Reaction mechanism of *h*OMPD and *Ca*AAD at atomic resolution

Dissertation

for the award of the degree
"Doctor rerum naturalium"
of the Georg-August-Universität Göttingen

within the doctoral program
"Biomolecules: Structure-Function-Dynamics"
of the Georg-August-University School of Science (GAUSS)

submitted by
Sören Rindfleisch
from Gehrden

Göttingen, 2018

Members of the Thesis Advisory Committee

Prof. Dr. Kai Tittmann	Department of Molecular Enzymology
	Georg-August University Göttingen
Prof. Dr. Holger Stark	Department of Structural Dynamics,
	Max-Planck-Institute for Biophysical Chemistry
Prof. Dr. Ricardo Mata	Department of Computational Chemistry and Biochemistry
	Georg-August University Göttingen

Members of the Examination Board

Prof. Dr. Simone Techert	Institute for X-Ray Physics
	Georg-August University Göttingen
Dr. Fabian Commichau	Department of General Microbiology
	Georg-August University Göttingen
Prof. Dr. Ulf Diederichsen	Department of organic and biomolecular chemistry
	Georg-August University Göttingen

Date of oral examination: 07.02.2019

Hereby I declare that I prepared the doctoral thesis in hand titled "Reaction mechanism of *h*OMPD and *Ca*AAD at atomic resolution" independently and with no sources and aids other than quoted. The thesis has not been submitted elsewhere.

Göttingen, 20.12.2018

Sören Rindfleisch

Contents

1 List of Figures

2 List of Tables

Abbrevations

A_{spec}	Specific catalytic activity [$^{U}/_{mg}$]
AA	Amino acid
***Ca*AAD**	Acetoacetate decarboxylase from *C. acetobutylicum*
AAD_{WT}	Acetoacetate decarboxylase
aaRS	Aminoacyl tRNA synthetase
AcAc	Acetoacetic acid
AcK	N^{ϵ}-acetyllysine
AcP	Acetonyl phosphonic acid
AcS	Acetonyl sulfonic acid
ADP	Atomic displacement parameter
Ala	Amino acid: Alanine (A)
Amido-UMP	6-amido-uridine 5'-monophosphate
APS	Ammonium persulfate
ATP	Adenosine 5'-triphosphate
Aza-UMP	6-azauridine-5'-monophosphate
BamHI	Restriction endonuclease BamHI
BFD	Benzoylformate decarboxylase
BMP	6-hydroxyuridine-5'-phosphate
C. acetobutylicum	*Clostridium acetobutylicum*
CD	Circular dichroism
CobB	NAD+-dependent deacetylases of *Escherichia coli*
CSS	Modified amino acid: S-Mercaptocysteine
CTP	Cytidine 5'-triphosphate
CV	Column volume
Da	Dalton
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
cDNA	Complementary DNA
DNaseI	Deoxyribonuclease I from bovine pancreas
dNTP	Deoxynucleotide triphosphate
dp	Differential power
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EG	Ethylene glycol
EO	1-(β-D-erythrofuranosyl)-orotic acid
ϵ_{280}	Molar extinction coefficient [M^{-1} cm^{-1}] at 280 nm
f.c.	Final concentration
f.sat.c.	Final saturation concentration
f.v.	Final volume
FEO	1-(β-D-erythrofuranosyl)-5-fluoroorotate
FO	Fluoroorotate
FPLC	Fast protein liquid chromatography

Gln	Amino acid: Glutamine (Q)
Glu	Amino acid: Glutamic acid (E)
Gly	Amino acid: Glycine (G)
GPDH	Glycerol-3-phosphate dehydrogenase
GSD	Ground-state destabilization
GST	Glutathion-S-transferase protein
GTP	Guanosine 5'-triphosphate
ddH_2O	Purified water
HbnThDP	2-(1-hydroxybenzyl) thiamin diphosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	Amino acid: Histidine (H)
His_6	Hexahistidine sequence for affinity chromatography
HM-UMP	6-hydroxymethyl-UMP
HP_i	Phosphite dianion
HRP	Horseradish peroxidase
IBE	Intrinsic phosphate binding energy
IEX	Ion exchange chromatography
Ile	Amino acid: Isoleucine (I)
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
k_{cat}	Turn over number [s^{-1}]
K_d	Dissociation constant [M^{-1}]
K_M	Michaelis constant [mM]
KinITC	Kinetic isothermal titration calorimetry
LB	Lysogeny broth
LBA	Lysogeny broth agar
LBHB	Low-barrier hydrogen bond
Leu	Amino acid: Leucine (L)
Lys	Amino acid: Lysine (K)
M. thermoautotrophicum	*Methanobacterium thermoautotrophicum*
MeAcAc	3-oxobutanoic acid methyl ester
Met	Amino acid: Methionine (M)
Mgy	Unit Mgray ($10^{6\,J}/_{kg}$)
MS	Mass spectrometry
MTh	α-mandelyl-thiamin
MW	Molecular weight
MWCO	Molecular weight cut-off
NAM	Nicotinamide
Nco1	restriction endonuclease NcoI
NMR	Nuclear magnetic resonance spectroscopy
OD_{600}	Optical density at 600 nm wavelength
ODU	Equivalent of 1 mL of OD_{600}
OMP	Orotidine 5'-monophosphate
***Ec*OMPD**	Orotidine-5'-monophosphate decarboxylase from *E. coli*

***Mt*OMPD**	Orotidine-5'-monophosphate decarboxylase from *M. thermoautotrophicum*
***Sc*OMPD**	Orotidine-5'-monophosphate decarboxylase domain from *S. cerevisiae*
***h*OMPD**	Human orotidine-5'-monophosphate decarboxylase domain of *h*UMPS
OMPD	Orotidine 5'-monophosphate decarboxylase
***Bs*OMPD**	Orotidine-5'-monophosphate decarboxylase from *Baillus subtilis*
***Pf*OMPD**	Orotidine-5'-monophosphate decarboxylase domain from *P.falciparum*
hOMPD	Genomic sequence of *h*OMPD
***hOMPD*$_{314\text{AMBER}}$**	Genomic sequence of *h*OMPD with UAG stop-codon on acsAA position 314
***h*OMPD**	Human orotidine-5'-monophosphate decarboxylase domain. Amino acids 224-480 of *h*UMPS
His$_6$-SUMO-*h*OMPD$_{224-480}$	Expression construct of *h*OMPD in *Escherichia coli*
***h*OMPD$_{314\text{AcK}}$**	*h*OMPD harboring AcK position 314
***h*OMPD$_{\text{WT}}$**	*h*OMPD wild type protein domain
His$_6$-GST-*h*OMPD	Expression construct of *h*OMPD in *Escherichia coli*
***h*OPRT**	Human orotate phosphoribosyltransferase
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PDC	Pyruvate decarboxylase
PEG 3350	Poly(ethylene glycol) with average Mn 3350 $^{\text{g}}/_{\text{mol}}$
PEG400	Poly(ethylene glycol) with average Mn 400 $^{\text{g}}/_{\text{mol}}$
PG	Propylene glycol
pH	*potentia Hydrogenii*
Phe	Amino acid: Phenylalanine (F)
PMSF	Phenylmethanesulfonylfluoride
PPG 400	Poly(propylene glycol) with average Mn 400 $^{\text{g}}/_{\text{mol}}$
Pro	Amino acid: Proline (P)
psi	Pound-force per square inch
rmsd	Root mean square deviation
RNA	Ribonucleic acid
rpm	Revolutions per minute
RT	Room temperature, 25 °C
S. cerevisiae	*Saccharomyces cerevisiae*
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
Ser	Amino acid: Serine (S)
SME	Modified amino acid: Methionine sulfoxide
SSHB	Short-strong hydrogen bond
SUMO	Small Ubiquitin-related modifier protein
TCA-UMP	6-thiocarboxamido-uridine 5'-monophosphate
TEMED	Tetramethylethylenediamine

TEV	Tobacco etch virus
TEV-protease	Tobacco Etch Virus nuclear-inclusion-a endopeptidase
ThDP	Thiamine diphosphate
Θ	Molar mean residual weight ellipticity [deg cm^2 $dmol^{-1}$]
Thr	Amino acid: Threonine (T)
TIM	Triose phosphate isomerase
TMAE	Trimethylammonium ethyl
Tris	Tris(hydroxymethyl)aminomethane
Trp	Amino acid: Tryptophan (W)
TSS	Transition state stabilization
TTP	Thymidine 5'-triphosphate
Tyr	Amino acid: Tyrosine (Y)
U	Enzymatic unit [$^{\mu mol}/_{min}$]
ULP1	Ubiquitin-like-specific protease 1 (SUMO-protease)
UMP	Uridine 5'-monophosphate
***h*UMPS**	Human uridine monophosphate synthetase
UV	Ultra violet
UV-Vis	Ultra violet and visible fraction of electromagnetic radiation
$^{v}/_{v}$	$^{volume}/_{volume}$
Val	Amino acid: Valine (V)
$^{w}/_{v}$	$^{weight}/_{volume}$
WT	Wild type
X-ray	X-radiation
XDS	X-ray detector software

3 Acknowledgements

First and foremost, I would like to thank Prof. Dr. Kai Tittmann for offering the opportunity to work on the projects presented on the following pages. I enjoy our discussions which always give helpful advise.

I would like to thank Prof. Dr. Holger Stark for participation in the Thesis Advisory Committee and all the insights he gave into cryo-electron microscopy. Thank you for the AAD-data which were unfortunately mentioned only very briefly due to temporal constrains in the last days of writing this thesis. Additionally, I really appreciated the invitations to the group retreats. Thank you Dr. Ashwin Chari for your help and expertise in crystallography. Especially, our lively discussions at the beamline were very helpful.

I want to thank Prof. Dr. Ricardo Mata, for being the third member of the Thesis Advisory Committee and all the help and explanations he offered during the time of the thesis. I admire what you and your group can achieve with the computational studies. Thanks Dr. Jon Uranga for all the discussions and the insights you gave into the electrostatic aspects of enzyme catalysis.

Thank you Prof. Dr. Simone Techert, Dr. Fabian Commichau and Prof. Dr. Ulf Diederichsen for being part the the thesis defence.

Dear Matthias Krull and Tobias Schmidt, I am very grateful for the inhibitors you synthesized and which were used in the crystallographic studies.

I would like to thank Prof. Dr. Ronald Kluger for the AcP and AcS samples.

For the access to the laboratories and instrumentation, I would like to thank Prof. Dr. Ralf Ficner. For the constant help with any crystallographic problem which appeared and the organisation of measurements, I would like to thank Dr. Piotr Neumann and Daniel Weinrich.

I would like to thank Dr. Heinz Neumann for the constructs to incorporate the acetyllysine residue into OMPD. Thank you, Dr. Oliver Valerius, for performing mass spectrometric analysis of the protein samples.

The synchrotron data was collected at EMBL Hamburg at the PETRA III storage ring (DESY). I would like to thank Dr. Gleb Bourenkov for the assistance in using the beamline.

Thank you, GGNB office, for managing the graduate school program and the perfect organisation.

I want to especially appreciate the help of and the discussions with my colleagues of the De-

partment of Molecular Enzymology. Furthermore, thanks a lot for proof-reading the manuscript and your valuable input.

Finally, I would like to thank my family and friends for all their support during the last four years. Thanks a lot, Britta, for your patience and for being there for me!

4 Materials and Devices

Chemicals:

Product	Supplier
Acetic acid	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Methyl acetoacetate	Sigma-Aldrich (Munich, Germany)
Acetonylphosphonic acid	Prof. Dr. Ronald Kluger
Acetonylsulfonic acid	Prof. Dr. Ronald Kluger
N^{ϵ}- Acetyllysine	Bachem Holding AG, (Bubendorf, Switzerland)
Acrylamide	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Adenosine 5'-triphosphate	Sigma-Aldrich (Munich, Germany)
Agar (bacteriology)	AppliChem GmbH (Darmstadt, Germany)
Agarose	AppliChem GmbH (Darmstadt, Germany)
Ammonium chloride	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Ammonium persulfate	Sigma-Aldrich (Munich, Germany)
Ammonium sulfate	AppliChem GmbH (Darmstadt, Germany)
L- Arabinose	Sigma-Aldrich (Munich, Germany)
Bovine serum albumin (BSA)	AppliChem GmbH (Darmstadt, Germany)
Bradford reagent	SERVA Electrophoresis GmbH (Heidelberg, Germany)
Bromphenole blue disodium salt	AppliChem GmbH (Darmstadt, Germany)
Calium chloride	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Carbenicillin disodium salt	AppliChem GmbH (Darmstadt, Germany)
Chloramphenicol	AppliChem GmbH (Darmstadt, Germany)
Coomassie Brilliant blue G-250	AppliChem GmbH (Darmstadt, Germany)
Cytidine 5'-triphosphate	ThermoFisher Scientific (Massachusetts, USA)
Dimethylsulfoxide (DMSO)	Sigma-Aldrich (Munich, Germany)
Dithiothreitol (DTT)	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Ethanol	THGeyer (Renningen, Germany)
Diethylether	Sigma-Aldrich (Munich, Germany)
Ethidium bromide	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Ethylene glycol	AppliChem GmbH (Darmstadt, Germany)
EDTA	AppliChem GmbH (Darmstadt, Germany)
D- Glucose	AppliChem GmbH (Darmstadt, Germany)

Glutathion	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Glycerol anhydrous	AppliChem GmbH (Darmstadt, Germany)
Glycine	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Guanidinium hydrochloride	AppliChem GmbH (Darmstadt, Germany)
Guanosine 5'-triphosphate	ThermoFisher Scientific (Massachusetts, USA)
Hellmanex II	Trademark of Hellma GmbH
Hydrochloric acid	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
HEPES	AppliChem GmbH (Darmstadt, Germany)
Imidazole	AppliChem GmbH (Darmstadt, Germany)
Isopropanol	AppliChem GmbH (Darmstadt, Germany)
IPTG	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Kanamycin sulfate	AppliChem GmbH (Darmstadt, Germany)
D- Lactose monohydrate	AppliChem GmbH (Darmstadt, Germany)
Magnesium chloride	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Magnesium sulfate	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Manganese(II) chloride	AppliChem GmbH (Darmstadt, Germany)
β- Mercaptoethanol	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
MES	AppliChem GmbH (Darmstadt, Germany)
Nickel(II) sulfate	AppliChem GmbH (Darmstadt, Germany)
Nicotinamide	Sigma-Aldrich (Munich, Germany)
N-methylglycine (Sarcosine)	Sigma-Aldrich (Munich, Germany)
Orotidine 5'-monophosphate	Prof. Dr. Ulf Diederichsen
Orotidine 5'-monophosphate	Sigma-Aldrich (Munich, Germany)
Phenylmethanesulfonylfluoride (PMSF)	AppliChem GmbH (Darmstadt, Germany)
PIPES	AppliChem GmbH (Darmstadt, Germany)
Poly(ethylene glycol) 3350	Hampton Research, (Aliso Viejo, USA)
Poly(ethylene glycol) 400	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Polysorbat 20 (Tween 20)	Trademark of Croda International PLC
Ponceau S	Sigma-Aldrich (Munich, Germany)
Potassium chloride	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Mono potassium phosphate	AppliChem GmbH (Darmstadt, Germany)
Di potassium phosphate	AppliChem GmbH (Darmstadt, Germany)

L- proline	Bachem Holding AG, (Bubendorf, Switzerland)
Propylene glycol	Sigma-Aldrich (Munich, Germany)
Sodium azide	AppliChem GmbH (Darmstadt, Germany)
Sodium chloride	AppliChem GmbH (Darmstadt, Germany)
Sodium dihydrogen phosphate	AppliChem GmbH (Darmstadt, Germany)
Sodium dodecyl sulfate	AppliChem GmbH (Darmstadt, Germany)
Sodium hydroxide	AppliChem GmbH (Darmstadt, Germany)
Mono sodium phosphate	AppliChem GmbH (Darmstadt, Germany)
Di sodium phosphate	AppliChem GmbH (Darmstadt, Germany)
Di sodium sulfate	AppliChem GmbH (Darmstadt, Germany)
Spectinomycin	Sigma-Aldrich (Munich, Germany)
Streptomycin sulfate	AppliChem GmbH (Darmstadt, Germany)
Tetracycline	AppliChem GmbH (Darmstadt, Germany)
TEMED	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Thymidine 5'-triphosphate	ThermoFisher Scientific (Massachusetts, USA)
TRIS hydrochloride	AppliChem GmbH (Darmstadt, Germany)
Tryptone	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
6-hydroxy- UMP	Prof. Dr. Ulf Diederichsen
6-aza- UMP	Prof. Dr. Ulf Diederichsen
6-amido- UMP	Prof. Dr. Ulf Diederichsen
6-thiocarboxamino- UMP	Prof. Dr. Ulf Diederichsen
Uridine 5'-monophosphate	Sigma-Aldrich (Munich, Germany)
Yeast extract	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)

Enzymes:

Product	Supplier
DNase I	AppliChem GmbH (Darmstadt, Germany)
FastAP	ThermoFisher Scientific (Massachusetts, USA)
Lysozyme	AppliChem GmbH (Darmstadt, Germany)
Taq-DNA polymerase	ThermoFisher Scientific (Massachusetts, USA)
SUMO-protease	provided by Dr. Stefan Lüdtke
T4-DNA-Ligase	ThermoFisher Scientific (Massachusetts, USA)

Bacterial strains:

Product	Supplier
E. coli-BL21 Star(DE3)	Invitrogen (Karlsruhe, Germany)
E. coli-SoluBL21(DE3)	Invitrogen, (Karlsruhe, Germany)
E. coli-XL1-Blue	Stratagene, (Heidelberg, Germany)

Vektors:

Product	Supplier
pET-SUMO	Invitrogen (Karlsruhe, Germany)
pET-28a	ThermoFisher Scientific (Massachusetts, USA)

Primers:

224+3_ SUMO	for	5'-	GGCGCCATGGAACTCAGCTTCGGTGCACG	-3'
seq_ rev	rev	5'-	TCAAACACCAAGTCTACTCAAATACG	-3'
314AMBER	for	5'-	CTTGATATTTGAAGACCGGTAGTTTGCAGATATAGGAAACACAG	-3'
314AMBER	rev	5'-	CTGTGTTTCCTATATCTGCAAACTACCGGTCTTCAAATATCAAG	-3'
T7Promotor	for	5'-	TAATACGACTCACTATAGGG	-3'
T7Terminator	rev	5'-	GCTAGTTATTGCTCAGCGG	-3'

Kits and solutions:

Product	Supplier
Champion pET-SUMO Protein Expression System	Invitrogen (Karlsruhe, Germany)
CloneJet PCR cloning Kit	ThermoFisher Scientific (Massachusetts, USA)
GC-buffer	ThermoFisher Scientific (Massachusetts, USA)
HF-buffer	ThermoFisher Scientific (Massachusetts, USA)
dNTP mix (10 mM)	ThermoFisher Scientific (Massachusetts, USA)
NucleoSpin Plasmid Kit	Macherey Nagel (Düren, Germany)
NucleoSpin Gel and PCR Clean-Up Kit	Macherey Nagel (Düren, Germany)
Bradford reagent, 5x concentrate	SERVA Electrophoresis GmbH (Heidelberg, Germany)

DNA and protein standards:

Product	Supplier
Gene Ruler 1kb DNA-Ladder	ThermoFisher Scientific (Massachusetts, USA)
PageRuler Unstained Protein Ladder	ThermoFisher Scientific (Massachusetts, USA)
Prestained Protein Molecular Weight Marker	ThermoFisher Scientific (Massachusetts, USA)
Unstained Protein Molecular Weight Marker	ThermoFisher Scientific (Massachusetts, USA)

Expression and purification

Product	Supplier
Incubation shaker, Unitron	Infors AG (Bottmingen, Switzerland)
Laminar flow Prettl-Telstar Bio-II-A	Telstar (Terrassa, Spain)
Microfluidizer, M-110S	Microfluidics (Newton, MA, USA)
ÄKTAprime plus	GE Healthcare
ÄKTApurifier	GE Healthcare
ÄKTApure	GE Healthcare
Ni-NTA	self packed
TMAE	self packed
HiPrep 26/10 desalting	GE Healthcare
Superdex 75 16/60	GE Healthcare
Superdex 200 HiLoad	GE Healthcare
Sephacryl HiPrep 26/60	GE Healthcare

Centrifuges and rotors:

Product	Supplier
Avanti HP-30I	Beckmann Coulter GmbH (Krefeld, Germany)
Rotor JA-10	Beckmann Coulter GmbH (Krefeld, Germany)
Rotor JA-30.50 Ti	Beckmann Coulter GmbH (Krefeld, Germany)
Avanti J-20XPI	Beckmann Coulter GmbH (Krefeld, Germany)
Rotor JLA-8.1000	Beckmann Coulter GmbH (Krefeld, Germany)
Centrifuge tubes	Beckmann Coulter GmbH (Krefeld, Germany)
Avanti JXN-26	Beckmann Coulter GmbH (Krefeld, Germany)
Rotor JA-25.50	Beckmann Coulter GmbH (Krefeld, Germany)
Eppendorf 5810R	Eppendorf AG (Wesseling-Berzdorf, Germany)
Rotor A-4-81	Eppendorf AG (Wesseling-Berzdorf, Germany)
Mikro 200	Hettich GmbH & Co. KG (Tuttlingen, Germany)
Rotor 2424 B	Hettich GmbH & Co. KG (Tuttlingen, Germany)
Universal 320R	Hettich GmbH & Co. KG (Tuttlingen, Germany)
Rotor 1420 A/B	Hettich GmbH & Co. KG (Tuttlingen, Germany)
Rotor 1617 A	Hettich GmbH & Co. KG (Tuttlingen, Germany)
Rotor 1620 A	Hettich GmbH & Co. KG (Tuttlingen, Germany)

Crystallography:

Product	Supplier
MicroMax-007 rotating anode X-ray generator	Rigaku Corp., (Michigan, USA)
X-stream 2000 Cryogenic Crystal Cooler	Rigaku Corp., (Michigan, USA)
Mar 345dtb image plate	marXperts GmbH (Norderstedt, Germany)
MicroMax-003 sealed tube X-ray generator	Rigaku Corp., (Michigan, USA)
800 Cryostream Cooler	Oxford Cryosystems (Oxford, UK)
PILATUS3 R 200K-A	DECTRIS Ltd. (Baden-Daetwill, CH)

Spectroscopy:

Product	Supplier
Chirascan plus CD spectrometer	Applied Photophysics Ltd., UK
Fluoromax 4 spectrofluorometer	HORIBA Europe GmbH, Oberursel
NanoDrop 2000	Thermo Scientific, USA
UV-vis spectrometer, V-650	Jasco GmbH, (Groß-Umstade, Germany)
UV-vis spectrometer, V-630	Jasco GmbH, (Groß-Umstade, Germany)
Suprasil precision cuvettes QS	Hellma GmbH & Co.KG (Mühlheim, Germany)

5 Methods

5.1 Polymerase chain reaction

In order to amplify specific DNA sequences and to introduce nucleotide sequence alterations, the polymerase chain reaction (PCR) was applied.[1] Phusion® DNA-polymerase was utilized and additional required components adjusted according to the manufacturer's manual. Designed DNA oligonucleotide primer sequences were purchased from Sigma-Aldrich (Sigma-Aldrich (Munich, Germany)) in desalted quality grade. The reactions were carried out in a thermocycler with 30 repeating cycles.

5.2 TA-cloning procedure

The TA-cloning procedure was used to incorporate a specific DNA-sequence into a linearized pET-SUMO vector. The required DNA-sequence was amplified applying PCR using *Thermus aquaticus* DNA-polymerase (*Taq*) according to the manufacturer's protocol. The *Taq*-specific addition of a 3'-deoxyadenosine nucleotide can anneal with a 3'-deoxythymidine flanking the linearized vector. Vector-Insert ligation was done with a freshly prepared amplification product in presence of T4 DNA Ligase overnight at 15 °C. The ligation reaction was used to transform chemical competent *E. coli* XL1-Blue cells and plated on selective 1.5 % LBA-plates (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl, 15 g/L agar, 50 µg/mL Kanamycin). A single colony was used to inoculate 10 mL LB (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl) and incubated overnight at 30 °C. The plasmid was isolated and stored at -20 °C (see. 5.5).

5.3 DpnI digestion of methylated DNA

To remove the methylated bacterial DNA which served as template for the amplification of a specific genomic DNA-region, the samples were treated with the restriction enzyme DpnI. Digestion was done according to the manufacturer's protocol.

5.4 Site-directed mutagenesis

To obtain amino acid substitution variants, site specific nucleotide sequence changes were introduced. The site directed mutagenesis reaction was performed according to the QuickChange™-protocol. An isolated and purified plasmid encoding the template sequence was used to generate a transcript including the sequence alteration. The template was amplified applying PCR with

complementary overlapping primers harbouring the desired nucleotide exchange. The amplification product was subjected to DpnI-digestion to remove the template plasmid. Inactivated reactions were directly used and introduced into chemically competent *E. coli* cells. The transformation reaction was plated on selective LB-plates, grown colonies cultivated and the plasmids isolated. The sequence was verified with Sanger-sequencing by a commercial company (GATC, Eurofins Genomics GmbH).

5.5 Isolation of plasmid DNA

Plasmids were isolated using the NucleoSpin® plasmid purification kit according to the manufacturer's manual (Macherey-Nagel).

5.6 Determination of DNA concentration

DNA concentrations were determined photometrically at 260 nm using a NanoDrop™ 2000 (ThermoFisher Scientific (Massachusetts, USA)) applying an average extinction coefficient and the calculation routines of the instrument.

5.7 Agarose gel electrophoresis

In order to obtain a size and conformation dependent separation of DNA-fragments, agarose gel electrophoresis was utilized. The required agarose gel (40 mM Tris/HCl, 20 mM acetic acid, 1 mM EDTA, pH 7.0; 1.5 %(w/v) agarose (low EEO grade)) was poured into a gel chamber. The DNA-sample was mixed with 6x DNA loading dye (ThermoFisher Scientific (Massachusetts, USA)) in 1:6 ratio, applied to the gel and electrophoresis was done at constant voltage (100 V) in TAE buffer (40 mM Tris/HCl, 20 mM acetic acid, 1 mM EDTA, pH 7.0). The gel was stained with ethidium bromide (2 μg/mL) to visualize DNA-fragments. The corresponding fragment size was determined with the GeneRuler 1 kb DNA-ladder (ThermoFisher Scientific (Massachusetts, USA)). For preparative purposes, the desired DNA fragment was extracted from the stained agarose gel and purified using a PCR-product clean-up kit.

5.8 Heat-shock transformation of chemical competent *E. coli*-cells

For the introduction of purified plasmid DNA into chemical competent *E. coli*-cells (BL21(DE3), BL21Star™(DE3), SoluBL21™(DE3)), 100 μL of frozen aliquots were thawed on ice. 5-10 ng of plasmid DNA were added, the cells and DNA carefully mixed and incubated for 15 min on

ice. The cells were heat-shocked at 42 °C for 45 s and incubated for 5 min on ice. 750 µL pre-warmed SOC-medium[2] (2 %(w/v) tryptone, 0.5 %(w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 20 mM $MgSO_4$, 20 mM glucose, pH 7.5) were added and the transformation reaction incubated for 60 min at 37 °C and 200 rpm. The cells were plated onto a LBA-plate (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl, 15 g/L agar) containing the required antibiotic and incubated for 16 h at 37 °C. 10 grown colonies were selected to inoculate a pre-culture of 250 mL LB-medium (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl) containing the required antibiotic.

5.9 Secondary structure determination using circular dichroism

To determine the secondary structure fingerprint of *h*OMPD of the wild type enzyme and variants, circular dichroism measurements were performed. The protein sample was transferred into potassium phosphate buffer (10 mM KH_2PO_4/KOH, pH 7.4). The sample was diluted in buffer to yield a concentration of approx. 0.1 mg/mL and the exact protein concentration was determined using the absorbance at 280 nm and the corresponding molar extinction coefficient. The measurement was done at 20 °C with 1 s accumulation time per data point and 30 repeating cycles from 180 - 280 nm with a cuvette path length of 1 mm. The resulting spectra were averaged. A buffer spectrum was recorded accordingly and subtracted from the protein sample. The molar mean residue weight ellipticity $[\Theta]_{MRW}$ was calculated from the measured ellipticity Θ, the molecular protein weight (MW), number of amino acids (n(AA)), the protein concentration c and the path length of the cuvette d according to:

$$[\Theta]_{MRW}(\text{deg cm}^2\ \text{dmol}^{-1}) = \frac{\Theta(\text{mdeg})}{1000} \cdot \frac{\frac{\text{MW(Da)}}{\text{n(AA)}-1}}{10 \cdot d(\text{cm}) c(\text{g/mL})} \quad (1)$$

$[\Theta]_{MRW}$ is plotted from 185 – 260 nm.[3] The corresponding absorbance spectrum was obtained from the identical protein sample.

5.10 Expression of the His_6-GST-*h*OMPD$_{224\text{-}480}$ construct

The published protein purification protocol was adjusted to obtain a higher yield of soluble protein.[4] SoluBL21™ (DE3) *E. coli*-cells harbouring the His_6-GST-*h*OMPD construct in pETM-30 were used to inoculate a 250 mL LB-medium (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl) pre-culture and incubated for 16 h at 30 °C and 200 rpm. The pre-culture was used to inoculate 500 mL M9-minimal medium (1 x M9 salt solution (10 x: 478 mM Na_2HPO_4, 220 mM KH_2PO_4, 86 mM NaCl, 187 mM NH_4Cl, pH 7.0), 0.1 mM $CaCl_2$, 1 mM $MgSO_4$, 0.2 x trace element solution

(1000 x: 50 mM $FeCl_3$, 20 mM $CaCl_2$, 10 mM $MnCl_2$, 10 mM $ZnSO_4$, 2 mM $CoCl_2$, 2 mM $CuCl_2$, 2 mM $NiCl_2$, 2 mM Na_2MoO_4, 2 mM Na_2SeO_3, 2 mM H_3BO_3), 0.3 %(v/v) glycerol) main cultures to an OD_{600} of 0.1. The cultures were incubated at 30 °C and 200 rpm until an OD_{600} of 0.6-0.8 was reached (approx. 5 h). Protein expression was induced by addition of IPTG to a f.c. of 0.1 mM. The cells were cultivated for 16 h at 30 °C and 200 rpm and harvested by centrifugation (Avanti JXN-26, JLA-8.1000, 4800 rpm, 4 °C). The cell pellet was frozen in $N_{2(l)}$ and stored at -80 °C.

5.11 Purification of His_6-GST-*h*OMPD$_{224\text{-}480}$ and *h*OMPD$_{WT}$

In order to purify the recombinant His_6-GST-*h*OMPD fusion protein, frozen cell pellets of 10 L M9-cultures were used. The pellet was resuspended in lysis buffer (50 mM Tris/HCl, 100 mM NaCl, pH 7.5) at 6 °C for 1 h. Once the pellet was completely resuspended, a tip of a spatula lysozyme, DNaseI (f.c. 5 $\mu g/mL$) and protease inhibitor PMSF (f.c. 1 mM) was added and incubated for 10 min. The cell disruption was performed using a fluidizer (Microfluidics) in 3 subsequent cycles (15000 psi) under ice cooled conditions. The cell lysate was clarified by centrifugation (Avanti JXN-26, JA-25.50, 30 min, 75000 xg, 10 °C) and the supernatant applied onto a GSTrap (GE Healthcare) affinity chromatography column. The sample was loaded with lysis buffer and the His_6-GST-*h*OMPD protein eluted with glutathion (50 mM Tris/HCl, 100 mM NaCl, 20 mM glutathion, pH 7.5). Elution fractions were pooled and diluted to 1 – 2 mg/mL with lysis buffer if necessary. The protein concentration was estimated spectrophotometrically at 280nm (NanoDrop™ 2000, ThermoFisher Scientific (Massachusetts, USA)). To remove the His_6-GST affinity tag, TEV-protease (Sigma-Aldrich (Munich, Germany), engineered catalytic domain of the NIa) was applied to a 0.5 %(w/w) ratio and incubated for 16 h at 6 °C.

The cleaved His_6-GST affinity tag and the labelled protease were removed with a HisTrap affinity chromatography column (GE Healthcare). The sample was loaded with lysis buffer and the His_6-GST protein eluted with imidazole (50 mM Tris/HCl, 100 mM NaCl, 300 mM imidazole, pH 7.5). The flow-through fractions were pooled, concentrated by ultra-filtration (Corning Spin-X UF, MWCO: 10k) and applied onto a S200 gel filtration column (HiLoad™ 16/60 Superdex™ 200, GE Healthcare) equilibrated with S200 buffer (20 mM Na/HEPES, pH 7.4).

5.12 Generation of the His_6-SUMO-*h*OMPD$_{224\text{-}480}$ expression construct

The sequence information of the *h*OMPD wild type protein domain was established in previous studies from a cDNA library (Swiss-Prot entry P11172).[4,5] To bypass the low solubility of the

applied His$_6$-GST-*h*OMPD construct, the coding sequence of the human OMPD-domain was transferred into a different expression system. As template for the amplification reaction using PCR served the His$_6$-GST-*h*OMPD$_{224\text{-}480}$ expressions construct. The nucleotide sequence of three additional N-terminal AAs necessary for and emerging from TEV-protease cleavage of the His$_6$-GST-*h*OMPD construct were appended to the *h*OMPD coding sequence to prevent attenuation of the crystallization propensity. The GAM-*h*OMPD protein is referred to as *h*OMPD$_{WT}$ in the following sections.

To generate the His$_6$-SUMO-*h*OMPD$_{224-480}$ expression construct, a TA-cloning procedure was applied.[6] The coding region of *h*OMPD$_{WT}$ was amplified applying PCR with *Taq* DNA polymerase and suitable primers (see: 5.2). The amplified fragment was treated with DpnI and ligated into the linearized vector pET SUMO according to the manufacturer's manual (Champion™ pET SUMO Expression System, ThermoFisher Scientific (Massachusetts, USA)).

The ligation reaction mixture was introduced into *E. coli* XL1-Blue cells using the heat-shock transformation protocol (see: 5.8). To select for the incorporation of the circularized pET SUMO plasmid, the transformation reaction was plated on selective 1.5 % LBA (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl, 15 g/L Agar, 50 µg/mL Kanamycin) and incubated overnight at 37 °C.

Grown colonies were used to inoculate 5 mL LB-medium (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl) and incubated overnight at 37 °C. After cell cultivation, plasmids were isolated and the sequence verified by sequencing (GATC, Eurofins Genomics GmbH). A sequence verified plasmid pool was introduced into *E. coli* SoluBL21 chemical competent cells using heat-shock transformation. The transformation reaction mixture was plated onto selective 1.5 %(w/v) LBA (50 µg/mL Kanamycin) and incubated overnight at 37 °C.

5.13 Generation of His$_6$-SUMO-*h*OMPD$_{224\text{-}480}$-variants

Protein variants were generated using the QuickChange™ protocol for site directed mutagenesis (see. 5.4).

5.14 Preparation of bacterial glycerol stocks

10 *E. coli* colonies containing the expression construct were used to inoculate a 250 mL Lysogeny broth-medium pre-culture and incubated overnight at 30 °C. A main culture was prepared and incubated according to the corresponding protein expression protocol (see: 5.15, 5.17, 5.32). At an OD_{600} of 0.6 - 0.8, 250 µL of the culture were mixed with a glycerol stock buffer (65 %(v/v) glycerol, 25 mM Tris, 0.1 M $MgSO_4$, pH 8.0) in 1:2 ratio and shock-frozen in $N_{2(l)}$. Glycerol

stocks were stored at -80 °C.

5.15 Expression of His_6-SUMO-*h*$OMPD_{224-480}$

To obtain functional *h*$OMPD_{WT}$ protein, *E. coli* SoluBL21(DE3) cells, containing the pET-SUMO vector harbouring the His_6-SUMO-*h*$OMPD_{224-480}$ construct were cultivated.[7,8] Based on the encoded T7-promotor sequence, the translation can be induced with the expression of the T7-polymerase. 10 colonies from the transformation reaction (see: 5.8) were used to inoculate a 250 mL LB-medium pre-culture and incubated overnight at 30 °C. For protein expression, 1 L of ZYM-5052 auto-induction medium (10 g/L tryptone, 5 g/L yeast extract, 1 mM $MgSO_4$, 25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 0.5 %(w/v) glycerol, 0.05 %(w/v) glucose, 0.2 %(w/v) α-lactose, 50 µg/mL Kanamycin) in a non-baffled 2 L flask was inoculated to an OD_{600} of 0.1.[9,10] The cultures were incubated at 37 °C under constant shaking (200 rpm) for 2 hours and then cooled to 30 °C and incubated for 48 h at 200 rpm. The cells were harvested by centrifugation (1 L, 4800 rpm, 4 °C, 15 min), washed in lysis buffer (50 mM Tris/HCl, 100 mM NaCl, 20 mM imidazole, pH 7.5) and centrifuged (8965 xg, 4 °C, 15 min) in a 50 mL tube. The cell pellets were frozen in $N_{2(l)}$ and stored at -20 °C.

5.16 Purification of *h*$OMPD_{224-480}$ and variants

To obtain purified *h*OMPD protein, frozen cell pellets (approx. 10 g) were resuspended in 45 mL lysis buffer (50 mM Tris/HCl, 100 mM NaCl, 20 mM imidazole, pH 7.5). Once the pellet is completely resuspended, a tip of a spatula lysozyme, DNaseI (f.c. 5 µg/mL) and protease inhibitor PMSF (f.c. 1 mM) were added and incubated for 10 min.

Cell disruption was performed using the fluidizer (Microfluidics) in three subsequent cycles (15000 psi). The cell lysate was clarified by centrifugation (30 min, 75000 xg, 10 °C) and loaded onto a His_6-affinity chromatography column (Ni-NTA FF 16/10, GE Healthcare) equilibrated with lysis buffer. Weak binding interactions were removed with 20 % elution buffer (50 mM Tris/HCl, 100 mM NaCl, 300 mM imidazole, pH 7.5) and His_6-SUMO-*h*$OMPD_{224-480}$ was eluted with 100 % elution buffer.

To prepare the elution fraction for further purification, the imidazole was removed using a desalting column (HiPrep™ 26/10, GE Healthcare) equilibrated with lysis buffer. Protein containing fractions were pooled and the total protein concentration determined photometrically (ϵ_{280}(His_6-SUMO-*h*$OMPD_{224-480}$)= 41.662 $M^{-1}cm^{-1}$).[11–14] The SUMO-tag was removed by addition of ULP1 in a 4000:1 m/m-ratio for 60 min at 6 °C under mild shaking. To remove insolubil-

ities after the cleavage process, the protein solution was filtered with a syringe tip filter (45 μm). ULP1 and SUMO were removed by His$_6$-affinity chromatography. The flow through fractions were pooled, concentrated by ultrafiltration to approx. 2 mL and applied to a gel filtration column (HiLoad™ 16/60 Superdex™ 75, GE Healthcare) equilibrated with S75-buffer (20 mM HEPES/NaOH, pH 7.4). Pure *h*OMPD fractions were pooled, concentrated by ultra-filtration to 30 $^{mg}/_{mL}$ and stored on ice.

5.17 Expression of His$_6$-SUMO-*h*OMPD$_{314AcK}$

In order to obtain His$_6$-SUMO-*h*OMPD$_{314AcK}$ protein, the amber-suppression system vector pCDF-AcKRS3/PylT and pET-SUMO-*h*OMPD$_{314AcK}$ were introduced into BL21(DE3)star cells. The transformation reaction was transferred to a LBA-plate supplemented with 75 $^{\mu g}/_{mL}$ spectinomycin and 50 $^{\mu g}/_{mL}$ kanamycin and incubated for 16 h at 37 °C. If possible, 10 colonies were used to inoculate 250 mL LB-medium containing 75 $^{\mu g}/_{mL}$ spectinomycin and 50 $^{\mu g}/_{mL}$ kanamycin and incubated for 16 h at 37 °C and 200 rpm. The pre-culture was used to inoculate 1 L LB (2 L non-baffled flask, blue cap) supplemented with 75 $^{\mu g}/_{mL}$ spectinomycin and 50 $^{\mu g}/_{mL}$ kanamycin to an OD_{600} of 0.1. The cultures were incubated at 37 °C and 200 rpm until an OD_{600} of 0.6 - 0.8 was reached.

Nicotinamide was added to a f.c. of 50 mM to inhibit potential lysine deacetylation by CobB and the cultures incubated for additional 30 min. To induce protein expression, IPTG and AcK were added to a f.c. of 1 mM and 10 mM, respectively. The cultures were incubated at 30 °C and 200 rpm for 24 h.

The cells were harvested by centrifugation (1 L, 4800 rpm, 4 °C, 15 min), washed in lysis buffer (50 mM Tris/HCl, 100 mM NaCl, 20 mM imidazole, pH 7.5) and centrifuged (8965 xg, 4 °C, 15 min) in a 50 mL tube. The cell pellets were frozen in $N_{2(l)}$ and stored at -20 °C.

5.18 Western blot analysis of *h*OMPD$_{314AcK}$ expression and solubility

In order to detect even low expression levels of the His$_6$-SUMO-*h*OMPD$_{314AcK}$ construct, western blot analysis was applied.[15] The complex protein samples from expression and purification were analysed using SDS-PAGE (see. 5.19) and a pre-stained protein marker (Pierce™ pre-stained protein MW marker). All components required for the western blotting procedure were thoroughly soaked in transfer buffer (25 mM Tris ultrapure, 125 mM glycine, 20 %($^{v}/_{v}$) ethanol). The SDS-PAGE gels (9.5 × 6.5 cm) were placed onto a nitrocellulose transfer membrane (Roti®-NC, Carl Roth GmbH + Co. KG (Karlsruhe, Germany)) and both sides were covered with chro-

matography paper (3 mm CHR, Whatman). The blot apparatus was assembled according to the manufacturer's guidelines (Mini-PROTEAN® Tetra System, Bio-rad) and filled with transfer buffer. To prevent undue heating during electrophoresis, a cooling element and a magnetic stirrer were added. The blotting procedure (80 V, 2 h) was performed at 6 °C under constant stirring.

To ensure protein transfer, the nitrocellulose membrane was stained with Ponceau S staining solution (0.1 %(w/v) Ponceau S, 1 %(v/v) acetic acid) which was removed with water afterwards. The polyacrylamide gel was stained with Coomassie (Brilliant blue G-250, AppliChem GmbH (Darmstadt, Germany)) to verify complete protein transfer.

The nitrocellulose membrane was blocked with 50 mL T-TBS buffer (10 mM Tris/HCl, 150 mM NaCl, 0.05 %(v/v) Tween20, pH 7.9) supplemented with 3 %(w/v) milk powder for 1 h. To specifically detect the His_6-SUMO-*h*$OMPD_{314AcK}$ construct, the membrane was incubated with a His_6-specific antibody(Pierce™ 6x-His tag monoclonal antibody, mouse IgG2b, ThermoFisher Scientific (Massachusetts, USA)) solved in T-TBS and 3 %(w/v) milk powder (50 mL, f.c. 1 $\mu g/mL$) at 6 °C for 16 h under mild shaking. The membrane was washed 4 × with 50 mL T-TBS and 3 %(w/v) milk powder and incubated with the secondary antibody (Pierce™ goat anti-mouse IgG (H+L) secondary antibody, HRP, ThermoFisher Scientific (Massachusetts, USA)) solved in T-TBS and 3 %(w/v) milk powder (50 mL, f.c. 80 ng/mL) at RT for 1 h. The membrane was washed 2× with T-TBS and 4× with TBS (10 mM Tris/HCl, 150 mM NaCl, pH 7.9) to remove free and non-specific secondary antibody.

To visualize the HRP-reaction, the membrane was packed into a transparent plastic foil and the substrates added (Pierce™ western blotting substrate, ThermoFisher Scientific (Massachusetts, USA)) depending on membrane size (0.03 mL/cm^2). A colorimetric image was taken and the chemiluminescence was recorded using a specific imaging system (ChemiDoc XRS+, Bio-rad) and signal accumulation (1 - 60 s, 6 images). To enable band size determination, the colorimetric image was merged with the chemiluminescence signal using the manufacturer's program (Image Lab™ software, Bio-rad).

5.19 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Discontinuous gels consisting of a stacking phase (5 % Acrylamide, 0.125 M Tris/HCl pH 6.8, 0.1 %(w/v) SDS) and a separation phase (12 % Acrylamide, 1.4 M Tris/HCl pH 8.8, 0.1 %(w/v) SDS) were poured between a glass plate and a ceramic plate in a gel caster (Hoefer, Massachusetts).[16]

5.20 Protein concentration determination

An inserted comb in the upper phase generated the sample wells. For loading, samples were mixed with 2x loading buffer (0.1 M Tris/HCl pH 6.8, 25 %(v/v) glycine, 2 %(w/v) SDS, 0.02 %(w/v) bromphenol blue, 2 %(v/v) mercaptoehanol) and applied into the sample wells. Gels ran in running buffer (0.25 M Tris, 2 M glycine, 1 %(w/v) SDS) with 35 A/Gel and max. 300 V. Gels were stained with coomassie stainer (0.25% Coomassie Blue, 30% ethanol, 6% acetic acid) for 10 min, subsequently destained in destaining solution (30% ethanol, 10% acetic acid) for 10 min, and afterwards kept in water. Gels were digitalized using CanoScan N650U and Adobe Photoshop CS3.

5.20 Concentration determination of purified protein samples

The protein concentration of purified protein samples was determined spectrophotometrically based on the intrinsic absorbance of tryptophane, cysteine and tyrosine residues. The absorbance at a wavelength of 280 nm (A_{280}) in a cuvette with path length (d) of 1 cm was recorded and the protein concentration calculated according to Lambert-Beer:[17–19]

$$A_{280} = c \cdot \epsilon \cdot d \tag{2}$$

The calculated molecular extinction coefficient was obtained from ProtParam.[11,13,14]

Table 1: Molecular weights and molar extinction coefficients of purified protein samples.

	SUMO-*h*OMPD$_{WT}$	*h*OMPD$_{WT}$	*h*OMPD$_{314AcK}$	*Ca*AAD
MW	41.663 kDa	28.265 kDa	28.296 kDa	27.255 kDa
ϵ_{280}	19940 $M^{-1}cm^{-1}$	18450 $M^{-1}cm^{-1}$	18450 $M^{-1}cm^{-1}$	26360 $M^{-1}cm^{-1}$

5.21 Preparation of *h*OMPD glycerol stocks

For long term protein storage, glycerol stocks were prepared. Freshly purified *h*OMPD protein was concentrated by ultra-filtration to 40 mg/mL and mixed with 50 %(v/v) glycerol (diluted with 20 mM HEPES/NaOH, pH 7.4) in a 1:2 ratio (25 %(v/v) glycerol). The mixture was split into 500 µL aliquots and frozen in $N_{2(l)}$ and stored at -80 °C.

5.22 UMP binding analysis using ITC

*h*OMPD$_{WT}$ has a certain affinity to its natural reaction product UMP. In order to analyse the binding affinity and to obtain insights into the binding mode of *h*OMPD$_{WT}$ to UMP, ITC binding

assays were performed. The protein storage stock (approx. 30 mg/mL) was centrifuged to remove any insolubilities (21382 xg, 4 °C, 30 min). The S75 buffer (20 mM HEPES/NaOH, pH 7.4) from size exclusion chromatography, which is required for instrument rinsing/washing, reactant dissolution and protein dilution steps, was pre-heated to 25 °C and thoroughly degased. UMP disodium salt (Sigma-Aldrich (Munich, Germany)) was dissolved in S75 buffer to obtain a f.c. of 10 mM (pH not adjusted) and constantly kept on ice. The OMPD concentration was determined using the spectrophotometric assay (see: 5.20) and adjusted to an active site concentration of 500 μM. Before use, all components required for a measurement were placed in the vacuum unit of the ITC and degased at 25 °C for 15 min. Titrations were performed with 20 or 30 injections using an ITC_{200} or VP-ITC, respectively.

The differential power signal was baseline corrected and integrated using the peak-shape analysis software NITPIC.[20,21] Further analysis was done using a python based script simulating titration reactions.[**Adair1925**, 22,23]

5.23 Photometric determination of enzyme kinetic constants of OMPD

To analyse enzyme catalysed OMP decarboxylation under steady-state conditions, a direct photometric assay was applied. The reactions were prepared in S75-buffer (20 mM HEPES/NaOH, pH 7.4) with varying concentrations of the OMP substrate. The reaction progression was monitored by the absorbance depletion at 285 nm (Δ_{285} = -1650 M^{-1} cm^{-1}).[5] The reactions were performed at 25 °C for 300 s and the absorbance measured using a spectrophotometer (Jasco V-650) in a quartz cuvette (Hellma Analytics, QS, 10 mm, 200 μL). In case of the wild type protein a final concentration of approx. 0.15 μmol of enzyme was used. The reaction was started by adding enzyme to the substrate solution in a 1:10 ratio. For $h\mathrm{OMPD}_{314\mathrm{AcK}}$, the enzyme concentration was 7.6 μmol. The aquired data points were fitted according to the Michaelis and Menten equation:[24]

$$v = \frac{v_{\max} \cdot [\mathrm{S}]}{K_m + [\mathrm{S}]} \tag{3}$$

5.24 Determination of enzyme kinetic constants using ITC

To determine the enzyme kinetic constants and to better resolve the reaction velocity of low substrate concentrations of $h\mathrm{OMPD}_{\mathrm{WT}}$ with OMP, isothermal titration calorimetry was utilized.[25,26] The protein storage stock (approx. 30 mg/mL) was centrifuged to remove any insolubilities

(21382 xg, 4 °C, 30 min). The S75 buffer (20 mM HEPES/NaOH, pH 7.4) from size exclusion chromatography which, is required for rinsing/washing, dissolution, dilution steps, was pre-heated to 25 °C and thoroughly degased. The OMP substrate (orotidine 5′-monophosphate trisodium salt, Sigma-Aldrich (Munich, Germany)) was dissolved in S75 buffer to obtain a f.c. of 10 mM. The pH-value was adjusted to 7.4. Stock solutions of 100 µL were prepared and frozen in $N_{2(l)}$.

The $h\text{OMPD}_{\text{WT}}$ concentration was determined using the spectrophotometric assay (see: 5.20) and adjusted to an active site concentration of 20 µM. The protein solution was loaded into the syringe and 10 µL were injected into the reaction cell containing 1 mM OMP substrate. The complete reaction was monitored until the differential power signal returned to the baseline value. A second injection of 10 µL enzyme ensured complete substrate consumption (detailed titration scheme in table: 14).

Data points were fitted according to the Hill-equation:

$$v = \frac{v_{\max} \cdot [\text{S}]^n}{{K_m}^n + [\text{S}]^n} \tag{4}$$

5.25 Crystallisation of *h*OMPD

In order to perform X-ray-crystallography studies and to obtain three dimensional structural information, protein crystals were prepared. In case of $h\text{OMPD}_{\text{WT}}$, both, freshly prepared protein and frozen protein stocks showed the propensity to crystallize. Preferably, fresh protein preparations were used.

Crystallisation of *h*OMPD was performed in VDX_{m} hanging-drop crystallisation plates (Hampton Research, (Aliso Viejo, USA)) according to a modified protocol from Wittmann *et al.*, 2007.[4] The purified enzyme, stored in S75 buffer (20 mM HEPES/NaOH, pH 7.4), was centrifuged to remove potential aggregates (30 min, 21282 xg, 4 °C), diluted to 5 $^{\text{mg}}/_{\text{mL}}$ and mixed with reservoir solution. Each reservoir well contained 400 µL mother liquor with four replicates on each plate and six increasing ammonium sulfate concentrations. 2 drops were placed on a 18 mm siliconized circular cover slide (Jena Bioscience) and placed to cover the reservoir well. Preparation of the crystal plates and crystal growth was done at 20 °C in a climate controlled environment.

To obtain resting state crystals, 2 µL reservoir solution (100 mM Tris/HCl pH 8.0, 1.8 - 2.0 M $(NH_4)_2SO_4$, 10 mM freshly prepared glutathion, 5 %(v/v) glycerol) were mixed with protein solution in a 1:2 ratio. Crystals appeared after approx. 3 days with a size of 100 – 200 µm. The crystallization plates were incubated for approx. 10 days at 20 °C to enable crystal growth to the

final size.

To prepare the crystals for plunge freezing in $N_{2(l)}$, they were transferred into a glycerol containing cryo-protection solution using a crystal loop (Molecular Dimensions). As stabilisation procedure, the crystal drop was supplemented with 2 µL stabilisation solution (100 mM Tris/HCl pH 8.0, 2.0 M $(NH_4)_2SO_4$, 10 mM freshly prepared glutathion, 5 %(v/v) glycerol). The transfer into the cryo-protection (100 mM Tris/HCl pH 8.0, 2.0 M $(NH_4)_2SO_4$, 10 mM freshly prepared glutathion, 25 %(v/v) glycerol) was conducted applying a linear 4 step procedure with mixtures of stabilisation and cryo-protection solution. The crystals were fished using a crystal loop (LithoLoop, Molecular Dimensions) and plunged into $N_{2(l)}$.

Alternatively, a L-proline based cryo protection solution was used for plunge freezing protein crystals.[27] For stabilization, a resting state crystal containing drop was supplemented with 2 µL proline cryo solution (100 mM Tris/HCl pH 8.0, 2.0 M $(NH_4)_2SO_4$, 10 mM freshly prepared glutathion, 5 %(v/v) glycerol, 1 M L-proline) and incubated for 1 min. Subsequntly, suitable crystals were transferred into pure proline cryo solution, the cover slide placed back onto the reservoir well, and incubated for 5 min. The crystals were fished and plunged in $N_{2(l)}$.

To crystallize the UMP product complex, *h*OMPD was mixed with UMP solved in 20 mM Na/HEPES pH 7.4 (100 mM stock from UMP disodium salt) to obtain a f.c. of 25 mM and incubated for 15 min at 20 °C. 2 µL protein solution were supplemented with reservoir solution (100 mM Tris/HCl pH 7.8, 1.6 – 1.8 M $(NH_4)_2SO_4$, 10 mM freshly prepared glutathion) in a 1:1 ratio. Crystals appeared after approx. 3 days with a size of 100 – 200 µm. To obtain final crystal size, the crystallisation plates were incubated for approx. 10 days at 20 °C. The drops were supplemented with 2 µL stabilization solution (100 mM Tris/HCl pH 7.8, 1.8 M $(NH_4)_2SO_4$, 25 mM UMP, 10 mM freshly prepared glutathion) and the crystals directly transferred into cryo-protection solution (100 mM Tris/HCl pH 7.8, 1.8 M $(NH_4)_2SO_4$, 25 mM UMP, 10 mM freshly prepared glutathion, 20 %(v/v) glycerol).

Alternatively, a L-proline based cryo protection solution was used (100 mM Tris/HCl pH 7.8, 1.8 M $(NH_4)_2SO_4$, 25 mM UMP, 10 mM freshly prepared glutathion, 5 %(v/v) glycerol, 1.5 M L-proline) and crystals treated identical to the resting state. The crystals were fished using a crystal loop (LithoLoop, Molecular Dimensions) and plunged into $N_{2(l)}$.

In order to obtain BMP inhibitor complex crystals, *h*OMPD was mixed with BMP dissolved in 20 mM Na/HEPES pH 7.4 to obtain a f.c. of 10 mM and incubated for 15 min at 20 °C. The co-crystallisation set up, behaviour and post crystallisation treatment was identical to the UMP product complex crystals.

5.25 Crystallisation of hOMPD

Since the crystal growth was impaired in presence of certain inhibitors, the respective ligands were soaked into resting state crystals as a post-crystallisation treatment. *h*OMPD resting state crystals were prepared as described above. In case of 6-amido-UMP, crystals with a size of approx. 100 – 200 μm were selected and transferred into an amido-UMP containing cryo-protection solution (12.5 mM 6-amido-UMP, 100 mM Tris/HCl pH 8.0, 2.0 M $(NH_4)_2SO_4$, 10 mM freshly prepared glutathion, 5 %(v/v) glycerol, 1 M L-proline) in a four step procedure. Each step was incubated for 30 min covering the reservoir well.

To obtain 6-aza-UMP containing protein complexes, drops with resting state crystals with a size of approx. 100 – 200 μm were mixed with 2 μL stabilization solution (12.5 mM 6-aza-UMP, 100 mM Tris/HCl pH 8.0, 2.0 M $(NH_4)_2SO_4$, 10 mM freshly prepared glutathion, 5 %(v/v) glycerol, 1 M L-proline) and incubated for 1 min covering the reservoir well. The crystals were transferred into cryo-protection solution in a single step and incubated covering the reservoir well for 30 min. Afterwards, the crystals were fished and frozen in $N_{2(l)}$.

Protein crystals complexed with 6-thiocarboxamido-UMP were obtained in an identical procedure as described for 6-aza-UMP. The cryo-protection solution (12.5 mM 6-TCA-UMP, 100 mM Tris/HCl pH 8.0, 2.0 M $(NH_4)_2SO_4$, 10 mM freshly prepared glutathion, 5 %(v/v) glycerol, 1 M L-proline) was used for stabilization and freezing.

In order to generate *h*OMPD$_{314AcK}$ resting state crystals, the reservoir solution with an increased ammonium sulfate concentration was used (100 mM Tris/HCl pH 8.0, 1.9 – 2.1 M $(NH_4)_2SO_4$, 10 mM freshly prepared glutathion, 5 %(v/v) glycerol). Crystal growth was comparable to the *h*OMPD$_{WT}$ protein. Crystals with a size of approx. 100 – 200 μm were transferred into cryo-protection solution (100 mM Tris/HCl pH 8.0, 2.0 M $(NH_4)_2SO_4$, 10 mM freshly prepared glutathion, 5 %(v/v) glycerol, 1 M L-proline) and incubated for 5 min covering the reservoir well. Crystals were fished and frozen in $N_{2(l)}$.

For the UMP-product complex, *h*OMPD$_{314AcK}$ was mixed with a UMP-solution (100 mM stock from UMP disodium salt dissolved in 20 mM HEPES/NaOH, pH 7.4) to a f.c. of 12.5 mM and incubated for 15 min at RT. 2 μL of protein solution were mixed with the reservoir (100 mM Tris/HCl pH 7.8, 1.6 – 1.8 M $(NH_4)_2SO_4$, 10 mM freshly prepared glutathion, 5 %(v/v) glycerol) and incubated according to the *h*OMPD$_{WT}$ protocol. Crystals with a size of approx. 100 – 200 μm were transferred into cryo-protection solution (100 mM Tris/HCl pH 7.8, 1.8 M $(NH_4)_2SO_4$, 10 mM freshly prepared glutathion, 5 %(v/v) glycerol, 1.5 M L-proline) and incubated for 1 min. Crystals were fished and frozen in $N_{2(l)}$.

OMP soaks into *h*OMPD$_{314AcK}$ resting state crystals were performed without cryo-protection

solution. Crystals were prepared as described above and soaked with an OMP-solution (50 mM OMP, 100 mM Tris/HCl pH 8.0, 2.0 M $(NH_4)_2SO_4$, 10 mM freshly prepared glutathion, 5 %(v/v) glycerol) with a duration of 1 – 10 min. Crystals were directly frozen in $N_{2(l)}$.

5.26 In-house X-ray cryo-crystallography measurements

To asses crystal diffraction quality, cryo-protection suitability and to generate initial models for phasing of synchrotron data, frozen crystals were tested using a rotating anode or a sealed tube X-ray generator (Rigaku, MicroMax™-003, MicroMax™-007: Department of Molecular Structural Biology (MSB)).

5.27 Synchrotron X-ray diffraction data collection and processing of *h*OMPD

All datasets were measured under cryogenic conditions at the "Deutsches Elektronen-Synchrotron (DESY)" in Hamburg at the beamlines P13 and P14. To prevent radiation induced damage of $h\text{OMPD}_{\text{WT}}$ protein crystals, the radiation dose was determined before each measurement and the total radiation dose adjusted to 1 MGy. Protein crystals of $h\text{OMPD}_{314}$ in complex with OMP were exposed to a total dose of 0.5 MGy to prevent radiation induced decarboxylation of the substrate molecule.

X-ray diffraction raw images of single crystals were collected under cryogenic temperatures (100 K). The synchrotron SAXS/MX data was collected at beamline P13/P14 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany).[28] The protein crystals were exposed to max. 1 MGy radiation for a complete data-set. Diffraction data analysis, integration, correction, scaling, post-refinement and space-group assignment was done with the XDS program-suite.[29,30] The automated space-group determination was evalutated using "pointless" and "phenix.xtriage".[31,32] The resolution cut-off decision was mainly made on the $CC_{(1/2)}$ (> 50 %), $I/\sigma(I)$ (> 1) and the completeness of data (> 95 %).[33–35] Data was converted to the MTZ reflection format using "XDSCONV" and "f2mtz" and R-free reflections were introduced using "XDSCONV" or "uniqueify".[29,36]

5.28 Phasing and initial structure determination of OMPD

As starting models for $h\text{OMPD}_{\text{WT}}$ and variant protein crystal data phasing, the PDB-entries 2QCD and 2QCE were used for data-sets refined to space-group $C222_1$ and $P2_1$, respectively.[4,5,37].

In most cases, the unit cell constants deviated $\leq 1\,\%$ from the input model parameters and the initial phasing was done with a rigid-body refinement.[38,39]

Data-sets with larger deviations from the deposited input structures were phased applying molecular replacement with "molrep" and the monomer of structure 2QCD.[36,40]

5.29 Model building and structure refinement of *h*OMPD

The obtained 3D crystal structures were corrected and adjusted manually against the diffraction data using the "COOT" software suite.[41] Potential ligands were drawn and the CIF-files generated using "Jligand".[42] Model adjustment and data refinement was done in an iterative manner using "COOT" and "refmac5".[38,39,41] All structures were refined with anisotropic atomic displacement parameters (ADP) and hydrogen atoms in riding positions. To obtain the estimated standard deviation values of bond length and angles, a least-square refinement with SHELXL was done using the final model as input structure.[43,44]

Protein structure representation and image generation was done with "PyMol" and "Chimera".[45,46] Dataset and refinement statistics were generated with phenix.table_one.[32]

5.30 Analysis of expression efficiency and protein solubility

In order to analyse the quantity of soluble protein of bacterial cultures, the cells were lysed and the protein distribution examined via SDS-PAGE. Therefore, 150 ODU were transferred into a 50 mL reaction tube and centrifuged (10 min, 4 °C, 8965 xg). The pellet was resuspended with 10 mL of the corresponding lysis buffer and the cells lysed using sonification (1 min, 40 % pulse). The lysate was clarified by centrifugation (30 min, 21282 xg, 4 °C) and the protein distribution of supernatant and pellet analysed using SDS-PAGE.

5.31 Generation of AAD expression construct

The coding nucleotide sequence of acetoacetate decarboxylase (*Ca*AAD) from *Clostridium thermoautotrophicum* was obtained from the Protein Data Bank entry 3BH2 by reverse translation.[47] The *E. coli* optimized nucleotide sequence was generated and synthesized by the GeneArt® Gene Synthesis service (ThermoFisher Scientific (Massachusetts, USA)). Due to the initial cloning strategy of the published expression construct including 5'-processing with Nco1, the applied expression construct differs from the WT sequence on position 4 (C → G). The resulting amino acid exchange (Leu → Val) was kept in order to prevent changes in the crystallization propensity.

The synthesized insert was introduced into pET-28a using the BamHI and Nco1 restriction sites.

Sequence verified plasmids were introduced into *E. coli* BL21(DE3)star chemical competent cells using heat-shock transformation (see 5.8). The transformation reaction mixture was plated onto selective 1.5 % LBA (50 $\mu g/mL$ Kanamycin) and incubated overnight at 37 °C.

10 grown colonies were used to inoculate a 250 mL Lysogeny broth-medium pre-culture and were incubated overnight at 30 °C.

5.32 Expression of *Ca*AAD

A small amount of a glycerol stock culture of transformed *E. coli* BL21(DE3)star cells was used to inoculate a 250 mL Lysogeny broth-medium pre-culture. The medium was incubated overnight at 30 °C at constant shaking with 200 rpm. For protein expression, 1 L of ZYM-5052 auto-induction medium (10 g/L tryptone, 5 g/L yeast extract, 1 mM $MgSO_4$, 25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 0.5 %(w/v) glycerol, 0.05 %(w/v) glucose, 0.2 %(w/v) α-lactose) in 2 L non-baffled flasks was inoculated to an OD_{600} of 0.1.[9,10] The cultures were incubated at 37 °C under constant shaking (200 rpm) for 2 hours and then cooled to 30 °C. Final incubation was performed for 48 h at 30 °C and 200 rpm. The cells were harvested by centrifugation (30 min, 4000 rpm, rotor: JLA-8.1000 4 °C), washed in lysis buffer (50 mM KH_2PO_4, pH 5.9), and centrifuged (15 min, 8965 xg, 4 °C). The pellets were frozen in $N_{2(l)}$ and stored at -20 °C.

5.33 Purification of *Ca*AAD

The purification procedure was adjusted from published methods.[48] Frozen cell pellets were resuspended in lysis buffer (50 mM KH_2PO_4, pH 5.9). 5 mL of lysis buffer were utilized to resuspend 5 g cells for 1 h at 6 °C to reach a homogeneous suspension. To prepare the cells for disruption, a tip of a spatula lysozyme and DNaseI (f.c. 5 $\mu g/mL$) were added under constant stirring for 10 min. Before cell lysis, the protease inhibitor PMSF (f.c. 1 mM) was added and incubated for 10 min. Cell disruption was performed using the fluidizer (Microfluidics) in 3 subsequent cycles (80 psi). The cell lysate was clarified by centrifugation (30 min, 75,000 xg, 10 °C) and the volume of the supernatant determined.

To remove impurities, $(NH_4)_2SO_4$ was added to a f.sat.c. of 40 % (233.3 mg/mL) and incubated at 6 °C and 375 rpm for 15 min. To remove the insoluble fraction, the lysate was centrifuged (30 min, 75,000 xg, 10 °C) and the pellet discarded. The *Ca*AAD protein was precipitated by addition of $(NH_4)_2SO_4$ to a f.sat.c. of 50 % (60.2 mg/mL). The insoluble fraction containing *Ca*AAD was removed from the suspension by centrifugation (30 min, 75,000 xg, 10 °C) and re-dissolved in lysis

buffer (60 min, 6 °C) under mild shaking.

To prepare the protein solution for the TMAE-IEX, residual ammonium sulfate was removed with dialysis against lysis buffer. The protein containing fractions were further diluted with lysis buffer to a f.v. of 150 mL, due to the low affinity of *Ca*AAD for the TMAE matrix. The protein was eluted from the column with approx. 5 % of TMAE-IEX-elution buffer (50 mM KH_2PO_4, 300 mM $(NH_4)_2SO_4$, pH 5.9. The almost pure elution fractions were dialysed against crystallisation and storage buffer (5 mM KH_2PO_4, 2 mM DTT, pH 5.95).

5.34 Determination of *Ca*AAD protein concentration

The concentration of purified protein samples was determined spectrophotometrically with a molecular weight of 27.255 kDa and a molar extinction coefficient of 26360 $M^{-1}cm^{-1}$ (see: 5.20).

5.35 Synthesis of AcAc from MeAcAc

In order to determine the reaction kinetic constants of *Ca*AAD, the substrate AcAc was chemically synthesised. The precursor substance MeAcAc was purchased from Sigma-Aldrich (Munich, Germany) and cleaved into AcAc and methanol using a basic saponification protocol. 5 g of MeAcAc (43.1 mmol) were weighted into a 100 mL round bottom flask. 1.9 g NaOH (47.5 mmol) were dissolved in 38.5 mL ddH_2O and slowly titrated into the round bottom flask (RT, 20 min). The colour slowly changed to light yellow. The reaction mixture was stirred for 18 h at RT. The hydrolysis reaction was quenched by addition of 35 g $(NH_4)_2SO_4$.

To isolate AcAc from the reaction mixture, acetoacetate has to be protonated to obtain the corresponding acid. 37 % HCl was slowly titrated into the flask until pH 2 was reached (pK_a(AcAc): 3.58). To prevent spontaneous decarboxylation of AcAc, the acidification process was done under constant stirring and ice cooling. 30 mL of the quenched and acidified reaction mixture were applied to a separating funnel and AcAc isolated from the aqueous phase with three subsequent extractions using 30 mL of diethyl ether. The diethyl ether fractions were collected and pooled.

To remove residual water, water-free $MgSO_4$ was added to the AcAc solution until no clumps appeared any more and the added magnesium sulfate remained snowy. The $MgSO_4$ was removed by filtration. The organic solvent was removed in a rotary evaporator under reduced pressure. The diethyl ether phase was applied to a 500 mL round bottom flask in a water bath and the solvent evaporated until approx. 5 mL remained (20 °C, 700 mbar). The residual volume was transferred into a 50 mL round bottom flask and the solvent completely evaporated (20 °C, 2 h, 700 mbar).

The viscosity of pure diethyl ether free AcAc is comparable to sun flower oil and appears as a pale yellow oil, which is a mixture of the keto-enol tautomers. The volume of the isolated AcAc from saponification of 5 g MeAcAc is approx. 1.5 ml.

5.36 Determination of AcAc concentration

To determine the concentration of AcAc for enzyme kinetic measurements and cryo-crystallography, spectrophotometric absorbance measurements were performed. In order to account for solvent polarity effects and pH dependencies of the molar extinction coefficient of AcAc, the pure stock solution (approx. 10 M) was diluted 1:10 in lysis buffer (50 mM KH_2PO_4, pH 5.9) and the pH adjusted to 5.9 using 10 M KOH. The absorbance was measured at a wavelength of 210 nm to avoid potential interference with acetone from spontaneous decarboxylation. The concentration was determined according to Lambert-Beer's-law:[17–19]

$$A_\lambda = \epsilon_\lambda \cdot c \cdot d \tag{5}$$

The concentration of acetoacetate (c(AcAc)) was determined from the measured absorbance at 210 nm A_{210}, the molar extinction coefficient at 210 nm (ϵ_{210}(AcAc)=420 $M^{-1}cm^{-1}$) and the optical path length (d).[49] 50 μL aliquots of the pH adjusted solutions were prepared, frozen in $N_{2(l)}$ and stored at -20 °C.

5.37 Determination of enzyme kinetic constants of *Ca*AAD

The concentration of the used AcAc stock-solution was determined directly before use, in order to account for the spontaneous self-decarboxylation of the β-keto acid. Thawed substrate stock solutions and prepared dilutions were constantly kept on ice and discarded at the end of the day.

The activity assays were performed in lysis buffer (50 mM KH_2PO_4, pH 5.9) at 25 °C utilizing a spectrophotometer (Jasco V-650, V-750). The assay relies on the absorbance depletion after AcAc decarboxylation to acetone and CO_2 at 270 nm.[49]

$$\Delta\epsilon_{270} = \epsilon_{270}(\text{AcAc}) - \epsilon_{270}(\text{Ac}) \tag{6}$$

ϵ_{270}(AcAc): molar extinction coefficient at 270 nm: 55 $M^{-1}cm^{-1}$

ϵ_{270}(Ac): molar extinction coefficient at 270 nm: 28.3 $M^{-1}cm^{-1}$

$$k_{\text{obs}} = \frac{\frac{\Delta \text{abs}_{270}}{60}}{\frac{\Delta \epsilon_{270} \cdot \text{d}}{\text{c(AAD)}}} \tag{7}$$

5.38 Crystallisation of *Ca*AAD$_{WT}$ and variants

In order to obtain information on *Ca*AAD 3-dimensional structure, X-ray cryo-crystallography was applied. Crystal formation was induced in VDX$_m$ hanging-drop crystallisation plates (Hampton Research, (Aliso Viejo, USA)) with a reservoir volume of 400 µL according to a modified protocol from Ho *et al.*, 2009.[47] Crystallisation was performed at 20 °C in a temperature controlled environment.

In order to obtain non-liganded resting state crystals of *Ca*AAD$_{WT}$ and variants, purified protein was concentrated to approx. 20 $^{mg}/_{mL}$ applying ultra-filtration (Spin-X® UF, 100k MWCO PES, Corning). The protein was centrifuged to remove potential aggregates (30 min, 21282 xg, 4 °C), diluted to 10 $^{mg}/_{mL}$ in crystallization and storage buffer (5 mM KH_2PO_4, 2 mM DTT, pH 5.95) and mixed with reservoir solution (40 mM KH_2PO_4 pH 5.95, 100 mM Sarcosine, 14 – 16 %($^{w}/_{v}$) PEG 3350, 20 %($^{v}/_{v}$) glycerol) on a 18 mm siliconized cover slide (Jena Bioscience). On a single cover slide, the total drop volumes were 3 µL and 4 µL with 2 µL of reservoir solution. Crystals appeared after 2 days and grew for approx. 10 days to the final size.

To obtain *Ca*AAD/AcP and *Ca*AAD/AcS inhibitor complex structures, 10 $^{mg}/_{mL}$ protein solution was supplemented with 50 mM AcP and AcS (100 mM stock in 5 mM KH_2PO_4, 2 mM DTT, pH 5.95), respectively. The protein-inhibitor mixture was incubated at RT for 15 min and used for crystallisation. The conditions and the procedures were identical to the resting state crystallisation.

5.39 Post-crystallisation treatments and cryo-protection of *Ca*AAD

In order to prepare AAD$_{WT}$ crystals for cryo-protection, the at room temperature grown crystals transferred to 6 °C. The crystallisation plates were placed in an approx 30×30×30 cm styrofoam box and placed in the cold room. To reduce the cooling rate, the boxes were supplied with temperature conditioned cooling elements (20 °C). The crystallization plates were equilibrated at 6 °C for 24 h.

A crystal containing drop was supplemented with 2 µL stabilization solution (40 mM KH_2PO_4 pH 5.95, 100 mM Sarcosine, 20 %($^{w}/_{v}$) PEG 3350, 20 %($^{v}/_{v}$) glycerol). The stabilised crystal was

transferred into cryo-protection solution (40 mM KH_2PO_4 pH 5.95, 100 mM Sarcosine, 20 %(w/v) PEG 3350, 20 %(v/v) glycerol, 10 %(v/v) EG, 10 %(v/v) PG, 10 %(v/v) PEG400) in a 4 step gradual transfer procedure of mixtures of stabilisation and cryo-protection solution. From the cryo-protection solution, the crystals were fished with a crystal loop (LithoLoop, Molecular Dimensions) and plunge-frozen in $N_{2(l)}$.

The inhibitor complex crystals were treated according to the resting state. For AAD_{WT}/AcP and AAD_{WT}/AcS crystals, the cryo-protection solution contained 50 mM AcP or AcS, respectively.

5.40 Synchrotron X-ray diffraction data collection and processing *Ca*AAD

X-ray diffraction raw images of single crystals were collected under cryogenic temperatures (100 K). The synchrotron diffraction data sets were collected at P14 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany).[28] Diffraction data analysis, integration, correction, scaling, post-refinement and space-group assignment was done with the XDS program-suite.[29,30] The automated space-group determination was evaluated using "pointless" and "phenix.xtriage".[31,32] The resolution cut-off decision was mainly made on the $CC_{(1/2)}$ (> 50 %), $I/\sigma(I)$ (> 1) and the completeness of data (> 95 %).[33–35] Data was converted to the MTZ reflection format using "XDSCONV" and "f2mtz" and R-free reflections were introduced using "XDSCONV" or "uniqueify".[29,36]

5.41 Phasing and initial structure determination of *Ca*AAD

As starting models for *Ca*AAD and variant protein crystal data phasing, the PDB-entry 3BH2 was used. The crystallographic space-group was assigned as $P2_12_12_1$ with two dodecamers in the asymmetric unit. The structure was solved with molecular replacement (Gleb Bourenkov, Ashwin Chari) using the program "molrep" and the resulting model used for all following data sets.[40] In most cases, the unit cell constants deviated ≤ 1 % from the input model parameters and the initial phasing was done with a rigid-body refinement.[38,39]

5.42 Model building and structure refinement of *Ca*AAD

The obtained 3D crystal structures were corrected and adjusted manually against the diffraction data using the "COOT" software suite.[41] Potential ligands were drawn and the CIF-files generated using "Jligand".[42] The covalent intermediate of the catalytic lysine with acetonyl phosphonic acid

and acetonyl sulfonic acid was modelled as independent peptide.

Model adjustment and data refinement was done in an iterative manner using "COOT" and "refmac5".[38,39,41] All datasets were refined applying a local non-crystallographic symmetry. Datasets with a resolution higher than 1.6 Å were refined with anisotropic atomic displacement parameters (ADP).

Protein structure representation and image generation was done with "PyMol" and "Chimera".[45,46] Dataset and refinement statistics were generated with phenix.table_one.[32]

6 Introduction: Orotidine 5'-monophosphate decarboxylase

A remarkable property of nature is the production of a variety of catalysts to facilitate the occurrence of chemical reactions at ambient temperatures with reaction rates we consider as "fast". Even though not the entirety of cellular metabolites requires the action of certain catalysts, the majority of chemical reactions in living cells are catalysed to enable reaction progression in reasonable time-spans.[50]

Decarboxylation reactions are important processes in organic chemistry and biochemical pathways. The removal of CO_2 from chemical compounds plays a major role in synthetic and degradative pathways and was studied from the beginning of the organic chemistry era to elucidate the fundamentals of reaction kinetics. In solution, decarboxylations proceed with unimolecular and bimolecular reaction mechanisms leading to the decomposition of acidic molecules. In both, the S_E1 and S_E2 mechanism, the decarboxylation is generally considered as irreversible electrophilic substitution reaction with a proton substituting the leaving carboxylate group. The heterolytic C-C bond fission of the carboxylate group and the subsequent dissociation of carbon dioxide generates a localised negative charge which is eliminated simultaneously or subsequently by a protonation.[51] However, many carboxylic acids are highly stable compounds with spontaneous decarboxylation rates which reach the limits of detection at room temperature. In striking contrast to the slow reaction rates in solution, the same reactions can occur within fractions of a second in the presence of a catalyst.[52]

Due to the slow spontaneous reaction rates in solution, a great amount of different catalysts evolved applying a variety of decarboxylation mechanisms ideally suited for specific substrate molecules.[53] Certain enzymes facilitating the decarboxylation of 2-keto acids such as pyruvate decarboxylase (PDC) utilize the cofactor thiamin diphosphate (ThDP) which is composed of a thiazolium ring, a 4-aminopyrimidine ring and a diphosphate anchor. The cofactor is bound to the enzyme by ionic interactions between a Ca^{2+}- or Mg^{2+}-ion and the intrinsic phosphoryl moiety.[54] The enzyme can activate the thiazolium ring by deprotonation to form a nucleophilic center which can subsequently attack the carbonyl carbon atom of the substrate.[55] The emerging negative charge after bond rupture can be resonance stabilized within the conjugated system of the cofactor to prevent the existence of a localised carbanionic species.[56] Furthermore, the enzyme can selectively stabilize certain intermediate species in order to favour a specific reaction path.[57]

Alternatively, 2-keto acid decarboxylation can be catalysed by a lysine residue of the enzyme which acts as a reactive nucleophile to attack the carbonyl carbon. Catalysis by the formation

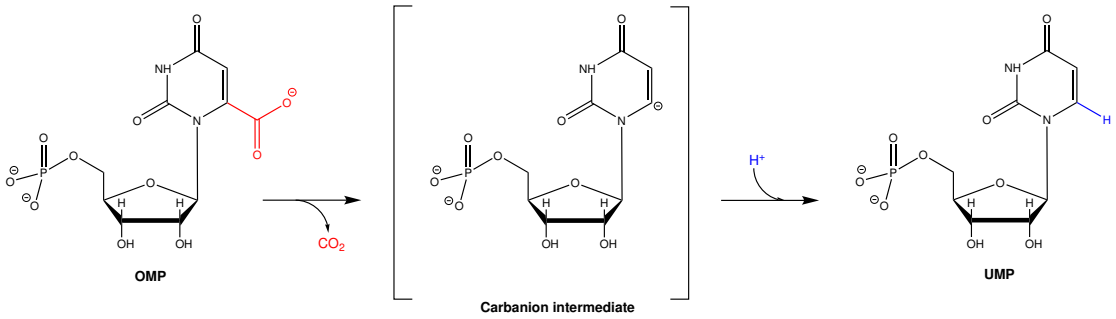


Figure 1: **Orotidine 5'-monophosphate decarboxylase mediated decarboxylation of orotidine 5'-monophosphate.** The structure in the brackets indicates a potential decarboxylation intermediate.

of a Schiff-base intermediate is observed for acetoacetate decarboxylase (AAD) mediated decarboxylation to yield CO_2 and acetone.[58] In order to facilitate decarboxylations of and other reactions on α-amino acids, the cofactor pyridoxal phosphate can be utilized.[59] Furthermore, additional cofactors, prosthetic groups and bivalent metal ions are recruited by evolution to facilitate challenging decarboxylation reactions.[60,61]

Among the heterogeneous group of decarboxylases, orotidine 5'-monophosphate decarboxylase (OMPD, EC 4.1.1.23) demands an extraordinary position. The enzyme facilitates the decarboxylation of orotidine 5'-monophosphate to yield uridine 5'-monophosphate and carbon dioxide without the help of any cofactors or metal ions (Figure 1).[62,63] Besides the notion of acting as a sole protein catalyst, it attracted the great attention of the scientific community due to the extraordinary reaction rate enhancement when comparing the catalysed and the uncatalysed reaction. The small chemical model compound 1,3-dimethylorotic acid has a half-life time ($t_{1/2}$) of 78 million years in neutral aqueous solution.[52] In the presence of the enzyme OMPD of *Saccharomyces cerevisiae*, the same decarboxylation reaction on the natural substrate OMP occurs with multiple turn-overs per second.[64,65] Thus, the yeast enzyme is capable of increasing the chemical reaction rate by a factor of 10^{17} and to bind the transition state according to the transition state stabilization theory with a dissociation constant of $5 \times 10^{-24}\,\mathrm{M}^{-1}$. The transition state is stabilized by $31\,{}^{\mathrm{kcal}}/_{\mathrm{mol}}$. In 1995, it was considered as most proficient enzyme known and applied to establish the term of catalytic proficiency ($\frac{k_{\mathrm{cat}}}{K_{\mathrm{M}}}/k_{\mathrm{uncat}}$). When the extraordinary catalytic prowess of OMPD was discovered, no information on the structural organisation of OMPD was present and a great amount of effort was undertaken to elucidate the mechanistic principles of the remarkable enzyme.

Besides the fascinating kinetic capabilities of OMPD, from 1950 and nowadays still, it is considered as promising target for cancer treatment.[66,67] Since vastly growing cancer cells require an excessive amount of nucleotides to maintain the cell growth, pharmaceuticals acting on nu-

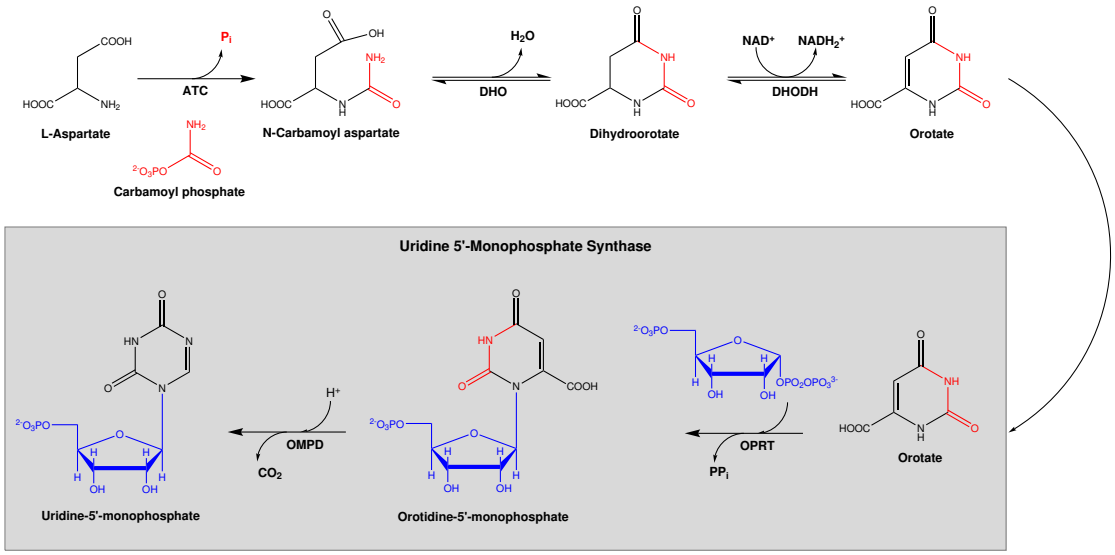


Figure 2: **Metabolic pathway of the *de novo* pyrimidine synthesis.** ATC: Aspartate carbamoyltransferase (EC 2.1.3.2), DHO: Dihydroorotase (EC 3.5.2.3), DHODH: Dihydroorotate dehydrogenase (EC 1.3.5.2), OPRT: Orotate phosphoribosyltransferase (EC 2.4.2.10), OMPD: Orotidine 5'-monophosphate decarboxylase (EC 4.1.1.23). The grey box indicates a bifunctional enzyme (UMPS: Uridine monophosphate synthase) facilitating both reactions in eucaryotes.

cleotide synthesis were intended to harness cancer cell division. OMPD catalyses the last step in the *de novo* pyrimidine biosynthesis pathway and is an essential component to supply nucleotides for DNA synthesis and maintenance (Figure 2). The enzyme carbamoyl phosphate synthetase catalyses the synthesis of carbamoyl phosphate from the precursor molecules bicarbonate, ammonia and ATP. The reaction product is transferred onto an aspartate residue by asparate carbamoyltransferase and dehydrated by dihydroorotase to form dihydroorotate. Oxidation of dihydroorotate by dihydroorotate dehydrogenase yields orotate which is linked to a furanosyl phosphate moiety by orotate phosphoribosyltransferase to obtain the assembled nucleotide. The decarboxylation of orotidine 5'-monophosphate is facilitated by OMPD.

In bacteria, OMPD is expressed as a single domain which forms a dimeric protein complex after translation.[68] In contrast to procaryotic organisms, in eucaryotes including *Plasmodium*, OMPD is part of the bifunctional uridine monophosphate synthetase (UMPS) machinery.[69] UMPS is composed of a functional orotate phosphoribosyltransferase (OPRT, EC 2.4.2.10) domain on the N-terminal end of the polypeptide sequence and the orotidine 5'-monophosphate decarboxylase subunit on the C-terminus. Consequently, the complex catalyses the two last reaction steps to produce uridine 5'-monophosphate.

The impressive catalytic characteristics of OMPD led to the formulation of a variety of different reaction mechanisms trying to rationalize the enormous reaction rate acceleration. The first proposals were based on chemical model reactions in aqueous solution. An emerging car-

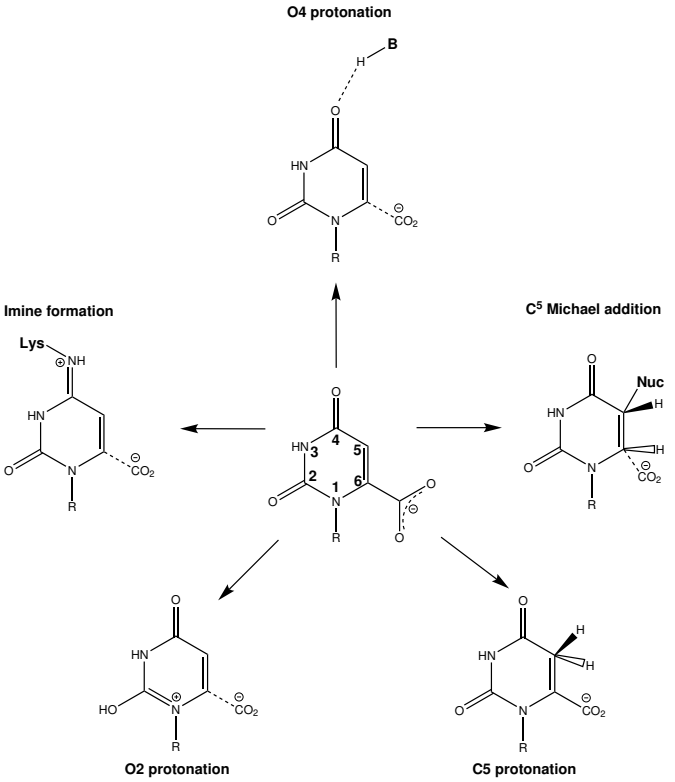


Figure 3: **Proposed reaction mechanisms for OMPD mediated decarboxylation of OMP.** The remarkable prowess of OMPD led to the formulation of reaction mechanism including O^4-protonation[74], a covalent intermediate on the C^5-atom[75], the C^6-atom protonation[76,77], an O^2-carbonyl protonation[70,71] and imine formation[78]. Most of the postulated reaction mechanisms can apply reactant-state destabilization and intermediate stabilization or both. R indicates the phosphoribosyl moiety. Figure is taken and adjusted from Stanton *et al.*, 2007.[79]

banion from the decarboxylation of picolinic acid could be stabilized by a protonation of the ring nitrogen to form a zwitterionic reaction intermediate.[51] Beak and Siegel applied the idea of ylide formation to OMPD mediated catalysis and stated a potential protonation of the C^2-carbonyl group leading to a positive charge on the N^1-atom of the pyrimidine ring.[70,71] They observed a 10^8 fold rate increase of spontaneous decarboxylation based on the methylation of the C^2-carbonyl group of the model compound 1-methylorotate. The methylation would lead to a positively charged nitrogen and could mimic a transient protonation of the carbonyl oxygen. The generated positive charge might stabilize the carbanion generated from heterolytic bond fission. The active site of OMPD could facilitate the C^2-carbonyl protonation by appropriately orienting a catalytic acid. Additionally, the active site would have to prevent the deprotonation of N^3 which would allow ylide formation. However, the chemical model compound citrazinic acid which lacks the N^1-atom of orotate undergoes decarboxylation with a comparable rate compared to 1,3-dimethylorotic acid.[72] Since zwitterion formation is not possible in citrazinic acid, the effect of ylide stabilization of the intermediate appears to be a minor contributor to decarboxylation in solution. Furthermore, ^{13}C kinetic isotope effects of 1,3-dimethylorotate and ^{15}N of picolinic acid and N-methyl picolinic acid argue against ylide formation.[73]

The very high binding affinity of the competitive inhibitors 1-(5'-phospho-β-d-ribofuranosyl)

barbituric acid (6-hydroxy-UMP, BMP) and 6-azauridine (Aza-UMP) towards OMPD indicated a binding mode as transition state analogues.[67,80,81] Since both ligands are considered to be able to exhibit a negative charge, they were proposed to resemble the carbanionic reaction intermediate of direct decarboxylation. The enzyme might place a positive charge in close proximity to the nascent carbanion and thus stabilize the transition state.[82] However, OMPD from *Methanobacterium thermoautotrophicum* and human also catalyse the conversion of 6-cyano-UMP which is likely to form a cationic reaction intermediate after CN^--release.[83] However, the results are difficult to interpret since covalent mechanisms also could lead to the product formation with a lysine residue as active nucleophile.[84] Intriguingly, the same reaction yields BMP in the bacterial enzyme and UMP in the human system.[83,84]

Based on the establishment of protein purification protocols for the yeast enzyme, the first kinetic and mechanistic investigations initiated systematic OMPD research.[68] The first proposals stated a potential reaction mechanism applying a catalytic zinc ion.[64] Based on the 10^5 fold higher affinity of OMPD towards 6-thiocarboxamido-UMP compared to 6-amido-UMP, a Zn^{2+}-ion in the active site was hypothesized. However, the proposal was rejected since absorption spectrocopy studies and activity measurements with zinc chelating reagents gave no indication of the presence of a zinc ion in the active site.[62,63,78]

An additional reaction mechanism claimed the protonation of the C^4-carbonyl group based on quantum chemical calculations.[74] Again, the protonation would generate a carbanionic intermediate with a positively charged OH^+ on position C^4. The zwitterionic intermediate would stabilize a carbanionic species from simultaneous decarboxylation in a hydrophobic active site environment. The formation of the mesomeric carbene structure could further stabilize the emerging negative charge from CO_2-release. According to the calculations, the active site architecture must provide two properties. A catalytic acid needs to be positioned in a place suitable for protonation in an environment of a low dielectric constant. However, different authors pointed out that the experimentally observed rate acceleration could not be reached with an exclusive protonation on O^2 or O^4.[85,86]

Another proposed decarboxylation mechanism reported a potential covalent linkage between an active site lysine and the C^5-position of the pyrimidine ring.[75] The protonation of the substrate's carboxylate group leads to a Michael addition to the C^5-atom of the pyrimidine ring followed by a decarboxylative elimination reaction. The general potential of an active site lysine to act as a reactive nucleophile was demonstrated, however, the rate limiting step in OMPD catalysis is likely to represent a proton transfer reaction which follows C-C bond cleavage.[84,87,88]

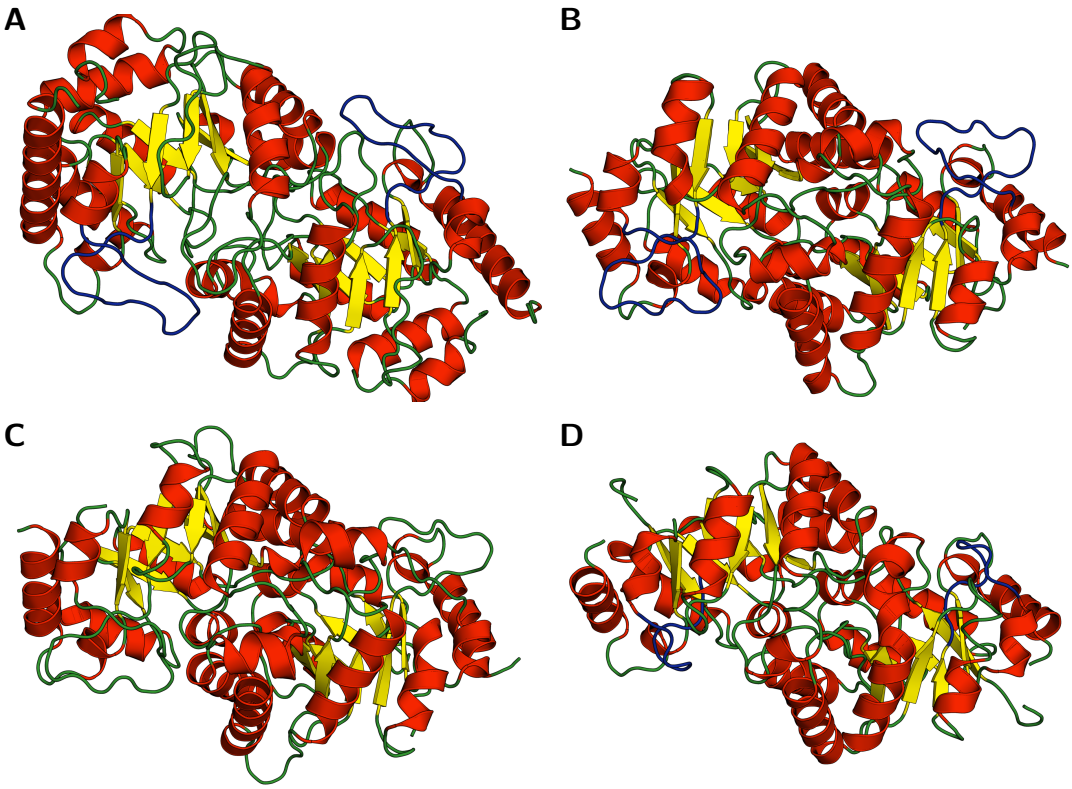


Figure 4: **Quaternary structure comparison of OMPD dimers from eucaryotes, procaryotes and *archea*.** Representative structures of *Saccharomyces cerevisiae*, *Escherichia coli*, *Bacillus subtilis* and *Archaeoglobus fulgidus* are depicted in the closed conformation (PDB entries: 1DQX, 1EIX, 1DBT, 4MUZ) with BMP in the active site.[76,90,91] α-helices, β-sheets and loop regions are indicated in red, yellow and green, respectively. The phosphate gripper loop is coloured in blue.

Aside from the potential nucleophilic attack of a lysine onto the C^5-position, a proton transfer to generate a sp^3-carbon was published.[89] Furthermore, a proton transfer to the C^6 simultaneously to decarboxylation was introduced.[76] The scientific community currently favours the generation of a stabilized vinyl carbanion as reaction intermediate of decarboxylation. However, no consensus on the exact reaction mechanism was obtained yet.

Since OMPD catalyses a fundamental reaction in nucleotide production, its activity is essential to all living organisms. Notably, certain catalytic residues are conserved and the overall active site architecture of OMPD is very similar across the three domains of life (Figure 4).[92] The first crystallographic analysis of OMPD was reported in 1991 describing crystallisation conditions. In 2000, the bacterial crystal structures of *B. subtilis* (PDB entry: 1DBT), *E. coli* (PDB entry: 1EIX) and *M. thermoautotrophicum* (PDB entry: 1DV7) and the eucaryotic OMPD-domain structures of *S. cerevisiae* (PDB entry: 1DQW) were published.[76,91,93–95] In the following years, a plethora of enzyme structures were released including a vast amount of amino acid substitution variants. Additionally, the human OMPD-domain (PDB entries: 2V30, 2QCE) and different

archeal (PDB entries: 4DF1, 4MUZ, 3VE9, 2CZF; no corresponding publication) protein structures were published.[5,84]

The protein monomer displays a TIM-barrel topology with the 8 characteristic parallel β-sheets forming the inner barrel structure.[96] The β-sheets are seperated by a variable number of α-helices and loop regions which cover the outside of the barrel structure to form a toroid. The alternating α/β motives form a robust particle but also allow catalysis via flexible loop regions of the fold. The obligate dimer of OMPD forms the active site at the interface of two monomers which are stacked with the C-terminal end of the β-barrels facing each other. Notably, ligand binding induces a movement of secondary structure elements and a structural rearrangement of the opened conformation to form the closed ligand complex.[97,98] The most prominent rearrangement upon ligand binding relates to the rigidification of a flexible loop which is disordered in the resting state enzyme (Figure 5).

The loop region, referred to as "phosphate gripper loop", folds over the phosphoribosyl moiety of the nucleotide in the active site pocket. The phosphoryl group of the ligand is coordinated by an arginine residue to the rigid part of the enzyme and a well conserved glutamine on the phosphate gripper loop. The glutamine interconnects both ends of the ligand by forming hydrogen bonds from the N^{ϵ} hydrogen atoms to the phosphate moiety and the C^2-carbonyl oxygen of the pyrimidine part. Furthermore, a tyrosine or an arginine in the loop region can be present and establish contacts to the phosphoryl group. The loop rearrangement covers the binding site and sequesters the ligand from the bulk solvent (Figure 6).

Additionally, a smaller loop segment moves towards the ligand's pyrimidine part and is making contact to the phosphate gripper loop in the closed conformation. The loop, referred to as "pyrimidine umbrella", tightens the active site pocket and connects both loops in the liganded complex. A characteristic serine (alternative threonine) residue forms a hydrogen bond to the carbonyl oxygen of the glutamine of the phosphate gripper loop as hydrogen bond acceptor. An additional hydrogen bond to the N^3-atom of the pyrimidine ring of the ligand is likely to be formed as hydrogen bond donor. Thus, both residues, the glutamine from the phosphate gripper loop and the serine of the pyrimidine umbrella stabilize the closed enzyme conformation.[100,101] Loop closure is generally considered to enclose the substrate in a water occluded protein cage with a decreased effective dielectric constant in order to increase the strength of electrostatic interactions with the reaction transition state and to specifically modify pK_a-values of certain residues.[102,103]

Notably, the loop size and the loop hinge flexibility is tightly adjusted to the environmental

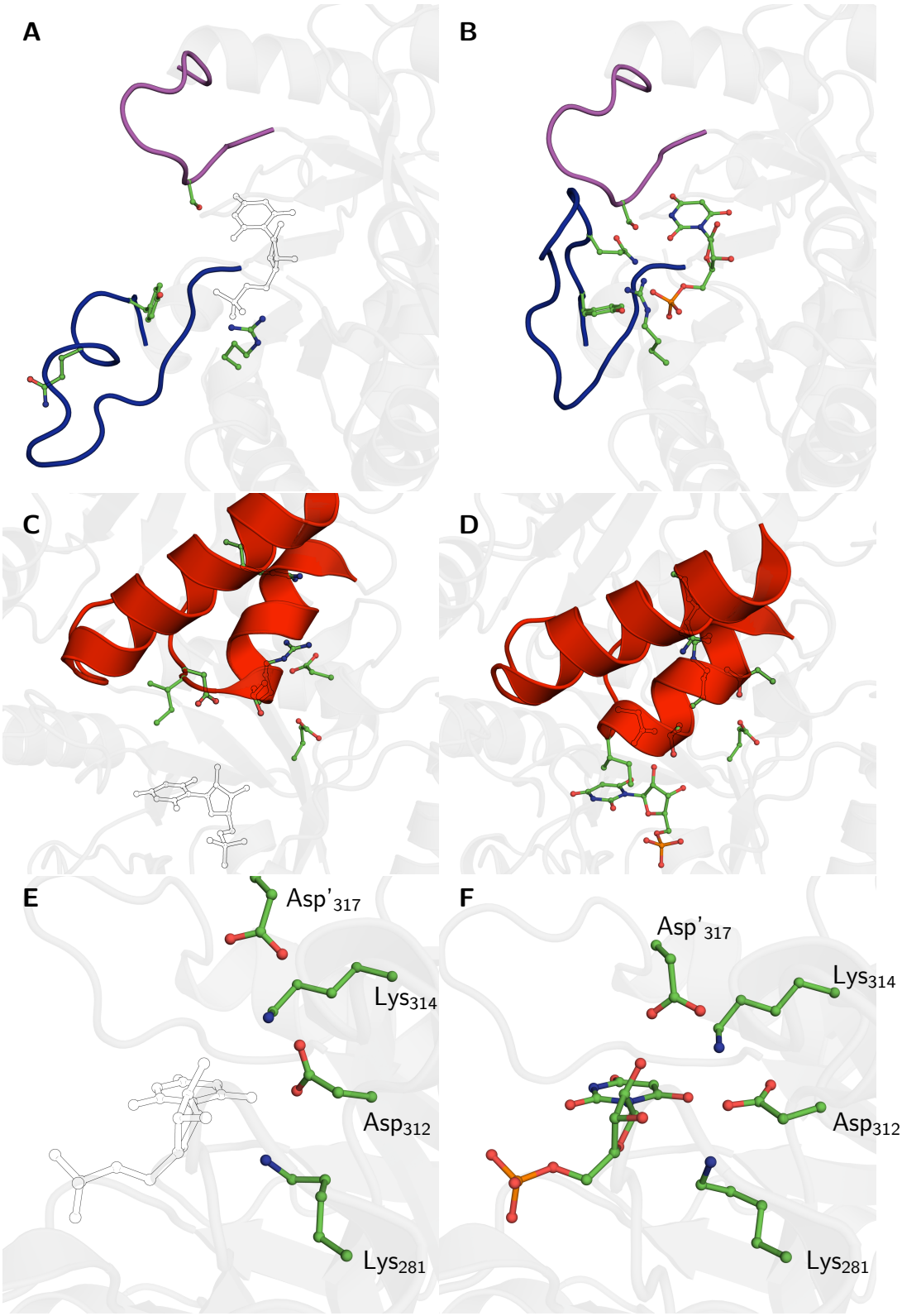


Figure 5: **Ligand binding induced conformational change to form the closed complex.** Partially covered and hidden residues are indicated with outlines. In the open conformation, the putative BMP inhibitor position is indicated with outlines for comparison reasons. Coordinates of the open and closed enzyme conformations are taken from PDB entries 3GDK and 1DQX of the yeast enzyme, respectively.[90,94] Residue numbering according to the human protein.

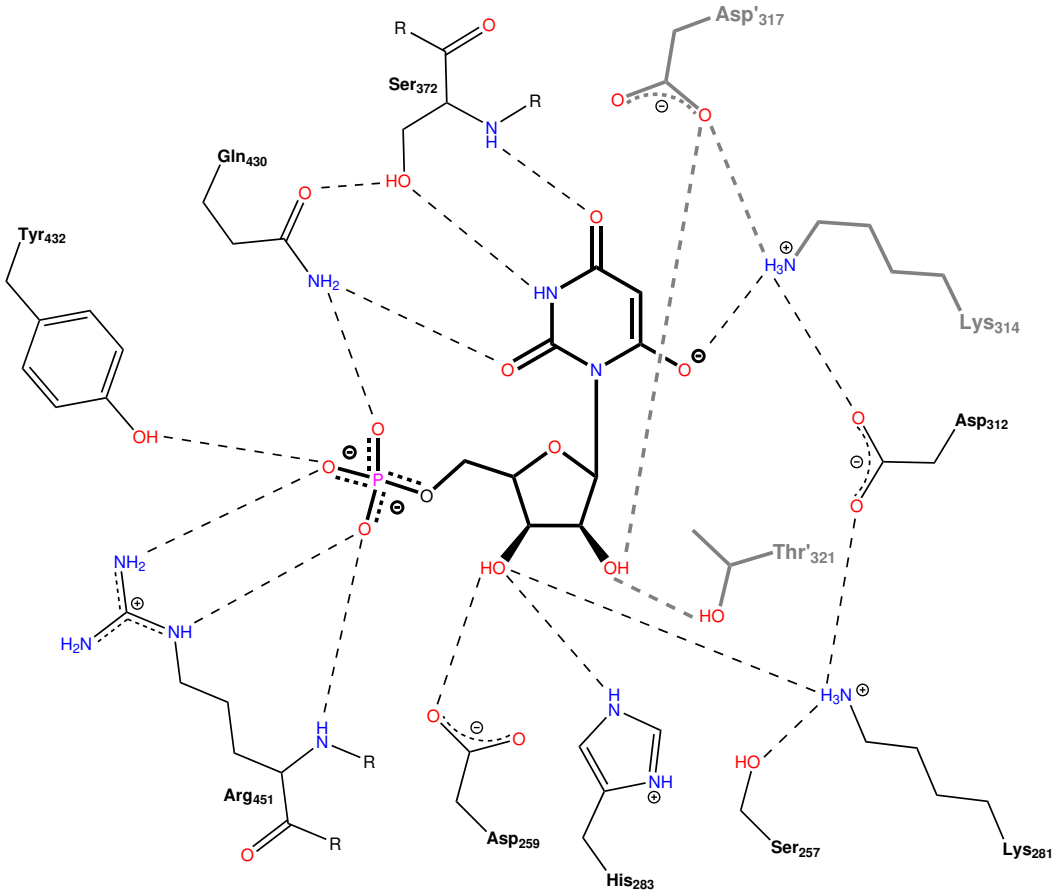


Figure 6: **Hydrogen bond network surrounding the ligand BMP in the active site of OMPD.** Main chain atoms are omitted except for Ser_{372} and Arg_{451}. Residues from the neighbouring subunit are labelled as [residue]'. Potential hydrogen bonds are indicated with dashed lines. Residues in grey are located above the pyrimidine plane. Residue numbering according to the human OMPD-domain. Modified and extended PoseView image of PDB-entry 3EX4.[84,99]

conditions. The phosphate gripper loop of *S. cerevisiae* with a length of 19 residues is 10 amino acids longer than the corresponding region in *M. thermoautotrophicum*. The intrinsic phosphodianion binding energy utilized to activate the enzyme is comparable in both organisms but slightly lower in the thermophilic organism. However, the entropic cost for loop closure is higher for the larger loop segment of *Sc*OMPD compared to *Mt*OMPD with values of -16.0 $^{cal}/_{K \cdot M}$ and -6.0 $^{cal}/_{K \cdot M}$, respectively. The unfavourable entropy of a larger protein loop is compensated by a higher enthalpic contribution to loop closure.[104]

The active site of the enzyme displays four prominent amino acid residues which are absolutely conserved throughout all domains of life (Figure 7).[106] The alternating charge network, referred to as "catalytic tetrad", is composed of two aspartate and two lysine residues (Lys-Asp-Lys-Asp). The latter residues of the catalytic tetrad are organized in a DXKXXD amino acid motive. The main residue of the amino acid array is a lysine in the middle of two asparates potentially

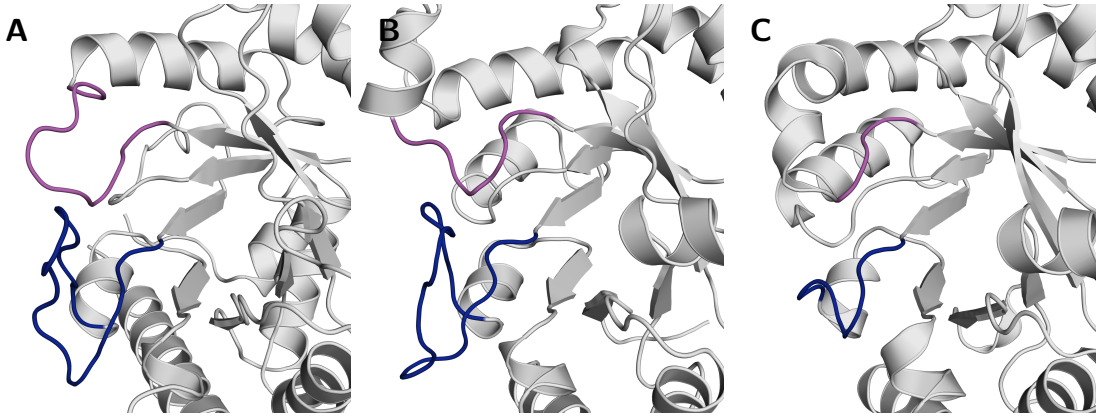


Figure 7: **Comparison of the length of the flexible loops of OMPD from mesophilic and thermophilic organisms.** View onto the active site of subunit A of the protein dimer in the closed conformation. For the sake of simplicity, residues of subunit B and bound ligands are omitted. The phosphate gripper loops and pyrimidine umbrella regions are indicated in blue and magenta, respectively. Structures of *Saccharomyces cerevisiae*, *Eschericha coli* and *Methanobacterium thermoautotrophicum* were taken from the protein data bank (PDB entries 3GDL, 1EIX, 3LTP).[5,91,105]

forming hydrogen bonds. The so called catalytic lysine residue is generally believed to protonate the carbanionic species generated from the decarboxylation reaction.[76] The analysis of the $^{k_{cat}}/_{K_M}$ curve of pH-depended decarboxylation indicated two ionisable groups with pK_a-values of 6.1 and 7.7. The higher pK_a-value was assigned to the active site lysine whereas the lower value was considered to correspond to the protonation of the phosphoryl moiety of the substrate.[87,107]

All tetrad residues are essential for efficient catalysis and an alanine substitution of any of these leads to nearly complete loss of catalytic activity (residues Lys_{95}, Asp_{91}, Lys_{93} and Asp_{96} in the yeast enzyme corresponding to K281, D312, K314 and D317 in the human protein. An exhaustive list of the active site residue numbering of different organisms can be found in the appendix (Table 12). The following sections will use the human numbering scheme if not explicitly indicated otherwise). The turn-over number for the decarboxylation of OMP of the yeast enzyme was determined to be $44\,s^{-1}$ (approx. 10-40 fold higher than the human enzyme).[5,63,107] Notably, the protein variants K281A, D312A, K314A and D317A display a relative catalytic rate reduction of 2×10^{-3}, 7.3×10^{-6}, 2×10^{-6} and 5.2×10^{-6}, respectively.[106] The special role of Lys_{314} was demonstrated in a study utilising a cysteine substitution variant (K314C). The catalytic activity of the mutant enzyme was reduced by a factor of approx. 2×10^{-8} but could partially be restored by a reaction with 2-bromoethylamine to form S-(2-aminoethyl)cysteine. The S-(2-aminoethyl)cysteine variant displayed an activity recovery rate of 5×10^{-5}.[82]

In the crystal structures, the catalytic tetrad residues are found to surround the pyrimidine moiety of the substrate on the opposite face of the phosphate gripper loop. The aspartate residue situated directly above the pyrimide plane is provided by the other subunit of the protein dimer

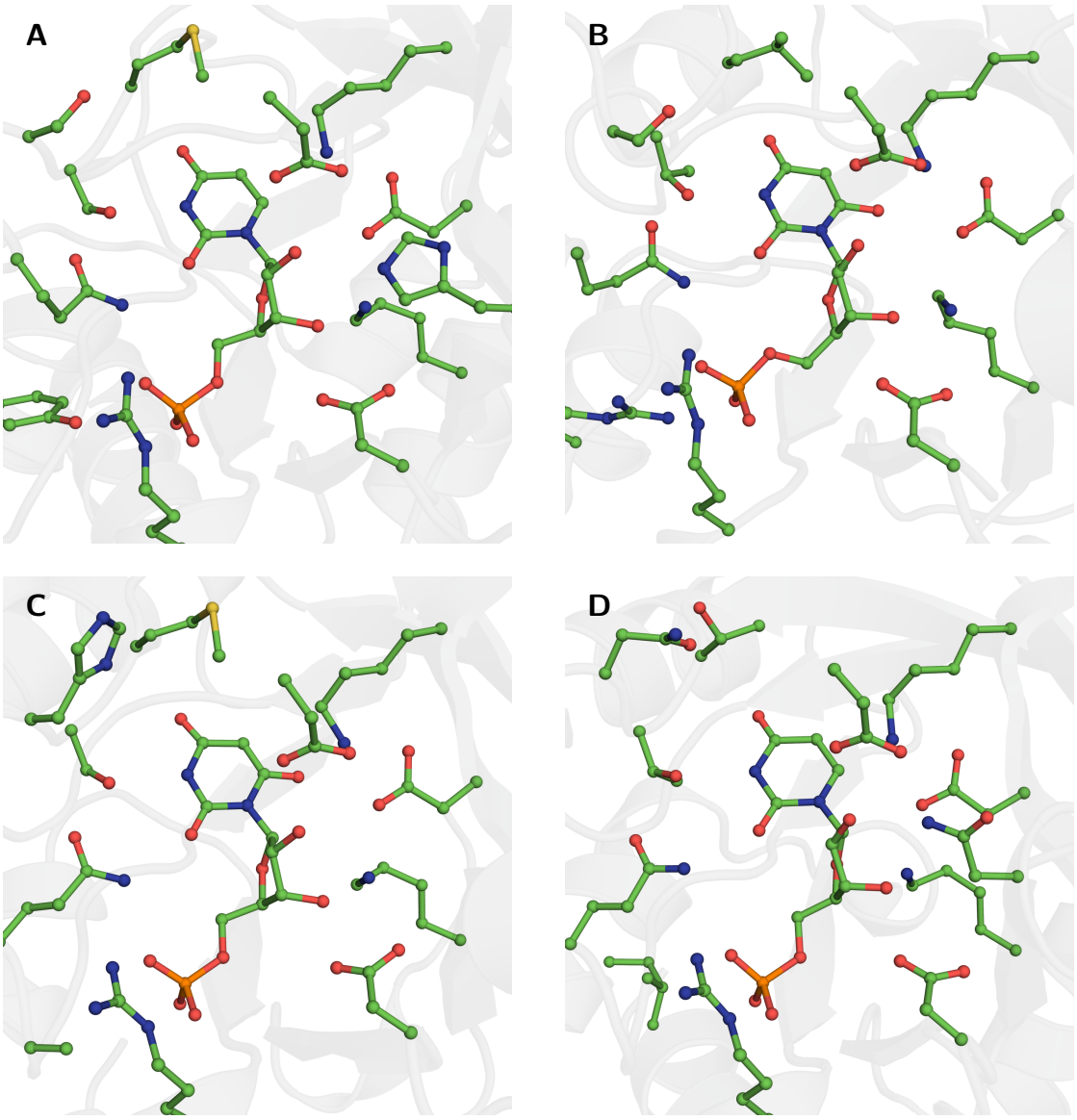


Figure 8: **OMPD active site conservation between different species.** Protein structures of A: *Homo sapiens*, B: *Escherichia coli*, C: *Methanobacterium thermoautotrophicum* and D: *Plasmodium falciparum* were taken from the protein data bank with entries 2QCD, 1EIX, 3LTP and 6DSR, respectively.[5,91,92,105,108]

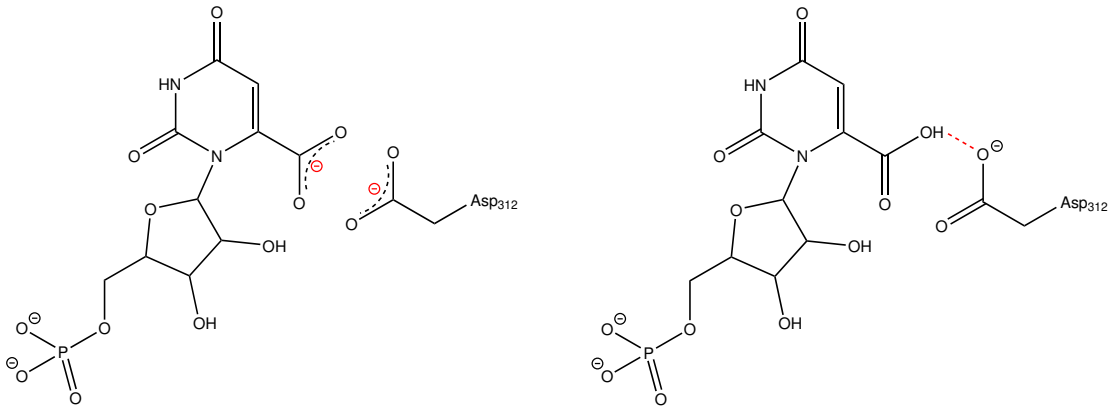


Figure 9: **Ground state destabilization (GSD) vs. transition state stabilization (TSS) in the active site of OMPD**

(residues from the neighbouring monomer are indicated as [residue]' and should not be confused with atoms of the ligands ribosyl moiety). Except for Asp'$_{317}$ which is placed on a helix forming after ligand binding, they are part of the rigid core of the enzyme on the C-terminal side of the β-barrel. The spatial proximity of the alternating aspartate and lysine residues appear to favour the ionized state of these residues. However, the pK_a-value of the reactive site lysine was determined to be significantly lower than in solution.[87,107]

Besides the description of the catalytic tetrad, the phosphate gripper loop, the pyrimidine umbrella and the open/close transition of OMPD, the crystal structures revealed an active site pocket which appears to be too narrow to accommodate the substrate molecule OMP. Most notably, assuming a comparable binding mode as observed for the product complex, with a nucleotide in the *syn*-conformation, the substrate's carboxylate group would sterically clash with Asp$_{312}$ of the enzyme. The spatial proximity of the potentially negatively charged groups led to the formulation of the electrostatic repulsion hypothesis.[109] Based on the crystallographic data and computational studies, Wu *et al.* claimed that the catalytic rate acceleration of OMPD originates from substrate binding in an energetically unfavourable manner. The intrinsic phosphodianion binding energy (IBE) is thought to be utilized to force the substrate's carboxylate group in close proximity to a negative charge. In order to prevent the electrostatic repulsion, the substrate molecule decarboxylates to form the uncharged carbon dioxide (Figure 9).

The mechanism of ground state destabilization (GSD) as catalytic principle was introduced by Jencks in 1975 and is referred to as "Circe Effect" in the literature.[110] It states that "the intrinsic binding energy [...] of a specific substrate with the active site of the enzyme is considerably larger than is generally believed. An important part of this binding energy may be utilized to provide the driving force for catalysis, so that the observed binding energy represents only what

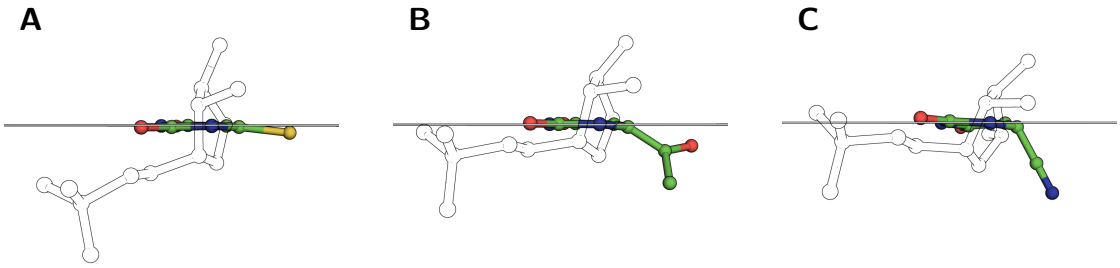


Figure 10: **Ligand distortion in the active site of OMPD.** Ligands with varying C_6-substituents bound to the active site of *h*OMPD. A: 6-mercapto-UMP. B: 6-acety-UMP. C: 6-cyano-UMP. Phosphoribosyl moiety is indicated as outlines. Coordinates were taken from PDB-entries 3L0N, 3L0K and 3EX2.[84]

is left over after this utilization".[110] However, the principle of GSD in natural biocatalysts is under constant debate and challenged by the transition state stabilization TSS proposal.[111–114] In 1948, Pauling stated "that enzymes are molecules that are complementary in structure to the activated complexes of the reactions that they catalyse, that is, to the molecular configuration that is intermediate between the reacting substances and the products of reaction for these catalysed processes."[115] Consequently, the binding affinity of the enzyme towards the transition state has to exceed the binding affinity for the enzyme/substrate complex. Thus, reactant state stabilization would contribute the majority of the catalytic prowess.

Since no structural data is available on the native enzyme/substrate complex, a variety of inhibitor studies were performed. Aided by the increasing resolution of the experimental data and ligands with different C_6-substituents on the pyrimidine ring, a pronounced ligand distortion was observed.[5,84,116,117] The substituents display a rotation along the C_6-X_7-axis and an out-of-plane distortion with respect to the pyrimidine ring towards the C-terminal end of the β-barrel (Figure 10). The origin of the ligand distortion remains unknown since it appears not to be exclusively from steric requirements of the ligand itself. Single and double alanine substitution variants of the reactive site residues Lys_{314} and Asp_{312} having a spacious active site architecture also display a pronounced ligand distortion.[116,118] It remains to be elucidated if ligand distortion contributes to a substantial amount to reduce the reaction barrier for decarboxylation.[117]

The most obvious characteristic during OMPD mediated decarboxylation is the open to closed conformational change induced by ligand binding. Upon substrate/ligand binding, the solvent exposed active site is covered by the phosphate gripper loop and the pyrimidine umbrella. Based on the open to closed transition observed in crystal structures of the *S.cerevisiae* enzyme, the ligand-induced structural transitions were analysed (Figure 5). In yeast, concomitantly to loop closure, two helical elements of the neighbouring subunit partially fold and move closer to the active site pocket. The catalytic tetrad residue Asp'_{317} and the neighbouring hydrophobic amino

acid Ile'$_{318}$ are positioned directly above the pyrimidine ring. Aided by Pro$_{417}$ which is located underneath the base, the ligand is clamped between hydrophobic residues. The phosphate gripper loop residues Gln$_{430}$ and Tyr$_{432}$ and residue Ser$_{372}$ from the pyrimidine umbrella form hydrogen bonds to the ligand in the active site. Arg$_{451}$ of the rigid protein core adopts an alternative conformation to bind to the phosphoryl moiety. In the closed conformation, a hydrophobic pocket is formed directly underneath the tetrad residue Asp$_{312}$. The pocket is formed by Pro$_{417}$, Ile$_{448}$, Phe$_{310}$, Ile$_{401}$, Leu$_{415}$ and Leu$_{366}$ and proposed to accommodate the nascent carbon dioxide after substrate decarboxylation.[119,120]

Furthermore, ligand binding induces a compression of the protein dimer. A large movement can be observed for the interface residue Arg$_{313}$ (sequence neighbour to the catalytic lysine residue) which forms an arginine stack.[121] The interaction is established with its own symmetry mate of the second subunit of the protein dimer. Additionally, Asp$_{285}$ forms a hydrogen bond to Arg'$_{313}$ and is connected to Glu$_{311}$ bridged by a water molecule. The same water is triangled between the N$^{\delta}$-atom of His$_{283}$ and the two mentioned acidic side chains. Loop closure is assisted by mostly hydrophobic interactions between the loop region and the static part of the enzyme molecule. These interactions located "remote" from the active site stabilize the rarely populated closed enzyme conformation and effectively increase the fraction of catalytically competent activated enzyme/substrate complex.[105,122]

Upon ligand binding, the molecule is sequestered from solvent in the interior of the molecule core. In order to elucidate the influence of the formation of the closed enzyme complex, a variety of studies analysed the mechanism and driving forces for loop closure.[98,101–103,105,122–129] The intrinsic phosphate binding energy for triosephosphate isomerase, glycerol-3-phosphate dehydrogenase and OMPD were analysed.[130–132] The substrate for the enzymatic reaction was split into a truncated version lacking the phosphoryl moiety.[133] The second order rate constants of the whole substrates and the phosphodianion truncated substrate piece were related to obtain the total phosphodianion binding energies. The obtained total dianion binding energy of 11 - 13 $^{\mathrm{kcal}}/_{\mathrm{mol}}$ can be divided into the anchoring contribution of the substrate's phosphodianion to the enzyme and the transition state stabilization portion. Analysing the restoration of the catalytic activity of the enzyme for turnover of the truncated substrate by addition of phosphite dianions revealed a specific transition state stabilizing contribution of 6 - 8 $^{\mathrm{kcal}}/_{\mathrm{mol}}$.[131,132,134] The remaining fraction of the total intrinsic phosphodianion binding energy can be attributed to the anchoring effect of the phosphodianion part and the contribution of the covalent linkage between the truncated substrate and the phosphoryl moiety. Thus, the major fraction of the phosphodi-

anion binding energy is specifically utilized during transition state stabilization. Furthermore, the binding energy induces the conformational change to form the catalytically active closed conformation.

Additionally, the split-substrate experiments were applied to clarify that the rate limiting step of the catalytic reaction is not a conformational change.[135] For this, the artificial 5-fluoro derivative of OMP was used in studies which compare the whole substrate and the reaction of substrate pieces.[78,135] 5-fluoro-OMP displays an approx. 30 fold higher turn-over number compared to the natural substrate OMP due to the electron withdrawal effect of the neighbouring fluoro-substituent. Thus, the substrate derivative is stabilizing the carbanionic intermediate emerging from decarboxylation. A comparison of the reaction rate reconstitution of the truncated substrates (1-(β-d-erythrofuranosyl)-5-fluoroorotic acid (FEO) and 1-(β-d-erythrofuranosyl)orotic acid (EO)) with phosphite dianion showed a smaller stabilization of the reaction intermediate of FEO compared to EO ($0.8\,^{\mathrm{kcal}}/_{\mathrm{mol}}$). The small phosphite dianion activation of FEO was attributed to a rate limiting conformational change which is either substrate binding or product release.[135,136]

A combination of experiments using the 5-fluoro derivatives of the product molecule UMP (whole product and phosphodianion truncated molecule) were further utilized to analyse the acidity of the C^6-carbon in the active site of OMPD. Measurements of the hydrogen/deuterium exchange of the C^6-proton of UMP in deuterated solvent indicate a pK_a-suppression from 34 (1,3-dimethyluracil) or 29 (uridine) to ≤ 22 in the active site pocket of OMPD.[137–139] The carbanionic intermediate is stabilized by approx. $13\,^{\mathrm{kcal}}/_{\mathrm{mol}}$ in the active site of OMPD and the 5-fluoro substituent of 5-F-UMP reduces thermodynamic barrier for proton transfer further by approx. $6\,^{\mathrm{kcal}}/_{\mathrm{mol}}$.[135] However, the proton abstraction reaction in the catalytic site requires a proton acceptor to facilitate the exchange reaction. The catalytic base is likely to be the tetrad residue Lys_{314} in its deprotonated form. Furthermore, the generated ammonium group after proton abstraction has to rotate to enable the deuteration of the ligand. The catalytic lysine has to act as catalytic base for deuterium exchange of UMP but has to serve as general acid in the decarboxylation reaction. However, the pK_a-suppression for deprotonation of the carbon atom is comparable to the extent observed in thiamin diphosphate enzyme activation.[140]

7 Results: Orotidine 5'-monophosphate decarboxylase

7.1 Purification of *h*OMPD$_{WT}$

In order to obtain the functional and pure orotidine 5'-monophosphate decarboxylase domain (*h*OMPD$_{224–480}$) of the human uridine monophosphate synthetase (*h*UMPS), recombinant protein expression systems were used. The initial expression and purification strategy was adopted from previous studies in our laboratory based on an existing protocol.[4] The plasmid pETM-30 allows the expression of the gene of interest as a fusion construct, harbouring a hexa-histidine glutathion-S-transferase (His$_6$-GST) protein tag separated by a tobacco etch virus (TEV) cleavage site. The produced His$_6$-GST-*h*OMPD protein construct can be cleaved by TEV-protease, leaving three additional amino acids N-terminal on the target protein (Glu-Ala-Met; Glu: TEV cleavage, Ala-Met: cloning procedure). The resulting modified GAM-*h*OMPD$_{224–480}$ protein domain was used for all studies in this work and is referred to as *h*OMPD in the following sections.

The protein isolation procedures yielded highly pure and functional enzyme based on published catalytic activity values.[5] However, the total amount of obtained protein was rather small and incompatible with material demanding experimental techniques such as protein crystallography. In order to increase the soluble enzyme fraction after protein expression, the *E. coli* SoluBl21(DE3) strain was used as expression host which was generated to specifically increase protein production of human gene sequences. Due to the limitations of the His$_6$-GST-*h*OMPD purification protocol, the expression system was changed after initial trials and the GST-based purification is only described briefly.

The first steps during the isolation of His$_6$-GST-*h*OMPD included the immobilization of the complete construct on a glutathione-sepharose column, elution with 20 mM glutathion containing buffer, protein cleavage with TEV-protease, Ni-NTA affinity chromatography to remove the free protein tag and S200 size exclusion chromatography.

Due to the very low protein yield of the His$_6$-GST-*h*OMPD expression system with protein mainly found in the pellet fractions, a new expression and purification protocol was established. To increase the expression efficiency and the amount of enzyme in the soluble fraction after cell disruption, the yeast protein SMT3, a homolog of the mammalian SUMO-1 protein (referred to as SUMO), was chosen to enhance protein solubility.[141] It was intended to substitute the applied GST-tag to increase the protein purification yield and enable affinity chromatography. The additional three amino acid residues N-terminal of the *h*OMPD domain which are present in

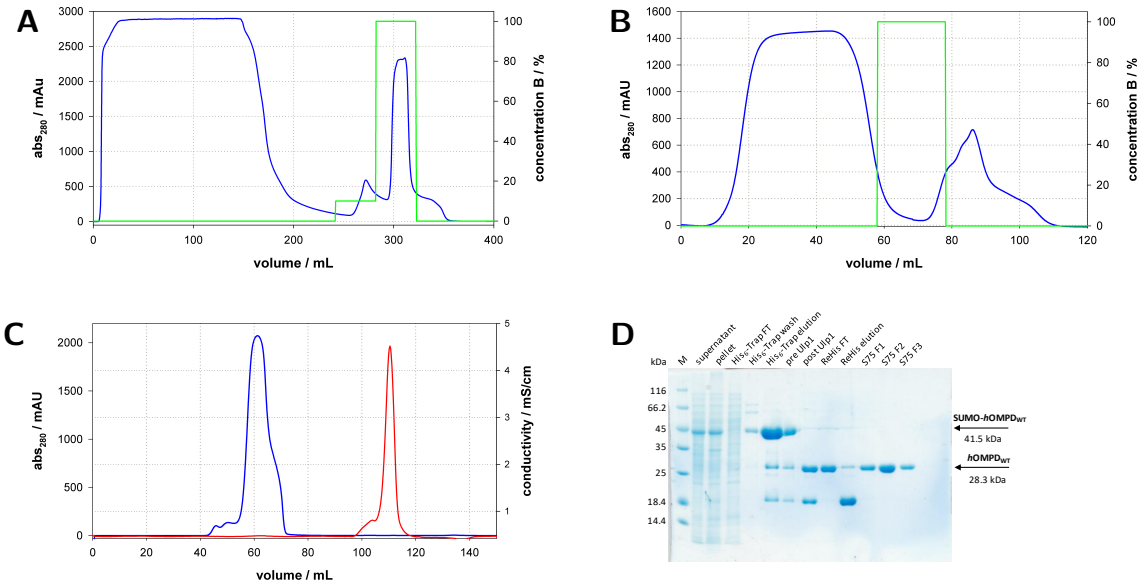


Figure 11: **Purification of *h*OMPD$_{WT}$ utilizing a FPLC-device.** A & B: Ni-NTA affinity chromatography. C: S75 Gel filtration. Blue, green and red traces indicate the measured intrinsic protein absorbance, the relative buffer B (300 mM imidazole) concentration and the buffer conductivity, respectively. D: SDS-PAGE analysis of each purification step.

the His$_6$-GST-*h*OMPD construct were kept to prevent potential negative effects of the removal on the crystallisation propensity. The fusion construct was expressed in *E. coli* SoluBL21(DE3) for 48 h at 30 °C. The sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) of the cleared supernatant of lysed bacterial cells shows a distinguishable band at approx. 40 kDa which corresponds well to the calculated value of 41.7 kDa of the protein monomer. The fusion protein was isolated from the cell lysate by Ni-affinity chromatography and subjected to protein cleavage by the ubiquitin-like protein protease (ULP1) which recognizes the tertiary structure of the SUMO-domain. The protein fragments were applied to an additional step of Ni-affinity chromatography. The labelled SUMO protein and the His$_6$-tagged protease were selectively removed. Potential remaining impurities and incorrectly folded enzymes were removed and the buffer exchanged using a S75 size exclusion chromatography column. The resulting protein sample has a high purity without any detectable contaminations as indicated by a single band at approx. 25 kDa in SDS-PAGE analysis. The *h*OMPD monomer has a calculated mass of 28.3 kDa. The purified protein was consumed within three days or supplemented with 25 % (w/v) glycerol and frozen in liquid nitrogen. Storage of protein in a glycerol stock solution at -80 °C did not influence measured activity values.

7.2 Kinetic characterisation of *h*OMPD

In order to determine the macroscopic catalytic constants k_{cat} and K_M of the OMPD catalysed decarboxylation of OMP, continuous steady-state assays were performed. The formation of the catalytically active enzyme-substrate complex can be analysed using the intrinsic protein fluorescence when excited with light with a wavelength of 280 nm. Protein fluorescence quenching after addition of substrate was utilized as an indicator for the relative abundance of the ES-complex.[107] However, the exact protein species and the mechanism of fluorescence quenching are not determined yet.

Alternatively, the conversion of OMP to UMP can be determined spectrophotometrically. The decarboxylation can directly be monitored based on the differential light absorption of the substrate OMP and the reaction product UMP at 285 nm.[5,107] Time-course measurements of the absorbance show a linear decrease over time (Figure 12). The initial slope of each steady-state reaction was determined, the reaction velocity calculated and plotted against the corresponding substrate concentration. The data points were fitted according to Michaelis and Menten assuming a protein dimer with two independent active sites.[24] The fitting parameters give a value of k_{cat}=3.2 ± 0.04 s^{-1} and K_m=1.5 ± 0.08 μM. Notably, the K_m-value is lower than the first measured data-point and consequently not very reliable. Lower values of substrate concentrations could not be determined due to the fast reaction completion within few seconds and the application of a manual bench-top spectrophotometer.

The spectrophotometric enzyme assay is well suited to determine the catalytic turn-over rate. However, to obtain information on a potential cooperativity between the two active sites of the protein dimer, the assay has severe limitations. At very low substrate concentrations, the decrease in absorbance is very small and the reaction is finished within seconds or less. To differentiate between two independent active sites and active site cooperativity, a high resolution of the reaction rates in low substrate concentrations is necessary. A deviation from a hyperbolic dependence can indicate a communication between both binding sites. To overcome the limitations of the spectrophotometric assay, the decarboxylation was monitored using an isothermal titration calorimeter (ITC_{200}) and the single enzyme injection method.

Applying isothermal titration calorimetry (ITC) offers the opportunity to analyse the substrate dependent reaction rate at very low OMP concentrations.[25,26,142,143] The substrate is loaded into the reaction cell of the instrument and the solution is equilibrated to 25 °C. The reaction is initiated by the addition of a specific amount of enzyme into the measurement cell and the progress followed until reaction completion (Figure 13). After reaction initiation, a following

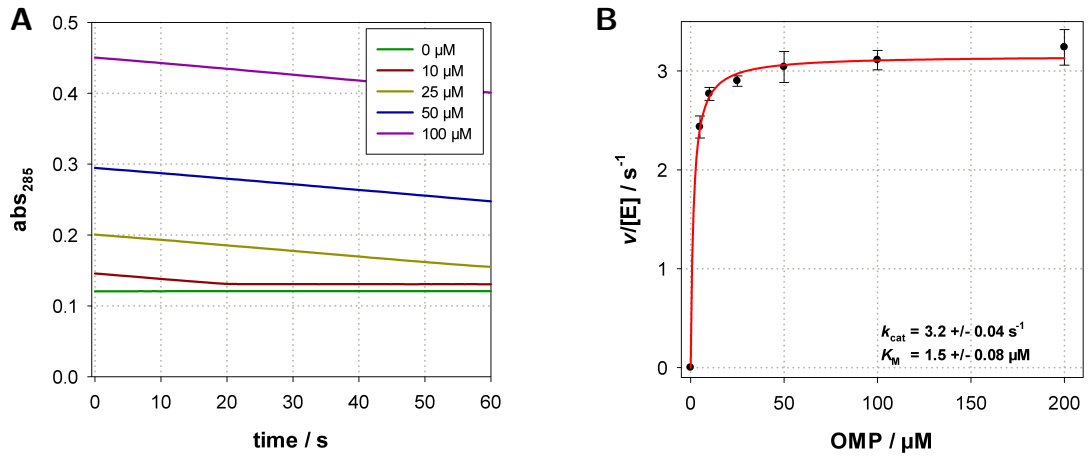


Figure 12: **Photometric determination of hOMPD$_{WT}$ catalysed OMP decarboxylation.** A: Raw-data traces showing linear decrease in absorbance due to different molar extinction coefficients of OMP and UMP ($\Delta\epsilon_{285}$=1650 $M^{-1}cm^{-1}$). B: Analysis of the decarboxylation rate in dependence on substrate concentration. Data points are fitted according to Michaelis and Menten.[24] Measured at 25 °C and 285 nm in 20 mM HEPES/NaOH, pH 7.4 with 0.164 μM hOMPD.

equilibration period can be observed with constantly decreasing differential power values. The mixing to obtain a homogeneous protein concentration requires approx. 100 s and a constant heat-release is observed. The negative differential power value indicates an exothermic reaction which lasts for about 500 s until the signal returns back to the baseline value. A second injection of enzyme into the reaction cell ensures complete substrate consumption to exclude a potential decay of enzyme activity over time.

The total amount of released heat directly correlates to the total amount of converted substrate. The determined apparent molar reaction enthalpy of the decarboxylation reaction (ΔH_{app}) is -6 $^{kcal}/_{mol}$. The ΔH_{app}-value is highly buffer dependent and was determined individually for every measurement. Consequently, the molar reaction enthalpy ($\Delta H_{M,app}$) can be calculated and translates differential power values to actual substrate conversion rates. Assuming a 1:1 stoichiometry of substrate consumption and product release, the OMP and UMP concentrations can be obtained from the decarboxylation rates. The determined turn-over rates were plotted against the corresponding substrate concentration and fitted according to the Hill-equation.[144] The determined catalytic constants are k_{cat}=3.46 $\pm$ 0.01 s^{-1} and K_m=30 $\pm$ 0.2 μM with a cooperativity factor n=1.66 indicating positive cooperativity.

In contrast to the spectrophotometric assay, the calorimetric determination of the decarboxylation rates resolves a substrate concentration range from 0 to 30 μM. It gives the opportunity to analyse reaction velocities at substrate concentrations around and below the K_m-value. The plot of the substrate dependent turn-over rates shows a deviation from a hyperbola with a smaller

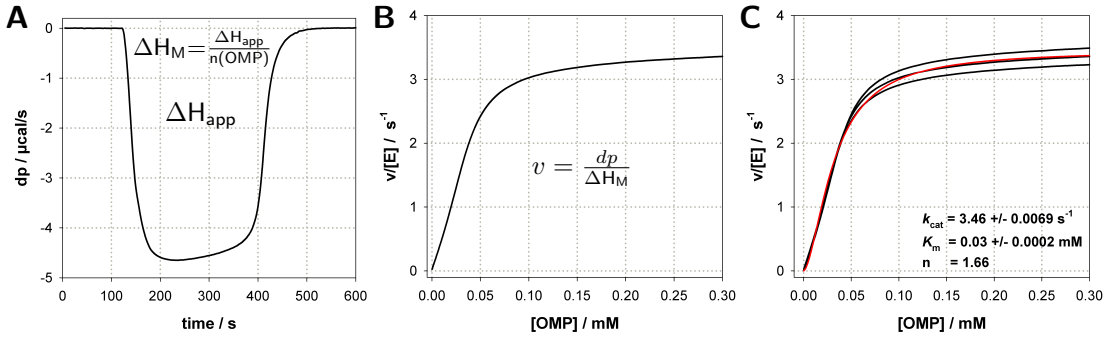


Figure 13: **Analysis of substrate turn-over of *h*OMPD$_{WT}$ using isothermal titration calorimetry.** A: Raw-data thermogram of a representative reaction. B: Calculated reaction rates of OMP decarboxylation in dependence on substrate concentration corresponding to A. C: Substrate turnover fitted according to Hill.[144] Measured at 25 °C in 20 mM HEPES/NaOH, pH 7.4 with 1 μM *h*OMPD.

slope at low OMP concentrations. To account for the sigmoidal curve progression, the Hill-equation is applied with the exponent n indicating cooperativity between the two active sites of the protein dimer. Multimeric subunit assemblies and even monomeric proteins can express synergistic (positive) or interfering (negative) substrate binding and/or turn-over effects.[145–147] In case of ligand binding and a dimeric protein, positive cooperativity increases the affinity of the second binding site once the first site is occupied. Alternatively, the binding affinity of the second site is reduced in case of negative cooperativity. Extreme interfering cooperative interactions can result in half-site reactivity.[148] An interaction between protein subunits effecting catalytic turn-over rates can also be viewed as influencing the affinity for the reaction transition-state.[149]

7.3 UMP binding

The steady-state analysis of substrate turn-over of *h*OMPD indicated a functional interconnection of both subunits of the enzyme dimer. The Hill-coefficient with a value larger than 1 indicates an increase in binding affinity of the second active site once the first site is occupied. Previous studies determined a dissociation constant K_d for the catalytically competent homodimer of 0.25 μM and <16 nM for yeast and human OMPD, respectively. Thus, the presence of momomeric enzyme species in the steady-state measurements is unlikely. Consequently, the values suggest a constitutive dimer formation and it is assumed that ligand binding is not necessary for or inducing dimerization.

In order to further characterise the substrate/ligand binding interactions of the protein dimer, the corresponding intrinsic association constants were analysed using ITC.[5,107] The general ligand

binding reaction is described according to:

$$\mathrm{E} + \mathrm{L} \underset{}{\overset{\mathrm{K_{a1}}}{\rightleftharpoons}} \mathrm{EL} \underset{}{\overset{\mathrm{K_{a2}}}{\rightleftharpoons}} \mathrm{EL_2} \tag{8}$$

$\mathrm{K_{a1}}$ and $\mathrm{K_{a2}}$ refer to the intrinsic association constants ($\mathrm{K_a} = {}^1/_{K_\mathrm{d}}$) for ligand (L) binding to the free enzyme (E) molecule and the mono-liganded EL-complex, respectively. In case of a homodimer and $\mathrm{K_{a1}} = \mathrm{K_{a2}}$, both binding sites are equivalent and independent. If $\mathrm{K_{a2}} > \mathrm{K_{a1}}$, the binding sites behave cooperatively in a positive manner in which ligand binding to the first site increases the binding affinity of the second site. Negative cooperativity occurs when $\mathrm{K_{a1}} > \mathrm{K_{a2}}$ and ligand binding decreases the binding affinity of the second site. The extreme form of non-independent binding sites with negative cooperativity and very low $\mathrm{K_{a2}}$-values shows half-site reactivity. Ligand binding to the first binding site reduces the binding affinity of the second site to an extent that it effectively remains unliganded or inactive.[148] Physiologically, multimeric proteins with positive cooperatvity have a higher sensitivity in a narrower substrate concentration range.**Koshland1996**

In contrast to the concerted Monod, Wyman and Changeux (MWC) model in which the cooperativity effect is expressed simultaneously to ligand binding, the Koshland, Nemethy and Filmer (KNF) model proposes a stepwise association mechanism.[150,151] In case of a dimeric protein, ligand binding to the first side leads to structural rearrangements of the second subunit. The structural differences can either increase or decrease the ligand binding affinity of the second subunit establishing positive or negative cooperativity, respectively.**Koshland1996** Ligand binding can be described as:

$$\mathrm{AA} + \mathrm{L} \underset{\mathrm{L}}{\overset{\mathrm{K_{a1}}}{\rightleftharpoons}} \mathrm{BCL} \underset{\mathrm{L}}{\overset{\mathrm{K_{a2}}}{\rightleftharpoons}} \mathrm{LDDL} \tag{9}$$

AA indicates the symmetric dimer of the apo-protein with two subunits and identical binding affinity. Once the ligand binds to the first subunit, structural rearrangements lead to an asymmetric particle exhibiting unequal binding sites with different ligand binding affinities. Association with the second ligand molecule forms again a symmetric protein. Alternatively, protein allostery can be introduced by structural dynamics effects in which ligand binding to the first binding site increases or decreases protein motions of the second subunit establishing entropically driven cooperativity.**Popovych2006**

However, ligand binding of a macromolecule with multiple binding sites can be described in

stepwise or global association scheme. The stepwise mechanism with separated ligand binding events directly determines the intrinsic association constants of each binding site (K_{ai}). For a dimeric protein ligand binding is described with:

$$E + L \rightleftharpoons^{K_{a1}} EL \tag{10}$$

$$EL + L \rightleftharpoons^{K_{a2}} EL_2 \tag{11}$$

In contrast to a stepwise binding mechanism, the over-all binding constants (β_i) represent the binding constants for a binding event with i ligands. For a dimeric protein ligand binding can be described with:

$$E + L \rightleftharpoons^{\beta_1} EL \tag{12}$$

$$E + 2\,L \rightleftharpoons^{\beta_2} EL_2 \tag{13}$$

The stepwise and global association constants are related by:

$$\beta_i = \prod_{i=1}^{j} K_i \tag{14}$$

In order to determine the specific association constants, the binding affinity of *h*OMPD towards UMP was measured using isothermal titration calorimetry. Due to the low binding affinity and the necessity to apply very high protein concentrations, initial titrations were performed using an ITC_{200}. The differential power values were integrated, related to the injected ligand molecule amount and plotted against the molar ratio of UMP and *h*OMPD-dimer concentration (Figure: 14A). The plot indicates a biphasic ligand binding progression with two inflection points at an approx. molar ratio of one and two. The obtained binding isotherm with a static stoichiometry of 1 defined the values of K_{a1}, K_{a2}, ΔH_1, ΔH_2, $T\Delta S_1$ and $T\Delta S_2$. Both association constants were determined with an error larger than the actual value and considered as very imprecise. To obtain a better resolution, the same titration was performed using a VP-ITC with 10 additional injections. Based on the larger reaction cell, the better signal-to-noise ratio and the additional injection steps, the isotherm was expected to fit the binding isotherm more precisely. Again, both association constants display larger errors than the determined values. In order to

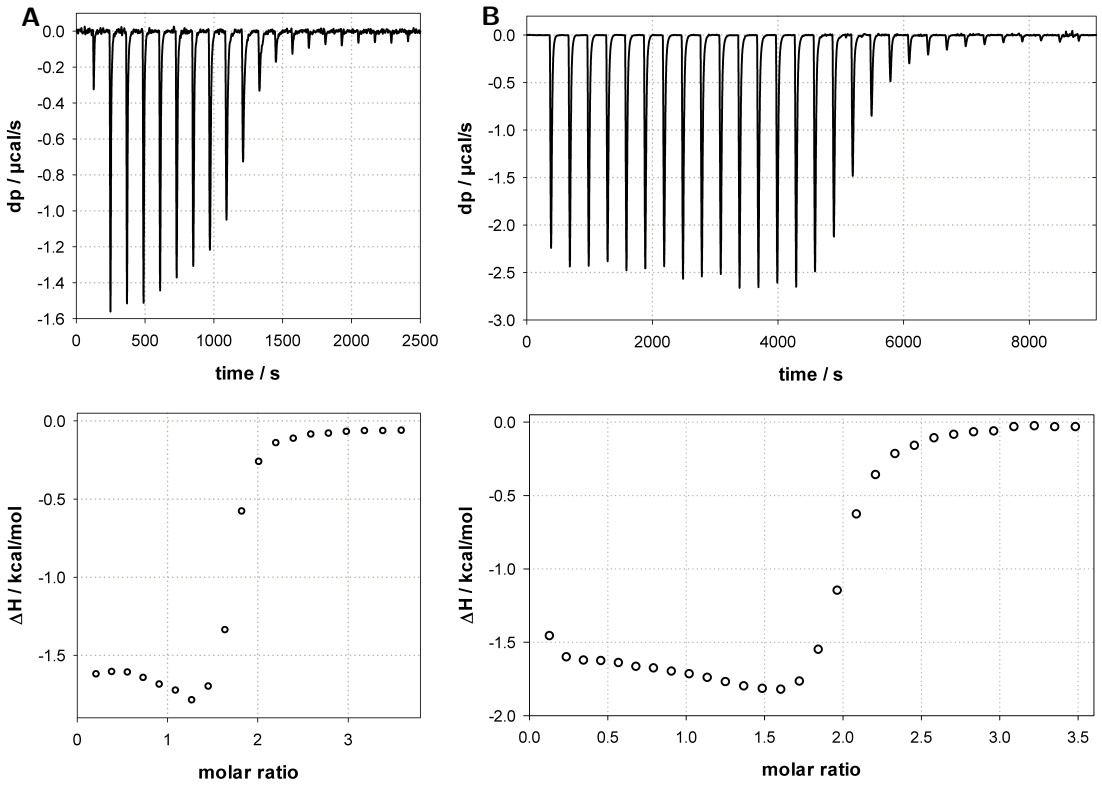


Figure 14: **Determination of UMP binding affinity using isothermal titration calorimetry.** Reactions were recorded with 19 injections of 2 μL (A) in an ITC200 and 29 injection of 10 μL (B) using a VP-ITC. Measured at 25 °C in 20 mM HEPES/NaOH, pH 7.4 with 250 μM *h*OMPD dimer and 4.4 mM UMP.

analyse the the origin of the large errors in the value determination, the titration experiment was simulated.[22,23,152,153] Ligand binding to various ligand molecules (i) can be represented with the general binding parameter which gives the average number of bound ligands to a macromolecule (n_{LB}):

$$n_{LB} = \frac{[L]_b}{[M]_t} = \frac{\sum_{i=0}^{n} i[ML_i]}{\sum_{i=0}^{n} [ML_i]} \tag{15}$$

The concentration of each species can be calculated with the corresponding association constant and the free ligand concentration.

$$\beta_i = \frac{[\mathrm{ML}_i]}{[\mathsf{M}][\mathsf{L}]^i} \tag{16}$$

$$[\mathrm{ML}_i] = \beta_i[\mathsf{M}][\mathsf{L}]^i \tag{17}$$

Substituting the specific species concentrations in the general binding term, the Adair's equation is obtained:[Adair1925]

$$\mathrm{n_{LB}} = \frac{\sum_{i=0}^{n} i\beta_i[\mathrm{L}]^i}{\sum_{i=0}^{n} \beta_i[\mathrm{L}]^i} \tag{18}$$

Based on the Adair's equation, the complete titration experiment was simulated over a wide range of association constants. The corresponding molar reaction enthalpy was calculated for each set of association constants based on two points (injection three and 14 in the VP-ITC experiment) of the experimental data. A simple RMSD-based comparison of the simulated and experimental data yielded solutions of the binding experiments (Figure 68). The obtained values for the binding enthalpy for each binding site are $\Delta\mathrm{H}_1$=-1.624 $^{\mathrm{kcal}}/_{\mathrm{mol}}$ and $\Delta\mathrm{H}_2$=-1.620 $^{\mathrm{kcal}}/_{\mathrm{mol}}$. The values for the dissociation constants for the binding sites are $\mathrm{K_{a1}}$=0.15 mM and $\mathrm{K_{a2}}$=4.4 mM. Thus, the binding affinity of the second binding site is considerably lower than the affinity of the first binding site. Importantly, the determined constants are to be treated very carefully, since they only represent the lower limit of potential solutions for the titration experiment. Reaction simulations assuming higher affinities for the association constants represent the data to the same extent as the solution introduced here as long as they follow $^{\mathrm{K_{a2}}}/_{K_{\mathrm{a1}}} > 10$. Consequently, the assumption that the enzyme displays negative cooperativity can be considered as likely even though the determined absolute dissociation constants might be wrong. The determination of the molar reaction enthalpies is very robust but depends on the buffer composition of each experiment.

7.4 Analysing protein-ligand complexes of *h*OMPD$_{\mathsf{WT}}$ applying cryo-crystallography

In order to obtain insights into the reaction mechanism of *h*OMPD and potential reaction intermediates, protein cryo-crystallography was applied. The purified wild-type protein was crystal-

lized using the hanging-drop vapour diffusion method with ammonium sulfate as main precipitant. The crystallization conditions were adopted from Wittmann *et al.* and modified to obtain high resolution diffraction data.[4,5] However, the initial crystallization trials with the newly generated His_6-SUMO-*h*OMPD$_{224-480}$ fusion construct for improved protein expression and purification failed to produce measurable crystals and formed solely precipitant. Varying precipitant concentrations and pH-values were not influencing the crystallization propensity positively.

The initial published purification protocol utilized a hexa-histidine tagged GST-fusion construct and protein preparations from GST-*h*OMPD expressions yielded moderate crystals. Since the poly-peptide sequences of both constructs were identical after proteolytic digestion and the affinity tag removal, protein modifications introduced during the purification procedure might influence crystal formation. The main difference between the purification strategies is the first step of affinity chromatography. The GST-construct is eluted with reduced glutathion whereas the SUMO-fusion protein is eluted with imidazole. Notably, the previously published crystal structures (PDB entries 2QCD, 2QCE) of human OMPD showed two potential positions for a glutathion induced modification. Two cysteine residues on position 255 and 278 of the amino acid sequence have a distance of approx. 4 Å. Assuming a certain degree of structural dynamics and flexibility of the protein in solution might allow the formation of a disulfide-bridge between these two residues. The covalently linked residues could form crystallization incompetent protein fractions. In the published structures, a disulfide bond was not observed and treatment of the purified protein with a reducing reagent might increase structural dynamics to obtain a protein species capable of crystal formation. Alternatively, structural data indicates the presence of a modified cysteine residue directly exposed to the surrounding solvent. Residue Cys_{304} displays a prominent electron density next to the sulfur atom. It was interpreted as S-mercaptocysteine (CSS) in the crystal structure but the origin, function and identity of the modification is unknown.[5]

Both, the potential disulfide bond and the modified cysteine residue indicated a glutathion induced modification during the purification procedure of GST-*h*OMPD. Addition of 10 mM reduced glutathion to the crystallisation reservoir solution restored the crystallisation propensity of the protein purification derived from the SUMO-fusion construct. The vapour-diffusion method in presence of glutathion enabled crystal formation to obtain comparable results as with the GST-construct. However, the actual role of reduced glutathion remains elusive.

The obtained resting-state crystals of the wild type protein diffracted to a resolution of 0.95 Å at cryogenic temperatures. The atomic model was refined to the experimental data with R_{work}= 10.8 % and R_{free}= 12.0 % in space group $C222_1$ with one monomer in the asymmetric unit. The

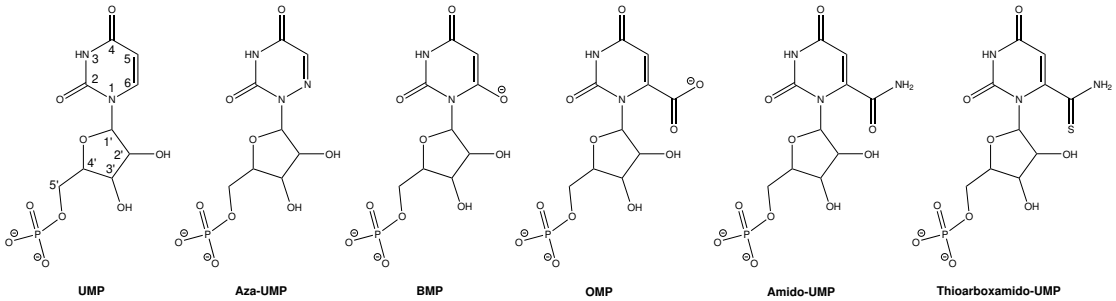


Figure 15: **Chemical compounds utilized to analyse different species of the catalytic reaction trajectory.** OMP and UMP represent the reaction substrate and product, respectively. Aza-UMP and BMP represent proposed transition state analogue inhibitors. Amido-UMP and Thiocarboxamido-UMP are intended to act as isosteric substrate mimics. Atom numbers are indicated in the UMP structure and correspond to the likewise oriented ligands. Compounds were synthesized by Matthias Krull und Tobias Schmidt, Georg-August-Universität Göttingen, Institut für Organische und Biomolekulare, Prof. Dr. Ulf Diederichsen.

overall quaternary structure is identical to previous published models.[5] The structure displays a TIM-barrel fold and the monomer forms a homodimeric complex with the symmetry equivalent subunit. The phosphate gripper loop is closed and covers the active site pocket which is occupied with a sulfate ion, a glycerol molecule in three different conformations and a chloride ion (Figure 17A). The sulfate ion forms potential hydrogen bonds with side- and main chain nitrogen atoms of Arg_{451}, with Tyr_{432} and Gln_{430} of the phosphate gripper loop and with two water molecules as likely hydrogen bond acceptors. The occupancy of the sulfate ion is adjusted to 80 % and water molecules cover the oxygen positions of an absent ion. The flexible glycerol in the active site resembles the bound ribosyl moiety of product and inhibitor complex structures (shown in the following paragraphs). Different conformations of Ser_{257}, Asp_{259}, Lys_{281}, His_{283}, Thr'_{321}, Asp'_{317} and a water molecule form contacts to the glycerol molecule. A chloride ion and two waters cover the positions of the uracyl-moiety of the product complex. The catalytic tetrad establishing the potential alternating charge network composed of Lys_{281}, Asp_{312}, Lys_{314} and Asp_{317} occupies the same positions as observed in the product- and inhibitor-complex structures.

Thus, the obtained resting-state model does not show the real apo-protein structure of the enzyme but more closely resembles the product complex (Figure 17B).

The product complex shows a slightly more compressed structure. The nucleobase contacts the hydroxyl group and the amide nitrogen of Ser_{372}. The UMP-molecule shows two distinct conformations with minor changes in the uracyl-group but with larger displacements in the phosphoryl moiety.

In order to describe the orientation of nucleosides and nucleotides free in solution and bound to the active site of *h*OMPD, three major parameters are used.[154] 1) The sugar puckering (pseu-

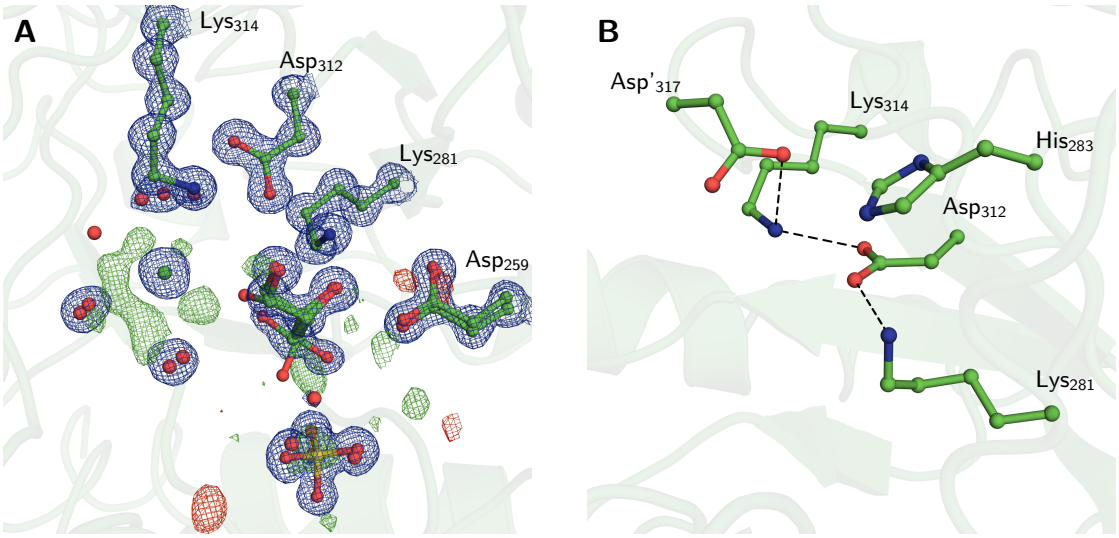


Figure 16: **Active site architecture of the resting state of *h*OMPD at 0.95 Å resolution.** A: Active site overview showing residues of the essential catalytic tetrad. $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with a contour level of 1.5 and $\pm 3\,\sigma$, respectively. Diffraction data was obtained with a resolution of 0.95 Å. B: Entire catalytic tetrad and His_{283}.

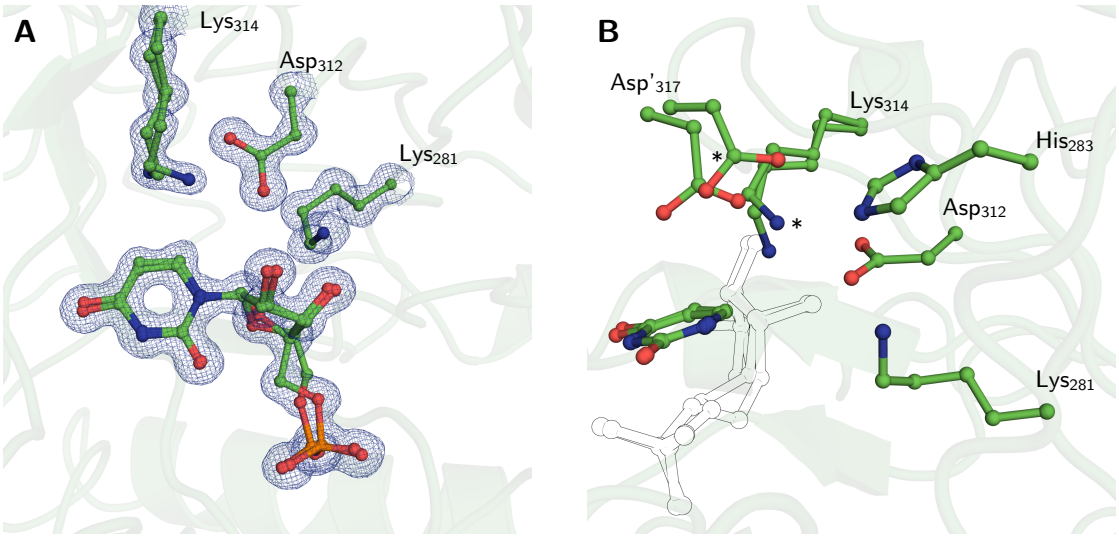


Figure 17: ***h*OMPD$_{WT}$ in complex with the decarboxylation product UMP.** Alternative residue conformations are indicated by the presence of multiple amino acid side chains. A: Active site overview showing residues of the essential catalytic tetrad. $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with a contour level of 1.5 and $\pm 3\,\sigma$, respectively. Diffraction data was obtained with a resolution of 1.0 Å. B: Entire catalytic tetrad and His_{283}. Ribosylphosphate moiety is indicated with outlines. Dominant conformational species are indicated with an asterisk and an occupancy of 70 %.

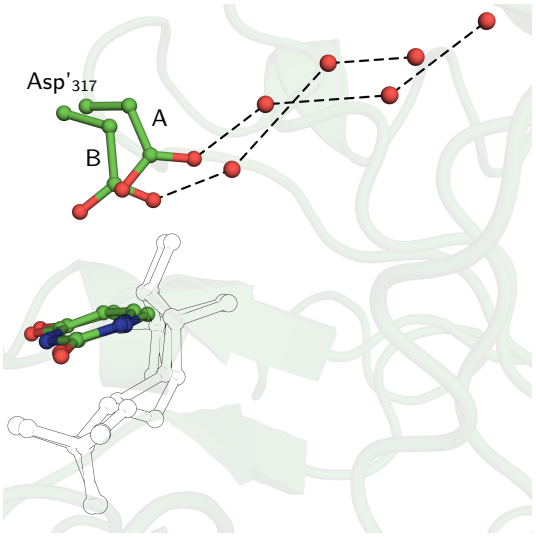


Figure 18: **Orchestrated water displacement of the alternative conformations of Asp$_{317}$.** Dashed lines indicate potential hydrogen bonds between the aspartate orientation and the corresponding water molecules. Asp$_{317}$ and the pyrimidin moiety of UMP are depicted as ball and stick models and the ribosylphosphate part is indicated with outlines. Conformation B and the corresponding waters are modelled with an occupancy of 30 %.

dorotation) indicates the carbon atom positions of the furanose ring with respect to the purine or pyrimidine base. 2) The glycosyl conformation describes the spatial orientation of the base relative to the sugar moiety given by the O'^{4}, C'^{1}, N^1, C^2 torsion angle for pyrimidine bases. And 3) the conformation of the exocyclic C'^{4},C'^{5}-bond.

Experimental diffraction data of the *h*OMPD/UMP product complex was recorded at a resolution of 1.0 Å and the atomic model build with R_{work}= 10.8 % and R_{free}= 12.0 % in space group C222$_1$ with one monomer in the asymmetric unit. Tyr$_{432}$ of the phosphate gripper moves closer to the phosphoryl part of UMP closing the loop more tightly. The N$^{\epsilon}$-atom of Gln$_{430}$ arrests the nucleotide by forming hydrogen bonds to the phosphoryl part and the C^2 carbonyl group of the uracyl moiety. The phosphoryl group interacts with Arg$_{451}$ of the rigid part of the enzyme and two water molecules which are coordinated by Asp$_{259}$ and the main chain interactions of Ile$_{448}$ and Gly$_{418}$. The binding mode of the dominantly populated UMP conformer in the active site of *h*OMPD is C^3-exo (C^2-endo) in the energetically unfavourable *syn*-conformation and gauche-trans with respect to the ribose.

Two residues of the catalytic tetrad occupy two different orientations. Lys$_{314}$ adopts a conformation in which the N$^{\epsilon}$-atom rotates approx. 90 ° towards the C-terminal end of the β-sheet barrel facing the water-filled hydrophobic pocket. The conformation is coordinated by Asp$_{312}$ and Asp'$_{317}$ and identical to the resting-state lysine. The position was determined to be the main conformation with an occupancy of 60 %. The alternative orientation has an occupancy of 40 % and two water molecules in the hydrophobic pocket adopt positions corresponding to the lysine movements. In order to compare the orientation of Lys$_{314}$ with special emphasis on the N$^{\epsilon}$-position with other lysine conformations, the dominant position observed in the UMP-complex

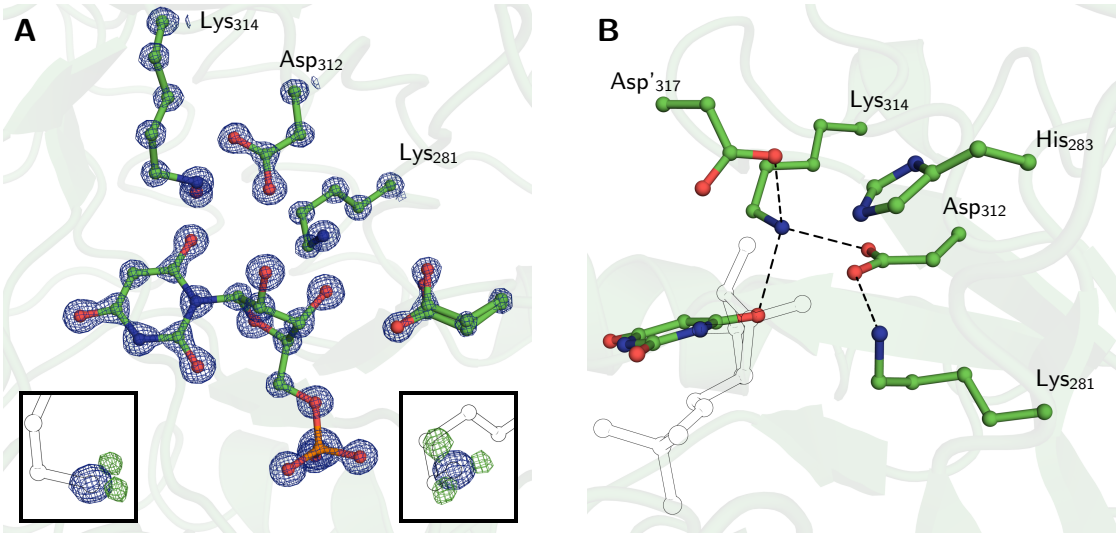


Figure 19: **Atomic resolution crystal structure of *h*OMPD$_{WT}$ in complex with the inhibitor BMP.** A: Active site of *h*OMPD with emphasis on selected residues of the catalytic tetrad. $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with a contour level of 4 and $\pm 3\,\sigma$, respectively. The determined resolution is 0.95 Å. Left inset: Hydrogen omit map of the N$^{\epsilon}$-position of Lys$_{314}$. Right inset: Hydrogen omit map of the N$^{\epsilon}$-position of Lys$_{281}$. The orientation is identical to the main image. B: Hydrogen bond network of the catalytic tetrad. Black dashed lines depict atomic distances of maximal 3 Å forming potential hydrogen bonds. Solid yellow lines indicate angles of selected residues onto the ligand's C^6-position.

is referred to as elongated lysine with a centered N$^{\epsilon}$-atom. The two Asp'$_{317}$ conformations have a distribution of 70 and 30 % with the orientation identical to the resting-state structure as the main population. The carboxylate group of the alternative aspartate moves approx. 1.7 Å closer to C^6-position of the pyrimidine ring. Two water molecules appear to follow the aspartate's movements and adopt positions in hydrogen bonding distance (Figure 18).

In order to obtain insights into the reaction mechanisms of *h*OMPD, structural studies with different enzyme inhibitors were performed (Figure 15). Since the desired ligand molecules are not purchasable, the inhibitor synthesis was performed by Matthias Krull and Tobias Schmidt of the Department Organic and Biomolecular Chemistry of Prof. Dr. Ulf Diederichsen.

The crystal structures with 1 -(5'-Phospho-/β-D-ribofuranosyl) barbituric acid (6-hydroxy-UMP, BMP) were obtained by co-crystallization. Based on the extraordinary high affinity of *h*OMPD towards BMP with a K_d-value of $9 \cdot 10^{-12}$M, it is considered as a potential transition state analogue.[80,81] The strong but reversible BMP-binding indicated a catalytic reaction path involving the development and stabilization of a localized negative charge and the formation of a carbanionic species. The keto-enol tautomers of BMP might form a molecular species resembling the anionic decarboxylation intermediate with a potential oxyanion on position O^6. In barbituric acid the enol tautomer is the thermodynamically favoured entity and the charged enolate form

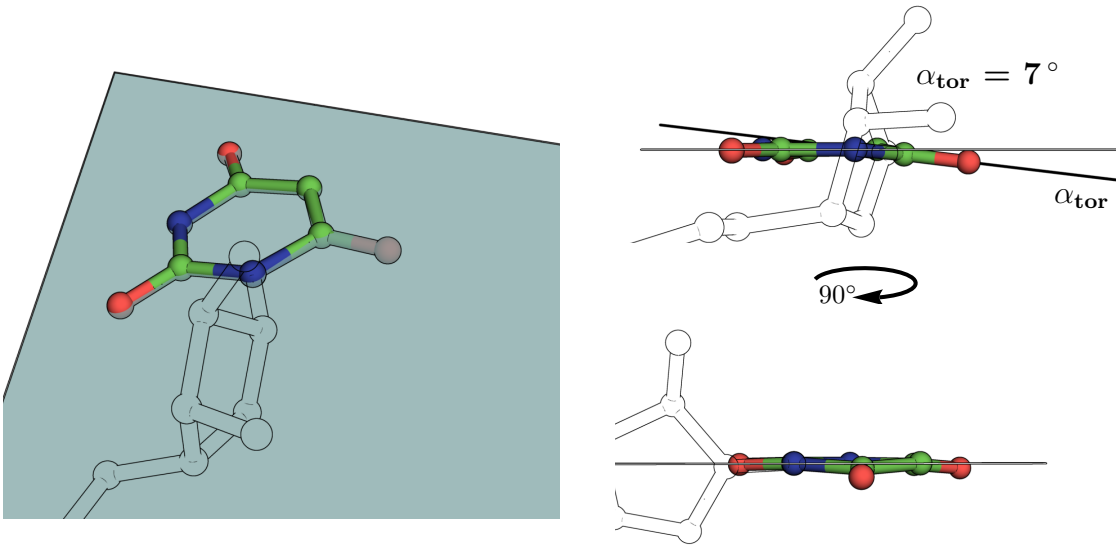


Figure 20: **Out-of-plane distortion of the carbonyl group of the inhibitor BMP in complex with hOMPD$_{WT}$.** Pyrimidin ring plane formed by N^1, N^3 and C^3-atoms. The ribosylphophate moiety is indicated as outlines.

of BMP is suspected to bind with an affinity comparable to the enzyme stabilized transition state.[155] The pK_a-value of BMP is approx. 4.5.[81] Since BMP is the inhibitor of highest binding affinity, it serves as a reference for the description of ligand binding and the analysis of transition state stabilization.

Crystallographic diffraction data of the hOMPD/BMP complex was collected to a resolution of 0.95 Å. The atomic model was build to correspond to the experimental data with R_{work}= 9.9 % and R_{free}= 11.2 % in space group $C222_1$ and one monomer in the asymmetric unit. The phosphate gripper loop is closed and covers the ligand bound to the active site pocket. Gln_{430} and Tyr_{432} adopt the typical positions of a closed enzyme conformation. The active site of hOMPD is identical to the dominantly populated species of the product complex. The residues of the catalytic tetrad display unambiguous positions without the appearance of alternate conformations. Only Asp_{312} shows small flexibility. The catalytic Lys_{314} is coordinated by Asp_{312}, Asp'_{317} and the O^6-atom of the BMP inhibitor and displays an elongated conformation with a centered N^ϵ-atom (Figure 19). The potential carbonyl group of the inhibitor is slightly bend out of the pyrimidine plane with an angle of 7 ° (Figure 24).

To estimate the ionization states of the catalytic tetrad residues, hydrogen omit maps were generated for the potential three hydrogen atoms on the N^ϵ-position of Lys_{314} and Lys_{281}. Two protonation sites could be determined for Lys_{314} in the mF$_o$-DF$_c$ difference electron density hydrogen omit map at 3 σ forming hydrogen bonds to Asp'_{317} and Asp_{312}. In case of Lys_{281}, three hydrogen atom positions were found forming bonds to Asp_{312}, BMP's O'^3-atom and a

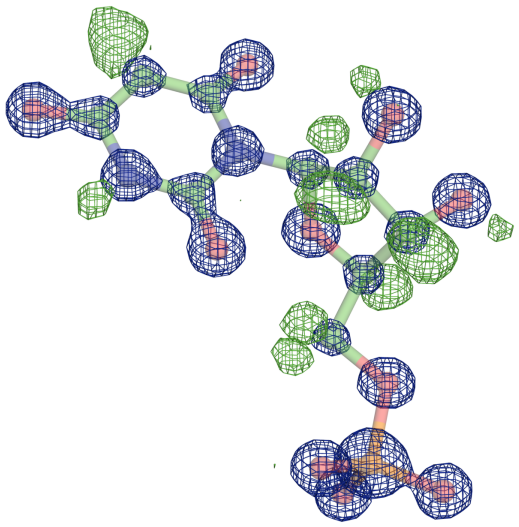


Figure 21: **Hydrogen omit map of BMP bound to the active site of *h*OMPD**. $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with a contour level of 4 and $\pm 3\,\sigma$, respectively. Obtained at a determined resolution of 0.95 Å.

water molecule. The functional groups of Asp'$_{317}$ and Asp$_{312}$ appear ionized with a symmetrical $2mF_o$-DF_c electron density distribution over the carboxylate group and equal C-O-bond lengths.

An analysis of the bond length and the electron density distribution of the pyrimidine moiety indicates the presence of three equivalent functional groups in the pyrimidine ring system. The refined model of the BMP molecule indicates a length of the carbon-oxygen bonds on position C^2-O^2, C^4-O^4 and C^6-O^6 of 1.24 Å ± 0.006 Å, 1.26 Å ± 0.006 Å and 1.25 Å ± 0.006 Å, respectively. The delocalized electron density in the $2mF_o$-DF_c maps indicates the presence of three carbonyl groups on position C^2, C^4 and C^6. A localized negative charge of an ionized O^6-oxyanion would manifest in a larger atomic distance of 1.44 Å and a distinct electron density distribution comparable to the hydroxy groups of the ribosyl moiety. Based on the high resolution of the obtained crystal structure, certain hydrogen atom positions of BMP could be identified in the mF_o-DF_c difference electron density hydrogen omit map at $3\,\sigma$ (Figure 21). The positive difference electron density on the N^3- and C^5-positions of the base indicates the presence of a hydrogen atom. The protonation of the hydroxy groups on the C'^{2}- and C'^{3}-position of the ribosyl moiety could be determined forming hydrogen bonds to Asp'$_{317}$ and Asp$_{259}$, respectively. All possible hydrogen positions on carbon atoms are detectable.

Due to the conformational rigidity of the obtained crystal structures, a distinct interconnection pathway between both active sites of the protein dimer could be detected (Figure 22). A network of water molecules links the catalytic residues of the neighbouring subunit. The waters occupy positions on and close to the crystallographic symmetry axis. Two main water channels can be observed starting from the active site residues Asp'$_{317}$ and His$_{283}$ of subunit A and ending

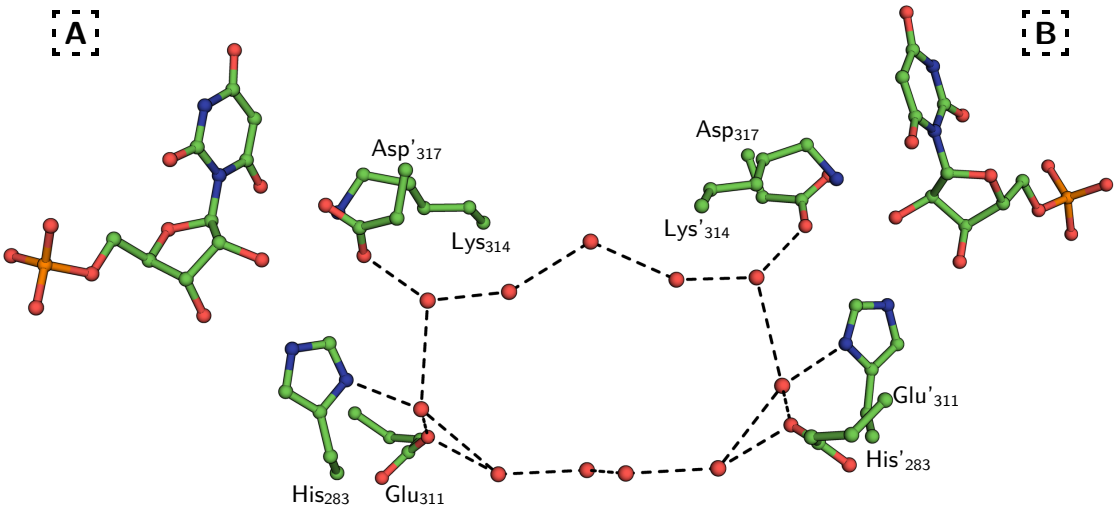


Figure 22: **"Crown-shaped" conformation of the water network interconnecting both active sites of the protein dimer.** Water molecules are indicated as red spheres. Distances below 4 Å are depicted with dashed lines. The dashed boxes indicate the subunit of the protein dimer.

on the corresponding positions of the crystallographic symmetry mate. Both channels are itself interconnected at the initializing water molecules forming a ring-shaped structure referred to as crown-conformation of the water channel. Asp'$_{317}$ of subunit B adopts a position corresponding to the main conformer of the same residue in the OMPD/UMP-complex. five water molecules are observed in the water channel starting from Asp'$_{317}$ with the central position on the symmetry axis. The second interconnection initialized by His_{283} involves five water molecules and Glu_{311} with the central positions as alternative conformers with an equal distribution.

The second analysed ligand in this study is 6-azauridine 5'-monphosphate (Aza-UMP) which has a nitrogen atom on position 6 in the pyrimidine ring. It is a compound discovered early in the *h*OMPD research and was intended as therapeutic drug for cancer treatment.[67,68,81] Aza-UMP has high binding affinity towards *h*OMPD with a K_i-value of 6.4×10^{-9} M. Aza-UMP was found to be most effective when the N^3-atom is deprotonated and the triazine ring (pK_a=7.0) is ionized.[67,68] The high binding affinity of Aza-UMP was explained with two potential binding modes. Resonance structures of the anionic triazine ring might lead to the formation of an oxyanion on position C^2 of the nucleobase. Aza-UMP binding in the *anti*-conformation would lead to a binding mode similar to BMP with a potential localised negative charge on O^2.[81] Aza-UMP binding in the *anti*-conformation was shown in *S. cerevisiae*.[90] Alternatively, the ligand could bind in the *syn*-conformation exposing the N^6-atom to the catalytic tetrad which was observed for *E. coli* and *M. thermoautotrophicum*.[90,109,118] The nitrogen atom with the two free electron pairs might mimic the carbanionic decarboxylation intermediate of *h*OMPD catalysis.

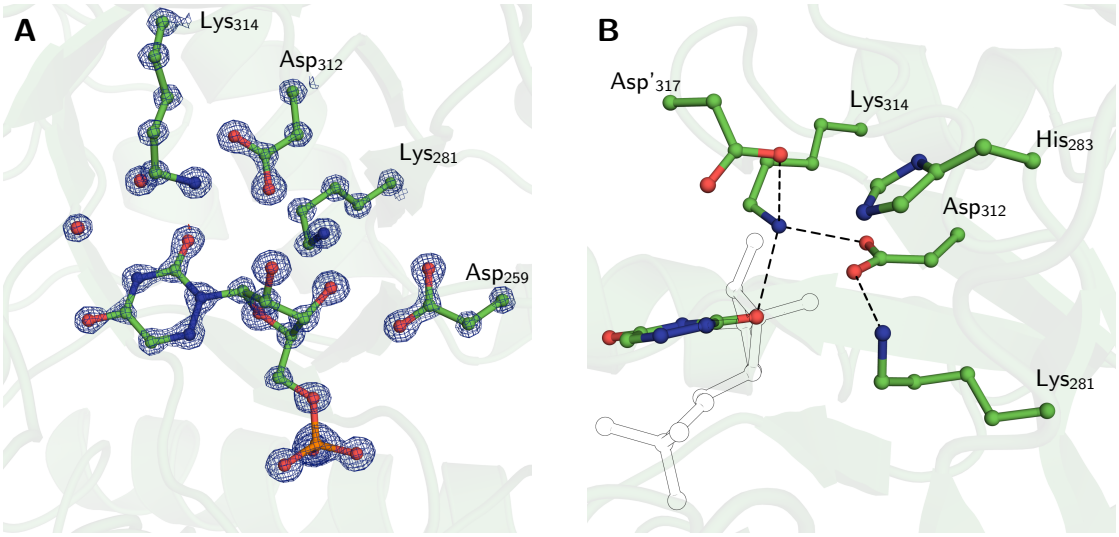


Figure 23: **Active site architecture of hOMPD$_{WT}$ complexed with Aza-UMP at atomic resolution.** Selected residues of the catalytic tetrade are indicated with the corresponding electron density maps. The $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with a contour level of 4 and $\pm 3\,\sigma$, respectively. The determined resolution is 1.0 Å. B: Hydrogen bond network of the catalytic tetrade. Black dashed lines depict atomic distances of maximal 3 Å forming potential hydrogen bonds.

In order to determine the binding mode of Aza-UMP in the human OMPD enzyme, the ligand was soaked into resting state crystals. Crystallographic diffraction data was collected with a resolution cut-off at 1.0 Å. The model was built and refined to represent the experimental electron density with a correlation of R_{work}=11.7 % and R_{free}=13.4 %. The space group of the crystal system was determined to be $C222_1$ with one monomer in the asymmetric unit. The ligand binding mode is very similar to the BMP-complex (Figure 23). The overall structure shows the closed active site architecture with a ligand occupancy of 100 %. The residues of the catalytic tetrad or responsible for ligand binding are positioned as observed in the BMP-complex. Notably, to bind in a conformation in which the carbonyl is oriented towards the catalytic lysine reside, the pyrimidine ring has to adopt the *anti*-conformation with respect to the ribosylphosphate moiety. The flip of the pyrimidine ring leads to the loss or elongation of the hydrogen bond to Gln_{430} which interconnects both ends of the reactive and binding part of the ligand in the BMP-structure. The N^3-nitrogen of the base which interacts with Ser_{372} in the BMP-complex coordinates a water molecule on the catalytic face of the active site. In total, two waters are detected close to the C^2-carbonyl group in hydrogen bond distance to each other.

A comparison of the bond length and electron density distribution of the 2 potential carbonyl groups of the pyrimidine ring in the hOMPD/Aza-UMP complex indicates the presence of two non-equal functional groups. The C-O bond on position C^2 has a length of 1.24 Å ± 0.01 Å and

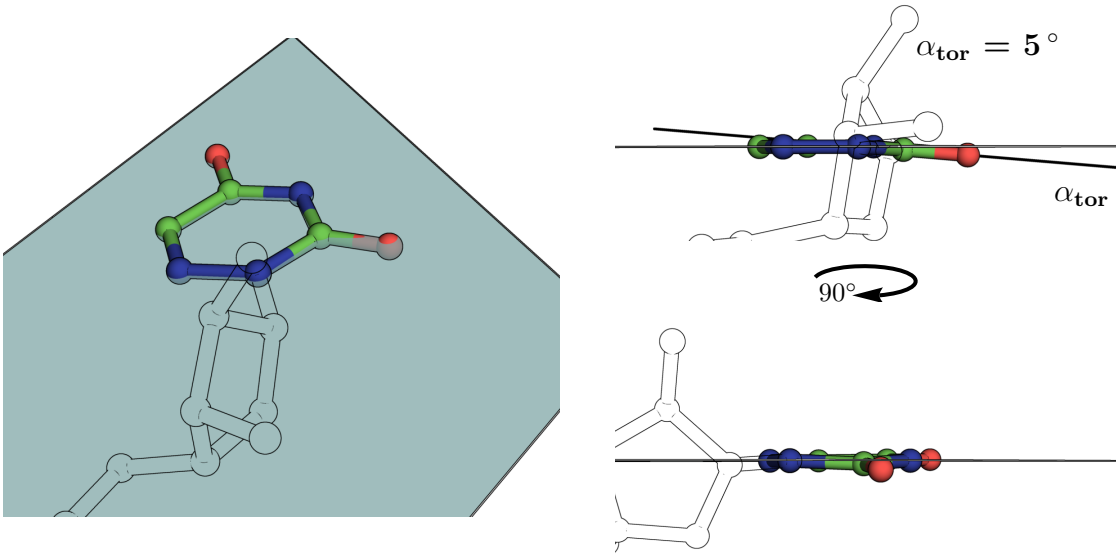


Figure 24: **Out-of-plane distortion of the carbonyl group of the inhibitor Aza-UMP in complex with *h*OMPD$_{\mathbf{WT}}$ bound in the *anti*-conformation.** The pyrimidine ring plane is formed by N^1, N^3 and C^3-atoms. The ribosylphophate moiety is indicated as outlines.

a continuous $2mF_o$-DF_c electron density distribution over the bond at $4\,\sigma$. Both characteristics indicate the presence of a carbonyl group on position C^2 in the dominantly populated ligand fractions. On the C^4 position of the pyrimidine ring, the C-O bond length was determined to be $1.31\,\text{Å} \pm 0.012\,\text{Å}$ with a distinct $2mF_o$-DF_c electron density for the carbon and oxygen atom at $4\,\sigma$ indicating a functional group with hydroxylate character. The mF_o-DF_c difference electron density hydrogen omit map for the potential hydrogen positions of the pyrimidine ring indicates the protonation of the C^5 and N^3-atoms. As observed in the BMP-complex structure, the 3 hydrogen positions of the N^{ϵ}-atom of Lys_{281} and 2 hydrogen atoms of Lys_{314} are detectable.

Both, BMP and Aza-UMP are very strong inhibitors of *h*OMPD activity. Even though the enzymatic transition state stabilization exceeds the BMP binding strength by 12 order of magnitude, the high binding affinity indicates a binding mode as transition state analogue.[156] In order to structurally examine the potential Michaelis-complex with the substrate in the catalytic site, stable substrate analogues were required. 6-amidouridine 5'-monophosphate (Amido-UMP) serves as a stable substrate analogue which is nearly isosteric to the OMP substrate. The carboxylate group of OMP is substituted by a stable amide function to mimic the enzyme bound pre-decarboxylation intermediate. Notably, the amide group is neutral in nature and not negatively charged as is the substrate's carboxylate group. It can serve either as a hydrogen bond donor or acceptor based on the orientation of the carbonyl and amine group, respectively. Amido-UMP has a moderate affinity towards *Sc*OMPD with a K_i-value of $6.2 \times 10^{-4}\,\text{M}$ which is approx. 3 orders of magnitude lower than the substrate OMP.[64]

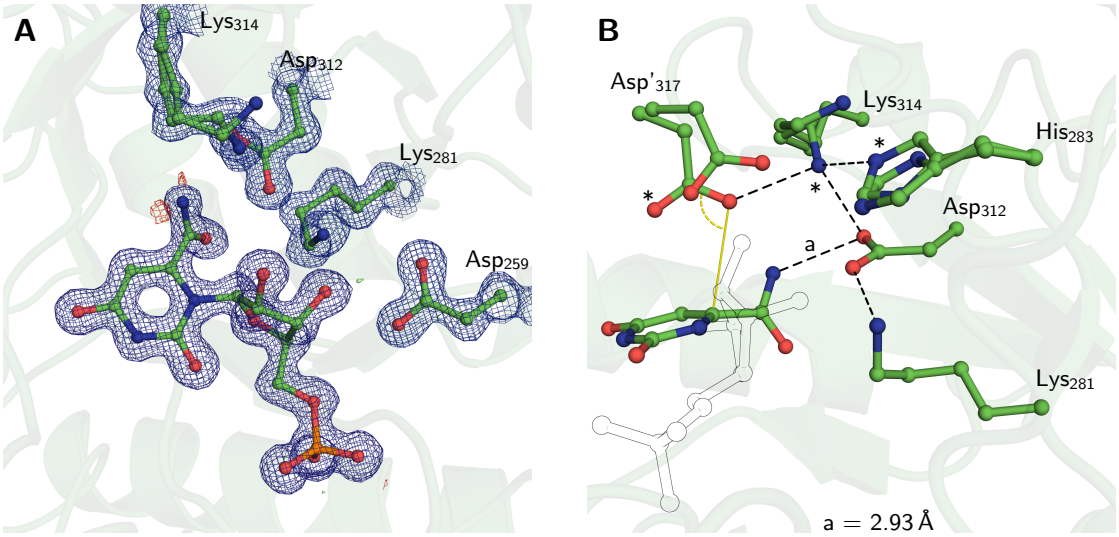


Figure 25: **Crystal structure of *h*OMPD$_{WT}$ complexed with Amido-UMP.** A: Active site overview with emphasis on the charged network of selected residues of the catalytic tetrad. The $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with a contour level of 1.5 and $\pm 3\,\sigma$, respectively. The atomic coordinates are determined at a resolution of 1.2 Å with a ligand occupancy of 90 %. B: Hydrogen bond network of the catalytic tetrad. Black dashed lines depict atomic distances of maximal 3 Å forming potential hydrogen bonds. The solid yellow line indicates the angle of the carboxy group of Asp'$_{317}$ to the ligand's C^6-position. Dominant structural conformations are indicated with an asterisk.

To obtain structural insights into ligand binding, resting state crystals of *h*OMPD were soaked with Amido-UMP. In contrast to previous soaking attempts with Aza-UMP, the ligand concentration was successively increased over four steps to prevent crystal decomposition. The crystals showed an experimental diffraction to 1.2 Å at cryogenic conditions. The ligand complex was indexed in space group C222$_1$ with one monomer in the asymmetric unit. The atomic model was refined to represent the electron density and corresponds with R_{work}=11.9 % and R_{free}=14.7 % to the experimental data. The overall fold is identical to previous obtained structures. The enzyme is locked in the closed conformation with a closed phosphate gripper loop (Figure 25).

An unambiguous electron density indicates the presence of Amido-UMP in the catalytic site and was modelled with 90 % occupancy. The ligand is bound in the *syn*-conformation with the amido group pointing towards the tetrad residue Asp$_{312}$. Notably, three active site residues show an alternative conformation, rotation or displacement with respect to the high affinity inhibitors. In the BMP-complex, Lys$_{314}$ is coordinated by the ligand's carbonyl function and the carboxylate groups of Asp'$_{317}$ and Asp$_{312}$. In the Amido-UMP bound structure, the main conformation of the catalytic lysine's N$^\epsilon$-atom adopts a position between the carboxylate group and the backbone carbonyl of Asp$_{312}$, the carboxylate group of Asp'$_{317}$ and the N$^\epsilon$-atom of a flipped His$_{283}$ conformation. The C$^\delta$- and C$^\gamma$-atoms of the side chain are displaced in order to allow the approx. 2.3 Å rearrangement of the N$^\epsilon$-atom. Furthermore, a small fraction of the

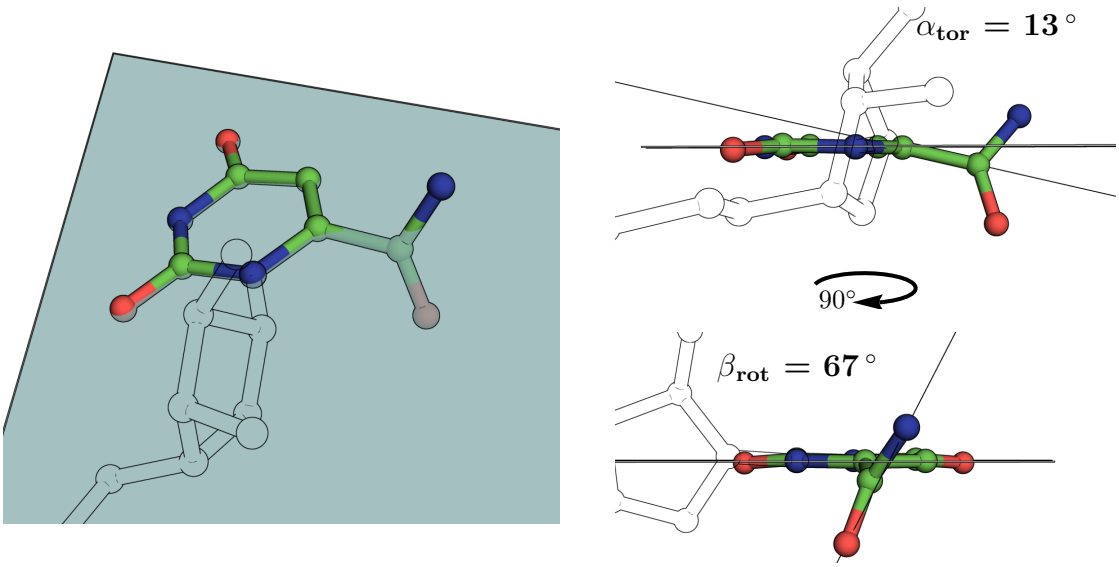


Figure 26: **Out-of-plane distortion and rotation of the ligand's carboxamide group in complex with $h\mathbf{OMPD_{WT}}$.**

Lys_{314} orientations (30 %) adopts a conformation pointing away from the ligand molecule into the water channel connecting the two subunits.

The main conformation of the catalytic Asp'_{317} (occupancy: 70 %) moves approx. 2 Å into a conformation corresponding to the position of the minor populated species Asp'_{317B} in the UMP-complex. The reorientation results in an angle of the carboxylate group of Asp'_{317} towards the C_6-position of the ligand of approx. 109 °. Without the side chain displacement of Asp'_{317}, a reorganisation of Lys_{314} into the observed dominating conformation would not be possible due to spatial constrains. A minor fraction of Asp'_{317} is observed in a position corresponding to Asp'_{317A} in the UMP-complex.

Additionally, the side chain of His_{283} is mainly flipped (70 %) and the N^{ϵ}-atom of the imidazole ring is pointing towards the N^{ϵ}-atom of Lys_{314}. The histidine flip results in the loss of a potential hydrogen bond to the ribosyl-moiety of the ligand but would allow the formation of new hydrogen bonds by the N^{ϵ}- and N^{δ}-atoms of His_{283} towards Lys_{314} and Asp_{259}, respectively. Since Lys_{314} is adopting a previously unrecognised position and to distinguish the lysine orientation in the Amido-UMP complex from the other inhibitor bound structures, it is referred to as "arching lysine". The catalytic tetrad consisting of Lys_{281}, Asp_{312}, Lys_{314} and Asp'_{317} is expanded by the interactions with His_{283} and Asp_{259}. The previously unknown active site architecture is referred to as "extended catalytic tetrad" conformation.

The ligand's amide group shows an out-of-plane torsion and rotation with angles of approx. 13 ° and 67 ° with respect to the pyrimidine plane, respectively (Figure 26). The amide binds

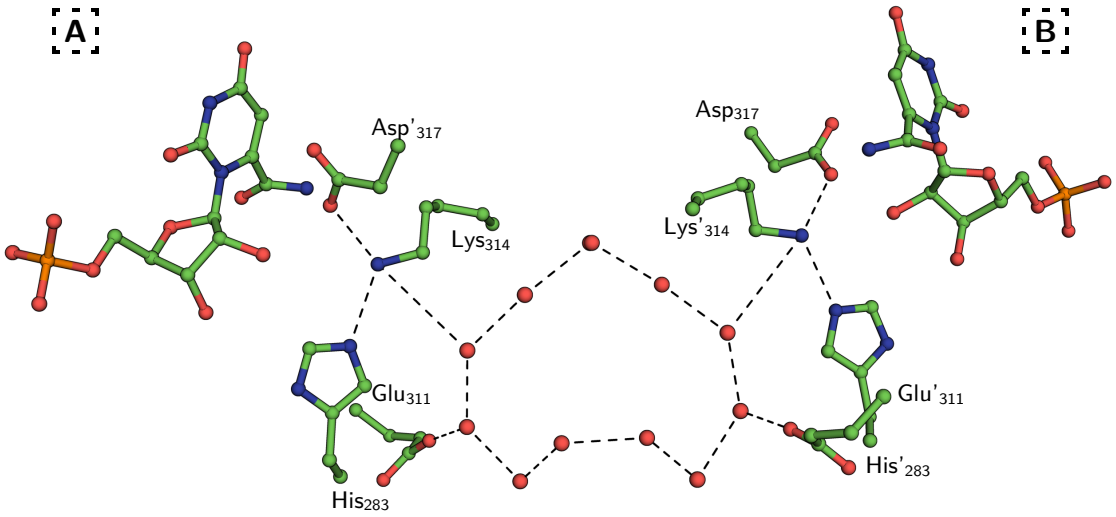


Figure 27: **"Heart-shaped" water channel conformation in the Amdio-UMP complex.** Water molecules are indicated with red spheres. Distances below 4 Å allowing hydrogen bond formation are indicated with dashed lines. The individual subunits of the protein dimer are marked with boxed letters. Only dominant structural conformations are shown.

in direct proximity to Asp_{312} with the amine facing the $O^{\delta 1}$-atom of the carboxylate group. The N-O-distance is 2.93 Å and would allow a hydrogen bond formation with the amine group as hydrogen bond acceptor. The ligand's carbonyl group is oriented towards the hydrophobic pocket, without a potential hydrogen bonding partner nearby. The determined $2mF_o$-DF_c map and the density distribution along the functional group do not indicate the presence of a rotamer of the amide group.

The lysine reorientation into the "arching conformation" results in a reorganization of the two water channels interconnecting both active sites of the subunits (Figure 27). The N^{ϵ}-atom of the mainly populated species of Lys_{314} adopts a conformation between Asp'_{317} and the water network. The connection of His_{283} to the channel is broken since it is present in the flipped conformation. The two mostly separated water channels which were observed in the BMP-complex (initiated by Asp'_{317} and His_{283}) are merged to form a ring shaped structure. To distinguish both water channel arrangements, the network observed in the Amido-UMP bound structure is referred to as being in the "heart-shaped" conformation.

Based on the insights into the new binding mode of *h*OMPD in complex with the potential substrate analogue Amido-UMP, a very similar ligand was synthesized and soaked into resting state crystals. 6-thiocarboxamido-UMP (TCA-UMP) is sterically very similar to Amido-UMP. The carbonyl oxygen of Amido-UMP is substituted by a sulfur group which introduces a high electron density. The binding affinity was determined to be 5 orders of magnitude higher than Amido-UMP with a K_d-value of 3.5×10^{-9} M.[64] However, the binding mode was expected

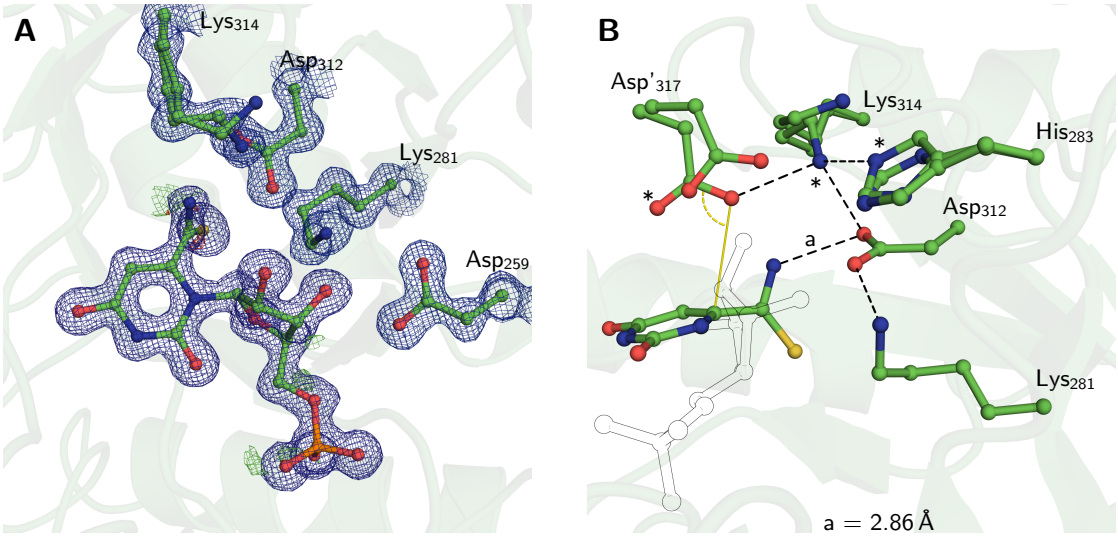


Figure 28: **Crystal structure of hOMPD$_{WT}$ complexed with Thiocarboxamido-UMP.** A: Active site overview with emphasis on the charged network of selected residues of the catalytic tetrad. The $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with a contour level of 1.5 and $\pm 3\,\sigma$, respectively. The atomic coordinates are determined at an experimental resolution of 1.2 Å. B: Hydrogen bond network of the catalytic tetrad. Black dashed lines depict atomic distances of maximal 3 Å forming potential hydrogen bonds. The solid yellow line indicates the angle of the carboxy group of Asp$_{317}$ to the ligand's C^6-position. Dominant positional species are marked with an asterisk.

to be very similar to Amido-UMP. No structural data about the ligand complex is available explaining the discrepancy in binding affinity. As described for the Amido-UMP complex, the amine group of TCA-UMP is expected to serve as hydrogen bond acceptor in the interaction with the carboxylate group of Asp$_{312}$.

The crystals were prepared and soaked as described for the Amido-UMP complex. A stepwise transfer with increasing ligand concentrations was not required and the crystals were soaked and cryo-protected within a single step. The soaking time-points varied between 1 min and 16 h to detect a potential loss of the thiocarboxamide group over time and synchrotron diffraction data-sets were recorded with a total radiation dose of 1 MGy. Experimental diffraction data was collected to a resolution of 1.2 Å at 12.7 KeV. The crystals were processed in space-group $C222_1$ with a single monomer in the asymmetric unit. The model was built and refined to fit the experimental data with R_{work}=12.1 % and R_{free}=15.2 %.

The overall binding mode of TCA-UMP is very similar to Amido-UMP (Figure 28). The enzyme is arrested in the closed conformation with the inhibitor bound in the *syn*-conformation to the active site. The active site architecture underwent the conformational change to form the extended catalytic tetrad with the potential conjugated hydrogen bond network between Asp'$_{317}$, Lys$_{314}$, Asp$_{312}$, Lys$_{281}$, His$_{283}$ and Asp$_{259}$. The conventional tetrad residues Asp'$_{317}$, Lys$_{314}$ and His$_{283}$ of the extended tetrad conformation are found in two distinct conformations.

The positioning of the major conformation of Asp'$_{317}$ corresponds to orientation B in the UMP complex and is occupied with 70 %. It interacts with the dominant conformer of Lys$_{314}$ which is in the arching position and in hydrogen bonding distance to His$_{283}$. As observed in the Amido-UMP complex, the main conformation of His$_{283}$ with an occupancy of 70 % is facing Lys$_{314}$ and Asp$_{259}$ with the N$^{\epsilon}$ and N$^{\delta}$-atom, respectively.

The ligand backbone corresponding to the product UMP is bound in the conventional binding mode. The molecule was refined with an occupancy of 90 % to account for a negative mF$_o$-DF$_c$ difference electron density peak. Analysis of crystal structures obtained from ligand soaking times between 1 min and 16 h under identical soaking conditions indicate a constant population of ligand bound to the active site. There is no indication of enzymatic substrate conversion over the course of 16 h. Since sulfur atoms are very sensitive to radiation damage, even the very low dose of 1 MGy might lead to a small fraction of damaged ligand.[157,158] However, no other sulfur atom or carboxylate group in the vicinity of the active site indicates an effect of radiation induced density loss. Furthermore, the occupancy of TCA-UMP correlates well with Amido-UMP which was refined with 90 %.

The amine group of the thiocarboxamide function is placed in hydrogen bonding distance to Asp$_{312}$ with an N-O-distance of 2.86 Å. The distance is slightly shorter than the one observed in the Amido-UMP complex. Again, the thiocarboxamide group is displaced out of the pyrimidine plane with a torsion angle of 12 °. The rotation of the functional group with respect to the pyrimidine plane is 82 ° and more pronounced than in the Amido-UMP structure. It might be attributed to the larger Van-der-Waals radius of sulfur compared to oxygen with 180 pm to 152 pm, respectively. Higher B-factor values and anisotropic displacement parameters indicate a higher flexibility of the nitrogen of the amine group compared to the sulfur atom.

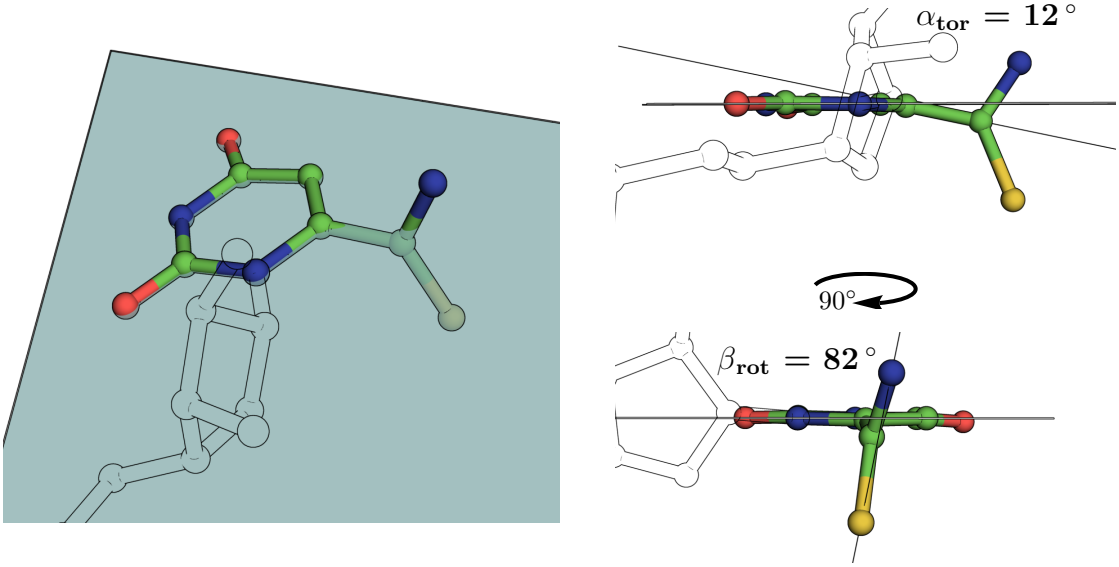


Figure 29: **Out of plane distortion and rotation of the ligand's thiocarboxamide group in complex with $h\mathbf{OMPD_{WT}}$.**

7.5 Purification of *h*OMPD$_{314AcK}$

The key to understand OMPD mediated catalysis is the elucidation of reaction pathway intermediates. The differentiation between the two conflicting proposals of ground state destabilization and reactant state stabilization depends on crystal structures of OMPD with the native substrate in the active site. In 2002, Miller and Wolfenden claimed: "We consider it likely that the substantial binding affinity of OMP in the ground state conceals important repulsive or distortion effects that are relieved in the transition state but which have yet to be clearly identified. Thus, the structure of the enzyme-OMP complex is now of most pressing interest."[156] Until now, crystal structures have been obtained for the resting state enzyme and with a variety of different ligand molecules bound to the active site. But no structural data of the wild type enzyme in complex with reaction intermediates is available. All attempts to soak substrate into resting state crystals failed or yielded the product complex. Only two existing crystal structures showed OMP in the active site but both structures required the generation of an enzyme variant lacking Asp$_{312}$.[5,118]

Previous work demonstrated the importance of the conserved active site lysine (Lys$_{314}$) for the catalytic activity of *h*OMPD. Protein variants of *Sc*OMPD harbouring an alanine substitution on position 93 (equivalent to human position 314) display a 200,000 fold reduction in the turnover rate and low affinity towards OMP.[63,106] Alanine substitution variants might incorporate a water molecule occupying a position in spatial proximity to the former lysine-N$^{\epsilon}$ which could compensate for the mutation effects. Thus, we were aiming for a lysine substitution variant closely resembling the wild type protein without the propensity to act as an acid-base or covalent catalyst and without the positive charge for a potential dipole-dipole interaction. Since the naturally occurring amino-acids do not fulfil our requirements, a system to introduce unnatural amino acids was applied (generously supplied by Dr. Heinz Neumann).

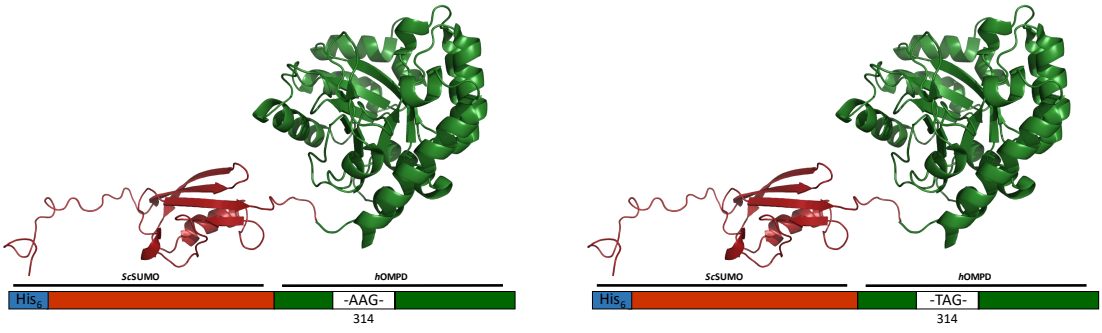


Figure 30: **Schematic representation of the His$_6$-SUMO-GAM-*h*OMPD$_{WT}$ and His$_6$-SUMO-GAM-*h*OMPD$_{314AcK}$ expression construct.**

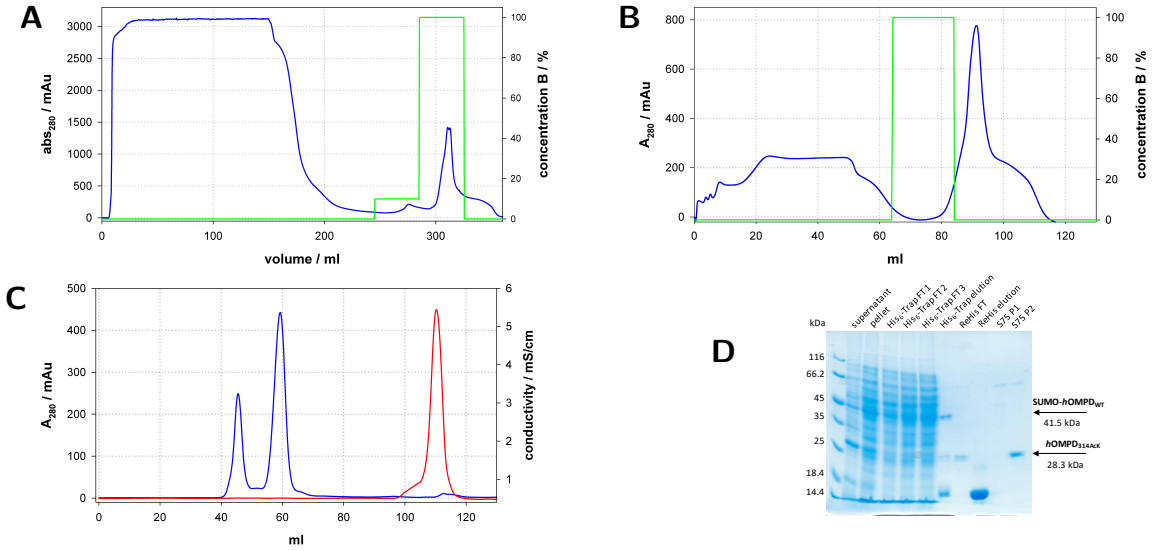


Figure 31: **Affinity chromatography based purification of $hOMPD_{314AcK}$.** A & B: Ni-NTA affinity chromatography. C: S75 Gel filtration. Blue, green and red traces indicate the measured intrinsic protein absorbance, the relative buffer B (300 mM imidazole) concentration and the buffer conductivity, respectively. D: SDS-PAGE analysis of each purification step.

The system to incorporate unnatural amino acids into recombinant protein molecules utilizes the cell's intrinsic translation machinery and relies on the presence of an independent aminoacyl tRNA synthetase/tRNA pair. Certain methanogenic bacteria, including *Methanosarcina barkeri*, naturally facilitate the incorporation of pyrrolysine into methyltransferase enzymes.[159] Pyrrolysine represents a rarely incorporated derivative of commonly utilized lysine and this system is part of the remarkable genetic code expansion in methanogenic *archaea* and *bacteria*. It requires the expression of a specific aminoacyl-tRNA synthetase (aaRS, *Mb*PylS) charging the corresponding $tRNA^{CUA}$ (*Mb*PylT) with pyrrolysine. The incorporation of pyrrolysine into the nascent polypeptide is governed by the TAG- or UAG-codon (AMBER-codon) present in the DNA or mRNA sequence, respectively. Thus, the natural UAG-stop codon induces the translation machinery to incorporate pyrrolysine resulting in AMBER-codon suppression. To facilitate the incorporation of a variety of unnatural amio-acids into a polypeptide chain, a directed evolution approach modified the aminoacyl-tRNA synthetase to accept different substrates. The aaRS variant AcKRS3 accepts N^{ϵ}-acetyllysine (AcK) and charges PylT to specifically incorporate AcK at an UAG-stop codon position.[160,161]

The His_6-SUMO-GAM-$hOMPD_{314AcK}$ expression construct was modified to obtain the TAG-nucleotide triplet on the position coding for Lys_{314}. The protein was expressed in presence of the AcKRS3/PylT-pair and $^{\epsilon}$-acetyllysine. Due to the presence of the natural translation termination pathway of *E. coli*, the AMBER-codon is still recognized as stop-codon which leads to frequent

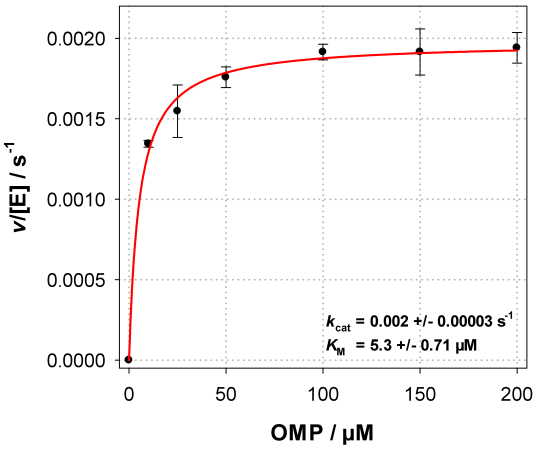


Figure 32: **Photometric determination of hOMPD$_{314AcK}$ catalysed OMP decarboxylation.** Analysis of the decarboxylation rate in dependence of substrate concentration determined by OMP consumption ($\Delta\epsilon_{285}$=1650 $M^{-1}cm^{-1}$). Data points fitted according to Michaelis and Menten.[24] Measured at 25 °C and 285 nm in 20 mM HEPES/NaOH, pH 7.4 with 7.6 μM hOMPD$_{314AcK}$.

abortion of desired polypeptide synthesis. A prominent band in the SDS-PAGE analysis at approx. 25 kDa indicates the presence of the N-terminal translation termination fragment with a calculated molecular weight of 23.8 kDa (Figure 31). However, a second prominent band at approx. 40 kDa most likely represents the full-length expression construct since it can well be isolated from the cell lysate applying affinity chromatography. Treatment of the isolated protein with SUMO-protease yields two polypeptide fractions with a size of approx. 10 and 25 kDa which likely indicates the presence of free SUMO-tag and hOMPD$_{314AcK}$.

The presence of a hexahistidine tagged protein was confirmed by western-blot analysis with antibodies against the affinity tag epitope. After sample digestion, the detected protein fragment is co-localized with the smaller band in the SDS-PAGE analysis. Furthermore, the acetyllysine modified peptide fragment was detected using mass spectrometry after trypsin digestion (Data supplied by Dr. Oliver Valerius, Service Unit LCMS Proteinanalytics). The overall yield of hOMPD$_{314AcK}$ purification was 0.46 $^{mg}/_{g}$ which corresponds to approx. 8 % of the wild type construct.

7.6 Steady-state kinetics of hOMPD$_{314AcK}$

The aim of the sequence specific introduction of $^{\epsilon}$-acetyllysine into hOMPD was the generation of a catalytically impaired enzyme to visualize reaction pathway intermediates. In order to determine the catalytic competence of hOMPD$_{314AcK}$, the substrate dependent decarboxylation rate was analysed. Due to the slow reaction rates, isothermal titration calorimetry failed to detect a measurable signal. The required heat-release for a deviation of the differential power value from the baseline level was too low. The very slow reaction rate and the necessity of a reaction to end with complete substrate consumption rendered the usage of ITC problematic. Very high protein

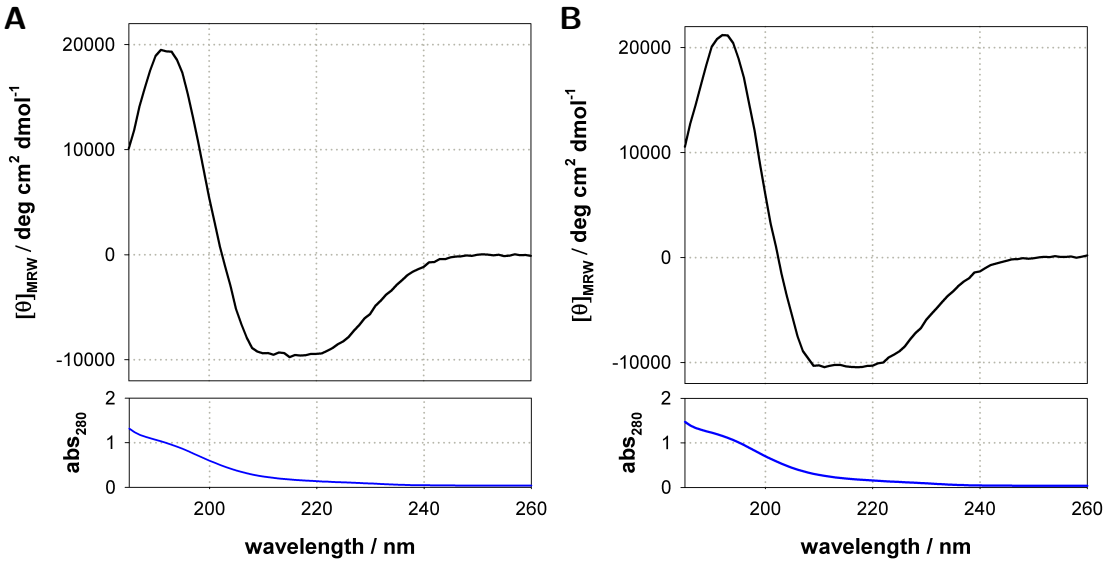


Figure 33: **Protein secondary structure analysis of hOMPD$_{314AcK}$ applying circular dichroism.** Upper panel: Molar mean residue weight ellipticity. Lower panel: Absolute absorbance at 280 nm of the protein sample. A: hOMPD wild-type protein. B: hOMPD$_{314AcK}$ substitution variant. Spectra were recorded of approx. 0.1 $^{mg}/_{mL}$ protein in 10 mM KH_2PO_4 pH 7.4, at 25 °C with 30 averaged repeats.

concentrations and long reaction periods would be needed to obtain analysable data. Since the calorimetric assay is not suitable, the spectrophotometric steady-state decarboxylation assay was performed.

The substrate dependent reaction rate shows a hyperbolic progression from 10 - 200 mM OMP in solution and was fitted according to Michaelis and Menten.[24] The turnover rate $k_{cat}=2 \cdot 10^{-3} \pm 3 \cdot 10^{-5}\,s^{-1}$ and Michaelis constant $K_M=5.3 \pm 0.71\,\mu M$ were determined. The concentration region close to the K_M-value is very poorly resolved, based on the very slow reaction rates. Additionally, the high protein concentration leads to a situation close to equimolaritiy at the first data point with 10 μM substrate in solution violating steady-state conditions. To obtain information on the substrate conversion below the determined K_M-value, the protein concentration would need to be reduced. However, the protein concentration has to be high enough to obtain a measurable decrease in absorbance over time.

7.7 Analysing secondary structure elements of *h*OMPD$_{314AcK}$ using circular dichroism

The macroscopic kinetic constants of hOMPD$_{314AcK}$ indicate a severely impaired enzyme in terms of substrate turnover. To ensure that the reduced catalytic reaction rates are the result of the silencing of Lys$_{314}$ function and not an effect of incorrect protein folding, the secondary

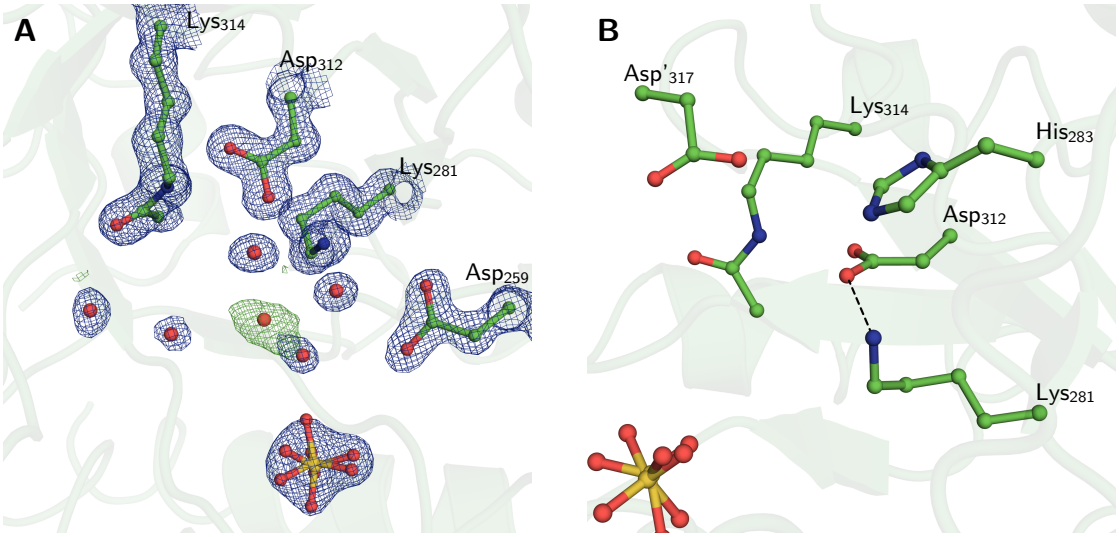


Figure 34: **Crystal structure of hOMPD$_{314AcK}$ resting state.** A: Active site overview with emphasis on the charged network of selected residues of the catalytic tetrad. The $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with a contour level of 1.5 and $\pm 3\,\sigma$, respectively. The atomic coordinates are determined at a resolution of 1.2 Å. B: Hydrogen bond network of the catalytic tetrad. Black dashed lines depict atomic distances of maximal 3 Å forming potential hydrogen bonds.

structure elements were analysed using circular dichroism (CD, Figure 33). Absorbance spectra of the wild-type enzyme and the lysine substitution variant were recorded from 185 - 260 nm and the molar mean residue weight ellipticity ($[\Theta]_{MRW}$) calculated.[3] The observed traces of both proteins indicate the presence of α-helical structural elements which would be expected for the TIM-barrel fold. Both spectra are very similar and the hOMPD$_{314AcK}$ protein appears to be properly folded. The subtle differences in the two obtained CD spectra could arise from a different folding of small protein regions or shifts in the equilibrium of protein conformations e.g. the open/close conformations of the phosphate gripper loop. Since the secondary structure determination using circular dichroism is not sensitive enough to obtain detailed structural characteristics of the enzyme variant, X-ray cryo-crystallography was applied.

7.8 Analysing the hOMPD$_{314AcK}$ protein variant applying cryo-crystallography

The secondary structure of the protein variant appeared similar to the wild-type protein. To obtain more detailed structural information, hOMPD$_{314AcK}$ was crystallized according to the established protocol and the diffraction patterns recorded at cryogenic conditions. The resting state enzyme crystallized morphologically identical to the wild-type protein. A slightly increased precipitant concentration was needed to grow crystals with a size of approx. $200 \times 200\,\mu$m.

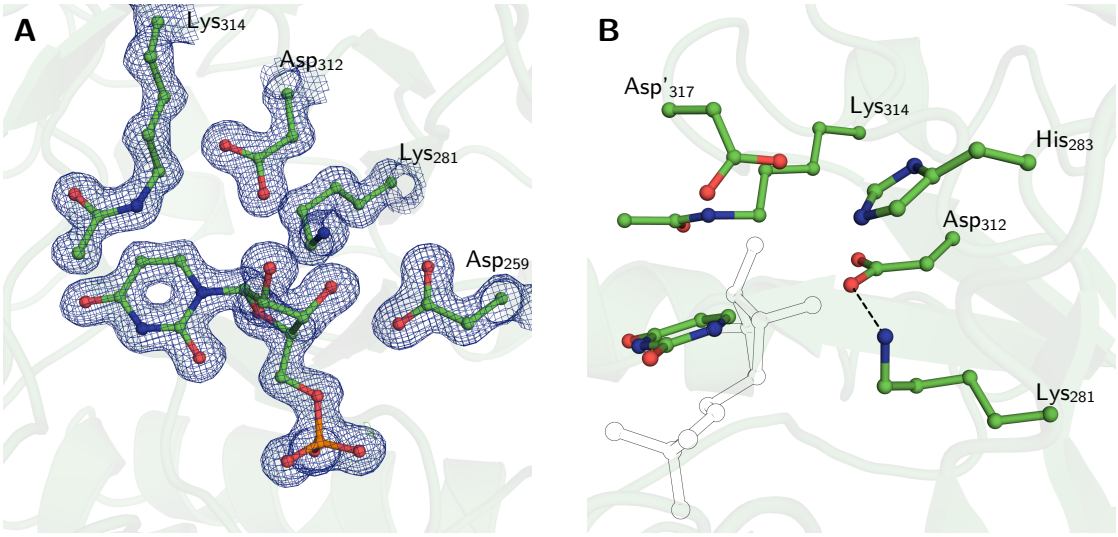


Figure 35: **Crystal structure of hOMPD$_{314AcK}$ complexed with UMP.** A: Active site overview with emphasis on the charged network of selected residues of the catalytic tetrad. The $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with a contour level of 1.5 and $\pm 3\,\sigma$, respectively. The atomic coordinates are determined at a resolution of 1.2 Å. B: Hydrogen bond network of the catalytic tetrad. Black dashed lines depict atomic distances of maximal 3 Å forming potential hydrogen bonds.

Experimental diffraction data was collected to a resolution of 1.2 Å and the model refined to reach an R_{work} and R_{free}-value of 13.8 % and 16.2 %, respectively. The protein crystallized in space group $C222_1$ with a single monomer in the asymmetric unit. The tertiary structure of the resting state crystal structure of hOMPD$_{314AcK}$ is very similar to the wild-type protein. The rigid TIM-barrel fold shows a root-mean-square deviation (rmsd) of 0.2 Å. Notably, the structure shows the opened active site conformation with no clear electron density for the phosphate gripper loop which spans the residues 423 - 432. A slightly rotated sulfate ion is located in the phosphate binding site of the enzyme. Arg_{459} has two alternative conformations orienting to and away from the sulfate ion. Glycerol cannot be detected in the active site as observed in the closed wild-type complex.

As a reference for the structural organization of the active site in the hOMPD$_{314AcK}$ variant, product complex crystal structures were analysed (Figure 35). The introduction of the acetyl group and the corresponding spatial requirements of the lysine derivative could lead to structural alterations in the active site architecture of the enzyme. Protein crystals were obtained by co-crystallization of hOMPD$_{314AcK}$ with the product molecule UMP. Synchrotron diffraction images were collected to a resolution of 1.25 Å and model adjusted to correspond to the experimental data with R_{work} = 13.8 % and R_{free} = 16.2 %. The asymmetric unit contains one monomer of the obligate dimer in space-group $C222_1$. The tertiary structure very closely resembles the closed product complex of the wild-type enzyme with an rmsd of 0.01 Å (Figure 36). Most active

site residues are placed accordingly except for Met_{371}. The N^{ϵ}-atom of AcK_{314} is rotated and pointing towards the neighbouring monomer of the dimer. The bulky acetyl group on the N^{ϵ}-atom of Lys_{314} hovers over the pyrimide ring of UMP displacing the C^{ϵ}-atom of Met_{371}. The catalytic Asp'_{317} adopts the conformation A as observed in the wild-type complex, even though Lys_{314} cannot serve as a hydrogen bonding partner. The remaining residues of the catalytic tetrad are found positioned likewise to the wild-type structure. Consequently, the introduction of acetyllysine into the active site pocket and the elimination of the positive charge of Lys_{314} does not prevent protein folding and active site formation.

The $h\text{OMPD}_{314\text{AcK}}$ appears suitable to analyse potential reaction intermediates of the enzyme catalysed decarboxylation. Importantly, the active site Asp_{312} is oriented properly to exert electrostatic stress onto the carboxy-group of the substrate if reactant state destabilization is applied as major driving force for catalysis. The variant's affinity for OMP is moderately high with 5 μM and would allow substrate binding in the crystal-soaking trials. There is no conclusive information present concerning saoking durations to obtain full active site occupancy. Consequently, several soaking time-points from 30 s to 30 min were chosen to analyse substrate binding and substrate conversion. Complete substrate occupancy was observed after 2 min of incubation in the substrate soaking solution and flash cooling in liquid nitrogen (Figure 37). The crystals grew in the C-centered orthorhombic space-group $C222_1$ with one monomer in the asymmetric unit. The native protein dimer can be obtained by the application of symmetry laws. The active site architecture is identical to the UMP product complex. The acetllysine is placed above the pyrimidine ring of OMP displacing the C^{ϵ}-atom of Met_{371} with the acetyl group.

The carboxy group of the substrate molecule is indicated by an unambiguous electron density close to the C^6-carbon of the pyrimidine ring. All functional groups of the ligand can be assigned and the corresponding positions determined. The substrate molecule is bound in the *syn*-conformation placing the carboxy-group in direct spatial proximity to Asp_{312}. With 2.45 Å, the O-O-distance of the two carboxy groups is relatively short and can be considered as "short, strong" hydrogen bond.[162–164] There is a clear asymmetry between the two acidic functions. The carboxy group of OMP shows a distinct $2mF_o$-DF_c electron density located on the oxygen atom contacting Asp_{312} and a continuous signal over the C^7 atom and the distal oxygen. The electron density indicates the presence of a carbonyl facing the hydrophobic pocket and a hydroxy group interacting with Asp_{312}. The carboxy group of Asp_{312} is likely to be ionized based on the symmetric electron density distribution over the functional group. In the crystallographic data and the calculated electron density maps, the hydrogen atom between the two functional groups is

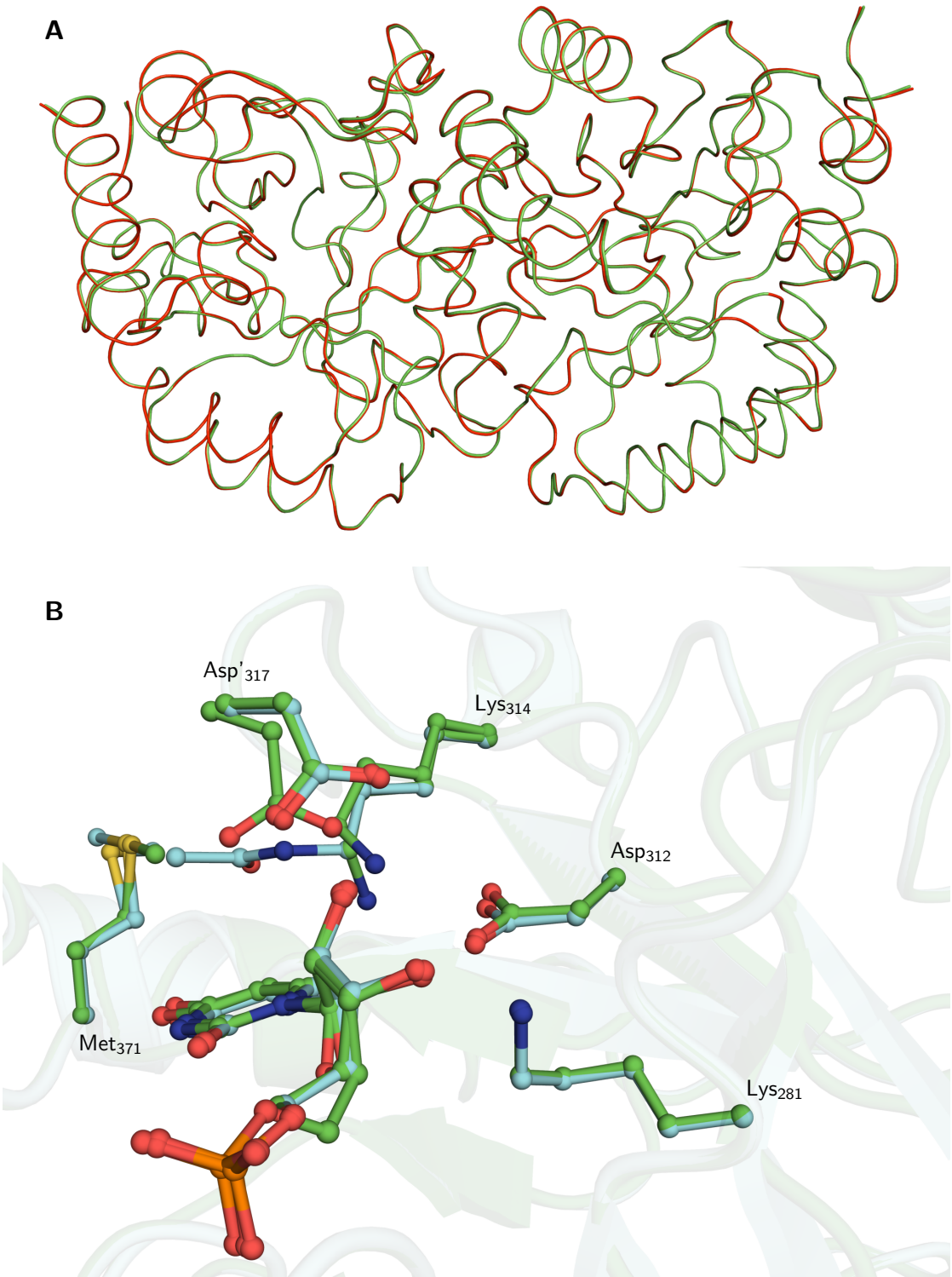


Figure 36: **Alignment of the UMP product-complex crystal structures of hOMPD$_{\mathbf{WT}}$ and hOMPD$_{\mathbf{314AcK}}$.** A: Ribbon representation of the C^{α}-atoms of the wild-type enzyme (green) and the acetyllysine harbouring protein variant (red). rmsd = 0.01 Å B: Active site view with emphasis on the catalytic tetrad and the Met_{371} displacement. The wild-type enzyme and the acetyllysine harbouring protein variant are depicted in green and cyan, respectively.

not visualized. However, the short distance between the two groups indicates a protonation of the substrate and the likely formation of a carboxylic acid to prevent the electrostatic repulsion between two negative charges. A structural alignment of the substrate and product complex shows a movement of Asp_{312} even closer to the carboxy group of OMP indicating a productive interaction. The angle of 114 ° appears suitable for the hydrogen bond establishment (Figure 36).

The substrate's carboxy group is placed directly over a previously described hydrophobic pocket.[105] The side chain of Ile_{401} is located directly underneath the carboxy group of OMP and the C^{γ} and C^{δ}-atoms are placed nearly parallel to it. However, the CO^2-leaving group is not in plane with the pyrimidine ring and bend towards the hydrophobic pocket. The out-of-plane distortion results in a displacement of the C^7-atom from the ring plane (N^1, N^3, C^5) with an angle of 27 ° towards the hydrophobic pocket residues (Figure 38). The determined distortion is in between the published values of 16 ° and 37 ° for OMP bound to active site of enzyme mutants.[84,109] Consequently, the carboxy group is located entirely underneath the pyrimidine ring plane of OMP. Additionally, a rotation of 43 ° can be observed.

The two catalytic subunits are interconnected by a water channel as observed in the other ligand complex structures. The residue Asp'_{317} adopts a position corresponding to the main conformation A in the UMP complex. The water tunnels are found in the crown conformation without the participation of Lys_{314} which is facing the substrate molecule (Figure 39).

In the data-series of different soaking times, durations less than 2 min showed only a partial binding site saturation with the substrate molecule. Collecting data-sets of cryo-cooled protein crystals soaked with OMP for 2 min showed full occupancy for the uridine part of OMP but a slightly smaller signal for the C^6-carboxy group. Thus, the substrate appears to be partially decarboxylated with a minor fraction of UMP product present in the active site. The weaker signal for the functional group might be based on the short period of being in the protein cage of $h\text{OMPD}_{314\text{AcK}}$ and a result of enzymatic catalysis. Alternatively, x-ray induced radiation damage could have resulted in the decarboxylation of the substrate molecule as it is routinely observed for carboxy-groups of amino acids of crystals exposed to high doses of high energy photons.[165,166] To exclude the possibility of radiation induced decarboxylation, a dose series of iterative illumination was performed. The identical 200 ° wedge of the protein crystal was used for subsequent data-set collection with a calibrated amount of energy introduced during exposure with a top-heat beam profile. Five data-sets were collected with a dose of 0.71 Mgy to obtain a total dose of 3.55 Mgy in the last data-collection. The image frames were analysed and the

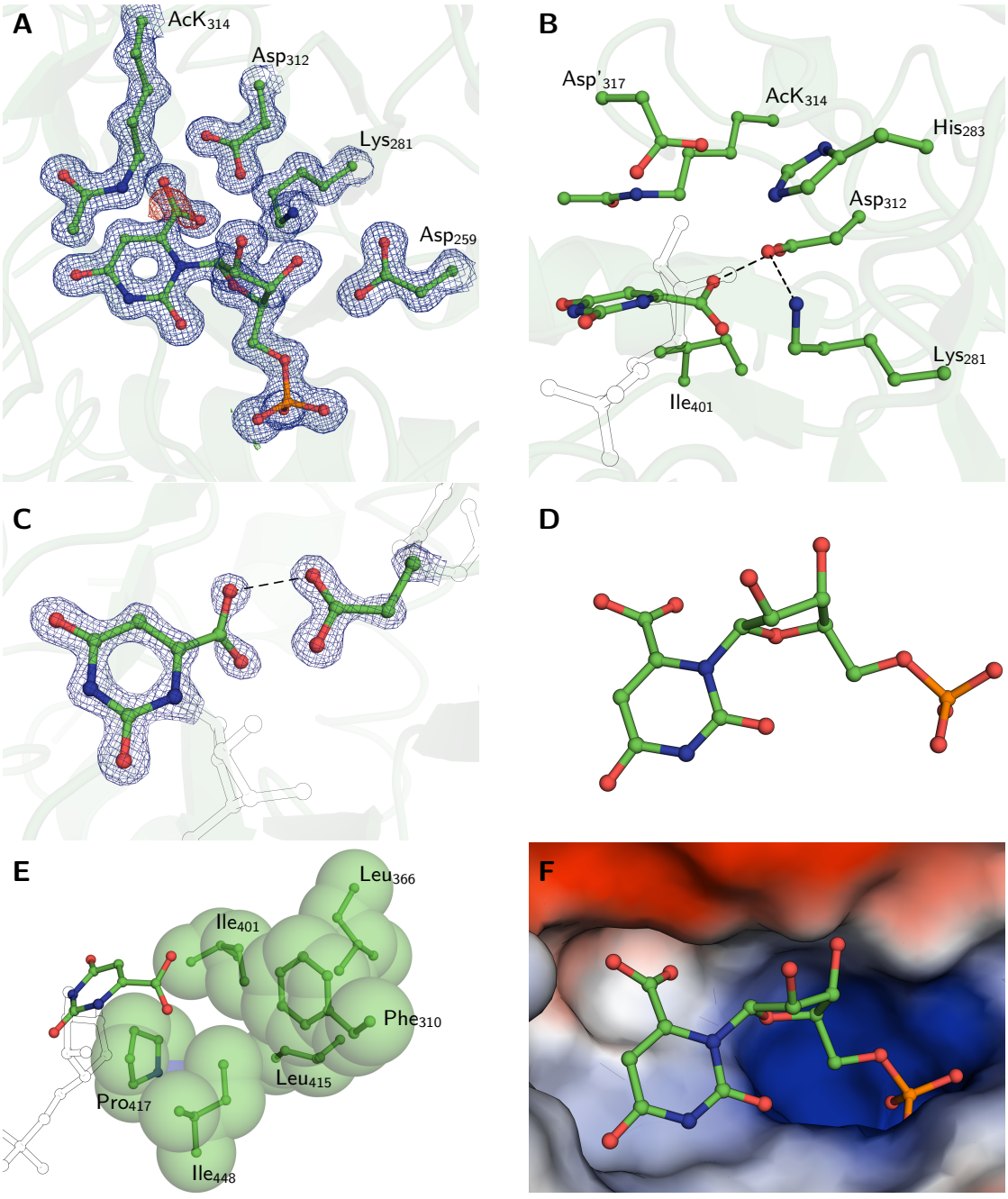


Figure 37: **Crystal structure of hOMPD$_{314AcK}$ in complex with the natural substrate OMP at 1.2 Å.** A: Active site overview with emphasis on the catalytic tetrad. The $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with a contour level of 1.5 and $\pm 3\,\sigma$, respectively. B: Hydrogen bond network of the catalytic tetrad. Black dashed lines depict atomic distances of maximal 3 Å forming potential hydrogen bonds. C: Short hydrogen bond between the substrate's carboxy group and Asp$_{312}$. The $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with a contour level of 3.5 and $\pm 3\,\sigma$, respectively. The occupancy of the substrate's carboxy group is adjusted to 85 %. D: Binding mode of OMP in the active site of hOMPD$_{314AcK}$. E: Hydrophobic pocket residues in direct proximity to the carboxy group of OMP. F: Electrostatic surface potentials of the active site cavity.

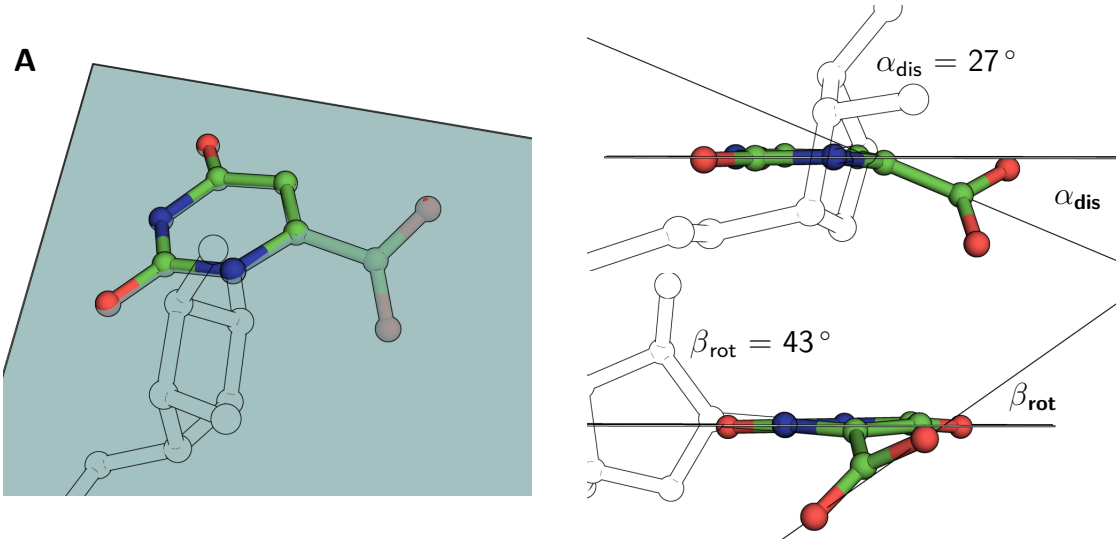


Figure 38: **Out-of-plane distortion and rotation of the substrate's carboxy group in the active site of *h*OMPD$_{314AcK}$.** Decarboxylation observed *in crystallo* after 2 min crystal soaking at 20 °C

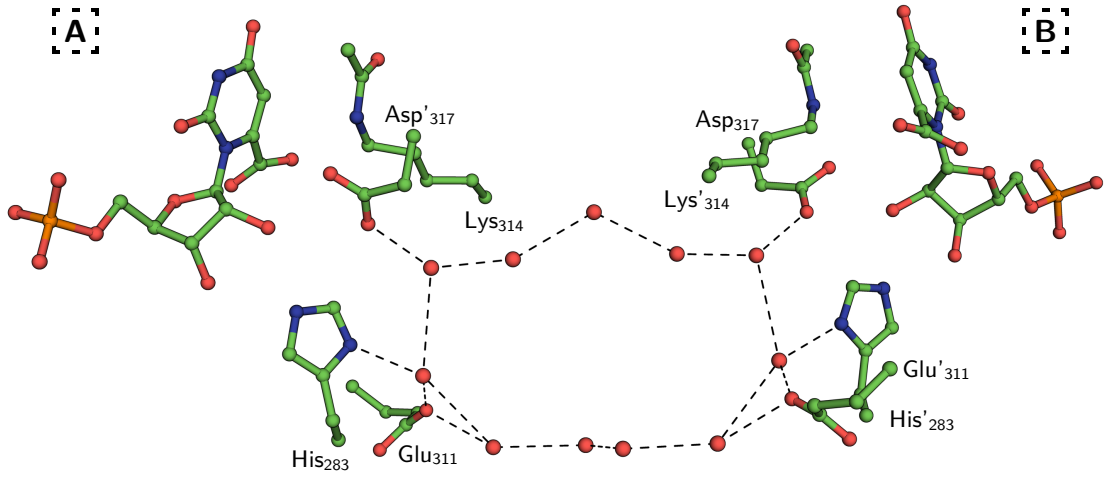


Figure 39: **"Crown-shaped" conformation of the water network interconnecting both active sites of the protein dimer in the *h*OMPD$_{314AcK}$/OMP complex.** Water molecules are indicated as red spheres. Distances below 4 Å are depicted with dashed lines. The dashed boxes indicate the subunit of the protein dimer.

models built and refined to a resolution suitable for all data-sets of 1.3 Å.

The crystal packing was determined to be C222$_1$ with one monomer in the asymmetric unit. The space-group assignment indicates the presence of two identical monomers with a true crystallographic symmetry axis generating the native dimeric protein assembly. Importantly, later time-points of the soaking trials could be indexed in space-group P2$_1$ only, with the catalytically competent protein dimer in the asymmetric unit. Consequently, the C2 crystallographic symmetry is broken leading to the presence of two enzyme subunits in the asymmetric unit. The symmetry of all crystals indexed in C222$_1$ could be expanded and the data refined in space-group P2$_1$ to obtain the protein dimer. In ambiguous cases, both possible crystallographic lattices were used for space-group assignment and data refinement to resolve subtle differences in the subunit architecture. The radiation damage analysis was done in space-group P2$_1$ with the dimer in the asymmetric unit.

The occupancy refinement of the C^6-carboxy group indicates a constant degree of decarboxylation over all measured datasets with an individual occupancy of 95 and 90 % for subunit A and B, respectively (Figure 40). The uridine portion of the ligand was refined to an occupancy of 100 % indicating complete active site saturation. The constant values for the ligand's carboxy group indicate an enzymatic decarboxylation process since the increasing introduction of high energy photons does not lead to further density loss for the carboxy group. The measurable rest activity of *h*OMPD$_{314AcK}$ is 0.002 s^{-1} ($^1/_{(10\,\mathrm{min})}$) and might lead to the observed presence of product in the active site of *h*OMPD$_{314AcK}$. However, the induced radiation damage seems to result in a bond elongation in the active site of subunit B between the substrate's carboxy group and the functional group of Asp$_{312}$. The initial distance in both subunits expands from 2.51 Å and 2.46 Å to 2.54 Å and 2.60 Å in subunit A and B, respectively. The bond length expansion appears to be constrained to the active site and does not manifest in the unit cell constants which remain nearly constant over the radiation series. In order to avoid radiation induced decarboxylation and to obtain radiation dose unaffected bond distance measurements, a dose limit of 0.5 MGy was defined. All datasets were measured with a total dose limit of the defined radiation threshold per data-collection.

Testing different soaking durations and, consequently, varying incubation periods of the protein crystal with the substrate solution, the complete enzymatic decarboxylation reaction was monitored. All collected datasets were refined in space-group P2$_1$ to prevent averaging of the two subunits and a potential masking of a differential catalytic behaviour. The early and late datasets of the soaking time-course were indexed as C222$_1$ in XDS and forced to P2$_1$ for com-

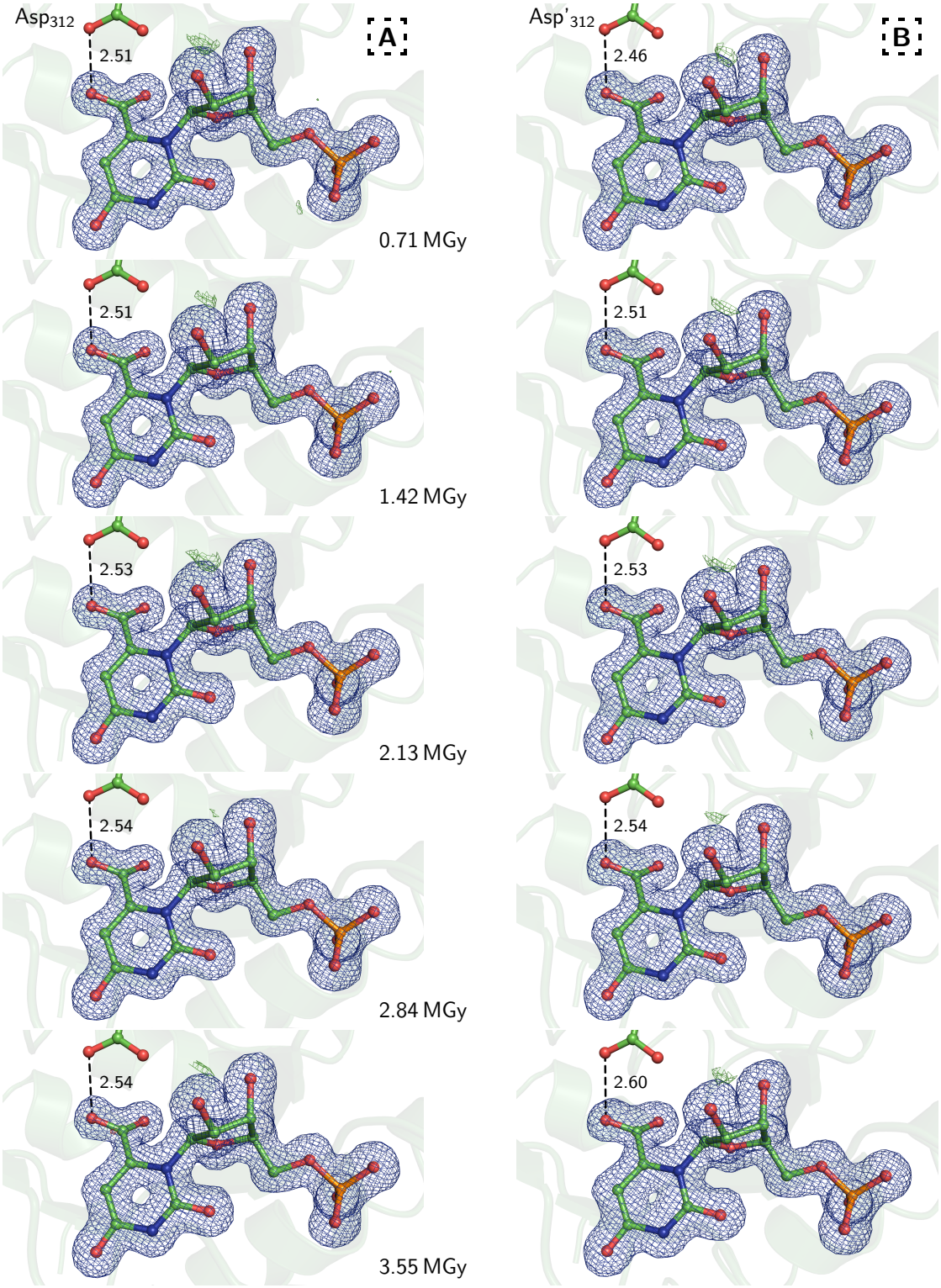


Figure 40: **Radiation induced bond length expansion between the OMP substrate and the catalytic $\mathbf{Asp_{312}}$ of $\mathbf{\textit{h}OMPD_{314AcK}}$.** Dashed boxes indicate the corresponding subunit. Iterative dataset collection of the identical 200 ° segment of a protein crystal. The $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with a contour level of 1.0 and $\pm 3\,\sigma$, respectively. Bond length between the carboxy groups of OMP and Asp_{312} is indicated as a dotted line and measured in Å. 2 min substrate soak at a common resolution of 1.3 A. Occupancy of the C^6-carboxy group refined with 95 and 90 % in subunit A and B, respectively.

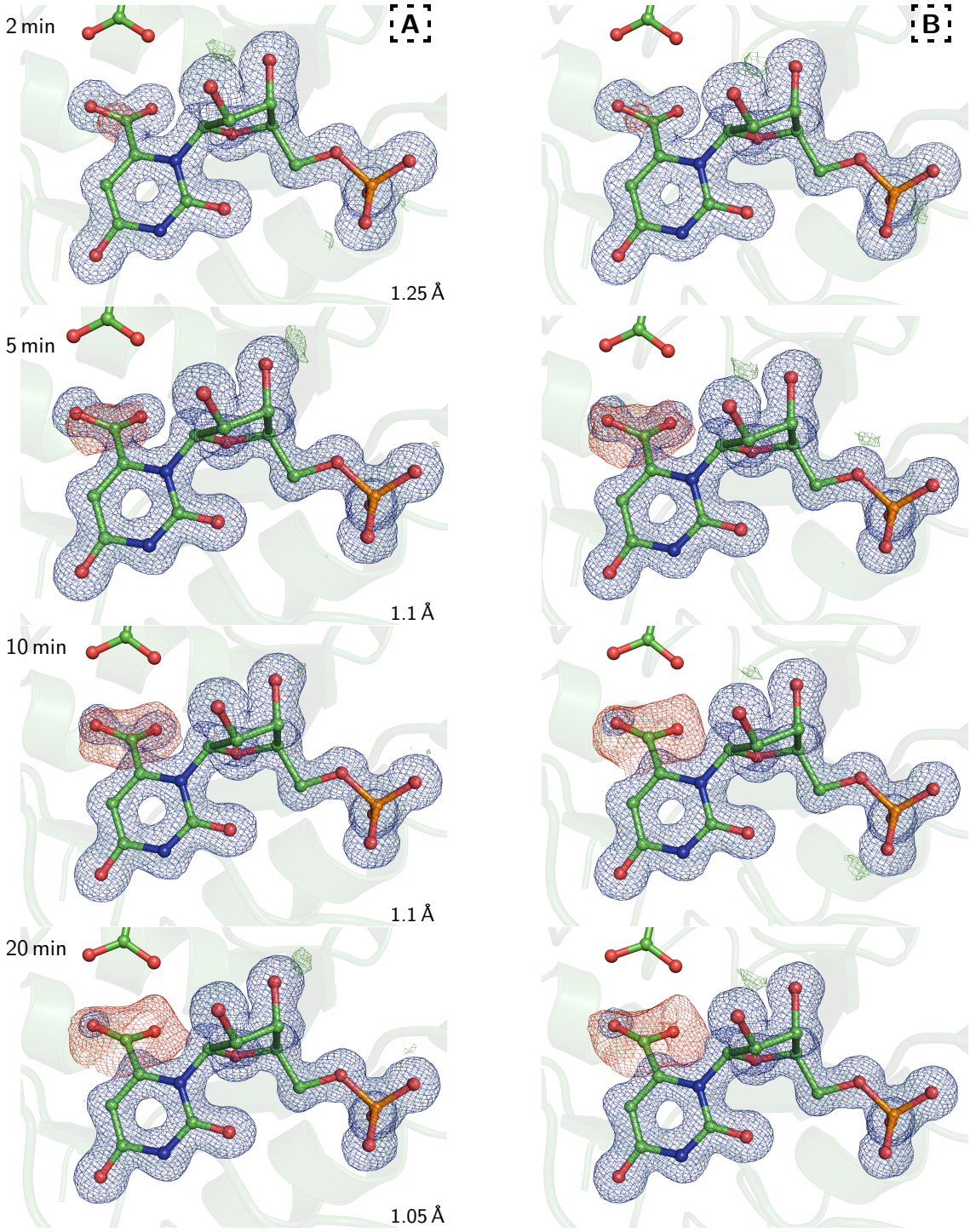


Figure 41: **Time resolved OMP decarboxylation catalysed by hOMPD$_{314AcK}$ *in crystallo*.** Active site view of both subunits of the native protein dimer with emphasis on the reaction substrate and Asp_{312}. Dashed boxes indicate the corresponding subunit. The $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with a contour level of 1.0 and $\pm 3\,\sigma$, respectively. Refinement with the OMP molecule and 100 % occupancy. The resolution of the dataset is indicated in Å.

parative purposes. Crystals were placed in an OMP solution and frozen after 2, 5, 10, 15, 20, 25, 30 and 35 min. After 2 min, full ligand occupancy was observed with a slight negative difference electron density peak on the C^6-carboxy group. As analysed in the radiation series, the loss of the carboxy group signal stems from enzymatic activity and not from radiation induced decarboxylation. The fuctional group of Asp_{312} shows full occupancy over all measured time-points.

The electron density loss of the substrate's carboxy group becomes more pronounced with increasing soaking durations. After 5 min substrate soaking and flash freezing, the negative $m\mathrm{F_o}$-$D\mathrm{F_c}$ difference electron density increases and the $2m\mathrm{F_o}$-$D\mathrm{F_c}$ density starts to vanish. No further structural rearrangements in the active site are observed except for the disappearing carboxy group. Notably, the observed decarboxylation is not identical in the two catalytic sites. In all measured data-sets, one subunit shows a weaker signal for the carboxy group indicating a more rapid decarboxylation. The faster active site is referred to as subunit B in the following sections.

After 10 min of substrate exposure, the $2m\mathrm{F_o}$-$D\mathrm{F_c}$ electron density and the negative difference electron density on the carboxy group of subunit B indicate the complete loss of the functional group. A water molecule occupies a position approx. 1 Å distant from to the carboxy oxygen previously interacting with Asp_{312}. $2m\mathrm{F_o}$-$D\mathrm{F_c}$ electron density can still be observed for the substrate's carboxy group in subunit A. An incubation of a protein crystal in an OMP solution of 20 min leads to the complete decarboxylation of the substrate in both active sites. Again, a water molecule is located in a position close to the vanished carboxy group which is not present in the early datasets complexed with OMP. The sum of the occupancy values of the carboxy group and the newly found water is less than one indicating a slow water diffusion into the crystal. Since the water molecule appears after the enzyme catalysed decarboxylation occurred, a participation in the reaction mechanism is unlikely. Furthermore, even though the enzyme crystals are incubated in a substrate solution during the soaking trials, no exchange of product against substrate seems to occur.

In order to get insights into the reaction mechanism of $h\mathrm{OMPD}_{314\mathrm{AcK}}$ mediated OMP decarboxylation, certain time-points of the time-series were analysed in space-group $P2_1$. The asymmetric unit is composed of the native protein dimer and the X-ray diffraction recorded at a resolution of 1.25 Å. The model was built to represent the experimental data with $R_{\mathrm{work}} = 13.2\,\%$ and $R_{\mathrm{free}} = 16.8\,\%$ (Figure 42). The substrate complex structures obtained from 2 min soaking durations do not show a considerable asymmetry in the active site architecture of the enzyme. Both ligands are bound with 100 % saturation but show a slight negative $m\mathrm{F_o}$-$D\mathrm{F_c}$ difference

electron density peak on the carboxy groups. The functional group of OMP was refined with an occupancy of 0.9 in both subunits. The O-O-distance between the substrate's carboxy group and Asp_{312} is 2.49 Å and 2.45 Å in active site A and B, respectively. Both ligands are likely to form short and strong hydrogen bonds to the protein even though the hydrogen atom is not resolved in the difference density maps. Furthermore, the C-O-distances between the C^7-atom and O^{71} and O^{72} deviate from each other with 1.31 Å and 1.25 Å in subunit A and 1.32 Å and 1.18 Å in active site B. The asymmetry of the carboxy groups and the $2mF_o$-DF_c electron density distribution indicate the presence of a hydroxy and a carbonyl function on C^7. The hydroxy group is likely to be protonated and serves as hydrogen bond acceptor since the carboxy group of Asp_{312} appears ionized.

Notably, the carboxy function of OMP is bent out of the pyrimidine plane generated by N^1, N^3 and C^5 (Figure 43). The observed out-of-plane distortion shows comparable values of 25 ° and 27 ° in subunit A and B, respectively. The C^6-atom is likely to have a sp^2 hybridisation and is stabilized by the aromatic pyrimidine ring. The distortion would result in at least a partial breakage of aromaticity of the ring system. The complete carboxy group is located entirely underneath the pyrimidine plane with O_{71} nearly in plane. Furthermore, the carboxy group is rotated by 44 ° and 50 ° in subunit A and B, respectively.

The time resolved datasets indicate a differential substrate processing in the two active sites of the protein dimer. Subunit B shows a higher product content compared to subunit A. After 4 min of substrate soaking at 20 °C and flash freezing, a substrate occupancy of 95 % and 60 % was observed in subunit A and B, respectively (Figure: 44). Thus, the asymmetric unit is composed of the native protein dimer and the X-ray diffraction recorded at a resolution of 1.1 Å. The model was built to represent the experimental data with $R_{work} = 13\,\%$ and $R_{free} = 16.1\,\%$ The O-O-distance of the substrate's carboxy group and Asp_{312} is 2.44 Å and shorter than observed in subunit B which is 2.5 Å. As observed in the initial substrate complex, the carboxy group of OMP shows a bond length asymmetry with a C-O-distance of 1.31 Å and 1.23 Å in both subunits. The bond length analysis and $2mF_o$-DF_c electron density distribution indicates the presence of a carbonyl function facing the hydrophobic pocket and a hydroxy group interacting with Asp_{312}. A potential hydrogen atom on O^{71} could not be located in the difference Fourier synthesis but is assumed to be present to form a short hydrogen bond. Since the O-O-distance of the carboxy function of OMP and Asp_{312} is shorter than the Van-der-Waals radii, the formation of a hydrogen bond is assumed. The loss of the carboxy group in subunit B is compensated by the accommodation of a water molecule close the the former O^{71} position.

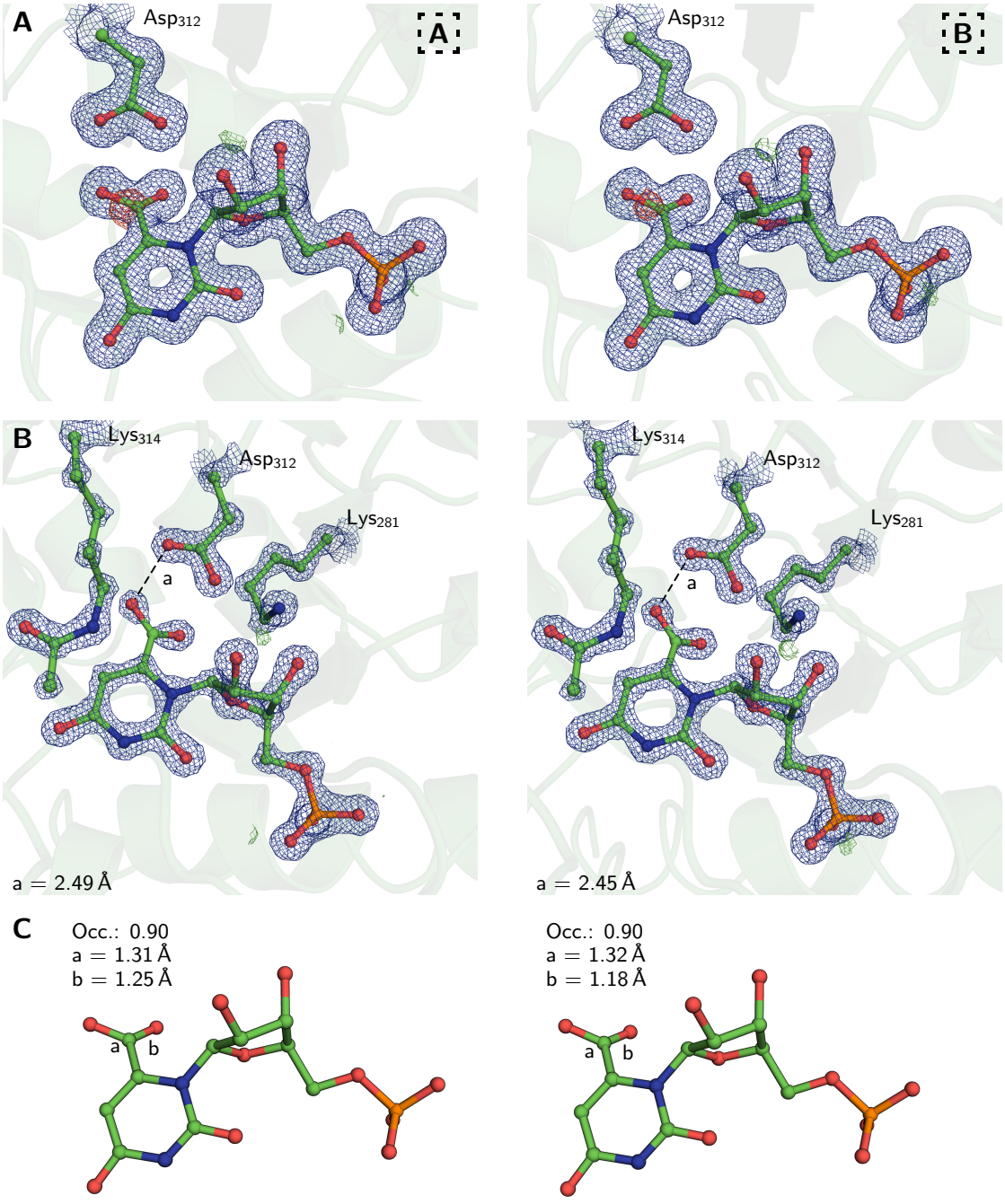


Figure 42: **Crystal structure of hOMPD$_{314AcK}$ in complex with OMP after 2 min soaking time at 20 °C.** Boxed letters A and B indicate the depicted subunit of the protein dimer. A: Substrate positioning and decarboxylation in spacial vicinity to the active site aspartate. $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with a contour level of 1 and $\pm$ 3 σ, respectively. Occupancy of every substrate atom is 1.0. B: Actite site view of substrate and residues from the catalytic tetrad. Occupancies of the OMP carboxylate atoms in subunit B are adjusted to comply with the electron density. The water molecule adopts the position of the former carboxylate group. $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with a contour level of 3 and $\pm$ 3 σ, respectively. C: Refinement results of the indicated input model of the ligand and resulting bond length and occupancy values of the substrate carboxy group.

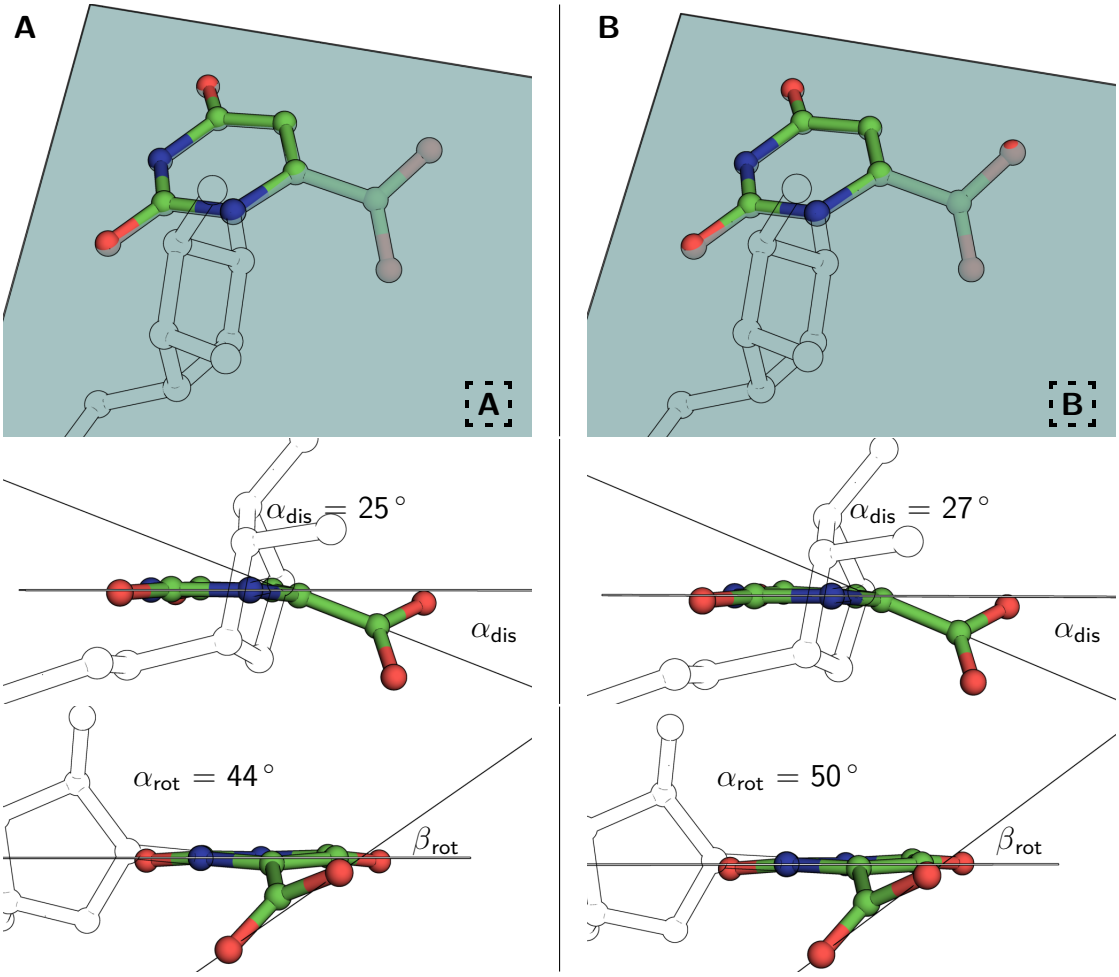


Figure 43: **Out-of-plane distortion and rotation of the substrate's carboxy group in the active site of $hOMPD_{314AcK}$.** OMP substrate complex after 2 min soaking time at 20 °C. Boxed letters A and B indicate the depicted subunit of the protein dimer. The occupancy of the carboxy group is 90 % in both subunits. The phosphoribosyl moiety is indicated with outlines.

The alignment of the obtained substrate and product complex structures revealed a very high degree of similarity. The active site architecture and the tetrad residues show identical positions in both structures. Only Asp_{312} rotates slightly to adopt a position spatially closer to the carboxy group of OMP favouring a productive interaction between the two functional groups. Furthermore the pyrimidine ring of the product UMP shifts slightly upwards towards the modified lysine residue.

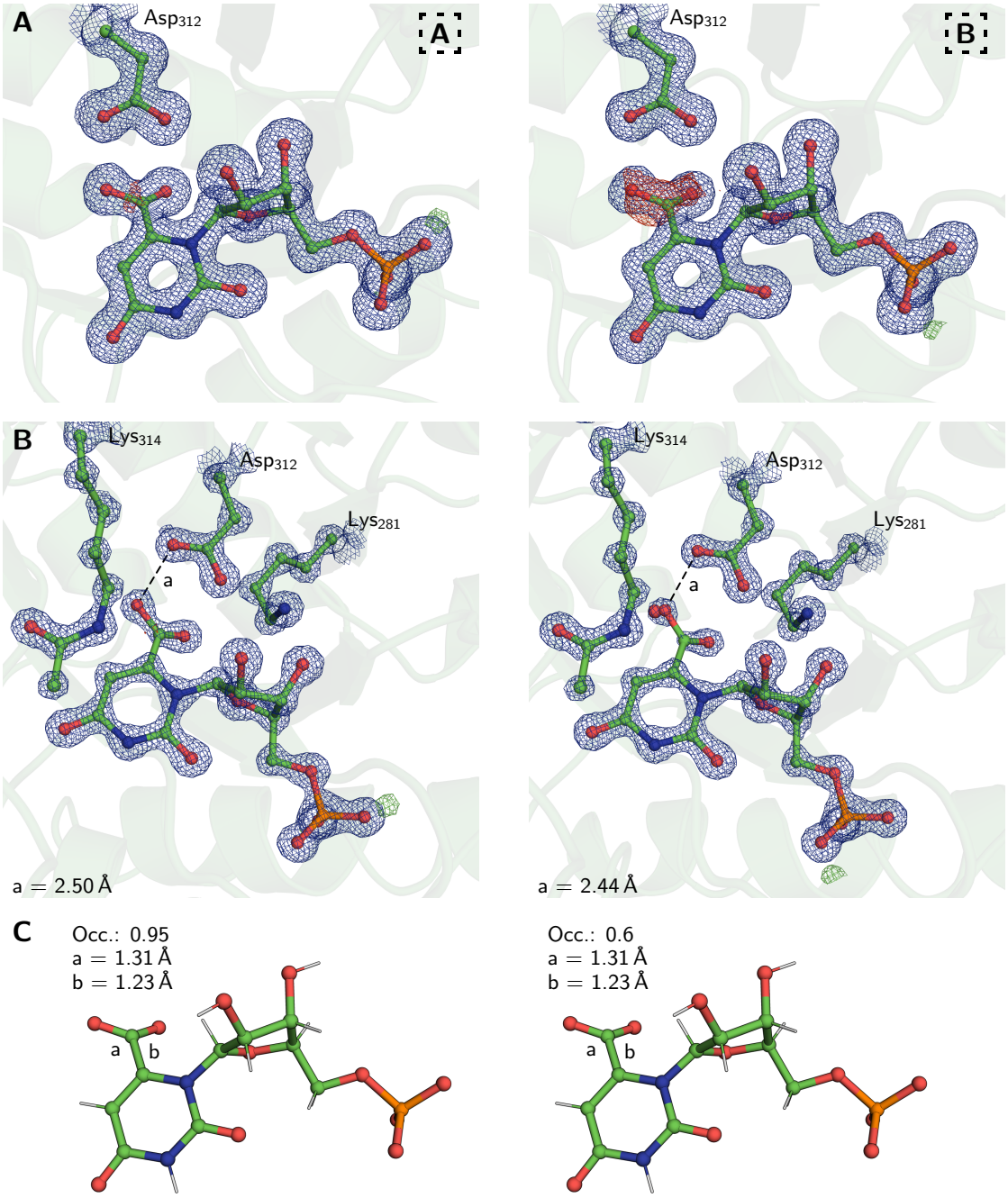


Figure 44: **Differential substrate processing of hOMPD$_{314AcK}$ *in crystallo* after 4 min soaking time at 20 °C.** Boxed letters A and B indicate the depicted subunit of the protein dimer. A: Substrate positioning and decarboxylation in spacial vicinity to the active site aspartate. $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with contour levels of 1 and ± 3 σ, respectively. Occupancy of every substrate atom is 1.0. B: Active site view of substrate and residues from the catalytic tetrade. Occupancies of the OMP carboxylate atoms in subunit B are adjusted to comply with the electron density. The water molecule adopts the position of the former carboxylate group $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with contour levels of 3 and ± 3 σ. C: Refinement results of the indicated input model of the ligand and resulting bond length and occupancy values of the substrate carboxy group.

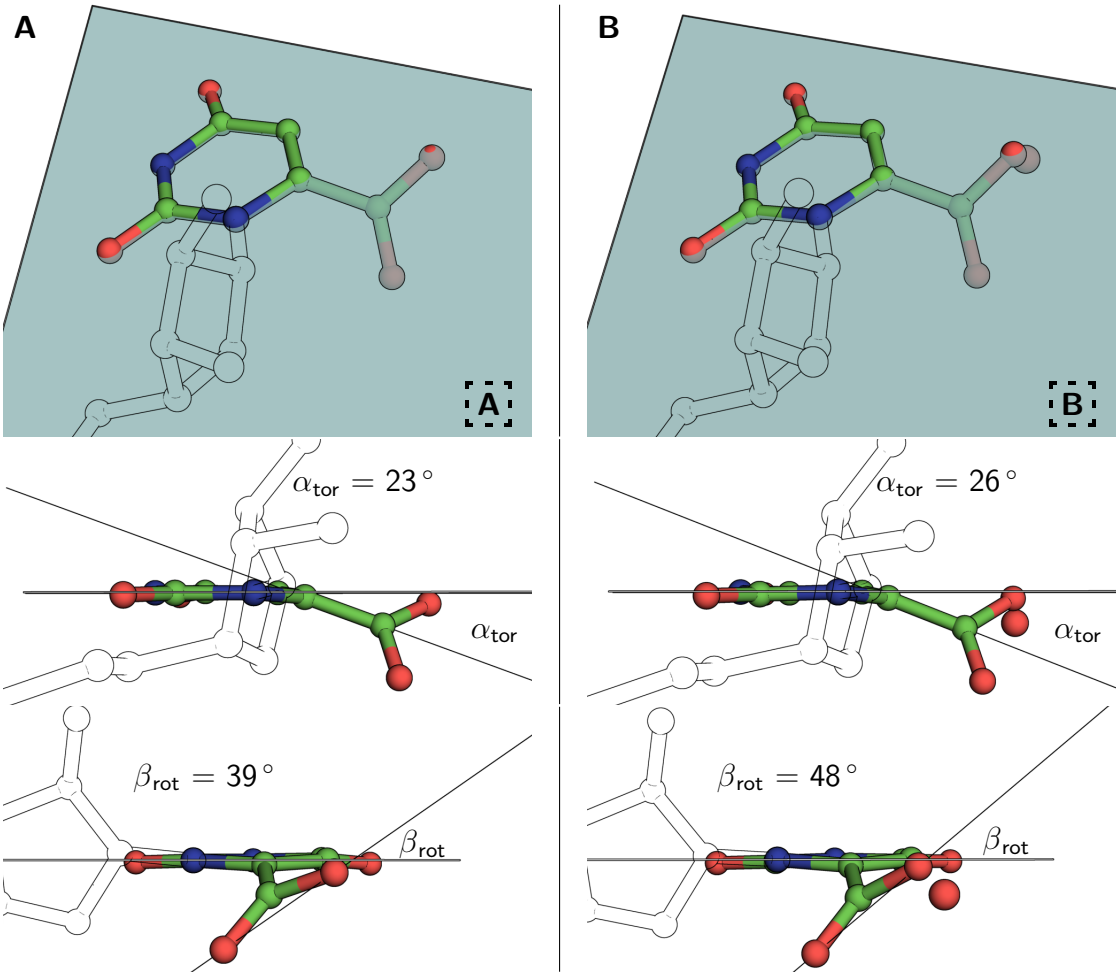


Figure 45: **Out-of-plane distortion and rotation of the substrate's carboxy group in the active site of hOMPD$_{314AcK}$.** Decarboxylation of OMP *in crystallo* after 4 min soaking time at 20 °C. Boxed letters A and B indicate the depicted subunit of the protein dimer. The occupancy of the carboxy group is 90 % in both subunits. The phosphoribosyl moiety is indicated with outlines.

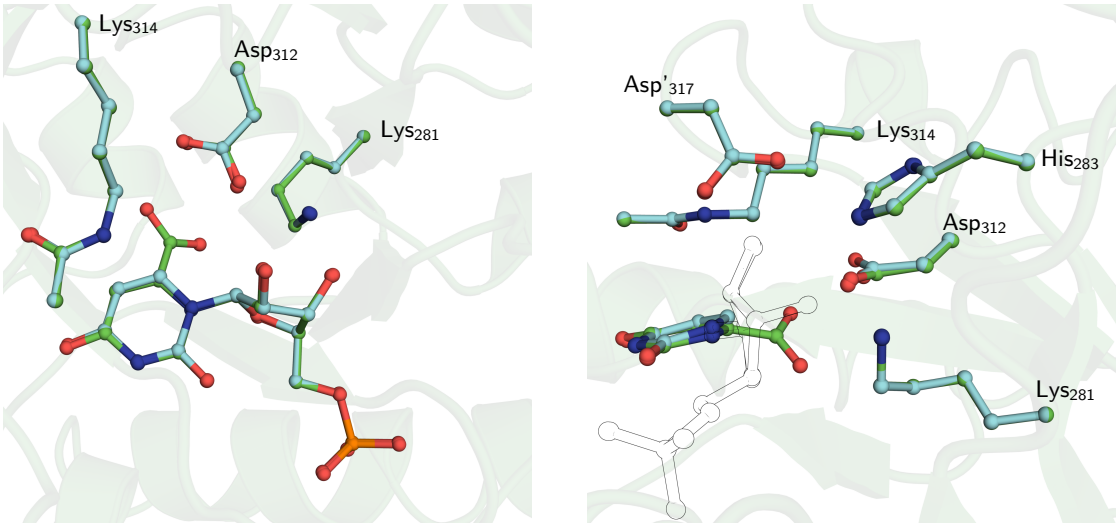


Figure 46: **Structural alignment of the active site of *h*OMPD$_{314AcK}$ in complex with the reaction substrate and product.** The OMP complex originated from a 2 min substrate soak into a resting state crystal. The product complex was obtained after complete *in crystallo* substrate conversion after 30 min. A: Active site overview showing residues of the essential catalytic tetrad. B: Entire catalytic tetrad and His$_{283}$. Ribosylphosphate moiety is indicated with outlines.

8 Discussion: Orotidine 5'-monophosphate decarboxylase

8.1 The active site of *h*OMPD$_{WT}$ favours uncharged species

The proposed reaction mechanism for the decarboxylation of OMP catalysed by *h*OMPD includes the formation of a localized vinyl carbanion after heterolytic bond fission.[73,98,137–139,167,168] The emerging charge after the departure of CO_2 remains on the C^6-atom of the pyrimidine base and is neutralized by protonation from the N^{ϵ}-ammonium group of Lys_{314} acting as catalytic acid. Consequently, as reaction intermediate, the vinyl carbanion has to be stabilized by electrostatic effects within the active site of *h*OMPD to perform efficient catalysis. The very high binding affinity of the non-covalent inhibitors BMP and Aza-UMP was explained by their resemblance to the transition state acting as transition state analogues. In case of BMP, the molecule was assumed to expose the potential C^6-oxyanion to the active site and to tightly interact with the positively charged Lys_{314}. The localized charge of the oxyanion could mimic the vinyl carbanion intermediate and the stable molecule tightly binds to the enzyme. Assuming a comparable binding mode, Aza-UMP was expected to bind in the *syn*-conformation positioning the N^6-atom directing to Lys_{314}. Again, the high electron density was expected to mimic the vinyl carbanion of the reaction trajectory and to be selectively stabilized by the enzyme environment.

However, no indication of a preferred localization of a negative charge in the active site of *h*OMPD could be found. The three potential carbonyl/hydroxy functional groups in the inhibitor BMP appear to be equivalent in the atomic resolution crystal structure. Clearly, the C^6-oxygen does not form the expected localized negatively charged group. Both, the bond length analysis and the electron density distribution do not indicate the formation of a C^6-oxyanion. However, since there is no protonation detected on the C^5-atom forming a sp^3-hybridized carbon, the negative charge has to be distributed within the aromatic system of the pyrimidine ring (Figure 47).

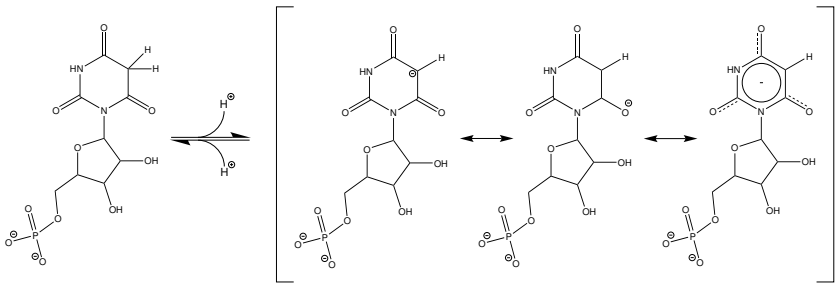


Figure 47: **Ionization states and mesomeric structures of the strong inhibitor BMP.** A negative charge can be stabilized within the pyrimidine ring system.

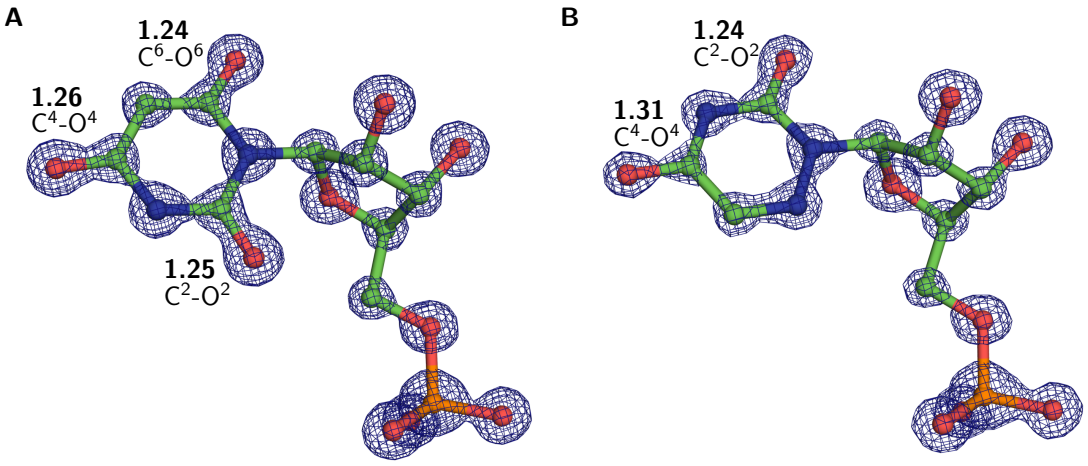


Figure 48: **Comparison of the electron density distribution of BMP and Aza-UMP in the active site on *h*OMPD.** Both inhibitors are depicted in the orientation as observed in the crystal structure. A: BMP is coordinated in the *syn*-conformation in the active site of *h*OMPD. B: Aza-UMP is bound in the anti-conformation. The C^2-C^2-group occupies the position of the C^6-O^6-atom in the BMP-complex. $2mF_o$-DF_c electron density is indicated as blue mesh with an isosurface contouring level of $4\,\sigma$. C-O bond length are indicated in Å with coordinate errors of 0.006 and 0.01 Å for BMP and Aza-UMP, respectively.

Interestingly, the second inhibitor, Aza-UMP, was expected to bind in a comparable way as hypothesised for BMP. It was noted that Aza-UMP could bind in *syn*- and/or *anti*-conformation to expose either the N^6-nitrogen with the lone electron pair to the catalytic tetrad or the C^2-carbonyl. Both conformations could mimic the carbanionic intermediate species. Again, as observed for BMP, Aza-UMP exposes a carbonylic group towards the active site lysine. Furthermore, the potential carbonyl groups of the inhibitor indicates an ionization of the C^4-O^4 group indicated by the bond length analysis and the $2mF_o$-DF_c electron density distribution. Consequently, both inhibitors do not form a localized negative charge to resemble an anionic intermediate close the C^6-position (Figure 48).

8.2 The intrinsic phosphate binding energy is utilized to induce phosphate gripper loop closure of *h*OMPD

The active site architecture of OMPD is well conserved from procaryotes to eucaryotes and the catalytic tetrad plays a distinguished role during catalysis. However, an additional structural element of OMPD mediated catalysis is located distant from the actual binding site. The phosphate gripper loop (423 – 432) folds over the liganded active site forming the catalytically competent caged enzyme/substrate complex.[98,105] Several well studied examples demonstrate the importance of flexible protein regions which are able to rigidify upon substrate encounter. The process

of evolution shaped the active site loops in size and flexibility to adopt to the environmental conditions and to enable efficient catalysis.[104] In 1975, William Jencks introduced the "Circe-effect" concerning enzyme mediated catalysis which proposes "the utilization of strong attractive forces to lure a substrate into a site in which it undergoes an extraordinary transformation of form and structure".[110] Beside the "Circe-effect", Jencks introduced the idea that the binding energy of a distant functional group can be utilized to compensate for an unfavourable interaction in a distinct part of the protein/ligand.[110] Thus, substrate binding could invoke ground state destabilizing effects like the reduction of the conformational, rotational or translational entropy which are compensated by favourable protein-ligand interactions. Furthermore, the binding enthalpy could be compensated by the introduction of steric and conformational strain or electrostatic destabilization and desolvation of the substrate molecule.[110] In the specific case of OMPD, Wu considered the electrostatic repulsion of the substrate's carboxylate group with the ionized Asp_{312} as major driving force for catalysis.[109] The very strong binding of the phosphodianion portion of OMP forces the two negative charges into close spatial proximity leading to the formation of non-polar CO_2.

The idea, or more precisely the extent, of reactant and ground state destabilization (GSD) initially introduced by Jencks is challenged by the concept of transition state stabilization (TSS) in a pre-organized active site and the observation of purely entropically driven catalysis.[86,115,169,170] Arieh Warshel argued that the required ground state destabilization of 32 $^{kcal}/_{mol}$ for catalysis could not be reached with two facing acidic groups in equilibrium. A destabilizing effect of approx. 6 $^{kcal}/_{mol}$ would lead to the protonation of one of the carboxylate groups.[111,171] Thus, OMPD specifically stabilizes the transition state of the reaction pathway to reduce the barrier to product formation. In order to determine the influence of loop closure on the decarboxylation, the group of J.P.Richard extensively studied the contribution of the phosphodianion binding energy to enzyme catalysis (reviewed in Amyes *et al.*, 2016).[129] The approach compares the differences in the second order rate constants of the whole substrate and a substrate lacking the phosphodianion group (Figure 49). In case of the examined enzymes OMPD, TIM and GPDH, the intrinsic phosphodianion binding energy accounts for approx. 12 $^{kcal}/_{mol}$. The addition of phosphite dianion can partially restore the catalytic efficiency and is characterised by a contribution of approx. 6 - 8 $^{kcal}/_{mol}$. The residual 4 - 6 $^{kcal}/_{mol}$ binding energy is utilized to anchor the whole substrate to the active site of the enzyme. Consequently, the majority of the intrinsic phosphodianion binding energy is utilized to drive a conformational change of the enzyme from the open (E_o) to the closed (E_c) complex and not to provoke electrostatic repulsion between the

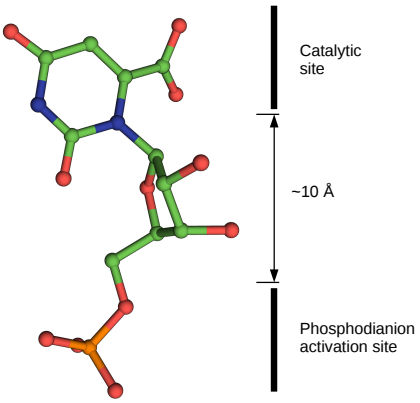


Figure 49: **The intrinsic phosphodianion binding energy is utilized to mediate enzymatic catalysis by inducing loop closure.** The reaction substrate can be divided into a catalytic and a phosphodianion portion separated by approx. 10 Å. Binding of the ligand induces phosphate gripper loop closure and formation of the caged and catalytic active substrate complex (E_o).

two carboxylate groups.[124]

The utilized 6 - 8 $^{kcal}/_{mol}$ intrinsic binding energy is proposed to account for the water extrusion from the active site during the loop closure to form the closed enzyme complex. Ligand binding positions and rigidifies the phosphate gripper loop and stabilizes the catalytically competent closed conformation relative to the open state. The closed conformation is further stabilized by certain hydrophobic and hydrogen bond interactions which are located distant from the phosphate binding site. In *Mt*OMPD, those "remote" residues were found to influence the present population of closed enzyme species and amino acid substitution variants influence solely the K_M-value and not the catalytic competence once the caged complex is formed.[105,122,125]

As proposed by Warshel, even though the pK_a-values of the carboxylic acid groups of OMP and Asp_{312} are relatively low with 2.1 and 3.7 in aqueous solution, a protonation of one of each would prevent a large effect introduced by electrostatic repulsion.[111] However, near atomic resolution crystal structures of OMP in the active site of the enzyme would be needed to clarify a potential proton transfer.

As observed in the crystal structures of the wild type enzyme with the inhibitors BMP and Aza-UMP, the active site appears to prefer uncharged species. Furthermore, the proposed protonation can be observed in the enzyme variant with an catalytically incompetent lysine residue. The acetyllysine substituted enzyme $h\text{OMPD}_{314\text{AcK}}$ unambiguously reveals the presence of the substrate's carboxy group in direct proximity to Asp_{312}. The observed $2mF_o$-DF_c electron density distribution indicates the presence of two unequal carboxy groups. Asp_{312} displays a symmetrical density around the carboxy group indicating the ionization/deprotonation. In contrast to Asp_{312}, the carboxy group of OMP has a localized electron density on the oxygen atom closest to Asp_{312}. Furthermore, Asp_{312} is located closer to the substrate's carboxy group in the $h\text{OMPD}_{314\text{AcK}}$/OMP-complex as it would be expected from the resting state structures.

Thus, the electron density distribution, the O-O-bond distance and the attraction between both functional group indicate the protonation of the substrate carboxy group in the closed enzyme complex.

A severe OMP destabilization generated by the spatial proximity of the two mentioned carboxy groups would favour the strong binding of the product molecule. Since all interactions of the substrate molecule are also present in the product complex, except for the C^6-carboxy group but the electrostatic stress exposed onto the molecule is missing, a higher affinity of the enzyme towards the product would be expected. However, the substrate affinity of *h*OMPD is approx. three orders of magnitude higher than the enzyme affinity towards the product. Substitution of the proximal asparate with an isosteric but neutral asparagine or a glycine would prevent the clash of two negative charges. Notably, these protein variants show a reduced affinity towards the substrate OMP arguing against ground state destabilization.[90]

However, the precise effect of loop closure remains elusive but is proposed to stabilize the vinyl carbanion.[100,122,124] The reduced dielectric constant in the interior of the protein cage might modulate the pK_a-values of the functional groups in the active site or the substrate molecule. In the enzyme triosephosphate isomerase, loop closure leads to an increased pK_a-value of the active site glutamate to serve as the catalytic base for proton abstraction.[102,129]

8.3 Potential Michaelis-complex of *h*OMPD

A major problem of crystallographic studies on the reaction mechanism of enzymes is the inability to obtain specific states of the catalytic trajectory. A very short-lived enzymatic intermediate might be missed simply due to the temporal requirements of the researcher to soak a crystal with a ligand/substrate and to subsequently freeze the treated sample in liquid nitrogen. A certain degree of experience in crystal handling can improve experimental data accessibility but will reach a limit when it comes to soaking trials below a few seconds. Additionally, substrate or ligand binding might induce a conformational change of the protein molecule which might lead to the disintegration of the crystal packing. Even small structural rearrangements can result in increased crystal disorder and reduce diffraction data quality. Consequently, specific reaction intermediates are not accessible directly using macromolecular cryo-crystallography.

In case of *h*OMPD, the substrate complex of the wild type protein has not been obtained yet. Different structural approaches tried to overcome the scarce information content on the Michaelis-complex but were not able to illustrate the active site conformation when the substrate is bound. Two protein variants of bacterial and human OMPD were crystallized and soaked with OMP.

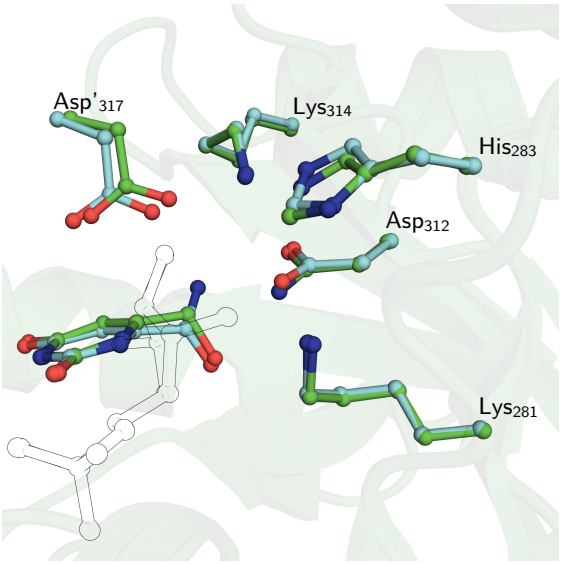


Figure 50: **Extended catalytic tetrad conformation observed after binding of HM-UMP and Amido-UMP** Alignment of the ligand crystal structures of 6-hydroxymethyl-UMP (PDB entry 2QCM, green) in complex with hOMPD$_{\mathrm{D312N}}$ and Amido-UMP (cyan) bound to hOMPD$_{\mathrm{WT}}$. For the sake of simplicity, only dominant conformations of the catalytic tetrad of the Amido-UMP complex and the ligand of the HM-UMP complex are shown. Phosphoribosyl moiety is indicated with outlines.

E.coli OMPD harboured two alanine substitutions for Lys$_{314}$ and Aps$_{312}$ and showed OMP in the active site (numbering according to hOMPD).[118] The human enzyme variant has an aspartate to asparagine substitution on position 312 and also shows the substrate molecule in the active site pocket.[5] Both enzymes have a severely impaired active site architecture and did not lead to additional insights into OMPD catalysis.

In 2008, a structure of the wild type *Plasmodium falciparum* OMPD was published which claimed to show OMP in the active site pocket.[92] Since OMPD was found to be a suitable target for the treatment of malaria, the *P. falciparum* enzyme became of considerable interest.[172] However, the obtained resolution of 2.6 Å is not sufficient and the observed electron density does not indicate the presence of a carboxylate group on position C^6 of the pyrimidine ring. A re-refinement clarified the actual presence of the reaction product in the active site of *Pf*OMPD.[108]

In order to obtain structures of the Michaelis-complex, resting state and product complex crystals were soaked with an OMP solution. Various soaking trials were performed with soaking times ranging from "as fast as possible" (approx. 5 s) to 1 h. However, the few resulting structures showed only the resting state enzyme or the product complex. The difficulties in resolving the substrate complex of the wild type enzyme led to the synthesis of stable substrate analogues with an amido and a thiocarboxamido group on C$_6$ (Kindly synthesized by Matthias Krull and Tobias Schmidt, AG Prof. Dr. Ulf Diederichsen). The nearly isosteric analogues resemble OMP in terms of the spatial requirements but are not negatively charged as it is expected for the carboxylate group of the substrate (pK_{a}(OMP) approx. 0.5 in water).[52] However, both substrate analogues were intended to test the binding mode of a ligand closely resembling OMP.

8.3 Potential Michaelis-complex of hOMPD

As expected, the amino groups of the two substrate analogues Amido-UMP and Thiocarboxamido-UMP bind to the active site of *h*OMPD acting as hydrogen bond acceptors from Asp_{312}. The substrate analogue structures display a previously observed but neglected active site rearrangement and Lys_{314} displacement (Figure 50).[5] As observed in a ligand complex of *h*OMPD with 6-hydroxymethyl-UMP (HM-UMP), the active site Lys_{314} moves into a position referred to as "arching lysine", centred between Asp'_{317}, His_{283} and Asp_{312}.[5] The flipped histidine residue is appended to the hydrogen bond network of the catalytic tetrad to form a conformation introduced as "extended catalytic tetrad". The conformational change upon ligand binding places the tetrad residue Asp'_{317} in a favourable position for proton donation to the C^6-atom of the pyrimidine ring. Considering the formation of the extended catalytic tetrad in a substrate complex, the emerging carbanion originating from the heterolytic bond cleavage of the C^6-C^7-bond of OMP could be quenched by a pre-oriented acid Asp'_{317}. Both, the substrate analogue structures and the HM-UMP complex indicate a potential role of Asp'_{312} as general catalytic acid for the decarboxylation reaction.

The main concern utilizing the designed substrate analogues was considered to be the presence of the amino group which acts as hydrogen bond acceptor for Asp_{312}. In comparison to the substrate's protonation state in solution, the different hydrogen bonding behaviour might lead to a change of the hydrogen bond network in the active site of *h*OMPD. Notably, structures of the OMP complex with the $hOMPD_{314AcK}$ protein variant showed that the substrate itself binds as an hydrogen bond acceptor and forms a short hydrogen bond to Asp_{312}. The binding mode of OMP to $hOMPD_{314AcK}$ indicates that the utilized substrate analogues might bind in the same manner as the natural substrate. The unexpected binding mode of OMP to $hOMPD_{314AcK}$ changes the picture of the substrate analogues and the positioning of the amino group. The interaction of OMP with Asp_{312} very closely resembles the substrate analogue binding mode in the wild type enzyme. Consequently, the observed active site architecture (with the formation of the extended catalytic tetrad) might be the actual conformation of the natural substrate complex. Thus, the Michaelis-complex might be composed of OMP and Asp_{312} of the active site in the conformation observed in the $hOMPD_{314AcK}$ complex and the active site architecture (Asp'_{317}, Lys_{314}, His_{283}, Lys_{281}) observed in the Amido-UMP or TCA-UMP complex (Figure 51).

A possibility to analyse the binding mode of a substrate carboxylate analogue could be bromo- or nitro-derivatives of OMP. These inhibitors would mimic the the negatively charged carboxylate of the substrate OMP even in the active site pocket. However, it is likely to assume that negatively charged ligands might not bind to the active site when a protonation is not possible.

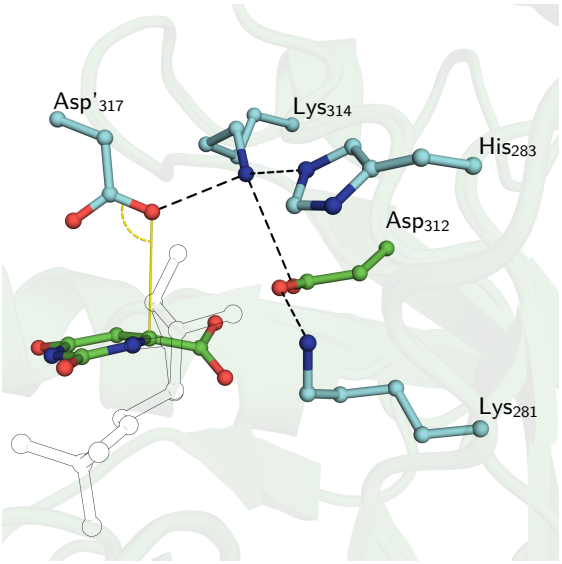


Figure 51: **Hypothetical Michaelis-complex of hOMPD$_{WT}$.** Extended catalytic tetrad of an artificial Michaelis-complex. Structures of the hOMPD$_{314AcK}$ substrate complex (green) and hOMPD$_{WT}$/Amido-UMP complex (cyan) were aligned and merged. Representative amino acids of the potential wild-type substrate complex were chosen to visualize a possible binding mode of OMP. For the sake of simplicity, only dominant conformations of the catalytic tetrad of the Amido-UMP complex are shown. Phosphoribosyl moiety is indicated with outlines.

The negatively charged inhibitors could introduce structural distortions in order to accommodate the nitro- or bromo-group or might not bind at all. However, the general ability to bind neutral, positively charged, negatively charged and bulky ligands was reported previously.[173,174]

8.4 Decarboxylations and the re-association of carbon dioxide

The current reaction mechanism of hOMPD mediated decarboxylation of OMP proposes the formation of a vinyl carbanion with a localized negative charge on the C^6-position of the pyrimidine ring in the transition state.[73,82,175] The considerations concerning the reaction mechanism mainly focus on the formation and stabilization of a carbanionic species in the active site of hOMPD. However, the role and reactivity of the nascent carbon dioxide is mostly neglected.[176] Studies of decarboxylation reactions in solution indicate that CO_2 can act as a very reactive electrophile.[177] The heterolytic bond cleavage leads to the formation of a reactive nucleophile and a strong electrophile in spatial proximity and the forward reaction (release of CO_2) competes with the reverse reaction (re-association of carbon dioxide to the carbanion). In case of hOMPD, a computational study with 1-methylorotate as model compound proposed the decarboxylation of OMP without a barrier for the re-association of carbon dioxide in free solution and in the enzyme active site. The computed potential of mean force (PMF) as a function of C^6-C^7-distance indicates no free energy minimum upon bond rupture distance of a carbon-carbon bond.[109,178] A comparable result for the computed PMF for the decarboxylation of N-methyl picolinate was obtained in water.[179] To favour the forward reaction, an enzyme can decelerate the reverse reaction and prevent the re-association of the product molecules.[57] Consequently, the spatial separation

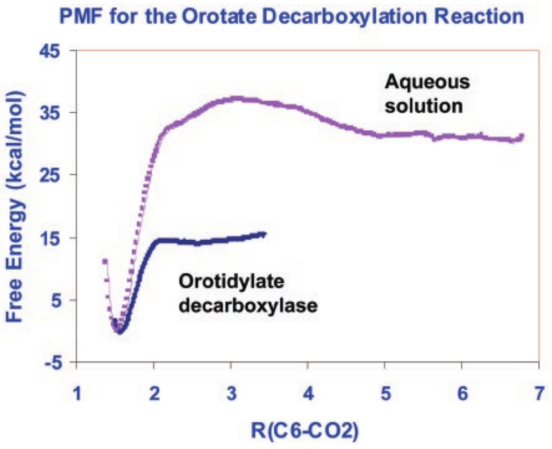


Figure 52: **Calculated potentials of mean force as a function of C^6-C^7-distance of 1-methylorotate.** Free energy profile of the decarboxylation of 1-methylorotate or OMP to 1-methyluracil or UMP and CO_2, respectively. Calculated free in aqueous solution (1-methylorotate) and the active site cavity of OMPD (OMP). Taken from Wu *et al.*, 2000.[109]

of carbon dioxide and the emerging carbanion can lead to a catalytic rate acceleration.

In order to overcome the problems of the separation of reactant molecules in decarboxylation reactions, enzymes evolved pathways to prevent the re-association of the reaction products.[176,180,181] The decarboxylation of trichloroacetate requires the hydration of the carboxylate group to release bicarbonate instead of CO_2.[180] The hydration prior to bond cleavage and the release of bicarbonate can overcome the problems with product separation and CO_2 solvation.

Another decarboxylation strategy which can facilitate the direct release of CO_2 uses the pre-association of a Brønsted acid in a favourable position to donate a proton to the emerging carbanion. The charge of the carbanion from the heterolytic bond cleavage can be quenched by the pre-associated acid which offers a competing pathway to the re-association of CO_2. In benzoylformate decarboxylase (BFD), benzoylformate reacts with the activated cofactor molecule thiamin diphosphate (ThDP) to form α-mandelyl-thiamin (MTh) which decarboxylates to yield 2-(1-hydroxybenzyl) thiamin diphosphate (HbnThDP) in the presence of a catalytic acid to quench the carbanion. Interestingly, in the absence of a catalytic acid, HbnThDP is subject to a very rapid spontaneous reaction that splits the ThDP cofactor into the pyrimidine and thiazolium moieties and irreversibly inactivates the compound. The fragmentation reaction ($10^4\ s^{-1}$) is approx. 100 times faster than the turn over rate of BFD. To prevent the cofactor inactivation BFD most probably utilizes a histidine residue in the active site which is favourably oriented to neutralize the emerging carbanion and prevents cofactor disintegration.[182,183] The pre-association of the Brønsted acid facilitates both, the suppression of the recombination of CO_2 and the carbanion and the prevention of ThDP inactivation. Pre-localization of a catalytic acid to compete with the reverse reaction was estimated to accelerate the decarboxylation rate by a factor of approx. 10^5.[56]

Considering the reaction mechanism of *h*OMPD, the problematic situation of CO_2 formation in direct proximity to the nascent vinyl carbanion could critically interfere with the decarboxylation reaction. A hydration of the reaction intermediate and subsequent release of bicarbonate instead of carbon dioxide is unlikely as it was shown in stopped-flow measurements with *Sc*OMPD.[107] In order to facilitate the diffusion of CO_2 and to prevent the recombination of the reaction products, the emerging carbanion could be stabilized by tautomerism. Protonation of the O^4-atom of the pyrimidine ring could lead to the formation of the carbene tautomer.[74,79] Since the emerging negative charge is distributed over the conjugated ring system, the tautomerization reaction would stabilize the carbanionic intermediate and enable the dissociation of carbon dioxide. However, in the direct proximity of the C^4-carbonyl, no suitable acid is located in a preferable position for protonation.

Applying the concept of the pre-association of a catalytic acid close to the nascent carbanion, two potential residues might facilitate proton donation to compete with the recombination reaction of carbon dioxide. In the conventional catalytic tetrad, Lys_{314} could quench the negative charge. However, the lysine residue is not well oriented for proton transfer. Additionally, in a potential substrate complex, the catalytic tetrad lysine would be tightly packed between the three carboxy groups of Asp'_{317}, Asp_{312} and OMP. The relatively strong basicity of the amino group of Lys_{314} might be further increased due to the neighbouring negative charges. An increased pK_a-value of the triangled lysine residue could impair its functionality as a general catalytic acid. On the contrary, the observed interconnection of both active sites might enable a proton transfer from the second subunit to protonate and neutralize carboxylate groups in subunit A. The altered active site electrostatic environment could also decrease the pK_a-value of Lys_{314} to enhance the acidic character. As observed in the Amido-UMP and TCA-UMP complex structures, the catalytic tetrad can also adopt a previously undescribed conformation. In the extended catalytic tetrad conformation, Lys_{314} is displaced and additionally interacts with His_{283}. Notably, the reoriented Asp'_{317} is located in a favourable position to donate a proton to a potential C^6-carbanion. Since every reaction cycle requires one proton to substitute the carboxy group, the interaction with the acid-base catalyst His_{283} could be beneficial for a proton transfer reaction. The histidine residue is also part of the interconnection between both subunits of the dimer and might mediate proton transfer between both subunits or serve as a proton buffer in its charged state. The actual occurrence and necessity for two potential conformational changes for efficient catalysis as well as how the vinyl carbanion is stabilized without the positive charge in close proximity to the C^6-atom remains elusive.

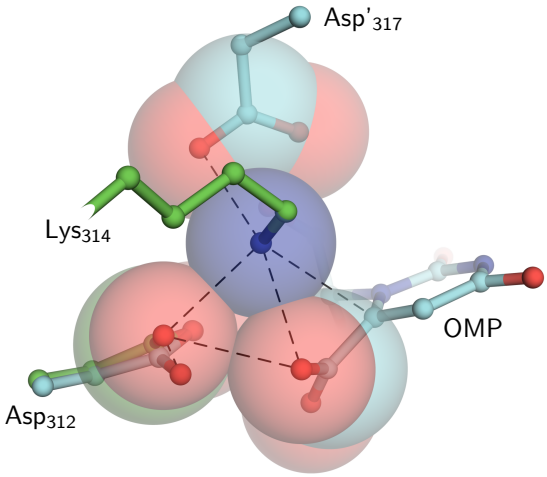


Figure 53: **Hypothetical transition state structure of *h*OMPD.** Residues of the wild type protein (Asp_{312} and Lys_{314}, green) are merged into the active site of $h\text{OMPD}_{314\text{AcK}}$ (cyan) to mimic a potential substrate-encounter complex.

8.5 Conventional decarboxylation proposals

Based on the plethora of experimental and computational data on the reaction mechanism of OMPD from different organisms, a variety of potential decarboxylation strategies emerged. In the last decade of OMPD research, the main proposal claims the formation of a vinyl carbanion which is protonated by Lys_{314} without major conformational changes of the active site.[73,137] The catalytic pocket is considered to be arranged as observed in the BMP inhibitor complex or the UMP product structures which are mostly identical at mediocre resolution. The catalytic Lys_{314} is oriented by hydrogen bonding interactions established with Asp'_{317} and Asp_{312} to point inwards into the active site cleft. Once the substrate binds, the decarboxylation reaction can occur and the emerging carbanion is neutralized by a protonation reaction facilitated by Lys_{314}.[82] However, even though the importance of all four catalytic tetrad residues is well documented, only an explicit role for Lys_{314} is reported which is going beyond residue positioning effects.

Based on the obtained structural and kinetic data, the conventional concept of *h*OMPD mediated decarboxylation of OMP and the function of involved residues can be expanded (Figure 54). Considering an enzymatic active site pocket which requires phosphate gripper loop closure for substrate conversion, the main catalytic tetrad residues Asp'_{317}, Asp_{312} and Lys_{314} can enable decarboxylation and carbanion elimination forming a proton relay system. Once OMP binds to the active site cleft, it triggers the closure of the phosphate gripper loop and forms the short hydrogen bond to Asp_{312} as observed in the $h\text{OMPD}_{314\text{AcK}}$/OMP complex. In order to avoid the electrostatic repulsion between the two carboxylate groups of the substrate and the aspartate residue, Lys_{314} is located in a suitable position for the protonation of the substrate.

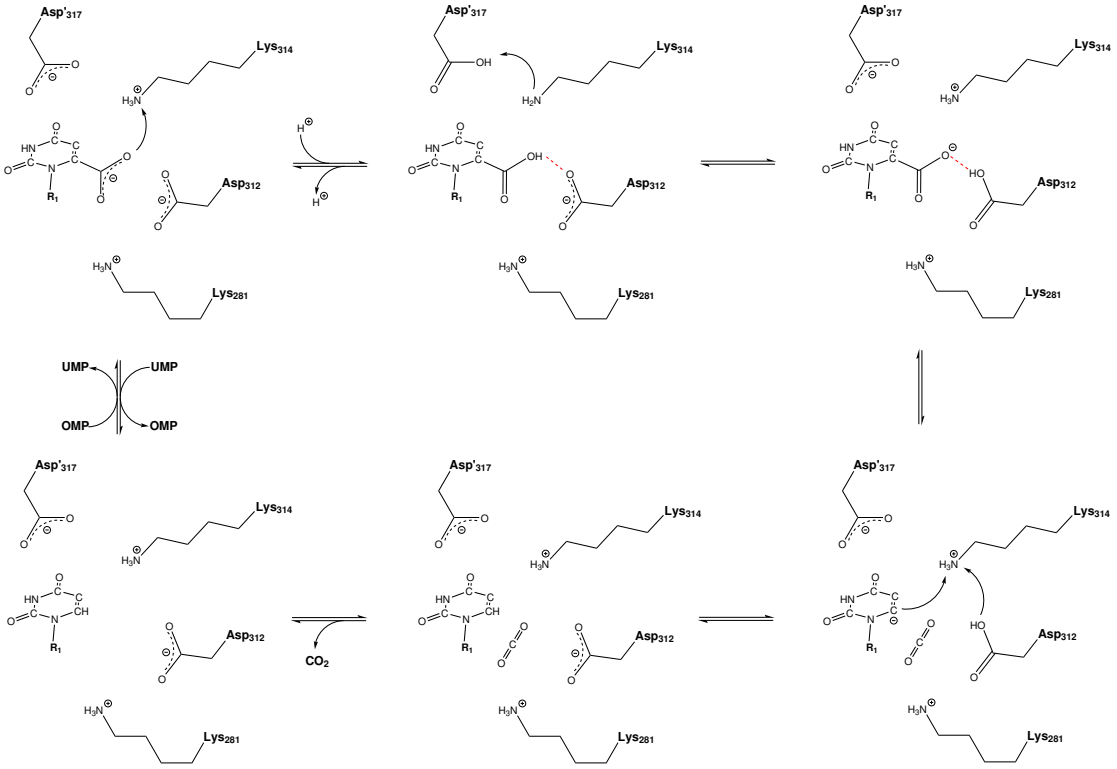


Figure 54: **Potential reaction mechanism of *h*OMPD catalysed decarboxylation of OMP.**

However, to mediate the electrostatic stabilization of the carbanionic reaction intermediate, the positive charge of the N^{ϵ}-atom of Lys_{314} has to be re-established. A suitable candidate for the proton transfer to the catalytic lysine might be Asp'_{317} which is directly connected to the water tunnel bridging both active sites of the protein dimer. The cooperativity of substrate turn-over which was observed in steady-state assays and *in crystallo* indicates a reaction mechanism with two non-independent active sites. Notably, no mechanistic step includes a proton abstraction reaction from a certain source which could be utilized to drive a proton shuffling mechanism between the neighbouring subunits. Since the enzyme displays the highest activity at pH 8.0, a possible proton source could be the binding of the the phosphate mono-anion of the substrate (pK_{a2} approx. 7.2, phosphate mono-anion in water). As observed in the high resolution crystal structures of BMP and UMP, once the active site forms the closed catalytic complex, the di-anion is observed in the active site pocket. How a proton could be transferred to the reactive part of the substrate remains unknown.

However, a proton transfer between the two active sites of the protein dimer could also utilize the asynchronous decarboxylation of both subunits in the protein dimer. An open subunit with a solvent exposed Asp_{317} could serve as proton source and initiate the transfer via the interconnecting water tunnel network. The second subunit considered as being substrate-bound and in

the closed and active conformation could serve as proton acceptor and enable the neutralization of the carbanion intermediate emerging from heterolytic bond cleavage.

Asp'$_{317}$ is located in a well suited position to mediate proton transfer since it directly interacts with the catalytic lysine. If the proton transfer reaction is established via Asp'$_{317}$, it could directly transfer the positive charge to Lys$_{314}$ to form the required N$^{\epsilon}$-ammonium group for catalysis. Thus, the active site of the enzyme is in the decarboxylation competent conformation once the alternate charge network of the catalytic tetrad is re-established. The short hydrogen bond between the substrate and Asp$_{312}$ might enable the proton abstraction from the substrate and transfer to the enzyme. Subsequently, the C^6-C^7-bond of substrate's carboxylate group can be cleaved with a stabilization of the vinyl-carbanion by Lys$_{314}$. To enable product separation, the catalytic lysine can quench the carbanion acting as general acid to obtain the reaction products UMP and CO_2. Due to the basic nature of the amino-group, Lys$_{314}$ could be protonated simultaneously by Asp$_{312}$ in a proton relay system.

Since Asp'$_{317}$ would be a better acid compared to the lysine residue, it could also neutralize the carbanion directly. However, the mechanism would require an additional proton to generate the aspartic acid residue and the N$^{\epsilon}$-ammonium group of Lys$_{314}$. The proton source for the reaction remains elusive.

8.6 A short hydrogen bond is formed between OMP and *h*OMPD

In order to analyse the potential Michaelis-complex structure of OMPD, the substrate analogues Amido-UMP and Thiocarboxamido-UMP were synthesised by the Department of Organic and Biomolecular Chemistry of Prof. Dr. Diederichsen. The chemical compounds are intended to represent stable substrate analogues with a preferably isosteric and isoelectronic functional group substituting the labile carboxy group of OMP in the active site of OMPD. Both inhibitors have an amino group which might serve as electron bond acceptor and a carbonyl or thiocarbonyl group which can serve as hydrogen bond donor. The duality of the functional group of the inhibitors is not expected for OMP since the pKa-value of the carboxy group of OMP is approx. 0.5 and ionized at neutral pH-values. However, both inhibitors have a binding affinity toward the enzyme comparable or higher than the reaction product. So far, no structural data on the binding mode of potential substrate analogues is present.

The obtained crystal structures of Amido-UMP and Thiocarboxamido-UMP unambiguously show the ligand bound the active site of hOMPD$_{\mathrm{WT}}$. The potential substrate analogues bind in a manner comparable to the transition state mimics BMP and Aza-UMP. The observed *syn*-

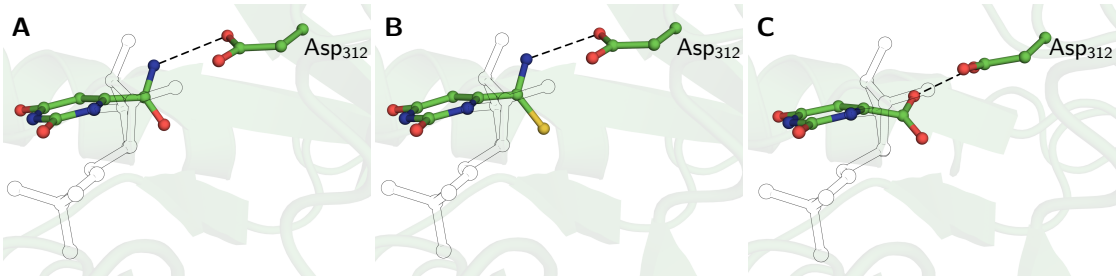


Figure 55: **The substrate analogues and the real substrate OMP serve as hydrogen bond acceptor in the active site of OMPD.** Active site view with the ligand represented as ball and stick model. The ribosylphosphate moiety indicated with outlines. A & B: Inhibitor complex structures of Amido-UMP and Thiocarboxamido-UMP, respectively. C: hOMPD$_{314AcK}$ substrate complex structure.

conformation of the nucleoside monophosphates with bulky C^6-substituents represents the dominant conformational species in aqueous solution.[154,184] The structures of both molecule complexes show a binding mode with the amino group in hydrogen bonding distance to the carboxylate of Asp$_{312}$ (Figure 55). They likely function as hydrogen bond acceptors for Asp$_{312}$ which is ionized in all obtained high resolution structures. Interestingly, a very similar organisation of OMP in the active site of the lysine variant hOMPD$_{314AcK}$ was observed. The substrate's carboxy group is placed in direct proximity to Asp$_{312}$. The functional group of the aspartate residue rotates slightly towards the OMP and the O-O-distance is reduced compared to the N-O-distances of the substrate analogues. The very short distance and the heterogenous electron density distribution of the substrate's carboxy group clearly indicates the protonation of OMP and the existence of a charge assisted short hydrogen bond (COOH-$^-$OOC).[185] The protonation of the substrate's carboxy group is further corroborated by the measured C-O-distances that clearly deviate from 1.26 Å which would be expected for a carboxylate group due to resonance stabilization.[186] The measured distances indicate the presence of a carbonylic function and a hydroxy group which is interacting with the Asp$_{312}$ carboxylate group.

The function of OMP's carboxy group as hydrogen bond donor is further corroborated by observation of a water molecule in the product complex structures of hOMPD$_{WT}$ and hOMPD$_{314AcK}$. The water molecule occupies a position in close proximity to the substrate's hydroxy group interacting with Asp$_{312}$ in the hOMPD$_{314AcK}$/OMP complex and is likely to mimic the hydrogen bond donor. Since the water molecule is not supposed to form a short hydrogen bond, its position is shifted slightly following the displacement of Asp$_{312}$ introduced with the lysine substitution. The positioning of the water molecule in the wild type complex might indicate a comparable substrate binding mode as observed in the hOMPD$_{314AcK}$/OMP complex also in the native protein.

The productive interaction between the two carboxy groups falsifies the proposed electrostatic

destabilization proposed after the advent of the first protein structures. Additionally, it explains the higher binding affinity of the substrate compared to the reaction product and the weak binding affinity of aspartate variants with a Asp_{312} substitution to neutral amino acids. However, the residue acting as catalytic acid for the protonation reaction is unknown. Assuming a proton transfer after the formation of the caged enzyme/substrate complex in the wild type active context, only two lysine residues could facilitate the donation of a proton. Lys_{281} and Lys_{314} are likely to be positioned on both termini of the carboxy group after substrate encounter. Notably, no protein structure is available with an intact active site and substrate bound in the reactive site pocket. In a simple approach, the structures of the $h\text{OMPD}_{\text{WT}}$ pseudo resting state and the $h\text{OMPD}_{314\text{AcK}}$/OMP complex were aligned and the coordinates of the wild type Lys_{314} modelled into the acetyllysine truncated substrate complex structure (Figure 56). The N^{ϵ}-atom of Lys_{314} tightly fits into a triangular space directly between the three carboxy groups of Asp_{312}, Asp'$_{317}$ and the substrate molecule. To prevent the electrostatic repulsion of the two carboxylate groups, Lys_{314} might act as Brønsted-Lowry acid and donate a proton to the carboxylate group of the substrate. The angle of the ammonium group of Lys_{314} towards the carboxylate group of 130 ° would allow the proton abstraction but would require a rotation of the ammonium hydrogen atoms. No water molecules are located in the OMP bound state to facilitate proton donation.

Arieh Warshel and others used a theoretical approach including computer simulations and general energy assumptions to analyse OMPD mediated catalysis.[86,111,171] The reacting system was composed of an orotate molecule oriented as observed in crystal structures and in close proximity to the ammonium group of the active site lysine.[109] After heterolytic C^6-C^7 bond rupture and C^6-carbanion formation, the transition state and the emerging negative charge is stabilized by an increased dipole moment and neutralized by proton transfer. The mechanism assumes the catalytic lysine to act as catalytic general acid to complete the decarboxylation reaction. Based on the obtained crystal structures, and the proposed protonation of the substrate in the encounter complex by Lys_{314}, the previously positively charged ammonium group is neutral by nature after proton abstraction. To comply with the productive dipole interaction mechanism, the lysine residue needs to be protonated from a currently unknown source. A potential residue to replenish the lysine with a proton to regenerate the positive ammonium group might be His_{283}. In the $h\text{OMPD}_{314\text{AcK}}$/OMP complex, the histidine residue is approx. 6 Å distant from the N^{ϵ}-atom of the catalytic lysine. However, in the wild type complex, the catalytic lysine residue displays a high degree of flexibility and can be located in hydrogen bonding distance to His_{283} suitable for proton transfer. Additionally, the water tunnel connecting the two active sites of the

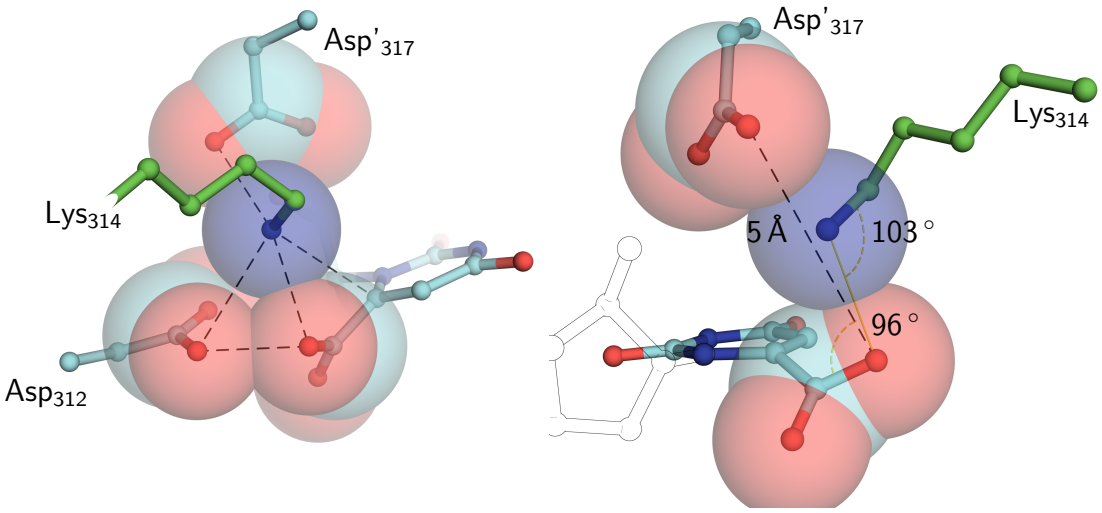


Figure 56: **Potential lysine mediated protonation of the carboxylate group of OMP** Merged structure of the active site Lys_{314} of the resting state complex (green) into the hOMPD$_{314AcK}$ substrate-complex structure (light blue). The acetyllysine residue is omitted and replaced by the resting state lysine. Specific Van-der-Waals radii are indicated with transparent spheres and the ribosylphosphate moiety with outlines.

protein dimer interacts directly with the histidine residue and might also play a certain role to recharge the catalytic lysine residue with a positive charge.

Studies on the H/D-exchange on position C^6 of the pyrimidine ring of the reaction product UMP and small model compounds in the active site of the protein and in aqueous solution, respectively, indicate the general role of the catalytic lysine to act as an acidic catalyst.[137,139] The emerging vinyl carbanion of C^6-H deprotonation is proposed to be generated and neutralized by the catalytic lysine residue which is held in place by hydrogen bonds to Asp'_{317} and Asp_{312}. The orientation of the N^{ϵ}-ammonium group would be restricted in movement and suitable for proton abstraction/donation.

Since the protonated species of OMP in the active site of hOMPD was observed in the acetyllysine substitution variant which is not capable of acid/base catalysis by AcK_{314}, the proton origin in the specific substrate complex can not be determined. A compensating proton source might serve as catalytic acid to prevent the electrostatic stress and reactant state destabilization. A potential candidate could be Lys_{281} which is approx. 4 Å distant from the substrate's carboxy group.

The formation of a short strong hydrogen bond between OMP and Asp_{312} imposes the problem of the decarboxylation of a protonated carboxy group. In the static crystal structure of hOMPD$_{314AcK}$ in complex with OMP, the protonated substrate would be stabilized since CO_2 is not accessible as leaving group. In order to facilitate decarboxylation, the proton has to be re-

moved from the functional group of OMP. In the equlibrium of a short hydrogen bond, a certain fraction of protein molecules might be protonated on Asp_{312} with the carboxylate group on the substrate which would be the competent fraction for catalysis. Since the pK_a of the substrate's carboxylate group is approx. 0.4 (1-methyl orotate) in solution, the acidity of OMP or Asp_{312} has to be severely reduced or increased in the protein active site, respectively.[52] However, in the substrate-complex structure, the orotate moiety is protonated, a compound which in solution is more acidic than aspartic acid.

Corroborating the importance of hydrogen bond formation, a protein variant with an asparagine substitution for Asp_{312} in *Methanothermobacter thermautotrophicus* OMPD (D70N in *Mt*) displays a severe deficiency in OMP decarboxylation but no influence of C^6-deuterium exchange which might result from the inability to abstract the proton from the OMP's carboxy group leading to a mechanism of substrate stabilization in the active site.[90] Furthermore, the productive interaction between the substrate's carboxy group and Asp_{312} is corroborated by the loss of binding affinity of OMP towards an enzyme variant lacking the tetrad aspartate.[106] If electrostatic repulsion between the two functional groups would lead to decarboxylation, a variant lacking the repulsive group would bind the substrate even tighter compared to the wild-type protein.

Importantly, the hydrogen bonding interaction between substrate and protein was observed in the acetyllysine protein variant. The $h\text{OMPD}_{314\text{AcK}}$/OMP complex clearly shows the asymmetry of the substrate's carboxy group with the carbonyl facing the hydrophobic pocket and the hydroxy group interacting with Asp_{312}. The side chain of Asp_{312} displays equal C-O-distances and a symmetric $2mF_o$-DF_c electron density distribution over the carboxylate group indicating ionization. The very short O-O-distance between the carboxy groups would allow the formation of a low-barrier hydrogen bond (LBHB).[162–164,187] LBHBs are characterized by the short interaction distance and matching pK_a-values of the involved carboxy groups to evenly share a proton. They were described for different enzymatic systems including peptidases and observed by the low-field proton NMR shift of 18-20 ppm.[164] A regular hydrogen bond could be modified to form a strong LBHB during the reaction progression to specifically stabilize the catalytic transition state. However, the pK_a-values of both carboxy groups in the $h\text{OMPD}_{314\text{AcK}}$/OMP complex appear not to be matching since the protonation position can be indirectly assigned to the substrate. The observed interaction represents a short-strong hydrogen bond (SSHB) and not a LBHB.[185] Since the structure was obtained in the context of the acetyllysine substitution, the hydrogen bonding behaviour might be different in the native substrate complex. The vari-

ant lacks the positive charge of the N^{ϵ}-ammonium group of Lys_{314} in close proximity to both functional groups. The presence of the positive charge of the lysine might equalize both pK_a-values to from a LBHB in the wild type complex. LBHB formation could be accompanied with a reduced barrier for proton abstraction from the substrate and transfer to Asp_{312} facilitating decarboxylation.

In order to further analyse the binding mode of the substrate and to clarify if the substrate binds as a hydrogen bond acceptor or donor, substrate analogues with an isosteric and negatively charged carboxylate mimic would be required. Based on the obtained structural information, the 6-dithioacid- or 6-nitro-derivatives of UMP would be of high interest as applied in comparable studies.[61,188] Unfortunately, they appear not to be synthetically accessible or to unstable to be used for crystallographic investigations. However, if the short-strong hydrogen bond between substrate and enzyme is formed in the native protein, a substrate analogue with a sustainable negative charge on the functional group would be expected to poorly bind to the active site of *h*OMPD.

8.7 Structural implications of the potential Michaelis-complex

The obtained substrate-analogue structures showed an active-site conformation which was previously observed but not recognized as potential catalytic rearrangement. The conformation referred to as "extended catalytic tetrad" is characterized by a Lys_{314} reorganization towards the interface of the protein dimer and the His_{283} flip to interact with Lys_{314} (Figure 57). The catalytic tetrad reside Asp'_{317} is displaced and positioned directly above the pyrimidine plane and in suitable position to donate a proton to an emerging C^6-vinyl carbanion. However, $2mF_o$-DF_c electron density distribution and C-O-distance measurements indicate the presence of a carboxylate group. Thus, in the observed crystal structures of Amido-UMP and TCA-UMP, the function of Asp'_{317} as general acid during catalysis is only assumed based on the side chain displacement induced by Lys_{314}. If the aspartate residue acts as the actual proton donor to neutralize the negative charge from heterolytic bond cleavage, the protonation of Asp'_{317} itself would be very transient and only a small fraction can serve as catalytic acid. On the contrary, the aspartate is directly interacting with the water channel in the conventional catalytic tetrad and indirectly in the extended conformation and could obtain a proton shuffled between the two active sites. The role of Asp'_{317} as catalytic acid might explain the observed cooperative behaviour during substrate conversion which was observed in the steady-state assay and *in crystallo*.

In *Sc*OMPD, a genetically introduced alanine substitutions of the corresponding catalytic

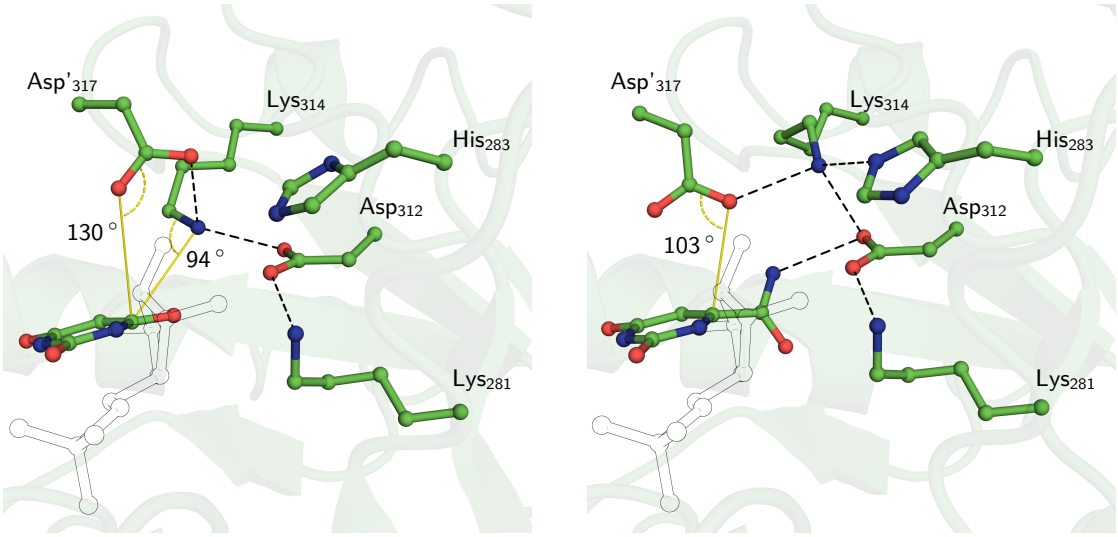


Figure 57: **Traditional and extended catalytic tetrad in OMPD catalysis** A: Conventional catalytic tetrad composed of Lys_{281}, Asp_{312}, Lys_{314} and Asp'_{317} observed in the hOMPD$_{WT}$/BMP complex structure. Lys_{314} is centred between Asp'_{317}, Asp_{312} and the inhibitor's carbonyl group. B: Extended catalytic tetrad observed in the Amido-UMP complex structure of hOMPD$_{WT}$. Only the dominant conformations of the depicted residues are shown. His_{283} is flipped to append to the catalytic tetrad in which Lys_{314} adopts the "arching" conformation and displaces Asp'_{317}. An additional hydrogen bond might be formed to the amide carboxyl of Asp_{312} (not shown). The angles of potential proton donors for the emerging C^6-carbanion to quench the negative charge are indicated in yellow. Potential hydrogen bonds are indicated with black dashed lines.

tetrad residues Asp'_{317} ($\geq 1.9 \times 10^5$ fold), Lys_{314} (5×10^5), Asp_{312} ($\geq 1.3 \times 10^5$) and Lys_{281} ($\geq 4.7 \times 10^2$) result in a tremendous impairment of catalytic competence.[106] The sole role of Asp'_{317} to orient Lys_{314} for proton transfer according to the conventional reaction mechanism proposal can not explain the severe loss in activity of the variant. A function as acid-base catalyst during the reaction progression appears likely. Notably, in the introduced extended catalytic tetrad hypothesis, the importance of Lys_{314} can not clearly be attributed to a specific function except for the positioning of Asp'_{317} in a suitable position for proton donation and the potential transfer of a proton to the aspartate itself. Lys_{314} could serve as a member in a proton-relay system to transfer a proton from the substrate's carboxylic acid group via Asp_{312} to Asp'_{317} to facilitate carbanion quenching (Figure 60).

The role of Lys_{314} to act as the catalytic acid to neutralize the emerging negative charge after decarboxylation was corroborated by C^6-H/D exchange experiments in deuterated water. OMPD also catalyses the hydrogen/deuterium exchange reaction of the reaction product UMP which indicates of the formation of a carbanionic species in the reaction trajectory. The pK_a-suppression for the of C^6-deprotonation was determined to be 10 pK_a-units.[139] The pK_a of UMP in the active site of *Sc*OMPD was determined as ≤ 22 and was compared to the pK_a of 1,3-dimethyluracil in aqueous solution (approx. 30-34). In the H/D exchange experiments, Lys_{314} might be the acid-

base catalyst since it is closest to the C^6-position in the catalytic tetrad observed in the product complex. However, if the substrate complex reorients to form the extended catalytic tetrad, the conformational changes alter the active site architecture and the acid-base catalyst could be Asp'$_{317}$. Additionally, the lysine residue is supposed to be protonated based on structural data and computational analysis.[111] Lys$_{314}$ is not well suited to catalyse deuterium exchange without an additional catalytic base to abstract the proton. Consequently, the proton source to quench the emerging carbanion could be different in the UMP H/D-exchange experiments and OMP decarboxylation or it could be facilitated by Asp'$_{317}$. The high resolution crystal structures of the *h*OMPD$_{WT}$/product complex show a certain degree of flexibility of the catalytic tetrad residues and Asp'$_{317}$ is also determined in a position observed in the substrate analogue complex structures and the extended catalytic tetrad. The previously undetected movement positions the aspartate in close proximity to the C^6-hydrogen and could facilitate proton abstraction. However, the angle for the abstraction of an in-plane (sp^2)-proton is not optimal.

Additionally, the observed catalytic cooperativity might also originate from the formation of the extended catalytic tetrad directly. Considering the extended catalytic tetrade formation in the wild-type protein after substrate binding, the active site rearrangements in the first/liganded subunit might interfere with substrate binding or conversion in the second subunit of the protein dimer. Since Lys$_{314}$ integrates into the interconnecting water tunnel, it could interfere with a potential proton shuffling mechanism. Interestingly, both substrate analogues are modelled with an active site occupancy of 90 %. Even though the reported affinities towards OMPD deviate by a factor of approx. 5.8×10^5, both ligands show the same active site occupancy.[64] The specific dissociaton constants of 6×10^{-4} M and 3.5×10^{-9} M for Amido-UMP and TCA-UMP, respectively, should be sufficiently low for complete saturation in the soaking trials (Ligand concentrations of Amido-UMP and TCA-UMP were identical in the soaking solutions). Careful analysis of the introduced radiation damage during crystallographic data collection and precise dose calibration exclude the possibility of photon induced loss of the C_6-substituents.[165,166,189] A possible explanation for the incomplete active site saturation could be the formation of the extended catalytic tetrad after binding of the substrate analogue to the first catalytic pocket. The conformational rearrangement could alter and reduce the affinity of the second binding site which might not completely saturate at the ligand concentrations in the crystal soaking conditions. The applied crystallographic symmetry could lead to an averaged partial ligand occupation in the crystallographic data and the corresponding model. A cooperative binding model is corroborated by the bi-phasic inhibition curves obtained for TCA-UMP of *Sc*OMPD in

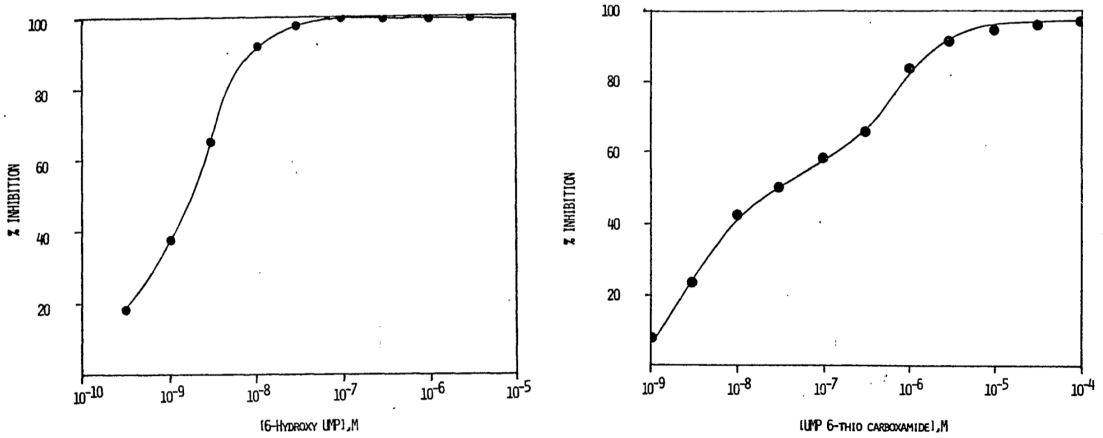


Figure 58: **Inhibition of *Sc*OMPD by BMP and TCA-UMP**. Assay contained 10^{-4} M [^{14}C-carboxy] OMP and recorded $^{14}C^2$ release. Taken from Landesman, PhD Thesis, 1982.[64]

the original publication (Figure 58).[64] The bi-phasic inhibition of OMPD was not observed for BMP or Aza-UMP which do not form the extended catalytic tetrad but rather stabilize Lys_{314} in the conventional active-site architecture.

The implied negative cooperativity was previously observed for various enzymes and could ultimately establish half-of-sites reactivity as observed in ThDP-dependent enzymes.[148] In *Mt*OMPD, a substitution of the lysine residue corresponding to Lys_{314}, leads to a tremendous loss of enzyme activity but an increase in binding affinity of the reaction product UMP towards OMPD.[117,190,191] The alanine substitution variant of *Mt*OMPD binds UMP with an K_d-reduction of 5 order of magnitude (WT: 4.2×10^{-4} M, K72A: 0.3×10^{-9} M). Consequently, without the sterically clashing catalytic tetrad lysine, OMPD would display a severe product inhibition. Since the substrate molecule has higher spatial requirements due to the carboxylate group, the displacement of the lysine residue and the formation of the extended catalytic tetrad could prevent atomic clashes and enable catalysis. After decarboxylation, Lys_{314} could compete with UMP for the resting state position and actively repel the product molecule.

However, since the carbanionic decarboxylation intermediate is supposed to be stabilized by the positive charge of Lys_{314}, the stabilization in the extended catalytic tetrad conformation appears weaker due to the increased distance between the N^{ϵ} of Lys_{314} and the C^6 of the pyrimidine ring. Furthermore, catalysis with the formation of the extended catalytic tetrad would require two conformational changes to facilitate decarboxylation. First, ligand binding induces the closure of the phosphate gripper loop to form the caged enzyme/substrate complex. The closed conformation undergoes simultaneously or subsequently an additional conformational change to form the extended catalytic tetrad as observed in the substrate analogue complex structures.

Catalysis might occur in the newly observed conformation and once the product is formed, the conventional catalytic tetrad is re-established. However, catalysis would require many conformational rearrangements and protein dynamics which conflicts with the pre-organization concept of enzyme active site.[86,112,113] However, the importance for conformational dynamics and protein rearrangements was proven for a variety of enzyme molecules.[192,193] Consequently, it is not clear if the observed amino acid rearrangements in the substrate analogue structures and the determined active site architecture represent the real Michaelis- or catalytic complex conformations.

8.8 Substrate deformation by OMPD

Crystal structures of OMPD in complex with ligands harbouring bulky C^6-substituents on the pyrimidine ring display a certain degree of rotation and torsion of the functional group.[84,116] In this study, the effect of ligand distortion can also be observed in protein structures complexed with the transition state mimics BMP and Aza-UMP to a minor extent and more severely with the substrate analogues Amido-UMP and TCA-UMP. The latter two ligands display a rotation of the amido and thiocarboxamido group out of the pyrimidine plane which is slightly larger for TCA-UMP. Both ligands are mainly populated in the *syn*-conformation in solution and are likewise observed in the crystal structures which represents the energetically favourable conformation for pyrimidine nucleotides with a large C^6-substituent.[154,194] Furthermore, the out-of-plane rotation is a spatial necessity to avoid atomic clashes of the molecule. However, the observed out-of-plane distortion of the C^6-C^7-bond measured as C^2-N^1-C^6-C^7 torsion angle of both inhibitors can not be explained from spatial requirements.

The same rotation and distortion was observed for the substrate complex in the $h\text{OMPD}_{314\text{AcK}}$ protein variant structures. The out-of-plane distortion (43 °) is even more pronounced compared to the substrate analogue structures. However, the distortion does not seem to originate from spatial restrains imposed by the protein environment since the present two additional substrate complex structures (*Ec*OMPD: D312A K314A double mutant, *hs*OMPD D312N) also show the ligand distortion (Figure 59). Based on the distortion angle of OMP in the enzyme/substrate complex, the C^6-atom might be partially sp^3-hybridized. Thus, the torsion angle for the out-of-plane displacement of the C^6-C^7-bond is likely to originate from electrostatic effects and not from spatial restrains as observed in different enzymatic systems.[188,195–197]

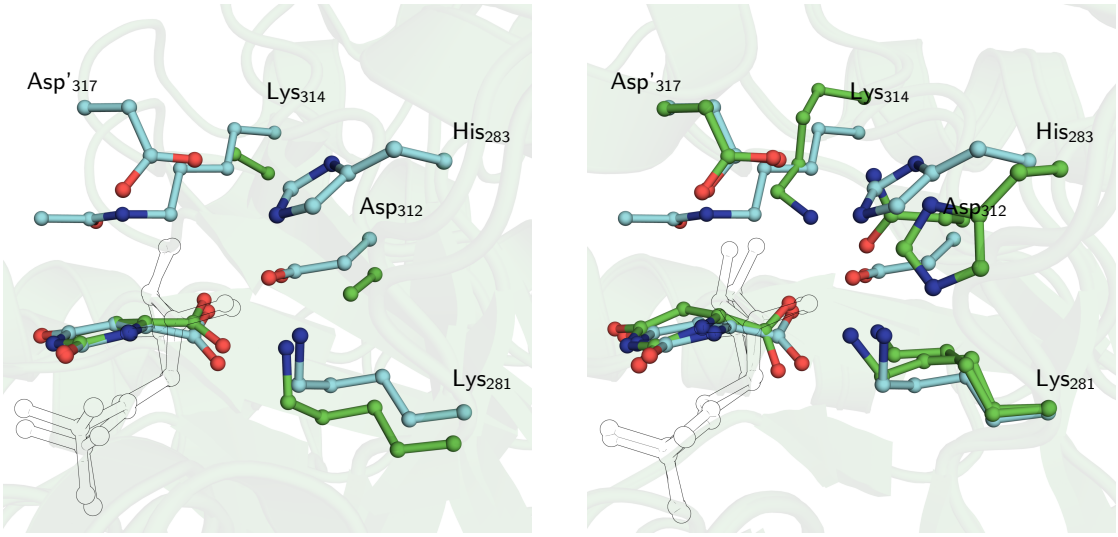


Figure 59: **Substrate distortion in the active site of OMPD.** Active-site view onto the catalytic tetrad residues. Enzyme variants *Ec*OMPD$_{D312A,K314A}$ (A) and *h*OMPD$_{D312N}$ (B) in complex with OMP (cyan). The *h*OMPD$_{314AcK}$ structure is depicted in green. Numbering according to the human enzyme. Coordinates taken from PDB entries 1KM6 and 2QCL.

8.9 A potential reaction mechanisms involving the formation of the "extended catalytic tetrad"

The observed structural re-arrangements induced by substrate analogue binding, led to the formulation of an additional proposal for the reaction mechanism. The conventional decarboxylation proposal postulates a rigid active site architecture in the conformation observed in the BMP liganded structure. The obtained lysine re-orientation in the newly introduced "extended catalytic tetrad" might represent a previously overseen active site conformation (Figure 60). Both substrate analogues, Amido-UMP and TCA-UMP in the wild type protein, and OMP in the *h*OMPD$_{314AcK}$ complex interact with Asp$_{312}$ as hydrogen bond acceptors. The structural data indicates that the substrate analogue structures closely mimic the native substrate complex. If the conformational changes are not considered as ligand binding artefacts, they might also occur during substrate binding and decarboxylation in the wild type enzyme.

The reaction mechanism involving the formation of the extended catalytic tetrad requires the protonation of the substrate's carboxylate group comparable to the conventional mechanistic proposal. Once the substrate enters the active site, the phosphate gripper loop closes over the catalytic pocket to form the caged enzyme/substrate complex. To prevent the electrostatic repulsion of the two carboxylate groups, Lys$_{314}$ protonates the substrate's carboxylate group which can subsequently form the short hydrogen bond to Asp$_{312}$ as hydrogen bond acceptor. Since the hydrogen bonding interaction to OMP is lost, Lys$_{314}$ repositions the N$^{\epsilon}$-amino group

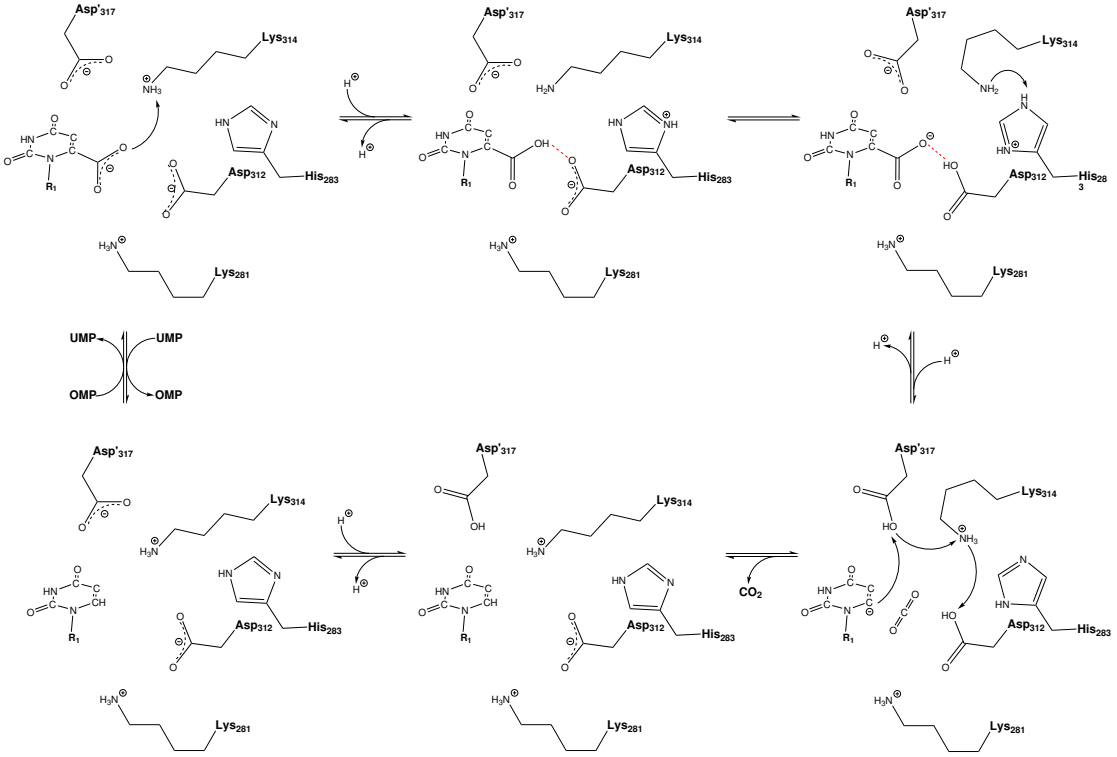


Figure 60: **Alternative reaction mechanism of *h*OMPD catalysed decarboxylation of OMP utilizing the extended catalytic tetrad formation.**

to form the extended catalytic tetrad conformation. The reorientation of the lysine residue induces charged His_{283} to flip and to transfer a proton to Lys_{314} to regenerate the N^{ϵ}-ammonium group. In addition to Asp'_{317}, the histidine residue is in direct contact to the interconnecting water tunnel bridging both subunits of the protein dimer. To neutralize the emerging negative charge, the lysine residue could directly act as catalytic acid and protonate the carbanionic intermediate. However, the reorientation the Lys_{314} placed Asp'_{317} in a favourable position for the protonation reaction.

For Asp'_{317} to act as the catalytic acid for carbanion neutralization, a further proton to form the corresponding acid would be required. As proposed for the conventional decarboxylation mechanism, the proton could be introduced via the inter-subunit water network prior to or after the conformational change to form the extended catalytic tetrad. Interestingly, due to the introduction of positive charges into the active site of the enzyme, the electrostatic environment changed from hydrophilic to likely hydrophobic with a nominal charge of +1. In equilibrium, a certain fraction of the short hydrogen bond between the substrate carboxy group and Asp_{312} will have a protonation on Asp_{312} with the carboxylate group located on the substrate. The substrate's carboxylate group is competent for decarboxylation and the emerging carbanionic

intermediate might be neutralized by a proton transfer from a proton relay system composed of the acids Asp'$_{317}$, Lys$_{314}$ and Asp$_{312}$. In a concerted transfer mechanism, Asp'$_{317}$ acts as the catalytic acid to quench the localized carbanion on position C^6 to generate the reaction product UMP. During the decarboxylation reaction, two protons were transferred to the active site pocket to establish the proton relay system. After product formation, the remaining proton can be fed back into the water network of His$_{283}$ for storage or transferred to the second catalytic site to aid catalysis.

In the reaction pathway involving the formation of the extended catalytic tetrad, His$_{283}$ serves as proton buffer and donor for the formation of the concerted proton transfer. Interestingly, the participation of the histidine residue in catalysis was not considered in the literature and no function has been assigned yet. It is generally considered as ribose binding partner and was not subject to mutational studies so far. The neglection of the residue is not surprising, since the amino acid is not very well conserved in enzymes of different domains of life. Sequence alignments indicate the substitution of His$_{283}$ with glycine (*Mt, Ec*) and glutamine (*Pf*). In the glycine variant a water molecule occupies the position of the histidine residue which is not possible in the glutamine protein variant.[92]

9 Summary: Orotidine 5'-monophosphate decarboxylase

The enzyme orotidine 5'-monophosphate decarboxylase is a remarkable catalyst. It is able to increase the chemical reaction rate compared to the in solution reaction by a factor of 10^{17} and was termed the "most proficient enzyme" known. A plethora of crystal structures of the enzyme was published of a variety of different organisms. Furthermore, structural and kinetic data of wild type enzymes and protein variants was obtained. However, the precise reaction mechanism is still under debate.

In order to obtain closer insights into the reaction mechanism, kinetic measurements were performed using isothermal titration calorimetry. The substrate dependent catalytic activity indicates a cooperative behaviour of both subunits of the protein dimer. Additionally, also binding of the product molecule is likely to occur in an interactive manner. The communication between the two catalytic sites could be established via a water tunnel connecting two active site aspartates.

To analyse the catalytic mechanisms, crystal structures of the human $\mathrm{OMPD_{WT}}$ enzyme were analysed in complex with potential substrate- and transition state analogues. In presence of the substrate analogues 6-amido-UMP and 6-thiocarboxamido-UMP, the active site architecture displays a previously unrecognised conformation. The reactive site lysine which is part of the essential and conserved catalytic tetrad ($\mathrm{Lys_{281}}$, $\mathrm{Asp_{312}}$, $\mathrm{Lys_{314}}$ and $\mathrm{Asp'_{317}}$) adopts an unusual position interacting with the nearby $\mathrm{His_{283}}$. The altered active site composition could indicate the structural organisation of the substrate encounter complex. Furthermore, the binding mode of two transition state analogues was analysed in complex with $h\mathrm{OMPD_{WT}}$. Both inhibitors were thought to exhibit a localized negative charge resembling the carbanionic decarboxylation intermediate. However, atomic resolution crystal structures revealed a binding mode with a large degree of delocalization of the negative charge and not the stabilization of a localized species.

Since structural information on the enzyme in complex with the natural substrate OMP is still lacking, substrate soaking trials were intended to determine the enzyme/substrate complex. Due to the rapid decarboxylation of OMP and potential conformational heterogeneity introduced by substrate conversion, a $\mathrm{N^{\epsilon}}$-acetyllysine (AcK) protein variant was generated ($h\mathrm{OMPD_{314AcK}}$). The substitution of the reactive site $\mathrm{Lys_{314}}$ reduces the catalytic activity by a factor of approx. 1000 but the general active site architecture remains intact. Notably, the modified lysine residue can not serve as acid/base-catalyst and is not positively charged as a natural lysine at pH 8.

Soaking of the substrate molecule into the resting state enzyme of $h\mathrm{OMPD_{314AcK}}$ revealed an

OMP binding mode in which the substrate's carboxy group forms a short hydrogen bond to the catalytic tetrad residue Asp_{312}. The very close distance between both acidic functions, bond length analysis and $2mF_o$-DF_c-electron density distribution of the fuctional groups indicate the protonation of the substrate. The two carboxy groups establish a prodcutive interaction between substrate and enzyme. These results argue against a proposed catalytic mechanism considering electrostatic repulsion between two carboxylate groups in the active site of OMPD as major driving force for catalysis.

10 Outlook: Orotidine 5'-monophosphate decarboxylase

The performed kinetic and structural studies elucidated new facets of OMPD mediated decarboxylation. However, the actual reaction mechanism and the origin of the catalytic prowess remains elusive. The enzyme might not stabilize a specific transition state like a localized vinyl carbanion but rather favour a complex ensemble of chemical states. The emerging carbanion from decarboxylation is likely to be stabilized within the entire pyrimidine ring structure, applying not only the positive charge of the catalytic lysine residue.

Furthermore it remains elusive if the "extended catalytic tetrad" is an actual catalytic conformation. Using conventional crystallographic techniques, one might not be able to determine a potential conformational change during catalysis. The substrate-soaking trials of the wild type enzyme demonstrated the limitations of the applied method. The diffraction of the protein crystals is abolished until the product complex is formed. The usage of micro-crystals could dampen the effect of introduced crystal disorder during substrate conversion. Furthermore, measurements of micro-crystals at room temperature with a precise and automated mixing procedure of protein crystals with the substrate solution could resolve conformational changes during catalysis.

Furthermore, nuclear magnetic resonance spectroscopy (NMR) could be applied to detect a potential lysine movement during catalysis. Notably, structural rearrangements during decarboxylation are expected to be rather subtle, since the wild-type crystals are able to perform catalysis without crystal decomposition. The catalytic reorganisation might include only the observed small changes in the substrate analogue structures. In order to elucidate the role of the reactive site His_{283}, enzyme variants should be prepared with glycine and glutamine substitutions as observed in the active sites of *E. coli* and *P. falciparum*, respectively. Additionally, the influence of the histidine residue on the catalytic cooperativity between both monomers is of high interest. No structural or kinetic data is present yet analysing the participation of the residue in the reaction mechanism.

During decarboxylation, the catalytic competent enyzme fraction displays a fluorescence quenching effect. Extensive studies of the yeast enzyme utilized the differential tryptophane fluorescence intensity to quantify the substrate dependent reaction rates. The effect is not observed in presence of UMP and BMP which indicates that these conformations are not equal to the activated [ES]-complex. Aza-UMP complex formation displays only a small influence on intrinsic protein fluorescence. Consequently, the origin of the reduced fluorescence is still unknown and could indicate that the catalytic conformation has yet to be discovered.

11 Introduction: Acetoacetate decarboxylase

A remarkable feature of enzyme catalysis is the dynamic modification of pK_a-values of certain ionizable groups. These acidic or basic functional groups can be amino acids in nature, e.g. aspartates, glutamates or lysines and histidines, or belong to bound cofactor molecules or substrates/ligands. Among other examples, the strength of pK_a-value modulation was demonstrated for triose phosphate isomerase (TIM) in which a glutamate residue is utilized for proton abstraction from the substrate.[102] Loop closure sequesters the acidic side chain from the accessible solvent. The environment of a low dielectric constant in the protein interior leads to an increase in the pK_a-value of the residue. The increased basicity of approx. 2.8 units is utilized to abstract a proton from the substrate molecule.[198]

In thiamin pyrophosphate dependend enzymes, the cofactor activation requires a proton transfer to obtain the reactive carbene species which subsequently acts as a nucleophile on carbonyl centres.[199] The basicity of the carbon acid of the thiazolium ring is decreased by approx. 8 pK_a-units.[140]

One of the first examples of a pK_a-value alteration by an enzyme molecule was introduced with the acetoacetate decarboxylase from *Clostridium acetobotylicum* (*Ca*AAD). The enzyme catalyses the irreversible decarboxylation of the 2-keto acid acetoacetate to yield acetone and carbon dioxide (Figure 62). With the development of sophisticated purification procedures, the first biochemical properties of *Ca*AAD and the molecular composition of the enzyme were elucidated.[200–209] Early in the field of biochemistry, in 1959, Hamilton and Westheimer proposed a potential Schiff-base formation with an enamine intermediate to enable CO_2 release.[210] The decarboxylation reaction showed an exchange of the carbonyl oxygen of the substrate with an oxygen of the aqueous solvent. Importantly, the reaction would require the deprotonation of a lysine residue in the active site of *Ca*AAD to act as a nucleophile for Schiff-base formation. However, the assumed pK_a-value of the N$^\epsilon$ ammonium group in solution would be too high for efficient catalysis.

The reaction was confirmed by the reduction of the Schiff-base intermediate with sodium boro-

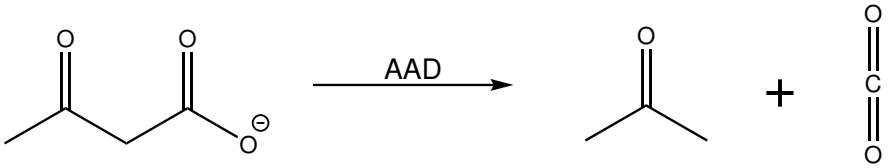


Figure 61: **Decarboxylation reaction catalysed by acetoacetate decarboxylase.**

hydride which led to the isolation of ϵ-N-iso propyllysine (IprLys).[202] The proteolytic digestion analysis of the labelled samples identified the peptide sequence Pro-IprLys-Lys as belonging to the catalytic site. They proposed that in order to increase the acidity of the active site lysine "it is tempting to speculate that the second lysine residue in acetoacetate decarboxylase is part of the positive binding site for the enzyme".[58] The positioning of two lysine residues in direct proximity would lead to electrostatic repulsion if both of them were charged which could suppress the pK_a-value of one of the residues.

Systematic studies focussing on the determination of the pK_a-value of the reactive site lysine in the interior of the protein revealed a pronounced increase in acidity. In elegant experiments, Westheimer and colleagues intended to elucidate the pK_a-value of the identified active site lysine. They utilized environmental sensitive pH-indicators with ionizable groups which were attached to the catalytic lysine residue. The pK_a-values of the compounds were titrated and the acidity of the reactants in the active site of the enzyme compared to the acidities measured in aqueous solution. The resulting pK_a-value shift resulting from accommodating the compound in the protein interior was assumed to apply to the reactive site lysine to the same extent. One of the utilized reactions was 5-nitrosalicylaldehyd which reacts with an amino group of the enzyme, likely the catalytic lysine, to form a Schiff-base product. The imine can be reduced with borohydride to yield N-substituted 2-hydroxy-5-nitrobenzyl.[211] The pK_a-value of the compound in the active site was determined to be 2.4 and is 3.4 pK_a-units lower than the value of the model compound N-methyl-2-hydroxy-5-nitrobenzylamine which was used as reference. However, the low pH-values required for pH-titration and the protein stability in acidic environments imposed severe difficulties on the experimental design.

In another study applying the same compounds, the catalytic lysine residue reacted with 5-nitrosalicylaldehyde and the Schiff-base from the condensation reaction was reduced with borohydride. The product of the reaction was used as photometric reporter group in the active site of AAD. The titration between pH 4.5 and 9.7 yielded small changes in the optical absorbance attributed to ammonium of the Schiff-base. The pK_a-value was determined by titration and compared to the value of N-methyl-2-hydroxy-5-nitrobenzylamine as model compound in solution. In the active site of AAD, the pK_a-value of the reporter group was determined to be 6 which indicated a pK_a-suppression of 4.7 units based on a pK_a-value of 10.7 of N-methyl-2-hydroxy-5-nitrobenzylamine in solution.[212]

Comparable studies utilizing different chemical reporter substances confirmed a pK_a-value of approx. 5.9 of the catalytic lysine of AAD.[213] The experiments demonstrated the strength

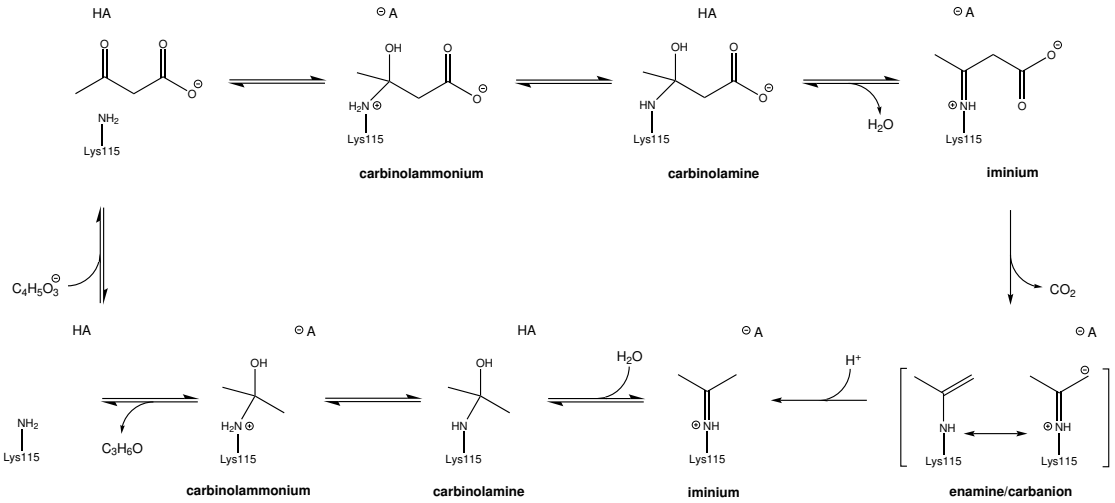


Figure 62: **Reaction pathway of acetoacetate decarboxylase mediated decarboxylation.**

of protein molecules to increase the acidity of an active site lysine, however, the origin of the pK_a-modulation remained elusive. Interestingly, the adjacent lysine (Lys_{116}) showed to have an impact on the acidity of the catalytic residue.[48] The amino acid substitution variants K116C and K116N showed a pK_a-value of >9.2 whereas the protein variant K116R with the conserved positive charge displayed an unperturbed acidity of Lys_{115}.[48]

With the release of the first crystal structures of *Ca*AAD, the proposed electrostatic destabilization of the ionized lysine residue in the active site of AAD was challenged.[47,214] The catalytically active enzyme is a 365 kDa homododecameric protein. The monomers form dimers of tetramers which assemble into a hollow sphere. The structural data revealed that the two lysine residues are facing into opposite directions with a distance of the N$^\epsilon$-ammonium groups of 14.8 Å. The active site pocket is composed of the hydrophobic residues Phe_{26}, Leu_{71}, Met_{96}, Leu_{98} and Leu_{223}. The tyrosine residues Tyr_{74} and Tyr_{113} are oriented perpendicular to and co-aligned with the catalytic lysine, respectively. A positive charged Arg_{29} is proposed to facilitate substrate binding establishing interactions with the acetoacetate's carboxylate group. Since the arginine residue is situated approx. 8 Å apart from the catalytic lysine, the influence of the residue's pK_a-value was assumed to be small. Additionally, approx. 4 Å apart from Lys_{115} the glutamate residue Glu_{76} is positioned which displayed two alternative conformations. The main conformation is facing towards the reactive site lysine and the second conformation is likely to form a hydrogen bond to the adjacent residue Glu_{61}.

Based on the crystal structure, the rather hydrophobic active site pocket was claimed as the origin for the pK_a-perturbation of Lys_{115}. Due to the distance of the N$^\epsilon$-ammonium groups

of Lys_{115} and Lys_{116}, an electrostatic destabilization effect was assumed unlikely. Aided by the atomic coordinates from the crystal structure, computational studies indicated an acidity increase of the reactive lysine of 8.3 pK_a-units from the desolvation of Lys_{115} in the active site. The pK_a-value decrease is partially compensated by the atomic charges of the surrounding residues leading to a decrease in acidity of 3.6 pK_a-units. The pK_a-values of the monomer and the dodecameric protein assembly were determined to be 5.73 and 5.37, respectively.[215]

The published crystal structures showed the resting state enzyme conformation and the covalent complex after reaction with 2,4-pentanedione (*Cv*AAC only).[47] Notably, the chosen inhibitor is not negatively charged as the natural substrate acetoacetate. In order to analyse the enzyme/inhibitor complex with a charged intermediate in the catalytic pocket and to analyse potential conformational changes induced by the presence of a negative charge, two new inhibitors, acetonyl phosphonic acid (AcP) and acetonyl sulfonic acid (AcS) were used. The Schiff-base forming inhibitors were synthesized by the Department of Chemistry, Toronto. The group of Ronald Kluger supplied the chemical compounds to be analysed bound to *Ca*AAD and to verify the newly introduced structural data.

12 Results: Acetoacetate decarboxylase

12.1 Expression and purification of *Ca*AAD

In order to analyse the pK_a-suppression of the catalytic lysine residue of the acetoacetate decarboxylase from *Clostridium acetobotylicum*, functional protein needed to be expressed and purified. Based on the crystallographic and kinetic studies performed by Ho *et al.*, a modified purification protocol was established.[47] To prevent potential complications concerning the intended crystallization and structure determination, the construct was designed to obtain an identical peptide sequence as published in PDB-entry 3BH2. The amino acid sequence was translated into a nucleotide sequence for expression in the bacterial *E. coli* expression system. The synthetic and codon optimized nucleotide sequence was ordered from GeneArt and sub-cloned into the pET-28a expression vector. No protein affinity tag was added in order to follow the native purification procedure. The protein was recombinantly expressed in *E. coli* cells.

The SDS-Page analysis of protein expression displays a clear band with an approximate size of 30 kDa which correlates with the calculated molecular monomer mass of 27 kDa of *Ca*AAD (referred to as AAD in the following sections if not indicated otherwise). According to the modified purification protocol, the AAD protein was precipitated from the cleared cell lysate with ammonium sulfate.[47,48] The cell pellet from the precipitation step was dissolved and thoroughly dialysed against lysis buffer to remove residual ammonium sulfate. The sample was excessively diluted with lysis buffer and applied to a tri-ethylammonium anion exchange chromatography column. Due to the very low affinity of the protein towards the column matrix and the presence of phosphate anions in the lysis buffer, the protein elutes very early in an ammonium sulfate gradient and has to be applied strongly diluted onto the column. However, AAD elutes from the matrix in a single and distinct peak with only a minor amount of contaminations. To remove remaining impurities, a S-500 gel-filtration step was applied.

Since kinetic measurements indicated a severe catalytic activity loss due to the application of the gel-filtration column, this step was omitted and replaced with an additional dialysis step for buffer exchange. The data shown in the following sections were obtained from protein preparation without the application of the S-500 gel-filtration column if not stated otherwise.

12.2 Determination of macroscopic kinetic constants of *Ca*AAD

In order to determine the macroscopic catalytic constants of *Ca*AAD, a continuous spectrophotometric steady-state assay was performed.[47,48] The reaction progression could directly be mea-

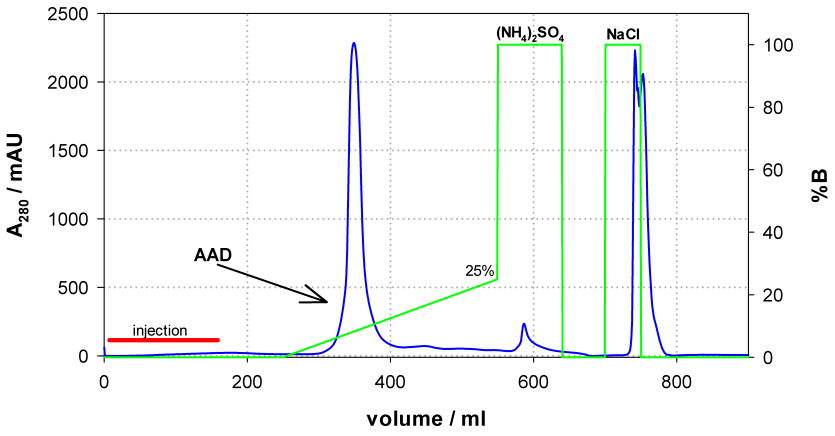


Figure 63: **TMAE-anion exchange chromatography of *Ca*AAD**. The blue line indicates the intrinsic protein absorbance. The green traces indicate the fraction of elution buffers containing ammonium sulfate (300 mM) and the column cleaned with NaCl (1 M)

sured based on the differential absorbance of acetoacetate and acetone, the reaction's substrate (enol form) and the product, respectively. Acetoacetate has two absorbance maxima at 210 nm and 270 nm (Figure 64). Conditions with high substrate concentrations were monitored at 270 nm and low substrate concentrations at 210 nm. Due to the instability of the substrate in aqueous solution, freshly synthesized acetoacetate stocks were used to prepare a 1 M solution in lysis buffer, the pH adjusted, small sample aliquots frozen in liquid nitrogen and stored at -20 °C. For a measurement, an aliquot was taken and the acetoacetate concentration determined at 210 nm.

A previously reported phenomenon describes the activation of the AAD activity by heat treatment.[216] To analyse the influence of a short heating period on enzyme catalysis, an untreated sample and a heat-treated sample was measured using the steady-state assay. Assuming independent active sites of the protein homododecamer, the data-points were fitted according to Michaelis and Menten.[24] The protein sample without heat-treatment displayed a k_{cat}-value of 174 ± 4 s^{-1} and a K_M-value of 4.5 ± 0.4 mM. The heat-activated sample showed a k_{cat}-value of 260 ± 25 s^{-1} and a K_M-value of 5.2 ± 0.7 mM. Thus, the heat-treatment results in an approx. 50 % increase in the catalytic activity. Notably, a small amount of protein precipitated during the heating step. However, the concentration loss was below 10 % and the removal of incorrectly or unfolded protein cannot account for the strong increase in catalytic activity. Notably, the substrate dependent activity of the pre-heated AAD showed a deviation from the hyperbolic curve progression. The deviation could indicate a violation of the conditions assumed by Michaelis and Menten. A potential cooperativity between the subunits of the dodecameric protein could result in a sigmoidal curve. In order to account for potential depended active sites, the experimental

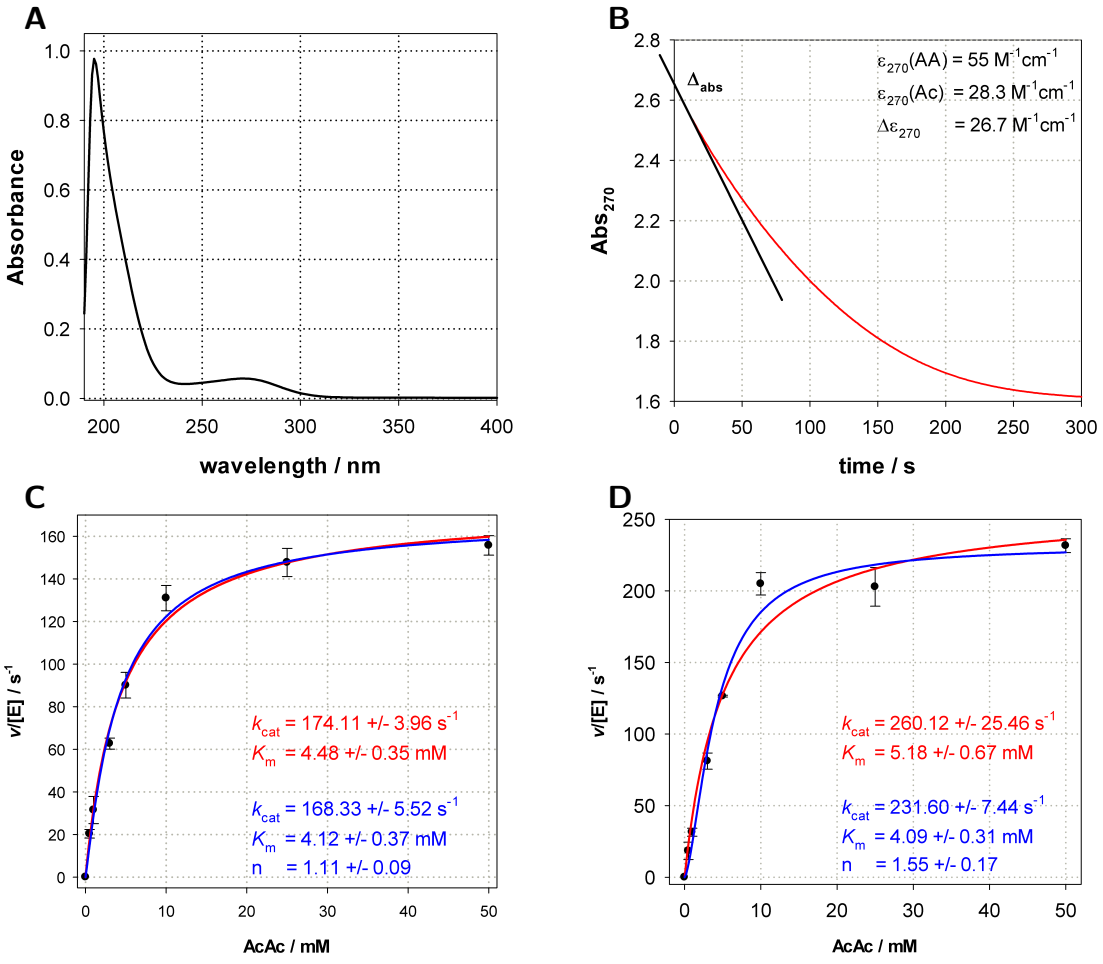


Figure 64: **Spectroscopic steady-state assay of *Ca*AAD.** A: Wavelength dependent absorbance of acetoacetate in aqueous solution. B: Representative raw-data curve of the enzymatic acetoacetate consumption. C: Substrate dependent turn-over rates of *Ca*AAD without the activation by heat. D: Substrate dependent turn-over rates of an activated enzyme sample.

data points were additionally fitted according to Hill.[144] The heat-treated protein sample showed a turnover number of 231 ± 7 s^{-1} and a K_M-value of 4.1 ± 0.3 mM with a cooperativity factor of n=1.55 ± 0.2.

12.3 Secondary structure analysis applying circular dichroism

In order to determine the differential absorbance of polarized light of *Ca*AAD and to obtain the protein specific secondary structure trace, circular dichroism was used. The CD-traces can be used to verify the structural integrity of protein preparations and the proper folding of enzyme variants. The CD-spectrogram shows a high content of anti-parallel β-sheets.[3] Based on the crystal structure, the enzyme monomer is mainly composed of β-sheets in an anti-parallel orientation. Only a minor portion of the protein folds into α-helical structures (44.3 % and 15.6 %

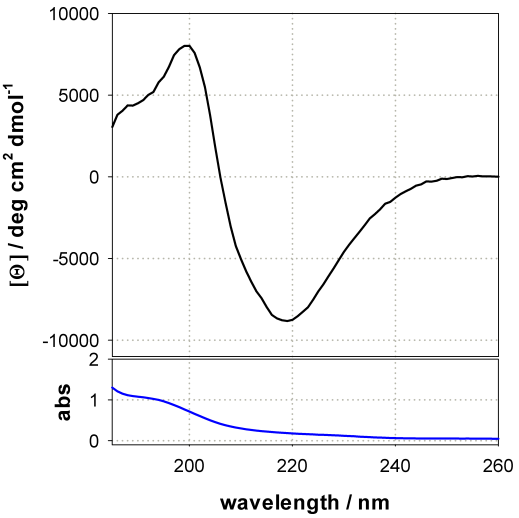


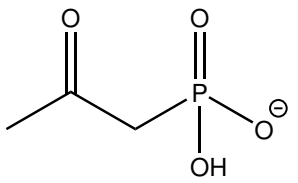
Figure 65: **Protein secondary structure analysis of *Ca*AAD applying circular dichroism.** Upper panel: Molar mean residue weight ellipticity. Lower panel: Absolute absorbance at 280 nm of the protein sample. Spectra were recorded of approx. 0.1 mg/mL protein in 10 mM KH_2PO_4 pH 5.95 at 25 °C with 30 averaged repeats.

β-strands and α-helices, respectively).[47]

12.4 Structural analysis of *Ca*AAD applying macromolecular cryo-crystallography

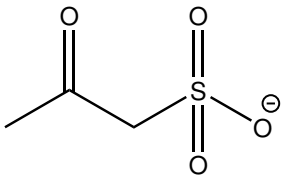
Since the quaternary structure of AAD was published previously, the main objective was to increase the obtainable resolution. The recorded resolution of 2.4 Å lacks information on most water molecules in the crystal structure which could contribute to the active site composition. Furthermore, the inhibitor complex structures of acetonyl phosphonic acid (AcP) and acetonyl sulfonic acid (AcS) were intended to be determined. The published crystallisation conditions were used to grow protein crystals of the resting state enzyme. The addition of AcP and AcS into the protein drop for crystallisation produced crystals with the same morphology. Due to the high polyethylene glycol and sarcosine concentration in the reservoir solution, plunge freezing in liquid nitrogen yielded non crystalline ice and could be used directly for synchrotron irradiation. However, in order to increase the diffraction resolution, additional cryo-protectants were screened. The best results were obtained with a mixture of ethylene glycol, propylene glycol and polyethylene glycol 400. The crystals were transferred into the cryo-solution in four steps with increasing concentrations. The transfer of at room temperature grown crystals into the cold-room (6 °C) further reduced the crystal mosaicity.

The obtained diffraction data was indexed as $P2_12_12_1$ with unit cell dimensions of A=104 Å, B=172 Å and 376 Å. Due to unexpectedly large unit cell constants and space group deviation from the published data, the structure was solved using molecular replacement (Gleb Bourenkov, Ashwin Chari). The asymmetric unit contains two dodecamers for all measured crystals.



acetonylphosphonic acid

pK_a = 6.3
K_i = 0.8 mM (pH 5.92)
intermediate not reduced with borohydride



acetonylsulfonic acid

pK_a = ~0
K_i = 8.0 mM (pH 5.92)
intermediate reducable with borohydride

Figure 66: **Inhibitors to analyse the covalent Schiff-base intermediate.** The substances were kindly provided by Prof. Dr. Ronald Kluger.

The resting state structure was determined to a resolution of 1.85 Å and the model adjusted to the experimental data to correlate with R_{work}=16.5 % and R_{free}=16.5 %. The active site architecture is comparable to the published crystal structures. However, the increase in resolution revealed further alternative conformations which were not detected before. Orchestrated side chain movements are observed for Glu_{76}, Met_{96}, Glu_{61} and Arg_{59}. The partially mutual exclusive conformations appear to exhibit a coordinated movement and form a path from the reactive site cavity to the exterior of the protein. The four amino acids are referred to as "residue zipper".

Experimental data of the inhibitor complex structures of AcP and AcS was collected to a resolution of 1.75 Å and 1.55 Å, respectively. The model of the AcP-complex was build with R_{work}=15.8 % and R_{free}=15.8 % and of the AcS-complex with R_{work}=15.5 % and R_{free}=15.5 %. Both ligands were modelled as modified lysine residue with an occupancy of 80 % for the non-lysine atoms. Notably, the main conformations of the zipper-residues occupy the opposing orientations. Additionally, Arg_{29} is detected to occupy two distinct conformations in the inhibitor complexes.

In the AcP-complex, Glu_{76} adopts the positions proximal to the catalytic lysine residue. The adjacent Met_{96} could spatially adopt both alternative conformations but is found mainly populated in the distal conformation with respect to Lys_{115}. Glu_{61} and Arg_{59} are facing to the protein exterior and no interaction between Glu_{76} and Glu_{61} can be found. The two glutamate residues are spatially separated by approx. 5 Å.

In the AcS-complex structure, Glu_{76} is oriented in the distal conformation with respect to the catalytic lysine and the sulfonate group. It is potentially forming a hydrogen bond with Glu_{61}.

Glu_{61} itself adopts the conformation which faces into the reactive site cavity. Together with Glu_{61}, Arg_{59} orients towards the protein interior. The methionine residue adjacent to Lys_{115} is positioned in the proximal conformation with respect to the lysine residue.

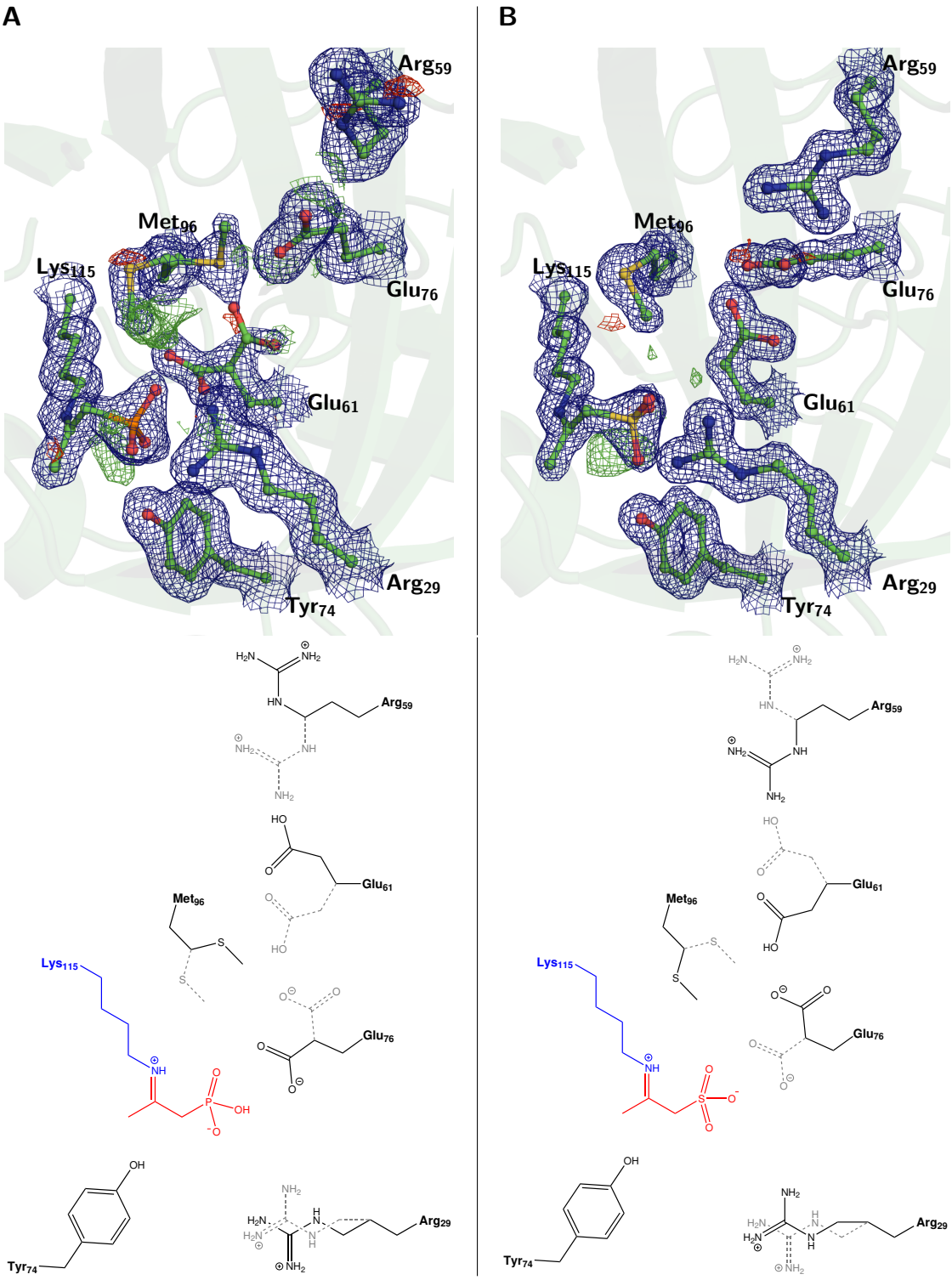


Figure 67: **Reactive site zipper positions of AAD in complex with the inhibitors AcP and AcS.** Upper panel: Covalent reaction intermediates forming a Schiff-base. A: Inhibitor complex with acetonyl phosphonic acid. The $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with a contour level of 1.0 and $\pm 3\,\sigma$, respectively. Major conformational populations modelled with 70 % occupancy. B: Acetonyl sulfonic acid complex. The $2mF_o$-DF_c electron density map (blue) and mF_o-DF_c difference electron density maps (pos: green; neg: red) are shown as meshes with a contour level of 1.5 and $\pm 3\,\sigma$, respectively. Lower panel: Schematic representation of the network of alternative conformations. The solid black and the dashed grey lines indicate the major and minor or not populated conformational species in the corresponding crystal structure, respectively.

13 Discussion: Acetoacetate decarboxylase

Acetoacetate decarboxylase served as prototypical example for the enzyme induced suppression of pK_a-values. The first mechanistic proposals were made without the knowledge of the 3-dimensional structure of the enzyme. Two conserved lysine residues which are on adjacent positions in the polypeptide chain were considered to interact to modify the acidity of the catalytic residue. However, when the first crystal structures were available, the spatial separation of the two residues was found to be too large for a repulsive interaction. The pK_a-suppression was attributed to the location of the lysine in a hydrophobic pocket which would favour the uncharged species. In order to further analyse the architecture of the active site pocket, a purification protocol was adopted and functional protein expressed and purified.[48]

The expression of *Ca*AAD in *E. coli* cells without an affinity-purification tag yielded high amounts of catalytically active protein. The activity of the protein samples can be directly measured using the absorbance of the substrate acetoacetate. However, first negative stain test images using electron microscopy showed a certain fraction of disintegrated particles. Since the dodecameric assembly might be sensitive to physical damage, the gel-filtration step was omitted. The samples were dialysed and only applied the to TMAE-ion exchange chromatography column. The purification procedure omitting the S500-column resulted in a protein purification with an increase in activity of 50 %.

Furthermore, the published heat-activation effect could be observed for the purified protein sample.[216] The incubation of AAD at 70 °C for 60 min yielded a protein sample with an activity of approx. 150 % compared to an untreated fraction. Only a small protein amount precipitated and could not account for the activity increase. The determined turn-over number of 230 s^{-1} is 35 % higher than the values published with the crystal structures.[47] The higher activity of the protein purification is likely to result from the heat-treatment since the step was omitted in the published purification procedure. Notably, all crystal structures were determined with untreated protein samples. The heat-treatment resulted in impaired crystal growth and diffraction quality.

The first crystallization trials produced crystals which already showed X-ray diffraction. The high PEG and sarcosine concentrations are sufficient for the cryo-protection of the protein crystal. However, a systematic screen of cryo-protection additives revealed conditions which were increasing the diffraction quality. Furthermore, the transfer of protein crystals grown at room temperature into 6 °C further reduced crystal mosaicity. Thus, the best structures obtained had a higher resolution compared to the published ones with an resolution increase of 1 Å. Importantly,

the obtained structures show an asymmetry in the dodecameric structure and crystallized as a dimer of dodecamers. The asymmetric unit of the published structure is composed of a tetramer. The obtained higher resolution might be sufficient to resolve the subtle asymmetry in the dodecameric particle.

The dynamic nature of the protein assembly is further indicated by the presence of flexible side chains in the path from the enzyme exterior to the active site cavity. The orchestrated movement of Glu_{76}, Glu_{61} and Arg_{59} might mediate a proton transfer from the outside of the protein into the interior. Furthermore, the alternative conformations could regulate the active site and influence catalysis. However, the actual role of the dynamic "zipper residues" remains elusive.

In order to analyse the pK_a-suppression, computational studies are required to elucidate the factors leading to the high acidity of Lys_{115}. Furthermore, single particle cryo-electron microscopy might determine the influence of protein dynamics on the reaction pathway.

14 References

[1] R. K. Saiki, D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. “Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase”. *Science* 239.4839 (1988), pp. 487–491.

[2] D. Hanahan. “Studies on transformation of Escherichia coli with plasmids.” *J Mol Biol* 166.4 (1983), pp. 557–580.

[3] Sharon M. Kelly, Thomas J. Jess, and Nicholas C. Price. “How to study proteins by circular dichroism.” *Biochim Biophys Acta* 1751.2 (2005), pp. 119–139.

[4] Julia G. Wittmann and Markus G. Rudolph. “Pseudo-merohedral twinning in monoclinic crystals of human orotidine-5’-monophosphate decarboxylase”. *Acta Cryst. D* 63.6 (2007), pp. 744–749.

[5] Julia G. Wittmann, Daniel Heinrich, Kathrin Gasow, Alexandra Frey, Ulf Diederichsen, and Markus G. Rudolph. “Structures of the human orotidine-5’-monophosphate decarboxylase support a covalent mechanism and provide a framework for drug design.” *Structure* 16.1 (2008), pp. 82–92.

[6] T. A. Holton and M. W. Graham. “A simple and efficient method for direct cloning of PCR products using ddT-tailed vectors.” *Nucleic Acids Res* 19.5 (1991), p. 1156.

[7] F. W. Studier and B. A. Moffatt. “Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes.” *J. Mol. Biol.* 189.1 (1986), pp. 113–130.

[8] B. A. Moffatt and F. W. Studier. “T7 lysozyme inhibits transcription by T7 RNA polymerase.” *Cell* 49.2 (1987), pp. 221–227.

[9] F. W. Studier. “Protein production by auto-induction in high density shaking cultures”. *Protein Expr. Purif.* 41.1 (2005), pp. 207–234.

[10] F. William Studier. “Stable Expression Clones and Auto-Induction for Protein Production in E. coli”. *Structural Genomics* (2013), pp. 17–32.

[11] H. Edelhoch. “Spectroscopic determination of tryptophan and tyrosine in proteins.” *Biochemistry* 6.7 (1967), pp. 1948–1954.

[12] C. N. Pace, F. Vajdos, L. Fee, G. Grimsley, and T. Gray. “How to measure and predict the molar absorption coefficient of a protein”. *Protein Sci.* 4.11 (1995), pp. 2411–2423.

[13] S. C. Gill and P. H. von Hippel. “Calculation of protein extinction coefficients from amino acid sequence data”. *Anal. Biochem.* 182.2 (1989), pp. 319–326.

[14] M. R. Wilkins, E. Gasteiger, A. Bairoch, J. C. Sanchez, K. L. Williams, R. D. Appel, and D. F. Hochstrasser. “Protein identification and analysis tools in the ExPASy server”. *Methods Mol. Biol.* 112 (1999), pp. 531–552.

[15] W. N. Burnette. “"Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate–polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A”. *Anal. Biochem.* 112.2 (1981), pp. 195–203.

[16] U. K. Laemmli. “Cleavage of structural proteins during the assembly of the head of bacteriophage T4”. *Nature* 227.5259 (1970), pp. 680–685.

[17] Pierre Bouguer. *Essai d’optique, Sur la gradation de la lumière*. 1729.

[18] Johann Heinrich Lambert. *Photometria, sive de mensura et gradibus luminis, colorum et umbrae*. 1760.

[19] August Beer. “Bestimmung der Absorption des rothen Lichts in farbigen Flüssigkeiten”. *Annalen der Physik und Chemie* (1852).

[20] S. Keller, C. Vargas, H. Zhao, G. Piszczek, C. A. Brautigam, and P. Schuck. “High-precision isothermal titration calorimetry with automated peak-shape analysis”. *Anal. Chem.* 84.11 (2012), pp. 5066–5073.

[21] Chad A Brautigam, Huaying Zhao, Carolyn Vargas, Sandro Keller, and Peter Schuck. "Integration and global analysis of isothermal titration calorimetry data for studying macromolecular interactions". *Nature Protocols* 11.5 (2016), pp. 882–894.

[22] Ernesto Freire, Arne Schön, and Adrian Velazquez-Campoy. "Isothermal Titration Calorimetry: General Formalism Using Binding Polynomials". *Methods in Enzymology* (2009), pp. 127–155.

[23] Adrian Velazquez-Campoy, Stephanie A. Leavitt, and Ernesto Freire. "Characterization of Protein-Protein Interactions by Isothermal Titration Calorimetry". *Protein-Protein Interactions* (), pp. 35–54.

[24] L. Michaelis and M.L. Menten. "Kinetic der Invertinwirkung". *Biochem. Zeitung* (1913).

[25] Ewa Poduch, Angelica M. Bello, Sishi Tang, Masahiro Fujihashi, Emil F. Pai, and Lakshmi P. Kotra. "Design of inhibitors of orotidine monophosphate decarboxylase using bioisosteric replacement and determination of inhibition kinetics." *J. Med. Chem.* 49.16 (2006), pp. 4937–4945.

[26] Mark K. Transtrum, Lee D. Hansen, and Colette Quinn. "Enzyme kinetics determined by single-injection isothermal titration calorimetry." *Methods* 76 (2015), pp. 194–200.

[27] T. A. Pemberton, B. R. Still, E. M. Christensen, H. Singh, D. Srivastava, and J. J. Tanner. "Proline: Mother Nature's cryoprotectant applied to protein crystallography". *Acta Crystallogr. D Biol. Crystallogr.* 68.Pt 8 (2012), pp. 1010–1018.

[28] M. Cianci, G. Bourenkov, G. Pompidor, I. Karpics, J. Kallio, I. Bento, M. Roessle, F. Cipriani, S. Fiedler, and T. R. Schneider. "P13, the EMBL macromolecular crystallography beamline at the low-emittance PETRA III ring for high- and low-energy phasing with variable beam focusing". *J Synchrotron Radiat* 24.Pt 1 (2017), pp. 323–332.

[29] Wolfgang Kabsch. "XDS". *Acta Crystallographica Section D Biological Crystallography* 66.2 (2010), pp. 125–132.

[30] Wolfgang Kabsch. "Integration, scaling, space-group assignment and post-refinement". *Acta Crystallographica Section D Biological Crystallography* 66.2 (2010), pp. 133–144.

[31] P. R. Evans. "An introduction to data reduction: space-group determination, scaling and intensity statistics". *Acta Crystallogr. D Biol. Crystallogr.* 67.Pt 4 (2011), pp. 282–292.

[32] Paul D. Adams, Pavel V. Afonine, Gábor Bunkóczi, Vincent B. Chen, Ian W. Davis, Nathaniel Echols, Jeffrey J. Headd, Li-Wei Hung, Gary J. Kapral, Ralf W. Grosse-Kunstleve, and et al. "PHENIX: a comprehensive Python-based system for macromolecular structure solution". *Acta Crystallographica Section D Biological Crystallography* 66.2 (2010), pp. 213–221.

[33] Kay Diederichs and P. Andrew Karplus. "Improved R-factors for diffraction data analysis in macromolecular crystallography". *Nature Structural Biology* 4.4 (1997), pp. 269–275.

[34] P. A. Karplus and K. Diederichs. "Linking Crystallographic Model and Data Quality". *Science* 336.6084 (2012), pp. 1030–1033.

[35] K. Diederichs and P. A. Karplus. "Better models by discarding data?" *Acta Crystallographica Section D Biological Crystallography* 69.7 (2013), pp. 1215–1222.

[36] Martyn D. Winn, Charles C. Ballard, Kevin D. Cowtan, Eleanor J. Dodson, Paul Emsley, Phil R. Evans, Ronan M. Keegan, Eugene B. Krissinel, Andrew G. W. Leslie, Airlie McCoy, and et al. "Overview of theCCP4 suite and current developments". *Acta Crystallographica Section D Biological Crystallography* 67.4 (2011), pp. 235–242.

[37] F. C. Bernstein, T. F. Koetzle, G. J. Williams, E. F. Meyer, M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi. "The Protein Data Bank: a computer-based archival file for macromolecular structures". *J. Mol. Biol.* 112.3 (1977), pp. 535–542.

[38] G. N. Murshudov, A. A. Vagin, and E. J. Dodson. "Refinement of Macromolecular Structures by the Maximum-Likelihood Method". *Acta Crystallographica Section D Biological Crystallography* 53.3 (1997), pp. 240–255.

[39] Garib N. Murshudov, Pavol Skubák, Andrey A. Lebedev, Navraj S. Pannu, Roberto A. Steiner, Robert A. Nicholls, Martyn D. Winn, Fei Long, and Alexei A. Vagin. "REFMAC5 for the refinement of macromolecular crystal structures". *Acta Crystallographica Section D Biological Crystallography* 67.4 (2011), pp. 355–367.

[40] A. Vagin and A. Teplyakov. "Molecular replacement with MOLREP". *Acta Crystallogr. D Biol. Crystallogr.* 66.Pt 1 (2010), pp. 22–25.

[41] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan. "Features and development of Coot". *Acta Crystallographica Section D Biological Crystallography* 66.4 (2010), pp. 486–501.

[42] Andrey A. Lebedev, Paul Young, Michail N. Isupov, Olga V. Moroz, Alexey A. Vagin, and Garib N. Murshudov. "JLigand: a graphical tool for theCCP4 template-restraint library". *Acta Crystallographica Section D Biological Crystallography* 68.4 (2012), pp. 431–440.

[43] G. M. Sheldrick. "A short history of SHELX". *Acta Crystallogr., A, Found. Crystallogr.* 64.Pt 1 (2008), pp. 112–122.

[44] G. M. Sheldrick. "Crystal structure refinement with SHELXL". *Acta Crystallogr C Struct Chem* 71.Pt 1 (2015), pp. 3–8.

[45] Schrödinger, LLC. "The PyMOL Molecular Graphics System" (2015).

[46] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin. "UCSF Chimera–a visualization system for exploratory research and analysis". *J Comput Chem* 25.13 (2004), pp. 1605–1612.

[47] M. C. Ho, J. F. Menetret, H. Tsuruta, and K. N. Allen. "The origin of the electrostatic perturbation in acetoacetate decarboxylase". *Nature* 459.7245 (2009), pp. 393–397.

[48] L. A. Highbarger, J. A. Gerlt, and G. L. Kenyon. "Mechanism of the reaction catalyzed by acetoacetate decarboxylase. Importance of lysine 116 in determining the pKa of active-site lysine 115". *Biochemistry* 35.1 (1996), pp. 41–46.

[49] I. Fridovich. "Inhibition of acetoacetic decarboxylase by anions. The Hofmeister lyotropic series". *J. Biol. Chem.* 238 (1963), pp. 592–598.

[50] M. A. Keller, G. Piedrafita, and M. Ralser. "The widespread role of non-enzymatic reactions in cellular metabolism". *Curr. Opin. Biotechnol.* 34 (2015), pp. 153–161.

[51] B. R. Brown. "The mechanism of thermal decarboxylation". *Q. Rev., Chem. Soc.* 5.2 (1951), p. 131.

[52] A. Radzicka and R. Wolfenden. "A proficient enzyme". *Science* 267.5194 (1995), pp. 90–93.

[53] Frank Jordan and Hetalben Patel. "Catalysis in Enzymatic Decarboxylations: Comparison of Selected Cofactor-dependent and Cofactor-independent Examples." *ACS Catal.* 3.7 (2013), pp. 1601–1617.

[54] C. F. Hawkins, A. Borges, and R. N. Perham. "A common structural motif in thiamin pyrophosphate-binding enzymes". *FEBS Lett.* 255.1 (1989), pp. 77–82.

[55] R. Breslow. "The mechanism of thiamine action: predictions from model experiments". *Ann. N. Y. Acad. Sci.* 98 (1962), pp. 445–452.

[56] Ronald Kluger and Kai Tittmann. "Thiamin diphosphate catalysis: enzymic and nonenzymic covalent intermediates". *Chem Rev* 108.6 (2008), pp. 1797–1833.

[57] Ronald Kluger. "Decarboxylation, CO 2 and the Reversion Problem". *Acc. Chem. Res.* 48.11 (2015), pp. 2843–2849.

[58] R. A. Laursen and F. H. Westheimer. "The active site of acetoacetate decarboxylase". *J. Am. Chem. Soc.* 88.14 (1966), pp. 3426–3430.

[59] A. C. Eliot and J. F. Kirsch. "Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations". *Annu. Rev. Biochem.* 73 (2004), pp. 383–415.

[60] Tingfeng Li, Lu Huo, Christopher Pulley, and Aimin Liu. "Decarboxylation mechanisms in biological system." *Bioorg. Chem.* 43 (2012), pp. 2–14.

[61] Shutong Xu, Wenjing Li, Junjun Zhu, Rong Wang, Zheng Li, Guo-Liang Xu, and Jianping Ding. "Crystal structures of isoorotate decarboxylases reveal a novel catalytic mechanism of 5-carboxyl-uracil decarboxylation and shed light on the search for DNA decarboxylase." *Cell Res* 23.11 (2013), pp. 1296–1309.

[62] W. Cui, J. G. DeWitt, S. M. Miller, and W. Wu. "No metal cofactor in orotidine 5'-monophosphate decarboxylase." *Biochem Biophys Res Commun* 259.1 (1999), pp. 133–135.

[63] B. G. Miller, J. A. Smiley, S. A. Short, and R. Wolfenden. "Activity of yeast orotidine-5'-phosphate decarboxylase in the absence of metals". *J. Biol. Chem.* 274.34 (1999), pp. 23841–23843.

[64] Paul Wayne Landesman. "Design, synthesis and evaluation of potential inhibitors of pyrimidine biosynthesis: A mechanistic approach". PhD thesis. State University of New York at Buffalo, 1982.

[65] J B Bell and M E Jones. "Purification and characterization of yeast orotidine 5'-monophosphate decarboxylase overexpressed from plasmid PGU2." *The Journal of biological chemistry* 266 (19 1991), pp. 12662–12667.

[66] R. E. Handschuhmacher and C. A. Pasternak. "Inhibition of orotidylic acid decarboxylase, a primary site of carcinostasis by 6-azauracil". *Biochim. Biophys. Acta* 30.2 (1958), pp. 451–452.

[67] R. E. Handschuhmacher. "Orotidylic acid decarboxylase: inhibition studies with azauridine 5'-phosphate". *J. Biol. Chem.* 235 (1960), pp. 2917–2919.

[68] R. S. Brody and F. H. Westheimer. "The purification of orotidine-5'-phosphate decarboxylase from yeast by affinity chromatography". *J. Biol. Chem.* 254.10 (1979), pp. 4238–4244.

[69] M. J. Yablonski, D. A. Pasek, B. D. Han, M. E. Jones, and T. W. Traut. "Intrinsic activity and stability of bifunctional human UMP synthase and its two separate catalytic domains, orotate phosphoribosyltransferase and orotidine-5'-phosphate decarboxylase." *J. Biol. Chem.* 271.18 (1996), pp. 10704–10708.

[70] B. Beak and B. Siegel. "Letter: Mechanism of decarboxylation of 1,3-dimethylorotic acid. A possible role for orotate decarboxylase." *J. Am. Chem. Soc.* 95.23 (1973), pp. 7919–7920.

[71] P. Beak and B. Siegel. "Mechanism of decarboxylation of 1,3-dimethylorotic acid. A model for orotidine 5'-phosphate decarboxylase". *J. Am. Chem. Soc.* 98.12 (1976), pp. 3601–3606.

[72] Weiming Wu, Ley-han Antonio, Freeman M. Wong, Travis J. Austin, and Susan M. Miller. "Decarboxylation of 1,3-dimethylorotic acid revisited: Determining the role of N-1". *Bioorganic & Medicinal Chemistry Letters* 7.20 (1997), pp. 2623–2628.

[73] M A Rishavy and W W Cleland. "Determination of the mechanism of orotidine 5'-monophosphate decarboxylase by isotope effects." *Biochemistry* 39 (16 2000), pp. 4569–4574.

[74] J. K. Lee and K. N. Houk. "A proficient enzyme revisited: the predicted mechanism for orotidine monophosphate decarboxylase." *Science* 276.5314 (1997), pp. 942–945.

[75] Richard B. Silverman and Michael P. Groziak. "Model chemistry for a covalent mechanism of action of orotidine 5'-phosphate decarboxylase". *J. Am. Chem. Soc.* 104.23 (1982), pp. 6434–6439.

[76] T. C. Appleby, C. Kinsland, T. P. Begley, and S. E. Ealick. "The crystal structure and mechanism of orotidine 5'-monophosphate decarboxylase." *Proc. Natl. Acad. Sci. USA* 97.5 (2000), pp. 2005–2010.

[77] T. P. Begley, T. C. Appleby, and S. E. Ealick. "The structural basis for the remarkable catalytic proficiency of orotidine 5'-monophosphate decarboxylase". *Curr. Opin. Struct. Biol.* 10.6 (2000), pp. 711–718.

[78] K. Shostak and M. E. Jones. "Orotidylate decarboxylase: insights into the catalytic mechanism from substrate specificity studies." *Biochemistry* 31.48 (1992), pp. 12155–12161.

[79] Courtney L. Stanton, I-Feng W. Kuo, Christopher J. Mundy, Teodoro Laino, and K. N. Houk. "QMMM metadynamics study of the direct decarboxylation mechanism for orotidine-5'-monophosphate decarboxylase using two different QM regions: acceleration too small to explain rate of enzyme catalysis." *J. Phys. Chem. B.* 111.43 (2007), pp. 12573–12581.

[80] B. W. Potvin, H. J. Stern, S. R. May, G. F. Lam, and R. S. Krooth. "Inhibition by barbituric acid and its derivatives of the enzymes in rat brain which participate in the synthesis of pyrimidine ribotides". *Biochem. Pharmacol.* 27.5 (1978), pp. 655–665.

[81] H. L. Levine, R. S. Brody, and F. H. Westheimer. "Inhibition of orotidine-5'-phosphate decarboxylase by 1-(5'-phospho-beta-d-ribofuranosyl)barbituric acid, 6-azauridine 5'-phosphate, and uridine 5'-phosphate." *Biochemistry* 19.22 (1980), pp. 4993–4999.

[82] Jeffrey A. Smiley and Mary Ellen Jones. "A unique catalytic and inhibitor-binding role for Lys93 of yeast orotidylate decarboxylase". *Biochemistry* 31.48 (1992), pp. 12162–12168.

[83] Masahiro Fujihashi, Angelica M. Bello, Ewa Poduch, Lianhu Wei, Subhash C. Annedi, Emil F. Pai, and Lakshmi P. Kotra. "An unprecedented twist to ODCase catalytic activity." *J. Am. Chem. Soc.* 127.43 (2005), pp. 15048–15050.

[84] Daniel Heinrich, Ulf Diederichsen, and Markus Georg Rudolph. "Lys314 is a nucleophile in non-classical reactions of orotidine-5'-monophosphate decarboxylase." *Chemistry* 15.27 (2009), pp. 6619–6625.

[85] M. Lundberg, M. R. Blomberg, and P. E. Siegbahn. "Density functional models of the mechanism for decarboxylation in orotidine decarboxylase". *J Mol Model* 8.4 (2002), pp. 119–130.

[86] A. Warshel. "Electrostatic origin of the catalytic power of enzymes and the role of preorganized active sites." *J. Biol. Chem.* 273.42 (1998), pp. 27035–27038.

[87] Jeffrey A. Smiley, Piotr Paneth, Marion H. O'Leary, Juliette B. Bell, and Mary Ellen Jones. "Investigation of the enzymic mechanism of yeast orotidine-5'-monophosphate decarboxylase using carbon-13 kinetic isotope effects". *Biochemistry* 30.25 (1991), pp. 6216–6223.

[88] Joel I. Ehrlich, Chi-Ching Hwang, Paul F. Cook, and John S. Blanchard. "Evidence for a Stepwise Mechanism of OMP Decarboxylase". *J. Am. Chem. Soc.* 121.29 (1999), pp. 6966–6967.

[89] T. S. Lee, L. T. Chong, J. D. Chodera, and P. A. Kollman. "An alternative explanation for the catalytic proficiency of orotidine 5'-phosphate decarboxylase." *J Am Chem Soc* 123.51 (2001), pp. 12837–12848.

[90] Kui K. Chan, B McKay Wood, Alexander A. Fedorov, Elena V. Fedorov, Heidi J. Imker, Tina L. Amyes, John P. Richard, Steven C. Almo, and John A. Gerlt. "Mechanism of the orotidine 5'-monophosphate decarboxylase-catalyzed reaction: evidence for substrate destabilization." *Biochemistry* 48.24 (2009), pp. 5518–5531.

[91] P. Harris, J. C. Navarro Poulsen, K. F. Jensen, and S. Larsen. "Structural basis for the catalytic mechanism of a proficient enzyme: orotidine 5'-monophosphate decarboxylase". *Biochemistry* 39.15 (2000), pp. 4217–4224.

[92] Keiji Tokuoka, Yukiko Kusakari, Sudaratana R Krungkrai, Hiroyoshi Matsumura, Yasushi Kai, Jerapan Krungkrai, Toshihiro Horii, and Tsuyoshi Inoue. "Structural basis for the decarboxylation of orotidine 5'-monophosphate (OMP) by Plasmodium falciparum OMP decarboxylase." *Journal of biochemistry* 143 (1 2008), pp. 69–78.

[93] J. B. Bell, M. E. Jones, and CW Carter Jr. “Crystallization of yeast orotidine 5’-monophosphate decarboxylase complexed with 1-(5’-phospho-beta-D-ribofuranosyl) barbituric acid.” *Proteins* 9.2 (1991), pp. 143–151.

[94] B. G. Miller, A. M. Hassell, R. Wolfenden, M. V. Milburn, and S. A. Short. “Anatomy of a proficient enzyme: the structure of orotidine 5’-monophosphate decarboxylase in the presence and absence of a potential transition state analog”. *Proc. Natl. Acad. Sci. U.S.A.* 97.5 (2000), pp. 2011–2016.

[95] B. G. Miller, A. M. Hassell, M. V. Milburn, and S. A. Short. “Crystallization of native and selenomethionyl yeast orotidine 5’-phosphate decarboxylase.” *Acta Crystallogr D Biol Crystallogr* 56.Pt 4 (2000), pp. 472–474.

[96] D. W. Banner, A. c. Bloomer, G. A. Petsko, D. C. Phillips, and I. A. Wilson. “Atomic coordinates for triose phosphate isomerase from chicken muscle”. *Biochem. Biophys. Res. Commun.* 72.1 (1976), pp. 146–155.

[97] P. Harris, J. C. Poulsen, K. F. Jensen, and S. Larsen. “Substrate binding induces domain movements in orotidine 5’-monophosphate decarboxylase”. *J. Mol. Biol.* 318.4 (2002), pp. 1019–1029.

[98] Sun Hur and Thomas C. Bruice. “Molecular dynamic study of orotidine-5’-monophosphate decarboxylase in ground state and in intermediate state: a role of the 203-218 loop dynamics.” *Proc. Natl. Acad. Sci. U S A* 99.15 (2002), pp. 9668–9673.

[99] K. Stierand and M. Rarey. “Drawing the PDB: Protein-Ligand Complexes in Two Dimensions”. *ACS Med Chem Lett* 1.9 (2010), pp. 540–545.

[100] J. P. Richard, T. L. Amyes, B. Goryanova, and X. Zhai. “Enzyme architecture: on the importance of being in a protein cage”. *Curr Opin Chem Biol* 21 (2014), pp. 1–10.

[101] A. C. Reyes, D. C. Plache, A. P. Koudelka, T. L. Amyes, J. A. Gerlt, and J. P. Richard. “Enzyme Architecture: Breaking Down the Catalytic Cage that Activates Orotidine 5’-Monophosphate Decarboxylase for Catalysis”. *J. Am. Chem. Soc.* (2018).

[102] M Merced Malabanan, Tina L. Amyes, and John P. Richard. “A role for flexible loops in enzyme catalysis.” *Curr. Opin. Struct. Biol.* 20.6 (2010), pp. 702–710.

[103] Archie C. Reyes, Xiang Zhai, Kelsey T. Morgan, Christopher J. Reinhardt, Tina L. Amyes, and John P. Richard. “The activating oxydianion binding domain for enzyme-catalyzed proton transfer, hydride transfer, and decarboxylation: specificity and enzyme architecture.” *J. Am. Chem. Soc.* 137.3 (2015), pp. 1372–1382.

[104] Krisztina Toth, Tina L. Amyes, B McKay Wood, Kui K. Chan, John A. Gerlt, and John P. Richard. “An examination of the relationship between active site loop size and thermodynamic activation parameters for orotidine 5’-monophosphate decarboxylase from mesophilic and thermophilic organisms.” *Biochemistry* 48.33 (2009), pp. 8006–8013.

[105] B McKay Wood, Tina L. Amyes, Alexander A. Fedorov, Elena V. Fedorov, Andrew Shabila, Steven C. Almo, John P. Richard, and John A. Gerlt. “Conformational changes in orotidine 5’-monophosphate decarboxylase: "remote" residues that stabilize the active conformation.” *Biochemistry* 49.17 (2010), pp. 3514–3516.

[106] B. G. Miller, M. J. Snider, R. Wolfenden, and S. A. Short. “Dissecting a charged network at the active site of orotidine-5’-phosphate decarboxylase.” *J. Biol. Chem.* 276.18 (2001), pp. 15174–15176.

[107] David J. T. Porter and Steven A. Short. “Yeast Orotidine-5’-Phosphate Decarboxylase: Steady-State and Pre-Steady-State Analysis of the Kinetic Mechanism of Substrate Decarboxylation”. *Biochemistry* 39.38 (2000), pp. 11788–11800.

[108] W. R. P. Novak, K. H. J. West, L. M. D. Kirkman, and G. S. Brandt. “Re-refinement of Plasmodium falciparum orotidine 5’-monophosphate decarboxylase provides a clearer picture of an important malarial drug target”. *Acta Crystallogr F Struct Biol Commun* 74.Pt 10 (2018), pp. 664–668.

[109] N. Wu, Y. Mo, J. Gao, and E. F. Pai. "Electrostatic stress in catalysis: structure and mechanism of the enzyme orotidine monophosphate decarboxylase." *Proc. Natl. Acad. Sci. USA* 97.5 (2000), pp. 2017–2022.

[110] William P. Jencks. "Binding Energy, Specificity, and Enzymic Catalysis: The Circe Effect". *Advances in Enzymology and Related Areas of Molecular Biology* (1975), pp. 219–410.

[111] Arieh Warshel, Marek Trajbl, Jordi Villa, and Jan Florián. "Remarkable Rate Enhancement of Orotidine 5'-Monophosphate Decarboxylase Is Due to Transition-State Stabilization Rather Than to Ground-State Destabilization". *Biochemistry* 39.48 (2000), pp. 14728–14738.

[112] A. Warshel, P. K. Sharma, M. Kato, Y. Xiang, H. Liu, and M. H. Olsson. "Electrostatic basis for enzyme catalysis". *Chem. Rev.* 106.8 (2006), pp. 3210–3235.

[113] G. Jindal and A. Warshel. "Misunderstanding the preorganization concept can lead to confusions about the origin of enzyme catalysis". *Proteins* 85.12 (2017), pp. 2157–2161.

[114] G. Jindal, B. Ramachandran, R. P. Bora, and A. Warshel. "Exploring the Development of Ground-State Destabilization and Transition-State Stabilization in Two Directed Evolution Paths of Kemp Eliminases". *ACS Catal* 7.5 (2017), pp. 3301–3305.

[115] L. Pauling. "Nature of forces between large molecules of biological interest." *Nature* 161.4097 (1948), pp. 707–709.

[116] Masahiro Fujihashi, Lianhu Wei, Lakshmi P. Kotra, and Emil F. Pai. "Structural Characterization of the Molecular Events during a Slow Substrate–Product Transition in Orotidine 5'-Monophosphate Decarboxylase". *Journal of Molecular Biology* 387.5 (2009), pp. 1199–1210.

[117] M. Fujihashi, T. Ishida, S. Kuroda, L. P. Kotra, E. F. Pai, and K. Miki. "Substrate distortion contributes to the catalysis of orotidine 5'-monophosphate decarboxylase". *J. Am. Chem. Soc.* 135.46 (2013), pp. 17432–17443.

[118] Ning Wu, Wanda Gillon, and Emil F. Pai. "Mapping the active site-ligand interactions of orotidine 5'-monophosphate decarboxylase by crystallography." *Biochemistry* 41.12 (2002), pp. 4002–4011.

[119] Vanessa Iiams, Bijoy J. Desai, Alexander A. Fedorov, Elena V. Fedorov, Steven C. Almo, and John A. Gerlt. "Mechanism of the orotidine 5'-monophosphate decarboxylase-catalyzed reaction: importance of residues in the orotate binding site." *Biochemistry* 50.39 (2011), pp. 8497–8507.

[120] Dayne West, Chae Un Kim, Chingkuang Tu, Arthur H. Robbins, Sol M. Gruner, David N. Silverman, and Robert McKenna. "Structural and kinetic effects on changes in the CO(2) binding pocket of human carbonic anhydrase II." *Biochemistry* 51.45 (2012), pp. 9156–9163.

[121] D. Lee, J. Lee, and C. Seok. "What stabilizes close arginine pairing in proteins?" *Phys Chem Chem Phys* 15.16 (2013), pp. 5844–5853.

[122] Bijoy J. Desai, B McKay Wood, Alexander A. Fedorov, Elena V. Fedorov, Bogdana Goryanova, Tina L. Amyes, John P. Richard, Steven C. Almo, and John A. Gerlt. "Conformational changes in orotidine 5'-monophosphate decarboxylase: a structure-based explanation for how the 5'-phosphate group activates the enzyme." *Biochemistry* 51.43 (2012), pp. 8665–8678.

[123] Shonoi A. Barnett, Tina L. Amyes, B McKay Wood, John A. Gerlt, and John P. Richard. "Activation of R235A mutant orotidine 5'-monophosphate decarboxylase by the guanidinium cation: effective molarity of the cationic side chain of Arg235." *Biochemistry* 49.5 (2010), pp. 824–826.

[124] Bogdana Goryanova, Tina L. Amyes, John A. Gerlt, and John P. Richard. "OMP decarboxylase: phosphodianion binding energy is used to stabilize a vinyl carbanion intermediate." *J. Am. Chem. Soc.* 133.17 (2011), pp. 6545–6548.

[125] Tina L. Amyes, Shonoi A. Ming, Lawrence M. Goldman, B McKay Wood, Bijoy J. Desai, John A. Gerlt, and John P. Richard. "Orotidine 5'-monophosphate decarboxylase: transition state stabilization from remote protein-phosphodianion interactions." *Biochemistry* 51.23 (2012), pp. 4630–4632.

[126] Krisztina Spong, Tina L. Amyes, and John P. Richard. "Enzyme architecture: the activating oxydianion binding domain for orotidine 5'-monophophate decarboxylase." *J. Am. Chem. Soc.* 135.49 (2013), pp. 18343–18346.

[127] Bogdana Goryanova, Lawrence M. Goldman, Tina L. Amyes, John A. Gerlt, and John P. Richard. "Role of a guanidinium cation-phosphodianion pair in stabilizing the vinyl carbanion intermediate of orotidine 5'-phosphate decarboxylase-catalyzed reactions." *Biochemistry* 52.42 (2013), pp. 7500–7511.

[128] Lawrence M. Goldman, Tina L. Amyes, Bogdana Goryanova, John A. Gerlt, and John P. Richard. "Enzyme architecture: deconstruction of the enzyme-activating phosphodianion interactions of orotidine 5'-monophosphate decarboxylase." *J. Am. Chem. Soc.* 136.28 (2014), pp. 10156–10165.

[129] T. L. Amyes, M. M. Malabanan, X. Zhai, A. C. Reyes, and Richard J.P. "Enzyme activation through the utilization of intrinsic dianion binding energy". *Protein Engineering, Design & Selection* (2016).

[130] T. L. Amyes, A. C. O'Donoghue, and J. P. Richard. "Contribution of phosphate intrinsic binding energy to the enzymatic rate acceleration for triosephosphate isomerase". *J. Am. Chem. Soc.* 123.45 (2001), pp. 11325–11326.

[131] Tina L. Amyes, John P. Richard, and James J. Tait. "Activation of orotidine 5'-monophosphate decarboxylase by phosphite dianion: the whole substrate is the sum of two parts." *J. Am. Chem. Soc.* 127.45 (2005), pp. 15708–15709.

[132] W. Y. Tsang, T. L. Amyes, and J. P. Richard. "A substrate in pieces: allosteric activation of glycerol 3-phosphate dehydrogenase (NAD+) by phosphite dianion". *Biochemistry* 47.16 (2008), pp. 4575–4582.

[133] Shonoi A. Barnett, Tina L. Amyes, Bryant M. Wood, John A. Gerlt, and John P. Richard. "Dissecting the total transition state stabilization provided by amino acid side chains at orotidine 5'-monophosphate decarboxylase: a two-part substrate approach." *Biochemistry* 47.30 (2008), pp. 7785–7787.

[134] T. L. Amyes and J. P. Richard. "Enzymatic catalysis of proton transfer at carbon: activation of triosephosphate isomerase by phosphite dianion". *Biochemistry* 46.19 (2007), pp. 5841–5854.

[135] Bogdana Goryanova, Krisztina Spong, Tina L. Amyes, and John P. Richard. "Catalysis by orotidine 5'-monophosphate decarboxylase: effect of 5-fluoro and 4'-substituents on the decarboxylation of two-part substrates." *Biochemistry* 52.3 (2013), pp. 537–546.

[136] B. M. Wood, K. K. Chan, T. L. Amyes, J. P. Richard, and J. A. Gerlt. "Mechanism of the orotidine 5'-monophosphate decarboxylase-catalyzed reaction: effect of solvent viscosity on kinetic constants". *Biochemistry* 48.24 (2009), pp. 5510–5517.

[137] Annette Sievers and Richard Wolfenden. "Equilibrium of formation of the 6-carbanion of UMP, a potential intermediate in the action of OMP decarboxylase." *J. Am. Chem. Soc.* 124.47 (2002), pp. 13986–13987.

[138] Wing-Yin Tsang, B McKay Wood, Freeman M. Wong, Weiming Wu, John A. Gerlt, Tina L. Amyes, and John P. Richard. "Proton transfer from C-6 of uridine 5'-monophosphate catalyzed by orotidine 5'-monophosphate decarboxylase: formation and stability of a vinyl carbanion intermediate and the effect of a 5-fluoro substituent." *J. Am. Chem. Soc.* 134.35 (2012), pp. 14580–14594.

[139] Tina L. Amyes, Bryant M. Wood, Kui Chan, John A. Gerlt, and John P. Richard. "Formation and stability of a vinyl carbanion at the active site of orotidine 5'-monophosphate decarboxylase: pKa of the C-6 proton of enzyme-bound UMP." *J. Am. Chem. Soc.* 130.5 (2008), pp. 1574–1575.

[140] D. Kern, G. Kern, H. Neef, K. Tittmann, M. Killenberg-Jabs, C. Wikner, G. Schneider, and G. Hubner. "How thiamine diphosphate is activated in enzymes". *Science* 275.5296 (1997), pp. 67–70.

[141] J. G. Marblestone, S. C. Edavettal, Y. Lim, P. Lim, X. Zuo, and T. R. Butt. "Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced expression and solubility with SUMO". *Protein Sci.* 15.1 (2006), pp. 182–189.

[142] Ewa Poduch, Lianhu Wei, Emil F. Pai, and Lakshmi P. Kotra. "Structural diversity and plasticity associated with nucleotides targeting orotidine monophosphate decarboxylase." *J. Med. Chem.* 51.3 (2008), pp. 432–438.

[143] Meena K. Purohit, Ewa Poduch, Lianhu William Wei, Ian Edward Crandall, Terrence To, Kevin C. Kain, Emil F. Pai, and Lakshmi P. Kotra. "Novel cytidine-based orotidine-5'-monophosphate decarboxylase inhibitors with an unusual twist." *J. Med. Chem.* 55.22 (2012), pp. 9988–9997.

[144] A. V. Hill. "The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curve". *J. Physiol.* (1910).

[145] Qiang Cui and Martin Karplus. "Allostery and cooperativity revisited". *Protein Science* 17.8 (2008), pp. 1295–1307.

[146] Alan Brown. "Analysis of Cooperativity by Isothermal Titration Calorimetry". *International Journal of Molecular Sciences* 10.8 (2009), pp. 3457–3477.

[147] Carol M. Porter and Brian G. Miller. "Cooperativity in monomeric enzymes with single ligand-binding sites". *Bioorganic Chemistry* 43 (2012), pp. 44–50.

[148] K. Schroder-Tittmann, D. Meyer, J. Arens, C. Wechsler, M. Tietzel, R. Golbik, and K. Tittmann. "Alternating sites reactivity is a common feature of thiamin diphosphate-dependent enzymes as evidenced by isothermal titration calorimetry studies of substrate binding". *Biochemistry* 52.15 (2013), pp. 2505–2507.

[149] S. G. Jacchieri. "Cooperative Binding of Transition States in Enzymatic Catalysis". *International Journal of Quantum Chemistry* (1989).

[150] J. Monod, J. Wyman, and J. P. changeux. "On the nature of allosteric transitions: A plausible Model". *J. Mol. Biol.* 12 (1965), pp. 88–118.

[151] D. E. Koshland, G. Nemethy, and D. Filmer. "Comparison of experimental binding data and theoretical models in proteins containing subunits". *Biochemistry* 5.1 (1966), pp. 365–385.

[152] Ilian Jelesarov and Hans Rudolf Bosshard. "Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition". *Journal of Molecular Recognition* 12.1 (1999), pp. 3–18.

[153] Sonia Vega, Olga Abian, and Adrian Velazquez-Campoy. "A unified framework based on the binding polynomial for characterizing biological systems by isothermal titration calorimetry". *Methods* 76 (2015), pp. 99–115.

[154] M. Sundaralingam. "Structure and conformation of nucleosides and nucleotides and their analogs as determined by x-ray diffraction". *Ann. N. Y. Acad. Sci.* 255 (1975), pp. 3–42.

[155] M. U. Schmidt, J. Bruning, J. Glinnemann, M. W. Hutzler, P. Morschel, S. N. Ivashevskaya, J. van de Streek, D. Braga, L. Maini, M. R. Chierotti, and R. Gobetto. "The thermodynamically stable form of solid barbituric acid: the enol tautomer". *Angew. Chem. Int. Ed. Engl.* 50.34 (2011), pp. 7924–7926.

[156] Brian G. Miller and Richard Wolfenden. "Catalytic proficiency: the unusual case of OMP decarboxylase." *Annu. Rev. Biochem.* 71 (2002), pp. 847–885.

[157] James M. Holton. "A beginner's guide to radiation damage." *J Synchrotron Radiat* 16.Pt 2 (2009), pp. 133–142.

[158] Kristin A. Sutton, Paul J. Black, Kermit R. Mercer, Elspeth F. Garman, Robin L. Owen, Edward H. Snell, and William A. Bernhard. "Insights into the mechanism of X-ray-induced disulfide-bond cleavage in lysozyme crystals based on EPR, optical absorption and X-ray diffraction studies." *Acta Crystallogr. D Biol. Crystallogr.* 69.Pt 12 (2013), pp. 2381–2394.

[159] B. Hao, W. Gong, T. K. Ferguson, C. M. James, J. A. Krzycki, and M. K. Chan. "A new UAG-encoded residue in the structure of a methanogen methyltransferase". *Science* 296.5572 (2002), pp. 1462–1466.

[160] Heinz Neumann, Sew Y. Peak-Chew, and Jason W. Chin. "Genetically encoding N(epsilon)-acetyllysine in recombinant proteins." *Nat Chem Biol* 4.4 (2008), pp. 232–234.

[161] Heinz Neumann, Susan M. Hancock, Ruth Buning, Andrew Routh, Lynda Chapman, Joanna Somers, Tom Owen-Hughes, John van Noort, Daniela Rhodes, and Jason W. Chin. "A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation." *Mol Cell* 36.1 (2009), pp. 153–163.

[162] J. A. Gerlt, M. M. Kreevoy, W. Cleland, and P. A. Frey. "Understanding enzymic catalysis: the importance of short, strong hydrogen bonds". *Chem. Biol.* 4.4 (1997), pp. 259–267.

[163] W. W. Cleland, P. A. Frey, and J. A. Gerlt. "The low barrier hydrogen bond in enzymatic catalysis". *J. Biol. Chem.* 273.40 (1998), pp. 25529–25532.

[164] W. W. Cleland. "Low-barrier hydrogen bonds and enzymatic catalysis". *Arch. Biochem. Biophys.* 382.1 (2000), pp. 1–5.

[165] Elspeth F. Garman and Colin Nave. "Radiation damage in protein crystals examined under various conditions by different methods." *J Synchrotron Radiat* 16.Pt 2 (2009), pp. 129–132.

[166] Elspeth F. Garman. "Radiation damage in macromolecular crystallography: what is it and why should we care?" *Acta Crystallogr D* 66.4 (2010), pp. 339–351.

[167] Krisztina Toth, Tina L. Amyes, Bryant M. Wood, Kui Chan, John A. Gerlt, and John P. Richard. "Product deuterium isotope effect for orotidine 5'-monophosphate decarboxylase: evidence for the existence of a short-lived carbanion intermediate." *J. Am. Chem. Soc.* 129.43 (2007), pp. 12946–12947.

[168] Jeremy L. Van Vleet, Laurie A. Reinhardt, Brian G. Miller, Annette Sievers, and W Wallace Cleland. "Carbon isotope effect study on orotidine 5'-monophosphate decarboxylase: support for an anionic intermediate." *Biochemistry* 47.2 (2008), pp. 798–803.

[169] Tina L. Amyes and John P. Richard. "Specificity in transition state binding: the Pauling model revisited." *Biochemistry* 52.12 (2013), pp. 2021–2035.

[170] Masoud Kazemi, Fahmi Himo, and Johan Åqvist. "Enzyme catalysis by entropy without Circe effect". *Proceedings of the National Academy of Sciences* 113.9 (2016), pp. 2406–2411.

[171] A. Warshel, J. Florián, M. Strajbl, and J. Villà. "Circe effect versus enzyme preorganization: what can be learned from the structure of the most proficient enzyme?" *Chembiochem.* 2.2 (2001), pp. 109–111.

[172] Angelica M. Bello, Ewa Poduch, Masahiro Fujihashi, Merhnaz Amani, Yan Li, Ian Crandall, Raymond Hui, Ping I. Lee, Kevin C. Kain, Emil F. Pai, and Lakshmi P. Kotra. "A potent, covalent inhibitor of orotidine 5'-monophosphate decarboxylase with antimalarial activity." *J. Med. Chem.* 50.5 (2007), pp. 915–921.

[173] Brian P. Callahan and Richard Wolfenden. "OMP decarboxylase: an experimental test of electrostatic destabilization of the enzyme-substrate complex." *J. Am. Chem. Soc.* 126.45 (2004), pp. 14698–14699.

[174] Charles A Lewis Jr and Richard Wolfenden. "Indiscriminate binding by orotidine 5'-phosphate decarboxylase of uridine 5'-phosphate derivatives with bulky anionic c6 substituents." *Biochemistry* 46.46 (2007), pp. 13331–13343.

[175] Alexandra Vardi-Kilshtain, Dvir Doron, and Dan Thomas Major. "Quantum and classical simulations of orotidine monophosphate decarboxylase: support for a direct decarboxylation mechanism." *Biochemistry* 52.25 (2013), pp. 4382–4390.

[176] Ronald Kluger, Graeme W. Howe, and Scott O.C. Mundle. "Avoiding CO2 in Catalysis of Decarboxylation". *Advances in Physical Organic Chemistry* (2013), pp. 85–128.

[177] Ronald Kluger and Steven Rathgeber. "Catalyzing separation of carbon dioxide in thiamin diphosphate-promoted decarboxylation." *FEBS J* 275.24 (2008), pp. 6089–6100.

[178] Jiali Gao, Shuhua Ma, Dan T. Major, Kwangho Nam, Jingzhi Pu, and Donald G. Truhlar. "Mechanisms and free energies of enzymatic reactions." *Chem. Rev.* 106.8 (2006), pp. 3188–3209.

[179] D. T. Major and J. Gao. "An Integrated Path Integral and Free-Energy Perturbation-Umbrella Sampling Method for Computing Kinetic Isotope Effects of Chemical Reactions in Solution and in Enzymes". *J Chem Theory Comput* 3.3 (2007), pp. 949–960.

[180] Graeme W. Howe and Ronald Kluger. "Decarboxylation without CO2: why bicarbonate forms directly as trichloroacetate is converted to chloroform." *J. Org. Chem.* 79.22 (2014), pp. 10972–10980.

[181] Graeme W. Howe, Adelle A. Vandersteen, and Ronald Kluger. "How Acid-Catalyzed Decarboxylation of 2,4-Dimethoxybenzoic Acid Avoids Formation of Protonated CO2". *Journal of the American Chemical Society* 138.24 (2016), pp. 7568–7573.

[182] E. S. Polovnikova, M. J. McLeish, E. A. Sergienko, J. T. Burgner, N. L. Anderson, A. K. Bera, F. Jordan, G. L. Kenyon, and M. S. Hasson. "Structural and kinetic analysis of catalysis by a thiamin diphosphate-dependent enzyme, benzoylformate decarboxylase". *Biochemistry* 42.7 (2003), pp. 1820–1830.

[183] A. Yep, G. L. Kenyon, and M. J. McLeish. "Saturation mutagenesis of putative catalytic residues of benzoylformate decarboxylase provides a challenge to the accepted mechanism". *Proc. Natl. Acad. Sci. U.S.A.* 105.15 (2008), pp. 5733–5738.

[184] V. Cody and I. Thomas Kalman. "Conformational Analysis of 6-Substituted Uridine Analogues: Crystal Structures of Uridine-6-Thiocarboxamide and 6-Cyanouridine". *Nucleosides and Nucleotides* (1985).

[185] J. Lin, E. Pozharski, and M. A. Wilson. "Short Carboxylic Acid-Carboxylate Hydrogen Bonds Can Have Fully Localized Protons". *Biochemistry* 56.2 (2017), pp. 391–402.

[186] R. A. Engh and R. Huber. "Accurate bond and angle parameters for X-ray protein structure refinement". *Acta Crystallographica Section A* 47.4 (1991), pp. 392–400.

[187] J. D. Graham, A. M. Buytendyk, D. Wang, K. H. Bowen, and K. D. Collins. "Strong, low-barrier hydrogen bonds may be available to enzymes". *Biochemistry* 53.2 (2014), pp. 344–349.

[188] Anna Vladimirova, Yury Patskovsky, Alexander A. Fedorov, Jeffrey B. Bonanno, Elena V. Fedorov, Rafael Toro, Brandan Hillerich, Ronald D. Seidel, Nigel G J. Richards, Steven C. Almo, and Frank M. Raushel. "Substrate Distortion and the Catalytic Reaction Mechanism of 5-Carboxyvanillate Decarboxylase." *J Am Chem Soc* (2015).

[189] Markus Gerstel, Charlotte M. Deane, and Elspeth F. Garman. "Identifying and quantifying radiation damage at the atomic level." *J Synchrotron Radiat* 22.2 (2015), pp. 201–212.

[190] M. Fujihashi, K. Mito, E. F. Pai, and K. Miki. "Atomic resolution structure of the orotidine 5'-monophosphate decarboxylase product complex combined with surface plasmon resonance analysis: implications for the catalytic mechanism". *J. Biol. Chem.* 288.13 (2013), pp. 9011–9016.

[191] Masahiro Fujihashi, Jagjeet S. Mnpotra, Ram Kumar Mishra, Emil F. Pai, and Lakshmi P. Kotra. "Orotidine Monophosphate Decarboxylase – A Fascinating Workhorse Enzyme with Therapeutic Potential". *Journal of Genetics and Genomics* 42.5 (2015), pp. 221–234.

[192] Arren Bar-Even, Elad Noor, Yonatan Savir, Wolfram Liebermeister, Dan Davidi, Dan S. Tawfik, and Ron Milo. "The Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters". *Biochemistry* 50.21 (2011), pp. 4402–4410.

[193] Arren Bar-Even, Ron Milo, Elad Noor, and Dan S. Tawfik. "The Moderately Efficient Enzyme: Futile Encounters and Enzyme Floppiness". *Biochemistry* 54.32 (2015), pp. 4969–4977.

[194] C. Altona and M. Sundaralingam. "Conformational analysis of the sugar ring in nucleosides and nucleotides. A new description using the concept of pseudorotation". *J. Am. Chem. Soc.* 94.23 (1972), pp. 8205–8212.

[195] S. Ludtke, P. Neumann, K. M. Erixon, F. Leeper, R. Kluger, R. Ficner, and K. Tittmann. "Sub-ångström-resolution crystallography reveals physical distortions that enhance reactivity of a covalent enzymatic intermediate". *Nat Chem* 5.9 (2013), pp. 762–767.

[196] P. Neumann and K. Tittmann. "Marvels of enzyme catalysis at true atomic resolution: distortions, bond elongations, hidden flips, protonation states and atom identities". *Curr. Opin. Struct. Biol.* 29 (2014), pp. 122–133.

[197] Anja Lehwess-Litzmann, Piotr Neumann, Christoph Parthier, Stefan Luedtke, Ralph Golbik, Ralf Ficner, and Kai Tittmann. "Twisted Schiff base intermediates and substrate locale revise transaldolase mechanism". *Nat. Chem. Biol.* 7.10 (2011), pp. 678–684.

[198] J. P. Richard. "The enhancement of enzymatic rate accelerations by Brønsted acid-base catalysis". *Biochemistry* 37.13 (1998), pp. 4305–4309.

[199] K. Tittmann. "Sweet siblings with different faces: the mechanisms of FBP and F6P aldolase, transaldolase, transketolase and phosphoketolase revisited in light of recent structural data". *Bioorg. Chem.* 57 (2014), pp. 263–280.

[200] F. Lederer, S. M. Coutts, R. A. Laursen, and F. H. Westheimer. "Acetoacetate decarboxylase. Subunits and properties". *Biochemistry* 5.3 (1966), pp. 823–833.

[201] P. P. Lee and F. H. Westheimer. "Acetoacetate decarboxylase. Photochemical oxidation induced by iodoacetate". *Biochemistry* 5.3 (1966), pp. 834–837.

[202] S. Warren, B. Zerner, and F. H. Westheimer. "Acetoacetate decarboxylase. Identification of lysine at the active site". *Biochemistry* 5.3 (1966), pp. 817–823.

[203] B. Zerner, S. M. Coutts, F. Lederer, H. H. Waters, and F. H. Westheimer. "Acetoacetate decarboxylase. Preparation of the enzyme". *Biochemistry* 5.3 (1966), pp. 813–816.

[204] M. H. O'Leary and F. H. Westheimer. "Acetoacetate decarboxylase. Selective acetylation of the enzyme". *Biochemistry* 7.3 (1968), pp. 913–919.

[205] W. Tagaki, J. P. Guthrie, and F. H. Westheimer. "Acetoacetate decarboxylase. Reaction with acetopyruvate". *Biochemistry* 7.3 (1968), pp. 905–913.

[206] W. Tagaki and F. H. Westheimer. "Acetoacetate decarboxylase. Reassociation of subunits". *Biochemistry* 7.3 (1968), pp. 891–894.

[207] W. Tagaki and F. H. Westheimer. "Acetoacetate decarboxylase. The molecular weight of the enzyme and subunits". *Biochemistry* 7.3 (1968), pp. 895–900.

[208] W. Tagaki and F. H. Westheimer. "Acetoacetate decarboxylase. Catalysis of hydrogen-deuterium exchange in acetone". *Biochemistry* 7.3 (1968), pp. 901–905.

[209] D. J. Petersen and G. N. Bennett. "Purification of acetoacetate decarboxylase from Clostridium acetobutylicum ATCC 824 and cloning of the acetoacetate decarboxylase gene in *Escherichia coli*". *Appl. Environ. Microbiol.* 56.11 (1990), pp. 3491–3498.

[210] G.A. Hamilton and F. H. Westheimer. “On the mechanism of the enzymatic decarboxylation of acetoacetate”. *J. Am. Chem. Soc.* (1959).

[211] P. A. Frey, F. C. Kokesh, and F. H. Westheimer. “A reporter group at the active site of acetoacetate decarboxylase. I. Ionization constant of the nitrophenol”. *J. Am. Chem. Soc.* 93.26 (1971), pp. 7266–7269.

[212] F. C. Kokesh and F. H. Westheimer. “A reporter group at the active site of acetoacetate decarboxylase. II. Ionization constant of the amino group”. *J. Am. Chem. Soc.* 93.26 (1971), pp. 7270–7274.

[213] D. E. Schmidt and F. H. Westheimer. “PK of the lysine amino group at the active site of acetoacetate decarboxylase”. *Biochemistry* 10.7 (1971), pp. 1249–1253.

[214] J. A. Gerlt. “Acetoacetate decarboxylase: hydrophobics, not electrostatics”. *Nat. Chem. Biol.* 5.7 (2009), pp. 454–455.

[215] H. Ishikita. “Origin of the pKa shift of the catalytic lysine in acetoacetate decarboxylase”. *FEBS Lett.* 584.15 (2010), pp. 3464–3468.

[216] M. S. Neece and I. Fridovich. “Acetoacetic decarboxylase. Activation by heat.” *J Biol Chem* 242.12 (1967), pp. 2939–2944.

15 Appendix

Table 2: **Data collection and model refinement statistics of measured *h*OMPD$_{WT}$ crystals in the pseudo resting-state and in complex with the decarboxylation product UMP**

	RS	UMP
Data collection		
Wavelength [Å]	0.7653	0.7653
Resolution range [Å]	42.45 - 0.95 (0.984 - 0.95)	42.46 - 1.0 (1.036 - 1.0)
Space group	$C222_1$	$C222_1$
Unit cell constants [Å]	77.396 116.562 61.958	77.389 116.572 61.968
Angles α, β, γ [°]	90, 90, 90	90, 90, 90
No. total reflections	736726 (58076)	385555 (36041)
No. unique reflections	172159 (17293)	146420 (14541)
Multiplicity	4.3 (3.4)	2.6 (2.5)
Completeness [%]	98.21 (99.37)	97.29 (97.56)
Mean $I/\sigma(I)$	18.51 (1.73)	15.00 (1.35)
Wilson B-factor	9.7	10.66
R-merge	0.03308 (0.6114)	0.03049 (0.7515)
R-meas	0.03715 (0.724)	0.03755 (0.9494)
R-pim	0.01638 (0.3777)	0.02153 (0.5721)
$CC_{1/2}$	0.999 (0.698)	1 (0.586)
CC*	1 (0.907)	1 (0.86)
Refinement		
Reflections used in refinement	172130 (17291)	146406 (14541)
Reflections used for R-free	8692 (904)	7365 (707)
R-work	0.1080 (0.3376)	0.1141 (0.4076)
R-free	0.1202 (0.3359)	0.1296 (0.4231)
CC(work)	0.974 (0.717)	0.982 (0.772)
CC(free)	0.971 (0.714)	0.979 (0.740)
Number of non-hydrogen atoms	2841	2955
macromolecules	2434	2506
ligands	30	70
solvent	377	379
Protein residues	258	257
RMS bond length [Å]	0.017	0.029
RMS angles [°]	1.92	2.65
Ramachandran favored [%]	98.42	98.81
Ramachandran allowed [%]	1.58	1.19
Ramachandran outliers [%]	0	0
Rotamer outliers [%]	1.87	3.26
Clashscore	14.09	16.57
Average B-factor	13.22	14.45
macromolecules	11.13	12.2
ligands	10.44	15.84
solvent	26.93	29.05

Table 3: **Data collection and model refinement statistics of measured *h*OMPD$_{WT}$ crystals in complex with 6-aza-UMP and 6-hydroxy-UMP**

	6-aza-UMP	6-hydroxy-UMP (BMP)
Data collection		
Wavelength [Å]	0.7293	0.7653
Resolution range [Å]	42.46 - 1.0 (1.036 - 1.0)	30.3 - 0.95 (0.984 - 0.95)
Space group	$C222_1$	$C222_1$
Unit cell constants [Å]	77.673 116.576 61.972	77.896 116.425 61.968
Angles α, β, γ [°]	90, 90, 90	90, 90, 90
No. total reflections	568296 (56271)	619037 (52593)
No. unique reflections	147568 (14767)	171326 (17215)
Multiplicity	3.9 (3.8)	3.6 (3.1)
Completeness [%]	97.67 (98.78)	97.14 (98.27)
Mean $I/\sigma(I)$	15.19 (1.59)	17.94 (2.12)
Wilson B-factor	9.36	8.26
R-merge	0.04599 (0.8462)	0.03575 (0.5231)
R-meas	0.05268 (0.9764)	0.04123 (0.632)
R-pim	0.02484 (0.4715)	0.01988 (0.3475)
$CC_{1/2}$	1 (0.588)	0.999 (0.76)
CC*	1 (0.861)	1 (0.929)
Refinement		
Reflections used in refinement	147521 (14764)	171183 (17204)
Reflections used for R-free	7338 (684)	8468 (905)
R-work	0.1166 (0.3173)	0.0992 (0.3037)
R-free	0.1337 (0.3286)	0.1121 (0.3217)
CC(work)	0.979 (0.757)	0.966 (0.859)
CC(free)	0.975 (0.698)	0.957 (0.867)
Number of non-hydrogen atoms	2782	3527
macromolecules	2352	2957
ligands	42	50
solvent	388	520
Protein residues	257	257
RMS bond length [Å]	0.01	0.01
RMS angles [°]	1.54	1.62
Ramachandran favored [%]	98.81	98.81
Ramachandran allowed [%]	1.19	1.19
Ramachandran outliers [%]	0	0
Rotamer outliers [%]	0.78	4.04
Clashscore	4.55	10.63
Average B-factor	13	11.61
macromolecules	10.73	9.52
ligands	17.91	17.92
solvent	26.2	22.88

Table 4: **Data collection and model refinement statistics of measured *h*OMPD$_{WT}$ crystals in complex with 6-amido-UMP in space group C222$_1$and P2$_1$**

	6-amido-UMP	6-amido-UMP
Data collection		
Wavelength [Å]		0.9763
Resolution range [Å]	44.81 - 1.2 (1.243 - 1.2)	44.82 - 1.2 (1.243 - 1.2)
Space group	$C222_1$	$P2_1$
Unit cell constants [Å]	77.735 116.302 62.174	70.012 62.151 69.843
Angles α, β, γ [°]	90, 90, 90	90, 112.475, 90
No. total reflections	569886 (46920)	572189 (46802)
No. unique reflections	85219 (7891)	161404 (14042)
Multiplicity	6.7 (5.9)	3.5 (3.3)
Completeness [%]	96.80 (90.55)	93.49 (81.73)
Mean $I/\sigma(I)$	13.40 (1.80)	12.95 (1.68)
Wilson B-factor	12.47	12.8
R-merge	0.07651 (0.9925)	0.05681 (0.7975)
R-meas	0.08301 (1.088)	0.06702 (0.9494)
R-pim	0.0318 (0.4362)	0.03518 (0.5078)
$CC_{1/2}$	0.999 (0.668)	0.999 (0.636)
CC*	1 (0.895)	1 (0.882)
Refinement		
Reflections used in refinement	85218 (7891)	161394 (14041)
Reflections used for R-free	4312 (406)	8090 (666)
R-work	0.1185 (0.2844)	0.1312 (0.2931)
R-free	0.1465 (0.3134)	0.1561 (0.2880)
CC(work)	0.973 (0.727)	0.969 (0.693)
CC(free)	0.962 (0.687)	0.966 (0.781)
Number of non-hydrogen atoms	2540	4711
macromolecules	2142	4140
ligands	38	48
solvent	360	523
Protein residues	257	516
RMS bond length [Å]	0.013	0.015
RMS angles [°]	1.77	1.85
Ramachandran favored [%]	98.41	98.43
Ramachandran allowed [%]	1.59	1.57
Ramachandran outliers [%]	0	0
Rotamer outliers [%]	1.28	1.57
Clashscore	8.12	6.85
Average B-factor	16.48	16.81
macromolecules	13.92	15.44
ligands	21.25	12.06
solvent	31.23	28.08

Table 5: **Data collection and model refinement statistics of measured *h*OMPD$_{314AcK}$ crystals in complex with 6-thiocarboxamido-UMP in space group C222$_1$and P2$_1$**

	6-thiocarboxamido-UMP	6-thiocarboxamido-UMP
Data collection		
Wavelength [Å]	0.9763	0.9763
Resolution range [Å]	42.43 - 1.2 (1.243 - 1.2)	42.41 - 1.2 (1.243 - 1.2)
Space group	$C222_1$	$P2_1$
Unit cell constants [Å]	77.485 116.145 62.133	69.806 62.125 69.754
Angles α, β, γ [°]	90, 90, 90	90, 112.579, 90
No. total reflections	220720 (18465)	572641 (47456)
No. unique reflections	84891 (8267)	168544 (16148)
Multiplicity	2.6 (2.2)	3.4 (2.9)
Completeness [%]	96.92 (95.37)	98.15 (94.50)
Mean $I/\sigma(I)$	12.01 (1.83)	13.71 (1.81)
Wilson B-factor	11.59	12.74
R-merge	0.04749 (0.4753)	0.05378 (0.6507)
R-meas	0.05898 (0.6122)	0.06398 (0.7978)
R-pim	0.03436 (0.3793)	0.03416 (0.4522)
$CC_{1/2}$	0.999 (0.747)	0.999 (0.699)
CC*	1 (0.925)	1 (0.907)
Refinement		
Reflections used in refinement	84885 (8267)	168531 (16148)
Reflections used for R-free	4235 (418)	8419 (787)
R-work	0.1209 (0.2696)	0.1301 (0.2907)
R-free	0.1521 (0.3057)	0.1589 (0.2975)
CC(work)	0.975 (0.798)	0.967 (0.695)
CC(free)	0.975 (0.794)	0.957 (0.664)
Number of non-hydrogen atoms	2567	4721
macromolecules	2196	4116
ligands	43	48
solvent	328	557
Protein residues	257	514
RMS bond length [Å]	0.014	0.015
RMS angles [°]	1.81	1.82
Ramachandran favored [%]	98.81	98.82
Ramachandran allowed [%]	1.19	1.18
Ramachandran outliers [%]	0	0
Rotamer outliers [%]	2.05	1.14
Clashscore	7.45	5.73
Average B-factor	15.36	16.88
macromolecules	13.3	15.34
ligands	19.84	12.11
solvent	28.56	28.69

Table 6: **Data collection and model refinement statistics of measured *h*OMPD$_{314AcK}$ crystals in the resting state and in complex with the reaction product UMP space group C222$_1$**

	RS	UMP
Data collection		
Wavelength [Å]	0.9763	0.8266
Resolution range [Å]	35.06 - 1.2 (1.243 - 1.2)	44.62 - 1.05 (1.088 - 1.05)
Space group	$C222_1$	$C222_1$
Unit cell constants [Å]	77.104 118.084 62.194	77.085 116.94 61.903
Angles α, β, γ [°]	90, 90, 90	90, 90, 90
No. total reflections	311835 (25427)	857397 (75482)
No. unique reflections	86575 (8463)	129563 (11454)
Multiplicity	3.6 (3.0)	6.6 (5.9)
Completeness [%]	97.61 (96.32)	97.79 (87.20)
Mean $I/\sigma(I)$	16.13 (1.50)	14.24 (1.40)
Wilson B-factor	18.01	11.15
R-merge	0.03392 (0.7343)	0.06871 (1.508)
R-meas	0.03953 (0.8787)	0.07456 (1.655)
R-pim	0.01974 (0.4735)	0.02863 (0.6698)
$CC_{1/2}$	0.999 (0.634)	0.999 (0.697)
CC*	1 (0.881)	1 (0.906)
Refinement		
Reflections used in refinement	86572 (8461)	127637 (11223)
Reflections used for R-free	4342 (405)	6327 (555)
R-work	0.1376 (0.3280)	0.1470 (0.4562)
R-free	0.1615 (0.3456)	0.1687 (0.4225)
CC(work)	0.960 (0.730)	0.968 (0.554)
CC(free)	0.951 (0.627)	0.963 (0.512)
Number of non-hydrogen atoms	2507	2589
macromolecules	2205	2218
ligands	22	44
solvent	280	327
Protein residues	255	256
RMS bond length [Å]	0.017	0.014
RMS angles [°]	1.96	1.89
Ramachandran favored [%]	97.17	98.79
Ramachandran allowed [%]	2.02	1.21
Ramachandran outliers [%]	0.81	0
Rotamer outliers [%]	2.49	0.41
Clashscore	12.59	8.1
Average B-factor	23.9	15.65
macromolecules	22.12	13.66
ligands	22.06	16.36
solvent	38.08	29.05

Table 7: **Data collection and model refinement statistics of measured *h*OMPD$_{314AcK}$ crystals in complex with OMP after 2 min soaking (Crystal09) in space group C222$_1$and P2$_1$**

	OMP-soaked-2min-09	OMP-soaked-2min-09
Data collection		
Wavelength [Å]	0.8266	0.8266
Resolution range [Å]	44.67 - 1.2 (1.243 - 1.2)	44.67 - 1.25 (1.295 - 1.25)
Space group	$C222_1$	$P2_1$
Unit cell constants [Å]	77.284 116.824 61.952	70.026 61.946 70.041
Angles α, β, γ [°]	90, 90, 90	90, 113.023, 90
No. total reflections	596526 (59205)	527656 (52396)
No. unique reflections	87513 (8675)	150342 (14943)
Multiplicity	6.8 (6.8)	3.5 (3.5)
Completeness [%]	99.89 (99.93)	98.78 (98.67)
Mean $I/\sigma(I)$	14.57 (1.41)	12.00 (1.24)
Wilson B-factor	12.6	13.5
R-merge	0.08956 (1.601)	0.07087 (1.181)
R-meas	0.09694 (1.732)	0.08376 (1.397)
R-pim	0.03678 (0.6552)	0.04416 (0.7379)
$CC_{1/2}$	0.999 (0.593)	0.999 (0.537)
CC*	1 (0.863)	1 (0.836)
Refinement		
Reflections used in refinement	87537 (8674)	150328 (14943)
Reflections used for R-free	4334 (415)	7589 (730)
R-work	0.1290 (0.3306)	0.1320 (0.3456)
R-free	0.1572 (0.3025)	0.1681 (0.3616)
CC(work)	0.974 (0.772)	0.975 (0.741)
CC(free)	0.970 (0.769)	0.959 (0.728)
Number of non-hydrogen atoms	2609	5145
macromolecules	2216	4362
ligands	67	126
solvent	326	657
Protein residues	256	512
RMS bond length [Å]	0.015	0.019
RMS angles [°]	1.91	1.99
Ramachandran favored [%]	98.79	98.39
Ramachandran allowed [%]	1.21	1.61
Ramachandran outliers [%]	0	0
Rotamer outliers [%]	0.83	1.49
Clashscore	7.18	7.64
Average B-factor	17.1	18.7
macromolecules	14.92	16.63
ligands	24.38	24.07
solvent	30.44	31.41

Table 8: **Data collection and model refinement statistics of measured *h*OMPD$_{WT}$ crystals in complex with OMP after 2 min soaking (Crystal10) in space group C222$_1$and P2$_1$**

	OMP-soaked-2min-10	OMP-soaked-2min-10
Data collection		
Wavelength [Å]	0.8266	0.8266
Resolution range [Å]	44.75 - 1.15 (1.191 - 1.15)	44.76 - 1.2 (1.243 - 1.2)
Space group	$C222_1$	$P2_1$
Unit cell constants [Å]	77.632 116.897 61.991	70.19 61.988 70.131
Angles α, β, γ [°]	90, 90, 90	90, 112.822, 90
No. total reflections	678641 (67133)	596763 (58811)
No. unique reflections	96450 (9373)	165270 (16155)
Multiplicity	7.0 (7.2)	3.6 (3.6)
Completeness [%]	96.44 (94.40)	95.55 (94.05)
Mean $I/\sigma(I)$	13.04 (1.38)	10.84 (1.29)
Wilson B-factor	12.26	13.26
R-merge	0.0841 (1.565)	0.06747 (1.152)
R-meas	0.0908 (1.687)	0.0792 (1.352)
R-pim	0.03387 (0.626)	0.04115 (0.7029)
$CC_{1/2}$	0.999 (0.547)	0.999 (0.495)
CC*	1 (0.841)	1 (0.814)
Refinement		
Reflections used in refinement	96440 (9372)	165228 (16150)
Reflections used for R-free	4794 (467)	8411 (897)
R-work	0.1334 (0.3327)	0.1282 (0.3376)
R-free	0.1607 (0.3473)	0.1605 (0.3483)
CC(work)	0.974 (0.716)	0.973 (0.574)
CC(free)	0.971 (0.637)	0.967 (0.576)
Number of non-hydrogen atoms	2505	5147
macromolecules	2151	4331
ligands	47	104
solvent	307	712
Protein residues	256	512
RMS bond length [Å]	0.015	0.016
RMS angles [°]	1.89	1.89
Ramachandran favored [%]	98.39	98.4
Ramachandran allowed [%]	1.61	1.6
Ramachandran outliers [%]	0	0
Rotamer outliers [%]	0	1.07
Clashscore	6.08	6.38
Average B-factor	16.34	18.52
macromolecules	14.5	16.04
ligands	17.12	21.1
solvent	29.17	33.25

Table 9: **Data collection and model refinement statistics of measured *h*OMPD$_{314AcK}$ crystals in complex with OMP after 4 min soaking (Crystal10) in space group P2$_1$**

	OMP-soaked-4min
Data collection	
Wavelength [Å]	0.7653
Resolution range [Å]	58.45 - 1.1 (1.139 - 1.1)
Space group	$P2_1$
Unit cell constants [Å]	70.138 62.016 70.263
Angles α, β, γ [°]	90, 112.724, 90
No. total reflections	774801 (76820)
No. unique reflections	222470 (22116)
Multiplicity	3.5 (3.5)
Completeness [%]	98.96 (98.81)
Mean $I/\sigma(I)$	12.44 (1.31)
Wilson B-factor	12.23
R-merge	0.0581 (0.9926)
R-meas	0.06884 (1.18)
R-pim	0.03649 (0.6294)
$CC_{1/2}$	0.999 (0.549)
CC*	1 (0.842)
Refinement	
Reflections used in refinement	222444 (22116)
Reflections used for R-free	10907 (1065)
R-work	0.1296 (0.3148)
R-free	0.1614 (0.3090)
CC(work)	0.976 (0.734)
CC(free)	0.970 (0.717)
Number of non-hydrogen atoms	5033
macromolecules	4260
ligands	102
solvent	671
Protein residues	512
RMS bond length [Å]	0.031
RMS angles [Å]	2.68
Ramachandran favored [%]	98.6
Ramachandran allowed [%]	1.4
Ramachandran outliers [%]	0
Rotamer outliers [%]	0.44
Clashscore	8.63
Average B-factor	16.4
macromolecules	14.27
ligands	19.14
solvent	29.5

Table 10: **Data collection and model refinement statistics of measured *Ca*AAD$_{WT}$ crystals in complex in the resting state and space group P2$_1$2$_1$2$_1$**

	RS
Data collection	
Wavelength [Å]	0.9763
Resolution range [Å]	156.1 - 1.85 (1.916 - 1.85)
Space group	$P2_12_12_1$
Unit cell constants [Å]	103.687 171.656 375.75
Angles α, β, γ [°]	90, 90, 90
No. total reflections	9847851 (991997)
No. unique reflections	563986 (55692)
Multiplicity	17.5 (17.8)
Completeness [%]	99.48 (99.02)
Mean $I/\sigma(I)$	17.07 (1.73)
Wilson B-factor	26.24
R-merge	0.1573 (1.9)
R-meas	0.162 (1.955)
R-pim	0.03846 (0.4588)
$CC_{1/2}$	0.999 (0.581)
CC*	1 (0.857)
Refinement	
Reflections used in refinement	563954 (55691)
Reflections used for R-free	28059 (2748)
R-work	0.1651 (0.2699)
R-free	0.1808 (0.2787)
CC(work)	0.967 (0.738)
CC(free)	0.963 (0.700)
Number of non-hydrogen atoms	52019
macromolecules	48241
ligands	1675
solvent	2103
Protein residues	5856
RMS bond length [Å]	0.012
RMS angles [°]	1.64
Ramachandran favored [%]	96.13
Ramachandran allowed [%]	3.46
Ramachandran outliers [%]	0.41
Rotamer outliers [%]	2.8
Clashscore	6.11
Average B-factor	30.17
macromolecules	29.37
ligands	49
solvent	33.6

Table 11: **Data collection and model refinement statistics of measured *Ca*AAD$_{WT}$ crystals in complex with acetonylphosphonic acid (AcP) and acetonylsulfonic acid (AcS) in space group $P2_12_12_1$**

	AcP	AcS
Data collection		
Wavelength [Å]	0.9763	0.9763
Resolution range [Å]	101.5 - 1.75 (1.813 - 1.75)	101.5 - 1.55 (1.605 - 1.55)
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell constants [Å]	103.942 172.242 376.613	103.979 172.054 377.15
Angles α, β, γ [°]	90, 90, 90	90, 90, 90
No. total reflections	11712607 (1175377)	16869744 (1653279)
No. unique reflections	673381 (66896)	968070 (96116)
Multiplicity	17.4 (17.6)	17.4 (17.2)
Completeness [%]	99.85 (99.91)	99.94 (99.90)
Mean $I/\sigma(I)$	15.42 (1.65)	14.39 (1.22)
Wilson B-factor	20.93	17.31
R-merge	0.175 (1.974)	0.1777 (2.34)
R-meas	0.1802 (2.032)	0.183 (2.412)
R-pim	0.04275 (0.4771)	0.04344 (0.5791)
$CC_{1/2}$	0.999 (0.588)	0.999 (0.471)
CC*	1 (0.861)	1 (0.8)
Refinement		
Reflections used in refinement	673337 (66895)	967958 (96114)
Reflections used for R-free	33547 (3243)	48822 (4837)
R-work	0.1577 (0.2657)	0.1547 (0.2707)
R-free	0.1766 (0.2785)	0.1748 (0.2841)
CC(work)	0.968 (0.797)	0.967 (0.755)
CC(free)	0.963 (0.758)	0.963 (0.735)
Number of non-hydrogen atoms	54185	52507
macromolecules	48157	47170
ligands	2451	2413
solvent	3577	2924
Protein residues	5856	5856
RMS bond length [Å]	0.014	0.008
RMS angles [°]	1.65	1.41
Ramachandran favored [%]	96.06	96.08
Ramachandran allowed [%]	3.52	3.5
Ramachandran outliers [%]	0.42	0.42
Rotamer outliers [%]	3.21	3.05
Clashscore	10.36	6.82
Average B-factor	24.37	19.85
macromolecules	22.92	18.69
ligands	39.44	32.39
solvent	33.54	28.13

Table 12: **Active site residue numbering of OMPD in different organisms.** As reference serves the human (*hs*) residue numbering. Potential amino acid substitutions of *Saccharomyces cerevisiae* (*Sc*), *Methanobacterium thermoautotrophicum* (*Mt*), *Escherichia coli* (*Ec*), and *Plasmodium falciparum* (*Pf*) are indicated in brackets.

	hs	*Sc*	*Mt*	*Ec*	*Pf*
Lys	314	93	72	73	138
Asp	312	91	70	71	136
Lys	281	59	42	44	102
Asp	259	37	20	22	23
Arg	451	235	203	222	294
His	283	61	44(Gly)	46(Gly)	104(Gln)
Pro	417	202	180	189	264
Met	371	153(Leu)	126	130(Leu)	194(Thr)
Ser	372	154	127	131(Thr)	195
Gln	430	215	185	201	269
Tyr	432	217	184(Ala)	192(Arg)	266(Ala)
Asp'	317	96	75	76	141
Ser	373	155(Cys)	128(His)	132	196(Asn)
Thr	321	100	79	80	145
Ser	257	35	18(Ala)	20(Ala)	21(Gly)
Ile'	318	97	76	77	142
Ile	401	183	155(Val)	167(Val)	240(Val)
Leu	366	150	123(Leu)	125(Ile)	89(Phe)
Phe	310	89	66(Ile)	69	134(Ile)
Ile	448	232	200	219(Val)	291(Asn)
Gly	347	126	102	103	169

Table 13: Comparison of recombinant protein expression systems used in this study to purify orotidine 5'-monophosphate decarboxylase

His_6-GST-*h*OMPD

Purification step	Measured entity	Value	Relative
Cell expression, LB	OD_{600}	4	
Cell culture	volume	6.00 L	
Cell pellet	cell mass	30 g	5 g/L
1. affinity chromatography	His_6-GST-*h*OMPD	75 mg	2.5 mg/g
2. affinity chromatography	*h*OMPD	14 mg	0.47 mg/g
S200	*h*OMPD	6 mg	0.2 mg/g

His_6-SUMO-*h*OMPD

Purification step	Measured entity	Value	Relative
Cell expression, ZYM-5052	OD_{600}	12	
Cell culture	volume	1.00 L	
Cell pellet	cell mass	10 g	10 g/L
1. affinity chromatography	His_6-SUMO-*h*OMPD	140 mg	14 mg/g
2. affinity chromatography	*h*OMPD	81 mg	8.1 mg/g
S200	*h*OMPD	53 mg	5.3 mg/g

Table 14: Instrument settings for KinITC measurements of hOMPD$_{WT}$

ITC$_{200}$		Injection	Volume [μL]	Spacing [s]
Cell Volume [mL]	0.207	1	10	1200
Cell Component	OMP	2	10	60
Syringe Volume [μL]	40			
Syringe Component	OMPD			
Total # Injections	2			
Cell temperature [°C]	25			
Reference Power [μCal/s]	7			
Initial Delay [s]	120			
Syringe Concentration [mM]	0.02			
Cell Concentration [mM]	1			
Stirring Speed [rpm]	750			
Feedback Mode/Gain	High			

Table 15: Instrument settings for OMPD/UMP binding analysis: VP-ITC

VP-ITC		Injection	Volume [μL]	Spacing [s]
Cell Volume [mL]	1.43	1	5	300
Cell Component	OMPD	2	10	300
			⋮	
Syringe Volume [μL]	300	30	10	300
Syringe Component	UMP			
Total # Injections	30			
Cell temperature [°C]	25			
Reference Power [μCal/s]	30			
Initial Delay [s]	60			
Syringe Concentration [mM]	4.4			
Cell Concentration [mM]	0.5			
Stirring Speed [rpm]	416			
Feedback Mode/Gain	High			
ITC Equilibration Options	Fast Eq, Auto			

Table 16: Instrument settings for OMPD/UMP binding analysis: ITC_{200}

ITC_{200}		Injection	Volume [μL]	Spacing [s]
Cell Volume [mL]	0.207	1	0.5	120
Cell Component	OMPD	2	2	120
			⋮	
Syringe Volume [μL]	40	20	2	120
Syringe Component	UMP			
Total # Injections	20			
Cell temperature [°C]	25			
Reference Power [μCal/s]	7			
Initial Delay [s]	60			
Syringe Concentration [mM]	4.4			
Cell Concentration [mM]	0.5			
Stirring Speed [rpm]	750			
Feedback Mode/Gain	High			

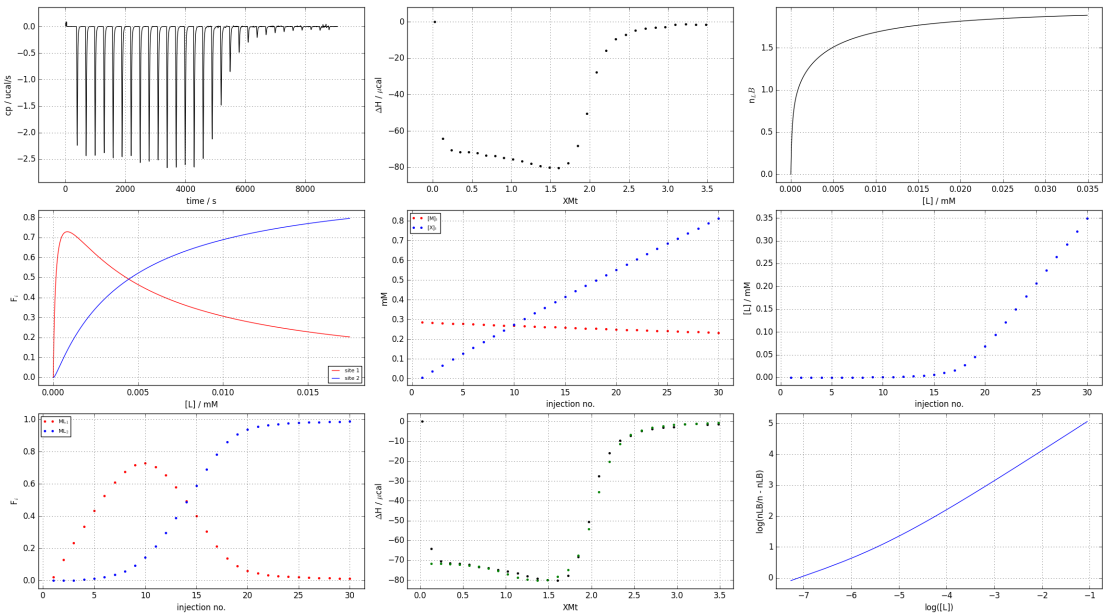


Figure 68: **Experimental data simulation of the UMP-titration of hOMPD$_{WT}$ in a VP-ITC.**

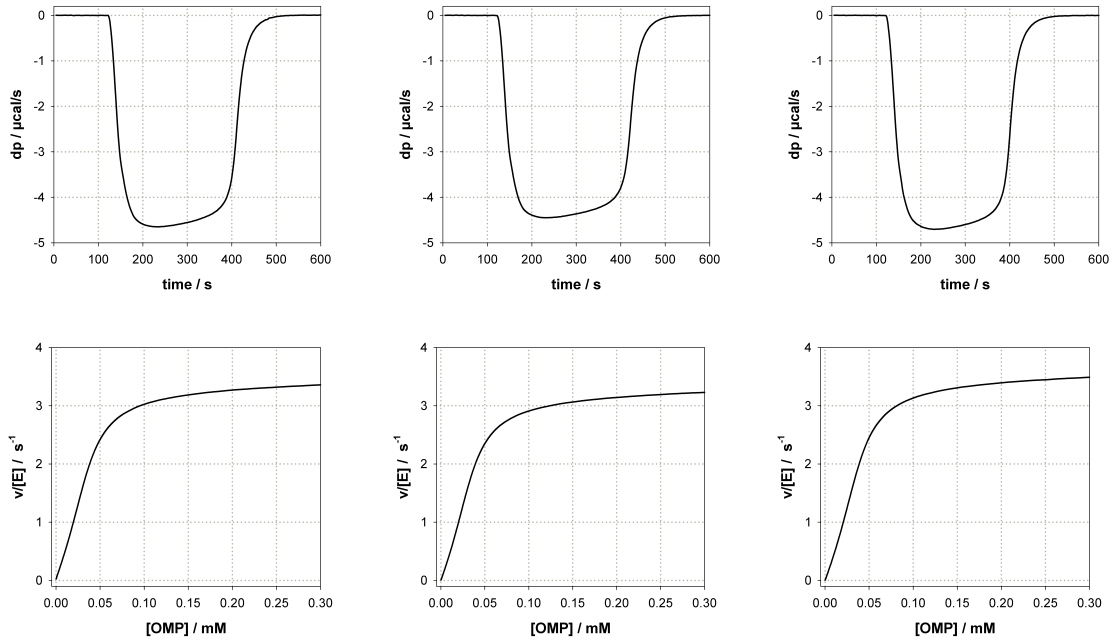


Figure 69: ITC raw data of hOMPD$_{\mathrm{WT}}$

16 Sequences

*Ca*AAD nucleotide sequence

plasmid: pET-28a
5'-introduction: NcoI
3'-introduction: BamHI

ATGGTAAAAGACGAGGTGATCAAACAAATTAGCACACCGCTGACCAGTCCGGCATTTCCGCGTGGTCCGT
ATAAATTCCATAATCGCGAATATTTCAACATCGTGTATCGCACCGATATGGATGCACTGCGTAAAGTTGT
GCCGGAACCGCTGGAAATTGATGAACCTCTGGTTCGTTTTGAAATTATGGCAATGCATGATACCAGCGGT
CTGGGTTGTTATACCGAAAGCGGTCAGGCAATTCCGGTTAGCTTTAATGGTGTTAAAGGCGATTATCTGC
ACATGATGTATCTGGATAATGAACCGGCAATTGCAGTTGGTCGTGAACTGAGCGCATATCCGAAAAAACT
GGGTTATCCGAAACTGTTTGTTGATAGCGATACCCTGGTTGGCACCCTGGATTATGGTAAACTGCGTGTT
GCAACCGCCACAATGGGTTATAAACATAAAGCACTGGATGCCAACGAAGCCAAAGATCAGATTTGTCGTC
CGAACTATATGCTGAAAATCATCCCGAATTATGATGGCAGTCCGCGTATTTGTGAACTGATTAATGCAAA
AATTACCGACGTGACCGTTCATGAAGCATGGACAGGTCCGACCCGTCTGCAGCTGTTCGATCATGCAATG
GCACCGCTGAATGATCTGCCGGTTAAAGAAATTGTTAGCAGCAGCCATATTCTGGCCGATATCATTCTGC
CTCGTGCCGAAGTTATTTATGACTACCTGAAATAA

*Ca*AAD peptide sequence

MVKDEVIKQISTPLTSPAFPRGPYKFHNREYFNIVYRTDMDALRKVVPEPLEIDEPLVRFEIMAMHDTSG
LGCYTESGQAIPVSFNGVKGDYLHMMYLDNEPAIAVGRELSAYPKKLGYPKLFVDSDTLVGTLDYGKLRV
ATATMGYKHKALDANEAKDQICRPNYMLKIIPNYDGSPRICELINAKITDVTVHEAWTGPTRLQLFDHAM
APLNDLPVKEIVSSSHILADIILPRAEVIYDYLK*

*h*UMPS nucleotide sequence

ATGGCGGTCGCTCGTGCAGCTTTGGGGCCATTGGTGACGGGTCTGTACGACGTGCAGGCTTTCAAGTTTG
GGGACTTCGTGCTGAAGAGCGGGCTTTCCTCCCCCATCTACATCGATCTGCGGGGCATCGTGTCTCGACC
GCGTCTTCTGAGTCAGGTTGCAGATATTTTATTCCAAACTGCCCAAAATGCAGGCATCAGTTTTGACACC
GTGTGTGGAGTGCCTTATACAGCTTTGCCATTGGCTACAGTTATCTGTTCAACCAATCAAATTCCAATGC
TTATTAGAAGGAAAGAAACAAAGGATTATGGAACTAAGCGTCTTGTAGAAGGAACTATTAATCCAGGAGA
AACCTGTTTAATCATTGAAGATGTTGTCACCAGTGGATCTAGTGTTTTGGAAACTGTTGAGGTTCTTCAG
AAGGAGGGCTTGAAGGTCACTGATGCCATAGTGCTGTTGGACAGAGAGCAGGGAGGCAAGGACAAGTTGC
AGGCGCACGGGATCCGCCTCCACTCAGTGTGTACATTGTCCAAAATGCTGGAGATTCTCGAGCAGCAGAA
AAAAGTTGATGCTGAGACAGTTGGGAGAGTGAAGAGGTTTATTCAGGAGAATGTCTTTGTGGCAGCGAAT
CATAATGGTTCTCCCCTTTCTATAAAGGAAGCACCCAAAGAACTCAGCTTCGGTGCACGTGCAGAGCTG
CCCAGGATCCACCCAGTTGCATCGAAGCTTCTCAGGCTTATGCAAAAGAAGGAGACCAATCTGTGTCTAT
CTGCTGATGTTTCACTGGCCAGAGAGCTGTTGCAGCTAGCAGATGCTTTAGGACCTAGTATCTGCATGCT
GAAGACTCATGTAGATATTTTGAATGATTTTACTCTGGATGTGATGAAGGAGTTGATAACTCTGGCAAAA
TGCCATGAGTTCTTGATATTTGAAGACCGGAAGTTTGCAGATATAGGAAACACAGTGAAAAAGCAGTATG
AAGGAGGTATCTTTAAAATAGCTTCCTGGGCAGATCTAGTAAATGCTCACGTGGTGCCAGGCTCAGGAGT
TGTGAAAGGCCTGCAAGAAGTGGGCCTGCCTTTGCATCGGGGGTGCCTCCTTATTGCGGAAATGAGCTCC
ACCGGCTCCCTGGCCACTGGGGACTACACTAGAGCAGCGGTTAGAATGGCTGAGGAGCACTCTGAATTTG
TTGTTGGTTTTATTTCTGGCTCCCGAGTAAGCATGAAACCAGAATTTCTTCACTTGACTCCAGGAGTTCA
GTTGGAAGCAGGAGGAGATAATCTTGGCCAACAGTACAATAGCCCACAAGAAGTTATTGGCAAACGAGGT
TCCGATATCATCATTGTAGGTCGTGGCATAATCTCAGCAGCTGATCGTCTGGAAGCAGCAGAGATGTACA
GAAAAGCTGCTTGGGAAGCGTATTTGAGTAGACTTGGTGTTTGA

*h*UMPS peptide sequence

MAVARAALGPLVTGLYDVQAFKFGDFVLKSGLSSPIYIDLRGIVSRPRLLSQVADILFQTAQNAGISFDT
VCGVPYTALPLATVICSTNQIPMLIRRKETKDYGTKRLVEGTINPGETCLIIEDVVTSGSSVLETVEVLQ
KEGLKVTDAIVLLDREQGGKDKLQAHGIRLHSVCTLSKMLEILEQQKKVDAETVGRVKRFIQENVFVAAN
HNGSPLSIKEAPKELSFGARAELPRIHPVASKLLRLMQKKETNLCLSADVSLARELLQLADALGPSICML
KTHVDILNDFTLDVMKELITLAKCHEFLIFEDRKFADIGNTVKKQYEGGIFKIASWADLVNAHVVPGSGV
VKGLQEVGLPLHRGCLLIAEMSSTGSLATGDYTRAAVRMAEEHSEFVVGFISGSRVSMKPEFLHLTPGVQ
LEAGGDNLGQQYNSPQEVIGKRGSDIIIVGRGIISAADRLEAAEMYRKAAWEAYLSRLGV

His_6-GST-*h*OMPD nucleotide sequence

plasmid: pETM-30
5'-introduction: NcoI
3'-introduction: EcoRI

ATGAAACATCACCATCACCATCACAACACTAGTAGCAATTCCATGTCCCCTATACTAGGTTATTGGAAAA
TTAAGGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTA
TGAGCGCGATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCT
TATTATATTGATGGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATATAGCTGACAAGCACA
ACATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAG
ATACGGTGTTTCGAGAATTGCATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTA
CCTGAAATGCTGAAAATGTTCGAAGATCGTTTATGTCATAAAACATATTTAAATGGTGATCATGTAACCC
ATCCTGACTTCATGTTGTATGACGCTCTTGATGTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTT
CCCAAAATTAGTTTGTTTTAAAAAACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAAATCCAGC
AAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAACTAGTG
GATCTGGTGGTGGTGGCGGATGGATGAGCGAGAATCTTTATTTTCAGGGCGCCATGGAACTCAGCTTCGG
TGCACGTGCAGAGCTGCCCAGGATCCACCCAGTTGCATCGAAGCTTCTCAGGCTTATGCAAAAGAAGGAG
ACCAATCTGTGTCTATCTGCTGATGTTTCACTGGCCAGAGAGCTGTTGCAGCTAGCAGATGCTTTAGGAC
CTAGTATCTGCATGCTGAAGACTCATGTAGATATTTTGAATGATTTTACTCTGGATGTGATGAAGGAGTT
GATAACTCTGGCAAAATGCCATGAGTTCTTGATATTTGAAGACCGGAAGTTTGCAGATATAGGAAACACA
GTGAAAAAGCAGTATGAAGGAGGTATCTTTAAAATAGCTTCCTGGGCAGATCTAGTAAATGCTCACGTGG
TGCCAGGCTCAGGAGTTGTGAAAGGCCTGCAAGAAGTGGGCCTGCCTTTGCATCGGGGGTGCCTCCTTAT
TGCGGAAATGAGCTCCACCGGCTCCCTGGCCACTGGGGACTACACTAGAGCAGCGGTTAGAATGGCTGAG
GAGCACTCTGAATTTGTTGTTGGTTTTATTTCTGGCTCCCGAGTAAGCATGAAACCAGAATTTCTTCACT
TGACTCCAGGAGTTCAGTTGGAAGCAGGAGGAGATAATCTTGGCCAACAGTACAATAGCCCACAAGAAGT
TATTGGCAAACGAGGTTCCGATATCATCATTGTAGGTCGTGGCATAATCTCAGCAGCTGATCGTCTGGAA
GCAGCAGAGATGTACAGAAAAGCTGCTTGGGAAGCGTATTTGAGTAGACTTGGTGTTTGA

His_6-GST-*h*OMPD peptide sequence

MKHHHHHHNTSSNSMSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLP
YYIDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLSKL
PEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSS
KYIAWPLQGWQATFGGGDHPPTSGSGGGGGWMSENLYFQGAMELSFGARAELPRIHPVASKLLRLMQKKE
TNLCLSADVSLARELLQLADALGPSICMLKTHVDILNDFTLDVMKELITLAKCHEFLIFEDRKFADIGNT
VKKQYEGGIFKIASWADLVNAHVVPGSGVVKGLQEVGLPLHRGCLLIAEMSSTGSLATGDYTRAAVRMAE
EHSEFVVGFISGSRVSMKPEFLHLTPGVQLEAGGDNLGQQYNSPQEVIGKRGSDIIIVGRGIISAADRLE
AAEMYRKAAWEAYLSRLGV*

His_6-SUMO-*h*OMPD$_{224-480}$ nucleotide sequence

ATGGGCAGCAGCCATCATCATCATCATCACGGCAGCGGCCTGGTGCCGCGCGGCAGCGCTAGCATGTCGG
ACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTT
AAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATG
GAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACTCCTTAAGATTCTTGTACGACGGTATTAGAATTC
AAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGATAACGATATTATTGAGGCTCACAGAGAACAGAT
TGGTGGTGGCGCCATGGAACTCAGCTTCGGTGCACGTGCAGAGCTGCCCAGGATCCACCCAGTTGCATCG
AAGCTTCTCAGGCTTATGCAAAAGAAGGAGACCAATCTGTGTCTATCTGCTGATGTTTCACTGGCCAGAG
AGCTGTTGCAGCTAGCAGATGCTTTAGGACCTAGTATCTGCATGCTGAAGACTCATGTAGATATTTTGAA
TGATTTTACTCTGGATGTGATGAAGGAGTTGATAACTCTGGCAAAATGCCATGAGTTCTTGATATTTGAA
GACCGGAAGTTTGCAGATATAGGAAACACAGTGAAAAAGCAGTATGAAGGAGGTATCTTTAAAATAGCTT
CCTGGGCAGATCTAGTAAATGCTCACGTGGTGCCAGGCTCAGGAGTTGTGAAAGGCCTGCAAGAAGTGGG
CCTGCCTTTGCATCGGGGGTGCCTCCTTATTGCGGAAATGAGCTCCACCGGCTCCCTGGCCACTGGGGAC
TACACTAGAGCAGCGGTTAGAATGGCTGAGGAGCACTCTGAATTTGTTGTTGGTTTTATTTCTGGCTCCC
GAGTAAGCATGAAACCAGAATTTCTTCACTTGACTCCAGGAGTTCAGTTGGAAGCAGGAGGAGATAATCT
TGGCCAACAGTACAATAGCCCACAAGAAGTTATTGGCAAACGAGGTTCCGATATCATCATTGTAGGTCGT
GGCATAATCTCAGCAGCTGATCGTCTGGAAGCAGCAGAGATGTACAGAAAAGCTGCTTGGGAAGCGTATT
TGAGTAGACTTGGTGTTTGA

His_6-SUMO-*h*OMPD$_{224-480}$ peptide sequence

MGSSHHHHHHGSGLVPRGSASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLM
EAFAKRQGKEMDSLRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGGGAMELSFGARAELPRIHPVAS
KLLRLMQKKETNLCLSADVSLARELLQLADALGPSICMLKTHVDILNDFTLDVMKELITLAKCHEFLIFE
DRKFADIGNTVKKQYEGGIFKIASWADLVNAHVVPGSGVVKGLQEVGLPLHRGCLLIAEMSSTGSLATGD
YTRAAVRMAEEHSEFVVGFISGSRVSMKPEFLHLTPGVQLEAGGDNLGQQYNSPQEVIGKRGSDIIIVGR
GIISAADRLEAAEMYRKAAWEAYLSRLGV*

Sören Rindfleisch

Curriculum Vitae

Tiergartenstraße 4
30926 Seelze
0173 5807027
soeren.rindfleisch@gmail.com

Professional Experience

03.01 - 07.01.18	**12th New Enzymology Kinetic Workshop 2018**, Prof. Dr. Kenneth Johnson
01.02. - 31.10.13	**iGem Team Göttingen 2013**, *The beast and its Achilles heel: A novel target to fight multi-resistant pathogenic bacteria*
01.02. - 31.03.13	**Lab assistant**, *J.F. Blumenbach Institute of Zoology and Anthropology, Göttingen*
01.02. - 31.07.12	**Study abroad**, *Università degli studi di Roma "Tor Vergata", Rome*
1.07. - 31.12.11	**Lab assistant**, *Department of Plant Cell Biology, Göttingen*

Education

01.02.2015 -	**PhD studies: "Biomolecules: Structure-Function-Dynamics"**, *Reaction mechanism of OMPD and AAD at atomic resolution*, Göttingen Graduate School for Neurosciences, Biophysics, and Molecular Biosciences (GGNB), Dept. Molecular Enzymology, Prof. Dr. Kai Tittmann
2014	**Master thesis**, *Reaction mechanism and inhibition of human glutaminyl cyclase*, Ref: Prof. Dr. Kai Tittmann, Corref: Prof. Dr. Jörg Stülke 1.0
2012 - 2014	**Master studies: "Genetics and Molecular Biology", Georg-August-Universität Göttingen**
2011	**Bachelor thesis**, *Infection analysis of A. thaliana with Verticillium longisporum and Verticillium dahliae and altered infection sensitivity due to mutations in pre- and post invasive defense mechanisms*, Ref: Prof. Dr. Volker Lipka, Corref: Dr. Thomas Teichmann 1.3
2008 - 2012	**Bachelor studies: "Biology", Georg-August-Universität Göttingen**
2001 - 2008	**High school Helene-Lange-Schule Hannover** Abitur: 1.9

Additional Skills

German	Mother tongue
English	Unicert III, Scientific English Level C1
Italian	Level B1

Publications

[1] M. Reusche, K. Thole, D. Janz, J. Truskina, S. Rindfleisch, C. Drubert, A. Polle, V. Lipka, and T. Teichmann. "Verticillium infection triggers VASCULAR-RELATED NAC DOMAIN7-dependent *de novo* xylem formation and enhances drought tolerance in Arabidopsis". *Plant Cell* 24.9 (Sept. 2012), pp. 3823–3837.