Diels-Alder Ligation of Peptides and Proteins

Zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
von der Fakultät für Chemie
der Universität Dortmund angenomene

Dissertation

Von
B.Sc.-Chemikerin
Aline Dantas de Araújo
aus Clevelândia (Brasilien)

March 2005

1. Gutachter: Prof. Dr. Herbert Waldmann
2. Gutachter: Prof. Dr. Martin Engelhard

Die vorliegende Arbeit wurde in der Zeit von April 2002 bis Februar 2005 am Max-Planck-Institut für molekulare Physiologie in Dortmund unter der Anleitung von Prof. Dr. Herbert Waldmann durchgeführt.

A minha família

Tudo vale a pena
Se a alma não é pequena.
Quem quer passar além do Bojador
Tem que passar além da dor.
Deus ao mar o perigo e o abismo deu,
Mas nele é que espelhou o céu.

It is worth while, all,
If the soul is not small.
Whoever means to sail beyond the Cape
Must double sorrow - no escape.
Peril and abyss has God to the sea given
And yet made it the mirror of heaven

Fernando Pessoa

CONTENTS

1. INTRODUCTION

The interface between chemistry and biology is certainly one of the most challenging and fruitful areas of research in life sciences at present. Biology has become more and more focused in understanding the natural processes of life at molecular level, with particular interest in determining the molecular structures of the biological entities and how they rule the intrinsic interactions with other biomolecules and small compounds. Therefore biology is moving towards chemistry and the frontiers between the two disciplines are intercalating. Chemistry not only plays an important role in the development of the pharmaceutical industry through drug design and delivery, but also provides exciting new ways to understand the mechanisms of complex living systems which could lead to future insights for the improvement of health care.

As classical biological techniques can not always supply the tools to study many biological phenomena in molecular details, the search of novel chemical methodologies that can address these issues is nowadays one of the topic research areas in biomedical science. In this field, the chemical-mediated synthesis and engineered modification of proteins have gained a lot of attention from the scientific community, reflecting new advances in biological research. Furthermore, whereas many biological assays employ specific non-covalent interactions that provide recognition between biomolecules in order to detect or localize a certain protein inside bioenvironments, the application of selective chemical reactions promises to emulate such high specificity by using small molecules that covalently bind to each other. In this sense, specificity is achieved by introduction of a reactive group into a particular protein that will uniquely modify a complementary chemical moiety of a target and then form a tight and irreversible junction between the two reaction partners.

This work describes the application of one of the most classical organic reactions, the Diels-Alder cycloadditions, for the strategic functionalization of complex biomacromolecules as proteins. The use of Diels-Alder reactions in protein science may appear paradoxical: the conditions in which these reactions are traditionally carried out during organic synthethic procedures (organic solvents, heating, use of catalysts) are contradictory with the type of reactions which are applied for protein chemistry (aqueous media, room temperature). Nevertheless Diels-Alder transformations can indeed be conducted in aqueous medium and the chemoselective nature of these reactions has been explored hereby to promote coupling of peptide segments and selective modification of proteins. The scope of the Diels-Alder ligation approach has been investigated using model peptides and proteins. The proposed methodology promises the possibility to equip a given protein (or polypeptide) with appropriate functionalities that may facilitate the investigation of a particular biological system.

2. THEORETICAL BACKGROUND

2.1. In the world of protein science[1-5]

An important objective of the biomedical science is to understand the molecular basis of proteins biological function. Since the explosive success of the genome-sequencing projects, this goal has been dramatically increased as hundreds of thousands of new proteins have been revealed, but only as predicted sequence data. A complete elucidation of the task of proteins within a biological system requires a full description of the protein structure and how their properties affect the inherent function as well as interactions with other molecules. To better understand these correlations, scientists are often confronted with the need of systematically altering the covalent structure of proteins with the intention to, for instance:

◈ introduce post-translational modifications, such as glycosilation, phosphorilation, lipidation, etc, which are fundamental transformations that rule protein activity inside and outside cells;

◈ incorporate appropriate biophysical probes into the protein molecule such as fluorescence markers or other reporter tags (biotin, epitopes, small ligands, etc) that permit protein detection and tracking within complex biological environments. Important biological events like protein-protein interactions, membrane/cytoplasm localization and cellular uptake of labeled proteins can be thus easily monitored in *in vitro* or *in vivo* assays;

◈ substitute strategically naturally occurring amino acids inside the protein structure by another amino acid or by an unusual scaffold which can then answer questions on the role of the specific amino acids in protein activity;

◈ alter protein properties in a specific manner to enhance activity or other physical characteristics like stability, solubitity, etc.

Proteins are a class of molecules characterized by both complexity and diversity, making their production a great challenge. However, for the last years, a range of biosynthetic and chemical synthetic approaches came out in the field of protein science which has led to a significant expansion of the spectrum of methods available for production of proteins either in their natural form or holding engineered modifications. An overview of the most important technologies that permit the covalent incorporation of unnatural molecules into proteins is illustrated in Figure 1 and discussed here in the next pages.

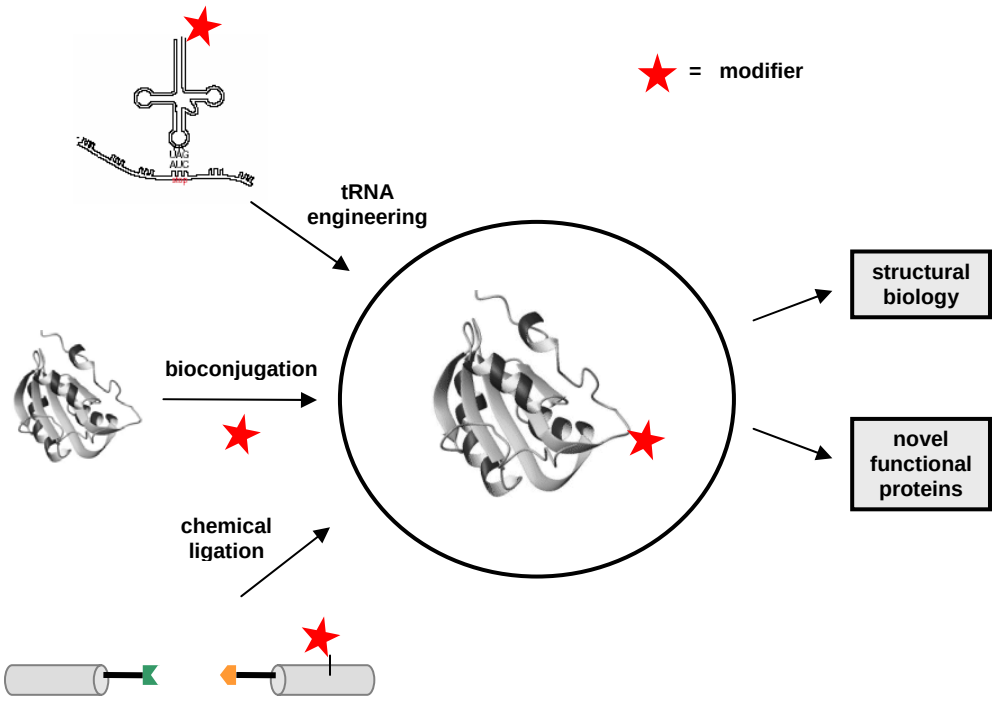


Figure 1. Methods for the covalent incorporation of non-natural molecules into proteins and their importance in biological sciences.

2.2. Biosynthetic methods for protein production[6-11]

Isolation of a particular protein among a myriad of other proteins and molecules found inside cells is not a trivial task either by laboratory synthesis or cell population purification. For the past 20 years, this task has been enormously facilitated by the development of biotechnological methods involving the recombinant DNA-based expression of proteins in genetically engineered cells. From its introduction until today, this powerful method revolutionized the study of proteins by enabling the production of large amounts of proteins of defined molecular composition and from different organism sources. It also allows the systematic variation of the peptide sequence of proteins with the diverse encoded amino acids (site-directed mutagenesis).

The ribosomal biosynthesis is usually performed employing the bacterium *Escherichia coli* as protein factory[12], although other organisms can be applied as well. Because the cell is used to manufacture the desired protein, the production of the macromolecule using this technology is limited to the 20 genetically encoded amino acids. Attempts to overcome this

limitation have been made to include noncoded amino acids as building blocks, in a method known as nonsense suppression mutagenesis or unnatural amino acid mutagenesis (Figure 2).[9] In this method, a desired point mutation is generated by replacing the codon of interest with an amber stop codon. Separetely, a tRNA is prepared that can recognize the amber stop codon. This tRNA is then charged with the desired amino acid derivative using both chemical and enzymatic coupling steps. Together these components are translated either *in vitro* or *in vivo*, and the unnatural amino acid is incorporated at the desired site. Over 100 different amino acids have been incorporated into dozens of soluble and transmembrane proteins using this technique.[6-8] Although this method promises to become a potential tool for production of artificial proteins, it is still in progress and, so far, can not find yet general applicability due to the low yields of mutant protein, technical demanding procedures and incompatibility of many unnatural amino acids with the ribosomal synthesis.

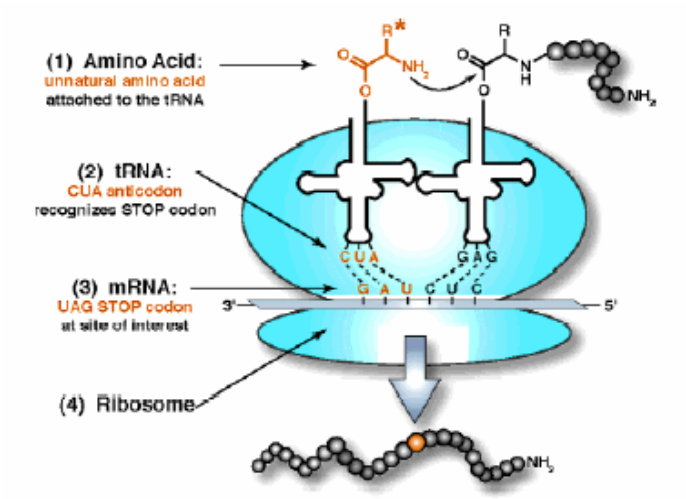


Figure 2. Key components for incorporation of unnatural amino acids into proteins using nonsense codon suppression.[8]

2.3. Chemical tools for assembly of proteins

The lack of general biotechnological tools to generate proteins enclosing noncoded modifications such as targeting probes, post-translational modifications, unnatural amino acids and other artificial modifiers has driven the development of different protein manufacturing approaches. In this scenario chemical synthesis has emerged as a powerful tool for protein

engineering. Because of its unmatched flexibility, chemical access to proteins provides the ability to incorporate unnatural alterations into proteins in a completely general fashion and has opened many new paths for the study of protein function.

2.3.1. Bioconjugation methods

Chemical reactions are the basis of the bioconjugation methods, a technology that has affected nearly every discipline in the life sciences, including scientific research, clinical diagnostics and human therapeutic markets. [13] Bioconjugation is the simplest and longest standing method for introduction of non-natural molecules into proteins. These methods make use of reactive functionalities found inside proteins to chemically connect a desired modification into the structure, thus creating unique conjugates that are able to interact with particular analytes in solution, cells or tissues. The amino, sulfhydryl and carboxylic acid groups present in the polypeptidyl molecules are the most used attachment points for protein targeting, due to their relative high reactivity in comparison with other groups present in amino acids. An overview of the most important reactions found in bioconjugation chemistry is outlined in Figure 3.

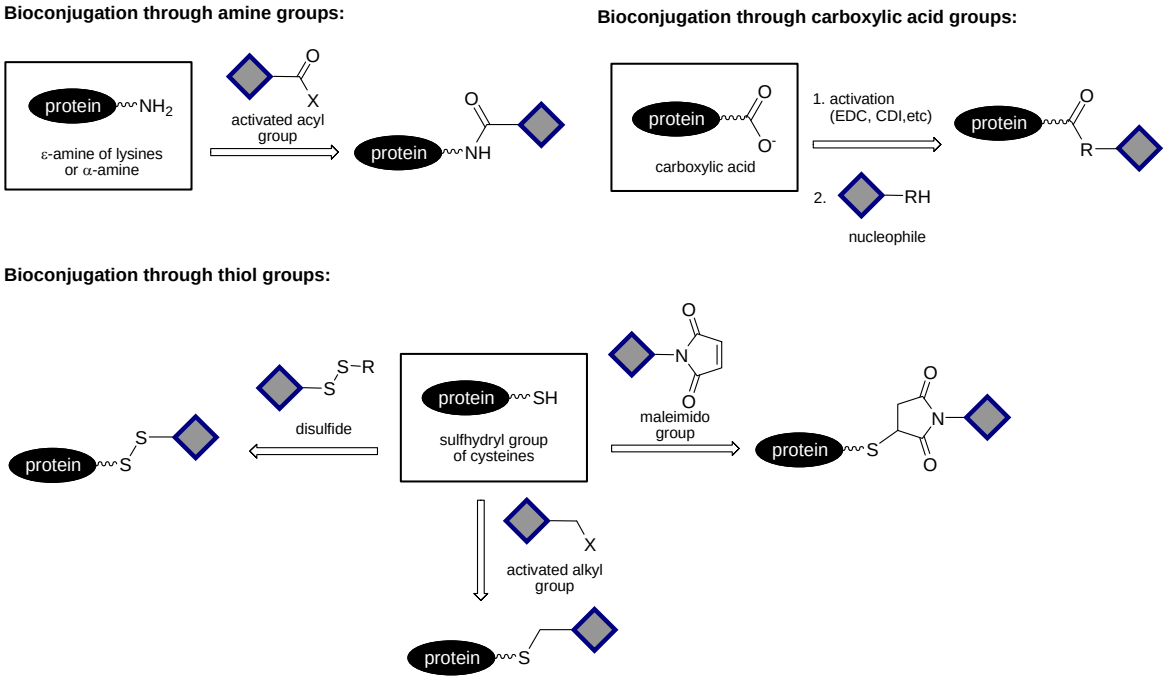


Figure 3. Overview of the most used bioconjugation reactions.

Despite their widespread application and versatility, the bioconjugation methods are quite limited in respect to the power of promoting site-specific protein modification. Uncontrolled alteration of a given protein molecule may perturb its structure in a way that it can not be longer functional. Also it may be very useful if one can introduce an appropriate target into a specific position of the protein structure in order to investigate the role of this particular region for the protein function. To overcome these problems, some approaches have been developed to provide moderate or even complete control over the bioconjugate derivatization, including:

◈ sequential combination of two or more bioconjugation steps which permits greater control over the conjugation process by using of heterobifunctional cross-linkers; [13]

◈ introduction of an unique cysteine into a particular position of the protein structure by site-directed mutagenesis, allowing the selective modification of this residue via thiol-reactive reagents; [14]

◈ moderate selective N-terminus derivatization by carrying out the amine-acylation reactions under slightly acidic conditions. [15]

The combination of biosynthetic methods for protein production and subsequent bioconjugation can not always provide the tools for all kinds of protein derivatization. A lot of efforts have been made in the last years to develop new technologies in this field. Scientists have found a way out to solve this problem by constructing protein molecules totally or partially by benchtop chemical reactions, permitting the production of proteins that possess either natural conformation or site-specific modifications.

2.3.2. Chemical protein synthesis

The first attempts to build up a protein molecule by chemical means were based on standard methods of peptide chemistry.[16] The application of classical solution-phase synthetic chemistry - where fully protected peptide segments are convergently condensed to form a large polypeptide - showed to be a quite demanding process. Most of the problems rely on the laborious preparation of the fully protected segments and their poor solubility in the coupling reaction medium. With the introduction of solid-phase peptide synthesis methods (SPPS), this area gained a great improvement. Together with the development of appropriate protecting groups and efficient coupling reagents, SPPS is nowadays capable to assemble polypeptides up

to 60 amino acid residues with relative facility. Moreover the method permits complete freedom to introduce modifications anywhere in the peptide sequence. Nevertheless, this number of amino acid residues that can be assembled via SPPS corresponds only to the very smallest proteins and protein domains. Different approaches involving the combination of solid-phase and solution-phase synthesis has been reported to construct larger proteins, yet they were achieved only after great efforts and laborious procedures. Noteworthy is also the development of enzymatic ligation methods for the preparation of proteins with enzymes specifically engineered to perform reverse proteolysis and to act as "ligases".[17] Nevertheless, despite some notable successes,[18-19] such methods have not found widespread use (yet) after the development of the more simple and versatile chemical ligation methods.

2.3.3. Chemical ligation methods[1,20-24]

The size of proteins which can be chemically synthesized has been increased considerably by the introduction of *chemical ligation* methods since early 1990´s in the area of protein chemistry. The innovative key concept of these methods is based on the linkage of peptide segments through chemoselective nonamide reactions in order to construct larger proteins. More specifically, these reactions involve the coupling in aqueous environment of two unprotected peptide segments bearing unique and complementary functional groups that are mutually reactive with each other, but unreactive with all other functionalities present in the segments to be coupled (Figure 4). Due to the chemoselectivity of the linking reaction, protection groups are unnecessary. The unprotected peptide segments to be used for the chemical ligation can be, in principle, easily synthesized either by chemical or biochemical means.

Because the mutually reactive groups (or at least one of them) are functionalities not normally found in peptides, the price to be paid for applying such chemoselective ligation is the formation of an unnatural structure at the ligation site. However, in practice, these unnatural structures are often well tolerated within the context of a folded protein and numerous examples of fully active proteins were prepared using these methods. Nevertheless, a very elegant approach called *native chemical ligation* has also been established which is able to generate true peptide bond-forming ligation.

Along the last years the chemical ligation methods have been shown to be a simple and highly effective strategy to construct large polypetides and proteins. Hundreds of

engineered proteins have been produced by means of this technology. The ligation strategies have found great application not just for the "clicking" of peptide segments to build a given protein macromolecule enclosing natural or unnatural properties; but also to tactically equip proteins with special functions that allow their discriminating targeting *in vitro* and *in vivo* experiments through chemoselective reactions. A variety of ligation chemistries has been used to perform such kind of reactions. Most of these reactions are based on imine or thiol chemistries, although new approaches involving the azido function have been recently developed. These techniques are briefly described in the next sections.

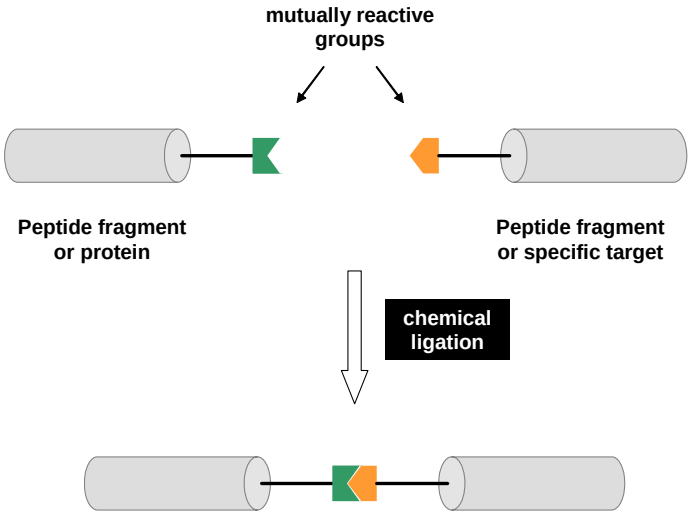


Figure 4. Principles of chemical ligation.

Aldehyde/ketone mediated ligations

The absence of aldehydes and ketones on the side chain of the naturally occurring amino acids makes this electrophilic functionality a candidate to perform unique reactions with a variety of nucleophiles. Selective peptide ligation has been obtained by reaction of aldehydes (or ketones) with hydroxylamines and hydrazines in the presence of protonated amino functions to form, respectively, oximes or hydrazones linkages (Figures 5A and 5B).[25-26] A recent application of this ligation method includes the synthesis of a glycoprotein human hormone erythropoietin protein polymer.[27] Because hydrazones are known to hydrolyze rather easily in water, the resulting backbone-engineered peptidyl hydrazone can be reduced with sodium cyanoborohydride to produce the more stable peptidyl hydrazide. Furthermore, the hydrazone ligation concept was explored for the development of a novel bioconjugation system (HydralinK^{TM}) which is based on the reaction of a 2-hydrazinopyridyl moiety with a

benzaldehyde moiety to yield a stable bis-aromatic hydrazone (Figure 5C).[28] The chemistry is highly selective and stable in solution, making it superior to conventional methods of bioconjugation such as maleimide/thiol and avidin/biotin.

Aldehydes can also react with β-amino thiols or alcohols (Cys, Ser or Thr) to form pseudoproline linkages (Figure 5D). The ligation involves an imine capture step that results from the coupling of the aldehyde group and the N-terminal amino group. Chemoselectivity is provided by the presence of –SH or –OH groups at the β-carbon position of the side chain, which permits the formation of a stable ring structure that can in turn rearrange to a pseudopropyl imide bond.[23] Tam and co-workers have demonstrated the effectiveness of the thiaproline ligation in the synthesis of analogues of TGF and HIV-1 protease.[29]

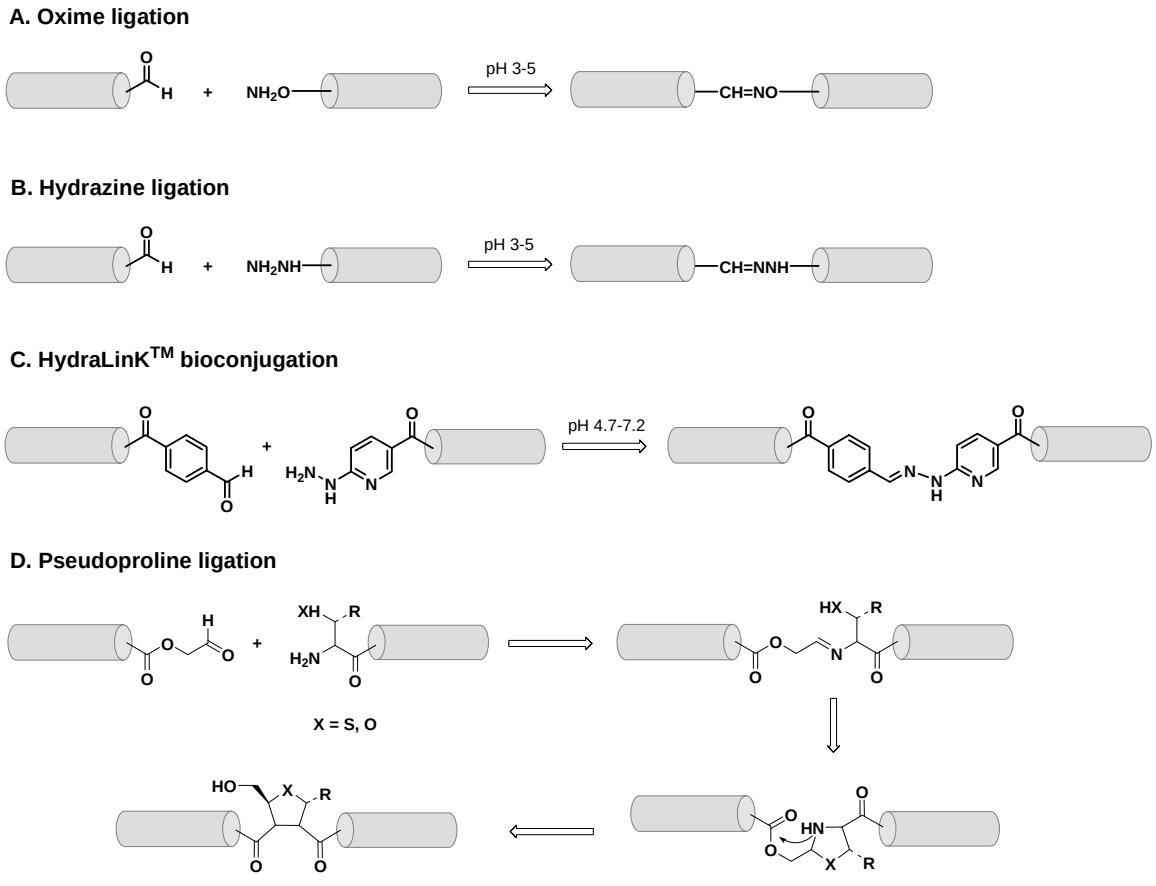


Figure 5. Aldehyde-based chemical ligations.

Thioacid and thioester mediated ligations

The thioester-forming ligation was the first example of backbone-engineered ligation, established by Schnolzer and Kent in 1992.[30] It is based on the generation of a thiocarboxyl

group at the C-terminus of a peptide segment that reacts at acidic pH with the N-terminal bromoacetyl group of the second peptide segment, forming a thioester moiety at the ligation site (Figure 6A). Under these conditions, all free amino groups are protonated and the reaction proceeds selectively. However the thioester linkage is only stable at pH range 3-6, being hydrolyzed at higher pH values. The synthesis of a HIV-1 protease analogue was the first example of application of this method.

The desire to assemble proteins with native backbone structures by chemoselective ligation reaction inspired Kent and co-workers to develop a novel thiol capture ligation approach[31] that generates amide bonds at the ligation site. The chemoselective step involves the reversible transthioesterifcation of a thioester modified C-terminus peptide with the thiol group of an N-terminal cysteine residue (Figure 6B). [32] A spontaneous, irreversible and rapid intramolecular S→N shift converts the thioester bond into a normal peptide bond, leading a cysteine residue at the ligation position. Internal Cys residues, if present in the peptide sequences, are not modified because the initial transthioesterification step is reversible and the S→N shift only occurs in the presence of the N-terminal amino group. To prevent oxidation of the N-terminal thiol, the reaction is carried out in the presence of thiols or other reducing reagents. While the presence of a Cys residue at the N-terminus is mandatory, almost all 20 amino acids can occupy the position of C-terminal thioester residue, excepting Val, Ile and Pro which react slowly and Asp and Glu which are prone for side-reactions.[33-34] This method, named native chemical ligation (NCL), is nowadays the most applied ligation strategy for the chemical-mediated construction of proteins.[1, 35]

Two variants of this approach were presented by Tam *et al.*.[36] Both methods are also based on the sequential capture and intramolecular acyl transfer principle where a thioacid is alkylated to form an intermediate thioester that will rearrange to give the cysteine at the ligation site (Figure 6C). Raines *et al.* have showed that the concept of NCL can also be extended to include selenocysteines.[37]

The main disadvantage of NCL is the necessity of a cysteine residue at the ligation site. The occurrence of this amino acid in proteins is very low and the insertion of additional Cys residues can alter the protein structure and thus its function by formation of unwanted disulfide bridges. Several approaches have been developed in the last years to circumvent this limitation. Cys-mimetic auxiliaries have been used to generate an amide bond leaving a glycine residue at the ligation position.[38-40] However, the native peptide conformation is only achieved after removal of the auxiliary under acidic or photolytic conditions. Furthermore, if

the presence of a Cys residue resultant from NCL is not desired, it can be transformed to an alanine residue by desulfurization mediated by palladium or Raney-nickel.[41]

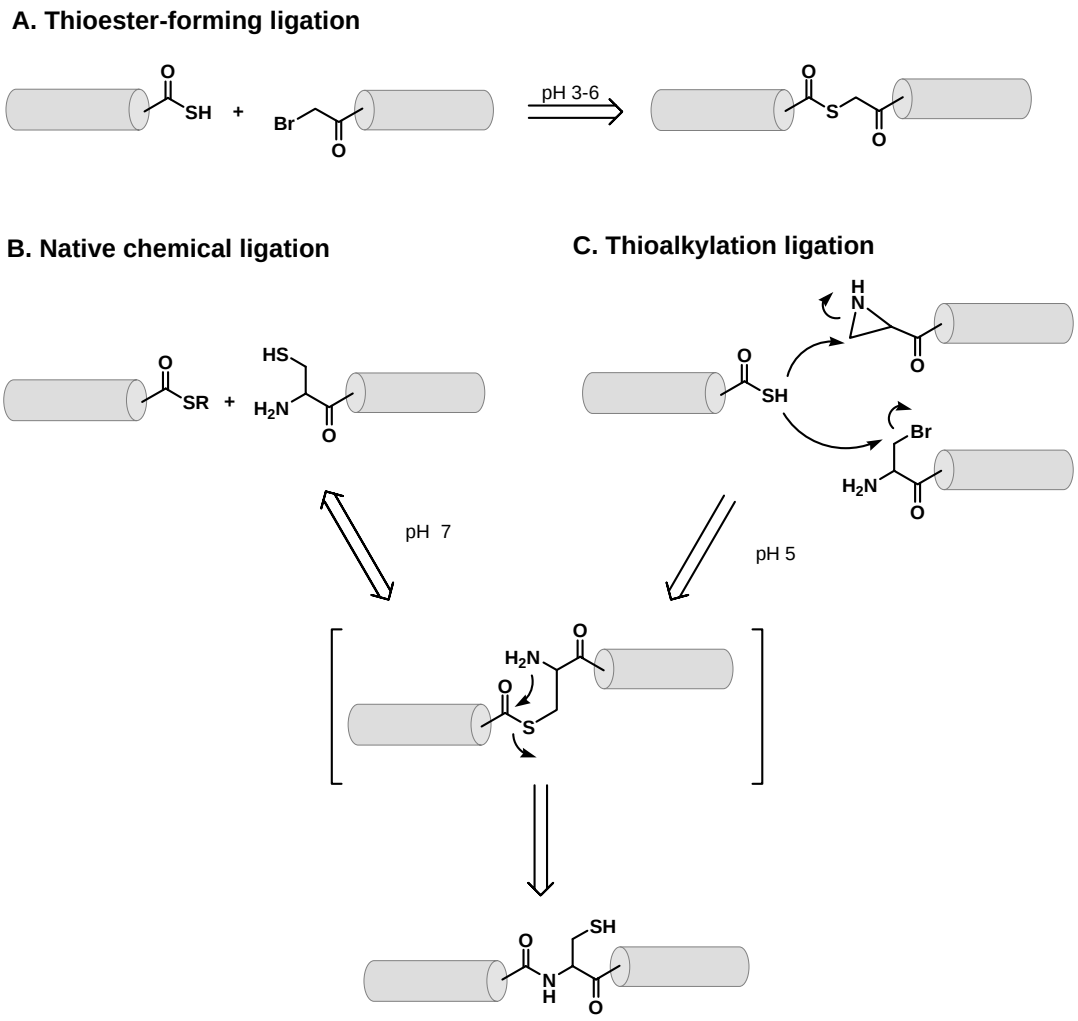


Figure 6. Thioester mediated ligations.

Azide mediated ligations

The first chemical ligation involving the azido group was based on the Staudinger reaction, where a phosphine reacts with an azide to form an aza-ylide intermediate that can rearrange to produce a stable amide bond (Figure 7A). In their pioneering studies involving the Staudinger ligation, Bertozzi and co-workers devised an appropriate phosphane ligand that allows effective coupling of this moiety with azido-derivatized molecules in aqueous media.[42] In just few years after its establishment, the Staudinger ligation approach has proved to be a valuable tool for the preparation of bioconjugates *in vitro* and for the targeting of biomolecules in the complex environment of living cells.[43] Furthermore, approaches to find a traceless

Staudinger ligation method, where a phosphane ligand is cleaved by hydrolysis thus leaving a native bond at the ligation site, are currently under development.[44-45] Although the Staudinger ligation can potentially be applied for noninvasive imaging and therapeutic targeting,[46] the reaction has some drawbacks. The required phosphines are susceptible to air oxidation and the optimization of their water solubility and increased reaction rate has proven to be synthetically challenging.

Sharpless *et al.* have demonstrated that the Huisgen [3+2] dipolar cycloaddition of azides and alkynes to give 1,2,3-triazoles are biocompatible and can be explored to promote selective linkage of proteins with chemical probes (Figure 7B).[47] The technique is nowadays known as "click chemistry" and has been applied for modification of virus particles, nucleic acids and proteins from complex tissues lysates.[48] The click ligation, however, requires the presence of a copper catalyst and other additives (reducing reagents and ligand) to be performed in reasonable reaction times. This condition may limit the general application of this methodology due to toxicity of the copper compounds and/or the additives to some biological systems. Nevertheless a recent approach reported by Bertozzi and co-workers promises to circumvent this problem by utilizing a strain-promoted azide-alkyne ligation, where the [3+2] cycloaddition reaction is driven by resultant ring stabilization of a strained cyclooctyne after ligation with an azido moiety.[49]

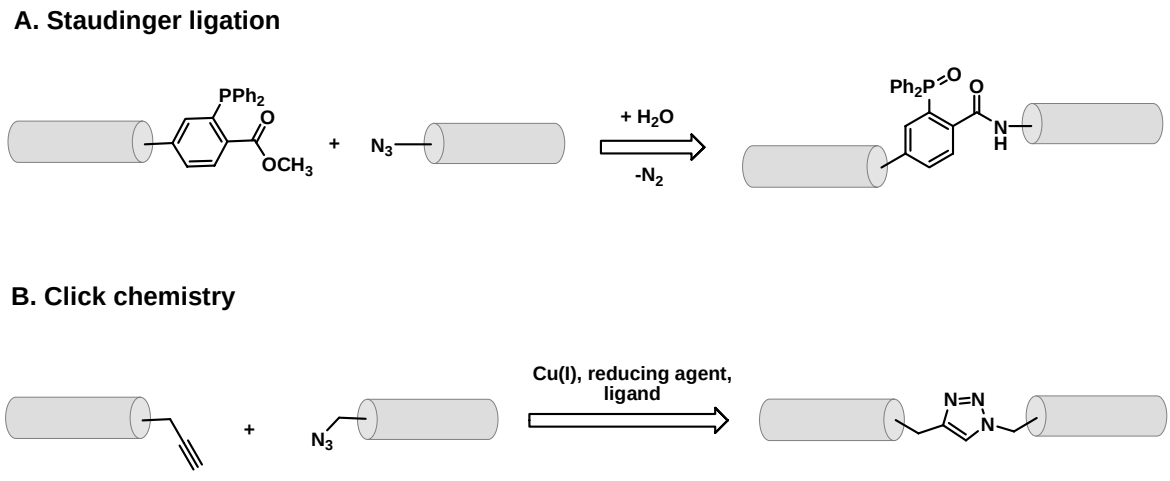


Figure 7. Azide-mediated ligations.

2.4. Combination of chemical ligation and biosynthetic methods

The combination of ligation techniques and solid-phase peptide synthesis has proven to be very useful for the synthesis of proteins up to 200 amino acids in length.[50] The assembly of larger proteins, however, requires the development of a multistep ligation procedure which can be rather technically difficult. The combination of chemical ligation and biosynthetic methods therefore is an attractive strategy to construct proteins of, in principle, unlimited size and of designed composition.

2.4.1. Expressed protein ligation

Some proteins undergo a process named *splicing*, in which two protein domains (exteins) are ligated with the concomitant elimination of the protein fragment (intein) between them (Figure 8). The intein itself is the catalyst of the splicing reaction and so far over 100 different inteins have been identified from diverse organisms.[51-52]

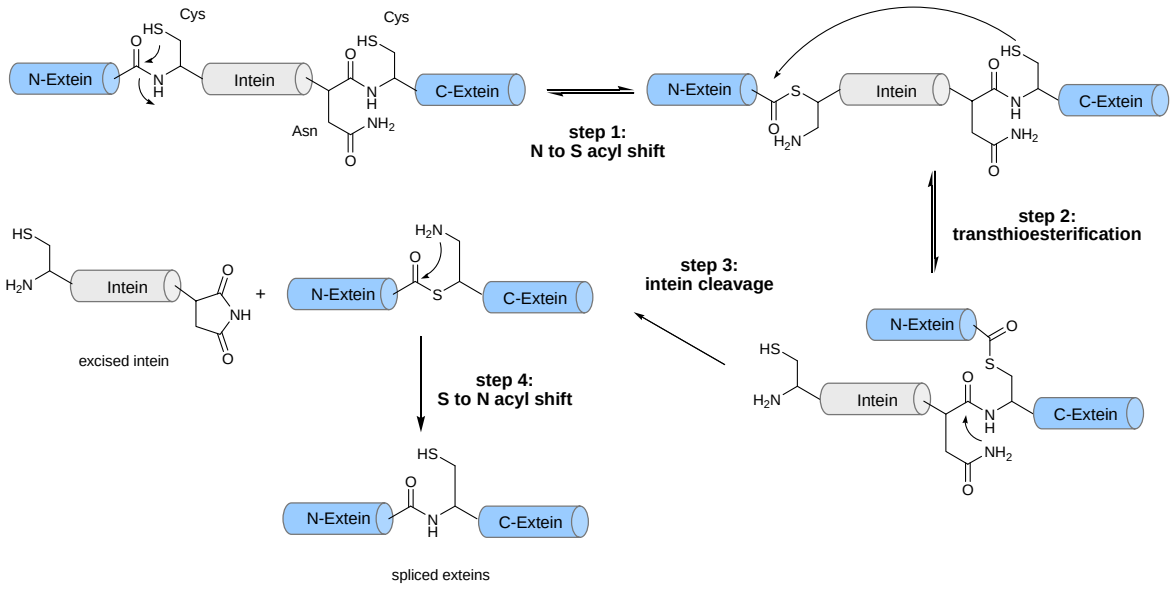


Figure 8. Mechanism of intein-mediated protein splicing.

Elucidations of the protein splicing mechanism have directed the design of engineered inteins that perform single splice-and-junction cleavage under specific conditions.[53] These

inteins, when fused to a particular protein either at its C or N terminus, may lead to the generation of a reactive C-terminal thioester or an N-terminal cysteine, respectively. In the case of the thioester formation (Figure 9), the strategy utilizes a mutation of the C-extein that prevents the splicing reaction to proceed after the initial acyl transfer reaction. The resulting thioester then becomes susceptible to undergo transthioesterification with added thiol reagents to release the intein and the thioester tagged protein. The C-terminus thioester in turn can be further modified by means of the native chemical ligation. The isolation of the protein-intein fusion complex after the expression step is facilitated by inclusion of an affinity tag (usually a chitin binding domain) in the intein fragment that permits immobilization of the fusion protein on a solid support before thiol-induced cleavage.

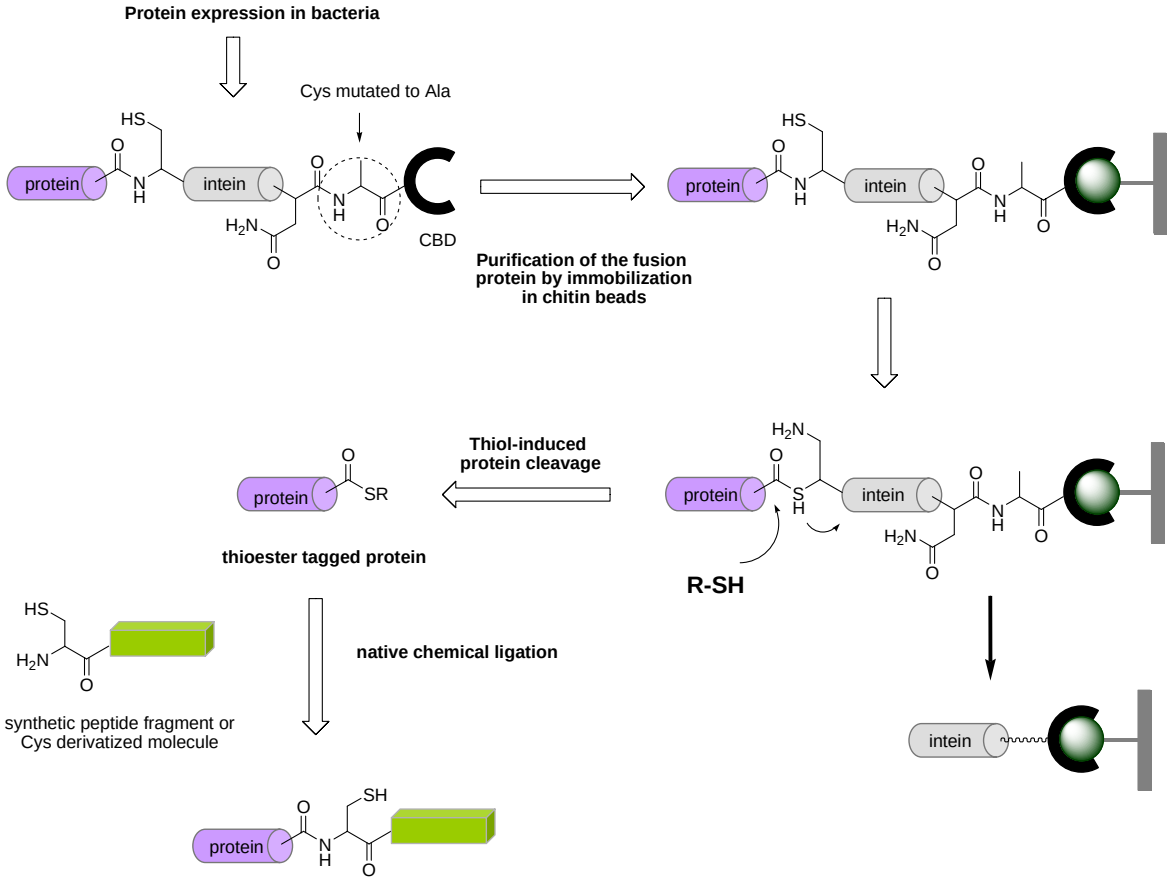


Figure 9. Principles of Expressed Protein Ligation. CBD = chitin binding domain.

This approach, known as expressed protein ligation (EPL), has found widespread applications since its introduction in 1998.[54] By allowing the controlled assembly of synthetic peptides and recombinant polypeptides, expressed protein ligation permits unnatural amino acids, biochemical probes, and biophysical probes to be specifically incorporated into

semisynthetic proteins.[2,55-57] Nevertheless, EPL (like NCL) is still limited mainly by the requirement of a Cys residue.

2.4.2. Unnatural amino acid site-mutagenesis and chemical ligations

As discussed before, suppressor tRNA techniques allow the use of the ribosomal machinery to insert a non-natural amino acid into proteins. Rather than just introduce the desired end-product amino acid, some recent approaches have instead demonstrated the incorporation of unique amino acid side chain featuring an orthogonal chemical functionality that can be further bioconjugated without interfering with other groups found inside the protein molecules. Using this approach, Schultz and co-workers have developed a method for the labeling of proteins in cells via hydrazone ligation using a ketone-modified protein.[58] The same group has reported recently a genetically-encoded incorporation of azide and acetylene tyrosine analogs into proteins that could be modified with dyes by copper(I)-catalysed click chemistry.[59] Furthermore, Bertozzi and co-workers demonstrated the modification of azidohomoalanine-labeled protein through Staudinger ligation with a phosphine reagent bearing an antigenic FLAG peptide (Figure 10). [60]

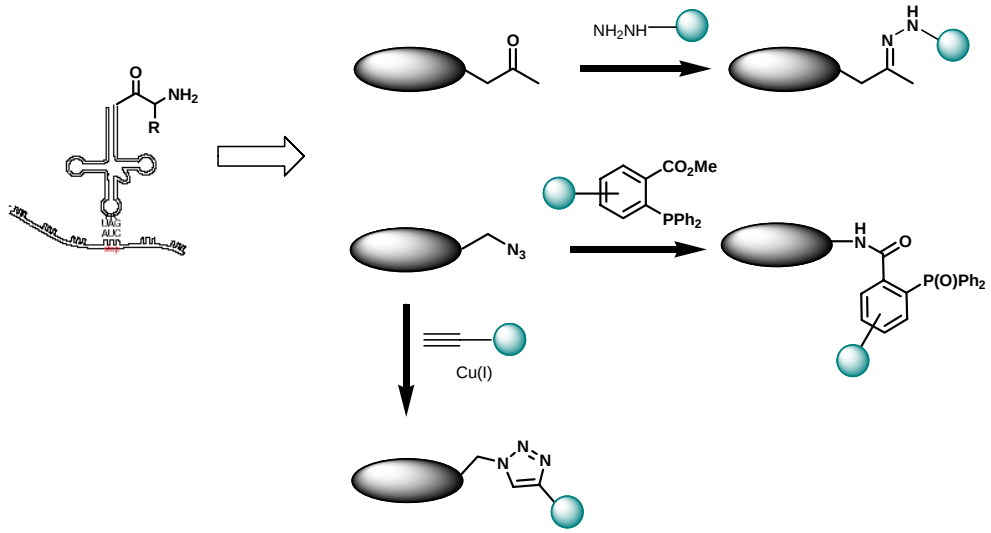


Figure 10. Combination of tRNA suppression technique and chemoselective bioconjugation for labeling of proteins.

2.5. Application of the chemical ligation methods for preparation of protein microarrays

Chemical ligation reactions can also be rationalized to be a valuable tool for the development of functional protein microarrays, an emerging branch of the proteomics field[61] that offers the possibility to simultaneously study a variety of proteins interactions in a microscale experiment. [62-66] By using a minute amount of sample, these miniaturized assays can be used for the high throughput analysis of interactions between proteins with other proteins, peptides, small molecules, oligosacharides or oligonucleotides. Nevertheless, the challenges when dealing with proteins microarrays are numerous and complex (in comparison with the established technology of DNA chips), requiring special manipulation and strategies to ensure appropriate spot uniformity, stable immobilization and preservation of desired protein activity in a microarray. Most of these aspects are dictated by the nature of the capture strategy in which the microarray is based. The development of appropriate capture agents is currently the most challenging bottleneck in protein microarray research.[67]

Glass slides have emerged to be a suitable surface to perform protein/peptide microarrays. They are inexpensive and possess great mechanical stability, low intrinsic fluorescence and a relatively homogeneous chemical surface. When used with appropriate bioconjugate chemistry, glass surfaces are capable of immobilizing biomolecules at very high densities. The surface of the glass slide is usually derivatized under specific conditions to generate functionalized layers. Immobilization of polypeptides is then subsequently carried out either by non-covalent or covalent linkage. Examples of non-covalent binding include: the interaction of antigens and antibody spotted surface; the binding of carbohydrates and nitrocellulose coated surfaces; the fixation of membrane proteins into lipidated surfaces.[64] Because the immobilized protein can adopt a variety of unpredictable orientations upon binding to the surface, these methods may lead to insufficient exposure of functional domains of a particular protein, rendering weak signals in further interactions with other analytes. The incorporation of recombinant affinity tags into specific sites of the protein molecule addresses the orientation issue (for instance, recombinant His-tagged proteins that binds to Ni-NTA-coated slides[68]). However the interactions of the tags, like the other approaches of non-covalent immobilization, are often reversible and may not be stable over the course of subsequent assays, resulting in graduate depletion of the protein from the microarray surface. More robust arrays are therefore obtained by covalent immobilization of the protein onto the glass surface. The first methods based on the covalent binding relied on the reaction of

chemical groups found within proteins (e.g. amines or thiols) with surfaces containing reactive groups (e.g. active esters, aldehydes, maleimides) using standard bioconjugation methods.[69] Here again the protein is attached to the surface in random orientations, which can often result in weaker signals because an unnecessary fraction of the biomolecules are immobilized with improper orientation, thus obstructing their binding with ligands.

Based on these facts, an attractive protein immobilization approach seems to involve the covalent binding of a protein onto a support surface via two unique and mutually reactive groups of small size, one present in a specific position of the protein and the other coated on the glass surface (Figure 11). Such type of linkage can be then fulfilled by chemical ligation strategies. While a number of research groups have demonstrated the use of chemical ligations to fix peptides, carbohydrates and other small biomolecules on glass surfaces (via aldehyde-mediated ligation,[70] native chemical ligation, [71] Staudinger ligation[72-73] and click chemistry[74]), only a few reports have shown a direct immobilization of an entire protein onto glass slides through chemoselective reaction. Coleman and co-workers have recently described the use of EPL for the creation of microarrays of proteins by covalent attachment of thioester tagged proteins onto a modified glass surface containing an N-terminal Cys poly(ethylene glycol) linker.[75] A reversed approach was developed by Yao´s group where proteins possessing a N-terminal Cys residue were immobilized on thioester functionalized glass surfaces. [76] EPL have been also employed to site-directed the immobilization of biotinylated proteins onto streptavidin coated surfaces[77] and of protein-nucleic acid conjugates onto DNA array containing capture oligonucleotides.[78]

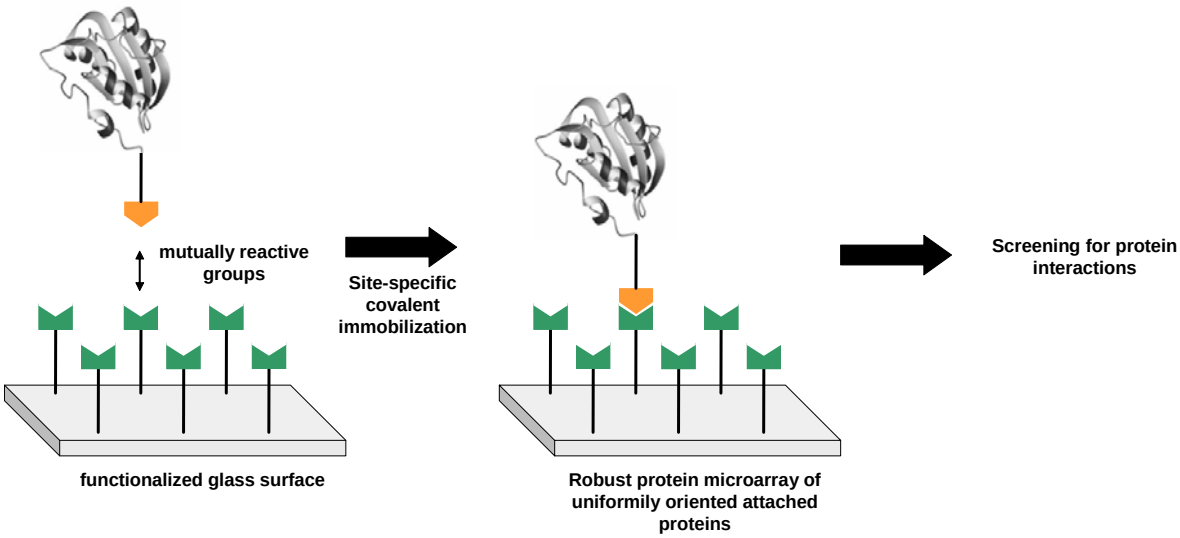


Figure 11. Principles of a protein microarray based on chemical ligation immobilization.

2.6. Other orthogonal methods for polypeptide and protein ligation

Thioester-, azide- and aldehyde-based orthogonal reactions have showed their practical value for the study of protein function. However the spectrum of application of these techniques is not unlimited. To broaden the applicability of the chemical approach for protein functionalization, the development of new bioorthogonal chemical linkages is required. In this scenario, Diels-Alder reactions appear to be an attracting alternative to perform covalent modifications with biomolecules.

2.6.1. Diels-Alder reactions in aqueous media

Named after the German chemists Otto Diels and Kurt Alder, who won the Nobel Prize in 1950 for their pioneering work on [4π + 2π] cycloadditions, the Diels-Alder (DA) reaction is one of the most important reactions in modern organic synthesis, featuring the formation in one step of two new carbon-carbon bonds in a chemoselective manner. Generally this reaction involves cycloaddition of a 1,3-conjugated double bond (diene) and an olefin equipped with electron-attracting groups (dienophile) to form a six-membered carbocycle (Figure 12).

Due to the hydrophobic nature of the reactants, organic solvents are the medium of choice for most synthetic Diels-Alder reaction procedures. However, since Breslow´s work in early 1980s,[79] many studies have shown that the Diels-Alder reaction often proceeds faster and with higher selectivity in water than in organic medium.[80-83] The origin for the water rate acceleration effect, although lacking complete understanding, seems to relay mainly on two effects: enforced hydrophobic interactions and activation of the dienophile by hydrogen bonding with water molecules.

Hydrophobic interactions between nonpolar parts of molecules in water are important non-covalent forces found in various biological systems. For instance, they participate in protein folding processes and enzyme-substrate interactions as well as play a crucial role in the aggregation of phospholipids and other lipidated compounds in biological cell membranes. In the context of the Diels-Alder reactions, the interaction between diene and dienophile is also benefited from the hydrophobic effect. When these two nonpolar entities react to form the cycloproduct molecule, the nonpolar surface area that is exposed to water is reduced during

the activation process (Figure 13). Therefore some of the water molecules that were before part of the hydration shell surrounding the reactants are now released to the bulky aqueous environment providing an additional driving force for the reaction.[84] The hydrophobic effect also explains why the preference for the formation of *endo*-cycloadduct is enhanced in water. The *endo* transition state is more compact than the extended *exo* transition state, thus allowing more water molecules to be removed from the hydration shells to the aqueous media, favoring the cycloaddition process.

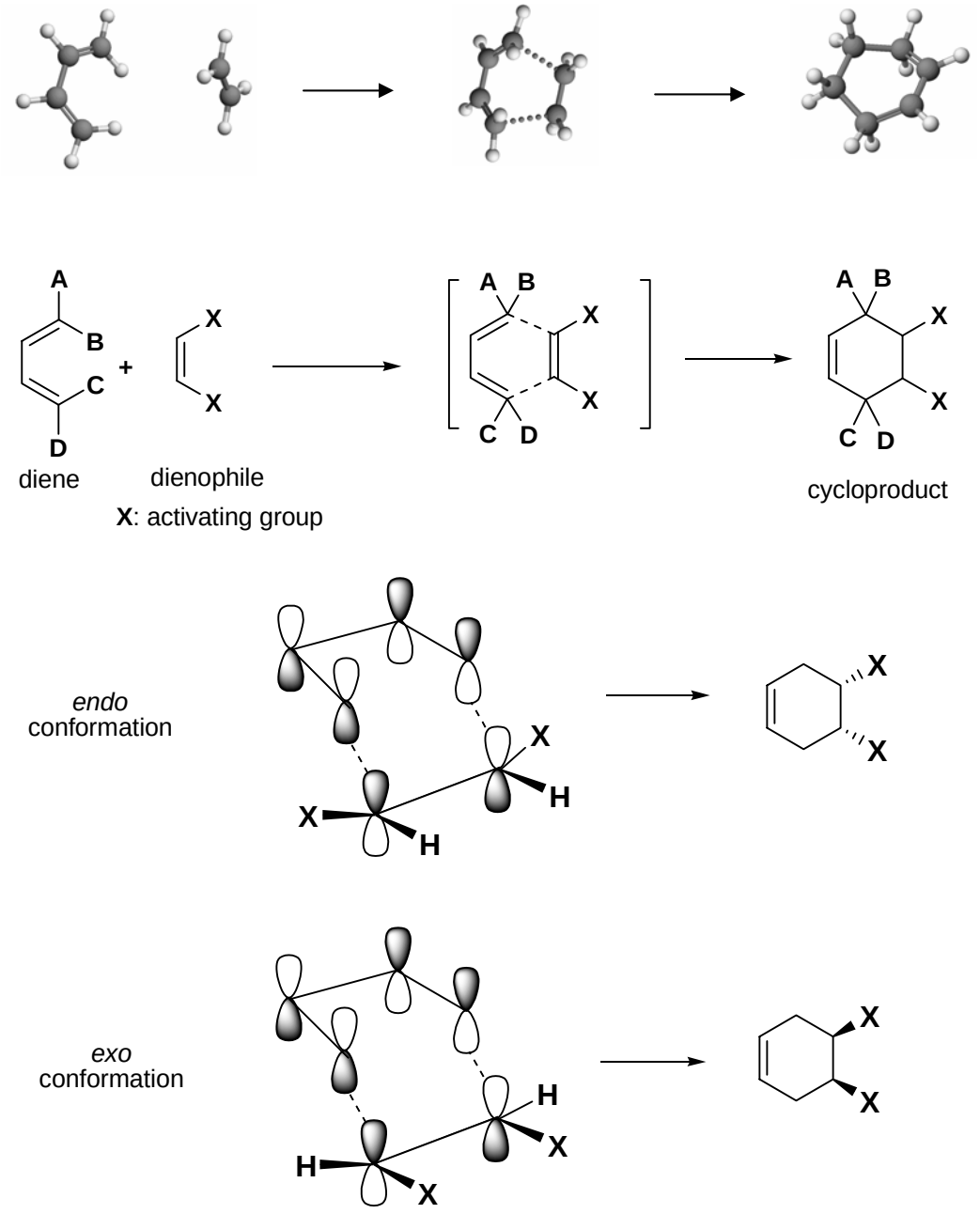


Figure 12. Principles of Diels-Alder cycloadditions.

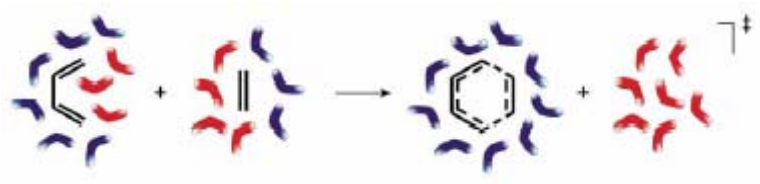


Figure 13. Schematic representation of the hydrophobic hydration shells surrounding starting materials and activated complex of a Diels–Alder reaction. The regions marked in red indicate parts of the hydration shell that are released into bulk solution upon reaction.[84]

Hydrogen bonding between water and the activating group of the dienophile (frequently a carbonyl group) is likely to be also responsible for the rate acceleration.[80-83] Similarly to the way Lewis acids activate Diels-Alder reactions in organic solvents, the water molecules coordinate with the carbonyl group to form an activated dienophile that is more electrophilic and thus more reactive towards cycloaddition with dienes (Figure 14).

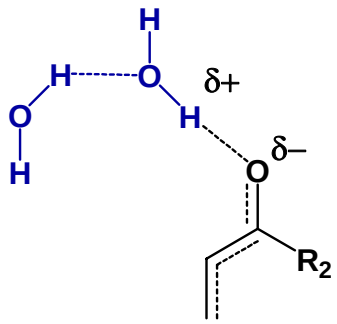


Figure 14. Hydrogen bond activation of the dienophile during Diels-Alder reactions in water.

2.6.2. Diels-Alder reactions as biotechnological tools

The ability to perform Diels-Alder cycloadditions in aqueous medium enables this highly chemoselective reaction to be regarded as a strong candidate to promote covalent modification of biomolecules, bringing up the opportunity to launch such versatile classical organic reactions into the biotechnological area. In fact, in the last few years, this idea has been explored by Sebesta and co-workers for the bioconjugation of nucleic acids. In a proof-of-concept approach, they demonstrated that synthetic oligonucleotides conjugated with a diene moiety could be targeted with dienophile derivatized fluorophores and biotinylated

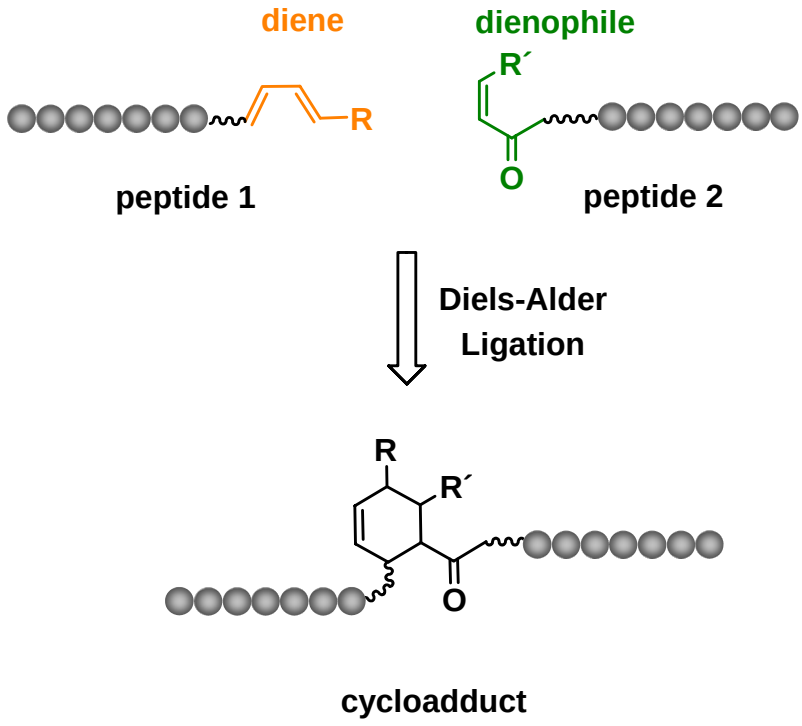
probes under aqueous conditions.[85] They also performed a few experiments to show that this method could be used to immobilize such oligonucleotides on surfaces.[86-87]

Pozsgay *et al.* also utilized the Diels-Alder reaction for the synthesis of a glyconjugate vaccine against gram-negative bacterium *Neisseria meningitidis* A.[88-89] Furthermore, Mrksich *et al.* have recently described in successive reports the development of biochips microarrays prepared by the Diels-Alder immobilization of monosaccharides[90] and small peptides[91-93] to self-assembled monolayers on gold-coated glass surfaces.

3. AIM OF THE THESIS

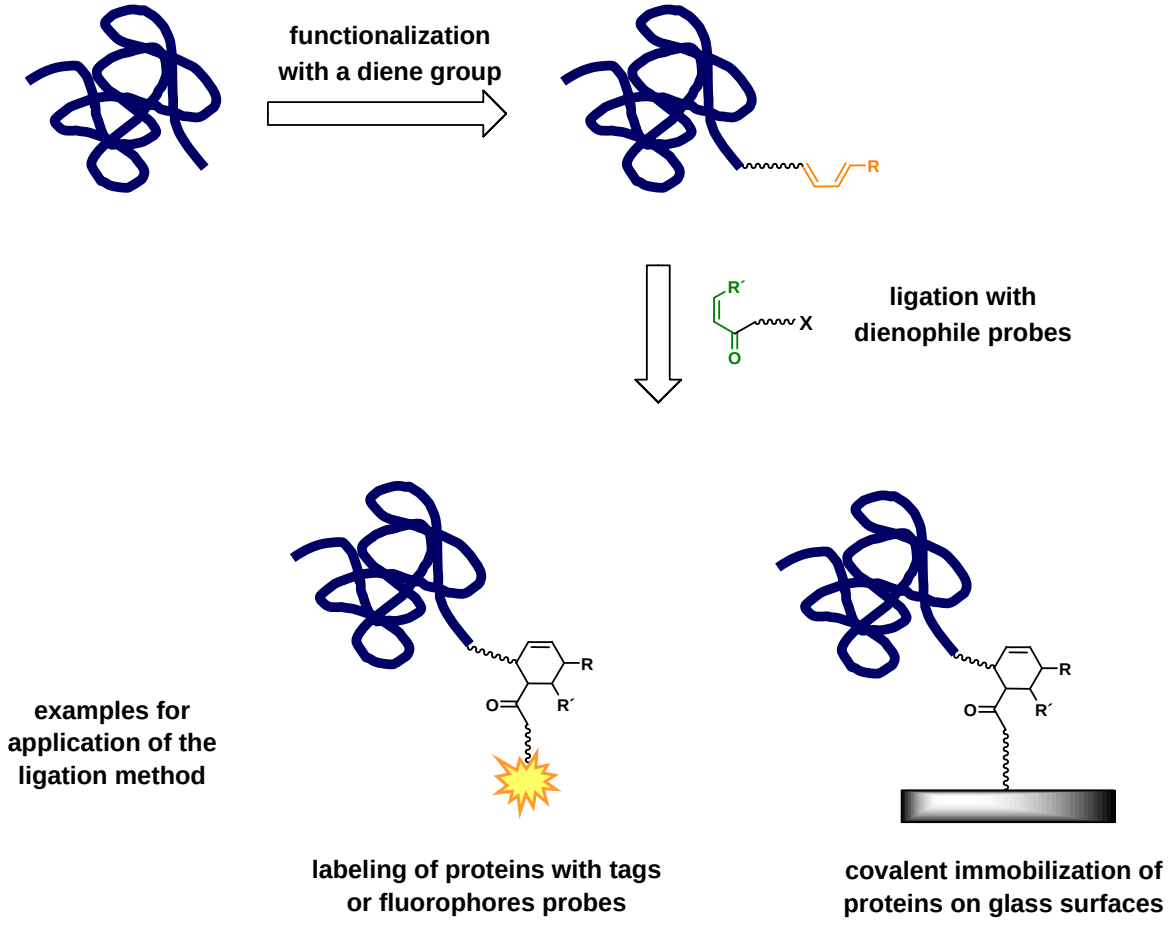
The massive number of applications in which chemical ligations have been used in the last few years reflects the importance of these technologies for the present and future of biology research. There is still a strong requirement for the development of other new chemical ligation approaches in order to expand the number of chemical tools for the creation of tailor-made proteins. A possibility to broaden this field is to find new suitable chemical reactions that can be carried out under physiological conditions and are orthogonal with respect to the functionalities present in polypeptides and proteins. Unfortunately these requirements are difficult to be fulfilled by most of the organic chemical reactions. Nonetheless, Diels-Alder cyloadditions can be considered as a promising alternative reaction to perform such kind of orthogonal linkages with biomolecules.

In light of these facts, this work is focused on the development of a novel chemical ligation method based on the Diels-Alder cycloadditions for the covalent modification of polypeptides and proteins. The first goal is to find suitable dienes and dienophile functions that could be easily incorporated into the peptide chain and could effectively and selectively undergo cycloaddition under mild aqueous conditions. To address these issues, a variety of diene- or dienophile-derived peptides are synthesized and the Diels-Alder ligation between these unprotected peptide segments is investigated in aqueous medium (Scheme 1).



Scheme 1. Diels-Alder ligation of peptides.

The next step is to investigate if the established Diels-Alder ligation method could also be applied for the covalent modification of entire proteins. The proposal is to equip a given protein with a diene unit that could later on be functionalized by Diels-Alder reactions with different dienophile probes. Examples of such type of protein derivatization are demonstrated herewith for the labeling of model proteins with fluorescent probes and for the protein immobilization on glass surfaces (Scheme 2).



Scheme 2. Functionalization of proteins by Diels-Alder reactions.

4. RESULTS AND DISCUSSION

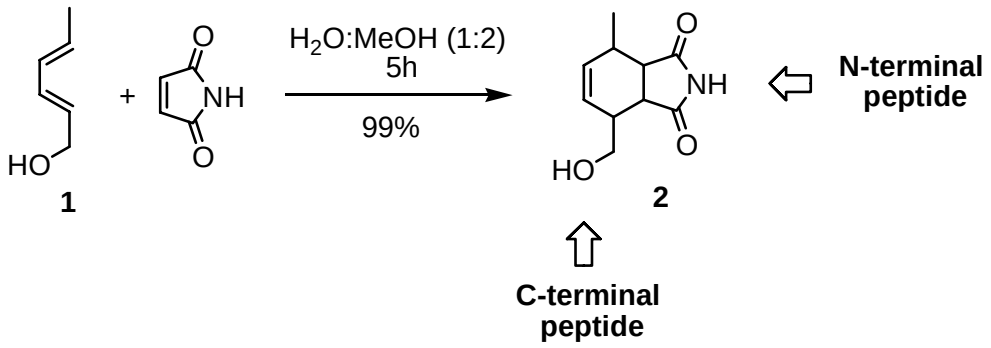
4.1. Peptide Ligation by Diels-Alder Reaction

4.1.1. First step: the choice of diene and dienophile

In order to exploit the Diels-Alder cycloaddition as a chemical reaction for peptide and protein modification two important requirements ought to be fulfilled:

- diene and dienophile functions should strongly react with each other under physiological conditions without need of catalysts or additional reagents;
- diene and dienophile groups are to be stable in aqueous medium and inert with respect to the range of functionalities found in proteins and biomolecules.

Based on these facts, hexadiene and maleimide were chosen as scaffolds for the first investigations of Diels-Alder ligation. Previous studies have shown that the acyclic hexadiene moiety is stable under aqueous environment and can undergo cycloaddition with reactive dienophiles.[94] Maleimido-compounds, on the other side, are among the most reactive dienophiles and yet extremely stabile under physiological conditions.[13] The effectiveness of the two selected functionalities in Diels-Alder cycloaddition was verified by reacting *trans*,*trans*-2,4-hexadienol **1** and maleimide in aqueous solution as illustrated in Scheme 3.



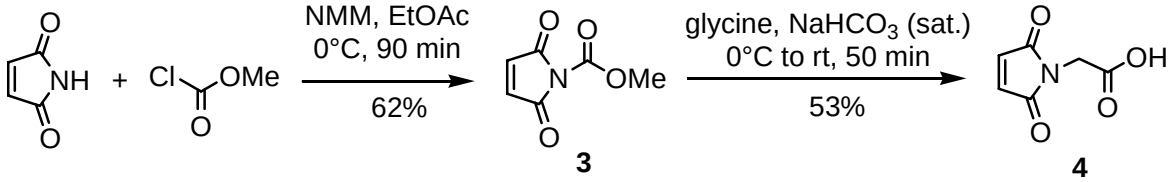
Scheme 3. Diels-Alder reaction between hexadienol and maleimide in aqueous solution and directions for peptide functionalization.

For the study of Diels-Alder peptide ligation, a set of diene- and dienophile-derived peptide segments was prepared: diene-peptides were constructed by C-terminal modification

with commercially available precursor **1**, whereas the dienophile handle was incorporated at the N-terminus (Scheme 3).

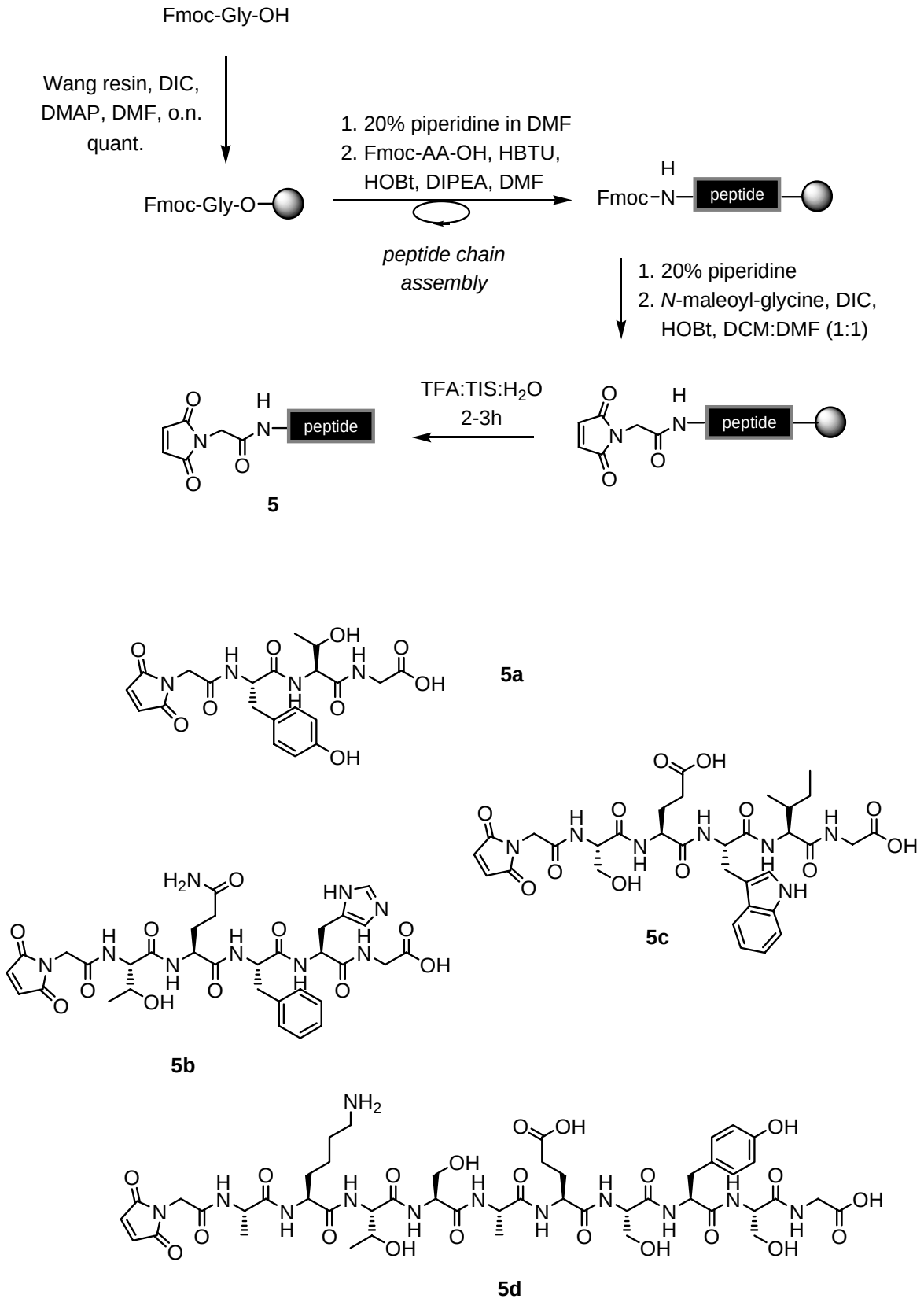
4.1.2. Preparation of the N-terminal dienophile peptides

The N-terminal maleimido-peptides were assembled by Fmoc/*t*Bu solid-phase strategy using Wang resin.[16] *N*-maleoyl-glycine was employed for introduction of the maleimide group in the last step of peptide sequencing. This compound was prepared by reaction of maleimide with methyl chloroformate and *N*-methylmorpholine in ethyl acetate, followed by conversion of *N*-methoxycarbonylmaleimide **3** with glycine in basic aqueous medium into *N*-maleoylglycine **4** (Scheme 4).[95]



Scheme 4. Synthesis of the building block *N*-maleoyl-glycine.

As revealed in Scheme 5, attachment of Fmoc-glycine, the first amino acid of the peptide sequence, was accomplished by DMAP catalyzed esterification of the hydroxyl-functionalized resin with diisopropylcarbodiimide (DIC) in DMF, giving quantitative resin loading as indicated by UV Fmoc-determination. The peptide chain was assembled by elongation cycles including Fmoc-protecting group removal with 20% piperidine in DMF followed by coupling of the next Fmoc-amino acid via HBTU/HOBt/DIPEA activation using standard procedures (4eq Fmoc-amino acid, 4eq HBTU, 4eq HOBt, 8eq DIPEA in DMF). Each residue coupling was monitored by Kaiser test. *N*-maleoyl-glycine **4** was coupled in the last cycle using DIC/HOBt activation in absence of base. Because the maleimide moiety is stable under acidic conditions,[95] side chain deprotection and cleavage were achieved trouble-free by treatment of the resin with TFA and scavengers, affording *N*-maleoyl peptides **5a-d** in 53-62% overall yield after lyophilization (Table 1). The identity of the four synthesized dienophile-peptides was confirmed by mass spectroscopy and NMR experiments (all products featured typical singlet signal at 6.8-7.0 ppm corresponding to the maleimido olefinic protons).



Scheme 5. Solid-phase synthesis of the *N*-maleoyl-peptides **5a-5d.** TIS: triisopropylsilane.

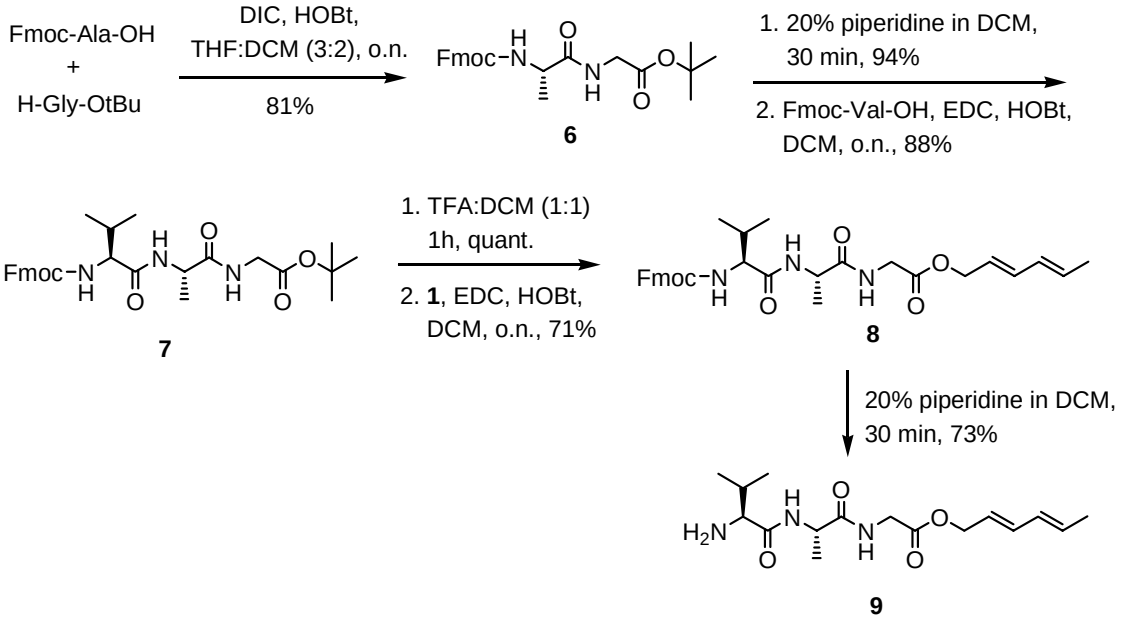
Table 1. Results for the synthesis of *N*-maleoyl-peptides **5**.

Maleimide	Peptide sequence	Overall yield (%)	ESI-MS for $[M+H]^+$	
			found	*calculated*
5a	YTG	62	477.0	477.1
5b	TQFHG	60	726.3	726.3
5c	SEWIG	53	728.1	728.3
5d	AKTSAESYSG	59	1137.5	1137.5

4.1.3. Preparation of the C-terminal diene peptides

For the synthesis of the diene-derived peptides in solution or solid phase, one should be concerned about the incompatibility of the diene functionality with acidic solid-phase cleavage conditions and protecting groups like *tert*-butoxycarbonyl (Boc), *tert*-butyl (*t*Bu) and benzyloxycarbonyl (Z). Base- or mild acid-labile protecting groups must then be applied.

The first template of a diene-modified peptide was prepared in solution as shown in Scheme 6.

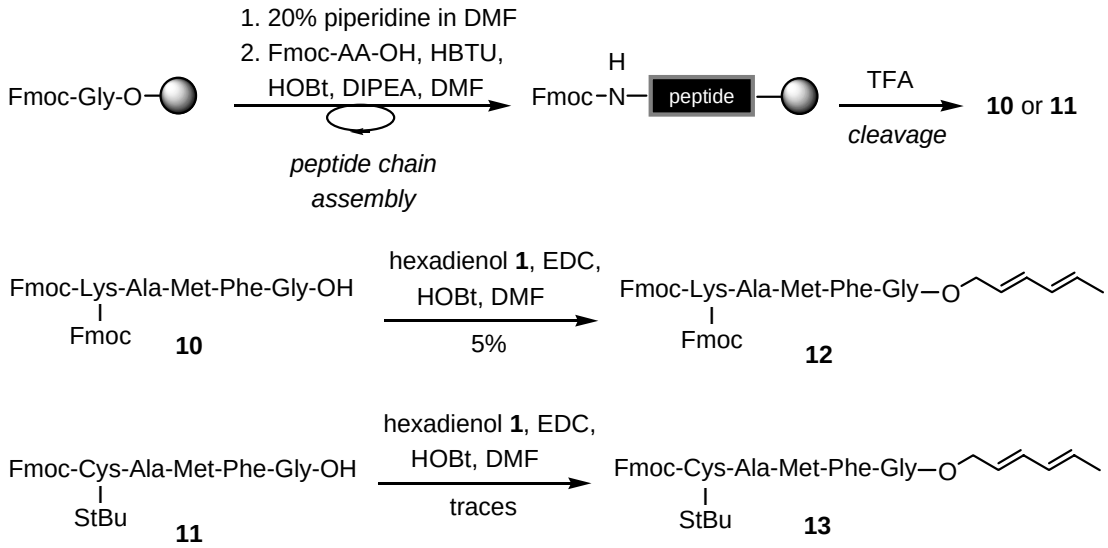


Scheme 6. Solution synthesis of tripeptide hexadienyl ester **9**.

The synthesis began with the preparation of protected dipeptide **6** using the DIC/HOBt activation method. After piperidine-mediated Fmoc-deprotection, the next amino

acid was coupled by the EDC/HOBt activation method to give tripeptide **7** in 67% yield over three steps (Scheme 6). Afterwards, the C-terminal *tert*-butyl protecting group was removed by acidolysis with TFA in DCM and the resulting free carboxylic acid was submitted to esterification with hexadienol **1** promoted by 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide.HCl (EDC) and DIPEA in DCM, to generate the Fmoc-peptide hexadienyl ester **8** in 71% yield. At last, **8** was converted into Val-Ala-Gly-hexadienyl ester **9** after removal of the Fmoc-group with 20% piperidine in DMF in 73% yield.

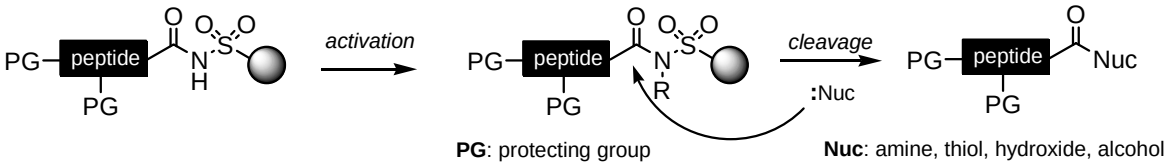
For the preparation of longer diene-peptide sequences, a solid-/solution-phase mixed approach was initially applied (Scheme 7). Fmoc-protected pentapeptides **10** and **11** were obtained by solid-phase synthesis following the Fmoc/*t*Bu methodology and using Wang resin as polymeric support. The fully protected peptides were then treated with hexadienol and coupling reagents using a similar esterification procedure as described above for the synthesis of hexadienyl ester **9**, but the reaction yielded very low amount of the desired peptide hexadienyl esters **12** and **13**.



Scheme 7. Attempts for the solution-phase esterification of fully protected pentapeptides.

Better results for the production of longer diene peptide segments were achieved by carrying out the complete synthesis on solid-phase using the sulfonamide-based safety-catch resin strategy. Initially introduced by Kenner *et al*,[96] the sulfonamide linker resin has been widely applied in SPPS for the preparation of carboxylic acid peptide derivatives. In this approach the linkage between the C-terminal residue and the resin linker is highly stable to the conditions of SPPS, but can be finally activated after peptide assembly by a mild chemical

reaction, resulting in an *N*-alkyl-*N*-acylsulfonamide which can then be cleaved by nucleophiles and provide compounds possessing a variety of carboxyl group modifications:



Scheme 8. Safety-catch strategy for C-terminus derivatization using sulfonamide linker resin.

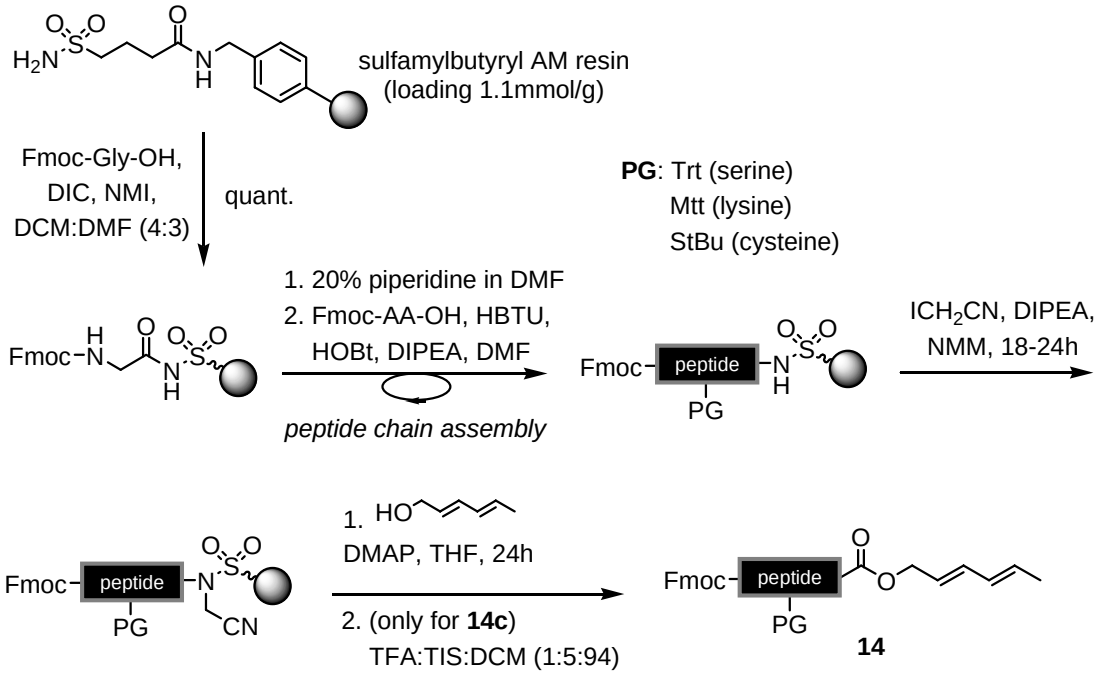
The diene-modified peptides were built on the sulfamylbutyryl linker resin (a modification of Kenner´s linker developed by Ellman and co-workers[97]) using Fmoc, trityl (Trt), 4-methyltrityl (Mtt) or *tert*-butylthiol (S*t*Bu) for protecting N-terminal and reactive side-chain groups of the amino acid residues during peptide chain growing and activation/cleavage steps (Scheme 9). Due to the poor nucleophilicity of the sulfonamide function, attachment of the first amino acid onto this resin is often a demanding process, giving low loading and possible racemization.[98] For this reason, simple glycine was selected as the C-terminal amino acid for all diene-peptide sequences. Still quantitative loading of Fmoc-glycine was only achieved using a large excess of amino acid and coupling reagents over extended time (Table 2).

Table 2. Loading of Fmoc-glycine onto sulfamylbutyryl resin.

Conditions	Loading
4eq Fmoc-Gly-OH, 4eq PyBOP, 8eq DIPEA, -20°C to rt, overnight	65%
4eq Fmoc-Gly-OH, 4eq PyBOP, 8eq DIPEA, -20°C to rt, overnight (2x)	81%
4eq Fmoc-Gly-OH, 4eq DIC, 4eq *N*-methylimidazole, rt, 26h	40%
8eq Fmoc-Gly-OH, 8eq DIC, 8eq *N*-methylimidazole, rt, 2x 18h	100%

Once the peptide chain synthesis was complete via the HBTU/HOBt activation protocol (in the same way as described for assembly of compounds **5**), the fully protected peptidyl sulfonamide was activated by alkylation with iodoacetonitrile and DIPEA in NMP for 18-24h (Scheme 9). The resulting *N*-cyanomethyl activated support was treated with a nucleophilic combination of hexadienol **1** and DMAP in dry THF for 1 day to release the protected peptide hexadienyl esters **14**. Figure 15 depicts a typical HPLC analysis of the

cleavage mixture. The crude cleavage products were purified by reversed-phase HPLC to eliminate excess of hexadienol and DMAP. When necessary, further trityl deprotection was obtained by combining the peptide with 1% trifluoroacetic acid (TFA):5% triisopropylsilane (TIS) in dichloromethane for 2h at room temperature (only for peptide **14c**). Hexadienyl ester peptides **14a-h** were isolated in 10-43 % overall yield after HPLC purification (Table 3). In general the only by-product observed after cleavage from the resin was the non-esterified peptide found in smaller amounts (approximately 15% in comparison to the ester formation).



Scheme 9. Solid-phase synthesis of the protected peptide hexadienyl esters using the sulfonamide safety-catch linker resin.

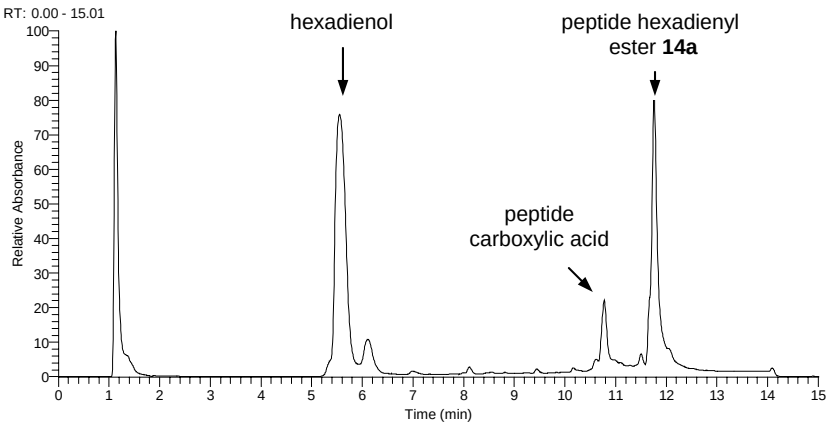


Figure 15. HPLC analysis of the cleavage mixture for the synthesis of peptide **14a**.

Table 3. Overall yield for the solid-phase synthesis of dienyl protected peptides **14** after HPLC purification.

Diene	Peptide sequence	Isolated yield (%)	MS $[M+H]^+$ *found*	MS $[M+H]^+$ *calculated*
14a	Fmoc-K^{Fmoc}PFLG	43	1085.3[a]	1085.5
14b	Fmoc-PC^{StBu}SMG	20	572.3[a]	572.3
14c	Fmoc-K^{Fmoc}LGFAG	33	1116.2	1116.5
14d	Fmoc-K^{Fmoc}LGK^{Mtt}AG	32	1353.4[a]	1353.7
14e	Fmoc-K^{Fmoc}C^{StBu}GVFG	22	1222.2[a]	1222.5
14f	Fmoc-K^{Fmoc}FPIGLFG	16[c]	1424.7[b]	1424.7
14g	Fmoc-K^{Fmoc}FPIGLGFG		1481.7[b]	1481.7
14h	Fmoc-K^{Fmoc}FPIK^{Mtt}LGK^{Mtt}AG	10	2135.3[b]	2134.2

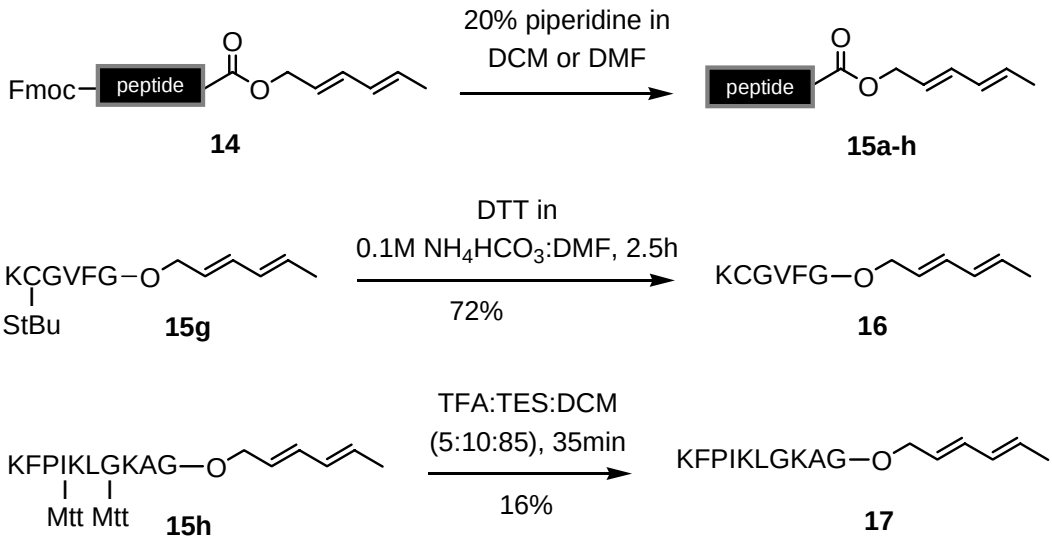
[a] MS measured by ESI-MS. [b] MS measured by MALDI-TOF. [c] During peptide chain assembly of **14g** one glycine residue coupling was not complete resulting in the formation of two products **14f** and **14g** in 12% and 4% yield, respectively.

Fmoc protecting groups were removed from compounds **14** with 20% piperidine in DCM or DMF (Scheme 10) to give, after HPLC purification and lyophilization, unprotected peptides **15a-h** in high purity (Table 4). Deprotection of the S*t*Bu group was accomplished by reduction of the disulfide bond of peptide **14e** with dithiothreitol (DTT) in ammonium bicarbonate medium for 2h, affording unprotected diene **16** (Scheme 10).

Surprisingly, in the case of compounds **15d** and **15h**, the Mtt group, which is known to be a very acid sensitive protecting group for amines,[99] could not be removed after treatment with 1%TFA/5% triethylsilane (or TIS) in DCM. While increasing the scavenger amount (up to 10%) seemed not to be relevant, the increment of TFA amount to 5% promoted partial removal of the Mtt group of peptide **15h**. However the extra acid addition resulted in considerable decomposition of the diene moiety. Therefore peptidyl diene **17** was isolated in poor yield (Scheme 10). No further conditions were investigated for optimal removal of this protecting group; for that reason peptide **15d** was subsequently employed having the Mtt-group at the lysine side-chain.

The synthesized peptides were characterized by mass spectroscopy and ^{1}H NMR analysis. The presence of the dienyl group at the C-terminus was validated for all peptide

hexadienyl esters by detection of characteristic olefin signals at 5.6, 5.8, 6.0 and 6.2 ppm and of the terminal methyl group at 1.7 ppm during proton-NMR experiments (Figure 19).



Scheme 10. Removal of the protecting groups of the dienyl peptides.

Table 4. Isolation yields after removal of the protecting groups.

Diene	Peptide sequence	Overall isolation yield (%)	ESI-MS $[M+H]^+$	
			found	*calculated*
15a	KPFLG	72	641.4[a]	641.4
15b	PC^{StBu}SMG	51	662.2[a]	662.3
15c	KLGFAG	83	671.4[a]	671.4
15d	KLGK^{Mtt}AG	73	909.6[b]	909.6
15e	KC^{StBu}GVFG	80	778.8[c]	778.4
15f	KFPIGLFG	91	958.7[a]	958.6
15g	KFPIGLGFG	92	1015.7[a]	1015.6
15h	KFPIK^{Mtt}LGK^{Mtt}AG	71	1651.6[a]	1651.0
16	KCGVFG	72	691.0[c]	690.4
17	KFPIKLGKAG	16	1138.9[a]	1138.7

MS measured by: [a] ESI-MS; [b] FAB-MS; [c] MALDI-TOF.

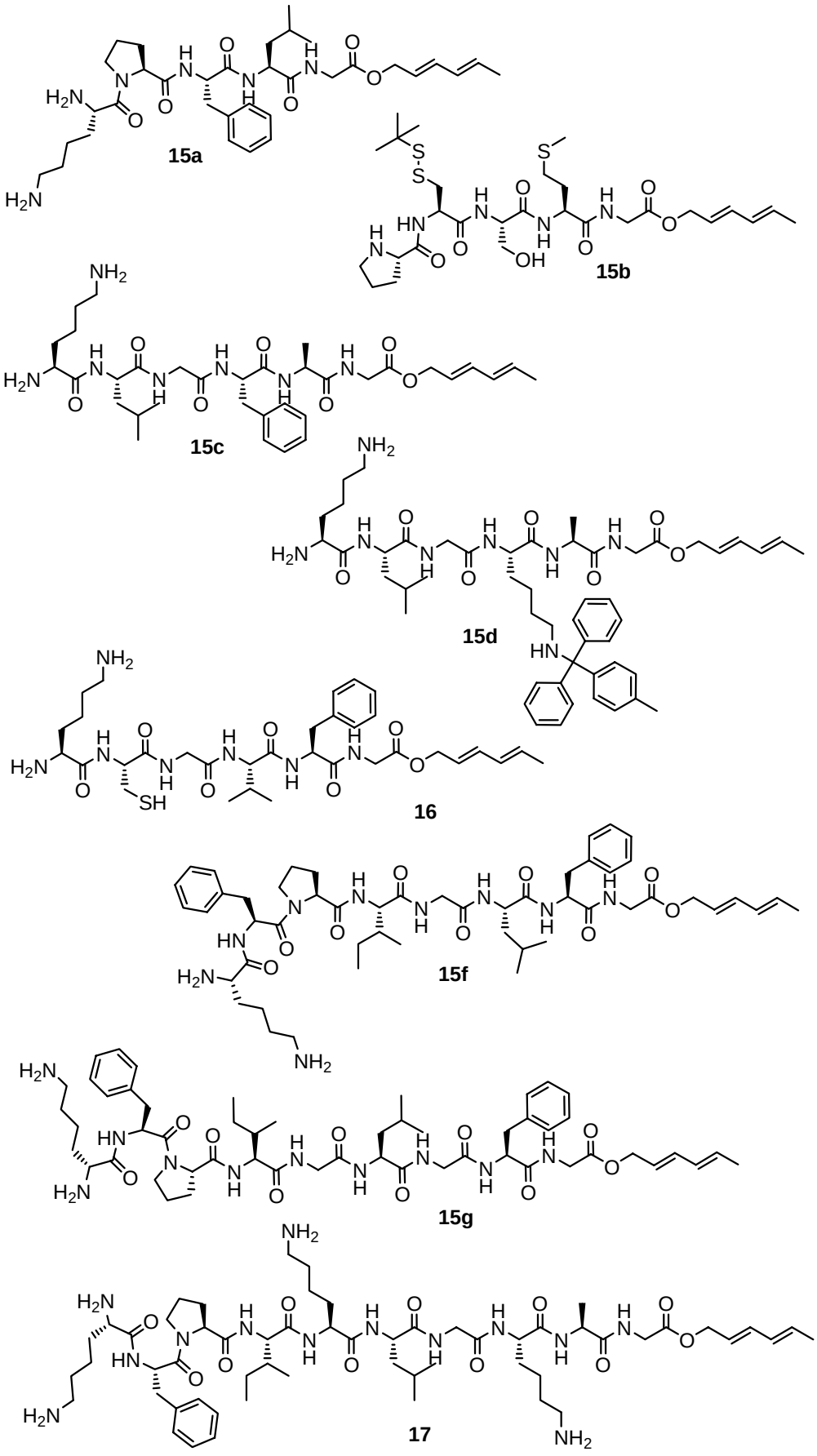
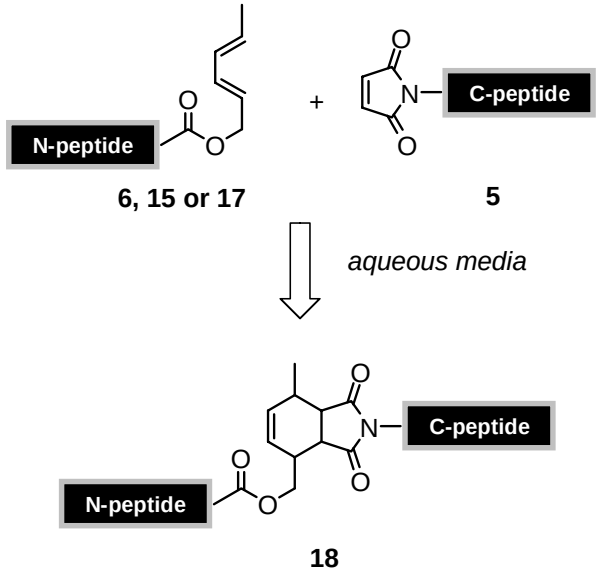


Figure 16. Overview of the final peptide hexadienyl esters structures.

4.1.4. Peptide Ligation via Diels-Alder Reaction

The Diels-Alder ligations of the diene and dienophile peptide segments to give the cycloadduct **18** (Scheme 11) were performed in aqueous solution at room temperature (Table 5). Diene and dienophile were mixed in equal amounts in most cases (usually at 10mM concentration) and allowed to react overnight. Eventually, methanol or DMF was added to help peptide solubilization in water.



Scheme 11. Diels-Alder ligation of dienyl and maleoyl-peptides.

Table 5. Diels-Alder ligation of dienyl and maleoyl-peptides.

Entry	Cyclo-product	Diene	Malei-mide	Solvent	Time	Ligation efficiency[a]
1	**18a**	**6**	**5a**[b]	H_2O:MeOH (10:3)	20h	~100%
2	**18b**	**15a**	**5a**	H_2O:MeOH (4:1)	24h	93%
3	**18c**	**15d**	**5b**	H_2O:MeOH (20:1)	24h	95%
4	**18d**	**15d**	**5c**	H_2O:DMF (4:1)	47h	84%
5	**18e**	**15b**	**5a**[c]	H_2O:MeOH (3:2)	24h	~100%
6	**18f**	**15f**	**5d**	H_2O	48h	93%
7	**18g**	**15g**	**5d**	H_2O	48h	92%
8	**18h**	**17**	**5d**	H_2O	48h	87%

[a] Based on the consumption of diene-peptide by analytical HPLC. Dienophile was added in excess: [b] 1.2 eq; [c] 2.4 eq.

A typical time course diagram for the Diels-Alder ligation of peptides **15** and **5** is given in Figure 17. After overnight reaction (ca 18h), ligation products were formed in 70-100% yield as revealed by HPLC analysis (except for entry 4, Table 5: at 18h, 52%). In some cases, consumption of the starting materials was mostly completed only after longer reaction time (entries 4, 6, 7, 8, Table 5). The use of DMF as co-solvent seems to slow down the rate of cycloaddition (entry 3 *vs* 4, Table 5). The utilization of an excess of the dienophile over the diene content considerably shortened the coupling time and led to total conversion of the hexadienyl peptide to the new cycloadduct (entries 1 and 5, Table 5).

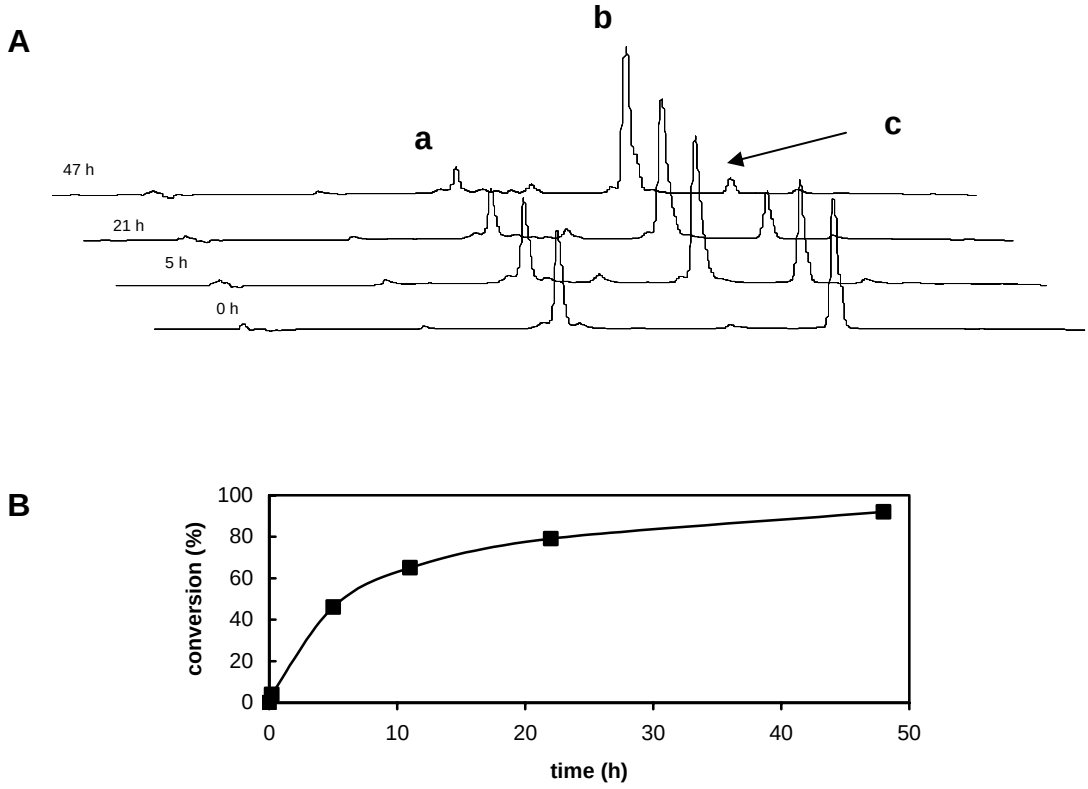


Figure 17. A. The time course for the ligation of peptides **15f** and **5d** followed by HPLC traces (entry 6, Table 5). **a =** maleimido-peptide **5d**. **b =** cycloadduct **18f**. **c =** diene-peptide **15g**. **B**. Plot of ligation efficiency versus time for reaction of peptide **15h** and **5d**.

All ligation products were isolated by HPLC purification and identified by mass spectroscopy (Figure 18). Most of the losses in yield (in comparison with the observed ligation efficiency) arose simply from HPLC recoveries (Table 6). The new ligated cycloadducts were clearly characterized by analysis of their ^{1}H NMR spectra: the most striking feature was the replacement of the typical diene four-vinyl-proton multiplets at 5.6-6.1 ppm and of the sharp maleimido-olefin singlet peak at ca 6.9 ppm by a broadened two-proton

signal at 5.7 ppm (Figure 19). In addition, new peaks appeared at 2.5, 2.7, 3.1 and 3.3 ppm which were attributed to the protons connected to the novel C-C bonds (H3, H4, H5 and H6 in Figure 19). Also the terminal methyl doublet of the diene group at 1.7 ppm was shifted to 1.4 ppm for all cycloadducts.

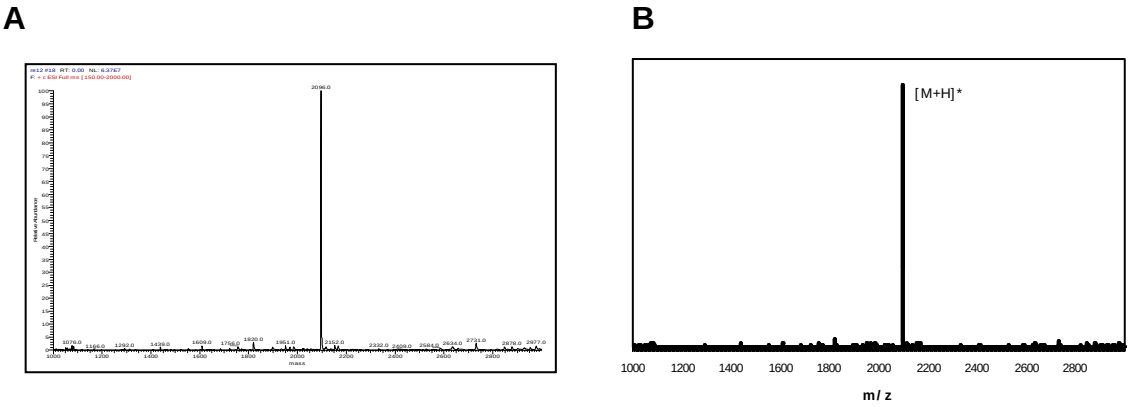


Figure 18. ESI-MS spectrum of ligated cycloproducts: **A**. **18f:** 2096.0 $[M+H]^+$ (calculated 2096.0); **B**. **18g:** 2153.0 $[M+H]^+$ (calculated 2153.1).

Table 6. Recoveries from ligated peptides after HPLC purification.

Cyclo product	Sequence	Yield (%)	MS $[M+H]^+$	
			found	*calculated*
18a	VAG-cyclo-GYTG	87	802.4[a]	802.4
18b	KPFLG-cyclo-GYTG	60	1117.6[a]	1117.6
18c	KLGK(Mtt)AG-cyclo-GTQFHG	74	1634.6/1378.7[a,c]	1634.7/1378.7
18d	KLGK(Mtt)AG-cyclo-GSEWIG	64	1636.6/1380.8[a,d]	1636.7/1380.7
18e	PC^{StBu}SMG-cyclo-GYTG	32	1138.4[a]	1138.4
18f	KFPIGLFG-cyclo-GAKTSAESYSG	69	2096.0[a]	2096.0
18g	KFPIGLGFG-cyclo-GAKTSAESYSG	67	2153.0[a]	2153.1
18h	KFPIKLGKAG-cyclo-GAKTSAESYSG	50	2277[b]	2276

Found mass peak for the cycloadducts measured by: [a] ESI-MS or [b] MALDI-TOF. The Mtt group was partially detached during reaction course and purification step, giving two final products: [c] **18c** (with Mtt, 43%) and **18c´** (without Mtt, 31%); [d] **18d** (with Mtt, 40%) and **18d´** (without Mtt, 24%).

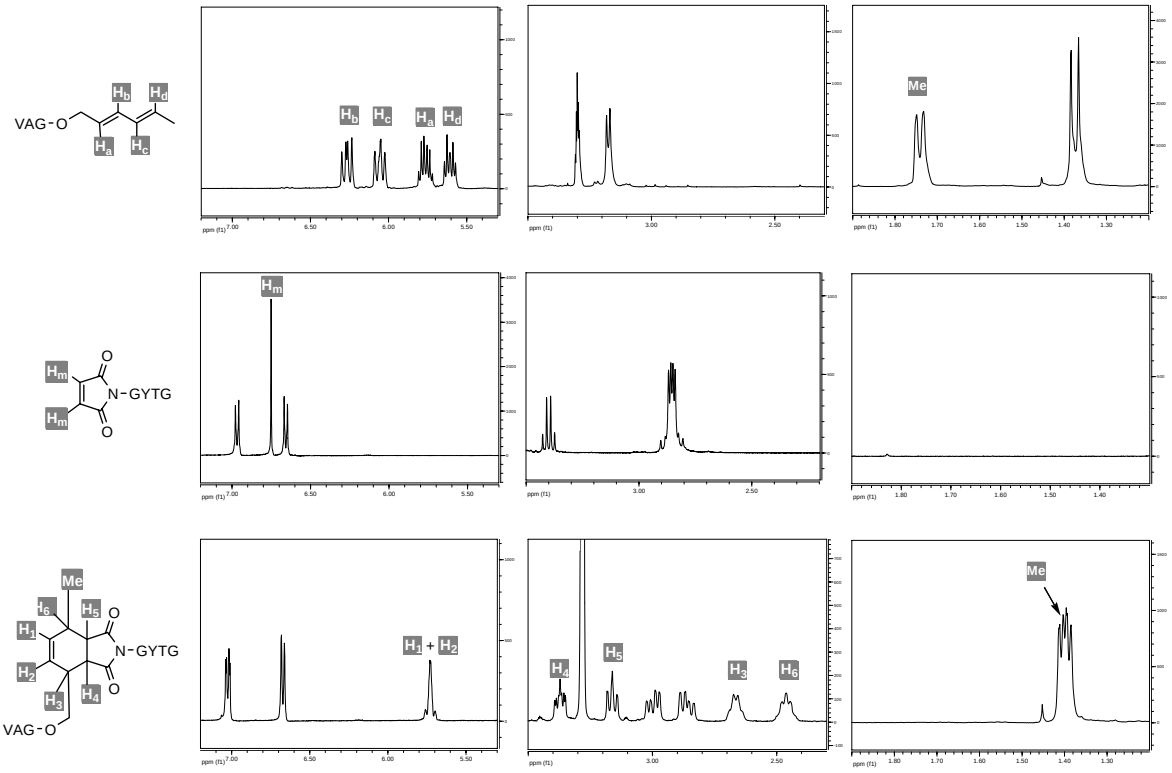
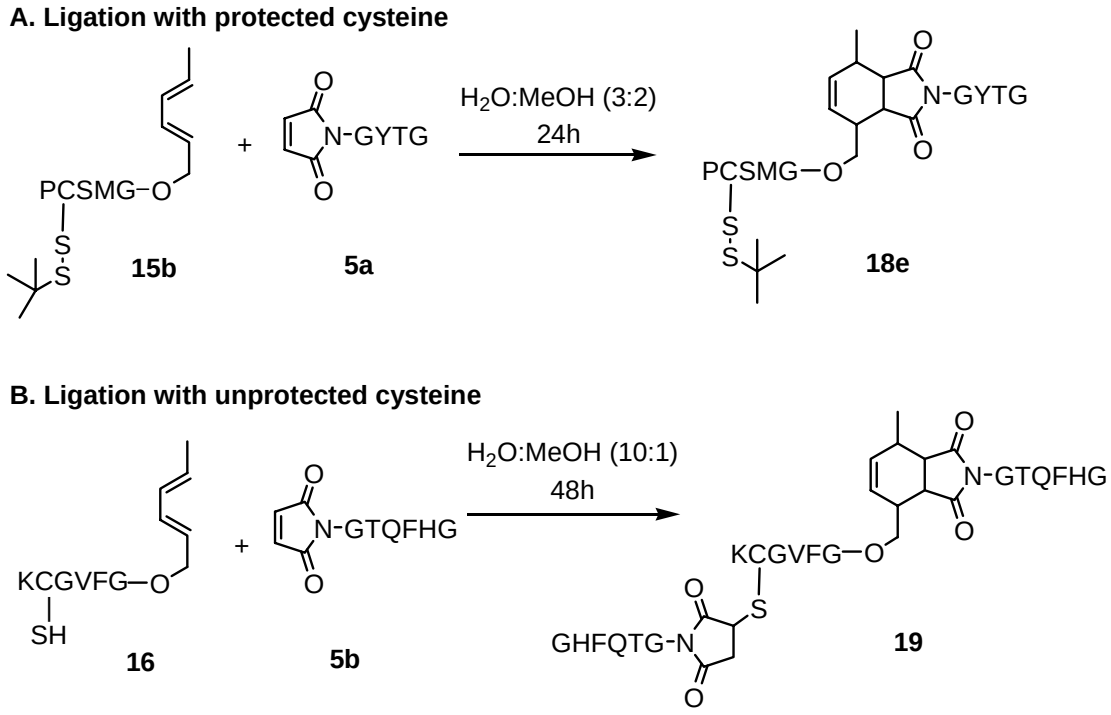


Figure 19. Comparison of the H-NMR spectra of peptidyl diene **9**, dienophile **5a** and cycloproduct **18a**.

These results revealed that the Diels-Alder ligation is chemoselective and compatible with reactive amino acids as lysine, histidine and tryptophan. Also no nucleophilic addition of the N-terminal amino group to the dienophile α,β-unsaturated double bond was observed under these conditions. However, due to the rather electrophilic nature of the maleimido moiety, we recognized the potentially troublesome reaction which can take place between this structure and the highly nucleophilic sulfhydryl group of cysteines. Michael addition to maleimide double bond by the mercaptan group of cysteines is a well established linkage method applied for bioconjugation of peptides and proteins.[13] Indeed, when peptide **16**, which possesses a free cysteine residue, was mixed with maleimido-peptide **5b**, the formation of product **19** resulting from both nucleophilic addition and cycloaddition reactions was detected (Scheme 12). To avoid this side-reaction, protection of the cysteine side-chain during Diels-Alder ligation is necessary, as illustrated for the combination of peptides **15b** and **5a** (entry 5, Table 5), where the sulfhydryl moiety is temporarily blocked by disulfide formation with a S*t*Bu group.



Scheme 12. Diels-Alder ligation involving peptides holding blocked and free cysteine residues.

4.1.5. Stereochemistry of the Diels-Alder ligation

The Diels-Alder reaction between an acyclic diene like **1** and maleimide normally gives the *endo* adduct as described earlier (Figure 20).[100-101]

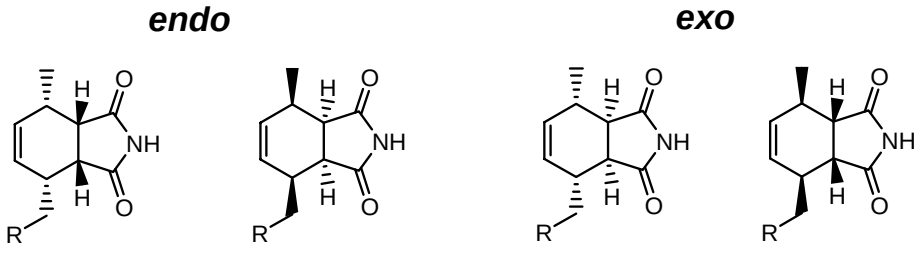
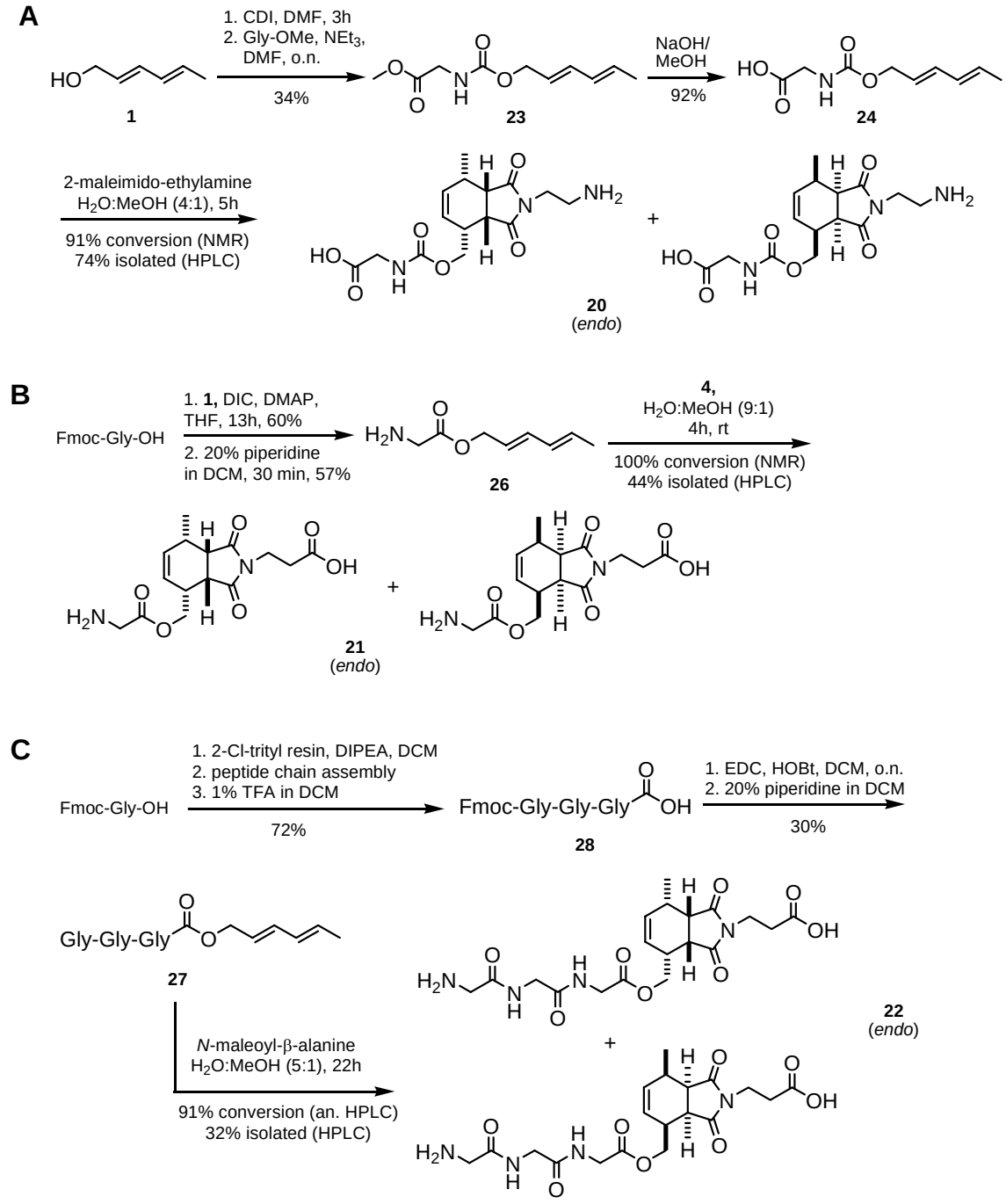


Figure 20. Structure of the *endo* and *exo* adducts (and the pair of enantiomers of each conformation) which can result from the cycloaddition between maleimide and diene **1**. The *endo* conformation is the favored one.

To check if the cycloaddition of peptidyl hexadienes and maleimides would also exhibit the expected selectivity, the proton NMR spectra of a set of peptide-like simple cycloadducts **20**-**22** and of the smallest ligation products **18a** and **18b** were examined. Spectroscopic and chromatographic analysis of the other ligation adducts did not permit a conclusive determination of the stereoselectivety.

The synthesis of the small cycloadducts is outlined in Scheme 13.



Scheme 13. Preparation of Diels-Alder adducts **20**, **21** and **22**. The cycloaddition yielded a pair of the *endo*-enantiomers.

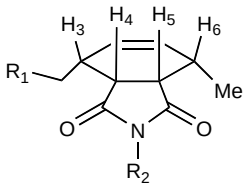
Conversion of hexadienol **1** via CDI activation to an intermediate imidazolyl carbamate that in turn reacts with glycine methylester in basic media afforded 2,4-hexadienyl-(methoxycarbonyl)-methylcarbamate **23**, which is then transformed into *N*-glycyl-2,4-hexadienyl carbamate **24** by saponification. This compound was further combined with 2-(maleimido)-ethanamine (see page 62 for synthesis of this building block) in H_2O:MeOH (4:1) producing the Diels-Alder adduct **20** (Scheme 13A). Cycloadduct **21** and **22** were obtained from reaction of glycine- and triglycine-hexadienyl ester **26** and **27**, respectively, with 3-maleimidopropionic acid in aqueous solution (Scheme 13B and 13C). Dienyl esters **26** and **27** were synthesized from esterification of precursors Fmoc-glycine and Fmoc-triglycine **28**, respectively.

As pointed out in Table 7, compounds **20, 21** and **22** showed clear chemical shifts for all protons belonging to the cyclohexene framework. Moreover the measured coupling constants are in good concordance with literature values (compounds **29** and **30**),[100] indicating the formation of the *endo* product (≥ 98% based on NMR) as expected. Additional NOE experiments supported the proposed stereochemistry (Figure 21).

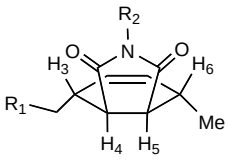
Table 7. ^{1}H-NMR data from compounds **20**, **21**, **22**, **18a** and **18b** in CD_3OD (experimental) and from **29** and **30** (literature).[101]

Cycloadduct:	**20**	**21**	**22**	**18a**	**18b**	**29**	**30**
Configuration (ratio)	*endo*	*endo*	*endo*	*endo I*/*endo II* (95:5)	*endo I*/*endo II* (50:50)	*endo*	*endo*
Chemical shift δ(ppm)							
H1 + H2	5.8	5.7	5.7	5.7	5.7		
H3	2.7	2.7	2.7	2.7	2.7		
H4	3.4	3.3	3.3	3.4/3.5	3.4/3.5		
H5	3.2	3.1	3.1	3.2/3.1	3.2/3.1		
H6	2.5	2.5	2.5	2.5	2.5		
Me	1.4	1.4	1.4	1.4	1.39/1.40		
Coupling constant *J*(Hz)							
J(H5,H6)	7.0	6.9	7.0	7.2/m[a]	7.2/7.3	6.9	6.6
J(H4,H5)	8.4	8.4	8.4	8.5/m[a]	8.4/8.5	8.4	8.4
J(H3,H4)	6.2	6.2	6.2	6.2/m[a]	6.2/6.3	6.0	-

[a] The observed multiplet signal was too weak for an appropriate *J* constant determination.



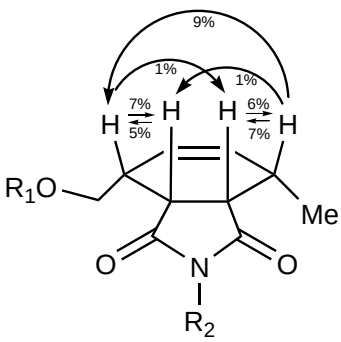
endo



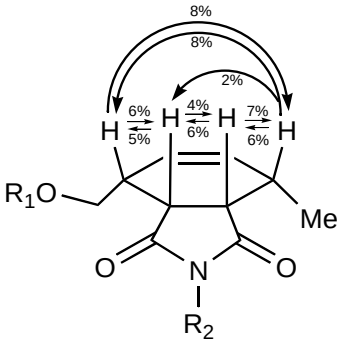
exo

29: R_1= -CH_2(2-PyMe_2Si); R_2 = Et
30: R_1= -CH_2OH; R_2 = Et

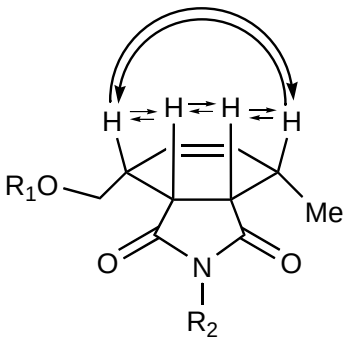
NOE signals for ligated peptides:



20

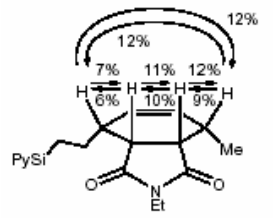


21

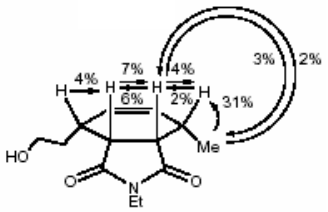


18a

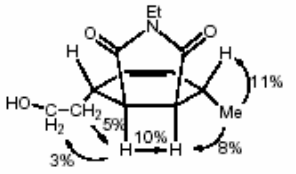
NOE (literature):



29



30 (*endo*)



30 (*exo*)

Figure 21. NOE signals enhancements observed for compounds **20**, **21** and **18a** in CD_3OD (experimental) and for **29** and **30** (literature).[100] The NOE spectrum of **18a** did not give strong signals for all proton-proton interactions. Thus the mentioned peak was observed but, due to its weakness (~1%), not quantified.

The Diels-Alder ligation process to give the peptides **18a** and **18b** led to the formation of the expected *endo*-products (Figure 22) as confirmed by comparison of their NMR values with the ones found for the model cycloadducts **20**-**22** (Table 7). Surprisingly, the cycloadduct **18a** was obtained as a single *endo*-stereoisomer (95%) whereas the cycloadduct **18b** was

obtained as a mixture of the two *endo*-isomers (50:50). These results suggest that a certain hydrogen bond pattern between the two reacting polypeptide chains may fix the diene and dienophile groups in a specific position that favors the *endo*-attack of only one of the faces.

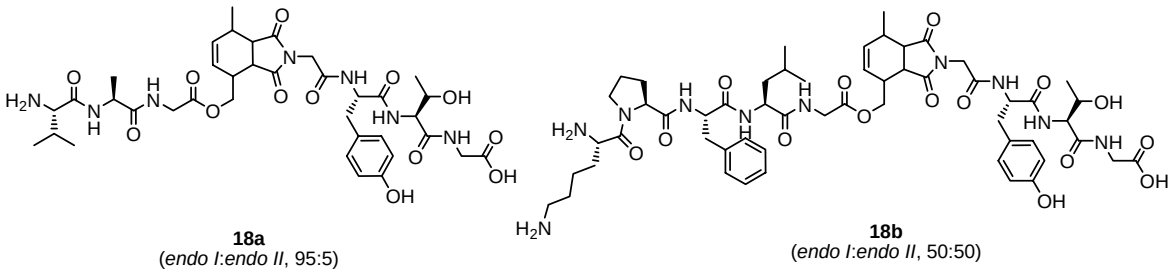
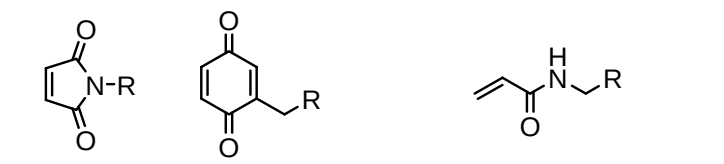
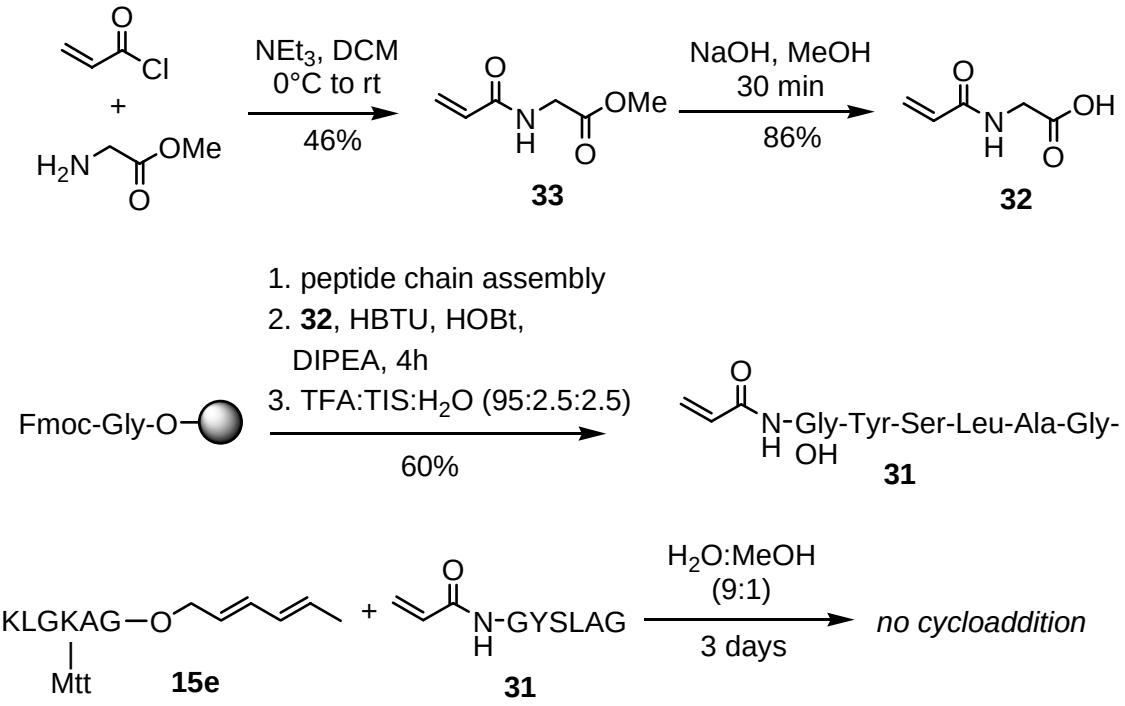


Figure 22. Structure of cycloproducts **18a** and **18b** and *endo* enantiomeric ratio.

4.1.6. Other diene and dienophile functionalities

In further experiments, the Diels-Alder peptide ligation was investigated using different dienophile building blocks. So far only maleimide and benzoquinone have been utilized as dienophiles for Diels-Alder assisted biomolecule modifications. [85-93] In order to determine the influence of dienophile double bond activation on the rate of cycloaddition, the N-terminal acryloyl peptide **31** was prepared by solid-phase synthesis following the same procedure as described for the synthesis of **5**, but replacing the maleoyl-glycine part for *N*-acryloylglycine **32** (Scheme 14). This acrylamide was obtained from reaction of glycine methyl ester with acryloyl chloride and subsequent C-terminal deprotection by saponification. The *N*-acryloylpeptide **31** was then combined with dienyl-peptide **15d** in aqueous solution; however no cycloadduct was formed even after prolongated reaction time. Consequently the presence of two activating groups in the dienophile structure seems to be necessary for effective Diels-Alder ligation to occur.





Scheme 14. Reactions involving *N*-acryloylpeptides.

Diverse other diene groups can also be regarded as reaction partners for the Diels-Alder ligation (Figure 23). Previous studies demonstrated comparable Diels-Alder reaction rate for biomolecules equipped with a hexadiene having acyclic **34** or cyclic configuration **35**, as well as enclosing internal **36** or terminal **34** diene moieties.[85,89] Likewise, application of anthracene **37** has been reported for construction of RNA conjugates.[102] Peptides incorporated with a cyclopentadiene moiety (**38**) presumably undergo rather faster cycloaddition, and this scaffold has been successfully employed for the preparation of peptide chips as described by Mrksich *et al.*[91-93]

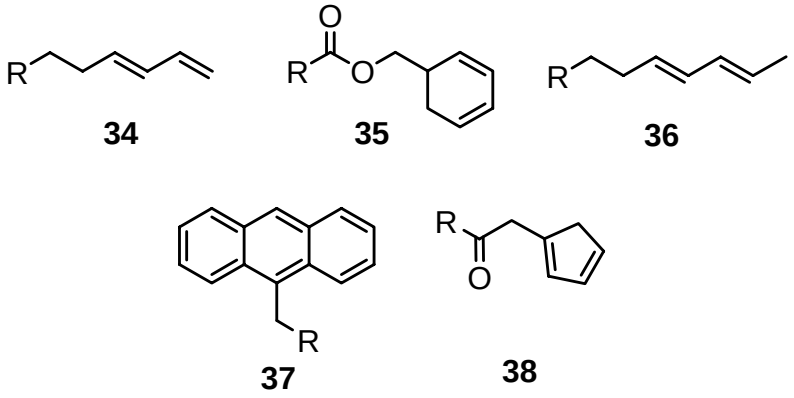
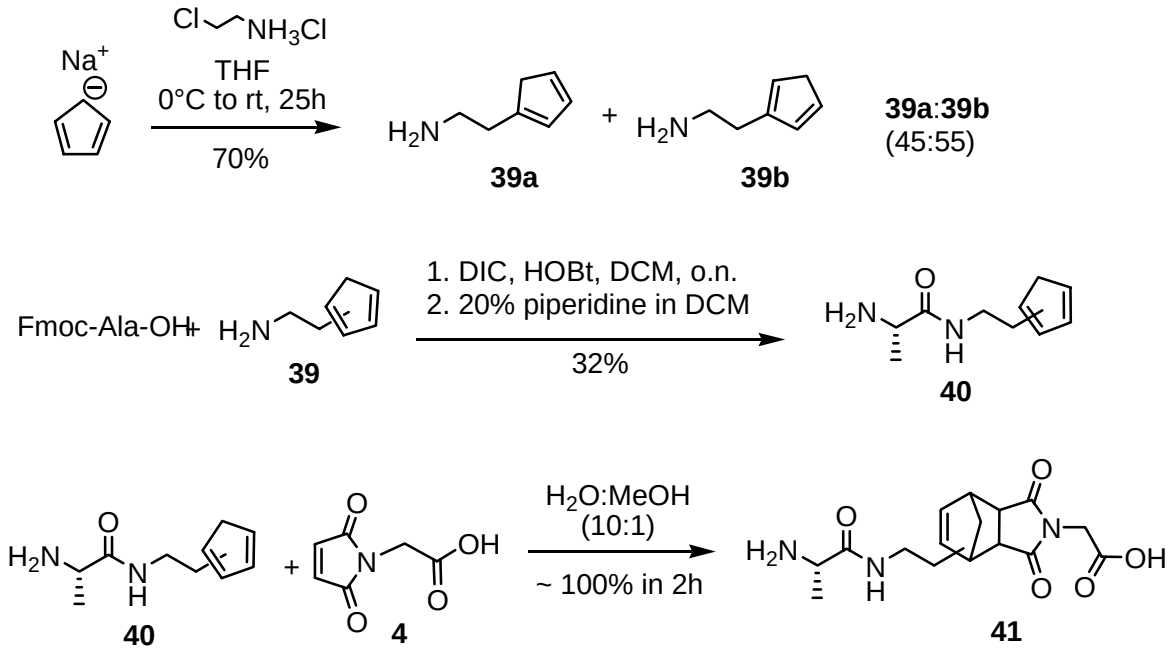


Figure 23. Structure of various dienes which undergo Diels-Alder with biomolecules.

Due to the great reactivity of cyclopentadiene towards cycloaddition, we were interested in exploring the peptide ligation using this diene group. For this purpose, the precursor cyclopentadienylethylamine **39** was synthesized from sodium cyclopentadienide as indicated in Scheme 15. At room temperature, the substituted cyclopentadiene underwent hydrogen migration and a mixture of rearranged compounds **39a** and **39b** was found.[103] Amine **39** was then coupled to the C-terminus of Fmoc-alanine and the resulting amide was deprotected with piperidine to generate the cyclopentadienyl modified peptide **40**. Afterwards the cyclopentadiene **40** was combined with maleoyl-glycine in aqueous media to give cycloadduct **41** within 2h.

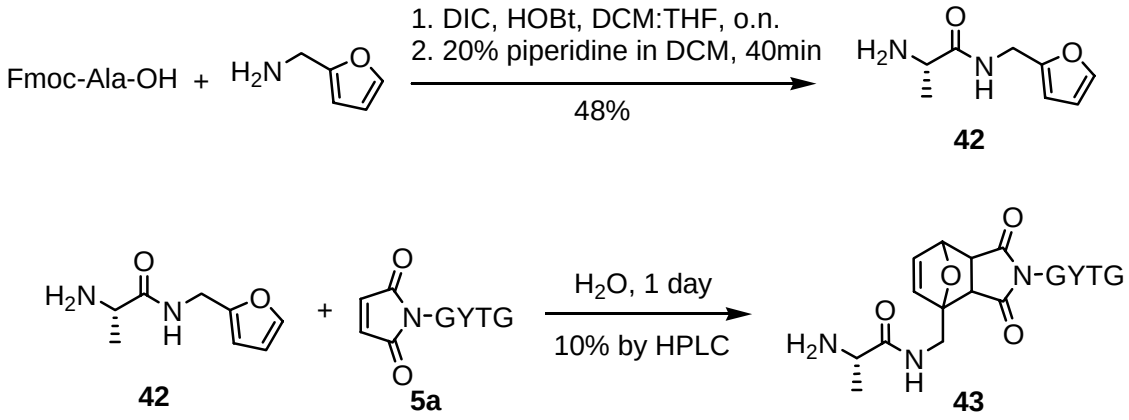


Scheme 15. Reactions with cyclopentadienyl-peptide.

Attempts to modify longer peptide sequences with the cyclopentadienyl precursor either by solution- or solid-phase synthesis proved to be a difficult task. The mayor problem was the facile dimerization of the cyclopentadienyl moiety under reaction conditions and peptide purification steps (HPLC). The Diels-Alder cycloaddition of cyclopentadiene to itself is actually accelerated in aqueous media[104] and, even though this group has the advantage of being quite more reactive, its instability is a considerable drawback for the utilization of this functionality for the Diels-Alder ligation. Nevertheless, for some particular applications where an excess of the diene-component is used over the amount of the dienophile during a ligation reaction, the level of dimerization may not be prejudicial for the efficiency of the Diels-Alder

ligation process and therefore the cyclopentadienyl group may be successfully applied. This is the case for the immobilization of biomolecules on glass slides, where a dienyl-derived substrate is spotted on a dienophile-functionalized surface.[92,93] This issue will be further discussed in Section 4.3.

A peptide possessing a furane group at the C-terminus has also been synthesized (Scheme 16). This functionality however was less reactive towards Diels-Alder cycloaddition using similar conditions applied for the acyclic diene/maleimide ligation system.



Scheme 16. Reactions with furane-peptide.

4.2. Protein Ligation by Diels-Alder Reaction

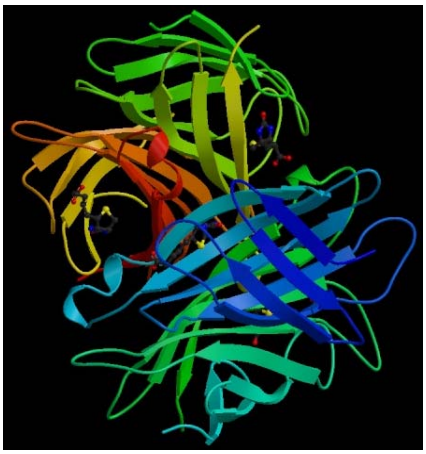
The stability of the hexadiene function in aqueous solutions and its compatibility with all functional groups present in coded amino acids opens up the opportunity to combine the Diels-Alder ligation method with other conjugation techniques to generate tailor-made proteins. In such a combined strategy the protein of interest is initially functionalized with a diene unit and then the resulting protein-derived diene can be further functionalized by Diels-Alder reactions with different dienophiles under very mild conditions. If the protein possesses reactive cysteine residues, the ligation conditions may be manipulated to avoid unspecific reactions between the dienophile unit and the mercapto group by temporary blocking of the sulfhydryl groups.

Three different approaches were explored for the incorporation of the diene functionality into proteins:

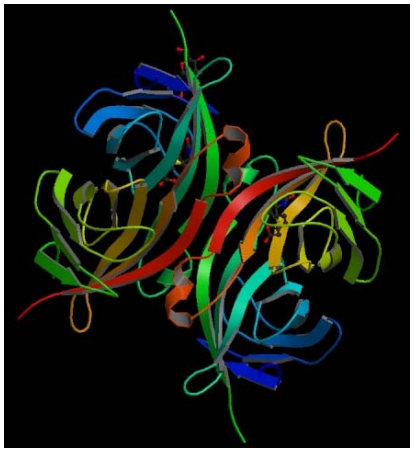
- *Method 1*: Labeling of a protein-ligand complex
- *Method 2*: Selective bioconjugation using a heterobifunctional cross-linker
- *Method 3*: Site-specific functionalization using Expressed Protein Ligation

4.2.1. Labeling of a protein-ligand complex

The scope of the Diels-Alder reaction for protein ligation employing the 1,3-hexadiene/maleimide method was initially investigated employing the biotin-binding tetrameric proteins streptavidin and avidin as model systems. These proteins were chosen as first model because they contain no free cysteines in their structure. Streptavidin has no cysteine at all in its sequence[105] and the cysteine residues found in avidin form disulfide bonds and are not reactive[106] (Figure 24). Both proteins have been applied extensively as probes in immunoassays and labeling of antibodies, enzymes and other molecules of interest.[13] Even though they bind biotin similarly, their primary structure differs considerably. Streptavidin is a 55 KDa protein isolated from bacteria *Streptomyces avidinii*, avidin is a 66 KDa glycoprotein found in egg white and tissues of birds, reptiles and amphibians.



Streptavidin



Avidin

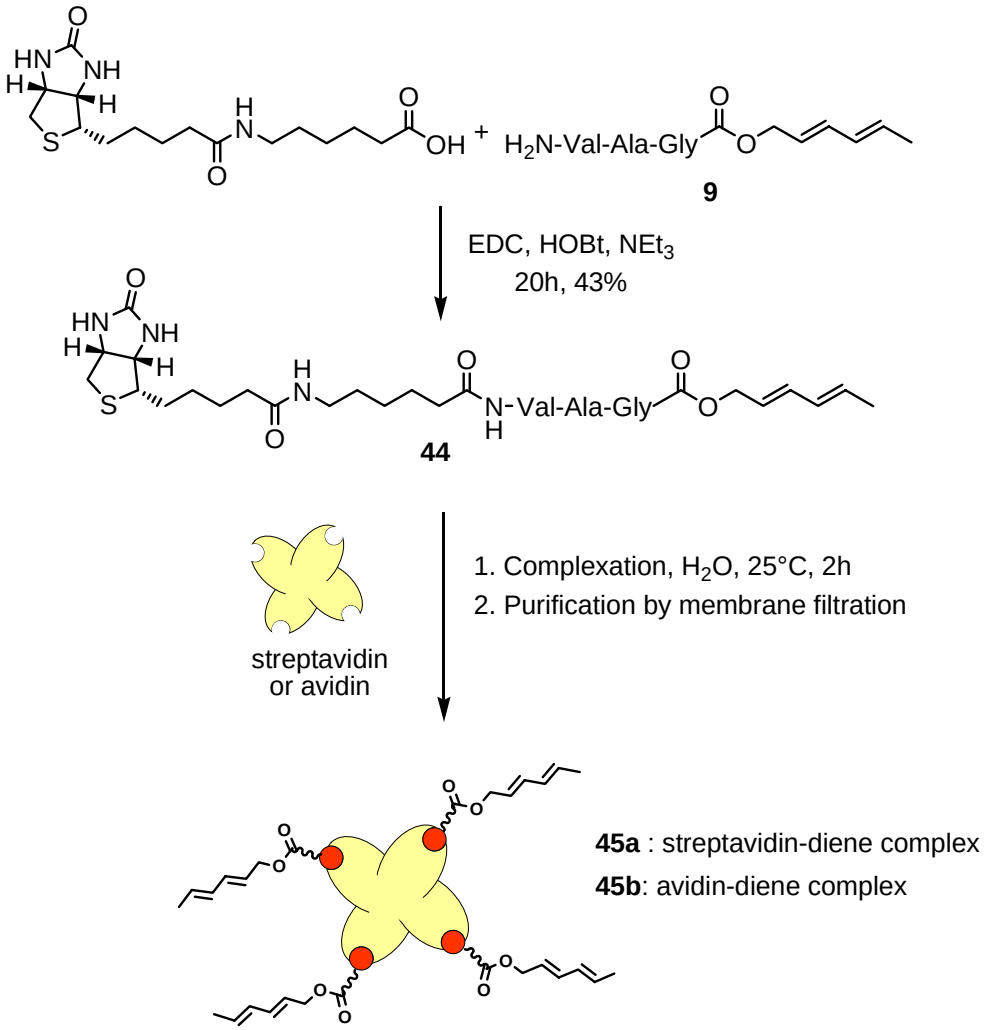
Streptavidin subunit sequence (127 AA):
AEAGITGTWYNQLGSTFIVTAGADGALTGTYESAVGNAESRYVLTGRYDSAPATDGSGTA
LGWTVAWKNNYRNAHSATTWSGQYVGGAEARINTQWLLTSGTTEANAWKSTLVGHDTFTK
VKPSAAS

Avidin subunit sequence (128 AA):
ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYTTAVTATSNEIKESPLHGTENTINKRT
QPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIF
TRLRTQKE

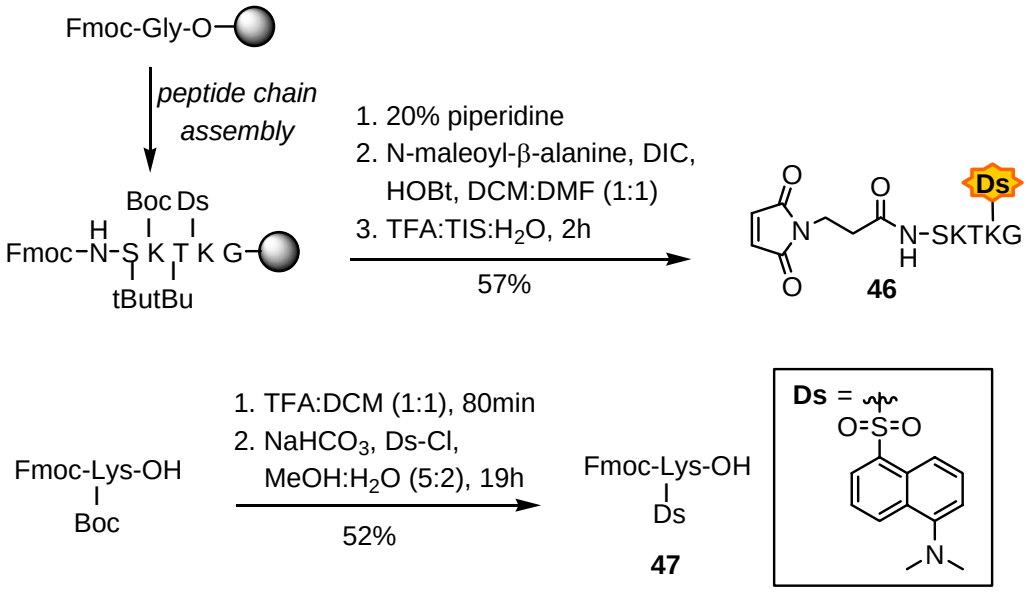
Figure 24. Structure details for streptavidin[105] and avidin.[106] The pictures show the proteins in complex with biotin.

In this first protein ligation study, the high affinity of streptavidin and avidin for biotin - which is the strongest known noncovalent biological interaction (K_d ~ 10^{-15} M) – was exploited for the preparation of a stable protein-diene complex. To this end, biotinylated hexadienyl-compound **44** was synthesized via EDC/HOBt-assisted coupling of biotin-6-amino-caproic acid and dienyl tripeptide **9** (Scheme 17). The biotinylated diene was then bound to streptavidin and avidin. A 5-fold excess of **6** over the biotin specific binding activity of the proteins was used to ensure occupation of all biotin-binding sites. Excess of this compound was posterior removed by diafiltration with a centrifugal membrane device to ensure that all diene-peptide present in the solution was linked to the protein.

Diels-Alder ligation using this protein-diene complex was performed applying the fluorescently labeled peptide **46** and maleimido-peptide **5d**. The first compound was synthesized in solid phase by the same method as for **5d** but utilizing maleimido-propionic acid for the insertion of the dienophile moiety (Scheme 18). The fluorescent dansyl group was introduced as a building block connected to the lysine residue as Fmoc-Lys(Dansyl)-OH **47**.

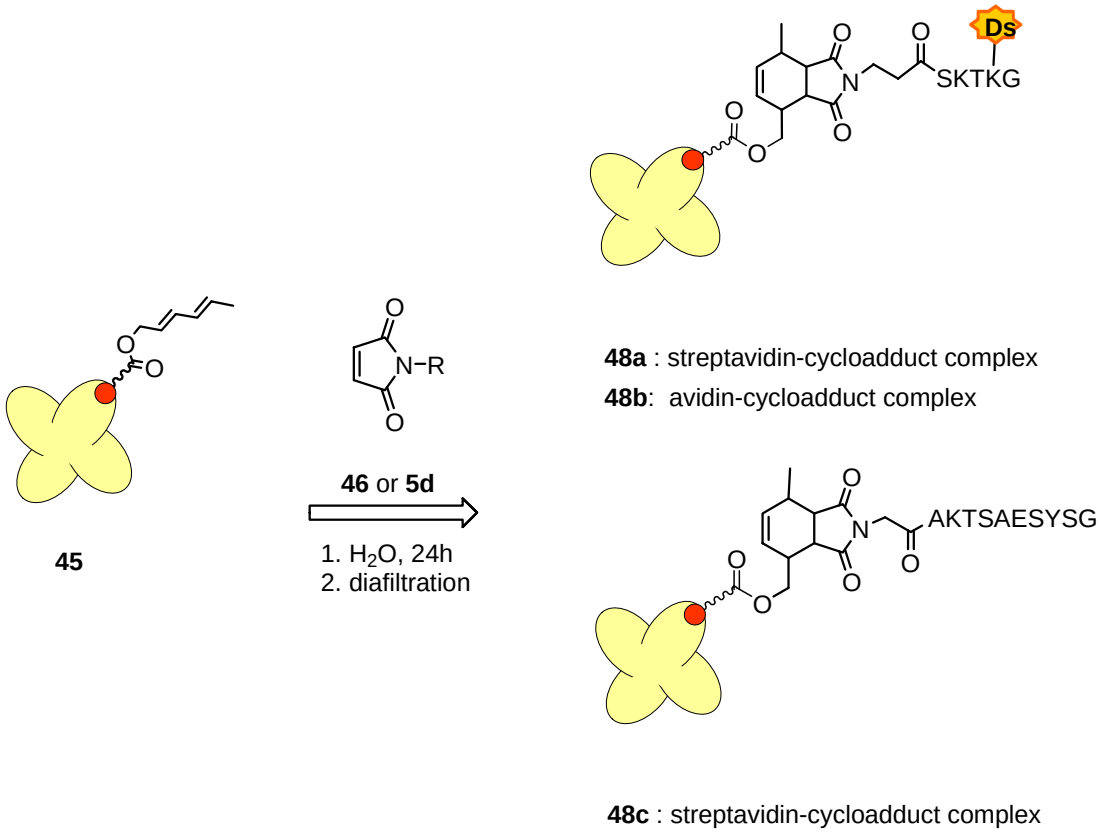


Scheme 17. Synthesis of the biotinylated peptide hexadienyl ester and complexation with streptavidin and avidin.



Scheme 18. Synthesis of the fluorescent labeled maleimido-peptide **46**.

The protein-hexadiene complex **45** at a concentration of 3.5 mg/mL (ca. 65µM of **45a**) or 8 mg/mL (ca. 120µM of **45b**) was treated with maleimido-peptides **46** and **5d** (50-fold dienophile relative to the estimated diene content) in water at room temperature during 1 day (Scheme 19). Subsequently unligated dienophile was removed by membrane ultracentrifugation. As control experiment, streptavidin/avidin was incubated with the same amount of dienophile and submitted to same procedure.



Scheme 19. Diels-Alder ligation with protein-diene complexes **45a** and **45b**. Only one binding site of the tetrameric protein is shown.

Concerning the case where the protein ligation was performed with labeled dienophile **46**, reaction vials (ligation and control) were scanned for fluorescence at 302nm. As shown in Figure 25A, only the solution involving the combination of streptavidin-diene complex and dienophile became fluorescent. The Diels-Alder ligation product was also detected by carrying out the reaction with the avidin complex **45b** and different amounts of dienophile **46** (Figure 25B). Further analyses were performed for all ligation reactions where the solution of complex **48** was heated until boiling in the presence of isopropanol in order to denature the protein, break the complex and release the biotinylated cycloadduct, whose mass was confirmed by

MALDI-TOF (Figure 26). Without denaturation of the protein, the ligation product could not be easily detected by mass spectrometry.

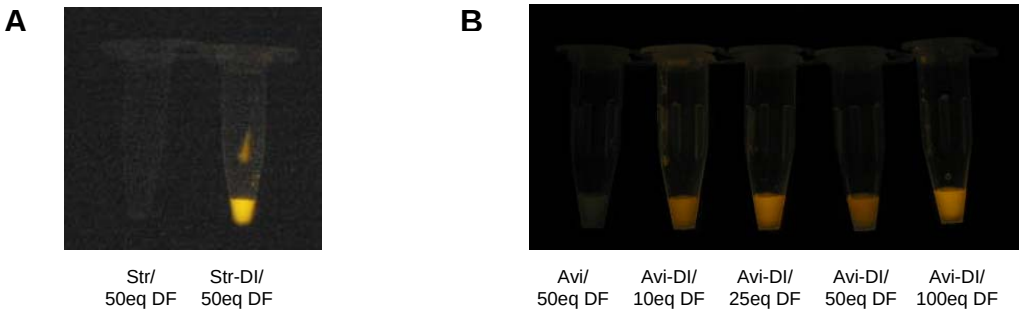


Figure 25. Fluorescence scan for the ligation reactions with peptide **46** (DF). **A.** With streptavidin (Str) and streptavidin-diene complex **45a** (Str-DI). **B.** With avidin (Avi) and avidin-diene complex **45b** (Avi-DI).

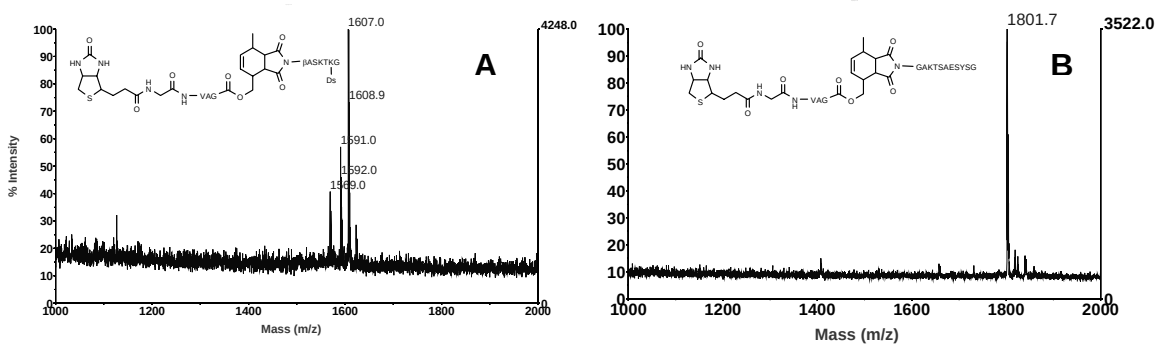


Figure 26. After denaturation, the expected mass for the Diels-Alder ligation cycloadduct between biotinylated diene **44** and dienophiles was observed: **A**. **44** and **47** ($C_{72}H_{109}N_{15}O_{20}S_2$, MW: 1568.85) 1569.0 $[M+H]^+$ (calc. 1568.7), 1591.0 $[M+Na]^+$ (calc. 1590.7), 1607.0 $[M+K]^+$ (calc. 1606.7)); **B**. **44** and **5d** ($C_{79}H_{120}N_{18}O_{28}S$, MW: 1801.97) 1801.7 $[M+H]^+$ (calc. 1801.8). No significant mass peaks were found for the biotinylated diene.

In additional experiments, the ligation involving peptide **46** was monitored by native polyacrylamide electrophoresis (Figure 27). Both fluorescent streptavidin- and avidin complex band were detected by ilumination of 302nm UV light upon treatment of **45a** or **45b** with **46**. Here again no reaction was observed in the control tests where the two proteins were

combined with the maleimide **46** (lane 3, Figure 27A and 27B)**,** thus the observed fluorescence of the complexes **48a** and **48b** was attributed to the Diels-Alder reaction which took place between the hexadiene group of the biotinylated compound and the dansyl-labeled maleimido-peptide.

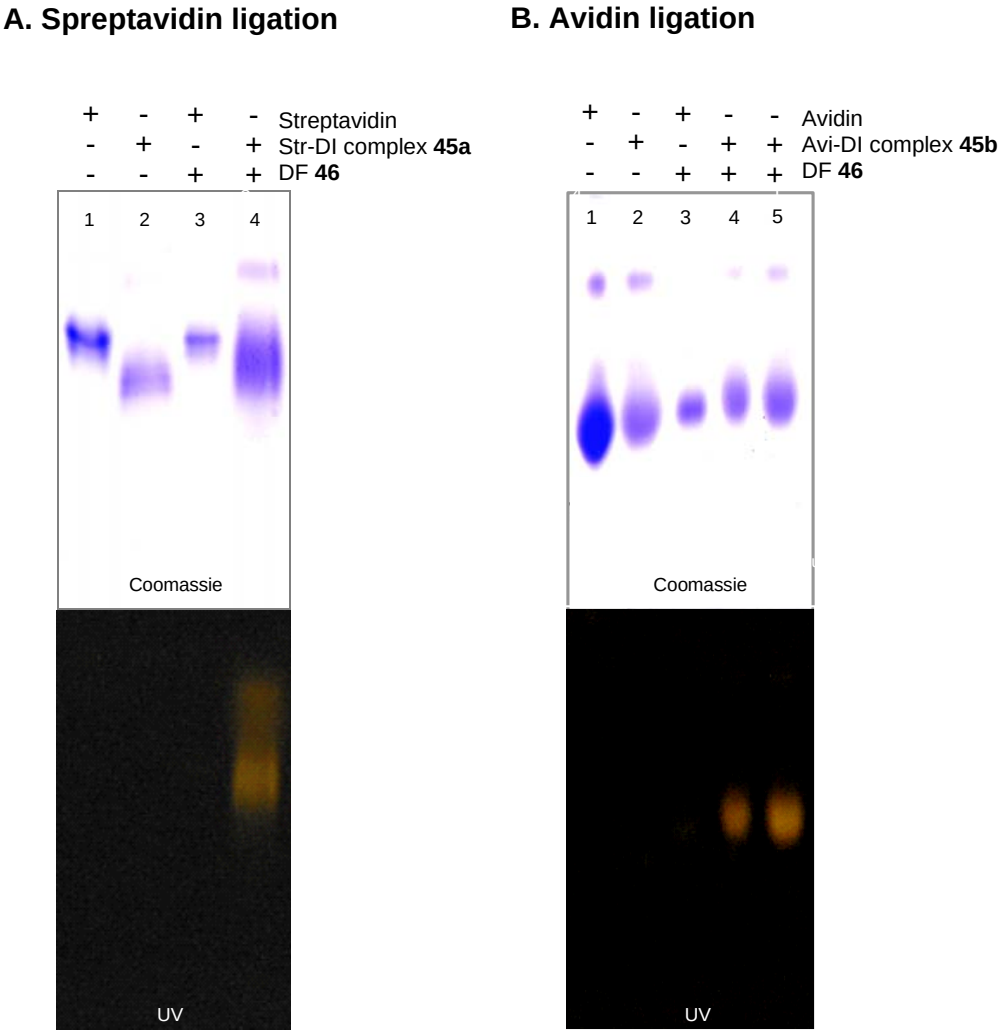
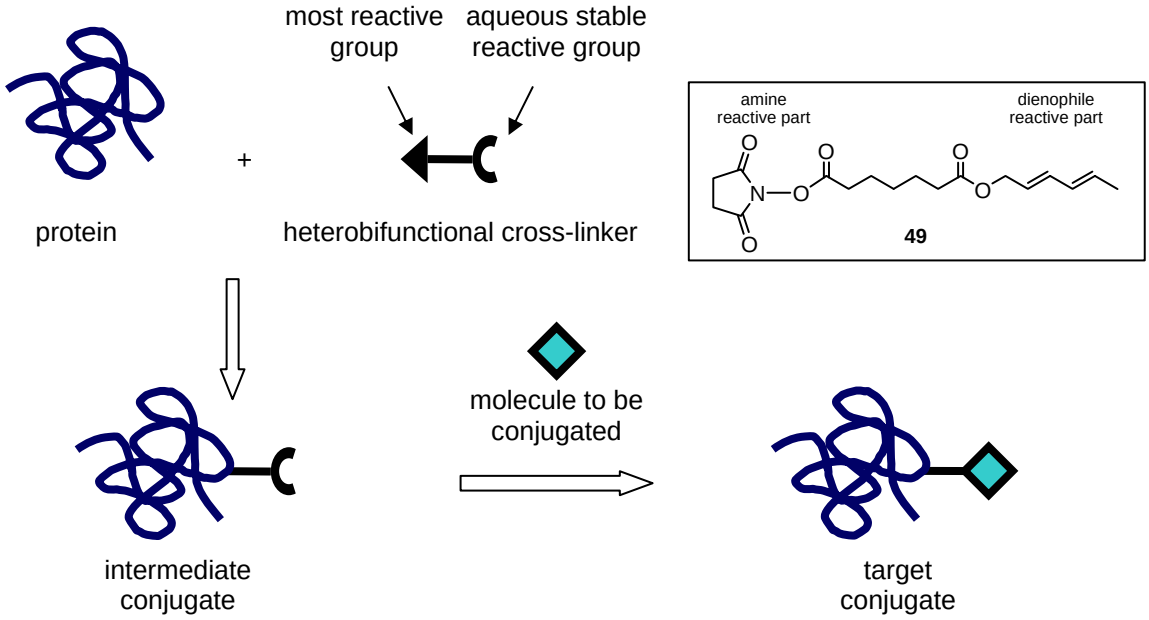


Figure 27. Analysis of the Diels-Alder ligation involving streptavidin and avidin complexes. **A.** Basic native PAGE of the Diels-Alder model using streptavidin. **B.** Acidic native PAGE of the Diels-Alder model using avidin.

4.2.2. Selective bioconjugation mediated by Diels-Alder Ligation

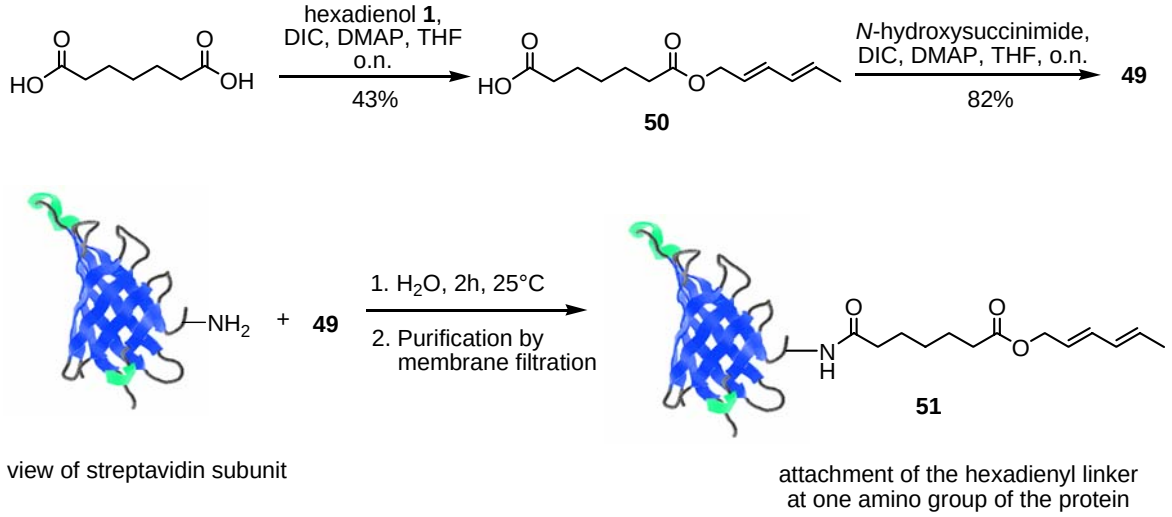
To introduce the Diels-Alder ligation as a method for protein bioconjugation, the heterobifunctional cross-linker **49**, which encloses a 1,3-hexadiene functionality and an amine-reactive part, was devised. Heterobifunctional conjugation reagents are used to cross-link proteins and other molecules in a two-step procedure (Scheme 20).[13] They contain one reactive group that displays extended stability in aqueous environments (in our case, the diene group), therefore allowing purification of an activated intermediate before adding the second molecule to be conjugate. The result is the ability to direct the cross-linking reaction to selected parts of target molecules and better control over the conjugation process.



Scheme 20. Bioconjugation strategy using a heterobifunctional cross-linker.

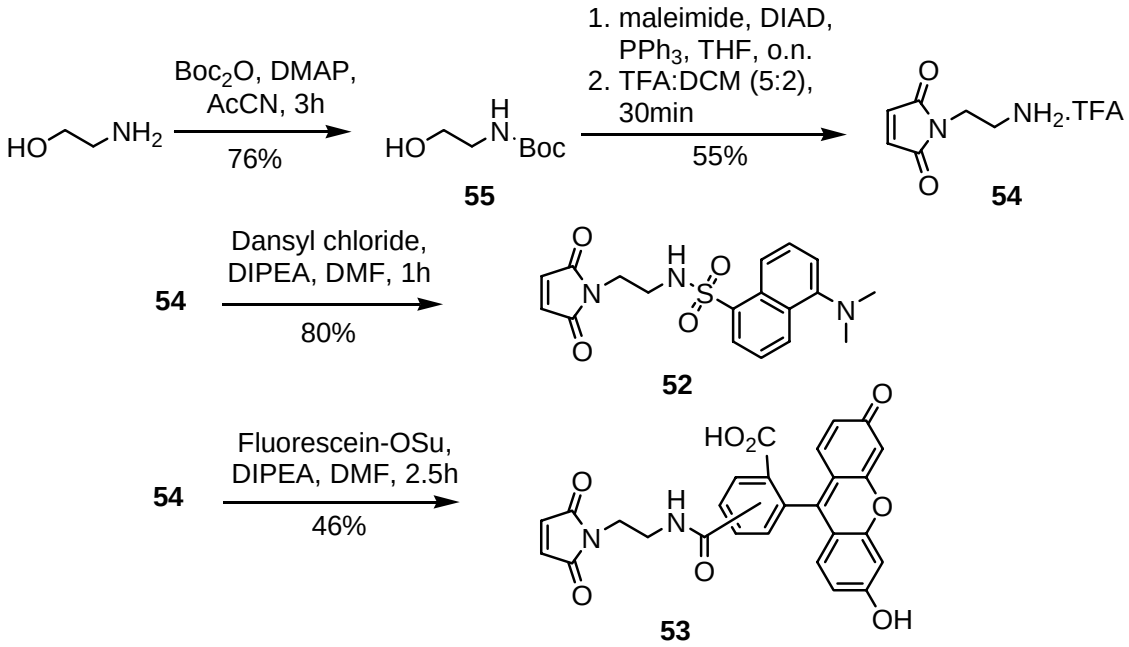
The cross-linker **49** was easily prepared in two steps from esterification of pimelic acid and *trans,trans*-2,4-hexadien-1-ol **1** followed by transformation of the diene **50** into the *N*-hydroxysuccimidyl ester **49**. Both reactions were mediated by the DIC/DMAP coupling method (Scheme 21). The linker **49** was first attached to streptavidin protein molecules by acylation of superficial lysine residues through the NHS-moiety (Scheme 22). The molar ratio between the cross-linker and streptavidin (6:1) and reaction conditions (H_2O, 2h, 25°C) were adjusted to provide low levels of protein modification. The diene-streptavidin conjugate **51** was finally purified using diafiltration with four changes of water. MALDI-TOF

measurements of the conjugated protein indicated that on average each streptavidin subunit was conjugated with one diene-linker (Figure 29B).



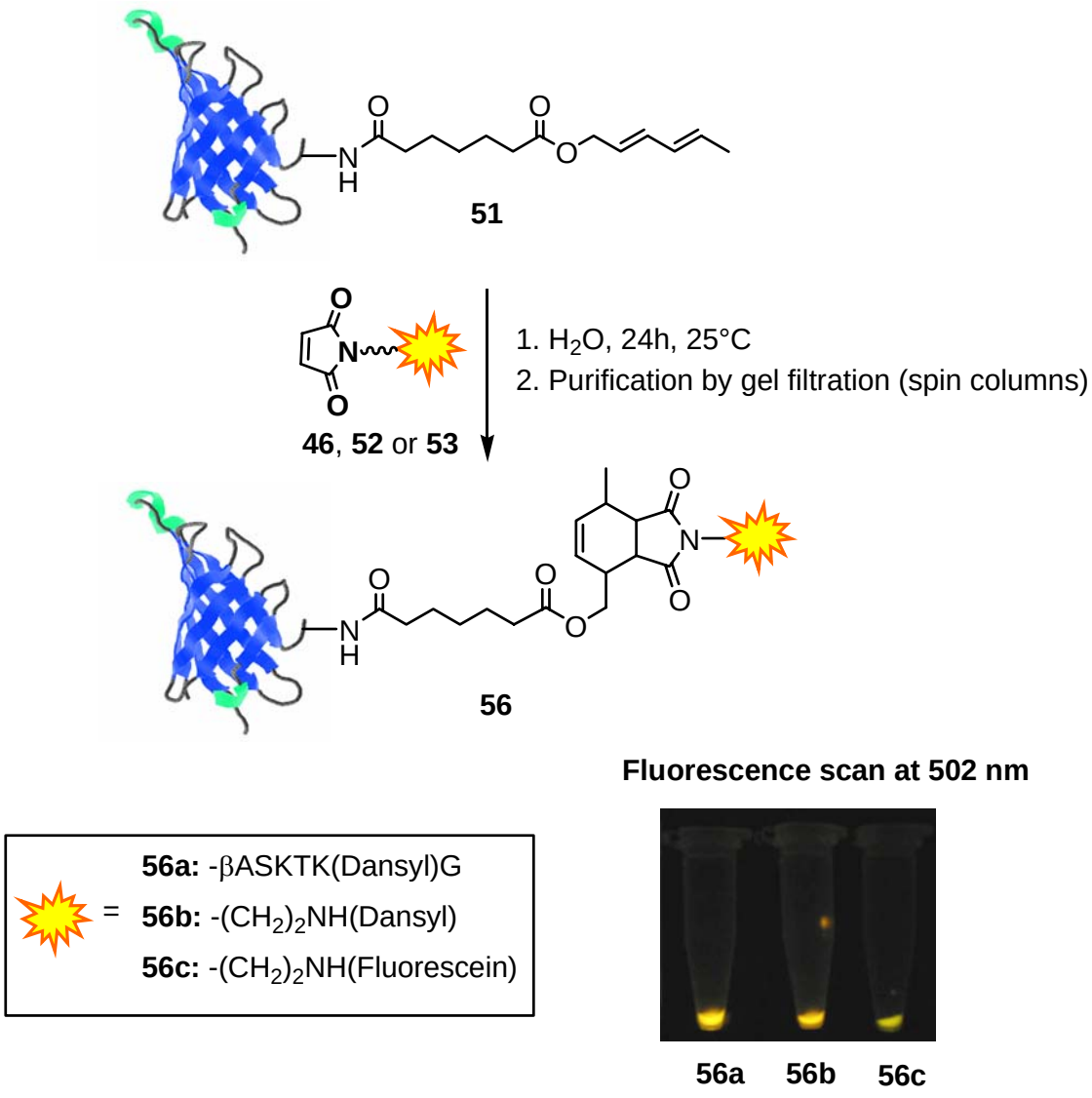
Scheme 21. Synthesis of the diene cross-linker and preparation of the hexadienyl conjugated streptavidin.

To illustrate the applicability of the DA conjugation the diene-modified protein was treated with three different fluorescent labeled maleimide probes **46**, **52** and **53**. Dansyl- (**52**) and fluorescein-maleimide (**53**) were synthesized from precursor **54**[107] as shown in Scheme 22.



Scheme 22. Synthesis of the maleimide fluorophores.

The conditions for the cycloaddition reactions were: 10µL diene-conjugated protein at 5 mg/ml in water (ca. 100µM concentration) was incubated with 2.7µL of a 10mM stock solution of the maleimide-compound (30-fold relative to protein, approximately 6-fold relative to diene quantity) at 25°C for 24h (Scheme 23). After removal of unligated dienophile by passing the reaction mixture through a spin gel filtration column (DyeEx columns, Qiagen), the presence of the new fluorescent protein **56** at 302 nm was verified by SDS-PAGE for all three conjugation reactions (Figure 28). In addition, the mass of the expected ligation cycloadduct was confirmed by MALDI-TOF analysis (Figure 29). Control tests involving the combination of streptavidin and maleimido-molecules **46**, **52** and **53** yielded no evidence of unspecific reactions of the dienophile with other functionalities of the protein (Figure 28).



Scheme 23. Diels-Alder conjugation of the diene-modified streptavidin with maleimide fluorophores.

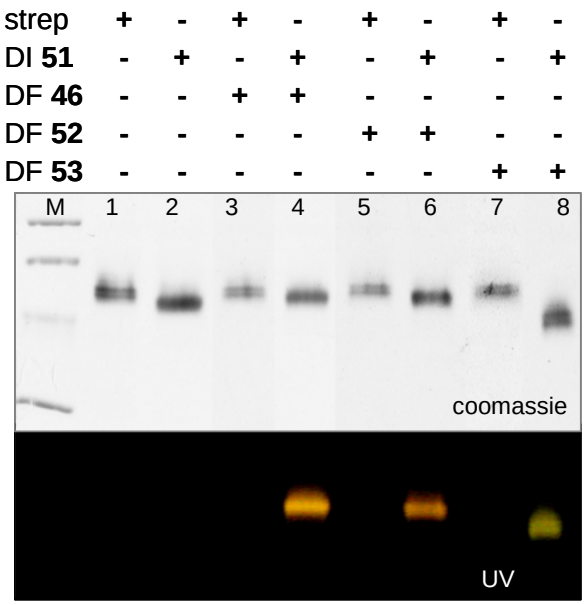


Figure 28. SDS-PAGE analysis of the Diels-Alder bioconjugation. *strep*: streptavidin, *DI*: diene-streptavidin conjugate, *DF*: dienophile (maleimide probes), *M*: molecular weight marker (up to down): 97, 67, 43 and 30 KDa.

Previous reports stated that the rate of the Diels-Alder bioconjugation employing nucleotides or carbohydrates is notably affected by the pH, proceeding much faster at pH 5.5-5.7 or in ion-exchanged water.[85,89] The next step was then to inspect the selectivity of the cycloaddition at different pH values. Anticipately the possibility to carry out these reactions at pH above 7 was discarded since nucleophilic addition of the dienophile double bond by amino groups can take place under these conditions.[13] Although this side reaction is remarkably slow at neutral and slightly acidic pH, one should be concerned about its possible occurrence owing to the relatively elongated reaction time and dienophile excess in which the cycloaddition needs to be performed for furnishing efficient Diels-Alder bioconjugation. For this reason, the ligation between streptavidin-diene conjugate **51** and a larger excess of dansylated maleimide-peptide **46** (100-fold relative to **51**, ca 25-fold relative to diene content) was investigated in sodium phosphate buffer in the pH range of 5.5 to 7.0 for 24h at 25°C. To verify any possible side-reaction, control experiments were carried out substituting the streptavidin-conjugate by pure streptavidin. As shown by SDS-PAGE analysis (Figure 30), the Diels-Alder cycloaddition is selective at pH 5.5 to 6.5, but not completely at pH 7.0. The observation of a fluorescent band for the reaction of streptavidin with **46** at this pH (lane 9,

Figure 30) indicated that the maleimido-compound binds unspecifically to the protein molecule under these conditions.

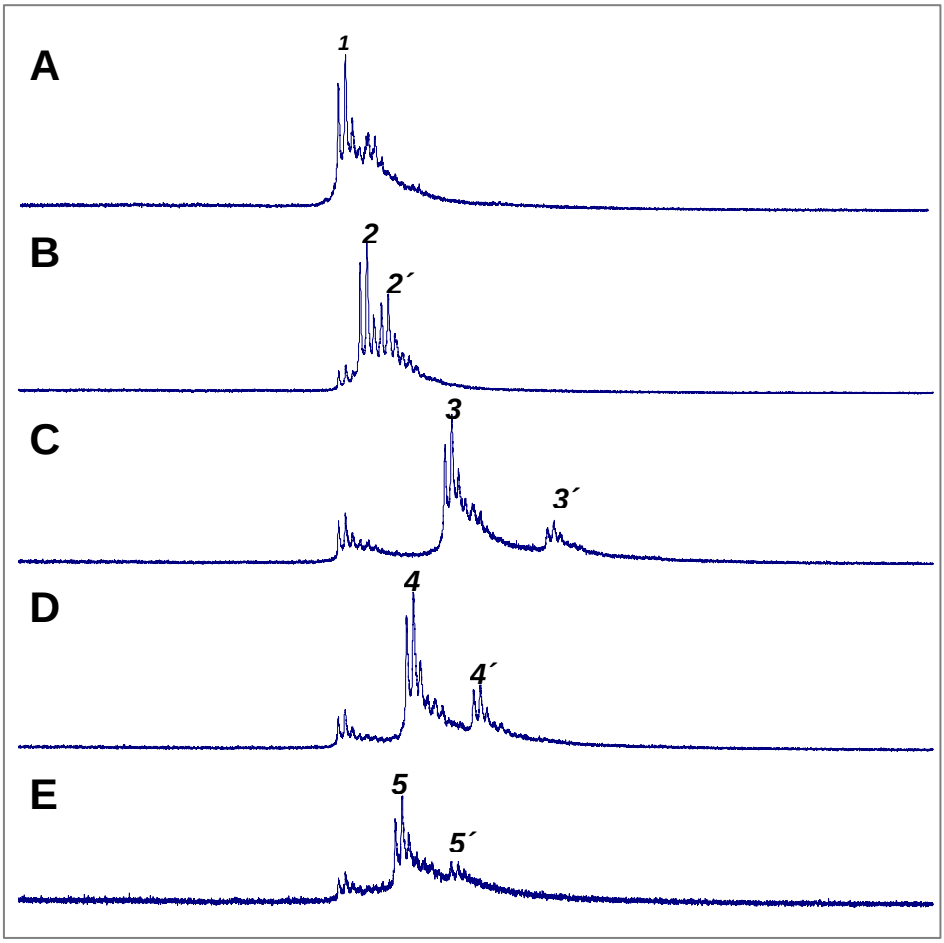


Figure 29. MALDI-TOF mass spectra for streptavidin and streptavidin conjugate subunits. Because the composition of native streptavidin subunits varies from 123 to 125 amino acids, a range of different mass peaks was found between 13115 and 14000 for the spectrum of streptavidin (**A**). The largest subunit peak, experimentally found at mass 13180 (± 6), was taken as reference for calculation of the expected protein conjugate mass. **A**. Streptavidin subunit: *1* = major peak 13180. **B**. Streptavidin-hexadiene conjugate **51**: *2* = one diene linker conjugated 13407 (calc. 13402), *2´* = two dienes linkers conjugated 13630 (calc. 13624). **C.** Streptavidin-cycloproduct **56a**: *3* = one cycloproduct conjugated 14307 (calc. 14306), *3´* = two cycloproducts conjugated 15432 (calc. 15432). **D**. Streptavidin-cycloproduct **56b**: *4* = one cycloproduct conjugated 13779 (calc. 13775), *4´* = two cycloproducts conjugated 14374 (calc. 14370). **E**. Streptavidin-cycloproduct **56c**: *5* = one cycloproduct conjugated 13900 (calc. 13900), *5´* = two cycloproducts conjugated 14623 (calc. 14620).

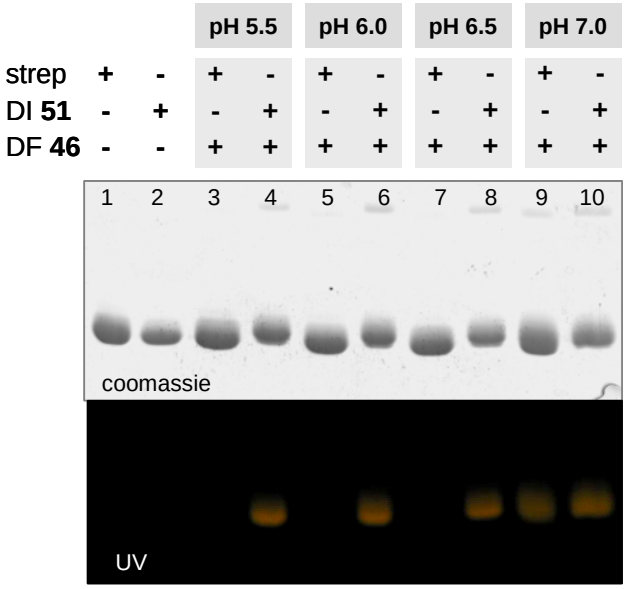
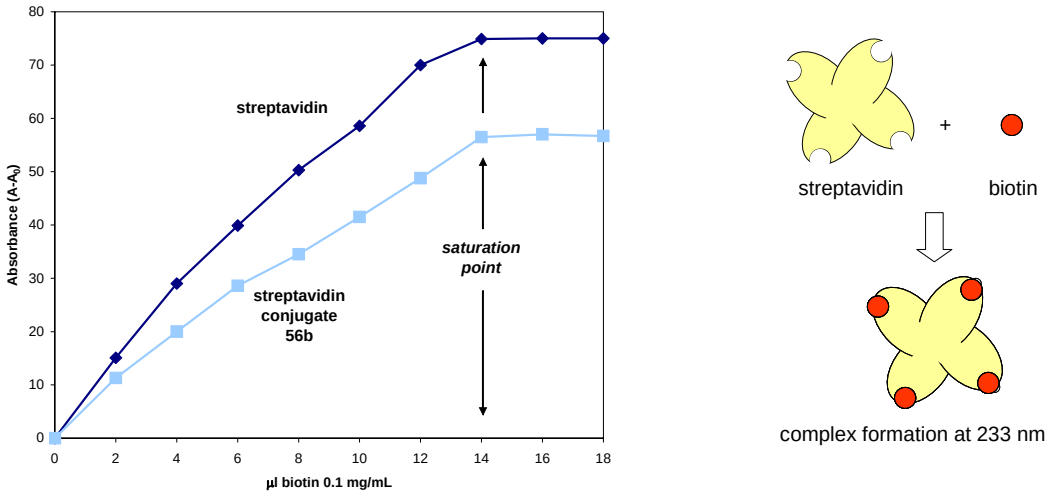


Figure 30. SDS-PAGE analysis for the Diels-Alder bioconjugation at pH 5.5 to 7.0. The observed protein bands represent streptavidin subunits as the tetrameric protein complex was denatured upon heating of the sample at 80°C for 3 minutes with denaturating loading buffer prior gel loading.

The last results revealed that the ligation between the diene derived protein and the dienophile compounds was an effective and selective reaction in pure water or under slightly acidic conditions. Subsequently it was also important to determine the consequences of the Diels-Alder labeling process for the proteins biological activity. Since the model protein employed here combines stoichiometrically with biotin, it was possible to use two standard spectrophotometric biotin-streptavidin complex assays to measure the biotin binding capacity of the new modified streptavidin proteins and then estimate any loss of activity.[108] In the first assay, the complex formation between streptavidin and biotin is monitored at 233nm. A red shift in absorption of the tryptophan residues of streptavidin that occurs upon complexation with biotin is responsible for the absorption change (Figure 31A). The second assay makes use of the dye 4´-hydroxyazobenzene-2-carboxylic acid (HABA) which also binds noncovalently streptavidin at the same site as biotin but with lower affinity (10^6 M^{-1}), forming a colorful complex that absorbs at 500 nm. The addition of biotin to a HABA-streptavidin complex results in displacement of HABA from the binding site by biotin and consequent decay on the absorbance of the complex at 500 nm (Figure 31B).

A. Direct measurement of streptavidin-biotin complexation



B. Indirect measurement of streptavidin-biotin complexation

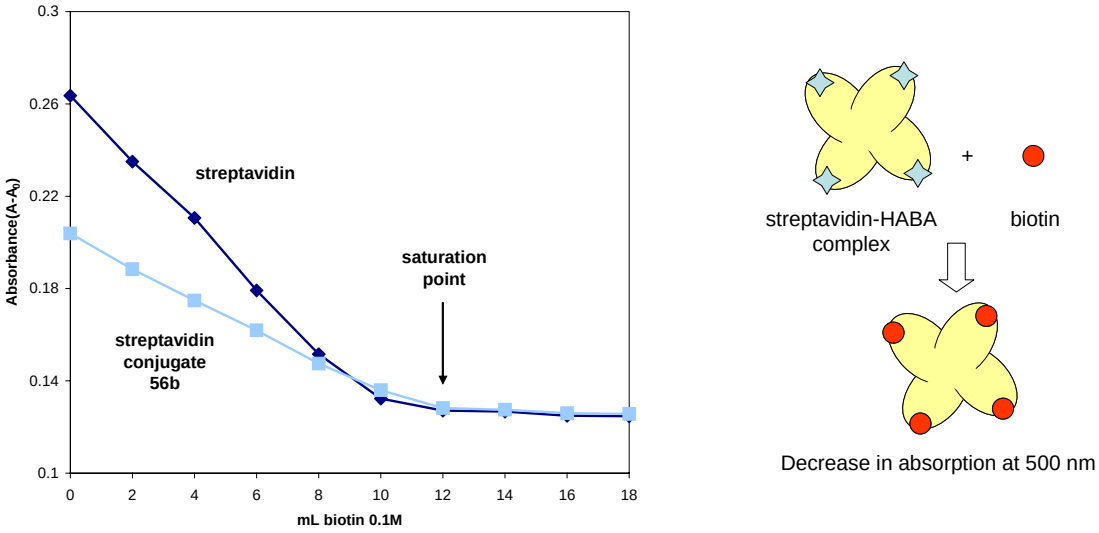
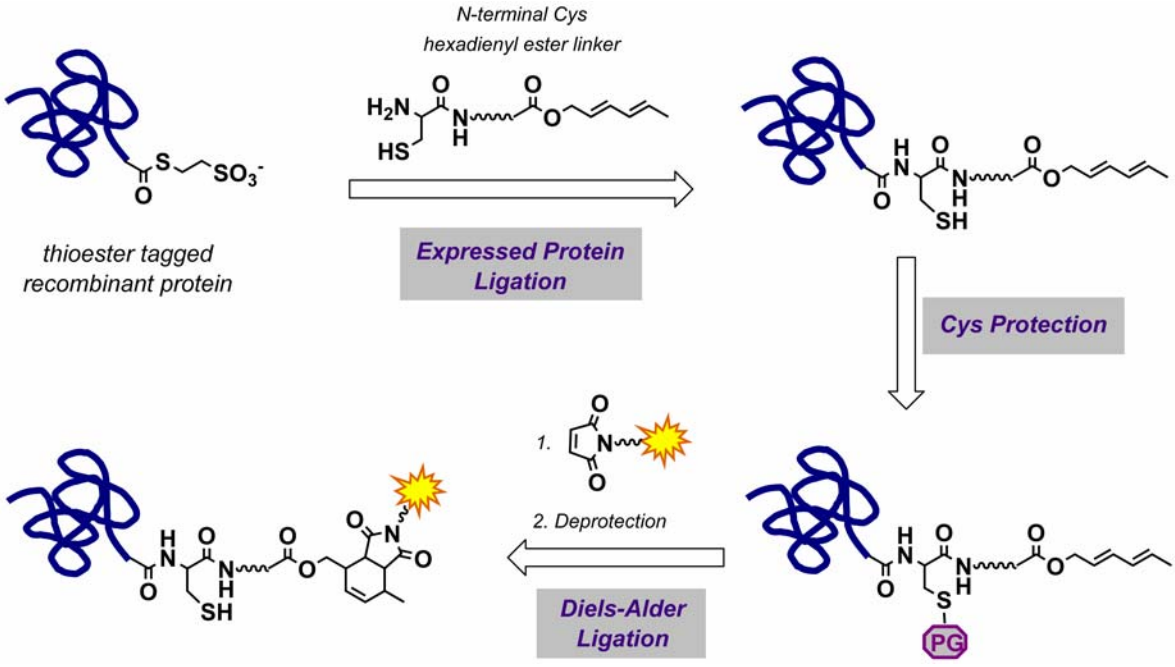


Figure 31. Assays for the determination of biotin binding capacity of streptavidin. Solutions of wild-type streptavidin and conjugate **56b** were assayed at concentration of 0.1 mg/mL. **A**. In this assay, the streptavidin sample was titrated with a biotin solution and the absorbance was recorded at 233nm. The amount of biotin at the inflection point indicates the biotin binding capacity. Because the biotin binding of both proteins is saturated with the same amount of biotin, they have the same activity (1mg protein binds 14µg biotin). **B**. In this assay, a solution containing a streptavidin-HABA complex was titrated with a solution of biotin and the absorbance was recorded at 500nm. If the protein samples show the same differential absorbance ($A-A_0$) at 500nm after the titration over the saturated point, then they have the same biotin binding activity.

In both titrimeric assays, wild-type streptavidin and streptavidin conjugate **56b** displayed similar biotin binding capacity (determined as 14 μg biotin per milligram protein, in concordance with the value given by the commercial supplier), indicating that the protein remained fully functional after the Diels-Alder conjugation.

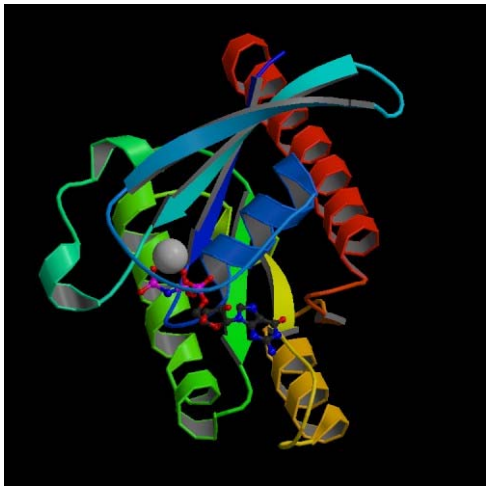
4.2.3. Site-specific labeling of Rab proteins by combination of Expressed Protein Ligation and Diels-Alder Ligation.

The Expressed Protein Ligation method is a suitable tool to equip a given protein with the diene functionality at a specific position.[2] For instance, the hexadiene group can be specifically located at the C-terminus of the protein by reaction of a thioester tagged recombinant protein with a particular linker enclosing a hexadienyl ester at one extremity and a cysteine amino acid at the other (Scheme 24). The EPL process generates a nucleophilic cysteine residue at the ligation site which should be temporary protected (along with other accessible cysteine side chains present in the protein) to avoid undesired modification of the mercapto group in the subsequent reaction with maleimido-probes. After the Diels-Alder ligation process, the masked cysteines can be transformed back into the free thiol form.



Scheme 24. Combination of EPL and Diels-Alder ligation.

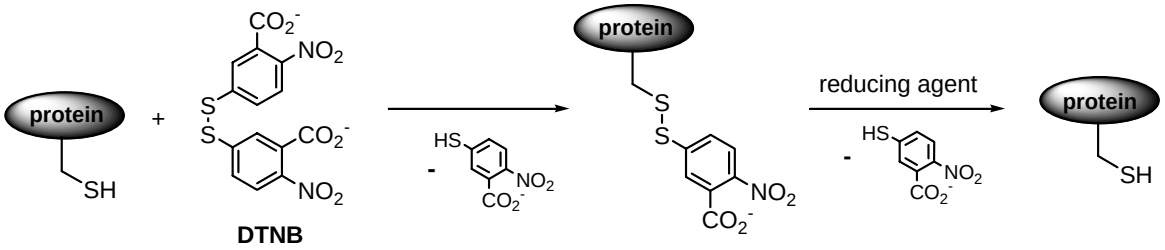
This approach was successfully implemented employing the Rab7 protein as a representative biologically relevant example (Figure 32). The Rab proteins are small GTP-binding proteins that mediate intracellular vesicular traffic.[109] Their function is important in the control of protein distribution after translation and processing, exocytosis, and endocytosis. Many aspects of the multi-step processes involving the Rab proteins are only partially understood, including, for instance, the timing of the cellular event and the localization of the proteins involved. Therefore, differently and site-specific labeled Rab proteins are considered versatile probes for the study of such biological events.[110-114]



Rab7 (207 AA):
MTSRKKVLLKVIILGDSGVGKTSLMNQYVNKKFSNQYKATIGADFLTKEVMVDDRLVTMQ
IWDTAGQERFQSLGVAFYRGADCCVLVFDVTAPNTFKTLDSWRDEFLIQASPRDPENFPF
VVLGNKIDLENRQVATKRAQAWCYSKNNIPYFETSAKEAINVEQAFQTIARNALKQETEV
ELYNEFPEPIKLDKNERAKASAESCSC

Figure 32. Structural details for Rab7 protein complexed with GNP and magnesium ion. Cysteine residues are marked in red.

Before implementing the proposed methodology for the labeling of Rab proteins, it was crucial to find an appropriate cysteine blocking agent that could efficiently bind to the mercapto groups of the protein and render them inert toward further reactions with maleimides. Ellmann´s reagent, 5,5´-dithio-bis-nitrobenzoic acid (DTNB), is a suitable reagent for this purpose.[115] It readily reacts with cysteines to form stable disulfide bonds and can later be cleaved by addition of reducing reagents to regenerate the free sulfhydryl groups (Scheme 25).



Scheme 25. Protection/deprotection of cysteine residues with the Ellman´s reagent. In both reactions, the chromogenic TNB anion (5-thio-2-nitrobenzoic acid) is released (absorbance at 412nm).[13]

To prove the efficiency of Ellmann´s reagent to mask the cysteine residues of Rab proteins, several control experiments were performed in which wild-type Rab7 was treated with dansylated maleimide **46** with or without prior DTNB protection. Rab7 has five Cys residues in its structure: three of them are located on the protein surface and can potentially be labeled by maleimides while the other two are buried in the core of the molecule and are not easily accessible as long as the protein remains folded.[114] Without blocking the Cys residues with the DTNB reagent, the protein was multi-labeled after overnight incubation with **46** at pH 6 as indicated by SDS gel electrophoresis analysis (Figure 33A). However, if the three Cys residues located on the protein surface are impeded to react by TNB disulfide formation blockage, no significant binding of the maleimide and Rab7wt was observed and the protein remained unlabeled. The other two Cys residues were mostly not reached by the Ellmann´s reagent as well as by the maleimide compound. These results were confirmed by ESI-MS measurements (Figure 33B).

Initial functionalization of Rab7 to generate the hexadienyl protein was achieved by means of expressed protein ligation. To this end, recombinant truncated Rab7 thioester from *Canis familiaris* (Rab7ΔC3-MESNA thioester) was prepared by members of the research group of Dr. Kirill Alexandrov at Department of Physical Biochemistry at MPI Dortmund.[110] The method uses a genetically engineered intein and a chitin binding domain (CBD) as fusion partners to express and purify the desired Rab thioester (Section 2.6.1) and has been successfully applied for the semi-synthesis of fluorescent labeled and lipidated Rab proteins by the groups of Dr. Kirill Alexandrov and Prof. Dr. Herbert Waldmann.[110-114]

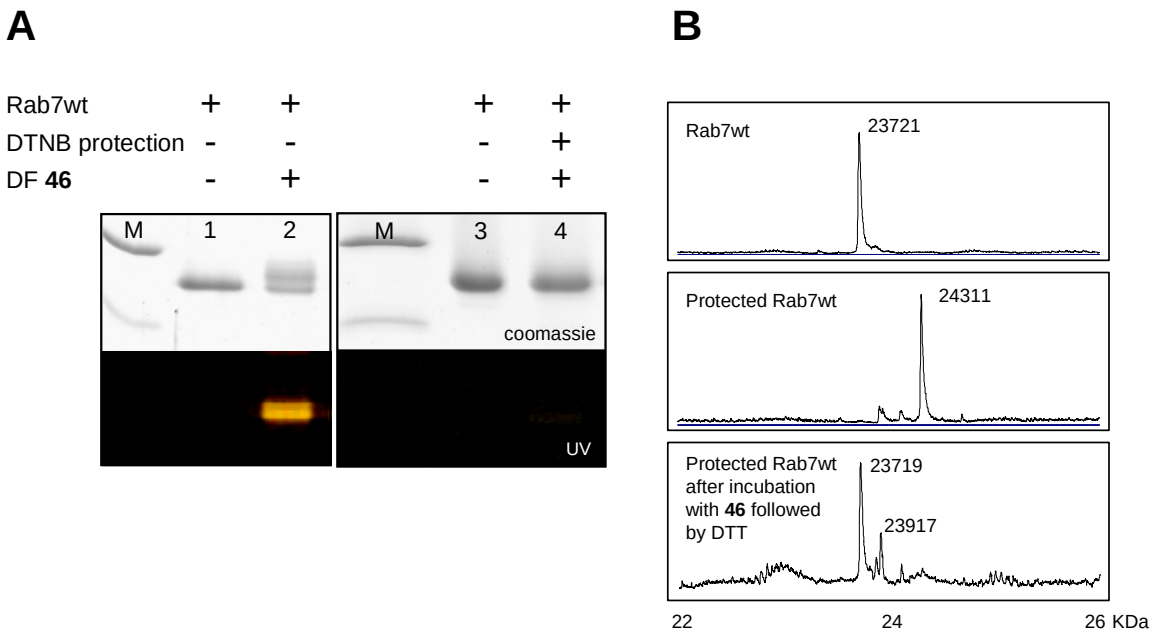
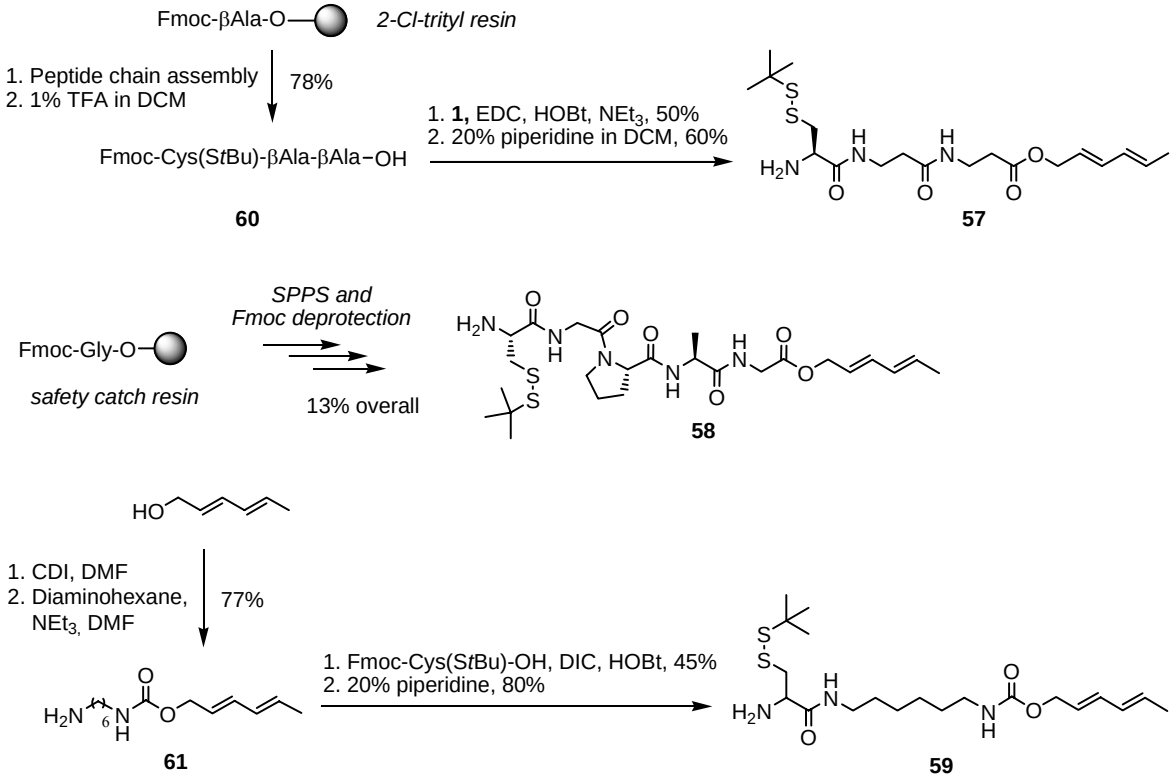


Figure 33. A. SDS-PAGE analysis of control reactions involving Rab7 wild type and maleimide **46** with or without prior cysteine DTNB protection. Labeled Rab7 appeared as a fluorescent protein band at ca 24-26 KDa. Key: *M*: molecular weight marker (up to down): 30, 20 KDa. B. ESI-MS for Rab7wt (experimental mass: 23721), TNB protected Rab7 (expected mass for Rab7 + 2TNB: 23415), and the product from the reaction of protected Rab7wt with maleimide **46** (20h, 25°C) followed by treatment with DTT (expected mass: 23721 (Rab7wt), 23918 (Rab7wt + 1TNB)).

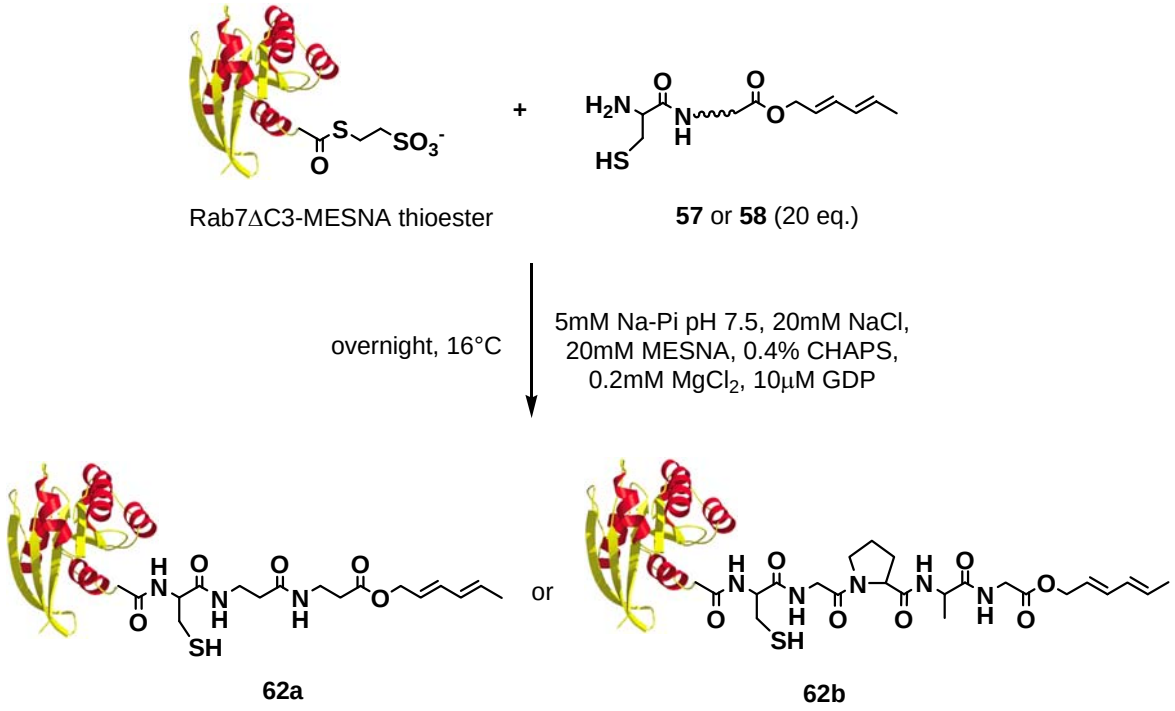
For the dienyl linker construction, three different cysteinyl hexadiene linkers **57**, **58** and **59** were synthesized (Scheme 26). Two of them have a peptide-like structure while the other embodies an alkyl chain. The reductively cleavable S*t*Bu group was chosen as the protecting group of the N-terminal cysteine to prevent oxidative reactions.[111] The deproctection of the cysteine is carried out in situ during protein ligation by using an excess of thiol agent. The synthesis of the linker **57** began with the assembly of Fmoc-protected compound **60** on solid phase using 2-chloro-trityl resin, followed by esterification of the resulting carboxylic peptide **60** with hexadienol **1** and final removal of the Fmoc protecting group (Section 4.1.3). Peptidyl hexadienyl ester **58** was synthesized on solid phase using the safety-catch resin strategy as described before (Section 4.1.3). The synthesis of the linker **59** started from the carbamate precursor **61**, which was prepared by reaction of hexadienol **1** and CDI and subsequent combination with 1,6-diamino-hexane in DMF. After that, the amino

group of **61** was coupled to Fmoc-Cys(S*t*Bu)-OH by the EDC activation method followed by Fmoc-deprotection to give the unprotected cystenyl compound **59**.



Scheme 26. Synthesis of the cysteinyl hexadiene linkers.

For the construction of the C-terminal hexadienyl Rab7 protein, recombinant truncated Rab7 thioester (Rab7ΔC3-thioester) was ligated with peptide hexadienyl ester under reducing conditions (MESNA) overnight at 16°C (Scheme 27).[111] An excess of 20 equivalents of the dienyl compound was used to ensure complete conversion. The ligation was carried out in the presence of guanosine 5´-diphosphate (GDP) and magnesium chloride to assure protein stabilization. Small amount of the detergent CHAPS were also included to increase protein and peptide solubilization under ligation conditions. While the ligation of dienes **57** and **58** was straightforward, the linker **59** could not be efficiently ligated to the protein because of its insolubility in the ligation buffer.



Scheme 27. Semi-synthesis of C-terminal Rab7 hexadienyl esters by EPL.

The ligation process generates a nucleophilic cysteine side chain and the truncated Rab proteins **62a-b** embody three further cysteines (Figure 34). The accessible cysteine side chains were protected as disulfide by treatment with Ellmann`s reagent immediately after the ligation reaction, without intermediate purification, to yield masked protein esters **63a** and **63b** (Scheme 28).

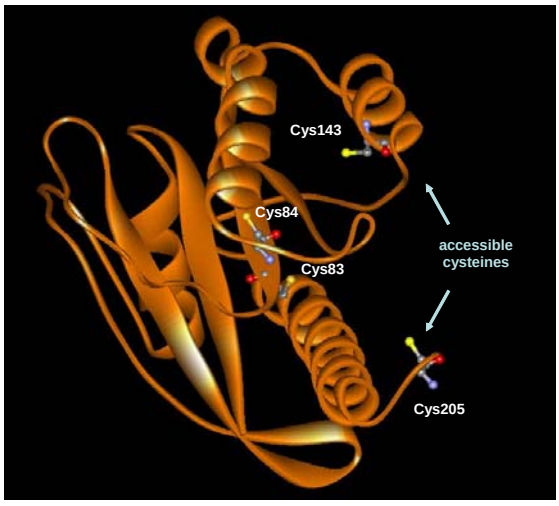
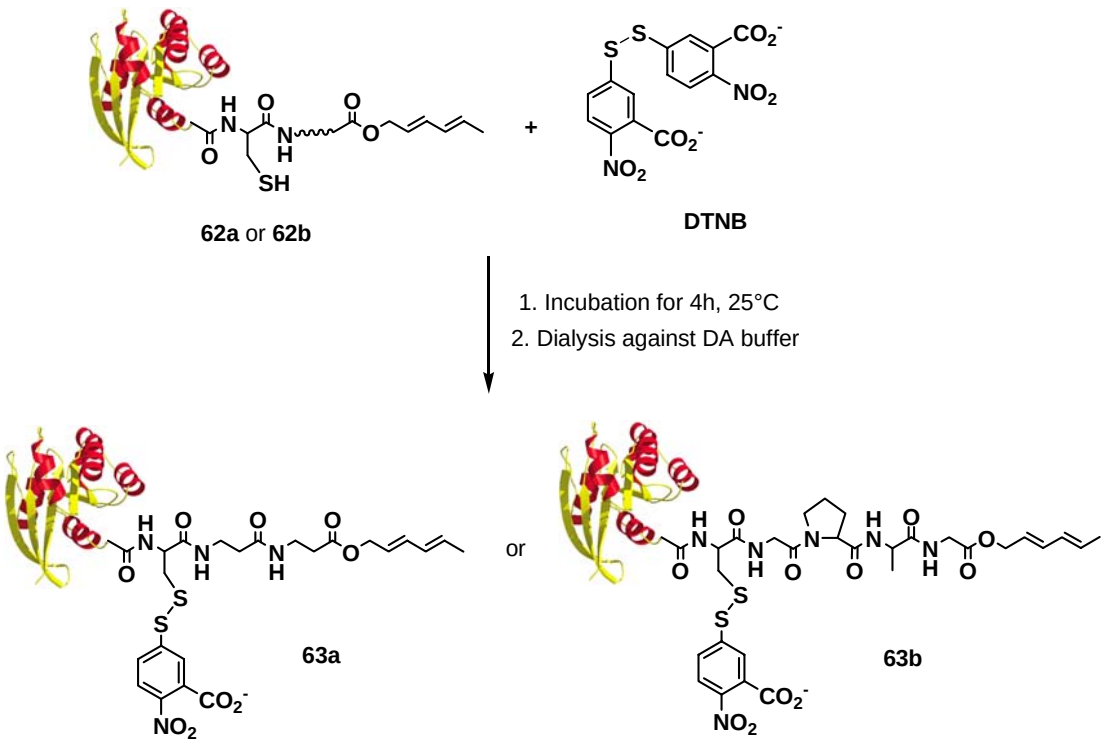


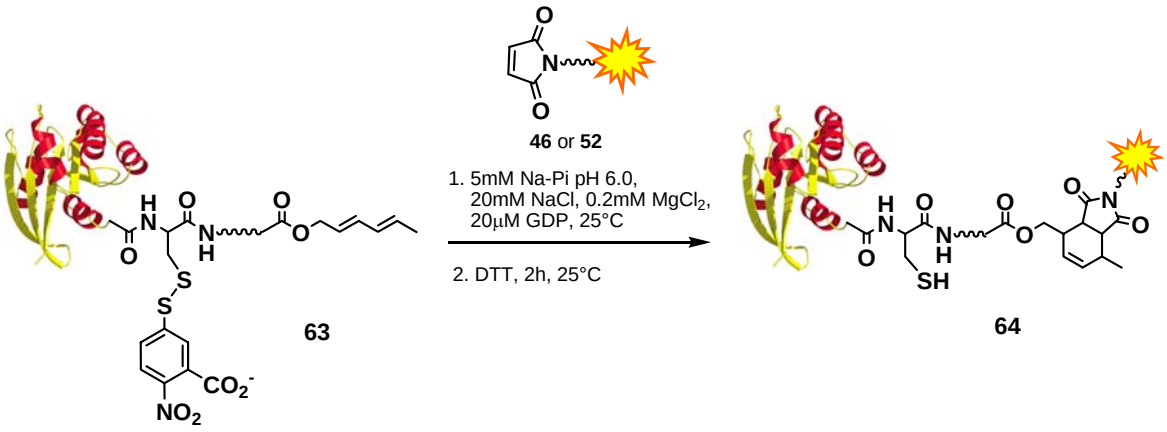
Figure 34. Position of the cysteine residues in the Rab7 hexadienyl ester protein.[114]

Protein hexadienyl ester was then dialyzed against DA buffer (5mM Na-Pi pH 6.0, 20mM NaCl, 0.2mM MgCl2, 20μM GDP) to remove all small molecules (MESNA, Ellmann´s reagent, dienyl peptide) and prepare the protein solution for the following ligation step. As determined by ESI-MS measurements (Figure 36), the two accessible cysteine residues of the truncated Rab7 protein **63** were in fact masked by the disulfide groups.



Scheme 28. Masking of the accessible cysteine residues of the Rab7 hexadienyl esters by treatment with DTNB.

Next the protected hexadienyl Rab7 protein **63** was subjected to Diels-Alder reaction with peptide-derived dienophile **46** and dansyl derivative **52** (Scheme 29). The cycloaddition ligation was carried out in buffer pH 6.0 at a protein concentration of approximately 40μM and at room temperature. A variety of other conditions were investigated for this ligation step as demonstrated in Figure 35. The coupling reactions were terminated by addition of excess dithiothreitol which traps the dienophile and simultaneously converts the disulfides into unmasked thiols. Because of the release of the chromogenic TNB group into solution, the ligation solution became yellowish at this point.



Labeled Rab7 protein	Rab7 hexadienyl ester	Maleimide
64a	63a	46
64b	63a	52
64c	63b	46
64d	63b	52

Scheme 29. Diels-Alder ligation of the masked Rab7 hexadienyl esters and maleimide compounds.

The extent of the three-step ligation process was followed by SDS-PAGE (Figure 35) and all intermediates and final products were identified by ESI-MS mass spectrometry (Figure 36). Based on gel quantification and LC spectra analyses, it appears that ca. 50% of the Rab7 hexadienyl ester **63** underwent cycloaddition with the maleimide probes after 20-24h incubation. Increasing the amount of added maleimide to 100-fold improves considerably the ligation efficiency to almost complete conversion after 24h reaction (Figure 35C). Undesired multiple labeling of Rab7 was not detected at significant levels.

The Diels-Alder conjugation of the hexadienyl Rab protein with the dansylated labeled dienophile rendered this protein insoluble in the reaction solution. The aggregation formation could not be avoided even when the ligation was carried out in the presence of cosolvent (glycerol) or detergents (CHAPS, Triton X-100 or Tween-20), or at lower protein concentration (10-20μM). This behavior was also formerly observed during the semisynthesis of lipidaded Rab proteins.[111-113]

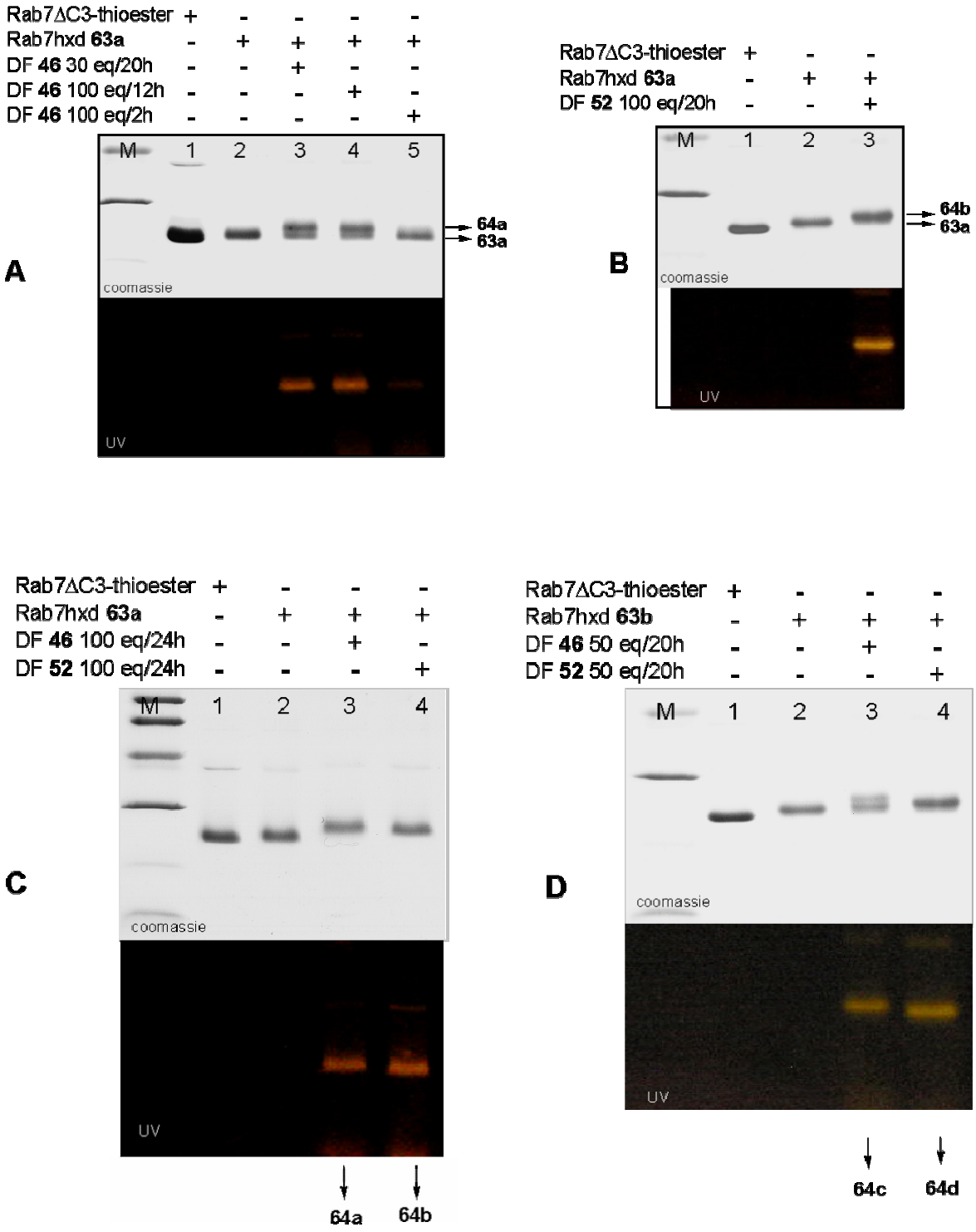


Figure 35. SDS-PAGE analysis of the Diels-Alder ligation between protected Rab7 hexadienyl ester **63** and maleimide probes **46** and **52**. Diels-Alder labeled Rab7 appeared as a unique fluorescent protein band at ca 24 KDa. Key: *Rab7hxd*: Rab7 hexadienyl ester. *M*: molecular weight marker (up to down): 43, 30, 20, 14 KDa (in A, B, D) and 94, 67, 43, 30, 20, 14 KDa (in C).

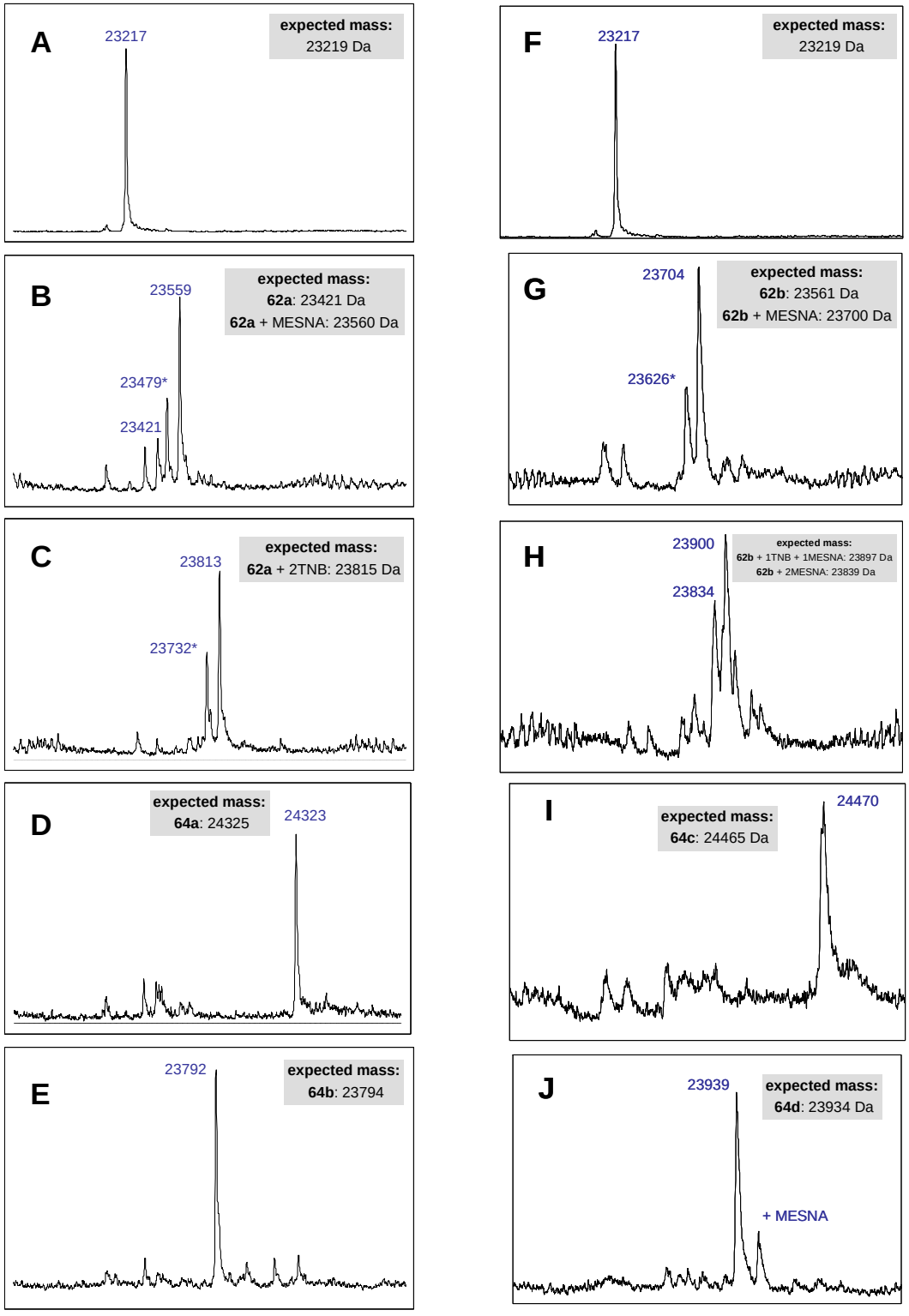


Figure 36. ESI-MS spectra of the Rab7ΔC3-thioester (A and F), Rab7 hexadienyl esters **62a** (B) and **62b** (G), masked Rab7 hexadienyl esters **63a** (C) and **63b** (H), and Diels-Alder modified Rab7 proteins **64a** (D), **64b** (E), **64c** (I) and **64d** (J). Key: * The ESI-MS spectrum of the hexadienyl esters showed a typical fragmentation peak of $[M\text{-}80]^+$ corresponding to the lost of the hexadienyl group during ionization.

The purification of the Rab protein **64** (Figure 37) was performed following the procedure described for the isolation of semisynthetic prenylated Rab proteins.[111] After incubation of the Rab hexadienyl ester **63** with the maleimide, the resulting protein precipitate was separated by centrifugation and washed with methanol in order to remove the excess of unligated dienophile. The resulting pellet was dissolved in buffer containing 6M guanidinium chloride and DTT and then refolded by a 25-fold dilution with renaturation buffer. Subsequently, REP-1 protein, a natural chaperone for Rab proteins,[114] was added and the formed complex was further purified by dialysis and concentrated by ultracentrifugation. Biochemical activity assays which will provide a measure of the functional activity of the fluorescent Rab7 **64** are currently being performed at Kirill Alexandrov´s group at MPI Dortmund.

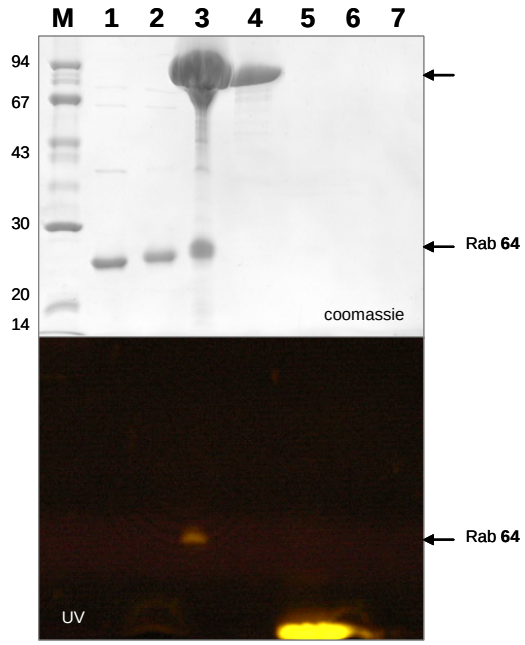
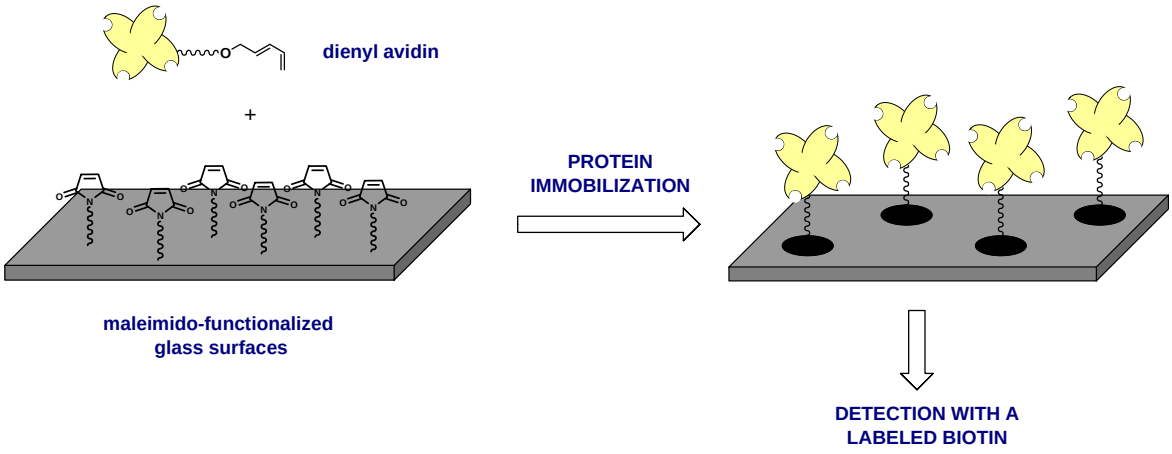


Figure 37. Purifcation of the ligation product (**64a**) resulting from the combination of the masked Rab7 hexadienyl ester **63a** with maleimide **46** (50-fold). After 24h incubation of protein and dienophile, the formed precipitate was separated from the supernadant (lane 5) by centrifugation. The pellet was washed twice with methanol (resulting washings: lanes 6 and 7), dissolved in denaturation buffer, refolded in renaturating buffer and complexed with REP-1 (lane 3). Other lanes: Rab7 MESNA thioester (lane 1), Rab hexadienyl ester **63a** (lane 2), REP-1 (lane 4), molecular weight marker (lane M).

4.3. Immobilization of proteins on glass surfaces by Diels-Alder ligation

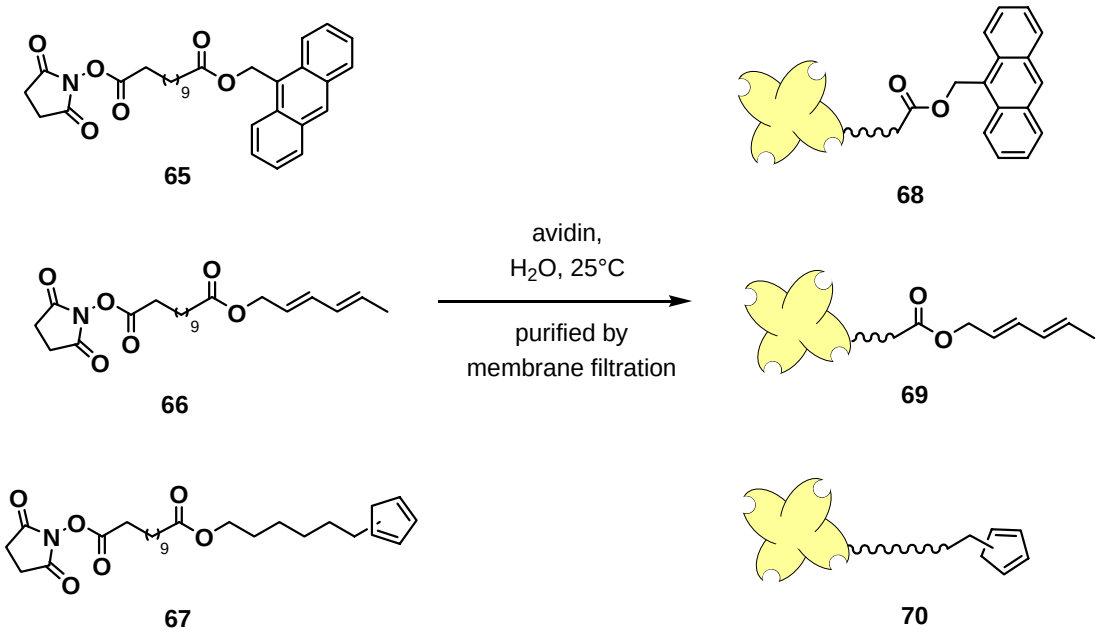
An attractive application of the developed Diels-Alder ligation method would be for the fabrication of protein biochips, enabling a robust and chemoselective covalent immobilization of proteins onto surfaces. Whereas the Diels-Alder reaction has been successfully employed for the surface attachment of small biomolecules in microarrays,[86,87,90-93] the challenge of immobilizing complex macromolecules such as proteins by this method still remains. Our initial efforts were focused on the use of avidin as a model protein in order to investigate the scope of the Diels-Alder reaction for immobilizing proteins on the surface of glass slides. For this aim, dienyl-avidin conjugates were prepared and spotted onto maleimide-functionalized glass slides. The level of immobilized protein was detected after binding with a fluorescently labeled biotin (Scheme 30).



Scheme 30. Immobilization of avidin protein by Diels-Alder ligation.

Three different types of diene conjugates were designed to assist the Diels-Alder protein immobilization (Scheme 31). One encloses the 2,4-hexadienyl group, the diene functionality that has been extensively studied in previous investigations. The other two conjugates possess an anthracene or a cyclopentadienyl moiety. These functions are expected to also display high cycloaddition reaction rates in aqueous media.[79,82] Nevertheless these functionalities are presumably more unstable under aqueous solutions than the acyclic diene,

being gradually desactivated by dimerization. Therefore, the relation between reactivity and stability of the dienyl conjugates are two important issues to be investigated for the immobilization of proteins by Diels-Alder ligation. For this purpose, the dienyl bifunctional linkers **65**, **66** and **67** were synthesized by Dr. José Palomo (Department of Chemical Biology, MPI Dortmund). As depicted in Scheme 31, avidin was bioconjugated with each linker employing essentially the same procedure as described previously (Section 4.2.2). On average, each avidin subunit was acylated with one dienyl linker as indicated by their mass spectrum (Figure 38). Conjugates **68**, **69** and **70** were kept at -80°C until use (within one week).



Scheme 31. Preparation of the dienyl avidin conjugates.

Dendrimer-activated glass slides were employed as support for the protein microarray and were kindly provided by Chimera Biotec GmbH (Dortmund).[116] The surface of these glass slides is modified with an intermediate layer of fourth generation polyamidoamine (PAMAM) starburst dendrimers, containing a large number of primary amino groups in the outer layer (Scheme 32). The presence of the dendrimeric interface leads to amplification of the signal intensity as well as to improvement of the physical-chemical stability and homogeneity of the biochips.[117,118] The functionalization of the dendrimeric slides with maleimido groups was performed by Maja Köhn (Department of Chemical Biology, MPI Dortmund) as indicated in Scheme 32.

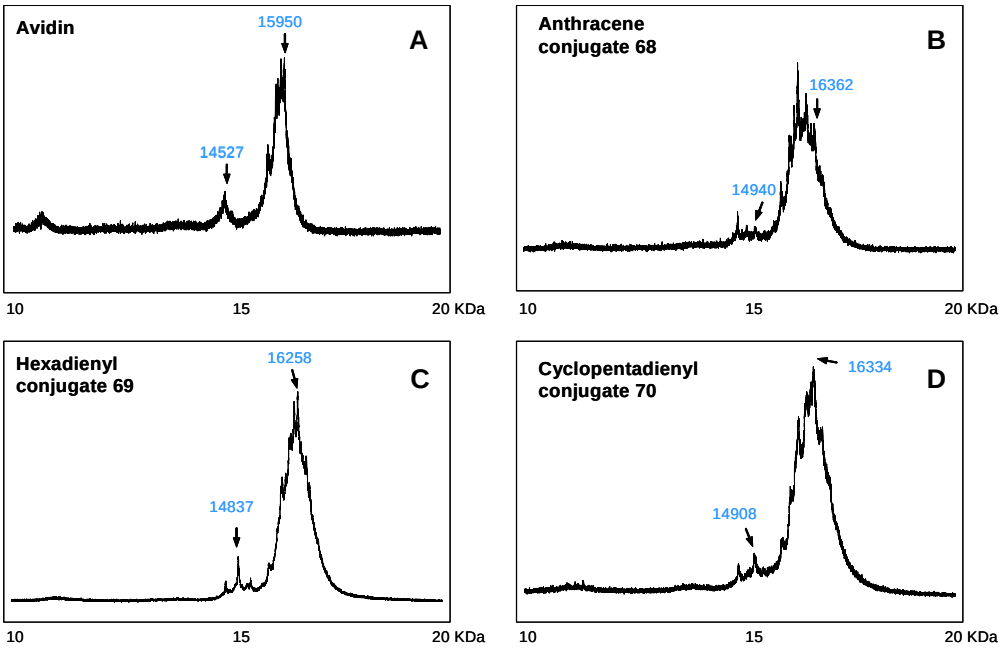
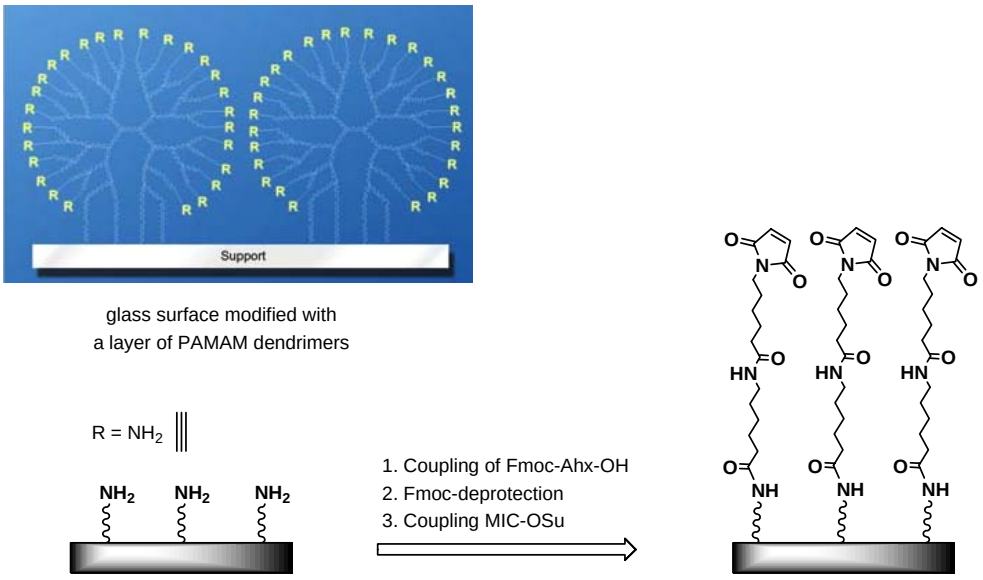
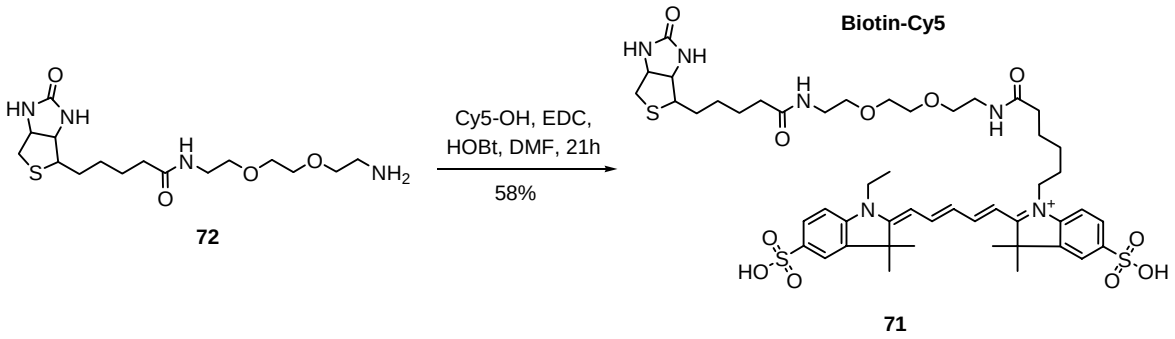


Figure 38. MALDI-TOF mass spectra for avidin and avidin conjugate subunits. Because the composition of the avidin subunits is variable, a range of different mass peaks were found between 14527 and 16126 for the spectrum of avidin. The two subunit peaks, experimentally found at mass 14527 and 15950 (± 10) Da (**A**), were taken as reference for calculation of the expected protein conjugate mass. **B**: expected mass 14947 and 16370. **C**: expected mass 14838 and 16261. **D**: expected mass: 14906 and 16329.



Scheme 32. Functionalization of dendrimer-modified glass surfaces with maleimido groups. MIC-OSu: Maleimido caproic acid-N-succinimidylester.

For the detection of the immobilized avidin molecules in the microarray experiment, the Cy5 derivatized biotin **71** was synthesized by acylation of the biotin derivative **72** with Cy5 mediated by the EDC/HOBt activation reagent in DMF (Scheme 33).

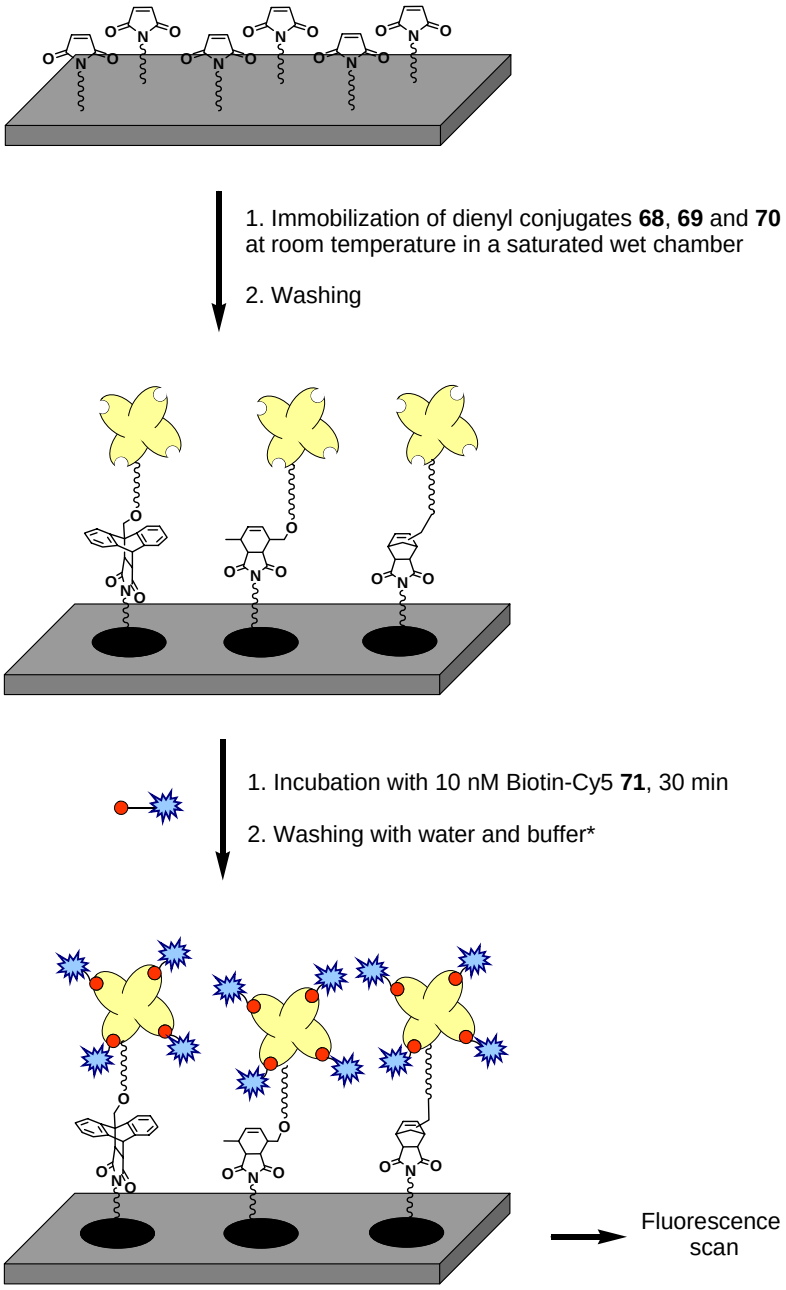


Scheme 33. Synthesis of the Cy5 labeled biotin.

The avidin array was prepared as demonstrated in Scheme 34. In the first experiments, the protein solutions were spotted manually on the maleimido-derivatized slide using an Eppendorf pipette (0.5-10μL capacity). After incubation of the spotted slide inside a saturated wet chamber at room temperature, the excess of protein solution was removed by rinsing the glass surface with water. The slide was then treated with a 10nM solution of biotin-Cy5 **71** for 30 minutes and subsequently washed extensively with water and buffer containing detergent. The degree of protein immobilization was evaluated by screening the slide for fluorescence.

The avidin conjugate solutions were initially spotted at different concentrations onto maleimido-modified slides and incubated overnight (Figure 39A). As a negative control, a solution of wild type avidin was also spotted at the same concentration on the same slide. As shown in Figure 39A, the intensity of the fluorescence signal is consistently influenced by the protein concentration for all three dienyl conjugates. The level of protein immobilization is also dependent on the incubation time (Figure 39A *vs* 39B).

In all cases, the cyclopentadienyl conjugate displayed higher reactivity as expected, giving a signal-to-noise ratio of 8.3 at the conditions indicated in Figure 40. The level of immobilization of the other dienyl conjugates has also showed significant ratio values of 4.4 for the anthracene **68** and 3.4 for the hexadiene **69**. These conjugates were stable enough for efficient protein immobilization within one week of use (longer times were not investigated).



Scheme 34. Immobilization of the dienyl-avidin conjugates on the maleimido-derivatized glass slides and interaction with biotinylated Cy5. * Washing buffer: 10mM sodium phosphate pH 7.5 + 0.05% Tween-20.

As indicated in Figure 39C, it was important to wash the spotted slide surface with a buffer containing detergent in order to lower the level of unspecific binding of the biotin-Cy5. To avoid unspecific binding of avidin to the maleimido-coated surface, it was essential to prevent the spots to be dried out during protein incubation as illustrated in Figure 39D.

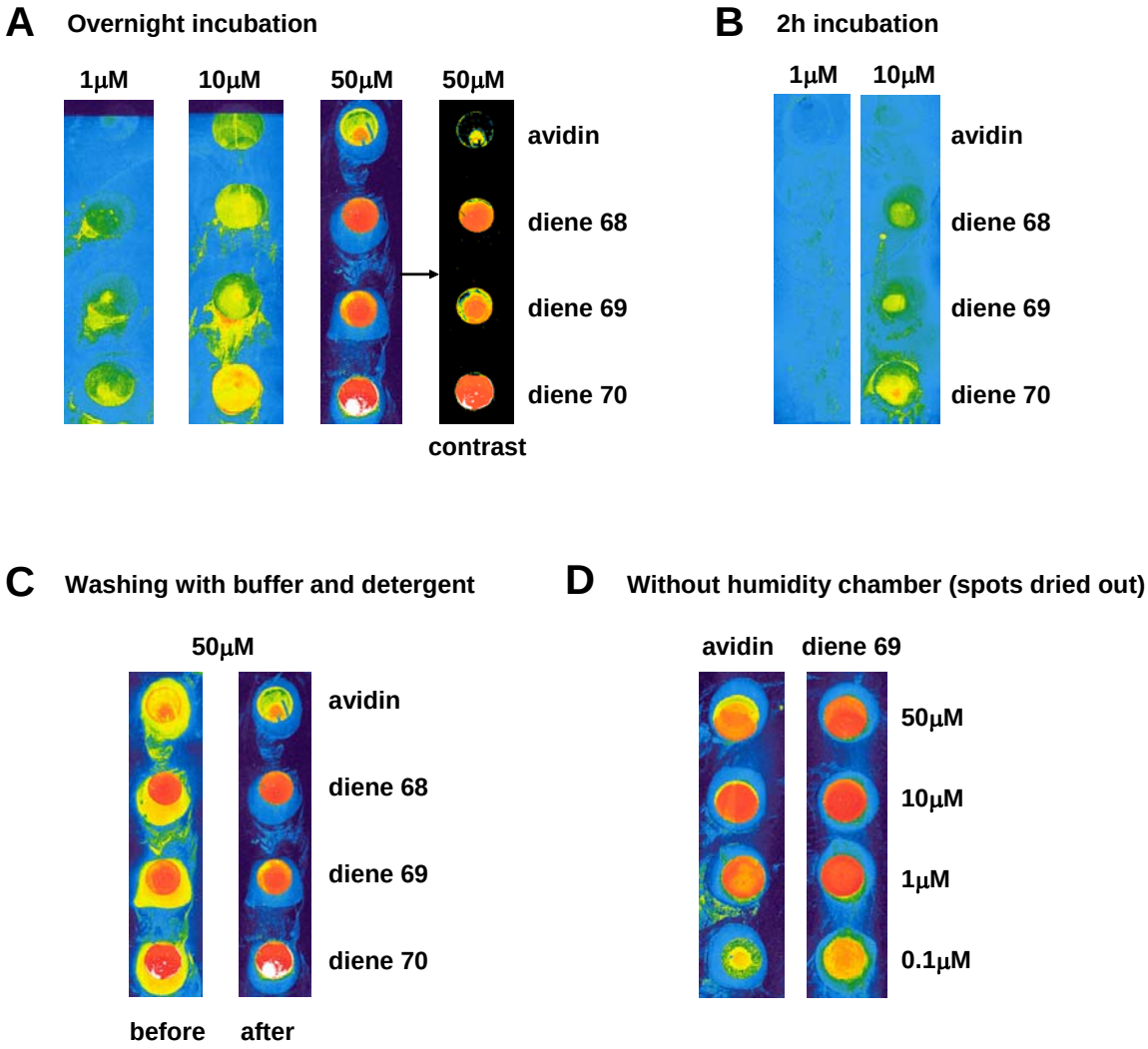


Figure 39. Fluorescence image of the avidin arrays after binding with biotin-Cy5. 5µL of the avidin solutions (in water) were spotted at the concentration shown. The biotin-Cy5 solution was added over each protein spot (25µL for each spot) or added throughout the entire glass slide surface (1mL). Slides were incubated with avidin solutions inside a humidity chamber (**A**, **B** and **C**) or not (**D**). After binding of biotin-Cy5, the slides were washed only with water (first slide in **C**) or with: water, 10mM sodium phosphate pH 5 buffer containing 0.05% Tween-20 and again water (**A**, **B**, **D** and second slide in **C**).

These first results have shown that the Diels-Alder reaction can potentially be applied for the preparation of protein arrays. However, many parameters are still to be investigated. The first issue to be optimized is the level of unspecific binding that still needs to be diminished in order to ensure the chemoselectivity of the protein immobilization. Moreover

the protein samples should be spotted at pico-scale volume using a spotting robot, for a more realistic model of a protein array. These issues are currently being addressed. Once the conditions for an effective chemoselective immobilization are established, the Diels-Alder reaction will be exploited for the site-specific immobilization of proteins in combination with the method described in Section 4.2.3. Future applications of this approach are intended for the preparation of Rab protein microarrays.

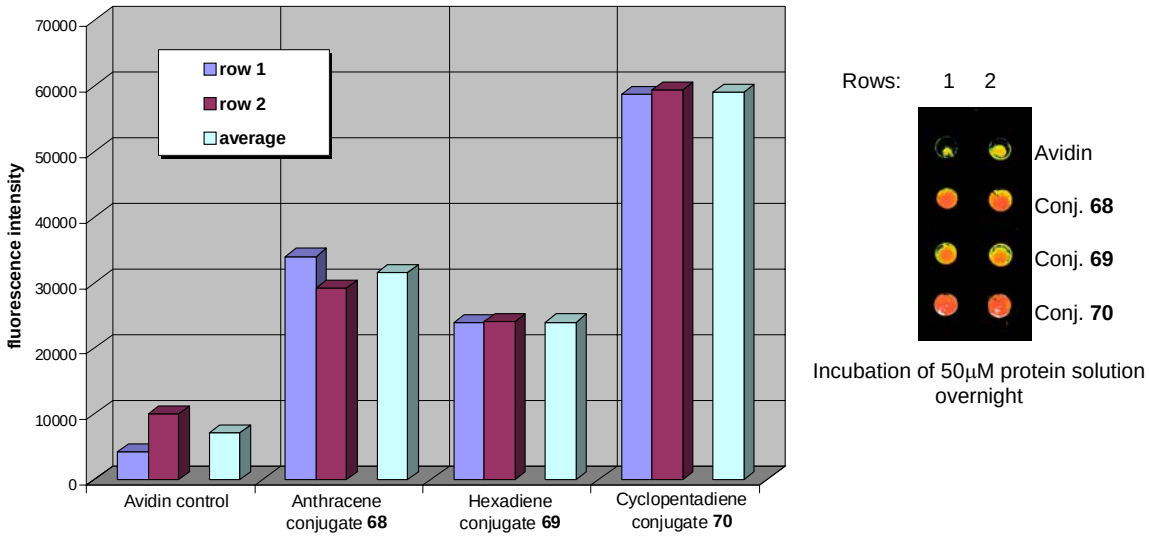


Figure 40. Fluorescence intensity of the proteins immobilized on the glass slide after binding with biotin-Cy5. Avidin and dienyl conjugates were spotted at 50μM concentration and incubated overnight (in duplicate).

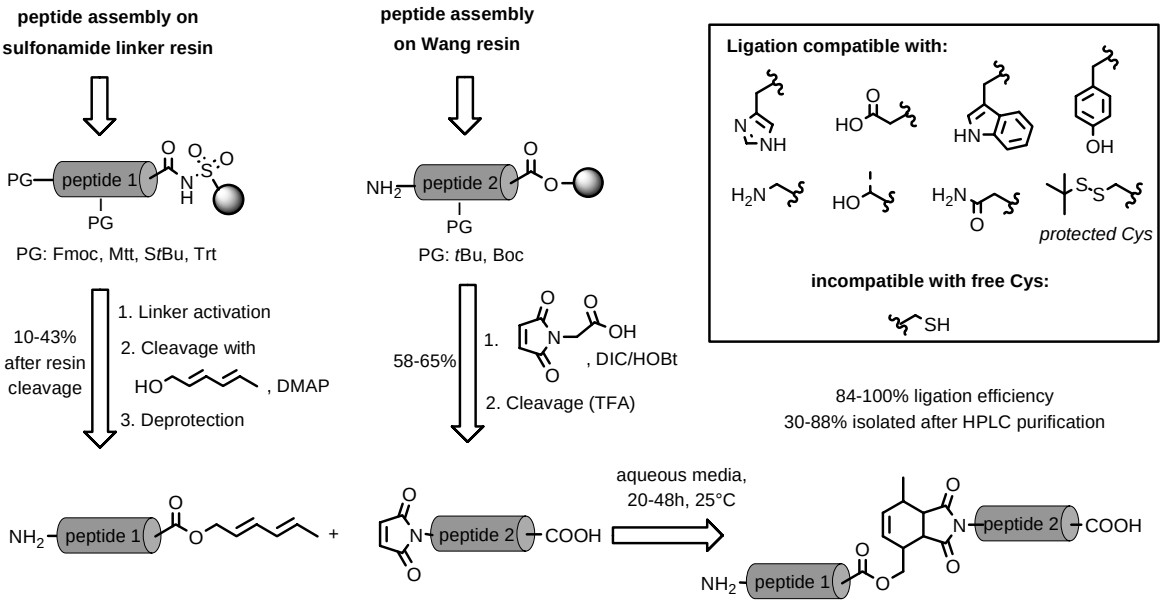
5. SUMMARY AND CONCLUSIONS

Summary

Diels-Alder Ligation of Peptides and Proteins

Over the last years, chemical ligation methods have been shown to be valuable tools for the study of protein function, opening up opportunities to selectively tailor a target protein with a range of natural and unnatural modifications for subsequent biochemical and biological investigations. In order to expand the spectrum of application of these methods, there is still a strong demand for the development of new chemical ligation approaches, providing alternatives to the already established techniques. The chemistry required must be compatible with the functional groups found in proteins and proceed chemoselectively under mild aqueous conditions. The Diels-Alder reaction is such a highly chemoselective transformation and it often proceeds faster and with higher selectivity in water than in organic medium.

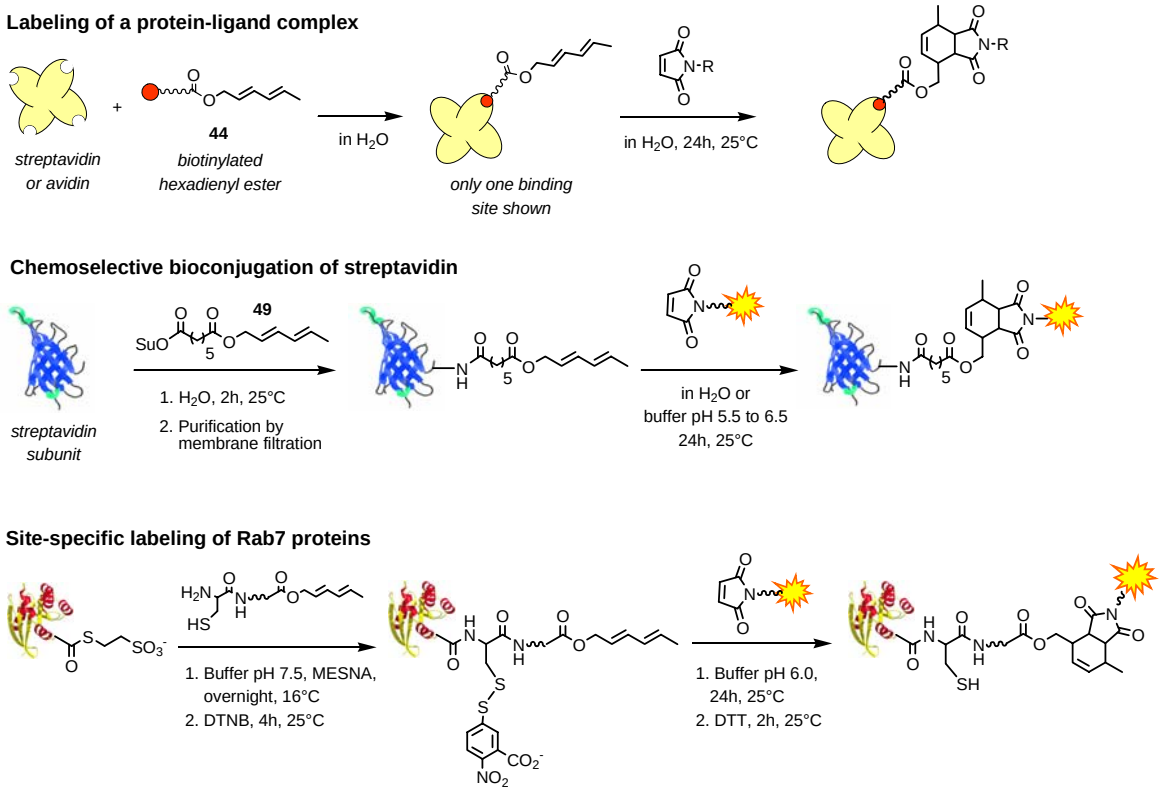
This work demonstrated that the Diels-Alder reaction is a suitable tool for chemical ligation reactions involving peptides and entire functional proteins. At first, several peptides composed of up to ten amino acid residues were synthesized by solid-phase synthesis and equipped with a diene group (*trans*,*trans*-2,4-hexadienyl) at the C-terminus or a dienophile group (maleimide) at the N-terminus (Scheme 35). The hexadienyl ester peptides were synthesized using the safety-catch strategy on sulfonamide resin, representing an advantageous example of application of this strategy to the preparation of peptide esters.



Scheme 35. Synthesis and Diels-Alder ligation of the hexadienyl- and maleimido-peptides.

Diene- and dienophile-peptides were mixed in water at room temperature and the desired cycloaddition proceeded smoothly with 84-100% ligation efficiency after 20-48h reaction (Scheme 35). The Diels-Alder ligation process was fully compatible with the nucleophilic side chains incorporated into the amino acids lysine, histidine and tyrosine. However, free cysteine SH-groups underwent conjugate addition to the maleimide group and, therefore, they had to be protected during the course of the Diels-Alder ligation. Spectroscopic analysis of small peptidyl cycloadducts demonstrated that the Diels-Alder ligation of peptide-derivatized hexadiene and maleimides is *endo*-selective and may proceed with a high degree of stereoselectivity depending on the amino acid sequence.

The stability of the hexadiene function in aqueous solutions and its compatibility with all functional groups present in proteins was explored for combination of the Diels-Alder ligation method with other conjugation techniques to generate tailor-made proteins. In this strategy a particular protein was initially functionalized with a diene unit and then the resulting protein-derived diene was further modified by Diels-Alder reactions with different dienophiles under very mild conditions. Three different approaches were employed to incorporate the diene function into proteins (Scheme 36).



Scheme 36. Strategies for Diels-Alder ligation of proteins.

In the first protein ligation model involving the labeling of a protein-ligand complex, streptavidin and avidin were chosen as model proteins because they do not embody reactive cysteine residues in their structure. A biotinylated hexadienyl ester **44** was synthesized and combined with streptavidin or avidin to form a stable protein-diene complex, which in turn was labeled with maleimido-derivatized peptides by means of the Diels-Alder cycloaddition. In the second approach, the Diels-Alder ligation was employed for the bioconjugation of streptavidin with fluorescent probes. To this end, streptavidin was acylated with the bifunctional linker **49** at lysine side chains to covalently introduce the hexadienyl unit into the protein molecule under controlled conditions. Subsequently the diene conjugate was combined with dansyl- or fluorescein-derivatized maleimides in aqueous solution for 24h at 25°C to yield fluorescent protein with high efficiency. Further investigations using this model showed that the Diels-Alder ligation is only chemoselective under slightly acidic conditions being better performed at pH 5.5 to 6.5 or in pure water.

Having established that the cycloaddition between the hexadiene and maleimide partners proceeded efficiently under the above mentioned conditions, the proposed Diels-Alder ligation was successfully implemented for the site-specific labeling of Rab7 proteins. Initial functionalization of Rab7 with the hexadiene unit was achieved by means of expressed protein ligation. To this end, recombinant truncated Rab7 thioester was ligated with a cysteinyl peptide hexadienyl ester under reducing conditions. The ligation process generated a nucleophilic cysteine side chain and the Rab protein embodies one further accessible cysteine. To avoid undesired modification of these mercapto groups in the subsequent reaction with the maleimide, the exposed cysteine side chains were protected as disulfide by treatment with Ellmann`s reagent immediately after the ligation reaction (Scheme 36). The masked hexadienyl protein was purified and subjected to Diels-Alder ligation with two different dansylated dienophile-derived compounds at pH 6. The coupling reactions were terminated by addition of excess dithiothreitol which traps the dienophile and simultaneously converts the disulfides into unmasked thiols. As indicated by mass spectrometry and gel electrophoresis analysis, the hexadienyl protein was converted into the desired fluorescently labeled cycloaddduct protein after incubation with the maleimide probes (100-fold) during 24h at room temperature. Unspecific multiple labeling of Rab7 could not be detected.

The applicability of the Diels-Alder ligation for the construction of protein microarrays was demonstrated for the immobilization of dienyl conjugated proteins onto maleimide-coated glass slides. As a simple model system, avidin was conjugated with a hexadiene, a cyclopentadiene or an anthracene unit under controlled conditions and spotted

onto the dienophile-functionalizated surface. The level of immobilized protein was detected after binding with fluorescently labeled biotin **71** (biotin-Cy5). Using avidin concentrations of 1 to 50μM, the first model arrays showed considerable signal-to-noise ratio when the protein incubation was performed under saturated humidity atmosphere conditions (spots not dried).

In conclusion this work has shown that the Diels-Alder ligation offers a new opportunity for the site-selective functionalization of proteins and peptides. The reaction partners of the Diels-Alder ligation, diene and dienophile, can be viewed as two entities which simply bind together when held in sufficiently proximity (driven by hydrophobic interactions) without need of any extra reagent or catalyst. The Diels-Alder coupling involving the maleimide segment should be carried out at slightly acidic conditions and in absence of reactive mercaptan moieties or other groups of similar nucleophilicity. If the protein possesses reactive cysteine residues in its structure, the ligation conditions must be manipulated to prevent unspecific reactions by carrying out a simple temporary blocking of the sulfhydryl groups. Although the Diels-Alder reaction leads to a non-traceless ligation site, the final cycloadduct skeleton cycloadduct is relatively small and should not significantly alter the protein structure. The 2,4-hexadiene moiety is stable under physiological conditions and can be easily incorporated chemically into biomolecules from the commercially available precursor *trans*,*trans*-2,4-hexadienol. Maleimide derivatives and probes are abundantly available from commercial suppliers. Other potent diene functions, such as cyclopentadiene, may also be appropriate employed for some particular applications, although their instability should be considered.

Combining the results herein described and the previous outcomes from the bioconjugation of oligonucleotides and saccharides, the Diels-Alder [4+2] cycloaddition has proven to be a feasible method to promote covalent biomolecule modification and a complementary alternative for the *in vitro* assembly of semisynthetic proteins and biopolymers.

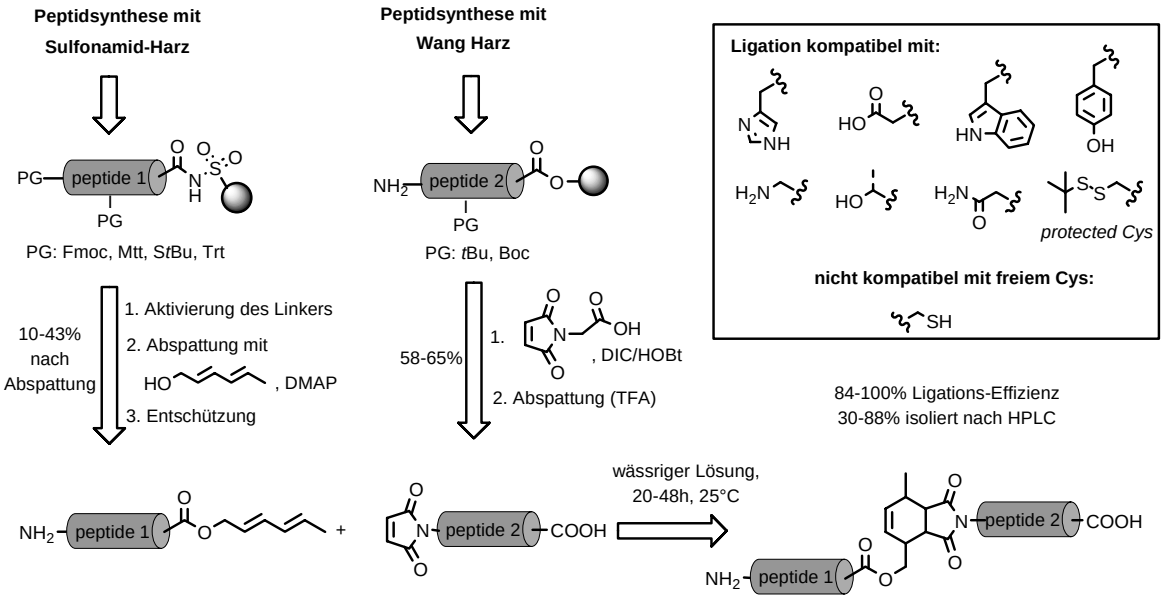
Zusammenfassung

Diels-Alder Ligation von Peptiden und Proteinen

In den letzten Jahren wurde gezeigt, dass Methoden der chemischen Ligation wertvolle Hilfsmittel sind, um Proteinfunktion zu studieren, indem sie Möglichkeiten eröffnen, selektiv ein Zielprotein mit einer Reichweite von natürlichen und unnatürlichen Modifikationen für darauf folgende biochemische und biologische Untersuchungen maß zu schneidern. Um das Spektrum der Anwendungen dieser Methoden zu erweitern, gibt es immer noch einen starken Bedarf für die Entwicklung neuer Ansätze der chemischen Ligation, als Alternativen zu den bereits etablierten Techniken. Die Chemie muss mit den in Proteinen vorkommenden funktionellen Gruppen kompatibel sein und chemoselektiv unter milden wässrigen Bedingungen verlaufen. Die Diels-Alder Reaktion ist solch eine hoch chemoselektive Transformation und verläuft in wässrigem Medium oft schneller und mit höherer Selektivität als im organischen Medium.

Diese Arbeit zeigte, dass die Diels-Alder Reaktion ein geeignetes Hilfsmittel für chemische Ligations-Reaktionen ist, die Peptide und gesamte funktionelle Proteine einschließt. Um zunächst die Tauglichkeit der Diels-Alder Cycloaddition für Peptid-Ligation zu untersuchen, wurden unterschiedliche Peptide bestehend aus bis zu zehn Aminosäuren an der festen Phase synthetisiert und mit einer Dien-Gruppe (*trans,trans*-2,4-hexadienyl) am C-Terminus oder einer Dienophil-Gruppe (maleinimid) am N-Terminus ausgestattet (Schema 37). Die Hexadienylester-Peptide wurden mit der safety-catch Strategie an Sulfonamid-Harz synthetisiert, was bis heute eines der sehr wenigen Beispiele der Anwendung dieser Strategie zur Darstellung von Peptidestern repräsentiert.

Die Dien- und Dienophil-Peptide wurden in Wasser bei Raumtemperatur gemischt und die erwünschte Cycloaddition verlief reibungslos mit 84-100% Ligations-Effizienz innerhalb 20-48 Stunden (Schema 37). Der Diels-Alder Ligations-Prozess war vollständig kompatibel mit den nucleophilen Seitenketten der Aminosäuren Lysin, Histidin, Tyrosin, usw. Die freien Cystein-SH-Gruppen unterlagen jedoch der konjugierten Addition an die Maleinimid-Gruppe, weshalb sie während des Verlaufs der Diels-Alder Reaktion geschützt werden mussten. Die spektroskopische Analyse kleiner Peptidyl-Cycloprodukte zeigte, dass die Diels-Alder Reaktion von Peptid-derivatisiertem Hexadien und Maleinimid mit einem hohen Grad an *endo*-Stereoselektivität verläuft.

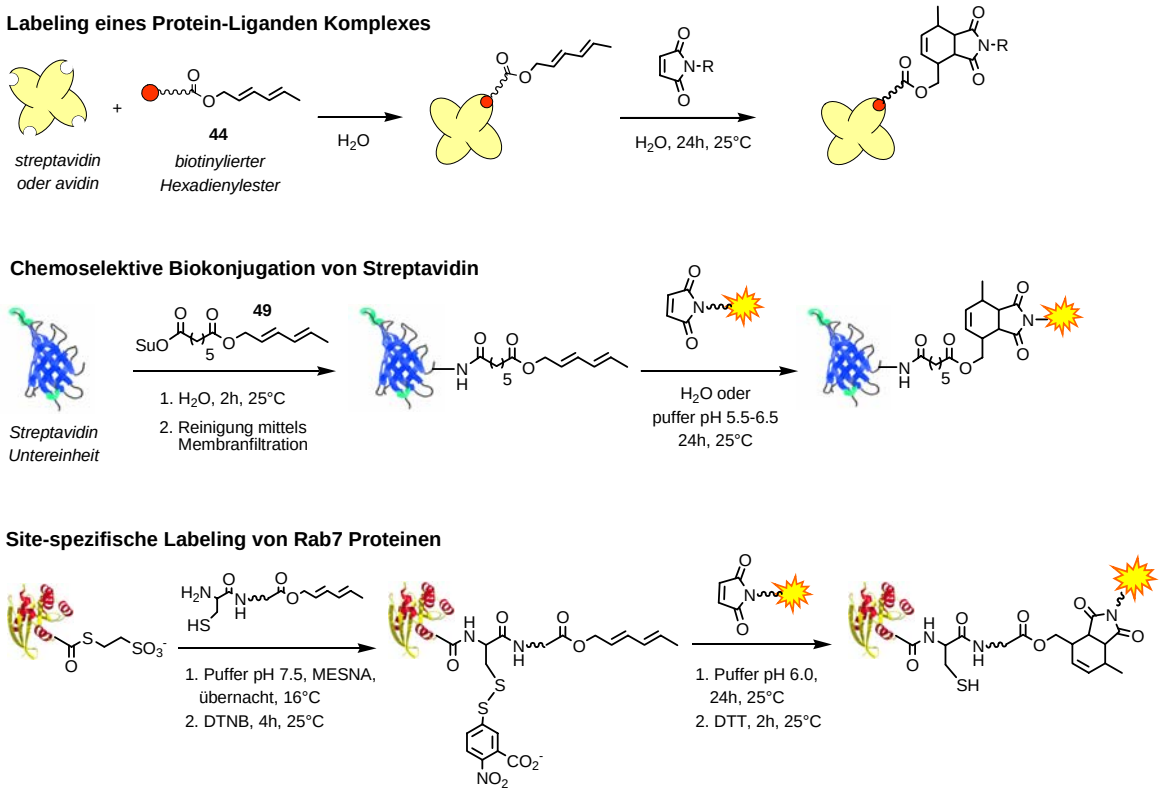


Schema 37. Synthese und Diels-Alder Ligation der Hexadienyl- und Maleinimido-Peptide

Die Stabilität der Hexadien-Einheit in wässriger Lösung und ihre Kompatibilität mit allen in Proteinen vorkommenden funktionellen Gruppen wurden dahingehend erforscht, die Diels-Alder Ligationsmethode mit anderen Konjugationstechniken zur Generierung maßgeschneiderter Proteine zu kombinieren. In dieser Strategie wurde ein bestimmtes Protein zunächst mit einer Dien-Einheit funktionalisiert und dann das resultierende Protein-abgeleitete Dien durch Diels-Alder Reaktion mit unterschiedlichen Dienophilen unter sehr milden Bedingungen weiter modifiziert.

Drei unterschiedliche Ansätze wurden verwendet, um Proteine mit der Dien-Funktion auszustatten (Schema 38). Das erste Protein Ligations-Modell umfasst das Labeling eines Protein-Liganden Komplexes durch eine Diels-Alder Reaktion. Zur Vereinfachung wurden Streptavidin und Avidin als Modell-Proteine gewählt, da sie keine reaktiven Cystein-Seitenketten in ihrer Struktur enthalten. Ein biotinylierter Hexadienylester **44** wurde synthetisiert und mit Streptavidin oder Avidin kombiniert, um einen stabilen Protein-Dien Komplex zu formen, der wiederum mit Maleinimido-derivatisierten Peptiden durch Diels-Alder Reaktion gelabelt wurde. Im zweiten Ansatz wurde die Diels-Alder Reaktion für die Biokonjugation von Streptavidin mit Fluoreszenz-Sonden eingesetzt. Zunächst wurde Streptavidin an Lysin-Seitenketten mit dem bifunktionellen Linker **49** acyliert, um die Hexadienyl-Einheit kovalent in das Protein unter kontrollierten Bedingungen einzuführen. Anschließend wurde das Dien-Konjugat mit Dansyl- oder Fluorescein-derivatisierten Maleinimiden in wässriger Lösung für 24 h bei 25 °C kombiniert, um das fluoreszente Protein

in hoher Effizienz zu erhalten. Weitere Untersuchungen mittels dieses Modells zeigten, dass die Diels-Alder Reaktion nur unter schwach sauren Bedingungen, am besten bei pH 5.5 bis 6.5 oder in reinem Wasser, chemoselektiv verläuft.



Schema 38. Strategien der Diels-Alder Ligation von Proteinen

Einmal etabliert, dass die Cycloaddition zwischen den Hexadien- und Maleinimid-Partnern unter den oben genannten Bedingungen effizient verläuft, wurde die vorgeschlagene Diels-Alder Ligation erfolgreich für das Site-spezifische Labeling von Rab7 Proteinen als ein repräsentatives biologisch relevantes Beispiel zur Anwendung gebracht. Die anfängliche Funktionalisierung von Rab7 mit der Hexadien-Einheit wurde durch Expressed-Protein-Ligation erreicht. Zu diesem Zweck wurde ein rekombinanter, verkürzter Rab7-Thioester mit einem Cysteinyl-Peptid Hexadienylester unter reduzierenden Bedingungen ligiert. Der Ligations-Prozess generiert eine nucleophile Cystein-Seitenkette und das Rab-Protein enthält ein weiteres zugängliches Cystein. Um ungewünschte Modifikationen dieser Mercaptogruppen in den folgenden Reaktionen mit dem Maleinimid zu verhindern, wurden die zugänglichen Cystein-Seitenketten als Disulfid durch Behandlung mit Ellman's Reagenz direkt nach der Ligations-Reaktion geschützt, um den geschützten Proteinester zu erhalten (Schema 38).

Dieses modifizierte Protein wurde aufgereinigt und der Diels-Alder Ligation mit zwei unterschiedlichen dansylierten Dienophil-abgeleiteten Verbindungen bei pH 6 unterworfen. Die Kupplungs-Reaktionen wurden durch Zugabe von Überschuss Dithiothreitol, welches das Dienophil abfängt, beendet und simultan die Disulfide in unmaskierte Thiole überführt. Wie durch Massenspektrometrie und Gel-Elektrophorese angedeutet, wurde das meiste Hexadienyl-Protein in das gewünschte fluoreszenz-markierte Cycloaddukt-Protein nach Inkubation mit der Maleinimid-Sonde (100fach) während 24 h bei Raumtemperatur umgesetzt. Unspezifisches mehrfaches Labeling von Rab7 konnte nicht detektiert werden.

Die Anwendbarkeit der Ligation via Diels-Alder Reaktion für den Aufbau von *Protein-Microarrays* wurde für die Immobilisierung von mit Dienyl-Einheiten verbundenen Proteinen auf Maleimid-bedeckten Glass-Trägern demonstriert. Als ein einfaches Modell-Beispiel wurde Avidin mit einer Hexadien-, Cylopentadien- oder Anthraceneinheit unter kontrollierten Bedingung verbunden uns auf die mit der Dienophil-funktionalisierten Oberfläche platziert. Die Menge an immobilisiertem Protein wurde nach der Umsetzung mit Biotin-Cy5, einem fluoreszenzmarkiertem Ligand, detektiert. Unter Verwendung von Avidin-Konzentrationen von eins bis 50 μM zeigten die ersten Modellversuche ein ansehnliches Signal-Rausch-Verhältnis, wenn die Einwirkung des Proteins in einer mit Feuchtigkeit gesättigten Atmosphäre stattfand.

Zusammenfassend hat diese Arbeit gezeigt, dass die Diels-Alder Ligation eine neue Möglichkeit für die Site-spezifische Funktionalisierung von Proteinen und Peptiden darstellt. Die Reaktionspartner der Diels-Alder Ligation, Dien und Dienophil, können als zwei Einheiten betrachtet werden, die sich einfach ohne Bedarf irgendwelcher zusätzlicher Reagenzien oder Katalysatoren miteinander vereinigen, wenn sie in ausreichende Nähe gebracht werden (getrieben durch hydrophobe Wechselwirkungen). Die Diels-Alder Kupplung mit dem Maleinimid-Segment sollte unter leicht sauren Bedingungen in Abwesenheit reaktiver Mercaptan-Einheiten oder anderer Gruppen ähnlicher Nucleophilie durchgeführt werden. Wenn das Protein reaktive Cystein-Seitenketten in seiner Struktur besitzt, können die Ligationsbedingungen durch einfaches temporäres Schützen der Sulfhydryl-Gruppen durch Disulfid-Bildung mit Ellman's Reagens variiert werden, um unspezifische Reaktionen zu verhindern. Obwohl die Diels-Alder Ligation zu einer nicht spurlosen Ligationsstelle führt, ist das endgültige Cycloaddukt-Grundgerüst relativ klein und sollte die Proteinstruktur nicht signifikant verändern. Die 2,4-Hexadien Einheit ist in physiologischem Medium stabil und kann mittels des kommerziell erhältlichen Precursor *trans,trans*-2,4-Hexadienol einfach über chemische Synthese in Biomoleküle eingeführt werden.

Maleinimid-Derivate und -Sonden sind in einer Vielzahl von kommerziellen Anbietern verfügbar. Andere wirksame Dien-Funktionen, so wie Cyclopentadien, können auch für einige spezielle Anwendungen geeignet eingesetzt werden, obwohl ihre Instabilität berücksichtigt werden sollte.

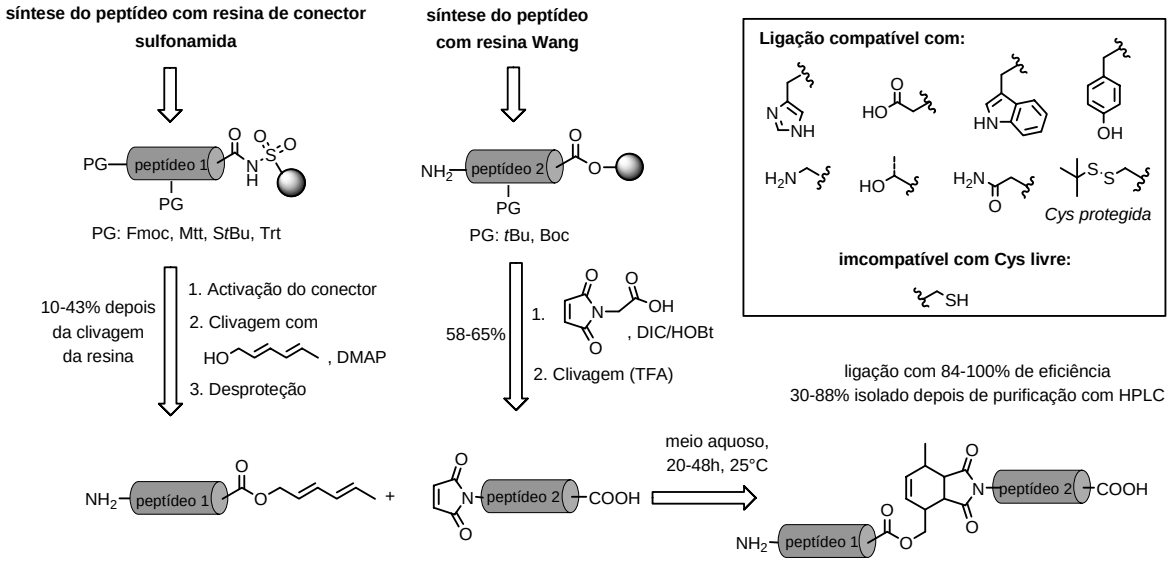
In Kombination der hier beschriebenen Ergebnisse mit vorherigen Resultaten der Biokonjugation von Oligonucleotiden und Sacchariden wurde gezeigt, dass die Diels-Alder [4+2]-Cycloaddition eine brauchare Methode ist, um kovalente Biomolekül-Modifikationen durchzuführen und eine komplementäre Alternative für die *in vitro*-Assemblierung von artifiziellen Proteinen und Biopolymeren darstellt.

Resumo

Ligação Diels-Alder de Peptídeos e Proteínas

Chemical ligations são poderosos métodos químicos que vêm sido amplamente aplicados ao longo dos últimos anos para o estudo das funções de proteínas, abrindo novos caminhos para modular seletivamente suas estruturas de modo a facilitar subsequentes investigações bioquímicas ou biológicas. Contudo ainda existe uma forte necessidade em desenvolver novos tipos de ligações, de modo a abranger o espectro de aplicações destes métodos e fornecer alternativas para as técnicas já existentes. As reações envolvidas nestes processos devem ser compatíveis com os grupos funcionais presentes em proteínas e proceder quimiosseletivamente em condições suaves e aquosas. As reações de Diels-Alder enquadram-se neste tipo de transformações seletivas e procedem, em geral, mais rapidamente e com maior seletividade em água do que em meio orgânico.

Este estudo demonstrou que as reações de Diels-Alder podem ser adequadamente empregadas para a ligação de peptídeos e de proteínas inteiramente funcionais. Para investigar primeiramente a eficácia destas reações em conectar cadeias de aminoácidos, vários peptídeos compostos de até dez aminoácidos foram sintetizados em fase sólida e funcionalizados com um grupo dieno (*trans*,*trans*-2,4-hexadienila) na posição C-terminal ou com um grupo dienófilo (maleimida) na N-terminal (Esquema 39).

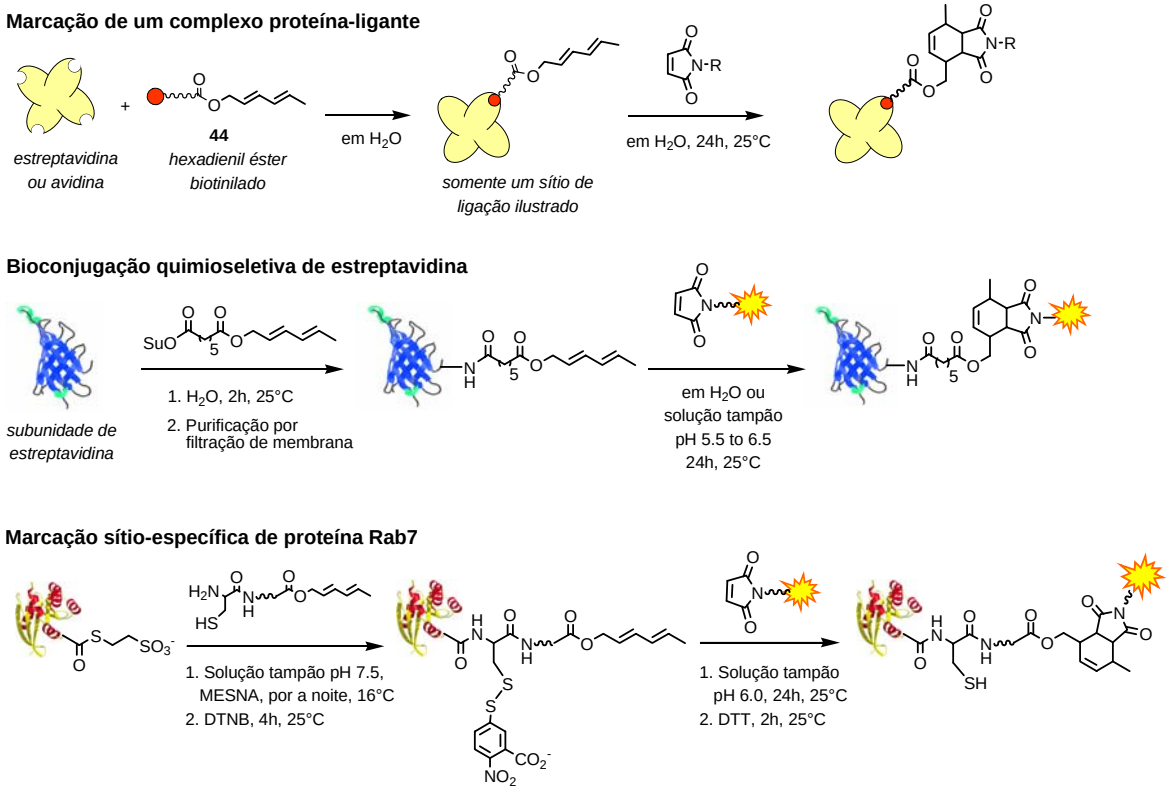


Esquema 39. Síntese e ligação Diels-Alder dos peptídeos de hexadienila e maleimida.

Os ésteres de hexadienila peptídicos foram sintetizados utilizando a estratégia de *safety-catch* em resina sulfonamida, representando um dos poucos exemplos de aplicação desta metodologia para a preparação de ésteres.

Peptídeos contendo as porções de dieno e de dienófilo foram combinados em água à temperatura ambiente e a esperada ligação por cicloadição ocorreu com eficiência de 84-100% depois de 20-48h de reação (Esquema 39). A reação de Diels-Alder mostrou ser totalmente compatível com as cadeias laterias de aminoácidos tais como lisina, histidina e tirosina. Contudo, os grupos –SH livres de resíduos de cisteínas sofreram adição conjugada com o grupo maleimida e, por isso, devem ser protegidos durante o curso da ligação de Diels-Alder. As análises espectroscópicas dos cicloprodutos de peptídeos mais simples demonstraram que a ligação de Diels-Alder é *endo*-seletiva e pode proceder com alto nível de esterosseletividade.

A estabilidade da função hexadieno em soluções aquosa e a sua compatibilidade com todos grupos funcionais encontrados em proteínas foram exploradas para a combinação da ligação Diels-Alder com outras técnicas de conjugação visando a produção de proteínas modificadas. Nesta estratégia, uma dada proteína é inicialmente funcionalizada com um grupo dieno e logo o derivado resultante é modificado com diferentes dienófilos sob condições suaves. Como ilustrado no Esquema 40, três procedimentos distintos foram empregados para incorporar a função dieno em proteínas.



Esquema 40. Estratégias de ligação Diels-Alder de proteínas.

O primeiro modelo de ligação de proteína envolveu a marcação de um complexo proteína-ligante através da reação de Diels-Alder. Por razões de simplicidade, estreptavidina e avidina foram escolhidas como proteínas modelo pois estas não contêm resíduos reativos de cisteína em suas estruturas. O éster de hexadienila biotinilado **44** foi sintetizado e combinado com streptavidina ou avidina para formar um complexo estável de proteína-dieno, que foi então marcado por meios da ligação Diels-Alder. No segundo método, a ligação Diels-Alder foi utilizada para a bioconjugação de streptavidina com grupos fluorescentes. Inicialmente, a streptavidina foi acilada com um conector bifuncional para introduzir covalentemente a unidade hexadieno na molécula de proteína sob condições controladas. Em seguida, o conjugado de dieno foi combinado com compostos de maleimida (incorporados com os grupos fluorescentes dansila ou fluoresceína) em solução aquosa a 25°C por 24h, produzindo uma proteína fluorescente com alta eficiência. Investigações adicionais deste modelo mostraram que a ligação Diels-Alder somente é quimiosseletiva sob condições ligeiramente ácidas sendo melhor conduzida em pH 5.5 a 6.5 ou em água pura.

Visto que a cicloadição entre os parceiros hexadieno e maleimida procedeu eficientemente sob as condições mencionadas, a ligação Diels-Alder foi implementada para a marcação sítio-específica de proteínas Rab7 como um exemplo biologicamente relevante de aplicação deste método. A funcionalização inicial da proteína Rab7 com o grupo hexadieno foi adquirida por meio do método EPL (*expressed protein ligation*). Para esse fim, um tioéster recombinante de Rab7 foi ligado, sob condições redutivas, com um peptídeo de éster de hexadienila que possui uma cisteína N-terminal. O processo de ligação gera um resíduo de cisteína e a proteína Rab7 incorpora uma outra cisteína que também é acessível a reações com o solvente. Para evitar a modificação indesejada desses grupos tióis em reações subsequentes com maleimidas, as cadeias laterais das cisteínas livres foram protegidas através da formação de dissulfeto pelo tratamento com o reagente de Ellmann imediatamente após a reação de ligação. A resultante proteína de hexadienila mascarada foi purificada e submetida à ligação Diels-Alder em solução tamponada a pH 6.0 com dois diferentes dienófilos marcados com o grupo fluorescente dansila. As reações de acoplamento foram terminadas pela adição de excesso de ditiotreitol que captura o dienófilo e simultaneamente desbloqueia os grupos tióis. Baseado em análises de espectrometria de massa e eletroforese em gel, comprovou-se que a maioria da proteína de hexadienila foi convertida no cicloproduto fluorescente desejado depois de sua incubação com os reagentes fluoróforos de maleimida durante 24h à temperatura ambiente. Marcação múltipla não-específica da proteína Rab7 não foi detectada.

A aplicabilidade da ligação Diels-Alder para a construção de *microarrays* de proteínas foi demonstrada para a imobilização de proteínas conjugadas de dienila sobre placas de vidro funcionalizadas com grupos dienófilo. Em um modelo simples, a avidina foi conjugada com um grupo hexadienila, um ciclopendienila ou um antraceno sob condições controladas e então fixada sobre um superfície vítrea modificada com grupos maleimida. O nível de proteína imobilizada foi detectado após complexação da avidina com um composto de biotina marcada com o fluoróforo Cy5. Usando concentrações de avidina entre 1 e 50μM, níveis significantes de imobilização seletiva dos conjugados de dienila foram obtidos quando a incubação da proteína foi realizada sob condições saturadas de humidade.

Em conclusão este estudo mostrou que a ligação de Diels-Alder oferece uma nova oportunidade para a funcionalização sítio-seletiva de proteínas e peptídeos. Os parceiros desta ligação, dieno e dienófilo, podem ser vistos como duas entidades que se combinam facilmente quando colocados em suficiente proximidade (dirigida por interações hidrofóbicas) sem precisar de nenhuma adição extra de reagentes ou catalisadores. O acoplamento de Diels-Alder envolvendo o segmento maleimida deve ser conduzido em meio ligeiramente ácido e na ausência de grupos mercaptanos reativos ou de qualquer outro grupo de igual nucleofilicidade. Caso a proteína possua resíduos reativos de cisteína, as condições da ligação podem ser manipuladas para prevenir reações não-específicas através de um simples bloqueio temporário dos grupos tióis através da formação de dissulfetos com reagente de Ellmann.

Apesar da reação de Diels-Alder gerar um grupo não-natural na molécula de proteína, o esqueleto do cicloproduto final é relativamente pequeno e não deve alterar significamente a sua estrutura. A função 2,4-hexadieno é estável em meio fisiológico e pode ser facilmente incorporada em proteínas a partir do precursor *trans*,*trans*-2,4-hexadienol. Os derivados de maleimida são abundantemente disponíveis comercialmente. Outras funções dieno como o ciclopentadieno e antraceno podem também ser apropriadamente empregadas para algumas aplicações, embora suas instabilidades em meio aquoso devem ser levadas em consideração.

Combinando os resultados aqui descritos com aqueles envolvendo a bioconjugação de oligonucleotídeos e sacarídeos reportados anteriormente, a cicloadição de Diels-Alder provou em ser um método adequado para ser empregado em modificações covalentes de biomoléculas e uma alternativa complementar para a construção *in vitro* de proteínas e outros biopolímeros artificiais.

6. EXPERIMENTAL SECTION

6.1. Materials, instruments and general methods for purification and analyses

Reagents

The reagents were purchased from Acros Chimica, Advanced Chemtech, Aldrich, Avocado, Bachem, Fluka, J.T. Baker, Merck, Novabiochem, Riedel de Haen, Roth, Senn Chemicals or Sigma. All solvents, when not purchased with necessary purity or dryness, were distilled using standard methods. Deionized water (Millipore Q-plus System) was used for all experiments.

Silica gel flash liquid chromatography

Purifications were performed using silica gel from J. T. Baker or Merck (particle size 40-60μm) under approximately 0.5 bar pressure.

Preparative reversed-phase high performance liquid chromatography (prep HPLC)

Purification of the compounds was performed on an Agilent preparative HPLC 1100 Series system using Nucleodur C18 Gravity column (Macherey-Nagel) and detection at 215 and 254nm. Linear gradients of solvent B (0.1% TFA in acetonitrile) in solvent A (0.1% TFA in water) were used at 25mL/min flow rate.

Nuclear magnetic resonance spectroscopy (NMR)

^{1}H- and ^{13}C-NMR spectra were recorded using a Varian Mercury 400 spectrometer (400MHz (^{1}H) and 100.6MHz (^{13}C)). Chemical shifts are expressed in parts per million (ppm) from internal trimethylsilane standard. Coupling constants (*J*) are given in Hertz (Hz) and the following notations indicate the multiplicity of the signals: s (singlet), d (doublet), t (triplet), dd (doublet of doublet), ddd (doublet of doublet of doublet), m (multiplet), br (broad signal).

Mass spectrometry (MALDI-TOF, ESI-MS and FAB-HR/LR)

Matrix assisted laser desorption ionization time-of-flight (MALDI-TOF) measurements were carried out with Voyager-DE Pro Biospectrometry Workstation from PerSeptive Biosystems using α-cyano-hydroxycinnammic acid (unless otherwise stated) as matrix. Electrospray mass spectrometric analyses (ESI-MS) were performed on a Finnigan LCQ spectrometer. Fast atom

bombardment (FAB) mass spectra were recorded on a Finnigan MAT MS 70 spectrometer, using m-nitrobenzylalcohol as matrix. Calculated masses were obtained using the program ChemDraw Ultra (CambridgeSoft Corporation).

Reversed-phase liquid chromatography – electronspray ionization mass spectrometry (LC-MS)

LC-MS measurements were carried out on a Hewlett Packard HPLC 1100/Finnigan LCQ mass spectrometer system using Nucleodur C18 Gravity, Nucleosyl 100-5 C18 Nautilus (Macherey-Nagel) or Jupiter C4 (Phenomenex) columns and detection at 215 and 254nm. Linear gradients of solvent B (0.1% formic acid in acetonitrile) in solvent A (0.1% formic acid in water) were used at 1mL/min flow rate.

Method A: C18 Nautilus, gradient (time/solvent B): 0min/0%B → 3min/0%B → 23min/70%B.

Method B: C18 Gravity, gradient (time/solvent B): 0min/10%B → 15min/90%B.

Method C: C4 Jupiter, gradient (time/solvent B): 0min/20%B → 35min/50%B.

Analytical reversed-phase high performance liquid chromatography (an. HPLC)

Analyses were performed on a Hewlett Packard HPLC 1100 system using Nucleodur C18 Gravity and Nucleosyl 100-5 C18 Nautilus columns (Macherey-Nagel) and detection at 215 and 254nm. Linear gradients of solvent B (0.1% TFA in acetonitrile) in solvent A (0.1% TFA in water) were used at 1mL/min flow rate.

Method A: C18 Nautilus, gradient (time/solvent B): 0min/0%B → 3min/0%B → 23min/70%B.

Method B: C18 Gravity, gradient (time/solvent B): 0min/10%B → 15min/90%B.

Thin layer chromatography (TLC)

TLC was carried out on Merck precoated silica gel plates (60F-254) using ultraviolet light irradiation at 254nm or the following solutions as developing agents:

Staining solution A: 25g molibdatophosphoric acid and 10g cerium (IV) sulfate in 60mL concentrated sulfuric acid and 940mL water.

Staining solution B (for detection of free amino groups): 300mg ninhydrin in 100mL ethanol and 3mL acetic acid.

Gas chromatography – mass spectrometry (GC-MS)

Spectra were obtained from a Hewlett Packard 6890 GC system coupled to a Hewlett Packard 5973 Mass Selective Detector. A HP 5TA capillary column (0.33μm x 25m x 0.2mm) and helium flow rate of 2mL/min were used.

Method A: temperature gradient: 0min (100°C) → 1min (100°C) → 6min (300°C) → 12min (300°C).

Method B: temperature gradient: 0min (50°C) → 2min (50°C) → 8min (300°C) → 12min (300°C).

Optical rotation

Optical rotations were measured in a Schmidt + Haensch Polartronic HH8 polarimeter at 589 nm. Concentrations are given in g/100mL solvent.

Ultraviolet Spectroscopy

UV measurements were achieved by using a Cary 50 UV/VIS spectrophotometer from Varian.

6.2. Chemical methods

6.2.1. General procedure for the peptide synthesis on solid support

Resins and reactors

Peptides were synthesized manually using glass synthesizers or plastic syringes. Resins were purchased from Novabiochem: Wang resin loading 1.2 mmol/g (200-400 mesh, 1% DVD), 4-sulfamylbutyryl AM resin loading 1.1 mmol/g (200-400 mesh, 1% DVD) and 2-chlorotrityl chloride resin loading 1.6 mmol/g (100-200 mesh, 1% DVD).

Peptides

Table 8. Fmoc-amino acid building blocks used for each solid-phase method.

Fmoc-amino acid	Wang resin	Sulfon amide resin	2-Cl-trityl resin	Fmoc-amino acid	Wang resin	Sulfon amide resin
Fmoc-Ala-OH	x			Fmoc-Lys(Fmoc)-OH		x
Fmoc-β-Ala-OH			x	Fmoc-Lys(Mtt)-OH		x
Fmoc-Cys(StBu)-OH		x	x	Fmoc-Met-OH		x
Fmoc-Gln(Trt)-OH	x			Fmoc-Phe-OH	x	x
Fmoc-Glu(Boc)-OH	x			Fmoc-Pro-OH		x
Fmoc-Gly-OH	x	x	x	Fmoc-Ser(tBu)-OH	x	
Fmoc-His(Trt)-OH	x			Fmoc-Ser(Trt)-OH		x
Fmoc-Ile-OH	x	x		Fmoc-Thr(tBu)-OH	x	
Fmoc-Leu-OH	x	x		Fmoc-Trp(Boc)-OH	x	
Fmoc-Lys(Boc)-OH	x			Fmoc-Tyr(tBu)-OH	x	
Fmoc-Lys(Dansyl)-OH (**46**)	x			Fmoc-Val-OH		x

Semi-quantitative estimation of first amino acid loading on resin

Approximately 1 mg of resin was placed into a 10 mm UV cuvette and treated with 3mL of a 20% piperidine in DMF solution during 5 minutes. A reference sample containing only the piperidine solution (3mL) was also prepared. Using UV spectrophotometer, the absorbance

(Abs) of the samples at 290nm was measured and the Fmoc amino acid loading estimated using the equation: $(Abs_{sample} – Abs_{ref})/(1.65 \times mg \text{ of resin})$.

Peptide chain assembly:

1) Removal of Fmoc protection group. The Fmoc-protected peptidyl resin was treated with a 20% piperidine in DMF solution (10mL/g resin) for 10 minutes. The resin was drained and treatment procedure repeated 1x. Finally, the resin was washed 5 times with DMF.

2) Coupling using HBTU/HOBt activation. The Fmoc amino acid to be coupled (4 equivalents relative to resin loading) was dissolved in dry DMF (approximately 10mL/g resin), followed by addition of HOBt (4 equivalents) and HBTU (4 equivalents). DIPEA (8 equivalents) was lastly included and the resulting solution was added immediately to the peptidyl resin. The mixture was shook at room temperature until the Kaiser test was negative (generally after 1-3h; occasionally some couplings were allowed to react overnight or repeated). Resin was filtered and washed with DMF (5 or more times).

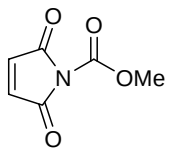
Kaiser test (ninhydrin test) for detection of primary amines

Some beads of resin were removed, washed with DMF, and treated with 2 drops of each of the three Kaiser test solutions A (5g in 100mL ethanol), B (2mL of KCN 1mM and 98mL pyridine) and C (80g phenol in 20mL ethanol). The mixture was heated to circa 120°C using a heating gun. If the beads became blue, free amino groups are present on resin, and thus indicated incomplete coupling. The test could not be used for peptides having Mtt groups in the side chain.

6.2.2. Preparation of the *N*-maleoyl-peptides

6.2.2.1. *Synthesis of the peptide building blocks used for the SPPS*

_N_-(methoxycarbonyl)-maleimide (3)



Maleimide (4.0g, 41mmol) and *N*-methylmorpholine (4.5mL, 41mmol) were dissolved in EtOAc (200mL) and cooled to 0°C. After 30 minutes methyl choroformate (3.2mL, 41mmol)

was added and the reaction turned violet. After 1 hour at 0°C, the precipitate was filtered off and washed with EtOAc. Filtrate and washings were washed 3x with water, dried over Na_2SO_4 and concentrated in vacuum. The product was crystallized from EtOAc/petroleum ether.

Yield: 3.92g (25mmol, 62%) light violet crystals

TLC: R_f = 0.58 (EtOAc)

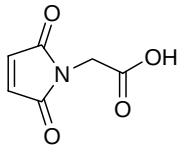
$C_6H_5NO_4$ (155.11)

^{1}H-NMR (D_2O, 400MHz): δ = 3.65 (3H, s, OCH_3), 6.23 (1H, d, *J* = 12.0 Hz, C*H*=CH) and 6.50 (1H, d, *J* = 12.0Hz, CH=C*H*) ppm.

^{13}C-NMR (D_2O, 100MHz): δ = 53.6 (OCH_3), 129.0 (*C*H=CH), 133.2 (CH=*C*H), 153.4 (C=O *methoxy*) and 169.0 (C=O *imide*) ppm.

GC-MS (Method B)**:** t_r = 5.12min; m/z 155 $[M]^+$, 125/124 $[M\text{-}OMe]^+$

N-maleoylglycine (4)



Glycine (375mg, 5mmol) was dissolved in $NaHCO_3$ saturated solution, cooled to 0°C in an ice-bath and *N*-methoxycarbonylmaleimide **3** (775mg, 5mmol) was added. After 10 minutes, the solution was diluted by adding 100mL water and stirred at room temperature for 40 minutes. The pH was brought to 6-7 by dropping concentrated H_2SO_4 and the solution was freeze-dried by lyophilization. 30mL water was added and the pH acidified to 2 by adding concentrated H_2SO_4. The product was extracted 4x 30mL EtOAc, washed with brine, dried over $MgSO_4$ and concentrated in vacuum. The crude product (~ 600mg) was dissolved in $CHCl_3$/5% acetic acid (10mL) passed through a column of silica gel (20g) and eluted with the same solvent. The fractions containing the desired product were collected, the eluate evaporated and residual acetic acid removed by co-evaporation with water.

Yield: 410mg (2.6mmol, 53%) colorless solid

TLC: R_f = 0.39 ($CHCl_3$/MeOH, 1:1)

$C_6H_5NO_4$ (155.11)

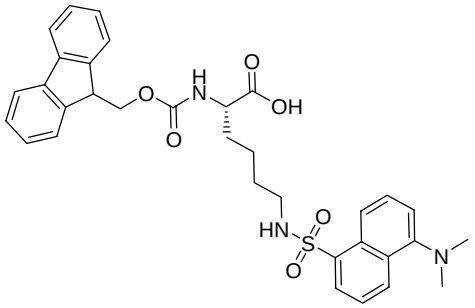
^{1}H-NMR (DMSO-d_6, 400 MHz): δ = 4.12 (2H, s, α-CH_2), 7.10 (2H, s, CH=CH) ppm.

^{13}C-NMR (DMSO-d_6, 100 MHz): δ = 39.4 (α-CH_2 *Gly*), 135.6 (CH=CH), 169.4 (CO_2H) and 170.9 (C=O *imide*) ppm.

GC-MS (Method B): t_r= 5.58 min; m/z: 155 $[M]^+$, 110/111 $[M\text{-}CO_2]^+$/+H

ESI-MS: 154.1 $[M\text{-}H]^-$ (calculated: 154.1)

Fmoc-Lys(Dansyl)-OH (47)



Fmoc-Lys(Boc)-OH (500mg, 1.1mmol) was treated with 10mL TFA:DCM (1:1) for 80 minutes at room temperature. Excess of TFA was removed by coevaporation with toluene and the peptide was dried under reduced pressure (colorless oil, Fmoc-Lys-OH.TFA). The side-chain unprotected peptide was dissolved in 35mL MeOH:H_2O (5:2), followed by addition of $NaHCO_3$ (270mg, 3.2mmol) and dansyl chloride (432mg, 1.6mmol). The reaction mixture was stirred for 19h at room temperature. The pH was adjusted to 2 by adding HCl 1M and the product was extracted 3x with DCM, washed with brine, dried over Na_2SO_4 and concentrated in vacuum. Purification was performed by flash silica gel chromatography, eluting first with DCM, followed by DCM:MeOH (10:1). Product showed characteristic dansyl-fluorescence.

Yield: 336mg (0.56mol, 52%) light yellow oil (fluorescent)

TLC: R_f = 0.10 (DCM:MeOH (10:1))

$C_{33}H_{35}N_3O_6S$ (601.71)

^{1}H-NMR (CD_3OD, 400 MHz): δ = 1.07-1.34 (4H, m, γ-CH_2 + δ-CH_2), 1.35-1.70 (2H, m, β-CH_2), 2.77-2.84 (2H, m, ε-CH_2), 2.79 (6H, s, *2x* CH_3 *Dan*), 3.95 (1H, dd, *J* = 4.5, 9.2 Hz, α-CH), 4.20 (1H, t, *J* = 6.7 Hz, CH *Fmoc*), 4.33 (2H, d, *J* = 6.0 Hz, CH_2 *Fmoc*), 7.06 (1H, d, *J* = 7.1 Hz, Ar *Dan*), 7.10-7.21 (2H, m, Ar *Dan*), 7.27 (2H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.40 (2H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.50 (2H, d, *J* = 7.0 Hz, Ar *Fmoc*), 7.64 (2H, d, *J* = 7.5 Hz, Ar *Fmoc*), 8.17 (1H, d, *J* = 7.2 Hz, Ar *Dan*), 8.34 (1H, d, *J* = 8.7 Hz, Ar *Dan*) and 8.51 (1H, d, *J* = 8.5 Hz, Ar *Dan*) ppm.

^{13}C-NMR (CD_3OD, 100 MHz): δ = 22.4, 29.1, 31.8, 42.8, 45.6, 47.3, 54.0, 67.2, 115.5, 119.3, 120.0, 123.5, 125.4, 127.1, 127.8, 128.5, 129.6, 129.8, 130.0, 130.5, 135.1, 141.4, 143.9, 144.1, 152.0, 156.8 and 174.8 ppm.

MALDI-TOF: m/z 600.4 $[M-H]^+$ (calc. 600.2), 624.4 $[M+Na]^+$ (calc. 624.2).

$[\alpha]_D^{20}$: - 5.5 (c= 0.4, $CHCl_3$)

6.2.2.2. *Solid-phase synthesis of the N-maleoyl-peptides using Wang resin*

Loading of Fmoc-glycine to Wang resin

4 equivalents of Fmoc-glycine, 0.1 equivalent of DMAP and 4 equivalents DIC were dissolved in DMF (10mL/g resin) and directly added to Wang resin (pre-swelled in DMF). The mixture was agitated overnight. The resin was washed with DMF (3x), DCM (3x) and MeOH (3x) and finally dried under high vacuum. Quantitative loading of Fmoc-glycine was obtained as indicated by UV measurements (0.9mmol Fmoc-glycine/g resin).

Peptide chain assembly

The general procedure was employed (Section 6.2.1).

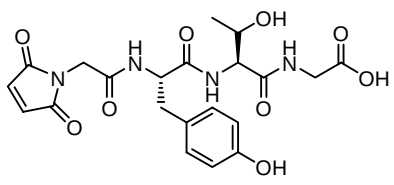
Coupling of the *N*-maleoyl-amino acid.

4 equivalents of *N*-maleoyl-glycine **4** (or *N*-maleoyl-β-alanine), of HOBt and of DIC were dissolved in DCM:DMF (1:1, 10mL/g resin) and added to resin. The mixture was agitated at room temperature and the coupling efficiency monitored by Kaiser test. The resin was washed with 5x DMF.

Peptide cleavage from Wang resin

The cleavage cocktail normally used was TFA:TIS:water (95:2.5:2.5). For peptide sequences longer than 10 amino acids, the amount of scavengers was increased. The peptidyl resin was treated with the cleavage cocktail (10mL/g resin) for 2-3h with gentle swirling. The resin was filtered off and washed with small portions of TFA. The filtrates were combined and the volume reduced to ca. 30% under reduced pressure. Ice-cooled diethylether was dropwise added to this solution (10-fold in volume) to promote peptide precipitation. The solid was separated by filtration, washed several times with diethylether, finally dissolved in a mixture of MeOH:water and freeze-dried. The desired peptides were obtained mostly in satisfactory purity. In a few cases, though, HPLC purification was required.

N-maleoyl-Gly-Tyr-Thr-Gly-OH (5a)



Starting from 301mg Fmoc-glycine loaded Wang resin (0.31mmol).

Yield: 91mg colorless solid (0.19mmol, 62%)

$C_{21}H_{24}N_4O_9$ (476.44)

^{1}H-NMR (D_2O, 400 MHz): δ = 0.99 (3H, d, *J* = 6.4 Hz, CH_3 *Thr*), 2.81-2.90 (2H, m, β-CH_2 *Tyr*), 3.77 (2H, s, α-CH_2 *Gly*), 4.02-4.07 (1H, m, α-CH *Thr*), 4.10 (2H, d, *J* = 2.8 Hz, α-CH_2 *Gly*), 4.17 (1H, d, *J* = 4.0 Hz, β-CH *Thr*), 4.50 (1H, t, *J* = 7.8 Hz, α-CH *Tyr*), 6.66 (2H, d, *J* = 8.4 Hz, Ar *Tyr*), 6.75 (2H, s, CH=CH) and 6.97 (2H, d, *J* = 8.4 Hz, Ar *Tyr*) ppm.

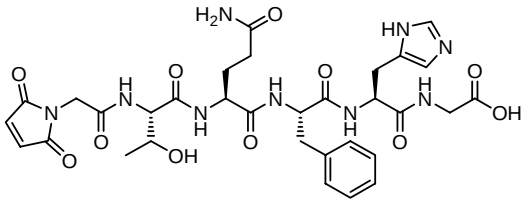
^{13}C-NMR (D_2O, 100 MHz): δ = 18.8, 36.4, 39.9, 41.4, 55.7, 58.9, 67.2, 115.7, 127.8, 130.7, 134.7, 154.5, 169.0, 171.7, 171.9, 172.7 and 173.2 ppm.

LC-MS (Method A): t_r = 13.9 min; m/z 477.0 $[M+H]^+$ (calc. 477.1), 499.0 $[M+Na]^+$ (calc. 499.1).

FAB-HRMS: m/z 499.1458 $[M+Na]^+$ (calc. 499.1441).

$[\alpha]_D^{20}$: +16.2 (c = 0.4, DMF)

N-maleoyl-Gly-Thr-Gln-Phe-His-Gly-OH (5b)



Starting from 276mg Fmoc-glycine loaded Wang resin (0.32mmol).

Yield: 140mg colorless solid (0.19mmol, 60%)

$C_{32}H_{39}N_9O_{11}$ (725.71)

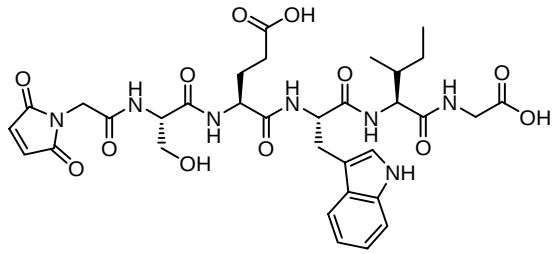
^{1}H-NMR (DMSO-d_6, 400 MHz): δ = 1.00 (3H, d, *J* = 6.2 Hz, CH_3 *Thr*), 1.61-1.68 (1H, m, β-CH_2 *Gln*), 1.80-1.88 (1H, m, β-CH_2 *Gln*), 1.99-2.11 (2H, m, γ-CH_2 *Gln*), 2.78 (1H, dd, *J* = 13.6, 9.4 Hz, β-CH_2), 2.83-3.01 (2H, m, *2x* 1H β-CH_2), 3.11 (1H, dd, *J* = 15.2, 5.4 Hz, β-CH_2), 3.68-3.84 (2H, m, α-CH_2 *Gly*), 3.91-3.98 (1H, m, β-CH *Thr*), 4.14-4.30 (4H, m, α-CH_2 *Gly* + *2x* α-CH), 4.43 (1H, dd, *J* = 8.8, 5.4 Hz, α-CH), 4.62 (1H, dd, *J* = 7.7, 6.0 Hz, α-CH), 6.82 (1H, br s, CH *His*), 7.08 (2H, s, CH=CH), 7.16-7.26 (5H, m, Ar *Phe*), 7.35 (1H, br s, CH *His*), 8.02 (1H, d, *J* = 7.6 Hz, CONH), 8.07 (1H, d, *J* = 7.5 Hz, CONH), 8.14 (1H, t, *J* = 5.8 Hz, CONH), 8.20 (1H, d, *J* = 8.1 Hz, CONH), 8.35 (1H, d, *J* = 8.5 Hz, CONH) and 8.95 (1H, s) ppm.

LC-MS (Method A): 12.2 min; m/z 726.3 $[M+H]^+$ (calc. 726.3).

MALDI-TOF: m/z 726.9 $[M+H]^+$ (calc. 726.3), 748.9 $[M+Na]^+$ (calc. 748.3) and 764.9 $[M+K]^+$ (calc. 764.3).

$[\alpha]_D^{20}$: - 20.2 (c = 1.0, DMF)

N-maleoyl-Gly-Ser-Glu-Trp-Ile-Gly-OH (5c)



Starting from 350mg Fmoc-glycine loaded Wang resin (0.40mmol).

Yield: 152mg colorless solid (0.21mmol, 53%)

$C_{33}H_{41}N_7O_{12}$ (727.72)

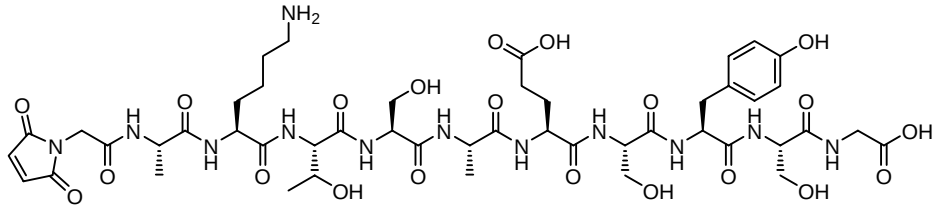
^{1}H-NMR ($CDCl_3$, 400 MHz): δ = 0.80 (3H, t, *J* = 7.4 Hz, CH_3 *Ile*), 0.84 (3H, d, *J* = 6.8 Hz, CH_3 *Ile*), 1.00-1.13 (1H, m, CH_2 *Ile*), 1.38-1.47 (1H, m, CH_2 *Ile*), 1.65-1.75 (2H, m, β-CH_2 *Glu*), 1.83-1.92 (1H, m, β-CH *Ile*), 2.19 (2H, t, *J* = 8.0 Hz, γ-CH_2 *Glu*), 2.89 (1H, dd, *J* = 14.8, 8.6 Hz, β-CH_2 *Trp*), 3.11 (1H, dd, *J* = 14.8, 5.4 Hz, β-CH_2 *Trp*), 3.49-3.59 (2H, m, β-CH_2 *Ser*), 3.65-3.80 (2H, m, α-CH_2 *Gly*), 4.12 (d, 2H, *J* = 5.5 Hz, α-CH_2 *Gly*), 4.17-4.28 (2H, m, 2x α-CH), 4.34 (1H, dd, *J* = 7.7, 5.9 Hz, α-CH), 4.59 (1H, dd, *J* = 8.1, 5.7 Hz, α-CH), 6.96 (1H, t, *J* = 8.0 Hz, Ar *Trp*), 7.04 (1H, t, *J* = 8.0 Hz, Ar *Trp*), 7.07 (s, 2H, CH=CH), 7.11 (1H, d, *J* = 2.3 Hz, Ar *Trp*), 7.30 (1H, d, *J* = 8.0 Hz, Ar *Trp*), 7.56 (1H, d, *J* = 7.8 Hz, Ar *Trp*), 7.86 (1H, d, *J* = 8.9 Hz, CONH), 8.02 (1H, d, *J* = 8.0 Hz, CONH), 8.10-8.12 (2H, m, 2x CONH), 8.33 (1H, d, *J* = 7.8 Hz, CONH) and 10.76 (1H, d, *J* = 2.0 Hz, NH *Trp*) ppm.

LC-MS (Method A): t_r = 16.9 min; m/z 728.1 $[M+H]^+$ (calc. 728.3).

MALDI-TOF: m/z 750.8 $[M+Na]^+$ (calc. 750.3) and 766.8 $[M+K]^+$ (calc. 766.3).

$[\alpha]_D^{20}$: - 24.5 (c = 0.7, DMF)

N-maleoyl-Gly-Ala-Lys-Thr-Ser-Ala-Glu-Ser-Tyr-Ser-Gly-OH (5d)



Starting from 285mg Fmoc-glycine loaded Wang resin (0.29mmol).

HPLC purification: gradient (time/solvent B): 0min (5%B) → 3min (5%B) → 15min (30%B), retention time: 10.1min.

Yield: 196mg colorless solid (0.17mmol, 59%)

$C_{47}H_{68}N_{12}O_{21}$ (1137.11)

^{1}H-NMR (D_2O, 400 MHz): δ = 1.05 (3H, d, *J* = 6.4 Hz, CH_3 *Thr*), 1.20-1.37 (2H, m, γ-CH_2 *Lys*), 1.26 (6H, d, *J* = 7.2 Hz, *2x* CH_3 *Ala*), 1.47-1.57 (2H, m, δ-CH_2 *Lys*), 1.58-1.78 (2H, m, β-CH_2 *Lys*), 1.79-1.97 (2H, m, β-CH_2 *Glu*), 2.31 (2H, t, *J* = 7.5 Hz, γ-CH_2 *Glu*), 2.82-2.92 (4H, m, ε-CH_2 *Lys* + β-CH_2 *Tyr*), 3.63-3.79 (6H, m, *3x* β-CH_2 *Ser*), 3.83 (2H, s, α-CH_2 *Gly*), 4.10-4.33 (11H, m, *8x* α-CH + α-CH_2 + β-CH *Thr*), 4.50 (1H, t, *J* = 7.5 Hz, α-CH *Tyr*), 6.67 (2H, d, *J* = 8.5 Hz, Ar *Tyr*), 6.80 (2H, s, CH=CH) and 7.00 (2H, d, *J* = 8.5 Hz, Ar *Tyr*) ppm.

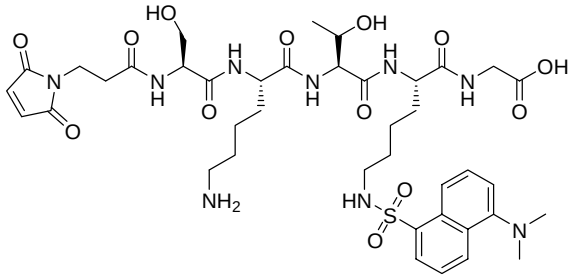
LC-MS (Method A): t_r = 11.4 min; m/z 1137.5 $[M+H]^+$ (calc. 1137.5).

MALDI-TOF: 1137.7 $[M+H]^+$ (calc. 1137.5), 1159.7 $[M+Na]^+$ (calc. 1159.5), 1175.6 $[M+K]^+$ (calc. 1175.5).

FAB-LRMS: 1136.45 $[M]^+$ (calc. 1136.46).

$[\alpha]_D^{20}$: - 7.7 (c = 0.3, DMF)

N-maleoyl-βAla-Ser-Lys-Thr- Lys(Dansyl)-Gly-OH (46)



Starting from 147mg Fmoc-glycine loaded Wang resin (0.13mmol)

Yield: 67mg of colorless solid (0.074mmol, 57%)

$C_{40}H_{57}N_9O_{13}S$ (904.0)

^{1}H-NMR (CD_3OD, 400 MHz): δ = 1.16 (3H, d, *J* = 6.4 Hz, CH_3 *Thr*), 1.22-1.43 (4H, m, *2x* γ-CH_2 *Lys*), 1.47-1.73 (6H, m, *2x* δ-CH_2 *Lys* + β-CH_2 *Lys*), 1.73-1.81 (1H, m, β-CH_2 *Lys*), 1.94-2.03 (1H, m, β-CH_2 *Lys*), 2.54 (2H, t, *J* = 6.7 Hz, CH_2 $^{\beta}$*Ala*), 2.84 (2H, t, *J* = 6.3 Hz, ε-CH_2 *Lys*), 2.94-3.00 (8H, m, ε-CH_2 *Lys* + *2x* CH_3 *Dan*), 3.68-3.83 (4H, m, CH_2 $^{\beta}$*Ala* + β-CH_2 *Ser*), 3.84-3.93 (2H, m, α-CH_2 *Gly*), 4.13-4.19 (1H, m, β-CH *Thr*), 4.25 (1H, dd, *J* = 9.1, 4.8 Hz, α-CH), 4.29-4.34 (2H, m, *2x* α-CH), 4.43 (1H, dd, *J* = 9.7, 4.7 Hz, α-CH), 6.79 (2H, s,

CH=CH), 7.39 (1H, d, *J* = 7.6 Hz, Ar *Dan*), 7.59-7.64 (2H, m, Ar *Dan*), 8.20 (1H, d, *J* = 7.3 Hz, Ar *Dan*), 8.42 (1H, d, *J* = 8.7 Hz, Ar *Dan*) and 8.53 (1H, d, *J* = 8.6 Hz, Ar *Dan*) ppm.

MALDI-TOF: m/z 904.9 $[M+H]^+$ (calc. 904.4), 926.9 $[M+Na]^+$ (calc. 926.4), 942.9 $[M+K]^+$ (calc. 943.3).

ESI-MS: m/z 904.7 $[M+H]^+$ (calc. 904.4)

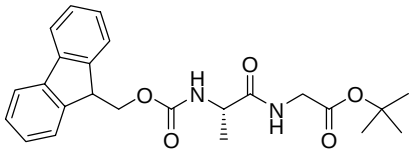
FAB-HRMS: m/z 904.3904 $[M+H]^+$ (calc. 904.3875).

$[\alpha]_D^{20}$**:** - 19.0 (c = 0.4, MeOH)

6.2.3. Synthesis of the hexadienyl ester peptides in solution phase

6.2.3.1. *First hexadienyl esters peptide templates*

Fmoc-Ala-Gly-O*t*Bu (6)



Fmoc-Ala-OH (3.79g, 12.2mmol) was dissolved in 25mL THF:DCM (3:2) and combined with a solution of H-Gly-O*t*Bu (2.04g, 12.2mmol) in 10mL DCM. After addition of HOBt (2.09g, 14.6mmol), the flask was cooled to 0°C and DIC (2.1mL, 13.4mmol) was added dropwise. After 1 hour, the reaction was allowed to warm to room temperature by stirring overnight. The resulting mixture was filtered to remove precipitated urea. The filtrate was washed 3x HCl 1M, 3x 10% $NaHCO_3$, 1x brine, dried over $MgSO_4$ and concentrated in vacuum. The product was recrystallized from EtOAc.

Yield: 4.2g colorless crystals (9.9mmol, 81%)

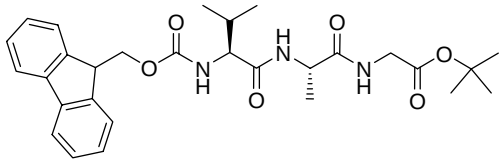
TLC: R_f = 0.46 (cHex/EtOAc, 1:2)

$C_{24}H_{28}N_2O_5$ (424.50)

^{1}H-NMR ($CDCl_3$, 400 MHz): δ = 1.41 (3H, d, *J* = 7.0 Hz, CH_3 *Ala*), 1.45 (9H, s, *tBu*), 3.95 (2H, d, α-CH_2 *Gly*), 4.18 (1H, t, *J* = 6.8 Hz, CH *Fmoc*), 4.30 (1H, m, α-CH *Ala*), 4.42 (2H, d, *J* = 5.5 Hz, CH_2 *Fmoc*), 5.36 (1H, br s, NH), 6.45 (1H, br s, NH), 7.32 (2H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.42 (2H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.60 (2H, d, *J* = 7.4 Hz, Ar *Fmoc*) and 7.78 (2H, d, *J* = 7.5 Hz, Ar *Fmoc*) ppm.

^{13}C-NMR ($CDCl_3$, 100 MHz): δ = 19.1 (CH_3 *Ala*), 28.4 (*tBu*), 42.5 (CH *Fmoc*), 47.5 (α-CH_2 *Gly*), 50.9 (α-CH *Ala*), 67.5 (*tBu*), 82.9 (CH_2 *Fmoc*), 120.4 (Ar *Fmoc*), 125.4 (Ar *Fmoc*), 127.5 (Ar *Fmoc*), 128.1 (Ar *Fmoc*), 141.7 (Ar *Fmoc*), 144.1 (Ar *Fmoc*), 156.3 (C=O), 169.1 (C=O) and 172.5 (C=O) ppm.

Fmoc-Val-Ala-Gly-O*t*Bu (7)



1) Fmoc deprotection: Fmoc-Ala-Gly-O*t*Bu **6** (556mg, 1.31mmol) was treated with 5mL 20% piperidine in DCM during 30 minutes. Piperidine was then removed by coevaporation with methanol (3x) and the product H-Ala-Gly-OtBu was purified by flash chromatography (DCM, DCM:MeOH (10:1)), affording 250mg colorless oil (1.23mmol, 94% yield). *TLC*: R_f = 0.17 (DCM/MeOH, 10:1). *^{1}H-NMR* ($CDCl_3$, 400 MHz): δ = 1.42 (3H, d, *J* = 7.0 Hz, CH_3 *Ala*), 1.46 (9H, s, $C(CH_3)_3$), 3.52 (1H, m, α-CH *Ala*), 3.85 (2H, m, α-CH_2 *Gly*) and 7.68 (1H, br s, NH) ppm.

2) Coupling: Fmoc-Val-OH (2.07g, 6.1mmol), H-Ala-Gly-O*t*Bu (1.24g, 6.1mmol) and triethylamine (846µl, 6.1mmol) were dissolved in DCM (30mL) and cooled to 0°C. HOBt (1.87g, 12.2mmol) and EDC (1.64g, 8.5mmol) were added and the reaction mixture was allowed to warm up to room temperature by stirring overnight. After 19 hours, the precipitate was separated by filtration, washed with DCM and the product crystallized from methanol (2.0g product was isolated). As some amount of product remained dissolved in the filtrate, this was washed 2x HCl 1M, 2x $NaHCO_3$ 1M, 1x water, dried over Na_2SO_4 and concentrated in vacuum. Product was again recrystallised from methanol (0.8g isolated, 88% yield).

Yield: 2.8g colorless solid (5.3mmol, 83% over two steps)

TLC: R_f = 0.51 (EtOAc)

$C_{29}H_{37}N_3O_6$ (523.62)

^{1}H-NMR (DMSO-d_6, 400 MHz): δ = 0.84 (3H, d, *J* = 6.8 Hz, CH_3 *Val*), 0.86 (3H, d, *J* = 6.9 Hz, CH_3 *Val*), 1.23 (3H, d, *J* = 7.0 Hz, CH_3 *Ala*), 1.38 (9H, s, *tBu*), 1.99 (1H, m, β-CH *Val*), 3.70 (2H, ddd, *J* = 44.4, 27.0, 5.5 Hz, α-CH_2 *Gly*), 3.90 (1H, t, *J* = 6.8 Hz, CH *Fmoc*), 4.22 (2H, d, *J* = 5.5 Hz, CH_2 *Fmoc*), 4.23 (2H, m, α-CH *Val* + α-CH *Ala*), 7.32 (2H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.41 (2H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.45 (1H, CONH), 7.74 (2H, d, *J* = 7.4 Hz, Ar *Fmoc*),

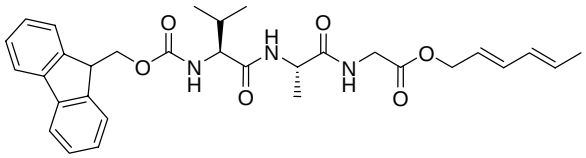
7.88 (2H, d, *J* = 7.5 Hz, Ar *Fmoc*), 8.00 (1H, d, *J* = 7.5 Hz, CONH) and 8.24 (1H, t, *J* = 5.7 Hz, CONH) ppm.

^{13}C-NMR (DMSO-d_6, 100 MHz): δ = 18.5 (CH_3 *Val*), 18.9 (CH_3 *Val*), 19.6 (CH_3 *Ala*), 28.0 ($C(CH_3)_3$), 30.7 (β-CH *Val*), 37.0 (CH *Fmoc*), 47.0 (α-CH_2 *Gly*), 48.2 (α-CH), 60.4 (α-CH), 66.0 ($C(CH_3)_3$), 80.9 (CH_2 *Fmoc*), 120.5 (Ar *Fmoc*), 125.7 (Ar *Fmoc*), 127.4 (Ar *Fmoc*), 128.0 (Ar *Fmoc*), 141.1 (Ar *Fmoc*), 144.1 (Ar *Fmoc*), 156.5 (C=O), 169.1 (C=O), 171.1 (C=O) and 172.8 (C=O) ppm.

ESI-MS: m/z 546.4 $[M+Na]^+$ (calc. 546.3)

$[\alpha]_D^{20}$: - 24.8 (c = 1.3, MeOH)

Fmoc-Val-Ala-Gly-hexadienylether (8)



1) *t*Bu deprotection: Fmoc-Val-Ala-Gly-O*t*Bu **7** (1.03g, 1.97mmol) was treated with 20mL TFA:DCM (1:1) for 1 hour at room temperature. Excess of TFA was removed by coevaporation with toluene and the peptide dried under reduced pressure (0.89g, colorless crystals, Fmoc-Val-Ala-Gly-OH).

2) Esterification: Fmoc-Val-Ala-Gly-OH (445mg, 0.95mmol) was dissolved in 24mL DCM:DMF (1:1) and combined with a solution containing *trans*,*trans*-2,4-hexadien-1-ol (95mg, 0.95mmol) and triethylamine (264µL, 1.90mmol) in 2mL DCM. After cooling the mixture to 0°C, HOBt (292mg, 1.90mmol) and EDC (255mg, 1.33mmol) were added. Next, the cooling bath was removed and reaction was stirred overnight at room temperature. After 18h, DCM was added and reaction mixture washed 3x 5% $NaHCO_3$, 1x water, dried over Na_2SO_4 and concentrated in high vacuum.

Yield: 385mg colorless solid (0.70mmol, 71%)

TLC: R_f = 0.60 (EtOAc)

$C_{31}H_{37}N_3O_6$ (547.64)

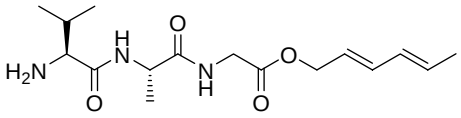
^{1}H-NMR (DMSO-d_6, 400 MHz): δ = 0.84 (3H, d, *J* = 6.8 Hz, CH_3 *Val*), 0.86 (3H, d, *J* = 6.8 Hz, CH_3 *Val*), 1.22 (3H, d, *J* = 7.0 Hz, CH_3 *Ala*), 1.72 (3H, d, *J* = 6.6 Hz, CH_3 *Hxd*), 1.92-2.02 (1H, m, β-CH *Val*), 3.78-3.93 (2H, ddd, *J* = 21.9, 17.4, 5.8 Hz, α-CH_2 *Gly* + 1H, CH *Fmoc*), 4.19-4.37 (4H, m, CH_2 *Fmoc* + *2x* α-CH), 4.55 (2H, d, *J* = 6.4 Hz, CH_2 *Hxd*), 5.56-5.63 (1H, m, CH=C*H*CH_3), 5.70-5.79 (1H, m, CH_2C*H*=CH), 6.02-6.09 (1H, dd, *J* = 15.1, 10.5 Hz,

C*H*=CHCH$_3$), 6.22-6.29 (1H, dd, *J* = 15.2, 10.4 Hz, CH$_2$CH=C*H*), 7.32 (2H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.39-43 (2H, t, *J* = 7.4 Hz, Ar Fmoc + 1H, CONH), 7.74 (2H, t, *J* = 6.8 Hz, Ar *Fmoc*), 7.88 (2H, d, *J* = 7.5 Hz, Ar *Fmoc*), 8.00 (1H, d, *J* = 7.5 Hz, CONH) and 8.32 (1H, t, *J* = 5.7 Hz, CONH) ppm.

LC-MS (Method B): t_r = 9.9 min; m/z 548.0 $[M+H]^+$ (calc. 548.3), 570.2 $[M+Na]^+$ (calc. 570.3).

$[\alpha]_D^{20}$: - 22.8 (c = 1.0, MeOH)

H-Val-Ala-Gly-OHxd (9)



152mg (0.29mmol) of the Fmoc-protected peptide hexadienyl ester **8** was treated with 20% piperidine in DMF (2.0mL) for 30 minutes at room temperature. After removal of piperidine was accomplished by coevaporation with methanol, the crude product was dried in vacuum and purified by HPLC and dried by lyophilization.

HPLC purification: gradient of 0min (10%B) → 16min (50%B), retention time: 10.7min.

Yield: 70mg colorless solid (0.21mmol, 73%)

$C_{16}H_{27}N_3O_4$ (325.40)

^{1}H-NMR (CD_3OD, 400 MHz): δ = 0.89 (3H, d, *J* = 6.8 Hz, CH$_3$ *Val*), 0.95 (3H, d, *J* = 6.8 Hz, CH$_3$ *Val*), 1.36 (3H, d, *J* = 7.1 Hz, CH$_3$ *Ala*), 1.72 (2H, d, *J* = 6.5 Hz, CH$_3$ *Hxd*), 1.93-2.03 (1H, m, β-CH *Val*), 3.16 (1H, d, *J* = 5.3 Hz, α-CH *Val*), 3.84-4.00 (2H, m, α-CH$_2$ *Gly*), 4.42 (1H, q, *J* = 7.1 Hz, α-CH *Ala*), 4.59 (2H, d, *J* = 6.5 Hz, CH$_2$ *Hxd*), 5.59 (1H, ddd, *J* = 14.1, 6.5, 6.5 Hz, CH=C*H*CH$_3$), 5.70-5.79 (1H, m, CH$_2$C*H*=CH), 6.04 (1H, dd, *J* = 15.0, 10.5 Hz, C*H*=CHCH$_3$) and 6.25 (1H, dd, *J* = 15.0, 10.6 Hz, CH$_2$CH=C*H*) ppm.

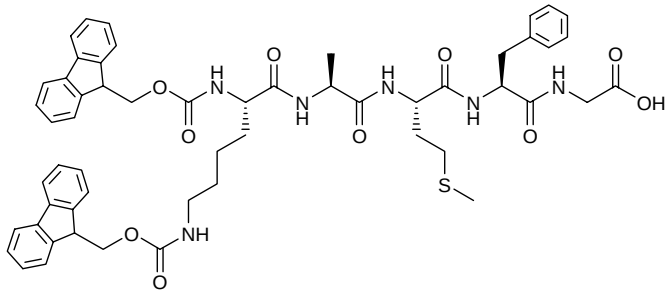
^{13}C-NMR (CD_3OD, 100 MHz): δ = 16.2 (CH$_3$ *Val*), 16.7 (CH$_3$ *Val*), 16.9 (CH$_3$ *Ala*), 18.3 (CH$_3$ *Hxd*), 31.8 (β-CH *Val*), 37.0 (α-CH$_2$ *Gly*), 40.5 (α-CH), 59.9 (α-CH), 65.1 (CH$_2$ *Hxd*), 123.0 (CH=*C*HCH$_3$), 130.3 (CH$_2$*C*H=CH), 130.7 (CH$_2$CH=*C*H), 134.8 (*C*H=CHCH$_3$), 169.4 (C=O), 173.8 (C=O) and 174.6 (C=O) ppm.

LC-MS (Method B): t_r = 1.5 min; m/z 326.1 $[M+H]^+$ (calc. 326.2).

MALDI-TOF: m/z 326.6 $[M+H]^+$ (calc. 326.2), 364.5 $[M+K]^+$ (calc. 364.2).

$[\alpha]_D^{20}$: - 18.5 (c = 0.5, MeOH)

Fmoc-Lys(Fmoc)-Ala-Met-Phe-Gly-OH (10)



Starting from 272mg Fmoc-glycine loaded Wang resin (0.31mmol). Due to the presence of a methionine residue, a different cleavage cocktail was used for this peptide: TFA: H_2O:EDT:TIS (94.5:2.5:2.5:1.0), which was isolated following the general procedure described above (Section 6.2.2).

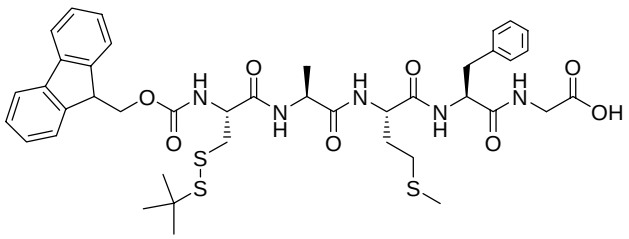
Yield: 165mg colorless solid (0.17mmol, 55%)

$C_{55}H_{60}N_6O_{10}S$ (997.16)

^{1}H-NMR (DMSO-d_6, 400 MHz): δ = 1.16 (3H, d, *J* = 7.0 Hz, CH_3 *Ala*), 1.17-1.45 (4H, m, γ-CH_2 *Lys* + δ-CH_2 *Lys*), 1.46-1.85 (4H, m, β-CH_2 *Lys* + β-CH_2 *Met*), 1.97 (3H, s, CH_3 *Met*), 2.26-2.37 (2H, m, γ-CH_2 *Met*), 2.79 (1H, dd, *J* = 13.9, 9.5 Hz, β-CH_2 *Phe*), 2.90-3.00 (2H, m, ε-CH_2 *Lys*), 3.03 (1H, dd, *J* = 13.9, 4.5 Hz, β-CH_2 *Phe*), 3.71-3.83 (2H, m, α-CH_2 *Gly*), 3.93-3.98 (1H, m, α-CH), 4.17-4.36 (8H, m, *2x* CH *Fmoc* + *2x* α-CH + *2x* CH_2 *Fmoc*), 4.53-4.58 (1H, m, α-CH *Phe*), 7.12-7.23 (5H, m, Ar *Phe*), 7.25-7.33 (5H, m, Ar *Fmoc* + CONH), 7.41 (4H, t, *J* = 7.5 Hz, Ar *Fmoc*), 7.49 (1H, d, *J* = 8.1 Hz, CONH), 7.65-7.73 (4H, m, Ar *Fmoc*), 7.87-7.91 (6H, m, Ar *Fmoc* + 2x CONH), 8.03 (1H, d, *J* = 7.1 Hz, CONH) and 8.33 (1H, t, *J* =5.7 Hz, CONH) ppm.

LC-MS (Method B): t_r = 10.2 min; m/z 997.2 $[M+H]^+$ (calc. 997.4).

Fmoc-Cys(*St*Bu)-Ala-Met-Phe-Gly-OH (11)



Starting from 272mg Fmoc-glycine loaded Wang resin (0.32 mmol). Due to the presence of a methionine residue, a different cleavage cocktail was used for this peptide: TFA:

H_2O:phenol:TIS (88:5:5:2), which was isolated following the general procedure described above (Section 6.2.2).

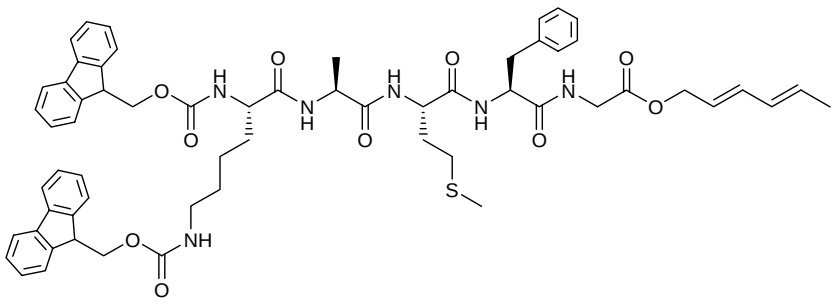
Yield: 106mg of a colorless solid (0.13mmol, 41%)

$C_{41}H_{51}N_5O_8S_3$ (838.07)

^{1}H-NMR (DMSO-d_6, 400 MHz): δ = 1.17 (3H, d, *J* = 7.0 Hz, CH_3 *Ala*), 1.29 (9H, s, *tBu*), 1.68-1.83 (2H, dm, β-CH_2 *Met*), 1.98 (3H, s, CH_3 *Met*), 2.28-2.36 (2H, m, γ-CH_2 *Met*), 2.79 (1H, dd, *J* = 14.0, 9.4 Hz, β-CH_2 *Phe*), 2.93 (1H, dd, *J* = 12.9, 9.9 Hz, β-CH_2 *Phe*), 3.02-3.08 (2H, m, β-CH_2 *Cys*), 3.72-3.83 (2H, m, α-CH_2 *Gly*), 4.20-4.33 (6H, m, CH *Fmoc* + *3x* α-CH + CH_2 *Fmoc*), 4.52-4.58 (1H, m, α-CH), 7.14-7.23 (5H, m, Ar *Phe*), 7.32 (2H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.41 (2H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.72 (3H, m, Ar *Fmoc* + 1H, NH), 7.89 (3H, m, Ar *Fmoc* + 1H, NH), 8.16 (1H, d, *J* = 7.3 Hz, NH) and 8.34 (1H, t, *J* = 5.6 Hz, NH) ppm.

LC-MS (Method B): t_r = 10.0 min; m/z 838.2 $[M+H]^+$ (calc. 838.3).

Fmoc-Lys(Fmoc)-Ala-Met-Phe-Gly-OHxd (12)



Starting from 78mg Fmoc-Lys(Fmoc)-Ala-Met-Phe-Gly-OH **10** (0.078mmol). Esterification procedure similar to the one performed for compound **9**. After overnight reaction, the mixture was directly submitted to reversed-phase HPLC purification (retention time: 18.4min). Fractions containing the desired product were combined and freeze-dried.

Yield: 5mg colorless solid (0.004mmol, 5%)

$C_{61}H_{68}N_6O_{10}S$ (1077.29)

^{1}H-NMR (DMSO-d_6, 400 MHz): δ = 1.16 (3H, d, *J* = 7.0 Hz, CH_3 *Ala*), 1.20-1.44 (4H, m, γ-CH_2 *Lys* + δ-CH_2 *Lys*), 1.45-1.85 (4H, m, β-CH_2 *Lys* + β-CH_2 *Met*), 1.71 (3H, d, *J* = 6.3 Hz, CH_3 *Hxd*), 1.97 (3H, s, CH_3 *Met*), 2.26-2.38 (2H, m, γ-CH_2 *Met*), 2.79 (1H, dd, *J* = 14.0, 9.5 Hz, β-CH_2 *Phe*), 2.91-3.00 (2H, m, ε-CH_2 *Lys*), 3.03 (1H, dd, *J* = 13.9, 4.5 Hz, β-CH_2 *Phe*), 3.80-3.90 (2H, m, α-CH_2 *Gly*), 3.91-3.98 (1H, m, α-CH), 4.17-4.29 (8H, m, *2x* CH *Fmoc* + *2x* α-CH + *2x* CH_2 *Fmoc*), 4.52-4.55 (1H, m, α-CH *Phe*), 4.57 (2H, d, *J* = 6.3 Hz, CH_2 *Hxd*), 5.59

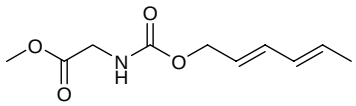
(1H, ddd, , J = 13.0, 6.3, 6.3 Hz, CH=C*H*CH_3), 5.70-5.79 (1H, m, CH_2C*H*=CH), 6.06 (1H, dd, J = 15.0, 10.8 Hz, C*H*=$CHCH_3$), 6.27 (1H, dd, J = 15.2, 10.5 Hz, CH_2CH=C*H*), 7.13-7.24 (5H, m, Ar *Phe*), 7.25-7.33 (5H, m, Ar *Fmoc* + CONH), 7.41 (4H, t, J = 7.4 Hz, Ar *Fmoc*), 7.49 (1H, d, J = 8.0 Hz, CONH), 7.66-7.72 (4H, m, Ar *Fmoc*), 7.87-7.94 (6H, m, Ar *Fmoc* + *2x* CONH), 8.03 (1H, d, J = 7.0 Hz, CONH) and 8.47 (1H, t, J = 5.6 Hz, CONH) ppm.

ESI-MS: m/z 1077.2 $[M+H]^+$ (calc. 1077.5)

MALDI-TOF: m/z 1099.8 $[M+Na]^+$ (calc. 1099.5), 1115.8 $[M+K]^+$ (calc. 1115.4).

6.2.3.2. *Peptides used for investigation of the stereochemistry*

2,4-hexadienyl-(methoxycarbonyl)-methylcarbamate (23)



1,1´-carbonyl-diimidazole (486mg, 3mmol) was added to a solution of *trans*,*trans*-2,4-hexadien-1-ol (294mg, 3mmol) in dry DMF (5mL) and the reaction solution was stirred at room temperature for 3 hours. A mixture of glycine-methyl ester hydrochloride (502mg, 4mmol) and triethylamine (1.3mL, 10mmol) in 10mL dry DMF was slowly added to the reaction solution and the resulting mixture was agitated overnight. DMF was removed in high vacuum at 55°C, followed by dilution of crude product in EtOAc:MeOH (20:1). The organic solution was washed with 3x $NaHCO_3$ 5%, 1x water, 1x brine, dried over Na_2SO_4 and concentrated under reduced pressure. The product was purified by flash silica gel chromatography, eluting with cHex:EtOAc (2:1).

Yield: 217mg colorless oil (1 mmol, 34%)

TLC: R_f = 0.37 (EtOAc)

$C_{10}H_{15}NO_4$ (213.23)

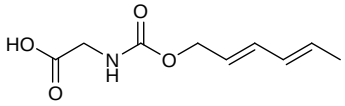
^{1}H-NMR ($CDCl_3$, 400MHz): δ = 1.74 (3H, d, J = 6.6 Hz, CH_3), 3.74 (3H, s, OCH_3), 3.95 (2H, d, J = 5.5 Hz, α-CH_2 *Gly*), 4.57 (2H, d, J = 6.5 Hz, CH_2 *Hxd*), 5.26 (1H, br s, NH), 5.61 (1H, ddd, , J = 13.7, 6.6, 6.6 Hz, CH=C*H*CH_3), 5.68-5.77 (1H, m, CH_2C*H*=CH), 6.03 (1H, dd, J = 15.0, 10.6 Hz, C*H*=$CHCH_3$) and 6.22 (1H, dd, J = 15.1, 10.5 Hz, CH_2CH=C*H*) ppm.

^{13}C-NMR ($CDCl_3$, 100MHz): δ = 18.3 (CH_3), 42.8 (α-CH_2 *Gly*), 52.5 (OCH_3), 66.0 (CH_2 *Hxd*), 124.3 (CH_2-*C*H=CH), 130.7 (CH_2-CH=*C*H), 131.4 (*C*H=CH-CH_3), 134.8 (CH_2-*C*H=CH), 156.5 (CONH) and 170.7 (*C*OOMe) ppm.

GC-MS (Method B)**:** t_r = 6.39 min; m/z: 213 $[M]^+$.

FAB-HRMS: m/z 213.0970 $[M]^+$ (calc. 213.1001).

N-glycyl-2,4-hexadienyl carbamate (24)



Compound **23** (189mg, 0.88mmol) was dissolved in 5mL MeOH and treated with NaOH 4M (443µl, 1.77mmol) for 30 minutes at room temperature. Methanol was removed from the solution under reduced pressure and water was added to the resulting mixture. The aqueous phase was washed with EtOAc and later acidified to pH 4 by adding HCl 1M. The product was extracted with EtOAc (3x), the organic phase then washed with brine, dried over Na_2SO_4 and concentrated/dried under vacuum. The resulting product showed satisfactory purity.

Yield: 162 mg colorless solid (0.81 mmol, 92%)

TLC: R_f = 0.16 (EtOAc)

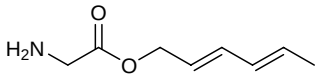
$C_9H_{13}NO_4$ (199.20)

^{1}H-NMR ($CDCl_3$, 400MHz): δ = 1.75 (3H, d, *J* = 6.7 Hz, CH_3), 4.01 (2H, d, *J* = 5.6 Hz, α-CH_2 *Gly*), 4.59 (2H, d, *J* = 6.5 Hz, CH_2 *Hxd*), 5.31 (1H, br s, NH), 5.61 (1H, ddd, , *J* = 13.4, 6.6, 6.6 Hz, CH=C*H*CH_3), 5.70-5.79 (1H, m, CH_2C*H*=CH), 6.04 (1H, dd, *J* = 15.0, 10.5 Hz, C*H*=CHCH_3), 6.24 (1H, dd, *J* = 15.1, 10.5 Hz, CH_2CH=C*H*) and 9.42 (1H, br s, COOH) ppm.

^{13}C-NMR ($CDCl_3$, 100MHz): δ = 18.3 (CH_3), 42.7 (α-CH_2 *Gly*), 66.8 (CH_2 *Hxd*), 124.0 (CH_2-*C*H=CH), 130.6 (CH_2-CH=*C*H), 131.5 (*C*H=CH-CH_3), 135.0 (CH_2-*C*H=CH), 157.6 (CONH) and 174.6 (COOH) ppm.

FAB-HRMS: m/z 200.0923 $[M+H]^+$ (calc. 200.0910).

Glycine-2,4-hexadienyl-1-ester (26)



1) Esterification: Fmoc-Gly-OH (6.89g, 23mmol) and *trans*,*trans*-2,4-hexadien-1-ol (2.27g, 23mmol) were dissolved in THF (100mL) and the solution cooled to 0°C. DMAP (280mg, 2.3mmol) and DIC (3.9mL, 25mmol) were added and the solution was allowed to warm up to room temperature and stirred overnight. After 13 hours, the reaction mixture was concentrated in vacuum, dissolved in DCM and $NaHCO_3$ saturated solution was added. The

urea precipitate was filtered off and the filtrate transferred to a separation funnel. The organic phase was washed again with $NaHCO_3$ saturated solution and brine, dried over $MgSO_4$ and concentrate in vacuum. The product was roughly purified by flash chromatography (flash master, cyclohexane:EtOAc) to give a colorless oil (7.1g, Fmoc-glycine-hexadienylester (60%) + hexadienol, determined by ^{1}H-NMR (data not shown)).

2) Fmoc deprotection: The Fmoc-glycine-hexadienylester (90mg, 60%, 0.14mmol) was treated with 20% piperidine in DCM (2.5mL) at room temperature for 30 minutes. Piperidine was removed by coevaporation with methanol and the product purified by flash chromatography (EtOAc, EtOAc:MeOH (95:5))

Yield: 15mg light yellow oil (0.10mmol, 34% over two steps)

TLC: R_f = 0.42 ($CHCl_3$/MeOH, 4:1)

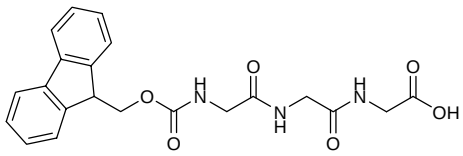
$C_8H_{13}NO_2$ (155.19)

^{1}H-NMR ($CDCl_3$, 400MHz): δ = 1.66 (2H, br s, NH_2), 1.74 (3H, d, *J* = 6.6 Hz, CH_3), 3.41 (2H, s, CH_2 *Gly*), 4.60 (2H, d, *J* = 6.7Hz, CH_2), 5.60 (1H, m, CH=C*H*-CH_3), 5.75 (1H, m, CH_2-C*H*=CH), 6.03 (1H, dd, *J* = 15.2, 10.4 Hz, C*H*=CH-CH_3) and 6.24 (1H, dd, *J* = 15.2, 10.4 Hz, CH_2-CH=C*H*) ppm.

^{13}C-NMR (DMSO-d_6, 100MHz): δ = 18.6 (*C*H_3), 44.4 (*C*H_2 *Gly*), 65.7 (*C*H_2), 123.4 (CH=*C*H-CH_3), 130.4 (CH_2-CH=*C*H), 131.8 (*C*H=CH-CH_3), 135.5 (CH_2-*C*H=CH) and 174.0 (*C*=O) ppm.

ESI-MS: m/z 155.9 $[M+H]^+$ (calc. 156.1)

Fmoc-Gly-Gly-Gly-OH (28)



1) Loading on 2-Cl trityl resin: 2 equivalents of the Fmoc-Gly-OH and 4 equivalents of DIPEA were combined in DCM (10mL per gram resin) and added to the resin (378mg, 0.60mmol, pre-swelled in DCM). After stirring the mixture for 4 hours, the resin was washed 3x with DCM/MeOH/DIPEA (17:2:1), then 3x DCM, 2x DMF, 2x DCM and finally dried under reduced pressure. The loading was quantitative (determined by UV measurement).

2) Peptide chain assembly: general procedure (section 6.2.1)

3) Cleavage: The peptidyl resin was treated four times with a solution of 1% TFA in DCM (approximately 10mL/g resin) for 5-10 minutes. The filtrates were combined and a solution of

10% pyridine in methanol was added (1/2 volume from peptide solution). The solvent was partially removed under reduced pressure and the peptide was precipitated by adding water. The solid was removed by filtration, dissolved in a mixture of methanol/water and finally dried by lyophilization.

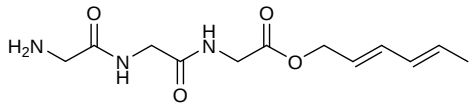
Yield: 177mg colorless solid (0.43mmol, 72%)

$C_{21}H_{21}N_3O_6$ (411.41)

^{1}H-NMR (CD_3OD, 400 MHz): δ = 3.81 (2H, br s, α-CH_2 *Gly*), 3.93 (4H, br s, *2x* α-CH_2 *Gly*), 4.22 (1H, t, *J* = 6.8 Hz, CH *Fmoc*), 4.38 (2H, d, *J* = 6.8 Hz, CH_2 *Fmoc*), 7.30 (2H, t, *J* = 7.5 Hz, Ar *Fmoc*), 7.39 (2H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.66 (2H, d, *J* = 7.4 Hz, Ar *Fmoc*) and 7.78 (2H, d, *J* = 7.5 Hz, Ar *Fmoc*) ppm.

MALDI-TOF: m/z 412.4 $[M+H]^+$ (calc. 412.2), 434.4 $[M+Na]^+$ (calc. 434.2), 450.4 $[M+K]^+$ (calc. 450.1).

H-Gly-Gly-Gly-OHxd (27)



1) Esterification: Starting from 72mg Fmoc-Gly-Gly-Gly-OH **28** (0.17mmol). Similar procedure was carried out as for compound **9**. Product was purified by silica gel flash chromatography eluting with ethyl acetate, to give 25mg colorless oil (0.051mmol, 30%). *TLC*: R_f = 0.18 (EtOAc). *^{1}H-NMR* ($CDCl_3$, 400 MHz): δ = 1.74 (3H, d, *J* = 6.7 Hz, CH_3 *Hxd*), 3.88 (2H, br s, α-CH_2 *Gly*), 3.98 (4H, br s, α-CH_2 *Gly*), 4.19 (1H, t, *J* = 6.8 Hz, CH *Fmoc*), 4.40 (2H, d, *J* = 6.6 Hz, CH_2 *Fmoc*), 4.57 (2H, d, *J* = 6.6 Hz, (CH_2 *Hxd*), 5.53 (1H, ddd, , *J* = 14.0, 6.7, 6.7 Hz, CH=C*H*CH_3), 5.69-5.77 (1H, m, CH_2C*H*=CH), 5.87 (1H, br s, CONH), 6.00 (1H, dd, *J* = 15.0, 10.6 Hz, C*H*=$CHCH_3$), 6.20 (1H, dd, *J* = 15.3, 10.7 Hz, CH_2CH=C*H*), 7.06 (1H, br s, CONH), 7.15 (1H, br s, CONH), 7.28 (2H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.38 (2H, t, *J* = 7.3 Hz, Ar *Fmoc*), 7.57 (2H, d, *J* = 7.2 Hz, Ar *Fmoc*) and 7.74 (2H, d, *J* = 7.5 Hz, Ar *Fmoc*) ppm.

2) Fmoc deprotection: Starting from 25mg Fmoc-Gly-Gly-Gly-O-hexadienyl ester (0.051mmol). Procedure was performed similarly to the synthesis of compound **6**.

Yield: 10mg colorless oil (0.037mmol, 73% for last step)

$C_{12}H_{19}N_3O_4$ (269.30)

^{1}H-NMR ($CDCl_3$, 400 MHz): δ = 1.74 (3H, d, *J* = 6.7 Hz, CH_3 *Hxd*), 3.74 (2H, s, α-CH_2 *Gly*), 3.97 (2H, s, α-CH_2 *Gly*), 3.99 (2H, s, α-CH_2 *Gly*), 4.62 (2H, d, *J* = 6.6 Hz, (CH_2 *Hxd*), 5.61

(1H, ddd, , J = 13.9, 6.6, 6.6 Hz, CH=C*H*CH_3), 5.74-5.81 (1H, m, CH_2C*H*=CH), 6.06 (1H, dd, J = 15.0, 10.5 Hz, C*H*=$CHCH_3$) and 6.27 (1H, dd, J = 15.3, 10.4 Hz, CH_2CH=C*H*) ppm. **MALDI-TOF**: m/z 270.5 $[M+H]^+$ (calc. 270.1), 292.5 $[M+Na]^+$ (calc. 292.1), 308.5 $[M+K]^+$ (calc. 308.1).

6.2.4. Preparation of the *C*-terminal hexadienyl ester peptides in solid phase using safety-catch strategy

6.2.4.1. *Solid-phase synthesis on sulfonamide linker resin*

Optimized loading of Fmoc-glycine to sulfamylbutyryl resin.

After swelling with DCM for 1h, the sulfamyl resin was treated overnight with a solution containing 7 equivalents of Fmoc-glycine, 7 equivalents *N*-methylimidazole and 7 equivalents DIC in DCM:DMF (4:3, 10mL/g resin). The resin was washed with DMF (3x), DCM (3x) and MeOH (3x) and finally dried under high vacuum. Estimation of amino acid attachment by UV measurements revealed a substitution level of 0.8 mmol Fmoc-glycine/g, conresponding to quantitative loading of the resin.

Peptide chain assembly

The general procedure was employed (Section 7.2.1).

Peptide cleavage from sulfamylbutyryl resin

1) Activation: The fully-protected peptidyl resin was swelled in DCM (1h) and washed 3 times with NMP. Iodoacetonitrile (25 equivalents relative to resin loading) and DIPEA (10 equivalents) were dissolved in NMP (4mL/mmol ICH_2CN) and filtered through basic alumina. The resulting solution was added to resin and the syringe was shielded from light. The mixture was shaken at room temperature for 18-24h. The resin was washed with NMP (5x) and THF (3x) and directly used for cleavage.

2) Cleavage: The activated resin was transferred to a round bottom flask and treated with a solution of *trans*,*trans*-2,4-hexadien-1-ol (20 equivalents) and DMAP (0.5 equivalent) in dry THF (10mL/g resin) during 1 day. The resin was removed by filtration and washed several times with THF. The filtrates were combined and THF was removed under reduced pressure.

The crude product was purified by reversed-phase HPLC (unless otherwise stated). Fractions containing the product (analysed by MALDI) were combined and dried by lyophilization.

Fmoc-Lys(Fmoc)-Pro-Phe-Leu-Gly-OHxd (14a)

Starting from 122mg Fmoc-Gly loaded sulfonamide resin (0.075mmol).

HPLC purification: gradient (time/solvent B) of 0min (10%B) → 10min (60%B) → 25min (100%B), retention time: 23.5min.

Yield: 35mg colorless solid (0.032mmol, 43%)

$C_{64}H_{72}N_6O_{10}$ (1085.29)

^{1}H-NMR ($CDCl_3$, 400 MHz): broad peaks: δ = 0.62-0.80 (6H, *2x* CH_3 *Leu*), 1.10-2.10 (16H, δ-CH_2 *Lys* + γ-CH_2 *Lys* + β-CH_2 *Leu* + γ-CH *Leu* + CH_3 *Hxd* + β-CH_2 *Lys* + β-CH_2 *Pro* + γ-CH_2 *Pro*), 2.9-3.2 (4H, β-CH_2 *Phe* + ε-CH_2 *Lys*), 3.45-3.70 (2H, ε-CH_2 *Pro*), 3.75-3.90 (2H, α-CH_2 *Gly*), 4.00-4.55 (12H, α-CH *Lys* + α-CH *Pro* + α-CH *Leu* + α-CH *Phe* + CH_2 *Hxd* + *2x* CH_2 *Fmoc* + *2x* CH *Fmoc*), 5.36-5.49 (1H, CH=C*H*CH_3), 5.60-5.68 (1H, CH_2C*H*=CH), 5.82-5.94 (1H, C*H*=$CHCH_3$), 6.06-6.13 (1H, CH_2CH=C*H*), 7.00-7.10 (5H, Ar *Phe*), 7.13-7.23 (4H, *2x* Ar *Fmoc*), 7.25-7.35 (4H, *2x* Ar *Fmoc*), 7.40-7.55 (4H, *2x* Ar *Fmoc*), 7.60-7.70 (4H, *2x* Ar *Fmoc*) ppm and amide protons gave weak signals.

LC-MS (Method B): t_r = 11.6 min, m/z 1085.3 $[M+H]^+$ (calc. 1085.5)

Fmoc-Pro-Cys(*St*Bu)-Ser-Met-Gly-OHxd (14b)

Starting from 151mg Fmoc-Gly loaded sulfonamide resin (0.10mmol). Product was cleaved from the resin as Fmoc-Pro-Cys(StBu)-Ser(Trt)-Met-Gly-OHxd and purified by RP-HPLC. Resulting fractions containing product were combined and concentrated in vacuum. After evaporation of the solvent, partial Trt-deprotection was detected (MALDI and LC-MS). In order to achieve complete removal of the Trt group, the crude product (58mg) was further treated with 1mL DCM:TFA:TIS (100:1:5) for 2h at room temperature. Solvents were removed in vacuum and product was purified by RP-HPLC.

HPLC purification: gradient (time/solvent B) of 0min (10%B) → 10min (60%B) → 25min (100%B), retention time: 14.5 min.

Yield: 18mg colorless solid (0.02mmol, 20%)

$C_{43}H_{57}N_5O_9S_3$ (884.14)

^{1}H-NMR (CD_3OD, 400 MHz): δ = 1.32 (9H, s, *tBu*), 1.73 (3H, d, *J* = 6.6 Hz, CH_3 *Hxd*), 1.86-2.00 (4H, m, γ-CH_2 *Pro* + β-CH_2 *Pro*), 2.06 (3H, s, CH_3 *Met*), 2.09-2.38 (2H, m, β-CH_2 *Met*), 2.44-2.65 (2H, m, 2H, γ-CH_2 *Met*), 3.3 (β-CH_2 *Cys* peak overlapped with solvent peak), 3.41-

3.70 (2H, m, ε-CH_2 *Pro*), 3.78-3.88 (2H, m, β-CH_2 *Ser*), 3.92 (2H, br s, α-CH_2 *Gly*), 4.18-4.56 (6H, m, CH *Fmoc* + CH_2 *Fmoc* + *3x* α-CH), 4.59 (2H, d, *J* = 6.5 Hz, CH_2 *Hxd*), 4.62-4.66 (1H, m, α-CH *Cys*), 5.57-5.63 (1H, m, CH=C*H*CH_3), 5.70-5.79 (1H, m, CH_2C*H*=CH), 6.01-6.08 (1H, dd, *J* = 15.0, 10.4 Hz, C*H*=$CHCH_3$), 6.21-6.28 (1H, dd, *J* = 15.0, 10.3 Hz, CH_2CH=C*H*), 7.32 (2H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.40 (2H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.65 (2H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.80 (2H, d, *J* = 7.4 Hz, Ar *Fmoc*) ppm and weak amide peaks.

LC-MS (Method B): t_r = 7.3 min, m/z 884.4 $[M+H]^+$ (calc. 884.3)

MALDI-TOF: m/z 906.8 $[M+Na]^+$ (calc. 906.3), 922.8 $[M+K]^+$ (calc. 922.4)

Fmoc-Lys(Fmoc)-Leu-Gly-Phe-Ala-Gly-OHxd (14c)

Starting from 51mg Fmoc-Gly loaded sulfonamide resin (0.043mmol).

HPLC purification: gradient (time/solvent B) of 0min (10%B) → 3min (50%B) → 17min (100%B), retention time: 14.3 min.

Yield: 16mg colorless solid (0.014mmol, 33%)

$C_{64}H_{73}N_7O_{11}$ (1116.3)

^{1}H-NMR (DMSO-d_6, 400 MHz): δ = 0.81 (3H, d, *J* = 6.5 Hz, CH_3 *Leu*), 0.82 (3H, d, *J* = 6.5 Hz, CH_3 *Leu*), 1.23 (3H, d, *J* = 7.1 Hz, CH_3 *Ala*), 1.22-1.65 (9H, m, *3x* CH_2 *Lys* + β-CH_2 *Leu* + γ-CH *Leu*), 1.70 (3H, d, *J* = 6.6 Hz, CH_3 *Hxd*), 2.71-2.77 (1H, m, β-CH_2 *Phe*), 2.92-2.99 (2H, m, CH_2 *Lys*), 2.99-3.05 (2H, m, β-CH_2 *Phe*), 3.55-3.71 (2H, m, α-CH_2 *Gly*), 3.77-3.91 (2H, m, α-CH_2 *Gly*), 3.94-4.00 (1H, m, α-CH), 4.14-4.35 (8H, m, *2x* CH *Fmoc* + *2x* α-CH + *2x* CH_2 *Fmoc*), 4.51-4.55 (1H, m, α-CH), 4.56 (2H, d, *J* = 6.2 Hz, CH_2 *Hxd*), 5.60 (1H, ddd, *J* = 12.8, 6.4, 6.4 Hz, CH=C*H*CH_3), 5.69-5.78 (1H, m, CH_2C*H*=CH), 6.04 (1H, dd, *J* = 14.9, 10.2 Hz, C*H*=$CHCH_3$), 6.27 (1H, dd, *J* = 15.0, 10.3 Hz, CH_2CH=C*H*), 7.13-7.26 (5H, m, Ar *Phe*), 7.31 (4H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.40 (4H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.66-7.71 (4H, m, Ar *Fmoc*), 7.84-7.86 (1H, m, NH), 7.87 (4H, d, *J* = 7.4 Hz, Ar *Fmoc*), 7.98 (1H, d, *J* = 8.2 Hz, NH), 8.05 (1H, t, *J* = 5.3 Hz, NH), 8.13 (1H, t, *J* = 5.7 Hz, NH) and 8.19 (1H, d, *J* = 7.6 Hz, NH) ppm.

LC-MS (Method B): t_r = 11.0 min, m/z 1116.2 $[M+H]^+$ (calc. 1116.5)

MALDI-TOF: m/z 1155.1 $[M+K]^+$ (calc. 1155.5)

Fmoc-Lys(Fmoc)-Leu-Gly-Lys(Mtt)-Ala-Gly-OHxd (14d)

Starting from 59mg Fmoc-Gly loaded sulfonamide resin (0.05mmol).

HPLC purification: gradient (time/solvent B) of 0min (10%B) → 3min (50%B) → 17min (100%B), retention time: 12.7min.

Yield: 21mg colorless solid (0.016mmol, 32%)

$C_{81}H_{92}N_8O_{11}$ (1353.64)

LC-MS (Method B): t_r = 9.6 min, m/z 1353.4 $[M+H]^+$ (calc. 1353.7)

MALDI-TOF: m/z 1392.2 $[M+K]^+$ (calc. 1391.8), 1136.1 $[(M-Mtt)+K]^+$ (calc. 1235.7) + Mtt peak.

FAB-LRMS: m/z 1352.79 $[M]^+$ (calc. 1352.69).

Fmoc-Lys(Fmoc)-Cys(StBu)-Gly-Val-Phe-Gly-OHxd (14e)

Starting from 164mg Fmoc-Gly loaded sulfonamide resin (0.056mmol).

HPLC purification: gradient (time/solvent B) of 0min (10%B) → 3min (50%B) → 15min (100%B), retention time: 11.9min.

Yield: 14mg colorless solid (0.012mmol, 22%)

$C_{67}H_{79}N_7O_{11}S_2$ (1222.52)

^{1}H-NMR (DMF-d_7, 400 MHz): δ = 0.92 (3H, d, *J* = 6.8 Hz, CH_3 *Val*), 0.94 (3H, d, *J* = 6.8 Hz, CH_3 *Val*), 1.46 (9H, s, *tBu*), 1.63-1.72 (4H, m, γ-CH_2 *Lys* + δ-CH_2 *Lys*), 1.88 (3H, d, *J* = 6.6 Hz, CH_3 *Hxd*), 1.90-2.10 (2H, m, β-CH_2 *Lys*), 2.16-2.25 (1H, m, β-CH *Val*), 3.28-3.50 (6H, m, ε-CH_2 *Lys* + β-CH_2 *Phe* + β-CH_2 *Cys*), 4.00-4.21 (4H, m, *2x* α-CH_2 *Gly*), 4.34-4.48 (8H, m, *2x* CH *Fmoc* + *2x* α-CH + *2x* CH_2 *Fmoc*), 4.78 (2H, d, *J* = 6.4 Hz, CH_2 *Hxd*), 4.85-4.90 (2H, m, *2x* α-CH), 5.78-5.85 (1H, m, CH=C*H*CH_3), 5.90-5.98 (1H, m, CH_2C*H*=CH), 6.22-6.28 (1H, dd, *J* = 15.0, 11.0 Hz, C*H*=CHCH_3), 6.45-6.52 (1H, dd, *J* = 15.0, 10.4 Hz, CH_2CH=C*H*), 7.34 (1H, t, *J* = 7.2 Hz, Ar *Phe*), 7.41 (2H, t, *J* = 7.2 Hz, Ar *Phe*), 7.47 (2H, d, *J* = 7.0 Hz, Ar *Phe*), 7.50 (4H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.60 (4H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.84-7.97 (8H, m, Ar *Fmoc* + *4x* CONH), 8.05 (4H, d, *J* = 7.4 Hz, Ar *Fmoc*), 8.35 (1H, t, *J* = 6.0 Hz, CONH), 8.49 (1H, t, *J* = 5.8 Hz, CONH) and 8.66 (1H, d, *J* = 7.5 Hz, CONH) ppm.

LC-MS (Method B): t_r = 11.8 min, m/z 1222.2 $[M]^+$ (calc. 1222.5), 1244.5 $[M+Na]^+$ (calc. 1244.5)

FAB-LRMS: 1221.7 $[M]^+$ (calc. 1221.5).

$[\alpha]_D^{20}$: - 13.5 (c = 0.2, MeOH)

Fmoc-Lys(Fmoc)-Phe-Pro-Ile-Gly-Leu-Phe-Gly-OHxd (14f) and Fmoc-Lys(Fmoc)-Phe-Pro-Ile-Gly-Leu-Gly-Phe-Gly-OHxd (14g)

Starting from mg Fmoc-Gly loaded sulfonamide resin (0.09mmol). The coupling of the second glycine residue was incomplete, yielding a mixture of two peptides **14f** and **14g**.

HPLC purification: gradient (time/solvent B) of 0min (10%B) → 3min (50%B) → 20min (100%B), retention time: 17.3min (**14f**) and 18.1min (**14g**).

Yield: 16% overall yield: 5.4mg of **14f** (0.010mmol, 4%) and 14.5mg of **14g** (0.0037mmol, 12%).

◈ **14f**: $C_{83}H_{98}N_{10}O_{14}$ (1459.73)

MALDI-TOF: m/z 1481.7 $[M+Na]^+$ (calc. 1481.7), 1497.7 $[M+K]^+$ (calc. 1497.7).

◈ **14g**: $C_{81}H_{95}N_9O_{13}$ (1402.67)

MALDI-TOF: m/z 1424.7 $[M+Na]^+$ (calc. 1424.7), 1440.7 $[M+K]^+$ (calc. 1440.7).

Fmoc-Lys(Fmoc)-Phe-Pro-Ile-Lys(Mtt)-Leu-Gly-Lys(Mtt)-Ala-Gly-OHxd (14h)

Starting from 59mg Fmoc-Gly loaded sulfonamide resin (0.05mmol).

HPLC purification: gradient (time/solvent B) of 0min (10%B) → 3min (50%) → 20min (100%B), retention time: 14.5min.

Yield: 9.6mg colorless solid (0.005mmol, 10%)

$C_{127}H_{147}N_{13}O_{15}$ (2095.6)

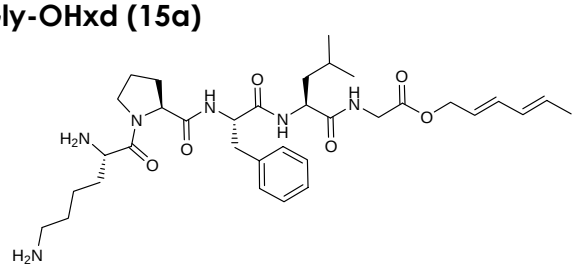
MALDI-TOF: m/z 2135.3 $[M+K]^+$ (calc. 2134.2), 1877.3 $[(M-Mtt)+K]^+$ (calc. 1877.1), 1621.2 $[(M-2Mtt)+K]^+$ (calc. 1621.0) + Mtt peak.

6.2.4.2. *Removal of the protecting groups of the C-terminal hexadienyl ester peptides*

General method for Fmoc deprotection of hexadienyl peptides

5-30mg Fmoc-protected peptide **14** was treated with 0.5-2.0mL 20% piperidine in DMF or DCM for 40 minutes at room temperature. The reaction mixture was coevaporated with methanol to remove excess of piperidine. The product was purified by reversed-phase HPLC, fractions containing the product (evaluated by MALDI-TOF) were combined and dried by lyophilization.

H-Lys-Pro-Phe-Leu-Gly-OHxd (15a)



33 mg of **14a** (0.032mmol) were treated with 1.5mL 20% piperidine in DCM.

HPLC purification: gradient (time/solvent B) of 0min (5%B) → 7min (5%B) → 15min (100%B), retention time: 12.1min.

Yield: 15mg colorless solid (0.023mmol, 72%)

$C_{34}H_{52}N_6O_6$ (640.81)

^{1}H-NMR (D_2O, 400 MHz):): δ = 0.91 (3H, d, *J* = 6.1 Hz, CH_3 *Leu*), 0.94 (3H, d, *J* = 6.2 Hz, CH_3 *Leu*), 1.46-1.82 (7H, m, δ-CH_2 *Lys* + γ-CH_2 *Lys* + β-CH_2 *Leu* + γ-CH *Leu*), 1.74 (3H, *J* = 6.7 Hz, CH_3 *Hxd*), 1.84-2.23 (6H, m, β-CH_2 *Pro* + γ-CH_2 *Pro* + β-CH_2 *Lys*), 2.90-3.00 (2H, m, ε-CH_2 *Lys*), 3.02 (1H, dd, *J* = 13.9, 7.7 Hz, β-CH_2 *Phe*), 3.13 (1H, dd, *J* = 13.9, 6.5 Hz, β-CH_2 *Phe*), 3.56-3.63 (1H, m ε-CH_2 *Pro*), 3.67-3.73 (1H, m, ε-CH_2 *Pro*), 3.81-3.96 (2H, α-CH_2 *Gly*), 4.25 (1H, t, *J* = 6.1 Hz, α-CH), 4.42 (1H, dd, *J* = 9.5, 5.5 Hz, α-CH), 4.50 (1H, t, *J* = 5.2 Hz, α-CH), 4.57 (1H, t, *J* = 6.6 Hz, α-CH), 4.61 (2H, d, *J* = 6.6 Hz, CH_2 *Hxd*), 5.61 (1H, ddd, *J* = 13.5, 6.5, 6.5 Hz, CH=C*H*CH_3), 5.72-5.81 (1H, m, CH_2C*H*=CH), 6.06 (1H, dd, *J* = 15.0, 10.4 Hz, C*H*=$CHCH_3$), 6.28 (1H, dd, *J* = 15.0, 10.4 Hz, CH_2CH=C*H*) and 7.17-7.30 (5H, m, Ar *Phe*) ppm.

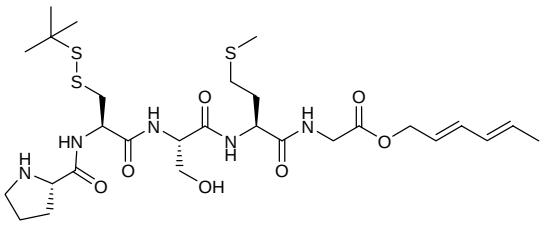
LC-MS (Method B): t_r = 5.1 min, m/z 641.4 $[M+H]^+$ (calc. 641.4)

MALDI-TOF: m/z 641.8 $[M+H]^+$ (calc. 641.4), 679.8 $[M+K]^+$ (calc. 679.5)

FAB-HRMS: m/z 640.3945 $[M]^+$ (calc. 640.3948).

$[\alpha]_D^{20}$: - 27.9 (c = 0.3, MeOH)

H-Pro-Cys(*St*Bu)-Ser-Met-Gly-OHxd (15b)



6 mg of **14b** (6.8μmol) were treated with 0.5mL 20% piperidine in DCM.

HPLC purification: gradient (time/solvent B) of 0min (5%B) → 7min (5%B) → 15min (100%B), retention time: 11.3min.

Yield: 2.3mg colorless solid (3.5μmol, 51%)

$C_{28}H_{47}N_5O_7S_3$ (661.90)

^{1}H-NMR (CD_3OD, 400 MHz): δ = 1.35 (9H, s, *tBu*), 1.75 (3H, d, *J* = 6.6 Hz, CH_3 *Hxd*), 1.90-2.21 (6H, m, γ-CH_2 *Pro* + β-CH_2 *Pro* + β-CH_2 *Met*), 2.09 (3H, s, CH_3 *Met*), 2.38-2.64 (2H, m,

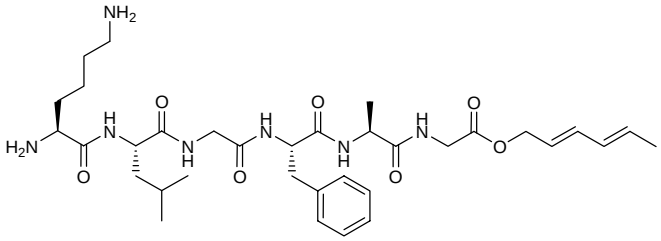
γ-CH_2 *Met*), 3.02 (1H, dd, *J* = 13.7, 9.1 Hz, β-CH_2 *Cys*), 3.23 (1H, dd, *J* = 13.6, 5.0 Hz, β-CH_2 *Cys*), 3.30-3.45 (2H, m, ε-CH_2 *Pro*), 3.75 (1H, dd, *J* = 10.9, 5.9 Hz, β-CH_2 *Ser*), 3.85 (1H, dd, *J* = 10.9, 5.4 Hz, β-CH_2 *Ser*), 3.94 (2H, m, α-CH_2 *Gly*), 4.31 (1H, dd, *J* = 8.4, 6.3 Hz, α-CH *Pro*), 4.40 (1H, t, *J* = 5.7 Hz, α-CH *Met*), 4.56 (1H, dd, *J* = 9.0, 4.9 Hz, α-CH *Ser*), 4.62 (2H, d, *J* = 6.7 Hz, CH_2 *Hxd*), 4.69 (1H, dd, *J* = 9.0, 5.1 Hz, α-CH *Cys*), 5.61 (1H, ddd, *J* = 13.2, 6.5, 6.5 Hz, CH=C*H*CH_3), 5.73-5.81 (1H, m, CH_2C*H*=CH), 6.06 (1H, dd, *J* = 15.0, 10.4 Hz, C*H*=$CHCH_3$) and 6.27 (1H, dd, *J* = 15.1, 10.3 Hz, CH_2CH=C*H*) ppm.

ESI-MS: m/z 662.2 $[M+H]^+$ (calc. 662.3), 684.3 $[M+Na]^+$ (calc. 684.2)

MALDI-TOF: m/z 662.8 $[M+H]^+$ (calc. 662.3), 684.8 $[M+Na]^+$ (calc. 684.2), 700.8 $[M+K]^+$ (calc. 700.4)

$[\alpha]_D^{20}$: - 17.1 (c = 0.1, MeOH)

H-Lys-Leu-Gly-Phe-Ala-Gly-OHxd (15c)



7 mg of **14c** (0.0063 mmol) were treated with 1.5mL 20% piperidine in DMF.

HPLC purification: gradient (time/solvent B) of 0min (5%B) → 7min (5%B) → 15min (100%B), retention time: 9.4 min.

Yield: 3.5 mg colorless solid (0.0052mmol, 83%)

$C_{34}H_{53}N_7O_7$ (671.83)

^{1}H-NMR (CD_3OD, 400 MHz):): δ = 0.95 (3H, d, *J* = 6.5 Hz, CH_3 *Leu*), 0.97 (3H, d, *J* = 6.5 Hz, CH_3 *Leu*), 1.38 (3H, d, *J* = 7.2 Hz, CH_3 *Ala*), 1.45-1.91 (9H, m, *3x* CH_2 *Lys* + β-CH_2 *Leu* + γ-CH *Leu*), 1.74 (3H, d, *J* = 6.4 Hz, CH_3 *Hxd*), 2.90-3.20 (4H, m, β-CH_2 *Phe* + CH_2 *Lys*), 3.88-3.97 (4H, m, *2x* α-CH_2 *Gly*), 4.34-4.42 (2H, m, *2x* α-CH), 4.58-4.61 (2H, m, *2x* α-CH), 4.62 (2H, d, *J* = 6.2 Hz, CH_2 *Hxd*), 5.62 (1H, ddd, *J* = 13.6, 6.6, 6.6 Hz, CH=C*H*CH_3), 5.72-5.81 (1H, m, CH_2C*H*=CH), 6.06 (1H, dd, *J* = 15.0, 10.5 Hz, C*H*=$CHCH_3$), 6.28 (1H, dd, *J* = 15.1, 10.3 Hz, CH_2CH=C*H*) and 7.18-7.29 (5H, m, Ar *Phe*) ppm.

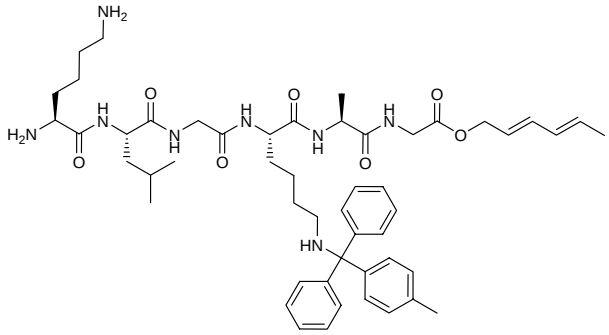
HPLC (Method A): 16.8 min.

ESI-MS: m/z 671.4 $[M]^+$ (calc. 671.4)

MALDI-TOF: m/z 672.9 $[M+H]^+$ (calc. 672.4), 694.9 $[M+Na]^+$ (calc. 694.4), 710.9 $[M+K]^+$ (calc. 710.4)

$[\alpha]_D^{20}$**:** - 7.2 (c = 0.3, MeOH)

H-Lys-Leu-Gly-Lys(Mtt)-Ala-Gly-OHxd (15d)



21 mg of **14d** (15.5µmol) were treated with 1.5mL 20% piperidine in DMF.

HPLC purification: gradient (time/solvent B) of 0min (10%B) → 3min (10%B) → 17min (100%B), retention time: 10.2min.

Yield: 10.2 mg colorless solid (11.2µmol, 73%)

$C_{51}H_{72}N_8O_7$ (909.17)

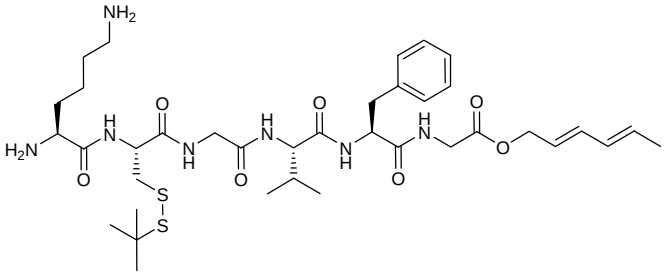
^{1}H-NMR (CD_3OD, 400 MHz): δ = 0.94 (3H, d, *J* = 6.4 Hz, CH_3 *Leu*), 0.96 (3H, d, *J* = 6.4 Hz, CH_3 *Leu*), 1.29-1.40 (2H, m, γ-CH_2 *Lys*), 1.39 (3H, d, *J* = 7.2 Hz, CH_3 *Ala*), 1.44-1.55 (2H, m, γ-CH_2 *Lys*), 1.58-1.93 (11H, m, *2x* δ-CH_2 *Lys* + β-CH_2 *Leu* + *2x* β-CH_2 *Lys* + γ-CH *Leu*), 1.74 (3H, d, *J* = 6.6 Hz, CH_3 *Hxd*), 2.38 (3H, s, CH_3 *Mtt*), 2.85-2.99 (4H, m, *2x* ε-CH_2 *Lys*), 3.77-3.96 (5H, m, *2x* α-CH_2 *Gly* + α-CH), 4.26-4.43 (3H, m, *3x* α-CH), 4.58 (2H, d, *J* = 6.6 Hz, CH_2 *Hxd*), 5.59 (1H, ddd, *J* = 13.3, 6.6, 6.6 Hz, CH=C*H*CH_3), 5.71-5.80 (1H, m, CH_2C*H*=CH), 6.05 (1H, dd, *J* = 15.0, 10.4 Hz, C*H*=$CHCH_3$), 6.28 (1H, dd, *J* = 15.0, 10.6 Hz, CH_2CH=C*H*), 7.18 (2H, d, *J* = 8.4 Hz, Ar *Mtt*), 7.27-7.33 (6H, m, Ar *Mtt*) and 7.43-749 (6H, m, Ar *Mtt*) ppm.

HPLC (Method A): t_r = 18.8 min

MALDI-TOF: m/z 672.9 $[M+H]^+$ (calc. 672.4), 694.9 $[M+Na]^+$ (calc. 694.4), 710.9 $[M+K]^+$ (calc. 710.5).

FAB-HRMS: m/z 909.5684 $[M+H]^+$ (calc. 909.5602).

H-Lys-Cys(*St*Bu)-Gly-Val-Phe-Gly-OHxd (15e)



11mg of **14e** (9.0μmol) were treated with 1mL 20% piperidine in DMF for 40 minutes.

HPLC purification: gradient (time/solvent B) of 0min (10%B) → 3min (10%B) → 15min (100%B), retention time 9.6min.

Yield: 5.6 mg colorless solid (7.2μmol, 80%)

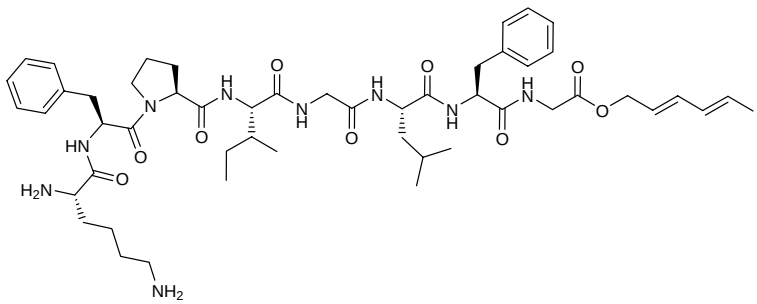
$C_{37}H_{59}N_7O_7S_2$ (778.04)

^{1}H-NMR (CD_3OD, 400 MHz): δ = 0.73 (3H, d, *J* = 6.8 Hz, CH_3 *Val*), 0.82 (3H, d, *J* = 6.8 Hz, CH_3 *Val*), 1.34 (9H, s, *tBu*), 1.45-1.55 (2H, m, γ-CH_2 *Lys*), 1.63-1.70 (2H, m, δ-CH_2 *Lys*), 1.74 (3H, d, *J* = 6.6 Hz, CH_3 *Hxd*), 1.84-1.95 (2H, m, β-CH_2 *Lys*), 2.01-2.10 (1H, m, β-CH *Val*), 2.94 (2H, t, *J* = 7.6 Hz, ε-CH_2 *Lys*), 2.98-3.07 (2H, m, β-CH_2 *Phe*), 3.20 (2H, dt, *J* = 13.9, 6.4 Hz, β-CH_2 *Cys*), 3.68 (1H, d, *J* = 16.1 Hz, α-CH *Val*), 3.94 (2H, s, α-CH_2 *Gly*), 3.98 (1H, t, *J* = 6.4 Hz, α-CH *Lys*), 4.08 (2H, t, *J* = 8.2 Hz, α-CH_2 *Gly*), 4.63 (2H, d, *J* = 6.6 Hz, CH_2 *Hxd*), 4.63-4.66 (2H, m, α-CH *Cys* + α-CH *Phe*), 5.58-5.65 (1H, m, CH=C*H*CH_3), 5.72-5.82 (1H, m, CH_2C*H*=CH), 6.03-6.10 (1H, dd, *J* = 15.0, 10.4 Hz, C*H*=$CHCH_3$), 6.25-6.32 (1H, dd, *J* = 15.1, 10.5 Hz, CH_2CH=C*H*) and 7.18-7.30 (m, 5H, Ar *Phe*) ppm.

MALDI-TOF: m/z 778.8 $[M+H]^+$ (calc. 778.4), 800.8 $[M+Na]^+$ (calc. 800.4), 816.7 $[M+K]^+$ (calc. 816.5)

$[\alpha]_D^{20}$: - 9.0 (c = 0.1, MeOH)

H-Lys-Phe-Pro-Ile-Gly-Leu-Phe-Gly-OHxd (15f)



14.5mg of **14f** (10.3μmol) were treated with 1mL 20% piperidine in DMF.

HPLC purification: gradient (time/solvent B) of 0min (5%B) → 15min (50%B) → 17min (100%B), retention time 14.2min.

Yield: 9.0mg colorless solid (9.4μmol, 91%)

$C_{51}H_{75}N_9O_9$ (958.20)

^{1}H-NMR (CD_3OD, 400 MHz): δ = 0.82 (3H, dd, *J* = 6.3, 4.3 Hz, CH_3 *Ile*), 0.87 (3H, d, *J* = 6.3 Hz, CH_3 *Ile*), 0.91-1.00 (6H, 2x CH_3 *Leu*), 1.21-1.71 (10H, m, CH_2 *Ile* + δ-CH_2 *Lys* + γ-CH_2 *Lys* + β-CH_2 *Leu* + β-CH_2 *Lys*), 1.74 (3H, *J* = 6.6 Hz, CH_3 *Hxd*), 1.83-2.14 (6H, m, β-CH_2 *Pro* + γ-CH_2 *Pro* + γ-CH *Leu* + β-CH *Ile*), 2.90-3.04 (4H, m, *2x* 1H β-CH_2 *Phe* + ε-CH_2 *Lys*), 3.17-3.26 (2H, m, *2x* 1H β-CH_2 *Phe*), 3.32-3.52 (1H, m, ε-CH_2 *Pro*), 3.76-3.95 (6H, m, ε-CH_2 *Pro* + *2x* α-CH_2 *Gly*), 4.08 (1H, t, *J* = 7.0 Hz, α-CH *Pro*), 4.22-4.28 (2H, m, *2x* α-CH), 4.53-4.67 (2H, m, *2x* α-CH), 4.62 (2H, d, *J* = 6.5 Hz, CH_2 *Hxd*), 5.61 (1H, ddd, *J* = 13.6, 6.4, 6.4 Hz, CH=C*H*CH_3), 5.72-5.81 (1H, m, CH_2C*H*=CH), 6.06 (1H, dd, *J* = 15.0, 10.5 Hz, C*H*=CHCH_3), 6.28 (1H, dd, *J* = 15.2, 10.5 Hz, CH_2CH=C*H*) and 7.17-7.34 (5H, m, Ar *Phe*) ppm.

HPLC (Method A): t_r = 18.9 min

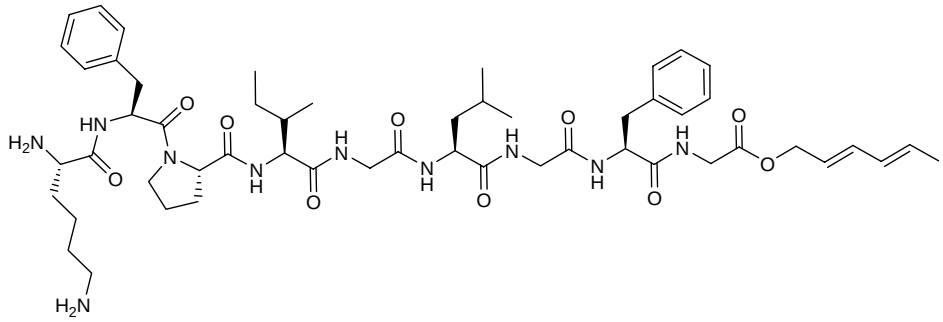
ESI-MS: m/z 958.7 $[M+H]^+$ (calc. 958.6)

MALDI-TOF: m/z 958.9 $[M+H]^+$ (calc. 958.6), 980.9 $[M+Na]^+$ (calc. 980.6), 996.8 $[M+K]^+$ (calc. 996.5).

FAB-LRMS: m/z 958.75 $[M]^+$ (calc. 958.58).

$[\alpha]_D^{20}$: - 28.6 (c = 0.3, MeOH)

H-Lys-Phe-Pro-Ile-Gly-Leu-Gly-Phe-Gly-OHxd (15g)



5.4mg of **14g** (3.7μmol) were treated with 1mL 20% piperidine in DMF.

HPLC purification: gradient (time/solvent B) of 0min (5%B) → 15min (50%B) → 17min (100%B), retention time: 14.3min.

Yield: 3.5 mg colorless solid (3.4μmol, 92%)

$C_{53}H_{78}N_{10}O_{10}$ (1015.25)

^{1}H-NMR (CD_3OD, 400 MHz): δ = 0.82 (3H, dd, *J* = 6.3, 3.1 Hz, CH_3 *Ile*), 0.87 (3H, d, *J* = 6.5 Hz, CH_3 *Ile*), 0.90-1.00 (6H, m, *2x* CH_3 *Leu*), 1.17-1.72 (10H, m, CH_2 *Ile* + δ-CH_2 *Lys* + γ-CH_2 *Lys* + β-CH_2 *Leu* + β-CH_2 *Lys*), 1.73 (3H, d, *J* = 6.6 Hz, CH_3 *Hxd*), 1.83-2.04 (5H, m, β-CH_2 *Pro* + γ-CH_2 *Pro* + γ-CH *Leu*), 2.04-2.13 (1H, m, β-CH *Ile*), 2.89-3.04 (4H, m, *2x* 1H β-CH_2 *Phe* + ε-CH_2 *Lys*), 3.17-3.26 (2H, m, *2x* 1H β-CH_2 *Phe*), 3.32-3.52 (1H, m, ε-CH_2 *Pro*), 3.72-4.02 (8H, m, ε-CH_2 *Pro* + *3x* α-CH_2 *Gly*), 4.10 (1H, m, α-CH), 4.21-4.28 (2H, m, *2x* α-CH), 4.49-4.56 (2H, m, *2x* α-CH), 4.62 (2H, d, *J* = 6.5 Hz, CH_2 *Hxd*), 5.62 (1H, ddd, *J* = 13.2, 6.7, 6.7 Hz, CH=C*H*CH_3), 5.71-5.82 (1H, m, CH_2C*H*=CH), 6.06 (1H, dd, *J* = 14.9, 10.5 Hz, C*H*=$CHCH_3$), 6.27 (1H, dd, *J* = 14.6, 10.0 Hz, CH_2CH=C*H*) and 7.19-7.35 (5H, m, Ar *Phe*) ppm.

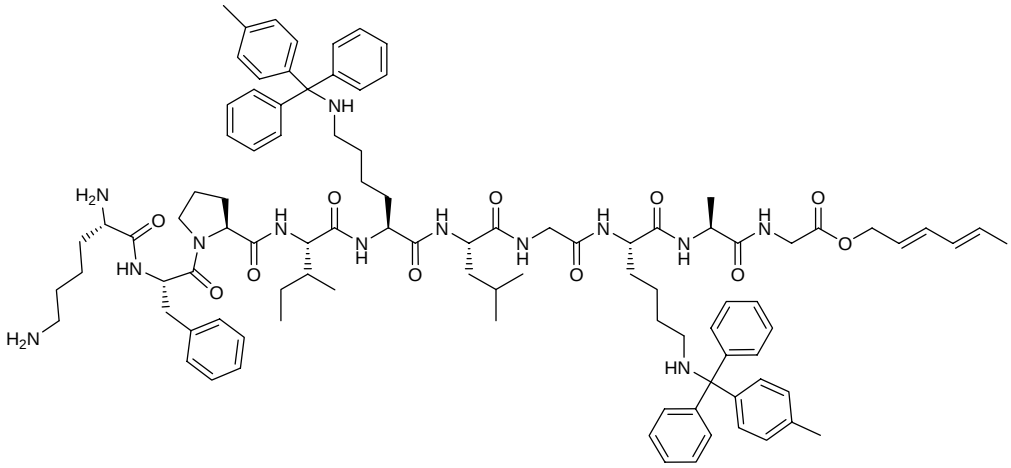
HPLC (Method A): t_r = 18.6 min

LC-MS (Method A): t_r = 15.7 min, m/z 1015.7 $[M+H]^+$ (calc. 1015.6)

MALDI-TOF: m/z 1015.9 $[M+H]^+$ (calc. 1015.6), 1037.8 $[M+Na]^+$ (calc. 1037.6), 1053.8 $[M+K]^+$ (calc. 1053.5).

$[\alpha]_D^{20}$: - 25.7 (c = 0.2, MeOH)

H-Lys-Phe-Pro-Ile-Lys(Mtt)-Leu-Gly-Lys(Mtt)-Ala-Gly-OHxd (15h)



9.6mg of **14h** (4.6μmol) were treated with 1mL 20% piperidine in DMF.

HPLC purification: gradient (time/solvent B) of 0min (5%B) → 15min (50%B) → 20min (100%B), retention time: 16.8min.

Yield: 5.3 mg colorless solid (3.2μmol, 71%)

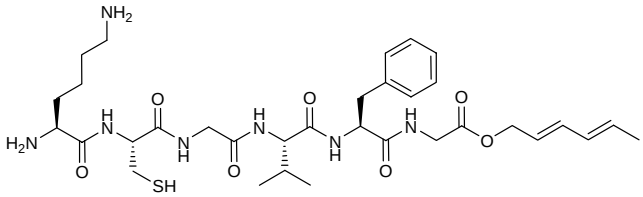
$C_{97}H_{127}N_{13}O_{11}$ (1651.13)

^{1}H-NMR (CD_3OD, 400 MHz): δ = 0.86-0.94 (12H, m, *2x* CH_3 *Ile* + 2x CH_3 *Leu*), 1.16-1.40 (4H, m, CH_2 *Ile* + γ-CH_2 *Lys*), 1.43-2.00 (23H, m, *2x* γ-CH_2 *Lys* + 3x δ-CH_2 *Lys* + *3x* β-CH_2 *Lys* + γ-CH *Leu* + β-CH_2 *Leu* + γ-CH_2 *Pro* + β-CH_2 *Pro*), 1.74 (3H, d, *J* = 6.6 Hz, CH_3 *Hxd*), 2.06-2.10 (1H, m, β-CH *Ile*), 2.36 (6H, br s, *2x* CH_3 *Mtt*), 2.85-3.02 (7H, m, β-CH_2 *Phe* + *3x* ε-CH_2 *Lys*), 3.18 (1H, dd, *J* = 14.0, 4.7 Hz, β-CH_2 *Phe*), 3.44-3.62 (2H, m, ε-CH_2 *Pro*), 3.75-3.96 (5H, m, *2x* α-CH_2 *Gly* + α-CH *Lys*), 4.09-4.20 (1H, m, α-CH *Pro*), 4.22-4.39 (4H, m, α-CH *Leu* + α-CH *Ile*, *2x* α-CH *Lys*), 4.49-4.52 (1H, m, α-CH *Ala*), 4.56 (2H, d, *J* = 6.5 Hz, CH_2 *Hxd*), 4.55-4.60 (1H, m, α-CH *Phe*), 5.57 (1H, ddd, *J* =14.1, 6.6, 6.6 Hz CH=C*H*CH_3), 5.70-5.79 (1H, m, CH_2C*H*=CH), 6.05 (1H, dd, *J* = 15.0, 10.5 Hz, C*H*=CHCH_3), 6.24 (1H, dd, *J* = 15.0, 10.4 Hz, CH_2CH=C*H*) and 7.12-7.20 (4H, m, Ar *Mtt*), 7.22-7.32 (12H, m, Ar *Mtt*), 7.35 (5H, m, Ar *Phe*), and 7.42-7.47 (12H, m, Ar *Mtt*) ppm.

ESI-MS: m/z 1651.6 $[M+H]^+$ (calc. 1651.0)

MALDI-TOF: m/z 1139.1 $[(M-2Mtt)+H]^+$ (calc. 1138.7), 1161.1 $[(M-2Mtt)+Na]^+$ (calc. 1160.7), 1177.0 $[(M-2Mtt)+K]^+$ (calc. 1176.8) + Mtt peak.

H-Lys-Cys-Gly-Val-Phe-Gly-OHxd (16)



4.2mg (5.4μmol) of peptide **15e** were dissolved in 800μl of a degassed solution of 0.1M ammonium bicarbonate: DMF (5:3) in argon atmosphere. 12μl of DTT (0.14mmol) dissolved in 200μl degassed DMF was added to the peptide solution and the mixture was stirred under argon at room temperature for 2.5h. The crude product was dried by lyophilization, redissolved in DMF, filtered and purified by reversed-phase HPLC.

Yield: 2.7 mg colorless solid (3.9μmol, 72%)

$C_{33}H_{51}N_7O_7S$ (689.87)

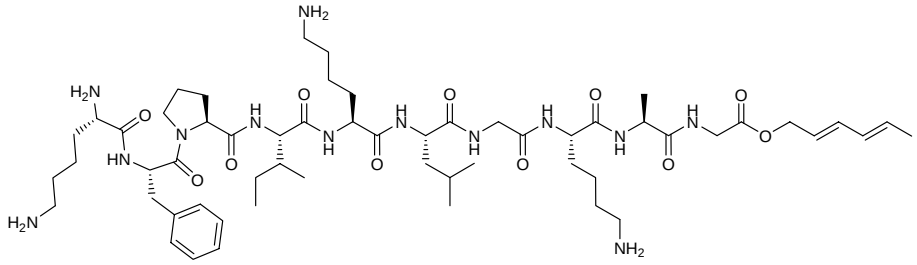
^{1}H-NMR (CD_3OD, 400 MHz): δ = 0.73 (3H, d, *J* = 6.8 Hz, CH_3 *Val*), 0.82 (3H, d, *J* = 6.8 Hz, CH_3 *Val*), 1.28-1.33 (1H, m, SH), 1.47-1.54 (2H, m, γ-CH_2 *Lys*), 1.63-1.72 (2H, m, δ-CH_2 *Lys*), 1.74 (3H, d, *J* = 6.6 Hz, CH_3 *Hxd*), 1.86-1.92 (2H, m, β-CH_2 *Lys*), 2.00-2.07 (1H, m, β-CH *Val*), 2.81-3.03 (4H, m, ε-CH_2 *Lys* + β-CH_2 *Phe*), 3.20 (2H, dt, *J* = 8.8, 5.0 Hz, β-CH_2 *Cys*), 3.77 (1H, d, *J* = 16.1 Hz, α-CH *Val*), 3.94 (2H, s, α-CH_2 *Gly*), 3.98 (1H, t, *J* = 6.4 Hz, α-CH

Lys), 4.07 (2H, t, J = 6.4 Hz, α-CH_2 *Gly*), 4.53 (2H, t, α-CH), 4.62 (2H, d, J = 6.6 Hz, CH_2 *Hxd*), 4.66 (2H, t, J = 5.0 Hz, α-CH), 5.58-5.65 (1H, m, CH=C*H*CH_3), 5.72-5.81 (1H, m, CH_2C*H*=CH), 6.03-6.10 (1H, ddd, J = 15.0, 10.4 Hz, C*H*=$CHCH_3$), 6.25-6.32 (1H, dd, J = 15.1, 10.5 Hz, CH_2CH=C*H*) and 7.18-7.30 (m, 5H, Ar *Phe*) ppm.

HPLC (Method A): t_r = 16.8 min

MALDI-TOF: m/z 691.0 $[M+H]^+$ (calc. 690.4), 713.0 $[M+Na]^+$ (calc. 712.4), 729.0 $[M+K]^+$ (calc. 728.3).

H-Lys-Phe-Pro-Ile-Lys-Leu-Gly-Lys-Ala-Gly-OHxd (17)



5.0mg of **15h** (3.0μmol) were treated with 0.5mL of DCM:TFA:TES (85:5:10) at room temperature for 35 minutes. The reaction was stopped by adding 27μl DIPEA in 1mL MeOH and directly purified by HPLC. Combined fractions containing the desired product were collected and lyophilized.

Yield: 0.53 mg colorless solid (0.47μmol, 16%)

$C_{57}H_{95}N_{13}O_{11}$ (1138.45)

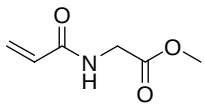
^{1}H-NMR (CD_3OD, 400 MHz): No Mtt signal

ESI-MS: m/z 1138.9 $[M+H]^+$ (calc. 1138.7)

MALDI-TOF: m/z 1138.1 $[M+H]^+$ (calc. 1138.7), 1161.1 $[M+Na]^+$ (calc. 1160.7), 1177.1 $[M+K]^+$ (calc. 1176.8)

6.2.5. Synthesis of other diene and dienophile peptides

N-acryloyl-glycine-methylester (33)



Glycine methyl ester (2.51g, 20mmol) and triethylamine (3.4mL, 24mmol) were dissolved in 100mL DCM. The solution was cooled to 0°C and acryloyl chloride (1.8mL, 22mmol) dissolved in 10mL DCM was added dropwise over 1 hour. The reaction mixture was allowed

to warm up to room temperature and stirred overnight. 200mL EtOAc were added and the formed solid was filtered off. The organic phase was washed 3x $KHSO_4$ 1M, 3x 5% $NaHCO_3$ and 1x brine, dried over $MgSO_4$ and concentrated in vacuum.

Yield: 1.3g (9.1mmol, 46%) colorless solid

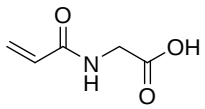
$C_6H_9NO_3$ (143.06)

^{1}H-NMR (CD_3OD, 400MHz): δ = 3.62 (3H, s, OCH_3), 3.91 (2H, d, *J* = 6.0 Hz, α-CH_2 *Gly*), 5.62 (1H, dd, *J* = 10.2, 2.1 Hz, *H*CH=CH), 6.09 (1H, dd, *J* = 17.1, 2.1 Hz, HC*H*=CH), 6.27 (1H, dd, *J* = 17.1, 10.2 Hz, CH_2=C*H*) and 8.50-8.56 (1H, m, CONH) ppm.

^{13}C-NMR (CD_3OD, 100MHz): δ = 41.3 (α-CH_2 *Gly*), 52.4 (OCH_3), 126.6 (*C*H_2=CH), 131.7 (CH_2=*C*H), 165.7 (C=O) and 171.0 (C=O) ppm.

ESI-MS: 144.07 $[M+H]^+$ (calc. 144.06)

N-acryloyl-glycine-OH (32)



N-acryloyl-glycine methyl ester **33** (0.42g, 2.9mmol) was dissolved in 7mL MeOH. 3mL NaOH 1M (3.0 mmol) were added and the mixture was stirred at room temperature during 30 minutes. MeOH was removed under reduced pressure and the resulting solution was neutralized by adding HCl 1M and lyophilized. The crude solid was washed many times with methanol. The solvent was removed under vacuum and the product was recrystallized from MeOH/Et_2O.

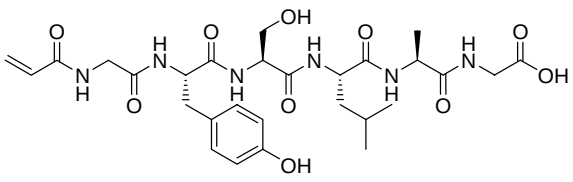
Yield: 0.32g (2.5mmol, 86%) colorless solid

$C_5H_7NO_3$ (129.11)

^{1}H-NMR (D_2O, 400MHz): δ = 3.71 (2H, br s, α-CH_2 *Gly*), 5.65 (1H, dd, *J* = 10.2, 0.6 Hz, *H*CH=CH), 6.08 (1H, dd, *J* = 17.2, 0.7 Hz, HC*H*=CH) and 6.19 (1H, dd, *J* = 17.2, 10.2 Hz, CH_2=C*H*) ppm.

^{13}C-NMR (D_2O, 100MHz): δ = 43.5 (α-CH_2 *Gly*), 127.6 (*C*H_2=CH), 130.7 (CH_2=*C*H), 168.5 (CONH) and 176.8 (CO_2H) ppm.

N-acryloyl-Gly-Tyr-Ser-Leu-Ala-Gly-OH (31)



Starting from 274mg Fmoc-glycine loaded Wang resin (0.31mmol). The same procedure for the synthesis of the maleimido-peptides was applied here (Section 6.2.2.2). *N*-acryloyl-glycine **32** was coupling using HBTU/HOBt activation.

Yield: 114mg light yellow solid (0.18mmol, 60%)

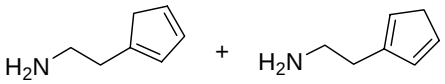
$C_{28}H_{40}N_6O_{10}$ (620.65)

^{1}H-NMR (DMSO-d_6, 400 MHz): δ = 0.83 and 0.86 (3H, d, *J* = 6.5 Hz, CH_3 *Leu*), 1.19 (3H, d, *J* = 7.1 Hz, CH_3 Ala), 1.40-1.52 (2H, m, β-CH_2 *Leu*), 1.55-1.65 (1H, m, γ-CH *Leu*), 2.59-2.67 and 2.86-2.90 (2H, dm, β-CH_2 *Tyr*), 3.53-3.82 (6H, m, α-CH_2 *Gly* + β-CH_2 *Ser* + α-CH_2 *N-acryloyl-Gly*), 4.22-4.72 (5H, m, *4x* α-CH + OH *Tyr*), 5.57 (1H, dd, *J* = 10.2, 1.9 Hz, C*H*H=CH), 6.05 (1H, dd, *J* = 17.1, 1.9 Hz, CH*H*=CH), 6.22-6.29 (1H, dd, *J* = 17.2, 10.2 Hz, CH_2=C*H*), 6.58 (2H, d, *J* = 8.4 Hz, Ar *Tyr*), 6.99 (2H, d, *J* = 8.4 Hz, Ar *Tyr*), 7.92 (1H, dd, *J* = 8.1, 3.0 Hz, CONH), 8.00-8.08 (3H, m, 3x CONH), 8.15 (1H, d, *J* = 7.7 Hz, CONH) and 8.30 (1H, t, *J* = 5.9 Hz, CONH) ppm.

LC-MS (Method A): t_r = 15.4 min; m/z: 621.1 $[M+H]^+$ (calc. 621.3).

FAB-HRMS: 643.2732 $[M+Na]^+$ (calc. 643.2704).

2-(cyclopentadienyl)-ethanamine (39)



a **b**

In an oven-dried flask under argon atmosphere, 4.64g of chloroethylamine hydrochloride (40mmol) was suspended in 100mL dry THF and the mixture was cooled to 0°C. Then 40mL of a 2M solution of sodium cyclopentadienide in THF (80mmol) was added via syringe into the reaction mixture, whose colour became clear pink. After reacting 4 hours at 0°C, the reaction was allowed to warm up to room temperature overnight. The reaction was quenched by adding 150mL of water. The organic compounds were extracted 3x EtOAc. The combined organic layers were extracted 3x 10% acetic acid. The combined acidic extractions were alkalinized from pH 5 to 13 by adding NaOH 4M. The product was extracted with 3x diethylether. The organic solution was washed with brine, dried over $MgSO_4$ and concentrated in medium vacuum (350mbar, 40°C).

Yield: 3.0g yellow oil (28mmol, 70% overall, 45% **39a** and 55% **39b** determined by NMR)

$C_7H_{11}N$ (109.17)

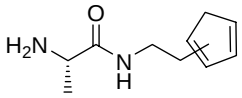
^{1}H-NMR ($CDCl_3$, 400MHz): δ = 1.41 (2H, br s, NH_2), 2.39-2.43 (2H of **39b**, m, CH_2), 2.44-2.48 (2H of **39a**, m, CH_2), 2.76-2.81 (2H, m, CH_2 + 2H of **39a**, m, CH_2 *Cp*), 2.86-2.88 (2H of

39b, m, CH_2 *Cp*), 5.97-5.99 (1H of **39b**, m, *Cp*), 6.11-6.13 (1H of **39a**, m, *Cp*), 6.16-6.19 (1H of **39a**, m, *Cp*) and 6.31-6.36 (2H of **39b**, m, *Cp* + 1H of **39a**, m, *Cp*) ppm.

^{13}C-NMR ($CDCl_3$, 100MHz): δ = 34.5/35.3, 41.6/41.8, 42.4/43.6, 127.7/128.0, 131.1/132.5, 134.2/134.5 and 144.6/146.9.

GC-MS (Method B): t_r= 3.67 min; m/z: 109 $[M]^+$

Glycine-(2-cyclopentadienyl)ethylamide (40)



1) Coupling: Fmoc-Ala-OH (1.93g, 6.2mmol) and 2-(cyclopentadienyl)-ethanamine **39** (0.61g, 5.6mmol) were dissolved in DCM (20mL) and the solution cooled to 0°C. HOBt (961mg, 6.7mmol) and DIC (0.97mL, 6.2mmol) were added and the solution was allowed to warm up to room temperature and stirred overnight. After 18 hours, the reaction mixture was concentrated in vacuum, and dissolved in EtOAc:cHex. The urea precipitate was filtered off and the filtrate transferred to a separation funnel. The organic phase was washed again with $NaHCO_3$ saturated solution and brine, dried over $MgSO_4$ and concentrated in vacuum. The product was purified by flash chromatography eluting with cHex:EtOAc (5:1 to 2:1) to give a colorless solid (742mg, 1.8mmol, 34% yield).

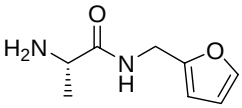
2) Fmoc deprotection: The Fmoc-peptide (251mg, 0.62mmol) was treated with 20% piperidine in DCM (18mL) at room temperature for 30 minutes. Piperidine was removed by coevaporation with methanol and the product purified by flash chromatography (DCM to DCM:MeOH (8:2))

Yield: 103mg light yellow oil (0.57mmol, 92% last step, 31% over the two steps)

$C_{10}H_{16}N_2O$ (180.25)

^{1}H-NMR ($DMSO\text{-}d_6$, 400MHz): δ = 1.17/1.18 (3H, *J* = 7.0 Hz, NH_2), 2.45-2.55 (2H, m, CH_2), 2.89-2.94 (2H, m, CH_2), 3.18-3.34 (2H, m, CH_2 *Cp*), 3.45 (1H, m, α-CH), 6.07-6.09, 6.20-6.21, 6.24-6.27, 6.38-6.43 and 6.45-6.47 (3H, m, *Cp*) and 8.14 (1H, t, *J* = 7.0 Hz, CONH) ppm.

Alanine-furfurylamide (42)



1) Coupling: Fmoc-Ala-OH (1.34g, 4.3mmol) and furfurylamine **39** (0.38mL, 4.3mmol) were dissolved in 30mL THF:DCM (2:1) and the solution cooled to 0°C. HOBt (0.74g, 5.2mmol)

and DIC (0.73mL, 4.7mmol) were added and the solution was allowed to warm up to room temperature and stirred overweekend. After that, the reaction mixture was concentrated in vacuum, and dissolved in EtOAc:cHex. The urea precipitate was filtered off and the filtrate transferred to a separation funnel. The organic phase was washed with 3x HCl 1M, 3x $NaHCO_3$ saturated solution and 1x brine, dried over $MgSO_4$ and concentrated in vacuum to give a colorless solid (1.5g, 3.8mmol, 88% yield).

2) Fmoc deprotection: The Fmoc-peptide (750mg, 1.9mmol) was treated with 20% piperidine in DCM (24mL) at room temperature for 30 minutes. Piperidine was removed by coevaporation with methanol and the product purified by flash chromatography (DCM to DCM:MeOH (8:2))

Yield: 175mg light yellow oil (1.0mmol, 55% last step, 51% over the two steps)

$C_8H_{12}N_2O_2$ (168.19)

^{1}H-NMR ($CDCl_3$, 400MHz): δ = 1.33 (3H, d, *J* = 7.0 Hz, CH_3 *Ala*), 3.50 (1H, m, α-CH), 4.36-4.42 (2H, m, CH_2), 6.20 (1H, dd, *J* = 3.2, 0.7 Hz, *Fur*), 6.30 (1H, dd, *J* = 1.9, 3.2 Hz, *Fur*), 7.33 (1H, dd, *J* = 1.8, 0.8 Hz, *Fur*) and 7.59 (1H, br s, CONH) ppm.

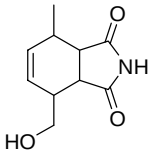
^{13}C-NMR ($CDCl_3$, 100MHz): δ = 21.9 (CH_3), 36.3 (CH_2), 50.9 (α-CH), 107.4 (*Fur*), 110.6 (*Fur*), 142.3 (*Fur*), 151.7 (*Fur*) and 175.7 (CO) ppm.

6.2.6. Diels-Alder ligation of peptides

General procedure for the Diels-Alder ligation

Diene- and dienophile-peptides were dissolved in water to an approximate concentration of 10mM. In some cases, methanol or DMF were added in minimum amounts to help peptide solubilization. After appropriate reaction time, the ligation product was directly purified by reversed phase HPLC and finally lyophilized.

HO-cyclo-NH (3)



Trans,*trans*-2,4-hexadien-1-ol (92mg, 0.94mmol) and maleimide (92mg, 0.94mmol) were dissolved in 2mL MeOH:H_2O (13:7) and stirred for 5h. The formed solid was separated by filtration, washed with water and diethylether and finally dried under reduced pressure.

Yield: 0.18g colorless solid (0.92mmol, 99% determined by NMR)

$C_{10}H_{13}NO_3$ (195.22)

^{1}H-NMR (CD_3OD, 400MHz): δ = 1.39 (3H, d, *J* = 7.4 Hz, CH_3), 2.40-2.49 (2H, m, H_6 + H_3), 3.10 (1H, dd, *J* = 8.5, 7.1 Hz, H_5), 3.30 (1H, m, H_4), 3.91 (1H, dd, *J* = 11.0, 7.9 Hz, CH_2), 4.03 (1H, dd, *J* = 11.0, 6.9 Hz, CH_2) and 5.74-5.86 (2H, m, CH_1=CH_2) ppm.

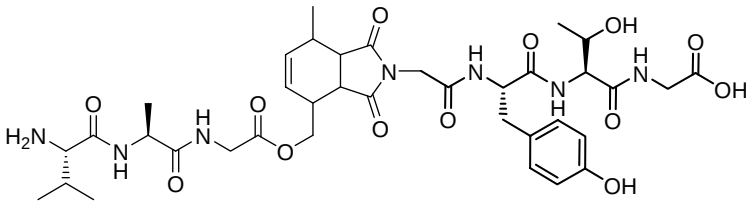
^{13}C-NMR (CD_3OD, 100MHz): δ = 16.0 (CH_3), 31.0 (C_6), 38.9 (C_3), 44.3 (C_4), 46.5 (C_5), 62.3 (CH_2), 130.1 (C_1), 134.6 (C_2), 179.7 (C=O) and 180.1 (C=O) ppm.

GC-MS (Method B)**:** t_r = 6.84; m/z: 165 $[M-30]^+$, t_r = 7.00; m/z: 195 $[M]^+$, 178 $[M-OH]^+$

ESI-MS: 196.3 $[M+H]^+$ (calc. 196.1)

FAB-HRMS: 196.0973 $[M+H]^+$ (calc. 196.0974).

Val-Ala-Gly-O-cyclo-N-Gly-Tyr-Thr-Gly-OH (18a)



Val-Ala-Gly-*trans*,*trans*-2,4-hexadienyl ester **9** (11mg, 0.033mmol) and *N*-maleoyl-Gly-Tyr-Thr-Gly-OH **5a** (19mg, 0.039mmol) were dissolved in 1mL H_2O:MeOH (10:3) and stirred at room temperature overnight. After 20h, complete consumption of the diene-peptide **9** was observed by HPLC. The Diels-Alder reaction product was purified by reversed-phase HPLC (gradient (time/solvent B) of 0min (5%B) → 5min (5%B) → 8min (10%B) → 15min (100%B), retention time: 11.5min). Fractions containing the desired product (evaluated by MALDI-TOF) were combined and lyophilized.

Yield: 23mg colorless solid (0.029mmol, 87% isolated)

$C_{37}H_{51}N_7O_{13}$ (801.84)

^{1}H-NMR (CD_3OD, 400MHz): δ = 1.03 (3H, d, *J* = 6.8 Hz, CH_3 *Val*), 1.06 (3H, d, *J* = 6.9 Hz, CH_3 *Val*), 1.14 (3H, d, *J* = 6.4 Hz, CH_3 *Thr*), 1.39 (3H, d, *J* = 7.2 Hz, CH_3 *Ala*), 1.40 (3H, d, *J* = 7.0 Hz, CH_3 *cyclo*), 2.14-2.24 (1H, m, β-CH *Val*), 2.44-2.52 (1H, m, H_6), 2.65-2.72 (1H, m, H_3), 2.88 (1H, dd, *J* = 13.7, 7.8 Hz, β-CH_2 *Tyr*), 3.02 (1H, dd, *J* = 13.9, 6.3 Hz, β-CH_2 *Tyr*), 3.12 (5%, m, H_5 *exo*), 3.18 (1H, dd, *J* = 8.4, 7.2 Hz , H_5 *endo*), 3.39 (1H, dd, *J* = 8.5, 6.1 Hz , H_4 *endo*), 3.47 (5%, m, H_4 *exo*), 3.67 (1H, d, *J* = 5.6 Hz, α-CH *Val*), 3.84-4.09 (6H, m, *3x* α-CH_2 *Gly*), 4.13-4.20 (1H, m, β-CH *Thr*), 4.31 (1H, t, *J* = 3.5 Hz, α-CH), 4.47 (1H, ddd, *J* = 14.3,

7.1, 3.2 Hz, CH_2 *cyclo*), 4.51-4.63 (3H, m, CH_2 *cyclo* + 2x α-CH), 5.71-5.79 (2H, m, CH=CH), 6.69 (2H, d, *J* = 7.4 Hz, Ar *Tyr*) and 7.04 (2H, d, *J* = 8.5 Hz, Ar *Tyr*) ppm.

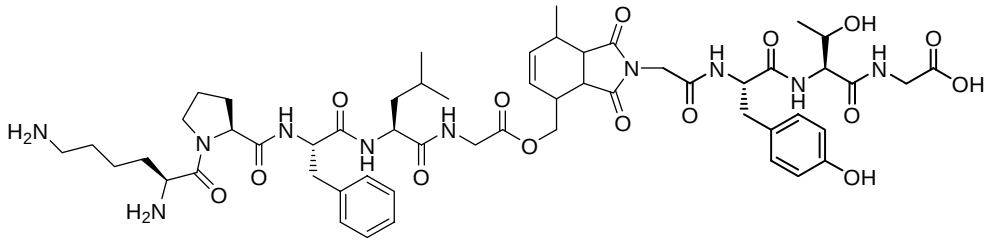
HPLC (Method A): t_r = 8.0 min

LC-MS (Method A): t_r = 16.0 min; m/z = 802.4 $[M+H]^+$ (calc. 802.4)

MALDI-TOF: 802.9 $[M+H]^+$ (calc. 802.4), 824.9 $[M+Na]^+$ (calc. 824.3), 840.9 $[M+K]^+$ (calc. 840.3).

$[\alpha]_D^{20}$: + 11.7 (c = 1.0, DMF)

Lys-Pro-Phe-Leu-Gly-O-cyclo-N-Gly-Tyr-Thr-Gly-OH (18b)



Hexadienyl peptide ester **15a** (10mg, 0.015mmol) and *N*-maleoyl-Gly-Tyr-Thr-Gly-OH **5a** (7mg, 0.015mmol) were dissolved in 1000µl H_2O:MeOH (4:1) and stirred at room temperature for 1 day. HPLC analysis revealed a ratio of 80:20 for the formation of the cycloadduct in comparison with the starting diene. The Diels-Alder reaction product was purified by reversed-phase HPLC (gradient (time/solvent B) of 0min (5%B) → 3min (5%B) → 5min (10%B) → 20min (100%B), retention time: 10.5min). Fractions containing desired product (evaluated by MALDI-TOF) were combined and lyophilized.

Yield: 10mg colorless solid (0.009mmol, 60% isolated)

$C_{55}H_{76}N_{10}O_{15}$ (1117.25)

^{1}H-NMR (CD_3OD, 400MHzδ = 0.90 (3H, d, *J* = 6.0 Hz, CH_3 *Leu*), 0.94 (3H, d, *J* = 6.0 Hz, CH_3 *Leu*), 1.14 (3H, d, *J* = 6.4 Hz, CH_3 *Thr*), 1.39/1.40 (ratio: 1/1, 3H, d, *J* = 7.3 Hz, CH_3 *cyclo*), 1.51-1.74 (7H, m, γ-CH_2 *Lys* + δ-CH_2 *Lys* + β-CH_2 *Leu* + γ-CH *Leu*), 1.87-2.09 (5H, γ-CH_2 *Pro* + β-CH_2 *Lys* + β-CH_2 *Pro*), 2.16-2.26 (1H, m, β-CH_2 *Pro*), 2.45-2.51 (1H, m, H_6), 2.64-2.73 (1H, m, H_3), 2.87 (1H, dd, *J* = 13.9, 7.8 Hz, β-CH_2 *Tyr*), 2.92-3.15 (5H, m, β-CH_2 *Tyr* + ε-CH_2 *Lys* + β-CH_2 *Phe*), 3.13/3.20 (ratio: 1/1, 1H, dd, *J* = 8.6, 7.2/8.2, 7.2 Hz, H_5), 3.38/3.46 (ratio: 1/1, 1H, dd, *J* = 8.4, 6.2/8.7, 6.3 Hz, H_4), 3.55-3.64 (1H, m, δ-CH_2 *Pro*) , 3.67-3.74 (1H, m, δ-CH_2 *Pro*), 3.83-4.11 (6H, m, *3x* α-CH_2 *Gly*), 4.13-4.19 (1H, m, β-CH *Thr*), 4.23-4.28 (1H, m, α-CH), 4.32 (1H, t, *J* = 3.7 Hz, α-CH), 4.39-4.46 (1H, m, CH_2 *cyclo*), 4.48-

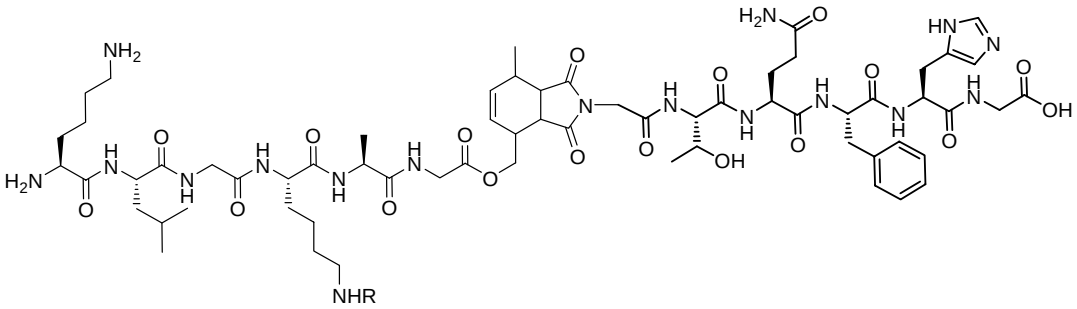
4.52 (1H, m, CH_2 cyclo), 4.54-4.65 (4H, m, *4x* α-CH), 5.70-5.78 (2H, m, CH=CH), 6.69 (2H, d, *J* = 8.4 Hz, Ar *Tyr*), 7.03 (2H, d, *J* = 8.5 Hz, Ar *Tyr*) and 7.18-7.31 (5H, m, Ar *Phe*) ppm.

ESI-MS: m/z 1117.6 $[M+H]^+$ (calc. 1117.6)

MALDI-TOF: m/z 1118.0 $[M+H]^+$ (calc. 1117.6), 1140.0 $[M+Na]^+$ (calc. 1139.5), 1156.0 $[M+K]^+$ (calc. 1155.5).

$[\alpha]_D^{20}$: - 23.2 (c = 0.3, MeOH)

Lys-Leu-Gly-Lys(Mtt)-Ala-Gly-O-cyclo-N-Gly-Thr-Gln-Phe-His-Gly-OH (18c)



Hexadienyl ester **15d** (2.0mg, 2.2μmol) and *N*-maleoyl-Gly-Thr-Gln-Phe-His-Gly-OH **5b** (1.6mg, 2.2μmol) were dissolved in 200μl H_2O:MeOH (10:1) and stirred at room temperature for 1 day. Approximately 95% of diene **15d** was consumed after 24h as shown by HPLC analysis. Partial removal of Mtt protection groups from lysine side chain took place during reaction, giving two cycloadducts **18c** (with Mtt) and **18c´**(without Mtt). These compounds were isolated by reversed-phase HPLC (gradient (time/solvent B) of 0min (5%B) → 3min (5%B) → 15min (60%B) → 17min (100%B), retention time: 17.0min (**18c**) and 13.0min (**18c´**)) and dried by lyophilization.

Overall yield: 74% cycloadduct isolated.

◈ **18c:** 1.3mg colorless solid (0.80μmol, 43%), compound with Mtt group

$C_{83}H_{111}N_{17}O_{18}$ (1634.87)

^{1}H-NMR (CD_3OD, 400MHz): δ = 0.94 (3H, d, *J* = 6.7 Hz, CH_3 *Leu*), 0.96 (3H, d, *J* = 6.7 Hz, CH_3 *Leu*), 1.20 (3H, d, *J* = 6.2 Hz, CH_3 *Thr*), 1.39 (6H, d, *J* = 7.1 Hz, CH_3 *Ala* + CH_3 *cyclo*), 1.3-2.3 (m, CH_2 *Lys* + CH_2 *Gln* + CH_2,CH *Leu*), 2.37 (3H, s, CH_3 *Mtt*), 2.44-2.51 (1H, m, H_6), 2.63-2.72 (1H, m, H_3), 2.9-3.4 (m, β-CH_2 *Phe* + β-CH_2 *His* + β-CH_2 *Lys* + H_5 + H_4), 3.7-4.7 (5H, m, α-CH_2 *Gly* + α-CH + CH_2 *cyclo*), 5.72-5.76 (2H, m, CH=CH), 7.17-7.49 (20H, m, Ar *Phe* + Ar *Mtt* + CH *His*) and 8.73 (1H, br s, NH *His*) ppm.

HPLC (Method A): t_r = 16.9 min

MALDI-TOF: m/z 1379.1 $[(M\text{-}Mtt)+H]^+$ (calc. 1378.7), 1417.1 $[(M\text{-}Mtt)+K]^+$ (calc. 1416.8), Mtt peak.

◈ **18c´**: 0.8mg of colorless solid (0.58μmol, 31%), compound without Mtt group

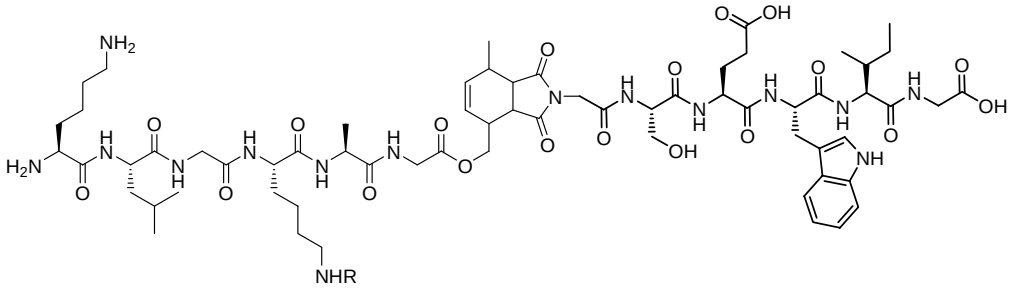
$C_{63}H_{95}N_{17}O_{18}$ (1378.53)

^{1}H-NMR (CD_3OD, 400MHz): δ = 0.95 (3H, d, *J* = 6.6 Hz, CH_3 *Leu*), 0.98 (3H, d, *J* = 6.6 Hz, CH_3 *Leu*), 1.20 (3H, d, *J* = 6.4 Hz, CH_3 *Thr*), 1.40 (6H, d, *J* = 7.5 Hz, CH_3 Ala + CH_3 *cyclo*), 1.4-2.3 (m, CH_2 *Lys* + CH_2 *Gln* + CH_2,CH *Leu*), 2.47-2.53 (1H, m, H_6), 2.67-2.74 (1H, m, H_3), 2.9-3.4 (m, β-CH_2 *Phe* + β-CH_2 *His* + β-CH_2 *Lys* + H_5 + H_4), 3.7-4.7 (m, α-CH_2 *Gly* + α-CH + CH_2 *cyclo*), 5.73-5.77 (2H, m, CH=CH), 7.35 (1H, br s, CH *His*), 7.18-7.29 (5H, m, Ar *Phe*) and 8.71 (1H, br s, NH *His*) ppm.

HPLC (Method A): t_r = 13.0 min.

MALDI-TOF: m/z 1379.5 $[M+H]^+$ (calc. 1378.7), 1401.4 $[M+Na]^+$ (calc. 1400.7), 1417.4 $[M+K]^+$ (calc. 1416.8).

Lys-Leu-Gly-Lys(Mtt)-Ala-Gly-O-cyclo-N-Gly-Ser-Glu-Trp-Ile-Gly-OH (18d)



Hexadienyl ester **15d** (2.0mg, 2.2μmol) and *N*-maleoyl-Gly-Ser-Glu-Trp-Ile-Gly-OH **5c** (1.6mg, 2.2μmol) were dissolved in 200μl H_2O:DMF (10:1) and stirred at room temperature for 2 days. Approximately 84% of diene **15d** was consumed after 47h as shown by HPLC analysis. Partial removal of Mtt protection groups from lysine side chain took place during the reaction, giving two cycloadducts **18d** (with Mtt) and **18d´** (without Mtt). These compounds were isolated by reversed-phase HPLC (gradient (time/solvent B) of 0min (5%B) → 3min (5%B) → 15min (60%B) → 17min (100%B), retention time: 17.1min (**18d**) and 16.1 (**18d´**)) and dried by lyophilization.

Overall yield: 64% cycloadduct isolated.

◈ **18d** : 1.2mg colorless solid (0.73μmol, 40%), compound with Mtt group

$C_{84}H_{113}N_{15}O_{19}$ (1636.89)

^{1}H-NMR (CD_3OD, 400MHz): found δ = 1.40 (3H, d, CH_3 *Hxd*), 2.38 (3H, s, *Mtt*), 5.67-5.75 (2H, m, CH=CH) 7.1-7.5 (*signals for Mtt group*) ppm and expected peaks for all amino acids.

HPLC (Method A): t_r = 18.4 min.

MALDI-TOF: 1381.5 $[(M\text{-}Mtt)+H]^+$ (calc. 1380.7), 1403.4 $[(M\text{-}Mtt)+Na]^+$ (calc. 1402.7), 1419.4 $[(M\text{-}Mtt)+K]^+$ (calc. 1418.7), Mtt peak.

◈ **18d´ :** 0.6mg of colorless solid (0.43μmol, 24%), compound without Mtt group

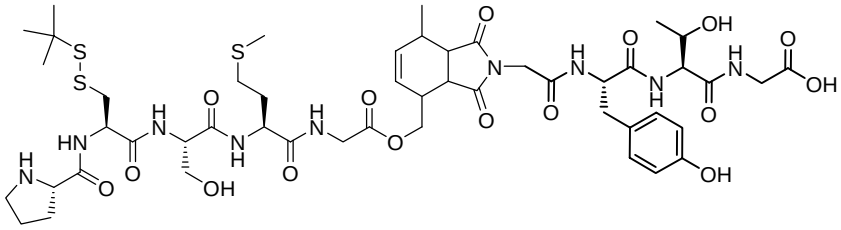
$C_{64}H_{97}N_{15}O_{19}$ (1380.54)

^{1}H-NMR (DMSO-d_6 400MHz): found δ = 1.2 (CH_3 *Hxd*) and 5.02-5.75 (2H, m, CH=CH) ppm and expected peaks for all amino acids.

HPLC (Method A): t_r = 15.0 min

MALDI-TOF: m/z 1381.3 $[M+H]^+$ (calc. 1380.7), 1403.2 $[M+Na]^+$ (calc. 1402.7), 1419.2 $[M+K]^+$ (calc. 1418.7).

Pro-Cys(StBu)-Ser-Met-Gly-O-cyclo-N-Gly-Tyr-Thr-Gly-OH (18e)



Hexadienyl peptide ester **15b** (2.0mg, 3.0μmol) and *N*-maleoyl-Gly-Tyr-Thr-Gly-OH **5a** (3.5mg, 7.3μmol) were combined in 500μl H_2O:MeOH (3:2) and stirred at room temperature for 24 hours. After this time, complete consumption of the diene-peptide **15b** was observed by HPLC. The cycloadduct was purified by reversed-phase HPLC (gradient (time/solvent B) of 0min (10%B) → 1min (30%B) → 4min (30%B) → 10min (50%B), retention time: 7.2min). Fractions containing desired product (evaluated by MALDI-TOF) were combined and lyophilized.

Yield: 1.1mg colorless solid (0.97μmol, 32% isolated)

$C_{49}H_{71}N_9O_{16}S_3$ (1138.33)

^{1}H-NMR (CD_3OD, 400MHz): δ = 1.14 (3H, d, *J* = 6.5 Hz, CH_3 *Thr*), 1.35 (9H, br s, tBu *Cys*), 1.40 (3H, d, *J* = 7.0 Hz, CH_3 *cyclo*), 1.89-2.21 (6H, m, γ-CH_2 *Pro* + β-CH_2 *Pro* + β-CH_2 *Met*), 2.08 (3H, s, CH_3 *Met*), 2.42-2.65 (3H, m, H_6 + γ-CH_2 *Met*), 2.66-2.72 (1H, m, H_3), 2.85-2.92 (1H, β-CH_2 *Tyr*), 2.97-3.05 (3H, m, β-CH_2 *Tyr* + β-CH_2 *Cys*), 3.16-3.22 (1H, m, H_5), 3.37-3.45 (1H, m, H_4), 3.57-4.19 (11H, m, δ-CH_2 *Pro* + β-CH_2 *Ser* + *3x* α-CH_2 *Gly* + β-CH *Thr*), 4.29-

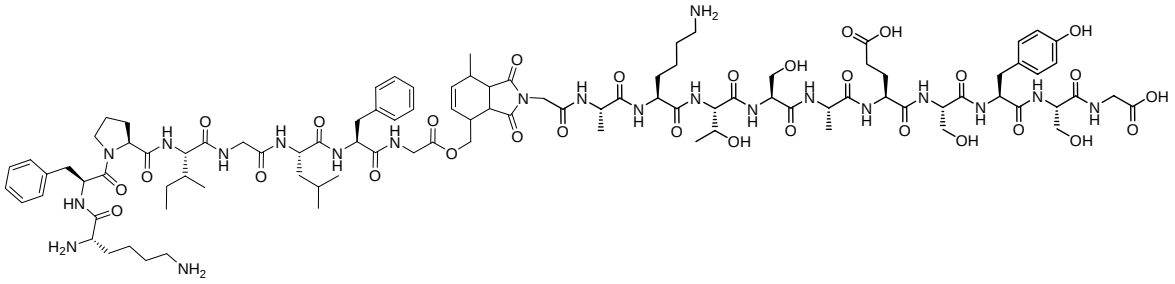
4.45 (3H, m, *3x* α-CH), 4.53-4.66 (4H, m, *2x* α-CH + CH_2 *cyclo*), 4.69-4.73 (1H, m, α-CH *Cys*), 5.68-5.80 (2H, m, CH=CH), 6.68 (2H, d, *J* = 8.4 Hz, Ar *Tyr*) and 7.03 (2H, d, *J* = 8.3 Hz, Ar *Tyr*) ppm.

ESI-MS: m/z 1138.4 $[M+H]^+$ (calc. 1138.4)

MALDI-TOF: m/z 1139.1 $[M+H]^+$ (calc. 1138.4), 1161.0 $[M+Na]^+$ (calc. 1160.4), 1177.0 $[M+K]^+$ (calc. 1176.4).

$[\alpha]_D^{20}$: - 31.5 (c = 0.1, MeOH)

Lys-Phe-Pro-Ile-Gly-Leu-Phe-Gly-O-cyclo-N-Gly-Ala-Lys-Thr-Ser-Ala-Glu-Ser-Tyr-Ser-Gly-OH (18f)



Hexadienyl peptide ester **15f** (2.8mg, 2.9μmol) and maleimide-undecapeptide **5d** (3.3mg, 2.9μmol) were dissolved in 300μl H_2O and the solution was stirred at room temperature for 2 days. 93% ligation production was formed after 48h reaction (in comparison with diene **15f** consumption) as indicated by HPLC. The cycloadduct was purified by reversed-phase HPLC (gradient (time/solvent B) of 0min (5%B) → 15min (50%B), retention time: 11.6min). Fractions containing desired product were combined and lyophilized.

Yield: 3.5mg colorless solid (1.7μmol, 69% isolated)

$C_{98}H_{143}N_{21}O_{30}$ (2095.31)

^{1}H-NMR (CD_3OD, 400MHz): δ = 0.81-1.00 (12H, m, *2x* CH_3 *Ile* + *2x* CH_3 *Leu*), 1.19 (3H, d, *J* = 6.3 Hz, CH_3 *Thr*), 1.19-1.77 (22H, m, *2x* β-/γ-/δ-CH_2 *Lys* + *2x* CH_3 *Ala* + CH_2 *Ile* + β-CH_2 *Leu*), 1.43 (3H, d, *J* = 7.3 Hz, CH_3 *cyclo*), 1.81-2.20 (8H, m, β-CH_2 *Glu* + β-CH_2 *Pro* + γ-CH_2 *Pro* + γ-CH *Leu* + β-CH *Ile*), 2.41-2.46 (2H, m, γ-CH_2 *Glu*), 2.47-2.53 (1H, m, H_6), 2.66-2.74 (1H, m, H_3), 2.89-3.26 (11H, m, *2x* ε-CH_2 *Lys* + β-CH_2 *Tyr* + 2x β-CH_2 *Phe* + H_5), 3.34-3.57 (3H, m, H_4 + ε-CH_2 *Pro*), 3.71-4.00 (14H, m, *3x* β-CH_2 *Ser* + *4x* α-CH_2 *Gly*), 4.06-4.68 (15H, m, 13x α-CH + CH_2 *cyclo*), 5.71-5.77 (2H, m, CH=CH), 6.68 (2H, d, *J* = 8.4 Hz, Ar *Tyr*), 7.08 (2H, d, *J* = 8.5 Hz, Ar *Tyr*) and 7.17-7.34 (10H, m, Ar *Phe*) ppm.

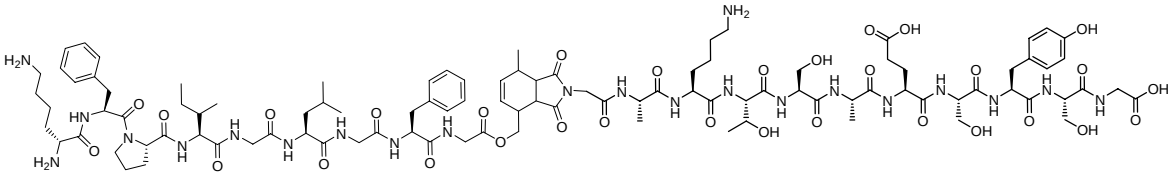
HPLC (Method A): t_r = 16.1 min.

LC-MS (Method A): t_r = 14.1 min; m/z 1048 $[M+H]^{++}$ (calc. 1048).

MALDI-TOF: m/z 2098 $[M+H]^+$ (calc. 2096), 2120 $[M+Na]^+$ (calc. 2118), 2135 $[M+K]^+$ (calc. 2134).

$[\alpha]_D^{20}$**:** - 22.9 (c = 0.1, MeOH)

Lys-Phe-Pro-Ile-Gly-Leu-Gly-Phe-Gly-O-cyclo-N-Gly-Ala-Lys-Thr-Ser-Ala-Glu-Ser-Tyr-Ser-Gly-OH (18g)



Hexadienyl peptide ester **15g** (2.8mg, 2.7μmol) and maleimide-undecapeptide **5d** (3.1mg, 2.7μmol) were dissolved in 270μl H_2O and the solution was stirred at room temperature for 2 days. 92% ligation production was formed after 48h reaction (in comparison with diene **15g** consumption) as indicated by HPLC. The cycloadduct was purified by reversed-phase HPLC (gradient (time/solvent B) of 0min (5%B) → 16min (50%B), retention time: 11.5min). Fractions containing desired product were combined and lyophilized.

Yield: 3.2mg colorless solid (1.5μmol, 67% isolated)

$C_{100}H_{146}N_{22}O_{31}$ (2152.36)

^{1}H-NMR (CD_3OD, 400MHz): δ = 0.82-1.00 (12H, m, *2x* CH_3 *Ile* + *2x* CH_3 *Leu*), 1.19 (3H, d, *J* = 6.3 Hz, CH_3 *Thr*), 1.20-1.74 (22H, m, *2x* β-/γ-/δ-CH_2 *Lys* + *2x* CH_3 *Ala* + CH_2 *Ile* + β-CH_2 *Leu*), 1.43 (3H, d, *J* = 7.3 Hz, CH_3 *cyclo*), 1.83-2.19 (8H, m, β-CH_2 *Glu* + β-/γ-CH_2 *Pro* + γ-CH *Leu* + β-CH *Ile*), 2.41-2.47 (2H, m, γ-CH_2 *Glu*), 2.47-2.53 (1H, m, H_6), 2.66-2.74 (1H, m, H_3), 2.89-3.26 (11H, m, *2x* ε-CH_2 *Lys* + β-CH_2 *Tyr* + *2x* β-CH_2 *Phe* + H_5), 3.32-3.57 (3H, m, H_4 + ε-CH_2 *Pro*), 3.70-4.02 (16H, m, *3x* β-CH_2 *Ser* + *5x* α-CH_2 *Gly*), 4.06-4.68 (16H, m, *14x* α-CH + CH_2 *cyclo*), 5.73-5.77 (2H, m, CH=CH), 6.68 (2H, d, *J* = 8.6 Hz, Ar *Tyr*), 7.06-7.09 (2H, m, Ar *Tyr*) and 7.17-7.33 (10H, m, Ar *Phe*) ppm.

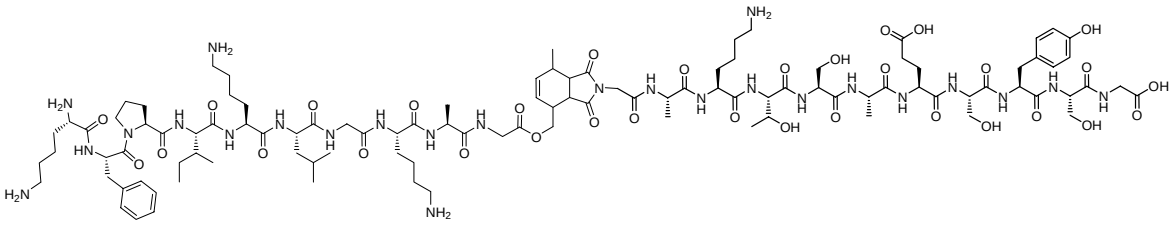
HPLC (Method A): t_r = 16.0 min

LC-MS (Method A): t_r = 14.0 min; m/z 1076 $[M+H]^{++}$ (calc. 1076)

MALDI-TOF: m/z 2155 $[M+H]^+$ (calc. 2153), 2177 $[M+Na]^+$ (calc. 2175), 2193 $[M+K]^+$ (calc. 2191).

$[\alpha]_D^{20}$**:** - 22.2 (c = 0.1, MeOH)

Lys-Phe-Pro-Ile-Lys-Leu-Gly-Lys-Ala-Gly-O-cyclo-N-Gly-Ala-Lys-Thr-Ser-Ala-Glu-Ser-Tyr-Ser-Gly-OH (18h)



An aliquot of 90µl of 5mM solution of diene decapeptide **17** (0.44µmol) were combined with 90µl of 5mM solution of maleimide undecapeptide **5d** (0.44µmol) and the reaction solution was mixed at room temperature. Past 48h, production of the Diels-Alder ligation peptide, in comparison with diene consumption, was of 87% determined by HPLC. The cycloadduct was isolated by HPLC (gradient (time/solvent B) of 0min (10%B) → 3min (10%B) → 15min (60%B), retention time: 9.2min), dried by lyophilization and analysed.

Yield: approximately 0.5mg colorless solid (0.2µmol, ca. 50%)

$C_{104}H_{163}N_{25}O_{32}$ (2275.56)

^{1}H-NMR (CD_3OD, 400MHz): found δ = 1.4 (CH_3 *cyclo*), 5.7 (C*H*=C*H cyclo*) ppm.

HPLC (Method A): t_r = 13.7 min

MALDI-TOF: m/z 2277 $[M+H]^+$ (calc. 2276), 2315 $[M+K]^+$ (calc. 2314).

Lys-Cys-Gly-Val-Phe-Gly-O-cyclo-N-Gly-Thr-Gln-Phe-His-Gly-OH (19)

An aliquot of 87µL of 10mM solution of diene hexapeptide **16** (0.87µmol) were added to 87µL of 10mM solution of maleimido-hexapeptide **5b** (0.87µmol) in H_2O:MeOH (4:1) and the reaction solution was agitated at room temperature under argon for 2 days. HPLC analysis of the reaction mixture after 48h showed that 93% of starting dienophile was consumed and two new products were formed in 73:27 ratio. Reserved-phase preparative HPLC was used to isolated the products (gradient (time/solvent B) of 0min (5%B) → 15min (50%B), retention time: 10.0min (major product) and 11.6min (minor product)), which were identified by mass spectroscopy.

◈ **Major product:** resulted from double ligation: nucleophilic adition and Diels-Alder reaction: KC(mal-GTDFHG)GVFG-O-cyclo-N-GTDFHG-OH (68% determined by HPLC).

$C_{97}H_{129}N_{25}O_{29}S$ (2141.28)

HPLC (Method A): t_r = 14.9 min.

MALDI-TOF: m/z 2143 $[M+H]^+$ (calc. 2142), 2180 $[M+K]^+$ (calc. 2180).

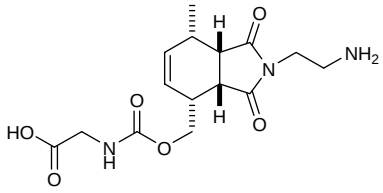
◈ **Minor product:** resulted from only one ligation (25% determined by HPLC).

$C_{65}H_{90}N_{16}O_{18}S$ (1415.57)

HPLC (Method A): t_r = 16.3 min.

MALDI-TOF: m/z 1416.5 $[M+H]^+$ (calc. 1415.6), 1454.4 $[M+K]^+$ (calc. 1453.6).

HO-Gly-HNC(O)-O-cyclo-*N*-ethylamine (20)



Glycine *tran,trans*-2,4-hexadienyl carbamate **24** (10mg, 0.051mmol) and *N*-maleoylthylamine **54** (13mg, 0.051mmol) were dissolved in 500µl H_2O:MeOH (4:1) and stirred at room temperature for 5h. The reaction mixture was concentrated using lyophilization, redissolved in deuterated methanol and analyzed by NMR. The spectrum indicated 91% Diels-Alder product formation. The cycloadduct was purified by reversed-phase HPLC (gradient (time/solvent B) of 0min (5%B) → 3min (5%B) → 16min (50%B), retention time: 8.6min). Fractions containing desired product (evaluated by MALDI-TOF) were combined and lyophilized.

Yield: 11mg colorless solid (0.032mmol, 74% isolated)

$C_{15}H_{21}N_3O_6$ (339.34)

^{1}H-NMR (CD_3OD, 400MHz): δ = 1.41 (3H, d, *J* = 7.4 Hz, CH_3), 2.47-2.56 (2H, m, H_6), 2.65-2.72 (1H, m, H_3), 3.06 (2H, t, *J* = 6.1 Hz, CH_2), 3.19 (1H, dd, *J* = 8.4, 7.0 4 Hz, H_5), 3.38 (1H, dd, *J* = 8.5, 6.2 Hz, H_4), 3.70 (2H, t, *J* = 5.8 Hz, CH_2), 3.82 (2H, br s, α-CH_2 *Gly*), 4.49 (1H, dd, *J* = 10.9, 8.6 Hz, CH_2 *cyclo*), 4.57 (1H, dd, *J* = 10.9, 6.6 Hz, CH_2 *cyclo*) and 5.72-5.83 (2H, m, CH=CH) ppm.

^{13}C-NMR (CD_3OD, 100MHz): δ = 15.9 (CH_3), 30.9 (C_6), 35.6 (CH_2), 36.0 (C_3), 37.7 (CH_2), 41.9 (CH_2), 42.8 (C_4), 45.2 (C_5), 62.2 (CH_2), 129.5 (C_1), 135.0 (C_2), 157.9 (HNCOO), 172.4 (COOH), 177.9 (C=O) and 178.1 (C=O) ppm.

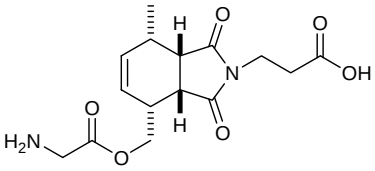
HPLC (Method A): t_r = 11.5 min.

ESI-MS: m/z 340.5 $[M+H]^+$ (calc. 340.1).

MALDI-TOF: m/z 340.5 $[M+H]^+$ (calc. 340.1), 362.5 $[M+Na]^+$ (calc. 362.1), 378.4 $[M+K]^+$ (calc. 378.1).

FAB-HRMS: m/z 340.1538 $[M+H]^+$ (calc. 340.1509).

Gly-O-cyclo-N-βAla (21)



Glycine *trans*,*trans*-2,4-hexadienyl ester **26** (6mg, 0.038mmol) and *N*-maleoyl-β-alanine (7mg, 0.038mmol) were dissolved in 400µl H_2O:MeOH (9:1) and stirred at room temperature for 4h. The reaction mixture was concentrated using lyophilization, redissolved in deuterated methanol and analyzed by NMR. The spectrum indicated complete consumption of diene **26**. The cycloadduct was purified by reversed-phase HPLC (gradient (time/solvent B) of 0min (5%B) → 3min (5%B) → 16min (30%B), retention time: 10.5min). Fractions containing desired product (evaluated by MALDI-TOF) were combined and freeze-dried.

Yield: 5mg colorless solid (0.015mmol, 44% isolated)

$C_{15}H_{20}N_2O_6$ (324.33)

^{1}H-NMR (CD_3OD, 400MHz): δ = 1.42 (3H, d, *J* = 7.4 Hz, CH_3), 2.44-2.51 (3H, m, H_6 + α-CH_2 *βAla*), 2.69-2.74 (1H, m, H_3), 3.11 (1H, dd, *J* = 8.4, 6.9 Hz, H_5), 3.32 (1H, dd, *J* = 8.5, 6.2 Hz, H_4), 3.64 (2H, t, *J* = 7.3 Hz, β-CH_2 *βAla*), 3.87 (2H, s, α-CH_2 *Gly*), 4.67 (1H, dd, *J* = 11.0, 8.1 Hz, CH_2 *cyclo*), 4.74 (1H, dd, *J* = 11.0, 7.2 Hz, CH_2 *cyclo*) and 5.73-5.78 (2H, m, CH=CH) ppm.

^{13}C-NMR (D_2O, 100MHz): δ = 16.1 (CH_3), 30.8 (C_6), 32.2 (CH_2), 34.7 (CH_2), 34.9 (C_3), 40.3 (CH_2), 42.7 (C_4), 45.1 (C_5), 66.6 (CH_2), 127.9 (C_1), 135.8 (C_2), 168.3 (C=O), 175.4 (COOH), 180.1 (C=O) and 180.3 (C=O) ppm.

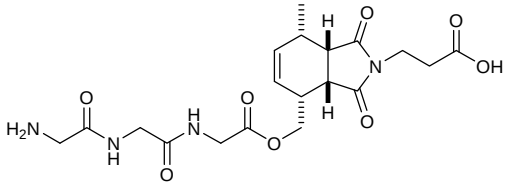
HPLC (Method A): t_r = 11.3 min.

ESI-MS: m/z 325.5 $[M+H]^+$ (calc. 325.1).

MALDI-TOF: m/z 325.5 $[M+H]^+$ (calc. 325.1), 347.5 $[M+Na]^+$ (calc. 347.1), 363.5 $[M+K]^+$ (calc. 363.1).

FAB-HRMS: m/z 325.1387 $[M]^+$ (calc. 325.1400).

Gly-Gly-Gly-O-cyclo-N-$^{\beta}$Ala (22)



Gly-Gly-Gly-*trans*-*trans*-2,4-hexadienyl ester **27** (7.6mg, 0.028mmol) and *N*-maleoyl-β-alanine (5.0mg, 0.028mmol) were dissolved in 300µl H_2O:MeOH (5:1) and stirred at room

temperature overnight. After 22h, the reaction was monitored by HPLC analysis, which indicated consumption of most starting diene-peptide (ratio: 91:9 (cycloadduct:diene). The Diels-Alder reaction product was purified by reversed-phase HPLC (gradient (time/solvent B) of 0min (5%B) → 3min (5%B) → 16min (30%B), retention time: 11.7min). Fractions containing desired product (evaluated by MALDI-TOF) were combined and lyophilized.

Yield: 3.7mg colorless solid (0.0084mmol, 32% isolated)

$C_{19}H_{26}N_4O_8$ (438.43)

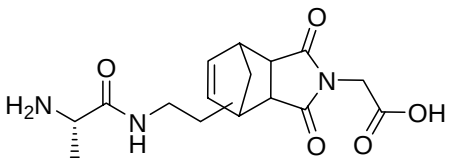
^{1}H-NMR (CD_3OD, 400MHz): δ = 1.40 (3H, d, *J* = 7.4 Hz, CH_3), 2.43-2.51 (3H, m, H_6 + α-CH_2 *βAla*), 2.65-2.70 (1H, m, H_3), 3.11 (1H, dd, *J* = 8.4, 7.1 Hz, H_5), ~3.30 (H_4, hidden by solvent peak), 3.64 (2H, t, *J* = 7.1 Hz, β-CH_2 *βAla*), 3.74 (2H, s, α-CH_2 *Gly*), 4.00 (4H, br s, *2x* α-CH_2 *Gly*), 4.56 (1H, dd, *J* = 10.9, 8.2 Hz, CH_2 *cyclo*), 4.64 (1H, dd, *J* = 10.9, 7.3 Hz, CH_2 *cyclo*) and 5.71-5.76 (2H, m, CH=CH) ppm.

HPLC (Method A): t_r = 11.9 min.

ESI-MS: m/z 439.5 $[M+H]^+$ (calc. 439.2).

MALDI-TOF: m/z 439.5 $[M+H]^+$ (calc. 439.2), 461.5 $[M+Na]^+$ (calc. 461.2), 477.4 $[M+K]^+$ (calc. 477.1).

Cycloadduct (41)



Cyclopentadienyl peptide **40** (5.9mg, 0.03mmol) and maleimide **4** (5.1mg, 0.03mmol) were dissolved in 0.5mL D_2O and stirred for 2h. The reaction mixture was directly analyzed by NMR.

Yield: complete consumption of the diene determined by NMR

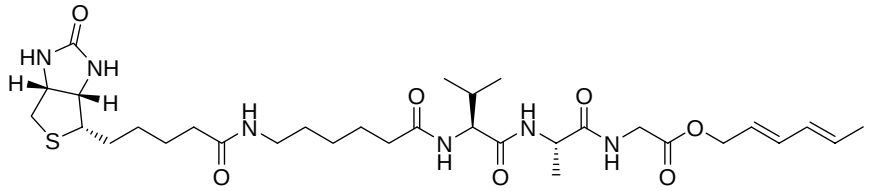
$C_{16}H_{21}N_3O_5$ (335.36)

^{1}H-NMR (D_2O, 400MHz): δ = signals for the new olefin group δ = 5.61 (br s), 5.85 (d) and 5.98-6.10 (m) ppm. No diene signals.

ESI-MS: 336.2 $[M+H]^+$ (calc. 336.2)

6.2.7. Synthesis of the biotinylated compounds

Biotin-ACA-Val-Ala-Gly-hexadienylether (44)



N-(+)-Biotinyl-6-aminocaproic acid (48mg, 0.13mmol) was dissolved in hot dry DMF (3mL), cooled down and added to a solution containing Val-Ala-Gly-hexadienylether **6** (0.10mmol in 1mL DMF). Afterwards, triethylamine (28µl, 0.20mmol), HOBt (31mg, 0.20mmol) and EDC (26mg, 0.13mmol) were added and the reaction mixture stirred for 20 hours at room temperature. Product was directly purified by reversed-phase HPLC (gradient (time/solvent B) of 0min (10%B) → 10min (50%B), retention time: 9.5min). Fractions containing the product were combined and lyophilized.

Yield: 28.5mg (0.043mmol, 43%) colorless solid

$C_{32}H_{52}N_6O_7S$ (664.86)

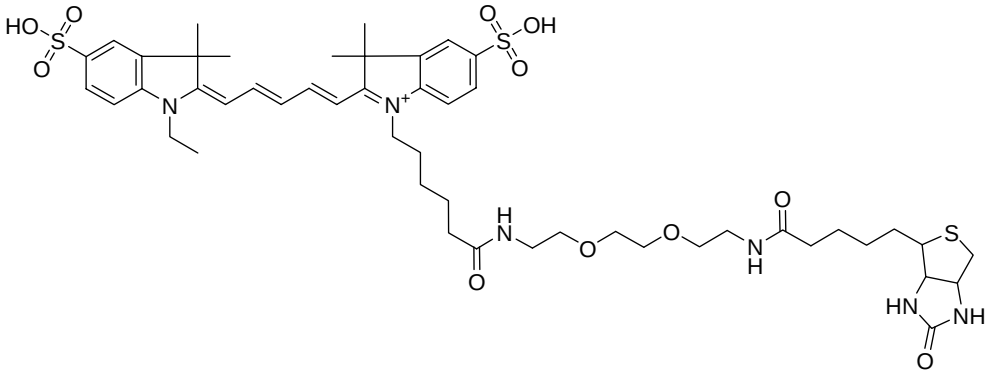
^{1}H-NMR (DMSO-d_6, 400 MHz): δ = 0.79 (3H, d, *J* = 6.8 Hz, CH_3 *Val*) 0.82 (3H, d, *J* = 6.8 Hz, CH_3 *Val*), 1.20 (3H, d, *J* = 7.1 Hz, CH_3 *Ala*), 1.21-1.64 (12H, m, *3x* CH_2 *ACA* + *3x* CH_2 *biot*), 1.72 (2H, d, *J* = 6.6 Hz, CH_3 *Hxd*), 1.90-1.97 (1H, m, β-CH *Val*), 2.02 (2H, t, *J* = 7.4 Hz, CH_2), 2.05-2.19 (2H, m, CH_2), 2.54 (1H, d, *J* = 12.4 Hz, CH_2S *biot*), 2.79 (1H, dd, *J* = 12.4, 5.0 Hz, CH_2S *biot*), 2.94-2.99 (2H, m, CH_2), 3.05-3.09 (1H, m, SCH *biot*), 3.72-3.91 (2H, m, α-CH_2 *Gly*), 4.09-4.16 (2H, m, *2x* α-CH), 4.26-4.31 (2H, m, *2x* CH *biot*), 4.54 (2H, d, *J* = 6.4 Hz, CH_2 *Hxd*), 5.59 (1H, ddd, *J* = 13.0, 6.4, 6.4 Hz, CH=C*H*CH_3), 5.69-5.78 (1H, m, CH_2C*H*=CH), 6.05 (1H, dd, *J* = 15.0, 10.6 Hz, C*H*=$CHCH_3$), 6.25 (1H, dd, *J* = 15.2, 10.6 Hz, CH_2CH=C*H*), 6.34 (1H, s, NH *urea*), 6.40 (1H, s, NH *urea*), 7.70 (1H, t, *J* = 5.5 Hz, CONH), 7.76 (1H, d, *J* = 8.8 Hz, CONH), 7.95 (1H, d, *J* = 7.4 Hz, CONH) and 8.25 (1H, t, *J* = 5.8 Hz, CONH) ppm.

MALDI-TOF: m/z 665.9 $[M+H]^+$ (calc. 665.4), 688.0 $[M+Na]^+$ (calc. 687.4), 704.0 $[M+K]^+$ (calc. 703.3).

FAB-HRMS: m/z 664.3636 $[M]^+$ (calc. 664.3618).

$[\alpha]_D^{20}$: - 16.8 (c = 0.7, MeOH)

Biotin-Cy5 (71)



Biotinylated amine **72** (13.5mg, 0.036mmol, provided by Maja Köhn, MPI Dortmund) and Cy5 (23.5mg, 0.036mmol) were dissolved in 4mL dry DMF. Then triethylamine (10µl, 0.072mmol), HOBt (11mg, 0.072mmol) and EDC (15.5mg, 0.054mmol) were added consecutively and the deep-blue reaction mixture was stirred for 21 hours at room temperature. Product was directly purified by reversed-phase HPLC (gradient (time/solvent B) of 0min (10%B) → 10min (20%B) → 15min (50%B), retention time: 12.5min). Fractions containing the product were combined and lyophilized.

Yield: 21mg intense blue solid (0.021, 58%)

$C_{49}H_{68}N_6O_{11}S_3$ (1013.29)

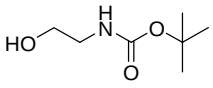
^{1}H-NMR (CD_3OD, 400 MHz): δ = 1.28-1.49 (9H, m, CH_3 *Cy5* + *2x* CH_2 *Cy5* + CH_2 *biot*), 1.50-1.88 (20H, m, *4x* CH_3 *Cy5* + *2x* CH_2 *Cy5* + 2x CH_2 *biot*), 2.16-2.23 (4H, m, CH_2 *Cy5* + CH_2 *biot*), 2.66 (1H, d, *J* = 12.8 Hz, CH_2S *biot*), 2.89 (1H, dd, *J* = 12.7, 4.9 Hz, CH_2S), 3.16-3.25 (1H, m, SCH *biot*), 3.30-3.35 (2H, m, CH_2 *Cy5*), 3.47-3.54 (4H, m, *2x* CH_2 *eg*), 3.56-3.64 (4H, m, 2x CH_2 *eg*), 4.09-4.20 (4H, m, 2x CH_2 *eg*), 4.29 (1H, dd, *J* = 7.8, 4.4 Hz, CHN *biot*), 4.47 (1H, dd, *J* = 7.8, 4.9 Hz, CHN *biot*), 6.31-6.37 (1H, m, *Cy5*), 6.64-6.73 (1H, m, *Cy5*), 7.32-7.35 (2H, m, *Cy5*), 7.86-7.90 (3H, m, *Cy5*) and 8.27-8.35 (2H, m, *Cy5*) ppm.

ESI-MS: m/z 1013.6 $[M+H]^+$ (calculated: 1013.4).

MALDI-TOF: m/z 1013.5 $[M+H]^+$ (calc. 1013.4), 1051.5 $[M+K]^+$ (calc. 1051.4).

6.2.8. Synthesis of the maleimide-derived fluorophores

2-(*tert*-butoxycarbonyl)ethanolamine (55)



Ethanolamine (1.53g, 25mmol) and 4-dimethylaminopyridine (334mg, 2.75mmol) were dissolved in 30mL acetonitrile. Di-*tert*-butyldicarbonate (6.0g, 25.5mmol) was added to this solution (gas evolution observed). The reaction mixture was stirred at room temperature for 3h. The solvent was removed by evaporation and the crude product was redissolved in EtOAc. The organic solution was washed with $KHSO_4$ 1M (2x), brine (1x), dried over Na_2SO_4 and concentrated under reduced pressure. No further purification step was performed.

Yield: 3.2g colorless oil (90% pure, 18mmol, 76%)

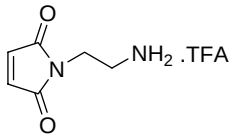
TLC: R_f = 0.54 (EtOAc)

$C_7H_{15}NO_3$ (161.2)

^{1}H-NMR ($CDCl_3$, 400MHz): δ = 1.38 (9H, s, $C(CH_3)_3$), 3.20 (2H, t, *J* = 5.1 Hz, CH_2), 3.32 (1H, br s, OH), 3.61 (2H, t, *J* = 5.1 Hz, CH_2) and 5.20 (1H, br s, NH) ppm.

^{13}C-NMR ($CDCl_3$, 100MHz): δ = 28.6 ($C(CH_3)_3$), 43.3 ($C(CH_3)_3$), 62.3 (CH_2), 82.6 (CH_2) and 156.9 (C=O) ppm.

2-(maleimido)ethylamine (54)



Step 1: A solution of triphenylphosphine (2.7g, 10.3mmol) in 20mL dry THF was cooled to -78°C and DIAD (2.0mL, 10.3mmol) was added slowly. Afterwards, a solution of maleimide (1.0g, 10.3mmol) and 2-(*tert*-butoxycarbonyl)ethanolamine **55** (2.0g, 11mmol) in 15mL dry THF was added via syringe and the reaction mixture was stirred for 10 minutes at -78°C and then allowed to reach room temperature by stirring overnight. The solvent was removed in vacuum and the product purified by flash silica gel chromatography, eluting with EtOAc:cHex (3:1). Fractions containing the intermediate product were collected and concentrated (2.25g colorless solid, GC-MS (Method A): = 4.27min; m/z: 240 $[M]^+$, 167 $[M\text{-}O\textit{t}Bu]^+$).

Step 2: 2.25g of *N*-(*tert*-butoxycarbonyl)-2-(maleimido)ethylamine were dissolved in 50mL DCM, cooled to 0°C and treated with 20mL trifluoroacetic acid for 30 minutes. After evaporation of DCM, excess of TFA was removed by coevaporation with toluene. The residue was dissolved in MeOH and the product was precipitated by adding Et_2O, separated by filtration, redissolved in MeOH:H_2O and dried by lyophilization.

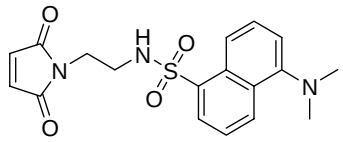
Yield: 1.42g colorless solid (5.6mmol, 55%)

$C_6H_8N_2O_2$.*TFA* (140.14 + *114.02*)

^{1}H-NMR ($CDCl_3$, 400MHz): δ = 3.17 (2H, t, *J* = 5.8 Hz, CH_2), 3.81 (2H, t, *J* = 5.7 Hz, CH_2) and 6.88 (2H, s, CH=CH) ppm.

^{13}C-NMR ($CDCl_3$, 100MHz): δ = 35.0 (CH_2), 38.7 (CH_2), 134.6 (CH=CH) and 171.2 (C=O) ppm.

***N*-dansyl-2-(maleimido)ethylamine (52)**



2-(maleimido)-ethylamine **54** (38mg, 0.15mmol), dansyl chloride (54mg, 0.20mmol) and DIPEA (78µl, 0.45mmol) were dissolved in 2.0mL dry DMF and the reaction solution stirred at room temperature for 1h. 120µl acetic acid was added and the solution concentrated under high vacuum. The residue was redissolved in DCM and the organic solution was washed with HCl 0.1M (3x), brine (1x), dried over Na_2SO_4 and concentrated. The product was purified by reversed-phase HPLC (gradient (time/solvent B) of 0min (10%B) → 20min (100%B), retention time: 7.1min) and freeze-dried.

Yield: 45mg light yellow solid (0.12mmol, 80%), fluorescent

$C_{18}H_{19}N_3O_4S$ (373.43)

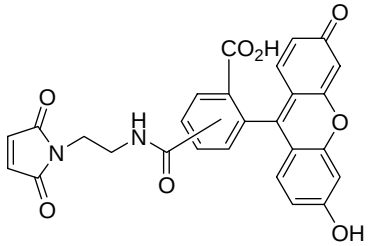
^{1}H-NMR (CD_3OD, 400MHz): δ = 3.09-3.13 (2H, m, CH_2), 3.11 (6H, s, $N(CH_3)_2$), 3.45 (2H, t, *J* = 5.7 Hz, CH_2), 6.45 (2H, s, CH=CH *maleoyl*), 7.56 (1H, d, *J* = 7.7 Hz, Ar), 7.62-7.68 (2H, m, Ar), 8.19 (1H, d, *J* = 7.3 Hz, Ar), 8.44 (1H, d, *J* = 8.6 Hz, Ar) and 8.48 (1H, d, *J* = 8.6 Hz, Ar) ppm.

^{13}C-NMR (CD_3OD, 100MHz): δ = 37.2 (CH_2), 40.3 (CH_2), 45.3 ($N(CH_3)_2$), 116.8 (Ar), 122.5 (Ar), 124.5 (Ar), 127.8 (Ar), 128.1 (Ar), 128.5 (Ar), 129.4 (Ar), 129.5 (Ar), 133.5 (CH=CH) and 170.9 (C=O) ppm.

MALDI-TOF: m/z 374.5 $[M+H]^+$ (calc. 374.1), 396.5 $[M+Na]^+$ (calc. 396.1), 412.5 $[M+K]^+$ (calc. 412.1).

FAB-HRMS: m/z 373.1098 $[M]^+$ (calc. 373.1096).

***N*-fluorescein-2-maleimido-ethylamine (53)**



2-(maleimido)ethylamine **54** (9mg, 0.035mmol), fluorescein succinimidyl ester (15mg, 0.028mmol) and DIPEA (16µl, 0.106mmol) were dissolved in 1.0mL dry DMF and the

reaction solution stirred at room temperature for 2.5h. 15µl acetic acid were added and the product directly purified by reversed-phase HPLC (gradient (time/solvent B) of 0min (5%B) → 3min (5%B) → 15min (40%B) → 22min (100%B), retention time: 15.5min).

Yield: 6.5mg yellow solid (0.013mmol, 46%), fluorescent

$C_{27}H_{18}N_2O_8$ (498.44)

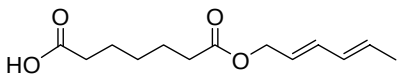
^{1}H-NMR (CD_3OD, 400MHz): δ = (I/II, ratio 1:0.65) 3.49/3.61 (2H, t, *J* = 6.1 Hz, CH_2), 3.67/3.78 (2H, t, *J* = 5.0 Hz, CH_2), 6.22-6.67 (2H, m, *Xan*), 6.71/6.82 (2H, s, CH=CH *maleoyl*), 6.71-6.82 (2H, m, *Xan*), 6.78-6.80 (2H, m, *Xan*), 7.31 (1H, d, *J* = 8.0 Hz, Ar [II]), 7.54 (1H, s, Ar [I]), 8.03 (1H, d, *J* = 8.0 Hz, Ar [I]), 8.11 (1H, d, *J* = 8.0 Hz, Ar [I+II]) and 8.34 (1H, s, Ar [II]) ppm.

ESI-MS: 499.5 $[M+H]^+$ (calc. 499.1)

MALDI-TOF: m/z 499.5 $[M+H]^+$ (calc. 499.1), 521.5 $[M+Na]^+$ (calc. 521.1), 537.5 $[M+K]^+$ (calc. 537.1).

6.2.9. Synthesis of the diene cross-linker

Pimeloyl diene ester (59)



To a solution of pimelic acid (3.2g, 20mmol), *trans*,*trans*-2,4-hexadien-1-ol (393mg, 4mmol) and DMAP (49mg, 0.4mmol) in 30mL THF was added a solution of DIC (619µl, 4mmol) in 5mL THF dropwise over a 30-minutes period at room temperature. The reaction mixture was stirred overnight. The solvent was then evaporated and urea co-product was removed by precipitation with EtOAc/cHex. The filtrate was concentrated under reduced pressure and product was purified by flash silica gel chromatography, eluting with cHex:EtOAc (10:3).

Yield: 401mg colorless oil (1.7mmol, 43%)

TLC: R_f = 0.16 (cHex:EtOAc, 10:3)

$C_{13}H_{20}O_4$ (240.3)

^{1}H-NMR ($CDCl_3$, 400MHz): δ = 1.33-1.41 (2H, m, CH_2), 1.60-1.68 (4H, m, *2x* CH_2), 1.75 (3H, d, *J* = 6.6 Hz, CH_3 *Hxd*), 2.32 (2H, t, *J* = 7.4 Hz, CH_2), 2.35 (2H, t, *J* = 7.4 Hz, CH_2), 4.56 (2H, d, *J* = 6.6 Hz, CH_2 *Hxd*), 5.61 (1H, ddd, , *J* =13.6, 6.4, 6.4 Hz, CH=C*H*CH_3), 5.70-

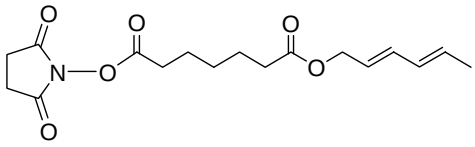
5.79 (1H, m, CH_2CH=CH), 6.04 (1H, dd, *J* = 15.1, 10.5 Hz, CH=$CHCH_3$) and 6.24 (1H, dd, *J* = 15.2, 10.4 Hz, CH_2CH=CH) ppm.

^{13}C-NMR ($CDCl_3$, 100MHz): δ = 18.3 (CH_3), 24.5 (CH_2), 24.8 (CH_2), 28.7 (CH_2), 34.0 (CH_2), 34.3 (CH_2), 65.1 (CH_2 *Hxd*), 123.9 (CH_2-CH=CH), 130.6 (CH_2-CH=CH), 131.5 (CH=CH-CH_3), 135.1 (CH_2-CH=CH), 173.6 (COOhxd) and 179.9 (COOH) ppm.

LC-MS (Method B): t_r = 8.36 min; m/z 263.1 $[M+Na]^+$ (calc. 263.1).

MALDI-TOF: 263.2 $[M+Na]^+$ (calc. 263.1), 279.2 $[M+K]^+$ (calc. 279.1).

Pimeloyl succinimidyl hexadienyl ester (49)



Pimeloyl diene ester **50** (160mg, 0.66mmol), *N*-hydroxysuccinimide (96mg, 0.83mmol) and 4-dimethylaminopyridine (8.6mg, 0.07mmol) were dissolved in 6mL dry THF. Then DIC (118µl, 0.76mmol) was added dropwise at room temperature and the reaction mixture allowed to react overnight. The solvent was removed under reduced pressure and urea co-product was precipitated by adding EtOAc/cHex solution and separated by filtration. The crude product was purified by flash silica gel chromatography using a gradient elution solvent system starting from cHex:EtOAc (3:1) to (1:1).

Yield: 184mg colorless oil (0.55mmol, 82%)

TLC: R_f = 0.31 (cHex:EtOAc, 1:1)

$C_{17}H_{23}NO_6$ (337.37)

^{1}H-NMR ($CDCl_3$, 400MHz): δ = 1.38-1.46 (2H, m, CH_2), 1.60-1.68 (2H, m, CH_2), 1.69-1.77 (2H, m, CH_2), 1.73 (3H, d, *J* = 6.6 Hz, CH_3 *Hxd*), 2.30 (2H, t, *J* = 7.5 Hz, CH_2), 2.58 (2H, t, *J* = 7.5 Hz, CH_2), 2.80 (4H, br s, *2x* CH_2 *NHS*), 4.54 (2H, d, *J* = 6.6 Hz, CH_2 *Hxd*), 5.59 (1H, ddd, , *J* = 14.0, 6.6, 6.6 Hz, CH=$CH$$CH_3$), 5.68-5.77 (1H, m, CH_2CH=CH), 6.02 (1H, dd, *J* = 15.0, 10.7 Hz, CH=$CHCH_3$) and 6.21 (1H, dd, *J* = 15.1, 10.5 Hz, CH_2CH=CH) ppm.

^{13}C-NMR ($CDCl_3$, 100MHz): δ = 18.3 (CH_3), 24.4 (CH_2), 24.5 (CH_2), 25.8 (*2x* CH_2 *NHS*), 28.4 (CH_2), 30.9 (CH_2), 34.1 (CH_2), 65.0 (CH_2 *Hxd*), 123.9 (CH_2-CH=CH), 130.6 (CH_2-CH=CH), 131.4 (CH=CH-CH_3), 135.0 (CH_2-CH=CH), 168.7 (CONHS), 169.4 (*2x* C=O *NHS*) and 173.4 (COOhxd) ppm.

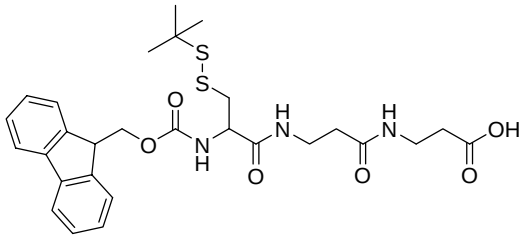
LC-MS (Method B): t_r = 9.09 min; m/z : 360.1 $[M+Na]^+$ (calc. 360.1).

MALDI-TOF: 360.2 $[M+Na]^+$ (calc. 360.1), 376.2 $[M+K]^+$ (calc. 376.1).

FAB-HRMS: 360.1448 $[M+Na]^+$ (calc. 360.1423).

6.2.10. Synthesis of the diene Cys-linkers to be used in EPL

Fmoc-Cys(*St*Bu)-βAla-βAla-OH (60)



1) Loading on 2-Cl trityl resin: 2 equivalents of the Fmoc-βAla-OH and 4 equivalents of DIPEA were combined in DCM (10mL per gram resin) and added to the resin (480mg, 0.77mmol, pre-swelled in DCM). After stirring the mixture for 4 hours, the resin was washed 3x with DCM/MeOH/DIPEA (17:2:1), then 3x DCM, 2x DMF, 2x DCM and finally dried under reduced pressure. The loading was quantitative (determined by UV measurement).

2) Peptide chain assembly: general procedure (section 7.2.1)

3) Cleavage: The peptidyl resin was treated four times with a solution of 1% TFA in DCM (approximately 10mL/g resin) for 5-10 minutes. The filtrates were combined and a solution of 10% pyridine in methanol was added (1/2 volume from peptide solution). The solvent was partially removed under reduced pressure and the peptide was precipitated by adding water. The solid was removed by filtration, dissolved in a mixture of methanol/water and finally dried by lyophilization.

Yield: 342mg colorless solid (0.60mmol, 78%)

$C_{28}H_{35}N_3O_6S_2$ (573.72)

^{1}H-NMR (CD_3OD, 400 MHz): δ = 1.32 (9H, s, *tBu*), 2.38 (2H, t, *J* = 6.7 Hz, α-CH_2 *βAla*), 2.49 (2H, t, *J* = 6.8 Hz, α-CH_2 *βAla*), 2.92 (1H, dd, *J* = 13.4, 8.8 Hz, β-CH_2 *Cys*), 3.12 (1H, dd, *J* = 13.4, 5.1 Hz, β-CH_2 *Cys*), 3.35-3.44 (4H, m, *2x* β-CH_2 *βAla*), 4.23 (1H, t, *J* = 6.7 Hz, CH *Fmoc*), 4.28-4.45 (3H, m, CH_2 *Fmoc* + α-CH *Cys*), 7.31 (2H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.39 (2H, t, *J* = 7.3 Hz, Ar *Fmoc*), 7.66 (2H, d, *J* = 7.4 Hz, Ar *Fmoc*) and 7.78 (2H, d, *J* = 7.5 Hz, Ar *Fmoc*) ppm.

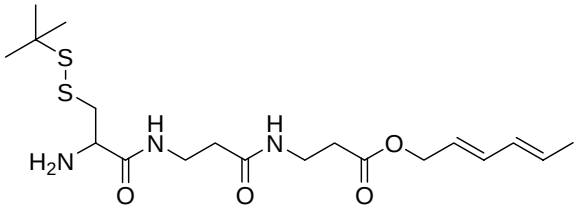
^{13}C-NMR (CD_3OD, 100 MHz): δ = 29.1, 33.5, 35.2, 35.3, 35.9, 42.0, 43.6, 54.9, 67.0, 119.7, 125.1, 127.0, 127.6, 141.4, 144.0, 157.0, 171.7, 172.5 and 174.8 ppm.

MALDI-TOF: m/z 574.3 $[M+H]^+$ (calc. 574.2), 596.3 $[M+Na]^+$ (calc. 596.2), 612.2 $[M+K]^+$ (calc. 612.2).

LC-MS (Method B): t_r = 8.8 min, m/z: 574.2 $[M+H]^+$ (calc. 574.2).

$[\alpha]_D^{20}$**:** - 65.5 (c = 0.7, MeOH)

H-Cys(*St*Bu)-βAla-βAla-hexadienylether (57)



1) Esterification: 100mg Fmoc-protected peptide **60** (0.17 mmol) were dissolved in 4mL DCM:DMF (3:1) and combined with a solution containing *trans*,*trans*-2,4-hexadien-1-ol (23mg, 0.23mmol) and triethylamine (48µl, 0.35mmol). After cooling the mixture to 0°C, HOBt (53mg, 0.35mmol) and EDC (43mg, 0.23mmol) were added. Next, the cooling bath was removed and reaction was stirred overnight at room temperature. After 16h, the solvents were removed under vacuum and the resulting solid was dissolved in 20mL chloroform and washed 3x 5% $NaHCO_3$, 3x water, 1x brine, dried over Na_2SO_4 and concentrated in high vacuum. The product was purified by flash chromatography eluting with ethylacetate:cyclohexane (2:1) to give 56mg of a colorless oil (0.086mmol, 50% yield). *TLC:* R_f = 0.33 (EtOAc). *^{1}H-NMR* ($CDCl_3$, 400 MHz): δ = 1.33 (9H, s, *tBu*), 1.74 (3H, d, *J* = 6.7 Hz, CH_3 *Hxd*), 2.37 (2H, t, *J* = 5.6 Hz, α-CH_2 *βAla*), 2.50 (2H, t, *J* = 5.9 Hz, α-CH_2 *βAla*), 3.03-3.13 (2H, m, β-CH_2 *Cys*), 3.40-3.59 (4H, m, *2x* β-CH_2 *βAla*), 4.21 (1H, t, *J* = 7.1 Hz, CH *Fmoc*), 4.32-4.45 (3H, m, α-CH *Cys* + CH_2 *Fmoc*), 4.56 (2H, d, *J* = 6.6 Hz, (CH_2 *Hxd*), 5.57 (1H, ddd, *J* = 13.8, 6.8, 6.8 Hz, CH=C*H*CH_3), 5.68-5.76 (1H, m, CH_2C*H*=CH), 5.92 (1H, br s, CONH), 6.01 (1H, dd, *J* = 15.0, 10.5 Hz, C*H*=$CHCH_3$), 6.21 (1H, dd, *J* = 15.2, 10.5 Hz, CH_2CH=C*H*), 6.40 (1H, br s, CONH), 7.15 (1H, br s, CONH), 7.30 (2H, t, *J* = 7.5 Hz, Ar *Fmoc*), 7.39 (2H, t, *J* = 7.5 Hz, Ar *Fmoc*), 7.59 (2H, d, *J* = 7.4 Hz, Ar *Fmoc*) and 7.75 (2H, d, *J* = 7.5 Hz, Ar *Fmoc*) ppm. *LC-MS* (Method B): t_r = 10.6 min; m/z : 676.4 $[M+Na]^+$ (calc. 676.2). $[\alpha]_D$: - 12.3 (c = 1.1, $CHCl_3$).

2) Fmoc- deprotection: 39mg Fmoc-protected peptide (0.06 mmol) were treated with 1.0mL 20% piperidine in DCM for 40 minutes at 0°C. The reaction mixture was coevaporated with methanol to remove excess of piperidine. The product was purified by reversed-phase-HPLC, fractions containing the product were combined and dried by lyophilization.

Yield: 16mg colorless oil (0.037mmol, 60% last step, 30% overall yield)

$C_{19}H_{33}N_3O_4S_2$ (431.61)

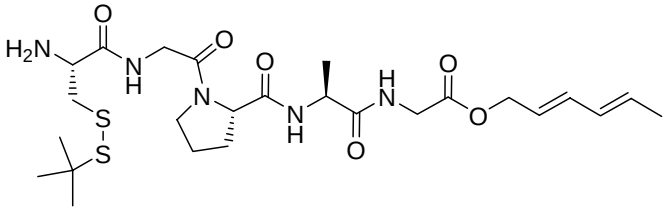
^{1}H-NMR (CD_3OD, 400 MHz): δ = 1.36 (9H, s, *tBu*), 1.74 (3H, d, *J* = 6.7 Hz, CH_3 *Hxd*), 2.42 (2H, t, *J* = 6.7 Hz, α-CH_2 *βAla*), 2.54 (2H, t, *J* = 6.7 Hz, α-CH_2 *βAla*), 3.07 (1H, dd, *J* = 14.2, 7.7 Hz, β-CH_2 *Cys*), 3.19 (1H, dd, *J* = 14.2, 5.8 Hz, β-CH_2 *Cys*), 3.41-3.50 (4H, m, 2x β-CH_2 *βAla*), 4.04 (1H, dd, *J* = 7.6, 5.8 Hz, α-CH *Cys*), 4.58 (2H, d, *J* = 6.6 Hz, CH_2 *Hxd*), 5.62 (1H, ddd, *J* = 13.8, 6.6, 6.6 Hz, CH=C*H*CH_3), 5.71-5.80 (1H, m, CH_2C*H*=CH), 6.06 (1H, dd, *J* = 15.0, 10.6 Hz, C*H*=$CHCH_3$) and 6.26 (1H, dd, *J* = 15.2, 10.4 Hz, CH_2CH=C*H*) ppm.

ESI-MS: m/z 432.2 $[M+H]^+$ (calc. 432.2)

MALDI-TOF: m/z 432.6 $[M+H]^+$ (calc. 432.2), 454.6 $[M+Na]^+$ (calc. 454.2), 470.6 $[M+K]^+$ (calc. 470.2).

$[\alpha]_D^{20}$: - 26.9 (c = 0.3, MeOH)

H-Cys(*St*Bu)-Gly-Pro-Ala-Gly-O-hexadienylester (58)



1) Solid-phase synthesis using sulfonamide resin: (see general procedure in 7.2.4.1.) Starting from 291mg Fmoc-Gly loaded sulfonamide resin (0.24mmol). Purified by flash chromatography, eluting with DCM and DCM:MeOH 100:1→20:1 to give 70mg of a colorless solid (0.09mmol, 37%) of Fmoc-Cys(StBu)-Gly-Pro-Ala-Gly-O-hexadienylester $C_{40}H_{51}N_5O_8S_2$ (793.99). *TLC:* R_f : 0.31 (DCM:MeOH, 10:1). *MALDI-TOF:* m/z 816.6 $[M+Na]^+$ (calc. 816.3), 833.5 $[M+K]^+$ (calc. 833.3).

2) Removal of the Fmoc protecting group: (see general procedure 7.2.4.2.) 70mg of Fmoc-protected peptide (0.09mmol) was treated with 1.5mL 20% piperidine in DCM.

HPLC purification: gradient (time/solvent B) of 0min (10%B) → 2min (10%B) → 15min (70%B), retention time 13.8min.

Yield: 18mg colorless solid (0.05mmol, 35%)

$C_{25}H_{41}N_5O_6S_2$ (571.75)

^{1}H-NMR (CD_3OD, 400 MHz): δ = 1.37 (9H, s, *tBu*), 1.38 (3H, d, *J* = 7.3 Hz, CH_3 *Ala*), 1.75 (2H, d, *J* = 6.7 Hz, CH_3 *Hxd*), 1.86-2.38 (4H, m, β-CH_2 + γ-CH_2 *Pro*), 3.07 (1H, dd, *J* = 8.5,

14.3 Hz, β-CH_2 *Cys*), 3.26 (1H, dd, *J* = 5.2, 14.1 Hz, β-CH_2 *Cys*), 3.52-3.74 (2H, m, δ-CH_2 *Pro*), 3.89-4.01 (2H, m, α-CH_2 *Gly*), 4.05-4.17 (3H, m, α-CH_2 *Gly* + α-CH *Cys*), 4.36-4.46 (2H, m, *2x* α-CH), 4.62 (2H, d, *J* = 6.6 Hz, CH_2 *Hxd*), 5.61 (1H, ddd, *J* = 13.8, 7.1, 7.1 Hz, CH=C*H*CH_3), 5.72-5.81 (1H, m, CH_2C*H*=CH), 6.06 (1H, dd, *J* = 15.1, 10.5 Hz, C*H*=CHCH_3) and 6.28 (1H, dd, *J* = 15.2, 10.4 Hz, CH_2CH=C*H*) ppm.

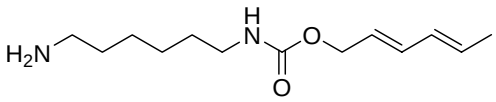
LC-MS (Method B): t_r = 5.7 min, m/z 572.3 $[M+H]^+$ (calc. 572.3), 594.3 $[M+Na]^+$ (calc. 594.2).

MALDI-TOF: m/z 572.7 $[M+H]^+$ (calc. 572.3), 594.7 $[M+Na]^+$ (calc. 594.2), 610.7 $[M+K]^+$ (calc. 610.2).

FAB-HRMS: m/z 571.2521 $[M]^+$ (calc. 571.2498).

$[\alpha]_D^{20}$: - 18.5 (c = 0.2, MeOH)

(6-aminohexyl)-2,4-hexadienyl carbamate (61)



CDI (486mg, 3mmol) was added to a solution of *trans*,*trans*-2,4-hexadien-1-ol (294mg, 3mmol) in dry DMF (5mL) and the reaction solution was stirred at room temperature for 3 hours. A mixture of 6-diaminehexane (1.74g, 15 mmol) and triethylamine (2.0mL, 15 mmol) in 10mL dry DMF was slowly added to the reaction solution (over 10 minutes) and the resulting solution was stirred for 80 minutes. 150mL DCM were added and the organic solution was washed with 3x 50mL sat. $NaHCO_3$, 1x 50mL brine, dried over Na_2SO_4 and concentrated under high vacuum at 55°C. The product was isolated with satisfactory purity.

Yield: 553mg colorless solid (2.3 mmol, 77%)

$C_{13}H_{24}N_2O_2$ (240.34)

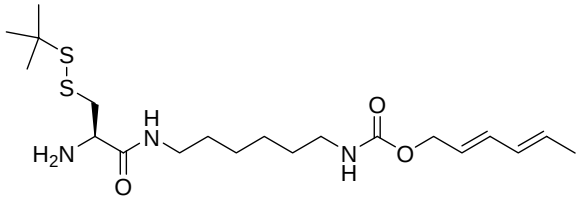
^{1}H-NMR ($CDCl_3$, 400MHz): δ = 1.21-1.33 (4H, m, *2x* CH_2), 1.37-1.47 (4H, m, *2x* CH_2), 1.71 (3H, d, *J* = 6.7 Hz, CH_3), 2.64 (2H, t, *J* = 7.1 Hz, CH_2), 3.10-3.15 (2H, m, CH_2), 4.51 (2H, d, *J* = 6.2 Hz, CH_2 *Hxd*), 4.85 (1H, br s, NH), 5.59 (1H, ddd, *J* = 13.7, 6.5, 6.5 Hz, CH=C*H*CH_3), 5.65-5.74 (1H, m, CH_2C*H*=CH), 6.01 (1H, dd, *J* = 14.9, 10.5 Hz, C*H*=CHCH_3) and 6.19 (1H, dd, *J* = 15.1, 10.5 Hz, CH_2CH=C*H*) ppm.

^{13}C-NMR ($CDCl_3$, 100MHz): δ = 18.3 (CH_3), 26.7 (2x CH_2), 30.1 (CH_2), 33.7 (CH_2), 33.7 (CH_2), 41.1 (CH_2), 42.8 (CH_2), 65.3 (CH_2 *Hxd*), 124.8 (CH_2-*C*H=CH), 130.8 (CH_2-CH=*C*H), 131.1 (*C*H=CH-CH_3), 134.5 (CH_2-*C*H=CH) and 156.7 (CONH) ppm.

ESI-MS: 241.4 $[M+H]^+$ (calc. 241.2)

FAB-HRMS: 241.1935 $[M+H]^+$ (calc. 241.1916).

H-Cys(*St*Bu)- (6-aminohexanyl)-2,4-hexadienyl carbamate (59)



1) Coupling: 100mg of carbamate **61** (0.42mmol) were dissolved in 3mL DCM:DMF (2:1) and combined with 180mg Fmoc-Cys(S*t*Bu)-OH (0.42mmol) in 5mL DCM. The mixture was cooled to 0°C and HOBt (77mg, 0.50mmol) and DIC (78µL, 0.50mmol) were added and the reaction was stirred at room temperature overnight. 20mL DCM were added and the solution was washed 1x $KHSO_4$ 1M, 3x $NaHCO_3$ 5%, 1x brine, dried over Na_2SO_4 and concentrated under reduced pressure. Urea was removed by precipitation in ethyl acetate/cyclohexane and the product was purified by flash chromatography eluting with cyclohexane:ethyl acetate (2:1), affording 121mg of colorless solid (0.19mmol, 45% yield). *^{1}H-NMR* ($CDCl_3$, 400MHz): δ = 1.23-1.38 (4H, m, *2x* CH_2), 1.34 (9H, s, *tBu*), 1.39-1.49 (4H, m, *2x* CH_2), 1.74 (3H, d, *J* = 6.6 Hz, CH_3 *Hxd*), 2.99-3.18 (4H, m, CH_2 + β-CH_2 *Cys*), 3.19-3.26 (2H, m, CH_2), 4.21 (1H, t, *J* = 7.0 Hz, CH *Fmoc*), 4.31-4.48 (3H, m, CH_2 *Fmoc* + NH), 4.54 (2H, d, *J* = 6.0 Hz, CH_2 *Hxd*), 4.82 (3H, m, α-CH *Cys*), 5.60 (1H, ddd, *J* = 13.5, 6.2, 6.2 Hz, CH=C*H*CH_3), 5.66-5.76 (1H, m, CH_2C*H*=CH), 5.86 (1H, d, *J* = 7.4 Hz, NH), 6.03 (1H, dd, *J* = 15.0, 10.6 Hz, C*H*=$CHCH_3$), 6.21 (1H, dd, *J* = 15.1, 10.4 Hz, CH_2CH=C*H*), 6.44 (1H, br s, NH), 7.29 (2H, t, *J* = 7.5 Hz, Ar *Fmoc*), 7.39 (2H, t, *J* = 7.4 Hz, Ar *Fmoc*), 7.58 (2H, d, *J* = 7.4 Hz, Ar *Fmoc*) and 7.75 (2H, d, *J* = 7.5 Hz, Ar *Fmoc*) ppm.

2) Fmoc-deprotection: Same procedure as for compound **57**. Starting from 90mg of Fmoc-protetcted peptide (0.14mmol).

Yield: 48mg of colorless solid (0.11mmol, 80%)

$C_{20}H_{37}N_3O_3S_2$ (431.66)

^{1}H-NMR (CD_3OD, 400MHz): δ = 1.30-1.41 (4H, m, *2x* CH_2), 1.36 (9H, s, *tBu*), 1.46-1.59 (4H, m, *2x* CH_2), 1.74 (3H, d, *J* = 6.7 Hz, CH_3 *Hxd*), 3.06-3.28 (6H, m, *2x* CH_2 + β-CH_2 *Cys*), 4.04 (1H, dd, J = 6.1, 7.1 Hz, α-CH *Cys*), 4.50 (2H, d, *J* = 6.2 Hz, CH_2 *Hxd*), 5.60 (1H, ddd, *J* = 6.3, 6.3, 14.7 Hz, CH=C*H*CH_3), 5.68-5.77 (1H, m, CH_2C*H*=CH), 6.05 (1H, dd, *J* = 10.5, 15.0 Hz, C*H*=$CHCH_3$) and 6.24 (1H, dd, *J* = 10.5, 15.0 Hz, CH_2CH=C*H*) ppm.

^{13}C-NMR ($CDCl_3$, 100MHz): δ = 17.0 (CH_3), 26.2 (2x CH_2), 28.8 (CH_2), 28.9 (*tBu*), 29.6 (CH_2), 39.5 (CH_2), 40.4 (CH), 41.3 (CH_2), 52.7, (CH_2), 64.8 (CH_2 *Hxd*), 124.8 (CH_2-*C*H=CH), 130.2 (CH_2-CH=*C*H), 130.8 (*C*H=CH-CH_3), 133.8 (CH_2-*C*H=CH), 157.8 (C=O) and 167.0 (C=O) ppm.

6.3. Protein ligation by Diels-Alder reaction

6.3.1. Material and general procedures

Proteins

Streptavidin from *Streptomyces avidinii* specific activity: 14μg biotin/mg protein (Sigma, product number S4762); avidin from egg white specific activity: 10.8μg biotin/mg solid (Sigma, product number A9275).

Rab7ΔC3-MESNA thioester, Rab7 wild type and REP-1 were expressed and purified by members of Dr. Kirill Alexandrov´s group at the MPI Dortmund.

Bradford assay for determination of protein concentration

Bradford Assay was employed to determine the total protein concentration of a sample. The assay is based on the observation that the absorbance maximum for an acidic solution of Coomassie Brilliant Blue G-250 shifts from 465 nm to 595 nm when binding to protein occurs.[119] The assay sample was prepared by combining 900μl of the dye stock solution with 100μl volume of H_2O containing 1-10μg of the protein to be quantified. The protein concentration of a test sample was determined by comparison to that of a series of streptavidin standards.

Analysis by polyacrylamide electrophoresis

Discontinuous PAGE was performed according to the method of LaemmLi. Native and SDS polyacrylamide gels were prepared using the recipes described in Table 9. Electrophoresis was performed in a Bio-Rad gel system. After the run was stopped, the gels were directly scanned for fluorescence using an ultraviolet illuminator (Reprostar II, Camag, 302nm wavelength). Finally, Coomassie staining of the gels was achieved by incubation with *Staining Solution* for ca 15 minutes and subsequent destaining.

Table 9. Recipes for preparation of polyacrylamide gels

	Basic gel PAGE (15%) (for preparation of 5 gels)	**Acidic gel PAGE (15%) (for preparation of 2 gels)**	**SDS-PAGE (15%) (for preparation of 1 gel)**
Separating gel	8.8mL water 9.0mL 1.5M Tris-Cl pH 8.8 18mL 30%acrylamide/0.8%bis-acrylamide 180µl 10% fresh APS 18µl TEMED	4.0mL 1.5M acetate-KOH pH 4.3 3.0mL 50% glycerol 6.7mL 30% acrylamide/0.8%bis-acrylamide 160µl 10% fresh APS 20µl TEMED	1.14mL water 1.25mL 1.5M Tris-Cl pH 8.8 2.50mL 30%acrylamide/0.8%bis-acrylamide 50µl SDS 30µl 10% fresh APS 10µl TEMED
Stacking gel	9.15mL water 3.75mL 0.5M Tris-Cl pH 6.8 2.0mL 30% acrylamide/0.8%bis-acrylamide 25µl 10% fresh APS 15µl TEMED	6.4mL water 2.5mL 0.25M acetate-KOH pH 6.8 1.0mL 30% acrylamide/0.8%bis-acrylamide 100µl 10% fresh APS 10µl TEMED	1.13mL water 0.5mL 0.5M Tris-Cl pH 6.8 0.33mL 30% acrylamide/0.8%bis-acrylamide 20µl SDS 10µl 10% fresh APS 5µl TEMED
Running buffer	(10x): 29g Tris Base 144g glycine water up to 1 liter	(1x): 25g β-alanine 6.4mL acetic acid water up to 800mL liter pH adjust to 4.3	(10x): 15g Tris Base 72g glycine 5g SDS water up to 500mL
Loading Buffer	(2x): 167µl 3.0M Tris-Cl pH 8.45 500µl 0.1% bromophenol blue 2.0mL glycerol 50% water up to 10mL	(5x): 1.45mL glycerol 50% 0.5mL 0.25M acetate-KOH pH 6.8 traces of methyl green	(2x): 2.5mL 0.5M Tris-Cl pH 6.8 4.0mL 10% SDS 500µl 0.1% bromophenol blue 1.0mL β-ME 2.0mL glycerol 50% water up to 10mL
Staining Solution	0.25% Coomassie Brilliant Blue in H_2O:MeOH:AcOH (4:5:1)		
Destaining Solution	Water, methanol 15%, acetic acid 10%		

Protein purification by ultracentrifugation

Depending on the amount of solution to be diafiltered, the following centrifugal filter devices were used: Amicon Ultra 15 low binding regenerated cellulose 10kDa MWCO (Milipore), Microcon YM-10 regenerated cellulose 10kDa MWCO (Milipore) or Vivaspin 500 polyethensulfan membrane 30kDa MWCO (Vivascience). Centrifugation was carried out at room temperature using Eppendorf centrifuge 5415D.

Protein purification by spin gel-filtration columns

An aliquot of 10-20µl of protein solution was loaded on the top of small spin gel filtration columns (DyeEx columns from Qiagen) and the columns were spinned for 3 minutes at 3000 g in an Eppendorf centrifuge 5415D.

Protein purification by dialysis

Up to 100µl of protein solution was pipeted inside small dialysis tubes (Slide-A-Lyser 7KDa MWCO from Pierce) and the solution was dialyzed against 1L of buffer at 4°C for approximately 24 hours with slow stirring.

6.3.2. Labeling of a (strept)avidin-biotin complex

6.3.2.1. *Biochemical Methods*

Preparation of the streptavidin-biotinylated diene complex (45a)

Streptavidin was dissolved in water at a concentration of 1mg/mL. 200µl of this solution (200µg protein, estimated binding capacity: 7.6µg biotinylated diene **44**) was combined with 5.7µl of 10mM biotinylated hexadienylester **44** solution in DMF (37.9µg, 5-fold relative to streptavidin binding capacity) in an 1.5mL Eppendorf tube. After shaking for 30 minutes at room temperature, the solution was placed into an Amicon diafiltration device and diafiltered using five changes of water. The final volume was approximately 0.20mL.

Preparation of avidin-biotinylated diene complex (45b)

Avidin (2.0mg, estimated binding capacity: 59µg diene **44**) was dissolved in 1mL water and treated with 48µl of 10mM solution of biotinylated diene **44** in DMF (323µg, 5-fold relative to avidin binding capacity) for 30 minutes at room temperature. After purification via diafiltration (same procedure as above), the resulting avidin-complex solution was approximately 0.25mL.

Preparation of streptavidin-biotinylated cycloadduct complex (48a)

Aliquots of 10µl streptavidin-biotinylated diene complex **45a** (concentration: 3.5mg/mL, 35µg of protein complex contains estimately 1.3µg (2.0nmol) biotinylated diene) were placed into a 0.5mL eppendorf tubes containing 3.0µl of water. To these solutions, 2.0µl of 50mM

dienophile **46** solution (100nmol, 50-fold related to diene content) was added and the combination allowed reacting for 48h at 25°C. After this time, the mixtures were transferred to a Vivaspin membrane ultrafiltration apparatus and diafiltered with three changes of water. The resulting solution was placed in an Eppendorf tube and submitted to fluorescence scan (Figure 11A). The ligation reaction was analyzed by basic native gel electrophoresis (Figure 13A) and MALDI experiments.

Preparation of avidin-biotinylated cycloadduct complex (48b)

Aliquots of 10µl avidin-biotinylated diene complex **45b** (concentration: 8mg/mL, 80µg of protein complex contains estimately 2.4µg (3.5nmol) biotinylated diene) were placed in eppendorf tubes and treated with the following amounts of a 50mM maleimido-fluorescently labeled peptide **46** solution: 0.78, 1.94, 3.88 and 7.76µl, corresponding respectively to addition of 10-, 25-, 50- and 100-fold of dienophile in comparison with diene content. The solutions were agitated for 2 days at 25°C and purified by a Microcon diafiltration device using three changes of water. The resulting solutions were transferred into eppendorf tubes and scanned for fluorescence detection (Figure 11B). Small samples from these solutions were analysed by acidic native gel electrophoresis (Figure 13B) and MALDI experiments.

Preparation of streptavidin-biotinylated cycloadduct complex (48c)

In a 1.5mL eppendorf tube, 100µl of the streptavidin-diene complex **45a** at 1mg/mL concentration (100µg of protein complex contains estimately 3.8µg (5.7nmol) biotinylated diene) was mixed with 57.1µl maleimide-peptide **5d** 5mM solution (285nmol, 50-fold) and allowed to react for 24 hours at 25°C with stirring. Unligated peptide was removed by membrane ultrafiltration (Amicon, five water changes) and the resultant solution was analyzed by mass spectroscopy.

Control experiments

Following the same procedure described above, all reactions were performed with wild-type streptavidin or avidin in replacement of the protein-diene complex for the investigation of unspecific reactions.

6.3.2.2. *Analysis of the complexes*

Preparation of samples for MALDI-TOF measurements

Aliquots of 2-5µl of sample solution were mixed in 1:1 ratio with isopropanol and heated to boiling for some seconds. 1µl of this solution was combined with 1µl sinapinic acid matrix (saturated sinapinic acid in acetonitrile/0.3% TFA 1:2), placed on the MALDI plate and analysed (Table 10).

Table 10. MALDI-TOF mass spectra of the ligands from the streptavidin-biotin avidin-biotin complexes.

Streptavidin (avidin)-biotinylated diene complex **45a** and **45b**	The expected mass peak for biotinylated diene **44** was found: 687.6 $[M+Na]^+$ (calc. 687.4), 703.5 $[M+K]^+$ (calc. 703.3).
Streptavidin-biotinylated cycloadduct **48a** and **48b**	The expected mass for the DA ligation product of diene **44** and dienophile **46** ($C_{72}H_{109}N_{15}O_{20}S_2$, MW: 1568.85) was observed: 1569.0 $[M+H]^+$ (calc. 1568.7), 1591.0 $[M+Na]^+$ (calc. 1590.7), 1607.0 $[M+K]^+$ (calc. 1606.7).
Streptavidin-biotinylated cycloadduct complex **48c**	The correct mass for the DA ligation product of biotinylated diene **44** and dienophile **5d** ($C_{79}H_{120}N_{18}O_{28}S$, MW: 1801.97) was observed (1801.8 $[M+H]^+$ (calc. 1801.8), 1824.8 $[M+Na]^+$ (calc. 1823.8), 1839.8 $[M+K]^+$ (calc. 1839.7)).

Native gel electrophoresis

Nondenaturating discontinuous polyacrylamide gel electrophoresis was applied to investigate the behavior of the streptavidin/avidin-biotinylated complexes upon ligation conditions. Streptavidin was analyzed by basic PAGE protocol, while acidic native electrophoresis was utilized for avidin due to its high basic nature (pI ~ 10.5). The recipes described in Table 9 were used for preparation of the gels. Before electrophoresis, samples were mixed with Loading Buffer and directly loaded into the gel wells. The basic gels were run at constant 125V while the acidic gels at constant 35mA.

6.3.3. Bioconjugation of streptavidin by Diels-Alder ligation

6.3.3.1. *Biochemical Methods*

Preparation of the streptavidin-diene conjugate (51)

Streptavidin (1.09mg, 21nmols) was dissolved in 450μl water and incubated with 2.5μl of a freshly prepared 50mM solution of the diene cross linker in DMF (125nmol, 6-fold relative to streptavidin) for 2h at 25°C. The reaction mixture was transferred to a Microcon centrifugal filtration device, diafiltered with four changes of water (centrifuged ca 30 minutes at 13 x g) and concentrated to a final volume of 55μl. Protein concentration for this solution was 19mg/mL. The protein solution was stored at -80°C.

Diels-Alder ligation between streptavidin-diene conjugate (51) and fluorescently labeled dienophiles

◈ **In water:** An aliquot of 10μL of streptavidin-diene conjugate **51** at 5mg/mL concentration (50μg, 0.9nmol) was combined with 2.7μL of a 10mM solution of dienophile **46**, **52** or **53** (27nmol, 30-fold) and kept at 25°C for 24 hours while shaking. After this time, excess of dienophile was removed by passing the reaction mixture through a DyeEx spin gel filtration column. The purified solution was placed into 0.5mL Eppendorf tubes and analyzed by ultraviolet radiation (see Scheme 22). Samples of the cycloadduct conjugate solution were taken for realization of SDS-PAGE and MALDI-TOF tests.

◈ **At different pH:** An aliquot of 8μl of conjugate solution **51** (at 5mg/mL concentration) was diluted with 2.5 μl sodium phosphate buffer 0.1M (at different pH: 5.5, 6.0, 6.5 and 7.0) and combined with 7.5 μl of a 10mM solution of maleimide **46.** After 24h, the ligated protein was purified as described above.

Control experiments

The same procedures described above were performed using wild type streptavidin instead of the streptavidin-diene conjugate as a negative control experiment.

6.3.3.2. *Analysis of the conjugates*

MALDI-TOF measurements

Mass spectra of the conjugated proteins were performed using sinapinic acid as matrix (saturated sinapinic acid in acetonitrile/0.3% TFA 1:2). Only the mass of monomeric streptavidin gave well-defined peaks. Because the composition of the streptavidin subunits varies from 123 to 125 amino acids, a range of different mass peaks were found between 13115 and 14000 for the spectrum of streptavidin (Figure 15A). The largest subunit peak, experimentally found at mass 13180 (± 6) Da, was taken as reference for calculation of the expected protein conjugate mass.

Denaturating SDS-PAGE

Discontinuous SDS-PAGE was performed according to the method of LaemmLi (1970). The gels were prepared following the recipes indicated in Table 9. Before electrophoresis, samples were mixed with Loading Buffer and loaded into the gel wells. In the cases where visualization of the streptavidin subunits was desired, the sample-mix was heated at 80°C for 5 minutes prior loading. Electrophoresis was run at constant 35mA.

Titrimetric assays for the determination of biotin binding capacity of streptavidin

◈ **Assay 1:** A sample of 1mL of a streptavidin at 0.1 mg/mL concentration in ammonium carbonate buffer 0.2 M pH 8.9 was added to a 1 mL UV-cuvette and the absorbance was read at 233 nm (A_0). Aliquots of d-biotin (0.1 mg/mL in ammonium carbonate buffer 0.2 M) were added in increments of 2 μL stepwise. The solution was stirred after each addition with the help of a magnetic stirrer and the absorbance recorded at 233 nm (A). The titration curve was built by plotting the differential absorbance ($A-A_0$) *vs* volume of biotin added (Figure 17A). The amount of biotin at the inflection point is divided by concentration of the streptavidin samples to give the specific activity:

$$\text{Units/mg} = V_1.C_1/V_2.C_2$$

where: V_1 = volume of biotin at inflection point, C_1 = concentration of the biotin solution, V_2 = volume of the streptavidin sample, C_2 = concentration of the streptavidin sample.

◈ **Assay 2:** To a 1 mL UV-cuvette, 1mL of a streptavidin sample at 0.1 mg/mL concentration in sodium phosphate buffer 0.1 M pH 7.0 was added together with 25μL of a solution of HABA 10mM in 10mM sodium hydroxide. The absorbance was read at 500 nm (A_0). Aliquots of d-biotin (2mM in sodium phosphate buffer 0.1 M) were added in increments

of 2 μL stepwise. The solution was stirred after each addition with the help of a magnetic stirrer and the absorbance recorded at 500 nm (A). The titration curve was built by plotting the differential absorbance ($A-A_0$) *vs* volume of biotin added (Figure 17A).

6.3.4. Site-specific labeling of Rab proteins

6.3.4.1. *Biochemical methods*

Cystein ligation of the Rab7ΔC3 thioester and peptides 57 or 58.

An aliquot of 100μL of a stock solution of Rab7ΔC3 thioester (6.8 mg/mL in buffer 25mM sodium phosphate pH 7.5, 25mM NaCl, 0.5 M MESNA, 10μM GDP, 2mM $MgCl_2$, 30nmol) was combined with 13μL of ligation buffer (25mM sodium phosphate pH 7.5, 100mM NaCl, 100mM MESNA, 2% CHAPS, 50μM GDP, 1mM $MgCl_2$) and 10μL of peptide **57** or **58** (60mM in methanol, 600nmol) was added. The final concentrations were 250μM for Rab7 thioester and 5mM for the peptide (approximately 20 equivalents). The ligation reaction mixture was incubated overnight at 16°C with slightly shaking. Small samples were removed for analysis by ESI-MS and SDS-PAGE. The resulting ligated protein solution was directly submitted to the next step without further purification.

Protection of cysteine residues with Ellmann´s reagent

An aliquot of 85μL of the dienyl Rab7 solution resulting from the cystein ligation (ca. 5.7 mg/mL) was mixed with 115μL of DTNB 30mM (in 60mM sodium phosphate pH 8) at 25°C for 4 hours. Next the yellowish reaction solution was dialyzed against DA buffer (5mM sodium phosphate buffer pH 6.0, 20mM NaCl, 0.2mM $MgCl_2$ and 20μM GDP). The final concentration was approximately 1.2-1.5 mg/mL of a colorless solution. Small samples were removed for analysis by ESI-MS and SDS-PAGE. The protein solution was frozen and stored at -80°C.

Diels-Alder ligation of Rab7 hexadienyl ester 62 and maleimide compounds

In different scale experiments, 10-300 μl of Rab7 hexadienyl ester **62** solution in DA buffer (5mM sodium phosphate buffer pH 6.0, 20mM NaCl, 0.2mM $MgCl_2$ and 20μM GDP) at concentration of approximately 1 mg/mL (ca. 40μM protein) were combined with 30 to 100

equivalents of maleimide compounds **46** or **52.** The ligation mixture was incubated at 25°C for 20-24h. Under these conditions, the ligated protein usually precipitated gradually during the reaction course.

◈ **Analysis:** The reaction was quenched by adding DTT 200mM (50 equivalents relative to the amount of dienophile added). The deprotection of the cystein residues was visually noticed by the development of a yellow coloration upon addition of DTT resulted from the release of the TNB groups into solution. After 2h, the reaction mixture was analyzed by SDS-PAGE electrophoresis (Figure 35) and ESI-MS (Figure 36).

◈ **Purification:** The reaction mixture was centrifugated and the supernadant was removed. The pellet was washed with methanol (2x) to remove excess of dienophile and redissolved in denaturating buffer (100mM Tris-HCl pH 8.0, 6M guanidinium-HCl, 100mM DTE, 1% CHAPS and 1mM EDTA) to a concentration of ca. 1 mg/mL (solution became yellowish because of the TNB group release) and incubated overnight at 4°C. The protein was refolded by diluting it 25-fold dropwise with folding buffer (100mM HEPES pH 7.5, 5mM DTE, 2mM $MgCl_2$, 100µM GDP, 1% CHAPS) and incubated at room temperature for 3h with slightly stirring. After that, an equimolar of the REP-1 was added and the solution was incubated overnight at 4°C. The resulting complex was dialysed against 25mM HEPES pH 7.5, 40mM NaCl, 3mM DTE, 2mM $MgCl_2$ and 20µM GDP, concentrated by ultracentrifugation (Amicon 10KDa cut-off) and stored at -80°C.

Control experiments with Rab7 wild type

◈ **Without DTNB blocking:** 20µL of a stock solution of Rab7wt[4] (10 mg/mL in 25mM HEPES pH 7.2, 40mM NaCl, 3mM DTE, 2mM $MgCl_2$ and 10µM GDP) was dialyzed against DA buffer (5mM sodium phosphate buffer pH 6.0, 20mM NaCl, 0.2mM $MgCl_2$ and 20µM GDP) to a final concentration of 5.2 mg/mL. 10µL of this solution (2.1nmol of protein) was treated with 6.3µL of maleimide **46** (10mM, 63nmol) at 25°C for 22h. The reaction was analyzed by SDS-PAGE (Figure 33A) and ESI-MS: found 23717 (Rab7wt: 23719), 25526 (calculated for Rab7wt + 2 addition of **46**: 25527) and 26430 (calculated for Rab7wt + 3 addition of **46**: 26431).

◈ **With DTNB blocking:** 50µL of the stock solution of Rab7wt was dialyzed against 5mM sodium phosphate buffer pH 7.5, 20mM NaCl, 0.2mM $MgCl_2$ and 20µM GDP to a final concentration of 8.7 mg/mL. 23µL of this solution was combined with 46µL of DTNB 30mM (in 60mM sodium phosphate buffer pH 8) for 2 hours at 25°C. The resulting yellowish

solution was dialyzed against DA buffer (5mM sodium phosphate buffer pH 6.0, 20mM NaCl, 0.2mM $MgCl_2$ and 20μM GDP) to a final protein concentration of 1 mg/mL. The protected Rab 7 was analyzed by ESI-MS: found 24311 (calculated for Rab7wt + 3 TNB groups: 24310) and 24703 (calculated for Rab7wt + 5 TNB groups: 24704). 25μL of the protected Rab 7 hexadienyl ester solution in DA buffer (25μg, 1nmol) was treated with 3μL of maleimide **46** (10mM, 30nmol) at 25°C for 20h. The reaction was quenched by addition of 8μl of DTT 200mM (solution became yellow) and analyzed by SDS-PAGE (Figure 33A) and ESI-MS (Figure 33B): found 23719 (Rab7wt: 23719).

6.3.4.2. *Analysis of the modified Rab proteins*

Denaturating SDS-PAGE

Discontinuous SDS-PAGE was performed according to the method of LaemmLi (1970). The gels were prepared following the recipes indicated in Table 9. Before electrophoresis, samples were mixed with Loading Buffer, heating at 75°C for 5 minutes and loaded into the gel wells. Electrophoresis was run at constant 35mA.

ESI-MS mass spectra

Approximately 20-30μg of protein solution sample was submitted to ESI-MS experiments (Method C). The spectra are illustrated in Figure 36.

6.4. Protein immobilization on glass slides

Avidin conjugates

Diene succinimidyl esters **65**, **66** and **67** were synthethized and provided by Dr. José Palomo (MPI Dortmund). 1mg of avidin (150nmols) was dissolved in 200µl water and incubated with 1.0 to 2.0µl of a freshly prepared 100mM solution of the diene cross linkers **65**, **66** or **67** in DMF (450 to 900nmol, 6- to 12-fold relative to avidin) for 4h at 25°C. The reaction mixture was centrifugated and the solution was transferred to a Microcon centrifugal filtration device and diafiltered with four changes of water (centrifuged ca 30 minutes at 13 x g). The conjugates solutions were analyzed by MALDI-TOF (Figure 37).

Glass Slides

PAMAM dendrimer-activated glass slides were kindly provided by Chimera Biotec GmbH (Dortmund).[116] The functionalization of these glass slides with maleimido groups was performed by Maja Köhn (Department of Chemical Biology, MPI Dortmund) as indicated in Scheme 32. The slides were stored at 4°C and used within two weeks.

Spotting and binding with Biotin-Cy5

5µL of each protein solution were spotted on the activated slides using an Eppendorf pipette and the slide was incubated overnight inside a satured wet chamber at room temperature. The slide was washed with water and dried under reduced pressure. After that, aliquots of a 10nM solution of biotin-Cy5 **71** was added over each protein spot (25 µL) or added throughout the entire glass slide surface (1 mL) and incubated for 30 minutes at room temperature. The slide was washed with buffer (10mM sodium phosphate buffer pH 7.5, 0.05% Tween-20) and water, dried under vacuum and scanned for fluorescence.

Fluorescence Scan

The fluorescence intensity of the spotted slides was measured using a microarray laser scanning system (Axon) at Chimera Biotec GmbH by Dr. Ron Wacker.

7. REFERENCES

[1] P. E. Dawson, S. B. Kent, "Synthesis of native proteins by chemical ligation." *Annu. Rev. Biochem.* **2000**, *69*, 923-960.

[2] T. W. Muir, "Semisynthesis of proteins by expressed protein ligation." *Annu. Rev. Biochem.* **2003**, *72*, 249-289.

[3] D. R. W. Hodgson, J. M Sanderson, "The synthesis of peptides and proteins containing non-natural amino acids." *Chem. Soc. Rev.* **2004**, *33*, 422-430.

[4] B. G. Davis, "Chemical modification of biocatalysts." *Curr. Opin. Biotech.* **2003**, *14*, 379-386.

[5] T. M. Hendrickson, V. Crécy-Lagard, P. Schimmel, "Incorporation of nonnatural amino acids into proteins." *Annu. Rev. Biochem.* **2004**, *73*, 147-176.

[6] N. Budisa, "Prolegonema to future experimental efforts on genetic code engineering by expanding its amino acid repertoire." *Angew. Chem. Int. Ed.* **2004**, *43*, 6426-6463.

[7] L. Wang, P. G. Schultz, "Expanding the genetic code." *Chem. Commun.* **2002**, 1-11.

[8] P. M. England, "Unnatural amino acid mutagenesis: a precise tool for probing protein structure and function." *Biochemistry* **2004**, *43*, 11623-11629.

[9] R.S. Goody, K. Alexandrov, M. Engelhard, "Combining chemical and biological techniques to produce modified proteins." *ChemBioChem* **2002**, *3*, 399-403.

[10] M. A. Shogren-Knaak, P. J. Alaimo, K. M. Shokat, "Recent advances in chemical approachesto the study of biological systems." *Annu. Rev. Cell. Dev. Biol.* **2001**, *17*, 405-433.

[11] A.J. Link, M. L. Mock, D. A. Tirell, "Non-canonical amino acids in protein engineering." *Curr. Opin. Biotech.* **2003**, *14*, 603-609.

[12] F. Baneyx, "Recombinant protein expression in *Eschericia coli*." *Curr. Opin. Biotechnol.* **1999**, *10*, 411-421.

[13] G.T. Hermanson, *Bioconjugate Techniques*, Academic Press, San Diego, **1996**.

[14] D. Qi, C. M. Tann, D. Haring, M. D. Distefano, "Generation of new enzymes via covalent modification of existing proteins." *Chem. Rev.* **2001**, *101*, 3081-3111.

[15] R. Wetzel, R. Halualani, J. T. Stults, C. Quan, "A general method for highly selective crosslinking of unprotected polypeptides via pH-controlled modification of N-terminal α-amino groups." *Bioconjugate Chem.* **1990**, *1*, 114-122.

[16] N. Sewald, H.-D. Jakubke, *Peptides: Chemistry and Biology*, Wiley-VCH Verlag, Weinheim, **2002**.

[17] F. Bordusa, "Proteases in organic synthesis." *Chem. Rev.* **2002**, *102*, 4817-4868.

[18] D. Y. Jackson, J. Burnier, G. Quan, M. Stanley, J. Tom, J. A. Wells, *Science* **1994**, *266*, 243-247.

[19] Z. Machova, R. Eggelkraut-Gottanka, N. Wehofsky, F. Bordusa, A. G. Beck-Sickinger, "Expressed enzymatic ligation for the semisynthesis of chemically modified proteins." *Angew. Chem. Int. Ed.* **2003**, *42*, 4916-4918.

[20] D. S. Y. Yeo, R. Srinivasan, G. Y. J. Chen, S. Q. Yao, "Expanded utility of the native chemical ligation reaction." *Chem. Eur. J.* **2004**, *10*, 4664-4672.

[21] J. P. Tam, J. Xu, K. D. Eom, "Methods and strategies of peptide ligation." *Biopolymers* **2001**, *60*, 194-205.

[22] H. Hang, C. R. Bertozzi, "Chemoselective approaches to glycoprotein assembly." *Acc. Chem. Res.* **2001**, *34*, 727-736.

[23] J. P. Tam, Q. Yu, Z. Miao, "Orthogonal ligation strategies for peptide and protein." *Biopolymers* **1999**, *51*, 311-332.

[24] C. J. A. Wallace, "Peptide ligation and semisynthesis." *Curr. Opin. Biotechnol.* **1995**, *6*, 403-410.

[25] H. F. Gaertner, K. Rose, R. Cotton, D. Timms, R. Camble, R. E. Offord, "Construction of protein analogues by site-specific condensation of unprotected peptides." *Bioconjugate Chem.* **1992**, *3*, 262-268.

[26] K. Rose, "Facile synthesis of homogeneous artificial proteins." *J. Am. Chem. Soc.* **1994**, *116*, 30.

[27] G.G. Kochendoerfer, S. Y. Chen, F. Mao, S. Cressman, S. Traviglia, H. Shao, C. L. Hunter, D. W. Low, E. N. Cagle, M. Carnevali, V. Gueriguian, P. J. Keogh, H. Porter, S. M. Stratton, M. C. Wiedeke, J. Wilken, J. Tang, J.J. Levy, L. P. Miranda, M. M. Crnogorac, S. Kalbag, P. Botti, J. Schindler-Horvat, L. Savatski, J. W. Adamson, A. Kung, S.B.H. Kent, J. W. Bradburne, "Design and chemical synthesis of a homogeneous polymer-modified erythropoiesis protein." *Science* **2003**, *299*, 884-887.

[28] New technology by Novabiochem (EMD Biosciences): www.emdbiosciences.com

[29] C.-F. Liu, J. P. Tam, "Peptide segment ligation strategy without protecting groups." *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 6584-6588.

[30] M. Schnolzer, S.H.B. Kent, "Constructing proteins by dovetailing unprotected synthethic peptides: backbone-engineered HIV protease." *Science* **1992**, *256*, 221.

[31] D. S. Kemp, R. I. Carey, "Synthesis of a 39-peptide and a 25-peptide by thiol capture ligations: observation of a 40-fold rate acceleration of the intramolecular O,N-acyl transfer reaction between peptide fragments bearing onl cysteine protecting groups." *J. Org. Chem.* **1993**, *58*, 2216-2222.

[32] P.E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, "Synthesis of native proteins by chemical ligation." *Science* **1994**, *266*, 776-779.

[33] T. M. Hackeng, J. H. Griffin, P. E. Dawson, "Protein synthesis by native chemical ligation: expanded scope by using straightforward methodology." *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10068-10073.

[34] M. Villain, H. Gaertner, P. Botti, "Native chemical ligation with aspartic and glutamic acids as C-terminal residues: scope and limitations." *Eur. J. Org. Chem.* **2003**, 3267-3272.

[35] S. Kent, "Total chemical synthesis of enzymes." *J. Peptide Science* **2003**, *9*, 574-593.

[36] J. P. Tam, Y. A. Lu, C.-F. Liu, J. Shao, "Peptide synthesis using unprotected peptides through orthogonal coupling methods." *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 12485-12489.

[37] R. J. Hondal, B. L. Nilsson, R. T. Raines, "Selenocysteine in native chemical logation and expressed protein ligation." *J. Am. Chem. Soc.* **2001**, *123*, 5140-5141.

[38] P. Botti, M. R. Carrasco, S.B.H. Kent, "Native chemical ligation using removable N^{α}-(1-phenal-2-mercaptoethyl) auxiliaries." *Tetrahedron Lett.* **2001**, *42*, 1831-1833.

[39] J. Offer, C.N.C. Boddy, P. E. Dawson, "Extending synthethic access to proteins with a removable acyl transfer auxiliary." *J. Am. Chem. Soc.* **2002**, *124*, 4642-4647.

[40] C. Marinzi, J. Offer, R. Longhi, P. E. Dawson, "An *o*-nitrobenzyl scaffold for peptide ligation: synthesis and applications." *Bioorg. Med. Chem.* **2004**, *12*, 2749-2757.

[41] L. Z. Yan, P. E. Dawson, "Synthesis of peptides and proteins without cysteines residues by native chemical ligation combined with desulfurization." *J. Am. Chem. Soc.* **2001**, *123*, 526-533.

[42] E. Saxon, C. R. Bertozzi, "Cell surface engineering by a modified Staudinger reaction." Science **2000**, *287*, 2007-2010.

[43] M. Köhn, R. Breinbauer, "The Staudinger ligation – a gift to chemical biology." *Angew. Chem. Int. Ed.* **2004**, *43*, 3106-3116.

[44] E. Saxon, J. I. Armstrong, C. R. Bertozzi, "A traceless Staudinger ligation for the chemoselective synthesis of amide bonds." *Org. Lett.* **2000**, *2*, 2141-2143.

[45] B. L. Nilsson, L. L. Kiessling, R. T. Raines, "High-yielding Staudinger ligation of a phosphinothioester and azide to form a peptide." *Org. Lett.* **2001**, *3*, 9-12.

[46] J. A. Prescher, D. H. Dube, C. R. Bertozzi, "Chemical remodeling of cell surfaces in living animals." *Nature* **2004**, *430*, 873-877.

[47] V. V. Rostovtsev, L. G. Green, V. V. Folkin, K. B. Sharpless, "A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective ligation of azides and terminal alkynes." *Angew. Chem. Int. Ed.* **2002**, *41*, 2596-2599.

[48] H. C. Kolb, K. B. Sharpless, "The growing impact of click chemistry on drug discovery." *Drug Discovery Today* **2003**, *8*, 1128-1137.

[49] N. J. Agard, J. A. Prescher, C. R. Bertozzi, "A strain-promoted [3+2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems." *J. Am. Chem. Soc.* **2004**, *126*, 15046-15047.

[50] D. Bang, N. Chopra, S.B. H. Kent, "Total chemical synthesis of Crambin." *J. Am. Chem. Soc.* **2004**, *126*, 1377-1383.

[51] H. Paulus, "Protein splicing and related forms of protein autoprocessing." *Annu. Rev. Biochem.* **2000**, *69*, 447-496.

[52] C. J. Noren, J. Wang, F. B. Perler, "Dissecting the chemistry of protein splicing and its applications." *Angew. Chem. Int. Ed.* **2000**, *39*, 450-466.

[53] S. Chong, F. B. Mersha, D. G. Comb, M. E. Scott, D. Landry, L. M. Vence, F. B. Perler, J. Benner, R. B. Kucera, C. A. Hirvonen, J. J. Pelletier, H. Paulus, M. Q. Xu, "Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element." *Gene* **1997**, *192*, 271-281.

[54] T. W. Muir, D. Sondhi, P. A. Cole, "Expressed protein ligation : a general method for protein engineering." *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6705-6710.

[55] R. David, M. P. O. Richter, A. G. Beck-Sickinger, "Expressed Protein Ligation - methods and applications." *Eur. J. Biochem.* **2004**, *271*, 663-677.

[56] R. M. Hoffman, T. W. Muir, "Recent advances in the application of expressed protein ligation to protein engineering." *Curr. Opin. Biochem.* **2002**, *13*, 297-303.

[57] N. Budisa, "Adding new tools to the arsenal of expressed protein ligation." *ChemBioChem* **2004**, *5*, 1176-1179.

[58] Z. W. Zhang, B. A. Smith, L. Wang, A. Brock, C. Cho, P. G. Schultz, "A new strategy for the site-specific modification of proteins in vivo." *Biochemistry* **2003**, *42*, 6735-6746.

[59] A. Deiters, T. A. Cropp, M. Mukherij, J. W. Chin, J. C. Anderson, P. G. Schultz, "Adding amino acids with novel reactivity to the genetic code of *Saccharomyces Cerevisiae*." *J. Am. Chem. Soc.* **2003**, *125*, 11782-11783.

[60] K. L. Kiick, E. Saxon, D. A. Tirell, C. R. Bertozzi, "Incorporation of azides into recombinant proteins for chemoselective modification by Staudinger ligation." *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 19-24.

[61] H. Zhu, M. Bilgin, M. Snyder, "Proteomics" *Annu. Rev. Biochem.* **2003**, *72*, 783-812.

[62] S. Venjatasubbarao, "Microarrays – status and prospects" *Trends Biotechnol.* **2004**, *22*, 630-637.

[63] P. F. Predki, "Functional protein microarrays: ripe for discovery." *Curr. Opin. Chem. Biol.* **2004**, *8*, 8-13.

[64] D. S. Y. Yeo, R. C. Panicker, L. Tan, S. Q. Yao, "Strategies for immobilization of biomolecules in a microarray." *Comb. Chem. High Throughput Screening* **2004**, *7*, 213-221.

[65] Q. Xu, K. S. Lam, "Protein and chemical microarrays – powerfull tools for proteomics." *J. Biomed. Biotechnol.* **2003**, *5*, 257-266.

[66] Y. Lee, M. Mrksich, "Protein chips; from concept to practice." *Trends Biotechnol.* **2002**, *20*, s14-s18.

[67] M. L. Phelan, S. Nock, "Generation of bioreagents for protein chips." *Proteomics* **2003**, *3*, 2123-2134.

[68] H. Zhu, M. Bilgin, R. Bangham, D. Hall, A. Casamayor, P. Bertone, N. Lan, R. Jansen, S. Bildlingmaier, T. Houfek, T. Mitchell, P. Miller, R. A. Dean, M. Gerstein, M. Snyder, "Global analysis of protein activities using proteome chips." *Science* **2001**, *293*, 2101-2105.

[69] G. Macbeath, S. L. Schreiber, "Printing proteins as microarrays for high-throughput function determination." *Science* **2000**, *289*, 1760-1762.

[70] J. R. Falsey, M. Renil, S. Park, S. Li, K. S. Lam, "Peptide and small molecule microarray for high throughput cell adhesion and functional assays." *Bioconjug. Chem.* **2001**, *12*, 346-353.

[71] M. L. Lesaicherre, M. Uttamchandani, G. Y. J. Chen, S. Q. Yao, "Developing site-specific immobilization strategies of peptides in a microarray." *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2079-2083.

[72] M. B. Soellner, K. A. Dickson, B. L. Nilsson, R. T. Raines, "Site-specific protein immobilization by Staudinger ligation." *J. Am. Chem. Soc.* **2003**, *125*, 11790-11791.

[73] M. Köhn, R. Wacker, C. Peters, H. Schröder, L. Soulere, R. Breinbauer, C. M. Niemeyer, H. Waldmann, "Staudinger ligation: a new immobilization strategy for the preparation of small-molecules arrays." *Angew. Chem. Int. Ed.* **2003**, *42*, 5830-5834.

[74] M. C. Bryan, F. Fazio, H. Lee, C. Huang, A. Chang, M. Best, D. A. Calarese, O. Blixt, J. C. Paulson, D. Burton, I. A. Wilson, C.-H. Wong, "Covalent display of oligossacharide arrays in microtiter plates." *J. Am. Chem. Soc.* **2004**, *126*, 8640-8641.

[75] J. A. Camarero, Y. Kwon, M. A. Coleman, "Chemoselective attachment of biological active proteins to surfsces by expressed protein ligation and its application for protein chip fabrication." *J. Am. Chem. Soc.* **2004**, *126*, 14730-14731.

[76] D.S. Y. Yeo, R. Srinisavan, M. Uttamchandani, G. Y. J. Chen, Q. Zhu, S. Q. Yao, *Chem. Commun.* **2003**, 2870-2871.

[77] R. Y. P. Lue, G. Y. J. Chen, Y. Hu, Q. Zhu, S. Q. Yao, "Versatile protein biotinylation strategies for potential high-throughput proteomics." *J. Am. Chem. Soc.* **2004**, *126*, 1055-1062.

[78] M. Lovrinovic, R. Seidel, R. Wacker, H. Schroeder, O. Seitz, M. Engelhard, R. S. Goody, C. M. Niemeyer, "Synthesis of protein-nucleic acids conjugates by expressed protein ligation." *Chem. Commun.* **2003**, 822-823.

[79] D. C. Rideout, R. J. Breslow, "Hydrophobic acceleration of Diels-Alder reaction." *J. Am. Chem. Soc.* **1980**, *102*, 7817-1818.

[80] S. Otto, J. B. F. N. Engberts, "Diels-Alder reactions in water." *Pure Appl. Chem.* **2000**, *72*, 1365-1372.

[81] U. M. Lindström, "Stereoselective organic reactions in water." *Chem. Rev.* **2002**, *102*, 2751-2772.

[82] P. A. Grieco, *Organic Reactions in Aqueous Media*, Wiley, New York, **1997**.

[83] F. Fringuelli, O. Piermatti, F. Pizzo, L. Vaccaro, "Recent advances in Lewis acid catalyzed Diels-Alder reactions in aqueous media." *Eur. J. Org. Chem.* **2001**, 439-455.

[84] S. Otto, J. B. F. N. Engberts, "Hydrophobic interactions and chemical reactivity." *Org. Biomol. Chem.* **2003**, *1*, 2809-2820.

[85] K. W. Hill, J. Tauton-Rigby, J. D. Carter, E. Kropp, K. Vagle, W. Pieken, D. P. C. McGee, G. M. Husar, M. Leuck, D. J. Anziano, D. P. Sebesta, "Diels-Alder bioconjugation of diene-modified oligonucleotides." *J. Org. Chem.* **2001**, *66*, 5352-5358.

[86] G. M. Husar, D. J. Anziano, M. Leuck, D. P. Sebesta, "Covalent modification and surface immobilization of nucleic acids via the Diels-Alder bioconjugation method." *Nucleos. Nucleot. Nuc. Acids Res.* **2001**, *20*, 559-566.

[87] H. A. Latham-Timmons, A. Wolter, J. S. Roach, R. Giare, M. Leuck, "Novel method for the covalent immobilization of oligonucleotides via Diels-Alder bioconjugation." *Nucleos. Nucleot. Nuc. Acids Res.* **2003**, *22*, 1495-1497.

[88] Berkin, B. Coxon, V. Pozsgay, "Towards a synthethic glyconjugate vaccine against *Neisseria meningitidis* A." *Chem. Eur. J.* **2002**, *8*, 4424-4433.

[89] V. Pozsgay, N. E. Vieira, A. Yergey, "A method for bioconjugation of carbohydrates using Diels-Alder cycloaddition." *Org. Lett.* **2002**, *4*, 3191-3194.

[90] B. T. Houseman, M. Mrksich, "Carbohydrate arrays for the evaluation of protein binding and enzymatic modification." *Chem. Biol.* **2002**, *9*, 443-454.

[91] B. T. Houseman, J. H. Huh, S. J. Kron, M. Mrksich, "Peptide chips for the quantitative evaluation of protein kinase activity." *Nature Biotech.* **2002**, *20*, 270-274.

[92] W. Yeo, M. N. Yousaf, M. Mrksich, "Dynamic interfaces between cells and surfaces: electroactive substrates that sequentially release and attach cells." *J. Am. Chem. Soc.* **2003**, *125*, 14994-14995.

[93] W. S. Dillmore, M. N. Yousaf, M. Mrksich, "A photochemical method for patterning the immobilization of ligands and cells to self-assembled monolayers." *Langmuir* **2004**, *20*, 7223-7231.

[94] T. M. Tarasow, S. L. Tarasow, B. E. Eaton, "RNA Diels-Alderases: relationships between unique sequences and catalytic function." *J. Am. Chem. Soc.* **2000**, *122*, 1015-1021.

[95] O. Keller, J. Rudinger, "Preparation and some properties of maleimido acids and maleoyl derivatives of peptides." *Helv. Chim. Acta* **1975**, *58*, 62-63.

[96] G. W. Kenner, J. R. McDermott, R. C. Sheppard, "The safety catch principle in solid phase peptide synthesis" *J. Chem. Soc. Chem. Commun.* **1971**, 636.

[97] B. J. Brackes, J. A. Ellman, "An alkanesulfonamide safety-catch linker for solid-phase synthesis." *J. Org. Chem.* **1999**, *64*, 2322-2330.

[98] R. Ingenito, D. Drezniak, S. Guffker, H. Wenschuh, "Efficient loading of sulfonamide safety-catch linkers by amino acid fluorides." *Org. Lett.* **2002**, *4*, 1187-1188.

[99] A. Aletras, K. Barlos, D. Gatos, S. Koutsogianni, P. Mamos, "Preparation of the very acid-sensitive Fmoc-Lys(Mtt)-OH." *Int. J. Pep. Protein Res.* **1995**, *45*, 488-496.

[100] K. Itami, T. Nokami, J. Yoshida, "2-Pyridyldimethylsilyl group as a removable hydrophilic group in aqueous organic reaction: formation of molecular aggregates ad dramatic rate enhancement in Diels-Alder reactions." *Adv. Synth. Catal.* **2002**, *344*, 441-451.

[101] J. F. W. Keana, A. P. Guzikowski, C. Morat, J. J. Volwerk, "Detergents containing a 1,3-diene group in the hydrophobic segment. Facile chemical modification by Diels-Alder reaction with hydrophilic dienophiles in aqueous solution." *J. Org. Chem.* **1983**, *48*, 2661-2666.

[102] B. Seelig, A. Jäschke, "A small catalytic RNA motif with Diels-Alderase activity." Chem. Biol. **1999**, *6*, 167-176.

[103] S. Malean, P. Haynes, "Hydrogen migration in cyclopentadienes." *Tetrahedron* **1965**, *21*, 2329-2342.

[104] R. Breslow, Z. Zhu, "Quantitative antihydrophobic effects as probes for transition state structures of Diels-Alder reactions." *J. Am. Chem. Soc.* **1995**, *117*, 9923-9924.

[105] B. A. Katz, "Binding of biotin to streptavidin stabilizes intersubunit salt bridges between Asp61 and His87 at low pH." *J. Mol. Biol.* **1997**, *274*, 776-800.

[106] O. Livnah, E. A. Bayer, M. Wilchek, J. L. Sussman, "Three-dimensional structures of avidin and the avidin-biotin complex." *Proc Natl Acad Sci U S A* **1993**, *90*, 5076-5080.

[107] C. Anzczak, B. Bauvois, C. Monneret, J.-C. Florent, "A new acivin prodrug designes for tumor-targeted delivery." *Bioorg. Med. Chem.* **2001**, *9*, 2843-2848.

[108] N. M. Gree, "Spectrophotometric determination of avidin and biotin." *Meth. Enzymol.* **1970**, 18, 418-427.

[109] F. Schimmöller, I. Simmon, S. R. Pfeffer, "Rab GTPases, directors of vesicle docking." *J. Biol. Chem.* **1998**, *273*, 22161-22164.

[110] A. Iakovenko, E. Rostkova, E. Merzlyak, A. M. Hillebrand, N. H. Thomä, R. S. Goody, K. Alexandrov, *FEBS Lett.* **2000**, *468*, 155-158.

[111] K. Alexandrov, I. Heinemann, T. Durek, V. Sidorovitch, R. S. Goody, H. Waldmann, "Intein-mediated synthesis of geranylgeranyl Rab7 protein in vitro." *J. Am. Chem. Soc.* **2002**, *124*, 5648-5649.

[112] A. Rak, O. Pylypenko, T. Durek, A. Watzke, S. Kushnir, L. Brunsveld, H. Waldmann, R. S. Goody, K. Alexandrov. "Structure of Rab GDP-dissociation inhibitor in complex with prenylated YPT1 GTPase." *Science* **2003,** *302*, 646-650.

[113] T. Durek, K. Alexandrov, R. S. Goody, A. Hildebrand, I. Heinemann, H. Waldmann, "Synthesis of fluorescently labeled mono- and diprenylated Rab7 GTPase." *J. Am. Chem. Soc.* **2004**, ASAP.

[114] A. Rak, O. Pylypenko, A. Niculae, K. Pyatkov, R. S. Goody, K. Alexandrov**,** "Structure of the Rab7:Rep-1 complex: Insights into the mechanism of Rab prenylation and choroideremia disease" *Cell* **2004**, *117*, 749-760.

[115] G. L. Ellman, "Tissue sulfhydryl groups." *Arch. Biochem. Biophys.* **1959**, *82*, 70-77.

[116] Chimera Biotec GmbH, Emil-Figge-Strasse 76a, 44227 Dortmund, Germany.

[117] R. Benters, C. M. Niemeyer, D. Wöhrle, "Dendrimer-activated solid supports for nucleic acid and protein microarrays" *ChemBioChem* **2001**, *2*, 686-694.

[118] R. Benters, C. M. Niemeyer, D. Drustschmann, D. Blohm, D. Wöhrle, "DNA microarrays with PAMAM dendritic linker system" *Nucleic Acids Res.* **2002**, *30*, e10.

[119] M. Bradford, "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." *Anal. Biochem.* **1976**, *72*, 248-254.

8. ABBREVIATIONS

AA	Amino acid
ACA	Aminocaprionic acid
APS	Ammonium persulfate
Ar	Aromatic
calc.	calculated
CBD	Chitin binding domain
CDI	1,1´-Carbonyl-diimidazole
CHAPS	3-(3-cholamidopropyl)-dimethyl-ammoniopropane sulfonate
cHex	Cyclohexane
Da	Dalton
DA	Diels-Alder
Dan	Dansyl
Dansyl	5-Dimethylaminonaphtalene-1-sulfonyl
DCM	Dichloromethane
DIC	N,N´-Diisopropylcarbodiimide
DIPEA	Diisopropylethylamine
DMAP	4-Dimethylaminopyridine
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DTNB	5,5´-dithio-bis-nitrobenzoic acid
DTT	Dithiolthreitol
DVD	Divinylbenzene
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride
EDT	Ethanedithiol
EPL	Expressed protein ligation
Eq.	Equivalent
ESI	Electron-spray inonization
FAB	Fast atom bombardment
Fmoc	9-Fluorenylmethoxycarbonyl

GC	Gas chromatography
GDP	Guanosine 5´-diphosphate
GNP	Phosphoaminophosphonic acid guanylate ester
HABA	2-(4-Hydroxyphenylazo)benzoic acid
HBTU	2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	Human immunodeficiency virus
HOBt	*N*-Hydroxybenzotriazole
HPLC	High performance liquid chromatography
Hxdn	2,4-Hexadienyl
K_d	Dissociation constant
LC-MS	Liquid chromatography – mass spectroscopy
MALDI-TOF	Matrix-assisted laser desorption ionization time-of-flight
ME	Mercaptoethanol
MESNA	2-mercapto-ethane sulfonate
MS	Mass spectrometry
Mtt	4-Methyltrityl
MW	Molecular weight
MWCO	Molecular weight cut-off
NCL	Native chemical ligation
NHS	N-Hydroxysuccinimide
NMP	*N*-Methylpyrrolidinone
NMR	Nuclear magnetic ressonance
NOE	Nuclear Overhauser effect
NTA	Nitrilotriacetic acid
o.n.	Overnight
OHxdn	2,4-Hexadienyl ester
PAGE	Polyacrylamide gel electrophoresis
PG	Protecting group
PyBOP	Benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluophosphate
Quant.	Quantitative
Rab	Ras-like protein from rat brain
RP	Reversed-phase
SDS	Sodium dodecylsulfate

SPPS	Solid-phase peptide synthesis
S*t*Bu	Thio-*tert*-butyl
***t*Bu**	*tert*-Butyl
TEMED	N,N,N',N'-Tetramethylethylenediamine
TES	Triethylsilane
TFA	Trifluoracetic acid
TGF	Transforming growth factor
THF	Tetrahydrofuran
TIS	Triisopropylsilane
TLC	Thin layer chromatography
TNB	5-thio-2-nitrobenzoic acid
Tris	Tris-hydroxymethyl-aminomethane
UV	Ultraviolet
wt	Wild type
GTP	Guanosine 5´-triphosphate
Na-Pi	Sodium phosphate buffer

Acknowledgements

I would like to express my sincerest thanks to the following people whose valuable assistance made possible the conclusion of this PhD work.

First I would like to thank my supervisor Prof. Herbert Waldmann for his support and constructive guidance and for providing all the necessary tools for the development of this work.

I would like to thank the International Max-Planck Institute Research School of Chemical Biology for financial support, in special to its organizers Dr. Jutta Rötter and Prof. Rolf Kinne for their friendly assistance whenever I needed.

I am grateful to Dr. Kirill Alexandrov for providing the Rab proteins that were used in this work and laboratory space for the performance of the protein ligation experiments. I also thank the people from Department of Physical Biochemistry at the MPI for helping me with the Rab ligation experiments, in special to Mahesh Kulharia.

My thanks also go to Dr. José Palomo, Maja Köhn and Dr. Ron Wacker for teaming up with me for the preparation of the avidin arrays, and to Chimera Biotec GmbH for providing the glass slides and analytical facilities for the execution of these experiments.

I would like also to express my gratitude to my colleagues at the Department of Chemical Biology who have contributed for a cooperative working environment along these last four years, in special to Markus Westhus and Stephanie Winkler for a pleasant lab atmosphere.

I am very much thankful to Dr. Joaquín Gomis and Dr. Ivan Corrêa Jr. their constructive comments on this manuscript, and Matthias Mentel for helping out with the German translation of the summary.

My deepest gratitude goes also to Dr. Miguel Sanz and Dr. Michele Leuenberger for their friendship, constant support and with whom I enjoyed valuable discussions about my work.

My warmest thanks go to Ivan Correa Jr., Sukanya Nad and the Spanish gang: José, Joaquín, Ana, Ester, Maria and Marta; for their encouragement at the last stage of my PhD work.

Back in time, starting from my earlier days of undergraduate studies until the beginning of my PhD work, two persons were very important in helping me to search for my goals: Dr. Joachim Demnitz and Dr. Luigi Lavecchia, thanks for your support. Without you I would not have achieved this far.

I want to express my thanks to my dear friends Priscila Sfredo, Peter Signore and Erica Hott with whom I shared so many great moments during my stay in Dortmund. Furthermore I would like to thank my friend Adriana de Souza for being always there for me here, in Switzerland, in Brazil, whenever I go. Thanks also to Bárbara Ramos, for your never-end motivation even from the other side of the world.

Finally, I would like to express my greatest gratitude to my family. Seu Fila, Dona Edina, and Mana Elisângela thanks for your love, help and guidance along my 26 years of life. You have given me the heart and the mind to start with.

Obrigada!!!

Aline

Curriculum vitae

Personal Information

Name:	Aline Dantas de Araújo
Date of birth:	06.10.1978 in Clevelândia (Brazil)
Nationality:	Brazilian

Education:

1985-1992	Primary school in Porto Alegre and Recife (Brazil)
02/1993 - 12/1996	Technical High School in Industrial Chemistry at Escola Técnica Federal de Pernambuco in Recife (Brazil)

Undergraduate Study

02/1996 - 01/2000	B.Sc. in Chemistry at Universidade Federal de Pernambuco in Recife (Brazil)
06/1999 - 12/1999	Diploma work under supervision of Prof. Joachim Demnitz at the Universidade Federal de Pernambuco, with the theme: "Synthesis of β-hidroxy esters by electrochemical reduction of isoxazolines"
02/2000 - 11/2000	Research training at Novartis Pharma in Basel (Switzerland) under supervision of Dr. Luigi Lavecchia, with the theme: "Asymmetric synthesis of β-aminoacids and γ-aminoalcohols by ozonolysis of chiral homoallylamines"

Doctoral Study

02/2001 - 02/2005	PhD work at Max-Planck Institut für molekulare Physiologie in Dortmund (Germany) under supervision of Prof. Herbert Waldmann, with the theme: "Diels-Alder reaction as a new tool for protein functionalization"
04/2002 - 12/2004	PhD studies at the International Max-Planck Research School in Chemical Biology in Dortmund

Dortmund, in February 2005