

Novel Approaches for the Creation of Artificial Metalloenzymes

Inauguraldissertation
zur Erlangung der Würde eines Doktors der Philosophie

vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Elisa Sofia Pereira Nogueira

*aus Vila Cova, Barcelos
Portugal*

Basel, 2013

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel
edoc.unibas.ch



Dieses Werk ist unter dem Vertrag "Creative Commons Namensnennung –Keine kommerzielle Nutzung –
Keine Bearbeitung – 2.5 Schweiz" lizenziert. Die vollständige Lizenz kann unter:
creativecommons.org/licenses/by-nc-nd/2.5/ch
eingesehen werden.

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von:

Prof. Dr Thomas R. Ward
Dr Gideon J. Grogan

Basel, den 26. März 2013

Prof. Dr Jörg Schibler
Dekan

In loving memory of my grandfather, Leonardo,

℘

To my family.

“A consciência da inconsciência da vida é o mais antigo imposto à inteligência. Há inteligências inconscientes – brilhos do espírito, correntes do entendimento, mistérios e filosofias – que têm o mesmo automatismo que os reflexos corpóreos, que a gestão que o fígado e os rins fazem de suas secreções.”

“The consciousness of life’s unconsciousness is the oldest tax levied on the intelligence. There are unconscious forms of intelligence – flashes of wit, waves of understanding, mysteries and philosophies – that are like bodily reflexes, that operate as automatically as the liver or kidneys handle their secretions.”

– Fernando Pessoa, *Livro do Desassossego*



Attribution-NonCommercial-NoDerivatives 3.0 Switzerland
(CC BY-NC-ND 3.0 CH)

You are free: to **Share** — to copy, distribute and transmit the work

Under the following conditions:



Attribution — You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).



Noncommercial — You may not use this work for commercial purposes.



No Derivative Works — You may not alter, transform, or build upon this work.

With the understanding that:

- **Waiver** — Any of the above conditions can be **waived** if you get permission from the copyright holder.
- **Public Domain** — Where the work or any of its elements is in the **public domain** under applicable law, that status is in no way affected by the license.
- **Other Rights** — In no way are any of the following rights affected by the license:
 - Your fair dealing or **fair use** rights, or other applicable copyright exceptions and limitations;
 - The author's **moral** rights;
 - Rights other persons may have either in the work itself or in how the work is used, such as **publicity** or privacy rights.
- **Notice** — For any reuse or distribution, you must make clear to others the license terms of this work. The best way to do this is with a link to this web page.

Impact of the work

The work presented herein was initiated and guided by Prof. Dr Thomas R. Ward, at the Department of Chemistry – University of Basel, from September 2008 to November 2012.

Excerpts from this work have been or will be published:

Nogueira, E. S., Schleier, T., Dürrenberger, M., Ballmer-Hofer, K., Ward, T. R., and Jaussi, R. “High level secretion of recombinant full-length streptavidin in *Pichia pastoris* and its application to enantioselective catalysis”, *Prot. Expr. Purif.*, **2014**, *93*, 54.

Monnard, F. W., Nogueira, E. S., Heinisch, T., Schirmer, T., and Ward, T. R. “Artificial imine reductase based on human carbonic anhydrase II as host protein”, *Chem. Sci.*, **2013**, *4*, 3269.

Zimbron, J. M., Heinisch, T., Schmid, M., Hamels, D., Nogueira, E. S., Schirmer, T., and Ward, T. R. “A dual anchoring strategy for the localization and activation of artificial metalloenzymes based on the biotin-streptavidin technology”, *J. Am. Chem. Soc.*, **2013**, *135*, 5384.

Bereau, T., Kramer, C., Monnard, F. W., Nogueira, E. S., Ward, T. R. and Meuwly, M. “Scoring multipole electrostatics in atomistic protein-ligand binding simulations”, *J. Phys. Chem. B.*, **2013**, *117*, 5460.

Schmid, M., Nogueira, E. S., Monnard, F. W., Ward, T. R. and Meuwly, M. “Arylsulfonamides as inhibitors for carbonic anhydrase: Prediction & validation”, *Chem. Sci.*, **2012**, *3*, 690.
Advance article + Front cover

Monnard, F. M., Heinisch, T., Nogueira, E. S., Schirmer, T., and Ward, T. R. “Human carbonic anhydrase II as host for piano-stool complexes bearing a sulfonamide anchor”, *Chem. Commun.*, **2011**, *47*, 8238.

Dürrenberger, M., Heinisch, T., Wilson, Y. M., Rossel, T., Nogueira, E. S., Knörr, L., Mutschler, A., Kersten, K., Zimbron, J. M., Pierron, J., Schirmer, T., and Ward, T. R. “Artificial transfer hydrogenases for the enantioselective reduction of cyclic imines”, *Angew. Chem. Int. Ed. Engl.*, **2011**, *50*, 3026.

Keywords: artificial metalloenzymes, high-throughput screening, chemo-genetic optimisation, streptavidin, human carbonic anhydrase II, bacterial and yeast expression systems.

Table of contents

Impact of the work	i
Table of contents	iii
Abbreviations	vii
Synopsis	ix
1 Artificial metalloenzymes	1
1.1 Concept and design of artificial metalloenzymes	1
1.1.1 Transition metal catalysts	3
1.1.2 Types of biomolecular scaffolds	4
1.1.3 Anchoring strategies	6
1.2 Catalytic scope	7
1.2.1 Asymmetric transfer hydrogenation of imines	7
1.2.2 Asymmetric reductive amination of α -keto acids	9
1.3 Biological scope	10
1.3.1 Streptavidin	10
1.3.2 Human carbonic anhydrase II	12
1.4 Tools for optimisation	16
1.4.1 Directed evolution	17
1.4.2 Designed evolution	18
1.4.3 Screening techniques	18
1.4.4 Expression systems	19
1.5 Scope of this thesis	27
1.6 References	28
2 New strategies for the purification of streptavidin	43
2.1 Introductory remarks	44
2.1.1 A long process: from the gene to the protein	44
2.1.2 Asymmetric transfer hydrogenation of imines	45
2.1.3 Asymmetric reductive amination of α -keto acids	47
2.1.4 Research project	49
2.2 Results & discussion	49

2.2.1	Screening on crude protein extracts	49
2.2.2	Screening on proteins purified by precipitation	50
2.2.3	Screening on proteins purified on a small-scale	55
2.2.4	Screening on crude protein extracts free of reduced glutathione	58
2.3	Conclusion & outlook	62
2.4	References	64
3	New platform for the expression of streptavidin	67
3.1	Introductory remarks	67
3.1.1	From bacterium to yeast	68
3.1.2	Research project	69
3.2	Results & discussion	70
3.2.1	Strain and genetic construct	70
3.2.2	Expression and detection of streptavidin in <i>Pichia pastoris</i>	73
3.2.3	Expression under different conditions	75
3.2.4	High-cell density fed-batch fermentation	77
3.2.5	Purification of streptavidin expressed in <i>Pichia pastoris</i>	78
3.2.6	Biochemical properties	79
3.3	Conclusion & outlook	83
3.4	References	85
4	New scaffold for the creation of an artificial metalloenzyme	89
4.1	Introductory remarks	90
4.1.1	Human carbonic anhydrase II as potential new biomolecular scaffold	90
4.1.2	Sulphonamides as inhibitors	91
4.1.3	Structure of human carbonic anhydrase II	91
4.1.4	Research project	94
4.2	Results & discussion	94
4.2.1	Production of human carbonic anhydrase II	94
4.2.2	Studies on human carbonic anhydrase II as potential biomolecular scaffold	103
4.2.3	Arylsulphonamides as inhibitors	109
4.2.4	Pseudo-contact shifts in solution-state NMR	113
4.3	Conclusion & outlook	117
4.4	References	119
5	Summary & outlook	123

6 Materials & methods	125
6.1 General experimental section	125
6.1.1 Standard methods & reagents	125
6.1.2 Equipment	125
6.2 Experimental section of Chapter 2	127
6.2.1 General procedure for the production of streptavidin	127
6.2.2 Purification procedure	132
6.2.3 Characterisation of recombinant streptavidin	134
6.2.4 Quantification methods	135
6.2.5 General procedure for the production of streptavidin in cell-free protein extracts	139
6.2.6 General procedure for enantioselective catalysis	140
6.3 Experimental section of Chapter 3	144
6.3.1 General procedure for the production of streptavidin	144
6.3.2 Purification procedures	151
6.3.3 Characterisation of recombinant streptavidin	152
6.4 Experimental section of Chapter 4	156
6.4.1 General procedure for the production of human carbonic anhydrase II	156
6.4.2 Purification procedures	165
6.4.3 Characterisation of recombinant human carbonic anhydrase II	167
6.5 References	170
Acknowledgements	173
Curriculum vitae	177
Appendices	183

Abbreviations

C	included in
α -MF	α -mating factor
Å	Angström
A	absorbance
amp	ampicillin
AOX	alcohol oxidase
APS	ammonium persulphate
ATH	asymmetric transfer hydrogenation
AU	arbitrary units
B4F	biotin-4-fluorescein
Biot	biotin
BSA	bovine serum albumin
CH ₂ Cl ₂	dichloromethane
cm	chloramphenicol
conv.	conversion
Cp*	pentamethylcyclopentadienyl
CV	column volume
D ₂ O	deuterated water
Da	Dalton
dcw	dry cell weight
DME	dimethoxyethane
DMF	<i>N,N</i> -dimethylformamide
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
DNSA	dansylamide
dNTP	any deoxyribonucleic triphosphate
dO ₂	dissolved oxygen
DOC	deoxycholate
DTT	dithiothreitol
ϵ	molar extinction coefficient
EC	enzyme classification
<i>ee</i>	enantiomeric excess
EDTA	ethylenediaminetetraacetic acid
<i>e.g.</i>	<i>exempli gratia</i> , for example
eq	equivalent
ESI-TOF	electron spray ionisation – time of flight
EtOH	ethanol
FPLC	fast performance liquid chromatography

Gdn-HCl	guanidinium chloride
GSH	reduced glutathione
GSSH	oxidised glutathione
h	hour(s)
hCAII	human carbonic anhydrase II
HPLC	high performance liquid chromatography
HTS	high-throughput screening
<i>i.e.</i>	<i>id est</i> , that is
<i>i</i> -PrOH	isopropanol
IPTG	isopropyl- β -L-galactoside
kb	kilo base
kDa	kilo Dalton
LB	lysogeny broth (LB-Miller)
MCS	multiple cloning site
MeOH	methanol
min	minute(s)
MOPS	3-(<i>N</i> -morpholino)propanesulphonic acid
MS	mass spectrometry
MUT	methanol utilisation pathway
MW	molecular weight
NMR	nuclear magnetic resonance
OD	optical density
PCR	polymerase chain reaction
PDB	Protein Data Bank
PNPA	<i>p</i> -nitrophenylacetate
ppm	parts per million
quant.	quantitative
rac.	racemic
RNA	ribonucleic acid
rpm	rotation per minute
RT	room temperature
RV	resin volume
Sav	streptavidin
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SSP	small-scale purification
t_R	retention time (HPLC)
TCA	trichloroacetic acid
TEMED	tetramethylethylenediamine
TON	turnover number
v/v	volume per volume
<i>vs</i>	<i>versus</i> , against
w/v	weight per volume
wcw	wet cell weight
wt	wild-type

Synopsis

Chapter 1. Artificial metalloenzymes In this Chapter, the concept and approaches followed in the design and optimisation of artificial metalloenzymes are presented. An overview on the art of creating artificial metalloenzymes is provided, with a particular focus on the biomolecular scaffolds and tools to optimise artificial metalloenzymes regarding the production of the host protein.

Chapter 2. New strategies for the purification of streptavidin The time-consuming production process of the host protein, streptavidin, is a bottleneck to high-throughput screening, and limits the application of directed evolution to artificial metalloenzymes. In this Chapter, the critical issue of sensitivity of the catalysts to cell-based poisons is assessed, and novel strategies to purify the protein in sufficient quantity and quality are presented.

Chapter 3. New platform for the expression of streptavidin Based on the advantages of *Pichia pastoris* expression system reported in literature, the primary objective of this study was to improve the expression levels of streptavidin through *Pichia pastoris* system relative to *Escherichia coli*, and subsequent purification steps. The high level of functional protein secreted would be suitable to perform high-throughput screening directly in the supernatant with no further steps of purification.

Chapter 4. New scaffold for the creation of an artificial metalloenzyme In this Chapter, the know-how on biotin-streptavidin technology was transferred to a new biomolecular scaffold, human carbonic anhydrase II. Applicable force field parameters amenable to molecular dynamics simulations of hCAII · inhibitor interactions were experimentally validated. X-ray crystal structure analyses of the artificial metalloenzyme provided insights on the coordination structures and non-covalent interactions between the metal moiety and protein scaffold. Isotopically labelled proteins were expressed for NMR studies.

Chapter 5. Summary & outlook In this last Chapter, achievements are summarised, and overall conclusions are drawn. Furthermore, lines of research are suggested.

Artificial metalloenzymes

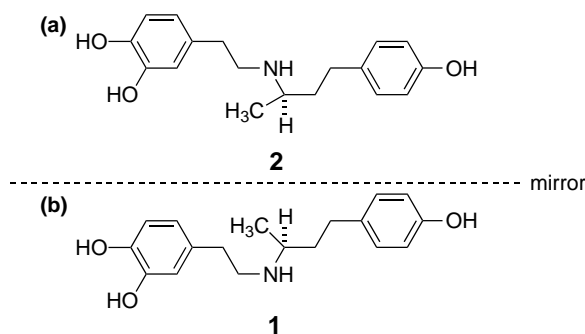
We must not forget that when radium was discovered, no one knew that it would prove useful in hospitals. The work was one of pure science. And this is a proof that scientific work must not be considered from the point-of-view of the direct usefulness of it. It must be done for itself, for the beauty of science, and then there is always the chance that a scientific discovery may become like the radium: a benefit for humanity.

Marie Skłodowska-Curie

1.1 Concept and design of artificial metalloenzymes

Production of enantiomerically pure (enantiopure) compounds is a major issue in organic chemistry in general, but also in pharmaceutical, flavour, and agrochemical industries, in particular. Although, in an achiral environment, pure enantiomers have identical physical and chemical properties (*e.g.* density, boiling point, and chemical reactivity) to their corresponding mirror images, they are different chemical compounds.^[1] Just like a left-handed person cannot use a right-handed baseball glove, one enantiomer may not fit in the binding pocket of an enzyme, where the other will. Thus, within biological systems, one enantiomer may often exhibit different pharmacological properties than the other enantiomer,^[2,3] since the molecules with which they interact are also optically active.^[4] Dobutamine, for example, belongs to this group. Dobutamine is a chiral drug with enantiomers being agonists at different receptors (Scheme 1.1). Whilst the (-)-isomer **1** has α 1-blocking agonist activity, the (+)-isomer **2** is a β 1-adrenergic agonist.^[5,6]

Scheme 1.1. (\pm)-Dobutamine enantiomers. (a) The (+)-isomer **2**, an β 1-adrenergic agonist. (b) The (-)-isomer **1**, an α 1-blocking agonist.



There are many ways to afford enantiopure compounds, and enantioselective catalysis – homogeneous and enzymatic – proved to be the most efficient way. In biocatalysis, the activity and (enantio)selectivity are due to the so-called “second coordination sphere”, which is the term used to define subtle combinations of secondary interactions (hydrogen bonding, hydrophobic interactions) provided by the protein.^[6] Whereas in metal-catalysed enantioselective catalysis, the activity and selectivity are almost exclusively dictated by the “first coordination sphere” provided by the chiral ligand.^[7–9]

Though their mode of action may be different, chemical catalysts and biocatalysts are on many aspects complementary: range of reactions and substrates, operating conditions or enantioselectivity. Furthermore, advantages (respectively disadvantages) of homogeneous catalysis are disadvantages (respectively advantages) in enzymatic catalysis.^[10] For example, biocatalysts have a limited substrate scope due to the lock-and-key specificity,^[11,12] and usually lead to a single enantiomer. Notwithstanding, they tend to be extremely selective and perform a wide range of reactions in aqueous conditions, with high turnover numbers. In a single step, biocatalysts can carry out transformations that might take two or more steps in a chemical process, and they can even effect reactions that cannot be done by chemical means at all. In comparison, traditional transition metal catalysts have a broad substrate scope, and both enantiomers can be accessed by homogeneous catalysis, as the optical antipodes of the chiral ligand are readily accessible. However, turnover numbers (TONs) are usually lower for metal catalysts than for enzymes, and chemical catalysts are often used in organic solvents.^[13] Considering the overall advantages and disadvantages of both systems, merging certain beneficial aspects of bio- and homogeneous catalysis seemed to be of interest, and one of the most promising approach.

The general concept of “artificial metalloenzymes”, in which a catalytically active transition metal complex is embedded into a host biomolecular scaffold (typically a protein^[14,15] or DNA^[16,17]), was originally introduced several decades ago. In the infancy period of the field of asymmetric catalysis, Wilson & Whitesides devised the generation of an artificial enzyme that relied on the incorporation of a metal-containing fragment into a host protein, thus creating a new supramolecular catalytic system.^[18,19] After years of neglect, the development of hybrid catalysts that combines the biological concepts for selective molecular recognition with those of transition metal catalysis found a new renaissance, when significant efforts were made to bridge the gap between homogeneous catalysis and biocatalysis.^[6,20] These efforts resulted in a remarkably versatile strategy that explores and combines the relative complementarities of both worlds: the efficiency, robustness and wide scope of reactions of synthetic catalysts, with the high selectivity under mild conditions of enzymes.^[10,21] Therefore, the introduction of a catalytic metal moiety, which ensures activity, in a chiral shell provided by the host biomolecule, is expected to create transition metal complexes with a well-defined second coordination sphere that display enzyme-like activities and selectivities. An attractive feature of

these hybrid catalysts is that their performance can be improved by “chemo-genetic” optimisation,^[22] *i.e.* by independently modify the homogeneous catalyst by chemical optimisation, and the biological scaffold by directed evolution or rational design (Section 1.4.1).^[23,24]

The incentives for the creation of active and selective artificial metalloenzymes are both practical and theoretical, as it holds key lessons for catalytic reactions mechanism and catalyst design in general.

In summary, in protein-based asymmetric catalysis, the chirality of the protein is used to induce enantioselectivity in a metal-catalysed reaction, by non-covalent binding of an achiral ligand to the protein. Hence, the catalytically active organometallic centre is brought into close proximity to the biomolecule binding pocket, which provides the chiral second coordination sphere and directs the catalysed reaction toward one of the enantiomers of the product, thus resulting in an enantiomeric excess (Figure 1.1). This concept has been demonstrated successfully in a plethora of classical asymmetric reactions, such as hydrolysis,^[25,26] hydrogenation,^[19,23,27–38] transfer hydrogenation,^[15,39,40] allylic alkylation,^[41] sulfoxidation,^[42] epoxidation,^[43–45] dihydroxylation,^[46] Diels-Alder reaction,^[23,47–50] transamination,^[51] and C-H activation.^[52]

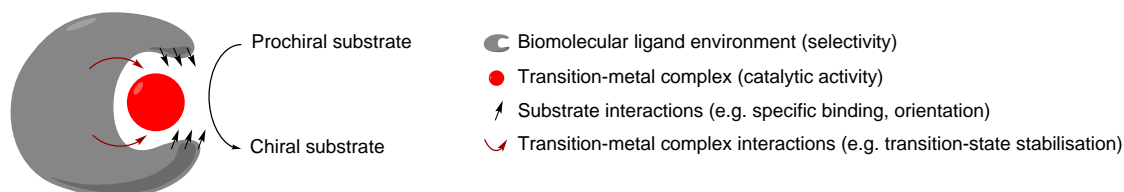


Figure 1.1. Concept of artificial metalloenzymes. Authorised reprint from John Wiley and Sons.^[20]

The design of enantioselective artificial metalloenzymes is based on three fundamental parameters:

- i. The transition metal catalyst;
- ii. The biomolecular scaffold;
- iii. The anchoring strategy.

The following sections describe these parameters in more detail.

1.1.1 Transition metal catalysts

The principal challenge in this approach is, without doubt, the development of selective catalysts that convert relevant substrates into desired products, with high (enantio)selectivity. The

choice of the transition metal catalyst is mainly driven by the reaction under consideration. The metal, its ligands (which perturb the electronic and steric properties of the metal centre), coordination number and geometries are all relevant to catalysis.^[53] For efficient asymmetric catalysis, the transition metal catalyst should fulfil the following requirements: (i) be orthogonal to the biomolecular scaffold, thus inert to the chemical functionalities presented by the biomolecule,^[6,20] (ii) be compatible with the biomolecular scaffold, (iii) be tolerant to water, as the use of artificial metalloenzymes implies working in aqueous solutions,^[54] and (iv) ideally, it should be inactive in its free form and active when embedded into the biomolecular host.^[6] The reaction conditions affect the reactivity of the transition metal catalyst, and therefore these must be cautiously selected and optimised.

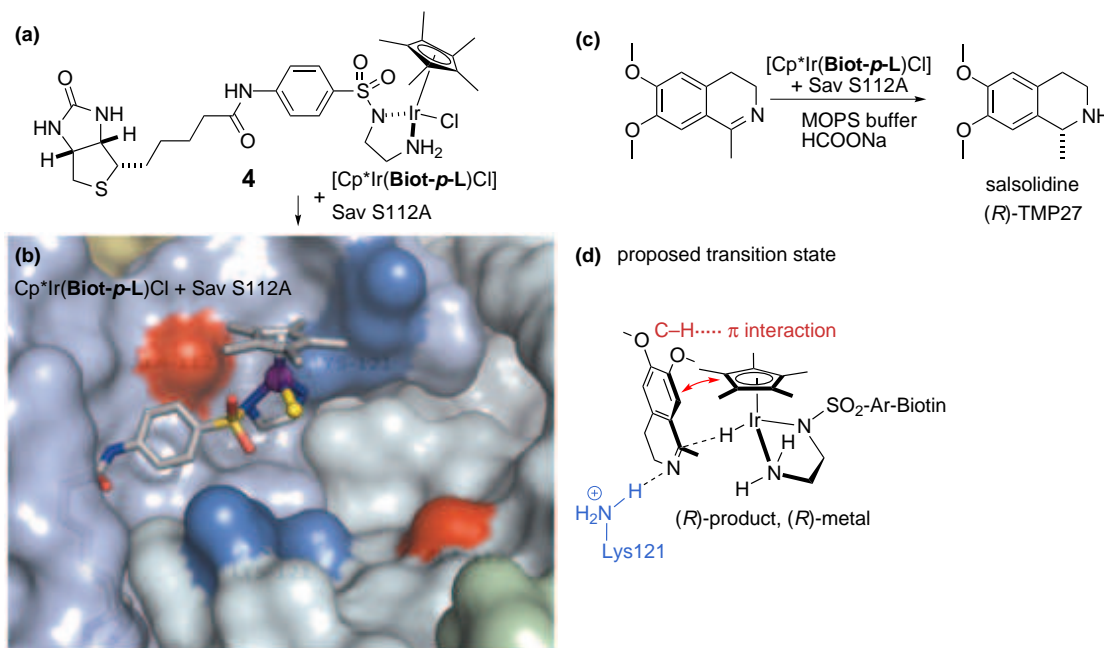
1.1.2 Types of biomolecular scaffolds

In many cases, the biomolecular scaffold not only induces selectivity in the reaction, but also affects the rate of the reaction; the kinetic effect can be either negative or positive. In some cases, significant rate accelerations have been observed^[46,52,55,56] whereas, in other cases, when the catalyst resides in the protein shell,^[16,27,57] the reaction is slower. Often the event is substrate-dependent, and hampered by substrate-binding preferences of the enzyme active site or by structural compatibility with certain activated complexes only.^[58] Therefore, the choice of biomolecular scaffolds is limited due to (i) the chemical properties of the scaffold, such as pH and temperature stability, overall charge, and tolerance to organic solvents,^[6] (ii) the size of the binding pocket, which has to be sufficiently large to accommodate both the transition metal catalyst and the substrates,^[54] (iii) the catalytic chemistry envisaged. For example, a DNA scaffold is susceptible to undergo oxidative DNA strand scission in a catalytic oxidation^[59,60] whilst this reaction is extensively explored with protein-based artificial metalloenzymes.

Additionally, artificial metalloenzymes can be based on a scaffold that comprises (i) an existing active site or binding pocket (*i.e.* proteins), which can be reengineered, or (ii) an active site *de novo* created (*i.e.* nucleotides), which expands greatly the number of scaffolds that can be used.

To date, proteins have been the most successful scaffold to achieve enantioselective catalysis. Avidin (Av),^[19,32] streptavidin (Sav),^[34] serum albumin (SA),^[44,61,62] apo-myoglobin (apo-Mb),^[37,38,63] papain,^[64–66] tHisf,^[14] photoactive yellow protein (PYP),^[67,68] and carbonic anhydrase (CA)^[43,54,69] are examples of protein scaffolds that have been used with considerable success. Using an existing active site or binding pocket represents an attractive approach because, after the initial design, the second coordination sphere can be in principle reengineered to optimise the performance of the catalyst (genetic optimisation). The versatility of these protein-based hybrid catalysts is illustrated by the artificial imine reductase

Scheme 1.2. Artificial imine reductase based on the biotin-avidin technology. (a) $[\text{Cp}^*\text{Ir}(\text{Biot-}p\text{-L})\text{Cl}]$ **4** pre-catalyst; (b) X-ray structure of an (*R*)-selective artificial ATH. The alanine residues at position 112 are highlighted in red, and the lysine residues at position 121, in blue; (c) The reaction conditions for the production of salsolidine **3**; (d) Proposed second coordination sphere mechanism, involving the protonation via K121 residue. Authorised reprint from the Royal Society of Chemistry.^[21]



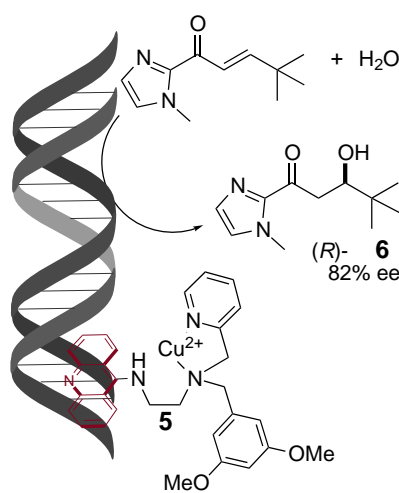
developed by Ward and co-workers. A three-legged piano stool complex, bearing an achiral aminosulphonamide ligand (Biot-*p*-L) tethered to a biotin-anchor **4**, was combined with streptavidin (Sav) for the asymmetric transfer hydrogenation (ATH) of imines (Scheme 1.2). This protein, which has an exceptionally strong affinity for biotin, contains a very large but shallow binding pocket that can accommodate both the catalyst and substrate. Chemo-genetic optimisation of the system allowed the identification of two active and selective ATHs for imine reduction, $[\text{Cp}^*\text{Ir}(\text{Biot-}p\text{-L})\text{Cl}] \subset \text{S112x}$ ($x = \text{A}$ or K). Introduction of a single point mutation at position S112 allowed to access both enantiomers of salsolidine **3** in 96% *ee* (*R*) with $[\text{Cp}^*\text{Ir}(\text{Biot-}p\text{-L})\text{Cl}] \subset \text{S112A}$ and 78 % *ee* (*S*) with $[\text{Cp}^*\text{Ir}(\text{Biot-}p\text{-L})\text{Cl}] \subset \text{S112K}$, respectively.^[40]

Although, for catalytic purposes, proteins are primarily favoured by Nature, oligonucleotides offer an attractive alternative to polypeptides as scaffolds for the incorporation of a catalytically active metal complex. A key example of this strategy is represented by the DNA-based artificial metalloenzyme introduced by Roelfes and Feringa, for the *syn* hydration of α,β -unsaturated ketones to yield enantioenriched 1,3-hydroxyketones (Scheme 1.3), a reaction that has no equivalent in conventional homogeneous catalysis. Roelfes and co-workers developed a copper complex, $[\text{Cu}(\text{diamine})]^{2+}$ **5** Lewis acid tethered to an intercalating agent (9-aminoacridine moiety), non-covalently bound to double-stranded DNA (ds-DNA). The cat-

alytic performance was optimised by variation of both the achiral ligand and the macromolecular structure.

The most selective catalyst was achieved with the sequence $d(\text{CAAAAATTTTGG})_2$ and $d(\text{GCGCTATAGCGC})_2$, and yielded the *syn* hydration product **6** in up to 82% *ee R*.^[70]

Scheme 1.3. A DNA-based artificial metalloenzyme for the *syn* hydration of enones. Authorised reprint from the Royal Society of Chemistry.^[21,70]



This approach confirmed that nucleic acids have the ability to form precise binding pockets for the specific recognition of substrates and cofactors, and to discriminate enantiomers of target molecules and bind them with high enantioselectivity.^[6]

The success in artificial metalloenzymes design relies mainly on the choice of the biomolecular scaffold, as it not only determines the type of reaction that can be implemented, but also influences the anchoring strategy.

1.1.3 Anchoring strategies

Three distinct methods concur to the design of hybrid catalysts, depending on whether the metal complex is incorporated into the chiral microenvironment provided by the biomolecular scaffold by (i) dative anchoring, *i.e.* modification of native metal cofactors,^[44,71] (ii) covalent anchoring to accessible reactive amino acid residues, typically cysteine or serine,^[18,72] and (iii) supramolecular anchoring, *i.e.* conjugation of metal ligands to native substrates,^[19,34,73] (Figure 1.2). A fourth anchoring strategy, dual anchoring, can be achieved by combination of two of the methods above described.^[74]

The aforementioned examples (Section 1.1.2) of artificial metalloenzymes based on the biotin/streptavidin interaction and DNA-based asymmetric catalysis involve supramolecular anchoring.

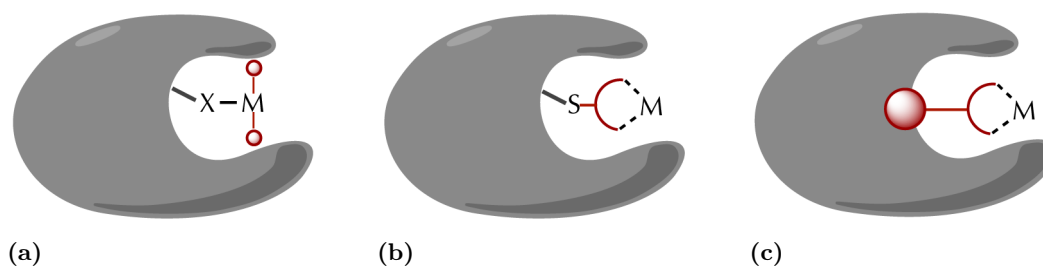


Figure 1.2. Representation of the three anchoring strategies: (a) Dative anchoring; (b) Covalent anchoring; and (c) Supramolecular anchoring. “M” denotes the catalytically active transition metal. The chemically synthesised first coordination sphere is highlighted in red, and the biomolecular scaffold in grey.

To unveil the influence of the anchoring of the organometallic moiety in the positioning of the metal cofactor and steric control of the substrate entrance will help to design better functional hybrid catalysts, with predicted catalytic activity and selectivity.^[75]

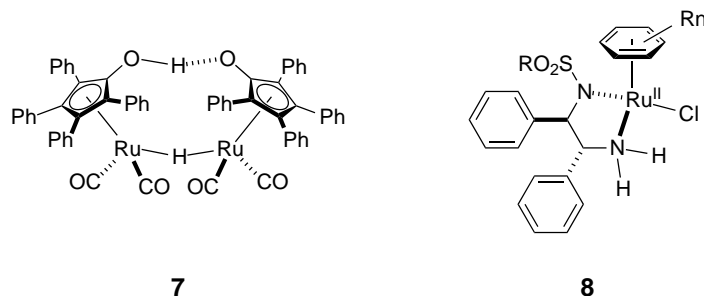
1.2 Catalytic scope

Countless reviews and book chapters covering the field of artificial metalloenzymes can be found in literature.^[6,20,21,76] This chapter section focuses on key examples for two reactions catalysed by traditional transition metal catalysts, and ultimately by artificial metalloenzymes.

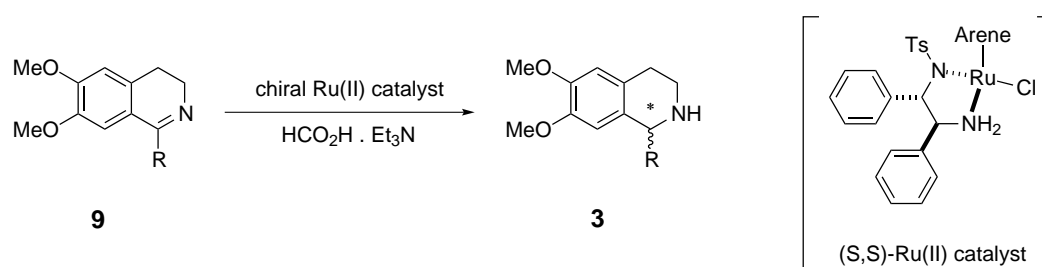
1.2.1 Asymmetric transfer hydrogenation of imines

Enantiopure amines are privileged compounds due to their application in pharmaceutical, agrochemical, food and fine chemical industries as *e.g.* synthetic intermediates in chemical syntheses.^[77] The asymmetric transfer hydrogenation (ATH) for reduction reactions has emerged as a powerful and efficient route to access such compounds. ATH is an efficacious strategy to reduce ketones or imines by a hydrogen donor other than hydrogen gas, with the aid of homogeneous transition metal catalysts.^[78–80]

The past two decades have witnessed the development of some of the most successful and general catalysts for the asymmetric transfer hydrogenation. Until the 1990s, enantioselectivity and conversion of the reported catalysts (*e.g.* Wilkinson’s catalyst, $[\text{Rh}(\text{PPh}_3)_3\text{Cl}]$)^[81–83] were, in general, low. Remarkable advancements in ATH emerged with Shvo’s diruthenium catalyst **7**^[84] and Noyori’s ruthenium(II)-based catalyst ($[\text{RuCl}(p\text{-cymene})[(S,S)\text{-TsDPEN}]$ **8**), Scheme 1.4.^[85]

Scheme 1.4. Shvo **7** and Noyori **8** complexes used for the transfer hydrogenation reaction.

Noyori's catalysts, Ru^{II} catalysts bearing monotosylated 1,2-diamines or amino alcohols, were successfully applied in the asymmetric hydrogenation of a wide range of imines, affording enantioselectivities up to 98%.^[85] Since then, efforts have been devoted in the last decade toward the development of catalytic systems for ATH of ketones and imines.^[86] Recent progress in the field has led to the discovery of related ligands and catalysts able to provide high efficiency (activity and enantioselectivity, in some cases up to 99% *ee*) with low catalyst loadings, in the asymmetric hydrogenation of imines. The main catalytic systems encountered for the hydrogenation of cyclic imines are based on an arene or tetramethylcyclopentadienyl group (Cp)-metal complex with a chiral bidentate ligand (monotosylated 1,2-diamine or amino-alcohols) and a halide ligand.^[87,88] Ruthenium is the most studied metal, but catalytic systems originally designed for transfer hydrogenation including iridium or rhodium^[89] were also applied in the reduction of cyclic imines. Alcohols (*e.g.* isopropanol) and the mixture of formic acid and triethylamine (T-F) have been the most popular solvents as they also act as hydrogen sources.^[90–92] Upon the increasing demand for efficient and environment-friendly syntheses, organometallic catalysis in aqueous media has been extensively studied and developed. Several groups have hitherto reported the ATH of ketones in aqueous formic acid/sodium formate, with water-soluble complexes.^[93] Furthermore, these studies on catalytic systems for the reduction of cyclic imines led to consider tetrahydroisoquinoline-based scaffolds as benchmark substrates for the implementation of new reactions. The synthesis of salsolidine^[94] **3** was considered to be a model substrate (Scheme 1.5).

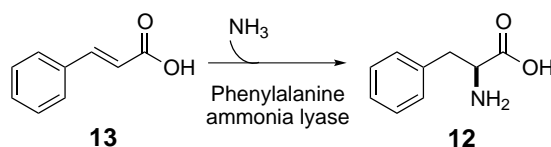
Scheme 1.5. Asymmetric transfer hydrogenation of cyclic imines: preparation of salsolidine **3**.

During the course of this thesis, a new highly active artificial metalloenzyme, $[\text{Cp}^*\text{Ir}(\text{Biot-}p\text{-L})\text{Cl}] \subset \text{streptavidin}$ was developed for the ATH of cyclic imines. After chemo-genetic optimisation, up to 96% *ee* (*R*) with $[\text{Cp}^*\text{Ir}(\text{Biot-}p\text{-L})\text{Cl}] \subset \text{S112A}$ and 78% *ee* (*S*) with $[\text{Cp}^*\text{Ir}(\text{Biot-}p\text{-L})\text{Cl}] \subset \text{S112K}$ were obtained for the reduction of a precursor of salsolidine.^[40]

1.2.2 Asymmetric reductive amination of α -keto acids

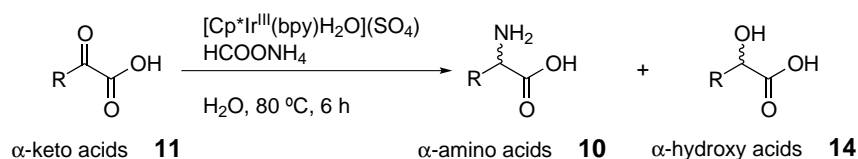
Optically active pure amino acids play an important role as intermediates in the pharmaceutical industry and agrochemistry, both of each which require a high degree of purity and large quantities of the compound(s).^[95,96] In a worldwide market predicted to hit US \$12.8 billions by 2017,^[97] the synthesis of α -amino acids **10** has attracted a lot of attention in recent years.^[98] Several synthetic routes have been reported in literature for the synthesis of enantiomerically pure amino acids from α -keto acids **11**, some of which involve multiple steps or use costly or hazardous reagents.^[99–101] No methodology consensus has yet emerged for the asymmetric synthesis of all non- and proteinogenic amino acids.^[96] For example, the synthesis of L-phenylalanine **12** can be achieved from the biotransformation of (*E*)-cinnamic acid **13**, with the enzyme phenylalanine ammonia lyase as catalyst (Scheme 1.6).^[102] However, on an industrial scale, L-phenylalanine **12** is obtained via an economical fermentation process, with phenylalanine over-producers.^[103]

Scheme 1.6. Lyase reaction for the production of L-phenylalanine **12**.



Asymmetric reductive amination represent a straightforward approach for accessing non-proteinogenic (in particular) α -amino acids from the corresponding α -keto carboxylic acids, in enantiomerically pure form,^[96,104–106] using chemo-catalysts as well as enzymes as catalyst components. The first non-enzymatic highly chemoselective synthesis of α -amino acids by reductive amination of α -keto acids, using HCOONH_4 in water, was reported by Fukuzumi *et al.*^[107] By employing an acid-stable mononuclear iridium hydride complex, all major types of α -amino acids were synthesised, in pH-controlled reactions (Scheme 1.7).

Scheme 1.7. Synthesis of α -amino acids by reductive amination of α -keto acids **11** with ammonia in water.



The breakthrough of this approach was the use of ammonium and formate as nitrogen and hydride sources, respectively, to produce α -amino acids in water.

1.3 Biological scope

The results of reactions catalysed by artificial metalloenzymes have proven that asymmetric catalysis can be achieved by selecting the proper protein as host.^[32,34] Thus, protein scaffolds should be stable, both over a wide pH range and at high temperature. In addition, crystal structures of the biomolecule are crucial for genetic optimisation.

1.3.1 Streptavidin

Streptavidin (Sav) is the bacterial counterpart of the biotin-binding protein, avidin. Sav is a homotetrameric protein, produced extracellularly by the bacterium *Streptomyces avidinii*. Its extremely high affinity to biotin and its analogs ($K_a \sim 10^{13} \text{ M}^{-1}$)^[108] is one of the strongest non-covalent interactions found in biological systems. The biological role of (strept)avidin appears to be that of a biotin scavenger, inhibiting bacterial proliferation. The biotin-streptavidin system has been commonly described as the molecular version of velcro, and it is widely applied in life science research, as well as in bio- and nanotechnology.^[109–111]

Structure

Although structurally and functionally comparable to avidin, streptavidin (Sav, MW [Da] 16,425, pI 6.2, 159 amino acids) is a non-glycosylated tetrameric protein with eight stranded β -barrels, which fold to give an antiparallel β -barrel tertiary structure (Figure 1.3.a).

The biotin-binding site is located at one end of each β -barrel, and is formed by inner amino acids of the barrel and a tryptophan residue from the neighbouring subunit.^[112] Ergo, streptavidin homotetramer acts as a dimer of dimers.^[113] The tryptophan from the adjacent monomer (Trp120) act as a hydrophobic lid, and is involved in inter-monomeric contacts that stabilise the tetrameric protein in the bound state.^[114] The biotin-binding pocket is located at the centre of each subunit in the tetrameric protein, at $\sim 9 \text{ \AA}$ below the protein surface, and is partially occupied by five water molecules in the absence of biotin, which are displaced upon biotin binding.^[115]

The hydrogen bonding network (formed between the ureido oxygen and the hydrophilic residues Asn23, Ser27 and Tyr43, and between the ureido nitrogen and Ser45 and Asp128), which along with the hydrophobic compartment created by aromatic amino acids (Trp79, Trp92, and Trp108, and Trp120 from the adjacent monomer, Figure 1.3.b) and a flexible binding loop (L3,4) give rise to tight biotin anchoring.^[116,117]

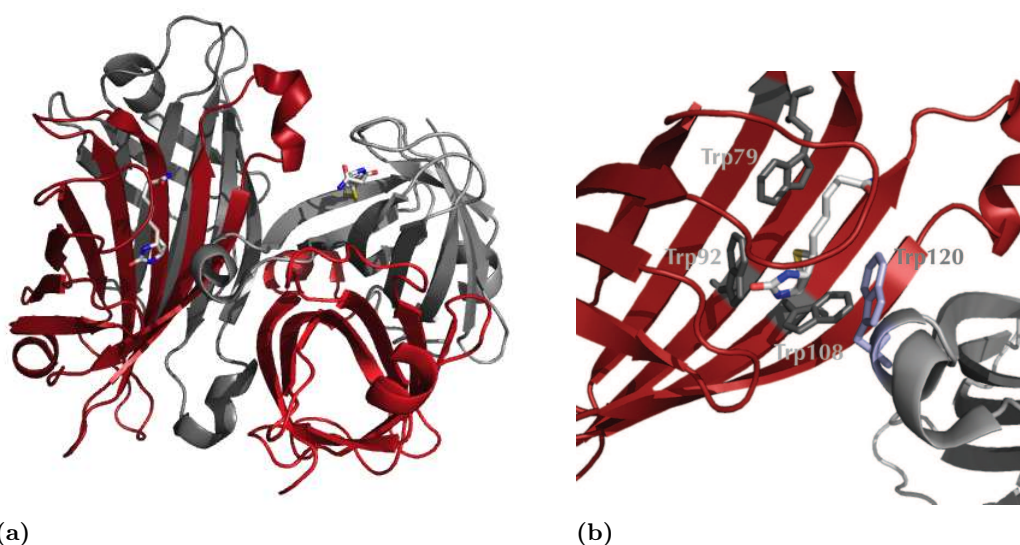
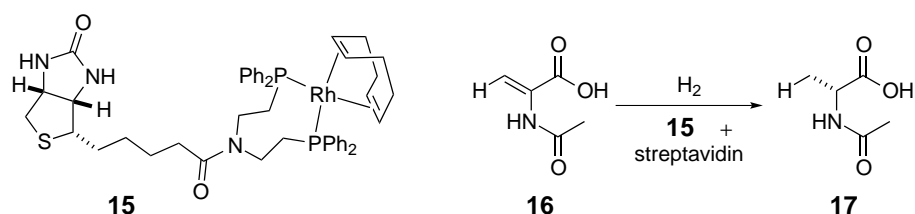


Figure 1.3. Ribbon diagram of (a) The binding pocket of streptavidin. Each monomer is represented in red or grey, and each dimer is formed by a pair of both coloured monomers (red/grey = dimer); the biotin molecule is shown as sticks, in two opposite monomers; (b) The hydrophobic interactions with biotin. The residues Trp79, Trp92 and Trp108 of the monomer · biotin bound (in red) are represented in grey sticks, and the residue Trp120 from the adjacent monomer is represented in blue stick, (PDB code: 2QCB).

Anchoring strategy

In the fledging period of the field of asymmetric catalysis, Whitesides & Wilson devised an innovative supramolecular catalytic system that took advantage of the very strong non-covalent interaction between avidin and biotin. The strength of the avidin-biotin interaction ensured quantitative binding of a biotin functionalised with a phosphine-based rhodium catalyst into the chiral biomolecular environment (supramolecular anchoring). Whitesides reported the asymmetric hydrogenation of an alkene, N-acetamidoacrylate, with catalytic amounts of the artificial metalloenzyme, that yielded 41% *ee* for (*S*)-N-acetamidoalanine, with full conversion.^[19] However this concept seemed to have been forgotten for almost twenty years until, in 1999, Chan and co-workers revived it again by linking a chiral Pырphos-Rh(I) to biotin for the hydrogenation of itaconic acid, and achieved moderate enantioselectivity for the preparation of methylsuccinic acids.^[25] Yet, the real breakthrough in the field of artificial metalloenzymes was delivered by Ward and co-workers, in 2003, when they reported on the generation of an artificial metalloenzyme based on the biotinylated rhodium-diphosphine complexes **15** in streptavidin as host protein, rather than the original host avidin (Scheme 1.8). A chemogenetic approach allowed the optimisation of the enantioselectivity for the hydrogenation of acetamidoacrylic acid **16** (up to 96% *ee* in favour of (*R*)-acetamidoalanine in Sav S112G, **17**).^[32,34] This seminal discovery opened the field of artificial metalloenzymes, and has led to the development of new hybrid catalysts based on the biotin-streptavidin system.

Scheme 1.8. Optimisation of Wilson & Whitesides approach, using a streptavidin adduct of **15** and site-specific mutagenesis.



Biomolecular scaffold

In summary, the biotin-streptavidin system is successful, for the following reasons: (i) the strong biotin-streptavidin affinity ($K_a \sim 10^{13} \text{ M}^{-1}$)^[108,118] allows unambiguous positioning of the organometallic moiety into the protein scaffold, and subsequent improvement of the catalytic performances;^[109] (ii) the biotin affinity is not dramatically affected by the derivatization of the valeric chain of biotin, by either introduction of linkers or modulation of the chelators;^[119,120] (iii) in the presence of biotin, streptavidin is exceptionally stable at extreme pH (> 1.5), high temperatures ($> 90 \text{ }^\circ\text{C}$), high concentrations of organic solvents ($> 50\%$ ethanol), and in the presence of surfactants (sodium dodecyl sulphate, SDS);^[55,114] (iv) the flexibility to modulate the catalyst reactivity by manipulation of the binding site of the protein; (v) the rapid estimation of free-biotin binding sites by titration of the protein with biotin-4-fluorescein;^[121] and (vi) Sav is easy to over-express in *E. coli* (about 200 mg per litre of culture) and to purify by affinity chromatography.^[114,122]

1.3.2 Human carbonic anhydrase II

Carbonic anhydrases (CAs, carbonate hydro-lyase, EC 4.2.1.1) are ancient metalloenzymes^[123] that are present all over the phylogenetic tree, with five distinct evolutionarily families: α -, β -, γ -, δ - and ϵ -carbonic anhydrase isoforms.^[124,125] The first enzyme was identified, in 1933, in red blood cells of cows by Meldrum and Roughton.^[126] Since then, CAs isozymes have been found to be abundant in all mammalian tissues (α -class), plants (β -class), algae and bacteria (γ -, δ - and ϵ -classes).^[124] Although α -CA genes are also found in many plants, algae and bacteria (*e.g.* in the bacterium *Neisseria sicca*, NsCA),^[127] they predominate among mammals, and are the only CA gene family expressed by vertebrates. The α -CA class is the best studied group, although recent reviews indicate a rapid expansion of knowledge (structural information, mechanism, and inhibitors) for other CA families.^[128–131]

In mammals, 16 α -CA isoforms were described so far,^[132] and from these, seven genetically distinct enzymes have been identified in humans. The presence of carbonic anhydrase isozymes in so many different tissues, organs and cells makes CAs of particular pharmaceutical interest

for the development of new therapeutic agents.^[133] The activation or inhibition of CAs can be used in the treatment of several diseases, such as obesity, osteoporosis, glaucoma and cancer.^[131,134,135] Although different in their sequences, all human carbonic anhydrases are catalytically active, and act as efficient catalysts for the reversible hydration of carbon dioxide to bicarbonate (Equation 1.1):

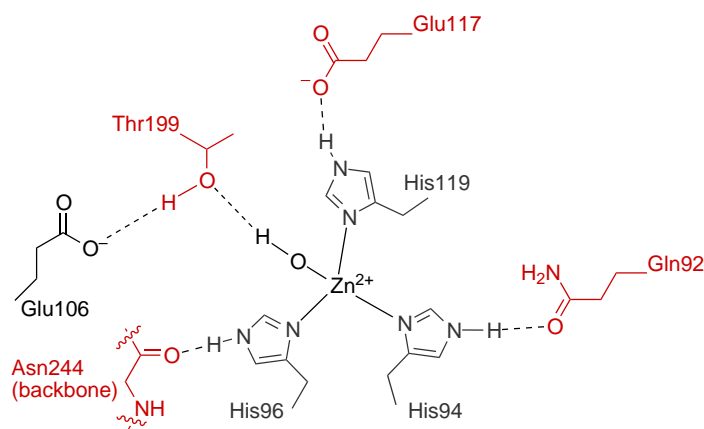


They are among the fastest enzymes known, with a maximum turnover number, k_{cat} , for the CO_2 hydration reaction that exceeds $1 \times 10^{-6} \text{ s}^{-1}$ (reviewed by^[136]).^[137] Therefore, they play a crucial role in a myriad of physiological processes as varied as respiration and transport, photosynthesis, ionic, acid-base and fluid balance, calcification, metabolism and cell growth, among others.^[136]

Structure

Human carbonic anhydrase II (hCAII, MW [Da] 29,227, pI 7.4, 259 amino acids) is a zinc-containing enzyme, with a predominantly β -sheet structure that encloses a large solvent sequestered hydrophilic cavity (approximately $55 \times 44 \times 39 \text{ \AA}$, Figure 1.4).^[138,139] The active site of hCAII is located in this deep, cone-shaped cleft that reaches almost to the centre of the enzyme.^[139] Near the bottom of the cavity lies a Zn^{II} ion tetrahedrally coordinated by three first-shell amino acid ligands (His94, His96 and His119, and referred as $\text{Zn}^{\text{II}}(\text{His})_3$) and a single solvent molecule, called the “zinc-water”, that ionises to a hydroxide ion at physiological pH (Figure 1.4.a).^[140–142] All three “direct ligands” are preordered by a hydrogen bond network. This zinc distorted coordination polyhedron is conserved among CA isozymes.^[143]

Scheme 1.9. Direct and indirect ligands of the active site zinc ion. In grey: first-shell ligands, and in red: second-shell ligands. These ligands govern the electrostatic environment of the Zn^{II} ion, and modulate the chemistry of the $\text{Zn}^{\text{II}}-\text{OH}^-$.^[144]



In the native enzyme, the zinc ligands are fully saturated by hydrogen bond (H-bond) networks, forming a second-shell of “indirect” ligands (Scheme 1.9). This conserved active site hydrogen bond network enhances the nucleophilicity of the zinc-water bound molecule, and orientates the substrate in a favourable location for the nucleophilic attack (Scheme 1.9).^[133,145,146] On one side of the active site of the protein lines predominantly hydrophobic residues and on the other, hydrophilic residues (Figure 1.4.b). Without a ligand bound, the pocket is filled by a network of water molecules, which provides direct contact to the solvent surrounding of the enzyme.

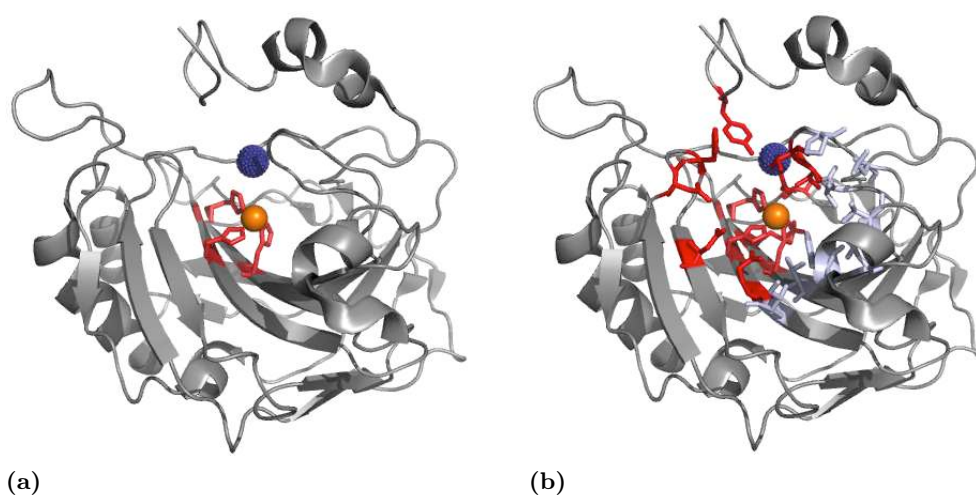


Figure 1.4. Ribbon diagram of (a) The active site of human carbonic anhydrase II. The active site zinc is shown as orange space-filled sphere. It is coordinated by three histidine residues (His94, His96, and His119, coloured in red) and by a water molecule (Wat263, coloured in blue). (b) The amphiphilic binding pocket of human carbonic anhydrase II. The zinc ion is shown as orange space-filled sphere; hydrophobic residues (Ile91, Val121, Phe131, Val135, Leu141, Val143, Leu198, Pro202, and Leu204) in blue, and hydrophilic residues (Tyr7, Asn62, His64, Asn67, Gln92, His94, His96, Glu106, Glu117, His119, Thr199, and Thr200) in red (PDB code: 1G54).

Anchoring strategy

Human carbonic anhydrase II has a narrow range of transition metals that yield an active form of the isozyme.^[147–149] Metal-substituted hCAII can be obtained by removal of the active site metal by chelators, and replacement with different divalent transition metal ions (*e.g.* Mn^{II} , Fe^{II} , Co^{II} , Cu^{II} , Ni^{II} or Mo^{II}).^[150] The coordination geometry of the surrogates in the protein plays an important catalytic role. Studies on the catalytic activity of metal-substituted wild-type hCAII for the hydration of carbon dioxide indicated that only metals able to easily adopt a tetrahedral coordination sphere, Zn^{II} and Co^{II} , provided significant catalytic enhancement.^[151] Co^{II} -bound protein retained approximately 50% of the activity compared to the native enzyme, whilst the other metal ions derivatives (Mn^{II} , Fe^{II} , Cu^{II} , and Ni^{II}) had no or negligible activity.^[147–149,152] This approach was more recently investigated

by Kazlauskas *et al.* to modify the catalytic activity of hCAII. The substitution of Zn^{II} (Figure 1.5.a) by Mn^{II} (Figure 1.5.b) originated a novel peroxidase able to enantioselectively epoxidise olefins conjugated to an aromatic or aliphatic carbon.^[43,69,153]

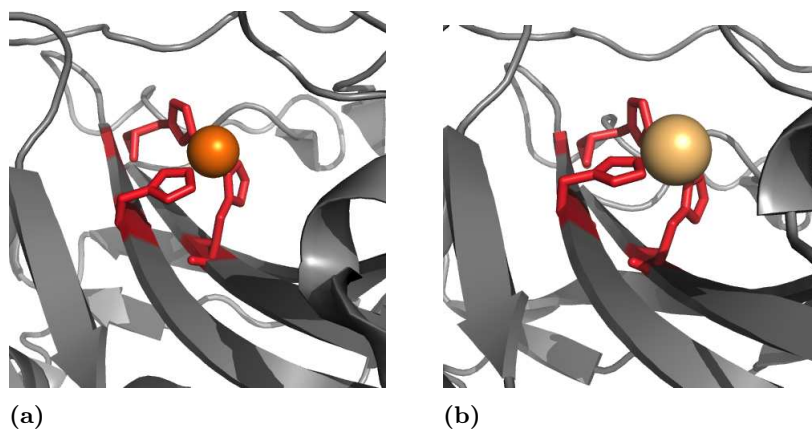
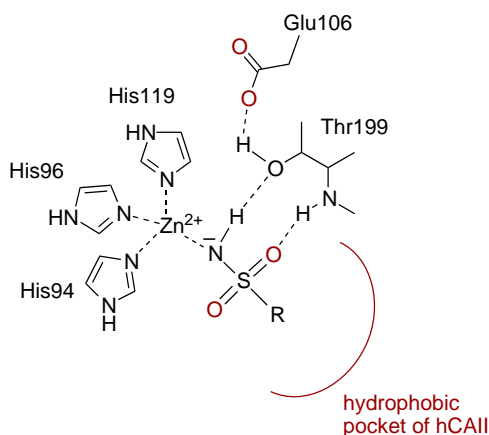


Figure 1.5. Ribbon diagram of (a) Native hCAII with Zn^{II} as active metal (PDB code: 1G54) (b) Modified hCAII with Mn^{II} as active metal (PDB code: 1RZD).

Inhibitors

Human carbonic anhydrase II is a target for therapeutic drugs^[133,154] such as sulphonamide derivatives,^[155,156] and inorganic and organic anions.^[131] In the case of sulphonamides, an ionised sulphonamide nitrogen displaces the zinc-bound hydroxide to form a stable enzyme-inhibitor complex with submicromolar to nanomolar affinity (Scheme 1.10).^[134,157–164] Thus, sulphonamide inhibitors can be extremely useful in the development of new metal-inhibitor assemblies.

Scheme 1.10. Tetrahedral adducts are generated by substitution of the non-protein ligand by unsubstituted sulphonamides and their bio-isosteres bound to the Zn^{II} ion of the enzyme.^[133]



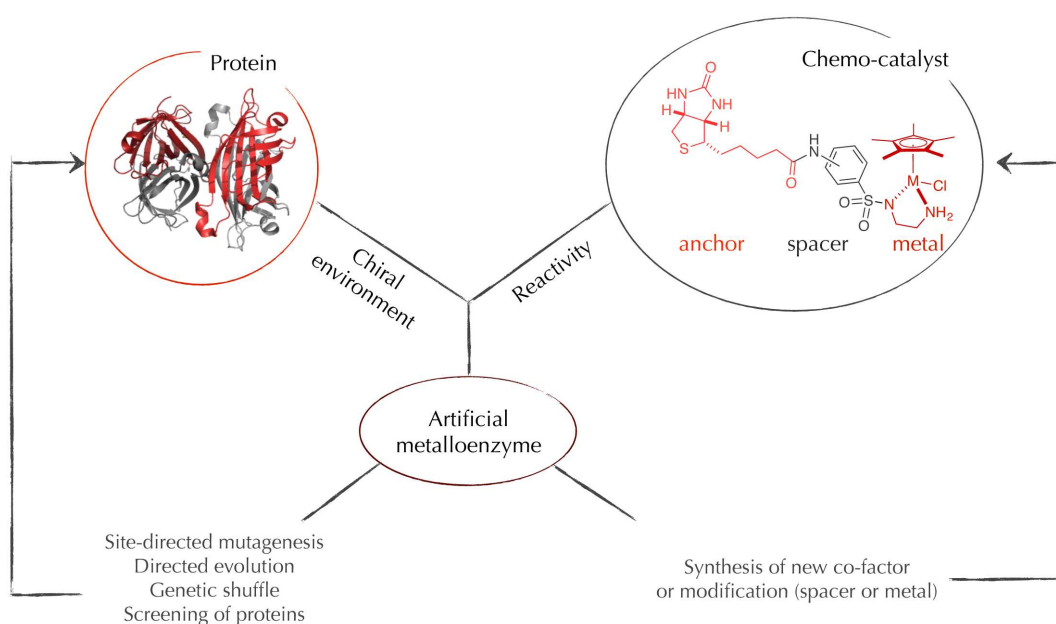
Biomolecular scaffold

Human carbonic anhydrase II is particularly well-suited for engineering a new artificial metalloenzyme, for the following reasons: (i) hCAII is a monomeric protein with a deep binding pocket in which the organometallic complex can be completely embedded; (ii) the flexibility to manipulate its active site via rational design,^[140,145] thus modulating the catalyst reactivity; (iii) the compatibility of human carbonic anhydrase II inhibitors with soft transition metals, the possibility to derivate them, and their low dissociation constant;^[119,165] (iv) hCAII is easy to over-express in *E. coli* and to purify;^[145] (v) its promiscuous esterase activity allows monitoring the rate of *p*-nitrophenyl acetate hydrolysis,^[166] and consequent binding profiles; (vi) x-ray determination of hCAII structure is well-established,^[167] and NMR studies may elucidate the structure of the protein · inhibitor complex in solution.^[168–170]

1.4 Tools for optimisation

The creation of artificial metalloenzymes has proven to be extremely versatile. The concept combines the catalytic power of a transition metal complex with the chiral architecture of a biomolecular host, and offers a strategy to improve or evolve independently one and the other. This strategy has been coined by Distefano and Häring as the “chemo-genetic approach” (Scheme 1.11).^[22]

Scheme 1.11. Achiral ligands and protein scaffolds as source of novel hybrid catalysts.^[64]



The chemo-genetic approach is defined as the optimisation of (i) the chemical catalyst by rational design, based on the structural information of the hybrid catalyst as a whole;^[31,35] and (ii) the biopolymer scaffold by rational design or in combination with evolutionary approaches, *e.g.* designed and directed evolution.^[9,76,171] These combinatorial approaches would benefit from a combination with high-throughput screening.

In this subsection, the various strategies that can be followed for the design and optimisation of the host protein are presented.

1.4.1 Directed evolution

Directed evolution or step-wise evolution is a strategy to rapidly evolve enzymes that exhibit new or improved properties for a specific application, mimicking Nature's evolution in a shorter time scale.^[172] A pre-requisite of this method is the capacity to generate large protein libraries by random mutagenesis in combination with high-throughput screening (HTS) to identify the best variants.^[173] In directed evolution, explicit understanding of either the structure or the mechanism of action of the enzyme is not required.^[174] Genetic diversity is achieved by performing iterative cycles of (i) random mutagenesis (*e.g.* by error-prone PCR); (ii) expression of the library of variants; and (iii) functional screening of the library and identification of variants with novel or improved properties.^[171,175] The successful outcome of a directed evolution experiment depends on the creation of a library of diverse variants. More rounds of randomisation and selection follow the screening and selection of the most active or selective mutants, until hybrid catalysts with improved properties are obtained.^[23,42]

Through directed evolution approach, positions remote from the binding site are also probed. On more than one occasion, it has been shown that subtle changes distanced from the binding pocket had a dramatic effect on the catalyst activity.^[176] However, this approach presents some drawbacks and challenges. As the library size increases, the project becomes increasingly dependent on high-throughput techniques, thus reliable screening methodologies should be developed and validated. Otherwise, the quality and fitness of the protein library, *i.e.* the number of hits and degree of hybrid catalyst improvement, can be a disappointment.^[171] Hence, highly efficient and chemo-selective anchoring of the organometallic complexes are required to avoid intermediate purification steps.^[76] To overcome some of the aforementioned pitfalls, Reetz and co-workers introduced the combinatorial active-site saturation test (CAST). This focused library technique was directed at the residues directly associated with the binding site. Reetz *et al.* demonstrated the potential of the CAST methodology by fine-tuning the enantioselectivity of an artificial metalloenzyme based on a biotinylated rhodium diphosphine complex embedded into streptavidin. After three iterative cycles of mutagenesis, the enantioselectivity was increased from 23 to 65%.^[177]

1.4.2 Designed evolution

Whereas in directed evolution, variations of a gene are created randomly, in rational design, mutations are inserted rationally built upon structural informations of the specific biomolecular host. Therefore, modification of the protein scaffold is mainly based on the understanding of the structural and mechanistic consequences of one specific change or set of changes. The current knowledge of structure-function relationships in proteins is still insufficient to make rational design a robust approach.^[178] However, computational design is increasingly applied in biocatalysis to obtain guidelines to direct mutagenesis efforts.^[179] Designed evolution^[180] combines rational design and combinatorial screening, which leads to the evolution of the protein.^[181] Guided by structural information of the hybrid metalloenzyme or the apo-scaffold, critical residues are identified and then subjected to saturation mutagenesis, either consecutively or simultaneously. The generated small-sized library is then combined with the organometallic moiety, and screened for the desired catalytic activity. Ward *et al.* have successfully applied this approach to the design of *e.g.* artificial hydrogenases, with improved enantioselectivities (96% *ee*).^[40,171] If intuition (rational design) and computational methodology can be applied to design desired (artificial) enzymes for more complex reactions, directed evolution can be used to fine-tune activity.^[182]

1.4.3 Screening techniques

In order to identify the best recombinant protein, a screening or selection procedure is required, which involves the development of a high-throughput assay that is sensitive to the properties targeted in the directed evolution process, and allows the identification of positive hits. Colorimetric or fluorometric high-throughput screening (HTS) techniques link enzyme activity to an easily detectable chemical response. They can be qualitative, if for example, the assay is colony based and the signal only delivers a “yes or no” answer. A successful colorimetric high-throughput colony-based solid phase assay has been developed by Turner and co-workers. The method was based on a coupled horseradish peroxidase (HRP) assay, using 3,3'-diaminobenzidine (DAB) as dye on the solid phase, and pyrogallol red (PGR) as dye in the liquid phase.^[183]

An optimum HTS approach, for artificial metalloenzymes based on biotin-streptavidin technology, can be described as a three-step process, with a feedback loop: (i) fast preparation of organometallic moieties and biomolecular scaffolds, with systematically varied properties; (ii) testing of the catalytic properties of the hybrid catalysts; and (iii) processing and evaluation of the experimental data, for modification and improvements of the next generation of hybrid metalloenzymes. The loop is repeated until certain criteria are fulfilled, and positive hits are scaled-up for detailed evaluation and characterisation.^[30] This approach is well-suited for

the design and production of small libraries of hybrid catalysts, as both chemical and genetic diversity can be exploited. In a larger scale, however, it is not applicable, since each variant is examined separately for the product formation using analytical methods such as HPLC, which is considered as low-throughput technique. The time-consuming sample preparation and analytical process is presently the only alternative, since a colorimetric assay is not available.

1.4.4 Expression systems

The choice of the expression system and an early assessment of process scalability issues have become the prime concern to minimise risk factors associated with protein production. To date, in laboratory-scale, microbial and yeast cell expression systems are the most common systems being utilised. Both of these systems have their own advantages and disadvantages for the production of functional streptavidin.

Escherichia coli

In many cases of heterologous protein expression, the best choice of host system are bacteria, because of their rapid growth rate, low demands on growth medium, and ease of genetic modification.^[184] An optimised protocol to produce soluble and functional T7-tag mature streptavidin in *Escherichia coli* has been developed by Ward *et al.* The use of BL21(DE)3pLysS cell strain, optimisation of critical parameters such as glucose concentration, pH and time of induction, and a single denaturing-renaturing step and affinity chromatography reproducibly yielded 230 mg/L of soluble protein.^[122] One disadvantage of bacterial cells in the production of streptavidin is the laborious and time-consuming process of purification, which seriously limits the applicability of directed evolution and high-throughput screening (HTS) of artificial metalloenzymes.

In the following subsections, an overview of the techniques developed to optimise the production process of streptavidin, namely the purification processes, are presented.

Crude protein The first step in the purification of a cytoplasmic recombinant protein expressed in *E. coli* is the lysis of the cell to release the target protein. The widely and routinely used cell lysis methods involve mechanical/physical or biological/chemical techniques. The recovery and yields from these techniques are often variable. The physical lysis of bacteria can be achieved by sonication, homogenisation (French press and Manton-Gaulin homogeniser), or freeze/thaw of the cells. Homogenisation and sonication generate heat and/or foaming, which are detrimental to many proteins. The freeze/thaw technique involves flash-freezing

the cell resuspension in a dry ice/ethanol or liquid nitrogen bath, and then thawing it at room temperature or 37 °C. This method can be lengthy since many cycles of freeze/thaw are necessary for efficient lysis, but it is effective for the release of recombinant proteins expressed in the cytoplasm of *E. coli*. Enzymatic lysis is achieved by digestion of the peptidoglycan layer of the bacterial cell wall upon addition of lysozyme or the use of a specific *E. coli* strain containing the T7 lysozyme gene, *e.g.* BL21(DE)3pLysS or Rosetta(DE)3pLysS, which provides additional stability to the target gene but also allows the lysis of the cells by quick freeze/thaw treatment. Osmotic lysis by addition of specifically developed detergent-based solutions, composed of particular types and concentrations of detergents, buffers, salts and reducing agents, is usually efficient in both lysing and solubilising the cells. However, most cell lysis or cellular disruption methods cause the release of biological molecules, including nucleic acids (DNA and RNA), organelles, native proteins and lipids from inside the bacterial cells. These contaminants have to be removed, as they can cause viscosity problems or interfere with subsequent purification steps or screening tests. DNA can be digested by addition of DNase I to the cell lysate or removed by precipitation with polyethyleneimine or protamine sulphate followed by centrifugation. The remaining unbroken cells, lipids, and particulate matter can be removed from the protein solution by an ultracentrifugation step.

Although efforts have been deployed to implement high-throughput screening of protein-based artificial metalloenzyme libraries using crude protein extracts, such efforts have been hampered by the presence of potential catalyst inhibitors, *e.g.* reduced glutathione, GSH.^[76]

Protein precipitation Until the beginning of the twentieth century, non-chromatographic methods (*e.g.* precipitation, crystallisation and partitioning) were the only available techniques for protein purification. Progress in chromatography revolutionised the approach to purify proteins, and nowadays, those early techniques are widely used, either separately or in combination with chromatography methods (*e.g.* ion exchange, affinity, hydrophobic interaction and reversed phase, immobilised metal affinity, and size exclusion), to ensure the quality, quantity and purity of the target protein. Selective precipitation of proteins can be used to recover the protein of interest from a crude extract.^[185] The solubility of the protein of interest can be reduced by ionic precipitation (addition of ammonium sulphate),^[186,187] pH (trichloroacetic acid (TCA) or trichloroacetic acid/sodium deoxycholate (TCA-DOC)),^[188–190] and organic solvents (cold ethanol (Cohn’s method), acetone or dimethoxyethane (DME)).^[191,192] Protein precipitation by organic solvents is mainly due to charge repulsion, whilst the mechanism of salting out is ascribed to the influence of the salt on the water structure and surface tension.^[193] The pH exerts its effect through the decrease of the protein solubility when at its isoelectric point (pI).^[194]

The first step for the purification of a cytoplasmic protein by precipitation is the preparation of a crude lysate, which contains a complex mixture of proteins, macromolecules, cofactors

and nutrients from the cell cytoplasm and culture medium. The crude protein extract is then obtained by removal of cellular debris, from the lysis of the host cells, by centrifugation and recovery of the supernatant. The second step involves addition of a precipitation agent, and recovery of the precipitated protein by centrifugation.

Small-scale purification A more recent approach, implemented by structural genomics initiatives and commonly used in pharmaceutical development, focuses on small-scale high-throughput protocols for cloning, stable expression, and purification of proteins. On a small-scale, the parallel processing typically involves the use of multi-well plates.

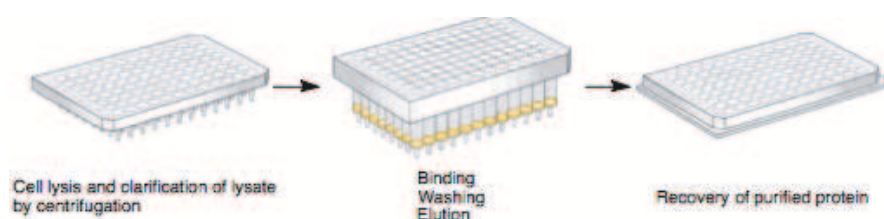
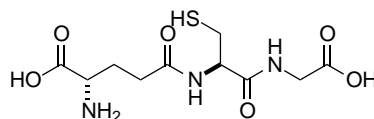


Figure 1.6. Processing flowchart small-scale purification of protein. The recombinant protein can be expressed in deep well plates, shake flasks, or in a fermentor.

It is a typical engineering solution that can be applied at any stage (up- or downstream) of the process. Concerning high-throughput purification, biotech companies have facilitated the process by commercialising ready-to-use filter spin plates pre-packed with affinity resins, that can be applied in a manual or automated platform. This method is commonly performed on *E. coli* cultures expressing the target protein(s) grown in 96 deep-well plates, generating *e.g.* a library of variants for mutational analysis. Cells can be lysed directly in their growth medium, and applied to a filter plate affinity resin, thus eliminating the centrifugation and pipetting steps. However, this system can also be used with larger cultures, following the standard cell lysis and soluble fraction separation steps. The crude extracts are pipetted into the pre-packed wells of the plate, and the wash and elution steps applied either by centrifugation or vacuum (Scheme 2.2).

Neutralisation of reduced glutathione Crude proteins preparations from bacterial cell cytosols may be laden with reduced glutathione (GSH, Scheme 1.12). This tripeptide (L- γ -glutamyl-L-cysteinyl-glycine) is one of the most abundant and ubiquitous non-protein thiol in prokaryotic and eukaryotic cells. In bacteria, it can accumulate to concentrations exceeding 10 mM.^[195–197] In addition to its key role in the regulation of cellular metabolism by maintaining the proper oxidation state of protein thiols, glutathione also acts as a scavenger of toxic compounds (hydroxyl radicals, chlorine compounds) and stresses (oxidative and osmotic).^[198–200]

Scheme 1.12. Structure of reduced glutathione.**18**

One way to minimise the deleterious interactions between the GSH present in the crude protein extract and the organometallic catalyst is to oxidise or neutralise the glutathione. This can be achieved using (i) chemical agents such as 1,4-benzoquinone, 1,1'-azobis(*N,N*-dimethylformamide) (*aka* diamide), and azoester, which oxidise GSH to glutathione disulphide, GSSH (the oxidised form of GSH)^[201–206] or (ii) electrophilic agents such as S_N2 electrophiles (aliphatic, allylic and benzylic halides, *e.g.* diethylmaleimide, *N*-ethylmaleimide (NEM), and bromobenzene),^[207,208] α,β -unsaturated carbonyl derivatives and vinyl sulphones (acting as Michael acceptors), and α -halogenated carbonyls,^[209–211] which form conjugates with GSH. However, these substances can be toxic and can have nonspecific effects on proteins and other components of the cells.^[212]

Advantages & disadvantages Most recombinant proteins can be cloned and expressed in *E. coli*, as it is a well-documented expression system for its advantages of (i) low cost; (ii) ease of genetic manipulation; (iii) rapid growth rate; and (iv) high protein yield. However, *E. coli* expression system has some limitations, as some target proteins (i) are insoluble and aggregate in inclusion bodies; (ii) require post-translational modifications in order to be completely functional, and the bacterium cells do not perform such modifications; and (iii) are toxic to *E. coli* cells (*e.g.* streptavidin), preventing to reach high-cell densities, thus impairing its yield in production processes.

In the context of artificial metalloenzymes, the secretion of the desired protein to either the cell periplasm or culture medium would be a step further toward *in vivo* catalysis with artificial metalloenzymes, by providing a high quality sample, free of unwanted foreign proteins and/or contaminants, albeit the overall yield would still be a bottleneck to high-throughput screening.

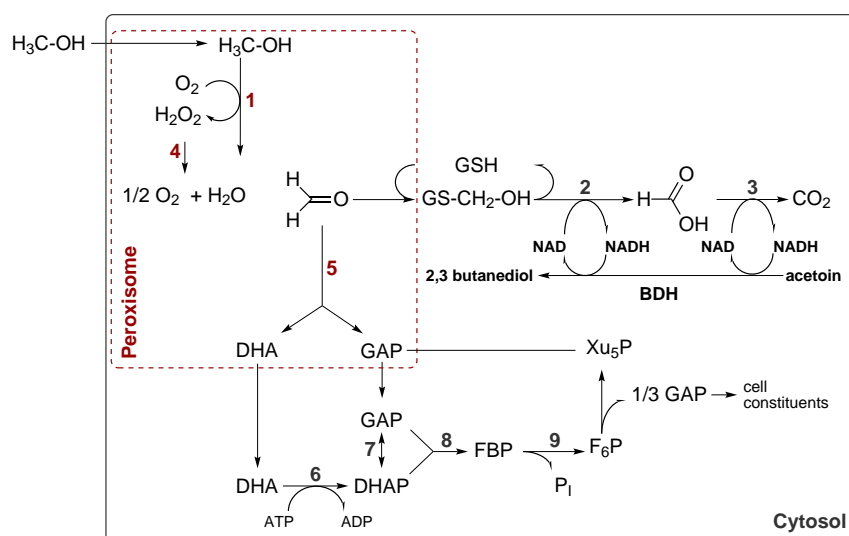
Pichia pastoris

To address the issues aforementioned, *Pichia pastoris* expression system was evaluated as expression platform for the production of streptavidin. *P. pastoris* is one of the known genera of methylotrophic yeast, which metabolises methanol as a sole carbon and energy source.

Methanol utilisation pathway The success of methylotrophic yeasts in the production of recombinant proteins relies mainly on their strong and tight regulated promoters of the methanol utilisation pathway (MUT pathway). Since the initial reactions of the MUT pathway are compartmentalised within peroxisomes, these subcellular organelles are strongly induced upon methanol as sole carbon source, and proliferate massively after induction, accounting for up to 80% of the cytoplasmic space.^[213] The peroxisomal localisation of the three key enzymes for methanol metabolism is essential to allow the methylotrophic yeast cells to grow on methanol.^[214]

In the first step of the methanol utilisation pathway, methanol is oxidised to formaldehyde and hydrogen peroxide by the enzyme alcohol oxidase (AOX, EC 1.1.3.13), and the hydrogen peroxide is broken down to water and oxygen by the enzyme catalase (CAT, EC 1.11.1.6). Both enzymes are sequestered in peroxisomes, due to the hydrogen peroxide toxicity (see Figure 3.1). Formaldehyde is either oxidised by two subsequent dehydrogenase reactions (dissimilation pathway) or assimilated in the cell metabolism by condensation with xylulose 5-phosphate (Xu₅P). The latter reaction is catalysed by dihydroxyacetone synthase (DAS, EC 2.2.1.3), which converts Xu₅P and formaldehyde into the C₃-compounds dihydroxyacetone (DHA) and glyceraldehyde 3-phosphate (GAP). These subsequent compounds are further metabolised in the cytosol.^[215–218] The AOX protein is only expressed during methanol metabolism, but then at very high levels.

Scheme 1.13. Methanol utilisation pathway of *Pichia pastoris*. **1:** Alcohol oxidase (AOX); **2:** Formaldehyde dehydrogenase; **3:** Formate dehydrogenase; **4:** Catalase (CAT); **5:** Dihydroxyacetone synthase (DAS); **6:** Dihydroxyacetone kinase; **7:** Triosephosphate isomerase; **8:** Fructose 1,6-bisphosphate aldolase; and **9:** Fructose-1,6-bisphosphatase. BDH: Butanediol dehydrogenase; GSH: Glutathione; DHA: Dihydroxyacetone. GAP: Glyceraldehyde-3-phosphate; DHAP: Dihydroxyacetone phosphate; FBP: Fructose-1,6-bisphosphate; F₆P: Fructose-6-phosphate; Xu₅P: Xylulose-5-phosphate; and Pi: Phosphate.



Metabolism Although several promoters, either constitutive or inducible, are available for *Pichia pastoris*, the two forms of alcohol oxidases (AOX1 and AOX2) has been the favoured choice in most studies and applications. Alcohol oxidase I promoter (P_{AOX1}) being the predominantly one expressed in *Pichia* cell. Its extraordinary strength and tight regulation by carbon sources, and availability in a commercial kit are the main reasons for its intensive use. Being tightly repressed by carbon sources such as glucose, glycerol or ethanol, transcription from the *AOX1* gene is highly induced upon shifting the cells to methanol as sole carbon source, whilst upon carbon starvation very low derepression of transcription occurs.^[218,219]

***Pichia pastoris* phenotypes** Three utilising-methanol phenotypes of *P. pastoris* (Mut^+ , Mut^S and Mut^-) are currently available for the production of heterologous proteins.^[220,221] The most used strain is the methanol utilisation plus phenotype, Mut^+ that grows on methanol at the wild-type rate. The expression cassette is inserted into HIS4 (histidinol dehydrogenase) locus, and both *AOX1* and *AOX2* genes are intact and active. The other two types of host strains, the methanol utilisation slow (Mut^S) and minus (Mut^-) phenotypes have respectively deletions in one or both *AOX* genes. In Mut^S , the expression cassette is inserted within the *AOX1* locus, thus the *AOX1* gene is knocked out and the cells must rely on the alternative and weaker *AOX2* gene, which slows down the methanol-uptake rate; in Mut^- , both *AOX1* and *AOX2* genes are deleted, thus lowering the growth rate.^[222] In general, Mut^+ strains, used in this study, are characterised by a higher growth rate and productivity when compared to Mut^S and Mut^- strains, although recent reports have shown that a strain with Mut^S phenotype was superior over a strain with Mut^+ phenotype, in both productivity and efficiency when expressing recombinant horseradish peroxidase *C1A*.^[223] To this date, for any given recombinant protein, it remains unclear which strain would perform better, and the information available in literature is conflicting in this respect.^[224]

Construct Upstream of the multiple cloning site (MCS), the vector contains the 5'*AOX1* promoter sequence followed by the secretion signal sequence, α -mating factor (α -MF), from *Saccharomyces cerevisiae*. By inserting the mutagenic vector in between the 5'*AOX1* promoter region and the *AOX1* locus in *P. pastoris* genome, the gene encoding the desired protein is transcribed at the same time as *AOX1* induction by methanol, and as a fusion product to the secretion signal. This secretion signal directs the expressed protein of interest out of the cells, and into the culture media, which greatly facilitates downstream processing, namely purification. Following the *MCS* is the native *AOX1* transcription termination and polyadenylation signal (TT). The vector also contains the *Sh ble* gene from *Streptoalloteichus hindustanus*, which confers resistance to the antibiotic Zeocin useful in both bacteria and yeast [Invitrogen, *Pichia Expression Kit* - Version M]. Zeocin resistance has been correlated with integrated vector copy number, where selection on increased antibiotic levels often indicates

the presence of multi-copy integrants.^[225]

Several procedures for molecular genetic manipulation of *Pichia pastoris* are similar to those of *Saccharomyces cerevisiae*, one of the most widely studied eukaryotic organism. For example, transformation by electroporation of a linearised expression plasmid, which stimulates homologous recombination between the chromosomal sequence and the plasmid, leads to the integration of the vector into the genome at a specific locus. This procedure is simplified by the significant set of *Pichia* strains and vectors commercially available.

Post-translational modifications *Pichia pastoris*, unlike bacterial expression systems, performs a variety of post-translational modifications typically associated with higher eukaryotes, including disulphide bond formation, processing of signal sequences, correct folding, and *O*- and *N*-glycosylation. Such modifications may critically influence the function of an expressed protein. Secreted proteins are likely to be glycosylated. The most common and best studied is *N*-linked glycosylation, where oligosaccharides are uniquely added to asparagine found in Asn-X-Ser/Thr recognition sequences in proteins, where X can be any amino acid except proline. Another type of glycosylation is *O*-linked glycosylation, which involves adding *O*-oligosaccharide chains, composed of solely mannose residues, to the hydroxyl group of serine and threonine of the secreted protein.^[226,227] Hyperglycosylation leads to heterogeneity of a recombinant protein product in both carbohydrate composition and molecular weight, and is often undesirable since it may complicate for example the purification of the protein.^[228]

Biotin requirements Biotin plays an essential role as cofactor for biotin-dependent carboxylases, which are involved in carboxylation and decarboxylation reactions of important metabolic pathways *e.g.* fatty acid biosynthesis, amino acid metabolism and gluconeogenesis.^[229] Animals rely on their dietary uptake of biotin, whereas most microorganisms and plants have the ability to synthesise *de novo* the vitamin. In contrast to the vast majority of microorganisms, some yeast species exhibit biotin auxotrophy, *i.e.* are unable to synthesise biotin themselves. One example is *Pichia pastoris*, which requires addition of high dosage of biotin during cultivation (4×10^4 g/L).^[230] Jungo *et al.* has extensively studied the biotin requirements for the production of recombinant avidin, the egg-white counterpart of streptavidin, in *Pichia pastoris*. Substitution of biotin by two unrelated compounds, aspartic acid and oleic acid, did not completely replace biotin, as wash-out occurred after six liquid residence times and overall protein productivity was lowered. However, low amounts of biotin (20 μ g/L) did have a positive effect on the stable chemostat cultures, as biotin-free avidin was obtained.^[231]

Proteolytic degradation In general, recombinant proteins expressed extracellularly in *Pichia pastoris* can potentially be proteolytically degraded in the culture medium by pro-

teases (cell-bound, intracellular protease from lysed cells and/or extracellular proteases). Due to proteolysis, several problems can be foreseen: (i) reduction or loss of biological activity by truncation of the product; (ii) low product yield due to degradation of the product; and (iii) contamination of the product by degradation intermediates in downstream processing due to their physicochemical similarity and/or affinity characteristics.^[232] Several strategies based on the modification at the cultivation, strains (protease-deficient strains, *e.g.* SMD 1163 and SMD 1168), and recombinant protein level can be employed to minimise proteolysis in *Pichia*.^[233,234] For example, in an attempt to minimise the proteolysis of an urokinase-type plasminogen activator, 0.01% (v/v) Triton X-100 was added to a feeding medium, which partially reduced the proteolysis and increased the secretion level.^[235] *P. pastoris* was incubated at temperatures as low as 15 °C, leading to reduced protease levels and greatly enhanced periods of scFv production,^[236] although incubation temperatures between 23 and 30 °C are recommended by Invitrogen. Maintaining the culture at low pH (3.0) during the methanol induction phase has been reported to be effective in protecting the product proteolysis of an insulin-like growth factor (IGF-1).^[237]

Advantages & disadvantages The *Pichia* system presents several advantages when compared to other expression systems. *Pichia pastoris* (i) is easy to genetically manipulate and is genetically stable; (ii) can grow on solely methanol as carbon and energy source, thus the fermentation processes are simple to scale-up without loss of yield and are inexpensive; (iii) reaches high cell densities (up to 130 g/L dry cell weight, dcw)^[238] on a simple basal salts medium; (iv) is capable of post-translation modifications (disulphide bonds, *N*- or *O*-glycosylations and proteolytic processing); and (v) produces high yield of proteins, either intra- or extracellularly with, in the case of secretion, high recombinant protein levels into an almost protein-free medium. Many examples of yields in the grams per litre range have been reported for several target proteins, including full-length antibodies.^[239–243] However, heterologous secretion is not guaranteed to work as it is more demanding than intracellular expression. Like any other system, *Pichia* is no panacea and has drawbacks. Although many remain unpublished, cases of failure or significantly low expression of recombinant proteins continue to accumulate.^[244–246] The most common issue encountered is the proteolysis of secreted proteins, though a number of ways to overcome this impediment are now available.^[247,248] Another common problem is related to truncated mRNA from yeast transcriptional terminators, which result in some genes not giving any detectable protein.^[249,250] The main disadvantages of *Pichia* are (i) *Pichia* cells have to be freshly prepared before use; (ii) the tedious and labour intensive screening process of different clones; (iii) the variation in expression yields between those clones; and (iv) the long induction phase, usually of several days (up to 12 days).

1.5 Scope of this thesis

Artificial metalloenzymes can be defined as the fusion of chemo- and enzymatic catalysts by insertion of non-specific achiral catalytic moieties into the chiral environment of a protein cavity, thereby broadening the scope of both fields. Presently, the field of artificial metalloenzymes is at an exciting stage, with interesting features already demonstrated^[251] and a growing reaction repertoire, but has not yet reached maturity.^[40,52] To exploit the full potential of catalytic reactions using hybrid catalysts is far from trivial and as the ability to screen for novel properties remains limited. Albeit computational simulations are advanced, directed/ designed evolution of such selective catalysts remains challenging, hence their development relies to a great extent on trial-and-error.^[13] Hitherto, screening and evaluation of ligands for activity or of proteins for selectivity still remains an elaborate process, as all library members have to be evaluated individually. High-throughput analysis to simultaneously evaluate large hybrid catalyst libraries within a short time-frame has therefore become increasingly important.

The key focuses of this research were the development of (i) novel purification strategies to simplify the challenge of producing functional protein (streptavidin), in sufficient quantity and appropriate purity for high-throughput screening of organometallic moieties; (ii) a reiterative approach to protein expression: re-design of streptavidin construct, moving to a more complex expression system, *Pichia pastoris*, to increase the production of streptavidin and simplify the purification process; and (iii) the design and synthesis of a novel biomolecular scaffold, human carbonic anhydrase II, for the creation of a new artificial metalloenzyme.

This thesis addresses the issues and feasibility of purification and expression processes and, highlights bottlenecks that arose using these technologies, key learning's and where to go from now.

Quoting Frances H. Arnold “*it is apparent that many solutions exist for any given problem, and there are often many paths that lead uphill, one step at a time.*”^[252]

1.6 References

- [1] Faber, K. *Biotransformations in organic chemistry - A textbook*; Springer Verlag Berlin Heidelberg: Berlin, Germany, 6th Edition ed.; 2011.
- [2] Darrow, J. J. *Stanford Technology Law Review* **2007**, *2*, 1–19.
- [3] Sweet, M. J. *Berkeley Technology Law Journal* **2007**, *24*, 129–147.
- [4] Islam, M. R.; Mahdi, J. G.; Bowen, I. D. *Drug Safety* **1997**, *17*, 149–165.
- [5] Landoni, M. F.; Soraci, A. L.; Delatour, P.; Lees, P. *Journal of Veterinary Pharmacology and Therapeutics* **1997**, *20*, 1–16.
- [6] Rosati, F.; Roelfes, G. *ChemCatChem* **2010**, *2*, 916–927.
- [7] Lu, Y. *Inorganic Chemistry* **2006**, *45*, 9930–9940.
- [8] Steinreiber, J.; Ward, T. R. *Bio-inspired catalysts*; volume 25 Springer Berlin Heidelberg: Berlin, Germany, 2009.
- [9] Pordea, A.; Ward, T. R. *Chemical Communications* **2008**, *36*, 4239–4249.
- [10] Ke, Z.; Abe, S.; Ueno, T.; Morokuma, K. *Journal of the American Chemical Society* **2012**, *134*, 15418–15429.
- [11] Fischer, E. *Berichte der deutschen chemischen Gesellschaft* **1894**, *27*, 2985–2993.
- [12] Koshland Jr., D. E. *Proceedings of the National Academy of Sciences of the United States of America* **1958**, *44*, 98–104.
- [13] Meeuwissen, J.; Reek, J. N. H. *Nature Chemistry* **2010**, *2*, 615–621.
- [14] Podtetenieff, J.; Taglieber, A.; Bill, E.; Reijerse, E. J.; Reetz, M. T. *Angewandte Chemie International Edition* **2010**, *49*, 5151–5155.
- [15] Letondor, C.; Ward, T. R. *ChemBioChem* **2006**, *7*, 1845–1842.
- [16] Oltra, N. S.; Roelfes, G. *Chemical Communications* **2008**, 6039–6041.
- [17] Megens, R. P.; Roelfes, G. *Organic & Biomolecular Chemistry* **2010**, *8*, 1387–1393.
- [18] Levine, H. L.; Nakagawa, Y.; Kaiser, E. T. *Biochemical and Biophysical Research Communications* **1977**, *76*, 64–70.
- [19] Wilson, M. E.; Whitesides, G. M. *Journal of the American Chemical Society* **1978**, *100*, 306–307.
- [20] Deuss, P. J.; den Heeten, R.; Laan, W.; Kamer, P. C. J. *Chemistry - A European Journal* **2011**, *17*, 4680–4698.
- [21] Ringenberg, M. R.; Ward, T. R. *Chemical Communications* **2011**, *47*, 8470–8476.

- [22] Häring, D.; Distefano, M. D. *Bioconjugate Chemistry* **2001**, *12*, 385–390.
- [23] Reetz, M. T.; Peyralans, J.-P.; Maichele, A.; Maywald, M. *Chemical Communications* **2006**, 4318–4320.
- [24] Kazlauskas, R. J.; Bornscheuer, U. T. *Nature Chemical Biology* **2009**, *5*, 526–529.
- [25] Lin, C.-C.; Lin, C.-W.; Chan, A. S. C. *Tetrahedron: Asymmetry* **1999**, *10*, 1887–1893.
- [26] Davies, R. R.; Distefano, M. D. *Journal of the American Chemical Society* **1997**, *119*, 11643–11652.
- [27] Yamaguchi, H.; Hirano, T.; Kiminami, H.; Taura, D.; Harada, A. *Organic & Biomolecular Chemistry* **2006**, *4*, 3571–3573.
- [28] de Vries, J. G.; Lefort, L. *Chemistry - A European Journal* **2006**, *12*, 4711–4734.
- [29] Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Maywald, M.; Maiwald, P.; Peyralans, J.-P.; Maichele, A.; Fu, Y.; Jiao, N.; Hollmann, F.; Mondière, R.; Taglieber, A. *Tetrahedron* **2007**, *63*, 6404–6414.
- [30] Thomas, C. M.; Ward, T. R. *Applied Organometallic Chemistry* **2005**, *19*, 35–39.
- [31] Creus, M.; Pordea, A.; Rossel, T.; Sardo, A.; Letondor, C.; Ivanova, A.; Letrong, I.; Stenkamp, R. E.; Ward, T. R. *Angewandte Chemie International Edition* **2008**, *47*, 1400–1404.
- [32] Collot, J.; Gradinaru, J.; Humbert, N.; Skander, M.; Zocchi, A.; Ward, T. R. *Journal of the American Chemical Society* **2003**, *125*, 9030–9031.
- [33] Collot, J.; Humbert, N.; Skander, M.; Klein, G.; Ward, T. R. *Journal of Organometallic Chemistry* **2004**, *689*, 4868–4871.
- [34] Skander, M.; Humbert, N.; Collot, J.; Gradinaru, J.; Klein, G.; Loosli, A.; Sauser, J.; Zocchi, A.; Gilardoni, F.; Ward, T. R. *Journal of the American Chemical Society* **2004**, *126*, 14411–14418.
- [35] Skander, M.; Malan, C.; Ivanova, A.; Ward, T. R. *Chemical Communications* **2005**, *14*, 4815–4817.
- [36] Klein, G.; Humbert, N.; Gradinaru, J.; Ivanova, A.; Gilardoni, F.; Rusbandi, U. E.; Ward, T. R. *Angewandte Chemie International Edition* **2005**, *44*, 7764–7767.
- [37] Ohashi, M.; Koshiyama, T.; Ueno, T.; Yanase, M.; Fujii, H.; Watanabe, Y. *Angewandte Chemie International Edition* **2003**, *115*, 1035–1038.
- [38] Ohashi, M.; Koshiyama, T.; Ueno, T.; Yanase, M.; Fujii, H.; Watanabe, Y. *Angewandte Chemie International Edition* **2003**, *42*, 1005–1008.
- [39] Pierron, J.; Malan, C.; Creus, M.; Gradinaru, J.; Hafner, I.; Ivanova, A.; Sardo, A.; Ward, T. R. *Angewandte Chemie International Edition* **2008**, *47*, 701–705.

- [40] Dürrenberger, M.; Heinisch, T.; Wilson, Y. M.; Rossel, T.; Nogueira, E.; Knörr, L.; Mutschler, A.; Kersten, K.; Zimbron, M. J.; Pierron, J.; Schirmer, T.; Ward, T. R. *Angewandte Chemie International Edition* **2011**, *50*, 3026–3029.
- [41] Pordea, A.; Creus, M.; Panek, J. J.; Duboc, C.; Mathis, D.; Novič, M.; Ward, T. R. *Journal of the American Chemical Society* **2008**, *130*, 8085–8088.
- [42] Reetz, M. T. *Tetrahedron* **2002**, *58*, 6595–6602.
- [43] Okrasa, K.; Kazlauskas, R. J. *Chemistry - A European Journal* **2006**, *12*, 1587–1596.
- [44] Kokubo, T.; Sugimoto, T.; Uchida, T.; Tanimoto, S.; Okano, M. *Journal of the Chemical Society, Chemical Communications* **1983**, 769–770.
- [45] Fernández-Gacio, A.; Codina, A.; Fastrez, J.; Riant, O.; Soumillion, P. *ChemBioChem* **2006**, *7*, 1013–1016.
- [46] Reetz, M. T.; Jiao, N. *Angewandte Chemie International Edition* **2006**, *45*, 2416–2419.
- [47] Roelfes, G.; Feringa, B. L. *Angewandte Chemie International Edition* **2005**, *44*, 3230–3232.
- [48] Roelfes, G.; Boersma, A. J.; Feringa, B. L. *Chemical Communications* **2006**, *6*, 635–637.
- [49] Roy, R. S. *Protein Engineering* **1997**, *10*, 691–698.
- [50] Roelfes, G. *Molecular BioSystems* **2007**, *3*, 126–135.
- [51] Dong, Z.; Wang, Y.; Yin, Y.; Liu, J. *Current Opinion in Colloid & Interface Science* **2011**, *16*, 451–458.
- [52] Hyster, T. K.; Knörr, L.; Ward, T. R.; Rovis, T. *Science* **2012**, *338*, 500–503.
- [53] Wilson, Y. M.; Dürrenberger, M.; Ward, T. R. *Organometallic chemistry in protein scaffolds*; volume 3 of *Protein Engineering Handbook* Wiley-VCH Verlag GmbH & Co.: Weinheim, Germany, 1st ed.; 2012.
- [54] Monnard, F. W.; Heinisch, T.; Nogueira, E. S.; Schirmer, T.; Ward, T. R. *Chemical Communications* **2011**, *47*, 8238–8240.
- [55] Carey, J. R.; Ma, S. K.; Pfister, T. D.; Garner, D. K.; Kim, H. K.; Abramite, J. A.; Wang, Z.; Guo, Z.; Lu, Y. *Journal of the American Chemical Society* **2004**, *126*, 10812–10813.
- [56] Rusbandi, U. E.; Lo, C.; Skander, M.; Ivanova, A.; Creus, M.; Humbert, N.; Ward, T. R. *Advanced Synthesis & Catalysis* **2007**, *349*, 1923–1930.
- [57] Ueno, T.; Suzuki, M.; Goto, T.; Matsumoto, T.; Nagayama, K.; Watanabe, Y. *Angewandte Chemie International Edition* **2007**, *43*, 2527–2530.
- [58] Qi, D.; Tann, C. M.; Haring, D.; Distefano, M. D. *Chemical Reviews* **2001**, *101*, 3081–3111.

- [59] Boersma, A. J.; Feringa, B. L.; Roelfes, G. *Organic Letters* **2007**, *9*, 3647–3650.
- [60] Boersma, A. J.; Klijn, J. E.; Feringa, B. L.; Roelfes, G. *Journal of the American Chemical Society* **2008**, *130*, 11783–11790.
- [61] Bertucci, C.; Botteghi, C.; Giunta, D.; Marchetti, M.; Paganelli, S. *Advanced Synthesis & Catalysis* **2002**, *344*, 556–562.
- [62] Coquière, D.; Bos, J.; Beld, J.; Roelfes, G. *Angewandte Chemie International Edition* **2009**, *48*, 5159–5162.
- [63] Satake, Y.; Abe, S.; Okazaki, S.; Ban, N.; Hikage, T.; Ueno, T.; Nakajima, H.; Suzuki, A.; Yamane, T.; Nishiyama, H.; Watanabe, Y. *Organometallics* **2007**, *26*, 4904–4908.
- [64] Panella, L.; Broos, J.; Jin, J.; Fraaije, M. W.; Janssen, D. B.; Jeronimus-Stratingh, M.; Feringa, B. L.; Minnaard, A. J.; de Vries, J. G. *Chemical Communications* **2005**, 5656–5658.
- [65] Talbi, B.; Haquette, P.; Martel, A.; de Montigny, F.; Fosse, C.; Cordier, S.; Roisnel, T.; Jaouen, G.; Salmain, M. *Dalton Transactions* **2010**, *39*, 5605–5607.
- [66] Haquette, P.; Talbi, B.; Barilleau, L.; Madern, N.; Fosse, C.; Salmain, M. *Organic & Biomolecular Chemistry* **2011**, *9*, 5720–5728.
- [67] Laan, W.; Muñoz, B. K.; den Heeten, R.; Kamer, P. C. J. *ChemBioChem* **2010**, *11*, 1236–1239.
- [68] den Heeten, R.; Muñoz, B. K.; Popa, G.; Laan, W.; Kamer, P. C. J. *Dalton Transactions* **2010**, *39*, 8477–8483.
- [69] Jing, Q.; Okrasa, K.; Kazlauskas, R. J. *Chemistry - A European Journal* **2009**, *15*, 1370–1376.
- [70] Boersma, A. J.; Coquière, D.; Geerdink, D.; Rosati, F.; Feringa, B. L.; Roelfes, G. *Nature Chemistry* **2010**, *2*, 991–995.
- [71] Rosenberg, R. C.; Root, C. A.; Bernstein, P. K.; Gray, H. B. *Journal of the American Chemical Society* **1975**, *97*, 2091–2096.
- [72] Wu, Z. P.; Hilvert, D. *Journal of the American Chemical Society* **1989**, *111*, 4513–4514.
- [73] Letondor, C.; Humbert, N.; Ward, T. R. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 4683–4687.
- [74] Zimbron, M. J.; Heinisch, T.; Hamels, D.; Nogueira, E. S.; Schirmer, T.; Ward, T. R. *Journal of the American Chemical Society* **2013**, *135*, 5384–5388.
- [75] Noyori, R.; Kitamura, M.; Ohkuma, T. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 5356–5362.

- [76] Köhler, V.; Wilson, Y. M.; Lo, C.; Sardo, A.; Ward, T. R. *Current Opinion in Biotechnology* **2010**, *21*, 744–752.
- [77] Aboul-Enein, H. Y.; Wainer, I. W. *The impact of stereochemistry on drug development and use*; Wiley: New York, USA, 1997.
- [78] Zassinovich, G.; Mestroni, G.; Gladiali, S. *Chemical Reviews* **1992**, *92*, 1051–1069.
- [79] Noyori, R.; Hashiguchi, S. *Accounts of Chemical Research* **1997**, *30*, 470–473.
- [80] Gladiali, S.; Alberico, E. *Chemical Society Reviews* **2006**, *35*, 226–236.
- [81] Descotes, G.; Sinou, D. *Tetrahedron Letters* **1976**, *17*, 4083–4086.
- [82] Ohkubo, K.; Setoguchi, M.; Yoshinaga, K. *Inorganic and Nuclear Chemistry Letters* **1979**, *15*, 235–238.
- [83] Grigg, R.; Mitchell, T. R. B.; Tongpenyai, N. *Synthesis* **1981**, *6*, 442–444.
- [84] Shvo, Y.; Czarkie, D.; Rahamim, Y.; Chodosh, D. F. *Journal of the American Chemical Society* **1986**, *108*, 7400–7402.
- [85] Hashiguchi, S. I.; Fujii, A.; Takehara, J.; Ikariya, T.; Noyori, R. *Journal of the American Chemical Society* **1995**, *117*, 7562–7563.
- [86] Wang, C.; Wu, X. F.; Xiao, J. *Chemistry - An Asian Journal* **2008**, *3*, 1750–1770.
- [87] Kayaki, Y.; Ikeda, H.; Tsurumaki, J.-I.; Shimizu, I.; Yamamoto, A. *Bulletin of the Chemical Society of Japan* **2008**, *81*, 1053–1061.
- [88] Martins, J. E. D.; Clarkson, G. J.; Wills, M. *Organic Letters* **2009**, *11*, 847–850.
- [89] Murata, K.; Ikariya, T.; Noyori, R. *The Journal of Organic Chemistry* **1999**, *64*, 2186–2187.
- [90] Wu, J.; Wang, F.; Ma, Y.; Cui, X.; Cun, L.; Zhu, J.; Deng, J.; Yu, B. *Chemical Communications* **2006**, 1766–1768.
- [91] Li, L.; Wang, F.; Liao, J.; Zhang, H.; Lian, C.; Zhu, J.; Deng, J. *Green Chemistry* **2007**, *9*, 23–25.
- [92] Wang, C.; Li, C.; Wu, X.; Pettman, A.; Xiao, J. *Angewandte Chemie International Edition* **2009**, *121*, 6646–6650.
- [93] Wu, X.; Li, X.; Hems, W.; King, F.; Xiao, J. *Organic & Biomolecular Chemistry* **2004**, *2*, 1818–1821.
- [94] Kaufman, T. S. *Tetrahedron: Asymmetry* **2004**, *12*, 1203–1237.
- [95] Kadyrov, R.; Riermeier, T. *Angewandte Chemie International Edition* **2003**, *115*, 5630–5632.

- [96] Breuer, M.; Ditrich, K.; Habicher, T.; Hauer, B.; Kessler, M.; Sturmer, R.; Zelinski, T. *Angewandte Chemie International Edition* **2004**, *43*, 788–824.
- [97] Chapman, P. *Amino acids market to hit \$12.8 billion by 2017*; Companies and Markets - Chemicals Vertical Edge Limited: London, United Kingdom, 2012.
- [98] Ages, D. J. *Amino acids, peptides and proteins in organic chemistry: Origins and synthesis of amino acid*; volume 1 Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2009.
- [99] Palmer, D. A.; Van Eldik, R. *Chemical Reviews* **1983**, *83*, 651–731.
- [100] Cooper, A. J. L.; Ginos, J. Z.; Meister, A. *Chemical Reviews* **1983**, *83*, 321–358.
- [101] Müller, U.; Hübner, S. *Advances in Biochemical Engineering/Biotechnology* **2003**, *79*, 137–170.
- [102] Yamada, S.; Nabe, K.; Izuo, N.; Nakamichi, K.; Chibata, I. *Applied and Environmental Microbiology* **1981**, *42*, 773–778.
- [103] Leuchtenberger, W. *Biotechnology Products of Primary Metabolism*; Wiley-VCH Verlag GmbH: Weinheim, Germany, 2nd Edition ed.; 2008.
- [104] Lin, Y.; Yamamoto, A. *Handbook of organopalladium chemistry for organic synthesis*; John Wiley & Sons: Hoboken, USA, 2002.
- [105] Hwang, B.-Y.; Cho, B.-K.; yun, H.; Koteswar, K.; Kim, B.-G. *Journal of Molecular Catalysis B* **2005**, *37*, 47–55.
- [106] Gröger, H.; May, O.; Werner, H.; Menzel, A.; Altenbuchner, J. *Organic Process Research & Development* **2006**, *10*, 666–669.
- [107] Ogo, S.; Uehara, K.; Abura, T.; Fukuzumi, S. *Journal of the American Chemical Society* **2004**, *126*, 3020–3021.
- [108] Green, N. M. *Advances in Protein Chemistry* **1975**, *29*, 65–133.
- [109] Wilchek, M.; Bayer, E. A. *Avidin-biotin technology*; volume 184 of *Methods in Enzymology* Academic Press: New York, USA, 1990.
- [110] Lesch, H. P.; Kaikkonen, M. U.; Pikkarainen, J. T.; Ylä-Herttuala, S. *Expert Opinion on Drug Delivery* **2010**, *7*, 551–564.
- [111] Laitinen, O. H.; Nordlund, H. R.; Hytönen, V. P.; Kulomaa, M. S. *Trends in Biotechnology* **2007**, *25*, 269–277.
- [112] Hendrickson, W. A.; Pähler, A.; Smith, J. L.; Satow, Y.; Merritt, E. A.; Phizackerley, R. P. *Proceedings of the National Academy of Sciences of the United States of America* **1989**, *5*, 2190–2194.
- [113] Green, N. M. *Biochemical Journal* **1963**, *89*, 599–609.

- [114] Sano, T.; Cantor, C. R. *Proceedings of the National Academy of Sciences of the United States of America* **1990**, *87*, 142–146.
- [115] Livnah, O.; Bayer, E. A.; Wilchek, M.; Sussman, J. L. *Proceedings of the National Academy of Sciences of the United States of America* **1990**, *90*, 5076–5080.
- [116] Chilkoti, A.; Tan, P. H.; Stayton, P. S. *Proceedings of the National Academy of Sciences of the United States of America* **1995**, *92*, 1754–1758.
- [117] Freitag, S.; Le Trong, I.; Chilkoti, A.; Klumb, L. A.; Stayton, P. S.; Stenkamp, R. E. *Journal of Molecular Biology* **1998**, *279*, 211–221.
- [118] Ballester, P.; Vidal-Ferran, A. *Supramolecular catalysis*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2008.
- [119] Siegel, J.; Zanghellini, A.; Lovick, H. M.; Kiss, G.; Lambert, A.; St. Clair, J. L.; Gallaher, J. L.; Hilvert, D.; Gelb, M.; Stoddart, B.; Houk, K.; Michael, F.; Baker, D. *Science* **2010**, *329*, 309–313.
- [120] Loosli, A.; Rusbandi, U. E.; Gradinaru, J.; Bernauer, K.; Schlaepfer, C. W.; Meyer, M.; Mazurek, S.; Novič, M.; Ward, T. R. *Inorganic Chemistry* **2006**, *45*, 660–668.
- [121] Gruber, H. J.; Kada, G.; Marek, K.; Kaiser, K. *Biochimica et Biophysica Acta* **1998**, *1381*, 203–212.
- [122] Humbert, N.; Schürmann, O.; Zocchi, A.; Neuhaus, J.-M.; Ward, T. R. *Methods in Molecular Biology* **2008**, *418*, 101–110.
- [123] Smith, K. S.; Jakubzick, C.; Whittam, T. S.; Ferry, J. G. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 15184–15189.
- [124] Tripp, B. C.; Smith, K.; Ferry, J. G. *The Journal of Biological Chemistry* **2001**, *276*, 48615–48618.
- [125] Xu, Y.; Feng, L.; Jeffrey, P. D.; Shi, Y.; Morel, F. M. *Nature* **2008**, *452*, 56–61.
- [126] Meldrum, N. U.; Roughton, F. J. W. *Journal of Physiology* **1933**, *80*, 113–142.
- [127] Blankenship, L. C.; Veitch, F. P. *Nature* **1963**, *197*, 76–77.
- [128] Rowlett, R. S. *Biochimica et Biophysica Acta* **2010**, *1804*, 362–373.
- [129] Supuran, C. T. *Frontiers in Pharmacology* **2011**, *2*, 1–6.
- [130] Ludwig, M. *Plant Cell and Environment* **2012**, *35*, 22–37.
- [131] De Simone, G.; Supuran, C. T. *Journal of Inorganic Biochemistry* **2012**, *111*, 117–129.
- [132] Montgomery, J. C.; Venta, P. J.; Tashian, R. E.; Hewett-Emmett, D. *Nucleic Acids Research* **1987**, *15*, 4687.

- [133] Boriack-Sjodin, P. A.; Zeitlin, S.; Chen, H. H.; Crenshaw, L.; Gross, S.; Dantanarayana, A.; Delgado, P.; May, J. A.; Dean, T.; Christianson, D. W. *Protein Science* **1998**, *7*, 2483–2489.
- [134] Baldwin, J. J.; Ponticello, G. S.; Anderson, P. S.; Christy, M. E.; Murcko, M. A.; Randall, W. C.; Schwam, H.; Sugrue, M. F.; Springer, J. P.; Gautheron, P. *Journal of Medicinal Chemistry* **1989**, *32*, 2510–2513.
- [135] Winum, J.-Y.; Scozzafava, A.; Montero, J. L.; Supuran, C. T. *Current Pharmaceutical Design* **2009**, *14*, 615–621.
- [136] Chegwidan, W. R.; Carter, N. D. *The Carbonic Anhydrases: New Horizons*; volume 90 of *Experientia Supplementum* Birkhäuser Basel: Basel, Switzerland, 2000.
- [137] Khalifah, R. G. *The Journal of Biological Chemistry* **1971**, *246*, 2561–2573.
- [138] Lindskog, S. *Zinc Enzymes*; John Wiley & Sons: New York, USA, 1983.
- [139] Lindskog, S. *Pharmacological Therapeutics* **1997**, *74*, 1–20.
- [140] Alexander, R. S.; Nair, S. K.; Christianson, D. W. *Biochemistry* **1991**, *30*, 11064–11072.
- [141] Silverman, D. N.; Lindskog, S. *Accounts of Chemical Research* **1988**, *21*, 30–36.
- [142] Liljas, A.; Kannan, K. K.; Bergstén, P. C.; Waara, I.; Fridborg, K.; Strandberg, B.; Carlbohm, U.; Järup, L.; Lövgren, S.; Petef, M. *Nature New Biology* **1972**, *235*, 131–137.
- [143] Sly, W. S.; Hu, P. Y. *Annual Review of Biochemistry* **1995**, *4*, 375–401.
- [144] Briesewitz, R.; Ray, G. T.; Wandless, T. J.; Crabtree, G. R. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 1953–1958.
- [145] Nair, S. K.; Calderone, T. L.; Christianson, D. W.; Fierke, C. *The Journal of Biological Chemistry* **1991**, *266*, 17320–17325.
- [146] Avvaru, B. S.; Kim, C. U.; Sippel, K. H.; Gruner, S. M.; Agbandje-McKenna, M.; Silverman, D. N.; McKenna, R. *Biochemistry* **2010**, *49*, 249–251.
- [147] Lindskog, S.; Coleman, J. E. *Proceedings of the National Academy of Sciences of the United States of America* **1973**, *70*, 2505–2508.
- [148] Thorslund, A.; Lindskog, S. *European Journal of Biochemistry* **1967**, *3*, 117–123.
- [149] Henderson, L. E.; Henriksson, D.; Nyman, P. O. *The Journal of Biological Chemistry* **1976**, *251*, 5457–5463.
- [150] Hunt, J. B.; Rhee, M.-J.; Storm, C. B. *Analytical Biochemistry* **1977**, *79*, 614–617.
- [151] Håkansson, K.; Wehnert, A.; Liljas, A. *Acta Crystallographica Section D - Biological Crystallography* **1994**, *D50*, 93–100.

- [152] Elder, I.; Han, S.; Tu, C.; Steele, H.; Laipis, P. J.; Viola, R. E.; Siverman, D. N. *Archives of Biochemistry and Biophysics* **2005**, *421*, 283–289.
- [153] Jing, Q.; Kazlauskas, R. J. *ChemCatChem* **2010**, *2*, 953–957.
- [154] Supuran, C. T.; Scozzafava, A. *Biorganic & Medicinal Chemistry* **2007**, *15*, 4336–4350.
- [155] Gitto, R.; Agnello, S.; Ferro, S.; De Luca, L.; Vullo, D.; Brynda, J.; Mader, P.; Supuran, C. T.; Chimirri, A. *Journal of Medicinal Chemistry* **2010**, *53*, 2401–2408.
- [156] Di Fiore, A.; Maresca, A.; Alterio, V.; Supuran, C. T.; De Simone, G. *Chemical Communications* **2011**, *47*, 11636–11638.
- [157] Vidgren, J.; Liljas, A.; Walker, N. P. C. *International Journal of Biological Macromolecules* **1990**, *12*, 342–344.
- [158] Prugh, J. D.; Hartman, G. D.; Mallorga, P. J.; McKeever, B. M.; Michelson, S. R.; Murcko, M. A.; Schwam, H.; Smith, R. L.; Sondey, J. M.; Springer, J. P. *Journal of Medicinal Chemistry* **1991**, *34*, 1805–1818.
- [159] Jain, A.; Whitesides, G. M.; Alexander, R. S.; Christianson, D. W. *Journal of Medicinal Chemistry* **1994**, *37*, 2100–2105.
- [160] Bunn, A. M. C.; Alexander, R. S.; Christianson, D. W. *Journal of the American Chemical Society* **1994**, *116*, 5063–5068.
- [161] Håkansson, K.; Liljas, A. *FEBS Letters* **1994**, *350*, 319–322.
- [162] Smith, G. M.; Alexander, R. S.; Christianson, D. W.; McKeever, B. M.; Ponticello, G. S.; Springer, J. P.; Randall, W. C.; Baldwin, J. J.; Habecker, C. N. *Protein Science* **1994**, *3*, 118–125.
- [163] Stams, T.; Chen, Y.; Boriack-Sjodin, P. A.; Hurt, J. D.; Liao, J.; May, J. A.; Dean, T.; Laipis, P. J.; Silverman, D. N.; Christianson, D. W. *Protein Science* **1998**, *7*, 556–563.
- [164] Grzybowski, B. A.; Ishchenko, A. V.; Kim, C.-Y.; Topalov, G.; Chapman, R.; Christianson, D. W.; Whitesides, G. M.; Shakhnovich, E. I. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 1270–1273.
- [165] Ward, T. R. *Angewandte Chemie International Edition* **2008**, *47*, 7802–7803.
- [166] Gould, S. M.; Tawfik, D. S. *Biochemistry* **2005**, *44*, 5444–5552.
- [167] Krishnamurthy, V. M.; Kaufman, G. K.; Urbach, A. R.; Gitlin, I.; Gudiksen, K. L.; Weibel, D. B.; Whitesides, G. M. *Chemical Reviews* **2008**, *108*, 946–1051.
- [168] Venters, R. A.; Huang, C. C.; Farmer, B. T.; Trolard, R.; Spicer, L. D.; Fierke, C. A. *Journal of Biomolecular NMR* **1995**, *5*, 339–344.
- [169] Venters, R. A.; Farmer, B. T.; Fierke, C. A.; Spicer, L. D. *Journal of Molecular Biology* **1996**, *264*, 1101–1116.

- [170] Häussinger, D.; Huang, J.-R.; Grzesiek, S. *Journal of the American Chemical Society* **2009**, *131*, 14761–14767.
- [171] Creus, M.; Ward, T. R. *Organic & Biomolecular Chemistry* **2007**, *5*, 1835–1844.
- [172] Hida, K.; Hanes, J.; Ostermeier, M. *Advanced Drug Delivery Reviews* **2007**, *59*, 1562–1578.
- [173] Jaeger, K. E.; Eggert, T. *Current Opinion in Biotechnology* **2004**, *15*, 305–313.
- [174] Arnold, F. H.; Wintrode, P. L.; Miyazaki, K.; Gershenson, A. *Trends in Biochemical Sciences* **2001**, *26*, 100–107.
- [175] Reetz, M. T. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 5716–5722.
- [176] Reetz, M. T.; Wu, S. *Journal of the American Chemical Society* **2009**, *131*, 15424–15432.
- [177] Reetz, M. T.; Peyralans, J.-P.; Maichele, A.; Fu, Y.; Maywald, M. *Chemical Communications* **2006**, 4318–4320.
- [178] Yuan, L.; Kurek, I.; English, J.; Keenan, R. *Microbiology and Molecular Biology Reviews* **2005**, *69*, 373–392.
- [179] Tiwari, M. K.; Singh, R.; Singh, R. K.; Kim, I.-W.; Lee, J.-K. *Computational and Structural Biotechnology Journal* **2012**, *2*, 1–13.
- [180] Petrounia, I. P.; Arnold, F. H. *Current Opinion in Biotechnology* **2000**, *11*, 325–330.
- [181] Bornscheuer, U. T.; Pohl, M. *Current Opinion in Chemical Biology* **2001**, *5*, 137–143.
- [182] Peisajovich, S. G.; Tawfik, D. S. *Nature Methods* **2007**, *4*, 991–994.
- [183] Alexeeva, M.; Carr, R.; Turner, N. J. *Organic & Biomolecular Chemistry* **2002**, *1*, 4133–4137.
- [184] Studier, F. W.; Rosenberg, A. H.; Dunn, J. J.; Dubendorf, J. W. *Methods in Enzymology* **1990**, *185*, 60–89.
- [185] Lovrien, R. E.; Matulis, D. *Current Protocols in Protein Science* **2001**, *4.5*, 1–36.
- [186] Hofmeister, F. *Archiv für experimentelle Pathologie und Pharmakologie* **1888**, *24*, 247–260.
- [187] Collins, K. D.; Washabaugh, M. W. *Quarterly Reviews of Biophysics* **1985**, *18*, 323–422.
- [188] Sivaraman, T.; Kumar, T. K. S.; Jayaraman, G.; Yu, C. *Journal of Protein Chemistry* **1997**, *16*, 291–297.
- [189] Link, A. J.; LaBaer, J. *Cold Spring Harbor Protocols* **2011**, *8*, 993–994.
- [190] Bensadoun, A.; Weinstein, D. *Analytical Biochemistry* **1976**, *70*, 241–250.

- [191] Cohn, E. J. *Chemical Reviews* **1936**, *19*, 241–273.
- [192] Sheldon, R. A. Cross-linked enzyme aggregates (CLEAs): stable and recyclable biocatalysts. In *7th International Conference on Protein Stabilisation*, Vol. 35; 2007.
- [193] Arakawa, T.; Kita, Y.; Shiraki, K.; Ohtake, S. *Global Journal of Analytical Chemistry* **2011**, *2*, 152–167.
- [194] Janson, J.-C.; Ryden, L. *Protein purification*; VCH Publishers, Inc.: New York, USA, 1989.
- [195] Fahey, R. C.; Brown, W. C.; Adams, W. B.; Worsham, M. B. *Journal of Bacteriology* **1978**, *133*, 1126–1129.
- [196] McLaggan, D.; Logan, T. M.; Lynn, D. G.; Epstein, W. *Journal of Bacteriology* **1990**, *172*, 3631–3636.
- [197] Loewen, P. C. *Canadian Journal of Biochemistry* **1979**, *57*, 107–111.
- [198] Meister, A.; Anderson, M. E. *Annual Review of Biochemistry* **1983**, *52*, 711–760.
- [199] Masip, L.; Veeravalli, K.; Georgiou, G. *Antioxidants & Redox Signaling* **2006**, *8*, 753–762.
- [200] Smirnova, G. V.; Muzyka, N.; Oktyabrsky, O. N. *Microbiological Research* **2012**, *167*, 166–172.
- [201] Lau, S. S.; Hill, B. A.; Highet, R. J.; Monks, T. J. *Molecular Pharmacology* **1988**, *34*, 829–836.
- [202] Kosower, N. S.; Kosower, E. M.; Wertheim, B.; Correa, W. S. *Biochemical and Biophysical Research Communications* **1969**, *37*, 593–596.
- [203] Kosower, N. S.; Song, K.-R.; Kosower, E. M. *Biochimica et Biophysica Acta* **1969**, *192*, 23–29.
- [204] Kosower, E. M.; Correa, W. S.; Kinon, B. J.; Kosower, N. S. *Biochimica et Biophysica Acta* **1972**, *264*, 39–44.
- [205] Kosower, E. M.; Miyadera, T. *Journal of Medicinal Chemistry* **1972**, *15*, 307–312.
- [206] DeLucia, A. J.; Mustafa, M. G.; Hussain, M. Z.; Cross, C. E. *The Journal of Clinical Investigation* **1975**, *55*, 794–802.
- [207] Monks, T. J.; Pohl, L. R.; Gillette, J. R.; Hong, M.; Highet, R. J.; Ferretti, J. A.; Hinson, J. A. *Chemico-Biological Interactions* **1982**, *41*, 203–216.
- [208] McLaggan, D.; Rufino, H.; Jaspars, M.; Booth, I. R. *Applied and Environmental Microbiology* **2000**, *66*, 1393–1399.
- [209] Santos, M. M.; Moreira, R. *Mini-Review in Medicinal Chemistry* **2007**, *7*, 1040–1050.

- [210] Böhme, A.; Thaens, D.; Paschke, A.; Schüürmann, G. *Chemical Research in Toxicology* **2009**, *22*, 742–750.
- [211] Schwöbel, J. A. H.; Wondrousch, D.; Koleva, Y. K.; Madden, J. C.; Cronin, M. T. D.; Schüürmann, G. *Chemical Research in Toxicology* **2010**, *23*, 1576–1585.
- [212] Masri, M. S.; Friedman, M. *Journal of Protein Chemistry* **1988**, *7*, 49–54.
- [213] Gleeson, M. A.; Sudbery, P. E. *Yeast* **2004**, *4*, 1–15.
- [214] Veenhuis, M.; Salomons, F. A.; Van der Klei, I. J. *Microscopy Research Technique* **2000**, *51*, 584–600.
- [215] Johnson, M. A.; Waterham, H. R.; Ksheminska, G. P.; Fayura, L. R.; Cereghino, J. L.; Stasyk, O. V.; Veenhuis, M.; Kulachkovsky, A. R.; Sibirny, A. A.; Cregg, J. M. *Genetics* **1999**, *151*, 1379–1391.
- [216] Stewart, M. Q.; Esposito, R. D.; Gowani, J.; Goodman, J. M. *Journal of Cell Science* **2001**, *114*, 2863–2868.
- [217] Yurimoto, H.; Sakai, Y.; Kato, N. *Hansenula polymorpha: biology and applications*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 1st ed.; 2002.
- [218] Hartner, F. S.; Glieder, A. *Microbial Cell Factories* **2006**, *5*, 1–21.
- [219] Tschopp, J. F.; Brust, P. F.; Cregg, J. M.; Stillman, C. A.; Gingeras, T. R. *Nucleic Acids Research* **1987**, *15*, 3859–3876.
- [220] Inan, M.; Meagher, M. M. *Journal of Bioscience and Bioengineering* **2001**, *92*, 585–589.
- [221] Inan, M.; Meagher, M. M. *Journal of Bioscience and Bioengineering* **2001**, *92*, 337–341.
- [222] Cereghino, J. L.; Cregg, J. M. *FEMS Microbiology Reviews* **2000**, *24*, 45–66.
- [223] Krainer, F. W.; Dietzsch, C.; Hajek, T.; Herwig, C.; Spadiut, O.; Glieder, A. *Microbial Cell Factories* **2012**, *11*, 1–14.
- [224] Pla, I. A.; Damasceno, L. M.; Vannelli, T.; Ritter, G.; Batt, C.; Shuler, M. L. *Biotechnology Progress* **2006**, *22*, 881–888.
- [225] Vassileva, A.; Chugh, D. A.; Swaminathan, S.; Khanna, N. *Protein Expression & Purification* **2001**, *21*, 71–80.
- [226] Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. *Molecular biology of the cell*; Garland Publishing, Inc.: New York, USA, 2nd ed.; 1989.
- [227] Cregg, J. M. *Pichia Protocols*; volume 389 of *Methods in Molecular Biology* Humana Press: Totowa, USA, 2nd ed.; 2007.
- [228] Macauley-Patrick, S.; Fazenda, M. L.; McNeil, B.; Harvey, L. M. *Yeast* **2005**, *22*, 249–270.
- [229] Streit, W.; Entcheva, P. *Applied Microbiology and Biotechnology* **2003**, *61*, 21–31.

- [230] Gasser, B.; Dragosits, M.; Mattanovich, D. *Metabolic Engineering* **2010**, *12*, 573–580.
- [231] Jungo, C.; Urfer, J.; Zocchi, A.; Marison, I.; von Stockar, U. *Journal of Biotechnology* **2007**, *127*, 703–715.
- [232] Jahic, M. *Process techniques for production of recombinant proteins with Pichia pastoris*, MSc thesis thesis, Royal Institute of Technology, 2003.
- [233] Clare, J. J.; Romanos, M. A.; Rayment, F. B.; Rowedder, J. E.; Smith, M. A.; Payne, M. M.; Sreekrishna, K.; Henwood, C. A. *Gene* **1991**, *105*, 205–212.
- [234] Gleeson, M. A.; White, C.; Meininger, D. P.; Komives, E. A. *Generation of protease-deficient strains and their use in heterologous protein expression*; volume 103 of *Pichia Protocols* Humana Press: Totowa, USA, 1998.
- [235] Tsujikawa, M.; Okabayashi, K.; Morita, M.; Tanabe, T. *Yeast* **1998**, *12*, 541–553.
- [236] Shi, X.; Karkut, T.; Chamankhah, M.; Alting-Mees, M.; Hemmingsen, S. M.; Hegedus, D. *Protein Expression & Purification* **2003**, *28*, 321–330.
- [237] Brierley, R. A. *Secretion of recombinant human insulin-like growth factor I (IGF-I)*; volume 103 of *Methods in Molecular Biology* Humana Press: Totowa, USA, 1998.
- [238] Wegner, E. H. “Biochemical conversions by yeast fermentation at high cell densities”, Patent US4414329A, United States Patent and Trademark Office, 1993.
- [239] Werten, W. W.; van den Bosch, T. J.; Wind, R. D.; Mooibroek, H.; de Wolf, F. A. *Yeast* **1999**, *15*, 1087–1096.
- [240] Potgieter, T. I.; Cukan, M.; Drummond, J. E.; Houston-Cummings, N.; Jiang, Y.; Li, F.; Lynaugh, H.; Mallem, M.; McKelvey, T. W.; Mitchell, T.; Nysten, A.; Rittenhour, A.; Stadheim, T.; Zha, D.; d’Anjou, M. *Journal of Biotechnology* **2009**, *139*, 318–325.
- [241] Barnard, G. C. *et al.* *Journal of Industrial Microbiology & Biotechnology* **2010**, *37*, 961–971.
- [242] Gurramkonda, C.; Polez, S.; Skoko, N.; Adnan, A.; Gaebel, T.; Chugh, D.; Swaminathan, S.; Khanna, N.; Tisminetzky, S.; Rinas, U. *Microbial Cell Factories* **2010**, *9*, 1–11.
- [243] Ye, J.; Ly, J.; Watts, K.; Hsu, A.; Walker, A.; McLaughlin, K.; Berdichevsky, M.; Prinz, B.; Kersey, D. S.; d’Anjou, M.; Pollard, D.; Potgieter, T. *Biotechnology Progress* **2011**, *27*, 1744–1750.
- [244] Medina-Godoy, S.; Valdez-Ortiz, A.; Valverde, M. E.; Paredes-López, O. *Biotechnology Journal* **2006**, *1*, 1085–1092.
- [245] Li, P.; Anumanthan, A.; Gao, X.-G.; Ilangovan, K.; Suzara, V. V.; Düzgüneş, N.; Renugopalakrishnan, V. *Applied Biochemistry and Biotechnology* **2007**, *142*, 105–124.

- [246] Kaewthai, N.; Harvey, A. J.; Hrmova, M.; Brumer, H.; Ezcurra, I.; Teeri, T. T.; Fincher, G. B. *Plant Biotechnology* **2010**, *27*, 251–258.
- [247] Jahic, M.; Gustavsson, M.; Jansen, A.-K.; Martinelle, M.; Enfors, S.-O. *Journal of Biotechnology* **2003**, *102*, 45–53.
- [248] Inan, M.; Meagher, M. M.; Plantz, B. A.; Sinha, J. *Biotechnology and Bioengineering* **2005**, *89*, 102–112.
- [249] Scorer, C. A.; Buckholz, R. G.; Clare, J. J.; Romanos, M. A. *Gene* **1993**, *136*, 111–119.
- [250] Lombardi, A.; Bursomanno, S.; Lopardo, T.; Traini, R.; Colombatti, M.; Ippoliti, R.; Flavell, D. J.; Flavell, S. U.; Ceriotti, A.; Fabbri, M. S. *The FASEB Journal* **2010**, *24*, 253–265.
- [251] Köhler, V.; Wilson, Y. M.; Dürrenberger, M.; Ghislieri, D.; Churakova, E.; Quinto, T.; Knörr, L.; Häussinger, D.; Hollmann, F.; Turner, N. J.; Ward, T. R. *Nature Chemistry* **2013**, *5*, 93–99.
- [252] Tracewell, C. A.; Arnold, F. H. *Current Opinion in Chemical Biology* **2009**, *13*, 3–9.

New strategies for the purification of streptavidin

We must have perseverance and above all, confidence in ourselves. We must believe that we are gifted for something.

Marie Skłodowska-Curie

Abstract

This Chapter reports on the attempts to develop a simple and straightforward technique to produce functional and pure streptavidin for high-throughput screening (HTS) of organometallic complexes. Three categories of protein precipitation techniques (organic solvent, acid and salt), with and without a previous step of dialysis against a chaotropic agent (guanidinium-chloride or urea), and small-scale affinity purification (SSP) in a 24-well plate format were tested as fast-purification methods. The “neutralisation” of reduced glutathione (GSH) present in crude extracts was also investigated. The potential of the methodologies developed was tested on the asymmetric transfer hydrogenation of imines and/or on the reductive amination of α -keto acids.

The pitfalls and practical issues surrounding high-throughput protein purification – in the context of screening protein-based artificial metalloenzymes – rely on the practicability and reproducibility of the method in parallel experiments, and on the final yield of the protein. Albeit promising and time-saving, protein precipitation was found to be inadequate when dealing with small amounts of protein, whereas small-scale purification was hindered by the amounts and final concentrations of protein obtained, though one positive hit for the production of (*S*)-phenylalanine was identified through this method.

The “neutralisation” of reduced glutathione by addition of a chemical agent, 1,1'-azobis(*N,N*-dimethylformamide), holds great potential for high-throughput pre-screening of hybrid catalysts, as it ultimately enabled catalysis on a scale that was no longer limited by sample preparation. The whole purification process of streptavidin was shorten from 12 days to just five.

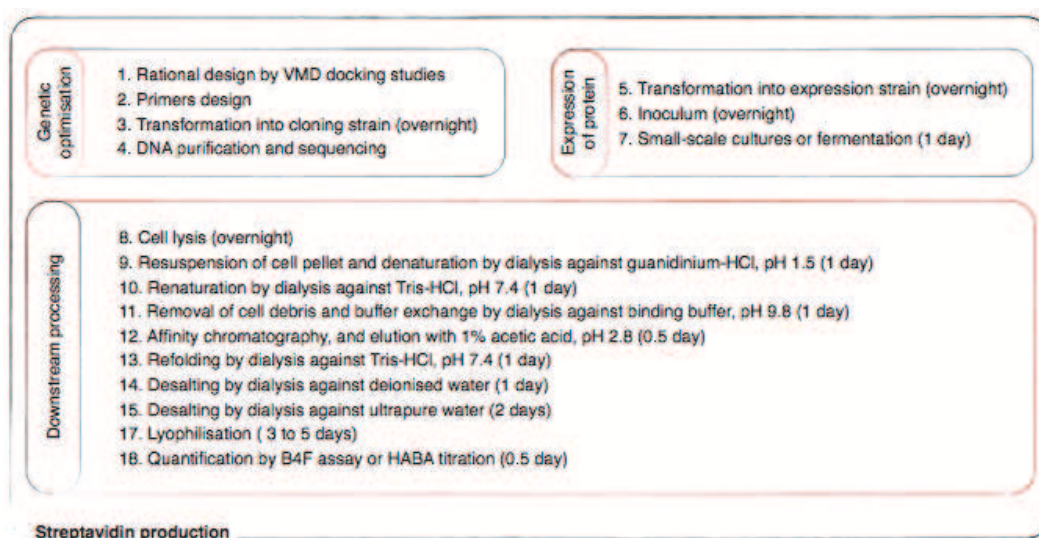
2.1 Introductory remarks

The basis of screening and selection methodologies for artificial metalloenzymes relies on the linkage between the host protein and a catalytically active organometallic moiety, and the reaction performed by this newly created artificial metalloenzyme. This system can be optimised via a chemo-genetic approach, that is the design of the primary coordination sphere (*i.e.* the ligands that directly coordinate to the metal ion) and of the second coordination sphere (*i.e.* the amino acid residues around the ligand).^[1-3] The latter plays an important role, as it modifies the local environment of the achiral metal moiety by providing a chiral environment around it, and induces enantioselectivity by interacting with the metal and/or the prochiral substrate through weak interactions.^[3-5] However, the number of variables at play in the expression and purification of a single protein dwarf those involved in synthesising new catalysts. To accelerate the screening process by constructing a large library of streptavidin mutants requires the development of protocols to reduce or simplify the challenge of producing protein in sufficient quantity and appropriate purity.

2.1.1 A long process: from the gene to the protein

Expressing and purifying proteins, especially in the work-horse host organism *Escherichia coli*, remains a major bottleneck in high-throughput screening.

Scheme 2.1. Processing flowchart of streptavidin expression in *E. coli*. The genetic optimisation based on VMD docking studies, the subsequent site-directed mutagenesis, and the expression of the protein are standard procedures, which cannot be optimised time-wise. Depending on external factors (*e.g.* outsourced services: time of delivery of primers and sequencing), these two processes can last at best between five to seven days. The downstream processing can take up to 12 days to be completed. The time length of the whole process is in average three to four weeks.

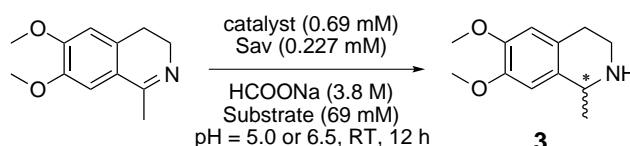


Historically, the production (expression and purification) of functional streptavidin involved a small number of selected variants, and a tedious and extremely time-consuming process, which resulted in screening on a low-throughput mode, and on a protein-by-protein basis. A complete rethinking of the production strategy was required to transform this process into a genuine high-throughput screening. Hence, an arsenal of methods (protein precipitation, small-scale purification, and neutralisation of reduced glutathione) was established to streamline steps in the purification process, identified as one of the bottlenecks of the pipeline (Scheme 2.1).

2.1.2 Asymmetric transfer hydrogenation of imines

The reduction of cyclic imines by asymmetric transfer hydrogenation (ATH) was selected as the model reaction to screen the potential of the three high-throughput purification methods developed herein. A new generation of biotinylated complexes developed and synthesised by Dr Jeremy Zimbron and Mr Marc Dürrenberger was used throughout this Chapter. The new scaffold involved d^6 -transition metal (iridium and rhodium) piano stool complexes that contained a pentamethylcyclopentadienyl moiety (Cp^*) tethered to biotin for incorporation into the biomolecular host, streptavidin. The resulting artificial metalloenzymes were tested in the ATH of imines, using the catalytic conditions described by Ogo *et al.* and the precursor of salsolidine **3**, 1-methyl-6,7-dimethoxy-3,4-dihydroisoquinoline, as substrate (Scheme 2.2).^[6]

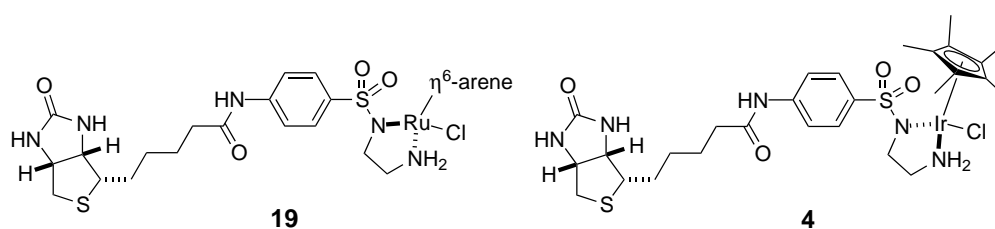
Scheme 2.2. Asymmetric transfer hydrogenation of imines for the production of salsolidine, **3**.



First generation of biotinylated complexes

The first generation of biotinylated complexes used for the asymmetric transfer hydrogenation was based on the highly active and selective catalyst $[\eta^6\text{-(arene)Ru(Biot-}q\text{-L)Cl}]$ **19** developed by Letondor *et al.* (Scheme 2.3).^[7]

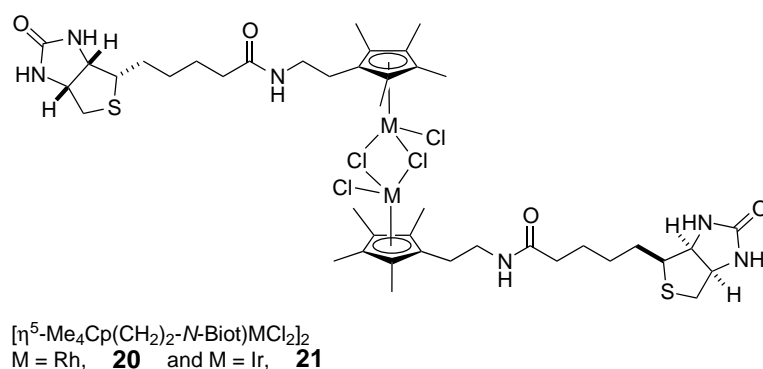
Scheme 2.3. Biotinylated complexes developed by Ward and co-workers for the asymmetric transfer hydrogenation of ketones **19** and imines **4**.^[7,8]



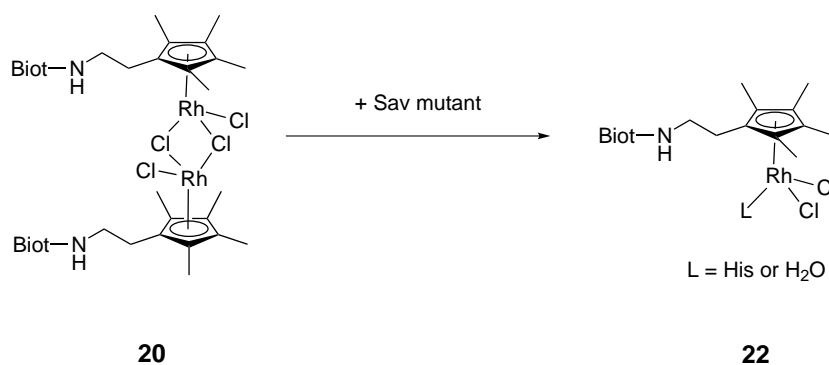
Second generation of biotinylated complexes

A second generation of biotinylated complexes with direct attachment of the biotin to the arene cap and an ethylene spacer, $[\eta^5\text{-Me}_4\text{Cp}-(\text{CH}_2)_2\text{-}N\text{-Biot})\text{MCl}_2]_2$ was developed by Dr Jeremy Zimbron (Scheme 2.4), and were used as starting material for the synthesis of diverse mononuclear biotinylated piano stool complexes bearing an additional bidentate ligand. To place the metal in close proximity to histidines (either in position 112 or 121), the ethylene spacer was used between the biotin anchor and the Cp*.^[9] The bioconjugation of the dimers with streptavidin mutants bearing a suitable histidine residue yielded dually anchored artificial metalloenzymes.^[9,10]

Scheme 2.4. Biotinylated dimeric rhodium **20** and iridium **21** complexes.



Scheme 2.5. Anticipated biotinylated monomeric rhodium **22** complex embedded into S112H or K121H.



Genetic optimisation

In silico docking experiments were performed by Dr Maurus Schmid on the structure of streptavidin against the biotinylated complexes, **20** and **22** (Figure 2.1). This simulation was a valid alternative to laborious and time-consuming large-scale mutagenesis, as it allowed the identification of potential residues that may influence the binding of the catalyst but also the type of spacer needed between the biotin anchor and the (Cp*)-M (M = Ir or Rh).

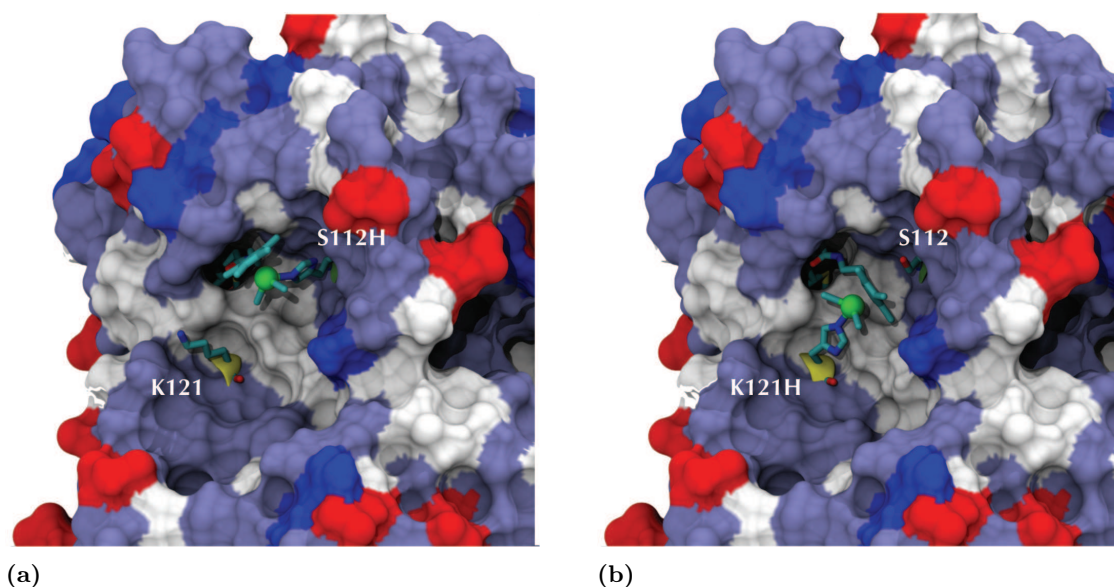


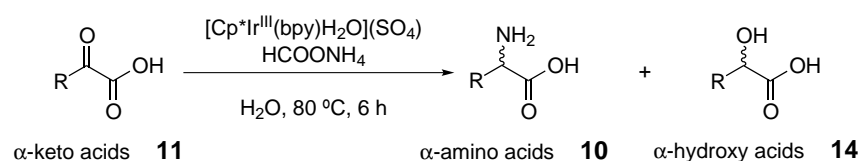
Figure 2.1. Docking simulation of the biotinylated monomeric rhodium (**22**) complex, embedded into (a) S112H or (b) K121H (the histidine is provided by the adjacent Sav monomer of the homotetrameric structure). Streptavidin is represented as solvent accessible surface (red: anionic-; blue: cationic-; white: apolar-; and grey: polar residues). The side-chains of streptavidin at positions 112 and 121 are represented as sticks. Pictures by MS.^[9]

As a result of this study, double mutants based on S112H and K121H templates were designed, expressed and “purified” using one of the methods aforementioned. Based on S112H template, mutations at residues N49x (x = A, S, W, F, D, E, R, and C), K121x (x = A, W, F, D, E, R, and C) and L124x (x = A) were introduced. Using K121H as template, a second round of mutagenesis was performed on the residue S112x (x = A, W, F, D, E, R, and C).

2.1.3 Asymmetric reductive amination of α -keto acids

Based on the previous work reported by Fukuzumi *et al.* for the synthesis of α -amino acids catalysed by bipyridine iridium complexes (Scheme 2.6), using HCOONH_4 as an amine and hydrogen source,^[11] two artificial metalloenzymes were developed (i) by combination of functionalised piano stool complexes associated with bidentate ligands or (ii) by hybrid catalysts formed with complex **22** \subset S112H and **22** \subset K121H, for the synthesis of a natural α -amino acid (phenylalanine **12**).^[10]

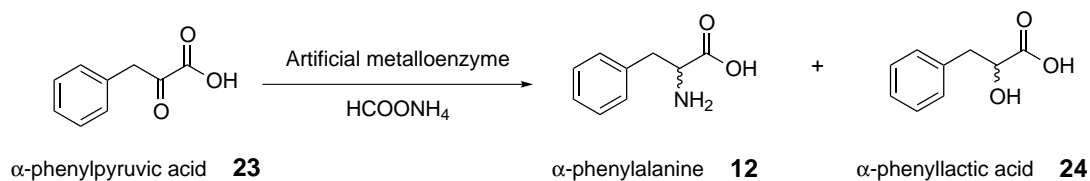
Scheme 2.6. Reductive amination of α -keto acids **11** for the synthesis of racemic α -amino acids **10**.



Substrate scope

The artificial metalloenzymes, **22** \subset Sav mutants, were tested on the enantioselective synthesis of the α -amino acid, phenylalanine **12** (Scheme 2.7).

Scheme 2.7. Reductive amination of α -phenylpyruvic acid **23**.



Genetic optimisation

To gain structural insight and to fully exploit the catalytic potential of the complex **20**, *in silico* screening of the biomolecular scaffold (K121H mutant as a template) against the organometallic moiety was carried out by Dr Maurus Schmid for a second round of genetic optimisation. The leucine at position 110 was identified as a possible hit, as it lies in close vicinity of the metal centre. Therefore, K121H-L110x mutants were expressed and purified via small-scale purification (SSP).

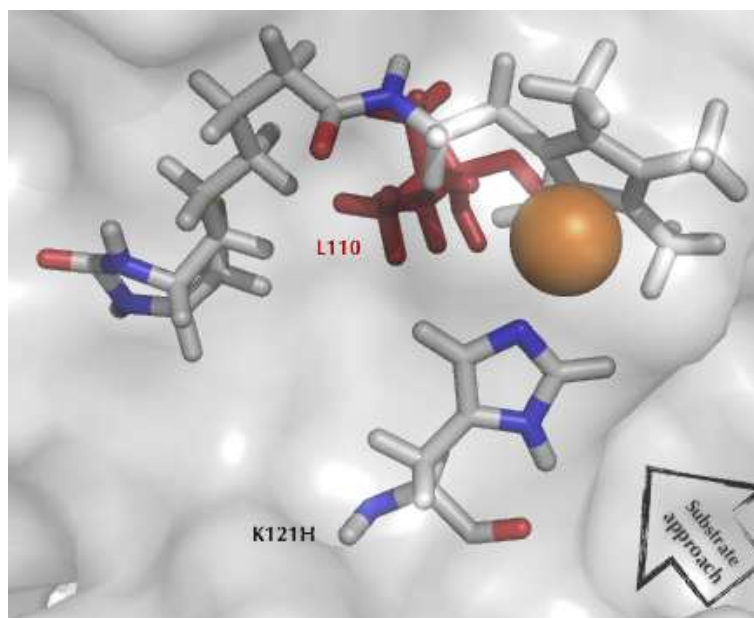


Figure 2.2. Model of the biotinylated rhodium catalyst **22** embedded into the host protein, Sav K121H. Residue leucine at position 110 is highlighted in red.

2.1.4 Research project

One of the major bottlenecks of high-throughput screening of hybrid catalysts is the lack of a simple, robust and fast system for protein purification. This was the impetus for improving the production process of mutant proteins, and downstream screening efficiency of artificial metalloenzymes in catalytic reactions. Three new experimental strategies were developed and tested with two newly synthesised d⁶-transition metal (iridium and rhodium) piano stool complexes, for the transfer hydrogenation of imines and/or the reductive amination of α -keto acids.

2.2 Results & discussion

Driven by the need to effectively and efficiently screen large libraries of artificial metalloenzymes, new protocols to produce streptavidin were developed. The results presented herein are a compilation of selected data obtained in collaboration with Dr Jeremy Zimbron and Dr Yvonne Wilson.

2.2.1 Screening on crude protein extracts

Crude protein extracts are prepared by removal of cellular debris generated by cell lysis, and contain a complex mixture of proteins from *E. coli* cells cytoplasm, and additional macromolecules such as cofactors and nutrients. Crude extract may be used for some applications in biotechnology, however, if purity is an issue, subsequent purification of the protein is performed. It has to be stressed that studies on screening organometallic moieties on crude protein extracts, *i.e.* non-purified protein, or on proteins purified by other means than by conventional affinity chromatography have been previously conducted in the research group of Prof. Ward.^[12-14] The first method developed to accelerate the optimisation process was based on the extraction-immobilisation with biotin-sepharose to capture streptavidin from crude cellular extracts. This straightforward protocol afforded > 92% *ee* for the enantioselective hydrogenation of *N*-protected dehydroamino acids, and > 90% *ee* for the transfer hydrogenation of prochiral ketones.^[12,13] The second method involved thermal treatment (with and without a dialysis step) as a rapid purification method of crude streptavidin. This protocol was tested on the palladium catalysed asymmetric allylic alkylation, and > 70% *ee* was obtained, although conversion was very low (< 5%).^[14] In order to assess the degree of inhibition of the catalyst **4**, [Cp*Ir(Biot-*p*-L)Cl], by contaminants present in crude protein extracts, preliminary experiments were performed with the hybrid catalysts, **4** \subset Sav (wt, S112A and S112K), for the asymmetric transfer hydrogenation of imines (Scheme 2.2).

Table 2.1. Results for the production of salsolidine **3** using the biotinylated piano stool complex **4** \subset Sav (wt, S112A or S112K).^a

Entry	Protein	Description	Conv. ^b [%]	ee ^b [%]
1	wild-type	purified by affinity chromatography	quant.	60 (<i>R</i>)
2	S112A	purified by affinity chromatography	99	81 (<i>R</i>)
3	S112K	purified by affinity chromatography	93	64 (<i>S</i>)
4	wild-type	crude protein extract, dialysed ^d	20	40 (<i>R</i>)
5	S112A	crude protein extract, dialysed ^d	15	46 (<i>R</i>)
6	S112K	crude protein extract, dialysed ^d	10	rac.
7	pET11b ^c	crude protein extract, dialysed ^d	17	rac.

^a The reactions were carried out at RT for 24 h, using 1 mol% complex **4** (690 μ M final concentration) and 0.33 mol% tetrameric streptavidin at pH 7.25 (MOPS buffer, 2.9 M) containing 3.65 M HCOONa. Total reaction volume: 200 μ L.

^b Determined by HPLC after extraction.

^c pET11b: Empty plasmid.

^d Crude protein extract, dialysed 1x against guanidinium-HCl, pH 1.5, 1x against 20 mM Tris-HCl, pH 7.4, and 2x against deionised water (dH₂O).

Having realised that the catalyst **4** was moderately active in dialysed crude protein extracts (Table 2.1, entries 4 to 7, conv. \leq 20%), the degree of protein purity was optimised by precipitating the cell-free extracts.

2.2.2 Screening on proteins purified by precipitation

In order to achieve an acceptable level of purity for screening, a series of protein precipitation techniques, preceded or not by dialysis, were tested (Figure 2.3).

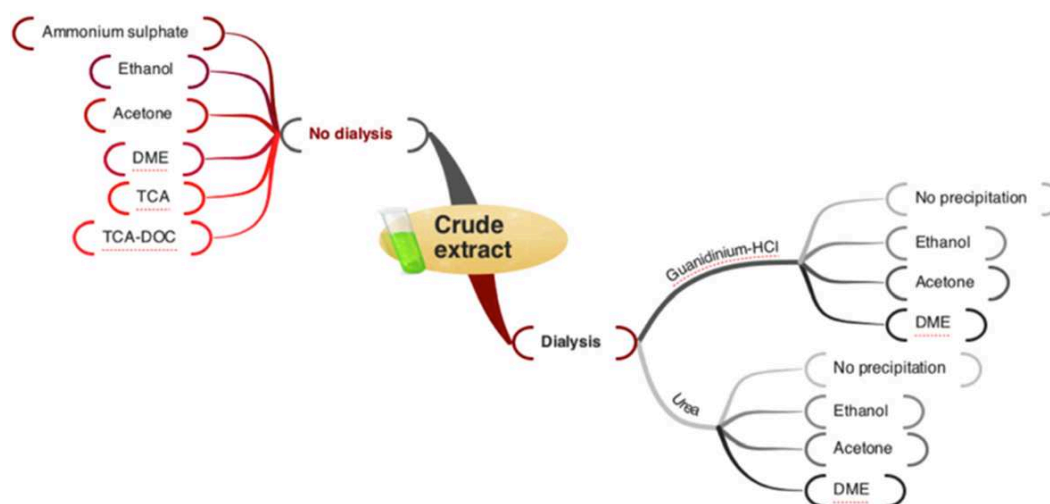


Figure 2.3. Mind map of the protein purification techniques used to purify streptavidin from crude protein extracts. The dialysis step consisted in a three-stage procedure: 24 h dialysis against guanidinium-HCl or urea, followed by 24 h dialysis against Tris-HCl, and a final dialysis for 48 h against deionised water (dH₂O). DME: dimethoxyethane; TCA: trichloroacetic acid; and TCA-DOC: trichloroacetic acid + sodium deoxycholate.

The viability and effectiveness of the methods were first assessed by SDS-PAGE analysis (Figure 2.4). The salting-out precipitation was carried out with ammonium sulphate at concentrations between 10 and 90% of total saturated salt solution. Best results were obtained at a concentration of 50%, as the protein was mainly in the precipitate, and none was detected in the supernatant (data not shown). Although TCA (trichloroacetic acid) precipitation is considered one of the most efficient protocol for the precipitation of proteins, recovery of streptavidin was not achieved using this technique, nor the addition of the carrier deoxycholate (DOC) improved the final yield (data not shown).^[15,16] These purification techniques were not pursued further, due to non-reproducibility of the method ((NH₄)₂SO₄) or to low recovery of the target protein (TCA and TCA-DOC). Precipitation by addition of organic solvents (ethanol, acetone and DME) proved to be successful. Best results were achieved when dialyses against guanidinium-HCl/Tris-HCl/dH₂O forego the precipitation step, as seen on Figure 2.4, lanes 7 to 9. The sample of lane 9 was partially loaded on the gel, as a jelly precipitate was formed and could not be completely dissolved. Dialysis against urea (instead of guanidinium-HCl) was also effective in removing contaminants (Figure 2.4, lanes 10 to 12), when compared to the samples treated only by precipitation (Figure 2.4, lanes 4 to 6). Thus, dialyses were effectual in removing small molecular weight species.

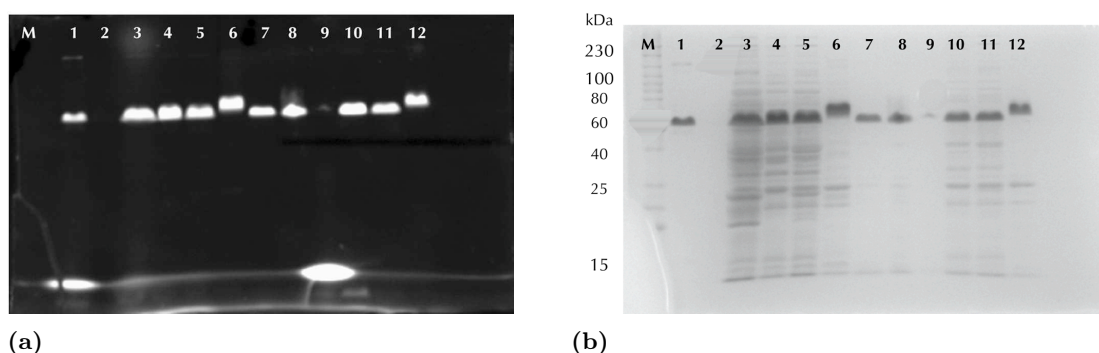


Figure 2.4. SDS-PAGE of different techniques used to precipitate wild-type streptavidin. **(a)** B4F analysis. **(b)** Coomassie Blue staining analysis. Lane 1: Sav purified by affinity chromatography; Lane 2: Empty lane; Lane 3: Sav unpurified; Lane 4: Sav precipitated in ethanol; Lane 5: Sav precipitated in acetone; Lane 6: Sav precipitated in DME; Lane 7: Sav dialysed in guanidinium-HCl and precipitated in ethanol; Lane 8: Sav dialysed in guanidinium-HCl and precipitated in acetone; Lane 9: Sav dialysed in guanidinium-HCl and precipitated in DME; Lane 10: Sav dialysed in urea and precipitated in ethanol; Lane 11: Sav dialysed in urea and precipitated in acetone; Lane 12: Sav dialysed in urea and precipitated in DME.

The two methods (dialysis combined with precipitation, or precipitation only) were taken a step further. Preliminary experiments were carried out to determine the catalytic activity of **4** in the presence of proteins purified by precipitation (wild-type streptavidin and empty plasmid) for the asymmetric transfer hydrogenation of imines (Scheme 2.2). The empty plasmid (pET11b) does not express Sav but contains all native proteins from *E. coli*, thus it was used as a negative control.

Table 2.2. Results for the production of salsolidine **3** using the biotinylated piano stool complex **4** \subset wt Sav or pET11b.^a

Entry	Protein	Protein concentration [μ M]	Conditions	Conv. ^b [%]	ee ^b [%]
1	pET11b	67 ^d	crude extract	– ^e	n.d. ^e
2	pET11b	67 ^d	dialysed ^c	17	rac.
3	pET11b	67 ^d	dialysed ^c , ethanol precipitation	– ^e	n.d. ^e
4	pET11b	67 ^d	dialysed ^c , acetone precipitation	quant.	rac.
5	pET11b	67 ^d	dialysed ^c , DME precipitation	16	rac.
6	pET11b	67 ^d	ethanol precipitation	– ^e	n.d. ^e
7	pET11b	67 ^d	acetone precipitation	– ^e	n.d. ^e
8	pET11b	67 ^d	DME precipitation	– ^e	n.d. ^e
9	wild-type	191	crude extract	– ^e	n.d. ^e
10	wild-type	67	purified	58	51 (<i>R</i>)
11	wild-type	340	purified	quant.	57 (<i>R</i>)
12	wild-type	178	dialysed ^c	20	40 (<i>R</i>)
13	wild-type	163	dialysed ^c , ethanol precipitation	83	51 (<i>R</i>)
14	wild-type	20	dialysed ^c , acetone precipitation	24	13 (<i>R</i>)
15	wild-type	39	dialysed ^c , DME precipitation	31	18 (<i>R</i>)
16	wild-type	63	ethanol precipitation	16	rac.
17	wild-type	63	acetone precipitation	– ^e	n.d. ^e
18	wild-type	40	DME precipitation	– ^e	n.d. ^e

^a The reaction were carried out at RT for 24 h, using 1 mol% complex **4** *vs* substrate (690 μ M final metal concentration, 69 mM substrate) and 0.33 mol% tetrameric streptavidin at pH 7.25 (MOPS buffer, 2.9 M) containing 3.65 M HCOONa. Total reaction volume: 200 μ L.

^b Determined by HPLC after extraction.

^c Crude extract, dialysed 1x against guanidinium-HCl pH 1.5, 1x 20 mM Tris-HCl pH 7.4, and 2x deionised water.

^d The concentration of pET11b is approximative, as there is no method to quantify the proteins present in *E. coli* cells.

^e n.d. not determined – peaks were too small to determine conversion or *ee*.

The positive control (wt Sav purified by conventional affinity chromatography) was performed at low and normal catalytic concentration (67 and 340 μ M, Table 2.2: entries 10 and 11, respectively) as it was anticipated that purification by precipitation would yield very low amounts of target protein. Initial results with the empty plasmid were not promising (entries 1 to 8), as very low conversions were obtained for all samples, except for the dialysed/acetone precipitated sample (entry 4). These results confirmed the sensitivity of the catalyst toward macromolecules present in *E. coli* cells. However, [Cp*Ir(Biot-*p*-L)Cl] \subset wt Sav yielded exceptional results (entries 9 to 18), even at low protein concentrations (20 to 178 μ M). Indeed and compared to the quantitative conversion and 57% *ee* (*R*) obtained with purified wt Sav at 690 μ M [Cp*Ir(Biot-*p*-L)Cl] concentration, nearly identical results was obtained with wt Sav purified by dialysis/ethanol precipitation (entry 13). This hybrid catalyst exhibited 51% *ee*

(*R*) and 83% conversion at low catalyst concentrations (85 μM). Another promising method was the dialysis/DME precipitation, which rendered 31% conv. and 18% *ee* (*R*) (entry 15) at a concentration of catalyst as low as 20 μM .

Complex **4** was screened with S112x ($x = \text{A}$ or K) to narrow down the four best techniques to one (Table 2.3, entries 1 to 12).

Table 2.3. Selected results for the production of salsolidine **3** using the biotinylated piano stool complex **4** C Sav S112A or S112K.^a

Entry	Protein	Protein concentration [μM]	Conditions	Conv. ^b [%]	<i>ee</i> ^b [%]
1	S112A	96	purified	78	64 (<i>R</i>)
2	S112A	69	dialysed ^c	20	48 (<i>R</i>)
3	S112A	23	dialysed ^c , ethanol precipitation	48	67 (<i>R</i>)
4	S112A	39	dialysed ^c , DME precipitation	51	55 (<i>R</i>)
5	S112A	80	ethanol precipitation	18	rac.
6	S112A	96	DME precipitation	– ^d	n.d. ^d
7	S112K	96	purified	40	44 (<i>S</i>)
8	S112K	79	dialysed ^c	– ^d	n.d. ^d
9	S112K	30	dialysed ^c , ethanol precipitation	21	36 (<i>S</i>)
10	S112K	39	dialysed ^c , DME precipitation	25	36 (<i>S</i>)
11	S112K	110	ethanol precipitation	– ^d	n.d. ^d
12	S112K	39	DME precipitation	– ^d	n.d. ^d

^a The reaction were carried out at RT for 24 h, using 1 mol% complex **4** (690 μM final concentration) and 0.33 mol% tetrameric streptavidin at pH 7.25 (MOPS buffer, 2.9 M) containing 3.65 M HCOONa. Total reaction volume: 200 μL .

^b Determined by HPLC after extraction.

^c Crude extract, dialysed 1x against guanidinium-HCl, pH 1.5, 1x 20 mM Tris-HCl, pH 7.4, and 2x deionised water.

^d n.d. not determined – peaks were too small to determine conversion or *ee*.

Even at low protein concentration ($\geq 30 \mu\text{M}$), the correct enantiomer was obtained for both mutants tested: (*R*) for S112A, and (*S*) for S112K. From this screening, it was confirmed that the techniques involving dialyses and precipitation with ethanol or DME were promising methods (entries 3, 4, 9 and 10). The whole purification process was shortened from 12 days down to five days. Both approaches yielded similar results: 67% *ee* (*R*) and 48% conversion *vs* 55% *ee* (*R*) and 51% conversion for S112A (entries 3 and 4), and 36% *ee* (*R*) and 21% conversion *vs* 36% *ee* (*R*) and 25% conversion for S112K purified by dialysis plus ethanol or DME precipitation, respectively (entries 9 and 10). Consequently, these strategies were performed on another system based on the catalyst **22** and Sav mutants, S112H and K121H (Table 2.4).

Table 2.4. Selected results for the production of salsolidine **3** using the biotinylated piano stool complex **22** \subset Sav S112H or K121H.^a

Entry	Protein	Protein concentration [μ M]	Conditions	Conv. ^b [%]	ee ^b [%]
1	S112H	442	purified	quant.	38 (<i>S</i>)
2	S112H	240	dialysed ^c , ethanol precipitation	30	37 (<i>S</i>)
3	S112H	238	dialysed ^c , DME precipitation	24	34 (<i>S</i>)
4	S112H	148	ethanol precipitation	– ^d	n.d. ^d
5	S112H	151	DME precipitation	– ^d	n.d. ^d
6	K121H	469	purified	99	78 (<i>R</i>)
7	K121H	228	dialysed ^c , ethanol precipitation	58	48 (<i>R</i>)
8	K121H	206	dialysed ^c , DME precipitation	27	33 (<i>R</i>)
9	K121H	157	ethanol precipitation	– ^d	n.d. ^d
10	K121H	162	DME precipitation	16	rac.

^a The reaction were carried out at 55 °C for 12 h, using 2 mol% complex **22** (680 μ M final concentration, 34 mM substrate) and 0.66 mol% tetrameric streptavidin at pH 6.5 (MOPS buffer, 3.1 M) containing 3.88 M HCOONa. Total reaction volume: 200 μ L.

^b Determined by HPLC after extraction.

^c Crude extract, dialysed 1x against guanidinium-HCl, pH 1.5, 1x 20 mM Tris-HCl, pH 7.4, and 2x deionised water.

^d n.d. not determined – peaks were too small to determine conversion or *ee*.

The use of simple precipitation, with ethanol or DME (without dialyses), was ruled out as purification methods, since they consistently yielded no/low conversion and no enantiomeric excess (entries 4, 5, 9 and 10). Again, no difference could be distinguished between the two methods involving dialysis and precipitation (EtOH or DME) as both exhibited moderate conversion and *ee* (entries 2, 3, 7 and 8) compared to the controls (entries 1 and 6), although for these hybrid catalysts (**20** \subset Sav S112H or K121H) higher concentrations of the host protein were needed.

A library of 19 mutants (expressed in 50 mL scale) was screened using this approach. The screening, however, did not yield any promising results (data not shown). Most likely due to the very low amount of protein obtained (as a sum of low protein expression in small-scale cultures and method of purification), conversions and enantioselectivities could not be determined.

Albeit promising and time-saving, the purification of proteins by precipitation was found to be inadequate for high-throughput screening. When dealing with small amounts of protein (*e.g.* crude extracts from small-scale cultures), the recovery of the precipitated pellet was painstaking and very labour-intensive. It would be (almost) impossible to perform this technique in a smaller scale, *i.e.* 24 or 96-well plate format. Furthermore, reproducibility of results (protein concentration *vs* catalysis) was low, when working at low protein concentrations with different systems. To implement high-throughput screening, the method of choice to produce

the biomolecular host should be well developed and reliable to avoid false “positive/negative” hits.

In another attempt to develop a steady method to purify the target protein in parallel, small-scale affinity purification in a 24-well plate was implemented (Section 2.2.3).

2.2.3 Screening on proteins purified on a small-scale

Based on the automated high-throughput purification of hexahistidines-tagged proteins, a 24-well affinity chromatography module was created by packing 2-iminobiotin sepharose resin (400 μL , wet volume) into each well of the plate. The purification procedure was optimised relatively to (i) the volume of beads in each well (400 μL); (ii) the wash and elution buffers (deionised water and formic acid, respectively); (iii) the number and volume of washing steps (four times 1 mL of deionised water); and (iv) the number and volume of elution steps (pre-elution: 150 μL of 100 mM formic acid, pH 2.34; elution: 250 μL of 200 mM formic acid, pH 2.11).

Based on docking studies, modifications of residues with close contacts with the metal centre were introduced, assuming that those have more influence on catalysis than distant ones. A library of 20 mutants was produced, by introducing a second mutation into the templates S112H (K121x and N49x) or K121H (L110x and aviloop). The loop L3-4 in streptavidin (Asn49 and Ser88) is three residues shorter compared to the analogous loop in avidin (Thr38, Ala39, Thr40, Ser73, and Ser75), which is why Sav has a slightly lower affinity for biotin compared with avidin.^[17] Mutating the loop of Sav with the longer loop from avidin should, in principle, result in higher binding affinity toward biotin.

Asymmetric transfer hydrogenation of imines

With the aim of testing this newly developed method, the library of genetically modified proteins was produced via small-scale purification (SSP). The biotinylated piano stool complex **22** was incorporated into Sav mutants, and the resulting hybrid catalysts were screened for activity and selectivity for the ATH of imines. Table 2.5 summarises the results using double mutants of S112H and K121H. As controls, the single mutants S112H and K121H were purified by conventional affinity chromatography (large scale) and via SSP, and were also tested on the ATH of imines.

For the designed mutants, a second mutation introduced into the templates S112H (K121x and N49x) or K121H (L110x and aviloop) inhibited the transfer hydrogenation catalysed by the artificial metalloenzymes (entries 3 to 13, and 10 to 22, respectively). Overall, very low conversions and enantioselectivities were obtained. However, even at low protein concentration, S112H bearing a second, potentially coordinating residue (glutamic acid and histidine) at position 121 (entries 3 to 5) revealed a modest reaction yield (43% and 22% conversion,

and 31% and 17% *ee*, respectively), and mutant S112H-K121W rendered the (*S*)-enantiomer in 27% conversion and 19% *ee*. Thus additional contacts between the substrate and the biomolecular scaffold may compete (matched or mismatched) with interactions between the substrate and the Cp*.

Table 2.5. Selected results for the production of salsolidine **3** using the biotinylated piano stool complex **22** \subset library of mutants. Sav S112H and K121H were used as template for a second round of mutagenesis.^a

Entry	Protein	Protein concentration [μ M]	Conv. ^b [%]	<i>ee</i> ^b [%]
1	S112H ^d	296	99	42 (<i>S</i>)
2	S112H ^e	60	53	rac.
3	S112H-K121E	125	43	31 (<i>S</i>)
4	S112H-K121W	125	27	19 (<i>R</i>)
5	S112H-K121H	125	22	17 (<i>S</i>)
6	S112H-K121D	165	17	rac.
7	S112H-K121F	125	– ^g	n.d. ^g
8	K121H ^d	275	quant.	78 (<i>R</i>)
9	K121H ^e	100	83	71 (<i>R</i>)
10	K121H-K121A	125	– ^g	n.d. ^g
11	K121H-N49S	180	21	12 (<i>S</i>)
12	K121H-N49A	40	– ^g	n.d. ^g
13	K121H-N49E	70	– ^g	n.d. ^g
14	K121H-N49F	120	– ^g	n.d. ^g
15	K121H-N49C	30	– ^g	n.d. ^g
16	K121H-S112A	125	17	rac.
17	K121H-L110A	70	– ^g	n.d. ^g
18	K121H-L110C	320	– ^g	n.d. ^g
19	K121H-L110K	80	– ^g	n.d. ^g
20	K121H-L110E	60	– ^g	n.d. ^g
21	K121H-L110D	40	– ^g	n.d. ^g
22	K121H, aviloop ^f	30	– ^g	n.d. ^g

^a The reaction were carried out at 55 °C for 12 h, using a fixed concentration of salsolidine precursor (68 mM with **20** \subset S112H, and 45.8 mM with **20** \subset K121H) at pH 5.0 (MOPS buffer, 3.1 M) containing 3.8 M HCOONa. Total reaction volume: 200 μ L.

^b Determined by HPLC after extraction.

^c Protein concentration determined on a NanoDrop.

^d Purified by conventional affinity chromatography.

^e Purified by small-scale affinity chromatography.

^f Aviloop = L3,4 loop of Sav exchanged with the corresponding loop of avidin.^[17]

^g n.d. not determined – peaks were too small to determine conversion or *ee*.

The overestimation of the protein concentration would be another plausible explanation for the results presented in Table 2.5. To ensure maximal protein concentration, the proteins were eluted in a minimal volume, *i.e.* 250 μ L, which precluded the titration by biotin-4-fluorescein. Hence, the protein concentration was quantified on a NanoDrop spectrophotometer (A_{280}), assuming four free-biotin binding sites. This overestimation might have lead to false “negative” hits. Moreover, it is known that protein expression in small-scale is often drastically lower than in large-scale culture. One example is K121H, which was expressed in a fermentor. When

purified by conventional or small-scale purification, K121H afforded high conversion and *ee* (entries 8 and 9). On the other hand, the protein S112H was expressed in a fermentor (entry 1) and in a 50 mL culture (entry 2), and moderate conversion (53%) and racemic product were obtained for the sample purified in small-scale. In this system, the final concentration of the organometallic catalyst is a key issue, as high concentrations (*e.g.* 50 – 500 μM) yield the most active and selective hybrid catalysts.^[14] Therefore, these results suggested that protein expression levels in small-scale cultures should be optimised since the application of this purification method is hindered by the amount and final concentrations of protein obtained.

During the course of this project, test experiments were conducted to express streptavidin in the auto-induction medium developed by Studier,^[18] but no recombinant expression was detected (data not shown). A modified medium, based on the same principle of temperature-induction, was tested and yielded large amounts of Sav (~ 130 mg/L), in a small-scale culture (Section 6.2.1). Nonetheless, further experiments using this medium, under different conditions, should be conducted as these preliminary results were promising (Section 6.2).

This purification system was further investigated by screening the biotinylated Rh complex **20** embedded into Sav isoforms against a more challenging enantioselective reaction, such as the reductive amination for the production of unprotected chiral α -amino acids.

Non-enzymatic reductive amination of α -keto acids

Based on *in silico* studies, the library of mutants, K121H-L110x (x = A, E, K, D, and C) and K121H-S112H, was screened for the preparation of phenylalanine **12** using the catalyst **22** (Table 2.6). The proteins were purified via SSP, in the same manner as the library screened for the asymmetric transfer hydrogenation of imines.

The catalysis experiments carried out with the library of double mutants afforded modest to good conversions, but racemic product. The hybrid catalyst **20** \subset K121H-S112H yielded racemic product (entry 8), possibly due to the competitive coordination of the metal by the two histidine residues. Notwithstanding, this screening effort led to the identification of one double mutant, K121H-L110C, which yielded nearly quantitative amounts of (*S*)-phenylalanine with 25% *ee* (entry 7).

To confirm the result of this screening, K121H-L110C was produced in large-scale and purified by conventional affinity chromatography. The new active artificial metalloenzyme, **20** \subset K121H-L110C was tested on the production of phenylalanine **12** under optimised catalytic conditions, by Dr Jeremy Zimbron. Quantitative conversion, 97% yield, 97% selectivity, and 24% *ee* for (*S*)-Phe were obtained, confirming the positive hit of the screening.

Table 2.6. Results for the production of phenylalanine **12** using the biotinylated piano stool complex **22** \subset library of mutants.^a

Entry	Protein	Protein concentration [μ M]	Conv. ^b [%]	Yield ^b [%]	Selectivity ^b [%]	ee ^b [%]
1	S112H	916 ^c	quant.	87	87	rac.
2	K121H	916 ^d	quant.	91	91	24 (<i>S</i>)
3	K121H-L110A	110 ^e	86	73	85	rac.
4	K121H-L110E	180 ^e	83	68	82	rac.
5	K121H-L110K	215 ^e	84	68	81	rac.
6	K121H-L110D	50 ^e	60	44	74	rac.
7	K121H-L110C	90 ^e	91	82	90	25 (<i>S</i>)
8	K121H-S112H	105 ^e	quant.	82	82	rac.

^a The reactions were carried out at 55 °C for 12 h, at pH 8.0 in 4.90 M HCOONH₄. Total reaction volume: 200 μ L. **Conversion:** quantity of α -keto acid transformed; **yield:** quantity of α -amino acid produced; and **(chemo)selectivity:** production of amino acids over α -hydroxy acids.

^b Determined by HPLC after extraction.

^c Using 1 mol% complex **22** vs substrate (680 μ M final metal concentration, 68 mM substrate and 0.33 mol% tetrameric S112H (obtained by conventional affinity chromatography) corresponding to 916 μ M biotin binding sites).

^d Using 1 mol% complex **22** vs substrate (458 μ M final metal concentration, 45.8 mM substrate and 0.5 mol% tetrameric K121H (obtained by conventional affinity chromatography) corresponding to 916 μ M biotin binding sites).

^e Using 2 mol% complex **22** vs substrate and 1 mol% tetrameric streptavidin obtained by small-scale affinity chromatography. Protein concentration determined using a NanoDrop spectrophotometer.

A third approach was developed with the aim of neutralising potential catalyst poisons present in crude protein extracts, thus excluding downstream purification.

2.2.4 Screening on crude protein extracts free of reduced glutathione

Previous studies indicated that the cytosolic pool of reduced glutathione (GSH) present in *E. coli* cells may play a critical role in inhibiting catalysis.^[19] To verify that the activity of the organometallic catalyst was inhibited by GSH present in crude protein extracts, a method to “neutralise” reduced glutathione was developed. Chemical agents such as 1,4-benzoquinone, 1,1'-azobis(*N,N*-dimethylformamide) (*aka* diamide), 4-(2-bromoacetyl)-benzoic acid and diazene, and electrophilic agents such as maleimide and phenyl vinyl sulphone (PVS) were tested. The asymmetric transfer hydrogenation of imines was chosen as model reaction, using the catalyst **4**. To analyse the effect of the six aforementioned agents on GSH, purified Sav S112A was spiked with 5 mM GSH, and treated with the agents at final concentrations ranging from 2.5 to 10 mM (Table 2.7). The purified protein was first incubated for 15 min with GSH, and then treated with the agent, for 24 h at RT. Elapsed the time, addition of the catalyst to the treated sample afforded the hybrid catalyst, and substrate was added to start the reaction. Reactions were incubated for another 24 h, at 30 °C.

Table 2.7. Results for the production of salsolidine **3** by “neutralisation” of reduced glutathione, using the biotinylated piano stool complex **4** \subset Sav S112A (purified/spiked).^a

Entry	“Neutralising” agent	Concentration [mM]	Conv. ^b [%]	ee ^b [%]
1	diamide	0	– ^c	n.d. ^c
2	diamide	2.5	94	57
3	diamide	5.0	95	81
4	diamide	10.0	95	80
5	phenyl vinyl sulphone	0	– ^c	n.d. ^c
6	phenyl vinyl sulphone	2.5	86	72
7	phenyl vinyl sulphone	5.0	82	72
8	phenyl vinyl sulphone	10.0	quant.	57
9	1,4-benzoquinone	0	– ^c	n.d. ^c
10	1,4-benzoquinone	2.5	56	75
11	1,4-benzoquinone	5.0	81	60
12	1,4-benzoquinone	10.0	60	72
13	diazene	0	– ^c	n.d. ^c
14	diazene	2.5	42	62
15	diazene	5.0	38	42
16	diazene	10.0	33	40
17	maleimide	0	– ^c	n.d. ^c
18	maleimide	2.5	37	20
19	maleimide	5.0	47	41
20	maleimide	10.0	18	rac.
21	4-(2-bromoacetyl)-benzoic acid	0	– ^c	n.d. ^c
22	4-(2-bromoacetyl)-benzoic acid	2.5	12	rac.
23	4-(2-bromoacetyl)-benzoic acid	5.0	19	rac.
24	4-(2-bromoacetyl)-benzoic acid	10.0	15	rac.

^a The reactions were carried out at 30 °C for 24 h, using 1 mol% complex **4** (500 μ M final concentration) and 0.33 mol% tetrameric streptavidin at pH 6.5 (MOPS buffer, 0.6 M) containing 3 M HCOONa. Catalyses were performed on purified S112A spiked with 5 mM GSH, and treated with increasing concentrations of chemical or electrophilic agents (0 to 10 mM). Total reaction volume: 200 μ L.

^b Determined by HPLC after extraction.

^c n.d. not determined – peaks were too small to determine conversion or *ee*.

These preliminary reactions confirmed that GSH acted as a catalyst poison. As shown in Table 2.7, in the presence of 5 mM GSH and no agent added (0 mM, entries 1, 5, 9, 13, 17 and 21), no catalytic activity was detected. However, at concentrations as low as 2.5 mM, moderate conversions ($\leq 45\%$) and *ee* ($\leq 50\%$ (*R*)) were obtained for the samples treated with 1,4-benzoquinone, diazene and maleimide (entries 10, 14 and 18). Above 85% conv. and 60% *ee* were obtained for samples treated with diamide and PVS (entries 2 and 6). Activity of the catalyst was not greatly enhanced after the addition of higher concentrations of agent, signifying that there should be a fine balance between the amount of neutralising agent added and the GSH present in the crude protein extracts. In excess, some of these substances are known to be toxic and to have non-specific effects on proteins. 4-(2-bromoacetyl)-benzoic acid had little effect on spiked protein (15% conv. and *ee*), thus it was excluded from further experiments. The same approach was used using Sav S112K as biomolecular scaffold. To further investigate their effect, the list of “neutralising” agents to test was narrowed down to the four best performing ones (Table 2.8).

Table 2.8. Results for the production of salsolidine **3** by “neutralisation” of reduced glutathione, using the biotinylated piano stool complex **4** \subset Sav S112K (purified/spiked).^a

Entry	“Neutralising” agent	Concentration [mM]	Conv. ^b [%]	ee ^b [%]
1	diamide	0	– ^c	n.d. ^c
2	diamide	2.5	89	94
3	diamide	5.0	64	70
4	diamide	10.0	60	71
5	phenyl vinyl sulphone	0	– ^c	n.d. ^c
6	phenyl vinyl sulphone	2.5	37	48
7	phenyl vinyl sulphone	5.0	68	61
8	phenyl vinyl sulphone	10.0	95	27
9	1,4-benzoquinone	0	– ^c	n.d. ^c
10	1,4-benzoquinone	2.5	60	73
11	1,4-benzoquinone	5.0	66	54
12	1,4-benzoquinone	10.0	74	28
13	diazene	0	– ^c	n.d. ^c
14	diazene	2.5	36	53
15	diazene	5.0	34	49
16	diazene	10.0	37	53

^a The reactions were carried out at 30 °C for 24 h, using 1 mol% complex **4** (500 μ M final concentration) and 0.33 mol% tetrameric streptavidin at pH 6.5 (MOPS buffer, 0.4 M) containing 3 M HCOONa. Catalyses were performed on purified S112K spiked with 5 mM GSH, and treated with increasing concentrations of chemical or electrophilic agents (0 to 10 mM). Total reaction volume: 200 μ L.

^b Determined by HPLC after extraction.

^c n.d. not determined – peaks were too small to determine conversion or *ee*.

This second round of experiments confirmed the results obtained with Sav S112A. Diamide was identified as the most promising “neutralising” agent, as it rendered 94% *ee* and 89% conv. at 2.5 mM (entry 2). Similar conversion but lower *ee* were obtained at higher concentration of PVS (10 mM, entry 8), 95% and 27%, respectively. The agent, 1,4-benzoquinone, yielded moderate conversion (30%) and *ee* (50%) at all concentrations tested (entries 9 – 12). Although good conversion and *ee* were obtained with the hybrid catalyst pre-treated with diazene (60% conv. and 60% *ee* (*S*) at 2.5 and 5 mM, entries 14 and 15, respectively), this agent was not used in further experiments as DMSO (final concentration 10% v/v) was required to solubilise it. Control experiments were carried out to assess the effect of DMSO on the catalyst. For S112A, the effect was minimal (97 vs 81% conversion, and 76 vs 75% *ee*, with and without DMSO, respectively), whereas for S112K, the catalyst shown reduced activity upon addition of the organic solvent (93 vs 26% conversion, and 72 vs 65% *ee*, with and without DMSO, respectively).

The correct enantiomer was obtained for both artificial metalloenzymes, *i.e.* (*R*) for S112A and (*S*) for S112K, thus excluding artefact reactions.

On the basis of these preliminary results, the three best agents were tested on crude extracts of Sav S112A and S112K (Table 2.9).

Table 2.9. Results for the production of salsolidine **3** by “neutralisation” of reduced glutathione, using the biotinylated piano stool complex **4** \subset Sav S112A or S112K (crude protein extracts).^a

Entry	Protein	“Neutralising” agent	Concentration [mM]	Conv. ^b [%]	ee ^b [%]
1	S112A	diamide	0	– ^c	n.d. ^c
2	S112A	diamide	2.5	35	45 (<i>R</i>)
3	S112A	diamide	5.0	39	52 (<i>R</i>)
4	S112A	diamide	10.0	31	43 (<i>R</i>)
5	S112A ^d	diamide	0	– ^c	n.d. ^c
6	S112A ^d	diamide	2.5	36	53
7	S112A ^d	diamide	5.0	95	81
8	S112A ^d	diamide	10.0	37	53
9	S112K	diamide	0	– ^c	n.d. ^c
10	S112K	diamide	2.5	57	75 (<i>S</i>)
11	S112K	diamide	5.0	57	75 (<i>S</i>)
12	S112K	diamide	10.0	51	74 (<i>S</i>)
13	S112K ^d	diamide	0	– ^c	n.d. ^c
14	S112K ^d	diamide	2.5	36	53
15	S112K ^d	diamide	5.0	64	69
16	S112K ^d	diamide	10.0	37	53
17	S112A	phenyl vinyl sulphone	0	– ^c	n.d. ^c
18	S112A	phenyl vinyl sulphone	2.5	20	26 (<i>R</i>)
19	S112A	phenyl vinyl sulphone	5.0	21	30 (<i>R</i>)
20	S112A	phenyl vinyl sulphone	10.0	45	41 (<i>R</i>)
21	S112K	phenyl vinyl sulphone	0	– ^c	n.d. ^c
22	S112K	phenyl vinyl sulphone	2.5	– ^c	n.d. ^c
23	S112K	phenyl vinyl sulphone	5.0	– ^c	n.d. ^c
24	S112K	phenyl vinyl sulphone	10.0	49	33 (<i>S</i>)
25	S112A	1,4-benzoquinone	0	– ^c	n.d. ^c
26	S112A	1,4-benzoquinone	2.5	20	39 (<i>R</i>)
27	S112A	1,4-benzoquinone	5.0	23	34 (<i>R</i>)
28	S112A	1,4-benzoquinone	10.0	37	44 (<i>R</i>)
29	S112K	1,4-benzoquinone	0	– ^c	n.d. ^c
30	S112K	1,4-benzoquinone	2.5	55	72 (<i>S</i>)
31	S112K	1,4-benzoquinone	5.0	64	69 (<i>S</i>)
32	S112K	1,4-benzoquinone	10.0	63	57 (<i>S</i>)

^a The reactions were carried out at 30 °C for 24 h, using 1 mol% complex **4** (500 μ M final concentration) and 0.33 mol% tetrameric streptavidin at pH 6.5 (MOPS buffer, 0.4 M) containing 3 M HCOONa. Catalyses were performed on Sav S112A and S112K crude protein extracts, treated with increasing concentrations of chemical or electrophilic agents (0 to 10 mM). Total reaction volume: 200 μ L.

^b Determined by HPLC after extraction.

^c n.d. not determined – peaks were too small to determine conversion or *ee*.

^d Purified, spiked with 5 mM GSH.

The catalysis experiments carried out with crude protein extracts confirmed the applicability of this method to neutralise potential cells contaminants/inhibitors, such as reduced glutathione. The most significant results emerging from this screening process, regarding conversion and enantioselectivity, were achieved with diamide and 1,4-benzoquinone for both metalloenzymes (35% conv. and 45% *ee* for S112A, and 58% conv. and 75% *ee* for S112K treated with 2.5 mM diamide; 21% conv. and 40% *ee* for S112A, and 55% conv. and 72% *ee* for S112K treated with 1,4-benzoquinone). PVS had similar effect on S112A scaffold as

1,4-benzoquinone (*i.e.* it yielded similar conv. and *ee*), but had almost no effect on S112K, except at the highest concentration (10 mM, 49% conv. and 33% *ee*).

The above results suggested that the optimal concentration of agent to be used should be 2.5 or 5 mM, as in most cases, best reaction yields were obtained with these concentrations. An excess of neutralising agent might be harmful to either the organometallic moiety and/or the biomolecular scaffold. At 10 mM, conversion and enantioselectivity tended to decrease 5 to 10% in terms of conversion and *ee*.

This study has proven that pre-treatment of crude protein extracts by chemical agents was an effective and reliable method. In Table 2.10, an overview of the results is given for Sav S112A and S112K, using diamide as “neutralising” agent.

Table 2.10. Summary of results for the production of salsolidine **3** using the biotinylated piano stool complex **4** \subset Sav S112H or S112K (purified, purified/spiked with GSH/treated and untreated, and crude protein extracts/treated and untreated).^a

Entry	Protein	Description	Conv. [%]	<i>ee</i> [%]
1	S112A	purified	99	81 (<i>R</i>)
2	S112A	purified, spiked with 5 mM GSH	– ^b	n.d. ^b
3	S112A	purified, spiked with 5 mM GSH, treated with 5 mM diamide	95	81 (<i>R</i>)
4	S112A	crude protein extract	– ^b	n.d. ^b
5	S112A	crude protein extract, treated with 5 mM diamide	39	52 (<i>R</i>)
6	S112K	purified	93	64 (<i>S</i>)
7	S112K	purified, spiked with 5 mM GSH	– ^b	n.d. ^b
8	S112K	purified, spiked with 5 mM GSH, treated with 5 mM diamide	64	69 (<i>S</i>)
9	S112K	crude protein extract	– ^b	n.d. ^b
10	S112K	crude protein extract, treated with 5 mM diamide	57	75 (<i>S</i>)

^a The reactions were carried out at 30 °C for 24 h using 1 mol% complex **4** (500 μ M final concentration), 5 mM substrate, 100 μ M free binding sites per monomeric Sav, in 0.4 M MOPS buffer (200 μ L total volume) containing 3 M HCOONa, pH 6.5.

^b n.d. not determined – peaks were too small to determine conversion or *ee*.

Although the results were usually inferior to those obtained with purified proteins, particularly in terms of conversion (entries 1, 3 and 5 for S112A; entries 6, 8 and 10, for S112K), such protein hosts, obtainable by a simple one-step “purification”, hold great potential for high-throughput pre-screening of the relative performance of the hybrid catalysts.

2.3 Conclusion & outlook

In the quest to apply Darwinian protocols to artificial metalloenzymes, any group working in the field faces the greatest challenge of producing and purifying, in parallel, hundreds of mutant proteins. Scouting for the best purification routine, three different experimental approaches – protein precipitation, small-scale purification and “neutralisation” of reduced

glutathione – were investigated to “purify” the highest amount of functional product per unit volume per unit time, to screen on artificial metalloenzymes.

The protein precipitation scheme yielded satisfactory results, as moderate conversions ($\leq 60\%$ conv.) and enantioselectivities ($\leq 50\%$ *ee*) were obtained using crude extracts dialysed against guanidinium-chloride and precipitated by addition of ethanol or dimethoxyethane. However, this procedure brought to light the pitfall of the amount of protein present in crude extracts, which was later on confirmed when implementing the small-scale purification strategy. A library of 20 mutants (expressed in 50 mL scale) prepared using both purification approaches was screened for ATH and reductive amination of α -keto acids. The library screened exerted no remarkable effect in activity and selectivity with one exception, K121H-L110C. This double mutant yielded 91% conversion and 25% *ee* for the production of (*S*)-phenylalanine, by reductive amination of α -keto acids. This positive hit was confirmed by producing the isoform in large-scale and testing it under the same reaction conditions, ergo validating the small-scale purification strategy when the concentration of the host protein is suitable for catalysis. Out of the three schemes presented herein, the “neutralisation” of reduced glutathione present in crude extracts by addition of 2.5 mM of diamide (1,1'-azobis(*N,N*-dimethylformamide)) met the most demanding requirements, for high-throughput screening of protein-based hybrid catalysts, with promising conversions ($> 40\%$) and good enantioselectivities ($> 50\%$). To consolidate this “purification” strategy, recombinant protein production in *E. coli* should be maximised, as poor protein concentration in the reaction leads to lower conversions compared to standard conditions.

Successful heterologous protein (over)production in *E. coli* involves many factors, and techniques to optimise them have been thoroughly explored. Reviews by Jana & Deb and Peti & Page summarise these optimisation strategies.^[20,21] Recently, Li *et al.* reported high yield production of triple-labelled and unlabelled proteins from 50 mL bacterial cell cultures. They developed a bacterial expression method that combined tightly controlled traditional isopropyl- β -D-1-thiogalactopyranoside (IPTG) induction expression with high-cell density of auto-induction expression, and routinely produced 14 to 25 mg and 15 to 35 mg of triple-labeled and of unlabeled proteins, respectively.^[22] If high yields can be achieved for the expression of isotopically labelled protein, which is known to typically yield very low amounts of target protein, then this method should be investigated in the production of unlabelled streptavidin to improve protein production in small-scale.

The presented data provided sufficient evidence that the “neutralisation” of reduced glutathione in crude protein extracts can be implemented as a purification method to fast-screen hybrid catalysts, in enantioselective reactions. The overall time of protein production was shortened to five days. High yield production of streptavidin allied with the “neutralisation” of glutathione by addition of diamide would pave the way toward high-throughput screening of artificial metalloenzymes.

2.4 References

- [1] Lu, Y. *Inorganic Chemistry* **2006**, *45*, 9930–9940.
- [2] Steinreiber, J.; Ward, T. R. *Bio-inspired catalysts*; volume 25 Springer Berlin Heidelberg: Berlin, Germany, 2009.
- [3] Pordea, A.; Ward, T. R. *Chemical Communications* **2008**, *36*, 4239–4249.
- [4] Thomas, C. M.; Ward, T. R. *Applied Organometallic Chemistry* **2005**, *19*, 35–39.
- [5] Panek, J. J.; Ward, T. R.; Jezierska-Mazzarello, A.; Novič, M. *Journal of Computer-Aided Molecular Design* **2010**, *24*, 719–732.
- [6] Ogo, S.; Makihara, N.; Kaneko, A.; Watanabe, Y. *Organometallics* **2001**, *20*, 4903–4910.
- [7] Letondor, C.; Humbert, N.; Ward, T. R. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 4683–4687.
- [8] Dürrenberger, M.; Heinisch, T.; Wilson, Y. M.; Rossel, T.; Nogueira, E.; Knörr, L.; Mutschler, A.; Kersten, K.; Zimbron, M. J.; Pierron, J.; Schirmer, T.; Ward, T. R. *Angewandte Chemie International Edition* **2011**, *50*, 3026–3029.
- [9] Zimbron, M. J.; Heinisch, T.; Hamels, D.; Nogueira, E. S.; Schirmer, T.; Ward, T. R. *Journal of the American Chemical Society* **2013**, *135*, 5384–5388.
- [10] Zimbron, M. J. *Engineering artificial metalloenzymes based on biotin-streptavidin technology for DNA recognition and asymmetric transfer hydrogenation catalysis*, Thesis, University of Basel, 2011.
- [11] Ogo, S.; Uehara, K.; Abura, T.; Fukuzumi, S. *Journal of the American Chemical Society* **2004**, *126*, 3020–3021.
- [12] Rusbandi, U. E.; Lo, C.; Skander, M.; Ivanova, A.; Creus, M.; Humbert, N.; Ward, T. R. *Advanced Synthesis & Catalysis* **2007**, *349*, 1923–1930.
- [13] Creus, M.; Pordea, A.; Rossel, T.; Sardo, A.; Letondor, C.; Ivanova, A.; Letrong, I.; Stenkamp, R. E.; Ward, T. R. *Angewandte Chemie International Edition* **2008**, *47*, 1400–1404.
- [14] Köhler, V.; Wilson, Y. M.; Lo, C.; Sardo, A.; Ward, T. R. *Current Opinion in Biotechnology* **2010**, *21*, 744–752.
- [15] Sivaraman, T.; Kumar, T. K. S.; Jayaraman, G.; Yu, C. *Journal of Protein Chemistry* **1997**, *16*, 291–297.
- [16] Bensadoun, A.; Weinstein, D. *Analytical Biochemistry* **1976**, *70*, 241–250.
- [17] Eisenberg-Domovich, Y.; Pazy, Y.; Nir, O.; Raboy, B.; Bayer, E. A.; Wilchek, M.; Livnah, O. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 5916–5921.

- [18] Studier, F. W. *Protein Expression & Purification* **2005**, *41*, 207–234.
- [19] Zimbron, M. Z.; Sardo, A.; Heinisch, T.; Wohlschlager, T.; Gradinaru, J.; Massa, C.; Schirmer, T.; Ward, T. R. *Chemistry A European Journal* **2010**, *16*, 12883–12889.
- [20] Jana, S.; Deb, J. K. *Applied Microbiology and Biotechnology* **2005**, *67*, 289–298.
- [21] Peti, W.; Page, R. *Protein Expression & Purification* **2007**, *51*, 1–10.
- [22] Murray, V.; Huang, Y.; Chen, J.; Wang, J.; Li, Q. *Protein NMR Techniques*; volume 831 of *Methods in Molecular Biology* Springer Science+Business Media: New York, USA, 2012.

New platform for the expression of streptavidin

Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less.

Marie Skłodowska-Curie

Abstract

This Chapter examines the use of the methylotrophic yeast *Pichia pastoris* as host for the production of recombinant mature streptavidin (Sav). The streptavidin gene was inserted into *P. pastoris* expression vector, pPICZ α A, in-frame with *Saccharomyces cerevisiae* α -mating factor (α -MF) secretion signal to extracellularly deliver the target protein. The recombinant plasmid of pPICZ α A-Sav was linearised by *SacI*, and transformed by electroporation into two different strains of *Pichia*, X-33 and KAI-3. Multi-copy insert transformants were selected, and cultivated in shake flasks. Secreted streptavidin, with a higher molecular weight size than Sav expressed in *Escherichia coli* was identified by SDS-PAGE and Western blot. Monomeric Sav displayed a molecular weight of 15.9 kDa, as assessed by ESI mass spectrometry. N-terminal amino acid sequencing indicated the presence of four residues (E-A-E-A), which indicated that the pro-sequence (α -MF) was partially cleaved. In a fed-batch fermentation, Sav was secreted at approximately 650 mg/L of culture supernatant. The secreted mature streptavidin displayed identical properties to streptavidin produced in *E. coli* for the creation of artificial imine reductase upon incorporation of a biotinylated piano stool catalyst.

3.1 Introductory remarks

Streptavidin (Sav) is a ~ 60 (4 x 15) kDa homotetrameric protein, isolated from the bacterium *Streptomyces avidinii*. Like its namesake avidin, streptavidin binds four equivalents of biotin per tetramer, with an affinity virtually unmatched in nature ($K_a \sim 10^{13} \text{ M}^{-1}$).^[1] It has been used for a variety of biochemical applications, *e.g.* immobilisation, cell-surface labelling, or delivery of diagnostic agents.^[2-5] Several homologous high-affinity biotin-binding proteins have been identified from a variety of organisms.^[6-11] In recent years, Sav has been utilised

as a host protein for biotinylated organometallic catalysts. Such hybrid catalysts have shown promising properties for enantioselective reactions.^[12,13] By combining chemical and genetic strategies (chemo-genetic optimisation), both activity and selectivity of the artificial metalloenzyme can be optimised and fine-tuned.^[14] From the biological point-of-view, developing new classes of artificial metalloenzymes requires a major effort, as large amounts of purified protein are needed. Several groups in the field have started to address this issue by investigating the potential of screening on either cells or semi-purified cell lysates.^[15–17] In 2006, Streu and Meggers reported the cleavage of allylcarbamates by [Cp*Ru(COD)Cl] complex, inside mammalian cells.^[15] Another approach was investigated by Reetz and co-workers who applied a simple heat treatment to purify in parallel several variants of the thermostable synthase, tHisF. Although conjugation with ligands, ligand/metal entities, and organocatalysts was achieved, no catalytic results were then reported.^[16] Catalytic results using tHisF were later on published for the asymmetric Diels-Alder reaction of azachalcone and cyclopentadiene (best results obtained with mutant Cu^{II}/HHD-4xala, with 73% conversion and 40% *ee*), using this time purified protein evolved by site-directed mutagenesis.^[18] More recently, Ward and co-workers investigated an alternative methodology, by purifying the protein extracts by ethanol precipitation. Sav mutant, S112A, yielded similar *ee* and conversion, at low catalyst concentration (down to 39 μ M), with both standard (purified by affinity chromatography) and precipitated proteins (64% *ee* (*R*) vs 61% (*S*) and 77% conversion vs 65%, respectively).^[17] Despite these promising results, alternative routes should be investigated, as screening on crude protein extracts or proteins purified by other means than the conventional chromatographic methods lead to *e.g.* lower conversions due to low protein concentrations and inhibition of catalysis by debris present in the cell lysates (DNA, thiols, lipids, and other proteins).^[17,19–21]

3.1.1 From bacterium to yeast

Hitherto, all screening studies conducted in the field of artificial metalloenzymes have been performed with purified proteins, which compromises the applicability of a Darwinistic approach. Therefore it is imperative to develop alternative and reliable methods for the production of functional proteins.

To this end and influenced by the several reasons that account for the rising popularity of *Pichia pastoris* expression system, the cDNA of Sav was cloned into the methylotrophic yeast pPICZ α A expression vector (Figure 3.1). As a eukaryote, *P. pastoris* has the ability to (i) produce correctly folded foreign proteins at high levels, either extracellularly or intracellularly; (ii) stably integrate expression plasmids at specific sites in the *P. pastoris* genome in either single or multiple copies; (iii) grow to a very high cell density in bioreactors; (iv) introduce post-translational modifications; and (v) have one of the strongest and most tightly

regulated eukaryotic promoters for controlled gene expression, the alcohol oxidase I (*AOX1*) promoter.^[22,23]

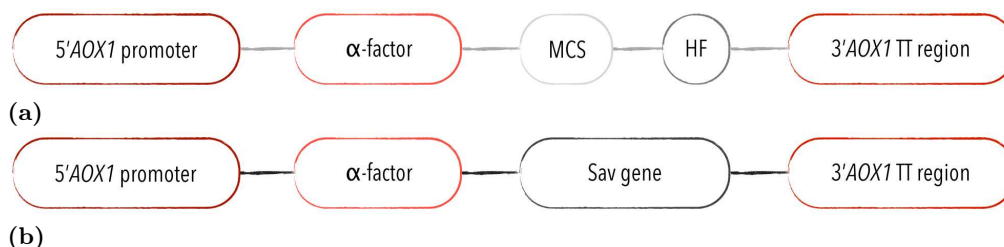


Figure 3.1. Diagrammatic representation of the construct made for the expression of streptavidin in *Pichia pastoris*. From left to right (5' to 3'), the vector contains the 5'-alcohol oxidase 1 promoter (*AOX1*), the α -factor (α -MF, the secretion signal from *S. cerevisiae*), either the multiple cloning site (MCS) or the insert of Sav gene, and finally the 3'*AOX1* TT represents the native *AOX1* transcription termination region. (a) pPICZ α A vector only. (b) Streptavidin insert in *Pichia* expression vector, without the polyhistidine affinity tag factor (HF) addition.

In this work, the vector chosen for expression in *Pichia pastoris* was the pPIZ α A vector from Invitrogen. It is a vector that uses the tightly regulated methanol-inducible *AOX1* promoter, and allows for secretion of the protein to the medium using the α -MF. This makes the subsequent purification simpler, since there is no need to disrupt the yeast cells to access the recombinant protein. Very low levels of endogenous proteins of *Pichia pastoris* are secreted to the medium, thus recombinant proteins expressed with an export signal will constitute the majority of the proteins in the culture supernatant. *Pichia pastoris* X-33 (Invitrogen) and KAI-3 (donated by Prof. Callewaert, University of Ghent, Belgium) were used as platform strains. Both strains are methanol utilisation plus. The *Pichia* strain KAI-3 was genetically engineered such that the genomic OCH1 (α -1,6-mannosyltransferase *Och1p* gene) is inactivated eliminating hypermannosylation, and an α -1,2-mannosidase is over-expressed to trim mannose residues from eight to a total of five, thereby better resembling the glycosylation structures of the human counterpart.^[24]

3.1.2 Research project

Several technical problems need to be solved before a Darwinian approach to enantioselective catalysis using artificial metalloenzymes can be put into practice. First, a very efficient expression system of the host protein is required to provide enough protein for subsequent purification, conjugation with introduction of a ligand/metal moiety and catalysis, since screening new ligands requires large quantities of the polypeptide.^[16] Second, because complexation between the biomolecular host and the organometallic catalyst can also occur with foreign proteins or be hampered by contaminants, an efficient separation of the host from other proteins present in crude extracts or supernatant prior to screening is imperative.^[31] To date, streptavidin has been recombinantly expressed in *E. coli*^[25-27] and in *B. subtilis* cells,^[28,29]

with acceptable yields for *E. coli* system (230 mg/L) and deceiving ones for *Bacillus system* (20 to 90 mg/L). More recently, functional core streptavidin has been expressed at high-levels in *Pichia pastoris* (4.0 g/L), using immobilised yeast cells, 2.0 M glycerol, a feeding flow rate of $0.11 \text{ mL} \cdot \text{min}^{-1}$, and aeration by air injection dispersed with a porous stone combined with agitation at 500 rpm.^[30]

Aiming to address these fundamental problems, the expression system *Pichia pastoris* was investigated to increase production of recombinant wild-type mature streptavidin and simplify the purification process.

3.2 Results & discussion

To differentiate between streptavidin expressed in *E. coli* and in *P. pastoris*, the latter was named Sav918, following an internal code from the Paul Scherrer Institute (Villigen PSI Ost, Switzerland), where this work was carried out.

3.2.1 Strain and genetic construct

The gene encoding a recombinant mature streptavidin (Sav), without the T7-tag, was successfully PCR amplified using synthetic oligonucleotides. Two stop codons (TAG and TGA) were added to the sequence, in order to express Sav without the C-terminal peptide containing the *c-myc* epitope and the hexahistidine tag (6xHis).

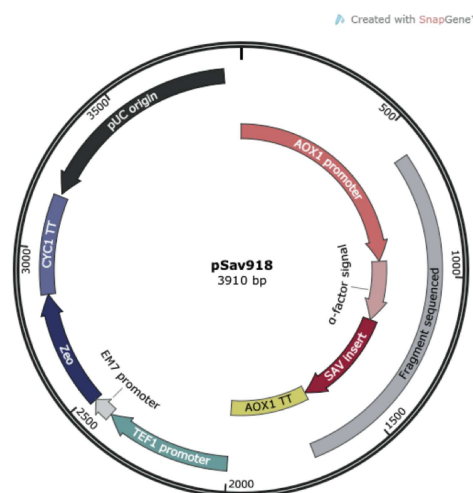


Figure 3.2. Plasmid map of streptavidin gene in pPICZ α vector (pSav918). The 3'- and 5'-nucleotide sequences (fragment sequenced, in grey) confirmed the insertion of Sav gene (dark pink) in frame with the pro-peptide of *S. cerevisiae* mating α -factor (in light pink) and downstream of the *AOX1* promoter (in orange).

The resulting 478 base pairs (bp) insert, encoding the proteolytic signals for *KEX2* and *STE13* cleavage fused to the mature streptavidin, was joined with the multiple cloning site region,

downstream of *Saccharomyces cerevisiae* α -mating factor secretion signal of the expression vector pPICZ α A (Figure 3.2 and Figure 3.3).

```

ATG [ $\alpha$ -factor] AAAAGAGAGGCTGAAGCTCGGGATCAGGCCGGCATCACCGGCACCTGGTACAAC
M           K R  $\blacktriangle$  E A E A  $\blacktriangle$  R D Q A G I T G T W Y N 23
                KEX2           STE13
CAGCTCGGCTCGACCTTCATCGTGACCGCGGGCGCCGACGGCGCCCTGACCGGAACCTACGAGTCGGCC
Q L G S T F I V T A G A D G A L T G T Y E S A 46
GTCGGCAACGCCGAGAGCCGCTACGTCTGACCGGTCGTTACGACAGCGCCCGGCCACCGACGGCAGC
V G N A E S R Y V L T G R Y D S A P A T D G S 69
GGCACCGCCCTCGGTTGGACGGTGGCCTGGAAGAATAACTACCGCAACGCCCACTCCGCGACACGTGG
G T A L G W T V A W K N N Y R N A H S A T T W 92
ACCGCCCAGTACGTCGGCGGCGCCGAGGCGAGGATCAACACCCAGTGGTGTCTACCTCCGGCCACCC
S G Q Y V G G A E A R I N T Q W L L T S G T T 115
GAGGCCAACGCCTGGAAGTCCACGCTGGTCGGCCACGACACCTTCACCAAGGTGAAGCCGTCCGCGCC
E A N A W K S T L V G H D T F T K V K P S A A 138
TCCATCGACGCGCGAAGAAGCGCGCTCAACAACGCCAACCGCTCGACCGCGTTCAGCAGTAGTGA
S I D A A K K A G V N N G N P L D A V Q Q * * 159

```

Figure 3.3. Nucleotide sequence of the artificial streptavidin reading frame (Sav918) and its translation. The vector-derived sequences are in italic, and in square brackets, the yeast α -factor part. The proteolytic cleavage sites, *KEX2* and *STE13*, are indicated by triangles. No consensus sites for *N*-glycosylation are present. In bold, the 30 potential *O*-glycosylation sites (serine or threonine residues). Asterisks represent the tandem stop codons, which are used to prevent read-through. The residues numbering of Sav expressed in *E. coli* was kept for ease of comparison.

The DNA sequence analysis (GATC, Germany) confirmed the presence of a single open reading frame coding for a translation product of 237 amino acids (aa), consisting of the yeast α -factor peptide signal (89 aa) and streptavidin (148 aa, Figure 3.4 and Figure 3.3).

```

Sav          MASMTGGQQMGRDQAGITGTWYNQLGSTFIVTAGADGALTGTYESAVGNAESRYVLTGRY 60
Sav918      -----RDQAGITGTWYNQLGSTFIVTAGADGALTGTYESAVGNAESRYVLTGRY 49
                *****

Sav          DSAPATDGSGTALGWTVAWKNNYRNAHSATTWSGQYVGGAEARINTQWLLTSGTTEANAW 120
Sav918      DSAPATDGSGTALGWTVAWKNNYRNAHSATTWSGQYVGGAEARINTQWLLTSGTTEANAW 109
                *****

Sav          KSTLVGHDTFTKVKPSAASIDAANKAGVNNGNPLDAVQQ 159
Sav918      KSTLVGHDTFTKVKPSAASIDAANKAGVNNGNPLDAVQQ 148
                *****

```

Figure 3.4. Alignment of the sequences of Sav918 (in pPICZ α A) and Sav (in pET11b), confirming the complete removal of the N-terminal peptide T7-tag (first eleven amino acids) [CLUSTAL 2.1 multiple sequence alignment].

Additionally, to confirm the successful construction of pSav918, a restriction analysis using five different restriction enzymes (Eco88I, EcoO109I, HindII, HinfI, and NaeI) was carried out (data not shown). The pPICZ α A vector has a single *SacI* restriction site in the AOX1 locus

that permits linearisation of the vector for efficient integration into the host's 5'-*AOX1* region (Figure 3.5). Hence, the linearised pSav918 was introduced by electroporation into two *Pichia* strains, X-33 (Invitrogen) and KAI-3^[24] (phenotype: Mut⁺). Positive yeast transformants were first screened on YPDS plates, based on the Zeocin resistance conferred by the vector, supplemented or not with biotin (Figure 3.6).

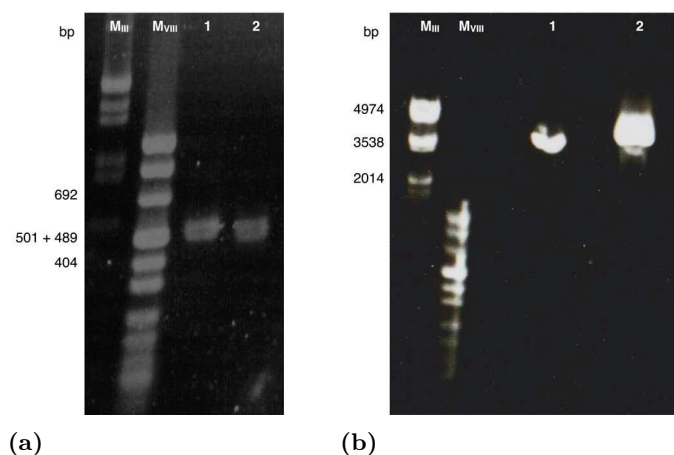


Figure 3.5. Thin agarose gels. (a) PCR product of the amplification of Sav gene, 478 bp (in duplicate, lanes 1 and 2). (b) Linearisation of Sav plasmid (pSav918, 3,910 bp). Lane 1: non-linearised construct, and lane 2: linearised construct by *SacI*. Independent run gel documented in 0.7% thin agarose gel in 0.5% SB buffer. Marker III (*EcoRI* + *HindIII*) and Marker VIII (Böhringer).

Selection of Zeocin resistant clones, which potentially carry multiple plasmid copies, was successful in both KAI-3 and X-33 *Pichia* strains. Several dozens of colonies were obtained in plates inoculated with 100 μ L of cells (Figure 3.6), and several hundred on plates with 300 μ L. Five clones of each construct (X-33 and KAI-3 strains), and from plates containing or not biotin were selected for growth on methanol medium, in small-scale (50 mL cultures), in a total of 35 clones. Clones were numbered from 1 to 20 and from A to J for KAI-3 and X-33 transformants, respectively. This numeration was used throughout this Chapter.

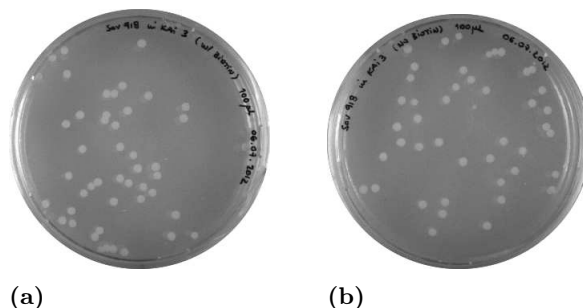


Figure 3.6. YPDS selective plates of pSav918 introduced into the KAI-3 strain. (a) With 0.2% biotin. (b) Without biotin. 100 μ L of electroporated cells were spread on both plates.

The sequence numbering of Sav918 differs from the one from Sav (expressed in *E. coli*), since

the first 11 amino acid residues referring to the T7-tag were removed. Therefore, residue 1 in Sav918 corresponds to residue 12 in Sav, and so on. For ease of comparison, the numbering of Sav expressed in *E. coli* was kept throughout this Chapter.

3.2.2 Expression and detection of streptavidin in *Pichia pastoris*

Recombinant Sav918 was expressed *in vitro* using pPICZ α A expression vector and *P. pastoris* X-33 and KAI-3 expression strains. A screening to find clones that produced recombinant protein at high level was performed in shake flasks. The recombinant protein was secreted into the culture medium, after induction by methanol (final concentration: 0.5% per 24 h). The inoculum culture (buffered minimal glycerol medium – BMGY, pH 6.0) was supplemented with 0.2% biotin, and cells were washed twice with ultrapure water to remove the excess of biotin, before resuspension in culture medium (buffered minimal methanol medium – BMMY, pH 6.0). The cultures were initially grown for 72 to 96 hours at 30 °C. Samples were collected directly from each of the cultures, at different time points (0, 3, 6, 9, 12, 24, 36, 48, 56, 72, and 96 h post induction, Figure 3.8).

Stable transformants were readily obtained by positive selection for vector-induced antibiotic resistance for both strains tested. Out of the 35 clones screened (data not shown), 20 isoforms were secreted as soluble proteins into the culture medium, all in functional form. Furthermore, production levels of these 20 isoforms were found to be variable, depending on the transformant, which further underscores the necessity of screening multiple clones for expression of functional protein (Figure 3.7 and Figure 3.8).

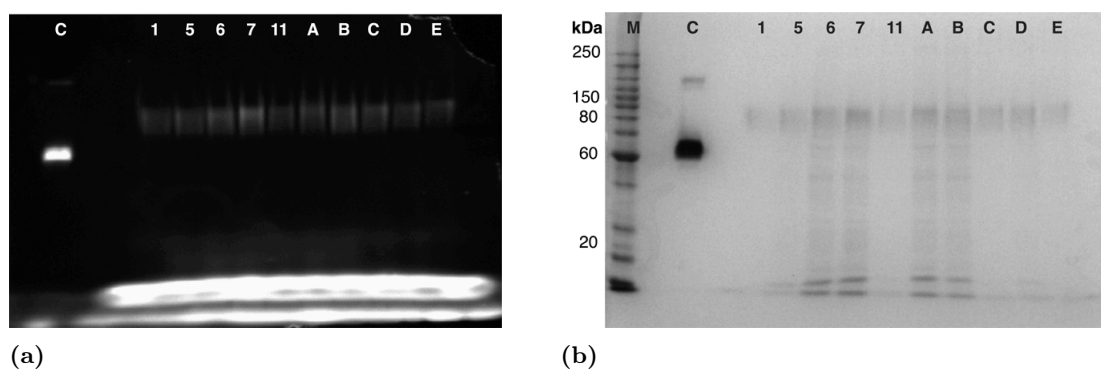


Figure 3.7. SDS-PAGE of the non-concentrated supernatant of ten clones screened for expression of streptavidin, after 96 h of induction time. Streptavidin wild-type expressed in *E. coli* (1 mg/mL sample) was used as positive control; the high molecular weight band corresponds to a oligomeric form of the protein (a) B4F analysis. (b) Coomassie Blue staining analysis. An equal volume (15 μ L) of samples was loaded on the gel.

To confirm the presence of Sav918, SDS-PAGE was performed and gels were revealed by B4F before being stained with silver (for lower expression levels, 1 – 10 ng), or Coomassie Blue (for higher levels of expression, 50 – 100 ng, Figure 3.7.b).

As displayed in Figure 3.7, a recombinant protein with an apparent molecular weight of ~ 80 kDa was observed. The difference of migration of Sav secreted by *P. pastoris* or expressed in *E. coli*, in its tetrameric form, is further discussed in 3.2.6. relates to the fact that Sav is not fully denatured by SDS, as highlighted by its binding to B4F (Figure 3.7.a). The upper band above 100 kDa on the control band (lane C, Figure 3.7) represent an oligomeric form of Sav. No streptavidin was found in the cell pellets as assessed by B4F. The clones that revealed high-level of expression were used for further expressions under different conditions, and allowed a first characterisation of recombinant streptavidin.

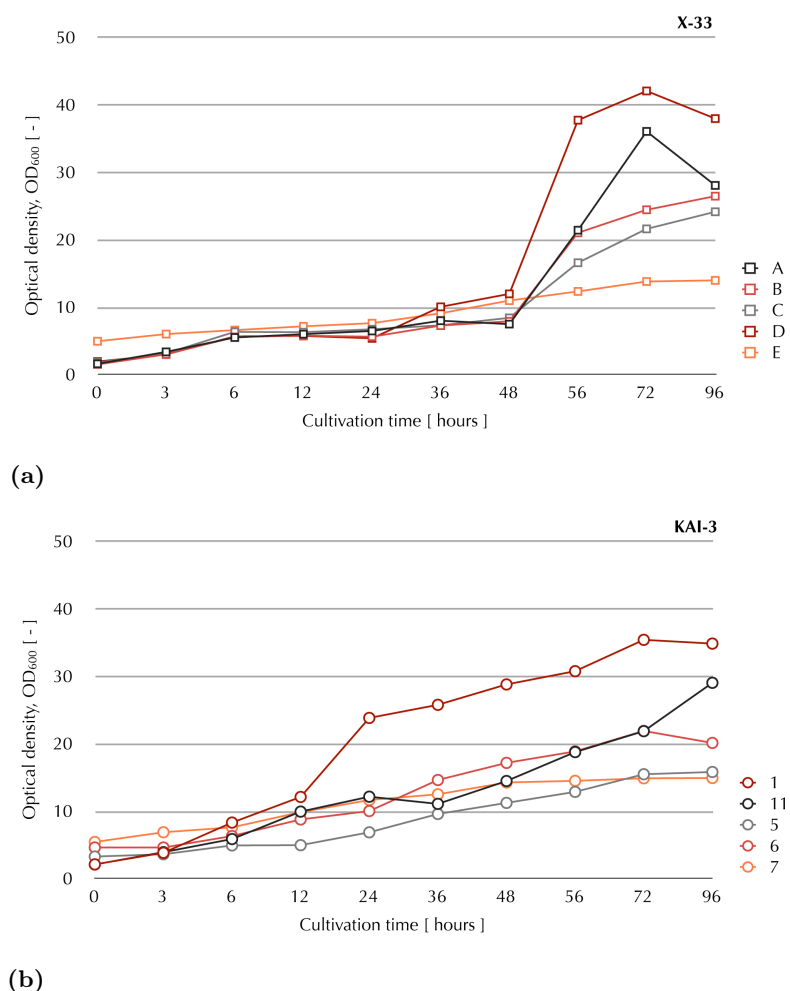


Figure 3.8. Representative growth curves for *Pichia pastoris* strains, KAI-3 and X-33, grown at 30 °C in small-scale cultures, following methanol induction. For ease of reading, only five clones are plotted. **(a)** X-33 clones. **(b)** KAI-3 clones. Cells were initially grown in BMGY prior to transfer to BMMY to enable methanol induction of the *AOX1* promoter. Cultures were incubated at 30 °C. Aerated at 225 rpm, and supplemented with 0.5% (v/v) MeOH every 24 h. Optical density (OD_{600}) readings were taken at different time points, in duplicate. OD levels represented are the mean value of the two readings. The numbering of clones reports to an internal numbering system.

Pichia growth was observed by increasing optical density (OD₆₀₀) values during induction, which indicated viability of the cells and consumption of the inducing agent, methanol (MeOH) (Figure 3.8). This, in turn, indicates induction of the *AOX1* promoter by which the gene of interest is driven. Each growth curve of KAI-3 and X-33 showed an increase in cell density over time, which was generally followed by a plateau after 72 hours. Streptavidin was detected 24 h after induction. The final cell densities ranged in OD₆₀₀ values from 11 to 38 for both strains.

From the 20 positive isoforms, six transformants that showed the highest expression level were chosen (clones 1, 7 and 11 from KAI-3, and A, C and D from X-33) for further cultivation studies.

3.2.3 Expression under different conditions

Given that the main aim of optimisation is to maximise the protein production, this process was only initiated once positive clones were identified. Variables that could affect the growth rate and productivity of *Pichia pastoris* cells or the quality of the recombinant protein were defined. Five factors: pH (unbuffered, and buffered to 5.0 or 6.0), addition of casamino acids (final concentration: 1% w/v) and biotin (final concentration: 0.2% w/v), concentration of methanol for induction (final concentration: 0.5 vs 1% v/v), and volume of the culture (200 vs 600 mL) and their effects, alone and/or in combination, were investigated. Temperature and pH are two of the most crucial parameters that influence protein production in various expression systems. Working with *Pichia pastoris* at low pH allows protease degradation to be avoided, and low temperature during the methanol induction phase can result in an increase in the production of recombinant protein. The effect of temperature, albeit important, was not investigated. Another important factor is the methanol concentration as *Pichia pastoris* metabolises methanol as a sole carbon and energy source, and the promoter regulating the production of alcohol oxidase is the one used to drive heterologous protein expression in *P. pastoris*. Thus, high levels of methanol can be toxic to the cells, and low level may not be enough to initiate transcription. The other parameters listed above were chosen to either increase the final yield or avoid proteolytic degradation.^[32-34]

Initial experiments of this study were carried out in BMGY supplemented with 0.2% (w/v) biotin, and BMMY, at pH 6.0 and 30 °C (Figure 3.9, Exp1). Pure methanol (final concentration 0.5%, v/v) was added to the medium twice daily, instead of once, in order to decrease concentration shifts in BMMY medium. The effect of biotin in the inoculum and culture medium was investigated. When no biotin was added to the inoculum, the biomass and optical density at 600 nm decreased greatly (Figure 3.9, Exp2). Therefore, biotin was added in excess (0.2%, v/v) to provide enough vitamin to the cells during biomass generation (inoculum). The cells were then washed twice with ultrapure water to remove the excess

of biotin prior to induction in methanol, this way preventing the binding of streptavidin to biotin and ensuring the expression of biotin-free streptavidin. The addition of casamino acids (1%, w/v) used to prevent proteolytic degradation had no significant effect on biomass nor optical density (Figure 3.9, Exp5, pH 6.0; Exp6, pH 5.0). Optimum growth was observed at pH 5.0 and unbuffered medium (Figure 3.9, Exp3, pH 5.0; Exp4, unbuffered). The final concentration of methanol in the medium had little to no effect on expression of Sav (data not shown). A 2- and 2.5-fold differences in OD₆₀₀ and biomass, respectively, were noted between using 10% and 30% of the volume capacity of the 2000 mL baffled shake flask (Figure 3.9, Exp7, 200 mL; Exp8, 600 mL).

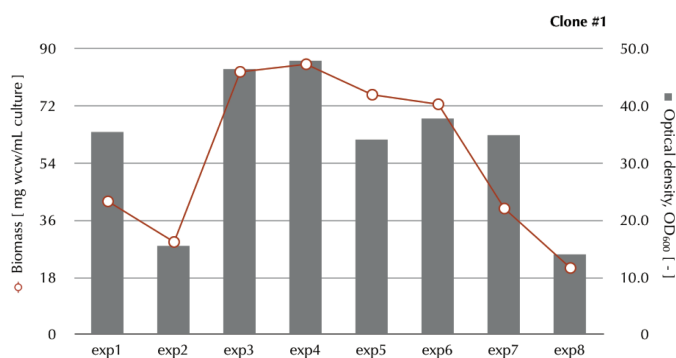


Figure 3.9. Representative growth curves (biomass *vs* optical density) for *Pichia pastoris*, under different cultivation conditions. The biomass and optical density (at 600 nm) report to final values, at the end of induction time (72 h). (Exp1) standard conditions: 0.2% (v/v) biotin in BMGY, 0.5% (v/v) MeOH in BMMY, at 30 °C, pH 6.0; (Exp2) no biotin in BMGY; (Exp3) pH 5.0; (Exp4) unbuffered; (Exp5) pH 6.0 + 1% casamino acids (w/v); (Exp6) pH 5.0 + 1% casamino acids (w/v); (Exp7) volume of culture: 200 mL, 10% of total volume capacity; (Exp8) volume of culture: 600 mL, 30% of total volume capacity. wcv = wet cell weight.

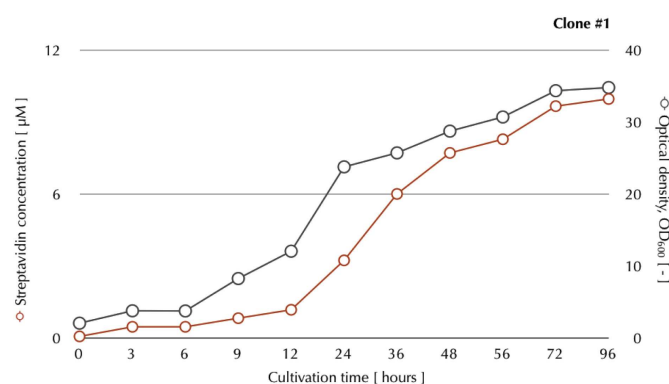


Figure 3.10. Cell density (OD₆₀₀) and expression pattern of streptavidin in the culture medium with *P. pastoris* KAI-3. The culture was carried out in BMMY, pH 5.0, at 30 °C for 72 h. Sav concentration was determined by fluorescence quenching with biotin-4-fluorescein as described in Section 6.2.4

After a series of experiments, optimal expression conditions for streptavidin were obtained as follows: optimal pH was 5.0 (Figure 3.10), optimal induction time points was on the third day, and methanol daily addition concentration of 0.5% (v/v). Under these conditions, high-level expression transformant of *P. pastoris* strain was obtained, purified and retained for further studies. Yields up to 50 mg/L and 150 mg/L were achieved from a 50 mL and a 200 mL culture, respectively.

In a first approach, some of the well-known avenues for the optimisation of protein production were assessed, allowing a rational optimisation. In a second approach, the production process should be further characterised and validated by some controls, in order to achieve a more comprehensive model and to demonstrate the robustness of the procedure, its reproducibility and reliability.

3.2.4 High-cell density fed-batch fermentation

One of the highest expression transformant (clone 1) was chosen for upscaled protein production in a 1.8 L working-volume fermentor. Long-limiting glycerol batch phase (~98 h) under oxygen-sufficient conditions not only increased volumetric productivity but also reduced the need of methanol and oxygen. When dissolved oxygen (dO_2) increased abruptly indicating that glycerol was exhausted, the glycerol feed was stopped for a starving phase of 1.5 h. The methanol feed started when the culture reached OD_{600} of 200, at pH 6.0, 30 °C, and dO_2 set at 20%. After induction with methanol, pH was set to 5.0 to inhibit proteases and samples were withdrawn every 12 h for SDS-PAGE analysis. Samples taken during growth on glycerol (0 to 96 h), lacked expression of streptavidin, which reflected repression of the *AOX1* promoter after substrate limited growth during the glycerol fed-batch phase, *i.e.* during biomass generation. Streptavidin expression was initiated upon change of the carbon source, from glycerol to methanol. The dry cell weight (dcw) increased from 54 g/L after the glycerol fed-batch phase to 235 g/L at the end of the methanol induction. The recombinant streptavidin secreted into the medium by *P. pastoris* continuously increased with methanol induction, and reached the peak at 50 h after induction. After 64 h of induction, an OD_{600} of 870 was reached, and the culture was so dense (more like a paste) that oxygen transfer was inefficient. The fermentation was stopped and the broth was harvested.

In contrast to small-scale cultures, large-scale fermentation in 2.7 L fermentor yielded higher production over a shorter fermentation time (150 mg/L in 72 h compared to 650 mg/L in 64 h). SDS-PAGE analysis of the culture supernatant of Sav918 indicated a major protein smeared band at a molecular weight of ~80 kDa (Figure 3.11). To undoubtedly identify the protein, a Western blot (WB) analysis was carried out. The WB showed a positive reaction of the Sav918 supernatant with polyclonal antibody against streptavidin, and this immunoreaction

became stronger after concentrating the sample (Figure 3.11.c).

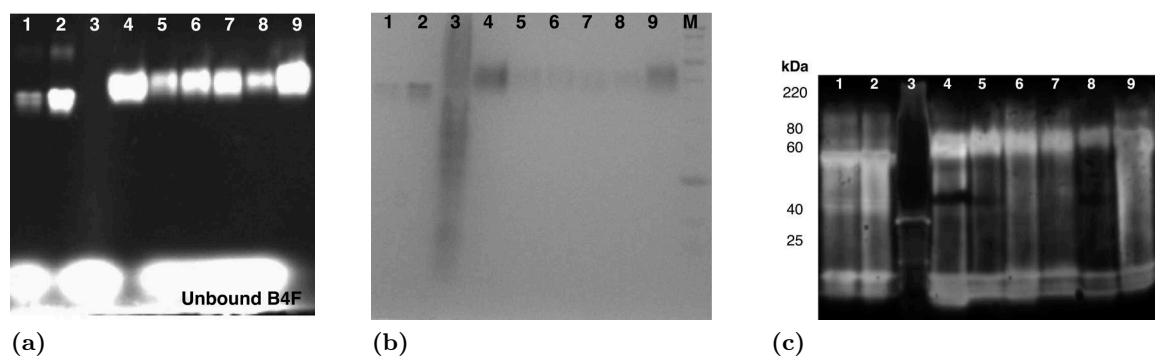


Figure 3.11. 10% SDS-PAGE and Western blot analysis of the fermentation of *P. pastoris* clone 1_Sav918. (a) B4F analysis. (b) Coomassie Blue staining. (c) Western blot analysis. Lanes 1 and 2: control, Sav expressed in *E. coli*, at 200 ng and 1 μg /well, respectively; Lane 3: supernatant, at time = 48 h (GY fed-batch); Lane 4: Sav918 concentrated, before purification; Lanes 5, 6 and 7: supernatants, at time = 100, 120 and 164 h, respectively (MeOH fed-batch); Lanes 8 and 9: 200 ng and 1 μg /well of purified Sav918. An equal volume (10 μL) of samples was loaded on the gel.

3.2.5 Purification of streptavidin expressed in *Pichia pastoris*

The recombinant protein was purified to homogeneity using a single affinity chromatography purification step (2-iminobiotin sepharose). The procedure prior purification for the recombinant Sav918 expressed in a fermentor with defined medium was the same as that employing complex medium in small-scale cultures. After cultivation, the broth was centrifuged, filtered and the supernatants of each clone were prepared for affinity chromatography by gravimetry for small-scale culture, and using an ÄKTA purifier for larger volumes, *i.e.* fermentation.

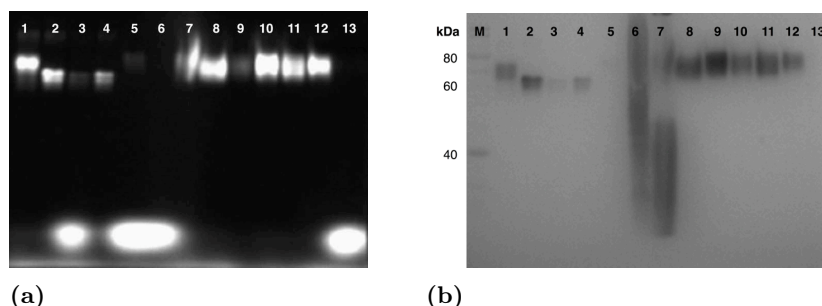


Figure 3.12. SDS-PAGE analysis of Sav918 after column purification. (a) B4F analysis. (b) Coomassie blue staining. Lane 1: supernatant at time = 164 h; Lanes 2 and 4: controls, Sav expressed in *E. coli*, 1 and 0.5 mg/mL, respectively; Lane 3: empty; Lane 5: flow-through; Lane 6: supernatant at time = 45 h; Lane 7: supernatant at time = 100 h; Lane 8: Sav918 before purification; Lanes 9, 10 and 11: pooled fractions of the third, second and first purifications, respectively; Lane 12: pooled fractions of gel filtration; Lane 13: waste. An equal volume (15 μL) of samples was loaded on the gel.

The crude Sav918 samples were concentrated by ultrafiltration, and subsequently equilibrated to pH 9.8 with the appropriate binding buffer, prior to their application to the chromatographic column. After chromatographic purification, the purity of proteins were analysed by SDS-PAGE (Figure 3.12).

The final waste fraction was analysed by SDS-PAGE (Lane 13 in Figure 3.12), along with the pooled fractions, to confirm that none of the recombinant protein remained unpurified. Overall, over 95% of the protein was recovered and purified. After visualisation with B4F and Coomassie Blue staining, the gel did show strong bands that indicated the presence of functional streptavidin (Figure 3.12).

Following these processes, a total of 975 mg purified Sav918 from 1.5 L fermentation broth was obtained, which corresponded to $\sim 11 \mu\text{M}$ of Sav918.

3.2.6 Biochemical properties

Glycosylation analysis

The molecular weight of Sav918 produced in *Pichia Pastoris* was greater than that expressed in *E. coli*. In complex medium (small-scale cultures), a broad and diffuse band between 60 and 80 kDa derived from SDS-PAGE analysis of crude and purified Sav918. This could suggest that the recombinant protein was *O*-glycosylated, since there was no putative *N*-glycosylation site but 30 potential *O*-glycosylation sites instead, in the deduced amino acid sequence of streptavidin (GlycoMod tool, Expasy).^[35] When glycosylated, the protein band is smeared due to the polydisperse nature of the glycan chains.

To determine whether the protein was glycosylated or not, a basic, simple and chromogenic method was used.^[36] Sav918 was resolved on a 12% SDS-PAGE, and the gel was stained for glycoproteins, using the periodic acid-Schiff (PAS) reaction, which reveals glycosylated proteins by magenta bands that begin to appear during the staining reaction and slowly intensify thereafter.

All three clones, expressed in small-scale (*i.e.* in complex medium), tested positive for *O*-glycosylation, as a faint magenta band was detected (Figure 3.13). Proteins with “poorer” glycan trees and fewer glycosylation sites have bands of low staining intensity when compared with heavily glycosylated proteins.^[37] However, Sav918 expressed in minimal medium (*i.e.* in a fermentor) appeared to not be glycosylated. Although bands were of higher molecular weight on the gels as the samples from shake flasks, bands were not smeared, and MS analysis confirmed that Sav918 expressed in a fermentor was not glycosylated (Figure 3.14). The difference of migration of tetrameric Sav secreted by *P. pastoris* or expressed in *E. coli* relates

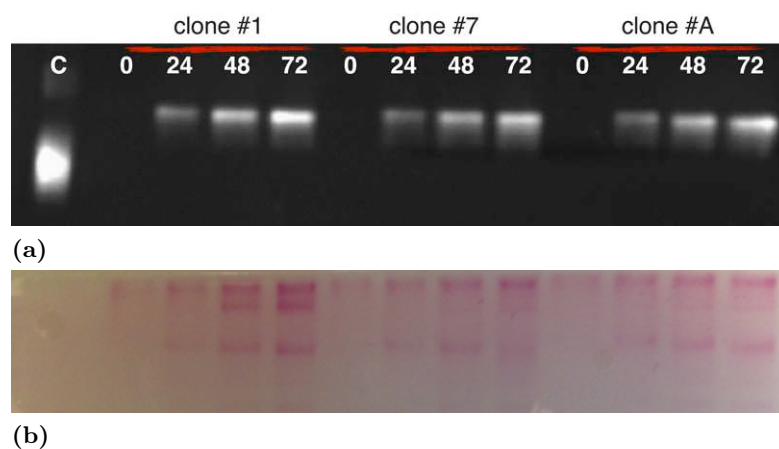


Figure 3.13. SDS-PAGE analysis of Sav918 *O*-glycosylation by periodic acid-Schiff (PAS) reaction. The gel was oxidised in a 1:3 (%) periodic acid:acetic acid solution, and reacted with Schiff's reagent, acidified fuchsin until a magenta colour developed. The gel was then reduced with 1% sodium metabisulfite. **(a)** B4F analysis. **(b)** PAS staining. The supernatant of three different clones were tested for glycosylation: two clones from KAI-3 strain (clones 1 and 7) and clone from X-33 (clone A). All three clones were positive on the test, by revealing magenta bands upon staining. Sav expressed in *E. coli* was used as negative control. An equal volume (15 μ L) of samples was loaded on the gel.

to the fact that Sav is not fully denatured by SDS, as highlighted by its binding to B4F (Figure 3.14). Therefore, folded Sav binds less SDS than it would in its denatured state. Less negative charge translates to slower migration in the gel. The higher mobility of the Sav from *E. coli* is probably due to the T7-tag, which can bind SDS. Recently, Gamboa-Suasnavart *et al.* have reported on the production of recombinant APA protein from *Mycobacterium tuberculosis* in *Streptomyces lividans*, and showed that *O*-mannosylation was affected by culture conditions in shake flasks. Depending on culture conditions (*e.g.* shear and oxygenation), carbohydrate composition differed from two mannose residues in conventional shake flasks to up to five mannose residues in coiled and baffled flasks.^[38] However, little is known about the influence of culture medium on oligosaccharide structures of *P. pastoris*-secreted proteins. Literature reports on the detection of lesser amounts of phosphorylated glycans in recombinant proteins grown in salts medium, in comparison with proteins obtained from complex medium. Neutral oligomannosides predominated in the glycosylation pattern.^[39–41] The reason why and how carbon sources modify glycosylation should be further studied.^[39,40]

Mass spectrometry analysis

Electron spray mass spectrometry analysis of the purified streptavidin detected a single component of 15,878 Da (Figure 3.14). This mass is precisely that predicted for the fully protonated encoded protein without N-terminal methionine (15,878.3 Da), taking into account the two protonated histidine residues. Importantly, this spectrum allowed the exclusion of any glycosylation on Sav. The complete report of the characterisation of Sav918 can be found in

the Appendices Section.

The N-terminal amino acid sequence was determined for the expressed Sav918 as E-A-E-A-R-D-Q-A, which meant that the proteolytic site *STE13* was retained.^[42] In *Pichia* pPICZ α A vector, the effective cleavage of the fusion protein formed by the α -factor signal and the recombinant protein sequences occurs in two steps: (i) the signal sequence Glu-Lys-Arg*Glu-Ala-Glu-Ala is preliminary cleaved by the *KEX2* gene product, between the arginine and the glutamine (site of cleavage marked with *); (ii) the Glu-Ala repeats are further cleaved by the *STE13* gene product.

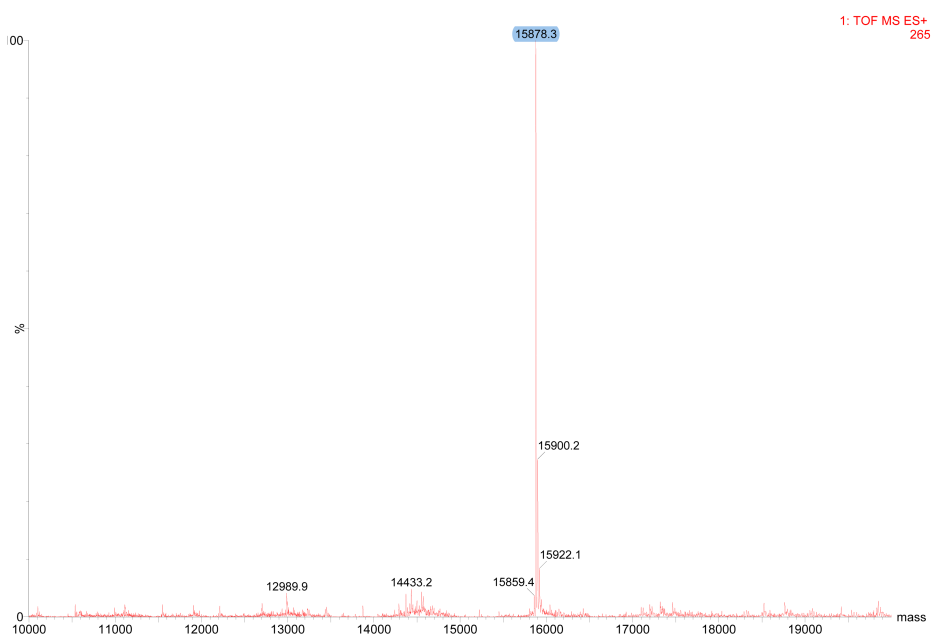


Figure 3.14. Positive ESI-MS spectrum of Sav918 in 1% acetic acid. Calculated mass of the protein: 15878.3 Da. Calculated mass of the main peak: 15878.3

There are some cases where *STE13* cleavage is inefficient, and Glu-Ala repeats are left on the N-terminal of the expressed protein of interest. This is generally dependent on the protein of interest, when the protein fold may impair the recognition and proteolytic cleavage by *STE13* or the amount of protein expressed exceeds the catalytic capacity of the protease.^[43] The complete report of the N-terminal sequencing can be found in the Appendices Section.

Biotin-binding analysis

The functionality of Sav918 and biotin-free binding sites were confirmed by titrating crude and purified proteins with biotin-4-fluorescein (B4F, Figure 3.15).

The titration with B4F confirmed that the presence of residual amounts of biotin in the

PTM1 solution added to the fermentation broth did not hindered the biotin-binding capacity of streptavidin produced in *Pichia pastoris*. Previous studies by Jungo *et al.* have shown that addition of 20 $\mu\text{g/L}$ of biotin (for a cell density of 8 g/L) resulted in stable chemostat cultures on methanol, with the production of recombinant biotin-free avidin.^[33] In this case, the final concentration of biotin in the culture medium was minimal (below 5 μmol). For the example of B4F titration given in Figure 3.15, the number of free-biotin binding sites (fbs) for purified Sav918 was 3.2, hence 0.8 binding sites might have been “blocked” by the biotin present in the medium.

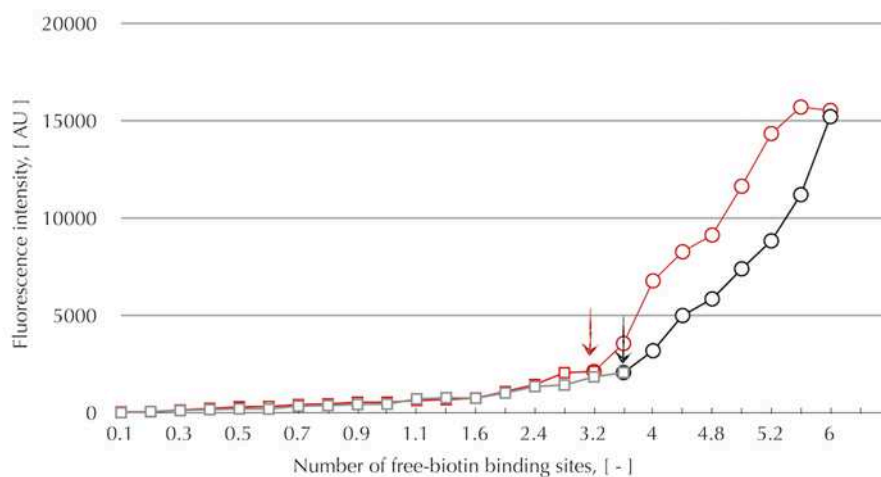


Figure 3.15. Determination of the biotin free-binding sites (fbs) in streptavidin expressed in *E. coli* ($MW_{tetramer} = 63,150$ Da) and in *P. pastoris* ($MW_{tetramer} = 63,510$ Da). The biotin-binding activity for recombinant Sav was detected by fluorescence quenching. Titration of 100 μM of purified samples of Sav produced in a fermentor. Sav control: Sav expressed in *E. coli* (grey lines); Sav918: Sav expressed in *P. pastoris* (red lines) The arrows indicate the equivalence point, 3.2 for Sav918 (red) and 3.6 for Sav control (grey).

Nevertheless, the protein obtained via fermentation was tested as biomolecular host for Cp*Ir moiety for the reduction of 1-methyl-6,7-dimethoxy-3,4-dihydroisoquinoline (Section 6.2.6). The biotinylated catalyst was incorporated into Sav918 (crude and purified forms), and tested in the asymmetric transfer hydrogenation of imines, using the catalytic conditions described in Scheme 3.1.

Considering that the fermentation medium does not correspond to the catalytic conditions, a buffer exchange against the catalysis buffer was performed on the crude sample of Sav918. The final concentration of the sample was calculated by B4F titration, and an equivalence point was observed at 4.4 nmoles of biotin-4-fluorescein, which corresponds to 11 μM streptavidin, assuming a tetrameric form of the protein, with all four binding sites.

Scheme 3.1. Asymmetric transfer hydrogenation of imines for the production of salsolidine. The reactions were carried out in a total volume of 200 μL , at room temperature for 24 h using the following final concentrations (when possible): 50 μM complex, 5 mM substrate, and 100 μM tetrameric streptavidin produced in *Pichia pastoris*, at pH 6.5 (MOPS buffer, 0.6 M) containing 3 M HCOONa.

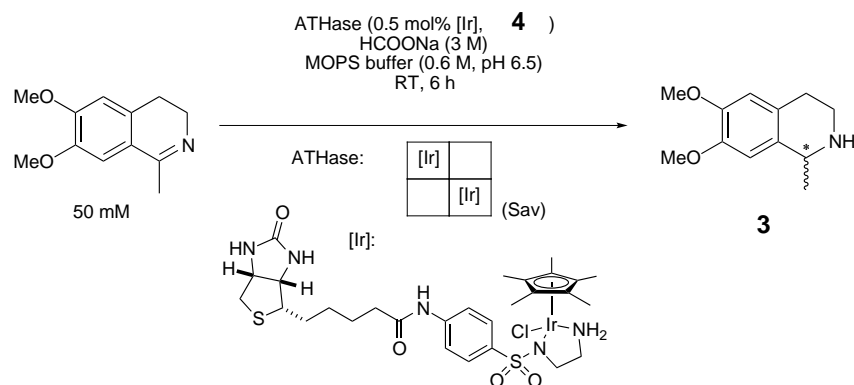


Table 3.1. Selected results for the production of salsolidine **3**, using the biotinylated piano stool complex **4** \subset Sav918.^a

Entry	Protein	Description ^a	Conv. ^b [%]	ee ^b [%]
1	Sav	produced in <i>E.coli</i> and purified by affinity chromatography	76	61 (<i>R</i>)
2	Sav918	purified by affinity chromatography	79	64 (<i>R</i>)
3	Sav918	concentrated supernatant from fermentation	16	25 (<i>R</i>)

^a The reaction was carried out at 30°C for 24 h, using 1 mol% complex **4**, 100 mol% substrate and 100 μM tetrameric streptavidin at pH 6.5 (MOPS buffer, 0.6 M) containing 3 M HCOONa. Total reaction volume: 200 μL .

^b Determined by HPLC after extraction.

The results presented in Table 3.1 confirmed that the artificial metalloenzyme created with Sav918 (entry 2) had similar catalytic performance as with Sav expressed in *E. coli* (entry 1), since similar conversion and *ee* were obtained. Regarding catalysis performed with supernatant (entry 3), further studies have to be conducted, namely optimisation of the reaction conditions in terms of working concentrations, as low protein concentration may lead to low conversion compared to standard conditions (100-fold more concentrated).

3.3 Conclusion & outlook

The very high reported levels of protein secretion in high-density cultures, such that the product can comprise over 80% of the protein in the medium have made *Pichia* an attractive

expression system for many researchers. Yet, *Pichia* is no “magic bullet” and secretion is complex and dependent not only on gene dosage and Mut phenotype, but also on signal sequence processing, proteolysis and glycosylation, which can affect the yield and quality of product.

Pichia pastoris has proven to be a suitable expression system for the heterologous expression of streptavidin, which was successfully expressed, in a fermentor, at 650 mg/L in the culture medium. Sufficient amounts of Sav for bioconjugation with an organometallic catalyst and subsequent catalysis were obtained through the expression of Sav in *P. pastoris*. However, for this application, the production in the prokaryotic host has to be further optimised to provide sufficient amounts of protein in the medium, in a shorter period of time.

Improvement of the expression levels and time of production in *P. pastoris* can be achieved by switching from the *AOX1* promoter to the constitutive strong promoter from glyceraldehyde-3-P dehydrogenase (*GAP*), which is readily available for heterologous expression of proteins in *Pichia pastoris*. In the *GAP* promoter-based system, the generation of biomass and production of protein occur simultaneously, in medium containing glycerol or glucose as sole carbon source. However, this constitutive system has one limitation: the protein of interest cannot be toxic to the host.^[44,45] Though (strep)avidin is believed to show cell toxicity due to its high affinity with biotin, Mattanovitch and co-workers engineered a biotin-prototrophic yeast strain, which was transformed with plasmids containing the protein genes of enhanced green fluorescent protein and porcine trypsinogen, under control of the *GAP* promoter, proving that both systems are compatible and could be used for the secretion of Sav.^[46]

However, a simpler approach to optimise this system would be to investigate the influence of pH, temperature, methanol concentration, cell density, medium composition or additives (casamino acids, sorbitol, EDTA) on protein expression in *P. pastoris*, since not all were tested during this project. Design of experiments (DoE) should be used to properly assess the effect of each parameter, on its own or in combination, on protein expression. A broader range of the factors should also be investigated, e.g. pH from 3.0 to 6.0, methanol concentration from 0.5 to 3%, and temperatures from 15 to 30 °C.

Another approach to improve artificial metalloenzymes based on streptavidin-biotin system would be the use of cell-free protein expression.^[47] Since the end of the 1990s, efforts have been directed in the development and optimisation of *E. coli* derived cell-free system. This system is suitable for expression of toxic proteins,^[48] and although yields are known to be lower than of cell-based protein production, cell-free expression systems are more convenient for screening of constructs. Thus implementation of high-throughput screening^[49,50] would be possible, and the presence of biotin in streptavidin binding sites would be no longer an issue.

3.4 References

- [1] Green, N. M. *Biochemical Journal* **1963**, *89*, 599–609.
- [2] Bayer, E. A.; Wilchek, M. *Journal of Chromatography A* **1990**, *510*, 3–11.
- [3] Wilchek, M.; Bayer, E. A. *Avidin-biotin technology*; volume 184 of *Methods in Enzymology* Academic Press: New York, USA, 1990.
- [4] Diamandis, E. P.; Christopolous, T. K. *Clinical Chemistry* **1991**, *37*, 625–636.
- [5] Sakahara, H.; Saga, T. *Advances in Drug Delivery Review* **1999**, *37*, 89–101.
- [6] Nordlund, H. R.; Hytönen, V. P.; Laitinen, O. H.; Kulomaa, M. S. *Journal of Biological Chemistry* **2005**, *280*, 13250–13255.
- [7] Laitinen, O. H.; Nordlund, H. R.; Hytönen, V. P.; Kulomaa, M. S. *Trends in Biotechnology* **2007**, *25*, 269–277.
- [8] Helppolainen, S. H.; Nurminen, K. P.; Määttä, J. A.; Halling, K. K.; Slotte, J. P.; Huhtala, T.; Liimatainen, T.; Ylä-Hertualla, S.; Airene, K. J.; Närvänen, A.; Jänis, J.; Vainiotalo, P.; Valjakka, J.; Kulomaa, M. S.; Nordlund, H. R. *Biochemistry Journal* **2007**, *405*, 397–405.
- [9] Hytönen, V. P.; Määttä, J. A.; Niskanen, E. A.; Huuskonen, J.; Helttunen, K. J.; Halling, K. K.; Nordlund, H. R.; Rissanen, K.; Johnson, M. S.; Salminen, T. A.; Kulomaa, M. S.; Laitinen, O. H.; Airene, T. T. *BMC Structural Biology* **2007**, *7*, 1–20.
- [10] Sardo, A.; Wohlschlager, T.; Lo, C.; Zoller, H.; Ward, T. R.; Creus, M. *Protein Expression and Purification* **2011**, *77*, 131–139.
- [11] Niederhauser, B.; Siivonen, J.; Määttä, J. A.; Jänis, J.; Kulomaa, M. S.; Hytönen, V. P. *Journal of Biotechnology* **2012**, *157*, 38–49.
- [12] Thomas, C. M.; Ward, T. R. *Applied Organometallic Chemistry* **2005**, *19*, 35–39.
- [13] Ward, T. R. *Accounts of Chemical Research* **2011**, *44*, 47–57.
- [14] Steinreiber, J.; Ward, T. R. *Bio-inspired catalysts*; volume 25 Springer Berlin Heidelberg: Berlin, Germany, 2009.
- [15] Streu, C.; Meggers, E. *Angewandte Chemie International Edition* **2006**, *45*, 5645–5648.
- [16] Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Taglieber, A.; Hollmann, F.; Mondiere, R. J. G.; Dickmann, N.; Hoecker, B.; Cerrone, S.; Haeger, M. C.; Sterner, R. *ChemBioChem* **2008**, *9*, 552–564.
- [17] Dürrenberger, M.; Heinisch, T.; Wilson, Y. M.; Rossel, T.; Nogueira, E.; Knörr, L.; Mutschler, A.; Kersten, K.; Zimbron, M. J.; Pierron, J.; Schirmer, T.; Ward, T. R. *Angewandte Chemie International Edition* **2011**, *50*, 3026–3029.

- [18] Podtetenieff, J.; Taglieber, A.; Bill, E.; Reijerse, E. J.; Reetz, M. T. *Angewandte Chemie International Edition* **2010**, *49*, 5151–5155.
- [19] Köhler, V.; Wilson, Y. M.; Lo, C.; Sardo, A.; Ward, T. R. *Current Opinion in Biotechnology* **2010**, *21*, 744–752.
- [20] Hyster, T. K.; Knörr, L.; Ward, T. R.; Rovis, T. *Science* **2012**, *338*, 500–503.
- [21] Sasmal, P. K.; Streu, C. N.; Meggers, E. *Chemical Communications* **2013**, *49*, 1581–1587.
- [22] Higgins, D. R.; Cregg, J. M. *Introduction to Pichia pastoris*; volume 103 of *Methods in Molecular Biology* Humana Press: Totowa, USA, 1998.
- [23] Cereghino, J. L.; Cregg, J. M. *FEMS Microbiology Reviews* **2000**, *24*, 45–66.
- [24] Verweken, W.; Callewaert, N.; Kaigorodov, V.; Geysens, S.; Contreras, R. *Pichia Protocols*; volume 389 of *Methods in Molecular Biology* Humana Press: New York, USA, Second ed.; 2007.
- [25] Sano, T.; Cantor, C. R. *Proceedings of the National Academy of Sciences of the United States of America* **1990**, *87*, 142–146.
- [26] Gallizia, A.; de Lalla, C.; Nardone, E.; Santambrogio, P.; Brandazza, A.; A., S.; Arosio, P. *Protein Expression & Purification* **1998**, *14*, 192–196.
- [27] Humbert, N.; Schürmann, O.; Zocchi, A.; Neuhaus, J.-M.; Ward, T. R. *Methods in Molecular Biology* **2008**, *418*, 101–110.
- [28] Nagarajan, V.; Ramaley, R.; Albertson, H.; Chen, M. *Applied and Environmental Microbiology* **1993**, *59*, 3894–3898.
- [29] Wu, S. C.; Hassan Qureshi, M.; Wong, S. L. *Protein Expression & Purification* **2002**, *24*, 348–356.
- [30] Casteluber, M. C.; Damasceno, L. M.; da Silveira, W. B.; Diniz, R. H.; Passos, F. J.; Passos, F. M. *Biotechnology Progress* **2012**, *28*, 1419–1425.
- [31] Ringenberg, M. R.; Ward, T. R. *Chemical Communications* **2011**, *47*, 8470–8476.
- [32] Inan, M.; Meagher, M. M.; Plantz, B. A.; Sinha, J. *Biotechnology and Bioengineering* **2005**, *89*, 102–112.
- [33] Jungo, C.; Urfer, J.; Zocchi, A.; Marison, I.; von Stockar, U. *Journal of Biotechnology* **2007**, *127*, 703–715.
- [34] Routledge, S. A.; Hewitt, C. J.; Bora, N.; Bill, R. M. *Microbial Cell Factories* **2011**, *10*, 1–11.
- [35] Cooper, C. A.; Gasteiger, E.; Packer, N. *Proteomics* **2001**, *1*, 340–349.
- [36] Zacharius, R. M.; Zell, T. E.; Morrison, J. H.; Woodlock, J. J. *Analytical Biochemistry* **1969**, *30*, 148–152.

- [37] Roth, Z.; Yehezkel, G.; Khalaila, I. *International Journal of Carbohydrate Chemistry* **2012**, *2012*, 1–10.
- [38] Gamboa-Suasnavart, R. A.; Valdez-Cruz, N. A.; Cordova-Dávalos, L. E.; Martínez-Sotelo, J. A.; Servaín-González, L.; Espitia, C.; Trujillo-Roldán, M. A. *Microbial Cell Factories* **2011**, *10*, 1–11.
- [39] Montesino, R.; Nimtz, M.; Quintero, O.; García, R.; Falcón, V.; Cremata, J. A. *Glycobiology* **1999**, *9*, 1037–1043.
- [40] Liu, H.; Pan, H. C.; Cai, S. X.; Chen, Z. W.; Zheng, X. F.; Yang, H. T.; Xiao, Z. Y. *Sheng Wu Gong Cheng Xue Bao* **2005**, *21*, 107–112.
- [41] Baumann, K.; Carnicer, M.; Dragosits, M.; Graf, A. B.; Stadlmann, J.; Jouhten, P.; Maahelmo, H.; Gasser, B.; Albiol, J.; Mattanovich, D.; Ferrer, P. *BMC Systems Biology* **2005**, *4*, 1–22.
- [42] Lombardi, A.; Bursomanno, S.; Lopardo, T.; Traini, R.; Colombatti, M.; Ippoliti, R.; Flavell, D. J.; Flavell, S. U.; Ceriotti, A.; Fabbrini, M. S. *The FASEB Journal* **2010**, *24*, 253–265.
- [43] Zocchi, A.; Jobé, A. M.; Neuhaus, J.-M.; Ward, T. R. *Protein Expression & Purification* **2003**, *32*, 167–174.
- [44] Waterham, H. R.; Digan, M. E.; Koutz, P. J.; Lair, S. V.; Cregg, J. M. *Gene* **1997**, *186*, 37–44.
- [45] Vassileva, A.; Chugh, D. A.; Swaminathan, S.; Khanna, N. *Journal of Biotechnology* **2001**, *88*, 21–35.
- [46] Gasser, B.; Dragosits, M.; Mattanovich, D. *Metabolic Engineering* **2010**, *12*, 573–580.
- [47] Zubai, G. *Annual Review on Genetics* **1973**, *7*, 267–287.
- [48] Szafranski, P.; Mello, C. M.; Sano, T.; Smith, C. L.; Kaplan, D. L.; Cantor, C. R. *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 142–146.
- [49] Kigawa, T.; Yabuki, T.; Matsuda, N.; Matsuda, T.; Nakajima, R.; Tanaka, A.; Yokoyama, S. *Journal of Structural and Functional Genomics* **2004**, *5*, 63–68.
- [50] Kim, T. W.; Kim, D. M.; Choi, C. Y. *Journal of Biotechnology* **2006**, *124*, 373–380.

New scaffold for the creation of an artificial metalloenzyme

I was taught that the way of progress was neither swift nor easy.

Marie Skłodowska-Curie

Abstract

The design and synthesis of novel biomolecular scaffolds, which can incorporate organometallic moieties, are an important goal in the discovery of new artificial metalloenzymes. The proposed research intended to investigate the versatility of human carbonic anhydrase II (hCAII) as host protein for the design of specific organometallic · protein assembly. This Chapter reports on the design and production of hCAII variants based on:

- i. the design and synthesis of metal complex scaffolds (work carried out by Dr Fabien Monnard) and the x-ray crystallographic studies of hCAII in complex with an organometallic moiety (work carried out by Dr Tillmann Heinisch);
- ii. the rational design of variants based on Molecular Dynamics Simulations (MDS) and Molecular Mechanics /Generalised Born Surface Area (MM-GBSA) calculations (work carried out by Dr Maurus Schmid);
- iii. and the study of pseudo-contact shifts in solution state NMR for the determination of the structure of an inhibitor · human carbonic anhydrase II in solution (work carried out by Mr Kaspar Zimmermann).

A second ligand generation of catalysts based on a 2-picolylamine ligand bearing a sulphamide anchor showed improved binding affinities toward wild-type hCAII. After one round of genetic optimisation, the novel hybrid catalyst afforded an enantiomeric excess of 29% (*S*) for the synthesis of salsolidine. The success of this chemo-genetic approach was only possible thanks to computational simulations and crystal structures. Force field parameters amenable to molecular dynamics simulations of hCAII · inhibitor interactions were experimentally validated. Five different single cysteine constructs were isotopically expressed, and site-specific labelling of hCAII with paramagnetic lanthanides was successfully achieved.

4.1 Introductory remarks

Human carbonic anhydrase II is a structurally and functionally well-defined protein; it is a zinc-containing enzyme that catalyses the interconversion of carbon dioxide and water into bicarbonate and protons, and one of the most active enzymes known, with a $k_{cat}/K_M \sim 1.5 \times 10^8 \text{ M}^{-1} \cdot \text{s}^{-1}$, approaching the limit of diffusion control.^[1–3] The ellipsoidal enzyme is a monomeric, single polypeptide chain of 259 amino acids, with a molecular mass of 29.3 kDa, and its x-ray structure has been resolved to 1.54 Å.^[4] The numbering system of human carbonic anhydrase I is used throughout this Chapter. In this system, residues Asn62, Asn67 and Thr200 correspond to residue Asn61, Asn66 and Thr199 in the numbering system of human carbonic anhydrase II.^[5]

4.1.1 Human carbonic anhydrase II as potential new biomolecular scaffold

A variety of highly enantioselective artificial metalloenzymes have been created by using covalent, dative, or supramolecular anchoring strategies.^[6–8] However, the choice of a protein scaffold has been limited by the necessity of a sufficiently large pocket to accommodate the catalyst and substrates.

Ergo, a new biomolecular scaffold featuring a deeper binding pocket in which the organometallic complex could be completely embedded, was investigated. The combination of human carbonic anhydrase II and a sulphonamide anchor as model system for the design of organometallic ·protein assemblies was based on Emil Fischer’s “lock-and-key” hypothesis,^[9] and the following criteria:

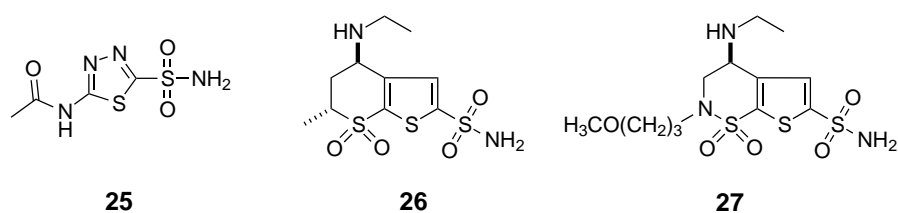
- i. the exceptionally large binding pocket (15 Å deep, and 15 Å diameter at its mouth) of the enzyme,
- ii. the well-established conserved mode of binding of sulphonamides to hCAII,
- iii. the compatibility of human carbonic anhydrase II inhibitors with soft transition metals, the possibility to derivate them, and their low dissociation constant,
- iv. the well-characterised assays to determine binding profiles,
- v. and the fact that carbonic anhydrase isozymes are over-expressed in certain forms of cancers, which makes this system interesting for therapeutic applications.^[10–12]

Furthermore, hCAII is easily over-expressed in *E. coli* cultures and purified by affinity chromatography.^[13] Therefore, hCAII mutants, with optimised reactive sites to which to couple ligands, can easily be generated.

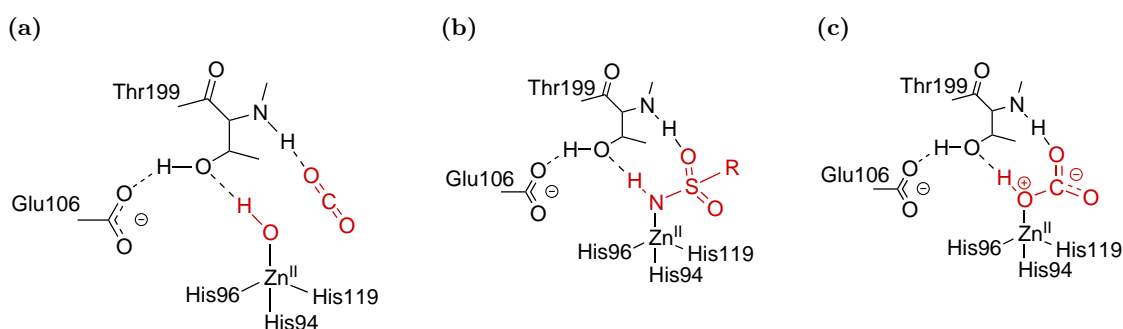
4.1.2 Sulphonamides as inhibitors

In 1940, Mann and Keilin discovered that sulphanilamides and certain related substances were powerful and specific inhibitors of mammalian carbonic anhydrase isozymes, at concentrations as low as 2×10^{-6} M.^[10,14] Most of known inhibitors containing a sulphonamide/sulphamate moiety (*e.g.* the clinically used derivatives acetazolamide, dorzolamide, and brinzolamide, Scheme 4.1) can coordinate the catalytic Zn^{II} ion of the enzyme active site, through their deprotonated nitrogen atom, which coordinates to the zinc ion and an extended network of hydrogen bonds, involving the “gatekeepers” residues Thr199 and Glu106 (Scheme 4.2). These residues also participate to the anchoring of the inhibitor molecule to the metal ion, whereas the organic part (heterocyclic/aromatic) of the inhibitor interacts with the hydrophobic and hydrophilic residues of the cavity.

Scheme 4.1. Structure of carbonic anhydrase inhibitors. All three inhibitors – acetazolamide (**25**), dorzolamide (**26**) and brinzolamide (**27**) – are used in the treatment of glaucomas. Acetazolamide is also used to treat epileptic seizures, cystinuria, idiopathic intracranial hypertension, and durla extasia. Dorzolamide, developed by Merck, was the first drug obtained from structure-based drug design.^[15]



Scheme 4.2. Diagram comparing (a) Carbon dioxide (putative interactions). (b) An arylsulphonamide; and (c) Bicarbonate bound to the active site of hCAII. The arylsulphonamide can be viewed as a transition state analogue of the hydratase reaction.^[16]



4.1.3 Structure of human carbonic anhydrase II

One of the keys to modern drug design is the understanding of biological phenomena at atomic resolution. In particular, knowledge of three-dimensional (3D) structures and conformational dynamics of proteins provides direct information on unique interactions with other

macromolecules, which is necessary to understand and design catalytic functions. The two conventional methods for the determination of protein structure are x-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy. Through these techniques, the coordination chemistry in a protein scaffold can be determined, and rational design and future applications for hybrid catalysts can be developed.

Crystallographic structures

In 1972, the first crystal structure of a mammalian α -CA, human carbonic anhydrase II (hCAII, previously called “CA C”) was described by Anders Liljas and co-workers.^[17] It was reviewed later on, in 2000, by Stams and Christianson.^[18]

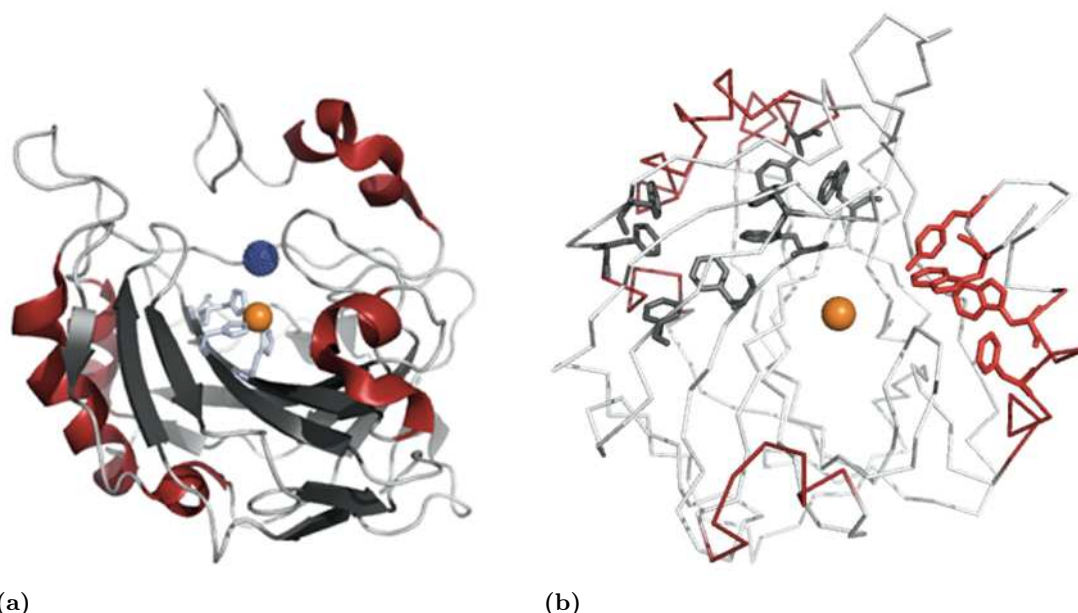


Figure 4.1. Ribbon diagram of the active site of human carbonic anhydrase II (PDB code: 1G54) (a) Tertiary structure is coloured as following: α -helices in red, β -sheets in dark grey and loops in light grey; the active site zinc is shown as orange space-filled sphere and it is coordinated by three histidine residues (His94, His96, and His119) coloured in grey and by a water molecule (Wat263, coloured in blue). (b) Stereo-drawing showing the two aromatic clusters in hCAII. In red, the first cluster (Trp5 and 16, Tyr7, and Phe20), and in dark grey, the second cluster (Phe66, 70, 93, 95, 176, 179, 226 and Trp97).

hCAII has a 16% helical structure (10 helices, 42 residues) and a dominating 10-fold β -sheet (29%, 18 strands and 77 residues) that extends throughout the entire molecule and is the predominant secondary element. Except for two pairs of parallel strands, the β -sheet is antiparallel. Through its tertiary structure determination, Liljas *et al.* showed that the protein exists of only one domain (Figure 4.1)^[18] devoid of disulphide bridges. Furthermore, all lysine residues are located at the surface of the enzyme, which are in close contact to the side chains of the neighbouring molecules. Two hydrophobic clusters formed by aromatic

residues were noted within the folding pattern of hCAII.^[19,20]

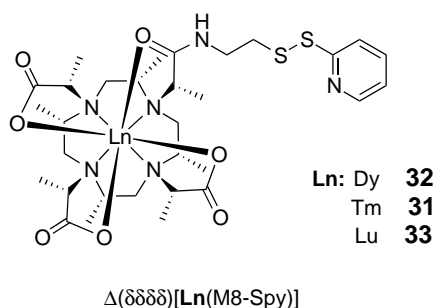
NMR structures

In February 2013, 9,740 3D NMR structures of proteins and nucleic acids in solution were registered in the Brookhaven Protein Data Bank (PDB), which accounted for only 10% of the total of structures deposited. The scarcity of 3D NMR structures of macromolecules reflects one of its major drawbacks - its size limitation to proteins with molecular weights between 20 and 30 kDa.^[21,22] However, recent advances in NMR spectroscopy have enabled solution 3D NMR studies of larger biomolecules.^[23] One of these promising methods relies on the measurement of pseudo-contact chemical shifts (PCS) induced by a covalently bound paramagnetic metal ion (Scheme 4.3). In order to measure the paramagnetic effects, two NMR spectra are recorded with and without the lanthanide chelating tag.^[24] By superimposition of the resulting NMR spectra, the PCS of nuclear spins can be easily measured as the difference (in ppm) of the chemical shifts between the two spectra. The paramagnetism originating from a single paramagnetic centre (dysprosium, lutetium or thulium) can be described in terms of a magnetic susceptibility tensor, χ , spanned by three principal axes (χ_x , χ_y , and χ_z) (Equations 4.1 and 4.2).

$$\Delta\chi_{\text{ax}} = \frac{\chi_z - (\chi_x + \chi_y)}{2} \quad (4.1)$$

$$\Delta\chi_{\text{rh}} = \chi_x - \chi_y. \quad (4.2)$$

Scheme 4.3. High-affinity lanthanide chelating tag, $\Delta(\delta\delta\delta\delta)$ [Ln(M8-Spy)]. “Ln” refers to thulium (Tm **31**), dysprosium (Dy **32**) or lutetium (Lu **33**), and “M8” to the eight methyl groups attached to rigidify the DOTA moiety.



This alternative structural technique provides the precise information on the location of the inhibitors/metal complexes in the protein scaffold, in solution.

4.1.4 Research project

The development of a novel class of enantioselective artificial metalloenzymes is one of the most attractive targets in the field of inorganic and catalytic chemistry, since these hybrid catalysts show remarkable chemo-selectivity and reactivity in aqueous media. For this purpose, covalent modification of the protein and cofactors have usually been utilised to attach a metal complex to a protein scaffold.

This Chapter focused on the dative insertion of metal complexes into human carbonic anhydrase II environment. It included the screening of stable metal complex/protein composites, crystal structures, and molecular design for regulating enantioselectivity of the target catalytic reactions. The purpose of this work was to create a library of hCAII variants focusing different objectives:

- i. in order to test hCAII as scaffold for artificial metalloenzymes, hCAII wild-type (wt hCAII) was expressed to evaluate its affinity toward complexes synthesised by Dr Fabien Monnard (**FM**), and monitored by protein crystallography (Dr Tillmann Heinisch, **TH**). In addition several other mutations were introduced into the active site replacing amino acids that are relevant for the binding of designed hCAII inhibitors.
- ii. the mutants hCAII L198x (x = A, Q, F, H) were created in the context of validating Molecular Dynamics Simulation (MDS) calculations performed by Dr Maurus Schmid (**MS**) and to evaluate their affinity toward benzenesulphonamide inhibitor.
- iii. the double mutants C206S-S50C, -S166C, -S173C, -S217C and -S220C were expressed (isotopically labelled forms) to investigate NMR pseudo-contact shifts (PCS) introduced by [M8-SPy] ligands (Mr Kaspar Zimmermann, **KZ**, and PD Dr Daniel Häußinger, **DH**) when tagged to the protein.

4.2 Results & discussion

The results presented herein are a compilation of data obtained in collaboration with Dr Fabien Monnard (**FM**), Dr Maurus Schmid (**MS**), Dr Tillmann Heinisch (**TH**), Kaspar Zimmermann (**KZ**), and PD Dr Daniel Häußinger (**DH**). It is noted throughout the text who performed which experiment.

4.2.1 Production of human carbonic anhydrase II

The plasmid pACA encoding for human carbonic anhydrase II (hCAII) was a generous gift from Prof. Dr Carol A. Fierke (University of Michigan, USA).^[13] This construct consists of

the hCAII gene^[25] behind a T7 RNA polymerase promoter, an f1 origin of replication,^[26,27] and ampicillin (amp^r) and chloramphenicol (cm^r) resistance genes in pMa5-8 vector (Figure 4.2).^[27] The construct of this plasmid has an alanine residue at position 2 instead of a serine, with no effect on protein expression or catalytic properties,^[28] and was used as a template for PCR.

Strain and genetic construct

The construct coding for hCAII (Figure 4.2) was amplified by transformation into *Escherichia coli* DH5 α cells, and the size of the plasmid was verified by restriction analysis. The results of the digestion of pACA plasmid were analysed on an agarose gel (Figure 4.3).

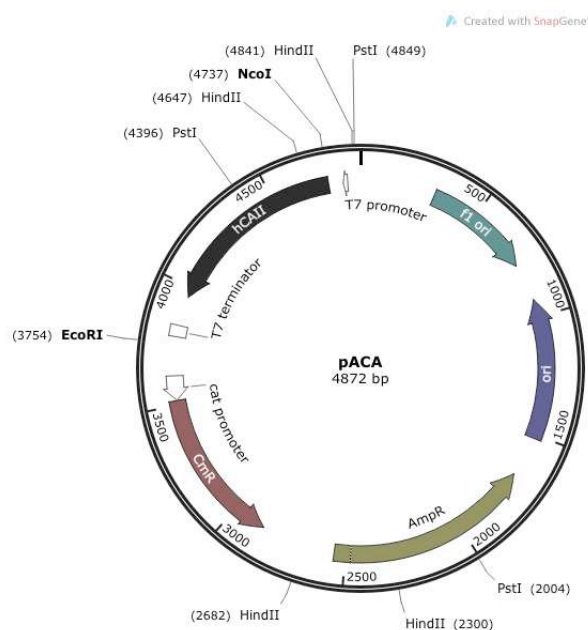


Figure 4.2. Plasmid map of pACA. Human carbonic anhydrase II gene (782 bp) was inserted behind a T7 promoter, into pMA5-8 vector (4089 bp) containing amp^r and cm^r markers for antibiotic resistance, and f1 origin of replication.

PstI generated three fragments of 4,849, 4,396 and 2,004 bp, and HindIII, two fragments of 4,853 and 3,969 bp (Table 4.1). The double digestion using NcoI and EcoRI generated one fragment of \sim 4,737 bp (although it seems that the plasmid DNA “broke” in between position 4,396 and 4,853 bp, generating a third, faint band), and a second of 3,754 bp. No negative control was performed, as the empty plasmid pMa5-8 was not available.

When uncut, the DNA of pACA migrated only \sim 4 kb, which might be an indication that

Table 4.1. Restriction map details of pACA plasmid.

Description	HindIII	PstI	NcoI	EcoRI
Sequence	AACGTT	CTGCAG	CCATGG	GAATTC
Site length	6	6	6	6
Overhang	5'	3'	5'	5'
Frequency	2	3	1	1
Cut position (bp)	3,969/4,853	2,004/4,396/4,849	4,737	3,754
Fragment size (kb)	0.9, 3.9	0.45, 2.0, 2.3	1.0	3.9

it was in Form II (nicked circle DNA)¹. When linearised by single or double digestions, the plasmid presented the correct size (see Figure 4.3, and details in Table 4.1).

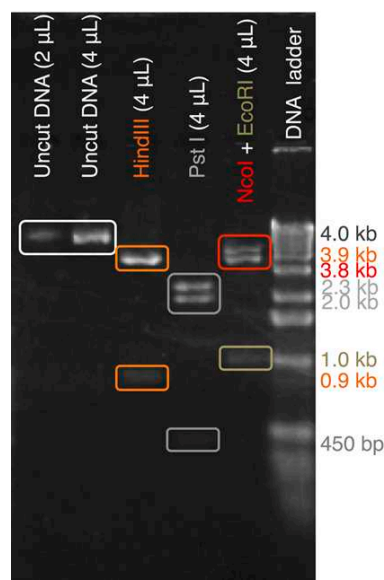


Figure 4.3. Analysis of restriction map of pACA plasmid. The DNA ladder is BenchTop 1kb Ladder (Promega AG - Dübendorf, CH). 2 and 4 μL of uncut DNA were loaded on the 1.2% agarose gel, and 4 μL for the other samples.

After confirmation of the correct construct to express human carbonic anhydrase II, pACA PCR product was transformed into *E. coli* XL-1 Blue super-competent cells, and sent to sequencing to confirm the authenticity of the construct (Figure 4.4).

¹Nicked circle DNA (Form II or “relaxed circle”) is the slowest conformation of uncut DNA. A nick can occur during isolation of the plasmid due to mechanical shearing of DNA or enzyme activity. In bacteria, topoisomerase I enzyme will nick one strand of the helix so that DNA polymerase have access to DNA for replication. The superhelical tension relaxes once one of the strands has been cut, and the tightly-wound ball becomes a floppy circle


```

ATGGCCCATCACTGGGGGTACGGCAACACACAAACGGACCTGAGCACCTGGCATAAGGACTTCCCCATTGCCAAGGGAGAGCGCCAGTCCCCT 30
M A H H W G Y G K H N G P E H W H K D F P I A K G E R Q S P
GTTGACATCGACACTCATAACAGCCAAGTATGACCCTTCCCTGAAGCCCTGTCTGTTTCCCTATGATCAAGCAACTTCCCTGAGGATCCTC 60
V D I D T H T A K Y D P S L K P L S V S Y D Q A T S L R I L
AACAAATGGTCAATGCTTTCAACGTGGAGTTTGATGACTCTCAGGACAAAGCAGTGCCTCAAGGGAGGACCCCTGGATGGCACTTACAGATTG 90
N N G H A F N V E F D D S Q D K A V L K G G P L D G T Y R L
ATTCAGTTTCACTTTCACTGGGGTTCACCTTGATGGACAAAGTTCAGAGCATACTGTGGATAAAAAGAAATATGCTGCAGAACTTCACTTG 120
I Q F H F H W G S L D G Q G S E H T V D K K K Y A A E L H L
GTTCACTGGAAACACCAATAATATGGGGATTTGGGAAAGCTGTGCAGCAACTGATGGACTGGCCGTTCTAGGTATTTTTTTGAAGGTTGGC 150
V H W N T K Y G D F G K A V Q Q P D G L A V L G I F L K V G
AGCGCTAAACCGGGCCCTCAGAAAGTTGTTGATGTGCTGGATTCCATTAAAACAAGGGCAAGAGTGTCTGACTTCACTAACTTCGATCCT 180
S A K P G L Q K V V D V L D S I K T K G K S A D F T N F D P
CGTGGCCTCCTTCTGAAATCCCTGGATTACTGGACCTACCCAGGCTCACTGACCACCCCTCCTCCTTCTGGAATGTGTGACCTGGATTGTG 210
R G L L P E S L D Y W T Y P G S L T T P P L L E C V T W I V
CTCAAGGAACCCATCAGCGTCAGCAGCGAGCAGGTGTTGAAATTCGTAACCTTAACCTCAATGGGGAGGGTGAACCCGAAGAAGTATGAT 240
L K E P I S V S S E Q V L K F R K L N F N G E G E P E E L M
GTTGACAACCTGGCGCCAGCTCAGCCACTGAAGAACAGGCAAAATCAAGCTTCCCTCAATAA 260
V D N W R P A Q P L K N R Q I K A S F K STOP

```

Figure 4.4. Sequence of pACA plasmid. In beige, loops; in grey, α -helices; in red, β -strands; in italic, “direct” and “indirect” binding site ligands; in bold, active site residues; and in red bold, “substrate binding” residues.

The expression system was optimised for high protein yields. Up to 400 mg pure protein was achieved from 1 L expression culture in shake flasks. Optimised site-directed mutagenesis^[29] allowed the creation of the desired mutants.

Expression and detection of human carbonic anhydrase II

Even though there is a trend to standardise as much as possible protein expression processes, it has been shown that even subtle changes would powerfully increase the productivity. Human carbonic anhydrase II expression was initially carried out following the protocol provided by Prof. Fierke, until improvement was required to circumvent the low expression yields obtained in the production of isotopically labelled protein. The know-how obtained in the expression of labelled hCAII was transferred to standard expressions, reducing the production time and achieving higher expression yields, in commonly employed shake flasks.

According to Fierke’s protocol, over-expression of hCAII can be obtained through a three-step process, (i) an inoculum (5 mL LB medium inoculated with one medium-sized colony (*E. coli* BL21(DE3) transformant), incubated for 6 to 7 h, at 37 °C and 250 rpm; (ii) a pre-culture (25 mL of fresh LB medium, inoculated with the 5 mL inoculum), incubated overnight growth at 37 °C and 250 rpm; and (iii) the culture carried out in minimal medium supplemented with ZnSO₄, and inoculated with 0.5% (v/v) of the pre-culture (or starting OD₆₀₀ = 0.0125). The temperature was kept constant at 37 °C, until induction time at which it was lowered to

30 °C. Cultures were incubated at 250 rpm for 3 hours (or until $OD_{600} = 0.8 - 1.0$), and were induced by addition of IPTG and $ZnSO_4$. Three hours later, 8 $\mu\text{g}/\text{mL}$ phenylmethylsulphonyl fluoride (PMSF, prepared in isopropanol) was added to the culture broth, to inhibit serine proteases growth. After additional growth for 3 h at 30 °C, cells were harvested (4,400 rpm, 4 °C for 15 min).

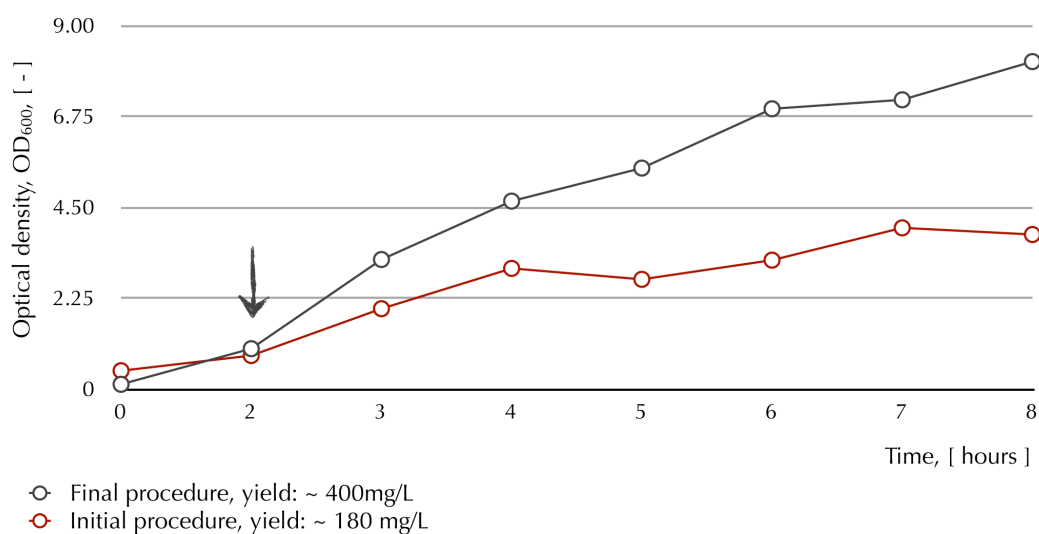


Figure 4.5. Example of the cellular growth of a 1 L culture, in 3 L non-baffled Erlenmeyer flasks, using the initial and final procedures. The arrow indicates the induction time. The “final procedure” and “initial procedure” cultures were induced at $OD_{600} \sim 0.8$, and were incubated at 250 rpm, for 6 h post-induction, and at 25 and 30 °C (post-induction), respectively.

Although this protocol was, initially, strictly followed, cells took often more than 3 hours to be at an optical density of 0.8 – 1.0, and to be inducible. Final OD_{600} did not exceed 2.5, and final wt hCAII yield was of ~ 25 mg per litre of culture. Some parameters, such as the *E. coli* expression strain (changed from BL21(DE3) to BL21(DE3)pLysS, to avoid the sonication step in the recovery of the enzyme and achieve tighter expression control), the culture temperature after induction (18, 25 and 30 °C), and the time of induction ($OD_{600} = 0.8 - 1.0$ or $OD_{600} = 1.0 - 1.3$), were investigated. The first round of optimisation yielded 180 mg of pure protein per litre of culture.

The optimised procedure for the expression of human carbonic anhydrase II is as described below. The main differences with Fierke’s protocol lied on the *E. coli* expression strain (BL21(DE3) vs BL21(DE)pLysS), the volume of the inoculum and pre-culture (inoculum: 5 mL vs 15 mL, and pre-culture: 15 mL vs 60 mL), centrifugation step before inoculation of the pre-culture and main culture (to remove β -lactamase, and minimise proteolysis), resuspension of main culture pellet in 20% glucose,^[30] and culture temperature after induction set to 25 °C instead of 30 °C (Figure 4.5 and Figure 4.6).

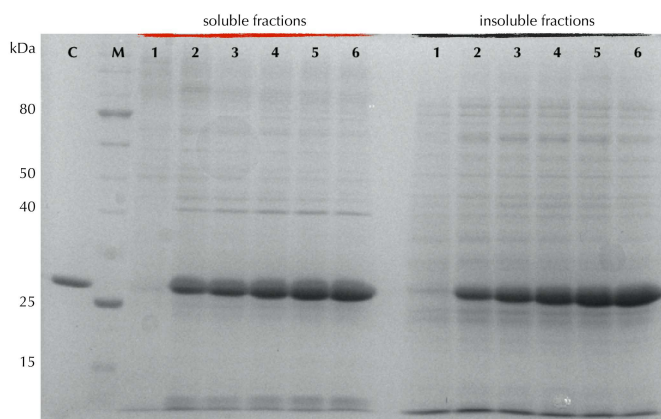
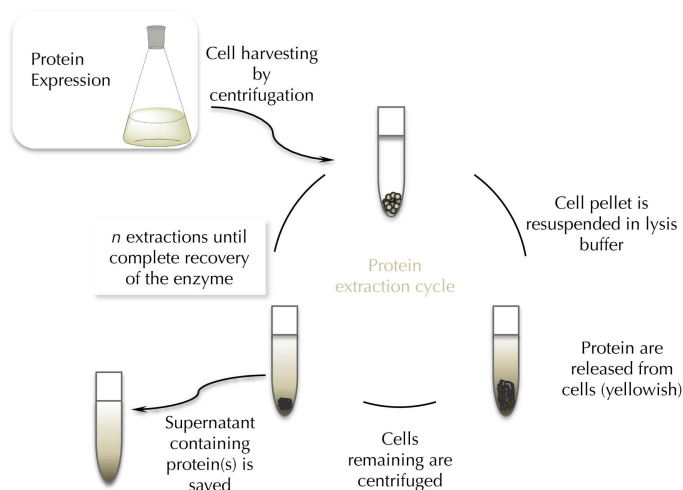


Figure 4.6. Visualisation of the increased wt hCAII yield after optimisation, by SDS-PAGE analysis. Soluble and insoluble fractions, at different stages of the expression, visualised by Coomassie Blue staining. The control “C” was purified wt hCAII, 200 $\mu\text{g}/\text{mL}$. The “M” reports to the molecular weight marker from NE Biolabs. Lane 1: “zero” hour (0 h) reports to a sample taken before induction, confirming no leaky expression in the absence of the inducer, IPTG; Lanes 2 to 6: 1, 2, 3, 4, and 6 hours post-induction, respectively.

To ascertain the maximum recovery of expressed enzyme from *E. coli* expression cells, several extractions were performed, as shown on Scheme 4.4. Five extractions were necessary for complete recovery of the protein from the cell pellet (Figure 4.7). The supernatants resulting from these extractions were pooled and purified as a single batch.

Scheme 4.4. Extraction of hCAII from *E. coli* expression cells. The pellet obtained from 1 L culture was resuspended in activity buffer (50 mM Tris- SO_4 , pH 8.6 and 0.5 mM ZnSO_4), supplemented with 10 $\mu\text{g}/\text{mL}$ PMSF and 1 $\mu\text{g}/\text{mL}$ DNase I. The resuspension was incubated under vigorous agitation, until complete reduction of the viscosity of the cell lysate. The resulting lysate was centrifuged at 12,000 rpm, for 20 min, 4 $^\circ\text{C}$. This procedure was repeated another five times (six extractions, in total).



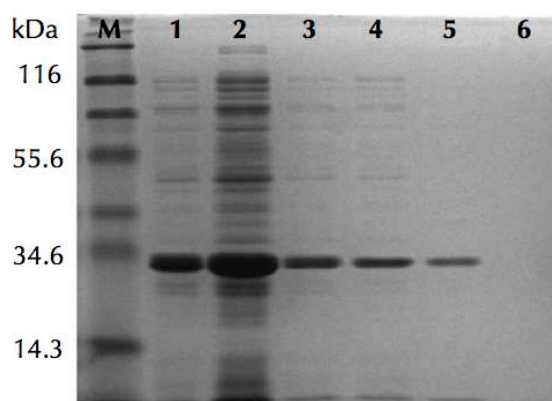


Figure 4.7. SDS-PAGE analysis of the extraction of hCAII from *E. coli* expression cells.

The first two extractions allowed almost complete recovery of wt hCAII. No enzyme was recovered after the fifth extraction (Figure 4.7, Lane 6).

Purification of human carbonic anhydrase II

The rapid purification of a recombinantly over-expressed protein from a complex mixture of proteins can be achieved by affinity chromatography, which takes advantage of the high affinity of a protein for specific ligands or chemical groups. The purification of human carbonic anhydrase II by affinity is based on the biological specificity that the enzyme has for sulphonamide inhibitors. *p*-amino-methylbenzene sulphonamide has identical inhibitory capacity as its derivative, toluene sulphonamide (a known inhibitor of CAs, $K_i = 0.1 \mu\text{M}$),^[31] and it can be covalently attached to agarose (*p*-amino-methylbenzene sulphonamide agarose, Sigma-Aldrich). The agarose-bound enzyme can be eluted by adding a high concentration of a competitive inhibitor, such as a monovalent anion. Two elution methods, using either SCN^- ^[32] or $\text{ClO}_4^-/\text{CH}_3\text{COO}^-$ ^[33] were tested for the purification of human carbonic anhydrase II, and the best result (analysed by SDS-PAGE) was implemented as the method of choice to purify hCAII.

Elution with sodium perchlorate and sodium acetate - protocol 1

The second method tested for the purification of hCAII was described by Gould and Tawfik.^[33] The column was equilibrated with activity buffer (50 mM Tris- SO_4 , pH 8.6 and 0.5 mM ZnSO_4), and the enzyme was loaded onto the sulphonamide column producing agarose-bound enzyme. The column was first washed with five CVs of buffer (50 mM Na_2SO_4 , 50 mM NaClO_4 , and 25 mM Tris at pH 8.8) to remove any unbound contaminants. The bound enzyme was eluted by addition of 200 mM NaClO_4 and 100 mM NaCH_3COO , at pH 5.6, and 5 mL eluted fractions were collected (Figure 4.8).

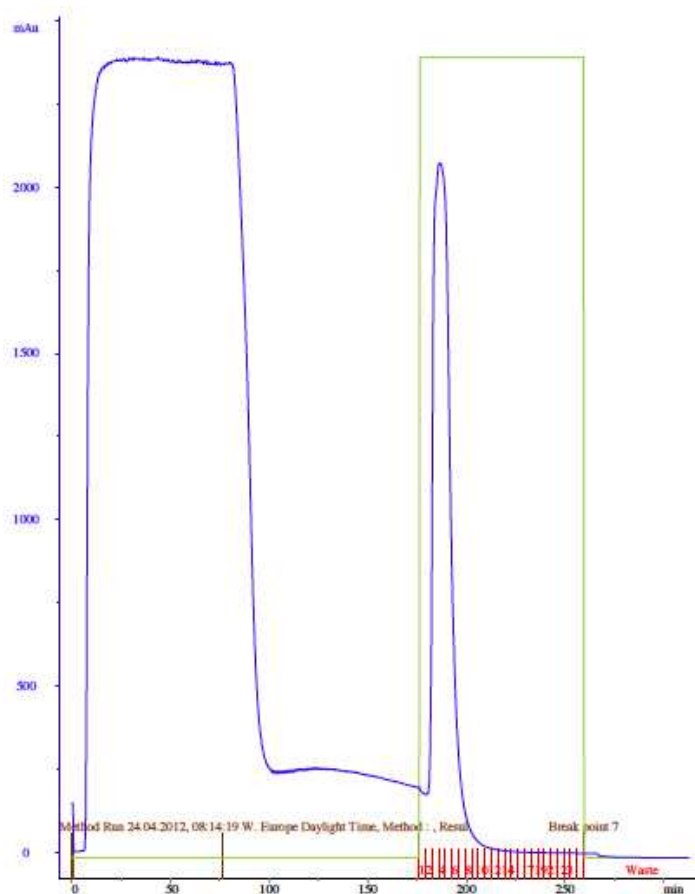


Figure 4.8. Elution profile of wt hCAII, eluted with sodium perchlorate/sodium acetate (**protocol 2**). After washing the column with 125 mL of 50 mM Na_2SO_4 , 50 mM NaClO_4 , and 25 mM Tris at pH 8.8, the enzyme was eluted with 200 mM NaClO_4 and 100 mM NaCH_3COO , at pH 5.6. The protein fractions from the main peak were pooled and analysed by SDS-PAGE.

Elution with potassium thiocyanide - protocol 2

Based on the work of Bering and Kuhns,^[32] wt hCAII, resuspended in activity buffer (50 mM Tris- SO_4 , pH 8.0 and 0.5 mM ZnSO_4), was slowly loaded at a flow rate of 1 mL/min onto a column packed with *p*-amino-methylbenzene sulphonamide agarose, previously equilibrated with the activity buffer. The column was subsequently washed with five column volumes (CVs) of Tris/ Na_2SO_4 buffer. The sulphonamide-bound enzyme was eluted with 10 CVs of 25 mM Tris, pH 8.3 and 0.4 M KSCN, and 5 mL fractions were collected.

The presence of wt hCAII, and purity of the fractions collected in both methods (protocols 1 & 2) were analysed on 12% SDS-PAGE (Figure 4.9).

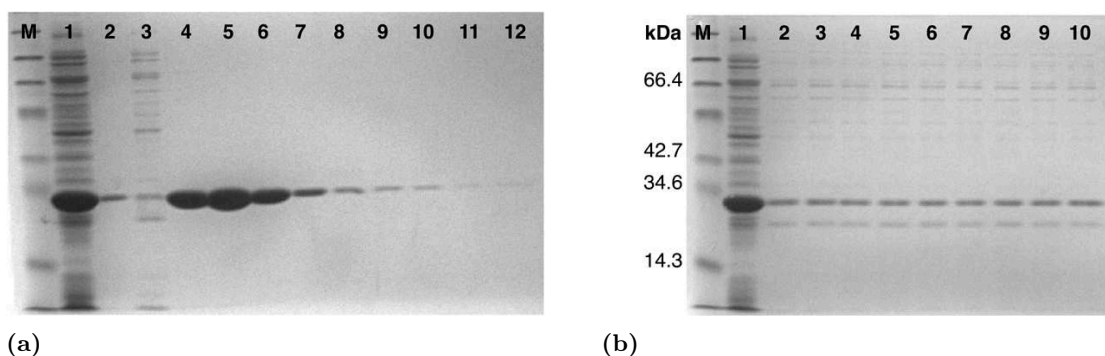


Figure 4.9. SDS-PAGE analysis of wt hCAII, (a) eluted with sodium perchlorate/sodium acetate (**protocol 1**) or (b) potassium thiocyanide (**protocol 2**), at different stages of purification. Electrophoresis was performed on 12% resolving gel and 4% stacking gel, and proteins were stained by Coomassie Blue. “M” stands for marker; (a) Lane 1: unpurified sample; Lane 2: sample from the wash step; Lanes 3 to 10: fractions collected, 1 to 8; (b) Lane 1: unpurified sample; Lane 2: empty lane; Lane 3: sample from the wash step; Lanes 4 to 12: fractions collected, 1 to 9.

As shown in Figure 4.9, potassium thiocyanide was not the best elution buffer compared to the one containing sodium perchlorate and sodium acetate. KSCN was a “weak” inhibitor to compete with sulphonamide, as the same amount of wt hCAII was present in the wash and elution fractions. The peak obtained by FPLC was broad (data not shown), and ten column volumes (250 mL total volume) were not sufficient to elute the enzyme completely. These results were somehow surprising as a trend is observed in anion inhibition, and thiocyanide ligand (SCN^-) is known to have a higher association constants with zinc than ClO_4^- and CH_3COO^- .^[34]

Mass spectrometry analysis

Figure 4.10 presents an ESI-MS mass spectrum measured for wt hCAII. The most abundant isotopic mass ($m^{m.a.}$) of the isozyme was determined to be 29,099.4 Da, which is in good agreement with the calculated value ($m^{m.a.} = 29,098.9$ Da, calculated with “Compute pI/Mw” from ExPASy),^[35] with the initial methionine removed and the second residue changed from serine to alanine.

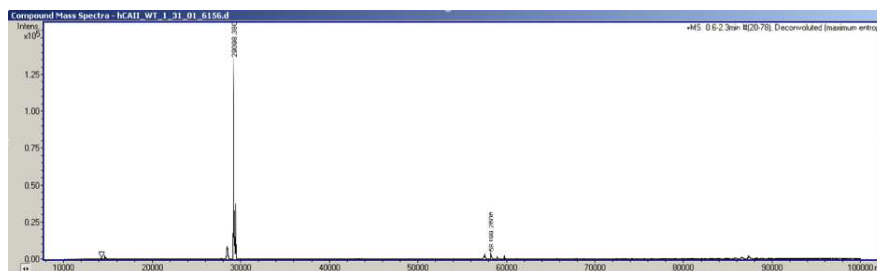


Figure 4.10. Positive ESI-MS spectrum of wt hCAII in methanol/formic acid (50:0.05, v/v), pH 3.0 – 4.0 and acetonitrile/acetic acid/TFA (50:0.1:11, v/v), pH 0.0 – 1.0. Calculated mass of the protein: 29,098.9 Da. Calculated mass of the main peak: 29,099.4 Da.

In Table 4.2 are listed the calculated and determined mass of hCAII variants.

Table 4.2. Calculated and determined masses of human carbonic anhydrase II variants. For the labelled proteins, the first mass corresponds to the selectively labelled sample (^{15}N Leu), and the second mass to the uniformly labelled one (^{15}N).

Protein	Calculated Mw [Da]	Determined Mw [Da]
wt	29,098.9	29,099.4
A2V-F131V	29,078.9	29,050.8
H64A	29,032.8	n.d. ^a
H64A-E106Q	29,031.8	n.d. ^a
I91A	29,056.8	29,052.0
Q92G	29,017.8	29,011.7
Q92G-V121G	29,116.9	29,119.6
V121G	29,056.8	29,056.00
F131A	29,022.8	29,014.7
K170A	29,041.8	29,044.0
L198A	29,188.0	29,156.4
L198F	29,132.9	29,133.0
L198H	29,122.8	29,123.6
L198Q	29,113.8	29,115.0
P202W	29,188.0	29,189.0
C206S-S50C	29,098.9	29,116.0/29,446.5 ^b
C206S-S166C	29,098.9	29,116.3/29,448.7 ^b
C206S-S173C	29,098.9	29,116.4/29,450.0 ^b
C206S-S217C	29,098.9	29,116.5/29,450.2 ^b
C206S-S220C	29,098.9	29,117.3/29,449.7 ^b

^a n.d. - Not determined.

^b Determined by **KZ**.

4.2.2 Studies on human carbonic anhydrase II as potential biomolecular scaffold

The high affinity of human carbonic anhydrase II toward sulphonamide inhibitors allows the insertion of achiral metal catalysts into the chiral environment of the protein when linked to a sulphonamide anchor. Thus, dative incorporation of an active organometallic moiety into the active site of human carbonic anhydrase II was explored.

The synthesis and characterisation of organometallic complexes, inhibition studies and catalysis were carried out by **FM**, who also performed docking experiments along with **MS**. X-ray crystallographic studies were performed by **TH**.

Library of mutants

In the absence of x-ray crystallographic information of the hybrid catalyst, the decomposition per residue of the binding free energies ($\Delta E_{\text{MM}} + \Delta G_{\text{solV}}$) offered the possibility to rationally design mutants (Figure 4.11).

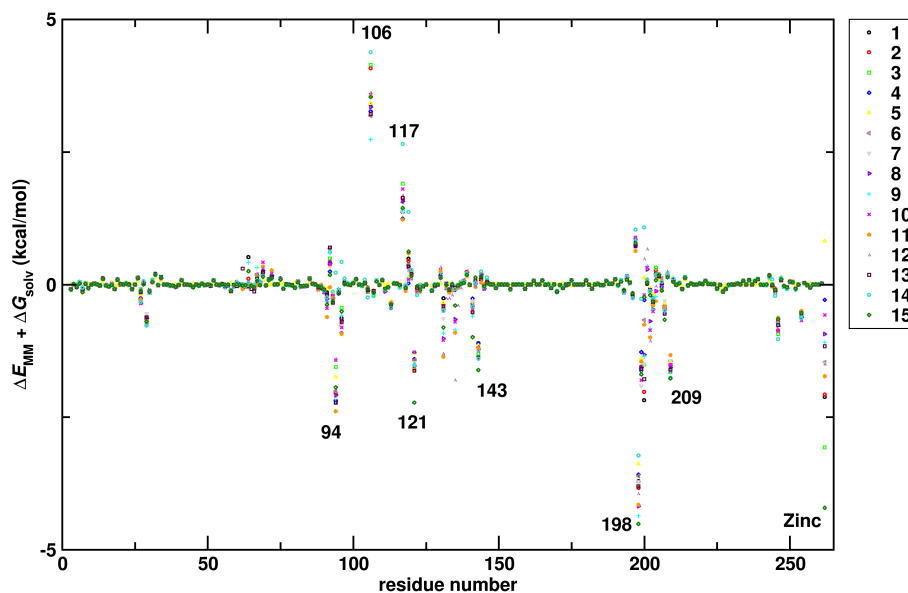


Figure 4.11. Per residue free energies $\Delta E_{\text{MM}} + \Delta G_{\text{solv}}$ for ligand \subset hCAII combinations. The list of ligands tested (1 – 15) can be found in Scheme 4.8. Picture by MS.

Based on docking calculations, a small library of mutants were designed, expressed and tested for their binding affinity.

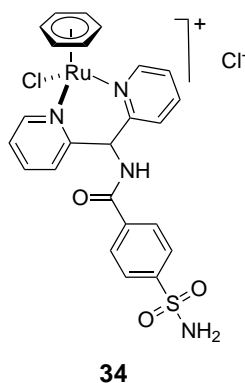
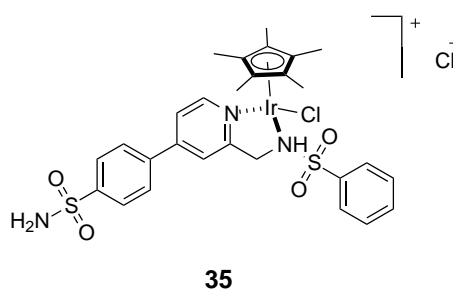
Library of piano stool complexes

The first approach relied on wild-type human carbonic anhydrase II expressed in *E. coli* as biomolecular scaffold and a d^6 -piano stool complex containing an arylsulphonamide anchor as a catalyst ($[(\eta^6\text{-arene})\text{Ru}(\text{bispy})\text{Cl}]^+$, Scheme 4.5). The conjugation of the Ru complexed with wt hCAII was confirmed by x-ray crystallography, and the catalytic activity of the artificial metalloenzyme ($[(\eta^6\text{-C}_6\text{Me}_6)\text{Ru}(\text{bispy})\text{Cl}]^+ \subset \text{wt hCAII}$) toward *p*-nitrophenyl acetate (PNPA) was evaluated.

The second approach relied on a small library of hCAII mutants constructed based on computational studies and a second generation of catalysts bearing a 2-picolyamine derivative (Scheme 4.6). The hybrid catalysts were screened for the transfer hydrogenation of 1-methyl-6,7-dimethoxy-3,4-dihydroisoquinoline.

Studies on the catalytic properties

The inhibition constant, K_i , was determined by steady-state kinetic experiments (**FM**). Ligands or complexes with inhibition constants on the order of mM were titrated against hCAII. Organometallic compounds with a d^6 electron count act as precursors of Lewis acid catalysts

Scheme 4.5. Bispyridine-type complex, $[(\eta^6\text{-C}_6\text{H}_6)\text{RuCl}_2]_2$.**Scheme 4.6.** 2-picolylamine-type complex, $[(\text{Cp}^*)\text{IrCl}_2]_2$.

in various organic reactions, and hold the potential of use as therapeutic agents.^[36] In this respect, the perspective of using optically active organometallic ligands, such as three-legged piano stool complexes, as chiral catalysts seemed appealing.

First generation of catalysts

A small library of three-legged Ru(II) piano stool complexes bearing an arylsulphonamide moiety – for anchoring purposes – was designed *in silico*. Due to the consistency of the binding mode found for arylsulphonamide \subset hCAII inhibitors,^[16] docking was performed manually with Maestro.^[37] For this reason, the structurally characterised sulphonamide of the model arylsulphonamide \subset hCAII structure (PDB code: 1G54)^[38] was deleted, the structure was refined and used as a template to superimpose the arylsulphonamide-bearing piano stool complexes. This docking process yielded two alternative structures with minimal van der Waals contacts for $[(\eta^6\text{-C}_6\text{Me}_6)\text{Ru}(\text{bispy})\text{Cl}]^+$. Molecular dynamic simulations were carried out on both structures (performed by **FM** and **MS**), and sampling of the simulations revealed a preference for the structure depicted in Figure 4.12, which was synthesised for further studies (performed by **FM**).

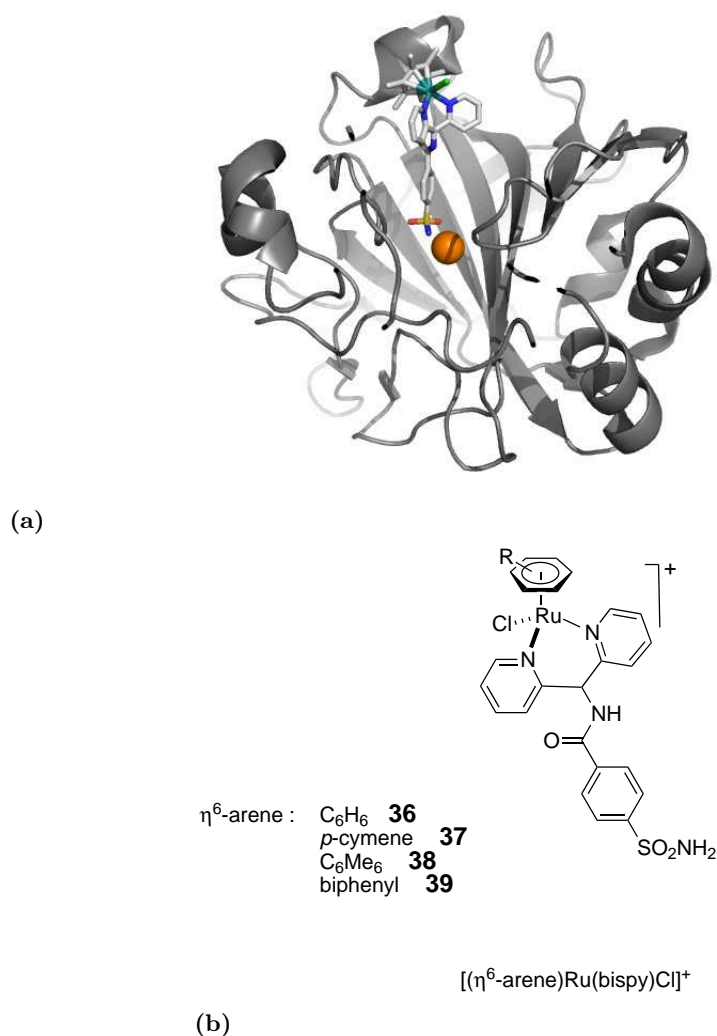


Figure 4.12. (a) Resulting structure of $[(\eta^6\text{-C}_6\text{Me}_6)\text{Ru}(\text{bispy})\text{Cl}]^+$ \subset wt hCAII, from docking experiments and molecular dynamic simulations (PDB code: 1G54). The zinc ion (orange filled sphere) interacts with the arylsulphonamide anchoring group of the three-legged piano stool complex (coloured by elements). (b) Piano stool complexes.

The binding profiles of $[(\eta^6\text{-arene})\text{Ru}(\text{bispy})\text{Cl}]^+$ toward wt hCAII were determined by measuring the initial rates of the hydrolysis of *p*-nitrophenyl acetate at 25 °C, pH 8.0.^[39–41] The formation of product was monitored spectrophotometrically at the isosbestic point for the corresponding nitrophenol and nitrophenolate ion (348 nm). The binding constants, K_i , were calculated using Equation 4.3 for the reaction with *p*-nitrophenyl acetate, and fitted using Gnuplot (v.4.2., least-squares method, **FM**). The non-enzymatic rates estimated from the blank reactions (all components except wt hCAII) were subtracted from the observed total initial rates (Table 4.3).

$$v = \frac{v_0 K_i}{K_i + \{[I]_t - 0.5(A - \sqrt{A^2 - 4[I]_t[E]_t})\}} \quad (4.3)$$

with v_0 being the initial rate of the enzyme-catalysed reaction in the absence of inhibitor, K_i the inhibition constant, $A = [I]_t + [E]_t + K_i$ the inhibitor total concentration, and $[E]_t$ the enzyme total concentration.^[42]

The data in Table 4.3 illustrates the interactions between hCAII and the metal moieties. Compared to the parent carboxylic acid 4-carboxybenzenesulphonamide, all complexes bearing the bispy ligand displayed increased affinity. Interestingly, $[(\eta^6\text{-biphenyl})\text{Ru}(\text{bispy})\text{Cl}]^+$ displayed the highest affinity toward hCAII compared to $[(\eta^6\text{-C}_6\text{H}_6)\text{Ru}(\text{bispy})\text{Cl}]^+$, which illustrates the subtle complementarity between the piano stool moiety and the binding pocket.

Table 4.3. Dissociation constants of $[(\eta^6\text{-arene})\text{Ru}(\text{bispy})\text{Cl}]^+ \subset \text{wt hCAII}$, with $\eta^6\text{-arene}$: benzene, *p*-cymene, C_6Me_6 , or biphenyl.

Inhibitor	K_i [nM]
$[(\eta^6\text{-benzene})\text{Ru}(\text{bispy})\text{Cl}]\text{Cl}$ 40	194 ± 19
$[(\eta^6\text{-}p\text{-cymene})\text{Ru}(\text{bispy})\text{Cl}]\text{Cl}$ 41	275 ± 13
$[(\eta^6\text{-C}_6\text{Me}_6)\text{Ru}(\text{bispy})\text{Cl}]\text{Cl}$ 42	329 ± 16
$[(\eta^6\text{-biphenyl})\text{Ru}(\text{bispy})\text{Cl}]\text{Cl}$ 43	145 ± 12

Crystallographic studies

X-ray crystal structure studies of the metal · protein hybrids provide valuable information on the interactions of the metal centre, to better understand and design the second coordination sphere. Structural insight of wt hCAII complexed with a ruthenium moiety bearing a bispy ligand, $[(\eta^6\text{-C}_6\text{Me}_6)\text{Ru}(\text{bispy})\text{Cl}]^+$, was obtained by soaking the enzyme in a solution of the metal-complex (Figure 4.13). Diffraction data were collected at the Synchrotron (Swiss Light Source, Paul Scherrer Institut - Villigen, Switzerland) to 1.3 Å resolution. After refinement of the crystal structure, it was established that, upon ligand binding, the phenyl group of residue Phe131 underwent conformational change, from tense (F131-T) to relaxed conformation (F131-R) to prevent clashes, and to form a CH/π interaction.^[43]

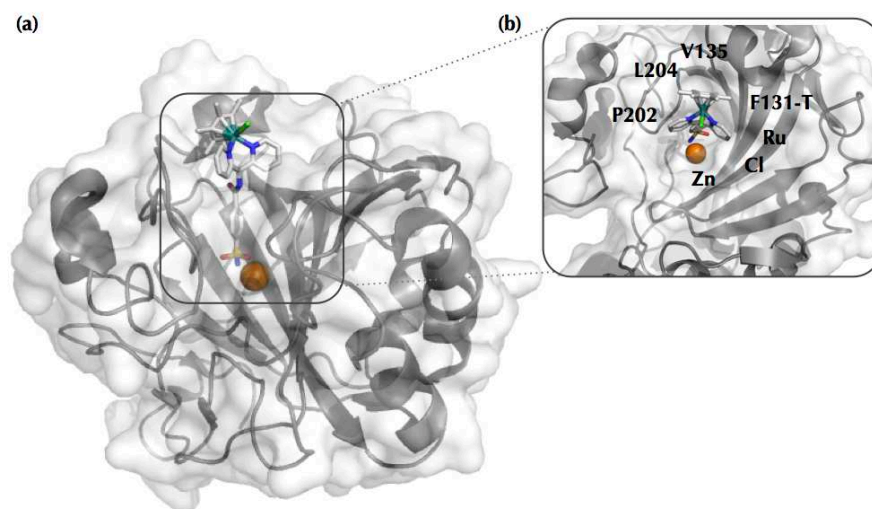


Figure 4.13. Crystal structure of $[(\eta^6\text{-C}_6\text{Me}_6)\text{Ru}(\text{bispy})\text{Cl}]^+ \subset \text{wt hCAII}$ (PDB code: 3PYK)^[43] (a) Overview of the complex embedded in the enzyme scaffold. The zinc ion (orange filled sphere) is deeply buried inside the binding site of hCAII and interacts with the sulphonamide group of the piano stool complex. (b) Close-up view of the piano stool complex. The complex is positioned at the entrance of the cone-shaped binding site and interacts with the hydrophobic wall side chains. Work carried out by **TH**.

The crystal structure indicated that the amino acids residues adjacent to the zinc active site provided a well-defined chiral environment to control the tacticity, and illustrated the subtle complementarity between the piano stool moiety and the cone-shaped cavity of hCAII (residues V121, F131, V135, L141, L198, P202, and L204).

Second generation of catalysts

The catalytic potential of $[(\text{Cp}^*)\text{IrCl}_2]_2 \subset \text{hCAII}$ was evaluated using the transfer hydrogenation of 1-methyl-6,7-dimethoxy-3,4-dihydroisoquinoline as a model system (Scheme 4.7) The reactions were performed with 1.8 mol% complex (0.35 mM final concentration), 20 mM substrate, 0.4 mM protein in 0.4 M MOPS buffer solution (200 μL , 5% DMSO) containing 3 M sodium formate, pH 7.5, for 44 h at 40 $^\circ\text{C}$ (work carried out by **FM**). Results are presented in Table 4.4.

Scheme 4.7. Asymmetric transfer hydrogenation of imines for the production of salsolidine **3**.

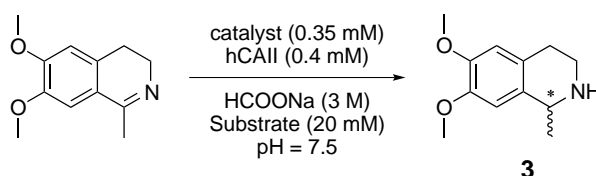


Table 4.4. Results obtained for the transfer hydrogenation of 1-methyl-6,7-dimethoxy-3,4-dihydroisoquinoline with hCAII variants.^a

Entry	Complex	Protein ^b	Conv. ^c [%]	TON	ee ^c [%]
1	35	wild-type	33	18	27 (<i>S</i>)
2	35	H64A	11	6	n.d.
3	35	I91A	99	55	29 (<i>S</i>)
4	35	K170A	59	33	30 (<i>S</i>)
5	35	E106Q-H64A	21	12	3 (<i>S</i>)
6	35	F131A	12	6	6 (<i>S</i>)
7	35	F131A-A2V	24	13	1 (<i>S</i>)
8	35	Q92G	15	8	3 (<i>S</i>)
9	35	Q92G-V121G	25	14	1 (<i>S</i>)
10	35	L198Q	11	6	23 (<i>S</i>)
11	35	L198A	6	3	6 (<i>R</i>)
12	35	L198F	13	7	23 (<i>S</i>)
13	35	L198H	13	7	23 (<i>S</i>)

^a The reaction was carried out at 40 °C for 20 h using 1.8 mol% complex (0.35 mM final concentration), 20 mM substrate, 0.4 mM protein, in 0.4 M MOPS buffer (200 μ L total volume, 5% DMSO) containing 3 M formate, pH 7.5.

^b ESI-MS of hCAII isozymes are reported in Table 4.2.

^c Determined by normal phase HPLC after extraction.

In comparison with wt hCAII (Table 4.4, entry 1), best results were obtained with mutants I91A and K170A (entries 3 and 4, respectively). The presence of a smaller residue (alanine) around the active metal catalyst affected only the catalyst activity and had a minor effect on the *ee* value. No significant enantioselectivity was obtained with the remaining mutants. Upon optimisation of the catalysis conditions (catalyst loading, temperature and time of the reaction), wt hCAII afforded 53% conversion and 56% *ee* in favour of the (*S*) enantiomer and mutant I91A, 75% conversion and 58% *ee* (*S*). These results were obtained for reactions carried out at room temperature for 20 h and under the same catalyst loading conditions as above described.^[44]

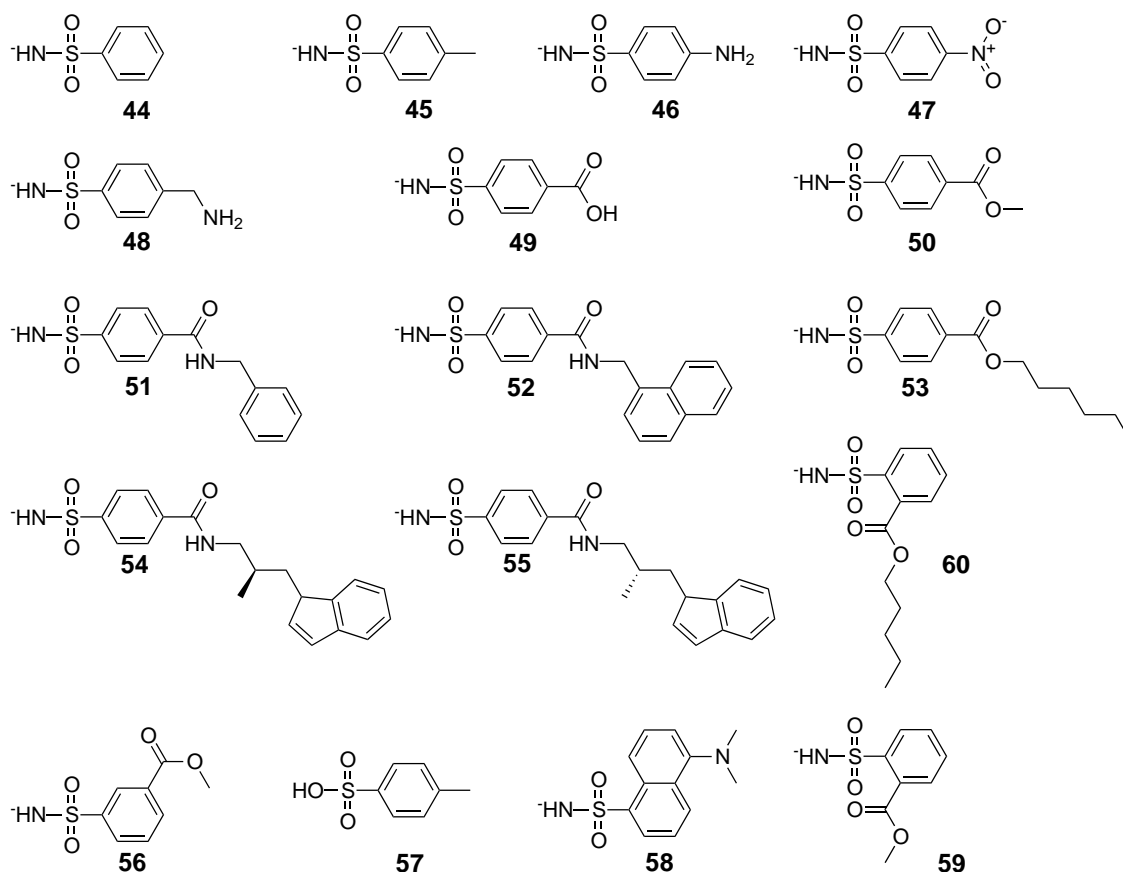
A crystal structure of these newly created artificial metalloenzymes gave structural insights on the guest-host interactions (PDB code: 3ZP9),^[44] and will allow further rational design of the biomolecular scaffold.

4.2.3 Arylsulphonamides as inhibitors

The application of docking methods and parameters to study inhibitors-enzyme interactions of human carbonic anhydrase II system was explored. More specifically, the docking of arylsulphonamides (Scheme 4.8), which constitute the best characterised and most common class of inhibitors.

Molecular Mechanics - Generalized Born Solvent Approximation (MM-GBSA) was selected to estimate ligand-binding free energies for a given set of arylsulphonamides. This simulation strategy was validated by comparison of the results with published biophysical data and with a simulation using a Quantum Mechanics/Molecular Mechanics (QM/MM) implementation with the Self-Consistent Charge Density-Functional Tight-Binding (SCCDFTB) module in CHARMM. Furthermore, point mutations of the amino acid residues identified as critical to binding of arylsulphonamides were computationally investigated, and binding free energies were determined (work carried out by **MS** and **FM**).

Scheme 4.8. Structures of arylsulphonamide inhibitors used for computational studies. These predictions carried out by **MS** were compared with experimental biophysical data (performed by **FM**) from hCAII variants expressed recombinantly in *E. coli*.



Library of hCAII L198x mutants

Leucine at position 198 was identified, by computation, as critical in terms of energetics contribution in the affinity of benzenesulphonamide (**44**) for wt hCAII, and three mutants

were designed and produced recombinantly in *E. coli*: L198A, L198F and L198Q. Human carbonic anhydrase II variants were expressed in 1 L shake flasks, and were purified by affinity chromatography. Between 200 and 300 mg of pure protein (> 95% purity) were obtained for each mutant, as confirmed by SDS-PAGE analysis and ESI-TOF MS.

Benzenesulphonamide (Scheme 4.8, **25**) was the inhibitor of choice to determine the affinity toward the identified protein, hCAII Leu198 and its variants (A, F and Q). Inhibition data for all four variants (wild-type, and L198x, x = A, F and Q) with benzenesulphonamide are presented in Table 4.5. The corresponding thermodynamics were determined using the esterase activity assay, and fitted raw data are presented in Figure 4.15.

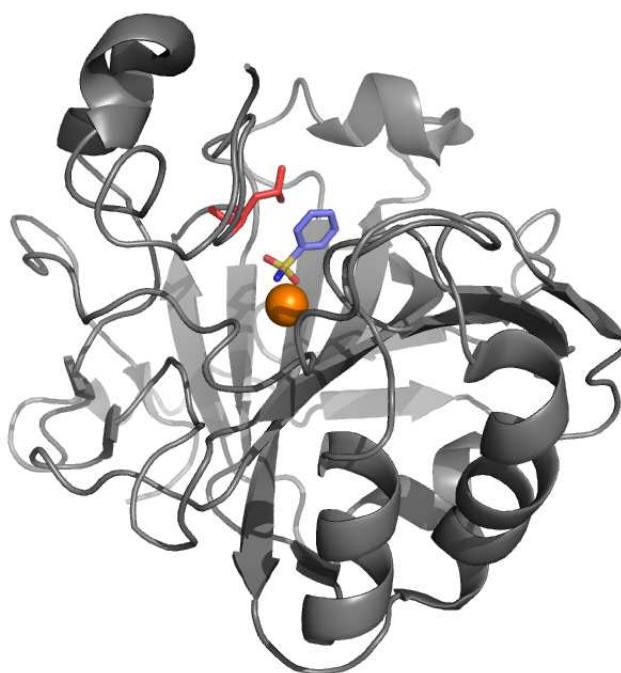


Figure 4.14. Benzenesulphonamide **44** docked into human carbonic anhydrase II. This inhibitor was chosen for the affinity studies, as it is commercially available. The residue chosen for mutagenesis, Leu 198 is represented in red sticks, the metal centre, zinc as an orange sphere, and the inhibitor in sticks coloured by elements.

An inhibition constant, K_i , of 1100 nM was determined for wt hCAII, which was in good agreement with the reported data ranging from 200 – 1500 nM.^[16] The experimentally measured dissociation constants of **44** for the L198x mutants are reported in Table 4.5.

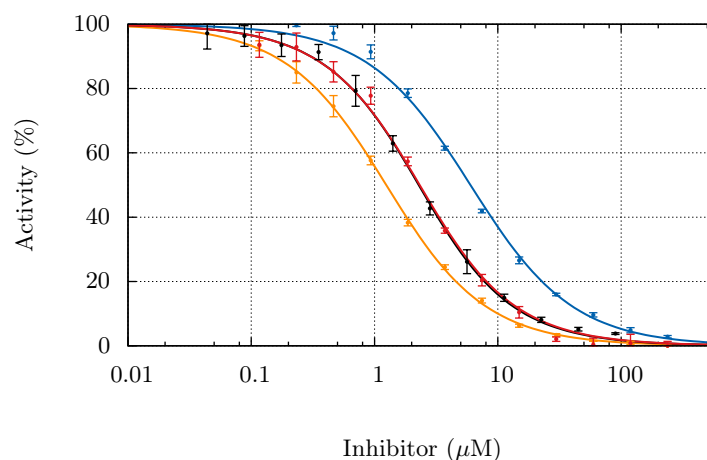


Figure 4.15. Steady-state kinetic data for the inhibition of hCAII mutated proteins by benzenesulfonamide (**44**). The initial rates of the enzyme-catalyzed hydrolysis of *p*-nitrophenyl acetate substrate were measured as a function of inhibitor concentration. [Enzyme] = 1 μ M, [*p*-nitrophenyl acetate] = 0.5 mM. The solid, smooth lines are the best fits of the data according to Equation 4.3 for the K_i of wt (\bullet), L198A (\bullet), L198F (\bullet), L198Q (\bullet). Picture by FM.

The experimentally measured inhibition constants of benzenesulfonamide **25** are summarised in Table 4.5.

Table 4.5. Final set of parameters for dissociation constant, K_d , for benzenesulfonamide **44** with asymptotic standard error for hCAII mutants (see Figure 4.15).^[42]

Entry	Protein	Dissociation constant [nM]	Published dissociation constant [nM]
1	wild-type	1100 \pm 40	200 – 1500 ^[16]
2	L198A	5500 \pm 270	-
3	L198F	1700 \pm 130	-
4	L198Q	1800 \pm 100	-

In spite of a slight over-estimation of the computed free binding energies between hCAII L198x mutants, the experimentally determined energies correlated well with the computed ones (see Figure 4.16).

As illustrated for hCAII L198x mutants, the influence of point mutations at a key position can be predicted from atomistic simulations. Albeit subtle when compared to the range of ligands tested, the influence of point mutations translates into an order of magnitude difference for the corresponding K_d , between wt hCAII and mutant L198Q. Thus, these results demonstrate the applicability of this theoretical model as working hypothesis for the reactive complex.

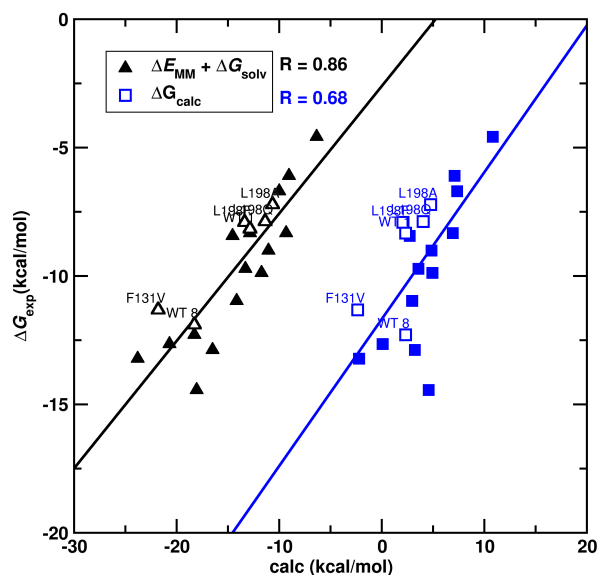


Figure 4.16. Correlation between computed $\Delta E_{MM} + \Delta G_{solv}$ (black triangles) and ΔG (blue squares) and experimentally determined ΔG for wt hCAII (full symbols) and hCAII mutants (empty symbols). Complex **25** was used for all variants except for F131V, which was tested *in silico* against **35**. Picture by MS.

4.2.4 Pseudo-contact shifts in solution-state NMR

A lanthanide chelating tag (Scheme 4.3) was used to attach paramagnetic lanthanide metals to human carbonic anhydrase II, introducing pseudo-contact shifts (PCS).^[24] PCS provide valuable information on the structure and the dynamics of proteins in solution.

The ^{15}N -HSQC and 1D ^{19}F spectra reported in this subsection were recorded and analysed by Mr Kaspar Zimmermann and PD Dr Daniel Häußinger.

Library of C206S-S50x mutants

Five serine residues, solvent-accessible but outside of the conical cleft of the enzyme, were chosen for mutagenesis (S50, S166, S173, S217 and S220, Figure 4.17). Single point mutating these residues to Cys generated a chemical handle for thiol-selective coupling.^[45,46] Human carbonic anhydrase II has an endogenous Cys at position 206, which was mutated to Ser in order to preclude side reaction at this site, thus generating double mutants for each Ser construct (Cys206→Ser, Serx (x = 50, 166, 173, 217 or 220)→Cys). Mutation of Cys206 to serine does not affect the activity nor the stability of the enzyme, as demonstrated by Krebs and Fierke, and Mårtensson *et al.* respectively.^[45,47]

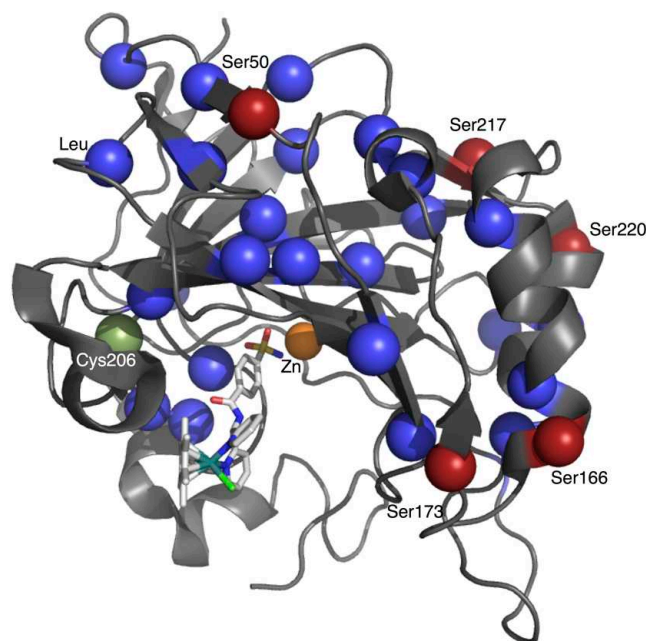


Figure 4.17. Human carbonic anhydrase II complexed with $[(\eta^6\text{-C}_6\text{Me}_6)\text{Ru}(\text{bispy})\text{Cl}]^+$ (PDB code: 3PYK). In red spheres, the solvent-accessible serine residues chosen for mutation to cysteine; in green sphere, the cysteine at position 206; in blue spheres, the 26 Leucine residues of hCAII; in orange sphere, the zinc II ion; and in sticks, the metal complex embedded into the protein scaffold.

The high number of Leucine residues in hCAII (26 in total) was decisive regarding to the choice for site-specific isotopic labelling.

All five single cysteine constructs were isotopically labelled (uniform ^{15}N and specific ^{15}N Leucine), and the construct C206S-S50C was also deuterium labelled, in *E. coli* expression system (Table 4.6). The latter form of labelling proved to be challenging as very low yields of protein were obtained from the two cultures that were carried out (~ 30 mg/L). In order to obtain maximum sensitivity in heteronuclear 3D experiments, ^2H , ^{13}C , ^{15}N labelled mutant was denatured and renatured to exchange all amide ^2H with ^1H . The efficiency of hCAII refolding limited the amount of protein that was recovered. About 50% of the protein was recovered, properly folded.

Table 4.6. Yields [mg/L] of human carbonic anhydrase II mutants (C206S-SyC), isotopically labelled. For uniformly ^{15}N labelled proteins, the cultures were carried out to a final volume of 2 L, and for specific ^{15}N Leu to a final volume of 3 L.

Protein	Uniform, ^{15}N [mg/L]	Specific, ^{15}N Leucine [mg/L]
C206S-S50C	95	230
C206S-S166C	95	140
C206S-S173C	85	180
C206S-S217C	60	180
C206S-S220C	80	165

HSQC spectra

Unambiguous ^{15}N -HSQC NMR spectra of selectively labelled hCAII were obtained (Figure 4.18.a), allowing the quick assignment of the PCS shifted peaks of the protein samples tagged with [Tm(M8-SPy)], and the measurement of the corresponding PCS.

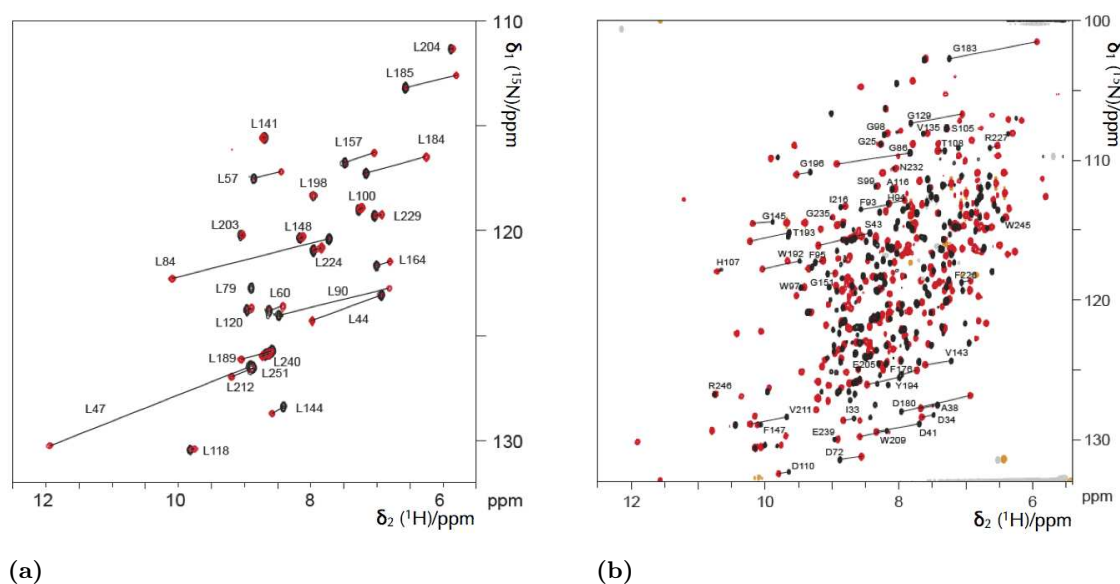


Figure 4.18. Overlay of 2D [^{15}N - ^1H]-HSQC spectra of hCAII C206S-S50C acquired with [Tm(M8-SPy)] (peaks shown in red) and [Lu(M8-SPy)] (peaks shown in black), at 298 K and pH 6.8. For illustrative purpose, solid lines connect selected pairs of diamagnetic and paramagnetic cross-peaks. (a) Selective ^{15}N Leu labelled. (b) Uniform ^{15}N labelled. Spectra by **KZ**.

These pseudo-contact shifts allowed the determination of the initial parameters of the χ -tensor (data not shown), and consequently the assignment of 80% of the amide resonances of the [Tm(M8-SPy)] tagged uniformly labelled samples. These assignments allowed the further refinement of the χ -tensor.

Below an example of the superimposition of the spectra of hCAII C206S-S50C, untagged and tagged with the dysprosium **32** and thulium **31** complexes (Figure 4.19).

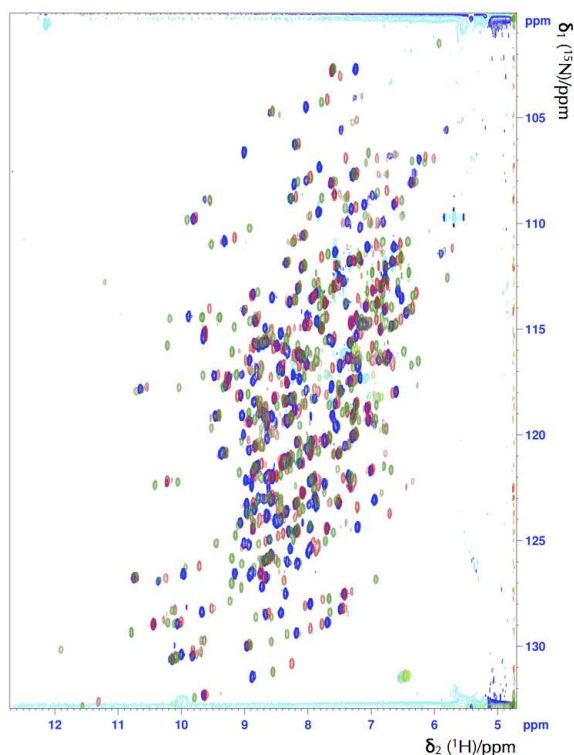
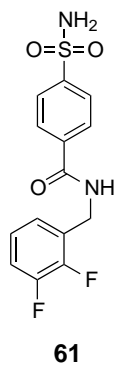


Figure 4.19. Manifestation of PCSs in ^{15}N - ^1H -HSQC NMR spectra of hCAII C206S-S50C. The example shows the superimposition of three ^{15}N -HSQC spectra of the protein, untagged and tagged with Dy^{3+} and Tm^{3+} , respectively. In blue, the untagged protein; in red, the protein tagged with $[\text{Dy}(\text{M8-SPy}) \mathbf{32}]$; and in green, the protein tagged with $[\text{Tm}(\text{M8-SPy}) \mathbf{31}]$. The spectra were recorded at 298 K and pH 6.8. Spectra by **KZ**.

To further investigate the interactions between the biomolecular scaffold and the catalyst in solution state, the pseudo-contact chemical shifts were recorded for a fluorinated inhibitor (Figure 4.20), *N*-(2,3-difluorobenzyl)-4-sulphamoylbenzamide (**61**), synthesised by **FM** (Scheme 4.9). This system was chosen as it had been previously reported by Christianson and co-workers (PDB code: 1G52).^[38]

Scheme 4.9. Structure of the fluorinated ligand-inhibitor used in NMR studies.



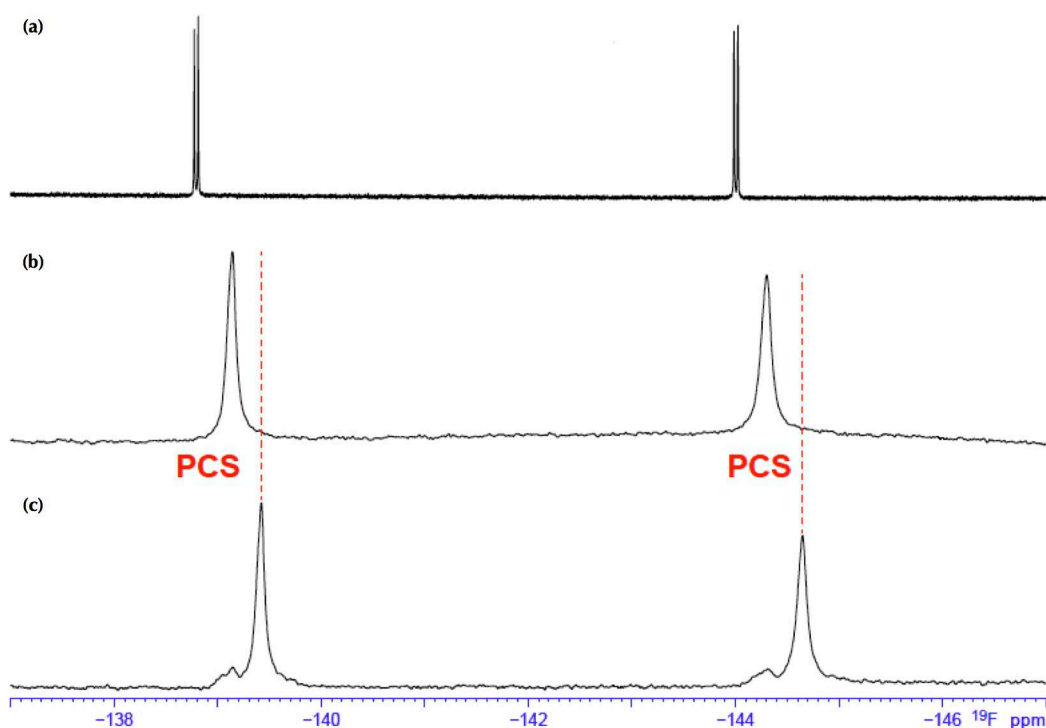


Figure 4.20. ${}^1\text{D } {}^{19}\text{F}$ spectra of inhibitor **61**. (a) The inhibitor without the protein. (b) The inhibitor in the presence of hCAII C206S-S166C. (c) The inhibitor in the presence of hCAII C206S-S166C tagged with [Tm(M8-SPy)]. The PCS are highlighted in red. Spectra by **KZ**.

The recorded 1D ${}^{19}\text{F}$ -NMR spectra of **61** \subset hCAII C206S-S166C labelled with [Tm(M8-SPy)] was in good agreement with the published x-ray structure. According to the experimental PCS and the initial calculations, the fluorine position deviate as much as 5 Å, from the expected position in the crystal structure.

4.3 Conclusion & outlook

Human carbonic anhydrase II was expressed in *E. coli* with the purpose of creating a new biomolecular scaffold to host an organometallic moiety. After optimisation of the production conditions (up- and downstream), up to 400 mg of pure protein per litre of culture was obtained in shake flasks. The screening of a small library of rationally designed mutants allowed the identification of I91 as an important residue for the enantioselective reduction of prochiral imines using $[(\text{Cp}^*)\text{IrCl}_2]_2$. X-ray crystal structure studies of the first generation of ligands confirmed that hCAII had a suitable hydrophobic cavity to anchor the complex. More recently, efforts were directed toward crystallographic investigations for the elucidation of the second generation of ligands embedded in hCAII. The structure provided more insight into the

origin of the observed enantioselectivity, and will allow further chemo-genetic modifications, in order to optimise the first and second coordination spheres of this novel hybrid catalyst. Docking experiments have proven to be a useful tool for rational design of the host protein. Indeed, the influence of mutations at a key position (L198x) in hCAII on ligand binding affinities were correctly predicted from atomistic simulations. The experimentally determined binding constants of the designed mutants for the reactive benzenesulphonamide correlated very well with the computed binding free energies. This indicates that docking could be used to some extent to construct hCAII variants with a designed specificity.

A rigid, high-affinity lanthanide chelating tag, [M8-SPy] was successfully bound to labelled hCAII. The pseudo-contact chemical shifts were determined by selectively ^{15}N Leucine labelled hCAII tagged with [Tm(M8-SPy)]. As a first approach to better understand the position and interactions between the inhibitor and the protein in solution, PCS of a difluorinated inhibitor bound to hCAII were obtained by simple 1D ^{19}F NMR spectra. The conformational information obtained from NMR with that obtained from x-ray crystallography indicated a deviation of 5 Å, between structures, confirming that crystal and solution structures of the inhibitor \subset protein differ. In order to further investigate this intriguing observation, an active fluorinated active catalyst was synthesised, and NMR studies are presently undergoing to solve the artificial metalloenzyme structure in solution. NMR and x-ray structure studies will glean more detailed and precise structural information on the position and orientation of the catalyst inside the protein cavity.

4.4 References

- [1] Lindskog, S.; Coleman, J. E. *Proceedings of the National Academy of Sciences of the United States of America* **1973**, *70*, 2505–2508.
- [2] Lindskog, S.; Liljas, A. *Current Opinion in Structural Biology* **1993**, *3*, 915–920.
- [3] Boriack-Sjodin, P. A.; Zeitlin, S.; Chen, H. H.; Crenshaw, L.; Gross, S.; Dantanarayana, A.; Delgado, P.; May, J. A.; Dean, T.; Christianson, D. W. *Protein Science* **1998**, *7*, 2483–2489.
- [4] Eriksson, A. E.; Jones, T. A.; Liljas, A. *Proteins* **1988**, *4*, 274–282.
- [5] Lindskog, S. *Zinc Enzymes*; John Wiley & Sons: New York, USA, 1983.
- [6] Reetz, M. T. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 5716–5722.
- [7] Mihovilovic, M. D. *Journal of Chemical Technology and Biotechnology* **2007**, *82*, 067–1071.
- [8] Steinberg, T. H. *Methods in Enzymology* **2009**, *463*, 541–563.
- [9] Cramer, F. *Perspectives in Supramolecular Chemistry*; volume 1 John Wiley & Sons, Ltd.: Chichester, United Kingdom, 2007.
- [10] Briesewitz, R.; Ray, G. T.; Wandless, T. J.; Crabtree, G. R. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 1953–1958.
- [11] Rami, M.; Winum, J.-Y.; Innocenti, A.; Montero, J. L.; Scozzafava, A.; Supuran, C. T. *Bioorganic & Medicinal Chemistry Letters* **2008**, *18*, 836–841.
- [12] Ang, W. H.; Parker, L. J.; De Luca, A.; Juillerat-Jeanneret, L.; Morton, C. J.; Lo Bello, M.; Parker, M. W.; Dyson, P. J. *Angewandte Chemie International Edition* **2009**, *48*, 3854–3857.
- [13] Nair, S. K.; Calderone, T. L.; Christianson, D. W.; Fierke, C. *The Journal of Biological Chemistry* **1991**, *266*, 17320–17325.
- [14] Mann, T.; Keilin, D. *Nature* **1940**, *146*, 164–165.
- [15] Supuran, C. T.; Scozzafava, A. *Bioorganic & Medicinal Chemistry* **2007**, *15*, 4336–4350.
- [16] Krishnamurthy, V. M.; Kaufman, G. K.; Urbach, A. R.; Gitlin, I.; Gudiksen, K. L.; Weibel, D. B.; Whitesides, G. M. *Chemical Reviews* **2008**, *108*, 946–1051.
- [17] Liljas, A.; Kannan, K. K.; Bergstén, P. C.; Waara, I.; Fridborg, K.; Strandberg, B.; Carlbom, U.; Järup, L.; Lövgren, S.; Petef, M. *Nature New Biology* **1972**, *235*, 131–137.
- [18] Stams, T.; Christianson, D. W. *The carbonic anhydrases: new horizons*; volume 90 of *Experimentia Supplementum* Birkhäuser Basel: Basel, Switzerland, 2000.

- [19] Alexander, R. S.; Nair, S. K.; Christianson, D. W. *Biochemistry* **1991**, *30*, 11064–11072.
- [20] Nayal, M.; Honig, B. *Proteins: Structure, Function and Bioinformatics* **2006**, *63*, 892–906.
- [21] Zuiderweg, E. R. P.; van Doren, S. R.; Kurochkin, A. V.; Neubig, R. R.; Majumdar, A. *Perspectives in Drug Discovery and Design* **1993**, *1*, 391–417.
- [22] Wider, G.; Wüthrich, K. *Current Opinion in Structural Biology* **1999**, *9*, 594–601.
- [23] Yu, H. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 332–334.
- [24] Häussinger, D.; Huang, J.-R.; Grzesiek, S. *Journal of the American Chemical Society* **2009**, *131*, 14761–14767.
- [25] Murakami, H.; Marelich, G. P.; Grubb, J. H.; Kyle, J. W.; Sly, W. S. *Genomics* **1987**, *1*, 159–166.
- [26] Studier, F. W.; Moffatt, B. A. *Journal of Molecular Biology* **1986**, *189*, 113–130.
- [27] Rosenberg, A. H.; Lade, B. N.; Chui, D. S.; Lin, S. W.; Dunn, J. J.; Studier, F. W. *Gene* **1987**, *56*, 125–135.
- [28] Henderson, L. E.; Henriksson, D.; Nyman, P. O. *The Journal of Biological Chemistry* **1976**, *251*, 5457–5463.
- [29] Zheng, L.; Baumann, U.; Reymond, J.-L. *Nucleic Acids Research* **2004**, *32*, e115.
- [30] Humbert, N. *Approche combinatoire de la production et de la purification de mutants de la streptavidine en vue de générer des métalloenzymes artificielles*, Thesis, University of Neuchâtel, 2005.
- [31] King, R. W.; Burgen, A. S. V. *Proceedings of the Royal Society of London. Series B, Biological Sciences* **1976**, *193*, 107–125.
- [32] Bering, C. L.; Kuhns, J. J. *Journal of Chemical Education* **1998**, *75*, 1021–1024.
- [33] Gould, S. M.; Tawfik, D. S. *Biochemistry* **2005**, *44*, 5444–5552.
- [34] Pocker, Y.; Stone, J. T. *Biochemistry* **1968**, *7*, 2936–2945.
- [35] Artimo, P. *et al. Nucleic Acids Research (webserver issue)* **2012**, *40*, W597–W603.
- [36] Hartinger, C. G.; Dyson, P. J. *Chemical Society Reviews* **2009**, *38*, 391–401.
- [37] Schrödinger, L. L. C. “Maestro”, version 9.3, 2011.
- [38] Kim, C.-Y.; Chang, J. S.; Doyon, J. B.; Baird Jr., T. T.; Fierke, C. A.; Jain, A.; Christianson, D. W. *Journal of the American Chemical Society* **2000**, *122*, 12125–12134.
- [39] Srivastava, D. K.; Jude, K. M.; Banerjee, A. L.; Haldar, M.; Manokaran, S.; Koren, J.; Mallik, S.; Christianson, D. W. *Journal of the American Chemical Society* **2007**, *129*, 5528–5537.

- [40] Drescher, D. G. *Analytical Biochemistry* **1978**, *90*, 349–358.
- [41] Yu, Z.; Xie, L.; Lee, S.; Zhang, R. *Comparative Biochemistry and Physiology, Part B* **2006**, *143*, 190–194.
- [42] Schmid, M.; Nogueira, E. S.; Monnard, F. W.; Ward, T. R.; Meuwly, M. *Chemical Science* **2012**, *3*, 690–700.
- [43] Monnard, F. W.; Heinisch, T.; Nogueira, E. S.; Schirmer, T.; Ward, T. R. *Chemical Communications* **2011**, *47*, 8238–8240.
- [44] Monnard, F. W.; Nogueira, E. S.; Heinisch, T.; Schirmer, T.; Ward, T. R. *Chemical Science* **2013**, *4*, 3269–3274.
- [45] Mårtensson, L. G.; Jonsson, B. H.; Freskgård, P. O.; Kihlgren, A.; Svensson, M.; Carlsson, U. *Biochemistry* **1993**, *32*, 224–231.
- [46] Burton, R. E.; Hunt, J. A.; Fierke, C. A.; Oas, T. G. *Protein Science* **2000**, *9*, 776–785.
- [47] Krebs, J. F.; Fierke, C. A. *The Journal of Biological Chemistry* **1993**, *268*, 948–954.

Summary & outlook

One never notices what has been done;
one can only see what remains to be
done.

Marie Skłodowska-Curie

Artificial metalloenzymes have emerged as a promising approach to merge beneficial aspects of bio- and homogeneous catalysis, and the scope of transformations enabled by metal-mediated reactions. Since activity and selectivity, including enantioselectivity, of natural (metallo)enzymes are due to the second coordination sphere interactions provided by the protein, artificial metalloenzymes aim at harnessing those interactions to create transition metal complexes that display catalytic activity for enantioselective syntheses. However, the design of new artificial metalloenzymes is rather challenging and extremely labour intensive, mainly due to the necessity to screen organometallic moieties on purified biomolecular scaffolds. Therefore, only very small mutant libraries (< 50 clones) can be handled in a slow and tedious overall process. More sophisticated screening methods are required for future reactions.

During the course of this thesis, various strategies to optimise production of streptavidin in *E. coli* were developed and tested as high-throughput screening methods for the asymmetric transfer hydrogenation and/or reductive amination reactions. The main aim was to reduce the time of protein production, and protect the organometallic catalyst from cell-based catalyst poisons (*e.g.* reduced glutathione). The first two approaches (protein precipitation and small-scale purification) rendered satisfactory results, with low to moderate overall reaction yields. The main obstacles, however, consisted of tedious individual purification of variants, and low amount of protein available from small-scale cultures. Presently, the best currently available expression system of streptavidin does not deliver enough protein to fulfil the requirements for efficient and extensive screening, in small-scale. Thus, it is not possible to implement these purification methods in *e.g.* microtiter plates. The last approach based on the neutralisation of glutathione present in cell-free protein extracts revealed very promising and reproducible results (> 40% conversion and > 50% enantioselectivities), and can be easily implemented in parallel screening. In order to develop a more efficient variation of this strategy, two options appeared to be logical: (i) improve the expression system for streptavidin in *E. coli* considerably, since low concentration of the biomolecular scaffold leads to lower conversions

compared to standard conditions; or (ii) turn to an alternative expression system that would yield high levels of secreted protein in the supernatant, allowing directed evolution, high-throughput expression, direct screening in the supernatant, and thus evolution of hybrid catalysts.

To fulfil the listed requirement, *Pichia pastoris* was tested as expression system for heterologous production of streptavidin. High-level of mature Sav was obtained in a fed-batch fermentation. Enantioselective catalysis performed by an artificial metalloenzyme created with crude supernatant of Sav expressed in *Pichia pastoris* was investigated and evaluated. Success of its application, however, was limited owing to the low-level of secreted protein before concentration. Similar catalysis results were obtained with purified protein expressed either in *E. coli* or *P. pastoris*. Further studies on Sav expressed in *P. pastoris*, mutagenesis of critical potential *O*-glycosylation sites, and optimisation of the culture parameters should be conducted. The use of cell-free based expression systems should also be investigated as it would allow to overcome processing bottlenecks, *e.g.* presence of GSH and biotin.

Human carbonic anhydrase II was investigated as potential scaffold for protein-based hybrid catalysts. The expression of hCAII was optimised up to a degree that high yield was achieved in shake flasks cultures in the order of 400 mg per litre, thus protein concentration was not an issue. A successful example of this system was the construction of a highly enantioselective imine reductase by chemo-genetic optimisation. A promising 29% *ee* was found for the imine reduction reaction of a salsolidine precursor using one of the created constructs. The combined use of rational, structural and computational design (designed evolution) had great advantages for the production of artificial metalloenzymes with novel and improved activity. Further development of these catalyst systems with use of both synthetic (*e.g.* optimisation of ligand structure) and biomolecular tools (*e.g.* optimisation of protein environment) for optimisation can lead, in the future, to very efficient and enantioselective conversions. Current efforts aim to solve the NMR structure of this hybrid catalyst to gain insights of the system in solution-state.

In summary, in order to generate and identify protein variants to enhance enantioselectivity in artificial metalloenzymes, directed evolution, screening and selection techniques have to be improved. State-of-the-art libraries have to be generated to make engineering of biomolecular scaffolds more efficient and effective. To this end, library construction techniques, which combine classical methods such as error-prone PCR and rational methods, have to be combined with computational, structural and statistical approaches. A colorimetric or fluorescence assay should be developed to ease the process of screening of large libraries accurately.

Advances in the artificial metalloenzyme field would bring the scientific community closer to realising the dream of tailor-made hybrid catalysts with high catalytic efficiency and selectivity for biotechnological and pharmaceutical applications.

Materials & methods

In Science, we must be interested in things, not in persons.

Marie Skłodowska-Curie

6.1 General experimental section

6.1.1 Standard methods & reagents

Standard methods, such as primers design, DNA extraction and purification, polyacrylamide gel electrophoresis, protein chromatography (affinity, anion exchange, gel filtration), and spectrophotometric quantification of cell growth were carried out according to published protocols, with minor adjustments (when necessary).

DNase I was purchased from Roche Diagnostics AG (Basel, Switzerland), pre-stained protein marker was either “Broad Range” from New England BioLabs Inc. (Allschwill, Switzerland) or “BenchMark” from Invitrogen (Carlsbad, CA USA), primers were ordered from Microsynth AG (Balgach, Switzerland), ampicillin and chloramphenicol were purchased from Applichem GmbH (Darmstadt, Germany) and Zeocin from Invitrogen (Carlsbad, CA USA), *Pfu* Turbo polymerase was from Stratagene (La Jolla CA, USA), isopropyl- β -D-1-thiogalactopyranoside (IPTG) and dithiothreitol (DTT) from Apollo Scientific (Stockport, UK) and biotin-4-fluorescein (B4F) from ANAWA Trading SA (Zurich, Switzerland). All other standard chemical and biological reagents were purchased from Sigma-Aldrich or Fluka (St. Louis, MO USA), Acros Organics (Geel, Belgium), Fischer Chemicals AG (Zurich, Switzerland), LuBio Science (Lucerne, Switzerland), Affymetrix (Santa Clara, CA USA), Applichem (Darmstadt, Germany), Merck (Billerica, MA USA) or Thermo Scientific (Waltham, MA USA). Suppliers of more specific kits and reagents are referenced throughout the text, in their respective section. Reagents and solvents were of the highest commercially available grade and used without further purification, and solutions were prepared in deionised water, unless otherwise indicated.

6.1.2 Equipment

Polymerase chain reactions (PCR) were performed in an Eppendorf Mastercycler Gradient (Hamburg, Germany). DNA concentration was determined on ThermoFisher Scientific Nan-

oDrop 1000 spectrophotometer (Waltham, MA USA). Electroporation was done using Bio-Rad GenePulser electroporator. Polyacrylamide (non- and denaturing) and agarose gel electrophoresis were carried out using Bio-Rad Mini-PROTEAN and Sub-Cell systems, respectively. Resulting gels were analysed with Molecular Image Gel Doc XR (Reinach, Switzerland). Western blot was carried out on Biometra Fastblot B34 system (Châtel-Saint-Denis, Switzerland) connected to a BioRad PowerPac HC system. Blots were analysed with GE Healthcare Life Sciences ImageQuant RT ECL imager (Glattbrugg, Switzerland).

Fermentations carried out at the University of Basel (Unibas) were performed in a 20 L working volume NLF22 fermentor, from BioEngineering AG (Wald, Switzerland). Fermentation conducted at the Paul Scherrer Institute (PSI Ost Villigen, Switzerland) was carried out in a 1.5 L working volume, custom-made fermentor also from BioEngineering AG. Shake flasks and 96-well plate cultures were carried out in either an Infors HT Ecotron (Bottmingen, Switzerland) or a New Brunswick Scientific INNOVA 44 (Edison NJ, USA).

Electronic absorption of cell growth were recorded on a Agilent | Varian Inc. Cary UV-Vis 5000 spectrophotometer (Englewood, CO USA), and electronic absorption and emission were recorded on Tecan Safire microplate reader (Männedorf, Switzerland) running SoftMax Pro software. Streptavidin expressed in *Pichia pastoris* was concentrated by tangential flow filtration, using Merck Millipore Prep/Scale Spiral Wound TFF-6 Module PLGC 10K Regenerated cellulose system (Billerica, MA USA) connected to a Verderflex peristaltic pump from Verder Ltd (Essex, UK). ÄKTAprime Plus (Unibas) and ÄKTAExpress (PSI) chromatography systems (FPLC), from GE Healthcare (Glattbrugg, Switzerland), were used in the purification of proteins. Purified proteins were lyophilised in a Labonco benchtop lyophiliser FreeZone 2.5 L (Kansas City, MO USA). The determination of the molecular mass of the proteins was performed by electrospray ionisation time-of-flight mass spectrometry, ESI-TOF MS, at Unibas Biozentrum (Basel, Switzerland). The instrument used was a Bruker Daltonics, micrOTOF, ESI/TOF MS benchtop type (Fällanden, Switzerland). The determination of the molecular mass and N-terminal sequencing of streptavidin expressed in *Pichia pastoris* was performed by the “Functional Genomics Center Zurich”, from the University of Zurich (Switzerland).

Catalysis reactions were performed in a magnetically stirred multireactor, RR 98072, Radleys Discovery Technologies (Essex, UK), a HLC Biotech heating thermomixer MR23 from Huber & Co, AG (Reinach, Switzerland), or in an incubator shaker (Infors HT Ecotron or New Brunswick Scientific INNOVA 44).

The conversion (conv.) and enantiomeric excess (*ee*) of the reduction products were determined by chiral phase HPLC, in an Agilent or Hewlett Packard 1100 Series systems equipped with a G1322A degasser, a G1312A binary pump, a G1316A column compartment, a G1315A diode array detector and a G1329A autosampler unit (Englewood, CO USA).

6.2 Experimental section of Chapter 2

6.2.1 General procedure for the production of streptavidin

Some methods described herein were developed and/or optimised by previous members of the Ward group, namely Dr Nicolas Humbert and Dr Alessia Sardo.^[1,2]

Cloning & mutagenesis

pET11b-Sav plasmid

The original pET11b-Sav plasmid was a generous gift from Prof. Santambrogio (University of Milan, Italy) who cloned streptavidin (Sav) coding sequence downstream of the T7 promoter^[3] and gene 10 leader sequences (Figure 6.1). Cloning the gene of interest using BamHI site resulted in the replacement of the first 13 amino acids of the N-terminal by an immunological marker, the T7 epitope tag, which increases the solubility of recombinant Sav in the cytoplasm of *Escherichia coli*.

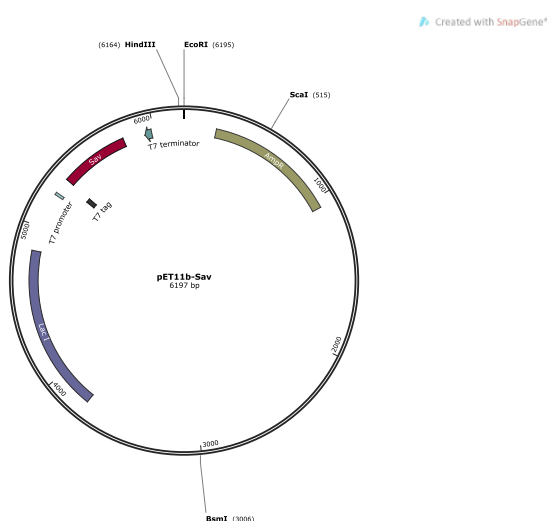


Figure 6.1. Plasmid map of streptavidin gene in pET11b vector. In the absence of the molecular mimic of allo-lactose, IPTG, the lac repressor (in purple) binds to the lac operon, preventing the binding of *E. coli* RNA polymerase and, consequently, the transcription of T7 RNA polymerase (blue) and streptavidin gene (red).

In this study, the construct of T7-tagged streptavidin was transformed into *Escherichia coli* BL21(DE3)pLysS strain (genotype: $F^- ompT lon hsdS_B(r_B^- m_B^-) dcm gal \lambda(DE3) [pLysS (cm^r)]$), produced in-house) for high-level expression of recombinant protein. This strain is deficient in proteases *lon* and *ompT*, and it confers increased stability and limits the proteolytic cleavage of the expressed protein. BL21(DE3)pLysS is lysogenic for λ -DE3, which

carries a chromosomal copy of the viral T7 bacteriophage gene I, encoding T7 RNA polymerase under the control of the lac UV5 promoter (inducible protein expression by addition of isopropyl- β -D-1-thiogalactopyranoside, IPTG). pLysS plasmid carries the gene encoding T7 lysozyme, which lowers the background expression level of target genes under the control of the T7 promoter without interfering with the level of expression achieved following induction by IPTG.

Site-directed mutagenesis

pET11b plasmid containing a BamHI cloning site as well as a gene encoding for streptavidin fused with a T7-tag was used as a template for PCR. Primers were designed following the method described by Zheng *et al.*,^[4] and tested *in silico* to minimise hairpin formation.^[5] Primers were obtained from Microsynth (Balgach, Switzerland). PCR reactions were prepared by addition of 5 μ L 10x *Pfu* buffer (Stratagene), 2 μ L of 10 mM dNTP, 2.5 μ L DMSO (final concentration 5%), 1.5 μ L *Pfu* Turbo polymerase (Stratagene), 1.5 μ L of 10 mM primers (sense and antisense), 35 μ L H₂O to 1 μ L of template. The cycle conditions were: initial denaturation (95 °C, 5 min), followed by 16 or 25 cycles of 1 min at 95 °C; 1 min at 60 °C (or adjusted to the melting temperature of the primers), and a final elongation at 68 °C for 1 h. The resulting PCR product was kept at 10 °C until further analysis. PCR products were analysed by 0.7% agarose gel electrophoresis (1.4 g agarose in 0.5% v/v Tris-borate-EDTA (TBE) and 10 μ L ethidium bromide (10 mg/mL stock solution)) after the initial DNA template (wild-type sequence) was digested by DpnI (4 h at 42 °C). The gel was analysed under fluorescent light, using Bio-Rad Gel Doc XR+ software. Ultra-competent XL1-Blue *E. coli* cells (genotype: *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lac1^q Z Δ M15 Tn10(Tet^r)*], produced in-house) were transformed with 5 μ L of PCR product. Plasmids were purified using Promega Wizard Plus SV Miniprep DNA purification system (Dübendorf, Switzerland), and were sequenced either by Starseq (Mainz, Germany) or Microsynth (Balgach, Switzerland). The list of primers (sense and antisense) used for each project of this Chapter, and their characteristics (melting temperature, silent mutation, etc.) are listed in Table 6.1 to Table 6.3.

Table 6.1. List of primers (s: sense; a: antisense) for the construction of Sav mutants, used as templates for a second round of mutagenesis. **wt Sav** was used as template (*e.g.* Sav S112A stands for the serine residue at position 112 mutated to an alanine). In lowercase, the mutation introduced, and in *italic*, a silent mutation (to avoid hairpin formation and/or self-annealing).

Mutant	Primers (5'→3')	T _m (°C)	Length (bases)
S112H	s G CTG ACC <i>cat GGt</i> ACC ACC GAG GCC AAC GCC TGG	72.9	34
	a GTT GTG GGT CAC CGA CGA CTG GGT <i>aCC</i> atg GTG G	70.4	34
K121H	s C GCC TGG <i>cac</i> TCC ACG CTG GTC GGC CAC GAC ACC	75.3	34
	a CGT GGA <i>gtg</i> CCA GGC GTT GGC CTC GGT GGT GCC	74.3	33

Table 6.2. List of primers (s: sense; a: antisense) the second round of mutagenesis. **Sav S112H** was used as template. In lowercase, the mutation introduced, and in *italic*, a silent mutation (to avoid hairpin formation and/or self-annealing).

Mutant		Primers (5'→3')	Tm (°C)	Length (bases)
K121W	s	GCC TGG <i>tg</i> g TCC ACG CTG GTC GGC CAC GAC ACC	74.7	33
	a	CGT GGA <i>cca</i> CCA GGC GTT GGC CTC GGT GGT GCC	74.7	33
K121A	s	GCC TGG <i>gc</i> g TCC ACG CTG GTC GGC CAC GAC ACC	75.6	33
	a	CGT GGA <i>cgc</i> CCA GGC GTT GGC CTC GGT GGT GCC	75.6	34
K121F	s	C GCC TGG <i>ttc</i> TCC ACG CTG GTC GGC CAC GAC ACC	74.1	34
	a	GT GGA <i>gaa</i> CCA GGC GTT GGC CTC GGT GGT GCC	72.1	32
K121D	s	C GCC TGG <i>gac</i> TCC ACG CTG GTC GGC CAC GAC ACC	75.3	34
	a	GT GGA <i>gtc</i> CCA GGC GTT GGC CTC GGT GGT GCC	73.4	34
K121C	s	C GCC TGG <i>tgt</i> TCC ACG CTG GTC GGC CAC GAC ACC	74.1	34
	a	CGT GGA <i>aca</i> CCA GGC GTT GGC CTC GGT GGT GCC	73.1	33
K121E	s	GCC TGG <i>gag</i> TCC ACG CTG GTC GGC CAC GAC ACC	74.3	33
	a	CGT GGA <i>ctc</i> CCA GGC GTT GGC CTC GGT GGT GCC	74.3	33
K121R	s	GCC TGG <i>cgg</i> TCC ACG CTG GTC GGC CAC GAC ACC	75.6	33
	a	CGT GGA <i>ccg</i> CCA GGC GTT GGC CTC GGT GGT GCC	75.6	33
N49A	s	GTC GGC <i>gcc</i> GCC GAG AGC CGC TAC GTC CTG	73.9	30
	a	T CTC GGC <i>ggc</i> GCC GAC GGC CGA CTC GTA GG	73.9	30
N49S	s	GTC GGC <i>tcc</i> GCC GAG AGC CGC TAC GTC CTG	72.6	30
	a	T CTC GGC <i>gga</i> GCC GAC GGC CGA CTC GTA GG	72.6	30
N49W	s	GTC GGC <i>tg</i> g GCC GAG AGC CGC TAC GTC CTG	73.2	30
	a	T CTC GGC <i>cca</i> GCC GAC GGC CGA CTC GTA GG	73.2	30
N49F	s	GTC GGC <i>ttt</i> GCC GAG AGC CGC TAC GTC CTG	70.7	30
	a	T CTC GGC <i>aaa</i> GCC GAC GGC CGA CTC GTA GG	70.7	30
N49D	s	GTC GGC <i>gat</i> GCC GAG AGC CGC TAC GTC CTG	71.9	30
	a	T CTC GGC <i>atc</i> GCC GAC GGC CGA CTC GTA GG	71.9	30
N49E	s	GTC GGC <i>gaa</i> GCC GAG AGC CGC TAC GTC CTG	71.9	30
	a	T CTC GGC <i>ttc</i> GCC GAC GGC CGA CTC GTA GG	71.9	30
N49R	s	GTC GGC <i>cgt</i> GCC GAG AGC CGC TAC GTC CTG	73.2	30
	a	T CTC GGC <i>acg</i> GCC GAC GGC CGA CTC GTA GG	73.2	30
N49C	s	GTC GGC <i>tgt</i> GCC GAG AGC CGC TAC GTC CTG	71.9	30
	a	T CTC GGC <i>aca</i> GCC GAC GGC CGA CTC GTA GG	71.9	30
L124A	s	GG AAG TCC ACG <i>gca</i> GTC GGC CAC GAC ACC TTC ACC	72.7	35
	a	G GCC GAC <i>g</i> gc CGT GGA CTT CCA GGC GTT GGC CTC GG	75.5	36

Table 6.3. List of primers (s: sense; a: antisense) the second round of mutagenesis. **Sav K121H** was used as template. In lowercase, the mutation introduced, and in *italic*, a silent mutation (to avoid hairpin formation and/or self-annealing).

Mutant		Primers (5'→3')	T _m (°C)	Length (bases)
S112A	s	G CTG ACC <i>gcc</i> <i>GGt</i> ACC ACC GAG GCC AAC GCC TGG	75.3	34
	a	GTT GTG GGT CAC CGA CGA CTG GGT <i>aCC</i> <i>ggc</i> GTG G	72.9	34
S112W	s	G CTG ACC <i>tgg</i> GGT ACC ACC GAG GCC AAC GCC TGG	74.1	34
	a	GGT GGT ACC <i>cca</i> GGT CAG CAG CCA CTG GGT GTT G	71.7	34
S112E	s	G CTG ACC <i>gag</i> GGC ACC ACC GAG GCC AAC GCC TGG	75.3	34
	a	GGT GGT GCC <i>ctc</i> GGT CAG CAG CCA CTG GGT GTT G	72.9	34
S112R	s	G CTG ACC <i>cgt</i> GGC ACC ACC GAG GCC AAC GCC TGG	75.3	34
	a	GGT GGT GCC <i>acg</i> GGT CAG CAG CCA CTG GGT GTT G	72.9	34
S112C	s	G CTG ACC <i>tgc</i> GGC ACC ACC GAG GCC AAC GCC TGG	75.3	34
	a	GGT GGT GCC <i>gca</i> GGT CAG CAG CCA CTG GGT GTT G	72.9	34
S112F	s	G CTG ACC <i>ttt</i> GGC ACC ACC GAG GCC AAC GCC TGG	73.1	34
	a	GGT GGT GCC <i>aaa</i> GGT CAG CAG CCA CTG GGT GTT G	70.6	34
S112D	s	G CTG ACC <i>gat</i> GGC ACC ACC GAG GCC AAC GCC TGG	74.3	34
	a	GGT GGT GCC <i>atc</i> GGT CAG CAG CCA CTG GGT GTT G	71.9	34
S112K	s	G CTG ACC <i>aag</i> GGC ACC ACC GAG GCC AAC GCC TGG	75.5	34
	a	GGT GGT GCC <i>ctt</i> GGT CAG CAG CCA CTG GGT GTT G	73.1	34
L110A	s	G TGG CTG <i>gca</i> ACC TCC GGC ACC ACC GAG GCC AAC	75.3	36
	a	GCC GGA GGT <i>cgc</i> CAG CCA CTG GGT GTT GAT CCT CGC	74.7	36
L110E	s	GG CTG <i>gag</i> ACC TCC GGC ACC ACC GAG GCC AAC GCC	76.1	35
	a	GCC GGA GGT <i>ctc</i> CAG CCA CTG GGT GTT GAT CCT CG	72.6	35
L110K	s	GG CTG <i>aag</i> ACC TCC GGC ACC ACC GAG GCC AAC GCC	75.0	35
	a	GC CTC GGT GGT GCC GGA GGT <i>ctt</i> CAG CCA CTG GG	74.1	34
L110W	s	G TGG CTG <i>gca</i> ACC TCC GGC ACC ACC GAG GCC AAC GC	75.5	36
	a	GCC GGA GGT <i>tgc</i> CAG CCA CTG GGT GTT GAT CCT CGC	73.5	36
L110F	s	G TGG CTG <i>ttc</i> ACC TCC GGC ACC ACC GAG GCC AAC GC	73.5	36
	a	GCC GGA GGT <i>gaa</i> CAG CCA CTG GGT GTT GAT CCT CGC	71.0	36
L110D	s	G TGG CTG <i>gat</i> ACC TCC GGC ACC ACC GAG GCC	74.3	36
	a	GCC GGA GGT <i>atc</i> CAG CCA CTG GGT GTT G	72.3	36
L110R	s	G TGG CTG <i>agg</i> ACC TCC GGC ACC ACC GAG GCC	75.5	36
	a	GCC GGA GGT <i>cct</i> CAG CCA CTG GGT GTT G	73.5	36
L110C	s	G TGG CTG <i>tgt</i> ACC TCC GGC ACC ACC GAG GCC AAC GC	74.3	36
	a	GCC GGA GGT <i>aca</i> CAG CCA CTG GGT GTT GAT CCT CGC	72.3	36

Expression in 50 mL cultures

This protocol describes how to express in parallel several Sav variants, in 50 mL culture, in suitable amounts for enantioselective catalysis screening.

For high gene expression, the plasmid containing Sav gene was transformed into ultra-competent *E. coli* BL21(DE3)pLysS cells.

After thawing on ice 100 μ L of BL21(DE3)pLysS cells, 8 μ L of dithiothreitol (DTT, 200 mM stock) and 3 μ L of plasmid (0.2 – 0.5 μ g of DNA) were added and mixed gently. The mixture was left on ice for 15 min, and then plated on pre-warmed lysogeny broth (LB-Miller) plates (containing 60 μ g/mL ampicillin (amp), 34 μ g/mL chloramphenicol (cm), and 2% w/v glu-

cose). The plates were inverted and incubated overnight at 37 °C, for plasmid multiplication. All solutions used were autoclaved (121 °C, 15 min) or filter-sterilised (0.22 µm sterile filter).

Expression in TP medium

[Inoculum] After overnight incubation of the plate at 37 °C, a medium sized colony was chosen to inoculate 5 mL TP medium (20 g/L bactotryptone, 15 g/L bacto yeast extract, 8 g/L NaCl, 2g/L Na₂HPO₄, and 1 g/L KH₂PO₄), containing the appropriate amounts of antibiotics (60 µg/mL amp and 34 µg/mL cm), and 2% w/v glucose. The 50 mL baffled shake flask was incubated overnight in an orbital shaker, at 37 °C and 250 rpm.

[Culture] A baffled shake flask containing 50 mL of cell culture medium (same composition as the inoculum medium, 60 µg/mL amp and 34 µg/mL cm, but no glucose) was inoculated with the overnight inoculum (1:10 dilution). Cells were grown at 37 °C and 415 rpm, for 3 h. Addition of 0.4 mM IPTG induced protein expression. A sample of the culture (1 mL) was taken just prior addition of IPTG, and before harvesting. The culture was centrifuged (4,400 rpm, 10 min, 4 °C), the supernatant deactivated (by addition of sodium hypochlorite) and discarded, and cell pellets were fast-frozen in liquid nitrogen and kept at -20 °C until further analysis/work-up.

[Glycerol stocks] Glycerol stocks of the variants were prepared by mixing 500 µL of the remaining pre-culture with 500 µL of 100% glycerol. The mixture was homogenised by vortexing, flash-frozen in liquid nitrogen, and stored at -80 °C for further use.

Expression in auto-induction medium

[Inoculum] After overnight incubation of the plate at 37 °C, a medium sized colony was chosen to inoculate 5 mL LB medium (10 g/L bactotryptone, 5 g/L bacto yeast extract, and 10 g/L NaCl), containing the appropriate amounts of antibiotic (60 µg/mL amp and 34 µg/mL cm), and 2% w/v glucose. The 50 mL baffled shake flask was incubated overnight in an orbital shaker, at 37 °C and 250 rpm.

[Culture] One litre culture medium was prepared as follow: 7.1 g/L Na₂HPO₄, 6.8 g/L KH₂PO₄, 20 g/L bactotryptone, 5 g/L bacto yeast extract, 5 g/L NaCl were dissolved in 960 mL deionised water, and autoclaved (liquid cycle, 121 °C, 15 min). The following solutions were prepared and autoclaved/filter-sterilised separately, and added to the culture medium prior to inoculation: 60% v/v glycerol (10 mL, final concentration: 0.1%), 10% w/v glucose (5 mL, final concentration: 0.05%), 8% w/v lactose (25 mL, final concentration: 0.2%), 1 M MgSO₄ (2 mL, final concentration: 2 mM), 60 µg/mL amp, and 34 µg/mL cm.

The autoinduction medium (50 mL) was inoculated with the whole overnight inoculum, and incubated at 37 °C until OD₆₀₀ reached 0.2 – 0.3. The shake flasks were then transferred to

another orbital shaker pre-heated to 23 °C (induction point)¹, and the culture incubated for another 35 h. Samples of the culture (1 mL) were taken just before temperature induction, and every 5 h until the end of the expression. The culture was centrifuged (4,400 rpm, 10 min, 4 °C), the supernatant deactivated (by addition of sodium hypochlorite) and discarded, and cell pellets were fast-frozen in liquid nitrogen, and kept at -20 °C until further analysis/work-up.

Expression in 20 L fermentor

[Seed culture] After overnight incubation of the plate at 37 °C, a medium sized colony or a tip dipped in the 25% glycerol stock inoculated 300 mL sterile TP medium (20 g/L bactotryptone, 15 g/L bacto yeast extract, 8 g/L NaCl, 2g/L Na₂HPO₄, and 1 g/L KH₂PO₄), containing the appropriate amounts of antibiotics (60 µg/mL amp and 34 µg/mL cm), and 2% w/v glucose. The 1 L baffled shake flask was incubated overnight in an orbital shaker, at 37 °C and 250 rpm.

[Fermentation] The fermentation was carried out in a 30 L capacity BioEngineering NFL22 fermentor (Wald, Switzerland). 20 L sterile TP medium (recipe aforementioned) were inoculated with the whole seed culture (*i.e.* 300 mL). Cells were grown at 37 °C and 1,000 rpm, for 3 hours or until OD₆₀₀ reached 1.8 – 2.2. Addition of 0.4 mM IPTG induced protein expression. A sample of the culture (1 mL) was taken just prior addition of IPTG, and every hour until harvesting. Three hours after induction, the broth was harvested, centrifuged (4,400 rpm, 10 min, 4 °C), the supernatant deactivated (by addition of sodium hypochlorite) and discarded, and cell pellets were fast-frozen in liquid nitrogen, and kept at -20 °C until further analysis/work-up.

6.2.2 Purification procedure

Preparation of samples

From a large-scale expression (20 L fermentation)

Streptavidin variants were expressed in *E. coli*, and the harvested cell pellets frozen at -20 °C for a minimum of 15 h. Wet cell pellets were thawed and resuspended in Tris-HCl buffer (20 mM, pH 7.4, 500 – 1000 mL for 150 – 450 g pellet). Two to 3 mg of DNase I (Roche Diagnostics, Switzerland) and 1 mM phenylmethylsulphonyl fluoride (PMSF) were added, and the resuspended cells were incubated at RT under vigorous shaking until total degradation of

¹To lower the temperature of an orbital shaker from 37 °C to 23 °C in a short period of time was not possible, thus compromising the effect of temperature (heat-shock) in the protein induction. The solution was to transfer the culture from one shaker to another, pre-warmed to 23 °C, thus ensuring the desired effect of temperature.

nucleic acids. The sample was dialysed against 6 M guanidinium hydrochloride pH 1.5 (24 h, at RT). The proteic extract was centrifuged (18,000 x *g*, 30 min, 4 °C) to remove cell debris. Two further dialyses were performed (20 mM Tris-HCl, pH 7.4, followed by 50 mM Na₂CO₃, 0.5 M NaCl, pH 9.8, 4 °C) to prepare the proteic extract for affinity chromatography.

From a small-scale expression (50 mL culture in shake flasks)

The pellet (~ 3 g from a 50 mL culture) was thawed and resuspended in deionised water (three times wet cell weight in mL). One microlitre of 10 mg/mL DNase I (Roche Diagnostics, Switzerland) and 1 mM PMSF were added, and the resuspended cells were incubated first at 37 °C for 2 h, and then at RT under vigorous shaking until total degradation of nucleic acids. The proteic extract was centrifuged (18,000 x *g*, 30 min, 4 °C) to remove cell debris. pH was adjusted to 9.8 by addition of 5 M Na₂CO₃, pH 9.8 to prepare the proteic extract for affinity chromatography.

By affinity chromatography

Affinity chromatography on ÄKTAprime Plus chromatography system

The dialysed protein extract was filtered (Whatman paper filter) and applied to a 2-iminobiotin sepharose (Affiland, Belgium) column equilibrated with three column volumes (CV) of binding buffer (50 mM Na₂CO₃ pH 9.8, and 500 mM NaCl). The column was washed with seven CV of binding buffer, and streptavidin was eluted with 1% acetic acid (5 CV). A final three CV wash was performed with binding buffer. Collected fractions (10 mL) were pooled and immediately dialysed against Tris-buffer (10 mM Tris-HCl, pH 7.4) for 24 h at 4 °C, distilled water for 24 h and finally two times against ultrapure water for 24 h, at 4 °C. The purified and dialysed protein was flash-frozen in liquid nitrogen, lyophilised, and stored at 4 °C until use.

Affinity chromatography on a 24-well plate format (Small-Scale Purification, SSP)

The extract (1.5 mL) was applied to 400 µL/well 2-iminobiotin sepharose (Affiland, Belgium) packed in a 24-deep well plate (in-house fabricated), and equilibrated at pH 9.8 (50 mM Na₂CO₃ pH 9.8, and 500 mM NaCl). Streptavidin was eluted with 1% v/v formic acid. The pH was immediately corrected with 5 M NaOH (to pH 5.0) or 32% NH₃ (to pH 8.0), for the imine reduction and reductive amination reactions, respectively. The purified protein was quantified on a Nanodrop spectrophotometer ($\epsilon = 41,490 \text{ M}^{-1}\text{cm}^{-1}$, MW_{monomer} calculated from ExpASy Protparam tool),^[6] and shock frozen using liquid nitrogen, lyophilised, and stored at 4 °C until use. The protein was resuspended in the appropriate buffer used in catalysis (MOPS/formate for the imine reduction and HCOONH₄ for the reductive amination),

and the final working concentration was confirmed by Nanodrop.

By protein precipitation

Resuspended cell-free protein extracts (2.5 mL) were transferred to fresh 15 mL tubes, and ice-cold pure ethanol (7.5 mL) was added at room temperature whereupon precipitation occurred. The tubes were left to stand at room temperature for 10 min before centrifugation (4,400 rpm, 10 min, 20 °C), and subsequent removal of the supernatant. The pellets were washed (1x 2 mL 90% ethanol, 1x 2 mL 80% ethanol), resuspended in the reaction buffer (400 μ L), and the solids separated by centrifugation (4,400 rpm, 10 min, 20 °C). The quantity of biotin-free binding sites in the resulting supernatants was determined by back-titration with biotin-4-fluorescein (Section 6.2.4).

6.2.3 Characterisation of recombinant streptavidin

Analysis by electrophoresis

The samples taken from the culture and/or fermentation were treated in the same way. All gels used in Chapter 2 were 12% SDS-PAGE, and were freshly cast for all experiments. Samples were not heated prior electrophoresis, nor did the loading buffer contain DTT.

The 12% resolving layer solution was prepared using 6 mL 30% acrylamide/bis solution, 5 mL ultrapure water, 3.8 mL resolving buffer made with 1.5 M Tris-HCl (pH 8.8), and 75 μ L 20% SDS. Polymerisation was initiated upon addition of 100 μ L 15% ammonium persulphate (APS, in dH₂O) and 6 μ L tetramethylethylenediamine (TEMED). The stacking layer (4% acrylamide final concentration) was prepared by mixing 1 mL 30% acrylamide/bis solution, 3.4 mL ultrapure water, 1.5 mL stacking buffer (0.5 M Tris-HCl, pH 6.8) and 30 μ L 20% SDS. Polymerisation was initiated by adding 40 μ L of 15% APS and 6 μ L TEMED. The stacking gel was poured on top of the polymerised resolving gel, the combs were inserted, and the gel was left for polymerisation, for 30 min. When ready, the gel was clamped to the chambers, and placed into the tank containing 1X SDS buffer (25 mM Tris-HCl, 0.192 M glycine, and 0.1% w/v SDS), to prevent dryness of the gel.

Preparation of samples

Cells were lysed by activating the gene encoding T7 lysozyme using three cycles of “freezing/thawing”, and were resuspended in a volume (in μ L) of ultrapure water equivalent to 40 times the optical density (600 nm) of the culture at the moment of sampling, supplemented with DNase I (1 μ L of 1 mg/mL in 5 mM Tris-HCl pH 7.5, 75 mM NaCl, 0.5 mM MgCl₂, and 50% v/v glycerol). The samples were vortexed, and incubated at RT, under vigorous shaking

(rotating shaker, 300 rpm), until complete digestion of the nucleic acids (30 – 60 min). The bacterial extracts were then centrifuged (14,000 rpm, 5 min at RT). The soluble fraction (supernatant) was transferred to a new microcentrifuge tube, and the insoluble fractions were resuspended by vortexing, in the same volume of ultrapure water as previously (*i.e.* in μL , $40\times \text{OD}_{600}$).

SDS-PAGE by B₄F detection

Twenty microlitre of each soluble and insoluble fractions, and a positive control (20 μL of pure protein diluted in ultrapure water, to a final concentration of 1 mg/mL) were mixed with 1 μL 0.6 mM biotin-4-fluorescein (B₄F, dissolved in DMSO), and incubated for 10 min at RT. B₄F dye bound to Sav makes the protein easily visible when exposed to UV light, and confirms the biotin-binding of the protein. 10 μL 3X loading buffer (50 mM Tris-HCl pH 6.8, 1% w/v SDS, 2% v/v β -mercaptoethanol, 10% w/v sucrose, 0.006% w/v bromophenol blue) were added to each samples, which were quickly vortexed to ensure homogeneous mixing of the sample and the loading buffer, and immediately charged on the gels. The gels ran at a constant voltage of 120 V, for 90 min. After running, the gels were removed from the electrophoresis apparatus, and placed directly onto the transilluminator, where they were exposed for a few seconds ($\sim 10 - 20$ sec) to UV light. The protein binding activity is easily identified by the reaction of B₄F in the presence of the Sav, which yields bands that emit sufficiently in the visible region.

SDS-PAGE by Coomassie Blue staining

After resolving the gels by UV light, they were transferred to a plastic box containing staining solution (100 mL, 0.25% w/v Coomassie Brilliant Blue R-250, 50% v/v methanol, and 7.5% v/v glacial acetic acid), and incubated for 2 h or overnight, under gentle rocking. The staining solution was discarded and replaced by destaining solution (100 mL, 20% v/v methanol, 10% v/v acetic acid), and the gel was incubated, on the rocker, for another 3 h or until the protein pattern started to be visible. The gel was left overnight in deionised water; a Kimwipe placed in the solution rapidly removed the excess of stain in the solvent. The destained gels were then photographed for analysis.

6.2.4 Quantification methods

B₄F titration

The number of free binding sites of streptavidin was measured by the specific quenching of biotin-4-fluorescein in the presence of the protein, which is detectable by measurement of its fluorescence at 485 nm excitation (12 nm slit) and 520 nm emission wavelength (12 nm slit), as described by Kada *et al.*^[7]

For the measurement of free binding sites, 100 μM of pure tetrameric Sav was resuspended in deionised water. A working solution of 2 μM Sav was prepared by dissolving the 100 μM protein stock solution in assay buffer (0.1 mg/mL BSA in 0.1 M phosphate buffer, pH 7.0). The 0.6 mM stock solution of B4F was prepared in pure DMSO, and kept at $-80\text{ }^\circ\text{C}$ until use. A working solution of 40 μM B4F was obtained by a 1:15 dilution in assay buffer. The Nunc 96 MicroWell black plate, from Thermo Scientific (Roskilde, Denmark) was prepared by first adding 100 μL of 2 μM Sav solution to each well of one lane. Then, 40 μM B4F in increments of 2 μL , from 8 to 30 μL , were added to each well. Finally, assay buffer was added to each well to a final volume of 130 μL per well. Each lane was performed in triplicate. The plate was shortly centrifuged (1 min, at 1,000 rpm) to remove air bubbles. The fluorescence of each well was measured with an UV microplate reader, with the following measurements parameters setup (Table 6.4):

Table 6.4. Measurement parameters setup.

Excitation wavelength: 485 nm	Fluorescence top gain: 50
Emission wavelength: 520 nm	Number of flash: 10
Excitation bandwidth: 12.0 nm	Lag time: 0 μs
Emission bandwidth: 12.0 nm	Integration time: 40 μs
Z-Position: 6900 μm	Plate definition file: NUN96ft.pdf
Unit: RFU	Shake duration (linear normal): 30 s

The evaluation of the number of free binding sites was done by identifying the quenching breakpoint (intersection of the linear regression of the lines, grey and red, Figure 6.2):

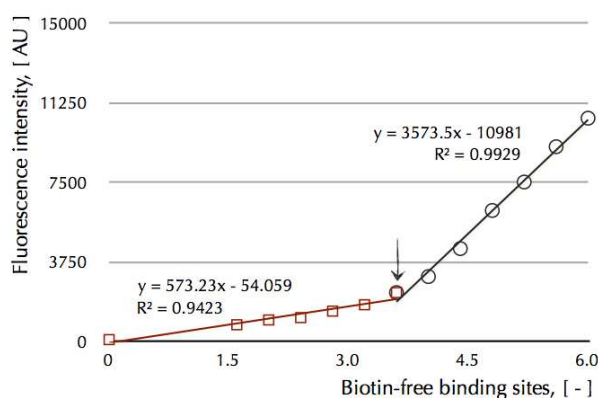


Figure 6.2. Example of a plot from a B4F titration (Sav S112A). The red line represents the first series of data, and the grey, the second series. On the x-axis, the number of free binding sites, and on the y-axis, the fluorescence intensity (in arbitrary units, AU).

$$y = mx + b \quad (6.1)$$

$$\frac{(m_1 - m_2)}{(b_2 - b_1)} = \text{quenching breakpoint} \quad (6.2)$$

which corresponds to the number of free binding sites, 3.7 from the plot example.

Per protein batch, three samples were weighed in, measured and the intersection calculated for each one. Then, the average of these three intersection values was taken as the number of free binding sites of the Sav variant analysed.

B4F back-titration

This protocol is a modified version of the protocol described above, and allows the calculation of monomeric Sav concentration in cell-free protein extracts (cfe).

The working solution of 40 μM B4F was prepared as described previously, and the solution of protein was used directly from the resuspension. The assay buffer was prepared by dissolving 0.1 mg/mL BSA in 0.1 M phosphate buffer, pH 7.0. The Nunc 96 MicroWell black plate from Thermo Scientific (Roskilde, Denmark) was prepared by first adding 10 μL of 40 μM B4F solution to each well of two lanes. Then, resuspended Sav was added to each well, in increments of 1 μL from 4 to 15 μL , of 5 μL from 15 to 50 μL , and finally of 10 μL from 50 to 100 μL . Finally, assay buffer was added to each well to a final volume of 120 μL per well. Each set of two lanes, covering the nominal concentration of 80 to 2000 μM , was performed in triplicate. The plate was shortly centrifuged (1 min, at 1,000 rpm) to remove air bubbles. The fluorescence of each well was measured with an UV microplate reader, with the same parameters setup as above (Table 6.4).

The evaluation of the nominal concentration of monomeric streptavidin was done by identifying the quenching breakpoint (intersection of the linear regression of the grey and red lines, Figure 6.3) knowing that 4×10^{-10} mol of B4F (amount of B4F added to each well) correspond to 4 μM monomeric Sav.

$$y = mx + b \quad (6.3)$$

$$\frac{(b_2 - b_1)}{(m_1 - m_2)} = \text{quenching breakpoint} \quad (6.4)$$

which corresponds to the nominal concentration of monomeric streptavidin. For the example given, 1642.38 μM of monomeric streptavidin, which corresponds to ~ 164 nmol free binding sites per mg of lyophilised cfe.

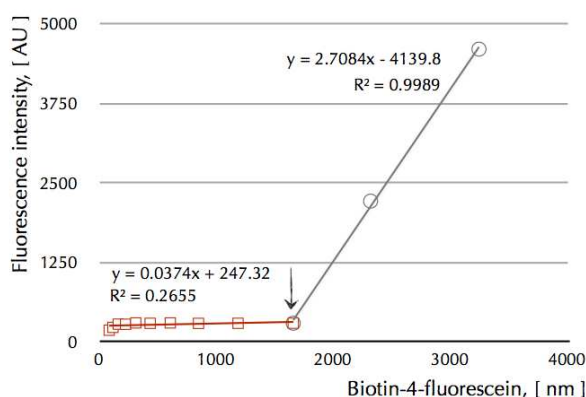


Figure 6.3. Example of a plot from a B4F back-titration (wt Sav cfe). The red line represents the first series of data, and the blue, the second series. On the x-axis, the nominal concentration of monomeric streptavidin, and on the y-axis, the fluorescence intensity (in arbitrary units, AU).

Analysis by mass spectrometry

The determination of the molecular mass of streptavidin and mutants was performed by electrospray ionisation (ESI) with a time-of-flight (TOF) mass spectrometry (MS) at the Biozentrum (University of Basel, Switzerland).

A stock solution of protein was prepared by dissolving 0.5 – 1 mg of protein in ultrapure water to a final concentration of 1 mg/mL. The analyte solution was diluted, in a plastic HPLC vial (100 μ L), to the proper final concentration of 0.1 mg/mL in 1:1 mixture of methanol:water, pH adjusted to 2.5 with formic acid, just before running the sample. Before analysing the protein sample by mass spectrometry, it is important to remove any precipitate and salts present in the sample. In case of precipitation, the sample was centrifuged for 1 min, at maximum speed, and the pellet discarded. The sample was desalted using Millipore ZipTip pipette tips (Billerica, MA USA), which could also be used to concentrate the sample (if necessary), following the supplier's protocol. The Agilent 1100 HPLC system, equipped with binary pump and solvent degasser, autosampler and column compartment, was equilibrated with a mixture of mobile phase, *i.e.* 50:50 acetonitrile:water with 0.1% formic acid, at 0.15 mL/min flow rate, until stabilisation of the signals (*i.e.* after 3 to 5 min). For positive-ion mode, 0.1% acid (either formic or acetic) was added into the analyte solution to enhance protonation, and increase sensitivity. The output of the HPLC separation was introduced into the ESI source of the Bruker Daltonic MicroTOF time-of-flight LC-MS system. Solvent A: 0.1% formic acid in water and solvent B: acetonitrile. The solvent composition was varied from 10% B to 95% B over 30 minutes. 20 μ L of the diluted sample was injected. The resulting data was analysed using Bruker Daltonics DataAnalysis software (Bruker Daltonics, Germany).

6.2.5 General procedure for the production of streptavidin in cell-free protein extracts

The work described herein was performed in collaboration with Dr Yvonne Wilson (**YW**) who designed the cell-free protein extracts and neutralisation of reduced glutathione experiments.

Preparation of cell-free protein extracts

Bacterial pellet from cells expressing streptavidin variants was resuspended in deionised water containing 1 mM PMSF and DNase I. The cell suspension was incubated (37 °C, 250 rpm) for 2 h before centrifugation (18,000 x *g*, 30 min, 4 °C). The concentration of biotin-binding sites in the supernatant was estimated by B4F back-titration before transfer to 50 mL PP tubes. The volume in each tube was noted, and these were frozen at -80 °C overnight. The samples were placed in a lyophiliser, and the water was removed under reduced pressure until completion. The mass of dried cell-free protein extracts in each tube was noted, allowing an approximate value of free biotin-binding sites per mg dry extracts to be calculated. As a negative control, cell-free protein extracts from the bacterial pellet of cells containing the pET11b plasmid with no Sav gene insert (empty plasmid) were prepared and lyophilised as described above.

Neutralisation of reduced glutathione

Several reagents (referred herein as “neutralising agents”) were tested in order to oxidise/neutralise reduced glutathione (GSH) present in cell-free protein extracts.

The lyophilised cell-free protein extract was resuspended in 1.33 M MOPS, pH 6.5, to a final concentration of 200 μ M monomer. A 40 mM stock solution of neutralising agent was prepared in catalysis buffer (1.33 M MOPS, 3.32 M HCOONa, pH 6.5), and was further diluted to obtain a 2x serial dilution of 20, 10, and 5 mM. The protein resuspension (100 μ L) was transferred to four different microcentrifuge tubes (2 mL), and to each different tube 100 μ L of neutralising agent was added, at different concentration (20, 10 and 5 mM). In the fourth tube, no neutralising agent was added (0 mM, negative control). The reaction tube was incubated for 24 h, at RT. The final concentrations in each tube were of 10, 5, 2.5, and 0 mM of neutralising agent, and 100 μ M Sav monomer. As negative control, purified protein (50 μ L, 400 μ M monomer) was spiked with 5 mM reduced glutathione (50 μ L of 10 mM stock solution in catalysis buffer, from Alfa-Aesar (Karlsruhe, Germany), and incubated for 2 h at RT. Then, each tube was treated as for the cell-free protein extracts samples, *i.e.* with 20, 10, 5, and 0 mM neutralising agent (100 μ L), incubated at RT for 24 h.

After 24 h, the catalysis reactions were carried out as described in Section 6.2.6, adding as positive control, purified protein.

6.2.6 General procedure for enantioselective catalysis

The conditions described herein are the optimised conditions developed by Dr Jeremy Zimbron (**JZ**) and Mr Marc Dürrenberger (**MD**), for their respective catalytic system.^[8,9] These conditions (temperature, pH, working concentrations) varied over time and after optimisation processes, and are not exhaustingly described in this thesis.

Experimental procedure for the transfer hydrogenation of imines

The work described herein was performed in collaboration with Dr Jeremy Zimbron (**JZ**) and Mr Marc Dürrenberger (**MD**), who synthesised the complexes and developed the analytical methods used in Chapters 2 and 3, respectively.^[8,9]

All experiments were carried out with degassed solutions (nitrogen flushed).

Preparation of stock solutions

For experiments performed in collaboration with **JZ**, the final concentrations of buffer were 4 M sodium formate and 3.2 M of MOPS in ultrapure water, and adjusted to pH to 5.0, by addition of concentrated formic acid (HCOOH) solution. A 2 M (final concentration) stock solution of substrate, 1-methyl-6,7-dimethoxy-3,4-dihydroisoquinoline, was used for the synthesis of salsolidine by asymmetric transfer hydrogenation (ATH).

The [Rh] dimeric complex was dissolved in 0.05 M DMSO, and stored under nitrogen, at 4 °C, until use.

For experiments performed in collaboration with **MD**, the final concentrations of buffer were 3 M sodium formate and 0.6 M of MOPS in ultrapure water, and adjusted to pH to 6.5, by addition of concentrated formic acid (HCOOH) solution. A 0.2 M (final concentration) stock solution of substrate, 1-methyl-6,7-dimethoxy-3,4-dihydroisoquinoline, was used for the synthesis of salsolidine by ATH. A stock solution of protein was prepared to a final concentration of 100 μM free binding sites per monomer. The final concentration of the biotinylated (Cp*)Ir complex **4** was 4 mM, and was prepared in pure DMSO.

Catalysis with protein purified by conventional affinity chromatography, JZ

Streptavidin was directly weighted into the reaction tubes (0.916 mM, final free binding sites concentration). A stock solution of 4 M sodium formate and 3.2 M MOPS (192 μL , pH 5.0) was added to the reaction tube, and the solution was stirred for 5 to 10 min, until complete dissolution of the protein. The dimer complex **20** (1.37/0.92 μL for a final metal concentration of 680/458 μM ; 0.75/0.50 equivalents [M] vs Sav S112H/Sav K121H free binding sites, respectively) was added to the reaction tube, and stirred for another 15 min, at RT.

Finally, the salsolidine precursor (6.8/4.58 μL for a final concentration of 68/45.8 mM for Sav S112H/Sav K121H, respectively) was added to the Sav complexed with the dimer complex **20**. The reaction tubes were placed in a magnetically stirred multireactor purged three times with nitrogen, heated to 55 $^{\circ}\text{C}$, and ran for 24 h.

Catalysis with protein purified by conventional affinity chromatography, MD

The volumes of solutions (protein, catalyst and substrate) varied based on the free binding sites concentration of the protein. Ideally, a final concentration of 100 μM free binding sites concentration per monomer, 50 μM catalyst **4** (1 mol%) and 5 mM salsolidine precursor (100 mol%) were used, which was translated in volumes of 192.5 μL of protein solution, mixed with 2.5 μL of catalyst for 10 min, and 5 μL substrate. The reactions were run in 2 mL microcentrifuge tubes, at 37 $^{\circ}\text{C}$ (or 30 $^{\circ}\text{C}$), for \pm 24 h, either in a shaking incubator or a heating thermomixer.

Catalysis with cell-free protein extracts obtained by precipitation, JZ

Depending on the final concentration of free-biotin binding sites, 189 – 198 μL of protein (equivalent to \sim 70 – 240 μM of protein) were pipetted into a reaction tube, mixed for 10 min with the appropriate volume for 0.5 equivalents of metal *vs* protein, and 1 mol% complex **20 vs** substrate. The reaction tube was placed in a magnetically stirred multireactor purged three times with nitrogen, heated to 55 $^{\circ}\text{C}$, and ran for 24 h.

Reaction work-up

After completion of the reaction, 500 μL of ultrapure water and 50 μL of 20% NaOH were added to the reaction mixture, which was extracted three times upon addition of 2 mL of pure CH_2Cl_2 or two times 1 mL of pure CH_2Cl_2 , in the case of the reactions performed with proteins purified by precipitation. The organic phase was dried over Na_2SO_4 , filtered through a 0.45 μm filter, and subjected to HPLC analysis.

Analytical method

The conversion (conv.) and enantiomeric excess (*ee*) of the reduction products were determined by chiral HPLC using Daicel Chemical Industries Chiralpak IC column - 20 μm , 250 x 4.5 mm (Tokyo, Japan), eluted with a mixture of $\text{CH}_2\text{Cl}_2/i\text{-PrOH/DEA}$ 98:2:0.1 at 1 mL/min (UV-detection at 280 nm). For the samples treated for oxidation of GSH, a mixture of $\text{CH}_2\text{Cl}_2/i\text{-PrOH/DEA}$ 99.8:0.2:0.06 was used for elution of the reduction products.

The absolute configurations were assigned by comparison of commercial enantiopure samples of salsolidine (6,7-dimethoxy-1-methyl-1,2,3,4-tetrahydroisoquinoline) and the salsolidine pre-

cursor (1-methyl-6,7-dimethoxy-3,4-dihydroisoquinoline):

(*R*)-6,7-dimethoxy-1-methyl-1,2,3,4-tetrahydroisoquinoline **3**, $t = 10.19$ min;

(*S*)-6,7-dimethoxy-1-methyl-1,2,3,4-tetrahydroisoquinoline **3**, $t = 6.86$ min;

1-methyl-6,7-dimethoxy-3,4-dihydroisoquinoline, $t = 8.37$ min.

The conversion was determined using the response factor (1.946) reported previously by Dr Thibaud Rossel.^[8]

Experimental procedure for the reductive amination of α -keto acids

The work described herein was performed in collaboration with Dr Jeremy Zimbron (**JZ**) who synthesised the complex and developed the analytical method used in Chapter 2.

Preparation of stock solutions

A stock solution of 5 M ammonium formate was prepared in ultrapure water, and pH adjusted to 8.0 by addition of concentrated ammonia solution (18 M). The dimer complex **20** was dissolved in 0.05 M DMSO and stored under nitrogen, at 4 °C, prior to use. Just before setting-up catalysis reactions, a stock solution of 2 M substrate solution was prepared by dissolving the α -keto acid in 5 M ammonium formate, and stirred for 1 h to favour the formation of the amine.

Catalysis with purified protein by standard affinity chromatography, JZ

The protein was directly weighted into the reaction tubes (final concentration of 0.916 mM free binding sites), dissolved in 194.5 μ L of 5 M HCOONH₄, and stirred until the dissolution of the protein was complete. The metal complex **20** was added in the same conditions as described in 6.2.6, for Sav S112H and Sav K121H. For other mutants screened, 0.5 equivalents [Rh] *vs* protein was used. The substrate was added to a final concentration of 45.8 mM. The reaction tubes were placed in a magnetically stirred multireactor purged three times with nitrogen, heated to 55 °C, and ran for 12 h.

Catalysis with purified protein by SSP, JZ

The lyophilised protein purified by SSP was resuspended in a minimum volume of 5 M HCOONH₄, depending on the initial concentration of the protein prior to lyophilisation (volume varied between 150 – 195 μ L), and the final working protein concentration was confirmed by Nanodrop ($\epsilon = 41,490 \text{ M}^{-1}\text{cm}^{-1}$, $MW_{monomer}$ calculated using ExpASy Prot-

Param tool).^[6] The catalyst and substrate volumes were adjusted to the concentration of the protein, in order to maintain (when possible) the same conditions as for the purified protein samples. The reaction tubes were placed in a magnetically stirred multireactor purged three times with nitrogen, heated to 55 °C, and ran for 12 h.

Reaction work-up

After completion of the reaction, 1.2 mL pure methanol was added to the reaction mixture, stirred for 15 min, and centrifuged at maximum speed (14,800 rpm, for 5 min). An aqueous solution containing 2 mM CuSO₄ was added in a 1:1 ratio to 600 μL of the reaction supernatant. The resulting solution was subjected to chiral HPLC analysis.

Analytical method

The conversion, selectivity and enantioselectivity of the reduction products were determined by chiral HPLC, using a Phenomenex Chirex 3126 (D)-penicillamine LC column - 5 μm, 50 x 4.6 mm (Torrance, CA USA). The yield was calculated by multiplying the selectivity by the conversion.

The enantiomer separation of (rac)-phenylalanine, phenylpyruvic acid and (rac)-phenyllactic acid was achieved by a gradient method (**line a**: 80% MeOH, 20% H₂O, and 2 mM CuSO₄, 22% from 0 to 14 min, and 37% from 14 to 40 min; **line b**: 2 mM H₂O/CuSO₄, 78% from 0 to 12 min, and 63% from 14 to 40 min), on a heated column at 40 °C, at a flow rate of 1.25 mL/min, and UV-detection at 247 nm.

The absolute configurations were assigned by comparison with commercial enantiopure samples of phenylpyruvic acid, phenylalanine and phenyllactic acid. Phenylpyruvic acid **12**, t = 5.93 min;

(*S*)-phenylalanine, t = 7.51 min;

(*R*)-phenylalanine, t = 10.11 min;

(*S*)-phenyllactic acid, t = 27.79 min;

and (*R*)-phenyllactic acid, t = 33.58 min.

The calculated response factors, at 247 nm, for phenylalanine was 1.6021, and for phenyllactic acid, 2.2601.^[8]

6.3 Experimental section of Chapter 3

6.3.1 General procedure for the production of streptavidin

The work described in Chapter 3 was carried out in the group of Prof. Dr Kurt Ballmer-Hofer, at the Laboratory of Biomolecular Research (LBR), Paul Scherrer Institute (PSI Ost Villigen, Switzerland). The molecular biology was performed under the direction of PD Dr Rolf Jaussi, and the initial screening of positive clones and fed-batch fermentation with the help and orientation of Mr Thomas Schleier.

Pichia pastoris KAI-3 strain^[10] was a generous gift from Prof. Callewaert (University of Ghent, Belgium).

Construction of the expression vector

In this Chapter, the yeast *Pichia pastoris* was selected for the expression of streptavidin wild-type. The expression vector pPICZ α was used as template. This vector (*c.a.* 3,600 base-pairs in size) contains a 5' *AOX* promoter region, which targets the plasmid integration to the *AOX1* locus in the *P. pastoris* genome, an α -factor-signal sequence that signals the secretion of the protein, the DNA sequence encoding for the protein of interest (wt Sav), and a Zeocin resistance used for selection.

The primers and construct numbering, used throughout this Chapter, reports to the internal database system of the LBR at the PSI.

Preparation of streptavidin insert

The primers were designed using Vector NTI software (Invitrogen, Carlsbad, CA USA). Two stop codons were introduced at the 3'-end to express the proteins without a C-terminal hexahistidine tag, which is encoded on pPICZ α A plasmid. The streptavidin insert was obtained by PCR using the following primers:

sense (5'-GAGAGGCTGAAGCTCGGGATCAGGCCGGCATC-3', T_m = 70.8 °C) and antisense (5'-CTAAGGCTACAAACTCACTACTGCTGAACGGCGTCG-3', T_m 67.9 °C), and a plasmid template of Sav, yielding a 478 bp fragment. The PCR reaction was prepared, on ice, by mixing 1 μ L of DNA template (10 ng), 1 μ L primers mix (1 μ L of sense primer + 1 μ L of antisense primer + 38 μ L dH₂O), 23 μ L dH₂O and, just before running the PCR reaction, 25 μ L of Phusion High Fidelity PCR Master Mix (Finnzymes, Espoo, Finland). The thermocycler block (Techne TC-312, Burlington, NJ USA) was pre-heated to 98 °C, for \sim 30 sec. The PCR conditions were: 2 min at 94 °C followed by 20 cycles of 15 s at 98 °C, stepdown cycle of 30 s at 60 °C (first cycle) – 50 °C (last cycle), and 30 s at 72 °C; and by

20 cycles of 15 s at 98 °C, 30 s at 55 °C, and 30 s at 72 °C, with a final extension of 1 min at 72 °C. The reaction was kept at 10 °C until DpnI digestion.

Preparation of Pichia pastoris vector

The vector was amplified using the primers

sense: 5'-TGAGTTTGTAGCCTTAGACATGACTG-3' and

antisense: 5'-AGCTTCAGCCTCTCTTTTCTCGAGAG-3',

and pPICZ α A (Invitrogen) as template, resulting in a fragment of 3463 bp. The crude PCR products were treated with 20 U DpnI enzyme (that disassembles only methylated DNA) for 18 h at 37 °C, and purified using MinElute Reaction Cleanup kit (Qiagen, Valencia, CA USA); in the case of the vector, PCR purification was performed only after linearisation (see below). The PCR products digested with DpnI were analysed on 0.7% w/v thin agarose gel (1.4 g agarose dissolved in 0.5X v/v sodium tetraborate (SB), and 3 μ L ethidium bromide. The gel ran at 225 V for 15 min).^[11] The vector was linearised with SacI (unique restriction site, from Fermentas GmbH, St. Leon-Rot, Germany). In an PCR tube, 50 μ L of DNA were mixed with 50 μ L of restriction enzyme buffer (1X NEBuffer1: 10 mM Bis-Tris-Propane-HCl, 10 mM MgCl₂, 1 mM DTT, pH 7.0), 370 μ L dH₂O, and 30 μ L SacI. The linearisation reaction was incubated at 37 °C, for 2h20.

Ligation of DNA insert into Pichia pastoris vector

The purified vector and insert DNA were then joined by CloneEZ reaction (GenScript, Piscataway, NJ USA), following the instructions of the supplier. In summary, 2 μ L of linearised vector was added to a mixture of 1 μ L of CloneEZ buffer, 3.5 μ L DNA insert, 1 μ L CloneEZ enzyme and 2.5 μ L of dH₂O.

Bacterial transformation

The recombined vector was used to transform *E. coli* One Shot Mach1-T1R chemically competent cells (genotype: F ϕ 80lacZ Δ M15 Δ lacX74 *hsdR*(r_K⁻m_K⁻) Δ *recA*1398 *endA*1 *tonA*). The ligation mixture from CloneEZ (5 μ L) was added to 50 μ L of thawed competent cells, containing 1 μ L of 10.8 M TMSO, and mixed gently. The transformation reaction was incubated on ice for 30 sec. The cells were heat-shock for 1 min at 42 °C, without shaking, and placed again on ice for 2 min. Pre-warmed SOC medium (450 μ L, 20g/L bactotryptone, 5 g/L bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, and 20 mM glucose) was added to the reaction vial, and incubated horizontally for 1 h, at 37 °C, 230 rpm in a shaking incubator. The different volumes of the transformation mix (50, 100 and 200 μ L) were spread on a pre-warmed low salt lysogeny broth (LB) agar Zeocin selective plate (10 g/L bactotryptone, 5 g/L NaCl, 5 g/L bacto yeast extract, 15 g/L agar + 25 μ g/mL Zeocin), and

incubated at 37 °C, overnight. For plasmid isolation, a single overnight-grown colony was grown in 5 mL of pre-warmed selective media (low salt LB containing 25 µg/mL Zeocin and saturating concentration of biotin [0.2% final concentration]). The cultures were shaken at 37 °C, overnight before isolating the plasmid (GenElute HP Plasmid kit, Sigma-Aldrich, St. Louis, MO USA). The construct bearing Sav gene was linearised with SacI, the digest was confirmed by 0.7% thin agarose gel electrophoresis before proceeding with electroporation.

Restriction enzyme analysis

The plasmid containing the insert was verified by restriction analysis. The chosen enzymes were Eco88I, Eco0109I, HindII, NaeI and HInfI. The first restriction enzyme cut the sequence twice (fragments of 1.3 and 2.5 kb), whilst Eco0109I, HindII, and NaeI cut three times (fragments of 0.14, 1.0 and 2.7; 0.6, 0.8 and 2.4; and 0.4, 1.1, and 2.3 kb, respectively), and HInfI five times (fragments of 0.3, 0.4, 0.6, 1.0 and 1.5 kb). The quantities of reactants and reaction conditions were used as described by the suppliers (Fermentas GmbH, and New England Biolabs Inc. for Eco0109I). The 1X Tango buffer was common to all four restriction enzymes from Fermentas GmbH, Eco88I, HindII, NaeI and HInfI (1X Tango: 33 mM Tris-acetate, 10 mM magnesium acetate, 66 mM potassium acetate, and 0.1 mg/mL BSA, pH 7.9); for Eco0109I, NEBuffer 4 (20 mM Tris-acetate, 10 mM magnesium acetate, 50 mM potassium acetate, 0.1 mg/mL BSA, and 1 mM DTT, pH 7.9) was used. The digestion reactions were prepared by mixing 1 µL DNA with 1 µL buffer, 0.5 µL enzyme and 7 µL dH₂O, and were incubated for 2 h at 37 °C. Restriction fragments were separated on 0.7% w/v thin agarose gel. The gel was analysed under fluorescent light. The construct was sequenced to confirm that Sav gene was in the correct orientation for expression, and cloned in frame with the α-factor signal sequence and the C-terminal peptide.

The correct sequence obtained was named "Sav918", and this annotation is used throughout this thesis.

Preparation and transformation of Pichia pastoris strains

Once the plasmid containing the correct insert (Sav918) was sequenced, enough plasmid DNA to transform *Pichia pastoris* cells (5 – 10 µg of plasmid per transformation) was generated. Two overnight cultures of 50 mL each (low salt LB medium containing 25 µg/mL Zeocin), and supplemented with 0.2% biotin (in 20 mM NaOH) were purified using GenElute HP Plasmid Midiprep kit from Sigma-Aldrich (St. Louis, MO USA). DNA was concentrated by isopropanol precipitation for integration into the genome of *Pichia pastoris* X-33 (wild-type, Invitrogen) and KAI-3^[10] (based on GlycoSwitch) strains. Sodium acetate (222 µL of 3 M, final concentration of 0.3M), 2 mL of DNA (resulting from the plasmid purification), and 1.6 mL 100% isopropanol (*i*-PrOH, 1:7 dilution) were placed in a microcentrifuge tube,

inverted 10x after the addition of *i*-PrOH, and incubated on ice for 15 min, and centrifuged for 15 min at maximum speed. The supernatant was carefully removed, and the pellet was washed two times with 70% ethanol (EtOH, 2 mL) at RT to remove the excess of salt. At each wash with EtOH, the tube was gently inverted two times, and centrifuged at maximum speed for 2 min. The ethanol was removed, the tube centrifuged again for 1 min, and dried at RT with the lid open. The pellet was finally resuspended in 100 μ L "T low E" buffer (10 mM Tris, pH 8.0 and 0.1 mM EDTA). The DNA precipitation by isopropanol allowed a five-fold concentration of the plasmid, *i.e.* initial concentration: 120 ng/ μ L *vs* final concentration: 570 ng/ μ L.

Both *Pichia pastoris* X-33 and KAI-3 strains were prepared and transformed in the same manner. The X-33 and KAI-3 cells (80 μ L) were thawed on ice and transferred to ice-cold 0.2 cm electrotransformation cuvettes from Bio-Rad (Munich, Germany). The linearised plasmid (10 μ L) was added to the cells, and incubated for 5 min on ice. The cuvette was tapped a few times on the bench to make sure that the cells sat evenly on the bottom of the cuvette, thus guaranteeing a correct electroporation. The cells were pulsed at 200 Ω (resistance), 25 μ F (capacitance) and 1.5 kV (electric potential). Immediately after pulsing, 1 mL of ice-cold 1 M sorbitol (1.82 mg in 10 mL deionised water, and filter-sterilised) was added, and the suspension transferred to a sterile 1 mL microcentrifuge tube. After regeneration for 2 h at 30 °C without shaking, aliquots were plated on YPDS agar selective plates (10 g/L bacto yeast extract, 20 g/L peptone, 20 g/L dextrose, 1 M sorbitol, 20 g/L agar, and 100 μ g/mL Zeocin, with and without 0.2% biotin), and grown for three days at 30 °C.

Culture conditions and expression

Media for small-scale experiments in shake flasks were prepared as described in Invitrogen User Manual, which required the preparation of several different stock solutions, as described below. All solutions were stored at 4 °C until use, and have a shelf-life of approximately one year, unless otherwise indicated. A saturated biotin stock solution (2%) was used instead of a 0.02% solution.

10X YNB: 67 g of yeast nitrogen base, with ammonium sulphate and without amino acids was dissolved in 500 mL of deionised water. The solution was heated to completely dissolve YNB. The solution was filter-sterilised. **10X D** (20% dextrose): 200 g of D-glucose was dissolved in 1000 mL of deionised water, and autoclaved. 1 M potassium phosphate buffer, pH 6.0: 132 mL of 1 M K_2HPO_4 and 868 mL of 1 M KH_2PO_4 were combined, and the pH adjusted to 6.0, if necessary by addition of phosphoric acid or KOH. The solution was sterilised by autoclaving and stored at room temperature. **50000X B** (2% biotin)²: 2 g of

²The concentration of the stock solution of biotin recommended by Invitrogen is 500X, which is equivalent

biotin was dissolved in 100 mL of 20 mM NaOH, pH 12.0. The addition of NaOH (at a final concentration in the medium of 0.04 mM) did not influence the final pH of the medium.

10X GY (10% glycerol): 100 mL of glycerol was mixed with 900 mL of deionised water, and autoclaved. **10X M** (5% methanol): 5 mL of pure methanol was mixed with 95 mL of deionised water, and filter-sterilised. The shelf-life of 10X M is approximately two months.

Inoculum and seed cultures were prepared using the buffered glycerol-complex medium (BM-GY), which contained (per litre) a combination of 10 g of bacto yeast extract, 20 g of peptone, 100 mM potassium phosphate buffer, pH 6.0 (100 mL of 1 M stock solution), 1X YNB (100 mL of 10X stock solution), 1X GY (100 mL of 10X stock solution), and 0.004% biotin (2 mL of 2% stock solution).

The main culture was prepared using the buffered methanol-complex medium (BMMY), which contained per litre: 10 g of bacto yeast extract, 20 g of peptone, 100 mM potassium phosphate buffer, pH 6.0 (100 mL of 1 M stock solution), 1X YNB (100 mL of 10X stock solution), 1X M (100 mL of 10X stock solution), and no biotin was added.

The defined basal salts medium (BSM) used for the fed-batch fermentation contained per litre: 26.7 mL of 85% phosphoric acid, 0.93 g of calcium sulphate dihydrate, 18.2 g of potassium sulphate, 14.9 g of magnesium sulphate heptahydrate, 4.13 g of potassium hydroxide, and 40 g of glycerol in 1 L ultrapure water. The solution was sterilised at 121 °C for 15 min, and after cooling down to room temperature, 4.3 mL of PTM1 mineral salts solution was added. The PTM1 mineral salts solution contained per litre: 6 g of cupric sulphate pentahydrate, 0.08 g of sodium iodide, 3 g of manganese sulphate hydrate, 200 mg of sodium molybdate dihydrate, 20 mg of boric acid, 500 mg of cobalt chloride, 20 g of zinc chloride, 65 g of ferrous sulphate heptahydrate, 200 mg of biotin, 5.0 mL of sulphuric acid, and ultrapure water to a final volume of 1 L. The solution was filter-sterilised and stored at room temperature.

50 mL expression control

From YPDS agar selective plates, five clones from X-33 and 15 clones from KAI-3 strains were chosen for a first production in a 50 mL format, in a total of 35 clones screened for streptavidin expression. In this first production, the two best producers of each strain were determined, and used for a scale-up culture of 200 mL. Different conditions were experimented (unbuffered medium, buffered at pH 5.0 and 6.0, and addition of 1% casamino acids) to overcome the proteolytic degradation of secreted recombinant protein, encountered in *Pichia pastoris* expression system.^[12,13] Early in the morning, for each clone, 5 mL BMGY was inoculated with a tip dipped in an isolated colony. The 50 mL PP tubes were properly labelled, and incubated at 30 °C, 225 rpm, for 12 h in the dark, as Zeocin is light sensitive

to 0.02% biotin.

(the doors of the incubator were covered with aluminium foil and the inside lamps covered, to minimise the degradation of the antibiotic). After the inoculation time of the first inoculum elapsed, the cells were spun down (4,000 rpm, 5 min) and resuspended in fresh 50 mL BMGY medium with 100 $\mu\text{g}/\text{mL}$ of Zeocin. This second inoculum was incubated overnight ($\pm 18 - 20$ h), at 30 °C and 225 rpm. Next day, cells were spun down (4,000 rpm, 5 min), and washed twice with ultrapure water to remove excess of biotin. To induce the protein expression, the medium was changed from BMGY to BMMY. The washed cells were resuspended in 50 mL BMMY medium with no antibiotic added, to a final OD_{600} of ~ 10 . Pure methanol was added twice a day to a final concentration of 0.5% per day, to reduce fluctuations in methanol concentration over time.

To follow the evolution of the protein expression, samples (500 μL) were taken at different time points (0, 24, 48, 72, and 96 h). After clearing the medium from cells, the supernatants were analysed by SDS-PAGE (Section 6.2.3).

Dimethyl sulphoxide (DMSO) and glycerol were used for the long-term storage of positive clones (final concentration of 7 and 50%, respectively). DMSO can be used as an alternative to glycerol for the freezing of cells, as it prevents the formation of ice crystals, thus preventing damage of the cells. In the case of stocks prepared in DMSO, 900 μL of the first inoculum was cooled down on ice before adding 100 μL DMSO. Cells were frozen slowly, in a -80 °C freezer. In the case of stocks prepared in glycerol, 50 μL of the first inoculum were mixed with 500 μL of glycerol and fast-frozen in liquid N_2 . To be noted, that when using stocks prepared in DMSO, cells have to be thawed very quickly, and washed once with medium before use.

200 mL scale-up expression

The growth of *Pichia pastoris* cells are limited by the oxygen mass transfer in baffled shake flasks. Invitrogen recommends a working volume of 10 to 30% of the total flask volume. As the only large volume baffled flasks available in our lab were of 2000 mL volume (BD Falcon - Erlenmeyer culture flasks baffled, BD Biosciences, Franklin Lanes, NJ USA), two expression volumes of scaled-up cultures were experimented: 200 and 600 mL, *i.e.* 10 and 30% of the total flask volume, respectively. Both working volume expressions followed the same procedure, thus only the 200 mL scale-up expression is described herein. As mentioned previously, the two best producers of each strain were used for a scale-up culture. The addition of antifoam was also investigated.^[14]

The inoculum was prepared by inoculating 25 mL of BMGY medium supplemented with 100 $\mu\text{g}/\text{mL}$ of Zeocin, in a 250 mL baffled flask, using a single colony. The culture was grown at 30 °C in a shaking incubator (250 rpm) until the culture reached an OD_{600} of $\sim 4 - 6$ (approximately 16 - 18 hours). The cells were in log-phase growth. The cells were harvested by centrifugation at 3,000 $\times g$ for 5 min at room temperature. The supernatant was decanted, and the cell pellet washed twice with ultrapure water. The cell pellet from the last wash was

resuspended to an OD_{600} of ~ 1.0 in 200 mL BMMY medium to induce expression. The culture flask was incubated at 30 °C and 250 rpm. Pure methanol to a final concentration of 1% methanol was added every 24 h to maintain induction, and avoid fluctuations of the inducer.

At different time points (in hours: 0, 6, 12, 24 (day 1), 36, 48 (day 2), 60, 72 (day 3), 84, and 96 (day 4)), 500 μ L of the expression culture was transferred to a 1.5 mL microcentrifuge tube. These samples were centrifuged at maximum speed for 2 – 3 min at room temperature. The supernatant was transferred to a separate tube, and both supernatant and pellet were frozen quickly in liquid N_2 . Samples were stored at -20 °C until ready to assay, *i.e.* to analyse expression levels, and determine the optimal time post-induction to harvest.

High-cell density fed-batch cultivation

The most common medium for high-cell density fermentation of methylotrophic yeast *P. pastoris* is the defined basal salt medium (BSM) along with trace salts medium (PTM1). The fermentation media were prepared according to the recommendations of Invitrogen Manual [Pichia Fermentation Process Guidelines, Version B, 053002], and to an internal protocol written by Mr Thomas Schleier [“Pichia Fermentation for Beginners”]. The pH value of the glycerol batch and fed-batch was adjusted to pH 6.0 with 25% ammonia solution and/or 3 M phosphoric acid, and to pH 5.0 for the methanol fed-batch.

The DMSO stock solution of clone 1 (50 μ L) was grown overnight in 5 mL BMGY medium with 25 μ g/mL of Zeocin, at 30 °C and 250 rpm. This inoculum was used to inoculate 200 mL of BMGY medium (0.25% (v/v), with 25 μ g/mL of Zeocin) in a 2000 mL baffled flask, and incubated at 30 °C and 180 rpm, until the culture reached an OD_{600} of 15. This inoculum seed was transferred into a 2.7 L fermentor (BioEngineering AG, Wald, Switzerland) containing 1.5 L defined basal salt medium (BSM) plus 4.35 mL/L trace salt solution (PTM1). After 21 h of generating biomass, the process control system detected a decrease in dissolved oxygen (dO_2) upon depletion of glycerol from the BSM medium, and a glycerol feed (50% w/v containing 12 mL/L PTM1) was started at a feed rate of 10 mL/h for the duration of 12 h. The temperature was kept at 30 °C, and pH was adjusted to 6.0 by automatic addition of 25% ammonia. Dissolved oxygen was maintained above air saturation ($>20\%$, 0.1 – 0.3 vvm) by adjusting the stirrer speed to 1,500 rpm. Antifoam 204 (1:20, Sigma-Aldrich) was added automatically on demand. The protein production was induced by addition of methanol, after 98 h inoculation, at an OD_{600} of 200. The glycerol-feed was stopped *c.a.* 1 h before starting the methanol-feed (starving phase). The methanol concentration was initially set to 0.5% v/v (feed rate of 3.6 mL/h, for 2 to 3 h) to adapt the cells to growth on methanol. The carrier gas (6 bar) was connected to maximise oxygen concentration in the medium ($>20\%$).

Once the culture was adapted to methanol, the inducer concentration was set at 1% v/v final concentration, at a constant feed rate of 16 mL/h for the remaining time of the fermentation. The first 500 mL of 100% methanol contained 12 mL/L PTM1, whilst the methanol added afterward did not contain any. At induction, pH was adjusted to 5.0. The entire methanol fed-batch lasted approximately 64 h, and was stopped at a final OD₆₀₀ of 890.

Due to the amount of glycerol and methanol added during the fed-batch phases, *c.a.* 800 mL of broth had to be removed twice from the fermentor.

6.3.2 Purification procedures

The purification of protein secreted by *Pichia pastoris* was only carried out for 200 and 600 mL cultures, and the fermentation.

All buffers were filtered through a Millipore pre-filter and a 0.22 μm filter (Type GVWP), and degassed under vacuum and strong stirring.

Preparation of samples

The yeast cells were removed by centrifugation (4,500 rpm, 30 min at 4 °C), and the cell-free medium (supernatant) further clarified through a Millipore Glass Fibre pre-filter (top) and a MF membrane filter 0.45 μm HA (bottom), using a Millipore stainless steel filter holder 90 mm, connected to a Watson Marlow 505U peristaltic pump (set at 140 rpm). Concentration of the protein and buffer exchange were performed on a Millipore Prep/Scale-TFF system. For example, the fermentation broth (1800 mL) was concentrated to 300 mL (void volume of the system: 200 mL), then topped-up to 1000 mL with binding buffer, and concentrated again to 300 mL. This procedure was repeated twice, and the pH verified to ascertain correct buffer exchange. To minimise the effect of proteases, 1 mL of 200 mM of PMSF (17 mg/mL in isopropanol) was added to the concentrated sample of Sav918.

The supernatant from the 200 mL scale-up expression was equilibrated in binding buffer by dialysis, MWCO: 10 kDa from SpectrumLabs (Irving, TX USA).

Gel filtration

A 10 mL sample of the recombinant protein produced by fermentation was purified via gel filtration, due to the uncertainty about the protein binding capacity for iminobiotin affinity chromatography (*i.e.* biotin-free binding sites to couple to the resin since the fermentation medium was supplemented with a residual amount of biotin).

The protein solution was loaded on a Superdex 200 10/300 GL gel filtration column (GE Healthcare) equilibrated with binding buffer containing 50 mM Tris-HCl pH 8.5 and 500 mM NaCl and 10 mM imidazole. The elution was carried out with the same buffer solution, at a flow rate of 1 mL/min, and 200 μ L fractions were collected. The peak fractions were analysed by SDS-PAGE, pooled and dialysed overnight at 4 °C once against Tris-HCl pH 7.4 and twice against ultrapure water. Concentration of the protein was calculated by measuring the optical density at 280 nm on a NanoDrop (from ExPASy Protparam tool: $MW_{monomer} = 15,478$ Da, $\epsilon = 41,940 \text{ M}^{-1} \cdot \text{cm}^{-1}$),^[6] and by weight after lyophilisation.

Affinity chromatography

Affinity chromatography was carried out as described in Section 6.2.2. Concentration of the protein was calculated by measuring the optical density at 280 nm on a NanoDrop (from ExPASy Protparam tool: $MW_{monomer} = 15,478$ Da, $\epsilon = 41,940 \text{ M}^{-1} \cdot \text{cm}^{-1}$),^[6] and by weight after lyophilisation.

6.3.3 Characterisation of recombinant streptavidin

Analysis by electrophoresis

Preparation of samples

The samples taken during cultivations were cleared of cells by centrifugation (5,500 rpm, 30 min, at 4 °C), the supernatant was collected and filtrated through a 0.45 mm membrane (Millipore). The samples from other processes (*e.g.* purification and concentration) were prepared as described in their respective sections. Samples were not heated prior electrophoresis (unless otherwise stated), nor the loading buffer contained DTT.

SDS-PAGE

Gel analysis of Sav918 was performed as described by Laemmli,^[15] using 10 or 12% acrylamide gels, followed by B4F detection and staining (Coomassie Brilliant Blue, silver, periodic acid/Schiff) or immunoblotting. For 12% gels, see description in Section 6.2.3. For 10% gels, the stacking layer (6% acrylamide final concentration) was prepared by mixing 2 mL 30% acrylamide/bis solution, 5.3 mL ultrapure water, 2.5 mL stacking buffer (0.5 M Tris-HCl, pH 6.8) and 50 μ L 20% SDS. Polymerisation was initiated by adding 65 μ L of 15% ammonium persulphate (APS, in dH₂O) and 10 μ L tetramethylethylenediamine (TEMED). The 10% resolving layer solution was prepared using 5.35 mL 30% acrylamide/bis solution, 6.35 mL

ultrapure water, 4 mL resolving buffer made with 1.5 M Tris-HCl (pH 8.8), and 80 μL 20% SDS. Polymerisation was initiated upon addition of 100 μL 15% APS and 16 μL TEMED.

Native-PAGE

Native gels (12% acrylamide) were prepared following the PSI internal recipe. The stacking layer (4% acrylamide final concentration) was prepared by mixing 6.4 mL 30% acrylamide/bis solution, 5.6 mL ultrapure water, and 4 mL stacking buffer (0.5 M Tris-HCl, pH 6.8). Polymerisation was initiated by adding 133 μL of 10% APS and 13.3 μL TEMED. The 12% resolving layer solution was prepared using 1 mL 30% acrylamide/bis solution, 5 mL ultrapure water, and 1 mL resolving buffer made with 1.5 M Tris-HCl (pH 8.8). Polymerisation was initiated upon addition of 30 μL 10% APS and 10 μL TEMED.

Analysis by B₄F detection

Analysis of SDS-PAGE by B₄F detection was performed as described in Section 6.2.3.

Analysis by Coomassie Blue staining

Analysis of SDS-PAGE by Coomassie Blue staining was performed as described in Section 6.2.3.

Analysis by silver staining

Silver staining is set as the standard of rigour for “ultra-sensitive” staining methods. The modified silver stain protocol (Bio-Rad Silver Stain, Cat No. 161-0443(7), LIT34 Rev B), optimised for mini gels ($\sim 7\text{ cm} \times 8\text{ cm} \times 0.75\text{ mm}$), was used as it gives clear backgrounds and consistent results in less time than the standard protocol.

After running the gels and removing them from the glass plates (and after B₄F detection), gels were transferred for 30 min or overnight in a fixative solution (40% methanol/10% acetic acid (v/v)). The gels were then incubated in Bio-Rad 1X oxidiser (10X stock solution dilute in deionised water), for 5 min, and washed several times afterward (6 to 7 times, for 15 min) with larges volumes of water, to leach the oxidiser out of the gels without removing it from the proteins. Followed incubation in silver reagent for 20 min, and a quick rinse with water for 30 sec maximum. Finally, the gels were developed for 30 sec or until a brown smokey precipitated appeared, with Bio-Rad developer solution, and this step was repeated as many times and for about 5 min, until the solution remained clear and the desired intensity achieved. The reaction was stopped by incubating the gels for 15 min in 5% acetic acid.

Analysis by periodic acid-Schiff staining

To determine if Sav918 was glycosylated, a simple and basic method was to resolve the protein on 12% SDS-PAGE, and to stain the gel for glycoproteins using one of the chromogenic gel staining procedure based on the periodic acid-Schiff (PAS) reaction.^[16] The gel was fixed in 50% methanol, washed and oxidised by immersion for an hour in 1% periodic acid dissolved in 3% acetic acid, and then thoroughly flushed with water for 1 h. The periodic acid oxidised two vicinal diol groups to form an aldehyde, which reacted with the Schiff's reagent (0.5% reduced, acidified fuchsin) to give a magenta colour within 1 h incubation. Stained gels were incubated in reducing agent, 1% sodium metabisulphite and stored in 3% acetic acid.^[17]

Analysis by Western blot

Roti-PDVF 0.45 μm membrane (Carl Roth GmbH Karlsruhe, Germany) were incubated for 5 min in 100% methanol, washed with ultrapure water and kept in water until use. After protein separation by 12% native-PAGE (150 V, for 90 min), gels were incubated with the membranes in transfer buffer (48 mM Tris-base, 39 mM glycine, 0.0375% SDS, and 20% methanol), for 15 min. The gels were transferred onto the membranes, using a semi-dry system (2 h, 30 mA/gel). After transfer, the membranes were washed in 1X TBST (50 mM Tris-HCl pH 8, 200 mM NaCl, 0.05% Tween 20) for 1 min, followed by a 30 min incubation in 3% BSA. The membranes were rolled "facing in", placed in a 50 mL PP tube, and incubated overnight at room temperature in 1:5,000 dilution of rabbit polyclonal anti-streptavidin antibody (Sigma-Aldrich) in 3% BSA, on a rocking/rolling table. Next day, the membranes were washed three times with TBST, for 10 min each time. The membranes were incubated in the the secondary antibody, 1:10,000 dilution in 3% BSA of goat anti-rabbit IgG alkaline phosphatase conjugated antibody (Southern Biotech, Birmingham, AL USA)), for one hour. Again, the membranes were washed three times with TBST, for 10 min each time. The blots were embedded in Lumi-Phos WB solution (ThermoScientific) on both sides, and analysed using ImageQuantRT ECL with 2 min exposure.

Quantification by B4F titration

Quantification of biotin-free binding sites was carried out as described in Section 6.2.4 for pure protein, and for the fermentation supernatant.

Analysis by mass spectrometry

The analyses described herein were performed in the Protein Analysis Group at the Functional Genomics Center Zurich (Switzerland). The mass determination of Sav918 was performed by Dr Serge Chesnov, and the N-terminal sequencing by Dr Peter Hunziker. The complete reports are presented in the Appendices Section.

A sample of Sav918 (in 1% acetic acid) was first analysed by ESI-MS. The detected mass was precisely that predicted for the fully protonated encoded protein, without the N-terminal methionine, and taking into account the two protonated histidine residues. The protein was also routinely digested by trypsin. The chymotryptic peptides were then desalted on Millipore C₁₈ ZipTips, and analysed by MALDI-MS in the reflectron-ion mode. The data was analysed using Bruker Daltonics flexAnalysis application.

The amino acids were sequentially cleaved from the N-terminal of a sample of Sav918 (in 1% acetic acid) via Edman Degradation. In this case, a Invitrogen ProSorb sample preparation cartridge was used to concentrate the sample onto a PVDF membrane, and to clean the sample from salts. The PVDF membrane was wet with 10 μL methanol, and 1.5 μL of sample was first diluted into 100 μL of 0.1% TFA before being added to the reservoir of the cartridge. The sample should be in contact with the the absorbent filter, and the fluid should transfer to the PVDF membrane. The membrane was washed once with 0.1% TFA, and dried before removal from the cartridge. The PVDF membrane was carefully cut out of the membrane with a razor blade, and the sample was then ready to be placed in the sequencer. The sequencing was analysed using SequencePro Data Analysis application (Invitrogen/Applied Biosystems).

Deglycosylation of sample

The deglycosylation of a sample of Sav918 was carried out using the Protein Deglycosylation Mix from New England Biolabs Inc., and following supplier's instructions.

For a denaturing reaction: *O*-glycosylated Sav918 (100 μg) was dissolved into 18 μL deionised water, and 10X glycoprotein denaturing buffer was added to a final volume of 20 μL . Sav918 was denatured by incubation at 100 °C, for 10 min. The denatured Sav918 was cooled down on ice, centrifuged for 10 sec, and 5 μL 10X G7 reaction buffer, 10% NP-40 and 15 μL dH₂O were added to the reaction. Before incubation at 37 °C for 4 h, 5 μL of deglycosylation enzyme cocktail was added and gently mixed.

For a non-denaturing reaction: *O*-glycosylated Sav918 (100 μg) was dissolved into 40 μL deionised water, and 5 μL of 10X glycoprotein denaturing buffer was added. To the reaction, 5 μL of deglycosylation enzyme cocktail was added and gently mixed, followed by incubation at 37 °C for 4 h.

A control (fetuin, provided by the supplier) was treated the same way for both reactions. The samples were analysed by SDS-PAGE.

6.4 Experimental section of Chapter 4

6.4.1 General procedure for the production of human carbonic anhydrase II

The plasmid pACA encoding for human carbonic anhydrase II (hCAII) was a generous gift from Prof. Carol A. Fierke (University of Michigan, USA).^[18] This construct consists of the hCAII gene^[19] behind a T7 RNA polymerase promoter, an f1 origin of replication,^[20] and an ampicillin (amp^r) and chloramphenicol (cm^r) resistance genes in pMa5-8 vector (Figure 6.4).^[21] The construct of this plasmid has an alanine residue at position 2 instead of a serine^[22] with no effect on protein expression or catalytic properties, and was used as a template for PCR.

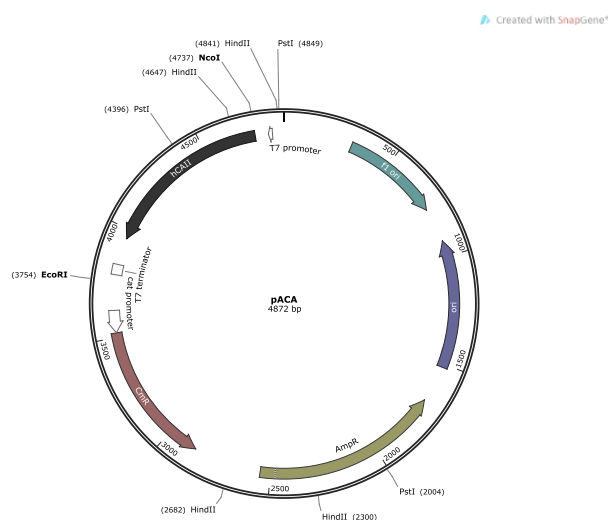


Figure 6.4. Plasmid map of hCAII gene in pMa5-8 vector. The heterologous gene expression of hCAII in *E. coli* is tightly controlled by the strong and efficiently regulatable T7 gene 10 promoter (in white). The transcriptional read-through from this promoter is minimised by the presence of a duplicated T7 transcription terminator sequence (in white), for the T7 RNA polymerase. Opposed to pET-vectors, read-through transcription from other plasmid promoters is minimised by the clockwise orientation of the T7 promoters relative to the anticlockwise orientation of the replication origin.

The numbering system of human carbonic anhydrase I was used throughout Chapter 4. In this system, residues Asn62, Asn67 and Thr200 correspond to residue Asn61, Asn66 and Thr199 in the numbering system of human carbonic anhydrase II.^[23]

Cloning & mutagenesis

pACA plasmid

The construct coding for hCAII was amplified by transformation into *E. coli* DH5 α cells (genotype: F⁻ Φ 80lacZ Δ M15 Δ (lacZYA-argF) U169 *recA1 endA1 hsdR17* (r_K^- , m_K^+) *phoA supE44* λ -*thi-1 gyrA96 relA1*), following Invitrogen protocol with minor changes. Fifty microlitre of

DH5 α chemically competent cells were thawed on ice. Ten nanograms of DNA were added to the cells, mixed gently, and incubated on ice for 30 min. Cells were heat-shocked for exactly 20 sec in a 42 °C water bath without shaking, and immediately placed on ice for 2 min. 950 μ L of pre-warmed SOC medium was added to the tube, followed by incubation at 37 °C for 1 h at 250 rpm in a shaking incubator. 20 μ L to 200 μ L from the transformation were spread on pre-warmed selective lysogeny broth (LB-Miller) plates (containing final concentrations of 60 μ g/mL ampicillin (amp) and 34 μ g/mL chloramphenicol (cm)). The different volumes was to ensure that at least one plate would have well-spaced colonies. The plates were inverted and incubated overnight at 37 °C. Three colonies were selected for a mini-plasmid preparation using Promega AG Wizard Plus SV Miniprep DNA purification system (Dübendorf, Switzerland), and analysed by sequencing at Microsynth AG (Balgach, Switzerland). The size of the plasmid was verified by restriction analysis. The chosen enzymes were NcoI, EcoRI, HindIII, and PstI. The first two restriction enzymes cut the sequence just once, whilst PstI and HindIII cut two and three times, respectively. The quantities of reactants are listed in Table 6.5, and follow the instructions given by the supplier (Promega AG).

Table 6.5. List of components for a restriction enzyme digestion [Promega technical manual]. The buffers used were buffer E (1X: 6 mM Tris-HCl, pH 7.4, 6 mM MgCl₂, 100 mM NaCl, and 1 mM DTT), and buffer H (90 mM Tris-HCl, pH 7.4, 10 mM MgCl₂, and 50 mM NaCl).

Component	Volume (μ L)
Sterile, ultrapure water	16.3
10 X buffer	2
Acetylated BSA, 10 μ g/ μ L	0.2
DNA, 1 μ g/ μ L	1.0
The samples were mixed by pipetting, followed by the addition of the restriction enzyme:	
Restriction enzyme, 10 U/ μ L	0.5
Final volume	20

For the simple digestions (HindIII and PstI), the buffer corresponding to the restriction enzyme was used (buffer E and H, respectively); for the double digestion (NcoI + EcoRI), buffer H was used as it is compatible with both enzymes. The digestions were incubated for 2 h at 37 °C. At the end of the reaction, 4 μ L of 6X loading buffer (24.4% v/v glycerol, 0.032% w/v bromophenol blue, 0.02% w/v xylene cyanol) were added to the reaction mixture, which was loaded on 0.7% w/v agarose gel (1.4 g of agarose in 0.5% v/v Tris-borate-EDTA (TBE) buffer and 10 μ L of ethidium bromide (10 mg/mL stock solution)). The gel was analysed under fluorescent light, using Bio-Rad Gel Doc XR+ software.

Site-directed mutagenesis

hCAII mutants were made by site-directed mutagenesis using the expression vector containing

wt hCAII (pACA) as template or, in the case of double mutations, the template used was one of the first mutation. Site-directed mutagenesis experiments were carried out following the procedure described by Zheng *et al.*,^[4] and tested *in silico* to minimise hairpin formation.^[5] The PCR primers were ordered to Microsynth (Balgach, Switzerland). The list of primers (sense and antisense) used for each subproject of this Chapter, and their characteristics are listed below (from Table 6.6 to Table 6.8).

Table 6.6. List of primers (s: sense; s: antisense) used for the construction of hCAII mutants (*e.g.* Q92G stands for glutamine residue at position 92 mutated to glycine), to test the affinity of ligands synthesised by FM. In lowercase, the mutation introduced, and in italic, a silent mutation (to avoid hairpin formation and/or self-annealing).

Mutant		Primers (5'→3')	T _m (°C)	Length (bases)	Comment
H4A	s	ATG GCC CAT <i>gca</i> TGG GGG TAC GGC AAA CAC AAC GG	70.2	35	Part of the histidine cluster at the entrance of the cavity ^[24]
	a	GCC CTA GGG CCA <i>tgc</i> ATG GGC CAT GGT ATA TC	68.5	32	
H64A	s	AAC AAT GGT <i>gct</i> GCT TTC AAC GTG GAG TTT G	62.2	31	Loss of catalytic proton shuttle ^[25]
	a	GTT GAA AGC <i>agc</i> ACC ATT GTT GAG GAT CCT C	63.5	31	
I91A	s	TAC AGA TTG <i>gct</i> CAG TTT CAC TTT CAC TGG GGT TC	64.5	35	Proton transfer in catalysis ^[26]
	a	GTG AAA CTG <i>agc</i> CAA TCT GTA AGT GCC ATC CAG GGG	67.7	36	
Q92G	s	C AGA TTG ATT <i>ggg</i> TTT CAC TTT CAC TGG GG	62.5	30	Indirect ligand-metal network ^[27]
	a	G AAA GTG AAA <i>ccc</i> AAT CAA TCT GTA AGT G	56.6	29	
V121G	s	CTT CAC TTA <i>ggt</i> CAC TGG AAC ACC AAA TAT GGG	63.2	33	Widens the hydrophobic pocket ^[18]
	a	GTT CCA GTG <i>acc</i> TAA GTG AAG TTC TGC AGC	63.7	30	
F131A	s	TAT GGG GAC <i>gcg</i> GGG AAA GCT GTG CAG CAA CCT G	70.6	34	Proton transfer in catalysis ^[26]
	a	AGC TTT CCC <i>cgc</i> GTC CCC ATA TTT GGT GTT CC	67.2	32	
K170A	s	CC ATT AAA ACA <i>gcc</i> GGC AAG AGT GCT GAC TTC A	65.7	34	Part of the active site cluster /involved in proton transfer ^[24]
	a	CT CTT GCC <i>ggc</i> TGT TTT AAT GGA ATC CAG C	63.7	30	
P202W	s	CTG ACC ACC CCT <i>tgg</i> CTT CTC GAA TGT GTG ACC TGG	71.0	36	Decreased folded-state stability, maintained activity ^[28]
	a	CA TTC GAG AAG <i>cca</i> AGG GGT GGT CAG TGA GCC TGG G	72.3	36	

Table 6.7. List of primers (s: sense; a: antisense) used for the construction of **hCAII L198x** variants (x = alanine, phenylalanine, glutamine and histidine), to test their affinity towards benzenesulphonamide and validate predictions from computational simulations (**MS**). In lowercase, the mutation introduced.

Mutant		Primers (5'→3')	Tm (°C)	Length (bases)
L198A	s	CC TAC CCA GGC TCA <i>gcg</i> ACC ACC CCT CCT CTT CCT G	72.7	35
	a	G AGG AGG GGT GGT <i>cgc</i> TGA GCC TGG GTA GGT CC	73.5	33
L198F	s	CC TAC CCA GGC TCA <i>ttt</i> ACC ACC CCT CCT CTT CCT G	70.2	35
	a	G AGG AGG GGT GGT <i>aaa</i> TGA GCC TGG GTA GGT CC	71.0	33
L198Q	s	CC TAC CCA GGC TCA <i>cag</i> ACC ACC CCT CCT CTT CCT G	71.5	35
	a	G AGG AGG GGT GGT <i>ctg</i> TGA GCC TGG GTA GGT CC	71.9	33
L198H	s	CC TAC CCA GGC TCA <i>cac</i> ACC ACC CCT CCT CTT CCT G	70.2	35
	a	G AGG AGG GGT GGT <i>gtg</i> TGA GCC TGG GTA GGT CC	68.5	33

Table 6.8. List of primers (s: sense; a: antisense) used for the construction of **hCAII C206x-SyC** variants (x = alanine or serine, and y = residue 50, 166, 173, 217 or 220), for chemical pseudo-contact shifts studies (**KZ**, **DH**). In lowercase, the mutation introduced, and in italic, a silent mutation (to avoid hairpin formation and/or self-annealing). C206A and C206S (marked with *) were used as templates for the second mutation of a serine residue. For expression, only C206S-SyX (marked with **) was used.

Mutant		Primers (5'→3')	Tm (°C)	Length (bases)
C206A*	s	CTT CTG GAG <i>gct</i> GTG ACC TGG ATT GTG CTC AAG	67.3	33
	a	CCA GGT CAC <i>agc</i> CTC CAG AAG AGG AGG GG	69.7	29
C206S**	s	CTT CTG GAG <i>gct</i> GTG ACC TGG ATT GTG CTC AAG	66.1	33
	a	CCA GGT CAC <i>aga</i> CTC CAG AAG AGG AGG GG	68.1	29
S50C	s	CTG TCT GTT <i>tgc</i> TAT GAT CAA GCG ACT TCC CTG	64.4	33
	a	TTG ATC ATA <i>gca</i> AAC AGA CAG GGG CTT CAG	62.5	30
S166C	s	GTG CTG GAT <i>tgc</i> ATT AAA ACA AAG GGG AAG AGT GC	64.5	35
	a	GT TTT AAT <i>gca</i> ATC CAG CAC ATC AAC AAC	58.2	29
S173C	s	AAG GGC AAA <i>tgc</i> GCT GAC TTC ACT AAC TTC G	63.5	31
	a	GAA GTC AGC <i>gca</i> TTT GCC CTT TGT TTT AAT G	61.0	31
S217C	s	AAG GAA CCC ATC <i>tgc</i> GTC AGC AGC GAG CAG GTG	69.8	33
	a	GCT GCT GAC <i>gca</i> GAT GGG TTC CTT GAG CAC AAT CC	69.0	35
S220C	s	AGC GTC AGC <i>tgc</i> GAG CAA GTG TTG AAA TTC CG	66.0	32
	a	CAC TTG CTC <i>gca</i> GCT GAC GCT GAT GGG TTC	67.8	30

PCR reactions were prepared by addition of 5 μ L 10x *Pfu* buffer, 2 μ L of 10 mM dNTP (final concentration 0.4 mM), 2.5 μ L DMSO (final concentration 5%), 1.5 μ L *Pfu* Turbo polymerase, 1.5 μ L of 10 μ M primers (sense and antisense), 35 μ L H₂O to 1 μ L of template. The cycle conditions were: initial denaturation (95 °C, 5 min), followed by 16 cycles of 1 min at 95 °C; 1 min at 60 °C; 15 min at 68 °C. The final elongation was performed at 68 °C for 1 h. PCR products were analysed by 1.2% agarose gel electrophoresis (2.4 g w/v agarose in 0.5% v/v TBE and 10 μ L ethidium bromide (10 mg/mL)).

The initial DNA template (wild-type sequence) was digested by DpnI (4 h at 42 °C). Five microlitre of PCR product was used to transform ultra-competent XL1-Blue *E. coli* cells (genotype: *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lac1^q ZΔM15 Tn10(Tet^r)*], produced in-house). Plasmids were purified using Promega Wizard Plus SV

Miniprep DNA purification system, and were sequenced either by Starseq (Mainz, Germany) or Microsynth (Balgach, Switzerland).

Expression in rich induction medium

The initial protocol to express human carbonic anhydrase II was kindly provided by Prof. Carol A. Fierke (University of Michigan, USA). It was, in time, adapted to the lab equipment and material available at the University of Basel, and continuously optimised, based on the best results obtained.

For high gene expression, the plasmid was transformed into *E. coli* BL21(DE3)pLysS cells (genotype: F⁻ *ompT lon hsdS_B(r_B⁻m_B⁻) dcm gal λ(DE3) [pLysS (cm^r), produced in-house). After thawing on ice 100 μL of ultra-competent *E. coli* BL21(DE3)pLysS cells, 8 μL of dithiothreitol (DTT, 200 mM stock) and 3 μL of plasmid (0.2 – 0.5 μg of DNA) were added to the ultra-competent BL21(DE3)pLysS cells and mixed gently. The mixture was left on ice for 15 min, and then plated on pre-warmed LB plates (containing final concentrations 60 μg/mL amp, 34 μg/mL cm, and 2% w/v glucose). The plates were inverted and incubated overnight at 37 °C, for plasmid multiplication.*

All solutions used were autoclaved (121 °C, 15 min) or filter-sterilised (0.22 μm sterile filter).

In microlitre scale

Prior to committing resources toward larger scale culture/fermentation, solubility and expression of hCAII variants were tested on a microlitre scale. The results of this analysis was used to guide decision making (temperature, time of induction, etc.) for larger scale production.

[Inoculum] In a 96-well plate, 300 μL of lysogeny broth (LB-Miller: 5 g/L bacto yeast extract, 10 g/L bactotryptone, and 10 g/L NaCl), containing 0.2% w/v glucose, 100 μg/mL amp and 34 μg/mL cm, were transferred to each well of the plate, and inoculated with one colony from a LB plate. The plate was covered with a breathable film and incubated at 37 °C, 250 rpm, overnight.

[Culture] In a 96 deep-well plate, 475 μL of rich induction medium (RI: 20 g/L bactotryptone, 10 g/L bacto yeast extract, and 5 g/L NaCl, and 1x M9 salts (6 g/L of Na₂HPO₄, 3 g/L of KH₂PO₄, 500 mg/L NaCl, 1 g/L NH₄Cl))³,^[29] containing 0.2% w/v glucose, 60 μM ZnSO₄, 100 μg/mL amp and 34 μg/mL cm, were transferred to each corresponding well, and inoculated with 25 μL of pre-culture. The tips used to transfer the pre-culture to the culture medium were left in the well to help with aeration. The plate was incubated at 37 °C, 250 rpm for 3 h. Due to the small volumes used in 96 well plates, optical density (OD₆₀₀) of cultures were not controlled over time, and induction was done arbitrarily after three hours.

³The rich induction medium and the 1X M9 salts are prepared separately and combined just prior to use.

Enzyme expression was induced by addition of 1.5 μL of RI containing 250 μM isopropyl- β -D-1-thiogalactopyranoside (IPTG) and 450 μM ZnSO_4 . The temperature was lowered to 18 $^\circ\text{C}$, to prevent the expression of the enzyme in inclusion bodies; the plate was incubated for another 15 h, at 250 rpm. The culture was stopped by centrifugation, at 4,400 rpm, 4 $^\circ\text{C}$. The supernatant was discarded, and the pellet was frozen at -20 $^\circ\text{C}$ until SDS-PAGE analysis.

In 1 L scale

Due to the high levels of hCAII obtained in a 1 L scale (up to 400 mg/L of 90 – 95% pure enzyme), most expressions were performed in this scale.

[Inoculum] After overnight incubation of the plate at 37 $^\circ\text{C}$, a medium sized colony was chosen to inoculate 15 mL LB medium (10 g/L bactotryptone, 5 g/L bacto yeast extract, and 10 g/L NaCl), containing the appropriate amounts of antibiotics (100 $\mu\text{g}/\text{mL}$ amp and 34 $\mu\text{g}/\text{mL}$ cm). The 75 mL baffled shake flask was incubated for 6 to 7 h in an orbital shaker, at 37 $^\circ\text{C}$ and 250 rpm.

[Pre-culture] After 6 to 7 hours incubation, the inoculum was centrifuged at 4,400 rpm, for 10 min, at 4 $^\circ\text{C}$. The supernatant was deactivated (by addition of sodium hypochlorite) and discarded, and the pellet was resuspended in 60 mL fresh LB medium, containing 0.2% w/v glucose, 100 $\mu\text{g}/\text{mL}$ amp and 34 $\mu\text{g}/\text{mL}$ cm. The 250 mL baffled shake flask was incubated overnight in an orbital shaker, at 37 $^\circ\text{C}$ and 250 rpm.

[Culture] A baffled shake flask containing 1 L of culture medium (620 mL of RI, 0.36X M9 salts, 0.4 % w/v glucose, 60 μM ZnSO_4 , 100 $\mu\text{g}/\text{mL}$ amp, and 34 $\mu\text{g}/\text{mL}$ cm) was inoculated with the pellet of the overnight cell culture (centrifuged at 4,400 rpm, 4 $^\circ\text{C}$ for 10 min). Cells were grown at 37 $^\circ\text{C}$ and 250 rpm, for 3 to 4 hours or until OD_{600} was between 0.8 and 1.0. Addition of 250 μM of IPTG supplemented with 500 μM ZnSO_4 induced protein expression, and ensured the correct folding of the expressed protein. The shaker temperature was set to 25 $^\circ\text{C}$, and shake for another 3 hours. A sample of the culture (1 mL) was taken just prior addition of IPTG, and then every hour. The samples were centrifuged (4,400 rpm, 5 min, 4 $^\circ\text{C}$), the supernatant was discarded, and the pellet was kept at -20 $^\circ\text{C}$ until SDS-PAGE analysis. After 3 hours of induction with IPTG/ ZnSO_4 , 8 $\mu\text{g}/\text{mL}$ PMSF (dissolved in DMSO) was added to the culture, to inhibit proteases action. After 6 h incubation at 25 $^\circ\text{C}$, the cells were harvested (4,400 rpm, for 15 min at 4 $^\circ\text{C}$). The supernatant was deactivated and discarded, and cell pellets were fast-frozen in liquid nitrogen and kept at -20 $^\circ\text{C}$ until further analysis.

[Glycerol stocks] Glycerol stocks of the variants were prepared by mixing 500 μL of the remaining first inoculum with 500 μL of 100% glycerol. The mixture was homogenised by vortexing, flash-frozen in liquid nitrogen, and stored at -80 $^\circ\text{C}$ until use.

In 20 L scale

Human carbonic anhydrase II wild-type was expressed only once in a 30 L fermenter (BioEngineering AG, Wald, Switzerland). Therefore, the fermentation conditions were not optimised, and the setup of the fermenter was based on the protocol in use for streptavidin fermentations.

[Seed culture] 150 mL of LB medium (10 g/L bactotryptone, 5 g/L bacto yeast extract, and 10 g/L NaCl), containing 0.2% w/v glucose and antibiotics at the same concentrations as in LB plates (100 $\mu\text{g}/\text{mL}$ amp, 34 $\mu\text{g}/\text{mL}$ cm), were inoculated with a single colony of transformed BL21(DE3)pLysS *E. coli* cells, and incubated overnight, in an orbital shaker (37 °C, 250 rpm).

[Fermentation] The whole seed culture was mixed with 0.4% w/v glucose, and then poured in the 30 L fermenter containing 20 L cell culture (RI medium: 20 g/L bactotryptone, 10 g/L bacto yeast extract, and 5 g/L NaCl; 1X M9 salts: 6 g/L of Na_2HPO_4 , 3 g/L of KH_2PO_4 , 500mg/L NaCl, 1 g/L NH_4Cl , 60 μM ZnSO_4 , 100 $\mu\text{g}/\text{mL}$ amp and 34 $\mu\text{g}/\text{mL}$ cm). The broth was fermented at 37 °C, under strong agitation (1,000 rpm), aeration (1.5 bar of air), and constant addition of Sigma antifoam 204 (Buchs, Switzerland). When optical density (OD_{600}) reached ~ 0.8 (approximately 4 h after inoculation), the cells were induced by addition of IPTG (250 μM) and ZnSO_4 (450 μM). The temperature of the vessel was set to 30 °C, and 1 mL sample was taken just prior addition of IPTG/ ZnSO_4 , and then every hour. The samples were centrifuged (4,400 rpm, 5 min, 4 °C), the supernatant was discarded, and the pellet was kept at -20 °C until SDS-PAGE analysis. Three hours after induction, 8 $\mu\text{g}/\text{mL}$ PMSF was added to the broth, to inhibit serine proteases action. When expression time was elapsed (~ 6 h after induction), bacterial cells were harvested (4,400 rpm, 4 °C, 15 min). The supernatant was deactivated (by addition of sodium hypochlorite) and discarded, and the pellet was kept at -20 °C until further analysis/work-up.

Expression in stable isotope medium

The first two steps (inoculum and pre-culture) to express labelled mutants were common to the uniform and specific isotopes labelling. Due to the use of deuterated water (D_2O) in the triple isotope labelling expression, the inoculum and pre-culture were prepared differently. Between the three methods of labelling, the main differences lie on the medium preparation, and time of expression.

All solutions used were autoclaved (121 °C, 15 min) or filter-sterilised (0.22 μm sterile filter).

Uniform ^{15}N isotope labelling

[Inoculum] Early morning, in a 75 mL baffled shake flask, 15 mL of LB medium (supplemented with 0.2% w/v glucose, 100 $\mu\text{g}/\text{mL}$ amp, and 34 $\mu\text{g}/\text{mL}$ cm) was inoculated with a

tip dipped into 50% glycerol stock of the variant to express, and grown for 6 - 7h at 37 °C, 250 rpm.

[Pre-culture] The inoculum was centrifuged at 4,400 rpm, for 10 min, 4 °C. The supernatant was discarded, and the pellet resuspended in 2.5 mL LB medium. Early in the evening (\pm 6 to 7 hours post-incubation of the inoculum), in a 250 mL baffled shake flask, 50 mL of LB medium (containing 0.2% w/v glucose, 100 μ g/mL amp, and 34 μ g/mL cm) were inoculated with 2.5 mL of resuspended inoculum, and incubated overnight at 37 °C, 250 rpm.

[Main culture] Uniformly 15 N-labelled variants was achieved by allowing the cells to grow in a defined medium consisting of 1X M9 salts preparation (2 g/L Na_2HPO_4 , 1 g/L KH_2PO_4 , and 500 mg/L NaCl), 1g/L $^{15}\text{NH}_4\text{Cl}$ (filter-sterilised), 2 mM MgSO_4 , 0.1 mM CaCl_2 , 0.2% w/v glucose, 0.5 mM ZnSO_4 , 1X FeSO_4 (prepared fresh and filter-sterilised just before use, as this solution is extremely prone to oxidation and precipitation), 1X vitamin mix (Sigma-Aldrich, Buchs, Switzerland), 100 μ g/mL amp, and 34 μ g/mL cm, in 1 L final volume. OD_{600} of the pre-culture was recorded before centrifugation ($\text{OD}_{600} \sim 6.0$, otherwise levels of expression were very low). The inoculum was centrifuged under sterile conditions, at 4,400 rpm, 4 °C for 5 min. The supernatant, was discarded, and the pellet resuspended in 50 mL of fresh culture medium (see above). The main culture was inoculated with the 50 mL of resuspended pre-culture, and incubated in the shaker, at 250 rpm and 37 °C, until $\text{OD}_{600} \sim 0.6$. The shaker temperature was set to 25 °C, and expression was induced with 250 μ M IPTG and 450 μ M ZnSO_4 , when OD_{600} reached 1.0 – 1.3. One millilitre sample was taken before induction, and afterwards every hour until the end of expression. The cells were incubated for another 4 to 6 hours, and OD_{600} was continuously monitored. To prevent leaking of labelling, the expression was not left for more than 6 hours post-induction, even if OD_{600} was still increasing. At the end of the expression, the culture flasks were placed on ice, and harvested at 4,400 rpm, 4 °C for 15 min. The supernatant was deactivated (by addition of sodium hypochlorite) and discarded, and the cell pellets were stored overnight at -20 °C. The samples were treated the same way as the main culture.

Specific ^{15}N Leucine isotope labelling

[Inoculum] See above for details.

[Pre-culture] See above for details.

[Main culture] OD_{600} of the pre-culture was recorded before centrifugation ($\text{OD}_{600} \sim 6.0$, otherwise levels of expression were very low). The pre-culture was centrifuged at 4,400 rpm, for 20 min, 4 °C. The supernatant was discarded, and the pellet resuspended in 20 mL of minimal medium from the 1 L main preparation. Incorporation of selectively ^{15}N -labelled Leucine was accomplished by growth in 1 L culture medium, containing 1X M9 salts (4.5 g/L

Na_2HPO_4 , 3 g/L KH_2PO_4 , and 500 mg/L NaCl)⁴, 1X amino acids mix (see Table 6.9), 0.2% w/v glucose, 2 mM MgSO_4 , 1X FeSO_4 , 0.1 mM CaCl_2 , 10 μM ZnSO_4 , 100 $\mu\text{g}/\text{mL}$ amp, and 34 $\mu\text{g}/\text{mL}$ cm. Fifty millilitre of the culture medium were used to dissolve 60 mg of ^{15}N Leu, which was filter-sterilised and reserved until induction time. Twenty millilitre of resuspended pellet from the inoculum were used to inoculate the main culture. The 3 L baffled shake flask was placed in the shaker, and incubated at 250 rpm, at 37 °C, until $\text{OD}_{600} \sim 0.8$. The shaker temperature was set to 25 °C, and expression was induced when OD_{600} reached 1.0 – 1.3, by addition with 250 μM IPTG, 450 μM ZnSO_4 , and 50 mL medium containing 60 mg of ^{15}N Leu. The cells were left to express for 4 to 6 hours (OD_{600} was continuously monitored, and samples pre- and post-induction were taken for SDS-PAGE analysis).

At the end of expression, the culture flask was placed on ice. Cells and samples were harvested at 4,400 rpm, 4 °C for 15 min. The supernatant was deactivated and discarded, and the pellets were stored at -20 °C.

Table 6.9. 1X amino acids recipe. The amino acids are listed in the order of which they have to be mixed. ^{15}N Leucine was added at the induction time.

Order	Amino acid	mg/L	Comments
1	Alanine	500	Dissolved well
2	Arginine	400	Dissolved well
3	Aspartic acid	400	Dissolved well
4	Glutamic acid	650	Dissolved well
5	Glycine	550	Dissolved well
6	Histidine	100	Dissolved well
7	Isoleucine	230	Dissolved well
8	Leucine (^{15}N labelled)	60	Floated on the surface, but dissolved well
9	Lysine-HCl	420	Floated on the surface, but dissolved well
10	Methionine	250	Floated on the surface, but dissolved well
11	Proline	100	
12	Serine	2100	
13	Threonine	230	
14	Valine	230	Floated on the surface, but dissolved well
15	Phenylalanine	130	Floated on the surface, but dissolved well
16	Tryptophan	50	Floated on the surface, but dissolved well
17	Asparagine	–	–
18	Cysteine-HCl	50	
19	Tyrosine	170	
20	Glutamine	400	Added at the end to avoid heating. Very unstable.

^2H , ^{13}C and ^{15}N isotopes labelling

Stock solutions were prepared in deuterated water (D_2O), and filter-sterilised. To minimise $^2\text{H}/^1\text{H}$ proton exchange, the media were used immediately after preparation, and were never autoclaved. To successfully implement complete replacement of protonated solvent with deuterated solvent and, subsequently obtain high levels of deuterium incorporation, *E. coli* cells containing hCAII C206S-S50C plasmid were directly diluted from dH_2O LB media into

⁴No ammonium chloride (NH_4Cl) salt was added in the medium used for specific labelling.

LB media containing high levels of D₂O (> 90%) without the need of intermediate steps of dilution.^[30]

[Inoculum] In early morning, in a 15 mL snap-cap PP tube, 5 mL of LB medium (containing no glucose, 100 µg/mL amp, and 34 µg/mL cm) were inoculated with a tip dipped into 50% glycerol stock of the variant to express. The inoculum was incubated for 6 – 7 h at 37 °C, 250 rpm.

[Pre-culture] Early in the evening (± 6 – 7 hours after beginning of the inoculum), the inoculum was centrifuged (4,400 rpm, 4 °C, 10 min), and resuspended in 50 mL of LB medium in 90% v/v D₂O (with 0.2% w/v ¹³C-glucose, 100 µg/mL amp, and 34 µg/mL cm). The pre-culture was incubated at 37 °C, 250 rpm, overnight.

[Main culture] Optical density of the inoculum was verified (OD₆₀₀ ~ 6.0) to guarantee an optimum level of expression. Upon confirmation of correct OD₆₀₀, the inoculum was centrifuged (4,400 rpm, 4 °C, 10 min) in a sterile PP tube. The pellet was resuspended in 15 mL of the culture medium (see below). In a 3 L baffled shake flask, 1X M9 salts (1.5 g/L Na₂HPO₄, 750 mg/L KH₂PO₄, and 375 mg/L NaCl) was combined with 1 g/L ¹⁵NH₄Cl, 0.2% w/v ¹³C-glucose, 2 mM MgSO₄, 1X FeSO₄, 0.1 mM CaCl₂, 10 µM ZnSO₄, 1X vitamin mix, 100 µg/mL amp, and 34 µg/mL cm, in 99% v/v D₂O.

The main culture was inoculated with 15 mL resuspended inoculum. The culture was incubated in a shaker, at 250 rpm and 37 °C, until OD₆₀₀ ~ 0.8. The shaker was cooled down to 25 °C, and expression was induced at OD₆₀₀ ~ 1.0 – 1.3, with 250 µM IPTG and 450 µM ZnSO₄. The cells were incubated for another 10 to 16 hours (16 hours post-induction was set as a limit, to prevent leaking of labelling and loss of plasmid). Optical density at 600 nm was continuously monitored, and samples were taken for SDS-PAGE analysis. Glucose levels were also monitored (“Roche Diagnostics Diabur-Test 5000”, Basel, Switzerland) as the culture should be stopped whenever the glucose in the medium was completely consumed. At the end of the expression, cells and samples were harvested at 4,400 rpm, 4 °C for 15 min. The cell pellets were stored at -20 °C overnight.

D₂O from pre-cultures and culture supernatants was recovered, and regenerated by distillation.

6.4.2 Purification procedures

Depending on the final use of the enzyme, isolation of human carbonic anhydrase II from *E. coli* lysates to apparent homogeneity was achieved by one, two or three steps purification procedure, involving anion exchange (DEAE Sepharose Fast Flow, GE Healthcare, Glattburg, Switzerland), immobilised metal affinity chromatography (HiTrapTM IMAC HP, GE Healthcare), and affinity chromatography (*p*-amino-methylbenzene sulphonamide agarose, Sigma-

Aldrich, Buch, Switzerland).

The three steps purification was only applied for crystallography purposes, as it rapidly yielded x-ray quality crystals, suitable for diffraction analysis.

Preparation of samples

The bacterial pellet from a culture and/or fermentation were treated the same way.

Cells were lysed by activating the gene encoding T7 lysozyme using three cycles of “freezing/thawing”, and were resuspended in 25 mL lysis buffer (50 mM Tris-SO₄ pH 8.6, 50 mM NaCl, and 0.5 mM ZnSO₄) containing the protease inhibitor phenylmethanesulfonyl fluoride (PMSF, 10 µg/mL). Cells resuspension was incubated under vigorous shaking (~ 300 rpm, rotating shaker) at room temperature (RT) for 30 min, DNase I (1 µg/L) was added, and cells were left for an hour under the same conditions as previously described, or until complete digestion of nucleic acids. The lysate was clarified by centrifugation (10,000 rpm, for 45 min at 4 °C). The supernatant was recovered, and the process (25 mL resuspension, shaking, and centrifugation) was repeated another two times (three extractions in total), to insure the maximum recovery of hCAII extract. Cell debris was deactivated and discarded, at the third and final centrifugation. The supernatant, resulting of the three extractions, was dialysed for 4 h against activity buffer (50 mM Tris-SO₄, pH 8.6, and 0.5 mM ZnSO₄), and filtered through a 0.45 µm filter.

Weak anion exchange chromatography

This purification step was done at room temperature, in a gravity column (“Glass Econo-Column”, BioRad, Reinach, Switzerland), packed with 25 mL chilled DEAE Sepharose Fast Flow resin (for 1 L culture). The anion exchange column was equilibrated with one resin volumes (RV) of 1 M Tris-SO₄, pH 8.6, followed by two RV of 0.5 M Tris-SO₄, pH 8.6, and finally four RV of activity buffer (50 mM Tris-SO₄, pH 8.6, 0.5 mM ZnSO₄) supplemented with 1mM DTT. The dialysed crude extract of hCAII was slowly loaded into the column. The first flow-through was collected, and the column was washed with one RV of activity buffer. The first wash was collected and kept separately from the first flow-through. The washing step was repeated twice, and each fraction were collected separately. An SDS-PAGE was carried out to identify the fractions containing hCAII. After confirmation of the presence of enzyme, fractions were pooled and dialysed against activity buffer, overnight at 4 °C.

The DEAE Sephacel resin was regenerated by performing one wash with two RV of 1 M Na₂SO₄, and three washes with two RV of 0.5 M Na₂SO₄, or until the flow-through was clear. The final wash was with two RV of 8 M urea, which was rinse out with deionised water (dH₂O). The column was stored in 50% ethanol.

Immobilised metal affinity chromatography

The first step of this purification is the removal of the copper ion from the HiTrapTM IMAC HP column and its replacement with zinc ion. The Co^{II} stripping was achieved by washing the column with 200 mL of stripping buffer (20 mM sodium phosphate, 500 mM NaCl, 50 mM EDTA, pH 7.4), followed by a wash with 200 mL binding buffer (20 mM sodium phosphate, 500 mM NaCl, and 1 mM imidazole), and a final wash with 200 mL of dH₂O. The column was charged with Zn^{II} ion by loading 10 mL 0.1 M ZnSO₄, and washed with 100 mL distilled water and 100 mL binding buffer, to adjust the pH. The column was then ready for use.

The dialysed suspension of hCAII was applied to a 15 mL (3x 5 mL) Sepharose-Zn^{II}-IDA affinity column (connected to an ÄKTAprime FPLC) equilibrated with activity buffer, at a flow-rate of 0.5 mL/min. The column was washed with five column volumes (CV) of activity buffer. The bound enzyme was eluted by a linear gradient of 1 – 200 mM imidazole in activity buffer. The peak fractions (4 to 5 mL) containing hCAII were pooled and dialysed two times against ultrapure water for 24 h, at 4 °C and against activity buffer overnight, at 4 °C.

Inhibitor affinity chromatography

The purification step by affinity chromatography (*p*-amino-methylbenzene sulphonamide agarose) can be performed on its own, as this procedure yielded 250 – 400 mg of protein, assessed by SDS-PAGE to be >95% pure. The XK 16 column (GE Healthcare, Glattburg, Switzerland), was packed with approximately 25 mL of resin, and equilibrated with five CV of activity buffer. The protein (either from previous steps of purification or directly from the first dialysis after extraction) was loaded onto the column at a very low rate (1 mL/min) to ensure complete binding to the resin. The affinity gel was washed with 5 CV of wash buffer (50 mM Na₂SO₄, 50 mM NaClO₄, and 25 mM Tris, pH 8.8). The bound enzyme was eluted with 10 CV of elution buffer (200 mM NaClO₄, 100 mM NaAc, pH 5.6). Collected fractions (10 mL) were pooled and dialysed at 4 °C against activity buffer for 24 h, dH₂O for another 24 h, and finally against ultrapure water, overnight. The final dialysed hCAII solution was flash-frozen in liquid nitrogen and lyophilised. After complete lyophilisation, the purified enzymes were kept at 4 °C until use.

6.4.3 Characterisation of recombinant human carbonic anhydrase II

Analysis by electrophoresis

Preparation of samples

The samples taken from the culture and/or fermentation were treated the same way. Cells were

lysed by activating the gene encoding T7 lysozyme using three cycles of “freezing/thawing”, and were resuspended in a volume (in μL) of ultrapure water equivalent to 40 times the optical density of the culture at the moment of sampling, supplemented with 1 μL of 1 mg/mL DNase I (in 5 mM Tris-HCl pH 7.5, 75 mM NaCl, 0.5 mM MgCl_2 , and 50% glycerol).^[29] The samples were vortexed, and incubated at RT, under vigorous shaking (rotating shaker, 300 rpm), until complete digestion of the nucleic acids (30 to 60 min). The bacterial extracts were then centrifuged (14,000 rpm, 5 min, RT). The soluble fraction (supernatant) was transferred to a new microcentrifuge tube, and the insoluble fractions were resuspended by vortexing, in the same volume of ultrapure water, as previously (*i.e.* in μL , 40x OD_{600}). 20 μL of each soluble and insoluble fractions, and a positive control (20 μL of pure protein diluted in ultrapure water, to a final concentration of 1 mg/mL) were mixed with 10 μL 3X loading buffer (50 mM Tris-HCl pH 6.8, 1% w/v SDS, 2% v/v β -mercaptoethanol, 10% w/v sucrose, 0.006% w/v bromophenol blue). The samples were quickly vortexed to ensure homogeneous mixing of the sample and loading buffer, and immediately charged on a 12% gel.

SDS-PAGE by Coomassie Blue staining

The 12% running gel was prepared by combining 5 mL of ultrapure water, 6 mL of acrylamide/bis-acrylamide (30% /0.8% w/v), 3.8 mL of 1.5M Tris-HCl pH 8.8, 75 μL of SDS (20% w/v), 100 μL of ammonium perchlorate (APS, 15% w/v), and 6 μL of tetramethylethylenediamine (TEMED). The 5% stacking gel was prepared by mixing 3.4 mL of ultrapure water, 1 mL of acrylamide/bis-acrylamide (30%/0.8% w/v), 1.5 mL of 0.5 M Tris-HCl pH 6.8, 30 μL of SDS (20% w/v), 40 μL of APS (15%), and 6 μL of TEMED. The stacking gel was poured on top of the polymerised running gel, the combs were inserted, and the gel was left for polymerisation, for 30 min. When ready, the gel was clamped to the chambers, and placed into the tank containing 1X SDS buffer (25 mM Tris-HCl, 0.192 M glycine, and 0.1% w/v SDS), to prevent dryness of the gel.

Twenty microlitre of each sample and 6 μL of the protein marker (“Prestained Protein Marker, Broad Range” from New England BioLabs Inc., Bioconcept, Allschwill, Switzerland) were loaded onto the gel, and ran at 200 V, until the blue colour of the bromophenol blue present in the 3X loading buffer reached the end of the gel (± 1.5 h). The gel was transferred to a plastic box containing staining solution (100 mL, 0.25% w/v Coomassie Brilliant Blue R-250, 50% v/v methanol, and 7.5% v/v glacial acetic acid), and incubated for one hour, under gentle rocking. The staining solution was discarded and replaced by destaining solution (100 mL, 20% v/v methanol, 10% v/v acetic acid), and the gel was incubated, on the rocker, for another 3 h or until the protein pattern started to be visible. The gel was left overnight in ultrapure water; a Kimwipe placed in the solution rapidly removed the excess of stain in the solvent.

Rapid staining of the gel was achieved by heating – almost to boiling – the staining solution, containing the gel, in a microwave oven. The gel was then placed for 5 min at RT, on the

rocker, during which the gel shrank. The staining solution was discarded and replaced by destaining solution. The gel was again heated – almost to boiling – in a microwave oven, and placed again under gentle rocking.^[31] The protein pattern was visible after 30 min, but overnight destaining was still necessary for image scanning. The gel was analysed using Bio-Rad Gel Doc XR+ transilluminator and software (Reinach, Switzerland).

Analysis by mass spectrometry

The determination of the molecular mass of human carbonic anhydrase II wild-type and mutants was performed as described in Section 6.2.4.

6.5 References

- [1] Humbert, N. *Approche combinatoire de la production et de la purification de mutants de la streptavidine en vue de générer des métalloenzymes artificielles*, Thesis, University of Neuchâtel, 2005.
- [2] Sardo, A. *Genetic optimization of biotin-binding proteins: Artificial metalloenzymes and beyond*, Thesis, University of Basel, 2010.
- [3] Studier, F. W.; Rosenberg, A. H.; Dunn, J. J.; Dubendorf, J. W. *Methods in Enzymology* **1990**, *185*, 60–89.
- [4] Zheng, L.; Baumann, U.; Reymond, J.-L. *Nucleic Acids Research* **2004**, *32*, e115.
- [5] Kibbe, W. A. “OligoCalc: an online oligonucleotide properties calculator”, <http://www.basic.northwestern.edu/biotools/OligoCalc.html>, 2007.
- [6] Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.; Appel, R.; Bairoch, A. “Protein identification and analysis tools on the ExPASy server”, <http://web.expasy.org/protparam>, 2012.
- [7] Kada, G.; Kaiser, K.; Falk, H.; Gruber, H. J. *Biochimica et Biophysica Acta* **1999**, *1427*, 44–48.
- [8] Zimbron, M. J. *Engineering artificial metalloenzymes based on biotin-streptavidin technology for DNA recognition and asymmetric transfer hydrogenation catalysis*, Thesis, University of Basel, 2011.
- [9] Dürrenberger, M.; Heinisch, T.; Wilson, Y. M.; Rossel, T.; Nogueira, E.; Knörr, L.; Mutschler, A.; Kersten, K.; Zimbron, M. J.; Pierron, J.; Schirmer, T.; Ward, T. R. *Angewandte Chemie International Edition* **2011**, *50*, 3026–3029.
- [10] Verweken, W.; Callewaert, N.; Kaigorodov, V.; Geysens, S.; Contreras, R. *Pichia Protocols*; volume 389 of *Methods in Molecular Biology* Humana Press: New York, USA, Second ed.; 2007.
- [11] Gordhart, J.; Gadella Jr., T. W. J. *Analytical Biochemistry* **2005**, *243*, 186–187.
- [12] Bora, N. *Large-scale production of secreted proteins in Pichia pastoris*; volume 866 of *Methods in Molecular Biology* Springer Science+Business Media: Dordrecht, The Netherlands, 2012.
- [13] Zhang, Y.; Liu, R.; Wu, X. *Annals of Microbiology* **2007**, *57*, 553–560.
- [14] Routledge, S. A.; Hewitt, C. J.; Bora, N.; Bill, R. M. *Microbial Cell Factories* **2011**, *10*, 1–11.
- [15] Laemmli, U. K. *Nature* **1970**, *227*, 680–685.
- [16] Steinberg, T. H. *Methods in Enzymology* **2009**, *463*, 541–563.
- [17] Clarke, J. T. *Annals of New York Academy of Sciences* **1964**, *121*, 428–436.

- [18] Nair, S. K.; Calderone, T. L.; Christianson, D. W.; Fierke, C. *The Journal of Biological Chemistry* **1991**, *266*, 17320–17325.
- [19] Murakami, H.; Marelich, G. P.; Grubb, J. H.; Kyle, J. W.; Sly, W. S. *Genomics* **1987**, *1*, 159–166.
- [20] Rosenberg, A. H.; Lade, B. N.; Chui, D. S.; Lin, S. W.; Dunn, J. J.; Studier, F. W. *Gene* **1987**, *56*, 125–135.
- [21] Studier, F. W.; Moffatt, B. A. *Journal of Molecular Biology* **1986**, *189*, 113–130.
- [22] Henderson, L. E.; Henriksson, D.; Nyman, P. O. *The Journal of Biological Chemistry* **1976**, *251*, 5457–5463.
- [23] Lindskog, S. *Zinc Enzymes*; John Wiley & Sons: New York, USA, 1983.
- [24] Briganti, F.; Mangani, S.; Orioli, P.; Scozzafava, A.; Vernaglione, G.; Supuran, C. T. *Biochemistry* **1997**, *36*, 10384–10392.
- [25] Fisher, Z.; Hernandez-Prada, J.; Tu, C.; Duda, D.; Yoshioka, C.; An, H.; Govindasamy, L.; Silverman, D. N.; McKenna, R. *Biochemistry* **2005**, *44*, 1097–1105.
- [26] Elder, I.; Han, S.; Tu, C.; Steele, H.; Laipis, P. J.; Viola, R. E.; Silverman, D. N. *Archives of Biochemistry and Biophysics* **2005**, *421*, 283–289.
- [27] Lesburg, C. A.; Christianson, D. W. *Journal of the American Chemical Society* **1995**, *117*, 6838–6844.
- [28] Krebs, J. F.; Fierke, C. A. *The Journal of Biological Chemistry* **1993**, *268*, 948–954.
- [29] Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular cloning: a laboratory manual*; volume 2 Cold Spring Harbor Laboratory Press: New York, USA, 2nd Edition ed.; 1982.
- [30] Venters, R. A.; Huang, C. C.; Farmer, B. T.; Trolard, R.; Spicer, L. D.; Fierke, C. A. *Journal of Biomolecular NMR* **1995**, *5*, 339–344.
- [31] Studier, F. W. *Protein Expression & Purification* **2005**, *41*, 207–234.

Acknowledgements

After all, science is essentially international, and it is only through lack of the historical sense that national qualities have been attributed to it.

Marie Skłodowska-Curie

My favourite Portuguese poet, Fernando Pessoa, once wrote “*there is a time when we must abandon our old clothes, which already have the shape of our body, and forget about our paths, the ones that takes us always to the same places. It is time to cross the river, and if we do not dare to do it, we will forever be at the margin of ourselves.*” It perfectly translates these past four and a half years in Basel. Somehow, I managed to leave my used clothes and *cross the river by feeling each stone*, and I owe it to the people who have contributed in so many (and twisted) ways.

I wish to thank Prof. Dr Thomas R. Ward for giving me the opportunity to carry out my PhD work in his pluridisciplinary group. It has been an insightful learning journey, both personally and professionally. I am also grateful to Dr Gideon J. Grogan (University of York, UK) for accepting to co-referee this thesis, and Prof. Dr Thomas Pfohl for accepting to be the chairman of a *viva voce* twice in one day.

This research project was funded by the European Commission Marie Curie Training Network (FP7-ITN-238531). Within the framework of the project, I did a three-months internship at Pfizer Process & Development Centre, in Loughbeg (Co. Cork, Ireland). It was a wonderful and enriching experience in rainy, damped and cold Ireland. Special thanks to all at the Loughbeg site, in particular Pat and Victor for accepting me as part of their team, Simon for his guidance, Siobhain and James for HPLC and GC analyses, and Carmel for her help with administrative paperwork. A warm hug to my office/coffee-break lads: Seb, Rob, and Louis “Dr Sheldon Cooper” (I too always think of you when iwastesomuchtime!). This internship would not have been possible without Roger who put me in contact with Pat. I owe you, big time!

A gigantic thank you to my peers in the “Biotrains” network, their PIs and the industrial partners. It was a wonderful and incredible expedition in the world of (bio)catalysis.

I share the credit of this work with Jeremy, Yvonne and Marc D. (**Chapter 2**), Rolf, Thomas and Prof. Dr Kurt Ballmer-Hofer from the PSI (Villigen, **Chapter 3**), and Fabien, Maurus, Kaspar, Tristan, Prof. Dr Markus Meuwly and Daniel (**Chapter 4**). This thesis relied greatly on those collaborations, and it is the sum of real teamwork.

I am indebted with the French Connection (Jeremy (*aka* Jéjé), Fabien (*aka* Fabinou dans le monde des bisounours), Cheikh, Sabi(na), Thib(aud), Didier (*aka* Dides), Greg(ory) and Michael, and Little Italy (Alessia, Tommaso and Gaetano) for their contribution, both personally and chemically. I learned a great deal with them, and words alone cannot convey how thankful I am for their friendship and help during my PhD, and the redaction of this manuscript. Their good vibes helped me to overcome some of the darkest moments of my postgraduate studies. A special thanks to Fabien, Sabina, Marc R. and Tommaso for drawing those strange and complex molecules on chemdraw. To my enantiomer, Fabien, who helped me get on the road with L^AT_EX, and was incredibly patient, even when I blew it off!

It is with immense gratitude that I acknowledge Marc C., Narasimha, Mark R., Yvonne and Stefan. Marc C. kept me entertained with his huge repertoire of anecdotes and stories, Narasimha taught me to have a wiser and calmer approach to the “situations”, and Yvonne was always a good listener and motivator to finish my thesis (btw, thanks for the pep talks). Mark R. was the funniest and loveliest character ever. I enjoyed every “American moment” with him. Thanks all for sharing your knowledge and insightful point-of-view with me!

My sincerest thanks to my TA team, Marko and Jörg. Enjoyable times we spent together during those long hours teaching. Again, thanks for the coaching, patience, and coffee/smoke-breaks!

The smooth running of the Department would not be possible without the professionalism of the technical staff, from the “Werkstatt” and the “Materialausgabe”. As well as efficiently taking care of the administrative chores, Bea provided me a friendly smile and unforgettable coffee-break moments. To Isa, the last addition of the Department, my deepest gratitude for the highly efficient and friendly service.

Beyond the Inorganic Chemistry Department, Stefan & Lena, Dirk & Eva, Sven, Diana & Vincent, and Smahan have been outstanding companiable lunch/dinner mates. Thanks for the open ear, and for making my stay in Switzerland an agreeable one.

To my medical team who helped me overcoming some of my “limitations”: Dr Stutz, Dr

Martin-Vogt, Anke, Christine and Prof. Dr Dr Valderrabano.

To my friends around the globe who never let me go – even when I disappeared for months, not to say years. With no special order, Sílvia, Anabela, Joana, Kikas, Irina, Paulo C., René O., Céline B., John B., Isabela, Danni, and Michael H.

One of the hardest decisions of life is deciding whether to walk away or try harder. “*And life is the hardest teacher, it gives the test first, and the lesson after*”. To my baby brother/cousin, Tiago, who is giving me **the** lesson of life. I hope that one day you will read this, and you will know how much I admire you, your strength, and your love for life. Thrive on!

A special thanks to my uncle and aunt, Bino and Augusta, for giving me a shelter/home when I moved to Portugal, and for making me go beyond my limits. To my uncle and godfather, Firmino, who taught me to never give up and was an example of a real fighter.

To Sabi, the one and only! I wonder if my liver will ever recover from our “Telma & Louise” nights out. Je te dois bien plus que je ne devrais. Merci de m’avoir remis les idées en place plus d’une fois, pour ton rire ravageur et nos longues conversations au H., autour d’une bonne bouteille de rouge ou de Porto. Tu vas me manquer, ma belle!

To André and Pierre – the *plus ones* of the family. Thanks for being part of our craziness!

To the Familie Schuster – Ich verspreche, dass ich eines Tages in der Lage sein werde, mit euch auf Deutsch zu kommunizieren, oder noch besser auf *Boarisch*. And many thanks for letting me be part of the team.

No one had a bigger role in this play than my partner in crime/life, Thomas. He was always able to talk me down from the verge of utter panic, to push me forward in my *donkey moments*, and to bring down my *monkey mind*. He was on my side in some of the worst and best moments of my life. My eternal gratitude to you, my *lxxe*, for your great patience, care and support over the past 3+ years. Let’s walk down the road together, wherever life takes us.

As always, my parents and my sisters have given me their unequivocal support and encouragement throughout, for which my mere expression of thanks does not suffice! To the Nogueira’s *adds-on*, my niece Matilde and nephew Tomás, without whom life would be meaningless. To my grandpa, Leonardo, wherever he is, I know that he is always watching over me. Estou eternamente grata por vos ter na minha vida, e palavras nunca poderão expressar o quanto vos amo.

Curriculum vitae

Elisa S. Nogueira

Burgfelderstrasse 5
CH - 4055 Basel

EDUCATION

PhD., Biochemistry Sept. 2008 to Apr. 2013

University of Basel, Basel – Switzerland,

Dissertation: “Novel approaches for the creation of artificial metalloenzymes”

Advisors: Prof. Dr Thomas R. Ward

Co-referee: Dr Gideon J. Grogan (University of York, UK)

MSc., Biochemistry Oct. 2006 to Aug.2008

University College Dublin, Dublin – Ireland,

Dissertation: “pH-dependence of the catalytic activity of hen-egg white lysozyme”

Advisors: Dr Jens Nielsen & Dr Damian Mooney

Bachelors thesis, Biological and Chemical Engineering Sept. 2003 to Mar. 2004

University of Maryland Baltimore County, Baltimore MA – USA

Dissertation: “Study of a novel protease secreted by *Teredinobacter turnirae*”

Advisors: Prof. Dr Antonio Moreira & Prof. Dr Antonio Teixeira

BSc., Bioengineering Sept. 1998 to Nov. 2004

University of Minho, Braga – Portugal

Major: Chemical and Food Technologies

EXPERIENCE

Post-Doctoral Researcher Jun. 2013 to Dec. 2013

University of Basel, Basel – Switzerland

Prof. Dr Wolfgang Meier & Prof. Dr Nico Bruns

Redesign of thermosome by circular permutation.

Trainee Research Scientist Nov. 2011 to Feb. 2012
Pfizer Process & Development Centre, Loughbeg Co. Cork – Ireland
Redesign the synthetic route of an existing active pharmaceutical ingredient by introducing a biocatalytic step.

Junior Consultant May 2006 to Oct. 2006
AEFafe, Fafe – Portugal
Implementation of quality (ISO 9001) and food safety (HACCP, ISO 22000) standards, in small companies.

Certified Coach Feb. 2005 to Oct. 2006
AIMinho, Braga – Portugal
Professional training of internal auditors, quality managers and production employees.

Junior Consultant and Auditor Feb. 2005 to May 2006
Process Advice, Braga – Portugal
Implementation and audit of quality (ISO 9001) and food safety (HACCP, ISO 22000) standards, in small/medium-size companies and industries.

Lab Technician Apr. 2003 to Aug. 2003
Eurotinge, Barcelos – Portugal
Preparation of dyed textile samples, and quality control in the laboratory and at the production line.
Development of spectrophotometric techniques to analyse textile dyes in fabric and industrial effluents.

SKILLS

Languages:

French: native

Portuguese: native

English: proficient

German: basic understanding

Informatics: Microsoft Office, iWork, \LaTeX , BibTex, EndNote, Scifinder, SnapGene, Py-mol, Sparky, and ExPASy.

Teaching:

Supervision of two Masters' students (six months), and four undergraduate students (Final project, six weeks).

Supervision and coordination of Analytical Chemistry wet lab practice during seven semesters (Freshers of Pharmaceutical Sciences).

Certified training in Quality standards.

Communication:

Expertise in presenting challenging scientific concepts clearly to general scientific audience, including group seminar series and international conferences (Poster presentation prize at Biotrans2012).

Proficient at generating scientific reports.

Experience in professional training (Quality standards) to professionals from diverse backgrounds and level of education.

Initiative:

Development of new techniques and methods not previously used in the group.

Acquisition and commissioning of equipment, and teaching.

Organisation of an international conference (Marie Curie ITN "Biotrains"), for 20 international students, their PI's and ten representative of renowned pharmaceutical and chemical companies.

OTHER PERTINENT SKILLS

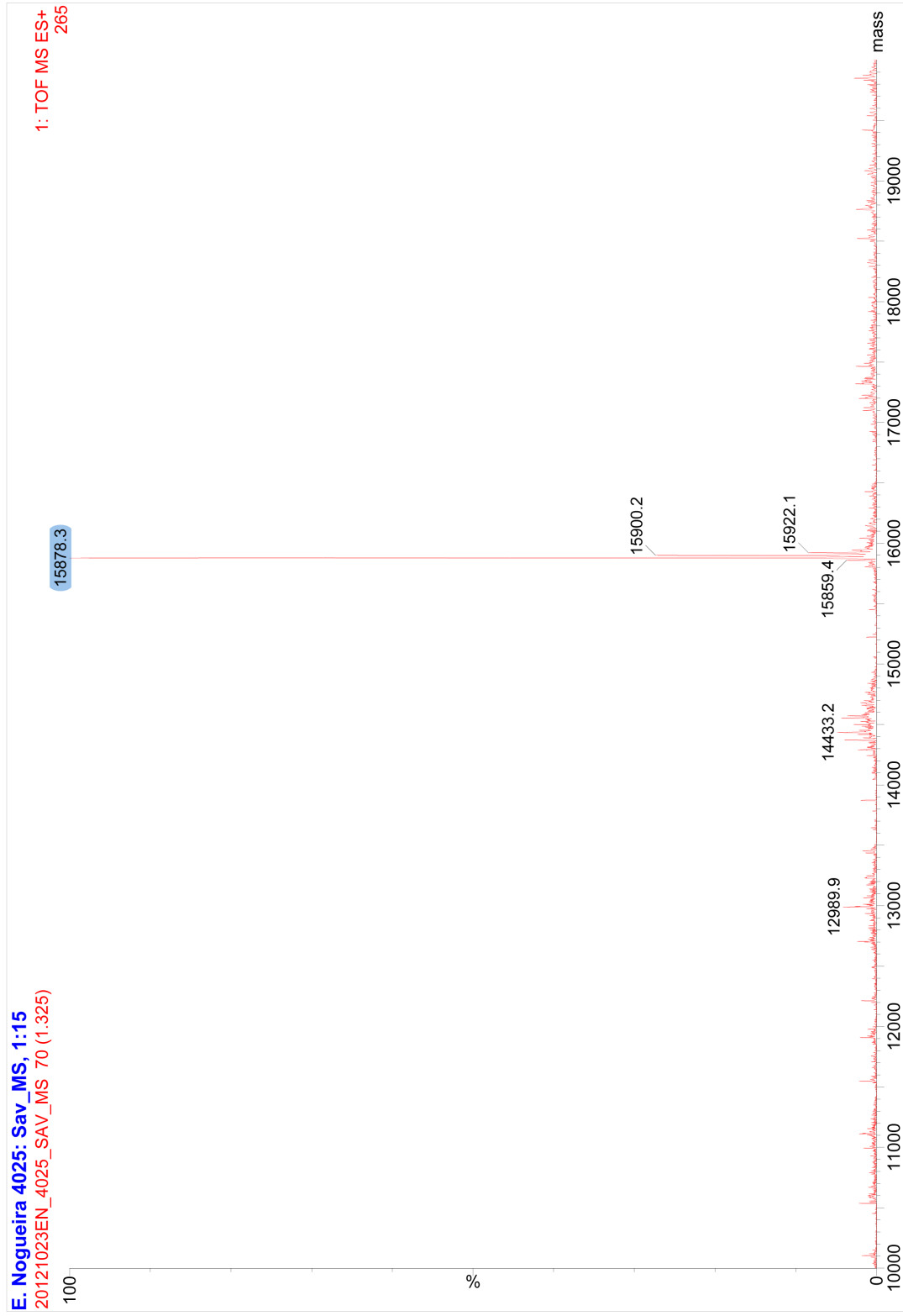
Certification in Quality Management tools (ISO 9001, ISO 22000, and HACCP standards), Clean-In-Place, cGMP, manager QA acceptor role, occupational hygiene, Key Performance Indicators (KPI), balanced scorecard, internal and external auditor.

Two years consulting and auditing experience.

INTERESTS

Blogger, "Language and foreign culture camp" monitor (2000-01), vegetarian cooking, traveling, hiking.

Appendices





functional genomics center zurich

Dr. Peter Hunziker
 Protein Analysis Group
 Functional Genomics Center Zurich
 Y 11 H 16
 Winterthurerstrasse 190
 CH-8057 Zürich

Tel. +41 44 635 39 31
 peter.hunziker@access.uzh.ch

Date: 01.11.2012
 Customer: Elisa Nogueira
 Order/Samples: 4028/Sav918
 Analysis ID: 20121031

Sample preparation

- Reduction/alkylation
- Proteolytic digestion
- ZipTip
- Precipitation
- Prosorb
- Electroblothing
- Other

Separation

- HPLC
- Gel electrophoresis

Mass determination

- Electrospray
- MALDI
- LC/ESI
- LC/MALDI

Protein/peptide identification/characterization

- PMF MALDI
- PMF LC/MS
- MALDI/MS/MS
- ESI/MS/MS
- LC/MALDI/MS/MS
- LC/ESI/MS/MS

Amino acid analysis

- Composition
- Concentration

Sequence/structural analysis

- Edman degradation
- ESI/MS/MS
- MALDI/MS/MS

Comments

For results see attached report. In order to see the comments and marks on a hardcopy of the attached report the option "Document and Markups" must be selected in the print window of the pdf reader program.

Procedure:**ProSorb:**

- 10 ul MeOH
- 1.5 ul sample + 100 ul 0.1% TFA loaded
- washed 1 x with 0.1% TFA

Result:

Clear sequence: EAEARD

SAMPLE INFORMATION

Sample Name:	20121031en_4025_	Std Amount:	2.000 pmols
ID Code:		Sample Amount:	2.000 pmols
Comments:	Sequence: EAEARD	Detector Scale:	0.005 AUFS

SEQUENCER INFORMATION

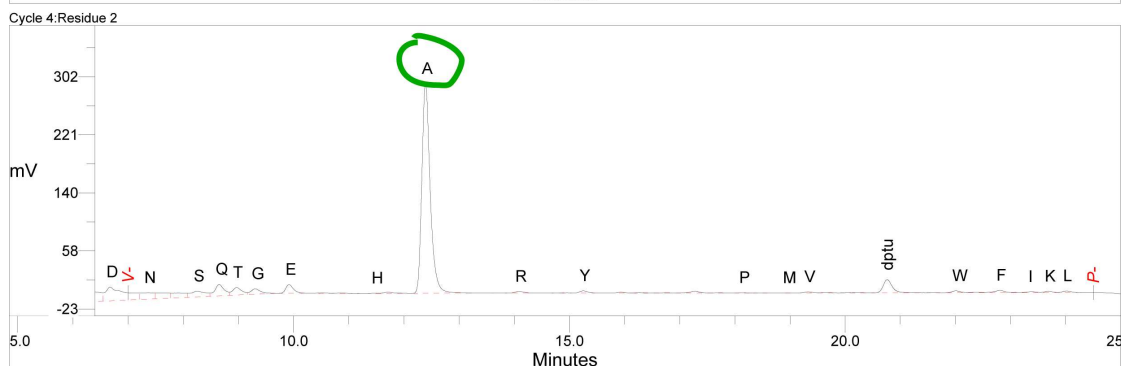
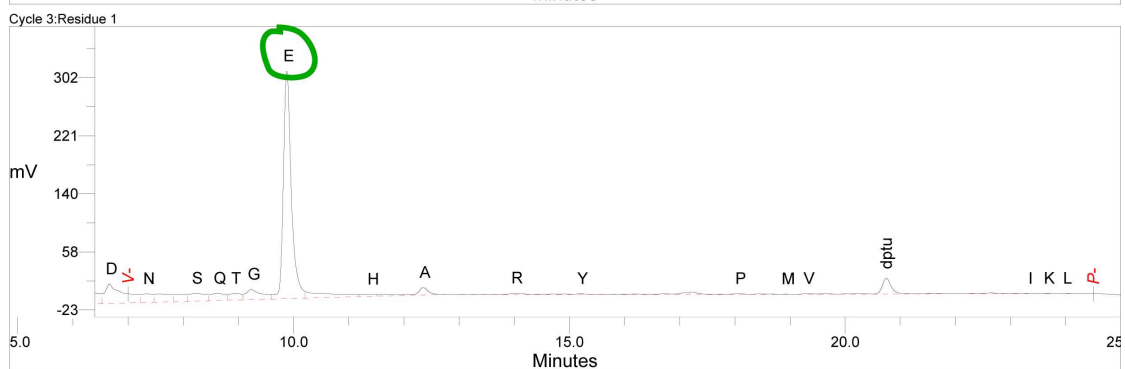
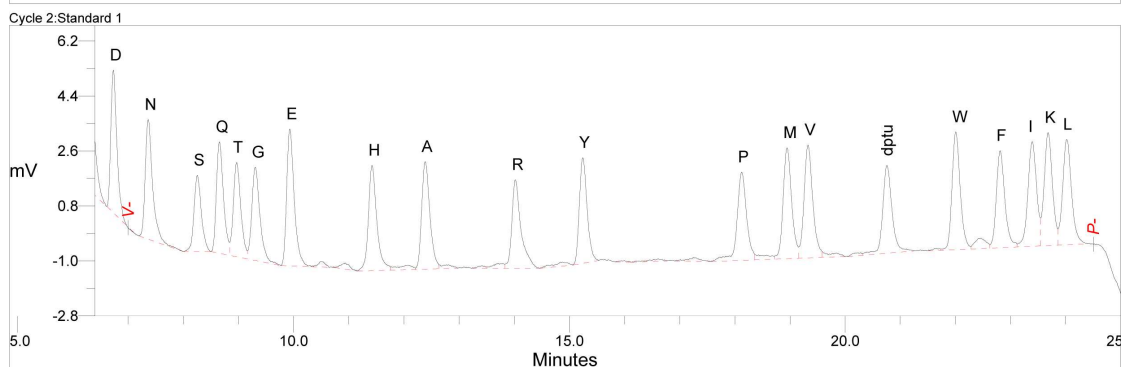
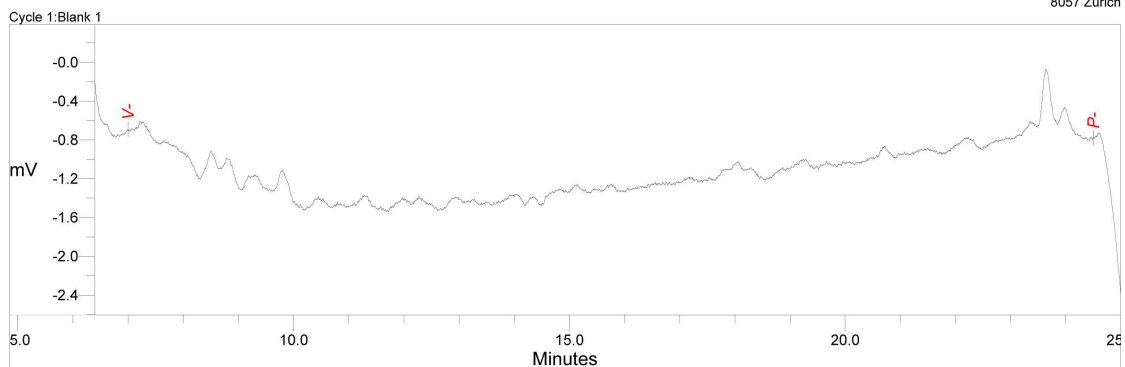
Name:	PROClSE	Model Number:	492
Method:	Pulsed-liquid PVDF cLC	Cartridge:	B
Operator:			

CHEMICAL INFORMATION

Run	Sample ID	Date	Cartridge	Std Amount	Sample Amount	Detector Scale
R1	1111174	18 October, 2012	X3	2.000 pmols	2.000 pmols	0.005 AUFS
R2	0607095	18 October, 2012	PTH Column	2.000 pmols	2.000 pmols	0.005 AUFS
R3	1110207	31 October, 2012	Solvent A3	2.000 pmols	2.000 pmols	0.005 AUFS
R4	1106133	18 October, 2012	Solvent B2	2.000 pmols	2.000 pmols	0.005 AUFS
R5	HP	18 October, 2012	Premix	2.000 pmols	2.000 pmols	0.005 AUFS
S1	1109498	30 March, 2012	Guard Column	2.000 pmols	2.000 pmols	0.005 AUFS
S2	1112713	31 October, 2012	Cartridge Seals	2.000 pmols	2.000 pmols	0.005 AUFS
S3	1103492	15 October, 2012	Glass Fiber Filter	2.000 pmols	2.000 pmols	0.005 AUFS
S4	1110100	31 October, 2012		2.000 pmols	2.000 pmols	0.005 AUFS
X1	LiChrosolv	31 October, 2012		2.000 pmols	2.000 pmols	0.005 AUFS
X2	LiChrosolv	31 October, 2012		2.000 pmols	2.000 pmols	0.005 AUFS
Total Cycles Count						

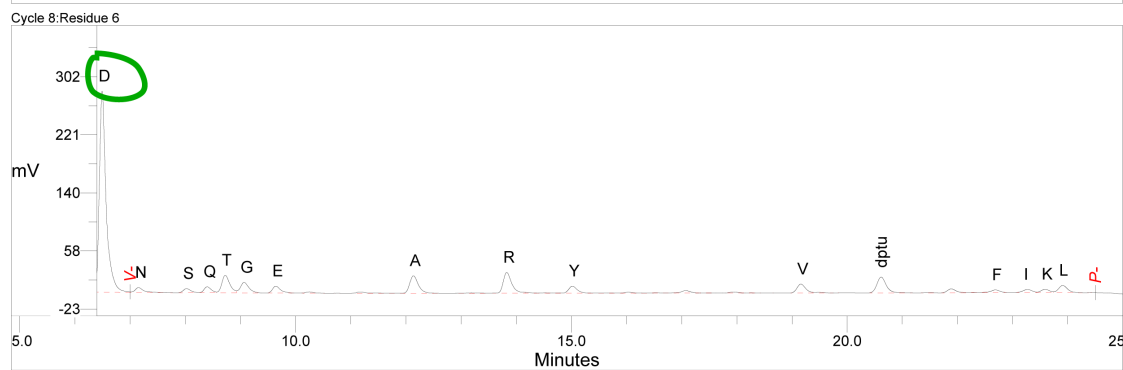
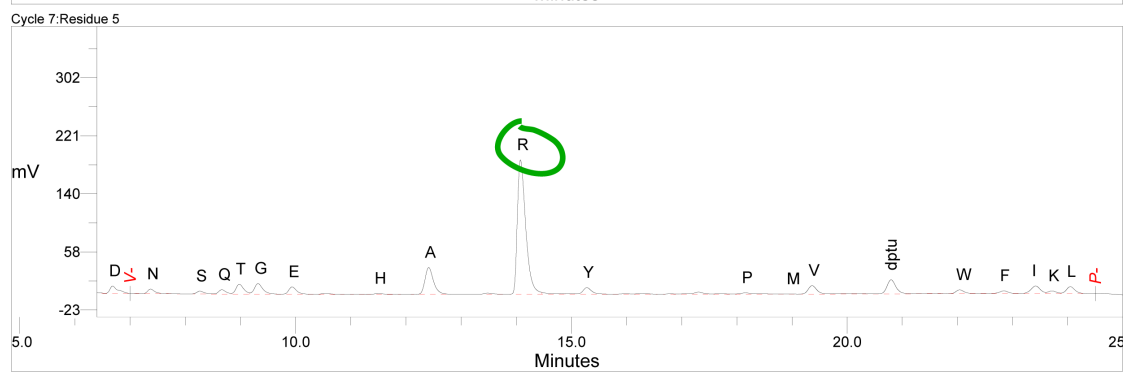
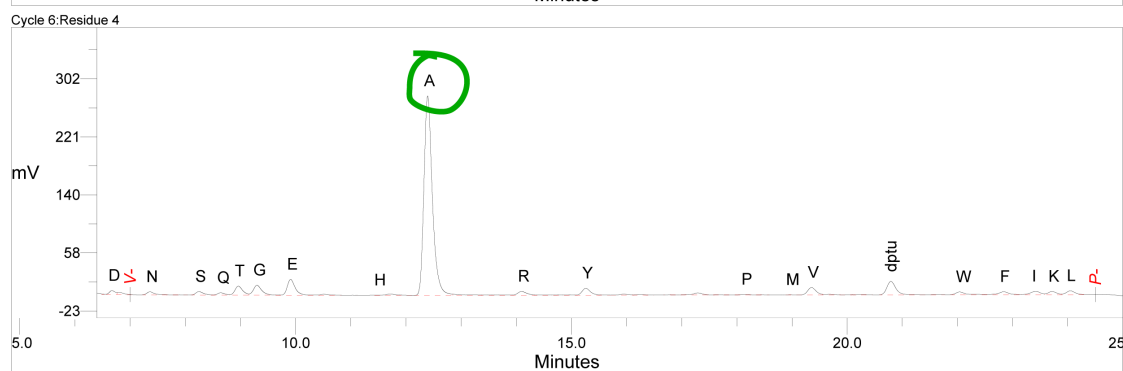
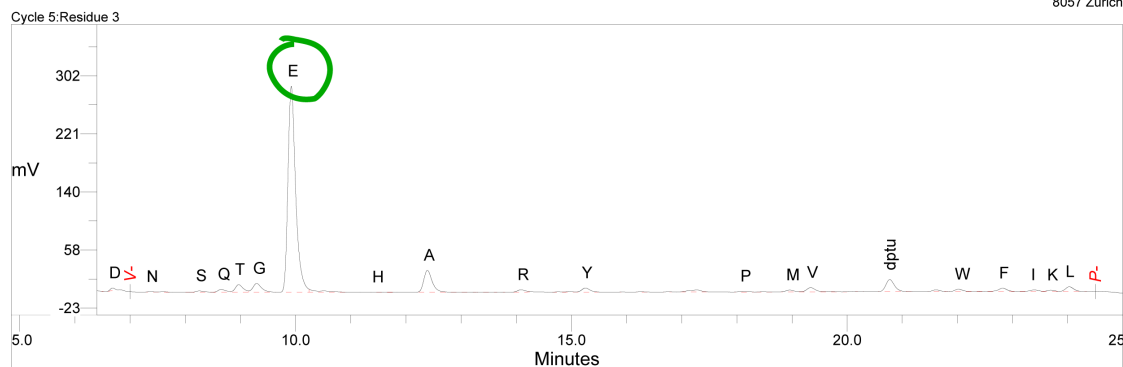
SequencePro™

Protein Analysis Group
 Functional Genomics Center
 Uni/ETH Zürich
 Winterthurerstr. 190
 8057 Zürich



SequencePro™

Protein Analysis Group
 Functional Genomics Center
 Uni/ETH Zürich
 Winterthurerstr. 190
 8057 Zürich



SequencePro™

Protein Analysis Group
Functional Genomics Center
Uni/ETH Zürich
Winterthurerstr. 190
8057 Zürich

Raw Pmol Table

	ASP	ASN	SER	GLN	THR	GLY	GLU	HIS	ALA	ARG	TYR	PRO	MET	VAL	TRP	PHE	ILE	LYS	LEU
1	9.631	2.008	1.067	0.743	0.807	4.691	139.5	0.122	5.443	1.150	0.405	0.398	0.114	0.353	0.000	0.000	0.195	0.289	0.292
2	8.311	4.557	6.438	8.840	6.821	4.964	5.454	0.198	167.3	1.543	1.827	0.212	0.100	0.803	1.391	2.069	0.781	0.987	1.337
3	1.591	0.264	1.431	2.168	6.866	8.029	128.6	0.092	17.48	2.193	3.338	0.712	1.457	3.271	1.804	3.032	1.345	1.077	3.980
4	2.310	2.213	3.931	1.719	8.113	8.984	9.800	0.125	159.2	3.719	5.651	0.624	0.136	5.671	2.022	2.508	2.702	2.324	3.105
5	4.582	3.422	3.227	3.337	8.941	9.660	4.542	0.661	21.36	130.0	5.337	1.189	0.057	6.392	2.920	2.657	6.536	2.119	5.765
6	120.9	3.563	4.551	4.505	15.86	9.635	4.318	0.000	13.98	20.13	5.746	0.000	0.000	6.737	0.000	2.419	2.622	2.576	5.946

“Adoramos a perfeição, porque não a podemos ter; repugná-la-íamos, se a tivéssemos. O perfeito é desumano, porque o humano é imperfeito.”

“We worship perfection because we can't have it; if we had it, we would reject it. Perfection is inhuman, because humanity is imperfect.”

Fernando Pessoa – *Autobiografia sem factos*