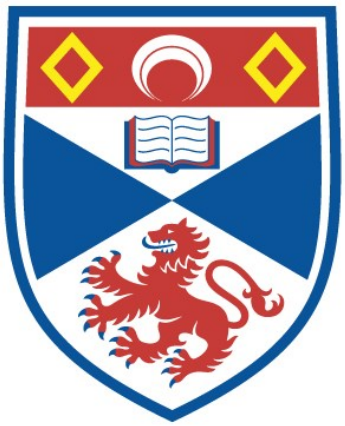
ARTIFICIAL METALLOENZYMES IN CATALYSIS

Lorenz Obrecht

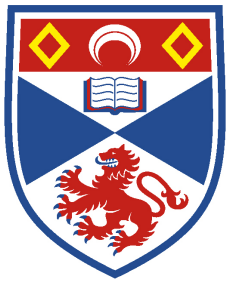
A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



2015

ARTIFICIAL METALLOENZYMES IN CATALYSIS

Lorenz Obrecht



University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of PhD
at the
University of St Andrews

2015

1. Candidate's declarations:

I, Lorenz Obrecht, hereby certify that this thesis, which is approximately 44000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in August, 2010 and as a candidate for the degree of Doctor in Philosophy in December, 2011; the higher study for which this is a record was carried out in the University of St Andrews between 2010 and 2015.

I, Lorenz Obrecht, received assistance in the writing of this thesis in respect of language, grammar, spelling and syntax, which was provided by my father Werner Obrecht

Date …… signature of candidate ……

2. Supervisor's declaration:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Doctor of Philosophy in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date …… signature of supervisor ……

3. Permission for publication: (*to be signed by both candidate and supervisor*)

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the

work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

PRINTED COPY

Embargo on all or part of print copy for a period of 5 years on the following ground: Publication would preclude future publication

ELECTRONIC COPY

Embargo on all or part of electronic copy for a period of 5 years on the following ground: Publication would preclude future publication

Date …… signature of candidate …… signature of supervisor ……

Abstract

This thesis describes the synthesis, characterisation and application of artificial metalloenzymes as catalysts. The focus was on two mutants of SCP-2L (SCP-2L A100C and SCP-2L V83C) both of which possess a hydrophobic tunnel in which apolar substrates can accumulate.

The crystal structure of SCP-2L A100C was determined and discussed with a special emphasis on its hydrophobic tunnel.

The SCP-2L mutants were covalently modified at their unique cysteine with two different *N*-ligands (phenanthroline or dipicolylamine based) or three different phosphine ligands (all based on triphenylphosphine) in order to increase their binding capabilities towards metals. The metal binding capabilities of these artificial proteins towards different transition metals was determined. Phenanthroline modified SCP-2L was found to be a promising scaffold for Pd(II)-, Cu(II)-, Ni(II)- and Co(II)-enzymes while dipicolylamine-modified SCP-2L was found to be a promising scaffold for Pd(II)-enzymes. The rhodium binding capacity of two additional phosphine modified protein scaffolds was also investigated. Promising scaffolds for Rh(I)- and Ir(I)-enzymes were identified.

Rh-enzymes of the phosphine modified proteins were tested in the aqueous-organic biphasic hydroformylation of linear long chain 1-alkenes and compared to the Rh/TPPTS reference system. Some Rh-enzymes were found to be several orders of magnitude more active than the model system while yielding comparable selectivities. The reason for this remarkable reactivity increase could not be fully elucidated but several potential modes of action could be excluded.

Cu-, Co-, and Ni-enzymes of *N*-ligand modified SCP-2L A100C were tested in the asymmetric Diels-Alder reaction between cyclopentadiene and *trans*-azachalcone. A promising 29% ee for the exo-product was found for the phenanthroline modified protein in the presence of nickel.

Further improvement of these catalyst systems by chemical means (e.g. optimisation of ligand structure) and bio-molecular tools (e.g. optimisation of protein environment) can lead to even more active and (enantio)selective catalysts in the future.

Table of contents

List of abbreviations

1	(Maleimido)propionic acid hydrazide
acac	Acetylacetonato
BSA	Bovine serum albumin
CD	Circular dichroism
COD	1,5-cyclooctadiene
CP	Cyclopenta-1,3-diene
CTAB	Hexadecyltrimethylammonium bromide
DA	Diels-Alder
DCM	Dichloromethane
DMF	*N*,*N*-Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease (DNase I from bovine pancreas)
DTT	Dithiothreitol
E. coli	Escheria coli
ee	Enantiomeric excess
EDTA	Ethylendiaminetetraacetic acid
EPR	Electron paramagnetic resonance
ESI	Electrospray ionisation
ESI^+	Electrospray ionisation, positive mode
eq	Equivalent
fig.	Figure
FMO	Frontier molecular orbital
GC	Gas chromatography
h	Hour(s)
HF	hydroformylation
hMFE	Human multifunctional enzyme
HOMO	Highest occupied molecular orbital
HPLC	High pressure liquid chromatography

HSA	Human serum albumin
ICP	Inductively coupled plasma
IPA	Isopropanol
IS	Internal standard
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kan	Kanamycin
KPi	Potassium phosphate (buffer)
LC-MS	Liquid chromatography-mass spectrometry
LB	Lysogeny broth
LUMO	Lowest unoccupied molecular orbital
l/b	Linear to branched ratio
MALDI	Matrix assisted laser desorbtion/ionisation
MeCN	Acetonitrile
MES	2-(*N*-morpholino)ethanesulfonic acid
min.	Minute(s)
MS	Mass spectrum
MFE	Multifunctional enzyme
MWCO	Molecular weight cut off
NMR	Nuclear magnetic resonance
OAc	Acetoxy group
O/N	Over night
PB	Production broth
PBS	Phosphored buffered saline
Phen	5-Maleiamido-1,10-phenanthroline
Picol	*N*,*N*-Dipicolyl-*N*'-maleoyl-3-aminopropanamide
ppb	Parts per billion
ppm	Parts per million
P(meta)	*meta*-Diphenylphosphino benzaldehyde
P(ortho)	*ortho*-Diphenylphosphino benzaldehyde
P(para)	*para*-Diphenylphosphino benzaldehyde
PYP	Photoactive yellow protein

RNA	Ribonucleic acid
rpm	Revolutions per minute
RT	Room temperature
SCP	Sterol carrier protein
SCP-2L	Sterol carrier protein 2-like domain of the human mult functional enzyme
SCP-2L A100C	SCP-2L in which amino acid alanine (A) at position 100 was replaced by cysteine (C)
SCP-2L V83C	SCP-2L in which amino acid valine (V) at position 83 was replaced by cysteine (C)
SDS	Sodium dodecyl sulfate
syn gas	Synthesis gas, a 1:1 mixture of H_2 and CO
TCEP	Tris(2-carboxyethyl)phosphine
TEV	Tobacco etch virus
TFA	Trifluoro acetic acid
THF	Tetrahydrofuran
TON	Turnover number(s)
TPP	Triphenylphosphine
TPPTS	Sodium triphenylphosphine trisulfonate
Tris	Tris(hydroxymethyl)aminoethane
TS	Transition state(s)
UV-VIS	Ultraviolet-visible

Artificial (metallo)enzyme abbreviation scheme:

The artificial (metallo)enzymes used within this thesis are abbreviated following the scheme

1) Protein Scaffold used
2) Mutations (if present)
3) Abbreviation of modification A (if present)
4) Abbreviation of modification B (if present)

5) Metal (if present)

Part 1) and 2) are separated by a space, all other parts are separated by "-".

Examples:

1) "SCP-2L A100C-1-P(para)-Rh" is the abbreviation for:
SCP-2L in which amino acid alanine (A) at position 100 was replaced by cysteine (C) which is modified (at the cysteine) with (maleimido)proppionic acid hydrazide (abbreviated as ”1”). A second modification (at the hydrazide) with *para*-diphenylphosphino-benzaldehyde (abbreviated as “P(para)”)is present. Rh is coordinating to it.

2) "SCP 2L V83C-*Phen*-Cu" is the abbreviation for:
SCP-2L in which amino acid valine (V) at position 83 was replaced by cysteine (C) which is modified (at the cysteine) with 5-Maleimido-1,10-phenanthroline (abbreviated as “*Phen*”). Cu is coordinating to it.

Chapter 1: Alternative approaches for the aqueous-organic biphasic hydroformylation of higher alkenes

1.1 Abstract

Many different approaches to enhance the low activity of the Rh/TPPTS-system used in the aqueous-organic biphasic hydroformylation of higher alkenes exist. In addition many alternative approaches trying to tackle the low activity while maintaining a high selectivity have been developed as well. All these systems are discussed in this chapter. Additionally, an outlook on this thesis is given.

This chapter was published before. Due to corrections it is slightly changed. The sub chapters “outlook on this thesis”, “Overview on the different aqueous-organic biphasic hydroformylation concepts”**,** as well as the “abstract” were not part of this publication.

L. Obrecht, P. C. J. Kamer, W. Laan, *Catal. Sci. Technol.*, 2013, **3**, 541
http://pubs.rsc.org/en/content/articlelanding/2013/cy/c2cy20538f#!divAbstract

1.2 Introduction

Hydroformylation (also known as oxo synthesis or oxo process) is one of the largest industrial applications of homogeneous catalysis, with a current world-wide production of approximately 9 million tons per year.[1] The process entails the transition-metal catalysed reaction of alkenes with carbon monoxide and hydrogen gas affording aldehydes. The aldehydes are easily hydrogenated to alcohols, which can be converted to important products such as detergents and plasticizers. Both branched and linear aldehydes can be formed, but their more prevalent use in downstream processes makes linear aldehydes the more desirable products.[2]

Catalyst separation is crucial for industrial processes. Catalyst recovery allows for the implementation of catalyst recycling strategies, thus reducing the overall process costs. Also contamination of the product with the catalyst or metal species can pose problems for downstream processes or applications. This issue has been very effectively tackled for the hydroformylation of propene and butene by the development of the aqueous biphasic Ruhrchemie/Rhône-Poulenc (RCH/RP) process, commercially used since 1984.[3] Water is a cheap, environmentally friendly and safe solvent, which renders it very attractive for green and sustainable production processes. The RCH/RP process uses the highly water-soluble ligand triphenylphosphine-3,3',3''-trisulfonate (TPPTS) to immobilize a rhodium catalyst in the aqueous phase, while the substrate and products form a second phase. The catalyst can be completely and quickly recovered by phase-separation. The process produces almost exclusively aldehydes with over 96% selectivity for the linear products under relatively mild conditions - “mild” in comparision to other hydroformylation setups. This technology is currently in operation at five plants worldwide, furnishing an annual production of 800000 tons of aldehydes. [4] The operational simplicity, robustness and the excellent economics of the process (loss of rhodium by leaching into the organic phase lies in the ppb range) make the RCH/RP a benchmark process in the field of aqueous biphasic transition-metal catalysis.

The solubilities of propene and C4 alkenes in the aqueous phase are sufficient to allow chemical reactions to occur at an acceptable rate without phase transfer limitations.

However, for longer alkenes, the reaction rate is too low to be economically viable. For example, hydroformylation of 1-octene under standard conditions (125 °C, 30 bar, [Rh] = 300 ppm) proceeds with a rate constant of $5.3\times10^{-4}\,min^{-1}$.[2]

For the hydroformylation of higher alkenes (>C5), homogeneous processes are used. For alkenes up to C8 commercial processes using phosphine-modified rhodium catalysts are known. For longer alkenes the distillation conditions required to separate the high-boiling products may lead to catalyst decomposition and concomitant loss of metal. Therefore cheaper cobalt catalysts are used, ensuring the processes are still economically viable. However, cobalt catalysts are less active than rhodium-catalysts, requiring higher reaction temperatures and pressures.[5]

For these reasons the development of a process for the hydroformylation of higher alkenes using rhodium catalysts under biphasic conditions, which allows for catalyst recycling by phase separation, is highly desirable. Using the unmodified Rh–TPPTS system, only minor improvements can be achieved by varying the reaction conditions such as the syngas pressure or the ligand–metal ratio.[6] Many approaches have been developed to tackle the issue of low space-time yield in biphasic aqueous–organic reaction systems involving poorly water-soluble substrates. In this perspective, we will discuss various alternative approaches which have been developed to overcome the mass-transfer limitations in the biphasic hydroformylation of higher 1-alkenes. Unless stated otherwise, the catalyst system applied in the studies discussed below is based on the Rh–TPPTS system.

Additives

Various additives are available to improve the solubility or the transport of hydrophobic substrates into the aqueous phase. To be suitable, an additive should (i) be inert with respect to the catalyst, substrate and reaction products; (ii) not increase the solubility of the catalyst in the product phase; (iii) exert a minimal influence on phase separation and (iv) show minimal leaching into the product phase.

Co-solvents

One of the first strategies employed to improve mass transfer in biphasic catalysis is the

use of co-solvents. Co-solvents increase the lipophilicity of the aqueous phase, thereby increasing the solubility of alkenes in the catalyst phase. For example, the solubility of 1-octene has been estimated to be 10^4 times higher in 50% ethanol than in pure water.[7] Indeed, the use of ethanol, acetone, acetonitrile or methanol as co-solvent in the hydroformylation of 1-octene leads to a significant increase of the reaction rate.[7–9] However, the addition of co-solvents results in a decrease in the linear selectivity, i.e. from 98% to 92% upon addition of 20% methanol to the alkene phase.[6] Transfer of the co-solvent from the aqueous to the organic phase due to the formation of nonanal has been reported.[7] This might affect the partition coefficient of the catalyst between the two phases. Also, ethanol was found to react with the product nonanal to yield acetals. A buffer solution of sodium carbonate and bicarbonate in the aqueous phase (pH=10) was found to be effective in suppressing acetal formation.[7]

The erosion of the linear selectivity, the fact that they react with the product, their leaching into the product phase as well as the potential leaching of catalyst make co-solvents an unattractive additive for commercial applications.

Surfactants

Above the critical micelle concentration surfactants form micelles, which can solubilize hydrophobic compounds within their hydrophobic cores. The formation of micelles increases the interfacial area in biphasic systems, thus increasing the phase transfer. The first reports of rate-enhancements in biphasic hydroformylation by using surfactants came from Johnson Matthey.[10] Despite their ability to encapsulate substrates, neutral and anionic surfactants show no activating effect on reaction rates.[6,11] In contrast, cationic surfactants greatly enhance reaction rates. For example, no aldehydes are formed during the hydroformylation of 1-dodecene in the presence of sodium dodecyl sulfate (SDS), while the use of cetyltrimethylammonium bromide (CTAB **1**, Fig. 1) afforded 61% conversion.[11] High reaction rates can be obtained using CTAB: turnover frequency (TOF) values up to 900 h^{-1} have been reported.[12]

The difference in the ability of anionic and cationic surfactants to enhance the reaction rate has been ascribed to electrostatic interaction between the surface of the micelle and the catalyst. Electrostatic repulsion between the negatively charged catalyst and the

surface of micelles formed by anionic surfactants prevents reaction of the catalyst with the substrate. In the case of cationic surfactants, however, electrostatic attraction between the negatively charged sulfonate group of the TPPTS ligand and the cationic headgroup of the surfactant leads to sequestering of the catalyst on the micelle surface, providing a high local concentration of catalyst in the vicinity of a high local concentration of substrate and thus causing an increase in reaction rate. Indeed, ^{31}P-NMR data suggest a strong interaction between the catalyst and CTAB micelles.[13] Light-scattering experiments confirmed that the rate increase of the hydroformylation of 1-dodecene in the presence of CTAB is due to solubilisation of the substrate inside the micelle and the binding of [HRh(CO)[TPPTS$_3$] to the micelle surface, leading to a 90 fold increase of the rhodium concentration in the micelle interface compared to the bulk water phase.[14]

The reported effects of surfactants on the linear selectivity are inconsistent. For instance, for the use of CTAB, both an increase[11] and a decrease[6] in the linear selectivity have been reported. Interestingly, Chen et al. observed a synergistic effect of Na-TPPTS and Na-TPPDS (disodium triphenylphosphine- 3,3'-disulfonate) on the regioselectivity of 1-dodecene hydroformylation in the presence of CTAB, increasing the ratio of linear to branched products (*l*/*b*) from 6.5 to 22.3 for TPPTS/ TTPDS = 2 : 1.[15]

Motivated by the high reaction rates obtained with CTAB, several other types of cationic surfactants have been explored. Li, Chen and co-workers found that gemini surfactants **2** (Fig. 1) containing alkyl bridges afford slightly higher rates and selectivities than CTAB in the hydroformylation of 1-dodecene.[16] Notably, shorter alkyl bridges lead to increased selectivities. The surfactants with shorter bridges were suggested to form more compact micelles, favouring the formation of the less crowded linear aldehyde. Also gemini-surfactants **3** and piperazine-based **4**, as well as trimeric triazine-based surfactant **6** (Fig. 1) were explored, with **4** performing the best, affording a TOF of 1845 h^{-1} in the hydroformylation of 1-decene.[17]

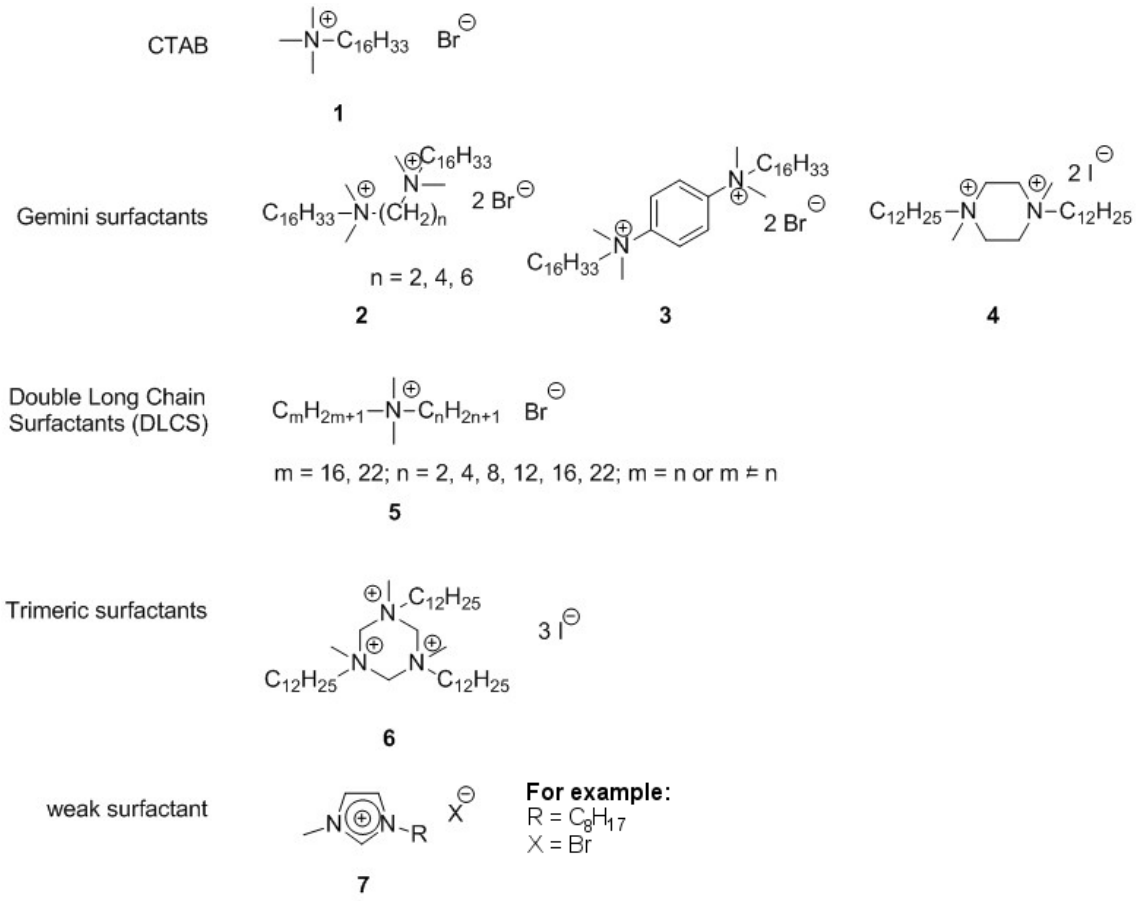


Fig. 1 Cationic surfactants applied in aqueous biphasic hydroformylation.

Nonetheless, the highest rate enhancements to date have been achieved by using double long-chain surfactants (DLCS, **5**). Using **5** (Fig. 1), Fu et al. achieved a TOF of 7472 h^{-1} for the biphasic hydroformylation of 1-dodecene, with an aldehyde selectivity of 94%; values which are comparable to those of the equivalent homogeneous reaction.[18] Threshold values for the length of both chains were determined: high reaction rates are obtained when $n \geq 8$ if $m = 22$ and $n \geq 12$ if $m = 16$ (see Fig. 1).[18]

Although the addition of surfactants provides high reaction rates and the leaching of metal into the product phase is generally low, it often makes reaction mixtures prone to emulsification, particularly at high stirring rates and conversions, rendering phase separation difficult.[19]

Desset et al. found that the weak surfactant 1-octyl-3-methylimidazolium bromide **7** ([Octmim]Br, Fig. 1; $X = Br^-$, $R = C_8H_{17}$) provides high reaction rates while at the same

time allows rapid and complete phase separation and good catalyst retention, albeit with a slight decrease in the linear selectivity. While imidazolium and triethylammonium salts with longer alkyl ‘‘tails’’ (R>C8) also provide large rate enhancements, these amphiphiles lead to stable emulsions.[20,21] It appears that [Octmim]Br and related molecules are promising surfactants for increasing the rate of aqueous biphasic hydroformylation without compromising the inherent strength characteristic of the RCH/RP process.

Polyethylene glycol

Polyethylene glycol (PEG) of medium molecular mass (400–1000) is an effective additive for increasing the reaction rate.[6] It significantly enhances the reaction rate, although it leads to a slight decrease in the *l*/*b* ratio. Leaching of the PEG and rhodium into the organic phase is very low, and phase separation is fast and straightforward.

Inverse phase transfer catalysts

Another approach to facilitate the migration of hydrophobic molecules from an organic to an aqueous phase is the use of inverse phase transfer catalysts. The two most widely studied classes of compounds are cyclodextrins[22,23] and calixarenes,[24] which are widely applied in aqueous biphasic catalysis. Cyclodextrins **8** (CD, Fig. 2) are cyclic oligosaccharides, and the three main types (α, β or γ) are composed of 6, 7 or 8 α-D-glucopyranoside units respectively. CDs are shaped like conical cylinders with a hydrophobic inner surface and a hydrophilic outer surface; the wider rim contains secondary hydroxyl-groups while primary hydroxyl groups occupy the narrower rim ($R_1 = R_2 = R_3 = OH$).

α- and γ-CD show no effect while β-CD only provides a moderate 2-fold increase in the reaction rate.[25,26] However, the effectiveness of CDs can be greatly improved by chemical modification of the hydroxyl-groups with hydrophobic or hydrophilic groups such as methyl, acetyl or 2-hydroxypropyl.

Randomly methylated β-CD (RAME-β-CD; on average 12.6 methyl groups added on positions 2, 3 and 6) is particularly effective.[25,26] This was initially ascribed to the solubility of the modified CD in both the aqueous and organic phase. Indeed, CDs

which are little soluble in either water or the organic phase have little influence on the reaction rate. Complementary experiments have demonstrated that modified CDs concentrate at the liquid–liquid interface, where substrate recognition leads to the formation of inclusion complexes which facilitate the reaction between the substrate and the water-soluble organometallic catalyst.[27,28] Interestingly, positive synergistic effects of binary mixtures of RAME-α-,β-, and γ-CDs have been observed, suggesting the formation of 2 : 1 ternary complexes.[29]

Unfortunately, the use of RAME-β-CD leads to a significant reduction of the linear selectivity. This decrease in selectivity stems from the formation of inclusion complexes between TPPTS and the CD. The formation of such complexes facilitates ligand dissociation, leading to catalyst species which afford lower linear selectivity.[30,31] To prevent this, various combinations of CD and ligands have been explored.

Due to their smaller cavity size, chemically modified α-CDs do not interact with the catalyst, and RAME-α-CD enhances the reaction rate while not affecting the *l*/*b* ratio.[32] Also α- and β-CDs bearing alkyl chains on the secondary face and sulfoalkyl- or poly(ethyleneoxide) chains on the primary face afford higher rates and regioselectivity than randomly methylated β-CD.[33–36]

Noteworthy is the use of the sulphonated xantphos ligand (**9**, Fig. 2) in combination with RAME-α and β-CDs. Although the CDs bind a ligand phenyl ring, ligand dissociation does not occur and improved chemo- and regioselectivities in the hydroformylation of 1-octene and 1-decene are obtained.[37] The ligand 1,3,5-triaza-7-phosphaadamantane **10** (PTA) and its *N*-benzylated derivative **11** (Fig. 2) were recently found not to interact with RAME-β-CD during the biphasic hydroformylation of 1-decene. Although this system affords poor linear selectivity, PTA may provide a suitable framework for ligand optimization.[38] In general, upon addition of CDs to biphasic hydroformylation reactions the phase-separation is still excellent, and leaching of the catalyst is minimal. Also catalyst recycling can be done without significant loss of activity. Even an increase in activity upon recycling has been reported; this was attributed to a gradual supramolecular organization of the rhodium complex, its solvation sphere and cyclodextrin in the interphase.[39]

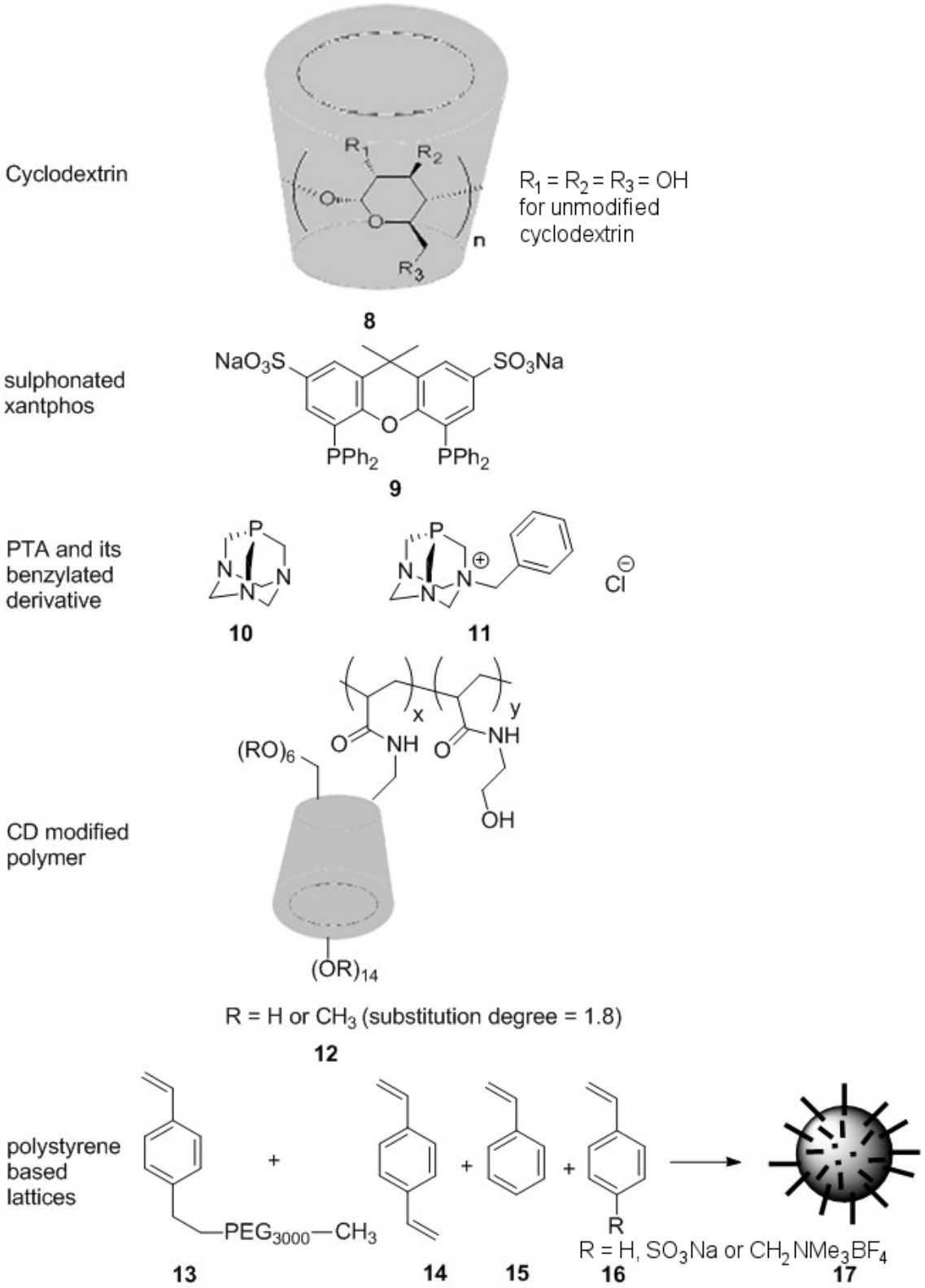


Fig. 2 Inverse phase transfer catalysts applied in aqueous biphasic hydroformylation.

Supramolecular organization has also been suggested to explain the performance of recently introduced β-CD functionalized polymers in biphasic hydroformylation.[40] Acrylamide polymers **12** with up to 50% degrees of β-CD substitution were synthesized (Fig. 2). Whereas highly substituted polymers were found to be nearly as effective as RAME-β-CD as mass-transfer promoters in the hydroformylation of 1-decene, the ineffectiveness of less substituted polymers was interpreted to stem from unfavourable organisation of the main-chains of the polymers at the phase-boundary, leading to shielding of the substrate from the CDs. With 1-hexadecene as substrate, highly substituted polymers were much more effective than RAME-β-CD. Whereas the substrate is too long to be efficiently transferred by a single β-CD cavity, cooperative multivalent substrate recognition by close-in-space polymer-bound CDs leads to efficient recognition and higher reaction rates.

Calixarenes are cyclic oligomers of substituted benzene units, synthesised from phenols and aldehydes. While calixarenes have been applied as phase-transfer agents in a number of biphasic catalysis reactions, to the best of our knowledge their use as additive in biphasic hydroformylation of linear alkenes has not been reported. Due to the performance of phosphine-modified Calixarenes as ligands (described further on) they are probably a suitable additive as well.

The group of Vogt recently applied polystyrene-based latices as phase transfer agents in the biphasic hydroformylation of 1-octene.[41] The latices **17**, obtained from the polymerization of PEGylated styrene **13**, divinylbenzene **14**, styrene **15** and styrylsalts **16** (Fig. 2; R = H, SO_3Na or $CH_2NMe_3BF_4$), consist of a lipophilic core and a hydrophilic shell. It was found that the rate-enhancement critically depends on the nature of the styrylsalt: the highest rate (TOF of 150 h^{-1}) was achieved using ammonium-modified styrene, while latices containing unmodified or sulfonated styrene had little effect. In analogy to the situation with cationic surfactants, this was explained by the association of the anionic catalyst with the cationic ammonium-group on the outer shell.

Activated carbon

Recently Monflier introduced activated carbon (e.g. Nuchar® WV-B) as a novel mass-transfer additive for aqueous biphasic catalysis.[42] Considerable rate-enhancements are achieved in biphasic hydroformylation using either cobalt/trisulfonated tris(biphenyl)phosphine (BiphTS) or the Rh–TPPTS system, with little effect on the *l/b* ratio. Also the aqueous phase catalyst and activated carbon can be recovered easily and recycled multiple times without significant loss of activity.[43,44] It was proposed that the activated carbon facilitates the mixing of the aqueous and organic phases, leading to a more effective interfacial area concomitant with confinement of the reactants and the catalyst.[42]

Ligand variation

TPPDS, TPPMS, TPP

Substituting TPPTS by the less water-soluble triphenylphosphine- 3,3'disulfonate (TTPDS) or triphenylphosphine-3- sulfonate (TPPMS) as ligand leads to a slight increase in reaction rate, but also leads to significant leaching of the ligand into the organic phase.[6,45]

Chaudhari et al. proposed the use of promoter ligands which are exclusively soluble in the organic phase. This was suggested to lead to the formation of mixed ligand complexes which would concentrate at the phase boundary, affording interfacial catalysis. Although inclusion of triphenylphosphine (TPP) in the hydroformylation of 1-octene with $HRh(CO)(TPPTS)_3$ leads to a 10–50 fold increase in the reaction rate,[46] it has been argued that TPP simply increases the solubility of the catalytic species in the organic phase, which may affect downstream processing.[6]

Amphiphilic ligands

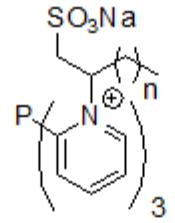
Another approach that has been explored is the use of ligands which combine surface-activity with water solubility. Analogous to surfactants, the aggregation of amphiphilic ligands may lead to the formation of micelles, which are capable of encapsulating substrates in their hydrophobic core. This increases the substrate solubility in the

aqueous phase and also brings the substrate and the catalyst in close proximity, thus increasing the reaction rate. Fell and Papadogianakis were the first to use a specifically designed amphiphilic ligand in biphasic hydroformylation.[47] Application of the zwitterionic trisulfoalkylated tris(2-pyridyl)-phosphine (**18**, Fig. 3) for the conversion of 1-tetradecene resulted in TOFs up to 340 h^{-1}. The best yields were obtained with the ligand where n = 5; longer chains afforded lower yields while stable emulsions were formed with n = 9, 11. The catalysts with n = 0–7 could be quantitatively recovered by a simple phase separation. Various other amphiphilic monodentate ligands superior to TPPTS in biphasic hydroformylation have been reported. Most of these ligands bear the hydrophilic group (either sulfonate or phosphate) and the phosphorous atom on the two ends of the alkyl chain, including phosphonatephosphines developed by Bischoff and the Hanson group's sulfonated tris(ω-phenylalkyl)phosphines.[8,48–51] The sodium salt of sulfonated *n*-$C_{12}H_{25}O$ $C_6H_4P(C_6H_4$-*p*$(CH_3O)_2$ (DMOPPS, **19**, Fig. 3) affords significant rate enhancements in the biphasic hydroformylation of long chain alkenes, not only providing TOFs up to 673 h^{-1} for 1-dodecene but even a TOF of 100 h^{-1} for 1-hexadecene, with *l/b* ratios around 2.4.[52]

Surface active ligands based on diphosphine frameworks such as 2,2'-bis(diphenylphosphino)-1,1'-biphenyl (BISBI) and 2,2'-bis(diphenylphosphino)-1,1'-binaphthyl (BINAP) have also been developed.[53] Sulfonated xantphos-derivatives containing long apolar chains (**20**, Fig. 3) spontaneously form large and thermally stable vesicles in water.[54] These vesicles were proven to increase the solubility of 1-octene in aqueous solution, and the activity of the vesicle-forming ligands in the biphasic hydroformylation was found to be 12–14 times higher than that of xantphos. It was shown that the aggregates stay intact during recycling, the rhodium is quantitatively retained in the aqueous phase and the TOF and linear selectivity remained the same in four consecutive runs.

Monflier and co-workers found that the performance of the surface-active ligands (**21**, **22**, Fig. 3) in the biphasic hydroformylation of 1-decene could be improved, by the addition of stoichiometric amounts of ionic β-cyclodextrins, leading to a 4-fold increase in reaction rate and an excellent phase separation. In contrast, the addition of neutral β-cyclodextrins led to a decrease in performance, which was shown to be due to

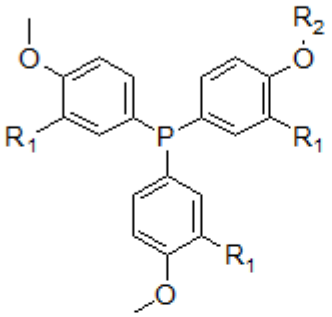
trisulfoalkylated
tris(2-pyridyl)phosphine



n = 0, 3, 5, 7, 9, 11

18

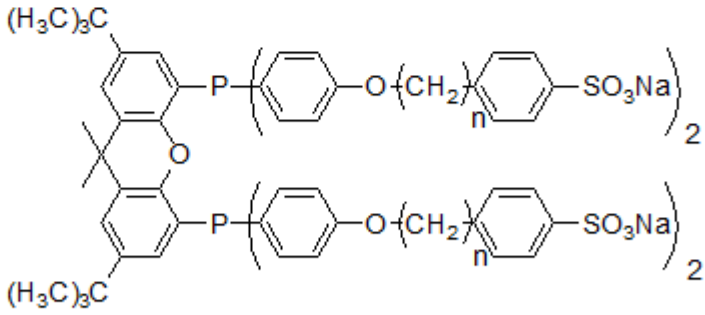
DMOPPS



$R_1 = SO_3Na$ or H
$R_2 = C_{12}H_{25}$

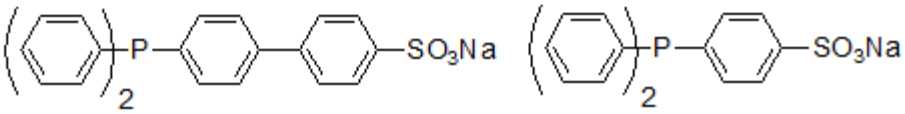
19

Xantphos
derivative



20

surface active
ligands used by
Monflier et al.



21 **22**

Fig. 3 Amphiphilic ligands used in the aqueous biphasic hydroformylation.

destruction of the micelles formed by the ligand.[55] This shows that for amphiphilic ligands, addition of CD with the right structure may provide an elegant way to further improve their performance in biphasic catalysis.

Ligands containing cyclodextrins or calixarenes

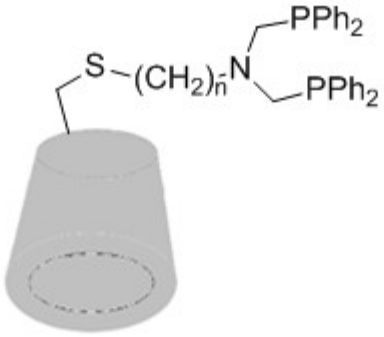
In addition to their use as additives in biphasic catalysis, cyclodextrins and calixarenes have also been incorporated as phase-transfer agents in ligand structures to develop dual-function catalysts.

Reetz reported the evaluation of a series of diphosphine-ligands modified with β-CD in the rhodium catalyzed biphasic hydroformylation of 1-octene (**23**, Fig. 4). Very high activities were achieved with *l*/*b* ratios of 3.2. Disappointingly, the activity dropped 50% upon reuse.[56] Also methylated α-cyclodextrins capped with a diphosphine/rhodium moiety were active catalysts in the biphasic hydroformylation of 1-octene, but significant leaching of the catalyst to the organic phase occurred over the course of the reaction.[57]

Using rhodium complexes of water-soluble calix[4]arene based ligands modified with two phosphines (**24**, Fig. 4) on the upper rim provided considerable rate-enhancements but slightly reduced *l*/*b* ratios in the biphasic hydroformylation of 1-octene. Remarkably, the system showed almost complete retention of activity and selectivity over two recycling runs.[58]

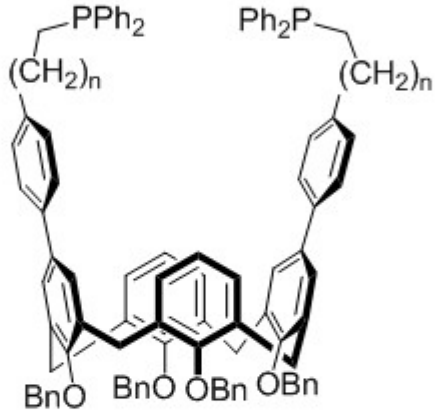
The use of hemispherical 1,3-calix[4]arene-diphosphites **25** (Fig. 4) as ligands in combination with upper rim sulfonated calix[4]arenes as surfactants in the biphasic hydroformylation of 1-octene and 1-hexene was recently reported by Matt and co-workers. In the absence of surfactant, catalysis takes place in the organic phase. In the presence of surfactant, the activity increased, and TOFs of 1160 h^{-1} and 750 h^{-1} were obtained for 1-octene and 1-hexene, respectively, with high *l*/*b* ratios (up to 61.8 for 1-octene using a 10-fold excess of ligand). However, stable emulsions were recovered after each reaction including surfactant. Although no recycling experiments were reported, this will probably hamper efficient catalyst recovery.[59]

β-CD modified diphosphines



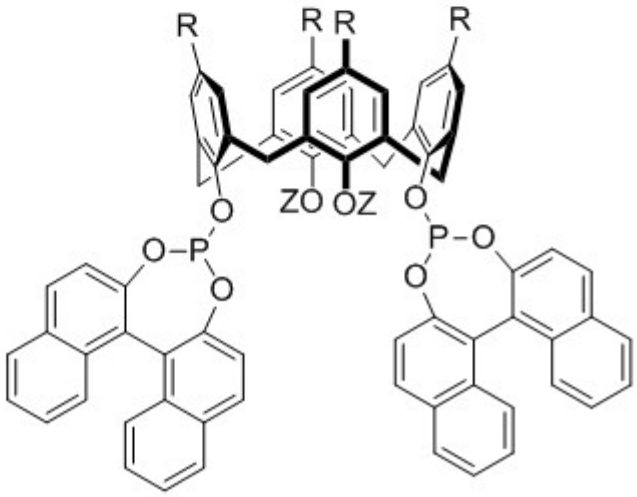
23

calix[4]arene based diphosphines



24

calix[4]arene based diphosphites



25

Fig. 4 Ligands containing cyclodextrins or calixarenes.

Polymer supported ligands

Water soluble polymers as support

In this approach the catalytically active metal is complexed to a water soluble polymer modified with donor groups. In most cases the coordinating groups are phosphorous based: mono-[60–66] or bidentate[67] phosphines, phosphine oxides,[68] phosphinites,[66] phosphonites[66] or bidentate P–N[69] ligands. There are also examples where bidentate N–N-ligands,[69] carbenes,[70] acetylacetonate[71] or glycolate[72] groups act as donor ligands for a rhodium metal, although these ligands afford less active hydroformylation catalysts.

These polymer bound catalysts are believed to act as phase transfer agents[64] and are successfully used in the aqueous biphasic hydroformylation of water insoluble alkenes. In general rhodium leaching to the organic phase is low and the catalyst can easily be reused. TOFs obtained are low (<100 h^{-1}) for these systems, while the *l*/*b* ratio varies from moderate to excellent (1.3–8.5) depending on the conditions, the polymer and the donor atoms used.[61,63–67,71]

There are two research groups who obtained remarkably high TOF (>500 h^{-1}) in the hydroformylation of long chain 1-alkenes when using water soluble polymers as support.

The group of Ritter used a polyethylene glycolate complex which was synthesised from rhodium trichloride and polyethylene glycol. The rhodium polyethylene glycolate complex was not definitively characterised. They assumed that the reaction takes place at the aqueous/organic interphase as the TOF is almost unaffected by the alkene chain length – and therefore by their water solubility; though they did not neglect the possibility of the catalyst being transferred into the organic phase.[72] For the hydroformylation of 1-dodecene the rhodium content in the organic phase after five catalytic cycles was measured as 1.9 ppm. The high TOF cannot be attributed to leached rhodium as upon addition of 4 equivalents of TPPTS to the reaction mixture the amount of leached rhodium drops below the detection limit of 0.1 ppm while the TOF even increased. Unfortunately their system is rather unselective, yielding a *l*/*b* ratio of ≈0.9.

Weberskirch and his group postulated that their tailor made amphiphilic block

copolymers form micelles encapsulating the olefin and the catalytic centre (see Fig. 5).[60,70,73,74]

All their catalysts were evaluated using 1-octene as substrate, resulting in TOFs as high as 3700 h^{-1} and a moderate selectivity (*l*/*b*≈3) in most cases. When rhodium leaching and catalyst recycling were investigated for one of their systems, the former was found to be low (0.4 ppm). Reusing this catalyst four times resulted in an increasing TOF in the first three cycles (from 1100 to 2185 h^{-1}), after which it remained constant in the next cycle. This effect was not clarified. On the other hand the *l*/*b* ratio of the aldehydes formed dropped from 2.57 to 1.33 during the first three cycles, indicating an increase in free rhodium, and remained constant in the last run.[70] Phase separation after the reaction is highly dependent on the polymer used. For some polymers a simple decantation can be applied while for others the separation was poor.

As a conclustion, the use of most water soluble polymers is unattractive due to low TOF. The use of polyethylene glycol as support gives high TOFs but low selectivities. The use of tailor made amphiphilic block copolymers is probably the most interesting approach as the TOFs are high, the *l*/*b* ratio can be tuned, rhodium leaching is small and recycling is possible.

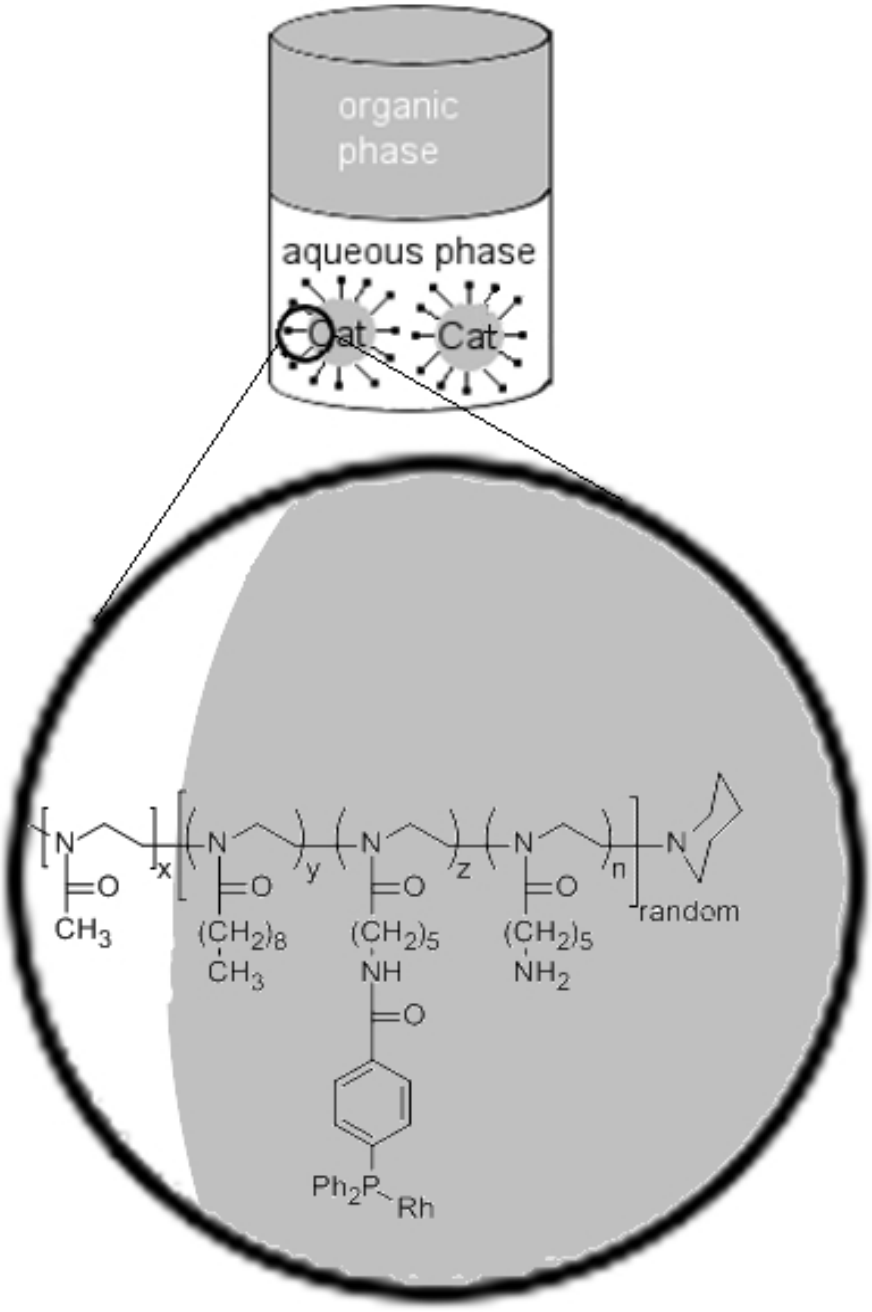


Fig. 5 Example of an amphiphilic block copolymer forming micelles.

Proteins as support

A promising approach was invented by the group of Marchetti. Instead of using rather poorly defined man-made polymers, they used proteins as support. Although the protein structure itself is well-defined, the protein–rhodium complexes synthesised are

undefined as the proteins contain numerous donor atoms (for example amines and carboxylates).

They investigated the proteins human serum albumin,[75] as well as papain and egg albumin[76] in combination with Rh(I) in the hydroformylation of 1-octene using an excess of rhodium (up to 120 eq.). The latter complexes are not stable under hydroformylation conditions but precipitate and release rhodium to the organic phase, while the rhodium–human serum albumin complex can be reused several times. Unfortunately rhodium leaching was quite significant in the range of 8 ppm.[75]

Supported aqueous phase catalysis

One approach to increase the surface area of the interphase between aqueous and organic phase and therefore the reaction rate is Supported Aqueous Phase Catalysis (SAPC), which was developed in 1989.[77]

In SAPC a thin aqueous layer containing a water soluble catalyst complex is adsorbed onto hydrophilic solids. The solid should have a large surface area and should be inert to thermal and mechanical stress and to all chemicals present in the reaction mixture. For hydroformylation most supports are based on silica, but other supports like cation exchange resin,[78,79] glass[80] and apatitic tricalcium phosphates[81] have been investigated as well. In most cases, the catalyst is based on rhodium but some examples for cobalt-[79,80,82] and platinum-based catalysts[82] exist. The reaction itself takes place at the interphase of the organic phase and the water layer absorbed to the particles (see Fig. 6), therefore the reaction rate is almost unaffected by the water solubility of the substrate. At the end of the reaction the organic phase containing the products can be separated from the support and its catalyst-containing water layer by filtration and the latter can be reused. SAPC not only depends on the nature of the metal-complex, it is also highly dependent on the inherent properties of the support such as particle size and surface area. Additionally the hydration of the support is important.[83] The influence of the hydration on the hydroformylation of long chain 1-alkenes was shown very clearly by Kalck et al. They varied the water content near the optimal hydration degree of their system in steps of 0.1%. Changing their optimal hydration by 0.2% in either way reduces the conversion by about 13%, while the selectivity remained unchanged.[84]

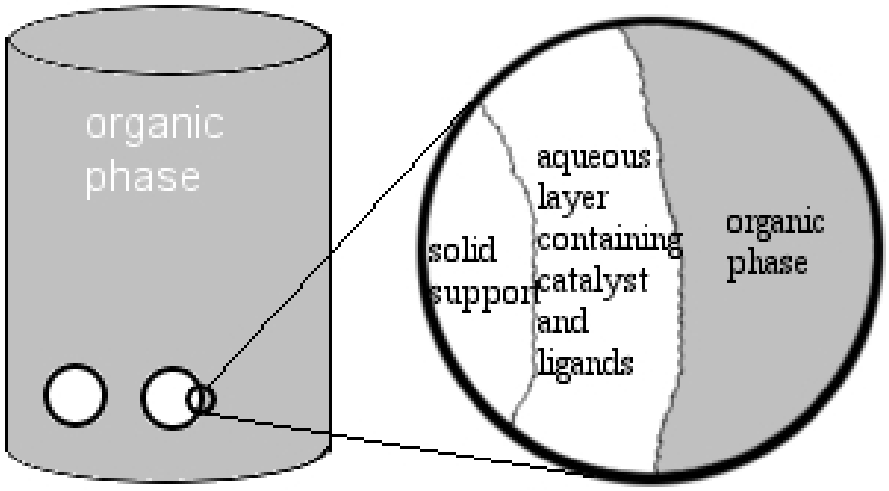


Fig. 6 SAPC concept.

In SAPC, the activity is low (TOF<100 h^{-1}) in most cases for both rhodium[82,85] and cobalt based[79,80] catalysts. The group of Hanson showed that in SAPC the reaction takes place at the phase-boundary: conversion of 1-octene, 1-decene and 1-dodecene occurred with similar TOFs of ≈45 h^{-1} affording the corresponding aldehydes with a *l*/*b* ratio of ≈2.3.[82] Ligand tuning is a powerful tool to influence the selectivity in SAPC. Using a xantphos-derivative as ligand results in a linear to branched selectivity of >30[86] which is almost as high as in biphasic catalysis.[87] Using TPPTS as ligand under almost the same conditions gives a much lower selectivity (*l*/*b* = 3).[85] There are some examples with high TOFs (>100 h^{-1}),[82,85,86,88,89] two of which use TPPTS as ligand and silica as support, the high TOFs are most likely due to optimised conditions.[82,85]

In one of the other successful approaches a mixture of polyethyleneglycol and water instead of pure water is adsorbed onto the particle.[88] Both the TOF (≈1000 h^{-1}) as well as the selectivity (*l*/*b* ≈ 16) are remarkably high in this case but it should be mentioned that the TOF was measured by the initial gas uptake of the reaction. This approach should give the same values but is prone to errors and TOF calculated this way tend to be higher than conventionally measured TOFs. No products were formed when reusing the organic phase, indicating that no rhodium had leached.

Another approach was published by Zhu and co-workers.[86] Their system is a combination of SAPC and heterogeneous catalysis as there are rhodium particles on the

support in addition to the adsorbed water-soluble catalyst (Fig. 7). Both TOF (>300 h^{-1}) and selectivity (*l/b* ≈ 6) are good and rhodium leaching is low (<0.1 ppm). The catalyst could be reused three times without significant erosion of performance.

Yuan et al. reported high TOF (up to >500 h^{-1}) and good selectivities (up to *l/b* ≈ 6) when using fumed silica as support.[89]

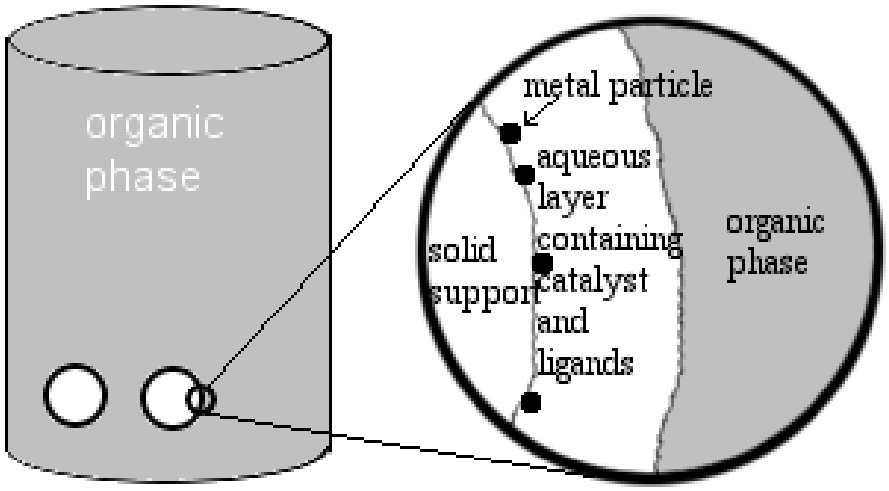


Fig. 7 Combining SAPC with heterogeneous catalysis.

Smart systems

In smart systems the aqueous biphasic system can be triggered to turn into a monophasic system during the reaction. Upon completion of the reaction it can be turned back into a biphasic system to allow for the separation of catalyst and products. There are small variations of this concept (Fig. 8) and a variety of different triggers exist.

Smart systems for biphasic hydroformylation were already reported in 1973. Using cobalt, rhodium or iridium complexes with basic aminophosphine ligands such as $P(CH_2CH_2CH_2NEt_2)_3$, the catalyst could be extracted either by aqueous solution of carbon dioxide under pressure or by using ‘‘standard’’ acids/ bases like sulphuric acid and sodium hydroxide. Reusing the catalyst afforded very similar conversions and selectivities.[90,91]

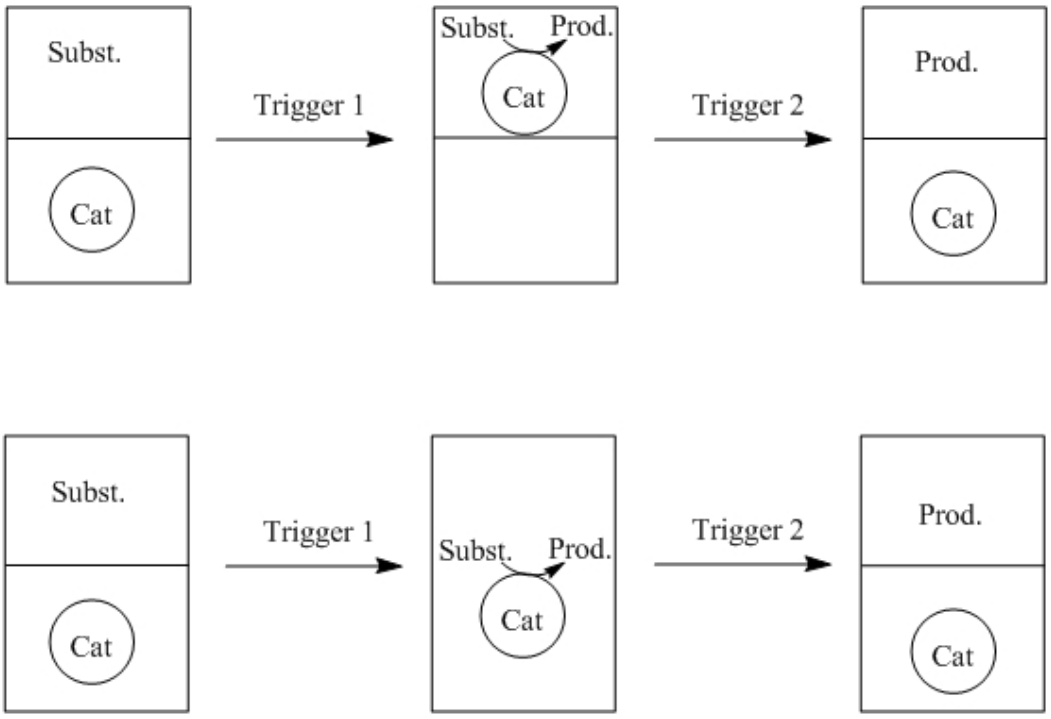
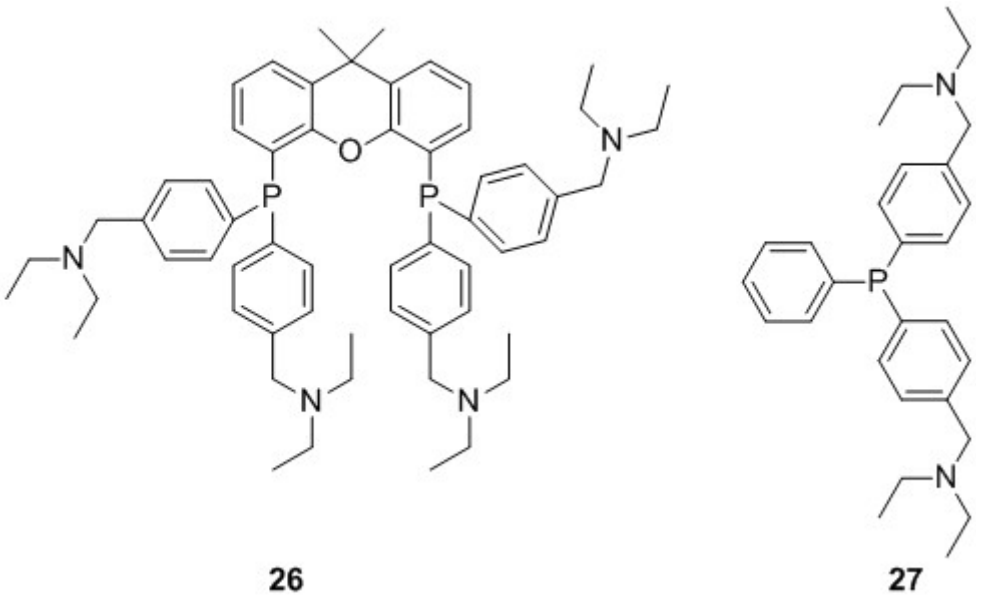


Fig. 8 Principle of smart systems.

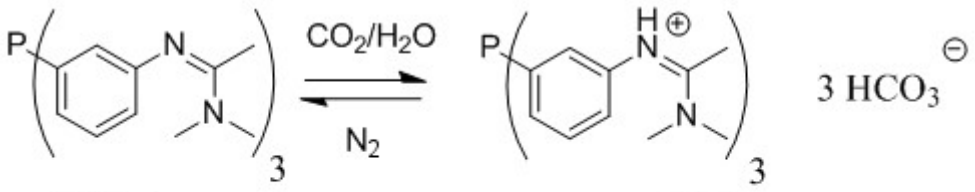
The group of van Leeuwen and Kamer reported several rhodium complexes with pH dependent solubility.[92–94] In neutral solutions these complexes are soluble in the organic layer while upon acidification they become water soluble. Acidic extraction into the water layer and/or backwash upon neutralisation into a new organic layer proved to be effective for two ligands (**26**, **27**, Fig. 9) of a series of mono- and bidentate ligands. In the case of these ligands 97–98% of rhodium could be transferred to a new organic layer. When the xantphos based ligand **26** was tested in the rhodium catalysed hydroformylation of 1-hexene, 1-octene and 1-dodecene a remarkably high selectivity (l/b = 48–52) and good TOF (≈200 h^{-1}) were obtained. They investigated the hydroformylation of 1-octene at a higher temperature (100 °C instead of 80 °C) and obtained a slightly lower selectivity (l/b ≈ 45) and slightly more isomerisation but a much higher activity (≈900 h^{-1}). The results are expected to be similar for different 1-alkenes. Unfortunately, for both ligands the activity in a new catalytic run dropped to ≈86%.[93,94] It was shown that this drop in activity was due to catalyst decomposition during the acidic extraction.[94] In addition to the significant rhodium leaching and catalyst decomposition, salt is formed in the acid–base reaction as undesired byproduct.



26 **27**

Fig. 9 pH dependent smart ligands.

The group of Cole-Hamilton used the pH change caused by bubbling carbon dioxide through the solution to obtain a water soluble ligand, e.g. **29**. Upon bubbling nitrogen through the solution carbon dioxide is released and the ligand becomes soluble in organic solvents again (Scheme 1).[95,96]

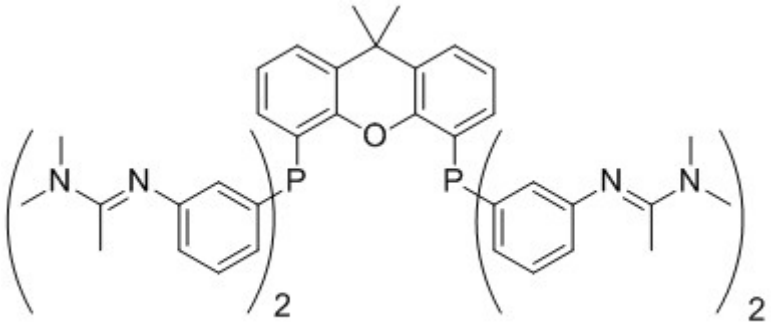


soluble in organic solvents **28** water soluble **29**

Scheme 1 A smart system using gases to change catalyst solubility.

The system has a remarkably high initial TOF (>10 000 h^{-1} measured by the initial gas uptake) and the linear to branched selectivity is highly dependent on the type of ligand used: when using ligand **28** the *l/b* ratio is ≈3. When using a xantphos based bidentate ligand (**30**, Fig. 10) the *l/b* ratio is ≈20 for the hydroformylation of 1-octene. Rhodium leaching of 1.9 ppm was measured for **28**. Although the level of rhodium leaching when

using **30** was not determined it was visually deemed to be very high. When **28** was reused the activity dropped to 92%. Upon reusing it for three times an activity of 91% was observed. The *l/b* ratio remains almost unchanged during these recycle experiments. In contrast to the approach of van Leeuwen, the use of gases to tune catalyst solubility does not suffer from the build-up of byproducts and the catalyst does not have to withstand harsh recycling conditions.



30

Fig. 10 Xantphos based smart ligand.

Although not strictly aqueous–organic biphasic reactions, some alternative approaches are also worth mentioning. In the Union Carbide process for the hydroformylation of higher alkenes, the reaction is carried out as a homogenous process using the TPPMS ligand in N-methylpyrrolidone (NMP) as solvent. The addition of water after the reaction brings about a phase separation, resulting in complete transfer of the catalyst to the aqueous phase, which is then recovered.[97] In another approach a homogeneous mixture of THF and water (7 : 3 (v : v)) was used as homogeneous reaction medium yielding a good TOF (≈400 h^{-1}) and moderate selectivity (*l/b* = 2.4).[98] At the end of the reaction carbon dioxide pressure is applied to induce phase separation. Different ligands were tested and it was shown that the aqueous layer containing the catalyst could be reused when using TPPMS as ligand for three times without changing the activity. Furthermore the rhodium content in the organic layer was below the detection limit of 1 ppm.

Using temperature as trigger to change the solubility was investigated by various groups.

The group of Monflier[99] used a thermo controlled cyclodextrin based system to obtain really high TOF (6000–7000 h^{-1}). At low temperatures an amphiphilic phosphine ligand is totally or partly located within the hydrophobic cavity of a modified cyclodextrin. Upon increasing the temperature the proportion of ligand within the cyclodextrin is reduced, consequently raising the ligand concentration at the aqueous/organic interphase. When the temperature is reduced this process is reversed, which gives rise to an easy phase separation. They investigated this concept for various 1-alkenes and the TOFs obtained were independent of the alkene chain length. Unfortunately the linear selectivity was rather poor (*l*/*b* = 1.7–2.0). They investigated the possibility to recycle the aqueous layer and found it could be recycled at least twice with little effect on conversion or selectivity. Unfortunately no data on the amount of rhodium or ligand leaching were published.

A further class of thermo regulated ligands used for the hydroformylation of long chain 1-alkenes was intensively investigated by the group of Jin.[100–106] The ligands are based on triphenylphosphines modified with a polyethylene chain **31**–**33** (Fig. 11) and have an inverse temperature dependent solubility in water. Good to excellent activities were found for the biphasic hydroformylation of various 1-alkenes (300–4000 h^{-1}), while the *l*/*b* ratio tends to be poor (*l*/*b* = 0.4–2.1). It was shown that the *l*/*b* ratio is highly temperature dependent: increasing the temperature from 80 to 120 °C caused the ratio to drop from 2.06 to 0.64. Though the selectivity is unsatisfying, the reusability of this type of ligands is very good as they could reuse one of their catalysts twenty times without significantly affecting its activity.[106]

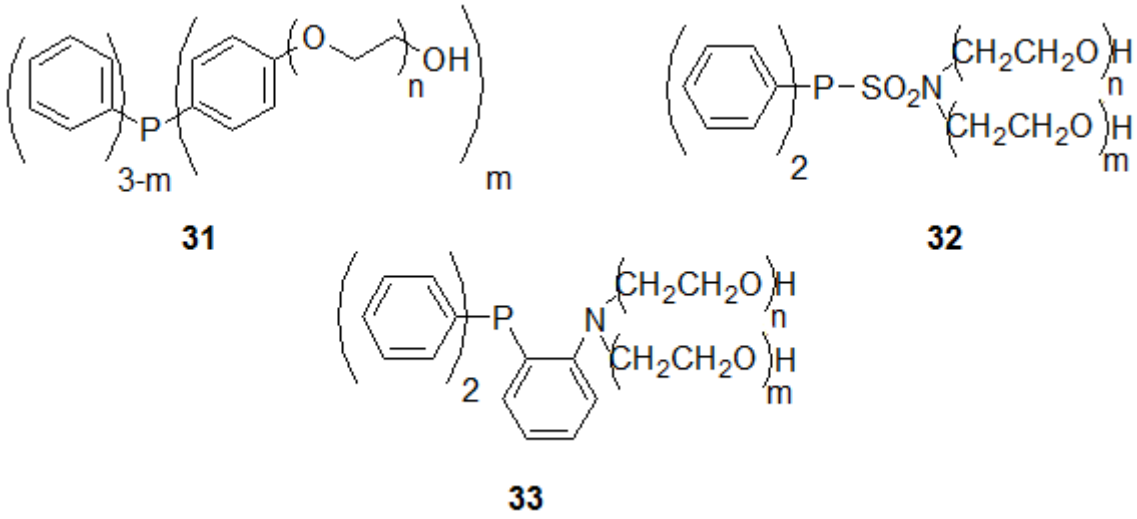


Fig. 11 Temperature-dependent smart ligands.

Overview on the different aqueous-organic biphasic hydroformylation concepts

It should be obvious that it is basically impossible to compare the methods presented in a single table. Anyway, one approach of every method introduced in this chapter is summarised in the following table to give a quick overview on activity, selectivity, leaching and reusability. The method with the highest TOF is listed and methods with TOF <500 were not taken into account.

Table 1 Summary of different aqueous-organic biphasic hydroformylation approaches

entry	Approach	substrate	TOF [h^{-1}]	l/b	Rh leaching	reusability
1[f]	Additives	1-dodecene	7472	15.7	0.5 ppm	Yes[a]
2[g]	Ligand varation	1-dodecene	673	2.4	0.01-0.1 ppm	Yes[b]
3[h]	Polymer supported	1-octene	3700	3	0.4 ppm	Yes[c]
4[i]	SAPC	1-octene	1000*	16	n.d.[d]	n.d.[d]
5[j]	“smart” system	1-octene	>10000 *	20	1.9 ppm	Yes[e]

a: An activity loss of 10 % over eight consecutive runs was observed. **b**: An activity loss of 10 % over four consecutive runs was observed. **c**: activity increase upon reuse, further reuse results in an activity drop. **d**: not determined. **e**: An activity loss of 10 % over three consecutive runs was observed. *: determined using the initial gas uptake reaction. **f:** Data taken from[18] **g**: Data taken from[52] **h**: Data taken from[60,70,73,74] **I**: Data taken from[88] **j**: Data taken from[95]

Outlook

Transferring the effectiveness and elegance of the RCH/RP process to the biphasic hydroformylation of higher alkenes has been a long-standing goal in transition-metal catalysis research. Despite intensive and creative research leading to a variety of approaches to overcome the mass-transfer limitations, most of these have drawbacks inhibiting their practical application, like reduced linear selectivity or poor catalyst recovery/recyclability. Nonetheless, some promising approaches like those based on the use of certain additives or smart ligands have been established, which provide considerable rate enhancements while preserving some of the inherent strength of the original process. To date, none of the concepts outlined above has been applied to the industrial hydroformylation of higher olefins. There is still a need for an approach that meets all of the strict requirements of a technical two-phase process, such as complete catalyst retention, high activity and stability, high aldehyde selectivity, simple phase separation, and low ligand costs in order to be economically competitive with the currently used processes.

1.3 Outlook on this thesis

Nature uses enzymes which are the most efficient and selective catalysts known. If their properties could be combined with the reactivity of transition metals it would be possible to form highly active and selective hybrid catalysts. Because of their chirality the resulting catalysts could be used for asymmetric reactions as well.

Regarding the design of hybrid catalysts the following two requirements were considered to be essential:

- In order to avoid catalysis by free (leached) metal the transition metal has to be firmly bound to the enzyme.
- In order to take full advantage of the tertiary structure of the enzyme the location of the transition metal should be in close proximity to the active site of the enzyme (e.g. a tunnel).

Covalent modification of enzymes allows the introduction of a metal binding site in

close proximity to the active site of a protein. This concept which is schematically depicted in fig. 12 was used in this thesis.

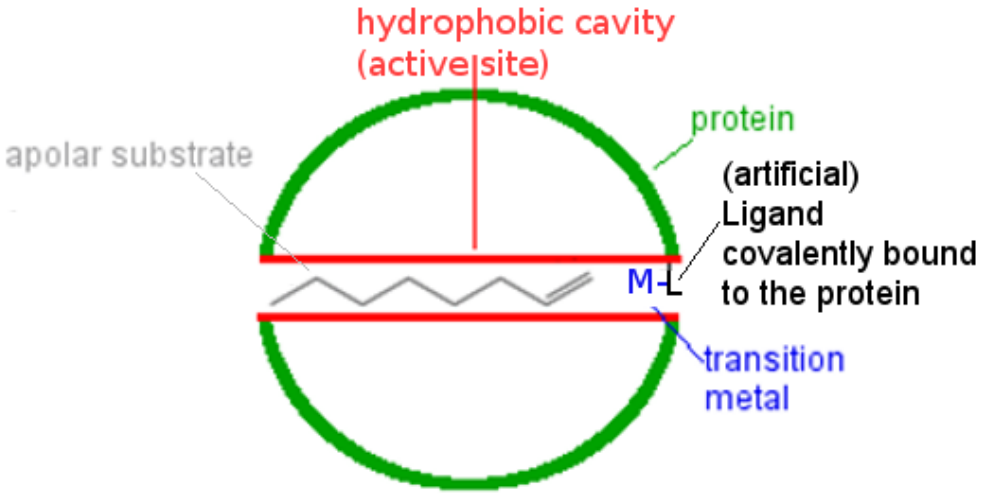


Fig. 12 Schematic representation of an ideal artificial metalloenzyme obtained by covalent modification.

The figure shows a protein with an active site (a hydrophobic tunnel) which is capable to bind the desired apolar substrate (as an example 1-octene is depicted). At a well defined site of the protein a ligand with a high affinity to a transition metal is covalently attached. If the binding capacity of the ligand is higher than other randomly distributed binding sites of the enzymes (e.g. amino acids) a transition metal is coordinated to this ligand. The modification is near the active site, so the introduced metal is in close proximity to the substrate. Due to high substrate concentrations within the tunnel next to the transition metal rate enhancements should be observed. In addition the protein environment should induce (enantio)selectivity.

In the design, synthesis and application of artificial metalloenzymes the basic concept outlined in fig. 12 is pursued throughout the whole thesis. The main focus of the research is on the two mutants SCP-2L A100C and SCP-2L V83C which were found to be suitable protein scaffolds.[107]

In the long run using artificially modified proteins has a lot of potential as it might use the high activities and selectivities of enzymes for reactions not catalysed by enzymes.

There are already many different reactions catalysed by artificial metalloenzymes – the ones relevant for this thesis will be discussed in the corresponding chapters. Additionally, a short general introduction about artificial metalloenzymes is given in chapter 3. The interested reader is refered to the many reviews[108–111] on artificial metalloenzymes for additional information.

Chapter 2:
The crystal structure of the SCP-2L A100C mutant was obtained and discussed. A special emphasis on its hydrophobic tunnel is taken and it is shown that the mutation has a small effect on this tunnel.
Additionally, the binding capability of the two SCP-2L mutants and their covalently modified counterparts towards different metal salts is investigated. Two different *N*-ligands (phenanthroline or dipicolylamine based) and three different phosphine ligands (all based on triphenylphosphine) were used for this study. Additionally, the rhodium binding capabilities of two protein scaffolds which were tested as a reference in the latter investigated hydroformylation reaction were determined as well.

Chapter 3:
This chapter focusses on the application of Rh-enzymes in the aqueous-organic biphasic hydroformylation of long chain 1-alkenes. Six different SCP-2L based phosphine modified Rh-enzymes were tested. In order to evaluate the high activities and selectivities observed the mode of action was further investigated.

Chapter 4:
This chapter describes the application of *N*-ligand modified SCP-2L A100C in the asymmetric Diels-Alder reaction. Cu(II)-, Ni(II)- and Co(II)-enzymes which were determined as promising artificial metalloenzymes in chapter 2 are tested. In these reactions the chiral environment of the protein is utilized to induce enantioselectivity in catalytic reactions.

Chapter 5:

Finally, the fifth chapter will give a summary, discussion and final conclusion on the approach and methodology applied in this thesis.

1.4 References

[1] S. K. Sharma, P. A. Parikh, R. V. Jasra, *J. Mol. Catal. A: Chem.* **2010**, *316*, 153–162.

[2] B. Cornils, W. A. Herrmann, *Aqueous-Phase Organometallic Catalysis*, Wiley-VCH, Weinheim, **2004**.

[3] C. W. Kohlpaintner, R. W. Fischer, B. Cornils, *Appl. Catal., A* **2001**, *221*, 219–225.

[4] E. Wiebus, B. Cornils, in *Catalyst Separation, Recovery and Recycling* (Eds.: D.J. Cole-Hamilton, R.P. Tooze), Springer, Dordrecht, **2006**, p. 105.

[5] P. W. N. M. Van Leeuwen, *Homogeneous Catalysis*, Kluwer Academic Publishers, Dordrecht, **2004**.

[6] H. Bahrmann, S. Bogdanovic, P. W. N. M. v. Leeuwen, in *Aqueous Phase Organometallic Catalysis* (Eds.: B. Cornils, W.A. Hermanns), Wiley-VCH, Weinheim, **2004**.

[7] P. Purwanto, H. Delmas, *Catal. Today* **1995**, *24*, 135–140.

[8] H. Ding, B. E. Hanson, T. Bartik, B. Bartik, *Organometallics* **1994**, *13*, 3761–3763.

[9] F. Monteil, R. Queau, P. Kalck, *J. Organomet. Chem.* **1994**, *480*, 177–184.

[10] M. J. Hayling, B. A. Murrer, (Johnson Matthey), **1982**, GB2085874A.

[11] H. Chen, Y. Li, J. Chen, P. Cheng, Y. He, X. Li, *J. Mol. Catal. A: Chem.* **1999**, *149*, 1–6.

[12] Y. Zhang, Z.-S. Mao, J. Chen, *Catal. Today* **2002**, *74*, 23–35.

[13] A. Riisager, B. E. Hanson, *J. Mol. Catal. A: Chem.* **2002**, *189*, 195–202.

[14] L. Wang, H. Chen, Y. He, Y. Li, M. Li, X. Li, *Appl. Catal., A* **2003**, *242*, 85–88.

[15] H. Chen, Y. Li, J. Chen, P. Cheng, X. Li, *Catal. Today* **2002**, *74*, 131–135.

[16] M. Li, H. Fu, M. Yang, H. Zheng, Y. He, H. Chen, X. Li, *J. Mol. Catal. A: Chem.* **2005**, *235*, 130–136.

[17] H. Fu, M. Li, H. Mao, Q. Lin, M. Yuan, X. Li, H. Chen, *Catal. Commun.* **2008**, *9*, 1539–1544.

[18] H. Fu, M. Li, H. Chen, X. Li, *J. Mol. Catal. A: Chem.* **2006**, *259*, 156–160.

[19] C. Yang, X. Bi, Z.-S. Mao, *J. Mol. Catal. A: Chem.* **2002**, *187*, 35–46.

[20] S. L. Desset, D. J. Cole-Hamilton, D. F. Foster, *Chem. Commun.* **2007**, 1933–1935.

[21] S. L. Desset, S. W. Reader, D. J. Cole-Hamilton, *Green Chem.* **2009**, *11*, 630–637.

[22] H. Bricout, F. Hapiot, A. Ponchel, S. Tilloy, E. Monflier, *Curr. Org. Chem.* **2010**, *14*, 1296–1307.

[23] F. Hapiot, A. Ponchel, S. Tilloy, E. Monflier, *C. R. Chim.* **2011**, *14*, 149–166.

[24] J. B. Simoes, D. L. da Silva, A. de Fatima, S. A. Fernandes, *Curr. Org. Chem.* **2012**, *16*, 949–971.
[25] J. Anderson, E. Campi, W. R. Jackson, *Catal. Lett.* **1991**, *9*, 55–58.
[26] E. Monflier, G. Fremy, Y. Castanet, A. Mortreux, *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 2269–2271.
[27] F. Hapiot, L. Leclercq, N. Azaroual, S. Fourmentin, S. Tilloy, E. Monflier, *Curr. Org. Synth.* **2008**, *5*, 162–172.
[28] L. Leclercq, H. Bricout, S. Tilloy, E. Monflier, *J. Colloid Interface Sci.* **2007**, *307*, 481–487.
[29] M. Ferreira, F.-X. Legrand, C. Machut, H. Bricout, S. Tilloy, E. Monflier, *Dalton Trans.* **2012**, *41*, 8643–8647.
[30] E. Monflier, H. Bricout, F. Hapiot, S. Tilloy, A. Aghmiz, A. M. Masdeu-Bultó, *Adv. Synth. Catal.* **2004**, *346*, 425–431.
[31] S. Tilloy, G. Crowyn, E. Monflier, P. W. N. M. van Leeuwen, J. N. H. Reek, *New J. Chem.* **2006**, *30*, 377–383.
[32] L. Leclercq, M. Sauthier, Y. Castanet, A. Mortreux, H. Bricout, E. Monflier, *Adv. Synth. Catal.* **2005**, *347*, 55–59.
[33] N. Badi, P. Guégan, F.-X. Legrand, L. Leclercq, S. Tilloy, E. Monflier, *J. Mol. Catal. A: Chem.* **2010**, *318*, 8–14.
[34] P. Blach, D. Landy, S. Fourmentin, G. Surpateanu, H. Bricout, A. Ponchel, F. Hapiot, E. Monflier, *Adv. Synth. Catal.* **2005**, *347*, 1301–1307.
[35] D. Kirschner, T. Green, F. Hapiot, S. Tilloy, L. Leclercq, H. Bricout, E. Monflier, *Adv. Synth. Catal.* **2006**, *348*, 379–386.
[36] D. Kirschner, M. Jaramillo, T. Green, F. Hapiot, L. Leclercq, H. Bricout, E. Monflier, *J. Mol. Catal. A: Chem.* **2008**, *286*, 11–20.
[37] L. Leclercq, F. Hapiot, S. Tilloy, K. Ramkisoensing, J. N. H. Reek, P. W. N. M. van Leeuwen, E. Monflier, *Organometallics* **2005**, *24*, 2070–2075.
[38] F.-X. Legrand, F. Hapiot, S. Tilloy, A. Guerriero, M. Peruzzini, L. Gonsalvi, E. Monflier, *Appl. Catal., A* **2009**, *362*, 62–66.
[39] M. Dessoudeix, M. Urrutigoïty, P. Kalck, *Eur. J. Inorg. Chem.* **2001**, *2001*, 1797–1800.
[40] J. Potier, S. Menuel, D. Fournier, S. Fourmentin, P. Woisel, E. Monflier, F. Hapiot, *ACS Catal.* **2012**, *2*, 1417–1420.
[41] K. Kunna, C. Müller, J. Loos, D. Vogt, *Angew. Chem., Int. Ed.* **2006**, *45*, 7289–7292.
[42] N. Kania, B. Léger, S. Fourmentin, E. Monflier, A. Ponchel, *Chem.--Eur. J.* **2010**, *16*, 6138–6141.
[43] A. A. Dabbawala, H. C. Bajaj, H. Bricout, E. Monflier, *Appl. Catal., A* **2012**, *413–414*, 273–279.
[44] J. Boulanger, A. Ponchel, H. Bricout, F. Hapiot, E. Monflier, *Eur. J. Lipid Sci. Technol.* **2012**, *114*, 1439–1446.
[45] P. J. Baricelli, E. Lujano, M. Modroño, A. C. Marrero, Y. M. García, A. Fuentes, R. A. Sánchez-Delgado, *J. Organomet. Chem.* **2004**, *689*, 3782–3792.
[46] R. V. Chaudhari, B. M. Bhanage, R. M. Deshpande, H. Delmas, *Nature* **1995**, *373*, 501–503.
[47] B. Fell, G. Papadogianakis, *J. Mol. Catal.* **1991**, *66*, 143–154.

[48] S. Bischoff, M. Kant, *Ind. Eng. Chem. Res.* **2000**, *39*, 4908–4913.
[49] E. Paetzold, G. Oehme, C. Fischer, M. Frank, *J. Mol. Catal. A: Chem.* **2003**, *200*, 95–103.
[50] Q. Peng, X. Liao, Y. Yuan, *Catal. Commun.* **2004**, *5*, 447–451.
[51] B. E. Hanson, H. Ding, C. W. Kohlpaintner, *Catal. Today* **1998**, *42*, 421–429.
[52] H. Fu, M. Li, J. Chen, R. Zhang, W. Jiang, M. Yuan, H. Chen, X. Li, *J. Mol. Catal. A: Chem.* **2008**, *292*, 21–27.
[53] H. Ding, J. Kang, B. E. Hanson, C. W. Kohlpaintner, *J. Mol. Catal. A: Chem.* **1997**, *124*, 21–28.
[54] M. S. Goedheijt, B. E. Hanson, J. N. H. Reek, P. C. J. Kamer, P. W. N. M. van Leeuwen, *J. Am. Chem. Soc.* **2000**, *122*, 1650–1657.
[55] M. Ferreira, H. Bricout, N. Azaroual, D. Landy, S. Tilloy, F. Hapiot, E. Monflier, *Adv. Synth. Catal.* **2012**, *354*, 1337–1346.
[56] M. T. Reetz, S. R. Waldvogel, *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 865–867.
[57] D. Armspach, D. Matt, *Chem. Commun.* **1999**, 1073–1074.
[58] S. Shimizu, S. Shirakawa, Y. Sasaki, C. Hirai, *Angew. Chem., Int. Ed.* **2000**, *39*, 1256–1259.
[59] L. Monnereau, D. Sémeril, D. Matt, L. Toupet, *Adv. Synth. Catal.* **2009**, *351*, 1629–1636.
[60] O. Nuyken, P. Persigehl, R. Weberskirch, *Macromol. Symp.* **2002**, *177*, 163–174.
[61] Y. Uozumi, M. Nakazono, *Adv. Synth. Catal.* **2002**, *344*, 274–277.
[62] E. A. Karakhanov, Y. S. Kardasheva, A. L. Maximov, E. A. Runova, T. A. Buchneva, M. A. Gaevskiy, A. Y. Zhuchkova, T. Y. Filippova, *Polym. Adv. Technol.* **2001**, *12*, 161–168.
[63] T. Malmström, C. Andersson, J. Hjortkjaer, *J. Mol. Catal. A: Chem.* **1999**, *139*, 139–147.
[64] E. A. Karakhanov, Y. S. Kardasheva, A. L. Maksimov, V. V. Predeina, E. A. Runova, A. M. Utukin, *J. Mol. Catal. A: Chem.* **1996**, *107*, 235–240.
[65] M. Ahlmann, O. Walter, M. Frank, W. Habicht, *J. Mol. Catal. A: Chem.* **2006**, *249*, 80–92.
[66] E. A. Karakhanov, Y. S. Kardasheva, E. A. Runova, V. A. Semernina, *J. Mol. Catal. A: Chem.* **1999**, *142*, 339–347.
[67] A. N. Ajjou, H. Alper, *J. Am. Chem. Soc.* **1998**, *120*, 1466–1468.
[68] L. Xiaozhong, L. Hongmei, K. Fanzhi, *J. Organomet. Chem.* **2002**, *664*, 1–4.
[69] B. C. E. Makhubela, A. Jardine, G. S. Smith, *Green Chem.* **2012**, *14*, 338–347.
[70] M. T. Zarka, M. Bortenschlager, K. Wurst, O. Nuyken, R. Weberskirch, *Organometallics* **2004**, *23*, 4817–4820.
[71] J. Chen, H. Alper, *J. Am. Chem. Soc.* **1997**, *119*, 893–895.
[72] T. Borrmann, H. W. Roesky, U. Ritter, *J. Mol. Catal. A: Chem.* **2000**, *153*, 31–48.
[73] B. Gall, M. Bortenschlager, O. Nuyken, R. Weberskirch, *Macromol. Chem. Phys.* **2008**, *209*, 1152–1159.
[74] M. Bortenschlager, N. Schöllhorn, A. Wittmann, R. Weberskirch, *Chem.--Eur. J.* **2007**, *13*, 520–528.
[75] M. Marchetti, G. Mangano, S. Paganelli, C. Botteghi, *Tetrahedron Lett.* **2000**, *41*, 3717–3720.
[76] C. Bertucci, C. Botteghi, D. Giunta, M. Marchetti, S. Paganelli, *Adv. Synth.*

Catal. **2002**, *344*, 556–562.
[77] J. P. Arhancet, M. E. Davis, J. S. Merola, B. E. Hanson, *Nature (London, U. K.)* **1989**, *339*, 454–455.
[78] R. T. Smith, R. K. Ungar, L. J. Sanderson, M. C. Baird, *Organometallics* **1983**, *2*, 1138–1144.
[79] M. K. Markiewicz, M. C. Baird, *Inorg. Chim. Acta* **1986**, *113*, 95–99.
[80] I. Guo, B. E. Hanson, I. Tóth, M. E. Davis, *J. Organomet. Chem.* **1991**, *403*, 221–227.
[81] M. Dessoudeix, U. J. Jáuregui-Haza, M. Heughebaert, A. M. Wilhelm, H. Delmas, A. Lebugle, P. Kalck, *Adv. Synth. Catal.* **2002**, *344*, 406–412.
[82] I. Tóth, I. Guo, B. E. Hanson, *J. Mol. Catal. A: Chem.* **1997**, *116*, 217–229.
[83] G. Frémy, E. Monflier, J.-F. Carpentier, Y. Castanet, A. Mortreux, *J. Catal.* **1996**, *162*, 339–348.
[84] P. Kalck, L. Miquel, M. Dessoudeix, *Catal. Today* **1998**, *42*, 431–440.
[85] A. J. Sandee, V. F. Slagt, J. N. H. Reek, P. C. J. Kamer, P. W. N. M. van Leeuwen, *Chem. Commun.* **1999**, 1633–1634.
[86] H. Zhu, Y. Ding, H. Yin, L. Yan, J. Xiong, Y. Lu, H. Luo, L. Lin, *Appl. Catal., A* **2003**, *245*, 111–117.
[87] M. Schreuder Goedheijt, P. C. . Kamer, P. W. N. . van Leeuwen, *J. Mol. Catal. A: Chem.* **1998**, *134*, 243–249.
[88] M. J. Naughton, R. S. Drago, *J. Catal.* **1995**, *155*, 383–389.
[89] Z. Li, Q. Peng, Y. Yuan, *Appl. Catal., A* **2003**, *239*, 79–86.
[90] G. Gregorio, A. Andretta, (Montecatini Edison Spa), **1973**, DE2313102 (A1).
[91] A. Andreetta, G. Barberis, G. Gregorio, *Chim. Ind. (Milan, Italy)* **1978**, 887–891.
[92] A. Buhling, J. W. Elgersma, S. Nkrumah, P. C. J. Kamer, P. W. N. M. van Leeuwen, *J. Chem. Soc., Dalton Trans.* **1996**, 2143–2154.
[93] A. Buhling, P. C. J. Kamer, P. W. N. M. van Leeuwen, J. W. Elgersma, *J. Mol. Catal. A: Chem.* **1997**, *116*, 297–308.
[94] A. Buhling, P. C. J. Kamer, P. W. N. M. van Leeuwen, J. W. Elgersma, K. Goubitz, J. Fraanje, *Organometallics* **1997**, *16*, 3027–3037.
[95] S. L. Desset, D. J. Cole-Hamilton, *Angew. Chem., Int. Ed.* **2009**, *48*, 1472–1474.
[96] P. Kalck, *Polyhedron* **1988**, *7*, 2441–2450.
[97] A. G. Abatjoglou, D. R. Bryant, (Union Carbide Corporation), **1988**, US 4.731.486.
[98] J. P. Hallett, J. W. Ford, R. S. Jones, P. Pollet, C. A. Thomas, C. L. Liotta, C. A. Eckert, *Ind. Eng. Chem. Res.* **2008**, *47*, 2585–2589.
[99] N. Six, A. Guerriero, D. Landy, M. Peruzzini, L. Gonsalvi, F. Hapiot, E. Monflier, *Catal. Sci. Technol.* **2011**, *1*, 1347–1353.
[100] Z. Jin, X. Zheng, B. Fell, *J. Mol. Catal. A: Chem.* **1997**, *116*, 55–58.
[101] X. Zheng, J. Jiang, X. Liu, Z. Jin, *Catal. Today* **1998**, *44*, 175–182.
[102] J. Jiang, Y. Wang, C. Liu, F. Han, Z. Jin, *J. Mol. Catal. A: Chem.* **1999**, *147*, 131–136.
[103] Y. Wang, J. Jiang, R. Zhang, X. Liu, Z. Jin, *J. Mol. Catal. A: Chem.* **2000**, *157*, 111–115.
[104] Y. Wang, J. Jiang, Q. Miao, X. Wu, Z. Jin, *Catal. Today* **2002**, *74*, 85–90.
[105] J. Jiang, Y. Wang, C. Liu, Q. Xiao, Z. Jin, *J. Mol. Catal. A: Chem.* **2001**, *171*, 85–

89.
[106] C. Liu, J. Jiang, Y. Wang, F. Cheng, Z. Jin, *J. Mol. Catal. A: Chem.* **2003**, *198*, 23–27.
[107] P. Deuss J., Artificial Metalloenzymes; Modified Proteins as Tuneable Transition Metal Catalysts, University of St Andrews, **2011**.
[108] J. C. Lewis, *ChemInform* **2014**, *45*, no–no.
[109] J. Steinreiber, T. R. Ward, *Coordination Chemistry Reviews* **2008**, *252*, 751–766.
[110] F. Rosati, G. Roelfes, *ChemCatChem* **2010**, *2*, 916–927.
[111] P. J. Deuss, R. den Heeten, W. Laan, P. C. J. Kamer, *Chem.--Eur. J.* **2011**, *17*, 4680–4698.

Chapter 2: Crystal structures and metal complexes

2.1 Abstract

In this chapter the crystal structure of SCP-2L A100C, which is a mutant of the sterol carrier protein-2 like domain of the human peroxisomal multifunctional enzyme type 2 is described and compared to its wild type. The mutant SCP-2L V83C is briefly discussed as well. A special emphasis is placed on the hydrophobic tunnel of these proteins.

The two SCP-2L mutants, lysozyme and a mutant of the photo-active yellow protein (PYP R52G) were modified with *P*-ligands based on triphenylphosphine. Additionally both SCP-2L mutants were modified with *N*-ligands based on either 1,10-phenanthroline or di-(2-picolyl)amine. The synthesised artificial enzymes as well as the unmodified proteins were treated with selected transition metals (twelve different metals were tested) and analysed by ESI-MS. The metal loading of most of these artificial metalloenzymes was also determined by ICP-MS.

Some of the phosphine modified proteins were identified as promising scaffolds for Ir(I)- and Rh(I)-enzymes as the artificial metalloenzymes could be detected by ESI-MS and a high metal loading was determined by ICP-MS.

The *N*-ligand modified proteins of SCP-2L A100C were also analysed by ICP-MS. The ICP-MS and ESI-MS data suggests that di-(2-picolyl)amine) modified SCP-2L is a promising scaffold for Pd(II)-enzymes and 1,10-phenanthroline modified SCP-2L is a promising scaffold for Pd(II)-, Cu(II)- Ni(II) and Co(II)-enzymes.

2.2 Introduction

SCP-2L

Multi functional enzyme type 2 (MFE-2) is one class of proteins in mammals, which catalyses β-oxidation of fatty acids. MFE-2 consists of three domains. The C-terminal domain shows similarities to sterol carrier protein type 2 (SCP-2) and is therefore referred to as sterol carrier protein 2-like (SCP-2L). SCP-2 proteins are known to have high binding affinity for a variety of different substrates like cholesterol (and other steroids), various phospholipids,[1] or long chain fatty acids.[2]

SCP-2L was described as almost spherical with both a large and a small hydrophobic tunnel and several large apolar cavities.[3] The crystal structure of SCP-2L is depicted in figure 1.

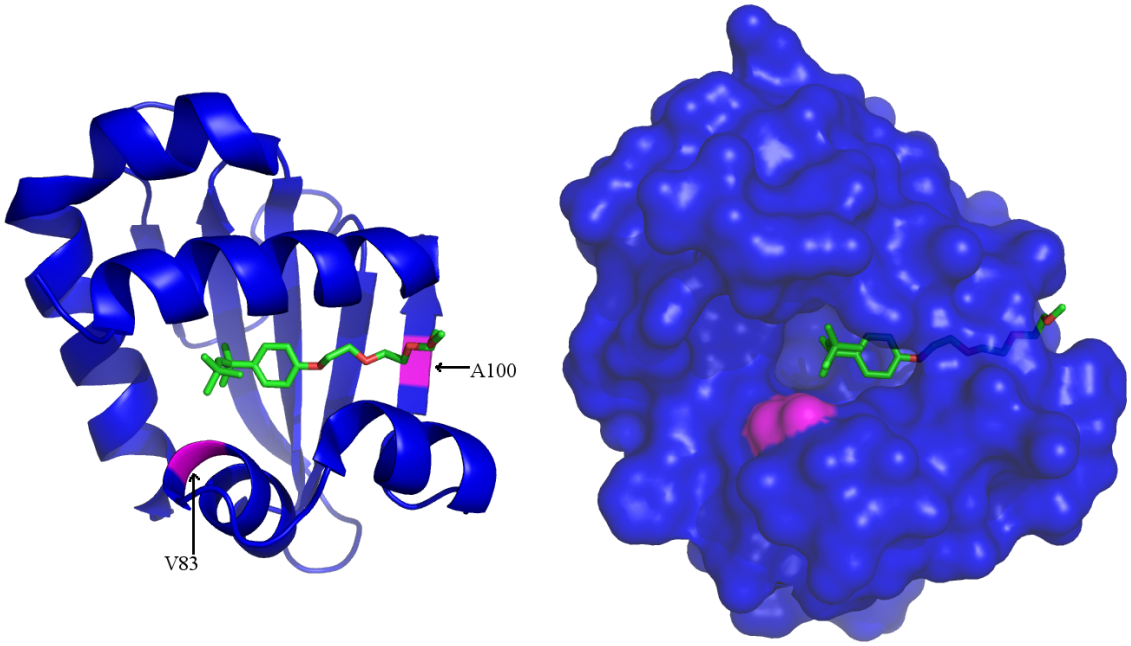
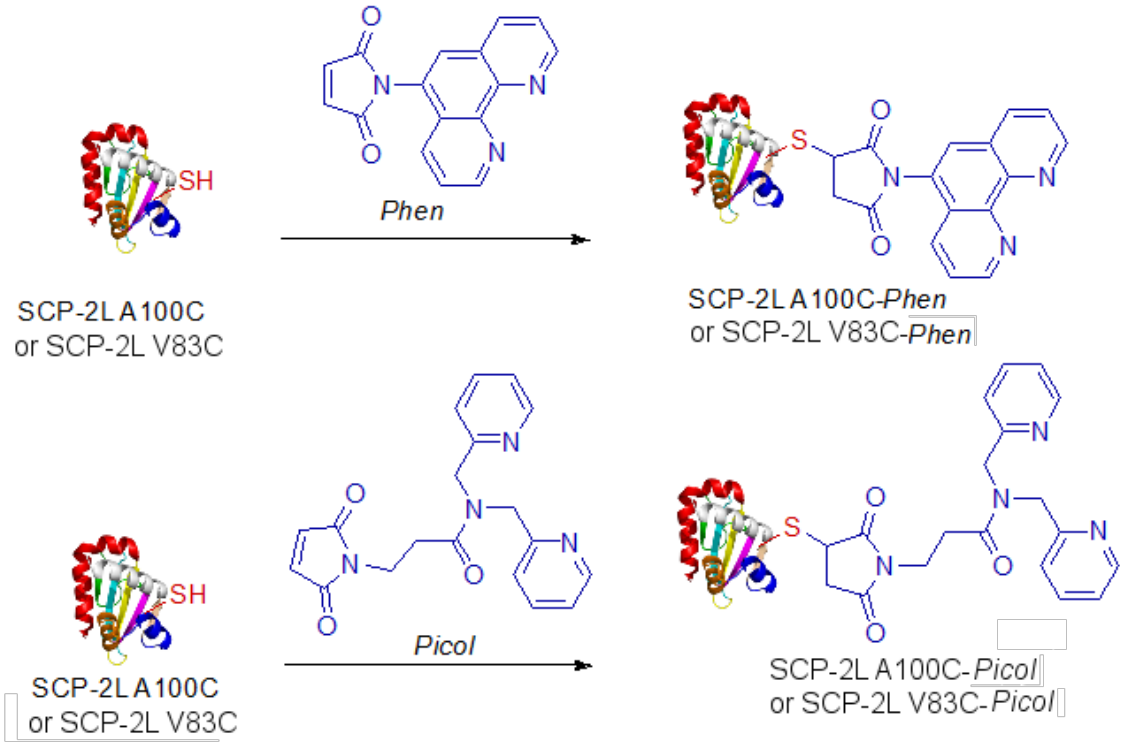


Fig. 1 Crystal structure of SCP-2L with ligand Triton X-100 within the (large) hydrophobic tunnel (pdb-code: RCSB013373). Triton X-100 is shown as ball and stick model with carbon in green, oxygen in red; hydrogens are omitted. Left: Cartoon representation of the protein, protein in blue except for the amino acid V83 and A100 which are coloured magenta. Right: Surface view with 20% transparency. Protein coloured in blue except for V83 which is coloured magenta.

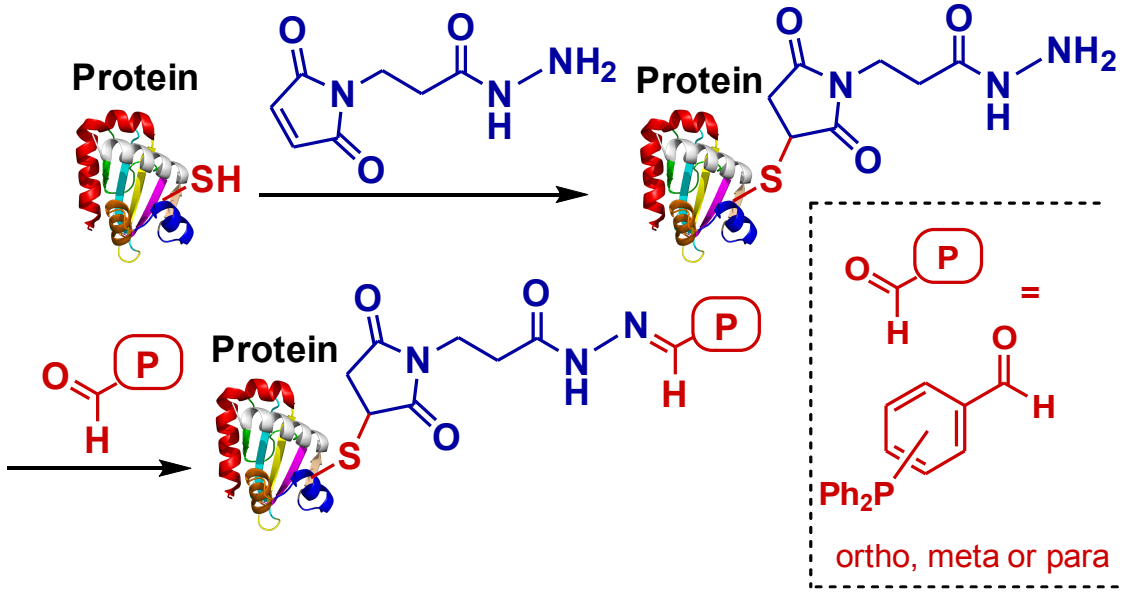
The most relevant part of the protein is the hydrophobic tunnel in which a molecule of Triton X-100 is located. Additionally the two amino acids V83 and A100 are highlighted. These amino acids are of importance because of their location at either end of the large hydrophobic tunnel (see fig. 1). They can be replaced with cysteine, which is not found elsewhere in the protein. Therefore research had been focussed on the two mutants SCP-2L A100C and SCP-2L V83C. Both mutants can be expressed in high yield, its purification is straightforward and both have a remarkable temperature stability.[4] These are important features of a potential artificial metalloenzyme. By covalent modification of the unique introduced cysteine several artificial enzymes bearing nitrogen ligands[5] (see scheme 1) were successfully synthesised.



Scheme 1 Protein modification with the *N*-ligands used in this thesis.

In this procedure the thiol group of the cysteine is site specifically modified with a *N*-ligand bearing maleimide. The introduced *N*-ligands are based on either the bidentate 1,10-phenanthroline structure (*Phen*) or the tridentate di-(2-picolyl)amine-based

structure (*Picol*). The synthesis of these ligands as well as the modification of the two SCP-2L mutants is straightforward but only copper-enzymes have been investigated so far.[4] This places the introduced *N*-ligand in close proximity to either end (dependent on the mutant used) of the large hydrophobic tunnel present within the SCP-2L *via* a covalent linkage (see scheme 1). A method for the site specific introduction of phosphine moieties has already been described in the literature[6] and used within this work (see scheme 2)



Scheme 2 Protein modification with phosphine ligands and the phosphines used in this thesis.

This procedure comprises a series of sequential reactions. In the first step the thiol group of the cysteine is reacted with a short commercially available hydrazide bearing maleimide linker (abbreviated as "1"). In the second step the phosphine is introduced by reacting the hydrazide with an aldehyde bearing phosphine. Three different mono-phosphines (*ortho*, *meta* and *para*-derivative) which are all based on triphenylphosphine were used in this thesis (see scheme 2).

The subsequent treatment of the modified SCP-2L with an appropriate metal resulted in artificial metalloenzymes which were tested as catalysts in the Diels-Alder-,[5] and the hydroformylation reaction[4] with promising results. The introduced metal was assumed to be in close proximity to the hydrophobic tunnel. This was used as rationalisation for

the rate accelerations observed when apolar substrates were used, as they would accumulate within the tunnel, in close proximity to the introduced metal.[4] CD-spectroscopy, ^{1}H NMR and binding studies did indicate that the tertiary structure of the protein was not influenced by neither the mutations nor the chemical modifications of the protein scaffold[4] but crystal structures as a definitive proof were not obtained so far.

Metalloenzymes

It is estimated that about 30% of all characterised proteins are associated with at least one metal.[7] A wide array of qualitative and quantitative characterisation methods for these are known and used.[8–10] The characterisation of artificial metalloenzymes is usually based on ultraviolet-visible[11–15] (UV-VIS) spectroscopy, circular dichroism[14–17,5,18,19] (CD) spectroscopy, mass spectroscopy,[6,11,13,14,16,18,20–23] usually electrospray ionisation (ESI), sometimes matrix assisted laser desorption/ionisation (MALDI)) and inductively coupled plasma[16,20,22,24] (ICP)-techniques. If applicable other techniques like fluorescence measurements[5,12,23], electron paramagnetic resonance[13,25] (EPR) and nuclear magnetic resonance[6,23] (NMR) are used as well.

Artificial metalloenzymes in this thesis

Various phosphine- and *N*-ligand-modified proteins were used within this thesis. SCP-2L based ones are most relevant, therefore it was aimed to obtain the crystal structures of the modified proteins, to gain insight on the protein environment surrounding the introduced ligand. This would make it possible to fine-tune the active site by biological or chemical means in respect to its substrate- or metal coordination. In order to obtain artificial metalloenzymes suitable for catalysis it is of utmost importance to determine the binding capabilities of the modified proteins towards transition metals before testing them in catalytic reactions. An artificial enzyme with a small binding capability towards a metal is an undesired catalyst, as a comparably high concentration of unligated metal would be present in solution. This unligated metal could outperform the catalytic activity of the artificial metalloenzyme itself, which would make it impossible to draw any conclusions on the performance of the artificial metalloenzyme. In theory the concentration of unligated metal could also be reduced by using a high artificial enzyme

to metal ratio, but due to the time consuming synthesis of artificial enzymes this is less desirable. In this chapter these issues are addressed.

2.3 Results and discussions

Crystal structure of SCP-2L A100C

Most of the artificial metalloenzymes investigated in this thesis are based on SCP-2L A100C and SCP-2L V83C, therefore it was of high importance to obtain the crystal structures of the two mutants. Additionally, attempts on the crystallisation of sulphur protected phosphine-bearing modifications were performed, which were unsuccessful (see experimental section). The crystallisation attempts on the mutants SCP-2L A100C and SCP-2L V83C were successful but only the refined crystal structure of SCP-2L A100C could be obtained so far.

Various aspects of the crystal structure of SCP-2L A100C in comparision to its wild-type are depicted in figure 2 and figure 3.

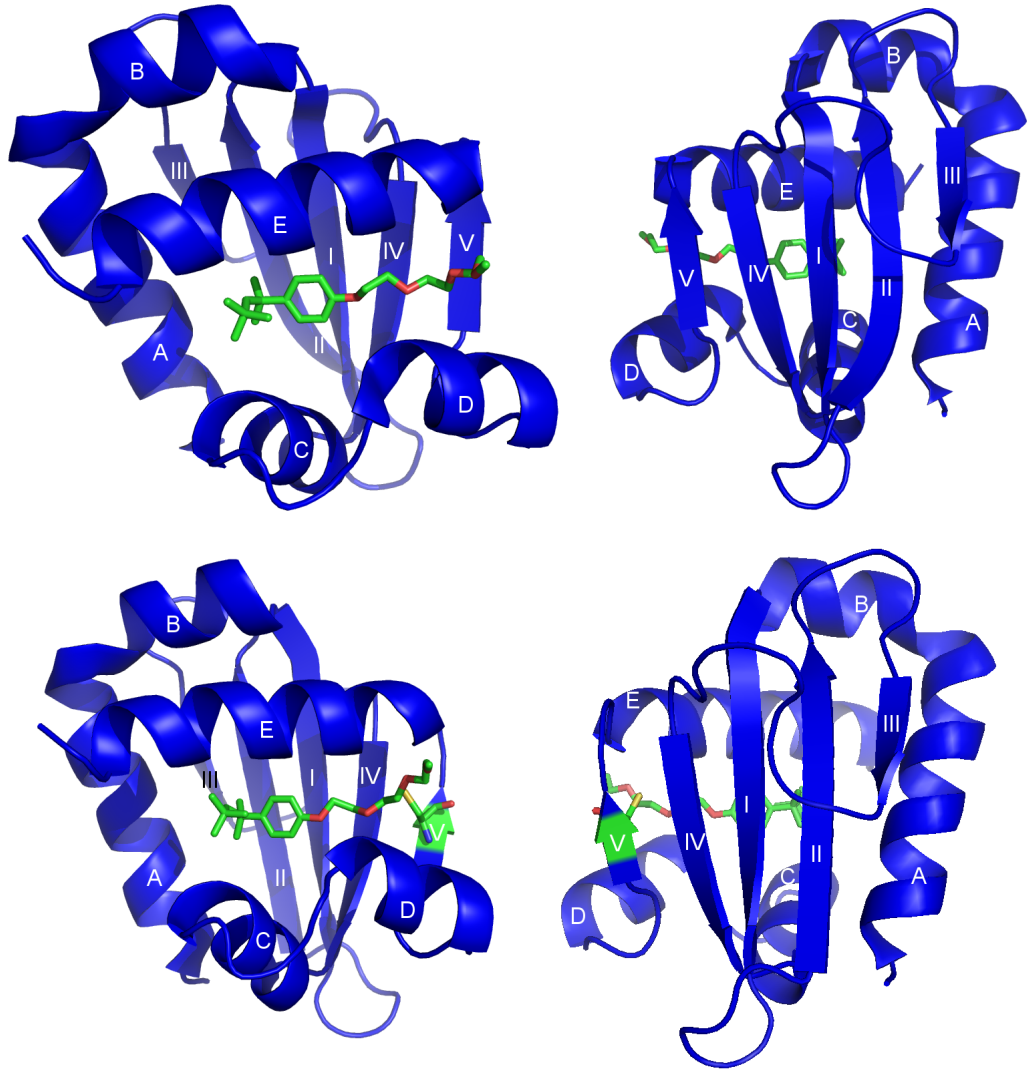


Fig. 2 Crystal Structure of SCP-2L and SCP-2L A100C with ligand Triton X-100. The ligand Triton X-100 within the large hydrophobic tunnel is shown as a ball and stick model with carbon in green, oxygen in red; hydrogens are omitted. Protein represented as cartoon and coloured in blue except for the single amino acid cysteine which is coloured green. Cysteine is shown as ball and stick model with the thiol as yellow. β-strands and α-helices are indicated with Roman numbers (I-IV) and Roman capitals (A-E) respectively.

Top: Crystal structure of SCP-2L

Bottom: Crystal structure of SCP-2L A100C

Figure 2 shows both SCP-2L (at the top) and the mutant SCP-2L A100C (at the bottom) as cartoon with labelled α-helices and β-strands as a “front” view and turned by 180°. The molecule Triton X-100 which is located within the tunnel is shown as ball and stick model. The introduced cysteine is located on β-strand “V” and coloured green. In case of the mutant SCP-2L A100C the amino acid cysteine is additionally shown as ball and stick with the sulphur coloured in yellow.
In figure 3 the surface potential of SCP-2L and SCP-2L A100C is depicted:

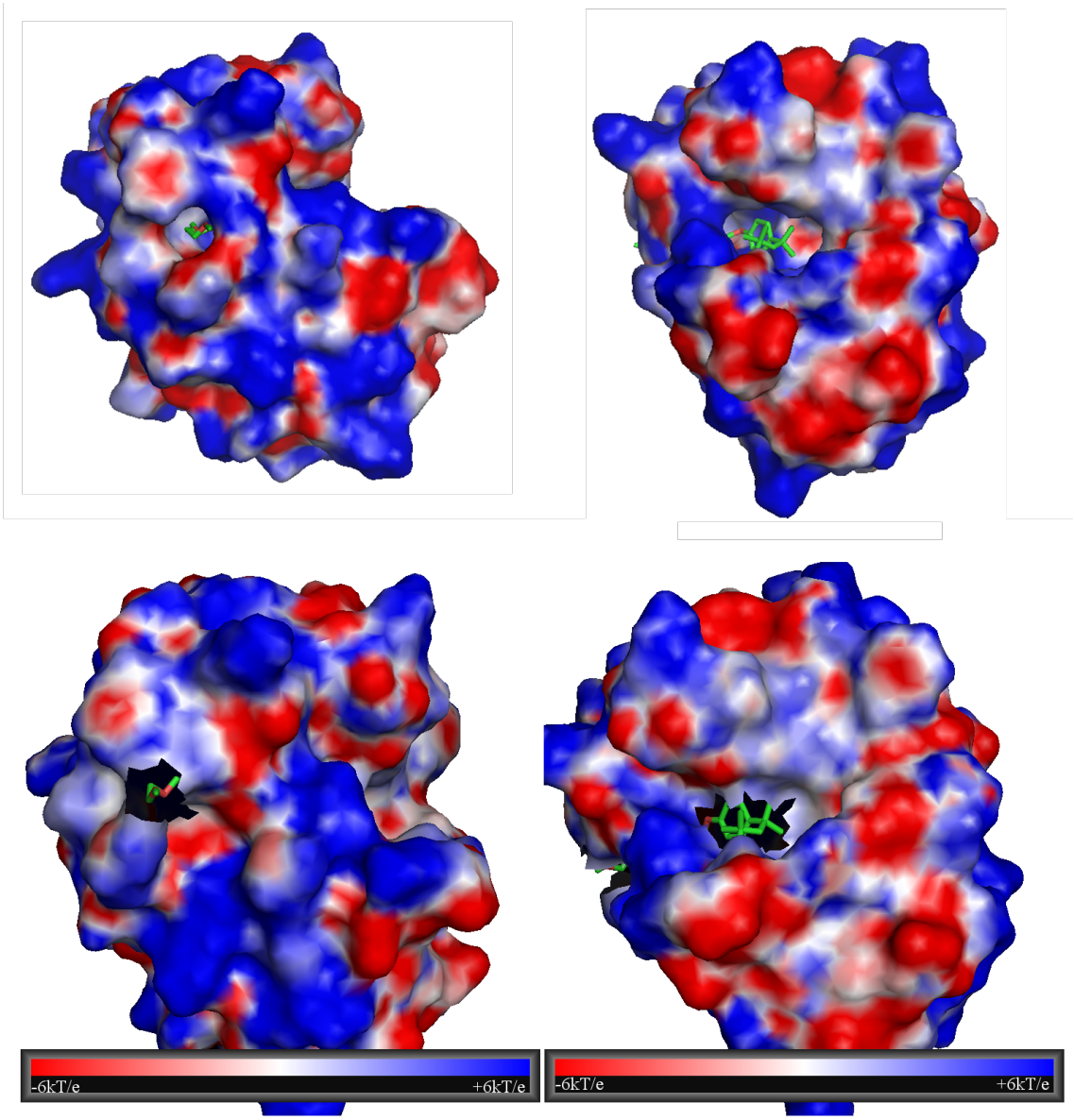


Fig. 3 Electrostatic surface potential of SCP-2L A100C and the wild-type protein SCP-2L with ligand Triton X-100. The ligand Triton X-100 within the large hydrophobic tunnel is shown as a ball and stick model with carbon in green, oxygen in red; hydrogens are omitted. Red corresponds to a negative, white to a neutral and blue to a positive surface potential. Either entrance of the large tunnel is shown.
Top: Electrostatic surface potential of SCP-2L
Bottom: Electrostatic surface potential of SCP-2L A100C

In Figure 3 the electrostatic surface potential of SCP-2L (top) and SCP-2L A100C (bottom) is depicted (calculated using APBS, see experimental section). The left side of

figure 3 is focussing on the electrostatic surface poteintial next to the side where the mutation took place. The tunnel entrance and the ethoxychain of the Triton X-100 is visible (Triton X-100 is represented as a ball and stick model). The electrostatic surface potential of the protein next to the tunnel is mostly neutral (represented by a white colour). Most of the surface visible in this figure has a positive surface potential (represented by a blue colour). The right half of the figure shows the electrostatic surface potential of the protein focussing on the other end of the tunnel. The surface potential around this end of the tunnel is -similar to most of the protein surface shown in this figure - mostly neutral. Although this side of the protein contains more areas with a negative surface potential these areas are quite scattered making the hydrophobic tunnel by far the major hydrophobic area of the mutant.

The crystal structure for the SCP-2L A100C mutant and its surface potential is very similar to the previously published crystal structure of its wild type.[3]

As the large tunnel was considered important for any artificial metalloenzyme based on this scaffold, it was further investigated by analysing the contact environment of Triton X-100. The schematic contact environment between Triton X-100 and the amino acids of SCP-2L A100C is shown in fig. 4.

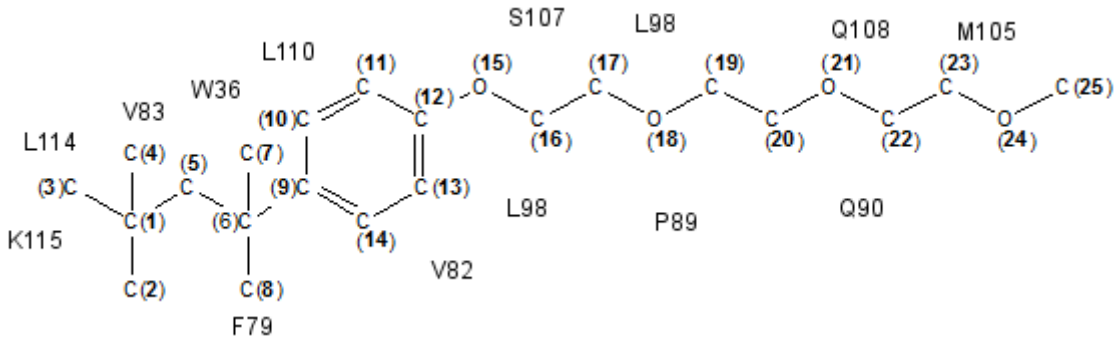


Fig. 4 Schematic contact environment between Triton X-100 and SCP-2L A100C. Amino acids are abbreviated using their 1-letter code, their residue number is given. All atoms of Triton X-100 are numbered.

The contact environment has many similarities compared to the contact environment between SCP-2L and Triton X-100. In order to perform a profound comparison between

the tunnel of the mutant and the wild type protein the distances between each atom of the Triton X-100 to each atom of both the wild type protein and the mutant (when either of it has a distance ≤4.0 Å) was compared (see Table 8 in the experimental section for all details). The results are quite similar for the wild type SCP-2L and its mutant SCP-2L A100C, but some minor differences were observed:

- The amino acid L114 is in close proximity to C7 of the Triton X-100 in the native protein. In case of the mutant SCP-2L A100C, L110 is in close proximity to C7 instead of L114. This shows the flexibility of the protein structure.
- The amino acids next to C11 are much further away in case of the mutant. Therefore the tunnel of SCP-2L A100C has a larger diameter at this place.
- The ethoxychain (especially C17, then from C20 to C25) is much closer to the protein in case of the wild-type than in the case of the SCP-2L A100C mutant. Therefore the tunnel of SCP-2L A100C has a larger diameter at these regions as well.

The single A100C is close to the end of the tunnel where the ethoxychain of the Triton X-100 is located. The cysteine side-chain is slightly larger than the alanine and an expansion of the tunnel diameter in this region can be observed.

Metal complexes of *N*-Ligand modified proteins

Artificial metalloenzymes covalently modified with *N*-Ligands were synthesised according to the previously described site-specific modification (see scheme 1).[5] The previously introduced mutants SCP-2L A100C and SCP-2L V83C were used as scaffolds. Subsequently, 2 eq. of the metal salts were added to the *N*-ligand modified proteins to obtain different artificial metalloenzymes. These were characterised by ESI-MS and ICP-MS. The acetate salts of the following metals were used: Pd(II), Zn(II), Mn(II), Fe(II), Cu(II), Ni(II) and Co(II).

ESI-MS data of *Picol*-modified proteins

The ESI-MS data regarding the *Picol*- modified proteins as well as the corresponding control reactions where the unmodified proteins SCP-2L A100C and SCP-2L V83C

were used showed no metal-modified protein adducts except in the case of Pd(II). When palladium(II)-acetate was used adducts between modified protein and the metal ion were detected for both *Picol*-modified protein scaffolds. In the reaction between SCP-2L V83C-*Picol* and Pd(II) the new peaks could be characterised as species with one, two or three Pd(II) coordinating to the scaffold (see fig. 1 in the appendix). In the reaction between SCP-2L A100C-*Picol* and Pd(II) the new species were mostly unassigned, but three peaks corresponding to species with one, two or three Pd(II) coordinating to the scaffold could be assigned as well (see fig. 2 in the appendix).

After treating the SCP-2L A100C-*Picol*-Pd mixture with EDTA and removal of EDTA-complexed metal by filtration (see experimental section for more details) the major species detected by ESI-MS is the 1:1 (Protein : Pd)-complex (see fig. 3 in the appendix). This indicates that one of the binding sites is much stronger than the other ones.

When unmodified protein SCP-2L A100C or SCP-2L V83C were treated with Pd(II) the resulting ESI-MS did not show any distinguishable mass peaks due to a low signal to noise ratio. Therefore it cannot be concluded whether the unmodified mutants posess any Pd(II) binding sites. From this data it is yet unclear whether the strong Pd(II)-binding observed is due to the introduced ligand or to the protein scaffold itself.

ESI-MS data of *Phen*-modified proteins

The ESI-MS data regarding the *Phen*-modified proteins as well as the corresponding control reactions where unmodified protein was used are listed in table 1.

Table 1 ESI-MS analysis of artificial metalloenzymes based on unmodified- and *Phen*-modified protein[a]

entry	metal salt	Protein scaffold			
		SCP-2L A100C	SCP-2L V83C	SCP-2L A100C-*Phen*	SCP-2L V83C-*Phen*
1	$Pd(OAc)_2$	-[b]	-[b]	+[c]	+[d]
2	$Zn(OAc)_2$	-	-	-	-
3	$Mn(OAc)_2$	-	-	-[e]	-
4	$Fe(OAc)_2$	-	-	-	-
5	$Cu(OAc)_2$	-	-	+[f]	+
6	$Ni(OAc)_2$	-	-	+[f]	+
7	$Co(OAc)_2$	-	-	-	+[g]

a: Reaction and analysis conditions: To 0.1 µmol protein in buffer 2-3 eq. of metal salt were added, reaction time 1 day. After buffer exchange a fifth was analysed by LC-MS. When SCP-2L A100C-*Phen* was used and the metal complex (or an unknown complex in case of Mn(II)) was detected the remaining solution was treated with EDTA and buffer exchanged. The EDTA treatment was repeated three times and the supernatant analysed by LC-MS. See experimental section for all details. **+**: A mass corresponding to the expected mass of the artificial metalloenzymes was detected, **-**: No masses higher than the (un)modified protein were detected (excluding sodium adducts). **B**: Low signal to noise ratio **C**: A mass corresponding to two metals coordinating were detected. After treatment with EDTA the noise to signal ratio is too low **D**: Peaks with masses corresponding to complexes with two and three metals coordinating were detected **E**: A species with a mass of 60-63 Da higher than the modified protein is detected. After treatment with EDTA only the modified protein can be detected by ESI-MS **F**: After treatment with EDTA only the modified protein can be detected by ESI-MS. **G**: The peak with the mass corresponding to the metal enzyme is broad and only slightly higher than the background signal, therefore it might be noise.

When *Phen*-modified proteins were treated with Palladium(II)-acetate (entry 1) the results are similar to the ones obtained when *Picol*-modified proteins were used. In the reaction between SCP-2L A100C-*Phen* and Pd(II) species corresponding to one and two Pd(II) coordinating were assigned (see fig. 4 in the appendix). In the reaction between SCP-2L V8C-*Phen* and Pd(II) the protein scaffold as well as species with one, two and three Pd(II) coordinating to the scaffold were detected (see fig. 6 in the appendix). When unmodified protein SCP-2L A100C or SCP-2L V83C were treated with Pd(II) the resulting ESI-MS did not show any distinguishable mass peaks due to a low signal to noise ratio (as mentioned before). After treating the SCP-2L A100C-*Phen*-Pd mixture with EDTA and removal of unbound metal by filtration (see experimental section for more details) the major species detected by ESI-MS is the mono-Pd-complex (see fig. 5

in the appendix).

When Zn(II)- (entry 2) or Fe(II)-acetate (entry 4) were used as metal precursors the masses corresponding to the artificial metalloenzymes could not be detected for the two control reactions and the *Phen* modified proteins.

When Mn(II)-acetate was used as metal precursor (entry 3) no masses corresponding to the artificial metalloenzymes could be detected for neither the unmodified proteins nor the SCP-2L V83C-*Phen* scaffold. After treating SCP-2L A100C-*Phen* with Mn(II) a peak with a mass of ≈61 Da higher than the modified protein was detected (see fig. 7 in the appendix). Due to the broadness of the peaks it might correspond to the Mn(II)-enzyme (M(Mn) = 55 Da). After EDTA treatment this species could not be detected anymore, strengthening this hypothesis (see fig. 7 in the appendix).

When Cu(II)- (entry 5) or Ni(II)-acetate (entry 6) were used as metal precursor no masses corresponding to the artificial metalloenzymes could be detected in the control reactions where unmodified protein was used. Both *Phen*- modified scaffolds did show the expected mass corresponding to the respective artificial metalloenzymes (see fig. 8-11 in the appendix). After treatment of both the Ni(II)- and Cu(II)-enzymes based on the SCP-2L A100C-*Phen* scaffold with EDTA and subsequent analysis by ESI-MS only the mass corresponding to the metal-free SCP-2L A100C-*Phen* could be detected (see fig. 10 for SCP-2L A100C-*Phen*-Ni(II) and fig. 11 for SCP-2L A100C-*Phen*-Cu(II); both in the appendix).

When Co(II)-acetate was used as metal precursor (entry 7) no masses corresponding to the artificial metalloenzymes could be detected for neither the unmodified proteins nor the SCP-2L A100C-*Phen* scaffold. The mass corresponding to SCP-2L V83C-*Phen*-Co could be detected (see fig. 12 in the appendix).

Summary on the ESI-MS data of *N*-ligand modified and unmodified artificial metalloenzymes

Addition of Pd(II) to unmodified mutants of SCP-2L resulted in low signal to noise ratios. Addition of Pd(II) to modified SCP-2L mutants bearing *Phen* or *Picol*-ligands resulted in up to three Pd(II)-ions coordinating to the modified proteins. This indicates high affinity of Pd(II) and probably unspecific coordination to exposed amino acids.

The 1:1 complexes of Pd(II) and SCP-2L A100C modified with either *Phen* or *Picol* could be detected by ESI-MS after EDTA treatment, indicating that one Pd(II) atom per modified protein is coordinating much stronger than the other ones.

In contrast to *Phen*-based artificial metalloenzymes, *Picol*-based ones (except for the Pd(II) discussed above) could not be detected by ESI-MS. This indicates a higher stability of the *Phen*-based artificial metalloenzymes in general. This finding is contrary to the expectation, as the equilibrium constants of the 1:1 complexes between Cu(II), Ni(II), Co(II) or Zn(II) and di-(2-picolyl)amine are much much higher than the equilibrium constants for the corresponding 1:1 complexes with 1,10-phenanthroline as ligand. The equilibrium constants of the 1:1 complex between Fe(II) or Mn(II) and 1,10-phenanthroline are higher than the corresponding constants of di-(2-picolyl)amine under the same conditions.[26–28] The equilibrium constants between Fe(II) or Mn(II) to the ligands are significantly smaller than the equilibrium constants of Ni(II), Cu(II), Co(II) or Zn(II) (the literature known complexation constants are listed in table 9 in the experimental section of this chapter). Even when the influence of the slightly acidic pH used in the protein complexation reactions is taken into account the phenomena cannot be explained.

The complexation constants of the *Phen*-modified proteins are higher than the corresponding *Picol*-modified proteins, which is basically the opposite of the literature known data of the "free" ligands. This might be explained by the coordination of multiple *Phen*-modified proteins to a single metal which seems unlikely as the introduced ligand is supposed to be buried within a pockt. It might also be possible that metal complexes of *Phen*-modified protein are subject to a higher stabilisation (compared to the *Picol*-modified proteins) due to amino acids coordinating to the metal. While *Picol* is a tridentate ligand in solution it might be possible that *Picol* bound to the protein can only act as a bidentate ligand due to steric hindrance caused by the bulky protein.

ICP-MS of SCP-2L A100C scaffolds

After metal treatment of the protein scaffolds SCP-2L A100C, SCP-2L A100C-*Phen* and SCP-2L A100C-*Picol* the coordinating metal was removed by EDTA and the EDTA and metal containing fractions combined, concentrated and analysed by ICP-MS to calculate the initial metal loading (see experimental section for details). The resulting data are listed in table 2.

Table 2 ICP-MS and metal loading of different artificial metalloenzymes.[a]

entry	metal salt	Protein scaffold SCP-2L A100C		SCP-2L A100C-*Picol*		SCP-2L A100C-*Phen*	
		ppb	% metal loading[b]	ppb	% metal loading[b]	ppb	% metal loading[b]
1	$Pd(OAc)_2$	1500	9	7900	93	3900	23
2	$Zn(OAc)_2$	-600[c]	-6[c]	1000	10	1000	10
3	$Mn(OAc)_2$	800	9	1200	13	500	5
4	$Fe(OAc)_2$	15600[d]	175[d]	3600	41	12000	70
5	$Cu(OAc)_2$	21000[e]	66[e]	1200	6	4900	23
6	$Ni(OAc)_2$	600	6	700	8	1000	11
7	$Co(OAc)_2$	200	3	300	3	2500	27

a: Reaction and analysis conditions: To 0.1 µmol protein in buffer 2 eq. of metal salt were added (in case of SCP-2L A100C 3 eq. were used), reaction time 1 day. After buffer exchange a fifth was analysed by LC-MS. The remaining solution was treated with EDTA and the buffer was exchanged. The EDTA treatment was repeated three times and the flow through collected, concentrated and its metal content determined by ICP-MS. See experimental section for all details. **b**: n(metal) / n(Protein), corrected for the 1/5th removed for ESI-MS analysis; 50% corresponds to 0.5 mol metal per 1 mol (un)modified protein, 100% corresponds to 1 mol metal per 1 mol (un)modified protein and so on. **c**: Rather high concentration of metal in the blank is probably the reason for the negative value **d**: Probably due to multiple metals coordinating to the scaffold. **e**: In a control reaction where SCP-2L V83C was reacted with $Cu(OAc)_2$ under the same conditions 950 ppb Cu (corrected for the blank) were determined by ICP-MS which corresponds to 18.7% metal.

It should be noted that SCP-2L A100C is not a perfect reference as its thiol is likely to coordinate metals, The wild type SCP-2L without the thiol would have been a better comparison. Comparing the metal loading of SCP-2L A100C-*Phen* and SCP-2L A100C-*Picol* still gives insight into the different binding capabilities of the two artificial enzymes.

SCP-2L A100C-*Picol* has a significantly higher metal loading (and therefore a higher

equilibrium constant) to Pd(II) (entry 1) and a slightly higher to Mn(II) (entry 3) than SCP-2L A100C-*Phen* and the unmodified mutant.

The metal loading for SCP-2L A100C-*Phen* and Fe(II) (entry 4), Cu(II) (entry 5), Ni(II) (entry 6) and Co(II) (entry 7) is significantly higher than the metal loading for SCP-2L A100C-*Picol* and the corresponding metal.

The metal loading between both modified proteins and Zn(II) (entry 2) is basically identical.

The metal loading of the modified proteins is in general higher than for the unmodified protein. The only exceptions are Fe(II) (entry 4) and Cu(II) (entry 5). The high metal loading of the unmodified protein in these cases might be due to the previously mentioned free thiol.

The differences in metal loading indicate the coordination of metals to the introduced ligand. Only in case of Fe(II), Cu(II) and Mn(II) no such conclusion can be drawn.

The metal loadings listed in the table were also used to calculate equilibrium constants between the *N*-ligand modified protein and each metal. Due to several simplifications the calculated equilibrium constants are most likely inaccurate (see experimental section for details).

Conclusion on the ESI-MS- and ICP-MS-analysis of *N*-ligand modified proteins

The Pd-enzyme derived from the *Picol*-modified SCP-2L A100C show high metal loading and could be detected by ESI-MS. This is a strong indication for a promising artificial metalloenzyme. The *Phen*-modified SCP-2L A100C mutants were found to be promising scaffolds for Pd(II), Cu(II) and Ni(II)-enzymes as they could be detected by ESI-MS indicating a strong binding and possessed a high metal loading according to ICP-MS as well. *Phen*-modified SCP-2L V83C showed the same promising ESI-MS results for these metals, the corresponding metal loadings were not determined but can be expected to be similar. The results for the Co-enzyme of the *Phen* modified SCP-2L mutants tested is ambivalent. While SCP-2L V83C-*Phen*-Co was detected by ESI-MS, SCP-2L A100C-*Phen*-Co could not be detected by ESI-MS but this Co-enzyme showed a high metal loading by ICP-MS.

Metal complexes of proteins modified with *P*-ligands

Synthesis and characterisation of phosphine containing proteins

In addition to the *N*-ligand modified proteins, phosphine-modified proteins were used for metal complexation reactions as well. Four cysteine containing enzymes were used as scaffold: SCP-2L A100C, SCP-2L V83C, a mutant of the photoactive yellow protein (PYP R52G) and lysozyme. The latter two proteins were meant to act as control reactions as they do not posess a hydrophobic tunnel.

In case of the SCP-2L based proteins, as well as of PYP R52G, the thiol group of the cysteine present in each protein was modified according to the literature known procedure introduced earlier (see scheme 2).[6]

Lysozyme contains eight cysteines, all of which are present as disulfides. In order to modify lysozyme the disulfides were reduced prior to the modification described above. For this reduction two different reducing agents were tested (TCEP and DTT). The linker modification after reduction with TCEP was more efficient as the reduction using DTT resulted in the degradation of lysozyme (see experimental section). The product obtained was a mixture of lysozyme species containing up to four linker moieties (the product mixture is refered to as lysozyme-$(1)_x$), where "1" represents the hydrazide linker (see fig. 18 in the appendix for the ESI-MS). The modified lysozyme was used as the feedstock for the subsequent modification reactions with the three different phosphines. The reaction between lysozyme-$(1)_x$ and *P(para)* as well as *P(meta)* were incomplete after long reaction times (about one week) and vast excess of phosphine (up to 25 eq.) (see fig. 19 and fig. 20 in the appendix for the ESI-MS). In both cases only the phosphine oxides were detected by ESI-MS indicating that these lysozyme based artificial enzymes are much more oxygen sensitive than the SCP-2L and PYP based ones. The modification with *P(ortho)* was virtually unsuccessful yielding only marginal quantities of phosphine modified lysozyme (according to ESI-MS, only the oxides were detected). The lysozyme-based partly phosphine modified artificial enzymes were used for metal complexation reactions with Rh(I) as well.

Metal complexes of *P*-ligand modified proteins

The synthesised artificial proteins bearing a phosphine were treated with different metal salts to obtain the corresponding artificial metalloenzymes. Table 3 gives an overview of all phosphine based artificial metalloenzymes synthesised and analysed in this study.

Table 3 List of artificial metalloenzymes bearing a phosphine synthesised in this thesis.[a]

entry	protein scaffold	metal precursor											
		Mn(II)	Fe(II)	Co(II)	Ni(II)	Cu(II)	Zn(II)	Ru(III)	Rh(I)	Pd(II)	Re(I)	Ir(I)	Pt(II)
1	SCP-2L A100C-1-P(para)[b]	+	+	+	+	+	+	+	+	+	+	+	+
2	SCP-2L A100C-1-P(meta)[b]	-	-	-	-	-	-	-	+	-	-	-	-
3	SCP-2L A100C-1-P(ortho)[b]	-	-	-	-	-	-	-	+	-	-	-	-
4	SCP-2L V83C-1-P(para)[c]	+	+	+	+	+	+	+	+	+	+	+	+
5	SCP-2L V83C-1-P(meta)[c]	-	-	-	-	-	-	-	+	-	-	-	-
6	SCP-2L V83C-1-P(ortho)[c]	-	-	-	-	-	-	-	+	-	-	-	-
7	Lysozyme-$(1)_x$-P(para)[d]	-	-	-	-	-	-	-	+	-	-	-	-
8	Lysozyme-$(1)_x$-P(meta)[d]	-	-	-	-	-	-	-	+	-	-	-	-
9	Lysozyme-$(1)_x$-P(ortho)[d]	-	-	-	-	-	-	-	+	-	-	-	-
10	PYP R52G-1-P(para)[e]	-	-	-	-	-	-	-	+	-	-	-	-

a: Slightly different reaction conditions for (almost) every metal complexation reaction. See experimental section for details. **b**: protein scaffold synthesised according to a published procedure:[4] **c**: protein scaffold synthesised according to a published procedure:[6] **d**: Synthesised according to the procedure outlined in the experimental section **e**: Synthesised by Dr. W. Laan according to a published procedure:[29] "**+**": Synthesis was performed "**-**" No attempt on synthesising the artificial metalloenzymes performed.

All resulting artificial metalloenzymes were analysed by ESI-MS and ICP-MS, except for PYP R52G-1-P(para)-Rh which was only characterised by ESI-MS.

Characterisation of phosphine based artificial metalloenzymes

ESI-MS of Rh-enzymes

The ESI-MS of the complex between SCP-2L A100C-1-P(para) (entry 1) and Rh(acac)$(CO)_2$ as well as the complex of SCP-2L V83C-1-P(para) (entry 4) and Rh(acac)$(CO)_2$ were published previously[4] and could be reproduced (see figures 21 and 22 in the appendix). The ESI-MS shows three peaks. The first peak corresponds to the phosphine modified protein. The second peak has a mass which is 16 Da higher than the first peak and corresponds to the oxidised phosphine modified protein (oxidation most likely due to the ESI-MS conditions which are not oxygen free). The third peak has a mass which

is 131 Da higher than the first peak. This difference was attributed to the presence of the fragment "Rh(CO)" (M (RhCO) = 131 Da). Two additional Rh(I) precursors were tested on SCP-2L A100C-1-P(para). When $[Rh(MeCN)_2COD]BF_4$ was used as precursor a new peak with a mass of 99 Da higher than the phosphine modified scaffold is observed (see fig. 13 in the appendix). Due to the broadness of the peaks and the resulting mass-inaccuracy this was attributed to a Rh-adduct (M (Rh) = 103 Da). When $RhH(PPh_3)_3CO$ was used as precursor no additional peaks beside the modified protein and its oxide could be detected. This indicates the importance of the metal precursor used in order to detect species by ESI-MS. The reaction between SCP-2L A100C-1-P(meta) (entry 2) and $Rh(acac)(CO)_2$ resulted in a MS with a high noise to signal ratio making it impossible to allocate peaks. The ESI-MS of the reaction between SCP-2L A100C-1-P(ortho) (entry 3) and $Rh(acac)(CO)_2$ contains the peaks corresponding to the phosphine and its oxide but no higher mass adducts. When SCP-2L V83C-1-P(meta) (entry 5) was treated with $Rh(acac)(CO)_2$ the Rh-enzyme was detected by LC-MS. Peaks corresponding to the modified enzyme, its oxide as well as a peak with a mass difference caused by a Rh(CO)-fragment were detected (see fig. 14 in the appendix). It should be noted that the signal to noise ratio was quite low. The ESI-MS obtained for the reaction between SCP-2L V83C-1-P(ortho) (entry 6) and $Rh(acac)(CO)_2$ showed three main peaks (see fig. 15 in the appendix). The first two correspond to SCP-2L V83C-1-P(ortho) and its oxide. The third peak has a mass of "SCP-2L V83C-1-P(ortho) + 100". Due to the broadness of the peaks this was attributed to a Rh-adduct (M (Rh) = 103 Da). The issues with the phosphine modification of linker modified lysozyme was mentioned before. When the lysozyme based phosphines as well as lysozyme-$(1)_x$ were treated with $Rh(acac)(CO)_2$ no rhodium adducts could be detected by ESI-MS analysis for any of these proteins (entries 7-9). The ESI-MS obtained when PYP R52G-1-P(para) was treated with $Rh(acac)(CO)_2$ results in a rather complicated spectrum (see fig. 16 in the appendix). Five peaks were identified: The first peak corresponds to the phosphine modified protein PYP R52G-1-P(para) (M = 16325 Da). The second and third peaks correspond to the mono and double oxidised species. The fourth peak (M = 16456 Da) corresponds to the desired product PYP R52G-1-P(para)-Rh(CO). The fifth peak corresponds to an oxidised species of the Rh-enzyme. The next group of three peaks

with masses of 16727 Da, 16745 Da and 16763 Da respectively cannot be accounted for but it seems plausible to assume that the latter two of this group are the oxidised and double oxidised species of the first.

ICP-MS of Rh-enzymes

After extensive washing of each Rh-enzyme using molecular weight cut-off technique to retain both the modified enzyme and the Rh-enzyme the rhodium content was determined by ICP-MS (see experimental section for details). The ICP-MS data for the reaction between the phosphine-modified proteins, unmodified protein and Rh(I) are listed in table 4.

Table 4 ICP-MS and metal loading of different artificial metalloenzymes.[a]

Protein scaffold	Rh(I) precursor					
	$Rh(acac)(CO)_2$		$[Rh(MeCN)_2COD]BF_4$		$RhH(PPh_3)_3CO$	
	ppb	% metal loading[b]	ppb	% metal loading[b]	ppb	% metal loading[b]
Lysozyme-$(1)_x$	7500	32	-[c]	-[c]	-[c]	-[c]
Lysozyme-$(1)_x$-P(para)	6000	31	-[c]	-[c]	-[c]	-[c]
Lysozyme-$(1)_x$-P(meta)	9300	39	-[c]	-[c]	-[c]	-[c]
Lysozyme-$(1)_x$-P(ortho)	7900	34	-[c]	-[c]	-[c]	-[c]
SCP-2L A100C-1-P(para)	5500	39	15400	226	3400	77
SCP-2L A100C-1-P(meta)	7800	96	-[c]	-[c]	-[c]	-[c]
SCP-2L A100C-1-P(ortho)	5500	57	-[c]	-[c]	-[c]	-[c]
SCP-2L V83C-1-P(para)	11500	47	-[c]	-[c]	-[c]	-[c]
SCP-2L V83C-1-P(meta)	10200	62	-[c]	-[c]	-[c]	-[c]
SCP-2L V83C-1-P(ortho)	3200	57	-[c]	-[c]	-[c]	-[c]

a: Slightly different reaction conditions for (almost) every metal complexation reaction. See experimental section for details. **b**: n(metal) / n(Protein). If appropriate: corrected for the 1/5th removed for ESI-MS analysis; 50% corresponds to 0.5 mol metal per 1 mol modified protein, 100% corresponds to 1 mol metal per 1 mol modified protein and so on. **c**: Reaction not performed

It has to be noted that almost every metal complexation used slightly different initial protein and metal concentrations. Therefore different metal concentrations in the ICP-MS analysis are to be expected. A direct comparison between the different rhodium concentrations and the calculated metal loadings is therefore not sensible but some general conclusions can be drawn.

The rhodium loading for the Rh-enzymes of linker modified lysozyme (entry 1) and linker-phosphine modified lysozymes (entries 2, 3, 4) are within the same range

regardless of the modification. This strongly suggests that rhodium is coordinating to (unspecific) amino acids of the protein scaffold, but not the phosphine ligands.
The rhodium loading of phosphine modified SCP-2L scaffolds is higher in most cases than for the lysozyme based scaffolds. This indicates a better coordination of these phosphine modified scaffolds. When SCP-2L A100C-1-P(para) was treated with different Rh(I) precursors the resulting metal loading is unexpected. When $[Rh(MeCN)_2(COD)]BF_4$ is used as precursor an average of 2.2 rhodium atoms are coordinating to every protein scaffold. Unless the protein scaffold itself undergoes some radical structural change in the presence of this rhodium precursor it is probably fair to assume that the precursor forms multimers (probably dimers and trimers) which are coordinating to the phosphine.

ESI-MS of phosphine based artificial metalloenzymes (excluding rhodium)

As listed in table 3 SCP-2L A100C-1-P(para) and SCP-2L V83C-1-P(para) were treated with a large variety of different metal precursors and analysed by LC-MS. As a comparison the unmodified SCP-2L A100C and SCP-2L V83C were treated with these metal salts as well (SCP-2L V83C was not treated with Ir(I), Pt(II), Ru(III) and Re(I)).
In most cases no higher masses than the (un) modified protein could be detected by ESI-MS indicating a rather weak interaction between the protein scaffold and the metal. In case of Pd(II) all spectra had a low signal to noise ratio making it impossible to draw any conclustions. Only when SCP-2L A100C-1-P(para) was treated with $(Ir(acac)(CO)_2$, which is the higher homologue of $Rh(acac)(CO)_2$, results similar to the reaction between $Rh(acac)(CO)_2$ and P(para)-modified proteins are observed: A new peak with a mass corresponding to a Ir(CO)-fragment is detected (see fig. 17 in the appendix) indicating strong coordination of iridium to the introduced phosphine.

ICP-MS of phosphine based artificial metalloenzymes (excluding rhodium)

After extensive washing of each of these metallo-enzymes using molecular weight cut-off technique to retain both the (un)modified enzyme and the metallo-enzyme the metal content was determined by ICP-MS (see experimental section for details). The ICP-MS data for the reaction between the P(para)-modified mutants SCP-2L V83C and SCP-2L

A100C as well as the unmodified protein SCP-2L A100C and the different metal salts are listed in table 5.

Table 5 ICP-MS and metal loading of different artificial metalloenzymes [a]

entry	metal salt	Protein Scaffold					
		SCP-2L A100C		SCP-2L A100C-1-P(Para)		SCP-2L V83C-1-P(para)	
		ppb	% metal loading[b]	ppb	% metal loading[b]	ppb	% metal loading[b]
1	$Ir(acac)(CO)_2$	23000	75	51000	166	189000	279
2	$Pd(OAc)_2$	1500	9	1200	7	3500	9
3	$Pt(OAc)_2$	1900	6	4700	15	11100	16
4	$Zn(OAc)_2$	-600[c]	-6[c]	500	4	500	2
5	$Mn(OAc)_2$	800	9	-100[c]	-1[c]	400	2
6	$Fe(OAc)_2$	15600	140	15000	168	12700	65
7	$RuCl_3$	100	1	500	3	22000	61
8	$Ru(acac)_3$	700	4	9500	59	200	1
9	$Re(CO)_5Br$	1000	3	1300	5	14000	21
10	$Cu(OAc)_2$	21000	66	3500	7	12000	18
11	$Ni(OAc)_2$	600	6	600	6	900	4
12	$Co(OAc)_2$	200	3	0[c]	0[c]	100	0

a: Reaction conditions for SCP-2L A100C and SCP-2L A100C-1-P(para): To 0.1 µmol (un)modified protein in 1.6 mL buffer 3 eq. of metal salt are added from a stock solution. In case of SCP-2L V83C-1-P(para) : To 0.22 µmol modified protein in 2.2 mL buffer 3 eq. of metal salt are added from a stock solution. After one day the buffer was exchanged, a fifth removed for ESI-MS analysis and the rest analysed by ICP-MS. See experimental section for all details **b**: n(metal) / n(Protein), corrected for the 1/5th removed for ESI-MS analysis; 50% corresponds to 0.5 mol metal per 1 mol (un)modified protein, 100% corresponds to 1 mol metal per 1 mol (un)modified protein and so on. **c**: Probably due to inaccuracies in the ICP-MS determination.

It should be kept in mind that SCP-2L A100C is not a perfect reference as its thiol is likely to coordinate metals, the wild type SCP-2L would have been a better reference. Additionally the reaction conditions were slightly different between the (un)modified SCP-2L A100C and SCP-2L V83C.

The high metal loading determined when Ir(I) was used as metal precursor (entry 1) are quite striking. Regardless of the protein scaffold used its coordination capacity regarding Ir(I) is high even for the unmodified SCP-2L A100C mutant. The metal loading of SCP-2L A100C-1-P(para) is roughly 100% higher than the metal loading between the unmodified SCP-2L A100C and Ir(I). This strongly indicates the high

binding capacity of the introduced phosphine ligand. When SCP-2L V83C-1-P(para) is used almost all iridium initially added is coordinating to the protein scaffold (almost 3 mol iridium per mol protein). These high iridium loadings of SCP-2L V83C-1-P(para) indicate the formation of (at least) one new iridium coordination site upon modification of the mutant. Alternatively the SCP-2L V83C mutant might contain (at least) one iridium coordination site which is not present in the SCP-2L A100C mutant.

The metal loadings are negligible when Zn(II) (entry 4), Mn(II) (entry 5), Ni(II) (entry 11) and Co(II) (entry 12) were used as metal precursors.

When Pd(II) is used as metal precursor (entry 2) the metal loadings detected are quite similar regardless of the scaffold used, indicating that the ion is coordinating to random amino acids of the prtotein.

When Pt(II) is used (entry 3) the metal loading is significantly higher for the Phosphine-modified scaffolds than for the unmodified protein, which is a good indication for coordination of Pt to the introduced phosphine ligand.

When Fe(II) (entry 6) is used the metal loading is only slightly higher for one of the phophine modified scaffolds and even lower for the second in comparison to the unmodified mutant which is a strong indication of Fe(II) coordination to the free thiol of the mutant.

The metal loading for Ru(III) (entries 7 & 8) are unexpected. The interaction of the unmodified protein to either of the Ru(III) precursors used is quite small (<5% metal loading). The metal loading when SCP-2L A100C-1-P(para) was treated with $RuCl_3$ is small as well but when it is treated with $Ru(acac)_3$ a high metal loading is detected. The metal loadings for SCP-2L V83C-1-P(para) treated with the same Ru(III)-salts is exactly the opposite: A high metal loading for $RuCl_3$ and a small metal loading for $Ru(acac)_3$ is detected. Probably the most convenient explanation is a mislabeling of samples as it seems unlikely that the different precursors behave so different regarding of the scaffold used.

When Re(I) (entry 9) is used the metal loading is small for the unmodified protein as well as for SCP-2L A100C-1-P(para). The metal loading is significantly higher for the scaffold SCP-2L V83C-1-P(para).

When Cu(II) (entry 10) is used the metal loading is high for the unmodified protein

while much smaller for the modified protein. This is a strong indication of Cu(II) coordination to the free thiol of the mutant. The Cu(II)-loading is significantly higher for the scaffold SCP-2L V83C-1-P(para) than for SCP-2L A100C-1-P(para).

2.4 Conclusion and future work

In this chapter it could be shown that the crystal structure of SCP-2L A100C is similar to its wild type. The hydrophobic tunnel present within this mutant is by far the largest hydrophobic area. By comparison with the wild type it could be shown that the hydrophobic tunnel of the mutant is slightly larger in diameter at the end where the mutation took place.

Only a small fraction of the artificial metalloenzymes analysed could be detected by ESI-MS indicating that most metals are only weakly coordinating to the protein scaffold. The analysis by ICP-MS was used to further quantify the metal loading of the artificial metalloenzymes synthesised. By combining the information obtained by ESI-MS and ICP-MS some promising new artificial metalloenzymes could be identified.

Several new promising Rh-protein adducts were identified by ESI-MS and showed high rhodium loadings. Phosphine modified SCP-2L V83C was found to be a suitable scaffold for Rh-enzymes regardless of the phosphine introduced. In case of phosphine modified SCP-2L A100C the P(para) modified protein is the most promising scaffold. PYP R52G-1-P(para) was also identified as a promising rhodium-binding scaffold. Additionally iridium was found to form stable complexes with the P(para)-modified SCP-2L A100C mutant (both by ESI-MS and ICP-MS), making it a promising metal complex for catalysis.

The *Phen*-modified SCP-2L A100C mutants were found to be promising scaffolds for Pd(II), Cu(II) and Ni(II)-enzymes as they could be detected by ESI-MS indicating a strong binding and possessed a high metal loading according to ICP-MS as well. *Phen*-modified SCP-2L V83C showed the same promising ESI-MS results for these metals as well. The corresponding metal loadings were not determined but can be expected to be similar. The results for the Co-enzyme of the *Phen* modified SCP-2L mutants tested is ambiguous. While SCP-2L V83C-*Phen*-Co was detected by ESI-MS, SCP-2L A100C-

Phen-Co could not be detected by ESI-MS although this Co-enzyme showed a high metal loading by ICP-MS. The comparison of the metal loading between proteins with different modifications as well as unmodified indicate metal binding to the introduced ligand. *Picol*-modified SCP-2L seems a less viable protein scaffold than the *Phen*-modified SCP-2L for the metals tested. Only the Pd-enzymes of the *picol*-modified SCP-2L mutants were deemed to be promising complexes for catalysis.

Future work should comprise the refinement and analysis of the crystal structure of SCP-2L V83C. Obtaining crystal structures for artificial metalloenzymes would be ideal, but it is probably more viable to aim for crystal structures of *N*-ligand modified proteins. These could be used to model corresponding artificial metalloenzymes and identify amino acids interacting with the metal. This would allow to fine-tune the artificial metalloenzymes by mutagenesis. A similar approach might be feasible for phosphine modified proteins. The crystal structures of the phosphine modified protein or of an artificial metalloenzyme of a phosphine modified protein will not be easy to obtain due to their air sensitivity. Therefore it seems more feasible to further investigate the approach used within this chapter and obtain crystal structures for sulphide (or otherwise) protected phosphines. In case these structures cannot be obtained it might be worthwhile to use the crystal structures of the mutants to determine the structure of the modified enzymes by computational means. Regardless of the approach these structures should then be used to model the corresponding artificial metalloenzymes. This would allow to gain insight into the structural properties next to the introduced ligand and fine-tune this active site in regard to its metal- and substrate coordination.

It might be worthwhile to further investigate the Ir-enzymes. One should investigate them by ^{31}P-NMR to verify the coordination of iridium to the phosphine. They should also be tested in a suitable reaction (like an asymmetric hydrogenation). As the *picol*- and *phen*- modified SCP-2L mutants were found to be promising scaffolds for Pd(II)-enzymes, it might be worthwhile to investigate these Pd-enzymes further and test them in a suitable reaction (for example a coupling reaction). Some of the promising Cu-, Co-, and Ni- enzymes were tested in a Diels-Alder reaction within the scope of this thesis (see chapter 4).

2.5 Experimental

General

Equipment, Chemicals and suppliers

4-(diphenylphosphino)benzaldehyde (P(para)) was synthesised within the Kamer workgroup by Dr. P. Deuss according to a literature known procedure.[30] P(meta) was syntesised within the workgroup by M. Drysdale according to a literature known procedure.[31] PYP R52G-1-P(para) was made available by Dr. W. Laan, following a literature known procedure.[6] The *N*-ligands used for protein conjugation (*Phen*[32,33] and *Picol*[5]) were synthesised by E. de Waard and M. Drysdale according to known literature procedures.

Acetonitrile was distilled over calcium hydride under nitrogen and degassed using three freeze/thaw cycles under argon. Methanol was distilled over Mg/I_2 under nitrogen and subjected to three freeze/thaw cycles under argon. Toluene was distilled over Na under nitrogen and subjected to three freeze/thaw cycles under argon. THF and DMF were subjected to three freeze/thaw cycles under argon before usage.

Programs

Protein images were processed by Pymol 0.99rc6. ESI-MS results were analysed by MassLynx V. 4.0 and its MaxEnt algorithm. General image manipulations (for example the addition of arrows as well as the information about "found mass" and "calculated mass" in MS spectra) were performed by GIMP 2.6.11. Distances between Triton X-100 and the amino acids were determined by UCSF Chimera 1.9.

Methods

Protein concentrations were determined using the Bradford assay.[34] Unless stated otherwise standard Schlenk technique was used. The separation by centrifugation performed after each protein modification step could not be performed under Schlenk conditions. ICP-MS data was either determined by Mrs. Sylvia Williamson at the University of St Andrews or by Dr. Lorna Eades at the University of Edinburgh.

Buffer exchange

Unless stated otherwise centrifugal concentrators with a molecular weight cut-off of 10k Da were used for the buffer exchange. A buffer exchange consists of three consecutive rounds of concentrating the solution from >10 mL to <1 mL and subsequent addition of buffer to a total volume >10 mL again. The concentrators can only be filled up to a maximum volume of 15 mL and solutions within cannot be concentrated to volumes <0.25 mL. A schematic presentation of a centrifugal concentrator is given in figure 5.

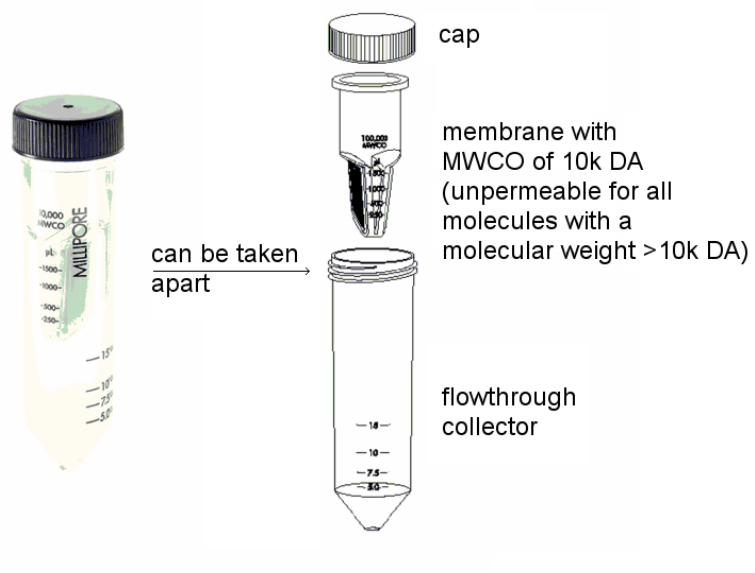


Fig. 5: Schematic representation of a centrifugal concentrator.

Media

The media LB, PB as well as PBS were prepared and sterilised by the technicians within the University of St. Andrews.

Production Broth (PB) medium contains:

20 g/L tryptone, 10 g/L yeast extract, 5 g/L dextrose, 5 g/L NaCl, 8.7 g/L K_2HPO_4, pH=7.0.

Lysogeny Broth (LB) medium contains:

10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl, pH=7.0.

Phosphated buffered saline (PBS) contains:

8.0 g/L NaCl, 0.2 g/L KCl, 1.44 g/L NaH_2PO_4, 0.2 g/L KH_2PO_4, pH=7.3.

The pH of buffers was adjusted using aqueous solutions of sodium hydoxide and / or hydrochloric acid.

Molecular weight and amino acid sequence of relevant proteins

Amino acids in blue are not visible in the crystal structure, mutation highlighted in red.

SCP-2L

MEGGKLQSTF VFEEIGRRLK DIGPEVVKKV NAVFEWHITK GGNIGAKWTI
DLKSGSGKVY QGPAKGAADT TIILSDEDFM EVVLGKLDPQ KAFFSGRLKA
RGNIMLSQKL QMILKDYAKL

M=13244 Da

SCP-2L A100C

GAMEGGKLQS TFVFEEIGRR LKDIGPEVVK KVNAVFEWHI TKGGNIGAKW
TIDLKSGSGK VYQGPAKGAA DTTIILSDED FMEVVLGKLD PQKAFFSGRL
KCRGNIMLSQ KLQMILKDYA KL

M=13405 Da

SCP-2L V83C

GAMEGGKLQS TFVFEEIGRR LKDIGPEVVK KVNAVFEWHI TKGGNIGAKW
TIDLKSGSGK VYQGPAKGAA DTTIILSDED FMEVCLGKLD PQKAFFSGRL
KARGNIMLSQ KLQMILKDYA KL

M=13377 Da

PYP R52G

MRGSHHHHHH GSDDDDKMEH VAFGSEDIEN TLAKMDDGQL DGLAFGAIQL
DGDGNILQYN AAEGDITGGD PKQVIGKNFF KDVAPCTDSP EFYGKFKEGV
ASGNLNTMFE YTFDYQMTPT KVKVHMKKAL SGDSYWVFVK RV

M=15761.4 Da

lysozyme

KVFGRCELAA AMKRHGLDNY RGYSLGNWVC AAKFESNFNT QATNRNTDGS
TDYGILQINS RWWCNDGRTP GSRNLCNIPC SALLSSDITA SVNCAKKIVS
DGNGMNAWVA WRNRCKGTDV QAWIRGCRL

M=14313.1 Da

Crystallisation

Crystallisations of SCP-2L A100C, SCP-2L A100C-1-P(para)=S, SCP-2L V83C and SCP-2L V83C-1-P(para)=S were performed by Dr. Branigan based on the crystallisation procedure for the wild type protein.[3] Crystals were obtained for the unmodified mutants SCP-2L V83C and SCP-2L A100C. Structure determination and refinement was performed by Dr. Branigan. While the refined structure of SCP-2L A100C is discussed in this chapter the structure of SCP-2L V83C was not completely refined.

ESI-MS of artificial metalloenzymes

Metal(X) salts were used as source (with X being a charge of "+1", "+2" or "+3"), therefore it is most likely that the metal coordinating to the protein will be present as metal(X) as well. It can be anticipated that the counterion is dissociating first in the MS conditions (positive ionisation), resulting in a cationic metallo-enzyme. If this assumption is correct the detected mass of all artificial metalloenzymes is X Da lower than the calculated mass, as the program (MassLynx MaxEnt algorithm) assumes the presence of neutral starting species when processing the raw data and calculates as if each positive charge results from "H^+".

The masses mentioned in text and figures only correspond to peak maxima. In general each peak has a width at half height which corresponds to about ±5 Da of this value. Additionally the LC-MS was only calibrated on an irregular basis therefore the obtained masses may deviate by up to 16 Da from the calculated mass. As all conclusions are based on the mass differences between two peaks this deviation is not relevant.

Besides the assigned peaks most spectra show additional peaks which can be assigned to oxidised species, sodium adducts (up to four Na^+ (instead of H^+ were detected) and combinations thereof. These peaks are usually not explicitly mentioned.

LC-MS settings used

LC-MS(ES+) used for analysis of protein and protein reactions was performed on a Waters Alliance HT 2795 equipped with a Micromass LCT-TOF mass spectrometer, using positive electrospray ionisation and applying a Waters MASSPREP® On-line Desalting 2.1x 10 mm cartridge. The software used were MassLynx V4.0 and the

MaxEnt algorithm of Masslynx for processing.

The gradient listed in table 6 was used:

Table 6 Gradient used for LC-MS analysis[a]

Time [min.]	Solvent A [%]	Solvent B [%]	Flow [mL/min]
0.00	2	98	0.050
0.50	2	98	0.050
2.50	98	2	0.050
4.50	98	2	0.050
4.60	2	98	0.050
12.00	2	98	0.050

a: Solvent A = 1% formic acid in acetonitrile, Solvent B = 1% formic acid in water.

The elution of the first 5 min. (containing salts which are potentially harmful to the detector) was automatically discarded. The elution from 5-30 minutes was analysed by ESI+, the (modified) protein usually eluting between minutes 15-25.

MS Tune setting for LC-MS and ESI-MS

The following MS-tune settings were used for both LC-MS and direct injection: Capillary = 3000 V, Sample Cone = 50 V, Extraction Cone = 3 V, RF Lens = 350, Dessolvation Temperature = 300 °C, Source Temperature = 100 °C, RF DC Offset 1 = 10, RF DC Ofset 2 = 8, Aperture = 10, Acceleration = 200, Focus = 0, Steering = 0.0, MCP Detector = 2850, Manual Pusher = Yes, Max. Flight time = 70, Ion Energy = 32, Tube Lens = 0, Grid 2 = 60, TOF flight tube = 4600, Reflection = 1788.

Electrostatic surface potential of SCP-2L A100C and SCP-2L

The electrostatic surface potential was calculated using the plug-in Adaptive Poisson-Boltzmann Solver ("APBS") of pymol. The plug-in is calculating the electrostatic surface by solving the Poisson-Boltzmann equation. Three amino acids of SCP-2L A100C had to be removed for the calculations. The amino acid sequence of SCP-2L A100C is listed below. The amino acids in blue are not visible in the crystal structure, the amino acids highlighted in red had to be removed for the calculation and are therefore not present in the figure showing the electrostatic surface potential. Neither blue nor red amino acids are therefore shown in the figure. The colour code in the

picture is ranging from -6 kT/e (in red) to +6 kT/e (in blue)

GAMEGGKLQS TFVFEEIGRR LKDIGPEVVK KVNAVFEWHI TKGGNIGAKW
TIDLKSGSGK VYQGPAKGAA DTTIILSDED FMEVVLGKLD PQKAFFSGRL
KCRGNIMLSQ KLQMILKDYA KL

Protein expression and purification

A simplified overview of the cell growth, protein expression and lysis is given in figure 6. The expression is performed under sterile conditions.

Fig. 6 Simplified overview of the cell growth, protein expression and lysis.

A small concentration of antibiotic resistant cells harbouring the desired plasmid were grown in the appropriate media. At a certain cell concentration the protein expression is triggered. In a last step the lysate (all proteins within the cells) was obtained.
The lysate containing the SCP-2L mutants was purified according to figure 7.

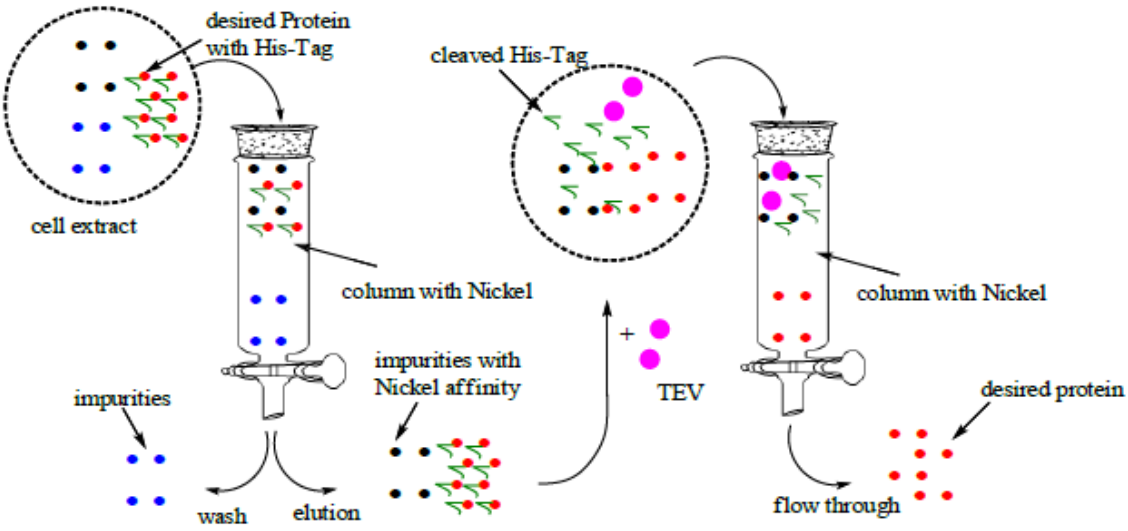


Fig. 7: Protein purification used for the purification of SCP-2L mutants.

The lysate is loaded on a Ni-column and washed. The His-tagged protein and impurities with high nickel affinity stick to the column while all other proteins are washed out. In a next step the His-Tagged protein and the impurities with nickel affinity are eluted and collected. TEV is added to cleave the His-Tag from the protein. The mixture obtained is applied to another column and the flow through which contains the desired protein is collected.

Protein Expression and purification of Tobacco Etch Virus (TEV)

Escherichia coli (*E. Coli*) cells harbouring the plasmid of the desired protein were grown overnight at 37°C in 100 mL Lysogeny Broth (LB) medium containing 50 µg/mL ampicillin. 10 mL of the overnight culture was used to inoculate 500 mL litre of LB medium containing 50 µg/mL ampicillin. The cells were allowed to grow at 37°C and 300 rpm to an optical density (at 600 nm) (OD_{600}) of ≈ 0.6. The temperature was reduced to 25°C and isopropyl-1-thio-β-D-galactopyranoside (IPTG) was added to a final concentration of 1 mM to initiate protein expression. The culture was left overnight (O/N) at 25°C and 300 rpm. The cells were harvested (4°C, 30 min., 7500 g) and stored at – 80°C. Pellets worth 3 L of culture were resuspended in 50 mL lysis

buffer (PBS, 300 mM NaCl, 10 mM imidazole, 1 mg DNase, 50 mg lysozyme and either 2 EDTA free protease inhibitor tablets or 0.5 mM benzamidine). The suspension was left for one hour at 4°C. Then it was sonicated at 4°C in portions of 10 mL with 0.5 second pulses at 90% for 1 min. (Hielscher 4P200S ultrasonic processor). The cell extract obtained after centrifugation (4°C, 50 min., 30000 g) was filtered (0.22 µm syringe filter, Millipore) and applied to a Ni-column (5 ml HisTrap FF columns, GE Healthcare) previously equilibrated (PBS, 300mM NaCl, 30mM imidazole). The elution was collected. The column was washed extensively (~10 column volumes) with the same buffer and the elution discarded. The protein was eluted (PBS, 300mM NaCl, 250mM imidazole) and collected. After equilibrating the column (PBS, 300mM NaCl, 30mM imidazole) the previously collected flow through was reapplied and the procedure repeated until no protein was detected in the elution using the 250 mM imidazole buffer. The buffer was exchanged either by dialysis O/N at 4°C against 4L buffer (50 mM Tris, 300 mM NaCl, pH = 7.5) or by using MWCO (including some glycerol to prevent precipitation). The protein was concentrated (in the presence of some glycerol) to ≈ 4 mg/mL. It was diluted with an equal volume of glycerol and stored at -80°C. Purity was >50% as determined by SDS-PAGE. Average yield ≈ 10 mg/L culture. The activity of the TEV obtained when benzamidine was used was much less than the activity of the TEV obtained when EDTA free protease inhibitor tablets were used (an activity difference of ≈ 0.1-0.5).

Expression and Purification of SCP-2L A100C and SCP-2L V83C

The expression and purification is based on a published procedure.[6]

PB medium with 50 µg/mL kanamycin and 34 µg/mL chloramphenicol (in ethanol) was used for protein production. 10 mL of an overnight culture (in case of SCP-2L A100C ≈ 36 hours were neccessary) of *E. coli* BL21 (Rosetta) cells harbouring the plasmid of the desired protein was used to inoculate 0.5 L of PB medium. The cells were allowed to grow at 37°C to an OD_{600} of 0.6 after which the temperature was lowered to 16°C. IPTG was added after one hour to a final concentration of 0.2 mM from a 0.2 M stock solution to initiate the expression of the protein. The culture was left overnight at 16°C. The cells were harvested by centrifugation (4°C, 20 min., 7500 g), washed with 500 mL of buffer

solution (16 mM K_2HPO_4, 120 mM NaCl, pH = 7.4) and centrifuged (4°C, 20 min., 7500 g). The pellet from 2 L cell culture was resuspended at 4°C in 100 mL of buffer solution (50 mM Tris · HCl, 20 mM imidazole, 150 mM NaCl, 0.5 mM benzamidine, pH = 8) and frozen at -80°C. After defrosting 20 mg of lysozyme, 1 mg of DNase and 1 ml of 1M $MgCl_2$ was added to the suspension. The suspension was left for 1 hour at 4°C and then sonicated (Hielscher 4P200S ultrasonic processor) in portions of 10 ml with 0.5 second pulses for 1 min. at 90%. The cell-extract, obtained after centrifugation (4°C, 50 min., 30000 g) was filtered (0.22 µm syringe filter, Millipore) and applied to a nickel column (5 ml HisTrap FF columns, GE Healthcare) equilibrated with 30 mM Tris · HCl, 20 mM imidazole, 150 mM NaCl, pH = 8 (wash buffer). The elution was collected. The column was washed with 5 column volumes of wash buffer, 5 column volumes of high salt buffer (wash buffer containing 1M NaCl) and another 5 column volumes of wash buffer. The protein was obtained by eluting with 6 column volumes elution buffer (wash buffer containing 330 mM imidazole) in an equal volume of wash buffer (to prevent precipitation of the protein). The previously collected elution was reloaded onto the equilibrated column and the procedure repeated until no protein could be detected when eluting with the 330 mM imidazole buffer. The combined fractions containing the proteins with nickel affinity were dialysed against 4 L buffer solution (30 mM Tris · HCl, 10 mM imidazole, 150 mM NaCl, pH = 8) at 4°C. If precipitate had formed the resulting mixture was centrifuged (30000 g, 50 min.) and the precipitate discarded. 0.014 equivalents of TEV-protease and final concentrations of 1 mM DTT and 0.5 mM EDTA were added to the solution. The mixture was left O/N at RT. The reaction was analysed by SDS-page or LC-MS and upon full cleavage any precipitate was removed by centrifugation (4°C, 50 min., 30000 g) before filtration. The solution was applied to a nickel column equilibrated with wash buffer. The elution containing pure protein was collected, concentrated and the buffer was exchanged to the storage and coupling buffer (degassed 20 mM MES 50 mM NaCl pH 6) using a centrifugal concentrator. Purity was >99% as determined by SDS-PAGE.

ICP-MS

All samples containing Pd, Fe, Pt, Ru, Re, Ir, both blanks as well as the following four samples: i) SCP-2L A100C + Zn, ii) SCP-2L A100C + Cu, iii) SCP-2L A100C-1-P(ortho) + Rh and iv) SCP-2L A100C-1-P(meta) + Rh were determined in Edinburgh by Dr. Eades using a Perkin Elmer Elan 6100 DRC Quad and a Thermo-Finnegan Element 2. Samples were diluted with an internal standard. The following isotopes were determined: ^{55}Mn, ^{56}Fe, ^{57}Fe, ^{59}Co, ^{60}Ni, ^{63}Cu, ^{66}Zn, ^{68}Zn, ^{99}Ru, ^{101}Ru, ^{103}Rh, ^{105}Pd, ^{185}Re, ^{191}Ir, ^{193}Ir, ^{194}Pt, ^{195}Pt. The average of three measurements of each isotope is used. in case two isotopes of the same element were determined their average is used. All other samples were determined by Mrs. Williamson in St Andrews.

All data was corrected for the metal content determined in the corresponding blank. When Protein-*Picol* and Protein-*Phen* were used as scaffold the metal concentrations determined were corrected by the "*N*-Ligand blank" value. When other protein scaffolds were used the concentrations obtained were corrected by the "metal blank" value.

Table 7 List of blank ICP-MS-values

entry	metal	N-ligand blank [ppb][a]	metal blank [ppb][b]
1	Mn	23	10
2	Fe	1083	874
3	Co	378	64
4	Ni	300	113
5	Cu	996	97
6	Zn	956	705
7	Ru		-1
8	Pd	1	3
9	Re		-1
10	Ir		1
11	Pt		0

a: N-ligand blank: 5 mL buffer (20 mM MES, 50 mM NaCl, pH = 6) was subjected to three buffer exchanges using 20 mM NH_4OAc, pH = 6 buffer. 4µL from an EDTA stock solution (0.5 M, pH = 8) was added. The mixture was buffer exchanged (using 20 mM NH_4OAc, pH = 6) and the flow through collected. This was repeated three times. The fractions were combined, concentrated to dryness, topped up with water to a known total volume (1.00 mL) and analysed by ICP-MS on the metals listed. **b**: Concentrated nitric acid and hydrogen peroxide solution (30 wt. %) (1 mL each) were mixed. The solution was carefully evaporated to dryness, topped up with water, evaporated to dryness. It was quantitative transferred to a 1 mL volumetric flask, topped up with water and analysed by ICP-MS on the metals listed.

Modification of lysozyme (from chicken egg white)

a) Using dithiothreitol (DTT) without stirring

141 mg (9.5 µmol) lysozyme and 45 mg (294 µmol, 29 eq) DTT were dissolved in 35 mL degassed buffer and incubated at 4 °C for 70h. The reaction mixture was subjected to a filtration removing any molecule with a molecular weight < 10k Da and 8.7 mg (4 eq.) of the linker maleimidpropionic acid hydrazide hydrochlorid (1) dissolved in a small amount of degassed buffer was added and the reaction incubated over night at room temperature. The solution was analysed by LC-MS and the two species lysozyme and lysozyme$(1)_2$ were detected in a ratio of about 100 : 20.

b) Using DTT with stirring

142 mg (9.5 µmol) lysozyme and 40 mg DTT (259 µmol, 27 eq) were dissolved in 25 mL degassed buffer and the reaction stirred over night. The reaction mixture was subjected to a filtration removing any molecule with a molecular weight < 10k Da and the protein solution filtered to remove precipitated protein. To this solution 33.8 mg (154 µmol, 16 eq) of the linker maleimidpropionic acid hydrazide hydrochloride (1) dissolved in a small amount of degassed buffer was added and the solution stirred over night. The solution was analysed by LC-MS and three species detected which correspond to lysozyme, lysozyme-1 and lysozyme-$(1)_2$ in a ratio of about 70:100:20. After stirring for another night and analysis by LC-MS the amount of modified lysozyme increased and lysozyme$(1)_3$ could be detected as well. Unfortunately a significant peak with a mass smaller than unmodified lysozyme could be detected as well, indicating the degradation of lysozyme.

c) Using Tris(2-carboxy-ethyl)phosphine hydrochloride (TCEP)

A solution of 61 mg (4.3 µmol) lysozyme and 32 mg (112 µmol, 26 eq) in 40 mL degassed buffer was incubated over night. The reaction mixture was subjected to a filtration removing any molecule with a molecular weight < 10k Da and 120 mg (546 µmol, 127 eq) of the linker maleimidpropionic acid hydrazide hydrochlorid (1) dissolved in a small amount of degassed buffer was added and the reaction incubated over night at room temperature. The solution was analysed by LC-MS and four species detected which correspond to lysozyme, lysozyme-$(1)_2$, lysozyme-$(1)_3$ and lysozyme-$(1)_4$ in a ratio of about 5:100:30:10. (Referred to as lysozyme-$(1)_x$). The buffer was

exchanged to remove any excess of TCEP prior to subsequent modifications (see fig. 18 in the appendix for the ESI-MS of lysozyme-$(1)_x$.

***N*-Ligand modification of proteins**

Protein modification of SCP-2L A100C and SCP-2L V83C with *Phen* or *Picol* were performed following a known procedure[5]: In a typical experiment 2 µmol protein in 40 mL buffer (20 mM MES, 50 mM NaCl, pH = 6) 7.2 mg (26 µmol, 13 eq) of the *N*-ligand derivative dissolved in a small volume of DMF was added. The reaction mixture was shaken gently and left at room temperature over night. When LC-MS showed full conversion to the modified protein, the filtered reaction mixture was buffer exchanged to remove the excess of the free *N*-ligand derivative. It was filtered to remove precipitated protein. On average 1.4 µmol (70%) of the modified protein was obtained.

Linker modification of protein (Protein + 1)

Protein modification of SCP-2L A100C, SCP-2L V83C with the linker "1" (maleimidopropionic acid hydrazide) were performed based on a known procedure.[6]

Modification of linker modified protein with phosphine [P(para), P(meta) or P(ortho)]

Protein modification of linker modified SCP-2L A100C, SCP-2L V83C and ysozyme with the phosphine were performed following a literature known procedure.[6]

As the modification of lysozyme-$(1)_x$ was virtually unsuccesfull using these standard conditions higher equivalents of phosphine (up to 80 eq.) as well as longer reaction times (up to eight days) were tested. The phosphine modified protein was only obtained in small quantities, when P(meta) was used the modification was virtually unsuccesful. The ESI-MS of the modification with P(para) and P(ortho) is depicted in the appendix (fig. 19 and fig. 20). Phosphine modified species were not detected under the analysis condtions but the corresponding phosphine oxides, indicating that the phosphine modified protein is much more oxygen sensitive than the SCP-2L and PYP based phosphines and readily oxidised under the analysis conditions.

***N*-Ligand modified protein + M[OAc]$_2$**

The following acetate salts were used: Zn(II), Cu(II), Mn(II), Pd(II), Fe(II), Co(II), Ni(II). (See appendix for purity and supplier)

This procedure was used for SCP-2L A100C-*Picol*, SCP-2L A100C-*Phen*, SCP-2L V83C-*Picol*, SCP-2L V83C-*Phen*. All possible combinations were sythesised.

DMF stock solutions of Pd(II) and Cu(II) were used. For the other metals stock solutions in buffer (20 mM MES, 50 mM NaCl, pH = 6) were prepared. All metal stocks had a concentration of about 25 µmol / mL.

To 0.1 µmol modified protein in 5 mL (c = 0.02 µmol/mL) buffer (20 mM MES, 50 mM NaCl, pH = 6) 2 eq. of metal salt was added from a stock. The mixture was occasionally shaken gently over the course of 1 day. The buffer was exchanged to 20 mM NH_4OAc, pH = 6 and a fifth analysed by ESI-MS.

4µL from an EDTA stock solution (0.5 M, pH = 8) was added. The mixture was buffer exchanged (using 20 mM NH_4OAc, pH = 6) and the flow through collected. This was repeated three times. The fractions were combined, concentrated to dryness, topped up with water to a known total volume (1.00 mL) and the metal content determined by ICP-MS. In case the ESI-MS previously contained an unidentified peak or a peak corresponding to the artificial metalloenzyme the supernatant after EDTA-treatment was analysed by ESI-MS again.

Unmodified protein (SCP-2L A100C & SCP-2L V83C) + M[OAc]$_2$

DMF stock solutions of Pd(II) and Cu(II) were used. For the other metals stock solutions in buffer (20 mM MES, 50 mM NaCl, pH = 6) were prepared. All metal stock solutions had a concentration of about 25 µmol / mL.

This procedure was used for both SCP-2L V83C and SCP-2L A100C. All possible combinations were prepared and analysed by ESI-MS. All metal complexes of SCP-2L A100C were analysed by ICP-MS as well.

To 0.1 µmol unmodified protein in 1.6 mL (c = 0.063 µmol/mL) buffer (20 mM MES, 50 mM NaCl, pH = 6) 3 eq. of metal salt was added from a stock solution. The mixture was occasionally shaken gently over the course of 1 day. The buffer was exchanged to 20 mM NH_4OAc, pH = 6 and a fifth analysed by ESI-MS.

4µL from an EDTA stock solution (0.5 M, pH = 8) was added. The mixture was buffer exchanged (using 20 mM NH_4OAc, pH = 6) and the flow through collected. This was repeated three times. The fractions were combined, concentrated to dryness, topped up with water to a known total volume (1.00 mL) and the metal content determined by ICP-MS.

SCP-2L A100C-1-P(para) + metal[X]

The following metal stock solutions were used (about 25 µmol/mL):

In degassed DMF: $Ir(acac)(CO)_2$**,** $Rh(acac)(CO)_2$, $Re(CO)_5Br$**,** $Ru(acac)_3$**,** $Pt(acac)_2$, $Pd(OAc)_2$

In degassed MeCN: $RuCl_3$, $[Rh(MeCN)_2COD]\ BF_4$, $[Rh(PPh_3)_3CO]H$

In degassed buffer (20 mM MES, 50 mM NaCl, pH = 6): $Mn(OAc)_2$, $Ni(OAc)_2$, $Zn(OAc)_2$, $Cu(OAc)_2$, $Fe(OAc)_2$, $Co(OAc)_2$.

The combinations mentioned in the text were prepared and analysed by LC-MS and ICP-MS.

The two Rh-salts $[Rh(MeCN)_2COD]\ BF_4$ and $[Rh(PPh_3)_3CO]H$ were only tested on SCP-2L A100C-1-P(para) using 0.07 µmol/mL modified protein and 3 eq of metal salt and subsequent analysis.

SCP-2L A100C-1-P(para): To 0.1 µmol modified protein in 1.6 mL (c = 0.063 µmol/mL) buffer (20 mM MES, 50 mM NaCl, pH = 6) 3 eq. of metal salt was added from a stock solution.

SCP-2L V83C-1-P(para): To 0.22 µmol modified protein in 2.2 mL (c = 0.1 µmol/mL) buffer (20 mM MES, 50 mM NaCl, pH = 6) 3 eq. of metal salt was added from a stock solution.

The mixture was occasionally shaken gently over the course of 1 day. The buffer was exchanged (3 times from >10 mL to <1 mL). A fifth of the concentrated solution was analysed by LC-MS.

The rest was carefully treated with concentrated nitric acid and hydrogen peroxide solution (30 wt. %) (1 mL each). The solution was carefully evaporated to dryness, topped up with water, evaporated to dryness. It was quantitative transferred to a 1 mL volumetric flask and topped up and analysed by ICP-MS.

In case of $Rh(acac)(CO)_2$ the following experimental settings were used: To 0.064 µmol/mL phosphine modified protein in 32 mL degassed buffer (2.06 µmol) 25 µL of a 9.8 µmol/mL $Rh(acac)(CO)_2$ stock solution in DMF was added (0.245 µmol). After buffer exchange the solution was concentrated to 3 mL. 0.5 mL of which were treated with nitric acid and hydrogen peroxide solution (30 wt. %) (1 mL each) (as above) and analysed by ICP-MS.

PYP R52G-1-P(para) + Rh

To 20 mL of a solution (20 mM MES 50 mM NaCl pH 7) of PYP R52G-1-P(para) (3.0 µmol) $Rh(acac)(CO)_2$ (1.5 µmol,0.5 eq) was added from a 3.9 mM stock solution in degassed DMF under argon. The pink solution turned quickly yellow and was stirred for three hours at room temperature at 200 rpm. No precipitate formed. The reaction mixture was washed three times with the same degassed buffer using MWCO equipment to remove uncomplexed rhodium and concentrated before LC-MS analysis.

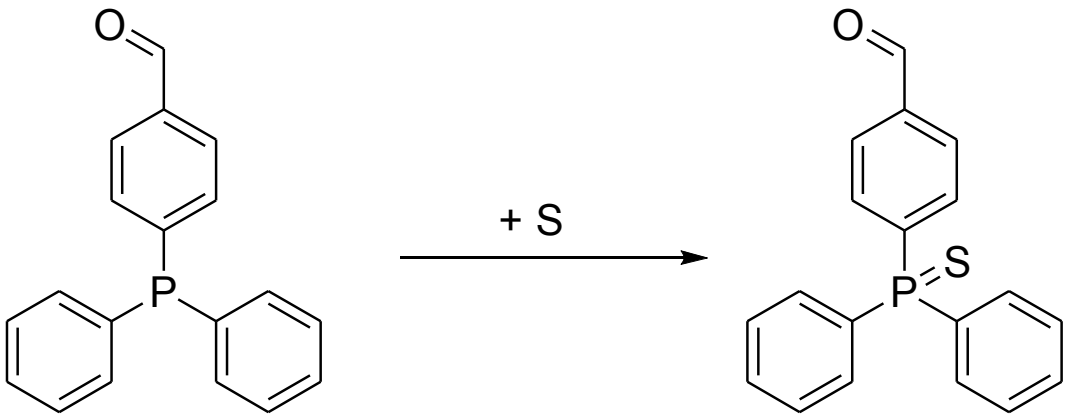
Rh-enzymes based on lysozyme

This procedure was used for lysozyme-$(1)_x$, lysozyme-$(1)_x$-P(para), lysozyme-$(1)_x$-P(ortho) and lysozyme-$(1)_x$-P(meta):

To 3.8 µmol modified protein in 30 mL degasses buffer (20 mM MES, 50 mM NacL, pH=6) 3.8 µmol $Rh(acac)(CO)_2$ dissolved in a small volume of degassed MeCN was added. The mixture turned yellow, was gently shaken and left over night. After buffer exchange the mixture was analysed by LC-MS and buffer exchanged.

Synthesis of 4-(diphenylphosphinsulphide)benzaldehyde [P(para)=S]

The synthesis of P(para)=S is shown in scheme 3.



Scheme 3: Synthesis of 4-(diphenylphosphinsulphide)benzaldehyde [P(para)=S]

3.0 mg (10.3 µmol) of the phosphine was dissolved in 1.0 mL dry DCM. 9.32 mg (291.3 µmol, 28 eq) of elemental sulphur was added and the reaction stirred at RT. The reaction was monitored by TLC and upon completion of the reaction it was used without further purification.

Synthesis of SCP-2L A100C-1-P(para)=S and SCP-2L V83C-1-P(para)=S

Protein modification of linker modified SCP-2L A100C and SCP-2L V83C were performed following a literature known procedure but using P(para)=S instead of P(para).[6] Upon completion (reaction followed by LC-MS) the excess of phosphine ligand (and excess sulphur still present) was removed by buffer exchange. The conversion of SCP-2L A100C-1 to the product was successful (see fig. 8).

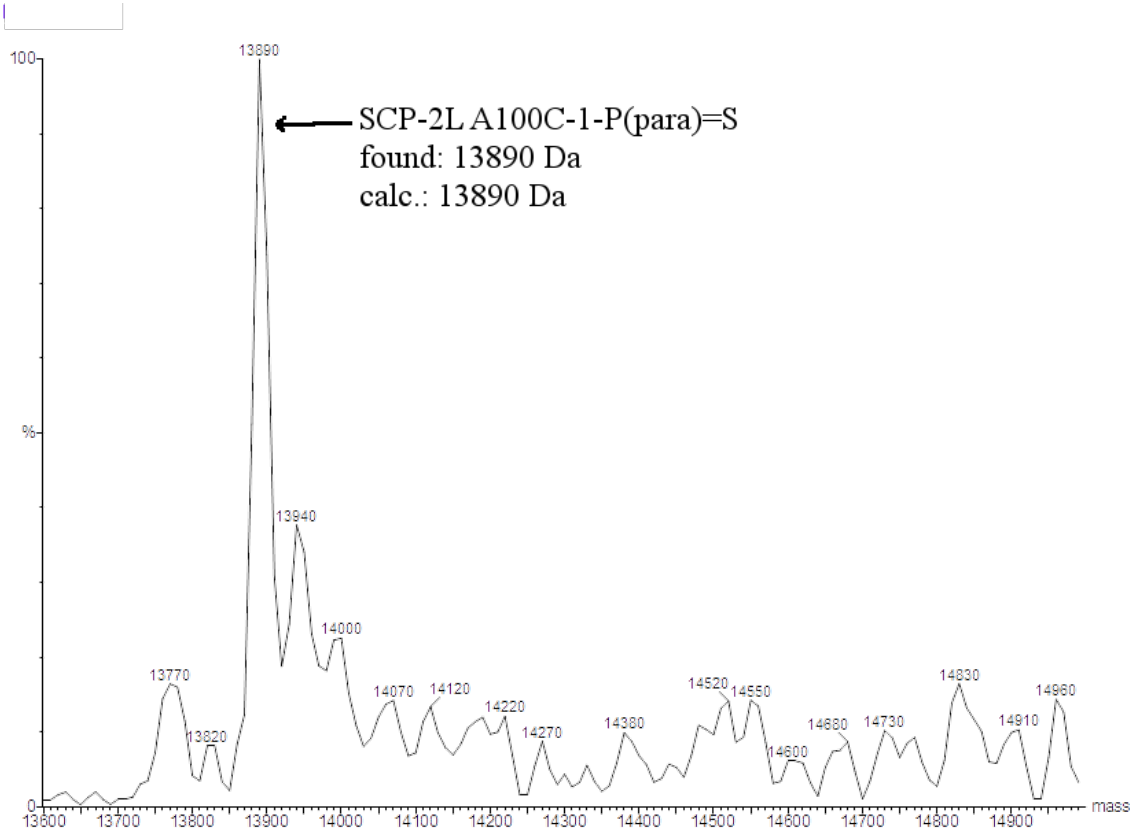


Fig. 8 ESI-MS of SCP-2L A100C-1-P(para)=S.

The conversion of SCP-2L V83C-1 with P(para)=S to the respective product was less successful. Besides the species with the expected mass several additional unassigned species with higher masses could be detected as well (see fig. 9). As the crystallisation of SCP-2L A100C-1-P(para)=S was unsuccessful no further attempt to synthesise this SCP-2L V83C-1-P(para)=S was tried.

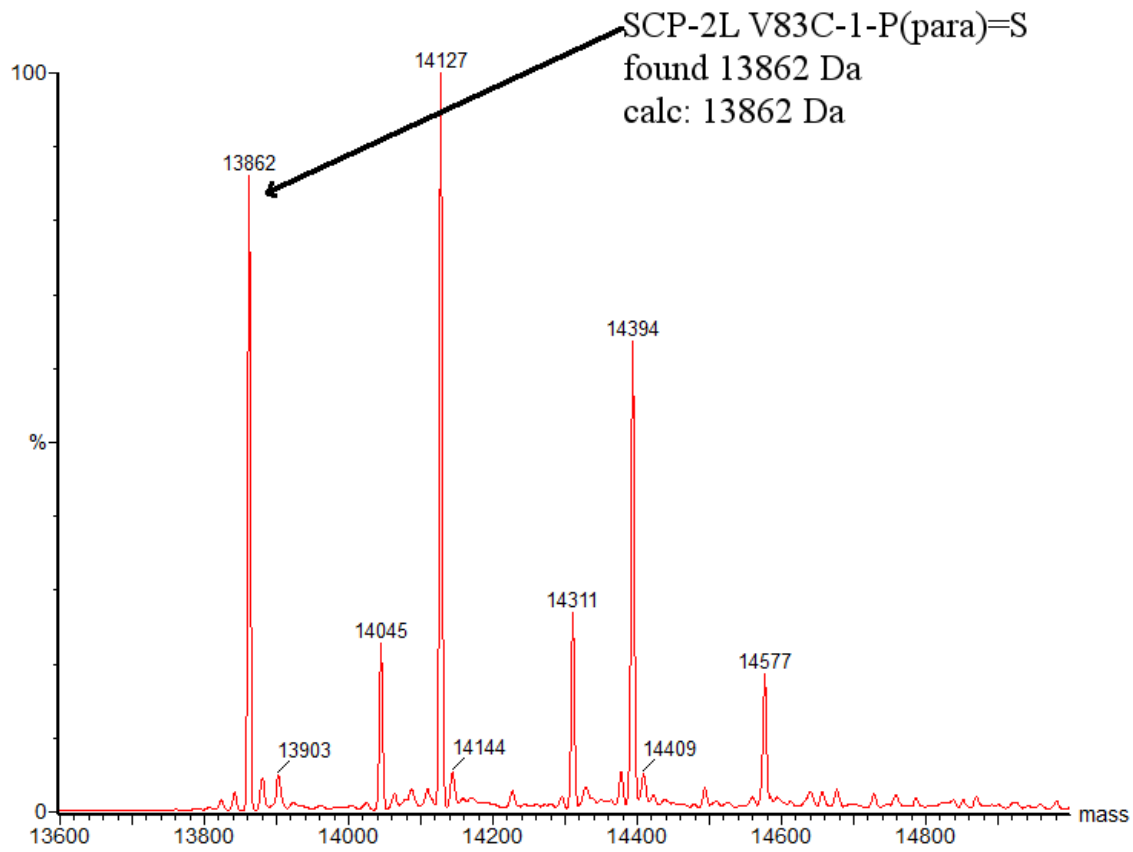


Fig. 9 ESI-MS of SCP-2L V83C-1-P(para)=S.

Distances between Triton X-100 and SCP-2L A100C / SCP-2L

The distances were derived with Chimera and are listed in the following table:

Table 8 Part 1 of 2: Distances between Triton X-100 and the wild type SCP-2L & distances between Triton X-100 and the mutant SCP-2L A100C.[a]

Triton X-100[b]	AA residue number, atom[c]	Distance wt[d]	Distance SCP-2L A100C[e]
C2	GLN 111, CG	**3.855**	4.210
C2	LYS 115, NZ	**3.990**	4.924
C3	LEU 114, C	4.389	**3.990**
C3	LEU 114, CB	4.009	3.812
C3	LYS 115, N	4.244	**3.716**
C4	VAL 83, CA	3.978	4.177
C4	VAL 83, CG2	3.990	3.979
C7	TRP 36, CH2	3.983	3.737
C7	TRP 36, CZ2	3.981	3.730
C7	LEU 110, CD2	4.369	**3.780**
C7	LEU 114, CD1	**3.675**	4.362
C8	PHE 79, CE2	3.714	3.691
C8	PHE 79, CZ	3.749	3.652
C8	VAL 83, CG2	3.850	3.893
C9	TRP 36, CH2	3.609	3.533
C10	TRP 36, CH2	3.692	3.870
C10	LEU 110, CD2	4.207	3.963
C11	TRP 36, CH2	**3.887**	4.230
C11	SER 107, O	**3.462**	3.775
C11	GLN 111, CB	**3.904**	4.242
C12	TRP 36, CH2	3.952	4.237
C13	TRP 36, CH2	3.902	3.916
C13	VAL 82, CG1	3.909	3.704
C14	TRP 36, CH2	3.706	3.550
C14	TRP 36, CZ3	4.045	3.965
C14	VAL 82, CG1	3.763	3.742
O15	PRO 89, CB	3.846	3.699
C16	PRO 89, CB	3.683	4.458
C16	GLN 90, NE2	**3.772**	4.721
C16	SER 107, C	4.180	**3.816**
C16	SER 107, CB	4.397	**3.639**
C16	SER 107, O	3.769	3.477
C16	SER 107, OG	4.397	**3.993**
C16	GLN 111, OE1	**3.181**	4.711

a: Distances between Triton X-100 and the wild type SCP-2L as well as the distance between Triton X-100 and SCP-2L A100C were determined using Chimera.[35] The crystal structure of SCP-2L was published before.[3] Distances between either of the protein atoms and its Triton X-100 atoms ≤4.0 Å were taken into account. **Bold values**: Distance between the same atoms in the mutant (or wild type) are 0.3 Å further away. Red values Distance is >4.0 Å. **e**: Distances to the SCP-2L A100C mutant in Å. **b**: Referring to the labelling scheme of Triton X-100 introduced in fig. 4. **c**: amino acid (AA) as well as the atom of the amino acid (using pdb file format nomenclature) are given. **d**: Distance between Triton X-100 to the wild type (wt) protein in Å. **e**: Distances between Triton X-100 to the mutant in Å.

Table 8 Part 2 of 2: Distances between Triton X-100 and the wild type SCP-2L & distances between Triton X-100 and the mutant SCP-2L A100C.[a]

Triton X-100[b]	AA residue number, atom[c]	Distance wt[d]	Distance SCP-2L A100C[e]
C17	GLN 90, NE2	**3.623**	4.628
C17	LEU 98, CD2	5.167	**3.736**
C17	SER 107, CB	3.616	3.727
C17	SER 107, OG	3.205	**3.552**
C17	SER 107, O	**3.786**	4.568
C17	SER 107, C	**3.724**	4.530
C17	GLN 108, N	**3.836**	4.924
C17	GLN 111, OE1	**3.841**	5.418
O18	GLN 90, NE2	**3.241**	3.572
O18	LEU 98, CD2	4.464	**3.903**
O18	SER 107, OG	3.941	4.216
C19	GLN 90, NE2	4.057	3.973
C19	SER 107, OG	3.525	3.595
C20	GLN 90, NE2	3.937	**3.393**
C20	PHE 93, CB	**3.583**	4.310
C20	PHE 93, CG	**3.569**	4.367
C20	PHE 93, CD2	**3.970**	4.277
C20	PHE 93, CD1	**3.986**	5.128
O21	GLN 90, NE2	**3.612**	4.024
O21	GLN 90, CG	**3.940**	4.522
O21	GLN 108, CG	4.772	**3.802**
O21	GLN 108, NE2	4.163	3.965
O21	GLN 108, OE1	**3.795**	4.616
C22	PHE 93, CD2	3.961	4.218
C22	GLN 108, OE1	**3.992**	4.686
C23	MET 105, CE	4.050	**3.635**
C23	MET 105, CD	**3.537**	3.964
C23	GLN 108, CD	**3.666**	4.413
C23	GLN 108, NE2	3.815	3.591
C23	GLN 108, OE1	**2.954**	3.845
O24	GLN 90, CG	**3.113**	3.871
O24	PHE 94, CE2	**3.932**	4.369
O24	GLN 108, CD	**3.405**	4.343
O24	GLN 108, NE2	3.078	3.342
O24	GLN 108, OE1	**3.027**	3.416
C25	GLN 90, CG	3.963	4.100
C25	GLN 108, CD	**3.732**	4.380
C25	GLN 108, NE2	3.413	3.457
C25	GLN 108, OE1	3.226	3.355

See legend on part 1 of the table

Calculation of the equilibrium constants

The steps involved from the addition of a metal salt to the determination of the metal loading is shown in figure 10.

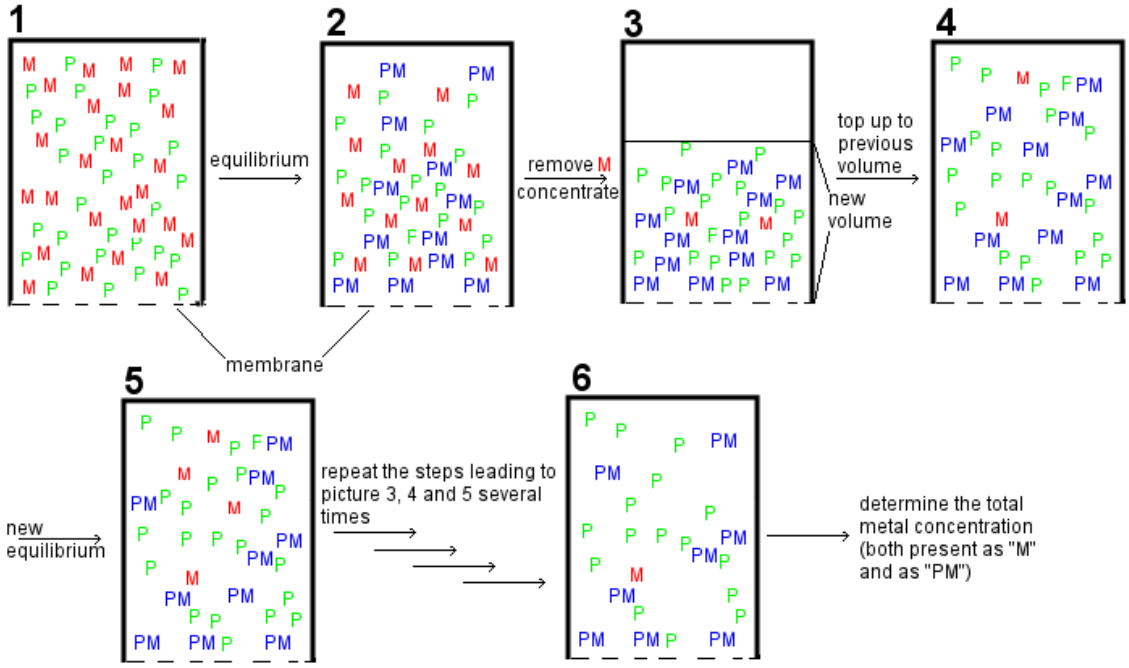


Fig. 10 Schematic overview from the addition of a metal to a protein up to the determination of the metal loading. "P"=(un) modified protein, "M"=metal salt, PM=protein-metal complex

In a certain volume protein ("P") and a metal salt ("M") are present (picture 1 in fig. 10). "P" and "M" are in equilibrium with the protein-metal complex "PM" After some time the equilibrium is reached (picture 2). Now the buffer is exchanged (pictures 3, 4 & 5). This means in a first step the volume is decreased. While metal "M" can pass through the membrane neither protein "P" nor the complex "PM" can pass through. Therefore the new volume has a higher concentration of "P" and "PM" while the concentration of "M" is constant (picture 3). The equilibrium is unaffected by this step as the concentrations of "P" and "PM" are increased by the same factor. In the next step the volume is increased by the addition of new buffer (picture 4 in fig. 10). As a consequence of the dilution a new equilibrium is obtained. The concentration and dilution steps are now repeated several times. The equilibrium obtained in the end is

analysed on its metal concentration - regardless if the metal is present as "M" or as "PM" (to be more precise only 4/5th was analysed on its metal content while 1/5th was analysed by ESI-MS). In the work carried out in this thesis the following information is known:

- The initial concentrations of "P" and "M" as well as the initial volume (initial concentration of "PM" is known but zero)
- the volume after concentrating it
- the volume after topping it up
- how many times the concentration / topping up was repeated
- the sum of the final concentrations of "M" and "PM"

In theory this information can be used to calculate the equilibrium constant under the following assumptions:

1. Total protein concentration stays constant, which is a simplification, as protein precipitates over time.
2. Only 1:1 metal-protein complexes are taking into account.
3. The initial equilibrium between "P" and "M" was reached after the reaction time of 1 day.
4. The temperature was considered as constant, which is a simplification as the ultracentrifugation took place in a cold room (4 °C), while all other steps took place at room temperature.
5. No "P" or "PM" is passing through the filter during the centrifugation.

The reaction between (modified) protein and metal can be written as:

For the reaction above the law of mass action can be written as:

Eq. (1): $$K = \frac{\bar{c}(PM)}{\bar{c}(P) \times \bar{c}(M)}$$

$\bar{c}(X)$	= equilibrium concentration of X
K	= complexation constant [L/mol]

When the following abbreviations

Eq. (2): $c_0(X) = \text{initial concentration of X}$

Eq. (3): $\bar{c}(P) = c_0(P) - \bar{c}(PM)$

Eq. (4): $\bar{c}(M) = c_0(M) - \bar{c}(PM)$

are used Eq. (1) can be rewritten as

Eq. (5): $$K = \frac{\bar{c}(PM)}{[c_0(P) - \bar{c}(PM)] \times [c_0(M) - \bar{c}(PM)]}$$

calculating the reciprocal value of equation 5 and subsequent multiplication by $\bar{c}(PM)$ results in equation 6:

Eq. (6): $$(c_0(P) - \bar{c}(PM)) \times (c_0(M) - \bar{c}(PM)) = \frac{\bar{c}(PM)}{K}$$

expandation of the left and subsequent substraction of the right hand side of the equation results in equation 7:

Eq. (7):

$$(c_0(P) \times c_0(M)) - (c_0(P) \times \bar{c}(PM)) - \bar{c}(PM) \times c_0(M) + (\bar{c}(PM))^2 - \frac{\bar{c}(PM)}{K} = 0$$

Equation (7) can be rewritten as:

Eq. (8): $$(\bar{c}(PM))^2 - \bar{c}(PM) \times ((c_0(P) + c_0(M) + \frac{1}{K}) + c_0(P) \times c_0(M) = 0$$

As the general solution for a quadratic equation of the form

Eq. (9): $$A^2 + Ax + c = 0$$

for unknown A is

Eq. (10): $$A = \pm\sqrt{\frac{1}{4}x^2 - c} + \frac{x}{2}$$

The general solution of Eq. (10) is used for Eq. (8) to obtain Eq. (11).

Eq. (11): $$\bar{c}(PM) = \pm\sqrt{\frac{1}{4}x^2 - c} + \frac{x}{2}$$

"x" in Eq. (11) is defined as:

Eq. (12): $$(c_0(P) + c_0(M) + \frac{1}{K}) = x$$

and c in Eq. (11) is defined as:

Eq. (13): $$c_0(P) \times c_0(M) = c$$

Combining Eq. (11), Eq. (12) and Eq (13) results in equation 14:

Eq. (14):

$$\bar{c}(PM) = \pm\sqrt{\frac{1}{4} \times (c_0(P) + c_0(M) + \frac{1}{K})^2 - c_0(P) \times c_0(M)} + \frac{1}{2} \times (c_0(P) + c_0(M) + \frac{1}{K})$$

The negative square root of Eq. (14) is giving the chemical meaningful results:

Eq. (15):

$$\bar{c}(PM) = -\sqrt{\frac{1}{4} \times (c_0(P) + c_0(M) + \frac{1}{K})^2 - c_0(P) \times c_0(M)} + \frac{1}{2} \times (c_0(P) + c_0(M) + \frac{1}{K})$$

Eq. (15) can be used to calculate the equilibration concentration of "PM". The equilibrium concentrations of "P" and "M" can be calculated afterwards using Eq. (3) and Eq. (4) respectively. These equilibrium concentrations correspond to picture 2 in figure 10.

In the subsequent buffer exchange "M" is removed from the reaction. There are two extreme cases for the equilibrium during the buffer exchange (picture 5 in fig. 10):
a) The equilibration of the reaction is much slower than the buffer exchange
b) The equilibration of the reaction is much faster than the buffer exchange

Additionally there are two extremes for the buffer exchange as the volume removed and added was within a certain range. For a buffer exchange the reaction mixture is concentrated from "X" mL to "Y" mL. Within this thesis "X" was a value between 10-15 and "Y" a value between 0.25-1.
The calculations for a slow equilibration (case a) is much more straightforward:
After calculation of the equilibration concentrations of "P", "M" and "PM" using Eq. (3), (4) and (15) [obviously starting with Eq. (15)!] the calculated equilibrium concentration of "M" is reduced due to the buffer exchange. $1/10^{th}$ to $1/60^{th}$ of "M" is removed per buffer exchange and three consecutive rounds of buffer exchange are performed, therefore $(1/10)^3$ to $(1/60)^3$ of the equilibrium concentration of "M" is remaining after the buffer exchange.

Eq. (16): $$c(M_{remaining}) = \frac{1}{60^3} \text{ to } \frac{1}{10^3} \times \bar{c}(M)$$

Additionally the final volume of the reaction mixture is different (0.25 mL in this example, see experimental section for details) from the initital volume (5.0 mL in this example, see experimental section for details). Therefore the previously calculated equilibration concentrations of "P" (Eq. 3), "PM" (Eq. 15), as well as the concentration of "$M_{remaining}$" (Eq. 16) need to be adapted by multiplication of the factor 20 (=5.0 mL/0.25 mL) to obtain the new theoretical concentrations:

Eq. (17): $$c(P_{Theory}) = 20 \times \bar{c}(P)$$
Eq. (18): $$c(PM_{Theory}) = 20 \times \bar{c}(PM)$$
Eq. (19): $$c(M_{Theory}) = 20 \times c(M_{remaining})$$

These concentrations are used to calculate the new equilibrium concentrations. Eq. (15)

can not be used directly, as it is only suitable if no "PM" is present. Therefore Eq. (20) and Eq. (21) are used to define new initial concentrations of "P" and "M" (just a mathematical trick, calculating as if a total back reaction of "PM" to "P" and "M" is happening so Eq. (15) can be used again)

Eq. (20): $$c_{0,\,new}(M) = c(M_{Theory}) + c(PM_{Theory})$$

Eq. (21): $$c_{0,\,new}(P) = c(P_{Theory}) + c(PM_{Theory})$$

With these initial concentrations Eq. (15) can be used to determine the new equilibrium concentration of "PM" as the initial concentration of "PM" is now (mathematically) zero.

The so calculated equilibrium concentration of "PM" can be used to calculate the equilibrium concentrations of "M" and "P" when the initial concentrations calculated from Eq. (20) and Eq. (21) are inserted into Eq. (4) and Eq. (5).

When the equation is solved for different K values (under certain reaction conditions) the corresponding concentrations of "P", "M" and "PM" in the final equilibrium (picture 6 in fig. 10) are obtained. These equilibrium concentrations can be depicted in a figure (see figure 11).

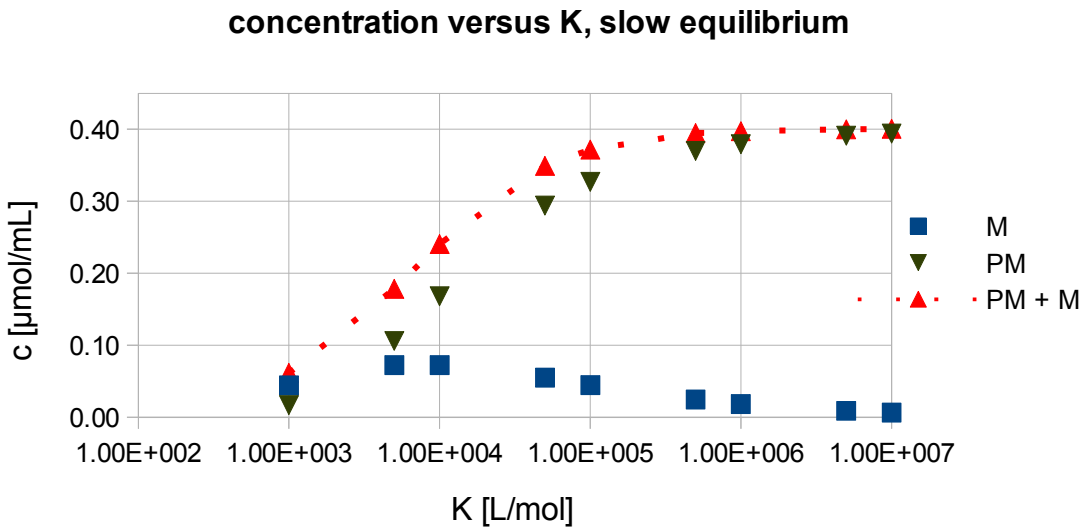


Fig. 11 Concentration over K for a slow equilibrium. Same conditions as used for A100C + Metal and A100C-1-P(para) + Metal reactions: V(initial) = 1.6 mL, V(final) = 0.25 mL, c_0 (P) = 0.0625 µmol/mL , c_0(M) = 0.1875 µmol/mL, three rounds of buffer exchange: Topped up to 10 mL, concentrated to 1 mL.

In case the re-equilibration of the reaction is much faster than the buffer exchange new equilibrium concentrations need to be calculated after every dilution and after every single buffer exchange, reducing the total metal content not just once as described above but three times. The concept and formulas described above remain the same though but all calculations need to be repeated three times. As a result a figure as depicted in figure 12 can be obtained.

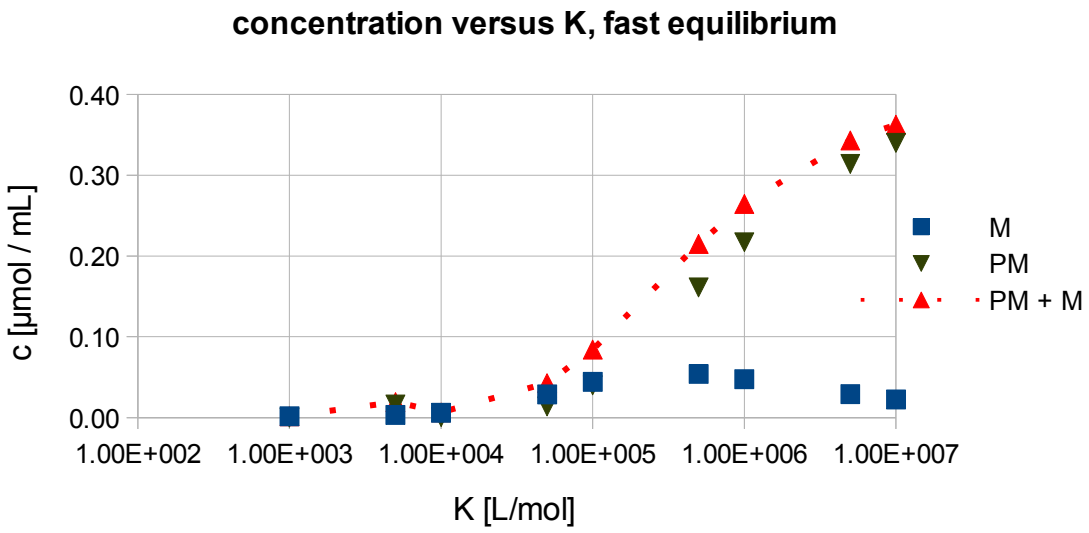


Fig. 12 Concentration over K for a fast equilibrium. Same conditions as used for A100C + Metal and A100C-1-P(para) + Metal: V(initial) = 1.6 mL, V(final) = 0.25 mL, c_0 (P) = 0.0625 µmol/mL , c_0(M) = 0.1875 µmol/mL, three rounds of buffer exchange: Topped up to 10 mL, concentrated to 1 mL.

The total metal content is determined by ICP-MS, due to different molecular weights these values are used to calculate a more general value - the "% metal" which is determined as

Eq. (22): $$\% \text{ metal} = \frac{\text{ppb metal determined by ICP}}{\text{ppb protein present}}$$

when an excess of metal to protein is added.

When substochiometric amounts of metal (to protein) are added it is defined as:

Eq. (23): $$\% \text{ metal} = \frac{\text{ppb metal determined by ICP}}{\text{ppb metal added}}$$

Equilibrium constants for different % metal values were calculated for all four cases:

- A fast equilibration with the minimal buffer exchange (from 10 mL to 1 mL)
- A fast equilibration with the maximal buffer exchange (from 15 mL to 0.25 mL)

- A slow equilibration with the minimal buffer exchange (from 10 to 1 mL)
- A slow equilibration with the maximal buffer exchange (from 15 mL to 0.25 mL).

The difference between the minimal and the maximal buffer exchange was found to be negligible. Different K-values can now be used to calculate the corresponding % metal values for all four cases. For a certain reaction setup figure 13 is obtained.

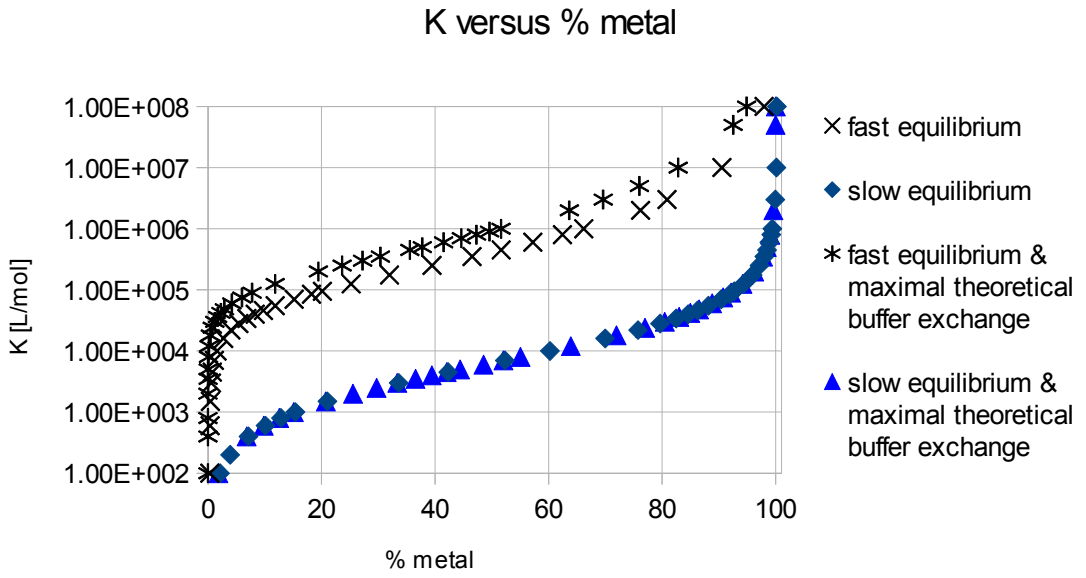


Fig. 13 Calculated values for K versus % metal. Same conditions as used for A100C + Metal and A100C-1-P(para) + Metal: V(initial) = 1.6 mL, V(final) = 0.25 mL, c_0 (P) = 0.0625 µmol/mL , c_0(M) = 0.1875 µmol/mL, three rounds of buffer exchange. for four different cases: The equilibration of the reaction is a) much faster than the buffer exchange calculating with a buffer exchange from 10 mL to 1 mL each time (the theoretical minimal), represented by "X" b) much faster than the buffer exchange calculating with a buffer exchange from 15 mL to 0.25 mL (the theoretical maximum), represented by a star c) much slower than the buffer exchange calculating with a buffer exchange from 10 mL to 1 mL each time (the theoretical minimal), represented by a square d) much slower than the buffer exchange calculating with a buffer exchange from 15 mL to 0.25 mL (the theoretical maximum), represented by a triangle.

The % metal values calculated from the ICP-MS concentrations detected can be used to derive the corresponding equilibrium constant for the complex formation between "P" and "M". For example a % metal value of 60 would be (roughly) in between 10^4-10^6 [$Lmol^{-1}$]

These calculations were performed for the ICP-MS data of the *N*-ligand modified proteins of SCP-2L A100C. The (range) of the equilibrium constants are listed in table 9. The equilibrium constants between (most) of the metals and 1,10-Phenanthroline as well as di-(2-picolyl)amine are literature known. These equilibrium constants are listed as well.

Table 9 calculated log K-data for *N*-modified SCP-2L A100C and different metals as well as the literature known log K data between *N*-ligands and different metals.[a]

entry	metal	SCP-2L A100C-*Phen*[b]	*Phen*[c]	SCP-2L A100C-*Picol*[b]	*Picol*[c]
1	Pd(II)	3.9-5.4	-[d]	5.8-7.8	-[d]
2	Zn(II)	3.4-5.1	6.55	3.5-5.1	7.63
3	Mn(II)	3.2-4.9	4.13	3.6-5.2	3.52
4	Fe(II)	5.0-6.5	>7.1[e]	4.3-5.8	6.15
5	Cu(II)	4.1-5.6	9.25	3.6-4.9	13.85
6	Ni(II)	3.5-5.1	8.80	3.3-5.0	9.30
7	Co(II)	4.0-5.5	7.25	2.9-4.7	8.05

a: Calculated for the conditions mentioned in fig. 13. Unit of K is in $Lmol^{-1}$ **b**: The range is given. **c**: Data taken from:[26–28] Conditions: 20 °C, ionic strength = 0.1 M ($NaNO_3$), *Phen* = 1,10-Phenanthroline, *Picol* = di-(2-picolyl)amine **d**: No literature data available **e**: The logarithm of the complexation constant of the 3:1 (ligand : metal) complex was determined to be 21.3. The logarithm of the complexation constant for the 1:1 complex is therefore >7.1.

The calculated range for the complexation constants of the *N*-ligand modified proteins are in general several orders of magnitude lower than the corresponding complexation constants between the free ligand and the corresponding metals. The only exceptions are the Mn(II)-complexes (entry 3) and SCP-2L A100C-*Phen* + Cu(II) (entry 5). The large differences cannot be attributed to the deviation of the different pH in the literature values as the complexation constant between *Phen* and H^+ was determined to be several orders of magnitude smaller than the complexation constant for the corresponding metal complexes except for Mn(II). In case of the Mn(II) complexes it was shown that the

complexation between the unmodified SCP-2L A100C and Mn(II) contained similar quantities of Mn(II) after the workup as the *N*-ligand modified SCP-2L A100C-scaffolds indicating that a large proportion of the complexation might be due to the protein scaffold itself and not the introduced *N*-ligand. Although it might be that the presence of the protein scaffold results in a reduction of the protein-metal complexation constant in comparison to the free ligand-metal complexes due to steric hindrance or other reasons it seems more likely that the approach for the determination of the complexation constants is too inaccurate (see the five assumptions made beforehand). Additionally the range of the calculated equilibrium constants is in between 1.5 to 2.0 orders of magnitude. Therefore no additional equilibrium constants were calculated.

2.6. References

[1] A. Frolov, T.-H. Cho, J. T. Billheimer, F. Schroeder, *J. Biol. Chem.* **1996**, *271*, 31878–31884.

[2] F. Schroeder, S. C. Myers-Payne, J. T. Billheimer, W. G. Wood, *Biochemistry* **1995**, *34*, 11919–11927.

[3] A. Haapalainen, D. van Aalten, G. Meriläinen, J. Jalonen, P. Pirilä, R. Wierenga, J. Hiltunen, T. Glumoff, *J. Mol. Biol.* **2001**, *313*, 1127—1138.

[4] P. Deuss J., Artificial Metalloenzymes; Modified Proteins as Tuneable Transition Metal Catalysts, University of St Andrews, **2011**.

[5] P. J. Deuss, G. Popa, A. M. Z. Slawin, W. Laan, P. C. J. Kamer, *ChemCatChem* **2013**, *5*, 1184–1191.

[6] P. J. Deuss, G. Popa, C. H. Botting, W. Laan, P. C. J. Kamer, *Angew. Chem., Int. Ed.* **2010**, *49*, 5315–5317.

[7] C. Andreini, I. Bertini, G. Cavallaro, G. Holliday, J. Thornton, *J. Biol. Inorg. Chem.* **2008**, *13*, 1205–1218.

[8] S. Mounicou, J. Szpunar, R. Lobinski, *Chem. Soc. Rev.* **2009**, *38*, 1119–1138.

[9] H. Sun, Z.-F. Chai, *Annu. Rep. Prog. Chem., Sect. A: Inorg. Chem.* **2010**, *106*, 20–38.

[10] J. S. Garcia, C. S. de Magalhães, M. A. Z. Arruda, *Talanta* **2006**, *69*, 1–15.

[11] J. R. Carey, S. K. Ma, T. D. Pfister, D. K. Garner, H. K. Kim, J. A. Abramite, Z. Wang, Z. Guo, Y. Lu, *J. Am. Chem. Soc.* **2004**, *126*, 10812–10813.

[12] R. R. Davies, H. Kuang, D. Qi, A. Mazhary, E. Mayaan, M. D. Distefano, *Bioorg. Med. Chem. Lett.* **1999**, *9*, 79–84.

[13] T. Ueno, M. Ohashi, M. Kono, K. Kondo, A. Suzuki, T. Yamane, Y. Watanabe, *Inorg. Chem.* **2004**, *43*, 2852–2858.

[14] F. Xie, D. E. K. Sutherland, M. J. Stillman, M. Y. Ogawa, *J. Inorg. Biochem.* **2010**, *104*, 261–267.

[15] M. Allard, C. Dupont, V. Munoz Robles, N. Doucet, A. Lledos, J.-D. Marechal, A. Urvoas, J.-P. Mahy, R. Ricoux, *ChemBioChem* **2012**, *13*, 240–251.
[16] A. Onoda, K. Fukumoto, M. Arlt, M. Bocola, U. Schwaneberg, T. Hayashi, *Chem. Commun.* **2012**, *48*, 9756–9758.
[17] C. Bertucci, C. Botteghi, D. Giunta, M. Marchetti, S. Paganelli, *Adv. Synth. Catal.* **2002**, *344*, 556–562.
[18] T. Matsuo, C. Imai, T. Yoshida, T. Saito, T. Hayashi, S. Hirota, *Chem. Commun. (Cambridge, U. K.)* **2012**, *48*, 1662–1664.
[19] A. N. Zaykov, B. V. Popp, Z. T. Ball, *Chemistry* **2010**, *16*, 6651–6659.
[20] N. Yokoi, Y. Miura, C.-Y. Huang, N. Takatani, H. Inaba, T. Koshiyama, S. Kanamaru, F. Arisaka, Y. Watanabe, S. Kitagawa, et al., *Chem. Commun.* **2011**, *47*, 2074–2076.
[21] L. Panella, J. Broos, J. Jin, M. W. Fraaije, D. B. Janssen, M. Jeronimus-Stratingh, B. L. Feringa, A. J. Minnaard, J. G. de Vries, *Chem. Commun.* **2005**, 5656–5658.
[22] S. Crobu, M. Marchetti, G. Sanna, *J. Inorg. Biochem.* **2006**, *100*, 1514–1520.
[23] R. den Heeten, B. K. Munoz, G. Popa, W. Laan, P. C. J. Kamer, *Dalton Trans.* **2010**, *39*, 8477–8483.
[24] Q. Jing, K. Okrasa, R. Kazlauskas, Springer Berlin Heidelberg, **2008**, pp. 1–17.
[25] J. Podtetenieff, A. Taglieber, E. Bill, E. J. Reijerse, M. T. Reetz, *Angew. Chem., Int. Ed.* **2010**, *49*, 5151–5155.
[26] G. Anderegg, E. Hubmann, N. G. Podder, F. Wenk, *Helv. Chim. Acta* **1977**, *60*, 123–140.
[27] G. Anderegg, *Helv. Chim. Acta* **1963**, *46*, 2813–2822.
[28] G. Anderegg, *Helv. Chim. Acta* **1963**, *46*, 2397–2410.
[29] P. J. Deuss, G. Popa, C. H. Botting, W. Laan, P. C. J. Kamer, *Angew. Chem., Int. Ed.* **2010**, *49*, 5315–5317.
[30] D. Rivillo, H. Gulyás, J. Benet-Buchholz, E. C. Escudero-Adán, Z. Freixa, P. W. N. M. van Leeuwen, *Angew. Chem., Int. Ed. Engl.* **2007**, *46*, 7247–7250.
[31] P. W. N. M. van Leeuwen, D. Rivillo, M. Raynal, Z. Freixa, *J. Am. Chem. Soc.* **2011**, *133*, 18562–18565.
[32] K. Binnemans, P. Lenaerts, K. Driesen, C. Gorller-Walrand, *J. Mater. Chem.* **2004**, *14*, 191–195.
[33] M. T. Reetz, M. Rentzsch, A. Pletsch, A. Taglieber, F. Hollmann, R. J. G. Mondière, N. Dickmann, B. Höcker, S. Cerrone, M. C. Haeger, et al., *ChemBioChem* **2008**, *9*, 552–564.
[34] M. M. Bradford, *Anal. Biochem.* **1976**, *72*, 248–254.
[35] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605–1612.

Chapter 3: Rh-enzymes as catalysts in the aqueous biphasic hydroformylation

3.1 Abstract

This chapter describes the catalytic performance of artificial rhodium containing enzymes (Rh-enzymes). The main focus is on Rh-enzymes based on the two mutants of SCP-2L described in the previous chapter. Their performance in the aqueous biphasic hydroformylation of long chain 1-alkenes is evaluated and compared to the Rh/TPPTS-system. In order to rationalise the comparably high activity and selectivity of the different Rh-enzymes their mode of action was investigated by supplementary experiments. In these experiments reaction parameters were systematically varied, additional Rh-enzymes and different substrates were tested.

3.2 Introduction

In the previous chapters the development and characterization of metal containing artificial metalloenzymes based on various enzymes were described. Also an overview of the aqueous biphasic hydroformylation of linear 1-alkenes was given and the structural features of the protein SCP-2L were presented. In this chapter this information will be merged with a study on the catalytic performance of artificial rhodium-enzymes in the aqueous biphasic hydroformylation of 1-alkenes with a special emphasis on two SCP-2L mutants which were used as protein scaffolds.

Artificial metalloenzymes as catalysts

Catalysis is usually categorized in terms of homogeneous-, heterogeneous- and bio-catalysis.

In homogeneous catalysis the catalyst and the substrate(s) are homogeneously dissolved in one phase – most often a liquid phase.[1] The inner reaction sphere of the catalyst is well defined by the ligands. The outer reaction sphere consists of randomly distributed counter ions and solvent molecules and is less well defined. As a consequence the outer reaction sphere has hardly any influence on the reaction.

In heterogeneous catalysis the catalyst, which in many cases is a metal is in general immobilised on a solid support. With heterogeneous catalysts it is almost impossible to tune steric and electronic effects of the inner reaction sphere. The activity of heterogeneous catalysts is usually lower than that of their homogeneous counterparts under the same conditions. Another major drawback in heterogeneous catalysis is the troublesome determination of the actual catalytic species and the mechanism of the catalytic process.[2]

Biocatalysis refers to all reactions catalysed by biological systems for example whole organisms, cells or purified enzymes.[3] In protein catalysed reactions the inner reaction sphere is represented by the active site of the protein. The outer reaction sphere consists of other parts of the protein and is much better defined than in man-made catalysts and it has a big influence on the catalytic activity as well as on the stereochemistry of the reaction products but it can also hamper the substrate scope.

As enzymes catalyse many different chemical reactions and often outperform synthetic catalysts in respect of stereo-, regio-, chemo- or enantioselectivity, low toxicity and high reaction rates especially at moderate temperatures, the number of enzymatic processes in industry is constantly increasing.[4] For some industrially relevant processes, as for example the hydroformylation, no catalytically active enzymes have been discovered, so far.

Artificial metalloenzymes consist of a combination of metal and protein which are non-existent in nature and might close the gap between enzymatic and homogeneous catalysis by "*merging the best of two worlds*"[5] - Even though the first studies using artificial metalloenzymes as catalysts were already published in the late 1970s[6,7] this emerging field at the borders of organic- and inorganic chemistry as well as biology only became popular at the end of the millennium.

The existing strategies of metal incorporation into enzymes as well as the use of artificial metalloenzymes in the hydroformylation reaction will be subsequently reviewed. Peptide-, antibody-, and DNA- based catalysts which are sometimes considered as artificial metalloenzymes [5,8] will not be covered.

The metal anchoring strategies used to obtain an artificial metalloenzyme are usually divided into the following subcategories: dative-, supramolecular- and covalent anchoring.[8]

Dative anchoring

The first strategy ever employed to obtain an artificial metalloenzyme was dative anchoring.[9] In this strategy the binding capability of amino acids towards metals is exploited. The most straight forward approach comprises the simple addition of metal salts to a protein scaffold yielding an artificial metalloenzyme. As most proteins have multiple binding sites the resulting metal containing species cannot be predicted. Also the precise location of the metal(s) is difficult to analyse.

In order to obtain better defined species additional and more sophisticated approaches towards enzymes suitable for dative anchoring are employed. The four approaches are outlined in figure 1.

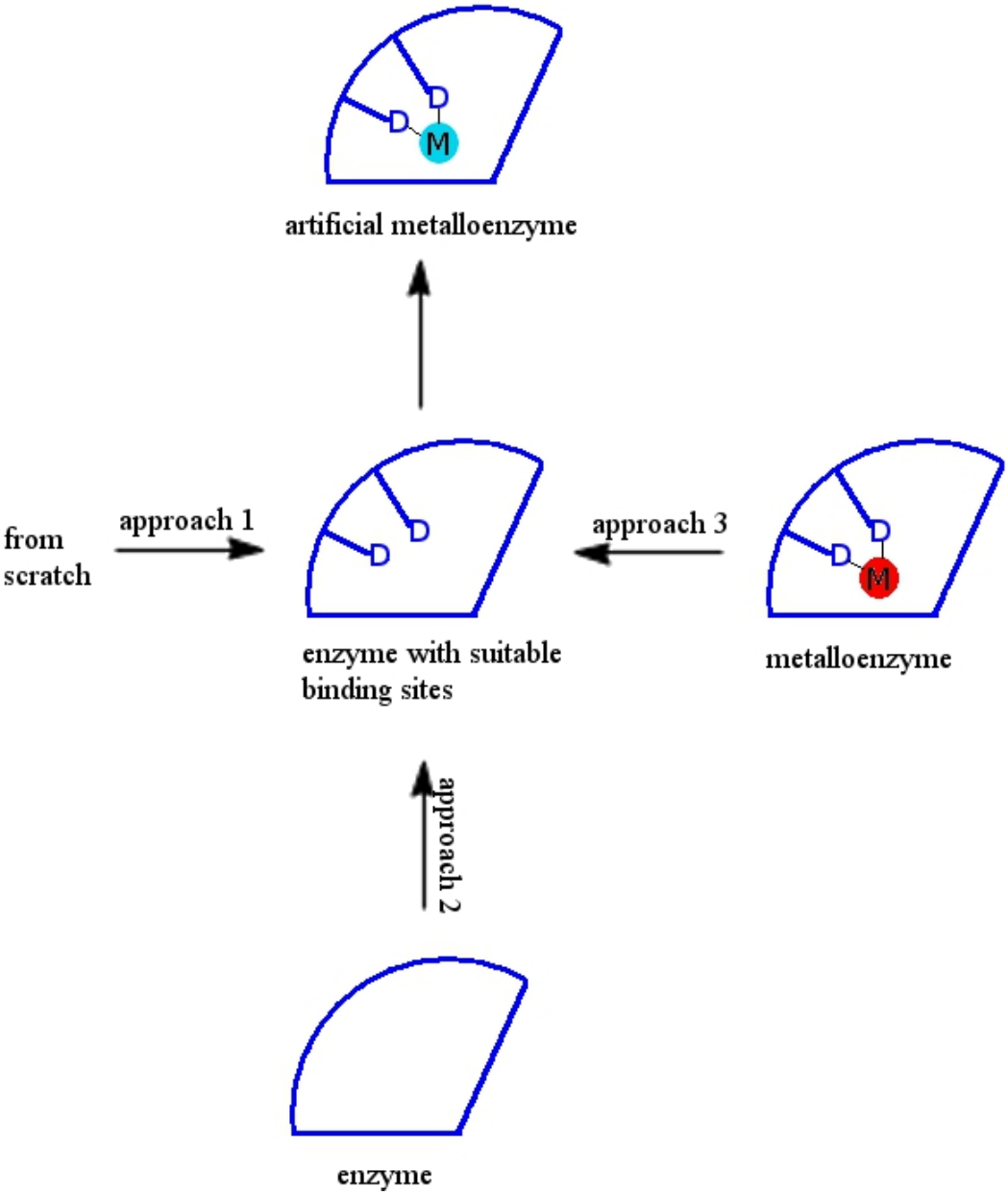


Figure 1 Approaches towards artificial metalloenzymes using dative anchoring. D = donor sites (oxygen, nitrogen, sulphur) present in the side chain of amino acids. M = transition metal

The first approach comprises the design and synthesis of a well-defined active site from scratch. The resulting active site is then incorporated into a suitable protein scaffold. Usually computational support is applied in this approach. So far, the enzymes obtained by this route only result in small rate accelerations in comparison to uncatalysed

reactions but they often prove as a decent starting point for further optimisations.[10]

In the second approach a suitable binding site is introduced into a protein for example by mutagenesis. This approach requires knowledge about the structure of the protein and usually uses a combination of chemical intuition, trial and error and computational methods.[11–13]

A third approach comprises the replacement of a metal present in a metalloenzyme by a different metal.[14,15]

Artificial metalloenzymes obtained by dative anchoring have been successfully used in various reactions like asymmetric epoxidation[14,15], hydrogenation[16], hydroformylation[17–19] and asymmetric Diels-Alder reaction[11]. A major advantage of dative anchoring strategies is that a wide array of proteins can be used as scaffold. A major drawback of this approach is that the actual environment of the metal is often unknown and most likely a mixture of many different species is obtained instead of one well defined one. To overcome this issue other potential binding sites need to be removed by chemical and/or biological means in order to obtain a single catalytically active species.[19]

Supramolecular anchoring

Supramolecular anchoring was first employed in the 1970s to obtain an artificial metalloenzyme.[20] This approach exploits the high binding constants between (some) proteins and smaller chemical molecules (from now on summarised as “cofactors”). Instead of using the genuine cofactor a modified cofactor which bears a metal or is capable of binding to a metal is used (schematic representation in figure 2).

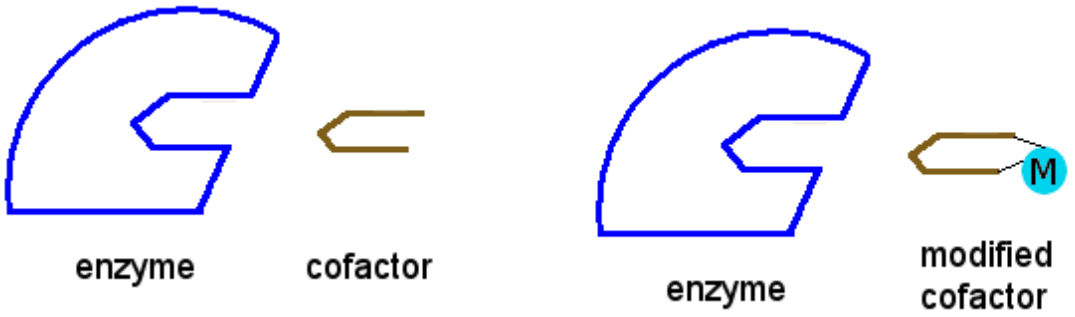


Figure 2 Principle of supramolecular anchoring. Left: Enzyme and corresponding cofactor. Right: Enzyme and a modified cofactor bearing a metal.

Probably the most exploited combination is the one between the modified cofactor biotin to either the protein avidin or streptavidin. The binding constant between biotin to (strept)avidin was determined as $\approx 10^{15}$ M^{-1}[21] which is one of the highest non covalent binding constants in nature. Artificial metalloenzymes obtained by supramolecular anchoring have been successfully used in various reactions like asymmetric hydrogenation[20,22,23], transfer hydrogenation[24–27], allylic alkylation[22] and sulfoxidation.[28] They are also known to catalyse C-H-activations[29] and metathesis[30] While the catalysts obtained by supramolecular anchoring possess a well defined outer reaction sphere the potential pool of enzymes is limited to specific sets of proteins and cofactor analogues.

Covalent modification

In covalent modification a binding site is covalently introduced at a specific position of the protein scaffold. This is usually achieved by the introduction of a single amino acid with a specific reactivity at the desired location by mutation of the wild type protein. In case the protein contains one ore more of these amino acids multiple modification will occur. Therefore the undesired amino acids have to be substituted by unreactive amino acids via mutagenesis in the first place. Virtually any protein can be used as scaffold in the covalent approach. Additionally a well defined species is obtained, therefore covalent modification combines the advantages of dative- and supramolecular anchoring. Supramolecular- and covalent anchoring have also been merged in dual

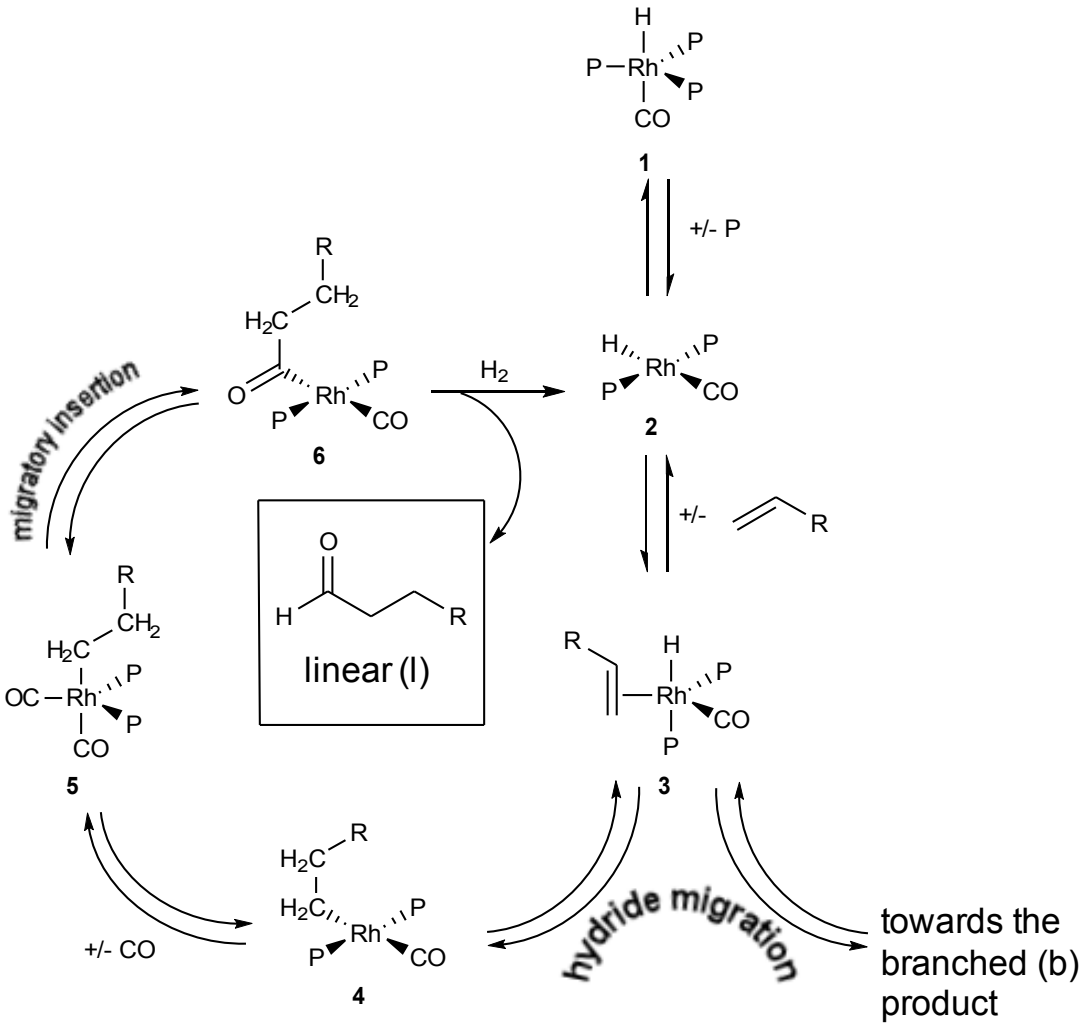
anchoring, which have the disadvantage that the optimisation of the resulting catalysts is not straightforward.[31] Artificial metalloenzymes obtained by covalent modification have been successfully tested in many different reactions like Diels-Alder[32], hydroformylation[33], sulfoxidation[34] and polymerisation[35]. Reaction schemes in which the covalent modification of a protein are outlined are depicted in various parts of this thesis.

Catalyst optimisation

The artificial metalloenzymes obtained by any of the methods mentioned are usually subjected to optimisation steps in order to improve key features like rate enhancement, enantioselectivity or stability. The optimisation methods comprise chemical,[36] or biological means.[37–42]

Mechanism of the rhodium catalysed hydroformylation

In 1970 Wilkinson and co-workers suggested a mechanism for the hydroformylation of 1-olefins (Wilkinson's dissociative mechanism) which is still accepted nowadays.[43–46] The reaction cycle is depicted in scheme 1.



Scheme 1 Simplified catalytic cycle for the Rh-catalysed hydroformylation of a 1-alkene ($P=PPh_3$).[43–46]

According to scheme 1 the first reaction step is the dissociation of a ligand from complex **1** or a similar precursor to form the square planar rhodium complex **2** which is considered as the catalytically active species. Upon coordination of a 1-alkene to its vacant site complex **3** is obtained. Hydride migration yields complex **4** or its branched analogue (not shown in the scheme). By the addition of CO complex **5** is formed. This species can undergo migratory insertion of the alkyl chain to form **6**. Upon hydrogenation of **6** the linear aldehyde and the catalytically active species **2** are obtained.

Instead of “P”-ligands “CO” (or other ligands) might coordinate to the metal as well. The prediction of the linear to branched ratios is not straightforward. The general accepted explanation for the formation of linear aldehydes is that the steric hindrance of the (substituents on the) phosphine favours the formation of linear aldehydes. [47]

Artificial metalloenzymes in hydroformylation

Due to the high toxicity of carbon monoxide to organisms unsurprisingly not a single naturally occurring enzyme is known to catalyse the hydroformylation and only a few artificial metalloenzymes have been reported as active catalysts in this reaction.[17–19,33,48] Due to its significantly increased hydroformylation activity in comparison to any other metal[39] rhodium is used as the active transition metal in all of these examples. The artificial Rh-enzymes were prepared by dative anchoring[17–19] and by covalent modification.[33,48] No artificial metalloenzymes employing supramolecular anchoring has been reported for this reaction so far. Research was mainly focused on the substrates styrene and linear 1-alkenes. In these studies the reaction was performed in an aqueous-organic biphasic type of reaction with the Rh-enzyme in the aqueous phase and the substrate in the organic phase. In the biphasic reactions rhodium can leach into the organic phase and catalyse the reaction. In order to account for this undesired catalysis it is important to determine the Rh-content of the organic phase after the reaction. Unfortunately this aspect was not considered in all publications. Table 1 gives a literature overview on artificial Rh-enzymes used in the hydroformylation of styrene and long chain linear 1-alkenes.

Table 1 Artificial metalloenzymes tested in hydroformylation.*

entry	catalyst	substrate	TON[a]	l / b[b]	leaching[c]
1[d]	HSA*Rh[e]	styrene	>600	5 / 95	8 ppm (1.3-1.9 %)
2[f]	HSA*Rh[e]	styrene	>250	n.d.[g]	n.d.[g]
3[h]	papain-phospite-Rh[i]	styrene	>680	n.d.[g]	n.d.[g]
4[j]	hcAII[k]	styrene	213	89 / 11	n.d.[g]
5[f]	papain*Rh[l]	styrene	≈250	8 / 92	n.d.[g]
6[f]	egg albumin*Rh[m]	styrene	>250	10 / 90	n.d.[g]
7[f]	HSA*Rh[e]	1-octene	>250	n.d.[g]	n.d.[g]
8[d]	HSA*Rh[e]	1-octene	>600	47 / 53	8 ppm (1.3-1.9 %)
9[f]	papain*Rh[l]	1-octene	≈25	55 / 45	n.d.[g]
10[f]	egg albumin*Rh[m]	1-octene	≈5	100 / 0	n.d.[g]
11[n]	SCP-2L A100C-1-P(para)-Rh[o]	1-octene	124	87 / 13	n.d.[g]
12[n]	SCP-2L V83C-1-P(para)-Rh[p]	1-decene	34	78 / 22	<0.5 ppm (<1.2 %)

***:** In case multiple examples within one reference were published the example in which Rh-leaching was determined was used. If no such information was given the example with the highest TON is listed. **a**: Turnover number (TON) – ratio of products to rhodium **b**: linear to branched product ratio **c**: Rhodium concentration in the organic phase after the reaction; percentage of the total initial rhodium amount in brackets. **d**: Data taken from [17] **e**: The protein human serum albumin (HSA) and $Rh(acac)(CO)_2$ in a molar ratio of ≈1:56. **f**: Data taken from [18] **g**: not determined **h**: Data taken from [48] **i**: The protein papain chemically modified with a phosphite at cysteine 25 (the only cysteine) and treated with $Rh(acac)(CO)_2$ in a molar ratio of ≈1:29. The solution is filtered prior to catalysis to remove the excess of rhodium. The TON is calculated on the assumption that no rhodium is removed by this filtration **j**: Data taken from [19] **k**: The protein mutant of human carbonic anhydrase II with the following mutations: H4R, H10R, H17F, The histidines on the surface were chemically modified with Diethylpyrocarbonate. A Rh:protein ratio of 1.4 was detected **l**: The protein papain treated with $Rh(acac)(CO)_2$ in a molar ratio of ≈1:20 **m**: The protein egg albumin treated with $Rh(acac)(CO)_2$ in a molar ratio of ≈1:38 **n**: Data taken from [33] **o**: The protein mutant SCP-2L A100C covalently modified with linker and P(para) (introduced earlier in this thesis) treated with $Rh(acac)(CO)_2$ in a molar ratio of ≈2:1 **p**: The protein mutant SCP-2L V83C covalently modified with linker and P(para) (introduced earlier in this thesis) treated with $Rh(acac)(CO)_2$ in a molar ratio of ≈2:1.

As can be seen in table 1 styrene is converted into the corresponding aldehydes with TON >600 when the Rh-enzyme obtained by combining the protein human serum albumin (HSA) with a large excess (≈56 eq.) of Rh(I) is used (entry 1). The rhodium concentration in the organic phase was determined as 8 ppm after the reaction indicating significant rhodium leaching. Covalent modification of the protein papain with a phosphite, treatment with an excess (≈29 eq.) of Rh(I) and subsequent filtration of the mixture to remove the excess of Rh(I) yields a TON >680 in the hydroformylation of styrene (entry 3). Unfortunately rhodium leaching was not determined.[48] All other

experiments in which styrene was used as substrate resulted in significantly smaller TON. The uncomplexed Rh(acac)$(CO)_2$ is known to give a linear to branched ratio of about 17 / 83[19] which is similar to the ratio obtained by most Rh-enzymes listed (entries 1, 5, 6). The linear selectivity totally differs when the protein human carbonic anhydrase is used as scaffold (entry 4) resulting in a preference for the linear product and a linear to branched ratio of 89 / 11. This is a strong indication for the influence of the protein scaffold.

The enantioselectivity of the reaction is only investigated in two of the examples. When the papain-phosphite was used as catalyst a racemate is obtained (entry 3), The HSA-Rh-complex (entry 1) yielded a "very low but definite enantiomeric excess".[17] This excess was not specified any further and was not mentioned in a subsequent publication anymore.[18]

When Rh-enzymes are used in the hydroformylation of linear 1-alkenes (1-octene and 1-decene) high TON (>600) are found when the HSA based Rh-enzyme is used (entry 8) while all other protein scaffolds yield much lower TON. The uncomplexed Rh(acac)$(CO)_2$ is known to give a linear to branched ratio of about 56 / 44[33] which is similar to the ratio obtained by most Rh-enzymes listed (entries 8, 9). The linear selectivity totally differs when the phosphine modified SCP-2L is used as scaffold (entries 11, 12) resulting in a profound preference for the linear product. This is a strong indication for the influence of the protein scaffold in the reaction.

Most of the artificial metalloenzymes are described as unstable under the reaction conditions used. The only exceptions are the HSA-Rh-complex (entries 1, 2, 7, 8) which is considered as stable.[16,17] For the phosphite modified papain the stability was not mentioned.[48]

Concept of SCP-2L as protein scaffold in the hydroformylation

The protein SCP-2L was considered as a protein scaffold suitable for the aqueous biphasic hydroformylation of long chain 1-alkenes due to its hydrophobic tunnel. This flexible tunnel is known to bind a great variety of different substrates. It was proven in the previous chapter that this tunnel is still intact within the mutant SCP-2L A100C. It was also pointed out that the tunnel within the mutant SCP-2L V83C seems to remain

intact as well. These two mutants possess their unique cysteine at either end of the tunnel. Therefore any covalently introduced modification will be in close proximity to either end of the tunnel. The following concept was envisioned (see figure 3):

- Under biphasic reaction conditions the apolar substrate will accumulate in the tunnel, resulting in a much higher concentration than in the surrounding aqueous phase.
- Due to the modification of the protein with a phosphine the transition metal coordinating to this phosphine will be in close proximity to the substrate.
- The bulky protein environment is expected to induce selectivity towards the linear product.

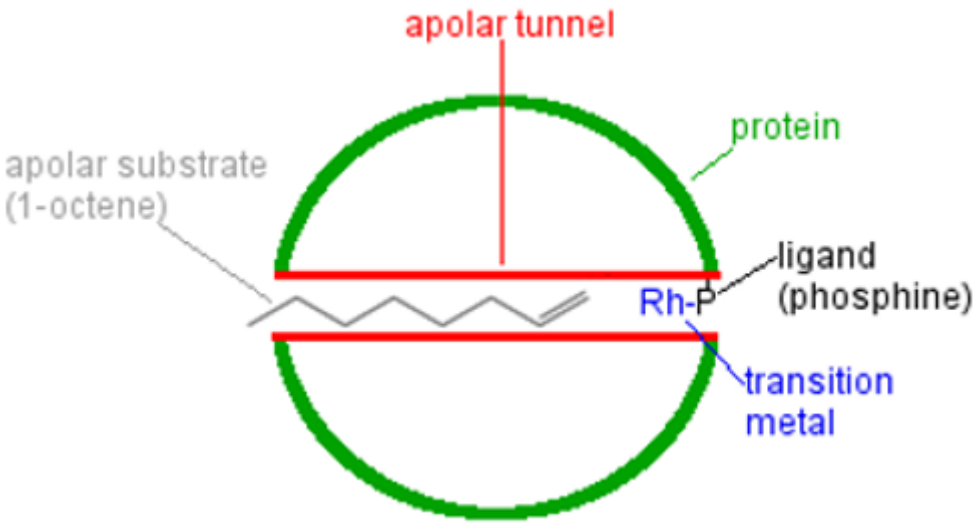


Figure 3 Envisioned concept for SCP-2L based Rh-enzymes used as hydroformylation catalysts.

3.3 Results and Discussion

Synthesis of Rh-enzymes

Both scaffolds SCP-2L A100C and SCP-2L V83C were modified with three different phosphines to obtain six different Rh-enzymes (see figure 4, the synthesis was discussed in chapter 2).

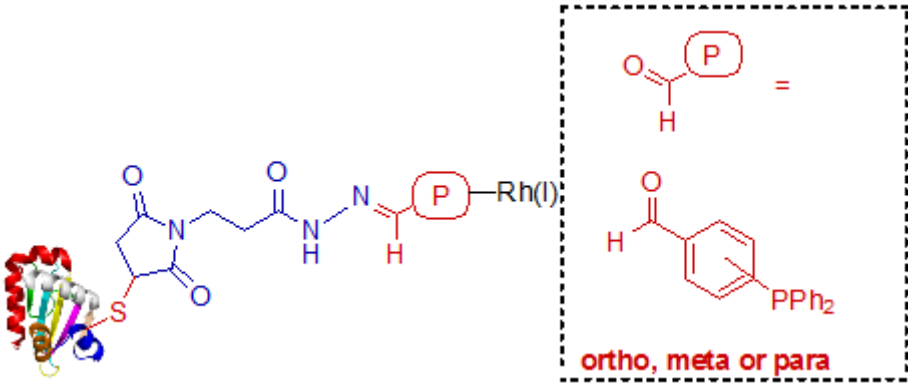
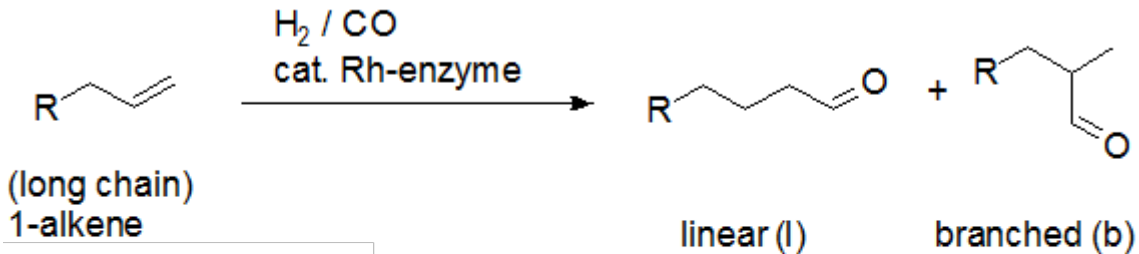


Figure 4 Synthesis of Rh-enzymes (detailed description in chapter 2).

The Rh-enzymes are tested in the hydroformylation of linear 1-alkenes (see scheme 2)



1-hexene,
1-octene,
1-decene,
1-dodecene or
1-octadecene

Scheme 2: Hydroformylation reaction with all substrates which were tested.

The Rh/TPPTS system

The Rh/TPPTS system was already briefly introduced in chapter 1. This system was used as a reference in order to compare the catalytic performance of the Rh-enzymes. The catalytic performance of the Rh/TPPTS-system in the aqueous biphasic

hydroformylation of long chain 1-alkenes was tested using reaction conditions similar to the ones later used for the Rh-enzymes. The results are given in table 2.

Table 2 Hydroformylation of 1-alkenes using Rh/TPPTS[a]

entry	catalyst[b]	substrate[c]	TON[d]	% linear selectivity[d]
1	Rh / TPPTS	1-octene	1 ± 1	72 ±7
2	Rh / TPPTS	1-decene	<1	73 ± 3
3	Rh / TPPTS	1-dodecene	-[e]	-[e]
4	Rh / TPPTS	1-octadecene	-[e]	-[e]

a: 80 bar syn gas (1:1 mixture of H_2 and CO), 48h, 35 °C, stirred at 625 rpm. 0.5 mL aqueous and organic solution each. Each data point represents the average of four experiments. A comparision with the literature known data is given in the experimental section **b**: Preformed Rh/TPPTS (1:30), 260.56 nanomol Rh in 0.5 mL 20 mM MES, 50 mM NaCl, pH=6 **c**: Substrate containing 9% (v/v) n-heptane and 1% (v/v) diphenylether **d**: determined by GC **e**: below the detection limit.

As known[33] and shown in table 2 the activity of the Rh/TPPTS-system is negligible for long chain 1-alkenes due to their low water solubility.[1] Minimal conversions could only be detected for the two shortest 1-alkenes (1-octene and 1-decene) tested with a linear selectivity of around 72%.

Evaluation of Rh-enzymes in the hydroformylation

The three Rh-enzymes based on the SCP-2L V83C mutant (see figure 4) were tested in the aqueous biphasic hydroformylation of 1-octene, 1-decene, 1-dodecene, and 1-octadecene. The results are listed in table 3.

Table 3 Hydroformylation of 1-alkenes using SCP-2L V83C-Phosphine-Rh[a]

entry	catalyst	substrate[b]	TON[c]	% linear selectivity[c]
1	SCP-2L V83C-1-P(ortho)-Rh[d]	1-octene	1 ± 0	79 ± 1
2	SCP-2L V83C-1-P(meta)-Rh[e]	1-octene	33 ± 2	74 ± 0
3	SCP-2L V83C-1-P(para)-Rh[f]	1-octene	74 ± 9	78 ± 0
4	SCP-2L V83C-1-P(ortho)-Rh[d]	1-decene	-[g]	-[g]
5	SCP-2L V83C-1-P(meta)-Rh[e]	1-decene	27 ± 7	72 ± 0
6	SCP-2L V83C-1-P(para)-Rh[f]	1-decene	39 ± 7	76 ± 0
7	SCP-2L V83C-1-P(ortho)-Rh[d]	1-dodecene	-[g]	-[g]
8	SCP-2L V83C-1-P(meta)-Rh[e]	1-dodecene	6 ± 3	73 ± 1
9	SCP-2L V83C-1-P(para)-Rh[f]	1-dodecene	13 ± 5	75 ± 2
10	SCP-2L V83C-1-P(ortho)-Rh[d]	1-octadecene	-[g]	-[g]
11	SCP-2L V83C-1-P(meta)-Rh[e]	1-octadecene	4 ± 1	76 ± 1
12	SCP-2L V83C-1-P(para)-Rh[f]	1-octadecene	2 ± 1	71 ± 3

a: 80 bar syn gas (1:1), 48h, 35 °C, stirred at 625 rpm. 0.5 mL aqueous and organic solution each. Each data point represents the average of four experiments **b**: Substrate containing 9% (v/v) n-heptane and 1% (v/v) diphenylether **c**: determined by GC **d**: SCP-2L V83C-1-P(ortho) : Rh = 5.8; 30.89 nanomol Rh (determined by ICP-MS) in 0.5 mL 20 mM MES, 50 mM NaCl, pH=6 **e**: SCP-2L V83C-1-P(meta) : Rh = 1.8; 99.12 nanomol Rh (determined by ICP-MS) in 0.5 mL 20 mM MES, 50 mM NaCl, pH=6 **f**: SCP-2L V83C-1-P(para) : Rh = 3.0; 46.58 nanomol Rh (determined by ICP-MS) in 0.5 mL 20 mM MES, 50 mM NaCl, pH=6 **g**: below the detection limit

Although the rhodium concentrations are about a factor three to eight lower than in the Rh/TPPTS system (table 2) considerably higher TON could be observed for most Rh-enzymes. The selectivity is virtually the same regardless of the Rh-enzyme and the substrate tested. All reactions (with conversion) result in about 75% linear selectivity. No products could be detected when 1-dodecene and 1-octadecene were used as substrates in the Rh/TPPTS-system (table 2). The same was observed for one of the Rh-enzymes listed in table 3 (entries 7, 10). The other two Rh-enzymes only yield small TON for these substrates (entries 8, 9, 11, 12) but they are clearly detectable.

The results can be summarised as a decreasing activity with increasing number of carbon atoms of the starting material as well an activity following the order P(para)>P(meta)>P(ortho). The least active Rh-enzymes based on SCP-2L V83C-P(ortho) show an activity and selectivity comparably to the Rh/TPPTS-system while the other Rh-enzymes are several orders of magnitude more active. A constant linear selectivity of about 75% is observed.

In the next step the three Rh-enzymes based on the SCP-2L A100C mutant (see figure

4) were tested in the same set of reactions. The results are listed in table 4.

Table 4 Hydroformylation of 1-alkenes using SCP-2L A100C-Phosphine-Rh[a]

entry	catalyst	substrate[b]	TON[c]	% linear selectivity[c]
1	SCP-2L A100C-1-P(ortho)-Rh[d]	1-octene	1 ± 1	70 ± 4
2	SCP-2L A100C-1-P(meta)-Rh[e]	1-octene	106 ± 2	70 ± 1
3	SCP-2L A100C-1-P(para)-Rh[f]	1-octene	409 ± 58	79 ± 6
4	SCP-2L A100C-1-P(ortho)-Rh[d]	1-decene	-[g]	-[g]
5	SCP-2L A100C-1-P(meta)-Rh[e]	1-decene	75 ± 8	66 ± 1
6	SCP-2L A100C-1-P(para)-Rh[f]	1-decene	135 ± 16	74 ± 1
7	SCP-2L A100C-1-P(ortho)-Rh[d]	1-dodecene	-[g]	-[g]
8	SCP-2L A100C-1-P(meta)-Rh[e]	1-dodecene	46 ± 9	64 ± 0
9	SCP-2L A100C-1-P(para)-Rh[f]	1-dodecene	66 ± 3	73 ± 0
10	SCP-2L A100C-1-P(ortho)-Rh[d]	1-octadecene	-[g]	-[g]
11	SCP-2L A100C-1-P(meta)-Rh[e]	1-octadecene	21 ± 4	64 ± 1
12	SCP-2L A100C-1-P(para)-Rh[f]	1-octadecene	20 ± 7	72 ± 2

a: 80 bar syn gas (1:1), 48h, 35 °C, stirred at 625 rpm. 0.5 mL aqueous and organic solution each. Each entry represents the average of four experiments **b**: Substrate containing 9% (v/v) n-heptane and 1% (v/v) diphenylether **c**: determined by GC **d**: SCP-2L A100C-1-P(ortho) : Rh = 3.5; 26.87 nanomol Rh (determined by ICP-MS) in 0.5 mL 20 mM MES, 50 mM NaCl, pH=6 **e**: SCP-2L A100C-1-P(meta) : Rh = 2.2; 37.72 nanomol Rh (determined by ICP-MS) in 0.5 mL 20 mM MES, 50 mM NaCl, pH=6 **f**: SCP-2L A100C-1-P(para) : Rh = 1.5; 22.98 nanomol Rh (determined by ICP-MS) in 0.5 mL 20 mM MES, 50 mM NaCl, pH=6 **g**: below the detection limit

The TON decreases with increasing numbers of carbon atoms of the 1-alkenes. In addition, there is a significant difference between the three phosphines. The activity ranking of the phosphines is the same as for the SCP-2L V83C mutant: P(para)>P(meta)>P(ortho). The activity of the SCP-2L A100C mutant modified with P(ortho) was so small that conversions could only be detected for the shortest (and therefore most active) 1-alkene investigated. The selectivity of the reaction expressed by the "% linear selectivity" in the table slightly decreases with the number of C-atoms and follows the order P(para)>P(meta)≈P(ortho).

The TON are by a factor 3-5 higher than the corresponding TON with SCP-2L V83C as scaffold (excluding those results in which either TON is ≤ 10). It needs to be emphasised that the activity of the most active Rh-enzyme (SCP-2L A100C-1-P(para) is several orders of magnitude higher in the hydroformylation of 1-octene and 1-decene than the corresponding Rh/TPPTS system (a detailed calculation is given in the

experimental section). While the Rh/TPPTS system requires high ligand to Rh ratios to obtain high selectivities for the Rh-enzymes a good linear selectivity is obtained at only a slight molar excess of the protein.

The results for the different SCP-2L based Rh-enzymes (table 3 and table 4) can be explained with the concept suggested in figure 3: The different activities for the different phosphines used could be explained by a mismatch of the coordinating rhodium to the substrate. A large mismatch in the case of P(ortho), which is slightly less in case of P(meta). In case of P(para) the coordinating rhodium might be in proximity to the substrate within the tunnel, resulting in excellent TON. It could also be due to different protein to rhodium ratios. The different activities of the substrates can be explained with a different binding preference within the tunnel. Finally the difference between the two mutants as scaffold can be explained by a better pairing between the complexed rhodium and the substrate. The linear selectivities observed can be explained by the bulky environment within the hydrophobic pocket of the protein. One might argue that the activity differences observed can be attributed to the different Protein : Rh-ratios used. This is not the case, as different SCP-2L A100C-P(para) : Rh ratios are investigated (fig. 7) and the activities are still high for high SCP-2L A100C-P(para) : Rh.

The mode of action of the Rh-enzymes

So far, the high catalytic activities of the P(para) modified SCP-2L scaffolds provide strong support for the validity of the envisioned concept put forward in figure 3. In order to exclude catalysis by artefacts additional experiments were performed. In the first step catalysis by rhodium leaching is addressed.

1) leaching

The mode of action investigated in this sub chapter is concerned with the catalysis by rhodium leached into the organic phase. Rhodium which migrates into the organic phase can either originate from a back reaction of the Rh-enzyme. The “back reactions” which might occure are the hydrolysis of the hydrazone, the retro michael addition, or the

decomplexation of rhodium. Leached rhodium would result in high activity and low selectivity but both pathways might result in leaching of a phosphine-rhodium-fragment which would result in a more selective reaction. The concept is depicted in figure 5.

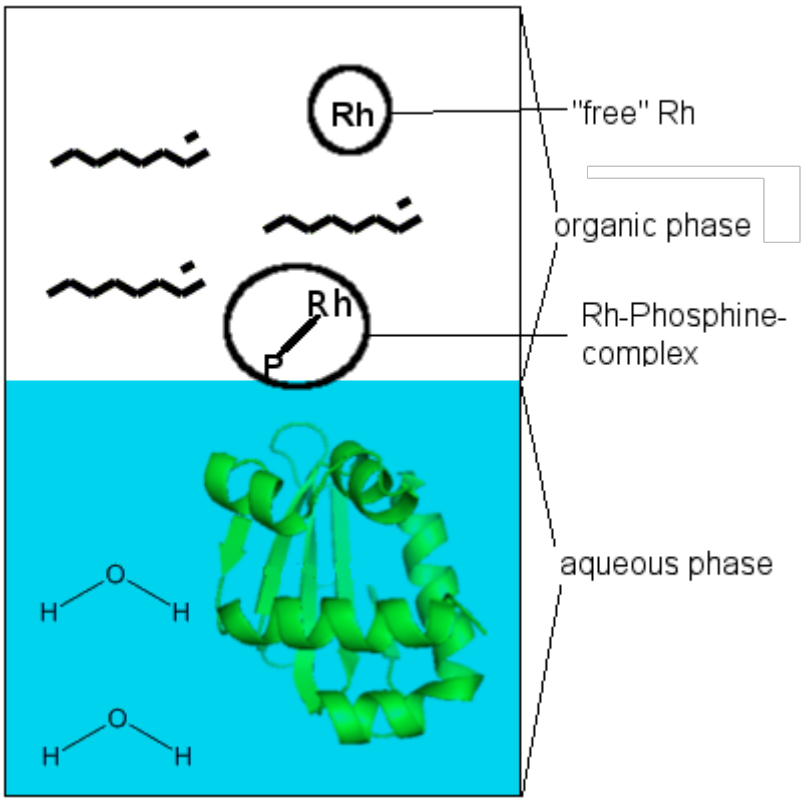


Figure 5 Hydroformylation catalysed by different rhodium-species leached into the organic phase.

In this figure the catalytically active species is either "free" rhodium (for example $HRh(CO)_4$) or a Rh-ligand fragment which is transferred into the organic phase. "Free" rhodium results in a linear selectivity of ≈58% under the conditions used. This is much lower than the linear selectivity observed when Rh-enzymes are used as catalysts. Therefore "free" rhodium cannot solely be the cause for the high activities observed. As it is also plausible to assume that the amount of Rh leached is the same regardless of the substrate used the comparably small conversions of 1-octadecene (for any Rh-enzyme tested) show that the total amount of free rhodium leached has to be quite small.

Study of potentially occuring "back reactions" of the Rh-enzyme

Analysis of SCP-2L A100C-1-P(para)-Rh by MALDI-TOF made it seem possible that the Rh-enzymes might release a P-Rh-fragment during the reaction. The MALDI-TOF spectrum of SCP-2L A100C-1-P(para)-Rh is depicted in figure 6.

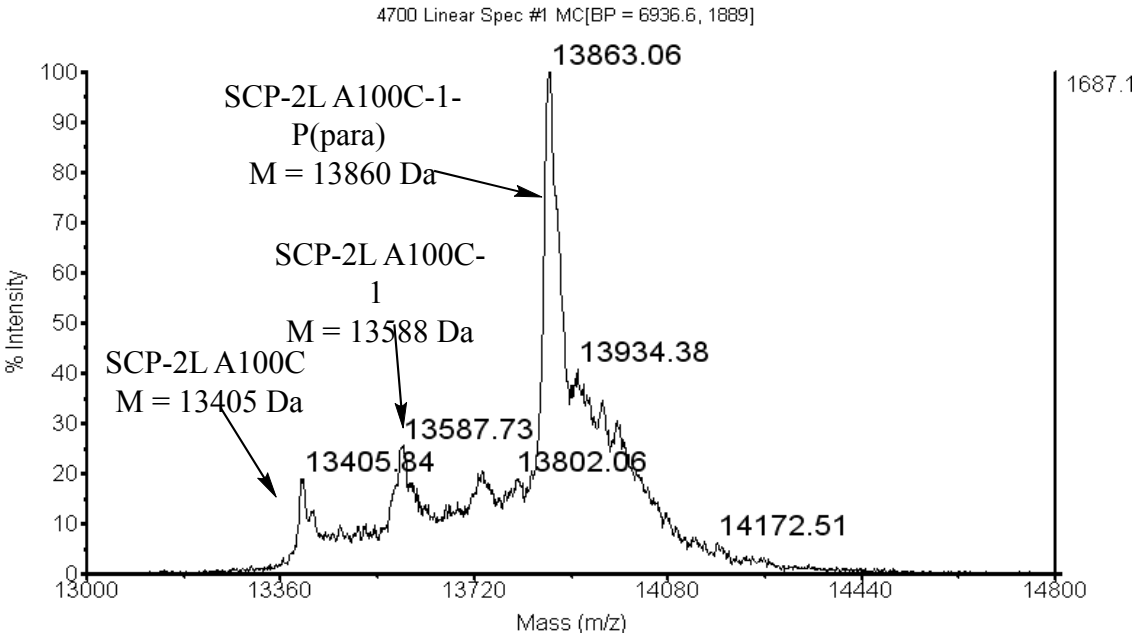


Figure 6 MALDI-TOF MS spectrum of SCP-2L A100C-1-P(para)-Rh before hydroformylation.

The spectrum contains three significant peaks. The peak at 13405 Da corresponds to the mass of the unmodified SCP-2L A100C. The second peak at 13588 Da corresponds to the mass of the linker modified SCP-2L A100C-1. The third peak at 13863 Da corresponds to the phosphine modified SCP-2L A100C-1-P(para). Various signals are present in the region in which the mass of the Rh-enzyme is expected but no distinguishable peak can be identified. The presence of both the unmodified protein as well as the linker modified protein in the spectrum might be due to back reactions of the Rh-enzyme or the phosphine modified enzyme during the analysis. It might also be possible that these species are always present due to incomplete modification but can only be detected by MALDI-TOF and not by the usually applied ESI-MS due to the higher sensitivity of the former one. This finding makes it seem likely that Rh-enzymes

can undergo the back reaction under the reaction conditions. The back reaction would release a P-Rh-fragment of the enzyme which can migrate to the organic phase and catalyse the reaction in a monophasic fashion causing the high activities and selectivities observed. In a previous publication the analysis of a Rh-enzyme (SCP-2L V83C-1-P(para)-Rh) by LC-MS recovered from the aqueous phase after the hydroformylation did show small quantities of both unmodified and linker modified protein.[33] In a reproduction of this experiment only the oxide of SCP-2L V83C-1-P(para) was detected when SCP-2L V83C-1-P(para)-Rh was analysed by ESI-MS after the reaction. The total concentrations of P-Rh-fragments released under the conditions are therefore most likely quite small.

Variation of SCP-2L A100C-1-P(para) to Rh ratios

It has been mentioned that due to its low selectivity "free" rhodium cannot be responsible for the total hydroformylation activity observed and is at the most responsible for a small fraction of the total conversion. In order to further evaluate this amount a reaction series was performed in which the ratio of SCP-2L A100C-1-P(para) to Rh was systematically varied. In this series the Rh-concentration in the reaction mixture was kept constant. The results of these experiments are depicted in figure 7.

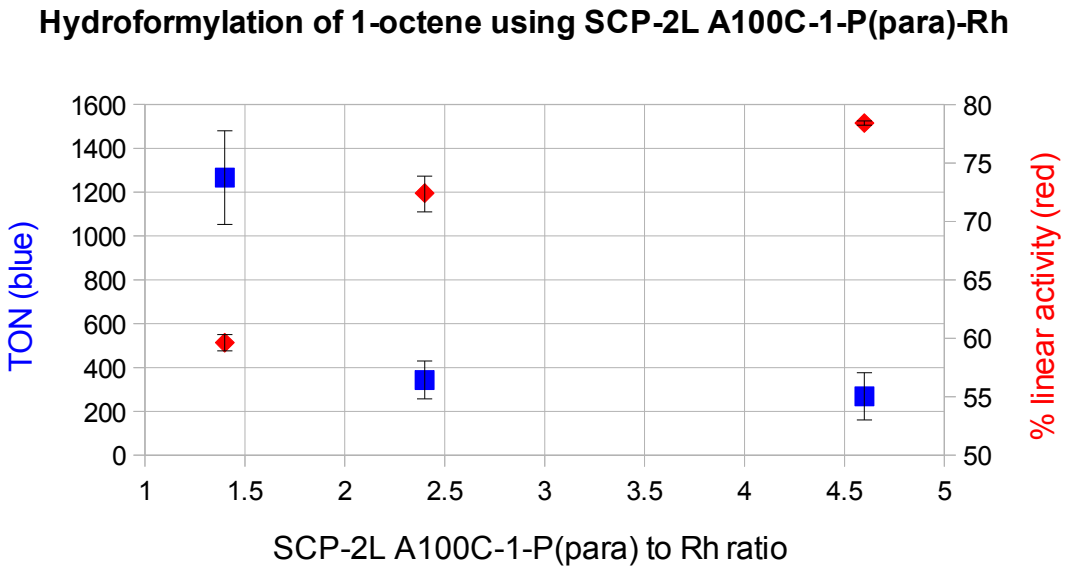


Figure 7 Variation of the SCP-2L A100C-1-P(para) / Rh-ratios in the hydroformylation of 1-octene (at constant rhodium concentrations).

Reaction conditions: 80 bar syn gas (1:1), 48h, 35 °C, stirred at 625 rpm. Organic solution: 0.5 mL 1-octene containing 9% (v/v) n-heptane and 1% (v/v) diphenylether. Aqueous solution: 0.5 mL 20 mM MES, 50 mM NaCl, pH=6. TON & linear selectivity determined by GC. 20.77 nanomol Rh (determined by ICP-MS). Each data point represents the average of two experiments.

For a SCP-2L A100C-1-P(para) to rhodium ratio ≈1.4 figure 7 shows a high activity and a low selectivity. Upon increasing this ratio the activity decreases and the selectivity increases. These dependencies can be explained by a decrease of "free" rhodium and an increase of the desired Rh-enzyme on increasing the ratio. The residual activity at high enzyme/Rh-ratios will be caused by the Rh-enzyme.

The observations could also be explained by the back reaction of the phosphine-modified-enzyme which releases a P-fragment. This is coordinating to leached rhodium and causing the selectivity observed. As the concentration of this P-fragment is proportional to the starting concentration of phosphine modified protein a higher selectivity is observed with increasing enzyme/Rh-ratios. On the basis of the experimental evidence given in figure 7 it is not possible to decide between these two possibilities.

Performance of P-Rh-fragments in the monophasic hydroformylation

In order to mimic the catalytic activity of a P-Rh-fragment, the hydroformylation of 1-octene was performed in the presence of rhodium and either of the three phosphine ligands P(ortho), P(meta), and P(para) used in the protein modification. In this experimental series the concentration of rhodium was kept constant and the ratios of P/Rh varied. The reactions were performed in a single organic phase (table 5).

Table 5 Hydroformylation of 1-octene with three different phosphines at varying P/Rh ratios (constant Rh)[a]

entry	catalyst[b]	Rh / P[c]	TON[d]	% linear selectivity[d]
1	P(ortho) + Rh	0.15 / 0	312 ± 36	56 ± 0
2	P(ortho) + Rh	0.15 / 0.08	728 ± 72	56 ± 0
3	P(ortho) + Rh	0.15 / 0.21	386 ± 47	56 ± 0
4	P(ortho) + Rh	0.15 / 1.6	1 ± 0	68 ± 1
5	P(meta) + Rh	0.15 / 0	530 ± 53	55 ± 1
6	P(meta) + Rh	0.15 / 0.05	691 ± 58	56 ± 0
7	P(meta) + Rh	0.15 / 0.30	311 ± 27	74 ± 1
8	P(meta) + Rh	0.15 / 1.7	521 ± 33	76 ± 0
9	P(para) + Rh	0.14 / 0.02	467 ± 73	57 ± 0
10	P(para) + Rh	0.14 / 0.08	486 ± 91	58 ± 1
11	P(para) + Rh	0.14 / 0.16	238 ± 17	74 ± 2
12	P(para) + Rh	0.14 / 1.5	487 ± 43	78 ± 0

a: 80 bar syn gas (1:1), 41h 15 min., 35 °C, stirred at 625 rpm. 0.5 mL 1-octene containing internal standard. Each entry represents the average of four experiments. **b**: P(ortho) = 2-(diphenylphosphino) benzaldehyde, P(meta) = 3-(diphenylphosphino) benzaldehyde, P(para) = 4-(diphenylphosphino) benzaldehyde **c**: Actual molar amounts used (in µmol) **d**: Determined by GC

When an excess of rhodium is present TON are about 300 - 700 and the % linear selectivity is in the range 54 - 59% which corresponds to the linear selectivity of free rhodium. Upon increasing the P/Rh ratio to values larger than 1 the linear selectivity increases from 56% to 68% for P(ortho), from 55% to 76% for P(meta), and from 57% to 78% for P(para). For the experiments with an excess of phosphine (except entry 4) TON is in the range of 200-500. The only exception (entry 4) is the experiment with the highest ratio of P(ortho) to Rh in which virtually no conversion took place.

Upon increasing the P/Rh ratio to values larger than 1 the selectivity increases to around 75% linear selectivity for P(para) and P(meta). In case of P(ortho) a larger excess of

phosphine is necessary to reach a linear selectivity of around 70%. About 200-500 TON are detected, the only exception is the experiment with the highest ratio of P(ortho) to Rh in which virtually no conversion took place.

For P(meta) and P(para) there are no systematic differences in activities. This is in total contrast with the Rh-enzymes investigated for which the activity strongly depended on the substitution pattern of the phosphines.

The activity and selectivity of the biphasic hydroformylation of 1-octene using SCP-2L A100C-1-P-Rh (with P being either P(para) or P(meta)) is similar to the results of the monophasic reaction when P(para) or P(meta) are used in a slight excess.

The selectivity of the biphasic hydroformylation of 1-octene using SCP-2L V83C-1-P-Rh (with P being either P(para) or P(meta)) is similar to the selectivity of the monophasic reaction when P(para) or P(meta) are used in a slight excess while the activity of the monophasic reaction is much higher.

The results obtained for the Rh-enzymes can be explained by fragmentation (or back reaction) of the Rh-enzyme and release of a P-Rh-fragment as the P-Rh-fragments investigated have a selectivity comparable with the selectivity observed for the Rh-enzymes. In case of SCP-2L A100C-1-P(para-Rh basically all of the Rh coordinating to the enzyme needs to be released as a P-Rh-fragment to obtain the detected activity. In case of the other Rh-enzymes a fraction of the protein bound rhodium released as a P-Rh-fragment would be sufficient to cause the activity observed in the Rh-enzyme catalysed reaction.

A different fraction of P-Rh-fragment release for SCP-2L A100C-1-P(para)-Rh and SCP-2L V83C-1-P(para)-Rh seems unlikely as the chemical modification in those two Rh-enzymes is the same. Additionally the back reaction should be the same regardless of the phosphine used in protein modification. This would result in the same quantities of P-Rh-fragments being released resulting in similar activities for the different Rh-enzymes regardless of the phosphine used. In reality the activity of the Rh-enzymes was highly dependant on the phosphine it was based upon.

From this background it is unlikely that the activities and selectivities of the Rh-enzymes are only caused by the degradation or by the back reaction of the Rh-enzyme and the subsequent release of a Rh-containing fragment into the organic phase.

However, one cannot exclude that these fragments are partially responsible for the catalysis.

ICP-MS analysis of the organic phase

The rhodium and phosphorus contents of the organic phase after catalysis with different Rh-enzymes were determined by ICP-MS. The phosphorus concentrations detected were quite low (up to 200 ppb, all details in the experimental section). This is a strong indication that any potentially existing P-Rh-fragments do not play an important role as catalysts.

The rhodium concentration was analysed in the same way but on more samples. No correlation between the 1-alkene used and the rhodium concentration leached into the organic phase could be found (details in experimental section). In table 6 rhodium concentrations from hydroformylation experiments in which different Rh-enzymes were used are summarized. The rhodium concentrations in the organic phase were determined by ICP-MS after the hydroformylation reaction. Also the ratios of rhodium concentrations in the organic phase over the corresponding initial rhodium concentrations in the aqueous phase before the reaction are listed (expressed in %).

Table 6 Rhodium content in the organic phase after hydroformylation.[a]

entry	Rh-enzyme	ppb Rh[b]	% Rh[c]	n[d]
1	A100C-1-P(para)-Rh	138±69	5.8±2.9	6
2	A100C-1-P(meta)-Rh	158±156	2.0±2.0	3
3	A100C-1-P(ortho)-Rh	15±1	0.3±0.0	2
4	V83C-1-P(para)-Rh	22±13	0.5±0.3	6
5	V83C-1-P(meta)-Rh	261±127	2.6±1.3	5
6	V83C-1-P(ortho)-Rh	43±-24	1.4±0.8	2

a: All details are described in the experimental section. **b**: determined in organic phase by ICP-MS. **c**: ratio of the Rh concentration in the organic phase after the reaction over the initial Rh concentration in the aqueous phase (as determined by ICP-MS) **d**: Number of experiments.

The rhodium concentrations detected in the organic phase seems to depend on the initial rhodium concentration in the aqueous phase as well as on the ratio between phosphine modified protein to rhodium (see experimental section, details on table 6). Neither of which was constant for the different Rh-enzymes listed above. The standard deviation

for the same reaction settings is quite large as well. The data obtained is insufficient for an in depth comparison but some conclusions can be drawn:

Only small quantities of rhodium were detected in the organic phase after the hydroformylation reactions. Although these quantities correspond to up to 6% of the initial amount of rhodium present in the aqueous phase they cannot solely explain the high conversion observed when SCP-2L based Rh-enzymes were used as catalysts unless the fragment leached is much more active than the one tested in the monophasic reaction or the polarity of the fragment released into the organic phase is pressure and/or temperature dependant resulting in a higher concentration during the reaction than after the reaction when the pressure is released and the temperature reduced (see "smart systems" in the introduction).

Investigation of the kinetics

This set of experiments investigates the potential existence of such a “smart” system mentioned before. Parallel reactions using the same batch of SCP-2L V83C-1-P(para) were used to catalyse the hydroformylation of 1-octene for different reaction times. If a temperature or pressure dependant P-Rh-fragment (a fragment which is soluble in the organic phase during the reaction but migrates into the aqueous solution at the end of the reaction) is causing the activity observed the activity needs to increase over time as the quantity of P-Rh fragment would accumulate in the organic phase. The activity might also stay constant if the P-Rh-fragment is formed instantly at the beginning of the reaction. If the protein scaffold is playing a role the activity is more likely to decrease over time, as the enzyme is precipitating over time. The relative activity and % linear selectivity for the different reaction times using SCP-2L V83C-1-P(para)-Rh as catalyst in the hydroformylation of 1-octene are depicted in figure 8.

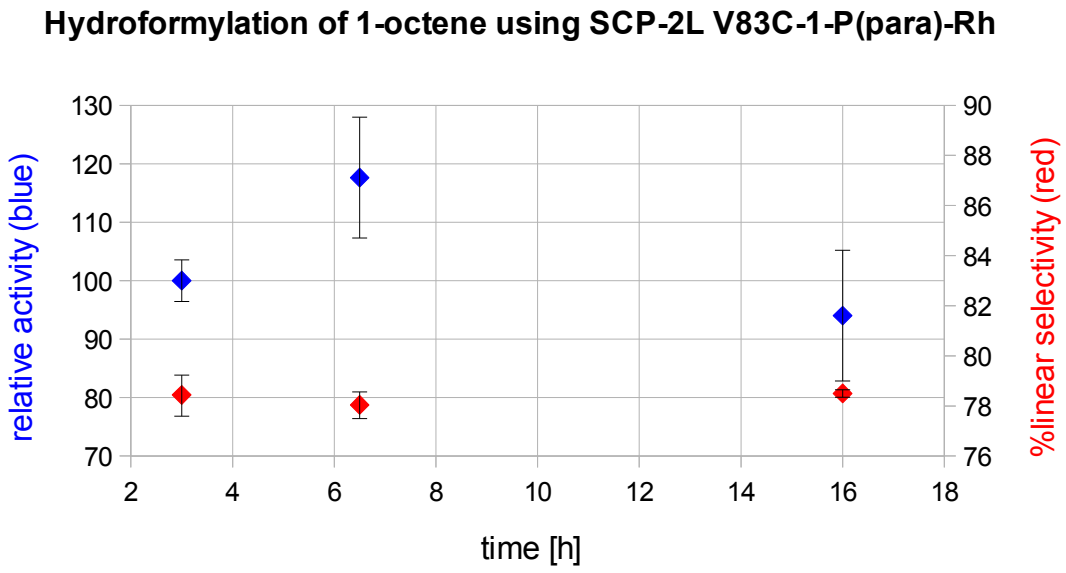


Figure 8 Dependency of the relative activity and selectivity in the hydroformylation of 1-octene over time using SCP-2L V83C-1-P(para)-Rh as catalyst.

Reaction conditions: 80 bar syn gas (1:1), 48h, 35 °C, stirred at 625 rpm. Organic solution: 0.5 mL 1-octene containing 9% (v/v) n-heptane and 1% (v/v) diphenylether. Aqueous solution: 0.5 mL 20 mM MES, 50 mM NaCl, pH=6. TON & linear selectivity determined by GC. 46.58 nanomol Rh (determined by ICP-MS). SCP-2L V83C-1-P(para) : Rh = 3.0; Each data point represents the average of three experiments. The first data point is set to a relative activity of 100.

The linear selectivity is around 78% for all measurement points. The relative activity is set to 100 for the reaction time of 3.0 h. The second reading point after a reaction time of 6.5h has a slightly higher activity while the last reading point after 16h reaction time shows a slightly lower activity. The increase in activity between the first two reaction times can be explained by an incubation period required to obtain a catalytically active species (see scheme 1 – mechanism of the hydroformylation). The decrease in relative activity between the second and third reaction time can be explained by a decrease in catalytically active species or by a decreasing 1-octene concentration if catalysis is faster than the phase transfer as the other factors (CO pressure, H_2 pressure, temperature) are virtually constant. The existence of a "smart system" can be excluded and an involvement of the Rh-enzyme in the catalytic cycle is quite likely.

Conclusion on this mode of action

When Rh-enzymes based on SCP-2L are used as catalysts in the hydroformylation reaction only a marginal fraction of the total hydroformylation activity can be attributed to either "free" rhodium or a Rh-phosphine-fragment leached into the organic phase. The Rh-enzymes do not act like a “smart system” but probably play a role in the catalytic cycle.

2) Enrichment of 1-alkenes within the hydrophobic tunnel

The concept envisioned for artificial metalloenzymes based on the SCP-2L scaffold was already discussed earlier (figure 3): The hydrophobic substrate does enrich within the tunnel during the reaction bringing it in close proximity to the metal centre. As a consequence of the increased concentration of 1-alkene an activity increase induced by the Rh-enzyme would be observed on the one hand. On the other hand a selectivity inducement by the whole protein environment would be expected.

Catalysis by different protein scaffolds

In the concept depicted in figure 3 the tunnel plays a decisive roll in the catalytic activity of the enzyme. In order to elucidate the significance of the tunnel a protein scaffold without a hydrophobic tunnel was used. Photoactive yellow protein (PYP) does not have a tunnel but contains cysteine. The PYP R52G mutant was used and modified to obtain the Rh-enzyme PYP R52G-1-P(para)-Rh. This Rh-enzyme was subsequently used as a catalyst in hydroformylation experiments in which the chain lengths of the 1-alkenes were varied between C6 to C18 (see table 7).

Table 7 Hydroformylation of 1-alkenes using SCP-2L V83C-Phosphine-Rha

entry	catalyst	substrate[b]	TON[c]	% linear selectivity[c]
1	PYP R52G-1-P(para)-Rh[d]	1-hexene	285 ± 15	78 ± 1
2	PYP R52G-1-P(para)-Rh[d]	1-octene	101 ± 6	76 ± 0
3	PYP R52G-1-P(para)-Rh[d]	1-decene	116	56
4	PYP R52G-1-P(para)-Rh[d]	1-dodecene	138	56
5	PYP R52G-1-P(para)-Rh[d]	1-octadecene	33	76

a: 80 bar syn gas (1:1), 48h, 35 °C, stirred at 625 rpm. 0.5 mL aqueous and organic solution each. **b**: Substrate containing 9% (v/v) n-heptane and 1% (v/v) diphenylether **c**: determined by GC **d**: ≤140 nanomol Rh (not determined by ICP-MS). PYP R52G-1-P(para) : Rh ≥ 2; the Conversion of 1-hexene and 1-octene are based on two experiments, all other entrys are based on one experiment.

The % linear selectivity is around 75% for 1-hexene, 1-octene and 1-octadecene. It is around 55% for 1-decene and 1-dodecene. The activity decreases with increasing chain lengths with 1-decene and 1-dodecene being the exception to this trend. The low selectivity and comparably high activity for 1-decene and 1-dodecene is most likely due to higher quantities of "free" rhodium in the organic phase for these reactions and will not be considered any further (very likely oxidation of the phosphine due to “old” batches of both 1-decene and 1-dodecene). The selectivity for this PYP based Rh-enzyme is similar to the selectivity obtained for the two corresponding SCP-2L based Rh-enzymes (tables 3 and 4) and its activity is in between them. Although neither PYP nor the mutant PYP R52G exhibit a hydrophobic tunnel as SCP-2L they possess a hydrophobic pocket next to their cysteine and therefore next to the introduced modification.[50]

In order to exclude that this hydrophobic pocket has a similar effect as the hydrophobic tunnel of the SCP-2L an attempt on the synthesis of another class of Rh-enzymes based on "Lysozyme" was performed. Unfortunately they could not be characterised successfully (see chapter 2) and when tested in the hydroformylation reaction the results obtained are quite disappointing as the conversion is barely detectable (<1 TON) turning any conclusions based on these results into sheer speculations.

Competitive binding

In order to verify the assumption that the high activities are due to the accumulation of 1-alkene within the hydrophobic tunnel of the protein competitive binding reactions

were performed to block the tunnel. Three different sets of experiments were performed in which either palmitic acid or pyrene decanoic acid were used. The experiments with palmitic acid were performed with the two Rh-enzymes SCP-2L V83C-1-P(para)-Rh and SCP-2L A100C-1-P(para)-Rh whereas in the series with pyrene decanoic acid only SCP-2L V83C-1-P(para)-Rh was used.

In the first set of experiments the Rh-enzyme SCP-2L V83C-1-P(para)-Rh was used and the amount of palmitic acid was varied (table 8). Although it is known that different SCP-2 enzymes bind palmitic acid inside the hydrophobic tunnel[51–57] the binding affinity of the modified proteins under reaction conditions is unknown. If the assumed principle is correct and the binding affinity towards palmitic acid is higher than to 1-octene increasing amounts of palmitic acid would decrease the hydroformylation activity as the tunnel is blocked. Additionally a lower selectivity would seem likely as well as the outer reaction sphere of the protein cannot influence the outcome of the reaction anymore. The influence of different quantities of palmitic acid in the biphasic hydroformylation of 1-octene with SCP 2L V83C-1-P(para)-Rh as catalyst are depicted in table 8.

Table 8 Hydroformylation of 1-octene using SCP-2L V83C-1-P(para)-Rh at different palmitic acid quantities.[a]

entry	eq. palmitic acid[b]	rel. activity[c,d]	% linear selectivity[d]
1	0	100±23	77±2
2	0.5	86	75
3	1	63	76
4	5	72±3	74±1
5	100	70	75

a: 80 bar syn gas (1:1), 48h, 35 °C, stirred at 625 rpm. Organic solution: Total of 0.5 mL 1-octene containing 9% (v/v) n-heptane and 1% (v/v) diphenylether. Aqueous solution: 0.5 mL 20 mM MES, 50 mM NaCl, pH=6. Maximum of 185 nanomol Rh per reaction (actual value after work-up not determined) SCP-2L V83C-1-P(para) : Rh ≥ 2.0. Entries 1 & 4 represent the average of two experiments, other entries based on single experiments **b**: equivalent palmitic acid, added from a stock in 1-octene and internal standard (total volume of organic solution is constant) **c**: relative activity in entry 1 is set to 100,**d**: determined by GC

In table 8 the highest activity is found when no palmitic acid is present. It is still high when sub-stoichiometric amounts are used but drops to an activity of about 60-70% when equimolar or higher equivalents of palmitic acid are present. The change in linear

selectivity is not significant.

A similar set of reactions was performed using pyrene decanoic acid instead of palmitic acid with the same Rh-enzyme (table 9). Pyrene decanoic acid is a fluorescent fatty acid analogue known to bind to the hydrophobic tunnel of the mutant SCP-2L V83C as well as to SCP-2L itself.[33]

Table 9 Hydroformylation of 1-octene using SCP-2L V83C-1-P(para)-Rh at different pyrene decanoic acid quantities.[a]

entry	eq.[b]	rel. activity[c,d]	% linear[d]
1	0	100	76
2	1	75	71
3	5	100	76
4	10	198	77

a: 80 bar syn gas (1:1), 48h, 35 °C, stirred at 625 rpm. Organic solution: Total of 0.5 mL 1-octene containing 9% (v/v) n-heptane and 1% (v/v) diphenylether, 40 µL methanol. Pyrene decanoic acid dissolved in the organic mixture. Aqueous solution: 0.5 mL 20 mM MES, 50 mM NaCl, pH=6. Maximum of 150 nanomol Rh per reaction (actual value after work-up not determined) SCP-2L V83C-1-P(para) : Rh ≥ 1.2. All entries based on a single experiments, **b**: Equivalent pyrene decanoic acid **c**: relative activity in entry 1 is set to 100, **d**: determined by GC

According to table 9 the selectivity is basically unaffected by the quantity of pyrene decanoic acid added. Additionally the relative activity is almost unchanged for up to 5 equivalents of pyrene decanoic acid. When 10 equivalents are added the relative activity increases by a factor two.

The influence of palmitic acid on the performance of SCP 2L A100C-1-P(para)-Rh in the hydroformylation of 1-octene is shown in table 10.

Table 10 Hydroformylation of 1-octene using SCP-2L A100C-1-P(para)-Rh at different palmitic acid quantities.[a]

entry	eq. palmitic acid[b]	rel. activity[c, d]	% linear selectivity[d]
1	0	100±2	80±0
2	0.1	84	89
3	0.5	107	82
4	1	81	80
5	5	166	84
6	10	164	82
7	100	165	81

a: 80 bar syn gas (1:1), 48h, 35 °C, stirred at 625 rpm. Organic solution: Total of 0.5 mL 1-octene containing 9% (v/v) n-heptane and 1% (v/v) diphenylether. Aqueous solution: 0.5 mL 20 mM MES, 50 mM NaCl, pH=6. ≤450 nanomol Rh per reaction (actual value after pre-formation and work-up not determined by ICP-MS) SCP-2L A100C-1-P(para) : Rh ≥ 2.0. Entry 1 represent the average of two experiments, other entries based on single experiments **b**: equivalent palmitic acid, added from a stock in 1-octene and internal standard (total volume of organic solution is constant) **c**: relative activity in entry 1 is set to 100 **d**: determined by GC

As can be seen from a comparison of the results from table 8 and 10 increasing amounts of palmitic acids cause opposite effects regarding the activity for the two tested Rh-enzymes. Pyrene decanoic in the presence of SCP-2L V83C-1-P(para)-Rh has a similar effect on the activity as palmitic acid on SCP-2L A100C-1-P(para)-Rh. These contradicting results make it impossible to draw conclusions on the potential mode of the tunnel within the protein.

Enantioselectivity

If the reaction takes place in the tunnel the chirality of the protein scaffold might induce chirality on the product. Unfortunately the hydroformylation of 1-alkene favours the formation of linear aldehydes while the (potentially chiral) branched aldehyde is only formed in small quantities. Therefore the hydroformylation of styrene using Rh-enzymes was performed as styrene is known to yield mainly the (potentially) chiral branched aldehyde under hydroformylation conditions.

When SCP-2L A100C-1-P(para)-Rh was used as catalyst extremely high branched selectivities of 98% were obtained. When SCP-2L V83C-1-P(para)-Rh was used as catalyst similarly high branched selectivities of around 94% were obtained. Unfortunately no enantiomeric excess of the formed 2-phenylpropionaldehyde was

found in either reaction.
It might be that the hydroformylation of styrene is totally different from the hydroformylation of 1-alkenes using Rh-enzymes with the former substrate not interacting with the tunnel while the latter ones are. It might also be that the reaction takes place within the tunnel but is not inducing any enantioselectivity.

Conclusion on this mode of action

One of the competitive binding experiments showed a decrease in activity with increasing concentration of competitive binder. This can be explained by competitive binding inside the tunnel. On the contrary the other two competitive binding experiments did not result in a decrease of the activities. Instead the activities increase with higher loadings of the competitive binder. This increase can be explained by the formation of micelles in the aqueous phase, by an increase of the stability of the Rh-enzymes or by the reaction taking place outside the enzyme / on the surface.
The tested Rh-enzyme based on PYP has a similar reaction rate as the SCP-2L based Rh-enzymes for most 1-alkenes tested. Although it does not possess a hydrophobic tunnel it is possible that its hydrophobic pocket has the same effect.
These results show that the concept envisioned in which the hydrophobic tunnel acts as a accumulation point for the substrates is possible.

3) Protein acting as amphiphile

The theory that the Rh-enzyme is acting as an amphiphile emerged during the workup of the Diels-Alder reactions (see chapter 4) where the phase separation after the reaction was problematic. This indicates the presence of surface active molecules. Due to this observation a special hydroformylation reaction was performed in which the reaction time was kept short and the autoclave was not put on ice when releasing the pressure (as usually). Some reaction vials of this experimental setup showed a bad phase separation, indicating that the Rh-enzymes might act as surface active compounds under the reaction conditions used. The most simplified approach is to consider the hydrophobic modification as oriented into the organic phase while the main part of the protein is

located in the aqueous phase (see figure 9). This was further investigated.

The decreasing activity with increasing chain length of the 1-alkene which was observed could also be explained by a decreasing concentration of the 1-alkenes with increasing chain length. as they were used neat at a constant total volume of the organic phase.

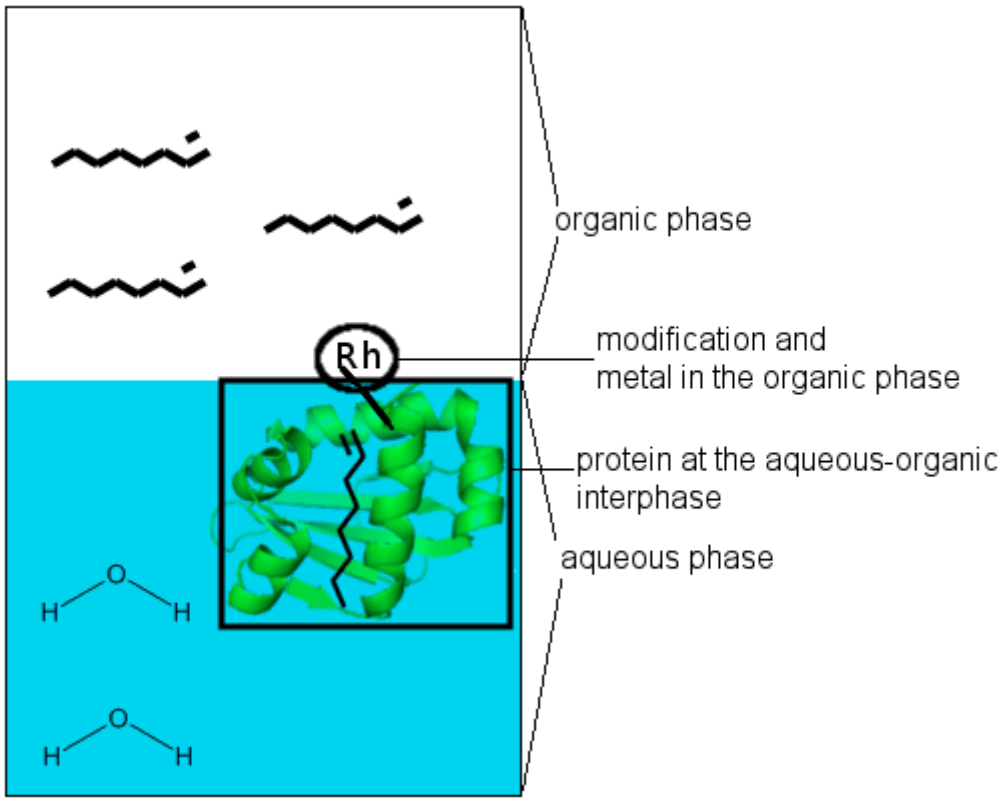


Figure 9 Rh-enzyme acting as amphiphile.

Variation of the Rh-enzyme concentration

In order to test this hypothesis increasing Rh-enzyme concentrations were used to catalyse the reaction. Once there is rhodium saturation at the interphase further increasing its concentration would not result in an increasing TON but a decreasing. The results for the hydroformylation of 1-octene using SCP-2L V83C-1-P(para) at different Rh-enzyme concentrations is depicted in figure 10.

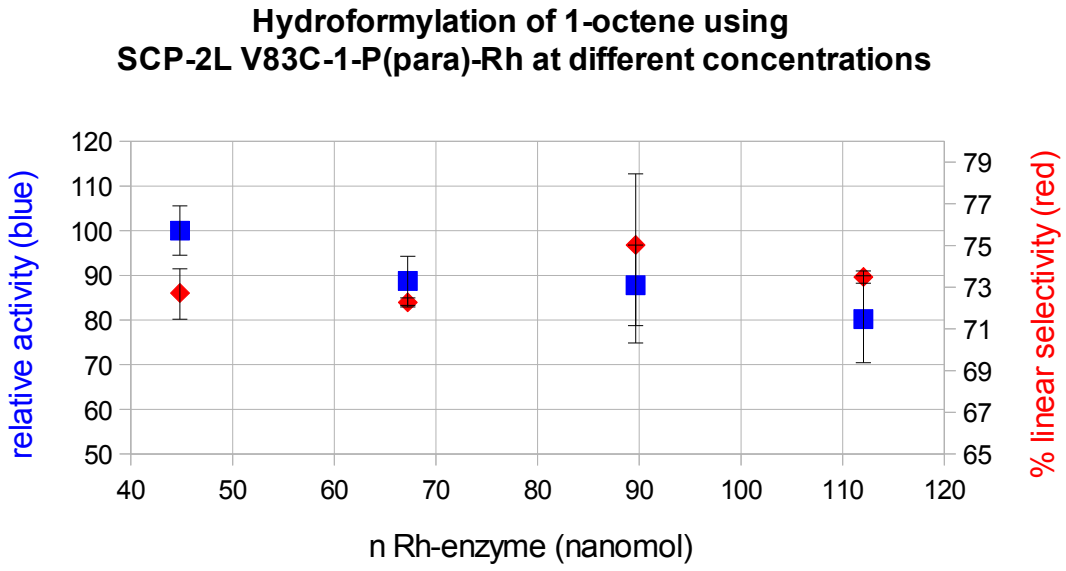


Figure 10 Hydroformylation of 1-octene using SCP-2L V83C-1-P(para)-Rh at different Rh-enzyme concentrations.

Reaction conditions: 80 bar syn gas (1:1), 48h, 35 °C, stirred at 625 rpm. Organic solution: Total volume of 0.5 mL. 1-octene containing 9% (v/v) n-heptane and 1% (v/v) diphenylether, aqueous solution: 0.5 mL 20 mM MES, 50 mM NaCl, pH=6. conversion & linear selectivity determined by GC. SCP-2L V83C-1-P(para) : Rh = 2.9. The Rhodium quantities used are (in nanomol): 44.82; 67.22; 89.63; 112.04; all data points are based on four experiment. The relative activity of the firs data point is set to 100.

The activity decreases with increasing concentration of the Rh-enzyme while the selectivity is basically unaffected. This could be explained with the saturation of Rh-enzymes at the interphase. Unfortunately, it was observed that the stability of the Rh-enzyme is decreasing with increasing concentration.

Competitive hydroformylation of 1-alkene mixtures

When 1-alkene mixtures with the same concentrations of each 1-alkene are used as feedstock the conversion should be unrelated to the chain length if the protein is acting as amphiphile. The results for the hydroformylation of 1-alkene mixtures with SCP-2L A100C-1-P(para)-Rh is listed in table 11.

Table 11 Hydroformylation of 1-alkene mixtures in an aqueous-organic biphasic environment using SCP-2L A100C-1-P(para)-Rh as catalyst.[a]

		1-octene		1-decene		1-dodecene		1-octadecene	
Starting mix[b]	c[mol/mL][c]	activity[d]	%linear	activity[d]	%linear	activity[d]	%linear	activity[d]	%linear
8&10	2.9	100±11	81.3±1.8	55±5	75.9±0.3				
8&12	2.6	98±4	82.0±0.3			29±4	73.1±1.5		
8&18	2.1	110±3	82.1±0.4					10±0	73.0±0.1
8&10&12&18	1.1	57±12	82.0±2.5	24±2	74.2±0.9	16±2	75.5±2.4	4±1	74.1±0.6

a: 80 bar syn gas (1:1), 48h, 35 °C, stirred at 625 rpm. Organic solution: Total of 0.5 mL 1-alkene mixture containing 9% (v/v) n-heptane and 1% (v/v) diphenylether. aqueous solution: 0.5 mL 20 mM MES, 50 mM NaCl, pH=6. ≤150 nanomol Rh per reaction (actual value after preformation and workup not determined by ICP-MS) SCP-2L A100C-1-P(para) : Rh ≥ 1.1. Each entry represents the average of two experiments. **b**: Carbon chain length of the linear 1-alkenes used **c**: concentration of each 1-alkene in the starting mixture **d**: Conversion of the 1-alkene to aldehydes in relation to the conversion of 1-octene to the corresponding aldehydes in the 1-octene & 1- decene mixture.

The relative activity for the hydroformylation of 1-octene is unchanged regardless of the second 1-alkene present in the starting mixture. Additonally the activity is unchanged when lower concentrations of 1-octene are used. The only exception when the starting mixture contains all four different 1-alkenes. In this case the activity of 1-octene decreases dramatically to almost half its initial activity, while the 1-octene concentration is at about 40% of the initial concentration. The activity of the second 1-alkene is highly dependant on the chain length of the second 1-alkene. Shorter 1-alkenes are more active than longer ones.

No conclusions could be drawn when SCP-2L V83C-1-P(para)-Rh was used as catalyst in a similar set of reactions as the conversion is much lower when this Rh-enzyme is used. An attempt on the hydroformylation of diluted 1-octene (in THF) using SCP-2L V83C-1-P(para)-Rh was inconclusive. Therefore no attempt on the hydroformylation of different 1-alkene mixtures diluted to the same concentration was performed.

Conclusion on this mode of action

The reactions performed make it seem highly unlikely that the protein is acting as an amphiphile with the reaction taking place at the interphase as the activity is affected by the 1-alkene chain length of the 1-alkene.

Various reactions

Protein stability

The aqueous solutions of the unmodified proteins SCP-2L V83C and SCP-2L A100C lack stability under the storage conditions used (for example 20 mM MES, 50 mM NaCl, pH=6, temperature either at room temperature or 4 °C) resulting in approximately 10-30% protein loss due to precipitation every week. The chemical modifications applied decrease the protein stability even further. Working at lower protein concentrations would increase the stability but is impractical. Parallel reactions showed that neither the high syn gas pressure (80 bar), nor the temperature (35 °C), nor the organic reactant present during the reaction have a noticeable effect on the stability. The mechanical influence of the stirring (with a stirring bar) results in a much faster precipitation of the enzymes (whether modified or not) than no such mechanical disturbance.

Influences of batch storage time and H_2/CO pressure

A good example for encountered problems are the hydroformylation experiments in which the ratio of reaction gases was varied and SCP-2L V83C-1-P(para)-Rh acted as catalyst. In the first set of experiments a ratio of 10 bar hydrogen to 40 bar CO for one autoclave and 20:40 for the other autoclave was used. After the reaction the same batch of Rh-enzyme (which had been stored at 4°C in the meantime) was immediately used to obtain data points for hydrogen/CO-ratios of 30:40 and 40:40. The results are depicted in figure 11.

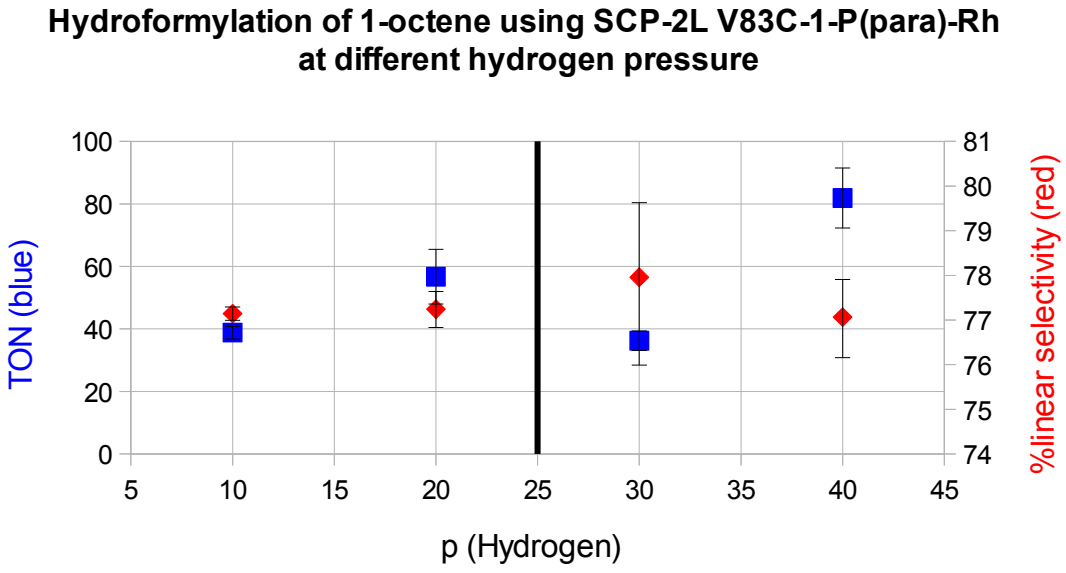


Figure 11 Hydroformylation of 1-octene using SCP-2L V83C-1-P(para)-Rh at different H_2 pressure. **Reaction conditions**: 40 bar CO topped up with the appropriate amount of H_2 to a total pressure of 50, 60, 70, 80 bar respectively, 16h 10 min. reaction time, 35 °C, stirred at 625 rpm. Organic solution: Total volume of 0.5 mL. 1-octene containing 9% (v/v) n-heptane and 1% (v/v) diphenylether, aqueous solution: 0.5 mL 20 mM MES, 50 mM NaCl, pH=6. conversion & linear selectivity determined by GC. SCP-2L V83C-1-P(para) : Rh = 4.9; 19.38 nanomol Rhodium; each data point represents the average of four experiments.

The % linear selectivity is constant at about 78% indicating that both the total pressure, as well as the increasing partial pressure of hydrogen do not influence linear selectivity. The activity on the other hand increases from 10 to 20 bar hydrogen as well as from 30 to 40 bar. The activity drop from 20 to 30 bar hydrogen is therefore most likely due to catalyst storage (1 day) in between the two runs resulting in a lower activity for the stored catalyst.

The influence of the partial CO pressure was investigated in the same way (see figure 12). One would expect a decreasing activity with increasing carbon monoxide pressure as the standard hydroformylation activity has a negative first order in carbon monoxide pressure. The results are depicted in figure 12.

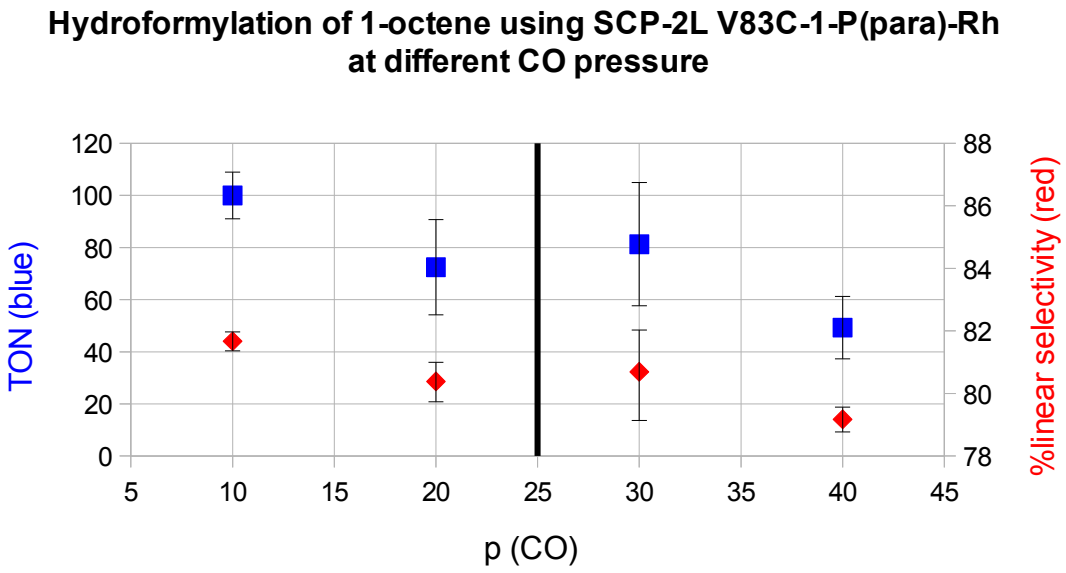


Figure 12 Hydroformylation of 1-octene using SCP-2L V83C-1-P(para)-Rh at different CO pressure.

Reaction conditions: 40 bar H_2 topped up with the appropriate amount of CO to a total pressure of 50, 60, 70, 80 bar respectively, 16h reaction time, 35 °C, stirred at 625 rpm. Organic solution: Total volume of 0.5 mL. 1-octene containing 9% (v/v) n-heptane and 1% (v/v) diphenylether, aqueous solution: 0.5 mL 20 mM MES, 50 mM NaCl, pH=6. conversion & linear selectivity determined by GC. SCP-2L V83C-1-P(para) : Rh = 4.9; 19.38 nanomol Rhodium; each data point represents the average of four experiments.

An increase in CO (and total) pressure slightly reduces the % linear selectivity from 82% to 79%. The total pressure does not influence the % linear selectivity (see figure 13), suggesting a weak inverse interaction between the CO pressure and the linear selectivity. The activity drops when increasing from 10 to 20 bar of CO and again when increasing from 30 to 40 bar CO. This time the activity increases from 20 to 30 bar, indicating that the catalyst stored for 1 day is more active which is exactly the opposite effect as seen before.

These two sets of experiments clearly highlight a major issue of the Rh-enzyme: It is almost impossible to compare different sets of experiments when the simple storage of the catalyst for such a short time already has a significant effect on the activity of the catalyst.

Reusability of Rh-enzymes

The reusability of Rh-enzymes was tested with SCP-2L V83C-1-P(para)-Rh. One portion of the Rh-enzyme batch was used in the hydroformylation while the other portion of the batch was stored at 4 °C. After the reaction the aqueous phase from sixteen reactions was combined, filtered, washed and reconcentrated to the same protein concentration as the portion of the stored batch. It was then reused as catalyst in a new round of hydroformylation experiments. The Rh-enzyme which had been stored at 4°C in the meantime was used as catalyst as well. LC-MS analysis performed on both batches prior to the second round of hydroformylation experiments showed the presence of Rh-enzyme in the catalyst solution which was stored while the catalyst solution which had been used in catalysis before only contained oxides of SCP-2L V83C-1-P(para). The hydroformylation results were in accordance with the LC-MS results. The reused Rh-enzyme yielded 68%-linear selectivity indicating high quantities of leached rhodium. The Rh-enzyme which was stored yielded a linear selectivity of 77%.

This clearly shows that the reusability of the Rh-enzymes is not only limited by the loss of Rh-enzyme due to precipitation but also by a lower selectivity, which might be attributed to oxidation of the phosphine during the workup.

3.4 Conclusions and future work

The aqueous organic hydroformylation of long chain 1-alkenes using different Rh-enzymes was investigated in this chapter. The main focus was on Rh-enzymes based on the SCP-2L scaffold. The best Rh-enzyme tested (SCP-2L A100C-1-P(para) showed rate accelerations of several orders of magnitude in comparison to the corresponding Rh/TPPTS system. In addition, a high linear selectivity was detected. The actual mode of action for these Rh-enzymes could not be identified but some potential modes were excluded: The leaching of either rhodium, or a Rh-P-fragment into the organic phase as well as the protein acting as an amphiphile. The previously envisioned role of the hydrophobic tunnel incorporated within the protein scaffold as a substrate rich environment which helps accelerating the reaction seems possible but could not be verified.

Future work should consider the investigation of a Rh-enzyme based on SCP-2L with a cysteine mutation neither near or inside the tunnel, nor next to a hydrophobic pocket. Testing a Rh-enzyme of such a mutant in the hydroformylation would determine whether the hydrophobic tunnel is indeed having the (significant) role in the reaction. It might also be worthwhile to investigate the (potential) formation of micelles under hydroformylation conditions in the presence of Rh-enzymes by light scattering.

If further studies on SCP-2L based catalysts are performed it will be worthwhile to address the stability of the protein. Besides time consuming biological approaches it might be feasible to address the mechanical disturbance by stirring. This was determined as the major factor causing protein precipitation. It might be worthwhile to investigate the stability of the Rh-enzymes when shaking instead of stirring. In case they are more stable under these conditions the general hydroformylation setup can be adjusted. In order to evaluate whether the tunnel within SCP-2L is playing any role in the most straightforward is the use of a new mutant, in which the introduced cysteine is not in proximity to either end of the tunnel.

3.5 Experimental

Protein expression, purification and modification

Protein expression and purification was performed as described in chapter 2.

Protein modification, synthesis of Rh-enzymes and the analysis of (un)modified proteins was performed as described in chapter 2.

ICP-MS

The rhodium and phosphorous concentrations of the organic phase after hydroformylation as well as the rhodium concentration in the aqueous solution before the hydroformylation were determined by ICP-MS (see chapter 2 for more details).

GC

A CE instrument GC8000 top equipped with a AS800 autosampler was used for GC analysis. The conversion of the 1-alkenes was determined by using calibrations of the linear aldehydes (heptanal, nonanal, undecanal, tridecanal, nonadecanal) against the internal standards heptane and diphenylether. The calibrations obtained were used for both the linear and the corresponding branched aldehydes. Chirality of the 2-Phenylpropionaldehyde was analysed by GC using a Supelco β-DEX 225 column.[58]

MALDI-TOF

MALDI-TOF analysis of the Rh-enzyme was performed by Dr. Catherine Botting on an ABSciex 4800 MALDI TOF/TOF Analyser

Calculation of the activity differences between Rh/TPPTS and Rh-enzymes

The rate constants for the aqueous biphasic conversion of 1-octene and 1-decene at 125°C, 80 bar of syn gas and a rhodium concentration of 300-400 ppm are literature known[59] (p. 394 within the reference). The rate constant for 1-octene is $4.6*10^{-4}$ $min.^{-1}$, the one for 1-decene is $1.50*10^{-4}$ $min.^{-1}$. The literature known rate constants at the given temperature can be used to calculate the (approximate) rate constant at a different temperature.

The rate constant (r) for any reaction can be written as (Eq. 1)

Eq. (1): $$r = A \times e^{\left(\frac{-E_A}{R \times T}\right)}$$

where A is the pre-exponential factor, E_A the activation energy, R the gas constant and T the temperature. The ratio of two different rate constants r_1 and r_2 for the same reaction but different temperatures (T_1 and T_2) can be written as (Eq. 2)

Eq. (2): $$\frac{r_1}{r_2} = \frac{e^{\left(\frac{-E_A}{RT_1}\right)}}{e^{\left(\frac{-E_A}{RT_2}\right)}}$$

which can be rewritten as Eq. 3:

Eq. (3): $$\frac{r_1}{r_2} = e^{\frac{-E_A \times (T_2 - T_1)}{R \times T_1 \times T_2}}$$

The literature known rate constants at 125°C for Rh/TPPTS are used to calculate the rate constant at 35°C (as this was the operating temperature of the Rh-enzymes) using Eq. (3). The activation energy of 1-dodecene for the Rh/TPPTS system was determined as 62-73 kJ/mol.[60] The same activation energy will be used for the subsequent calculations regarding 1-octene and 1-decene.

Eq. (3) is used to calculate the ratio between rate constants at T_1=398.15K and T_2=308.15K. 8.314 $Jmol^{-1}K^{-1}$ is used for the gas constant R. When an activation energy of 62000 J/mol (lower limit) is used the rate constant at 308.15K (35°C) is a factor 237 smaller than the rate constant at 398.15K (125 °C). When an activation energy of 73000 kJ/mol (upper limit) is used the rate constant at 35 °C is by a factor of 627 smaller.

The reaction rate of the hydroformylation is first order in the concentrations of both rhodium and 1-alkene. It is also of a negative first order in the concentration of CO. While the 1-alkene concentration and the CO concentration (as well as the hydrogen concentration) are the same for the literature process and the Rh-enzymes the rhodium

concentration is much smaller in case of the Rh-enzymes. To compare the two systems this has to be taken into account. As already mentioned the rhodium concentration in the Rh/TPPTS system was between 300-400 ppm. When SCP-2L A100C-1-P(para)-Rh was used as catalyst rhodium concentrations of ≈ 46 nanomol*mL^{-1} are used (see table 4). This corresponds to ≈ 4.7 ppm Rh. The rhodium concentration in the Rh/TPPTS-literature-system is therefore by a factor of 64 to 85 higher. As mentioned the reaction rate of the hydroformylation is 1st order in rhodium concentration, therefore the reaction rates of the literature Rh/TPPTS system need to be reduced by a factor 64 to 85 to account for its higher rhodium concentration.

Multiplying the factor of the higher temperature (factor 237 to factor 627) with the factor for the higher rhodium concentration (factor 64 to factor 85) does yield the factor by which the Rh/TPPTS system has to be reduced in order to compare it with the Rh-enzyme. This factor is between ≈15000 to ≈53000. After multiplication of the literature given rate constants [$min.^{-1}$] by 60 the rate constant per hour [h^{-1}] is obtained. This value is corrected for both the higher temperature and the higher rhodium concentration by division through 15000 (lower limit) or 53000 (upper limit) and the corrected rate constants are obtained:

rate constant (Rh/TPPTS at 35°C, 4.7 ppm Rh, 1-octene, h^{-1}) = $5.2*10^{-7}$ to $1.8*10^{-6}$

rate constant (Rh/TPPTS at 35°C, 4.7 ppm Rh, 1-decene, h^{-1}) = $1.7*10^{-7}$ to $6.0*10^{-7}$

When Rh-enzymes are used as hydroformylation catalysts in this thesis the concentration of 1-alkene can be considered as constant as only marginal quantities are converted. The other relevant factors stay constant as well. Therefore the values above can be compared with the TON obtained when any Rh-enzyme was used as catalyst. As an example the data for the hydroformylation of 1-octene and 1-decene with SCP-2L A100C-1-P(para)-Rh as catalyst is used (see table 4). The TON listed in table 4 correspond to a reaction time of 48h. After dividing these values by 48 the TOF [h^{-1}] is obtained.

TOF (SCP-2L A100C-1-P(para)-Rh at 35°C, 4.7 ppm, 1-octene, h^{-1}) = 8.5

TOF (SCP-2L A100C-1-P(para)-Rh at 35°C, 4.7 ppm, 1-octene, $48h^{-1}$) = 2.8

When these values for the Rh-enzyme are compared to the values calculated for the Rh/TPPTS-system the activity of (this) Rh-enzyme is (at least) 10^6 higher than the

Rh/TPPTS system under similar reaction conditions.

Standard Hydroformylation setup and conditions for aqueous biphasic reactions

This setup and conditions were used unless stated otherwise.

Hydroformylation reactions were carried out in stainless steel autoclaves (up to two simultaneously) containing up to eight glass reaction vials (volume of approximately 5 mL) each. The reaction vials were equipped with mini stirring bars as well as a septum cap pierced with a needle in order to allow contact with the reaction gases and labelled. The so prepared vials were placed in the autoclave which was flushed three times with >20 bar argon before adding the chemicals. While charging the vials a positive argon pressure was maintained by placing a tube with constant argon flow into the bottom of the autoclave and switching off the suction of the fumehood. 0.5 mL of an aqueous solution (20 mM MES, 50 mM NaCl, pH=6, in case of PYP based Rh-enzymes pH=7 was used) containing the Rh-enzyme was added to each vial. An aliquot was used to determine the protein concentration by Bradford assay. A Protein-1-P:Rh ratio of about 1.5-4 was aimed for (the actual concentration and ratio was only determined once the Rh-ICP data was known, which was in general months after the actual experiment). 0.5 mL organic solution consisting of the 1-alkene (or 1-alkene mixtures), 9% (v/v) heptane and 1% (v/v) diphenylether as internal standard freshly filtered over silica to remove peroxides was added. In case of competitive binding experiments stock solution of the competitive binder in the appropriate (filtered) 1-alkene containing internal standards were used instead (in case of pyrene decanoic acid 40 µL methanol were added as well). Unless stated otherwise experiments were run in quadruple. An aliquot of the catalytic solution was withdrawn for ICP-MS analysis to determine the actual rhodium concentration used in the experiment. After charging the vials the suction was switched back on and the autoclave was flushed three times with >20 bar of syn gas and subsequently charged to the desired pressure (80 bar of syn gas in general). The autoclave was then placed into an oil bath which was preheated to the desired temperature (35 °C in general) and preset to the desired stirring speed (625 rpm). After the reaction time (48 h in general) the reaction was stopped by putting the autoclave on ice and slowly releasing the pressure. The organic phase was either filtered over a small

plug of silica before GC analysis or not filtered and analysed by ICP-MS to determine its rhodium content.

Hydroformylation using Rh/TPPTS

2.4 mg (9.4 µmol) $Rh(acac)(CO)_2$ and 183 mg (285 µmol, 30 eq) TPPTS * 3 Na * 4 H_20 were dissolved in 18 mL distilled and degassed methanol at room temperature for 8 days. The solvent was removed and the solid redissolved in 11.0 mL degassed buffer (20 mM MES, 50 mM NaCl, pH = 6) and used as aqueous catalyst solution. 261 nanomol Rh per reaction vial using the standard hydroformylation setup outlined above.

Hydroformylation using $Rh(acac)(CO)_2$ / Phosphinoaldehydes

In a glove box sixteen small glass vials were equipped with stirring bars. The appropriate amounts of $Rh(acac)(CO)_2$ and the phosphinoaldehyde were added from stock solutions (mixtures of freshly filtered 1-octene and internal standard (9 volume-% heptane, 1 volume-% diphenylether)). All reaction vials were topped up with a mixture of 1-octene containing internal standard to a total volume of 400 µL. Two autoclaves were flushed three times with argon (>20 bar). The vials were sealed and transferred to the autoclaves. The lid of each vial was pierced with a needle to allow contact with the syn gas. The autoclaves were carefully flushed three times with syn gas (>20 bar), pressurised to 80 bar and put in a preheated oilbath (35 °C). The reaction was stirred at 625 rpm and stopped after 41h 15 min. by putting the autoclaves on ice and slowly releasing the pressure.

Experimental details on table 6 (rhodium & phosphorous content in the organic phase after hydroformylation).

The organic phase was analysed for Rhodium after standard hydroformylation settings by ICP-MS. Different Rh-enzymes and different 1-alkenes were used. The rhodium content in the aqueous phase before the reaction was determined by ICP-MS as well. Protein concentration before the reaction was determined by Bradford assay. The rhodium content (unit is "ppb") in the organic phase after the reaction as determined by ICP-MS. Additionally the 1-alkene and the initial modified protein to rhodium ratios are

given. Due to evaporation of the organic phase between submission and measurement of the samples (matter of several months) the rhodium concentrations listed below are most likely significantly higher than in the initial volume especially for the shortest 1-alkene 1-octene (due to evaporation over this period of time). The phosphorous and rhodium concentrations in the organic phase before the reaction (=blanks) were not determined.

SCP-2L A100C-1-P(para)-Rh: SCP-2L A100C-1-P(para) : Rh = 1.5; 23.0 nanomol Rh per vial (before the reaction).

Rh [ppb] (in the organic phase after the reaction): 1-octene: 66.6; 89.8; 1-decene: 249.9; 1-dodecene: 178.4; 1-octadecene: 150.7; 94.9.

P [ppb] (in the organic phase after the reaction): 1-octene: 89.8

SCP-2L A100C-1-P(ortho)-Rh: SCP-2L A100C-1-P(ortho) : Rh = 3.5; 26.9 nanomol Rh per vial (before the reaction).

Rh [ppb] (in the organic phase after the reaction): 1-octene: 16.1; 1-dodecene: 14.1

P [ppb] (in the organic phase after the reaction): 1-octene: 16.1

SCP-2L A100C-1-P(meta)-Rh: SCP-2L A100C-1-P(meta) : Rh = 2.2; 37.7 nanomol Rh per vial (before the reaction).

Rh [ppb] (in the organic phase after the reaction): 1-octene: 156.6; 1-dodecene: 315.3; 1-octadecene: 2.7

P [ppb] (in the organic phase after the reaction): 1-octene: 156.6

SCP-2L V83C-1-P(para)-Rh: SCP-2L V83C-1-P(para) : Rh = 3.0; 46.6 nanomol Rh per vial (before the reaction).

Rh [ppb] (in the organic phase after the reaction): 1-octene: 25.2; 42.4; 1-decene: 19.2; 1-dodecene: 27.1; 1-octadecene: 3.9; 11.6.

P [ppb] (in the organic phase after the reaction): 1-octene: 42.4; 1-octadecene: 11.6

SCP-2L V83C-1-P(ortho)-Rh: SCP-2L V83C-1-P(ortho) : Rh = 5.8; 30.9 nanomol Rh per vial (before the reaction).

Rh [ppb] (in the organic phase after the reaction): 1-octene: 26.1; 59.3

P [ppb] (in the organic phase after the reaction): 1-octene: 59.3

SCP-2L V83C-1-P(meta)-Rh: SCP-2L V83C-1-P(meta) : Rh = 1.8; 99.1 nanomol Rh per vial (before the reaction).

Rh [ppb] (in the organic phase after the reaction): 1-octene: 179.7; 201.0; 1-decene: 440.9; 1-dodecene: 344.6; 1-octadecene: 137.8.
P [ppb] (in the organic phase after the reaction): 1-octene: 201.0

Experimental details on hydroformylations using Rh-eznymes based on Lysozyme

Lysozyme-$(1)_x$-P(para)-Rh: Protein : Rh = 2.0; 58.67 nanomol Rh per vial. Stopped after 20h reaction time. Four reactions on each of the following 1-alkenes: 1-octene, 1-decene, 1-dodecene and 1-octadecene.

Lysozyme-$(1)_x$-P(meta)-Rh: Protein : Rh = 1.8; 99.12 nanomol Rh per vial. Stopped after 24h and 35 min. Due to the previous results involving Lysozyme based Rh-enzymes hydroformylation was only tested on 1-octene (four reactions)

Lysozyme-$(1)_x$-P(ortho)-Rh: Protein : Rh = 2.0; 76.944 nanomol Rh per vial. Due to the previous results involving Lysozyme based Rh-enzymes hydroformylation was only tested on 1-octene (four reactions). Stopped after 20h and 10 min.

Lysozyme-$(1)_x$-Rh: Protein : Rh = 2.0; 73.14 nanomol Rh per vial. Due to the previous results involving Lysozyme based Rh-enzymes hydroformylation was only tested on 1-octene (four reactions). Stopped after 20h and 10 min.

Hydroformylation of styrene using SCP-2L A100C-1-P(para)-Rh

Standard hydroformylation conditions were used. SCP-2L A100C-1-P(para) : Rh = 2.0. 48.08 nanomol rhodium per reaction vial. Styrene was freshly filtered over alumina. Four experiments using neat styrene and four experiments using 0.5 mL of a mixture containing 2.7 mL styrene and 3 μL diphenylether were used.

Hydroformylation of styrene using SCP-2L V83C-1-P(para)-Rh

Standard hydroformylation conditions were used. SCP-2L V83C-1-P(para) : Rh = 5.9. 45.67 nanomol rhodium per reaction vial. Styrene was freshly filtered over alumina. Four experiments using neat styrene and three experiments using 0.5 mL of a mixture containing 2.7 mL styrene and 3 μL diphenylether were used.

Hydroformylation of 1-alkene mixtures using SCP-2L V83C-1-P(para)-Rh

Two seperate set of reactions (using the general hydroformylation setup) were performed in order to obtain data for this setup. Both yielded inconclusive results. The two settings used were:

1) SCP-2L V83C-1-P(para) : Rh ≥ 1.1. Maximum of 150 nanomol rhodium per reaction vial, not determined by ICP-MS. The equimolar 1-alkene mixtures were prepared just prior to the reaction using stocks of the freshly filtered (over alumina) corresponding 1-alkene (containing internal standard).
2) SCP-2L V83C-1-P(para) : Rh ≥ 1.2. Maximum of 145 nanomol rhodium per reaction vial, not determined by ICP-MS. The equimolar 1-alkene mixtures were prepared just prior to the reaction using stocks of the freshly filtered (over alumina) corresponding 1-alkene (containing internal standard).

Hydroformylation of diluted 1-octene (in THF) using SCP-2L V83C-1-P(para)-Rh

Two seperate set of set of reactions (using the general hydroformylation setup) were performed. Both yielded inconclusive results. The two settings used were:

1) SCP-2L V83C-1-P(para) : Rh ≥ 1.2. Maximum of 150 nanomol Rh per reaction vial. Actual rhodium concentration not determined by ICP-MS.

   1-octene is filtered over activated alumina and the following two mixtures are prepared and used in hydroformylation experiments: 1) 7.2 µL 1-octene dissolved in 1.5 mL THF. 2) 72 µL 1-octene dissolved in 1.5 mL THF. Each mixture was tested in dublicate.
2) SCP-2L V83C-1-P(para) : Rh = 11.0. 10.94 nanomol Rh per reaction vial.

   To 1.5 mL 1-octene & internal standard 3.5 mL THF was added. Instead of 0.5 mL organic solvent the following volumes were added to the individual reaction vial (each mixture tested in quadruple): 3 µL, 6 µL, 9 µL, 12 µL.

Recycling of Rh-enzymes

Hydroformylation of 1-octene using SCP-2L V83C-1-P(para)-Rh: Reusability experiment.

SCP-2L V83C-1-P(para) : Rh ≥ 2.0. Maximum of 150 nanomol Rh per reaction vial.

Actual rhodium concentration not determined by ICP-MS. Standard hydroformylation setup was performed using sixteen vials while some of the Rh-enzyme was stored in a Schlenk at 4 °C in the meantime. After the reaction the aqueous phase of the sixteen vials was combined, the solid removed by centrifugation and filtration, washed, and concentrated to a protein concentration of 8.8 mg/mL and used as new catalyst in 3 reaction vials. The Rh-enzyme which had been stored was centrifuged and filtered to remove precipitated protein, washed and concentrated to 8.8 mg/mL and used as catalyst in four reaction vials. The reused catalyst had about half the activity of the stored catalyst and a % linear selectivity of about 68% was observed. The stored Rh-enzyme resulted in a % linear selectivity of about 77%.

3.6 References

[1] P. W. N. M. van Leeuwen, *"Homogenous Catalysis: Understanding the Art*, Kluwer Academic Publishers, **2004**.

[2] R. Weberskirch, Amphiphilic, Water-Soluble Copolymers - Their Synthesis and Application as Support Materials for Catalysis in Aqueous Media, Technische Universität München, **2005**.

[3] A. S. Bommarius, B. R. Riebel, *Biocatalysis: Fundamentals and Applications*, Wiley– VCH, **2004**.

[4] S. Sanchez, A. L. Demain, *Org. Process Res. Dev.* **2010**, *15*, 224–230.

[5] M. R. Ringenberg, T. R. Ward, *Chem. Commun.* **2011**, *47*, 8470–8476.

[6] M. E. Wilson, G. M. Whitesides, *J. Am. Chem. Soc.* **1978**, *100*, 306–307.

[7] K. Yamamura, E. T. Kaiser, *J. Chem. Soc., Chem. Commun.* **1976**, 830–831.

[8] P. J. Deuss, R. den Heeten, W. Laan, P. C. J. Kamer, *Chem.--Eur. J.* **2011**, *17*, 4680–4698.

[9] S. AKABORI, S. SAKURAI, Y. IZUMI, Y. FUJII, *Nature* **1956**, *178*, 323–324.

[10] G. Kiss, N. Çelebi-Ölçüm, R. Moretti, D. Baker, K. N. Houk, *Angew. Chem., Int. Ed.* **2013**, *52*, 5700–5725.

[11] J. Podtetenieff, A. Taglieber, E. Bill, E. J. Reijerse, M. T. Reetz, *Angew. Chem., Int. Ed. Engl.* **2010**, *49*, 5151–5155.

[12] S. Chakraborty, J. Reed, M. Ross, M. J. Nilges, I. D. Petrik, S. Ghosh, S. Hammes-Schiffer, J. T. Sage, Y. Zhang, C. E. Schulz, et al., *Angew. Chem., Int. Ed.* **2014**, *53*, 2417–2421.

[13] W. Zeng, A. Barabanschikov, N. Wang, Y. Lu, J. Zhao, W. Sturhahn, E. E. Alp, J. T. Sage, *Chem. Commun.* **2012**, *48*, 6340–6342.

[14] A. Fernández-Gacio, A. Codina, J. Fastrez, O. Riant, P. Soumillion, *ChemBioChem* **2006**, *7*, 1013–1016.

[15] Q. Jing, K. Okrasa, R. Kazlauskas, Springer Berlin Heidelberg, **2008**, pp. 1–17.

[16] Mauro Marchetti, Fabiola Minello, Stefano Paganelli, Oreste Piccolo, *Appl. Catal., A* **2010**, *373*, 76–80.
[17] M. Marchetti, G. Mangano, S. Paganelli, C. Botteghi, *Tetrahedron Lett.* **2000**, *41*, 3717–3720.
[18] C. Bertucci, C. Botteghi, D. Giunta, M. Marchetti, S. Paganelli, *Adv. Synth. Catal.* **2002**, *344*, 556–562.
[19] Q. Jing, R. J. Kazlauskas, *ChemCatChem* **2010**, *2*, 953–957.
[20] M. E. Wilson, G. M. Whitesides, *J. Am. Chem. Soc.* **1978**, *100*, 306–307.
[21] O. Livnah, E. A. Bayer, M. Wilchek, J. L. Sussman, *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 5076–5080.
[22] J. Steinreiber, T. Ward, in *Bio-Inspired Catalysts* (Ed.: T. Ward), Springer Berlin Heidelberg, **2009**, pp. 93–112.
[23] C.-C. Lin, C.-W. Lin, A. S. C. Chan, *Tetrahedron: Asymmetry* **1999**, *10*, 1887–1893.
[24] C. Letondor, N. Humbert, T. R. Ward, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 4683–4687.
[25] M. Dürrenberger, T. Heinisch, Y. M. Wilson, T. Rossel, E. Nogueira, L. Knörr, A. Mutschler, K. Kersten, M. J. Zimbron, J. Pierron, et al., *Angew. Chem., Int. Ed.* **2011**, *50*, 3026–3029.
[26] F. W. Monnard, E. S. Nogueira, T. Heinisch, T. Schirmer, T. R. Ward, *Chem. Sci.* **2013**, *4*, 3269–3274.
[27] M. V. Cherrier, S. Engilberge, P. Amara, A. Chevalley, M. Salmain, J. C. Fontecilla-Camps, *Eur. J. Inorg. Chem.* **2013**, *2013*, 3596–3600.
[28] A. Pordea, D. Mathis, T. R. Ward, *J. Organomet. Chem.* **2009**, *694*, 930–936.
[29] T. K. Hyster, L. Knörr, T. R. Ward, T. Rovis, *Science* **2012**, *338*, 500–503.
[30] C. Lo, M. R. Ringenberg, D. Gnandt, Y. Wilson, T. R. Ward, *Chem. Commun.* **2011**, *47*, 12065–12067.
[31] J. R. Carey, S. K. Ma, T. D. Pfister, D. K. Garner, H. K. Kim, J. A. Abramite, Z. Wang, Z. Guo, Y. Lu, *J. Am. Chem. Soc.* **2004**, *126*, 10812–10813.
[32] P. J. Deuss, G. Popa, A. M. Z. Slawin, W. Laan, P. C. J. Kamer, *ChemCatChem* **2013**, *5*, 1184–1191.
[33] P. Deuss J., Artificial Metalloenzymes; Modified Proteins as Tuneable Transition Metal Catalysts, University of St Andrews, **2011**.
[34] J.-L. Zhang, D. K. Garner, L. Liang, Q. Chen, Y. Lu, *Chem. Commun.* **2008**, 1665–1667.
[35] A. Onoda, K. Fukumoto, M. Arlt, M. Bocola, U. Schwaneberg, T. Hayashi, *Chem. Commun.* **2012**, *48*, 9756–9758.
[36] D. Häring, M. D. Distefano, *Bioconjugate Chem.* **2001**, *12*, 385–390.
[37] E. T. Farinas, U. Schwaneberg, A. Glieder, F. H. Arnold, *Adv. Synth. Catal.* **2001**, *343*, 601–606.
[38] M. Reetz, in *Bio-Inspired Catalysts* (Ed.: T. Ward), Springer Berlin Heidelberg, **2009**, pp. 63–92.
[39] M. T. Reetz, *J. Org. Chem.* **2009**, *74*, 5767–5778.
[40] N. J. Turner, *Nat. Chem. Biol.* **2009**, *5*, 567–573.
[41] D. H. Appella, *Nat. Chem. Biol.* **2010**, *6*, 87–88.
[42] C. A. Tracewell, F. H. Arnold, *Curr. Opin. Chem. Biol.* **2009**, *13*, 3–9.

[43] P. W. N. M. Van Leeuwen, *Homogeneous Catalysis*, Kluwer Academic Publishers, Dordrecht, **2004**.
[44] M. Yagupsky, C. K. Brown, G. Yagupsky, G. Wilkinson, *J. Chem. Soc. A* **1970**, 937–941.
[45] C. K. Brown, G. Wilkinson, *J. Chem. Soc. A* **1970**, 2753–2764.
[46] G. Yagupsky, C. K. Brown, G. Wilkinson, *J. Chem. Soc. A* **1970**, 1392–1401.
[47] M.-N. Birkholz (nee Gensow), Z. Freixa, P. W. N. M. van Leeuwen, *Chem. Soc. Rev.* **2009**, *38*, 1099–1118.
[48] L. Panella, Phosphoramidite Ligands and Artificial Metalloenzymes in Enantioselective Rhodium-Catalysis: Old Challenges and New Frontiers, Rijksuniversiteit Groningen, **2006**.
[49] M. Beller, Ed. , *Catalytic Carbonylation Reactions*, **2006**.
[50] W. Laan, B. K. Muñoz, R. den Heeten, P. C. J. Kamer, *ChemBioChem* **2010**, *11*, 1236–1239.
[51] M.-S. Kim, Q. Lan, *BMC Physiology* **2010**, *10*, 9.
[52] L. Viitanen, M. Nylund, D. M. Eklund, C. Alm, A.-K. Eriksson, J. Tuuf, T. A. Salminen, P. Mattjus, J. Edqvist, *FEBS J.* **2006**, *273*, 5641–5655.
[53] D. H. Dyer, I. Vyazunova, J. M. Lorch, K. T. Forest, Q. Lan, *Mol. Cell. Biochem.* **2009**, *326*, DOI 10.1007/s11010-008-0007-z.
[54] F. V. Filipp, M. Sattler, *Biochemistry* **2007**, *46*, 7980–7991.
[55] W. A. Stanley, K. Versluis, C. Schultz, A. J. R. Heck, M. Wilmanns, *Arch. Biochem. Biophys.* **2007**, *461*, 50–58.
[56] D. H. Dyer, S. Lovell, J. B. Thoden, H. M. Holden, I. Rayment, Q. Lan, *J. Biol. Chem.* **2003**, *278*, 39085–39091.
[57] D. H. Dyer, V. Wessely, K. T. Forest, Q. Lan, *J. Lipid. Res.* **2008**, *49*, 644–653.
[58] R. den Heeten, AMINO ACID MODIFIED PHOSPHINE LIGANDS FOR THE DEVELOPMENT OF ARTIFICIAL TRANSITION METALLOENZYMES, University of Amsterdam, **2009**.
[59] B. Cornils, W. A. Herrmann, *Aqueous-Phase Organometallic Catalysis*, Wiley-VCH, Weinheim, **2004**.
[60] Y. Zhang, Z.-S. Mao, J. Chen, *Catal. Today* **2002**, *74*, 23–35.

Chapter 4: Artificial metalloenzymes in the Diels-Alder reaction

4.1 Abstract

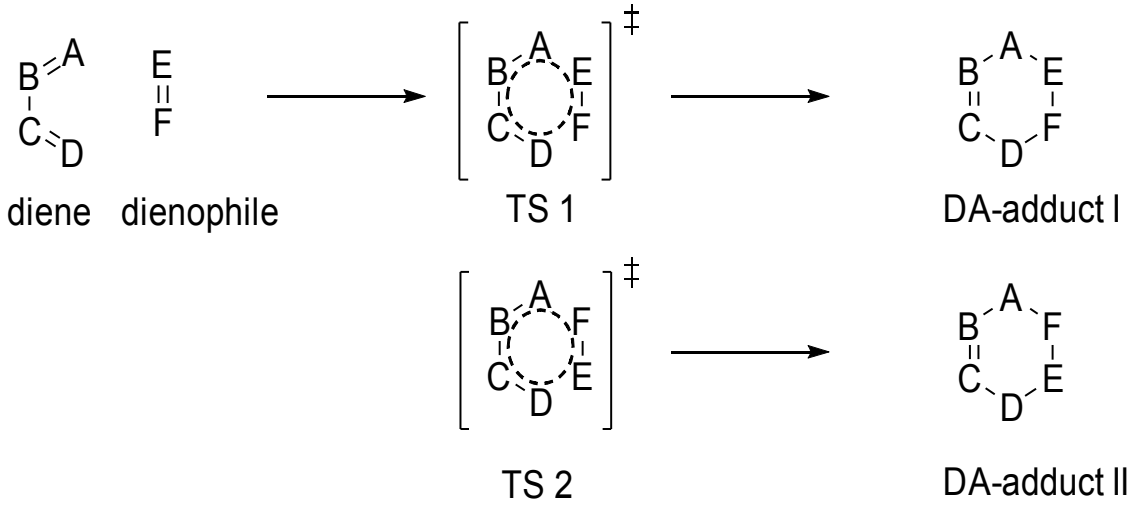
The catalytic performance of artificial metalloenzymes was tested in the asymmetric Diels-Alder (DA)-reaction between cyclopentadiene (Cp) and 2-azachalcone. Artificial cobalt-, copper- and nickel-enzymes based on the SCP-2L A100C mutant modified with either a phenanthroline-based (*Phen*) or a dipicolylamine-based (*Picol*) moiety were used. The copper and cobalt enzymes hardly influence the reaction in comparison to the corresponding free metal catalysed reactions. The nickel-enzymes have an influence on the endo/exo ratio in either way depending on the protein scaffold used. Additionally, while most tested catalysts showed no - or hardly any - enantioselectivity 13% ee of the endo and 29% ee of the exo product was observed when the reaction was catalysed with the Ni-enzyme of the *Phen*-modified protein.

4.2 Introduction

Diels-Alder reaction

The Diels-Alder (DA) reaction is named after the two German chemists Otto Diels and Kurt Alder who were not the first to discover the reaction,[1,2] but realised its importance and investigated it in great detail from the late 1920s onwards. They were awarded the Nobel Prize in 1950 "for their discovery and development of the diene synthesis"[3] as the DA reaction was referred to in the beginning.

The DA reaction is a [4+2] cycloaddition between a conjugated diene and a dienophile (usually a monoene, but an alkyne is possible as well) forming a six membered ring - the DA-adduct - in a concerted reaction.[2] The reaction is depicted in scheme 1.

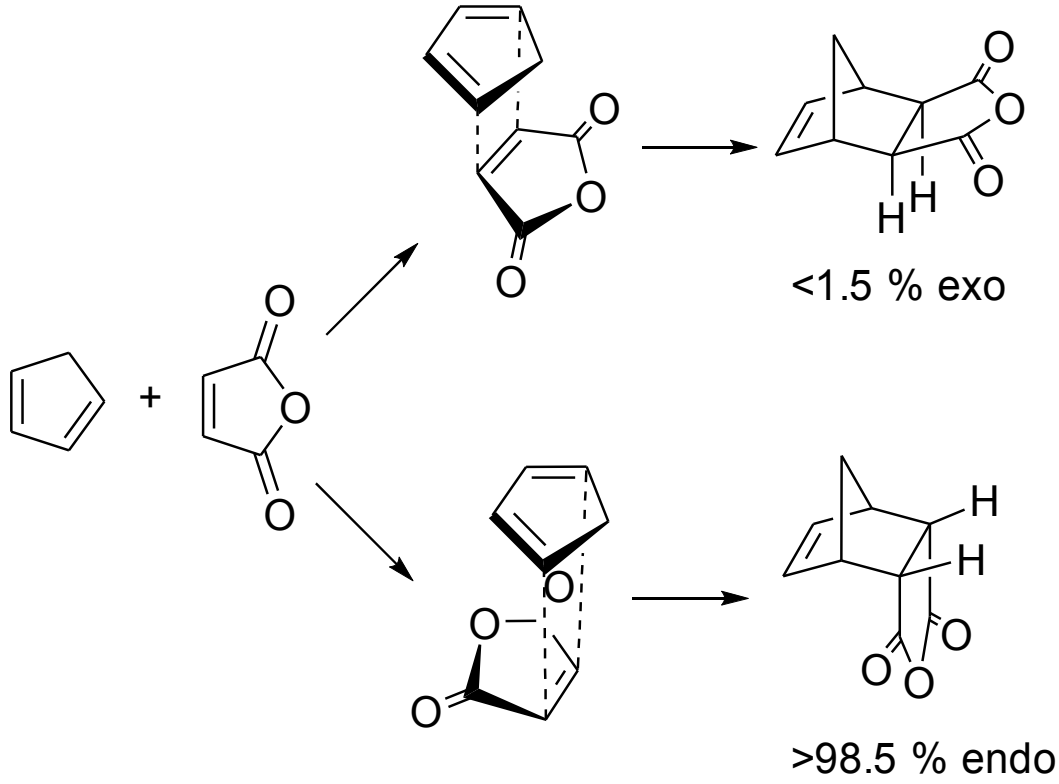


Scheme 1: Schematic representation of the DA-reaction with concerted transition state(s) (TS).

The versatility of the DA-reaction is indicated in the scheme above. As hetero atoms and substituents are allowed at any of the positions (A-F) two different regioisomers (DA-adduct I and DA-adduct II) can be obtained. If one (or more) of the involved atoms is a non carbon atom the reaction is considered a hetero Diels-Alder reaction and a heterocycle is obtained. If substituents are present up to four new stereogenic centres can be introduced in a single Diels-Alder reaction. The DA-reaction is stereospecific,

which means that conformations of the reacting double bonds are fully retained in the adduct. This observation was recognised very early and described as the "cis-principle". [4]

The DA-adduct can form either the "endo" or the "exo"-adduct. As an example the reaction between Cp and maleic anhydride is depicted in scheme 2.



Scheme 2: DA-reaction between Cp and maleic anhydride. Product distribution taken from[5]

The thermodynamically less stable "endo" product is the major product in most DA-reactions. ([6,7] are examples for exo-selective DA-reaction). The experimental observation was rationalised with the frontier molecular orbital (FMO)-theory. The FMO diagram for a normal electron demand DA reaction is depicted in figure 1. It shows the interaction of the HOMO of a diene with the LUMO of a dienophile.

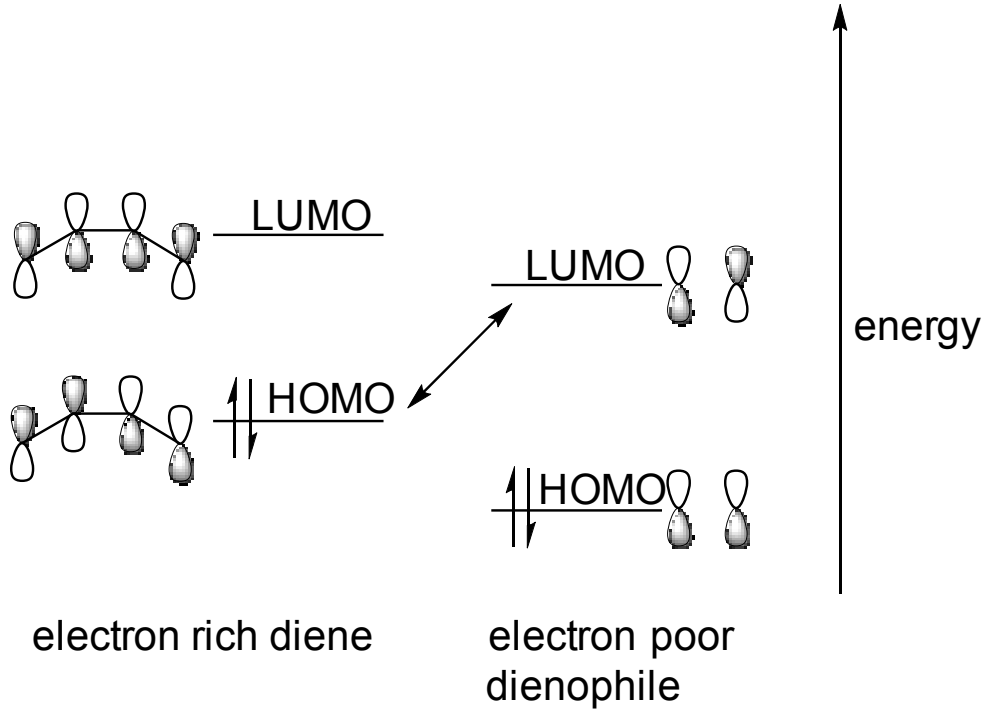


Fig. 1: Orbital correlation diagram for a normal DA reaction with with an electron rich diene and an electron poor dienophile.[2]

The preference of the endo-DS-adduct can be rationalised by taking secondary orbital interactions into account (see fig. 2)

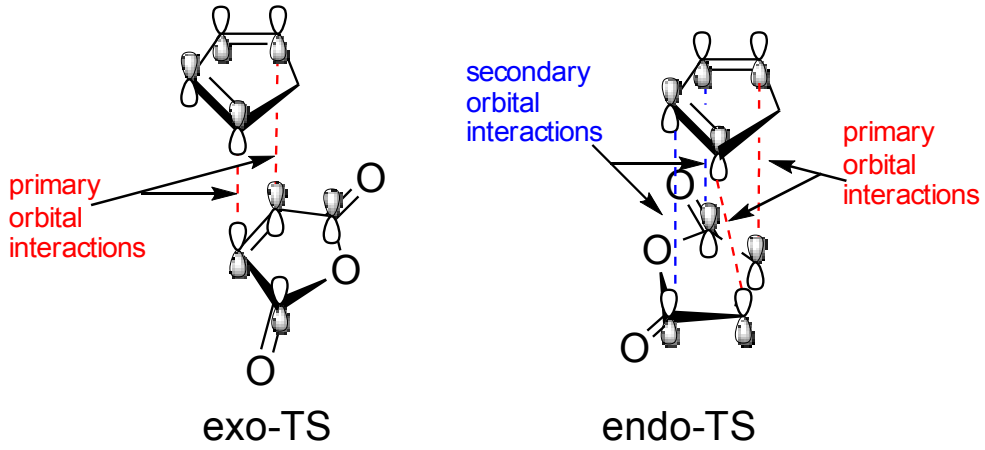


Fig. 2: Transition states (TS) of the reaction between Cp and maleic anhydride. Dashed lines represent interactions.

While the exo-TS only exhibits primary orbital interactions the endo-TS exhibits additional secondary orbital interactions.

Biochemical approaches to catalyse the Diels-Alder reaction

As the DA reaction is an important tool in synthetic chemistry, especially in the total synthesis of natural compounds.[1,9–11] it was proposed that enzymes catalysing the DA reaction (so called "Diels-Alderases") play an important role in their biosynthesis as well[12–15]. But so far only five potential Diels-Alderases have been discovered.[16] The biochemical approaches towards DA-catalysis will be briefly introduced.

The first RNA based Diels-Alderase was discovered in 1997 by in vitro selection of a large library (~10^{14}) of unique RNA sequences.[17] While one of the reactants is usually covalently tethered to RNA (usually via a polyethylene glycol linker[17–21]) RNA sequences were successfully used to catalyse the reaction between untethered reactants.[20] The RNA sequences tested might contain unnatural RNA to furnish additional hydrophobic and dipolar interactions as well as to provide additional metal coordination sites.[17] An increase in the reaction rate by up to four orders of magnitude[18] and up to 95% ee[20] was observed using this sort of catalysts.

The first antibody catalysed DA reaction was published in 1989.[22] In order to obtain an antibody which is (theoretically) capable of catalysing a specific DA reaction a TS-analogue of this reaction is needed. While most approaches use rigid TS-analogues[22–26] it was found that flexible ones can be successfully used, as well.[27] The synthesised analogue is transferred into a hapten, coupled to an appropriate immunogenic protein and used to immunise a host (mice in general). Up to 95% ee was observed using antibodies as catalysts in a DA reaction[27] showing the potential of this approach.

DNA-based catalysts ("DNAzymes") have been successfully tested in various DA reactions. The catalytically active Cu(II) is brought into close proximity to a DNA sequence by covalent modification of DNA with an appropriate ligand or by non covalent means.[28] Both the DNA-sequence[29] as well as the coordinating ligand[30] play an important role in the outcome of the reaction. The DNA is not only acting as an effective chiral scaffold yielding up to 99% ee[31] but rate accelerations of up to two orders of magnitude were observed as well.[31] The re-usability of this type of catalysts

was shown for solid supported DNA and only a slight decrease in ee was observed during ten consecutive runs indicating the stability of such DNAzymes.[32]

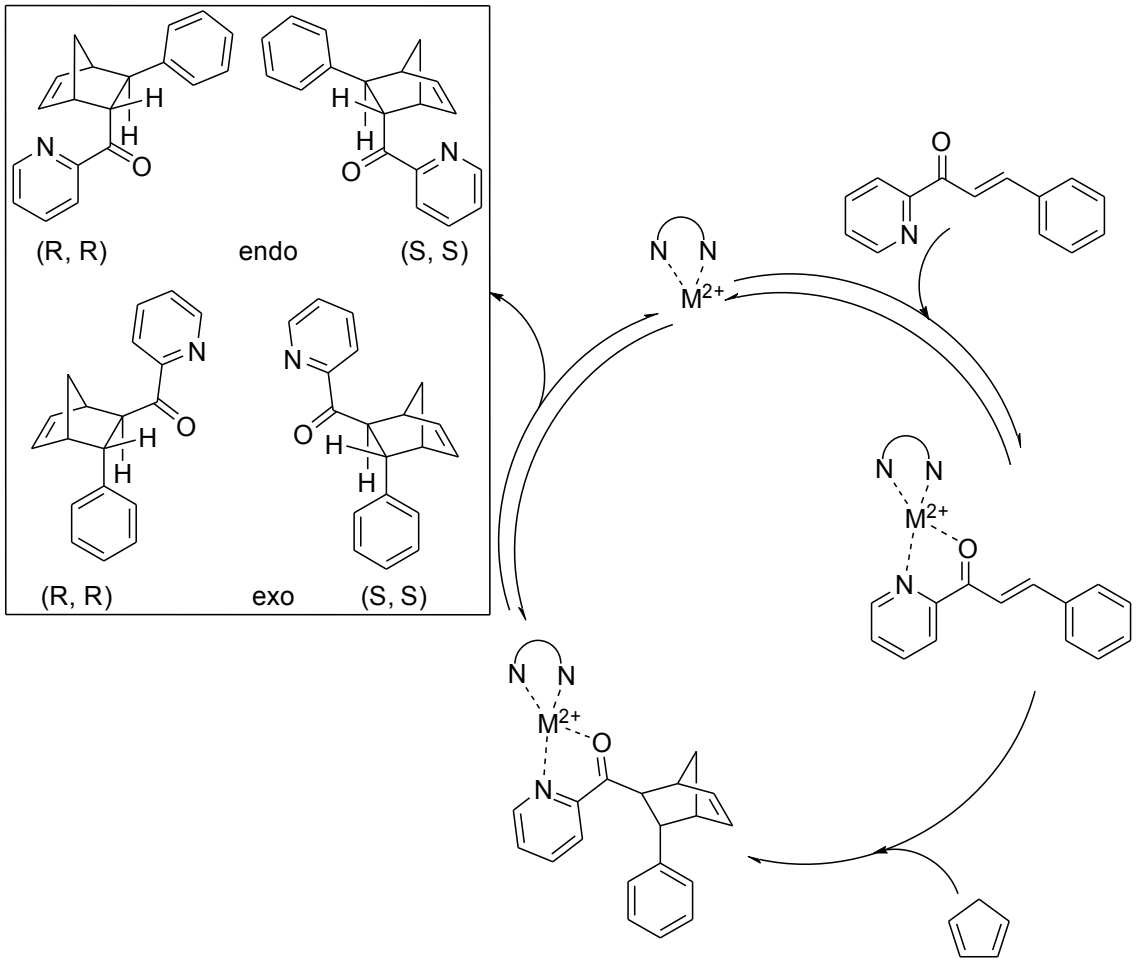
Copper(II) coordinating to polypeptides,[33] artificial metalloenzymes[11,34–37] or artificial Diels-Alderase[38,39] have been successfully tested as catalysts for the DA reaction. The metal is either coordinated to an appropriate ligand which is covalently or non-covalently bound to the scaffold or directly coordinated to the amino acids of the scaffold. High ee values obtained for different substrates show the versatility of such an approach.

Diels-Alder reaction between Cp and 2-azachalcone

The reaction between Cp and 2-azachalcone, which was investigated in this thesis has been widely tested for different bio-inspired approaches based on DNAzymes[28–32,40–50], artificial metalloenzymes[34–37,51,52] and polypeptides[33] using copper(II) as active metal centre. From this background this reaction can be regarded as a benchmark for bio-inspired catalysis. Up to 99.4% ee[28,31] and endo/exo ratios ranging from 1.7 to 24[37] have been reported for this reaction. Although copper(II) is more active than other transition metals it is known that nickel(II) and cobalt(II) can be used as catalysts as well.[2] In this chapter the performance of copper(II)-bound artificial metalloenzymes will be compared with the respective nickel(II)-, and cobalt(II)- derivatives.

The water insolubility of the substrates makes them perfect candidates to analyse the effect of the protein scaffold in catalysis. Additionally the results obtained can be compared with the many results published earlier to put the results obtained into persepctive.

The mechanism proposed for the metal catalyzed DA-reaction is shown in scheme 3.



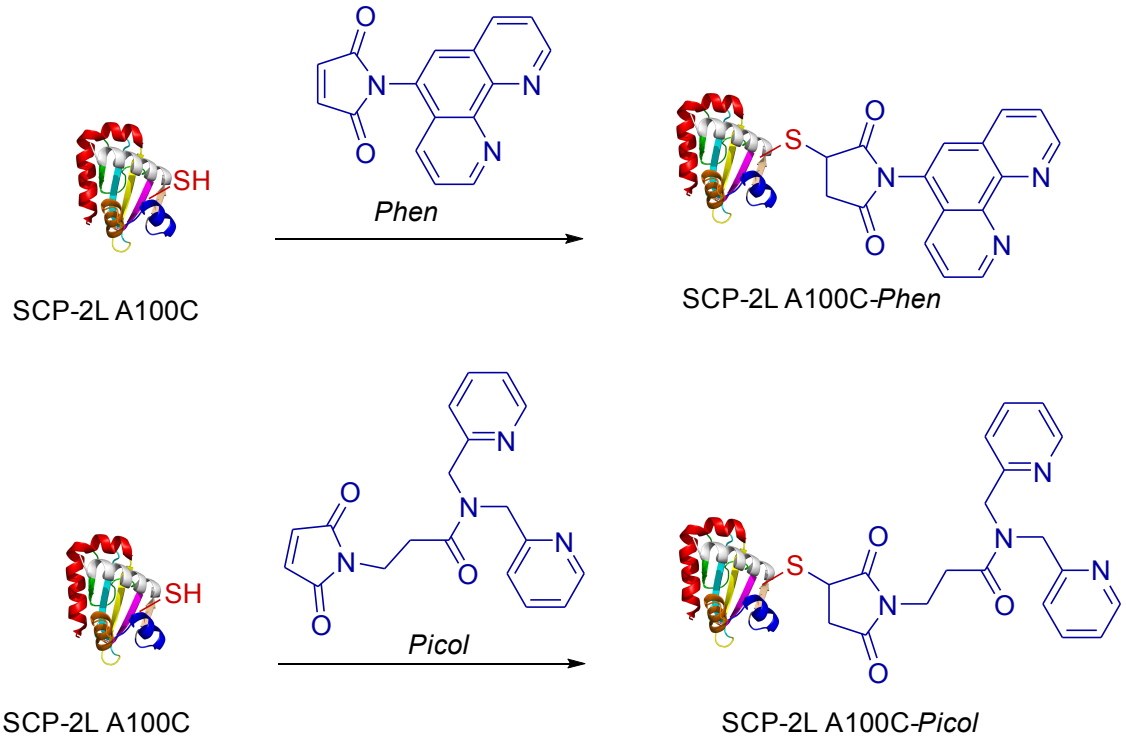
Scheme 3: Proposed mechanism the metal catalyzed DA-reaction between Cp and *trans*-azachalcone.[2]

First the (ligated) metal is coordinating to both the nitrogen and the oxygen of the azachalcone, decreasing the energy of the LUMO of the adjacent double bond and therefore activating it for the cycloaddition. The cycloaddition takes place in the next step when a molecule of Cp enters the cycle. In the last step the DA-adduct is released and the (ligated) metal can start the cycle all over again. Four different products (two endo and two exo) are formed.

4.3 Results and Discussion

Protein modification

For the catalysis of the DA reaction two modifications of the protein SCP-2L A100C were used. In either modification nitrogen containing ligands were attached to the thiol groups. The *N*-Ligands used are based on the bidentate phenanthroline-like structure (*Phen*) or the tridentate dipicolylamine-based structure (*Picol*). For both N-ligands tested, the synthesis of these ligands as well as the modification of the two SCP-2L mutants was found to be straightforward.[52] The scheme for the modification of SCP-2L A100C was introduced in chapter 2 and is depicted in scheme 4.



Scheme. 4: Protein modification scheme and N-donor-ligands used in this chapter.

Metal coordination

The capability of the (un)modified protein to coordinate to various metals was already discussed in chapter 2. In table 1 the metal contents of the unmodified SCP-2LA100C

are compared with the metal contents of the *Picol* and *Phen*-modified derivatives. It should be mentioned that the metal contents given in table 1 were determined after extensive washing. The data summarized in table 1 will be used for the discussion of the catalytic results.

Table 1: Metal loading of SCP-2L A100C, SCP-2L A100C-*Picol* and SCP-2L A100C-*Phen* after the addition of of Cu, Ni and Co and after extensive washing[a]

		Protein scaffold		
		SCP-2L A100C	**SCP-2L A100C-*Picol***	**SCP-2L A100C-*Phen***
entry	**metal salt**	% metal loading[b]	% metal loading[b]	% metal loading[b]
1	$Cu(OAc)_2$	66.3	5.8	22.9
2	$Ni(OAc)_2$	6.4	7.8	11.1
3	$Co(OAc)_2$	2.5	2.9	26.8

a: To 0.1 µmol protein (in 20 mM MES, 50 mM NaCl, pH=6) 2-3 eq. of metal salt were added, reaction time 1 day. Buffer exchanged to 20 mM NH4OAc, pH=6 and a fifth was removed for LC-MS analysis. 4µL from an EDTA stock solution (0.5 M, pH = 8) was added to the rest. The mixture was buffer exchanged (using 20 mM NH_4OAc, pH = 6) and the flow through collected. This was repeated three times. The fractions were combined, concentrated to dryness, topped up with water to a known total volume (1.00 mL) and the metal content determined by ICP-MS. See experimental section in Chapter 2 for all details. **b**: n(metal) / n(Protein), corrected for the 1/5th removed for ESI-MS analysis; 50% corresponds to 0.5 mol metal per 1 mol (un)modified protein, 100% corresponds to 1 mol metal per 1 mol (un)modified protein and so on. All details are described in the experimental section in chapter 2.

Based on the metal loading listed in table 1 it can be stated that unmodified SCP-2L A100C has a higher complexation affinity to Cu(II) and a lower to Ni(II) and Co(II) than the modified derivatives SCP-2L A100C-*Picol* and SCP-2L A100C-*Phen*. In addition, it can be derived from table 1 that SCP-2L A100C-*Phen* has a higher complexation affinity to the tested metals salts than SCP-2L A100C-*Picol*. After extensive washing the residual metal contents of SCP-2L A100C-*Phen* are almost four times higher regarding Cu(II), roughly 50% higher for Ni(II) and more than nine times higher for Co(II) than the corresponding metal contents of *SCP-2L A100C-Picol*. Therefore it can be expected that the metal derivatives of the SCP-2L A100C-*Phen* scaffold will show much more pronounced catalytic effects than the respective SCP-2L A100C-*Picol* scaffolds.

Catalytic reactions

The catalytic results for the tested DA reaction are listed in table 2.

Table 2: Diels-Alder reaction between 2-azachalcone and Cp in the presence of different catalysts[a]

Entry	catalyst	Conversion [%][b]	endo : exo[b]	Endo ee[b,c]	exo ee[b,c]	N[d]
1	Ni(II)	84 (±1)	7.4 (±0.2)	-	-	4
2	Cu(II)	97 (±1)	13.4 (±0.3)	-	-	3
3	Co(II)	42 (±2)	8.1 (±0.3)	-	-	3
4	Blank	n.d.[e]	6.1 (±0.1)	-	-	4
5	SCP-2L A100C + Ni(II)	91 (±1)	8.7 (±0.2)	4 (±1)	8 (±4)	3
6	SCP-2L A100C + Cu(II)	93 (±0)	13.3 (±0.0)	-	4 (±0)	2
7	SCP-2L A100C + Co(II)	55 (±5)	7.6 (±0.1)	-	-	2
8	SCP-2L A100C-*Picol* + Ni(II)	78 (±5)	8.5 (±0.6)	3 (±0)	3 (±0)	3
9	SCP-2L A100C-*Picol* + Cu(II)	98 (±0)	14.2 (±0.4)	3 (±0)	16 (±1)	3
10	SCP-2L A100C-*Picol* + Co(II)	35 (±12)	9.7 (±1.2)	-	-	3
11	SCP-2L A100C-*Phen* + Ni(II)	92 (±6)	5.4 (±0.1)	13 (±1)	29 (±2)	3
12	SCP-2L A100C-*Phen* + Cu(II)	79 (±2)	13.3 (±0.5)	-	6 (±1)	3
13	SCP-2L A100C-*Phen* + Co(II)	33 (±1)	7.0 (±0.7)	-	-	3

a: 50 nmol of (unmodified) protein, 45 nmol of metal(II)nitrate, 0.5 µmol azachalcone and 15 µmol Cp in 0.8 mL buffer (20 mM MES, 50 mM NaCl, pH = 6). Shaken for 72h at 4 °C **b**: determined by HPLC, see experimental section for details, standard deviations in brackets **c**: In case of ee: The first enantiomer to elute was in excess, "-" means no ee was detected **d**: number of experiments **e**: not determined

Metal catalysed reactions (entries 1-4)

In a first set of control reactions (entries 1-3) the Diels-Alder reaction was performed in the presence of one of the metal(II)-nitrates (Ni, Co and Cu). Also the reference reaction without metals was performed in this context (entries 4). The conversion follows the trend Cu>Ni>>Co. The Cu(II) catalysed reaction reaches almost full conversion (91%). Ni(II)-catalysis yields 80% and Co(II) catalysis about 40% conversion. The conversion of the uncatalysed reaction was not determined but is known to be low (≈10%) under these conditions.[35]

The endo/exo ratio follows the trend Cu>>Co>Ni>blank. The Cu(II) catalysed reaction gives a ratio of 13.4, the Co(II) catalysed a ratio of 8.1, the Ni(II) catalysed one a ratio

of 7.4 and the uncatalysed one a ratio of 6.1. As expected no enantioselectivity was observed in these reactions.

Metal catalysis in the presence of unmodified protein (entries 5-7)

In a second set of control reactions the DA-reaction was performed in the presence of both the unmodified protein SCP-2L A100C as well as one of the metal ions. Compared to the previous set of control reactions the conversions are similar except for some minor deviations. A slightly lower conversion for the Cu(II) and a slightly higher conversions for Ni(II) and Co(II) in the presence of protein compared to the metal(II) catalysed reactions without protein was detected. The endo/exo ratios are basically the same except for Ni(II)-catalysis in the presence of protein which yields a higher ratio than the free metal catalysed reaction (8.7 compared to 7.4). Interestingly a small ee was detected for both adducts when Ni(II) was used as catalyst in the presence of the protein (entry 5) and a small ee was detected for the exo-adduct when Cu(II) was used in the presence of protein (entry 6).

Metal catalysis in the presence of SCP-2L A100C-*Picol* (entries 8-10)

In a next set of experiments the reaction was performed in the presence of modified protein SCP-2L A100C-*Picol* and either one of the metal salts (entries 8-10). The differences for the actual metalloenzyme in comparison to the unmodified enzyme in presence of the metal (entries 5-7) in respect of both the conversion and the endo/exo ratio is in general small. The conversions for the Ni- and the Co-enzyme are slightly lower while the conversion of the Cu-enzyme is slightly higher. The endo/exo ratios in the presence of Cu and Co are slightly higher while it is unchanged in the presence of Ni when comparing it to either the free metal reactions (entries 1-3) or the reactions in which unmodified protein was used (entries 5-7). This is in full agreement with the expectations from table 1 as the affinity between SCP-2L A100C-*Picol* to Ni(II) and Co(II) is lower than the affinity of unmodified protein to the same metals (see table 1). Table 1 does show a much stronger interaction between Cu(II) and unmodified SCP-2L A100C than between Cu(II) and SCP-2L A100C-*Picol* therefore a more significant difference between these two catalytic systems could be expected. As the results are also

almost identical to the ones obtained when using "free" Cu(II) it might be that the quantity of "free" Cu(II) is still rather high when protein is present and this "free" Cu(II) determines the outcome of the reaction while the comparably small quantities of protein-bound Cu(II) hardly influence it.

When using SCP-2L A100C-*Picol*-Ni (entry 8) a small ee for both adducts was detected. The same ee was observed when the unmodified protein was used in the presence of Ni(II) (entry 5). In case of SCP-2L A100C-*Picol*-Cu (entry 9) a small ee could be detected for the endo-adduct and a quite significant ee was detected for the exo-adduct. When unmodified SCP-2L A100C and Cu(II) was used as catalyst no ee was detected for the endo adduct and only a small ee was found in the exo-adduct (entry 6).

Metal catalysis in the presence of SCP-2L A100C-*Phen* (entries 11-13)

In a next set of experiments the reaction was performed in the presence of modified protein SCP-2L A100C-*Phen* and either one of the metal salts (entries 11-13). This set of experiments was expected to be the most promising (see table 1) as the metal loadings are much higher for SCP-2L A100C-*Phen* than for SCP-2L A100C-*Picol* and SCP-2L A100C (SCP-2L A100C + Cu being an exception).

The results obtained for the Cu-enzyme (entry 12) are similar to the ones obtained when using Cu(II) as catalyst, although the conversion was lower than in previous experiments. Just like in the previous experiment where unmodified protein was present no ee of the adduct was detected. The results obtained for the Co-enzyme (entry 13 are similar to the other reactions with Co(II) as catalyst. In case of the Ni-enzyme (entry 11) a slightly higher conversion was detected. More interesting, the endo/exo ratio was much lower than in all previous reactions with Ni(II) as catalyst (entries 1, 5, 8). Additionally the highest ee for both products was detected (13% ee for the endo) with this Ni-enzyme as catalyst. The Ni-enzyme was the metalloenzyme in which the highest concentrations of unligated metal were expected (according to table 1). Nevertheless the reaction outcome indicates a high influence of the artificial metalloenzyme which makes it even more significant.

Rationalisation of the results

Regarding the combinations of modified protein and metal, three distinct species, which probably coexist are possible. Many intermediates of these extremes, as well as non-1:1 coordination might exist but all of these can be described as combinations of the three systems depicted in fig. 3.

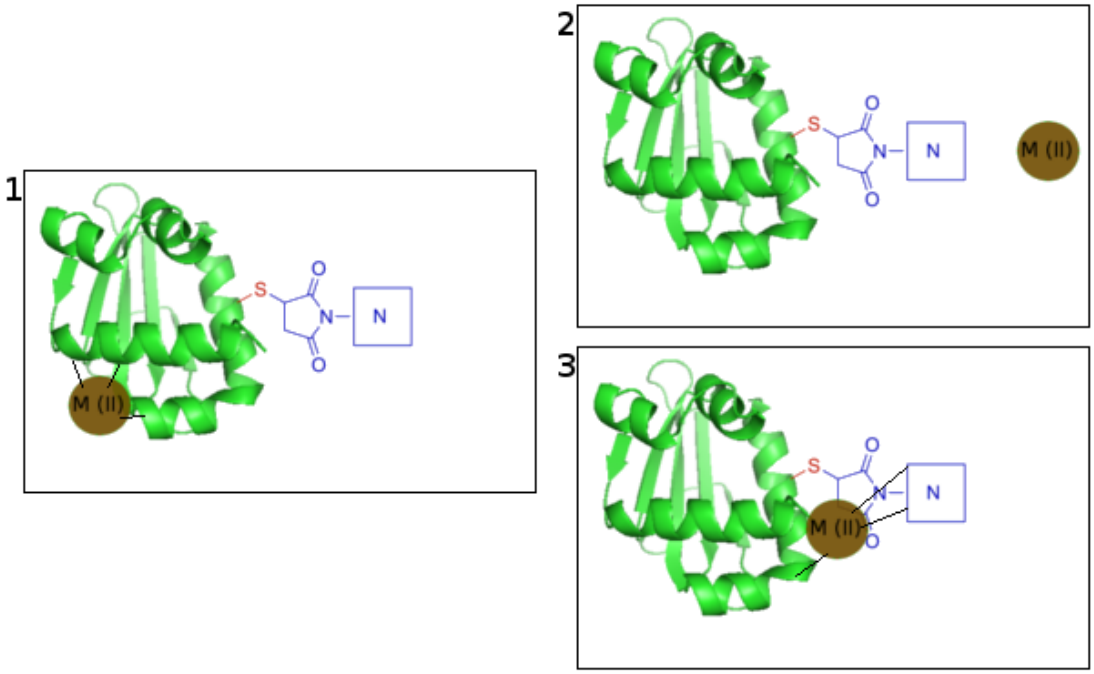


Figure 3: Potential combinations of a modified protein and a metal (II) [depicted as M(II)]. The introduced *N*-ligand is represented by a "N" in a box. **1**: unspecific coordination of metal to undefined amino acids (interactions indicated by black lines). **2**: Non protein-coordinated ("free") metal **3:** coordination of metal to the introduced ligand (interactions indicated by black lines).

In the first case depicted in figure 3 the metal is coordinated to amino acids rather than to the artificially introduced ligand. In the second case no interaction between metal and protein is present. In the third case the metal is coordinated to the introduced ligand. Each of these species might show catalytic activity and might yield different products. As the three species coexist it is impossible to clearly assign catalytic results to a distinct species. Therefore only differences in the results of comparative experiments can be discussed.

The catalytic results obtained make it seem likely that the outcome of the reaction is mostly determined by free metal (figure 3, species 2). It is also possible that the difference in the product distribution between protein coordinated metal (figure 3, species 3) and free metal is small. SCP-2L A100C-*Phen* has a much higher affinity to Co(II) and Ni(II) than SCP-2L A100C-*Picol* (see table 1) and results in lower conversions. This might be explained by a reduced activity of the protein bound metal in comparison to the free metal. SCP-2L A100C-*Phen*-Ni is the most interesting case. Its metal loading is comparably low (see table 1) which means high concentrations of unligated nickel will be present during reaction conditions. Nevertheless the effect on both the selectivity as well as the enantioselectivity is much more pronounced than for any other artificial metalloenzyme tested. This means that the protein bound nickel has to have a high preference for the exo product and must induce a high enantioselectivity as most of the catalysis must be determined by unligated nickel.

4.4 Conclusions and future work

The DA-reaction catalysed by Cu(II) in the presence of the protein scaffolds SCP-2L A100C (*Phen* and *Picol* modified) and SCP-2L V83C (*Phen* and *Picol* modified) has been described in a previous paper.[35] Under the same experimental conditions, catalysis with SCP-2L A100C (*Phen* and *Picol* modified) and Cu(II) yielded similar results as in this thesis. The ee of the endo-product obtained by copper-enzymes based on the protein mutant SCP-2L V83C was significantly higher.[35]

In this work the Ni- and Co-derivatives of the modified proteins have been tested for the first time. The ee induced by the combination of Ni(II) with SCP-2L A100C-*Phen* is much higher than for the Cu(II) and Co(II)-derivatives. From this background it might be worthwhile to test the Ni-derivative of the SCP-2L V83C-*Phen* scaffold. In order to reduce the undesired catalytic reaction of unligated metal its concentration has to be kept low. This can be achieved by performing catalysis at a high protein to metal ratio. Low concentrations of unligated metal can also be obtained by the use of preformed and washed metalloenzymes – which is essentially another possibility to obtain a high protein to metal ratio.

4.5 Experimental

Equipment, Chemicals and suppliers

The *N*-ligands used for protein conjugation (*Phen*[53,54] and *Picol*[55]) were synthesised by E. de Waard and M. Drysdale according to known literature procedures. Dicyclopentadiene was cracked just before the catalysis. 2-azachalcone was synthesised according to a literature known procedure.[56]

Protein expression, purification and modification

Protein expression, purification and modification was performed based on the procedure described in chapter 2.

Catalysis and analysis

Catalysis and analysis was performed based on a published procedure[35] :

50 nmol of ligand (whereas ligand refers to protein or modified protein) in buffer (20 mM MES, 50 mM NaCl, pH = 6) and 45 nmol of metal(II) nitrate (Cu or Co or Ni) in the same buffer was mixed in a 1.5 mL Eppendorf tube and topped up with the buffer to obtain a total volume of 794 µL. It was mixed for 1 h prior to the catalytic reactions. Protein concentrations were determined just before addition using Bradfords reagent. 2-azachalcone (5 µL from a 100 mM solution in MeCN) and freshly cracked cyclopentadiene (1 µL) were added to the reaction mixture. The obtained solution was shaken for 72 h at 4 °C. The reaction mixtures were extracted 5 times with diethyl ether (1 mL) and the combined organic layers were dried in vacuo. The obtained solid was dissolved in a suitable amount of 30% IPA in hexane (HPLC grade). HPLC analysis was performed by Dr. José Antonio Fuentes García using an OD-H column with a flow of 0.5 mL/min. of 2% IPA in hexanes (retention times: Cp: 6.7 min., exo products 12.7 min. and 14.2 min., endo products 15.8 min. and 20.4 min., azachalcone 22.9 min.). The chromatogram was obtained by UV analysis at 212 nm and 254 nm. The spectra which gave the highest absorption values, the sharpest peaks and the most consistent baseline is used for analysis. The conversion was calculated using the %-area values and are therefore approximate

4.6 References

[1] C. C. Nawrat, C. J. Moody, *Angew. Chem., Int. Ed.* **2014**, *53*, 2056–2077.
[2] S. Otto, CATALYSIS OF DIELS-ALDER REACTIONS IN WATER, Rijksuniversiteit Groningen, **1998**.
[3] http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1950/.
[4] K. Alder, G. Stein, *Angew. Chem.* **1937**, *50*, 510–519.
[5] J. Sauer, *Angew. Chem., Int. Ed. Engl.* **1967**, *6*, 16–33.
[6] T. S. Powers, W. Jiang, J. Su, W. D. Wulff, B. E. Waltermire, A. L. Rheingold, *J. Am. Chem. Soc.* **1997**, *119*, 6438–6439.
[7] B. Boren, J. S. Hirschi, J. H. Reibenspies, M. D. Tallant, D. A. Singleton, G. A. Sulikowski, *J. Org. Chem.* **2003**, *68*, 8991–8995.
[8] M., M. K. Boysen, in *Carbohydrates: Tools for Stereoselective Synthesis* (Ed.: M., M. K. Boysen), Wiley-VCH, **2013**, pp. 65–93.
[9] G. Muncipinto, in *Diversity-Oriented Synthesis* (Ed.: A. Trabocchi), Wiley-VCH, **2013**, pp. 59–70.
[10] S. Reymond, J. Cossy, *Chem. Rev. (Washington, DC, U. S.)* **2008**, *108*, 5359–5406.
[11] E. M. Stocking, R. M. Williams, *Angew. Chem., Int. Ed.* **2003**, *42*, 3078–3115.
[12] W. L. Kelly, *Org. Biomol. Chem.* **2008**, *6*, 4483–4493.
[13] H. J. Kim, M. W. Ruszczycky, H. Liu, *Curr. Opin. Chem. Biol.* **2012**, *16*, 124–131.
[14] H. Oikawa, T. Tokiwano, *Nat. Prod. Rep.* **2004**, *21*, 321–352.
[15] M. W. S. C. L. L. Hak Joong KimRuszczycky, *Nature (London, U. K.)* **2011**, *473*, 109–112.
[16] T. M. Tarasow, S. L. Tarasow, B. E. Eaton, *Nature (London, U. K.)* **1997**, *389*, 54–57.
[17] B. Seelig, A. Jäschke, *Chem. Biol.* **n.d.**, *6*, 167–176.
[18] T. M. Tarasow, S. L. Tarasow, C. Tu, E. Kellogg, B. E. Eaton, *J. Am. Chem. Soc.* **1999**, *121*, 3614–3617.
[19] B. Seelig, S. Keiper, F. Stuhlmann, A. Jaschke, *Angew. Chem. Int. Ed. Engl.* **2000**, *39*.
[20] T. M. Tarasow, S. L. Tarasow, B. E. Eaton, *J. Am. Chem. Soc.* **2000**, *122*, 1015–1021.
[21] D. Hilvert, K. W. Hill, K. D. Nared, M. T. M. Auditor, *J. Am. Chem. Soc.* **1989**, *111*, 9261–9262.
[22] A. C. Braisted, P. G. Schultz, *J. Am. Chem. Soc.* **1990**, *112*, 7430–7431.
[23] N. Bahr, R. Güller, J.-L. Reymond, R. A. Lerner, *J. Am. Chem. Soc.* **1996**, *118*, 3550–3555.
[24] A. A. P. Meekel, M. Resmini, U. K. Pandit, *J. Chem. Soc., Chem. Commun.* **1995**, 571–572.
[25] A. A. P. Meekel, M. Resmini, U. K. Pandit, *Bioorg. Med. Chem.* **1996**, *4*, 1051–1057.
[26] J. T. Yli-Kauhaluoma, J. A. Ashley, C.-H. Lo, L. Tucker, M. M. Wolfe, K. D. Janda, *J. Am. Chem. Soc.* **1995**, *117*, 7041–7047.
[27] A. J. Boersma, R. P. Megens, B. L. Feringa, G. Roelfes, *Chem. Soc. Rev.* **2010**, *39*, 2083–2092.

[28] G. Roelfes, B. L. Feringa, A. Meetsma, J. E. Klijn, A. J. Boersma, F. Rosati, *Chem.--Eur. J.* **2009**, *15*, 9596–9605.
[29] F. Rosati, G. Roelfes, *ChemCatChem* **2011**, *3*, 973–977.
[30] A. J. Boersma, J. E. Klijn, B. L. Feringa, G. Roelfes, *J. Am. Chem. Soc.* **2008**, *130*, 11783–11790.
[31] S. Park, K. Ikehata, H. Sugiyama, *Biomater. Sci.* **2013**, *1*, 1034–1036.
[32] D. Coquiere, J. Bos, J. Beld, G. Roelfes, *Angew. Chem., Int. Ed. Engl.* **2009**, *48*, 5159–5162.
[33] J. Podtetenieff, A. Taglieber, E. Bill, E. J. Reijerse, M. T. Reetz, *Angew. Chem., Int. Ed. Engl.* **2010**, *49*, 5151–5155.
[34] P. J. Deuss, G. Popa, A. M. Z. Slawin, W. Laan, P. C. J. Kamer, *ChemCatChem* **2013**, *5*, 1184–1191.
[35] M. T. Reetz, M. Rentzsch, A. Pletsch, M. Maywald, P. Maiwald, J. J.-P. Peyralans, A. Maichele, Y. Fu, N. Jiao, F. Hollmann, et al., *Tetrahedron* **2007**, *63*, 6404–6414.
[36] J. Bos, F. Fusetti, A. J. M. Driessen, G. Roelfes, *Angew. Chem., Int. Ed.* **2012**, *51*, 7472–7475.
[37] N. Preiswerk, T. Beck, J. D. Schulz, P. Milovnik, C. Mayer, J. B. Siegel, D. Baker, D. Hilvert, *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 8013–8018.
[38] C. B. Eiben, J. B. Siegel, J. B. Bale, S. Cooper, F. Khatib, B. W. Shen, F. Players, B. L. Stoddard, Z. Popovic, D. Baker, *Nat. Biotechnol.* **2012**, *30*, 190–192.
[39] G. Roelfes, B. L. Feringa, *Angew. Chem., Int. Ed.* **2005**, *44*, 3230–3232.
[40] S. Park, H. Sugiyama, *Angew. Chem., Int. Ed. Engl.* **2010**, *49*, 3870–3878.
[41] C. Wang, G. Jia, J. Zhou, Y. Li, Y. Liu, S. Lu, C. Li, *Angew. Chem., Int. Ed.* **2012**, *51*, 9352–9355.
[42] G. Roelfes, A. J. Boersma, B. L. Feringa, *Chem. Commun.* **2006**, 635–637.
[43] A. J. Boersma, B. de Bruin, B. L. Feringa, G. Roelfes, *Chem. Commun.* **2012**, *48*, 2394–2396.
[44] N. S. Oltra, G. Roelfes, *Chem. Commun.* **2008**, 6039–6041.
[45] L. Gjonaj, G. Roelfes, *ChemCatChem* **2013**, *5*, 1718–1721.
[46] R. P. Megens, G. Roelfes, *Org. Biomol. Chem.* **2010**, *8*, 1387–1393.
[47] A. J. Boersma, B. L. Feringa, G. Roelfes, *Org. Lett.* **2007**, *9*, 3647–3650.
[48] K. Ishihara, M. Fushimi, M. Akakura, *Acc. Chem. Res.* **2007**, *40*, 1049–1055.
[49] S. Roe, D. J. Ritson, T. Garner, M. Searle, J. E. Moses, *Chem. Commun. (Cambridge, U. K.)* **2010**, *46*, 4309–4311.
[50] M. T. Reetz, *Chem. Rec.* **2012**, *12*, DOI 10.1002/tcr.201100043.
[51] M. T. Reetz, N. Jiao, *Angew. Chem., Int. Ed.* **2006**, *45*, 2416–2419.
[52] P. Deuss J., Artificial Metalloenzymes; Modified Proteins as Tuneable Transition Metal Catalysts, University of St Andrews, **2011**.
[53] K. Binnemans, P. Lenaerts, K. Driesen, C. Gorller-Walrand, *J. Mater. Chem.* **2004**, *14*, 191–195.
[54] M. T. Reetz, M. Rentzsch, A. Pletsch, A. Taglieber, F. Hollmann, R. J. G. Mondière, N. Dickmann, B. Höcker, S. Cerrone, M. C. Haeger, et al., *ChemBioChem* **2008**, *9*, 552–564.
[55] P. J. Deuss, G. Popa, A. M. Z. Slawin, W. Laan, P. C. J. Kamer, *ChemCatChem* **2013**, *5*, 1184–1191.
[56] S. Otto, F. Bertoncin, J. B. F. N. Engberts, *J. Am. Chem. Soc.* **1996**, *118*, 7702–

7707.

Chapter 5: Summary, conclusion and outlook

This thesis describes the synthesis and application of artificial metalloenzymes for transition metal catalysed reactions not performed by natural enzymes. The artificial metalloenzyme are based on a concept which is depicted in figure 1.

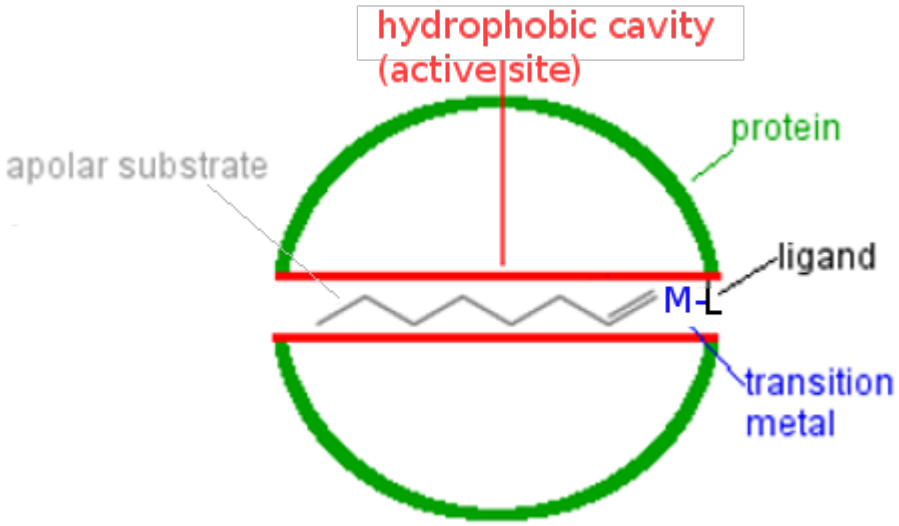


Fig. 1 Schematic representation of an ideal artificial metalloenzyme obtained by covalent modification.

The figure shows a protein with an active site (a hydrophobic tunnel) which is capable to bind the desired apolar substrate (1-octene as an example). At a well defined site of the protein a ligand with a high affinity to a transition metal is covalently introduced. For the method used in this thesis a unique cysteine needs to be present at the desired position to obtain such a species. The introduction of a unique amino acid at any position is a standard procedure in molecular biology. Finally, a transition metal is coordinated to this ligand. As the modification is located near the active site, the introduced metal is in close proximity to the substrate. Due to high substrate concentrations within the tunnel next to the transition metal rate enhancements should be observed. In addition, the protein environment should induce (enantio)selectivity.

The advantage of such an approach is that virtually any protein scaffold and any transition metal can be used. Additionally, the artificial metalloenzymes can be optimised by both chemical- (e.g. choice of metal and ligand structure) and biomolecular tools (e.g. optimisation of protein environment).

The main focus of the research was on the protein SCP-2L two mutants of which were used (SCP-2L A100C and SCP-2L V83C). Both mutants have a tunnel and bear the cysteine at opposite ends of the the tunnel. The two mutants were covalently modified with *N*- or P-ligands to introduce suitable binding sites for transition metals. The *N*-ligands were based on 1,10-phenanthroline (*Phen*)- or di-(2-picolyl)amine) (*Picol*)-entities, whereas the P-ligands were based on triphenylphosphine derivatives (fig.2).

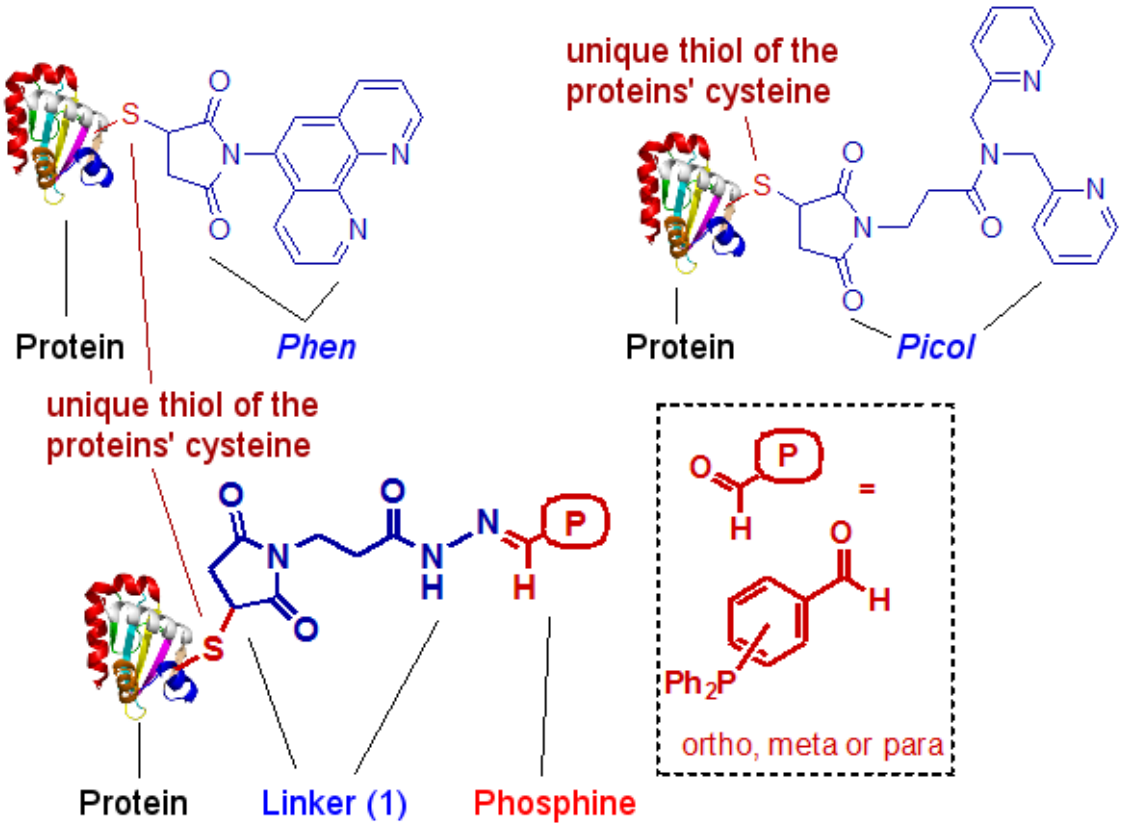


Fig. 2 Overview of the phosphine- and *N*-ligands used in the thesis.

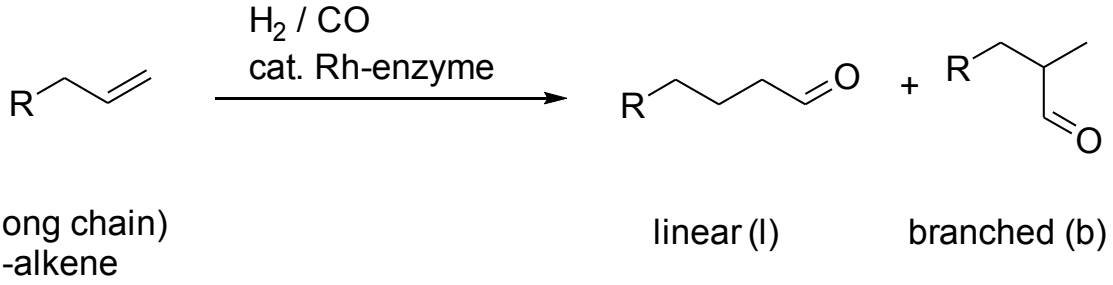
As any modification of a natural enzyme has an impact on its tertiary structure it would have been ideal to determine crystal structures of the artificial metalloenzymes to elucidate the protein environment next to the transition metal. This knowledge would

make it possible to fine-tune the active site by biological or chemical means in respect to its substrate- or metal coordination. It would have been almost as convenient but much more feasible to obtain crystal structures of the ligand modified proteins and use this structure to model the artificial metalloenzymes. It is also of utmost importance to determine the binding capabilities of artificial metalloenzymes before testing them in catalytic reactions. An artificial metalloenzyme with a small binding capability towards a metal is an undesired catalyst as a comparably high concentration of unligated metal would be present in solution. This unligated metal could outperform the catalytic activity of the artificial metalloenzyme itself which would make it impossible to draw any conclusions on the performance of the artificial metalloenzyme.
The topics of Chapter 2 are crystal structures and metal binding. The topic of chapter 3 is the catalytic performance of phosphine modified enzymes in the biphasic hydroformylation of long chain 1-alkenes. In chapter 4 the catalytic performance of Ni(II), Cu(II) and Co(II)-enzymes of *N*-ligand modified SCP-2L A100C in the Diels-Alder reaction between cyclopentadiene and *trans*-azachalcone is investigated.

In Chapter 2 the crystal structure of the unmodified SCP-2L A100C is discussed. The crystal structure shows that in comparison with its wild type the mutation has hardly any influence on the overall structure and the electrostatic surface potential and only a small effect on the structure of the tunnel. Unfortunately crystal structures of modified proteins could not be obtained.
By the addition of metal salts to the ligand modified proteins artificial metalloenzymes were obtained. For an evaluation of their catalytic capability the interaction of the ligand modified proteins towards different metal salts was investigated by ESI-MS and ICP-MS. It was found that the *Picol*-modified SCP-2L mutants are promising scaffolds for Pd(II)-enzymes while *Phen*-modified SCP-2L mutants are promising scaffolds for Pd(II)-, Cu(II),- Co(II) and Ni(II)-enzymes.
Phosphine modified SCP-2L V83C was found to be a suitable scaffold for Rh-enzymes regardless of the phosphine introduced. In case of phosphine modified SCP-2L A100C the P(para) modified protein is the most promising scaffold. PYP R52G-1-P(para) was also identified as a promising Rh-scaffold. Additionally SCP-2L A100C-1-P(para) was

found to be a promising scaffold for Ir-enzymes.

In Chapter 3 the previously synthesised Rh-enzymes are tested in the aqueous-organic biphasic hydroformylation of long chain 1-alkenes (see scheme 1).



Scheme 1: Hydroformylation of 1-alkene.

SCP-2L A100C and SCP-2L V83C modified with either of the three different phosphines were used as scaffold. The activity follows the following trends:

- In comparison with the known Rh/TPPTS catalyst system the Rh-enzymes were found to be several orders of magnitude more active than the reference while yielding comparable selectivities.
- artificial metalloenzymes based on SCP-2L A100C have a higher activity than SCP-2L V83C based ones.
- The activity of the phosphine modified proteins showed the following ranking: P(para) > P(meta) > P(ortho)
- The activity of the substrate is decreasing with increasing chain length.

The linear selectivity is high (≈ 75% linear selectivity) for all Rh-enzymes tested.
In order to rationalise the high activities and selectivities observed the mode of action was further investigated. The actual mode of action could not be determined, some could be excluded and the initially envisioned mode of action (see fig. 1) could not be excluded.
In Chapter 4 *N*-ligand modified SCP-2L A100C (see fig. 2) was used as scaffold for Ni-, Cu- and Co-enzymes. The artificial metalloenzymes are tested in the Diels-Alder reaction between cyclopentadiene and *trans*-azachalcone (see scheme 2).

Appendices

Chemicals and supplier

Acetonitrile	Aldrich, HPLC grade
Aluminum oxide (neutral)	Sigma-Aldrich
Ampicillin	Formedium
Argon gas	Boc
Benzamidine	Acros, 98%
Bradford reagent	Sigma-Aldrich
Calcium hydride	Acros, 93 %
Carbon monoxide gas	Boc
Chloramphenicol	Calbiochem
$Co(OAc)_2$	Aldrich, 99.9995%
Cobalt (II) nitrate * 6 H_2O	BDH, 97-101%
Copper (II) nitrate * 3 H_2O	Fluka, 99.0-104%
$Cu(OAc)_2$ * H_2O	Fisons Scientific, >99.0 %
C0mplete EDTA free protease inhibitor tablets	Roche Applied Science
Dicyclopentadiene	Aldrich
Dithiothreitol (DTT)	Formedium, 99.5%
Dimethyl formamide (DMF)	Acros, >99%
Diphenylether	Fluka, 99.9 %
DNase (DNase I from bovine pancreas)	Sigma-Aldrich
1-Decene	Sigma-Aldrich, 94%
1-Dodecene	Sigma-Aldrich, 94%
Ethanol	Fisher Scientific, analytical grade
$Fe(OAc)_2$	Aldrich, 95%
Glycerol	Sigma-Aldrich, 99%
1-Hexene	Fisher Scientific, general purpose grade
Hydrochloric acid	Fisher Scientific
Hydrogen gas	Boc
Hydrogen peroxide	Fisher Scientific, >30% (w/v)
Imidazole	Sigma-Aldrich, ≥98.5%
Iodine	Fisher Scientific
$Ir(acac)(CO)_2$	Strem, 99%
Isopropyl-1-thio-b-D-galactopyranoside (IPTG)	Formedium, 99.8%
Kanamycin	Formedium
Lysozyme (from chicken egg white)	Sigma-Aldrich
Magnesium	BDH
3-Maleimidopropionic acid hydrazide * HCl (linker, **1**)	Shanghai Speed Chemicals, >97%
MES	TCI, >98.0%

Methanol	Sigma-Aldrich, HPLC grade
$Mn(Oac)_2 * 4 H_2O$	≥99.0 %
Nickel (II) nitrate * $6 H_2O$	Sigma-Aldrich, ≥97.0 %
$Ni(OAc)_2 * 4 H_2O$	Fisons Laboratory, >98 %
Nitric acid	Acros, concentrated
Nitrogen gas	Boc
1-Octene	Acros, ≥99%
1-Octadecene	Sigma-Aldrich, >95%
Palmitic acid	BDH, 98%
Potassium phosphate	Fisher Scientific, analytical grade
$Pt(acac)_2$	97%
$Pd(OAc)_2$	Acros
$Re(CO)_5Br$	Strem, 98%
Sodium	Sigma-Aldrich
Sodium chloride	Fisher Scientific or Sigma-Aldrich
Styrene	Acros
Syn gas	Boc
Tetra hydro furane (THF)	Fisher Scientific, HPLC grade
Tris(hydroxymethyl) aminomethane (Tris)	Sigma-Aldrich, ≥99.9%
Sodium hydroxide	Fisher Scientific, analytical grade
$Rh(acac)(CO)_2$	Strem, 99%
$[Rh(MeCN)_2COD] BF_4$	Aldrich
$[Rh(Ph_3)_3CO] H$	Strem, 98%
$Ru(acac)_3$	Aldrich
Toluene	Sigma-Aldrich
Tris(3-sulfophenyl)phosphine trisodium salt hydrate (TPPTS)	J & K chemical LTD, 95%
Triton X-100	Sigma-Aldrich
$Zn(OAc)_2 * 2 H_2O$	Alfa Aesar, >99 %

ESI-MS data

ESI-MS of artificial metalloenzymes

Metal(X) salts were used as source (with X being a charge of "+1", "+2" or "+3"), therefore it is most likely that the metal coordinating to the protein will be present as metal(X) as well. It can be anticipated that the counterion is dissociating first in the MS conditions (positive ionisation), resulting in a cationic metallo-enzyme. If this assumption is correct the detected mass of all artificial metalloenzymes is X Da lower than the calculated mass as the program (MassLynx MaxEnt algorithm) assumes the presence of neutral starting species when processing the raw data and calculates as if each positive charge results from "H^+".

The masses mentioned in text and figures only correspond to peak maxima. In general each peak has a width at half height which corresponds to about ±5 Da of this value. Additionally the LC-MS was only calibrated on an irregular basis therefore the obtained masses may deviate by up to 16 Da from the calculated mass. As all conclusions are based on the mass differences between two peaks this deviation is not relevant.

Besides the assigned peaks most spectra show additional peaks which can be assigned to oxidised species, sodium adducts (up to four Na^+ (instead of H^+ were detected) and combinations thereof. These peaks are usually not explicitly mentioned.

Reaction conditons

The reaction conditions for the following figures are specified within the experimental section of the corresponding chapter. All reactions shown took place in 20 mM MES, 50 mM NaCl, pH=6. Except for PYP R52G, where the same buffer but at a pH of 7 was used.

The *Picol* and *Phen*-modified proteins were treated with 2 eq. of metal salt and left for 1 day. After 1 day the buffer was exchanged to 20 mM NH_4OAc (10000 Da MWCO) and the solution analysed by ESI-MS. Subsequently they were treated with EDTA and analysed by ESI-MS again

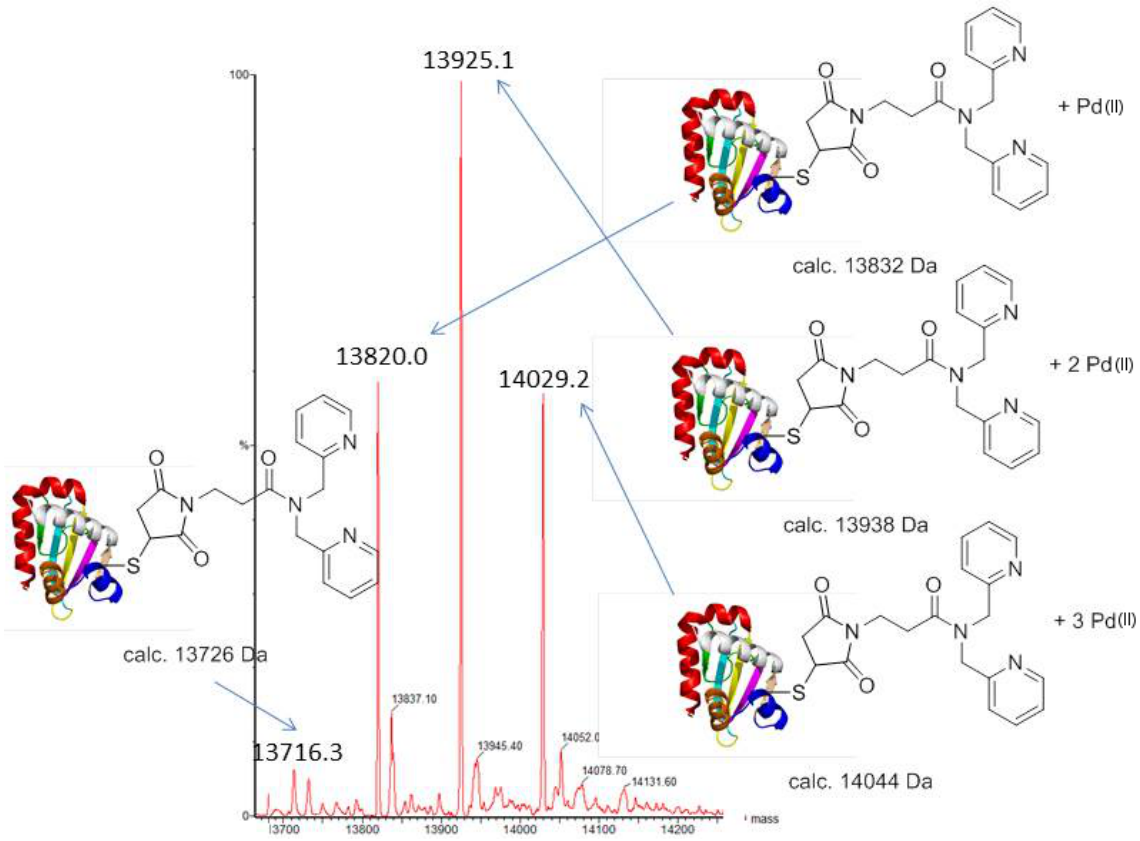


Fig. 1 ESI-MS of SCP-2L V83C-*Picol* + Pd(II)

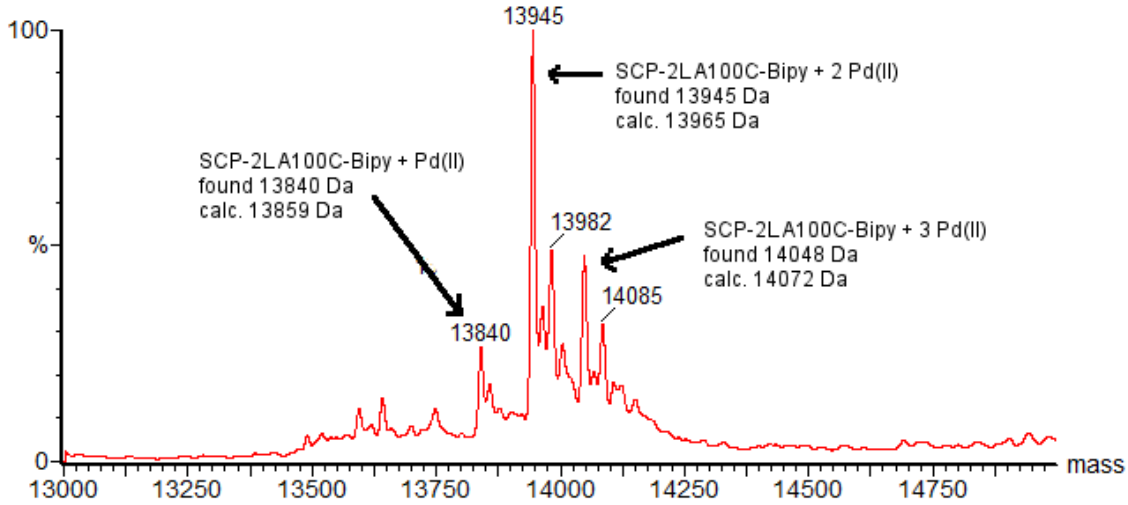


Fig. 2 ESI-MS of SCP-2L A100C-*Picol* + Pd(II).

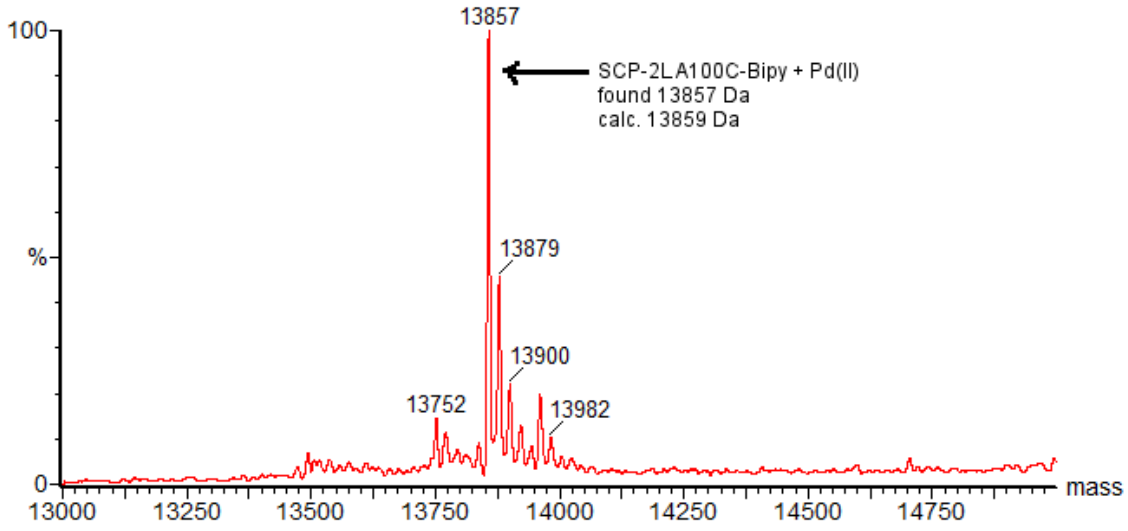


Fig. 3 ESI-MS of SCP-2L A100C-*Picol* + Pd(II) after EDTA-treatment.

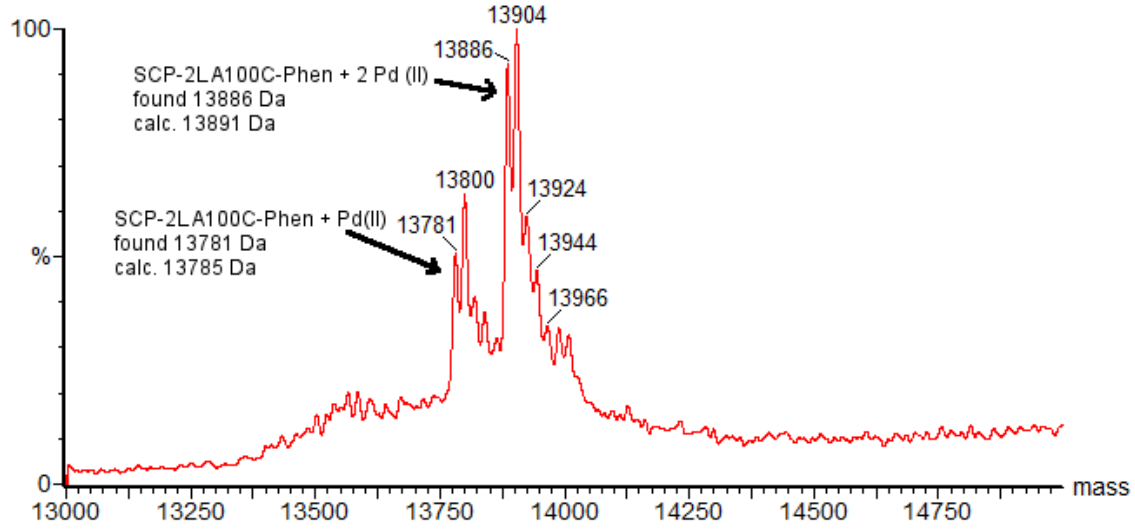


Fig 4 ESI-MS of A100C-*Phen* + Pd(II)

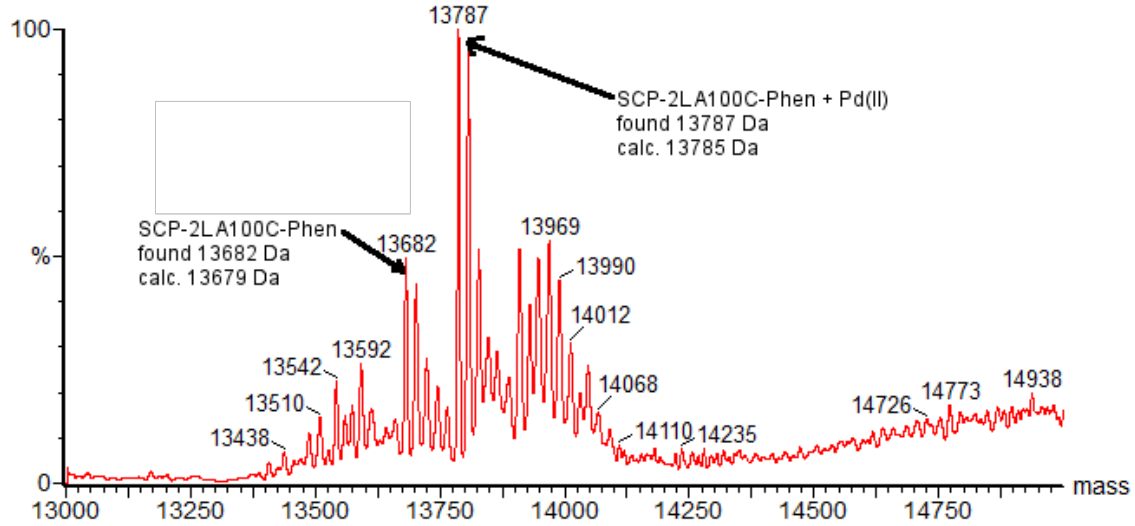


Fig 5 ESI-MS of SCP-2L A100C-*Phen* + Pd(II) after EDTA treatment.

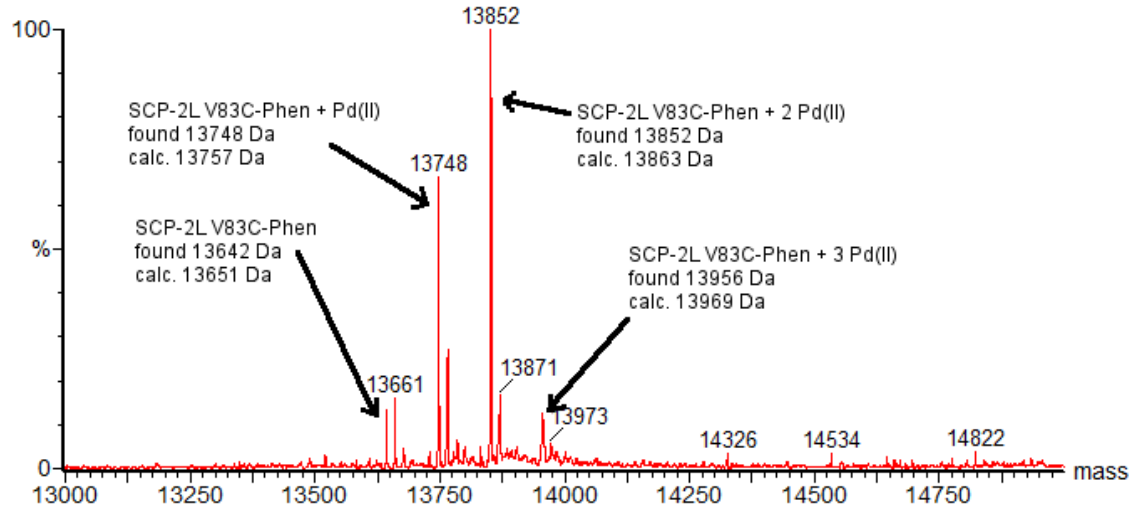


Fig. 6 ESI-MS of SCP-2L V83C-*Phen* + Pd(II)

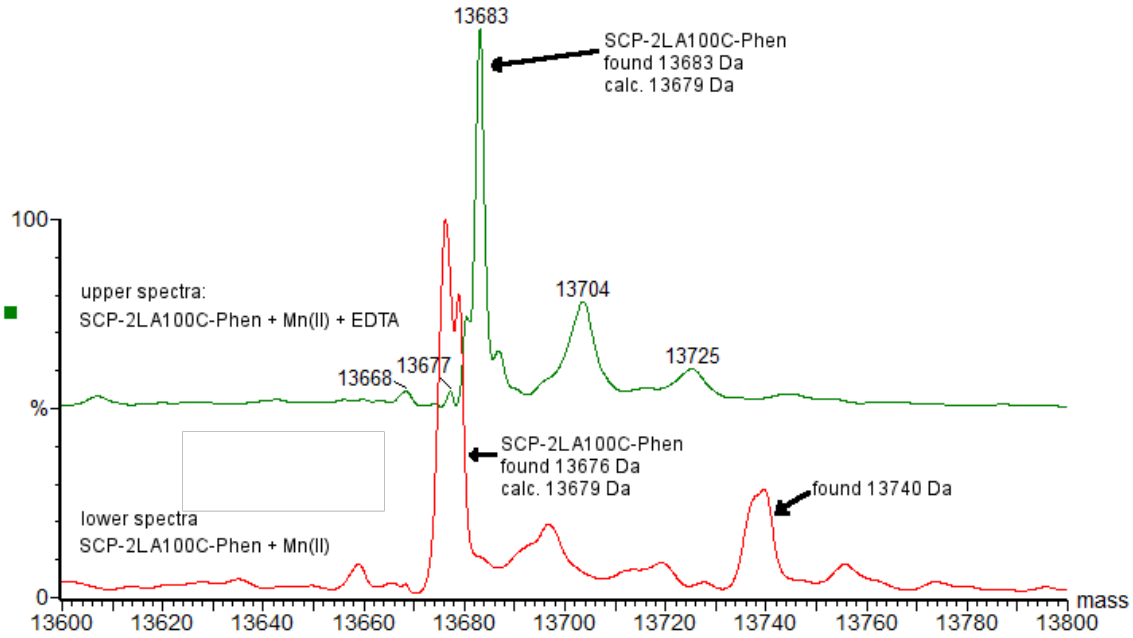


Fig. 7 top: ESI-MS of SCP-2L A100C-*Phen* + Mn(II) after EDTA treatment

bottom: ESI-MS of SCP-2L A100C-*Phen* + Mn(II)

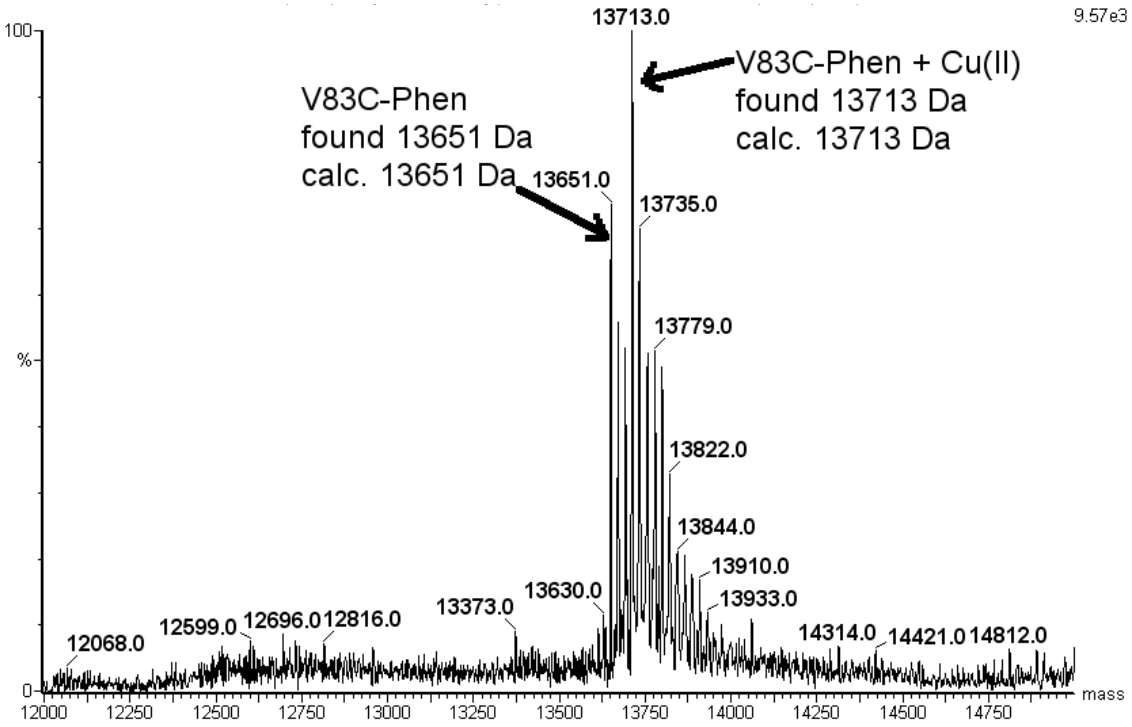


Fig. 8 ESI-MS of SCP-2L V83C-*Phen* + Cu(II).

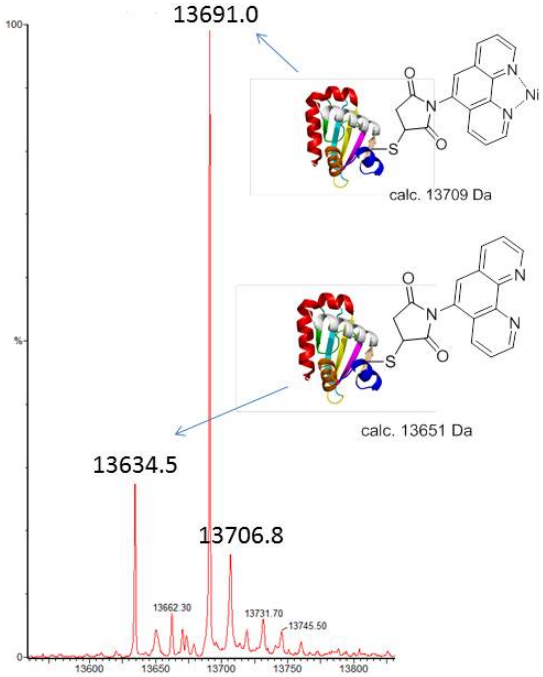


Fig. 9 ESI-MS of SCP-2L V83C-*Phen* + Ni(II).

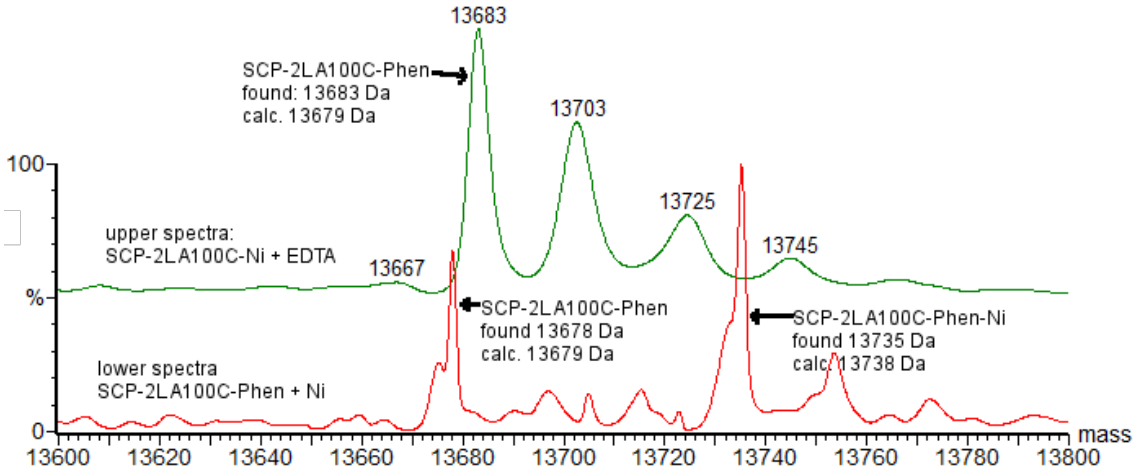


Fig. 10 top: ESI-MS of SCP-2L A100C-*Phen* + Ni(II) after EDTA treatment

bottom: ESI-MS of SCP-2L A100C-*Phen* + Ni(II)

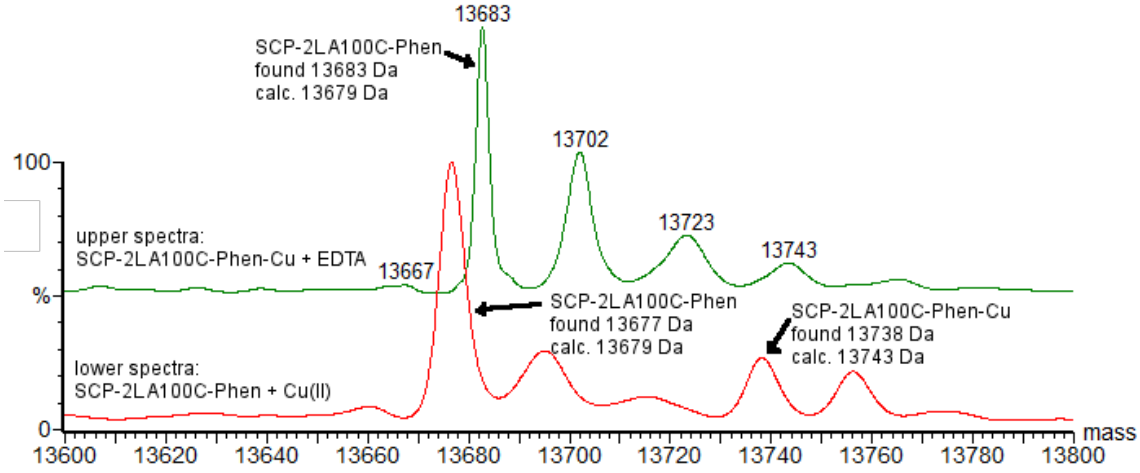


Fig. 11 top: ESI-MS of SCP-2L A100C-*Phen* + Cu(II) after EDTA treatment

bottom: ESI-MS of SCP-2L A100C-*Phen* + Cu(II)

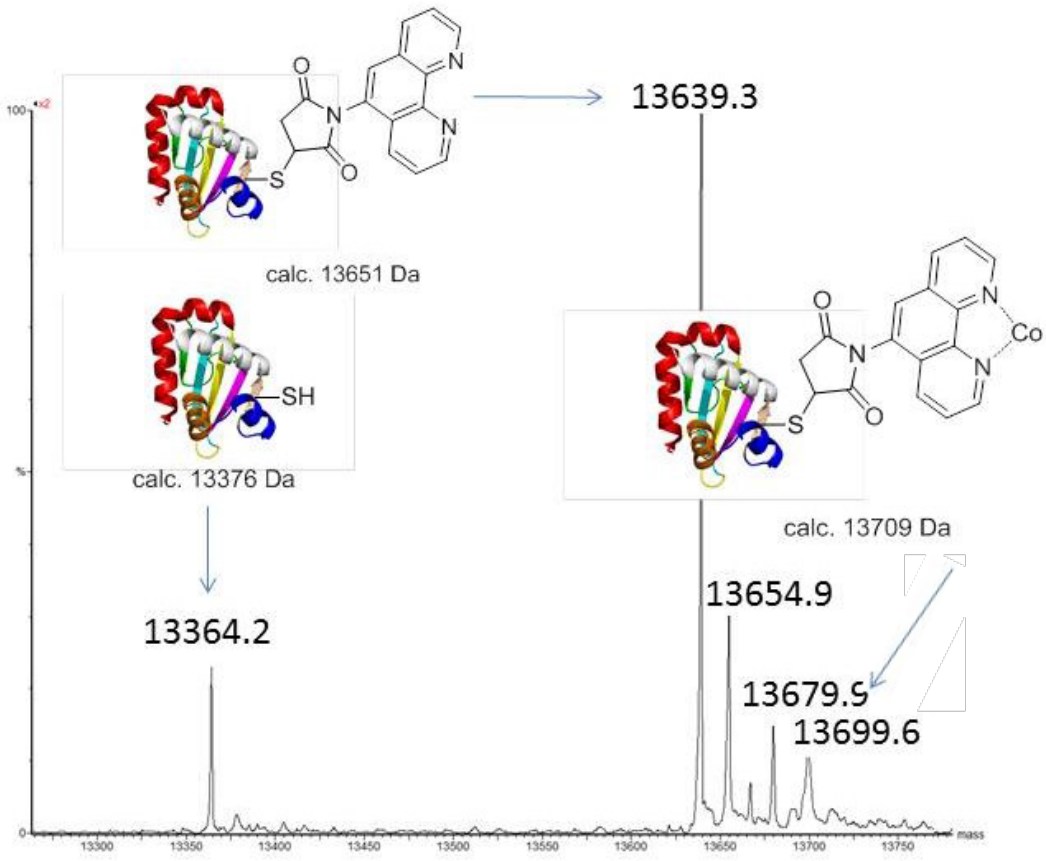


Fig. 12 ESI-MS of SCP-2L V83C-*Phen* + Co(II).

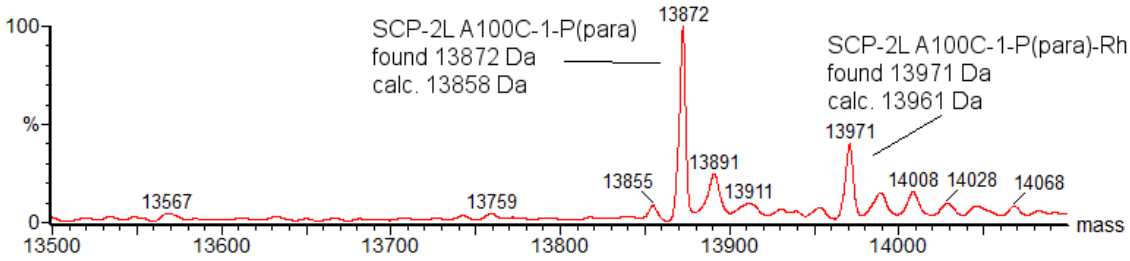


Fig. 13 ESI-MS of SCP-2L A100C-1-P(para) + $[Rh(MeCN)_2COD]BF_4$

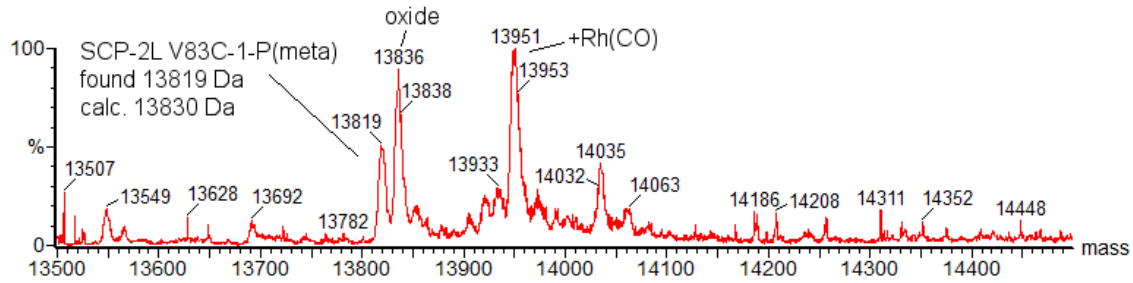


Fig. 14 ESI-MS of SCP-2L V83C-1-P(meta) + $Rh(acac)(CO)_2$.

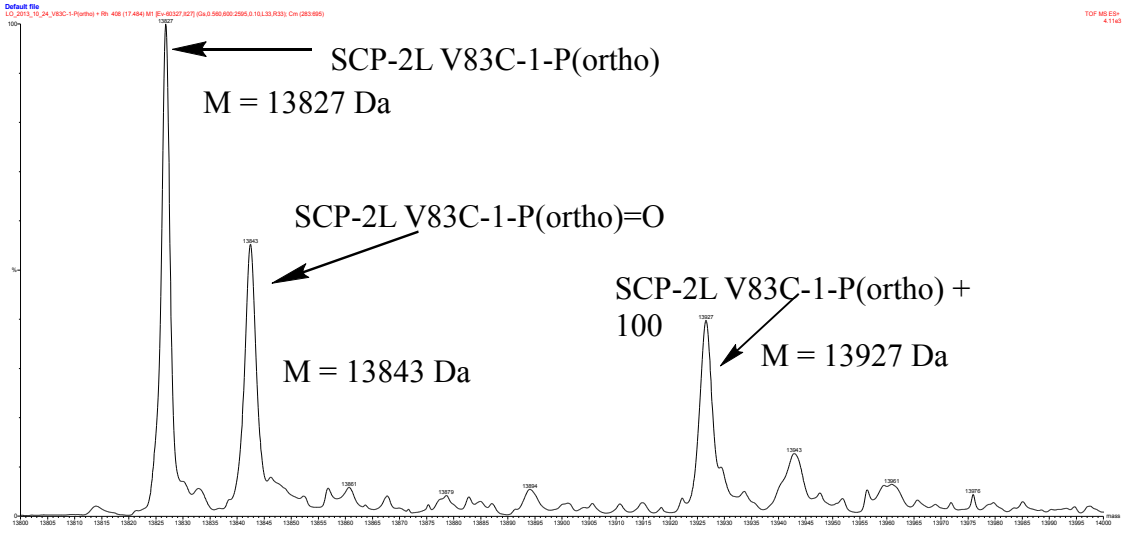


Fig. 15 ESI-MS of SCP-2L V83C-1-P(ortho) + $Rh(acac)(CO)_2$.

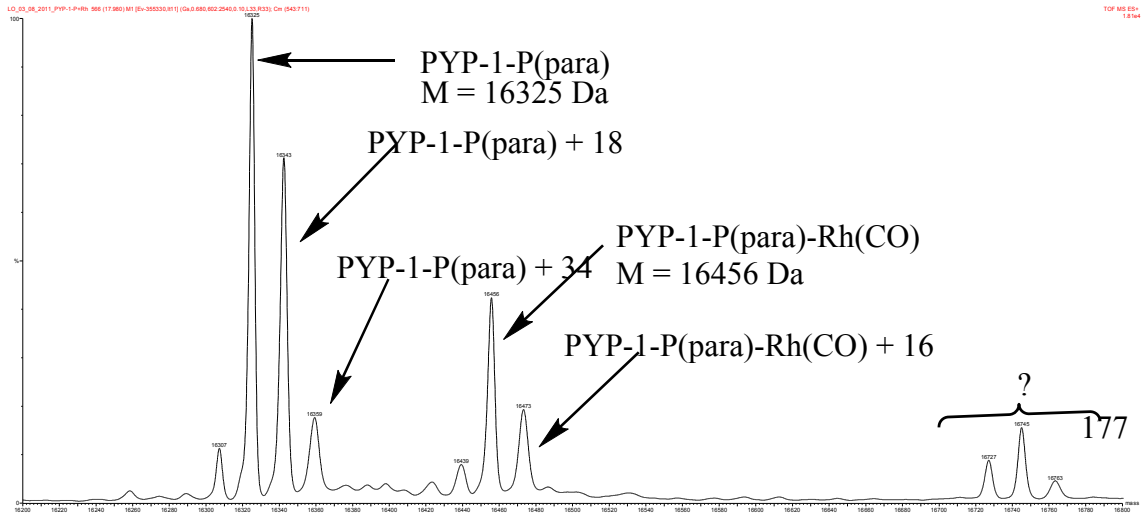


Fig. 16 ESI-MS of PYP R52G-1-P(para) + $Rh(acac)(CO)_2$.

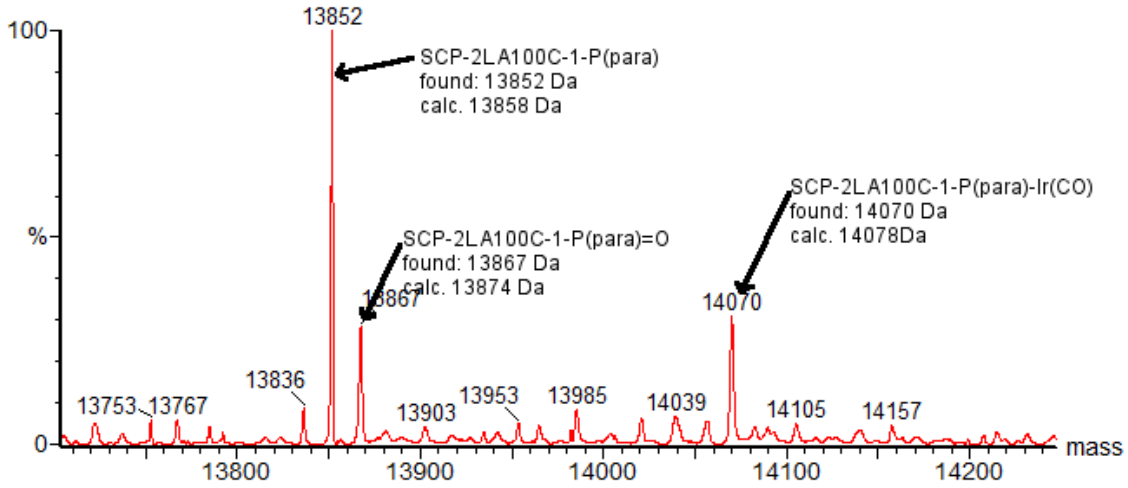


Fig. 17 ESI-MS SCP-2L A100C-1-P(para) + $Ir(acac)(CO)_2$.

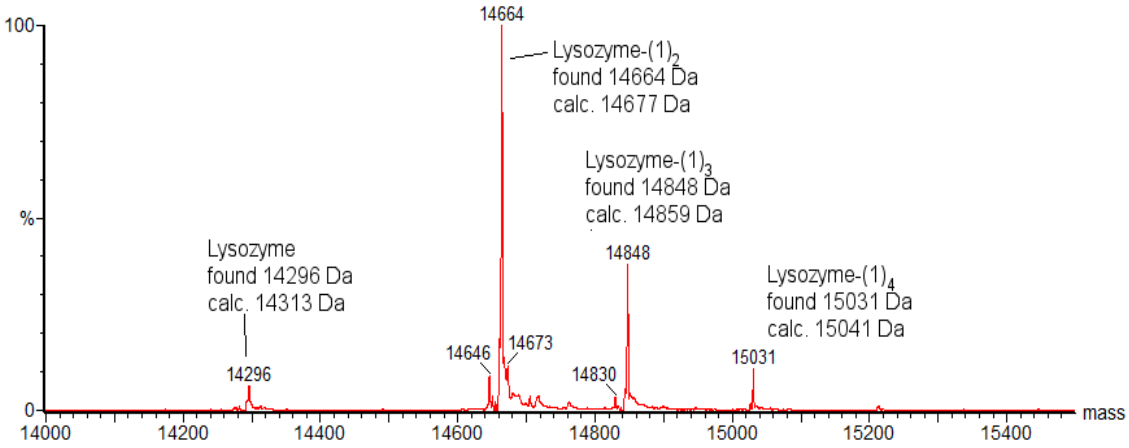


Fig. 18 ESI-MS Lysozyme-$(1)_x$ (TCEP pre treatment). Used as feedstock in further reactions and modifications.

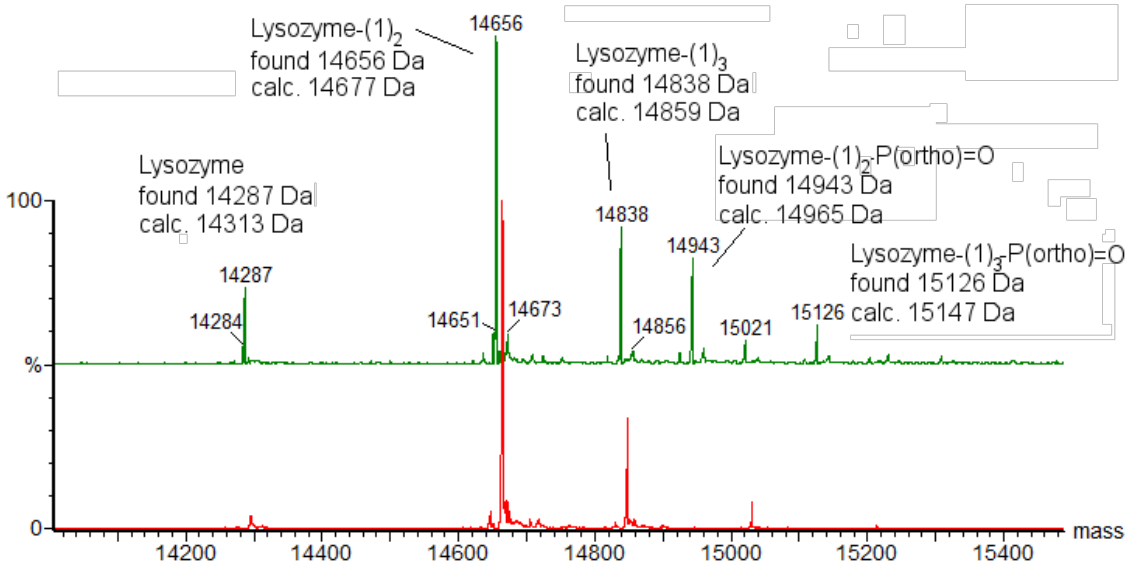


Fig. 19 top: ESI-MS of Lysozyme-$(1)_x$ + P(ortho). 3 eq. of phosphine added and the mixture left for 2 days. Due to low conversions another 3.8 eq. of the phosphine were added and the mixture left for 2 days. Due to low conversions another 3.8 eq. of the phosphine were added and the mixture left for 2 days (no significant reaction progress could be detected anymore)

Bottom: ESI-MS of Lysozyme-$(1)_x$ feedstock.

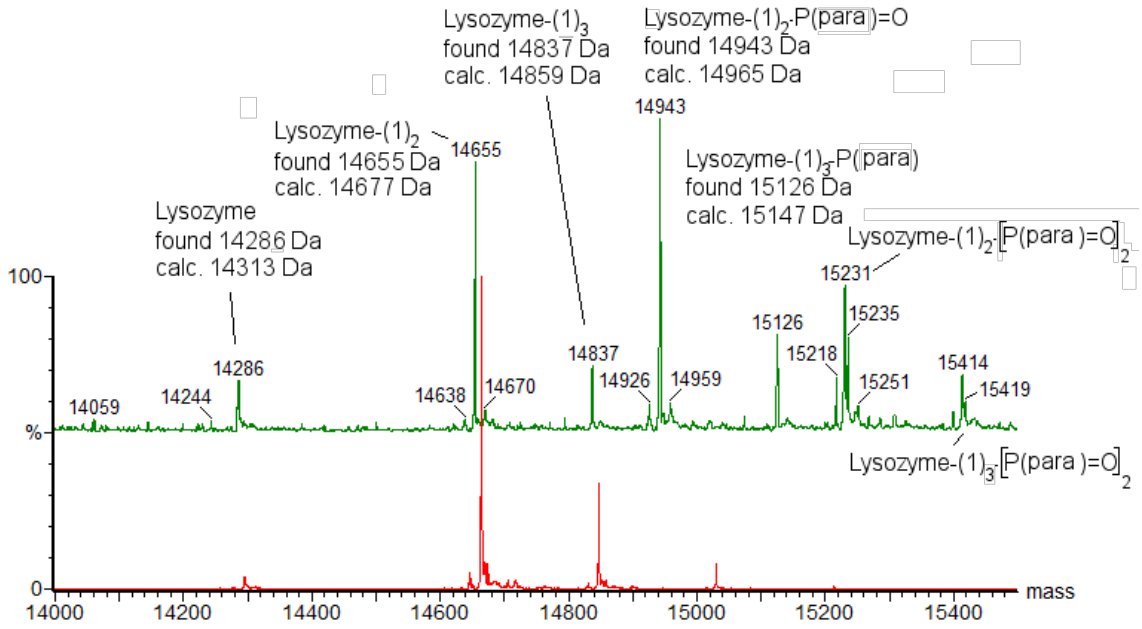


Fig. 20 top: ESI-MS of Lysozyme-$(1)_x$ + P(para). 5 eq. of phosphine added and the mixture left for 1 day. Due to low conversions another 10 eq. of the phosphine were added and the mixture left for 2 days. Due to low conversions another 10 eq. of the phosphine were added and the mixture left for 3 days (no significant reaction progress could be detected anymore)

Bottom: ESI-MS of Lysozyme-$(1)_x$ feedstock.

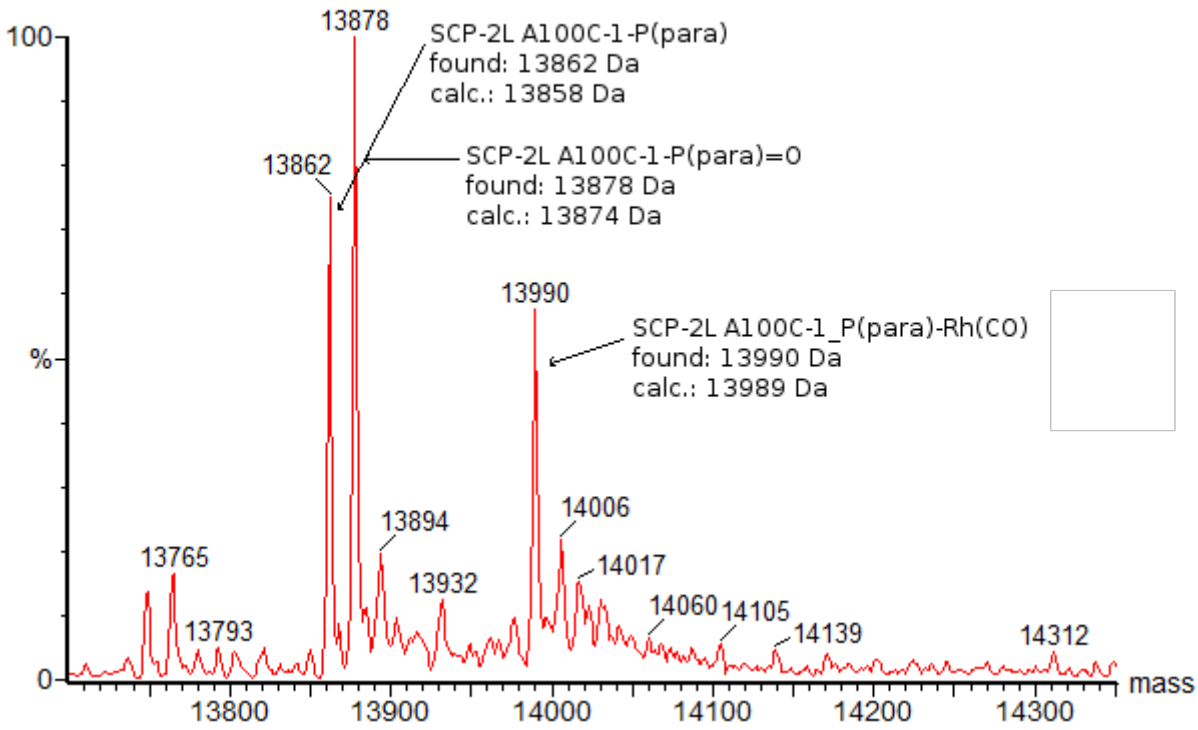


Fig. 21 ESI-MS of SCP-2L A100C-1-P(para) + $Rh(acac)(CO)_2$.

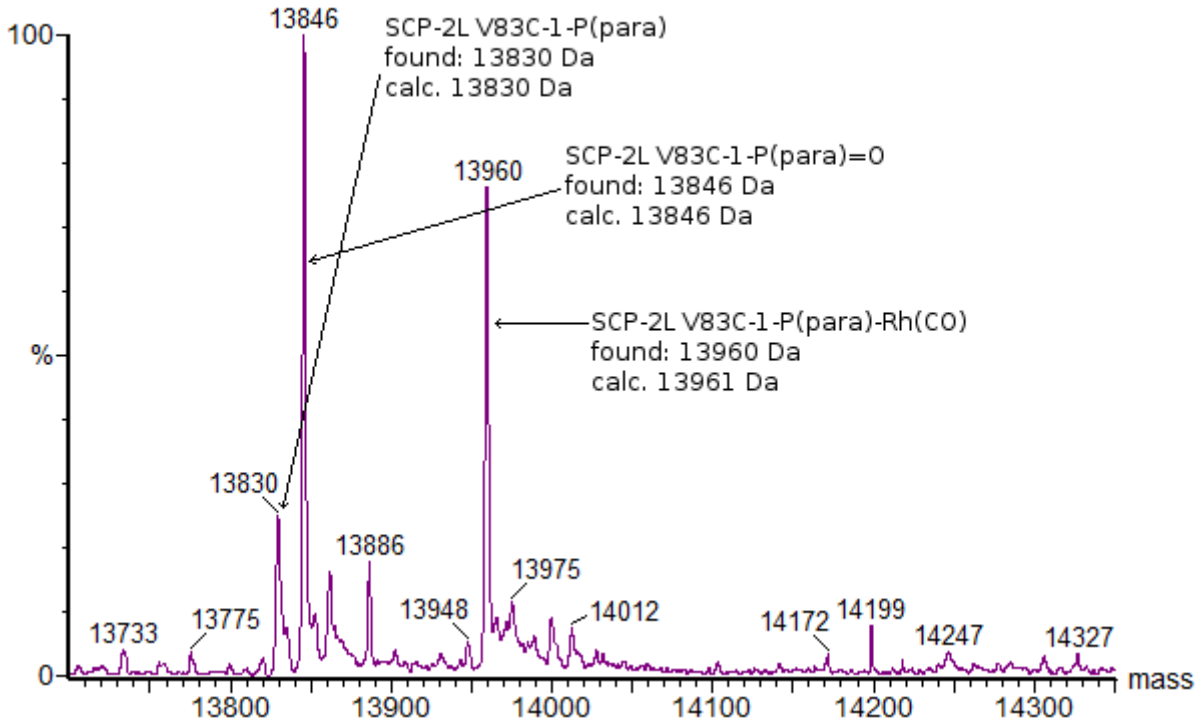
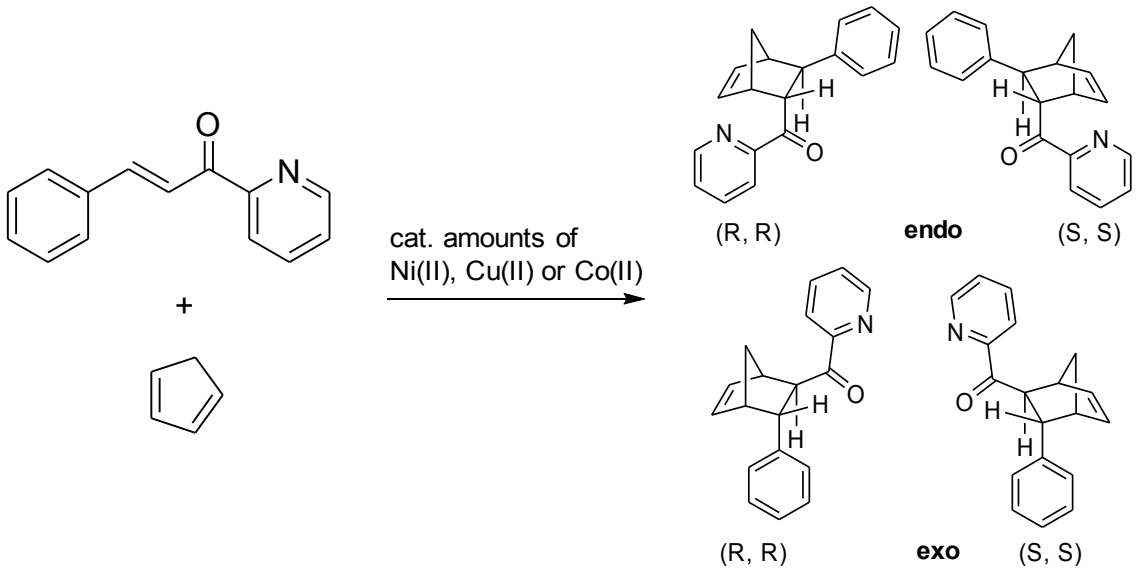


Fig. 22 ESI-MS of SCP-2L V83C-1-P(para) + $Rh(acac)(CO)_2$.



Scheme 2: Diels-Alder reaction between cyclopentadiene and *trans*-azachalcone.

Previous investigations (chapter 2) showed a high affinity between the *Phen*-modified protein to Cu(II) and Co(II), while the affinity to Ni(II) was smaller. The binding capability of the *Picol*-modified protein to each of these metal ions was significantly smaller.

The most promising catalyst was the Ni-enzyme of the *Phen*-modified protein though. In spite of its (comparably) low affinity towards nickel and therefore a respectively high concentration of unligated nickel this Ni-enzyme had the most pronounced effect on the (enantio)selectivity of the reaction. In comparision to the unligated metal a much smaller endo/exo-product ratio as well as 13% ee for the endo-adduct and 29% ee for the exo-adduct was detected. This means that the protein bound nickel has a high preference for the exo product and must induce a high enantioselectivity as a large quantity of the catalysis must be determined by unligated nickel. This Ni-enzyme is therefore worthwhile to be further investigated.

This thesis demonstrates the potential of new artificial metalloenzymes based on the covalent modification. Besides the optimisation of the tested artificial metalloenzymes it might be worthwhile to further investigate the potentially interesting catalysts identified in this thesis.

Acknowledgements

First of all I would like to thank my supervisor Paul Kamer who gave me the opportunity to work on this interesting project. I am really grateful that he let me (do my) work on this project - Thank you Paul!
I am also grateful to the University of St Andrews for a good working environment.
Thanks to all group members I got to know for all the fruitful (sometimes even about chemistry) discussions (mostly) during lunch & tea breaks.
Thanks to the following people with whom I shared the lab at some point:
Wouter ("the devil") Laan - although you tried to poison us once you have been a great labmate, really helpful and one of the most unstressed people I know! Peter (my predecessor) for giving me advice and useful input for the project from the beginning till the end. Phil (the penguin): The best labmate I will probably ever have (although you tried to kill me – paranoid people would probably begin to see a pattern). Don't grow up! Some people refer to the next group of people as "research students" - but I don't like this euphemism and I guess they got used to be called "slaves". Thanks for all your input & output. Ian - if you aren't a millionaire yet: Go for the Jackpot! Jaap: All the best. Esther: Remember the two rules I taught and *bamm* don't be so jumpy. Michael: Your music taste will get you far.
Besides "my" lab there's always been the large PCJK-lab – thanks to:
Michiel – Congratulations to your title. Jason ("of course") - keep inventing more useful equations! Christine – hope you finally listened to my advice and stopped working 24/7. Tanja – good luck back in the Netherlands and all the best in finding a job! Upu – Stop drinking that much booze! Oh, you still claim not to drink? Well, then stop sighing. Thanks for the free food (although way too mild, I like it hot..... NOT). Amanda, Megan & Andy – All the best to the "new generation" and good luck with those nasty artificial metalloenzymes! Dave: Seems I am the overall Looooooser as you already finished. Dave's "slave" Annika: Thanks for the castle trips. Roy: Good luck in Glasgow. The "group slaves": Torben, Lee & Paul S.: Your presence forced me to read quite a lot about frustrated lewis pairs. I enjoyed it. Special thanks to Frank & Amanda (although she's not been part of the "other lab") for a great last year with wisky, B&B, nice trips

and chemistry.
I always considered the DJCH-group as part of our group. Especially while Peter was still around I saw them roughly as often as the rest of my group (during lunch & tea breaks). Jan: Hope I managed to improve your table foottball skills. Ruben: Are you proud of your young Padawan? Thanks for chess, B&B and a nice time in general. Hope you know how to pronounce “Apfelmus” by now. Tina (although technically not DJCH): Take a deep breath and relax. Marc: Chess, B&B, castles, herring, golf,...,Jenny – Unlike most of the rest I will still see you on a regular basis I guess – thanks anyway! “Holiday”-Stu: Football (both active and passive), interesting literature and a nice last year in general. Special thanks to Peter. Your presence and all the (funny) stories during tea breaks were highly appreciated! Barthel (+wife) for a nice trip to the isle of Skye - stop snoring ;-) A big THANK YOU to Patrizia, Jacorien (+ William) – you were some of the few people in the ever changing environment who have been around during my whole time in St Andrews (although that's not completely accurate it felt as if). Also thanks to Dominik, Rosa, Thomas, Sabrina and those I forgot to list.
I also spent a lot of time in the biology labs and would like to thank everyone with whom I got in contact. Special thanks to Helen, Lei, Jude & Louise who were always helpful. Extra super überspecial thanks to Jane – I cannot imagine another person half as friendly, nice and helpful.
A very **big** thank you to Emma who managed to grow and get (refined) crystal structures of the mutant(s) (I can only imagine how much work it must have been) – also thanks for being at work during the weirdest hours (so I felt less weird). I will not mention any stories involving seagulls.
While writing about weird working hours I also have to mention and thank Stephanie – no matter when I was in it was highly likely that you were around as well.
Thanks Bobby - in the end there was only a small time gap before I actually understood what you just said. I guess you were just having pity with the retarded German and spoke with the poshest british accent possible. Whenever I had a serious problem you found the time to fix it as soon as possible – I won't mention the not so serious problems. ICP-MS: Sylvia & Lorna. Thank you so much for the hard work you put into the ICP analysis. I really appreciate it! LC/ESI-MS: Sally & Catherine: You were

always friendly and helpful - thanks a lot! Dr. José Antonio Fuentes García is greatly acknowledged for his help and supervision regarding the HPLC analysis. I am also grateful to the following groups for providing me with some of the metal precursors used in this thesis: SPN, PK, MC and DJCH. I am grateful to the following biology groups for letting me use their equipment: JHN, TKS, GLT, TMG, RRR, US and MFW. Special thanks to the GLT-group whose workspace & equipment I was using & blocking throughout all these years. Nevertheless you always treated me really nice.

Almost forgot them - but thanks to the USIC-organisaton team: Brian, César, Phil. Special Thanks to Siobhan who "translated" and corrected all of my grant applications. It was an awful lot of work but it was also a great conference which was mainly thanks to our (plenary) speakers: Professor Fraser Armstrong, Dr. Victor F. Quiroga Norambuena. and Professor Dominic Wright. In addition to the plenary speakers I am grateful to Paul Bane, Professor David J. Cole-Hamilton and Professor Derek Woollins who did entertain us (or at least me) during all of our conference dinners

Thanks to all the people who managed to visit me during my diaspora: Flock, Thorsten, my brother and my parents, Flo, Jelly, Hanna, Raf, Jenny.

I also want to thank everyone who will actually read my thesis and everyone I forgot to mention.

While lab-work was mostly fun and enjoyable I would have struggled to actually write this thesis without the following people (and companies): Ritter Sport & Choceur (chocolate companies) as well as various wisky companies (especially Lagavullin DE). Without these necessities I would have gone crazy! Any errors you spot result from too little wisky consumed while writing. Special thanks to Flock who managed to distract me on a regularly basis (which was not too complicated I have to admit). I am also grateful to the people who developed/improved the following (free to use) software: Pymol, APBS, Gimp, Open Office, Zotero, Wincoot, Firemath and everyone who wrote tutorials and such on how to use certain features. Thanks to my parents for providing a place to sleep and food. Last but not least I want to thank my dad for proofreading the fragments in the beginning – as well as the final thesis.