Evolutionary Insights into Oxygen Sensing



Tongri Liu

Corpus Christi College, Oxford

Trinity Term, 2018

A thesis submitted to the Board of the Material, Physical and Life Sciences Division at the University of Oxford in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

Abstract: Evolutionary Insights into Oxygen Sensing

Tongri Liu, Corpus Christi College **Trinity Term, 2018**

In humans, cellular oxygen sensing employs the Hypoxia Inducible Factor (HIF) system in which the prolyl hydroxylase domain containing (PHD) and Factor Inhibiting HIF (FIH) oxygenases play a key role. The oxygen sensing mechanism enabled by the HIF hydroxylases has been extensively studied in humans and other mammals; how 'lower' organisms, such as protists and bacteria sense oxygen is less well understood. The social amoeba, *Dictyostelium discoideum*, and the protozoan parasite, *Toxoplasma gondii*, are representatives of pre-animal eukaryotes that do not possess the PHD-HIF oxygen sensing system. A different, yet HIF related, mechanism of cellular oxygen sensing regulates the growth of these species, i.e. the Prolyl Hydroxylase (PHD/P4H/PhyA) and S-Phase Kinase Associated Protein 1 (Skp1) system. Skp1 acts as an adaptor protein of the SCF (Skp1-Cul1-F-box) E3-ubiquitin ligase complex. In *Dictyostelium discoideum* and *Toxoplasma gondii* Skp1 is prolyl-4-hydroxylated and subsequently glycosylated, resulting in a change of stability of the SCF E3-ubiquitin ligase complex, that in turn affects the activity of the ubiquitination complex.

A major part of the work described in this thesis focused on the *in vitro* characterisation of the enzymes, *Dictyostelium discoideum* Prolyl-4-Hydroxylase (DdP4H) and *Toxoplasma gondii* Prolyl-Hydroxylase A (TgPhyA), which catalyse hydroxylation of Skp1 protein substrates, i.e. *Dictyostelium discoideum* Skp1 (DdSkp1) and *Toxoplasma gondii* Skp1 (TgSkp1). Methods were optimised for the production of recombinant full-length and truncated forms of DdP4H and TgPhyA on a milligram scale. Purified DdP4H was crystallised. An X-ray crystal structure of truncated DdP4H (residues 60-284) shows the conserved modified core double stranded β-helix (DSBH) fold, which is characteristic of 2OG oxygenases, along with elements involved in substrate binding that surround the active site. Structurally informed sequence analyses suggest DdP4H is homologous to human PHD2.

Biochemical studies of DdP4H and TgPhyA, using full-length DdSkp1 and TgSkp1 proteins as substrates, revealed high substrate K_m values, suggesting the hydroxylation efficiencies for these primitive enzymes are lower compared with the human PHDs (at least in terms of substrate binding). A high DdP4H K_m value for oxygen indicates DdP4H has potential to act as an oxygen sensor in *Dictyostelium discoideum,* as the PHDs do in humans. The unexpectedly low TgPhyA K_m values for oxygen may reflect the complicated life cycle of the parasite, which resides mostly in human tissues with a low oxygen level environment.

Inhibition studies were carried out on DdP4H and TgPhyA with known inhibitors of the human PHDs and other 2OG oxygenases, including compounds in current clinical trials. Most PHD2 inhibitors showed similar inhibition profile towards DdP4H and TgPhyA, consistent with the related active sites of these enzymes. Thus, PHD inhibitors have potential as (starting points) for inhibition of DdP4H and TgPhyA.

Cross organism prolyl hydroxylation assays were performed to investigate evolutionary relationships between the prolyl hydroxylases. Assays including the prolyl-hydroxylases and their corresponding substrates from *Dictyostelium discoideum*, *Toxoplasma gondii*, *Trichoplax adhaerens* (the simplest free-living animal), the Gram-negative gammaproteobacteria, *Pseudomonas putida*, and the single-celled eukaryotic choanoflagellate, *Monosiga brevicollis*, provided insight into the evolution of hypoxia sensing systems.

In summary, the results provide further insights into a chemical understanding of the evolutionary origins of the oxygen sensing systems and may enable a better understanding of the HIF system in humans. Studies on TgPhyA may also help reveal the new treatment for toxoplasmosis, infection caused by *Toxoplasma gondii*.

Acknowledgements

I would like to give my utmost gratitude to my supervisor, Professor Christopher Schofield, for full support and thoughtful guidance provided throughout my D.Phil study. I want to express my admiration for his creative ideas and his enthusiastic efforts on my study and my research. I would also like to give my appreciation to Dr Christopher West (University of Georgia) and Professor Ira J. Blader (University at Buffalo, the State University of New York) for our collaboration. I also thank Corpus Christi College and the China Scholarship Council for their financial support over the three years of my study.

I would like to thank all members of the CJS group for providing a friendly and enjoyable environment to work and to communicate. I would like to pay particular thanks to Dr Michael McDonough and Dr Adam Hardy for their constant support over the past 4 years. Much of the work in this thesis could not be achieved without their guidance and help. I would also like to give my appreciation to Dr Michael McDonough for introducing the X-ray crystallography process. I am grateful for Dr Adam Hardy, who trained me in a number of techniques when I first arrived in Oxford. Huge thanks belong to Dr Martine I. Abboud and Dr Christopher Lohans for teaching me how to perform NMR assays. I would like to give my appreciation to Dr Emily Flashman and Dr Hanna Tarhonskaya for training me on kinetic studies, to Dr Anthony Tumber for assisting me to perform inhibition assays using RapidFire-MS, and to Dr Rasheduzzaman Chowdhury for help with crystallography. I am also grateful for the help in various aspects from Dr Gareth Langley, Dr James P Holt-Martyn, Dr Jürgen Brem, Dr Sarah Wilkins, Dr Suzana Markolovic, Dr Ilaria Pettinati and Dr Ian Clifton. I would like to express my warmest thanks to all members of the Mass Spectrometry Facility at the Chemistry Research Laboratory, including Dr James Wickens, Ms Elisabete Pires, Dr Victor Mikhailov and Prof James McCullagh.

Most importantly, my particular thanks go to my parents for their constant supports and love through my whole life to share the happiness and to help me overcome all the difficulties.

Abbreviation

2,4-PDCA	2,4-Pyridinedicarboxylic acid
2D TOCSY	2 Dimensional Total Correlation Spectroscopy
2OG	2-Oxoglutaric acid
aa-tRNA	Aminoacyl-tRNA
AIDS	Acquired immunodeficiency syndrome
ATP	Adenosine-triphosphate
bHLH	Basic Helix-Loop-Helix
BSA	Bovine serum albumin
CTAD	C-terminal Trans-Activation Domain
cAMP	Cyclic adenosine monophosphate
CBP	CREB-binding protein
CL	Cell lysate
CNS	Central nervous system
CODD	C-Terminal oxygen-dependent degradation domain
CPH	Collagen prolyl hydroxylase
CTRL	Control
Cul1	Cullin 1
DdP4H	*Dictyostelium discoideum* P4H
DdSkp1	*Dictyostelium discoideum* S-phase kinase associated protein 1
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DSBH	Double-Stranded β-Helix fold
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EC_{50}	Concentration for 50% of maximal effect
EDTA	Ethylenediaminetetraacetic acid
EF-Tu	Elongation factor thermo unstable
EPO	Erythropoietin
ESI	Electrospray ionization
EST	Expression sequence tag
FAS	Ferrous ammonium sulphate
FBP	F-box binding protein
FIH	Factor Inhibiting HIF
Fuc	L-Fucose
Gal	D-Galactose
GDP	Guanosine-diphosphate
GlcNAc	N-aceltyl-D-glalactosamine
Gnt1	α-N-acetylglucosaminyltransferase
GTP	Guanosine-triphosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIF	Hypoxia Inducible Factor
His_6-ZB	Polyhistidine-ZBasic-tag
HRE	Hypoxia response elements
HyPro	Hydroxylated Proline
IC_{50}	Half maximal inhibitory concentration

IEC	Ion Exchange Chromatography
IPTG	Isopropyl β-D-Thiogalactoside
JmjC	Jumonji C
L-AA	L-Ascorbic Acid
LC-MS	Liquid chromatography-mass spectrometry
LIC	Ligation Independent Cloning
MALDI-MS	Matrix-Assisted Laser Desorption/Ionization-Mass Spectrometry
MbP4H	Monosiga brevicollis prolyl-hydroxylase
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
NEB	New England Biolabs
NMR	Nuclear Magnetic Resonance
NODD	N-terminal oxygen-dependent degradation domain
NOG	N-oxalylglycine
O.D.	Optimal density
OGFOD1	2-oxoglutarate and iron-dependent oxygenase domain containing 1
P4H	Prolyl 4-hydroxylase
PAS	Per-Arnt-Sim
PCR	Polymerase chain reaction
PHD	Prolyl hydroxylase domain
PPHD	*Pseudomonas* Prolyl-Hydroxylase Domain
PTM	Post-translational modifications
pVHL	von-Hippel Lindau protein
QToF	Quadrupole Time of Flight
Rbx1	RING-box protein 1
RF-MS	RapidFire-Mass Spectrometry
RNA	Ribonucleic acid
rpm	Rounds per minute
SCF	Skp1-Cul1-F-box
SDM	Site-Directed Mutagenesis
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size Exclusion Chromatography
Skp1	S-phase kinase associated protein 1
SPE	Solid phase extraction
TaODD	*Trichoplax adhaerens* oxygen-dependent degradation domains
TaPHD	*Trichoplax adhaerens* PHD
TCA	Tricarboxylic acid
TCEP	Tris(2-carboxyethyl)phosphine
TEV	Tobacco Etch Virus
TgPhyA	*Toxoplasma gondii* prolyl-4-hydroxylase A
TgPhyB	*Toxoplasma gondii* prolyl-4-hydroxylase B
TgSkp1	*Toxoplasma gondii* S-phase kinase associated protein 1
T_m Shift	Thermal Shift
Tris	Tris(hydroxymethyl)aminomethane
UV	Ultraviolet
VEGF	Vascular endothelial growth factor
WC	Whole cell

Table of Contents

Chapter 1 Introduction 1

1.1 Oxygen and the Hypoxic Response in Humans 1

1.1.1 Hypoxia Inducible Factor (HIF) Pathway 2

1.1.2 Domain Architecture of HIF 5

1.1.3 Proposed Mechanism for Prolyl Hydroxylation by PHDs 6

1.1.4 Available Crystal Structures of PHD2 8

1.1.5 PHDs Act as Oxygen Sensors 10

1.1.6 Therapeutic Inhibitors Targeting HIF Prolyl Hydroxylases 11

1.2 HIF System in Metazoans 11

1.3 Oxygen Sensing System in Other Organisms 13

1.4 Oxygen Sensing System in *Dictyostelium discoideum* 14

1.4.1 The Life Cycle of *Dictyostelium discoideum* 14

1.4.2 The Oxygen Sensing System of *Dictyostelium discoideum* 15

1.4.3 *Dictyostelium discoideum* S-Phase Kinase Associated Protein 1 (DdSkp1) 16

1.4.4 *Dictyostelium discoideum* Prolyl 4-Hydroxylase (DdP4H) 20

1.5 Toxoplasma gondii 20

1.5.1 The Life Cycle of *Toxoplasma gondii* 20

1.5.2 Toxoplasmosis 21

1.5.3 The Oxygen Sensing System of *Toxoplasma gondii* 22

1.6 Prolyl-Hydroxylases in Other Organisms 24

1.6.1 The Structure of *Pseudomonas* Prolyl-Hydroxylase Domain (PPHD) and its Substrate, Elongation Factor Thermo Unstable (EF-Tu) 24

1.6.2 The Function of PPHD and EF-Tu 25

1.6.3 *Monosiga brevicollis* Prolyl-4-Hydroxylase (MbP4H) 27

1.7 Evolution of Oxygen Sensing Systems 28

1.8 Skp1 in Humans 29

1.9 Skp1 in Different Organisms 32

1.10 Thesis Objectives 34

1.11 References 36

Chapter 2 Production of DdP4H and TgPhyA 47

2.1 Summary 47
2.2 Introduction 47
2.3 Initial Expression Trials for DdP4H in pET-15 and TgPhyA in pET-15b 48
2.3.1 Expression Trials for DdP4H in pET-15 48
2.3.2 Expression Trials for TgPhyA in pET-15b 49
2.4 Expression Optimization Trials for TgPhyA Using ArcticExpress (DE3) Competent Cells 50
2.5 Ligation-independent cloning (LIC) of DdP4H and TgPhyA 52
2.6 Expression Trials of DdP4H in pNIC-ZB and TgPhyA in pNIC-ZB 54
2.6.1 Expression Trials of DdP4H in pNIC-ZB 54
2.6.2 Expression Trials of TgPhyA in pNIC-ZB 56
2.7 Expression and Purification of DdP4H and TgPhyA 57
2.7.1 Expression and Purification of DdP4H 57
2.7.2 Expression and Purification of TgPhyA 59
2.8 Differential Scanning Fluorimetry (T_m Shift) Assays 60
2.9 Truncated Constructs, DdP4H (60-284) and TgPhyA (53-264) 64
2.10 LIC on DdP4H, DdP4H (60-284), and TgPhyA, TgPhyA (53-264) 66
2.11 Expression Optimization Trials of DdP4H, DdP4H (60-284), and TgPhyA, TgPhyA (53-264) 67
2.12 Producing and purifying DdP4H, DdP4H (60-284) 68
2.13 Producing and Purifying TgPhyA (53-264) 70
2.14 References 72

Chapter 3 Structural Studies of DdP4H (60-284) 74

3.1 Summary 74
3.2 Introduction 74
3.3 Crystallisation Screening and Optimisation of DdP4H (60-284) 74
3.4 X-ray Crystallography of DdP4H (60-284) 77
3.5 Overall Structure of DdP4H (60-284) 79
3.6 Active Site of DdP4H (60-284) 80
3.7 Comparison of Structures of DdP4H (60-284), PPHD and PHD2 81
3.8 Comparison of the Active Sites of DdP4H (60-284), PHD2 and PPHD 86
3.9 Crystal Structure of DdP4H (60-284) with a Model of the DdSkp1 Substrate 87
3.10 Disulfide-bond Crosslinking to Crystallise DdP4H (60-284) and DdSkp1 Complex 89

3.10.1 Using the DdP4H (60-284) Structure and DdSkp1 Model for Crosslinking Design 89
3.10.2 Site-Directed Mutagenesis of DdP4H (60-284) 90
3.10.3 Expression and Purification of DdP4H (60-284) R192C 91
3.10.4 Synthesis of DdSkp1 Peptides 93
3.10.5 Time Course Assays of Disulfide-Bond Crosslinking 93
3.11 Discussions and Future Plans 94
3.11.1 Crystal Structure of DdP4H (60-284) in Complex with DdSkp1 Substrate 94
3.11.2 Attempted Crystallisation of TgPhyA (56-264) 95
3.12 References 95

Chapter 4 DdP4H and TgPhyA Activity Assay Development 97

4.1 Summary 97
4.2 Introduction 97
4.3 Synthesis of DdSkp1, TgSkp1, HsSkp1, and HsSkp1 E147P Peptides 97
4.4 MALDI-MS Based Hydroxylation Activity Assay of DdP4H and TgPhyA with DdSkp1 and TgSkp1 98
4.5 Time Course Hydroxylation Assays 99
4.6 Kinetic Study of DdP4H with DdSkp1 Peptide 100
4.7 LIC on DdSkp1 and TgSkp1 102
4.8 Expression Trials of DdSkp1 and TgSkp1 103
4.9 Purification of DdSkp1 and TgSkp1 105
4.10 Purification of HsSkp1 and HsSkp1 E147P 107
4.10.1 Site-Directed Mutagenesis to Make HsSkp1 E147P Construct from HsSkp1 Plasmid 107
4.10.2 Expression Trials of HsSkp1 108
4.10.3 Expression and Purification of HsSkp1 and HsSkp1 E147P 109
4.11 Mass Spectrometry Based Hydroxylation Assays 110
4.12 Hydroxylation Assays of DdP4H and TgPhyA with Full-Length DdSkp1 and TgSkp1 Proteins 112
4.12.1 Truncated DdP4H, DdP4H (60-284) Hydroxylated Full-length DdSkp1 Protein 114
4.13 TgSkp1 Exhibits Non-Specific Oxidation 116
4.14 The Non-Specific Oxidation of TgSkp1 Does Not Affect Kinetic Studies 119
4.15 Kinetic Studies of DdP4H Using Full-Length DdSkp1 Protein Substrate 121

4.16 Comparison of DdP4H Kinetics Using DdSkp1 Peptide and Full-length Protein Substrates 122
4.17 DdP4H Kinetic Parameters for 2OG 122
4.18 RapidFire Mass Spectrometry (RF-MS) Hydroxylation Assay Development 124
4.19 Kinetic Study of DdP4H with 2OG Using RF-MS 125
4.20 Kinetic Studies of DdP4H with Oxygen 126
4.21 Kinetic Studies of TgPhyA with Full-Length TgSkp1 Protein Substrate 128
4.22 Kinetic Studies of TgPhyA with 2OG 129
4.23 Kinetic Studies of TgPhyA with Oxygen 130
4.24 Discussion 132
4.25 MS/MS Analysis of Proline/Hydroxyproline in Full-Length DdSkp1 Protein 133
4.25.1 Trypsin Digestion of Samples 133
4.25.2 MS/MS Analysis Using a Nanoscale Liquid Chromatography Coupled to Tandem Mass Spectrometry (Nano LC-MS/MS) 134
4.25.3 MS/MS Analysis of Proline/Hydroxyproline in Full-Length DdSkp1 Protein 135
4.25.4 MS/MS Analysis of Proline/Hydroxyproline in Full-Length TgSkp1 Protein 138
4.26 Regiochemistry of the Hydroxyproline in DdSkp1 and TgSkp1 142
4.27 2OG Turnover Assays 145
4.27.1 2OG Turnover Assays Conditions 145
4.27.2 2OG Turnover Assays Results 147
4.27.3 2OG Turnover of DdP4H with Various Substrates 148
4.27.4 2OG Turnover of TgPhyA with Various Substrates 149
4.27.5 2OG Turnover of PHD2 with Various Skp1s 150
4.27.6 A Further Investigation of the Uncoupled 2OG Turnover by TgPhyA 151
4.28 References 153

Chapter 5 Inhibition Studies on DdP4H and TgPhyA 155

5.1 Summary 155
5.2 Introduction 155
5.3 Known Prolyl Hydroxylase Inhibitors for DdP4H and TgPhyA 155
5.3.1 Potential 2OG Oxygenase Inhibitors and 2OG Analogues 156
5.3.2 PHD2 Inhibitors 157
5.3.3 Other 2OG Oxygenase Inhibitors 159

5.3.4 Endogenous Inhibitors of the 2OG Oxygenases and the Analogues 160
5.3.5 Others 161
5.3.6 Compounds from Dr. James P Holt -Martyn 162
5.4 Inhibition Screening of DdP4H Using Full-length DdSkp1 as the Substrate Analysis by RF-MS 163
5.5 Inhibitor Screening of TgPhyA Using TgSkp1 Peptide as the Substrate Analysed by RF-MS 167
5.6 Inhibitor Screening of DdP4H Using DdSkp1 Peptide as the Substrate Analysed by RF-MS 170
5.7 Discussion of Inhibitor Screening Results 172
5.7.1 Comparison of IC_{50}s Using Full-Length DdSkp1 Protein and DdSkp1 Peptide as the Substrate for DdP4H 172
5.7.2 Comparison of IC_{50}s of DdP4H and TgPhyA 173
5.8 Inhibitor Screening of DdP4H and TgPhyA Using Novel PHD2 Inhibitors 174
5.9 References 176

Chapter 6 Cross Organism Hydroxylation Studies 180

6.1 References 180
6.2 Summary 180
6.3 Cross Organism Hydroxylation Between DdP4H, TgPhyA, Human PHD2 181
6.4 Screening of DdP4H and TgPhyA Against HIF Substrates 183
6.5 Cross Organism Hydroxylation Assays 184
6.6 *Toxoplasma gondii* PhyB, TgPhyB 187
6.7 TgPhyB Assays 189
6.8 Summary of Cross Organism Hydroxylation Results 191
6.9 References 192

Chapter 7 Summary 194

References 199

Chapter 8 Methods and Materials 202

8.1 Molecular Biology Techniques 202
8.1.1 Preparation of Competent Cells 202
8.1.2 Plasmids 203

8.1.3 Site-Directed Mutagenesis (SDM) 203
8.1.4 Bacterial Transformations 205
8.1.5 Ligation Independent Cloning (LIC) 205
8.1.6 Agarose Gel Electrophoresis 207
8.2 Protein Production 208
8.2.1 2TY Growth Medium 208
8.2.2 Glycerol Stock Preparation 208
8.2.3 Starter Cultures 208
8.2.4 Expression Trials 208
8.2.5 SDS-PAGE Electrophoresis 209
8.2.6 Large-Scale Growth 210
8.3 Protein Purification 211
8.3.1 Ni(II) Affinity Chromatography 211
8.3.2 Ion Exchange Chromatography (IEC) 212
8.3.3 Ni(II) Affinity Chromatography to Re-purify the Protein Samples 213
8.3.4 Size Exclusion Chromatography (SEC) 214
8.3.5 Purification Strategies 215
8.4 Mass Spectrometry 216
8.4.1 Hydroxylation Assay and Matrix-Assisted Laser Desorption Ionization Mass Spectrometry (MALDI-MS) 216
8.4.2 Kinetic Assays of Prolyl Hydroxylases with MALDI-MS 217
8.4.3 Liquid Chromatography-Mass Spectrometry (LC-MS) 218
8.4.4 RapidFire Mass Spectrometry (RF-MS) 218
8.4.5 Nanoscale Liquid Chromatography Coupled to Tandem Mass Spectrometry (Nano LC-MS/MS) 218
8.5 Crystallography 219
8.5.1 Crystallography Trials 219
8.6 Peptide Synthesis 222
8.7 Construct Information 223
8.8 References 228

1. Introduction

1.1 Oxygen and the Hypoxic Response in Humans

Around 1,500 million years ago, it is proposed that atmospheric oxygen levels began to increase due to the growing number of photosynthetic organisms. The energy in the dioxygen molecule was likely toxic to many single-cell prokaryotic life forms [1, 2], but new organisms are proposed to have evolved that they were able to harness the chemical properties of dioxygen and the use of dioxygen to generate cellular ATP became much more efficient than anaerobic metabolism. This new energy source is proposed to have driven the rapid boost in diversification of species during the Cambrian era 542–488 million years ago when aerobic organisms were born. The ability to harness dioxygen appears to have been a major contribution to the development and diversification of aerobic species, ultimately including humans [3]. Relatively simple organisms, such as flatworms and nematodes harness dioxygen by direct diffusion [4]. More complex organisms have developed different systems in order to ensure the transportation of dioxygen into every single cell. These systems include the tracheal respiratory mechanism in insects, and the circulatory systems in vertebrates coupled with either gills or lungs [5].

In the human body, dioxygen is acquired from the environment through the respiratory system. After entering the body, dioxygen then binds to haemoglobin in red blood cells and is then carried throughout the body before finally being transferred to every cell where it takes part in respiration and other chemical reactions [6, 7]. The sensing of dioxygen relies on the sensory organ, carotid bodies [8], which function to sense acute changes of arterial dioxygen tension to elicit hyperventilation [9]. The molecular mechanisms in carotid body dioxygen sensing are not fully clear, but there is emerging evidence that HIF-α isoforms may play a key role [10], which will be further discussed below.

Normally in healthy people, hypoxia can occur when dioxygen (hereafter simply referred to as 'oxygen') supply is lower than normal, for example, when people reach high altitude, or on prolonged breath holding during underwater diving, or when blood vessels supplying tissues become blocked. People normally recover when returning back to a normal-oxygen-level environment. Chronic hypoxia occurs when being exposed in a low oxygen level environment for longer than weeks. Chronic hypoxia may be induced by low levels of haemoglobin (anaemia) or when there is insufficient blood flow to tissues (ischaemia) or in some solid tumours. The consequences of chronic hypoxia include the abnormal cellular respiration, energy insufficiency, and abnormal metabolism [11]. Accordingly, animals have evolved a complex system to allow them to adapt to chronic periods of low oxygen; this 'hypoxic response' is substantially controlled by the Hypoxia Inducible Factor (HIF) system.

1.1.1 Hypoxia Inducible Factor (HIF) Pathway

In most, if not all, metazoans, the hypoxic response is regulated by the Hypoxia Inducible Factor (HIF) system. The α,β-heterodimeric HIF proteins function as transcription factors. HIF-α isoforms are located in the cytoplasm. The HIF-α, but not HIF-β, isoforms are degraded in an oxygen dependent manner. In normoxia, HIF-α is constantly expressed and degraded. When oxygen levels are sufficiently high, proline residues (Pro402 and Pro564) in the N- and C-terminal oxygen dependent domains (NODD and CODD) of HIF-α are efficiently hydroxylated by prolyl hydroxylase domain (PHD) containing proteins [12]. In humans, there are three PHDs (PHD1-3) that catalyse HIF-α prolyl hydroxylation. The PHDs are Fe(II) and 2-oxoglutaric acid (2OG) dependent oxygenases [13]. The detailed mechanisms of prolyl hydroxylation by PHDs are discussed in Chapter 1.1.3. HIF-α prolyl hydroxylation promotes its degradation in the cytoplasm [14]. Hydroxylated HIF-α has increased

binding affinity for the von Hippel-Lindau tumour suppressor protein (pVHL) and is rapidly captured, ubiquitinated, and destroyed by a pVHL E3 ligase complex [15].

HIF-β is normally located in the nucleus of cells. Under hypoxic conditions, HIF-α translocates to nucleus and dimerises with HIF-β. The HIF-α,β heterodimer binds to the penta-nucleotide sequence 3'-(A/G)CGTG-5' hypoxia response elements (HRE) within DNA sequences, in order to promote transcription of subsequent target genes. Examples of the hundreds of HRE containing genes include erythropoietin (EPO) [16] which promotes hematopoiesis, vascular endothelial growth factor (VEGF) which stimulates angiogenesis, and genes involved in glucose transport and metabolism [17], all of which are involved in the chronic hypoxia response [18].

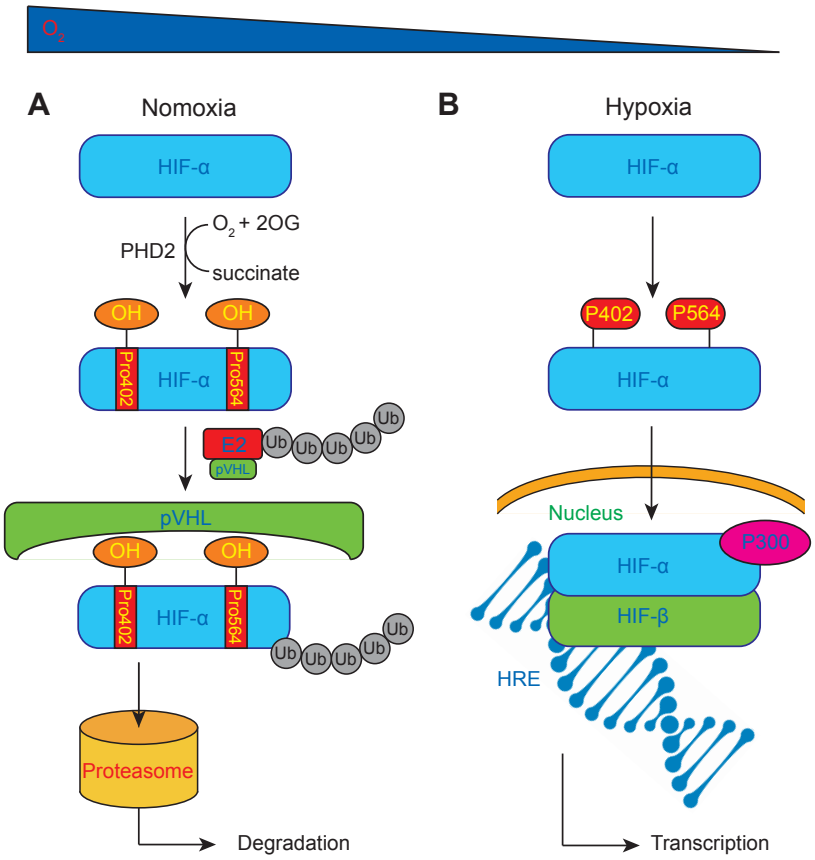


Fig. 1.1 Overview of the HIF system. (A) In normoxic conditions, HIF-α is rapidly degraded. Proline residues (Pro402 and Pro564) in the N- and C-terminal oxygen dependent domains (NODD and

CODD) regions of HIF-α are hydroxylated by PHDs. The hydroxylated HIF-α is then recognized and bound by the von-Hippel Lindau protein (pVHL), which is a part of an E3 ubiquitin ligase complex, and HIF-α is ubiquitinated by a ubiquitin conjugating enzyme (E2). Ubiquitinated HIF-α is then subject to proteasomal degradation. (B) Under hypoxia, when oxygen levels are low, HIF-α is not as sufficiently hydroxylated by the PHDs. The unhydroxylated HIF-α translocates to the nucleus, and dimerises with HIF-β leading to a transcriptionally active complex. The complex then binds to the hypoxia response elements (HRE) of DNA and triggers downstream gene transcription.

The hypoxic response in humans is also regulated by Factor Inhibiting HIF (FIH), another 2OG oxygenase, which recognises and hydroxylates HIF Asp803 under sufficiently normoxic conditions [19]. The hydroxylation of Asp803 significantly decreases the HIF-α binding affinity to the transcriptional co-activator complex CBP/p300, resulting in reduced HIF mediated transcription [20]. Biochemical studies support the proposed hydroxylation of proline residues in CODD or NODD by PHD2, and hydroxylation of Asn803 in CTAD (CAD) domain catalysed by FIH [21] (Table 1.1) [22]. It is proposed that the K_m^{app} values for oxygen of the HIF hydroxylases reflect this role as hydroxia sensors.

A

Enzyme	Substrate	K_m^{app}(2OG) / μM	K_m^{app}(Peptide) / μM	$K_m^{app}(O_2)$ / μM	K_{cat}^{app}/ s^{-1}
FIH	HIF-1α $CAD_{35\text{-}mer}$	110 ± 20	180 ± 30	110 ± 30	0.56 ± 0.04
FIH	HIF-2α $CAD_{35\text{-}mer}$	19 ± 6	315 ± 40	110 ± 10	0.049 ± 0.005
PHD2	HIF-1α $CODD_{19\text{-}mer}$	13 ± 2	10 ± 6	460 ± 30	0.060 ± 0.006
PHD2	HIF-1α $NODD_{19\text{-}mer}$	30 ± 9	11 ± 2	> 450	0.028 ± 0.001
PHD2	HIF-2α $CODD_{19\text{-}mer}$	9 ± 2	34 ± 10	> 450	0.069 ± 0.006
PHD2	HIF-2α $NODD_{19\text{-}mer}$	17 ± 5	50 ± 8	410 ± 80	0.013 ± 0.001

B

Peptide	Sequence
HIF-1α CAD (789-822)	DESGLPQLTSYDCEVNAPIQGSRNLLQGEELLRAL
HIF-2α CAD (832-866)	ESYLLPELTRYDCEVNVPVLGSSTLLQGGDLLRAL
HIF-1α CODD (556-574)	DLDLEMLAPYIPMDDDFQL
HIF-1α NODD (395-413)	DALTLLAPAAGDTIISLDF
HIF-2α CODD (523-541)	ELDLETLAPYIPMDGEDFQ
HIF-2α NODD (398-416)	EELAQLAPTPGDAIISLDF

Table 1.1 Reported steady-state kinetic results for FIH and PHD2 catalysis. (A) Data for PHD2 and FIH were derived from (Tarhonskaya 2015) [22]. (B) Sequences of HIF-1,2α CODD, NODD and CAD peptides.

Therefore, in total (at least) three sites of HIF are regulated by oxygen levels, but there is evidence that the sensitivity to oxygen level changes is in the order of Pro402 > Pro564 > Asn803 [22], which means the hydroxylation occur quicker in such order when oxygen level increases from hypoxia. This finding suggests FIH pathway may be for fine tuning oxygen sensing, while the PHDs are the primary oxygen sensors [23].

1.1.2 Domain Architecture of HIF

There are three isoforms of HIF-α in humans, HIF-1α (Gene ID: Q16665.1), HIF-2α (Gene ID: Q99814.3), and HIF-3α (Gene ID: AAD22668.1). All HIFs are basic helix–loop–helix (bHLH) and PAS (Per-Arnt-Sim) domain containing proteins. The bHLH and PAS domains take part in DNA binding and dimerisation with HIF-β. The PHD targeted prolines are located in the N- and C-terminal oxygen dependent degredation domains (NODD and CODD). There is only one ODD and CTAD in human HIF-3α. HIF-1α Asp803 is located in the C-terminal Trans-Activation Domain (CTAD) region of HIF. The ODDs and CTAD regions are involved in the degradation and transactivation of HIFs.

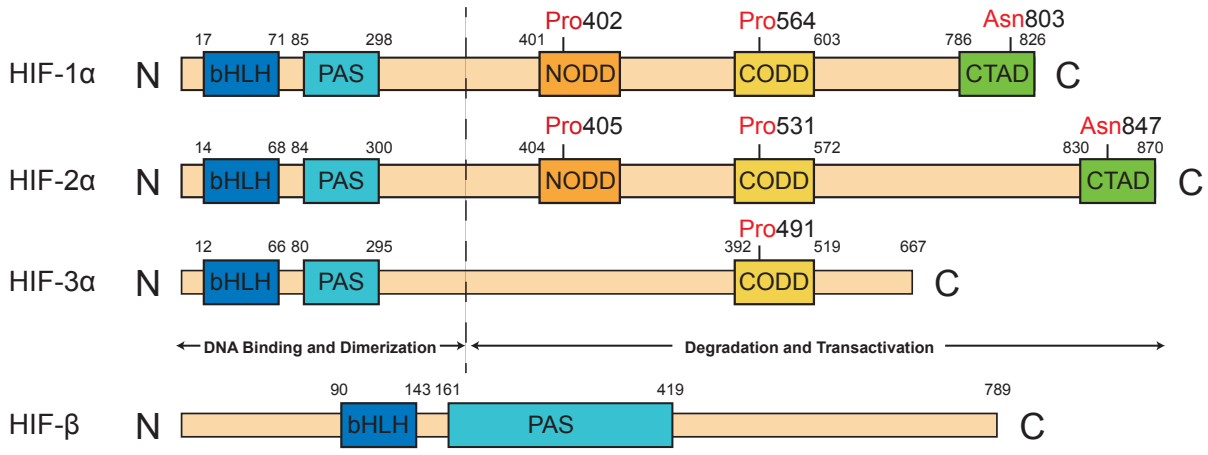


Fig. 1.2 Domain architectures of human HIF-1α, -2α, -3α and HIF-β. There are three isoforms of HIF-α and one HIF-β in humans. HIF-α contains bHLH and PAS domains which take part in DNA binding and dimerisation. HIF-1α and HIF-2α contain NODD, CODD and CTAD domains which take

part in oxygen sensing and functions in HIF degradation and transactivation. HIF-3α contains one ODD domain and no CTAD domain. HIF-β also contains bHLH and PAS domains, which function to bind DNA and dimerise with HIF-α.

1.1.3 Proposed Mechanism for Prolyl Hydroxylation by PHDs

The hydroxylation reaction catalysed by PHDs involves iron cofactor mediated oxidative decarboxylation of the co-substrate (2-oxoglutaric acid, 2OG) and the transfer of an oxygen atom from dioxygen to the 4-position of the proline residues of the HIF substrate [21, 24, 25]. In the active site of the PHDs, 2OG coordinates to the central Fe(II) in a bidentate manner via its oxalate group. Fe(II) is also coordinated by the side chains of His374, His313, Asp315 and one water molecule to form an octahedral coordination complex. In the standard 2OG oxygenase mechanism, the substrate binds after the binding of 2OG. Molecular oxygen then binds to the active site iron (in the core of PHD2 [24]) displacing a tightly bound water molecule. Oxidative decarboxylation of 2OG presumably via a peroxo-intermediate then occurs, releasing CO_2 and forming succinate, along with a reactive Fe(IV)=O intermediate. The peptide substrate is hydroxylated by the Fe(IV)=O intermediate with the release of succinate and CO_2. Another 2OG molecule can then bind, initiating another catalytic cycle.

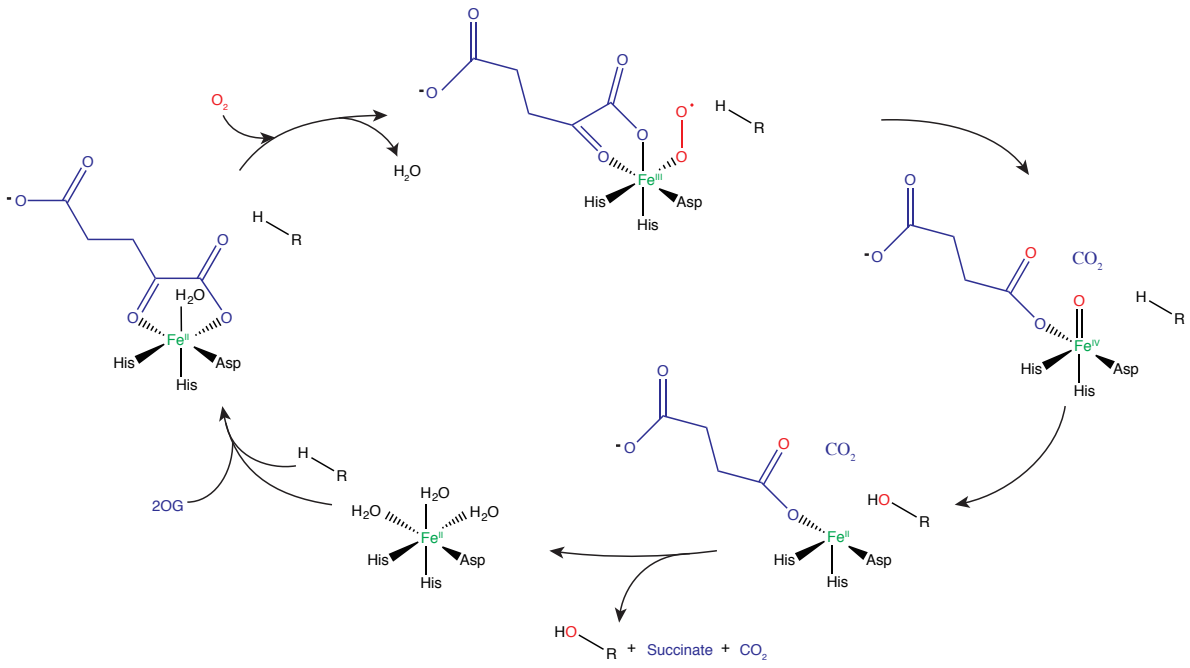


Fig. 1.3 Outline scheme for a general 2OG oxygenase mechanism [24]. In the active site of 2OG oxygenases, an active site iron is coordinated by a conserved triad of residues [HX(D/E)...H motif] and three water molecules to form an octahedral coordination complex. Binding of 2OG and oxygen, thus, displace the metal bound water molecules. Oxidative decarboxylation of 2OG is then catalysed by the active site iron-oxygen complex and is coupled with hydroxylation of the substrate. The active site iron is restored to its original state upon completion of the catalytic cycle.

1.1.4 Available Crystal Structures of PHD2

There are at least 3 PHDs (PHD1-3) in most mammals, including humans. However, the different PHDs are expressed differently in different tissues in a temporal and tissue specific manner and are localised to different subcellular regions [26]. Fluorescence microscopy shows that PHD1 is only localised in the nucleus, PHD2 and FIH are mainly distributed in the cytoplasm, and PHD3 is homogeneously present in both cytoplasm and nucleus[27]. Although the specific roles of the different PHDs are not well characterised, PHD2 is reported to be the most important of the PHDs in response to hypoxia [28]. The PHD2 knock-out mouse results in a high rate of developmental defects

and premature lethality, although apparent angiogenic defects are not observed for PHD1 or PHD3 knock-out mice [29].

Therefore, PHD2 has been the more focus of studies than PHD1 and PHD3. The crystal structure of PHD2 (Figure 1.4) provides the first description of the structure of human prolyl hydroxylases [30, 31]. Like other 2OG dependent oxygenases, the central structure of PHD2 possesses a core double-stranded β-helix fold (DSBH) [32]. The DSBH fold consists of 8 anti parallel β-strands, with 4 β-strands forming the major β-sheet and another 4 β-strands forming the minor β-sheet. In PHD2, the core DSBH fold supports three iron binding ligands: His374, His313 and Asp315, which form the conserved HXD. . .H metal binding motif [31]. The active site iron sits at the more open end of the DSBH barrel. This open end provides a route for cofactors, co-substrate and prime substrate binding. Several α-helices buttress the major sheet. Additional parallel and anti-parallel β-strands to those of the DSBH and several loops are involved in substrate binding [30].

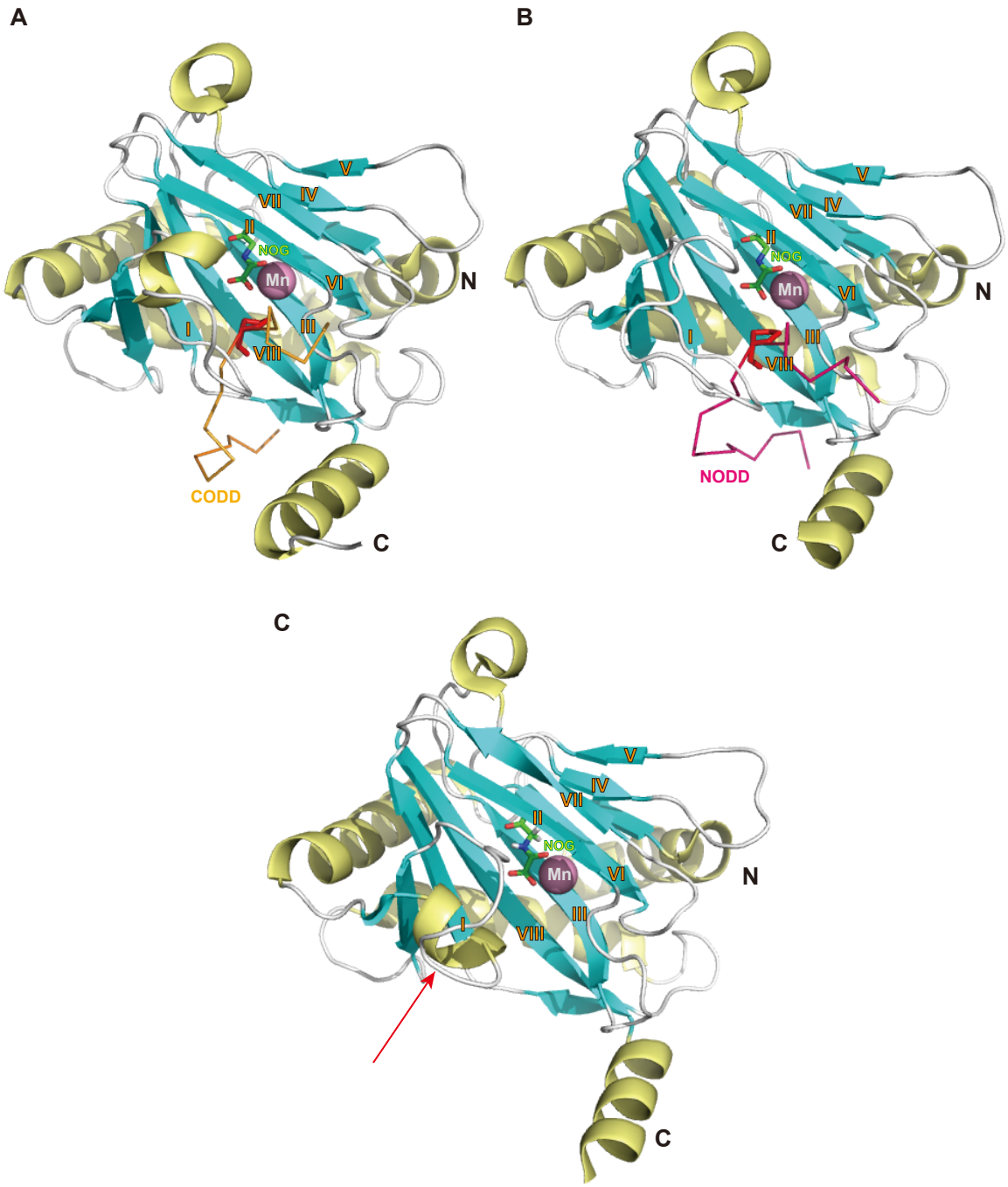


Fig. 1.4 Views from crystal structures of PHD2 in complex with substrates. (A) Cartoon representation of a structure of PHD2:Mn^{2+}:NOG:CODD complex (PDB ID: 3HQR). (B) Cartoon representation of a structure of PHD2:Mn^{2+}:NOG:NODD complex (PDB ID: 5L9V). (C) Cartoon representation of a structure of PHD2:Mn^{2+}:NOG (PDB ID: 5L9R). β-sheets are coloured in cyan. Roman numerals indicate β-strands comprising the core DSBH. α-helices are coloured in pale yellow, loops in white. The 2OG analogue, N-oxalylglycine (NOG), is shown as green sticks residing in the active site surrounded by the core DSBH. The metal ion, Mn^{2+}, is shown as a pink sphere. CODD peptide is represented as ribbons coloured in orange in (A). NODD peptide is represented as

ribbons coloured in red pink in (B). The target proline residue is shown as sticks coloured in red. The red arrow in (C) indicates the loop involved in substrate binding. The β2-β3 loop closes when CODD or NODD peptide is bound in the active site.

1.1.5 PHDs Act as Oxygen Sensors

Biochemical studies on the PHDs have supported the function of these proteins as oxygen sensors. *In vitro* kinetic studies indicate a high apparent K_m of oxygen for PHD2 [$K_m^{app}(O_2) > 450$ μM] [22]. The K_m of oxygen is higher than the atmospheric oxygen concentrations. Thus, even small changes of O_2 concentration can in principle influence the activity of PHDs and thus down regulate HIF-α. This, in turn, can influence the transcription of hypoxia responsive HIF target genes. Consistent with the sensitivity to changes in O_2 concentration and slow reaction with O_2 [33], PHD2, being the most important of the PHDs, is proposed to biochemically function as an oxygen sensor in many animals [34].

Biological studies also imply the importance of PHD2 in oxygen sensing. *In vivo* studies show that inhibition of PHD2 results in elevated EPO production and erythropoiesis [17], which is a feedback response to help the organism cope with low oxygen availability of hypoxia conditions. The gene knock out of PHD2 leads to pulmonary arterial hypertension in mice and human [35] and lethality in mice [36]. PHD2 is also responsible for oxygen-induced retinopathy, by regulating retinal microvessels genesis [37]. It is also reported that PHD2 plays an important role in the immune response [38] and tumour development. Dysfunction of PHD2 correlates with tumour angiogenesis [39], which may explain the feature of tumour cells of adapting to low oxygen or hypoxic conditions. All these findings are consistent with the proposed functions of PHD2 as an oxygen sensor and the HIF pathway as a key regulator in oxygen homeostasis [11].

1.1.6 Therapeutic Inhibitors Targeting HIF Prolyl Hydroxylases

Inhibitors targeting prolyl hydroxylases have been shown to be effective in late stage clinical trials for treatment of anaemia in patients with chronic kidney disease by enhancing endogenous erythropoietin production [40, 41]. Patients suffering more generally from ischaemia related diseases might also benefit from HIF prolyl hydroxylase inhibitor treatment [42, 43]. Moreover, the activation of HIFs by the inhibition of HIF prolyl hydroxylases has been shown to have effects on the immune system and inflammatory cells [43, 44]. However, a complete understanding of the effects of HIF prolyl hydroxylase inhibitors as well as the long-term safety issues are lacking; further research is required.

A series of, at least partially selective, HIF prolyl hydroxylases inhibitors have been developed and tested, including FG-2216 [45-47], FG-4592 [48, 49], IOX2 [50], BIOX2 [51], IOX4 [52], Vadadustat [53, 54], GSK1278863 [55], and FG-0041 [56]. All the compounds bind to the active site iron and compete with 2OG. The chemical structures of inhibitors and their inhibitory properties are further discussed in Chapter 5.

1.2 HIF System in Metazoans

Being the central system that coordinates the chronic hypoxia response, the HIF system apparently exists in most metazoans and is thus highly conserved [57]. A functional HIF system can even be found in the simplest animal, *Trichoplax adhaerens* [58]. *Trichoplax adhaerens*, belonging to Placozoans (literally, 'flat animals'), are small, disc-shaped creatures. *Trichoplax adhaerens* is recognised as the simplest free-living animal and has a primitive metazoan form [59]. Bioinformatics studies suggested the existence of a conserved PHD in *Trichoplax adhaerens* (TaPHD) and a HIF substrate with an oxygen-dependent degradation domain (TaODD). Indeed, in the presence of 2OG, oxygen and Fe(II),

TaPHD hydroxylates TaODD in the same manner as PHD2 hydroxylates HIF-α. TaPHD is sensitive to oxygen levels, and therefore, is likely involved in a HIF regulatory system that coordinates the hypoxic response in *Trichoplax adhaerens*. Cross-species hydroxylation assays also show that TaPHD can hydroxylate human CODD and NODD peptide fragments, and that human PHD2 hydroxylates TaODD (these experimental results can be found in Dr. Kerstin Lippl's thesis [60], Department of Chemistry, Chemistry Research Laboratory, University of Oxford, OX1 3TA), suggesting the conservation of the oxygen sensing systems between *Trichoplax adhaerens* and humans and by implication in most animals [58].

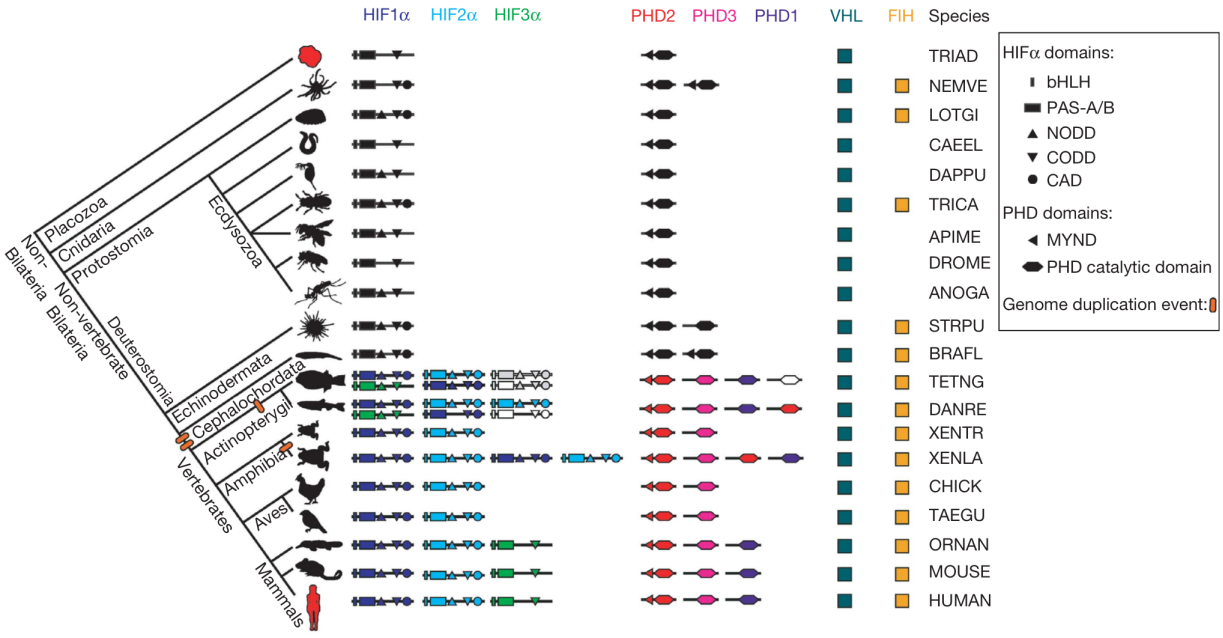


Fig. 1.5 Phylogenetic domain analysis of the hypoxia-inducible transcription factor system in Metazoans [58]. Being the simplest animal (coloured in red on the top of the figure), *Trichoplax adhaerens* has one copy of TaODD, which is the homologue of human HIF-1α, and one copy of TaPHD, which is the homologue of human PHD2. Animals at higher evolution levels tends to have more than one copy of PHD and HIF-α. The highest level of biological species, mammals, have a full set of HIF system, including 3 PHD homologues, 3 HIF-α homologues, as well as VHL and FIH. The figure is taken from (Loenarz 2011); permission acquired.

1.3 Oxygen Sensing System in Other Organisms

The HIF, PHD and VHL system is apparently not present in non-metazoans [57, 58]. Therefore, the evolution involving the development of oxygen sensing hydroxylation is of interest. Several PHD-related oxygenases are present in other life forms, including protists and prokaryotes, which include but are not limited to *Monosiga brevicollis*, *Dictyostelium discoideum*, *Toxoplasma gondii*, *Pseudomonas putida* and *Pseudomonas aeruginosa*. Although HIF homologues are not likely present in the genomes of these organisms, one or several PHD homologues exist, each of which will be discussed in the following sections. The substrates for these PHD homologues varies. To date, S-Phase Kinase Associated Protein 1, Skp1, or elongation factor thermo unstable, EF-Tu, have been identified as PHD homologues substrates in *Dictyostelium discoideum* (Skp1) [61], *Toxoplasma gondii* (Skp1) [62] and *Pseudomonas aeruginosa* (EF-Tu) [63], respectively.

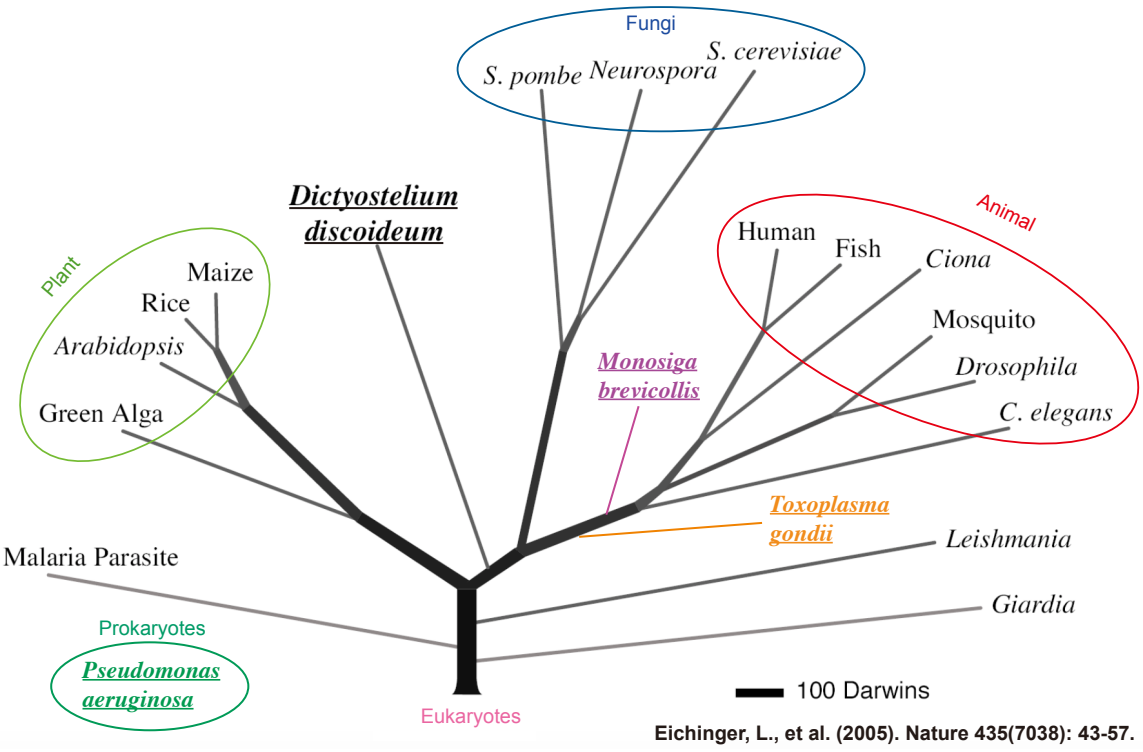


Fig. 1.6 Proteome-based Phylogenetic Tree [64]. Apart from the three main Kindoms, animals, fungi, and plants, selected other organisms are listed. The evolutionary relationships of these organisms

are obscure. The species of interest in this thesis are highlighted with underlining. These species include protists, *Monosiga brevicollis*, *Dictyostelium discoideum*, *Toxoplasma gondii* and prokaryotes, *Pseudomonas aeruginosa*. This figure is modified from a figure in (Eichinger 2005).

1.4 Oxygen Sensing System in *Dictyostelium discoideum*

1.4.1 The Life Cycle of *Dictyostelium discoideum*

Over the last 50 years, *Dictyostelium discoideum* has been used as a model organism for a wide range of biological studies, including developmental signaling [65], cell motility, chemotaxis, phagocytosis, endocytic vesicle traffic, cell adhesion, pattern formation, caspase-independent cell death, autophagy, and social evolution studies [66].

Dictyostelium discoideum is a social amoeba, but also has a single cell stage during growth and initial development. When facing starvation situations, unicellular *Dictyostelium discoideum* aggregates into multicellular moulds in a process involving secretion of cyclic adenosine monophosphate (cAMP) as a signal to recruit surrounding cells [66, 67]. Thousands of cells aggregate and develop to form a migrating slug. The slugs can migrate and are sensitive to light, temperature and oxygen. The movement is regulated by the anterior tip, which functions as an organizer. Ultimately, the *Dictyostelium discoideum* slugs form a fruiting body, which has sub structures, including a basal disc, a stalk, and a spore cyst. This step is described as culmination and it is reported to be exquisitely determined by oxygen level (> 10%) [68]. The fruiting body enters into a germination state, releasing matured spores. In nutrient-rich habitats, the spores germinate and resume proliferation as individual amoebae [66, 67].

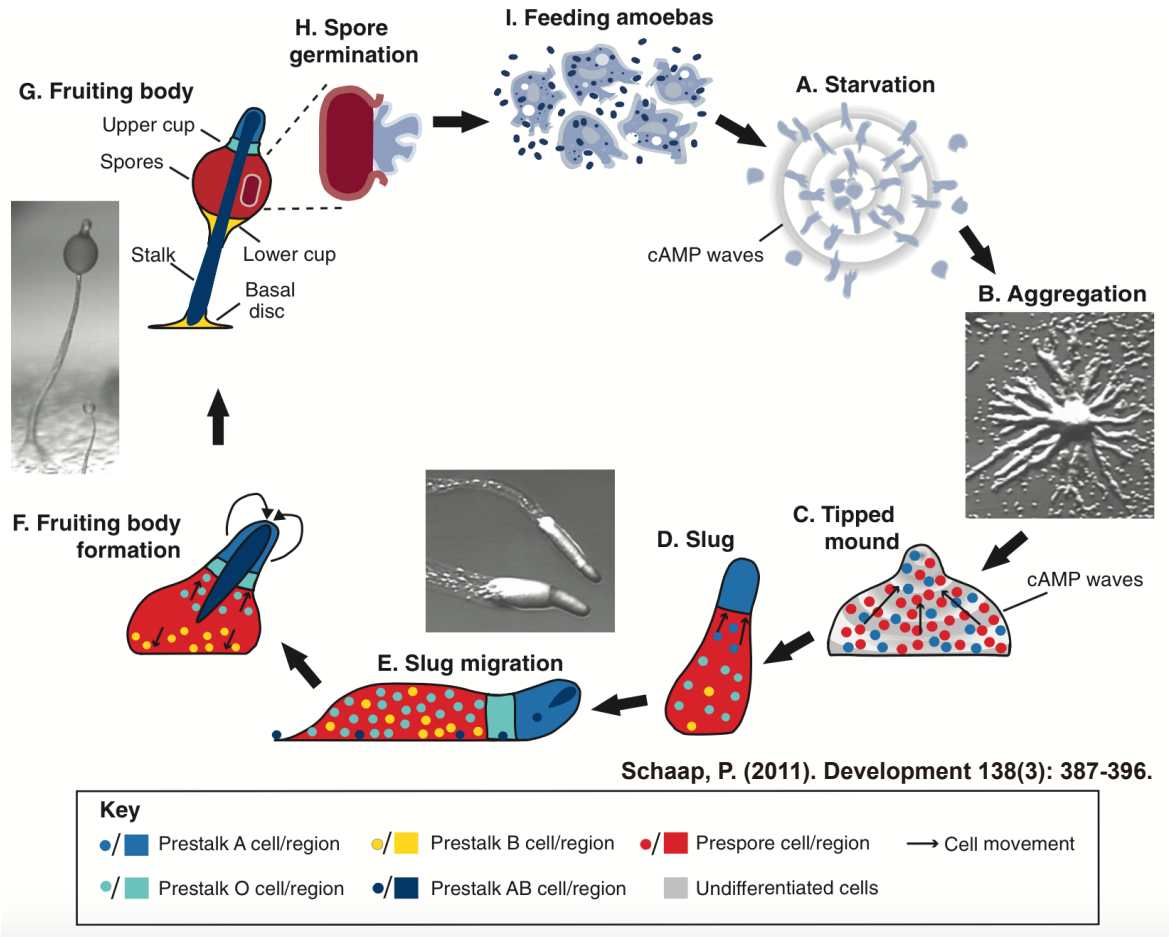


Fig. 1.7 Life cycle of *Dictyostelium discoideum* [67]. (A) Starved *Dictyostelium discoideum* secrete cAMP, indicated by grey rings to gather other cells. (B) Cells aggregate responding to the chemoattractant cAMP. (C) Aggregates form a tipped mound with cells differentiated into Prestalk A cells and Prespore cells. (D-E) A mound then forms a moving slug with further cellular differentiation. (F) The slug undergos fruiting body formation. (G) Different sub structures develop, including a basal disc, a stalk and a spore cyst. (H-I) Spores are released, which germinate and reform individual amoebae [65, 67]. The figure is taken from (Schaap 2011); permission acquired.

1.4.2 The Oxygen Sensing System of *Dictyostelium discoideum*

Dictyostelium discoideum exhibits oxygen sensing features during its development [69]. Although systematic genetic screens have not been done for genes that disrupt the effects of hypoxia, the signalling pathway that mediate oxygen sensing has long been identified [70]. A homologue of human PHD2, P4H (prolyl 4-hydroxylase) has been reported to be the oxygen sensor in *Dictyostelium discoideum*. Unlike the HIF system in humans, *Dictyostelium discoideum* lacks HIF-α. Instead,

Dictyostelium discoideum P4H (DdP4H, previously known as DdP4H1) modifies DdSkp1 (*Dictyostelium discoideum* S-phase kinase associated protein 1), an adaptor protein of the SCF (Skp1-Cul1-F-box) E3-ubiquitin ligase complex [70]. Although there is little information about the function of SCF E3-ubiquitin ligase complex in *Dictyostelium discoideum*, it can be speculated that *Dictyostelium discoideum* SCF E3-ubiquitin ligase complex may share similar function as its orthologue, human SCF E3-ubiquitin ligase complex, which is reported to play a very important role in the cell cycle, by ubiquitination of several cell cycle proteins to trigger their degradation [71].

1.4.3 *Dictyostelium discoideum* S-Phase Kinase Associated Protein 1 (DdSkp1)

In eukaryotes, Skp1 is expressed universally in the cytoplasm and nucleus, and like pVHL in humans, is a linker protein, in this case for a SCF E3-ubiquitin ligase complex. Skp1 can regulate the activity of SCF by binding to another part of the SCF complex [72]. A detailed explanation of human Skp1 structure and function is given in Chapter 1.9.

In *Dictyostelium discoideum*, Skp1 is modified by the hydroxylation of Pro143 as catalysed by DdP4H, followed by the attachment of a pentasaccharide unit to the hydroxyproline by glycosyltransferases to form Galactose-Galactose-Fucose-Galactose-N-Acetyl-D-galactosamine-HyPro143-DdSkp1 [73-77]. This pathway is reported to be related to the oxygen regulated development of *Dictyostelium discoideum* [76, 78], which may provide a clue to the evolutionary origin of an oxygen sensing system in pre-metazoan species.

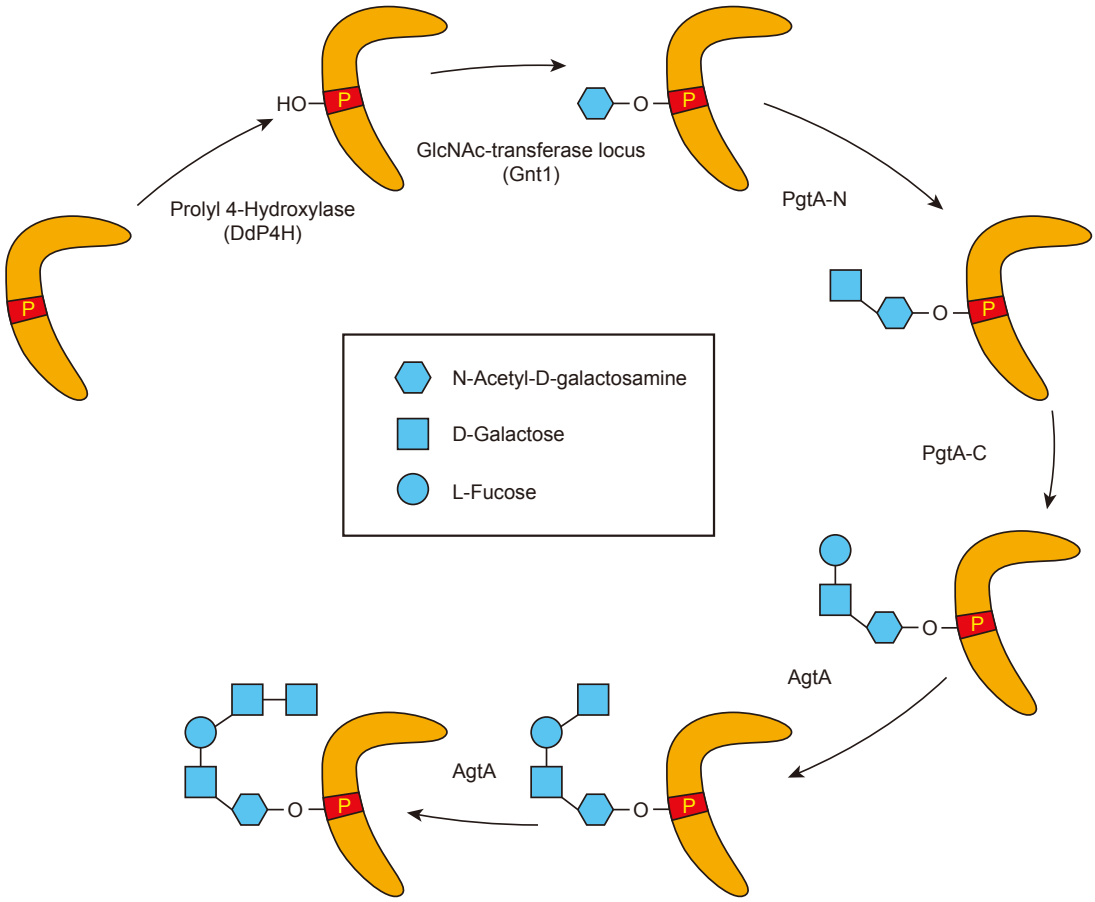


Fig. 1.8 The *Dictyostelium discoideum* Skp1 (DdSkp1) modification pathway. In normoxic conditions, DdSkp1 is hydroxylated on Pro143 by DdP4H. The hydroxyproline is then modified by attachment of an N-aceltyl-D-glalactosamine (GlcNac) by the α-N-acetylglucosaminyltransferase (Gnt1). DdP4H and Gnt1 have related mechanisms for recognizing and modifying DdSkp1, although Gnt1 does not activate or inhibit DdP4H [79]. The second and third sugars are then added to the GlcNac by PgtA, a dual function β3-galactosyltransferase / α2-fucosyltransferase [73]. This structure forms the type I blood group H antigen (Fucα1,2Galβ1,-3GlcNAc-) [80]. Addition of the fourth and the fifth sugar is catalysed by AgtA, resulting in modification of fucose with two other galactoses in a 1,3-linkage [81].

The hydroxylation and glycosylation of DdSkp1 is a highly regulated pathway with feedbacks [82]. DdSkp1 is the functional target of DdP4H in oxygen regulation [82]. The prolyl hydroxylation step acts as the gatekeeper reaction for subsequent glycosylation by PgtA. The second and third sugars are then added to the GlcNac by PgtA, a dual function β3-galactosyltransferase / α2-fucosyltransferase

[73] forming the type I blood group H antigen (Fucα1,2Galβ1,-3GlcNAc-) structure [80]. The C-terminal WD40 repeat domain of AgtA catalyses the attachment of the last two galactoses while inhibiting the recognition of DdSkp1 by DdP4H and Gnt1 [82]. AgtA mediated glycosylation was firstly thought to be regulated by oxygen, since it was crucial for oxygen-dependent terminal development [82]. Both P4H knock-out strain and AgtA knock-out strain need higher oxygen level to culminate [82]. However, it was later found that it is PgtA that exhibits a DdP4H-dependent antagonistic function, which can be reversed by AgtA [73]. AgtA also inhibits the binding of F-box binding protein (FBP) to the unhydroxylated DdSkp1. After being hydroxylated and glycosylated, DdSkp1 binds more efficiently to FBP forming a more stable SCF complex [74, 83].

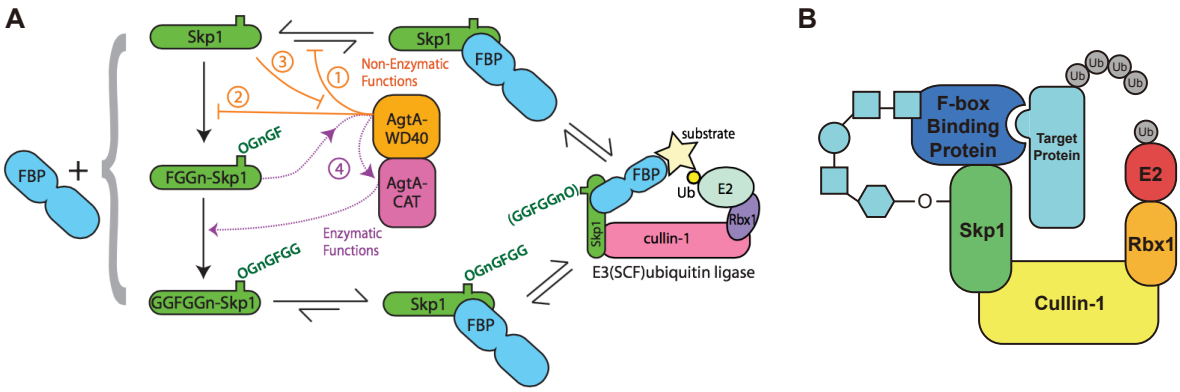


Fig. 1.9 The feedback pathway of Skp1 modification. (A) Proposed impact of AgtA on Skp1 interactions [82]. AgtA is both a Skp1-binding protein and an enzyme that catalyses the galactosylation to form the pentasaccharide attached to Skp1. FBP prefer binding to the glycosylated Skp1 to the unhydroxylated Skp1. Figure is taken from (Schafer 2014), permission acquired. (B) Model of SCF complex in the context of the binding to E2 ubiquitin ligase and a target ubiquitinated protein.

Dictyostelium discoideum has two Skp1 genes, Skp1A and Skp1B, which differ by one intron, comprising 33 nucleotides, and one amino acid (Ser39 in DdSkp1A, Ala39 in DdSkp1B) in the N-terminal region [81, 84, 85]. DdSkp1A and DdSkp1B are present both in the cytoplasm and the nucleus, the genes for them are expressed and transcribed at the same level throughout development, as

shown by reverse transcription-polymerase chain reaction (RT-PCR) [86]. The difference in function between the two isoforms is unclear. In this thesis DdSkp1 represents DdSkp1A, unless otherwise specified.

Being the element of SCF ubiquitin ligase complex, the main function of DdSkp1 is to bind to the F-box protein. It was predicted that there are 55 FBP isoforms in *Dictyostelium discoideum* [87]. Two FBPs (FbxD and FbxA) prefer binding to the glycosylated DdSkp1 rather than the unmodified DdSkp1 [74]. Circular dichroism and small-angle X-ray scattering studies support the tendency of glycosylated DdSkp1 binding to the F-box protein with a consequent decrease of the tendency of the Skp1 to dimerise [83]. It is possible that an equilibrium between the modified and unmodified DdSkp1 exists, which affects the function of SCF ubiquitin ligase complex by regulating the binding affinity of DdSkp1 to the F-box protein.

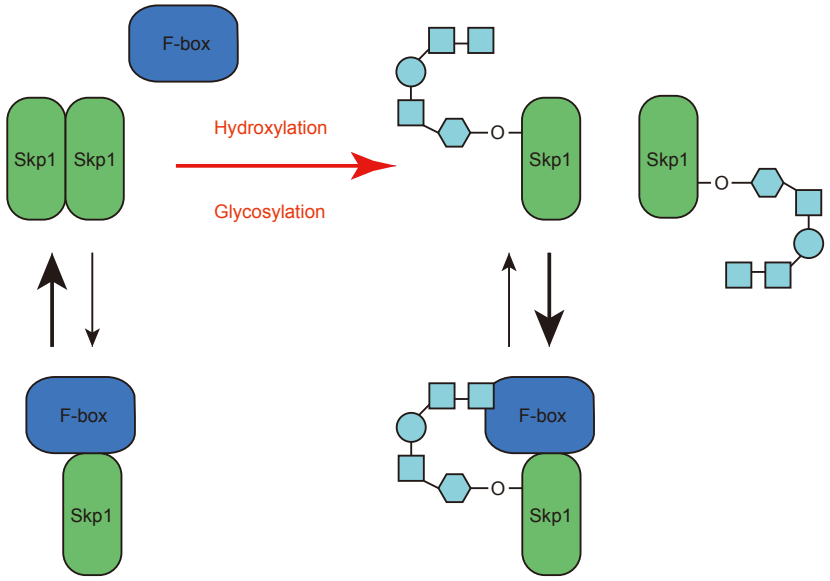


Fig. 1.10 The binding state of DdSkp1. In the unhydroxylated state, DdSkp1 tends to dimerise in an antiparallel form, with a lower affinity for the F-box protein. After being hydroxylated and glycosylated by DdP4H and the glycosyltransferases, the modified DdSkp1 detaches from the dimer, preferring to bind to the F-box protein, forming a more stable SCF ubiquitin ligase complex [83].

1.4.4 *Dictyostelium discoideum* Prolyl 4-Hydroxylase (DdP4H)

DdP4H is reported to hydroxylate Pro143 of DdSkp1, with its activity being limited by oxygen levels. Cellular studies indicate that the K_m for oxygen of DdP4H corresponds to 40% oxygen [70]. DdP4H knock-out strain needs higher oxygen level to culminate [81]. Over expressing DdP4H can complement the DdP4H deficiency in DdP4H knock-out strains, but is dependent on the hydroxylation activity [61]. At a physiological oxygen level, DdP4H regulates the culmination of the *Dictyostelium discoideum* slug phase. The DdP4H sequence is 57 % similar (29 % identical) to the catalytic domain of human PHD2. A structure for DdP4H had not been solved at the onset of my work.

1.5 *Toxoplasma gondii*

1.5.1 The Life Cycle of *Toxoplasma gondii*

Toxoplasma gondii is a protozoan parasite with a complex life cycle (Figure 1.11). In feline definitive hosts, *Toxoplasma gondii* undergoes a sexual cycle, when oocysts are produced. The oocysts can infect intermediate hosts, including humans by ingestion [88]. The oocysts can also be transmitted between hosts either vertically (mother–fetus) or horizontally (carnivorism). After entering the intermediate hosts, they develop into bradyzoites. Bradyozoites are the slow-growing, transmissible, and encysted form of the parasite, with resistance to gastric peptidases in the host stomach [89]. After entering the small intestine, *Toxoplasma gondii* turns into a rapid cellular dividing state, Tachyzoites [89]. Tachyzoites in pregnant women can infect the fetus. Eventually, tachyzoites localise to muscle tissues and the central nervous system (CNS) and result in diseases [90]. In host bodies, bradyzoites and tachyzoites are interconvertable [89].

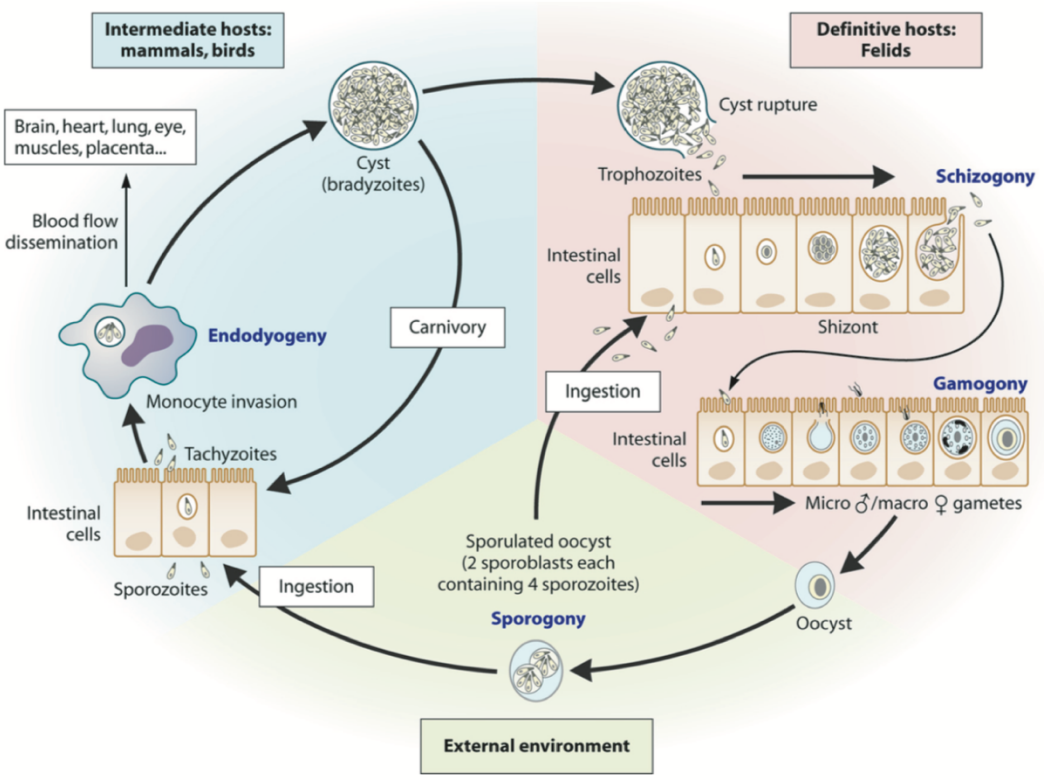


Fig. 1.11 Life cycle of *Toxoplasma gondii* [91]. The biology, infection, and replication of the three infective stages of the parasites. The figure is taken from (Robert-Gangneux 2012); permission acquired.

1.5.2 Toxoplasmosis

Infection with the protozoan parasite *Toxoplasma gondii* has a widespread distribution, causing a disease called toxoplasmosis. It has been estimated that 25 to 30% of the world's human population is infected by *Toxoplasma gondii* with prevalence varying between regions [91]. The seroprevalence is affected by multiple factors, including climatic factors [91]. Because toxoplasmosis can be transmitted by ingesting oocysts, dietary habits, culture habits, quality of water, sanitation coverage can also contribute to the spread of *Toxoplasma gondii* [91]. The infection of *Toxoplasma gondii* can trigger human immune responses, resulting in enlarged lymph nodes, fever, fatigue, muscle pain, sore throat and headache [90]. Although the condition may be mild in most humans, toxoplasmosis can be

severe and fatal for infants and patients with acquired immunodeficiency syndrome (AIDS), resulting in encephalitis [90]. Toxoplasma during pregnancy can be dangerous resulting in stillbirth, miscarriage or congenital toxoplasmosis. Approximately 30% of AIDS patients in Europe are estimated to die from toxoplasmosis [90], making it a high-ranking disease leading to lethality for AIDS patients.

Most healthy people don't require toxoplasmosis treatment, because in most cases, the parasite will not survive the adaptive immune response. In severe cases, however, several drugs are used the treatment of toxoplasmosis. The typical treatment choice of drugs includes pyrimethamine in combination with sulphonamide, clindamycin, atovaquone, spiramycin and prophylactic treatment. Trimethoprim/sulfamethoxazole has been used for *Toxoplasma gondii* seropositive patients with AIDS [92].

1.5.3 The Oxygen Sensing System of *Toxoplasma gondii*

By analogy to *Dictyostelium* Skp1, the reported process of *Toxoplasma* Skp1 (TgSkp1) hydroxylation and glycosylation suggests that it may be the oxygen sensing system for *Toxoplasma gondii* in a mechanism involving regulation of the assembly of the SCF complex [93]. Pro154 of TgSkp1, the subunit of the Skp1/Cullin-1/F-box protein (SCF)-class of E3-ubiquitin ligases in *Toxoplasma gondii*, is a natural substrate of the *Toxoplasma* prolyl-4-hydroxylase (TgPhyA) enzyme [62]. The hydroxylation requires 2OG and Fe(II) similarly to the reported hydroxylation of DdSkp1 by DdP4H [62]. Like its *Dictyostelium* counterpart, the *Toxoplasma* TgSkp1 hydroxyproline is then modified by a similar pentasaccharide by *Toxoplasma* homologues of Gnt1 and PgtA forming a Hexose-Hexose-Fucose-Galactose-N-Acetylglucosamine-Hydroxyproline154-TgSkp1 modification [93]. The mechanism

of attachment of the two terminal hexose residues was unknown because of the apparent absence of a homologue for AgtA in the *Toxoplasma* genome [93].

The remarkable similarity of the hydroxylation and glycosylation pathway between *Dictyostelium discoideum* and *Toxoplasma gondii* suggests that TgPhyA and TgSkp1 may play the role of the oxygen sensor system in *Toxoplasma gondii*. It is reported through gene-disruption experiments that TgPhyA is required for Skp1 glycosylation in *Toxoplasma gondii* and that disrupted parasites grow slowly at physiological oxygen levels [62]. However, biochemical analysis suggests that TgPhyA exhibits an unexpectedly higher affinity for oxygen compared to DdP4H [62], which means that only at low oxygen levels would the hydroxylation efficiency of TgPhyA be affected. Such a difference between *Dictyostelium discoideum* and *Toxoplasma gondii* may be due to their different life environments. *Dictyostelium discoideum* lives in the soil where the oxygen level fluctuates between 5 – 21% [62]. Whereas *Toxoplasma gondii*, as an animal parasite, resides most of the time in host cells or tissues as tachyzoites or bradyzoites, where the oxygen concentration can be lower than 1% [94].

TgPhyA is reported to exhibit a similar K_m for 2OG as DdP4H and can function in place of DdP4H in O_2-dependent culmination of *Dictyostelium discoideum* [62]. It has been proposed that TgPhyA can be regulated by oxygen in an indirect way through 2OG regulation, since the production of 2OG through TCA cycle is also regulated by oxygen levels [87].

Infection by *Toxoplasma gondii* can activate host HIF-1α through an activin-like receptor signalling pathway possibly by increasing the stability of HIF-1α and inactivating the activity of host PHD2 [95]. On the other hand, HIF-1α promotes parasite growth in a way not yet understood[87]. *Toxoplasma* growth and survival was severely reduced in HIF-1α knock-out cells at 3% oxygen [96]. Therefore, the connection between the growth of the parasite and the host oxygen sensing system regulation may not be simply linked by the reduced oxygen levels in the microenvironment caused by the parasite

oxygen consumption, but rather by more complicated pathways, which may evolve from the prolonged attack and defence process between the parasite and the host [89, 96, 97].

1.6 Prolyl-Hydroxylases in Other Organisms

1.6.1 The Structure of *Pseudomonas* Prolyl-Hydroxylase Domain (PPHD) and its Substrate, Elongation Factor Thermo Unstable (EF-Tu)

Pseudomonas is a Gram-negative aerobic gammaproteobacteria, belonging to the family *Pseudomonadaceae* [98]. Well studied *Pseudomona* species include the opportunistic human pathogen, *Pseudomonas aeruginosa*, the plant pathogen, *Pseudomonas syringae*, the soil bacteria, *Pseudomonas putida*, and the plant growth-promoting *Pseudomonas fluorescens* [99].

Pseudomonas aeruginosa has a PHD homologue (PPHD, 29% identity to PHD2), for which its substrate is Elongation Factor Termo Unstable, EF-Tu [31, 100]. A crystal structure of *Pseudomonas putida* PPHD:EF-Tu complex has been obtained; the structure shows a highly conserved core DSBH fold of PPHD similar to that of other prolyl-hydroxylases [63]. In a PPHD:EF-Tu protein-protein complex crystal structure, PPHD:EF-Tu was co-crystallised with Mn(II) and NOG to stabilize the complex by interacting with PPHD; Mg(II) and GDP were used to stabilize EF-Tu. The active site metal of PPHD is chelated by a conserved HXD...H metal binding motif. The switch I loop of EF-Tu threads through a tunnel formed by the PPHD β2 - β3 finger and β5 - β6 thumb loops resulting in a close distance (~4.5 Å) between the catalytically metal ion and the target proline residue [63]. This structure suggests tight binding affinity to PPHD which may result in efficient hydroxylation.

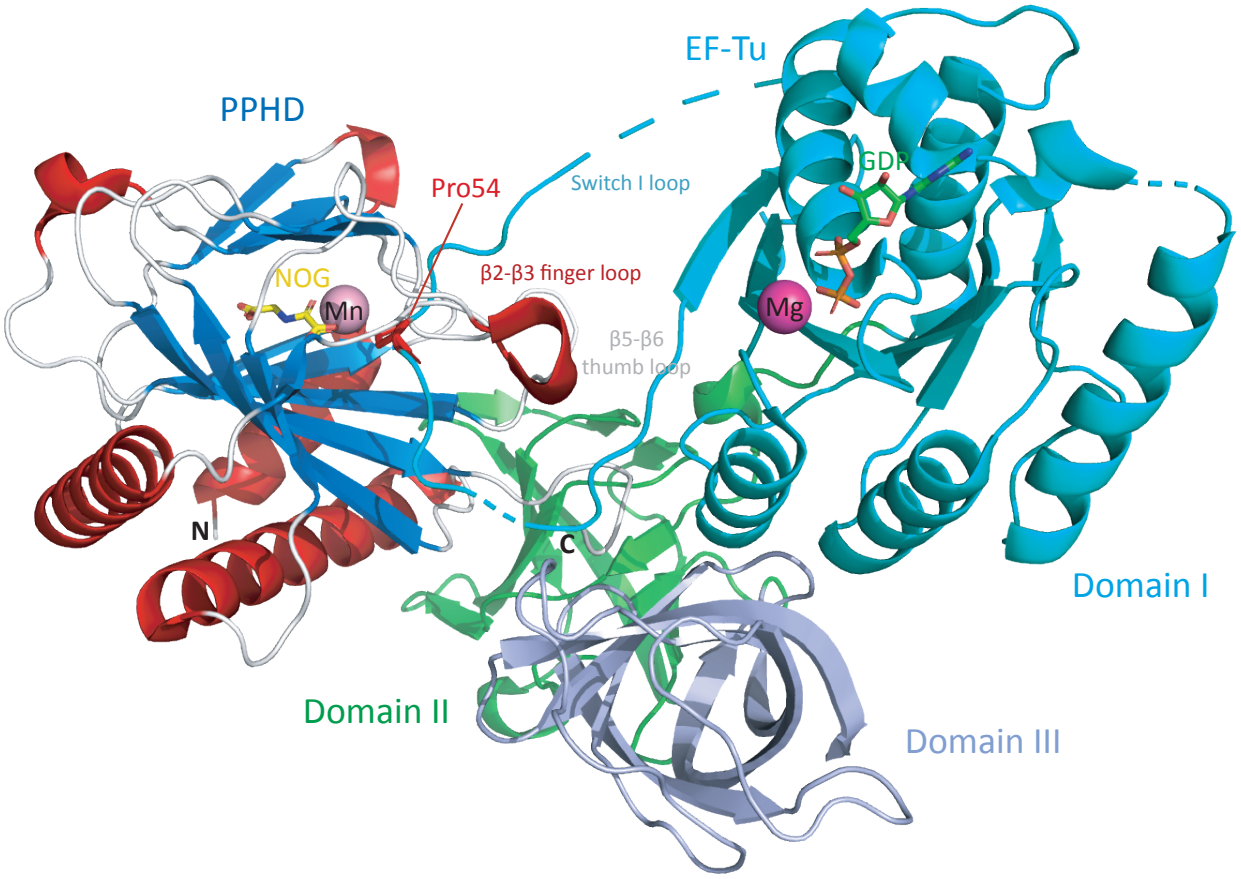


Fig. 1.12 View from a crystal structure of the PPHD:EF-Tu complex. Cartoon representation of the structure of PPHD:EF-Tu complex (PDB ID: 4IW3). β-Sheets of PPHD are coloured in cyan, α-Helices in red, loops in white. 2OG analogue, N-oxalylglycine (NOG), is shown as sticks residing in the active site surrounded by DSBH. The metal ion, Mn^{2+}, is shown as a pink sphere. EF-Tu is coloured in three colours (cyan, light blue and green) for its three different domains. The switch I loop of Domain I threads through the tunnel formed by β2-β3 finger loop and β5-β6 thumb loop. The target proline 54 is shown as sticks in red. The metal ion, Mg^{2+}, shown as a purple sphere, sits at the GTP catalytic site. A GDP molecule is shown as sticks chelating the central Mg^{2+}.

1.6.2 The Function of PPHD and EF-Tu

The function of PPHD has been partly investigated. A PPHD knock-out strain of *Pseudomonas aeruginosa* showed an increase of the level of secreted pyocyanin, a redox-active phenazine [100]. With their redox properties, pyocyanin functions in cellular signaling, iron acquisition, and redox homeostasis [101]. In addition, the growth of a *Pseudomonas aeruginosa* PPHD insertional mutant strain was more sensitive to levels of 2,2-bipyridine, an iron chelator, indicating the function of

PPHD may in part be related to iron regulation [63, 100]. Pyocyanin has been shown to down regulate genes involved in ferric iron acquisition [102]. 2OG is reported to be used as an electron donor by pyruvate and 2OG dehydrogenase complexes to directly catalyse the reduction of phenazine in *Pseudomonas aeruginosa* [101]. Also being the co-substrate of PPHD, 2OG levels may act as a link between EF-Tu hydroxylation and increased pyocyanin production. In addition, PPHD is reported to regulate the host neutrophil response to acute pseudomonal infection, probably through the secretion of phenazine metabolites, particularly pyocyanin [103]. This finding also hints that the prolyl hydroxylation, pyocyanin secretion, and iron regulation may all be closely related [63, 100].

Kinetic studies of PPHD and EF-Tu have been carried out. Using an EF-Tu 20mer oligopeptide (residues 44-63) as substrate, PPHD exhibits K_m^{app} for substrate (47.2 ± 8.5) μM, for 2OG (184 ± 22) μM, for oxygen (47.5 ± 9.2) μM [100]. A relatively low K_m^{app} for oxygen compared to human PHD2 [$K_m^{app}(O_2)$ > 450 μM] seems to suggest PPHD may not be an oxygen sensor, but it is worthy to note that *Pseudomonas aeruginosa* is a human pathogen. In some human tissues, the oxygen concentration can be lower than 1% [94]. Further, unpublished results suggest that the full-length EF-Tu protein is a better substrate than EF-Tu oligopeptide, resulting in a much lower K_m for substrate (personal communication from Patrick Trenkle [104]). This result indicates that PPHD binds the full-length EF-Tu more tightly than the 20mer peptide. Therefore, the K_m^{app} for oxygen using full-length EF-Tu as the substrate needs to be performed before concluding the potential function of PPHD as an oxygen sensor or not.

EF-Tu also has a catalytic function, participating in protein synthesis, the mechanism of which has been well studied for more than 20 years [105]. EF-Tu takes part in the delivery of aminoacyl-tRNAs — the primary substrate of the ribosome, relying on the formation of a ternary complex with EF-Tu and GTP [105-107]. The aminoacyl-tRNA (aa-tRNA) firstly accumulates on the A-site of the ribosome and

is later released by the EF-Tu after GTP hydrolysis [107], resulting in elongation of the synthesized peptide.

In the PPHD:EF-Tu complex structure (Fig. 1.12), the active site of EF-Tu has a GDP and a Mg^{2+} bound [63]. It was previously reported that EF-Tu undergoes a drastic conformational change between its active and inactive forms when bound to GTP and GDP, respectively [106]. Also, the switch 1 loop of EF-Tu shows significant local change in its tertiary and secondary structures upon activation [108, 109]. These findings for EF-Tu reveal that EF-Tu is a multifunctional protein. The two different reactions observed for EF-Tu may be linked or even regulated by each other.

1.6.3 *Monosiga brevicollis* Prolyl-4-Hydroxylase (MbP4H)

Monosiga brevicollis is one of the choanoflagellates, a group of microscopic, heterotrophic single-celled colony-forming eukaryotes. Because its full genome has been sequenced [110], *Monosiga brevicollis* is a useful organism for studying the early evolutionary relationships to multicellular animals. Analysis of mitochondrial genomes and expression sequence tags (ESTs) from various protists support the theory that choanoflagellates are the closest living unicellular organism to animals [111, 112]. A BLAST search using the human PHD2 amino acid sequence yields a homologue of PHD2 in *Monosiga brevicollis* [113]. The result also shows that MbP4H processes more similarity with PHD2 than that of DdP4H, which is consistent with the position of *Monosiga brevicollis* in phylogenetic tree (Fig 1.6). The *Monosiga* prolyl-4-hydroxylase (MbP4H) was then produced and studied. Although MbP4H can hydroxylate prolines in *Trichoplax adhaerens* HIF oxygen-dependent degradation domains (TaODD), a substrate for MbP4H in *Monosiga brevicollis* has not yet been identified [113]. The absence of an ODD sequence in the *Monosiga brevicollis* genome suggests that it may not possess a HIF system [113], which supports the hypothesis that HIF system is only conserved

in animals [58]. Because of the limited numbers of studies of *Monosiga brevicollis*, there is a lack of knowledge of its oxygen sensing system.

1.7 Evolution of Oxygen Sensing Systems

Animals exhibit a highly conserved oxygen sensing system, the HIF system (Section 1.2). Hypoxic regulation by PHDs and HIFs likely exists in most animals, including the simplest animal, *Trichoplax* [58]. However, even the closest relative to animals, *Monosiga brevicollis*, appears not to possess a HIF system. As described above, the PHD/P4H – Skp1 system is reported to be an alternative oxygen sensing system for protists [114]. In animals, normoxia results in the hydroxylation and degradation of HIF-α by the E3 ubiquitin ligase [21, 58]. In the protist pathway, where HIF-α does not exist, oxygen sensing is related to the S-phase kinase associated protein 1 (Skp1), the adaptor protein of the SCF (Skp1-Cul1-F-box) E3-ubiquitin ligase complex, which is hydroxylated under normoxic conditions, resulting in a subsequent series of glycosylations of HO-Skp1; glycosylation affects the conformation and catalytic efficiency of the SCF (Skp1-Cul1-F-box) E3 ubiquitin ligase [115].

Therefore, it has been proposed that although many organisms have long developed mechanisms to cope with and utilize oxygen for metabolism throughout the history of life, including the P4H-Skp1 oxygen sensing system in protists, metazoans have evolved a different, yet possibly related, PHD-HIF system. This PHD-HIF system lead to a more specialised function for both Skp1 and HIF, potentially resulting in more efficient and less redundant systems [116, 117]. This difference may provide an example of functional diversity in protein/animal evolution [117].

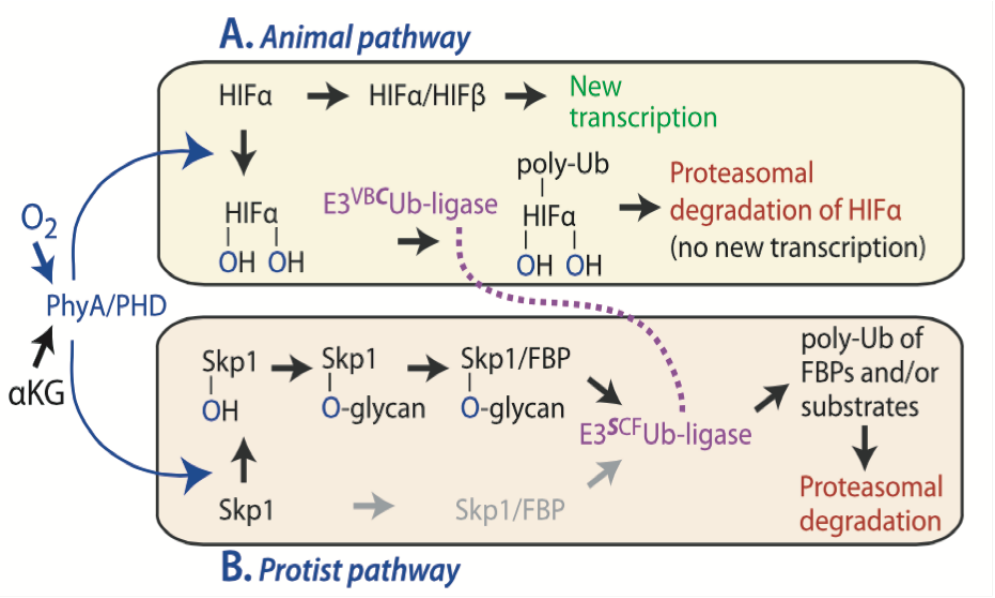


Fig. 1.13 The comparison between the oxygen sensing systems in both animals and protists [114]. Figure is taken from (Xu 2012), permission acquired. (A) Animal oxygen sensing system. HIF-α plays a central role in oxygen sensing by normoxic hydroxylation by PHD2 (in human) resulting in its proteasomal degradation. (B) The protist oxygen sensing system. Skp1 is a substrate of the prolyl hydroxylases and plays a role of the oxygen sensor. Instead of being affected by stability, hydroxylated Skp1, together with its subsequent glycosylation, affects the binding to the F-box protein and the activity of the SCF E3 ubiquitin ligase.

1.8 Skp1 in Humans

Human Skp1 helps to form the SCF E3 ubiquitin ligase complex [118]. The SCF E3 ubiquitin ligase relies on one of its components — an F-box protein (FBP) — for its specificity, thus, controlling the ubiquitin-mediated degradation of many cell-cycle-regulatory proteins and transcription factors [119]. Ubiquitination usually results in the formation of a covalent bond between ubiquitin and a lysine residue of the target protein. The mechanism requires the action of three enzymes: (a) an activating enzyme (E1) that activates the C terminus of ubiquitin for nucleophilic attack; (b) a conjugating enzyme (E2) that transiently carries the activated ubiquitin molecule; and (c) a ligase (E3) that transfers the activated ubiquitin from the E2 to the substrate (or ubiquitin) lysine residue. This

three-step mechanism initiates all known ubiquitination reactions, independent of whether the substrate-bound ubiquitin(s) signal in proteasomal proteolysis, endocytosis, or some other fates [120]. Among all the SCF ubiquitin ligases, the SCF^{Skp2} ubiquitin ligase complex, of which a crystal structure has been determined [121], regulates the G1/S transition in mammalian cells by controlling the abundance of the cyclin-dependent kinase inhibitor p27Kip1 [122]. Cullin 1 (Cul1) interacts with RING-box protein 1 (Rbx1) through its C terminus and with Skp1 through its N terminus. The interaction of the F-box protein with Skp1 occurs through the F-box motif, an ~40-amino acid motif. The F-box protein contains additional protein interaction domains that bind to the target proteins [121-124].

Three types of multisubunit E3s with a small RING finger protein as a crucial component has been identified. The RING finger proteins are named Multisubunit RING E3s [120]. These enzymes are SCF E3s, the anaphase-promoting complex (APC), and the VCB E3s (pVHL-elongin C-elongin B). In the SCF E3s, substrate-specific F-box proteins are recruited to SCF complexes with Skp1 through the F-box motif. The substrate-recognition subunit of VCB is pVHL, recruited to the complex through interactions with the heterodimeric adaptor Elongin B/C [120]. Although elongin C has 56 % sequence similarity to Skp1, it lacks the C-terminal region where the target proline residue is positioned in *Dictyostelium discoideum* and *Toxoplasma gondii*. In comparison, APC is less well understood, since no high-resolution crystallographic structure of the APC has been acquired, and the functions of most other APC subunits are not clear [118].

Interestingly, unlike the Skp1s in *Dictyostelium discoideum* or *Toxoplasma gondii*, human Skp1 does not have an analogous proline residue in its C-terminal region (Fig. 1.15). Instead, the equivalent residue at the proline-position is a glutamate [Fig 1.14 (B)], which is unlikely to be hydroxylated by prolyl hydroxylases. It is reported that in humans, the instability of the F-box proteins is a mechanism for rapid remodeling of SCF complexes in response to changes in cell cycle stage or

metabolic status [120]. We can infer that the hydroxylation and glycosylation of the Skp1s in the F-box motif in *Dictyostelium discoideum* and *Toxoplasma gondii* may affect the stability of the SCF ubiquitin ligase complex, adding another layer of regulation to the ubiquitination process through the SCF ubiquitin ligase complex.

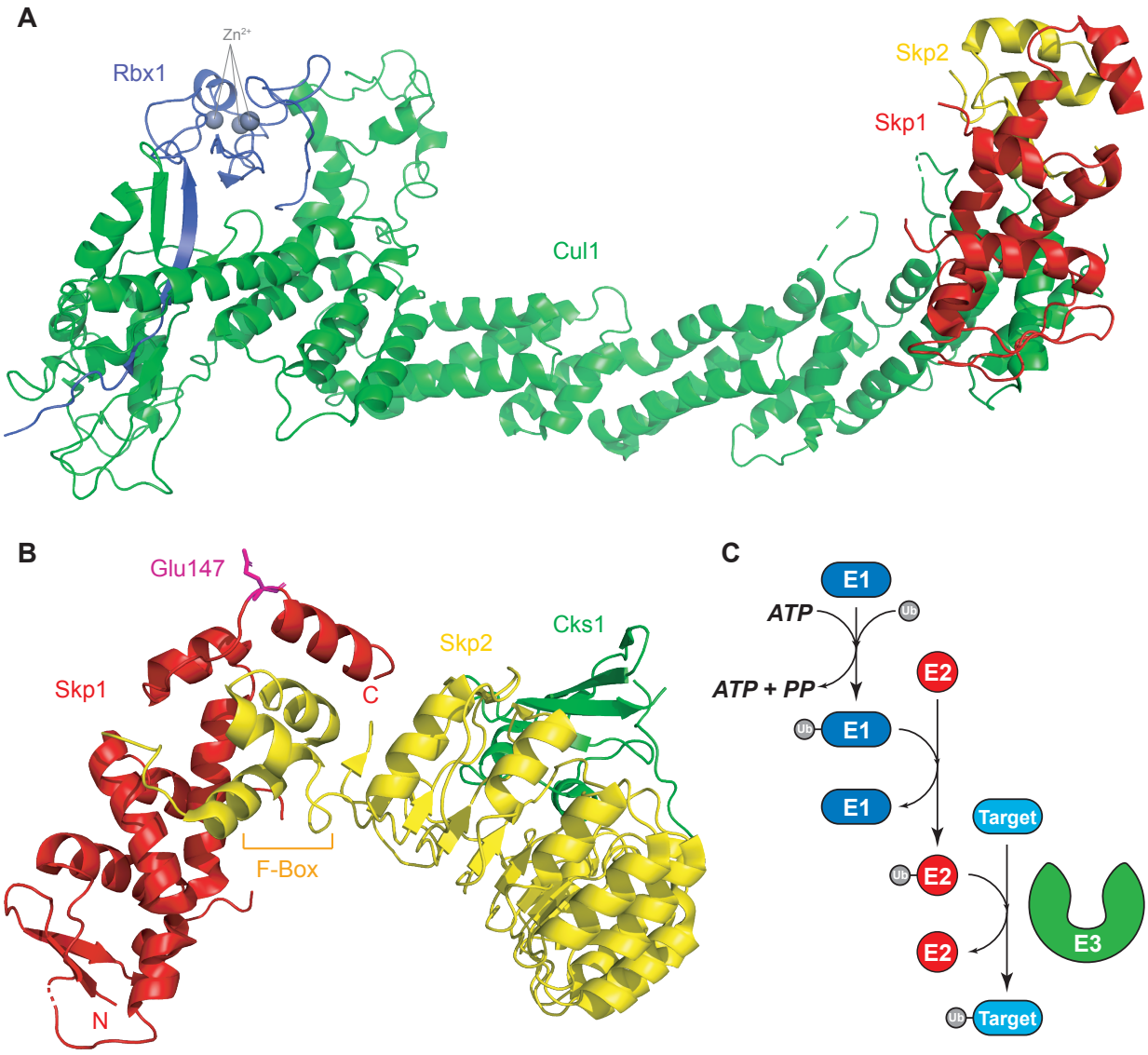


Fig. 1.14 The role of human Skp1 in the SCF ubiquitin ligase complex. (A) Cartoon representation of the structure of Cul1–Rbx1–Skp1–F-box^{Skp2} SCF ubiquitin ligase complex (PDB ID: 1LDK) [121]. Skp1 is coloured in red, Skp2 in yellow, cullin 1 (Cul1) in green, and RING-box protein 1 (Rbx1) in cyan. Three Zn^{2+} ions are labelled as grey spheres. Skp1 in this structure lacks the C-terminal sequence of 20 amino acids. (B) Cartoon representation of the structure of Skp1-Skp2-Cks1 complex (PDB ID: 2AST) [124]. Skp1 is coloured in red, Skp2 in yellow, cyclin-dependent kinases regulatory subunit (Cks1)

in green. The ~40-residue F-box motif, which is the common feature of the F-box proteins, is highlighted [122]. Glutamate 147 is shown as magenta sticks. In *Dictyostelium discoideum* and *Toxoplasma gondii*, the amino acid at this position is the target proline, which can be hydroxylated by DdP4H and TgPhyA, respectively. (C) Steps in protein ubiquitination [118]. Protein ubiquitination begins with the hydrolysis of ATP and the attachment of the ubiquitin protein to the ubiquitin-activating enzyme (E1). Then the ubiquitin is transferred to the ubiquitin-conjugating enzyme (E2), before attaching to the target protein by the collaboration with the target-specific ubiquitin ligase complex (E3). Additional ubiquitins are then added following the same step. The ubiquitin chains are then recognised by the 26S proteasome, leading to the degradation of the target protein.

1.9 Skp1 in Different Organisms

As described above, by contrast with protist Skp1s, human Skp1 does not have the proline residue in the C terminal region, making Skp1 an unlikely substrate for prolyl-hydroxylases. Thus, it is worth investigating Skp1s in different organisms to learn more about their oxygen sensing systems.

From the sequence alignment of Skp1s from different organisms (Fig. 1.15), one can observe that the protist hydroxylated proline in Skp1 is only conserved in protists, *Trichoplax adhaerens*, as well as some insects. In most animals, however, Skp1 does not have the proline in the C-terminal region.

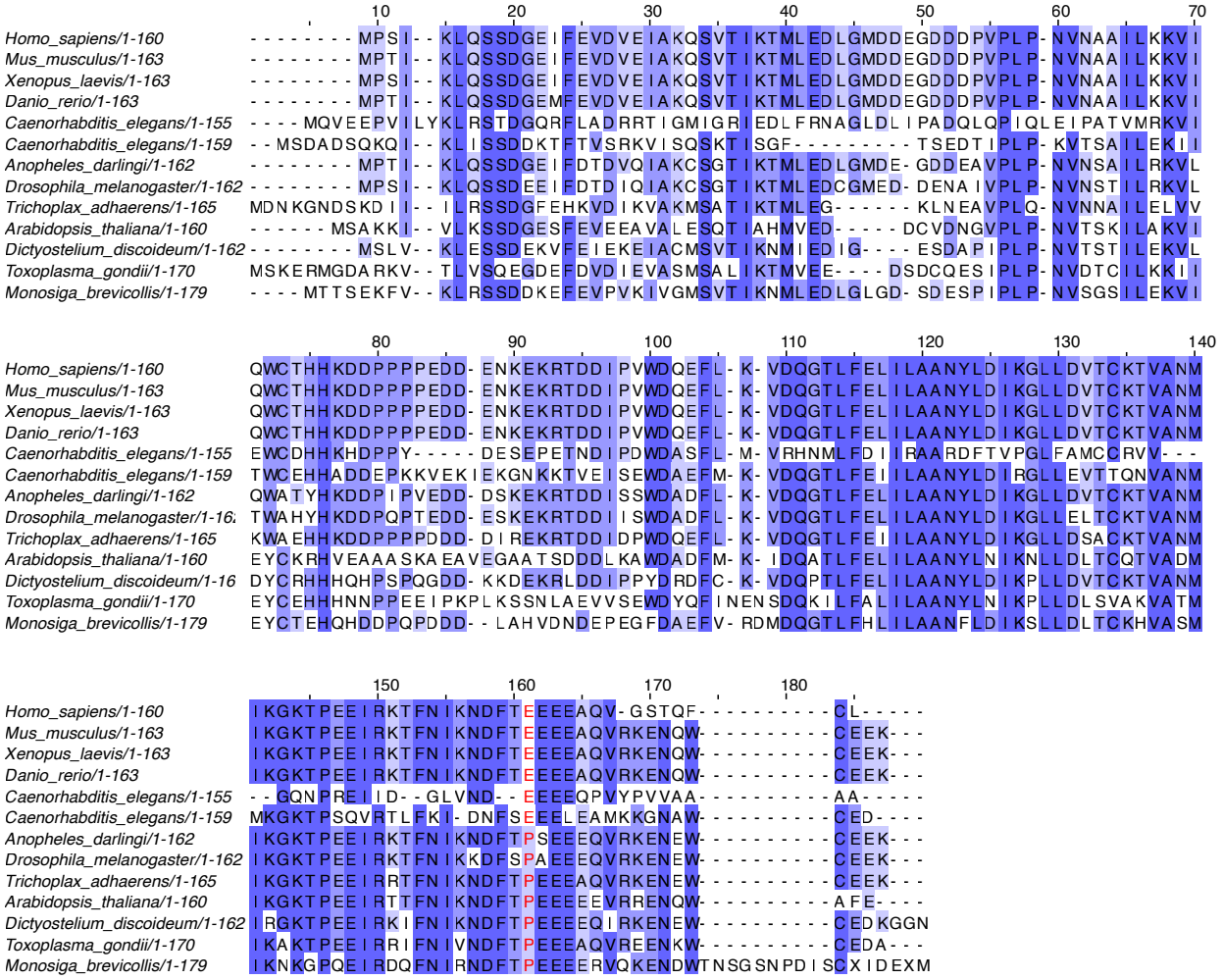


Fig. 1.15 Sequence alignment of Skp1s in different organisms. Organisms include human (*Homo Sapiens*, Gene ID: NP_008861.2), mouse (*Mus musculus*, Gene ID: NP_035673.3), frog (*Xenopus laevis*, Gene ID: AAF65619.1), zebra fish (*Danio rerio*, Gene ID: NP_957037.1), roundworm (*Caenorhabditis elegans*, Gene ID: NP_510193.4), roundworm (*Caenorhabditis* elegans, Gene ID: NP_507857.1), mosquito (*Anopheles darlingi*, Gene ID: ETN58923.1), fruitfly (*Drosophila melanogaster*, Gene ID: NP_477390.1), *Trichoplax adhaerens* (Gene ID: EDV18896.1), *Arabidopsis thaliana* (Gene ID: AEE35780.1), *Dictyostelium discoideum* (Gene ID: XP_644826.1), *Toxoplasma gondii* (Gene ID: CAJ20499.1), *Monosiga brevicollis* (Gene ID: EDQ90844.1, predicted protein). The amino acid locus at which the proline can be hydroxylated by PHD/P4H in *Dictyostelium discoideum* and *Toxoplasma gondii* is labelled in red.

In terms of oxygen sensing systems, only the PHD/P4H-Skp1 pathway has been identified in *Dictyostelium discoideum* and *Toxoplasma gondii* [87]. As described above, a different system, the PHD-HIF system, is the main chronic oxygen sensor for animals and it is highly conserved even in the

simplest animal, *Trichoplax adhaerens* (Section 1.2) [58]. Although a hypothetical Skp1 protein, with a proline residue in the C terminus, exists in *Trichoplax adhaerens*, whether it can be recognized and further hydroxylated by TaPHD was unknown prior to my work. Another protist, *Monosiga brevicollis*, has an apparent prolyl-hydroxylase (MbP4H), which is able to hydroxylate TaODD [113]. From the evolutionary perspective, it is of interest to know if the putative MbSkp1, which contains a C-terminal proline, is hydroxylated by MbP4H, possibly in a hypoxia sensing manner for this organism. *Pseudomonas* has neither Skp1 nor HIF. Instead, the substrate of PPHD is the elongation factor, EF-Tu [63], a reaction which may be involved in the iron regulation process. Whether the PPHD/EF-Tu is the oxygen sensor for this early prokaryote remains a mystery.

1.10 Thesis Objectives

The PHD-HIF oxygen sensing system is apparently only conserved in animals. Therefore, validating potential oxygen sensing systems in other organisms is important for understanding the evolution of oxygen sensing from a molecular perspective. One question addressed in this thesis is if Skp1 hydroxylation by P4H is a conserved process among protists and if it participates in the sensing of O_2. Cellular work on *Toxoplasma* and *Dictyostelium* suggests that Skp1 hydroxylation and glycosylation appeared in ancestral eukaryotes prior to their potential loss in fungi, higher plants, animals, and select protists [87]. These questions were probed using biophysical and biochemical methods.

Furthermore, it was considered that structural analyses may show the extent to which DdP4H and TgPhyA are conserved with the human PHDs. A thorough understanding of the structure and mechanism of *Dictyostelium* and *Toxoplasma* Skp1 hydroxylation will in turn help increase our understanding of human PHDs. Therefore, two objectives of the work described in this thesis were:

- To obtain structures of DdP4H and TgPhyA using crystallography.
- To define the kinetic properties of DdP4H-DdSkp1 and TgPhyA-TgSkp1 in order to investigate their possible oxygen sensing roles in early eukaryotes.

The inhibition of human 2OG oxygenases has been thoroughly studied in the context of the PHDs, and is of major current medical interest. Inhibition assays may also help our understanding of the structures and mechanisms of all prolyl hydroxylases. By comparing the inhibition results for different prolyl hydroxylases, it was hoped to gain a better understanding of the inhibitors of human oxygenases and gain insights into the specificity of drugs destined for the clinic. Moreover, a potent and specific inhibitor of TgPhyA may be useful as a probe for cellular and physiological studies as well as for development of potential drugs for treatment of toxoplasmosis. A third objective of the work described in this thesis was therefore:

- To identify potent inhibitors of DdP4H and TgPhyA to use as chemical probes to aid cellular and physiological biologists studying the prolyl hydroxylation pathways.

Lastly, by comparing the prolyl hydroxylation process and the structural information of different prolyl hydroxylases from different organisms, it was hoped to obtain a deeper understanding of the hydroxylase specificity, kinetic, and structural features of these enzymes, and so improve our understanding of the evolution of oxygen sensing systems.

1.11 References

1. Dole, M. (1965) The natural history of oxygen, *J Gen Physiol 49*, Suppl:5-27.

2. Taylor, C. T., and Pouyssegur, J. (2007) Oxygen, hypoxia, and stress, *Ann N Y Acad Sci 1113*, 87-94.

3. Taylor, C. T., and McElwain, J. C. (2010) Ancient atmospheres and the evolution of oxygen sensing via the hypoxia-inducible factor in metazoans, *Physiology (Bethesda) 25*, 272-279.

4. Fisher, S. A., and Burggren, W. W. (2007) Role of hypoxia in the evolution and development of the cardiovascular system, *Antioxid Redox Signal 9*, 1339-1352.

5. Costa, K. M., Accorsi-Mendonca, D., Moraes, D. J., and Machado, B. H. (2014) Evolution and physiology of neural oxygen sensing, *Front Physiol 5*, 302.

6. McLafferty, E., Johnstone, C., Hendry, C., and Farley, A. (2013) Respiratory system. Part 2: Gaseous exchange, *Nurs Stand 27*, 35-42.

7. McLafferty, E., Johnstone, C., Hendry, C., and Farley, A. (2013) Respiratory system part 1: pulmonary ventilation, *Nurs Stand 27*, 40-47.

8. Prabhakar, N. R. (2000) Oxygen sensing by the carotid body chemoreceptors, *J Appl Physiol (1985) 88*, 2287-2295.

9. Gonzalez, C., Agapito, M. T., Rocher, A., Gomez-Nino, A., Rigual, R., Castaneda, J., Conde, S. V., and Obeso, A. (2010) A revisit to O2 sensing and transduction in the carotid body chemoreceptors in the context of reactive oxygen species biology, *Respir Physiol Neurobiol 174*, 317-330.

10. Prabhakar, N. R., and Semenza, G. L. (2016) Regulation of carotid body oxygen sensing by hypoxia-inducible factors, *Pflugers Arch 468*, 71-75.

11. Michiels, C. (2004) Physiological and pathological responses to hypoxia, *Am J Pathol 164*, 1875-1882.

12. Fong, G. H., and Takeda, K. (2008) Role and regulation of prolyl hydroxylase domain proteins, *Cell Death Differ 15*, 635-641.

13. Markolovic, S., Wilkins, S. E., and Schofield, C. J. (2015) Protein Hydroxylation Catalyzed by 2-Oxoglutarate-dependent Oxygenases, *J Biol Chem 290*, 20712-20722.

14. Marxsen, J. H., Stengel, P., Doege, K., Heikkinen, P., Jokilehto, T., Wagner, T., Jelkmann, W., Jaakkola, P., and Metzen, E. (2004) Hypoxia-inducible factor-1 (HIF-1) promotes its degradation by induction of HIF-alpha-prolyl-4-hydroxylases, *Biochem J 381*, 761-767.

15. Jaakkola, P., Mole, D. R., Tian, Y. M., Wilson, M. I., Gielbert, J., Gaskell, S. J., von Kriegsheim, A., Hebestreit, H. F., Mukherji, M., Schofield, C. J., Maxwell, P. H., Pugh, C. W., and Ratcliffe, P. J. (2001) Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation, *Science 292*, 468-472.

16. Jelkmann, W. (2007) Erythropoietin after a century of research: younger than ever, *Eur J Haematol 78*, 183-205.

17. Ozolins, T. R., Fisher, T. S., Nadeau, D. M., Stock, J. L., Klein, A. S., Milici, A. J., Morton, D., Wilhelms, M. B., Brissette, W. H., and Li, B. (2009) Defects in embryonic development of EGLN1/PHD2 knockdown transgenic mice are associated with induction of Igfbp in the placenta, *Biochem Biophys Res Commun 390*, 372-376.

18. Semenza, G. L. (2001) Hypoxia-inducible factor 1: control of oxygen homeostasis in health and disease, *Pediatr Res 49*, 614-617.

19. McNeill, L. A., Hewitson, K. S., Claridge, T. D., Seibel, J. F., Horsfall, L. E., and Schofield, C. J. (2002) Hypoxia-inducible factor asparaginyl hydroxylase (FIH-1) catalyses hydroxylation at the beta-carbon of asparagine-803, *Biochem J 367*, 571-575.

20. Zhang, N., Fu, Z., Linke, S., Chicher, J., Gorman, J. J., Visk, D., Haddad, G. G., Poellinger, L., Peet, D. J., Powell, F., and Johnson, R. S. (2010) The asparaginyl hydroxylase factor inhibiting HIF-1alpha is an essential regulator of metabolism, *Cell Metab 11*, 364-378.

21. Chowdhury, R., Hardy, A., and Schofield, C. J. (2008) The human oxygen sensing machinery and its manipulation, *Chem Soc Rev 37*, 1308-1319.

22. Tarhonskaya, H., Hardy, A. P., Howe, E. A., Loik, N. D., Kramer, H. B., McCullagh, J. S., Schofield, C. J., and Flashman, E. (2015) Kinetic Investigations of the Role of Factor Inhibiting Hypoxia-inducible Factor (FIH) as an Oxygen Sensor, *J Biol Chem 290*, 19726-19742.

23. Zheng, X., Linke, S., Dias, J. M., Zheng, X., Gradin, K., Wallis, T. P., Hamilton, B. R., Gustafsson, M., Ruas, J. L., Wilkins, S., Bilton, R. L., Brismar, K., Whitelaw, M. L., Pereira, T., Gorman, J. J., Ericson, J., Peet, D. J., Lendahl, U., and Poellinger, L. (2008) Interaction with factor inhibiting HIF-1 defines an additional mode of cross-coupling between the Notch and hypoxia signaling pathways, *Proc Natl Acad Sci U S A 105*, 3368-3373.

24. Clifton, I. J., McDonough, M. A., Ehrismann, D., Kershaw, N. J., Granatino, N., and Schofield, C. J. (2006) Structural studies on 2-oxoglutarate oxygenases and related double-stranded beta-helix fold proteins, *J Inorg Biochem 100*, 644-669.

25. Tarhonskaya, H., Chowdhury, R., Leung, I. K., Loik, N. D., McCullagh, J. S., Claridge, T. D., Schofield, C. J., and Flashman, E. (2014) Investigating the contribution of the active site environment to the slow reaction of hypoxia-inducible factor prolyl hydroxylase domain 2 with oxygen, *Biochem J 463*, 363-372.

26. Yasumoto, K., Kowata, Y., Yoshida, A., Torii, S., and Sogawa, K. (2009) Role of the intracellular localization of HIF-prolyl hydroxylases, *Biochim Biophys Acta 1793*, 792-797.

27. 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., and Fandrey, J. (2003) Intracellular localisation of human HIF-1 alpha hydroxylases: implications for oxygen sensing, *J Cell Sci 116*, 1319-1326.

28. Appelhoff, R. J., Tian, Y. M., Raval, R. R., Turley, H., Harris, A. L., Pugh, C. W., Ratcliffe, P. J., and Gleadle, J. M. (2004) Differential function of the prolyl hydroxylases PHD1, PHD2, and PHD3 in the regulation of hypoxia-inducible factor, *J Biol Chem 279*, 38458-38465.

29. Takeda, K., Cowan, A., and Fong, G. H. (2007) Essential role for prolyl hydroxylase domain protein 2 in oxygen homeostasis of the adult vascular system, *Circulation 116*, 774-781.

30. McDonough, M. A., Loenarz, C., Chowdhury, R., Clifton, I. J., and Schofield, C. J. (2010) Structural studies on human 2-oxoglutarate dependent oxygenases, *Curr Opin Struct Biol 20*, 659-672.

31. 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., and Schofield, C. J. (2006) Cellular oxygen sensing: Crystal structure of hypoxia-inducible factor prolyl hydroxylase (PHD2), *Proc Natl Acad Sci U S A 103*, 9814-9819.

32. Aik, W., McDonough, M. A., Thalhammer, A., Chowdhury, R., and Schofield, C. J. (2012) Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases, *Curr Opin Struct Biol 22*, 691-700.

33. Flashman, E., Hoffart, L. M., Hamed, R. B., Bollinger, J. M., Jr., Krebs, C., and Schofield, C. J. (2010) Evidence for the slow reaction of hypoxia-inducible factor prolyl hydroxylase 2 with oxygen, *FEBS J 277*, 4089-4099.

34. Hirsila, M., Koivunen, P., Gunzler, V., Kivirikko, K. I., and Myllyharju, J. (2003) Characterization of the human prolyl 4-hydroxylases that modify the hypoxia-inducible factor, *J Biol Chem 278*, 30772-30780.

35. Dai, Z., Li, M., Wharton, J., Zhu, M. M., and Zhao, Y. Y. (2016) Prolyl-4 Hydroxylase 2 (PHD2) Deficiency in Endothelial Cells and Hematopoietic Cells Induces Obliterative Vascular Remodeling and Severe Pulmonary Arterial Hypertension in Mice and Humans Through Hypoxia-Inducible Factor-2alpha, *Circulation 133*, 2447-2458.

36. Minamishima, Y. A., Moslehi, J., Bardeesy, N., Cullen, D., Bronson, R. T., and Kaelin, W. G., Jr. (2008) Somatic inactivation of the PHD2 prolyl hydroxylase causes polycythemia and congestive heart failure, *Blood 111*, 3236-3244.

37. Duan, L. J., Takeda, K., and Fong, G. H. (2011) Prolyl hydroxylase domain protein 2 (PHD2) mediates oxygen-induced retinopathy in neonatal mice, *Am J Pathol 178*, 1881-1890.

38. Guentsch, A., Beneke, A., Swain, L., Farhat, K., Nagarajan, S., Wielockx, B., Raithatha, K., Dudek, J., Rehling, P., Zieseniss, A., Jatho, A., Chong, M., Santos, C. X., Shah, A. M., and Katschinski, D. M. (2017) PHD2 Is a Regulator for Glycolytic Reprogramming in Macrophages, *Mol Cell Biol 37*.

39. Chan, D. A., and Giaccia, A. J. (2010) PHD2 in tumour angiogenesis, *Br J Cancer 103*, 1-5.

40. Chan, M. C., Holt-Martyn, J. P., Schofield, C. J., and Ratcliffe, P. J. (2016) Pharmacological targeting of the HIF hydroxylases--A new field in medicine development, *Mol Aspects Med 47-48*, 54-75.

41. Haase, V. H. (2013) Regulation of erythropoiesis by hypoxia-inducible factors, *Blood Rev 27*, 41-53.

42. Bao, W., Qin, P., Needle, S., Erickson-Miller, C. L., Duffy, K. J., Ariazi, J. L., Zhao, S., Olzinski, A. R., Behm, D. J., Pipes, G. C., Jucker, B. M., Hu, E., Lepore, J. J., and Willette, R. N. (2010) Chronic inhibition of hypoxia-inducible factor prolyl 4-hydroxylase improves ventricular performance, remodeling, and vascularity after myocardial infarction in the rat, *J Cardiovasc Pharmacol 56*, 147-155.

43. Eltzschig, H. K., Bratton, D. L., and Colgan, S. P. (2014) Targeting hypoxia signalling for the treatment of ischaemic and inflammatory diseases, *Nat Rev Drug Discov 13*, 852-869.

44. Barrett, T. D., Palomino, H. L., Brondstetter, T. I., Kanelakis, K. C., Wu, X., Yan, W., Merton, K. P., Schoetens, F., Ma, J. Y., Skaptason, J., Gao, J., Tran, D. T., Venkatesan, H., Rosen, M. D., Shankley, N. P., and Rabinowitz, M. H. (2015) Prolyl hydroxylase inhibition corrects functional iron deficiency and inflammation-induced anaemia in rats, *Br J Pharmacol 172*, 4078-4088.

45. Hsieh, M. M., Linde, N. S., Wynter, A., Metzger, M., Wong, C., Langsetmo, I., Lin, A., Smith, R., Rodgers, G. P., Donahue, R. E., Klaus, S. J., and Tisdale, J. F. (2007) HIF prolyl hydroxylase inhibition results in endogenous erythropoietin induction, erythrocytosis, and modest fetal hemoglobin expression in rhesus macaques, *Blood 110*, 2140-2147.

46. Hong, Y. R., Kim, H. T., Lee, S. C., Ro, S., Cho, J. M., Kim, I. S., and Jung, Y. H. (2013) [(4-Hydroxyl-benzo[4,5]thieno[3,2-c]pyridine-3-carbonyl)-amino]-acetic acid derivatives; HIF prolyl 4-hydroxylase inhibitors as oral erythropoietin secretagogues, *Bioorg Med Chem Lett 23*, 5953-5957.

47. Chowdhury, R., Leung, I. K., Tian, Y. M., Abboud, M. I., Ge, W., Domene, C., Cantrelle, F. X., Landrieu, I., Hardy, A. P., Pugh, C. W., Ratcliffe, P. J., Claridge, T. D., and Schofield, C. J. (2016) Structural basis for oxygen degradation domain selectivity of the HIF prolyl hydroxylases, *Nat Commun 7*, 12673.

48. Besarab, A., Provenzano, R., Hertel, J., Zabaneh, R., Klaus, S. J., Lee, T., Leong, R., Hemmerich, S., Yu, K. H., and Neff, T. B. (2015) Randomized placebo-controlled dose-ranging and pharmacodynamics study of roxadustat (FG-4592) to treat anemia in nondialysis-dependent chronic kidney disease (NDD-CKD) patients, *Nephrol Dial Transplant 30*, 1665-1673.

49. Besarab, A., Chernyavskaya, E., Motylev, I., Shutov, E., Kumbar, L. M., Gurevich, K., Chan, D. T., Leong, R., Poole, L., Zhong, M., Saikali, K. G., Franco, M., Hemmerich, S., Yu, K. H., and Neff, T. B. (2016) Roxadustat (FG-4592): Correction of Anemia in Incident Dialysis Patients, *J Am Soc Nephrol 27*, 1225-1233.

50. Chowdhury, R., Candela-Lena, J. I., Chan, M. C., Greenald, D. J., Yeoh, K. K., Tian, Y. M., McDonough, M. A., Tumber, A., Rose, N. R., Conejo-Garcia, A., Demetriades, M., Mathavan, S., Kawamura, A., Lee, M. K., van Eeden, F., Pugh, C. W., Ratcliffe, P. J., and Schofield, C. J. (2013) Selective small molecule probes for the hypoxia inducible factor (HIF) prolyl hydroxylases, *ACS Chem Biol 8*, 1488-1496.

51. (2016) Burzlaff, N. (Friedrich Alexander University Erlangen-Nürnberg).

52. Chan, M. C., Atasoylu, O., Hodson, E., Tumber, A., Leung, I. K., Chowdhury, R., Gomez-Perez, V., Demetriades, M., Rydzik, A. M., Holt-Martyn, J., Tian, Y. M., Bishop, T., Claridge, T. D., Kawamura, A., Pugh, C. W., Ratcliffe, P. J., and Schofield, C. J. (2015) Potent and Selective Triazole-Based Inhibitors of the Hypoxia-Inducible Factor Prolyl-Hydroxylases with Activity in the Murine Brain, *PLoS One 10*, e0132004.

53. Pergola, P. E., Spinowitz, B. S., Hartman, C. S., Maroni, B. J., and Haase, V. H. (2016) Vadadustat, a novel oral HIF stabilizer, provides effective anemia treatment in nondialysis-dependent chronic kidney disease, *Kidney Int 90*, 1115-1122.

54. Martin, E. R., Smith, M. T., Maroni, B. J., Zuraw, Q. C., and deGoma, E. M. (2017) Clinical Trial of Vadadustat in Patients with Anemia Secondary to Stage 3 or 4 Chronic Kidney Disease, *Am J Nephrol 45*, 380-388.

55. Holdstock, L., Meadowcroft, A. M., Maier, R., Johnson, B. M., Jones, D., Rastogi, A., Zeig, S., Lepore, J. J., and Cobitz, A. R. (2016) Four-Week Studies of Oral Hypoxia-Inducible Factor-Prolyl Hydroxylase Inhibitor GSK1278863 for Treatment of Anemia, *J Am Soc Nephrol 27*, 1234-1244.

56. Franklin, T. J., Morris, W. P., Edwards, P. N., Large, M. S., and Stephenson, R. (2001) Inhibition of prolyl 4-hydroxylase in vitro and in vivo by members of a novel series of phenanthrolinones, *Biochem J 353*, 333-338.

57. Rytkonen, K. T., Williams, T. A., Renshaw, G. M., Primmer, C. R., and Nikinmaa, M. (2011) Molecular evolution of the metazoan PHD-HIF oxygen-sensing system, *Mol Biol Evol 28*, 1913-1926.

58. Loenarz, C., Coleman, M. L., Boleininger, A., Schierwater, B., Holland, P. W., Ratcliffe, P. J., and Schofield, C. J. (2011) The hypoxia-inducible transcription factor pathway regulates oxygen sensing in the simplest animal, *Trichoplax adhaerens*, *EMBO Rep 12*, 63-70.

59. Srivastava, M., Begovic, E., Chapman, J., Putnam, N. H., Hellsten, U., Kawashima, T., Kuo, A., Mitros, T., Salamov, A., Carpenter, M. L., Signorovitch, A. Y., Moreno, M. A., Kamm, K., Grimwood, J., Schmutz, J., Shapiro, H., Grigoriev, I. V., Buss, L. W., Schierwater, B., Dellaporta, S. L., and Rokhsar, D. S. (2008) The *Trichoplax* genome and the nature of placozoans, *Nature 454*, 955-960.

60. Lippl, K. (2018) Structural, Biochemical and Inhibition Studies on Hypoxia-Inducible Factor (HIF) Prolyl Hydroxylases.

61. West, C. M., van der Wel, H., and Wang, Z. A. (2007) Prolyl 4-hydroxylase-1 mediates O2 signaling during development of *Dictyostelium*, *Development 134*, 3349-3358.

62. Xu, Y., Brown, K. M., Wang, Z. A., van der Wel, H., Teygong, C., Zhang, D., Blader, I. J., and West, C. M. (2012) The Skp1 protein from Toxoplasma is modified by a cytoplasmic prolyl 4-hydroxylase associated with oxygen sensing in the social amoeba *Dictyostelium*, *J Biol Chem 287*, 25098-25110.

63. Scotti, J. S., Leung, I. K., Ge, W., Bentley, M. A., Paps, J., Kramer, H. B., Lee, J., Aik, W., Choi, H., Paulsen, S. M., Bowman, L. A., Loik, N. D., Horita, S., Ho, C. H., Kershaw, N. J., Tang, C. M., Claridge, T. D., Preston, G. M., McDonough, M. A., and Schofield, C. J. (2014) Human oxygen sensing may have origins in prokaryotic elongation factor Tu prolyl-hydroxylation, *Proc Natl Acad Sci U S A 111*, 13331-13336.

64. Eichinger, L., Pachebat, J. A., Glockner, G., Rajandream, M. A., Sucgang, R., Berriman, M., Song, J., Olsen, R., Szafranski, K., Xu, Q., Tunggal, B., Kummerfeld, S., Madera, M., Konfortov, B. A., Rivero, F., Bankier, A. T., Lehmann, R., Hamlin, N., Davies, R., Gaudet, P., Fey, P., Pilcher, K., Chen, G., Saunders, D., Sodergren, E., Davis, P., Kerhornou, A., Nie, X., Hall, N., Anjard, C., Hemphill, L., Bason, N., Farbrother, P., Desany, B., Just, E., Morio, T., Rost, R., Churcher, C., Cooper, J., Haydock, S., van Driessche, N., Cronin, A., Goodhead, I., Muzny, D., Mourier, T., Pain, A., Lu, M., Harper, D., Lindsay, R., Hauser, H., James, K., Quiles, M., Madan Babu, M., Saito, T., Buchrieser, C., Wardroper, A., Felder, M., Thangavelu, M., Johnson, D., Knights, A., Loulseged, H., Mungall, K., Oliver, K., Price, C., Quail, M. A., Urushihara, H., Hernandez, J., Rabbinowitsch, E., Steffen, D., Sanders, M., Ma, J., Kohara, Y., Sharp, S., Simmonds, M., Spiegler, S., Tivey, A., Sugano, S., White, B., Walker, D., Woodward, J., Winckler, T., Tanaka, Y., Shaulsky, G., Schleicher, M., Weinstock, G., Rosenthal, A., Cox, E. C., Chisholm, R. L., Gibbs, R., Loomis, W. F., Platzer, M., Kay, R. R., Williams, J., Dear, P. H., Noegel, A. A., Barrell, B., and Kuspa, A. (2005) The genome of the social amoeba *Dictyostelium discoideum*, *Nature 435*, 43-57.

65. Schaap, P. (2016) Evolution of developmental signalling in Dictyostelid social amoebas, *Curr Opin Genet Dev 39*, 29-34.

66. Bozzaro, S. (2013) The model organism *Dictyostelium discoideum*, *Methods Mol Biol 983*, 17-37.

67. Schaap, P. (2011) Evolutionary crossroads in developmental biology: *Dictyostelium discoideum*, *Development 138*, 387-396.

68. Sandona, D., Gastaldello, S., Rizzuto, R., and Bisson, R. (1995) Expression of cytochrome c oxidase during growth and development of *Dictyostelium*, *J Biol Chem 270*, 5587-5593.

69. Srinivas, U. K., and Katz, E. R. (1980) Oxygen utilization in *Dictyostelium discoideum*, *FEMS Microbiology Letters 9*, 53-55.

70. van der Wel, H., Ercan, A., and West, C. M. (2005) The Skp1 prolyl hydroxylase from *Dictyostelium* is related to the hypoxia-inducible factor-alpha class of animal prolyl 4-hydroxylases, *J Biol Chem 280*, 14645-14655.

71. Ang, X. L., and Wade Harper, J. (2005) SCF-mediated protein degradation and cell cycle control, *Oncogene 24*, 2860-2870.

72. Mathias, N., Johnson, S., Byers, B., and Goebl, M. (1999) The abundance of cell cycle regulatory protein Cdc4p is controlled by interactions between its F box and Skp1p, *Mol Cell Biol 19*, 1759-1767.

73. Wang, Z. A., van der Wel, H., Vohra, Y., Buskas, T., Boons, G. J., and West, C. M. (2009) Role of a cytoplasmic dual-function glycosyltransferase in O_2 regulation of development in *Dictyostelium*, *J Biol Chem 284*, 28896-28904.

74. Sheikh, M. O., Xu, Y., van der Wel, H., Walden, P., Hartson, S. D., and West, C. M. (2015) Glycosylation of Skp1 promotes formation of Skp1-cullin-1-F-box protein complexes in *Dictyostelium*, *Mol Cell Proteomics 14*, 66-80.

75. West, C. M., Van Der Wel, H., Sassi, S., and Gaucher, E. A. (2004) Cytoplasmic glycosylation of protein-hydroxyproline and its relationship to other glycosylation pathways, *Biochim Biophys Acta 1673*, 29-44.

76. West, C. M., Wang, Z. A., and van der Wel, H. (2010) A cytoplasmic prolyl hydroxylation and glycosylation pathway modifies Skp1 and regulates O_2-dependent development in *Dictyostelium*, *Biochim Biophys Acta 1800*, 160-171.

77. Teng-umnuay, P., Morris, H. R., Dell, A., Panico, M., Paxton, T., and West, C. M. (1998) The cytoplasmic F-box binding protein SKP1 contains a novel pentasaccharide linked to hydroxyproline in *Dictyostelium*, *J Biol Chem 273*, 18242-18249.

78. Zhang, D., van der Wel, H., Johnson, J. M., and West, C. M. (2012) Skp1 prolyl 4-hydroxylase of *Dictyostelium* mediates glycosylation-independent and -dependent responses to O_2 without affecting Skp1 stability, *J Biol Chem 287*, 2006-2016.

79. van der Wel, H., Johnson, J. M., Xu, Y., Karunaratne, C. V., Wilson, K. D., Vohra, Y., Boons, G. J., Taylor, C. M., Bendiak, B., and West, C. M. (2011) Requirements for Skp1 processing by cytosolic prolyl 4(trans)-hydroxylase and alpha-N-acetylglucosaminyltransferase enzymes involved in O(2) signaling in *Dictyostelium*, *Biochemistry 50*, 1700-1713.

80. Van Der Wel, H., Fisher, S. Z., and West, C. M. (2002) A bifunctional diglycosyltransferase forms the Fucalpha1,2Galbeta1,3-disaccharide on Skp1 in the cytoplasm of *Dictyostelium*, *J Biol Chem 277*, 46527-46534.

81. Wang, Z. A., Singh, D., van der Wel, H., and West, C. M. (2011) Prolyl hydroxylation- and glycosylation-dependent functions of Skp1 in O2-regulated development of *Dictyostelium*, *Dev Biol 349*, 283-295.

82. Schafer, C. M., Sheikh, M. O., Zhang, D., and West, C. M. (2014) Novel regulation of Skp1 by the *Dictyostelium* AgtA alpha-galactosyltransferase involves the Skp1-binding activity of its WD40 repeat domain, *J Biol Chem 289*, 9076-9088.

83. Sheikh, M. O., Schafer, C. M., Powell, J. T., Rodgers, K. K., Mooers, B. H., and West, C. M. (2014) Glycosylation of Skp1 affects its conformation and promotes binding to a model f-box protein, *Biochemistry 53*, 1657-1669.

84. Kozarov, E., van der Wel, H., Field, M., Gritzali, M., Brown, R. D., Jr., and West, C. M. (1995) Characterization of FP21, a cytosolic glycoprotein from *Dictyostelium*, *J Biol Chem 270*, 3022-3030.

85. West, C. M., Kozarov, E., and Teng-umnuay, P. (1997) The cytosolic glycoprotein FP21 of *Dictyostelium discoideum* is encoded by two genes resulting in a polymorphism at a single amino acid position, *Gene 200*, 1-10.

86. Sassi, S., Sweetinburgh, M., Erogul, J., Zhang, P., Teng-Umnuay, P., and West, C. M. (2001) Analysis of Skp1 glycosylation and nuclear enrichment in *Dictyostelium*, *Glycobiology 11*, 283-295.

87. West, C. M., and Blader, I. J. (2015) Oxygen sensing by protozoans: how they catch their breath, *Curr Opin Microbiol 26*, 41-47.

88. Kim, K., and Weiss, L. M. (2004) Toxoplasma gondii: the model apicomplexan, *Int J Parasitol 34*, 423-432.

89. Blader, I. J., and Saeij, J. P. (2009) Communication between *Toxoplasma gondii* and its host: impact on parasite growth, development, immune evasion, and virulence, *APMIS 117*, 458-476.

90. Hill, D., and Dubey, J. P. (2002) *Toxoplasma gondii*: transmission, diagnosis and prevention, *Clin Microbiol Infect 8*, 634-640.

91. Robert-Gangneux, F., and Darde, M. L. (2012) Epidemiology of and diagnostic strategies for toxoplasmosis, *Clin Microbiol Rev 25*, 264-296.

92. Halonen, S. K., and Weiss, L. M. (2013) Toxoplasmosis, *Handb Clin Neurol 114*, 125-145.

93. Rahman, K., Zhao, P., Mandalasi, M., van der Wel, H., Wells, L., Blader, I. J., and West, C. M. (2016) The E3 Ubiquitin Ligase Adaptor Protein Skp1 Is Glycosylated by an Evolutionarily Conserved Pathway That Regulates Protist Growth and *Development*, *J Biol Chem 291*, 4268-4280.

94. Sakadzic, S., Roussakis, E., Yaseen, M. A., Mandeville, E. T., Srinivasan, V. J., Arai, K., Ruvinskaya, S., Devor, A., Lo, E. H., Vinogradov, S. A., and Boas, D. A. (2010) Two-photon high-resolution measurement of partial pressure of oxygen in cerebral vasculature and tissue, *Nat Methods 7*, 755-759.

95. Wiley, M., Sweeney, K. R., Chan, D. A., Brown, K. M., McMurtrey, C., Howard, E. W., Giaccia, A. J., and Blader, I. J. (2010) *Toxoplasma gondii* activates hypoxia-inducible factor (HIF) by stabilizing the HIF-1alpha subunit via type I activin-like receptor kinase receptor signaling, *J Biol Chem 285*, 26852-26860.

96. Spear, W., Chan, D., Coppens, I., Johnson, R. S., Giaccia, A., and Blader, I. J. (2006) The host cell transcription factor hypoxia-inducible factor 1 is required for *Toxoplasma gondii* growth and survival at physiological oxygen levels, *Cell Microbiol 8*, 339-352.

97. Laliberte, J., and Carruthers, V. B. (2008) Host cell manipulation by the human pathogen *Toxoplasma gondii*, *Cell Mol Life Sci 65*, 1900-1915.

98. Huang, C. R., Lien, C. Y., Tsai, W. C., Lai, W. A., Hsu, C. W., Tsai, N. W., Chang, C. C., Lu, C. H., Chien, C. C., and Chang, W. N. (2018) The clinical characteristics of adult bacterial meningitis caused by non-*Pseudomonas* (Ps.) *aeruginosa Pseudomonas* species: A clinical comparison with Ps. aeruginosa meningitis, *Kaohsiung J Med Sci 34*, 49-55.

99. Sutter, V. L. (1968) Identification of *Pseudomonas* species isolated from hospital environment and human sources, *Appl Microbiol 16*, 1532-1538.

100. Scotti, J. S. (2014) STRUCTURAL, FUNCTIONAL AND EVOLUTIONARY STUDIES ON PROLYL-3-HYDROXYLASES University of Oxford.

101. Glasser, N. R., Wang, B. X., Hoy, J. A., and Newman, D. K. (2017) The Pyruvate and alpha-Ketoglutarate Dehydrogenase Complexes of Pseudomonas aeruginosa Catalyze Pyocyanin and Phenazine-1-carboxylic Acid Reduction via the Subunit Dihydrolipoamide Dehydrogenase, *J Biol Chem 292*, 5593-5607.

102. Dietrich, L. E., Price-Whelan, A., Petersen, A., Whiteley, M., and Newman, D. K. (2006) The phenazine pyocyanin is a terminal signalling factor in the quorum sensing network of Pseudomonas aeruginosa, *Mol Microbiol 61*, 1308-1321.

103. Dickinson, R. S., Murphy, F., Doherty, C., Williams, S., Mirchandani, A., Willson, J., Scotti, J. S., Preston, G., Schofield, C. J., Whyte, M. K. B., and Walmsley, S. R. (2017) *Pseudomonas* expression of an oxygen sensing prolyl hydroxylase homologue regulates neutrophil host responses in vitro and in vivo, *Wellcome Open Res 2*, 104.

104. Trenkle, P. (2016) Investigations on the Evolutionary Related PHD-Like Enzymes *Pseudomonas* PHD and *Monosiga brevicollis* P4H, Unpublished.

105. Clark, B. F., and Nyborg, J. (1997) The ternary complex of EF-Tu and its role in protein biosynthesis, *Curr Opin Struct Biol 7*, 110-116.

106. Andersen, G. R., Nissen, P., and Nyborg, J. (2003) Elongation factors in protein biosynthesis, *Trends Biochem Sci 28*, 434-441.

107. Nilsson, J., and Nissen, P. (2005) Elongation factors on the ribosome, *Curr Opin Struct Biol 15*, 349-354.

108. Berchtold, H., Reshetnikova, L., Reiser, C. O., Schirmer, N. K., Sprinzl, M., and Hilgenfeld, R. (1993) Crystal structure of active elongation factor Tu reveals major domain rearrangements, *Nature 365*, 126-132.

109. Kjeldgaard, M., Nissen, P., Thirup, S., and Nyborg, J. (1993) The crystal structure of elongation factor EF-Tu from Thermus aquaticus in the GTP conformation, *Structure 1*, 35-50.

110. King, N., Westbrook, M. J., Young, S. L., Kuo, A., Abedin, M., Chapman, J., Fairclough, S., Hellsten, U., Isogai, Y., Letunic, I., Marr, M., Pincus, D., Putnam, N., Rokas, A., Wright, K. J., Zuzow, R., Dirks, W., Good, M., Goodstein, D., Lemons, D., Li, W., Lyons, J. B., Morris, A., Nichols, S., Richter, D. J., Salamov, A., Sequencing, J. G., Bork, P., Lim, W. A., Manning, G., Miller, W. T., McGinnis, W., Shapiro, H., Tjian, R., Grigoriev, I. V., and Rokhsar, D. (2008) The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans, *Nature 451*, 783-788.

111. Lang, B. F., O'Kelly, C., Nerad, T., Gray, M. W., and Burger, G. (2002) The closest unicellular relatives of animals, *Curr Biol 12*, 1773-1778.

112. Ruiz-Trillo, I., Roger, A. J., Burger, G., Gray, M. W., and Lang, B. F. (2008) A phylogenomic investigation into the origin of metazoa, *Mol Biol Evol 25*, 664-672.

113. Boleininger, A. (2012) Evolution of PHDs as oxygen sensors: mechanistic and structural studies of the PHD of Trichoplax adhaerens, the simplest animal, and mechanistic studies of a PHD-like enzyme of the protist *Monosiga brevicollis*, D.Phil. Thesis, University of Oxford.

114. Xu, Y., Wang, Z. A., Green, R. S., and West, C. M. (2012) Role of the Skp1 prolyl-hydroxylation/glycosylation pathway in oxygen dependent submerged development of *Dictyostelium*, *BMC Dev Biol 12*, 31.

115. Karunaratne, C. V., Weldeghiorghis, T. K., West, C. M., and Taylor, C. M. (2014) Conformational changes associated with post-translational modifications of Pro(143) in Skp1 of *Dictyostelium*--a dipeptide model system, *J Am Chem Soc 136*, 15170-15175.

116. Siddiq, M. A., Hochberg, G. K., and Thornton, J. W. (2017) Evolution of protein specificity: insights from ancestral protein reconstruction, *Curr Opin Struct Biol 47*, 113-122.

117. Khersonsky, O., and Tawfik, D. S. (2010) Enzyme promiscuity: a mechanistic and evolutionary perspective, *Annu Rev Biochem 79*, 471-505.

118. Morgan, D. O. (2007) *The Cell Cycle: Principles of Control*, New Science Press.

119. Cardozo, T., and Pagano, M. (2004) The SCF ubiquitin ligase: insights into a molecular machine, *Nat Rev Mol Cell Biol 5*, 739-751.

120. Pickart, C. M. (2001) Mechanisms underlying ubiquitination, *Annu Rev Biochem 70*, 503-533.

121. 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., and Pavletich, N. P. (2002) Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex, *Nature 416*, 703-709.

122. Schulman, B. A., Carrano, A. C., Jeffrey, P. D., Bowen, Z., Kinnucan, E. R., Finnin, M. S., Elledge, S. J., Harper, J. W., Pagano, M., and Pavletich, N. P. (2000) Insights into SCF ubiquitin ligases from the structure of the Skp1-Skp2 complex, *Nature 408*, 381-386.

123. Jin, J., Cardozo, T., Lovering, R. C., Elledge, S. J., Pagano, M., and Harper, J. W. (2004) Systematic analysis and nomenclature of mammalian F-box proteins, *Genes Dev 18*, 2573-2580.

124. Hao, B., Zheng, N., Schulman, B. A., Wu, G., Miller, J. J., Pagano, M., and Pavletich, N. P. (2005) Structural basis of the Cks1-dependent recognition of p27(Kip1) by the SCF(Skp2) ubiquitin ligase, *Mol Cell 20*, 9-19.

2. Production of DdP4H and TgPhyA

2.1 Summary

The work described in this chapter concerns the optimization of methods for the production of recombinant forms of DdP4H and TgPhyA full-length proteins on a relatively large scale. The full-length proteins were used for all the assays described in later chapters. In the latter part of the chapter, procedures for the production of truncated recombinant enzymes aimed at providing material suitable for crystallography studies, are described.

2.2 Introduction

In order to study the structures of DdP4H and TgPhyA, developing methods to produce large amounts of highly purified recombinant enzymes from *Escherichia coli* (*E. coli*) were required. Molecular biological cloning and expression optimization procedures were performed to produce the desired expression constructs encoding for DdP4H and TgPhyA genes, and to optimize the expression conditions (Chapter 2.3 – Chapter 2.6). Full-length DdP4H and TgPhyA were then produced on a large scale and their purification strategies were optimized (Chapter 2.7). After acquiring large amounts of the purified enzymes, differential scanning fluorimetry (T_m Shift) assays were performed to investigate the metals and inhibitors that stabilize the enzymes for crystallisation (Chapter 2.8). Initial crystallisation trials using the optimized conditions did not produce protein crystals. Truncated versions of the enzymes [DdP4H (60-284) and TgPhyA (53-264)] were designed using the online sequence secondary structure and disorder predictions as a guide (Chapter 2.9). Cloning, expression trials, and purification optimization studies were then performed on the newly designed constructs, to finally produce the desired highly purified truncated recombinant enzymes in *E. coli* (Chapter 2.10 – Chapter 2.13).

2.3 Initial Expression Trials for DdP4H in pET-15 and TgPhyA in pET-15b

2.3.1 Expression Trials for DdP4H in pET-15

The original constructs of DdP4H in pET-15 [1] and TgPhyA in pET-15b [1] (sequences and plasmid maps can be found in Chapter 7) were kindly provided by Dr. Adam Hardy (Department of Chemistry, Chemistry Research Laboratory, University of Oxford, OX1 3TA). The expression optimization trials were done using these constructs in BL21-Gold (DE3) *E. coli* competent cells [2]. Variation in induction times, IPTG concentrations and induction temperatures were tested. The *E. coli* pellets were then sonicated. Samples of various induction conditions were analysed using SDS-PAGE (Figure 2.1). The theoretical mass of His tagged DdP4H is around 35.9 kDa (https://web.expasy.org/protparam/). Either one of the two bands in the SDS-PAGE gels, marked by the red arrows, could represent DdP4H.

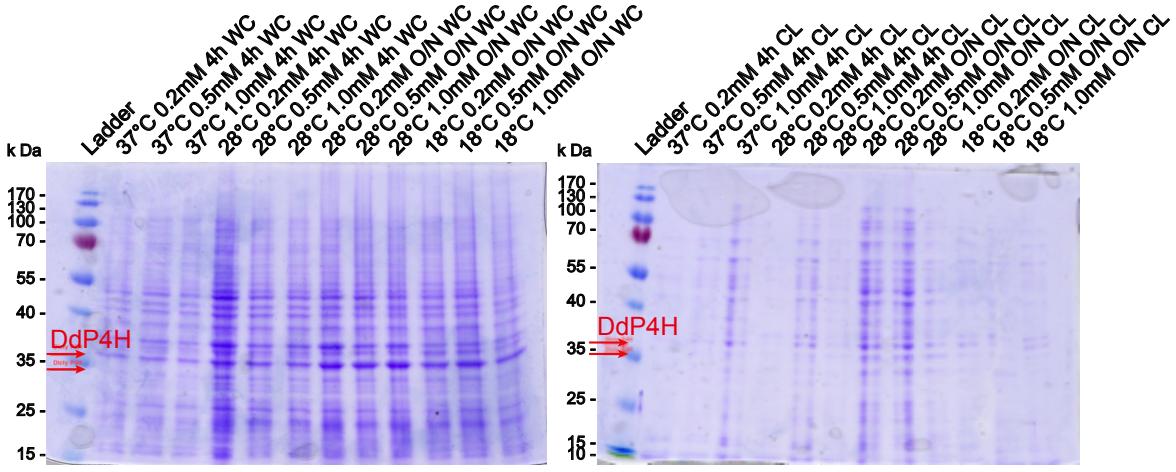


Fig. 2.1 SDS-PAGE analysis of expression optimization trials results for DdP4H in pET-15 with BL21-Gold (DE3) competent cells. Compared to whole cell (WC) gel (left), there was soluble DdP4H in the cell lysate (CL) after sonicating the cells. The expression conditions from left to right were: Ladder (PageRuler from Thermo Scientific); 37°C, 0.2 mM IPTG, 4 hours induction; 37 °C, 0.5 mM IPTG, 4 hours induction; 37 °C, 1 mM IPTG, 4 hours induction; 28 °C, 0.2 mM IPTG, 4 hours induction; 28 °C, 0.5 mM IPTG, 4 hours induction; 28 °C, 1 mM IPTG, 4 hours induction; 28 °C, 0.2 mM IPTG, overnight (O/N) induction; 28 °C, 0.5 mM IPTG, overnight induction; 28 °C, 1 mM IPTG, overnight induction; 18 °C, 0.2 mM IPTG, overnight induction; 18 °C, 0.5 mM IPTG, overnight induction; 18 °C, 1 mM IPTG, overnight induction. Either one of the red arrows indicates potential DdP4H protein.

The results indicate that the gene encoding DdP4H was overexpressed under some conditions using this construct. However, no intense protein bands were observed in the cell lysates, indicating the DdP4H may be insoluble under these conditions. An alternative method was thus needed to produce the recombinant enzyme.

2.3.2 Expression Trials for TgPhyA in pET-15b

Expression trials of the gene encoding TgPhyA were performed, screening different induction times, IPTG concentrations, and induction temperatures. The *E. coli* pellets were then sonicated and analysed using SDS-PAGE electrophoresis (Figure 2.2).

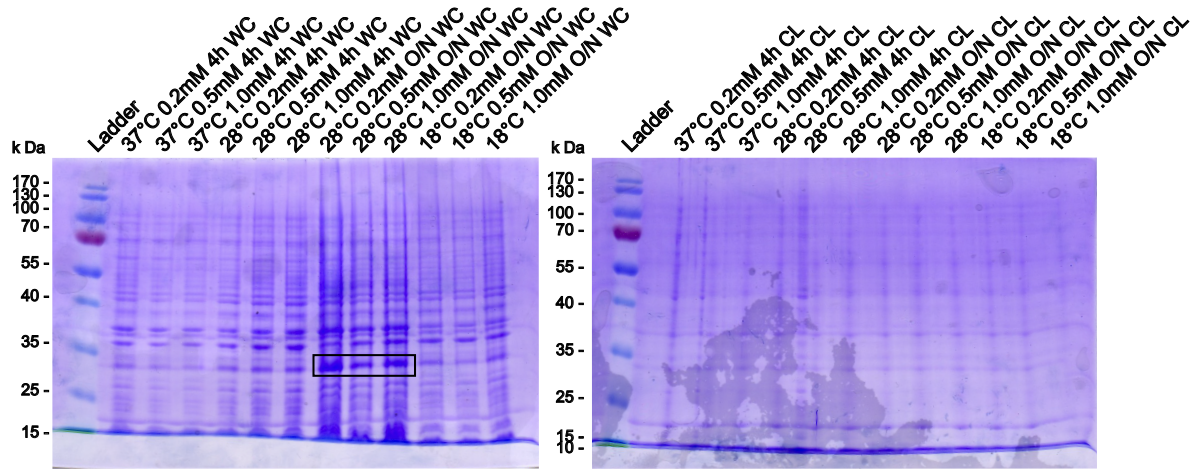


Fig. 2.2 SDS-PAGE analysis of expression optimization trials results for TgPhyA in pET-15b with BL21-Gold (DE3) competent cells. Compared to the whole cell (WC) gel (left), there was little TgPhyA in the cell lysate (CL) after sonication of the cells. The expression conditions from left to right were: 37 °C, 0.2 mM IPTG, 4 hours induction; 37 °C, 0.5 mM IPTG, 4 hours induction; 37 °C, 1 mM IPTG, 4 hours induction; 28 °C, 0.2 mM IPTG, 4 hours induction; 28 °C, 0.5 mM IPTG, 4 hours induction; 28 °C, 1 mM IPTG, 4 hours induction; 28 °C, 0.2 mM IPTG, overnight (O/N) induction; 28 °C, 0.5 mM IPTG, overnight induction; 28 °C, 1 mM IPTG, overnight induction; 18 °C, 0.2 mM IPTG, overnight induction; 18 °C, 0.5 mM IPTG, overnight induction; 18 °C, 1 mM IPTG, overnight induction. The intense bands in the box may indicate the overexpressed TgPhyA.

The results indicate that the TgPhyA gene might be overexpressed as indicated by the intense bands in the box, but the protein was likely not soluble under these conditions. Other methods were thus needed to produce the recombinant enzyme in a soluble form.

2.4 Expression Optimization Trials for TgPhyA Using ArcticExpress (DE3) Competent Cells

The results of original expression trials for DdP4H and TgPhyA (Section 2.3) indicated that the genes were being well expressed (Figure 2.1 and 2.2), but the recombinant proteins were not apparent, at least substantially, in the cell lysate solutions. It was considered that the insolubility of the enzymes may be related to the high efficiency of expression, by not providing enough time for the protein to fold independently, leading to the formation of highly aggregated protein which is packed as inclusion bodies by *E. coli* [3-6]. Therefore, attempts were made to reduce the efficiency of expression.

A less rich media (LB broth) did not help to reduce the amount of insoluble proteins [7, 8]. A different competent cell line, ArcticExpress (DE3) *E. coli* competent cells [9], was then used to try to produce soluble TgPhyA protein. The construct of pET-15b encoding for TgPhyA was transformed into ArcticExpress (DE3) *E. coli* competent cells (Agilent #230192). Expression trials were then carried out. Different IPTG concentrations, and cell densities were tested, induced at 10 °C for 24 hours. The recombinant *E. coli* pellets were then sonicated and analysed using SDS-PAGE electrophoresis (Figure 2.3).

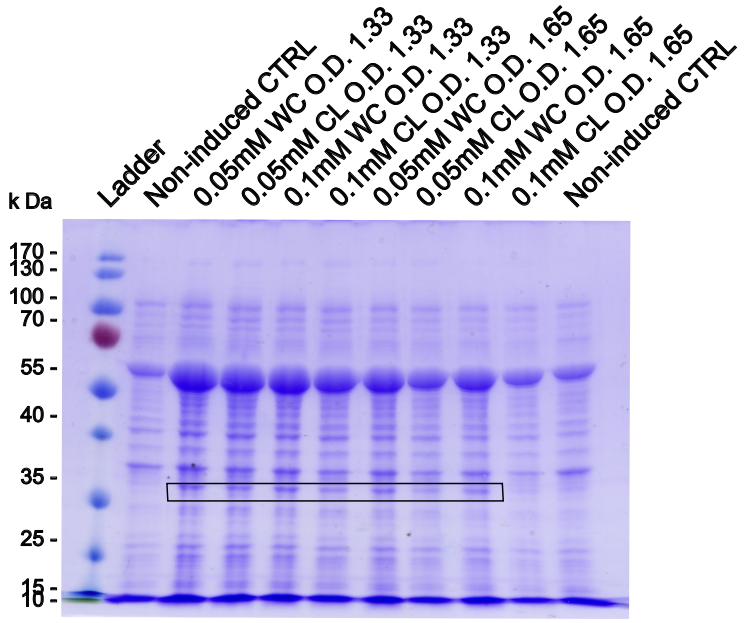


Fig. 2.3 SDS-PAGE analysis of expression optimization trials results for TgPhyA in pET-15b with ArcticExpress (DE3) competent cells. In both whole cell (WC) and cell lysate (CL) samples, soluble TgPhyA was present. The lanes from left to right were ladder; Non-induced control (CTRL); 0.05 mM, Optimal Density (O.D.) 1.33, whole cell (WC); 0.05 mM, O.D. 1.33, cell lysate (CL); 0.1 mM, O.D. 1.33, WC; 0.1 mM, O.D. 1.33, CL; 0.05 mM, O.D. 1.56, WC; 0.05 mM, O.D. 1.56, CL; 0.1 mM, O.D. 1.56, WC; 0.1 mM, O.D. 1.56, CL; Non-induced CTRL. The bands of TgPhyA are indicated by a box.

The observation of intense bands at around 55 kDa in the SDS-PAGE gel indicates the overexpression of the gene encoding for a chaperone [9], but TgPhyA (around 35 kDa) was not apparently produced (at least at a high level) under either conditions as indicated by the box in Figure 2.3.

Overall, neither genes encoding for DdP4H or TgPhyA were over expressed in the trials using the original vectors. Other methods therefore had to be pursued for the production of the recombinant proteins. Ligation-independent cloning (LIC) was then performed to clone the genes encoding for DdP4H and TgPhyA into different vectors, i.e. pNIC-ZB [10, 11] and pNH-TrxT [11, 12]. These vectors provide comparatively heavy N-terminal tags attached to the target

recombinant DdP4H and TgPhyA. Larger proteins may reduce the expression speed and may result in more soluble proteins.

2.5 Ligation-independent cloning (LIC) [13, 14] of DdP4H and TgPhyA

The primers designed specifically for DdP4H for LIC were:

Forward, 5' – TACTT-CCAATCCATGGATATTTCTAACTTACCC – 3';

Reverse, 5' – TATCCACCTTTACTGTTAATAAATCCAAGTT-GTAATTG – 3'.

The primers for TgPhyA were:

Forward, 5' – TACTTCCAATCCATGGCTCTTGA-GGACCTG – 3';

Reverse, 5' – TATCCACCTTTACTGTCACGATCCCGATTGACC – 3'.

The primers were designed with the following 5' extensions: TACTTCCAATCCATG for the forward primer and TATCCACCTTTACTG for the reverse primer. A stop codon (TTA for DdP4H and TCA for TgPhyA) was added after the sequence TATCCACCTTTACTG for the reverse primer depending on the sequence of DdP4H or TgPhyA. The DNA was amplified by polymerase chain reaction (PCR) using a thermal cycler using the following conditions:

A

PCR Reagent	Volumn Per Tube / µL
5x Q5 buffer (NEB #B9027S)	13.5
$MgSO_4$ (50 mM, Bioline BIO-37026)	5
dNTP (NEB #N0447S)	0.5
Q5 polymerase (NEB #M0491S)	0.75
Primer mix (10 µM)	0.5
Template DNA (5 ng/µL)	2.5
MilliQ water	2.5
Total	25

B

PCR Step	Temperature / °C	Incubation Time
1	95	1 min
2	94	30 s
3	52	30 s
4	68	6.5 min
5	Repeat Step 2 – 4 for 30 cycles	
6	68	5 min
7	15	Pause

Table 2.1 Procedures for LIC to clone the genes encoding for DdP4H and TgPhyA into different vectors. (A) Sample components for PCR. (B) PCR conditions.

The samples were then treated with DpnI (New England BioLabs #R0176S) for 1 hour at 37°C. The PCR products were purified and samples were analysed by agarose gel electrophoresis (Figure 2.4). The bands with the expected mass were detected.

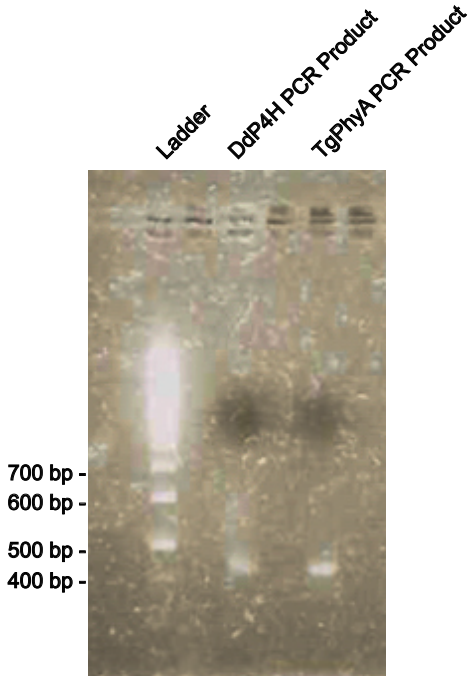


Fig. 2.4 Agarose Gel Electrophoresis results for DdP4H and TgPhyA PCR products.

The samples were digested by T4 polymerase to generate a cohesive end [14, 15].

T4 Digestion Reagents	Volumn Per Tube / µL
DNA	5
10 X T4 polymerase buffer (NEB #M0203S)	1
dCTP (25 mM NEB #N0447S)	1
DTT (100 mM)	0.5
BSA (10 mg/mL NEB B9001S)	0.1
T4 DNA polymerase (5 ng/µL NEB #M0203S)	0.25
MilliQ water	2.15
Total	10

Table 2.2 Sample components for T4 DNA polymerase digestion to generate cohesive ends.

The inserts were annealed into the pNIC-ZB [10, 14] and pNH-TrxT [12] vectors. The new constructs were transformed into XL10 Gold *E. coli* competent cells [16] and were further purified and stored for future use.

2.6 Expression Trials of DdP4H in pNIC-ZB [10, 14] and TgPhyA in pNIC-ZB

2.6.1 Expression Trials of DdP4H in pNIC-ZB

The expression trials for DdP4H were done with a new construct, pNIC-ZB, using BL21-Gold (DE3) *E. coli* competent cells. Different induction times, IPTG concentrations, and induction temperatures were tested. The recombinant *E. coli* pellets were then sonicated and were analysed using SDS-PAGE electrophoresis (Figure 2.5).

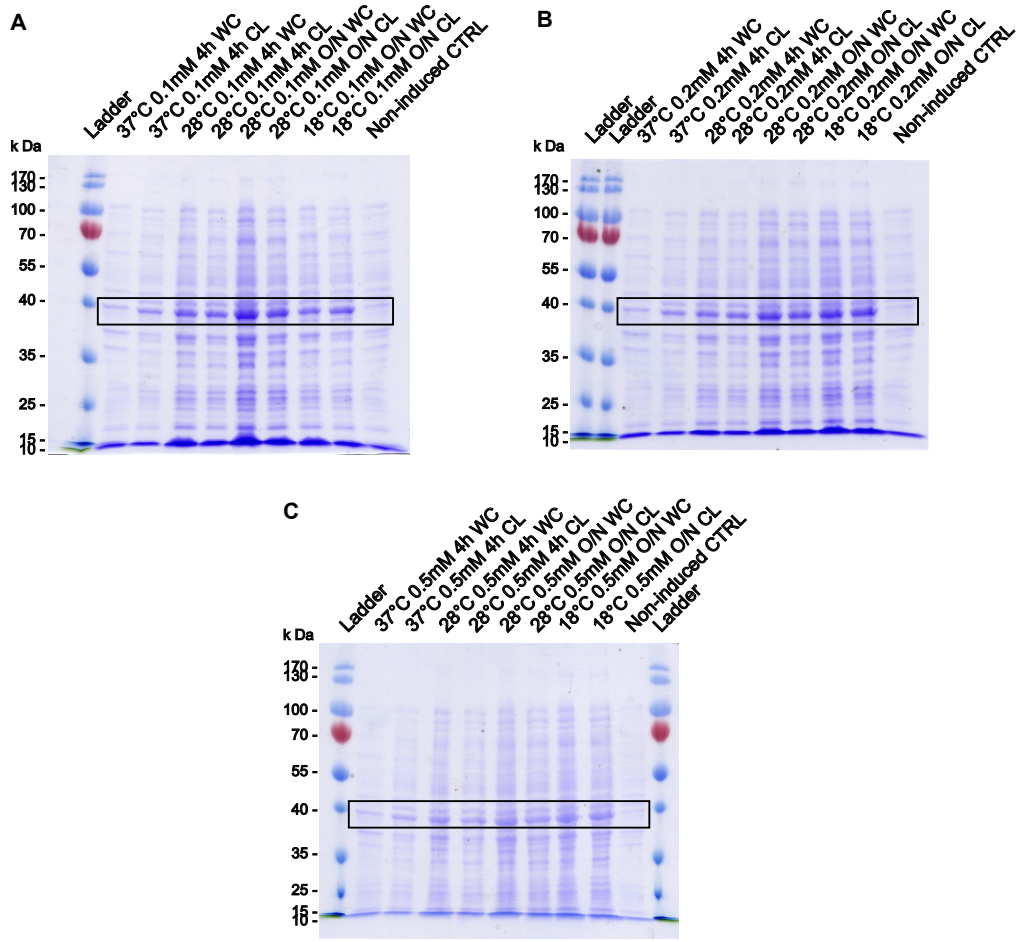


Fig. 2.5 SDS-PAGE analysis of expression optimization trials result for DdP4H in pNIC-ZB using BL21-Gold (DE3) competent cells. The lanes from left to right are: (A) ladder (PageRuler from Thermo Scientific); 0.1 mM IPTG, 4 hours induction at 37 °C whole cell (WC); 0.1 mM IPTG, 4 hours induction at 37 °C cell lysate (CL); 0.1 mM IPTG, 4 hours induction at 28 °C WC; 0.1 mM IPTG, 4 hours induction at 28 °C CL; 0.1 mM IPTG, overnight (O/N) induction at 28 °C WC; 0.1 mM IPTG, O/N induction at 28 °C CL; 0.1 mM IPTG, O/N induction at 18 °C WC; 0.1 mM IPTG, O/N induction at 18 °C CL; Non-induced control (CTRL). (B) 0.2 mM IPTG induction. (C) 0.5 mM IPTG induction. The bands of His_6-ZB-DdP4H are indicated by boxes.

SDS-PAGE gels indicate a high level of expression of the gene encoding for DdP4H, importantly with a high level of soluble DdP4H recombinant protein in cell lysates. The most intense DdP4H band was apparent when using 0.2 mM IPTG induced at 18°C overnight growth. This condition was therefore selected to express DdP4H in further protein production and purification.

2.6.2 Expression Trials of TgPhyA in pNIC-ZB [10, 14]

The expression trials of the gene encoding for TgPhyA were done with the construct, pNIC-ZB, in BL21-Gold (DE3) *E. coli* competent cells as for expression of the DdP4H gene. Pellets were lysed and analysed using SDS-PAGE electrophoresis (Figure 2.6).

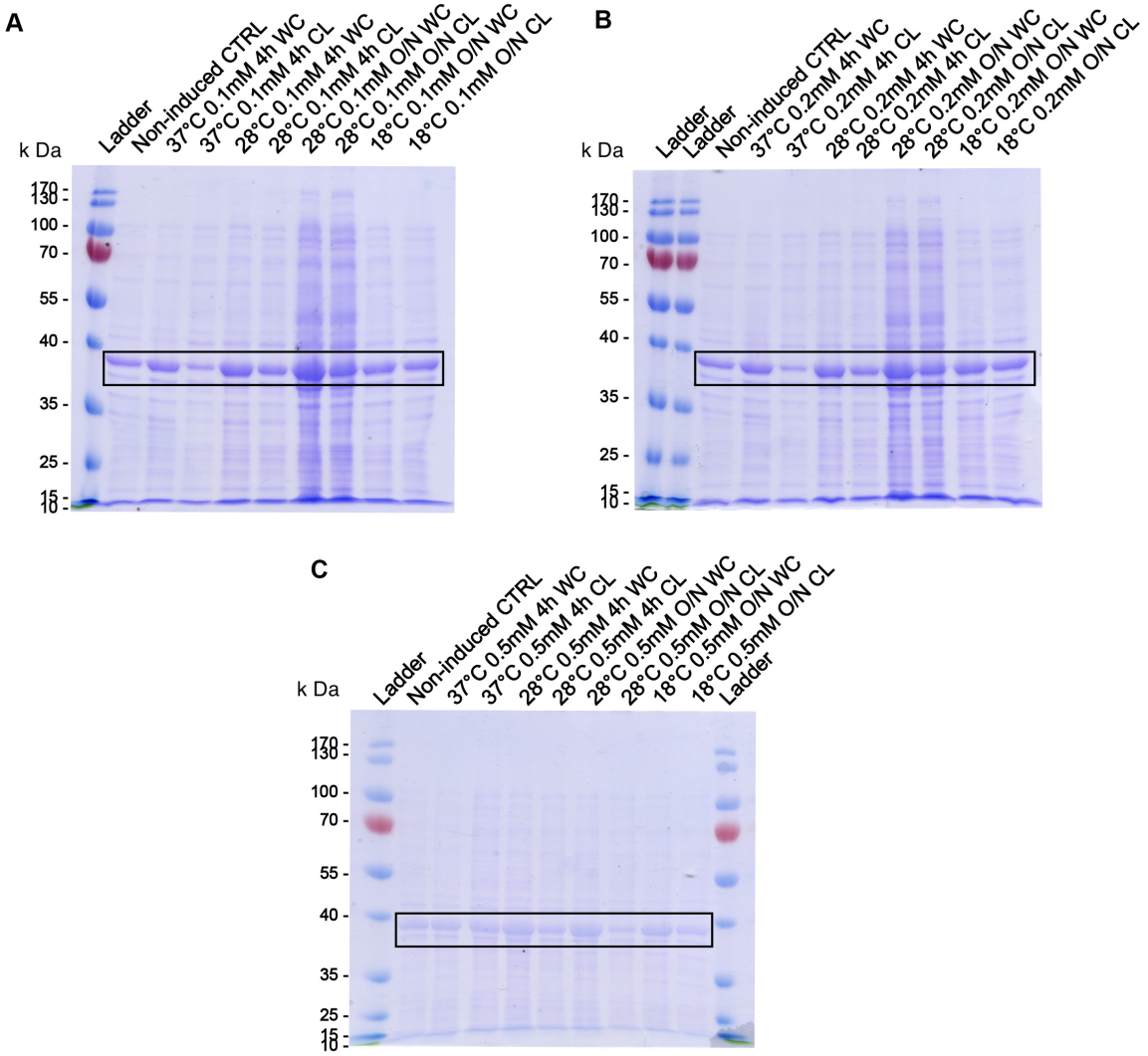


Fig. 2.6 SDS-PAGE analysis of expression optimization trials result for the TgPhyA gene in pNIC-ZB using BL21-Gold (DE3) competent cells. The lanes from left to right are: (A) ladder (PageRuler from Thermo Scientific); Non-induced control (CTRL); 0.1 mM IPTG, 4 hours induction at 37 °C whole cell (WC); 0.1 mM IPTG, 4 hours induction at 37 °C cell lysate (CL); 0.1 mM IPTG, 4 hours induction at 28 °C WC; 0.1 mM IPTG, 4 hours induction at 28 °C CL; 0.1 mM IPTG, overnight (O/N) induction at 28 °C WC; 0.1 mM IPTG, O/N induction at 28 °C CL; 0.1 mM IPTG, O/N induction at 18 °C WC; 0.1 mM IPTG, O/N induction at 18 °C CL. (B) 0.2 mM IPTG induction. (C) 0.5 mM IPTG induction. The bands of His_6-ZB-TgPhyA are indicated by boxes.

SDS-PAGE analysis showed a high level of expression of the gene encoding for TgPhyA and a high level production of soluble recombinant TgPhyA protein in the cell lysates. The most intense band of TgPhyA was apparent using 0.1 mM IPTG induced at 28 °C with growth overnight. This condition was therefore selected as the optimal condition to produce DdP4H for further protein production and purification work.

2.7 Expression and Purification of DdP4H and TgPhyA

2.7.1 Expression and Purification of DdP4H

Expression optimisation trials for the DdP4H gene demonstrated that the gene encoding for His_6-ZB-DdP4H was efficiently expressed in BL21 Gold (DE3) *E. coli* cells, when induced with 0.2 mM IPTG (overnight, 18 °C). DdP4H was then produced on a large scale using these conditions. BL21 Gold (DE3) cells, containing DdP4H plasmid was incubated until optimal density (O.D.) reached 1.2. Cells were then induced by adding IPTG and were incubated overnight at 18 °C. A 60 g wet cell mass was harvested. His_6-ZB-DdP4H was purified using Ni(II) affinity chromatography. The enzyme was then treated with the tobacco etch virus (TEV) protease and EDTA (50 mM) overnight and re-purified by Ni(II) affinity chromatography to remove the *N*-terminal tag. The flow-through was purified using size exclusion chromatography. The final untagged *apo*-DdP4H (20 mg) was produced in 95% purity (as judged by SDS-PAGE analysis, Figure 2.7 E). This protein was then used in broad screen crystallisation trials and hydroxylation assays.

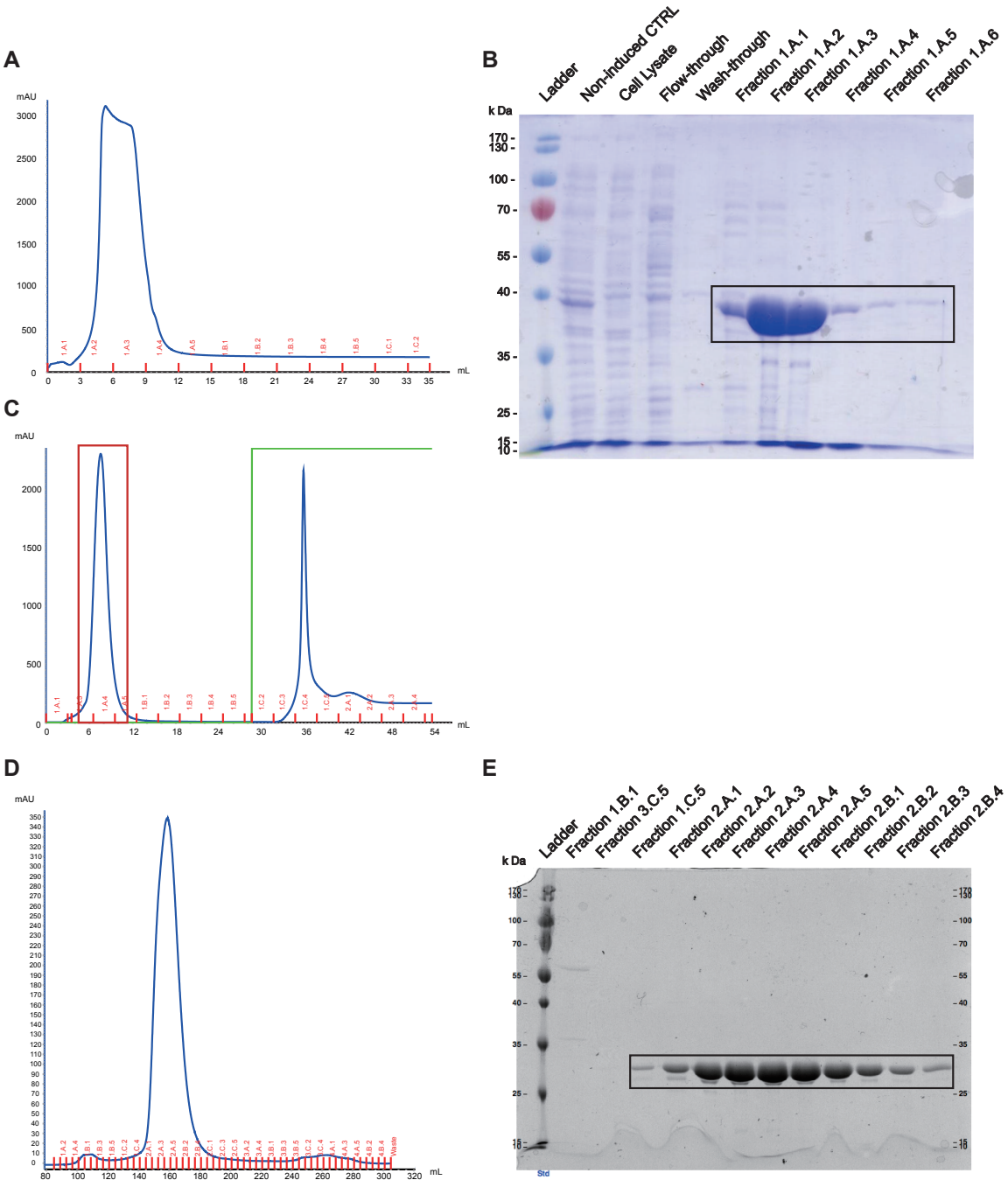


Fig. 2.7 Purification of DdP4H. (A) Ni(II) affinity chromatography UV chromatogram elution file. (B) SDS-PAGE analysis. The lanes from left to right are: Ladder (PageRuler from Thermo Scientific); non-induced control (CTRL); cell lysate; flow-through; wash-through; fractions 1.A.1 – 1.A.6. The bands corresponding to His_6-ZB-DdP4H are indicated by a box. (C) Ni(II) affinity chromatography UV chromatogram after TEV and EDTA treatment overnight. The flow-through (0 – 28 mL) and eluted volume (28 – 54 mL) were all collected in fractions. The green line showed the transition from 100% binding buffer (5 mM imidazole, 0.5 M NaCl, 50 mM HEPES, pH 7.6) to 100% elution buffer (500 mM imidazole, 0.5 M NaCl, 50 mM HEPES, pH 7.6). Fractions indicated by a red box were collected and were further purified using S75 size exclusion chromatography. The final storage buffer was 50 mM HEPES, pH 7.6. (D) Size exclusion UV

chromatogram. (E) SDS-PAGE analysis. The bands corresponding to untagged *apo*-DdP4H are indicated by a box.

2.7.2 Expression and Purification of TgPhyA

The tag-cleaved *apo*-TgPhyA (30mg) was expressed and purified using the same method as for the purification of *apo*-DdP4H. This protein was then used in broad screen crystallisation trials and hydroxylation assays.

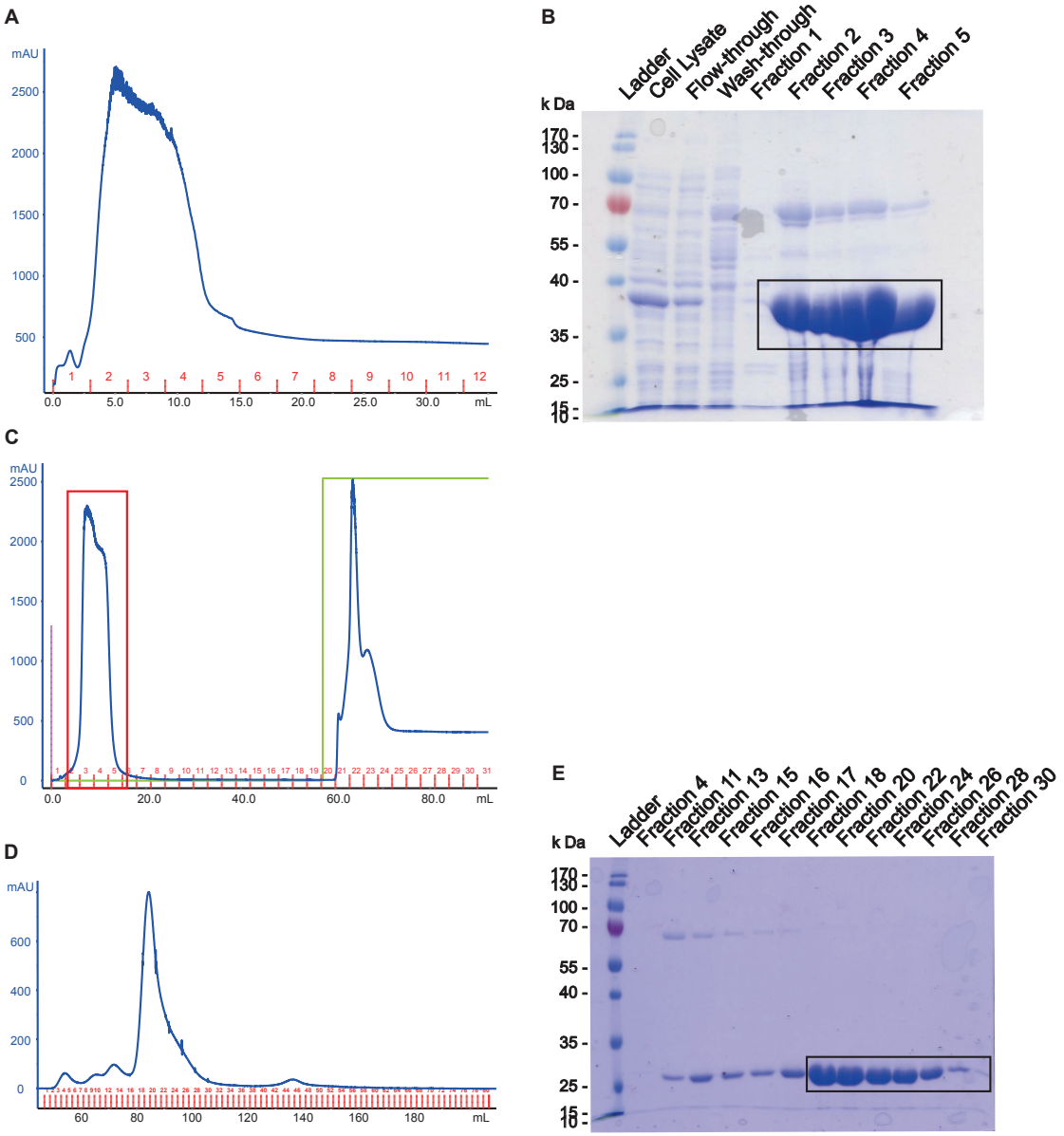


Fig. 2.8 Purification of TgPhyA. (A) Ni(II) affinity chromatography UV chromatogram elution profile. (B) SDS-PAGE gel analysis. The lanes from left to right are: Ladder (PageRuler from Thermo Scientific); cell lysate; flow-through; wash-through; fractions 1 - 5. The bands

correspondent to His_6-ZB-TgPhyA are indicated by a box. (C) Ni(II) affinity chromatography UV chromatogram after TEV and EDTA treatment overnight. The flow-through (0 – 58 mL) and eluted volume (58 – 94 mL) were all collected in fractions. The green line showed the transition from 100% binding buffer (5 mM imidazole, 0.5 M NaCl, 50 mM HEPES, pH 7.6) to 100% elution buffer (500 mM imidazole, 0.5 M NaCl, 50 mM HEPES, pH 7.6). Fractions indicated by a red box were collected and were further purified using S75 size exclusion chromatography. The final storage buffer was 50 mM HEPES, pH 7.6. (D) Size exclusion UV chromatogram. (E) SDS-PAGE analysis. The bands corresponding to the untagged *apo*-DdP4H are indicated by a box.

2.8 Differential Scanning Fluorimetry (T_m Shift) Assays [17, 18]

Based on sequence comparison and structural models of DdP4H and TgPhyA, these two enzymes show substantial similarity with other 2OG dependent prolyl-hydroxylases (Figure 2.11). Hydroxylation assays (described in Chapter 3) showed their requirement for Fe(II) and 2OG to be present for substrate hydroxylation to occur. Therefore, it is reasonable to assume that the presence of a suitable metal ion and cofactor will stabilize the enzymes in a manner beneficial for crystallisation, as observed for other prolyl hydroxylases.

Differential Scanning Fluorimetry (T_m shift) assays were used to investigate the optimal conditions of buffer, pH, metal, and cofactors that stabilise the protein. With increases in temperature, most proteins gradually unfold. The hydrophobic parts of proteins are revealed upon unfolding. A fluorescent dye (5,000 x SYPRO Orange), which increases fluorescence emission in a non-polar environment, can bind to the exposed hydrophobic parts of unfolded protein and emit fluorescence [17]. The observed fluorescence can increase sigmoidally with the increase of temperature. The mid point of the sigmoidal denaturation curve, which denotes 50% unfolded protein, is defined as the T_m [18]. Different buffer conditions or ligands can stabilize or destabilize the protein and change the apparent T_m of the protein. T_m shift assays have also been used in the analysis of the selective inhibitors towards the 2OG dependent Jumonji C (JmjC) histone demethylases and PHD2 [17, 19].

To investigate the optimal conditions for stabilizing DdP4H and TgPhyA, T_m shift assays were performed.

Different metals (Mn^{2+}, Fe^{2+}, Zn^{2+}, Mg^{2+}, Ni^{2+}, Co^{2+}, and Cd^{2+}) and cofactors 2OG or the analogue N-oxalylglycine (NOG) were screened in the T_m shift assays for both DdP4H and TgPhyA using a CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad).

The purified enzymes were mixed with metal chloride salt and cofactors. Non-metal or non-cofactor conditions were tested as negative controls. The SYPRO Orange dye was added last. The mixed solutions were incubated at RT for 10 minutes before measurements were taken. The fluorescence readings were taken every 1 °C increment from 25 – 80 °C. Each experiment was tested in triplicate for an accurate T_m. The florescence curves were fitted to Boltzmann curves using GraphPad Prism 5. The T_m were determined as EC_{50} between the minimum and maximum of fluorescence intensity.

Assay Reagent	Final Conc. / µM	Stock Conc. / µM	Volumn Per Assay / µL
2OG or NOG	50	500	5
Enzyme	4	20	10
Metal	50	1000	2.5
Buffer	1 µL SYPRO Orange to 2.5 µL Buffer		32.5
Total	-		50

Table 2.3 Assay conditions for the T_m shift assays of the enzyme, DdP4H and TgPhyA. Buffer condition was 100 mM HEPES, pH 7.6.

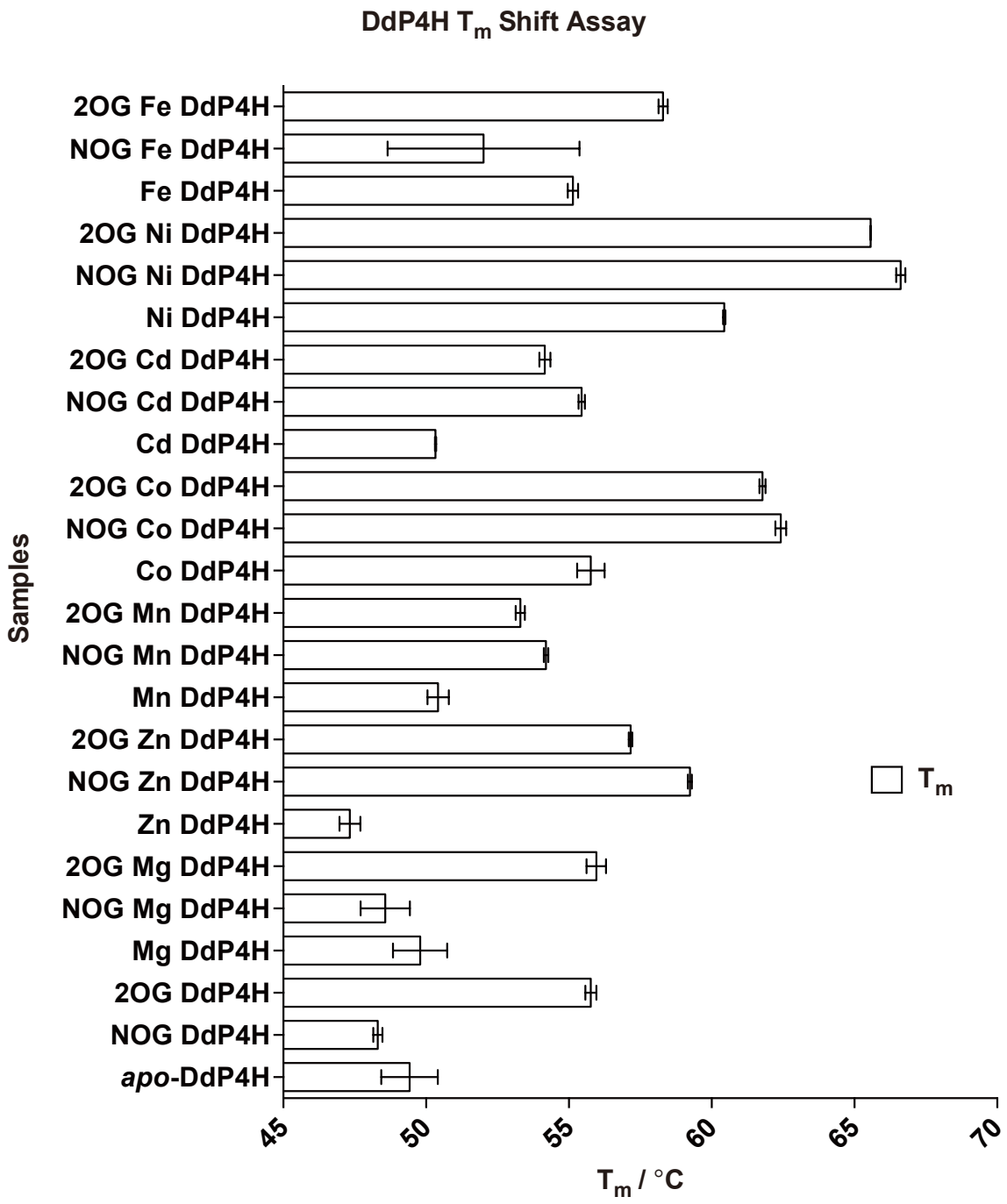


Fig. 2.9 T_m shift assays on DdP4H. Results shown are the mean ± s.d. of T_m in triplicate. An *apo*-DdP4H was used as a reference sample.

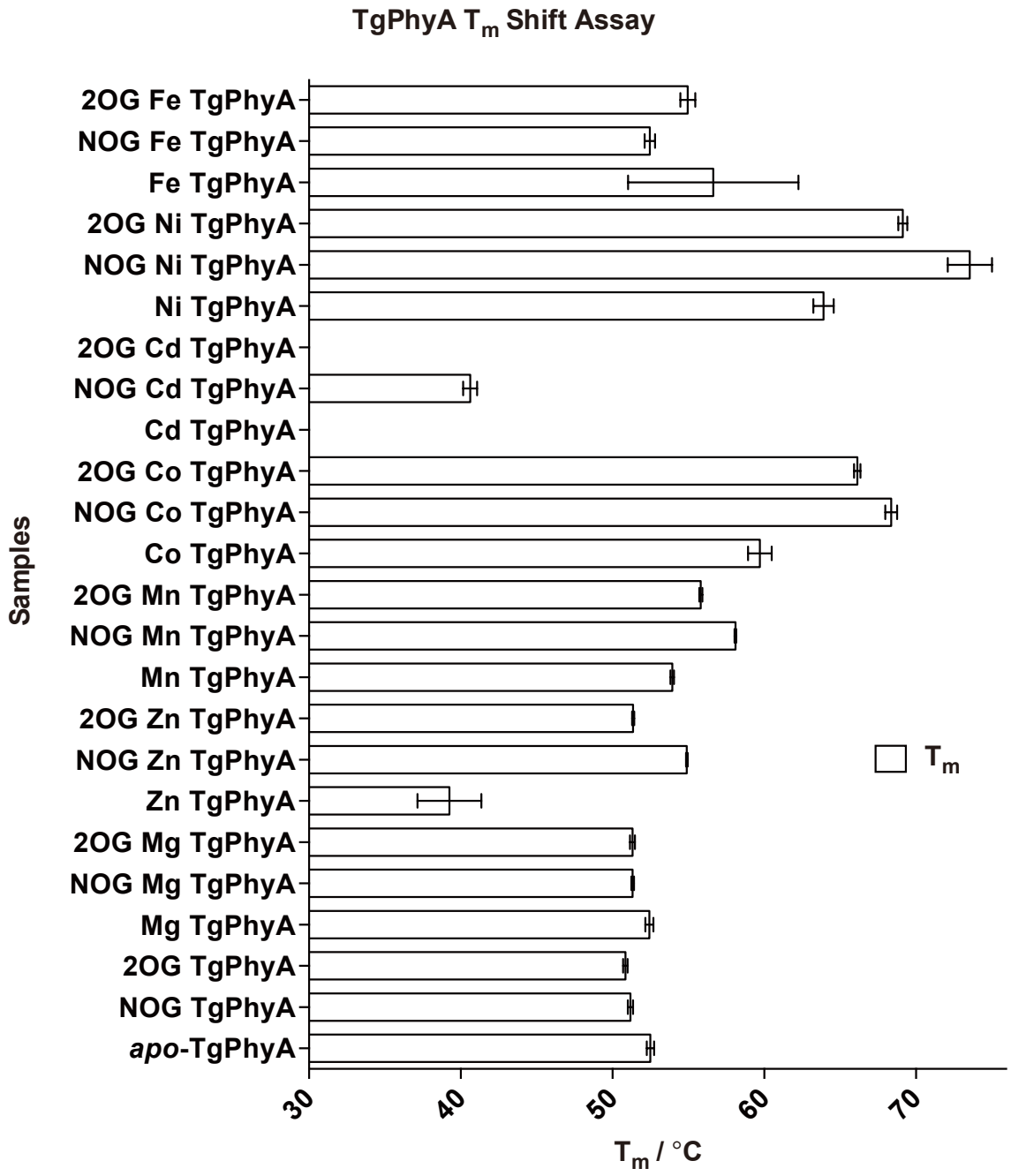


Fig. 2.10 T_m shift assays on and TgPhyA. Results shown are the mean ± s.d. of T_m in triplicate. An *apo*-TgPhyA was used as a reference sample.

For both DdP4H and TgPhyA, the addition of NOG and $NiCl_2$ increased the T_m most among the tested compounds, indicating they stabilised the enzymes. Crystallisation trials of the purified DdP4H and TgPhyA were thus pursued using NOG and $NiCl_2$.

2.9 Truncated Constructs, DdP4H (60-284) and TgPhyA (53-264)

Initial crystallisation trials were carried out with *apo*-DdP4H and *apo*-TgPhyA in the presence of stabilizing components identified in the T_m shift assays (2 mM $NiCl_2$, 2 mM NOG, 100 mM HEPES, pH 7.6). Commercial sparse matrix crystallisation screens were used in trials to crystallise full-length DdP4H and TgPhyA. The commercial screens included INDEX (Hampton Research), Wizard (Molecular Dimensions), JCSG+ (Molecular Dimensions), PACT Premier (Molecular Dimensions), SaltRx (Hampton Research) and PEG/Ion (Hampton Research), and MIDAS (Molecular Dimensions).

Crystallisation trials were performed using the sitting drop vapour diffusion method. The trials were set up in a 93-3 low profile INTELLI-PLATE (Art Robbins Instruments, Catalog #: 102-0001-13) using a Phoenix RE liquid dispensing robot (Art Robbins Instruments). For each crystal condition, 3 sub-conditions were set up with the condition of Sub Well 1: 100 nL / 200 nL, Sub Well 2: 100 nL / 100 nL, Sub Well 3: 200 nL / 100 nL. The volume of the reservoir solution was 80 µL. The dispensed plates were sealed with StarSeal Advanced Polyolefin Film (STARLAB, E2796-9795) and were incubated at 20 °C and 4 °C.

No protein crystals were observed after one month's incubation. Alternative ways were therefore explored to obtain crystal structures of the enzymes [20, 21]. Disorder and secondary structure predictions of the enzymes were then carried out aiming to find truncations more amenable to crystallisation.

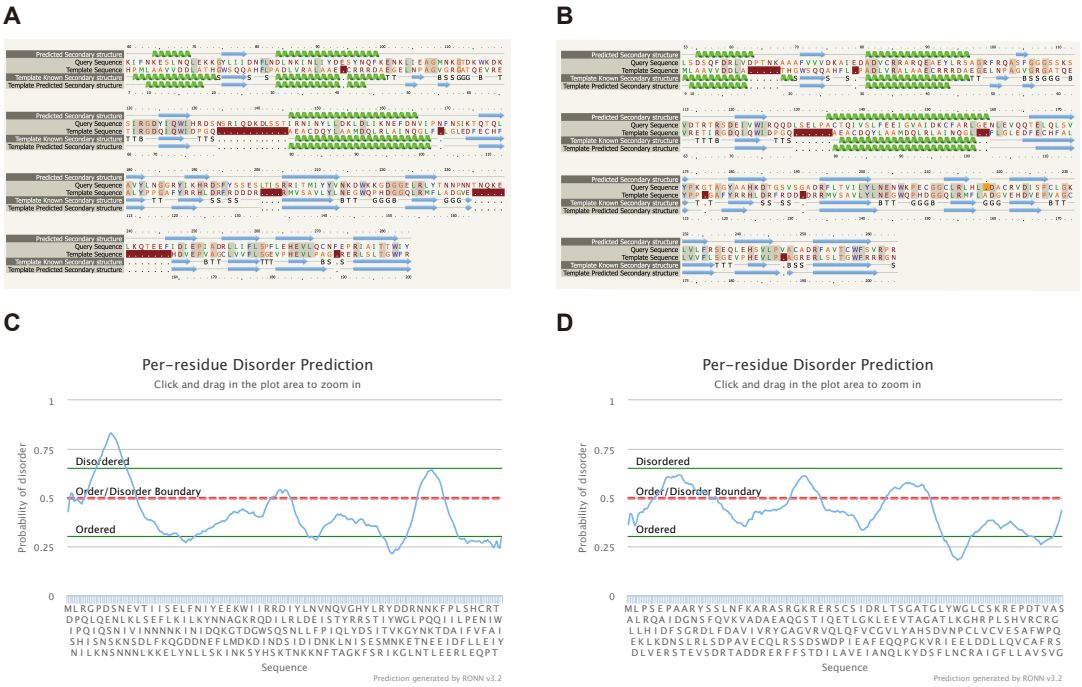


Fig. 2.11 Secondary structure prediction and disorder predictions of DdP4H and TgPhyA. Secondary structure prediction using Phyre2 modelled DdP4H (A) and TgPhyA (B) with (apparent) 100 % confidence by the highest scoring template, PPHD. Confidence scores represent the probability (0 – 100%) that the match between the sequence and the template is a true homology. They do not represent the expected accuracy of the models. Disorder predictions by RONN predicted the disordered sequence at the N-termini of DdP4H (C) and TgPhyA (D).

Disorder prediction analysis (http://www.strubi.ox.ac.uk/RONN) [22] indicated the N-terminal regions of DdP4H and TgPhyA were both disordered. Secondary structure predictions based on the structure of *Pseudomonas putida* prolyl-4-hydroxylase (http://www.sbg.bio.ic.ac.uk/~phyre2) [23] also indicated the secondary structure of the N-terminal regions of DdP4H and TgPhyA are hard to predict. It was thus proposed that the apparently disordered N-terminal regions might interfere with the crystallisation process. Therefore, truncated DdP4H and TgPhyA constructs based on this analysis were made using LIC.

2.10 LIC on DdP4H, DdP4H (60-284), and TgPhyA, TgPhyA (53-264)

DNA encoding for a truncated form of DdP4H, DdP4H (60-284), and a truncated form of TgPhyA, TgPhyA (53-264), were cloned into pNIC-ZB and pNIC28-Bsa4 vectors, respectively, using LIC.

The primers designed for DdP4H (60-284) for LIC were:

Forward, 5' – TACTTCCAATCCATGAAAATCTTCAACAAAGAATC – 3';

Reverse, 5' – TATCCACCTTTACTGTTAAT-AAATCCAAGTTGTAATTG – 3'.

The primers for TgPhyA (53-264) were:

Forward, 5' – TACTT-CCAATCCATGCTTAGCGATTCGCAATTTG – 3';

Reverse, 5' – TATCCACCTTTACTGTCAACGAGGGC-GAACGGAG – 3'.

The primers had the following 5' extensions: TACTTCCAATCCATG for the forward primer and TATCCACCTTTACTG for the reverse primer. A stop codon (TTA for DdP4H and CAA for TgPhyA) was added after the sequence TATCCACCTTTACTG for the reverse primer depending on the sequence of DdP4H (60-284) or TgPhyA (53-264).

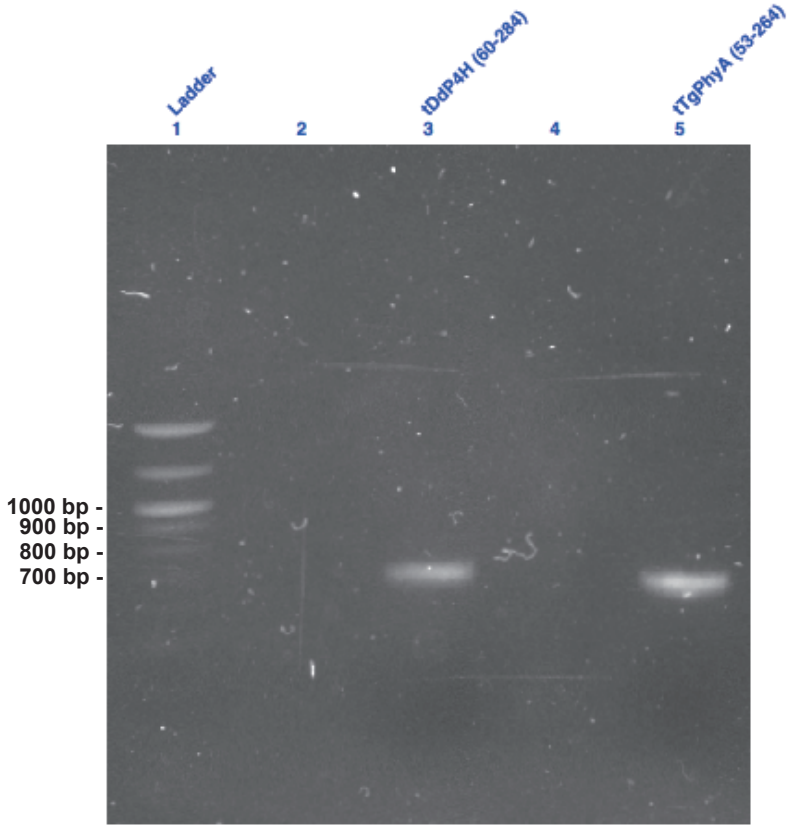


Fig. 2.12 Agarose Gel Electrophoresis results of DdP4H (60-284) and TgPhyA (53-264) PCR product.
The inserts were treated with T4 Polymerase and were annealed into the pNIC-ZB and pNIC28-Bsa4 vectors, respectively. The new constructs were transformed into XL10 Gold competent cells and were purified and stored for future use.

2.11 Expression Optimization Trials of DdP4H, DdP4H (60-284), and TgPhyA, TgPhyA (53-264)

Expression optimization trials of DdP4H (60-284) were done using the same conditions described in Chapter 2.3. DdP4H (60-284) was best expressed using the pNIC-ZB vector with 0.1 mM IPTG induced overnight at 18 °C. The best condition for TgPhyA (53-264) expression was using the pNIC-28 Bsa4 plasmid with 0.1 mM IPTG induced overnight at 18 °C.

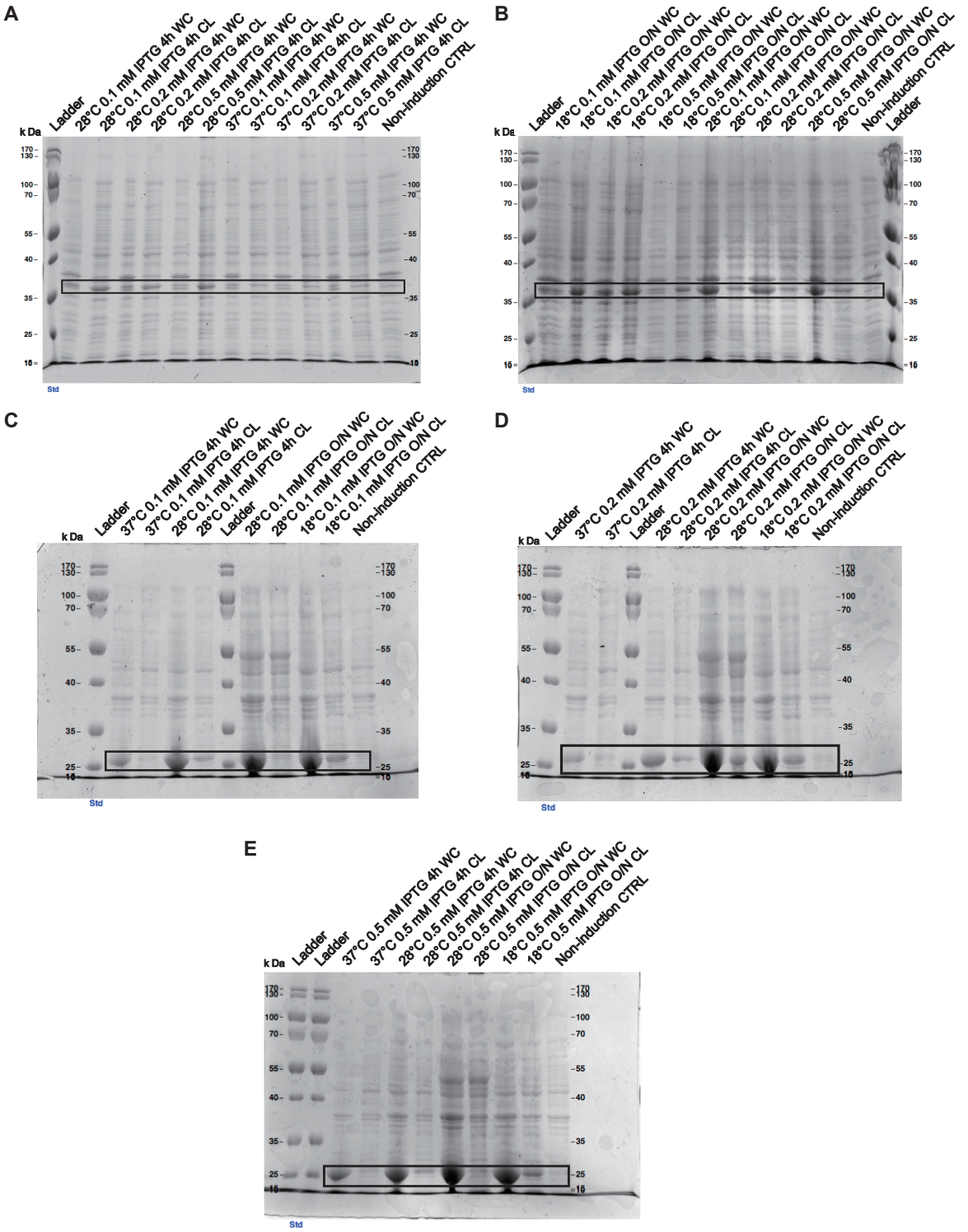


Fig. 2.13 SDS-PAGE analysis of expression optimization trials result for DdP4H (60-284) in pNIC-ZB with BL21-Gold (DE3) Competent Cells (A and B) and SDS-PAGE analysis of expression trials result for TgPhyA (53-264) in pNIC-28 Bsa4 with BL21-Gold (DE3) competent cells (C, D and E). The conditions were labelled above each lane. The bands corresponding to the target proteins are indicated by boxes.

2.12 Producing and purifying DdP4H, DdP4H (60-284)

His_6-ZB-DdP4H (60-284) was produced on a large scale (45 g wet cell mass) and was purified using Ni(II) affinity chromatography. The enzyme was then treated with the tobacco etch virus (TEV) protease (1mg / 50 mg target protein) and EDTA (50 mM) overnight at 4 °C and re-purified by Ni(II) affinity chromatography to remove the *N*-terminal tag. The flow-through was purified using size exclusion chromatography. The final untagged *apo*-DdP4H (60-284) (15 mg) produced was 99% pure (as judged by SDS-PAGE, Figure 2.14 D). The hydroxylation ability of truncated DdP4H (60-284) was not affected (Chapter 4.11.1). This protein was then used in broad screen crystallisation trials (Chapter 3).

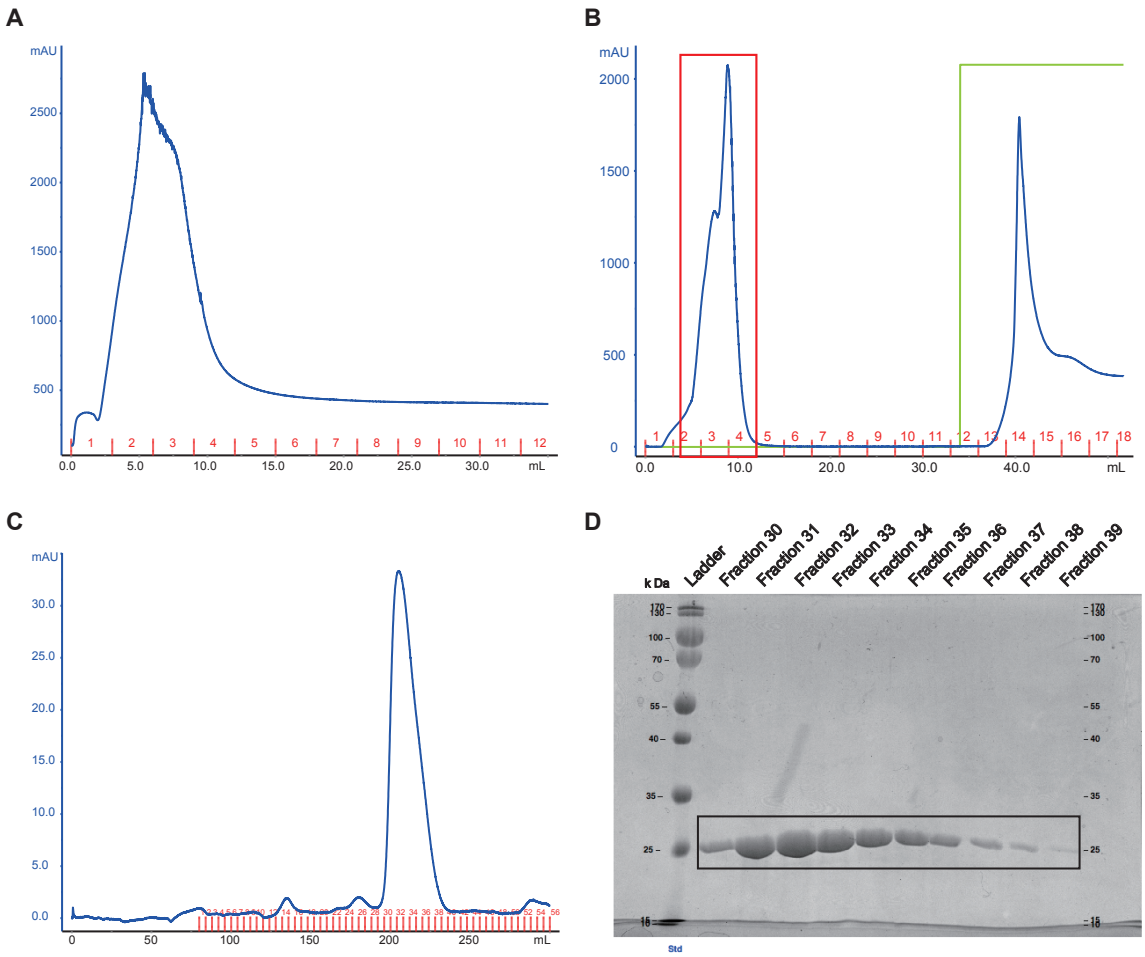


Fig. 2.14 Purification of DdP4H (60-284). (A) Ni(II) affinity chromatography UV chromatogram elution profile. (B) Ni(II) affinity chromatography UV chromatogram after TEV and EDTA

treatment overnight. Flow-through (0 – 35 mL) and eluted volume (35 – 55 mL) were all collected as fractions. The green line showed the transition from 100 % binding buffer (5 mM imidazole, 0.5 M NaCl, 50 mM HEPES, pH 7.6) to 100% elution buffer (500 mM imidazole, 0.5 M NaCl, 50 mM HEPES, pH 7.6). Fractions indicated by a red box were collected and purified using S75 size exclusion chromatography. The final storage buffer was 50 mM HEPES, 100 mM NaCl, pH 7.6. (C) Size exclusion UV chromatogram. (D) SDS-PAGE analysis. The bands corresponding to the untagged *apo*-DdP4H (60-284) are indicated by boxes.

2.13 Producing and Purifying TgPhyA (53-264)

His_6-ZB-TgPhyA (53-264) was produced on a large scale and purified using Ni(II) affinity chromatography. The enzyme was then treated with the tobacco etch virus (TEV) protease (1mg / 50mg target protein) and EDTA (50 mM) for 4 hours at room temperature and was re-purified by Ni(II) affinity chromatography to remove the *N*-terminal tag. The flow-through was purified using size exclusion chromatography. The final untagged *apo*-TgPhyA (53-264) (5 mg) produced was 99% pure (as judged by SDS-PAGE analysis, Figure 2.15 F). This protein was then used in broad screen crystallisation trials. (Chapter 3)

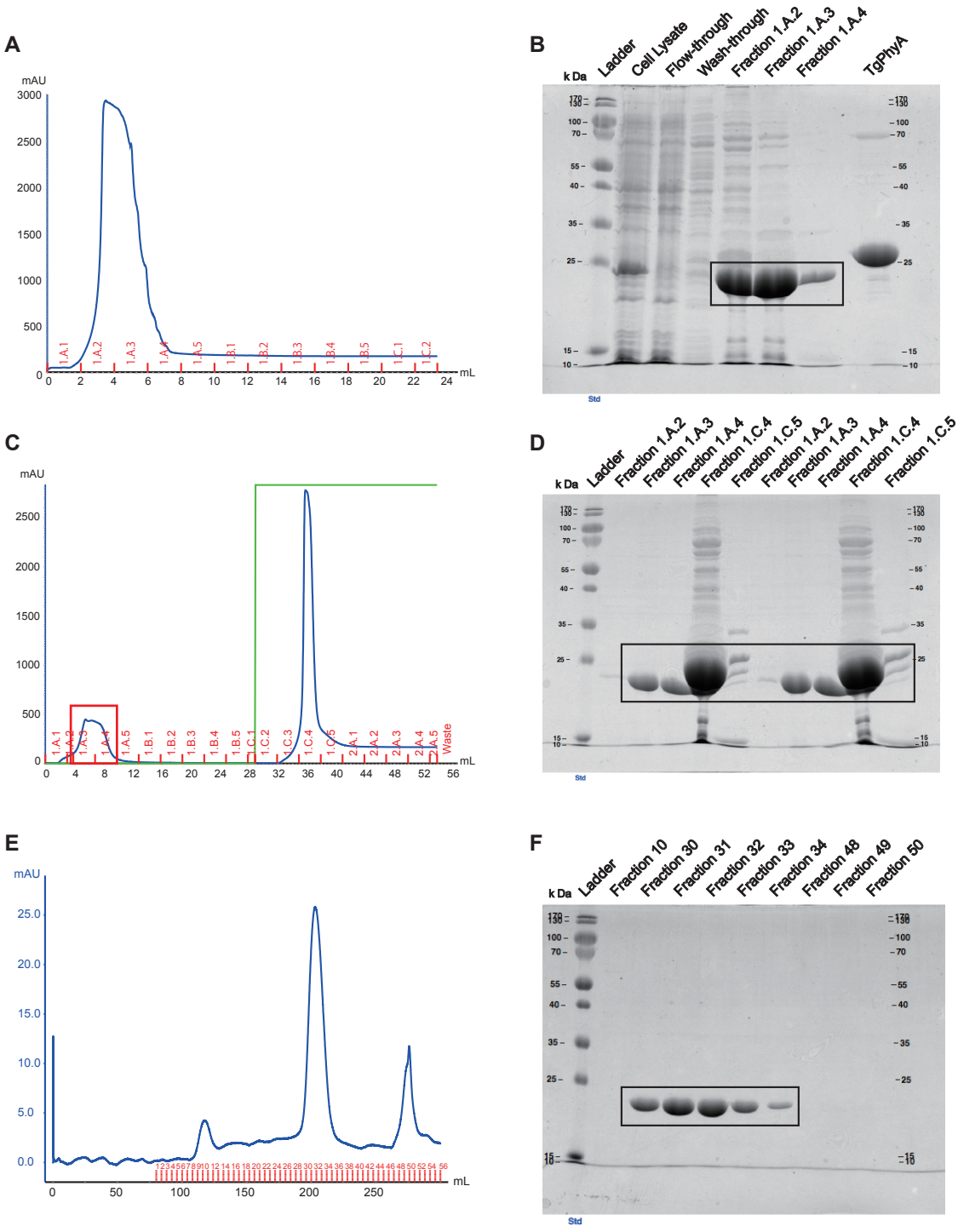


Fig. 2.15 Purification of TgPhyA (53-264). (A) Ni(II) affinity chromatography UV chromatogram elution profile. (B) SDS-PAGE analysis. The lanes from left to right are: Ladder (PageRuler from Thermo Scientific), cell lysate, flow-through, wash-through, fractions 1.A.2 – 1.A.4, untagged full-length TgPhyA. The bands corresponding to the His_6-ZB-TgPhyA (53-264) are indicated by a box. (C) Ni(II) affinity chromatography UV chromatogram after TEV and EDTA treated overnight. Flow-through (0 – 29 mL) and eluted volume (29 – 54 mL) were all collected in fractions. The green line showed the transition from 100 % binding buffer (5 mM imidazole, 0.5 M NaCl, 50 mM HEPES, pH 7.6) to 100 % elution buffer (500 mM imidazole, 0.5 M NaCl, 50 mM HEPES, pH

7.6). Fractions indicated by a red box were collected and were further purified using S75 size exclusion chromatography. (D) SDS-PAGE analysis. The lanes from left to right are: Ladder (PageRuler from Thermo Scientific), fractions 1.A.2 – 1.A.4, 1.C.4 – 1.C.5. The bands of untagged TgPhyA (53-264) are indicated by a box. (E) Size exclusion UV chromatogram. (D) SDS-PAGE analysis. The bands corresponding to the untagged *apo*-TgPhyA (53-264) are indicated by a box. The final storage buffer was 50 mM HEPES, 100 mM NaCl, pH 7.6.

2.14 References

1. pET-15b vector (Cat. No. 69661-3), (Novagen, Ed.), Novagen.

2. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

3. Singh, S. M., and Panda, A. K. (2005) Solubilization and refolding of bacterial inclusion body proteins, *J Biosci Bioeng 99*, 303-310.

4. Sambrook, J., and Russell, D. W. (2006) Purification of expressed proteins from inclusion bodies, *CSH Protoc 2006*.

5. Garcia-Fruitos, E., Vazquez, E., Diez-Gil, C., Corchero, J. L., Seras-Franzoso, J., Ratera, I., Veciana, J., and Villaverde, A. (2012) Bacterial inclusion bodies: making gold from waste, *Trends Biotechnol 30*, 65-70.

6. Rinas, U., Garcia-Fruitos, E., Corchero, J. L., Vazquez, E., Seras-Franzoso, J., and Villaverde, A. (2017) Bacterial Inclusion Bodies: Discovering Their Better Half, *Trends Biochem Sci 42*, 726-737.

7. de Groot, N. S., and Ventura, S. (2006) Effect of temperature on protein quality in bacterial inclusion bodies, *FEBS Lett 580*, 6471-6476.

8. Garcia-Fruitos, E. (2010) Inclusion bodies: a new concept, *Microb Cell Fact 9*, 80.

9. ArcticExpress Competent Cells and ArcticExpress (DE3) Competent Cells, (Agilent, Ed.).

10. pNIC-ZB (Plasmid #26107).

11. Savitsky, P., Bray, J., Cooper, C. D., Marsden, B. D., Mahajan, P., Burgess-Brown, N. A., and Gileadi, O. (2010) High-throughput production of human proteins for crystallization: the SGC experience, *J Struct Biol 172*, 3-13.

12. pNH-TrxT (Plasmid #26106).

13. Aslanidis, C., and de Jong, P. J. (1990) Ligation-independent cloning of PCR products (LIC-PCR), *Nucleic Acids Res 18*, 6069-6074.

14. Gileadi, O., Burgess-Brown, N. A., Colebrook, S. M., Berridge, G., Savitsky, P., Smee, C. E., Loppnau, P., Johansson, C., Salah, E., and Pantic, N. H. (2008) High throughput

production of recombinant human proteins for crystallography, *Methods Mol Biol 426*, 221-246.

15. Pheiffer, B. H., and Zimmerman, S. B. (1983) Polymer-stimulated ligation: enhanced blunt- or cohesive-end ligation of DNA or deoxyribooligonucleotides by T4 DNA ligase in polymer solutions, *Nucleic Acids Res 11*, 7853-7871.

16. XL10 Gold Ultracompetent Cells, (Agilent, Ed.).

17. Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability, *Nat Protoc 2*, 2212-2221.

18. Groftehauge, M. K., Hajizadeh, N. R., Swann, M. J., and Pohl, E. (2015) Protein-ligand interactions investigated by thermal shift assays (TSA) and dual polarization interferometry (DPI), *Acta Crystallogr D Biol Crystallogr 71*, 36-44.

19. Woon, E. C., Tumber, A., Kawamura, A., Hillringhaus, L., Ge, W., Rose, N. R., Ma, J. H., Chan, M. C., Walport, L. J., Che, K. H., Ng, S. S., Marsden, B. D., Oppermann, U., McDonough, M. A., and Schofield, C. J. (2012) Linking of 2-oxoglutarate and substrate binding sites enables potent and highly selective inhibition of JmjC histone demethylases, *Angew Chem Int Ed Engl 51*, 1631-1634.

20. Malawski, G. A., Hillig, R. C., Monteclaro, F., Eberspaecher, U., Schmitz, A. A., Crusius, K., Huber, M., Egner, U., Donner, P., and Muller-Tiemann, B. (2006) Identifying protein construct variants with increased crystallization propensity--a case study, *Protein Sci 15*, 2718-2728.

21. Mooij, W. T., Mitsiki, E., and Perrakis, A. (2009) ProteinCCD: enabling the design of protein truncation constructs for expression and crystallization experiments, *Nucleic Acids Res 37*, W402-405.

22. Yang, Z. R., Thomson, R., McNeil, P., and Esnouf, R. M. (2005) RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins, *Bioinformatics 21*, 3369-3376.

23. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis, *Nat Protoc 10*, 845-858.

3. Structural Studies of DdP4H (60-284)

3.1 Summary

This chapter describes X-ray crystallographic studies on a truncated form of DdP4H (residues 60-284). A crystal structure was obtained and analysed in comparison with crystal structures of other prolyl hydroxylases.

3.2 Introduction

After the establishment of optimal production and purification strategies for the truncated recombinant enzyme, DdP4H (60-284), crystallisation screening and subsequent optimisation were performed with the aim of obtaining diffraction quality crystals (Chapter 3.3). Crystals of DdP4H (60-284) in complex with Ni(II) and NOG were obtained and were flash cooled in cryoprotectant before placing in liquid nitrogen prior to diffraction data being collected at a synchrotron (Chapter 3.4). A crystal structure of DdP4H(60-284).Ni(II).NOG complex was obtained and analysed in detail (Chapter 3.5). Subsequently, the structure of DdP4H (60-284) was compared with structures of different prolyl hydroxylases, including *Psudomonas* PHD (PPHD) and human PHD2 (Chapter 3.6 – 3.8). At the end of this chapter, attempts to crystallise the DdP4H (60-284) and DdSkp1 complex using a disulfide-bond crosslinking protocol is described (Chapter 3.9 - 3.10).

3.3 Crystallisation Screening and Optimisation of DdP4H (60-284)

Preliminary screenings of DdP4H (60-284) for crystallisation were done with the kind help of Dr. Michael McDonough (Department of Chemistry, Chemistry Research Laboratory, University of Oxford, OX1 3TA).

After the successful production and purification of the truncated recombinant DdP4H (60-284), a large amount of pure *apo*-DdP4H (60-284) was purified for crystallisation trials. Crystallisation screening experiments on DdP4H (60-284) using multiple commercial crystallisation screens were carried out at both 4°C and 20°C, using the INDEX (Hampton Research), Wizard (Molecular Dimensions), JCSG+ (Molecular Dimensions), PACT Premier (Molecular Dimensions), SaltRx (Hampton Research) and PEG/Ion (Hampton Research) screens.

The *apo*-DdP4H (60-284) protein buffer solution contained the components in Table 3.1:

Assay Conditions	Concentration
DdP4H (60-284)	~ 23 mg/mL
$NiCl_2$	2 mM
NOG	2 mM
Buffer	50 mM HEPES, 100 mM NaCl, pH 7.6

Table 3.1 Ingredients and conditions of protein solutions for crystallisation.

Initial crystals were observed in the JCSG+ and PACT Premier screens. Many of these crystals had a hexagonal rod morphology (Figure 3.1):

Commercial Block	Well	Condition
JCSG+	G4.1	0.2 M TMAO, 0.1 M Tris, pH 8.5, 20% w/v PEG 2000 MME
JCSG+	G5.1	0.005 M cobalt(II) chloride hexahydrate, 0.005 M cadmium chloride hemi(pentahydrate), 0.005 M magnesium chloride hexahydrate, 0.005 M nickel(II) chloride hexahydrate, 0.1 M HEPES, pH 7.5, 12% w/v PEG 3350
JCSG+	G5.2	Same
JCSG+	G5.3	Same
Pact Premier	C7.1	0.2 M sodium chloride, 0.1 M HEPES, pH 7.0, 20% w/v PEG 6000
Pact Premier	C7.2	Same
Pact Premier	C7.3	Same
Pact Premier	C9.2	0.2 M lithium chloride, 0.1 M HEPES, pH 7.0, 20% w/v PEG 6000
Pact Premier	C12.1	0.01 M zinc chloride, 0.1 M HEPES, pH 7.0, 20% w/v PEG 6000
Pact Premier	D7.1	0.2 M sodium chloride, 0.1 M Tris, pH 8.0, 20% w/v PEG 6000
Pact Premier	D7.2	Same as above
Pact Premier	D7.3	Same as above
Pact Premier	D9.1	0.2 M lithium chloride, 0.1 M Tris, pH 8.0, 20 % w/v PEG 6000
Pact Premier	D9.2	Same as above
Pact Premier	D9.3	Same as above
Pact Premier	E1.1	0.2 M sodium fluoride, 20 % w/v PEG 3350
Pact Premier	E2.1	0.2 M sodium bromide, 20 % w/v PEG 3351

Pact Premier	E3.2	0.2 M sodium iodide, 20 % w/v PEG 3352
Pact Premier	E3.3	Same as above

Table 3.2 List of DdP4H (60-284) crystals obtained in the initial crystallisation screens.

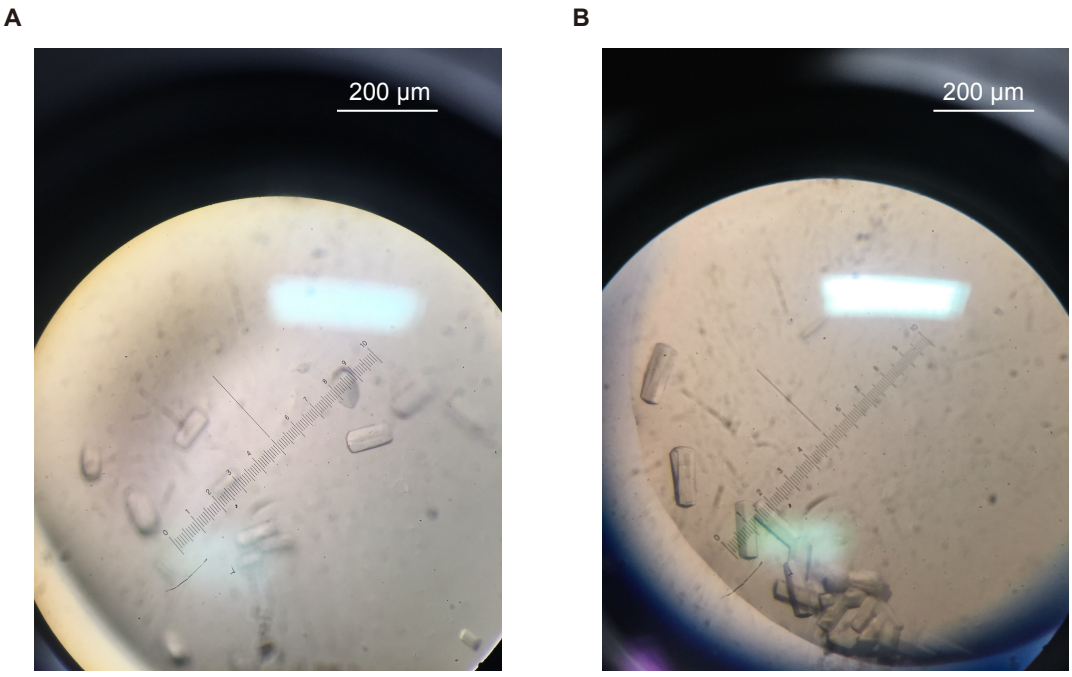


Fig. 3.1 Micrographs of DdP4H (60-284) crystals. (A) Crystals from Pact Premier, well D9.3: 0.2 M lithium chloride, 0.1 M Tris, pH 8.0, 20 % w/v PEG 6000. (B) Pact Premier, well D7.2: 0.2 M sodium chloride, 0.1 M Tris, pH 8.0, 20% w/v PEG 6000.

Based on the crystal distribution in the initial crystallisation screens, optimisation blocks were designed varying salt type (NaCl and LiCl), concentration of salt (0.1 M, 0.15 M, 0.2 M, 0.25 M), concentration of PEG 6000 (11%, 14%, 17%, 20%, 23%, 25%), buffer (Tris and HEPES) and pH (6.8, 7.2, 7.6, 8.0). A large number of crystals were obtained from the optimisation trials (Table 3.3). The quality of crystals improved in both size and amount relative to those obtained in the initial screens. Further improvement was attempted using additive screens: Silver Bullet HR2-088, Silver Bullet HR2-096 and Detergent Screen HR2-406 (Hampton Research).

Blocks	Wells	Conditions
Pact Premier	C7.2	0.2 M Sodium Chloride, 0.1 M HEPES pH 7.0, 20% PEG 6000
Pact Premier	C9.2	0.2 M Lithium Chloride, 0.1 M HEPES pH 7.0, 20% PEG 6000
Pact Premier	D9.2	0.2 M Lithium Chloride, 0.1 M Tris pH 8.0, 20% PEG 6000

Pact Premier	C12.2	0.01 M Zinc Chloride, 0.1 M HEPES pH 7.0, 20% PEG 6000
Opt Tray	A3.3	0.1 M Sodium Chloride, 0.1 M HEPES pH 6.8, 17% PEG 6000
Opt Tray	A4.1	0.1 M Sodium Chloride, 0.1 M HEPES pH 6.8, 20% PEG 6000
Opt Tray	A4.2	0.1 M Sodium Chloride, 0.1 M HEPES pH 6.8, 20% PEG 6000
Opt Tray	A4.3	0.1 M Sodium Chloride, 0.1 M HEPES pH 6.8, 20% PEG 6000
Opt Tray	A9.1	0.1 M Sodium Chloride, 0.1 M HEPES pH 7.2, 17% PEG 6000
Opt Tray	B3.2	0.15 M Sodium Chloride, 0.1 M HEPES pH 6.8, 17% PEG6000
Opt Tray	B4.3	0.15 M Sodium Chloride, 0.1 M HEPES pH 6.8, 20% PEG 6000
Opt Tray	B5.1	0.15 M Sodium Chloride, 0.1 M HEPES pH 6.8, 23% PEG 6000
Opt Tray	B9.2	0.15 M Sodium Chloride, 0.1 M HEPES pH 7.2, 17% PEG 6000
Opt Tray	C3.2	0.2 M Sodium Chloride, 0.1 M HEPES pH 6.8, 17% PEG 6000
Opt Tray	C3.3	0.2 M Sodium Chloride, 0.1 M HEPES pH 6.8, 17% PEG 6000
Opt Tray	C4.1	0.2 M Sodium Chloride, 0.1 M HEPES pH 6.8, 20% PEG 6000
Opt Tray	C4.2	0.2 M Sodium Chloride, 0.1 M HEPES pH 6.8, 20% PEG 6000
Opt Tray	C4.3	0.2 M Sodium Chloride, 0.1 M HEPES pH 6.8, 20% PEG 6000
Opt Tray	D3.3	0.25 M Sodium Chloride, 0.1 M HEPES pH 6.8, 17% PEG6000
Opt Tray	D4.1	0.25 M Sodium Chloride, 0.1 M HEPES pH 6.8, 20% PEG 6000
Opt Tray	D4.2	0.25 M Sodium Chloride, 0.1 M HEPES pH 6.8, 20% PEG 6000
Opt Tray	D4.3	0.25 M Sodium Chloride, 0.1 M HEPES pH 6.8, 20% PEG 6000

Table 3.3 List of DdP4H (60-284) crystals obtained from the crystallisation optimisation trials.

3.4 X-ray Crystallography of DdP4H (60-284)

The X-ray crystallography work in this part of the thesis was conducted with Dr. Rasheduzzaman Chowdhury (Department of Chemistry, Chemistry Research Laboratory, University of Oxford, OX1 3TA).

Crystals of DdP4H(60-284).Ni(II).NOG complexes [23 mg/mL DdP4H (60-284), 2 mM $NiCl_2$, 2 mM NOG] were obtained in sitting drops at 25 °C in buffer conditions containing 50 mM HEPES, 100 mM NaCl, pH 7.6. Crystals were cryo-protected with 25% glycerol before flash-cooling in liquid nitrogen. Data were collected using single crystals at 100 K at the Diamond Light Source (DLS) beam line I03 (λ 0.97623 Å) and processed using XDS [1] in the $C222_1$ space group (~2.0 Å resolution) as described in Table 3.4. A DdP4H (60-284) structure was solved by molecular replacement using PHASER [2] locating 3 molecules in the asymmetric unit and an ensemble model based on the PPHD (PDB: 4J25) and PHD2 (PDB: 5L9R) crystal structures. The structure was refined using CNS [3] / PHENIX [4] until R_{factor}/R_{free} converged with iterative cycles model

building using COOT [5]. MOLPROBITY [6] was used to monitor the geometric quality of the models between refinement cycles. Water molecules were added to peaks (> 1.5 σ) in $2F_oF_c$ electron density maps that were within H-bond distance to protein with reasonable geometry to form H-bonds.

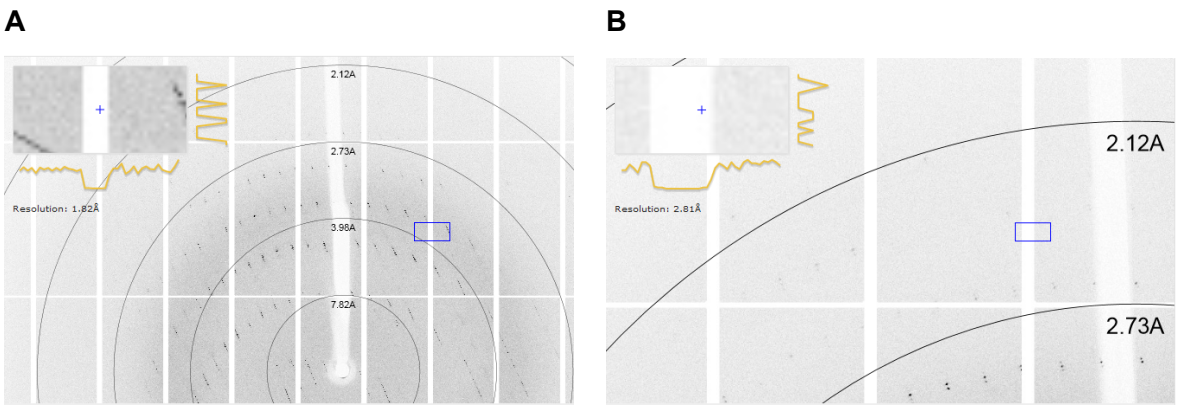


Fig. 3.2 Diffraction patterns of DdP4H (60-284).

Data Collection	
Beamline (Wavelength, Å)	Diamond Light Source IO3 (0.97623)
Detector	PILATUS 6M-F
Data Processing	XDS
Space Group	C 2 2 2_1
Cell Dimensions a, b, c (Å)	45.21, 81.74, 379.07
No. of Molecules / ASU	3
Resolution (Å)	47.38 – 2.02 (2.13 – 2.02)*
No. of Unique Reflections	47025 (6792)*
Completeness (%)	100.0 (100.0)*
Redundancy	9.1 (9.1)*
$^{**}R_{sym}$ / R_{merge}	0.103 (1.426)*
Mean I/σ(I)	12.2 (1.7)*
Refinement	
$^{‡}R_{factor}$	0.216
$^{‡}R_{free}$	0.256
R.m.s. Deviation	
Bond Length, Å	0.004
Bond Angle, °	0.685
Ramachandran Plot	
Most Favoured Region	96.59
Additionally Allowed Region	3.41
Disallowed Region	0.0

Table 3.4 Crystallographic data processing and refinement statistics.

*Highest resolution shell shown in parenthesis.
**$R_{sym}/R_{merge} = \Sigma|I - \langle I \rangle| / \Sigma I$, where I is the intensity of an individual measurement and $\langle I \rangle$ is the average intensity from multiple observations.
‡$R_{factor} = \Sigma_{hkl}||F_{obs}(hkl)| - k|F_{calc}(hkl)|| / \Sigma_{hkl} |F_{obs}(hkl)|$ for the working set of reflections; R_{free} is the R_{factor} for ~5% of the reflections excluded from refinement.

3.5 Overall Structure of DdP4H (60-284)

Consistent with the structures of other 2OG oxygenases, the DdP4H (60-284) crystal structure contains a core double stranded β-helix (DSBH) motif, which is composed of 8 β-strands (I - VIII). The DSBH β-strands pack to form major and minor β-sheets that surround the metal, Ni^{2+}, and NOG active site. Six antiparallel β-strands [β1 - β2 - β3(I) – β5(III) – β8(VI) – β10(VIII)] form the major β-sheet and four antiparallel β-strands [β4(II) – β6(IV) – β7(V) – β9(VII)] form the minor β-sheet. Three helices (α1 - α3) pack against the major β-sheet. Two disordered loops connect N134 – L142 and N230 - E244, respectively.

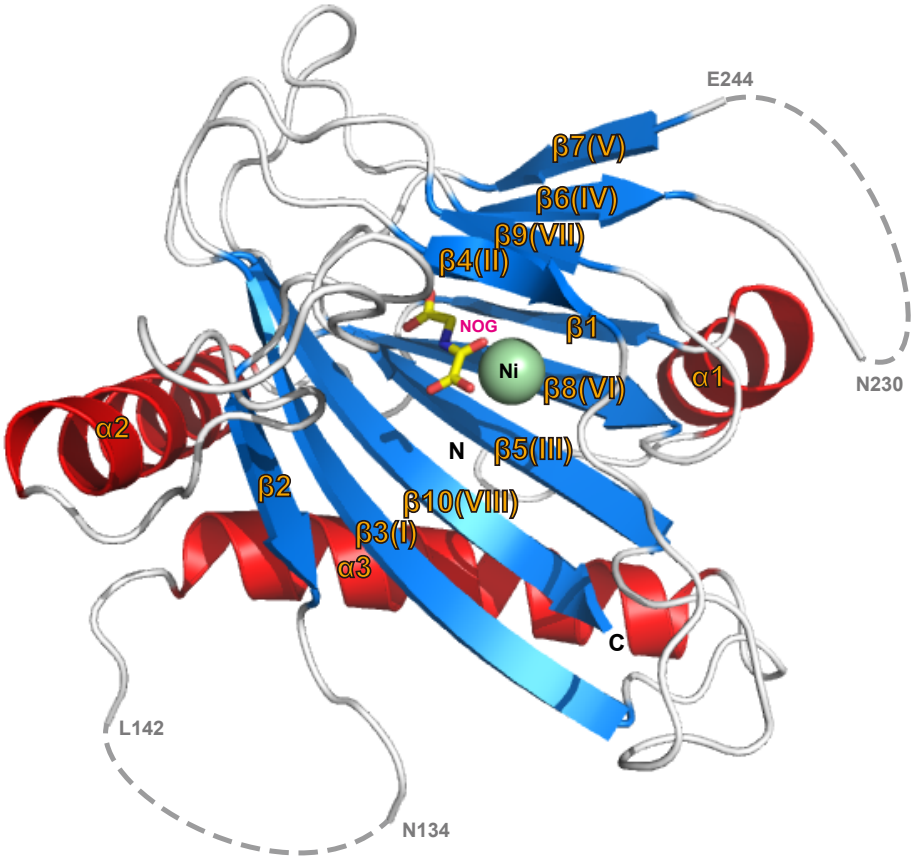


Fig 3.3 View of the overall structure of DdP4H (60-284). Cartoon representation of the overall structure of DdP4H (60-284). β-strands are coloured in blue. Roman numerals indicate the core DSBH strands. α-helices are in red, loops in white. NOG is shown as sticks and resides in the active site within the DSBH. The metal ion, Ni^{2+}, is shown as a light green sphere.

3.6 Active Site of DdP4H (60-284)

The active site of DdP4H (60-284) is found within the DSBH motif and contains three metal binding ligands: two histidine residues and an aspartate residue. These three residues form the conserved HXD. . .H motif, which is present in most other 2OG oxygenases [7]. They are His191, Asp193 and His266 in DdP4H (60-284). His191 and Asp193 are positioned on strand β4 (II) and the loop between strand β5 (III) of the DSBH. His266 is positioned on strand β9 (VII). The metal ion (Ni^{2+}) is also chelated by the oxalyl group of NOG to complete an octahedral metal coordination. NOG is positioned to the Arg276 in a formation of two salt bridges, between the carboxyl group of NOG and the guanidino group of Arg276.

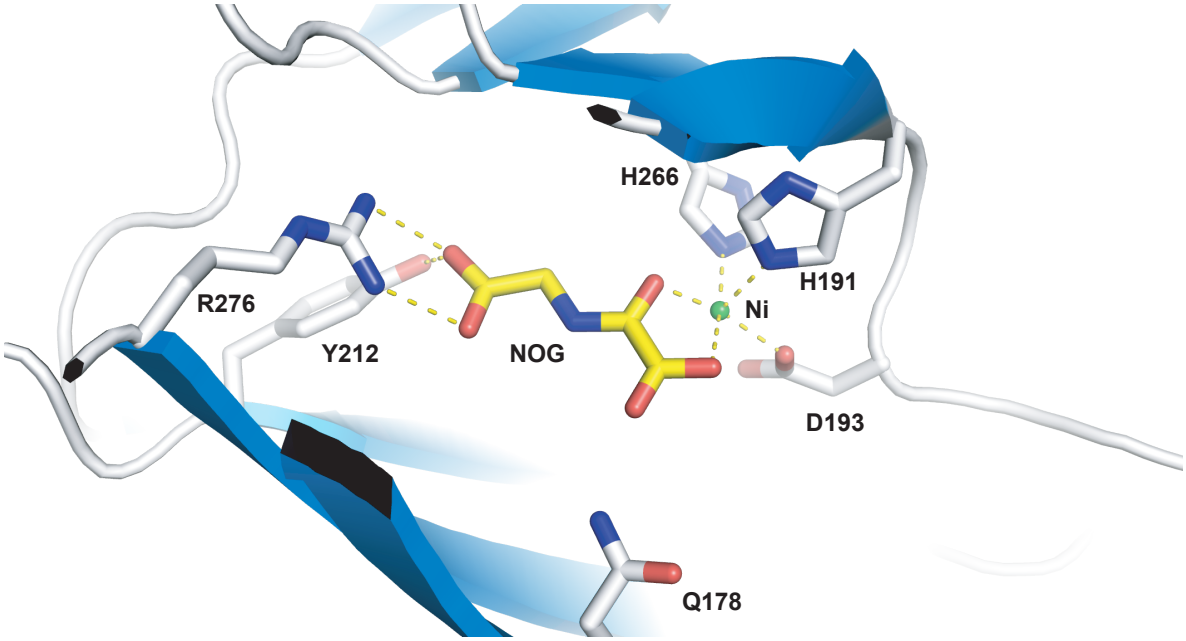


Fig 3.4 View of the active site of DdP4H (60-284). Stick representation of the active site of DdP4H (60-284). β-strands are coloured in marine. The metal ion, Ni^{2+}, is shown as a light green sphere. Dashed lines indicate coordination bonds surrounding Ni^{2+}, and hydrogen bonds between Arg276, Tyr212 and NOG.

3.7 Comparison of Structures of DdP4H (60-284), PPHD and PHD2

Superimposition of DdP4H (60-284) and human PHD2 structures (PDB ID: 5L9R [8]) reveals that their overall structures are similar (Figure 3.5). However, differences do exist between them; these differences help to rationalise their different functions:

1. The C-terminus of DdP4H (60-284) ends after the last β-strand, β10(VIII), yet the C-terminus of PHD2 contains an α-helix (α4); NMR studies have shown PHD2 α4 to be involved in substrate recognition and inhibitor binding [9]. Notably, the DdP4H (60-284) construct is N-terminally truncated. The C-terminus of DdP4H (60-284) may reflect the actual structure of the full-length DdP4H.
2. The N-terminus of DdP4H (60-284) is similar to that of PHD2, consisting of an α-helix (α1), which lies against the major sheet of the DSBH. A BLAST search of the truncated region, DdP4H (1-59), does not yield any similar sequence in the database. The truncated region may have different functions other than prolyl hydroxylation or would lead to more efficient hydroxylation by binding to the substrate. The truncated part of DdP4H (1-59) may be flexible and disordered, which is consistent with the secondary structure prediction (Chapter 2.8).
3. The β3 (I)-strand of DdP4H (60-284) appears much longer than in PHD2, and may play a role in substrate binding.
4. Two α-helices found in PHD2, including the one between α2 - β3 and the one between β5 (III) – β6 (IV), are not present in DdP4H (60-284).
5. Two disordered loops connecting N134 – L142 and N230 - E244 are longer compared with analogous regions in PHD2. These loops may be involved in substrate binding in DdP4H (as they are in PHD2).

6. When the substrate CODD/NODD binds to PHD2, the β2/β3 loop closes to promote substrate binding. Such a long loop also exists in DdP4H, where it may also play an important role in substrate binding.

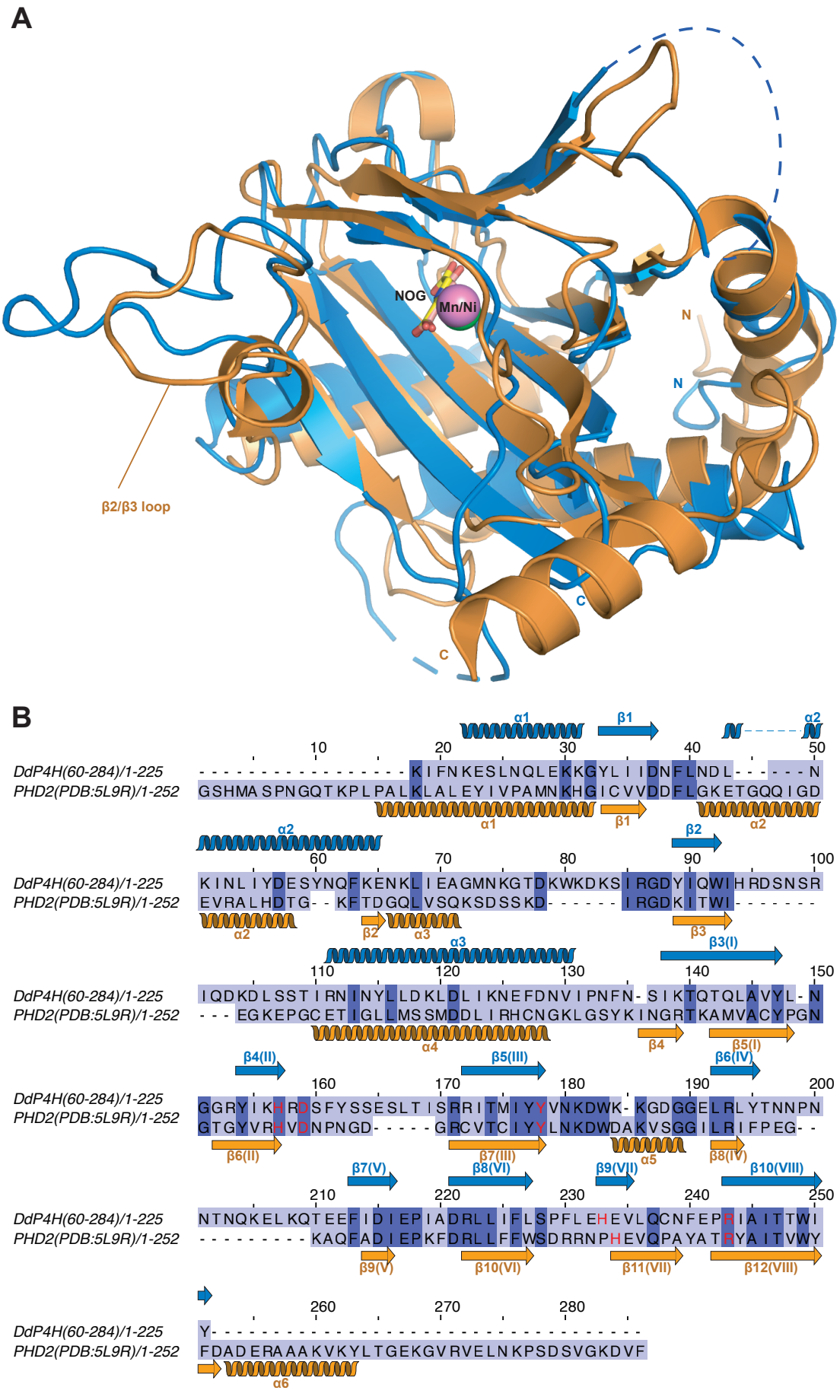

A
NOG
Mn/Ni
N
N
β2/β3 loop
C
C
B
DdP4H(60-284)/1-225
PHD2(PDB:5L9R)/1-252
α1
β1
α2
β2
α3
β3
β3(I)
α4
β4
β5(I)
β4(II)
β5(III)
β6(IV)
β6(II)
β7(III)
α5
β8(IV)
β7(V)
β8(VI)
β9(VII)
β10(VIII)
β9(V)
β10(VI)
β11(VII)
β12(VIII)
α6

Fig 3.5 Superimposition of structures for DdP4H (60-284) and human PHD2 (PDB ID: 5L9R). (A) The crystal structure of DdP4H (60-284) is coloured in marine. NOG is shown as yellow sticks residing in the active site surrounded by DSBH. The metal ion, Ni^{2+}, is shown as a light green sphere. PHD2 is coloured in orange. NOG is shown as orange sticks. The metal ion, Mn^{2+}, is shown as a pink sphere. (B) Sequence alignment of DdP4H (60-284) and PHD2. Amino acid residues taking part in the active site binding are coloured in red.

The structure and function of PPHD is discussed in Chapter 1.6.

The DdP4H (60-284) and PPHD structures (PDB ID: 4J25 [10]) are also similar overall by comparing the superimposition of these two protein structures (Figure 3.6). However, several notable differences are present:

1. Two additional α-helices exist in PPHD, including one between β6(III) – β7(IV) and one between β7(IV) – β8(V).
2. Two disordered loops of DdP4H (60-284) connecting N134 – L142 and N230 - E244 are longer compared to the same regions in PPHD.
3. When the substrate EF-Tu binds to PPHD, the β2/β3 loop closes to promote substrate binding. Such a long loop also exists in DdP4H, where it may also play an important role in substrate binding.

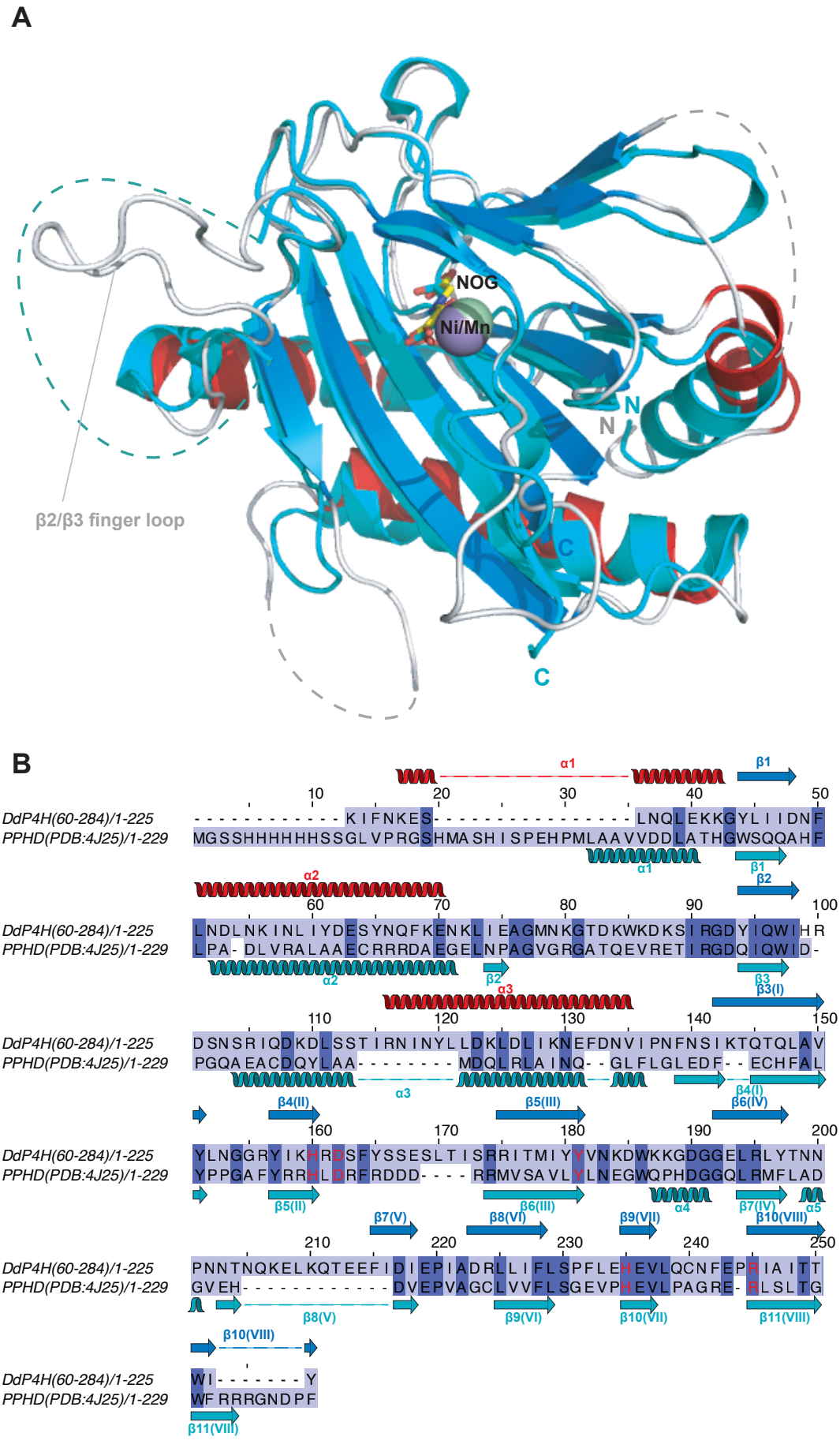

A
NOG
Ni/Mn
N
N
C
C
β2/β3 finger loop
B
DdP4H(60-284)/1-225
PPHD(PDB:4J25)/1-229
α1
β1
α2
β2
α3
β3
β3(I)
β4(I)
β4(II)
β5(III)
β6(IV)
β5(II)
β6(III)
α4
β7(IV)
α5
β7(V)
β8(VI)
β9(VII)
β10(VIII)
β8(V)
β9(VI)
β10(VII)
β11(VIII)

Fig 3.6 Superimposition of structures for DdP4H (60-284) and PPHD (PDB ID: 4J25). (A) In DdP4H (60-284), β-strands are coloured in marine, α-helices in red, loops in white. NOG is shown as yellow sticks in the active site surrounded by DSBH. The metal ion, Ni^{2+}, is shown as a light green sphere. PPHD is coloured in light blue. NOG is shown as light blue sticks. The metal ion, Mn^{2+}, is shown as a pink sphere. (B) Sequence alignment of DdP4H (60-284) and PPHD. Amino acid residues taking part in the active site binding are coloured in red.

In summary, the crystallised structure of DdP4H (60-284) shows more disordered regions than PHD2. The DdP4H (60-284) structure is similar to PPHD (29% identity based on Phyre2 [11]) and PHD2 (28% identity). This result might imply a close evolutionary relationship of prolyl hydroxylases between *Dictyostelium discoideum*, *Psudomonas putida* and Human. The more disordered regions observed in DdP4H (60-284) may explain its lower hydroxylation efficiency compared to PHD2 (Chapter 4.23). The regions close to the open end of DSBH barrel are less similar between the three enzyme structures, which may be consistent with a role for this region in different substrate recognition and binding likely manifest between the enzymes.

3.8 Comparison of the Active Sites of DdP4H (60-284), PHD2 and PPHD

The active site of DdP4H (60-284) superimposes well with those of PHD2 (Figure 3.7 A) and PPHD (Figure 3.7 B), showing a very conserved core structure. $His266_{DdP4H}$ / $His374_{PHD2}$ / $His183_{PPHD}$, $His191_{DdP4H}$ / $His313_{PHD2}$ / $His124_{PPHD}$, $Asp193_{DdP4H}$ / $Asp315_{PHD2}$ / $Asp126_{PPHD}$ bind to the central metal ion in a conserved manner. $Arg276_{DdP4H}$ / $Arg383_{PHD2}$ / $Arg192_{PPHD}$ and $Tyr212_{DdP4H}$ / $Tyr392_{PHD2}$ / $Tyr141_{PPHD}$ also support the binding of NOG (and likely 2OG) in the active site through hydrogen bonds and salt bridges.

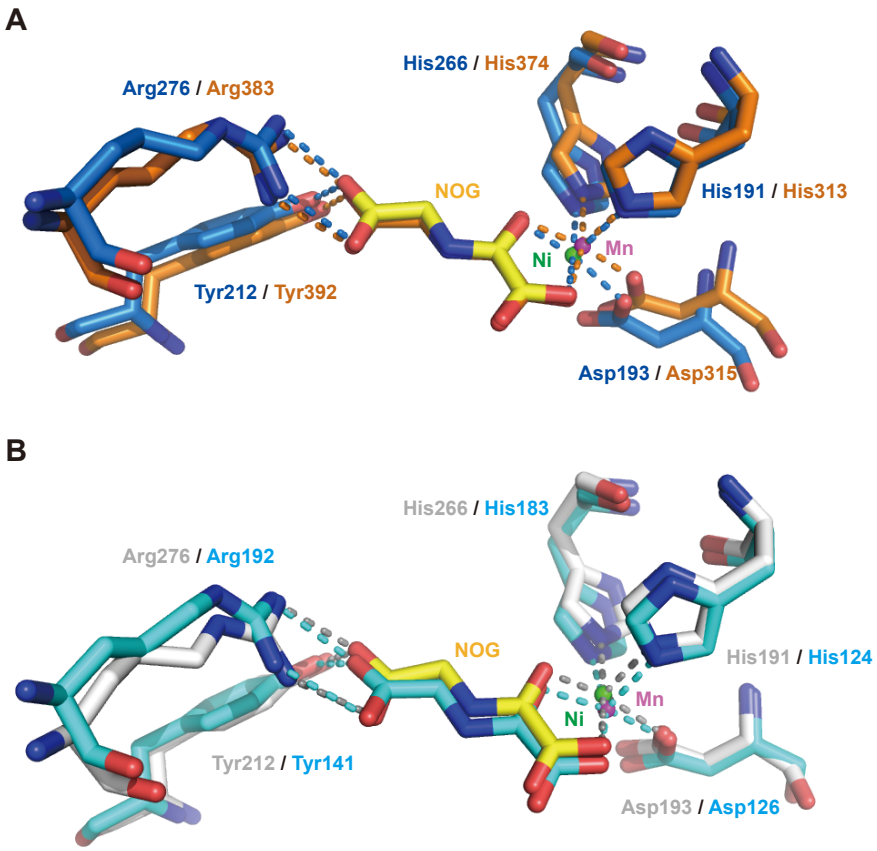


Fig 3.7 Superimposition of the active site views for DdP4H (60-284) and PHD2 (PDB ID: 5L9R), DdP4H (60-284) and PPHD (PDB ID: 4J25). (A) Stick representations of active site residues of DdP4H (60-284) (marine) and PHD2 (orange). The metal ions, Ni^{2+} and Mn^{2+}, are shown as light green and pink spheres. NOG is shown as yellow and orange sticks. Metal coordination, salt bridges and hydrogen bonds are indicated as dashes. (B) Stick representations of DdP4H (60-284) (white) and PHD2 (light blue) active sites. The metal ions, Ni^{2+} and Mn^{2+}, are shown as light green and pink spheres. NOG is shown as yellow and light blue sticks. Metal coordination, salt bridges and hydrogen bonds are indicated as dashes.

3.9 Crystal Structure of DdP4H (60-284) with a Model of the DdSkp1 Substrate

DdSkp1 is hydroxylated by DdP4H at Pro143 as described in Chapter 4.11 and Chapter 4.24. It was therefore of interest to obtain a structure of DdP4H complexed with its substrate, DdSkp1. Attempts to crystallise full-length DdSkp1 on its own and attempts to co-crystallise DdP4H (60-284) with full-length DdSkp1 by setting up crystallisation trials using the mixture of purified DdP4H (60-284) and purified full-length DdSkp1 were unsuccessful. Therefore, a model of

DdSkp1 was generated using the Protein Homology/analogY Recognition Engine V 2.0 (Phyre2, http://www.sbg.bio.ic.ac.uk/phyre2/) with the help of Dr. Rasheduzzaman Chowdhury (Department of Chemistry, Chemistry Research Laboratory, University of Oxford, OX1 3TA). The model of DdSkp1 was used to manually docked to the structure of DdP4H (60-284) with target Pro143 of DdSkp1 being placed in the active site, with the C-4 position closest to the central metal ion (Figure 3.8).

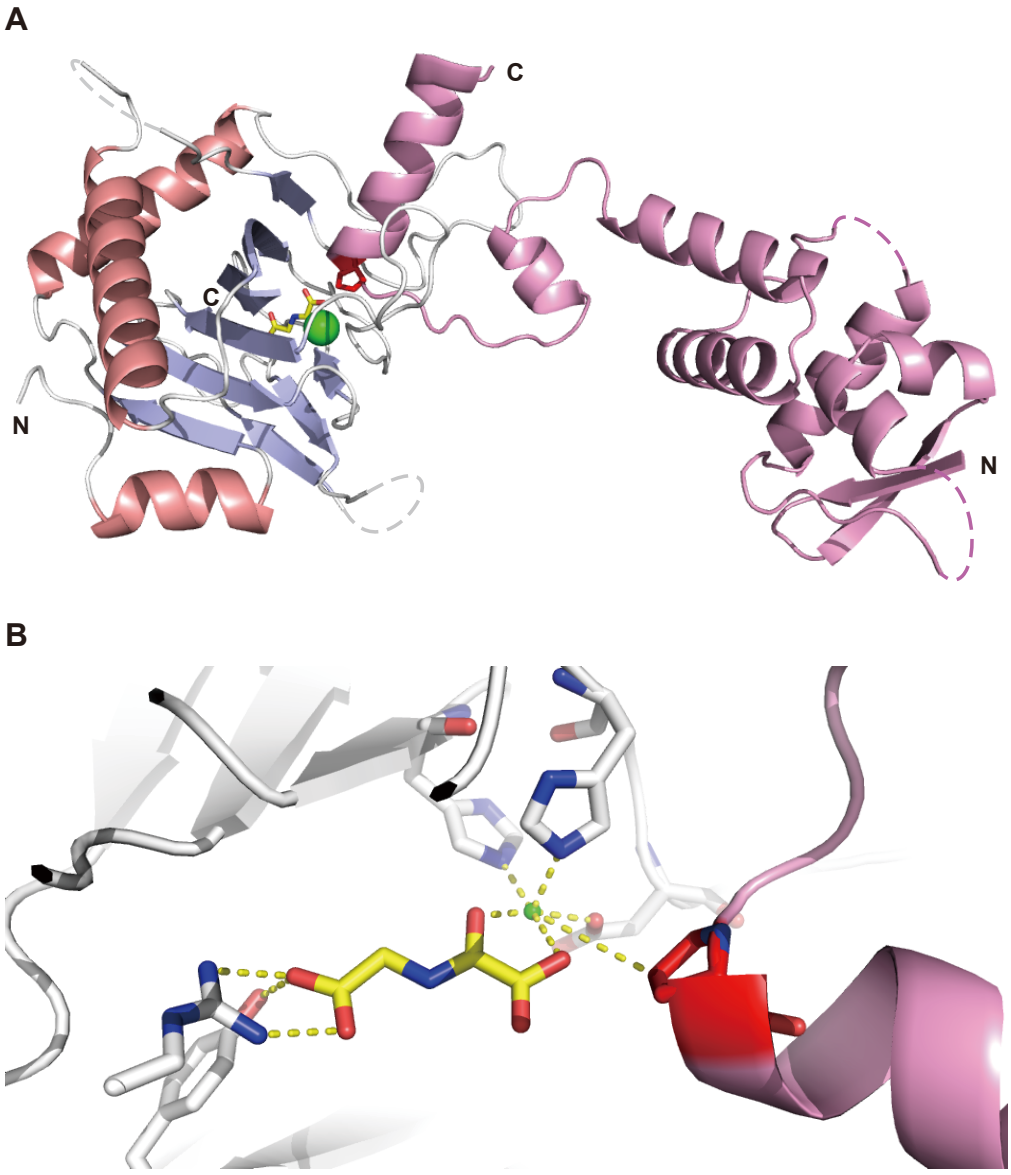


Fig 3.8 View from the structure for DdP4H (60-284) with a model of DdSkp1 in the active site. (A) Cartoon representation of the crystal structure of DdP4H (60-284). β-Strands are in light blue. α-Helices are in salmon, loops in white. NOG is shown as yellow sticks in the active site. The

metal ion, Ni^{2+}, is shown as a light green sphere. The DdSkp1 is represented as pink cartoons. The target Pro143 is in red. (B) Stick representation of the active site of DdP4H (60-284) in white. The metal ion, Ni^{2+}, is shown as a light green sphere. NOG is shown as yellow sticks. The DdSkp1 model is shown in pink. The target Pro143 is shown in red. Dashed lines indicate Ni^{2+} coordination, hydrogen bonds/salt bridges between Arg276 and NOG, and the distance between C-4 of Pro143 and the central Ni^{2+}.

3.10 Disulfide-bond Crosslinking to Crystallise DdP4H (60-284) and DdSkp1 Complex

Attempts to co-crystallise DdP4H (60-284) with DdSkp1 peptides did not result in crystals. Kinetic studies also showed that DdSkp1 was not a good substrate of DdP4H with lower hydroxylation efficiency than CODD or NODD peptide of HIF-1α with PHD2 (Chapter 4.23). A crosslinking strategy was therefore explored in order to covalently link DdSkp1 peptide to DdP4H (60-284) aiming to improve crystallisation of the DdP4H (60-284): DdSkp1 complex.

Cysteines located on different proteins or different domains of the same protein can form a disulfide bond in an oxidising environment when positioned close enough to each other, a strategy that has been useful in obtaining 2OG oxygenase substrate complexes [12, 13]. It was proposed that this strategy may help to stabilise the DdP4H (60-284) and DdSkp1 complex for crystallography [14].

The crystal structure of DdP4H (60-284) shows that the side chain of the only cysteine residue, Cys271, is buried inside of the enzyme. Therefore, no sulfhydryl groups are exposed on the surface of DdP4H (60-284), providing a potentially optimal situation for cysteine substitution by site-directed mutagenesis and disulfide-bond crosslinking with DdSkp1.

3.10.1 Using the DdP4H (60-284) Structure and DdSkp1 Model for Crosslinking Design

Using the structure of the DdP4H (60-284) DdSkp1 modelled in Chapter 3.9, amino acid residues in the loop between β4 (II) and β5 (III), and the residues in DdSkp1 around the target Pro143 were

examined for suitability to form crosslinks with cysteines on a substrate. Considering the orientation of the side chains, the amino acid residues, Arg192 and Phe195 in DdP4H (60-284), were selected to be substituted to cysteine. DdSkp1 peptides with potentially suitable cystesines were also made. The detailed structure of DdP4H (60-284) with DdSkp1 model is shown in Figure 3.9.

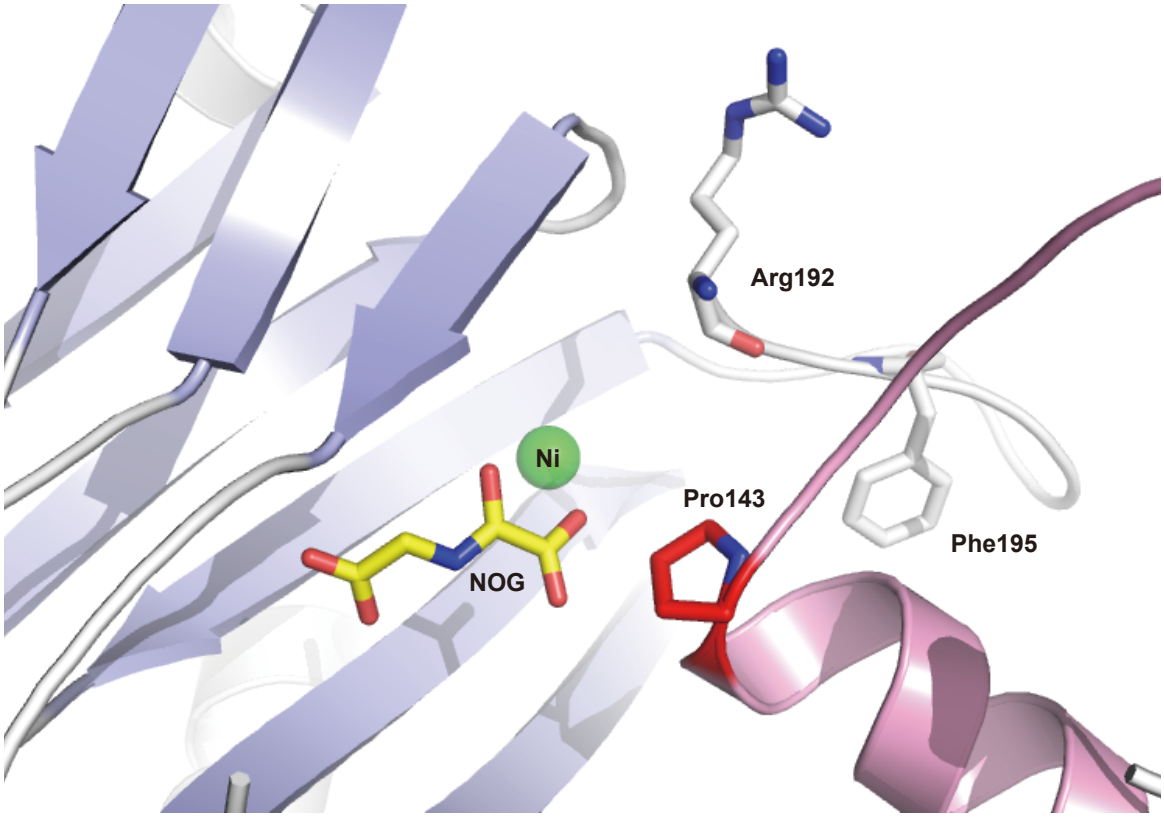


Fig 3.9 The structure of DdP4H (60-284) with DdSkp1 modelled in. Cartoon representation of the crystal structure of DdP4H (60-284). β-Strands are in light blue, loops in white. NOG is shown as yellow sticks in the active site. The metal ion, Ni^{2+}, is shown in light green as a sphere. The modelled DdSkp1 structure is represented as pink ribbons. The target Pro143 is shown as red sticks. The Arg192 and Phe195, aimed to be mutated to cysteines, are shown as white sticks.

3.10.2 Site-directed Mutagenesis of DdP4H (60-284)

Work in this part of the thesis was conducted with the help of Dr. Rasheduzzaman Chowdhury.

Site-directed mutagenesis was conducted to substitute Arg192 and Phe195 of DdP4H (60-284) for cysteine residues.

The primers designed specifically for mutating R192C were:

Forward, 5’ – GTAGATATATTAAACATTGTGATTCATTTTATTCCTC – 3’;

Reverse, 5' – GAGGAATAAAATGAATCACAATGTTTAATATATCTAC – 3'.

The primers designed specifically for mutating F195C were:

Forward, 5' – CATAGAGATTCATGTTATTCCTCAGAATC – 3';

Reverse, 5' – GATTCTGAGGAATAACATGAATCTCTATG – 3'.

The DNA was amplified by polymerase chain reaction (PCR) using the following protocol:

A

PCR Reagents	Volume Per Tube / µL
5x Q5 buffer (NEB #B9027S)	10
dNTP (NEB #N0447S)	1
Q5 polymerase (NEB #M0491S)	1
Primer mix (10 µM)	4
Template DNA (40 ng – 100 ng)	1
MilliQ water	2.5
Total	50

B

PCR Steps	Temperature / °C	Incubation Time
1	98	2 min
2	98	30 s
3	50	1 min
4	72	8 min 40 s
5	Repeat Step 2 – 4 for 16 cycles	
6	72	2 min
7	4	Pause

Table 3.5 Procedures for site-directed mutagenesis to generate DdP4H (60-284) R192C and DdP4H (60-284) F195C. (A) Sample contents for PCR. (B) PCR protocol.

The new constructs were transformed into XL10 Gold *E. coli* competent cells [15] and were further purified and stored for future use.

3.10.3 Expression and Purification of DdP4H (60-284) R192C

The DdP4H (60-284) R192C construct was transformed into BL21-Gold (DE3) *E. coli* competent cells and the gene encoding for the recombinant protein was expressed under the same conditions as for DdP4H (60-284) (Chapter 2.11). The untagged *apo*-DdP4H (60-284) R192C (~30 mg) was purified using the same method as that used for purification of

apo-DdP4H (60-284) described in Chapter 2.11. The purification result is shown in Figure 3.10.

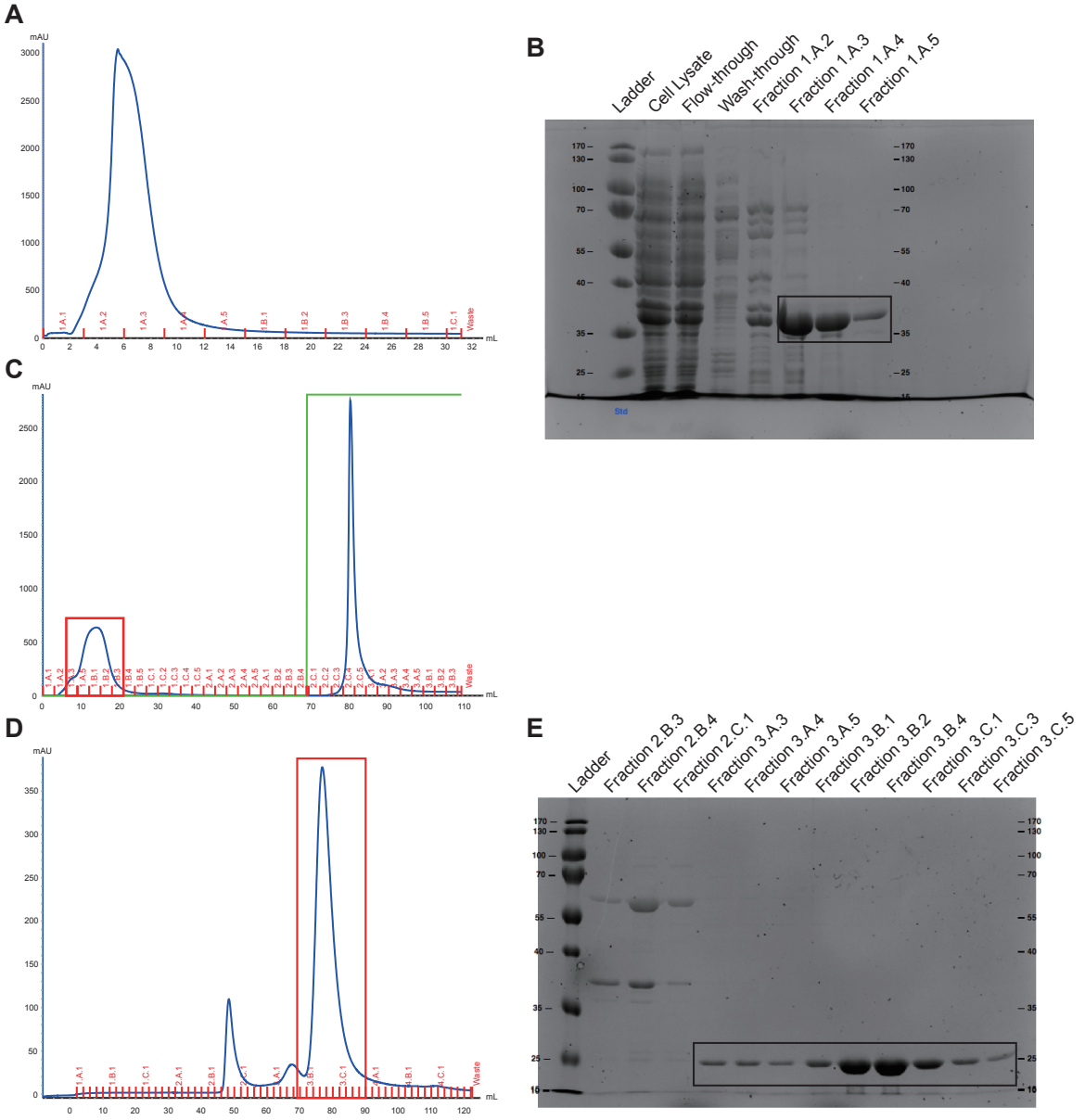


Fig. 3.10 Purification of DdP4H (60-284) R192C. (A) Ni(II) affinity chromatography UV chromatogram elution profile. (B) SDS-PAGE analysis. The lanes from left to right are: Ladder (PageRuler from Thermo Scientific), cell lysate, flow-through, wash-through, fractions 1.A.2 – 1.A.5. The bands corresponding to His_6-ZB-DdP4H (60-284) R192C are indicated by a box. (C) Ni(II) affinity chromatography UV chromatogram after TEV and EDTA treatment overnight. Flow-through (0 – 40 mL) were collected in fractions. The green line showed the transition from 100% binding buffer (5 mM imidazole, 0.5 M NaCl, 50 mM HEPES, pH 7.6) to 100% elution buffer (500 mM imidazole, 0.5 M NaCl, 50 mM HEPES, pH 7.6). Fractions in red box were collected and were further purified using S75 size exclusion chromatography. The final storage buffer was 50 mM HEPES, pH 7.6. (D) Size exclusion UV chromatogram. (E) SDS-

PAGE analysis. The bands corresponding to the untagged *apo*-DdP4H (60-284) R192C are boxed.

3.10.4 Synthesis of DdSkp1 Peptides

The DdSkp1 peptides with cysteine variants were synthesised using a Liberty Blue™ Automated Microwave Peptide Synthesiser (CEM Corporation), based on a fragment of the original DdSkp1 sequence, FNIKNDFTPEEEEQIRKENE.

Peptide Name	Sequence	Peptide Mass / Da
DdSkp1 F135C	CNIKNDFTPEEEEQIRKENE	2465.63
DdSkp1 N136C	FCIKNDFTPEEEEQIRKENE	2498.70
DdSkp1 I137C	FNCKNDFTPEEEEQIRKENE	2499.65
DdSkp1 K138C	FNICNDFTPEEEEQIRKENE	2484.63
DdSkp1 N139C	FNIKCDFTPEEEEQIRKENE	2498.70
DdSkp1 D140C	FNIKNCFTPEEEEQIRKENE	2497.72
DdSkp1 F141C	FNIKNDCTPEEEEQIRKENE	2465.63
DdSkp1 T142C	FNIKNDFCPEEEEQIRKENE	2511.70

Table 3.6 Sequences of synthesised peptides.

3.10.5 Time Course Assays of Disulfide-Bond Crosslinking

The potential formation of a disulfide bond between the purified DdP4H (60-284) R192C protein and DdSkp1 cysteine variant peptides was then invesigated. The reactions were set up under the conditions described in Table 3.8. Samples were incubated at 4 °C until being quenched with 1% (v/v) formic acid at different time points [1 h, 2 h, 4 h, 8 h, O/N (24h)]. Quenched samples were stored at -80 °C before being analysed using RapidFire Mass Spectrometry (RF-MS). Detailed acquisition parameters are discussed in Chapter 4.17. Data were analysed using MassHunter (Agilent).

	Stock Concentration	Volume To Add	Final Concentration
DdP4H (60-284) R192C	744 µM	5 µL	300 µM
Peptide	4 mM	7.5 µL	1500 µM
Ni(II)	50 mM	0.8 µL	2 mM
NOG	100 mM	0.4 µL	2 mM
Buffer (100mM HEPES, pH 7.6)	-	6 µL	-

Table 3.7 Ingredients of samples for disulfide-bond crosslinking time course assays.

The RapidFire results showed no evident covalently linked complex for all samples, although peaks of DdP4H (60-284) R192C could be observed after de-convoluting the spectra. However, no peaks representing enzyme + peptide complexes were observed, suggesting that no disulfide-bonds had formed even after 24-hour incubation.

3.11 Discussions and Future Plans

Crystals of DdP4H (60-284) in complex with Ni(II) and NOG were obtained and diffraction data were successfully collected at the synchrotron. The X-ray crystal structure of DdP4H (60-284).Ni(II).NOG complex was successfully determined. DdP4H (60-284) was found to share some common features with other prolyl hydroxylases, including *Psudomonas* PHD (PPHD) and human PHD2. They all contain a core double stranded β-helix (DSBH) motif, and the active sites all contain the HXD. . .H motif, forming a central iron and 2OG binding site.

The findings lead to more work needs to be accomplished in the future:

3.11.1 Crystal Structure of DdP4H (60-284) in Complex with DdSkp1 Substrate

The initial trials screening for disulfide-bond crosslinking using DdP4H (60-284) R192C were unsuccessful. No disulfide bond formation after the incubation was observed with the tested cysteines. However, it is worth repeating the experiment including using different recombinant mutant enzymes, e.g. DdP4H (60-284) F195C. Other strategies to form a more stable enzyme-substrate complex could be explored, e.g. using inhibitors that interact with both enzyme and substrate.

3.11.2 Attempted Crystallisation of TgPhyA (56-264)

Although the TgPhyA (56-264) construct was designed (Chapter 2.8) and purified (Chapter 2.12) with the aim of obtaining a crystal structure of TgPhyA. Initial crystallisation trials of TgPhyA (56-264) were attempted in the presence of stabilising components identified in T_m shift assays (2 mM $NiCl_2$, 2 mM NOG, 100 mM HEPES, pH 7.6) using commercial sparse matrix crystallisation screens, including INDEX (Hampton Research), Wizard (Molecular Dimensions), JCSG+ (Molecular Dimensions), PACT Premier (Molecular Dimensions), SaltRx (Hampton Research) and PEG/Ion (Hampton Research), and MIDAS (Molecular Dimensions) screens. No protein crystals were observed after one month's incubation. Further optimisation, include of different constructs, needs to be carried out in the future.

3.12 References

1. Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement, *Acta Crystallogr D Biol Crystallogr 66*, 133-144.

2. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software, *J Appl Crystallogr 40*, 658-674.

3. 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., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr D Biol Crystallogr 54*, 905-921.

4. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr D Biol Crystallogr 66*, 213-221.

5. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr D Biol Crystallogr 60*, 2126-2132.

6. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom

structure validation for macromolecular crystallography, *Acta Crystallogr D Biol Crystallogr 66*, 12-21.

7. Aik, W., McDonough, M. A., Thalhammer, A., Chowdhury, R., and Schofield, C. J. (2012) Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases, *Curr Opin Struct Biol 22*, 691-700.

8. Chowdhury, R., Leung, I. K., Tian, Y. M., Abboud, M. I., Ge, W., Domene, C., Cantrelle, F. X., Landrieu, I., Hardy, A. P., Pugh, C. W., Ratcliffe, P. J., Claridge, T. D., and Schofield, C. J. (2016) Structural basis for oxygen degradation domain selectivity of the HIF prolyl hydroxylases, *Nat Commun 7*, 12673.

9. McDonough, M. A., Loenarz, C., Chowdhury, R., Clifton, I. J., and Schofield, C. J. (2010) Structural studies on human 2-oxoglutarate dependent oxygenases, *Curr Opin Struct Biol 20*, 659-672.

10. Scotti, J. S., Leung, I. K., Ge, W., Bentley, M. A., Paps, J., Kramer, H. B., Lee, J., Aik, W., Choi, H., Paulsen, S. M., Bowman, L. A., Loik, N. D., Horita, S., Ho, C. H., Kershaw, N. J., Tang, C. M., Claridge, T. D., Preston, G. M., McDonough, M. A., and Schofield, C. J. (2014) Human oxygen sensing may have origins in prokaryotic elongation factor Tu prolyl-hydroxylation, *Proc Natl Acad Sci U S A 111*, 13331-13336.

11. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis, *Nat Protoc 10*, 845-858.

12. Rhoads, D. M., Brunner-Neuenschwander, B., Levings, C. S., 3rd, and Siedow, J. N. (1998) Cross-linking and disulfide bond formation of introduced cysteine residues suggest a modified model for the tertiary structure of URF13 in the pore-forming oligomers, *Arch Biochem Biophys 354*, 158-164.

13. Gao, Y., and Mehta, K. (2001) Interchain disulfide bonds promote protein cross-linking during protein folding, *J Biochem 129*, 179-183.

14. Heinz, D. W., and Matthews, B. W. (1994) Rapid crystallization of T4 lysozyme by intermolecular disulfide cross-linking, *Protein Eng 7*, 301-307.

15. XL10 Gold Ultracompetent Cells, (Agilent, Ed.).

4. DdP4H and TgPhyA Activity Assay Development

4.1 Summary

This chapter describes work on the biochemical characterisation of the recombinant DdP4H and TgPhyA enzymes. Mass spectrometry based kinetics assays were established to investigate the biochemical properties of DdP4H and TgPhyA, including K_m values for 2OG, oxygen, peptide and full-length protein substrates, for comparison with other prolyl hydroxylases.

4.2 Introduction

Kinetic studies were first attempted using synthetic peptides as substrates and the recombinant DdP4H and TgPhyA enzymes (Chapter 4.3 – 4.6). The inability to acquire K_m values suggested that peptides were poor substrates and not suitable for kinetic studies. Procedures were then established to produce full-length protein substrates (Chapter 4.7 – 4.9). Kinetic studies were carried out using recombinant DdP4H and TgPhyA enzymes, and the recombinant DdSkp1 and TgSkp1 protein substrates. K_m values for substrates, 2OG, and oxygen were obtained and compared with other prolyl hydroxylases, including PHD2 and PPHD (Chapter 4.10 – 4.23). Tandem MS/MS analyses were carried out to confirm that proline hydroxylation of DdSkp1 and TgSkp1 was catalysed by DdP4H and TgPhyA (Chapter 4.24). 2D TOCSY NMR experiments were carried out to confirm the regiochemistry of the prolyl hydroxylations catalysed by DdP4H and TgPhyA (Chapter 4.25). Lastly, 2OG turnover catalysed by DdP4H, TgPhyA and PHD2 was analysed by NMR (Chapter 4.26).

4.3 Synthesis of DdSkp1, TgSkp1, HsSkp1, and HsSkp1 E147P Peptides

It was of interest to characterise the biochemical properties of the DdP4H and TgPhyA enzymes. A series of peptide fragments of DdSkp1 and TgSkp1 containing the target proline were

synthesised. HsSkp1 and HsSkp1 E147P were synthesised in parallel for inhibition experiments and cross species hydroxylation tests. HsSkp1 E147P is a variant of HsSkp1 changing a glutamate, which at the same position in DdSkp1 and TgSkp1 is a proline, to proline. The peptides were synthesised with C-terminal amides using a Liberty Blue™ Automated Microwave Peptide Synthesiser (CEM Corporation).

Protein	Sequence	Number of Amino Acid Residues
DdSkp1	FNIKNDFTPEEEEQIRKENE	20
	EIRKIFNIKNDFTPEEEEQIRKENE	25
	FNIKNDFTPEEEEQIRKENEWCEDK	25
	RKIFNIKNDFTPEEEEQIRKENEWCE	25
HsSkp1	FNIKNDFTEEEEAQVRKENQ	20
	FNIKNDFTEEEEAQVRKENQWCEEK	25
HsSkp1 E147P	FNIKNDFTPEEEAQVRKENQ	20
	FNIKNDFTPEEEAQVRKENQWCEEK	25
TgSkp1	FNIVNDFTPEEEAQVREENK	20
	FNIVNDFTPEEEAQVREENKWCEDA	25
	EIRRIFNIVNDFTPEEEAQVREENK	25

Table 4.1 List of synthesised peptides.

4.4 MALDI-MS Based Hydroxylation Activity Assay of DdP4H and TgPhyA with DdSkp1 and TgSkp1

The detection of hydroxylation of DdSkp1 peptide by DdP4H was initially performed using Matrix-Assisted Laser Desorption/ Ionization Mass Spectrometry (MALDI-MS). In the presence of Fe (II), 2OG, ascorbate and oxygen, the mass of DdSkp1 peptide exhibited a + 16 Da mass shift after 1-hour incubation as analysed by MALDI-MS, indicating that the DdSkp1 peptide was hydroxylated. Both Fe(II) and ascorbate were required to promote catalytic activity of DdP4H. TgSkp1 peptide was also hydroxylated by TgPhyA under the same conditions. The mass spectra are shown in Figure 4.1.

Assay Conditions	Final Concentration	Volume
2-oxoglutarate disodium salt (Sigma-Aldrich, K3752)	500 µM	2 µL
Sodium L-ascorbate (Sigma-Aldrich, 11140)	1 mM	2 µL
Ammonium iron(II) sulfate hexahydrate (Sigma-Aldrich, 215406)	50 µM	2 µL
DdSkp1 or TgSkp1 peptide substrate	100 µM	2 µL
DdP4H or TgPhyA enzyme	2 µM	2 µL
Buffer (100 mM HEPES, pH 7.6)	-	10 µL
Total volume	-	20 µL

Table 4.2 Assay conditions of the hydroxylation assay.

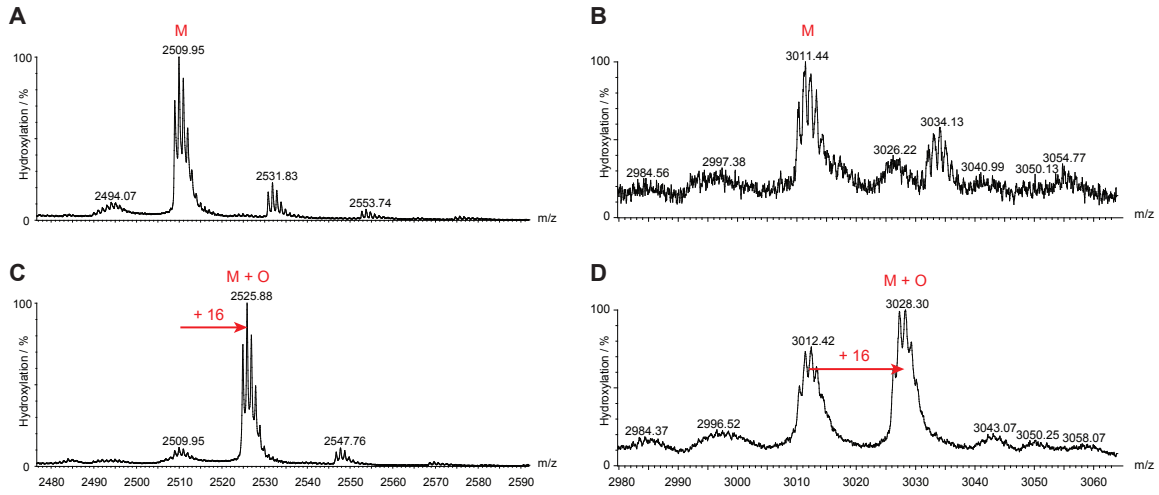


Fig. 4.1 Detection of hydroxylase activity for DdP4H and TgPhyA with DdSkp1 and TgSkp1 peptide substrates by MALDI-MS. (A) No enzyme control of DdSkp1 peptide (100 µM) in the presence of Fe(II) (50 µM), 2OG (500 µM), and ascorbate (1 mM) in 100 mM HEPES, pH 7.6 incubated at 37 °C for 1 hour. (B) Under the same conditions, no enzyme control of TgSkp1 peptide (100 µM). (C) DdSkp1 peptide (100 µM) underwent a +16 Da modification in the presence of 2 µM DdP4H incubated using the same conditions as above. (D) TgSkp1 peptide (100 µM) underwent a +16 Da modification in the presence of 2 µM TgPhyA incubated using the same conditions as above. Reactions were all quenched with 1% (v/v) aqueous formic acid before analysis by MALDI-MS.

4.5 Time Course Hydroxylation Assays

To investigate the kinetic properties of DdP4H, three different peptide fragments of DdSkp1 were assayed using MALDI-MS. The reactions were performed using the same conditions as above in triplicate. Reactions were quenched at different time points and the results plotted.

TgPhyA was assayed using two different TgSkp1 under the same conditions as above in singlicate, due to insufficient materials. The time course curves are shown in Figure 4.2.

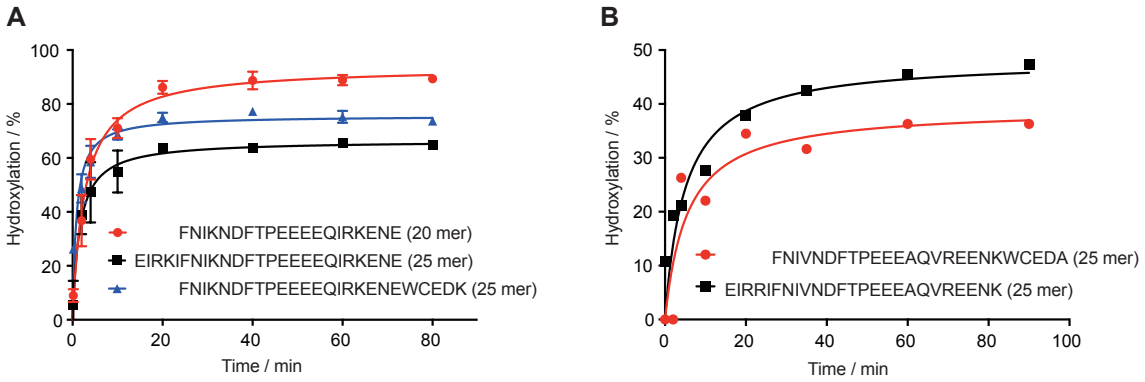


Fig. 4.2 Time course assays of DdSkp1 and TgSkp1 hydroxylation by MALDI-MS. (A) Percentage of hydroxylated DdSkp1 (100 µM) peptide products hydroxylated by DdP4H (2 µM) over time in the presence of Fe(II) (50 µM), 2OG (500 µM), and ascorbate (1 mM) in 100 mM HEPES, pH 7.6 at 37 °C. Time course assays were performed in triplicate. (B) Percentage of TgSkp1 (100 µM) peptide products hydroxylated by TgPhyA (2 µM) over time under the same assay conditions. Time course assays were performed in singlicate.

The efficiency of DdP4H towards hydroxylation of the three different peptides was not significantly different, so FNIKNDFTPEEEEQIRKENEWCEDK (25 mer) was selected for further investigation. FNIVNDFTPEEEAQVREENKWCEDA (25 mer) was selected as the optimal TgSkp1 peptide for further investigation.

4.6 Kinetic Study of DdP4H with DdSkp1 Peptide

An initial kinetic study was performed using DdP4H and its DdSkp1 peptide substrate. To obtain a valid K_m value for substrate, reactions were quenched before substrates were substantially depleted. A time course was performed again with more time points collected at early time to determine the linear range of the reaction. The Fe(II) cofactor and 2OG co-substrate were kept saturating. The time course curve remained linear until 8 minutes. Therefore, subsequent Michaelis-Menten kinetics of the peptide substrate, DdSkp1 were investigated. Assays were prepared using the same condition, while varying the concentration of DdSkp1 peptide. All

assays were initiated by addition of DdP4H and were quenched using 1% (v/v) aqueous formic acid after 8 minutes. The time course curves and Michaelis-Menten curve are shown in Figure 4.3.

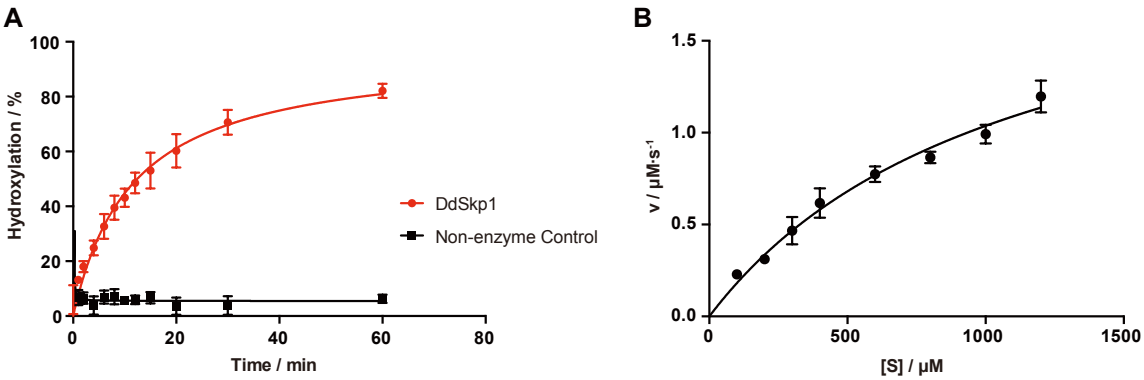


Fig. 4.3 Kinetic study of DdSkp1 catalysed by DdP4H using MALDI-MS. (A) Time course of the hydroxylation of DdSkp1 (100 μM) being hydroxylated by DdP4H (2 μM) in the presence of Fe(II) (50 μM), 2OG (500 μM), and ascorbate (1 mM) in 100 mM HEPES, pH 7.6 at 37 °C. Samples were quenched with 1% (v/v) aqueous formic acid. Time course assays were performed in triplicate. (B) K_m determination for substrate, DdSkp1 peptide. Samples were prepared using the same condition as for time course assays, while varying the concentration of DdSkp1 peptide (100 μM, 200 μM, 300 μM, 400 μM, 600 μM, 800 μM, 1000 μM, 1200 μM). Reactions were all quenched with 1% (v/v) aqueous formic acid at 8 minutes. Assays were performed in triplicate.

The results shown in Figure 4.3 led to the determination of initial K_m and V_{max} values, which were (1099 ± 182) μM and (2.2 ± 0.2) μM·s^{-1}, respectively. Notably, the Michaelis-Menten curve did not reach a plateau even at high substrate concentration. This suggests a high K_m value for the hydroxylation. Under Michaelis-Menten steady-state assumption, in the reaction equation, E + S ⇌ ES → E + P, $K_m = (k_{-1} + k_2) / k_1$. The high value of K_m may result from the low k_1, which possibly reflecting a reduced binding ability of the substrate due to the shorter fragment length of the peptide. Full-length DdSkp1 substrate was then used for kinetic studies.

4.7 LIC on DdSkp1 and TgSkp1

Constructs of full-length DdSkp1 in pET-19b, full-length TgSkp1 in pET-15 and full-length FLAG-TgSkp1 in pET-15 were kindly provided by Ms. Hanke van der Wel (Department Head of Biochemistry & Molecular Biology, University of Georgia). Proteins expressed in these vectors all contained a His_6 tag, which were only cleavable by the protease thrombin. In order to reduce the influence of His_6 tags towards the kinetic studies of Skp1s, and to produce them easily for the lab, Skp1 genes were cloned into pNIC28-Bsa4 plasmid, which introduced TEV protease cleavable tags.

The genes of DdSkp1 and TgSkp1 were cloned into pNIC28-Bsa4 vector using LIC.

The primers designed specifically for DdSkp1 for LIC in this case were:

Forward, 5’ – TACTTCCAATCCATGTCTTTAGTTAAATTAGAATCTT – 3’;

Reverse, 5’ - TATCCACCTTTACTGTTAGTTTCCACCTTTATCTTC – 3’.

The primers for TgSkp1 were:

Forward, 5’ – TACTTCCAATCCATGTCAAAAGAAAGAATGG – 3’;

Reverse, 5’ – TATCCACCTTTACTGTCATGCATCTTCGCAC – 3’.

The primers are designed with the following 5' extensions: TACTTCCAATCCATG for the forward primer and TATCCACCTTTACTG for the reverse primer. A stop codon (TTA for DdSkp1 and TCA for TgSkp1) was introduced after the sequence TATCCACCTTTACTG for the reverse primer depending on the sequence of DdSkp1 or TgSkp1. PCR products were analysed by agarose gel electrophoresis (Fugure 4.4).

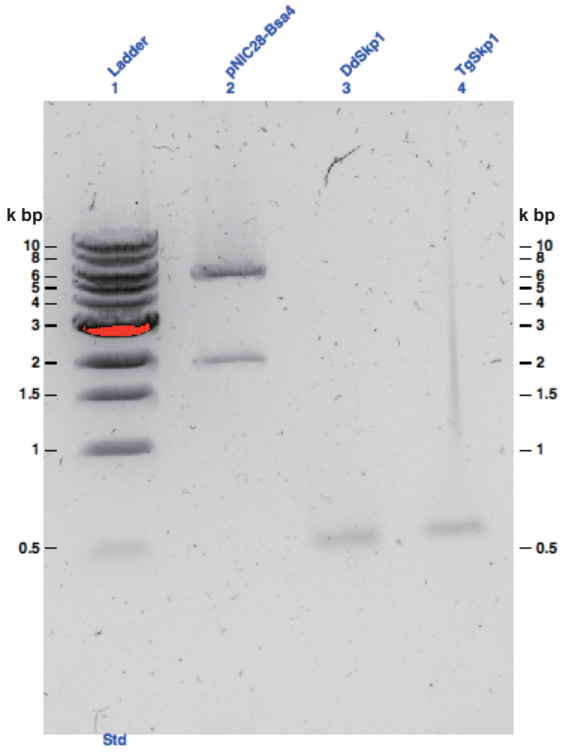


Fig. 4.4 Agarose Gel Electrophoresis results of DdSkp1, TgSkp1, and pNIC28-Bsa4 PCR product. Red area indicates saturated exposure.

The inserts were treated with T4 Polymerase and were annealed into pNIC28-Bsa4 vector. The new constructs were transformed into *E. coli* XL10 Gold competent cells and were further purified for future use.

4.8 Expression Trials of DdSkp1 and TgSkp1

Expression trials indicate DdSkp1 in pNIC28-Bsa4 vector expresses best using 0.2 mM IPTG induced overnight at 28°C as shown in Figure 4.5 A, B and C. The optimal condition for TgSkp1 expression was 0.2 mM IPTG inducing overnight at 28°C as shown in Figure 4.5 D and E.

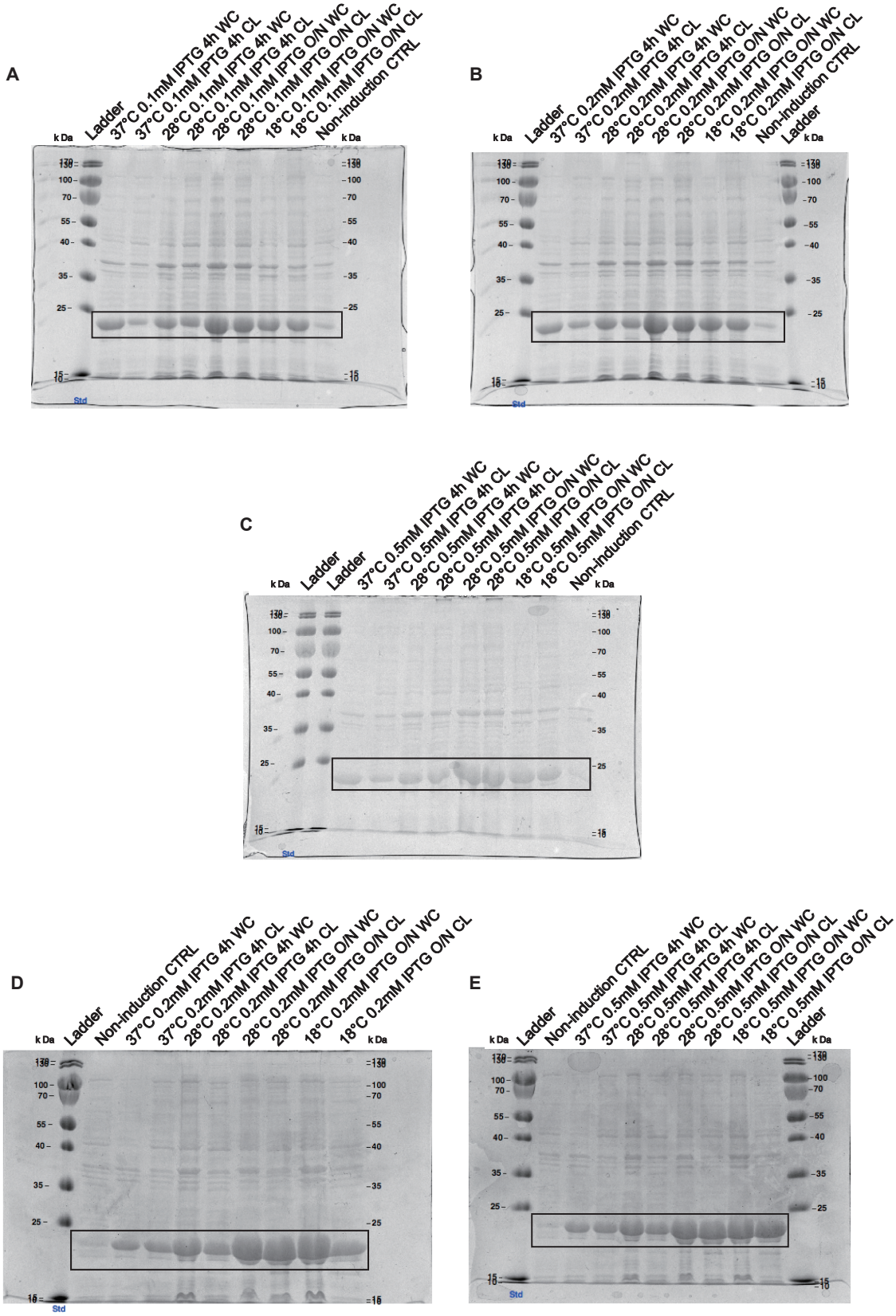


Fig. 4.5 Expression optimisation trials for DdSkp1 in pNIC28-Bsa4 using *E. coli* BL21-Gold (DE3) competent cells (A, B and C) and expression optimisation trials for TgSkp1 in pNIC28-Bsa4 using *E. coli* BL21-Gold (DE3) competent cells (D and E). The conditions are labelled above each lane. The bands corresponding to the target proteins are indicated by boxes.

4.9 Purification of DdSkp1 and TgSkp1

DdSkp1 was then produced on a large scale. BL21-Gold (DE3) *E. coli* competent cells, containing DdSkp1 plasmid were incubated until the optical density (O.D.) reached 1.2 at 600 nm wavelength. Expression was then induced by adding IPTG and incubated overnight at 28°C. A total of 70 g wet cell mass was harvested. His_6-DdSkp1 was purified using Ni(II) affinity chromatography. The purified enzyme was then treated with the tobacco etch virus (TEV) protease (1 mg TEV protease per 50 mg DdSkp1) overnight and re-purified by Ni(II) affinity chromatography to remove the cleaved *N*-terminal tag. The flow-through was buffer exchanged into the storage buffer (100 mM HEPES, pH 7.6) using a PD-10 column (GE Healthcare Life Sciences). The final untagged *apo*-DdSkp1 (60mg) was 95% pure as judged by SDS-PAGE and stored at -80°C for future use.

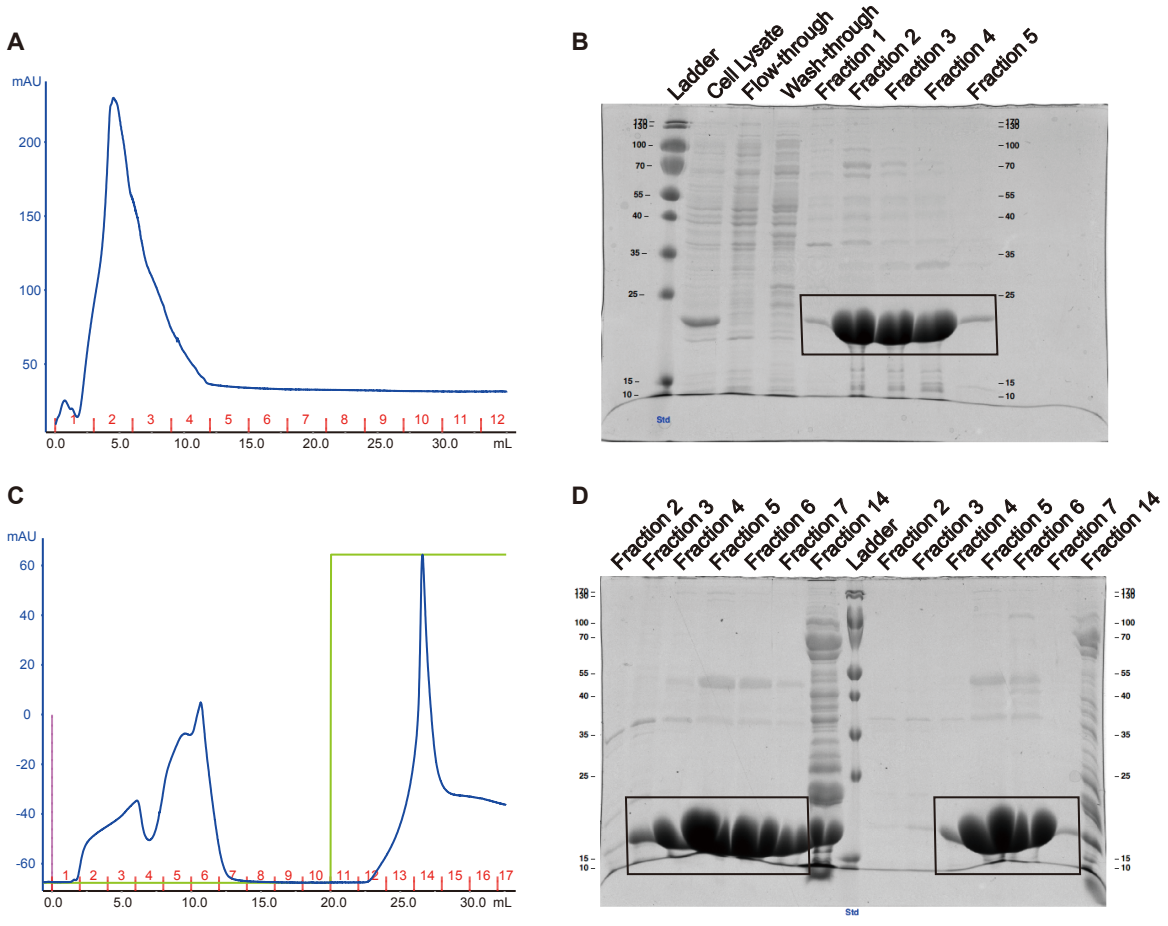


Fig. 4.6 Purification of DdSkp1. (A) Ni(II) affinity chromatography UV chromatogram elution profile. (B) Associated SDS-PAGE gel. The lanes from left to right: ladder (PageRuler from Thermo Scientific), cell lysate, flow-through, wash-through, fractions 1 – 5. The bands of His_6-DdSkp1 are indicated in a box. (C) Ni(II) affinity chromatography UV chromatogram after TEV treated overnight. (D) Associated SDS-PAGE gel. The lanes from left to right: Fractions 2 – 7, and 14, ladder. The bands of untagged DdSkp1 are indicated in a box.

The untagged TgSkp1 (30mg) was expressed and purified using the same method as per DdSkp1

and then stored at -80 °C for use in future hydroxylation assays.

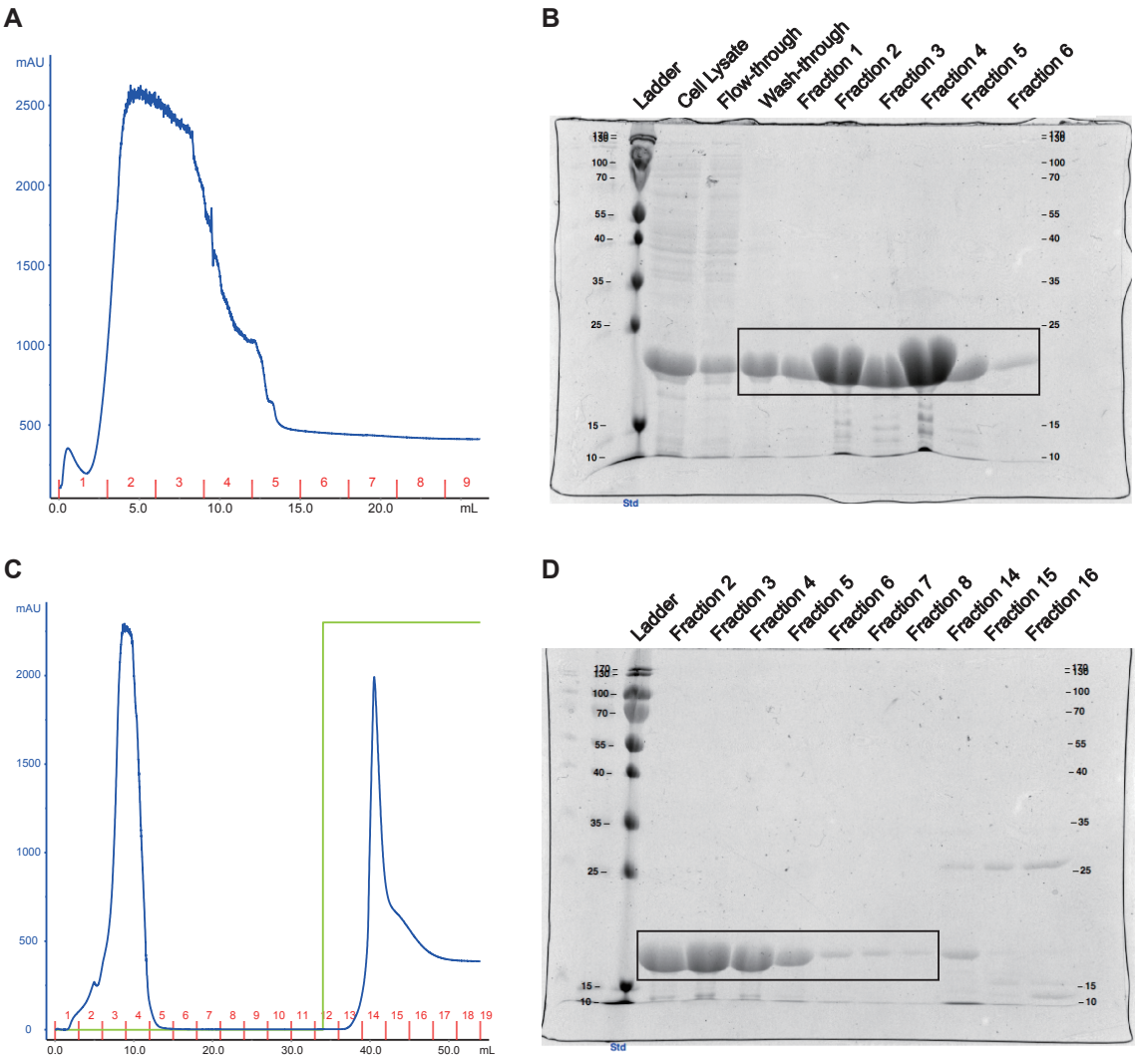


Fig. 4.7 Purification of TgSkp1. (A) Ni(II) affinity chromatography UV chromatogram elution profile. (B) Associated SDS-PAGE gel. The lanes from left to right: Ladder (PageRuler from Thermo Scientific), cell lysate (CL), flow-through, wash-through, fractions 1 – 6. The bands of His_6-TgSkp1 are indicated in a box. (C) Ni(II) affinity chromatography UV chromatogram after

TEV treated overnight. (D) Associated SDS-PAGE gel. The lanes from left to right: Ladder, fractions 2 – 8, and 14 – 16. The bands of untagged TgSkp1 are indicated in a box.

4.10 Purification of HsSkp1 and HsSkp1 E147P

4.10.1 Site-Directed Mutagenesis to Make HsSkp1 E147P Construct from HsSkp1 Plasmid

The construct encoding for HsSkp1 in the pET-21 vector was kindly provided by Dr. Adam Hardy (Department of Chemistry, Chemistry Research Laboratory, University of Oxford, OX1 3TA). The gene of HsSkp1 was connected with a C-terminal His_6-tag. Site-directed mutagenesis was conducted to mutate the Glu147 of HsSkp1 to a proline.

The primers designed specifically for mutating E147P were:

Forward, 5' –CAATATCAAAAATGACTTTACTCCGGAGGAGGAAGCCCAGGTACGC– 3';

Reverse, 5' –GCGTACCTGGGCTTCCTCCTCCGGAGTAAAGTCATTTTTGATATTG– 3'.

The DNA was amplified by polymerase chain reaction (PCR) using a thermal cycler by subjecting to the following condition:

A

PCR Conditions	Volumn Per Tube / µL
5x Q5 Buffer (NEB #B9027S)	5
dNTP (NEB #N0447S)	2.5
Q5 Polymerase (NEB #M0491S)	0.5
Primer Mix (10 µM)	2.5
Template Dna (40 ng – 100 ng)	1
MilliQ Water	13.5
Total	25

B

PCR Steps	Temperature / °C	Incubation Time
1	94	2 min
2	94	30 s
3	55	1 min
4	68	12 min
5	Repeat Step 2 – 4 for 18 cycles	
6	68	20 min
7	4	Pause

Table 4.3 Procedures for site-directed mutagenesis to genrate HsSkp1 E147P. (A) Sample ingredient for PCR. (B) PCR conditions.

The samples were then treated with 1 µL DpnI (New England BioLabs #R0176S) for 3 hours at 37 °C to digest the non-mutated plasmid. The digested products were then transformed into XL10 Gold *E. coli* competent cells cultured and were further purified and stored for future use.

4.10.2 Expression Trials of HsSkp1

Expression trials with HsSkp1 were done using the pET-21 construct in BL21 Gold (DE3) *E. coli* competent cells. Different induction times, IPTG concentrations, and induction temperatures were tested. The *E. coli* pellets were then sonicated. The various induction conditions were analysed using SDS-PAGE electrophoresis (Figure 4.8).

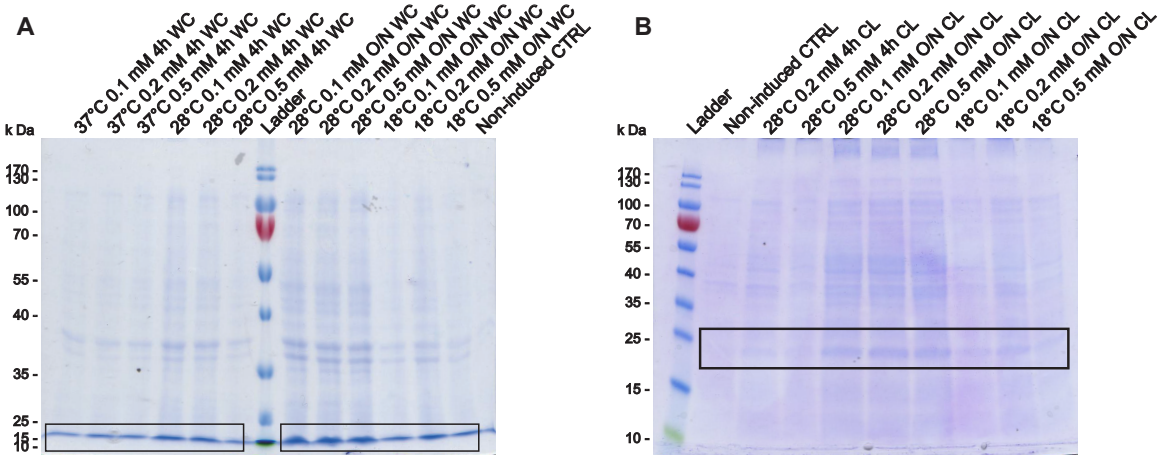


Fig. 4.8 SDS-PAGE analysis of expression trials for HsSkp1 in pET-21 with BL21 Gold (DE3) competent cells. (A) Whole cell (WC) gel. The expression conditions from left to right were: 37 °C, 0.1 mM IPTG, induced 4 hours; 37 °C, 0.2 mM IPTG, induced 4 hours; 37 °C, 0.5 mM IPTG, induced for 4 hours; 28 °C, 0.1 mM IPTG, induced for 4 hours; 28 °C, 0.2 mM IPTG, induced 4 hours; 28 °C, 0.5 mM IPTG, induced for 4 hours; ladder (PageRuler from Thermo Scientific); 28 °C, 0.1 mM IPTG, induced overnight (O/N); 28 °C, 0.2 mM IPTG, induced overnight; 28 °C, 0.5 mM IPTG, induced overnight; 18 °C, 0.1 mM IPTG, induced overnight; 18 °C, 0.2 mM IPTG, induced overnight; 18 °C, 0.5 mM IPTG, induced overnight and non-induced control (CTRL). (B) Cell lysate (CL) gel. The expression conditions from left to right were: Ladder (PageRuler from Thermo Scientific); non-induced control (CTRL); 28 °C, 0.2 mM IPTG, induced 4 hours; 28 °C, 0.5 mM IPTG, induced 4 hours; ladder (PageRuler from Thermo Scientific); 28 °C, 0.1 mM IPTG, induced overnight (O/N); 28 °C, 0.2 mM IPTG, induced overnight; 28 °C, 0.5 mM IPTG, induced overnight; 18 °C, 0.1 mM IPTG, induced overnight; 18 °C, 0.2 mM IPTG, induced overnight; 18 °C, 0.5 mM IPTG, induced overnight. The bands of HsSkp1 are indicated in boxes.

SDS-PAGE gels showed certain levels of soluble HsSkp1 recombinant protein were present in the cell lysates. The most intense band of HsSkp1 was apparent when using 0.1 mM IPTG induced at 28 °C overnight growth. This condition was therefore selected to express HsSkp1 for further protein production and purification.

4.10.3 Expression and Purification of HsSkp1 and HsSkp1 E147P

The constructs of HsSkp1 and of HsSkp1 E147P were transformed into BL21 Gold (DE3) *E. coli* competent cells and the recombinant proteins were produced on a large scales using the optimised conditions identified above.

In brief, BL21 Gold (DE3) cells, containing HsSkp1 plasmid were incubated until the optical density (O.D.) reached 1.2 (600 nm wavelength). Cells were induced by addition of IPTG and incubated overnight at 18 °C. A 30 g wet cell mass was harvested by centrifugation. HsSkp1-His_6 was purified using Ni(II) affinity chromatography (Figure 4.9a). The sample was then concentrated to 15 mg/mL and was buffer exchanged to the final storage buffer (100mM HEPES, pH 7.6). The final yield of the purification was ~220 mg from 6 L of cell culture. This protein was then used in hydroxylation assays.

HsSkp1 E147P-His_6 was purified using the same strategy. The final yield of the purification was ~45 mg from 6 L of cell culture.

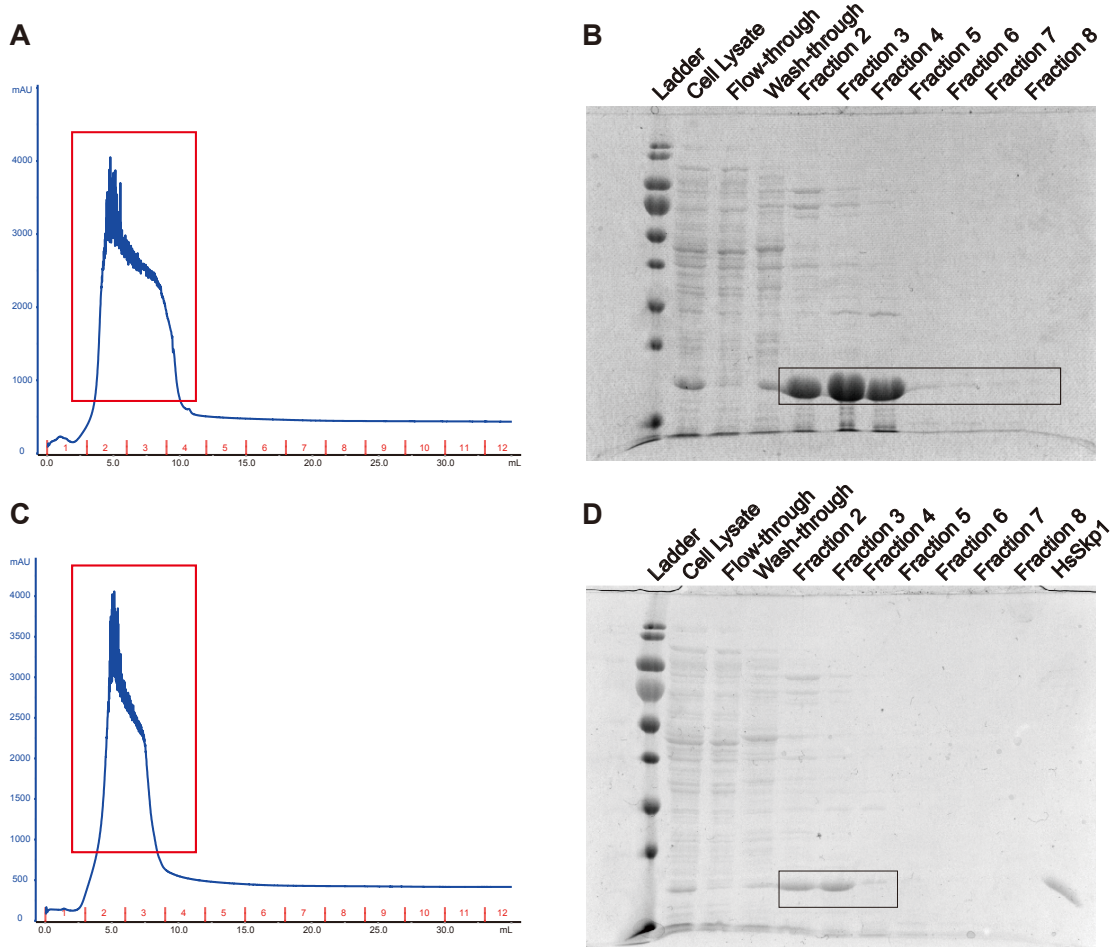


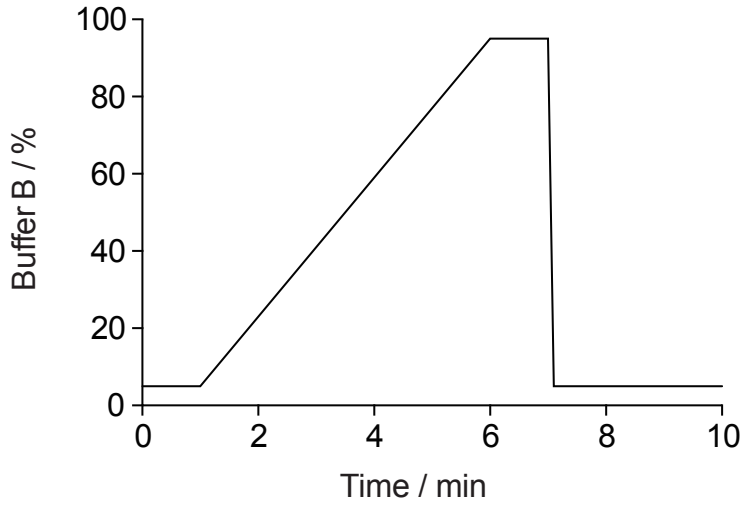
Fig. 4.9 Purification of HsSkp1-His$_6$ and HsSkp1 E147P-His$_6$. (A) Ni(II) affinity chromatography UV chromatogram elution profile for the purification of HsSkp1-His$_6$. (B) SDS-PAGE analysis. The lanes from left to right are: Ladder (PageRuler from Thermo Scientific), cell lysate, flow-through, wash-through, fractions 2 – 8. The bands corresponding to HsSkp1-His$_6$ are indicated in a box. (C) Ni(II) affinity chromatography UV chromatogram for the purification of HsSkp1 E147P-His$_6$. (D) SDS-PAGE analysis. The lanes from left to right are: Ladder (PageRuler from Thermo Scientific), cell lysate, flow-through, wash-through, fractions 2 – 8, HsSkp1-His$_6$ sample. The bands corresponding to HsSkp1 E147P-His$_6$ are indicated in a box.

4.11 Mass Spectrometry Based Hydroxylation Assays

The hydroxylation samples described in this chapter were all analysed using a Xevo G2-S QToF mass spectrometer (Waters) equipped with an electrospray source and coupled to an Acquity (Waters) ultra performance liquid chromatography instrument unless otherwise specified. Instrument control and data processing were performed using MassLynx 4.1 software

(Waters, Wilmslow UK). The column, Aeris^{TM} 3.6 µm WIDEPORE C4 200 Å (Phenomenex), was used for separation. Solvents and gradient protocol are shown below:

A



B

Time (min)	Flow (mL / min)	A (%)	B (%)
0.0	0.3	95	5
1.0	0.3	95	5
6.0	0.3	5	95
7.0	0.3	5	95
7.1	0.3	95	5
10.0	0.3	95	5

Fig. 4.10 Xevo G2-S QToF mass spectrometer solvent chromatography. (A) LC-MS reversed-phase chromatography was used, starting with high percentage of aqueous solution (95 % v/v buffer A), eluting with a gradient of organic solvent until 95 % (v/v buffer B) for 10 minutes. Buffer A: water + 0.1 % (v/v) aqueous formic acid. Buffer B: acetonitrile + 0.1 % (v/v) aqueous formic acid. (B) Chromatography protocol shown in tabulated form.

4.12 Hydroxylation Assays of DdP4H and TgPhyA with Full-Length DdSkp1 and TgSkp1 Proteins

Hydroxylation assays using full-length DdSkp1 protein substrate and DdP4H enzyme were performed and analysed using the LC-MS (Xevo G2-S) machine. In the presence of DdP4H, Fe (II), 2OG, ascorbate and oxygen, full-length DdSkp1 protein was analysed for hydroxylation after incubation at 37 °C for 1 hour, the reaction was quenched with 1 % (v/v) aqueous formic acid. The hydroxylation assays of full-length TgSkp1 protein substrate with TgPhyA enzyme used the same conditions.

Assay Conditions	Final Concentration	Volume
2-oxoglutarate disodium salt (Sigma-Aldrich, K3752)	500 µM	2 µL
Sodium L-ascorbate (Sigma-Aldrich, 11140)	1 mM	2 µL
Ammonium iron(II) sulfate hexahydrate (Sigma-Aldrich, 215406)	50 µM	2 µL
Full-length DdSkp1 or TgSkp1 protein substrate	100 µM	2 µL
DdP4H or TgPhyA enzyme	1 µM	2 µL
Buffer (100 mM HEPES, pH 7.6)	-	10 µL
Total volume	-	20 µL

Table 4.4 Assay conditions of hydroxylation reaction mixtures.

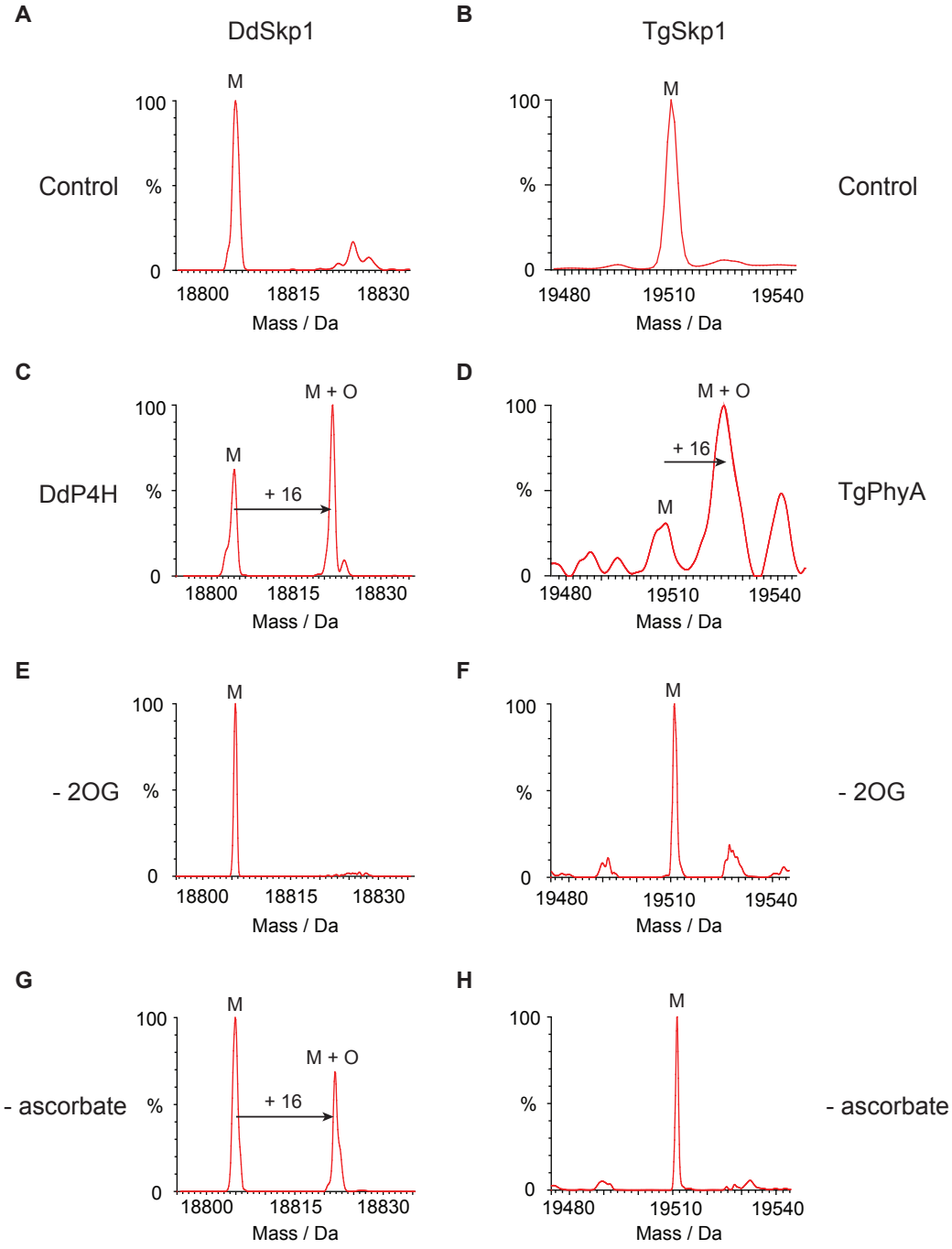


Fig. 4.11 Mass spectra of hydroxylation assays analysed by LC-MS. (A) No enzyme control of full-length DdSkp1 protein (100 µM) in the presence of Fe(II) (50 µM), 2OG (500 µM), and ascorbate (1 mM) in buffer (100 mM HEPES, pH 7.6) incubated at 37 °C for 1 hour. (B) No enzyme control of full-length TgSkp1 protein (100 µM). (C) DdSkp1 protein (100 µM) underwent a + [O] modification in the presence of 1 µM DdP4H incubated with the same condition as above. (D) TgSkp1 protein (100 µM) underwent a + [O] modification in the presence of 1 µM TgPhyA incubated with the same condition as above. (E) Hydroxylation sample of DdSkp1 with DdP4H in the absence of 2OG. (F) TgSkp1 with TgPhyA in the absence of 2OG. (G) DdSkp1 with DdP4H in

the absence of ascorbate. (H) TgSkp1 with TgPhyA in the absence of ascorbate. All reactions were all quenched with 1% (v/v) aqueous formic acid before analysed by LC-MS.

Compared to the no enzyme control, the mass of full-length DdSkp1 protein exhibits a + [O] mass shift with the ingredient of 2OG, ascorbate, Fe(II), and the enzyme, DdP4H. TgSkp1 was hydroxylated by TgPhyA under the same conditions. Both 2OG and ascorbate promoted the catalytic activity.

4.12.1 Truncated DdP4H, DdP4H (60-284) Hydroxylated Full-Length DdSkp1 Protein

The hydroxylation of full-length DdSkp1 protein by DdP4H (60-284) was performed using LC-MS (Xevo G2-S) with the same condition as in Chapter 4.11. The results (Figure 4.12) showed that DdP4H (60-284) was active and was able to hydroxylate full-length DdSkp1 protein. DdP4H (60-284) was used for crystallisation, as described in Chapter 3.

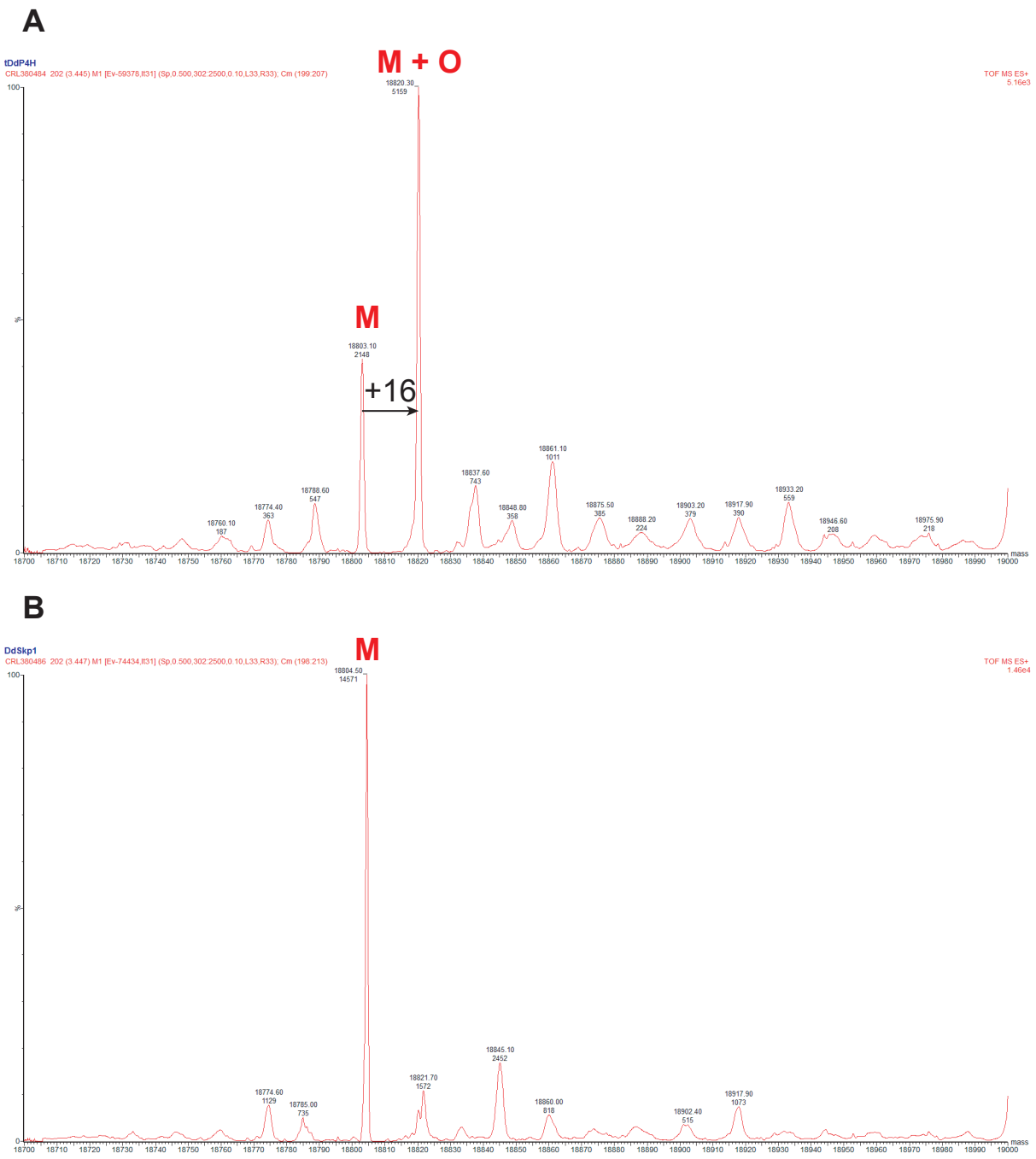


Fig. 4.12 Mass spectra of DdSkp1 hydroxylated by DdP4H (60-284) analysed by LC-MS. (A) Hydroxylated full-length DdSkp1 protein (100 µM) with DdP4H (60-284) (1 µM) in the presence of Fe(II) (50 µM), 2OG (500 µM), and ascorbate (1 mM) in 100 mM HEPES, pH 7.6, incubated at 37 °C for 1 hour. (B) No enzyme control of full-length DdSkp1 protein (100 µM) incubated under the same conditions as above.

4.13 TgSkp1 Exhibits Non-Specific Oxidation

During the hydroxylation assay experiments of TgSkp1 with TgPhyA, a series of + 16 Da peaks were observed after long-time scale incubation at 37 °C (> 30 min), shown in Figure 4.13 A and C. The same observation existed in no enzyme control samples. After incubating for more than 30 minutes in the presence of 2OG, ascorbate and Fe(II), TgSkp1 was hydroxylated even in the absence of the enzyme, TgPhyA (Figure 4.13 B and D). It was hypothesised that non-specific oxidation (possibly methionine oxidation) occurred due to the oxidising conditions of the enzyme; Fe(II), ascorbate, 2OG and oxygen are also present and can potentially enable oxidising reactions.

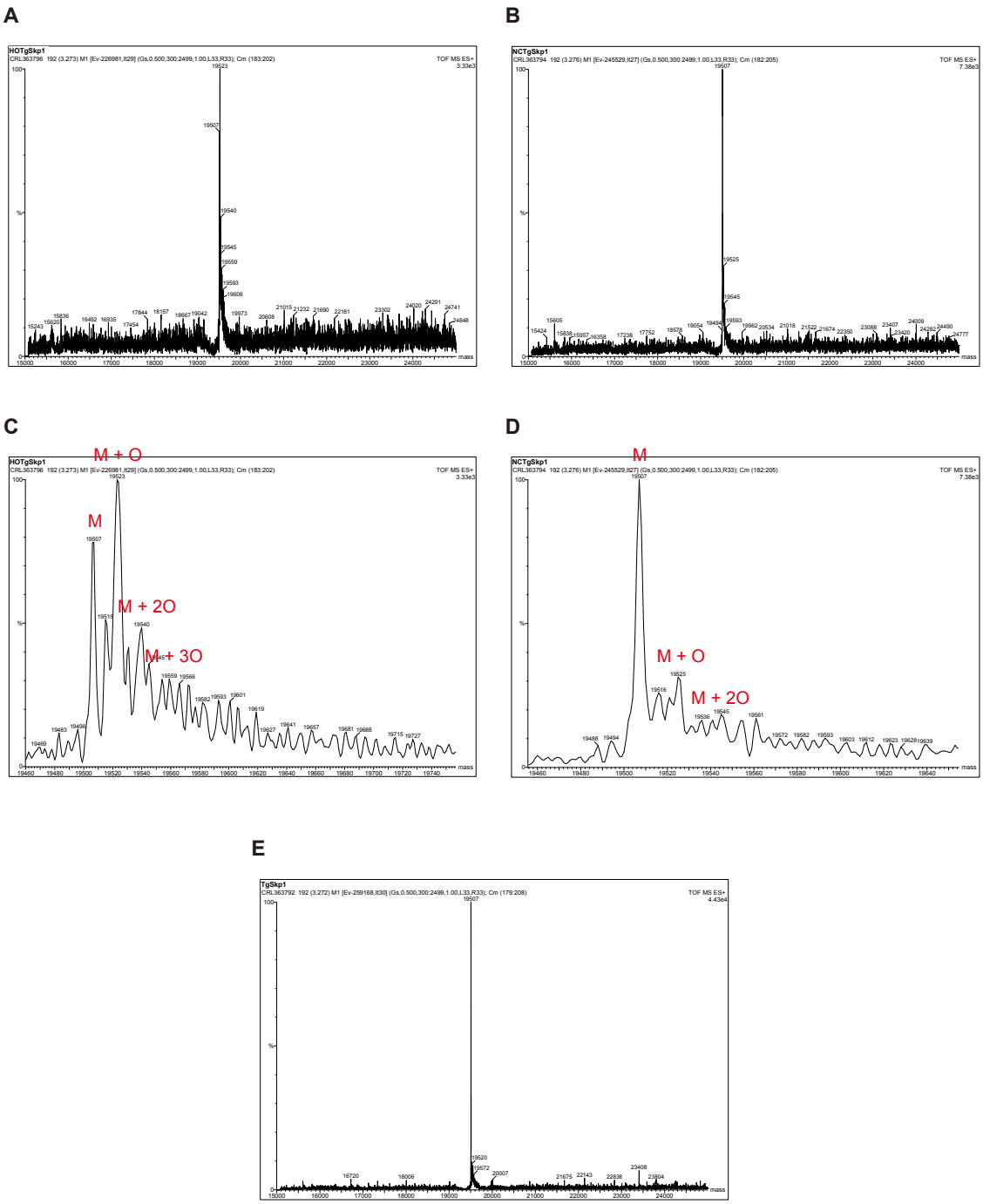


Fig. 4.13 Mass spectra of TgSkp1 analysed by LC-MS. (A) Hydroxylated full-length TgSkp1 protein (100 µM) with TgPhyA (1 µM) in the presence of Fe(II) (50 µM), 2OG (500 µM), and ascorbate (1 mM) in 100 mM HEPES, pH 7.6, incubated at 37 °C for 1 hour. Sample name: HOTgSkp1. (B) No enzyme control of full-length TgSkp1 protein (100 µM) incubated under the same condition as above. Sample name: NCTgSkp1. (C) Zoomed view of the spectrum HOTgSkp1. (D) Zoomed view of the spectrum NCTgSkp1. The peaks of unhydroxylated TgSkp1 and the following series peaks of hydroxylated TgSkp1 were all labelled in red. (E) Spectrum of purified TgSkp1 in buffer, 100 mM HEPES, pH 7.6.

To reduce formation of potential peroxide intermediates, which might cause the non-specific oxidation in solution, freshly made reducing reagents were added in the no enzyme reactions, including 1,4-dithiothreitol (DTT) and tris (2-carboxyethyl) phosphine (TCEP). The LC-MS spectra are shown in Figure 4.14.

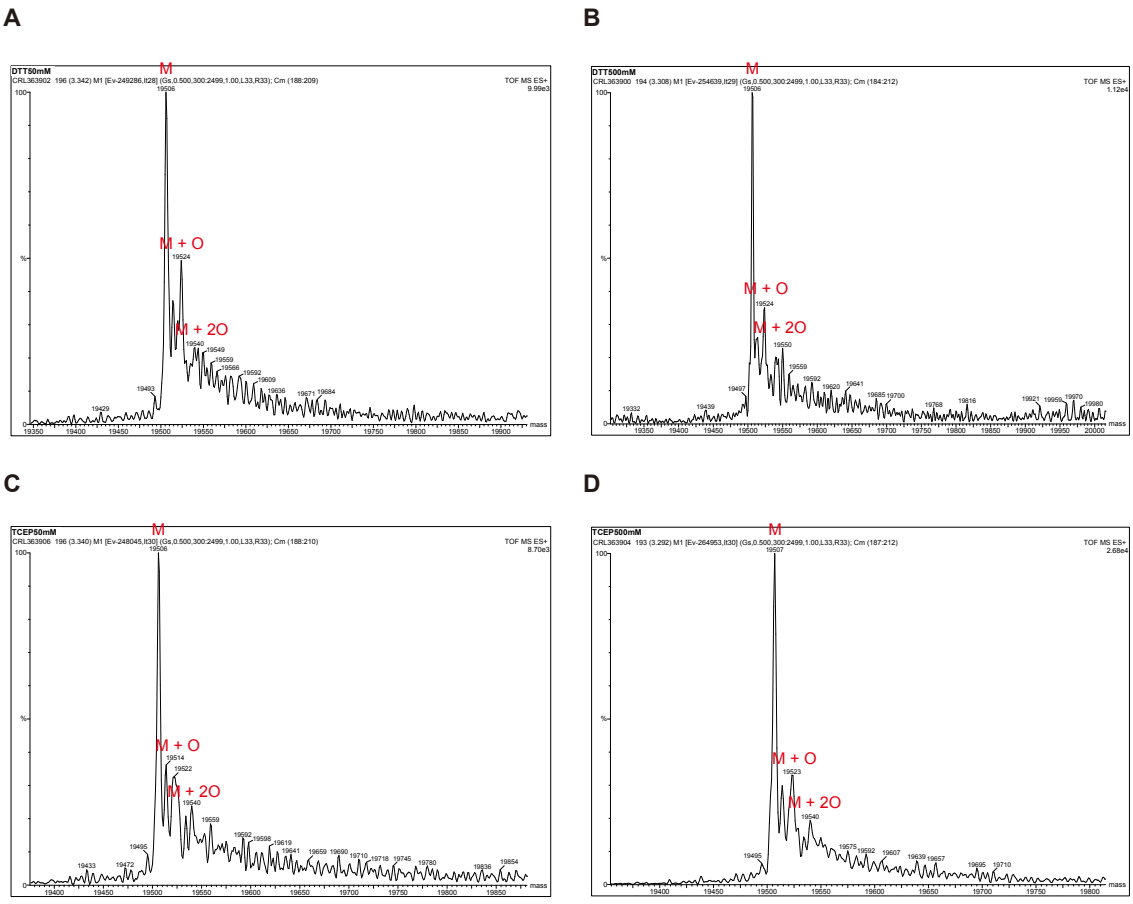


Fig. 4.14 Mass spectra of TgSkp1 analysed by LC-MS. (A) No enzyme control of full-length TgSkp1 protein (100 µM) with DTT (50 mM). Sample name: DTT50mM. (B) TgSkp1 with DTT (500 mM). Sample name: DTT500mM. (C) TgSkp1 with TCEP (50 mM). Sample name: TCEP50mM. (D) TgSkp1 with TCEP (500 mM). Sample name: TCEP500mM. Samples were all prepared fresh and were incubated in the presence of Fe(II) (50 µM), 2OG (500 µM), and ascorbate (1 mM) in 100 mM HEPES, pH 7.6, incubated at 37 °C for 1 hour before being quenched with 1% (v/v) aqueous formic acid.

[M + O] peaks existed in all no enzyme control samples even when reducing reagents were added to the solution, indicating the effect of non-specific oxidation could not be removed by their presence.

Although oxidation could be observed in the no enzyme control samples, the [M + O] peak (30%) was much lower compared to the samples in the presence of TgPhyA (60%). Therefore, it could be confirmed that TgPhyA catalysed hydroxylation of TgSkp1 occurs. Subsequent MS/MS results support this conclusion, as described in Chapter 4.24.4.

4.14 The Non-Specific Oxidation of TgSkp1 Does Not Affect Kinetic Studies

With the existence of non-specific oxidation, kinetic studies using full-length TgSkp1 protein were complicated. It was impossible to separate the TgPhyA catalysed prolyl hydroxylation from the non-specific oxidation from the MS spectra. The [M + O] peak contained oxidised TgSkp1 caused by both hydroxylation and non-specific oxidation. If the peak height was used directly for kinetic studies, the apparent value would result in a higher reaction velocity and a lower K_m value. Therefore, further investigation on the non-specific oxidation was required. A time course of non-specific oxidation was thus performed (Figure 4.15).

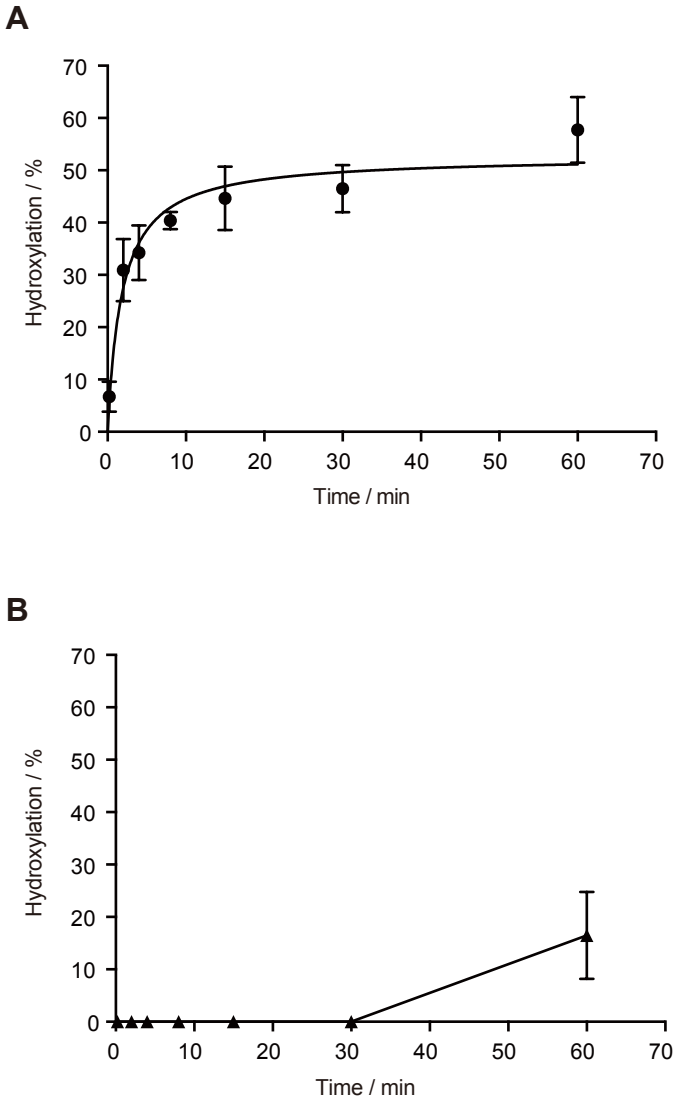


Fig. 4.15 Time course of the hydroyxation of TgSkp1 analysed by LC-MS. (A) Time course of the hydroxylation of TgSkp1 (100 µM) being catalysed by TgPhyA (1 µM) in the presence of Fe(II) (50 µM), 2OG (500 µM), and ascorbate (1 mM) in 100 mM HEPES, pH 7.6 at 37 °C. Samples were quenched with 1% (v/v) aqueous formic acid. Time course assays were performed in triplicate. (B) Time course of the hydroxylation of no enzyme control samples of TgSkp1 (100 µM) under the same condition as above. Time course assays were performed in triplicate.

It can be observed that in no enzyme controls, oxidation only appears after 30 minutes from the initiation of the reaction. For kinetic studies, hydroxylation reactions had to be quenched before substrates were depleted, within the linear range of the time course. The time course of the hydroxylation of TgSkp1 with TgPhyA showed the curve remained linear to 2 minutes,

considerably shorter than 30 minutes, when non-specific oxidation appeared. Therefore, using full-length TgSkp1 protein and TgPhyA to perform kinetic studies was assessed to be feasible.

4.15 Kinetic Studies of DdP4H Using Full-Length DdSkp1 Protein Substrate

Kinetic studies were performed with DdP4H and its full-length protein substrate, DdSkp1. A time course assay was analysed to determine the linear range of the curve. The Fe(II) cofactor and 2OG co-substrate were kept above saturation [50 µM ammonium iron(II) sulfate hexahydrate (Sigma-Aldrich), 500 µM 2OG (2-oxoglutaric acid disodium salt from Sigma-Aldrich)]. The time course curve remained linear to 4 minutes. Therefore, subsequent Michaelis-Menten kinetics of the protein substrate, DdSkp1 were tested. Different samples were prepared under the same condition, except the concentration of DdSkp1. All assays were initiated at the same time by adding DdP4H into the solution and were quenched after 4 minutes. The hydroxylation samples were analysed using a Xevo G2-S QToF LC-MS, the method of which was discussed in Chapter 4.10. Time course curves and Michaelis-Menten curve are shown in Figure 4.16.

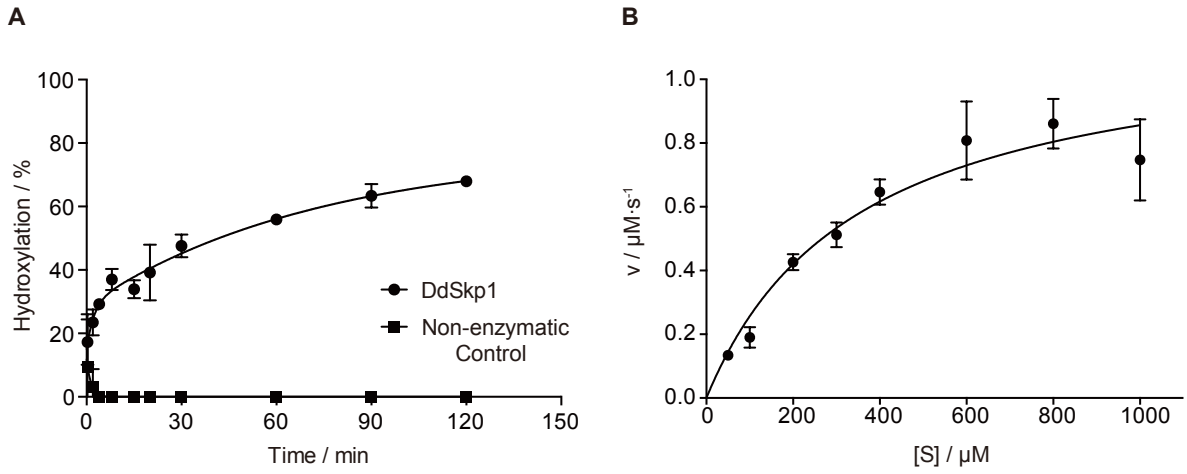


Fig. 4.16 Kinetic study of full-length DdSkp1 protein by LC-MS. (A) Time course of full-length DdSkp1 protein (100 µM) being hydroxylated by DdP4H (1 µM) in the presence of Fe(II) (50 µM), 2OG (500 µM), and ascorbate (1 mM) in 100 mM HEPES, pH 7.6 at 37 °C. Samples were quenched with 1% (v/v) aqueous formic acid. Time course assays were performed in triplicate. (B) K_m for substrate, full-length DdSkp1 protein. Samples were prepared with the same

conditions as for time course assays, excepting various concentration of DdSkp1 protein (50 µM, 100 µM, 200 µM, 300 µM, 400 µM, 600 µM, 800 µM, 1000 µM). Reactions were all quenched with 1% (v/v) aqueous formic acid 4 minutes after initiation. Assays were performed in triplicate.

Subsequent studies led to the determination of initial K_m and V_{max} values, which were (349.9 ± 77.3) µM and (1.2 ± 0.1) µM·s^{-1}, respectively.

4.16 Comparison of DdP4H Kinetics Using DdSkp1 Peptide and Full-Length Protein Substrates

Kinetic studies were initially performed using DdSkp1 peptide (residue 136-160) (Chapter 4.6). However, using the peptide substrate, the Michaelis-Menten curve did not reach a plateau (saturation) even at high substrate concentrations, so that a K_m value could not be accurately determined. By using the full-length DdSkp1 protein as a DdP4H substrate, it was possible to determine a more valid K_m value, reflecting that full-length DdSkp1 protein was a better substrate than the 25-mer DdSkp1 peptide tested. This observation may be because other regions other than the active site of DdSkp1 protein take part in substrate binding.

4.17 DdP4H Kinetic Parameters for 2OG

Kinetic studies defining the parameters for 2OG were carried out using the full-length protein substrate, DdSkp1. A time course was carried out to determine the linear range of the curve. The Fe(II) cofactor was kept above saturation [50 µM ammonium iron(II) sulfate hexahydrate from Sigma-Aldrich]. Full-length DdSkp1 protein substrate was kept at 600 µM for all samples. The time course curve remained linear for the first 15 minutes. Subsequently, Michaelis-Menten kinetics of 2OG were tested using different samples prepared with the same conditions, varying the concentration of 2OG. All assays were initiated at the same time by the addition of DdP4H and were quenched after 6 minutes using 1% (v/v) aqueous formic acid. Samples were analysed

using the Xevo G2-S QToF LC-MS. Time course curves and Michaelis-Menten curve are shown in Figure 4.17.

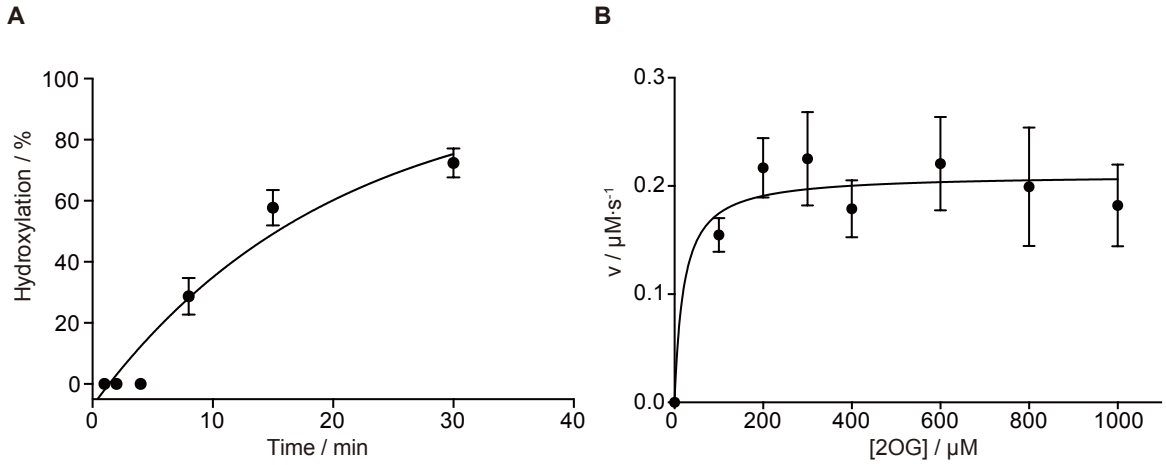


Fig. 4.17 DdP4H K_m determination for 2OG by LC-MS. (A) Time course of the hydroxylation of full-length DdSkp1 protein (600 µM) being hydroxylated by DdP4H (1 µM) in the presence of Fe(II) (50 µM), 2OG (200 µM), and ascorbate (1 mM) in 100 mM HEPES, pH 7.6 at 37 °C. Samples were quenched with 1% (v/v) aqueous formic acid. Time course assays were performed in triplicate. (B) Apparent K_m for 2OG with DdP4H. Samples were prepared with the same conditions as for time course assays, excepting various concentration of 2OG (100 µM, 200 µM, 300 µM, 400 µM, 600 µM, 800 µM, 1000 µM). Reactions were all quenched with 1% (v/v) aqueous formic acid 6 minutes after initiation. Assays were performed in triplicate. Samples were analysed using the Xevo G2-S QToF LC-MS.

Subsequent studies led to the determination of initial K_m and V_{max} values for 2OG, which were (20.6 ± 20.0) µM and (0.21 ± 0.02) µM·s^{-1}, respectively. Xevo G2-S QToF LC-MS determined data had large errors to the values. The apparent V_{max} value was significantly lower than the V_{max} value for substrate (Figure 4.16). This may be due to the low detection accuracy of the LC-MS, when high concentration of DdSkp1 substrate had to be used. In addition, the Michaelis-Menten curve lacks data points in the accelerating range, which is beyond the limits of detection for the instrument. Therefore, a different instrument, RapidFire Mass Spectrometry (RF-MS), was pursued for determination of the K_m value of 2OG.

4.18 RapidFire Mass Spectrometry (RF-MS) Hydroxylation Assay Development

All RapidFire Mass Spectrometry (RF-MS) kinetic studies were carried out with the help of Dr. Anthony Tumber (Department of Chemistry, Chemistry Research Laboratory, University of Oxford, OX1 3TA).

RF-MS is a high throughput MS based robotic platform that incorporates solid phase extraction technology to remove non-volatile buffer components from assay samples [1]. Samples are aspirated under vacuum and loaded onto a C4 solid phase extraction (SPE) cartridge. The cartridge is then washed with an aqueous solvent, typically 0.1% (v/v) aqueous formic acid in LC-MS grade water and then analytes of interest eluted from the SPE with 85% acetonitrile containing 0.1% (v/v) aqueous formic acid onto an Agilent 6530 accurate mass QToF. The SPE cartridge is then re-equilibrated with 0.1% (v/v) aqueous formic acid in water. The whole cycle of loading, wash, elution and re-equilibration takes typically 7 – 12 seconds, enabling a 384-well plate to be sampled in 1.5 – 2.0 hours.

All assays were performed in 384-well polypropylene v-bottom plates (Greiner Bio-One). All assay steps were performed in assay buffer (50 mM Tris.Cl, 50 mM NaCl, pH 7.5). Ferrous ammonium sulphate (FAS) was from Sigma Aldrich (Code 12304) and was prepared fresh each day by dissolving 50 – 100 mg in 10 mM HCl to a concentration of 400 mM and then further diluted to 1 mM in deionized water. 2-oxoglutarate (2OG) was from Sigma Aldrich (K3752) and was prepared fresh each day by dissolving 1.0 – 2.0 mg in deionized water to a concentration of 10 mM. L-ascorbic acid (L-AA) was from Sigma Aldrich (Code A5960) and was prepared fresh each day by dissolving 8 – 10 mg in deionized water to a concentration of 50 mM. Reagent transfers (enzyme in assay buffer and substrate in assay buffer) were performed using a

multidrop combi dispenser (Thermo Scientific, Code 5840300) to which was attached a dispensing cassette (Thermo Scientific, Code 24073290).

Plates were then transferred to a RapidFire RF360 high throughput sampling robot (Agilent) connected to an Agilent accurate mass 6530 QTOF. Samples were aspirated under vacuum for 500 ms and loaded onto a C4 SPE cartridge. The cartridge was then washed for 5000 ms with 0.1% (v/v) aqueous formic acid in water at a flow rate of 1.5 mL/min after which the protein was eluted from the SPE cartridge with 85% acetonitrile in water + 0.1% (v/v) aqueous formic acid for 5000 ms at a flow rate of 1.25 mL/min. The SPE cartridge was then re-equilibrated in 0.1% (v/v) aqueous formic acid in water for 800 ms at a flow rate of 1.5 mL/min. Skp1s (DdSkp1 and TgSkp1) were detected in positive ESI mode with a fragmentor voltage of 200 V. Extracted ion chromatogram data for the protein were integrated and % conversion of non-hydroxylated Skp1 to hydroxylated Skp1 was calculated using the equation:

$$\% \text{ Conversion} = 100 \times \text{Hydroxylated} / (\text{Hydroxylated} + \text{Non-Hydroxylated})$$

Data were determined from nonlinear regression curve fit using GraphPad Prism 5.

4.19 Kinetic Study of DdP4H with 2OG Using RF-MS

Samples were prepared using the same conditions as Chapter 4.16 varying the concentration of 2OG, quenched after 6 minutes with 1% (v/v) aqueous formic acid, before being analysed by RF-MS. Michaelis-Menten curve was shown in Figure 4.18.

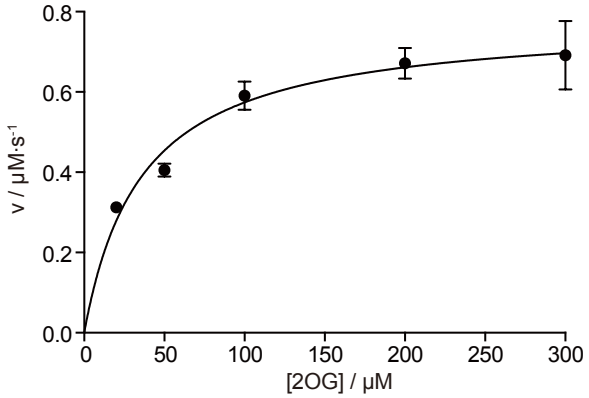


Fig. 4.18 Kinetic study of DdP4H with 2OG using RF-MS. Apparent K_m for 2OG with DdP4H. Samples were prepared with the same conditions as for time course assays, varying concentration of 2OG (20 µM, 50 µM, 100 µM, 200 µM, 300 µM). Reactions were all quenched with 1% (v/v) aqueous formic acid 6 minutes after initiation. Assays were performed in triplicate. Samples were analysed using the RF-MS.

The apparent K_m and V_{max} values for 2OG were (35.9 ± 5.9) µM and (0.78 ± 0.03) µM·s^{-1}, respectively. The values were consistent with the results from the Xevo G2-S QToF LC-MS, but more accurate.

4.20 Kinetic Studies of DdP4H with Oxygen

In order to investigate whether DdP4H can play a role in oxygen sensing, kinetic studies of oxygen with DdP4H were performed. Samples were prepared in sealed cone bottomed MS vials (Macherey-Nagel, REF: 702006). Full-length DdSkp1 protein, final concentration of which was described in the table below, was incubated for 10 minutes at different ratios of O_2/N_2 gas constantly being blown into the vials at 37 °C. Then 2 µL of ascorbate + 2OG mix was added, before 2 µL of Fe(II) being added. Lastly, 2 µL of 1 mM DdP4H enzyme was added marking the reaction initiation. Reactions were carried out in a 37 °C water bath, before being quenched with 1% (v/v) aqueous formic acid at the 8.25-minute time point, determined by time course

assay (Figure 4.19 A). Samples were analysed using the Xevo G2-S QToF LC-MS. The result is shown in Figure 4.19 B.

Assay Conditions	Final Concentration	Volume
2-oxoglutarate disodium salt (Sigma-Aldrich, K3752)	500 µM	2 µL
Sodium L-ascorbate (Sigma-Aldrich, 11140)	1 mM	
Ammonium Iron(II) sulfate hexahydrate (Sigma-Aldrich, 215406)	50 µM	2 µL
Full-length DdSkp1 protein substrate	600 µM	94 µL
DdP4H enzyme	1 µM	2 µL
Buffer (100 mM HEPES, pH 7.6)	-	0
Total volume	-	100 µL

Table 4.5 Assay conditions of hydroxylation assay.

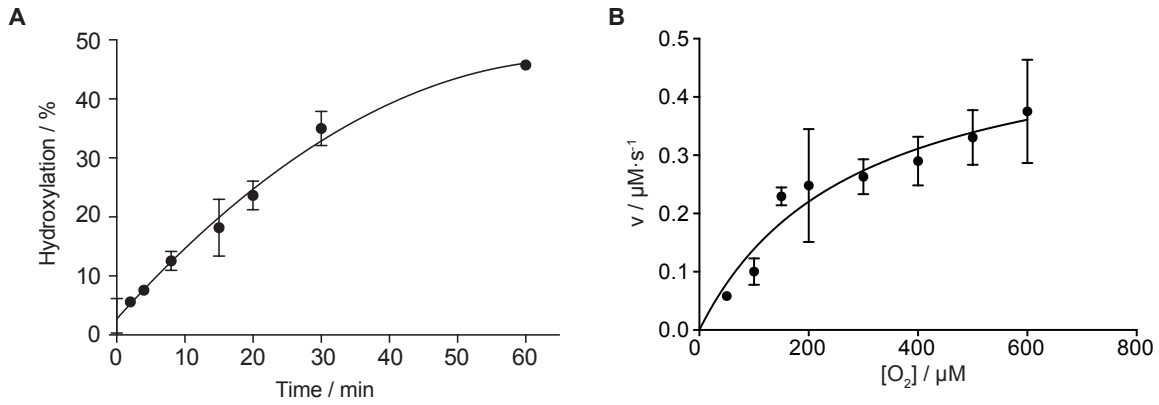


Fig. 4.19 Kinetic study of DdP4H with oxygen using LC-MS. (A) Time course of the hydroxylation of full-length DdSkp1 protein (600 µM) being hydroxylated by DdP4H (1 µM) in the presence of Fe(II) (50 µM), 2OG (500 µM), and ascorbate (1 mM) in 100 mM HEPES, pH 7.6 at 37 °C. Samples were quenched with 1% (v/v) aqueous formic acid. Time course assays were performed in triplicate. (B) Apparent K_m for oxygen with DdP4H and DdSkp1. Samples were prepared under the same condition as described in the table. Different percentage of oxygen was used to incubate the samples: 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%. Reactions were all quenched with 1% (v/v) aqueous formic acid 8.25 minutes after initiation. Assays were performed in triplicate. Samples were analysed using the Xevo G2-S QToF LC-MS.

The rate of reaction (V) was significantly lower than the V_{max} value for substrate (Figure 4.16). This may be due to the low detection accuracy of the LC-MS, when high concentration of DdSkp1 substrate had to be used. Or this may be because of different real enzyme concentration in different aliquots for these two assay experiments.

The activity of DdP4H increased with O_2 concentration up to attainable limits (60 % O_2 in gas). Maximal rates of activity were not definitively obtained. Therefore, it was not possible to

accurately calculate K_m values for O_2. The estimated K_m^{app} (O_2) is > 60 % (≈ 450 µM under the similar buffer and temperature condition, tested and reported by Dr. Rebecca Louise Hancock [2]). The high K_m value for oxygen observed indicates that oxygen is not saturating during the hydroxylation reaction under atmospheric levels. Full-length DdSkp1 protein was used as the substrate, thus, mimicking a more authentic *in vivo* condition. The $K_m(O_2)$ value determined is likely to be translated to the *in vivo* state. The high $K_m(O_2)$ value represents a relative ability to respond to changes in environmental O_2 availability, supporting the evidence that DdP4H is a potential oxygen sensor for *Dictyostelium discoideum*.

4.21 Kinetic Studies of TgPhyA with Full-Length TgSkp1 Protein Substrate

Kinetic studies were performed with TgPhyA and its full-length protein substrate, TgSkp1. A time course was performed to determine the linear range of the curve. The Fe(II) cofactor and 2OG co-substrate were kept above saturation, as described in Chapter 4.11. The time course curve remained linear during the first 2 minutes, for both 0.5 µM or 1 µM enzyme. Subsequent Michaelis-Menten kinetics analysis of TgPhyA using the TgSkp1 protein substrate was performed varying the concentration of TgSkp1. All assays were initiated at the same time by addition of TgPhyA and were quenched after 2 minutes. The samples were analysed using a Xevo G2-S QToF LC-MS machine, the method of which is discussed in Chapter 4.10. Time course curves and Michaelis-Menten curve are shown in Figure 4.20.

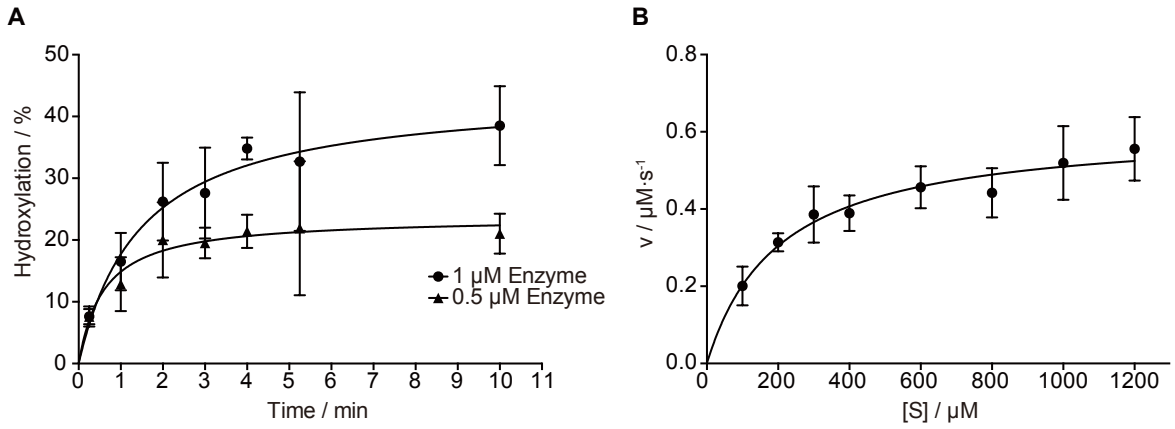


Fig. 4.20 Kinetic studies of TgPhyA with full-length TgSkp1 protein by LC-MS. (A) Time course of the hydroxylation of full-length TgSkp1 protein (100 µM) being hydroxylated by TgPhyA (1 µM and 0.51 µM) with the presence of Fe(II) (50 µM), 2OG (500 µM), and ascorbate (1 mM) in 100 mM HEPES, pH 7.6 at 37 °C. Samples were quenched with 1% (v/v) aqueous formic acid. Time course assays were performed in triplicate. (B) K_m for substrate, full-length TgSkp1 protein. Samples were prepared using the same conditions as for the time course assay, varying concentration of TgSkp1 protein (100 µM, 200 µM, 300 µM, 400 µM, 600 µM, 800 µM, 1000 µM, 1200 µM). Reactions were all quenched with 1 % (v/v) aqueous formic acid 2 minutes after initiation. All assays were performed in triplicate.

Samples were quenched 2 minutes after reaction initiation. The reaction time was less than the time when non-specific oxidation was observed (30 minutes) in Chapter 4.13. Therefore, non-specific oxidation was considered negligible on the time scale of the experiment.

Subsequent studies led to the determination of initial K_m and V_{max} values for TgSkp1, which were (202.4 ± 46.5) µM and (0.61 ± 0.04) µM·s^{-1}, respectively. The Michaelis-Menten curve reaches a plateau at high substrate concentration.

4.22 Kinetic Studies of TgPhyA with 2OG

Kinetic studies of TgPhyA using 2OG and its full-length TgSkp1 protein substrate were performed under the same conditions described in Chapter 4.20. However, no valid K_m was obtained, although several attempts were tried. It was found in later 2OG turnover assays that TgPhyA exhibits a high rate of uncoupled 2OG turnover, which is discussed in detail in Chapter

4.26.4. The high uncoupled 2OG turnover makes it impossible to determine the K_m value of 2OG for TgPhyA.

4.23 Kinetic Studies of TgPhyA with Oxygen

Samples were prepared in sealed cone bottom MS vials (Macherey-Nagel, REF: 702006) in the same manner as for DdP4H. The concentrations of the components were kept the same as described in Chapter 4.19. Reactions were carried out in a 37 °C water bath, and were quenched using 1 % (v/v) aqueous formic acid 2 minutes after initiation of the reactions, because the time course curve remained linear during the first 2 minutes (Figure 4.20 A). Samples were prepared in triplicate and were analysed using the Xevo G2-S QToF LC-MS machine. The Michaelis-Menten curve lacks the data points in the accelerating range, which were beyond the limits of detection for the instrument. Therefore, a second attempt was performed using lower oxygen percentage. Because lower amount of TgSkp1 was hydroxylated under low oxygen conditions, the RF-MS instrument with higher sensitivity for hydroxylated TgSkp1 protein was used to analyse the samples. Results of kinetic assays for oxygen are shown in Figure 4.21.

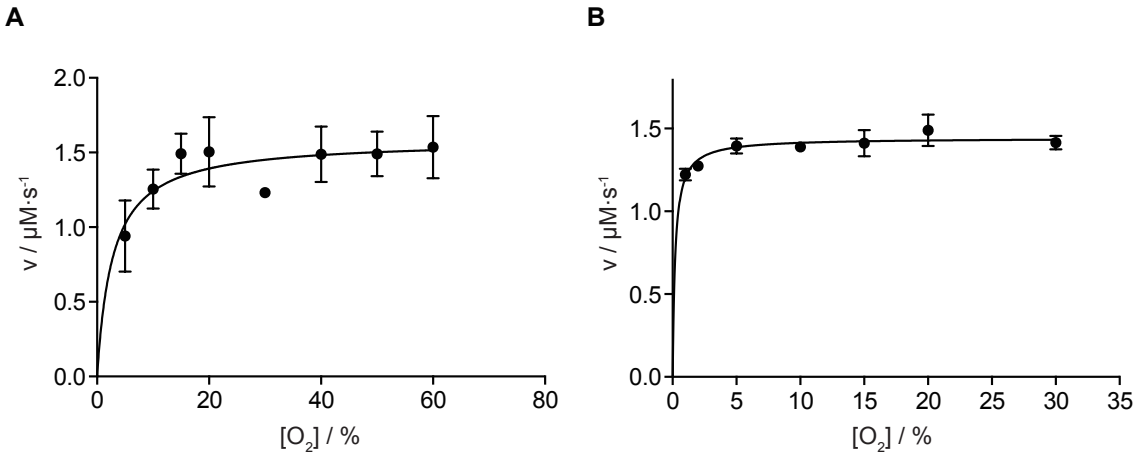


Fig. 4.21 Kinetic studies of TgPhyA with oxygen. (A) Assaying apparent K_m for oxygen with TgPhyA and TgSkp1 using the Xevo G2-S QToF LC-MS. Samples were prepared under the same condition as for the experiment of DdP4H. Different concentrations of oxygen were used to incubate the samples: 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%. Reactions were all quenched

with 1% (v/v) aqueous formic acid 2 minutes after initiation. Assays were performed in triplicate. (B) Assaying apparent K_m for oxygen with TgPhyA and TgSkp1 using RF-MS. Concentrations of oxygen used to incubate the samples were: 1%, 2%, 5%, 10%, 15%, 20%, 30%. Assays were performed in triplicate.

Using Xevo G2-S QToF LC-MS, the apparent TgPhyA K_m and V_{max} values for oxygen were (2.8 ± 0.9) % and (1.6 ± 0.07) $\mu M \cdot s^{-1}$, respectively. However, at low oxygen levels, hydroxylation was so low after 2 minutes that the detection of hydroxylated TgSkp1 using Xevo G2-S QToF LC-MS became inaccurate.

Using RF-MS, it was possible to detect hydroxylated TgSkp1 even at 1% oxygen. The apparent TgPhyA K_m and V_{max} values for oxygen were (0.2 ± 0.04) % and (1.4 ± 0.02) $\mu M \cdot s^{-1}$, respectively. The Michaelis-Menten curve still lacks the points in the accelerating range, but it was not possible to set a lower-oxygen level due to instrument limits.

The apparent V_{max} value was significantly higher than the V_{max} value for substrate (Figure 4.20). This may be due to the different batches of TgPhyA used for these two experiments, in which enzyme activities might vary.

The results indicate that the apparent K_m value resides in the range of 0 – 0.2 µM, which is consistent with the value produced using the Xevo G2-S QToF LC-MS. A low K_m value for oxygen was unexpected, taking into consideration that the K_m of oxygen with DdP4H and full-length DdSkp1 protein was too high to determine.

It needs to be stressed that a full-length TgSkp1 protein was used as the substrate, possibly mimicking a more authentic *in vivo* condition. The $K_m^{app}(O_2)$ value determined might translate to the *in vivo* situation, representing that the response to changes in O_2 availability is significant only at very low oxygen levels. Considering *Toxoplasma gondii* is a parasite, residing most of the time as tachyzoites or bradyzoites in host cells or tissues, where the oxygen concentration can be lower than 1% [3], TgPhyA may play a role as the oxygen sensor for *Toxoplasma gondii* in low

oxygen environments. This result is consistent with the reported *in vivo* result, in which the activity of His_6-TgPhyA was still highest at the 1.0 % oxygen level [4].

4.24 Discussion

Comparing the kinetic results of prolyl hydroxylation from different organisms, it is revealed that DdP4H and TgPhyA exhibit a similar substrate K_m value as the human asparaginyl hydroxylase FIH, and much higher than the human PHD2, [5-8] suggesting that even the full-length Skp1 protein substrates bind weakly to these enzymes in the assay. The apparent K_m value for DdP4H for 2OG was in a similar range as other prolyl hydroxylases. Kinetic studies of TgPhyA with 2OG was not possible to get due to the high rate of uncoupled 2OG turnover, which is discussed in detail in Chapter 4.26.4. The estimated K_m^{app} (O_2) for DdP4H is > 450 µM (60%). A high K_m value was also observed in human PHD2 [5-7], representing the relative ability of both DdP4H and PHD2 to respond to relatively small changes in atmospheric O_2 levels, indicating that DdP4H may potentially be an oxygen sensor for *Dictyostelium discoideum*. K_m value for TgPhyA with oxygen was (0.2 ± 0.04) %, consistent with the reported *in vivo* result [4]. However, a low K_m value for oxygen does not imply that TgPhyA is not an oxygen sensor in *Toxoplasma gondii*. The difference may be due to the complicated life cycle of *Toxoplasma gondii* as a parasite.

Enzyme	Substrate	K_m^{app}(2OG) / µM	K_m^{app}(Substrate) / µM	$K_m^{app}(O_2)$ / µM
FIH	HIF-1α CAD	110 ± 20	180 ± 30	110 ± 30
FIH	HIF-2α CAD	19 ± 6	315 ± 40	110 ± 10
PHD2	HIF-1α CODD	13 ± 2	10 ± 6	460 ± 30
PHD2	HIF-1α NODD	30 ± 9	11 ± 2	> 450
PHD2	HIF-2α CODD	9 ± 2	34 ± 10	> 450
PHD2	HIF-2α NODD	17 ± 5	50 ± 8	410 ± 80
DdP4H	DdSkp1 Protein	35.9 ± 5.9	349.9 ± 77.3	> 450 (60%)
TgPhyA	TgSkp1 Protein	--	202.4 ± 46.5	(0.2 ± 0.04) %

Table 4.6 Comparison of steady-state kinetic parameters of hydroxylases from different oxygen sensing systems. Data for PHD2 and FIH were derived from (Tarhonskaya, 2015) [8]. Sequences of HIF-1,2α CODD, NODD, and CAD peptides are shown in Table 1.1.

4.25 MS/MS Analysis of Proline/Hydroxyproline in Full-Length DdSkp1 Protein

4.25.1 Trypsin Digestion of Samples [9, 10]

Hydroxylated DdSkp1 protein was prepared using the conditions shown Table 4.7. The sample was incubated at 37 °C for 1 hour before being quenched with 1% (v/v) aqueous formic acid. A no enzyme control sample was prepared the same way except for the absence of DdP4H.

Assay Conditions	Final Concentration	Volume
2-oxoglutarate disodium salt (Sigma-Aldrich, K3752)	500 µM	10 µL
Sodium L-ascorbate (Sigma-Aldrich, 11140)	1 mM	10 µL
Ammonium iron(II) sulfate hexahydrate (Sigma-Aldrich, 215406)	50 µM	10 µL
Full-length DdSkp1 protein substrate	100 µM	10 µL
DdP4H enzyme	1 µM	10 µL
Buffer (100 mM HEPES, pH 7.6)	-	50 µL
Total volume	-	100 µL

Table 4.7 Components of samples for produced hydroxylated DdSkp1 prior to tryptic digestion and MS/MS analysis.

The protein (around 180 µg DdSkp1) was then precipitated following a Chloroform / Methanol Extraction protocol [11]:

800 µL methanol was added to the 200 µL protein sample. The sample was then vortexed thoroughly and 200 µL chloroform was added. 600 µL water was added and the sample was vortexed again. Three layers of liquid were apparent after centrifugation of the sample: a large aqueous layer on top, precipitated protein in the interphase, and a smaller chloroform layer at the bottom. The top aqueous layer was removed carefully. Another 800 µL methanol was then added to the sample and vortexed. The protein pelleted at the bottom of the solution after 5-minute centrifugation at 20,000 x g. The methanol was then carefully removed using a pipette. The precipitated protein was dried thoroughly using a Speed Vacuum machine (Concentrator plus from Eppendorf).

30 µL 6 M urea buffer was then added to the protein pellet and was vortexed for 2 minutes. 5 µL of DTT, a reducing reagent (200 mM DTT in 0.4 M Tris buffer, pH 7.8), was then added to the solution. The resuspended mixture was incubated for 60 minutes at room temperature while vortexing. 10 µL of iodoacetamide alkylating reagent (200 mM iodoacetamide in 0.4 M Tris buffer, pH 7.8) was added to the mixture and incubated for 60 minutes at room temperature while vortexing. Another 5 µL of DTT reducing reagent (200 mM DTT in 0.4 M Tris buffer, pH 7.8) was added to the solution while vortexing for 60 minutes at room temperature. The urea concentration was then reduced by diluting the reaction mixture with 250 µL MilliQ-H_2O. 20 µL trypsin (0.2 µg/µL, Sigma-Aldrich) was added into the sample. The solution was mixed carefully and incubated overnight at 37 °C. The next day, the reaction was stopped by adding concentrated acetic acid until a pH < 6 was reached.

A no enzyme control sample was prepared the same way.

4.25.2 MS/MS Analysis Using a Nanoscale Liquid Chromatography Coupled to Tandem Mass Spectrometry (Nano LC-MS/MS)

The trypsin digested samples were analysed with the help of Ms. Elisabete Pires (Department of Chemistry, Chemistry Research Laboratory, University of Oxford, OX1 3TA).

The trypsin digested peptides were analysed using a NanoAcquity-UPLC system (Waters®) coupled to an Orbitrap Elite™ mass spectrometer (Thermo Fischer Scientific™) [12] possessing an EASY-Spray nano-electrospray ion source (Thermo Fischer Scientific™). The peptides were trapped on a packed guard column (75 µm i.d. x 20 mm, Acclaim Pepmap100 C18, 3µm, 120 Å) using solvent A [0.1% (v/v) aqueous formic acid]. 2µL of injected peptides were separated on an EASY-spray Acclaim PepMap® analytical column (75 µm i.d. × 15mm, RSLC C18, 3 µm, 100 Å) using a linear gradient [length: 68 minutes, 3 % to 97 % solvent B (0.1% (v/v) aqueous formic acid in acetonitrile, flow rate: 300 nL/min]. Data-dependent MS acquisition full-scan MS spectra

were collected (scan range 350–1500 m/z, resolution 120 000, AGC target 1E6, maximum injection time 100 ms). After the MS scans, the 10 most intense peaks were selected for CID fragmentation at 35 % of normalised collision energy. CID spectra were acquired in the Iontrap (AGC target 1E3, maximum injection time 100 ms). The results were analysed using PEAKS Studio (Bioinformatics Solutions Inc.) [13, 14].

4.25.3 MS/MS Analysis of Proline/Hydroxyproline in Full-Length DdSkp1 Protein

The raw data from the Nano LC-MS/MS was analysed using a *de novo* sequencing and database search to maximise peptide identification. Following the identification of peptides with MS/MS spectra, the resulting peptide sequences were used for alignment with the original DdSkp1 protein sequence. Post-translational modifications (PTM) were then identified using PEAKS Studio (Bioinformatics Solutions Inc.) [13].

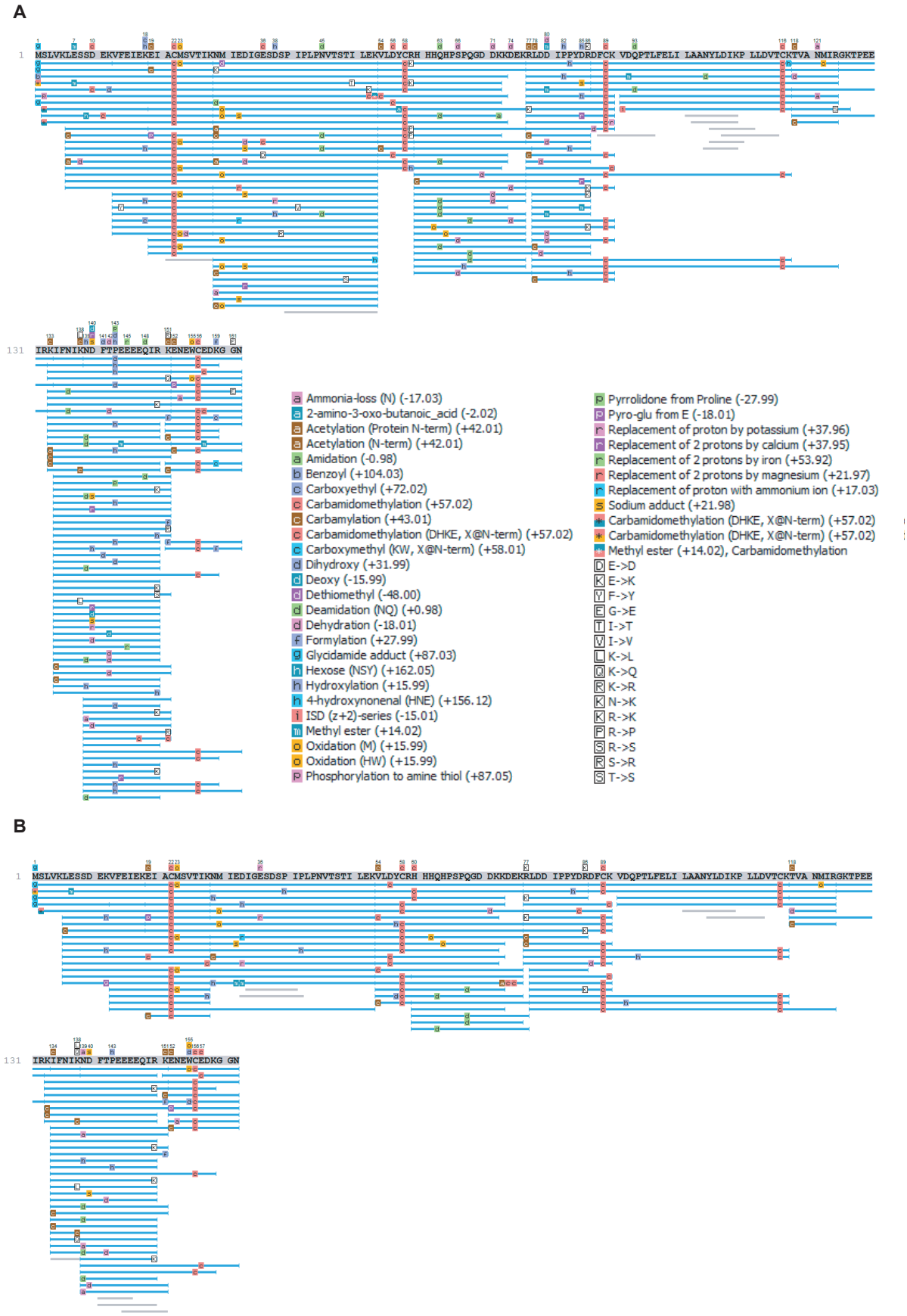

A
MSLVKLESSD EKVFEIEKEI ACMSVTIKNM IEDIGESDSP IPLPNVTSTI LEKVLDYCRH HHQHPSPQGD DKKDEKRLDD IPPYDRDFCK VDQPTLFELI LAANYLDIKP LLDVTCKTVA NMIRGKTPEE
IRKIFNIKND FTPEEEEQIR KENEWCEDKG GN
Ammonia-loss (N) (-17.03)
2-amino-3-oxo-butanoic_acid (-2.02)
Acetylation (Protein N-term) (+42.01)
Acetylation (N-term) (+42.01)
Amidation (-0.98)
Benzoyl (+104.03)
Carboxyethyl (+72.02)
Carbamidomethylation (+57.02)
Carbamylation (+43.01)
Carbamidomethylation (DHKE, X@N-term) (+57.02)
Carboxymethyl (KW, X@N-term) (+58.01)
Dihydroxy (+31.99)
Deoxy (-15.99)
Dethiomethyl (-48.00)
Deamidation (NQ) (+0.98)
Dehydration (-18.01)
Formylation (+27.99)
Glycidamide adduct (+87.03)
Hexose (NSY) (+162.05)
Hydroxylation (+15.99)
4-hydroxynonenal (HNE) (+156.12)
ISD (z+2)-series (-15.01)
Methyl ester (+14.02)
Oxidation (M) (+15.99)
Oxidation (HW) (+15.99)
Phosphorylation to amine thiol (+87.05)
Pyrrolidone from Proline (-27.99)
Pyro-glu from E (-18.01)
Replacement of proton by potassium (+37.96)
Replacement of 2 protons by calcium (+37.95)
Replacement of 2 protons by iron (+53.92)
Replacement of 2 protons by magnesium (+21.97)
Replacement of proton with ammonium ion (+17.03)
Sodium adduct (+21.98)
Carbamidomethylation (DHKE, X@N-term) (+57.02)
Carbamidomethylation (DHKE, X@N-term) (+57.02)
Methyl ester (+14.02), Carbamidomethylation
E->D
E->K
F->Y
G->E
I->T
I->V
K->L
K->Q
K->R
N->K
R->K
R->P
R->S
S->R
T->S
B
MSLVKLESSD EKVFEIEKEI ACMSVTIKNM IEDIGESDSP IPLPNVTSTI LEKVLDYCRH HHQHPSPQGD DKKDEKRLDD IPPYDRDFCK VDQPTLFELI LAANYLDIKP LLDVTCKTVA NMIRGKTPEE
IRKIFNIKND FTPEEEEQIR KENEWCEDKG GN

Fig. 4.22 Sequence alignment of peptide sequences from Nano LC-MS/MS with DdSkp1 sequence. (A) Trypsin digested peptide sequences aligned with DdSkp1 sequence from the hydroxylated sample containing DdP4H, Fe(II) cofactor, and 2OG co-substrate. (B) The no enzyme control sample peptide sequences aligned with DdSkp1 sequence. Peptide sequences with modifications of a higher confidence to align with DdSkp1 sequence lie closer to the top.

The sequence alignments shown in Figure 4.22 indicate that proline 143 was hydroxylated, whereas no hydroxylation was observed at the proline 143 in the no enzyme control sample. Detailed MS/MS spectra of the peptide containing the target proline was examined. The peaks were assigned to the theoretical mass of the fragments of DdSkp1 peptide (Figure 4.23).

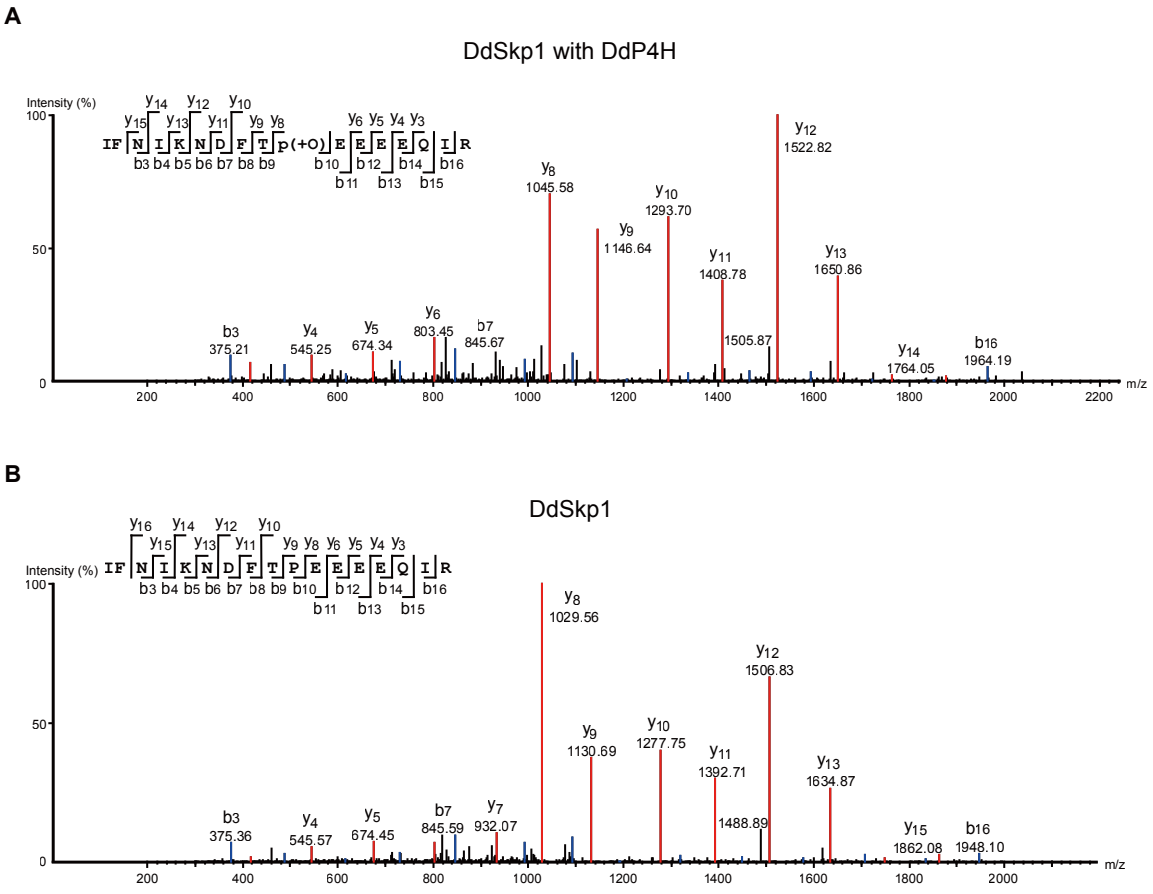


Fig. 4.23 Assignment of modification sites of trypsin digested DdSkp1 by Nano LC-MS/MS. (A) MS/MS analysis of the peptide containing the hydroxylated proline residue generated from in-solution trypsin digested full-length DdSkp1 after being incubated with DdP4H with Fe(II), 2OG, and ascorbate at 37 °C for 1 h. (B) MS/MS analysis of the peptide containing the target proline residue generated from in-solution trypsin digested full-length DdSkp1 after being incubated with Fe(II), 2OG, and ascorbate at 37 °C for 1 h without enzyme.

The spectra of the peptides from the two samples, containing the target proline, are examined, the sequence of which is IFNIKNDFTPEEEEQIR. By comparing the spectra of the two samples, we observe that all masses for the fragments of y_8 – y_{15} and b_{10} – b_{16} exhibits + 16 Da shift, indicating that hydroxylation modification occurs at the proline 143 residue in the hydroxylated sample. Therefore, it can be confirmed that DdP4H hydroxyates the full-length DdSkp1 protein at the proline 143 residue in the presence of 2OG, Fe(II) and ascorbate.

4.25.4 MS/MS Analysis of Proline/Hydroxyproline in Full-Length TgSkp1 Protein

In order to confirm the hydroxylation site in TgSkp1 treated by TgPhyA, three samples were prepared, the hydroxylated sample containing full-length TgSkp1 protein, TgPhyA, Fe(II) cofactor and 2OG co-substrate; the no enzyme control sample containing only TgSkp1 protein, Fe(II) cofactor and 2OG co-substrate; and a sample containing only full-length TgSkp1 protein. The concentrations of the components were kept the same as Table 4.7. Samples were incubated at 37 °C for 1 hour before being quenched with 1% (v/v) aqueous formic acid.

The samples were digested with trypsin using the same protocol as for digesting DdSkp1 (Chapter 4.24.1). The digested samples were analysed using the Nano LC-MS/MS, the method of which was described in Chapter 4.24.2.

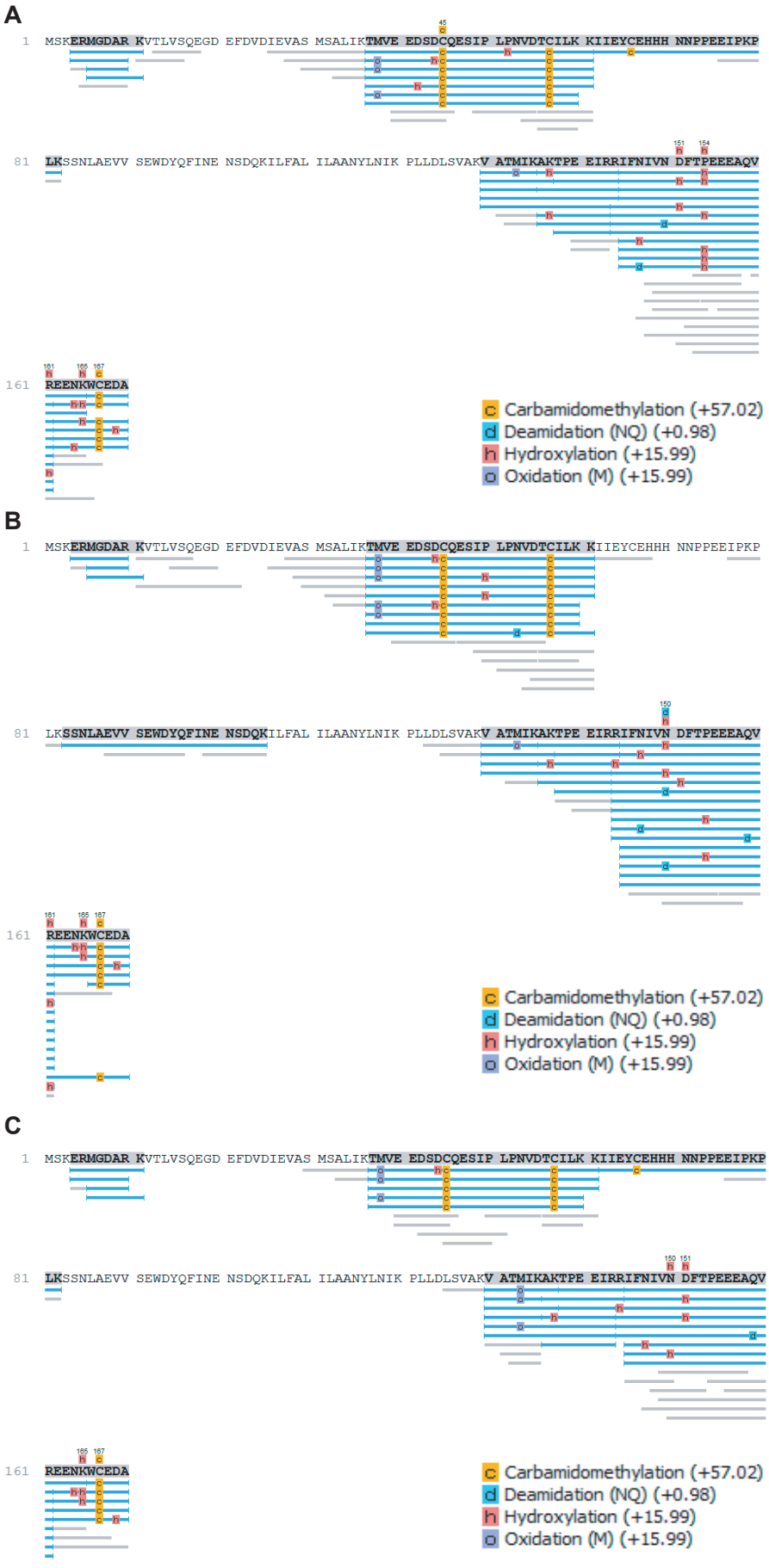

A
1 MSKERMGDAR KVTLVSQEGD EFDVDIEVAS MSALIKTMVE EDSDCQESIP LPNVDTCILK KIIEYCEHHH NNPPEEIPKP
81 LKSSNLAEVV SEWDYQFINE NSDQKILFAL ILAANYLNIK PLLDLSVAKV ATMIKAKTPE EIRRIFNIVN DFTPEEEAQV
161 REENKWCEDA
c Carbamidomethylation (+57.02)
d Deamidation (NQ) (+0.98)
h Hydroxylation (+15.99)
o Oxidation (M) (+15.99)
B
1 MSKERMGDAR KVTLVSQEGD EFDVDIEVAS MSALIKTMVE EDSDCQESIP LPNVDTCILK KIIEYCEHHH NNPPEEIPKP
81 LKSSNLAEVV SEWDYQFINE NSDQKILFAL ILAANYLNIK PLLDLSVAKV ATMIKAKTPE EIRRIFNIVN DFTPEEEAQV
161 REENKWCEDA
c Carbamidomethylation (+57.02)
d Deamidation (NQ) (+0.98)
h Hydroxylation (+15.99)
o Oxidation (M) (+15.99)
C
1 MSKERMGDAR KVTLVSQEGD EFDVDIEVAS MSALIKTMVE EDSDCQESIP LPNVDTCILK KIIEYCEHHH NNPPEEIPKP
81 LKSSNLAEVV SEWDYQFINE NSDQKILFAL ILAANYLNIK PLLDLSVAKV ATMIKAKTPE EIRRIFNIVN DFTPEEEAQV
161 REENKWCEDA
c Carbamidomethylation (+57.02)
d Deamidation (NQ) (+0.98)
h Hydroxylation (+15.99)
o Oxidation (M) (+15.99)

Fig. 4.24 Sequence alignment of peptide sequences from Nano LC-MS/MS with TgSkp1 sequence. (A) Tryptic digest peptide sequences aligned with TgSkp1 sequence from the hydroxylated sample containing TgPhyA, Fe(II) cofactor, and 2OG co-substrate. (B) The no enzyme control sample peptide sequences aligned with TgSkp1 sequence. (C) The TgSkp1 control sample. Peptide sequences with modifications of a higher confidence to align with TgSkp1 sequence lie closer to the top.

The sequence alignments of tryptic digest peptides in Figure 4.24 showed that the proline 154 is hydroxylated, whereas no hydroxylation is observed at the proline 154 in the no enzyme control. By comparing the sequencing alignment of the no enzyme control sample and the TgSkp1 control sample, likely non-specific oxidation can be observed, which possibly occurs at the methionine 38, methionine 133, asparagine 150, arginine 161, or lysine 165 residue. The discovery that non-specific oxidation does not occur at the proline 154 residue confirmed utility of performing kinetic studies using full-length TgSkp1 protein.

Detailed MS/MS spectra of the peptide containing the target proline154 was examined. The peaks were assigned to the theoretical mass of the fragments of TgSkp1 peptide, as shown in Figure 4.25.

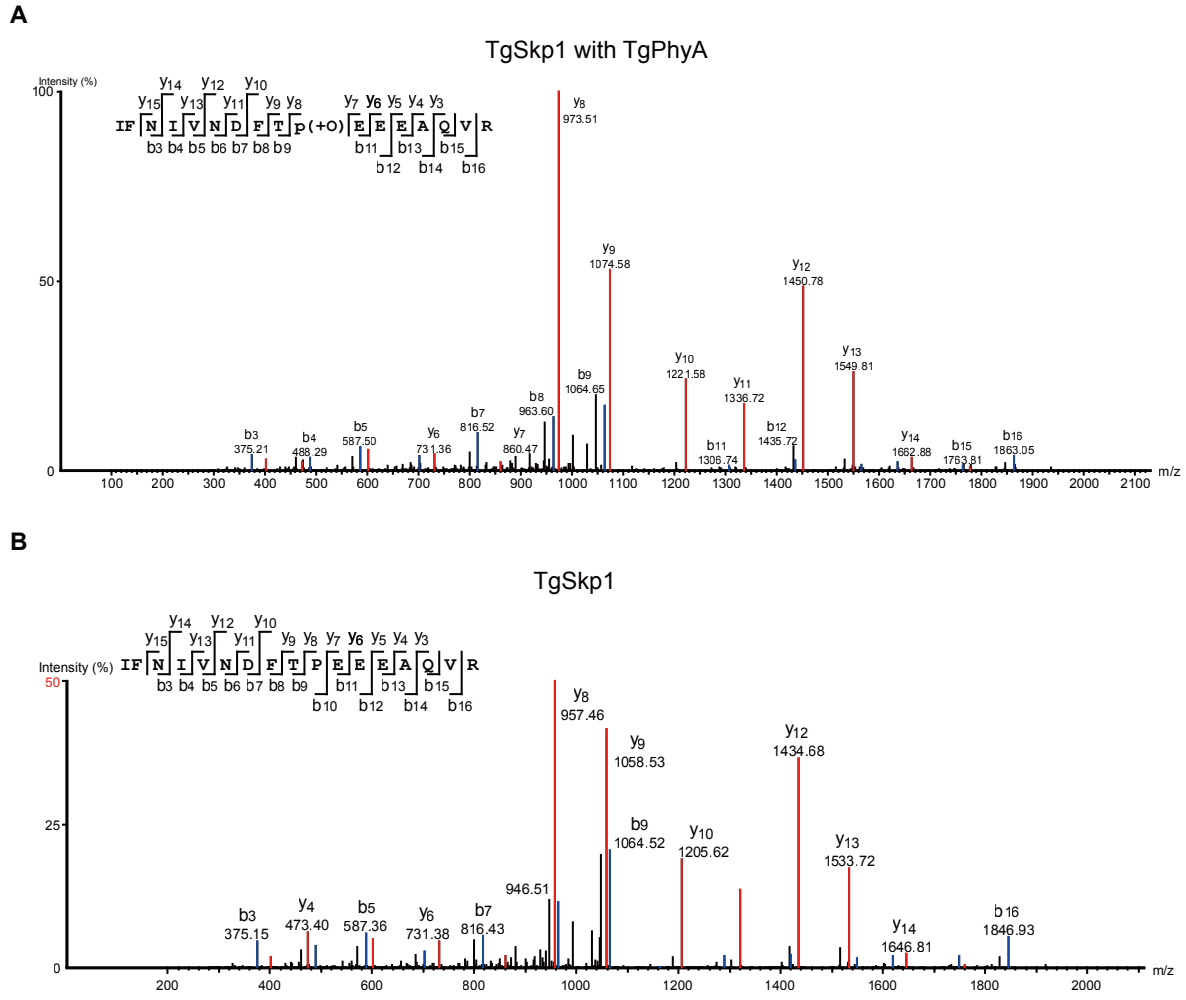


Fig. 4.25 Assignment of modification sites of trypsin digested TgSkp1 by Nano LC-MS/MS. (A) MS/MS analysis of the peptide containing the hydroxylated proline residue generated from in-solution trypsin digested full-length TgSkp1 after being incubated with TgPhyA with 2OG and Ascorbate at 37 °C for 1 h. (B) MS/MS analysis of the peptide containing the target proline residue generated from in-solution trypsin digested full-length TgSkp1 after being incubated with 2OG and Ascorbate at 37 °C for 1 h without the enzyme, TgPhyA.

The spectra of the peptides from the hydroxylated sample and the no enzyme control sample, containing the target proline 154, are examined, the sequence of which is IFNIVNDFTPEEEAQVR. By comparing the spectra of the two samples, it can be observed that all the masses of the fragments of y_8 – y_{15} and b_{10} – b_{16} exhibit + 16 Da shift, indicating hydroxylation occurs at the proline 154 residue in the hydroxylated sample. Therefore, it can be confirmed that TgPhyA

catalysed the hydroxylation of the full-length TgSkp1 protein at the proline 154 residue in the presence of 2OG, Fe(II), and ascorbate.

The LC-MS/MS analysis confirms the DdP4H and TgPhyA hydroxylation position in DdSkp1 and TgSkp1 to be in the proline residue. Notably, full-length DdSkp1 or TgSkp1 protein were used as substrates in the hydroxylation assays, trypsin digestion and LC-MS/MS analysis, more closely representing *in vivo* conditions than when using peptide substrates.

From the combined LC-MS results, it was revealed that hydroxylation of DdSkp1 and TgSkp1 by DdP4H or TgPhyA is (at least) reasonably sequence specific. Although multiple proline residues exist in DdSkp1 or TgSkp1, only the target prolines (proline 143 in DdSkp1 and proline 154 in TgSkp1) were observed to be hydroxylated.

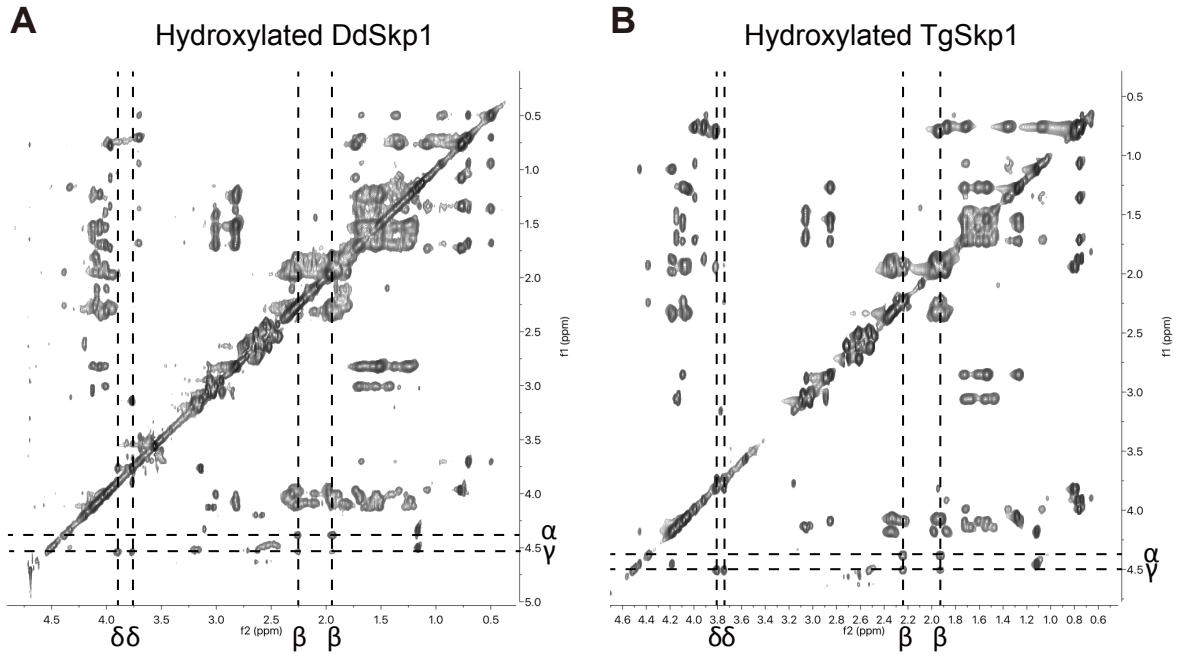
4.26 Regiochemistry of the Hydroxyproline in DdSkp1 and TgSkp1

This part of work was conducted with the help of Dr. Christopher Lohans (Department of Chemistry, Chemistry Research Laboratory, University of Oxford, OX1 3TA).

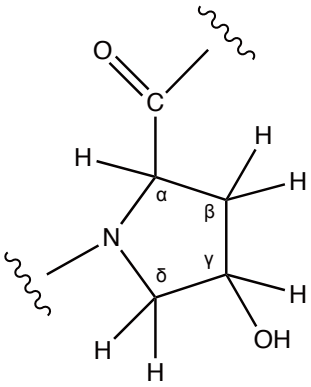
Amino acid analyses were performed to determine the regiochemistry and stereochemistry of the hydroxyproline in DdSkp1 and TgSkp1 upon treatment with DdP4H and TgPhyA. Several attempts by either acid or enzymatic hydrolysis methods, aiming to hydrolyse the amide bonds of the proteins were tried. In addition, either full-length proteins or peptides were tested in the experiments.

Unfortunately, hydroxyproline was not detected, suggesting amino acid analysis attempts to determine the stereochemistry of hydroxylated proline in DdSkp1 or TgSkp1 were unsuccessful. Possible reasons may include the spatial structure of the amide bond of the hydroxyproline in DdSkp1 or TgSkp1 inaccessible to HCl (6 M) or the hydrolysis of the enzyme, even at high temperature, not sufficient to break the amide bond of the hydroxyproline.

Therefore, NMR spectroscopy was used to determine the regiochemistry of the hydroxylated proline in DdSkp1 and TgSkp1. Hydroxylation assays were carried out under the following conditions: DdP4H (10 μM) or TgPhyA (10 μM), DdSkp1 peptide (100 μM) or TgSkp1 peptide (100 μM), $(NH_4)_2Fe(II)(SO_4)_2$ (10 μM), sodium L-ascorbate (100 μM) and 2OG (50 μM), in HEPES (100 mM), pH 7.6. The reactions were incubated at 37 °C for 1 hour and quenched with 1% (v/v) aqueous formic acid. The hydroxylated peptides were purified through a Sep-Pak Cartridge (Waters®) to remove enzymes and salts. The samples were then dried under vacuum and dissolved in D_2O to a concentration of 4 mM prior to analysis on a Bruker AV700 MHz NMR spectrometer equipped with an inverse TCI 1H/13C/15N CryoProbe. Analysis of the results are shown in Figure 4.26.



C



D

Proline Chemical Shift

Position	Average Chemical Shift of Proline	Hydroxylated Proline of DdSkp1	Hydroxylated Proline of TgSkp1
H^{α}	4.39	4	4.3
$H^{\beta 2}$	2.08	1.9	1.9
$H^{\beta 3}$	2.00	2.3	2.2
$H^{\gamma 2}$	**1.92**	**4.5**	**4.5**
$H^{\gamma 3}$	**1.90**		
$H^{\delta 2}$	3.65	3.7	3.7
$H^{\delta 3}$	3.61	3.9	3.8

Fig. 4.26 Investigation of the regiochemistry of proline hydroxylation of DdSkp1 and TgSkp1 by DdP4H and TgPhyA, respectively, by NMR spectroscopy. (A) 2D total correlation spectroscopy (TOCSY) spectrum (700 MHz) of hydroxylated DdSkp1 peptide (FNIKNDFTPEEEEQIRKENEWCEDK-NH_2) containing a 4-hydroxyproline. 2 scans were made at the resolution of 512 data points in the indirect detected dimension. Mixing time: 100 ms. (B) 2D TOCSY spectrum (700 MHz) of TgSkp1 peptide (FNIVNDFTPEEEAQVREENK-NH_2) containing a 4-hydroxyproline. The signals corresponding to the protons of the hydroxyproline are indicated with lines. The distinctive chemical shifts of these protons are consistent in the presence of 4-hydroxyproline in both hydroxylated DdSkp1 and TgSkp1. 4 scans were made at the resolution of 512 data points in the indirect detected dimension. Mixing time: 100 ms. (C) Chemical structure of a 4-hydroxyproline-residue with the carbon atoms labelled. (D) Table showing the average proton chemical shifts of unhydroxylated proline from the Biological Magnetic Resonance Data Bank as compared to the chemical shifts of the protons of the DdSkp1 and TgSkp1 4-hydroxyproline.

2D TOCSY spectra (700 MHz) of the hydroxylated DdSkp1 and TgSkp1 peptides showed the hydroxyproline was hydroxylated at the C-4 position by DdP4H and TgPhyA, respectively.

4.27 2OG Turnover Assays

This part of work was conducted with the help of Dr. Martine I. Abboud (Department of Chemistry, Chemistry Research Laboratory, University of Oxford, OX1 3TA).

4.27.1 2OG Turnover Assays Conditions

Knowing that DdP4H, TgPhyA are able to hydroxylate DdSkp1 and TgSkp1, we next aimed to investigate the turnover of 2OG with these enzymes. ^{1}H-NMR spectroscopy was used to monitor the conversion of 2OG to succinate during hydroxylation. The enzymes used include: DdP4H, TgPhyA, and human PHD2. The substrates tested include: Full-length DdSkp1, full-length TgSkp1, full-length HsSkp1-His_6 and full-length HsSkp1 E147P-His_6. Proteins were first buffer exchanged into 50 mM Tris-D_2, pH 7.5. Stock solutions were made in the same buffer.

A

Assay Conditions	Stock Concentration	Final Concentration
Enzyme	> 100 µM	20 µM
Substrate	> 800 µM	400 µM
2-oxoglutarate disodium salt (Sigma-Aldrich, K3752)	5 mM	500 µM
Sodium L-ascorbate (Sigma-Aldrich, 11140)	10 mM	1 mM
Ammonium iron(II) sulfate hexahydrate (Sigma-Aldrich, 215406)	500 µM	50 µM

B

Sample	Sample Number	Enzyme	Enzyme Volume / uL	Substrate	Substrate Volume / uL	2OG / uL	Ascorbate / uL	Fe(II) / uL	Tris-D_2 / uL	D_2O / uL	Whole Volume / uL
1	AV700: 1-22	DdP4H (355.9 µM)	28	DdSkp1 (1393 µM)	143.6	50	50	50	128.4	50	500
2	AV700: 31-52		28	TgSkp1 (968.6 µM)	206.5	50	50	50	65.5	50	500
3	AV700: 91-112		28	HsSkp1 E147P (1104 µM)	181.1	50	50	50	90.9	50	500
4	AV700: 121-142		28	HsSkp1 (1730 µM)	115.6	50	50	50	156.4	50	500

5	AV700: 61-82		28	—	0	50	50	50	272	50	500
6	AV700: 401-422	TgPhyA (332.6 µM)	30	DdSkp1 (1393 µM)	143.6	50	50	50	126.4	50	500
7	AV700: 431-452		30	TgSkp1 (968.6 µM)	206.5	50	50	50	63.5	50	500
8	AV700: 461-482		30	HsSkp1 E147P (1104 µM)	181.1	50	50	50	88.9	50	500
9	AV700: 491-512		30	HsSkp1 (1730 µM)	115.6	50	50	50	154.4	50	500
10	AV700: 521-542		30	—	0	50	50	50	270	50	500
11	AV600: 71-106	PHD2	25.7	DdSkp1 (1393 µM)	143.6	50	50	50	130.7	50	500
12	AV600: 111-145		25.7	TgSkp1 (968.6 µM)	206.5	50	50	50	67.8	50	500
13	AV600: 151-185		25.7	HsSkp1 E147P (1104 µM)	181.1	50	50	50	93.2	50	500
14	AV600: 191-225		25.7	HsSkp1 (1730 µM)	115.6	50	50	50	158.7	50	500
15			25.7	—	0	50	50	50	274.3	50	500
16	AV700: 151-171	—	0	DdSkp1 (1393 µM)	7.18	50	50	50	292.82	50	500
17	AV700: 181-202		0	TgSkp1 (968.6 µM)	10.3	50	50	50	289.7	50	500
18	AV700: 211-228		0	HsSkp1 E147P (1104 µM)	5.78	50	50	50	294.22	50	500
19	AV700: 241-259		0	HsSkp1 (1730 µM)	9	50	50	50	291	50	500
20	AV700: 271-332		0	—	0	50	50	50	300	50	500

Table 4.8 (A) 2OG turnover assay conditions. (B) Records of 2OG turnover assay samples.

NMR spectra were recorded using either a Bruker AVIII 700 MHz NMR spectrometer equipped with a 5-mm inverse cryoprobe or a Bruker AVIII 600 MHz NMR spectrometer equipped with a broad band BB-$^{19}F/^{1}H$ Prodigy N_2 cryoprobe using 5 mm diameter NMR tubes (Norell) or 3 mm MATCH NMR tubes (Cortectnet). Data were processed with Bruker 3.1 software.

^{1}H excitation sculpting suppression NMR spectra were obtained at 150 s intervals and integrated using absolute intensity scaling to monitor changes in the intensity of signals of interest [15]. The reaction was initiated by addition of 2OG. The reactions were carried out at room temperature.

4.27.2 2OG Turnover Assays Results

Examples of spectra of 2OG turnover can be seen below in Figure 4.27. For the 3-H of the molecule of 2OG, there are two hydrogen atoms attached to the adjacent carbon, exhibiting a triplet peak in the NMR spectrum. The four hydrogen atoms in succinate are equivalent, exhibiting a single peak in the NMR spectrum. The heights of the peaks can represent the comparative ratio of 2OG and succinate in the samples, reflecting the percentage of 2OG turnover.

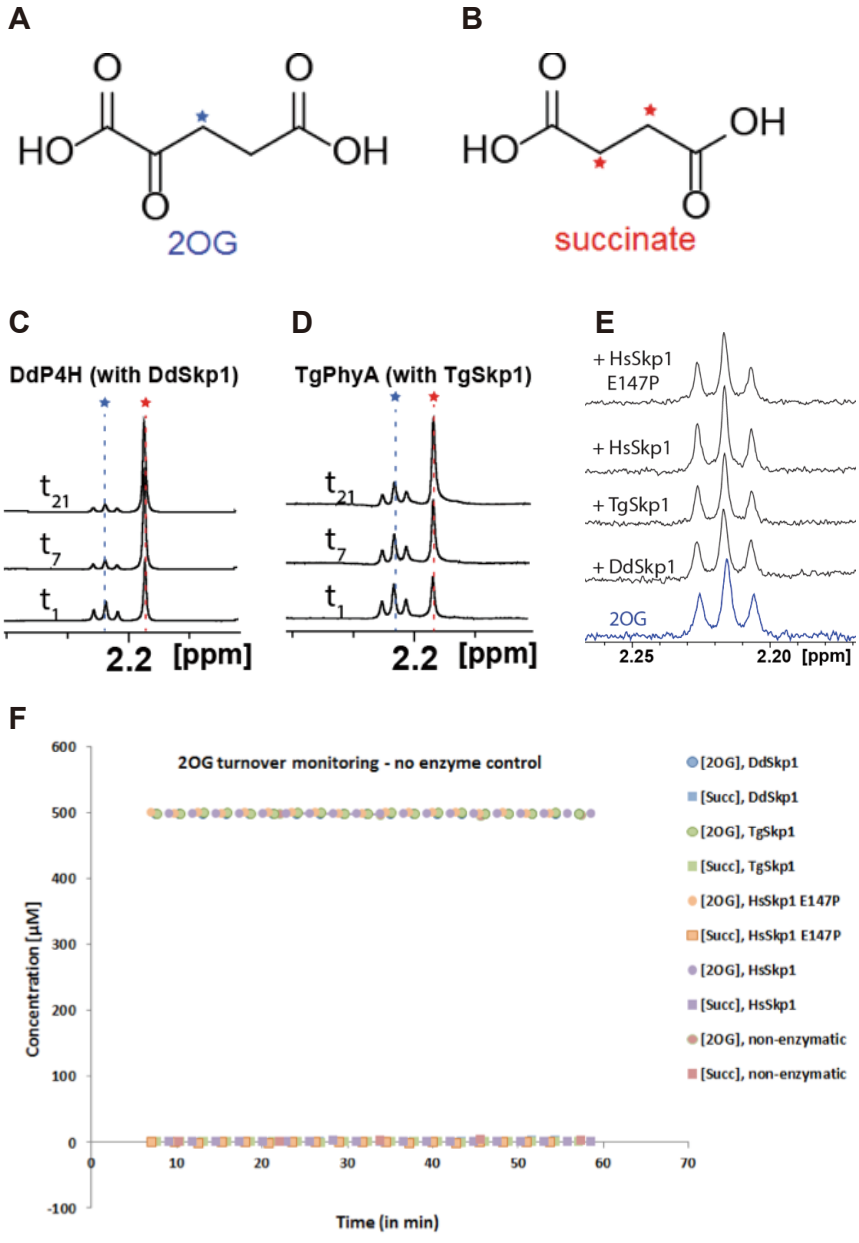


Fig. 4.27 2OG turnover to succinate of DdP4H and TgPhyA using ^{1}H NMR by Bruker AV600 and AV700 MHz Spectrometer. (A) Structure of 2OG. The hydrogen atoms monitored by NMR are labelled as a blue star. (B) Structure of succinate. The hydrogen atoms monitored by NMR are labelled as red stars. (C) Examples of spectra of 2OG turnover of DdP4H in the presence of the substrate, DdSkp1. (D) Examples of spectra of 2OG turnover of TgPhyA in the presence of the substrate, TgSkp1. (E) Investigation on the binding of 2OG to Skp1. Spectra of 2OG turnover with Skp1s after 1 h incubation. Peaks corresponding to the hydrogen atoms of 2OG and succinate are indicated by the stars by the same colours as figure A and B. (F) Time course of 2OG turnover with various Skp1s. Concentration of 2OG is indicated as dots. Concentration of succinate is indicated as squares. Time course of various substrates are distinguished by different colours, including blue (DdSkp1), green (TgSkp1), yellow (HsSkp1 E147P), purple (HsSkp1) and red (no substrate).

During the reaction incubation, the peak indicating succinate increased, while the peaks indicating 2OG decreased, suggesting a turnover of 2OG to succinate. This observation occurred in both DdP4H + DdSkp1 and TgPhyA + TgSkp1 samples. The 2OG turnover did not occur in the absence of enzyme in DdSkp1, TgSkp1 HsSkp1 or HsSkp1 E147P.

4.27.3 2OG Turnover of DdP4H with Various Substrates

A time course of 2OG turnover of DdP4H with various substrates was generated.

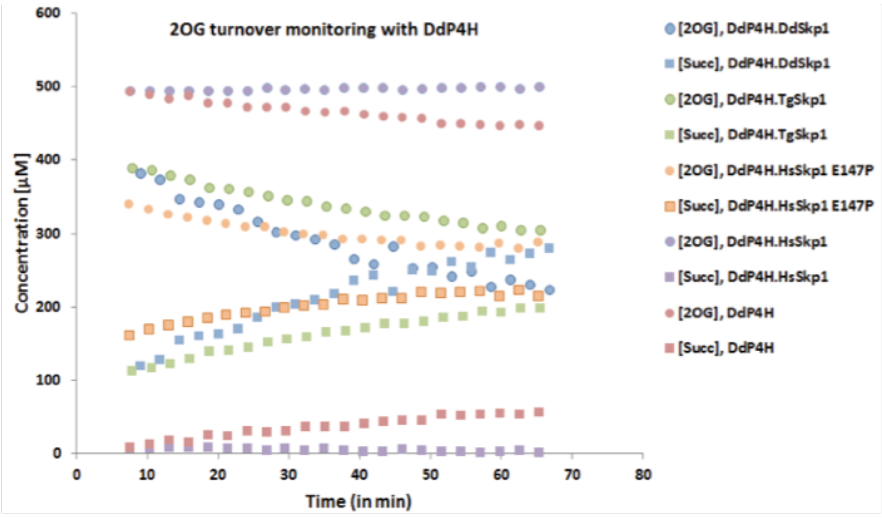


Fig. 4.28 2OG turnover to succinate by DdP4H using ^{1}H NMR by Bruker AV600 and AV700 MHz Spectrometer. Concentration of 2OG is indicated as dots. Concentration of succinate is indicated as squares. Time course of DdP4H with various substrates are distinguished by different colours, including blue (DdSkp1), green (TgSkp1), yellow (HsSkp1 E147P), purple (HsSkp1) and red (no substrate).

Some interesting results can be seen from the result of 2OG turnover:

- DdP4H exhibits a slow substrate uncoupled turnover of 2OG, with ~10% of 2OG being turned into succinate in 1 hour without the substrate, Skp1.
- The presence of DdSkp1, TgSkp1, or HsSkp1 E147P result in increased 2OG turnover for DdP4H.
- The turnover rate is the fastest with the substrate DdSkp1, suggesting DdSkp1 is a better substrate than other Skp1s for DdP4H.
- 2OG turnover decreased to ~ 0% (within detection limits) with the presence of HsSkp1. This might because HsSkp1 binds to DdP4H to block the active site of DdP4H preventing the turnover of 2OG.

4.27.4 2OG Turnover of TgPhyA with Various Substrates

A time course of 2OG turnover of TgPhyA in the presence of various substrates.

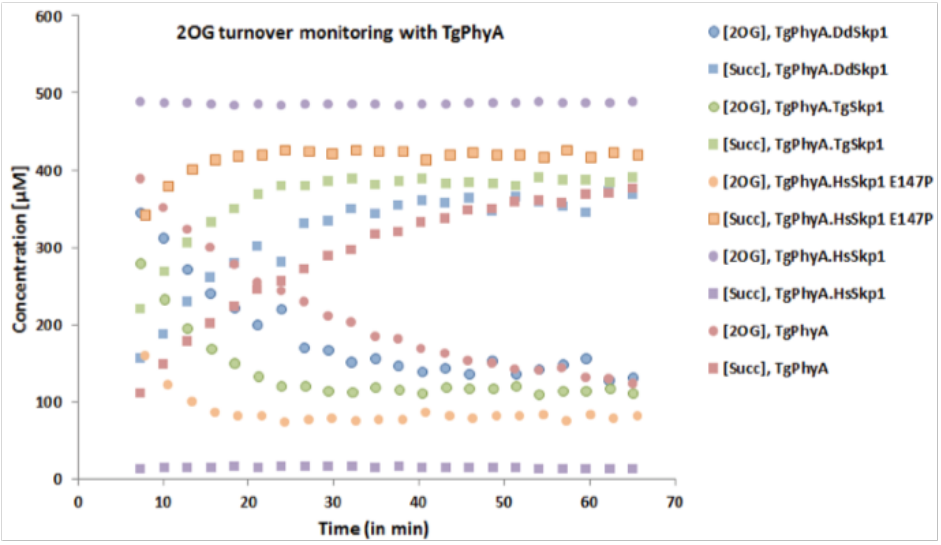


Fig. 4.29 2OG turnover to succinate by TgPhyA using ^{1}H NMR by Bruker AV600 and AV700 MHz Spectrometer. Concentration of 2OG is indicated as dots. Concentration of succinate is indicated as squares. Time course of TgPhyA with various substrates are distinguished by

different colours, including blue (DdSkp1), green (TgSkp1), yellow (HsSkp1 E147P), purple (HsSkp1) and red (no substrate).

- TgPhyA exhibits a significantly rapid substrate uncoupled turnover of 2OG, with > 60% of 2OG turned into succinate in 1 hour without the substrate, Skp1.
- The presence of DdSkp1, TgSkp1 or HsSkp1 E147P result in increased 2OG turnover for TgPhyA.
- The turnover rate was the fastest with the substrate HsSkp1 E147P, suggesting it may be a better substrate than other Skp1s for TgPhyA, including its original substrate TgSkp1.
- 2OG turnover decreased to ~ 0% (within detection limits) with the presence of HsSkp1. This might because HsSkp1 binds to TgPhyA to block the active site of TgPhyA preventing the turnover of 2OG.

4.27.5 2OG Turnover of PHD2 with Various Skp1s

A time course of 2OG turnover by PHD2 with various Skp1s was generated.

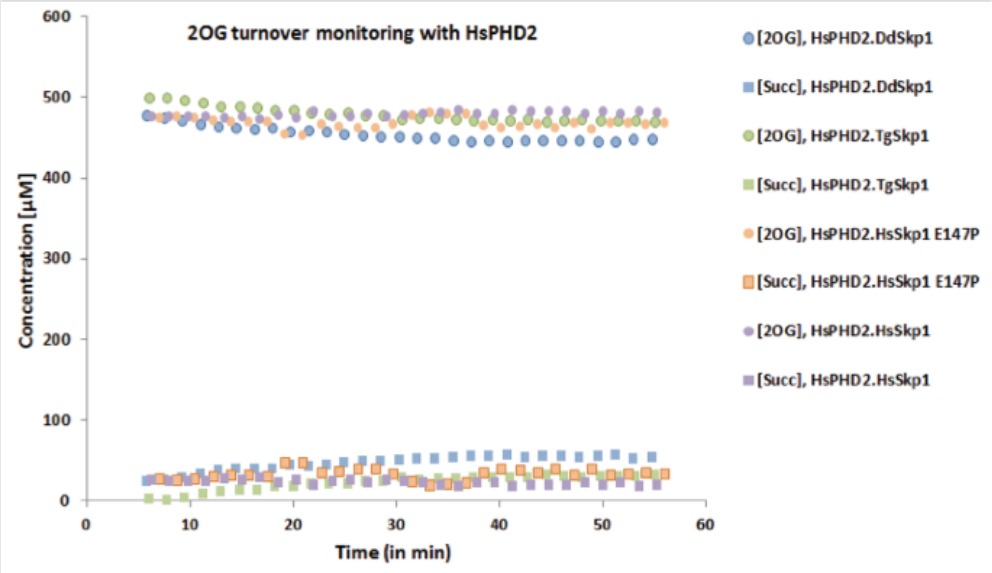


Fig. 4.30 2OG turnover to succinate by PHD2 using ^{1}H NMR by Bruker AV600 and AV700 MHz Spectrometer. Concentration of 2OG is indicated as dots. Concentration of succinate is

indicated as squares. Time course of PHD2 with various substrates are distinguished by different colours, including blue (DdSkp1), green (TgSkp1), yellow (HsSkp1 E147P) and purple (HsSkp1).

- PHD2 exhibits a small amount of substrate uncoupled turnover of 2OG (as reported [16]), with < 10% of 2OG turned into succinate in 1 hour with DdSkp1, TgSkp1, HsSkp1 E147P.
- Comparing the turnover of 2OG in the cases of DdP4H and TgPhyA, the amount of 2OG turnover of PHD2 with various Skp1s was much lower. The result is consistent with the discovery that PHD2 were not able to hydroxylate DdSkp1, TgSkp1 or HsSkp1 E147P. (Chapter 7)
- 2OG turnover decreased to ~ 0% (within detection limits) in the presence of HsSkp1. This might because HsSkp1 is able to bind to PHD2 to prevent the turnover of 2OG.

4.27.6 A Further Investigation of the Uncoupled 2OG Turnover by TgPhyA

The time courses of uncoupled 2OG turnover by DdP4H and TgPhyA were extracted to compare. TgPhyA exhibits a much rapid of uncoupled 2OG turnover. The effect of ascorbate in the process of 2OG turnover was further investigated.

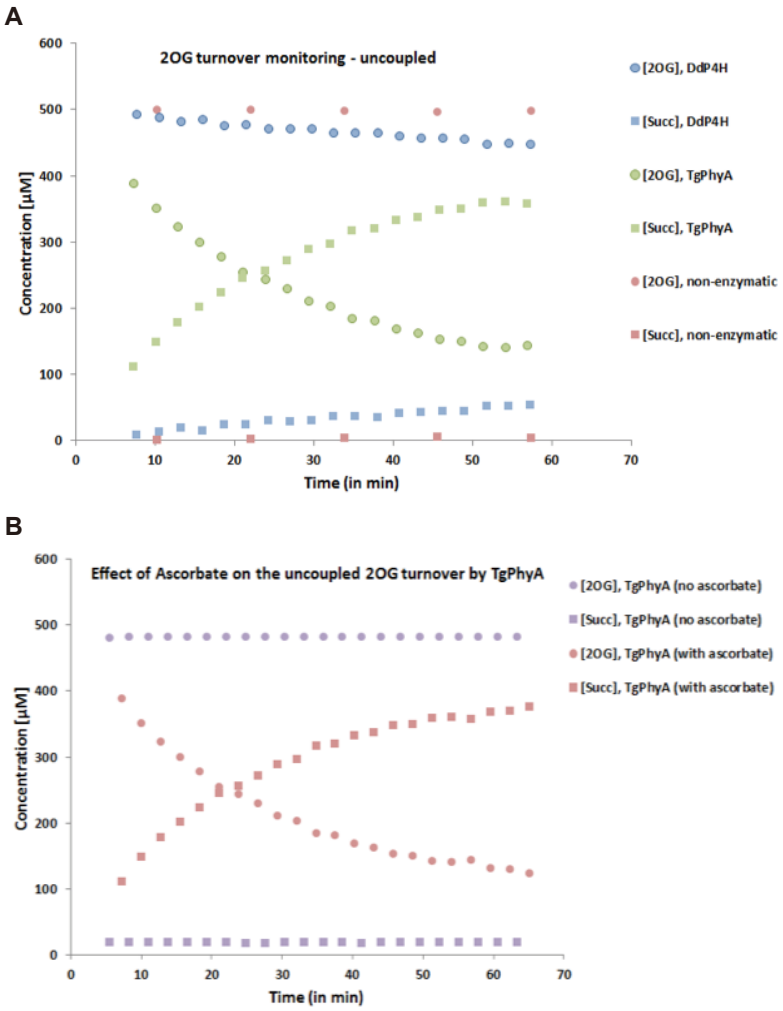


Fig. 4.31 Time course of uncoupled 2OG turnover to succinate using ^{1}H NMR by Bruker AV600 and AV700 MHz Spectrometer. (A) Comparison of uncoupled 2OG turnover between DdP4H (blue) and TgPhyA (green). (B) The uncoupled 2OG turnover by TgPhyA with (red) and without (purple) ascorbate.

Without ascorbate, the uncoupled 2OG turnover by TgPhyA decreased to ~ 0% (within detection limits), suggesting the need of ascorbate in the process of uncoupled 2OG turnover. The function and the mechanism of ascorbate affecting 2OG turnover is still unclear.

4.28 References

1. Leveridge, M., Buxton, R., Argyrou, A., Francis, P., Leavens, B., West, A., Rees, M., Hardwicke, P., Bridges, A., Ratcliffe, S., and Chung, C. W. (2014) Demonstrating enhanced throughput of RapidFire mass spectrometry through multiplexing using the JmjD2d demethylase as a model system, *J Biomol Screen 19*, 278-286.

2. Hancock, R. L. (2016) Investigating the effect of hypoxia on the JmjC histone lysine demethylase KDM4A, University of Oxford.

3. Sakadzic, S., Roussakis, E., Yaseen, M. A., Mandeville, E. T., Srinivasan, V. J., Arai, K., Ruvinskaya, S., Devor, A., Lo, E. H., Vinogradov, S. A., and Boas, D. A. (2010) Two-photon high-resolution measurement of partial pressure of oxygen in cerebral vasculature and tissue, *Nat Methods 7*, 755-759.

4. Xu, Y., Brown, K. M., Wang, Z. A., van der Wel, H., Teygong, C., Zhang, D., Blader, I. J., and West, C. M. (2012) The Skp1 protein from *Toxoplasma* is modified by a cytoplasmic prolyl 4-hydroxylase associated with oxygen sensing in the social amoeba Dictyostelium, *J Biol Chem 287*, 25098-25110.

5. Flashman, E., Hoffart, L. M., Hamed, R. B., Bollinger, J. M., Jr., Krebs, C., and Schofield, C. J. (2010) Evidence for the slow reaction of hypoxia-inducible factor prolyl hydroxylase 2 with oxygen, *FEBS J 277*, 4089-4099.

6. Tarhonskaya, H., Chowdhury, R., Leung, I. K., Loik, N. D., McCullagh, J. S., Claridge, T. D., Schofield, C. J., and Flashman, E. (2014) Investigating the contribution of the active site environment to the slow reaction of hypoxia-inducible factor prolyl hydroxylase domain 2 with oxygen, *Biochem J 463*, 363-372.

7. Tarhonskaya, H., Rydzik, A. M., Leung, I. K., Loik, N. D., Chan, M. C., Kawamura, A., McCullagh, J. S., Claridge, T. D., Flashman, E., and Schofield, C. J. (2014) Non-enzymatic chemistry enables 2-hydroxyglutarate-mediated activation of 2-oxoglutarate oxygenases, *Nat Commun 5*, 3423.

8. Tarhonskaya, H., Hardy, A. P., Howe, E. A., Loik, N. D., Kramer, H. B., McCullagh, J. S., Schofield, C. J., and Flashman, E. (2015) Kinetic Investigations of the Role of Factor Inhibiting Hypoxia-inducible Factor (FIH) as an Oxygen Sensor, *J Biol Chem 290*, 19726-19742.

9. Gundry, R. L., White, M. Y., Murray, C. I., Kane, L. A., Fu, Q., Stanley, B. A., and Van Eyk, J. E. (2009) Preparation of proteins and peptides for mass spectrometry analysis in a bottom-up proteomics workflow, *Curr Protoc Mol Biol Chapter 10*, Unit10 25.

10. Huynh, M. L., Russell, P., and Walsh, B. (2009) Tryptic digestion of in-gel proteins for mass spectrometry analysis, *Methods Mol Biol 519*, 507-513.

11. Jiang, L., He, L., and Fountoulakis, M. (2004) Comparison of protein precipitation methods for sample preparation prior to proteomic analysis, *J Chromatogr A 1023*, 317-320.

12. Michalski, A., Damoc, E., Lange, O., Denisov, E., Nolting, D., Muller, M., Viner, R., Schwartz, J., Remes, P., Belford, M., Dunyach, J. J., Cox, J., Horning, S., Mann, M., and Makarov, A. (2012) Ultra high resolution linear ion trap Orbitrap mass spectrometer (Orbitrap Elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes, *Mol Cell Proteomics 11*, O111 013698.

13. Zhang, J., Xin, L., Shan, B., Chen, W., Xie, M., Yuen, D., Zhang, W., Zhang, Z., Lajoie, G. A., and Ma, B. (2012) PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification, *Mol Cell Proteomics 11*, M111 010587.

14. McCafferty, J., Reid, R., Spencer, M., Hamp, T., and Fodor, A. (2012) Peak Studio: a tool for the visualization and analysis of fragment analysis files, *Environ Microbiol Rep 4*, 556-561.

15. Walport, L. J., Hopkinson, R. J., Chowdhury, R., Schiller, R., Ge, W., Kawamura, A., and Schofield, C. J. (2016) Arginine demethylation is catalysed by a subset of JmjC histone lysine demethylases, *Nat Commun 7*, 11974.

16. Flashman, E., Bagg, E. A., Chowdhury, R., Mecinovic, J., Loenarz, C., McDonough, M. A., Hewitson, K. S., and Schofield, C. J. (2008) 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*, 3808-3815.

5. Inhibition Studies on DdP4H and TgPhyA

The experiments in this chapter were conducted with the help of Dr. Anthony Tumber. (Department of Chemistry, Chemistry Research Laboratory, University of Oxford, OX1 3TA).

5.1 Summary

The work described in this chapter concerns studies on the inhibition of recombinant DdP4H and TgPhyA full-length proteins. A series of known and potential 2OG oxygenase inhibitors were tested. The IC_{50}s are obtained, compared and discussed. The effects of using full-length DdSkp1 protein or DdSkp1 peptide fragment as substrates to perform inhibition assays are also discussed.

5.2 Introduction

Inhibition of the 2OG oxygenase superfamily has been studied for many years [1]. Many of the identified small molecule inhibitors target the active site and directly inhibit normal catalytic function [2]. Some of the reported inhibitors are non cell-penetrating acids but can be made cell permeable by forming esters [2]. 2OG oxygenase inhibitors are being developed into drugs for therapeutic applications, and are of interest as tools for probing the biological function of the target proteins. Inhibition can help in 2OG oxygenase crystallisation and provide valuable information for further inhibitor design. The development of effective inhibitors targeting TgPhyA may also be a clinically useful treatment for the disease toxoplasmosis.

5.3 Known Prolyl Hydroxylase Inhibitors for DdP4H and TgPhyA

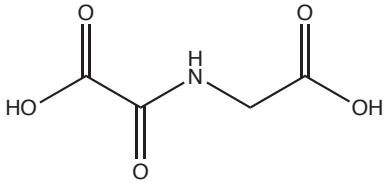
The inhibition of the human PHDs has been studied extensively. PHD inhibitors are currently in Phase III clinical trials for the treatment of anaemia in chronic kidney desease [3-5], and have

potential in other therapeutic areas [6, 7]. Sequence alignments show that PHD2 and DdP4H, PHD2 and TgPhyA share approximately 28% and 29% sequence identity (Phyre2 [8]), respectively. The high K_m values for DdP4H suggests that DdP4H may act as an oxygen sensor, like PHD2 does (Chapter 4). The crystallographic studies of DdP4H show that PHD2 and DdP4H are structurally similar (Chapter 3). Given the similarity of these proteins, it was of interest to test DdP4H and TgPhyA using known PHD inhibitors.

5.3.1 Potential 2OG Oxygenase Inhibitors and 2OG Analogues

As can be seen from the crystal structure of DdP4H and the sequence alignment of TgPhyA, both DdP4H and TgPhyA maintain the conserved core DSBH structure, as in other 2OG dependent oxygenases (Chapter 3.5 and 2.8). The hydroxylation assays with DdSkp1 or TgSkp1 also showed a requirement for Fe(II) and 2OG in DdP4H and TgPhyA, suggesting that inhibitors targeting the Fe(II) and 2OG binding sites should be tested.

A number of broad spectrum 2OG oxygenase inhibitors have been described, many of which are 2OG analogues that chelate the active site Fe(II) and occupy the 2OG binding site. These inhibitors include N-oxalylglycine (NOG) [9], 2,4-pyridinedicarboxylic acid (2,4-PDCA) [10] and IOX1 [11]. Other 2OG analogues may also have the potential to inhibit DdP4H and TgPhyA, including D,L-α-hydroxyglutaric acid, D,L-4-hydroxy-2-ketoglutaric acid, L-α-hydroxyglutaric acid, D-α-hydroxyglutaric acid, and 2,6-dipicolinic acid. [9-11]

Inhibitor	Structure	Comment
NOG		An amide analogue of 2OG is an inhibitor of multiple 2OG dependent enzymes, including JMJD2A, JMJD2C, and JMJD2E, as well as prolyl hydroxylases, including PHD1 and PHD2. [9]

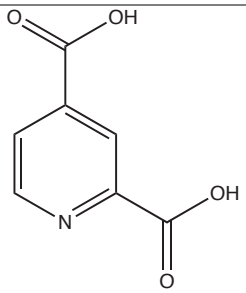
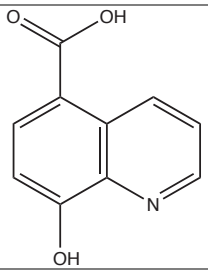
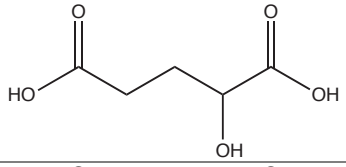
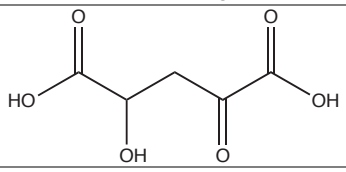
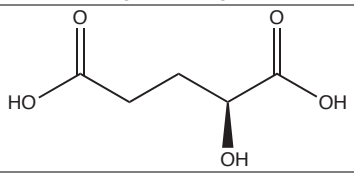
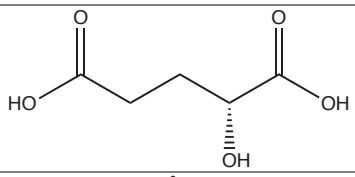
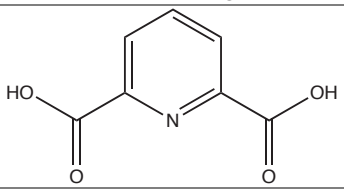
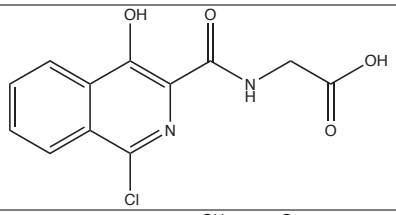
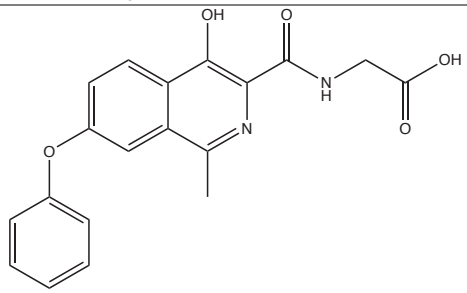
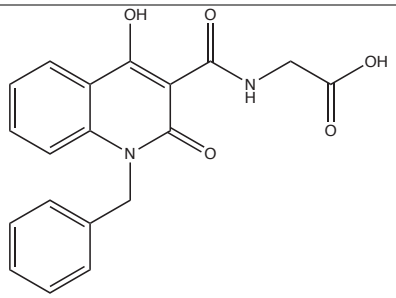
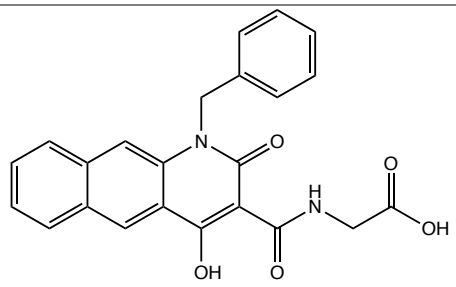
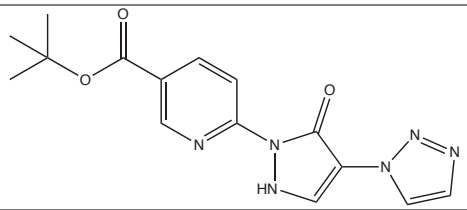
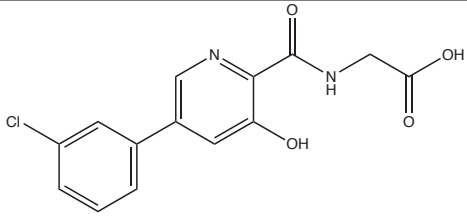
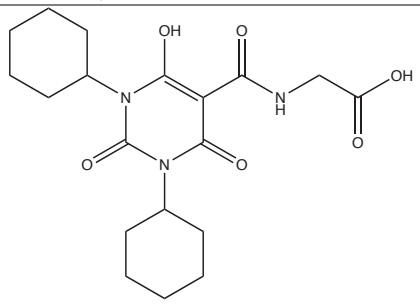
2,4-PDCA		Broad spectrum inhibitor of the 2OG oxygenases, including PHDs, collagen prolyl hydroxylase (CPH), 2-oxoglutarate and iron-dependent oxygenase domain containing 1 (OGFOD1), and histone lysine demethylases. [10]
IOX1		Broad spectrum inhibitor of 2OG oxygenases, including JmjC demethylases. [11]
DL-α-hydroxyglutaric acid		2OG analogue
DL-4-hydroxy-2-ketoglutaric acid		2OG analogue
L-α-hydroxyglutaric acid		2OG analogue
D-α-hydroxyglutaric acid		2OG analogue
2,6-dipicolinic acid		2OG analogue

Table 5.1 Structures of potential 2OG oxygenase inhibitors and 2OG analogues.

5.3.2 PHD2 Inhibitors

Potent PHD2 inhibitors include FG-2216, FG-4592, IOX2, IOX4, Vadadustat, GSK1278863, BIOX2, FG-0041 and deferoxamine mesylate. [3, 12-21]

Inhibitor	Structure	Comment
FG-2216		PHD2 inhibitor in clinical trials for treatment of anaemia. [3, 12, 13]
FG-4592		Analogue of 2OG that acts as prolyl-hydroxylase inhibitor in Phase III clinical trials for treatment of renal anaemia. [14, 15]
IOX2		Inhibitor for HIF prolyl-hydroxylases. [16]
BIOX2		Homolog of IOX2, which also inhibits HIF prolyl-hydroxylases. [17]
IOX4		Prolyl hydroxylase inhibitor, which increases HIF-1α levels in cells. OGFOD1 inhibitor. [18]
Vadadustat		Prolyl hydroxylase inhibitor. OGFOD1 inhibitor. [19, 20]
GSK1278863		Prolyl hydroxylase inhibitor, which is in clinical trials, currently being investigated as a treatment for renal anemia. OGFOD1 inhibitor. [21]

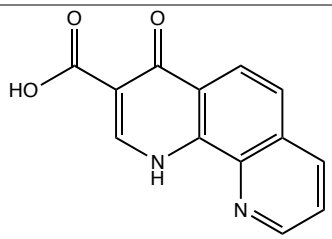
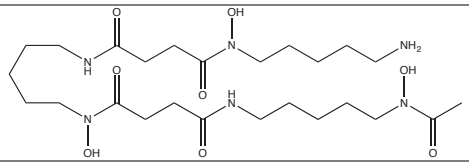
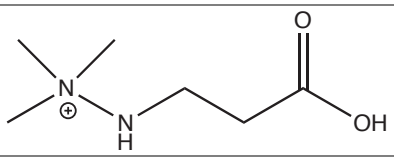
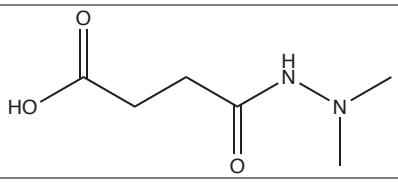
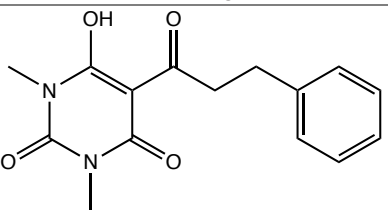
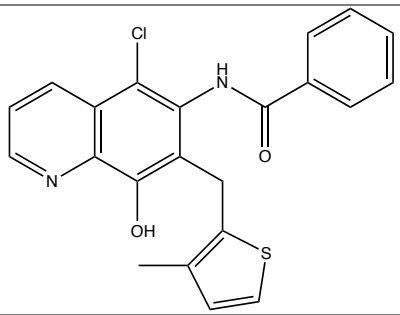
FG-0041		Prolyl hydroxylase inhibitor. [22]
Deferoxamine mesylate		Iron chelator which stabilises HIF-1α and inhibits prolyl hydroxylases (PHDs).

Table 5.2 Structures of PHD2 inhibitors.

5.3.3 Other 2OG Oxygenase Inhibitors

Potent inhibitors for other 2OG oxygenases, i.e. BBOX, OGFOD1 and histone demethylases, include mildronate, daminozide, CCT6-140, CCT2-66 and GSK-J1. [23-29]

Inhibitor	Structure	Comment
Mildronate		Competitive substrate of γ-butyrobetaine hydroxylase (BBOX). [23]
Daminozide		Human KDM2/7 histone demethylase inhibitor. [24, 25]
CCT6-140		Inhibitor of 2OG dependent prolyl-3-hydroxylase OGFOD1. [26]
CCT2-66		Histone demethylase inhibitor. [27]

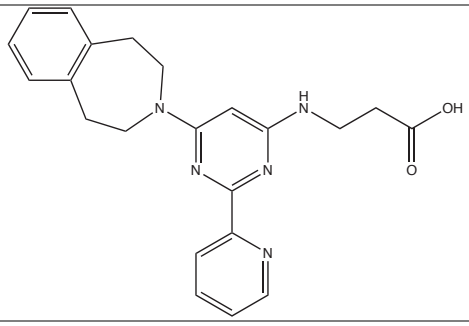
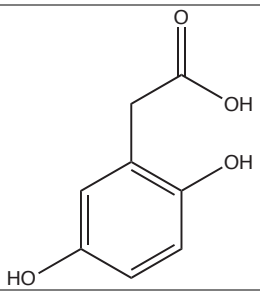
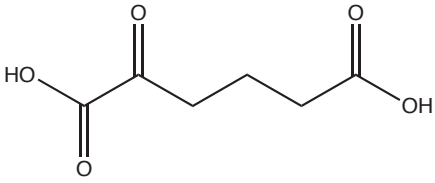
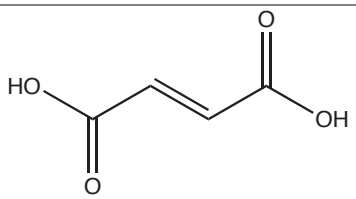
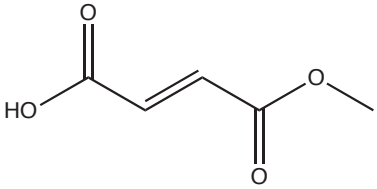
GSK-J1		Potent selective jumonji H3K27 demethylase inhibitor, which is selective for the KDM6 subfamily members JMJD3 and UTX. [28, 29]

Table 5.2 Structures of inhibitors for other 2OG oxygenases.

5.3.4 Endogenous Inhibitors of the 2OG Oxygenases and the Analogues

It was considered that cellular metabolic intermediates may also act as inhibitors for 2OG oxygenases. These inhibitors include homogentisic acid, 2-oxoadipic acid, fumaric acid, oxaloacetic acid, succinic acid and itaconic acid. The analog of fumaric acid, mono-methyl fumarate, is also a potential inhibitor. [30-33]

Inhibitor	Structure	Comment
Homogentisic acid		Intermediate of the metabolic breakdown of tyrosine and phenylalanine. [30]
2-oxoadipic acid		Important metabolite between the TCA cycle and lysine biosynthesis. Of interest for research on mitochondrial metabolite transporters. [31]
Fumaric acid		Intermediate in the TCA Cycle. [31]
Mono-methyl fumarate		Analogue of fumaric acid. [32]

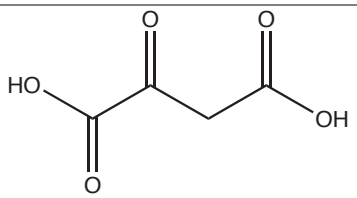
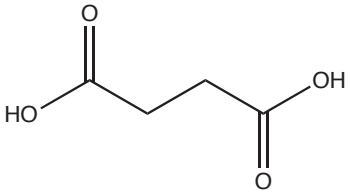
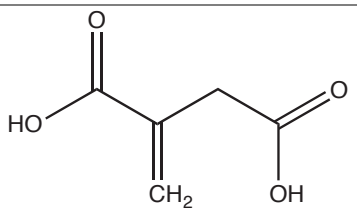
Oxaloacetic acid		Intermediate in the TCA Cycle. [31]
Succinic acid		Intermediate in the TCA Cycle. [31]
Itaconic acid		Intermediate in the C5-branched dibasic acid metabolism, a substrate for the enzyme succinate-CoA ligase. [33]

Table 5.3 Structures of endogenous inhibitors and analogues.

5.3.5 Others

Other inhibitors that were of interest to test included glutamic acid, D-(−)-isoascorbic acid and Meloxicam.

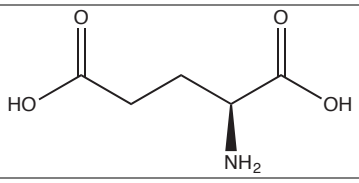
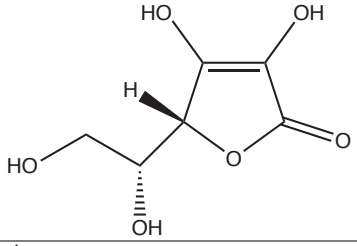
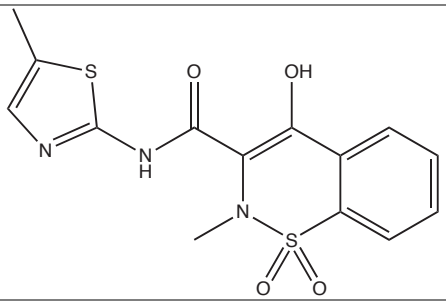
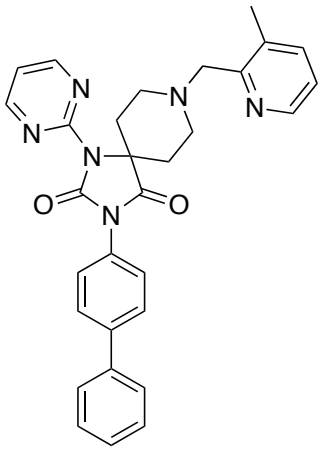
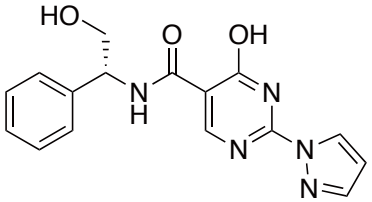
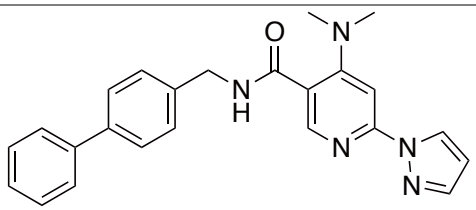
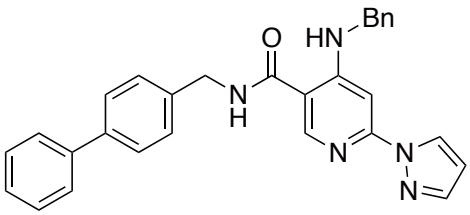
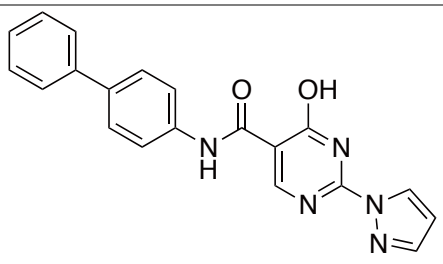
Name	Structure	
Glutamic acid		Amino acid. [34]
D-(−)-isoascorbic acid		Stereoisomer of ascorbic acid. [35]
Meloxicam		Cyclooxygenase (COX) inhibitor. COX is the enzyme responsible for converting arachidonic acid into prostaglandin. [36]

Table 5.4 Structures of other tested inhibitors.

5.3.6 Compounds from Dr. James P Holt -Martyn

Other compounds, which are potent PHD2 inhibitors or potential PHD2 activators in the Oxford laboratory, were kindly provided by Dr. James Holt-Martyn [37].

Inhibitor	Structure	Comment
2-167 (Labelled as Compound 10 in Chapter 3 of Dr. James P Holt –Martyn's Thesis)		Potent PHD2 inhibitor (IC_{50} = 2.1 nM) with strong selectivity over structurally similar 2OG oxygenases.
2-174 (Labelled as Compound 33 in Chapter 2 of Dr. James P Holt –Martyn's Thesis)		Inhibitor for PHD2 (IC_{50} = 0.329 µM) and vCPH (IC_{50} = 1.1 µM).
377 (Labelled as Compound 84 in Chapter 3 of Dr. James P Holt –Martyn's Thesis)		Potential PHD2 activator with a result of 61 % increase in substrate (CODD) hydroxylation by PHD2.
388 (Labelled as Compound 83 in Chapter 3 of Dr. James P Holt –Martyn's Thesis)		Inhibitor for PHD2 (IC_{50} = 14.0 µM).
3-201 (Labelled as Compound 32 in Chapter 2 and Chapter 13 of Dr. James P Holt –Martyn's Thesis)		Potent PHD2 inhibitor (IC_{50} = 0.258 nM) with strong selectivity over structurally similar 2OG oxygenases.

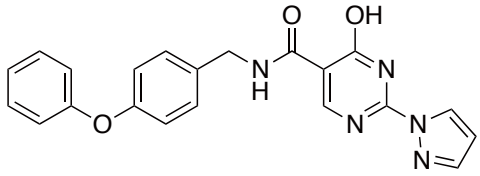
3-202 (Labelled as Compound 72 in Chapter 3 of Dr. James P Holt –Martyn's Thesis)		Compound that no longer maintains PHD2 inhibition.

Table 5.5 Structures of compounds from Dr. James Holt-Martyn [37].

5.4 Inhibition Screening of DdP4H Using Full-length DdSkp1 as the Substrate Analysis by RF-MS

IC_{50} assays were carried out under the conditions shown in Table 5.6. The acquisition parameters for the RF-MS machine are described in Chapter 4.17. In the presence of Fe (II), 2OG, ascorbate and oxygen, full-length DdSkp1 protein was observed to be hydroxylated by adding DdP4H. An 11-point and 3-fold (half log) dilution series was prepared from a 20 mM stock of inhibitor in DMSO. 250 nL of each dilution was dispensed in duplicate by acoustic transfer into a 384-well plate using an ECHO 550 acoustic dispenser (Labcyte). A DMSO control was dispensed into column 12 (16 wells) and a well characterised PHD2 inhibitor, FG-4592 (10 mM) [3, 12, 13], was transferred into column 24 (no enzyme control). The DMSO concentration was kept constant with a final assay concentration of 0.5 % (v/v) across the plate. DdP4H was diluted to 400 nM in assay buffer and 25 µL was added to each well. The plate was then pulse centrifuged for 3 seconds (1000 RPM), sealed and incubated 15 minutes at room temperature. The enzyme reaction was initiated by adding assay buffer containing cofactors and Skp1 substrates at 2 X final concentration (Table 5.6) and the plate centrifuged briefly and sealed. The enzyme reaction was allowed to progress for 45 minutes at room temperature and then stopped by adding 5 µL of 10 % (v/v) aqueous formic acid across the plate.

Plates were then transferred to a RapidFire RF360 high throughput sampling robot (Agilent) connected to an Agilent accurate mass 6530 QTOF. Samples were aspirated under vacuum for

500 ms and loaded onto a C4 SPE cartridge. The cartridge was then washed for 5000 ms with 0.1 % (v/v) aqueous formic acid in water at a flow rate of 1.5 mL/min after which Skp1 was eluted from the SPE cartridge with 85 % (v/v) acetonitrile in water + 0.1 % (v/v) aqueous formic acid for 5000 ms at a flow rate of 1.25 mL/min. The SPE cartridge was then re-equilibrated in 0.1 % (v/v) aqueous formic acid in water for 800 ms at a flow rate of 1.5 mL/min. Skp1 was detected in positive ESI mode with a fragmentor voltage of 200 V. Extracted ion chromatogram data for the Skp1 were integrated and % conversion of non-hydroxylated Skp1 to hydroxylatel Skp1 was calculated using the equation:

$$\% \text{ conversion} = 100 \times \text{hydroxylated} / (\text{hydroxylated} + \text{non-hydroxylated})$$

Data were analysed using Agilent MassHunter Workstation and GraphPad Prism 5. Assays were carried out in duplicate.

Assay Conditions	Final Concentration
2-oxoglutarate disodium salt (Sigma-Aldrich, K3752)	50 µM
Sodium L-ascorbate (Sigma-Aldrich, 11140)	100 µM
Ammonium iron(II) sulfate hexahydrate (Sigma-Aldrich, 215406)	10 µM
Full-length DdSkp1 protein substrate	5 µM
DdP4H enzyme	200 nM
Buffer (50 mM Tris, 50mM NaCl, pH 7.5)	-

Table 5.6 Assay conditions of DdP4H inhibitor screening samples.

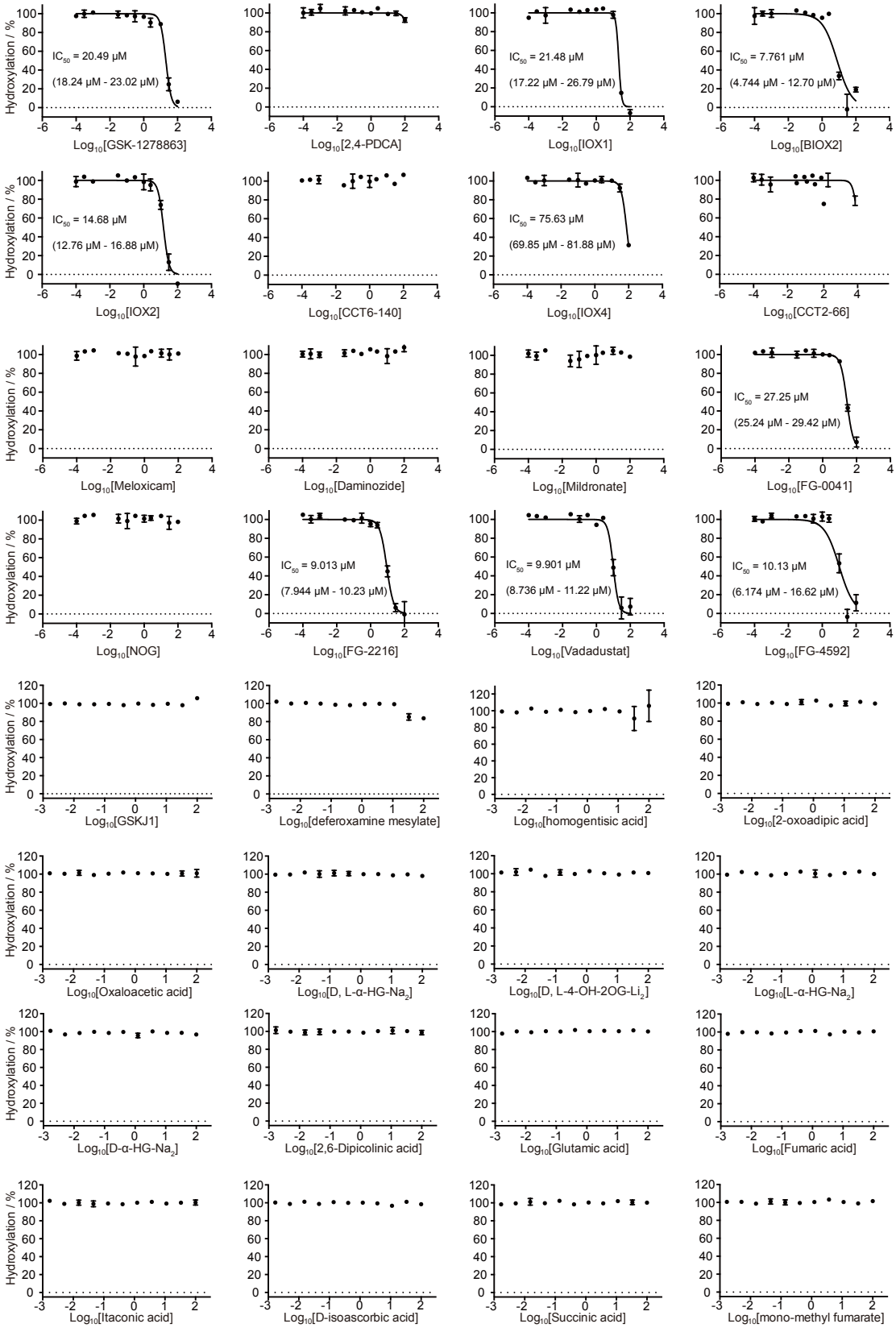


Fig. 5.1 IC_{50} values of DdP4H using full-length DdSkp1 as a substrate. IC_{50} values and 95 % confidence interval range are indicated.

Key observations from the results from of inhibitor screening are as follows:

- GSK-1278863 (IC_{50} = 20.49 μM), IOX1 (IC_{50} = 21.48 μM), BIOX2 (IC_{50} = 7.76 μM), IOX2 (IC_{50} = 14.68 μM), FG-0041 (IC_{50} = 27.25 μM), FG-2216 (IC_{50} = 9.01 μM), Vandadustat (IC_{50} = 9.90 μM) and FG-4592 (IC_{50} = 10.13 μM) showed relatively strong inhibition of DdP4H. By contrast, IOX4 was a weaker inhibitor (IC_{50} > 75.6 μM).
- Other 2OG analogues, including NOG, 2,4-PDCA, D,L-α-hydroxyglutaric acid, D,L-4-hydroxy-2-ketoglutaric acid, L-α-hydroxyglutaric acid, D-α-hydroxyglutaric acid, 2,6-dipicolinic acid, did not inhibit DdP4H (IC_{50} >> 100 μM).
- Potential endogenous inhibitors of the 2OG oxygenases and the analogues did not inhibit DdP4H. These inhibitors included homogentisic acid, 2-oxoadipic acid, fumaric acid, oxaloacetic acid, succinic acid, itaconic acid and mono-methyl fumarate.
- Other non prolyl hydroxylase 2OG oxygenase inhibitors, including mildronate, daminozide, CCT6-140, CCT2-66 and GSK-J1, did not inhibit DdP4H, suggesting selectivity of DdP4H for particular inhibitor types.
- Glutamic acid, D-(−)-isoascorbic acid or Meloxicam did not inhibit DdP4H.
- All the potent inhibitors for PHD2, except IOX4, which is a potent inhibitor for PHD2 (IC50 = 0.003 μM [38]), showed strong inhibition of DdP4H, consistent with the crystallographically observed similarity of DdP4H to PHD2.

5.5 Inhibitor Screening of TgPhyA Using TgSkp1 Peptide as the Substrate Analysed by RF-MS

The mass spectra data for full length TgSkp1 protein were of poor quality when using RF-MS. A second apparent oxidation of full length TgSkp1 could be observed in the spectra further complicating the analysis (Figure 4.13). Therefore, a synthetic TgSkp1 20-mer peptide (Severn Biotech Ltd.) was used for inhibition assays with TgPhyA to avoid non-specific oxidation observed when using full-length TgSkp1 protein. The sequence of TgSkp1 peptide was: FNIVNDFTPEEEAQVREENK (Mass: 2409 Da) with N-terminal amide.

IC_{50} assays were carried out using the same conditions as described in Chapter 5.4, except for the enzyme concentration, which is highlighted in Table 5.7. Samples were analysed by RF-MS. Data were analysed using Agilent MassHunter Workstation and GraphPad Prism 5. Assays were carried out in duplicate.

Assay Conditions	Final Concentration
2-oxoglutarate disodium salt (Sigma-Aldrich, K3752)	50 µM
Sodium L-ascorbate (Sigma-Aldrich, 11140)	100 µM
Ammonium iron(II) sulfate hexahydrate (Sigma-Aldrich, 215406)	10 µM
TgSkp1 peptide substrate	5 µM
TgPhyA enzyme	1 µM
Buffer (50 mM Tris, 50mM NaCl, pH 7.5)	-

Table 5.7 Assay conditions of TgPhyA inhibitor screening samples.

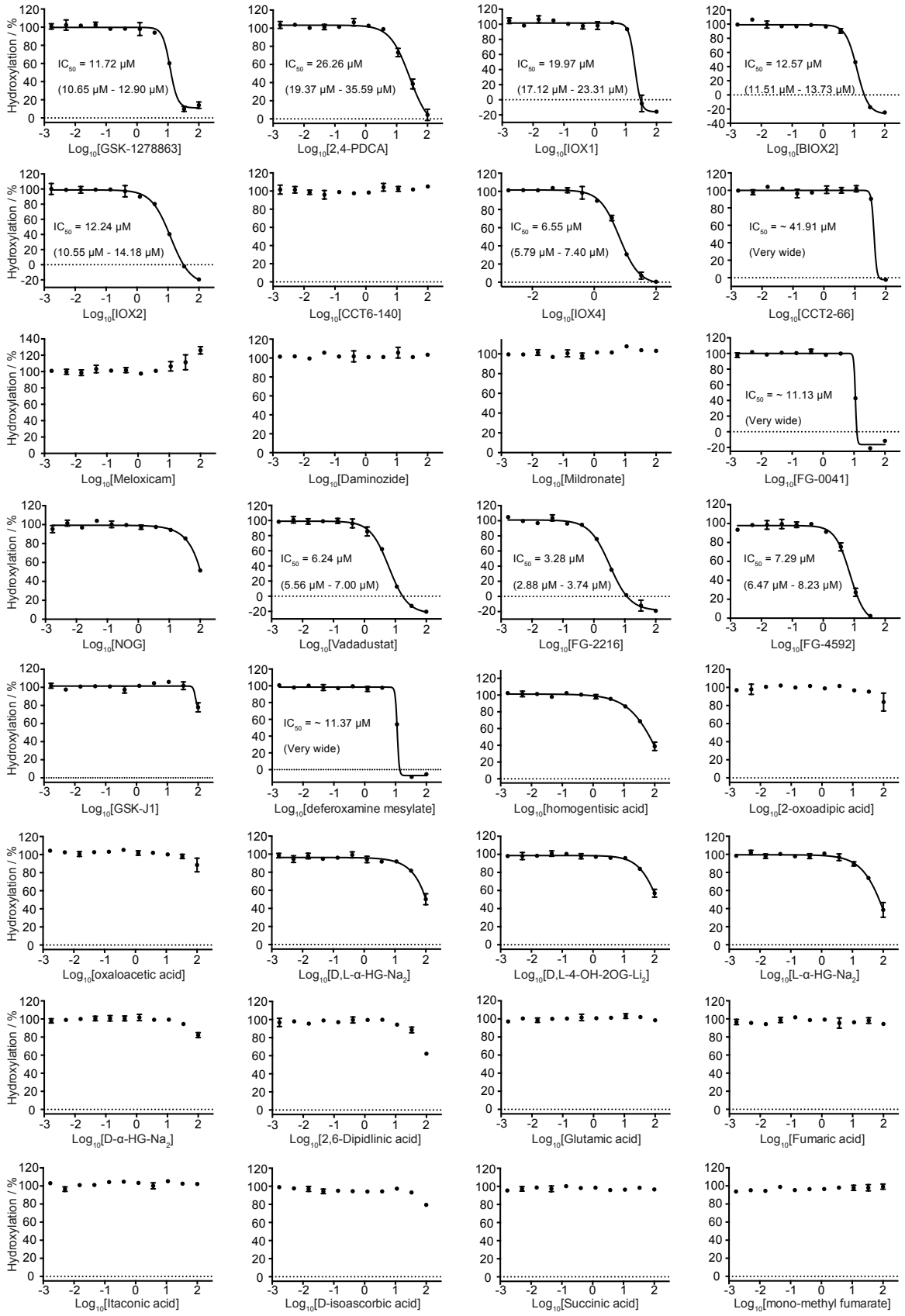


Fig. 5.2 IC_{50} values of TgPhyA using TgSkp1 peptide as a substrate. IC_{50} values and 95 % confidence interval range are indicated in each figure.

Key observations arising from the in results from the inhibitor screening are as follows:

- GSK-1278863 (IC_{50} = 11.72 μM), 2,4-PDCA (IC_{50} = 26.26 μM), IOX1 (IC_{50} = 19.97 μM), BIOX2 (IC_{50} = 12.57 μM), IOX2 (IC_{50} = 12.24 μM), IOX4 (IC_{50} = 6.55 μM), FG-2216 (IC_{50} = 3.28 μM), Vandadustat (IC_{50} = 6.24 μM) and FG-4592 (IC_{50} = 7.29 μM) showed strong inhibition to TgPhyA.
- CCT2-66 (IC_{50} = 41.91 μM), FG-0041 (IC_{50} = 11.13 μM), deferoxamine mesylate (IC_{50} = 11.37 μM) also showed strong inhibition, although accurate IC_{50}s could not be acquired due to the inability to perform curve fitting using software Prism.
- Some other 2OG analogues, including NOG, D,L-α-hydroxyglutaric acid, D,L-4-hydroxy-2-ketoglutaric acid, L-α-hydroxyglutaric acid, D-α-hydroxyglutaric acid, 2,6-dipicolinic acid, only show inhibition at high inhibitor concentration (IC_{50} > 100 μM).
- Endogenous inhibitors and the analogues of 2OG did not inhibit TgPhyA. These inhibitors include homogentisic acid, 2-oxoadipic acid, fumaric acid, oxaloacetic acid, succinic acid, itaconic acid and mono-methyl fumarate (IC_{50} >> 100 μM).
- Other 2OG oxygenase inhibitors, including mildronate, daminozide, CCT6-140, and GSK-J1, except CCT2-66, did not inhibit TgPhyA, suggesting strong selectivity of these inhibitors.
- Glutamic acid, D-(−)-isoascorbic acid or meloxicam did not inhibit TgPhyA, although meloxicam showed activation of TgPhyA at high concentration.
- All the potent inhibitors for PHD2 showed strong inhibition of TgPhyA, suggesting that TgPhyA has high structural similarity to PHD2.

5.6 Inhibitor Screening of DdP4H Using DdSkp1 Peptide as the Substrate Analysed by RF-MS

The inhibitor screening assays of DdP4H were tested using full-length DdSkp1 as the substrate, whereas the inhibitor screening assays of TgPhyA were conducted using TgSkp1 peptide. To understand if the results of the two sets of IC_{50}s are comparable, it was of interest to investigate whether the inhibition observed was different between using a peptide substrate or a protein substrate. Therefore, a synthetic DdSkp1 peptide (Severn Biotech Ltd.) was used for inhibition assays of DdP4H. The sequence of the DdSkp1 peptide used was: FNIKNDFTPEEEEQIRKENEWCEDK (Mass: 3170 Da) with a C-terminal amide.

IC_{50} assays were carried out under the same conditions as described in Chapter 5.4. Samples were analysed by RF-MS. Data were analysed using Agilent MassHunter Workstation and GraphPad Prism 5. Assays were carried out in duplicate.

Assay Conditions	Final Concentration
2-oxoglutarate disodium salt (Sigma-Aldrich, K3752)	50 µM
Sodium L-ascorbate (Sigma-Aldrich, 11140)	100 µM
Ammonium iron(II) sulfate hexahydrate (Sigma-Aldrich, 215406)	10 µM
DdSkp1 peptide substrate	5 µM
DdP4H enzyme	200 nM
Buffer (50 mM Tris, 50mM NaCl, pH 7.5)	-

Table 5.8 Assay conditions of DdP4H inhibitor screening samples with DdSkp1 peptide as the substrate.

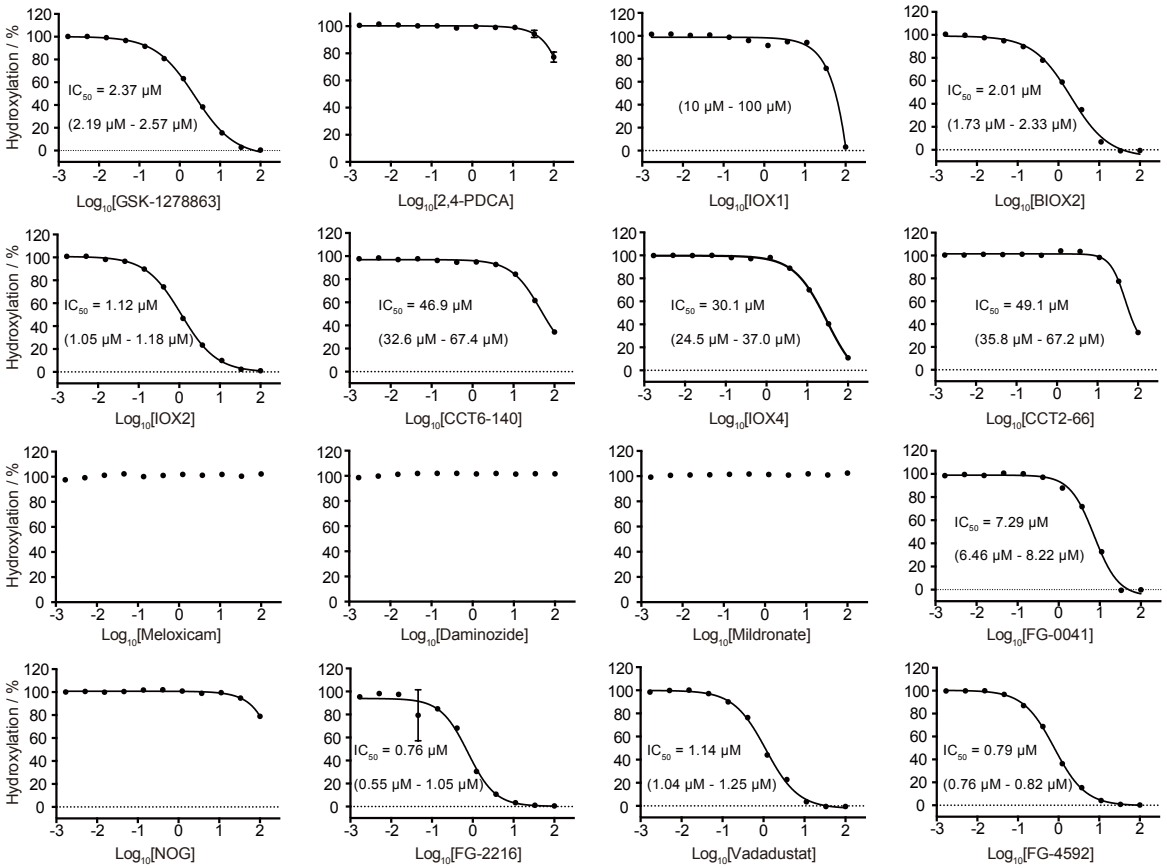


Fig. 5.3 Inhibition assays of DdP4H using DdSkp1 peptide as substrate. IC_{50} values and 95 % confidence interval range are indicated in each figure.

Key observation in results from these inhibitor screens are as follows:

- GSK-1278863 (IC_{50} = 2.37 μM), BIOX2 (IC_{50} = 7.76 μM), IOX2 (IC_{50} = 1.12 μM), FG-0041 (IC_{50} = 7.29 μM), FG-2216 (IC_{50} = 0.76 μM), Vandadustat (IC_{50} = 1.14 μM) and FG-4592 (IC_{50} = 0.79 μM) showed strong inhibition to DdP4H. IOX4 (IC_{50} > 30.1 μM), CCT2-66 (IC_{50} > 46.9 μM), CCT6-140 (IC_{50} > 49.1 μM) exhibited minor inhibition effect at high concentration.
- IOX1 (IC_{50} > 10 μM) also showed strong inhibition. However, an accurate IC_{50} could not be acquired due to the inability to perform curve fitting using software Prism
- Other 2OG analogues, including NOG and 2,4-PDCA did not inhibit DdP4H.
- Mildronate, daminozide or meloxicam did not inhibit DdP4H.

5.7 Discussion of Inhibitor Screening Results

5.7.1 Comparison of IC_{50}s Using Full-Length DdSkp1 Protein and DdSkp1 Peptide as the Substrate for DdP4H

The IC_{50}s using full-length DdSkp1 protein and DdSkp1 peptide as the substrates for DdP4H are compared in parallel in Table 5.9.

Enzyme Substrate	DdP4H (200 nM) DdSkp1 protein	DdP4H (200 nM) DdSkp1 peptide (5 μM)
GSK1278863	20.5 (18.2-23.0)	2.37 (2.19-2.57)
2,4-PDCA	×	×
IOX1	21.5 (17.2-26.8)	inaccurate
BIOX2	7.76 (4.74-12.70)	2.01 (1.73-2.33)
IOX2	14.7 (12.8-16.9)	1.12 (1.05-1.18)
CCT6-140	×	46.9 (32.6-67.4)
IOX4	75.6 (69.9-81.9)	30.1 (24.5-37.0)
CCT2-66	×	49.1 (35.8-67.2)
Meloxicam	×	×
Daminozide	×	×
Mildronate	×	×
FG-0041	27.3 (25.2-29.4)	7.29 (6.46-8.22)
NOG	×	×
FG-2216	9.01 (7.94-10.23)	0.757 (0.55-1.05)
Vadadustat	9.9 (8.7-11.2)	1.14 (1.04-1.25)
FG-4592	10.1 (6.2-16.6)	0.785 (0.76-0.82)

Table 5.9 IC_{50}s using full-length DdSkp1 protein and DdSkp1 peptide as the substrates for DdP4H. The units of the IC_{50}s are μM. IC_{50}s and 95 % confidence interval range for each inhibitor are compared in the same rows.

Key observation arising from the comparison results in Table 5.9 are:

- Almost all inhibitors show a similar inhibition pattern towards DdP4H when using full-length protein or peptide fragment as the substrate with peptide substrate having 10 fold lower IC_{50}s in general, including GSK1278863, IOX2, FG-2216 and FG-4592. The only exceptions are CCT2-66 and CCT6-140, which begin to exhibit inhibition at high inhibitor concentration when using DdSkp1 peptide as the substrate.
- The observation that inhibitors are more effective when using the DdSkp1 peptide substrate indicates that full-length DdSkp1 protein is a better substrate than DdSkp1

peptide. With the full-length protein substrate, a higher concentration of inhibitors was required to inhibit the hydroxylation reaction. This is consistent with the results of the kinetic assays described in Chapter 4.6 and Chapter 4.15, that a more valid K_m value was able to determine when using full-length DdSkp1 protein as the substrate, which may because other regions other than the active site of DdSkp1 protein take part in substrate binding, possibly reflecting a reduced binding ability of the peptide substrate due to the shorter fragment length.

- The results also suggest that we should be careful when doing *in vitro* inhibition assays using peptide fragments as substrates. A full-length protein substrate will naturally mimic the *in vivo* condition better than a peptide substrate. Higher IC_{50} values are likely observed using full-length substrates, more relevant to the *in vivo* situation and will be more successful for identifying useful inhibitors.

5.7.2 Comparison of IC_{50}s of DdP4H and TgPhyA

The IC_{50}s of DdP4H and TgPhyA are compared in Table 5.10.

Enzyme Substate	DdP4H (200 nM) DdSkp1 peptide (5 µM)	TgPhyA (1 µM) TgSkp1 peptide (5 µM)
GSK1278863	20.5 (18.2-23.0)	11.7 (10.7-12.9)
2,4-PDCA	×	26.3 (19.4-35.6)
IOX1	21.5 (17.2-26.8)	20.0 (17.1-23.3)
BIOX2	7.76 (4.74-12.70)	12.6 (11.5-13.7)
IOX2	14.7 (12.8-16.9)	12.2 (10.6-14.2)
CCT6-140	×	×
IOX4	75.6 (69.9-81.9)	6.55 (5.79-7.40)
CCT2-66	×	~42 (very wide)
Meloxicam	×	×
Daminozide	×	×
Mildronate	×	×
FG-0041	27.3 (25.2-29.4)	~11.1 (very wide)
NOG	×	×
FG-2216	9.01 (7.94-10.23)	3.28 (2.88-3.74)
Vadadustat	9.9 (8.7-11.2)	6.24 (5.56-7.00)
FG-4592	10.1 (6.2-16.6)	7.29 (6.47-8.23)

Table 5.10 Comparison of IC_{50}s of DdP4H and of TgPhyA using DdSkp1 / TgSkp1 peptide respectively as the substrates. IC_{50}s and 95 % confidence interval range for each inhibitor are compared in the same rows.

Key observations arising from the comparison results in Table 5.10 are as follows:

- Most inhibitors showed similar inhibition towards DdP4H and TgPhyA. GSK1278863, IOX1, BIOX2, IOX2, FG-0041, FG-2216, Vadadustat, and FG-4592 showed potent inhibition for both enzymes. IOX4 is only potent for TgPhyA. Meloxicam, daminozide, mildronate and NOG showed no inhibition for the enzymes. The only exceptions were 2,4-PDCA and CCT6-140. 2,4-PDCA showed no inhibition towards DdP4H. CCT6-140 showed no inhibition to TgPhyA. The similarity of IC_{50}s between DdP4H and TgPhyA suggest the structural similarity of these two enzymes.
- IC_{50}s numbers between these two enzymes may not be comparable due to different enzyme concentration. Because the concentration of DdP4H is 5 fold less than TgPhyA, most IC_{50}s of DdP4H are lower than TgPhyA. The exceptions are IOX1, IOX4 and CCT2-66, suggesting these inhibitors exhibit stronger inhibition effect towards TgPhyA than DdP4H.

5.8 Inhibitor Screening of DdP4H and TgPhyA Using Novel PHD2 Inhibitors

Inhibition assays of compounds synthesised by Dr. James Holt-Martyn were carried out under the conditions shown in Table 5.11 [37]. In the presence of Fe (II), 2OG, ascorbate and oxygen, full-length DdSkp1 / TgSkp1 protein, the reactions were initiated by adding DdP4H or TgPhyA. Two inhibitor concentrations (10 μM and 100 μM) were tested. The reactions were incubated at room temperature for 1 hour and were quenched with 1% (v/v) aqueous formic acid, before being subjected to analysis by Xevo G2-S QToF mass spectrometer. The acquisition parameters of Xevo G2-S QToF mass spectrometer are described in Chapter 4.10. Data were analysed using

MassLynx 4.1 software (Waters, Wilmslow UK) and GraphPad Prism 5. Assays were carried out in triplicate.

Assay Conditions	Final Concentration
α-ketoglutaric acid disodium salt (Sigma-Aldrich, K3752)	500 µM
Sodium L-ascorbate (Sigma-Aldrich, 11140)	1 mM
Ammonium iron(II) sulfate hexahydrate (Sigma-Aldrich, 215406)	50 µM
Full-length DdSkp1 / TgSkp1 protein substrate	100 µM
DdP4H / TgPhyA enzyme	1 µM
Buffer (100 mM HEPES, pH 7.6)	-

Table 5.11 Assay conditions of inhibition assay samples.

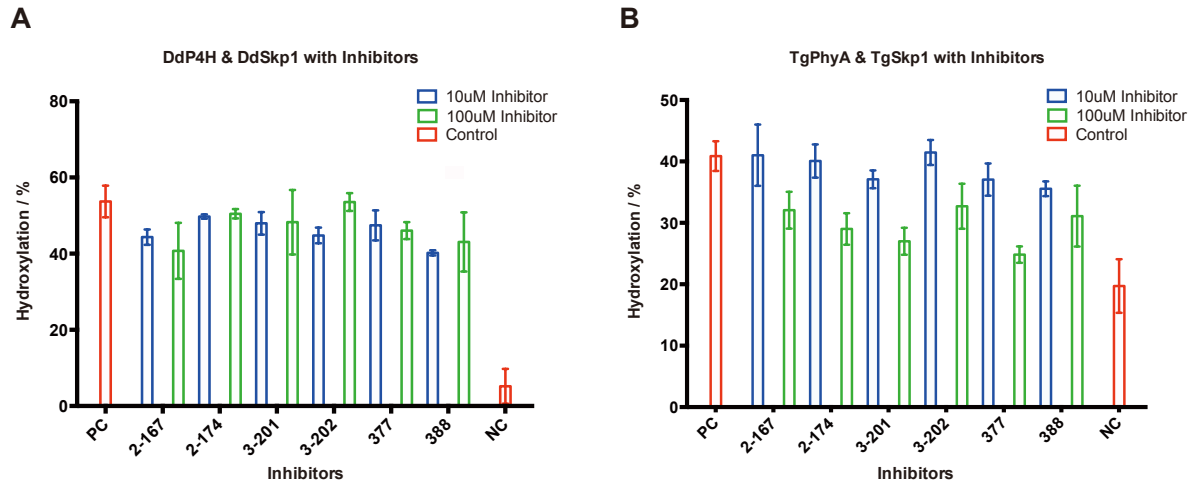


Fig. 5.4 Inhibitor screen of DdP4H and TgPhyA using the inhibitors from Dr. James Holt-Martyn. (A) Inhibitor screening using full-length DdSkp1 protein as the substrate for DdP4H. (B) Inhibitor screening using full-length TgSkp1 protein as the substrate for TgPhyA. Hydroxylation percentage is indicated as bars at two different inhibitor concentrations. Positive Control (PC, red) and Negative Control (NC, red) were performed at the same time with no inhibitor and 100 µM GSK1278863.

The results shown Fig. 5.4 (A) indicate that none of the compounds inhibit DdP4H, even at a concentration of 100 µM. Considering the potent inhibition effects or activation effects of these inhibitors on PHD2, this result reinforces the strong selectivity of these compounds for human PHD2.

Fig. 5.4 (B) shows most compounds are less potent agonist of TgPhyA at a compound concentration of 100 µM. Although compound 377 was reported (personal communication, Dr. James Holt-Martyn) to be an activator of PHD2, it inhibits TgPhyA, suggesting distinct structural

differences exist between PHD2 and TgPhyA. This result also reinforces the selectivity compound 377 has for PHD2 over other prolyl hydroxylases.

Note: around 20% of TgSkp1 was hydroxylated in the negative control sample, due to non specific oxidation of TgSkp1. More on this phenomenon was discussed in Chapter 4.12. Further optimisation is required for accurate IC_{50}s if the full-length TgSkp1 protein is to be used as the substrate of TgPhyA.

5.9 References

1. Rose, N. R., McDonough, M. A., King, O. N., Kawamura, A., and Schofield, C. J. (2011) Inhibition of 2-oxoglutarate dependent oxygenases, *Chem Soc Rev 40*, 4364-4397.

2. Schofield, C. J., and Zhang, Z. (1999) Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes, *Curr Opin Struct Biol 9*, 722-731.

3. Hsieh, M. M., Linde, N. S., Wynter, A., Metzger, M., Wong, C., Langsetmo, I., Lin, A., Smith, R., Rodgers, G. P., Donahue, R. E., Klaus, S. J., and Tisdale, J. F. (2007) HIF prolyl hydroxylase inhibition results in endogenous erythropoietin induction, erythrocytosis, and modest fetal hemoglobin expression in rhesus macaques, *Blood 110*, 2140-2147.

4. Fraisl, P., Aragones, J., and Carmeliet, P. (2009) Inhibition of oxygen sensors as a therapeutic strategy for ischaemic and inflammatory disease, *Nat Rev Drug Discov 8*, 139-152.

5. Bao, W., Qin, P., Needle, S., Erickson-Miller, C. L., Duffy, K. J., Ariazi, J. L., Zhao, S., Olzinski, A. R., Behm, D. J., Pipes, G. C., Jucker, B. M., Hu, E., Lepore, J. J., and Willette, R. N. (2010) Chronic inhibition of hypoxia-inducible factor prolyl 4-hydroxylase improves ventricular performance, remodeling, and vascularity after myocardial infarction in the rat, *J Cardiovasc Pharmacol 56*, 147-155.

6. Semenza, G. L. (2003) Targeting HIF-1 for cancer therapy, *Nat Rev Cancer 3*, 721-732.

7. Masoud, G. N., and Li, W. (2015) HIF-1alpha pathway: role, regulation and intervention for cancer therapy, *Acta Pharm Sin B 5*, 378-389.

8. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis, *Nat Protoc 10*, 845-858.

9. Kawamura, A., Munzel, M., Kojima, T., Yapp, C., Bhushan, B., Goto, Y., Tumber, A., Katoh, T., King, O. N., Passioura, T., Walport, L. J., Hatch, S. B., Madden, S., Muller, S., Brennan, P. E., Chowdhury, R., Hopkinson, R. J., Suga, H., and Schofield, C. J. (2017)

Highly selective inhibition of histone demethylases by de novo macrocyclic peptides, *Nat Commun 8*, 14773.

10. Tschank, G., Raghunath, M., Gunzler, V., and Hanauske-Abel, H. M. (1987) Pyridinedicarboxylates, the first mechanism-derived inhibitors for prolyl 4-hydroxylase, selectively suppress cellular hydroxyprolyl biosynthesis. Decrease in interstitial collagen and Clq secretion in cell culture, *Biochem J 248*, 625-633.

11. Hopkinson, R. J., Tumber, A., Yapp, C., Chowdhury, R., Aik, W., Che, K. H., Li, X. S., Kristensen, J. B. L., King, O. N. F., Chan, M. C., Yeoh, K. K., Choi, H., Walport, L. J., Thinnes, C. C., Bush, J. T., Lejeune, C., Rydzik, A. M., Rose, N. R., Bagg, E. A., McDonough, M. A., Krojer, T., Yue, W. W., Ng, S. S., Olsen, L., Brennan, P. E., Oppermann, U., Muller-Knapp, S., Klose, R. J., Ratcliffe, P. J., Schofield, C. J., and Kawamura, A. (2013) 5-Carboxy-8-hydroxyquinoline is a Broad Spectrum 2-Oxoglutarate Oxygenase Inhibitor which Causes Iron Translocation, *Chem Sci 4*, 3110-3117.

12. Hong, Y. R., Kim, H. T., Lee, S. C., Ro, S., Cho, J. M., Kim, I. S., and Jung, Y. H. (2013) [(4-Hydroxyl-benzo[4,5]thieno[3,2-c]pyridine-3-carbonyl)-amino]-acetic acid derivatives; HIF prolyl 4-hydroxylase inhibitors as oral erythropoietin secretagogues, *Bioorg Med Chem Lett 23*, 5953-5957.

13. Chowdhury, R., Leung, I. K., Tian, Y. M., Abboud, M. I., Ge, W., Domene, C., Cantrelle, F. X., Landrieu, I., Hardy, A. P., Pugh, C. W., Ratcliffe, P. J., Claridge, T. D., and Schofield, C. J. (2016) Structural basis for oxygen degradation domain selectivity of the HIF prolyl hydroxylases, *Nat Commun 7*, 12673.

14. Besarab, A., Provenzano, R., Hertel, J., Zabaneh, R., Klaus, S. J., Lee, T., Leong, R., Hemmerich, S., Yu, K. H., and Neff, T. B. (2015) Randomized placebo-controlled dose-ranging and pharmacodynamics study of roxadustat (FG-4592) to treat anemia in nondialysis-dependent chronic kidney disease (NDD-CKD) patients, *Nephrol Dial Transplant 30*, 1665-1673.

15. Besarab, A., Chernyavskaya, E., Motylev, I., Shutov, E., Kumbar, L. M., Gurevich, K., Chan, D. T., Leong, R., Poole, L., Zhong, M., Saikali, K. G., Franco, M., Hemmerich, S., Yu, K. H., and Neff, T. B. (2016) Roxadustat (FG-4592): Correction of Anemia in Incident Dialysis Patients, *J Am Soc Nephrol 27*, 1225-1233.

16. Chowdhury, R., Candela-Lena, J. I., Chan, M. C., Greenald, D. J., Yeoh, K. K., Tian, Y. M., McDonough, M. A., Tumber, A., Rose, N. R., Conejo-Garcia, A., Demetriades, M., Mathavan, S., Kawamura, A., Lee, M. K., van Eeden, F., Pugh, C. W., Ratcliffe, P. J., and Schofield, C. J. (2013) Selective small molecule probes for the hypoxia inducible factor (HIF) prolyl hydroxylases, *ACS Chem Biol 8*, 1488-1496.

17. (2016) Burzlaff, N. (Friedrich Alexander University Erlangen-Nürnberg).

18. Chan, M. C., Atasoylu, O., Hodson, E., Tumber, A., Leung, I. K., Chowdhury, R., Gomez-Perez, V., Demetriades, M., Rydzik, A. M., Holt-Martyn, J., Tian, Y. M., Bishop, T., Claridge, T. D., Kawamura, A., Pugh, C. W., Ratcliffe, P. J., and Schofield, C. J. (2015)

Potent and Selective Triazole-Based Inhibitors of the Hypoxia-Inducible Factor Prolyl-Hydroxylases with Activity in the Murine Brain, *PLoS One 10*, e0132004.

19. Pergola, P. E., Spinowitz, B. S., Hartman, C. S., Maroni, B. J., and Haase, V. H. (2016) Vadadustat, a novel oral HIF stabilizer, provides effective anemia treatment in nondialysis-dependent chronic kidney disease, *Kidney Int 90*, 1115-1122.

20. Martin, E. R., Smith, M. T., Maroni, B. J., Zuraw, Q. C., and deGoma, E. M. (2017) Clinical Trial of Vadadustat in Patients with Anemia Secondary to Stage 3 or 4 Chronic Kidney Disease, *Am J Nephrol 45*, 380-388.

21. Holdstock, L., Meadowcroft, A. M., Maier, R., Johnson, B. M., Jones, D., Rastogi, A., Zeig, S., Lepore, J. J., and Cobitz, A. R. (2016) Four-Week Studies of Oral Hypoxia-Inducible Factor-Prolyl Hydroxylase Inhibitor GSK1278863 for Treatment of Anemia, *J Am Soc Nephrol 27*, 1234-1244.

22. Franklin, T. J., Morris, W. P., Edwards, P. N., Large, M. S., and Stephenson, R. (2001) Inhibition of prolyl 4-hydroxylase in vitro and in vivo by members of a novel series of phenanthrolinones, *Biochem J 353*, 333-338.

23. Leung, I. K., Krojer, T. J., Kochan, G. T., Henry, L., von Delft, F., Claridge, T. D., Oppermann, U., McDonough, M. A., and Schofield, C. J. (2010) Structural and mechanistic studies on gamma-butyrobetaine hydroxylase, *Chem Biol 17*, 1316-1324.

24. Rademacher, W. (2000) GROWTH RETARDANTS: Effects on Gibberellin Biosynthesis and Other Metabolic Pathways, *Annu Rev Plant Physiol Plant Mol Biol 51*, 501-531.

25. Rose, N. R., Woon, E. C., Tumber, A., Walport, L. J., Chowdhury, R., Li, X. S., King, O. N., Lejeune, C., Ng, S. S., Krojer, T., Chan, M. C., Rydzik, A. M., Hopkinson, R. J., Che, K. H., Daniel, M., Strain-Damerell, C., Gileadi, C., Kochan, G., Leung, I. K., Dunford, J., Yeoh, K. K., Ratcliffe, P. J., Burgess-Brown, N., von Delft, F., Muller, S., Marsden, B., Brennan, P. E., McDonough, M. A., Oppermann, U., Klose, R. J., Schofield, C. J., and Kawamura, A. (2012) Plant growth regulator daminozide is a selective inhibitor of human KDM2/7 histone demethylases, *J Med Chem 55*, 6639-6643.

26. Thinnes, C. C. (2014) Chemical and biological studies on human oxygenases, In *Oxford University, UK.*

27. Thinnes, C. C., Tumber, A., Yapp, C., Scozzafava, G., Yeh, T., Chan, M. C., Tran, T. A., Hsu, K., Tarhonskaya, H., Walport, L. J., Wilkins, S. E., Martinez, E. D., Muller, S., Pugh, C. W., Ratcliffe, P. J., Brennan, P. E., Kawamura, A., and Schofield, C. J. (2015) Betti reaction enables efficient synthesis of 8-hydroxyquinoline inhibitors of 2-oxoglutarate oxygenases, *Chem Commun (Camb) 51*, 15458-15461.

28. Kruidenier, L., Chung, C. W., Cheng, Z., Liddle, J., Che, K., Joberty, G., Bantscheff, M., Bountra, C., Bridges, A., Diallo, H., Eberhard, D., Hutchinson, S., Jones, E., Katso, R., Leveridge, M., Mander, P. K., Mosley, J., Ramirez-Molina, C., Rowland, P., Schofield, C. J., Sheppard, R. J., Smith, J. E., Swales, C., Tanner, R., Thomas, P., Tumber, A., Drewes, G., Oppermann, U., Patel, D. J., Lee, K., and Wilson, D. M. (2012) A selective

jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response, *Nature 488*, 404-408.

29. Yapp, C., Carr, A. J., Price, A., Oppermann, U., and Snelling, S. J. (2016) H3K27me3 demethylases regulate in vitro chondrogenesis and chondrocyte activity in osteoarthritis, *Arthritis Res Ther 18*, 158.

30. Hegedus, Z. L., and Nayak, U. (1994) Homogentisic acid and structurally related compounds as intermediates in plasma soluble melanin formation and in tissue toxicities, *Arch Int Physiol Biochim Biophys 102*, 175-181.

31. Koivunen, P., Hirsila, M., Remes, A. M., Hassinen, I. E., Kivirikko, K. I., and Myllyharju, J. (2007) Inhibition of hypoxia-inducible factor (HIF) hydroxylases by citric acid cycle intermediates: possible links between cell metabolism and stabilization of HIF, *J Biol Chem 282*, 4524-4532.

32. Yao, Y., Miao, W., Liu, Z., Han, W., Shi, K., Shen, Y., Li, H., Liu, Q., Fu, Y., Huang, D., and Shi, F. D. (2016) Dimethyl Fumarate and Monomethyl Fumarate Promote Post-Ischemic Recovery in Mice, *Transl Stroke Res 7*, 535-547.

33. Steiger, M. G., Blumhoff, M. L., Mattanovich, D., and Sauer, M. (2013) Biochemistry of microbial itaconic acid production, *Front Microbiol 4*, 23.

34. Walker, M. C., and van der Donk, W. A. (2016) The many roles of glutamate in metabolism, *J Ind Microbiol Biotechnol 43*, 419-430.

35. Hornig, D., and Weiser, H. (1976) Interaction of erythorbic acid with ascorbic acid catabolism, *Int J Vitam Nutr Res 46*, 40-47.

36. Auvinet, B., Ziller, R., Appelboom, T., and Velicitat, P. (1995) Comparison of the onset and intensity of action of intramuscular meloxicam and oral meloxicam in patients with acute sciatica, *Clin Ther 17*, 1078-1098.

37. Holt-Martyn, J. (2018) Novel and Selective Small Molecule Inhibitors and Activators For The Prolyl Hydroxylase Domain Enzymes, In *Department of Organic Chemistry*, University of Oxford.

38. Yeh, T. L., Leissing, T. M., Abboud, M. I., Thinnes, C. C., Atasoylu, O., Holt-Martyn, J. P., Zhang, D., Tumber, A., Lippl, K., Lohans, C. T., Leung, I. K. H., Morcrette, H., Clifton, I. J., Claridge, T. D. W., Kawamura, A., Flashman, E., Lu, X., Ratcliffe, P. J., Chowdhury, R., Pugh, C. W., and Schofield, C. J. (2017) Molecular and cellular mechanisms of HIF prolyl hydroxylase inhibitors in clinical trials, *Chem Sci 8*, 7651-7668.

6. Cross Organism Hydroxylation Studies

6.1 Introduction

It is reported that the TgPhyA protein can prolyl hydroxylate both the *Toxoplasma gondii* and *Dictyostelium discoideum* Skp1s [1]. In addition, overexpression of TgPhyA in DdP4H-null *Dictyostelium discoideum* leads to hydroxylation of Skp1 and also apparently restores the O_2 sensing system [1]. These findings suggest a close evolutionary relationship between the prolyl hydroxylation mechanisms in *Toxoplasma gondii* and *Dictyostelium discoideum*. This chapter describes the results of studies with cross organism hydroxylation assays, which aimed to explore the potential evolutionary relationships of prolyl hydroxylation systems from different organisms. It was proposed that the differences in substrates for the prolyl hydroxylases may provide clues as to how evolution of oxygen sensing systems occurred.

6.2 Summary

Cross-organism hydroxylation assays were first performed between *Dictyostelium discoideum*, *Toxoplasma gondii* and human prolyl hydroxylases and their substrates (Chapter 6.3 - 6.4). Cross organism hydroxylation assays were widened to include prolyl hydroxylases and substrates from *Pseudomonas putida* (prolyl hydroxylase: PPHD, substrate: EF-Tu), *Monosiga brevicollis* (prolyl hydroxylase: MbP4H, hypothesized substrate: MbSkp1), and *Trichoplax adhaerens* (prolyl hydroxylase: TaPHD, substrate: TaODD and possibly TaSkp1) (Chapter 6.5), Although the results should be regarded as preliminary, due to the lack of detailed kinetic analysis, they suggest some interesting evolutionary possibilities.

6.3 Cross Organism Hydroxylation Between DdP4H, TgPhyA, Human PHD2

Cross organism hydroxylation assays were initially performed between three different organisms; *Dictyostelium discoideum*, *Toxoplasma gondii* and human prolyl hydroxyases with their native or potential substrates. In the presence of Fe (II), 2OG, ascorbate and oxygen, the full-length substrates, DdSkp1, TgSkp1 and HsSkp1 E147P, were tested with DdP4H, TgPhyA and PHD2. The reactions were incubated at 37 °C for 1 hour before being quenched with 1% (v/v) aqueous formic acid. Both Fe(II) and ascorbate were added to promote the catalytic activity. Samples were analysed using an LC-MS machine (Xevo G2-S).

Assay Conditions	Final Concentration	Volume
2-oxoglutarate disodium salt (Sigma-Aldrich, K3752)	500 µM	2 µL
Sodium L-ascorbate (Sigma-Aldrich, 11140)	1 mM	2 µL
Ammonium iron(II) sulfate hexahydrate (Sigma-Aldrich, 215406)	50 µM	2 µL
Full-length protein of DdSkp1, TgSkp1 or HsSkp1 E147P	100 µM	2 µL
DdP4H, TgPhyA or PHD2 enzyme	10 µM	2 µL
Buffer (100 mM HEPES, pH 7.6)	-	10 µL
Total volume	-	20 µL

Table 6.1 Assay conditions of hydroxylation assay samples.

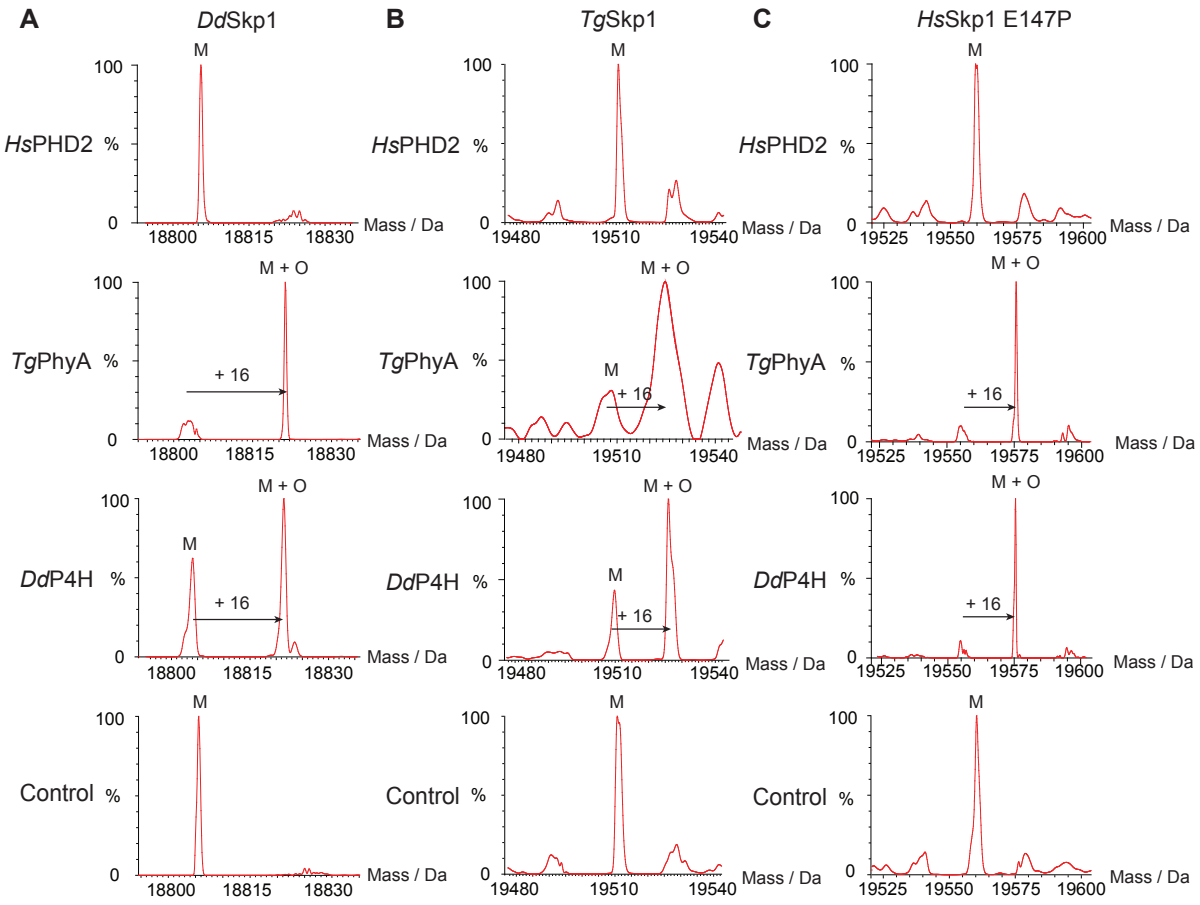


Fig. 6.1 Cross organism hydroxylation assays between *Dictyostelium discoideum*, *Toxoplasma gondii* and human. Prolyl hydroxylases from the three different organisms were incubated with the three different substrates for 1 h at 37 °C with the cofactor Fe(II) and the co-substrate 2OG before being quenched by 1 % (v/v) aqueous formic acid. Samples were analysed by MALDI-MS. (**A**) Full-length DdSkp1 protein tested with the three prolyl hydroxylases in comparison with the non-enzyme control sample below. (**B**) Full-length TgSkp1 protein. (**C**) Full-length HsSkp1 E147P protein.

DdP4H and TgPhyA are active on all three full-length protein Skp1s, although these substrates originate from different organisms. This result suggests a relatively weak substrate specificity of DdP4H and TgPhyA. However, it may also reflect the highly conserved sequence similarity of the Skp1s, especially in the proline-containing region (Figure 1.15 in Chapter 1). Human PHD2 did not accept Skp1 as a substrate, supporting the distant evolutionary relationships between the oxygen sensing systems of humans and these two organisms.

6.4 Screening of DdP4H and TgPhyA Against HIF Substrates

This part of the work was carried out in collaboration with Dr. Kerstin Lippl.

DdP4H and TgPhyA were further tested using additional substrates. These include TgSkp1 peptide (as the positive control), HIF-1α CODD, TaODD (TaHIF), HIF-1α NODD and CODD 2x M-A variant substrates tested with DdP4H and TgPhyA, in the presence of Fe (II), 2OG, ascorbate and oxygen. The assay conditions were the same as described in Table 6.1. The reactions were incubated at 37 °C for 1 hour before being quenched with 1% (v/v) aqueous formic acid. Samples were analysed by MALDI-MS.

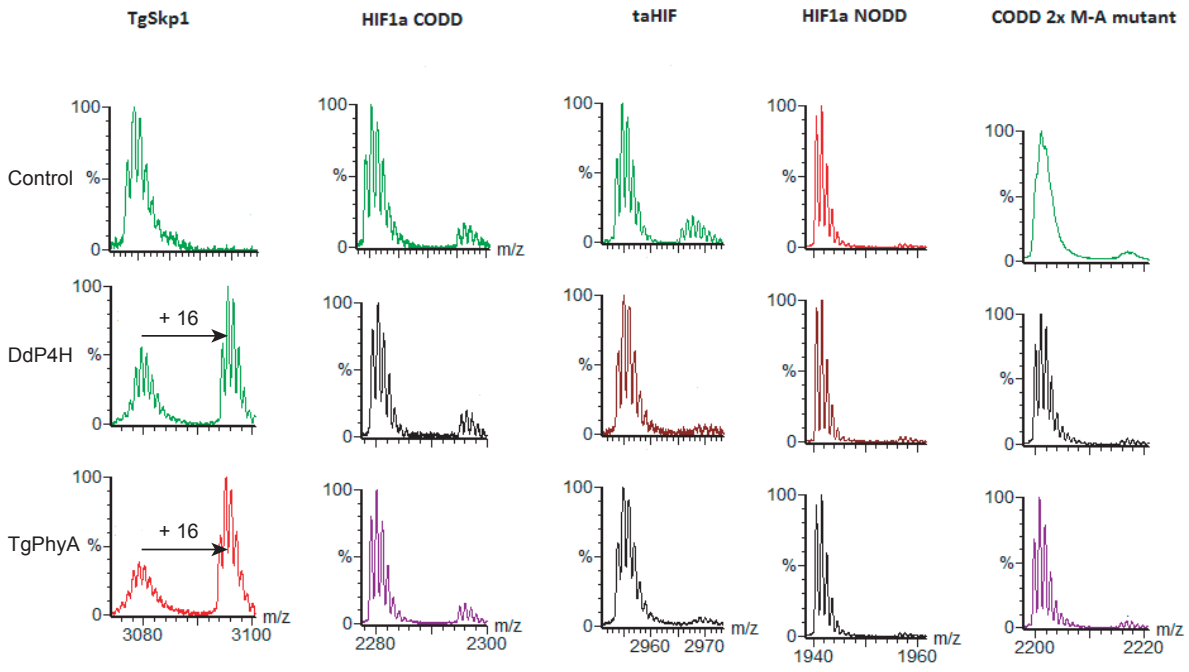


Fig. 6.2 DdP4H and TgPhyA assays using HIF substrates. DdP4H and TgPhyA were incubated with different substrates respectively for 1 h at 37 °C with cofactors before being quenched by 1% (v/v) aqueous formic acid. Samples were analysed by MALDI-MS. The figure is adapted from (Lippl, 2018) (permission acquired) [2].

The presence of a +16 Da peak in the samples with DdP4H + TgSkp1 or TgPhyA + TgSkp1 indicates that both enzymes were active. However, neither DdP4H nor TgPhyA was able to hydroxylate any of the HIF peptide substrates.

6.5 Cross Organism Hydroxylation Assays

This part of the work was carried out in collaboration with Dr. Kerstin Lippl.

Standard *in vitro* hydroxylation assays were carried out using the reaction mixture as shown in Table 6.2 with a final volume of 50 µL with each enzyme, as listed in Table 6.3, and substrates as detailed in Table 6.4. The enzyme to substrate ratio was 1:10, except for the MbP4H CODD and NODD hydroxylation assays, which were performed with an enzyme to substrate ratio of 1:4. Samples were quenched by 1% (v/v) aqueous formic acid after 1-hour incubation before being analysed by MALDI-MS. Assays were performed in triplicate.

Assay Conditions	Final Concentration	Volume
2-oxoglutarate acid disodium salt (Sigma-Aldrich, K3752)	300 µM	2.5 µl
Sodium L-ascorbate (Sigma-Aldrich, 11140)	4 mM	2 µL
Ammonium iron(II) sulfate hexahydrate (Sigma-Aldrich, 215406)	50 µM	1 µL
Substrates	100 µM	5 µL
Enzymes	10 µM (25 µM for MbP4H)	5 µL
Buffer (50 mM Tris pH 7.5)		34.5 µL
Total volume	-	50 µL

Table 6.2 Cross organism assay conditions [3].

Enzyme	Reported natural substrates
MbP4H	Unknown
PPHD	EF-Tu
TgPhyA	TgSkp1
DdP4H	DdSkp1
TaPHD*	TaHIF (TaODD)
PHD2*	HIF (NODD/CODD)

*Kindly provided by Dr. Kerstin Lippl.

Table 6.3 Enzymes used in cross organism assays [3].

Peptides and Hydroxylation Sites	Sequence	MW / Da
20mer EF-Tu-P54	IVEFDKIDSA**P**EEKARGITI-NH_2	2231.5
30mer_1 EF-Tu-P54	IVEFDKIDSA**P**EEKARGITINTAHVEYNST-NH_2	3348.6
30mer_2 EF-Tu-P54	VFGSAIVEFDKIDSA**P**EEKARGITINTAHV-NH_2	3215.6
30mer_3 EF-Tu-P54	RVCSEVFGSA IVEFDKIDSA**P**EEKARGITI-NH_2	3267.7
HIF-1α CODD-P564*	DLDLEMLA**P**YIPMDDDFQL-NH_2	2254.6
HIF-1α NODD-P402	DALTLLA**P**AAGDTIISLDF-NH_2	1916.2
MbSkp1-P150	IRDQFNIRNDFT**P**EEEERVQKENDW-NH_2	3208.4
TaSkp1-P153	IRRTFNIKNDFT**P**EEEAQVRKENEW-NH_2	3150.4
TaHIF-p486 (TaODD)	PINEKEDYDDLA**P**FVPPPSFDNRLY-NH_2	2950.4

*Kindly provided by Dr. Kerstin Lippl.

Table 6.4 Substrate peptides [3].

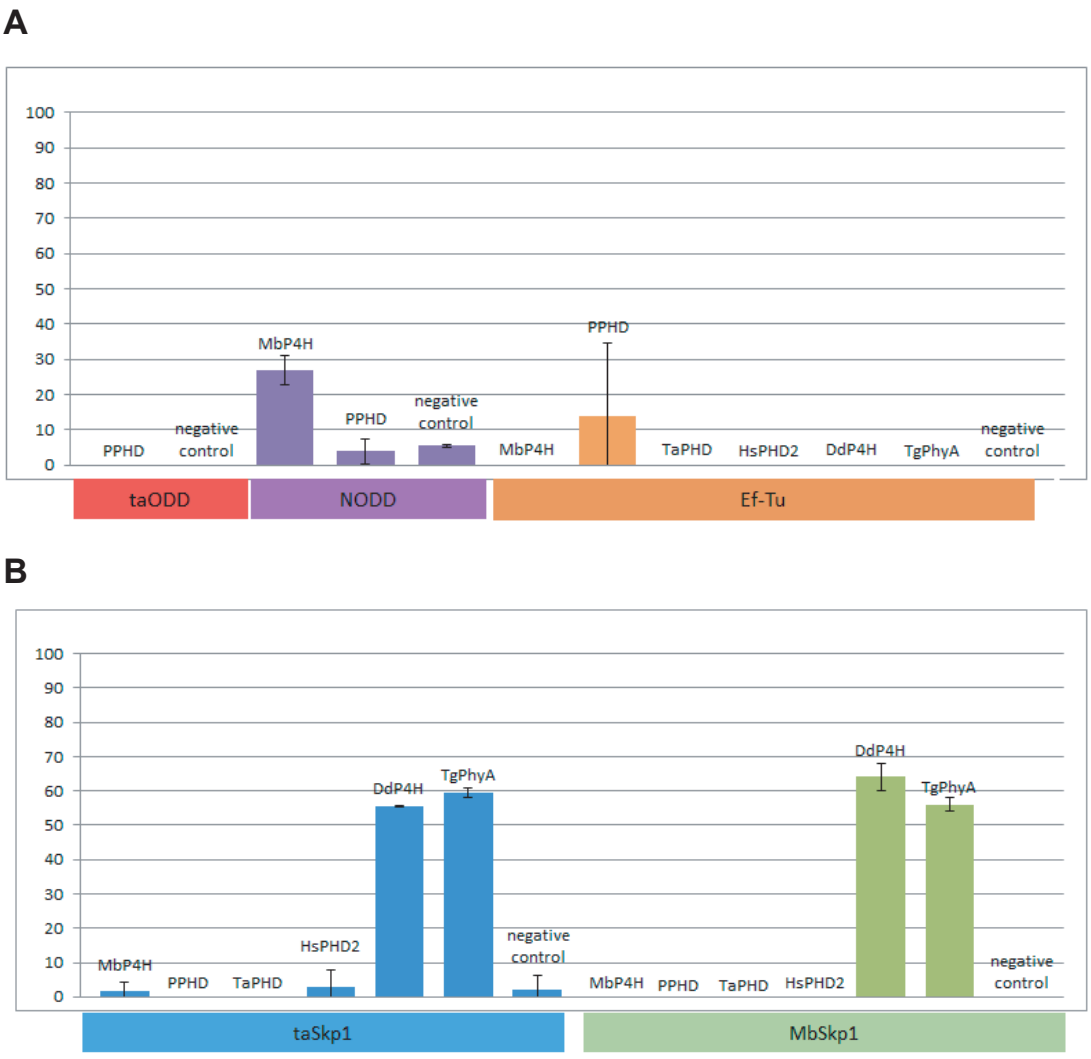


Fig. 6.3 Cross organism hydroxylation assays. PHD and PHD-like enzymes from various organisms were incubated with different peptides for 1 h at 37 °C with Fe(II) cofactor and 2OG co-substrate before being quenched by 1% (v/v) aqueous formic acid. Samples were analysed by MALDI-MS. (**A**) Hydroxylation results of TaODD tested with PPHD and without enzyme (red); NODD tested with MbP4H, PPHD and without enzyme (violet); EF-Tu tested with MbP4H, PPHD, TaPHD, HsPHD2, DdP4H, TgPhyA and without enzyme (orange). (**B**) Hydroxylation results of TaSkp1 tested with MbP4H, PPHD, TaPHD, HsPHD2, DdP4H, TgPhyA and without enzyme (blue); MbSkp1 tested with MbP4H, PPHD, TaPHD, HsPHD2, DdP4H, TgPhyA and without enzyme (green). The figure was adapted from (Trenkle, Investigations on the Evolutionary Related PHD-Like Enzymes Pseudomonas PHD and Monosiga brevicollis P4H, 2017) (permission acquired) [3].

Some interesting results are apparent from the data in Figure 6.2:

- DdP4H and TgPhyA are able to hydroxylate both TaSkp1 and MbSkp1, but not TaODD, NODD or EF-Tu.
- *Pseudomonas* EF-Tu is only recognised and hydroxylated by the *Pseudomonas* PPHD.
- *Monosiga brevicollis* MbP4H can hydroxylate NODD (human HIF-1α), but not other substrates.

6.6 *Toxoplasma gondii* PhyB, TgPhyB

There is another putative prolyl hydroxylase enzyme, named *Toxoplasma gondii* prolyl hydroxylase B (TgPhyB) in *Toxoplasma gondii*. The predicted structure of TgPhyB using Phyre2 (www.sbg.bio.ic.ac.uk/phyre2/) [4] modelled based on human PHD2 indicates a central DSBH core structure as in other prolyl hydroxylases and a sequence identity of 33% with the structure of PHD2 (Phyre2). Sequence alignment of TgPhyB with PHD2 also indicate the conserved HXD. . .H motif. The prediction strongly suggests that TgPhyB may be another prolyl hydroxylase in *Toxoplasma gondii*.

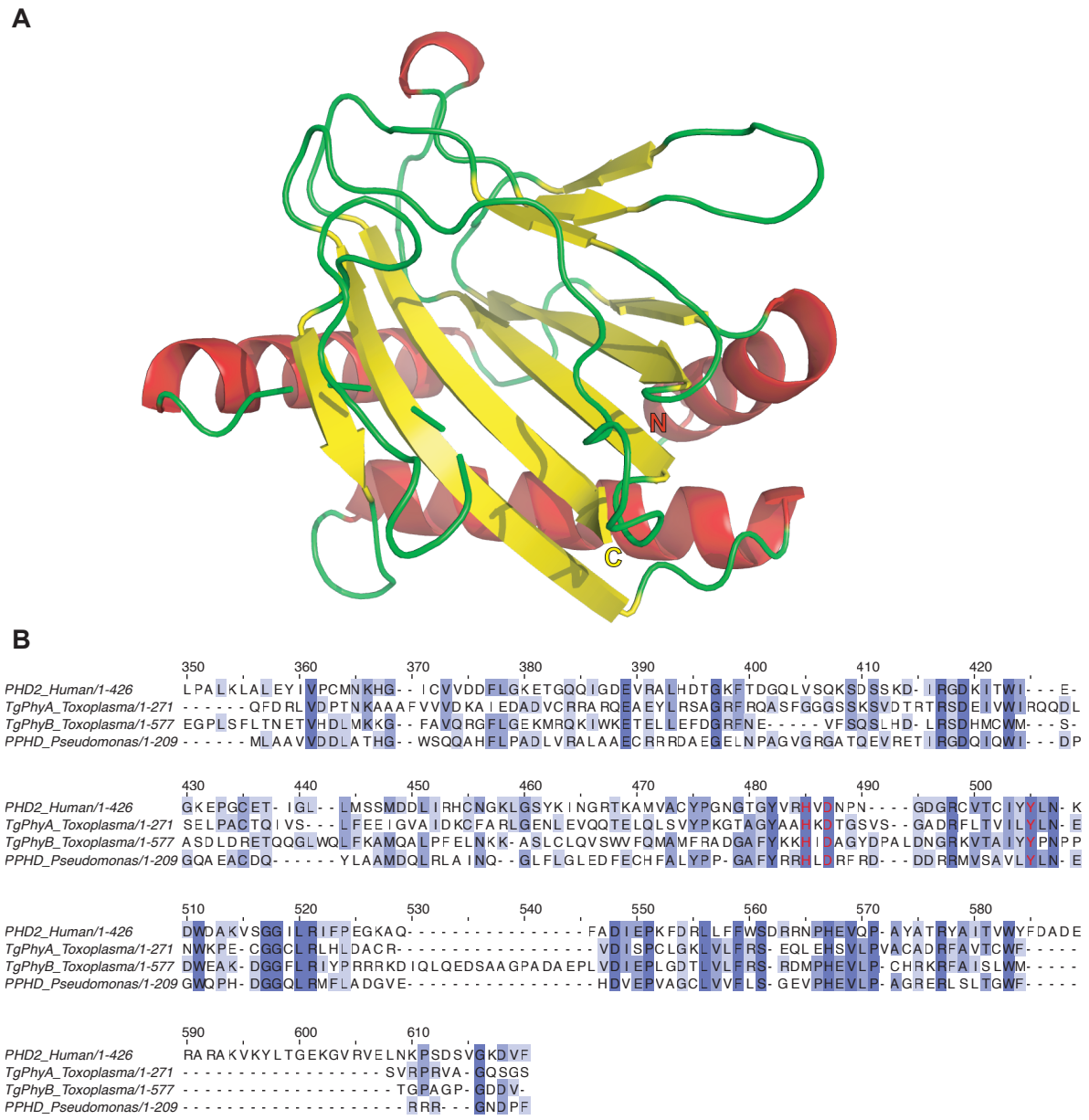


Fig. 6.4 Predicted structure of TgPhyB (361-576). (A) Cartoons representation of the predicted structure of TgPhyB (361-576). β-sheets are colored in yellow. α-helices are in red, loops in green. (B) Sequence alignment of PHD2 (Gene ID: Q9GZT9.1), TgPhyA (Gene ID: TGGT1_114560), TgPhyB (Gene ID: KYF42913.1) and *Pseudomonas* PPHD (PDB: 4J25). Amino acid residues of the conserved HXD. . .H motif are labelled in red.

6.7 TgPhyB Assays

One aliquot of His_6-TgPhyB [buffer condition: 50 mM Tris-HCl, pH 7.0 (at RT), 50 mM NaCl, 10% (v/v) glycerol, 2 mM DTT, 5-10 µM $FeSO_4$, 10 µg/ml leupeptin & aprotinin] was kindly provided by Hanke van der Wel and Dr. Christopher West (Department of Biochemistry and Molecular Biology, University of Georgia). TgPhyB was tested for hydroxylation activity. In the presence of Fe (II), 2OG, ascorbate and oxygen, full-length substrates (DdSkp1, TgSkp1 and HsSkp1 E147P), and peptide substrates [HIF1-CODD, HIF1-NODD, HIF2-CODD, HIF2-NODD, HIF3-CODD, TaODD, TaSkp1, MbSkp1 (kindly provided by Dr. Kerstin Lippl)] were all tested in triplicate with the conditions described in Table 6.5. The reactions were incubated at 37 °C for 1 hour before being quenched with 1% (v/v) aqueous formic acid. Protein samples were analysed by LC-MS (Xevo G2-S). Peptide samples were analysed by MALDI-MS.

Assay Conditions	Final Concentration	Volume
2-oxoglutarate disodium salt (Sigma-Aldrich, K3752)	500 µM	2 µL
Sodium L-ascorbate (Sigma-Aldrich, 11140)	1 mM	2 µL
Ammonium iron(II) sulfate hexahydrate (Sigma-Aldrich, 215406)	50 µM	2 µL
Substrates (peptides or proteins)	100 µM	2 µL
TgPhyB	10 µM	2 µL
Buffer (100 mM HEPES, pH 7.6)	-	10 µL
Total volume	-	20 µL

Table 6.5 Assay conditions of hydroxylation assay samples.

No hydroxylation of these substrates was observed, which suggests TgPhyB may not act as a prolyl hydroxylase (at least with the tested susbrtates). It is also possible that TgPhyB lost its activity either during expression, purification, transportation or because of an un-optimised assay condition. Therefore, in order to study TgPhyB further, effective in-house purification strategies and optimisation of assay conditions are needed.

The sequence of TgPhyB is aligned with other prolyl hydroxylases and a phylogenetic tree is generated to analyse TgPhyB in relation to these prolyl hydroxylases. The result is shown in Figure 6.5.

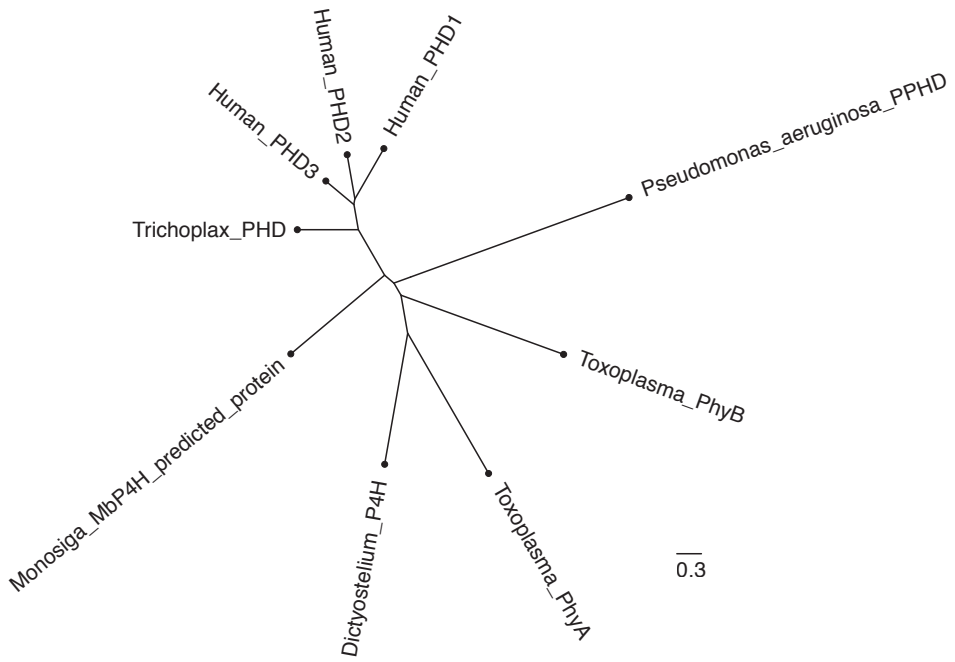


Fig. 6.5 Phylogenetic tree of prolyl hydroxylases from different organisms. Prolyl hydroxylases sequences include: human PHD1 (UniProtKB/Swiss-Prot: Q96KS0.1), human PHD2 (UniProtKB/Swiss-Prot: Q9GZT9.1), human PHD3 (UniProtKB/Swiss-Prot: Q9H6Z9.1), *Trichoplax adhaerens* PHD (GenBank: AFM37574.1), predicted *Monosiga brevicollis* MbP4H (GenBank: EDQ86867.1), *Pseudomonas aeruginosa* PPHD (GenBank: PUA00460.1), *Dictystelium discoideum* DdP4H (UniProtKB/Swiss-Prot: Q86KR9.1), *Toxoplasma gondii* TgPhyA (Gene ID: TGGT1_232960), and *Toxoplasma gondii* TgPhyB (Hanke van der Wel and Dr. Christopher West from Department of Biochemistry and Molecular Biology, University of Georgia). Sequence alignment and phylogenetic tree is generated using Seaview v4.7 [5, 6] and FigTree v1.4.3 [7]. The tree is computed using PhyML v3.0 algorithm based on the maximum-likelihood principle [8].

Phylogenetic analysis suggests the sequence of TgPhyB is most similar as TgPhyA and DdP4H. Therefore, TgPhyB may be another copy of prolyl hydroxylase sharing the similar function as TgPhyA in *Toxoplasma gondii*.

6.8 Summary of Cross Organism Hydroxylation Results

Overall, the results of the cross organism hydroxylation assays (Table 6.6) reveal some interesting findings:

- DdP4H and TgPhyA hydroxylate not only their own Skp1, but also other Skp1s, including MbSkp1, TaSkp1 and HsSkp1E147P. This result suggests that these enzymes are not specific for the Skp1 from their own organism. This might be due to the high similarity of the prolyl-containing region of all the Skp1s. DdP4H and TgPhyA seem to specifically recognise Skp1s as their substrates, as no +16 Da mass shift could be detected for the other organism peptide substrates.
- Similarly, no further cross organism hydroxylation by PHD2 could be detected other than the previously known hydroxylation of TaODD by PHD2.
- PPHD is specific for *Pseudomonas* EF-Tu and was not observed to accept prolyl hydroxylase substrates from human or other organisms.
- TaPHD hydroxylates CODD and NODD suggesting close structural and functional relationships with the human PHDs.
- MbP4H exhibits significant hydroxylation of human HIF-1α NODD, but does not hydroxylate human HIF-1α CODD [data can be found in (Trenkle, Investigations on the Evolutionary Related PHD-Like Enzymes Pseudomonas PHD and Monosiga brevicollis P4H, 2017)] [3]. The observation that NODD is hydroxylated by a more distantly related prolyl hydroxylase may seem to contradict the hypothesis that "an ODD more closely related to CODD evolved before NODD" [9]. However, *Monosiga brevicollis* likely does not contain HIF, so more experiments need to be performed to investigate this interesting observation.

	PPHD	MbP4H	DdP4H	TgPhyA	TgPhyB	TaPHD	PHD2
EF-Tu	✓	✗	✗	✗	✗	✗	✗
MbSkp1	✗	✗	✓	✓	✗	✗	✗
DdSkp1 (full-length protein)	✗	✗	✓	✓	✗	✗	✗
TgSkp1 (full-length protein)	✗	✗	✓	✓	✗	✗	✗
TaSkp1	✗	✗	✓	✓	✗	✗	✗
HsSkp1E147P (full-length protein)	✗	✗	✓	✓	✗	✗	✗
TaODD	✗	✓	✗	✗	✗	✓[9]	✓[2]
HIF1-CODD	✗	✗	✗	✗	✗	✓[3]	✓[2]
HIF1-NODD	✗	✓	✗	✗	✗	✓[3]	✓[2]
HIF2-CODD	–	–	–	–	✗	–	✓[2]
HIF2-NODD	–	–	–	–	✗	–	✓[2]
HIF3-CODD	–	–	–	–	✗	–	✓[2]

Table 6.6 Summary of cross organism hydroxylation assays. Ticked boxes suggest that hydroxylation was observed in the assay. Crossed boxes indicate no hydroxylation was observed.

6.9 References

1. Xu, Y., Brown, K. M., Wang, Z. A., van der Wel, H., Teygong, C., Zhang, D., Blader, I. J., and West, C. M. (2012) The Skp1 protein from *Toxoplasma* is modified by a cytoplasmic prolyl 4-hydroxylase associated with oxygen sensing in the social amoeba Dictyostelium, *J Biol Chem 287*, 25098-25110.

2. Lippl, K. (2018) Structural, Biochemical and Inhibition Studies on Hypoxia-Inducible Factor (HIF) Prolyl Hydroxylases, D.Phil. Thesis, Department of Chemistry, University of Oxford.

3. Trenkle, P. unpublished results, Investigations on the Evolutionary Related PHD-Like Enzymes Pseudomonas PHD and *Monosiga brevicollis* P4H, Department of Chemistry, University of Oxford.

4. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis, *Nat Protoc 10*, 845-858.

5. Galtier, N., Gouy, M., and Gautier, C. (1996) SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny, *Comput Appl Biosci 12*, 543-548.

6. Gouy, M., Guindon, S., and Gascuel, O. (2010) SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building, *Mol Biol Evol 27*, 221-224.

7. Rambaut, A. http://tree.bio.ed.ac.uk/, Institute of Evolutionary Biology, University of Edinburgh.

8. Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0, *Syst Biol 59*, 307-321.

9. Loenarz, C., Coleman, M. L., Boleininger, A., Schierwater, B., Holland, P. W., Ratcliffe, P. J., and Schofield, C. J. (2011) The hypoxia-inducible transcription factor pathway regulates oxygen sensing in the simplest animal, *Trichoplax adhaerens*, *EMBO Rep 12*, 63-70.

7. Summary

Research over the past two decades or so has shown that the HIF oxygen sensing mechanism is highly conserved in nearly all animals. [1] However, the oxygen sensing systems of more diverse organisms, like protists and bacteria, is not yet fully understood. In humans, the response to chronic hypoxia principally relies on the HIF systems [2, 3, 4]. In the HIF system, hydroxylases act as hypoxia sensors; the most extensively studied and probably the most important HIF hydroxylase is PHD2. The interactions between PHD2 and its HIF-1α CODD/NODD substrates have been extensively explored via crystal structures [5-8], but also via biochemical investigations aiming to investigate if its kinetic features are suited to its oxygen sensing role [9-12]. Furthermore, inhibitors of human PHDs have been developed, and are presently being tested in the clinic for the therapeutic treatment of anaemia [13, 14].

Dictyostelium discoideum and *Toxoplasma gondii* are two model early eukaryotic organisms, whose oxygen sensing systems are currently being investigated. A different yet HIF related pathway for oxygen sensing has been revealed in these protists. The protist sensing system involves a prolyl hydroxylase (PHD/P4H/PhyA) - Skp1 pathway, in which Skp1 is hydroxylated at a proline residue under normoxic conditions. Such hydroxylation enables glycosylation of the proline hydroxyl group [15-17]. When hydroxylation and glycosylation occurs, the E3 ubiquitin ligase, in which Skp1 plays a key role, becomes more stable, but how the efficiency of the subsequent ubiquitination changes is unknown [18, 19]. At the onset of my work, neither structural analyses of the protist prolyl hydroxylases or their Skp1substrates, nor *in vitro* kinetic studies of the enzymes using the full-length Skp1s as the substrates, had been carried out.

The work described in the first part of this thesis (**Chapter 2**), focused on optimisation of methods for the production of recombinant forms of DdP4H and TgPhyA full-length proteins on a relatively large scale. The full-length proteins were then used for biochemical and biophysical assays aiming to mimic *in vivo* conditions as close as possible. Truncated recombinant enzymes were used for crystallographic studies.

Chapter 3 describes work leading to an X-ray crystal structure of a truncated DdP4H (60-284), in complex with Ni(II) and NOG, T_m shift assays were used to confirm that Ni^{2+} and NOG stabilised the enzyme. The DdP4H (60-284) crystal structure was analysed in comparison with crystal structures of other prolyl hydroxylases. The structure of DdP4H (60-284) revealed a conserved distorted DSBH core fold, which is present in all studied members of the 2OG oxygenase super family [5-7]. In the active site of DdP4H (60-284), the highly conserved conserved HXD. . .H motif supports binding of the metal ion. Comparison of the DdP4H (60-284) structure with that of human PHD2 (PDB ID: 5L9R [20]) and *Pseudomonas putida* PPHD (PDB ID: 4J25 [21]), indicates that, at least as crystallised, the structure of DdP4H (60-284) has more disordered regions than PHD2. The DdP4H (60-284) structure is more similar to PPHD than PHD2 (Phyre2), implying a closer evolutionary relationship between *Dictyostelium discoideum* and *Pseudomonas putida* than *Dictyostelium discoideum* and the human PHDs. The relatively disordered regions in DdP4H (60-284) may in part explain its lower hydroxylation efficiency compared to PHD2 (**Chapter 4**) and / or reflect the folded nature of the DdP4H Skp1 substrate compared to the disordered ODD PHD substrates. The regions at the open end of DSBH barrel are less similar between the three enzymes, which may explain differences in substrate recognition and binding.

Attempts at co-crystallising DdP4H (60-284) with DdSkp1 peptides and disulfide-bond crosslinking methods to crystallise DdP4H (60-284) with DdSkp1 were carried out, although they

did not yield crystals. The difficulty encountered in obtaining an enzyme-substrate complex may in part be due to the low binding affinity between DdP4H and its substrate DdSkp1, which was further investigated in **Chapter 4**. Obtaining a structure of TgPhyA was more challenging than DdP4H. Several crystallisation trials were unsuccessfully carried out using both full-length TgPhyA and truncated TgPhyA (53-264). The structure of TgPhyA will be helpful for understanding the biochemical behaviour of this enzyme and for further comparing the differences observed between different prolyl hydroxylases. More TgPhyA crystallographic experiments need to be carried out in the future.

Chapter 4 focuses on the biochemical characterization of DdP4H and TgPhyA. Similar to other prolyl hydroxylases, DdP4H and TgPhyA hydroxylate DdSkp1 and TgSkp1 respectively, requiring the presence of 2OG, ascorbate and Fe^{2+}. LC-MS assay results suggest that full-length DdSkp1 and TgSkp1 proteins are better substrates than the 25-mer DdSkp1 and TgSkp1 peptide substrates for DdP4H and TgPhyA, respectively. This observation may be because other regions than the active site of DdSkp1 are required to take part in substrate binding. Therefore, the full-length protein substrates were used to carry out kinetic studies for the first time. The two organisms both show comparatively high K_m values [(349.9 ± 77.3) µM for DdSkp1, and (202.4 ± 46.5) µM for TgSkp1] for the substrate, even using full-length Skp1s, indicating a weak binding affinity for these substrates. In *Dictyostelium discoideum*, the K_m value for 2OG was (35.9 ± 5.9) µM. The K_m value for oxygen was too high to be determined, representing a relative ability to respond to changes in O_2 availability, supporting previous reports that DdP4H is a potential oxygen sensor in *Dictyostelium discoideum* [17, 22, 23]. In *Toxoplasma gondii*, however, the K_m value for 2OG was impossible to acquire, due to a high rate of substrate uncoupled 2OG turnover (**Chapter 4.26.4**). K_m value for oxygen was unexpectedly low (0.2 ± 0.04 %). Considering that *Toxoplasma gondii*, as a parasite, resides most of the time in host cells or tissues as tachyzoites

or bradyzoites, where the oxygen concentration can be lower than 1%, TgPhyA might act as an oxygen sensor.

MS/MS analysis and regiochemistry studies confirmed that hydroxylation occurs at the C-4 position of the target proline. 2OG turnover assays provide several unexpected results. Firstly, DdP4H exhibited slow uncoupled turnover. The turnover rate was fastest with DdSkp1, suggesting DdSkp1 is a better substrate than other Skp1s for DdP4H. TgPhyA exhibited significantly rapid uncoupled turnover, with ascorbate as a requirement, implying its active site may be different from DdP4H. However, HsSkp1 was not able to be hydroxylated due to the lack of a proline residue in the corresponding region, but prevents uncoupled turnover, possibly by binding to the enzyme substrate binding site as a competitive inhibitor. The uncoupled TgPhyA 2OG turnover rate was the fastest using the substrate HsSkp1 E147P, suggesting it might even be a better substrate than its native substrate TgSkp1. For PHD2, 2OG turnover decreased to ~ 0 % in the presence of HsSkp1. This indicates that HsSkp1 may act as an inhibitor by binding to PHD2 to prevent the hydrolysis of 2OG (and hence hydroxylation of HIF). Therefore, the *in vivo* kinetics of hydroxylation of HIF-1α by PHD2 may be affected by HsSkp1.

The work of **Chapter 5** aimed to test known 2OG oxygenase inhibitors of DdP4H and TgPhyA in order to characterise their effects on these prolyl hydroxylases. Most PHD2 inhibitors showed similar inhibition effects against DdP4H and TgPhyA, reflecting the structural similarity of the enzymes. The more potent inhibitors include GSK1278863, IOX1, BIOX2, IOX2, IOX4, FG-0041, FG-2216, Vadadustat and FG-4592. Interestingly, the exceptions were 2,4-PDCA, CCT6-140 and NOG. Remarkably, the (at least, normally) broad spectrum 2OG oxygenase inhibitor 2,4-PDCA showed no inhibition of DdP4H under the conditions tested. 2,4-PDCA is normally a 2OG dependent oxygenase inhibitor that mimics 2OG. CCT6-140 did not inhibit TgPhyA. And surprisingly, the normally broad spectrum inhibitor NOG also showed no inhibition towards either enzyme. During the inhibition studies of DdP4H, it was observed that the inhibitors

appear more effective when DdSkp1 peptide was used as a substrate rather than full-length DdSkp1 protein. This finding likely reflects the evidence that full-length DdSkp1 protein is a better substrate than the DdSkp1 peptide, which is consistent with the results of the kinetic assays described in **Chapter 4**. Some of the innovative highly potent PHD inhibitors or activators (provided by Dr. James Holt-Martyn) were ineffective against DdP4H and TgPhyA. The results for DdP4H and TgPhyA suggest these compounds have high selectivity and specificity towards PHD2 when compared to other prolyl hydroxylases.

In the last part of this thesis (**Chapter 6**), cross-organism hydroxylation assays were analysed with the different prolyl hydroxylases. The results of this work may provide insight into the evolution of oxygen sensing systems among different organisms. DdP4H and TgPhyA were able to hydroxylate their own Skp1 and other Skp1s, including MbSkp1, TaSkp1 and HsSkp1E147P. These results showed that the enzymes are promiscuous towards Skp1s from different organisms. This may reflect the proposition that more basal organisms tend to have enzymes with higher promiscuity [24]. A more likely possibility revolves around the high-degree of conservation within the prolyl-containing region of the Skp1s. PHD2 was not able to take any of the Skp1 proteins as substrates, supporting the comparatively distant relationship of the oxygen sensing systems between human (HIF) and distant eukaryote (such as Skp1). PPHD is (apparently) highly specific for EF-Tu as it did not take any of the other tested substrates. TaPHD was able to hydroxylate human CODD and NODD in addition to its natural substrate TaODD. No DdP4H substrates were identified for the *Monosiga brevicollis* oxygenase, MbP4H, although it exhibited significant hydroxylation activity towards human NODD.

Overall, the results of this thesis provide insight understanding of the *Dictyostelium discoideum* oxygen sensing system and *Toxoplasma gondii* TgPhyA as a potential oxygen sensor. In addition, the study of homologues of prolyl hydroxylases in different organisms revealed complicated, yet

fascinating, evolutionary paths for oxygen sensing systems. The investigations on prolyl hydroxylases and substrates across different organisms implies evolution of various features, including structure, types of substrate, kinetics and biological functions. The results suggest that the evolution of oxygen sensing systems may have not originated from a single rapid change, but rather multiple divergent and convergent processes. It is hoped that a chemical understanding of the evolutionary origins of oxygen sensing systems will enable a better understanding of the HIF system in humans. The studies on TgPhyA may also one-day help enable prosecution of a novel target for the treatment of toxoplasmosis.

References

1. Loenarz, C., Coleman, M. L., Boleininger, A., Schierwater, B., Holland, P. W., Ratcliffe, P. J., and Schofield, C. J. (2011) The hypoxia-inducible transcription factor pathway regulates oxygen sensing in the simplest animal, *Trichoplax adhaerens*, *EMBO Rep 12*, 63-70.

2. Jelkmann, W. (2007) Erythropoietin after a century of research: younger than ever, *Eur J Haematol 78*, 183-205.

3. Semenza, G. L. (2001) Hypoxia-inducible factor 1: control of oxygen homeostasis in health and disease, *Pediatr Res 49*, 614-617.

4. Chowdhury, R., Hardy, A., and Schofield, C. J. (2008) The human oxygen sensing machinery and its manipulation, *Chem Soc Rev 37*, 1308-1319.

5. Aik, W., McDonough, M. A., Thalhammer, A., Chowdhury, R., and Schofield, C. J. (2012) Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases, *Curr Opin Struct Biol 22*, 691-700.

6. McDonough, M. A., Loenarz, C., Chowdhury, R., Clifton, I. J., and Schofield, C. J. (2010) Structural studies on human 2-oxoglutarate dependent oxygenases, *Curr Opin Struct Biol 20*, 659-672.

7. Clifton, I. J., McDonough, M. A., Ehrismann, D., Kershaw, N. J., Granatino, N., and Schofield, C. J. (2006) Structural studies on 2-oxoglutarate oxygenases and related double-stranded beta-helix fold proteins, *J Inorg Biochem 100*, 644-669.

8. Schofield, C. J., and Zhang, Z. (1999) Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes, *Curr Opin Struct Biol 9*, 722-731.

9. Tarhonskaya, H., Hardy, A. P., Howe, E. A., Loik, N. D., Kramer, H. B., McCullagh, J. S., Schofield, C. J., and Flashman, E. (2015) Kinetic Investigations of the Role of Factor Inhibiting Hypoxia-inducible Factor (FIH) as an Oxygen Sensor, *J Biol Chem 290*, 19726-19742.

10. Tarhonskaya, H., Rydzik, A. M., Leung, I. K., Loik, N. D., Chan, M. C., Kawamura, A., McCullagh, J. S., Claridge, T. D., Flashman, E., and Schofield, C. J. (2014) Non-enzymatic chemistry enables 2-hydroxyglutarate-mediated activation of 2-oxoglutarate oxygenases, *Nat Commun 5*, 3423.

11. Tarhonskaya, H., Chowdhury, R., Leung, I. K., Loik, N. D., McCullagh, J. S., Claridge, T. D., Schofield, C. J., and Flashman, E. (2014) Investigating the contribution of the active site environment to the slow reaction of hypoxia-inducible factor prolyl hydroxylase domain 2 with oxygen, *Biochem J 463*, 363-372.

12. Flashman, E., Hoffart, L. M., Hamed, R. B., Bollinger, J. M., Jr., Krebs, C., and Schofield, C. J. (2010) Evidence for the slow reaction of hypoxia-inducible factor prolyl hydroxylase 2 with oxygen, *FEBS J 277*, 4089-4099.

13. Chowdhury, R., Candela-Lena, J. I., Chan, M. C., Greenald, D. J., Yeoh, K. K., Tian, Y. M., McDonough, M. A., Tumber, A., Rose, N. R., Conejo-Garcia, A., Demetriades, M., Mathavan, S., Kawamura, A., Lee, M. K., van Eeden, F., Pugh, C. W., Ratcliffe, P. J., and Schofield, C. J. (2013) Selective small molecule probes for the hypoxia inducible factor (HIF) prolyl hydroxylases, *ACS Chem Biol 8*, 1488-1496.

14. Rose, N. R., McDonough, M. A., King, O. N., Kawamura, A., and Schofield, C. J. (2011) Inhibition of 2-oxoglutarate dependent oxygenases, *Chem Soc Rev 40*, 4364-4397.

15. West, C. M., and Blader, I. J. (2015) Oxygen sensing by protozoans: how they catch their breath, *Curr Opin Microbiol 26*, 41-47.

16. Zhang, D., van der Wel, H., Johnson, J. M., and West, C. M. (2012) Skp1 prolyl 4-hydroxylase of *Dictyostelium* mediates glycosylation-independent and -dependent responses to O2 without affecting Skp1 stability, *J Biol Chem 287*, 2006-2016.

17. Xu, Y., Wang, Z. A., Green, R. S., and West, C. M. (2012) Role of the Skp1 prolyl-hydroxylation/glycosylation pathway in oxygen dependent submerged development of *Dictyostelium*, *BMC Dev Biol 12*, 31.

18. Sheikh, M. O., Xu, Y., van der Wel, H., Walden, P., Hartson, S. D., and West, C. M. (2015) Glycosylation of Skp1 promotes formation of Skp1-cullin-1-F-box protein complexes in *Dictyostelium*, *Mol Cell Proteomics 14*, 66-80.

19. Sheikh, M. O., Schafer, C. M., Powell, J. T., Rodgers, K. K., Mooers, B. H., and West, C. M. (2014) Glycosylation of Skp1 affects its conformation and promotes binding to a model f-box protein, *Biochemistry 53*, 1657-1669.

20. Chowdhury, R., Leung, I. K., Tian, Y. M., Abboud, M. I., Ge, W., Domene, C., Cantrelle, F. X., Landrieu, I., Hardy, A. P., Pugh, C. W., Ratcliffe, P. J., Claridge, T. D., and Schofield,

C. J. (2016) Structural basis for oxygen degradation domain selectivity of the HIF prolyl hydroxylases, *Nat Commun 7*, 12673.

21. Scotti, J. S., Leung, I. K., Ge, W., Bentley, M. A., Paps, J., Kramer, H. B., Lee, J., Aik, W., Choi, H., Paulsen, S. M., Bowman, L. A., Loik, N. D., Horita, S., Ho, C. H., Kershaw, N. J., Tang, C. M., Claridge, T. D., Preston, G. M., McDonough, M. A., and Schofield, C. J. (2014) Human oxygen sensing may have origins in prokaryotic elongation factor Tu prolyl-hydroxylation, *Proc Natl Acad Sci U S A 111*, 13331-13336.

22. West, C. M., van der Wel, H., and Wang, Z. A. (2007) Prolyl 4-hydroxylase-1 mediates O_2 signaling during development of *Dictyostelium*, *Development 134*, 3349-3358.

23. Wang, Z. A., Singh, D., van der Wel, H., and West, C. M. (2011) Prolyl hydroxylation- and glycosylation-dependent functions of Skp1 in O_2-regulated development of *Dictyostelium*, *Dev Biol 349*, 283-295.

24. Carbonell, P., Lecointre, G., and Faulon, J. L. (2011) Origins of specificity and promiscuity in metabolic networks, *J Biol Chem 286*, 43994-44004.

8. Methods and Materials

8.1 Molecular Biology Techniques

8.1.1 Preparation of Competent Cells [1]

Escherichia coli cells were streaked onto an LB-agar plate and were incubated at 37 °C overnight. One colony was picked to inoculate in 100 mL sterile 2TY growth medium and was grown in a shaking incubator (220 rpm) at 37 °C overnight. 200 µL of this starter culture was added into 100 mL sterile growth medium, and was incubated at 37 °C to O.D. (at 600 nm wavelength) 1.2. Cells were centrifuged at 750 × g, 4 °C for 5 minutes using a pre-chilled sterile 50 centrifuge tube, and were gently suspended in 25 mL sterile 100 mM $CaCl_2$, 20 % (v/v) glycerol. Cells were then kept on ice for 20 minutes before being centrifuged at 750 × g, 4 °C for 5 minutes and gently resuspend in 1 mL 100 mM $CaCl_2$, 20 % (v/v) glycerol. Cells were aliquoted and stored immediately at -80°C.

8.1.2 Plasmids

The plasmids used in my work are listed in Table 8.1.

Plasmids	Gene	Tag	Source
pET-15b-TgPhyA	*Toxoplasma gondii* PhyA	*N*-Terminal His-Tag	Dr Emily Flashman
pET-15-DdP4H	*Dictyostelium discoideum* P4H	*N*-Terminal His-Tag	Dr Christoph Loenarz
pNIC-ZB-TgPhyA	*Toxoplasma gondii* PhyA	*N*-Terminal His-ZB-Tag	Ligation Independent Cloning (LIC)
pNIC28-Bsa4-TgPhyA (53-264)	*Toxoplasma gondii* PhyA	*N*-Terminal His -Tag	Ligation Independent Cloning (LIC)
pNIC-ZB-DdP4H	*Dictyostelium discoideum* P4H	*N*-Terminal His-ZB-Tag	Ligation Independent Cloning (LIC)
pNIC-ZB-DdP4H (60-284)	*Dictyostelium discoideum* P4H	*N*-Terminal His-ZB-Tag	Ligation Independent Cloning (LIC)
pNIC-ZB-DdP4H (60-284) R192C	*Dictyostelium discoideum* P4H	*N*-Terminal His-ZB-Tag	Site-directed mutagenesis
pET-21a-HsSkp1	Human Skp1b	*C*-Terminal His-Tag	Dr Thomas Ayrton Murray-Rust
pET-21a-HsSkp1-E147P variant	Human Skp1b E147P variant	*C*-Terminal His-Tag	Site-directed mutagenesis
pET-19b-DdSkp1	*Dictyostelium discoideum* Skp1	No Tag	Dr Hanke van der Wel
pNIC28-Bsa4-DdSkp1	*Dictyostelium discoideum* Skp1	*N*-Terminal His-Tag	Ligation Independent Cloning (LIC)
pET-15-TgSkp1	*Toxoplasma gondii* Skp1	No Tag	Dr Hanke van der Wel
pNIC28-Bsa4-TgSkp1	*Toxoplasma gondii* Skp1	*N*-Terminal His-Tag	Ligation Independent Cloning (LIC) [18]
pET-15-FLAG-TgSkp1	*Toxoplasma gondii* Skp1	N-Terminal FLAG-Tag	Dr Hanke van der Wel

Table 8.1 Vector constructs.

8.1.3 Site-Directed Mutagenesis (SDM) [2]

Site-directed mutagenesis (SDM) was performed using the QuikChange® Site-Directed Mutagenesis Kit (New England Biolabs, Catalog: M0491S), using the forward primer and the reverse primer. After the PCR was completed, 1 µL of DpnI restriction enzyme (10 U/µL) was added to the reaction solution and incubated for 3 hours at 37 °C. Following incubation, 2 µL of the reaction solution was transformed into XL10 Gold *E. coli* competent cells. Plasmids were

purified using a GeneJET Plasmid Miniprep Kit (Thermo Scientific) and sequenced by Source BioScience for verification.

A

PCR Reagents	Volume Per Tube / µL
5x Q5 buffer (NEB #B9027S)	10
dNTP (NEB #N0447S)	1
Q5 polymerase (NEB #M0491S)	1
Primer mix (10 µM)	4
Template DNA (40 ng – 100 ng)	1
MilliQ water	2.5
Total	50

B

PCR Steps	Temperature / °C	Incubation Time
1	98	2 min
2	98	30 s
3	50	1 min
4	72	8 min 40 s
5	Repeat Step 2 – 4 for 16 cycles	
6	72	2 min
7	4	Pause

C

Protein Name	Primer Sequence
DdP4H (60-284) R192C	Forward, 5’ – GTAGATATATTAAACATTGTGATTCATTTTATTCCTC – 3’
	Reverse, 5’ – GAGGAATAAAATGAATCACAATGTTTAATATATCTAC – 3’
DdP4H (60-284) F195C	Forward, 5’ – CATAGAGATTCATGTTATTCCTCAGAATC – 3’
	Reverse, 5’ – GATTCTGAGGAATAACATGAATCTCTATG – 3’
HsSkp1 E147P	Forward, 5’ –CAATATCAAAAATGACTTTACTCCGGAGGAGGAAGCCCAGGTA-CGC– 3’
	Reverse, 5’ –GCGTACCTGGGCTTCCTCCTCCGGAGTAAAGTCATTTTTGATAT-TG– 3’

Table 8.2 Procedures for site-directed mutagenesis to genrate DdP4H (60-284) R192C and DdP4H (60-284) F195C. (A) Sample contents for PCR. (B) PCR protocol. (C) Primers designed specifically for site-directed mutagenesis.

8.1.4 Bacterial Transformations [1, 3]

Plasmid DNA was transfected into XL10 Gold or BL21 Gold (DE3) *E. coli* competent cells. Bacterial cells (20 µL) were thawed in a 50 mL Falcon tube on ice and added to 1.0 - 2.0 µL of bacterial plasmid DNA (100 ng/µL). After resting for 10 - 15 minutes on ice, the bacteria were heat shocked at 42 °C for 30 s (XL10 Gold *E. coli* competent cells) or 20 s [BL21 Gold (DE3) *E. coli* competent cells]. 2TY growth media (300 µL) was added to the mixture, and solution was incubated at 37 °C for 1 hour. Afterwards, 40 µL of the mixture was spread onto LB-agar plates with relevant antibiotic and allowed to grow at 37 °C overnight.

8.1.5 Ligation Independent Cloning (LIC) [4, 5]

The primers were designed with the following 5' extensions: TACTTCCAATCCATG for the forward primer and TATCCACCTTTACTG for the reverse primer. A stop codon (TTA for DdP4H and TCA for TgPhyA) was added after the sequence TATCCACCTTTACTG for the reverse primer depending on the sequence of DdP4H or TgPhyA.

The DNA was amplified by the polymerase chain reaction (PCR) using a thermal cycler using the following conditions:

A

PCR Reagent	Volumn Per Tube / µL
5x Q5 Buffer (NEB #B9027S)	13.5
$MgSO_4$ (50mm, Bioline BIO-37026)	5
dNTP (NEB #N0447S)	0.5
Q5 Polymerase (NEB #M0491S)	0.75
Primer Mix (10 µM)	0.5
Template DNA (5 ng/µL)	2.5
MilliQ Water	2.5
Total	25

B

PCR Step	Temperature / °C	Incubation Time
1	95	1 min
2	94	30 s
3	52	30 s
4	68	6.5 min
5	Repeat Step 2 – 4 for 30 cycles	
6	68	5 min
7	15	Pause

Table 8.3 Procedures for LIC to clone the genes encoding for DdP4H and TgPhyA into different vectors. (A) Sample ingredient for PCR. (B) PCR conditions.

The samples were then treated with Dpn1 (New England BioLabs #R0176S) for 1 hour at 37°C. The PCR products were purified using a GeneJET PCR Purification Kit (Thermo Scientific #K0701) and samples were analysed by agarose gel electrophoresis. The samples were digested by T4 polymerase (New England Biolabs #M0203) to generate a cohesive end [6, 7].

T4 Digestion Reagents	Volumn Per Tube / µL
DNA	5
10 X T4 polymerase buffer (NEB #M0203S)	1
dCTP (25 mM NEB #N0447S)	1
DTT (100 mM)	0.5
BSA (10 mg/mL NEB B9001S)	0.1
T4 DNA polymerase (5 ng/µL NEB #M0203S)	0.25
MilliQ water	2.15
Total	10

Table 8.4 Sample conditions for T4 DNA polymerase digestion to generate cohesive ends.

The inserts were annealed into the pNIC-ZB [6, 8] and pNH-TrxT [9] vectors. The new constructs were transfected into XL10 Gold *E. coli* competent cells [10] and were further purified using a GeneJET Plasmid Miniprep Kit (Thermo Scientific #K0502). Purified plasmid was stored for future use.

Protein Name	Primer Sequence
DdP4H	Forward, 5’ – TACTT-CCAATCCATGGATATTTCTAACTTACCC – 3’
	Reverse, 5’ – TATCCACCTTTACTGTTAATAAATCCAAGTT-GTAATTG – 3’

TgPhyA	Forward, 5’ – TACTTCCAATCCATGGCTCTTGA-GGACCTG – 3’
	Reverse, 5’ – TATCCACCTTTACTGTCACGATCCCGATTGACC – 3’
DdP4H (60-284)	Forward, 5’ – TACTTCCAATCCATGAAAATCTTCAACAAAGAATC – 3’
	Reverse, 5’ – TATCCACCTTTACTGTTAAT-AAATCCAAGTTGTAATTG – 3’
TgPhyA (53-264)	Forward, 5’ – TACTT-CCAATCCATGCTTAGCGATTCGCAATTTG – 3’
	Reverse, 5’ – TATCCACCTTTACTGTCAACGAGGGC-GAACGGAG – 3’
DdSkp1	Forward, 5’ – TACTTCCAATCCATGTCTTTAGTTAAATTAGAATCTT – 3’
	Reverse, 5’ - TATCCACCTTTACTGTTAGTTTCCACCTTTATCTTC – 3’
TgSkp1	Forward, 5’ – TACTTCCAATCCATGTCAAAAGAAAGAATGG – 3’
	Reverse, 5’ – TATCCACCTTTACTGTCATGCATCTTCGCAC – 3’

Table 8.5 Primers designed specifically for ligation independent cloning.

8.1.6 Agarose Gel Electrophoresis [11]

Agarose gels (1 %) were prepared by heating 0.5 g agarose in 50 mL TAE buffer [40 mM Tris- HCl, 20 mM sodium acetate, 1mM ethylenediaminetetraacetic acid (EDTA)] in a SuperWave 750 microwave (Tecnolec) until dissolved. Then 10 µL SYBR SafeTM DNA gel stain (Invitrogen) was added. Gels were cast in a H2-SET gel tray (Anachem) were kept at room temperature for 20 min. DNA Loading Dye (New England Biolabs, #B7024) was added to the samples before loading them onto the gel. The gels were run on a Bio-Rad system at 100 V for approximately 45 min. DNA bands were visualised by UV light illumination using a Bio Rad gel imaging system. Samples were run against a GeneRuler 1 kb DNA Ladder (Thermo Fisher Scientific) for molecular weight determinations.

8.2 Protein Production

8.2.1 2TY Growth Medium

Reagent	Amount Per 1L
NaCl	5.0 g
Yeast extract	10.0 g
Tryptone	16.0 g

Table 8.6 2TY growth medium.

8.2.2 Glycerol Stock Preparation [12]

Glycerol stocks were made by gently mixing 0.75 mL of starter culture with 0.25 mL of sterile glycerol in a sterile cryotube. Glycerol stocks were stored at -80°C.

8.2.3 Starter Cultures

Starter cultures were made one day before induction trial or large-scale growth. A bacterial colony from a LB plate or small amount of a glycerol stock was added into 100 mL of 2TY growth media. The mixture was then inoculated overnight at 37 °C, 180 rpm.

8.2.4 Expression Trials

Starter cultures were transferred into fresh 2TY media at 1:100 ratios. Antibiotics were then added into the media and the bacteria were grown at 37 °C, 180 rpm until optimal density (O.D. at 600 nm wavelength) reached 1.2. Induction was carried out with various conditions with different temperature, induction time, and concentrations of IPTG (Table 8.7).

Sample	Induction time / h	IPTG concentration / mM	Induction temperature / C°
Non-induction control	0	0	0
1		0.1	37
2	4	0.2	37
3		0.5	37
4		0.1	28
5	4	0.2	28
6		0.5	28
7		0.1	28
8	Overnight	0.2	28
9		0.5	28
10		0.1	18
11	Overnight	0.2	18
12		0.5	18

Table 8.7 Induction trial conditions.

8.2.5 SDS-PAGE Electrophoresis [13, 14]

SDS-PAGE gels were prepared before electrophoresis (Table 8.8, 8.9). 15 µL of protein sample was taken out into an Eppendorf tube. Samples were mixed with 15 µL 2 × SDS-PAGE Loading Buffer (Table 8.10) before being placed into a hot block to boil for 10 minutes. A 10µl of mixture was added into each well of the SDS-PAGE gel. The gel was run at 200 V in running buffer [200 mM glycine, 25 mM Tris, 10 % (v/v) SDS] for 70 minutes. The gel was then stained in stain buffer [40 % (v/v) methanol, 10 % (v/v) acetic acid, 0.05 % (w/v) Commassie Blue R] for 10 minutes subsequently destained [40 % (v/v) methanol, 10 % (v/v) acetic acid] overnight.

Reagents	Quantities at acrylamide content			
Percentage gel	7.5 % (w/v)	10 % (w/v)	12.5 % (w/v)	15 % (w/v)
Tris-HCl (1 M, pH 7.5)	2.50 mL	2.50 mL	2.50 mL	2.50 mL
Tris-HCl (0.5 M, pH 6.8)	0 mL	0 mL	0 mL	0 mL
Milli-Q water	4.80 mL	4.00 mL	3.14 mL	2.30 mL
30 % (w/v) acrylamide / 1.034 % (w/v) bis-acrylamide	2.50 mL	3.32 mL	4.20 mL	5.00 mL
10 % (w/v) SDS	100 µL	100 µL	100 µL	100 µL
10 % (w/v) APS	100 µL	100 µL	100 µL	100 µL
TEMED	8 µL	8 µL	8 µL	8 µL

Table 8.8 Resolving gel (quantities for 2 gels).

Reagents	Amount
Tris-HCl (0.5 M, pH 6.8)	1.26 mL
Milli-Q water	3.14 mL
30 % (w/v) acrylamide / 1.034 % (w/v) bis-acrylamide	0.50 mL
10 % (w/v) SDS	50 µL
10 % (w/v) APS	50 µL
TEMED	5 µL

Table 8.9 Stacking gel (quantities for 2 gels).

Reagent	Amount
10 % (w/v) SDS solution	0.90 mL
100 % (v/v) glycerol	0.45 mL
1 M Tris (pH 6.8)	0.27 mL
Bromophenol blue	0.005 g
β-mercaptoethanol	0.25 mL
Milli-Q water	0.63 mL

Table 8.10 2 × SDS-PAGE loading buffer (2.5 mL).

8.2.6 Large-Scale Growth

Starter cultures were transferred into 12 x 600 mL fresh 2TY media at a ratio of 1 : 100 culture : fresh media. Antibiotics were then added into the media and the bacteria were grown at 37 °C, 180 rpm until O.D. (600 nm wavelength) reached 1.2. Induction was carried out at an optimised temperature, an optimized induction time, and an optimized concentration of IPTG (Table M12). Cells were harvested by centrifugation (8,000 rpm, 10 min, 10 °C), and stored at -80 °C.

Protein	Antibiotics	Induction time	IPTG concentration (mm)	Induction temperature (C°)
His_6-ZB-DdP4H	Kanamycin	Overnight	0.2	18
His_6-ZB-TgPhyA	Kanamycin	Overnight	0.1	18
His_6-ZB-DdP4H (60-284)	Kanamycin	Overnight	0.1	18
His_6-TgPhyA (53-264)	Kanamycin	Overnight	0.1	18
His_6-ZB-DdP4H (60-284) R192C	Kanamycin	Overnight	0.1	18
His_6-ZB-DdSkp1	Kanamycin	Overnight	0.2	28
His_6-ZB-TgSkp1	Kanamycin	Overnight	0.2	28
His_6-TEV	Ampicillin	Overnight	0.5	18
HsSkp1-His_6	Ampicillin	Overnight	0.2	28
HsSkp1 E147P-His_6	Ampicillin	Overnight	0.2	28

Table 8.11 Protein expression conditions.

8.3 Protein Purification

8.3.1 Ni(II) Affinity Chromatography [15]

All protein purification procedures were carried out at 4 °C with a flow rate of 1 mL· min^{-1} using an AKTA FPLC system. Several Buffers were prepared for Ni(II) affinity chromatography as in Table 8.12.

Buffer name and amount	Quantities
1 × Charge Buffer	100 mM of $NiSO_4$
1× Binding Buffer	5 mM imidazole
	0.5 M NaCl
	50 mM HEPES
1 × Wash Buffer	30 mM imidazole
	0.5 M NaCl
	50 mM HEPES
1 × Elution Buffer	0.5 M imidazole
	0.5 M NaCl
	50 mM HEPES
1 × Strip Buffer	100 mM EDTA
	0.5 M NaCl
	50 mM HEPES

Table 8.12 Buffers for Ni(II) Affinity Chromatography.

Cell pellets (~30 g) were resuspended in 120 mL Binding Buffer (Table 8.12), and were gently stirred until the cells were fully resuspended. A small amount of DNaseI was then added into the mixture. The cells were then sonicated on ice by sonication (Sonics Vibracell VCX750 Ultrasonic Cell Disruptor) at 60 % amplitude with 10 × (30 s sonication + 40 s rest).

A 5 mL HisTrap HP column (GE Healthcare) was charged with Charge Buffer and equilibrated with Binding Buffer (Table 8.12). The lysate from sonication was centrifuged at 23,000 rpm, 4 °C for 30 minutes. The supernatant was filtered through a Minisart 0.45 µm syringe filter (Sartorius Stedim Biotech) before loading onto the column. The column was then washed with 100 mL Wash Buffer (Table 8.12) to remove non-specifically bound proteins. The protein of interest was later eluted into fractions with Elution Buffer (Table 8.12). Afterwards, the HisTrap HP column

was cleaned with Strip Buffer. Several samples were taken out of different fractions to run SDS-PAGE to identify the target protein.

8.3.2 Ion Exchange Chromatography (IEC) [16, 17]

Buffer name and amount	Quantities
Buffer A	100 mM MES hydrate
	pH 5.8
	100 mM MES hydrate
Buffer B	2 M NaCl
	pH 5.8

Table 8.13 Buffer A and Buffer B Preparation.

Proteins were further purified by cation exchange chromatography using a 15 mL SP-Sepharose column. Buffer A and Buffer B (Table 8.13) were prepared beforehand. The proteins were buffer exchanged into Buffer A using a PD-10 desalting column (GE Healthcare) before loading onto the column. The column was firstly washed with 20 mL Milli-Q water and pre-equilibrated with 180 mL Buffer A. The proteins were then loaded onto the column and the column was washed with 30 mL Buffer A to remove the unbound proteins. Afterwards, the proteins were eluted using a gradient from 100 % Buffer A to 100 % Buffer B. The elutions were collected in fractions. Lastly, the SP-Sepharose column was cleaned with 15 mL Milli-Q water and was stored in 20 % (v/v) ethanol. Several samples were taken out of different fractions to run SDS-PAGE to identify the target protein.

8.3.3 Ni(II) Affinity Chromatography to Re-purify the Protein Samples

Proteins were treated with the His_6-TEV protease (1 mg / 50 mg target protein) and 50 mM ethylenediaminetetraacetic acid (EDTA) overnight at -4 °C to cleave the His-ZB tag and to remove Fe(II) from the protein [18]. The solution was then buffer exchanged into Binding Buffer using a PD-10 desalting column (GE Healthcare) before loading onto a 5 mL HisTrap HP column (GE Healthcare). The flow-through was collected into a Falcon tube and a sample was taken out for SDS-PAGE analysis to identify the target protein. The HisTrap HP column was then cleaned with Strip Buffer (Table 8.12). The purity of the protein reached 95%. It was concentrated using a 15 mL centrifugal filter (Amicon®) and the concentration was tested using the absorbance at 280 nm as measured by a NanoDrop ND-1000 spectrophotometer (Thermo Scientific). Extinction coefficients were calculated by Expasy ProtParam (http://web.expasy.org/protparam/). The protein solution was then aliquoted and stored at -80 °C.

8.3.4 Size Exclusion Chromatography (SEC) [19-21]

The protein was further purified by size exclusion chromatography (SEC) using a Superdex 75 300 mL column (GE Healthcare). The column was washed with 1 column volume of de-gassed Milli-Q water and pre-equilibrated with 1 column volume of de-gassed 100 mM HEPES buffer. The protein was loaded onto the column. The protein was eluted with the same buffer and collected into fractions. SDS-PAGE was used to identify fractions containing the protein of interest. Protein concentration was tested with a NanoDrop ND-1000 spectrophotometer (Thermo Scientific). Once the protein solution had reached a concentration of 6 mg/ml, the sample was centrifuged (12,000 rpm, 4 °C, 10 min) to remove any precipitate. The supernatant was aliquoted and stored at -80 °C.

8.3.5 Purification Strategies

Protein	Final form	Purification strategy	Storage buffer
PHD2 (181-426)	PHD2 (181-426)	1. Ion Exchange Chromatography 2. Size Exclusion Chromatography	100 mM HEPES 150 mM NaCl pH 7.5
His_6-ZB-DdP4H	DdP4H	1. Ni(II) affinity chromatography 2. TEV cleavage overnight 3. Ni(II) Affinity Chromatography 4. Size Exclusion Chromatography	100 mM HEPES pH 7.6
His_6-ZB-TgPhyA	TgPhyA	1. Ni(II) affinity chromatography 2. TEV cleavage overnight 3. Ni(II) Affinity Chromatography 4. Size Exclusion Chromatography	100 mM HEPES pH 7.6
His_6-ZB-DdP4H (60-284)	DdP4H (60-284)	1. Ni(II) affinity chromatography 2. TEV cleavage overnight 3. Ni(II) Affinity Chromatography 4. Size Exclusion Chromatography	100 mM HEPES 50 mM NaCl pH 7.6
His_6-TgPhyA (53-264)	TgPhyA (53-264)	1. Ni(II) affinity chromatography 2. TEV cleavage overnight 3. Ni(II) Affinity Chromatography 4. Size Exclusion Chromatography	100 mM HEPES 50 mM NaCl pH 7.6
His_6-ZB-DdP4H (60-284) R192C	DdP4H (60-284) R192C	1. Ni(II) affinity chromatography 2. TEV cleavage overnight 3. Ni(II) Affinity Chromatography 4. Size Exclusion Chromatography	100 mM HEPES 50 mM NaCl pH 7.6
His_6-DdSkp1	DdSkp1	1. Ni(II) affinity chromatography 2. TEV cleavage overnight 3. Ni(II) Affinity Chromatography 4. Size Exclusion Chromatography	100 mM HEPES pH 7.6
His_6-TgSkp1	TgSkp1	1. Ni(II) affinity chromatography 2. TEV cleavage overnight 3. Ni(II) Affinity Chromatography 4. Size Exclusion Chromatography	100 mM HEPES pH 7.6
His_6-TEV	His_6-TEV	1. Ni(II) affinity chromatography 2. Size Exclusion Chromatography	100 mM HEPES 50 mM NaCl pH 7.6
HsSkp1-His_6	HsSkp1-His_6	1. Ni(II) affinity chromatography 2. Size Exclusion Chromatography	100 mM HEPES pH 7.6
HsSkp1 E147P-His_6	HsSkp1 E147P-His_6	1. Ni(II) affinity chromatography 2. Size Exclusion Chromatography	100 mM HEPES pH 7.6

Table 8.14 Protein purification conditions.

8.4 Mass Spectrometry

8.4.1 Hydroxylation Assay and Matrix-Assisted Laser Desorption Ionization Mass Spectrometry (MALDI-MS) [22]

Stock reagents were prepared: 5 mM 2OG, 10 mM ascorbate, 0.5 mM diammonium iron sulphate, 1 mM substrate peptide and 10 μM enzyme. The reagents were dissolved in 100 mM HEPES buffer (pH 7.6). Reactions were prepared according to Table 8.15.

Assay Conditions	Final Concentration	Volume
α-ketoglutaric acid disodium salt (Sigma-Aldrich, K3752)	500 μM	2 μL
Sodium L-ascorbate (Sigma-Aldrich, 11140)	1 mM	2 μL
Ammonium iron(II) sulfate hexahydrate (Sigma-Aldrich, 215406)	50 μM	2 μL
Substrate	100 μM	2 μL
Enzyme	1 μM	2 μL
Buffer (100 mM HEPES, pH 7.6)	-	10 μL
Total volume	-	20 μL

Table 8.15 Reactions for MALDI-MS.

Reagents were mixed immediately and placed in 37 °C incubator for 1 h. 20 μL formic acid was then added to the reaction mixture. The mixture was diluted into 100 μL using Milli-Q water. 1 μL sample was taken out, mixed with 1 μL MALDI-MS matrix (α-Cyano-4-hydroxycinnamic acid, Sigma-Aldrich®) and dropped onto a MALDI-MS plate. The sample was dried in air and subsequently analysed using Waters MALDI Micro MX.

8.4.2 Kinetic Assays of Prolyl Hydroxylases with MALDI-MS

For kinetic assays, each reaction was performed in triplicate. A triplicate of no enzyme control was performed at the same time.

Reagents were freshly prepared: 5 mM 2OG, 10 mM ascorbate, 0.5 mM diammonium iron sulphate, 1 mM substrate peptide and 10 µM enzyme. The reagents were all dissolved in 100 mM HEPES buffer (pH 7.6). Keep the concentration of one of the substrates/co-substrates (peptide/protein, or 2OG) various and maintaining fixed saturating concentrations of the other components. The reagents, except enzyme, were added into PCR tubes and mixed (Table 8.16). The enzyme was added into separate PCR tubes. The tubes were kept at 37 °C using a PCR instrument for 5 minutes to equilibrate.

Assay conditions	Final concentration	Volume
α-ketoglutaric acid disodium salt (Sigma-Aldrich, K3752)	500 µM	4 µL
Sodium L-ascorbate (Sigma-Aldrich, 11140)	1 mM	4 µL
Ammonium iron(II) sulfate hexahydrate (Sigma-Aldrich, 215406)	50 µM	4 µL
Substrate	100 µM	4 µL
Enzyme	1 µM	4 µL
Buffer (100 mM HEPES, pH 7.6)	-	20 µL
Total volume	-	40 µL

Table 8.16 Reaction Prepared for Kinetic Assays.

6 µL of the enzyme solutions were transferred into the mix tubes using a multi-channel pipette. A 2 µL sample of the reaction mixture was removed and quenched at time points 0 min, 1 min, 2 min, 4 min, 6 min, 8 min, 10 min, 12 min, 15 min, 20 min, 30 min and 60 min using 1 % (v/v) formic acid solution. 1 µL sample was taken out, mixed with 1 µL MALDI matrix (α-Cyano-4-hydroxycinnamic acid, Sigma-Aldrich®) and spotted onto MALDI plate. The sample was dried in air and subsequently analysed using Waters MALDI Micro MX.

The hydroxylation ratio of each sample was calculated and analysed using SciDAVis 1.D008.

8.4.3 Liquid Chromatography-Mass Spectrometry (LC-MS) [23]

Hydroxylation samples were analysed using a Xevo G2-S QToF mass spectrometer (Waters) equipped with an electrospray source and coupled to an Acquity (Waters) ultra performance liquid chromatography instrument unless otherwise specified. Instrument control and data processing were performed using MassLynx 4.1 software (Waters, Wilmslow UK). The column, Aeris^{TM} 3.6 µm WIDEPORE C4 200 Å (Phenomenex), was used for separation using a 10 min separation method with a gradient of solution A [Milli-Q water + 0.1 % (v/v) formic acid and solution B [acetonitrile + 0.1 % (v/v) formic acid] under standard calibration.

8.4.4 RapidFire Mass Spectrometry (RF-MS) [24]

Hydroxylation samples were analysed using RapidFire Mass Spectrometry (RF-MS). RF-MS is a high throughput MS based robotic platform that incorporates solid phase extraction technology to remove non-volatile buffer components from assay samples. All assays were performed in 384-well polypropylene v-bottom plates (Greiner Bio-One). Plates were then transferred to a RapidFire RF360 high throughput sampling robot (Agilent) connected to an Agilent accurate mass 6530 QTOF. Samples were aspirated under vacuum for 500 ms and loaded onto a C4 solid phase extraction (SPE) cartridge.

Data were determined from a nonlinear regression curve fit using GraphPad Prism 5.

8.4.5 Nanoscale Liquid Chromatography Coupled to Tandem Mass Spectrometry (Nano LC-MS/MS) [25, 26]

Tryptic digest peptides were analysed on a NanoAcquity-UPLC system (Waters®) coupled to an Orbitrap Elite™ mass spectrometer (Thermo Fischer Scientific™) [26] possessing an EASY-Spray nano-electrospray ion source (Thermo Fischer Scientific™). The peptides were trapped on a packed guard column (75 µm i.d. x 20 mm, Acclaim Pepmap100 C18, 3µm, 120 Å) using solvent

A [0.1% (v/v) formic acid]. 2µL of injected peptides were separated on an EASY-spray Acclaim PepMap® analytical column (75 µm i.d. × 15mm, RSLC C18, 3 µm, 100 Å) using a linear gradient [length: 68 minutes, 3 % to 97 % solvent B (0.1% (v/v) formic acid in acetonitrile, flow rate: 300 nL/min]. Data-dependent MS acquisition full-scan MS spectra were collected (scan range 350–1500 m/z, resolution 120 000, AGC target 1E6, maximum injection time 100 ms). After the MS scans, the 10 most intense peaks were selected for CID fragmentation at 35 % of normalised collision energy. CID spectra were acquired in the Iontrap (AGC target 1E3, maximum injection time 100 ms). Results were analysed using PEAKS Studio (Bioinformatics Solutions Inc.) [27, 28].

8.5 Crystallography

8.5.1 Crystallography Trials

95% pure protein solution was concentrated to ~20 mg/mL. N-Oxalylglycine (NOG, final concentration of 2 µM) and nickel(II) chloride (final concentration of 2 µM) was added into the protein solution. The sample was centrifuged (13,000 rpm, 4 °C, 10min) to remove any precipitation. The protein solution and the commercial crystallisation screen solutions were mixed and dispersed onto crystallography plates (HR3-119 CrystalMation Intelli-Plate 96-3 low-profile, Hampton Research) using The Phoenix RE Liquid Handling System (Art Robbins Instruments). The plates were then sealed using X-Clear Advanced Polyolefin StarSeal (STARLAB Ltd). The plates were stored at either 20 °C or 4 °C and were inspected weekly.

Crystallography screens	Company
JCSG-*PLUS*™	Molecular Dimensions Limited
SaltRx HT™	Hampton Research
PEG / ion HT ™	Hampton Research
INDEX HT	Hampton Research
Wizard	Emerald BioStructures
PACT Premier™ HT-96	Molecular Dimensions Limited
Silver Bullet HR2-088	Hampton Research
Silver Bullet HR2-096	Hampton Research
Detergent Screen HR2-406	Hampton Research

Table 8.17 Crystallography Screens.

A

0.1 M Tris pH 11 % PEG 6000 0.10 M NaCl	0.1 M Tris pH 14 % PEG 6000 0.10 M NaCl	0.1 M Tris pH 17 % PEG 6000 0.10 M NaCl	0.1 M Tris pH 20 % PEG 6000 0.10 M NaCl	0.1 M Tris pH 23 % PEG 6000 0.10 M NaCl	0.1 M Tris pH 26 % PEG 6000 0.10 M NaCl
0.1 M Tris pH 11 % PEG 6000 0.15 M NaCl	0.1 M Tris pH 14 % PEG 6000 0.15 M NaCl	0.1 M Tris pH 17 % PEG 6000 0.15 M NaCl	0.1 M Tris pH 20 % PEG 6000 0.15 M NaCl	0.1 M Tris pH 23 % PEG 6000 0.15 M NaCl	0.1 M Tris pH 26 % PEG 6000 0.15 M NaCl
0.1 M Tris pH 11 % PEG 6000 0.20 M NaCl	0.1 M Tris pH 14 % PEG 6000 0.20 M NaCl	0.1 M Tris pH 17 % PEG 6000 0.20 M NaCl	0.1 M Tris pH 20 % PEG 6000 0.20 M NaCl	0.1 M Tris pH 23 % PEG 6000 0.20 M NaCl	0.1 M Tris pH 26 % PEG 6000 0.20 M NaCl
0.1 M Tris pH 11 % PEG 6000 0.25 M NaCl	0.1 M Tris pH 14 % PEG 6000 0.25 M NaCl	0.1 M Tris pH 17 % PEG 6000 0.25 M NaCl	0.1 M Tris pH 20 % PEG 6000 0.25 M NaCl	0.1 M Tris pH 23 % PEG 6000 0.25 M NaCl	0.1 M Tris pH 26 % PEG 6000 0.25 M NaCl

B

0.1 M HEPES pH 11 % PEG 6000 0.10 M NaCl	0.1 M HEPES pH 14 % PEG 6000 0.10 M NaCl	0.1 M HEPES pH 17 % PEG 6000 0.10 M NaCl	0.1 M HEPES pH 20 % PEG 6000 0.10 M NaCl	0.1 M HEPES pH 23 % PEG 6000 0.10 M NaCl	0.1 M HEPES pH 26 % PEG 6000 0.10 M NaCl
0.1 M HEPES pH 11 % PEG 6000 0.15 M NaCl	0.1 M HEPES pH 14 % PEG 6000 0.15 M NaCl	0.1 M HEPES pH 17 % PEG 6000 0.15 M NaCl	0.1 M HEPES pH 20 % PEG 6000 0.15 M NaCl	0.1 M HEPES pH 23 % PEG 6000 0.15 M NaCl	0.1 M HEPES pH 26 % PEG 6000 0.15 M NaCl
0.1 M HEPES pH 11 % PEG 6000 0.20 M NaCl	0.1 M HEPES pH 14 % PEG 6000 0.20 M NaCl	0.1 M HEPES pH 17 % PEG 6000 0.20 M NaCl	0.1 M HEPES pH 20 % PEG 6000 0.20 M NaCl	0.1 M HEPES pH 23 % PEG 6000 0.20 M NaCl	0.1 M HEPES pH 26 % PEG 6000 0.20 M NaCl
0.1 M HEPES pH 11 % PEG 6000 0.25 M NaCl	0.1 M HEPES pH 14 % PEG 6000 0.25 M NaCl	0.1 M HEPES pH 17 % PEG 6000 0.25 M NaCl	0.1 M HEPES pH 20 % PEG 6000 0.25 M NaCl	0.1 M HEPES pH 23 % PEG 6000 0.25 M NaCl	0.1 M HEPES pH 26 % PEG 6000 0.25 M NaCl

C

0.1 M Tris pH 11 % PEG 6000 0.10 M LiCl	0.1 M Tris pH 14 % PEG 6000 0.10 M LiCl	0.1 M Tris pH 17 % PEG 6000 0.10 M LiCl	0.1 M Tris pH 20 % PEG 6000 0.10 M LiCl	0.1 M Tris pH 23 % PEG 6000 0.10 M LiCl	0.1 M Tris pH 26 % PEG 6000 0.10 M LiCl
0.1 M Tris pH 11 % PEG 6000 0.15 M LiCl	0.1 M Tris pH 14 % PEG 6000 0.15 M LiCl	0.1 M Tris pH 17 % PEG 6000 0.15 M LiCl	0.1 M Tris pH 20 % PEG 6000 0.15 M LiCl	0.1 M Tris pH 23 % PEG 6000 0.15 M LiCl	0.1 M Tris pH 26 % PEG 6000 0.15 M LiCl
0.1 M Tris pH 11 % PEG 6000 0.20 M LiCl	0.1 M Tris pH 14 % PEG 6000 0.20 M LiCl	0.1 M Tris pH 17 % PEG 6000 0.20 M LiCl	0.1 M Tris pH 20 % PEG 6000 0.20 M LiCl	0.1 M Tris pH 23 % PEG 6000 0.20 M LiCl	0.1 M Tris pH 26 % PEG 6000 0.20 M LiCl
0.1 M Tris pH 11 % PEG 6000 0.25 M LiCl	0.1 M Tris pH 14 % PEG 6000 0.25 M LiCl	0.1 M Tris pH 17 % PEG 6000 0.25 M LiCl	0.1 M Tris pH 20 % PEG 6000 0.25 M LiCl	0.1 M Tris pH 23 % PEG 6000 0.25 M LiCl	0.1 M Tris pH 26 % PEG 6000 0.25 M LiCl

D

0.1 M HEPES pH 0.2 M NaCl 17.5 % PEG6k 0.002 M NiCl2 0 % Glycerol	0.1 M HEPES pH 0.2 M NaCl 18.0 % PEG6k 0.002 M NiCl2 0 % Glycerol	0.1 M HEPES pH 0.2 M NaCl 18.5 % PEG6k 0.002 M NiCl2 0 % Glycerol	0.1 M HEPES pH 0.2 M NaCl 19.0 % PEG6k 0.002 M NiCl2 0 % Glycerol	0.1 M HEPES pH 0.2 M NaCl 19.5 % PEG6k 0.002 M NiCl2 0 % Glycerol	0.1 M HEPES pH 0.2 M NaCl 20.0 % PEG6k 0.002 M NiCl2 0 % Glycerol	0.1 M HEPES pH 0.2 M NaCl 20.5 % PEG6k 0.002 M NiCl2 1 % Glycerol	0.1 M HEPES pH 0.2 M NaCl 21.0 % PEG6k 0.002 M NiCl2 2 % Glycerol	0.1 M HEPES pH 0.2 M NaCl 21.5 % PEG6k 0.002 M NiCl2 3 % Glycerol	0.1 M HEPES pH 0.2 M NaCl 22.0 % PEG6k 0.002 M NiCl2 4 % Glycerol	0.1 M HEPES pH 0.2 M NaCl 22.5 % PEG6k 0.002 M NiCl2 5 % Glycerol	0.1 M HEPES pH 0.2 M NaCl 23.0 % PEG6k 0.002 M NiCl2 6 % Glycerol

8.6 Peptide Synthesis [29, 30]

Peptides were prepared by solid phase peptide synthesis using a Liberty Blue™ Automated Microwave Peptide Synthesiser (CEM Corporation) and N,N'-diisopropylcarbodiimide (DIC) as a coupling reagent. Peptides were cleaved from the Rink Amide MBHA resin (AGTC Bioproducts Ltd.) and side chain protecting groups removed by treatment with 600 µL of 97.5 % (v/v) trifluoroacetic acid, 2.5 % (v/v) triisopropylsilane. The peptides were then precipitated by the addition of 4 mL of cold ether. The mixture was centrifuged and the supernatant discarded. Dried peptides were resuspended in 1 mL of water (with the addition of acetonitrile if the peptides were insoluble in water) and lyophilised on a freeze dryer to remove residual traces of trifluoroacetic acid. Peptides were stored at -20 °C. The synthesized peptides are listed below:

Protein	Sequence	Number of amino acid residues
DdSkp1	FNIKNDFTPEEEEQIRKENE	20
	EIRKIFNIKNDFTPEEEEQIRKENE	25
	FNIKNDFTPEEEEQIRKENEWCEDK	25
	RKIFNIKNDFTPEEEEQIRKENEWCE	25
HsSkp1	FNIKNDFTEEEEAQVRKENQ	20
	FNIKNDFTEEEEAQVRKENQWCEEK	25
HsSkp1 E147P	FNIKNDFTPEEEAQVRKENQ	20
	FNIKNDFTPEEEAQVRKENQWCEEK	25
TgSkp1	FNIVNDFTPEEEAQVREENK	20
	FNIVNDFTPEEEAQVREENKWCEDA	25
	EIRRIFNIVNDFTPEEEAQVREENK	25
DdSkp1 F135C	CNIKNDFTPEEEEQIRKENE	20
DdSkp1 N136C	FCIKNDFTPEEEEQIRKENE	20
DdSkp1 I137C	FNCKNDFTPEEEEQIRKENE	20
DdSkp1 K138C	FNICNDFTPEEEEQIRKENE	20
DdSkp1 N139C	FNIKCDFTPEEEEQIRKENE	20
DdSkp1 D140C	FNIKNCFTPEEEEQIRKENE	20
DdSkp1 F141C	FNIKNDCTPEEEEQIRKENE	20
DdSkp1 T142C	FNIKNDFCPEEEEQIRKENE	20

Table 8.19 Peptide synthesis.

8.7 Construct Information

PHD2 (181-426)

```
        10          20          30          40          50          60
PNGQTKPLPA LKLALEYIVP CMNKHGICVV DDFLGKETGQ QIGDEVRALH DTGKFTDGQL

        70          80          90         100         110         120
VSQKSDSSKD IRGDKITWIE GKEPGCETIG LLMSSMDDLI RHCNGKLGSY KINGRTKAMV

       130         140         150         160         170         180
ACYPGNGTGY VRHVDNPNGD GRCVTCIYYL NKDWDAKVSG GILRIFPEGK AQFADIEPKF

       190         200         210         220         230         240
DRLLFFWSDR RNPHEVQPAY ATRYAITVWY FDADERARAK VKYLTGEKGV RVELNKPSDS


VGKDVF
```

DdP4H

```
        10          20          30          40          50          60
MHHHHHHSSG VDNKFNKERR RARREIRHLP NLNREQRRAF IRSLRDDPSQ SANLLAEAKK

        70          80          90         100         110         120
LNDAQPKGTE NLYFQSMDIS NLPPHIRQQI LGLISKPQQN NDESSSSNNK NNLINNEKVS

       130         140         150         160         170         180
NVLIDLTSNL KIENFKIFNK ESLNQLEKKG YLIIDNFLND LNKINLIYDE SYNQFKENKL

       190         200         210         220         230         240
IEAGMNKGTD KWKDKSIRGD YIQWIHRDSN SRIQDKDLSS TIRNINYLLD KLDLIKNEFD

       250         260         270         280         290         300
NVIPNFNSIK TQTQLAVYLN GGRYIKHRDS FYSSESLTIS RRITMIYYVN KDWKKGDGGE

       310         320         330         340         350         360
LRLYTNNPNN TNQKELKQTE EFIDIEPIAD RLLIFLSPFL EHEVLQCNFE PRIAITTWIY
```

DdP4H (60-284)

```
         10          20          30          40          50          60
MHHHHHHSSG VDNKFNKERR RARREIRHLP NLNREQRRAF IRSLRDDPSQ SANLLAEAKK

         70          80          90         100         110         120
LNDAQPKGTE NLYFQSMKIF NKESLNQLEK KGYLIIDNFL NDLNKINLIY DESYNQFKEN

        130         140         150         160         170         180
KLIEAGMNKG TDKWKDKSIR GDYIQWIHRD SNSRIQDKDL SSTIRNINYL LDKLDLIKNE

        190         200         210         220         230         240
FDNVIPNFNS IKTQTQLAVY LNGGRYIKHR DSFYSSESLT ISRRITMIYY VNKDWKKGDG

        250         260         270         280         290         300
GELRLYTNNP NNTNQKELKQ TEEFIDIEPI ADRLLIFLSP FLEHEVLQCN FEPRIAITTW


IY
```

DdP4H (60-284) R192C

```
         10          20          30          40          50          60
MHHHHHHSSG VDNKFNKERR RARREIRHLP NLNREQRRAF IRSLRDDPSQ SANLLAEAKK

         70          80          90         100         110         120
LNDAQPKGTE NLYFQSMKIF NKESLNQLEK KGYLIIDNFL NDLNKINLIY DESYNQFKEN

        130         140         150         160         170         180
KLIEAGMNKG TDKWKDKSIR GDYIQWIHRD SNSRIQDKDL SSTIRNINYL LDKLDLIKNE

        190         200         210         220         230         240
FDNVIPNFNS IKTQTQLAVY LNGGRYIKHC DSFYSSESLT ISRRITMIYY VNKDWKKGDG

        250         260         270         280         290         300
GELRLYTNNP NNTNQKELKQ TEEFIDIEPI ADRLLIFLSP FLEHEVLQCN FEPRIAITTW


IY
```

DdSkp1

```
         10          20          30          40          50          60
MHHHHHHSSG VDLGTENLYF QSMSKERMGD ARKVTLVSQE GDEFDVDIEV ASMSALIKTM

         70          80          90         100         110         120
VEEDSDCQES IPLPNVDTCI LKKIIEYCEH HHNNPPEEIP KPLKSSNLAE VVSEWDYQFI

        130         140         150         160         170         180
NENSDQKILF ALILAANYLN IKPLLDLSVA KVATMIKAKT PEEIRRIFNI VNDFTPEEEA

        190
QVREENKWCE DA
```

TgPhyA

```
         10          20          30          40          50          60
MHHHHHHSSG VDNKFNKERR RARREIRHLP NLNREQRRAF IRSLRDDPSQ SANLLAEAKK

         70          80          90         100         110         120
LNDAQPKGTE NLYFQSMALE DLLLKLPRHL VSQIKEEADD RPIFNSADSS TAGGLERNRR

        130         140         150         160         170         180
VYSDLSSFLS DSQFDRLVDP TNKAAAFVVV DKAIEDADVC RRARQEAEYL RSAGRFRQAS

        190         200         210         220         230         240
FGGGSSKSVD TRTRSDEIVW IRQQDLSELP ACTQIVSLFE EIGVAIDKCF ARLGENLEVQ

        250         260         270         280         290         300
QTELQLSVYP KGTAGYAAHK DTGSVSGADR FLTVILYLNE NWKPECGGCL RLHLDACRVD

        310         320         330         340
ISPCLGKLVL FRSEQLEHSV LPVACADRFA VTCWFSVRPR VAGQSGS
```

TgPhyA (56-264)

```
         10          20          30          40          50          60
SMLSDSQFDR LVDPTNKAAA FVVVDKAIED ADVCRRARQE AEYLRSAGRF RQASFGGGSS

         70          80          90         100         110         120
KSVDTRTRSD EIVWIRQQDL SELPACTQIV SLFEEIGVAI DKCFARLGEN LEVQQTELQL

        130         140         150         160         170         180
SVYPKGTAGY AAHKDTGSVS GADRFLTVIL YLNENWKPEC GGCLRLHLDA CRVDISPCLG

        190         200         210
KLVLFRSEQL EHSVLPVACA DRFAVTCWFS VRPR
```

TgSkp1

```
         10          20          30          40          50          60
MHHHHHHSSG VDLGTENLYF QSMSKERMGD ARKVTLVSQE GDEFDVDIEV ASMSALIKTM

         70          80          90         100         110         120
VEEDSDCQES IPLPNVDTCI LKKIIEYCEH HHNNPPEEIP KPLKSSNLAE VVSEWDYQFI

        130         140         150         160         170         180
NENSDQKILF ALILAANYLN IKPLLDLSVA KVATMIKAKT PEEIRRIFNI VNDFTPEEEA

        190
QVREENKWCE DA
```

HsSkp1

```
        10          20          30          40          50          60
MPSIKLQSSD GEIFEVDVEI AKQSVTIKTM LEDLGMDDEG DDDPVPLPNV NAAILKKVIQ

        70          80          90         100         110         120
WCTHHKDDPP PPEDDENKEK RTDDIPVWDQ EFLKVDQGTL FELILAANYL DIKGLLDVTC

       130         140         150         160         170
KTVANMIKGK TPEEIRKTFN IKNDFTEEEE AQVRKENQWC EEKLEHHHHH H
```

HsSkp1 E147P

```
        10          20          30          40          50          60
MPSIKLQSSD GEIFEVDVEI AKQSVTIKTM LEDLGMDDEG DDDPVPLPNV NAAILKKVIQ

        70          80          90         100         110         120
WCTHHKDDPP PPEDDENKEK RTDDIPVWDQ EFLKVDQGTL FELILAANYL DIKGLLDVTC

       130         140         150         160         170
KTVANMIKGK TPEEIRKTFN IKNDFTPEEE AQVRKENQWC EEKLEHHHHH H
```

8.8 References

1. Chan, W. T., Verma, C. S., Lane, D. P., and Gan, S. K. (2013) A comparison and optimization of methods and factors affecting the transformation of Escherichia coli, *Biosci Rep 33*.

2. Bachman, J. (2013) Site-directed mutagenesis, *Methods Enzymol 529*, 241-248.

3. Austrian, R. (1952) Bacterial transformation reactions, *Bacteriol Rev 16*, 31-50.

4. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

5. Aslanidis, C., and de Jong, P. J. (1990) Ligation-independent cloning of PCR products (LIC-PCR), *Nucleic Acids Res 18*, 6069-6074.

6. Gileadi, O., Burgess-Brown, N. A., Colebrook, S. M., Berridge, G., Savitsky, P., Smee, C. E., Loppnau, P., Johansson, C., Salah, E., and Pantic, N. H. (2008) High throughput production of recombinant human proteins for crystallography, *Methods Mol Biol 426*, 221-246.

7. Pheiffer, B. H., and Zimmerman, S. B. (1983) Polymer-stimulated ligation: enhanced blunt- or cohesive-end ligation of DNA or deoxyribooligonucleotides by T4 DNA ligase in polymer solutions, *Nucleic Acids Res 11*, 7853-7871.

8. pNIC-ZB (Plasmid #26107).

9. pNH-TrxT (Plasmid #26106).

10. XL10 Gold Ultracompetent Cells, (Agilent, Ed.).

11. Koontz, L. (2013) Agarose gel electrophoresis, *Methods Enzymol 529*, 35-45.

12. Howard, D. H. (1956) The preservation of bacteria by freezing in glycerol broth, *J Bacteriol 71*, 625.

13. Walker, J. M. (1994) Gradient SDS polyacrylamide gel electrophoresis of proteins, *Methods Mol Biol 32*, 35-38.

14. Sambrook, J., and Russell, D. W. (2006) SDS-Polyacrylamide Gel Electrophoresis of Proteins, *CSH Protoc 2006*.

15. Bornhorst, J. A., and Falke, J. J. (2000) Purification of proteins using polyhistidine affinity tags, *Methods Enzymol 326*, 245-254.

16. Jungbauer, A., and Hahn, R. (2009) Ion-exchange chromatography, *Methods Enzymol 463*, 349-371.

17. Cummins, P. M., Rochfort, K. D., and O'Connor, B. F. (2017) Ion-Exchange Chromatography: Basic Principles and Application, *Methods Mol Biol 1485*, 209-223.

18. Polayes, D. A., Parks, T. D., Johnston, S. A., and Dougherty, W. G. (1998) Application of TEV Protease in Protein Production, *Methods Mol Med 13*, 169-183.

19. Irvine, G. B. (2001) Determination of molecular size by size-exclusion chromatography (gel filtration), *Curr Protoc Cell Biol Chapter 5*, Unit 5 5.

20. Duong-Ly, K. C., and Gabelli, S. B. (2014) Gel filtration chromatography (size exclusion chromatography) of proteins, *Methods Enzymol 541*, 105-114.

21. Fekete, S., Beck, A., Veuthey, J. L., and Guillarme, D. (2014) Theory and practice of size exclusion chromatography for the analysis of protein aggregates, *J Pharm Biomed Anal 101*, 161-173.

22. Singhal, N., Kumar, M., Kanaujia, P. K., and Virdi, J. S. (2015) MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis, *Front Microbiol 6*, 791.

23. Chen, G., and Pramanik, B. N. (2008) LC-MS for protein characterization: current capabilities and future trends, *Expert Rev Proteomics 5*, 435-444.

24. Rohman, M., and Wingfield, J. (2016) High-Throughput Screening Using Mass Spectrometry within Drug Discovery, *Methods Mol Biol 1439*, 47-63.

25. Huynh, M. L., Russell, P., and Walsh, B. (2009) Tryptic digestion of in-gel proteins for mass spectrometry analysis, *Methods Mol Biol 519*, 507-513.

26. Michalski, A., Damoc, E., Lange, O., Denisov, E., Nolting, D., Muller, M., Viner, R., Schwartz, J., Remes, P., Belford, M., Dunyach, J. J., Cox, J., Horning, S., Mann, M., and Makarov, A. (2012) Ultra high resolution linear ion trap Orbitrap mass spectrometer (Orbitrap Elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes, *Mol Cell Proteomics 11*, O111 013698.

27. Zhang, J., Xin, L., Shan, B., Chen, W., Xie, M., Yuen, D., Zhang, W., Zhang, Z., Lajoie, G. A., and Ma, B. (2012) PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification, *Mol Cell Proteomics 11*, M111 010587.

28. McCafferty, J., Reid, R., Spencer, M., Hamp, T., and Fodor, A. (2012) Peak Studio: a tool for the visualization and analysis of fragment analysis files, *Environ Microbiol Rep 4*, 556-561.

29. Jensen, K. J. (2013) Solid-phase peptide synthesis: an introduction, *Methods Mol Biol 1047*, 1-21.

30. Behrendt, R., White, P., and Offer, J. (2016) Advances in Fmoc solid-phase peptide synthesis, *J Pept Sci 22*, 4-27.