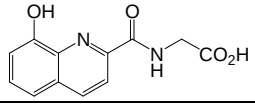
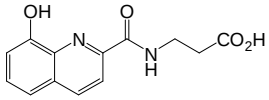
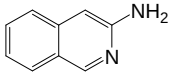
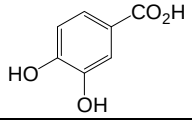
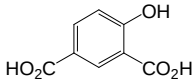
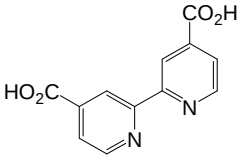
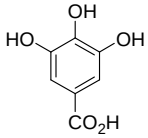
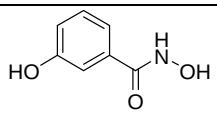
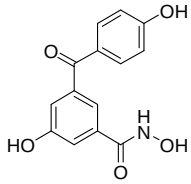
3,5-PDCA		$C_7H_5NO_4$	167.12
Prohexadione		$C_{10}H_{12}O_6$	212.20

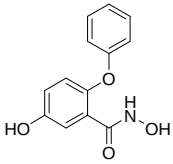
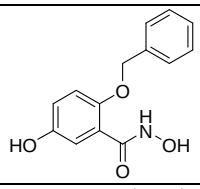
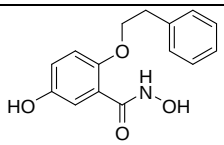
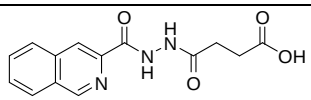
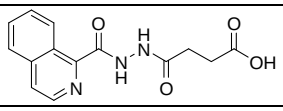
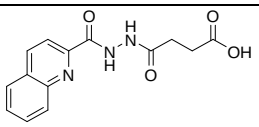
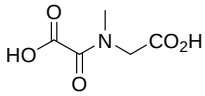
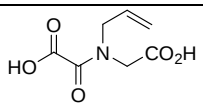
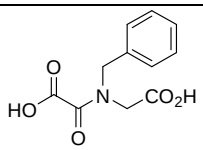
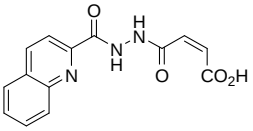
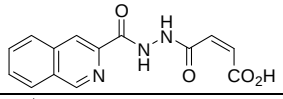
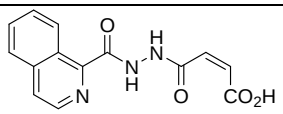
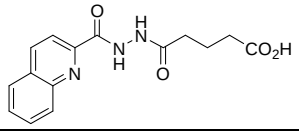
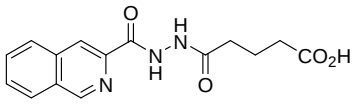
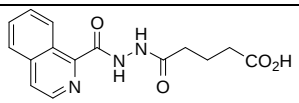
ACG-106a		$C_{11}H_9NO_3S$	235.26
ACG-109a		$C_{13}H_{11}NO_3S$	261.30
ACG-113a		$C_{15}H_{14}N_2O_4S$	318.35
ACG-114c		$C_{13}H_{12}N_2O_3S$	276.31
ACG-137a		$C_{12}H_{11}NO_3S$	249.29
ACG-144a		$C_{12}H_{11}NO_4S$	265.29
ACG-138a		$C_{17}H_{11}NO_5S$	314.34
ACG-148a		$C_{13}H_{11}NO_4$	245.23
ACG-181a		$C_{13}H_{11}NO_4$	245.23
ACG-180a		$C_{12}H_{10}N_2O_4$	246.22
ACG-191a		$C_{12}H_{10}N_2O_4$	246.22
TB088		$C_{13}H_{11}NO_5$	261.23
2,4-DHB		$C_7H_6O_4$	154.12
2,5-DHB		$C_7H_6O_4$	154.12
N-Hydroxy phthalimide		$C_8H_5NO_3$	163.03

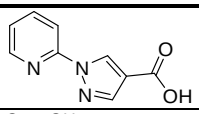
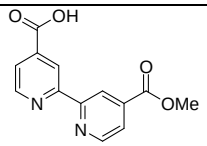
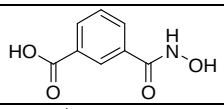
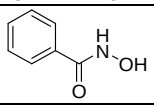
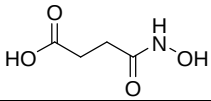
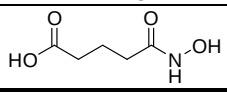
Phthalimide		$C_8H_5NO_2$	147.13
BB103		$C_{13}H_8N_2O_3$	240.05
BB105		$C_{13}H_8N_2O_3$	239.23
BB110H		$C_{10}H_7NO_3$	189.17
BB169		$C_{11}H_6N_2O_3$	214.18
BB204		$C_{18}H_{14}N_2O_4$	322.31
BB210		$C_{13}H_{12}O_3$	216.23
BB212		$C_9H_{12}O_3$	168.19
BB209		$C_8H_{10}O_3$	154.16
BB222		$C_{12}H_9NO_5$	247.20
BB224		$C_{12}H_{10}N_2O_4$	246.22
NGP		$C_{10}H_7NO_4$	205.17
BB237		$C_{12}H_{10}N_2O_4$	246.22
Hydralazine		$C_8H_8N_4.HCl$	196.64

BB236		$C_{10}H_7NO_4$	205.17
3A2MQC		$C_{11}H_{10}N_2O_2$	202.21
HQNO		$C_9H_7NO_2$	161
28AHQ		$C_9H_8N_2O$	160.17
28CHQ		$C_{10}H_7NO_3$	189.17
Fumarate		$C_4H_4O_4$	116.1
ACG263		$C_{12}H_9ClN_2O_4$	280.66

ACG-198a **SG00241**		$C_9H_9NO_5$	211.17
ACG-257b **SG00242**		$C_9H_9NO_5$	211.17
ACG-237a **SG00243**		$C_9H_9NO_5$	211.17
ACG-236a **SG00244**		$C_9H_9NO_5$	211.17
ACG-255d **SG00245**		$C_9H_{10}N_2O_4$	210.19
ACG-253a **SG00246**		$C_9H_8N_2O_6$	240.17
ACG-227a **SG00247**		$C_9H_8ClNO_4$	229.62
ACG-223a **SG00248**		$C_9H_8ClNO_4$	229.62
ACG-226a **SG00249**		$C_{10}H_{11}NO_4$	209.20
ACG-228a **SG00250**		$C_{10}H_{11}NO_4$	209.20
BB-287		$C_{13}H_{12}N_2O_3$	244.3
BB-294		$C_{13}H_{11}BrN_2O_3$	323.1
BB-315		$C_{14}H_{14}N_2O_3$	258.3
BB-286		$C_{13}H_{12}N_2O_3$	244.3
BB-316		$C_{13}H_{12}N_2O_3$	244.3

BB-324		$C_{11}H_{10}N_2O_3S$	250.3
BB-323		$C_{12}H_{12}N_2O_3S$	264.1
BB-307		$C_9H_9BrN_2O_3$	273.1
BB-308		$C_9H_9ClN_2O_3$	228.6
BB-317		$C_{10}H_{12}N_2O_3$	208.2
BB-318		$C_9H_9BrN_2O_3$	273.1
BB-328		$C_{11}H_{13}BrN_2O_3$	301.1
BB-237		$C_{12}H_{10}N_2O_4$	246.2
BB-270		$C_{13}H_{12}N_2O_4$	260.3
3-AIQ		$C_9H_8N_2$	144.2
3,4-DHBA		$C_7H_6O_4$	154.1
2-HIPA		$C_8H_6O_6$	198.1
2,2'-ByP-4,4'-DCA		$C_{12}H_8N_2O_4$	244.2
Gallate		$C_7H_6O_5$	170.1
LBE332		$C_7H_7NO_5$	153.1
LBE153		$C_{14}H_{11}NO_5$	273.2

LBE314		$C_{13}H_{11}NO_4$	245.2
LBE329		$C_{14}H_{13}NO_4$	259.3
LBE526		$C_{15}H_{15}NO_4$	273.3
JM87		$C_{14}H_{13}N_3O_4$	287.3
JM110		$C_{14}H_{13}N_3O_4$	287.3
JM104		$C_{14}H_{13}N_3O_4$	287.3
HM-Oxalo sarcosine		$C_5H_7NO_5$	161.1
HM-N-allylNOG		$C_7H_9NO_5$	187.2
HM-N-phenylNOG		$C_{11}H_{11}NO_5$	237.2
JM106		$C_{14}H_{11}N_3O_4$	285.3
JM134		$C_{14}H_{11}N_3O_4$	285.3
JM124		$C_{14}H_{11}N_3O_4$	285.3
JM103		$C_{15}H_{15}N_3O_4$	301.3
JM102		$C_{15}H_{15}N_3O_4$	301.3
JM109		$C_{15}H_{15}N_3O_4$	301.3

BMRL08012		$C_9H_7N_3O_2$	189.2
BMRL08014		$C_{13}H_{10}N_2O_4$	258.2
LBE096		$C_8H_7NO_4$	181.2
BHA		$C_7H_7NO_2$	137.1
LBE043		$C_4H_7NO_4$	133.1
LBE060		$C_5H_9NO_4$	147.1

Appendix 3

Hydrogen bonds and hydrophobic interactions between tPHD2 and CODD as apparently observed in the tPHD2.CODD complex structure.

Substrate residue atom(s)	Protein residue atom(s)	Distance (Å)[f]
Leu 559 Cδ1	Val 314 Cγ2	3.9
Leu 559 Cδ1	Pro 317 Cβ	3.5
Leu 559 Cδ1	Arg 370 Cζ	3.3
Leu 559 Cβ	Pro 317 Cβ	3.5
Leu 559 Cβ	Pro 317 Cγ	3.8
Glu 560 Cγ	Ser 242 Cβ	3.3
Glu 560 Oε1	Ser 242 Oγ	3.6
Glu 560 O	Ser 242 NH	3.4
Met 561 Cβ	Lys 244 Cβ	3.8
Met 561 Sδ	Lys 244 Cδ	3.3
Met 561 Cε	Lys 244 Cε	3.3
Leu 562 Cδ2	Ile 251 Cδ1	3.4
Leu 562 Cδ2	Ile 251 Cγ2	3.7
Leu 562 Cδ2	Val 311 Cγ2	4.0
Leu 562 Cβ	His 313 Cβ	3.5
Leu 562 Cα	Ile 251 Cγ2	3.7
Leu 562 O	Tyr 310 OH	2.5
Ala 563 Cβ	Pro 317 Cδ	3.5
Ala 563 O	Val 241 NH	2.8
Ala 563 Cβ	Val 241 Cγ2	3.8
Pro 564 NH	Tyr 310 OH	2.9
Pro 564 Cδ	Asp 315 C	3.9
Pro 564 Cδ	Asp 315 Cα	3.3
Pro 564 Cδ	His 313 Cδ2	3.5
Pro 564 Cγ	Trp 389 Cζ2	3.4
Pro 564 Cβ	Trp 389 CH2	3.7
Pro 564 O	Arg 322 NH1	2.7
Pro 564 O	Arg 322 NH2	2.9
Tyr 565 N	Gln 239 O	2.7
Tyr 565 Cβ	Gln 239 C	3.1
Tyr 565 Cβ	Leu 240 C	3.3
Tyr 565 Cβ	Gln 239 Cβ	3.7
*Tyr 565 O	H_2O-Arg 252 NH1	2.3 & 2.7
Ile 566 Cγ2	Trp 258 CH2	3.8
Ile 566 Cγ2	Trp 389 Cζ3	3.6
Ile 566 Cγ2	Trp 389 Cε3	3.6
Ile 566 Cγ2	Arg 322 Cζ	3.6
Ile 566 Cδ1	Phe 391 Cε2	3.5
Ile 566 Cδ1	Phe 391 Cζ	3.6
Pro 567 Cδ	Trp 258 CH2	3.9
Pro 567 O	Arg 396 NH2	2.9

Substrate residue atom(s)	Protein residue atom(s)	Distance (Å)[f]
Met 568 Cβ	Tyr 403 Cδ2	3.2
Met 568 Cβ	Tyr 403 Cε2	3.2
Met 568 Cβ	Tyr 403 Cζ	3.8
Met 568 Cγ	Tyr 403 Cδ2	3.7
Met 568 Cγ	Tyr 403 Cγ	3.8
Met 568 Cε	Tyr 403 Cβ	3.7
Met 568 Cε	Leu 404 Cδ2	3.7
Met 568 Sδ	Arg 396 NH1	3.6
Met 568 O	Arg 396 NH2	3.4
Asp 570 Cγ	Lys 297 Cε	3.7
Asp 571 Oδ2	Arg 396 NH2	2.8
Asp 571 Oδ1	Arg 396 Nε	2.6
Asp 571 Oδ1	Lys 400 Nζ	3.0
Phe 572 NH	Arg 295 O	2.6
Phe 572 O	Arg 295 NH	2.8
[§]Phe 572 (π)	Arg 295 NH2/NH1/Cζ/Nε	3.7
Leu 574 NH	Asn 293 O	2.9
Leu 574 Cβ	Asn 293 C	3.7
Leu 574 Cγ	Gly 294 Cα	3.5
Leu 574 Cδ2	Gly 294 Cα	3.4
Leu 574 Cδ2	Arg 295 Cγ	3.7
Leu 574 Cδ1	Tyr 390 Cδ1	3.8
Leu 574 Cδ1	Tyr 390 Cε1	3.7
Leu 574 Cδ1	Ile 292 Cγ2	3.9
Leu 574 Cδ1	Ile 280 Cγ2	3.9
Leu 574 O	Arg 281 Nε	3.0

* A hydrogen bond network between Tyr 565 O-H_2O-Arg 252 NH1.

[f] A 4.0 Å cut-off was used for hydrophobic interactions.

[§] Cation-π interaction (distance was determined using the programme PROTEIN INTERACTIONS CALCULATOR)[1] with E(es) = -4.24 kcal/mol and E(vdw) = -2.50 kcal/mol (electrostatic and van der Waals bond energies calculated by CAPTURE)[2]

References

1. Sathyapriya, R. & Vishveshwara, S. Interaction of DNA with clusters of amino acids in proteins. *Nucleic Acids Res.* **32**, 4109-4118 (2004).
2. Gallivan, J.P. & Dougherty, D.A. Cation-π interactions in structural biology. *Proc. Natl. Acad. Sci. USA* **96**, 9459-9464 (1999).

Appendix 4

Table 1. Summary of results for 2OG turnover and hydroxylation assays with NODD and CODD with tPHD2 and tPHD2 variants (using an enzyme-substrate ratio of 1:25).

Wildtype/ Variant tPHD2	Equivalent residues in PHD1-3 isoforms			% 2OG turned over relative to tPHD2*		% Hydroxylation	
	PHD1	PHD2	PHD3	CODD	NODD	CODD	NODD
tPHD2				100.0	100.0	98.5 ± 0.7	78.5 ± 0.7
1. Variants relating to homotrimeric crystal packing							
T236A	Arg 220	Thr 236	Arg 57	104.4	101.7	99.5 ± 5.0	75.5 ± 0.7
D394A	Lys 378	Asp 394	Glu 216	110.5	120.3	98.0 ± 4.2	97.0 ± 2.8
S289K	Ser273	Ser289	Lys111	101.5	83.3	99.5 ± 0.7	65.5 ± 3.5
R398A	Ala 224	Arg 398	Glu 220	128.8	119.8	98.5 ± 0.7	98.0 ± 2.8
2. Variants relating to β2/β3 loop							
tPHD2Δloop	-	-	-	212.4	41.7	<10	n.d.
tPHD2-PHD3	-	-	-	102.2	101.1	99.5 ± 5.0	89.0 ± 4.2
R252A	Arg 236	Arg 252	Arg 74	3.8	3.1	19.5 ± 0.7	n.d.
D254A	Arg 238	Arg 254	Arg 76	116.6	112.3	16.5 ± 2.1	n.d.
D254K	Asp 238	Asp 254	Asp 76	131.4	103.5	21.0 ± 1.4	n.d.
R252C/D254C	Arg 236 Asp 238	Arg 252 Arg 254	Arg 74 Asp 76	107.5	102.8	22.5 ± 2.1	n.d.
3. Clinically observed and related mutations							
P317R	Pro 301	Pro 317	Pro 139	66.1	20.2	96.5 ± 0.7	n.d.
P317E	Pro 301	Pro 317	Pro 139	N.T.	N.T.	95.5 ± 2.1	29.0 ± 4.2
R371H	Arg 355	Arg 371	Arg 193	106.6	89.8	99.0 ± 2.8	69.0 ± 2.8
4. Variants relating to 2OG and Fe(II) interacting residues							
M299V	Met 283	Met 299	Met 121	88.4	42.8	94.0 ± 2.8	53.0 ± 2.8
M299E	Met 283	Met 299	Met 121	7.8	3.7	56.5 ± 3.5	39.5 ± 2.1
Y303F	Tyr 287	Tyr 303	Tyr 125	105.2	N.T.	97.0 ± 2.8	N.T.
Y303A	Tyr 287	Tyr 303	Tyr 125	171.4	N.T.	22.5 ± 3.5	N.T.
D315G	Asp 299	Asp 315	Asp 137	51.8	N.T.	22.5 ± 3.5	N.T.
T325A	Thr 309	Thr 325	Thr 147	107.6	N.T.	97.5 ± 2.1	N.T.
Y329F	Tyr 313	Tyr 329	Tyr 151	89.8	N.T.	97.0 ± 1.4	N.T.
F366Y	Phe 350	Phe 366	Phe 188	101.2	N.T.	96.0 ± 2.8	N.T.
R383A	Arg 367	Arg 383	Arg 205	2.5	N.T.	<10	N.T.
T387A	Thr 271	Thr 387	Thr 209	113.4	N.T.	99.0 ± 1.4	N.T.
5. Variants relating to LXXLAP binding residues							
S242G	Ser 226	Ser 242	Gly 63	N.T.	N.T.	97.5 ± 0.7	47.5 ± 0.7
K244R	Arg 228	Lys 244	Arg 65	N.T.	N.T.	71.5 ± 0.7	28.0 ± 1.4
I251L	Ile 235	Ile 251	Leu 73	N.T.	N.T.	86.5 ± 0.7	41.0 ± 0.7
I251G	Ile 235	Ile 251	Leu 73	N.T.	N.T.	54.5 ± 0.7	15.0 ± 2.8
K244R/I251L	Arg 228 Ile 235	Lys 244 Ile 251	Arg 65 Leu 73	112.0	91.2	88.5 ± 2.1	43.0 ± 1.4
S242G/K244R/I251L	Ser 226 Arg 228 Ile 235	Ser 242 Lys 244 Ile 251	Gly 63 Arg 65 Leu 73	104.8	103.5	87.0 ± 2.8	77.0 ± 1.4
S242G/K244R/I251L /P264E	Ser 226 Arg 228 Ile 235 Pro 248	Ser 242 Lys 244 Ile 251 Pro 264	Gly 63 Arg 65 Leu 73 Glu 86	107.2	108.8	87.5 ± 3.5	71.5 ± 2.1

Wildtype/ Variant tPHD2	Equivalent residues in PHD1-3 isoforms			% 2OG turned over relative to tPHD2*		% Hydroxylation	
S242G/K244R/I251L /A228Q/K234A/T236 R/P264E	Ala 212 Arg 218 Arg 220 Ser 226 Arg 228 Ile 235 Pro 248	Ala 228 Lys 234 Thr 236 Ser 242 Lys 244 Ile 251 Pro 264	Gln 49 Ala 55 Arg 57 Gly 63 Arg 65 Leu 73 Glu 86	102.1	97.9	90 ± 2.8	86.5 ± 2.1
6. Variants of other residues involved in substrate binding/ selectivity							
R281L	Arg 265	Arg 281	Leu 103	92.2	51.7	97.5 ± 0.7	46.0 ± 1.4
I280V/I292V/R281L	Ile 264 Arg 265 Ile 276	Ile 280 Arg 281 Ile 292	Val 102 Leu 103 Val 114	112.3	60.2	98.5 ± 2.1	84.5 ± 3.5
I280C	Ile 264	Ile 280	Val 102	80.2	N.T.	90.0 ± 2.8	50.0 ± 2.8
K297I	Lys 281	Lys 297	Lys 119	107.8	73.3	N.T.	N.T.
Y310F	Tyr 294	Tyr 310	Tyr 132	97.5	105.5	98.5 ± 2.1	85.5 ± 0.7
R396A	Arg 380	Arg 396	Arg 218	29.3	90.1	28.5 ± 5.0	96.5 ± 0.7
7. Cysteine variants							
C201A	Cys 185	Cys 201	Cys 22	97.0	94.7	99.5 ± 0.7	86.5 ± 0.71
C208A	Cys 192	Cys 208	Cys 29	129.4	96.6	99.5 ± 0.7	80.5 ± 1.41
C266A	Cys 250	Cys 266	Cys 88	128.7	86.5	91.4 ± 3.5	70.0 ± 2.12
C283A	Cys 267	Cys 283	Cys 105	99.3	100.7	98.5 ± 2.1	76.5 ± 2.83
C302A	Cys 286	Cys 302	Cys 124	150.9	145.0	98.5 ± 1.4	85.5 ± 3.54
C323A	Cys 307	Cys 323	Cys 145	97.8	89.8	98.0 ± 2.8	70.5 ± 2.12
C326A	Cys 310	Cys 326	Cys 148	124.6	134.0	93.5 ± 2.8	69.0 ±4.95

* Activities measured in nmoles of 2OG turnover (mean ± S.D.) were converted into percentages relative to tPHD2 and hence standard deviations are not given.

n.d. = not detected under the experimental conditions; N.T. = Not tested.

NODD = HIF-1$\alpha_{395\text{-}413}$; CODD = HIF-1$\alpha_{556\text{-}574}$

Table 2. Summary of results for hydroxylation of HIF-1α ODDs (NODD and CODD) at variable E/S ratios (1:50 and 1:25) using tPHD2/the important tPHD2 variants in individual and competition substrate experiments.

CODD	**Individual (1:50)**	**Individual (1:25)**	**Competition (1:50)**	**Competition (1:25)**
tPHD2	97.5 ± 0.7	98.5 ± 0.7	99.5 ± 0.7	99.5 ± 0.7
T236A	98.0 ± 0.0	99.5 ± 5.0	99.0 ± 0.0	100.0 ± 1.4
S242G	96.5 ± 0.7	97.5 ± 0.7	N.T.	98.5 ± 0.7
K244R	64.0 ± 1.4	71.5 ± 0.7	N.T.	63.5 ± 3.5
I251L	69.0 0 ±1.4	86.5 ± 0.7	N.T.	N.T.
I251G	36.0 0 ±2.8	54.5 ± 0.7	N.T.	46.0 ± 2.8
R252A	17.0 ± 1.4	19.5 ± 0.7	19.5 ± 0.7	21.5 ± 0.7
D254A	15.5 ± 2.1	16.5 ± 2.1	21.0 ± 5.7	23.5 ± 0.7
D254K	20.0 ± 1.4	21.0 ± 1.4	20.5 ± 3.5	24.5 ± 0.7
Y310F	94.0 ± 1.4	98.5 ± 2.1	98.5 ± 0.7	98.0 ± 2.8
P317R	93.0 ± 1.4	96.5 ± 0.7	96.5 ± 0.7	97.5 ± 0.7
P317E	90.5 ± 2.1	95.5 ± 2.1	97.5 ± 2.1	96.0 ± 1.4
R371H	98.5 ± 3.5	99.0 ± 2.8	95.5 ± 0.7	97.5 ± 2.1
D394A	95.5 ± 0.7	98.0 ± 4.2	94.5 ± 0.7	100.5 ± 2.1
R396A	29.0 ± 4.2	28.5 ± 5.0	11.5 ± 0.7	24.5 ± 2.1
R398A	96.5 ± 0.7	98.5 ± 0.7	91.5 ± 0.7	96.5 ± 3.5
K244R/I251L	86.5 ± 0.7	88.5 ± 2.1	96.5 ± 0.7	98.5 ± 0.7
S242G/K244R/I251L	90.5 ± 2.1	87.0 ± 2.8	97.5 ± 2.1	99.0 ± 1.4
I280V/R281L/I292V	90.5 ± 2.1	98.5 ± 2.1	93.5 ± 0.7	96.5 ± 6.4
R281L	89.0 ± 1.4	97.5 ± 0.7	97.0 ± 1.4	100.0 ± 0.0
tPHD2-PHD3	91.5 ± 0.7	99.5 ± 5.0	93.5 ± 2.1	99.0 ± 1.4

NODD	**Individual (1:50)**	**Individual (1:25)**	**Competition (1:50)**	**Competition (1:25)**
tPHD2	76.0 ± 1.4	78.5 ± 0.7	9.5 ± 0.7	50.0 ± 2.8
T236A	75.0 ± 1.4	75.5 ± 0.7	25 ± 1.4	40.5 ± 0.7
S242G	33.5 ± 0.7	47.5 ± 0.7	N.T.	24.0 ± 0.7
K244R	22.0 ± 2.8	28.0 ± 1.4	N.T.	n.d.
I251L	28.0 ± 0.7	41.0 ± 0.7	N.T.	N.T.
I251G	n.d.	n.d.	N.T.	n.d.
R252A	n.d.	n.d.	n.d.	n.d.
D254A	n.d.	n.d.	n.d.	n.d.
D254K	n.d.	n.d.	n.d.	n.d.
Y310F	78.0 ± 2.8	85.5 ± 0.7	47.5 ± 0.7	51.5 ± 2.1
P317R	n.d.	n.d.	n.d.	n.d.
P317E	11.0 ± 1.4	29.0 ± 4.2	n.d.	<10
R371H	55.0 ± 1.4	69.0 ± 2.8	n.d.	35.5 ± 2.1
D394A	79.0 ± 1.4	97.0 ± 2.8	50.5 ± 0.7	57.5 ± 0.7
R396A	74.0 ± 2.8	96.5 ± 0.7	81.5 ± 0.7	88.5 ± 2.1
R398A	85.5 ± 0.7	98.0 ± 2.8	54 ± 1.4	63.0 ± 4.2
K244R/I251L	43.0 ± 1.4	43.0 ± 1.4	n.d.	39.5 ± 0.7
S242G/K244R/I251L	69.0 ± 1.4	77.0 ± 1.4	n.d.	n.d.
I280V/R281L/I292V	69.0 ± 1.4	76.5 ± 0.7	19 ± 2.8	32.0 ± 2.8
R281L	29.0 ± 0.00	46.0 ± 1.4	n.d.	9.5 ± 0.7
tPHD2-PHD3	87.50 ± 2.1	89.0 ± 4.2	n.d.	n.d.

n.d. = not detected under the experimental conditions;
NODD = HIF-1$\alpha_{395-413}$; CODD = HIF-1$\alpha_{556-574}$

Appendix 5

Table 1. Hydrogen bonds/ ionic interactions as apparently observed in tPHD2.NODD models (see Chapter 4, Fig. 4.8).

PHD1 Equivalent Residue	PHD3 Equivalent Residue	PHD2 Residue Atom	tPHD2.NODD Model 2		tPHD2.NODD Model 3	
			NODD Residue Atom	Distance/ Å	NODD Residue Atom	Distance / Å
Arg 354	Arg 192	Arg 370 NH1	-	-	Asp 395 Oδ1	2.6
Arg 354	Arg 192	Arg 370 NH2	-	-	Asp 395 Oδ2	2.8
His 302	Asn 140	Asn 318 Oδ1	Thr 398 Oγ1	2.9	-	-
Ser 226	Gly 63	Ser 242 NH	-		Thr 398 O	3.0
Tyr 294	Tyr 132	Tyr 310 OH	Leu 400 O	2.8	Leu 400 O	2.8
Val 225	Ala 62	Val 241 NH	Ala 401 O	3.0	Ala 401 O	2.8
Arg 306	Arg 144	Arg 322 NH1	Pro 402 O	2.8	Pro 402 O	2.8
Arg 306	Arg 144	Arg 322 NH2	Pro 402 O	3.0	Pro 402 O	3.0
Gln 223	Gln 60	Gln 239 O	-		Ala 403 NH	3.4
Arg 380	Arg 218	Arg 396 NH1	Gly 405 O	2.7	-	-
Arg 380	Arg 218	Arg 396 NH2	Gly 405 O	3.4	Thr 407 Oγ1	3.0
Arg 380	Arg 218	Arg 396 NH2	Asp 406 O	3.5	-	-
Arg 380	Arg 218	Arg 396 NH2	Thr 407 O	3.0	-	-
Arg 277	Arg 117	Arg 295 O	Ser 410 NH	2.8	Ile 409 NH	3.1
Arg 277	Arg 117	Arg 295 NH	Ser 410 O	2.7	Ile 409 O	3.1
Arg 277	Arg 117	Arg 295 NH2	Asp 412 O	2.7	-	
Asn 275	Lys 115	Asn 293 O	-	-	Leu 411 NH	3.1
Arg 263	Leu 103	Arg 281 Nε	-	-	Asp 412 Oδ2	2.8
Arg 263	Leu 103	Arg 281 NH1	Asp 412 Oδ1	2.6	Asp 412 Oδ1	2.7
Arg 263	Leu 103	Arg 281 NH2	Asp 412 Oδ2	2.8	-	-
Arg 263	Leu 103	Arg 281 NH1	Leu 411 O	3.0	-	-
Arg 263	Leu 103	Arg 281 NH2	Asp 412 O	2.7	-	-

Table 2. Hydrophobic interactions as apparently observed in tPHD2.NODD models (see Chapter 4, Fig. 4.8).

PHD1 Equivalent Residue	PHD3 Equivalent Residue	PHD2 Residue Atom	tPHD2.NODD Model 2 NODD Residue	tPHD2.NODD Model 3 NODD Residue
Pro 301	Pro 13	Pro 317	-	Leu 397
Pro 301	Pro 13	Pro 317	-	Thr 398
Arg 354	Arg 192	Arg 370	Leu 397	-
Val 298	Val 136	Val 314	Leu 397	Leu 397
Val 225	Ala 62	Val 241	Thr 398	Ala 401
Gln 227	Pro 64	Gln 243	Leu 399	-
Arg 233	Arg 65	Lys 244	Leu 399	Leu 399
Ile 235	Leu 73	Ile 251	Leu 400	Leu 400
Pro 301	Pro 13	Pro 317	Ala 401	Ala 401
Trp 373	Trp 211	Trp 389	Pro 402	Pro 402
Tyr 375	Tyr 213	Phe 391	Ala 404	-
Arg 306	Arg 144	Arg 322	Ala 404	-
Thr 278	Ser 118	Thr 296	Ala 404	Ala 404
Thr 278	Ser 118	Thr 296	-	Ile 408
Lys 279	Lys 119	Lys 297	-	Ile 408
Arg 277	Arg 117	Arg 295	-	Ile 409
Tyr 374	Tyr 212	Tyr 390	-	Leu 411
Asp 261	Asp 99	Asp 277	-	Leu 411
Gly 276	Glu 116	Gly 294	-	Leu 411
Arg 380	Arg 218	Arg 396	Thr 407	-
Arg 263	Leu 103	Arg 281	Asp 412	Asp 412

Appendix 6

MS/MS data on tPHD2Me-I tryptic fragments showing peptide mass fingerprints, sequence coverage and MS/MS ion scores.

Match to: **Q9GZT9** (Prolyl hydroxylase domain-containing protein 2) Score: **2527**
Instrument type: ESI-QUAD-TOF; Software used: MASCOT (http://www.matrixscience.com/)
Sequence Coverage: **54%;** Matched peptides shown in **Bold Red**

```
   1 MANDSGGPGG PSPSERDRQY CELCGKMENL LRCSRCRSSF YCCKEHQRQD
  51 WKKHKLVCQG SEGALGHGVG PHQHSGPAPP AAVPPPRAGA REPRKAAARR
 101 DNASGDAAKG KVKAKPPADP AAAASPCRAA AGGQGSAVAA EAEPGKEEPP
 151 ARSSLFQEKA NLYPPSNTPG DALSPGGGLR PNGQTKPLPA LKLALEYIVP
 201 CMNKHGICVV DDFLGKETGQ QIGDEVRALH DTGKFTDGQL VSQKSDSSKD
 251 IRGDKITWIE GKEPGCETIG LLMSSMDDLI RHCNGKLGSY KINGRTKAMV
 301 ACYPGNGTGY VRHVDNPNGD GRCVTCIYYL NKDWDAKVSG GILRIFPEGK
 351 AQFADIEPKF DRLLFFWSDR RNPHEVQPAY ATRYAITVWY FDADERARAK
 401 VKYLTGEKGV RVELNKPSDS VGKDVF
```

The table contains the following terms: **Start – End**, the residues that bracket the peptide sequence in the protein; **Observed**, experimental *m/z* value; **Mr (expt)**, experimental *m/z* transformed to a relative molecular mass; **Mr (calc)**, relative molecular mass calculated from the matched peptide sequence; **Delta**, difference (error) between the experimental and calculated masses; **Miss**, number of missed cleavage sites; **Sequence** of the peptide; **Ions score**, the ion score is a measure of how well the observed MS/MS spectrum matches to the stated peptide. A higher score indicates a more confident match. Modifications found in the peptide are given in the right.

Start - End	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Sequence
142 - 146	530.3088	529.3015	528.2907	1.0108	0	E.AEPGK.E Dimethyl (K) (Ions score 3)
198 - 227	840.1806	3356.6933	3355.6727	1.0206	2	Y.IVPCMNKHGICVVDDFLGKETGQQIGDEVR.A 2 Dimethyl (K) (Ions score 110)
198 - 227	840.1806	3356.6933	3355.6727	1.0206	2	Y.IVPCMNKHGICVVDDFLGKETGQQIGDEVR.A 2 Dimethyl (K) (Ions score 110)
202 - 227	977.5103	2929.5091	2929.4426	0.0664	2	C.MNKHGICVVDDFLGKETGQQIGDEVR.A Dimethyl (K); Methyl (K) (Ions score 48)
202 - 227	977.5103	2929.5091	2929.4426	0.0664	2	C.MNKHGICVVDDFLGKETGQQIGDEVR.A Dimethyl (K); Methyl (K) (Ions score 45)
202 - 227	977.8469	2930.5189	2929.4426	1.0762	2	C.MNKHGICVVDDFLGKETGQQIGDEVR.A Dimethyl (K); Methyl (K) (Ions score 28)
228 - 246	701.6965	2102.0677	2102.0644	0.0033	2	R.ALHDTGKFTDGQLVSQKSD.S 2 Dimethyl (K) (Ions score 62)
228 - 252	698.3585	2789.4049	2788.4355	0.9694	3	R.ALHDTGKFTDGQLVSQKSDSSKDIR.G 2 Dimethyl (K) (Ions score 23)
228 - 252	561.5365	2802.6461	2802.4512	0.1949	3	R.ALHDTGKFTDGQLVSQKSDSSKDIR.G 2 Dimethyl (K); Methyl (K) (Ions score 30)
228 - 252	564.2947	2816.4371	2816.4668	-0.0297	3	R.ALHDTGKFTDGQLVSQKSDSSKDIR.G 3 Dimethyl (K) (Ions score 29)
228 - 252	705.1183	2816.4441	2816.4668	-0.0227	3	R.ALHDTGKFTDGQLVSQKSDSSKDIR.G 3 Dimethyl (K) (Ions score 64)
228 - 252	564.2961	2816.4441	2816.4668	-0.0227	3	R.ALHDTGKFTDGQLVSQKSDSSKDIR.G 3 Dimethyl (K) (Ions score 40)
228 - 252	705.1244	2816.4685	2816.4668	0.0017	3	R.ALHDTGKFTDGQLVSQKSDSSKDIR.G 3 Dimethyl (K) (Ions score 39)
228 - 252	939.8405	2816.4997	2816.4668	0.0328	3	R.ALHDTGKFTDGQLVSQKSDSSKDIR.G 3 Dimethyl (K) (Ions score 55)
228 - 252	567.4970	2832.4486	2830.4461	2.0025	3	R.ALHDTGKFTDGQLVSQKSDSSKDIR.G Acetyl (K); 2 Dimethyl (K) (Ions score 27)
230 - 252	527.4729	2632.3281	2632.3457	-0.0175	3	L.HDTGKFTDGQLVSQKSDSSKDIR.G 3 Dimethyl (K) (Ions score 36)
230 - 252	659.0976	2632.3613	2632.3457	0.0156	3	L.HDTGKFTDGQLVSQKSDSSKDIR.G 3 Dimethyl (K) (Ions score 43)
231 - 252	624.8382	2495.3237	2495.2867	0.0370	3	H.DTGKFTDGQLVSQKSDSSKDIR.G 3 Dimethyl (K) (Ions score 63)

241 - 252	469.2546	1404.7420	1404.7572	-0.0153	2	L.VSQKSDSSKDIR.G 2 Dimethyl (K) (Ions score 44)
253 - 272	743.7176	2228.1310	2228.1398	-0.0089	2	R.GDKITWIEGKEPGCETIGLL.M Acetyl (K); Dimethyl (K) (Ions score 43)
253 - 281	816.6610	3262.6149	3262.6287	-0.0138	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H 2 Dimethyl (K) (Ions score 137)
253 - 281	1088.5583	3262.6531	3262.6287	0.0243	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H 2 Dimethyl (K) (Ions score 77)
253 - 281	820.1597	3276.6097	3276.6080	0.0017	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H Acetyl (K); Dimethyl (K) (Ions score 95)
253 - 281	1099.2100	3294.6082	3292.6029	2.0053	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H Acetyl (K); Oxidation (M); Dimethyl (K) (Ions score 45)
253 - 281	1099.2100	3294.6082	3292.6029	2.0053	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H Acetyl (K); Oxidation (M); Dimethyl (K) (Ions score 52)
253 - 281	1104.5616	3310.6630	3308.5978	2.0651	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H Acetyl (K); 2 Oxidation (M); Dimethyl (K) (Ions score 18)
253 - 281	1104.5616	3310.6630	3308.5978	2.0651	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H Acetyl (K); 2 Oxidation (M); Dimethyl (K) (Ions score 20)
255 - 281	1037.1718	3108.4936	3106.5388	1.9547	2	D.KITWIEGKEPGCETIGLLMSSMDDLIR.H Acetyl (K); Oxidation (M); Methyl (K) (Ions score 26)
256 - 276	776.7345	2327.1817	2326.0895	1.0922	1	K.ITWIEGKEPGCETIGLLMSSM.D 2 Oxidation (M) (Ions score 13)
256 - 281	975.1593	2922.4561	2920.4384	2.0177	1	K.ITWIEGKEPGCETIGLLMSSMDDLIR.H Methyl (K) (Ions score 44)
266 - 295	680.3450	3396.6886	3395.6822	1.0065	3	G.CETIGLLMSSMDDLIRHCNGKLGSYKINGR.T Oxidation (M); 2 Dimethyl (K) (Ions score 10)
268 - 295	810.6478	3238.5621	3236.6468	1.9153	3	E.TIGLLMSSMDDLIRHCNGKLGSYKINGR.T Carbamidomethyl (C); 2 Oxidation (M); 2 Dimethyl (K) (Ions score 20)
273 - 281	1069.3308	1068.3235	1066.4787	1.8448	0	L.MSSMDDLIR.H (Ions score 59)
275 - 281	849.4321	848.4248	848.4062	0.0186	0	S.SMDDLIR.H (Ions score 18)
282 - 295	534.9466	1601.8180	1601.8460	-0.0280	2	R.HCNGKLGSYKINGR.T 2 Dimethyl (K) (Ions score 73)
284 - 295	455.2612	1362.7618	1361.7779	0.9839	2	C.NGKLGSYKINGR.T 2 Dimethyl (K) (Ions score 61)
296 - 312	902.4570	1802.8994	1802.8444	0.0551	1	R.TKAMVACYPGNGTGYVR.H Oxidation (M) (Ions score 18)
296 - 312	908.4061	1814.7976	1814.8808	-0.0831	1	R.TKAMVACYPGNGTGYVR.H Dimethyl (K) (Ions score 104)
296 - 312	908.4061	1814.7976	1814.8808	-0.0831	1	R.TKAMVACYPGNGTGYVR.H Dimethyl (K) (Ions score 99)
296 - 312	605.9683	1814.8831	1814.8808	0.0023	1	R.TKAMVACYPGNGTGYVR.H Dimethyl (K) (Ions score 108)
296 - 312	909.9340	1817.8534	1816.8600	0.9934	1	R.TKAMVACYPGNGTGYVR.H Oxidation (M); Methyl (K) (Ions score 47)
296 - 312	909.9340	1817.8534	1816.8600	0.9934	1	R.TKAMVACYPGNGTGYVR.H Oxidation (M); Methyl (K) (Ions score 42)
298 - 312	779.8756	1557.7366	1557.7068	0.0298	0	K.AMVACYPGNGTGYVR.H (Ions score 54)
300 - 312	678.8273	1355.6400	1355.6292	0.0108	0	M.VACYPGNGTGYVR.H (Ions score 91)
303 - 312	542.2706	1082.5266	1082.5145	0.0121	0	C.YPGNGTGYVR.H (Ions score 94)
304 - 312	460.7343	919.4540	919.4512	0.0029	0	Y.PGNGTGYVR.H (Ions score 47)
304 - 312	460.7345	919.4544	919.4512	0.0033	0	Y.PGNGTGYVR.H (Ions score 50)
308 - 312	595.2916	594.2843	594.3126	-0.0282	0	G.TGYVR.H (Ions score 8)
313 - 322	540.7369	1079.4592	1079.4744	-0.0152	0	R.HVDNPNGDGR.C (Ions score 54)
313 - 322	540.7369	1079.4592	1079.4744	-0.0152	0	R.HVDNPNGDGR.C (Ions score 60)
313 - 322	540.7409	1079.4672	1079.4744	-0.0072	0	R.HVDNPNGDGR.C (Ions score 17)
313 - 322	540.7448	1079.4750	1079.4744	0.0006	0	R.HVDNPNGDGR.C (Ions score 58)
313 - 322	541.2387	1080.4628	1079.4744	0.9884	0	R.HVDNPNGDGR.C (Ions score 58)
323 - 344	859.0855	2574.2347	2572.3182	1.9165	2	R.CVTCIYYLNKDWDAKVSGGILR.I 2 Dimethyl (K) (Ions score 40)
323 - 344	863.1083	2586.3031	2586.3338	-0.0307	2	R.CVTCIYYLNKDWDAKVSGGILR.I Dimethyl (K); Trimethyl (K) (Ions score 20)
323 - 344	863.7669	2588.2789	2586.3338	1.9451	2	R.CVTCIYYLNKDWDAKVSGGILR.I Dimethyl (K); Trimethyl (K) (Ions score 22)
327 - 344	723.0709	2166.1909	2166.1837	0.0072	2	C.IYYLNKDWDAKVSGGILR.I 2 Dimethyl (K) (Ions score 28)
328 - 344	685.3731	2053.0975	2053.0996	-0.0022	2	I.YYLNKDWDAKVSGGILR.I 2 Dimethyl (K) (Ions score 55)
329 - 344	631.0171	1890.0295	1890.0363	-0.0068	2	Y.YLNKDWDAKVSGGILR.I 2 Dimethyl (K) (Ions score 68)
333 - 344	672.8683	1343.7220	1343.7197	0.0023	1	K.DWDAKVSGGILR.I Dimethyl (K) (Ions score 69)
345 - 362	722.0221	2163.0445	2163.1364	-0.0919	2	R.IFPEGKAQFADIEPKFDR.L 2 Dimethyl (K) (Ions score 77)
345 - 362	541.7911	2163.1353	2163.1364	-0.0011	2	R.IFPEGKAQFADIEPKFDR.L 2 Dimethyl (K) (Ions score 51)
345 - 362	1082.5922	2163.1698	2163.1364	0.0334	2	R.IFPEGKAQFADIEPKFDR.L 2 Dimethyl (K) (Ions score 115)
346 - 362	513.5165	2050.0369	2050.0524	-0.0155	2	I.FPEGKAQFADIEPKFDR.L 2 Dimethyl (K) (Ions score 54)
346 - 362	684.3670	2050.0792	2050.0524	0.0268	2	I.FPEGKAQFADIEPKFDR.L 2 Dimethyl (K) (Ions score 65)
347 - 362	635.3302	1902.9688	1902.9839	-0.0152	2	F.PEGKAQFADIEPKFDR.L 2 Dimethyl (K) (Ions score 48)
347 - 362	635.3315	1902.9727	1902.9839	-0.0113	2	F.PEGKAQFADIEPKFDR.L 2 Dimethyl (K) (Ions score 41)
347 - 362	635.9890	1904.9452	1902.9839	1.9612	2	F.PEGKAQFADIEPKFDR.L 2 Dimethyl (K) (Ions score 70)
348 - 362	602.9863	1805.9371	1805.9312	0.0059	2	P.EGKAQFADIEPKFDR.L 2 Dimethyl (K) (Ions score 59)
349 - 362	559.9723	1676.8951	1676.8886	0.0065	2	E.GKAQFADIEPKFDR.L 2 Dimethyl (K) (Ions score 30)
358 - 362	345.6946	689.3746	689.3860	-0.0114	1	E.PKFDR.L Dimethyl (K) (Ions score 16)
363 - 370	542.2795	1082.5444	1082.5549	-0.0105	0	R.LLFFWSDR.R (Ions score 36)
363 - 370	542.2924	1082.5702	1082.5549	0.0153	0	R.LLFFWSDR.R (Ions score 19)
363 - 371	620.3198	1238.6250	1238.6560	-0.0310	1	R.LLFFWSDRR.N (Ions score 20)
363 - 383	651.5874	2602.3205	2602.3193	0.0012	2	R.LLFFWSDRRNPHEVQPAYATR.Y (Ions score 47)
364 - 371	563.7822	1125.5498	1125.5720	-0.0221	1	L.LFFWSDRR.N (Ions score 24)

365 - 370	857.3929	856.3856	856.3868	-0.0012	0	L.FFWSDR.R (Ions score 18)
371 - 383	513.5715	1537.6927	1537.7749	-0.0823	1	R.RNPHEVQPAYATR.Y (Ions score 62)
371 - 383	513.5715	1537.6927	1537.7749	-0.0823	1	R.RNPHEVQPAYATR.Y (Ions score 45)
371 - 383	513.6038	1537.7896	1537.7749	0.0146	1	R.RNPHEVQPAYATR.Y (Ions score 18)
371 - 383	513.9049	1538.6929	1537.7749	0.9179	1	R.RNPHEVQPAYATR.Y (Ions score 62)
371 - 383	513.9049	1538.6929	1537.7749	0.9179	1	R.RNPHEVQPAYATR.Y (Ions score 45)
371 - 383	514.2233	1539.6481	1537.7749	1.8731	1	R.RNPHEVQPAYATR.Y (Ions score 63)
371 - 383	514.2233	1539.6481	1537.7749	1.8731	1	R.RNPHEVQPAYATR.Y (Ions score 44)
372 - 383	691.8413	1381.6680	1381.6738	-0.0058	0	R.NPHEVQPAYATR.Y (Ions score 46)
372 - 383	691.8468	1381.6790	1381.6738	0.0052	0	R.NPHEVQPAYATR.Y (Ions score 48)
372 - 383	691.8468	1381.6790	1381.6738	0.0052	0	R.NPHEVQPAYATR.Y (Ions score 50)
374 - 383	586.3043	1170.5940	1170.5782	0.0159	0	P.HEVQPAYATR.Y (Ions score 33)
379 - 383	581.3095	580.3022	580.2969	0.0053	0	P.AYATR.Y (Ions score 23)
384 - 396	824.8657	1647.7168	1647.7569	-0.0401	0	R.YAITVWYFDADER.A (Ions score 32)
384 - 396	824.8844	1647.7542	1647.7569	-0.0027	0	R.YAITVWYFDADER.A (Ions score 90)
384 - 396	824.8846	1647.7546	1647.7569	-0.0023	0	R.YAITVWYFDADER.A (Ions score 93)
384 - 396	824.8878	1647.7610	1647.7569	0.0041	0	R.YAITVWYFDADER.A (Ions score 81)
384 - 396	824.8879	1647.7612	1647.7569	0.0043	0	R.YAITVWYFDADER.A (Ions score 73)
384 - 396	824.8901	1647.7656	1647.7569	0.0087	0	R.YAITVWYFDADER.A (Ions score 94)
384 - 396	824.8901	1647.7656	1647.7569	0.0087	0	R.YAITVWYFDADER.A (Ions score 94)
384 - 396	824.8908	1647.7670	1647.7569	0.0101	0	R.YAITVWYFDADER.A (Ions score 93)
384 - 396	824.8909	1647.7672	1647.7569	0.0103	0	R.YAITVWYFDADER.A (Ions score 93)
384 - 396	824.8912	1647.7678	1647.7569	0.0109	0	R.YAITVWYFDADER.A (Ions score 89)
384 - 396	824.8912	1647.7678	1647.7569	0.0109	0	R.YAITVWYFDADER.A (Ions score 93)
384 - 396	824.8912	1647.7678	1647.7569	0.0109	0	R.YAITVWYFDADER.A (Ions score 71)
384 - 396	824.8915	1647.7684	1647.7569	0.0115	0	R.YAITVWYFDADER.A (Ions score 75)
384 - 396	824.8919	1647.7692	1647.7569	0.0123	0	R.YAITVWYFDADER.A (Ions score 67)
384 - 396	824.8948	1647.7750	1647.7569	0.0181	0	R.YAITVWYFDADER.A (Ions score 52)
384 - 396	824.8962	1647.7778	1647.7569	0.0209	0	R.YAITVWYFDADER.A (Ions score 56)
384 - 396	824.8965	1647.7784	1647.7569	0.0215	0	R.YAITVWYFDADER.A (Ions score 94)
399 - 411	511.6212	1531.8418	1531.9450	-0.1032	3	R.AKVKYLTGEKGVR.V 3 Dimethyl (K) (Ions score 28)
399 - 411	383.9871	1531.9193	1531.9450	-0.0257	3	R.AKVKYLTGEKGVR.V 3 Dimethyl (K) (Ions score 48)
399 - 411	383.9876	1531.9213	1531.9450	-0.0237	3	R.AKVKYLTGEKGVR.V 3 Dimethyl (K) (Ions score 45)
399 - 411	511.6540	1531.9402	1531.9450	-0.0048	3	R.AKVKYLTGEKGVR.V 3 Dimethyl (K) (Ions score 46)
399 - 411	766.9895	1531.9644	1531.9450	0.0195	3	R.AKVKYLTGEKGVR.V 3 Dimethyl (K) (Ions score 67)
399 - 411	511.9628	1532.8666	1531.9450	0.9216	3	R.AKVKYLTGEKGVR.V 3 Dimethyl (K) (Ions score 26)
399 - 411	512.2752	1533.8038	1531.9450	1.8588	3	R.AKVKYLTGEKGVR.V 3 Dimethyl (K) (Ions score 26)
399 - 411	384.5005	1533.9729	1531.9450	2.0279	3	R.AKVKYLTGEKGVR.V 3 Dimethyl (K) (Ions score 53)
401 - 411	435.9282	1304.7628	1304.7816	-0.0188	2	K.VKYLTGEKGVR.V 2 Dimethyl (K) (Ions score 24)
412 - 426	563.9448	1688.8126	1688.8985	-0.0859	1	R.VELNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 70)
412 - 426	845.4363	1688.8580	1688.8985	-0.0405	1	R.VELNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 86)
412 - 426	563.9701	1688.8885	1688.8985	-0.0100	1	R.VELNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 38)
412 - 426	563.9734	1688.8984	1688.8985	-0.0001	1	R.VELNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 74)
412 - 426	845.4616	1688.9086	1688.8985	0.0101	1	R.VELNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 47)
413 - 426	795.9131	1589.8116	1589.8301	-0.0184	1	V.ELNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 67)
413 - 426	795.9230	1589.8314	1589.8301	0.0014	1	V.ELNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 60)
413 - 426	795.9235	1589.8324	1589.8301	0.0024	1	V.ELNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 35)
414 - 426	731.3640	1460.7134	1460.7875	-0.0740	1	E.LNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 54)
414 - 426	731.3764	1460.7382	1460.7875	-0.0492	1	E.LNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 46)
414 - 426	731.4038	1460.7930	1460.7875	0.0056	1	E.LNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 17)
414 - 426	731.9892	1461.9638	1460.7875	1.1764	1	E.LNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 19)
415 - 426	674.8706	1347.7266	1347.7034	0.0232	1	L.NKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 39)
417 - 426	539.7797	1077.5448	1077.5343	0.0106	1	K.PSDSVGKDVF.- Dimethyl (K) (Ions score 19)
417 - 426	540.2886	1078.5626	1077.5343	1.0284	1	K.PSDSVGKDVF.- Dimethyl (K) (Ions score 19)
420 - 426	779.4512	778.4439	778.4225	0.0214	1	D.SVGKDVF.- Dimethyl (K) (Ions score 17)

Appendix 7

MS/MS data on tPHD2Me-II tryptic fragments showing peptide mass fingerprints, sequence coverage and MS/MS ion scores.

Match to: **Q9GZT9** (Prolyl hydroxylase domain-containing protein 2) Score: **2817**

Instrument type: ESI-QUAD-TOF; Software used: MASCOT (http://www.matrixscience.com/)

Sequence Coverage: **50%;** Matched peptides shown in **Bold Red**

```
   1 MANDSGGPGG PSPSERDRQY CELCGKMENL LRCSRCRSSF YCCKEHQRQD
  51 WKKHKLVCQG SEGALGHGVG PHQHSGPAPP AAVPPPRAGA REPRKAAARR
 101 DNASGDAAKG KVKAKPPADP AAAASPCRAA AGGQGSAVAA EAEPGKEEPP
 151 ARSSLFQEKA NLYPPSNTPG DALSPGGGLR PNGQTKPLPA LKLALEYIVP
 201 CMNKHGICVV DDFLGKETGQ QIGDEVRALH DTGKFTDGQL VSQKSDSSKD
 251 IRGDKITWIE GKEPGCETIG LLMSSMDDLI RHCNGKLGSY KINGRTKAMV
 301 ACYPGNGTGY VRHVDNPNGD GRCVTCIYYL NKDWDAKVSG GILRIFPEGK
 351 AQFADIEPKF DRLLFFWSDR RNPHEVQPAY ATRYAITVWY FDADERARAK
 401 VKYLTGEKGV RVELNKPSDS VGKDVF
```

The table contains the following terms: **Start – End**, the residues that bracket the peptide sequence in the protein; **Observed**, experimental *m/z* value; **Mr (expt)**, experimental *m/z* transformed to a relative molecular mass; **Mr (calc)**, relative molecular mass calculated from the matched peptide sequence; **Delta**, difference (error) between the experimental and calculated masses; **Miss**, number of missed cleavage sites; **Sequence** of the peptide; **Ions score**, the ion score is a measure of how well the observed MS/MS spectrum matches to the stated peptide. A higher score indicates a more confident match. Modifications found in the peptide are given in the right.

Start - End	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Sequence
211 - 227	974.9904	1947.9662	1947.9174	0.0488	1	V.DDFLGKETGQQIGDEVR.A Acetyl (K) (Ions score 33)
217 - 227	616.3005	1230.5864	1230.5840	0.0024	0	K.ETGQQIGDEVR.A (Ions score 89)
217 - 227	616.3005	1230.5864	1230.5840	0.0024	0	K.ETGQQIGDEVR.A (Ions score 89)
228 - 244	624.9967	1871.9683	1871.9741	-0.0059	1	R.ALHDTGKFTDGQLVSQK.S Dimethyl (K) (Ions score 59)
228 - 244	624.9967	1871.9683	1871.9741	-0.0059	1	R.ALHDTGKFTDGQLVSQK.S Dimethyl (K) (Ions score 73)
228 - 252	564.2907	2816.4171	2816.4668	-0.0497	3	R.ALHDTGKFTDGQLVSQKSDSSKDIR.G 3 Dimethyl (K) (Ions score 27)
228 - 252	939.8316	2816.4730	2816.4668	0.0061	3	R.ALHDTGKFTDGQLVSQKSDSSKDIR.G 3 Dimethyl (K) (Ions score 66)
228 - 252	567.3012	2831.4696	2830.4461	1.0235	3	R.ALHDTGKFTDGQLVSQKSDSSKDIR.G Acetyl (K); 2 Dimethyl (K) (Ions score 21)
228 - 255	780.1710	3116.6549	3116.6102	0.0447	4	R.ALHDTGKFTDGQLVSQKSDSSKDIRGDK.I 3 Dimethyl (K) (Ions score 14)
230 - 252	527.4692	2632.3096	2632.3457	-0.0360	3	L.HDTGKFTDGQLVSQKSDSSKDIR.G 3 Dimethyl (K) (Ions score 34)
230 - 252	659.1030	2632.3829	2632.3457	0.0372	3	L.HDTGKFTDGQLVSQKSDSSKDIR.G 3 Dimethyl (K) (Ions score 47)
241 - 252	469.2566	1404.7480	1404.7572	-0.0093	2	L.VSQKSDSSKDIR.G 2 Dimethyl (K) (Ions score 34)
253 - 281	1079.2231	3234.6475	3234.5974	0.0500	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H 2 Methyl (K) (Ions score 84)
253 - 281	813.1566	3248.5973	3248.6131	-0.0158	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H Dimethyl (K); Methyl (K) (Ions score 139)
253 - 281	813.1566	3248.5973	3248.6131	-0.0158	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H Dimethyl (K); Methyl (K) (Ions score 104)
253 - 281	1083.8889	3248.6449	3248.6131	0.0318	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H Dimethyl (K); Methyl (K) (Ions score 145)

253 - 281	1088.5403	3262.5991	3262.6287	-0.0297	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H 2 Dimethyl (K) (Ions score 51)
253 - 281	816.6641	3262.6273	3262.6287	-0.0014	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H 2 Dimethyl (K) (Ions score 135)
253 - 281	1088.8868	3263.6386	3262.6287	1.0098	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H 2 Dimethyl (K) (Ions score 70)
253 - 281	1089.8862	3266.6368	3264.6080	2.0288	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H Oxidation (M); Dimethyl (K); Methyl (K) (Ions score 53)
253 - 281	1094.2084	3279.6034	3277.6033	2.0001	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H Carbamidomethyl (C); Methyl (K) (Ions score 81)
253 - 281	1094.5582	3280.6528	3278.6237	2.0291	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H Oxidation (M); Methyl (K); Trimethyl (K) (Ions score 35)
253 - 281	1094.5582	3280.6528	3278.6237	2.0291	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H Oxidation (M); Methyl (K); Trimethyl (K) (Ions score 34)
253 - 281	824.6535	3294.5849	3292.6029	1.9820	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H Acetyl (K); Oxidation (M); Dimethyl (K) (Ions score 11)
253 - 281	824.6535	3294.5849	3292.6029	1.9820	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H Acetyl (K); Oxidation (M); Dimethyl (K) (Ions score 17)
253 - 281	1099.8875	3296.6407	3294.5822	2.0585	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H Acetyl (K); 2 Oxidation (M); Methyl (K) (Ions score 20)
253 - 281	1099.8875	3296.6407	3294.5822	2.0585	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H Acetyl (K); 2 Oxidation (M); Methyl (K) (Ions score 25)
253 - 281	1104.5654	3310.6744	3308.5978	2.0765	2	R.GDKITWIEGKEPGCETIGLLMSSMDDLIR.H Acetyl (K); 2 Oxidation (M); Dimethyl (K) (Ions score 23)
255 - 281	1031.1935	3090.5587	3090.5439	0.0147	2	D.KITWIEGKEPGCETIGLLMSSMDDLIR.H Acetyl (K); Methyl (K) (Ions score 92)
255 - 281	1031.1935	3090.5587	3090.5439	0.0147	2	D.KITWIEGKEPGCETIGLLMSSMDDLIR.H Acetyl (K); Methyl (K) (Ions score 96)
256 - 281	979.1638	2934.4696	2934.4541	0.0155	1	K.ITWIEGKEPGCETIGLLMSSMDDLIR.H Dimethyl (K) (Ions score 108)
256 - 281	979.1638	2934.4696	2934.4541	0.0155	1	K.ITWIEGKEPGCETIGLLMSSMDDLIR.H Dimethyl (K) (Ions score 132)
256 - 281	989.8246	2966.4520	2964.4282	2.0237	1	K.ITWIEGKEPGCETIGLLMSSMDDLIR.H Acetyl (K); Oxidation (M) (Ions score 71)
256 - 281	989.8246	2966.4520	2964.4282	2.0237	1	K.ITWIEGKEPGCETIGLLMSSMDDLIR.H Acetyl (K); Oxidation (M) (Ions score 58)
260 - 281	808.7421	2423.2045	2421.1589	2.0455	1	I.EGKEPGCETIGLLMSSMDDLIR.H Dimethyl (K) (Ions score 43)
260 - 281	808.7587	2423.2543	2421.1589	2.0953	1	I.EGKEPGCETIGLLMSSMDDLIR.H Dimethyl (K) (Ions score 44)
260 - 281	814.0762	2439.2068	2437.1539	2.0529	1	I.EGKEPGCETIGLLMSSMDDLIR.H Oxidation (M); Dimethyl (K) (Ions score 44)
260 - 281	814.0785	2439.2137	2437.1539	2.0598	1	I.EGKEPGCETIGLLMSSMDDLIR.H Oxidation (M); Dimethyl (K) (Ions score 46)
262 - 281	1135.0077	2268.0008	2267.0847	0.9161	1	G.KEPGCETIGLLMSSMDDLIR.H 2 Oxidation (M); Dimethyl (K) (Ions score 24)
263 - 281	1086.0668	2170.1190	2167.9799	2.1391	0	K.EPGCETIGLLMSSMDDLIR.H Carbamidomethyl (C); 2 Oxidation (M) (Ions score 19)
273 - 281	534.2482	1066.4818	1066.4787	0.0031	0	L.MSSMDDLIR.H (Ions score 27)
273 - 281	1083.5503	1082.5430	1082.4736	0.0694	0	L.MSSMDDLIR.H Oxidation (M) (Ions score 44)
274 - 281	936.4493	935.4420	935.4382	0.0038	0	M.SSMDDLIR.H (Ions score 54)
275 - 281	849.4457	848.4384	848.4062	0.0322	0	S.SMDDLIR.H (Ions score 37)
275 - 281	850.4218	849.4145	848.4062	1.0083	0	S.SMDDLIR.H (Ions score 24)
275 - 281	850.5751	849.5678	848.4062	1.1616	0	S.SMDDLIR.H (Ions score 27)
275 - 295	811.7372	2432.1898	2432.2416	-0.0519	3	S.SMDDLIRHCNGKLGSYKINGR.T 2 Dimethyl (K) (Ions score 20)
278 - 295	540.5274	2158.0805	2156.1636	1.9169	3	D.DLIRHCNGKLGSYKINGR.T Carbamidomethyl (C); 2 Dimethyl (K) (Ions score 26)
282 - 295	534.9439	1601.8099	1601.8460	-0.0361	2	R.HCNGKLGSYKINGR.T 2 Dimethyl (K) (Ions score 60)
282 - 295	534.9439	1601.8099	1601.8460	-0.0361	2	R.HCNGKLGSYKINGR.T 2 Dimethyl (K) (Ions score 50)
282 - 295	534.9543	1601.8411	1601.8460	-0.0049	2	R.HCNGKLGSYKINGR.T 2 Dimethyl (K) (Ions score 11)
282 - 312	683.9534	3414.7306	3412.6955	2.0352	4	R.HCNGKLGSYKINGRTKAMVACYPGNGTGYVR.H Acetyl (K); 2 Dimethyl (K) (Ions score 18)
287 - 312	567.3012	2831.4696	2831.4574	0.0122	3	K.LGSYKINGRTKAMVACYPGNGTGYVR.H 2 Dimethyl (K) (Ions score 16)
287 - 312	723.5870	2890.3189	2890.4582	-0.1393	3	K.LGSYKINGRTKAMVACYPGNGTGYVR.H Carbamidomethyl (C); Oxidation (M); Dimethyl (K); Methyl (K) (Ions score 21)
288 - 295	461.7679	921.5212	921.5032	0.0181	1	L.GSYKINGR.T Dimethyl (K) (Ions score 41)
288 - 295	461.7734	921.5322	921.5032	0.0291	1	L.GSYKINGR.T Dimethyl (K) (Ions score 44)
288 - 295	461.7762	921.5378	921.5032	0.0347	1	L.GSYKINGR.T Dimethyl (K) (Ions score 44)
296 - 312	908.4114	1814.8082	1814.8808	-0.0725	1	R.TKAMVACYPGNGTGYVR.H Dimethyl (K) (Ions score 93)
296 - 312	908.4114	1814.8082	1814.8808	-0.0725	1	R.TKAMVACYPGNGTGYVR.H Dimethyl (K) (Ions score 104)
296 - 312	605.9660	1814.8762	1814.8808	-0.0046	1	R.TKAMVACYPGNGTGYVR.H Dimethyl (K) (Ions score 56)
296 - 312	605.9686	1814.8840	1814.8808	0.0032	1	R.TKAMVACYPGNGTGYVR.H Dimethyl (K) (Ions score 86)
296 - 312	606.6342	1816.8808	1814.8808	2.0000	1	R.TKAMVACYPGNGTGYVR.H Dimethyl (K) (Ions score 86)
296 - 312	909.9353	1817.8560	1816.8600	0.9960	1	R.TKAMVACYPGNGTGYVR.H Oxidation (M); Methyl (K) (Ions score 41)
296 - 312	909.9353	1817.8560	1816.8600	0.9960	1	R.TKAMVACYPGNGTGYVR.H Oxidation (M); Methyl (K) (Ions score 46)
296 - 312	611.3020	1830.8842	1830.8757	0.0085	1	R.TKAMVACYPGNGTGYVR.H Oxidation (M); Dimethyl (K) (Ions score 91)
296 - 312	916.4550	1830.8954	1830.8757	0.0198	1	R.TKAMVACYPGNGTGYVR.H Oxidation (M); Dimethyl (K) (Ions score 90)
296 - 322	720.0949	2876.3505	2876.3446	0.0059	2	R.TKAMVACYPGNGTGYVRHVDNPNGDGR.C Dimethyl (K) (Ions score 31)
302 - 312	594.8123	1187.6100	1185.5237	2.0864	0	A.CYPGNGTGYVR.H (Ions score 12)
303 - 312	542.2688	1082.5230	1082.5145	0.0085	0	C.YPGNGTGYVR.H (Ions score 90)
304 - 312	460.7161	919.4176	919.4512	-0.0335	0	Y.PGNGTGYVR.H (Ions score 49)
304 - 312	460.7312	919.4478	919.4512	-0.0033	0	Y.PGNGTGYVR.H (Ions score 49)
304 - 312	460.7331	919.4516	919.4512	0.0005	0	Y.PGNGTGYVR.H (Ions score 38)
313 - 322	540.7315	1079.4484	1079.4744	-0.0260	0	R.HVDNPNGDGR.C (Ions score 52)
313 - 322	540.7625	1079.5104	1079.4744	0.0360	0	R.HVDNPNGDGR.C (Ions score 55)

313 - 322	541.7443	1081.4740	1079.4744	1.9996	0	R.HVDNPNGDGR.C (Ions score 52)
323 - 344	863.7562	2588.2468	2586.2974	1.9493	2	R.CVTCIYYLNKDWDAKVSGGILR.I Acetyl (K); Dimethyl (K) (Ions score 20)
326 - 344	1136.5719	2271.1292	2269.1929	1.9364	2	T.CIYYLNKDWDAKVSGGILR.I 2 Dimethyl (K) (Ions score 14)
326 - 344	1136.5719	2271.1292	2269.1929	1.9364	2	T.CIYYLNKDWDAKVSGGILR.I 2 Dimethyl (K) (Ions score 16)
328 - 344	685.3721	2053.0945	2053.0996	-0.0052	2	I.YYLNKDWDAKVSGGILR.I 2 Dimethyl (K) (Ions score 44)
345 - 362	722.0505	2163.1297	2163.1364	-0.0067	2	R.IFPEGKAQFADIEPKFDR.L 2 Dimethyl (K) (Ions score 25)
346 - 362	513.5180	2050.0429	2050.0524	-0.0095	2	I.FPEGKAQFADIEPKFDR.L 2 Dimethyl (K) (Ions score 55)
346 - 362	684.3687	2050.0843	2050.0524	0.0319	2	I.FPEGKAQFADIEPKFDR.L 2 Dimethyl (K) (Ions score 46)
347 - 362	635.3375	1902.9907	1902.9839	0.0067	2	F.PEGKAQFADIEPKFDR.L 2 Dimethyl (K) (Ions score 27)
347 - 362	635.9881	1904.9425	1902.9839	1.9585	2	F.PEGKAQFADIEPKFDR.L 2 Dimethyl (K) (Ions score 70)
348 - 362	602.9872	1805.9398	1805.9312	0.0086	2	P.EGKAQFADIEPKFDR.L 2 Dimethyl (K) (Ions score 68)
349 - 362	559.9732	1676.8978	1676.8886	0.0092	2	E.GKAQFADIEPKFDR.L 2 Dimethyl (K) (Ions score 29)
363 - 370	542.2804	1082.5462	1082.5549	-0.0087	0	R.LLFFWSDR.R (Ions score 37)
363 - 371	620.3378	1238.6610	1238.6560	0.0050	1	R.LLFFWSDRR.N (Ions score 20)
363 - 383	651.5873	2602.3201	2602.3193	0.0008	2	R.LLFFWSDRRNPHEVQPAYATR.Y (Ions score 47)
364 - 371	563.7954	1125.5762	1125.5720	0.0043	1	L.LFFWSDRR.N (Ions score 25)
365 - 370	429.2001	856.3856	856.3868	-0.0011	0	L.FFWSDR.R (Ions score 30)
371 - 383	513.5775	1537.7107	1537.7749	-0.0643	1	R.RNPHEVQPAYATR.Y (Ions score 61)
371 - 383	513.5775	1537.7107	1537.7749	-0.0643	1	R.RNPHEVQPAYATR.Y (Ions score 47)
371 - 383	513.9091	1538.7055	1537.7749	0.9305	1	R.RNPHEVQPAYATR.Y (Ions score 61)
371 - 383	513.9091	1538.7055	1537.7749	0.9305	1	R.RNPHEVQPAYATR.Y (Ions score 47)
371 - 383	514.2303	1539.6691	1537.7749	1.8941	1	R.RNPHEVQPAYATR.Y (Ions score 61)
371 - 383	514.2303	1539.6691	1537.7749	1.8941	1	R.RNPHEVQPAYATR.Y (Ions score 48)
371 - 398	679.9314	3394.6206	3394.6595	-0.0389	3	R.RNPHEVQPAYATRYAITVWYFDADERAR.A (Ions score 65)
371 - 398	679.9314	3394.6206	3394.6595	-0.0389	3	R.RNPHEVQPAYATRYAITVWYFDADERAR.A (Ions score 53)
372 - 383	691.8534	1381.6922	1381.6738	0.0184	0	R.NPHEVQPAYATR.Y (Ions score 48)
372 - 383	691.8534	1381.6922	1381.6738	0.0184	0	R.NPHEVQPAYATR.Y (Ions score 50)
372 - 396	1004.8281	3011.4625	3011.4202	0.0423	1	R.NPHEVQPAYATRYAITVWYFDADER.A (Ions score 70)
384 - 396	824.8309	1647.6472	1647.7569	-0.1097	0	R.YAITVWYFDADER.A (Ions score 94)
384 - 396	824.8865	1647.7584	1647.7569	0.0015	0	R.YAITVWYFDADER.A (Ions score 94)
384 - 396	824.8876	1647.7606	1647.7569	0.0037	0	R.YAITVWYFDADER.A (Ions score 54)
384 - 396	824.8878	1647.7610	1647.7569	0.0041	0	R.YAITVWYFDADER.A (Ions score 93)
384 - 396	824.8882	1647.7618	1647.7569	0.0049	0	R.YAITVWYFDADER.A (Ions score 72)
384 - 396	824.8882	1647.7618	1647.7569	0.0049	0	R.YAITVWYFDADER.A (Ions score 88)
384 - 396	824.8883	1647.7620	1647.7569	0.0051	0	R.YAITVWYFDADER.A (Ions score 62)
384 - 396	824.8887	1647.7628	1647.7569	0.0059	0	R.YAITVWYFDADER.A (Ions score 71)
384 - 396	824.8893	1647.7640	1647.7569	0.0071	0	R.YAITVWYFDADER.A (Ions score 59)
384 - 396	824.8907	1647.7668	1647.7569	0.0099	0	R.YAITVWYFDADER.A (Ions score 72)
384 - 396	824.8908	1647.7670	1647.7569	0.0101	0	R.YAITVWYFDADER.A (Ions score 94)
384 - 396	824.8909	1647.7672	1647.7569	0.0103	0	R.YAITVWYFDADER.A (Ions score 74)
384 - 396	824.8909	1647.7672	1647.7569	0.0103	0	R.YAITVWYFDADER.A (Ions score 88)
384 - 396	824.8909	1647.7672	1647.7569	0.0103	0	R.YAITVWYFDADER.A (Ions score 62)
384 - 396	824.8931	1647.7716	1647.7569	0.0147	0	R.YAITVWYFDADER.A (Ions score 61)
384 - 396	824.8955	1647.7764	1647.7569	0.0195	0	R.YAITVWYFDADER.A (Ions score 73)
384 - 396	824.8959	1647.7772	1647.7569	0.0203	0	R.YAITVWYFDADER.A (Ions score 61)
384 - 396	824.8959	1647.7772	1647.7569	0.0203	0	R.YAITVWYFDADER.A (Ions score 78)
384 - 396	824.8971	1647.7796	1647.7569	0.0227	0	R.YAITVWYFDADER.A (Ions score 41)
384 - 398	625.9792	1874.9158	1874.8951	0.0206	1	R.YAITVWYFDADERAR.A (Ions score 80)
397 - 411	440.7751	1759.0713	1759.0468	0.0245	4	R.ARAKVKYLTGEKGVR.V Acetyl (K); Dimethyl (K); Methyl (K) (Ions score 76)
399 - 411	511.6216	1531.8430	1531.9450	-0.1020	3	R.AKVKYLTGEKGVR.V 3 Dimethyl (K) (Ions score 50)
399 - 411	766.9939	1531.9732	1531.9450	0.0283	3	R.AKVKYLTGEKGVR.V 3 Dimethyl (K) (Ions score 67)
399 - 411	511.9545	1532.8417	1531.9450	0.8967	3	R.AKVKYLTGEKGVR.V 3 Dimethyl (K) (Ions score 48)
399 - 411	512.2768	1533.8086	1531.9450	1.8636	3	R.AKVKYLTGEKGVR.V 3 Dimethyl (K) (Ions score 48)
399 - 411	384.5018	1533.9781	1531.9450	2.0331	3	R.AKVKYLTGEKGVR.V 3 Dimethyl (K) (Ions score 53)
412 - 426	845.4429	1688.8712	1688.8985	-0.0273	1	R.VELNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 82)
412 - 426	563.9713	1688.8921	1688.8985	-0.0064	1	R.VELNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 72)
412 - 426	563.9713	1688.8921	1688.8985	-0.0064	1	R.VELNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 76)
412 - 426	563.9730	1688.8972	1688.8985	-0.0013	1	R.VELNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 46)

412 - 426	845.4625	1688.9104	1688.8985	0.0119	1	R.VELNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 43)
412 - 426	852.4526	1702.8906	1702.8778	0.0129	1	R.VELNKPSDSVGKDVF.- Acetyl (K); Dimethyl (K) (Ions score 25)
412 - 426	569.3042	1704.8908	1702.8778	2.0130	1	R.VELNKPSDSVGKDVF.- Acetyl (K); Dimethyl (K) (Ions score 31)
412 - 426	569.3042	1704.8908	1702.8778	2.0130	1	R.VELNKPSDSVGKDVF.- Acetyl (K); Dimethyl (K) (Ions score 24)
413 - 426	795.9063	1589.7980	1589.8301	-0.0320	1	V.ELNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 65)
413 - 426	795.9191	1589.8236	1589.8301	-0.0064	1	V.ELNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 65)
413 - 426	795.9264	1589.8382	1589.8301	0.0082	1	V.ELNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 62)
414 - 426	731.3706	1460.7266	1460.7875	-0.0608	1	E.LNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 57)
414 - 426	731.3793	1460.7440	1460.7875	-0.0434	1	E.LNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 45)
414 - 426	731.4006	1460.7866	1460.7875	-0.0008	1	E.LNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 17)
414 - 426	732.3977	1462.7808	1460.7875	1.9934	1	E.LNKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 46)
415 - 426	674.8587	1347.7028	1347.7034	-0.0006	1	L.NKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 22)
415 - 426	674.8657	1347.7168	1347.7034	0.0134	1	L.NKPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 46)
416 - 426	617.8477	1233.6808	1233.6605	0.0203	1	N.KPSDSVGKDVF.- 2 Dimethyl (K) (Ions score 17)
417 - 426	539.7772	1077.5398	1077.5343	0.0056	1	K.PSDSVGKDVF.- Dimethyl (K) (Ions score 20)

Appendix 8

MS/MS data on tPHD2-SNO tryptic fragments showing peptide mass fingerprints, sequence coverage and MS/MS ion scores.

Match to: **Q9GZT9** (Prolyl hydroxylase domain-containing protein 2) Score: **3243**

Instrument type: ESI-QUAD-TOF; Software used: MASCOT (http://www.matrixscience.com/)

Sequence Coverage: **54%**; Matched peptides shown in **Bold Red**

```
   1 MANDSGGPGG PSPSERDRQY CELCGKMENL LRCSRCRSSF YCCKEHQRQD
  51 WKKHKLVCQG SEGALGHGVG PHQHSGPAPP AAVPPPRAGA REPRKAAARR
 101 DNASGDAAKG KVKAKPPADP AAAASPCRAA AGGQGSAVAA EAEPGKEEPP
 151 ARSSLFQEKA NLYPPSNTPG DALSPGGGLR PNGQTKPLPA LKLALEYIVP
 201 CMNKHGICVV DDFLGKETGQ QIGDEVRALH DTGKFTDGQL VSQKSDSSKD
 251 IRGDKITWIE GKEPGCETIG LLMSSMDDLI RHCNGKLGSY KINGRTKAMV
 301 ACYPGNGTGY VRHVDNPNGD GRCVTCIYYL NKDWDAKVSG GILRIFPEGK
 351 AQFADIEPKF DRLLFFWSDR RNPHEVQPAY ATRYAITVWY FDADERARAK
 401 VKYLTGEKGV RVELNKPSDS VGKDVF
```

The table contains the following terms: **Start – End**, the residues that bracket the peptide sequence in the protein; **Observed**, experimental *m/z* value; **Mr (expt)**, experimental *m/z* transformed to a relative molecular mass; **Mr (calc)**, relative molecular mass calculated from the matched peptide sequence; **Delta**, difference (error) between the experimental and calculated masses; **Miss**, number of missed cleavage sites; **Sequence** of the peptide; **Ions score**, the ion score is a measure of how well the observed MS/MS spectrum matches to the stated peptide. A higher score indicates a more confident match. Modifications found in the peptide are given in the right.

Start - End	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Sequence
180 - 192	711.3680	1420.7214	1418.8358	1.8857	0	L.RPNGQTKPLPALK.L (Ions score 59)
180 - 192	711.3980	1420.7814	1418.8358	1.9457	0	L.RPNGQTKPLPALK.L (Ions score 58)
180 - 192	474.6119	1420.8139	1418.8358	1.9781	0	L.RPNGQTKPLPALK.L (Ions score 29)
180 - 216	816.6242	4078.0846	4077.1730	0.9117	2	L.RPNGQTKPLPALKLALEYIVPCMNKHGICVVDDFLGK.E (Ions score 2)
181 - 192	632.3377	1262.6608	1262.7347	-0.0738	0	R.PNGQTKPLPALK.L (Ions score 73)
181 - 192	632.3377	1262.6608	1262.7347	-0.0738	0	R.PNGQTKPLPALK.L (Ions score 63)
181 - 192	421.8961	1262.6665	1262.7347	-0.0682	0	R.PNGQTKPLPALK.L (Ions score 52)
181 - 192	421.9144	1262.7214	1262.7347	-0.0133	0	R.PNGQTKPLPALK.L (Ions score 14)
181 - 192	422.2364	1263.6874	1262.7347	0.9527	0	R.PNGQTKPLPALK.L (Ions score 14)
181 - 192	633.3868	1264.7590	1262.7347	2.0244	0	R.PNGQTKPLPALK.L (Ions score 73)
181 - 192	633.3868	1264.7590	1262.7347	2.0244	0	R.PNGQTKPLPALK.L (Ions score 63)
181 - 192	633.3873	1264.7600	1262.7347	2.0254	0	R.PNGQTKPLPALK.L (Ions score 73)
182 - 192	583.8546	1165.6946	1165.6819	0.0128	0	P.NGQTKPLPALK.L (Ions score 66)

183 - 192	526.8287	1051.6428	1051.6390	0.0039	0	N.GQTKPLPALK.L (Ions score 10)
183 - 192	526.8325	1051.6504	1051.6390	0.0115	0	N.GQTKPLPALK.L (Ions score 55)
193 - 204	697.3483	1392.6820	1392.7145	-0.0324	0	K.LALEYIVPCMNK.H (Ions score 33)
193 - 216	898.4279	2692.2619	2692.3427	-0.0808	1	K.LALEYIVPCMNKHGICVVDDFLGK.E Oxidation (M) (Ions score 84)
193 - 218	989.1456	2964.4150	2964.4183	-0.0034	2	K.LALEYIVPCMNKHGICVVDDFLGKET.G 2 Nitrosyl (C) (Ions score 19)
193 - 218	989.1456	2964.4150	2964.4183	-0.0034	2	K.LALEYIVPCMNKHGICVVDDFLGKET.G 2 Nitrosyl (C) (Ions score 19)
193 - 218	994.4758	2980.4056	2980.4133	-0.0077	2	K.LALEYIVPCMNKHGICVVDDFLGKET.G Oxidation (M); 2 Nitrosyl (C) (Ions score 12)
198 - 216	716.3452	2146.0138	2144.0581	1.9557	1	Y.IVPCMNKHGICVVDDFLGK.E Carbamidomethyl (C) (Ions score 4)
198 - 216	740.0164	2217.0274	2217.0745	-0.0471	1	Y.IVPCMNKHGICVVDDFLGK.E 2 Carbamidomethyl (C); Oxidation (M) (Ions score 15)
198 - 216	740.0164	2217.0274	2217.0745	-0.0471	1	Y.IVPCMNKHGICVVDDFLGK.E 2 Carbamidomethyl (C); Oxidation (M) (Ions score 14)
198 - 216	1109.5453	2217.0760	2217.0745	0.0016	1	Y.IVPCMNKHGICVVDDFLGK.E 2 Carbamidomethyl (C); Oxidation (M) (Ions score 7)
198 - 216	740.6758	2219.0056	2217.0745	1.9311	1	Y.IVPCMNKHGICVVDDFLGK.E 2 Carbamidomethyl (C); Oxidation (M) (Ions score 15)
200 - 216	631.2824	1890.8254	1890.8791	-0.0537	1	V.PCMNKHGICVVDDFLGK.E Oxidation (M) (Ions score 3)
217 - 227	616.2628	1230.5110	1230.5840	-0.0730	0	K.ETGQQIGDEVR.A (Ions score 89)
217 - 227	616.3056	1230.5966	1230.5840	0.0126	0	K.ETGQQIGDEVR.A (Ions score 48)
217 - 227	616.3149	1230.6152	1230.5840	0.0312	0	K.ETGQQIGDEVR.A (Ions score 68)
217 - 227	617.3117	1232.6088	1230.5840	2.0248	0	K.ETGQQIGDEVR.A (Ions score 89)
218 - 227	551.7832	1101.5518	1101.5415	0.0104	0	E.TGQQIGDEVR.A (Ions score 86)
228 - 234	371.1975	740.3804	740.3817	-0.0013	0	R.ALHDTGK.F (Ions score 36)
228 - 244	922.9527	1843.8908	1843.9428	-0.0520	1	R.ALHDTGKFTDGQLVSQK.S (Ions score 114)
228 - 244	922.9877	1843.9608	1843.9428	0.0180	1	R.ALHDTGKFTDGQLVSQK.S (Ions score 110)
228 - 244	616.3042	1845.8908	1843.9428	1.9479	1	R.ALHDTGKFTDGQLVSQK.S (Ions score 107)
228 - 249	588.0449	2348.1505	2348.1608	-0.0103	2	R.ALHDTGKFTDGQLVSQKSDSSK.D (Ions score 36)
228 - 249	783.7350	2348.1832	2348.1608	0.0223	2	R.ALHDTGKFTDGQLVSQKSDSSK.D (Ions score 27)
230 - 244	554.2808	1659.8206	1659.8217	-0.0011	1	L.HDTGKFTDGQLVSQK.S (Ions score 46)
235 - 244	561.7848	1121.5550	1121.5717	-0.0167	0	K.FTDGQLVSQK.S (Ions score 45)
235 - 244	561.7968	1121.5790	1121.5717	0.0073	0	K.FTDGQLVSQK.S (Ions score 44)
240 - 244	574.3643	573.3570	573.3486	0.0084	0	Q.LVSQK.S (Ions score 17)
245 - 252	454.2286	906.4426	906.4407	0.0020	1	K.SDSSKDIR.G (Ions score 55)
252 - 281	845.6431	3378.5433	3378.6622	-0.1189	3	I.RGDKITWIEGKEPGCETIGLLMSSMDDLIR.H Oxidation (M) (Ions score 11)
252 - 281	845.6431	3378.5433	3378.6622	-0.1189	3	I.RGDKITWIEGKEPGCETIGLLMSSMDDLIR.H Oxidation (M) (Ions score 14)
252 - 281	1127.2160	3378.6262	3378.6622	-0.0360	3	I.RGDKITWIEGKEPGCETIGLLMSSMDDLIR.H Oxidation (M) (Ions score 24)
252 - 281	1127.2160	3378.6262	3378.6622	-0.0360	3	I.RGDKITWIEGKEPGCETIGLLMSSMDDLIR.H Oxidation (M) (Ions score 29)
253 - 262	573.7733	1145.5320	1145.6081	-0.0760	1	R.GDKITWIEGK.E (Ions score 54)
253 - 262	573.7733	1145.5320	1145.6081	-0.0760	1	R.GDKITWIEGK.E (Ions score 57)
253 - 262	574.8124	1147.6102	1145.6081	2.0022	1	R.GDKITWIEGK.E (Ions score 54)
253 - 262	574.8124	1147.6102	1145.6081	2.0022	1	R.GDKITWIEGK.E (Ions score 57)
256 - 262	423.7392	845.4638	845.4647	-0.0008	0	K.ITWIEGK.E (Ions score 23)
257 - 262	733.3788	732.3715	732.3806	-0.0091	0	I.TWIEGK.E (Ions score 35)
257 - 281	942.4772	2824.4098	2822.3289	2.0809	1	I.TWIEGKEPGCETIGLLMSSMDDLIR.H Nitrosyl (C) (Ions score 1)
259 - 281	1255.1420	2508.2694	2506.2117	2.0577	1	W.IEGKEPGCETIGLLMSSMDDLIR.H (Ions score 25)
260 - 281	1197.5845	2393.1544	2393.1276	0.0268	1	I.EGKEPGCETIGLLMSSMDDLIR.H (Ions score 57)
263 - 281	1054.9967	2107.9788	2107.9588	0.0201	0	K.EPGCETIGLLMSSMDDLIR.H Nitrosyl (C) (Ions score 106)
263 - 281	1055.9476	2109.8806	2107.9588	1.9219	0	K.EPGCETIGLLMSSMDDLIR.H Nitrosyl (C) (Ions score 101)
265 - 281	942.9673	1883.9200	1881.8634	2.0566	0	P.GCETIGLLMSSMDDLIR.H Nitrosyl (C) (Ions score 28)
270 - 291	625.0629	2496.2225	2494.2130	2.0095	2	I.GLLMSSMDDLIRHCNGKLGSYK.I Carbamidomethyl (C) (Ions score 0)
270 - 295	588.2952	2936.4396	2934.4626	1.9770	3	I.GLLMSSMDDLIRHCNGKLGSYKINGR.T Carbamidomethyl (C) (Ions score 4)
272 - 281	590.7895	1179.5644	1179.5628	0.0017	0	L.LMSSMDDLIR.H (Ions score 55)
272 - 281	1182.5067	1181.4994	1179.5628	1.9367	0	L.LMSSMDDLIR.H (Ions score 56)
272 - 281	1182.6880	1181.6807	1179.5628	2.1180	0	L.LMSSMDDLIR.H (Ions score 56)
273 - 281	542.2487	1082.4828	1082.4736	0.0092	0	L.MSSMDDLIR.H Oxidation (M) (Ions score 69)

274 - 281	936.4144	935.4071	935.4382	-0.0311	0	M.SSMDDLIR.H (Ions score 33)
274 - 281	936.6179	935.6106	935.4382	0.1724	0	M.SSMDDLIR.H (Ions score 33)
282 - 291	568.3164	1134.6182	1134.5240	0.0942	1	R.HCNGKLGSYK.I Nitrosyl (C) (Ions score 22)
286 - 312	738.1065	2948.3969	2948.4749	-0.0780	4	G.KLGSYKINGRTKAMVACYPGNGTGYVR.H Oxidation (M); Nitrosyl (C) (Ions score 8)
298 - 312	779.8635	1557.7124	1557.7068	0.0056	0	K.AMVACYPGNGTGYVR.H (Ions score 51)
298 - 312	780.8588	1559.7030	1557.7068	1.9962	0	K.AMVACYPGNGTGYVR.H (Ions score 66)
298 - 312	780.8754	1559.7362	1557.7068	2.0294	0	K.AMVACYPGNGTGYVR.H (Ions score 84)
298 - 312	787.8619	1573.7092	1573.7017	0.0075	0	K.AMVACYPGNGTGYVR.H Oxidation (M) (Ions score 57)
298 - 312	794.3624	1586.7102	1586.6970	0.0133	0	K.AMVACYPGNGTGYVR.H Nitrosyl (C) (Ions score 31)
298 - 312	809.4268	1616.8390	1614.7283	2.1108	0	K.AMVACYPGNGTGYVR.H Carbamidomethyl (C) (Ions score 29)
302 - 312	622.3110	1242.6074	1242.5452	0.0623	0	A.CYPGNGTGYVR.H Carbamidomethyl (C) (Ions score 36)
304 - 312	920.4657	919.4584	919.4512	0.0072	0	Y.PGNGTGYVR.H (Ions score 38)
313 - 322	540.7448	1079.4750	1079.4744	0.0006	0	R.HVDNPNGDGR.C (Ions score 58)
313 - 322	541.2386	1080.4626	1079.4744	0.9882	0	R.HVDNPNGDGR.C (Ions score 40)
313 - 322	541.2396	1080.4646	1079.4744	0.9902	0	R.HVDNPNGDGR.C (Ions score 52)
323 - 332	610.2957	1218.5768	1218.5777	-0.0008	0	R.CVTCIYYLNK.D (Ions score 54)
323 - 337	917.9194	1833.8242	1833.8430	-0.0187	1	R.CVTCIYYLNKDWDAK.V (Ions score 55)
325 - 332	509.2588	1016.5030	1016.5001	0.0030	0	V.TCIYYLNK.D (Ions score 8)
346 - 350	577.3040	576.2967	576.2908	0.0060	0	I.FPEGK.A (Ions score 19)
351 - 359	509.7524	1017.4902	1017.5131	-0.0228	0	K.AQFADIEPK.F (Ions score 33)
351 - 359	509.7524	1017.4902	1017.5131	-0.0228	0	K.AQFADIEPK.F (Ions score 34)
351 - 359	510.2474	1018.4802	1017.5131	0.9672	0	K.AQFADIEPK.F (Ions score 33)
351 - 359	510.7542	1019.4938	1017.5131	1.9808	0	K.AQFADIEPK.F (Ions score 33)
352 - 359	474.2451	946.4756	946.4760	-0.0003	0	A.QFADIEPK.F (Ions score 28)
360 - 370	751.3652	1500.7158	1500.7514	-0.0355	1	K.FDRLLFFWSDR.R (Ions score 56)
360 - 370	751.3652	1500.7158	1500.7514	-0.0355	1	K.FDRLLFFWSDR.R (Ions score 56)
360 - 370	751.3883	1500.7620	1500.7514	0.0107	1	K.FDRLLFFWSDR.R (Ions score 22)
360 - 371	553.2817	1656.8233	1656.8525	-0.0292	2	K.FDRLLFFWSDRR.N (Ions score 33)
360 - 371	553.2828	1656.8266	1656.8525	-0.0259	2	K.FDRLLFFWSDRR.N (Ions score 45)
360 - 371	553.9444	1658.8114	1656.8525	1.9589	2	K.FDRLLFFWSDRR.N (Ions score 35)
360 - 383	756.1389	3020.5265	3020.5158	0.0107	3	K.FDRLLFFWSDRRNPHEVQPAYATR.Y (Ions score 20)
361 - 371	504.2669	1509.7789	1509.7841	-0.0052	2	F.DRLLFFWSDRR.N (Ions score 30)
363 - 370	542.2875	1082.5604	1082.5549	0.0055	0	R.LLFFWSDR.R (Ions score 25)
363 - 371	620.3314	1238.6482	1238.6560	-0.0078	1	R.LLFFWSDRR.N (Ions score 27)
363 - 383	651.5777	2602.2817	2602.3193	-0.0376	2	R.LLFFWSDRRNPHEVQPAYATR.Y (Ions score 48)
363 - 383	651.5777	2602.2817	2602.3193	-0.0376	2	R.LLFFWSDRRNPHEVQPAYATR.Y (Ions score 45)
363 - 383	652.0848	2604.3101	2602.3193	1.9908	2	R.LLFFWSDRRNPHEVQPAYATR.Y (Ions score 48)
363 - 383	652.0848	2604.3101	2602.3193	1.9908	2	R.LLFFWSDRRNPHEVQPAYATR.Y (Ions score 45)
364 - 371	563.7877	1125.5608	1125.5720	-0.0111	1	L.LFFWSDRR.N (Ions score 23)
371 - 383	513.5973	1537.7701	1537.7749	-0.0049	1	R.RNPHEVQPAYATR.Y (Ions score 23)
371 - 383	769.8997	1537.7848	1537.7749	0.0099	1	R.RNPHEVQPAYATR.Y (Ions score 76)
371 - 383	513.9028	1538.6866	1537.7749	0.9116	1	R.RNPHEVQPAYATR.Y (Ions score 62)
371 - 383	514.2218	1539.6436	1537.7749	1.8686	1	R.RNPHEVQPAYATR.Y (Ions score 63)
371 - 396	792.8862	3167.5157	3167.5213	-0.0056	2	R.RNPHEVQPAYATRYAITVWYFDADER.A (Ions score 39)
372 - 383	691.8258	1381.6370	1381.6738	-0.0368	0	R.NPHEVQPAYATR.Y (Ions score 54)
372 - 383	691.8258	1381.6370	1381.6738	-0.0368	0	R.NPHEVQPAYATR.Y (Ions score 54)
372 - 383	692.3485	1382.6824	1381.6738	1.0086	0	R.NPHEVQPAYATR.Y (Ions score 33)
372 - 383	692.8483	1383.6820	1381.6738	2.0082	0	R.NPHEVQPAYATR.Y (Ions score 55)
372 - 398	810.6487	3238.5657	3238.5584	0.0073	2	R.NPHEVQPAYATRYAITVWYFDADERAR.A (Ions score 64)
375 - 383	517.7437	1033.4728	1033.5192	-0.0464	0	H.EVQPAYATR.Y (Ions score 14)
379 - 398	813.7367	2438.1883	2437.1815	1.0068	2	P.AYATRYAITVWYFDADERAR.A (Ions score 28)

384 - 396	824.8828	1647.7510	1647.7569	-0.0059	0	R.YAITVWYFDADER.A (Ions score 94)
384 - 396	824.8840	1647.7534	1647.7569	-0.0035	0	R.YAITVWYFDADER.A (Ions score 94)
384 - 396	824.8859	1647.7572	1647.7569	0.0003	0	R.YAITVWYFDADER.A (Ions score 59)
384 - 396	824.8870	1647.7594	1647.7569	0.0025	0	R.YAITVWYFDADER.A (Ions score 86)
384 - 396	824.8878	1647.7610	1647.7569	0.0041	0	R.YAITVWYFDADER.A (Ions score 94)
384 - 396	824.8881	1647.7616	1647.7569	0.0047	0	R.YAITVWYFDADER.A (Ions score 94)
384 - 396	824.8884	1647.7622	1647.7569	0.0053	0	R.YAITVWYFDADER.A (Ions score 78)
384 - 396	824.8885	1647.7624	1647.7569	0.0055	0	R.YAITVWYFDADER.A (Ions score 77)
384 - 396	824.8887	1647.7628	1647.7569	0.0059	0	R.YAITVWYFDADER.A (Ions score 94)
384 - 396	824.8889	1647.7632	1647.7569	0.0063	0	R.YAITVWYFDADER.A (Ions score 98)
384 - 396	824.8897	1647.7648	1647.7569	0.0079	0	R.YAITVWYFDADER.A (Ions score 93)
384 - 396	824.8901	1647.7656	1647.7569	0.0087	0	R.YAITVWYFDADER.A (Ions score 64)
384 - 396	824.8903	1647.7660	1647.7569	0.0091	0	R.YAITVWYFDADER.A (Ions score 58)
384 - 396	824.8903	1647.7660	1647.7569	0.0091	0	R.YAITVWYFDADER.A (Ions score 94)
384 - 396	824.8908	1647.7670	1647.7569	0.0101	0	R.YAITVWYFDADER.A (Ions score 63)
384 - 396	824.8910	1647.7674	1647.7569	0.0105	0	R.YAITVWYFDADER.A (Ions score 94)
384 - 396	824.8914	1647.7682	1647.7569	0.0113	0	R.YAITVWYFDADER.A (Ions score 93)
384 - 396	824.8915	1647.7684	1647.7569	0.0115	0	R.YAITVWYFDADER.A (Ions score 94)
384 - 396	824.8916	1647.7686	1647.7569	0.0117	0	R.YAITVWYFDADER.A (Ions score 94)
384 - 396	824.8920	1647.7694	1647.7569	0.0125	0	R.YAITVWYFDADER.A (Ions score 76)
384 - 396	824.8930	1647.7714	1647.7569	0.0145	0	R.YAITVWYFDADER.A (Ions score 76)
384 - 396	824.8961	1647.7776	1647.7569	0.0207	0	R.YAITVWYFDADER.A (Ions score 60)
384 - 396	824.8961	1647.7776	1647.7569	0.0207	0	R.YAITVWYFDADER.A (Ions score 75)
384 - 396	824.8994	1647.7842	1647.7569	0.0273	0	R.YAITVWYFDADER.A (Ions score 77)
384 - 398	625.9706	1874.8900	1874.8951	-0.0052	1	R.YAITVWYFDADERAR.A (Ions score 73)
384 - 398	938.4619	1874.9092	1874.8951	0.0141	1	R.YAITVWYFDADERAR.A (Ions score 35)
386 - 398	821.4039	1640.7932	1640.7947	-0.0014	1	A.ITVWYFDADERAR.A (Ions score 23)
391 - 398	980.4955	979.4882	978.4519	1.0363	1	Y.FDADERAR.A (Ions score 21)
391 - 398	980.6317	979.6244	978.4519	1.1725	1	Y.FDADERAR.A (Ions score 21)
401 - 408	469.2704	936.5262	936.5280	-0.0018	1	K.VKYLTGEK.G (Ions score 41)
412 - 423	424.8859	1271.6359	1271.6721	-0.0362	0	R.VELNKPSDSVGK.D (Ions score 78)
412 - 423	637.3384	1272.6622	1271.6721	0.9901	0	R.VELNKPSDSVGK.D (Ions score 34)
412 - 426	545.2667	1632.7783	1632.8359	-0.0576	1	R.VELNKPSDSVGKDVF.- (Ions score 70)
412 - 426	545.2667	1632.7783	1632.8359	-0.0576	1	R.VELNKPSDSVGKDVF.- (Ions score 71)
412 - 426	817.4041	1632.7936	1632.8359	-0.0423	1	R.VELNKPSDSVGKDVF.- (Ions score 73)
412 - 426	817.9215	1633.8284	1632.8359	0.9925	1	R.VELNKPSDSVGKDVF.- (Ions score 23)
413 - 423	587.2755	1172.5364	1172.6037	-0.0673	0	V.ELNKPSDSVGK.D (Ions score 83)
413 - 423	587.2755	1172.5364	1172.6037	-0.0673	0	V.ELNKPSDSVGK.D (Ions score 83)
413 - 423	587.2958	1172.5770	1172.6037	-0.0267	0	V.ELNKPSDSVGK.D (Ions score 69)
413 - 423	588.3171	1174.6196	1172.6037	2.0159	0	V.ELNKPSDSVGK.D (Ions score 67)
413 - 423	588.3190	1174.6234	1172.6037	2.0197	0	V.ELNKPSDSVGK.D (Ions score 82)
413 - 423	588.3190	1174.6234	1172.6037	2.0197	0	V.ELNKPSDSVGK.D (Ions score 81)
413 - 426	767.8931	1533.7716	1533.7675	0.0042	1	V.ELNKPSDSVGKDVF.- (Ions score 39)
414 - 423	522.7775	1043.5404	1043.5611	-0.0207	0	E.LNKPSDSVGK.D (Ions score 53)
414 - 426	703.3619	1404.7092	1404.7249	-0.0157	1	E.LNKPSDSVGKDVF.- (Ions score 43)
415 - 423	466.2470	930.4794	930.4771	0.0024	0	L.NKPSDSVGK.D (Ions score 21)
416 - 423	409.2224	816.4302	816.4341	-0.0039	0	N.KPSDSVGK.D (Ions score 38)

Appendix 9

List of lysis buffer conditions used for PHD3 expression trials.

	Buffer	Salt	Stabiliser/ reducing agent	Chelating agent	Detergent/ Others	Final pH
1.	50 mM Tris –HCl	-	-	-	-	7.5
2.	100 mM Tris-HCl	-	10% (v/v) glycerol	-	-	7.5
3.	100 mM Tris-HCl	50 mM LiCl	-	-	-	7.5
4.	100 mM Tris-HCl	50 mM NaCl, 5 mM Ca-acetate	-	-	-	7.5
5.	100 mM Tris-HCl	1 M $(NH_4)_2SO_4$	10 mM DTT	-	-	8.2
6.	100 mM Tris-HCl	50 mM NaCl	-	-	100 mM Urea	8.2
7.	100 mM Tris-HCl	100 mM KCl	-	2 mM EDTA	1% Triton X-100	8.2
8.	100 mM Tris-HCl	100 mM NaCl	-	1 mM EDTA	-	8.5
9.	100 mM Tris-HCl	50 mM KCl	25% (v/v) glycerol	-	0.1% (v/v) Deoxycholate	7.5
10.	100 mM Tris-HCl	50 mM NaCl	-	-	10% isopropanol	8.2
11.	100 mM Tris-HCl	100 mM Na glutamate	10 mM DTT	-	-	8.2
12.	100 mM HEPES-Na	100 mM KCl	-	-	-	7.0
13.	100 mM HEPES-Na	50 mM (NH4)2SO4	10% glycerol	-	-	7.0
14.	100 mM HEPES-Na	50 mM LiCl	-	-	0.1% CHAPS	7.0
15.	100 mM HEPES-Na	100 mM Na glutamate	5 mM DTT	-	-	7.0
16.	100 mM HEPES-Na	50 mM LiCl	-	-	0.1% Deoxycholate	7.0
17.	100 mM HEPES-Na	1 M MgSO4	-	-	-	7.0
18.	250 mM NaH2PO4/ Na2HPO4	-	-	-	0.1% CHAPS	6.0
19.	100 mM Na_2HPO4/ NaH2PO4	50 mM (NH4)2SO4	-	-	1% Triton X-100	6.0
20.	100 mM Triethanolamine	100 mM KCl	10 mM DTT	-	-	8.5
21.	100 mM Triethanolamine	50 mM LiCl	-	5 mM EDTA	-	8.5
22.	100 mM Triethanolamine	50 mM (NH4)2SO4, 10 mM MgSO4	-	-	-	8.5
23.	100 mM Triethanolamine	100 mM sodium glutamate	10% glycerol	-	0.02% n-octyl-ß-D-glucoside	8.5
24.	30 mM MES	10 mM NaCl, 5 mM $CaCl_2$	5 mM DTT	-	0.25% (w/v) Brij-35, 0.005% (w/v) BSA, 3% (v/v) DMSO	6.0

Appendix 10

Summary of the PHD3 constructs/ proteins produced in the current study.

Construct/Protein	**Expression Vector**	**Soluble expression level (%)**	**Purification method**	**Protein yield***
PHD3-NH (1-239)	pET28a	<<1	-	-
PHD3-CH (1-239)	pET29b	~ 1	IMAC (Ni^{2+}/ Co^{2+})	0.7
PHD3-NH (1-226)	pET28a	<<1	-	-
PHD3-NH (13-239)	pET28a	<<1	-	-
PHD3-NH (13-232)	pET28a	<<1	-	-
PHD3-NH (13-226)	pET28a	<<1	-	-
PHD3-NH (13-215)	pET28a	<<1	-	-
PHD3-NH C42Q	pET28a	~1	IMAC (Ni^{2+})	0.7
PHD3-CH C42Q	pET29b	2	IMAC (Ni^{2+})	0.7
PHD3-CH C42Q/C52D (dcys)	pET29b	>2	IMAC (Ni^{2+})	1.0
PHD3-NH ΔR71	pET28a	<<1	-	-
PHD3-NH ΔH72	pET28a	<<1	-	-
PHD3-NH C29A	pET28a	<<1	-	-
PHD3-CH dcys R57T	pET29b	>2	-	-
PHD3-CH dcys Y30Q	pET29b	>2	-	-
PHD3-CH dcys V38E	pET29b	>2	-	-
PHD3-CH dcys H72D	pET29b	>2	-	-
PHD3-CH dcys E45D/K48R/R65K/R100D	pET29b	>3	-	-
PHD3-CH dcys E45D/K48R/R65K/R100D/L103R/Y104H	pET29b	>3	-	-
PHD3-CH dcys E45D/K48R/R65K/R100D/L103R/Y104H/E 220R	pET29b	>3	-	-
PHD3-CH dcys E45D/K48R/R65K/R100D/L103R/Y104H/E 216D /E220R	pET29b	>4	-	-
PHD3-CH dcys E45D/K48R/R65K/R100D/L103R/Y104H/E 220R /K223V	pET29b	>4	-	-
PHD3 dcys E45D/K48R/R65K/R100D/L103R/Y104H/E 216D/E220R (**MM-3**)	pET29b	>5	-	-
PHD3-CH R65K/H72D/K111S/Y113K	pET29b	>1	-	-
PHD3-CH R57T/R65K/H72D/K111S/Y113K	pET29b	>1	-	-
PHD3-CH R65K/H72D/K111S/Y113K/E220R	pET29b	>1	-	-
PHD3-CH R57T/R65K/H72D/K111S/Y113K/E220R/K 223V (**MM-1**)	pET29b	>3	-	-
PHD3-CH R57T/R65K/H72D/K111S/Y113K/E216D/E 220R	pET29b	>3	-	-
PHD3-CH R57T/R65K/H72D/K111S/Y113K/E216D/E 220R/K223V (**MM-2**)	pET29b	>3	-	-

* Average protein yield (>90% homogeneity) was calculated as mg of protein per g of wet cell pellet.

Appendix 11

List of primers used for generation of tPHD2 variants.

Mutant	Primer	Sequence (5' to 3')
C201A	forward	TAC ATC GTG CCG GCC ATG AAC AAG CAC GGC
	reverse	GTG CTT GTT CAT GGC CGG CAC GAT GTA CTC
C208A	forward	AAG CAC GGC ATC GCT GTG GTG GAC GAC TTC
	reverse	GTC GTC CAC CAC AGC GAT GCG CCG TGC TTG TTC
A228Q	forward	GAC GAG GTG CGC CAA CTG CAC GAC ACC GGG
	reverse	GGT GTC GTG CAG TTG GCG CAC CTC GTC GCC
K234A/T236R	forward	CTG CAC GAC ACC GGG GCG TTC AGG GAC GGG CAG CTG
	reverse	CAG CTG CCC GTC CCT GAA CGC CCC GGT GTC GTG CAG
T236A	forward	CC GGG AAG TTC GCG GAC GGG CAG CTG
	reverse	G CTG CCC GTC CGC GAA CTT CCC GGT G
S242G	forward	GAC GGG CAG CTG GTC GGC CAG AAG AGT GAC TCG
	reverse	CGA GTC ACT CTT CTG GCC GAC CAG CTG CCC GTC
K244R	forward	CAG CTG GTC AGC CAG AGG AGT GAC TCG TCC AAG
	reverse	CTT GGA CGA GTC ACT CCT CTG GCT GAC CAG CTG
S242G/K244R	forward	GAC GGG CAG CTG GTC GGC CAG AGG AGT GAC TCG
	reverse	CGA GTC ACT CCT CTG GCC GAC CAG CTG CCC GTC
I251L	forward	GAC TCG TCC AAG GAC CTC CGA GGC GAT AAG
	reverse	GAT CTT ATC GCC TCG GAG GTC CTT GGA CG
R252A	forward	CG TCC AAG GAC ATC GCA GGC GAT AAG ATC ACC
	reverse	CA GGT GAT CTT ATC GCC TGC GAT GTC CTT GGA C
R252C	forward	CG TCC AAG GAC ATC TGT GGC GAT AAG ATC ACC
	reverse	CA GGT GAT CTT ATC GCC ACA GAT GTC CTT GGA C
R252C/D254C	forward	CG TCC AAG GAC ATC TGT GGC TGT AAG ATC ACC
	reverse	CA GGT GAT CTT ACA GCC ACA GAT GTC CTT GGA C
D254A	forward	G GAC ATC CGA GGC GCT AAG ATC ACC TGG ATC
	reverse	CCA GGT GAT CTT AGC GCC TCG GAT GTC CTT G
D254K	forward	G GAC ATC CGA GGC AAA AAG ATC ACC TGG ATC
	reverse	C GAT CCA GGT GAT CTT TTT GCC TCG GAT GTC
K255I	forward	C ATC CGA GGC GAT ATG ATC ACC TGG ATC GAG G
	reverse	C CTC GAT CCA GGT GAT CAT ATC GCC TCG GAT G
P264E	forward	C GAG GGC AAG GAG GAA GGC TGC GAA ACC
	reverse	GGT TTC GCA GCC TTC CTC CTT GCC CTC G
C266A	forward	AAG GAG CCC GGC GCC GAA ACC ATT GGG CTG
	reverse	CCC AAT GGT TTC GGC GCC GGG CTC CTT GCC
I280C	forward	GC ATG GAC GAC CTG TGT CGC CAC TGT AAC GGG
	reverse	CTT CCC GTT ACA GTG GCG GCA CAG GTC GTC CAT G
R281L	forward	G GAC GAC CTG ATA CTC CAC TGT AAC GGG AAG
	reverse	CCC GTT ACA GTG GAG TAT CAG GTC GTC CAT G
C283A	forward	CTG ATA CGC CAC GCT AAC GGG AAG CTG GGC
	reverse	CAG CTT CCC GTT AGC GTG GCG TAT CAG GTC
S289K	forward	GGG AAG CTG GGC AAA TAC AAA ATC AAT GGC
	reverse	CC ATT GAT TTT GTA TTT GCC CAG CTT CCC G
K297I	forward	C AAT GGC CGG ACG ATA GCC ATG GTC GCT TG
	reverse	CA AGC GAC CAT GGC TAT CGT CCG GCC ATT G
M299V	forward	CGG ACG AAA GCC GTG GTT GCT TGT TAT CCG
	reverse	GG ATA ACA AGC AAC CAC GGC TTT CGT CCG G
M299E	forward	CGG ACG AAA GCC GAG GTT GCT TGT TAT CCG
	reverse	GG ATA ACA AGC AAC CTC GGC TTT CGT CCG G
C302A	forward	G AAA GCC ATG GTT GCT GCT TAT CCG GGC AAT GG
	reverse	CGT TCC ATT GCC CGG ATA AGC AGC AAC CAT GGC
Y303F	forward	GCC ATG GTT GCT TGT TTT CCG GGC AAT GGA AC
	reverse	GT TCC ATT GCC CGG AAA ACA AGC AAC CAT GGC
Y303A	forward	GCC ATG GTT GCT TGT GCT CCG GGC AAT GGA AC
	reverse	GT TCC ATT GCC CGG AGC ACA AGC AAC CAT G

Mutant	**Primer**	**Sequence (5’ to 3’)**
Y310F	forward	GGC AAT GGA ACG GGT TTT GTA CGT CAT GTT GAT
	reverse	C AAC ATG ACG TAC AAA ACC CGT TCC ATT GCC
P317R	forward	CGT CAT GTT GAT AAT CGA AAT GGA GAT GGA AG
	reverse	CA TCT TCC ATC TCC ATT TCG ATT ATC AAC ATG
P317E	forward	CGT CAT GTT GAT AAT GAA AAT GGA GAT GG
	reverse	CT TCC ATC TCC ATT TTC ATT ATC AAC ATG
C323A	forward	CCA AAT GGA GAT GGA AGA GCT GTG ACA TGT ATA TAT TAT C
	reverse	G ATA ATA TAT ACA TGT CAC AGC TCT TCC ATC TCC ATT TGG
T325A	forward	GAT GGA AGA TGT GTG GCA TGT ATA TAT TAT C
	reverse	G ATA ATA TAT ACA TGC CAC ACA TCT TCC ATC
C326A	forward	GGA AGA TGT GTG ACA GCT ATA TAT TAT CTT AAT AAA G
	reverse	CTT TAT TAA GAT AAT ATA TAG CTG TCA CAC ATC TTC C
Y329F	forward	GTG ACA TGT ATA TAT TTT CTT AAT AAA GAC
	reverse	C CCA GTC TTT ATT AAG AAA ATA TAT ACA TG
F366Y	forward	GA CTG CTG TTT TAC TGG TCT GAC CGT CGC
	reverse	G ACG GTC AGA CCA GTA AAA CAG CAG TCT ATC
R371H	forward	GG TCT GAC CGT CAC AAC CCT CAT GAA GTA CAA CC
	reverse	G TAC TTC ATG AGG GTT GTG ACG GTC AGA CC
R383A	forward	CCA GCA TAT GCT ACA GCG TAC GCA ATA ACT G
	reverse	C AGT TAT TGC GTA CGC TGT AGC ATA TGC TGG
T387A	forward	CA AGG TAC GCA ATA GCT GTT TGG TAT TTT G
	reverse	GC ATC AAA ATA CCA AAC AGC TAT TGC GTA CC
R396A	forward	GAT GCA GAT GAG GCA GCA CGA GCT AAA G
	reverse	C TTT AGC TCG TGC TGC CTC ATC TGC ATC
D394A	forward	GG TAT TTT GAT GCA GCT GAG AGA GCA CG
	reverse	GC TCG TGC TCT CTC AGC TGC ATG AAA ATA C
R398A	forward	GAT GAG AGA GCA GCA GCT AAA GTA AAA TAT C
	reverse	G ATA TTT TAC TTT AGC TGC TGC TCT CTC ATC TGC
V401A	forward	GCA CGA GCT AAA GCA AAA TAT CTA ACA G
	reverse	GT TAG ATA TTT TGC TTT AGC TCG TGC TC
loopΔDAAI	forward	GAC ACC GGG AAG TTC ACG GAC GCA GCA ATC CGA GGC GAT AAG ATC ACC TGG
	reverse	CCA GGT GAT CTT ATC GCC TCG GAT TGC TGC GTC CGT GAA CTT CCC GGT GTC

Appendix 12

List of primers used for generation of PHD3 variants.

Wt/Variant	Primer	Sequence (5' to 3')
PHD3-CH	forward	GAT ATA CAT ATG CCC CTG GGA CAC ATC ATG AGG CTG GAC CTG
	reverse	GTA ATT CTC GAG GGA ACC GCG TGG CAC CAG GTC TTC AGT GAG
PHD3-NH (13-239)	forward	CTG CCG CAT ATG AAA ATT GCC CTG GAG TAC
	reverse	GTG GTG CTC GAG TCA GTC TTC AGT GAG GGC
PHD3-NH (13-232)	forward	CT AGG AAA ACT TAA TCT GCC CTC ACT G
	reverse	C TTC AGT GAG GGC AGA TTA AGT TTT CCT AG
PHD3-NH (13-226)	forward	GCC AAA AAG AAA TTC AGG TAG TTA ACT AGG AAA AC
	reverse	C AGT TTT CCT AGT TAA CTA CCT GAA TTT CTT TTT G
PHD3-NH (13-215)	forward	G TAC TTT GAT GCT TAA GAA AGG GCA GAA G
	reverse	C TTC TGC CCT TTC TTA AGC ATC AAA GTA CCA G
PHD3 C29A	forward	CAC GAG GTG GGC TTC GCC TAC CTG GAC AAC TTC
	reverse	GAA GTT GTC CAG GTA GGC GAA GCC CAC CTC GTG
PHD3 Y30Q	forward	GTG GGC TTC TGC CAG CTG GAC AAC TTC CTG
	reverse	GAA GTT GTC CAG CTC GCA GAA GCC CAC CTC
PHD3 V38E/C42Q	forward	C TTC CTG GGC GAG GAG GTG GGC GAC CAA
	reverse	GAC TTG GTC GCC CAC CTC CTC GCC CAG
PHD3 C42Q	forward	GAG GTG GTG GGC GAC CAA GTC CTG GAG CGC GTC
	reverse	GAC GCG CTC CAG GAC TTG GTC GCC CAC CAC CTC
PHD3 E45D/K48R	forward	GGC GAC CAA GTC CTG GAC CGC GTC AGG CAG CTG CAC GAC ACC
	reverse	GGT GTC GTG CAG CTG CCT GAC GCG GTC CAG GAC TTG GTC GCC
PHD3 C52D	forward	C AAG CAG CTG CAC GAC ACC GGG GCC CTG CGG
	reverse	CCG CAG GGC CCC GGT GTC GTG CAG CTG CTT G
PHD3 R57T	forward	ACC GGG GCC CTG ACG GAC GGC CAG CTG GCG GG
	reverse	CC CGC CAG CTG GCC GTC CGT CAG GGC CCC GGT
PHD3 R65K	forward	CAG CTG GCG GGG CCG AAG GCC GGC GTC TCC AAG
	reverse	CTT GGA GAC GCC GGC CTT CGG CCC CGC CAG CTG
PHD3 ΔR71	forward	GCC GGC GTC TCC AAG CAC CTG CGG GGC GAC
	reverse	GTC GCC CCG CAG GTG CTT GGA GAC GCC GGC
PHD3 ΔH72	forward	GGC GTC TCC AAG CGA CTG CGG GGC GAC CAG
	reverse	CTG GTC GCC CCG CAG TCG CTT GGA GAC GCC
PHD3 H72D	forward	GGC GTC TCC AAG CGA GAC CTG CGG GGC GAC
	reverse	GTC GCC CCG CAG GTC TCG CTT GGA GAC GCC
PHD3 R100D/L103R/ Y104H	forward	G TCC CTC ATC GAC GAC CTG GTC CGC CAC TGC
	reverse	GCA GTG GCG GAC GAG GTC GTC GAT GAG GGA C
PHD3 L103R/Y104H	forward	CC CTC ATC GAC AGG CTG GTC CGC CAC TGC GGG AGC CGG CTG GGC
	reverse	GCC CAG CCG GCT CCC GCA GTG GCG GAC CAG CCT GTC GAT GAG
PHD3 K111S/Y113K	forward	GGG AGC CGG CTG GGC AGT TAC AAG GTC AAG GAG AGG TCT AAG
	reverse	CTT AGA CCT CTC CTT GAC CTT GTA ACT GCC CAG CCG GCT CCC
PHD3 E216D/E220R	forward	GG TAC TTT GAT GCT GAT GAA AGG GCA AGA
	reverse	TCT TGC CCT TTC ATC AGC ATC AAA GTA CC
PHD3 E220R/K223V	forward	GCA AGA GCC AAA GTG AAA TTC AGG AAT TTA AC
	reverse	GT TAA ATT CCT GAA TTT CAC TTT GGC TCT TGC
PHD3 E220R	forward	GAA GAA AGG GCA AGA GCC AAA AAG AAA TTC
	reverse	GAA TTT CTT TTT GGC TCT TGC CCT TTC TTC
PHD3K223V	forward	GCA GAA GCC AAA GTG AAA TTC AGG AAT TTA AC
	reverse	GT TAA ATT CCT GAA TTT CAC TTT GGC TTC TGC