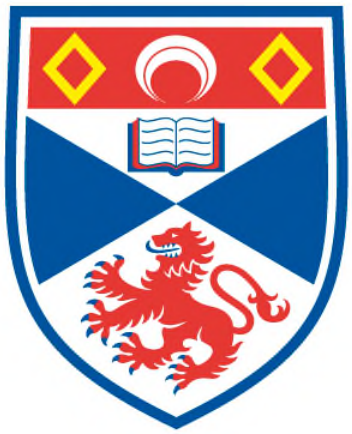
DEVELOPMENT OF ARTIFICIAL METALLOENZYMES VIA COVALENT MODIFICATION OF PROTEINS

Gina Popa

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



2010

University of St Andrews

School of Chemistry

DEVELOPMENT OF ARTIFICIAL METALLOENZYMES VIA COVALENT MODIFICATION OF PROTEINS

PhD thesis

2010

Gina Popa

1. Candidate's declarations:

I, GINA POPA, hereby certify that this thesis, which is approximately 45,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in October 2005 and as a candidate for the degree of PhD in October 2005; the higher study for which this is a record was carried out in the University of St Andrews between 2005 and 2010.

Date 01/09/2010 signature of candidate ……GINA POPA…

2. Supervisor's declaration:

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

Date 01/09/2010 signature of supervisor …….

3. Permission for electronic 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 electronic publication of this thesis:

Embargo on both of printed copy and electronic copy for the same fixed period of 3 years (*maximum five*) on the following ground(s):

- publication would preclude future publication.

Date 01/09/2010
signature of candidate …GINA POPA signature of supervisor …Prof. PAUL C.J. KAMER

.

Table of contents

Acknowledgements

Firstly, I would like to thank my PhD supervisor, Prof. Paul Kamer for entrusting me with this interesting and challenging project, for his encouragement, help and academic supervision.
I am also grateful to Prof. Ron Wever and Dr Paul Wright for agreeing to be the co-examiners in my dissertation committee.
Special thanks to Dr Wouter Laan for his guidance during these years, for being a very demanding, rigorous and constructive scientific supervisor and for kindly proof-reading the thesis.

I would like to thank Prof Garry Taylor for allowing the use of laboratory space and equipment at the Centre for Biomolecular Sciences – St Andrews.
I would like to acknowledge the BSRC Mass Spectrometry and Proteomics Department team, in particular Dr Catherine Botting and Dr Sally Shirran for helpful discussions and technical support in performing the LC-MS measurements and also Mr. Alex Houston for the MALDI TOF measurements.
I would also like to thank the Solution-phase NMR Spectroscopy Service team, Dr Tomas Lebl and Mrs Melanja Smith for their help for the NMR measurements. I am thankful to Prof. Malcolm White for allowing us to perform the fluorescence measurements in his laboratory.
I acknowledge the EPSRC National Mass Spectrometry Service Centre – Swansea for the accurate mass measurements performed there.
I thank Prof. David A. Bernlohr (University of Minnesota) for kindly providing the pRSET-ALBP plasmid, as well as Dr Julio J. Caramelo (Fundación Instituto Leloir and Instituto de Investigaciones Bioquímicas de Buenos Aires), for kindly providing the pET22b(+) Δ98Δ IFAB plasmid.
This work was financially supported by European Union Marie Curie Excellence Grants (MEXT-2004- 014320).

I am very thankful to all former and current colleagues from the Hom-Cat group. I have to start with thanks for our special colleague Bert, for his joyfulness and

brightness. Special thanks for Serghei, Bianca, Gregorio, Peter, Arnald, Debbie and Jason for their fruitful discussions and help with the organic chemistry project but also for being great friends and exceptional lab mates. I am indebted also to Yanmei and Nikos for the synthetic chemistry part. I want to thank Christine, Tanja and Michiel for the good and efficient time we spent together as a team.
I am especially grateful to the Applecross girls for the breath of fresh air… provided every day when it was crucially needed. I feel very lucky and fulfilled for having the chance to meet interesting people and to make great friends here in the Chemistry Department, and I must thank Jenny, Bianca, Marzia, Aga, Jacorien and Louise for their permanent encouragements, support and friendship. Special thanks also for Elena, Alma, Cristian, Dragos and Andrei, for bringing into St Andrews the flavour of home.

Many thanks also to Prof. Mircea Bulancea and Prof. Rodica Segal for their trust and encouragements at the beginning of my career. Without their initial support I would not be here today.

Most importantly, I am grateful to my very special and loving family, for their permanent support. Special thanks to my sister for finding the strength to encourage and help me whilst she has to deal with her own life threatening problems, for my brother for being such a helpful, reliable and honest person and to my parents for teaching me through their personal examples the real values of life. Particular thanks to my daughter, who bravely started a new life in harsh conditions. I have to thank you for being so mature and determined in your work and for being such a good, patient and understanding person during my worst and best days here. What does not kill you makes you stronger and we are both alive!

Last but not least, I want to thank God for giving me a second chance in life!

Abstract

Development of selective artificial metalloenzymes by combining the biological concepts for selective recognition with those of transition metal catalysis has received much attention during the last decade. Targeting covalent incorporation of organometallic catalysts into proteins, we explored site-selective covalent coupling of phosphane and N–containing ligands. The successful approach for incorporation of phosphane ligands we report herein consists of site-specific covalent coupling of a maleimide functionalized hydrazide into proteins, followed by coupling of aldehyde functionalized phosphanes via a hydrazone linkage. Site selective incorporation of N–containing ligands was obtained by coupling maleimide functionalized N–ligands to proteins via Michael addition to the maleimide double bond. These two methods can be easily applied to virtually any protein displaying a single reactive cysteine and allows a wide range of possibilities in terms of cofactor design.

Site-specific covalent incorporation of transition metal complexes of phosphane ligands into proteins was successfully obtained. The success of the approach is influenced by several factors like the metal precursor, the phosphane type and the protein scaffold.

Metal complexes of 5–maleimido–1,10–phenanthroline modified proteins were formed *in situ*, via addition of a metal precursor to the phenanthroline modified proteins or by coupling preformed metal complexes to proteins via Michael addition of the thiol group from a cysteine residue to the maleimide double bond of the N-ligand.

These successful coupling methods enable the use of a wide range of protein structures as templates for the preparation of artificial transition metalloenzymes, which opens the way to full exploration of the power of selective molecular recognition of proteins in transition metal catalysis.

List of Abbreviations

ALBP	adipocyte lipid-binding protein
ANS	1-anilino-8-naphthalene sulfonate
Apo-Mb	apomyoglobin
Bpy	2,2'-bipyridine
BSA	bovine serum albumin
CDI	1,1'-carbonyldiimidazole
CE	capillary electrophoresis
COD	1,5-cyclooctadiene
Cp*	1,2,3,4,5-pentamethylcyclopentadienyl
CPA	carboxypeptidase A
DABCO	1,4-diazabicyclo[2.2.2]octane
DCC	*N,N′*-dicyclohexylcarbodiimide
DCM	CH_2Cl_2-dichloromethane
DKR	dynamic kinetic resolution
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DTNB (Ellman's reagent)	5-(3-carboxy-4-nitrophenyl)disulfanyl-2-nitrobenzoic acid
DTT	(*2S,3S*)-1,4-bis-sulfanylbutane-2,3-diol
e.e.	enantiomeric excess
EDTA	2,2',2",2"'-(ethane-1,2-diyldinitrilo)tetraacetic acid
EPR	electron paramagnetic resonance
eq.	equivalent
ESI-MS	electrospray ionization mass spectrometry
EtOAc	ethylacetate;
FABPs	fatty acid-binding proteins
g	gram
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GC	gas chromatography

GuCl	guanidinium chloride
h	hour
HABA	2-(4'-hydroxyphenyl-azo)benzoate
HOBt	1-hydroxybenzotriazole
HPLC	high performance liquid chromatography
HSA	human serum albumin
IFABP	Intestinal Fatty Acid Binding Protein
IPTG	isopropyl β-D-1-thiogalactopyranoside
l	liter
LB	Lysogeny broth
LC-MS	liquid chromatography-mass spectrometry
m	meter; multiplet
M	mol/l; molar; transition metal
MALDI-TOF	matrix-assisted laser desorption/ionization - time-of-flight
Mb	myoglobin
MeCN	acetonitrile
MeOD	deuterated methanol
MeOH	methanol
min	minute
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MWCO	Molecular Weight Cut Off
N∩N	N,N bidentate chelating ligand
NHS	*N*-hydroxysuccinimide
NMR	nuclear magnetic resonance
OD_{600}	optical density measured at a wavelength of 600 nm
PB	Protein Broth medium
PCR	polymerase chain reaction
Phen	1,10-phenanthroline
ppm	parts per milion
PSA	pig serum albumin
Px	pyridoxamine
r.t.	room temperature

rpm	rotations per minute
RSA	rabbit serum albumin
SDS-PAGE	sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SSA	sheep serum albumin
TBHP	*tert*-butyl hydroperoxide
tBu	*tert*-butyl
TCEP	*tris*(2-carboxyethyl)phosphine
THF	tetrahydrofuran
Tris	2-amino-2-hydroxymethyl-propane-1,3-diol

I. INTRODUCTION

I.1 Catalysis and sustainable chemistry

Waste-free, environmentally friendly production processes are an absolute necessity for a sustainable society. The atom economic synthesis of many natural products has been optimized by nature over millions of years and is realized by efficient biocatalysts, enzymes. However, for the synthesis of many important and widely used pharmaceuticals, agrochemicals, fine and bulk chemicals there are no enzymes available. In industrial synthesis, several biological or bioinspired catalytic strategies have been developed lately.[1-2] Still, the number and diversity of these applications are modest and the demand for new strategies remains high.

Catalysis is widely applied, from research laboratories to the chemical industry (petrochemistry, bulk chemicals and fine chemicals industry). Also food processing and many biological processes rely on catalysis. It is estimated that 90% of all chemical products involve catalysis during different phases of their production.[3] Catalytic synthesis is superior to stoichiometric organic synthesis not only from an economic, but especially from an environmental point of view. The amount of waste generated is reduced as a result of improved selectivties and energy consumption is lowered by reduction of reactions time and mildness of the reactions conditions.[4] Therefore, a solution to the environmental problems generated by chemical synthesis is the development of new, highly selective catalysts, capable to catalyse a wide range of chemical reactions and to improve reaction rates and yields.[5]

The idea of catalysed processes was first introduced by Jöns Jakob Berzelius in 1835, who investigated systematically the previously recorded observations about changes in substances when they were brought in contact with small amounts of species called “ferments”.[6-7] Wilhelm Ostwald, in 1894, was the first to give definition for a catalyst: a substance that changes the rate of a chemical reaction but which is not consumed by the reaction.[8] According to Ostwald’s definition, a catalyst can also slow down a reaction. Currently, a catalyst is defined as a substance that increases the rate of a reaction while remaining itself unchanged. The catalyst affects only the rate of the reaction; it does not

change the thermodynamics of the reaction and therefore not the equilibrium composition.[9]

Catalysis can be divided into three distinct categories:

- heterogeneous catalysis – the area of catalysis involving substrates and catalyst components in different phases. A heterogeneous catalytic reaction takes place usually at the surface provided by the catalyst. A schematic mechanism is represented by diffusion and adsorption of one or more reactants to the catalytic surface, when the reactants bonds become weaker and allow product bonds formation. This is followed by product desorption from the surface at the end of the reaction, facilitated by the weak bonds existing between catalyst and product, and diffusion of the product into the reactor environment.
- homogeneous catalysis - the substrates for reactions and the catalyst components are in the same phase, most often the liquid phase. Homogeneous catalysis can be further divided, corresponding to the nature of the catalysts, into nucleophilic catalysis (using Lewis bases, Brønsted bases as catalysts), electrophilic catalysis (using Lewis acids, Brønsted acids as catalysts), organocatalysis, which uses simple organic molecules as catalysts and organometallic catalysis which deals with transition metal complexes.
- enzymatic catalysis - catalysis that occurs due to the presence of enzymes or complete microorganisms.

Homogeneous catalysis has been extensively used in organic synthesis. However, limitations like a low selectivity or a reduced life time shown by some synthetic catalytic systems, converted at the end in a low efficiency, require for major improvements. The possibility of ligand tuning requires experience, a good knowledge of the reactions mechanism and of the molecular modelling techniques, but also relies on intuition and a great deal of trial and error.[10] Moreover, in order to optimize the catalyst performance, not only the steric and electronic properties of the ligand should be considered, but also subtle experimental parameters like temperature, solvent or presence of additives.[10] Therefore, the methodologies used to develop new catalysts for modern organic synthesis have to deal with all these problems.

Industrial application of biocatalysis for organic transformations has increased during the last 20 years, mainly because of the high activity, chemo-, regio- and enantioselectivity displayed by enzymes. The possibility to use some of them in organic

solvents expanded also their applicability.[11] The rapid progress in genomics, directed enzyme evolution and the entire area of biotechnological engineering surpassed many limitations of enzymes applicability.[12] However, solutions still have to be found for its remaining limitations like operational stability of enzymes and limited substrate and reaction scope.[12]

An ideal catalyst seems to be the one capable of combining the best properties of both areas: biocatalysis and homogeneous catalysis. It should demonstrate not only good selectivity and high activity, but also high atom efficiency, low E-factor and industrial applicability. An answer to these challenges could be represented by artificial metalloenzymes, capable of combining the two domains and allowing using man-made catalysts in organic synthesis under biological catalytic efficiency.

I.2 Enantioselective catalysis

The term enantiomers refer to stereoisomers that are nonsuperimposable complete mirror images of each other. The differences in properties between enantiomers are not displayed in regular environments, where enantiomers have identical physico-chemical properties, but in asymmetric environments, like biological systems. The building blocks of biological systems, amino acids and carbohydrates, are chiral and exist mainly in single enantiomeric forms, consequently the chirality of the biological environments is formed by them. In these chiral biological and living systems, chiral compounds like drugs, fragrances, crop protectors, flavours etc. are recognized stereoselectively. The difference in properties of two possible enantiomers could be the result of this stereoselective interaction with chiral systems.[13] Consequently, there is great scientific and commercial interest for the preparation of enantiomerically pure compounds. For example, the chiral drugs produced by the pharmaceutical industry needs to be tested for the influence of chirality on all of their pharmacological, pharmacokinetic and toxicological properties. This requirement is now a standard part of the drug research and development process[14] and is controlled by strict regulations.[15] When a new chiral drug is discovered, the development and commercialization of either a racemate or a single enantiomer needs to be established from the first stages. Most often a single enantiomer is developed, the opposite enantiomer being considered as an impurity that is discarded when being detected.[13]

A prime example of the impact of chirality on drugs safety is the thalidomide case. This drug was widely prescribed to pregnant women in the late 1950's and early 1960's to combat morning sickness. Before being released, the influence of chirality on drug activity and safety was insufficiently tested. Thalidomide was used as a racemate, but while the (*R*) enantiomer is effective against morning nausea, the (*S*) enantiomer is teratogenic, resulting in severe birth malformations for the children of women who had taken thalidomide during their pregnancies.[16] The two enantiomers can interconvert *in vivo*, therefore it is not possible to prevent the teratogenic effect by administration of only one enantiomer.[17]

Considering the importance of chirality, efficient enantioselective technologies have been developed for production and analysis of enantiomerically pure compounds.[13] Three fundamentally different approaches are applied in order to obtain enantiomerically pure compounds:[18-19]

- Resolution: separation of enantiomers by chemical and physical means – e.g. by crystallization or liquid chromatography using chiral stationary phases. Also to be mentioned is the interest during the last years in the development and application of new miniaturized separation techniques, like electrochromatography or nanotechnologies.[13]
- Chiral pool synthesis – the final chiral molecule is obtained from chiral substrates, usually originating from natural products. The method can be applied on a large scale only depending on the commercial availability of the starting material.[20]
- Enantioselective synthesis, either stoichiometric or catalytic – preparation of chiral compounds from achiral precursors, using chiral reagents or chiral catalysts. Enantioselective catalysis is one of the most efficient methods in terms of atom economy.[21]

I.2.a. Enantioselective homogeneous catalysis

Enantioselective organo- and organometallic catalytic reactions have been studied intensively during the last forty years because of their high efficiency in producing enantiopure compounds. The importance of work accomplished in this area was recognized by a Nobel Prize awarded in 2001 to W.S. Knowles[22] and R. Noyori[23] for

their contribution on chirally catalysed hydrogenation reactions and to K. B. Sharpless[24], for his contribution on chirally catalyzed oxidation reactions.

The history of enantioselective homogeneous catalysis starts in 1965 with the use of an achiral catalyst: chlorotris-(triphenylphosphane)rhodium(I) - [$RhCl(PPh_3)_3$] by Wilkinson and co-workers, as catalyst for the hydrogenation of olefins, demonstrating for the first time the activation of molecular hydrogen by a metal complex (Figure 1). Although homogeneous catalysts had been reported before, this was the first one having rates comparable with those reported for the heterogeneous counterparts.[25]

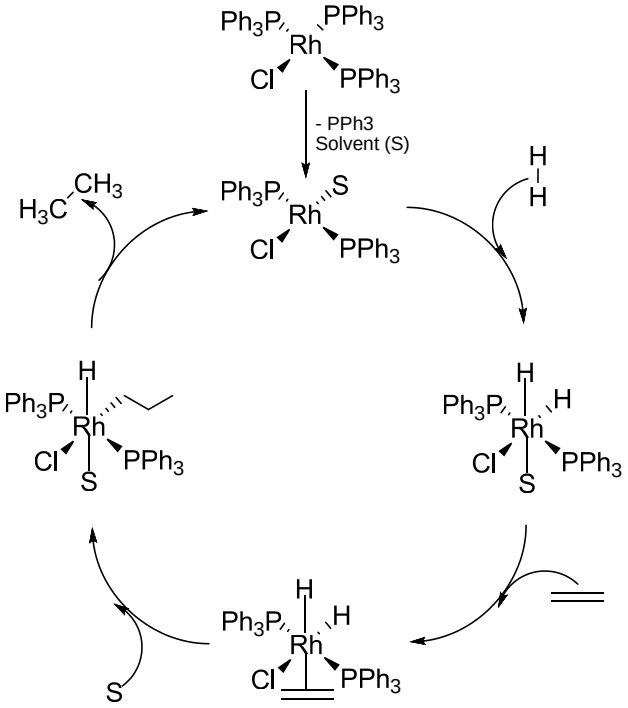


Figure 1 Wilkinson's hydrogenation cycle (mechanism investigated by Halpern and co-workers)[5,26-27]

In 1968, Knowles replaced the triphenylphosphane of Wilkinson's catalyst with a chiral equivalent, methyl(propyl)(phenyl)phosphane, and used this ligand for the rhodium-catalysed hydrogenation of a prochiral olefin, α-phenylacrylic acid. The 15% enantiomeric excess (e.e.) reported, although modest, confirmed the possibility of carrying out enantioselective catalytic reactions by using a metal (Rh) placed within a chiral environment.[28]

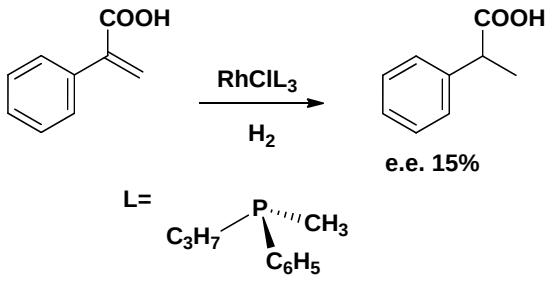


Figure 2 The first catalytic hydrogenation using a chiral phosphane ligand[28]

In the same year, Noyori and Nozaki discovered another example of asymmetric catalysis, while trying to elucidate the mechanism of carbene reactions. Chiral carbenoid species were obtained via decomposition of diazo compounds catalysed by a chiral Cu(II) metal complex. One example is the decomposition of ethyl diazoacetate in styrene in presence of bis[N-(R)-α-phenethylsalicylaldiminato] copper(II) or bis[N-(L)-α-phenethylsalicylaldiminato] copper(II). An optically active mixture of *cis* and *trans* cyclopropanecarboxylates is formed, with an enantiomeric excess of 6% and 10% respectively (Figure 3).[29]

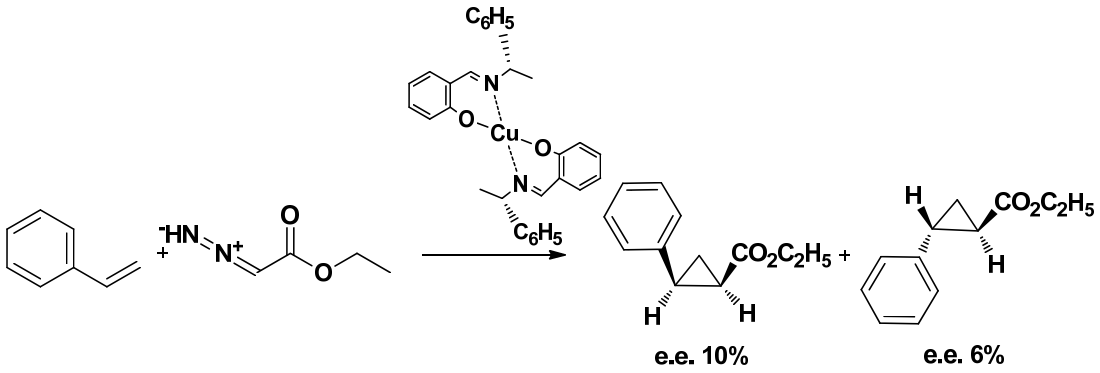


Figure 3 Asymmetric cyclopropanation of styrene using a chiral Cu(II) metal complex as catalyst[29]

Distinctive from Noyori's catalyst, where the chiral centre is a carbon atom, the Knowles catalyst contains a phosphorus stereogenic atom. When phosphanes having two identical chiral alkyl side chains were tested, the enantioselectivity in hydrogenation reactions was very low. Therefore it was hypothesized at this stage that, to induce high selectivities, the chirality of the ligands should be located directly on phosphorus.[22]

The substantial demand for the chemical L-dopa, a drug used in the treatment of Parkinson disease, provided strong impetus for the further development of chiral

phosphanes. The classical synthesis of this compound involved hydrogenation of a prochiral enamide intermediate to protected DL-dopa, the racemic intermediate being further resolved and deprotected to obtain L-dopa (Figure 4).[22,30]

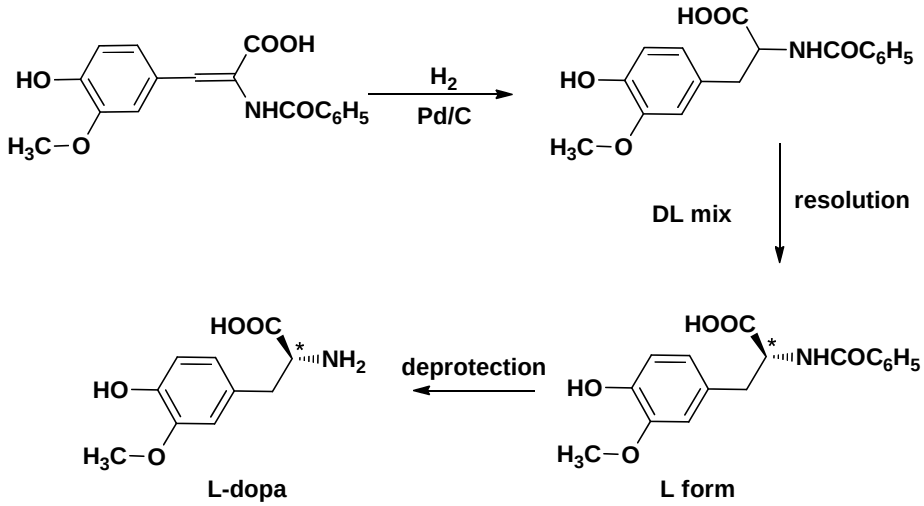


Figure 4 The former pathway of L-dopa synthesis (Hoffman-LaRoche process)

The next target was to find the right structure of a chiral phosphane able to catalyse the hydrogenation of the enamine intermediate with very high enantioselectivity.

The enantioselective Rh-catalysed hydrogenation of a phenylalanine intermediate was used as test reaction to study the efficiency and the structure-activity relationship of the tested chiral phosphanes (Figure 5). The chirality induced by the initial ligand, methyl(propyl)(phenyl) phosphane for this reaction was only 28%, but this provided a promising starting point for a structure-activity study.

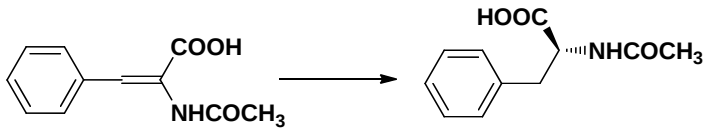


Figure 5 Test reaction used to study the structure-activity relationship of chiral phosphanes developed for the synthesis of L-dopa[22]

Initially the alkyl groups on phosphorus were varied by replacing the *n*-propyl with more hindered isopropyl or cyclohexyl groups, but this did not yield a significant improvement in enantioselectivity. The first enhancement resulted from the introduction of the *o*-anisyl group, the enantioselectivity induced by PAMP - methyl(phenyl)(*o*-anisyl)phosphane being increased up to 58%. Next, CAMP - methyl(cyclohexyl)(*o*-anisyl)phosphane gave up to 88% e.e. for the test reaction, but also for the

hydrogenation of the L-dopa precursor (Figure 6). Therefore its rhodium complex $[Rh(CAMP)_2(COD)]^+BF_4^-$ was used at the beginning in the Monsanto's industrial synthesis of L-dopa starting from vanillin.[31] Knowles used an $RhClL_3$ metal complex for the first hydrogenation reaction studied. The replacement of chlorine by cyclooctadiene (COD) in Monsanto's process was justified by the improved solubility and stability in air of the final metal complex.[22]

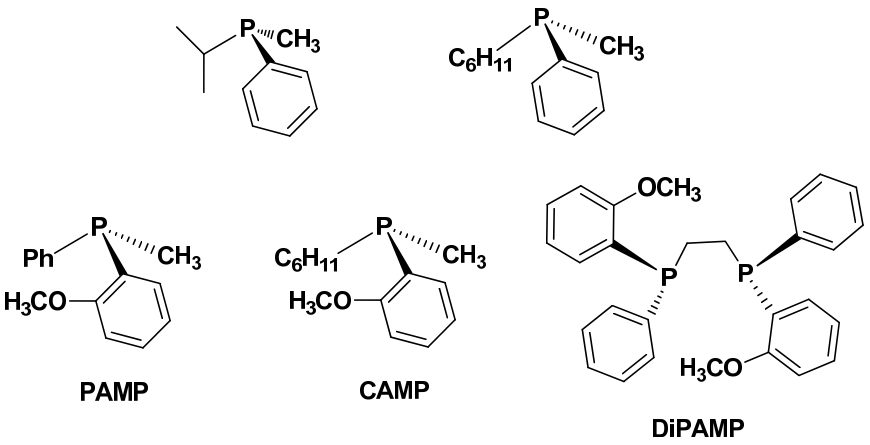


Figure 6 Series of chiral phosphane ligands synthesized by Knowles and co-workers by varying the substituents on the phosphorus atom[22]

Initially it was hypothesized that the phosphane catalysts having the chirality on the phosphorus are definitely superior. However, in 1972 Kagan introduced a diphosphane ligand, which he called (*R,R*)-DIOP, with similar catalytic performances as CAMP, but with the chirality provided by the carbon backbone (Figure 7).[32] This discovery introduced a new possibility to develop a whole series of diphosphane ligands with the chirality located at the carbon backbone.

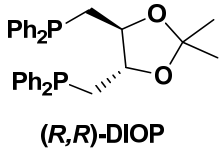


Figure 7 (*R,R*)-DIOP – phosphane ligand having the chirality induced by the carbon backbone[32]

Shortly afterwards, Knowles and co-workers reported the synthesis of a chelating diphosphane ligand, called DIPAMP (Figure 6), with the chirality on the phosphorus atom. This ligand replaced CAMP in the Monsanto L-dopa process where it displayed enantioselectivity as high as 95% e.e.[31] DIPAMP not only showed higher

enantioselectivity, but it was easier to be synthesized than CAMP and it is a crystalline, air-stable solid.[22]

The area of chiral phosphane ligands has expanded significantly after that and a large number of versatile ligands are now readily available from commercial suppliers. For example BINAP (2,2'-bis(diphenylphosphanyl)-1,1'-binaphtyl), a ligand with axial chirality (Figure 8) introduced in the 1980s by Noyori and co-workers, was used initially in combination with rhodium for catalytic hydrogenation of olefins,[33] but it has been used on a multitude of other reactions ever since.

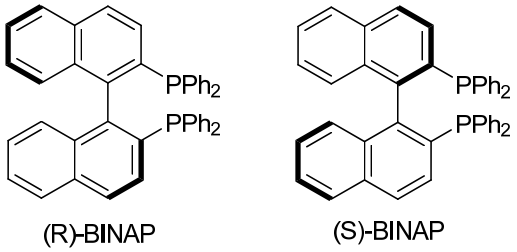


Figure 8 BINAP - ligand presenting axial chirality

During the last two decades the area of asymmetric transition metal catalysis has been developed significantly, becoming a common type of catalysis and allowing access to a variety of compounds. However, a moderate number of asymmetric catalytic reactions are used on an industrial scale today (e.g. hydrogenation, epoxidation, dihydroxylation, isomerization, epoxide opening and addition reactions).[34] This is due to several problems associated with the production of chiral compounds on industrial scale. One issue is the enantiopurity of the ligand, and consequently the enantioselectivity of a chiral catalyst, which must be higher than 99% for pharmaceuticals and more than 80% for agrochemicals and food products. The high price of chiral ligands and metal precursors is also a limiting factor for their application on a large scale.

Generally, synthesis and optimization of a new catalyst involves a cyclic process which includes three steps: (i) conception of a new catalyst, (ii) synthesis of the new catalyst and (iii) testing its catalytic properties. The catalyst optimization consists of analysis of the results obtained at the end of each cycle, followed by selection of the catalyst with improved properties and the enhancement of its conception and synthesis for further improvement. To obtain a highly active and selective catalyst, a very large number of iterations may be required, making it a time-consuming process. These optimization

procedures have generally been applied to individual catalysts, known as the traditional "one-catalyst-at-a-time" approach.[10]

In order to reduce the time and the costs of catalyst optimization, methods like combinatorial chemistry and parallel synthesis can be applied in catalyst development. These methods are used for simultaneous synthesis of different but structurally related molecules. Combinatorial chemistry is a collection of methods that allow the simultaneous chemical synthesis of a large number of compounds using a diversity of starting materials. Multiple parallel synthesis is a related method used to prepare a selected subset of molecules. The set of compounds prepared by multiple parallel synthesis is more focused and less diverse than the large library resulting from combinatorial techniques.[35]

Using the combinatorial chemistry and parallel synthesis methods, the synthesis and optimization cycles can be applied simultaneously to a library of potential asymmetric catalysts, rather than to each catalyst individually.[10] In order to be applied, three main conditions must be achieved:[21,36]

- existence of a source of diversity – represented by a library of ligands, of metal precursors or a variety of reaction conditions;
- availability of methods to synthesize the catalysts fast and with a high purity;
- availability of methods for a high-throughput screening of the catalysts performance.

The first condition, existence of a source of diversity, is not a limiting factor anymore. A large variety of metal precursors and ligands is currently commercially available.[37] However, even if methods for high-throughput screening of catalysts performance are available nowadays (NMR, GC, HPLC or less often CE), this aspect still represents a limiting factor not only presently, but probably also in the near future.[10]

Despite its significant progress, the area of asymmetric transition metal catalysis still has important limitations. The energy difference involved in transition states leading to both enantiomers is small. For that reason it is difficult to model or to calculate which enantiomer of a desired product will be obtained for a catalytic reaction.[38] In addition, several subtle experimental parameters (solvent, counter ion, added salts etc), included under the general term of second coordination sphere,[39-41] exert an important influence over the outcome of a reaction in homogeneous catalysis. However, the control over the second coordination sphere is limited for the common enantiopure ligands because of

their relative small size. They are able to exert a considerable control only over the hemisphere located around the catalytically active metal (known as the chiral ligand hemisphere).[39,41] Therefore, to induce a high enantioselectivity a catalyst should display a higher control over the largest part of the second coordination sphere. In this way the catalyst "shields" the active site and stabilizes the subtle experimental parameters which influence the enantioselectivity of the reaction.[41] However, even for catalysts displaying high control over the second coordination sphere it still remains hard to predict the enantioselectivity in a catalytic reaction.[41]

I.2.b. Enantioselective enzymatic catalysis

An alternative to enantioselective organometallic catalysis for the synthesis of enantiomerically pure compounds is enzymatic catalysis. Enzymes are proteins that possess catalytic activity and typically exhibit excellent regio-, chemo- and enantioselectivity. Like all natural proteins, they are assembled from L-amino acids and therefore they are chiral catalysts. Structurally an enzyme contains an active site which is directly involved in selectively binding the substrate, creating the chiral environment and carrying out the catalytic reactions. Frequently enzymes require cofactors (non-protein organic or inorganic compounds) for catalytic activity. Cofactors can be divided in two large groups: prosthetic groups and coenzymes, although there is not a sharp division between these two.[42-43] Generally, prosthetic groups form a permanent part of an enzyme structure, being tightly bound to the apoenzyme, while coenzymes are loosely bound organic molecules which carry chemical groups between enzymes and can be released during the catalytic cycle. Two examples of compounds acting as prosthetic group for some enzymes and coenzyme for others are thiamine diphosphate (ThDP) and nicotinamide adenine dinucleotide (NAD^+). ThDP is tightly bound in transketolase or pyruvate decarboxylase, and less tightly bound in pyruvate dehydrogenase. NAD^+ also can be tightly bound in some enzymes, while it is loosely bound in others.[43]

The catalytic mechanism of enzymes is explained by Pauling's principle and consists of lowering the activation energy of reactions by confining substrates in binding sites that stabilize the transition state.[44-46]

The use of biocatalysts provides several advantages. Enzymes display high selectivity (chemo, regio and enantioselectivity) and high rates, resulting in a high quality product. They supply a chiral first coordination sphere and a high control over the second coordination sphere, being structurally able to shield the reactants and thus to prevent possible side reactions.[41] The molecular biology techniques could be applied in order to obtain enzymes which satisfy supplementary requirements (e.g. broader reaction scope, higher thermostability). Enzymes can be made readily available by expression in adequate microorganisms, using recombinant DNA technology. Usually they act under mild conditions and are used in aqueous media (but not always), therefore some inherent problems of metal-catalysis (e.g. requirement for special equipments, undesired side-reactions) are minimized. They are used in environmentally benign synthesis and are biodegradable. The enzymatic reactions require few steps and in general there is no need for protection of functional groups.

However, to fully exploit the advantages provided by enzymatic catalysis, it is necessary to find solutions to overcome limitations of this area. One limitation is represented by the small substrate scope exhibited in general by enzymes. Instability under severe process conditions, substrate or product inhibition, high cost of production, limited availability and lifetime durability, problems with product isolation due to possible product contamination with catalyst are problems which can also be encountered in biocatalysis. Enzymes have low volumetric productivities in general and this requires high dilutions and complicated down-stream processing.

There are two ways to obtain enantiopure compounds using enzymes. One option is biocatalysed asymmetric synthesis, in which an achiral substrate is converted to a chiral product via enantioselective enzymatic conversion. The second option is the kinetic resolution (KR) of a racemic mixture. In this process, one enantiomer from a racemic mixture is converted, via enzymatic catalysis, with a higher reaction rate than the other enantiomer. This means that the maximum yield is limited to 50%. A yield higher than 50% means that some of the other enantiomer has also reacted. However, Backvall and co-workers demonstrated that full conversion of a racemic substrate to an enantiomerically pure product can be achieved by combining enzyme catalysis for the kinetic resolution with transition metal catalysis for racemization. By using a racemization catalyst, the preferentially converted substrate enantiomer is replenished during the course of the reaction, a process named dynamic kinetic resolution (DKR).[47-

[49] This approach has been applied by several groups, for example for the chemoenzymatic DKR of alcohols and amines.[50-60]

Currently enzymatic catalysis has become a valuable method for the synthesis of chiral compounds not only in research laboratories, but also in pharmaceutical or biotechnological industry. Commonly used enzymes are lipases for reactions involving ester formation or hydrolysis,[61-64] nitrilases for hydrolysis of cyanide group to acids[65] and dehydrogenases for oxydoreduction of alcohols and ketones.[66] The use of dehydrogenases is limited sometimes by the cofactor (NAD(P)H or FADH) regeneration.[67-69]

In other applications, like the enzymatic hydrogenation of enones and nitrolefins, it was possible to use the whole microorganism which expresses the enzyme as catalyst instead of the pure enzyme.[70-71] This approach is more advantageous as the entire process of enzyme purification is avoided. Another alternative is the combination of reactions catalysed by whole microorganisms with reactions catalysed by enzymes. An example, applied on industrial scale, is the two-step synthesis of vanillin from glucose, by coupling a reaction catalysed by a microorganism (*E. coli*) with an enzymatic reaction (Figure 9).[72]

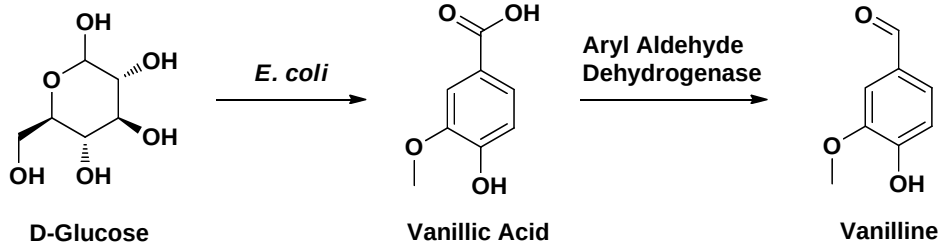


Figure 9 Synthesis of vanillin from glucose using *E. coli* (whole cells) and a pure enzyme[72]

Although the conventional use of enzymes at mild conditions in aqueous media brings several advantages, it also presents limitations. The main disadvantages of using aqueous media for synthesis reactions are the requirement of water soluble substrates, occurrence of side reactions induced by water, degradation of common organic reagents in aqueous media and the difficulties encountered sometimes in product recovery.[11] Therefore a significant number of studies have been devoted to enzymatic reactions in non-aqueous systems like organic solvents,[11,73] ionic liquids, supercritical fluids,[74] gases,[75-76] eutectic mixtures[77] or liquid crystals.[11] Several enzymes have been used as biocatalysts in biphasic systems like ionic liquids/supercritical fluids,[78] ionic

liquids/water,[79] aqueous/oil phase[80] or aqueous/organic phase.[73] Enzymes have been also used for conversions in solvent free systems.[81-82]

The long-standing dogma that enzymes have limited applicability in organic solvents was initially justified by the conventional theory that they are denaturated in such solvents.[83] This theory holds when enzymes are investigated in aqueous-organic mixtures, where water acts as a molecular lubricant.[84-85] In such conditions, enzymes have both the tendency to denature and also the conformational flexibility to do so. However, in the absence of water enzymes become very rigid. As a result even if their tendency to denature is higher in dry organic solvents, they don't have the pliability for denaturation.[86]

In general enzymes exhibit lower catalytic activity when they are used in organic solvents.[87] In hydrophobic solvents the enzymatic activity is usually higher than in hydrophilic ones. This is because the hydrophilic solvents are capable of stripping the tightly bound water, which is essential for catalytic activity.[88] However, in organic solvents enzymes can develop new properties like the ability to catalyse reactions which are impossible or only marginal in water, solvent dependent selectivity and improved thermostability. For example lipases and esterases are incapable of hydrolysing amides but they are able to catalyse direct asymmetric acylation of racemic amines in organic solvents,[89-90] hence the industrial applicability for the preparation of chiral amines and amides.[91] Lipases and esterases have also been used for the synthesis of enantiomerically pure alcohols, carboxylic acids or esters via enzymatic resolution of their racemates in organic solvents.[92-93] Regarding enzymes selectivity in organic solvents, in particular substrate,[94-97] enantio-[98] and prochiral[99] selectivity, was proved to change remarkably when switching from one solvent to another.[100-101] Enzymes showed also improved thermostability in organic solvents. Three tested enzymes: porcine pancreatic lipase,[102] ribonuclease[103] and α-chymotrypsin[104] exhibited half-lives of several hours in anhydrous solvents at 100ºC, while their deactivation in water requires seconds at the same temperature.

Although the enzymatic catalysis in organic solvents has overcome some problems encountered in the classical aqueous biocatalysis, the work in this area has involved so far only a number of hydrolytic enzymes.[93,105]

I.2.b.1 Directed evolution of natural enzymes

The approach of natural enzyme optimization through evolution was applied in biochemistry as the directed evolution concept, to optimize and diversify natural enzymes which are subsequently used as catalysts. The concept was first demonstrated by Reetz in 1997 and several successful studies have been reported since then by both academic and industrial groups.[106-109] It is based on the appropriate combination of random gene mutagenesis, expression of the enzyme mutants and high or medium throughput screening for improvement of the desired properties[10,110-115] (e.g. activity, selectivity,[107] organic solvents tolerance or improved thermostability).[106] Importantly, the concept can be applied to enzymes for which no structural information is available. Together, these three steps represent a round of directed evolution. At the end of each round the gene for the enzyme variant (hit) revealing the best properties is used as a template for a new round of mutagenesis/expression/screening, a process which exerts evolutionary pressure until the desired level for the improved properties is reached.[109,112,115-118]

PCR (polymerase chain reaction) is a technique used to copy and amplify the complementary strands of a DNA molecule without using living organisms. Several techniques based on PCR have been used to create libraries of mutants. The most often applied technique for mutagenesis is error-prone PCR (epPCR),[119-120] in addition to other methods like saturation mutagenesis[121-122] and DNA shuffling.[123-124]

Error-prone PCR is based on introduction of random mutations, without being able to control the location and the nature of the mutations. This is usually realised by using error prone polymerases in the PCR process,[125] but can also be realized by varying the concentration of magnesium during the reaction, by using manganese salts or by using an excess of certain oligonucleotides.[126] Another possibility to introduce random mutations without controlling their location and nature is to transform and grow the plasmid containing the enzyme gene in “mutator” bacteria (bacteria with elevated mutation frequencies).[127-131]

Saturation mutagenesis is a molecular biology method which restricts randomization to predetermined sites in the enzyme. The result is represented by focused libraries, created by replacing amino acid residues one by one, via mutagenesis, by the other 19 natural amino acids.[128,132-133] The iterative saturation mutagenesis (ISM) method is a more systematic approach for generation of focused libraries and it is a symbiosis of rational design and combinatorial saturation mutagenesis. Using structural information

of an enzyme, two or three amino acids are selectively chosen for saturation mutagenesis, based on rational considerations.[134-135]

Two forms of ISM were developed, Combinatorial Active-site Saturation Test (CAST), for controlling substrate scope and/or enantioselectivity,[107-108,133] and the B-factor saturation test (B-FIT) for increasing thermostability.[106,135] Using these techniques, two or three amino acids "hot spots" whose side chains are located next to each other and close to the active site of the enzyme, are identified and randomized resulting in libraries of mutants. The process is performed all around the active site of the enzyme. The scheme can be applied in repetitive cycles, starting with mutagenesis in one selected library, selecting the improved mutants and continuing successively to several other libraries derived from improved mutants of the previous library. A powerful tool to construct more focused enzyme libraries that requires considerably less screening effort is the choice of the optimal codon degeneracy when modelling saturation mutagenesis experiments. In this way, the size of focused libraries generated by targeted mutagenesis is decreased and the quality, in terms of frequency of favourable variants and the degree of catalyst improvement, is increased.[136]

In contrast with PCR, DNA shuffling is PCR without synthetic primers. It begins with fragmentation of families of related genes with enzymes. The gene fragments then are heated up to separate them into single-stranded templates. Some of these fragments will bind to other fragments that share complementary DNA regions, which in some cases are from other family members. At the end, new full length, shuffled genes are constructed and multiplied by PCR.[123]

Even if important properties of enzymes (stability, selectivity, organic solvent tolerance, substrate scope) could be engineered by applying the directed/designed evolution tools, the overall process is time consuming, requires sophisticated equipment for high-throughput screening and still has a very limited applicability.

I.2.c. Enantioselective hybrid catalysis

Biocatalysis and homogeneous asymmetric catalysis are both employed for the production of enantiomerically pure compounds and these two areas of catalysis are complementary in many essential aspects (Table 1).

Table 1 Comparison of typical characteristics of enzymatic and homogeneous catalysis[38,137]

	Enzymatic catalysis	**Homogeneous catalysis**
Reaction repertoire	Small	Large
Turnover number	Large	Modest
Optimization	Genetic	Chemical
Second coordination sphere	Well-defined	Ill-defined
Substrate scope	Limited	Large
Enantiomers	Single (usually)	Both
Reaction medium	Aqueous (mostly)	Organic (mostly)
Reaction conditions	Mild	Harsh
Recycling	Immobilization protocols	Rare

Hybrid catalysis is a relatively new area of asymmetric catalysis, aiming to combine the most desirable properties of both parts, enzymatic and homogeneous catalysis. Artificial metalloenzymes, as defined by Ward (2007), are catalysts resulting from the introduction of an organometallic catalyst into a macromolecular host such as a protein, DNA or an antibody, which provide a well-defined second coordination sphere and thus induce the selectivity of the reaction.[41]

For a hybrid catalyst it is possible to optimize both the activity and the selectivity using a chemogenetic approach.[138] The activity can be optimized via a combinatorial approach or parallel synthesis for the organometallic catalyst. The selectivity can be optimized via molecular biology techniques for the host biomolecule.

The methods used to introduce a catalytic moiety into a host protein should be optimized for the following aspects:

– *Selectivity*. All biomacromolecules possess a multitude of functional groups, therefore the synthesis of well-defined systems by selective modification of a specific site is not a trivial task.

– *Compatibility with the host protein.* The modified protein should be structurally defined, therefore introduction of a ligand should not result in unfolding or denaturation of the protein. The proper selection criteria should consider proteins presenting a robust scaffold and ligands which are also active when they are located within the protein environment.

– *Compatibility with coupling conditions and catalysis*. Organometallic and organic catalysts often require reaction conditions which are not compatible with proteins (high temperatures, high pressure and presence of organic solvents).
A strategy is more beneficial when it can be applied for more than one protein and more than one metal- or organocatalyst.

The strategies used to incorporate an artificial catalytic moiety within a biomacromolecule host are the following:

- *Supramolecular anchoring.* The catalytic moiety acts like an anchor or is coupled to an anchor which has supramolecular affinity for the host biomolecule. To ensure the adequate localization of the ligand inside the binding pocket of the host molecule, a very strong non-covalent guest-host couple should be selected. This condition comes as a restriction of the method because of the limited number of suitable guest-host pairs available.
- *Dative anchoring.* The artificial metalloenzyme is formed via coordination complex formation of a metal ion of a catalytic moiety with amino acid residues of a protein, which act as Lewis bases.
- *Covalent anchoring* of a catalytic moiety to a single, accessible functional group of the host biomolecule. Typically the thiol group of a cysteine or the hydroxyl group of a serine residue is used for covalent modification when proteins are used as host biomolecule. However, selective incorporation of unnatural amino acids into proteins, strategy developed mainly during the last ten years, allows further enhancement for the covalent anchoring method. A variety of unnatural amino acids were incorporated into recombinant proteins using specifically engineered bacterial strains.[139-143] Those non-natural amino acids possessing distinctive chemical reactivity create new possibilities for bio-orthogonal conjugation of proteins with organic ligands. For example an azide containing amino acid incorporated in a target protein has been selectively modified via Staudinger ligation with triarylphosphane reagents. Importantly, this was performed using a cell lysate mixture, in presence of all the other cellular proteins.[144]
- *Cotranslational incorporation* of multidentate metal-binding amino acids into proteins. Schultz and co-workers reported the cotranslational incorporation of (2,2'-bipyridin-5-yl)alanine into proteins. This approach opens additional possibilities to the classical methods (covalent, dative and supramolecular anchoring) for artificial metalloenzymes engineering.[145]

I.3 Supramolecular anchoring

Biotin-(strept)avidin system

The most successful systems used for synthesis of artificial metalloenzymes based on supramolecular anchoring of the catalyst are represented by biotin-avidin and biotin-streptavidin couples. Both proteins, avidin and streptavidin, bind the vitamin biotin with very high affinity and specificity. The affinity of streptavidin is lower ($K_a \approx 2.5x10^{13}$ M^{-1}) than that of avidin ($K_a \approx 1.5x10^{15}$ M^{-1}), nevertheless these are the highest affinities known in nature for a noncovalent interaction.[146-147] Streptavidin has a deeper binding pocket and it is a better chiral inducer than avidin,[148] as host protein, and is very stable in presence of denaturating agents.[149-151]

The biotin-avidin system was first used in 1978 by Wilson and Whitesides for creation of artificial metalloenzymes. Their hybrid catalyst, formed by an achiral biotinylated rhodium-diphosphane complex non-covalently coupled to avidin, was tested for the hydrogenation of acetamidoacrylic acid. The 41% e.e. in favour of (S)-acetamidoalanine was induced by the protein chirality.[152]

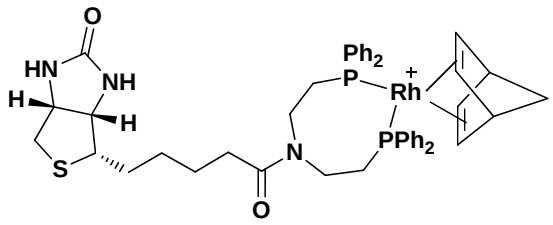


Figure 10 Whitesides diphosphane-rhodium biotin complex[152]

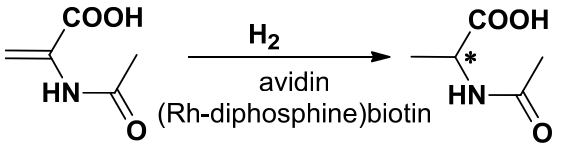


Figure 11 Enantioselective hydrogenation of acetamidoacrylic acid using a diphosphane-rhodium biotin complex anchored to avidin[152]

Wilson and Whitesides achievement was strongly improved by Ward and co-workers. They produced several efficient artificial metalloenzymes using initially avidin and later streptavidin as host protein, in combination with biotinylated catalysts.[148,153-165]

As starting point the Whitesides rhodium-diphosphane biotin complex anchored into the wild-type streptavidin was used for the hydrogenation of α-acetamidoacrylic and α-

acetamidocinnamic acid. Two steps of chemogenetic engineering were applied. Eighteen biotinylated rhodium-diphosphane complexes were screened throughout the chemical optimization step and twenty mutants of streptavidin were tested during the genetic optimization step. As a result, the enantioselectivity was improved to 94% e.e. for the (R)-α-acetamidophenylalanine*)* and could be inverted to 88% e.e. for the (S)-α-acetamidophenylalanine.[155-156]

Additional enhancements of the artificial metalloenzyme characteristics (selectivity, reaction rates, solvent tolerance) were realized by applying the chemogenetic approach[138] again, in an extensive manner. The chemical optimization was done by introducing chiral amino acid spacers (proline or phenylalanine) between the biotin anchor and the diphosphane moiety and by varying the components of the metal complex (metal, diphosphane ligand) and the chiral spacer. The genetic optimization was done by introducing point mutations combined with saturation mutagenesis in the protein scaffold in certain selected spots (Figure 12).

In this way the enantioselectivity was increased to 96% for the (R)-α-acetamidoalanine and 95% for the (S)-α-acetamidoalanine.[148,157] In addition, the hybrid catalyst containing the proline spacer showed enhanced stability towards organic solvents (DMSO and ethyl-acetate, the latter under biphasic conditions).[157-159]

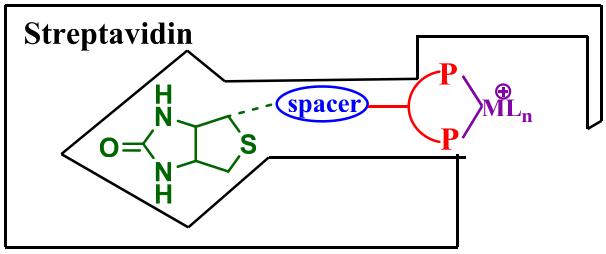


Figure 12 Schematic representation of the artificial metalloenzymes developed by Ward and co-workers in enantioselective hydrogenation reactions[148,155]

Ward and co-workers were the first to genetically modify an artificial metalloenzyme. For the optimization of the hybrid catalyst they used a combination of rational design and combinatorial screening, an approach named *designed evolution.*[137,166] The rational design involved rational choices for the protein scaffold and for the mutation sites based on structural information of the protein-biotin complex, and also for the chemical components. The combinatorial screening included variation of the chemical components and saturation mutagenesis for the protein.

Reetz applied the concept of *directed evolution* of an artificial metalloenzyme,[167-168] a process fundamentally different from Ward's designed evolution, for the optimization of the hybrid catalysts based on biotin-streptavidin in hydrogenation of α-acetamido-acrylic acid-methyl ester. By applying three cycles of random mutagenesis/expression/screening, using Whitesides' system as the starting point, hybrids with improved and/or inverted selectivity were obtained. Starting from 23% e.e. in favour of the (R) product before optimization, the enantioselectivity was enhanced up to 65% e.e. for the (R) product and up to 7% e.e. for the (S) product).[12,134] However, the concept proved to be more difficult to apply for the hybrid catalysts than for the natural enzymes. This was mainly due to problems associated with the protein expression system and the air sensitivity of the metal complexes.

Following the application of the biotin-streptavidin based artificial metalloenzymes in hydrogenation reactions, Thomas Ward's group has focused on more challenging reactions like transfer hydrogenation of ketone derivatives,[153,160-161] oxidation of secondary alcohols,[162] allylic alkylation[163] and sulfoxidation.[164-165]

Hybrid catalysts containing achiral biotinylated d^6 piano-stool ruthenium complexes [η^6-(arene)Ru(biotin-ligand)Cl] bound to streptavidin were employed for enantioselective transfer hydrogenation of ketones. After a first chemogenetic optimization step, further improvements were realized by applying an additional chemogenetic optimization based on the X-ray structure of a hybrid catalyst variant.[161,169] The nature of the capping arene (benzene or *p*-cymene) was found to have a substantial influence on the reaction enantioselectivity.[153,160]

A similar approach based on biotinylated d^6 piano-stool complexes embedded in strept(avidin) was used for the reverse reaction, oxidation of racemic secondary alcohols using *tert*butylhydroperoxide (TBHP). In this case not only Ru-complexes were tested, but also Ir and Rh-complexes, combined with both avidin and streptavidin as host protein. Avidin and its mutants showed higher yields than streptavidin.[162]

Two types of streptavidin based hybrid catalysts were used by Ward and co-workers for enantioselective sulfoxidation reactions: biotinylated manganese-salen complexes incorporated in streptavidin as well as vanadyl-streptavidin.

Four achiral biotinylated manganese-salen complexes incorporated in streptavidin were used as hybrid enantioselective catalysts for the aqueous sulfoxidation of thioanisole with hydrogen peroxide, yielding moderate conversion (up to 56%) and low enantioselectivities (maximum of 13% e.e.).[165]

Although being optimized by nature to bind biotin with high affinity by combining both hydrogen bonding and hydrophobic interactions, streptavidin can also bind a variety of other ligands: (2-(4'-hydroxyphenyl-azo)benzoate (HABA), 1-anilino-8-naphthalene sulfonate (ANS), different oligopeptides).[146,150,164,170-171] Therefore, Ward and co-workers considered the possibility of binding a catalytically active, small polar coordination compound, the vanadyl ion, into the biotin-binding pocket of streptavidin. Electron paramagnetic resonance (EPR) spectroscopy, docking studies, chemical (addition of biotin) and genetic modification of the host protein suggested that interactions between the active precatalyst $[VO(H_2O)_5]^{2+}$ and the biotin binding pocket are only via second coordination sphere. The resulting artificial metalloenzyme was tested in enantioselective oxidation of prochiral sulfides using TBHP as oxidizing agent. Good enantioselectivities were observed for both dialkyl and alkyl-aryl substrates (up to 93% e.e. for the sulfoxidation of methyl-2-naphtylsulfide).[164]

Serum albumins

Serum albumins have also attracted considerable interest in the area of artificial enzymes based on supramolecular anchoring. Functioning as transport proteins, they have a significant affinity for a variety of hydrophobic compounds like fatty acids, steroids, thyroxine, porphyrins etc.[172] Remarkably, serum albumins by themselves, and in particular BSA, have been reported to induce moderate to good enantioselectivities in several reactions, including sulfoxidation,[173-177] epoxidation,[178-180] reduction,[181-182] and Diels-Alder cycloaddition reactions.[183]

The first artificial metalloenzyme created via supramolecular incorporation of a metal complex into serum albumin was reported by Watanabe in 2003. Two types of artificial metalloenzymes containing a [Cr^{III}-5,5'-tBu_2-salophen] complex, one based on BSA and the second on a myoglobin mutant, were tested for the enantiomeric sulfoxidation of thioanisole. The hybrid catalyst based on BSA contained the Cr^{III} complex bound non-specifically, while in the myoglobin mutant it was bound via hydrogen bonding and histidine (His 93) coordination to Cr^{III}. The enantioselectivity and rate achieved with [Cr^{III}-5,5'-tBu_2-salophen]-BSA were very low compared with those obtained using the same Cr^{III}-5,5'-tBu_2-salophen complex incorporated within the myoglobin mutant. This was assumed to be due to the non-specific binding of the metal complex to BSA and to the absence of histidine coordination. No further characterization of the Cr^{III}-5,5'-tBu_2-salophen binding positions to BSA has been reported.[184]

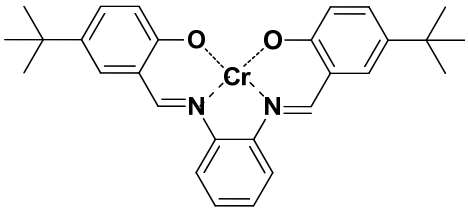


Figure 13 Cr^{III}-**5,5'-tBu_2**-salophen complex inserted into BSA[184]

In 2009 Ménage and co-workers characterized three hybrid catalysts based on Mn-salen complexes and HSA. As the hybrids crystallization was not successful, the localization of the Mn-salen complexes within the cavity of HSA was assigned based on spectroscopic characterization and on several competitive binding experiments.[185] The artificial metalloenzymes were tested in sulfoxidation of thioanisole by NaOCl. In comparison with HSA alone, the artificial enzymes displayed higher efficiency (conversion yield and rates) and chemoselectivity for sulfoxide over sulfone, but no enantioselectivity was reported. In contrast with the metal complex alone, the hybrid catalysts displayed inverted chemoselectivity. No oxidation was observed in presence of hydrogen peroxide as oxidant. Considering the observation that for the epoxidation reactions performed with hydrogen peroxide as oxidant, the reactivity of Mn-salen complexes is highly improved when imidazole is coordinated in an apical position,[186] this observation confirms that no histidine residue is involved in a dative interaction.[185]

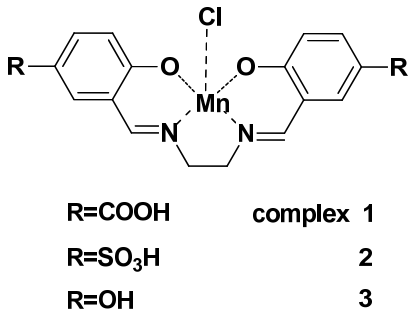


Figure 14 Mn-salen complexes coupled to HSA via supramolecular interactions[185]

Antibodies

Another strategy to create artificial enzymes is based on non-covalent introduction of catalytic activity into antibodies. This method exploits the capacity of the immune system to generate specific monoclonal antibodies for particular molecules of interest.

In 1988 Schultz and Shokat reported reduction of Safranine T using an antibody-flavin complex, a reaction that does not occur with free flavin. Antibodies were elicited

against both the oxidized and the reduced form of the flavin cofactor. However the oxidized flavin was bound with a $4x10^4$ higher affinity than the reduced form to one of the monoclonal antibodies. Based on the observation that enzymes could direct substrate binding by non-covalent interactions,[187] the idea was coined to incorporate substrate binding sites adjacent to flavins to allow stereocontrolled chemical reductions.[188]

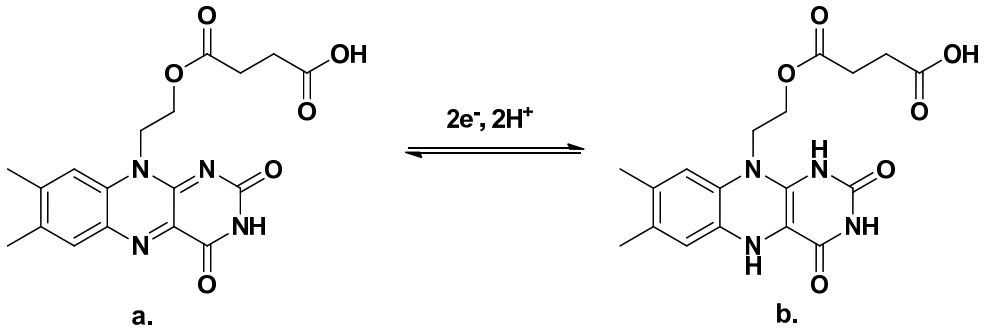


Figure 15 Flavin cofactor used to generate antibody-flavin complexes: (a)oxidized form; (b)reduced form[188]

In 1999 Keinan and Nimri were the first to report a catalytic antibody-porphyrin assembly capable of catalyzing oxygenation reactions. Notably, this system exhibited typical enzyme characteristics: oxidant and substrate selectivity, enantioselectivity and Michaelis –Menten saturation kinetics. A structural analogue of the transition state of sulfoxidation reactions was used as hapten for the production of monoclonal antibodies. The highest enantioselectivity reported was 43% in favour of the (S)-sulfoxide enantiomer, achieved for the sulfoxidation of thioanisole with iodosyl benzene (PhIO) in presence of antibody SN37.4 and a Ru(II)porphyrin cofactor (Figure 16).[189]

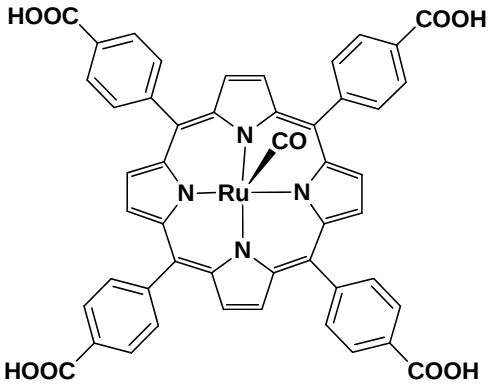


Figure 16 Rutenium (II) porphyrin cofactor coupled to antibody SN73.4 and used for enantioselective sulfoxidation of thioanisole[189]

Several other hemoabzymes (catalytic antibodies with a metalloporphyrin cofactor) have been reported,[190] exhibiting either peroxidase activity or monooxygenase activity. However, all systems showed a relative low catalytic efficiency compared to natural hemoenzymes. A possible explanation could be the fact that the selected antibodies lack an amino acid residue like histidine or cysteine, which can coordinate axially to iron and thereby enhances its redox potential. To address this problem the heme octapeptide microperoxidase 8 (MP8), obtained by hydrolytic digestion of horse heart cytochrome c, was used as hapten. MP8 contains the heme prosthetic group and also the amino acid residues 14-21 of the cytochrome. The imidazole side chain of histidine 18 plays a critical role by acting as a fifth axial ligand for iron. MP8 binding to amino acids of the binding site occurs via four carboxylate substituents, two from propionate side chains of the heme and two from the C-terminal glutamic acid Glu 21 of the octapeptide (Figure 17).[190]

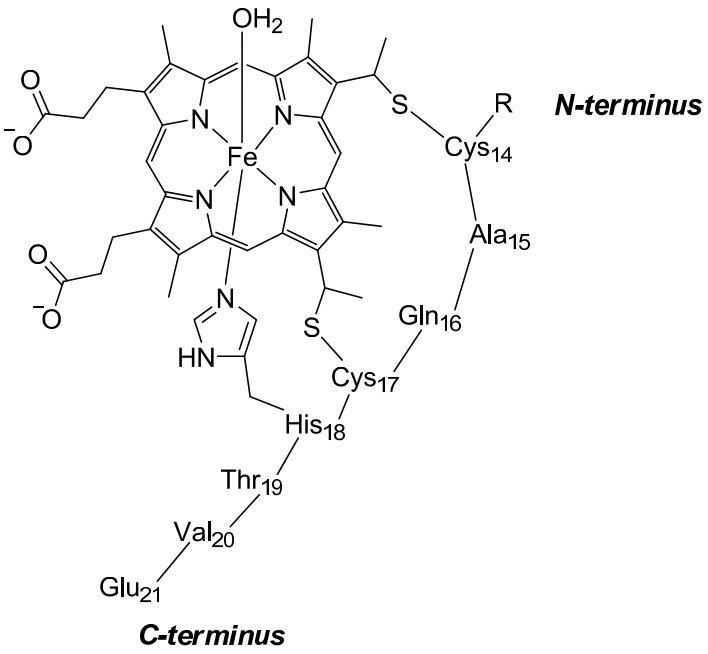


Figure 17 Structure of microperoxidase 8

Using the catalytic system formed by the 3A3 monoclonal antibody and microperoxidase 8 (3A3-MP8) for sulfoxidation of thioanisole using hydrogen peroxide as oxidant, the enantiomeric excess was 45% in favour of the (R)-sulfoxide enantiomer. This enantioselectivity is not only slightly higher than the one reported by Keinan et al.

in 1999 for the same reaction, but is also inverted, as the (R) enantiomer is favoured now.[191]

In 2006 Harada and coworkers elicited monoclonal antibodies against an achiral Rh cyclooctadiene phosphane complex (Figure 18).

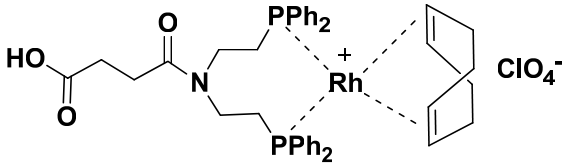


Figure 18 Structure of the achiral Rh complex used as hapten by Harada and co-workers[192]

Four catalytic systems based on four different monoclonal antibodies complexed with this hapten were tested in asymmetric hydrogenation of amino acid precursors. One combination (1G8-Rh complex) gave the enantiomer (S)-N-acetyl-L-alanine in 98% e.e. for the hydrogenation of 2-acetamidoacrylic acid. In absence of antibodies, the achiral Rh complex converted the substrate to racemic N-acetyl-alanine. The high enantioselectivity obtained indicates that the attack over the double bond of the substrate occurs only at one side of the alkene, the other side being blocked by the antibody.

The Rh catalyst-antibody 1G8 complex also displayed substrate specificity, yielding no conversion for larger substrates (2-acetamido-cinnamic acid and 2-acetamido-4-methylpentenoic acid). This is caused by the control exerted by the bound antibody, which does not allow accessibility of larger substrates to the Rh catalyst.[192]

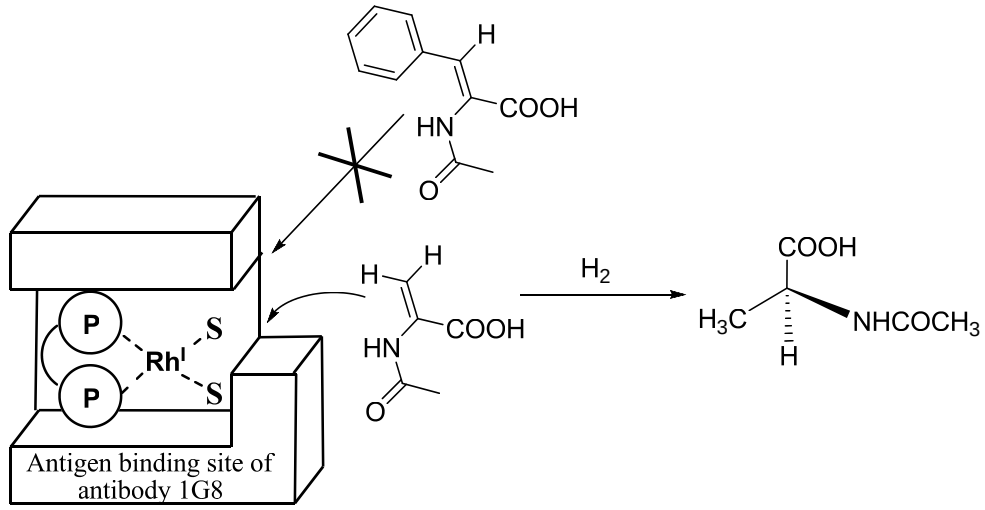


Figure 19 Schematic representation of the antibody 1G8 - Rh catalyst complex showing the substrate specificity (P - diphenyl phosphane ligand; S - solvent molecules)[192]

The previously reported antibody-based hybrid catalysts were based on cofactors existing also in natural enzymes (flavins, porphyrins). In contrast, Harada's catalytic system employs a transition metal complex which is not found in natural enzymes, but which has previously been applied in homogeneous catalysis for a variety of reactions such as hydrogenation, hydroformylation and isomerization and was also used for the synthesis of (strept)avidin based artificial metalloenzymes.[148,152,155-156,158]

I.4 Dative anchoring

The first artificial metalloenzyme created via dative anchoring strategy by Kaiser and co-workers (1976) was based on the concept of replacing the native metal bound to the active site of a metalloenzyme with a different one.[193] By applying this concept, the problems associated with synthesis of ligands and their couplings to proteins are avoided. However, the high potential for unspecific binding of metals to proteins, most probably to the side chains of cysteines, lysines and histidines, limits this approach.[194-196] For example, the Rh(II) complex [Rh_2(carboxylate)$_4$] binds weakly ($K \approx 10^3$) to serum albumin in a molar ratio of about 8 rhodium atoms per serum albumin.[197-198] Several years after the first studies about the direct linkage of Rh(I) and Rh(II) complexes to albumins, a complex formed by serum albumin and 30-fold molar excess of rhodium (I) was tested in hydroformylation reaction, nevertheless no improved selectivity was reported.[199]

Kaiser's artificial metalloenzyme involved the carboxypeptidase A (CPA) as protein host. This zinc containing metalloenzyme exhibits peptidase and esterase activity. By replacing the active-site zinc by copper the CPA was transformed into an effective catalyst for the oxidation of ascorbic acid to dehydroascorbic acid. The approach was termed chemical mutation. The new artificial metalloenzyme showed neither peptidase nor esterase activity, but was capable of binding the typical CPA inhibitors, such as β-phenylpropionate and *(rac)*-α-benzylsuccinate. This property indicates the preservation of the native structure of the protein binding site.[193]

Later on, two different groups created an enantioselective oxidase by replacing the active-site zinc in carbonic anhydrase with manganese. The hybrid catalyst was used for enantioselective epoxidation of styrene with hydrogen peroxide. The reported

enantioselectivity was moderate, but it was comparable to that of natural peroxidases. Moreover, no unspecific binding of Mn ion was detected.[200-201]

A similar strategy was utilized by Sheldon and co-workers. They replaced the active site zinc of the protease thermolysin with a tungstate, molybdate or selenate oxoanion instead of another metal. This strategy differs from a simple replacement, as zinc and the oxoanions have opposite charge and a different size. The resulting tungstate–thermolysin hybrid catalyst was moderately active but not enantioselective in sulfoxidation reaction.[202]

All methods described so far involved modification of the active site of metalloproteins. Sheldon and co-workers introduced also a different approach based on incorporation of metal ions in non-metalloproteins, in certain specific binding sites. Their strategy exploited the structural similarity of vanadate to phosphate for the efficient inhibition of the hydrolytic activity of a phytase and for generation of a vanadium-incorporated phytase. Initially they used a phytase from *Aspergillus ficuum*, an enzyme which mediates the hydrolysis of phosphate esters *in vivo*. The vanadium based semisynthetic enzyme has been used for the enantioselective sulfoxidation of thioanisole. The choice of the reaction was motivated by the phosphatase-like activity shown by natural vanadium chloroperoxidases and also based on the structural similarity of vanadium chloroperoxidases and the acid phosphatases (Figure 20).[203] Although the metal free phytase also catalyses this reaction, the vanadium-incorporated phytase increased the conversion to 100% and the e.e. up to 66% (Figure 21).[204] Furthermore, this semisynthetic peroxidase was stable for more than three days at relative high concentration of hydrogen peroxide, in contrast to heme-dependent peroxidases which have the disadvantage of low operational stability. Slightly higher enantioselectivities were observed when the same hybrid catalyst was tested for the enantioselective sulfoxidation of ethyl-phenyl-sulphide and *meta* and *para* substituted thioanisoles under the same conditions. Several other transition metal oxoanions were incorporated in the active site of phytase and the catalytic properties of the hybrid catalysts were tested in the same conditions, but only the vanadate-incorporated phytase showed genuine peroxidase activity.[205]

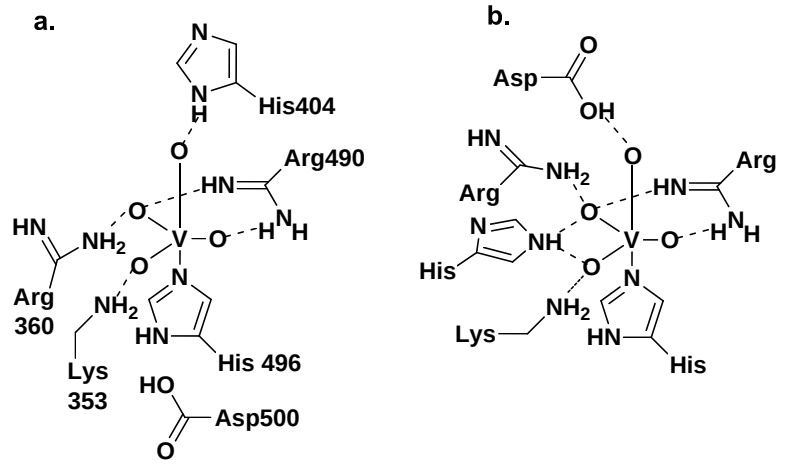
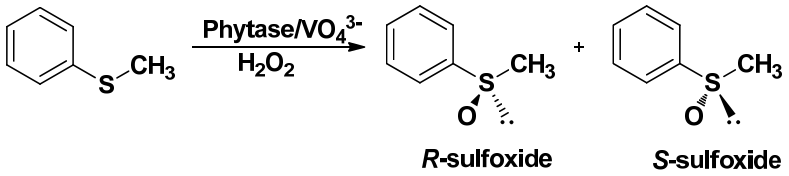


Figure 20 Schematic representation of the active sites of vanadium chloroperoxidase from *C. inaequalis* (a) and vanadate modified phytase from *A. ficuum* (b)[205]



Conversion: >99%
Selectivity: >99% (sulfoxide)
e.e. up to: 66% (*S*-enantiomer)

Figure 21 Sulfoxidation of thioanisole catalyzed by vanadium-modified phytase[205]

Following the phytase from *A. ficuum*, other phytases from different sources, but also other enzymes (acid phosphatases, aminoacylases, albumins, phospholipases, sulfatases) were tested for peroxidase activity with and without vanadate incorporation. In the presence of vanadate, several enzymes exhibited peroxidase activity: the phytases from *A. fumigatus* and *A. nidulans*, sulfatase from *Helix pomatia* and phospholipase D from cabbage. In the absence of vanadate, only phytase from *A. ficuum* gave catalytic enantioselective sulfoxidation of thioanisole.[206]

The first attempt to create a semisynthetic hydrogenase via the dative anchoring strategy was performed by Zhu and co-workers (1993).[207] Following Kaiser's procedure, Zhu and co-workers were successful in replacing the zinc ion of CPA with cobalt, nickel and rhodium, metals known to catalyse hydrogenation reactions. However, Co-, Ni-, Rh-CPAs and the native Zn-CPA did not show any activity as hydrogenation catalysts.[207]

Kazlauskas and co-workers created a semisynthetic hydrogenase capable of direct hydrogenation of substrates with hydrogen, by replacing the active-site zinc of carbonic anhydrase with rhodium. The problem of cofactor regeneration was surpassed by using

molecular hydrogen. The Rh(I) – carbonic anhydrase was reported as an active and selective catalyst for direct hydrogenation of olefins, showing a 20:1 selectivity for hydrogenation of cis-stilbene over trans-stilbene. The enantioselectivity was less than 10% in the products, for preliminary experiments with prochiral substrates. Two approaches were efficiently applied to minimize non-specific binding of rhodium to protein while allowing its binding to the active site: site-directed mutagenesis and combined site directed mutagenesis with chemical modification of histidines.[196]

Serum albumins

Serum albumins are a good choice for generation of enantioselective hybrid catalysts because they are robust proteins, readily accessible and easy to handle.[172] The first enantioselective semisynthetic enzyme based on dative incorporation of a metal moiety into a protein environment was based on BSA and it was reported by Kokubo and co-workers (1983). They used a 1:1 complex of osmium tetroxide and BSA as catalyst in *cis* dihydroxylation of alkenes using *tert*butylhydroperoxide as oxidant. The anchoring of OsO_4 inside BSA was speculated to occur via two dative bonds with lysine residues. The highest enantioselectivity achieved was 68% *(S)* for the dihydroxylation of α-methylstyrene.[208]

It has been reported that sodium salts of di-, tri- and tetrasulfonic acid derivatives of porphyrins, phthalocyanines and corroles bind strongly to serum albumins in a non-covalent manner.[209-213] The first assignment for structural localization of a hemin group inside HSA was reported by Curry and co-workers.[214]

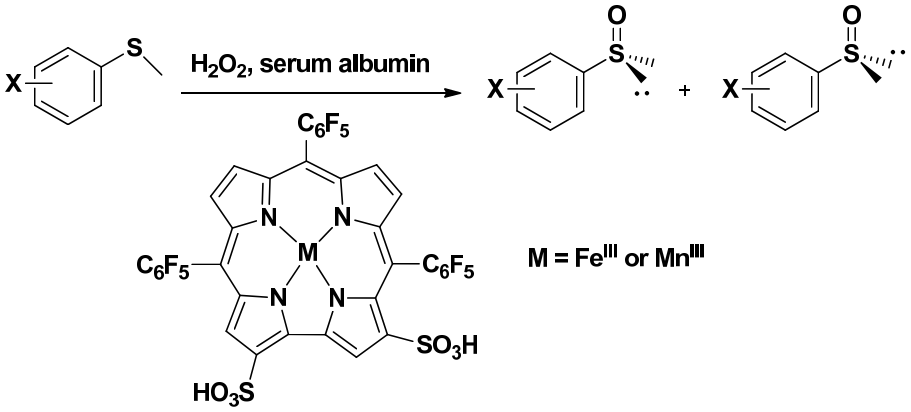


Figure 22 Albumin-conjugated corrole metal complexes as enantioselective catalysts in asymmetric sulfoxidation[215]

Mahammed and Gross anchored amphiphilic bis-sulfonated Mn^{III} and Fe^{III} corroles to several serum albumins. The resulting hybrids were used for asymmetric oxidation of prochiral sulfides to sulfoxides by hydrogen peroxide, with up to 74% e.e (Figure 22). The albumin source had a significant influence on enantioselectivity. Five sources of serum albumin were tested, the highest e.e. being obtained for conjugates of BSA. Throughout the series of substrates and albumins, the results obtained with Mn^{III} corrole conjugates were significantly superior to those with Fe^{III} corrole conjugates regarding enantioselectivity, chemical yield and stability of the catalyst. The limited stability, reflected by catalyst inactivation, is the main disadvantage for the related porphyrin-based systems.[216] However, no catalyst inactivation and/or protein oxidation was observed for the oxidation reactions catalysed by albumin-conjugated manganese corroles in presence of hydrogen peroxide.[215]

Reetz and co-workers used copper – phthalocyanine conjugates of serum albumins to catalyze Diels-Alder reaction of azachalcones with cyclopentadiene in aqueous conditions. High *endo*-selectivities (91-95%) and enantioselectivities (85-98% e.e.) were reported (Figure 23). Among the six different serum albumins tested, the best selectivities were achieved with BSA.[217]

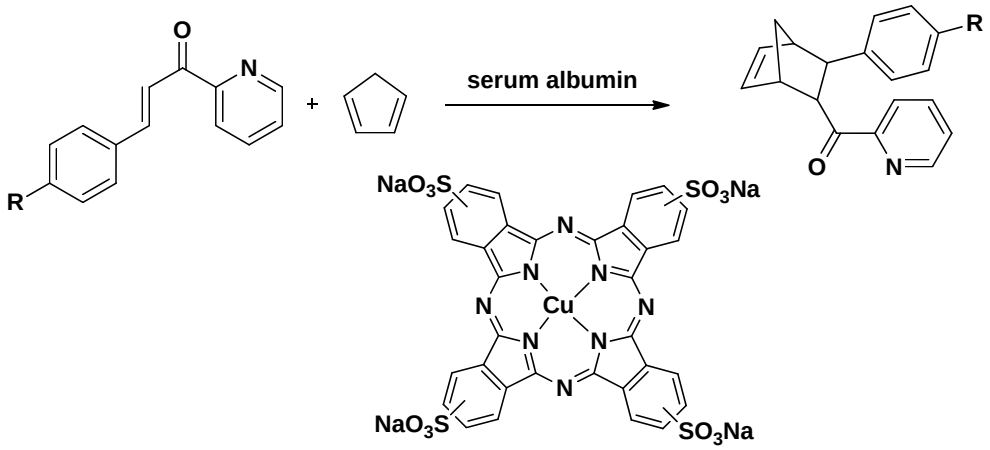


Figure 23 Serum albumin conjugated copper-phthalocyanines as enantioselective catalysts for Diels-Alder reaction[217]

Serum albumins without additional metal complexes were previously used as catalysts for sulfoxidation and Diels-Alder reaction (see sub-chapter I.3). However, for the particular reactions tested by Gross and Reetz the catalytic activity of serum albumins alone was negligible under identical conditions.[217]

Curry and co-workers confirmed the binding of a hemin complex to HSA through axial coordination of tyrosine Tyr161 to the iron ion.[214] Therefore it is believed that binding of Mn^{III} and Fe^{III} corroles (Figure 22) and Cu-phtalocyanine (Figure 23) to serum albumins is realized via this tyrosine residue as well.[137]

Apomyoglobin

A novel strategy based on reconstitution of apomyoglobin (apo-Mb) and several apomyoglobin mutants with symmetric M^{III} Schiff base complexes (M=Cr and Mn) was reported by Watanabe and co-workers.[184] Apo-Mb is an excellent candidate for construction of artificial metalloenzymes because it possesses a cavity of 10Å which accommodates the heme prosthetic group. The heme group is bound non-covalently to the protein via hydrogen bonds and hydrophobic interactions and also via coordination of histidine His93 to the heme iron.[218] Additionally the apo-Mb reconstitution with heme has been well studied.[219-222]

M^{III} Schiff base complexes (Figure 24) have been used as catalysts for several oxidation reactions in organic solvents. On this basis the resulting artificial metalloenzymes were used for enantioselective sulfoxidation of thioanisole with hydrogen peroxide. The reactivity and enantioselectivity reported for the M^{III}-Schiff base complexes coupled to the wild type apo-Mb were low.

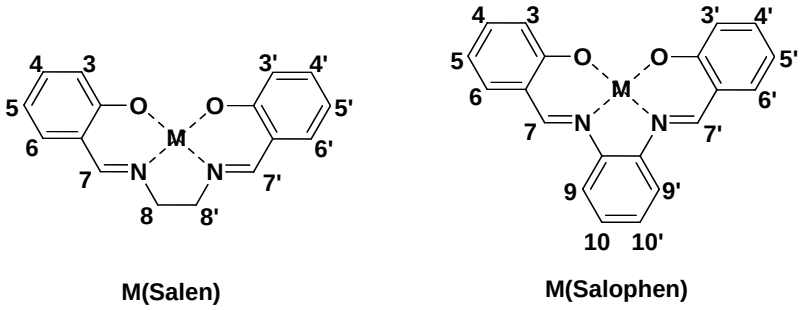


Figure 24 Schiff base complexes inserted into apomyoglobin[223]

Using a rational design approach based on the X-ray crystal structure data of metal complex-protein hybrids, the reactivity and enantioselectivity were improved either by site directed mutagenesis or by changing the position of the metal ion.[224] The position of the metal complex in apo-Mb is regulated by noncovalent interactions between the ligand and surrounding protein and by the coordination of histidine His93 to the metal ion. In addition, hydrophobic interactions between isoleucine Ile107 and the 3 and 3'-

substituents on the salen/salophen ligand control the location of the Schiff base ligand in the active site (Figure 25). Therefore it was possible to successfully control the enantioselectivity by changing the size of the substituents at 3 and 3' positions. Moreover, the enantioselectivity was inverted by replacing the 3 and 3' - methyl groups with more bulky ethyl and *n*-propyl groups.[223]

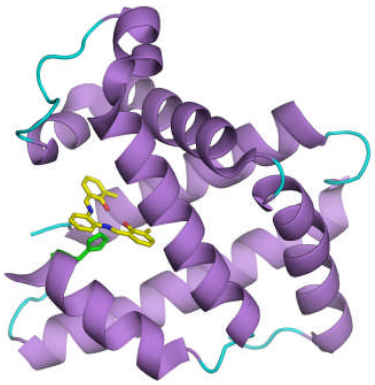


Figure 25 Cartoon representation of the crystal structure of Cr^{III}(3,3'-Me_2-salophen)/apo-A71G Mb (PDB ID 1J3F)[224] (His 93 of apo-A71G Mb is represented in green**)**

The importance of the anchoring strategy for the catalytic properties of an artificial metalloenzyme was highlighted by Lu and co-workers. To optimize the low e.e. of sulfoxidation reactions catalyzed by M^{III} Schiff base complexes non-covalently attached to the wild type myoglobin[184,225] it was necessary to modify either the myoglobin binding pocket by site directed mutagenesis or the Schiff base ligand by changing its substituents.[223] Using these strategies, the e.e. was increased up to 33% *(S)* and 27%*(R)*.[223] Another strategy which increased slightly the enantioselectivity (12%) was the covalent incorporation, by single point attachment, of a Mn-salen complex into the Y103C apo-Mb mutant.[225] A similar e.e. (<10%) was reported by Reetz and co-workers for a Mn-salen complex attached covalently to papain.[12] The low enantioselectivities reported for the single point covalent attachment strategy suggest that the metal complex can adopt multiple conformations inside the binding site of the protein. This hypothesis was corroborated by the significant increase in e.e. (51%) when the Mn-salen complex (Figure 26) was incorporated into the L72C/Y103C myoglobin mutant via two disulfide bridges. The dual anchoring limits the number of possible conformational states of the metal complex inside the protein cavity and as a result improves the e.e. The high enantioselectivity reported for the dual anchoring strategy

was obtained without additional genetic mutations of myoglobin or modification of the salen group. This suggests that the e.e. could be even further improved via chemogenetic optimization.[225]

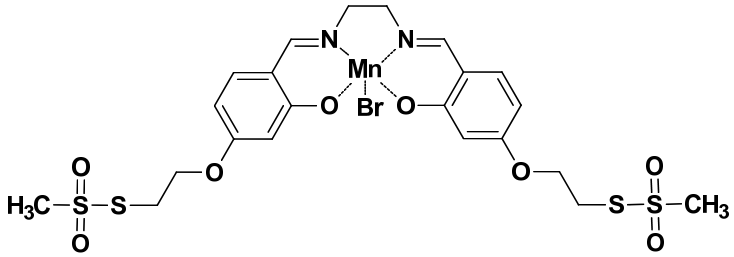


Figure 26 Mn-salen complex coupled to cysteines by methane thiosulfonate groups[225]

I.5 Covalent anchoring

Covalent anchoring is an alternative approach to create artificial metalloenzymes. The most used method for covalent protein functionalization is the modification of the thiol group of a cysteine. Modifications of other amino acid residues have also been described. A number of groups (Reetz, van Koten) reported the synthesis of hybrid catalysts based on modification of an activated serine residue of lipase.[12,226-228] Lysine covalent modification of proteins (BSA, lysozyme, avidin) with several organotransition metal complexes was reported by Salmain and co-workers. However, protein labelling was not site-selective and involved several binding sites, resulting in mixtures of modified proteins.[229-230] Specific labelling of a lysine residue of a protein (the aldolase antibody 38C2) was reported by Janda and co-workers.[231] Kaiser and co-workers used the asparagine residues Asp-52 and Asp-101 of lysozyme for the synthesis of two flavolysozyme isoenzymes via ester bond formation betwteen flavin analogues and asparagines.[232]

Selective cysteine modification is a particularly useful bioconjugation reaction. It exploits the high reactivity of the sulfhydryl group of cysteine but also the relatively modest occurrence (1.7%) of cysteine in peptides and proteins.[233-234] As a consequence, the requirement of a unique cysteine for a selected protein can be easily accomplished via site directed mutagenesis. The cysteine location and consequently the position of the catalytic moiety within the protein cavity can also be adjusted via site directed

mutagenesis. This can be a useful tool for controlling the enantioselectivity of a hybrid catalyst.

Papain

Kaiser and co-workers were the first to report a semisynthetic enzyme created by covalent modification of a cysteine residue of papain with an isoalloxazine derivative.[235] The strategy of the initial concept was to combine the substrate specificity of the enzyme with the characteristic chemical reactivity of the coenzyme analogue. By alkylation of the active site cysteine-25 of papain with the flavin analogue 7α-(bromoacetyl)-10-methylisoalloxazine (Figure 27- a) a semisynthetic oxidoreductase for the oxidation of dihydronicotinamides was synthesized (Figure 27- d). The novel synthetic enzyme exhibited saturation kinetics at low substrate concentrations and significant rate accelerations relative to the model flavin 7-acetyl-10-methylisoalloxazine.[235-236]

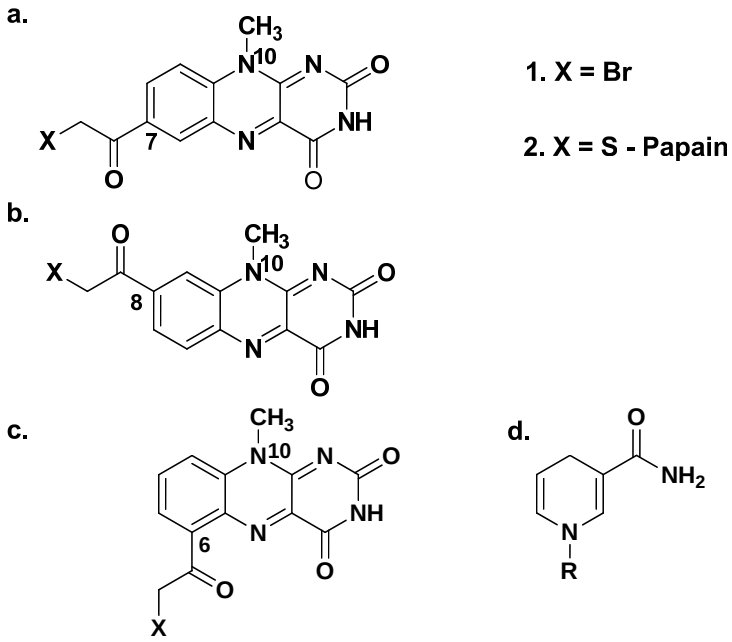


Figure 27 Isoalloxazine derivatives covalently coupled to papain to generate semisynthetic flavoenzymes for the oxidation of dihydronicotinamides[237-238]

Subsequently, two other flavopapain isoenzymes were created by alkylation of Cys-25 with 6α- and 8α-(bromoacetyl)-10-methylisoalloxazine (Figure 27- b and c). Even though the three parent isomers (6, 7 and 8-acetyl-10-methylisoalloxazine) have similar chemical reactivity, the resulting flavopapains showed completely different rate enhancements for catalysis of the oxidation of dihydronicotinamides (Figure 27- d) by

oxygen. While flavopapain (c) showed no rate enhancement and flavopapain (a) exceeds about 10 times the oxidation rate of the flavin prostethic group, flavopapain (b) shows a catalytic rate enhancement of 100-fold to almost 1000-fold, which is comparable to the activity exhibited by naturally occurring flavin containing oxidoreductases.[238]

The three flavoenzymes were used also as catalysts for the oxidation of various dithiols to disulfides, under anaerobic conditions. The most efficient catalyst was again the flavopapain (b), with a catalytic rate enhancement of 100-fold relative to the model flavin.[239-240]

The substantial differences in efficiency between the three flavopapains were assumed to be the result of the differences in the flavin geometry at the active site. The flavin orientation is determined by the hydrogen interactions to the neighbouring residues in the protein backbone.

After Kaiser's report, several other protein templates were covalently modified with isoalloxazine derivatives to generate semisynthetic flavoenzymes.

Two flavolysozyme isoenzymes were created via ester bond formation between bromoacetyl derivatives of isoalloxazines and the asparagine Asp-101 and respectively Asp-52 of lysozyme. In this manner Kaiser's initial concept was extended to proteins that lack a highly reactive amino acid.[232]

A semisynthetic flavohemoglobin capable of carrying out aniline hydroxylation in a similar mode with microsomal cytocrome P-450, was created by covalently attaching an isoalloxazine moiety to hemoglobin in the vicinity of heme. The activity of NADPH-dependent flavohemoglobin Fl-Hb^{3+} was similar to that of ferric hemoglobin Hb^{3+} reconstituted with NADPH-cytochrome P-450 reductase for this hydroxylation reaction.[241]

Another protein used to generate semisynthetic flavoenzymes was glyceraldehyde-3-phosphate dehydrogenase (GAPDH).[242] Two types of GAPDH were employed. The first one was the rabbit muscle GADPH. The resulting flavo-enzyme was used as catalyst for oxidation of NADH, NADPH and 1,4-dihydronicotinamides, showing a rate acceleration of 120-fold for the oxidation of NADH.[242] The second one was the GADPH from *Bacillus stearothermophilus*. The subsequent flavo-GADPH showed 6000-fold rate acceleration for the oxidation of NADH by molecular oxygen at low substrate concentration. It was also capable of catalysing the oxidation of dihydronicotinamides at temperatures as high as 55°C.[243]

Following Kaiser's study on flavopapains, several other chemical modifications at the active site sulfhydryl of Cys-25 of papain were reported, with coenzyme analogues or with organic ligands.

Zhu and co-workers reported thiazolium modified papains as the first semisynthetic enzymes capable of mediating carbon-carbon bond formation. The main reaction was the dimerisation of 6-oxoheptanal (a) to 7-hydroxytetradecane-2,8,13-trione (b) and the secondary reaction was cyclisation of 6-oxoheptanal (a) to 2-methylcyclohex-2-enone (c) (Figure 28).[244]

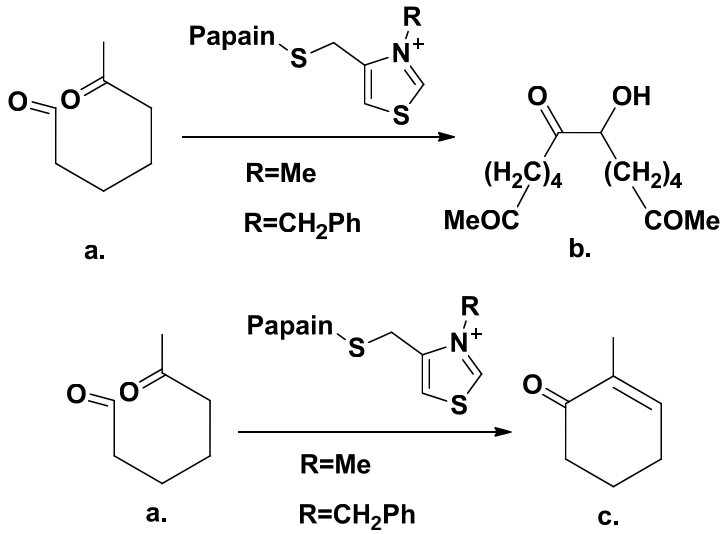


Figure 28 Thiazolium modified papains as hybrid catalysts for dimerisation and cyclisation of 6-oxoheptanal[244]

De Vries and co-workers reported a papain-hybrid catalyst synthesized by covalently coupling a monodentate phosphite ligand to the cysteine Cys-25 (Figure 29). The artificial metalloenzyme was active in the hydrogenation of methyl 2-acetamidoacrylate with hydrogen (100% conversion). However, no enantioselectivity was reported, presumably because of the large flexibility of the catalyst or because of the catalyst location far from the chiral environment of the enzyme. This study shows clearly that active catalysts can be obtained via site specific modification of a cysteine residue in proteins with transition metal complexes.[245]

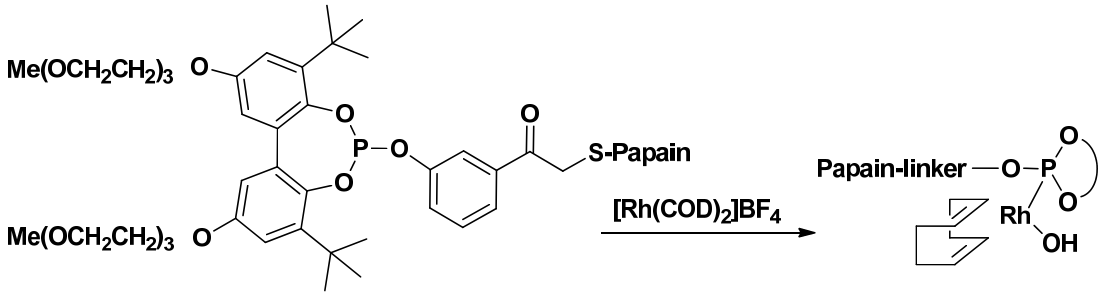


Figure 29 Papain based hybrid catalysts synthesized via covalent modification of Cys-25 with a monodentate phosphite ligand and complexation with rhodium[245]

Reetz and co-workers modified Cys-25 of papain with a manganese-salen complex (Figure 30- a) and with dipyridyl moieties (Figure 30- b) complexing copper, palladium and rhodium. The Mn-salen hybrid catalyst was tested in epoxidation reactions and the Rh-catalyst in hydrogenation, but in all cases the enantioselectivities were not higher than 10%, confirming once again the lack of stereocontrol of the papain based hybrid catalysts.[12,246-247]

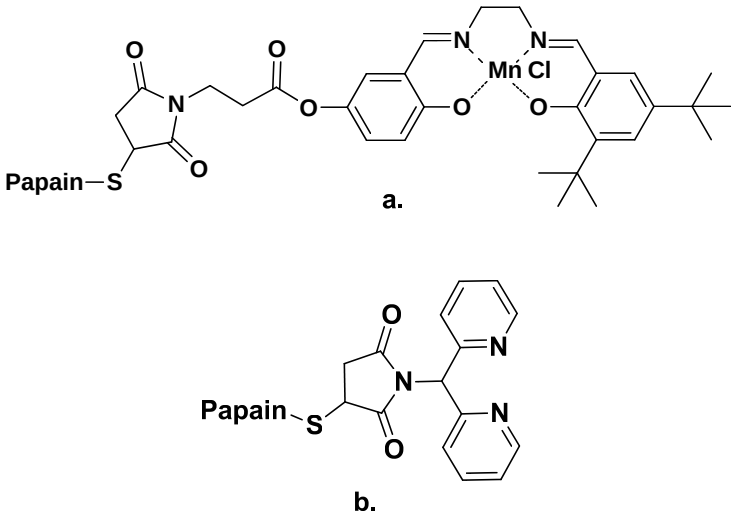


Figure 30 Covalent modification of the thiol group of Cys-25 of papain with a Mn-salen complex (a) and with a dipyridyl moiety (b)[12,247]

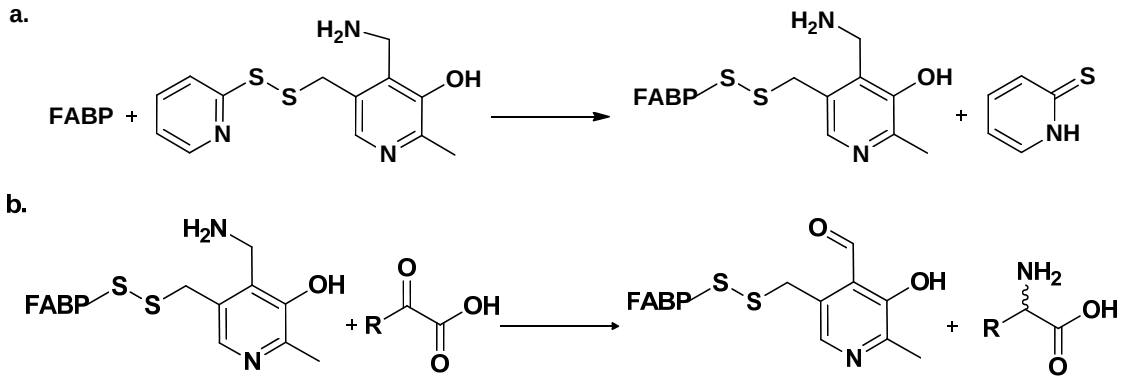
Three organometallic maleimides (Figure 31) derived from ferrocene, ruthenocene and cyrhetrene (η^5-cyclopentadienyl tricarbonyl rhenium(I)) were synthesized and covalently coupled to the Cys-25 of papain by Jaouen and co-workers. The transition metal organic complexes were investigated as heavy-metal reagents for protein structural determination by X-ray crystallography. This work confirms once again the

possibility of chemospecific modification of proteins with transition metal complexes via cysteine modification.[248]

Figure 31 Organometallic *N*-substituted maleimides coupled covalently to Cys-25 of papain[248]

FABPs (fatty acid binding proteins)

Kaiser's approach for creation of semisynthetic enzymes by combining the substrate specificity of naturally occurring enzymes with the characteristic reactivity of natural cofactors was amended by Distefano and co-workers. Instead of using natural enzymes, they employed carrier proteins to synthesize pyridoxamine (Px) conjugates (Figure 32- a). Two fatty acid carrier proteins (FABP) had been tested, the adipocyte lipid binding protein (ALBP) and the intestinal fatty acid binding protein (IFABP). The ALBP-Px conjugate contains the pyridoxamine cofactor coupled via a disulfide bridge to the unique cysteine residue Cys117. This artificial enzyme was shown to be an active and enantioselective catalyst (up to 94% e.e.) for the reductive amination of a number of alkyl, aryl and side-chain functionalized α-keto acids to α-amino acids (Figure 32- b).[249] Moreover, it was possible to tune the rate, the enantioselectivity and the substrate specificity of the semisynthetic transaminases. This was done by varying the position of pyridoxamine inside the protein cavity, through changing the position of cysteine via site directed mutagenesis.[250] To increase the turnover numbers it was necessary to overcome the pyridoxamine regeneration, the factor limiting turnover. This was solved by using a second amino acid as an amine source (Figure 32- c).[251]



c.

Figure 32 FABP-Px used as hybrid catalysts for transamination reactions[249-251]

By replacing pyridoxamine with N-methylated pyridoxamine, a permanently positive charge was placed on the cofactor. In this way a native transamination site was mimicked, by stabilizing a cationic pyridoxamine ring system.[252]

A major step forward was achieved by the synthesis of 1,10-phenanthroline conjugate of FABPs, thus enabling coordination of a metal ion within the defined protein cavity (Figure 33- a). The ALBP-Phen-Cu(II) conjugate catalyzed the enantioselective hydrolysis of several unactivated amino acid esters (with e.e.'s ranging from 31% to 86%) and the hydrolysis of amides (Figure 33- b and c).[253]

The rate and the selectivity of phenanthroline derivatives were enhanced by modifying the location of phenanthroline ligand attachment within the protein cavity to three different positions. In this way the cofactor orientation and the environment around the catalytic metal centre were modified .[251]

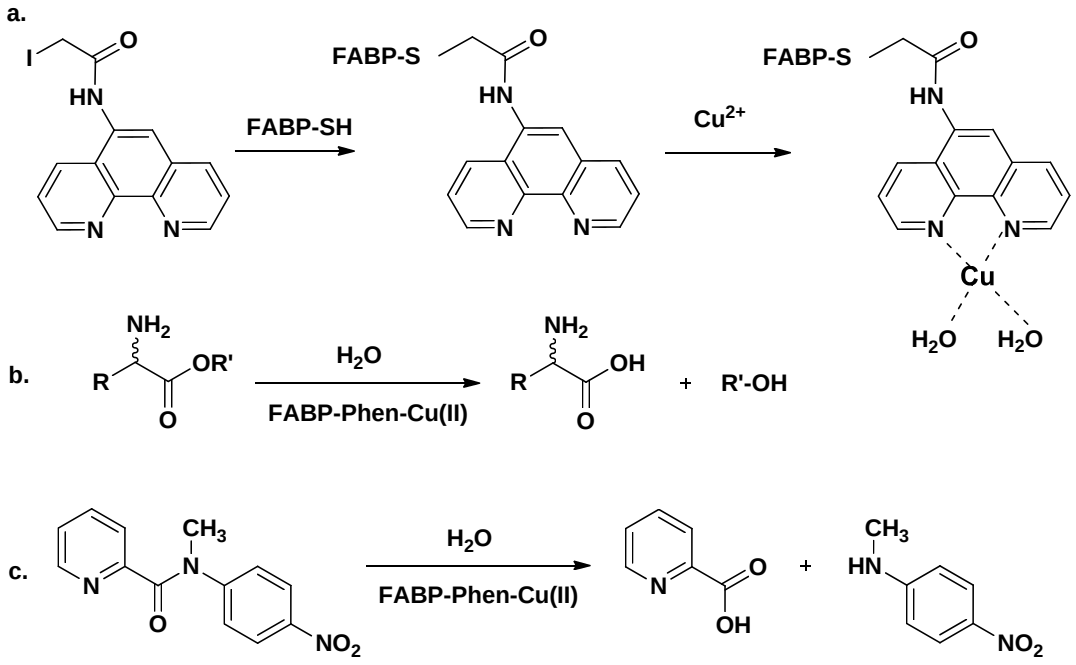


Figure 33 Cu(II)-Phenanthroline conjugates of fatty acid binding proteins used as hybrid catalysts for ester and amide hydrolysis[253]

Figure 34 Cartoon representation of the crystal structure of ALBP-phenanthroline (PDB ID 1A18)[254]

Lipases

The hydroxyl group of serine has also been exploited for site specific modification with organic ligands. To circumvent multiple modifications of serines, a successful idea was to use lipases as protein backbone and phosphonate inhibitors as anchoring units, in order to achieve a single site-directed modification. The phosphonate inhibitors are capable of reacting covalently at the catalytically active serine site of a lipase, selectively in the presence of other serine residues in the enzyme. The concept was applied first by Reetz and co-workers. By using a dinitrophenol phosphonate diphosphane, they incorporated a bidentate diphosphane ligand in the lipase from *Bacillus subtilis* (Figure 35). However, the resulting hybrid was unstable and after 24 h the inhibitor was cleaved hydrolytically in water[167,246].

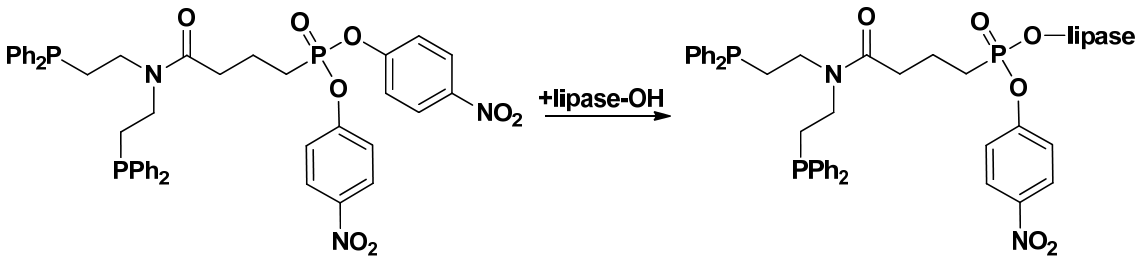


Figure 35 Bidentate diphosphane covalently coupled to the reactive serine residue in the active site of lipase via a phosphonate inhibitor[167,246]

To prevent the hydrolysis of the second *p*-nitrophenol leaving group, van Koten and co-workers replaced it by ethanol. A small series of phosphonate esters were synthesised and covalently coupled to different ECE-pincer-type (ECE=[$C_6H_3(CH_2E)_2$-2,6]; E= NR_2 or SR) metal complexes through a propyl tether (Figure 36- a). The organometallic phosphonate-pincer-metal complexes were successfully coupled to the serine residue

Ser[120] of the lipase cutinase. The resulting hybrids were stable in aqueous buffer solutions and tolerated aqueous dialysis purification conditions, but so far they were not tested in catalysis.[226]

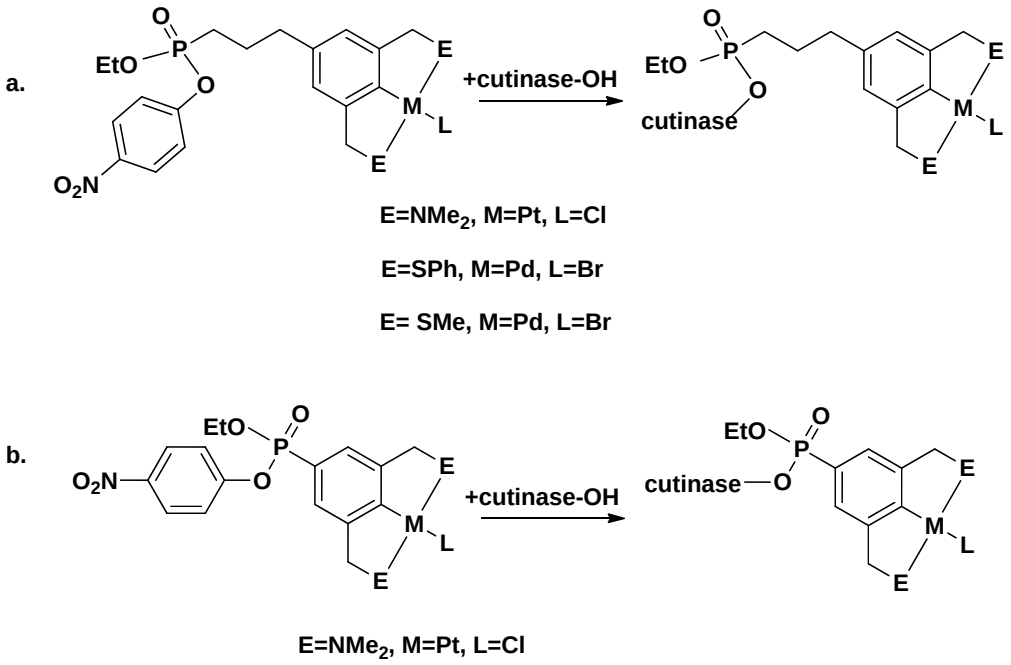


Figure 36 Selective covalent modification of cutinase with different organometallic phosphonate-ECE pincer-metal complexes[226-227]

The presence of the propyl linker placed the metal centre of the ECE pincer out of the cutinase molecule, as was shown by the crystal structures of the modified cutinase.[228] The location outside of the metal centre allowed different orientations of the metal complex. Consequently, the chirality induced by the protein environment in the catalytic reactions of the pincer-metal complex is anticipated to be low. To achieve a closer contact between the pincer-metal group and the chiral protein environment, different pincer-metal cutinase hybrids were synthesized in which the C_3 linker was removed and the pincer-metal complex was directly attached to the phosphonate group (Figure 36-b).[227]

The nature of the ECE-pincer-metal group is also important, as was shown by the crystal structures of the covalently modified cutinase. The ECE-pincer-platinum and palladium inhibitors occupy different pockets in the active site of cutinase and impose different stereochemistry on the phosphorus atom.[228]

Seleno-subtilisin

Already in 1966, Polgar and Bender reported the preparation of a semisynthetic enzyme, a thiol-subtilisin, by transforming the reactive serine residue Ser221 in the active site of subtilisin to a cysteine residue.[255-256] By applying this approach Hilvert and co-workers achieved the synthesis of a seleno-subtilisin, through converting the active site serine Ser221 into cysteine and subsequently into selenocysteine.[257] The new semisynthetic enzyme displayed novel acyl transferase properties as it favoured aminolysis over hydrolysis, compared to the native subtilisin. In addition, it was capable of catalysing the reduction of hydroperoxides by thiols, similar to the natural selenoprotein glutathione peroxidase.[258-259] Further investigations of Schreier and co-workers reported enantioselective reduction of a variety of racemic hydroperoxides catalysed by seleno-subtilisin. The catalytic efficiency of seleno-subtilisin was comparable to that of native peroxidases. However, the semisynthetic enzyme exhibited the opposite sense in enantioselectivity, suggesting that it may optimally complement the naturally available peroxidases.[260] Cross-linked microcrystals of seleno-subtilisin were synthesised using glutardialdehyde as cross-linker. The microcrystals are insoluble in water and organic solvents, can be easily recovered by centrifugation or filtration and recycled. Moreover, the e.e.'s obtained using microcrystals are similar with those showed by the amorphous seleno-subtilisin. The high activity of microcrystals was retained in organic solvents too and remained high when recycled, due to their stability.[261]

Antibodies

The first covalent anchoring of an organic ligand to an antibody was reported in 2002 by Janda and co-workers. They synthesized an artificial metalloenzyme by coupling a bis-imidazole cofactor to the aldolase antibody 38C2 (Figure 37- a). This antibody was selected for modification because of its large hydrophobic pocket containing a highly nucleophilic lysine residue, Lys H^{93}, suitable for covalent modification. The choice of the metal binding bis-imidazole cofactor was inspired by the active-site structures of a number of natural metalloenzymes which possess a coordination sphere with two or more histidine-derived imidazole ligands. The resulting 38C2 - bis-imidazole derivative - Cu hybrid displayed esterase activity, being an active catalyst for the hydrolysis of a picolinate ester (Figure 37- b).[231]

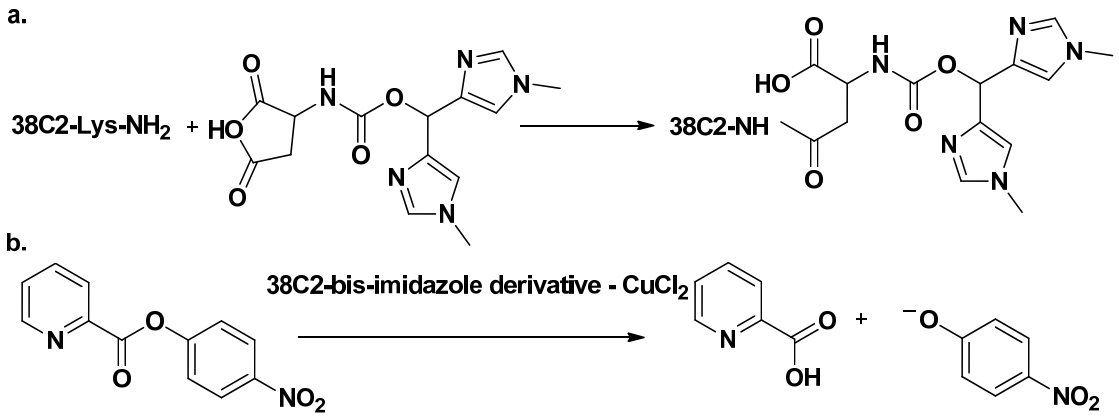


Figure 37 Covalent modification of 38C2 antibody with a bis-imidazole cofactor (a) to generate a Cu-complexed hybrid catalyst for the hydrolysis of 4-nitrophenyl picolinate (b)[231]

I.6 DNA-based asymmetric catalysis

Proteins are not the only source of chirality which have been used for generation of enanioselective artificial metalloenzymes. Roelfes and Feringa reported the synthesis of hybrid catalysts formed via non-covalent incorporation of copper complexes into DNA. Two families of ligands have been investigated. Initially a metal complex of a non-chiral ligand was connected through a spacer to a DNA intercalating moiety.[262] Later on, in an improved approach, bidentate ligands known to bind to DNA were used as DNA binding moieties, omitting the requirement of a spacer.[263-265] Enantioselectivities up to 99% were obtained using Cu(II) complexes coupled to salmon testes DNA for the Diels-Alder reaction of azachalcone with cyclopentadiene[263] and for Michael reactions.[264] In addition, the role of DNA is not only as a chiral scaffold, as it was proved for the Diels-Alder reaction. Both the enantioselectivity and the rate enhancement of the reaction are dependent on the DNA sequence, as shown by replacing the salmon DNA with synthetic oligonucleotides of defined sequence.[265]

Covalent modification of aminodeoxyuridines and amino-functionalized oligonucleotides with phosphane ligands,[266-267] but also with a Pd-phosphane complex were recently reported.[267] Another strategy used for DNA covalent modification consists of Pd catalysed coupling of diphenylphosphane to 5-iodo-nucleosides and 5-iodo-oligonucleotides.[268] Covalent coupling of N-ligand complexes to DNA was also reported. A Cu(II) complex of (2,2'-bipyridine)-5-carboxylic acid was covalently coupled through an alkyl linker to a DNA double strand and was used as catalyst for the Diels-Alder reaction.[269]

I.7 Scope and outline of the thesis

As is clear from the overview above, the combination of biological concepts for selective recognition with those of transition metal catalysis is a very promising approach for the development of selective hybrid catalysts for important asymmetric catalytic reactions for which there are no enzymes available. The long term objective of this work is to synthesize novel artificial metalloenzymes based on the covalent site-selective incorporation of organometallic catalysts into proteins, to use them for asymmetric reactions and to tune their properties via chemogenetic optimization.

Targeting the development of transition metal complexes coupled covalently to proteins, the initial objective was the introduction of metal binding ligands into proteins. Phosphane ligands are an important class of ligands, not only in academic organometallic chemistry, but also in catalytic reactions on industrial scale. Chiral phosphane ligands are widely used in transition metal catalyzed asymmetric synthesis. However, for a number of reactions the enantioselectivity achieved with these ligands remains modest. Therefore, the need for novel, highly selective chiral phosphane-based catalysts for asymmetric reactions remains high. However, high-throughput synthesis of structurally diverse phosphane ligands toward catalysts optimization remains a challenging and time-consuming process. By merging relatively simple achiral phosphanes with protein scaffolds as chiral inducing environment, the powerful methods of chemogenetic engineering and directed/designed evolution can be applied to the optimization of phosphane-based catalysts.

Chapter II describes several strategies explored for covalent site-specific incorporation of phosphane ligands into proteins. The first studied strategy was a thioester bond formation between the thiol group of a unique cysteine of a protein and a carboxylic group of a phosphane ligand. Problems encountered with the specificity of this method led us to the second strategy, formation of a thioether bond between the thiol group of cysteine and the double bond of a maleimide functionalized ligand. This method demanded the use of protected phosphane ligands, as phosphanes react with the double bond of maleimides. Several problems were encountered with phosphane deprotection. Therefore, the successful strategy was to incorporate a maleimide functionalized hydrazide into proteins and to couple several aldehyde functionalized phosphanes via hydrazone linkage.

Nitrogen ligands have also a large applicability in homogeneous catalytic reactions, due to their distinctive advantages. Chapter III describes strategies explored for site-specific covalent incorporation of this class of ligands into proteins. Covalent incorporation of a carboxylic acid functionalized bipyridine was achieved after NHS (N-hydroxysuccinimide) activation of the carboxylic group. The activated ligand was coupled via thioester bond formation with a thiol group of a cysteine residue, however low selectivity towards cysteine was observed for the proteins tested. Succesful site selective incorporation of N-containing ligands was obtained by coupling maleimide functionalized N-ligands to proteins via Michael addition to the maleimide double bond.

The final aim is construction of artificial metalloenzymes by coupling transition organometallic catalysts into proteins. Therefore the fourth chapter contains several strategies which have been explored in order to reach this target. Site-specific covalent incorporation of transition metal complexes (Rh, Pd) of phosphane ligands into proteins was successfully obtained.[270-271] However, the success of the approach is influenced by several factors like the metal precursor, the phosphane type and the protein scaffold. Pd and Rh metal complexes of phosphane ligands proved to induce protein instability for two particular proteins we investigated, ALBP and Δ98Δ. These problems were not encountered when metal complexes of maleimide functionalized N-ligands were tested. Efficient and site selective covalent modification of a three different protein scaffolds with metal complexes of 5-maleimido-1,10-phenathroline was obtained.

REFERENCES

1. Schoemaker, H. E.; Mink, D.; Wubbolts, M. G. *Science* **2003**; *299*; 1694
2. Anastas, P. T.; Zimmerman, J. B. *Environ. Sci. Technol.* **2003**; *37*; 94a.
3. *R&D Magazine* **September 2005**; 20.
4. Zapf, A.; Beller, M. *Topics in Catalysis* **2002**; *19*; 101.
5. van Leeuwen, P. W. N. M. *Homogeneous Catalysis*; Kluwer Academic Dordrecht, The Nederlands, 2004.
6. Green, S. *Industrial Catalysis* Macmillan Company: New York, 1928.
7. Berzelius, J. J. *Årsberättelsen om framsteg i fysik och kemi*; Royal Swedish Academy of Sciences, 1835.
8. Schwab, G. M. *Catalysis Science and Technology*; Springer: New York, 1981; Vol. 2.
9. McNaught, A. D.; Wilkinson, A. *IUPAC. Compendium of Chemical Terminology*; Blackwell Scientific: Oxford, 1997.
10. Reetz, M. T. *Angew. Chem. Int. Ed. Engl.* **2001**; *40*; 284.
11. Klibanov, A. M. *Nature* **2001**; *409*; 241.
12. Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Maywald, M.; Maiwald, P.; Peyralans, J. J. P.; Maichele, A.; Fu, Y.; Jiao, N.; Hollmann, F.; Mondiere, R.; Taglieber, A. *Tetrahedron* **2007**; *63*; 6404.
13. Sancho, R.; Minguillo, C. *Chem. Soc. Rev* **2009**; *38*; 797.
14. Francotte, E.; Lindner, W. *Chirality in Drug Research*; Wiley-VCH: Weinheim, 2006; Vol. 33.
15. In *International Conference on Harmonisation (ICH) (http://www.ich.org)* Nov. 2000.
16. Castilla, E. E.; Ashton-Prolla, P.; Barreda-Mejia, E.; Brunoni, D.; Cavalcanti, D. P.; Correa-Neto, J.; Delgadillo, J. L.; Dutra, M. G.; Felix, T.; Giraldo, A.; Juarez, N.; Lopez-Camelo, J. S.; Nazer, J.; Orioli, I. M.; Paz, J. E.; Pessoto, M. A.; Pina-Neto, J. M.; Quadrelli, R.; Rittler, M.; Rueda, S.; Saltos, M.; Sanchez, O.; Schuler, L. *Teratology* **1996**; *54*; 273.
17. Teo, S. K.; Colburn, W. A.; Tracewell, W. G.; Kook, K. A.; Stirling, D. I.; Jaworsky, M. S.; Scheffler, M. A.; Thomas, S. D.; Laskin, O. L. *Clinical Pharmacokinetics* **2004**; *43*; 311.
18. Eliel, E. L.; Wilen, S. H.; Mander, L. N. *Stereochemistry of Organic Compounds*; Wiley: New York, 1994.
19. Keith, J. M.; Larrow, J. F.; Jacobsen, E. N. *Adv. Synth.&Catal.* **2001**; *343*; 5.
20. Klar, U.; Rohr, B.; Kuczynski, F.; Schwede, W.; Berger, M.; Skuballa, W.; Buchmann, B. *Synthesis-Stuttgart* **2005**; 301.
21. Jacobsen, E. N.; Pfaltz, A.; Yamamoto, H. *Comprehensive Asymmetric Catalysis*; Springer, Berlin, 1999.
22. Knowles, W. S. *Angew. Chem. Int. Ed. Engl.* **2002**; *41*; 1999.
23. Noyori, R. *Angew. Chem. Int. Ed. Engl.* **2002**; *41*; 2008.
24. Sharpless, K. B. *Angew. Chem. Int. Ed. Engl.* **2002**; *41*; 2024.
25. Young, J. F.; Osborn, J. A.; Jardine, F. H.; Wilkinson, G. *Chem. Commun. (London)* **1965**; 131
26. Halpern, J.; Wong, C. S. *J. Chem. Soc. - Chem. Commun.* **1973**; 629.
27. Halpern, J.; Okamoto, T.; Zakhariev, A. *J. Molec. Catal.* **1976**; *2*; 65.
28. Knowles, W. S.; Sabacky, M. J. *Chem. Commun.* **1968**; 1445.

29. Nozaki, H.; Takaya, H.; Moriuti, S.; Noyori, R. *Tetrahedron* **1968**; *24*; 3655
30. Humphrey, C. E.; Furegati, M.; Laumen, K.; La Vecchia, L.; Leutert, T.; Constanze, J.; Muller-Hartwieg, D.; Vogtle, M. *Organic Process R&D* **2007**; *11*; 1069.
31. Knowles, W. S. In *Asymmetric Catalysis on Industrial Scales*; Blaser, H. U.; Schmidt, E., Eds.; Wiley-VCH Weinheim, 2004; Vol. 1, p 23.
32. Kagan, H. B.; Dang, T. P. *J. Am. Chem. Soc.* **1972**; *94*; 6429.
33. Miyashita, A.; Yasuda, A.; Takaya, H.; Toriumi, K.; Ito, T.; Souchi, T.; Noyori, R. *J. Am. Chem. Soc.* **1980**; *102*; 7932.
34. Blaser, H. U. *Chem. Commun. (Camb)* **2003**; 293.
35. Mitscher, L. A.; Dutta, A. In *Burger's Medicinal Chemistry and Drug Discovery* Burger, A.; Abraham, D. J., Eds.; John Wiley & Sons: 2003; Vol. 2: Drug Discovery and Drug Development, p 1.
36. Hughes, P. F.; Graham, T. H.; Mendoza, J. S. *J. Comb. Chem.* **2004**; *6*; 308.
37. Gennari, C.; Piarulli, U. *Chem. Rev.* **2003**; *103*; 3071.
38. Ward, T. R. *Chem. Eur. J.* **2005**; *11*; 3798.
39. Lehn, J. M. *Supramolecular Chemistry*; Wiley-VCH: Weinheim, 1995.
40. Loeb, S. J. *Comprehensive Supramolecular Chemistry*; Pergamon, Oxford, 1996.
41. Creus, M.; Ward, T. R. *Org. Biomol. Chem.* **2007**; *5*; 1835.
42. Bryce, C. F. A. *Trends Biochem. Sci.* **1979**; *4*; 62.
43. Metzler, D. E. *Biochemistry. The chemical reactions of living cells*; 2nd ed.; Harcourt: San Diego, 2001.
44. Pauling, L. *Chem. Eng. News* **1946**; *161*; 707.
45. Pauling, L. *Nature* **1948**; *161*; 707.
46. Zhang, X. Y.; Houk, K. N. *Acc. Chem. Res.* **2005**; *38*; 379.
47. Yokozeki, K.; Nakamori, S.; Eguchi, C.; Yamada, K.; Mitsugi, K. *Agric. Biol. Chem.* **1987**; *51*; 355.
48. Fulling, G.; Sih, C. J. *J. Am. Chem. Soc.* **1987**; *109*; 2845.
49. Noyori, R.; Ikeda, T.; Ohkuma, T.; Widhalm, M.; Kitamura, M.; Takaya, H.; Akutagawa, S.; Sayo, N.; Saito, T.; Taketomi, T.; Kumobayashi, H. *J. Am. Chem. Soc.* **1989**; *111*; 9134.
50. Kim, M. J.; Ahn, Y.; Park, J. *Curr. Opin. Biotechnol.* **2002**; *13*; 578.
51. Larsson, A. L. E.; Persson, B. A.; Backvall, J. E. *Angew. Chem. Int. Ed. Engl.* **1997**; *36*; 1211.
52. Persson, B. A.; Larsson, A. L. E.; Le Ray, M.; Backvall, J. E. *J. Am. Chem. Soc.* **1999**; *121*; 1645.
53. Verzijl, G. K. M.; de Vries, J. G.; Broxterman, Q. B. 2001; Vol. PCT Int. Appl. WO 01/90396 Al, p 43.
54. Martin-Matute, B.; Edin, M.; Bogar, K.; Backvall, J. E. *Angew. Chem. Int. Ed. Engl.* **2004**; *43*; 6535.
55. Paetzold, J.; Backvall, J. E. *J. Am. Chem. Soc.* **2005**; *127*; 17620.
56. Gastaldi, S.; Escoubet, S.; Vanthuyne, N.; Gil, G.; Bertrand, M. P. *Org. Lett.* **2007**; *9*; 837.
57. Ahn, Y.; Ko, S. B.; Kim, M. J.; Park, J. *Coord. Chem. Rev.* **2008**; *252*; 647.
58. Hoben, C. E.; Kanupp, L.; Backvall, J. E. *Tetrahedron Lett.* **2008**; *49*; 977.
59. Karvembu, R.; Prabhakaran, R.; Tamizh, M. M.; Natarajan, K. *Comptes Rendus Chimie* **2009**; *12*; 951.

60. Thalen, L. K.; Zhao, D. B.; Sortais, J. B.; Paetzold, J.; Hoben, C.; Backvall, J. E. *Chem. Eur. J.* **2009**; *15*; 3403.
61. Liese, A.; Filho, M. V. *Curr. Opin. Biotechnol.* **1999**; *10*; 595.
62. Csjernyik, G.; Bogar, K.; Backvall, J. E. *Tetrahedron Lett.* **2004**; *45*; 6799.
63. Ghanem, A.; Aboul-Enein, H. Y. *Tetrahedron-Asymmetry* **2004**; *15*; 3331.
64. Edin, M.; Backvall, J. E.; Cordova, A. *Tetrahedron Lett.* **2004**; *45*; 7697.
65. DeSantis, G.; Zhu, Z. L.; Greenberg, W. A.; Wong, K. V.; Chaplin, J.; Hanson, S. R.; Farwell, B.; Nicholson, L. W.; Rand, C. L.; Weiner, D. P.; Robertson, D. E.; Burk, M. J. *J. Am. Chem. Soc.* **2002**; *124*; 9024.
66. Koeller, K. M.; Wong, C. H. *Nature* **2001**; *409*; 232.
67. Bommarius, A. S.; Drauz, K.; Hummel, W.; Kula, M. R.; Wandrey, C. *Biocatalysis* **1994**; *10*; 37.
68. Bommarius, A. S.; Schwarm, M.; Drauz, K. *J. Molec. Catal. B-Enzymatic* **1998**; *5*; 1.
69. Ma, S. K.; Gruber, J.; Davis, C.; Newman, L.; Gray, D.; Wang, A.; Grate, J.; Huisman, G. W.; Sheldon, R. A. *Green Chem.* **2010**; *12*; 81
70. Dauphin, G.; Gramain, J. C.; Kergomard, A.; Renard, M. F.; Veschambre, H. *J. Chem. Soc. - Chem. Commun.* **1980**; 318.
71. Sakai, K.; Nakazawa, A.; Kondo, K.; Ohta, H. *Agric. Biol. Chem.* **1985**; *49*; 2331.
72. Li, K.; Frost, J. W. *J. Am. Chem. Soc.* **1998**; *120*; 10545.
73. Carrea, G.; Riva, S. *Organic Synthesis with Enzymes in Non-Aqueous Media*; Wiley InterScience, 2008.
74. Chaudhary, A. K.; Beckman, E. J.; Russell, A. J. *J. Am. Chem. Soc.* **1995**; *117*; 3728.
75. Barzana, E.; Klibanov, A. M.; Karel, M. *Appl. Biochem. Biotech.* **1987**; *15*; 25.
76. Bárzana, E. In *Advances in Biochemical Engineering/Biotechnology*; Springer Berlin / Heidelberg: 1996; Vol. 53, p 1.
77. Gill, I.; Vulfson, E. *Trends Biotechnol.* **1994**; *12*; 118.
78. Itoh, T. *Chorinkai Saishin Gijutsu* **2007**; *9*; 1.
79. Hu, J.; Nian, B.; Wu, J.; Xu, G.; Yang, L. *Huaxue Fanying Gongcheng Yu Gongyi* **2007**; *23* 207.
80. Chen, X.; Du, W.; Liu, D. H. *World J. Microbiol. & Biotech.* **2008**; *24*; 2097.
81. Hilterhaus, L.; Thum, O.; Liese, A. *Organic Process R&D* **2008**; *12*; 618.
82. Hilterhaus, L.; Minow, B.; Muller, J.; Berheide, M.; Quitmann, H.; Katzer, M.; Thum, O.; Antranikian, G.; Zeng, A. P.; Liese, A. *Bioprocess Biosyst. Eng.* **2008**; *31*; 163.
83. Nelson, D. L.; Cox, M. M. *Lehninger Principles of Biochemistry*; 3rd ed.; Worth: New York 2000; Vol. 192.
84. Kuntz, I. D.; Kauzmann, W. *Adv. Protein Chem.* **1974**; *28*; 239.
85. Rupley, J. A.; Careri, G. *Adv. Protein Chem.* **1991**; *41*; 37.
86. Griebenow, K.; Klibanov, A. M. *J. Am. Chem. Soc.* **1996**; *118*; 11695.
87. Klibanov, A. M. *Trends Biotechnol.* **1997**; *15*; 97.
88. Zaks, A.; Klibanov, A. M. *J Biol Chem* **1988**; *263*; 3194.
89. Zaks, A.; Klibanov, A. M. *Proc. Nat. Acad. Sci. USA* **1985**; *82*; 3192.
90. Kitaguchi, H.; Fitzpatrick, P. A.; Huber, J. E.; Klibanov, A. M. *J. Am. Chem. Soc.* **1989**; *111*; 3094.

91. Gutman, A. L.; Meyer, E.; Kalerin, E.; Polyak, F.; Sterling, J. *Biotechnol. Bioeng.* **1992**; *40*; 760.
92. Kirchner, G.; Scollar, M. P.; Klibanov, A. M. *J. Am. Chem. Soc.* **1985**; *107*; 7072.
93. Carrea, G.; Riva, S. *Angew. Chem. Int. Ed. Engl.* **2000**; *39*; 2226.
94. Zaks, A.; Klibanov, A. M. *J. Am. Chem. Soc.* **1986**; *108*; 2767.
95. Wescott, C. R.; Klibanov, A. M. *J. Am. Chem. Soc.* **1993**; *115*; 1629.
96. Gaertner, H.; Puigserver, A. *Eur. J. Biochem.* **1989**; *181*; 207.
97. Ryu, K.; Dordick, J. S. *Biochemistry* **1992**; *31*; 2588.
98. Wescott, C. R.; Noritomi, H.; Klibanov, A. M. *J. Am. Chem. Soc.* **1996**; *118*; 10365.
99. Ke, T.; Wescott, C. R.; Klibanov, A. M. *J. Am. Chem. Soc.* **1996**; *118*; 3366.
100. Carrea, G.; Ottolina, G.; Riva, S. *Trends in Biotechnology* **1995**; *13*; 63.
101. Wescott, C. R.; Klibanov, A. M. *Biochim. Biophys. Acta - Protein Structure and Molecular Enzymology* **1994**; *1206*; 1.
102. Zaks, A.; Klibanov, A. M. *Science* **1984**; *224*; 1249.
103. Volkin, D. B.; Staubli, A.; Langer, R.; Klibanov, A. M. *Biotechnol. Bioeng.* **1991**; *37*; 843.
104. Zaks, A.; Klibanov, A. M. *J. Biol. Chem.* **1988**; *263*; 3194.
105. *Enzymatic Reactions in Organic Media*; Blackie-Pergamon: London, 1996.
106. Reetz, M. T.; D Carballeira, J.; Vogel, A. *Angew. Chem. Int. Ed. Engl.* **2006**; *45*; 7745.
107. Reetz, M. T.; Carballeira, J. D.; Peyralans, J.; Hobenreich, H.; Maichele, A.; Vogel, A. *Chem. Eur. J.* **2006**; *12*; 6031.
108. Reetz, M. T.; Wang, L. W.; Bocola, M. *Angew. Chem. Int. Ed. Engl.* **2006**; *45*; 1236.
109. Reetz, M. T. *Proc. Nat. Acad. Sci. USA* **2004**; *101*; 5716.
110. Reetz, M. T. In *Methods in Molecular Biology*; Humana: Totowa, 2003; Vol. 230, p 259.
111. In *Enzyme Assays: High-throughput Screening, Genetic Selection and Fingerprinting*; Reymond, J. L., Ed.; Wiley-VCH: Weinheim, 2006, p 41.
112. Reetz, M. T.; Zonta, A.; Schimossek, K.; Liebeton, K.; Jaeger, K. E. *Angew. Chem. Int. Ed. Engl.* **1997**; *36*; 2830.
113. Reetz, M. T.; Zonta, A.; Schimossek, K.; Liebeton, K.; Jager, K. 2003; Vol. EP 1,002,100 B1.
114. Reetz, M. T.; Wilensek, S.; Zha, D. X.; Jaeger, K. E. *Angew. Chem. Int. Ed. Engl.* **2001**; *40*; 3589.
115. Reetz, M. T. In *Advances in Catalysis*; Gates, B. C.; Knozinger, H., Eds.; Elsevier: San Diego, 2006; Vol. 44, p 1.
116. Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Taglieber, A.; Hollmann, F.; Mondiere, R. J. G.; Dickmann, N.; Hocker, B.; Cerrone, S.; Haeger, M. C.; Sterner, R. *ChemBioChem* **2008**; *9*; 552.
117. *Directed enzyme evolution: Screening and Selection Methods*; Humana: Totowa, 2003; Vol. 230.
118. Rubin-Pitel, S. B.; Zhao, H. M. *Combinatorial Chemistry & High Throughput Screening* **2006**; *9*; 247.
119. Leung, D. W.; Chen, E.; Goeddel, D. V. *Technique* **1989**; *1*; 11.
120. Cadwell, R. C.; Joyce, G. F. *PCR Methods Appl* **1992**; *2*; 28.
121. Hogrefe, H. H.; Cline, J.; Youngblood, G. L.; Allen, R. M. *Biotechniques* **2002**; *33*; 1158.
122. Casali, N.; Preston, A. In *Methods in Molecular Biology*; Humana Press: Totowa, 2003; Vol. 235, p 209.

123. Stemmer, W. P. C. *Nature* **1994**; *370*; 389.
124. Powell, K. A.; Ramer, S. W.; del Cardayre, S. B.; Stemmer, W. P. C.; Tobin, M. B.; Longchamp, P. F.; Huisman, G. W. *Angew. Chem. Int. Ed. Engl.* **2001**; *40*; 3948.
125. Camps, M.; Naukkarinen, J.; Johnson, B. P.; Loeb, L. A. *Proc. Nat. Acad. Sci. USA* **2003**; *100*; 9727.
126. Vartanian, J. P.; Henry, M.; WainHobson, S. *Nucleic Acids Research* **1996**; *24*; 2627.
127. Chopra, I.; O'Neill, A. J.; Miller, K. *Drug Resistance Updates* **2003**; *6*; 137.
128. Horsman, G. P.; Liu, A. M. F.; Henke, E.; Bornscheuer, U. T.; Kazlauskas, R. J. *Chem. Eur. J.* **2003**; *9*; 1933.
129. Bornscheuer, U. T.; Enzelberger, M. M.; Altenbuchner, J.; Meyer, H. H. *Strategies* **1998**; *11*; 16.
130. Bornscheuer, U. T.; Altenbuchner, J.; Meyer, H. H. *Biotechnol. Bioeng.* **1998**; *58*; 554.
131. Bornscheuer, U. T.; Altenbuchner, J.; Meyer, H. H. *Bioorg. Med. Chem.* **1999**; *7*; 2169.
132. May, O.; Nguyen, P. T.; Arnold, F. H. *Nature Biotechnology* **2000**; *18*; 317.
133. Reetz, M. T.; Bocola, M.; Carballeira, J. D.; Zha, D. X.; Vogel, A. *Angew. Chem. Int. Ed. Engl.* **2005**; *44*; 4192.
134. Reetz, M. T.; Peyralans, J. J. P.; Maichele, A.; Fu, Y.; Maywald, M. *Chem. Commun.* **2006**; 4318.
135. Reetz, M. T.; Carballeira, J. D. *Nature Protocols* **2007**; *2*; 891.
136. Reetz, M. T.; Kahakeaw, D.; Lohmer, R. *ChemBioChem* **2008**; *9*; 1797.
137. Steinreiber, J.; Ward, T. R. *Coord. Chem. Rev.* **2008**; *252*; 751.
138. Qi, D. F.; Tann, C. M.; Haring, D.; Distefano, M. D. *Chem. Rev.* **2001**; *101*; 3081.
139. Wang, L.; Schultz, P. G. *Angew. Chem. Int. Ed. Engl.* **2005**; *44*; 34.
140. Liu, C. C.; Mack, A. V.; Tsao, M. L.; Mills, J. H.; Lee, H. S.; Choe, H.; Farzan, M.; Schultz, P. G.; Smider, V. V. *Proc. Nat. Acad. Sci. USA* **2008**; *105*; 17688.
141. Cellitti, S. E.; Jones, D. H.; Lagpacan, L.; Hao, X.; Zhang, Q.; Hu, H.; Brittain, S. M.; Brinker, A.; Caldwell, J.; Bursulaya, B.; Spraggon, G.; Brock, A.; Ryu, Y.; Uno, T.; Schultz, P. G.; Geierstanger, B. H. *J. Am. Chem. Soc.* **2008**; *130*; 9268.
142. de Graaf, A. J.; Kooijman, M.; Hennink, W. E.; Mastrobattista, E. *Bioconjug. Chem.* **2009**; *20*; 1281.
143. Ai, H. W.; Shen, W.; Brustad, E.; Schultz, P. G. *Angew Chem Int Ed Engl* **2009**; *49*; 935.
144. Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. *Proc. Nat. Acad. Sci. USA* **2002**; *99*; 19.
145. Xie, J. M.; Liu, W. S.; Schultz, P. G. *Angew. Chem. Int. Ed. Engl.* **2007**; *46*; 9239.
146. Weber, P. C.; Wendoloski, J. J.; Pantoliano, M. W.; Salemme, F. R. *J. Am. Chem. Soc.* **1992**; *114*; 3197.
147. Wilchek, M.; Bayer, E. A. *Biomolec. Eng.* **1999**; *16*; 1.
148. Collot, J.; Gradinaru, J.; Humbert, N.; Skander, M.; Zocchi, A.; Ward, T. R. *J. Am. Chem. Soc.* **2003**; *125*; 9030.
149. Gonzalez, M.; Argarana, C. E.; Fidelio, G. D. *Biomolec. Eng.* **1999**; *16*; 67.
150. Green, N. M. *Adv. Protein Chem.* **1975**; *29*; 85.
151. Sano, T.; Pandori, M. W.; Chen, X. M.; Smith, C. L.; Cantor, C. R. *J. Biol. Chem.* **1995**; *270*; 28204.
152. Wilson, M. E.; Whitesides, G. M. *J. Am. Chem. Soc.* **1978**; *100*; 306.
153. Letondor, C.; Pordea, A.; Humbert, N.; Ivanova, A.; Mazurek, S.; Novic, M.; Ward, T. R. *J. Am. Chem. Soc.* **2006**; *128*; 8320.

154. Letondor, C.; Ward, T. R. *ChemBioChem* **2006**; *7*; 1845.
155. Skander, M.; Humbert, N.; Collot, J.; Gradinaru, J.; Klein, G.; Loosli, A.; Sauser, J.; Zocchi, A.; Gilardoni, F.; Ward, T. R. *J. Am. Chem. Soc.* **2004**; *126*; 14411.
156. Klein, G.; Humbert, N.; Gradinaru, J.; Ivanova, A.; Gilardoni, F.; Rusbandi, U. E.; Ward, T. R. *Angew. Chem. Int. Ed. Engl.* **2005**; *44*; 7764.
157. Rusbandi, U. E.; Lo, C.; Skander, M.; Ivanova, A.; Creus, M.; Humbert, N.; Ward, T. R. *Adv. Synth.&Catal.* **2007**; *349*; 1923.
158. Skander, M.; Malan, C.; Ivanova, A.; Ward, T. R. *Chem. Commun.* **2005**; 4815.
159. Rusbandi, U. E.; Skander, M.; Ivanova, A.; Malan, C.; Ward, T. R. *Comptes Rendus Chimie* **2007**; *10*; 678.
160. Letondor, C.; Humbert, N.; Ward, T. R. *Proc. Nat. Acad. Sci. USA* **2005**; *102*; 4683.
161. Creus, M.; Pordea, A.; Rossel, T.; Sardo, A.; Letondor, C.; Ivanova, A.; Le Trong, I.; Stenkamp, R. E.; Ward, T. R. *Angew. Chem. Int. Ed. Engl.* **2008**; *47*; 1400.
162. Thomas, C. M.; Letondor, C.; Humbert, N.; Ward, T. R. *J. Organomet. Chem.* **2005**; *690*; 4488.
163. Pierron, J.; Malan, C.; Creus, M.; Gradinaru, J.; Hafner, I.; Ivanova, A.; Sardo, A.; Ward, T. R. *Angew. Chem. Int. Ed. Engl.* **2008**; *47*; 701.
164. Pordea, A.; Creus, M.; Panek, J.; Duboc, C.; Mathis, D.; Novic, M.; Ward, T. R. *J. Am. Chem. Soc.* **2008**; *130*; 8085.
165. Pordea, A.; Mathis, D.; Ward, T. R. *J. Organomet. Chem.* **2009**; *694*; 930.
166. Petrounia, I. P.; Arnold, F. H. *Curr. Opin. Biotechnol.* **2000**; *11*; 325.
167. Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Maywald, M. *Chimia* **2002**; *56*; 721.
168. Reetz, M. T. *Tetrahedron* **2002**; *58*; 6595.
169. Pordea, A.; Ward, T. R. *Chem. Commun.* **2008**; 4239.
170. Schmidt, T. G. M.; Skerra, A. *Prot. Eng.* **1993**; *6*; 109.
171. Dixon, R. W.; Radmer, R. J.; Kuhn, B.; Kollman, P. A.; Yang, J.; Raposo, C.; Wilcox, C. S.; Klumb, L. A.; Stayton, P. S.; Behnke, C.; Le Trong, I.; Stenkamp, R. *J. Org. Chem.* **2002**; *67*; 1827.
172. Peters, T. *All about Albumin: Biochemistry, Genetics and Medical Applications* Academic Press, San Diego, California, 1996.
173. Colonna, S.; Gaggero, N. *Tetrahedron Lett.* **1989**; *30*; 6233.
174. Ogura, K.; Fujita, M.; Iida, H. *Tetrahedron Lett.* **1980**; *21*; 2233.
175. Sugimoto, T.; Kokubo, T.; Miyazaki, J.; Tanimoto, S.; Okano, M. *J. Chem. Soc. - Chem. Commun.* **1979**; 402.
176. Sugimoto, T.; Kokubo, T.; Miyazaki, J.; Tanimoto, S.; Okano, M. *Bioorg. Chem.* **1981**; *10*; 311.
177. Colonna, S.; Banfi, S.; Fontana, F.; Sommaruga, M. *J. Org. Chem.* **1985**; *50*; 769.
178. Colonna, S.; Manfredi, A. *Tetrahedron Lett.* **1986**; *27*; 387.
179. Colonna, S.; Manfredi, A.; Annunziata, R.; Spadoni, M. *Tetrahedron* **1987**; *43*; 2157.
180. Colonna, S.; Gaggero, N.; Manfredi, A.; Spadoni, M.; Casella, L.; Carrea, G.; Pasta, P. *Tetrahedron* **1988**; *44*; 5169.
181. Sugimoto, T.; Matsumura, Y.; Tanimoto, S.; Okano, M. *J. Chem. Soc. - Chem. Commun.* **1978**; 926.
182. Baba, N.; Matsumura, Y.; Sugimoto, T. *Tetrahedron Lett.* **1978**; 4281.

183. Colonna, S.; Manfredi, A.; Annunziata, R. *Tetrahedron Lett.* **1988**; *29*; 3347.
184. Ohashi, M.; Koshiyama, T.; Ueno, T.; Yanase, M.; Fujii, H.; Watanabe, Y. *Angew. Chem. Int. Ed. Engl.* **2003**; *42*; 1005.
185. Rousselot-Pailley, P.; Bochot, C.; Marchi-Delapierre, C.; Jorge-Robin, A.; Martin, L.; Fontecilla-Camps, J. C.; Cavazza, C.; Menage, S. *ChemBioChem* **2009**; *10*; 545.
186. Pietikainen, P. *Tetrahedron* **1998**; *54*; 4319.
187. *Enzyme Catalysis in Organic Synthesis*; 2nd ed.; Wiley - VCH Weinheim, 2002.
188. Shokat, K. M.; Leumann, C. J.; Sugasawara, R.; Schultz, P. G. *Angew. Chem. Int. Ed. Engl.* **1988**; *27*; 1172.
189. Nimri, S.; Keinan, E. *J. Am. Chem. Soc.* **1999**; *121*; 8978.
190. Ricoux, R.; Sauriat-Dorizon, H.; Girgenti, E.; Blanchard, D.; Mahy, J. P. *J. Immunol. Methods* **2002**; *269*; 39.
191. Ricoux, R.; Lukowska, E.; Pezzotti, F.; Mahy, J. P. *Eur. J. Biochem.* **2004**; *271*; 1277.
192. Yamaguchi, H.; Hirano, T.; Kiminami, H.; Taura, D.; Harada, A. *Org. Biomol. Chem.* **2006**; *4*; 3571.
193. Yamamura, K.; Kaiser, E. T. *J. Chem. Soc. - Chem. Commun.* **1976**; 830.
194. Trynda-Lemiesz, L.; Pruchnik, F. P. *J. Inorg. Biochem.* **1997**; *66*; 187.
195. Trynda, L.; Pruchnik, F. *J. Inorg. Biochem.* **1995**; *58*; 69.
196. Jing, Q.; Okrasa, K.; Kazlauskas, R. J. *Chem. Eur. J.* **2009**; *15*; 1370.
197. Esposito, B. P.; Faljoni-Alario, A.; de Menezes, J. F. S.; de Brito, H. F.; Najjar, R. *J. Inorg. Biochem.* **1999**; *75*; 55.
198. Esposito, B. P.; de Oliveira, E.; Zyngier, S. B.; Najjar, R. *J. Brazil. Chem. Soc.* **2000**; *11*; 447.
199. Bertucci, C.; Botteghi, C.; Giunta, D.; Marchetti, M.; Paganelli, S. *Adv. Synth.&Catal.* **2002**; *344*; 556.
200. Okrasa, K.; Kazlauskas, R. J. *Chem. Eur. J.* **2006**; *12*; 1587.
201. Fernandez-Gacio, A.; Codina, A.; Fastrez, J.; Riant, O.; Soumillion, P. *ChemBioChem* **2006**; *7*; 1013.
202. Bakker, M.; van Rantwijk, F.; Sheldon, R. A. *Canad. J. Chem.-Revue Canadienne De Chimie* **2002**; *80*; 622.
203. Hemrika, W.; Renirie, R.; Dekker, H. L.; Barnett, P.; Wever, R. *Proc. Nat. Acad. Sci. USA* **1997**; *94*; 2145.
204. van de Velde, F.; Konemann, L.; van Rantwijk, F.; Sheldon, R. A. *Chem. Commun.* **1998**; 1891.
205. van de Velde, F.; Arends, I. W. C. E.; Sheldon, R. A. *J. Inorg. Biochem.* **2000**; *80*; 81.
206. van de Velde, F.; Konemann, L.; van Rantwijk, F.; Sheldon, R. A. *Biotechnol. Bioeng.* **2000**; *67*; 87.
207. Aitken, D. J.; Alijah, R.; Onyiriuka, S. O.; Suckling, C. J.; Wood, H. C. S.; Zhu, L. M. *J. Chem. Soc., Perkin Trans.* **1993**; 597.
208. Kokubo, T.; Sugimoto, T.; Uchida, T.; Tanimoto, S.; Okano, M. *J. Chem. Soc. - Chem. Commun.* **1983**; 769.
209. Mahammed, A.; Gray, H. B.; Weaver, J. J.; Sorasaenee, K.; Gross, Z. *Bioconjug. Chem.* **2004**; *15*; 738.
210. Gantchev, T. G.; Ouellet, R.; van Lier, J. E. *Arch. Biochem. Biophys.* **1999**; *366*; 21.

211. Chatterjee, S.; Srivastava, T. S. *Journal of Porphyrins and Phthalocyanines* **2000**; *4*; 147.
212. Ding, Y.; Lin, B.; Huie, C. W. *Electrophoresis* **2001**; *22*; 2210.
213. Kragh-Hansen, U.; Chuang, V. T. G.; Otagiri, M. *Biol. Pharm. Bull.* **2002**; *25*; 695.
214. Zunszain, P. A.; Ghuman, J.; Komatsu, T.; Tsuchida, E.; Curry, S. *BMC Struct. Biol.* **2003**; *3*; 1.
215. Mahammed, A.; Gross, Z. *J. Am. Chem. Soc.* **2005**; *127*; 2883.
216. Meunier, B. *Chem. Rev.* **1992**; *92*; 1411.
217. Reetz, M. T.; Jiao, N. *Angew. Chem. Int. Ed. Engl.* **2006**; *45*; 2416.
218. *Handbook of Metalloproteins*; John Wiley & Sons: Chichester, 2001; Vol. 1.
219. Neya, S.; Funasaki, N.; Imai, K. *J. Biol. Chem.* **1988**; *263*; 8810.
220. Hayashi, T.; Hisaeda, Y. *Acc. Chem. Res.* **2002**; *35*; 35.
221. Hamachi, I.; Shinkai, S. *Eur. J. Org. Chem.* **1999**; 539.
222. Tamura, M.; Asakura, T.; Yonetani, T. *Biochim. Biophys. Acta* **1973**; *295*; 467.
223. Ueno, T.; Koshiyama, T.; Abe, S.; Yokoi, N.; Ohashi, M.; Nakajima, H.; Watanabe, Y. *J. Organomet. Chem.* **2007**; *692*; 142.
224. Ueno, T.; Koshiyama, T.; Ohashi, M.; Kondo, K.; Kono, M.; Suzuki, A.; Yamane, T.; Watanabe, Y. *J. Am. Chem. Soc.* **2005**; *127*; 6556.
225. Carey, J. R.; Ma, S. K.; Pfister, T. D.; Garner, D. K.; Kim, H. K.; Abramite, J. A.; Wang, Z. L.; Guo, Z. J.; Lu, Y. *J. Am. Chem. Soc.* **2004**; *126*; 10812.
226. Kruithof, C. A.; Casado, M. A.; Guillena, G.; Egmond, M. R.; van der Kerk-van Hoof, A.; Heck, A. J. R.; Gebbink, R. J. M. K.; van Koten, G. *Chem. Eur. J.* **2005**; *11*; 6869.
227. Kruithof, C. A.; Dijkstra, H. P.; Lutz, M.; Spek, A. L.; Egmond, M. R.; Gebbink, R. J. M. K.; van Koten, G. *Eur. J. Inorg. Chem.* **2008**; 4425.
228. Rutten, L.; Wieczorek, B.; Mannie, J. P. B. A.; Kruithof, C. A.; Dijkstra, H. P.; Egmond, M. R.; Lutz, M.; Klein Gebbink, R. J. M.; Gros, P.; van Koten, G. *Chem. Eur. J.* **2009**; *15*; 4270.
229. Salmain, M.; Jaouen, G. *Comptes Rendus Chimie* **2003**; *6*; 249.
230. Salmain, M. In *Bioorganometallics*; Jaouen, G., Ed.; Wiley-VCH: Weinheim, 2005, p 181.
231. Nicholas, K. M.; Wentworth, P.; Harwig, C. W.; Wentworth, A. D.; Shafton, A.; Janda, K. D. *Proc. Nat. Acad. Sci. USA* **2002**; *99*; 2648.
232. Rokita, S. E.; Kaiser, E. T. *J. Am. Chem. Soc.* **1986**; *108*; 4984.
233. *Bioconjugation: Protein Coupling Techniques for the Biomedical Sciences*; Macmillan London, 1998.
234. Hermanson, G. T. *Bioconjugate techniques*; Academic Press: San Diego, 1996.
235. Levine, H. L.; Nakagawa, Y.; Kaiser, E. T. *Biochem. Biophys. Res. Commun.* **1977**; *76*; 64.
236. Levine, H. L.; Kaiser, E. T. *J. Am. Chem. Soc.* **1978**; *100*; 7670.
237. Levine, H. L.; Kaiser, E. T. *J. Am. Chem. Soc.* **1980**; *102*; 343.
238. Slams, J. T.; Radziejewski, C.; Oruganti, S. R.; Kaiser, E. T. *J. Am. Chem. Soc.* **1984**; *106*; 6778.
239. Fried, H. E.; Kaiser, E. T. *J. Am. Chem. Soc.* **1981**; *103*; 182.
240. Stewart, K. D.; Radziejewski, C.; Kaiser, E. T. *J. Am. Chem. Soc.* **1986**; *108*; 3480.
241. Kokubo, T.; Sassa, S.; Kaiser, E. T. *J. Am. Chem. Soc.* **1987**; *109*; 606.
242. Hilvert, D.; Kaiser, E. T. *J. Am. Chem. Soc.* **1985**; *107*; 5805.

243. Hilvert, D.; Hatanaka, Y.; Kaiser, E. T. *J. Am. Chem. Soc.* **1988**; *110*; 682.
244. Suckling, C. J.; Zhu, L.-M. *Bioorg. Med. Chem. Lett.* **1993**; *3*; 531.
245. Panella, L.; Broos, J.; Jin, J. F.; Fraaije, M. W.; Janssen, D. B.; Jeronimus-Stratingh, M.; Feringa, B. L.; Minnaard, A. J.; de Vries, J. G. *Chem. Commun.* **2005**; 5656.
246. Maywald, M.; Ruhr-Universität Bochum; 2005.
247. Pletsch, A.; Ruhr-Universität Bochum; 2004.
248. Haquette, P.; Salmain, M.; Svedlung, K.; Martel, A.; Rudolf, B.; Zakrzewski, J.; Cordier, S.; Roisnel, T.; Fosse, C.; Jaouen, G. *ChemBioChem* **2007**; *8*; 224.
249. Kuang, H.; Brown, M. L.; Davies, R. R.; Young, E. C.; Distefano, M. D. *J. Am. Chem. Soc.* **1996**; *118*; 10702.
250. Kuang, H.; Davies, R. R.; Distefano, M. D. *Bioorg. Med. Chem. Lett.* **1997**; *7*; 2055.
251. Davies, R. R.; Kuang, H.; Qi, D. F.; Mazhary, A.; Mayaan, E.; Distefano, M. D. *Bioorg. Med. Chem. Lett.* **1999**; *9*; 79.
252. Haring, D.; Lees, M. R.; Banaszak, L. J.; Distefano, M. D. *Prot. Eng.* **2002**; *15*; 603.
253. Davies, R. R.; Distefano, M. D. *J. Am. Chem. Soc.* **1997**; *119*; 11643.
254. Ory, J. J.; Mazhary, A.; Kuang, H.; Davies, R. R.; Distefano, M. D.; Banaszak, L. J. *Prot. Eng.* **1998**; *11*; 253.
255. Polgar, L.; Bender, M. L. *J. Am. Chem. Soc.* **1966**; *88*; 3153.
256. Polgar L., M. L. B. *Biochemistry* **1967** *6*; 610.
257. Wu, Z. P.; Hilvert, D. *J. Am. Chem. Soc.* **1989**; *111*; 4513.
258. Wu, Z. P.; Hilvert, D. *J. Am. Chem. Soc.* **1990**; *112*; 5647.
259. Bell, I. M.; Hilvert, D. *Biochemistry* **1993**; *32*; 13969.
260. Haring, D.; Schuler, E.; Adam, W.; Saha-Moller, C. R.; Schreier, P. *J. Org. Chem.* **1999**; *64*; 832.
261. Haring, D.; Schreier, P. *Angew. Chem. Int. Ed. Engl.* **1998**; *37*; 2471.
262. Roelfes, G.; Feringa, B. L. *Angew. Chem. Int. Ed. Engl.* **2005**; *44*; 3230.
263. Roelfes, G.; Boersma, A. J.; Feringa, B. L. *Chem. Commun.* **2006**; 635.
264. Coquiere, D.; Feringa, B. L.; Roelfes, G. *Angew. Chem. Int. Ed. Engl.* **2007**; *46*; 9308.
265. Boersma, A. J.; Klijn, J. E.; Feringa, B. L.; Roelfes, G. *J. Am. Chem. Soc.* **2008**; *130*; 11783.
266. Caprioara, M.; Fiammengo, R.; Engeser, M.; Jaschke, A. *Chemistry* **2007**; *13*; 2089.
267. Nuzzolo, M.; Grabulosa, A.; Slawin, A. M.; Meeuwenoord, N. J.; Van der Marel, G. A.; Kamer, P. C. *Eur. J. Org. Chem.* **2010**; *in press*.
268. Ropartz, L.; Meeuwenoord, N. J.; van der Marel, G. A.; van Leeuwen, P. W.; Slawin, A. M.; Kamer, P. C. *Chem. Commun. (Camb)* **2007**; 1556.
269. Oltra, N. S.; Roelfes, G. *Chem. Commun. (Camb)* **2008**; 6039.
270. Laan, W.; Munoz, B. K.; Den Heeten, R.; Kamer, P. C. *ChemBioChem* **2010**; *in press*.
271. Deuss, P. J.; Popa, G.; Botting, C. H.; Laan, W.; Kamer, P. C. J. *Angew. Chem. Int. Ed. Engl.* **2010**; *in press*.

II. Site-selective bioconjugation of phosphane ligands to structurally diverse protein hosts

Abstract

A method for the site-selective introduction of phosphane-ligands in protein hosts has been developed. Initially, carboxylic acid functionalized phosphanes were coupled to proteins, but the chemoselectivity proved to be low for most proteins. Borane protected maleimido-phosphanes were synthesised and coupled site-selective to proteins. Unexpectedly, borane deprotection was not achievable under conditions compatible with proteins in solution. An alternative approach reported here consists of covalent coupling of 3-(maleimido)propionic acid hydrazide to the thiol group of a cysteine residue of a protein, followed by coupling of aldehyde functionalized phosphanes to the hydrazide group via a hydrazone linkage. This method constitutes a valuable tool for the development of phosphane-containing artificial metalloenzymes.

II.1 Introduction

Natural enzymes have evolved over millions of years to catalyze biologically relevant chemical reactions under mild reaction conditions with excellent rates and enantio-, regio-, and chemoselectivity. These properties render enzymes attractive synthetic tools as biocatalyst for the chemical industry. Moreover, the powerful methods of directed evolution allow for significant improvements of the useful properties of enzymes.[1-2] However, for many synthetically important, transition-metal catalyzed transformations, *e.g.*, hydrogenation, allylic substitution and hydroformylation, enzymes do not exist. Therefore these reactions are still confined to the realm of homogeneous catalysis, relying on man-made organometallic catalysts. Hybrid catalysts, consisting of a protein scaffold functionalized with a synthetic transition metal catalyst, have emerged as a potential solution to this limitation. This approach aims to combine the catalytic activity of transition-metal complexes with the significant potential for selectivity-inducing interactions between proteins and substrates. Inspired by the pioneering work of Kaiser[3] and Whitesides,[4] several groups have developed artificial metalloenzymes, either by non-covalent or covalent incorporation of catalytic moieties within a macromolecular host. A particularly successful example of the non-covalent approach is the work of Whitesides, Chan and particularly Ward, based on the well-developed biotin-avidin technology. Chemical modification of biotin with phosphorous ligands allows the introduction of transition metal catalysts in (strept)avidin. Subsequent tuning of both the ligand structure and the protein sequence has resulted in highly selective catalysts for the rhodium catalyzed asymmetric hydrogenation of amino-acid precursors,[5-11] transfer hydrogenation[12-15] and more recently, palladium catalyzed asymmetric C-C bond formation.[16] Moreover, Reetz subjected this system to directed evolution, demonstrating that this method can also be applied to improve the performance of hybrid catalysts.[17-18] Other examples of the non-covalent assembly of hybrid catalysts include antibodies raised against transition metal complexes,[19-20] and the reconstitution of apo-myoglobin with Cr^{III} Schiff base complexes.[21]

The covalent approach relies on the chemical modification of reactive residues in proteins. Distefano modified adipocyte lipid binding protein with a Cu^{II} 1,10-phenantroline complex by alkylation of a unique cysteine residue. The resulting

conjugate catalyzed the enantioselective hydrolysis of amino-acid esters.[22] The cysteine of papain has been modified with maleimide modified manganese-salen and rhodium-dipyridine complexes; the resulting hybrids were active as epoxidation and hydrogenation catalysts, but showed no enantioselectivity.[18] A hybrid of papain and a rodium-phosphite complex was also active as a hydrogenation catalyst, but again no stereocontrol was observed.[23] The strategy of modifying the active site serine of lipases has been employed to introduce palladium and platinum complexes of nitrogen- and sulfur containing pincer ligands,[24] and a free bidentate diphosphane ligand.[25-26]

Phosphane ligands are ubiquitous in transition metal chemistry and afford extremely reactive and versatile homogeneous catalysts. Several hybrid catalysts containing phosphane-ligands have been created with the non-covalent approach, using the biotin-avidin system[4] and antibodies.[20] To the best of my knowledge the only example of covalent, site-selective introduction of a phosphane-ligand is the aforementioned modification of the active site serine of a lipase with a diphosphane coupled to a phosphonate inhibitor.[25-26] However, the resulting linkage was hydrolytically unstable. A general method for stable, covalent, site-selective introduction of phosphane-ligands is desirable, as this would allow the functionalization of structurally diverse proteins with this important class of ligands. Because of the nucleophilic reactivity of its thiol side-chain and relative low abundance in protein sequences, cysteine is a widely employed target for the site-selective modification of proteins. Frequently used methods for the functionalization of cysteine are disulfide bridge formation using thiols, and using electrophilic moieties such as haloacetamides and maleimides.[27] All three methods have been employed in the synthesis of ligand modified proteins, e.g. Reetz introduced a small library of nitrogen based chelating ligands and organocatalysts in a protein host using maleimides.[28]

Because phosphanes are known to reduce disulfide bridges (*e.g.* tris(2-carboxyethyl)phosphane is a well-known reducing agent), modification based on the formation of a disulfide bridge is incompatible with phosphane ligands. The latter two methods are incompatible with unprotected phosphanes due to their nucleophilicity.

Because of the limitations of using the standard methods for cysteine modification outlined above there is a need for alternative coupling methods for the introduction

of phosphanes and transition metal phosphane complexes into proteins. Herein I report on our exploration of different strategies, which led us to the development of a highly efficient and selective method for site-selective covalent introduction of unprotected phosphane ligands in different protein hosts.

II.2 Results and discussion

II.2.1 Engineering of protein hosts

Selectivity can be imparted on catalytic reactions by placing an achiral catalyst in the constrained environment of a chiral cavity. Cavities that completely encapsulate the catalyst will afford the highest potential for selectivity imparting interactions, but may conversely impede access of substrates to the catalysts.

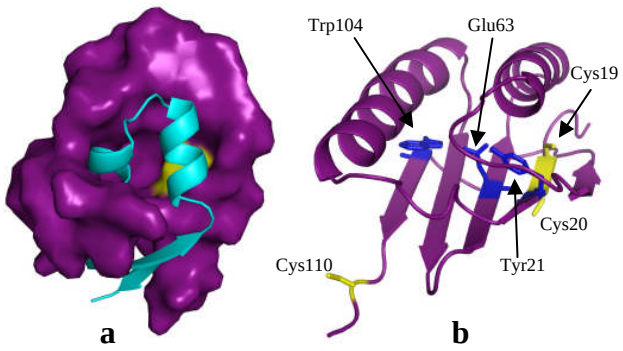


Figure 1 Structures of protein hosts used in this study.
(a) X-ray structure of I-FABP. The structural elements deleted in IFABP Δ98Δ are highlighted in cyan, the position of the introduced cysteine is highlighted in yellow. ALBP has the same fold as I-FABP. **(b)** NMR structure of the BLUF domain of AppA.[29] The mutated cysteines are highlighted in yellow; the residues mutated to cysteines for modification are highlighted in blue. Residues 113-125 are omitted for clarity.

Therefore proteins with different cavity structures, as hosts for artificial metalloenzyme development, were used: (i) a protein containing a virtually completely enclosed cavity; (ii) a protein containing a cavity with a large opening, and (iii) a protein containing a deep cleft. For ease of purification of the proteins I choose nickel-affinity chromatography as the purification procedure, using proteins containing a hexa histidine-tag.

Fatty Acid Binding Proteins (FABPs) are a class of proteins that possess a large apolar cavity. The cavity is formed by a β-barrel structure, capped by a α-helical lid, and completely encapsulates bound lipids. Thus FABPs contain a true protein cavity.

Two FABPs were used as hosts for modification with phosphane ligands: murine adipocyte lipid binding protein (ALBP) and an engineered variant of rat intestinal fatty acid binding protein (I-FABP), Δ98Δ. Both ALBP and I-FABP have previously been shown to be suitable scaffolds for the development of hybrid catalyst by chemical modification with catalytic cofactors,[22,30-32] but have so far not been modified with phosphane-ligands.

Wild-type ALBP contains two native cysteines, at positions 1 and 117. Cys117 is located inside the protein cavity and thus at a suitable position for modification. Cys1, located on the surface of the protein, was replaced by alanine. Initially, ALBP Cys1Ala containing an N-terminal histidine-tag was produced using a reported expression system.[30] However, the low yield (1-2 mg/l) in our hands prompted us to investigate alternative expression conditions. Cloning of the DNA coding for ALBP Cys1Ala into the vector pQE30 and expression in *E. coli* M15 pREP4 yielded a satisfactory 80 mg of N-terminally his-tagged protein per litre of bacterial culture.

IFABP Δ98Δ is a genetically engineered all β-sheet variant of IFABP, which lacks the helices covering the cavity in the parent protein (Figure 1a). The protein has a stable, well folded structure and has retained the ability to bind fatty acids. Deletion of the helical cap has resulted in a cavity with a large opening, and thus easier access for ligands and substrates compared to ALBP.[33] IFABP Δ98Δ contains no native cysteine; therefore the IFABP Δ98Δ Tyr90Cys mutant was created, which contains a cysteine at a position structurally equivalent to ALBP Cys117. In addition, a stop-codon, present in the initial expression construct between the IFABP gene and the vector encoded His-tag was replaced by alanine, resulting in expression of the protein with a C-terminal His-tag (hereafter referred to as Δ98Δ).

The BLUF domain of the blue-light sensing flavoprotein AppA binds a flavin non-covalently in a 12 Å deep cleft formed by two α-helices lying on top of a β-sheet (Figure 1b).[34-36] I reasoned that if the flavin could be removed, anchoring catalysts to the bottom of this cleft would ensure they are in close proximity to the chiral environment of the protein. The BLUF domain used in this study (residues 5-125, hereafter referred to as AppA) contains three native cysteines: Cys19, located buried under a helix; Cys20, located on the solvent exposed surface of the β-sheet and Cys110, located in a loop outside the flavin-binding region. Based on sequence comparison with other BLUF domains, I exchanged these cysteines for leucine,

threonine and alanine, respectively.[37] To introduce a cysteine at the bottom of the cleft, the residues Tyr21, Glu63 and Trp104 were selected to be replaced independently by cysteine, based on their location inside the protein cavity. In addition, replacement of these amino acid residues is expected to disrupt the hydrogen bonding interaction of the protein with the flavin, facilitating its removal.[36] Three mutant proteins: AppaA Y21C, AppA Q63C and AppA W104C, suitable for chemical modification, resulted after cysteine introduction. All mutants were purified with flavin bound, suggesting the mutations did not result in significant structural changes**.** Coupling reactions to the cysteine should therefore be performed under conditions which also result in flavin removal.

II.2.2 Conjugation of phosphanes to protein hosts

Our group reported recently a method for the site-selective conjugation of phosphanes to the cysteine of PYP, an additional protein scaffold we employ in our research program to develop artificial metalloenzymes.[38] Here I explored the suitability of this method for the site-selective conjugation of phosphane ligands to the other protein hosts.

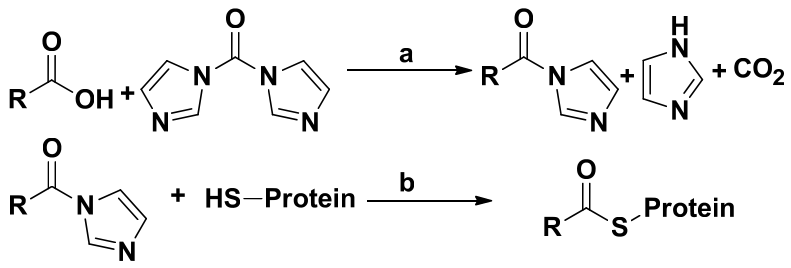


Figure 2 Site-selective protein modification using activated carboxylic acids. (a) DMF, o.n. (b) 50 mM Tris-HCl/H_2O, pH 8.0

The conjugates were analysed by mass spectrometry (MS) and the position of the modification was confirmed by tryptic digestion. Due to their nucleophilicity, free phosphanes react similarly with reagents used for the colorimetric detection of thiol groups as free thiols (e.g. Ellman's reagent). This prohibited the use of such assays to determine the specificity and efficiency of the coupling reactions. Instead, I relied on MS to determine the extent of modification. To correct the MS data for differences in ionization efficiency between unmodified and modified proteins, this difference was assessed by comparing the MS spectra of samples of modified protein with and

without the addition of 1 equivalent of unmodified protein. It should be noted that although the MS measurements were done in air, very little oxidation of the phosphane ligands occurred.

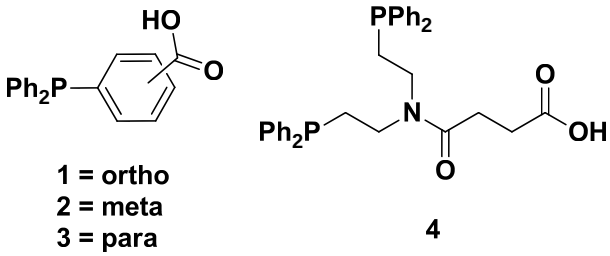


Figure 3 Structures of phosphane-carboxylic acids **1-4**

ALBP was efficiently modified with ligand **3**, but no coupling was observed for ligands **1, 2** and **4**. The latter ligands are possibly sterically more demanding than ligand **3** and may be unable to enter the cavity to conjugate to the cysteine. According to LC-MS analysis and considering the difference in ionization between the modified and unmodified protein, the coupling efficiency of **3** to ALBP is higher than 95%.

Purified Δ98Δ readily forms dimers via disulfide bridge formation, necessitating a reduction step prior to the conjugation reactions.[39] In contrast to ALBP, Δ98Δ could be modified with all four ligands. This may reflect the fact that the latter protein contains a more accessible cavity, and hence a more accessible cysteine. However, the coupling efficiency was limited to 30-50% for single modification; attempts to improve the modification levels by increasing the excess of ligand in the coupling reactions resulted in the coupling of multiple ligands to the protein. The additional sites of modification were identified as lysines, by trypsin digestion and MS analysis. To prevent modification of lysines during the conjugation reactions, I sought to protect their amine groups via amide or imine bond formation.[40-44] Reaction of the protein with aldehydes, anhydrides or succinimide derivatives prior to coupling of the phosphanes afforded some protection, but did not lead to a sufficient increase in selectivity for phosphane-coupling. Also the replacement of the most reactive lysines (K74 and K104, identified by tryptic digest analysis) with glutamine did not yield a significant improvement in chemoselective modification of the protein.

It was previously reported that the presence of high concentrations of imidazole

(more than 2M) appears to cause the release of the flavin from AppA.[45] Therefore to facilitate the flavin removal during the conjugation reactions, phosphanes **1-4** were coupled to AppA Y21C in presence of 3M imidazole. Again, it was not possible to achieve efficient monomodification. Attempts to increase the coupling efficiency by using a larger excess of ligands **1-4** led to unspecific incorporation of ligands. So the chemoselectivity of bioconjugation of phosphanes via thioester formation is highly dependent on the protein employed, prohibiting its implementation as a general method for coupling of phosphanes to the cysteine of proteins.

Table 1 Coupling of activated carboxylic acids to protein hosts

	Calculated Mass (Da)	**LC-MS Mass found (Da)**
ALBP - **3**[a]	16893.1	16893.0±0.5
Δ98Δ-**1**[a]	13711.3	13711.2±0.9
Δ98Δ-**2**[a]	13711.3	13711.5±0.7
Δ98Δ-**3**[a]	13711.3	13710.4±0.5
Δ98Δ-**4**[a]	13946.2	13945.7±0.4

[a] Incomplete conversion

In search of a more chemoselective and efficient method I turned to the coupling of phosphanes using maleimides, a standard strategy for highly selective cysteine conjugation.[46-47] However, since phosphanes are known to react with maleimides leading to the formation of phosphonium-salts,[48-49] the use of phosphane-protecting groups is essential.

Initially I investigated a sulfur protected phosphane, diphenylphosphane sulfide benzyl maleimide (Figure 4), which was incorporated with excellent chemoselectivity into ALBP and Δ98Δ. However, the standard Raney nickel procedure for sulfur deprotection was not successful for the phosphane coupled to protein, based on ^{31}P NMR and MS data.[50-51] No further methods for sulfur deprotection were tested.

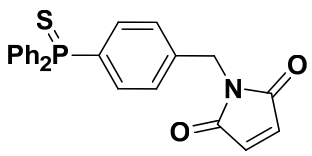


Figure 4 N-(4-(diphenylphosphane sulfide)benzyl) maleimide

A recent report on the synthesis of functionalized phosphanes using Cu(I)-catalyzed azide-alkyne "click" cycloaddition, employing borane-protected phosphanes,

prompted us to explore the synthesis of maleimide-functionalized phosphane-ligands via click-chemistry.[52]

A number of alkyne and azide functionalized "building blocks" for the synthesis of maleimide-functionalized phosphane ligands were synthesized (Figure 5).

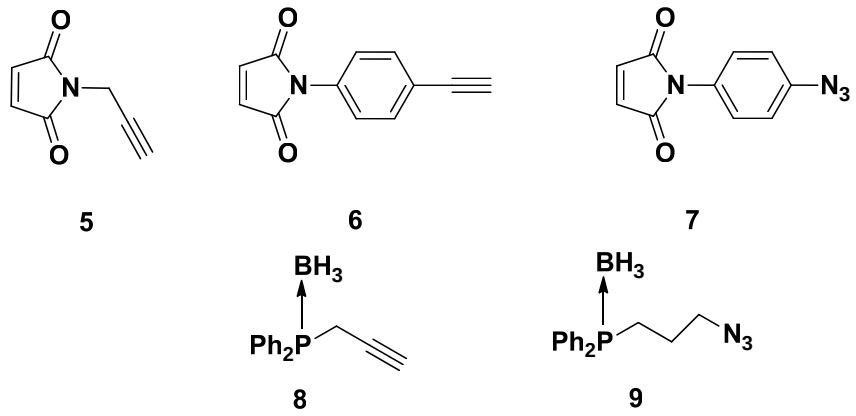


Figure 5 Structures of azide and alkyne functionalized "building blocks" for the synthesis of borane-protected maleimido-phosphanes

Using compounds **5-9** as building blocks, two strategies were available: (I) coupling of maleimide alkynes and azide **5-7** to proteins first, followed by incorporation of borane phosphanes via azide-alkyne click-reaction or (II) synthesis of maleimide functionalized borane phosphanes via click-reaction followed by their covalent incorporation into proteins and subsequent deprotection.

The maleimide functionalized alkynes **5, 6** and azide **7** were successfully incorporated in three proteins: ALBP, Δ98Δ and AppA Y21C. Moreover, Michael addition to maleimide functionalized ligands was highly chemoselective. Complete selective modification with **5-7** was achieved for AppA Y21C and Δ98Δ. The incomplete modification (≈70%) of ALBP could be caused by the reduced accessibility of the thiol group of cysteine located deep inside the enclosed cavity of the protein (Table 2).

The maleimide derivatives **5-9** were coupled to AppA Y21C in presence of 3M imidazole, to help the flavin removal and also to improve the coupling efficiency. Surprisingly, imidazole proved also to reduce unspecific coupling of maleimides. This phenomenon initially observed for AppA, also applied to Δ98Δ, but ALBP was not stable in presence of 3M imidazole. When a large molar excess of maleimide derivatives was used (25x or 50x), selective cysteine modification was only observed when the reactions were performed in presence of 3M imidazole. Under

similar conditions but without imidazole, multiple couplings were observed for AppA Y21C and Δ98Δ. The mechanism underlying this apparent protection against lysine-modification is unclear at present.

An increase of the masses of proteins modified with ligands **5-7** with 18 Da and/or 16 Da was observed in time. The increase of 18 Da is likely due to hydrolytic opening of the maleimide ring, affording maleamic acid adducts. This process is known to be accelerated at basic pH.[53] In addition the oxidation of methionine residues can occur under Mass Spec conditions resulting in addition of 16 or 32 Da.

Table 2 Coupling of alkyne and azide maleimides **5-7** to proteins

	Calculated Mass (Da)	**LC-MS Mass found (Da)**
ALBP-**5**[a]	16739.9	16757.0±0.7
ALBP-**6**[a]	16801.9	16819.2±0.4
ALBP-**7**[a]	16818.9	16836.4±0.4
Δ98Δ-**5**	13558.1	13575.5±0.5
Δ98Δ-**6**	13620.2	13638.5±0.4
Δ98Δ-**7**	13637.2	13654.9±0.6
AppA Y21C -**5**	15530.2	15530.2±1.8
AppA Y21C -**6**	15592.3	15592.2±1.7
AppA Y21C -**7**	15609.3	15609.2±1.7

[a]Incomplete conversion

After coupling of alkyne/azide functionalized maleimides **5-7** to proteins, the coupling of borane-protected phosphanes **8** and **9** to the modified proteins via copper-catalyzed azide-alkyne [3+2] "click" cycloaddition was attempted. Although the click-reaction has previously been applied for bioorthogonal labelling of proteins,[54-55] our attempts to couple borane protected phosphanes to azide/alkyne functionalized proteins via click-reaction proved unsuccessful. When the reactions were performed in the absence of a ligand for the copper catalyst, no protein signals were obtained during subsequent LC-MS analysis. Residual copper, possibly coordinated to the protein, likely suppresses protein ionization. Extensive washing procedures did not resolve this issue. Protein signals were restored when ligands commonly used for coordination of copper during the click-reaction were employed,[55-56] but no protein modification was observed. This is likely due to the fact that the azide or alkyne functionalities are located inside the protein cavities, making them inaccessible for the copper catalysts. Indeed, the vast majority of reports on bioconjugation using click-chemistry deal with labelling protein surfaces

with fluorescent probes, involving reactions with limited steric demand and which are typically solely analysed by fluorescence.[54-55,57]

Three maleimide functionalized borane phosphanes **10**-**12** were synthesized via azide-alkyne click-reaction, using compounds **5**-**9** as building blocks.

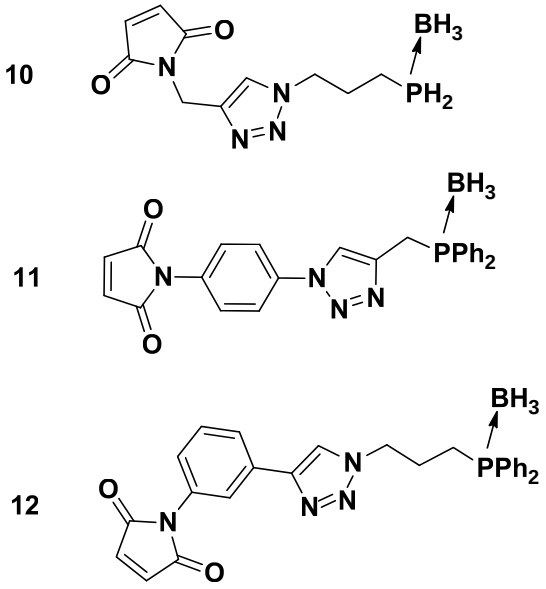


Figure 6 Maleimide functionalized borane-phosphanes synthesized via copper catalysed azide-alkyne [3+2] "click" cycloaddition

All three click-products **10**-**12** were coupled chemo-specifically to proteins (Table 3). Once more, an addition of 18 and 16 Da to the masses of the modified proteins was observed in time.

Table 3 Coupling of click-compounds **10**-**12** to proteins

	Calculated Mass (Da)	**LC-MS Mass found (Da)**
ALBP-**10**	17023.0	17022.2±0.9
ALBP-**11**[a]	17057.0	17074.5±0.8
ALBP-**12**[a]	17085.1	17103.0±1.1
Δ98Δ-**10**	13841.2	13840.1±0.5
Δ98Δ-**11**	13875.2	13893.0±0.6
Δ98Δ-**12**[a]	13903.3	13903.1±0.6
AppA Y21C -**10**	15813.3	15812.3±0.9
AppA Y21C -**11**	15847.3	15864.9±0.6
AppA Y21C -**12**	15875.4	15893.1±0.8

[a]Incomplete conversion

Several reagents reported in literature for decomplexation of phosphane-borane complexes (amines, KCN, CO)[58-62] were tested for their ability to deprotect the bioconjugated phosphane-boranes in aqueous conditions. The decomplexation

conditions screened were only those deemed compatible with proteins in solution. For all conditions tested removal of the BH_3 protecting group from the ligands was unsuccessful.

The phosphane-borane deprotection was initially assessed by LC-MS analysis of the modified proteins after treatment. However, the MS resolution is not high enough to clearly distinguish a mass difference of 2 Da for proteins of several kDa, so I could not immediately differentiate between unprotected proteins and proteins that were successfully deprotected but subsequently oxidized.

Previous studies reported the synthesis of Pd(0) or Rh(I) phosphane complexes by reacting Pd(II) or Rh(III) salts with phosphane boranes, without previous decomplexation, making use of the reductive properties of the borane group.[63-64] Although reacting the phosphane-boranes **10-12** with $RhCl_3$ or $Pd(OAc)_2$ was successful regarding borane decomplexation and metal complex formation, this was accompanied by the reduction of the maleimide double bond, rendering the complexes unsuitable for bioconjugation. Experiments are ongoing for the synthesis of Pd(0) and Rh(I) metal complexes in situ, with borane protected maleimide-phosphanes in presence of borane trapping agents, in order to protect the maleimide double bond.

Previous studies reporting synthesis of protein conjugates via intermolecular hydrazone linkage[65-67] led us to a successful approach for phosphane ligand incorporation into proteins.

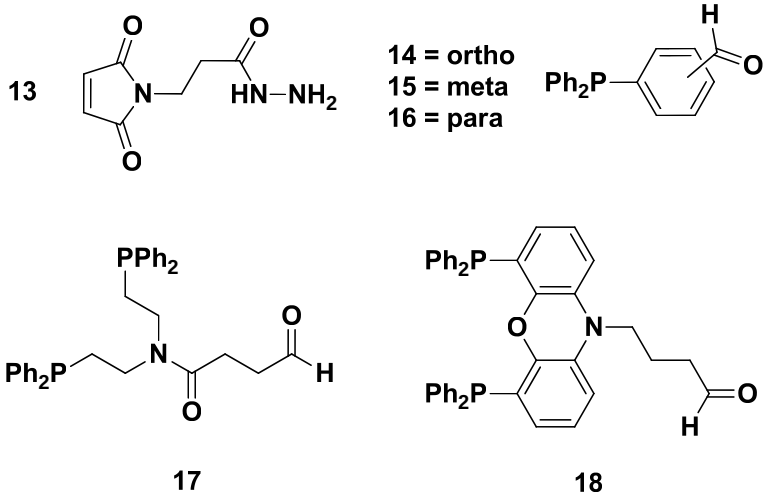


Figure 7 Structures of maleimide propionic acid hydrazide **13** and phosphane-aldehydes **14-18**

First the maleimide propionic acid hydrazide**13** was coupled covalently to ALBP,

Δ98Δ, AppA Y21C and AppA Q63C, followed by incorporation of phosphane-aldehydes **14-18** via hydrazone formation. Phosphane ligands incorporation was confirmed by MS analysis (Table 4) and chemoselective modification of cysteine was confirmed by partial digestion with trypsin followed by MS.

The hydrazone ligation of aldehyde functionalized phosphanes **14-16** to the four proteins containing a hydrazide functionality was proved to be highly selective (Table 4). The coupling efficiency was 100% and no secondary modification of the NH_2 group of lysines occurred, as determined by trypsin mediated partial digestion combined with MS analysis. All the couplings conditions required were compatible with proteins in solutions.

Phosphanes **17** and **18** were successfully coupled to AppA Y21C, AppA Q63C and Δ98Δ, without being possible to reach complete modification. Once again it was not possible to couple these ligands to ALBP, supporting the results obtained for ligand **4**.

Table 4 Covalent incorporation of **13** into proteins via Michael addition, followed by incorporation of phosphanes **14-18** via hydrazone ligation to **13**

	Calculated Mass (Da)	**LC-MS Mass found (Da)**
ALBP - **13**	16787.9	16787.2±0.7
ALBP - **14**	17060.2	17060.2±0.65
ALBP - **15**	17060.2	17060.1±0.5
ALBP - **16**	17060.2	17060.7±0.6
Δ98Δ-**13**	13606.1	13605.8±0.5
Δ98Δ-**14**	13878.4	13878.2±0.65
Δ98Δ-**15**	13878.4	13878.2±0.6
Δ98Δ-**16**	13878.4	13878.2±0.4
Δ98Δ-**17**[a]	14113.7	14114.0±1.9
AppA Y21C -**13**	15578.2	15578.2±0.6
AppA Y21C -**14**	15850.5	15850.6±0.6
AppA Y21C -**15**	15850.5	15850.9±0.6
AppA Y21C -**16**	15850.5	15850.0±0.5
AppA Y21C -**17**[a]	16085.8	16083.4±2.8
AppA Q63C -**13**	15613. 3	15612.8±0.6
AppA Q63C -**14**	15885.5	15900.2±0.8
AppA Q63C -**15**	15885.5	15884.8±0.6
AppA Q63C -**16**	15885.5	15884.6±0.8
AppA Q63C -**17**[a]	16120.8	16120.2±1.1
AppA Q63C -**18**[a]	16216.9	16216.1±0.9

[a] Incomplete conversion

II.3 Conclusions

In conclusion a method for site specific covalent incorporation of phosphane ligands into structurally different proteins was developed. The strategy is based on the intermolecular hydrazone linkage between an aldehyde functionalized phosphane and a hydrazide – maleimide precoupled to the thiol group of a cysteine residue of a protein.

This concept of hydrazone ligation for incorporation of phosphane ligands into proteins is a very powerful tool, making it possible to combine a large variety of phosphane ligands and virtually any protein.

Next step will be the formation of metal complexes of hybrid proteins either by complexation of metal precursors to the hybrid protein modified with phosphane or by hydrazone ligation of a preformed metal complex of an aldehyde phosphane.

This first site-selective covalent coupling method of phosphane ligands enables the modification of a wide range of structurally different proteins. Moreover, proteins with improved properties like thermostability or organic solvents tolerance can be used, potentially leading to artificial metalloenzymes with improved stability and activity.

II.4 Experimental

II.4.1 General remarks

Chemicals were purchased from Aldrich and Fluka unless otherwise stated, and were used as received. 3-maleimidopropionic acid hydrazide hydrochloride **13** was purchased from Apollo Scientific. N,N-dimethylformamide was degassed and stored on molsieves under an argon atmosphere. All air- and water-sensitive reactions were carried out under dry, air free conditions using dry degassed solvents and standard Schlenk techniques under an atmosphere of purified argon. NMR spectra were recorded at room temperature on Bruker Avance 300, 400 and 500 spectrometers. Positive chemical shifts (δ) are given (in ppm) for high-frequency shifts relative to a

TMS reference (^{1}H and ^{13}C) or an 85% H_3PO_4 reference (^{31}P). ^{13}C and ^{31}P spectra were measured with ^{1}H decoupling.

Protein stability (apo and modified proteins) was assessed by Bradford assay[68] and SDS-PAGE .

Modified proteins were analysed by LC-MS (ES+), partial digestion with trypsin followed by MS and Ellman Assay[69] (where possible).

II.4.2 Whole protein mass spectrometry using ESI-TOF

The protein sample (20ul, 5 pM/ul) was desalted on-line through a XTerra MS C8 2.1 x 10 mm column, eluted with an increasing acetonitrile concentration (2% acetonitrile, 98% aqueous 1% formic acid to 98% acetonitrile 2% aqueous 1% formic acid) and delivered to an electrospray ionisation mass spectrometer (LCT, Micromass, Manchester, U.K.) which had previously been calibrated using myoglobin. An envelope of multiply charged signals was obtained and deconvoluted using MaxEnt1 software to give the molecular mass of the protein. Standard deviations were calculated manually from the raw MS data.

The sequence of ALBP reported by Bernlohr *et al*[70] was used as reference for the mass of ALBP C1A his-tagged mutant. For the mass of the Δ98Δ mutant, the Δ98Δ sequence reported by Delfino *et al*[33] was used as reference and for the mass of AppA mutants the AppA sequence reported by Laan *et al*[71] was used.

II.4.3 Protein tryptic digestion

The protein samples (5 µl, 10 pmoles/µl) were dialysed into ammonium bicarbonate (50 mM, pH 8.0) using a membrane filter (Millipore, Billerica, MA) and trypsin (0.5 µl, 0.1 µg, Promega, Madison, WI) was added, followed by incubation at 37°C for 16 hours.

II.4.4 Δ98Δ - reduction of disulfide bridges was performed every time before modification with carboxylic acid or maleimide functionalized compounds. Prior to the reduction reaction the protein solution was degassed by purging with argon for 8 hours. 6 µmol (6 eq.) of DTT were added from a freshly prepared stock solution of 0.5 M in water to a solution of 1 µmol of protein (1 eq.) in Tris·HCl (50 mM, 10 ml, pH 8.0) in a Schlenk tube. The reaction mixture was stirred at room temperature for 1 hour. The reduced protein solution was used for the coupling reactions without

further purification. For each coupling reaction the molar excess of modification reagent was higher than 12-fold, to assure the right amount of modification reagent for the potential modification of both the thiol groups of DTT and the thiol group of the target cysteine.

II.4.5 Synthesis of modification reagents

Borane protected 3-(diphenylphosphane)benzenesulfonic acid

The borane protection of 3-(diphenylphosphane)benzenesulfonic acid was done based on a literature procedure.[72] A Schlenk tube containing a solution of 3-(diphenylphosphane)benzenesulfonic acid (68.41 mg, 0.2 mmol) in dimethylformamide (1 ml) was placed on an ice-bath (0ºC) and a 2M solution of BH_3SMe_2 (759.5 mg, 10 mmol) in tetrahydrofuran was added slowly via syringe, under stirring, into the ice cooled flask. After that the solution was allowed to reach room temperature and the stirring was continued for 16 hours, when the reaction was complete. The excess of BH_3SMe_2 was removed by adding toluene and evaporating the solvent under vacuum; this process was repeated 3 times. The product was obtained as a white solid and was used without further purification. Yield 98% (69.81 mg). ^{31}P $\{^1H\}$ (161.9 MHz, D_2O): δ = 20.64 ppm (br). ESI-MS (ES+) m/z=356.14 ([M] $C_{18}H_{18}BO_3PS$ requires 356.18)

Oxidation of 3-(diphenylphosphane)benzenesulfonic acid - was performed according to a literature procedure.[73]

3-(diphenylphosphane)benzoic acid (2) - this compound was prepared according to a literature procedure[74] and it was kindly provided by Dr Rene de Heeten (University of Amsterdam).

4-(bis(2-(diphenylphosphane)-ethyl)amino-4-oxobutanoic acid (4) - this compound was kindly provided by Dr Rene de Heeten (University of Amsterdam).[75]

CDI activation of phosphane carboxylic acids 1-4 – was performed according to a literature procedure.[76]

N-(4-(diphenylphosphane sulfide)benzyl) maleimide – this compound was kindly provided by Dr Yanmei Zhang (University of St Andrews).

N-(propargyl) maleimide (5) - this compound was prepared according to a literature procedure[77] and was kindly provided by Dr Gregorio Guisado (University of St Andrews).

N-(3-ethynylphenyl) maleimide (6) - this compound was kindly provided by Dr Nikos Tsoureas (University of St Andrews).

N-(*p*-azidophenyl)-maleimide (7) - this compound was prepared according to a literature procedure[78] and was kindly provided by Dr Gregorio Guisado (University of St Andrews).

Diphenyl(propargyl)phosphane borane complex (8) - this compound was prepared according to a literature procedure.[52]

(3-Azidopropyl)diphenyl phosphane borane complex (9) - this compound was kindly provided by Dr Gregorio Guisado (University of St Andrews).

Click-compounds 10-12

a. N-(1-((1-(3-(diphenyl(borane)phosphane)propyl)-1H-1,2,3-triazol-4-yl)methyl) maleimide (10)

This compound was synthesized based on a literature procedure.[79-80] To a solution of **(9)** (94.41 mg, 0.333 mmol) in $CHCl_3$ (1 ml) a solution of $CuSO_4$x $5H_2O$ (22.18 mg, 0.139 mmol) in water (0.5 ml) and a solution of sodium ascorbate (137.63 mg, 0.695 mmol) in water (0.5 ml) were added. Finally a solution of N-(propargyl)-maleimide (37.55 mg, 0.278 mmol) in $CHCl_3$ (1 ml) was added. The heterogeneous reaction mixture was stirred at room temperature for 5 hours. After addition of H_2O (10 ml) the aqueous phase was extracted with CH_2Cl_2 (3x10 ml) and the combined organic layers were washed with water (10 ml) and brine (10 ml). The organic phase was dried over anhydrous Na_2SO_4, filtered and concentrated. Flash chromatography on silica gel (CH_2Cl_2:EtOAc 4:1) afforded the product as a light yellow solid in 76% yield (88.36 mg).

^{1}H NMR (400 MHz, $CDCl_3$): δ=7.3-7.65 (m, ar, 10H), 6.7 (s, 2H), 4.8 (s, 2H), 4.37 (t, J=6.44 Hz, 2H), 2.0-2.25 (m, 4H), 0.87 ppm (br m, 3H-BH_3). ^{31}P {^{1}H} NMR (161.9 MHz, $CDCl_3$): δ=15.75 ppm (br). ^{13}C {^{1}H} NMR (75 MHz, $CDCl_3$): δ= 170.0 (2CO), 142.8 (2C ar diφ), 134.3 (2CH, maleimide), 132.1, 132.0, 131.5, 129.0, 128.9 (10CH, ar diφ), 128.2 (1C triazole), 122.6 (1CH, triazole), 50.6-50.4 (d, $1CH_2$), 32.85 ($1CH_2$ – propargyl), 24.1 ($1CH_2$), 22.8-22.3 ppm (d, $1CH_2$). HRMS (FTMS +): m/z calc. for $C_{22}H_{25}N_4O_2BP$ [$M+H^+$]: 419.1803; found 419.1806.

b. N-(1-(4-(4-((diphenyl(borane)phosphane)methyl)-1H-1,2,3-triazol-1-yl)phenyl) maleimide (11)

This compound was synthesized based on a literature procedure.[52,79] To a solution of

(7) (81.85 mg, 0.382 mmol) in $CHCl_3$ (2ml), a solution of $CuSO_4x5H_2O$ (30.48 mg, 0.191 mmol) in water (1 ml) and a solution of sodium ascorbate (189.18 mg, 0.955 mmol) in water (1 ml) were added. Finally the solution of **(8)** (109.19 mg, 0.458 mmol) in $CHCl_3$ (2 ml) was added. The heterogeneous reaction mixture was stirred at room temperature for 5 hours. After addition of H_2O (10 ml) the aqueous phase was extracted with CH_2Cl_2 (3x20 ml) and the combined organic layers were washed with water (20 ml) and brine (20 ml). The organic phase was dried over anhydrous Na_2SO_4, filtered and concentrated. Flash chromatography on silica gel (CH_2Cl_2:EtOAc 4:1) afforded the product as light yellow solid in 92% yield (158.94 mg).

^{1}H NMR (400 MHz, $CDCl_3$): δ=7.4-7.8 (m, 15H), 6.9 (s, 2H), 3.8 (d, J=11.0 Hz, 2H), 0.88 ppm (br m, 3H-BH_3) . ^{31}P {^{1}H} NMR (161.9 MHz, $CDCl_3$): δ=16.3 ppm (br). ^{13}C {^{1}H} NMR (75 MHz, $CDCl_3$): δ= 169.0 (2CO), 140.8 (2C ar diφ), 134.4 (2CH, maleimide), 132.5, 132.4 (4CH, ar diφ), 131.6 (2CH, ar), 129.0, 128.9 (6CH, ar diφ), 127.7 (3C, 2 from ar, 1 from triazole), 126.9 (2CH, ar), 120.9 (1CH, triazole ring), 24.9-24.4 ppm (d, 1 CH_2). MALDI-MS m/z = 452.15 ([M] $C_{25}H_{22}BN_4O_2P$ requires 452.25)

c. N-(1-(3-(1-(3-(diphenyl(borane)phosphane)propyl)-1H-1,2,3-triazol-4-yl)phenyl) maleimide (12)

The compound was synthesized based on a literature procedure.[79-80] To a solution of **(9)** (249.82 mg, 0.882 mmol) in $CHCl_3$ (3 ml) a solution of $CuSO_4x$ $5H_2O$ (58.65 mg, 0.367 mmol) in water (1.5 ml) and a solution of sodium ascorbate (364 mg, 1.837 mmol) in water (1.5 ml) were added. Finally a solution of N-(3-ethynylphenyl)-maleimide (145 mg, 0.735 mmol) in $CHCl_3$ (3 ml) was added. The heterogeneous reaction mixture was stirred at room temperature for 6 hours. After addition of H_2O (15 ml) the aqueous phase was extracted with CH_2Cl_2 (3x30 ml) and the combined organic layers were washed with water (30 ml) and brine (30 ml). The organic phase was dried over anhydrous Na_2SO_4, filtered and concentrated. Flash chromatography on silica gel (CH_2Cl_2:EtOAc 4:1) afforded the product as cream coloured, oily solid in 53% yield (187.10 mg).

^{1}H NMR (300 MHz, $CDCl_3$): δ=7.25-7.9 (m, 15H), 6.86 (s, 2H), 4.46 (t, J=6.03 Hz, 2H), 2.1-2.3 ppm (m, 4H). ^{31}P {^{1}H} NMR (121 MHz, $CDCl_3$): δ=15.31 ppm (br). ^{13}C {^{1}H} NMR (75 MHz, $CDCl_3$): δ=169.4 (2CO), 147.0 (1C-triazole ring), 134.4 (2CH), 132.2 (2CH), 132.1 (2CH), 131.8 (d, 2C), 131.6 (1CH-triazole ring), 129.8 (1CH),

129.0 (6CH), 128.3 (2C), 125.8 (1CH), 125.3 (1CH), 123.5 (1CH), 50.7 (d, $1CH_2$), 24.27 ($1CH_2$), 22.5 ppm (d, 1 CH_2). MALDI-MS m/z = 480.17 ([M] $C_{27}H_{26}BN_4O_2P$ requires 480.31)

3-(diphenylphosphane)benzaldehyde (15) - this compound was prepared according to a literature procedure[81] and was kindly provided by Peter Deuss (University of St Andrews).

4-(diphenylphosphane)benzaldehyde (16) - this compound was prepared according to a literature procedure[82] and was kindly provided by Peter Deuss (University of St Andrews).

N,N-bis(2-(diphenylphosphane)ethyl)-4-oxobutanamide (17) – this compound was kindly provided Peter Deuss (University of St Andrews).[83]

4-(4,6-bis(diphenylphosphane)-phenoxazin-10-yl)butanal (18) - this compound was kindly provided Peter Deuss (University of St Andrews).[83]

II.4.6 Protein engineering, expression and purification

Bacterial strains and media

Cloning was performed using *E. coli* XL1-Blue competent cells grown in LB medium, following established protocols.[84] For heterologous overexpression, *E. coli* BL21(DE3) or M15 pREP4 cells harbouring the appropriate expression construct were grown in PB medium (20 g l^{-1} tryptone, 10 g l^{-1} yeast extract, 5 g l^{-1} glucose, 5 g l^{-1} NaCl, 8.7 g l^{-1} K_2HPO_4, pH 7.0). Kanamycin and ampicilin were used at 50 and 100 mg l^{-1}, respectively.

Site-directed mutagenesis

Mutations were introduced using the Quick-Change Site-Directed Mutagenesis Kit (Stratagene). The following plasmids were used as templates for the PCR reactions: pQE-$AppA_{5\text{-}125}$ for AppA,[71] pRSET-ALBP for ALBP,[30,85] pET22b(+) Δ98Δ IFABP for Δ98Δ.[33]

Following PCR reaction, the *dam* methylated plasmid templates were digested using the *Dpn I* restriction enzyme (Stratagene) and the vectors containing the new mutations were transformed into *E. coli* XL1-Blue and purified using QIAprep

Miniprep kit (Qiagen). All mutations were verified by sequencing (The Sequencing Service, Dundee, UK). The ALBP and AppA mutants verified by sequencing were transformed to *E. coli* M15-pREP4 for protein overexpression, and the Δ98Δ mutants were transformed to *E. coli* BL21(DE3).

The mutagenic primers are specified in Tables 5-8.

Table 5 Sequences of primers used o introduce mutations in AppA. Mutated bases are indicated in bold.

Primer	Mutation acquired	Sequence
Primer1	C19L/C20T	5' GCTCGGATCTGGTTTCC**CTGACC**TACCGCAGCCTGGCGGC 3'
Primer 2	TGCTGC/CTGACC	5' GCCGCCAGGCTGCGGTA**GGTCAG**GGAAACCAGATCCGAGC 3'
Primer 3	C110A	5' CACATGCAGCTCTCC**GCG**TCGGAGGCCGACATG 3'
Primer 4	TGC/GCG	5' CATGTCGGCCTCCGA**CGC**GGAGAGCTGCATGTG 3'
Primer 5	Y21C	5' GTTTCCCTGACC**TGC**CGCAGCCTGGCG 3'
Primer 6	TAC/TGC	5' CGCCAGGCTGCG**GCA**GGTCAGGGAAAC 3'
Primer 7	Q63C	5' CAGGGCGTCTTCTTC**TGC**TGGCTCGAAGGCCAC 3'
Primer 8	CAG/TGC	5' GTGGCCTTCGAGCCA**GCA**GAAGAAGACGCCCTG 3'
Primer 9	W104C	5' CGCTTTGCGGGA**TGC**CACATGCAGCTC 3'
Primer 10	TGG/TGC	5' GAGCTGCATGTG**GCA**TCCCGCAAAGCG 3'
Primer 11	L65R	5' CTTCCAGTGG**CGT**GAAGGCCACC 3'
Primer 12	CTC/CGT	5' GGTGGCCTTC**ACG**CCACTGGAAG 3'
Primer 13	V75R	5' CGTGGCGGAG**CGT**ATGAGCCACATC 3'
Primer 14	GTC/CGT	5' GATGTGGCTCAT**ACG**CTCCGCCACG 3'

Table 6 AppA mutants created by site-directed mutagenesis

AppA-BLUF domain mutant	(Short) Name
$AppA_{5-125}$ C19L/C20T/Y21C/C110A	AppA **Y21C**
$AppA_{5-125}$ C19L/C20T/Q63C/C110A	AppA **Q63C**
$AppA_{5-125}$ C19L/C20T/W104C/C110A	AppA **W104C**

Table 7 Sequences for mutagenic primers used for C1A mutation in ALBP. Mutated bases are indicated in bold.

Primer	Mutation acquired	Sequence
ALBP C1A-F	C1A	5' CCTCGAAGGTTTACAAAATG**GCC**GATGCCTTTGTGGGAACC 3'
ALBP C1A-R	TGT/GCC	5' GGTTCCCACAAAGGCATC**GGC**CATTTTGTAAACCTTCGAGG 3'

Table 8 Sequences for mutagenic primers used for single point mutations in Δ98Δ. Mutated bases are indicated in bold.

Primer	Mutation acquired	Sequence
Primer SC-F	Stop codon replaced with alanine	5' GAGTGGAGGCCAAGCGG**GCC**GGATCCGAATTCGAGC 3'
Primer SC-R	TGA/GCC	5' GCTCGAATTCGGATCC**GGC**CCGCTTGGCCTCCACTC 3'
Primer RG-F	G base deleted	5' GCCAAGCGGGCCGATCCGAATTCGAG 3'
Primer RG-R	GCC**Δ**GAT	5' CTCGAATTCGGATCGGCCCGCTTGGC 3'

Primer 74 F	K74Q	5’ CGTGTAGACAATGGA**CAG**GAGCTGATTGCTGTC 3’
Primer 74 R	AAG/CAG	5’ GACAGCAATCAGCTC**CTG**TCCATTGTCTACACG 3’
Primer 104 F	K104Q	5’ GCTCCGTCGAC**CAG**CTTGCGGCCGC 3’
Primer 104 R	AAG/CAG	5’ GCGGCCGCAAG**CTG**GTCGACGGAGC 3’

Cloning

To clone the ALBP gene into pQE30, cDNA corresponding to the ALBP C1A mutant was amplified by PCR using pRSET-ALBP C1A as template. The PCR product was purified from agarose gel (Wizard SV Gel and PCR Clean-Up System, Promega), digested with BamHI (Fermentas) and PstI (Fermentas) and ligated into BamH I/Pst I digested pQE30 vector (Qiagen) using T4 DNA ligase (Fermentas). The construct was verified by sequencing (The Sequencing Service, Dundee, UK).

Table 9 Primers used for cloning of the ALBP C1A segment. Restriction sites (BamH I or Pst I) are underlined, the enterokinase site is given in italics.

Primer	**Sequence**
Sc 1	5’-GCACGGATCC*GATGACGATGACAAA* ATGGCCGATGCCTTTGTGG-3’
Sc 2	5’-GCACCTGCAG TCATGCCCTTTCATAAACTCTTG-3’

Protein expression and purification

AppA

The AppA Y21C and AppA Q63C mutants were produced and purified according to a literature procedure.[71] Immobilized nickel affinity chromatography was performed using a 5ml HisTrap™ FF column (GE Healthcare Life Sciences).

ALBP

E. coli M15 pREP4 transformed with pQE-ALBP C1A was grown at 37°C in PB containing ampicillin and kanamycin. When the OD_{600} reached 3, protein expression was induced by adding IPTG to a final concentration of 0.5 mM and growth was continued for 16 hours at 20°C. The cells were then harvested by centrifugation (4°C, 20 min, 3000 g) and resuspended in ice-cold lysis buffer (50 mM phosphate buffer, 50 mM NaCl, 10% glycerol, 0.5% Nonidet P-40,[86] pH 8.0) containing protease inhibitor (Complete EDTA free, Roche). The cell-suspension was treated with deoxyribonuclease (30 μg ml^{-1}), ribonuclease (30 μg ml^{-1}), and lysozyme (1 mg

ml^{-1}), held on ice for 30 min, sonicated (10 minutes with a 50% duty cycle) and centrifuged (4°C, 30 min, 27000 g). The soluble fraction was applied to a 5ml HisTrap™ FF column (GE Healthcare Life Sciences), the column was washed first with washing buffer (50 mM phosphate, 150 mM NaCl, 20 mM imidazole, 10% glycerol, pH 8.0) containing 0.5% detergent (Nonidet P-40),[86] and then with an equivalent volume of washing buffer without detergent. The protein was eluted with elution buffer (50 mM phosphate, 150 mM NaCl, 500 mM imidazole, pH 8.0) and dialyzed overnight at 4ºC against 50 mM phosphate, 50 mM NaCl, pH 8.0 and stored at -20°C. Protein concentration was determined using the Bradford assay.[68] Purity of the samples was checked by SDS-PAGE.

Δ98Δ

E. coli BL21(DE3), transformed with pET-22b(+)-Δ98Δ IFABP was grown at 37°C in PB containing ampicillin. When the OD_{600} reached 0.6, the protein expression was induced by adding IPTG (0.1 mM) and growth was continued for 16 hours at 25°C. The cells were then harvested by centrifugation (4°C, 20 min, 3000 g) and resuspended in ice-cold lysis buffer (50 mM Tris-HCl, 150 mM NaCl, 0.1% Triton-X 100, 15% glycerol, pH 8.0) containing protease inhibitors (Pefabloc, Roche). The cell-suspension was treated with deoxyribonuclease (30 μg ml^{-1}), ribonuclease (30 μg ml^{-1}), and lysozyme (1 mg ml^{-1}), held on ice for 30 min, sonicated (10 minutes with a 50% duty cycle) and centrifuged (4°C, 30 min, 27000 g). The soluble fraction was applied to a 5ml HisTrap™ FF column (GE Healthcare Life Sciences), the column was washed first with washing buffer (50 mM Tris-HCl, 500 mM NaCl, 20 mM imidazole, pH 8.0) containing 0.5% detergent (Nonidet P-40),[86] and then with an equal volume of the same washing buffer without detergent. The protein was eluted using elution buffer (50 mM Tris-HCl, 500 mM NaCl, 500 mM imidazole, pH 8.0) and dialyzed at 4°C against 10 mM Tris-HCl, pH 8.0. After dialysis the protein was stored at -20°C, in buffer (10mM Tris-HCl, pH 8.0) containing 10% glycerol. Protein concentration was determined using the Bradford assay[68] and the purity by SDS-PAGE.

II.4.7 Chemical modification of proteins

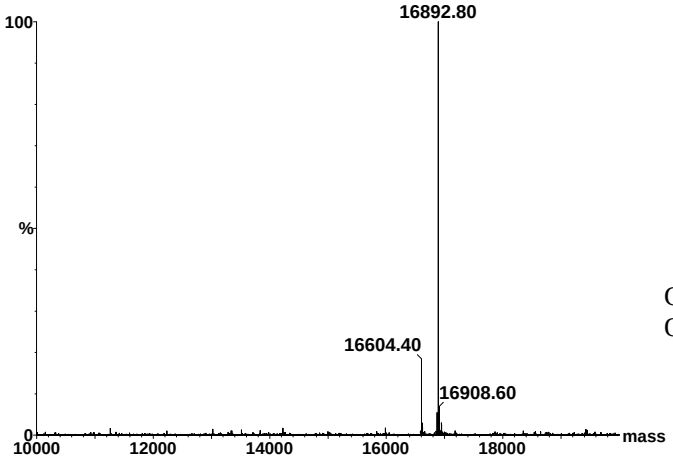
a. Conjugation with carboxylic acid functionalized phosphanes 1-4

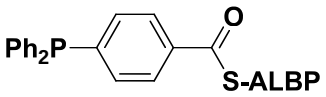
ALBP

1. Conjugation with 3

Prior to coupling the protein solution was degassed by purging with argon for 8 hours.

2.5 µmol (25 eq.) of CDI activated **3** in 50 µl of dimethylformamide was added slowly to a solution of 100 nmol of protein (1 eq.) in Tris·HCl (50 mM, 1 ml, pH 8.0) in a Schlenk tube. The cloudy reaction mixture was stirred at room temperature for 16h. The insoluble fraction of the excess modification reagent was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with degassed Tris·HCl (50 mM, pH 8.0) and concentrated to 0.5 ml. This step was repeated four times after which the modified protein was concentrated to 1 ml and stored at 4°C.

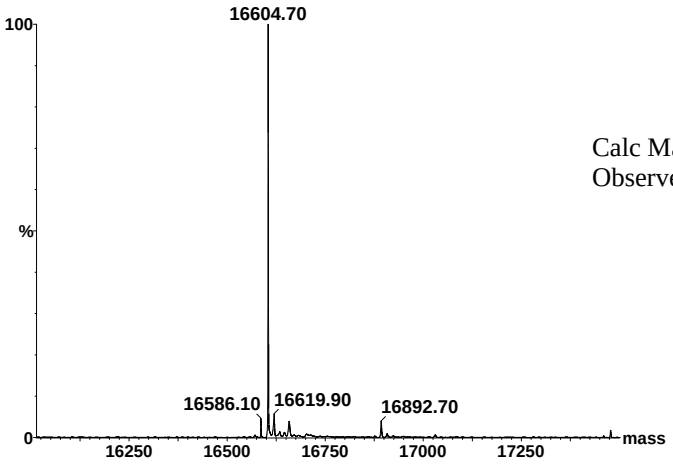




	ALBP	ALBP – **3**
Calc Mass	16604.8	16893.1
Observed Mass	16604.4	16892.8

Figure 8 LC-MS (ES+) spectrum of ALBP -**3**

The difference in ionization efficiency between unmodified and modified ALBP was assessed by comparing the MS spectra of modified protein with and without the addition of 1 equivalent of unmodified protein. The ionization efficiency of unmodified ALBP was found to be approximately 10x higher then the ionization efficiency of the ALBP-**3**.



	ALBP	ALBP – **3**
Calc Mass	16604.8	16893.1
Observed Mass	16604.7	16892.7

Figure 9 LC-MS (ES+) spectrum of ALBP-**3** after addition of 1 equivalent of unmodified protein

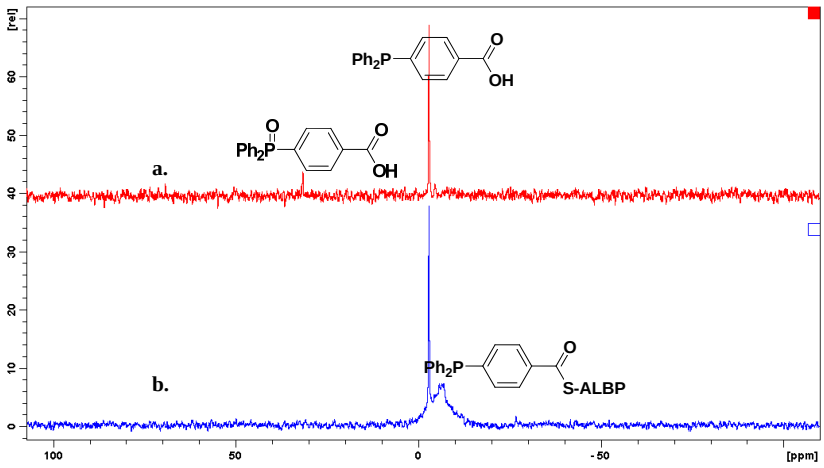


Figure 10 ^{31}P NMR {1H} NMR (202 MHz, D_2O) spectrum of: (a) ligand **3** in 50 mM Tris·HC (pH 8); (b) ALBP-**3** in 50 mM Tris·HC (pH 8)

In Figure 10- a, the sharp peak at 31.8 ppm is likely generated by traces of phosphane oxide, while the sharp peak at -2.8 ppm is generated by the free phosphane **3** in 50 mM Tris·HCl (pH 8). In Figure 10- b, the sharp peak at -2.8 ppm is likely generated by traces of free phosphane **3** left in solution after washing, while the broad peak from -15 ppm to 2 ppm is generated by the phosphane **3** coupled to ALBP.

Δ98Δ

1. Conjugation with 1

290 nmol (14.5 eq.) of CDI-activated **1** in 5.8 µl of dimethylformamide was added slowly to a solution of 20 nmol of protein (1 eq.) in Tris·HCl (50 mM, 0.4 ml, pH 8.0) containing DTT (6 eq.) in a Schlenk-tube. The reaction time required to avoid

multiple modification of the protein was 9 hours. The insoluble fraction of the excess modification reagent was removed by centrifugation at 49,000 g at 4°C for 10 min. The supernatant was transferred to a centrifugal concentrator (Vivaspin 2, 10,000 MWCO), diluted to 2 ml with degassed Tris·HCl (50 mM, pH 8.0) and concentrated to 0.1 ml. This step was repeated four times after which the modified protein was concentrated to 0.4 ml and stored at 4°C.

2. Conjugation with 2

Conjugation with **2** was performed as described for **1**. 260 nmol (13 eq.) of CDI-activated **2** in 5.2 µl of dimethylformamide were used for the coupling reaction. The reaction time was 4 hours, in order to achieve monomodification of the protein with **2**.

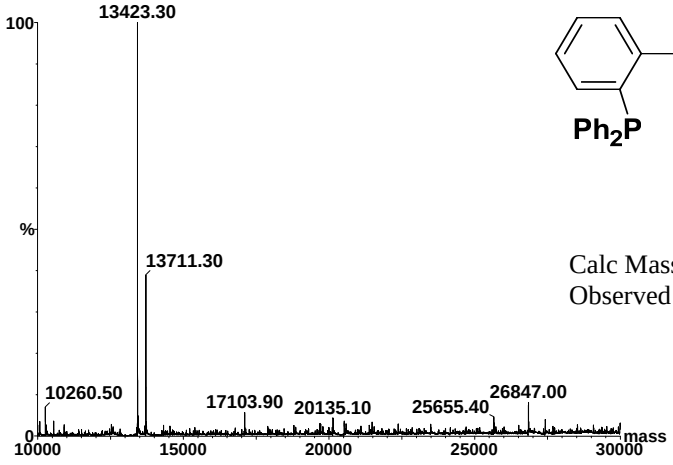
3. Conjugation with 3

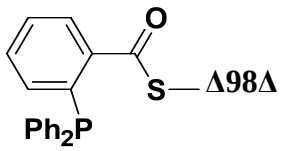
Conjugation with **3** was performed as described for **1**. 280 nmol (14 eq.) of CDI-activated **3** in 5.6 µl of dimethylformamide were used for the coupling reaction. The reaction time required to avoid multiple modification of the protein was 9 hours.

4. Conjugation with 4

Conjugation with **4** was performed as described for **1**. 390 nmol (19.5 eq.) of CDI-activated **4** in 7.8 µl of dimethylformamide were used for the coupling reaction. The reaction time required to avoid multiple modification of the protein was 9 hours.

It was not possible to accurately compare the ionization efficiency of unmodified and modified Δ98Δ, due to dimer formation observed by LC-MS (ES+).





	Δ98Δ	Δ98Δ-**1**
Calc Mass	13423.0	13711.3
Observed Mass	13423.3	13711.3

Figure 11 LC-MS (ES+) spectrum of Δ98Δ modified with **1**

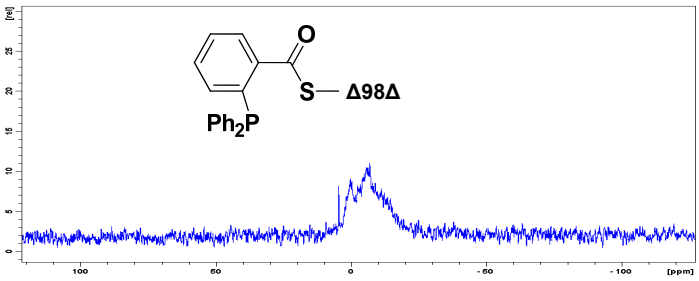
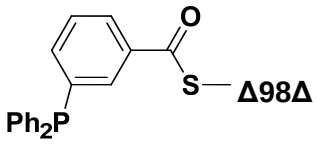


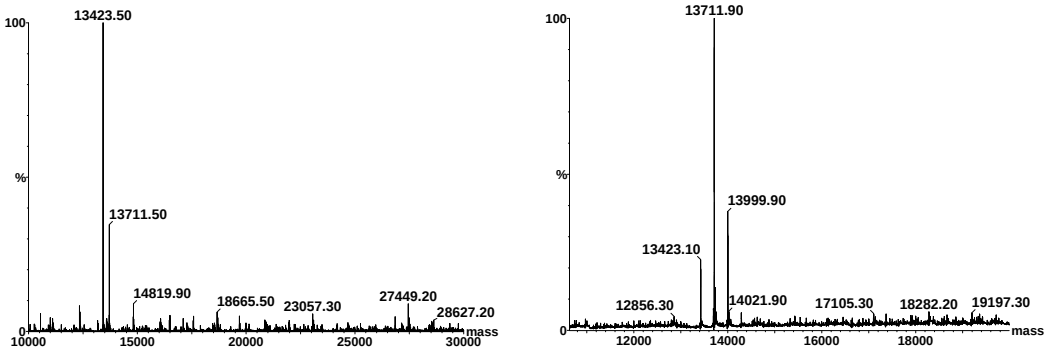
Figure 12 ^{31}P {^{1}H} NMR (202 MHz, D_2O) spectrum of Δ98Δ-**1** in 50 mM Tris·HCl (pH 8)

The sharp peak at +4.6 ppm is likely generated by traces of free phosphane **1** left in solution after washing. The two broad peaks (around 0 ppm and -10 ppm) could be generated either by the free phosphane **1** present inside the hydrophobic cavity of the protein and by the phosphane **1** covalently coupled to protein, or by the migration of the ligand **1** from the SH group to the NH_2 group of a lysine residue. However the LC-MS (ES+) spectrum of Δ98Δ modified with **1** revealed only monomodification of the protein.



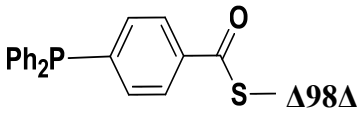
	Δ98Δ	Δ98Δ-**2**
Calc Mass	13423.0	13711.3
Observed Mass	13423.5	13711.5

	Δ98Δ	Δ98Δ-**2**	Δ98Δ-(**2**)$_2$
Calc Mass	13423.0	13711.3	13999.6
Observed Mass	13423.1	13711.9	13999.9



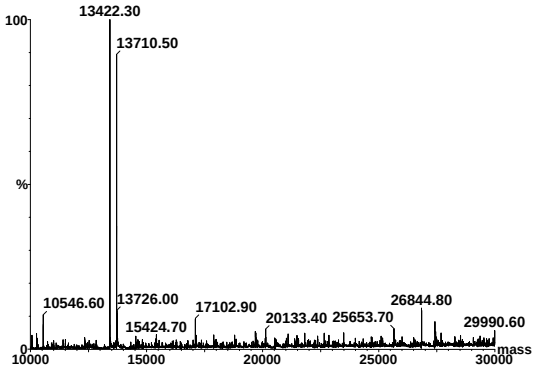
(a) (b)

Figure 13 LC-MS (ES+) spectrum of Δ98Δ-**2** (a, b). Unspecific modifications were obtained when high excess ligand was used to improve the coupling efficiency (b).

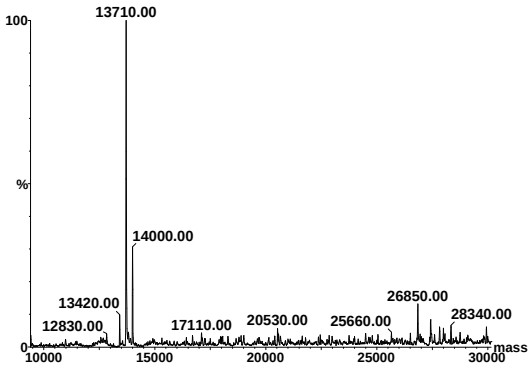


	Δ98Δ	Δ98Δ-**3**
Calc Mass	13423.0	13711.3

	Δ98Δ	Δ98Δ-**3**	Δ98Δ-(**3**)$_2$
Calc Mass	13423.0	13711.3	13999.6
Observed Mass	13420.0	13710.0	14000.0

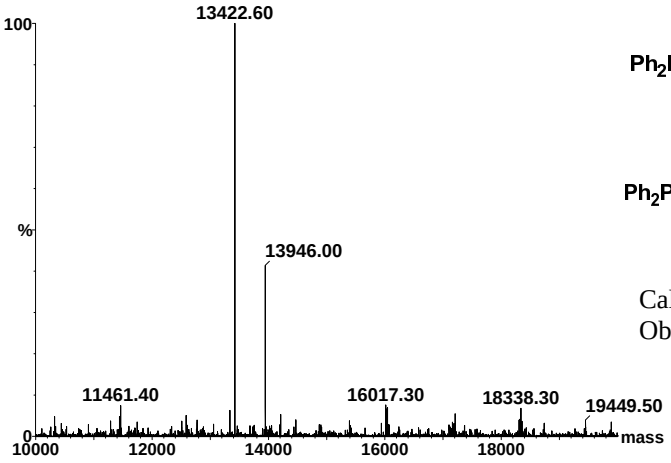


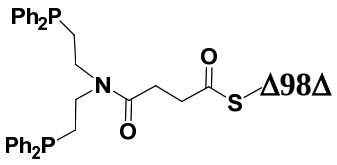
(a)



(b)

Figure 14 LC-MS (ES+) spectrum of Δ98Δ-**3** (a, b). The attempt to improve coupling efficiency by increasing the excess ligand led to unspecific modifications (b).





	Δ98Δ	Δ98Δ-**4**
Calc Mass	13423.0	13946.2
Observed Mass	13422.6	13946.0

Figure 15 LC-MS (ES+) spectrum of Δ98Δ-**4**

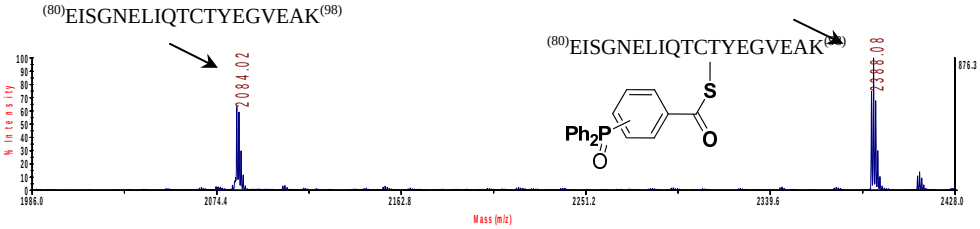
5. MALDI-MS analysis of tryptic peptides resulted from digestion of Δ98Δ-1, Δ98Δ-2 and Δ98Δ-3

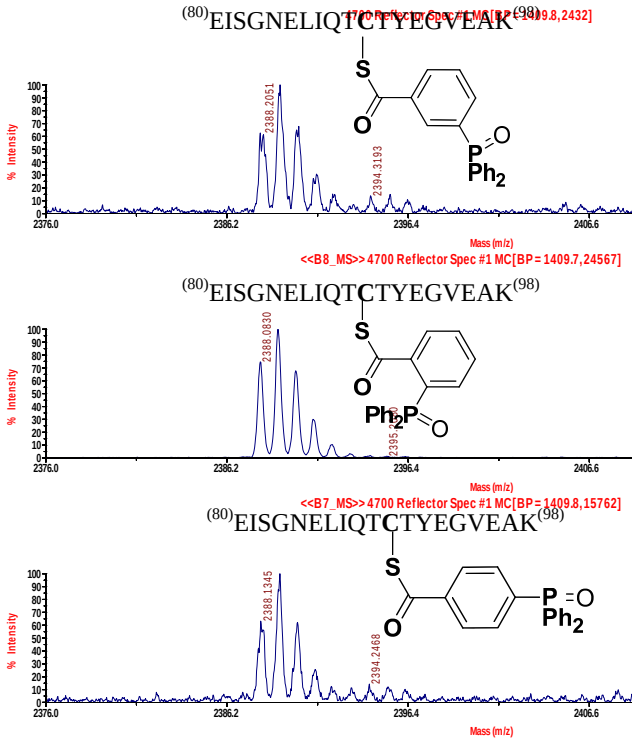
The digest solution (0.5 µl) was applied to the MALDI target along with alpha-cyano-4-hydroxycinnamic acid matrix (0.5 µl, 10 mg/ml in 50:50 acetonitrile:0.1% TFA) and 0.1% TFA (0.5 µl) and allowed to dry. MALDI MS was acquired using a 4800 MALDI TOF/TOF Analyser (Applied Biosystems, Foster City, CA) equipped with a Nd:YAG 355 nm laser and calibrated using a mixture of peptides. The spot was initially analysed in positive MS mode between 800 and 4000 m/z, by averaging 1000 laser spots. The

ion at m/z 2083.8 was associated to the mass of the unmodified peptide $^{(80)}$EISGNELIQTCTYEGVEAK$^{(98)}$ and the ion at m/z 2388.08 was found to correspond to the peptide $^{(80)}$EISGNELIQTCTYEGVEAK$^{(98)}$ modified with *o*, *m* or *p*-(diphenylphosphinyl)-benzoic acid (Figure 16), proving that Δ98Δ is modified with the ligands **1**-**3** at cysteine 90.

	Calc Mass	Observed Mass
Glu$^{(80)}$-Lys$^{(98)}$	2083.86	2084.02
Glu$^{(80)}$-Lys$^{(98)}$+ *o*, *m* or *p*-(diphenylphosphinyl)-benzoic acid	2388.16	2388.08

(a)





(b)

Figure 16 MALDI MS spectrum of peptides resulting from trypsin digestion of Δ98Δ+**1, 2** or **3** (a). Ions at m/z = 2388, corresponding to the peptide Glu$^{(80)}$-Lys$^{(98)}$ modified with *o*, *m* or *p*-(diphenylphosphinyl)-benzoic acid were identified for the tryptic peptide mixture of Δ98Δ modified with **1, 2** or **3** (b).

Protection of the NH_2 group of lysines (to prevent multiple modifications of Δ98Δ with phosphanes 1-4)

NH_2 protection via imine bond formation

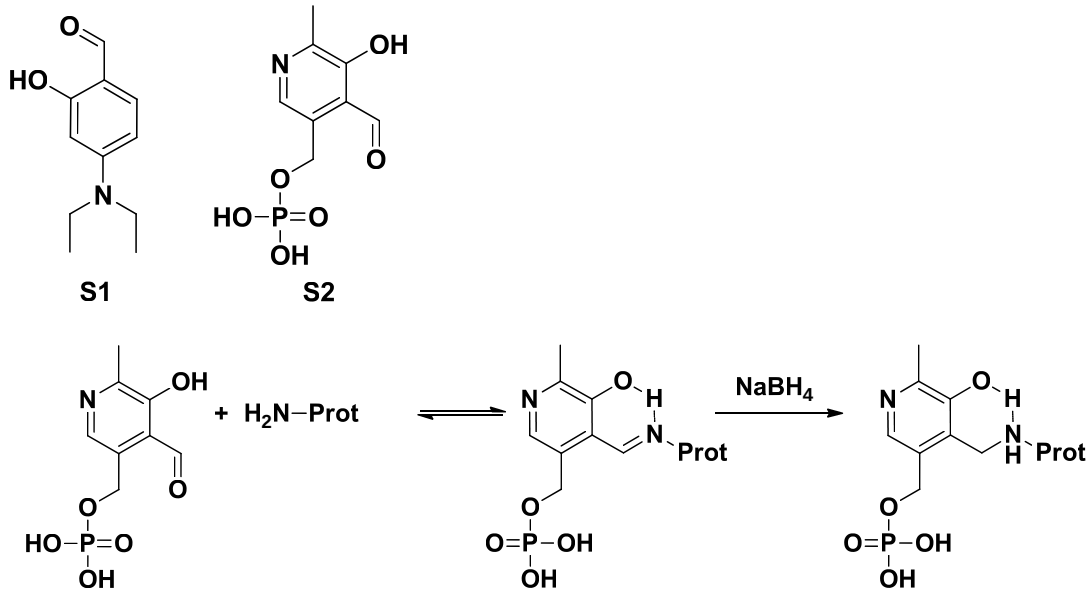


Figure 17 Lysines NH_2 group protection via imine bond formation[40]

Reaction between the primary amino-group of a lysine and the carbonyl group of 4-diethylamino 2-hydroxy benzaldehyde (**S1**) or pyridoxal 5'-phosphate (**S2**) leads to formation of an imine bond,[40] thereby protecting it from reacting with the activated phosphino-carboxylic acids.

1 ml of protein solution (100 µM in phosphate buffer at pH 7.5) was treated with a 10x, 25x or 50x molar excess aldehyde **S1** for 16 hours, at room temperature, under air, followed by centrifugation and washing with phosphate buffer (pH 7.5) to remove the excess **S1**. Subsequently, coupling of carboxylic acid functionalized phosphanes **1-4** was performed as described above for the unmodified protein.

The procedure for NH_2 protection using **S2** was performed as described for **S1**.

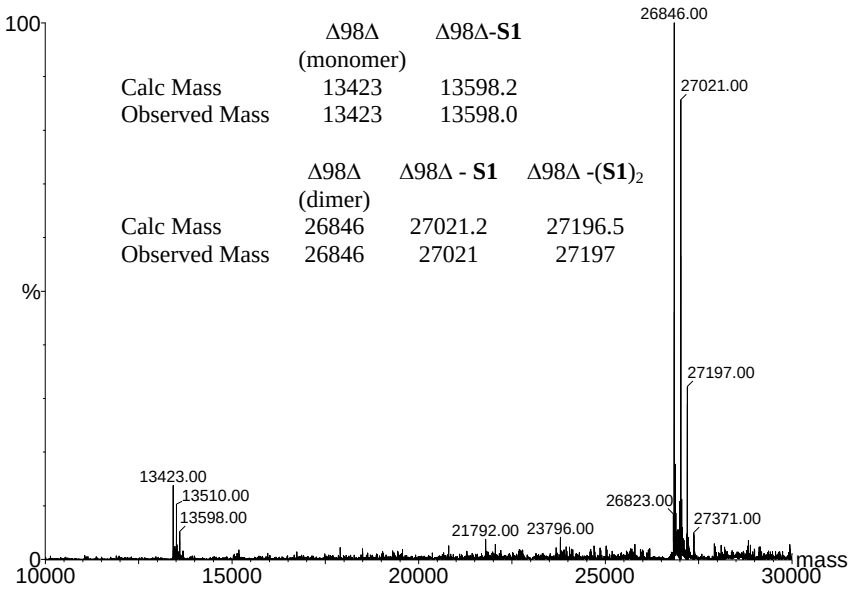


Figure 18 LC-MS (ES+) spectrum of Δ98Δ modified with **S1**

	Calc Mass	Observed Mass		Calc Mass	Observed Mass
Δ98Δ-**3**	13711.3	13710.7	Δ98Δ+**S1**+**3**	13886.5	13885.3
Δ98Δ-(**3**)$_2$	13999.6	14000.0	Δ98Δ+**S1**+(**3**)$_2$	14175.0	14175.0
Δ98Δ-(**3**)$_3$	14287.9	14287.4	Δ98Δ+**S1**+(**3**)$_3$	14463.3	14439.6

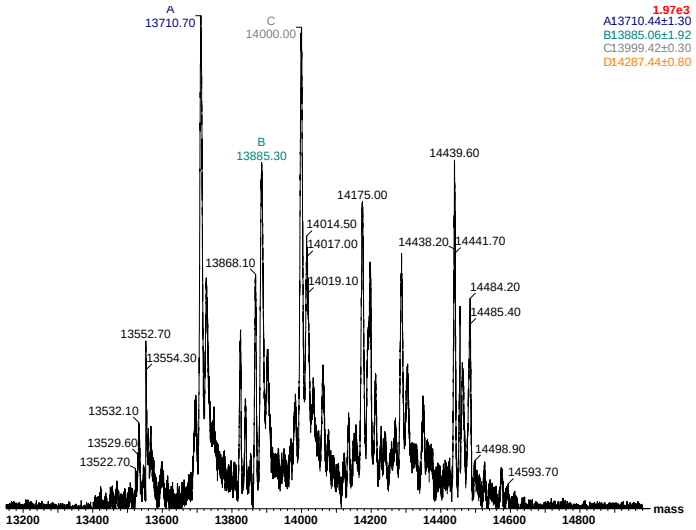
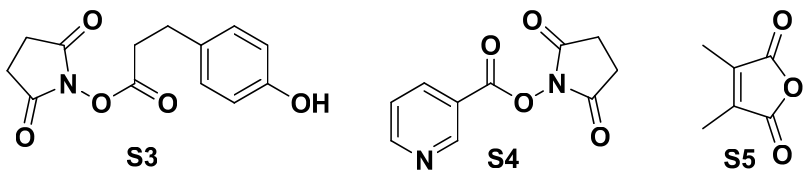


Figure 19 LC-MS (ES+) spectrum of Δ98Δ modified with **S1** and ligand **3**

NH_2 protection via amide bond formation



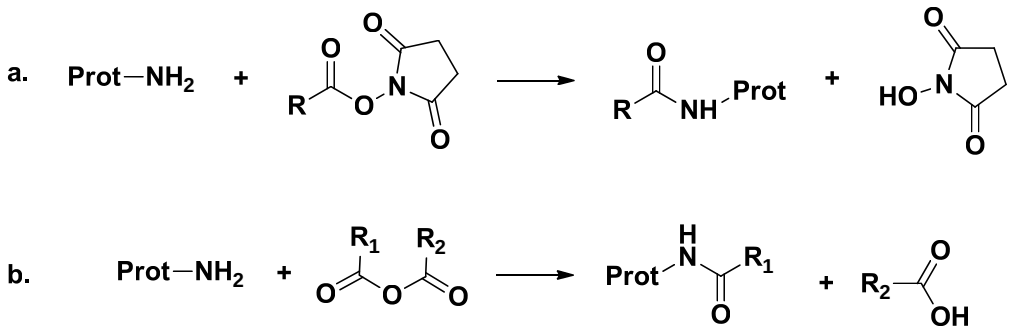


Figure 20 Lysines NH_2 group protection via amide bond formation (a) with N-hydroxysuccinimide ester derivatives and (b) with anhydrides

The amide bond was formed using either N-hydroxysuccinimide ester derivatives or anhydrides. Two esters were tested: 3-(4-hydroxyphenyl)-propionic acid N-hydroxysuccinimide ester (**S3**)[41] and nicotinic acid N-hydroxysuccinimide ester (**S4**),[42] and two anhydrydes: acetic anhydride[43] and 2,3 dimethyl maleic anhydride (**S5**).[44] NH_2 protection via amide bond formation was performed as described for **S1**.

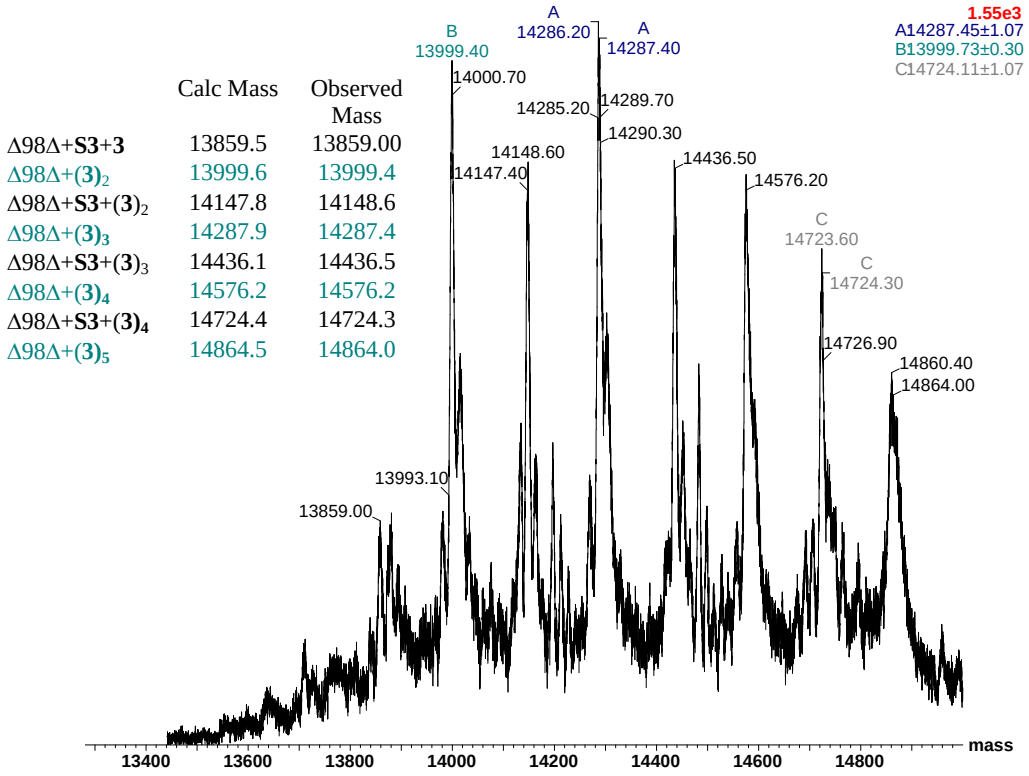


Figure 21 LC-MS (ES+) spectrum of Δ98Δ protected with **S3** (10x molar excess) and subsequently modified with ligand **3** (5x molar excess)

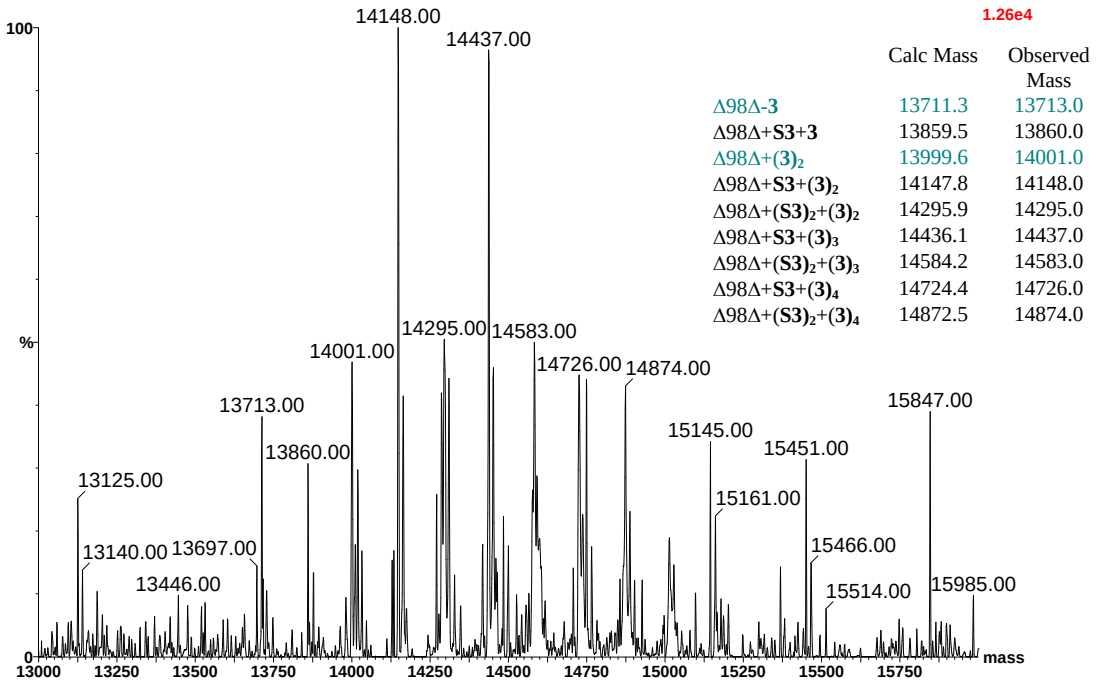


Figure 22 LC-MS (ES+) spectrum of Δ98Δ protected with **S3** (25x molar excess) and subsequently modified with ligand **3** (5x molar excess)

All tested compounds **S1-S5** afforded proteins with varying number of modifications and modification of all the NH_2 groups could not be achieved. Moreover, when the phosphane ligand was coupled to the protected protein, still multiple couplings of the ligand were observed, in addition to the multiple imine/amide modifications (Figure 19, Figure 21, Figure 22).

Replacement of the most reactive lysines (K74 and K104) with glutamine

This procedure was followed also by incomplete or multiple modifications of the protein with phosphanes **1-4**.

AppA Y21C

Prior to couplings the protein solution was degasses by purging with argon for 8 hours.

1. Conjugation with 1

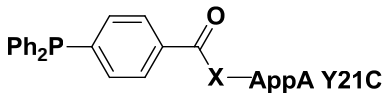
500 nmol (10 eq.) of CDI activated **1** in 10 µl of dimethylformamide was added to a solution of 50 nmol of protein (1 eq.) in buffer (50 mM Tris·HCl, 1 ml) containing 3M imidazole at pH 7.0. The reaction mixture was stirred at room temperature for

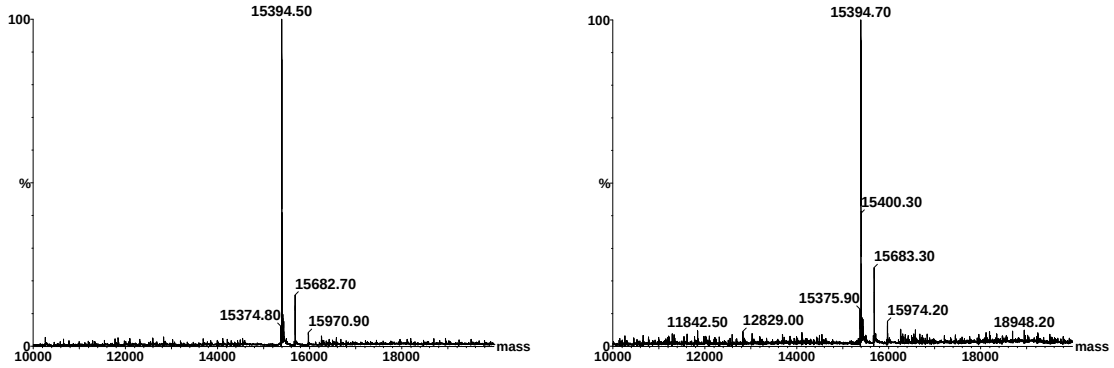
16h. The insoluble fraction of the excess modification reagent was removed by centrifugation at 49,000 g at 4°C for 10 min. The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with degassed Tris·HCl (50 mM, pH 7.0) and concentrated to 0.5 ml. This step was repeated four times after which the modified protein was concentrated to 1 ml and stored at 4°C.

2. Conjugation with 2, 3 and 4

Conjugation with **2, 3** and **4** was performed as described for **1**.

	Calc Mass	Observed Mass **(a)**	Observed Mass **(b)**
AppA	15395.1	15394.5	15394.7
AppA-**3**	15683.4	15682.7	15683.3
AppA-(**3**)$_2$	15971.7	15970.9	15974.2



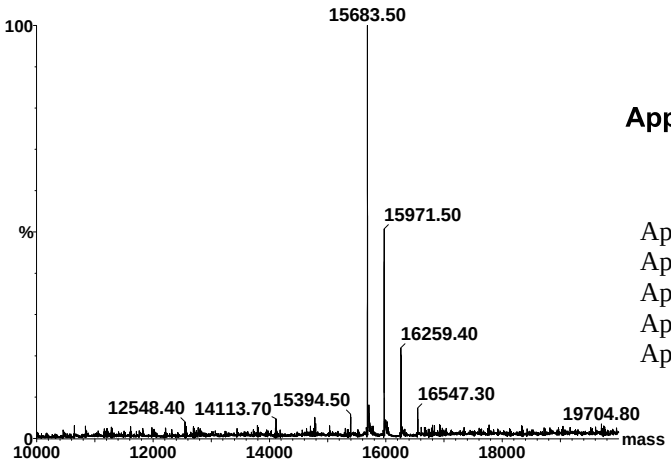


(a) (b)

Figure 23 LC-MS (ES+) spectrum of AppA Y21C - **3**. Coupling reaction done in absence of imidazole, using a 10x (a) and a 50x (b) molar excess of ligand **3**

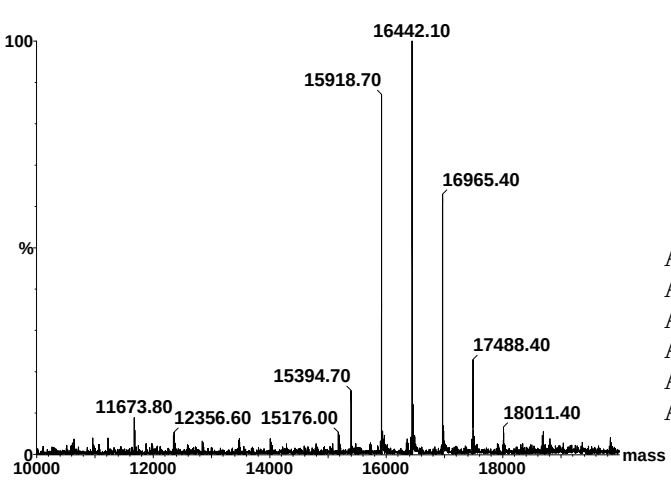
When conjugation with carboxylic acid functionalized phosphanes was done in absence of imidazole, very low coupling efficiency was observed, even when the molar excess of ligand was as high as 50. Attempts to increase the conjugation efficiency by increasing the excess ligand resulted in multiple unspecific modifications of the protein (Figure 23).

The conjugation efficiency of ligands **1-4** was improved by performing the coupling reactions in presence of 3M imidazole.[45] However it was not possible to achieve complete monomodification of the protein for any ligand.



	Calc Mass	Observed Mass
AppA	15395.1	15394.5
AppA-**3**	15683.4	15683.5
AppA-(**3**)$_2$	15971.7	15971.5
AppA-(**3**)$_3$	16260	16259.4
AppA-(**3**)$_4$	16548.3	16547.3

Figure 24 LC-MS (ES+) spectrum of AppA Y21C - **3.**
Coupling reaction done in presence of 3M imidazole using 10x molar excess of ligand **3**



	Calc Mass	Observed Mass
AppA	15395.1	15394.7
AppA-**4**	15918.7	15918.7
AppA-(**4**)$_2$	16442.2	16442.1
AppA-(**4**)$_3$	16965.8	16965.4
AppA-(**4**)$_4$	17489.3	17488.4
AppA-(**4**)$_5$	18012.9	18011.4

Figure 25 LC-MS (ES+) spectrum of AppA Y21C - **4**
Coupling reaction done in presence of 3M imidazole using 10x molar excess of ligand **4**

b. Covalent modification with N-(4-(diphenylphosphane sulfide)benzyl) maleimide

ALBP

5 µmol (50 eq.) of phosphane sulfide in 0.1 ml of dimethylformamide was added to a solution of 100 nmol of protein (1 eq.) in phosphate buffer (50 mM, 1ml, pH 7). The cloudy reaction mixture was stirred at room temperature for 16 hours. The insoluble excess modification reagent was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with degassed phosphate buffer (50 mM, pH 7.0) and concentrated to 0.5 ml. This step was repeated four times after which the

modified protein was concentrated to 1 ml and stored at 4°C.

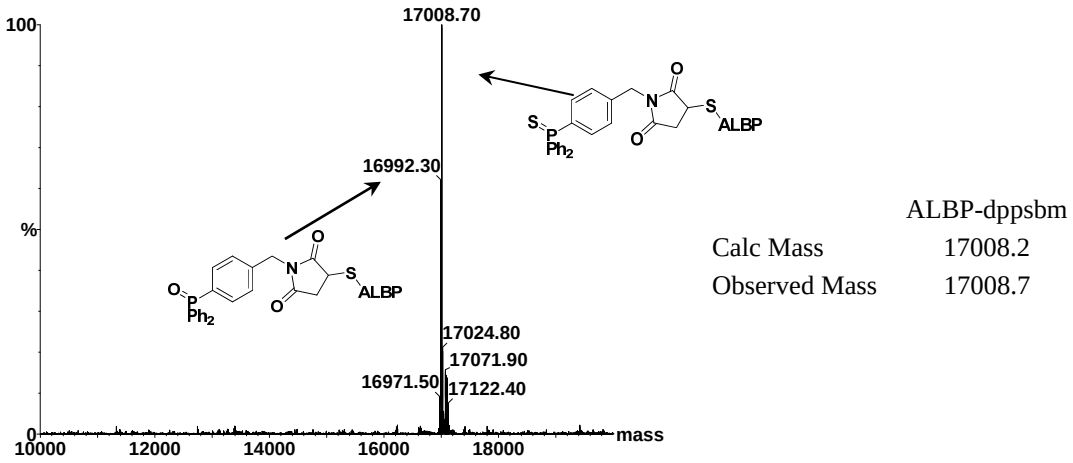


Figure 26 LC-MS (ES+) spectrum of ALBP modified with N-(4-(diphenylphosphane sulfide)benzyl) maleimide (dppsbm)

The LC-MS (ES+) spectrum of ALBP modified with N-(4-(diphenylphosphane sulfide)benzyl) maleimide (Figure 26) indicates species corresponding to the protein modified with the phosphane sulfide (17008.7), but also to the protein modified with the phosphane oxide (16992.3). Most probable the sulfur deprotection and the consequent phosphane oxidation occur during LC-MS analysis. This is suggested also by the NMR analysis of Δ98Δ modified with the phosphane sulfide, which shows only the peak corresponding to the phosphane sulfide modified protein, but no peak corresponding to the phosphane oxide (Figure 27).

Raney nickel deprotection of N-(4-(diphenylphosphane sulfide)benzyl) maleimide coupled to ALBP was tested according to a reported literature procedure.[51,87] No mass for the protein modified with the free phosphane was detected by LC-MS (ES+) after the deprotection treatment.

Δ98Δ

Prior to coupling the protein solution was degassed by purging with argon for 8 hours. Reduction of protein disulfide bridges was performed using 6x molar excess of TCEP.[88] The reduction reaction proceeded for 1 hour at room temperature and the protein solution was used directly for the next step, without further purification.
100 nmol (1 eq.) of phosphane sulfide in 2 µl of dimethylformamide was added to a solution of 100 nmol of protein (1 eq.) in Tris·HCl (50 mM, 1 ml, pH 7) in a

Schlenk-tube. The reaction mixture was stirred at room temperature for 16 hours. The insoluble excess modification reagent was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with degassed Tris·HCl (50 mM, pH 7.0) and concentrated to 0.5 ml. This step was repeated four times after which the modified protein was concentrated to 1 ml and stored at 4°C.

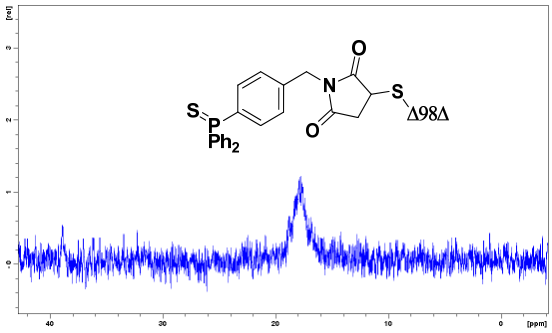


Figure 27 ^{31}P {^{1}H} NMR spectrum of Δ98Δ modified with N-(4-(diphenylphosphane sulfide)benzyl) maleimide

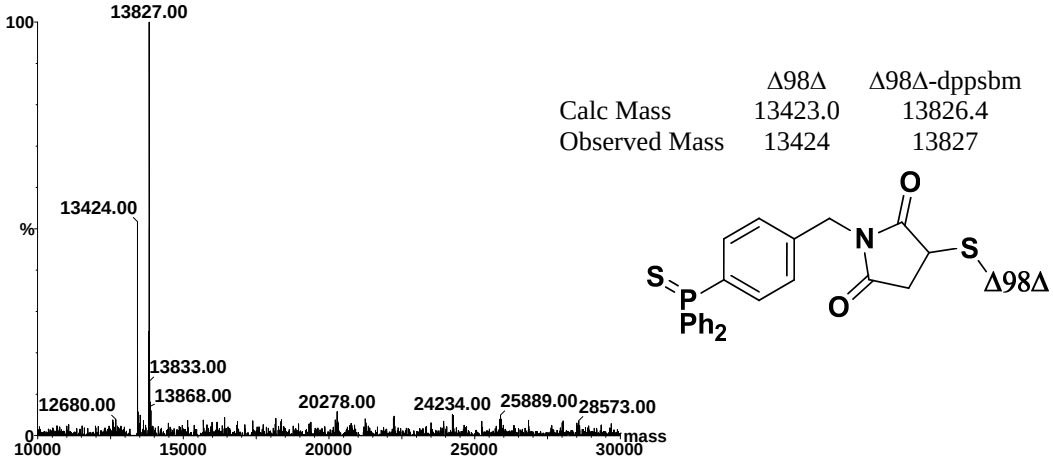
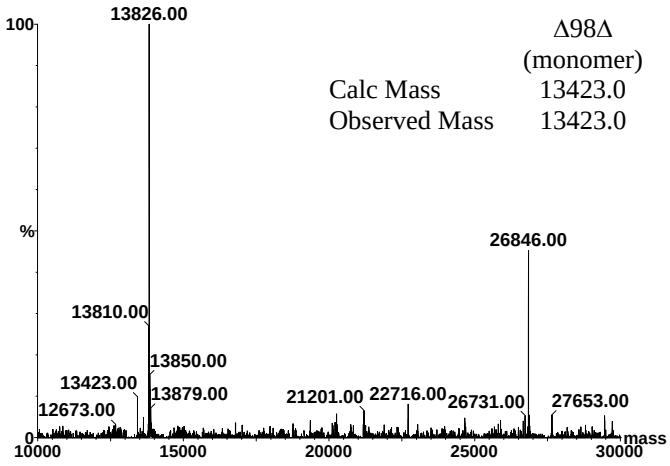


Figure 28 LC-MS (ES+) spectrum of Δ98Δ modified with N-(4-(diphenylphosphane sulfide)benzyl) maleimide (dppsbm)

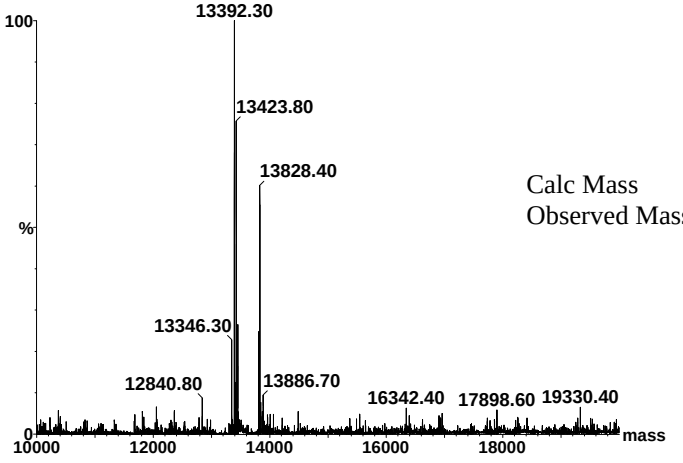
Raney nickel deprotection of N-(4-(diphenylphosphane sulfide)benzyl) maleimide coupled to Δ98Δ was tested according to a literature procedure.[51,87] No mass for the protein modified with the free phosphane was detected by LC-MS (ES+). The mass at 13810, observed after the deprotection attempt, corresponds to the protein modified with phosphane oxide, most probable formed during LC-MS analysis, as observed also for ALBP (Figure 26).



	Δ98Δ (monomer)	Δ98Δ-dppsbm	Δ98Δ (dimer)
Calc Mass	13423.0	13826.4	26846.0
Observed Mass	13423.0	13826.0	26846.0

Figure 29 LC-MS (ES+) spectrum of Δ98Δ modified with N-(4-(diphenylphosphane sulfide)benzyl) maleimide, after Raney Nickel deprotection

To check if the sulfur deprotection of the ligand coupled to the protein was inhibited by the protein environment, Raney nickel deprotection of N-(4-(diphenylphosphane sulfide)benzyl) maleimide coupled to Δ98Δ was tested also under denaturating conditions (in presence of 0.25M, 0.5M or 2M GuCl). No modification of the mass of unmodified or modified protein was observed when the deprotection reaction was done in presence of 0.25M or 0.5M GuCl. When 2M GuCl was used, the LC-MS spectrum showed a new peak at 13392.3 (Figure 30) which corresponds to the mass of unmodified protein lacking a sulphur atom. The partial unfolding of protein could facilitate the cysteine desulfurization by Raney nickel.[50]



	Δ98Δ	Δ98Δ-dppsbm
Calc Mass	13423.0	13826.4
Observed Mass	**13392.3**/13423.8	13828.4

Figure 30 LC-MS (ES+) spectrum of Δ98Δ modified with N-(4-(diphenylphosphane sulfide)benzyl maleimide, after Raney Nickel deprotection in presence of 2M GuCl

c. Coupling of maleimide functionalized azide and alkynes 5-7

ALBP

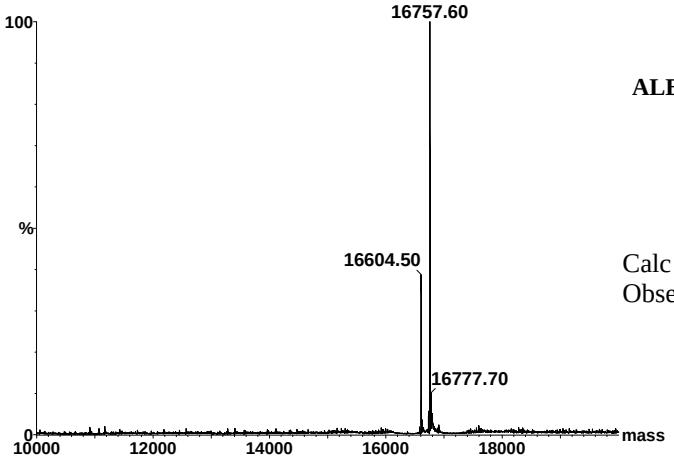
General procedure for coupling of maleimide derivatives **5-7** to ALBP

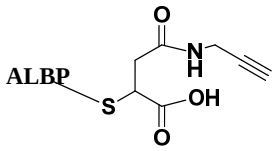
300 nmol (3 eq.) of maleimide derivative **5** in 6 µl of dimethylformamide was added to a solution of 100 nmol of protein (1 eq.) in Tris·HCl (50 mM, 1ml, pH 8). The cloudy reaction mixture was stirred at room temperature for 16 h. The insoluble excess modification reagent was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with 50 mM Tris·HCl (pH 8.0) and concentrated to 0.5 ml. This step was repeated four times after which the modified protein was concentrated to 1 ml and stored at 4°C.

Coupling of maleimides **6** and **7** was performed as described for **5**.

To compare the ionization efficiency of the modified and unmodified protein, 1 eq. of unmodified protein was added to the reaction mixture after removal of excess maleimides **5-7**. No difference in ESI-MS ionization efficiency between the modified and unmodified protein was observed for all three compounds.

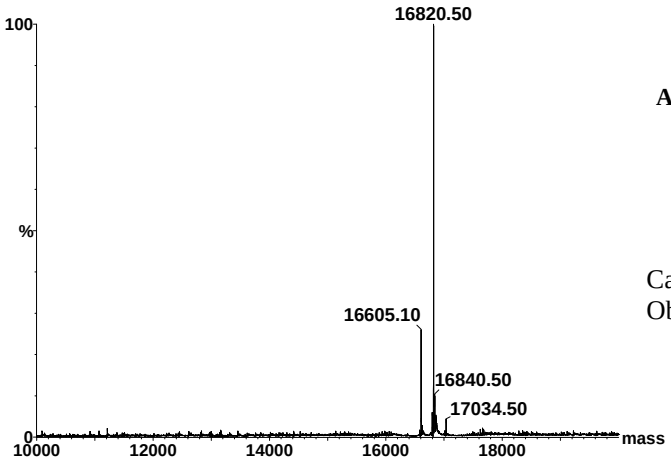
An increase of 18 Da of the masses of ALBP-**5**, ALBP-**6** and ALBP-**7** was observed by LC-MS (ES+). This is most probably caused by the hydrolytic opening of the maleimide ring.

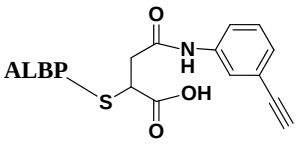




	ALBP	ALBP-**5**
Calc Mass	16604.8	16739.9
Observed Mass	16604.5	16757.6

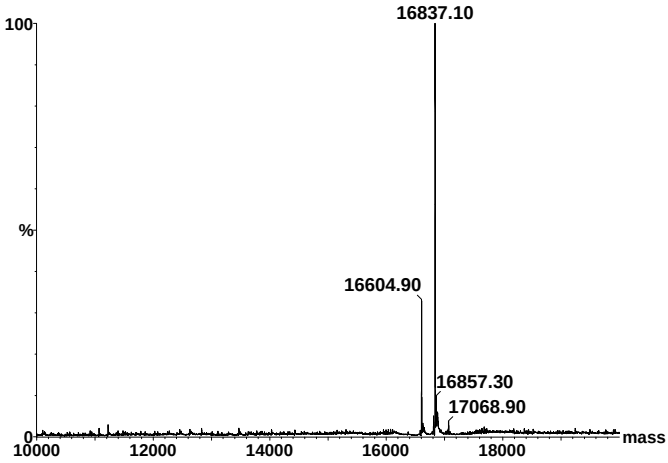
Figure 31 LC-MS (ES+) spectrum of ALBP - **5**

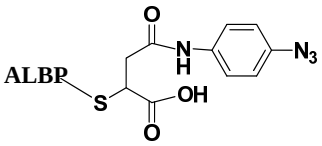




	ALBP	ALBP-**6**
Calc Mass	16604.8	16801.9
Observed Mass	16605.1	16820.5

Figure 32 LC-MS (ES+) spectrum of ALBP - **6**





	ALBP	ALBP-**7**
Calc Mass	16604.8	16818.9
Observed Mass	16604.9	16837.1

Figure 33 LC-MS (ES+) spectrum of ALBP - **7**

MALDI-MS analysis of the peptides resulted from digestion of ALBP-5 with trypsin

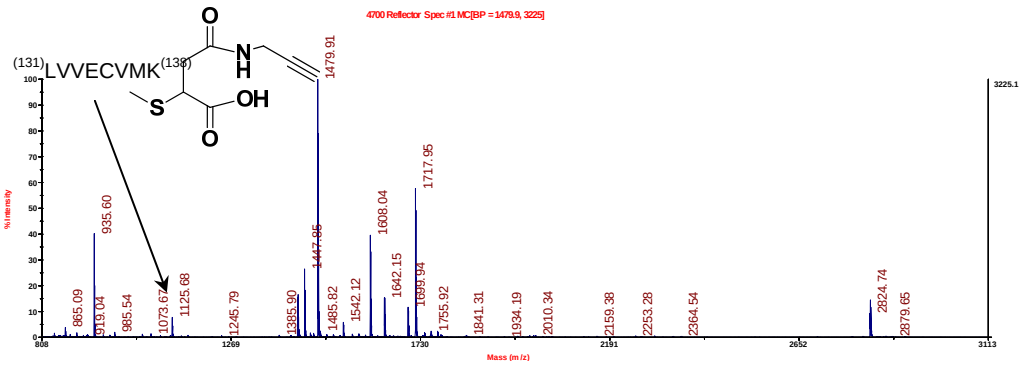


Figure 34 MALDI MS spectrum of tryptic peptides resulted from digestion of ALBP-**5**

The digest solution (0.5 µl) was applied to the MALDI target along with alpha-cyano-4-hydroxycinnamic acid matrix (0.5 µl, 10 mg/ml in 50:50 acetonitrile:0.1% TFA) and 0.1% TFA (0.5 µl) and allowed to dry. MALDI MS was acquired using a 4800 MALDI TOF/TOF Analyser (Applied Biosystems, Foster City, CA) equipped with a Nd:YAG 355 nm laser and calibrated using a mixture of peptides. The spot was initially analysed in positive MS mode between 800 and 4000 m/z, by averaging 1000 laser spots. The ion at m/z 1073.67 found in the spectrum of digested ALBP-**5** was assigned to the peptide $^{(131)}$LVVECVMK$^{(38)}$ modified with the maleamic acid adduct of **5** (Figure 34).

MALDI-MS/MS analysis of the peptide at m/z 1073.67

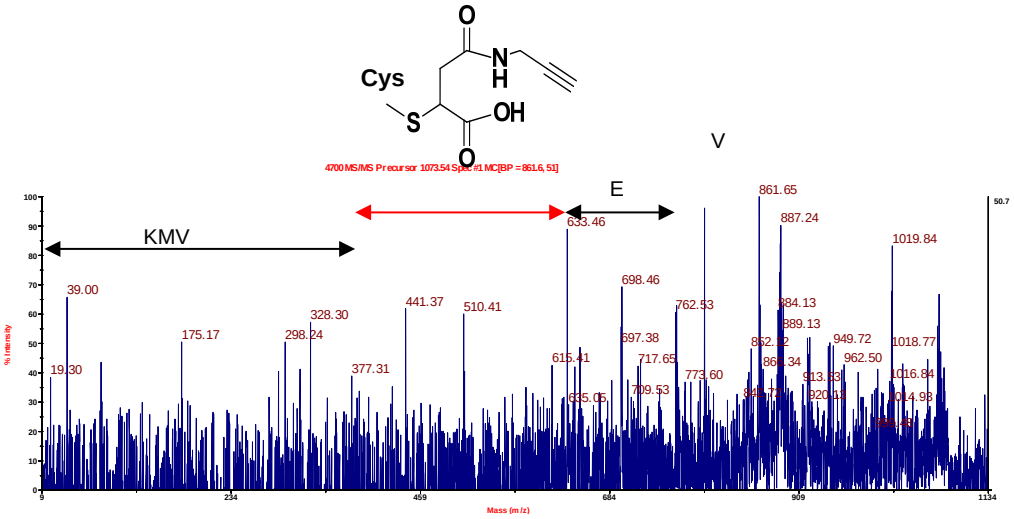


Figure 35 MALDI MS/MS spectrum of peptide Leu^{131}-Lys^{138}-**5**

The peptide Leu^{131}-Lys^{138}-**5** was selected for MS/MS (Figure 35) to confirm the cysteine modification. The MS/MS analysis was acquired to a maximum of 3000 laser shots or until the accumulated spectrum reached a S/N ratio of 35. All MS/MS data were acquired using 1 keV collision energy.

The combined MS and MS/MS data were analysed, using GPS Explorer (Applied Biosystems) to interface with the Mascot 2.1 search engine (Matrix Science, London, UK), against the UniProt (Swiss-Prot and TREMBL combined) database (April 2009). No species restriction was applied. The data was searched with tolerances of 100 ppm for the precursor ions and 0.5 Da for the fragment ions, trypsin as the cleavage enzyme,

assuming up to one missed cleavage, carbamidomethyl modification of cysteines as a fixed modification and methionine oxidation selected as a variable modification.

MALDI-MS analysis of the peptides resulted from digestion of ALBP-6 with trypsin

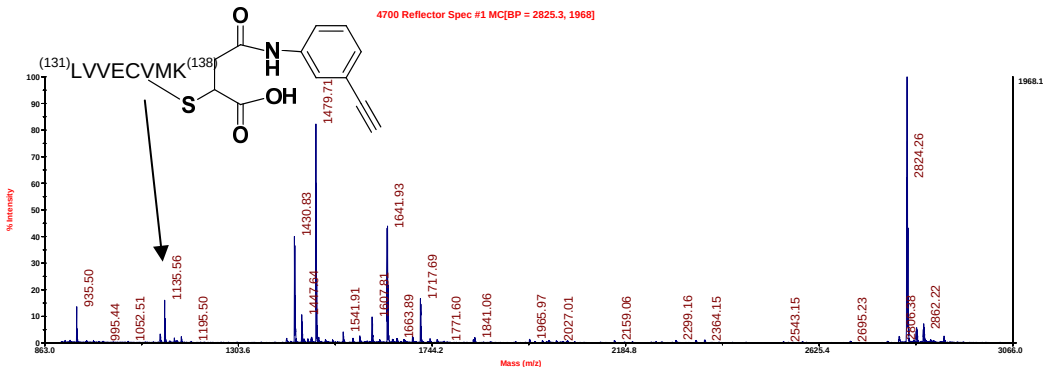


Figure 36 MALDI MS spectrum of tryptic peptides from ALBP-**6**

MALDI MS of the ALBP-**6** digest solution was performed as described for ALBP-**5**. The ion at m/z 1135.6 found in the spectrum of digested ALBP-**6** was assigned to the peptide $^{(131)}$LVVECVMK$^{(138)}$ modified with the maleamic acid adduct of **6** (Figure 36).

MALDI-MS/MS analysis of the peptide at m/z 1135.6

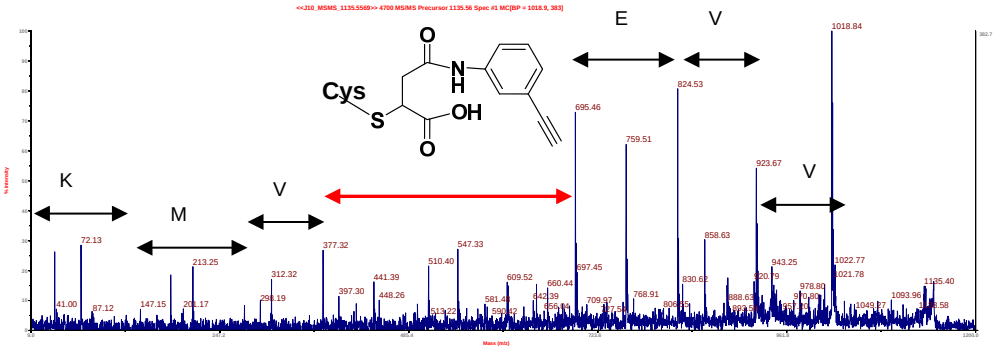


Figure 37 MALDI MS/MS spectrum of peptide Leu^{131}-Lys^{138}-**6**

The MALDI MS/MS analysis of the peptide Leu^{131}-Lys^{138}-**6** (performed as described for Leu^{131}-Lys^{138}-**5**) confirmed cysteine modification (Figure 37).

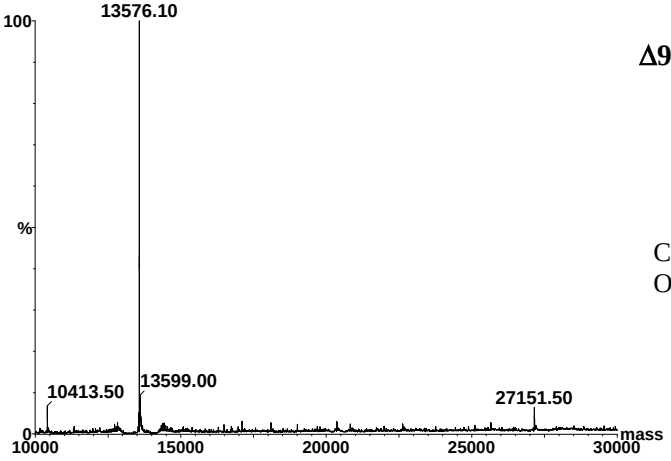
Δ98Δ

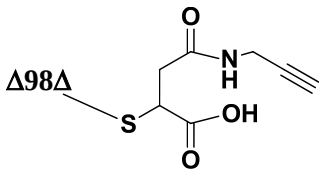
General procedure for coupling of maleimide derivatives **5-7** to Δ98Δ

1.5 µmol (15 eq.) of maleimide derivative **5** in 30 µl of dimethylformamide was added to a solution of 100 nmol of protein (1 eq.) in Tris·HCl (50 mM, 1 ml, pH 8) containing DTT (6 eq.) in a Schlenk-tube, followed by stirring at room temperature for 16 hours. The insoluble excess modification reagent was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with degassed 50 mM Tris·HCl (pH 7.0) and concentrated to 0.5 ml. This step was repeated four times after which the modified protein was concentrated to 1 ml and stored at 4°C.

Coupling of maleimides **6** and **7** was performed as described for **5**.

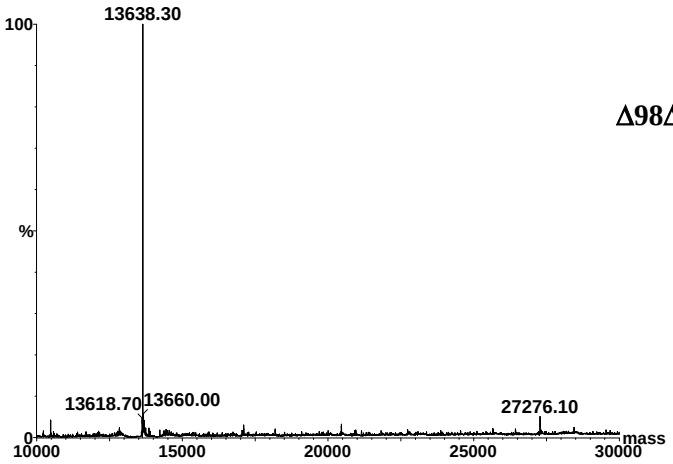
An increase of 18 Da of the mass of Δ98Δ modified with the ligands **5-7** was observed by LC-MS (ES+), most probably caused by the hydrolytic opening of the maleimide ring.





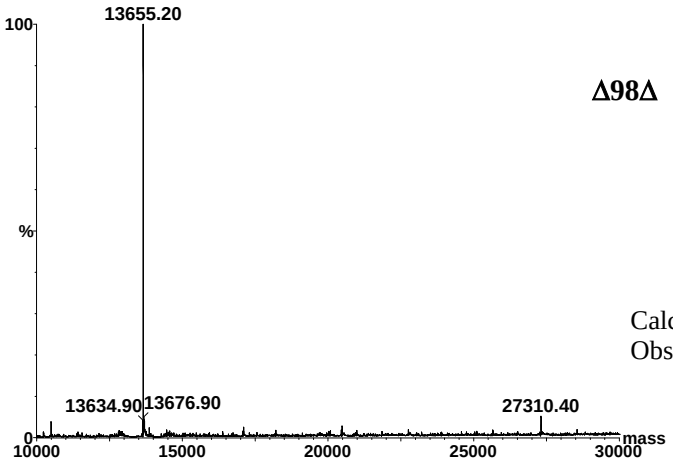
	Δ98Δ-**5**
Calc Mass	13558.1
Observed Mass	13576.1

Figure 38 LC-MS (ES+) spectrum of Δ98Δ - **5**



	Δ98Δ-**6**
Calc Mass	13620.2
Observed Mass	13638.3

Figure 39 LC-MS (ES+) spectrum of Δ98Δ - **6**



	Δ98Δ-**7**
Calc Mass	13637.2
Observed Mass	13655.2

Figure 40 LC-MS (ES+) spectrum of Δ98Δ - **7**

MALDI-MS analysis of the peptides resulted from digestion of Δ98Δ-5 with trypsin

MALDI MS of the digest solution was performed as described for ALBP-**5**. The ion at m/z 2218.1 found in the spectrum of digested Δ98Δ-**5** was assigned to the peptide $^{(80)}$EISGNELIQTCTYEGVEAK$^{(98)}$ modified with **5** and the ion at m/z 2236.8 was assigned to the same peptide modified with the maleamic acid adduct of **5** (Figure 41).

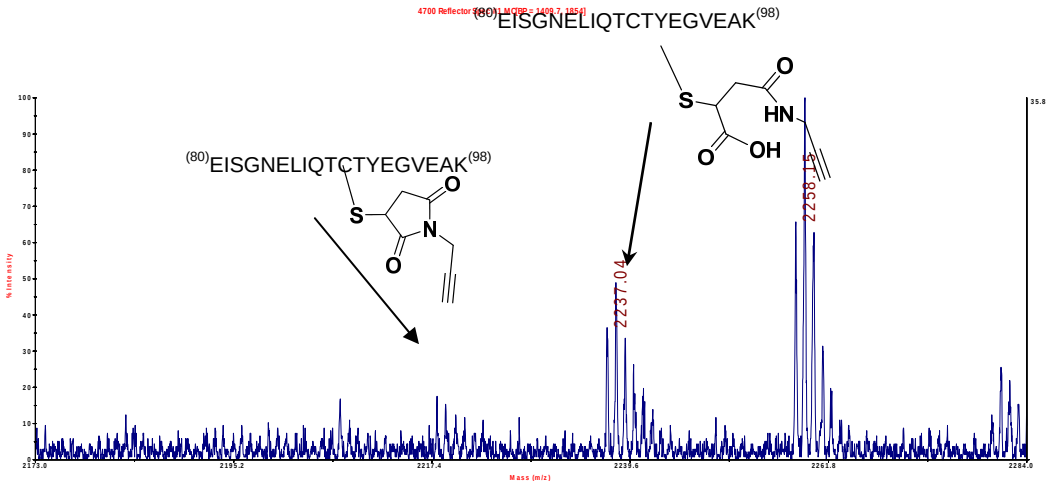


Figure 41 MALDI MS spectrum of tryptic peptides resulted from digestion of Δ98Δ-**5**

MALDI-MS/MS analysis of the peptide at m/z 2236.8

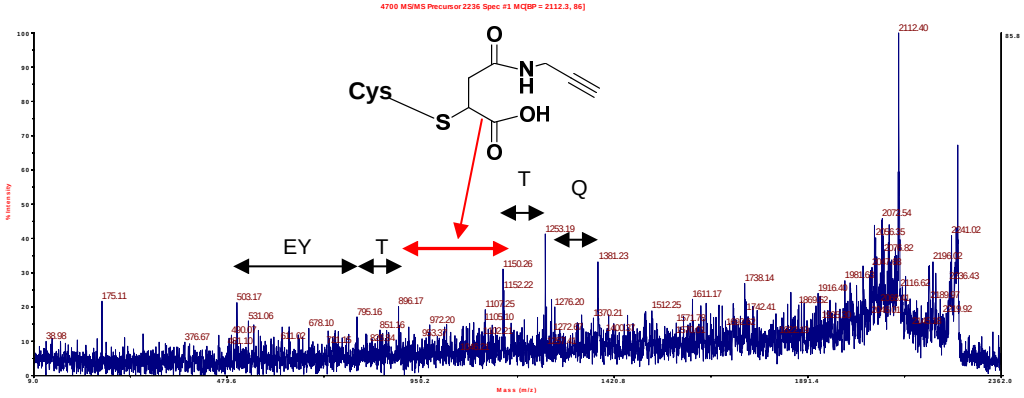


Figure 42 MALDI MS/MS spectrum of the peptide at m/z 2236.8, Glu^{80}-Lys^{98}-**5**

Cysteine modification was confirmed by the MALDI MS/MS analysis of the peptide at m/z 2236.8, Glu^{80}-Lys^{98}-**5** (Figure 42), performed as described for ALBP-**5**.

Highly selective cysteine modification with maleimide functionalized compounds and prevention of multiple modifications induced by imidazole

AppA Y21C

General procedure for coupling of maleimide derivatives **5-7** to AppA Y21C in presence of 3M imidazole

2.5 µmol (25 eq.) of maleimide derivative **5** in 50 µl of dimethylformamide was

added to a solution of 100 nmol of protein (1 eq.) in buffer (Tris·HCl 50 mM, 1ml) containing 3M imidazole at pH 7.0. The cloudy reaction mixture was stirred at room temperature for 4 h. The insoluble excess modification reagent was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with 50 mM Tris·HCl (pH 7.0) and concentrated to 0.5 ml. This step was repeated four times after which the modified protein was concentrated to 1 ml and stored at 4°C.

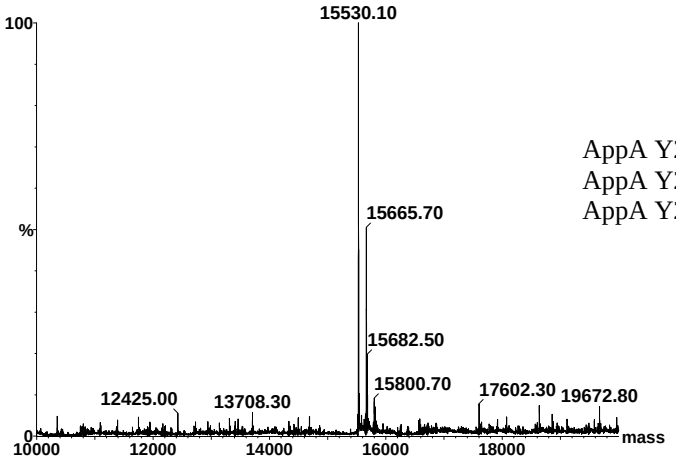
Coupling of maleimides **6** and **7** was performed as described for **5**.

Δ98Δ

General procedure for coupling of maleimide derivatives **5**-**7** to Δ98Δ in presence of 3M imidazole

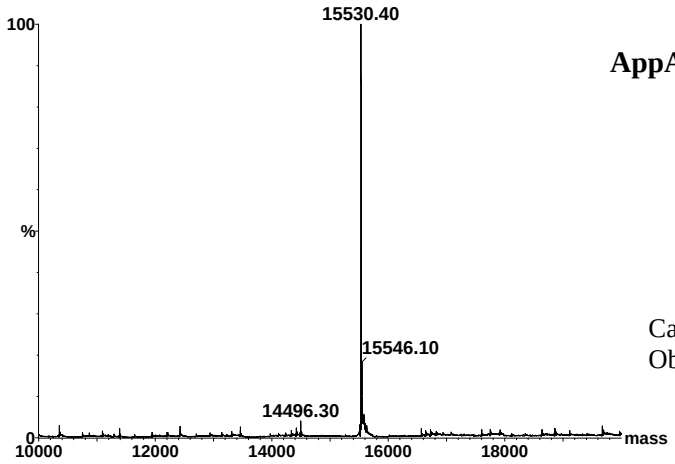
3.7 µmol (37 eq.) of maleimide derivative **5** in 74 µl of dimethylformamide was added to a solution of 100 nmol of protein (1 eq.) in buffer (Tris·HCl, 50 mM, 1 ml) containing 3M imidazole, at pH 7, in a Schlenk-tube. The reaction mixture was stirred at room temperature for 16 hours. The insoluble excess modification reagent was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with degassed 50 mM Tris·HCl (pH 7.0) and concentrated to 0.5 ml. This step was repeated four times after which the modified protein was concentrated to 1 ml and stored at 4°C.

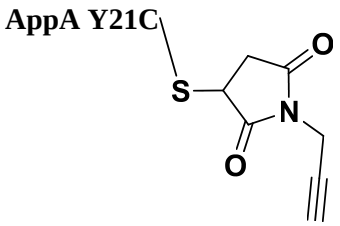
Coupling of maleimides **6** and **7** was performed as described for **5**.



	Calc Mass	Observed Mass
AppA Y21C-**5**	15530.2	15530.1
AppA Y21C-(**5**)$_2$	15665.4	15665.7
AppA Y21C-(**5**)$_3$	15800.6	15800.7

(a)

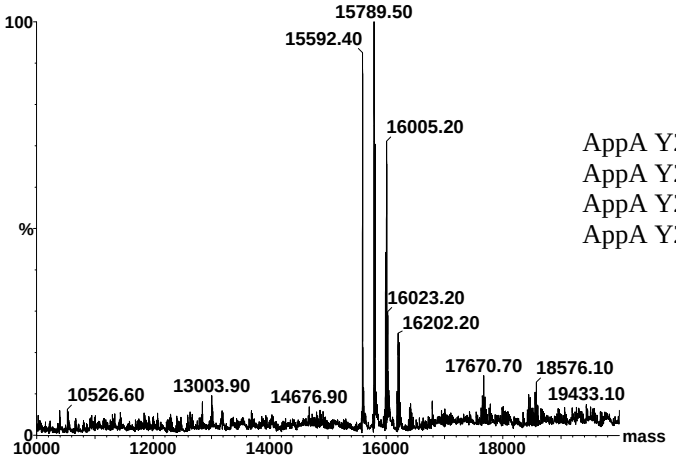




	AppA Y21C-**5**
Calc Mass	15530.2
Observed Mass	15530.4

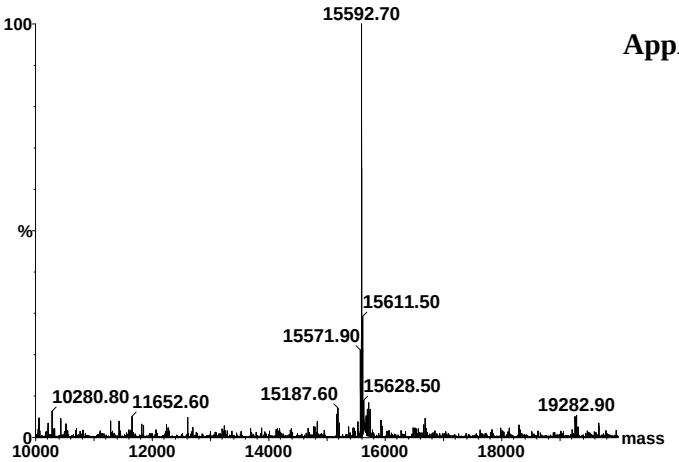
(b)

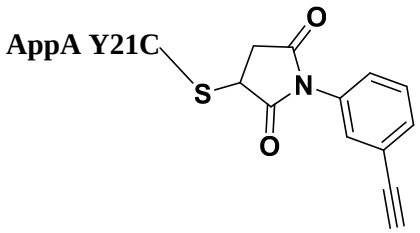
Figure 43 LC-MS (ES+) spectrum of AppA Y21C - **5** in absence (a) and in presence (b) of 3M imidazole



	Calc Mass	Observed Mass
AppA Y21C-**6**	15592.3	15592.4
AppA Y21C-$(\mathbf{6})_2$	15789.5	15789.5
AppA Y21C-$(\mathbf{6})_3$	15986.7	16005.2
AppA Y21C-$(\mathbf{6})_4$	16183.8	16202.2

(a)

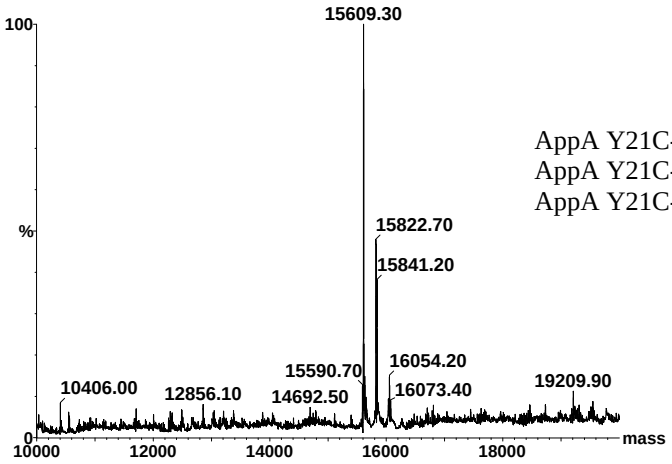




	AppA Y21C-**6**
Calc Mass	15592.3
Observed Mass	15592.7

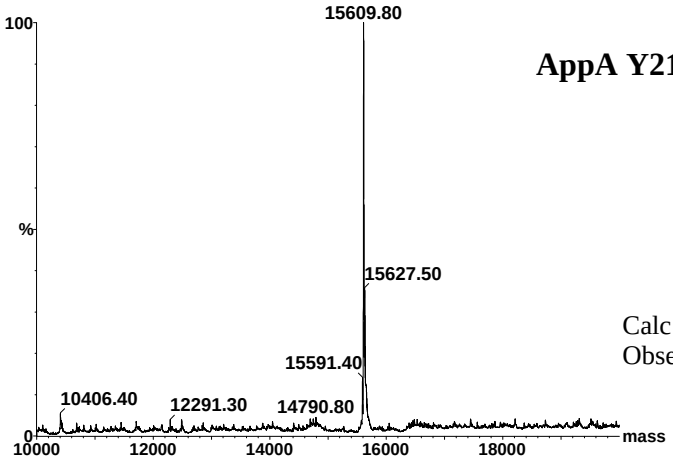
(b)

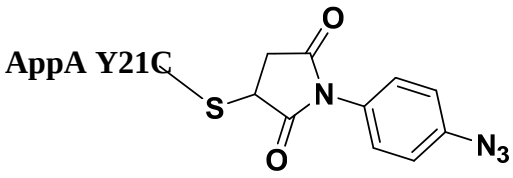
Figure 44 LC-MS (ES+) spectrum of AppA Y21C - **6** in absence (a) and in presence (b) of 3M imidazole



	Calc Mass	Observed Mass
AppA Y21C-**7**	15609.3	15609.3
AppA Y21C-(**7**)$_2$	15823.5	15822.7
AppA Y21C-(**7**)$_3$	16037.7	16054.2

(a)

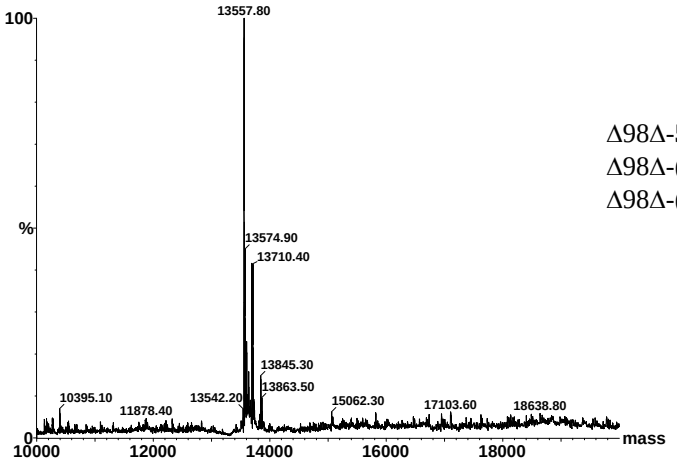




	AppA Y21C-**7**
Calc Mass	15609.3
Observed Mass	15609.8

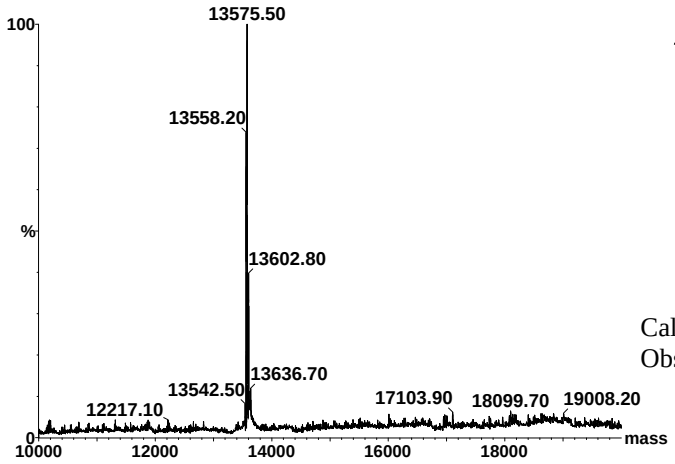
(b)

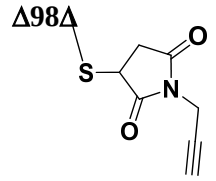
Figure 45 LC-MS (ES+) spectrum of AppA Y21C - **7** in absence (a) and in presence (b) of 3M imidazole



	Calc Mass	Observed Mass
Δ98Δ-**5**	13558.1	13557.8
Δ98Δ-(**5**)$_2$	13693.3	13710.4
Δ98Δ-(**5**)$_3$	13828.5	13845.3

(a)

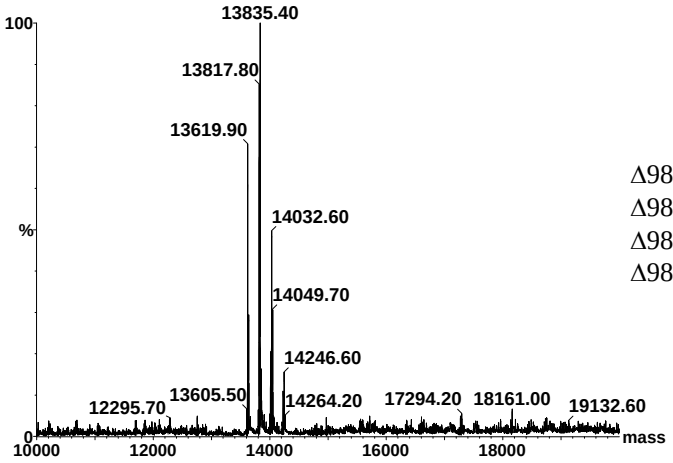




	Δ98Δ-**5**
Calc Mass	13558.1
Observed Mass	13558.2/13575.5

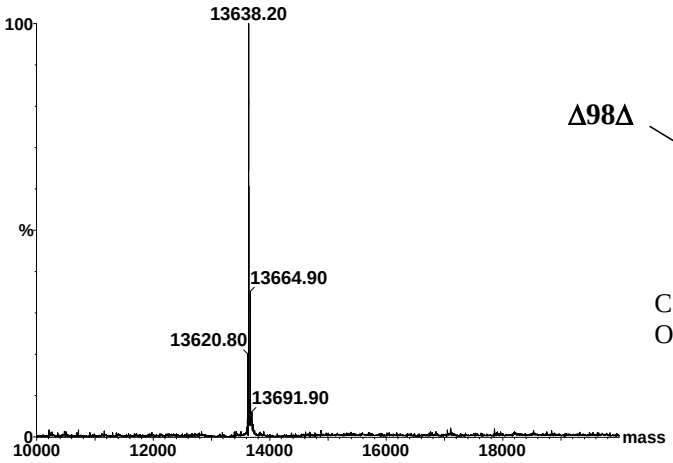
(b)

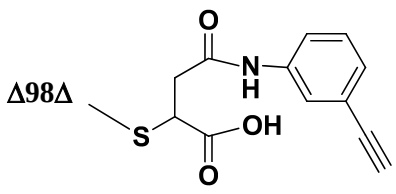
Figure 46 LC-MS (ES+) spectrum of Δ98Δ - **5** in absence (a) and in presence (b) of 3M imidazole



	Calc Mass	Observed Mass
Δ98Δ-**6**	13620.2	13619.9
Δ98Δ-$(\mathbf{6})_2$	13817.4	13817.8
Δ98Δ-$(\mathbf{6})_3$	14014.58	14032.6
Δ98Δ-$(\mathbf{6})_4$	14211.77	14246.6

(a)





	Δ98Δ-**6**
Calc Mass	13620.2
Observed Mass	13638.2

(b)

Figure 47 LC-MS (ES+) spectrum of Δ98Δ - **6** in absence (a) and in presence (b) of 3M imidazole

Multiple couplings to AppA Y21C were detected by LC-MS (ES+) when a 25-fold molar excess of **5-7** was used for the coupling reactions. However, when the reactions were done in presence of 3M imidazole, selective monomodification of cysteine was observed, despite of the high molar excess of maleimide.

To check the influence of imidazole in multiple modifications prevention, the maleimides **5-7** were coupled to Δ98Δ in presence and in absence of 3M imidazole, using also a large excess of modification reagent. Once again selective monomodification of Δ98Δ was observed when the reactions were performed in presence of imidazole, while multiple couplings were noticed for the same excess of modification reagent in absence of imidazole.

Similar results were found when a different maleimide functionalized compound, tert-butyl phenyl maleimide, was coupled to AppA Y21C and Δ98Δ in presence and in absence of imidazole using a large excess of modification reagent.

d. Attempts to introduce BH_3 protected phosphanes into alkyne and azide functionalized proteins via click- reaction

Several reported literature procedures were adapted in order to modify our library of alkyne/azide functionalized proteins via click- reaction.

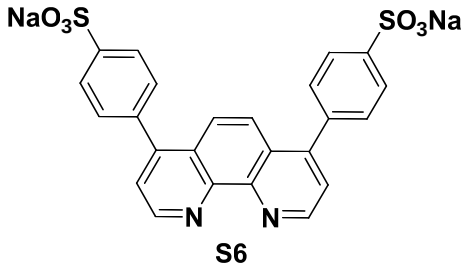
1. A procedure reported by Hatzakis and co-workers[89] for protein labelling was adapted to test the incorporation of alkyne functionalized phosphane borane **8** to ALBP-**7** and Δ98Δ-**7**.

0.5 μmol (20 eq.) of phosphane-borane **8** in 50 μl of tetrahydrofurane was added to a solution of 25 nmol of azide modified protein (1 eq.) in phosphate buffer (20 mM, 0.5 ml, pH 7.2), followed by the addition of catalyst solution (2.5 μmol of $CuSO_4x5H_2O$, 12.5 μmol of Na-ascorbate) in phosphate buffer (20 mM, 0.2 ml, pH 7.2). The reaction mixture was stirred for 24 hours. The insoluble excess modification reagent and Cu(II) salts were removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with degassed 20 mM phosphate buffer (pH 7.2) and concentrated to 0.5 ml. This

step was repeated four times after which the modified protein was concentrated to 0.75 ml and stored at 4°C.

LC-MS was performed before and after the washing step, but no protein signal could be detected. 50% of the initial amount of Δ98Δ and 25% of the initial amount of ALBP was found in solution after the washing step (Bradford assay[68]).

2. A procedure reported by Finn and co-workers involves the water-soluble ligand **S6** (sulfonated bathophenanthroline) as promoter of Cu(I) mediated azide-alkyne cycloaddition.[56] This procedure was tested for incorporation of phosphane-boranes **8** and **9** to ALBP and Δ98Δ modified with **5**, **6** or **7**.



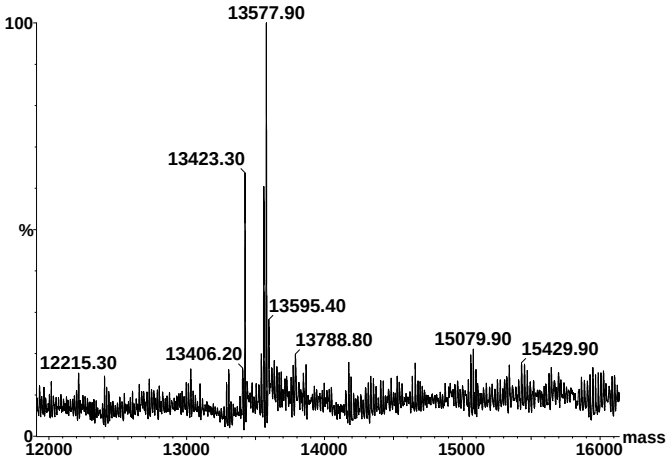
S6

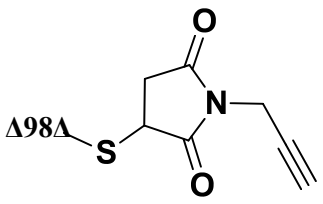
0.8 μmol (10 eq.) of phosphane-borane **8** in 5 μl of tetrahydrofurane was added to a solution of 80 nmol of azide modified protein (1 eq.) in 50 mM Tris·HCl (1.232 ml, pH 8), followed by the addition of preformed catalyst solution, synthesized by mixing 40 μmol of CuBr in 0.288 ml acetonitrile with 120 μmol of ligand **S6** in 0.480 ml Tris·HCl, pH 8, under argon. The reaction mixture was stirred for 16 hours.

Proteins stability was verified at the end of the reaction by SDS-PAGE and Bradford assay. Both proteins proved to be stable under the reaction conditions. However, LC-MS of the reaction mixture did not indicate any protein signal.

The precipitated Cu(II) salts and the insoluble excess phosphane were removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was desalted on a PD 10 column (GE Healthcare) and eluted with demineralised water (3.5 ml). The eluent was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with 20 mM EDTA solution[90] in 50 mM Tris·HCl (pH 8) and concentrated to 0.5 ml. This step was repeated two times. After that the protein solution was diluted twice to 10 ml with 50 mM Tris·HCl (pH 8), concentrated to 0.5 ml and characterised by Bradford assay (64% of the initial amount of protein was left in solution), SDS-Page and LC-MS (ES+).

The attempt to couple the phosphane-borane **9** via click-reaction to alkyne modified proteins was performed as described for **8**.



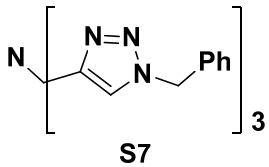


	Δ98Δ	Δ98Δ-**5**
Calc Mass	13423	13558.1
Observed Mass	13423.3	13577.9

Figure 48 LC-MS (ES+) spectrum of Δ98Δ - **5**, after the click-reaction done in presence of **S6** followed by washing with EDTA

The only masses detected by LC-MS (ES+) after the click reaction done according to this procedure were the ones corresponding to the mass of apo-proteins and of the protein modified with the maleimide functionalized compounds **5-7** (Figure 48).

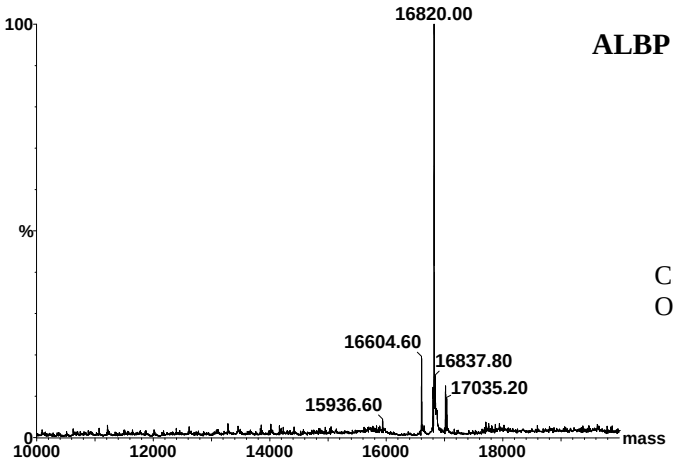
3. A procedure used by Finn and co-workers for the modification of an alkyne/azide functionalized virus via Cu(I) catalyzed azide-alkyne [3+2] cycloaddition is involving the ligand **S7** (tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine).[55] I adapted this procedure in an attempt to incorporate the phosphane-borane **9** to ALBP modified with **5** or **6.**

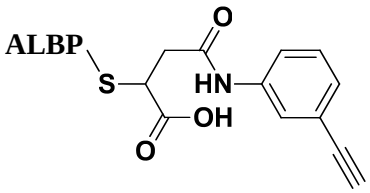


S7

400 nmol (10 eq) of phosphane-borane **9** in 4 µl of tetrahydrofurane was added to a solution of 40 nmol of alkyne modified ALBP (1 eq.) in 50 mM Tris·HCl (1 ml, pH 8), followed by the addition of 2 µmol of TCEP in 0.200 ml Tris·HCl and 2 µmol of **S7** in 40 µl dimethyl sulfoxide. 1 µmol of $CuSO_4$ in Tris·HCl (250 µl, pH 8) was then added. The reaction mixture was stirred for 16 hours. The precipitated Cu(II) salts and the insoluble excess phosphane were removed by centrifugation (49,000 g, 4°C, 10 min).

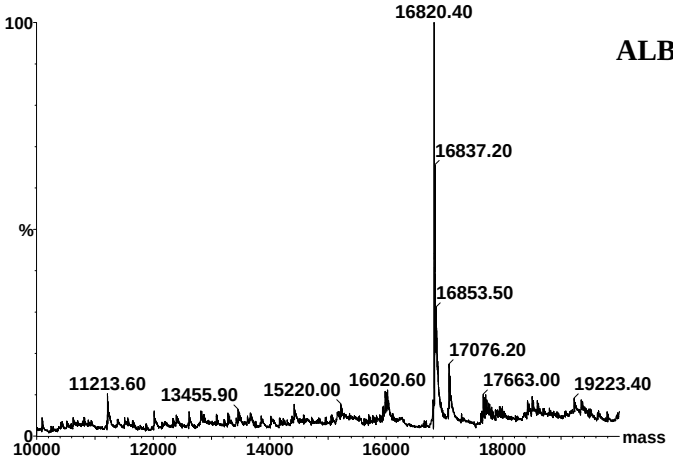
The supernatant was desalted on a PD 10 column (GE Healthcare) and eluted with demineralised water (3.5 ml). The eluent was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with 20 mM EDTA solution[90] in 50 mM Tris·HCl (pH 8) and concentrated to 0.5 ml. This step was repeated two times. After that the protein solution was diluted twice to 10 ml with 50 mM Tris·HCl (pH 8) and concentrated to 0.5 ml. At the end the protein was characterised by Bradford Assay (46% of the initial amount of protein was left in solution), SDS-PAGE and LC-MS (ES+). LC-MS spectra showed only the peak corresponding to the alkyne modified protein.

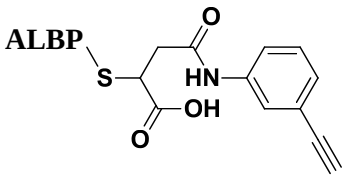




	ALBP-**6**
Calc Mass	16801.9
Observed Mass	16820.0

(a)





	ALBP-**6**
Calc Mass	16801.9
Observed Mass	16820.4

(b)

Figure 49 LC-MS (ES+) spectrum of ALBP - **6** before (a) and after (b) the click-reaction done in presence of **S7**

In conclusion, when the attempts to couple the phosphane-boranes **8** and **9** to the modified proteins via click-reaction were performed in presence of Cu(I) it was not

possible to asses the product formation via LC-MS, due to possible suppression of the protein signal by Cu(I). The use of nitrogen ligands capable to coordinate and stabilize the metal appears to prevent this suppression during the LC-MS analysis, however masses corresponding to the click modified proteins were not observed. This suggests that either the click reaction did not work for the tested conditions, or the click modified proteins were not stable for the tested reaction conditions.

e. Covalent modification with maleimide functionalized phosphane-boranes 10-12

ALBP

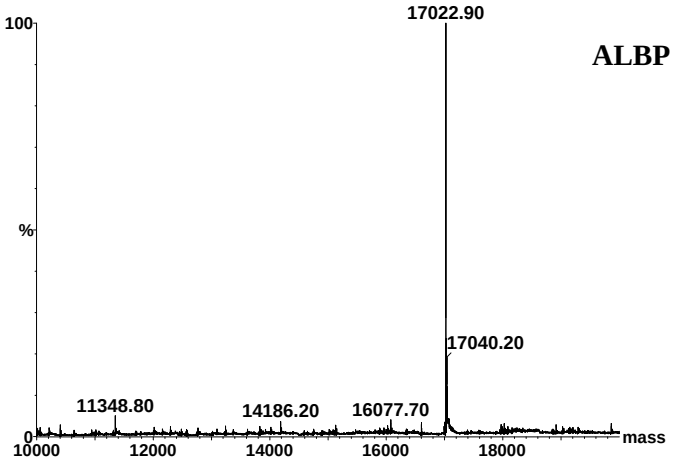
Prior to coupling the protein solution was degassed by purging with argon for 8 hours.

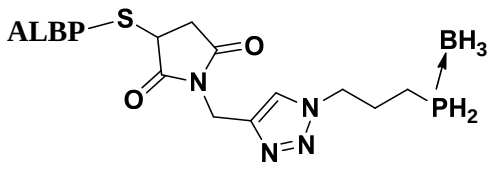
1. Conjugation with 10

312.5 nmol (25 eq.) of click-compound **10** in 6.25 µl of dimethylformamide was added to a solution of 12.5 nmol of protein (1 eq.) in Tris·HCl (50 mM, 0.25 ml, pH 8). The reaction mixture was stirred at room temperature for 5h. The insoluble fraction of excess ligand was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Vivaspin 2, 10,000 MWCO), diluted to 2 ml with degassed Tris·HCl (50 mM, pH 8.0) and concentrated to 0.1 ml. This step was repeated four times after which the modified protein was concentrated to 0.25 ml and stored at 4°C.

2. Conjugation with 11 and 12

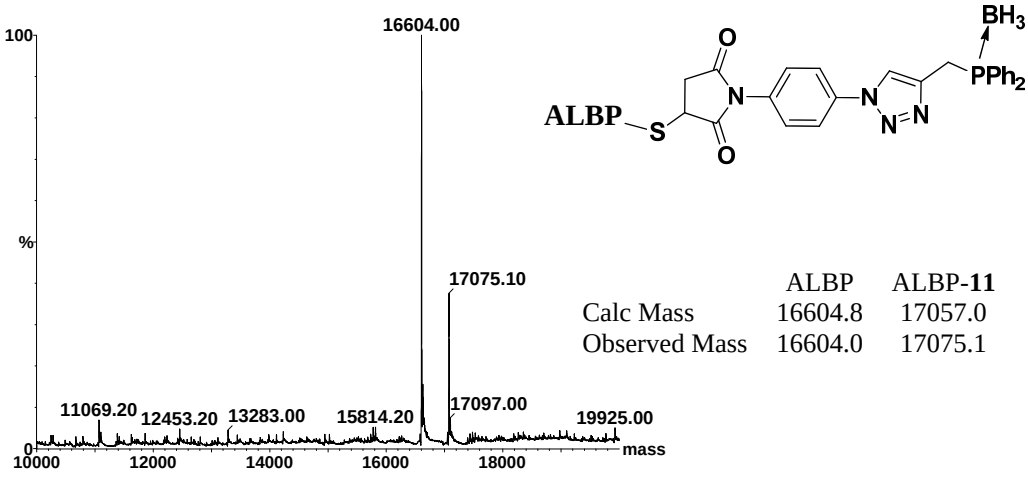
Conjugation with **11** and **12** was performed as described for **10**.





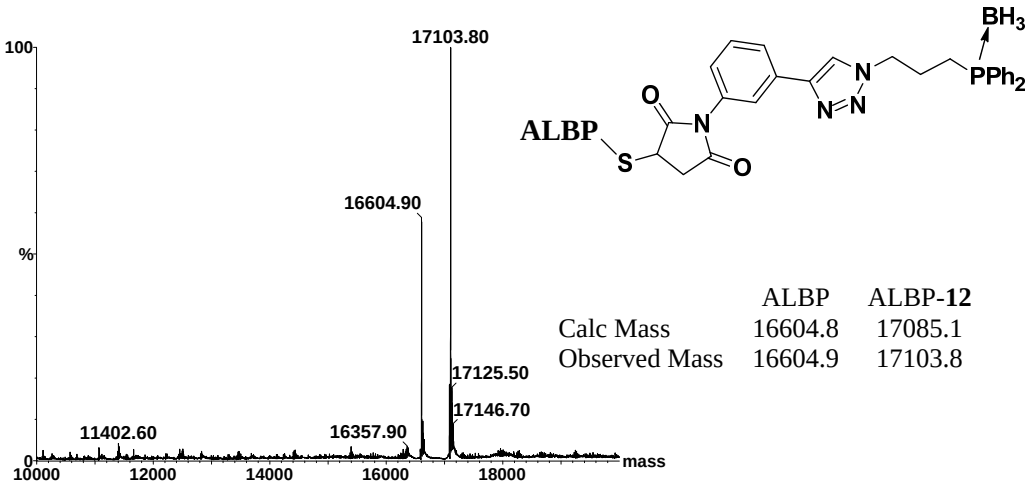
	ALBP-**10**
Calc Mass	17023.0
Observed Mass	17022.9

Figure 50 LC-MS (ES+) spectrum of ALBP - **10**



	ALBP	ALBP-**11**
Calc Mass	16604.8	17057.0
Observed Mass	16604.0	17075.1

Figure 51 LC-MS (ES+) spectrum of ALBP - **11**



	ALBP	ALBP-**12**
Calc Mass	16604.8	17085.1
Observed Mass	16604.9	17103.8

Figure 52 LC-MS (ES+) spectrum of ALBP - **12**

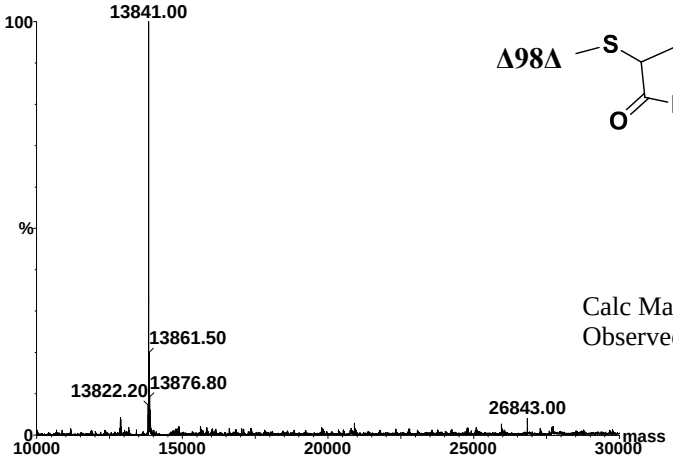
Δ98Δ

1. Conjugation with 10

1.85 µmol (37 eq.) of click-compound **10** in 37 µl of dimethylformamide was added to a solution of 50 nmol of protein (1 eq.) in Tris·HCl (50 mM, 1 ml, pH 8.0) containing DTT (6 eq.) in a Schlenk-tube, followed by stirring at room temperature for 9 h. The insoluble excess modification reagent was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with degassed Tris·HCl (50 mM, pH 8.0) and concentrated to 0.5 ml. This step was repeated four times after which the modified protein was concentrated to 1 ml and stored at 4°C.

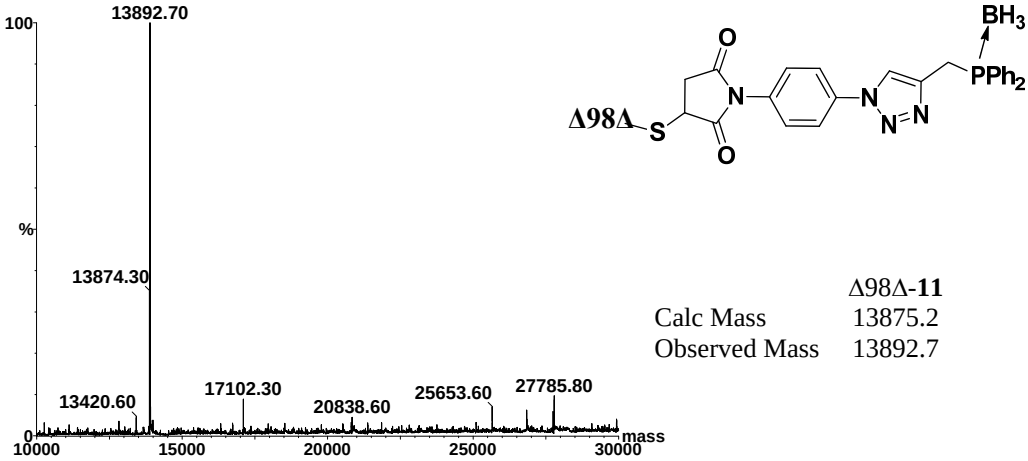
2. Conjugation with 11 and 12

Conjugation with **11** and **12** was performed as described for **10**.



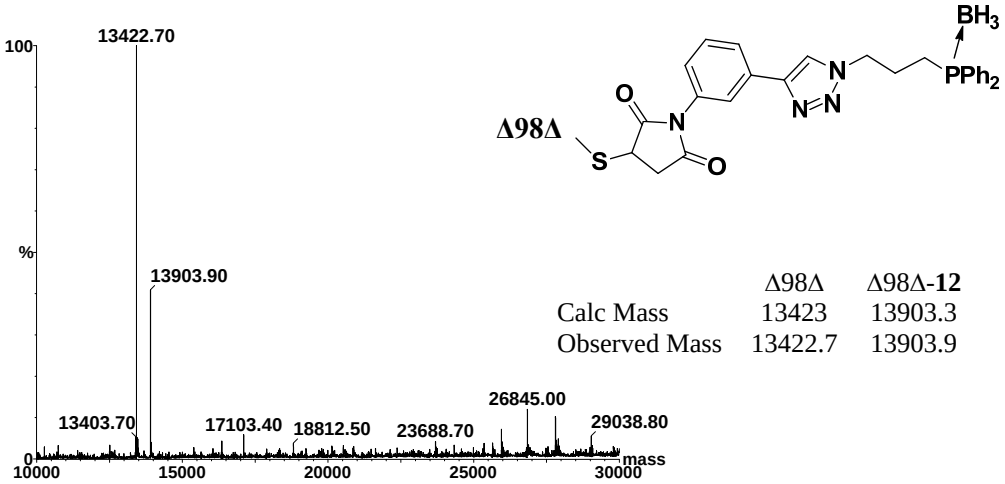
	Δ98Δ-**10**
Calc Mass	13841.2
Observed Mass	13841.0

Figure 53 LC-MS (ES+) spectrum of Δ98Δ - **10**



	Δ98Δ-**11**
Calc Mass	13875.2
Observed Mass	13892.7

Figure 54 LC-MS (ES+) spectrum of Δ98Δ - **11**



	Δ98Δ	Δ98Δ-**12**
Calc Mass	13423	13903.3
Observed Mass	13422.7	13903.9

Figure 55 LC-MS (ES+) spectrum of Δ98Δ - **12**

AppA Y21C

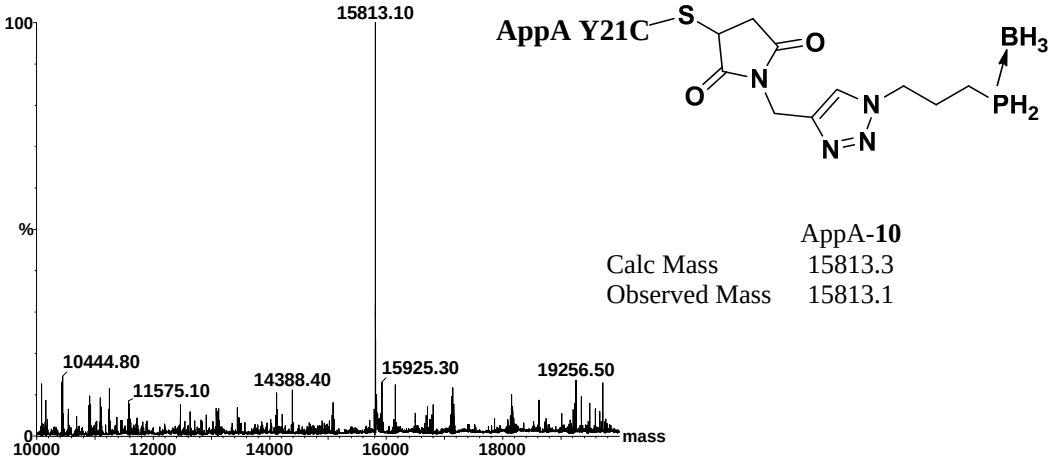
Prior to coupling the protein solution was degassed by purging with argon for 8 hours.

1. Conjugation with 10

312.5 nmol (25 eq.) of click-compound **10** in 6.25 µl of dimethylformamide was added to a solution of 12.5 nmol of protein (1 eq.) in buffer (Tris·HCl 50 mM, 0.25 ml) containing 3M imidazole at pH 7.0. The reaction mixture was stirred at room temperature for 5h. The insoluble fraction of excess ligand was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Vivaspin 2, 10,000 MWCO), diluted to 2 ml with degassed Tris·HCl (50 mM, pH 7.0) and concentrated to 0.1 ml. This step was repeated four times after which the modified protein was concentrated to 0.25 ml and stored at 4°C.

2. Conjugation with 11 and 12

Conjugation with **11** and **12** was performed as described for **10**.



AppA-**10**	
Calc Mass	15813.3
Observed Mass	15813.1

Figure 56 LC-MS (ES+) spectrum of AppA Y21C - **10**

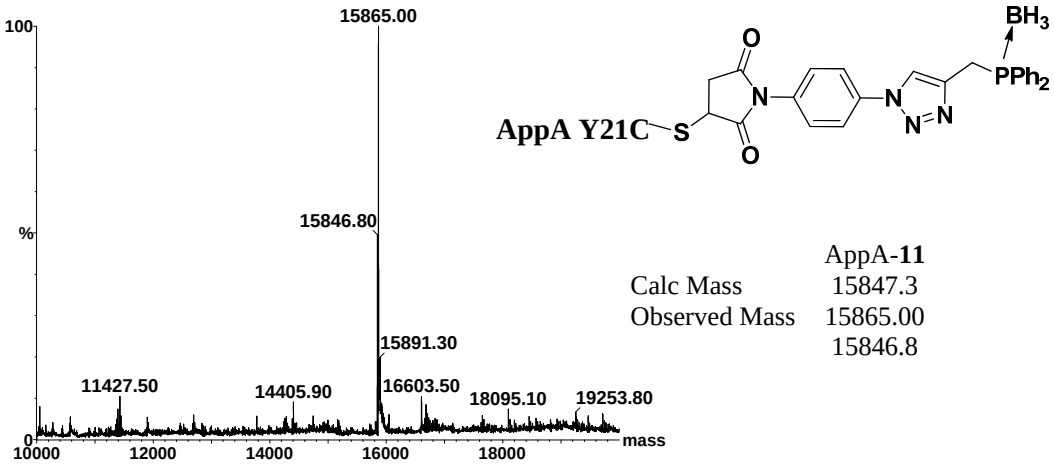


Figure 57 LC-MS (ES+) spectrum of AppA Y21C – **11**

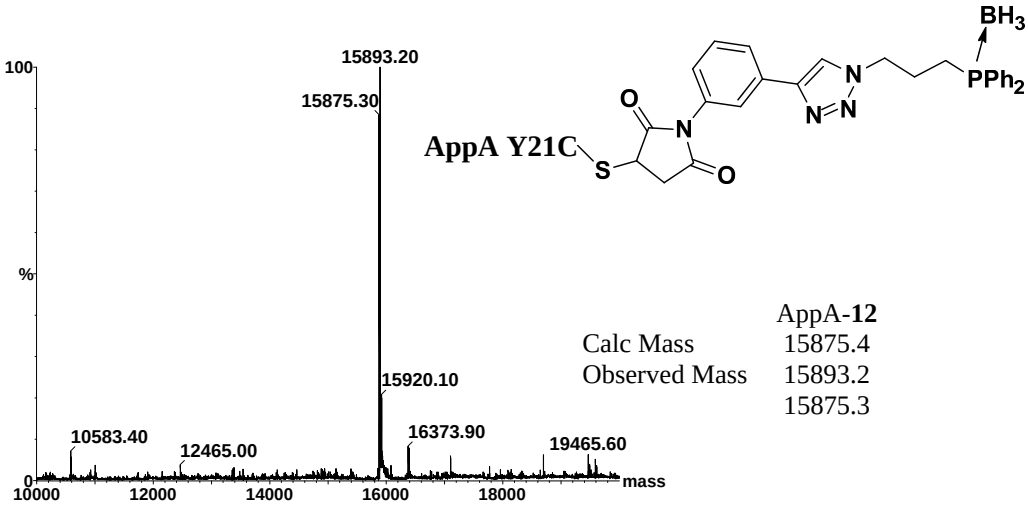


Figure 58 LC-MS (ES+) spectrum of AppA Y21C - **12**

f. Procedures tested for deprotection of phosphane-boranes coupled to proteins

<u>Ethanolamine</u>[58]

ALBP

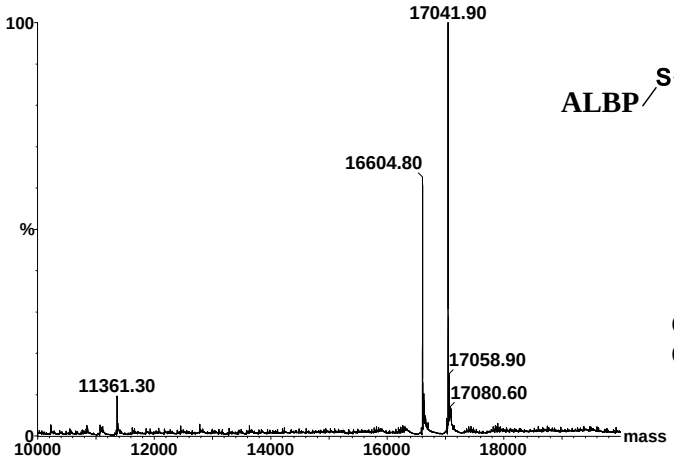
a. A solution of 12.5 nmol of protein modified with **10** in 50 mM Tris·HCl (0.25 ml, pH 8) was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 15 ml with degassed 20 mM ethanolamine in 50mM Tris-HCl (pH 8) and concentrated to 0.5 ml. The procedure was repeated and the protein solution was finally concentrated to 0.25 ml and analysed by LC-MS (ES+).

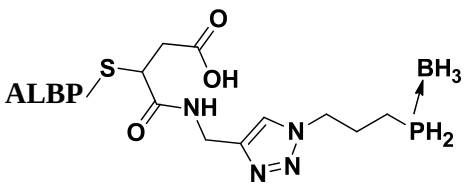
b. 15 ml of degassed 100 mM ethanolamine in 50mM Tris-HCl (pH 8) were

added to a solution of 12.5 nmol of protein modified with **10** in 50 mM Tris·HCl (0.25 ml, pH 8) in a Schlenk-tube. The reaction was stirred for 2 hours at room temperature. The protein solution was concentrated to 0.25 ml and analysed by LC-MS (ES+).

c. 15 ml of degassed 200 mM ethanolamine in 50mM Tris-HCl (pH 8) were added to a solution of 12.5 nmol of protein modified with **10** in 50 mM Tris·HCl (0.25 ml, pH 8). The reaction was stirred for 24 hours at room temperature. The protein solution was concentrated to 0.25 ml and analysed by LC-MS (ES+).

No mass corresponding to the protein modified with free phosphane was observed by LC-MS after all the tested procedures.





	ALBP-**10**
Calc Mass	17023.0
Observed Mass	17041.9

Figure 59 LC-MS (ES+) spectrum of ALBP - **10** after decomplexation reaction with 200 mM ethanolamine**.**

TCEP

ALBP

a. A solution of 12.5 nmol of protein modified with **10** in Tris·HCl (50 mM, 0.25 ml, pH 8) was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 15 ml with degassed 20 mM TCEP in Tris-HCl (50mM, pH 8) and concentrated to 0.5 ml. The procedure was repeated and the protein solution was finally concentrated to 0.25 ml and analysed by LC-MS (ES+).

b. 15 ml of degassed 100 mM TCEP in Tris-HCl (50mM, pH 8) were added to a solution of 12.5 nmol of protein modified with **10** in Tris·HCl (50 mM, 0.25 ml, pH 8) in a Schlenk-tube, followed by stirring for 2 hours at room

temperature. The protein solution was finally concentrated to 0.25 ml and analysed by LC-MS (ES+).

c. 15 ml of degassed 200 mM TCEP in Tris-HCl (50mM, pH 8) were added to a solution of 12.5 nmol of protein modified with **10** in Tris·HCl (50 mM, 0.25 ml, pH 8) in a Schlenk-tube, followed by stirring for 24 hours at room temperature. The protein solution was finally concentrated to 0.25 ml and analysed by LC-MS (ES+).

No mass corresponding to protein modified with free phosphane was observed by LC-MS for any of the tested procedures.

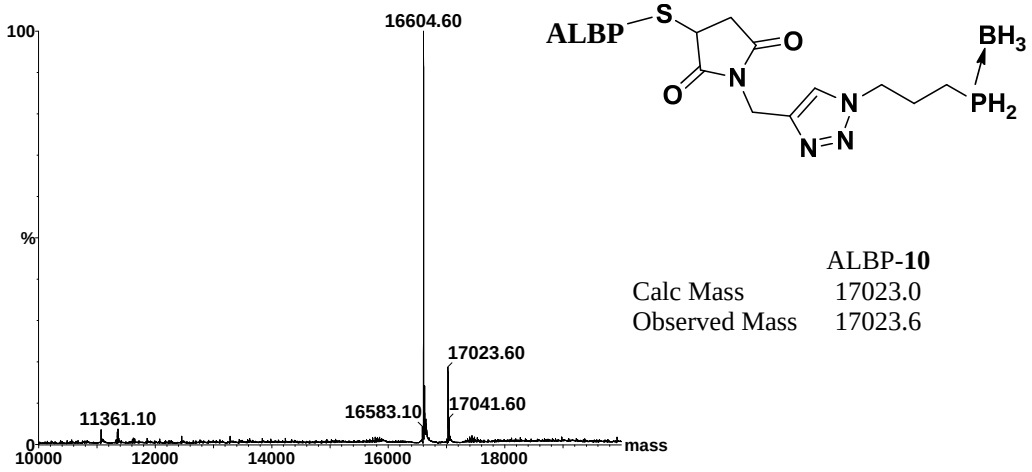


Figure 60 LC-MS (ES+) spectrum of ALBP - **10** after decomplexation reaction with TCEP (200 mM).

DABCO[59]

ALBP, Δ98Δ

a. 15 ml of degassed 100 mM DABCO in phosphate buffer (50mM, pH 9.5) were added to a solution of 12.5 nmol of protein modified with **10** in Tris·HCl (50 mM, 0.25 ml, pH 8) in a Schlenk-tube, followed by stirring for 5 hours at room temperature. The protein solution was finally concentrated to 0.25 ml and analysed by LC-MS (ES+).

b. The same procedure was tested using a 200 mM DABCO in phosphate buffer (50mM, pH 9.9).

No mass corresponding to protein modified with free phosphane was observed by LC-MS after the tested procedures.

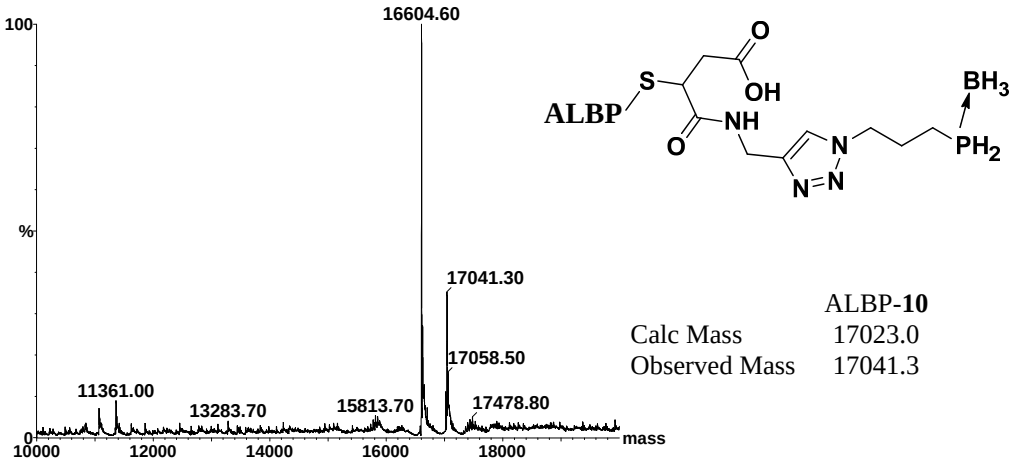


Figure 61 LC-MS (ES+) spectrum of ALBP - **10** after decomplexation reaction with DABCO

DABCO treatment under denaturating conditions[59]

ALBP

15 ml of degassed 200 mM DABCO and 6M GuCl in phosphate buffer (50mM, pH 9.9) were added to a solution of 12.5 nmol of protein modified with **10** in Tris·HCl (50 mM, 0.25 ml, pH 8) in a Schlenk-tube, followed by stirring for 16 hours at room temperature. The protein solution was finally concentrated to 0.25 ml and analysed by LC-MS (ES+). No mass corresponding to protein modified with free phosphane was observed.

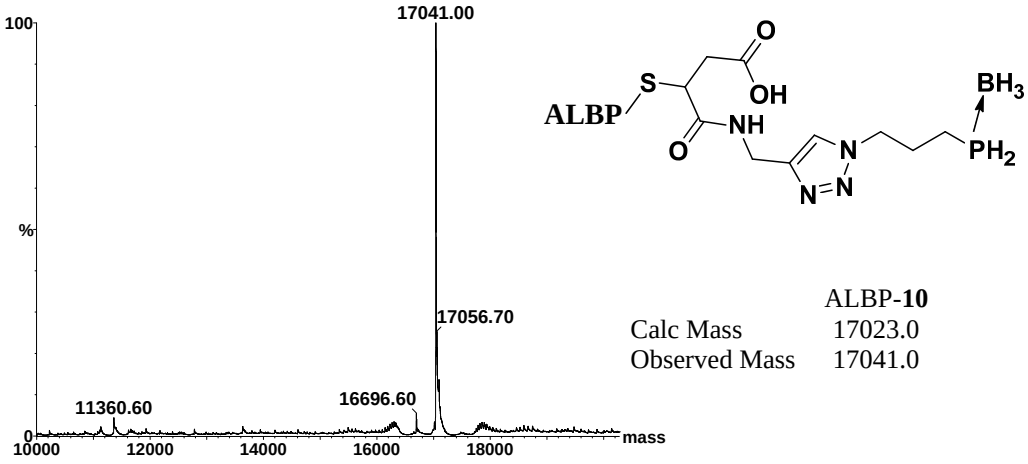


Figure 62 LC-MS (ES+) spectrum of ALBP - **10** after decomplexation reaction with DABCO under denaturating conditions (6M GuCl)

DABCO treatment in presence of KF[60,91]

ALBP

15 ml of 200 mM DABCO and 6M GuCl in phosphate buffer (50mM, pH 9.9) and 0.15 ml of saturated KF in water were added to a solution of 12.5 nmol of protein modified with **10** in Tris·HCl (50 mM, 0.25 ml, pH 8) in a Schlenk-tube, followed by stirring for 16 hours at room temperature. The protein solution was finally concentrated to 0.25 ml and analysed by LC-MS. No mass corresponding to protein modified with free phosphane was observed.

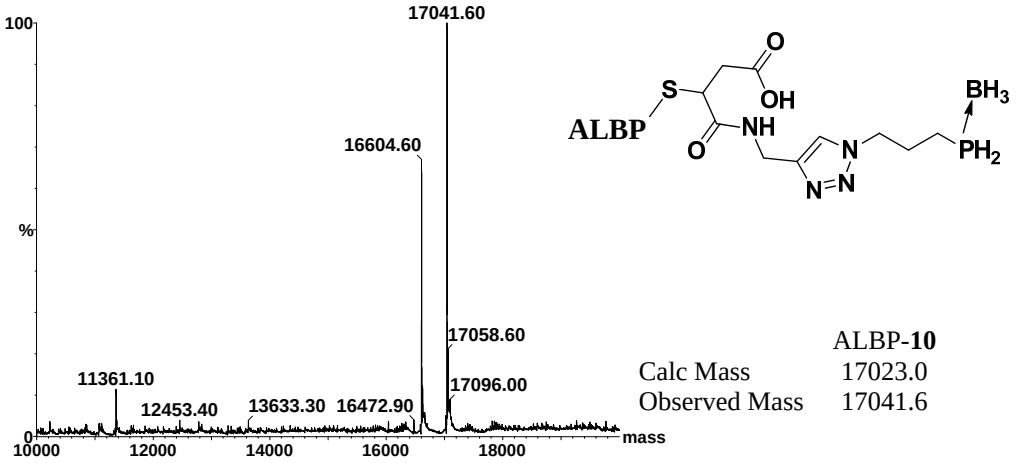


Figure 63 LC-MS (ES+) spectrum of ALBP - **10** after decomplexation with DABCO under denaturating conditions, in presence of KF

Imidazole

ALBP, Δ98Δ

a. 15 ml of degassed 0.5 M imidazole in Tris-HCl (50 mM, pH 8.5) were added to a solution of 12.5 nmol of protein modified with **10** in Tris·HCl (50 mM, 0.25 ml, pH 8) in a Schlenk-tube, followed by stirring for 16 hours at room temperature. The protein solution was finally concentrated to 0.25 ml and analysed by LC-MS (ES+).
b. The same procedure was tested for 1M imidazole concentration (in Tris-HCl, 50 mM, pH 8.5).

No mass corresponding to the protein modified with free phosphane was observed in both cases.

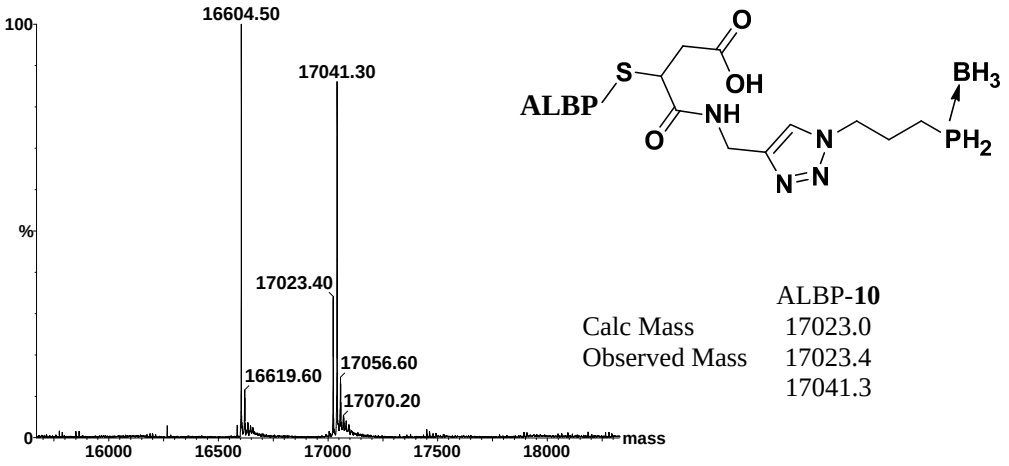


Figure 64 LC-MS (ES+) spectrum of ALBP - **10** after decomplexation treatment with 1 M imidazole.

Tris(hydroxymethyl)phosphane

ALBP

15 ml of degassed 100 mM $P(CH_2OH)_3$ in Tris-HCl (50mM, pH 8) were added to a solution of 12.5 nmol of ALBP modified with **10** in Tris·HCl (50 mM, 0.25 ml, pH 8.0) in a Schlenk-tube, followed by stirring for 5 hours at room temperature. The protein solution was finally concentrated to 0.25 ml.

This same procedure was tested also using a 200 mM $P(CH_2OH)_3$ solution, as described for the 100 mM $P(CH_2OH)_3$ solution.

No protein (modified or unmodified) was present in solution at the end of the decomplexation reaction (Bradford assay[68]).

Hydroxylamine

ALBP

15 ml of degassed 100 mM hydroxylamine in Tris-HCl (50mM, pH 8) were added to a solution of 12.5 nmol of ALBP modified with **10** in Tris·HCl (50 mM, 0.25 ml, pH 8) in a Schlenk-tube, followed by stirring for 16 hours at room temperature. The protein solution was finally concentrated to 0.25 ml and analysed by LC-MS (ES+).

This same procedure was tested also using a 200 mM hydroxylamine solution, as described for the 100 mM hydroxylamine.

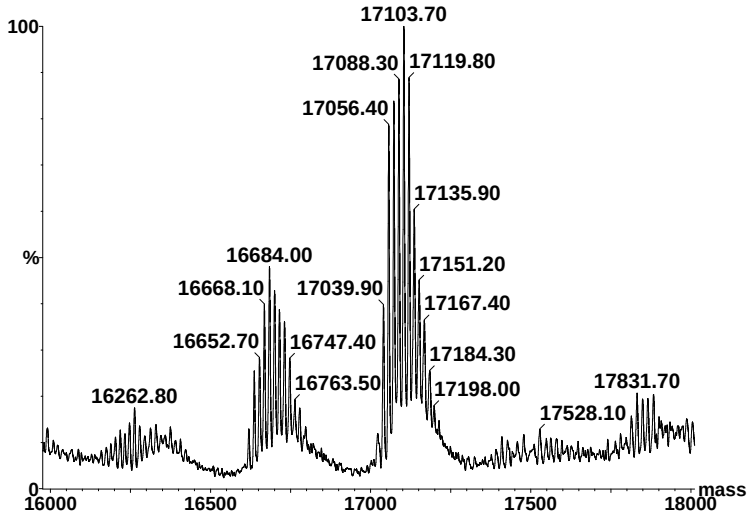


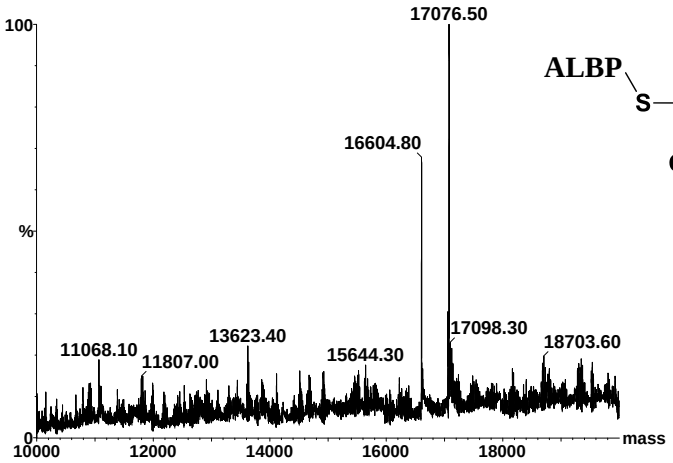
Figure 65 LC-MS (ES+) spectrum of ALBP - **10** after decomplexation attempt with hydroxylamine (200 mM).

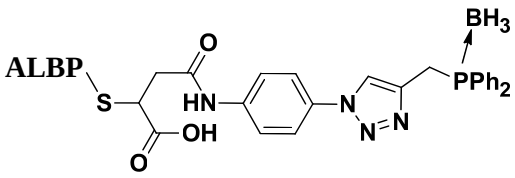
Multiple additions of a mass varying from 14 to 16 to the mass of both unmodified and modified protein were observed after the decomplexation attempt.

Morpholine (0.5 M)[61]

ALBP, Δ98Δ, AppA Y21C

15 ml of degassed 0.5 M morpholine in Tris-HCl (50mM, pH 9) were added to a solution of 12.5 nmol of protein modified with click-phosphane in Tris·HCl (50 mM, 0.25 ml, pH 8) in a Schlenk-tube, followed by stirring for 16 hours at room temperature. The protein solution was finally concentrated to 0.25 ml and analysed by LC-MS ES(+).





	ALBP-**11**
Calc Mass	17057.0
Observed Mass	17076.5

Figure 66 LC-MS (ES+) spectrum of ALBP - **11** after decomplexation attempt with 0.5 M morpholine

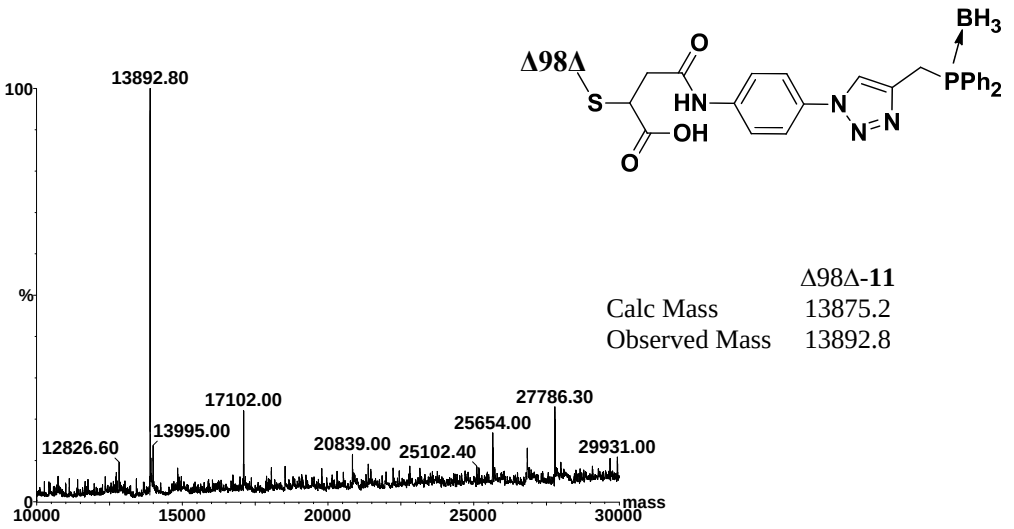


Figure 67 LC-MS (ES+) spectrum of Δ98Δ - **11** after decomplexation treatment with 0.5M morpholine

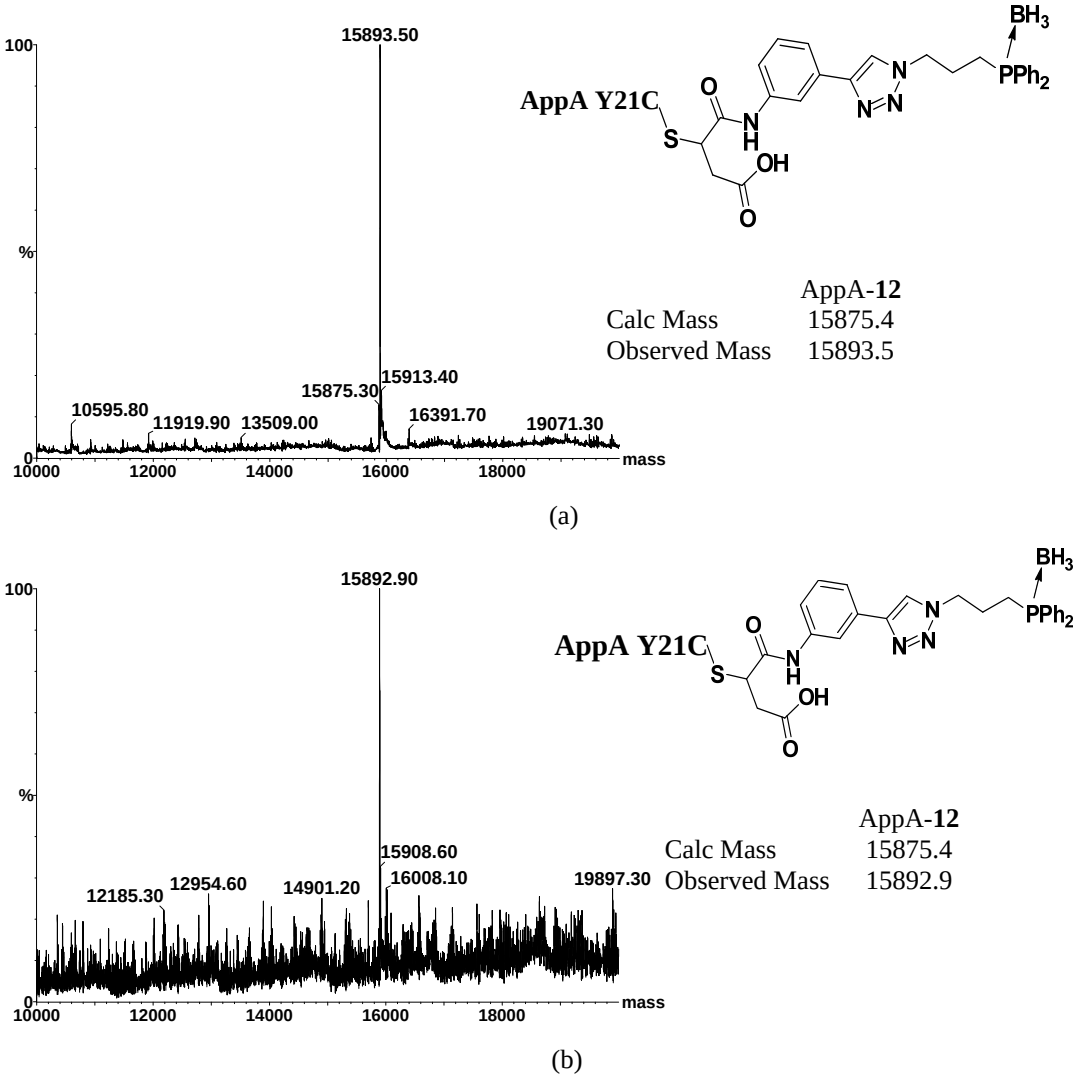


Figure 68 LC-MS (ES+) spectrum of AppA Y21C modified with **12** after decomplexation treatment with 0.5 M morpholine (a) and neat morpholine (b)

No mass of the protein modified with free phosphane was observed after the decomplexation treatment, for all the three proteins.

Morpholine (neat)[61]

AppA Y21C

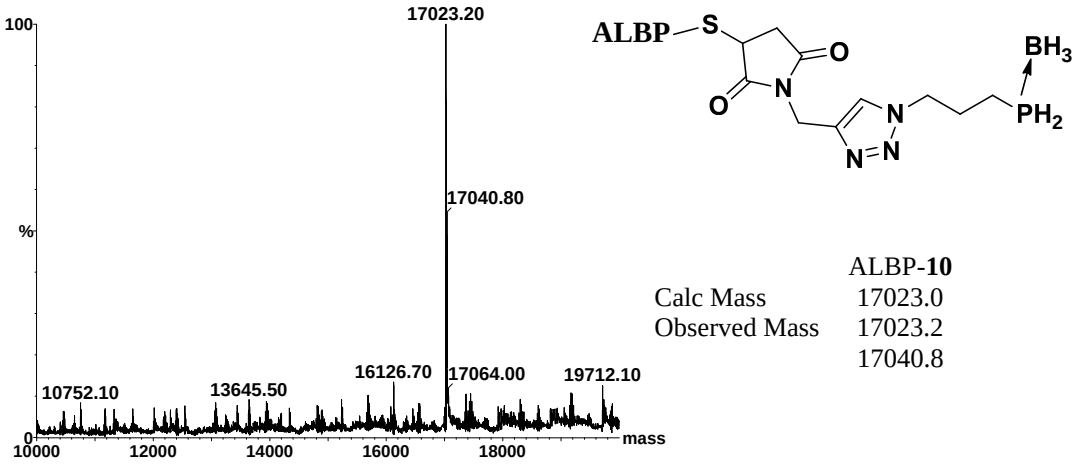
The borane decomplexation with neat morpholine was also tested. 15 ml of degassed morpholine were added to 12.5 nmol of AppA Y21C modified with phosphane **12** in Tris·HCl (50 mM, 0.25 ml), in a Schlenk tube. The reaction was stirred for 3 hours at room temperature. After that the reaction mixture was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), concentrated to 0.5 ml, diluted to 10 ml with Tris-HCl (50 mM, pH 8.0) and concentrated to 0.2 ml. This step was repeated four times after which the modified protein was concentrated to 0.2 ml and analysed by LC-MS (ES+). No mass corresponding to protein modified with free phosphane was observed (Figure 68b).

KCN

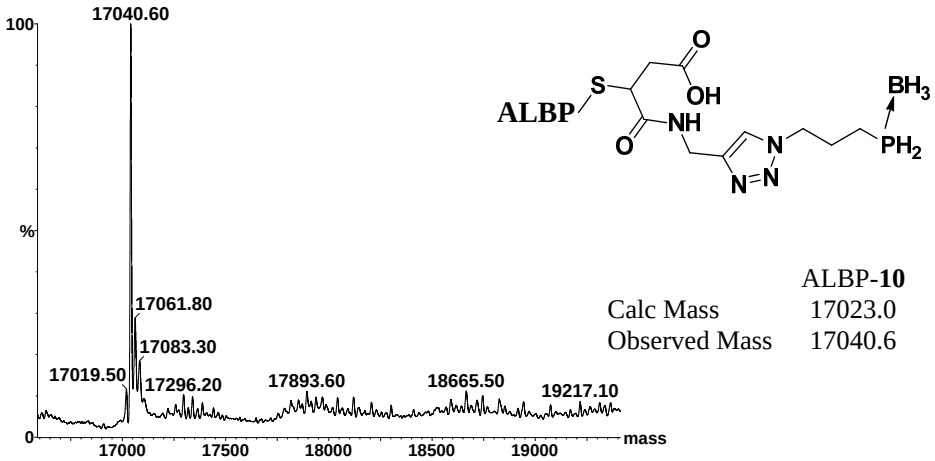
ALBP, Δ98Δ, AppA Y21C

The BH_3 decomplexation attempt using KCN was performed based on a literature procedure.[58]

5 ml of degassed 100 mM KCN in Tris-HCl (50mM, pH 8) were added to a solution of 10 nmol of protein modified with **10** in Tris·HCl (50 mM, 0.2 ml, pH 8) in a Schlenk-tube, followed by stirring for 16 hours at room temperature. After that the reaction mixture was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with degassed Tris-HCl (50 mM, pH 8.0) and concentrated to 0.2 ml. This step was repeated four times after which the modified protein was concentrated to 0.2 ml and analysed by LC-MS ES(+). No mass corresponding to the protein modified with free phosphane was detected by LC-MS after decomplexation treatment.



(a)



(b)

Figure 69 LC-MS (ES+) spectrum of ALBP - **10** before (a) and after (b) decomplexation attempt with KCN

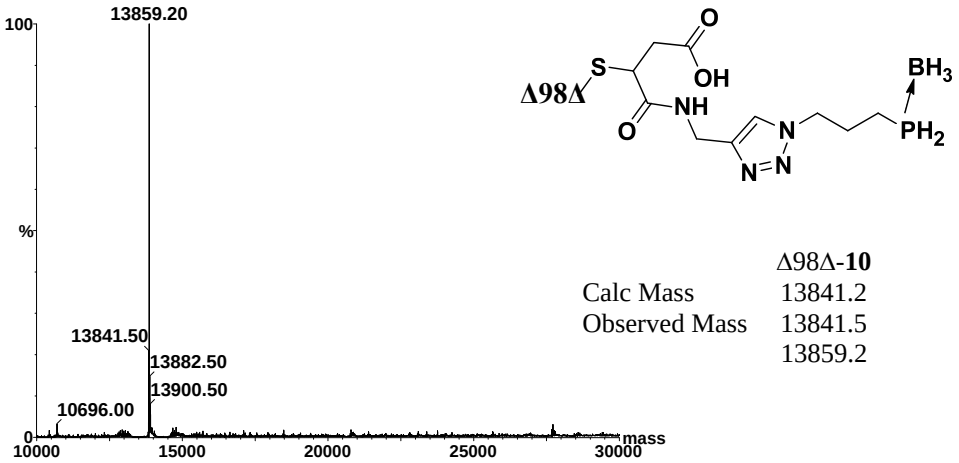


Figure 70 LC-MS (ES+) spectrum of Δ98Δ - **10** after decomplexation with KCN

CO

The attempt to deprotect the phosphane-borane **10** coupled to ALBP, Δ98Δ and AppA Y21C using CO was performed following a modified literature procedure.[62]

A solution of 25 nmol of protein modified with **10** in Tris·HCl (50 mM, 0.5 ml, pH 8.0) was maintained for 16 hours under 30 bars CO pressure at room temperature.

LC-MS (ES+) spectrum of modified ALBP, treated with CO for borane deprotection, shows an addition of a mass of 56 to the mass of modified protein (Figure 71), which could correspond to the mass of two molecules of CO. When one equivalent of unmodified protein was added to the modified protein after CO decomplexation, the same addition of a mass of 56 was observed for the freshly added apo-protein (Figure 72).

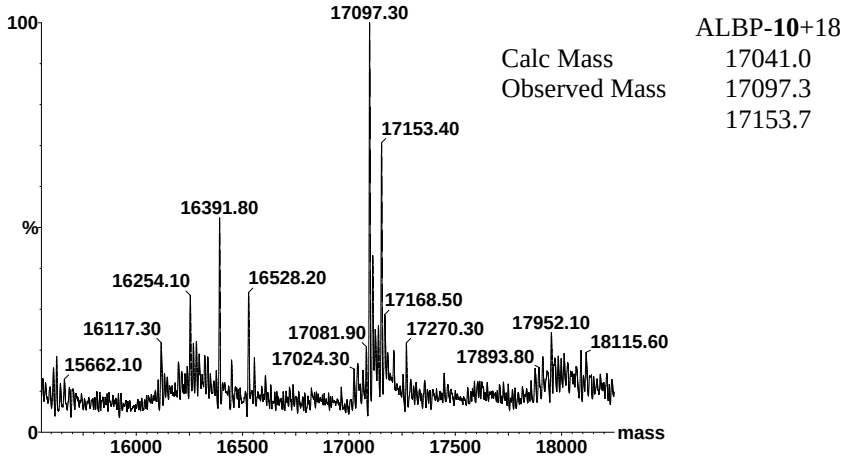


Figure 71 LC-MS (ES+) of ALBP - **10** after decomplexation with CO

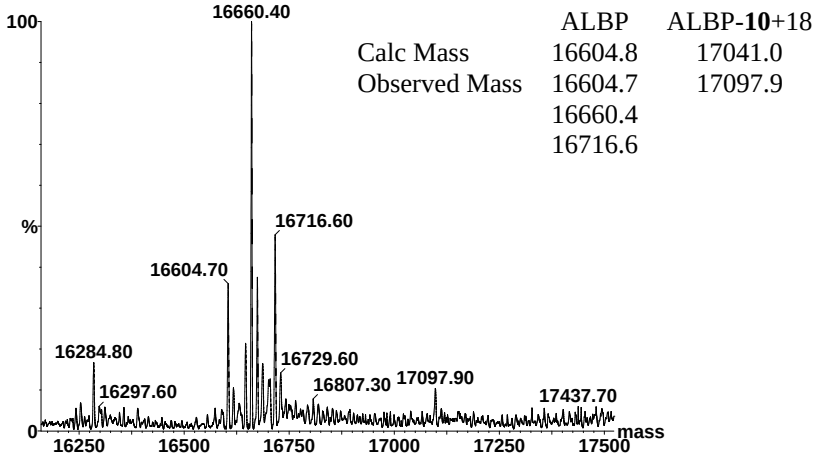


Figure 72 LC-MS (ES+) of ALBP - **10** after decomplexation with CO and addition of one equivalent of unmodified protein

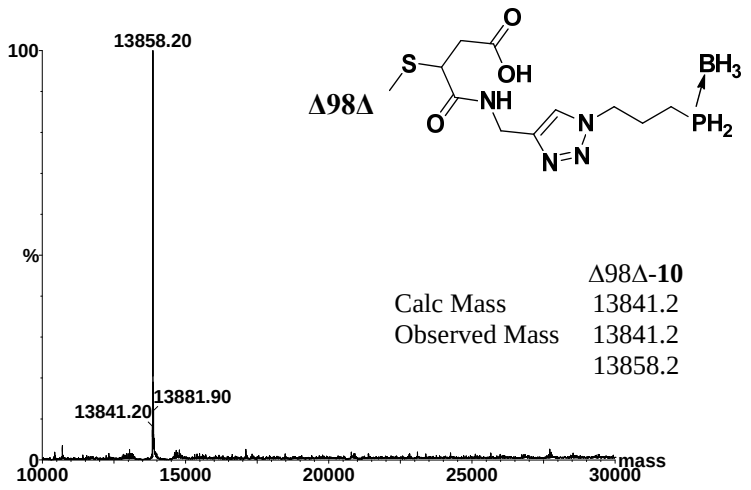


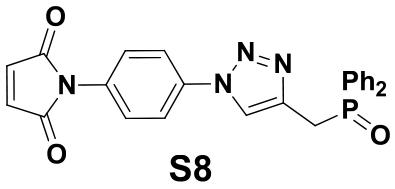
Figure 73 LC-MS (ES+) of Δ98Δ - **10** after decomplexation treatment with CO

However, the addition of 56 was not observed for Δ98Δ and AppA Y21C, were no changes in the mass of the modified protein were noticed after the CO decomplexation attempt.

The question was posed if the protein environment, in particular the NH_2 groups of the lysine residues, have the ability to induce BH3 deprotection and consequent oxidation of phosphane-boranes in aqueous solutions. To verify this hypothesis I chose 3-(diphenylphosphane)benzenesulfonic acid, a phosphane available on large scale. I tested the borane decomplexation of its BH_3 complex in aqueous solutions, in presence and in absence of ALBP. The decomplexation was followed by ^{31}P $\{^1H\}$ NMR. According to the ^{31}P NMR spectra, after 16 hours only phosphane-borane and phosphane oxide were observed and no free phosphane. Also the ratio between phosphane-borane and phosphane oxide was the same for both solutions, with or without protein, and remained the same after 24 hours. Therefore I concluded that for this particular borane-phosphane the protein environment did not promote BH3 deprotection and consequent phosphane oxidation.

The mass difference of only 2 Da between the proteins modified with phosphane oxide and with phosphane borane made difficult to evaluate the borane deprotection of precoupled phosphanes to proteins via LC-MS. To be able to compare between the masses of the two types of proteins, the maleimide functionalized phosphane oxide **S8** was synthesized via click-reaction and coupled to all three proteins.

After a statistically relevant number of coupling reactions of the phosphane-boranes **10-12** and of the phosphane-oxide **S8** to proteins it was possible to distinguish between the masses of the two types of modified protein. In general, for the proteins modified with phosphane-borane the observed mass is 0.5-0.9 Da higher than the calculated mass. In the case of phosphane-oxide, an average mass of the modified protein 0.03 Da higher than the calculated mass (protein+phosphane-oxide) was obtained, thus 2.03 Da higher than the calculated mass for phosphane-borane modified protein.



S8

Figure 74 Maleimide functionalized phosphane-oxide synthesized via azide-alkyne cycloaddition

The mass difference between the modified protein after borane deprotection and the calculated mass of the protein modified with phosphane-borane, observed by MS analysis, was in general no higher than 1 Da, but every time lower than 2 Da. Therefore I concluded that the decomplexation of phosphane-boranes precoupled to proteins, using the procedures outlined above, was unsuccessful for the tested conditions.

g. Coupling of phosphane ligands via hydrazone linkage

Conjugation with 3-maleimidopropionic acid hydrazide

ALBP

2.5 μmol (10 eq.) of **13** in 50 μl of dimethylformamide was added to a solution of 250 nmol of protein (1 eq.) in potassium phosphate buffer (50 mM, 5 ml, pH 7), followed by stirring at room temperature, for 16 hours, under air. To remove the unreacted compound **13** the reaction mixture was dialyzed for 16 hours against potassium phosphate buffer (50mM, pH 7) after which the modified protein was

stored at 4°C.

Δ98Δ

5.5 µmol (22 eq.) of **13** in 0.11 ml of dimethylformamide was added to a solution of 250 nmol of protein (1 eq.) in Tris·HCl (50 mM, 5 ml, pH 8.0) containing DTT (6 eq.) in a Schlenk-tube, followed by stirring for 16 hours at room temperature. To remove the unreacted compound **13** and the side products, the reaction mixture was dialyzed for 16 hours against potassium phosphate buffer (50 mM, pH 7) after which the modified protein was stored at 4°C.

AppA Y21C

2.5 µmol (10 eq.) of 3-maleimidopropionic acid hydrazide in 50 µl of dimethylformamide was added to a solution of 250 nmol of protein (1 eq.) in Tris·HCl (50 mM, 5 ml) containing 3M imidazole at pH 7, followed by stirring at room temperature under air for 16 hours. To remove the imidazole and the unreacted compound **13**, the reaction mixture was dialyzed for 16 hours against potassium phosphate buffer (50 mM, pH 7) after which the modified protein was stored at 4°C.

AppA Q63C

Coupling of **13** to AppA Q63C was performed as described for coupling of **13** to AppA Y21C.

Conjugation with aldehyde functionalized phosphanes 14-17 via intermolecular hydrazone ligation

ALBP, Δ98Δ, AppA Y21C, AppA Q63C

Prior to coupling reactions the solution of the proteins modified with 3-maleimidopropionic acid hydrazide were degassed by purging with argon for 8 hours.

1. Conjugation with 14

500 nmol (10 eq.) of aldhyde functionalized phosphane **14** in 10 µl of dimethylformamide was added to a solution of 50 nmol of protein modified with 3-maleimidopropionic acid hydrazide (1 eq.) in potassium phosphate buffer (50 mM, 1 ml, pH 7) . The reaction mixture was stirred at room temperature for 24 hours. The insoluble fraction of the excess phosphane ligand was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal

concentrator (Amicon Ultra 15, 10,000 MWCO) diluted to 10 ml with degassed potassium phosphate buffer (50 mM, pH 7.0) and concentrated to 0.5 ml. This step was repeated four times after which the modified protein was concentrated to 1 ml and stored at 4°C.

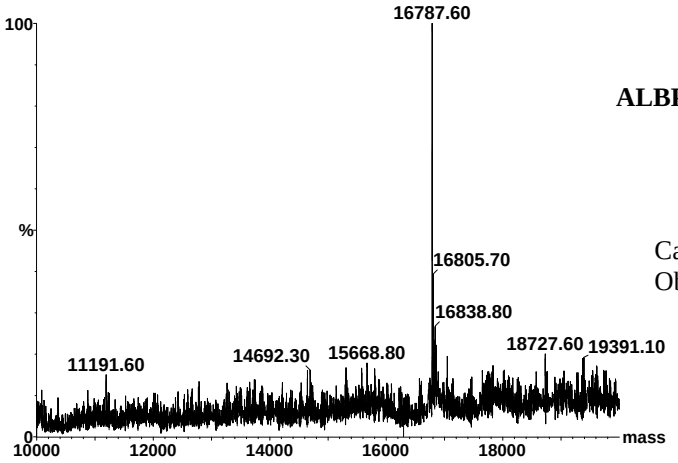
2. Conjugation with 15-17

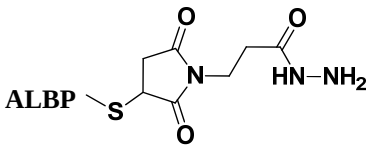
Conjugation with phosphanes **15-17** was performed as described for **14**.

2. Conjugation with 18

Conjugation with phosphane **18** was performed as described for **14**. 1 µmol (20 eq) of phosphane aldehyde **18** in 40 µl dimethylformamide was used for the coupling reaction.

ALBP





	ALBP-**13**
Calc Mass	16787.9
Observed Mass	16787.6

Figure 75 LC-MS (ES+) spectrum of ALBP - **13**

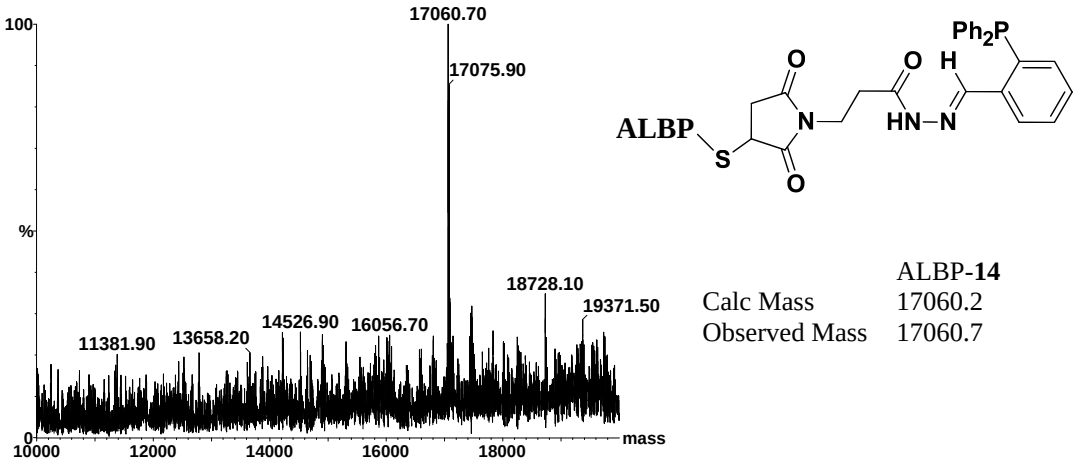


Figure 76 LC-MS (ES+) spectrum of ALBP modified with **14** via hydrazone linkage

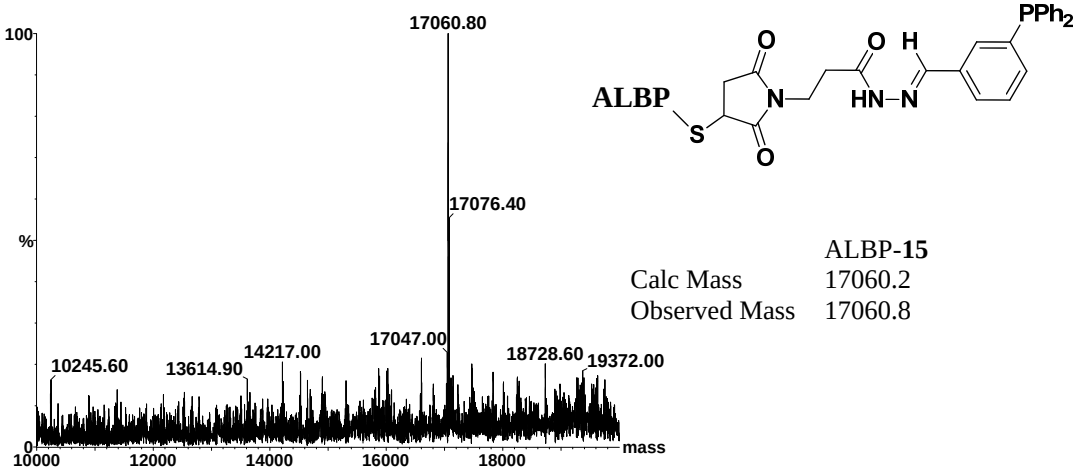


Figure 77 LC-MS (ES+) spectrum of ALBP modified with **15** via hydrazone linkage

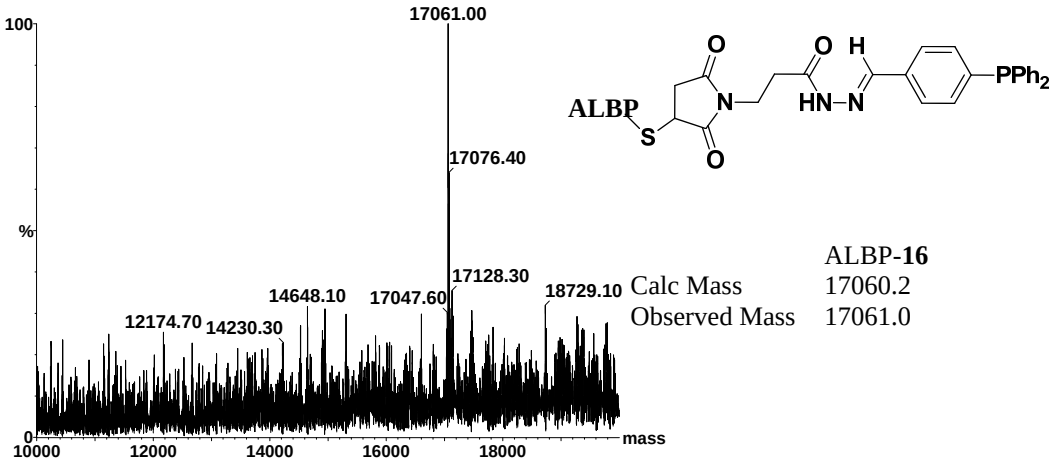


Figure 78 LC-MS (ES+) spectrum of ALBP modified with **16** via hydrazone linkage

nLC-ESI-MS/MS analysis of the peptides resulted from digestion of ALBP (apo-protein and protein modified with (13+15)) with trypsin

The peptides were acidified with TFA and then separated using an UltiMate nanoLC (LC Packings, Amsterdam) equipped with a PepMap C18 trap & column. A 60 minutes gradient of increasing acetonitrile concentration containing 0.1 % formic acid was used (5-35% acetonitrile in first 35 minutes, 35-50% in a further 20 minutes, followed by 95% acetonitrile to clean the column). The eluent was sprayed into a Q-Star XL tandem mass spectrometer (Applied Biosystems, Foster City, CA) and analysed in Information Dependent Acquisition (IDA) mode, performing 1 sec of MS followed by 3 sec MS/MS analyses of the 2 most intense peaks seen by MS. These masses are then excluded from analysis for the next 60 sec. MS/MS data for doubly and triply charged precursor ions was converted to centroid data, without smoothing, using the Analyst QS1.1 mascot.dll data import filter with default settings. The MS/MS data file generated was analysed using the Mascot 2.1 search engine (Matrix Science, London, UK) against an internal database containing the protein sequence of interest. The data was searched with tolerances of 0.2 Da for the precursor and fragment ions, trypsin as the cleavage enzyme, one missed cleavage, and (oxidized) ligand modification of cysteine, and lysine and methionine oxidation, selected as variable modifications.

An ion at m/z 460.2 corresponding to the double protonated peptide Leu^{131}-Lys^{138} was observed in the MS spectrum of digested ALBP (Figure 80a) and ALBP-(**13+15**). An ion at m/z 705.3 corresponding to the double protonated peptide Leu^{131}-Lys^{138} modified with the maleamic acid adduct of oxidised hydrazone phosphane (**13**+**15**) was observed in the MS spectrum of digested ALBP-(**13+15**) (Figure 79b, Figure 80b). MS/MS of peptide Leu^{131}-Lys^{138}-(**13+15**) confirmed cysteine modification (Figure 81).

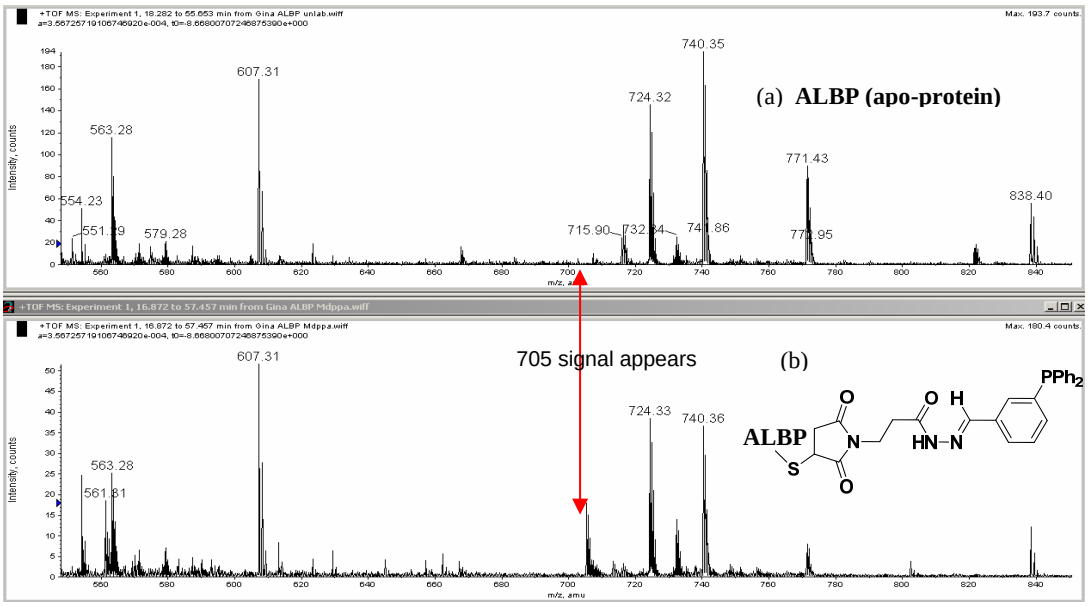
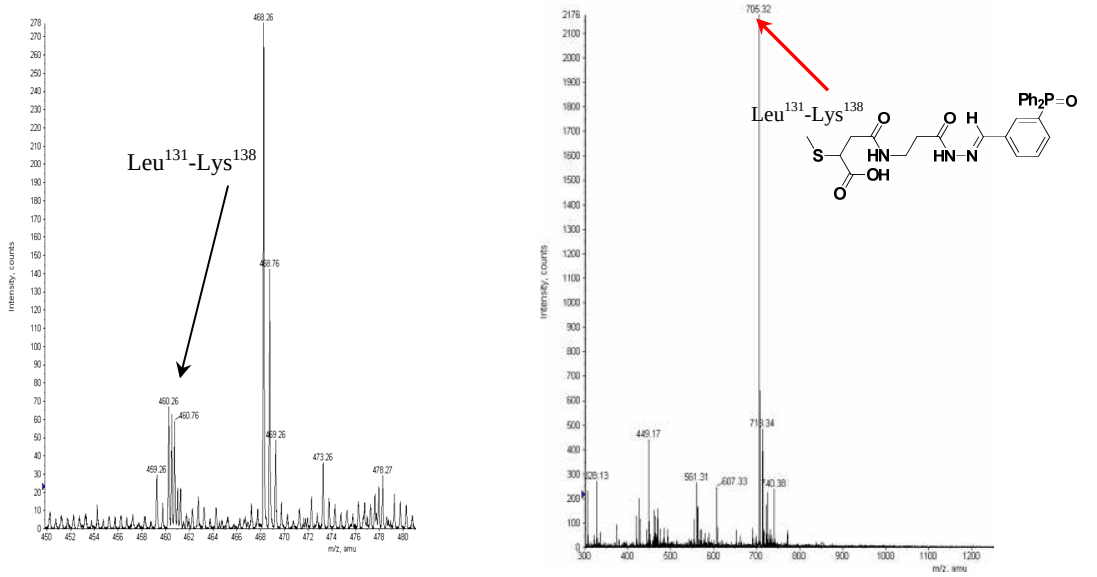
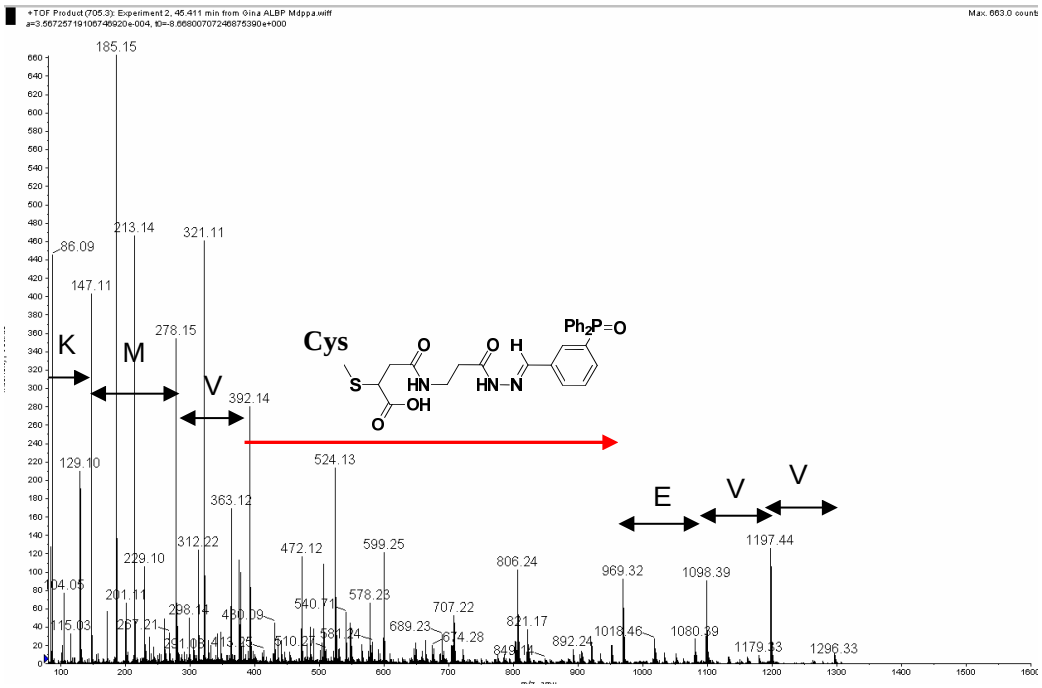


Figure 79 ESI-MS spectrum of tryptic peptides from ALBP (a) and ALBP+(**13**+**15**) (b)

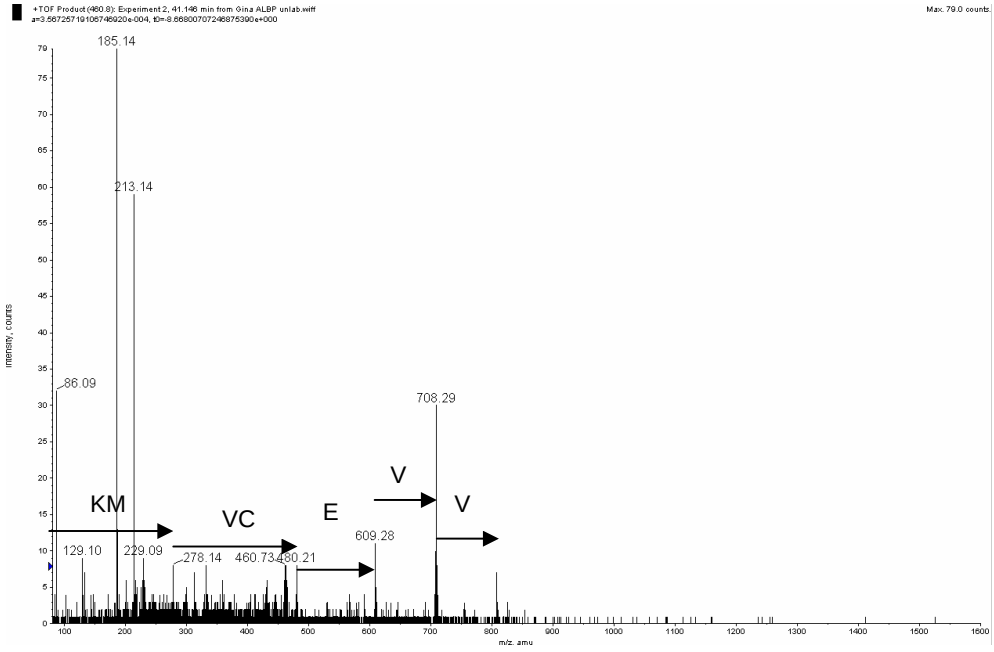


(a) (b)

Figure 80 Zoomed ESI-MS spectrum of tryptic peptides showing the ion at m/z 460.2 corresponding to the peptide Leu^{131}-Lys^{138} (a) and the ion at m/z 705.3 corresponding to Leu^{131}-Lys^{138} modified with maleamic acid adduct of oxidised hydrazone phosphane (**13**+**15**)



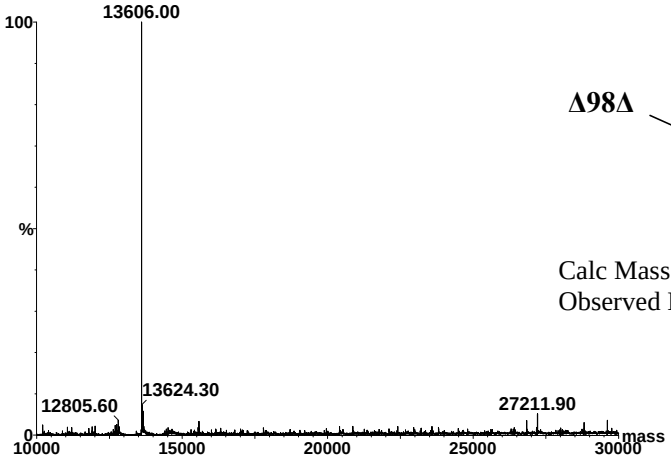
(a)

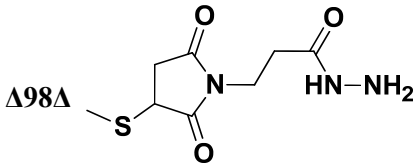


(b)

Figure 81 nLC-ESI-MS/MS of the peptide at m/z 705.3 (a) and of the peptide at m/z 460.2 (b)

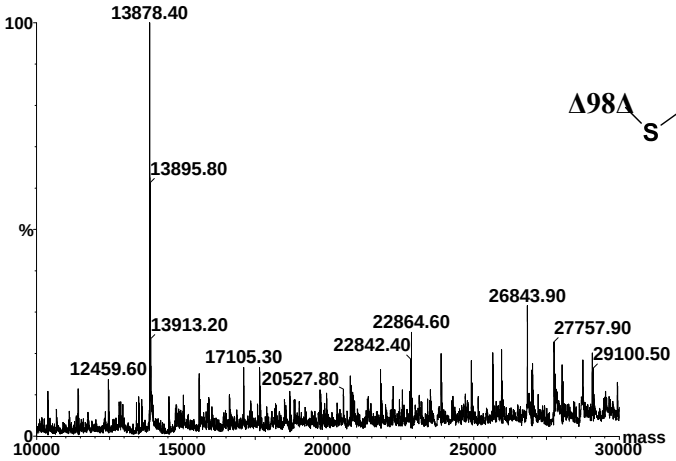
Δ98Δ

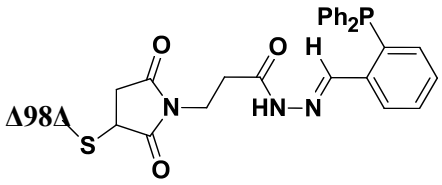




	Δ98Δ-**13**
Calc Mass	13606.1
Observed Mass	13606.0

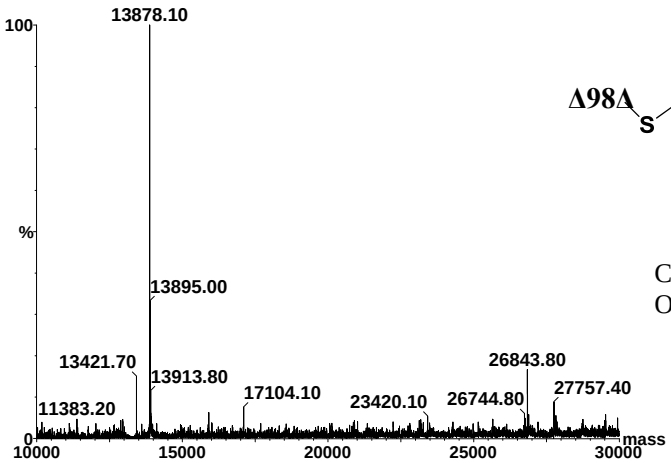
Figure 82 LC-MS (ES+) spectrum of Δ98Δ - **13**

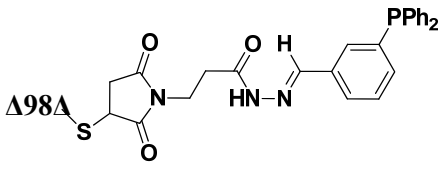




	Δ98Δ-**14**
Calc Mass	13878.4
Observed Mass	13878.4

Figure 83 LC-MS (ES+) spectrum of Δ98Δ modified with **14** via hydrazone linkage





	Δ98Δ-**15**
Calc Mass	13878.4
Observed Mass	13878.1

Figure 84 LC-MS (ES+) spectrum of Δ98Δ modified with **15** via hydrazone linkage

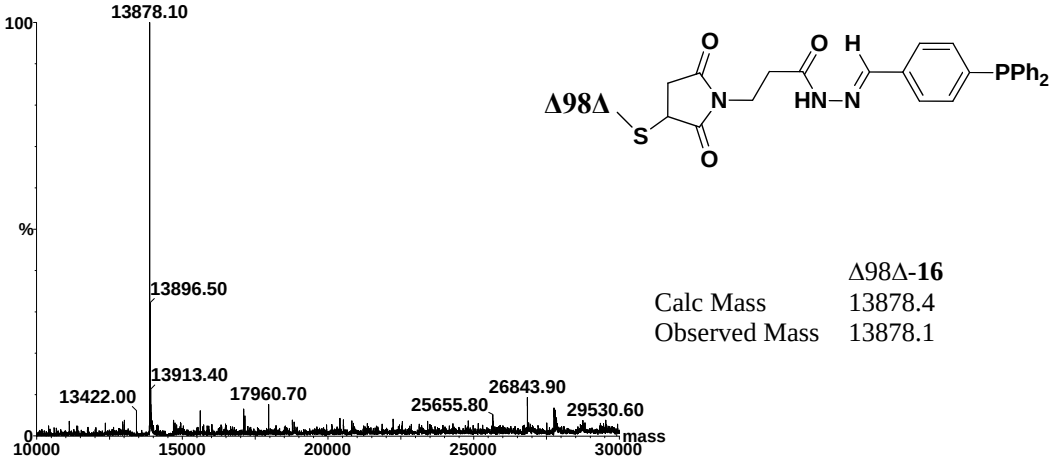


Figure 85 LC-MS (ES+) spectrum of Δ98Δ modified with **16** via hydrazone linkage

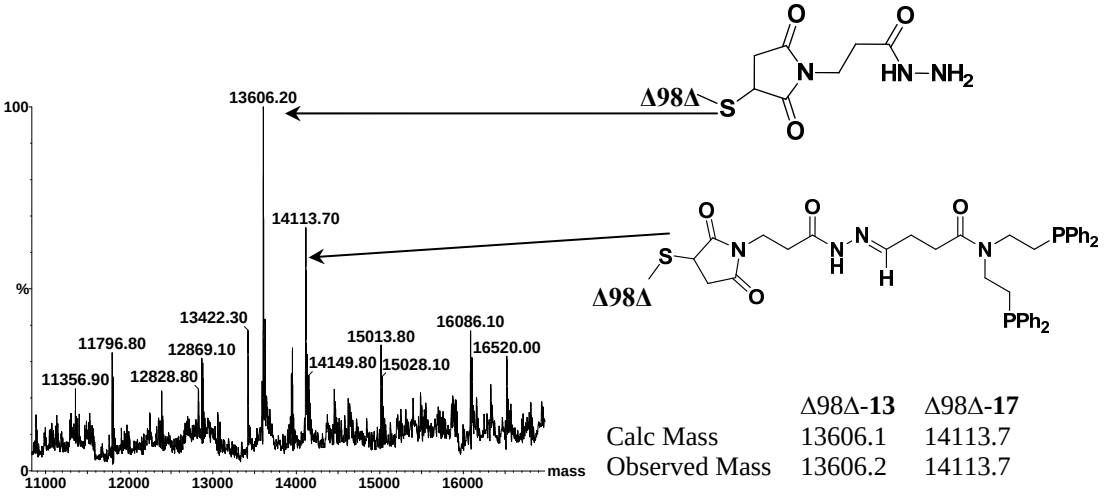


Figure 86 LC-MS (ES+) spectrum of Δ98Δ modified with **17** via hydrazone linkage

MALDI-MS analysis of the peptides resulted from digestion of Δ98Δ (apo-protein and protein modified with (13+15) and (13+16)) with trypsin

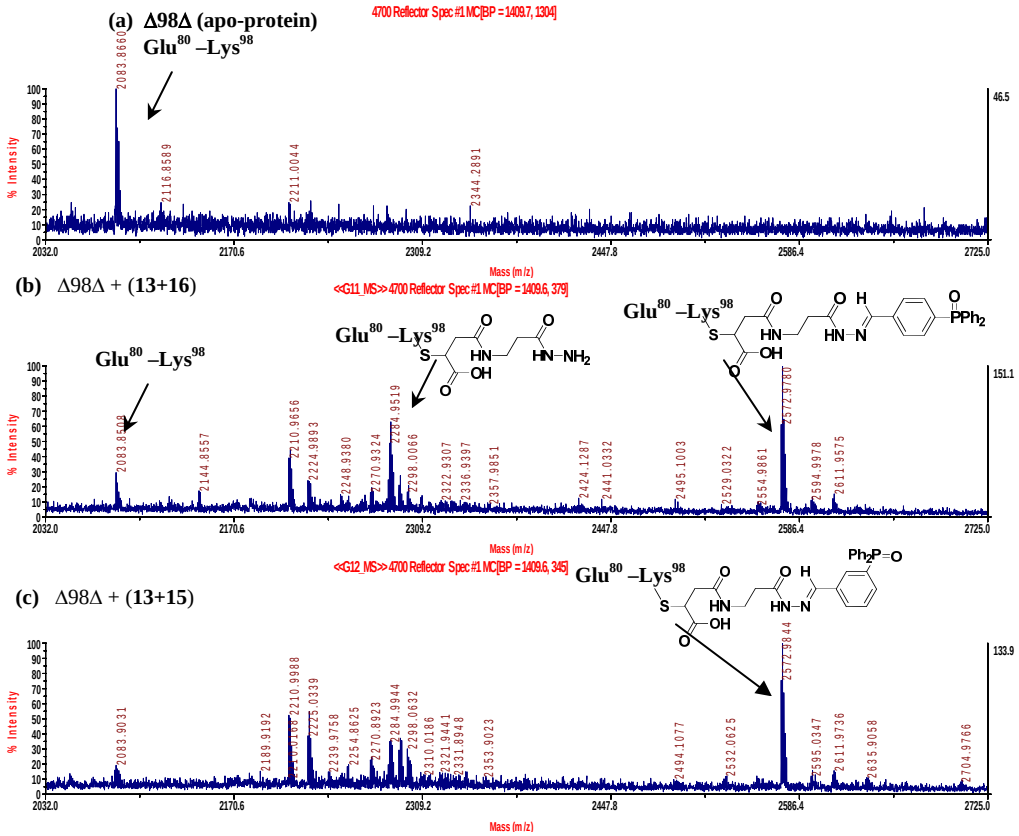
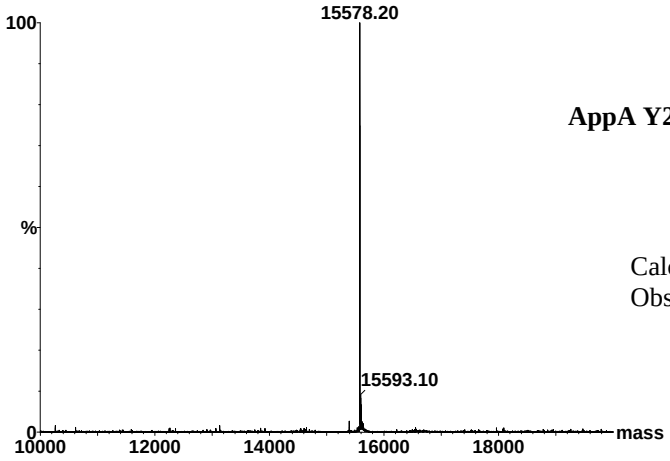


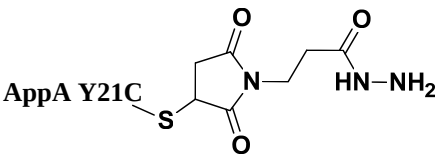
Figure 87 MALDI MS spectrum of tryptic peptides resulted from digestion of Δ98Δ (a), Δ98Δ-(**13**+**16**) (b) and Δ98Δ-(**13**+**15**) (c) with trypsin

MALDI MS of the digest solution was performed as described for ALBP-**5**. The spot was analysed in positive MS mode between 800 and 5000 m/z. The ion at m/z 2083.8 was found in the spectrum of the apo- and modified Δ98Δ, corresponding to the peptide $^{(80)}$EISGNELIQTCTYEGVEAK$^{(98)}$. The ions at m/z 2284.9 and 2572.98 were found in the spectrum of modified Δ98Δ. The ion at m/z 2284.9 corresponds to the peptide Glu^{80} –Lys^{98} modified with the maleamic acid adduct of **13**. The ion at m/z 2572.98 corresponds to the peptide Glu^{80} –Lys^{98} modified with the maleamic acid adduct of the oxidised hydrazone phosphanes (**13**+**15**) or (**13**+**16**).

MALDI MS/MS analysis of peptide at m/z 2572.98 confirmed cysteine modification.

AppA Y21C



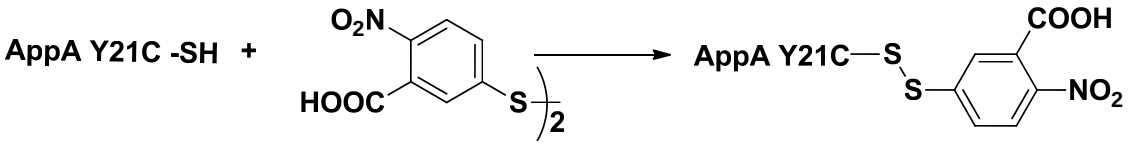


	AppA Y21C-**13**
Calc Mass	15578.2
Observed Mass	15578.2

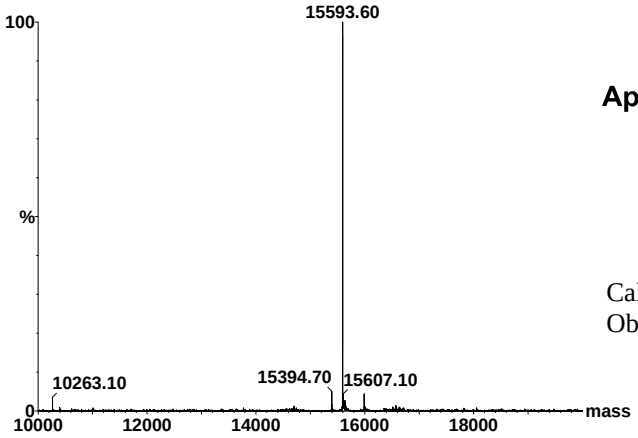
Figure 88 LC-MS (ES+) spectrum of AppA Y21C - **13**

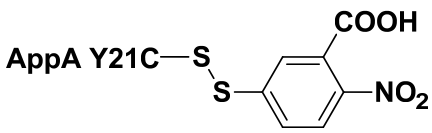
Control of AppA Y21C-13 cysteine modification (control reaction between AppA Y21C-13 and DTNB)[92]

AppA Y21C and Ellman's reagent (DTNB)



500 nmol (10 eq.) of 5,5'-dithiobis-(2-nitrobenzoic acid) in 50 µl of Tris·HCl (50 mM, pH 7) was added to a solution of 50 nmol of AppA Y21C (1 eq.) in buffer (50 mM Tris·HCl, 1 ml), containing 3M imidazole, at pH 7, followed by stirring at 4°C for 5 minutes. Modified protein was analysed directly by LC-MS (ES+). Only the mass of protein modified with 5-thio-(2-nitro benzoic acid) was detected (Figure 89).

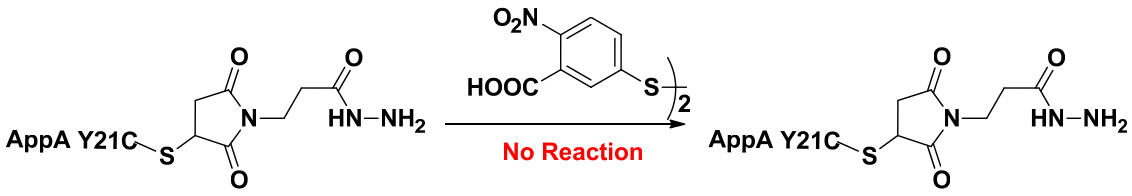




	AppA Y21C-**TNB**
Calc Mass	15592.3
Observed Mass	15593.6

Figure 89 LC-MS (ES+) spectrum of AppA Y21C - **TNB**

AppA Y21C-**13** and Ellman's reagent (**DTNB**)



500 nmol (10 eq.) of 5,5'-dithiobis-(2-nitrobenzoic acid) in 50 µl of Tris·HCl (50 mM, pH 7) was added to a solution of 50 nmol of AppA Y21C-**13** (1 eq.) in buffer (50 mM Tris·HCl, 1 ml), containing 3M imidazole, at pH 7, followed by stirring at 4°C for 5 minutes. Modified protein was analysed directly by LC-MS (ES+). Only one protein species was detected by LC-MS, corresponding to the AppA Y21C modified with **13**. This result indicates that all cysteines of the apo-protein are modified with maleimide **13**.

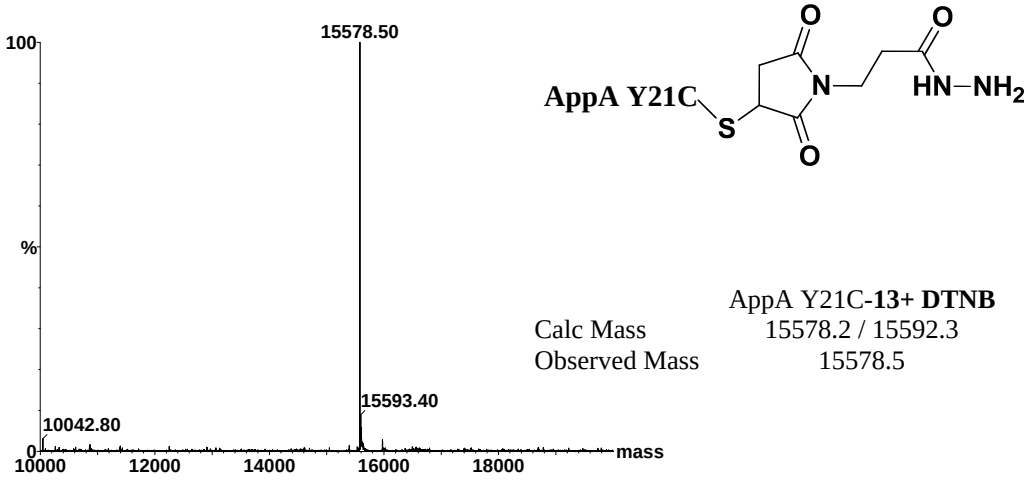


Figure 90 LC-MS (ES+) spectrum of AppA Y21C -**13** after the treatment with **DTNB**

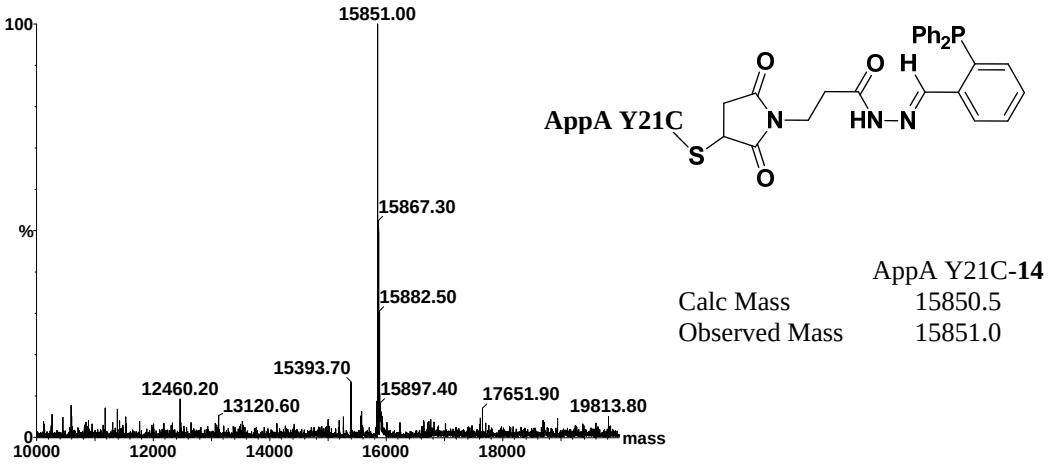
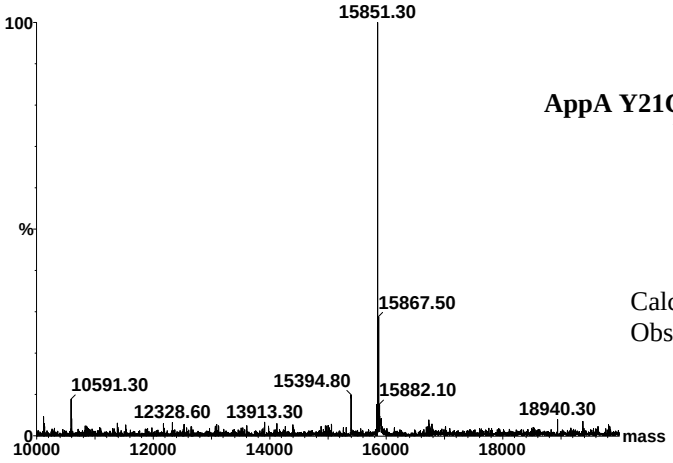
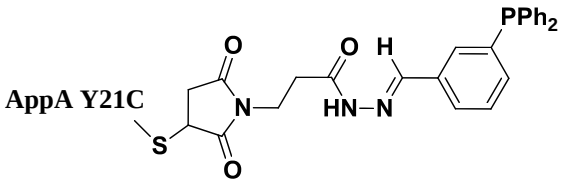


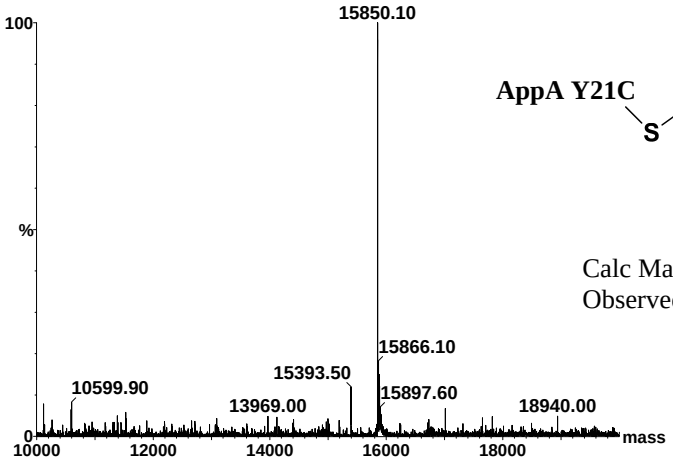
Figure 91 LC-MS (ES+) spectrum of AppA Y21C modified with **14** via hydrazone linkage

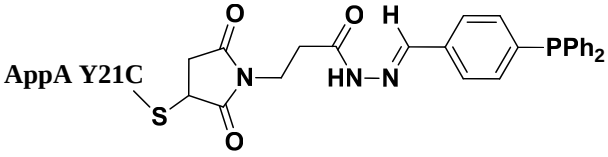




	AppA Y21C-**15**
Calc Mass	15850.5
Observed Mass	15851.3

Figure 92 LC-MS (ES+) spectrum of AppA Y21C modified with **15** via hydrazone linkage

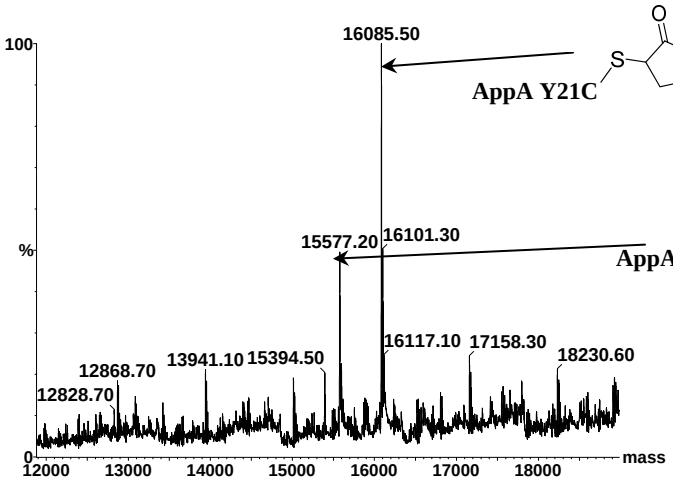


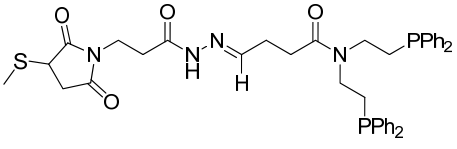


	AppA Y21C-**16**
Calc Mass	15850.5
Observed Mass	15850.1

Figure 93 LC-MS (ES+) spectrum of AppA Y21C modified with **16** via hydrazone linkage

	AppA Y21C-**13**	AppA Y21C-**17**
Calc Mass	15578.2	16085.8
Obs. Mass	15577.2	16085.5





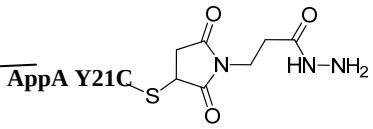
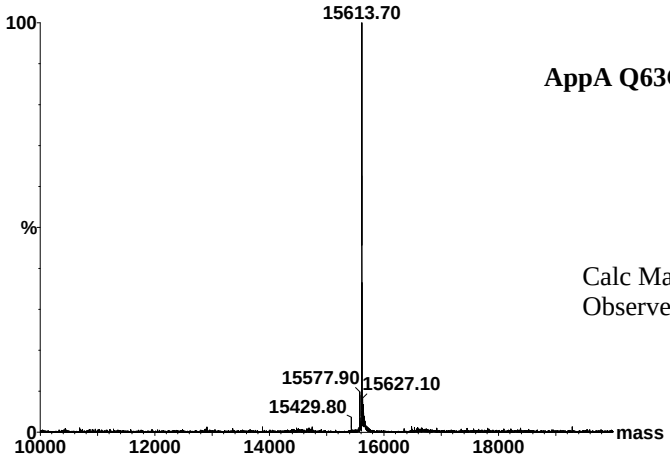
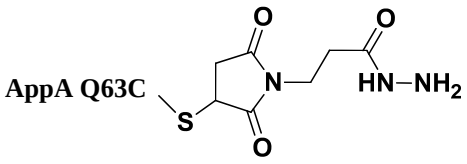


Figure 94 LC-MS (ES+) spectrum of AppA Y21C modified with **17** via hydrazone linkage

AppA Q63C



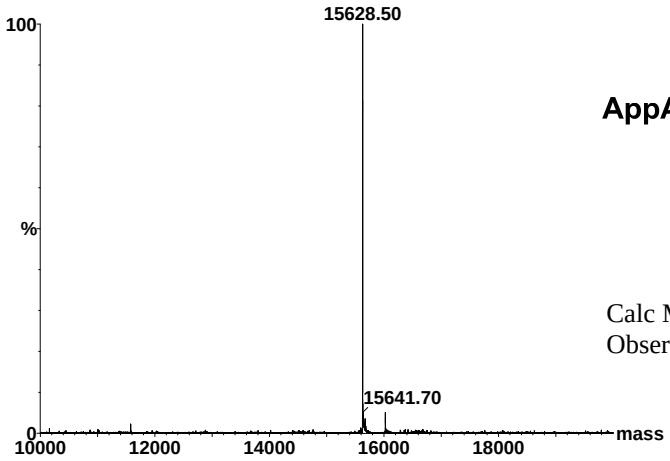


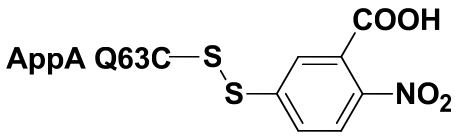
	AppA Q63C-**13**
Calc Mass	15613.3
Observed Mass	15613.7

Figure 95 LC-MS (ES+) spectrum of AppA Q63C - **13**

Control of AppA Q63C -13 cysteine modification (control reaction between AppA Q63C-13 and DTNB)[92]

AppA Q63C and Ellman's reagent (DTNB) – this reaction was performed as described for AppA Y21C and Ellman's reagent. LC-MS (ES+) spectrum indicates the presence of only one species of protein, AppA Q63C modified with 5-thio-(2-nitro benzoic acid) (Figure 96).



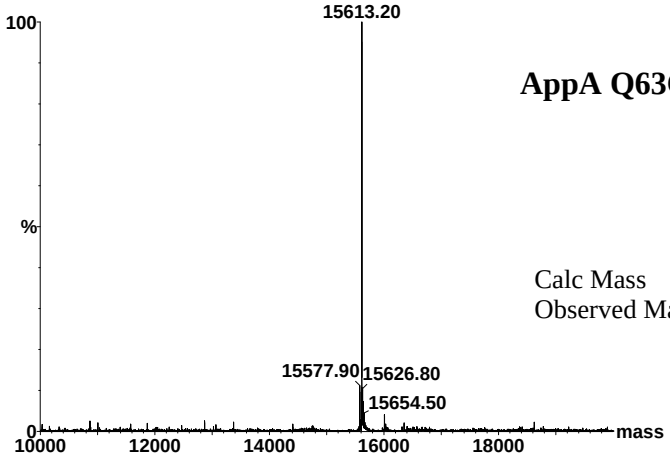


	AppA Q63C-**TNB**
Calc Mass	15627.3
Observed Mass	15628.5

Figure 96 LC-MS (ES+) spectrum of AppA Q63C - **TNB**

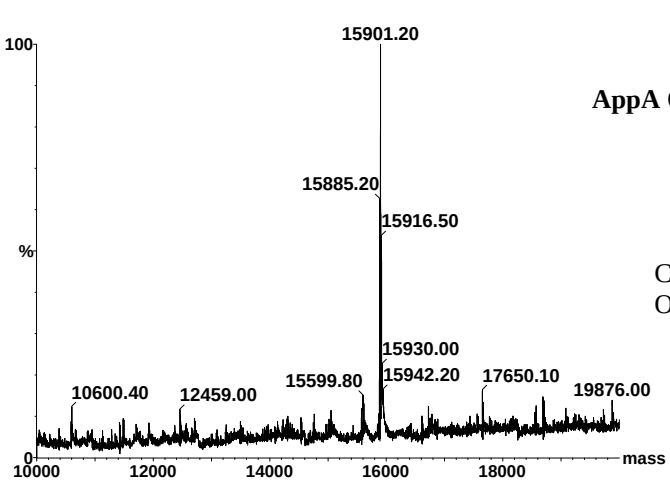
AppA Q63C-**13** and Ellman's reagent (DTNB) – this reaction was performed as described for AppA Y21C–**13** and Ellman's reagent. Only the mass of the protein

modified with hydrazide (AppA Q63C-**13**) was detected by LC-MS (ES+). This indicates complete cysteine modification with **13**.



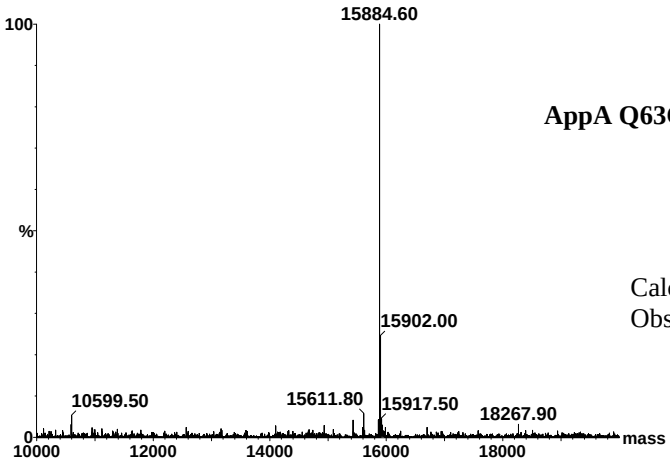
	AppA Q63C-**13**+**DTNB**
Calc Mass	15613.3 / 15627.3
Observed Mass	15613.2

Figure 97 LC-MS (ES+) spectrum of AppA Q63C - **13** after the treatment with **DTNB**



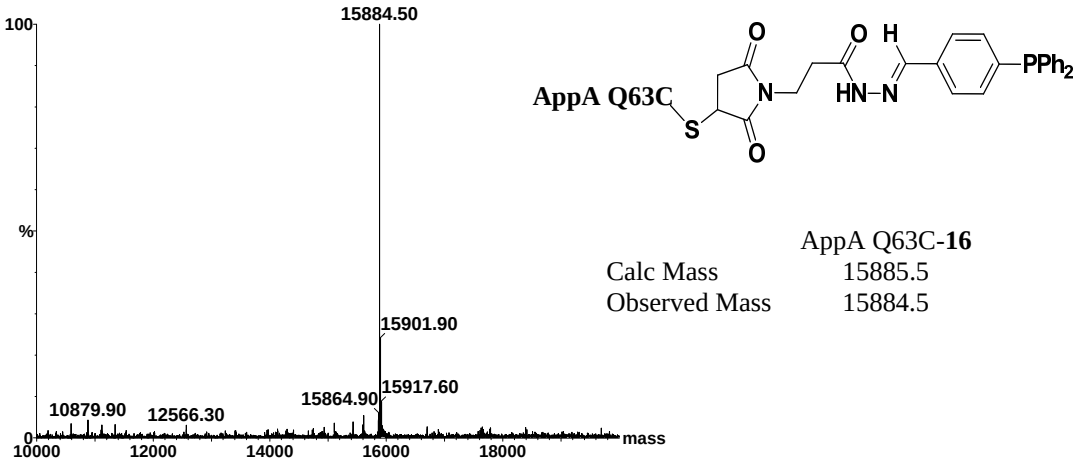
	AppA Q63C-**14**
Calc Mass	15885.5
Observed Mass	15885.2
	15901.2
	15916.5

Figure 98 LC-MS (ES+) spectrum of AppA Q63C modified with **14** via hydrazone linkage



	AppA Q63C-**15**
Calc Mass	15885.5
Observed Mass	15884.6

Figure 99 LC-MS (ES+) spectrum of AppA Q63C modified with **15** via hydrazone linkage



	AppA Q63C-**16**
Calc Mass	15885.5
Observed Mass	15884.5

Figure 100 LC-MS (ES+) spectrum of AppA Q63C modified with **16** via hydrazone linkage

	AppA Q63C-**13**	AppA Q63C-**17**
Calc Mass	15613.3	16120.8
Observed Mass	15613.4	16120.4

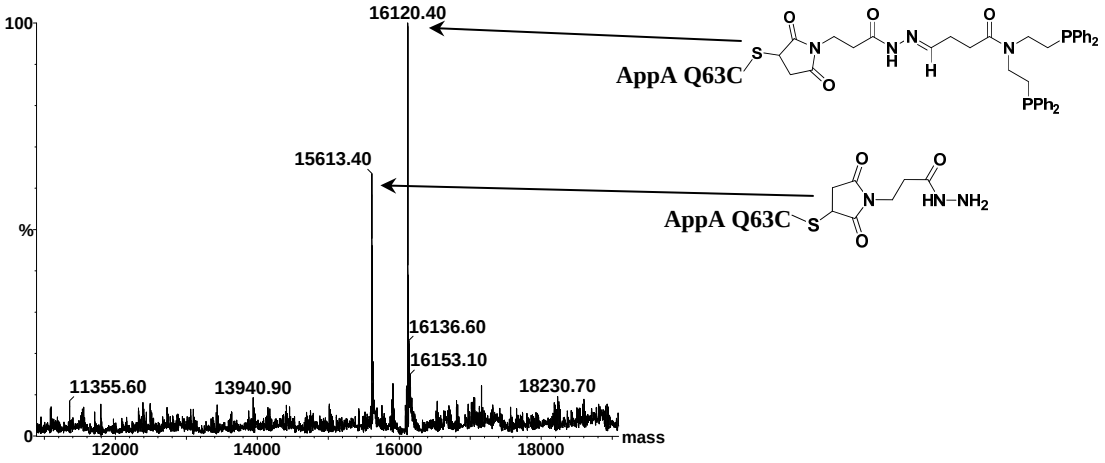


Figure 101 LC-MS (ES+) spectrum of AppA Q63C modified with **17** via hydrazone linkage

	AppA Q63C-**13**	AppA Q63C-**18**
Calc Mass	15613.3	16216.9
Observed Mass	15613.0	16216.2

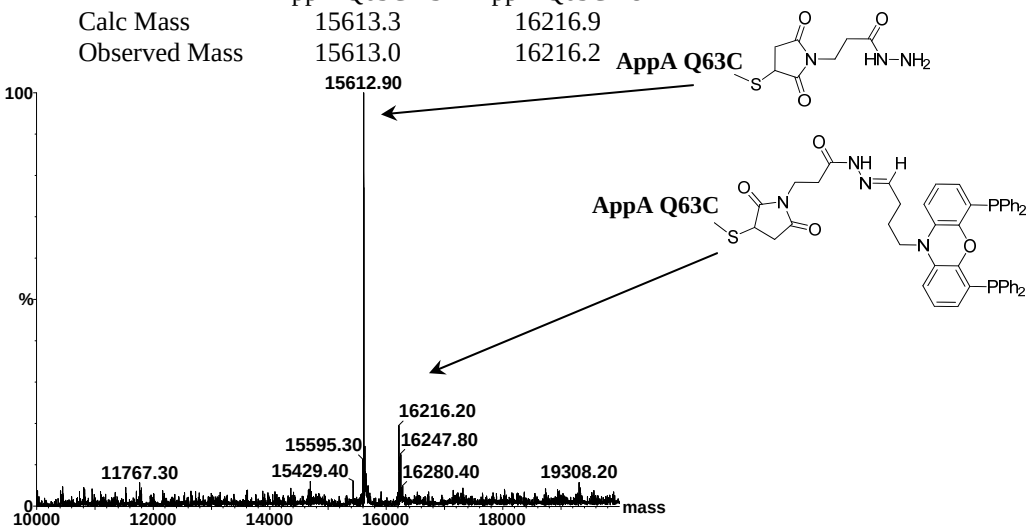


Figure 102 LC-MS (ES+) of spectrum AppA Q63C modified with **18** via hydrazone linkage

REFERENCES

1. Reetz, M. T.; Zonta, A.; Schimossek, K.; Liebeton, K.; Jaeger, K. E. *Angew. Chem. Int. Ed. Engl.* **1997**; *36*; 2830.
2. Reetz, M. T.; Zonta, A.; Schimossek, K.; Liebeton, K.; Jager, K. 2003; Vol. EP 1,002,100 B1.
3. Slams, J. T.; Radziejewski, C.; Oruganti, S. R.; Kaiser, E. T. *J. Am. Chem. Soc.* **1984**; *106*; 6778.
4. Wilson, M. E.; Whitesides, G. M. *J. Am. Chem. Soc.* **1978**; *100*; 306.
5. Collot, J.; Gradinaru, J.; Humbert, N.; Skander, M.; Zocchi, A.; Ward, T. R. *J. Am. Chem. Soc.* **2003**; *125*; 9030.
6. Skander, M.; Humbert, N.; Collot, J.; Gradinaru, J.; Klein, G.; Loosli, A.; Sauser, J.; Zocchi, A.; Gilardoni, F.; Ward, T. R. *J. Am. Chem. Soc.* **2004**; *126*; 14411.
7. Klein, G.; Humbert, N.; Gradinaru, J.; Ivanova, A.; Gilardoni, F.; Rusbandi, U. E.; Ward, T. R. *Angew. Chem. Int. Ed. Engl.* **2005**; *44*; 7764.
8. Qi, D. F.; Tann, C. M.; Haring, D.; Distefano, M. D. *Chem. Rev.* **2001**; *101*; 3081.
9. Rusbandi, U. E.; Lo, C.; Skander, M.; Ivanova, A.; Creus, M.; Humbert, N.; Ward, T. R. *Adv. Synth.&Catal.* **2007**; *349*; 1923.
10. Skander, M.; Malan, C.; Ivanova, A.; Ward, T. R. *Chem. Commun.* **2005**; 4815.
11. Rusbandi, U. E.; Skander, M.; Ivanova, A.; Malan, C.; Ward, T. R. *Comptes Rendus Chimie* **2007**; *10*; 678.
12. Letondor, C.; Pordea, A.; Humbert, N.; Ivanova, A.; Mazurek, S.; Novic, M.; Ward, T. R. *J. Am. Chem. Soc.* **2006**; *128*; 8320.
13. Letondor, C.; Humbert, N.; Ward, T. R. *Proc. Nat. Acad. Sci. USA* **2005**; *102*; 4683.
14. Creus, M.; Pordea, A.; Rossel, T.; Sardo, A.; Letondor, C.; Ivanova, A.; Le Trong, I.; Stenkamp, R. E.; Ward, T. R. *Angew. Chem. Int. Ed. Engl.* **2008**; *47*; 1400.
15. Pordea, A.; Ward, T. R. *Chem. Commun.* **2008**; 4239.
16. Pierron, J.; Malan, C.; Creus, M.; Gradinaru, J.; Hafner, I.; Ivanova, A.; Sardo, A.; Ward, T. R. *Angew. Chem. Int. Ed. Engl.* **2008**; *47*; 701.
17. Reetz, M. T.; Peyralans, J. J. P.; Maichele, A.; Fu, Y.; Maywald, M. *Chem. Commun.* **2006**; 4318.
18. Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Maywald, M.; Maiwald, P.; Peyralans, J. J. P.; Maichele, A.; Fu, Y.; Jiao, N.; Hollmann, F.; Mondiere, R.; Taglieber, A. *Tetrahedron* **2007**; *63*; 6404.
19. Nimri, S.; Keinan, E. *J. Am. Chem. Soc.* **1999**; *121*; 8978.
20. Yamaguchi, H.; Hirano, T.; Kiminami, H.; Taura, D.; Harada, A. *Org. Biomol. Chem.* **2006**; *4*; 3571.
21. Ohashi, M.; Koshiyama, T.; Ueno, T.; Yanase, M.; Fujii, H.; Watanabe, Y. *Angew. Chem. Int. Ed. Engl.* **2003**; *42*; 1005.
22. Davies, R. R.; Distefano, M. D. *J. Am. Chem. Soc.* **1997**; *119*; 11643.
23. Panella, L.; Broos, J.; Jin, J. F.; Fraaije, M. W.; Janssen, D. B.; Jeronimus-Stratingh, M.; Feringa, B. L.; Minnaard, A. J.; de Vries, J. G. *Chem. Commun.* **2005**; 5656.
24. Kruithof, C. A.; Casado, M. A.; Guillena, G.; Egmond, M. R.; van der Kerk-van Hoof, A.; Heck, A. J. R.; Gebbink, R. J. M. K.; van Koten, G. *Chem. Eur. J.* **2005**; *11*; 6869.
25. Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Maywald, M. *Chimia* **2002**; *56*; 721.
26. Maywald, M.; Ruhr-Universität Bochum; 2005.

27. Hermanson, G. T. *Bioconjugate techniques*; Academic Press: San Diego, 1996.
28. Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Taglieber, A.; Hollmann, F.; Mondiere, R. J. G.; Dickmann, N.; Hocker, B.; Cerrone, S.; Haeger, M. C.; Sterner, R. *ChemBioChem* **2008**; *9*; 552.
29. Grinstead, J. S.; Hsu, S. T.; Laan, W.; Bonvin, A. M.; Hellingwerf, K. J.; Boelens, R.; Kaptein, R. *Chembiochem* **2006**; *7*; 187.
30. Kuang, H.; Brown, M. L.; Davies, R. R.; Young, E. C.; Distefano, M. D. *J. Am. Chem. Soc.* **1996**; *118*; 10702.
31. Davies, R. R.; Kuang, H.; Qi, D. F.; Mazhary, A.; Mayaan, E.; Distefano, M. D. *Bioorg. Med. Chem. Lett.* **1999**; *9*; 79.
32. Haring, D.; Lees, M. R.; Banaszak, L. J.; Distefano, M. D. *Prot. Eng.* **2002**; *15*; 603.
33. Curto, L. M.; Caramelo, J. J.; Delfino, J. M. *Biochemistry* **2005**; *44*; 13847.
34. Laan, W.; Bednarz, T.; Heberle, J.; Hellingwerf, K. J. *Photochemical & Photobiological Sciences* **2004**; *3*; 1011.
35. Gomelsky, M.; Kaplan, S. *J. Biol. Chem.* **1998**; *273*; 35319.
36. Anderson, S.; Dragnea, V.; Masuda, S.; Ybe, J.; Moffat, K.; Bauer, C. *Biochemistry* **2005**; *44*; 7998.
37. Gomelsky, M.; Klug, G. *Trends Biochem. Sci.* **2002**; *27*; 497.
38. Laan, W.; Munoz, B. K.; Den Heeten, R.; Kamer, P. C. *ChemBioChem* **2010**; *in press*.
39. Cleland, W. W. *Biochemistry* **1964**; *3*; 480.
40. Masuda, T.; Nobuyuki Ide, N.; Kitabatake, N. *Chem. Senses* **2005**; *30*; 253.
41. Ishikawa, Y.; Yamamoto, Y.; Otsubo, M.; Theg, S. M.; Tamura, N. *Biochemistry* **2002**; *41*; 1972.
42. Miyagi, M.; Nakao, M.; Nakazawa, T.; Kato, I.; Tsunasawa, S. *Rapid Commun. Mass Spectrom.* **1998**; *12*; 603.
43. Yang, J.; Gitlin, I.; Krishnamurthy, V. M.; Vazquez, J. A.; Costello, C. E.; Whitesides, G. M. *J. Am. Chem. Soc.* **2003**; *125*; 12392.
44. de la Escalera, S.; Palacián, E. *Biochem. Cell Biol.* **1989**; *67*; 63.
45. Laan, W.; Gauden, M.; Yeremenko, S.; van Grondelle, R.; Kennis, J. T. M.; Hellingwerf, K. J. *Biochemistry* **2006**; *45*; 51.
46. Tsao, T. C.; Bailey, K. *Biochim. Biophys. Acta* **1953**; *11*; 102.
47. Smyth, D. G.; Blumenfeld, O. O.; Konigsberg, W. *Biochem. J.* **1964**; *91*; 589.
48. Schonberg, A.; Ismail, A. F. A. *J. Chem. Soc., Perkin Trans.* **1940**; 1374.
49. Hedaya, E.; Theodoropulos, S. *Tetrahedron* **1968**; *24*; 2241.
50. Cooley, S. L.; Wood, J. L. *Arch. Biochem. Biophys.* **1951**; *34*; 372.
51. Greenfield, S. J.; Agarkov, A.; Gilbertson, S. R. *Org. Lett.* **2003**; *5*; 3069.
52. Detz, R. J.; Heras, S. A.; de Gelder, R.; van Leeuwen, P. W. N. M.; Hiemstra, H.; Reek, J. N. H.; van Maarseveen, J. H. *Org. Lett.* **2006**; *8*; 3227.
53. Hermanson, G. T. *Bioconjugate techniques*; Academic Press: San Diego, 1996, pp 184.
54. Link, A. J.; Tirrell, D. A. *J. Am. Chem. Soc.* **2003**; *125*; 11164.
55. Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. *J. Am. Chem. Soc.* **2003**; *125*; 3192.

56. Gupta, S. S.; Kuzelka, J.; Singh, P.; Lewis, W. G.; Manchester, M.; Finn, M. G. *Bioconjug. Chem.* **2005**; *16*; 1572.
57. Speers, A. E.; Adam, G. C.; Cravatt, B. F. *J. Am. Chem. Soc.* **2003**; *125*; 4686.
58. Bradley, D.; Williams, G.; Lombard, H.; van Niekerk, M.; Coetzee, P. P.; Holzapfel, C. W. *Phosphorus, Sulfur, and Silicon and the Related Elements* **2002**; *177*; 2799
59. Brisset, H.; Gourdel, Y.; Pellon, P.; Le Corm, M. *Tetrahedron Lett.* **1993**; *34*; 4523.
60. Wright, S. W.; Hageman, D. L.; McClure, L. D. *J. Org. Chem.* **1994**; *59*; 6095.
61. Imamoto, T.; Oshiki, T.; Onozawa, T.; Kusumoto, T.; Sato, K. *J. Am. Chem. Soc.* **1990**; *112*; 5244.
62. de Bod (Lombard), H.; Rand Afrikaans University; 2003.
63. Darcel, C.; Kaloun, E. B.; Merdes, R.; Moulin, D.; Riegel, N.; Thorimbert, S.; Genet, J. P.; Juge, S. *J. Organomet. Chem.* **2001**; *624*; 333.
64. Riegel, N.; Darcel, C.; Stephan, O.; Juge, S. *J. Organomet. Chem.* **1998**; *567*; 219.
65. Heitzmann, H.; Richards, F. M. *Proc. Nat. Acad. Sci. USA* **1974**; *71*; 3537.
66. Itaya, K.; Gahmberg, C. G.; Hakomori, S. I. *Biochem. Biophys. Res. Commun.* **1975**; *64*; 1028.
67. Orr, G. A.; Rando, R. R. *Nature* **1978**; *272*; 722.
68. Bradford, M. M. *Anal. Biochem.* **1976**; *72*; 248.
69. Riener, C. K.; Kada, G.; Gruber, H. J. *Analytical and Bioanalytical Chemistry* **2002**; *373*; 266.
70. Matarese, V.; Bernlohr, D. A. *J. Biol. Chem.* **1988**; *263*; 14544.
71. Laan, W.; van der Horst, M. A.; van Stokkum, I. H.; Hellingwerf, K. J. *Photochemistry and Photobiology* **2003**; *78*; 290.
72. Naghipour, A.; Sabounchei, S. J.; Morales-Morales, D.; Hernandez-Ortega, S.; Jensen, C. M. *J. Organomet. Chem.* **2004**; *689*; 2494.
73. Larpent, C.; Patin, H. *Tetrahedron* **1988**; *44*; 6107.
74. Herd, O.; HeBler, A.; Hingst, M.; Tepper, M.; O., S. *J. Organomet. Chem.* **1996**; *522*; 69.
75. den Heeten, R. In *PhD Dissertation: Amino acid modified phosphine ligands for the development of artificial transition metalloenzymes*; University of Amsterdam: Amsterdam, 2009, p 105.
76. den Heeten, R. In *PhD Dissertation: Amino acid modified phosphine ligands for the development of artificial transition metalloenzymes*; University of Amsterdam: Amsterdam, 2009, p 133.
77. Karlen, B.; Lindeke, B.; Lindgren, S.; Svensson, K. G.; Dahlbom, R.; Jenden, D. J.; Giering, J. E. *J. Med. Chem.* **1970**; *13*; 651.
78. Trommer, W.; Hendrick, M. *Synthesis* **1973**; 484.
79. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem. Int. Ed. Engl.* **2002**; *41*; 2596.
80. Dolhem, F.; Johansson, M. J.; Antonsson, T.; Kann, N. *J. Comb. Chem.* **2007**; *9*; 477.
81. Hon, Y. S.; Lee, C. F.; Chen, R. J.; Szu, P. H. *Tetrahedron* **2001**; *57*; 5991.
82. Rivillo, D.; Gulyas, H.; Benet-Buchholz, J.; Escudero-Adan, E. C.; Freixa, Z.; van Leeuwen, P. W. N. M. *Angew. Chem. Int. Ed. Engl.* **2007**; *46*; 7247.
83. Deuss, P. J.; Popa, G.; Botting, C. H.; Laan, W.; Kamer, P. C. J. *Angew. Chem. Int. Ed. Engl.* **2010**; *in press*.

84. Sambrook, J.; Russell, D. W. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, 2001.
85. Xu, Z. H.; Buelt, M. K.; Banaszak, L. J.; Bernlohr, D. A. *J. Biol. Chem.* **1991**; *266*; 14367.
86. Jenkins-Kruchten, A. E.; Bennaars-Eiden, A.; Ross, J. R.; Shen, W. J.; Kraemer, F. B.; Bernlohr, D. A. *J. Biol. Chem.* **2003**; *278*; 47636.
87. Gilbertson, S. R.; 5541289, U. P. **1996**.
88. Levison, M. E.; Josephson, A. S.; Kirschenbaum, D. M. *Experientia* **1969**; *25*; 126.
89. Hatzakis, N. S.; Engelkamp, H.; Velonia, K.; Hofkens, J.; Christianen, P. C. M.; Svendsen, A.; Patkar, S. A.; Vind, J.; Maan, J. C.; Rowan, A. E.; Nolte, R. J. M. *Chem. Commun.* **2006**; 2012.
90. van Kasteren, S. I.; Kramer, H. B.; Jensen, H. H.; Campbell, S. J.; Kirkpatrick, J.; Oldham, N. J.; Anthony, D. C.; Davis, B. G. *Nature* **2007**; *446*; 1105.
91. Castanet, A. S.; Colobert, F.; Broutin, P. E.; Obringer, M. *Tetrahedron: Asymmetry* **2002**; *13*; 659.
92. Chalker, J. M.; Wood, C. S.; Davis, B. G. *J. Am. Chem. Soc.* **2009**; *131*; 16346.

III. Site-selective bioconjugation of nitrogen-containing ligands to structurally diverse protein hosts

Abstract

A strategy for site-specific covalent incorporation of nitrogen-containing ligands into several protein hosts is described.

Covalent incorporation of 2,2'-bipyridine-5-carboxylic acid into three protein hosts was accomplished after N-hydroxysuccinimide activation of the carboxylic acid group. However the lack of specificity towards cysteine resulted in unspecific protein modifications and incomplete conversion. Site-specific covalent modification of cysteine was achieved by using maleimide functionalized nitrogen-ligands. Complete site-selective chemical modification of three structurally different protein hosts with two maleimide functionalized nitrogen-containing ligands (5-maleimido-1,10-phenanthroline and 3-(N-maleimido)-(N,N-bis(2-picolyl))-propanamide) was obtained via Michael addition to the maleimide double bond.

III.1 Introduction

The wide application scope of nitrogen-containing ligands (N-ligands) in asymmetric catalysis stems from several distinctive advantages of this class of ligands. They are largely available in enantiomerically pure form, either from the chiral pool or as cheap industrial intermediates.[1] Even if chirality on nitrogen atom is difficult to obtain due to instantaneous epimerization of a chiral nitrogen atom at room temperature, formation of a stable chiral centre on a nitrogen atom could be obtained by using bicyclic structures.[2] Unlike phosphines, N-ligands are very stable in aqueous solutions and do not react with atmospheric oxygen. Consequently, their preparation and long term storage are simplified. The large diversity of this class of ligands, sometimes used as part of chelating systems which also involve other donor atoms, such as C, O, P or S, allows their adaptation to different types of reactions and metals.[2] In particular the interactions with transition metals could be varied by involving X-type ligands (amides, sulfonamides) or L-type ligands (amines, imines).[2-3] Amines are considered as "hard" in classical hard/soft acid/base theory[4] and usually bind less polarizable metal centres in higher oxidation states. Imines and pyridines are both good σ donors and π acceptors, therefore are of "borderline" hardness. Consequently they have a great flexibility in binding a wide variety of metals, in a range of oxidation states. However, phosphines and carbenes could be better suited ligands for some catalytic processes involving low oxidation states of the transitional metal centre.[5] The effectiveness of N-ligands for asymmetric catalytic reactions with less expensive transition metals makes them applicable for distinctive reactions, complementary to those catalysed by phosphine-based catalysts. In addition, they could afford alternative catalysts for reactions where the efficiency of phosphines has already been established.[6]

Among the N,N chelating ligands, the derivatives of 2,2'-bipyridine and the structurally similar 1,10-phenanthroline have received special attention, mainly because of their particular coordination chemistry.[7-8] Over the last decades, 2,2'-bipyridine complexes of virtually every transition metal in the periodic table have been described.[7-8] Moreover, transition-metal complexes of chiral 2,2'-bipyridines have found widespread use in asymmetric homogeneous catalysis.[9] 1,10-phenanthroline is of particular importance for organometallic chemistry as a very rigid ligand capable of coordinating various metal ions.[10-11] Metal complexes of 1,10-phenanthroline have been used for a broad range of catalytic reactions, such us aryl and alkenyl aminations and amidations,[12-16] oxidation of

alcohols,[17] conjugate addition reactions,[18-19] and cyclopropanations[20]. Applications of chiral 1,10-phenanthroline derivatives in asymmetric catalysis are still modest, but include Pd-catalysed allylic alkylation of allylic acetates,[10,21-23] Cu-catalysed cyclopropanation of styrenes,[24-26] Rh-catalysed asymmetric hydrosilylation[27] and transfer hydrogenation of ketones,[28] Cu-catalysed allylic oxidation of alkenes[29], Sc and Ir catalysed ring-opening reactions of *meso*-epoxides and Cu-catalysed amination of β-keto esters.[30]

Aqua complexes of 2,2'-bipyridine and 1,10-phenanthroline[31-34] have the particular advantage of displaying good stability, activity and solubility for the conditions required by proteins in solution (aqueous media, mild temperatures, atmospheric pressure). Ogo and co-workers reported the synthesis of water soluble $[Cp^*\text{-}Ir^{III}(bpy)(OH_2)]^{2+}$ and $[(\eta^6\text{-}C_6Me_6)\text{-}Ru^{II}\text{-}(bpy)(H_2O)]^{2+}$ catalyst precursors for transfer hydrogenation of carbonyl compounds in aqueous conditions.[31-32] Notably, the activity of these catalysts displayed a strong pH dependence. The catalysts developed by Ogo were exploited by Ward and co-workers to synthesize artificial transfer hydrogenases, via supramolecular anchoring of biotinylated $[\eta^6\text{-}(C_6Me_6)Ru(Biot\text{-}bipy)H_2O]^{2+}$ into streptavidin, for the asymmetric reduction of acetophenone in water.[35] Also for pH-dependent transfer hydrogenation of carbonyl compounds, Süss-Fink and co-workers explored the catalytic potential of seven $[(\eta^6\text{-arene})Ru(N\cap N)(OH_2)]^{2+}$ complexes, involving 1,10-phenanthroline or its 5-NO_2 or 5-NH_2 derivatives as $N\cap N$ ligand.[33-34] So far, these organometallic catalysts have not been used in the synthesis of artificial metalloenzymes.

Artificial metalloenzymes based on nitrogen-containing metal complexes and DNA[36-44] or proteins[45-64] as chiral environment have been previously reported in literature. Mainly two types of nitrogen-ligands have been used for the synthesis of artificial metalloenzymes. One category is formed by large heterocyclic tetradentate ligands like symmetric M^{III} Schiff-base complexes or ligands related structurally to porphyrins (corroles, phthalocyanines).[49-59,64] Their metal complexes were introduced either to myoglobin, a protein possessing a heme prosthetic group suitable for replacement or to serum albumins, proteins which possess a natural high affinity for hydrophobic compounds.

M^{III} Schiff-base complexes (Figure 1) were incorporated into apo-myoglobin mutants, via dative anchoring[50,57-58] or covalent anchoring.[59] Even if the enantioselectivities

obtained for sulfoxidation of thioanisole using these artificial metalloenzymes were not very high (up to 33% e.e. for dative incorporation and up to 51% e.e. for covalent anchoring), these results proved that asymmetric reactions could be performed by employing chiral protein cavities and symmetric metal complexes.

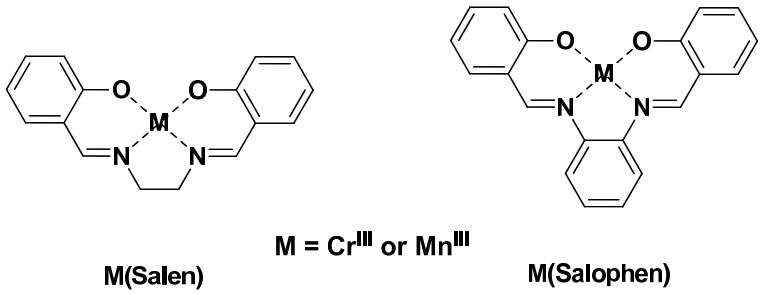


Figure 1 Schiff-base complexes inserted into apo-myoglobin mutants[50,57-59]

A drawback for dative incorporation of symmetric M^{III} Schiff-base complexes into proteins is the requirement of histidine coordination to the metal ion, in order to control and stabilize the metal complex location inside the protein cavity. This requirement added to the requirement of supramolecular affinity between the protein scaffold and symmetric M^{III} Schiff-base complexes tremendously limits the number of suitable proteins for this strategy.

Artificial metalloenzymes based on serum albumins and porphyrin related ligands (Mn^{III} and Fe^{III} corroles,[55] copper-phthalocyanines[56]) displayed up to 74% e.e. for asymmetric oxidation of prochiral sulfides to sulfoxides by H_2O_2[55] and 85-98% e.e. for Diels-Alder reactions.[56] However, the applicability of this strategy is also limited.

The second category of nitrogen-ligands used for the synthesis of artificial metalloenzymes is represented by bidentate chelating nitrogen ligands, for example derivatives of 2, 2'-bipyridine, sulfonated diamines or 1,10-phenanthroline. Their metal complexes were coupled via covalent or supramolecular anchoring to proteins[35,45-48,60-62] or DNA.[36-44]

Iodoacetamido-1,10-phenanthroline was used by Distefano and co-workers to modify the cysteine residue of two proteins: ALBP and IFABP (Figure 2). Enantioselectivities varying from 31% to 86% e.e. were reported for the hydrolysis of several unactivated amino acid esters and the hydrolysis of amides catalysed by ALPB-Phen-Cu(II).[61] Genetic optimization - site directed mutagenesis applied in order to change the cysteine location inside protein's cavity - was used to enhance the rate and selectivity of the hybrid catalysts.[62]

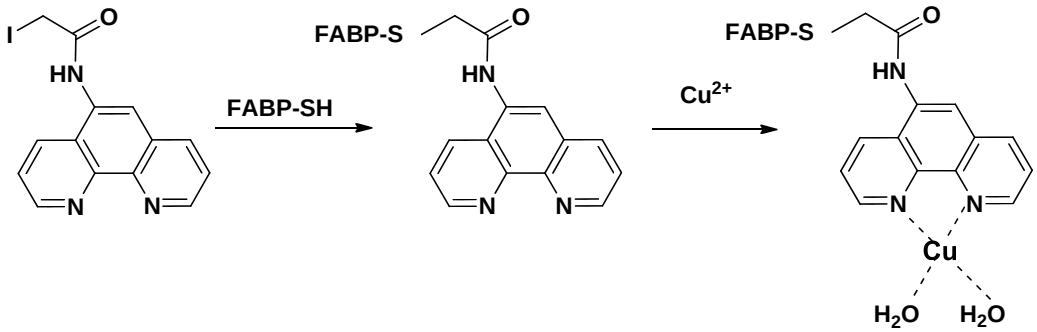


Figure 2 Cu(II)-Phenanthroline conjugates of fatty acid binding proteins used as hybrid catalysts for ester and amide hydrolysis[61]

Reetz and co-workers reported covalent incorporation of a Pd complex of *N*-(di(pyridin-2-yl)methyl)-maleimide (Figure 3a) and a Rh-complex of 5-maleimido-1,10-phenanthroline (Figure 3b) into the protein tHisF-C9AD11C, a thermostable enzyme from *Thermotoga maritima*.[65-66] The artificial metalloenzyme containing the Rh complex was tested for the reduction of benzoyl formic acid to mandelic acid, but low activity was obtained. Moreover, unspecific modifications of the protein host with the rhodium complex were observed, most probably due to undesired direct interaction of the protein with the rhodium centre rather than unspecific covalent modification of protein with the metal complex.[65] The Pd complex of N-(di(pyridin-2-yl)methyl)-maleimide (Figure 3a) was coupled to papain. The resulting artificial metalloenzyme was used as catalyst for the hydrogenation of dimethyl itaconate, but low activity and no enantioselectivity were obtained.[60]

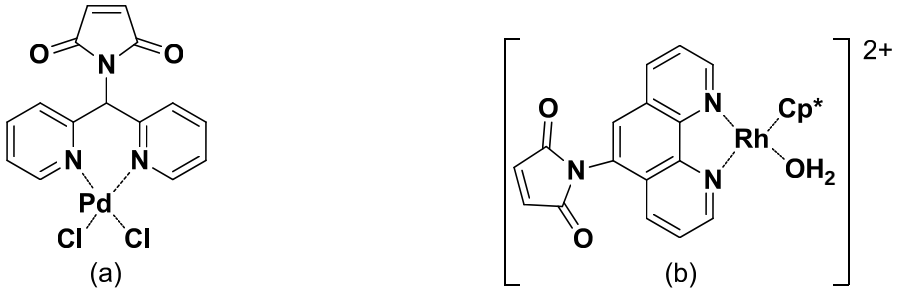


Figure 3 Pd complex of N-(di(pyridin-2-yl)methyl)-maleimide (a) and Rh-complex of 5-maleimido-1,10-phenanthroline (b) covalently coupled to tHisF-C9AD11C[65-66]

A biotinylated Rh complex of a di(pyridin-2-yl)methane derivative (Figure 4) was coupled to streptavidin and tested by Reetz and co-workers for the enantioselective transfer hydrogenation of acetophenone, ethyl pyruvate and α,α,α trifluoro-acetophenone. The catalyst was active, but low enantioselectivities (<5% e.e.) were obtained for the tested reactions.[67]

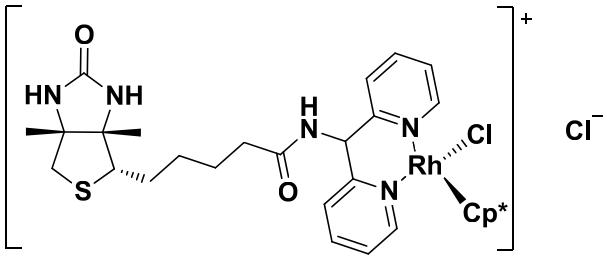


Figure 4 Biotinylated Rh complex of di(pyridin-2-yl)methanamine which was coupled to streptavidin and used as catalyst for transfer hydrogenation reactions[67]

Ward and co-workers synthesized and chemogenetically optimised transfer hydrogenases based on biotinylated d^6-piano stool complexes of N-ligands [η^n-(C_nR_n)M(**Biot-L**)] (M=Ru, Rh or Ir) incorporated to streptavidin mutants, for the enantioselective transfer hydrogenation of prochiral ketones.[45-46,68-69] The first studies involved biotinylated Ru-bipyridine complexes. However, high temperature and low pH were required in order to obtain good activity and enantioselectivity with these complexes.[35] The problem was overcome by replacing the bipyridine ligand with *ortho*-, *meta*- and *para*- biotinylated sulfonamides (Figure 5), which afford metal complexes highly active and enantioselective under conditions compatible with proteins in solution.[45-46,68-69]

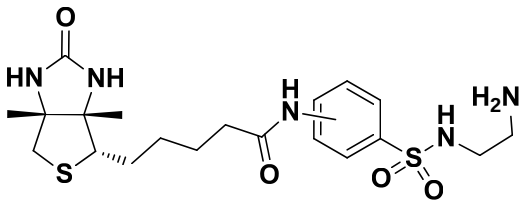


Figure 5 *Ortho*-, *meta* and *para*- biotinylated amino-sulfonamides used for the synthesis of streptavidin based artificial metallo-enzymes[45-46,68-69]

DNA has also been exploited as chiral biomolecule for the development of hybrid catalysts of bidentate chelating nitrogen ligands. The first generation of DNA-based catalysts was based on ligands containing a metal-binding domain linked through a spacer to a DNA intercalator.[36] For the second generation of DNA catalysts, the DNA binding moiety was integrated into the metal-binding domain.[41] Cu(II) complexes of 2,2'-bipyridines and 1,10-phenanthroline, known to bind to DNA in a non-covalent manner, were employed for the synthesis of this type of catalysts, displaying high enantioselectivities (up to 99%) for Diels-Alder reactions[37] (Figure 6) and Michael additions[38] in water. The scope of DNA-based catalysis was extended and high

enantioselectivities were reported for Friedel-Crafts alkylations,[42] fluorination reactions[43] and epoxide hydrolysis.[44]

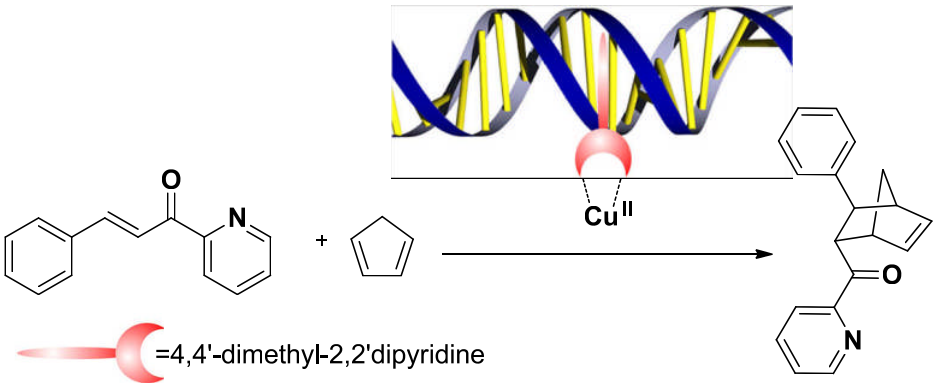


Figure 6 Cu-(4,4'-dimethyl-2,2'dipyridine$(NO_3)_2$)-DNA catalyzed Diels-Alder cycloaddition[37]
Reprinted from E.W. Dijk, B.L. Feringa and G. Roelfes, *Tetrahedron: Asymmetry*, **2008**, *19 (20)*, 2374-2377. Copyright 2008, with permission from Elsevier.

A Cu(II) complex of (2,2'-bipyridine)-5-carboxylic acid covalently coupled through an alkyl linker to a DNA double strand was used as enantioselective catalyst for the same Diels-Alder reaction (Figure 6). The hybrid catalyst containing the Cu-bipyridine complex covalently coupled to DNA showed similar conversion, but significantly higher enantioselectivity than the one having the Cu(II) complex coupled non-covalently, for the same DNA strand. This clearly demonstrates the advantage of the covalent coupling strategy in the synthesis of a well defined hybrid catalyst.[40]

As observed, 1,10-phenanthroline, 2,2'-bipyridine and di(pyridin-2-yl)methane derivatives have been used for the synthesis of protein or DNA based hybrid catalysts and tested for enantioselective catalytic reactions. In particular the hybrid catalysts formed by 1,10-phenathroline-FABPs and 1,10-phenanthroline- and 2,2'-bipyridine -DNA proved to be good chiral inducers for several catalytic reactions. The wide applicability of this type of N-ligands in homogeneous catalytic reactions allows the applicability of the resulting class of hybrid catalysts to a potential large number of reactions. Moreover, covalent anchoring of this type of ligands into proteins or DNA allows a large flexibility in terms of varying the protein/DNA scaffold or of the ligand/metal complex design. Nevertheless, the number and applicability of the artificial metalloenzymes based on N-ligands reported has been limited so far.

Considering the promising results previously reported for hybrid catalysts based on 1,10-phenanthroline and 2,2'-bipryridine ligands, I explored covalent incorporation of this type of ligands into three different protein structures (ALBP, Δ98Δ and AppA). Two strategies were tested for covalent incorporation of N-ligands: via thioester bond

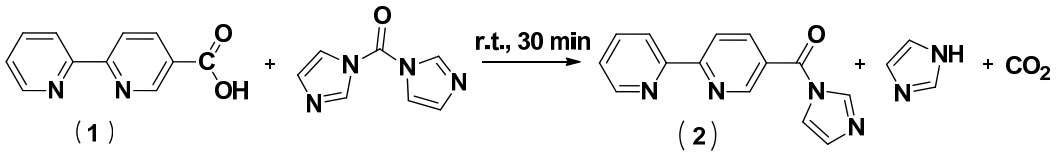
formation between a thiol group of a cysteine residue and an activated carboxylic acid group of an N-ligand and via Michael addition to the maleimide double bond of a maleimide functionalized ligand.

III.2 Results and Discussion

III.2.1 Coupling of carboxylic acid functionalised N-ligands

Ward and co-workers were the first to show incorporation of bipyridine ligands into a protein scaffold and to investigate the catalytic properties of the resulting artificial metalloenzyme. They synthesized artificial transfer hydrogenases by supramolecular anchoring of biotinylated $[\eta^6\text{-}(C_6Me_6)Ru(\text{Biot-bipy})H_2O]^{2+}$ into streptavidin.[35] Inspired by these studies and also considering the wide range of applications of the bipyridine ligands,[9,70] I explored the covalent incorporation of carboxylic acid functionalized 2,2'-bipyridine in several protein scaffolds.

The three proteins previously used for covalent incorporation of phosphine ligands (mutants of ALBP, Δ98Δ and AppA, see Chapter II) were also used for incorporation of (2,2'-bipyridine)-5 carboxylic acid via thioester bond formation. The choice of the proteins was based again on the possibility of comparing the activity and the selectivity of artificial metalloenzymes having different cavity structures and subsequently different degrees of encapsulation of the organometallic catalyst into the protein scaffold.

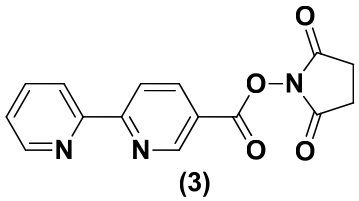


Scheme 1 Carbonyl diimidazole activation of 2,2'-bipyridine-5-carboxylic acid

The first method tested for covalent coupling of (2,2'-bipyridine)-5 carboxylic acid was the reaction between the active imidazolide derivative **2** formed upon CDI activation of **1** (Scheme 1) and the thiol group of a selected protein. Still, the coupling efficiency of an imidazolyl derived from a pyridine based carboxylic acid was very low, compared to the benzoic acid derived imidazolyl compounds, used previously for phosphine

incorporation (see Chapter II). No significant differences in coupling efficiency were observed regarding the protein scaffold involved. The highest efficiency, about 20% according to LC-MS (ES+) analysis, was found for AppA Y21C. These values were not corrected for the difference in ionization efficiency between the modified and unmodified proteins. Attempts to quantify this difference were not successful, because of the low coupling efficiency.

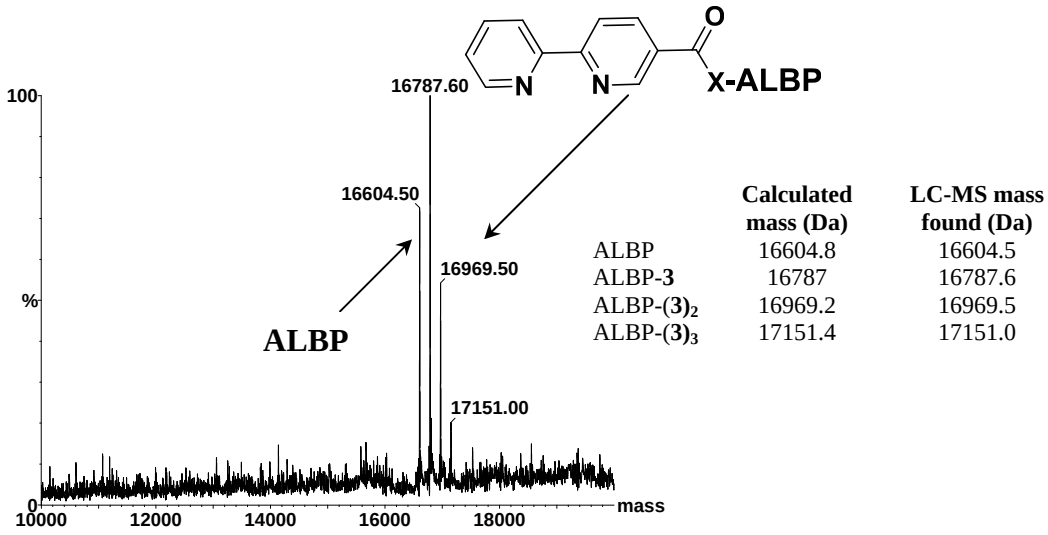
When proteins were modified covalently with CDI activated benzoic acid derivatives (see Chapter II), good coupling efficiencies were obtained.[71] This result allowed us to conclude that the rate of hydrolysis of CDI activated benzoic acids is slower than the rate of their coupling to proteins, which is an overnight reaction. However, considering the very low coupling efficiency of CDI activated bipyridine-carboxylic acids, two hypotheses were formulated. The first possibility was that the imidazolide **2** formed as product of the reaction between 2,2'-bipyridine 5-carboxylic acid and CDI in dimethylformamide is highly reactive and, although it is stable in dimethylformamide, hydrolyses rapidly in aqueous conditions. Consequently its coupling efficiency is very low. A second possibility was that the hydrolysis rate of **2** is lower than its coupling rate to proteins and the coupling efficiency to proteins is high, but the modified proteins are not stable and the bipyridine-carboxylic acid **1** hydrolyses rapidly after its coupling to proteins. Therefore the stability of the imidazolide **2** in presence of water was investigated. After addition of 0.5 ml of D_2O to CDI activated 2,2'-bipyridine 5-carboxylic acid in 0.5 ml of DMF-d_6, 40% of the imidazolide **2** was hydrolysed during the first three minutes, as observed by ^1H NMR, complete hydrolysis being achieved after the first hour of the reaction. Therefore I concluded that the hydrolysis rate of the imidazolide **2** is higher than the rate of the covalent modification of protein with **2**, which is usually an overnight reaction, causing the very low efficiency of protein modification.



(3)

Figure 7 Structure of the N-succinimidyl ester of 2,2'-bipyridine-5-carboxylic acid

The N-succinimidyl ester **3** of the ligand **1** is more stable under aqueous conditions than the imidazolide derivative **2** and therefore its hydrolysis rate might be lower than the rate of coupling to proteins. Consequently, its coupling efficiency to proteins is significantly improved. However, incorporation of the NHS activated carboxylic acid bipyridine into proteins was not selective toward cysteine modification. It was not possible to obtain complete protein mono-modification and the attempts to improve the coupling efficiency by increasing the molar excess of modification reagent resulted in multiple modifications of the protein host (Figure 8).

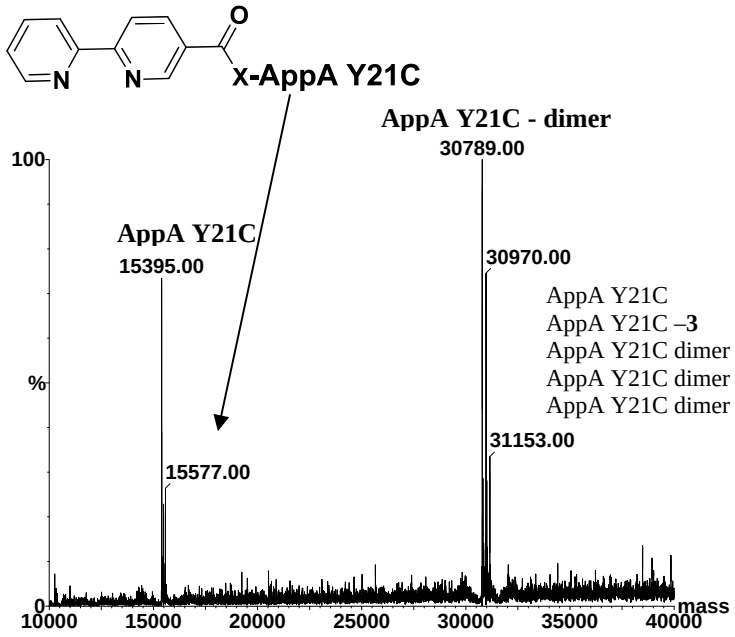


	Calculated mass (Da)	LC-MS mass found (Da)
ALBP	16604.8	16604.5
ALBP-**3**	16787	16787.6
ALBP-(**3**)$_2$	16969.2	16969.5
ALBP-(**3**)$_3$	17151.4	17151.0

Figure 8 LC-MS (ES$^+$) spectrum of ALBP - **3**
(unspecific modifications of the protein were obtained when higher molar excess of modification reagent was used for the coupling reaction)

Considering the efficient coupling of **3** to ALBP in 50 mM Tris buffer at pH 7-8, I can assume that the hydrolysis rate of **3** for the conditions tested is lower than protein modification. However, when coupling of **3** to the AppA mutants was done in 50 mM Tris buffer, in presence of 3M imidazole required to facilitate the flavin cofactor removal from the enzyme active site, no coupling was observed irrespective of the molar excess of ligand used. One cause could be the rapid hydrolysis of the NHS activated acid induced by the high concentration of nucleophilic imidazole, in addition to the increased reactivity of the ester bond induced by the electron withdrawing effect of pyridine ring. Indeed, when coupling of **3** to AppA was done in absence of imidazole, the coupling efficiency was highly improved. The AppA mutants were first maintained in 3M imidazole buffer solution to facilitate the flavin removal, and then the imidazole was removed by over-night dialysis against 50 mM Tris buffer. Subsequent

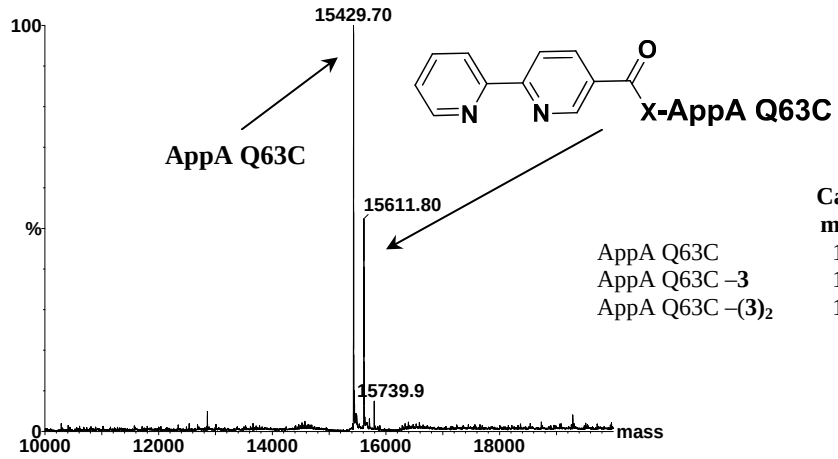
coupling of **3** resulted in efficient, but unspecific coupling of the NHS-activated ligand, for both Y21C and Q63C mutants of AppA.



	Calculated mass (Da)	LC-MS mass found (Da)
AppA Y21C	15395.1	15395.2
AppA Y21C –**3**	15577.3	15577.1
AppA Y21C dimer	30790.2	30789
AppA Y21C dimer –(**3**)	30972.4	30970
AppA Y21C dimer –$(\mathbf{3})_2$	31154.6	31153

Figure 9 LC-MS (ES^+) spectrum of AppA Y21C - **3**, after imidazole treatment of the unmodified protein

LC-MS (ES^+) analysis revealed formation of the dimer of AppA Y21C. Potential changes in the mutant folding generated by flavin removal following the imidazole treatment and dialysis, could expose the thiol of cysteine 21 to the surface and facilitate the dimer formation via a disulfide bridge. Consequently the dimer possesses no free cysteine. Nevertheless, efficient coupling of **3** to the AppA Y21C dimer was observed by LC-MS (ES^+), proving the lack of specificity towards cysteine of the method (Figure 9).



	Calculated mass (Da)	LC-MS mass found (Da)
AppA Q63C	15430.1	15429.7
AppA Q63C –**3**	15612.3	15611.8
AppA Q63C –$(\mathbf{3})_2$	15794.5	15793.9

Figure 10 LC-MS (ES+) spectrum of AppA Q63C - **3**, after imidazole treatment of the unmodified protein

The dimer formation was not observed for the AppA Q63C mutant. Yet incomplete conversion and multiple modification of the protein with **3** were observed (Figure 10) when 7.5-fold excess ligand was used for the coupling reaction.

Flavin cofactor removal might be also possible to achieve after its reduction, firstly, followed by its removal, based on reported literature procedures,[72] in the reduced form.

In conclusion, although CDI activation of benzoic and aliphatic carboxylic acid functionalized phosphine ligands is compatible with aqueous coupling conditions,[71] this method was not suitable for the bipyridine ligand having the carboxylic group attached to a pyridine ring. The reactivity of the imidazolide intermediate of a carboxylic acid functionalized pyridine is higher than that of a benzoic or aliphatic carboxylic acid derivative. Consequently, hydrolysis of pyridine based carboxylic acid-imidazolides is faster than that of benzoic acid imidazolides or aliphatic carboxylic acid imidazolides.

On the other hand, the reactivity of NHS activated carboxylic acid functionalised pyridines is lower than the reactivity of the CDI activated ones, and subsequently their hydrolysis occurs at a lower rate. As a result a higher coupling efficiency of the NHS activated ligands was obtained.

Regarding the nature of the ligand (phosphine or nitrogen-containing), for the same molar excess ligand (e.g. 25-fold excess for ALBP modification), monomodification with *p*-diphenylphosphino-benzoic acid was obtained, while multiple modifications were obtained when 2,2'-bipyridine-5-carboxylic acid was coupled. The electron withdrawing effect of pyridine is higher compared to that of the benzoic acid ring. As a result, the acidity of (bi)pyridine carboxylic acids is higher than that of benzoic acids and consequently the reactivity of activated bipyridine carboxylic acids is higher. This can explain the reduced selectivity towards cysteine of the activated bipyridine carboxylic acids, which are more reactive than phosphine carboxylic acids and, therefore, more prone to react with the NH_2 group of lysines.

As observed, the main drawback of the covalent incorporation of carboxylic acid functionalized ligands in proteins via thioester bond formation with a cysteine residue seems to be the lack of specificity. Nevertheless, selective and efficient protein modification with benzoic acid and aliphatic acid functionalized phosphine ligands and their metal complexes was reported in our group.[71] These results suggest that the selectivity of this method depends on the host-protein structure, in particular on the number and the reactivity of the lysine residues.

III.2.2 Coupling of maleimide functionalised N-ligands

The lack of selectivity towards cysteine of the carboxylic acid functionalized nitrogen-containing ligands prompted us to explore a different method for covalent incorporation of this class of ligands into proteins. Among the common methods used for selective cysteine modification, Michael addition to maleimide functionalized ligands was selected. This method was previously applied for covalent incorporation of a library of N-ligands and their metal complexes into proteins by Reetz and co-workers.[60,65-66]
Applications of chiral phenanthroline derivatives in asymmetric catalysis have been previously reported.[30] Based on these facts, but also considering the successful covalent incorporation of 1,10-phenanthroline derivatives in ALBP[61] and tHisF-C9AD11C[65] and the good enantioselectivities induced by ALBP as protein scaffold for the hydrolytic reactions catalysed by the ALPB-Phen-Cu(II) complex,[61] I investigated the covalent incorporation of 5-maleimido-1,10-phenanthroline to ALBP, Δ98Δ and AppA mutants.
Ru-complexes of the type $[(\eta^6\text{-arene})Ru(N\cap N)(Cl)]^+$ containing N,N-chelating dipyridylamine ligands (Figure 11) were reported as efficient catalysts for transfer hydrogenation of aryl ketones in water.[73]

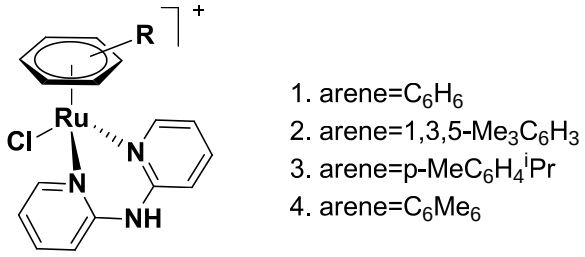


Figure 11 Structure of the dipyridylamine-Ru complexes used as catalyst for transfer hydrogenation of aryl ketones[73]

Inspired by this study, the 3-(N-maleimido)-(N,N-bis(2-picolyl))-propanamide ligand (Figure 12) was synthesized in our group and its covalent incorporation in the above mentioned proteins was tested. The choice for a dipicolylamine over a dipyridylamine derivative was motivated by the higher flexibility of the third N atom, which would allow its potential coordination to the metal centre and subsequent stabilization of metal complexes. In addition, the synthesis of ligand **5** could be easily achieved via DCC/HOBt coupling of 3-maleimidopropionic acid to the commercially available di-(2-picolyl)amine.

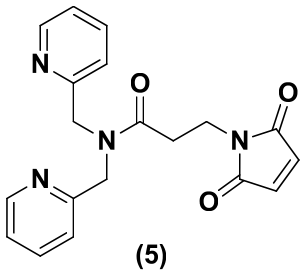


Figure 12 Structure of the 3-(N-maleimido)-(N',N'-bis(2-picolyl))-propamide

The chemical modification of proteins was performed according to standard conditions reported in literature for other systems[66,74] and the bioconjugates were characterized by LC-MS (ES^+).

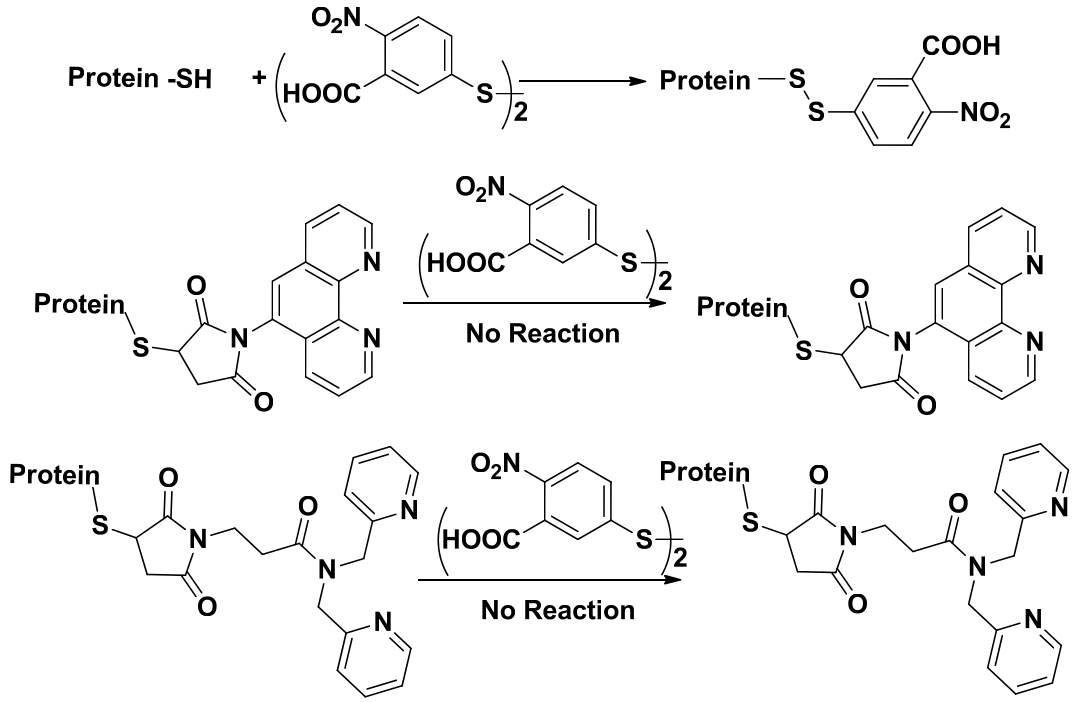


Figure 13 Reaction of (modified) proteins with Ellman's reagent (DTNB) as a method for the quantification of the cysteine modification via Michael addition to maleimide functionalized N-ligands[75]

Cysteine conversion upon Michael addition to maleimide functionalized N-ligands was tested by subsequent treatment with Ellman's reagent[75-76] followed by LC-MS (ES^+) analysis. Under the same reaction conditions, treatment with Ellman reagent[76] converts the unmodified proteins to the corresponding 5-thio-(2-nitro benzoic acid) conjugates (Table 1), while the modified proteins, displaying no free cysteine, show no modification with **DTNB** (Figure 13, Figure 14, Figure 15).

Table 1 Chemical modification of apo-proteins with **DTNB**

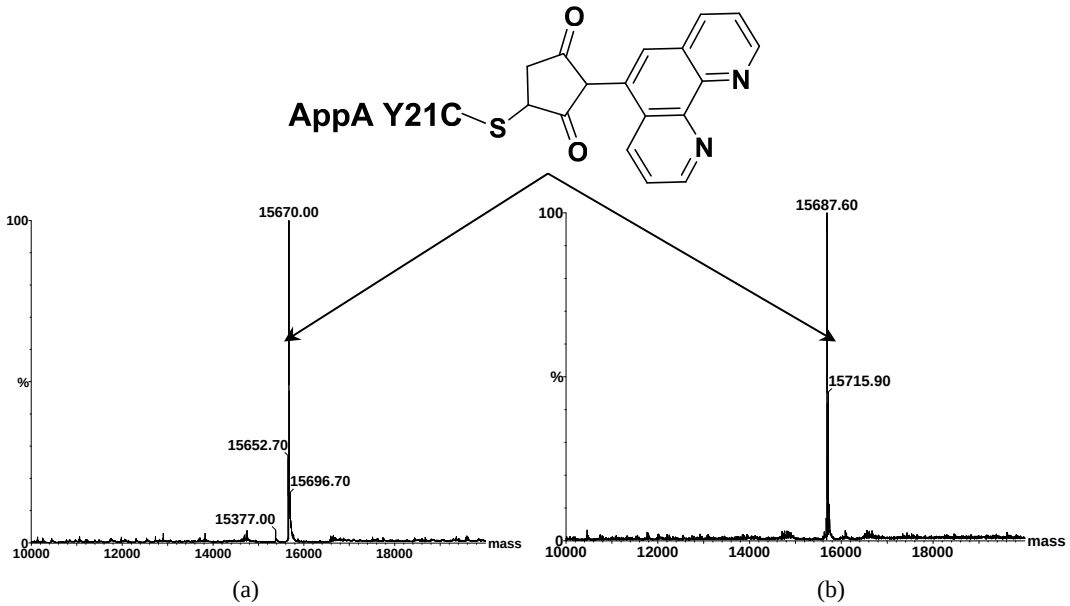
	Calculated mass (Da)	LC-MS mass found (Da)
ALBP-**TNB**	16802.0	16802.0
D98D-**TNB**	13620.2	13620.2
AppA Q63C-**TNB**	15627.3	15627.2
AppA Y21C-**TNB**	15592.3	15592.2

Contrary to the covalent coupling of carboxylic acid functionalized ligands, the incorporation of maleimide functionalized ligands into proteins via thioether bond proved to be highly selective towards cysteine. This was evidenced by the complete mono-modification of proteins obtained for the ligands **4** and **5** (Table 2 and Table 3). The high selectivity was observed for all the four proteins tested for chemical modification.

In conclusion, Michael addition to maleimide functionalized ligands turns out to be a suitable method for selective modification of virtually any protein displaying a single accessible, reactive cysteine.

Table 2 Proteins chemical modification with **4**

	Calculated mass (Da)	LC-MS mass found (Da)
ALBP-**4**	16880.1	16880.1±0.5
Δ98Δ-**4**	13698.3	13698.6±0.6
AppA Y21C-**4**	15670.4	15670.0±0.5
AppA Q63C-**4**	15705.4	15704.9±0.6

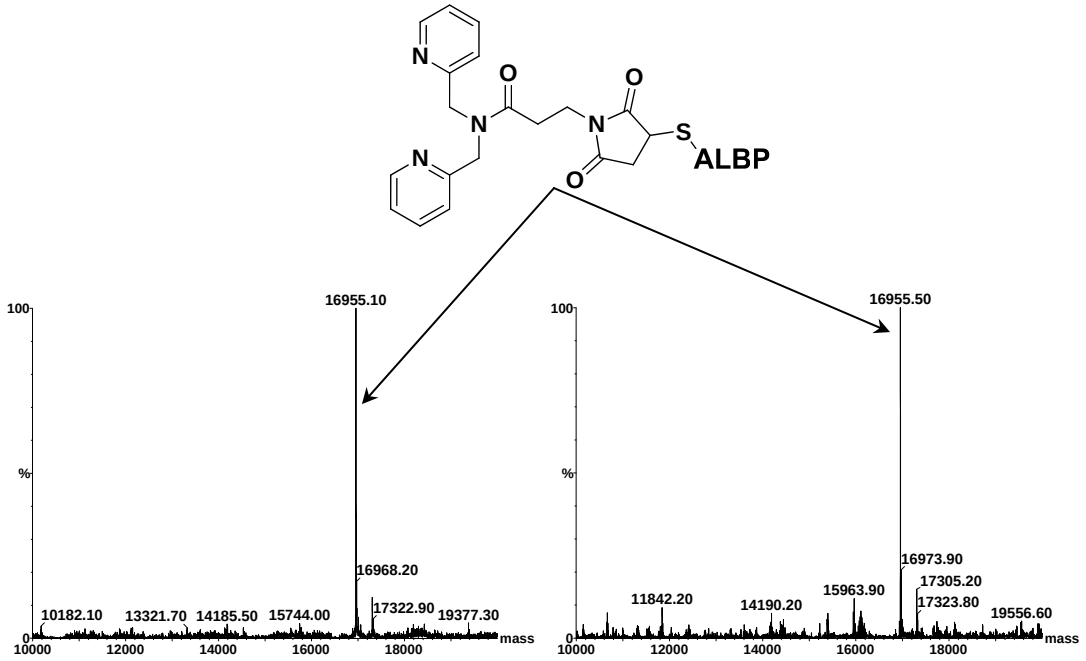


(a) (b)

Figure 14 (a) LC-MS (ES+) spectrum of AppA Y21C - **4; (b)** LC-MS (ES+) spectrum of AppA Y21C-**4** after treatment with Ellman's reagent

Table 3 Chemical modification of proteins with **5**

	Calculated mass (Da)	LC-MS mass found (Da)
ALBP-**5**	16955.2	16955.0±0.2
Δ98Δ-**5**	13773.4	13773.5±0.9
AppA Y21C-**5**	15745.5	15745.4±0.4
AppA Q63C-**5**	15780.4	15779.5±0.6



(a) (b)

Figure 15 (a) LC-MS (ES+) spectrum of ALBP - **5**; **(b)** LC-MS (ES+) spectrum of ALBP-**5** after the reaction with Ellman's reagent

III.3 Conclusions

Incorporation of carboxylic acid functionalized 2,2'-bipyridine into proteins, in aqueous conditions and at room temperature, could be easily achieved by NHS activation of the carboxylic group followed by covalent coupling of the activated ligand. However, the selectivity towards thioester bond formation is low, multiple chemical modifications of proteins or incomplete conversions were obtained. The low selectivity could be explained by the increased reactivity of the activated bipyridine carboxylic acids compared to the benzoic acids, due to the difference in the inductive effect of the two aromatic rings. The NHS activated 2,2'-bipyridine 5-carboxylic acid is stable in aqueous conditions, but in presence of high concentrations of the nucleophilic imidazole, the stability appears to be low.

CDI activation of carboxylic acid functionalized bipyridines proved not to be an efficient method for coupling this class of ligands to proteins. The imidazolide intermediate of 2,2'-bipyridine 5-carboxylic acid formed upon activation with CDI in dimethylformamide is highly reactive and hydrolyses rapidly under aqueous conditions, before coupling to proteins.

Bioconjugation with maleimide functionalized N-ligands can be easily achieved and it is very selective towards cysteine modification, as was demonstrated for the three structurally different proteins. This method can be easily applied to virtually any protein displaying a single reactive cysteine and allows a wide range of possibilities in terms of cofactor design. Herein, selective incorporation of two maleimide functionalized N-ligands having different heterocyclic structures: a 1,10-phenanthroline derivative and a tridentate N,N-bis(2-picolyl))-amide derivative was demonstrated.

The coordination properties of 1,10-phenanthrolines are well-known, however their applications as chiral ligand templates in asymmetric catalysis are still limited.[11] The typical way for synthesis of chiral 1,10-phenanthrolines is the introduction of chiral groups in the peripheral region of the phenanthroline heterocycle. However, by using this method the stereogenic centre is located far from the aromatic ring and consequently far from the metal complex centre. This disadvantage is potentially overcome by embedding an achiral phenanthroline heterocycle within a chiral protein environment, thus having a higher control over the first and second coordination sphere of the metal complex and subsequently over the enantioselectivity. As a result of this project, the number of proteins modified with 1,10 phenanthroline derivatives, initially limited to FABP[61-62] and tHisF-C9AD11C,[66] is now enriched with another three different protein scaffolds: Δ98Δ, AppA Y21C and AppA Q63C. The synthesis of Cu(II) complexes of 1,10-phenanthroline modified proteins is straightforward[61-62] and their catalytic properties could be subsequently tested for the hydrolytic reactions reported by Distefano and co-workers[61] or for other Cu(II) catalysed reactions like Michael additions,[38] Diels-Alder reactions,[37] Friedel-Crafts alkylations[42] or epoxide hydrolysis.[44] The library of metal complexes of modified proteins and their use in asymmetric homogeneous catalytic reactions can be further expanded to other enantioselective reactions catalysed by metal complexes of chiral bipyridines or phenanthrolines. I reported successful incorporation of a Ru complex of 1,10 phenanthroline to ALBP and AppA mutants (see Chapter IV).

III.4 Experimental

III.4.1 General remarks

Chemicals were purchased from Aldrich and Fluka and were used as received. NMR spectra were recorded at room temperature on Bruker Avance 400 and 300 spectrometers. Positive chemical shifts (δ) are given (in ppm) for high-frequency shifts relative to a TMS reference (^{1}H and ^{13}C). ^{13}C spectra were measured with ^{1}H decoupling. Multiplicities are indicated by: s (singlet), d (doublet), dd (doublet of doublets), and m (multiplet).

Protein stability (apo and modified proteins) was assessed by Bradford assay[77] and SDS-PAGE. Modified proteins were analysed by LC-MS (ES+). Quantification of cysteine modification was done by reaction with Ellman's reagent (DTNB).[75]

III.4.1.a Whole protein mass spectrometry using ESI-TOF

The protein sample (20 μL, 5 pM/μL) was desalted on-line through a XTerra MS C8 2.1 × 10 mm column, eluted with an increasing acetonitrile concentration (2 % acetonitrile, 98 % aqueous 1 % formic acid to 98 % acetonitrile 2 % aqueous 1 % formic acid and delivered to an electrospray ionisation mass spectrometer (LCT, Micromass, Manchester, U.K.) which had previously been calibrated using myoglobin. An envelope of multiply charged signals was obtained and deconvoluted using MaxEnt1 software to give the molecular mass of the protein.

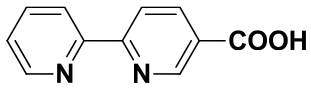
The sequence of ALBP reported by Bernlohr *et al*[78] was used as reference for the mass of ALBP C1A his-tagged mutant. For the mass of the Δ98Δ mutant, the Δ98Δ sequence reported by Delfino *et al*[79] was used as reference and for the mass of AppA mutants the AppA sequence reported by Laan *et al*[80] was used.

II.4.1.b Δ98Δ - reduction of disulfide bridges was performed every time before modification with carboxylic acid or maleimide functionalized compounds. Prior to the reduction reaction the protein solution was degassed by purging with argon for 8 hours. 6 μmol (6 eq.) of DTT were added from a freshly prepared stock solution of 0.5 M in water to a solution of 1 μmol of protein (1 eq.) in Tris·HCl (50 mM, 10 ml, pH 8.0) in a Schlenk tube. The reaction mixture was stirred at room temperature for 1 hour. The reduced protein solution was used for the coupling reactions without further purification. For each coupling reaction the molar excess of modification reagent was higher than 12-fold, to assure the right amount of modification reagent

for the potential modification of both the thiol groups of DTT and the thiol group of the target cysteine.

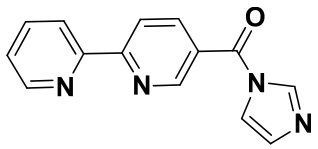
III.4.2 Synthesis of modification reagents

2,2' – bipyridine-5-carboxylic acid (1)



This compound was synthesised according to a reported literature procedure.[81]

[2,2'-bipyridin]-5-yl(1H-imidazol-1-yl)methanone **(2)**



[2,2'-bipyridin]-5-yl(1H-imidazol-1-yl)methanone was synthesised according to a modified literature procedure.[82-83]

In a Schlenk-tube, 2,2'–bipyridine-5-carboxylic acid (10mg, 50 μmol) and N,N'-carbonyldiimidazole (24.3 mg, 150 μmol) were dissolved in 1 ml dimethylformamide and the reaction mixture was stirred for 30 min at room temperature. The CDI activated ligand thus obtained was used directly for the chemical modification of proteins, without further purification.

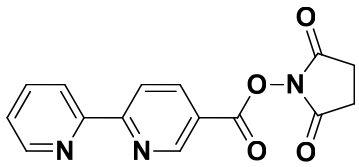
^{1}H NMR (400 MHz, d_7-DMF): 9.17 (d, J=1.9Hz, 1H), 8.81-8.82 (d, J=4.5Hz, 1H), 8.68-8.70 (d, J=8.2Hz, 1H), 8.57-8.59 (d, J=7.8Hz, 1H), 8.47-8.49 (dd, J=2.3, 8.1Hz, 1H), 8.42 (s, 1H), 8.13 (d, 1H), 8.12 (d, 1H), 8.06-8.11 (m, 1H), 7.57-7.60 (m, 1H); ^{13}C NMR (101 MHz, d_7-DMF): δ=165.0, 159.2, 154.5, 150.4, 149.9, 138.8, 137.6, 128.5, 125.2, 124.8, 121.5, 120.2, 118.2 ppm

[2,2'-bipyridin]-5-yl(1H-imidazol-1-yl)methanone hydrolysis

To follow the ligand **2** hydrolysis in time, in presence of water, the 2,2' – bipyridine-5-carboxylic acid was activated with CDI, in d_7-dimethylformamide, as described above. After that, 0.5 ml of D_2O was added to the CDI activated ligand in d_7-dimethylformamide (1ml) and the hydrolysis reaction was followed by ^{1}H and ^{13}C NMR.

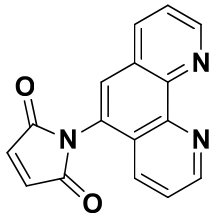
As a result of imidazolide hydrolysis, three minutes after D_2O addition the following NMR peaks corresponding to **2** were shifted to the corresponding peaks of the acid **1**: ^{1}H NMR (400 MHz, d_7-DMF, D_2O): δ=9.07 (d, J=1.3, 1H) was shifted to δ=9.10 (d, J=1.4, 1H); 8.72(d, J=4.8Hz, 1H) was shifted to 8.65(d, J=4.2Hz, 1H); 8.49 (d, J=7.9Hz, 1H) was shifted to 8.40 (d, J=8.2 Hz, 1H); 8.44(dd, J=2.0,8.3Hz, 1H) and 8.39(d, 1H) were shifted to 8.22-8.27 (m, 2H); 8.04 (m, 1H) was shifted to 7.98 (m, 1H); 7.60 (m, 1H) was shifted to 7.52 ppm (m, 1H). ^{13}C NMR (101 MHz, d_7-DMF, D_2O): δ=165.4 was shifted to 168.9; 157.3 was shifted to 154.4; 155.9 was shifted to 152.9; 148.5 was shifted to 148.3; 147.8 was shifted to 147.5; 137.7 was shifted to 136.9; 136.8 was shifted to 136.6; 129.3 was shifted to 131.5; 128.5 remained at 128.5; 124.0 was shifted to 123.2; 120.8 was shifted to 120.2 ppm.

N-hydroxysuccinimide 2,2-bipyridine-5-carboxylate **(3)**



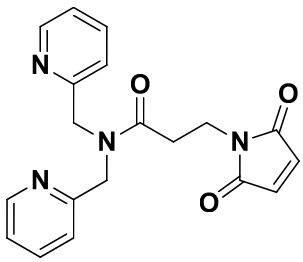
This compound was synthesised according to a reported literature procedure.[40]

5-maleimido-1,10-phenanthroline (4)



This compound was synthesised according to a reported literature procedure [65,84].

3-(N-maleimido)-(N,N-bis(2-picolyl))-propanamide (5)



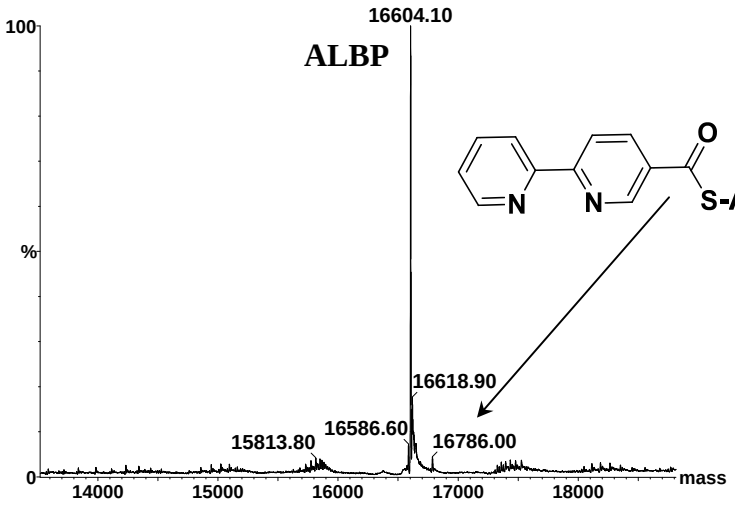
This compound was kindly provided by Peter Deuss (University of St Andrews, UK).

III.4.3 Chemical modification of proteins

Chemical modification with [2,2'-bipyridin]-5-yl(1H-imidazol-1-yl)methanone

ALBP

1.25 µmol (25 eq.) of **2** in 25 µl of dimethylformamide was added slowly to a solution of 50 nmol of protein (1 eq.) in potassium phosphate buffer (50 mM, 1 ml, pH 7.0). The cloudy reaction mixture was stirred at room temperature for 16h. The insoluble excess modification reagent was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), concentrated to 0.5 ml and then diluted to 10 ml with potassium phosphate buffer (50 mM, pH 7.0). This step was repeated four times after which the protein solution was concentrated to 1 ml and stored at 4°C.

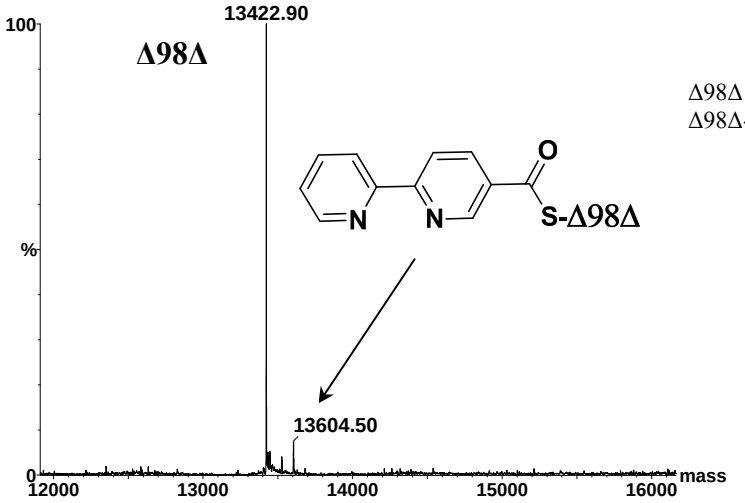


	Calculated Mass (Da)	LC-MS mass found (Da)
ALBP	16604.8	16604.1
ALBP-**2**	16787	16786

Figure 16 LC-MS (ES+) spectrum of ALBP after reaction with **2**

Δ98Δ

1.85 µmol (37 eq.) of **2** in 37 µl of dimethylformamide was added slowly to a degassed solution of 50 nmol reduced Δ98Δ (1 eq.) in 50 mM Tris·HCl (1ml, pH 8.0) containing DTT (6 eq.), in a Schlenk-tube, followed by stirring at room temperature for 16h under inert atmosphere. The insoluble fraction from the excess modification reagent was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO) and concentrated to 0.5 ml, followed by dilution to 10 ml with 50 mM Tris·HCl (pH 8.0). This step was repeated four times after which the protein solution was concentrated to 1 ml and stored at 4°C.

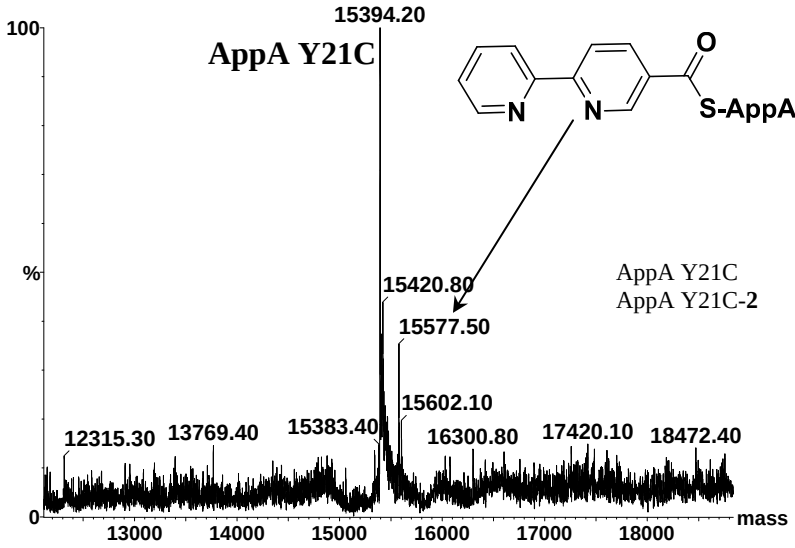


	Calculated Mass (Da)	LC-MS mass found (Da)
Δ98Δ	13423	13422.9
Δ98Δ-**2**	13605.2	13604.5

Figure 17 LC-MS (ES+) spectrum of Δ98Δ after the reaction with **2**

AppA Y21C

1.25 µmol (25 eq.) of **2** in 25 µl of dimethylformamide was added slowly to a solution of 50 nmol of AppA Y21C (1 eq.) in buffer (Tris·HCl 50 mM, 1ml) containing 3M imidazole, at pH 7.0. The cloudy reaction mixture was stirred at room temperature for 16h. The insoluble excess modification reagent was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO) and concentrated to 0.5 ml, followed by dilution to 10 ml with Tris·HCl (50 mM, pH 7.0). This step was repeated four times after which the protein solution was concentrated to 1 ml and stored at 4°C.



	Calculated Mass (Da)	LC-MS mass found (Da)
AppA Y21C	15395.1	15394.2
AppA Y21C-**2**	15577.3	15577.5

Figure 18 LC-MS (ES+) spectrum of AppA Y21C after the reaction with **2**

Chemical modification with N-hydroxysuccinimide 2,2-bipyridine-5-carboxylate

ALBP

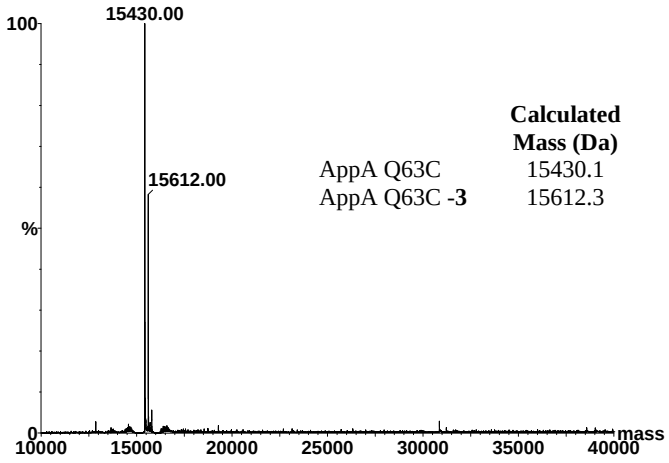
150 nmol (3 eq.) of **3** in 3.0 µl of dimethylformamide was added slowly to a solution of 50 nmol of protein (1 eq.) in potassium phosphate buffer (50 mM, 1 ml, pH 7.0). The cloudy reaction mixture was stirred at room temperature for 5h. The insoluble excess modification reagent was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), concentrated to 0.5 ml and then diluted to 10 ml with potassium phosphate buffer (50 mM, pH 7.0). This step was repeated four times after which the protein solution was concentrated to 1 ml and stored at 4°C.

AppA Y21C

50 nmol of protein (1 eq.) was maintained for 4 hours in buffer (Tris·HCl 50 mM, 1ml) containing 3M imidazole, at pH 7.0 and was further dialysed against Tris·HCl (50 mM, pH 7.0) to remove the imidazole. 375 nmol (7.5 eq.) of **3** in 7.5 µl of dimethylformamide was added to the protein solution after dialysis. The reaction mixture was stirred at room temperature for 5h. The insoluble excess modification reagent was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO) and concentrated to 0.5 ml, followed by dilution to 10 ml with Tris·HCl (50 mM, pH 7.0). This step was repeated four times after which the protein solution was concentrated to 1 ml and stored at 4°C.

AppA Q63C

Coupling of **3** to AppA Q63C was performed as described for coupling of **3** to AppA Y21C.



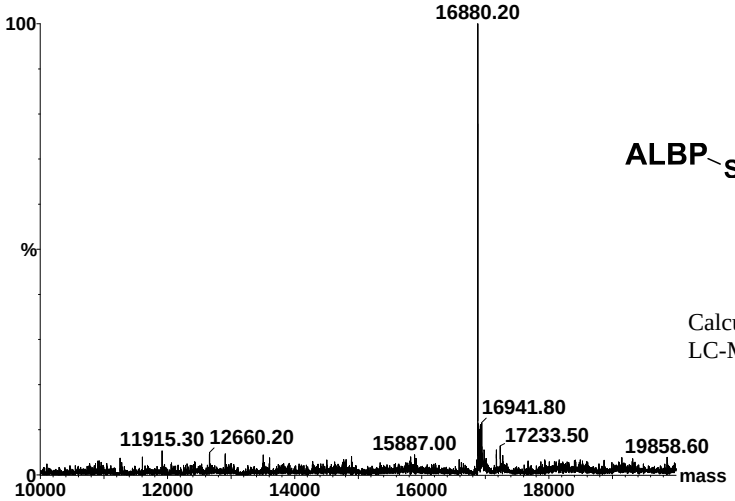
	Calculated Mass (Da)	LC-MS mass found (Da)
AppA Q63C	15430.1	15430.0
AppA Q63C -**3**	15612.3	15612.0

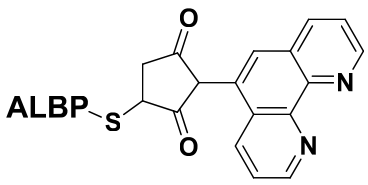
Figure 19 LC-MS (ES+) spectrum of AppA Q63C - **3**, after imidazole treatment of unmodified protein

Chemical modification with 5-maleimido-1,10-phenanthroline (4)

ALBP

500 nmol (10 eq.) of **4** in 10 µl of dimethylformamide was added slowly to a solution of 50 nmol of ALBP (1 eq.) in Tris·HCl (50 mM, 1ml, pH 8.0). The cloudy reaction mixture was stirred at room temperature for 5h. The insoluble fraction of the excess modification reagent was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO) diluted to 10 ml with Tris·HCl (50 mM, pH 8.0) and concentrated to 0.5 ml. This step was repeated four times after which the modified protein was concentrated to 1 ml and stored at 4°C. Yield: 100%.





Calculated Mass (Da)	16880.1
LC-MS mass found (Da)	16880.2

Figure 20 LC-MS (ES+) spectrum of ALBP - **4**

Δ98Δ

850 nmol (17 eq.) of **4** in 17 µl of dimethylformamide was added slowly to a degassed solution of 50 nmol of reduced Δ98Δ (1 eq.) in Tris·HCl (50 mM, 1ml, pH 8.0) containing DTT (6 eq.), in a Schlenk-tube, followed by stirring at room temperature for 5h under inert atmosphere. The insoluble fractions from the excess modification reagent were removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO) and concentrated to 0.5 ml, followed by dilution to 10 ml with Tris·HCl (50 mM, pH 8.0). This step was repeated four times after which the modified protein was concentrated to 1 ml and stored at 4°C. Yield: 100%.

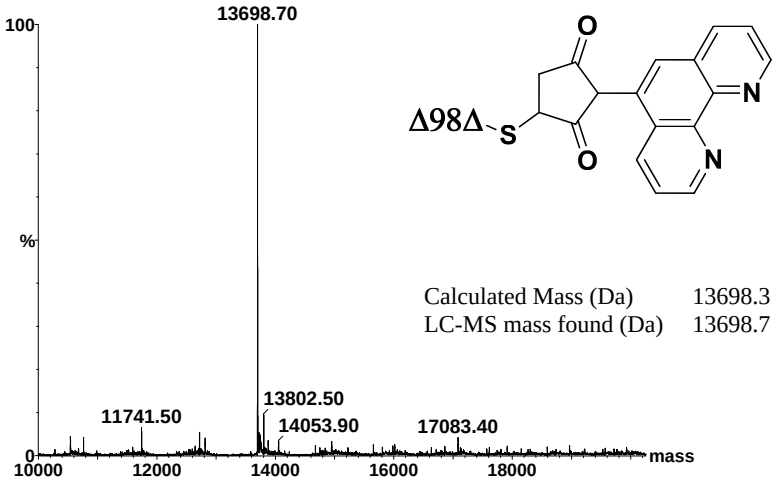


Figure 21 LC-MS (ES+) spectrum of Δ98Δ - **4**

AppA Y21C

1.25 µmol (25 eq.) of **4** in 25 µl of dimethylformamide was added slowly to a solution of 50 nmol of AppA Y21C (1 eq.) in buffer (Tris·HCl 50 mM, 1ml) containing 3M imidazole, at pH 7.0. The cloudy reaction mixture was stirred at room temperature for 5h. The insoluble excess modification reagent was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO) and concentrated to 0.5 ml, followed by dilution to 10 ml with Tris·HCl (50 mM, pH 7.0). This step was repeated four times after which the modified protein was concentrated to 1 ml and stored at 4°C. Yield: 100%.

AppA Q63C

Coupling of **4** to AppA Q63C was performed as described for coupling of **4** to AppA Y21C. Yield: 100%.

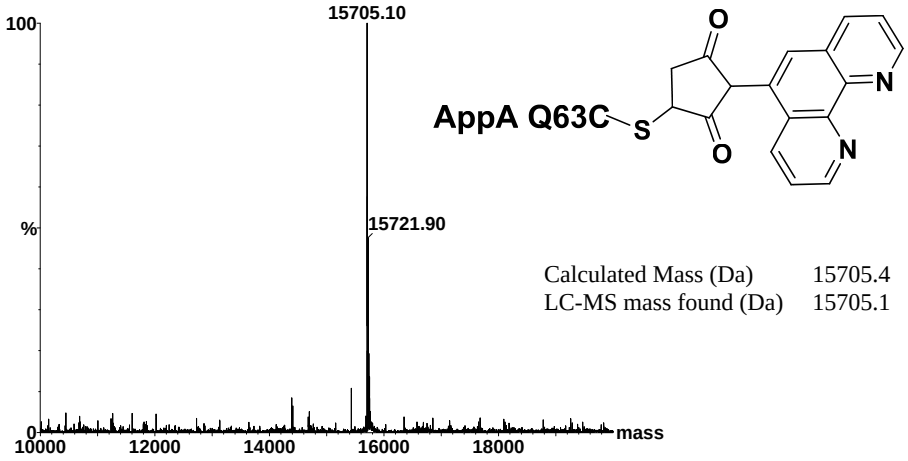


Figure 22 LC-MS (ES+) spectrum of AppA Q63C - **4**

Chemical modification with 3-(N-maleimido)-(N,N-bis(2-picolyl))-propanamide

ALBP

31.25 nmol (2.5 eq.) of **5** in 0.625 µl of dimethylformamide was added to a solution of 12.5 nmol of ALBP (1 eq.) in Tris·HCl (50 mM, 0.25 ml, pH 8.0). The reaction mixture was stirred at room temperature for 5h. To remove the excess modification reagent, the reaction mixture was transferred to a centrifugal concentrator (Vivaspin 2, 10,000 MWCO), diluted to 2 ml with Tris·HCl (50 mM, pH 8.0) and concentrated to 0.1 ml. This step was repeated four times after which the modified protein was concentrated to 0.25 ml and stored at 4°C. Yield: 100%.

Δ98Δ

187.5 nmol (15 eq.) of **5** in 3.75 µl of dimethylformamide was added to a degassed solution of 12.5 nmol of reduced Δ98Δ (1 eq.) in Tris·HCl (50 mM, 0.25 ml, pH 8.0) containing DTT (6 eq.), in a Schlenk-tube. The reaction mixture was stirred at room temperature for 5h, under inert atmosphere. The reaction mixture was transferred to a centrifugal concentrator (Vivaspin 2, 10,000 MWCO), diluted to 2 ml with Tris·HCl (50 mM, pH 8.0) and concentrated to 0.1 ml. This step was repeated four times after which the modified protein was concentrated to 0.25 ml and stored at 4°C. Yield: 100%.

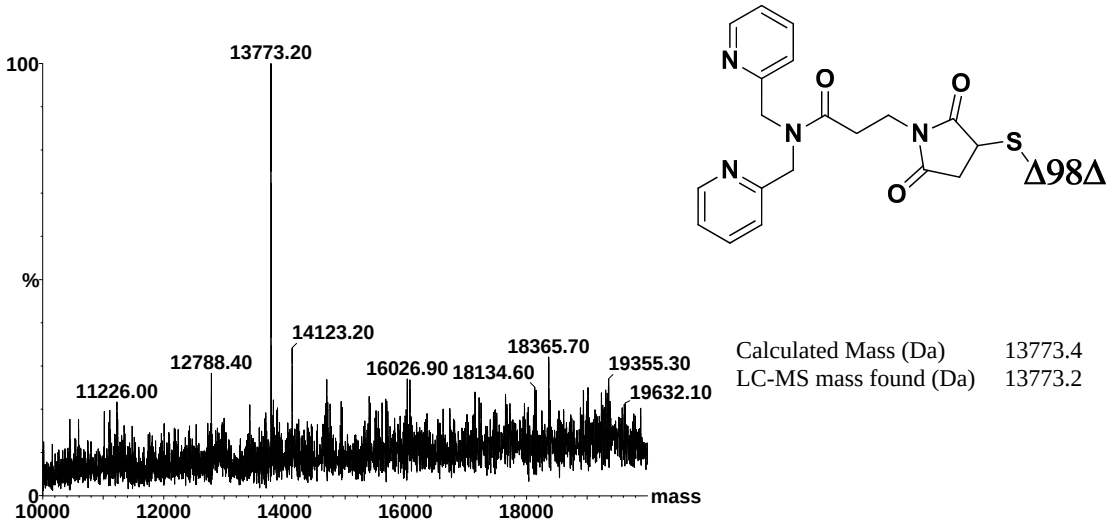


Figure 23 LC-MS (ES+) spectrum of Δ98Δ - **5**

AppA Y21C

312.5 nmol (25 eq.) of **5** in 6.25 µl of dimethylformamide was added to a solution of 12.5 nmol of AppA Y21C (1 eq.) in buffer (Tris·HCl 50 mM, 0.25 ml) containing 3M imidazole, at pH 7.0. The reaction mixture was stirred at room temperature for 5h. To remove the excess modification reagent, the reaction mixture was transferred to a centrifugal concentrator (Vivaspin 2, 10,000 MWCO), diluted to 2 ml with Tris·HCl (50 mM, pH 7.0) and concentrated to 0.1 ml. This step was repeated four times after which the modified protein was concentrated to 0.25 ml and stored at 4°C. Yield: 100%.

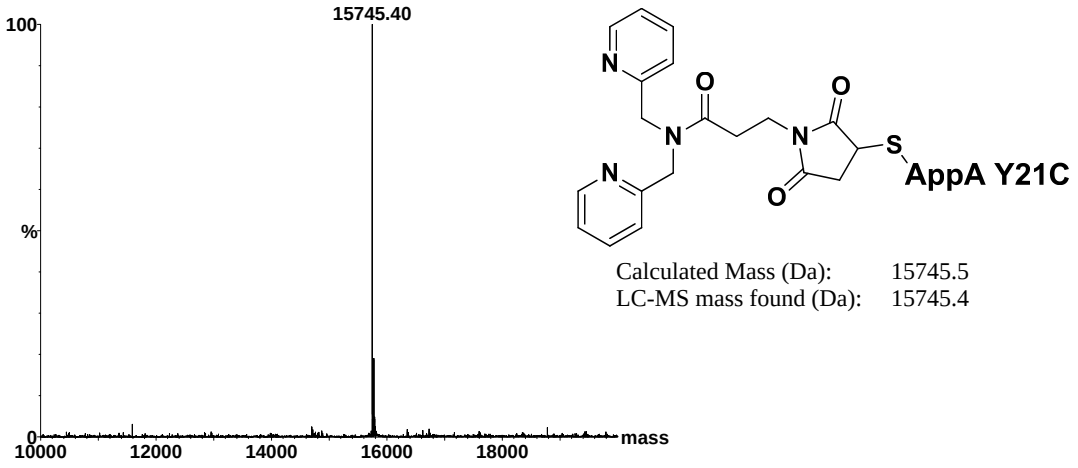


Figure 24 LC-MS (ES+) spectrum of AppA Y21C - **5**

AppA Q63C

Coupling of **5** to AppA Q63C was performed as described for coupling of **5** to AppA Y21C. Yield: 100%.

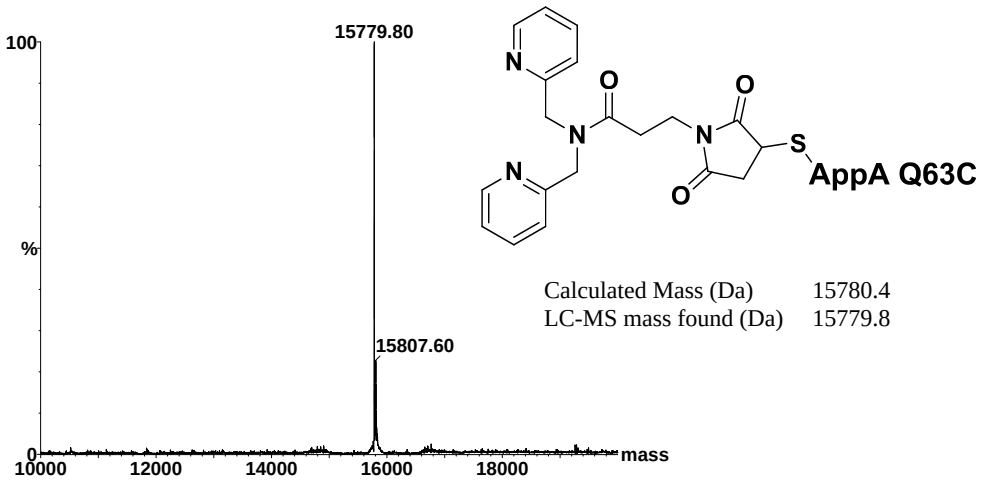


Figure 25 LC-MS (ES+) spectrum of AppA Q63C - **5**

Chemical modification with Ellman's reagent (DTNB)

ALBP

500 nmol (10 eq.) of 5,5'-dithiobis-(2-nitrobenzoic acid) in 50 µl of Tris·HCl (50 mM, pH 8) was added to a solution of 50 nmol of ALBP (1 eq.) in Tris·HCl (50 mM, 1ml, pH 8.0), followed by stirring at room temperature for 5 minutes. Modified protein was analyzed directly by LC-MS (ES+). Only the mass of protein modified with 5-thio-(2-nitro benzoic acid) was detected.

Δ98Δ

1.1 µmol (22 eq.) of 5,5'-dithiobis-(2-nitrobenzoic acid) in 0.11 ml of Tris·HCl (50 mM, pH 8) was added to a degassed solution of 50 nmol of reduced Δ98Δ (1 eq.) in Tris·HCl (50 mM, 1 ml, pH 8.0) containing DTT (6 eq.), in a Schlenk-tube. The reaction mixture was stirred at room temperature for 5 minutes, under inert atmosphere. Modified protein was analyzed directly by LC-MS (ES+). Only the mass of protein modified with 5-thio-(2-nitro benzoic acid) was detected.

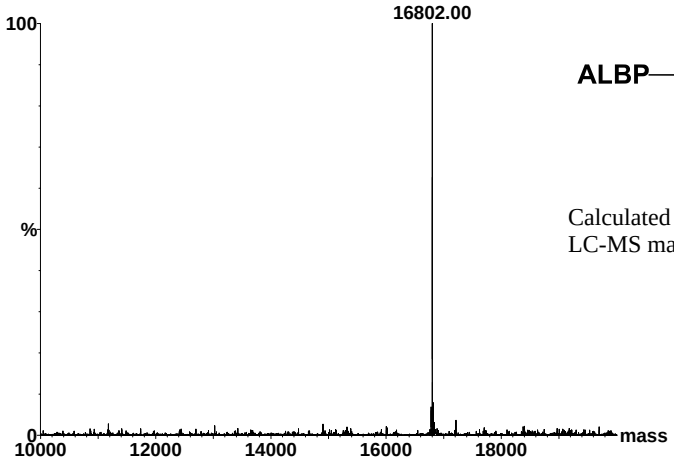
AppA Y21C

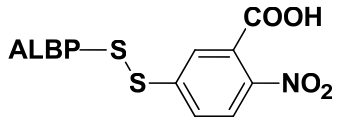
500 nmol (10 eq.) of 5,5'-dithiobis-(2-nitrobenzoic acid) in 50 µl of Tris·HCl (50 mM, pH 7) was added to a solution of 50 nmol of AppA Y21C (1 eq.) in buffer (Tris·HCl 50 mM, 1 ml) containing 3M imidazole, at pH 7.0, followed by stirring at room

temperature for 5 minutes. Modified protein was analyzed directly by LC-MS (ES+). Only the mass of protein modified with 5-thio-(2-nitro benzoic acid) was detected.

AppA Q63C

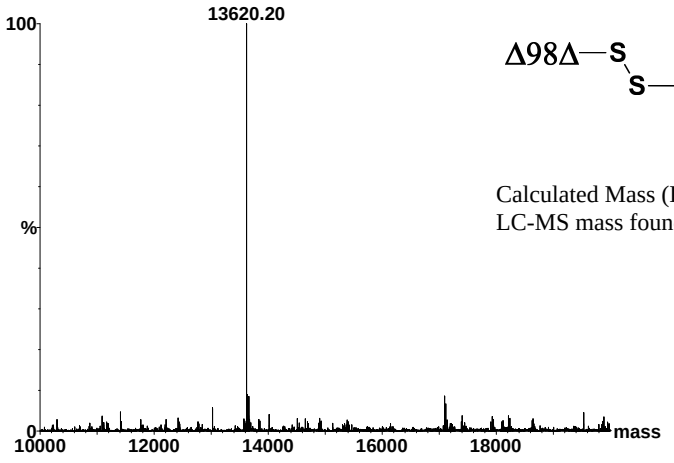
Coupling of 5,5'-dithiobis-(2-nitrobenzoic acid) to AppA Q63C was performed as described for the chemical modification of AppA Y21C with 5,5'-dithiobis-(2-nitrobenzoic acid). Only the mass of protein modified with 5-thio-(2-nitro benzoic acid) was detected.





Calculated Mass (Da)	16802.0
LC-MS mass found (Da)	16802.0

(a)



Calculated Mass (Da)	13620.2
LC-MS mass found (Da)	13620.2

(b)

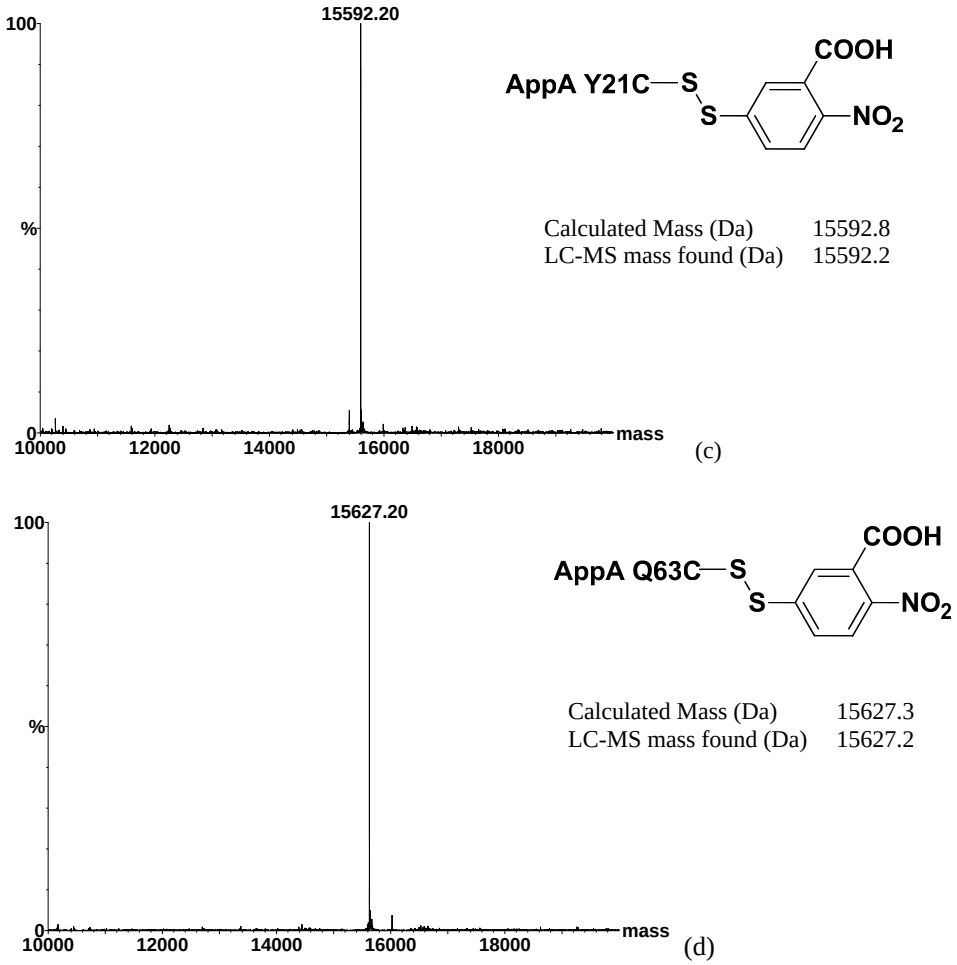
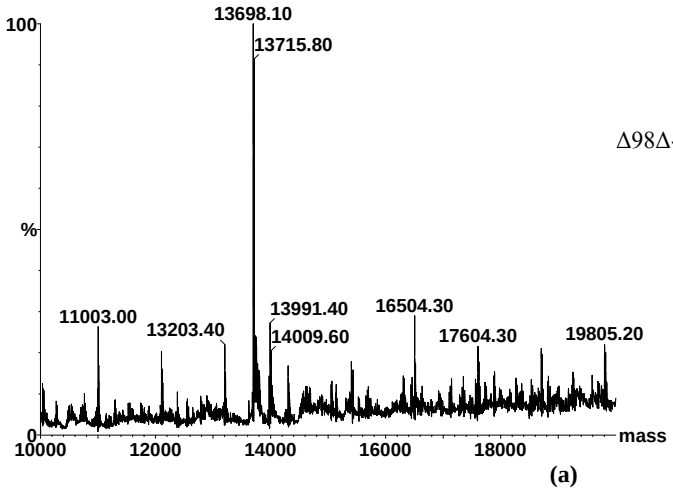


Figure 26 LC-MS (ES+) spectrum of (a) ALBP - **TNB**, (b) Δ98Δ - **TNB**, (c) AppA Y21C – **TNB** and (d) AppA Q63C - **TNB**

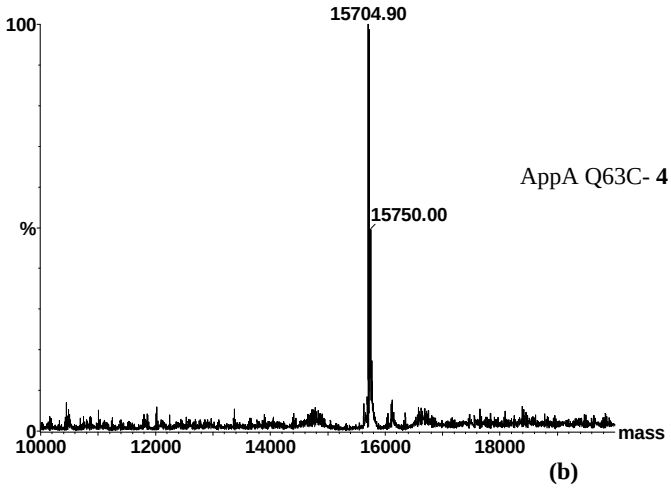
Control reaction between 4-modified proteins and DTNB

The control reaction between Ellman's reagent and **4-**modified proteins was performed as described for chemical modification of apo-proteins with **DTNB**. Only one protein species was detected by LC-MS ES(+), corresponding to the **4-**modified proteins. This result indicates that all cysteines of the apo-proteins are modified with the maleimide **4**.



(a)

	Calculated Mass (Da)	LC-MS mass found (Da)
Δ98Δ- **4**	13698.3	13698.1 13715.8



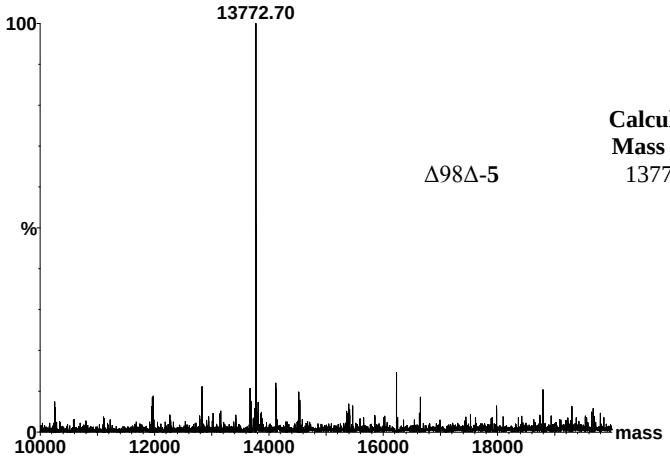
(b)

	Calculated Mass (Da)	LC-MS mass found (Da)
AppA Q63C- **4**	15705.4	15704.9

Figure 27 LC-MS (ES+) spectrum of (a) Δ98Δ-**4** and (b) AppA Q63C-**4** after the reaction with Ellman's reagent

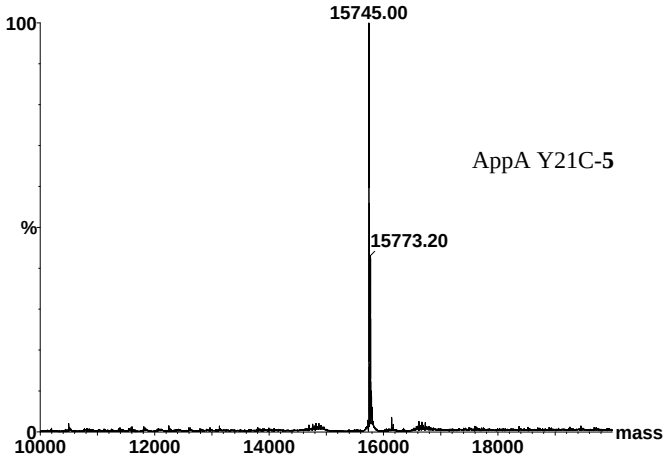
Control reaction between 5-modified proteins and DTNB

The control reaction between **DTNB** and **5**-modified proteins was performed as described for chemical modification of apo-proteins with **DTNB**. Only one protein species was detected by LC-MS ES(+), corresponding to the **5-**modified proteins. This result indicates that all cysteines of the apo-proteins are modified with the maleimide **5**.



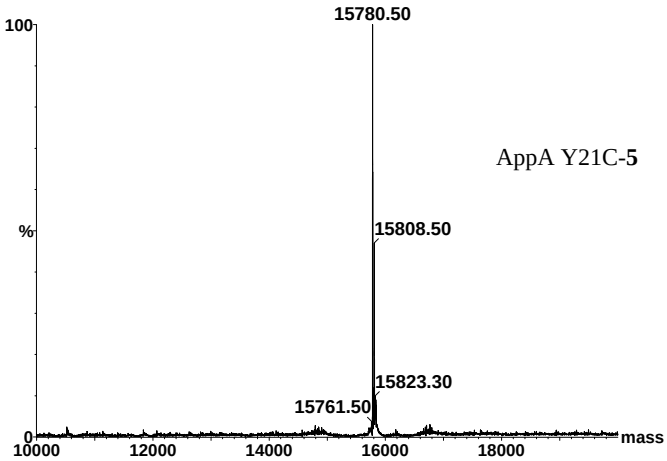
	Calculated Mass (Da)	LC-MS mass found (Da)
Δ98Δ-**5**	13773.4	13772.7

(a)



	Calculated Mass (Da)	LC-MS mass found (Da)
AppA Y21C-**5**	15745.5	15745.0

(b)



	Calculated Mass (Da)	LC-MS mass found (Da)
AppA Y21C-**5**	15780.4	15780.5

(c)

Figure 28 LC-MS (ES+) spectrum of (a) Δ98Δ-**5,** (b) AppA Y21C-**5** and (c) AppA Q63C –**5** after the reaction with Ellman reagent

REFERENCES

1. Fonseca, M. H.; Konig, B. *Adv Synth Catal* **2003**; *345*; 1173.
2. Fache, F.; Schulz, E.; Tommasino, M. L.; Lemaire, M. *Chem. Rev.* **2000**; *100*; 2159.
3. Green, M. L. H. *J. Organomet. Chem.* **1995**; *500*; 127.
4. Pearson, R. G. *J. Am. Chem. Soc.* **1963**; *85*; 3533.
5. Togni, A.; Venanzi, L. M. *Angew. Chem. Int. Ed. Engl.* **1994**; *33*; 497.
6. Fache, E.; Dunjic, B.; Gamez, P.; Lemaire, M. *Topics in Catalysis* **1997**; *4*; 201.
7. Reedijk, J. *Comprehensive Coordination Chemistry* Pergamon Press: Oxford, 1987.
8. Constable, E. C.; Steel, P. J. *Coord. Chem. Rev.* **1989**; *93*; 205.
9. Fletcher, N. C. *J. Chem. Soc., Perkin Trans.* **2002**; 1831.
10. Chelucci, G.; Saba, A.; Sanna, G.; Soccolini, F. *Tetrahedron-Asymmetry* **2000**; *11*; 3427.
11. Schoffers, E. *Eur. J. Org. Chem.* **2003**; 1145.
12. Kiyomori, A.; Marcoux, J. F.; Buchwald, S. L. *Tetrahedron Lett.* **1999**; *40*; 2657.
13. Goodbrand, H. B.; Hu, N. X. *J. Org. Chem.* **1999**; *64*; 670.
14. Wolter, M.; Klapars, A.; Buchwald, S. L. *Org. Lett.* **2001**; *3*; 3803.
15. Shen, R. C.; Lin, C. T.; Bowman, E. J.; Bowman, B. J.; Porco, J. A. *Org. Lett.* **2002**; *4*; 3103.
16. Shen, R. C.; Lin, C. T.; Porco, J. A. *J. Am. Chem. Soc.* **2002**; *124*; 5650.
17. Marko, I. E.; Giles, P. R.; Tsukazaki, M.; Chelle-Regnaut, I.; Gautier, A.; Brown, S. M.; Urch, C. J. *J. Org. Chem.* **1999**; *64*; 2433.
18. Soai, K.; Hayasaka, T.; Ugajin, S. *J. Chem. Soc. - Chem. Commun.* **1989**; 516.
19. Tzalis, D.; Knochel, P. *Tetrahedron Lett.* **1999**; *40*; 3685.
20. Loffler, F.; Hagen, M.; Luning, U. *Synlett* **1999**; 1826.
21. PenaCabrera, E.; Norrby, P. O.; Sjogren, M.; Vitagliano, A.; DeFelice, V.; Oslob, J.; Ishii, S.; ONeill, D.; Akermark, B.; Helquist, P. *J. Am. Chem. Soc.* **1996**; *118*; 4299.
22. Chelucci, G.; Pinna, G. A.; Saba, A.; Sanna, G. *J. Molec. Catal. a-Chemical* **2000**; *159*; 423.
23. Chelucci, G.; Saba, A. *Tetrahedron-Asymmetry* **1998**; *9*; 2575.
24. Puglisi, A.; Benaglia, M.; Annuniziata, R.; Bologna, A. *Tetrahedron Lett.* **2003**; *44*; 2947.
25. Chelucci, G.; Gladiali, S.; Sanna, M. G.; Brunner, H. *Tetrahedron-Asymmetry* **2000**; *11*; 3419.
26. Chelucci, G.; Cabras, M. A.; Saba, A. *J. Molec. Catal. a-Chemical* **1995**; *95*; L7.
27. Gladiali, S.; Pinna, L.; Delogu, G.; Graf, E.; Brunner, H. *Tetrahedron-Asymmetry* **1990**; *1*; 937.
28. Gladiali, S.; Pinna, L.; Delogu, G.; Demartin, S.; Zassinovich, G.; Mestroni, G. *Tetrahedron-Asymmetry* **1990**; *1*; 635.
29. Chelucci, G.; Loriga, G.; Murineddu, G.; Pinna, G. A. *Tetrahedron Lett.* **2002**; *43*; 3601.
30. Nandakumar, M. V.; Ghosh, S.; Schneider, C. *Eur. J. Org. Chem.* **2009**; 6393.
31. Ogo, S.; Abura, T.; Watanabe, Y. *Organometallics* **2002**; *21*; 2964.
32. Abura, T.; Ogo, S.; Watanabe, Y.; Fukuzumi, S. *J. Am. Chem. Soc.* **2003**; *125*; 4149.
33. Canivet, J.; Karmazin-Brelot, L.; Suss-Fink, G. *J. Organomet. Chem.* **2005**; *690*; 3202.
34. Stepnicka, P.; Ludvik, J.; Canivet, J.; Suss-Fink, G. *Inorg Chim Acta* **2006**; *359*; 2369.
35. Letondor, C.; Université de Neuchâtel; 2006.

36. Roelfes, G.; Feringa, B. L. *Angew. Chem. Int. Ed. Engl.* **2005**; *44*; 3230.
37. Roelfes, G.; Boersma, A. J.; Feringa, B. L. *Chem. Commun.* **2006**; 635.
38. Coquiere, D.; Feringa, B. L.; Roelfes, G. *Angew. Chem. Int. Ed. Engl.* **2007**; *46*; 9308.
39. Boersma, A. J.; Klijn, J. E.; Feringa, B. L.; Roelfes, G. *J. Am. Chem. Soc.* **2008**; *130*; 11783.
40. Oltra, N. S.; Roelfes, G. *Chem. Commun. (Camb)* **2008**; 6039.
41. Rosati, F.; Boersma, A. J.; Klijn, J. E.; Meetsma, A.; Feringa, B. L.; Roelfes, G. *Chemistry* **2009**; *15*; 9596.
42. Boersma, A. J.; Feringa, B. L.; Roelfes, G. *Angew. Chem. Int. Ed. Engl.* **2009**; *48*; 3346.
43. Shibata, N.; Yasui, H.; Nakamura, S.; Toru, T. *Synlett* **2007**; 1153.
44. Dijk, E. W.; Feringa, B. L.; Roelfes, G. *Tetrahedron-Asymmetry* **2008**; *19*; 2374.
45. Letondor, C.; Pordea, A.; Humbert, N.; Ivanova, A.; Mazurek, S.; Novic, M.; Ward, T. R. *J. Am. Chem. Soc.* **2006**; *128*; 8320.
46. Letondor, C.; Humbert, N.; Ward, T. R. *Proc. Nat. Acad. Sci. USA* **2005**; *102*; 4683.
47. Creus, M.; Pordea, A.; Rossel, T.; Sardo, A.; Letondor, C.; Ivanova, A.; Le Trong, I.; Stenkamp, R. E.; Ward, T. R. *Angew. Chem. Int. Ed. Engl.* **2008**; *47*; 1400.
48. Thomas, C. M.; Letondor, C.; Humbert, N.; Ward, T. R. *J. Organomet. Chem.* **2005**; *690*; 4488.
49. Pordea, A.; Mathis, D.; Ward, T. R. *J. Organomet. Chem.* **2009**; *694*; 930.
50. Ohashi, M.; Koshiyama, T.; Ueno, T.; Yanase, M.; Fujii, H.; Watanabe, Y. *Angew. Chem. Int. Ed. Engl.* **2003**; *42*; 1005.
51. Rousselot-Pailley, P.; Bochot, C.; Marchi-Delapierre, C.; Jorge-Robin, A.; Martin, L.; Fontecilla-Camps, J. C.; Cavazza, C.; Menage, S. *ChemBioChem* **2009**; *10*; 545.
52. Nimri, S.; Keinan, E. *J. Am. Chem. Soc.* **1999**; *121*; 8978.
53. Ricoux, R.; Sauriat-Dorizon, H.; Girgenti, E.; Blanchard, D.; Mahy, J. P. *J. Immunol. Methods* **2002**; *269*; 39.
54. Ricoux, R.; Lukowska, E.; Pezzotti, F.; Mahy, J. P. *Eur. J. Biochem.* **2004**; *271*; 1277.
55. Mahammed, A.; Gross, Z. *J. Am. Chem. Soc.* **2005**; *127*; 2883.
56. Reetz, M. T.; Jiao, N. *Angew. Chem. Int. Ed. Engl.* **2006**; *45*; 2416.
57. Ueno, T.; Koshiyama, T.; Abe, S.; Yokoi, N.; Ohashi, M.; Nakajima, H.; Watanabe, Y. *J. Organomet. Chem.* **2007**; *692*; 142.
58. Ueno, T.; Koshiyama, T.; Ohashi, M.; Kondo, K.; Kono, M.; Suzuki, A.; Yamane, T.; Watanabe, Y. *J. Am. Chem. Soc.* **2005**; *127*; 6556.
59. Carey, J. R.; Ma, S. K.; Pfister, T. D.; Garner, D. K.; Kim, H. K.; Abramite, J. A.; Wang, Z. L.; Guo, Z. J.; Lu, Y. *J. Am. Chem. Soc.* **2004**; *126*; 10812.
60. Pletsch, A.; Ruhr-Universität Bochum; 2004.
61. Davies, R. R.; Distefano, M. D. *J. Am. Chem. Soc.* **1997**; *119*; 11643.
62. Davies, R. R.; Kuang, H.; Qi, D. F.; Mazhary, A.; Mayaan, E.; Distefano, M. D. *Bioorg. Med. Chem. Lett.* **1999**; *9*; 79.
63. Nicholas, K. M.; Wentworth, P.; Harwig, C. W.; Wentworth, A. D.; Shafton, A.; Janda, K. D. *Proc. Nat. Acad. Sci. USA* **2002**; *99*; 2648.

64. Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Maywald, M.; Maiwald, P.; Peyralans, J. J. P.; Maichele, A.; Fu, Y.; Jiao, N.; Hollmann, F.; Mondiere, R.; Taglieber, A. *Tetrahedron* **2007**; *63*; 6404.
65. Taglieber, A.; Ruhr-Universität Bochum; 2007.
66. Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Taglieber, A.; Hollmann, F.; Mondiere, R. J. G.; Dickmann, N.; Hocker, B.; Cerrone, S.; Haeger, M. C.; Sterner, R. *ChemBioChem* **2008**; *9*; 552.
67. Maywald, M.; Ruhr-Universität Bochum; 2005.
68. Letondor, C.; Ward, T. R. *ChemBioChem* **2006**; *7*; 1845.
69. Pordea, A.; Ward, T. R. *Chem. Commun.* **2008**; 4239.
70. Caputo, C. A.; Jones, N. D. *Dalton Trans* **2007**; 4627.
71. Laan, W.; Munoz, B. K.; Den Heeten, R.; Kamer, P. C. *ChemBioChem* **2010**; *in press*.
72. Hefti, M. H.; Vervoort, J.; van Berkel, W. J. *Eur. J. Biochem.* **2003**; *270*; 4227.
73. Romain, C.; Gaillard, S.; Elmkaddem, M. K.; Toupet, L.; Fischmeister, C.; Thomas, C. M.; Renaud, J. *Organometallics* **2010**; *29*; 1992.
74. Hermanson, G. T. *Bioconjugate techniques*; Academic Press: San Diego, 1996.
75. Chalker, J. M.; Wood, C. S.; Davis, B. G. *J. Am. Chem. Soc.* **2009**; *131*; 16346.
76. Riener, C. K.; Kada, G.; Gruber, H. J. *Analytical and Bioanalytical Chemistry* **2002**; *373*; 266.
77. Bradford, M. M. *Anal. Biochem.* **1976**; *72*; 248.
78. Matarese, V.; Bernlohr, D. A. *J. Biol. Chem.* **1988**; *263*; 14544.
79. Curto, L. M.; Caramelo, J. J.; Delfino, J. M. *Biochemistry* **2005**; *44*; 13847.
80. Laan, W.; van der Horst, M. A.; van Stokkum, I. H.; Hellingwerf, K. J. *Photochemistry and Photobiology* **2003**; *78*; 290.
81. Panetta, C. A.; Kumpaty, H. J.; Heimer, N. E.; Leavy, M. C.; Hussey, C. L. *J. Org. Chem.* **1999**; *64*; 1015.
82. Weizman, H.; Libman, J.; Shanzer, A. *J. Am. Chem. Soc.* **1998**; *120*; 2188.
83. den Heeten, R. In *PhD Dissertation: Amino acid modified phosphine ligands for the development of artificial transition metalloenzymes*; University of Amsterdam: Amsterdam, 2009, p 133.
84. Trammell, S. A.; Goldston, H. M.; Tran, P. T.; Tender, L. M.; Conrad, D. W.; Benson, D. E.; Hellinga, H. W. *Bioconjug. Chem.* **2001**; *12*; 643.

IV. Site selective bioconjugation of transition metal complexes to structurally different protein hosts

Abstract

Site-selective covalent incorporation of metal complexes of CDI activated carboxylic acid functionalized phosphane ligands into protein scaffolds can be efficiently accomplished.[1-2] However, for some protein structures, problems related to modified proteins' stability or unspecific multiple modifications were encountered. Highly efficient and selective coupling of phosphane ligands to proteins was achieved via a hydrazone linkage between an aldehyde functionalized phosphane and 3-(maleimido)propionic acid hydrazide modified proteins. Using this strategy, hybrid transition metal protein adducts can be formed *in situ* from metal precursors and phosphane modified proteins,[3] or by coupling preformed metal complexes of aldehyde functionalized phosphanes to 3-(maleimido)propionic acid hydrazide modified proteins.
Cu and Ru complexes of 5-maleimido-1,10-phenanthroline were incorporated highly efficiently and site selectively into proteins via Michael addition of the thiol group from a cysteine residue to the maleimide double bond. Cu(II) complexes were formed *in situ*, by adding $CuSO_4$ to phenanthroline modified ALBP and Δ98Δ. Preformed Ru complexes of 5-maleimido-1,10-phenanthroline were efficiently and selectively coupled to ALBP, AppA Y21C and AppA Q63C.

IV.1 Introduction

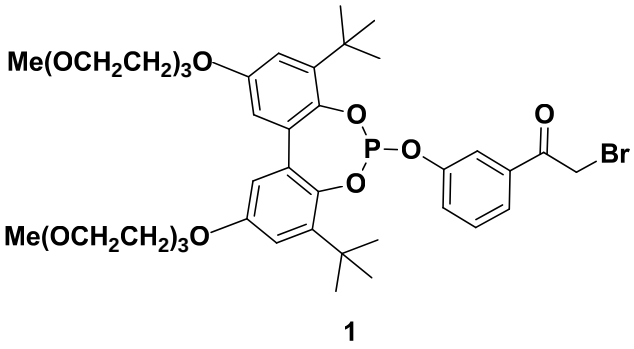
It has been known for decades that artificial metalloenzymes can be synthesised through covalent or non-covalent incorporation of transition metal complexes into protein hosts.[4-5] In this way, the chirality of proteins could be transferred to achiral transition metal catalysts with important applications in organic synthesis.

Covalent incorporation of catalytic functionalities into proteins, first applied by Kaiser and co-workers,[6] and by several other groups after that,[5,7-19] allows a wide range of possibilities for choosing the protein structure or the catalyst. The catalytic moiety inside the protein cavity can be attached at a well-defined position, due to the precise location of a unique cysteine of the protein. This is a particular advantage of the covalent anchoring strategy. Moreover, catalytic systems having different catalytic properties can be created via covalent anchoring, using the same protein scaffold – catalyst couple, simply by varying the location of cysteine inside protein's cavity, as it was proved by Distefano and co-workers.[19] In other words, genetic optimization is an efficient and easy to apply tool for the covalent approach. Also by using the covalent strategy, an organometallic catalyst can be coupled to virtually any protein scaffold displaying a unique, reactive and accessible cysteine.

The non-covalent incorporation of transition metal complexes into proteins has been applied for the synthesis of several artificial metalloenzymes, but this concept has the disadvantage of a limited structural diversity of the protein. So far this concept has been successfully and extensively applied for the biotin-(strept)avidin couple.[4,20-33] It was also applied for serum-albumins,[34-35] displaying high affinity for a variety of hydrophobic compounds, for apomyoglobin[34,36-37] capable of binding symmetric Schiff-base complexes non-covalently and for antibodies[38-42] raised against compounds displaying a similar structure to the specific antigens.

Amongst the common methods reported for chemoselective thiol modification, maleimide and haloacetamide reagents have been used for covalent anchoring of organometallic catalysts into proteins, for the synthesis of artificial metalloenzymes. The haloacetamide derivatives have good chemoselectivity towards cysteine at slightly alkaline pH when low excess of modification reagent relative to the number of sulfhydryl groups present is used.[43] De Vries and co-workers applied this method for selective covalent incorporation of a phosphite ligand into papain via a phenacyl bromide moiety (Figure 1). The resulting artificial metalloenzyme, formed by treating

the modified protein with excess $[Rh(COD)_2]BF_4$, was an active catalyst for the hydrogenation of methyl 2-acetamidoacrylate, but no enantioselectivity was obtained.[14]



1

Figure 1 Monodentate phosphite ligand coupled covalently to papain via a phenacyl bromide moiety[14]

The same method was used by Davies and Distefano for chemoselective covalent incorporation of 5-iodoacetamido-1,10-phenanthroline into two structurally related fatty acid binding proteins (FABPs): ALBP[5] and IFABP[19] (Figure 2a). The Cu(II) complex of the 1,10-phenanthroline modified proteins was subsequently formed. Good enantioselectivities of 31-86% were reported for the hydrolysis of several unactivated amino acid esters (Figure 2b) using the FABP-Cu^{II}-Phen,[5] which were enhanced to 94% e.e. after genetic optimization of the catalyst. Three different catalytic systems were created by placing the phenanthroline catalyst into three different locations inside IFABP cavity. The three catalytic systems displayed different rates and selectivities for the hydrolysis reaction. This proves that the catalytic properties of an artificial metalloenzyme created via covalent anchoring can be modified by changing the attachment site of the metal complex into the protein cavity.[19]

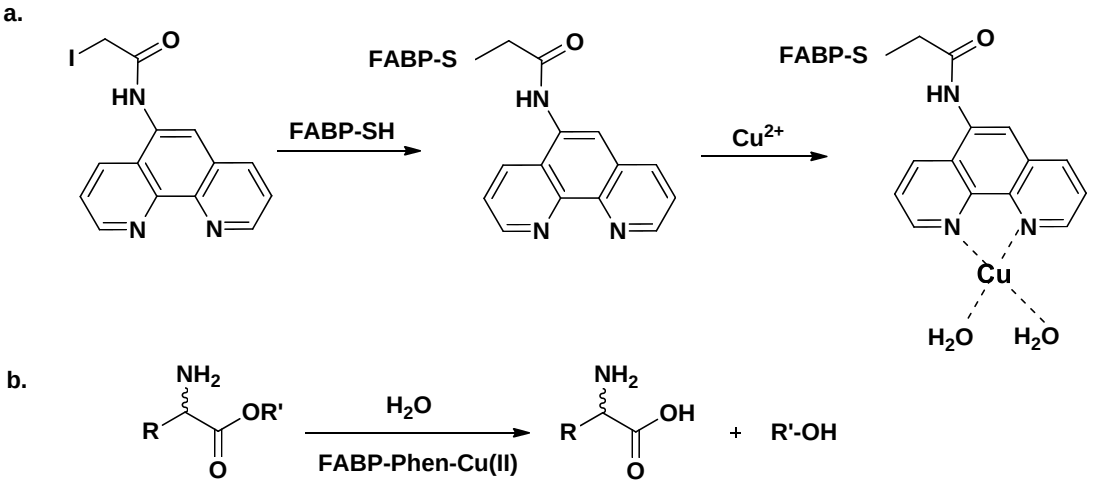


Figure 2 Cu(II)-Phenanthroline conjugates of fatty acid binding proteins (a) used as hybrid catalysts for ester hydrolysis (b)

Distefano's work shows clearly the possibility of synthesizing highly active and selective artificial metalloenzymes by attaching transition metal complexes covalently to proteins. FABPs proved to be good chiral inducers as protein scaffold, probably because of the good encapsulation of the organocatalyst into the protein cavity. This observation was an incentive for us to test ALBP for covalent coupling of transition metal complexes of phosphorus and nitrogen containing ligands, which are very attractive for the synthesis of artificial metalloenzymes, as they are able to catalyse efficiently a variety of important homogeneous catalytic reactions.

The maleimide reagents display high selectivity toward thiol modification at neutral pH. The reaction rates of maleimides with sulfhydryls at pH 7 is reported to be 1000 greater than the reaction with amines.[43] Chemoselective modification of papain and of the protein tHisF-C9AD11C (a thermostable enzyme from *Thermotoga maritime*) via Michael addition to the maleimide double bond of a Pd complex of N-(di(pyridine-2-yl)methyl)maleimide (Figure 3a) was reported by Reetz and co-workers.[16] The protein tHisF-C9AD11C was also used as host for covalent incorporation of a Rh-complex of 5-maleimido-1,10-phenanthroline (Figure 3b) via Michael addition to the maleimide group.[44] The catalytic properties of the resulting artificial metalloenzymes were tested in hydrogenation reactions, but low activities and enantioselectivities were obtained.[16,44]

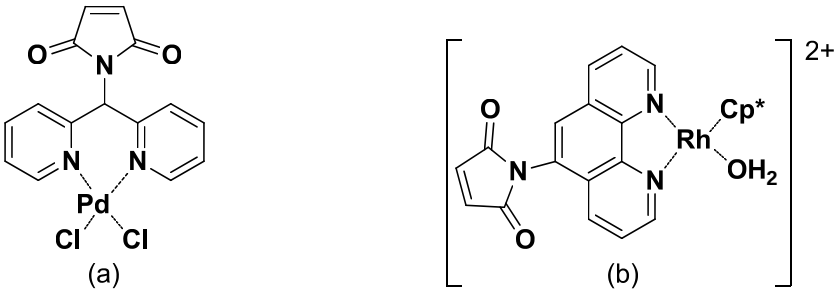


Figure 3 (a) Pd complex of N-(di(pyridine-2-yl)methyl)maleimide coupled via Michael addition to papain and tHisF-C9AD11C;[16] (b) Rh-complex of 5-maleimido-1,10-phenanthroline coupled via Michael addition to tHisF-C9AD11C;[44]

Considering the high affinity of maleimides towards thiol modification, I used this method to explore the efficient, site selective incorporation of metal complexes of nitrogen and phosphane ligands intro three different protein scaffolds: ALBP, Δ98Δ and AppA mutants. The choice of proteins was determined, as mentioned in Chapter II, by their different cavity structure and by a different degree of catalyst encapsulation.

Various routes for covalent incorporation of Pd and Rh metal complexes of phosphane and nitrogen ligands into structurally diverse proteins were explored. Site selective incorporation of metal complexes of phosphane ligands for the efficient synthesis of artificial metalloenzymes has become feasible recently.[1-3] A Rh complex formed through coordinating a phosphane ligand, precoupled to a protein via hydrazone ligation, to a Rh precursor was reported by our group.[3] In addition, I report herein coupling of a preformed Rh–complex of an aldehyde functionalized phosphane to a hydrazide modified protein, via hydrazone ligation. Synthesis of artificial metalloproteins containing metal complexes of 5-maleimido-1,10-phenathroline could be accomplished via Michael addition of cysteine to 5-maleimido-1,10-phenathroline.

IV.2 Results and discussion

IV.2.a Metal complexes of phosphane ligands

The first approach to incorporate phosphane ligands into proteins was by coupling CDI activated phosphane-carboxylic acids to the unique cysteine of a protein. By applying this approach, covalent mono-modification of ALBP with phosphane **4** and covalent mono-modification of Δ98Δ with phosphanes **2-5** (Chapter II) was obtained. These results prompted us to explore subsequent incorporation of transition metal complexes of these phosphanes into the two proteins. Catalysts comprised of phosphane ligands coordinating Rh or Pd have been used for synthetically important transition-metal catalysed transformations like hydrogenation, hydroformylation and allylic substitution. With these types of reactions in mind, I set out to synthesise Rh and Pd based artificial metalloenzymes.

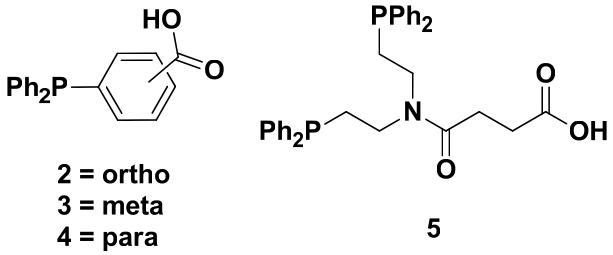
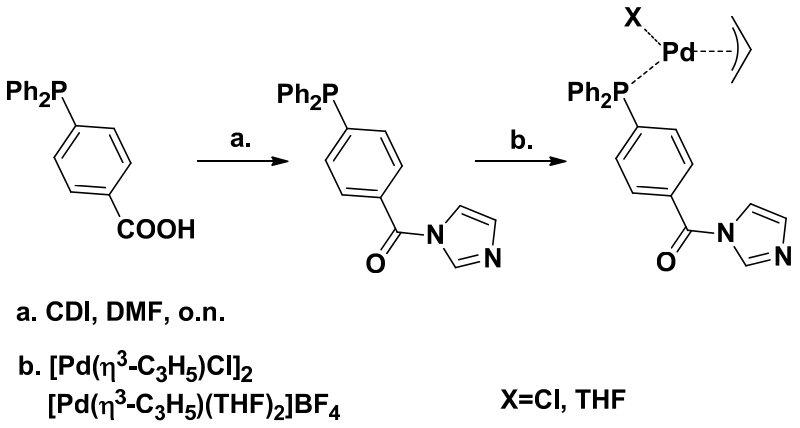


Figure 4 Structures of phosphane-carboxylic acids **2-5**

Two methods were tested in order to achieve selective cysteine modification with transition metal complexes in presence of the multiple functional groups of protein.

I initially explored *in-situ* formation of transition metal complexes of the phosphane-modified proteins, by adding the metal precursor to the proteins modified with phosphane ligands. $[Pd(\eta^3\text{-}C_3H_5)Cl]_2$, $[Pd(\eta^3\text{-}C_3H_5)(THF)_2]BF_4$, and $[Rh(cod)(MeCN)_2]BF_4$ were tested for potential coordination to the phosphane ligand **4** coupled to proteins. However, the modified ALBP and Δ98Δ proved to be unstable in presence of these metal precursors. Formation of a precipitate soluble in 8M GuCl (a common denaturating reagent which facilitates solubilization of precipitated proteins) was observed after the metal addition. In each case, only species corresponding to the unmodified proteins were observed by LC-MS (ES^+). Additives to prevent protein precipitation (glycerol, NaCl, $MgCl_2$), or the use of a co-solvent (acetonitrile) to improve the solubility of modified protein did not solve the problem.

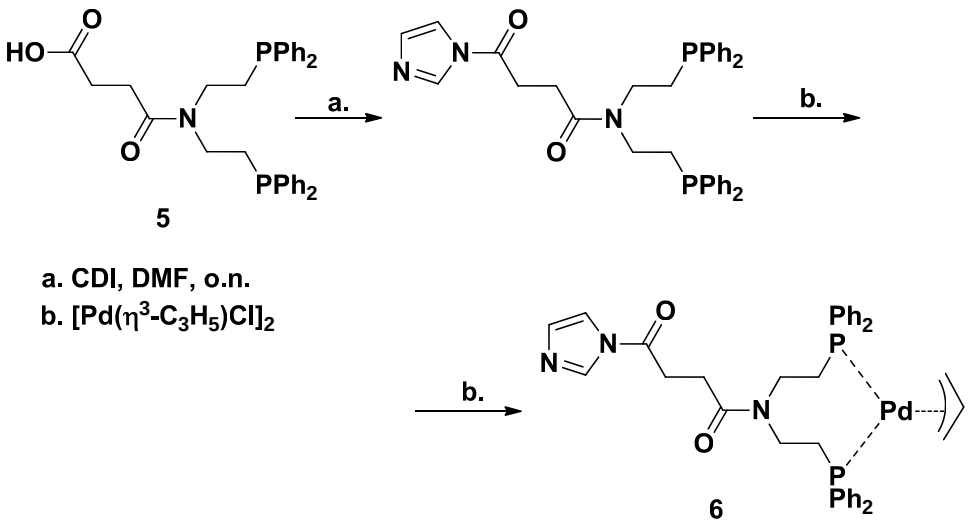
A possible cause of protein precipitation is structural modification and subsequent instability induced by metal coordination to donor sites present in the protein. To avoid this, I tried to couple preformed metal-complexes of ligands **4** and **5** to the unmodified proteins. First, preformed Pd-**4** metal complexes, synthesized from $[Pd(\eta^3\text{-}C_3H_5)Cl]_2$ or $[Pd(\eta^3\text{-}C_3H_5)(THF)_2]BF_4$ and CDI activated ligand **4** (Scheme 1) were coupled to ALBP and Δ98Δ. Once again formation of a yellow precipitate soluble in 8M GuCl was observed during the coupling reaction. Protein precipitation was observed also in presence of the previously mentioned additives (glycerol, NaCl, $MgCl_2$). Moreover, at the end of the coupling reactions of preformed Pd-**4** metal complexes less than 25% of the initial amount of protein was present in solution (Bradford assay), for both proteins tested. No species having the mass corresponding to the protein modified with phosphane ligand or with the metal complex were observed by LC-MS (ES^+).



Scheme 1 Synthesis of Pd-$(\eta^3\text{-}C_3H_5)$ complexes of *p*-(diphenylphosphane)benzoic acid

Coupling of CDI activated carboxylic acid-phosphanes to proteins in buffered solutions did not lead to protein precipitation for any of the proteins investigated. Therefore the preliminary hypothesis was that the instability of modified proteins in presence of transition metal complexes is caused by possible subtle structural modifications resulting from coupling of these complexes. To determine if the instability of the proteins is caused by the metal precursor or by the metal complex, the stability of the apo-proteins in presence of several metal precursors: $[Pd(\eta^3\text{-}C_3H_5)Cl]_2$, $[Pd(\eta^3\text{-}C_3H_5)(THF)_2]BF_4$, $Pd_2(dba)_3$, Na_2PdCl4 and $[Rh(cod)MeCN]BF_4$ was investigated. Both apo-proteins, ALBP and Δ98Δ, were stable in presence of all the tested metal precursors. In all situations, the small amount of precipitate observed after centrifugation of the apo-proteins and metal complexes mixtures was not soluble in 8M GuCl and the amount of residual proteins in supernatant was higher than 70%. Therefore I concluded that even though covalent coupling of the carboxylic acid-phosphane **4** to ALBP and Δ98Δ does not induce structural changes leading to protein precipitation, coupling of Rh and Pd metal complexes of **4** to the two proteins induces their precipitation.

Chelating bidentate phosphane ligands can form structurally well defined transition metal complexes. In addition, they display better selectivities in several metal catalysed reactions, compared to mono-phosphane ligands. For this reason I synthesized the Pd complex **6** and the Rh complex **7** of the CDI activated bidentate phosphane ligand **5** and I attempted to couple them to Δ98Δ. Coupling of CDI activated ligand **5** to ALBP was not successful, probably because of the large steric requirements of this ligand, but also because of the enclosed cavity of ALBP, partially blocked by the α-helical lid (see Chapter II). Consequently, coupling of these two metal complexes could be obtained for Δ98Δ, but not for ALBP. CDI activated ligand **5** was coupled to Δ98Δ, however incomplete conversion to the monomodified protein with **5** was obtained (see Chapter II).



Scheme 2 Synthesis of Pd-(η^3-C_3H_5) complexes of CDI activated diphosphane ligand **5**

Complex **6** was synthesized by reacting CDI activated ligand **5** with $[Pd(\eta^3\text{-}C_3H_5)Cl]_2$. The attempt to couple the complex **6** to Δ98Δ resulted in addition of the allyl group to the protein. A mass corresponding to the allyl modified protein was observed by LC-MS (ES+), but not other masses corresponding to the apo-protein or of the protein modified with the metal complex were observed (Figure 5).

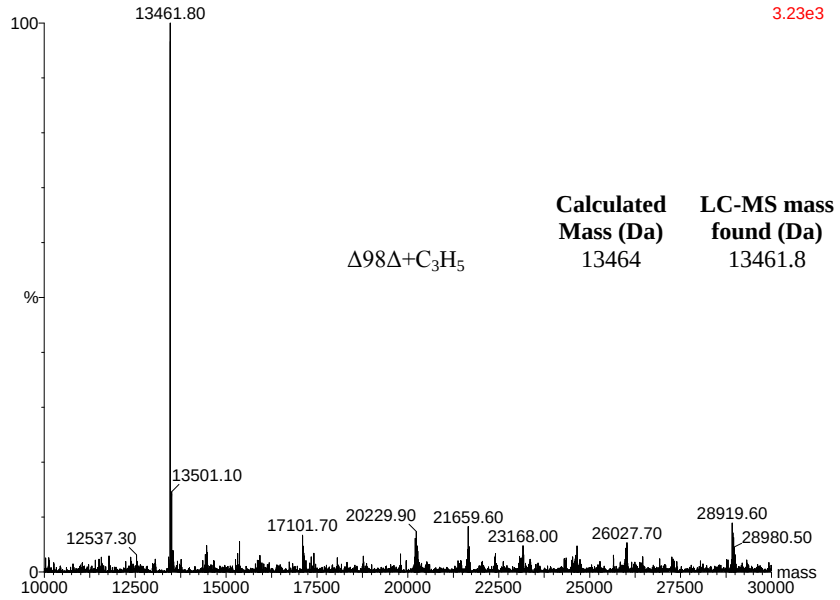
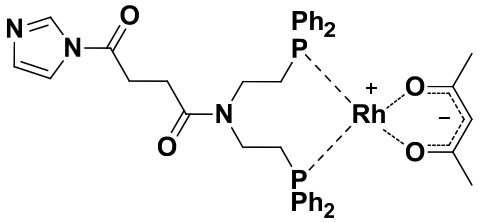


Figure 5 LC-MS (ES+) spectrum of the reaction mixture of Δ98Δ with the complex **6**

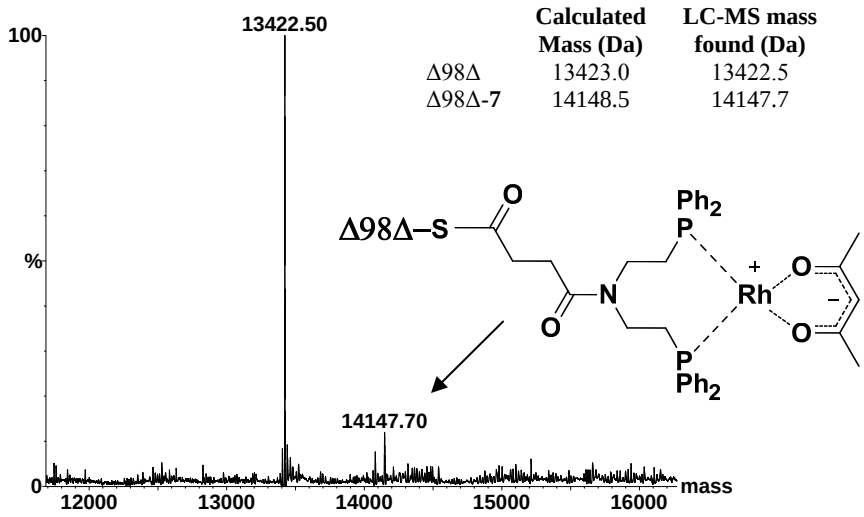
Rh complex **7** was synthesized by reacting CDI activated ligand **5** with $[Rh(acac)(CO)_2]$ (Figure 6) and was coupled to Δ98Δ. Very low coupling efficiency of complex **7** was observed by LC-MS (ES^+) (Figure 7). Considering the incomplete conversion obtained when free ligand **5** was coupled to Δ98Δ, the low coupling efficiency of the complex **7**

could be caused also by the large steric requirements of complex **7.** Alternatively, the instability of the Rh-modified protein could be a cause as the coupling results were not reproducible regarding the coupling efficiency. Less than 50% of the initial amount of Δ98Δ was present in solution after the coupling reaction with complex **7** (Bradford assay), suggesting again the instability of the modified protein in solution.



7

Figure 6 Structure of the Rh(acac) complex of CDI activated ligand **5**



	Calculated Mass (Da)	LC-MS mass found (Da)
Δ98Δ	13423.0	13422.5
Δ98Δ-**7**	14148.5	14147.7

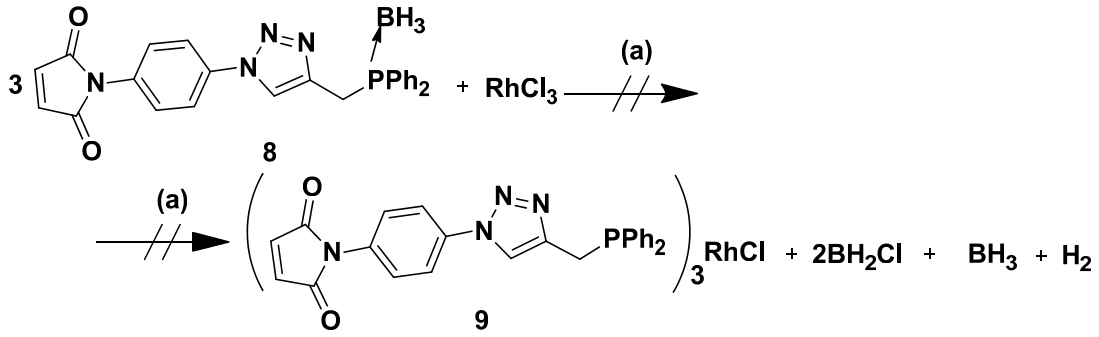
Figure 7 LC-MS (ES+) spectrum of the reaction mixture of Δ98Δ with the complex **7**

Next I considered the synthesis of Pd-complexes of phosphane sulfides. N-(4-(diphenylphosphane sulfide)benzyl) maleimide was successfully coupled to ALBP and Δ98Δ, but attempts to subsequently remove the S protecting group failed (see Chapter II).

Axially chiral P,S and S,S heterodonor ligands have previously been shown to be suitable ligands for asymmetric Pd catalysed allylic substitution reactions.[45-47] Inspired by these studies, I attempted to synthesize *in-situ* the Pd complex of N-(4-(diphenylphosphane sulfide)benzyl) maleimide coupled to Δ98Δ, using $[Pd(\eta^3\text{-}C_3H_5)Cl]_2$ as metal precursor. After treating the phosphane sulfide modified protein with the metal-precursor, no species having the mass of the protein modified with the

metal complex were observed by LC-MS (ES^+). Attempts to determine whether palladium was coordinated to the phosphane sulfide by ^{31}P-NMR required higher concentration of the protein, which led to precipitation of the protein in presence of the metal precursor. Therefore from the preliminary results I could not conclude if coordination of the metal to the N-(4-(diphenylphosphane sulfide)benzyl) maleimide occurred or not.

After incorporation of the BH_3 protected click-phosphanes into ALBP, Δ98Δ and AppA (see Chapter II), the next attempt was to synthesize the corresponding transition metal-hybrids. Previous studies reported synthesis of low oxidation state catalytic systems from phosphane borane complexes. In particular, Rh(I) and Pd(0) phosphane complexes were synthesized from Rh(III) or Pd(II) salts by making use of the reductive properties of the borane protecting group.[48-49] Inspired by these studies, I explored the synthesis of Pd(0) and Rh(I) complexes of click-phosphane **8** using $Pd(OAc)_2$ and $RhCl_3$ as metal precursors (Scheme 3). To avoid reduction of the maleimide double bond by the BH_3 protecting group, the metal complexes synthesis was performed in presence of acrylic acid as BH_3 trapping agent.

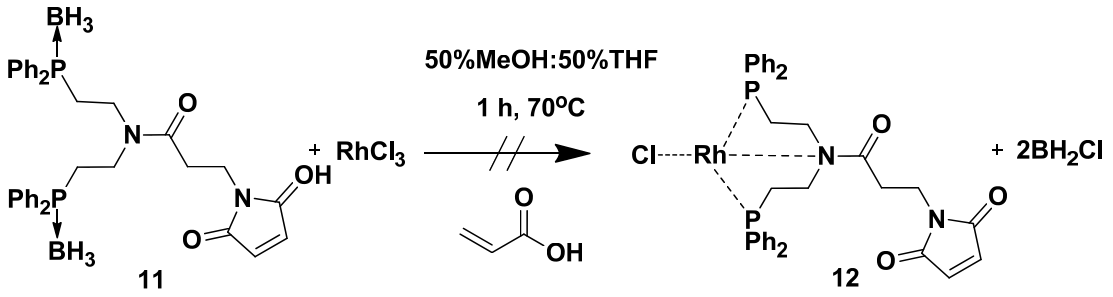


Scheme 3 Attempt of synthesis of the Rh(I) complex of BH_3 protected click phosphane from a Rh(III) precursor. (a) 50%MeOH:50%THF, 1h, 70°C

^{31}P NMR analysis of the crude reactions indicated BH_3 deprotection of the phosphane and possible formation of the Pd or Rh metal complexes, without reaching complete conversion of the BH_3 protected phosphane to the corresponding metal complex. About a third part of the BH_3 protected phosphane remained unreacted, as detected by ^{31}P NMR. The presence of unreacted phosphane suggests a possible coordination of the nitrogen atom of the triazole ring to the metal centre. From the ^{1}H NMR spectra of the crude reactions it was not possible to conclude whether the maleimide double bond was intact.

The potential problem of the nitrogen atom coordinating to the metal centre may be avoided by synthesizing the metal complex of diphenyl(propargyl)phosphane, followed by coupling of the preformed metal complex to the corresponding N-(*p*-azidophenyl)-maleimide via a click-reaction.[50]

Difficulties were encountered in characterization of the Pd and Rh complexes of the click heterodonor ligands, probably due to the presence of several possible isomeric structures, stemming from coordination of the N atom of the triazole ring to the metal centre. Therefore I explored the Rh(I) complex synthesis of the ligand **11**, aiming to obtain a well defined metal complex by coordinating the ligand **11** as a tridentate P-N-P ligand (Scheme 4). Again $RhCl_3$ was used as metal precursor for the synthesis of the metal complex. No evidence was obtained for the presence of an intact maleimide double bond after the attempts to synthesis of the Rh complex **9** and the Pd complex **10**. Therefore, to prevent the possible reduction of the maleimide double bond by the BH_3 group, I used acrylic acid as BH_3 trapping agent. To achieve a possible P,N coordination to Rh, the ligand:metal ratio used was 1:1.



Scheme 4 Attempt of synthesis of the Rh(I) complex of a BH_3 protected diphosphane from a Rh(III) precursor in presence of acrylic acid as BH_3 trapping agent

The preliminary analysis of the crude reaction indicates that BH_3 decomplexation and Rh-complex formation was successful, but the maleimide double bond was reduced. Therefore the resulting metal complex was unsuitable for protein modification. A solution could be the use of more efficient BH_3 trapping agents, e.g. cyclooctadiene.[49]

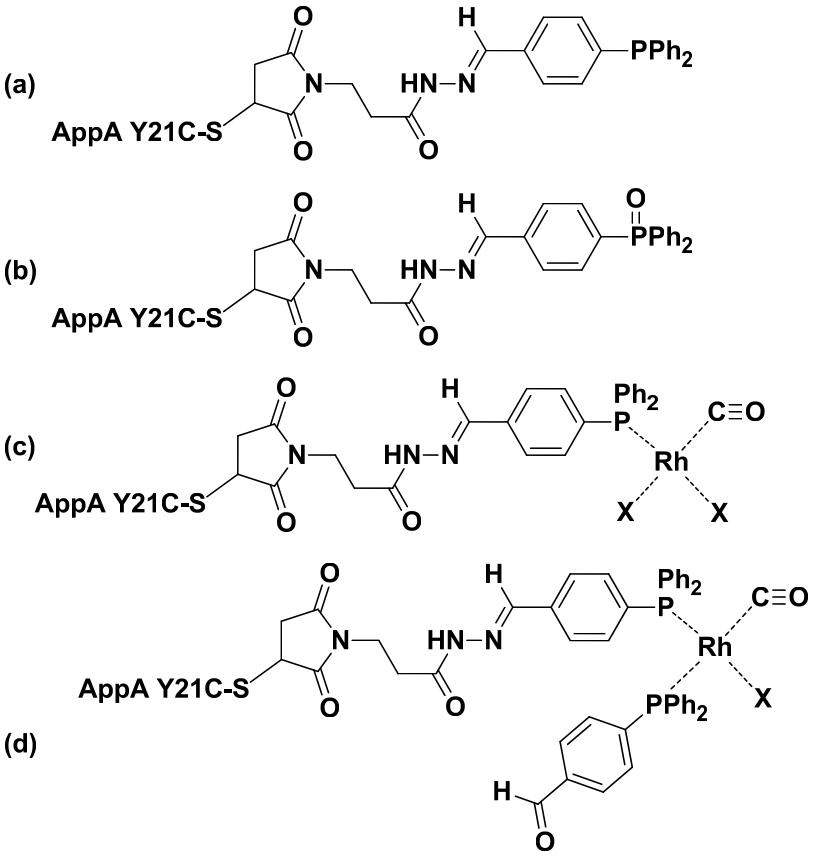
Following the successful incorporation of free phosphane ligands into proteins via hydrazone ligation (see Chapter II), I attempted to coordinate transition metals to the phosphane modified proteins. *In situ* formation of a Rh complex of *p*-(diphenylphosphane)benzaldehyde coupled via hydrazone linkage to the sterol carrier protein 2 like domain of MFE (SCP-2L V83C) was previously reported. The Rh

complex was formed by treating the phosphane modified SCP-2L V83C with one equivalent of $Rh(acac)(CO)_2$, and the metal complex formation was confirmed by mass spectrometry and ^{31}P-NMR.[3]

Preliminary results indicate successful incorporation of a preformed Rh complex of *p*-(diphenylphosphane)benzaldehyde to AppA Y21C. The Rh complex was formed by reacting one equivalent of $Rh(acac)(CO)_2$ with two equivalents of *p*-(diphenylphosphane)benzaldehyde and it was coupled to the AppA Y21C pre-modified with 3-(N-maleimido)propionic acid hydrazide.

Formation of a Rh complex having two monodentate phosphane ligands coordinating to the metal centre was preferred because of a resulting well defined, stable structure of the preformed metal complex.

LC-MS (ES^+) spectrum of the crude reaction showed masses corresponding to the protein modified with the phosphane ligand **(a)**, with the phosphane oxide **(b)**, with the Rh-monophosphane complex **(c)** and in addition, with the Rh-diphosphane complex **(d)** (Figure 8).



		Calculated mass (Da)	LC-MS mass found (Da)
Rh+CO	**(c)-(a)**	130	129.5
p-diphenylphosphino benzaldehyde	**(d)-(c)**	290	289.5
	(a)	15850.5	15848.8
	(b)	16866.5	15864.4
	(c)	15980.4	15978.3
	(d)	16270.3	16267.8

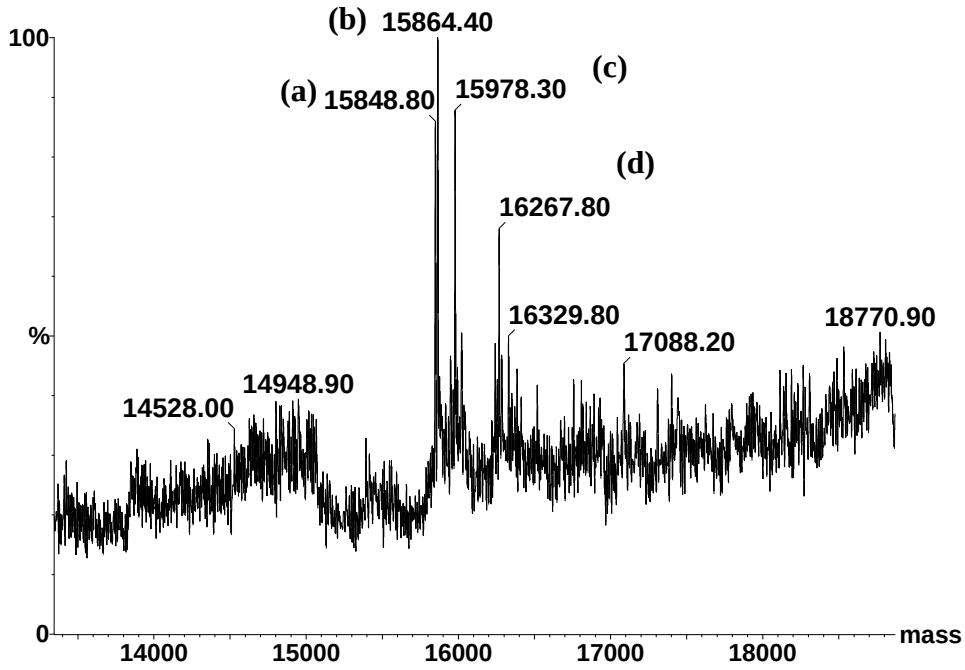


Figure 8 LC-MS (ES^+) of AppA Y21C modified with the Rh complex **19**

The attempt to couple this Rh-diphosphane metal complex to ALBP using the same two step procedure was followed by protein precipitation. LC-MS (ES^+) spectra of the crude coupling reaction revealed only the peak of ALBP modified with the corresponding *p*-(diphenylphosphoryl)benzaldehyde. As observed, different results were obtained for two different protein scaffolds: AppA and ALBP. This emphasizes again the utility of this general method of coupling metal complexes to proteins. Incompatibilities between a particular class of metal complexes and a protein scaffold can be surmounted by using a different type of protein scaffold.

Future work should be done towards structural optimization of this type of Rh metal complexes and subsequent characterization of the resulting Rh-hybrids. However this preliminary result shows clearly the viability of the concept of preformed metal complexes incorporation into protein scaffolds via hydrazone linkage.

IV.2.b Metal complexes of N-containing ligands

Coupling of a metal complex of 2,2'-bipyridine-5-carboxylic acid via thioester bond formation was not successful. No protein modification was obtained when I tried to couple the $[\eta^6$-(*p*-cymene)Ru(N∩N)$(OH_2)](OTf)_2$ complex, having as (N∩N) bidentate chelating ligand the NHS activated 2,2'-bipyridine-5 carboxylic acid, to ALBP. Efficient and selective incorporation of two maleimide functionalized N-ligands into three structurally different protein hosts was successful (see Chapter III). Subsequently, I explored the synthesis of artificial metalloproteins containing Cu(II) and Ru complexes of these N-ligands.

Selective incorporation of the Cu(II) complex of 1,10 phenanthroline into ALBP (lacking the His-tag sequence) was previously reported by Distefano and co-workers. Their results indicated ALBP as a robust scaffold for incorporation of this particular metal complex and as a good chiral inducer.[5] More recently, DNA hybrid catalysts containing Cu(II) complexes of 1,10-phenanthroline were used for Diels-Alder[51] and Michael reactions[52] in water. The cationic arene Ru complexes containing chelating 1,10-phenanthrolines were active catalysts for transfer hydrogenation reactions.[53] Based on that, I explored the synthesis of artificial metalloproteins containing Cu(II) and cationic *p*-cymene Ru complexes of 1,10-phenanthroline.

In addition, I studied the possible influence of the His-tag sequence on ALBP and Δ98Δ stability by comparing my results with those reported by Distefano for ALBP and IFABP lacking the His-tag sequence. I investigated the Cu(II) complex formation of 5-maleimido-1,10-phenantroline precoupled to ALBP and Δ98Δ. After selective covalent incorporation of 5-maleimido-1,10-phenanthroline into ALBP and Δ98Δ (see Chapter III), the corresponding Cu(II) complexes were synthesized following the protocol reported by Distefano and co-workers, by adding Cu(II) to a solution of modified protein. Metal complex formation was confirmed by fluorescence spectroscopy and the results obtained were in concordance with the results reported by Distefano and co-workers for ALBP and IFABP mutants.[5,19]

The fluorescence emission spectrum of unmodified ALBP has an emission maximum at 340 nm, attributed to tyrosine and tryptophan residues. Compared to that, the spectrum of ALBP modified with 5-maleimido-1,10-phenanthroline has the emission maximum at 340 nm substantially reduced, likely due to quenching by phenanthroline, and has also an additional emission maximum at around 420 nm attributed to the phenanthroline fluorophore. Both the phenanthroline fluorescence as well as the protein fluorescence

were quenched after addition of Cu(II) to the modified protein solution. Addition of EDTA to the solution of protein modified with the Cu(II) complex resulted in Cu(II) removal by EDTA and partial returning of phenanthroline fluorescence (Figure 9).

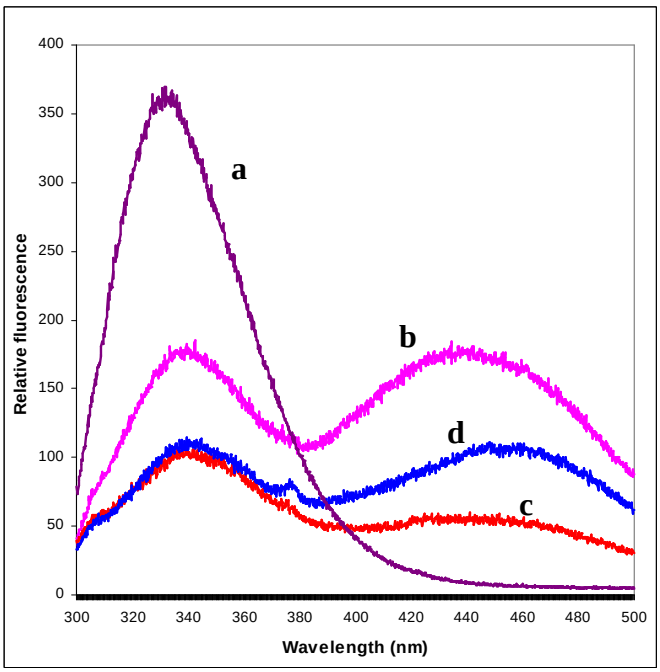


Figure 9 Fluorescence emission spectra of: (a) ALBP, (b) ALBP modified with 5-maleimido-1,10-phenanthroline, (c) ALBP modified with the Cu(II) complex of 5-maleimido-1,10-phenanthroline (d) ALBP modified with the Cu(II) complex of 5-maleimido-1,10-phenanthroline in presence of excess EDTA.

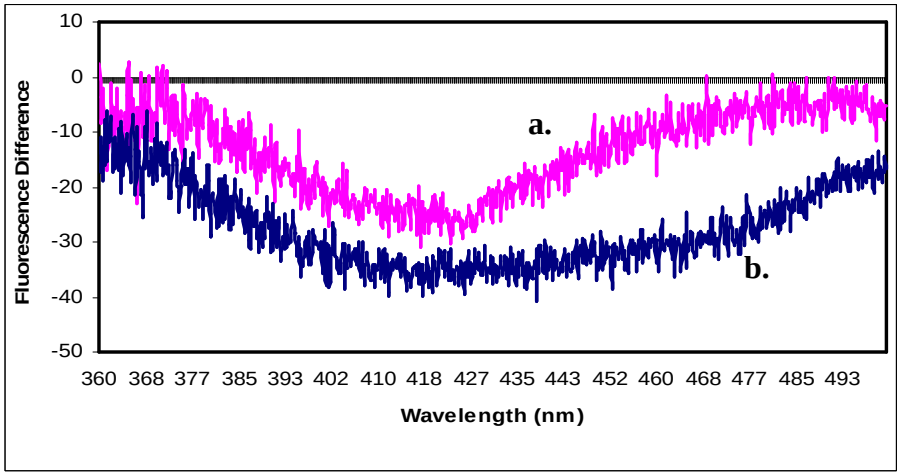


Figure 10 Difference fluorescence spectrum obtained by subtracting: (a) the fluorescence spectra of Δ98Δ from the spectra of Δ98Δ modified with 5-maleimido-1,10-phenanthroline; (b) the fluorescence spectra of Δ98Δ modified with Cu(II)-5-maleimido-1,10-phenanthroline from the spectra of Δ98Δ modified with Cu(II)-5-maleimido-1,10-phenanthroline in presence of excess EDTA.

Perturbation of the phenanthroline fluorophore coupled to Δ98Δ in presence of $CuSO_4$ was proven by comparing two fluorescence difference spectra. One spectrum was

obtained by subtracting the spectrum of Δ98Δ from the spectrum of Δ98Δ modified with 5-maleimido-1,10-phenanthroline and the second one by subtracting the spectrum of Δ98Δ modified with the Cu(II) complex of 5-maleimido-1,10-phenanthroline from the spectrum of Δ98Δ modified with the same Cu(II) complex in presence of excess EDTA. A similar shape was observed in both cases, in agreement with the results reported by Distefano and co-workers.[19] When the fluorescence properties of three IFABP mutants modified covalently with 1,10-phenanthroline were explored, the results indicated the IFABP mutants as less amenable for this sort of analysis. Only one modified mutant, IFABP L72C, exhibited significant phenanthroline fluorescence.[19] Δ98Δ is a genetically engineered variant of IFABP, which lacks the N-terminal 1-28 amino acids and the last five amino acids belonging to the C-terminus.[54] Belonging to the IFABP class, it is expected that the differences in the fluorescence spectra of apo and modified-Δ98Δ to be less evident regarding the phenanthroline fluorescence.

Although the His-tag sequence was present for both proteins, ALBP and Δ98Δ, the formation of the metal complex did not induce protein instability. These results show that the stability of an artificial metalloenzyme construct depends on the protein scaffold as well as on the metal and the ligand type. Several subtle interactions between metal complex and protein could lead to an unstable construct.

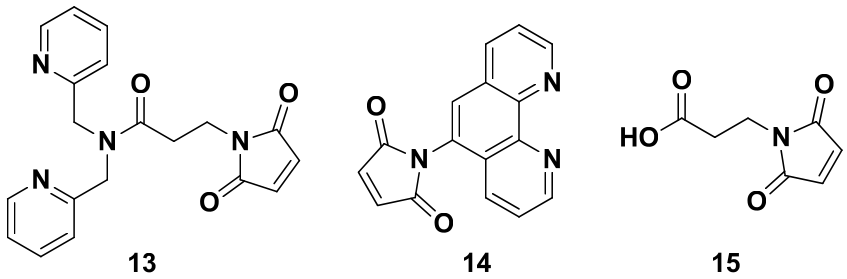


Figure 11 Structures of 3-(N-maleimido)-(N,N-bis(2-picolyl))-propanamide **(13)**, 5-maleimido-1,10-phenanthroline **(14)** and 3-(N-maleimido)-propionic acid **(15)**

After successful incorporation of the Cu(II)-1,10 phenanthroline into ALBP and Δ98Δ, I tried to incorporate [η^6-(*p*-cymene)Ru(N∩N)Cl]Cl and [η^6-(*p*-cymene)Ru(N∩N)(OH_2)] $(OTf)_2$ complexes, with N∩N being either the bidentate chelating ligand 3-(N-maleimido)-(N,N-bis(2-picolyl))-propanamide) or 5-maleimido-1,10-phenanthroline. This type of complexes has the advantage of being stable, soluble and catalytically active in aqueous conditions.[53]

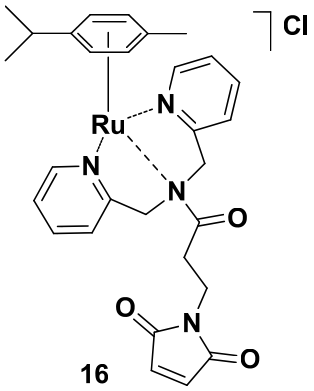
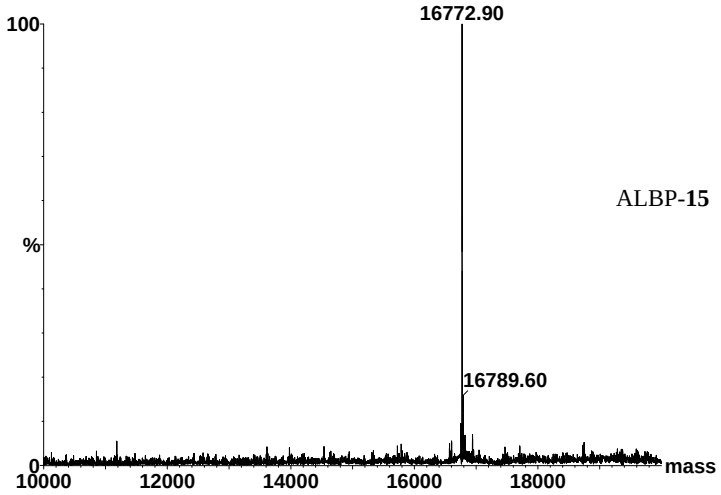


Figure 12 Possible structure of the η^6-(*p*-cymene)Ru complex of **13**

When metal complex **16** was allowed to react with ALBP and AppA Y21C, LC-MS (ES^+) spectra showed only the peaks corresponding to the proteins modified with 3-(N-maleimido)-propionic acid (Figure 18), which is the product of hydrolysis of the amide bond of **13**. The metal complex **16** was formed in dichloromethane and was also detected in the reaction mixture (LC-MS (ES^+)). However it is possible that the amide bond becomes more labile upon coordination of the central N atom to Ru centre and in aqueous conditions **15** is released as a result of hydrolysis.



	Calculated Mass (Da)	LC-MS mass found (Da)
ALBP-**15**	16773.9	16772.9

Figure 13 LC-MS (ES+) spectrum of the reaction mixture of ALBP with **16**

Metal complex **17** (Figure 14) was successfully incorporated into ALBP. However, slow hydrolysis of the coupled complex in the buffered aqueous environment was observed. A secondary mass corresponding to the protein modified with the resulting Ru aqua complex **18** was observed to appear in time (LC-MS (ES+)) (Figure 15). The peak corresponding to the mass of the protein modified with the complex **17** (17145.9)

was decreasing in time, being replaced by the peaks corresponding to the protein modified with complex **18** (17128.7) and the protein modified with the maleamic acid adduct of complex **17** (17166.1). To prove that the complex **18** is formed upon hydrolysis of the complex **17** coupled to ALBP in aqueous conditions, I synthesized and coupled the derived complex **18**, which is stable in water. This metal complex was successfully incorporated into ALBP and into the two AppA mutants: Y21C and Q63C (Table 1).

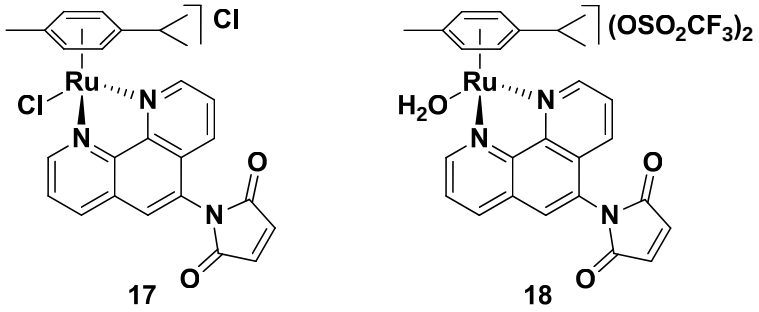
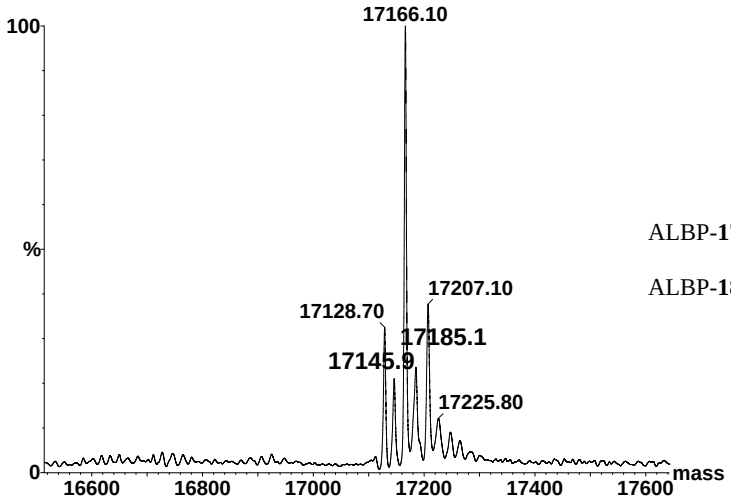


Figure 14 η^6-(*p*-cymene)Ru complexes of 5-maleimido-1,10-phenanthroline



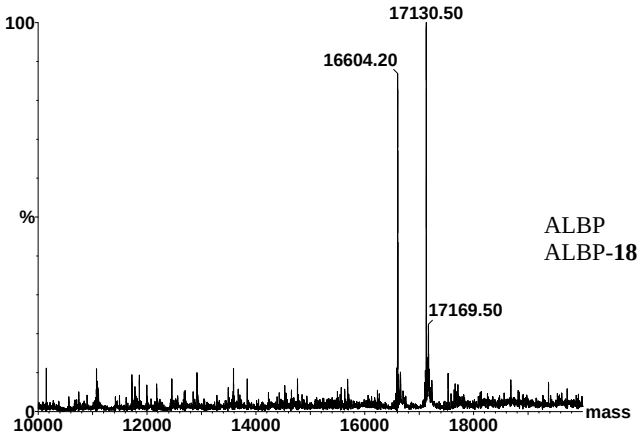
	Calculated Mass (Da)	LC-MS mass found (Da)
ALBP-**17**	17150.8	17145.9 17166.1
ALBP-**18**	17131.3	17128.7

Figure 15 LC- MS (ES+) of ALBP modified with complex **17**

Table 1 ALBP and AppA chemical modification with Ru complex **18**

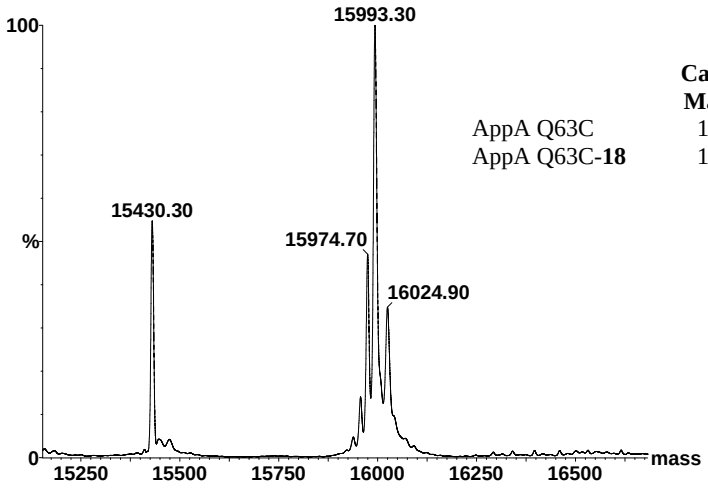
	Calculated Mass (Da)	LC-MS mass found (Da)
ALBP	16604.8	16604.2
ALBP -**18**	17131.3	17130.5
AppA Q63C	15430.1	15430.5
AppA Q63C -**18**	15956.6	15956.9
AppA Y21C	15395.1	15394.9
AppA Y21C -**18**	15921.6	15921.7

Incomplete modification of the three proteins with the Ru-complex **18** was obtained (Figure 16). Conversion was estimated at 62% for ALBP, 44% for AppA Y21C and 32% for AppA Q63C. These results were obtained after correction of the LC-MS data for the differences in ionization efficiency between the modified and unmodified proteins, which was estimated by adding half of equivalent of unmodified protein to the crude reaction mixture.



	Calculated Mass (Da)	LC-MS mass found (Da)
ALBP	16604.8	16604.2
ALBP-**18**	17131.3	17130.5

Figure 16 LC-MS (ES^+) spectrum of ALBP - **18**



	Calculated Mass (Da)	LC-MS mass found (Da)
AppA Q63C	15430.1	15430.3
AppA Q63C-**18**	15956.6	15974.7
		15993.3 (main peak)
		16024.9

Figure 17 LC-MS (ES+) of AppA Q63C modified with **18** in presence of 3M imidazole

The coupling efficiency of complex **18** to the AppA mutants was improved when the coupling was done in presence of 3M imidazole. However, an increase of 36 Da of the mass of proteins modified with the metal complex was observed (Figure 17). This addition could result from the formation of the maleamic acid adduct of the ligand (+18 Da) in presence of high concentration of imidazole and addition of another H_2O molecule to the metal complex (+18 Da). However, the MS resolution is not high enough to clearly distinguish a mass difference of 2 Da for proteins of several kDa,

therefore the addition of a mass of 16 Da caused by oxidation of a methionine residue of the protein should also be considered, instead of addition of water to the metal complex.

IV.3 Conclusions

Artificial metalloenzymes can be synthesized via covalent incorporation of transition metal complexes (Rh, Pd) of phosphane ligands into proteins.[1,3] However, the success of the approach might be influenced by several factors like the metal precursor, the class of phosphane ligand (mono- or diphosphane) and most important, the protein scaffold. Pd and Rh metal complexes of phosphane ligands proved to induce protein instability for the particular proteins tested (ALBP and Δ98Δ). A possible cause could be the presence of the His-tag sequence attached at the N-terminus of ALBP and at the C-terminus of Δ98Δ, displaying high affinity toward metal coordination. Future work on attaching the phosphane metal complexes after the enzymatic cleavage of the His-tag of AppA, ALBP and Δ98Δ might reveal different coupling results.

Herein I report a general method for incorporation of phosphane complexes into proteins, applicable for a wide protein structure and catalyst space. By using this method, the potential incompatibilities between a particular organometallic catalyst and a protein scaffold could be surpassed by varying the protein space. In particular, even if incorporation of Pd and Rh complexes was not compatible with ALBP and Δ98Δ protein structures, the approach has been successful for other protein scaffolds (PYP[1] and MFE SCP-2L V83C[3]).

A careful selection of the metal complexes and of the protein hosts is required. Robust proteins, displaying high stability over a wide range of pH, but also soluble and stable in organic solvents, could be a good choice.

Covalent coupling of Cu and Ru complexes of N-ligands to proteins proved to be less troublesome regarding the stability of modified proteins (for the particular case of the proteins studied in this work). Efficient and selective covalent modification of three protein scaffolds with Cu and Ru complexes of 5-maleimido-1,10phenathroline was obtained. Artificial metalloproteins containing Ru complexes of 1,10-phenanthroline have not been reported before. Both Cu and Ru hybrids are stable in buffered solutions

and are suitable to be tested for reactions catalysed by complexes of 1,10-phenanthroline (Michael additions,[52] Diels-Alder reactions,[55] transfer hydrogenation[56-57]).

IV.4 Experimental

IV.4.1 General methods

Chemicals were purchased from Aldrich, Fluka and Strem Chemicals and were used as received. NMR spectra were recorded at room temperature on Bruker Avance 400 and 300 spectrometers. Positive chemical shifts (δ) are given (in ppm) for high-frequency shifts relative to a TMS reference (^{1}H and ^{13}C). ^{13}C spectra were measured with ^{1}H decoupling. Multiplicities are indicated by: s (singlet), d (doublet), dd (doublet of doublets) and m (multiplet). All reactions involving phosphanes were performed under an argon atmosphere using degassed solvents and standard Schlenk techniques. Fluorescence measurements were performed using a Varian Cary Eclipse Fluorescence Spectrophotometer.

Protein stability (apo and modified proteins) was quantified by Bradford assay and SDS-PAGE.

Proteins chemical modification was assessed by LC-MS (ES^{+}). The protein sample (20 μL, 5 pM/μL) was desalted on-line through a XTerra MS C8 2.1 × 10 mm column, eluted with an increasing acetonitrile concentration (2 % acetonitrile, 98 % aqueous 1 % formic acid to 98 % acetonitrile 2 % aqueous 1 % formic acid and delivered to an electrospray ionisation mass spectrometer (LCT, Micromass, Manchester, U.K.) which had previously been calibrated using myoglobin. An envelope of multiply charged signals was obtained and deconvoluted using MaxEnt1 software to give the molecular mass of the protein.

The sequence of ALBP reported by Bernlohr *et al*[58] was used as reference for the mass of ALBP C1A his-tagged mutant. For the mass of the Δ98Δ mutant, the Δ98Δ sequence reported by Delfino *et al*[54] was used as reference and for the mass of AppA mutants the AppA sequence reported by Laan *et al*[59] was used.

$[Rh(cod)(CH_3CN)_2]BF_4$ – was prepared according to a literature procedure[60] and it was kindly provided by Dr Wouter Laan (University of St Andrews).

$[Pd(\eta^3\text{-}C_3H_5)(THF)_2]BF_4$ – was synthesized according to a literature procedure[61] and it was kindly provided by Dr Wouter Laan (University of St Andrews).

Attempts for *in-situ* formation of the Pd(η^3-C_3H_5)-4-protein complexes

ALBP

45 nmol (0.45 eq) of $[Pd(\eta^3\text{-}C_3H_5)Cl]_2$ in 4.5 µl of dimethylformamide was added dropwise to a degassed solution of 100 nmol of ALBP-**4** (1 eq.) in bicine (50 mM, 5 ml, pH 8.5), in a Schlenk tube. The reaction mixture was stirred at room temperature for 16 hours. The unreacted palladium was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with degassed Tris·HCl (50 mM, pH 8) and concentrated to 0.5 ml. This step was repeated four times after which the protein solution was concentrated to 1 ml and stored at 4°C.

The attempt to form *in-situ* the Pd complex of ALBP-**4** using $[Pd(\eta^3\text{-}C_3H_5)(THF)_2]BF_4$ as metal precursor was performed as described for $[Pd(\eta^3\text{-}C_3H_5)Cl]_2$. 90 nmol (0.9 eq.) of $[Pd(\eta^3\text{-}C_3H_5)(THF)_2]BF_4$ in 1.8 µl of tetrahydrofuran was used for metal complex formation.

Δ98Δ

The attempts to form *in-situ* the Pd(η^3-C_3H_5)-**4**-Δ98Δ complexes starting from $[Pd(\eta^3\text{-}C_3H_5)Cl]_2$ and $[Pd(\eta^3\text{-}C_3H_5)(THF)_2]BF_4$ and Δ98Δ-**4** were performed as described for ALBP-**4**.

Attempts for *in-situ* formation of the Rh-4-protein complexes

ALBP

100 nmol (1 eq.) of $[Rh(cod)(MeCN)_2]BF_4$ in 2 µl of acetonitrile was added dropwise to a degassed solution of 100 nmol of ALBP-**4** (1 eq.) in bicine (50 mM, 5 ml, pH 8.5), in a Schlenk tube. The reaction mixture was stirred for 16 hours at room temperature. The unreacted rhodium was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with degassed Tris·HCl (50 mM, pH 8.0) and concentrated to 0.5 ml. This step was repeated four times after which the protein solution was concentrated to 1 ml and stored at 4°C.

Δ98Δ

The attempt to form the Rh complex of Δ98Δ-**4** starting from $[Rh(cod)(MeCN)_2]BF_4$ was performed as described for ALBP-**4**.

The attempts for *in-situ* formation of the Rh and Pd **4**-protein complexes were also performed in presence of the following common additives used to prevent protein precipitation: 10% glycerol; 80mM $MgCl_2$; 500 mM NaCl; 25% acetonitrile; 10% glycerol and 500mM NaCl; 10% glycerol and 80 mM $MgCl_2$; 10% glycerol, 500mM NaCl and 80mM $MgCl_2$.

CDI activation of ligands 4 and 5 – was performed as described in Chapter II.

Synthesis of activated 4-Pd(η^3-C_3H_5)Cl

50 μmol (2.6 eq.) of CDI activated ligand **4** in 1 ml of dimethylformamide was added to 19 μmol (6.9 mg, 1 eq.) of $[Pd(\eta^3\text{-}C_3H_5)Cl]_2$. The yellow reaction mixture was left at room temperature for 30 min., after which the solution was used for protein modification.

Synthesis of activated 4-Pd(η^3-C_3H_5)(THF)

50 μmol (1.3 eq.) of CDI activated ligand **4** in 0.5 ml of dimethylformamide was added to 38 μmol (1 eq.) of $[Pd(\eta^3\text{-}C_3H_5)(THF)_2]BF_4$ in 0.5 ml of tetrahydrofuran. The yellow reaction mixture was left at room temperature for 30 min., after which the solution was used for protein modification.

Attempts to couple the activated 4-Pd(η^3-C_3H_5)Cl to ALBP

2.5 μmol (25 eq.) of activated [**4**-Pd(η^3-C_3H_5)Cl] was added slowly to a degassed solution of 100 nmol of ALBP (1 eq.) in 5 ml Tris·HCl (50 mM, pH 8), in a Schlenk tube. The reaction mixture was stirred for 16 hours at room temperature. The unreacted palladium complex was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with degassed Tris·HCl (50 mM, pH 8) and concentrated to 0.5 ml. This step was repeated four times after which the protein solution was concentrated to 1 ml and stored at 4°C.

Attempts to couple the activated 4-Pd(η^3-C_3H_5)Cl to Δ98Δ

A solution of 100 nmol of Δ98Δ (1 eq.) in Tris·HCl (50 mM, 5 ml, pH 8) was degassed by purging with argon for 8 hours, in a Schlenk tube. The disulfide bridges were reduced using 6-fold molar excess TCEP (12 μl of stock solution of 50 mM in water). The reduction reaction proceeded for 1 hour at room temperature. While stirring, 2.5 μmol (25 eq.) of activated [**4**-Pd(η^3-C_3H_5)Cl] was added slowly. The reaction mixture

was stirred for 16 hours at room temperature. The unreacted palladium complex and TCEP were removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with degassed Tris·HCl (50 mM, pH 8) and concentrated to 0.5 ml. This step was repeated four times after which the protein solution was concentrated to 1 ml and stored at 4°C.

Attempts to couple the activated 4-Pd(η^3-C_3H_5)THF to ALBP – was performed as described for the activated **4**-Pd(η^3-C_3H_5)Cl.

Attempts to couple the activated 4-Pd(η^3-C_3H_5)THF to Δ98Δ – was performed as described for the activated **4**-Pd(η^3-C_3H_5)Cl.

Synthesis of the complex 6

50 μmol (2.2 eq.) of CDI activated ligand **5** in 1 ml of dimethylformamide was added to 22.5 μmol (8.2 mg, 1 eq.) of [Pd(η^3-C_3H_5)Cl]$_2$. The yellow reaction mixture was left at room temperature for 30 min., after which the solution was used for protein modification.

Attempt to couple the complex 6 to Δ98Δ - was performed as described for the activated [**4**-Pd(η^3-C_3H_5)Cl].

Synthesis of the complex 7

20 μmol (1 eq.) of CDI activated ligand **5** in 0.4 ml of dimethylformamide was added to 18 μmol (4.64 mg, 0.9 eq.) of [Rh(acac)(CO)$_2$]. The yellow reaction mixture was left at room temperature for 30 min., after which the solution was used for Δ98Δ modification.

Coupling of the complex 7 to Δ98Δ - was performed as described for the activated [**4**-Pd(η^3-C_3H_5)Cl].

Attempt to synthesize the Pd complex of N-(4-(diphenylphosphane sulfide)benzyl) maleimide coupled to Δ98Δ

The synthesis of this Pd complex was attempted based on a reported literature procedure.[47]

0.45 μmol (0.16 mg, 0.9 eq.) of [Pd(η^3-C_3H_5)Cl]$_2$ in 9 μl of tetrahydrofuran was added dropwise to a degassed solution of 1 μmol (2 eq.) of Δ98Δ modified with N-(4-(diphenylphosphane sulfide)benzyl) maleimide in Tris·HCl (50 mM, 50 ml, pH 7), in a Schlenk tube. The reaction mixture was stirred for 4 hours at room temperature. The unreacted palladium was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), concentrated to 2 ml, diluted to 10 ml with degassed Tris·HCl (50 mM, pH 7)

and concentrated again to 2 ml. This step was repeated four times after which the protein solution was concentrated to 1 ml and stored at 4°C.

Attempt to synthesize the Rh complex 9

The synthesis of this Rh complex was attempted based on a reported literature procedure.[49]

30 μmol (13.57 mg, 3 eq.) of click-compound **8** and 10 μmol (2.62 mg, 1 eq.) of $RhCl_3$ were dissolved in 5 ml of anhydrous mixture methanol: tetrahydrofuran (1:1). The reaction mixture was heated at reflux for 1 hour. After cooling, the solvent was removed under reduced pressure and the residue was analysed by NMR spectroscopy.

^{31}P {^{1}H} NMR (121 MHz, MeOD): δ=17.1 (br), 38.64 (dd, J=108.12, 16.99), 46.82 (dd, J=110.04, 16.97) ppm

Attempt to synthesize the Rh complex 9 in presence of acrylic acid – was performed in similar conditions as described above. 100 μmol of acrylic acid (7.2 mg, 10 eq.) was added to the reaction mixture before refluxing.

Attempt to synthesize the Pd complex 10

The synthesis of this Pd complex was attempted based on a reported literature procedure.[49]

30 μmol (13.57 mg, 2 eq.) of click-compound **8** and 15 μmol (3.36 mg, 1 eq.) of $Pd(OAc)_2$ were dissolved in 1 ml of anhydrous tetrahydrofuran and the reaction mixture was stirred at room temperature for 30 min. Formation of a pale yellow precipitate was observed after the first 10 min. The solvent was removed under reduced pressure and the residue was analysed by NMR spectroscopy.

^{31}P {^{1}H} NMR (162 MHz, DMSO): δ=16.25 (br), 16.87 (br), 19.60 (br), 20.08 (br) ppm.

Synthesis of 11

<u>bis(2-(diphenylphosphane)ethyl)amine hydrochloride</u>

This compound was synthesized according to a literature procedure[62] and was kindly provided by Marzia Nuzzolo (University of St. Andrews).

<u>Borane protected bis(2-(diphenylphosphane)ethyl)amine hydrochloride</u>

The borane protection of bis(2-(diphenylphosphane)ethyl)amine hydrochloride was done based on a reported literature procedure.[63] A Schlenk tube containing the solution of bis(2-(diphenylphosphane)ethyl)amine hydrochloride (1.08 mmol, 517.55 mg, 1 eq.) in tetrahydrofuran (8 ml) was placed on ice-bath and a 2M solution of BH_3SMe_2 (2.27 mmol, 1.14 ml, 2.1 eq.) in tetrahydrofuran was slowly

added via syringe, under stirring, into the ice cooled flask. After that the solution allowed to reach room temperature and the stirring was continued for another 10 minutes, when the reaction was complete, according to the ^{31}P NMR spectrum of the crude. The excess of BH_3SMe_2 was removed by adding toluene and evaporating the solvent under vacuum, process repeated 3 times. The product was obtained as a white solid. Yield 100%.

^{1}H NMR (400 MHz, $CDCl_3$): δ=10.2 (s, br, 1H, N*H*), 7.50-7.37 (m, 12 H, ar), 7.70-7.62 (m, 8H, ar), 3.10 (m, 2H, CH_2), 2.81 (m, 2H, CH_2), 1.30-0.53 (br, 3H, BH_3) ppm. ^{31}P {^{1}H} NMR (162 MHz, $CDCl_3$): δ=12.92 ppm. ^{11}B {^{1}H} NMR (128 MHz, $CDCl_3$): δ=-40.1 ppm.

3-(N-maleimido)propionyl chloride – was synthesized according to a literature procedure [64] and was used for the next step without further purification.

Synthesis of **11**

The synthesis of the compound was done according to a modified literature procedure [62]. A Schlenk tube containing a mixture of borane protected bis(2-(diphenylphosphane) ethyl)amine (1.8 mmol, 924.25 mg, 1 eq.) and triethylamine (7.8 mmol, 786.14 mg, 1.1 ml, 4.25 eq.) in dichloromethane (15 ml) was placed on ice-bath and the 3-(N-maleimido)propionyl chloride (2.0 mmol, 377.22 mg, 1.1 eq.) in dichloromethane (1.5 ml) was added dropwise, under stirring, into the ice cooled flask. The solution allowed to reach room temperature and the stirring was continued for another 16 hours at room temperature. After that the solution was treated with 2M HCl (2x25 ml), the organic phase was dried over $MgSO_4$, filtered and concentrated under reduced pressure. The residue was purified by silica gel flash chromatography (10% ethyl acetate in dichloromethane) to afford **11** as white solid. Yield: 70% (793.7 mg).

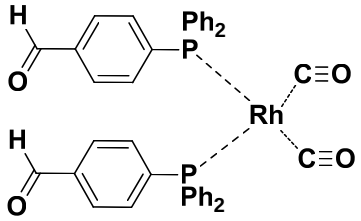
^{1}H NMR (400 MHz, $CDCl_3$): δ= 7.76-7.61 (m, 8H, ar), 7.50-7.33 (m, 12H, ar), 6.64 (s, 2H, *CH=CH* maleimide), 3.81-3.68 (t, J=7.36 Hz, 2H, CH_2), 3.54-3.40 (m, 2H, CH_2), 3.33-3.20 (m, 2H, CH_2), 2.63-2.30 (m, 6H, CH_2), 0.50-1.44 (m, 6H, BH_3) ppm. ^{13}C NMR (100.6 MHz, $CDCl_3$): *δ*= 170.4 (2x*CO*), 169.8 (1x*CO*), 134.3 (*CH=CH* maleimide), 132.0 (d, J=6.4, ar), 131.9 (d, J=6.4, ar), 131.7 (d, J=1.2, ar), 131.5 (d, J=1.2, ar), 129.2 (d, J=10.0, ar), 129.0 (d, J=10.0, ar), 128.6 (d, J=55.6, ar),127.8 (d, J=55.6, ar), 43.1 (d, J=6.9, CH_2), 41.5 (d, J=4.8, CH_2), 33.9 CH_2, 31.5 CH_2, 25.3 (d, J=34.5, CH_2), 23.7 (d, J=34.5, CH_2) ppm. ^{31}P {^{1}H} NMR (162 MHz, $CDCl_3$): δ=12.4, 11.9 ppm (br). ^{11}B {^{1}H} NMR (128 MHz, $CDCl_3$): δ= -39.8 ppm (br). HRMS (FTMS +): m/z calc. for $C_{35}H_{41}B_2N_2O_3P_2$ ($M+H^+$): 621.2785; found: 621.2780.

Attempt to synthesize the Rh complex 12

50 mmol (31 mg, 1 eq.) of ligand **11**, 0.3 mmol (20.57 µl, 6 eq.) of acrylic acid and 50 mmol (13.16 mg, 1eq.) of $RhCl_3$ were dissolved 10 ml of dry methanol. The mixture was heated at reflux for 1 hour. After cooling the solvent was removed under vacuum and the orange residue was analysed by NMR spectroscopy. ^{1}H NMR (300 MHz, MeOD): no peak at 6.68 corresponding to the C*H*=C*H* of maleimide; ^{31}P {^{1}H} NMR (121 MHz, MeOD): δ=52.73 (dd, J=127.4, J=23.58), 44.12 (dd, J=119.9, J=23.35).

ESI-MS (ES+) m/z= 697.03 – corresponding to the compound having chemical formula $C_{35}H_{36}N_2O_3P_2Rh$. ([M] $C_{35}H_{34}N_2O_3P_2Rh$ requires 695.5)

Synthesis of complex 19



19

78 µmol (22.5 mg, 2 eq.) of *p*-(diphenylphosphane)benzaldehyde in 0.65 ml anhydrous dimethylformamide was added to 39 µmol (19 mg, 1 eq.) of $Rh(acac)(CO)_2$. The yellow reaction mixture was left at room temperature for 30 min., after which the solution was used for protein modification.

Coupling of complex 19 to AppA Y21C modified with 3-(maleimido)propionic acid hydrazide

0.25 µmol (5 eq.) of complex **19** in 4.2 µl of dimethylformamide was added dropwise to a degassed solution of 50 nmol of AppA Y21C (1 eq.) modified with 3-(N-maleimido)propionic acid hydrazide in potassium phosphate buffer (50 mM, 1 ml, pH 7), in a Schlenk tube. The reaction mixture was stirred for 5 hours at room temperature. The unreacted rhodium complex was removed by centrifugation (49,000 g, 4°C, 10 min). The supernatant was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with degassed potassium phosphate buffer (50 mM, pH 7) and concentrated to 0.5 ml. This step was repeated four times after which the protein solution was concentrated to 1 ml and stored at 4°C.

Cu(II) complex formation of 5-maleimido-1,10-phenantroline modified proteins (ALBP and Δ98Δ) was performed according to a reported literature procedure.[5,19]

Fluorescence spectroscopy of apo- and modified proteins (ALBP and Δ98Δ) with 5-maleimido-1,10-phenanthroline and with the corresponding Cu(II) complexes was performed according to a reported literature procedure.[5,19]

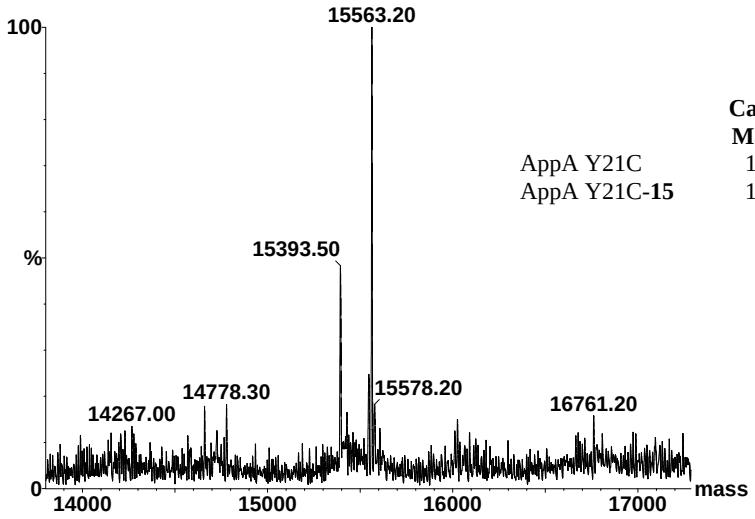
Complex 16 – was synthesized based on a reported literature procedure[53] and was kindly provided by Dr. Bianca Munoz (University of St Andrews).

Coupling of complex 16 to ALBP

A solution of 50 nmol of ALBP (1 eq.) in Tris·HCl (50 mM, 1 ml, pH 8) was degassed by purging with argon for 8 hours, in a Schlenk tube. While stirring, 2.50 μmol (50 eq.) of complex **16** in 50 μl of water was added slowly. The reaction mixture was stirred for 5 hours at room temperature. To remove the excess modification reagent, the reaction mixture was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with degassed Tris·HCl (50 mM, pH 8.0) and concentrated to 0.5 ml. This step was repeated four times after which the protein solution was concentrated to 1 ml and stored at 4°C.

Coupling of complex 16 to AppA Y21C

A solution of 50 nmol of AppA (1 eq.) in buffer (Tris·HCl 50 mM, 1 ml, pH 7) containing 3M imidazole was degassed by purging with argon for 8 hours, in a Schlenk tube. While stirring, 2.50 μmol (50 eq.) of complex **16** in 50 μl of water was added slowly. The reaction mixture was stirred for 5 hours at room temperature. To remove the excess modification reagent, the reaction mixture was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with degassed Tris·HCl (50 mM, pH 7.0) and concentrated to 0.5 ml. This step was repeated four times after which the protein solution was concentrated to 1 ml and stored at 4°C.



	Calculated Mass (Da)	LC-MS mass found (Da)
AppA Y21C	15395.1	15393.5
AppA Y21C-**15**	15564.2	15563.2

Figure 18 LC-MS (ES^+) spectrum of the reaction mixture of AppA Y21C with **16**

Complex 17

This complex was synthesized based on a reported literature procedure.[53]

41.3 mg (0.15 mmol, 2 eq.) of 5-maleimido-1,10-phenanthroline were added to a suspension of $[RuCl_2(p\text{-cymene})]_2$ (45.93 mg, 0.075 mmol, 1 eq.) in dichloromethane (20 ml) and the mixture was stirred for 3 hours at room temperature. After solvent removal under reduced pressure, the residue was dissolved in water, the solution was filtered and evaporated to dryness giving the product in quantitative yield as an orange solid. ^{1}H NMR (400 MHz, $CDCl_3$): δ= 0.96 (d, J = 6.9 Hz, 6H, *iso*-CH_3), 2.17 (s, 3H, CH_3), 2.8 (sept, J = 6.9 Hz, 1H, C*H*), 6.21 (d, J = 6 Hz, 1H, C_6H_4), 6.27 (d, J = 6.2 Hz, 1H, C_6H_4), 6.34 (d, J = 6 Hz, 1H, C_6H_4), 6.97 (s, 1H, CH maleimide), 6.98 (s, 1H, CH maleimide), 7.84 (s, 1H, C*H*= phen), 7.98 (m, 2H, C*H*= phen), 8.13 (d, J = 8.3, 1H, C*H*= phen), 8.46 (d, J = 8.3, 1H, C*H*= phen), 10.27 (d, J = 4, 2H, C*H*= phen).

Coupling of complex 17 to ALBP– was performed as described for complex **16**.

Complex 18

The ruthenium complex **17** (75.2 mg, 0.14 mmol, 1 eq.) was dissolved in distilled dichloromethane (20 ml) in a Schlenk tube covered with aluminium foil. 70.77 mg (0.28 mmol, 2 eq.) of $AgSO_3CF_3$ were added to the solution and the reaction was stirred for two hours at room temperature. The resulting suspension was filtered over celite, under argon, to remove the precipitated AgCl. The clear filtrate solution was collected, the solvent was removed under vacuum and the complex **18** was obtained quantitatively as an orange solid.

^{1}H NMR (400 MHz, D_2O): δ= 0.68 (d, J = 5.1 Hz, 6H, *iso*-C*H*$_3$), 1.98 (s, 3H, C*H*$_3$), 2.31 (sept, J = 5.1 Hz, 1H, C*H*), 5.99 (d, J = 4.7 Hz, 2H, C_6H_4), 6.21 (d, J = 4.7 Hz, 2H, C_6H_4), 7.01 (s, 2H, CH maleimide), 7.95 (m, 2H, C*H*= phen), 8.06 (s, 1H, C*H*= phen), 8.49 (d, J = 6.51, 1H, C*H*= phen), 8.69 (d, J = 6.2, 1H, C*H*= phen), 9.84 (d, J = 3.9, 2H, C*H*= phen),. ^{13}C NMR (100.6 MHz, D_2O): *δ*=17.6 (C*H*$_3$), 20.93 (CH(C*H*$_3$)$_2$), 20.96 (CH(C*H*$_3$)$_2$), 30.43 (C*H*(CH$_3$)$_2$), 83.3 (CH-C_6H_4), 86.5 (CH-C_6H_4), 86.6 (CH-maleimide), 101.6 (C-C_6H_4), 103.9 (C-C_6H_4), 126.6 (CH= phen), 127.0 (CH= phen), 127.2 (C= phen), 128.4 (C= phen), 128.7 (CH= phen), 130.08 (C= phen), 135.30 (C= phen), 136.3 (C= phen), 140.7 (CH= phen), 145.9 (C= phen), 146.6 (C= phen), 156.1 (CH= phen), 156.9 (CH= phen), 171.31 (C=O), 171.36 (C=O).^{19}F{^{1}H}NMR (δ, D_2O, ppm) -79.37 (CF_3SO_3).

MALDI-TOF m/z found: 529.084 (M.W. $C_{26}H_{25}N_3O_3Ru$ = 528.56)

Coupling of complex 18 to ALBP– was performed as described for complex **16**.

Coupling of complex 18 to AppA Y21C

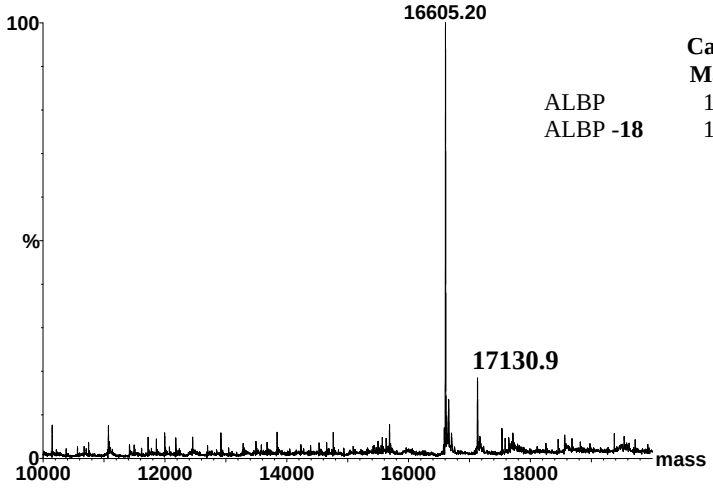
50 nmol protein (1 eq.) was maintained for 4 hours in buffer (Tris·HCl 50 mM, 1ml) containing 3M imidazole, at pH 7.0 and was further dialysed against Tris·HCl (50 mM, pH 7.0) to remove the imidazole and degassed by purging with argon for 8 hours. 2.50 μmol (50 eq.) of **18** in 50 μl of D_2O was added to the degassed protein solution (1 ml) after dialysis. The reaction mixture was stirred at room temperature for 5h. To remove the excess modification reagent, the reaction mixture was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with degassed Tris·HCl (50 mM, pH 7.0) and concentrated to 0.5 ml. This step was repeated four times after which the protein solution was concentrated to 1 ml and stored at 4°C.

Coupling of complex 18 to AppA Q63C – was performed as described for AppA Y21C.

Coupling of complex 18 to AppA Y21C in presence of 3M imidazole

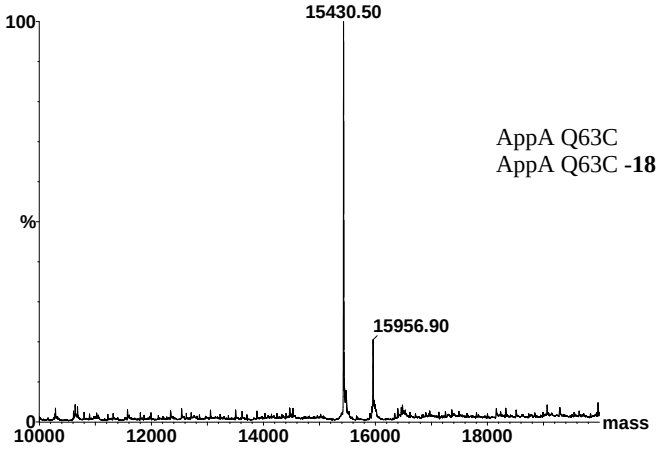
50 nmol protein (1 eq.) in buffer (Tris·HCl 50 mM, 1ml) containing 3M imidazole, at pH 7.0, was degassed by purging with argon for 8 hours. 2.50 μmol (50 eq.) of **18** in 50 μl of D_2O was added to the degassed protein solution. The reaction mixture was stirred at room temperature for 5h. To remove the excess modification reagent, the reaction mixture was transferred to a centrifugal concentrator (Amicon Ultra 15, 10,000 MWCO), diluted to 10 ml with degassed Tris·HCl (50 mM, pH 7.0) and concentrated to 0.5 ml. This step was repeated four times after which the protein solution was concentrated to 1 ml and stored at 4°C.

Coupling of complex 18 to AppA Q63C in presence of 3M imidazole– was performed as described for AppA Y21C.



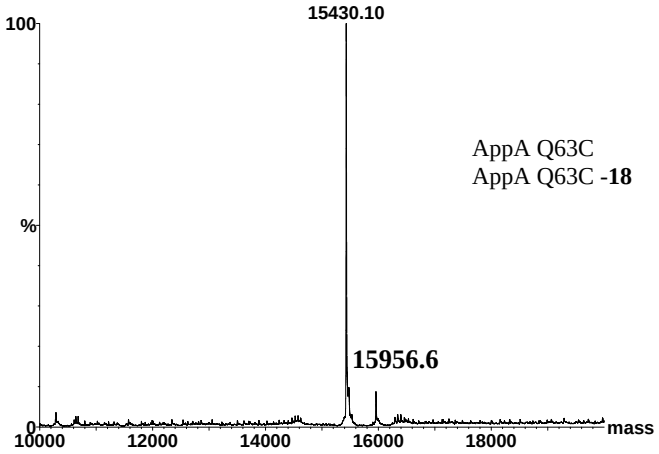
	Calculated Mass (Da)	LC-MS mass found (Da)
ALBP	16604.8	16605.2
ALBP -**18**	17131.3	17130.9

Figure 19 LC-MS (ES+) of ALBP - **18**, after addition of 0.5 equiv. of apo-protein



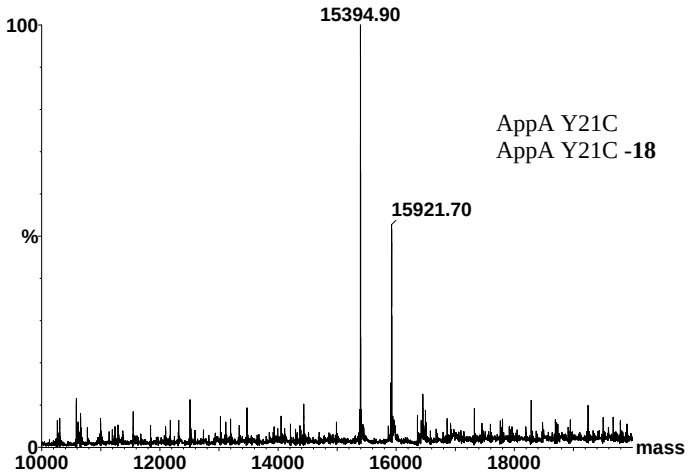
	Calculated Mass (Da)	LC-MS mass found (Da)
AppA Q63C	15430.1	15430.5
AppA Q63C -**18**	15956.6	15956.9

Figure 20 LC-MS (ES+) of AppA Q63C - **18**



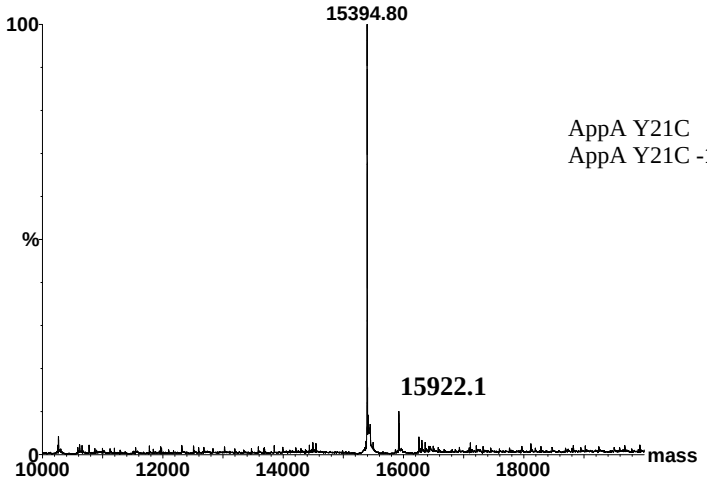
	Calculated Mass (Da)	LC-MS mass found (Da)
AppA Q63C	15430.1	15430.1
AppA Q63C -**18**	15956.6	15956.6

Figure 21 LC-MS (ES+) of AppA Q63C - **18** after addition of 0.5 equiv. of apo-protein



	Calculated Mass (Da)	LC-MS mass found (Da)
AppA Y21C	15395.1	15394.9
AppA Y21C -**18**	15921.6	15921.7

Figure 22 LC-MS (ES+) of AppA Y21C - **18**



	Calculated Mass (Da)	LC-MS mass found (Da)
AppA Y21C	15395.1	15394.8
AppA Y21C -**18**	15921.6	15922.1

Figure 23 LC-MS (ES+) of AppA Y21C - **18** after addition of 0.5 equiv. of apo-protein

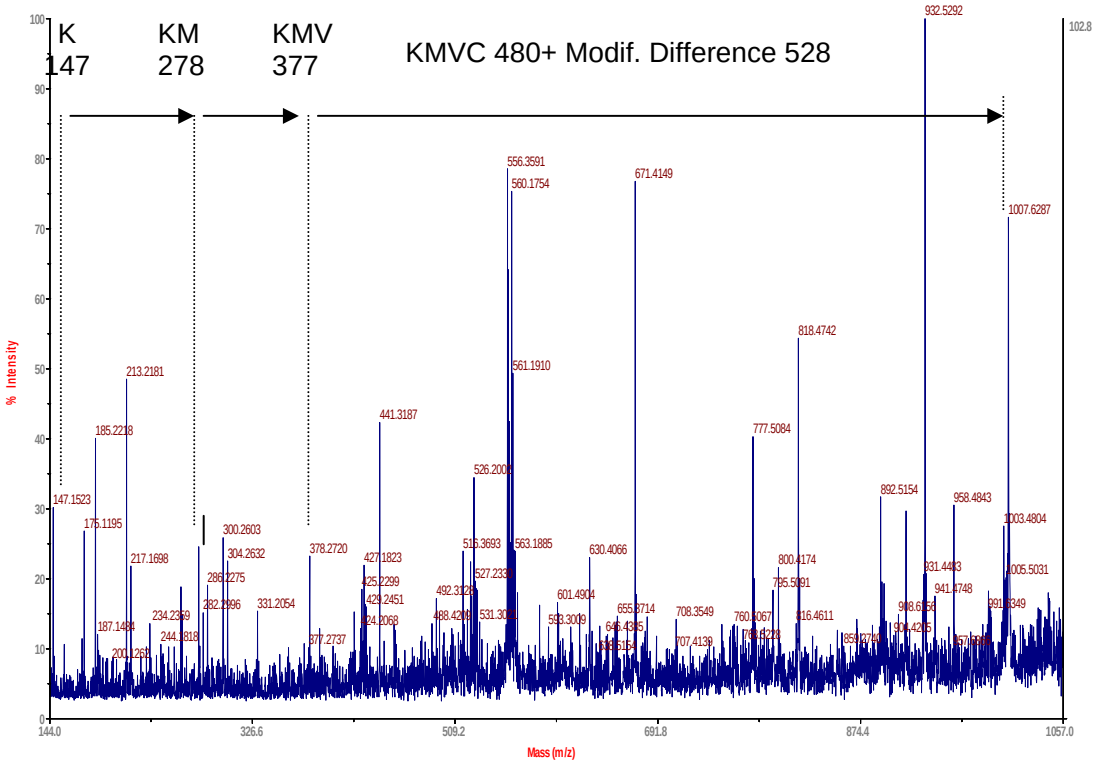


Figure 24 MALDI MS/MS spectrum of peptide Leu[131]-Lys[138]-**18** (at m/z=1448)

ALBP-18 tryptic digestion

The ALBP-**18** solution (5 µl, 10 pmoles/µl) was dialysed into ammonium bicarbonate (50 mM, pH 8.0) using a membrane filter (Millipore, Billerica, MA) and trypsin (0.5 µl, 0.1 µg, Promega, Madison, WI) was added. The sample was incubated at 37°C overnight.

MALDI-MS analysis of the peptides resulted from digestion of ALBP-18 with trypsin

The digest solution (0.5 µl) was applied to the MALDI target along with alpha-cyano-4-hydroxycinnamic acid matrix (0.5 µl, 10 mg/ml in 50:50 acetonitrile:0.1% TFA) and 0.1% TFA (0.5 µl) and allowed to dry. MALDI MS was acquired using a 4800 MALDI TOF/TOF Analyser (Applied Biosystems, Foster City, CA) equipped with a Nd:YAG 355 nm laser and calibrated using a mixture of peptides. The spot was initially analysed in positive MS mode between 800 and 4000 m/z, by averaging 1000 laser spots. The ion at m/z 1448 found in the spectrum of digested ALBP-**18** was assigned to the peptide $^{(131)}$LVVECVMK$^{(38)}$ modified with the complex **18**.

MALDI-MS/MS analysis of the peptide at m/z 1448

The peptide Leu^{131}-Lys^{138}-**18** was selected for MS/MS and the cysteine modification was confirmed.

The MS/MS analysis was acquired to a maximum of 3000 laser shots or until the accumulated spectrum reached a S/N ratio of 35. All MS/MS data were acquired using 1 keV collision energy.

The combined MS and MS/MS data were analysed, using GPS Explorer (Applied Biosystems) to interface with the Mascot 2.1 search engine (Matrix Science, London, UK), against the UniProt (Swiss-Prot and TREMBL combined) database (April 2009). No species restriction was applied. The data was searched with tolerances of 100 ppm for the precursor ions and 0.5 Da for the fragment ions, trypsin as the cleavage enzyme, assuming up to one missed cleavage, carbamidomethyl modification of cysteines as a fixed modification and methionine oxidation selected as a variable modification.

REFERENCES

1. Laan, W.; Munoz, B. K.; Den Heeten, R.; Kamer, P. C. *ChemBioChem* **2010**; *in press*.
2. den Heeten, R. In *PhD Dissertation: Amino acid modified phosphine ligands for the development of artificial transition metalloenzymes*; University of Amsterdam: Amsterdam, 2009.
3. Deuss, P. J.; Popa, G.; Botting, C. H.; Laan, W.; Kamer, P. C. J. *Angew. Chem. Int. Ed. Engl.* **2010**; *in press*.
4. Wilson, M. E.; Whitesides, G. M. *J. Am. Chem. Soc.* **1978**; *100*; 306.
5. Davies, R. R.; Distefano, M. D. *J. Am. Chem. Soc.* **1997**; *119*; 11643.
6. Levine, H. L.; Nakagawa, Y.; Kaiser, E. T. *Biochem. Biophys. Res. Commun.* **1977**; *76*; 64.
7. Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Maywald, M.; Maiwald, P.; Peyralans, J. J. P.; Maichele, A.; Fu, Y.; Jiao, N.; Hollmann, F.; Mondiere, R.; Taglieber, A. *Tetrahedron* **2007**; *63*; 6404.
8. Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Maywald, M. *Chimia* **2002**; *56*; 721.
9. Kruithof, C. A.; Casado, M. A.; Guillena, G.; Egmond, M. R.; van der Kerk-van Hoof, A.; Heck, A. J. R.; Gebbink, R. J. M. K.; van Koten, G. *Chem. Eur. J.* **2005**; *11*; 6869.
10. Kruithof, C. A.; Dijkstra, H. P.; Lutz, M.; Spek, A. L.; Egmond, M. R.; Gebbink, R. J. M. K.; van Koten, G. *Eur. J. Inorg. Chem.* **2008**; 4425.
11. Rutten, L.; Wieczorek, B.; Mannie, J. P. B. A.; Kruithof, C. A.; Dijkstra, H. P.; Egmond, M. R.; Lutz, M.; Klein Gebbink, R. J. M.; Gros, P.; van Koten, G. *Chem. Eur. J.* **2009**; *15*; 4270.
12. Nicholas, K. M.; Wentworth, P.; Harwig, C. W.; Wentworth, A. D.; Shafton, A.; Janda, K. D. *Proc. Nat. Acad. Sci. USA* **2002**; *99*; 2648.
13. Suckling, C. J.; Zhu, L.-M. *Bioorg. Med. Chem. Lett.* **1993**; *3*; 531.
14. Panella, L.; Broos, J.; Jin, J. F.; Fraaije, M. W.; Janssen, D. B.; Jeronimus-Stratingh, M.; Feringa, B. L.; Minnaard, A. J.; de Vries, J. G. *Chem. Commun.* **2005**; 5656.
15. Maywald, M.; Ruhr-Universität Bochum; 2005.
16. Pletsch, A.; Ruhr-Universität Bochum; 2004.
17. Kuang, H.; Brown, M. L.; Davies, R. R.; Young, E. C.; Distefano, M. D. *J. Am. Chem. Soc.* **1996**; *118*; 10702.
18. Kuang, H.; Davies, R. R.; Distefano, M. D. *Bioorg. Med. Chem. Lett.* **1997**; *7*; 2055.
19. Davies, R. R.; Kuang, H.; Qi, D. F.; Mazhary, A.; Mayaan, E.; Distefano, M. D. *Bioorg. Med. Chem. Lett.* **1999**; *9*; 79.
20. Letondor, C.; Pordea, A.; Humbert, N.; Ivanova, A.; Mazurek, S.; Novic, M.; Ward, T. R. *J. Am. Chem. Soc.* **2006**; *128*; 8320.
21. Letondor, C.; Ward, T. R. *ChemBioChem* **2006**; *7*; 1845.
22. Skander, M.; Humbert, N.; Collot, J.; Gradinaru, J.; Klein, G.; Loosli, A.; Sauser, J.; Zocchi, A.; Gilardoni, F.; Ward, T. R. *J. Am. Chem. Soc.* **2004**; *126*; 14411.
23. Klein, G.; Humbert, N.; Gradinaru, J.; Ivanova, A.; Gilardoni, F.; Rusbandi, U. E.; Ward, T. R. *Angew. Chem. Int. Ed. Engl.* **2005**; *44*; 7764.
24. Collot, J.; Gradinaru, J.; Humbert, N.; Skander, M.; Zocchi, A.; Ward, T. R. *J. Am. Chem. Soc.* **2003**; *125*; 9030.

25. Rusbandi, U. E.; Lo, C.; Skander, M.; Ivanova, A.; Creus, M.; Humbert, N.; Ward, T. R. *Adv. Synth.&Catal.* **2007**; *349*; 1923.
26. Skander, M.; Malan, C.; Ivanova, A.; Ward, T. R. *Chem. Commun.* **2005**; 4815.
27. Rusbandi, U. E.; Skander, M.; Ivanova, A.; Malan, C.; Ward, T. R. *Comptes Rendus Chimie* **2007**; *10*; 678.
28. Letondor, C.; Humbert, N.; Ward, T. R. *Proc. Nat. Acad. Sci. USA* **2005**; *102*; 4683.
29. Creus, M.; Pordea, A.; Rossel, T.; Sardo, A.; Letondor, C.; Ivanova, A.; Le Trong, I.; Stenkamp, R. E.; Ward, T. R. *Angew. Chem. Int. Ed. Engl.* **2008**; *47*; 1400.
30. Thomas, C. M.; Letondor, C.; Humbert, N.; Ward, T. R. *J. Organomet. Chem.* **2005**; *690*; 4488.
31. Pierron, J.; Malan, C.; Creus, M.; Gradinaru, J.; Hafner, I.; Ivanova, A.; Sardo, A.; Ward, T. R. *Angew. Chem. Int. Ed. Engl.* **2008**; *47*; 701.
32. Pordea, A.; Creus, M.; Panek, J.; Duboc, C.; Mathis, D.; Novic, M.; Ward, T. R. *J. Am. Chem. Soc.* **2008**; *130*; 8085.
33. Pordea, A.; Mathis, D.; Ward, T. R. *J. Organomet. Chem.* **2009**; *694*; 930.
34. Ohashi, M.; Koshiyama, T.; Ueno, T.; Yanase, M.; Fujii, H.; Watanabe, Y. *Angew. Chem. Int. Ed. Engl.* **2003**; *42*; 1005.
35. Rousselot-Pailley, P.; Bochot, C.; Marchi-Delapierre, C.; Jorge-Robin, A.; Martin, L.; Fontecilla-Camps, J. C.; Cavazza, C.; Menage, S. *ChemBioChem* **2009**; *10*; 545.
36. Ueno, T.; Koshiyama, T.; Abe, S.; Yokoi, N.; Ohashi, M.; Nakajima, H.; Watanabe, Y. *J. Organomet. Chem.* **2007**; *692*; 142.
37. Ueno, T.; Koshiyama, T.; Ohashi, M.; Kondo, K.; Kono, M.; Suzuki, A.; Yamane, T.; Watanabe, Y. *J. Am. Chem. Soc.* **2005**; *127*; 6556.
38. Shokat, K. M.; Leumann, C. J.; Sugasawara, R.; Schultz, P. G. *Angew. Chem. Int. Ed. Engl.* **1988**; *27*; 1172.
39. Nimri, S.; Keinan, E. *J. Am. Chem. Soc.* **1999**; *121*; 8978.
40. Ricoux, R.; Sauriat-Dorizon, H.; Girgenti, E.; Blanchard, D.; Mahy, J. P. *J. Immunol. Methods* **2002**; *269*; 39.
41. Ricoux, R.; Lukowska, E.; Pezzotti, F.; Mahy, J. P. *Eur. J. Biochem.* **2004**; *271*; 1277.
42. Yamaguchi, H.; Hirano, T.; Kiminami, H.; Taura, D.; Harada, A. *Org. Biomol. Chem.* **2006**; *4*; 3571.
43. Hermanson, G. T. *Bioconjugate techniques*; Academic Press: San Diego, 1996, pp 184.
44. Taglieber, A.; Ruhr-Universität Bochum; 2007.
45. Faller, J. W.; Wilt, J. C. *Organometallics* **2005**; *24*; 5076.
46. Faller, J. W.; Wilt, J. C. *Org. Lett.* **2005**; *7*; 633.
47. Zhang, W.; Xu, Q.; Shi, M. *Tetrahedron-Asymmetry* **2004**; *15*; 3161.
48. Riegel, N.; Darcel, C.; Stephan, O.; Juge, S. *J. Organomet. Chem.* **1998**; *567*; 219.
49. Darcel, C.; Kaloun, E. B.; Merdes, R.; Moulin, D.; Riegel, N.; Thorimbert, S.; Genet, J. P.; Juge, S. *J. Organomet. Chem.* **2001**; *624*; 333.
50. Detz, R. J.; Heras, S. A.; de Gelder, R.; van Leeuwen, P. W. N. M.; Hiemstra, H.; Reek, J. N. H.; van Maarseveen, J. H. *Org. Lett.* **2006**; *8*; 3227.

51. Boersma, A. J.; Klijn, J. E.; Feringa, B. L.; Roelfes, G. *J. Am. Chem. Soc.* **2008**; *130*; 11783.
52. Coquiere, D.; Feringa, B. L.; Roelfes, G. *Angew. Chem. Int. Ed. Engl.* **2007**; *46*; 9308.
53. Canivet, J.; Karmazin-Brelot, L.; Suss-Fink, G. *J. Organomet. Chem.* **2005**; *690*; 3202.
54. Curto, L. M.; Caramelo, J. J.; Delfino, J. M. *Biochemistry* **2005**; *44*; 13847.
55. Roelfes, G.; Boersma, A. J.; Feringa, B. L. *Chem. Commun.* **2006**; 635.
56. Ogo, S.; Abura, T.; Watanabe, Y. *Organometallics* **2002**; *21*; 2964.
57. Abura, T.; Ogo, S.; Watanabe, Y.; Fukuzumi, S. *J. Am. Chem. Soc.* **2003**; *125*; 4149.
58. Matarese, V.; Bernlohr, D. A. *J. Biol. Chem.* **1988**; *263*; 14544.
59. Laan, W.; van der Horst, M. A.; van Stokkum, I. H.; Hellingwerf, K. J. *Photochemistry and Photobiology* **2003**; *78*; 290.
60. Rodman, G. S.; Mann, K. R. *Inorganic Chemistry* **1988**; *27*; 3338.
61. Sameni, S.; Jeunesse, C.; Matt, D.; Toupet, L. *Chem. Eur. J.* **2009**; *15*; 10446.
62. Wilson, M. E.; Nuzzo, R. G.; Whitesides, G. M. *J. Am. Chem. Soc.* **1978**; *100*; 2269.
63. Naghipour, A.; Sabounchei, S. J.; Morales-Morales, D.; Hernandez-Ortega, S.; Jensen, C. M. *J. Organomet. Chem.* **2004**; *689*; 2494.
64. Mantovani, G.; Lecolley, F.; Tao, L.; Haddleton, D. M.; Clerx, J.; Cornelissen, J. J. L. M.; Velonia, K. *J. Am. Chem. Soc.* **2005**; *127*; 2966.