BIOLOGY AND CHEMISTRY APPROACHES TOWARD SITE-DIRECTED PROTEIN MODIFICATIONS AND A NOVEL SYNTHETIC METHOD TO RHODAMINE B DERIVATIVES FOR PROTEIN LABELING

CHEN XI

NATIONAL UNIVERSITY OF SINGAPORE

-2010-

BIOLOGY AND CHEMISTRY APPROACHES TOWARD SITE-DIRECTED PROTEIN MODIFICATIONS AND A NOVEL SYNTHETIC METHOD TO RHODAMINE B DERIVATIVES FOR PROTEIN LABELING

CHEN XI

(B.Sc., SCU)

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF CHEMISTRY

NATIONAL UNIVERSITY OF SINGAPORE

2010

ACKNOWLEDGEMENT

My utmost gratitude goes to my thesis supervisor—Prof. Tanja Weil for her comprehensive knowledge, her creativity, her discussion and her consistent encouragement!

Many thanks also give to my parents and my brother without whom I would never have been able to achieve so much!

Next, I have to express my appreciation to Madam Han Yanhui for teaching and kindly helping in NMR techniques and problems.

I would like to appreciate my friends also my lab members—Wu Yuzhou, Kuan Seahling, Wang Tao, Goutam Pramanik, Archana Mishra, David Ng as well as my German colleagues Joerg Wilhelmi and Anne Pfister for our scientific discussions!

Lastly, I should give my thanks to my close partner—Kasturi Mothossamy for her great support in the past eight months!

TABLE OF CONTENTS

SUMMARY

The modification of proteins and the further decoration of them using reactive chromophores represent a recent topic of scientific interest. In this thesis, novel approaches toward site-directed protein modifications will be explored *via* either biology or chemistry means. In conjunction with those novel protein modification approaches, reactive rhodamine B dyes were prepared as useful chromophores to decorate those functionalized proteins and peptide.

First of all, a biology method termed *selective pressure incorporation* (SPI) was developed to incorporate non-canonical amino acid (NCAA) into proteins in a site-directed fashion as exemplified by the incorporation of azidohomoalanine (AHA) into ribonuclease A (RNase A) in this thesis. AHA which is an analogue of methionine was synthesized with a considerably enhanced yield compared to recently reported protocols. This NCAA was efficiently incorporated into the protein RNase A *via* SPI, thus giving rise to a site-directed bi-azido functionalized RNase variant with two surface AHA located at AHA1 and AHA41 respectively.

Two chemistry approaches toward site-directed protein modifications were also established that target cysteine (Cys) and lysine (Lys) respectively. The first cysteine-targeting approach requires the protein substrates to bare a single accessible cysteine residue on its surface, such as serum albumin proteins, antibody monomers and *etc.*. Compared to this method, the second chemistry approach which addresses lysine residue displays a wider application scope as

lysine typically exists on most protein surfaces. The solvent accessibility of each lysine residue as well as the interaction between lysine and other residues in close vicinity play an important role to achieve the site-directionality. The success of this work was demonstrated by the mono-biotinlylation of RNase A, lysozyme C and somatostatin (SST) using a novel bioconjugation reagent—biotin-PEG-NHS (**49**).

The last part of this thesis is dedicated to the preparation of functionalized rhodamine B dyes as reactive chromophore to label functional protein/peptides, such as azido-antibody and SST. This is a one step approach toward functionalized rhodamine B dyes *via* Steglich esterification. Compared to the previously reported protocol, this route is more practical, concise and avoids the usage of highly reactive reagents. Aside from rhodamine B ethyl ester, three reactive rhodamine B derivatives were prepared. They are propargyl rhodamine B ester for targeting azido group, azidoethyl rhodamine B ester for targeting ethynyl group and monosulfone-OH rhodamine B ester for targeting accessible disulfide bonds. Finally, their applicability for labeling corresponding functionalized protein/peptide is demonstrated.

LST OF FIGURES

LIST OF SCHEMES

LIST OF SYMBOLS

1. **AB:** Antibody

2. **AHA:** Azidohomoalanine

3. **Boc:** *tert*-Butoxycarbonyl

4. **Boc-AHA:** *N*-Boc azidohomoalanine

5. **Boc-Dab:** *N*-Boc-2,4-diaminobutyric acid

6. **BPTI:** Bovine pancreatic tyrosine inhibitor

7. **CalB:** *Candida antarctica* lipase B

8. **CGP:** 2'-Deoxycytidine-2'-deoxyguanosine-3'-5'-monophosphate

9. **CuAAC:** Copper catalyzed azide-alkyne cycloaddition, also termed Huisgen 1,3-dipolar

    cycloaddition

10. **Cys:** Cysteine

11. **DCC:** *N,N'*-Dicyclohexylcarbodiimide

12. **DCU**: *N,N'*-Dicyclohexylurea

13. **DCM:** Dichloromethane

14. **DI:** Deionised

15. **DIEA:** *N,N*-Diisopropylethylamine

16. **DMAP:** 4-Dimethylaminopyridine

17. **DMF:** *N,N*-Dimethylformamide

18. **DMSO:** Dimethylsulfoxide

19. **E.** *coli***:** Escherichia. *coli*

20. **EDC:** 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide

21. **Equiv.:** Equivalent

22. **ESI:** Electron spray ionization

23. **FCS:** Fluorescence correlation spectroscopy

24. **FRET:** Fluorescence resonance energy transfer

25. **F.W.:** Formula weight

26. **Gly:** Glycine

27. **HRMS:** High resolution mass spectroscopy

28. **Hz:** Hertz

29. **IR:** Infrared spectroscopy

30. **K:** Lysine

31. **kD:** Kilo Dalton

32. **LC:** High performance liquid chromatography

33. **Lys:** Lysine

34. **MALDI-ToF-MS:** matrix assisted laser desorption ionisation time-of-flight mass spectrometry

35. **MOE:** Molecular operation environment

36. **MQ-H$_2$O:** Milli-Q water

37. **M.W.:** Molecular weight

38. **MWCO:** Molecular weight cut-off

39. **NHS:** *N*-Hydroxysuccinimidyl (ester)/ *N*-hydroxysuccinimide

40. **NMR:** Nuclear magnetic resonance

41. **NCAA:** Non-canonical amino acid

42. **N-Valine:** (R,S)-2-Amino-2-(dimethylamino)acetic acid

43. **PBS:** Phosphate buffered saline

44. **PEG:** Poly(ethylene oxide)

45. **PEGylation:** Convalent conjugation of poly(ethylene oxide)

46. **PPL:** Pyrodoxal-5-phosphate

47. **PreHPLC:** Preparative high performance liquid chromatography

48. **RB:** Rhodamine B

49. **RBF:** Round-bottomed flask

50. **RNase A:** Ribonuclease A

51. **RT:** Room temperature

52. **SA:** Sinapinic acid

53. **SPI:** Selective pressure incorporation

54. **SST:** Somatostatin

55. **TCEP:** *tris*(2-carboxyethyl)phosphine

56. **TFA:** trifluoroacetic acid

57. **TG II:** Transglutaminase

58. **UV:** Ultraviolet-visible spectroscopy

Chapter 1. INTRODUCTION

Modifications of proteins in a site-directed way and labeling proteins with chromophore(s) at distinct site(s) gained enormous interest from both biologists and chemists due to their attractive applications.[1~11] The dynamics of a protein can be investigated via fluorescence correlation spectroscopy (FCS) after attachment of a fluorescent chromophore.[1,2,3] If there are two photo-stable fluorophores located at two distinct positions on a protein within a proper range, conformational changes of a protein in solution such as the mechanism of proteolysis of an enzyme could be studied e.g. via fluorescence resonance energy transfer (FRET).[4,5,6] Conjugation of poly(ethylene oxide) chains (PEGylation) onto therapeutic proteins alters their pharmacokinetic profile and increases their half-life time in the body thus improving their therapeutic effects.[7,8,9,10] Furthermore, decoration of proteins with polymer chains in a defined manner could generate novel protein biohybrids with tailored properties different from the native ones.[11] Therefore, developing novel protein modification strategies that allow site-directed modification reactions on a protein surface and synthesis of novel fluorophores are highly appealing.

In the following, major published approaches for protein modification will be reviewed. Firstly, the biological method, *selective pressure incorporation* (SPI), which utilizes reactive non-canonical amino acids (NCAA) will be briefly introduced. Thereafter, different chemistry approaches for protein modifications will be discussed as well. In conjunction with protein modification, prominent functionalization strategies of the chromophore rhodamine B as useful protein decoration reagents will be summarized. Unfortunately, all these methods suffer from different limitations and the challenges behind them will be discussed as well. This thesis is dedicated in providing some feasible solutions to each problem and they will be discussed in detail in the *Results and Discussion* chapter.

1.1. **Protein Modification via *Selective Pressure Incorporation* (SPI)**

1.1.1. BACKGROUND

Living cells utilize 21 naturally occurring amino acids (aside from a few exceptions) as building blocks to produce a large variety of proteins. This process includes two major steps termed transcription and translation. Usually, the cell's own set of translational machineries work cooperatively to ensure the two processes proceed correctly.

However, it has been found that in living cells, the translational machineries can be manipulated in order to accept synthetic amino acids that are not produced by living organisms.[18] This discovery opened one possibility: certain cells can recognize exogenous natural amino acids analogues based on which novel proteins are produced that do not exist in nature. This new research branch was explored by several groups and significant progress was made during the past few years. The first approach, *nonsense codon suppression*, was developed by Schultz *et al*.[19,20,21,22] using a unique codon and a corresponding tRNA: aminoacyl-tRNA-synthetase pair. Via this method, NCAA can be inserted into a protein at a distinct position(s) either in prokaryotic or eukaryotic cells. However, this technique is complicated and technically demanding. Another strategy termed s*elective pressure incorporation* (SPI) was first developed by Tirrel *et al*.[23,24,25,26] This strategy makes use of auxotrophic E. *coli* strains that are not able to produce a certain amino acid and which can only grow if the nutrient media contains this amino acid. In the case that a very close analogue of the natural amino acid is added, and the cell cannot distinguish between the natural amino acid and the structurally close analogues during the translation process, the NCAA is inserted into the protein codon. As a result, a recombinant protein containing NCAA is produced. This biological method is generally considered less challenging as compared to *nonsense codon suppression*. This is mainly because this approach does not require the development of a new transfer-RNA: aminoacyl-tRNA-synthetase pair.

1.1.2. AN INTRODUCTION TO SPI

The SPI process is briefly described in figure 1 using a methionine auxotrophic E. *coli* strain for AHA-protein expression as an example. First of all, an auxotrophic E. *coli* strain is applied which is unable to self-generate methionine for growing. Subsequently, an engineered plasmid that can express a certain protein of interest is introduced into E. *coli* cells. Then, the E. *coli* cells are fed with a medium containing AHA, a close analogue of the natural amino acid methionine which allows them to grow and insert AHA into the protein sequence instead of methionine. This process is called *selective pressure incorporation* of AHA. When the growth of E. *coli* ceases, the cell culture is subjected to cell lyses in order to release the expressed AHA-protein.
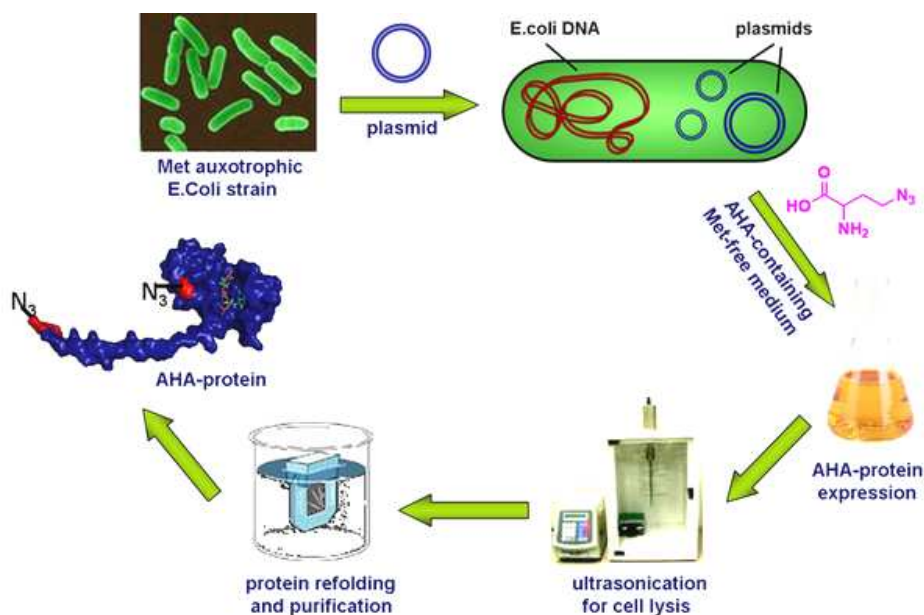
Figure1. The general steps that are involved in the incorporation of a NCAA (e.g. AHA (**2**)) into a protein (e.g. RNase A) *via* SPI.

1.1.3. TRANSLATIONALLY ACTIVE NCAAs IN SPI

In the past, several NCAAs were incorporated into a few proteins by the SPI approach. Using methionine auxotrophic bacteria strains, propargylalanine (**1**)[23] and AHA (**2**)[24] were efficiently incorporated into several proteins. Those proteins are attractive since the presence of azido or ethynyl groups allows for subsequent click reactions. For instance, recombinant proteins bearing azido groups could be used for site-specific Staudinger ligation by reacting with appropriate triarylphosphine reagents.[25] In addition, 4-ethynylphenylalanine (**4**),

4-azidophenylalanine (**5**) and 4-iodophenylalanine (**6**) were found to replace phenylalanine when using phenylalanine auxotrophic bacteria strains.[27] Regardless of the diversity of NCAA, only a few recombinant proteins were successfully expressed by this approach such as AHA-CalB, AHA-RNase A.[28] and ketone-T4L lysozyme.[12] Some typical NCAA that have been successfully applied to protein expression and that are translationally active are listed below (Figure 2).

a. methionine analogues
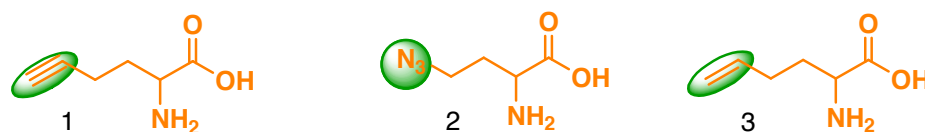
b. phenylalanine analogues
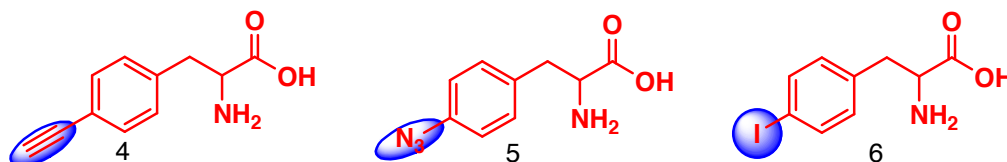
Figure2 a-b. Some typical translationally active NCAAs in SPI including: a. methionine analogues (propargylhomoglycine (**1**), AHA (**2**) and allylalanine (**3**)). b. phenylalanine analogues (4-ethynylpehynylalanine (**4**), 4-azidophenylalanine (**4**) and 4-iodophenylalanine (**6**)).

1.1.4. LIMITATIONS OF SPI

Although SPI is an excellent technique for producing proteins modified in a site-directed fashion, it suffers from several limitations. The major limitation is that one NCAA would replace all its corresponding native residues during protein expression thus affording proteins with multiple reactive NCAAs. If mono-functionalization is a key concern, one way to circumvent this limitation is via the selection of a proper residue that is the only one exposed on the protein surface. This strategy has been applied to the expression of AHA-CalB *etc*. Since methionine is often a rare residue and is usually buried inside the protein scaffold, it is often selected for preparing mono-functionalized proteins[28]. As for lysozyme, only one methionine out of five is located on the surface. Incorporation of AHA afforded AHA-CalB with only one accessible AHA on its surface[27]. The structure of CalB and the distribution of methionine in CalB are shown in figure 3. In case methionine is not present on the protein

surface, introduction of point mutation will be necessary.

Figure3. The crystal structure of CalB where all methionine residues are indicated by red sticks except the one located on the surface highlighted by yellow spheres (PDB file: 1tca).

1.1.5. SYNTHESIS of NCAAs

The development of synthetic strategies in order to achieve NCAA in high yields represents a challenge for chemists mainly due to the structure diversity of various NCAAs. In order to achieve sufficient quantities of a protein of interest bearing NCAA, usually gram scale quantities of the respective NCAA is required. In the following, the most prominent ways to prepare reactive NCAAs are summarized.

In order to introduce an azido group as present in AHA, there are two major routes that involve either diazotransfer or nucleophilic substitution reactions by an appropriate azide-reagent (Scheme 1). Triflic azide has been found to act as the initial successful diazo transfer reagent and has been mainly used until now and most of the recently reported synthetic routes to AHA still employ this reagent.[13,14,15] However, it also suffers from several severe limitations such as its highly explosive nature upon shaking, heating and concentrating, which hampers upscaling. In addition, triflic azide is toxic and decomposes quickly during storage. These drawbacks led to the development of the new diazotransfer reagent—imidazole-1-sulfonyl azide hydrochloride.[16] Compared to triflic azide, this compound is not explosive and can be stored for months without obvious loss of efficacy. In the second method, the NCAA precursor is treated with sodium azide. The two protocols are

compared below (Scheme 1).

a. *via* diazo-transfer
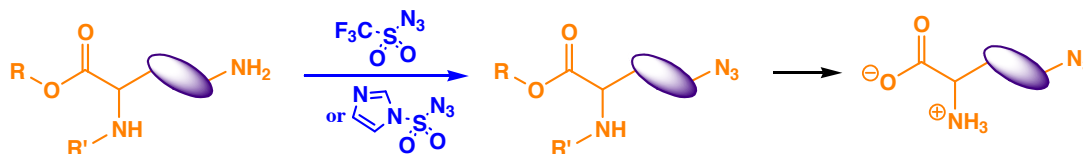
b. *via* nucleophilic substitution

X = Br, I, TsO, *et al.*
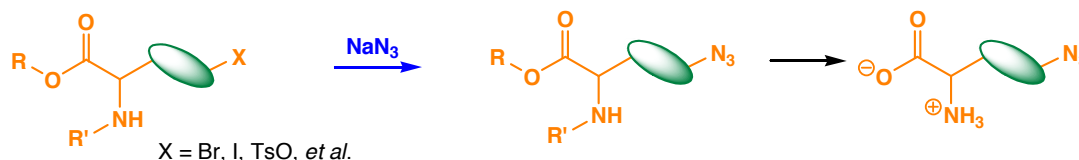
Scheme1 a-b. The two major routes toward the preparation of azido-containing NCAAs are via either diazotransfer (a) or nucleophilic substitution using sodium azide (b).

The Sonogashira coupling reaction represents a powerful approach to introduce ethynyl groups onto phenyl rings as demonstrated by the synthesis of 4-ethynylphenylalanine (**4**) (Scheme 2a).[17] Ethynyl-containing aliphatic amino acid can be prepared via Strecker synthesis using a ethynyl-containing precursor. For instance, after treatment with ammonium and sodium cyanide followed by hydrolysis, pent-4-ynal (**7**) is converted into propargylhomoglycine (**12**) (Scheme 2b).

a. *via* Sonogashira coupling
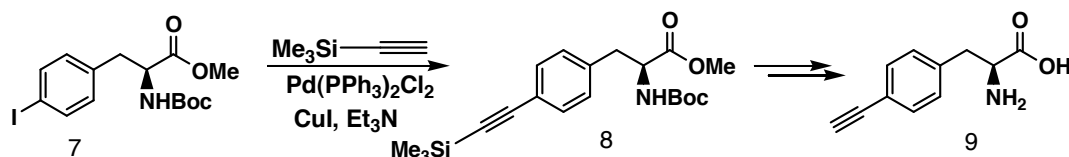
b. *via* Strecker synthesis
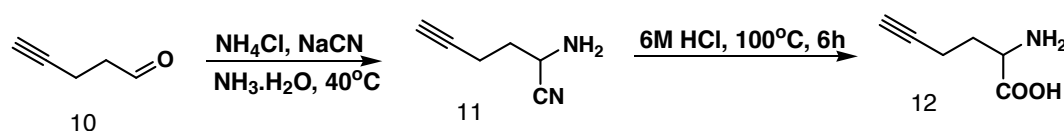
Scheme2 a-b. Two routes toward the synthesis of ethynyl-containing NCAAs. [45]

One important consideration of this synthesis scheme is related to the stereochemistry of the α-carbon of the NCAA. It has been demonstrated by experiments that racemic NCAA such as

(D,L)-AHA can be effectively incorporated into the amino acid codon so that no chiral separation of the racemic mixture is necessary. This is because the cell's translational machinery can selectively recognize L-NCAA. The D-amino acid will not affect protein expression but will remain in the culture solution.

1.2. Chemical Protein Modifications

In the following, chemistry approaches toward protein functionalization will be highlighted. The most prominent reaction schemes to functionalize proteins targeting cysteine or lysine substituents since they either contain the reactive thiol group or the primary amino group. Unfortunately, these two approaches usually yield a heterogeneous protein mixture. For instance, PEGylation strategies to improve the properties of therapeutic proteins targeting lysine often afford products with a broad degree of modification. These limitations motivate the discovery of new site-directed protein modification strategies that only yield homogeneously functionalized proteins. In the following sections, recently developed protein modification strategies, both site-directed and none site-directed, will be reviewed.

1.2.1. RESIDUE-SELECTIVE PROTEIN MODIFICATIONS

Some protein modification strategies are able to address the side chains of certain amino acids on proteins without interfering with other amino acid residues. These modification methods are called *residue-selective*. In this subsection, the two most recently developed modification techniques are summarized.

1.2.1.1. Tryptophan-Targeting Protein Modifications

Tryptophan is considered an inert amino acid due to the absence of very reactive functional groups in its side chain. However, in the presence of transition metal catalysts, the aromatic ring system of the tryptophan side chain can be activated and participate in certain reactions. Previously, rhodium carbenoids were found to efficiently react with tryptophan residue[29] (Scheme 3). Compared to conventional lysine and cysteine modifications, targeting

tryptophan is attractive due to its low abundance on the protein surface. Therefore, this strategy generally shows a higher specificity toward achieving a defined product and further more, it only requires low protein concentrations at the micromolar level.
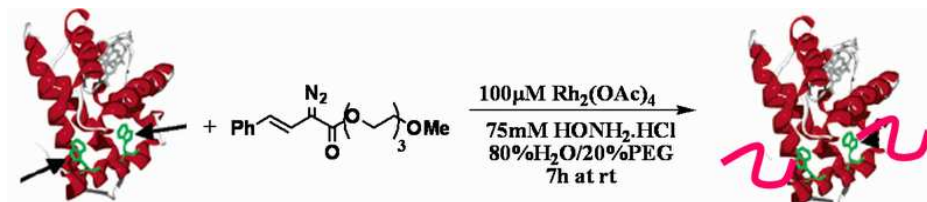
Scheme3. Tryptophan-selective protein modification using rhodium carbenoids.[29]

1.2.1.2. Tyrosine-Targeting Protein Modifications

The aromatic amino acid tyrosine can also be targeted due to its reactive phenol side chain. As reported by Francis, π-allylpalladium complexes were used to address tyrosine residues in a selective fashion[30] by using palladium acetate. During this reaction, the π-allylpalladium complex was formed and subsequently reacted with the protein of interest. One characteristic feature of this method includes the cleavage of a "disposable group" upon the allyl-modification. This strategy was successfully applied to attach a lipophilic C17-allyl chain onto the surface of myoglobin (Scheme 4).
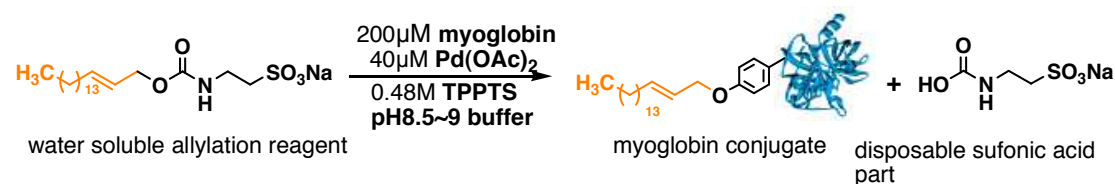
Scheme4. Attachment of a lipophilic allyl chain onto myoglobin by targeting tyrosine residues via a water soluble allylation reagent.[30]

1.2.2. SITE-DIRECTED CHEMICAL PROTEIN MODIFICATIONS

1.2.2.1. Targeting a Cysteine via Maleimide Reagents

Some proteins bear only one thiol group at their surface such as BSA, HSA and some antibody monomers. For example, BSA only has a single cysteine 34 on its surface; others are paired up to form disulfide bonds and are buried inside. Upon modification with a maleimide reagent, Cys 34 could be site-specifically targeted. Nevertheless, this approach

usually affords a mixture of modified and unmodified protein since the conversion is not quantitative. This problem needs to be circumvented if we wish to prepare a pure modified protein via this approach.

1.2.2.2. Targeting a Disulfide-Bond via Mono-/Bis-Sulfone Intercalation Reagents

Recently, another method was established by Brocchini *et al.*[31] using disulfide intercalation reagent containing mono-/bis-sulfone groups. Site-specific modification has been achieved in such cases when there is only one accessible disulfide bond on the protein surface. This approach is based on the reduction of the accessible disulfide bridge and the reaction of the thiol groups with an appropriate mono-sulfone reagent via a double Michael addition mechanism to afford a bridged disulfide bond (Scheme 5). The major advantage of this scheme is the mild reaction condition and therefore, the protein's tertiary structure and biological activity was maintained in most cases. This approach has been successfully used for site-specific PEGylation of cytokines, enzymes, antibody fragments and peptides[31]. A computational approach for the identification of the accessible disulfide bonds and prediction of the structural effects of such modifications are also available.[32]
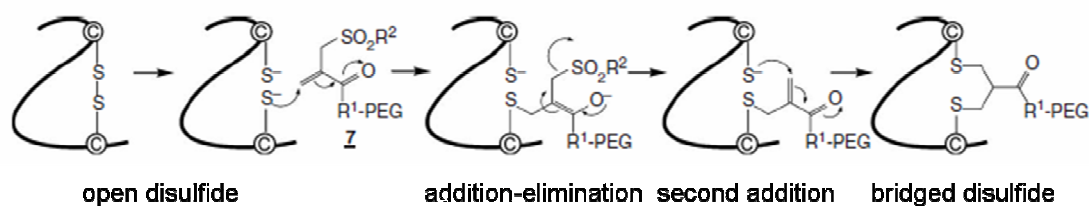
Figure4. Site-specific PEGylation of proteins using PEG-monosulfone reagent.[31]

1.2.2.3. *N*-Terminal Modification of Proteins via Transamination

In 2006, Gilmore *et al.* developed a relatively universal approach by targeting the *N*-terminal residue of proteins.[33] This method utilizes the unique reactivity of the *N*-terminal amine upon transamination. Therefore, *N*-terminal amino groups can be preferentially converted into a carbonyl group which can be afterwards conjugated with alkylhydrazine or alkoxyamine. The most efficient transamination reagent is pyrodoxal-5-phosphate (PPL). The two-step

modification process is shown below (Scheme 5).
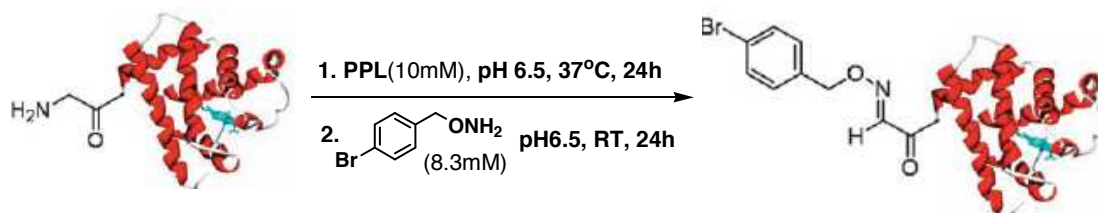
Scheme5. *N*-Terminal modification of proteins via transamination.[33]

Since most proteins possess a free *N*-terminal amino group, this approach benefits from a broad application scope. However, there also exist some limitations. For example, this method is not suitable for proteins with *N*-terminal serine, cysteine, threonine or tryptophan residues due to their known side reactions with aldehydes/ketones. In addition, mixtures of native and modified proteins are generated and the separation of them is challenging due to their close molecular weight (M.W.). Furthermore, most of the proteins which were studied gave poor conversion ratios.[33]

1.3. Enzyme-Mediated Site-Directed Protein Modification

Application of post-translational enzymes in order to achieve site-directed protein modifications is also one promising way. This possibility has been explored by using transglutaminase II (TG II) to decorate RNase A and bovine pancreatic tyrosine inhibitor (BPTI) .[34] TG II is a posttranslational enzyme which has a remarkable specificity toward glutamine (Q) and lysine (K). Using dansyl-*ε*-aminocaproyl-GlnGlnIleVal as substrate, TG II specifically attached this fluorescent Q probe onto the K1 residue of RNase A and to the K15 residue of BPTI. Interestingly, no K residues on lysozyme C could be modified using TG2 (Figure 5) which indicates that this approach somehow lacks predictability.
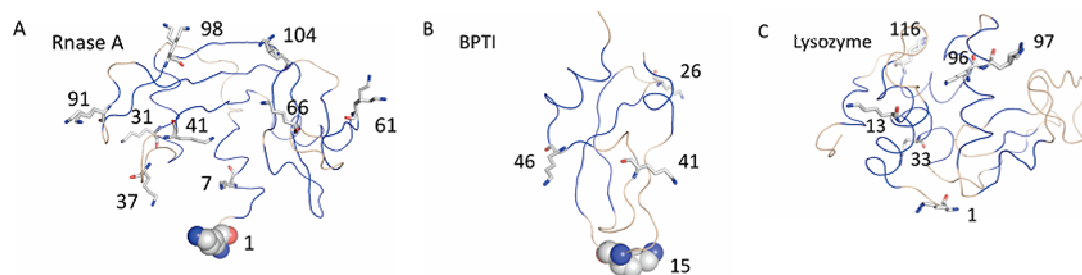
Figure5 a-c. K1 of RNase A and K15 of BPTI were selectively modified while no modification occurs on lysozyme under the catalysis of TG II.[34]

1.4. Preparation of Reactive Rhodamine Dyes as Useful Chromophores for Protein Decoration

In this section, functional rhodamine dyes that can be applied to decorate proteins and impart them with new functionalities will be discussed. The attachment of chromophores generally represents a straight-forward method to quickly assess how many substituents have been attached during a novel reaction scheme. In addition, chromophores have gained much interest for sophisticated dynamics studies of biomolecules in solution e.g. via FCS[18,19] and FRET[22,23] studies, for the development of chemosensors and cell imaging. Furthermore, chromophore-labeled proteins allow investigation of their dynamics in solution at the single molecule level[20], thus leading to an improved insight of e.g. structural rearrangements during enzyme catalysis or protein-protein interactions.

Rhodamine B is one of the most widely used chromophores since it is cheap, water soluble and relatively photostable. However, derivatization strategies are still limited and there are only a few, very costly derivatives commercially available.

1.4.1 FUNCTIONALIZATION OF RHODAMINE VIA RHODAMINE-PIPERAZIINE AMIDE INTERMEDIATES

Rhodamine dyes are ionic chromophores which are prone to intramolecular cyclization under basic conditions. These features render the derivatization of rhodamines challenging. One

approach to address this problem is to convert rhodamines into rhodamine-piperazine amides intermediate, which can be easily derivatized further afterwards (Scheme 6). Via this method, several functional groups including carboxylic, hydroxyl, *N*-hydroxysuccinimidyl (NHS), 2-chloroacetyl, 2-iodoacetyl and 2-thiolacetyl group have been introduced (Scheme 6).
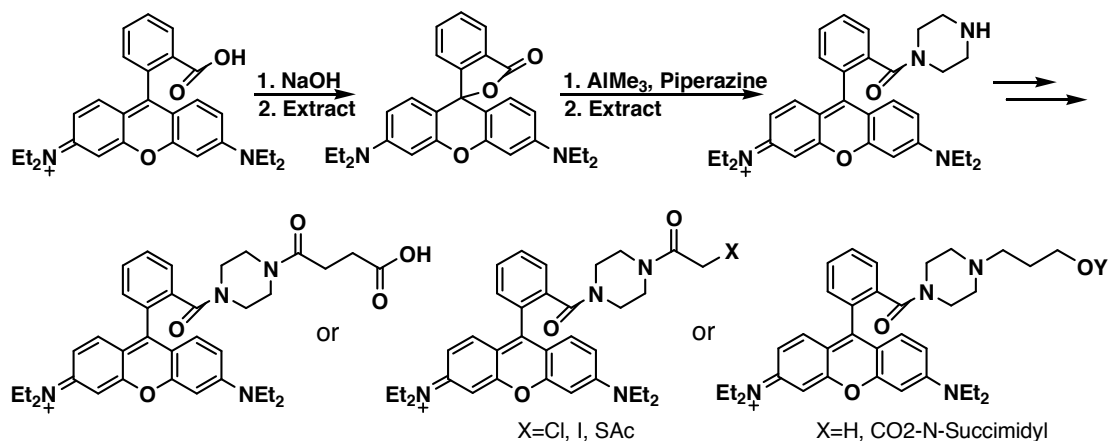
Scheme6. Derivatization of rhodamine B via rhodamine B-piperazine amide.[35]

One limitation of this approach is the multi-step procedure where rhodamine was first converted to the rhodamine base, then transformed to rhodamine-piperazine intermediate and finally derivatized. In addition, the lipophilic piperazine ring linker hampers the water solubility of rhodamine, which represents a key concern in protein chemistry where good water solubility of conjugation reagents is preferred. Furthermore, the usage of dangerous trimethylaluminum limits the application of this approach.

1.4.2. DERIVATIZATION OF RHODAMINE B TO RB-NHS ESTER (55)

A more recent method is one-step conversion of rhodamine B into rhodamine B NHS ester applying coupling reagent dicyclohexylcarbodimide (DCC) (Scheme 7).[36] This protocol is relatively simple to yield an amine-reactive rhodamine dye. However, more attractive derivatives such as azido rhodamine, ethynyl rhodamine, and maleimide rhodamine cannot be achieved via this method. Therefore, the development of a facile, safe and versatile method toward useful rhodamine dyes represents a key concern.
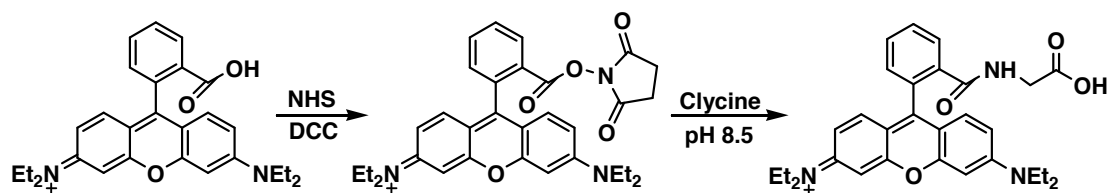
Scheme7. Derivatization of rhodamine B (**54**) via RB-NHS (**55**) intermediate.

In Results and Discussions section, our new research progresses in circumventing the above mentioned challenges in protein modification field as well as rhodamine derivatization area will be discussed accordingly. It includes a biology way toward site-directed protein modifications, two chemistry ways toward site-directed protein modifications and a facile approach for preparing a variety of useful rhodamine dyes for decorating those functionalized proteins.

Chapter 2. RESULTS AND DISCUSSIONS

2.1 The Biology SPI Approach towards Site-Directed Protein Modifications

In this section, the biology SPI approach was applied to yield a bis-azido RNase A modified in a site-directed manner. The NCAA—AHA (**2**) was synthesized prior to the incorporation into RNase A. Subsequently, this amino acid was successfully inserted into the amino acid sequence of RNase A, thus yielding bis-azido RNase A with two AHA residues located at AHA 1 and AHA 41.

2.1.1. SYNTHESIS OF AHA (2)

AHA has emerged as a widely used NCAA that has been successfully incorporated into the amino acid sequence of a few native proteins, such as *Candida antarctica* lipase B[13] and ribonuclease A.[11] In this thesis, an optimized synthetic strategy of the originally published procedure has been elaborated and applied to produce larger amounts of AHA.[14] Based on this protocol, a significantly higher yield was achieved, which enabled the expression of RNase A for site-directed click-reactions. The whole reaction scheme is described in scheme 8.
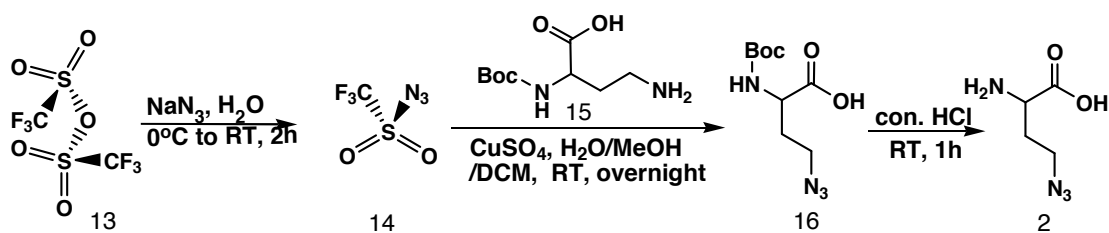
Scheme8. 3-Step synthesis of AHA (**2**) via diazo transfer reaction.

2.1.1.1 Synthesis of Triflic Azide (14)

The synthesis of triflic azide is the first key step since triflic azide acts as the diazo transfer reagent which converts a primary amino group into an azido group. The synthesis starts from

the reaction of triflic anhydride with sodium azide at 0 ^{o}C in an aqueous media. According to our experimental results, the purity of the starting material—triflic anhydride (**13**), plays a key role to ensure a good yield of triflic azide. In order to achieve this, triflic anhydride requires purification via a so called "vacuum transfer" set up. Briefly, P$_2$O$_5$ was employed as the desiccation reagent and triflic anhydride was "distilled" at 0 ^{o}C, then solidified in a liquid nitrogen bath under high vacuum. The whole process is elaborated in the experimental part. Via this method, purified triflic anhydride was obtained. Subsequently, the successful preparation of triflic azide is indicated by the observation of an organic layer which settled at the bottom of the reaction vial. This organic layer is the desirable product—triflic azide—and can easily be isolated by extraction using dichloromethane (DCM).

The major safety consideration in this synthesis is the highly reactive triflic azide as it is unstable, toxic, and extremely prone to explosion upon any slight shaking, grinding, pressurizing, concentrating or heating. Also, triflic azide must be used in a solution form rather than an isolated state. Therefore, the preparation and handling of this compound must be with great caution and a large scale synthesis of this reagent should be avoided. As a result, a separation funnel is not applicable for extraction but the combined use of test tubes and droppers is suitable.

Triflic azide is unstable and is only effective for the next diazotransfer reaction when it is used immediately after preparation. The azide-containing waste needs to be disposed according to standard operation procedure (SOP) since azide waste is highly toxic to the environment.

2.1.1.2 Synthesis of Boc-AHA (**16**)

The second key step is the synthesis of Boc-AHA as the precursor of AHA. In this process, the amino group of *N*-Boc-2,4-diaminobutyric acid (Boc-Dab **15**) was converted to an azido group by triflic azide under the catalysis of Cu^{2+} in basic conditions. Due to the different

solubility profile of Boc-Dab, triflic azide and cupper salt, a trinary solvent system was used which consisted of H_2O, MeOH and DCM in a ratio of 1:2:1. In such a way, cupper salt, hydrophilic Boc-Dab and lipophilic triflic azide could somehow stay in a same phase thus speeding up the reaction.

2.1.1.3 Deprotection of the Boc-AHA and Ion Exchange Purification of AHA (2)

The deprotection of Boc-AHA proceeds smoothly using concentrated hydrochloric acid. The resultant product is further purified via cationic exchange chromatography. It should be noted that the loading of the cationic exchange resin should be sufficient in order to avoid the loss of AHA during the ion exchange purification. The obtained AHA is a white crystal with an average yield of 77 % which is obviously higher than the reported various yield of 34~59 %.

2.1.1.4. Characterization of AHA (2)

The final product AHA is characterized by ESI-MS, IR, ^{1}H-NMR and ^{13}C NMR. ESI-MS gives peaks of m/e 144.9 [M+H]$^+$ and 143.0 [M-H]$^-$ which match the M.W. of AHA (144 g/mol). IR (KBr) confirms the presence of azido groups in AHA due to the characteristic peak at 2110 cm^{-1} (Figure 6). ^{1}H-NMR (D_2O) showed the three sets of proton in AHA (Figure 6) and ^{13}C-NMR (D_2O) further supported the structure of AHA (see appendix).
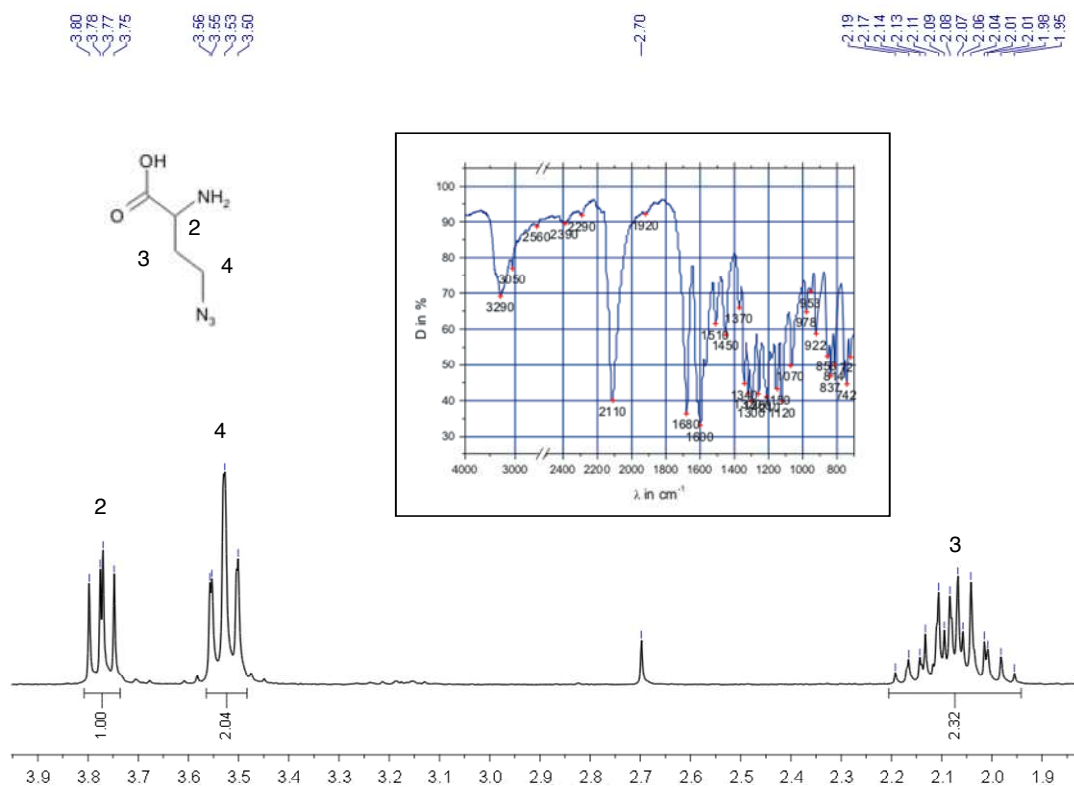
Figure6. IR and ^{1}H-NMR (D$_2$O) spectra of AHA (**2**).

2.1.2. INCORPORATION OF AHA (2) INTO RNASE A (18) VIA SPI

After successful preparation of AHA, this amino acid was incorporated as a methionine analogue into the amino acid sequence of RNase A using SPI. This piece of work was performed together with Joerg Wilhelmi and will be described in his thesis as well. After substitution of all methionine amino acids by AHA during SPI, an RNase with two accessible AHA-groups (called "bi-functional" RNase) was obtained. This RNase A variant was fully characterized by MALDI-ToF MS, gel electrophoresis and nanoLC-HR(ESI)MS/MS. The MALDI-ToF-MS spectrum of this protein reveals a characteristic signal with a M.W. of 15.77 kD (Figure 7a). In parallel, the amino acid sequence information was revealed by nanoLC-HR(ESI)MS/MS of the in-gel digested protein sample. It shows that the two accessible AHA residues are AHA 1 and AHA 41. AHA 41 was located in close proximity to the active centre and the other one (AHA 1) was positioned at the *N*-terminus. The rest of the AHA residues were buried inside protein and are considered non-accessible for chemical reactions. The fidelity of AHA 1 and AHA 41 are determined to be 99.6 % and 97.8 %,

repectively (Figure 7b).
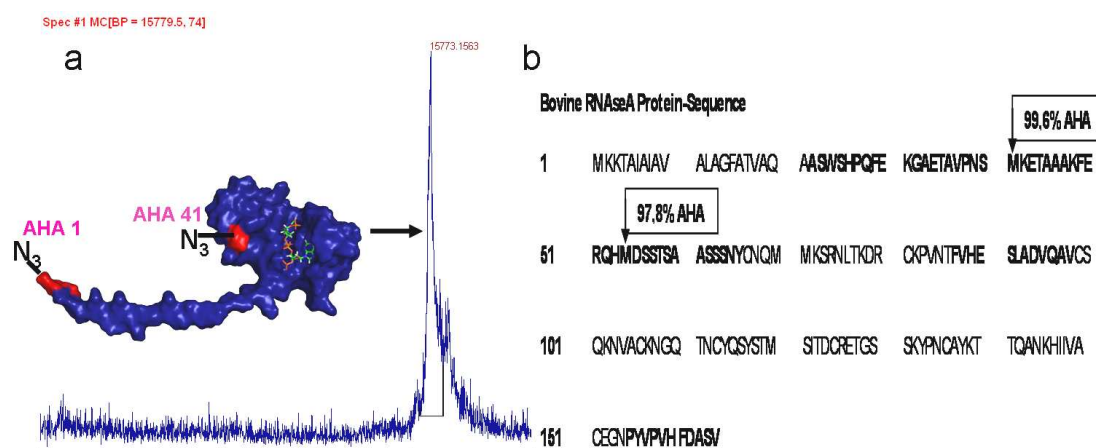
Figure7 a-b. MALDI-ToF-MS spectrum of bis-AHA-RNase A (**17**) (a) and its aminoacid sequence informatioin revealed by nanoLC-HR(ESI)MS/MS (b).

In summary, bis-azido RNase A with two AHA residues located at AHA1 and AHA41 was successfully prepared via the SPI appraoch. First of all, the NCAA, AHA, was synthesized with significantly enhanced yield and subseuqently, it was incorporated into the amino acid sequence of RNase A yielding the bis-azido RNase A variant. The M.W. of this variant was determined by MALDI-ToF-MS and its sequence information was retrieved from nanoLC-HR(ESI)MS/MS analysis of the in-gel digested protein sample.

2.1.3. ATTEMPTS TOWARD THE SYNTHESIS OF AN "N-VALINE" DERIVATIVE (**20**) FOR THE INCORPORATION INTO RNASE A (**18**)

Charges play a key role in fine-tuning the characteristics and the activity profile of proteins. For instance, it has been shown that the conversion of negatively charged carboxylic acid groups into primary amino groups yields a more "cationic" protein. Via this approach, the plasma protein BSA as well as RNase A were cationized, which strongly affected their ability to penetrate cell membranes.[29] In particular, cationic RNase A was able to cross cell membranes which enhanced its therapeutic profile as a promising anti-cancer agent.[37,38,39] In order to explore this possiblity, nitrogen containing amino acids (N-amino acids) were designed. They have a high likelyhood of being recognized by the translational machinery of

cell due to their high structural similarity to their corresponding native ones. In particular, the novel nitrogen-containing valine surrogate—"N-Valine" was designed. (Figure 8).
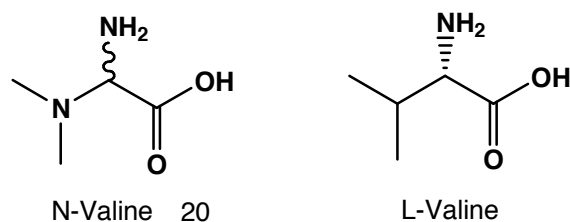
Figure8. Comparison of the structure between valine surrogate "N-valine" (**20**) and native L-valine.

2.1.3.1. Synthesis of 20 via Reductive Amination

The typical reductive amination involves the conversion of a carbonyl group into an imine intermediate. This reaction often takes place *in situ* where the formed imine was subsequently reduced into the amino group by a suitable reducing reagent. This reaction was considered worthwhile since the reduction agent selectively differentiates between the carbonyl starting material and the imine intermediate. In case a strong reducing reagent is applied such as lithium aluminum hydride, both the starting material and the imine intermediate will be reduced. Therefore, along with the formation of amino-products, an alcohol byproduct would be generated in parallel. One possibility to avoid this undesirable side reaction is to use a milder reducing agent that only attacks the more reactive imine-intermediate. Usually, sodium cyanoborohydride represents a good choice since the electron-withdrawing cyano group decreases the nucleophilicity of the hydride ion. In addition, water formed during the reaction needs to be removed in order to shift the equilibrium towards the product. Hence, activated 4-Å molecular sieve or other desiccating reagents are often added to the reaction mixture.

For our purpose, *N,N*-dimethyloxamic acid (**19**) was used as starting material. After the reduction of the carbonyl group, the target compound "N-Valine" (**20**) could be obtained in a one-step reaction according to the reaction scheme shown below (Scheme 9):
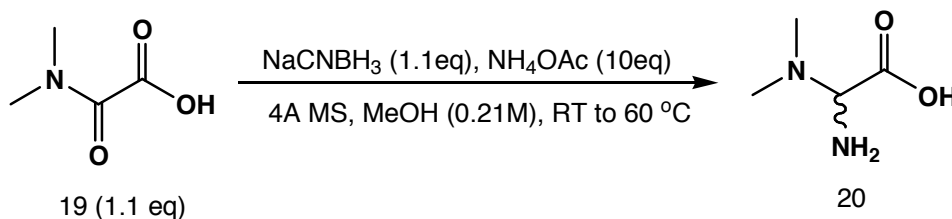
Scheme9. Attempts toward the synthesis of "N-Valine" (**20**) via reductive amination of *N,N*-dimethyloxamic acid (**19**).

Unfortunately, after an overnight reaction at room temperature (RT), no product was detected according to LC-ESIMS analysis. Therefore, the reaction temperature was increased to 60 °C, but still, no product was detected. Then, different ammonium sources such as ammonium hydroxide, ammonium chloride and even the strong reducing reagent sodium borohydride were tested. Unfortunately, none of these experiments turned out to be successful. Further attempts to achieve **20** via this method were abandoned as a new synthetic scheme was proposed (Scheme 10). Probably the presence of the dimethylamino group of **19** deactivates its adjacent carbonyl group via the conjugation electron-donating effect.

2.1.3.2. Synthesis of 20 via Bromination and Nucleophilic Substitution

It was reported in *Organic Synthesis* (1993)[40] that Boc-glycine *tert*-butyl ester was brominated by *N*-bromosuccimide under UV irradiation using CCl_4 as solvent. The brominated Boc-Gly-O*t*Bu could in principle be converted into the desired N-valine after nucleophilic substitution with dimethylamine (Scheme 10). As CCl_4 is a controlled chemical in Singapore, the possibility to use alternative solvents for this reaction was explored.
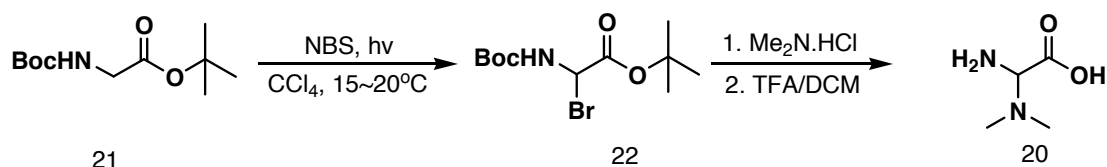
Scheme10. Synthesis of "N-Valine" (**20**) via radical bromination of Boc-Gly-O*t*Bu (**21**) followed by nucleophilic substitution by dimethylamine.

Five different solvents were screened including $CHCl_3$, DCM, trichloroethylene, benzene and

chlorobenzene and for all reactions the reported procedure was strictly obeyed. Anhydrous benzene was used as purchased; all other solvents were pre-treated with activated 4Å molecular sieve before use. Unfortunately, none of the reactions yielded the brominated product according to LC-ESIMS and ^{1}H-NMR analysis. Since reported similar bromination reactions all proceeded in CCl_4 only, this radical bromination reaction might only work in CCl_4. As a result, a third synthetic scheme employing phthaloglycine as a precursor was designed.

2.1.3.3. Synthesis of 20 Using Phthaloglycine (24) as a Precursor

Compared to *N*-Boc protected glycine (**21**), *N*-phthalo protected glycine (**24**) has a much higher chemical stability and the phthalo group can easily be deprotected under the treatment of con. HCl. The deprotected phthalo group is converted to phthalic acid precipitate which can easily be removed by filtration. In addition, the phthalimido group represents a tertiary amide that might not undergo any side reactions associated with the *N*-H bond on a secondary amide protection group.

Three approaches were proposed. The first route was based on the Hell-Volhard-Zelinsky bromination reaction (Scheme 11, **25→28**) whereby phthalimidoacetyl chloride (**25**) was treated with refluxing Br_2. Unfortunately, according to ^{1}H-NMR determination, the phthamidoacetyl chloride (**25**) remained unchanged. The second approach was the abstraction of the *α*-H of the phthaloglycine methyl ester (**26**) by sodium hydroxide followed by treatment with NBS. Similarly, this attempt was unsuccessful according to ^{1}H-NMR analysis. We predicted that sodium hydroxide might not be strong enough to abstract the *α*-H of phthaloglycine methyl ester (**26**). In addition, NBS might not be the best bromine donor. Hence, we finally tried NaH combined with Br_2 for the reaction. We expected that the NaH would be strong enough to abstract the *α*-H and the resulting anion would subsequently attack bromine to complete the bromination (Scheme 11, **26→27**). Surprisingly, the starting material was totally unchanged after an overnight reaction.
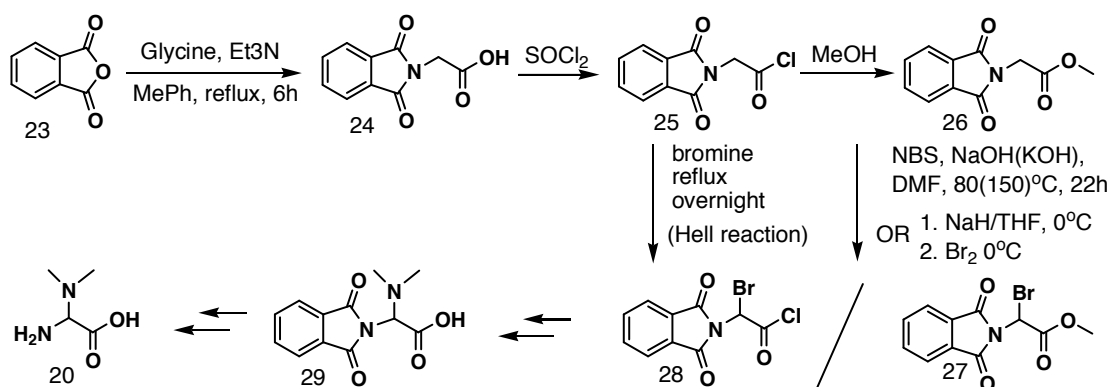
Scheme 11. Attempts toward the synthesis of "N-valine" (**20**) using phthaloglycine (**24**) as a precursor.

In conclusion, the possibility of preparing "N-Valine" via several distinct routes, such as reductive amination, bromination of Boc-Gly-O*t*Bu, Hell-Volhard-Zelinsky reaction and bromination of phthaloglycine methyl ester, were explored. Although these attempts seem to be not satisfactory, meaningful results were acquired that might be helpful for future experiments within this project. Site-directed modifications of proteins via e.g. genetic engineering are certainly feasible, but they also suffer from several limitations. In this context, both SPI or *nonsense codon suppression* require not only the design of the respective NCAA, but also long time periods for protein expression. Furthermore, the yield of the corresponding protein is variable and might not be feasible to produce sufficient quantities for further investigations. Based on the above understanding, we were motivated to develop chemistry methods for site-directed protein modification which are more concise, time-saving and might afford larger quantities.

2.2. Chemistry Approaches toward Site-Directed Protein Modification

In this section, chemistry approaches, i.e. cysteine-targeting and lysine-targeting methods, for introducing site-directed protein modifications will be introduced.

2.2.1. THE CYSTEINE-TARGETING APPROACH TOWARD THE SITE-DIRECTED

MONO-BIOTINYLATION OF HSA (33)

Human serum albumin (HSA) is the most abundant protein in human blood plasma produced in the liver. It acts as a transportation protein in the body to carry thyroid hormones, fatty acids, unconjugated bilirubin etc. It also plays a key role in maintaining oncotic pressure.[46] Due to these facts, it is attractive to modify HSA in order to tune and alter its function. Among the 585 amino acid residues, cysteine 34 is a unique residue since it is the only unpaired cysteine present in the sequence and it is exposed to the surface thus allowing further chemical modifications. Reactions of free cysteine residues often involve Michael addition reactions applying maleimide derivatives as thiol-reactive group. Therefore, site-specific mono-modification of HSA is possible.

In theory, the chemical reaction needs to proceed with quantitative yield since otherwise, a mixture of native and modified HSA will be formed. Unfortunately, conjugation with HSA using maleimides usually failed to reach 100 % conversion. Hence, the native HSA needs to be removed from the reaction mixture after chemical modification. One possibility to address this problem is based on the attachment of an affinity purification tag on the maleimide reagent. After the reaction, mono-modified HSA could be isolated via affinity purification and separated from the native HSA.

Biotin was chosen as the affinity tag since biotin-containing proteins can be purified via Avidin affinity column due to the "quasi" reversible binding between biotin and the monomeric protein Avidin. Based on this principle, biotin-LC-maleimide (sheme12) was designed and synthesized. This novel reagent was subsequently reacted with HSA yielding the biotinylated HSA, which was purified by Avidin column.
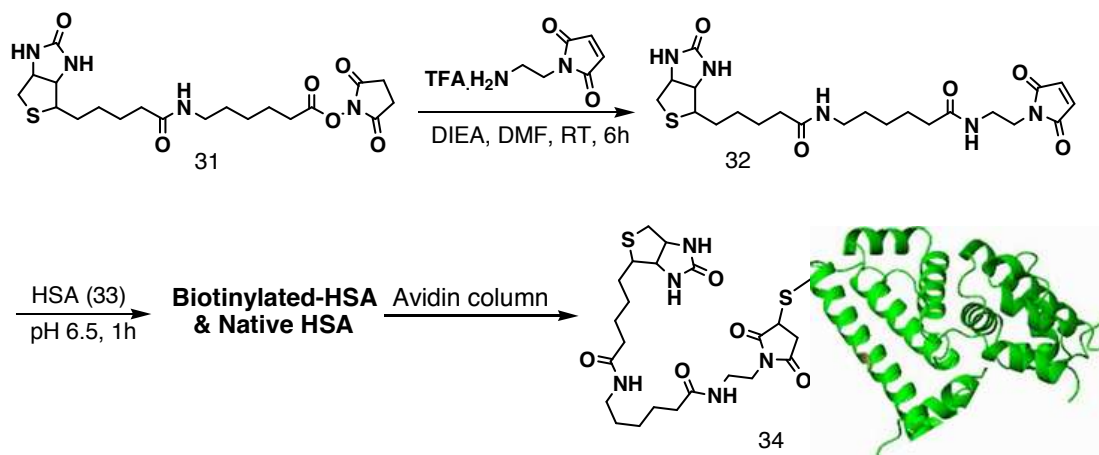
Scheme12. Synthesis of biotin-LC-maleimide (**32**) for mono-biotinylation of HSA (**33**).

The success of this approach was demonstrated by MALDI-ToF-MS using sinapinic acid (SA) as the matrix. As shown in figure 5a, native HSA gives a M.W. of 66618 Dalton. After treatment of biotin-LC-maleimide and Avidin column purification, the modified protein shows a M.W. of 67097 Dalton (figure 5b). The M.W. increase (478 Dalton) closely matches the M.W. of biotin-LC-maleimide (480 Dalton).
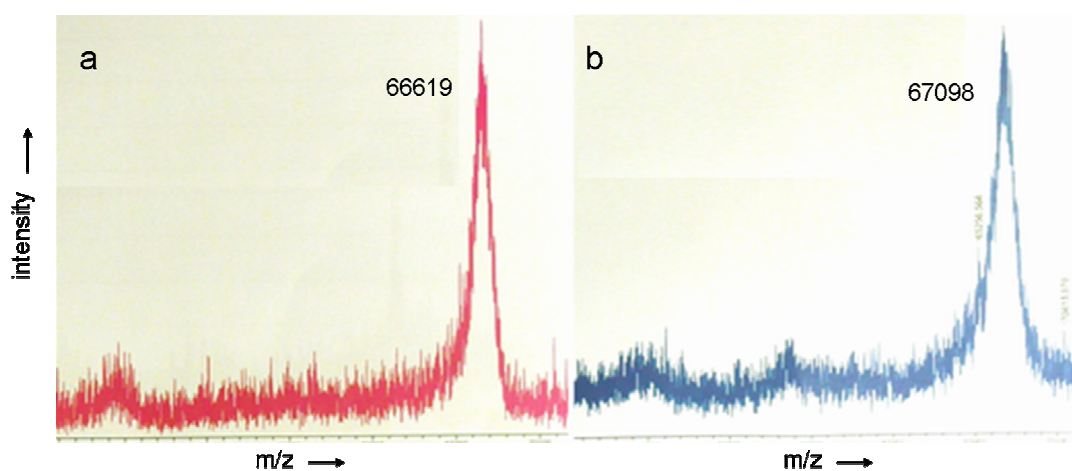
Figure9 a-b. MALDI-ToF-MS spectra of native HSA (**33**) (a) and mono-biotin-HSA (**34**) (b).

This approach was proven successful for site-specific functionalization of protein. Using the reported protein database for molecular modeling, the proteins which have only a unique cysteine on their surface were identified. This protein series included serum albumins, histone proteins and some antibodies. Proteins under this category can be used for site-specific

modification by this protocol. If one protein does not show a surface cysteine, it could be modified after introducing a single surface-accessible cysteine group at a distinct position via site-directed mutagenesis. For other proteins which bear multiply accessible cysteines, this approach might not be applicable.

However, this protocol suffers from the limitation that the scope of this method is relatively narrow. Hence, we explored the possibility of other more widely applicable functionalization protocols.

2.2.2. *N*-TERMINAL FUNCTIONALIZATION OF RNASE A (**18**) *VIA* TRANSAMINATION

2.2.2.1. Design of the Modification Protocol

As mentioned in *Introduction*, J M Gilmore and M B Francis have reported the *N*-terminal protein modification via a biomimetic transamination reaction.[33] The *N*-terminal *α* amino group was converted to a ketone or an aldehyde upon reaction with PPL. When a protein has an *N*-terminal glycine residue, the intermediate will be an aldehyde which allows additional reactions such as the Pictet–Spengler reaction.[34] We have tested three approaches for *N*-terminal modification of RNase A based on this rationale.

All three routes were based on the transamination of RNase A to afford the activated RNase A with a ketone tag at the *N*-terminus (Scheme 13a). This keto-RNase A was subjected to three different reactions—(1) hydrazone formation, (2) an Aldol reaction and (3) a Pictet-Spengler reaction. The correspondingly designed reaction schemes are shown below.

a. Transamination of native RNaseA (18) to afford keto-RNaseA (35)

Native RNaseA   18   keto-RNaseA   35   PPL 36
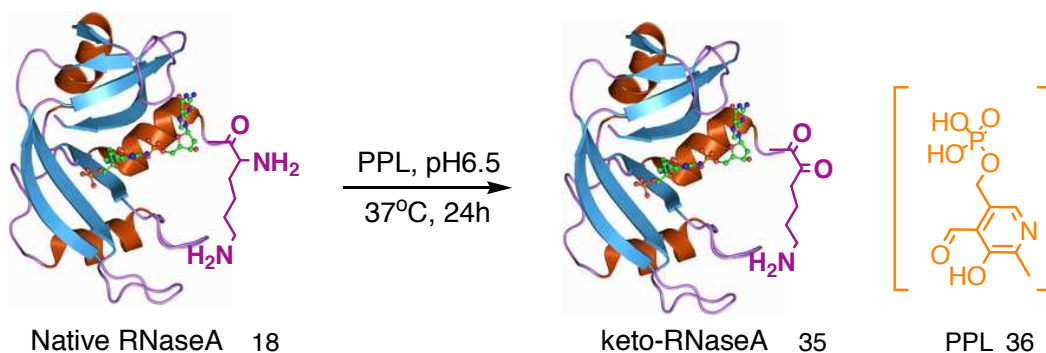
b. Introduce an ethynyl group *via* hydrozone formation

37   38   39
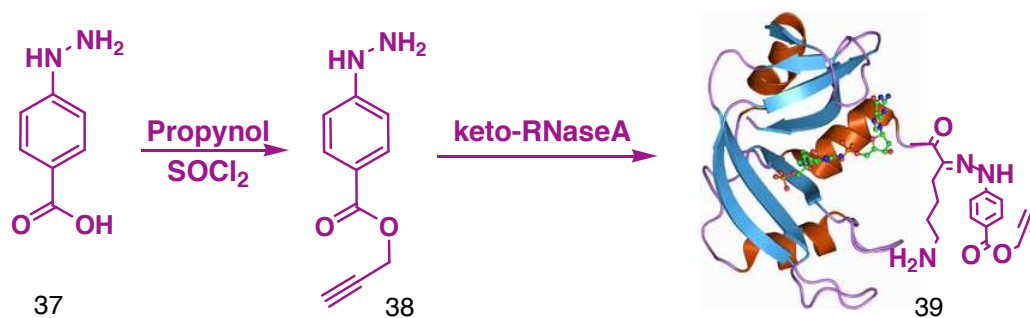
c. Introduce an ethynyl group *via* Aldol reaction

40   41   Williamson Etherification   42
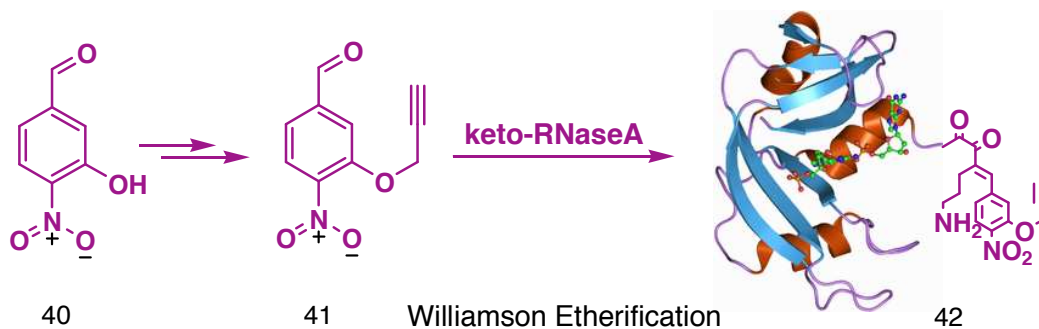
d. Introduce an ethynyl group *via* Pictet-Spengler reaction

43   Sonogashira Reaction   44   45
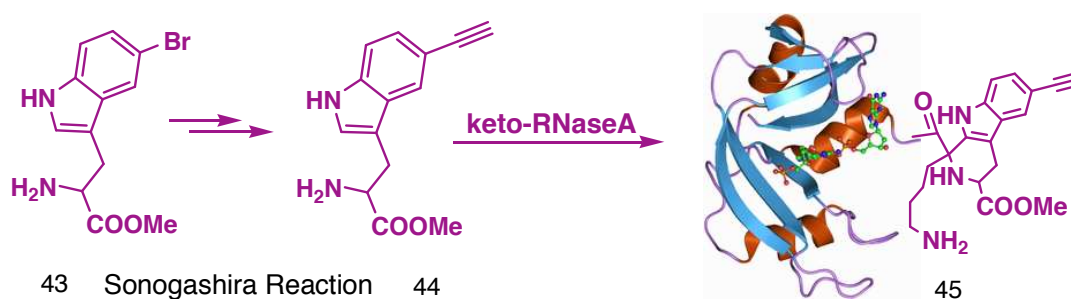
Scheme13a-d. Four designed routes toward *N*-terminal functionalization of RNase A

(**18**) *via* transamination reaction.

In scheme b, the 4-hydrazinylbenzoic acid (**37**) is first equipped with an ethynyl group after treatment with 2-propynol to afford 4-hydrazinylbenzoic acid propargyl ester (**38**). This reagent further reacts with keto-RNase A (**35**) in order to yield a ethynyl-RNase A (**39**) via a hydrazone formation reaction. In scheme c, 3-hydroxy-4-nitrobenaldehyde (**40**) is introduced with an ethynyl group via Williamson etherification to afford **41**. Subsequently, this ethynyl-benzaldehyde (**41**) reacts with keto-RNase A (**35**) to form the ethynyl conjugated RNase A (**42**) via an Aldol reaction. In the fourth scheme d, 5-bromotryptophan methyl ester (**43**) is converted to 5-ethynyltryptophan methyl ester (**44**) though a Sonogashira reaction. Then 5-ethynyltryptophan further reacts with keto-RNase A to form ethynyl-RNase A (**45**) via a Pictet-Spengler reaction.

2.2.2.2. Experimental Results

As a start, 4-hydrazinylbenzoic acid, 4-nitrobenzaldehyde and 5-bromotryptophan were employed to test the feasibility of all three routes. Both MALDI-ToF-MS and UV spectroscopy were used to assess product formation. Unfortunately, no protein was detected according to MALDI-ToF-MS or UV. According to the MALDI-ToF-MS spectrum, no protein peak was observed in the range of 14 kD (Figure 10 a). Similarly, in the UV absorption spectrum, no absorption was found indicating the absence of RNase A in the sample.
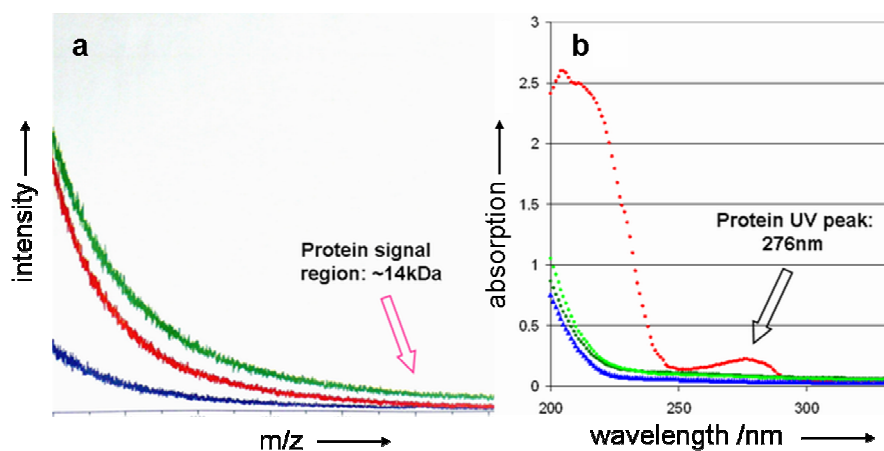
Figure10 a-b. a. MALDI-ToF-MS spectra of the three reaction solutions (green: hydrazone reaction; red: Aldol reaction; blue: Pictet-Spengler reaction); b. UV spectrum of the three reaction solutions together with native RNase A.

2.2.2.3. The Explanation of the Failure of This Approach

Based on these negative results, the protein activation step, i.e. the transamination reaction was reconsidered. A conversion of about 50% of RNase A during transamination was reported.[33] We hereby determined the UV and MALDI-ToF-MS of the keto-RNase A derivative (**35**), which was prepared strictly according to the reported procedure.[33] The results suggested that the first reaction step exhibits some problems. As a control, the MALDI-ToF-MS spectrum of a similar concentration of native RNase A (50mM) was also collected. Compared to the intensive mass peak at ~13.7kDa of native RNase A (Figure 11a), keto-RNase A only gives a faint mass peak at around 14 kDa (Figure 11b). Furthermore, the UV spectrum of keto-RNase A also shows the absence of the characteristic peak at 276nm (spectra not shown here). These data suggest that RNase A has decomposed during the transamination reaction. The failure of this method for the modification of RNase A might not occur in the functionalization of other proteins, for instance thioredoxin. According to the literature, thioredoxin shows a much higher conversion rate under the same conditions (~94 %). However, since this approach proved to be not successful on RNase A, an alternative strategy was proposed targeting lysine residues. .
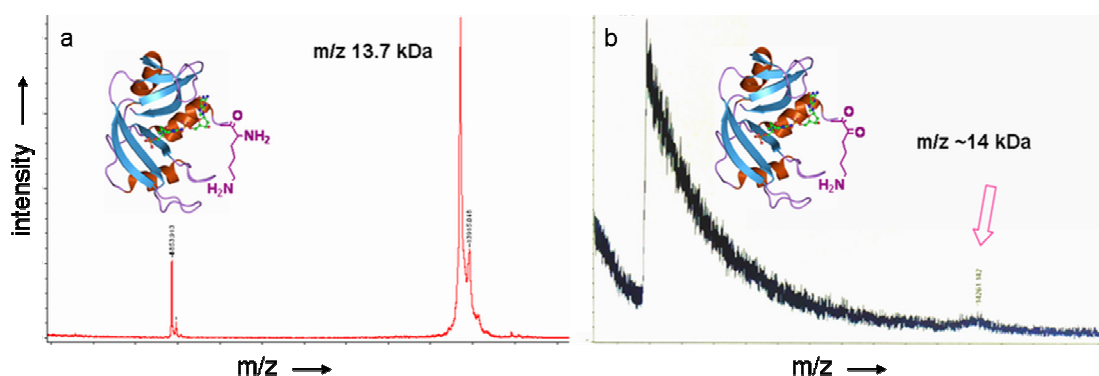
Figure11 a-b. MALDI-ToF-MS spectrum of native RNase A (**18**, 50mM) (a) and keto－RNase A (**35**, ~50mM) (b).

2.2.3. THE LYSINE-TARGETING APPROACH TOWARD THE MONO-BIOTINYLATION OF PROTEINS

As most proteins bare multiple lysine residues on their surface, targeting a single lysine

without interfering with others appears to be challenging. For instance, according to some published results, modification of proteins targeting lysine leads to a heterogeneous mixtures.[32] However, as the lysine targeting approach benefits from the fact that it only requires mild reaction conditions, wide application scope etc., we were motivated to explore this possibility.

First, we carefully examined the surface-exposed functional groups of proteins. With respect to lysine residue, we raised the question whether all surface lysine groups are equally exposed at the surface and whether they are equally accessible by amino-reactive reagents. In addition, we wondered whether all lysine residues had the same interaction with their environment, for instance H-H bonding. If not, surface lysines might display different activity toward amine-reactive reagents and the most reactive lysine might be selectively targeted using a tailored bioconjugation reagent.

2.2.3.1. Evaluation of the Solvent Accessibility of Lysine Residues of RNase A (18)

Again, RNase A has been chosen as a model protein since it is of medium size, is a well-characterized enzyme, is stable, and has been explored as a potential effective anti-tumor therapeutic.[37,41,42,43] First the distribution of lysine residues was examined applying the molecular modeling software MOE (molecular operation environment) (PDF file: 1RCA). As revealed by the 3D structure, RNase A possesses as many as ten lysines residues and nine out of ten are accessible on the surface except lysine 41 which is located in the active centre (Figure 12). Among these nine surface lysines, K1 exhibits the highest accessibility according to the calculation by MOE.
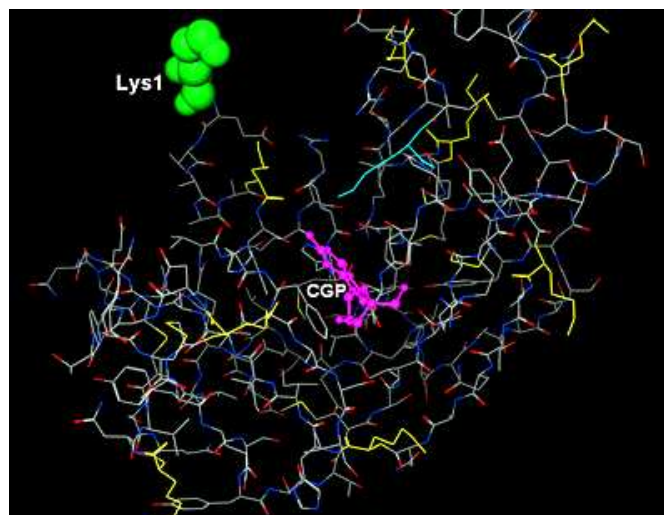
Figure12. The crystal structure of RNase A (**20**) (PDB file: 1RCA) where Lys 1 is highlighted by green spheres whereas other accessible lysine residues are indicated in yellow except the solely buried Lys 41 shown in bright blue; the substrate of RNase A, 2'-Deoxycytidine-2'-deoxyguanosine -3'-5'-monophosphate (CGP), is indicated in mauve.

Furthermore, we examined the interactions of K1 together with other surface lysines, with other residues in close vicinity based on its crystal structure. According to our study, no obvious interactions, such as H-H bonding were revealed. This precludes the possibility that residue interactions might affect the activity of lysine. Therefore, we predicted that the K1 residue might be preferentially targeted when RNase A is treated with an amine-reactive reagent such as NHS ester.

2.2.3.2. Several Considerations Concerning Site-Directed Modifications

There are several considerations for designing the modification reagent and choosing the appropriate reaction conditions. Firstly, the modification reagent needs to be active enough to react with the protein smoothly at low concentrations. Secondly, water–solubility of the bioconjugation-reagent is preferred for protein modifications in aqueous media. Thirdly, the modification reagent should be present in limited quantities during the reaction in order to ensure mono-functionalization of RNase A. A mixture of native RNase A and modified RNase A will be formed under such reaction conditions. Therefore, the presence of an affinity purification tag on the reagent is necessary for later separation.

2.2.3.3. Design and Synthesis of Biotin-PEG-NHS (49)

Based on the before-mentioned considerations, biotin-PEG-NHS (**49**, scheme 14) was designed and synthesized. The biotin moiety represents an affinity tag while the PEG chain contributes to water-solubility. The NHS-ester is known to be a very active reagent that selectively attacks amino-groups.
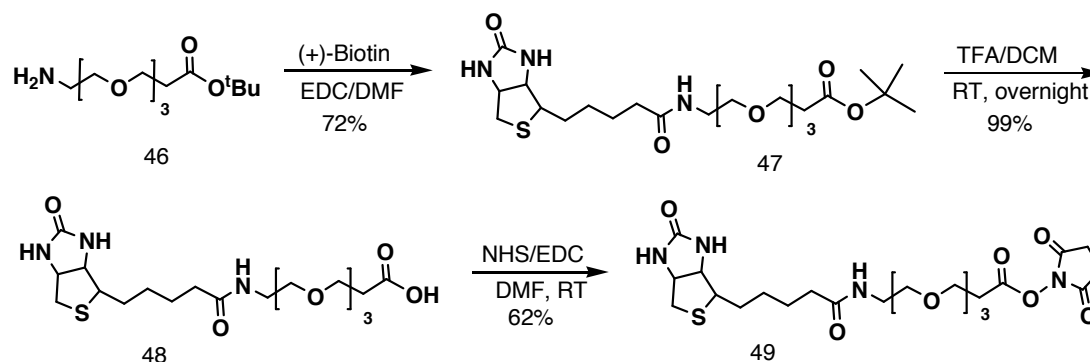
Scheme14. Synthetic scheme of biotin-PEG-NHS (**49**).

2.2.3.4. Mono-Biotinylation of RNase A (18), Lysozyme C (50) and SST (51)

RNase A, lysozyme C and Somatostatin (SST) were tested based on this method (Scheme 15). Basically, the protocol involves the addition of 0.5 equiv. of biotin-PEG-NHS solution in 100 portions into each protein solution stepwise (1 equiv.) during a two hours period. We expected that a mixture of native protein and mono-biotinylated protein would be formed.

Monobiotinylated Protein/Peptide
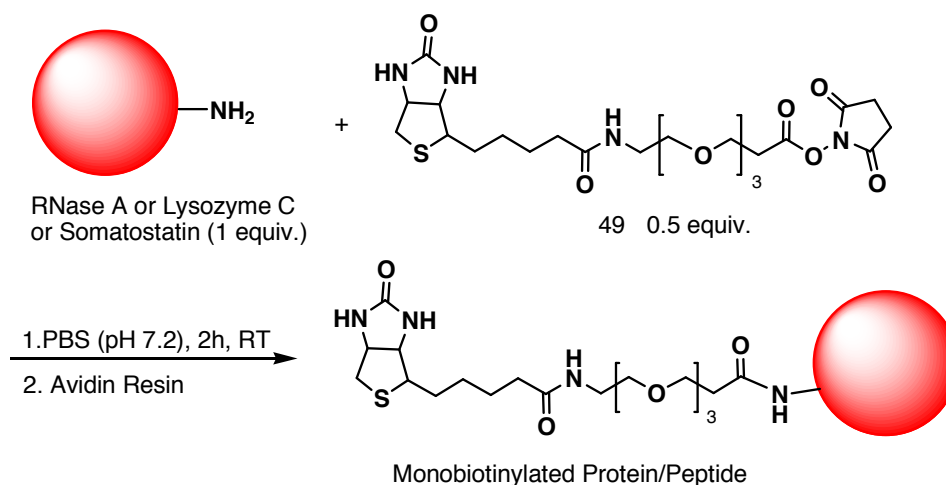
Scheme15. Mono-biotinylation of RNase A (**18**), lysozyme C (**50**) and SST (**51**) in the lysine-targeting approach.

The reaction solution of each protein was subjected to MALDI-ToF-MS analysis using SA as the matrix. In agreement with our hypothesis, all three proteins/peptides gave a mixture of native and mono-modified protein, as revealed by their MALDI-ToF-MS spectra (Figure 13)
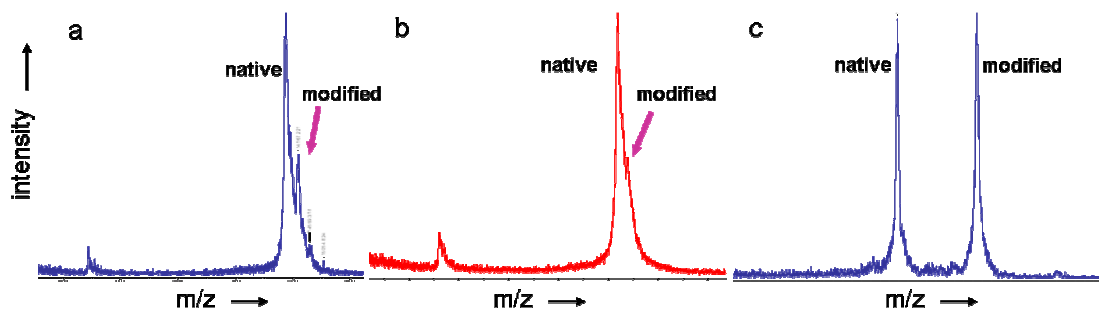
Figure13 a-c. MALDI-ToF-MS spectra of the reaction mixture of RNase A (**18**) (a), lysozyme C (**50**) (b) and SST (**51**) (c).

Subsequently, the reaction mixture was treated with ethanolamine to quench any remaining biotin-PEG-NHS. The resultant solutions were first purified by gel filtration to remove biotin-containing small molecules in the sample. Finally, the two proteins, biotin-PEG-RNase A and biotin-PEG-lysozyme C were further purified via affinity chromatography using Avidin resin. Purified proteins were characterized by MALDI-ToF-MS and biotin-PEG-RNase A displayed a single peak at m/z 14138 (calcd. m/z 14134). On the other hand, biotin-PEG-lysozyme C shows a major peak at m/z 14728 (calcd. m/z 14734) and a minor peak located at the left which indicates the presence of some native lysozyme C. This problem which is due to its non-specific binding to the Avidin column could be addressed in principle by using a fresh batch of Avidin resin. The final protein product can be further purified via a second gel filtration to remove biotin and salt and lyophilized to afford the powder form.
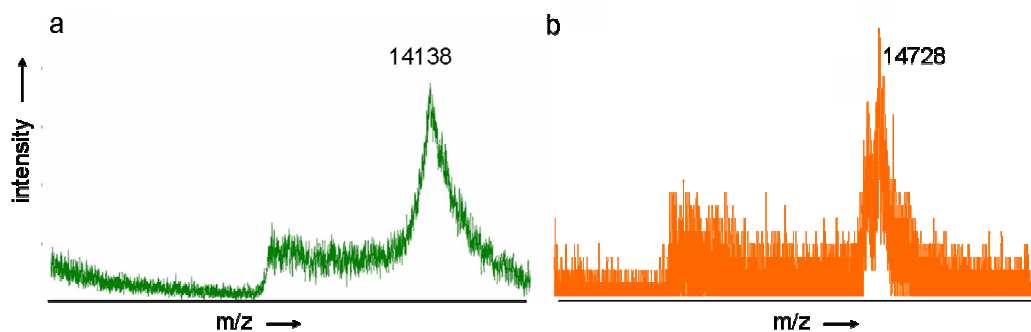
Figure14 a-b. Biotin-PEG-RNase A (**52**) shows a M.W. of 14.14 kD in the MALDI-ToF-MS spectrum and biotin-PEG-Lysozyme C (**53**) displays a M.W. of 14.73 kD.

In conclusion, a novel and facile protein mono-functionalization method has been developed which targets lysine residues. This approach contains four major steps as described in figure 15. First of all, the native protein was treated with a biotin-conjugation reagent to generate mono-biotinylated protein which contains native protein (Figure 15, step a). Subsequently, a quencher solution was added to decompose unreacted biotin-conjugation reagent followed by the first gel filtration to remove the deactivated biotin-containing small molecules (Figure 15, step b and c). Thirdly, mono-biotinylated protein was "fished" up from the solution via affinity chromatography based on the reverse binding between the biotin group and the monomeric Avidin, thus achieving the separation between modified and unmodified proteins. In the last step, the biotin and salt which stayed in the protein solution after affinity chromatography were removed via a second gel filtration and finally lyophilized to generate our product as a powder form. It is worthwhile to note that from a mixture of native and modified protein to biotinylated protein, native protein is also recovered and can participate in the next modification circle (Figure 15, dark blue arrow).
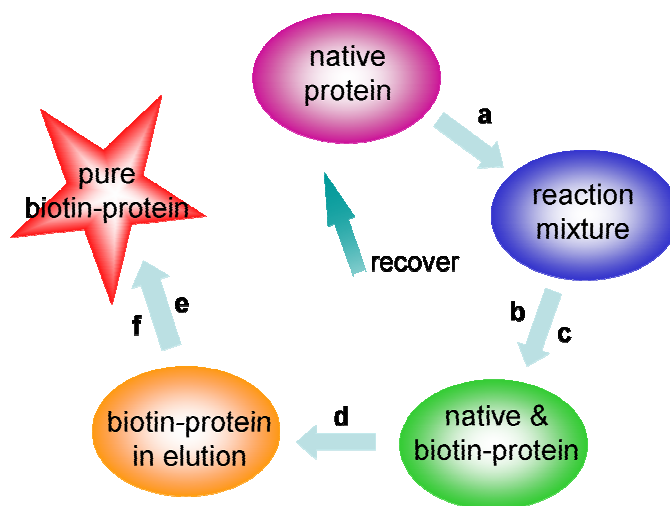
Figure15. The four major steps involved in the lysine-targeting approach for mono-modification of proteins.

In all, we expect that this method has a broad application scope since it targets surface-exposed lysine residues which typically exist on most protein surfaces. In addition, we predicted that the most reactive lysine residue was preferentially addressed during the mono-modification process based on the consideration of solvent-accessibility as well as the residue interactions in close vicinity. As a potential application, this method can be applied to introduce a single bioorthogonal functional group onto a protein via bioorthogonal conjugation reagents thus opening up the opportunity to further decorate proteins via click reactions.

2.3. A Novel Synthetic Method towards Rhodamine B Derivatives for Protein Labeling

After the successful introduction of functional groups into the protein scaffold in a site-directed fashion via three distinct approaches—the biology SPI approach, the cysteine targeting, and the lysine targeting chemistry approach—we were further motivated to develop reactive chromophores to label these modified proteins. This piece of work is scientifically attractive as these chromophores can impart new functionalities on proteins and enable further

investigation of them via FCS, FRET or other analytic techniques. In the following section, we focus on the derivatization of rhodamine B dye as it is commercially available and displays a higher photostability than some of other dyes such as fluorescien etc.

2.3.1. RB DERIVATIZATION BY REACTING WITH AMINES

It has been reported that rhodamine dyes can be derivatized via rhodamine-piperazine amide intermediate.[44] This method is generally applicable. However, due to the multiple reaction steps involved and usage of the hazardous reagent trimethylaluminum, the application of such method is still restricted. Hence, an alternative strategy has been suggested in this thesis, which is considered greener and safer.

2.3.1.1 Reaction of RB-NHS (55) with Primary Amines

Rhodamine B possesses a single free carboxylic acid group which could be targeted for derivatization reactions. One straightforward way is to directly let rhodamine B reacts with amines. Hence, our initial attempts are focused on the reaction of rhodamine B (**54**) with propargylamine to afford rhodamine B propargyl amide (**56**). To achieve this goal, rhodamine B was first activated to NHS ester via DCC coupling. The purified NHS ester was later treated with propargylamine in order to achieve ethynyl-rhodamine B (**56**, scheme 16).
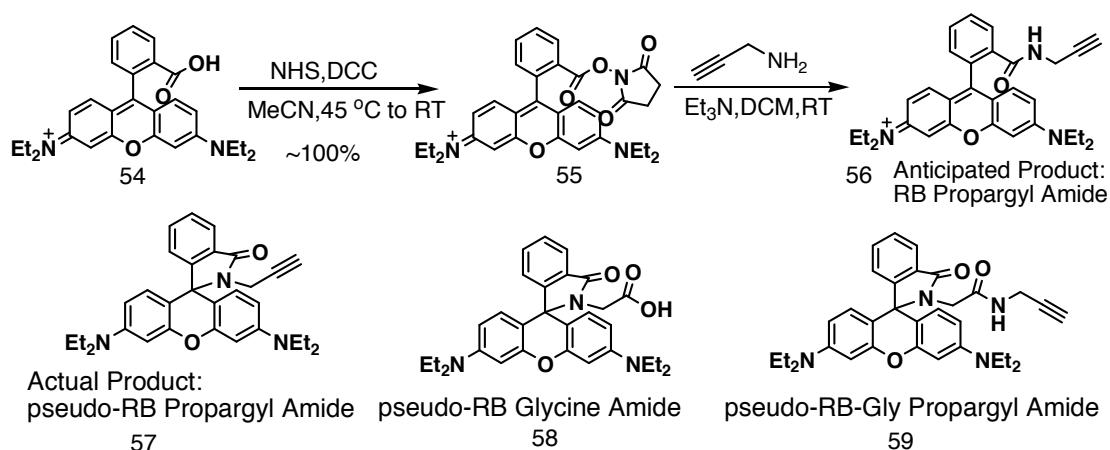
Scheme16. Synthesis of RB-propargyl amide (**56**) affords a pseudo product (**57**) due to the intramolecular cyclization.

A light pink powder was obtained after RB-NHS (**55**) reacted with propargylamine. The ^{1}H-NMR spectrum appeared to match with the anticipated product, RB propargyl amide (**56**) since both the rhodamine B moiety and the propargyl group could be identified individually in the ^{1}H-NMR spectrum. Also, the ESI-MS analysis seems to be in agreement with its M.W., m/z 480 [M$^+$] (calcd. m/z 480). Interestingly, two characteristic properties of rhodamine B were missing. The product became water insoluble which indicated the loss of its ionic nature. In addition, this compound had a light pink color rather than dark violet as we expected. This observation suggests the absence of the conjugated xanthene system. Hence, it can be concluded that the pseudo-rhodamine B propargyl amide (**57**) was formed instead of the anticipated product (**56**) via an intramolecular cyclization reaction (Scheme 14). Such a pseudo-rhodamine B propargyl amide (**57**) has an identical ESIMS value and a similar ^{1}H-NMR spectrum compared to rhodamine B propargyl amide (**56**) but it is not water soluble nor violet.

We also tried to couple RB-glycine (**72**) with propargylamine via DCC coupling. A colorless mixture was obtained whose ESI-MS data indicated that it was a mixture of pseudo-RB glycine amide (**58**, m/z 500 [M+H] $^+$) and pseudo RB-Gly propargyl amide (**59**, m/e 537 [M+H]$^+$) (Scheme 14).

All this experimental data suggested that the N-H group of the secondary amide bond in the expected product (e.g. **56**) tends to insert into the xanthene ring system to form the pseudo products via intramolecular cyclization. Therefore, we believe that replacement of a primary amine by a secondary amine as the starting material can circumvent this problem since a tertiary amide bond will be formed which will prevent the intramolecular cyclization. This approach will be discussed in the following section.

2.3.1.2. Reaction of RB-NHS (**55**) with Secondary Amines

Unlike *N*-propargylamine, *N*-methyl-*N*-propargylamine is a secondary amine which contains

one more methyl group on its nitrogen atom. Therefore, it was chosen to react with RB-NHS applying the same procedure in order to prepare ethynyl-rhodamine B (Scheme 17). This route proved to be successful since a dark violet solid was obtained and the LC-ESIMS indicated the formation of the targeted product.
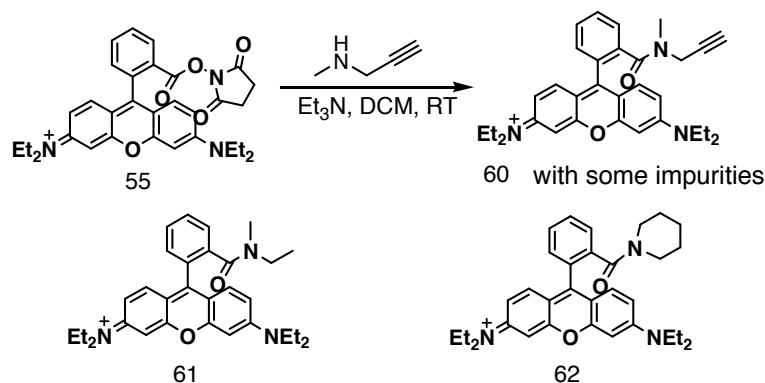
Scheme17. Secondary amines react with RB-NHS (**55**) to afford products **60**, **61** and **62** but with some impurities.

However, this product was found to contain some impurities which cannot easily be removed via a normal workup procedure and column chromatography techniques. The same situation was encountered when applying *N*-methyl-*N*-ethylamine and piperazine as the starting material. In principle, this problem could be overcome by using Prep-HPLC, but it will be costly and is uneasy to achieve large quantities. Because of this, a more successful and facile route was developed via one step Steglich esterification.

2.3.2. ONE-STEP FUNCTIONALIZATION OF RHODAMINE B VIA STEGLICH ESTERIFICATION

2.3.2.1. Synthesis of Rhodamine B Ethyl Ester (60)

Finally, we have explored the possibility of applying the Steglich esterification for rhodamine B derivatization. In the model reaction, rhodamine B was reacted with EtOH using DCC as the coupling reagent under the catalysis of DMAP (Scheme 18).
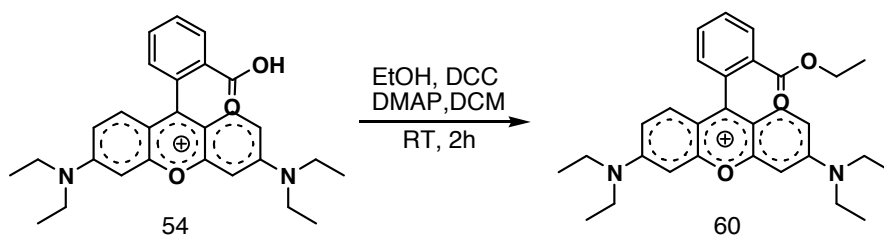
Scheme18. Synthesis of rhodamine B ethyl ester (**60**) via DCC-coupled Steglich esterification.

This DCC-coupled Steglich esterification proceeded very well in the synthesis of rhodamine B ethyl ester (**60**). A pure product was obtained as a violet crystal after workup and silica gel flash chromatography. LC-ESIMS gives 98.5 % purity (Figure 16) with m/z 471.15 [M$^+$] (calcd. 471.26). The structure of this rhodamine B ethyl ester was further confirmed by ^{1}H-NMR and ^{13}C-NMR.
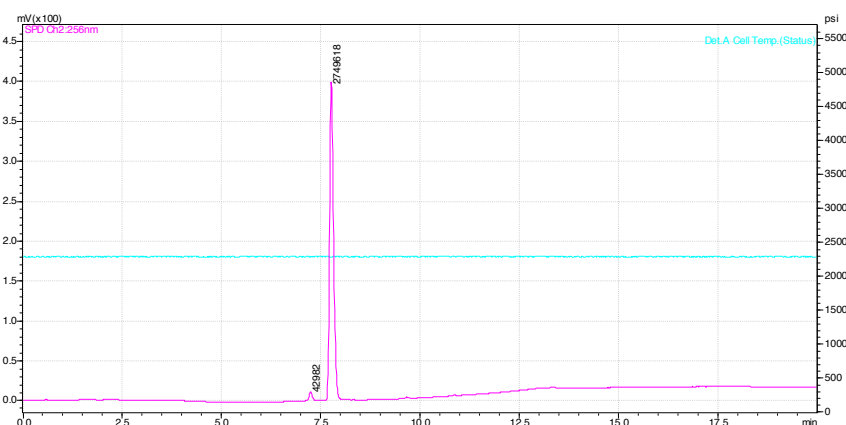
Figure16. HPLC spectrum of rhodamine B ethyl ester (**60**) reveals a purity of 98.5 % (256 nm).

2.3.2.2. Synthesis of Rhodamine B Propargyl Ester (**73**) for Antibody Labeling

With this protocol in hand, we first tried to synthesize rhodamine B propargyl ester for labeling the azido-proteins. Applying the same protocol, rhodamine B reacted with propargyl alcohol affording a pure rhodamine B propargyl ester in the form of a violet crystal.

This dye was used as a protein probe to label azido-3F11 which is a 25 kD antibody (monomer) functionalized with an azido group via copper catalyzed azide-alkyne cycloaddition (CuAAC, also known as Huisgen 1,3-dipolar cycloaddition) (Scheme 19a).

Gel electrophoresis results indicated the success of the labeling as the click product gave an intensive fluorescence band at 25 kD (bottom, monomer) and a faint fluorescent band at 50 kD (top, dimmer). In contrast, the azido-antibody (N_3-AB) band at neither 25 kD nor 50 kD showed no fluorescence at all.
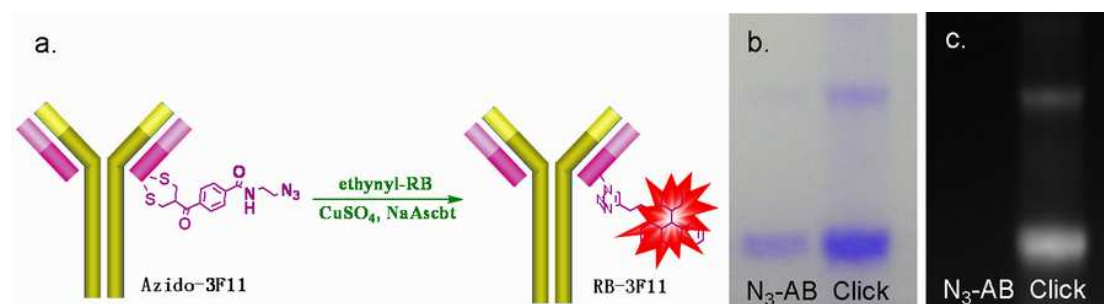
Scheme19 a-c. a. Label azido-3F11 using rhodamine B propargyl ester (**73**) via CuAAC; b. gel electrophoresis image of azido-3F11 (N_3-AB) and the click product (coomassie blue stain); c. fluorescence image of the electrophoresis gel of both N_3-AB and the click product.

2.3.2.3. Synthesis of Rhodamine Azidoethyl Ester (63) as an Ethynyl Reactive Chromophore

The third rhodamine B derivative synthesized is rhodamine B azidoethyl ester. 2-Bromoethanol was reacted with sodium azide to form 2-azidoethanol which was subsequently reacted with rhodamine B to afford azido-rhodamine B. The synthetic scheme is shown below (Scheme 20).
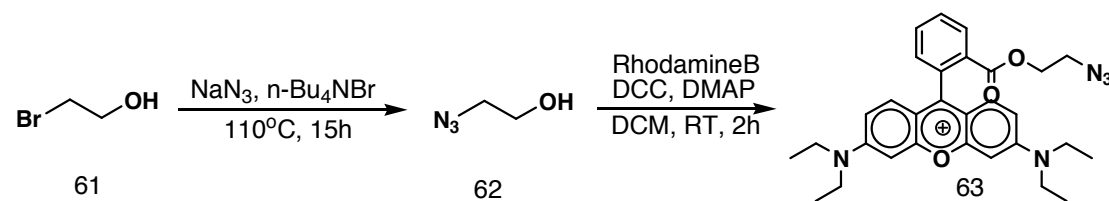
Scheme20. Synthetic scheme of rhodamine B azidoethyl ester (**63**).

Rhodamine B azidoethyl ester can be used to label ethynyl-proteins via a click reaction. We first tested such possibility by reaction this azido-rhodamine B with 5-ethynyltryptophan ethyl ester (Scheme 21). HR-ESIMS and ^{1}H-NMR confirmed the formation of the product. Applying a similar reaction condition, rhodamine B azidoethyl ester successfully labeled ethynyl-BSA (data not shown). This reaction has been applied to test the number of accessible

ethynyl groups on BSA by calculating the UV absorption ratio between the characteristic

peaks of rhodamine B and BSA.

a.

b.
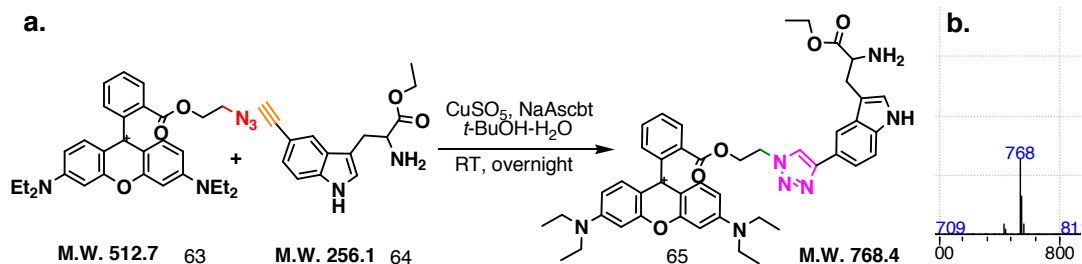
Scheme21 a-b. a. Huisgen 1,3-dipolar cycloaddition reaction between azido-rhodamine B (**63**) and 5-ethynyltryptophan ethyl ester (**64**); b. ESI-MS spectrum of the clicked product (**65**).

2.3.2.4. Synthesis of Monosulfone-OH Rhodamine B Ester (69) as a Disulfide-Reactive

Chromophore

The fourth rhodamine B derivative prepared was a disulfide-reactive chromophore, called

monosulfone-rhodamine B. This dye can be used to label proteins possessing accessible

disulfide bonds. In addition, by determining the mass increase of the labeled protein, the

number of accessible disulfide bonds could be assessed. The synthesis of this dye started with

the reaction of bisulfone-NHS ester with ethanolamine to afford monosulfone-ethanol. This

intermediate subsequently reacted with rhodamine B via Steglich esterification to give

monosulfone-rhodamine B in the form of a violet solid (Scheme 22). This chromophore was

successfully used to label SST which is a disulfide bridged cyclic peptide known as a growth

hormone-inhibiting hormone (data not shown).
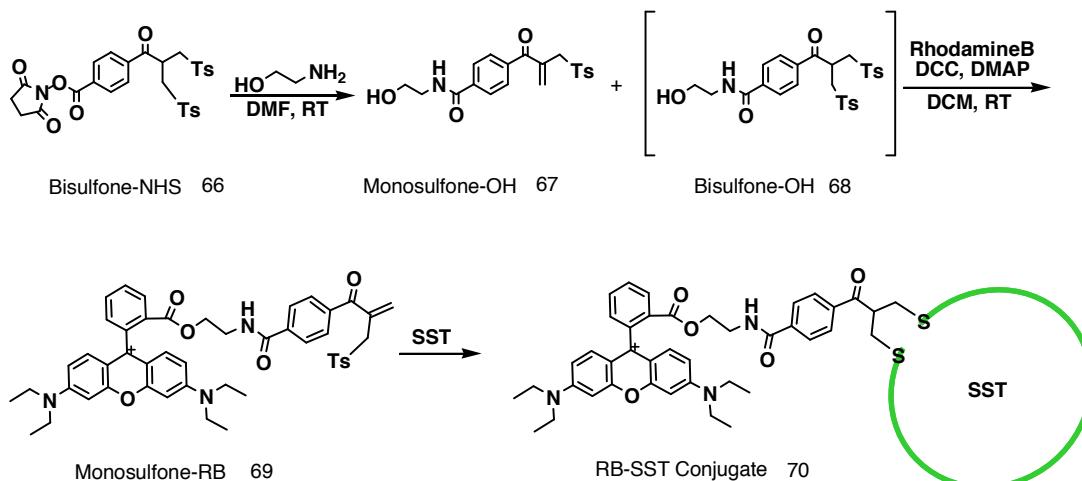
Scheme22. Synthesis of monosulfone-OH rhodamine B ester (**69**) for bioconjugation with the cyclic peptide SST (**51**).

In conclusion, four rhodamine B derivatives were prepared via one step Steglich esterification. They proved to be efficient protein probes for labeling functionalized proteins/peptide such as azido-antibodies and SST. This strategy is concise (one step only) and safe (it avoids the usage of dangerous reagents) compared to previously reported methods. We hope that our results could serve as bases to achieve rhodamine B bearing attractive functionalities suitable for bioconjugation reactions.

Chapter 3. EXPERIMENTAL SECTION

^{1}H and ^{13}C-NMR spectra were recorded with a Bruker AV 300, a Bruker AMX 500 or a Bruker DRX 500 NMR spectrometer operating at 300 MHz for ^{1}H and 75.48MHz for ^{13}C NMR. Chemical shifts were reported in ppm (δ scale) relative to the solvent signal, and coupling constant (J) values were reported in hertz (Hz). High-resolution mass spectra (HRMS) were recorded on a Finnigan MAT95XL-T mass spectrometer by direct infusion of the solution of each compound using electrospray ionization (ESI) in positive or negative mode. LC-ESIMS were recorded by a Schmadzu LC-20AD/SPD-20A/SIL-20AC/LCMS-2010EV using the C-18 column and MeCN/H_2O as eluent. UV spectra were recorded by a BioTEK SYNERGY 4 Microplate Reader. MALDI-ToF-MS was determined by a Bruker auto-Flex spectrometer. Chemicals were purchased from Sigma-Aldrich, Merck and Regent and used without further purification unless otherwise stated. Dowex 50 W×4 cationic exchange resin was purchased from Sigma-Aldrich (Fluka, Na+ form, 200-400 mesh, category No. 44473). And the PreHPLC machine was distributed from Gilson.

3.1. Triflic Azide (14) Solution

3.1.1. PURIFICATION OF TRIFLIC ANHYDRIDE (13)

It is a prerequisite that triflic anhydride (**13**) is purified prior to its usage in order to remove any hydrolyzed impurities. Since **13** is very sensitive and considered to be a very hazardous compound, the following "vacuum transfer" apparatus was set up, including a 50 ml RBF, a Schlenk flask, a stirring bar and a 3-way tube with 2 stopcocks.
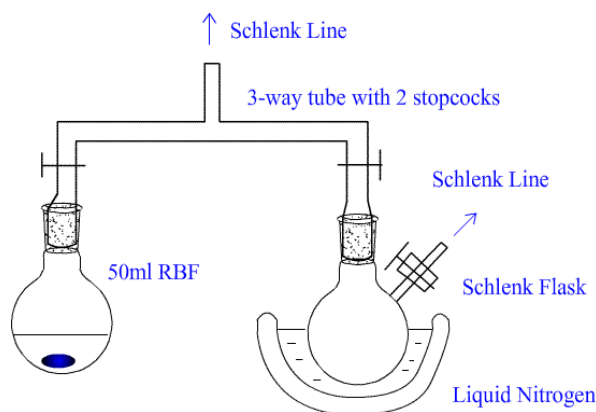
Figure17. The "vacuum transfer" setup includes a 50 ml RBF, a Schlenk flask, a stirring bar and a 3-way tube with 2 stopcocks.

Briefly, the whole apparatus was thoroughly heated and dried in vacuum in advance. Then P_2O_5 (0.25 g) and triflic anhydride (2.5 ml) were placed in a 50 ml RBF under Ar atmosphere and stirred for 10 hours overnight. The originally yellowish triflic anhydride became milk yellow upon adding the P_2O_5 then completely faded away after stirring overnight. Using such "vacuum transfer" methodology, triflic anhydride was solidified in the Schlenk flask which was cooled by a liquid nitrogen bath. About 1.7 ml freshly purified triflic anhydride was obtained with a yield of 68 %. This freshly prepared liquid was immediately used in the following reaction in order to avoid any delay.

3.1.2. TRIFLIC AZIDE (14) SOLUTION

The preparation of triflic azide (14) solution, Boc-AHA (16) and AHA (2) is a continuous and multiple step process. In favor of a clear description of the whole protocol, each single step has been numbered in the following.

[1] Prepare a 100 ml Schlenk flask with NaN_3 (1.3 g, 20 mmol) and H_2O (3.2 ml) inside, and then place it in an ice bath.

[2] Add dropwise the purified triflic anhydride (0.664 ml, 4 mmol) into the flask which contains sodium azide from step 1. Then remove the ice bath and allow the reaction mixture to stir at RT for two hours.

[3] At the completion of the reaction in step 2, there will be a layer settled at the bottom of the

RBF which indicates the formation of triflic azide. Transfer the reaction mixture to a big test tube and rinse the RBF with DCM (2 ml) which is combined with the reaction mixture in the big test tube subsequently. After separation, the DCM layer in the big test tube is isolated and the remaining aqueous phase is extracted twice more with DCM (2 ml × 2) using the same procedure.

[4] Combine the organic layers and wash them in a big test tube with brine (6.4 ml) and a saturated NaHCO$_3$ solution. Shake and allow the emulsion to separate. Collect the DCM layer by a dropper which is the triflic azide solution in DCM. As this solution is unstable at RT, it should be used in the following diazotransfer reaction as soon as possible.

According to the standard operation procedure (SOP) of NaN$_3$ waste, all aqueous waste generated in this reaction should be disposed accordingly as NaN$_3$ is a highly poisonous for the environment, particularly for aqueous organisms.

3.2. Boc-AHA (16)

[5] During the two hour period in step 3, add Boc-Dab (436 mg, 2 mmol), K$_2$CO$_3$ (416 mg, 3 mmol) and a catalytic amount of CuSO$_4$·5H$_2$O (5 mg, 0.02 mmol) into a clean 100 ml RBF equipped with a stir bar. To this RBF add 6.4 ml of H$_2$O to dissolve the reactants with ultrasonication. Finally, add 12.8 ml of MeOH to this RBF and shake; the resultant solution will be basically clear. Set this flask aside until the triflic azide is prepared.

[6] Add the TfN$_3$/DCM solution from step 5 dropwise to the flask prepared in Step 3. The reaction solution will be, in the main, clear with a light-blue color. The reaction mixture is stirred at RT overnight.

[7] The following day, remove the MeOH and DCM via rotary evaporation at 40 $^{\circ}$C leaving behind about 6.4 ml of aqueous solution. This solution will be bright blue and basic (pH 9~10)

[8] Add 6 M HCl by a dropper slowly to the solution until the pH reaches 6. Note that upon

the pH reaching 9 the solution begins to turn green; bubbles are released upon the pH reaching 8 and finally the solution will become slightly cloudy and yellowish with a little of green upon the pH reaching 6.

[9] Dilute the solution with 24 ml of phosphate buffer composed of 0.25 M KH_2PO_4 and 0.25 M K_2HPO_4 (pH 6.2). Readjust the pH of the entire mixture to 3 by adding 6 M HCl dropwise.

[10] Extract the aqueous solution in a separatory funnel twice with EtOAc (2×24 ml).

[11] Combine the EtOAc layers and wash them with brine (12 ml). Dry over anhydrous $MgSO_4$ (2.4 g) for 30 min, then filtrate and remove the EtOAc by rotary evaporation at 40 °C. This will afford 609 mg (on average) of semisolid as the product.

^{1}H-NMR (CDCl$_3$, 300MHz): δ 5.49(d, J=7.2Hz, br, 1H), 4.22(m, 1H), 3.35(t, J=6.7Hz, 2H), 1.95(m 2H), 1.35(s, 9H)
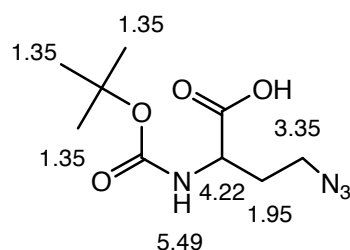
ESI-MS (-): $C_9H_{15}N_4O_4^-$ calcd. 243.1, found 243.0

3.3. AHA (**2**)

The obtained Boc-AHA (**16**) is treated with con. HCl to yield the target amino acid AHA (**2**).

[12] Dissolve the semisolid product from step 11 in 1.24 ml of concentrated HCl and stir at RT for 1 h.

[13] Dilute the solution with 11.2 ml of water.

[14] Pack an ion exchange column by adding Dowex cationic exchange resin (16.7 g) to a glass chromatography column of a suitable size under the assistance with MQ-H$_2$O.

[15] Condition the resin by washing it three times with 1 M NH$_4$OH (72 ml), once with DI-H$_2$O (890 ml), once with 1 M HCl (72 ml) and once with DI-H$_2$O (890 ml) again. The

column eluent should be slightly acidic.

[16] Add the solution from step 13 to the column, collect the flowthrough and pass it over the column twice more. If any AHA can be detected in the flowthrough by TLC/ninhydrin, a higher loading of cationic resin is required.

[17] Wash the column with water (595 ml) until the pH of the eluent remains constant (pH 6-7).

[18] Elute the amino acid from the column by adding 1 M $NH_3.H_2O$ (149 ml).

[19] Remove the aqueous NH_4OH using a rotary evaporator at 55 ^{o}C leaving behind 223 mg (on average) of colorless clumpy crystal as the product in a yield of 77 %.

^{1}H-NMR (D_2O, 300MHz): δ 3.85(t, 6.2Hz, 1H), 3.60(t, J=6.6Hz, 2H), 2.15(m, 2H)
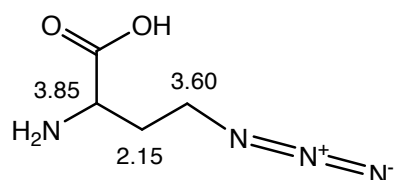
^{13}C-NMR (300MHz, D_2O): δ 174.0, 52.8, 47.5, 29.6
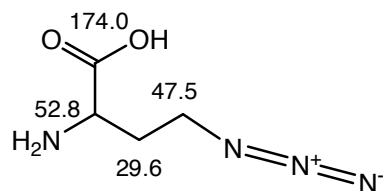
ESI-MS (-): $C_4H_7N_4O_2^-$ cacld. 143.1, found 143.0

IR (KBr/cm^{-1}): 2119.07 (azido group)

3.4. The Reaction Conditions in Reductive Amination of *N,N*-Dimethyloxamic Acid (**19**) to afford "N-Valine" (**20**)

The first reaction condition: *N,N*-dimethyloxamic acid (117 mg, 1 mmol), NaCNBH$_3$ (69.3 mg, 1.1 mmol), NH$_4$OAc (770 mg), dry MeOH (dried over 4 Å molecular sieve, 4.76 ml, 0.21 M) and 1 g activated 4 Å molecular sieve were added into a 25 ml Schlenk flask under Ar atmosphere. This reaction suspension was allowed to stir at RT overnight and was monitored

by TLC and LC-ESIMS.

The second reaction condition: an additional 1 g of 4 Å molecular sieve was added and the reaction temperature was increased to 60 ^{o}C and stirred for an additional 24 h.

The third reaction condition: this reaction was performed in a similar way but equal equivalent of $NaBH_4$ or NH_4Cl or NH_4OH (aq.) were used to replace the corresponding reagents.

3.5. The Reaction Condition in Bromination of *N*-Boc-Gly-O*t*Bu (**21**)

In a dried 25 ml-Schlenk flask *N*-Boc-Gly-O*t*Bu (100 mg, 0.433 mmol), NBS (81.3 mg, 0.457 mmol, freshly recrystallized from water), and dry solvent ($CHCl_3$, DCM, benzene, chlorobenzene or tetrachloroethylene, 1.75 ml) were combined under Ar atmosphere. The Schlenk flask was cooled in a 9 ^{o}C water bath and irradiated by a 300-W tungsten lamp. The temperature of the water bath was maintained by continuous addition of ice. The temperature inside the Schenk flask was maintained at 15 ^{o}C. This was controlled by a thermometer put inside. The reaction was allowed to proceed for 1 h and monitored by LC-ESIMS and ^{1}H-NMR.

3.6. Phthaloglycine (**24**)

Into a 250 ml RBF equipped with a Dean-Stark trap and a condenser, glycine (7.51 g, 100 mmol), phthalic anhydride (14.2 g, 100 mmol), Et_3N (1.6 ml, 10 mmol) and toluene (60 ml) were placed. The reaction mixture was heated to reflux for 6 h with azeotropic removal of water. Toluene was removed by evaporation under reduced pressure. The residue was dissolved in 150 ml H_2O and acidified by adding con. HCl (2 ml, 22 mmol). The product was collected by filtration, washed with water (2×10 ml) and dried in *vacuo* to give 19.3 g of phthaloglycine (**24**) as a white solid in a yield of 94.2 %.

[1]**H-NMR** ($CDCl_3$, 300MHz): *δ* 7.89(m, 2H), 7.75(m, 2H), 4.47(s, 2H)
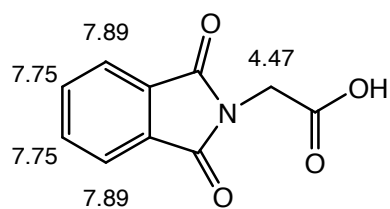
ESI-MS (-): $C_{10}H_6NO_4^-$ calcd. 204.0, found 203.9

3.7. The Reaction Condition in Bromination of Phthaloglycine (**24**) via the Hell-Volhard-Zelinsky Reaction

Into a 25 ml-Schenk flask, phthaloglycine (2.05 g, 10 mmol) and $SOCl_2$ (4 ml) were added under Ar atmosphere. The reaction mixture was heated to reflux for 1.5 h and excess of $SOCl_2$ was removed under vacuum. The ^{1}H-NMR spectrum confirmed the formation of the intermediate product—phthalimidoacetyl chloride (**25**). Liquid bromine (0.64 ml) was added to the RBF in an Ar atmosphere and the reaction mixture was heated to reflux overnight and monitored by ^{1}H-NMR.

3.8. Phthaloglycine Methyl Ester (**26**)

Into a 25 ml Schlenk flask equipped with a condenser, phthaloglycine (2.05 g, 10 mmol) and $SOCl_2$ (4 ml) were placed and refluxed for 1.5 h. Then, MeOH was added into the flask and heated to gentle reflux. After excess of $SOCl_2$ and MeOH were removed under vacuum the residue was taken up in the EtOAc. The organic layer was washed with sat. $NaHCO_3$, water and brine, dried over anhydrous Na_2SO_4, filtrated and the EtOAc was evaporated under reduced pressure. The product was dried in *vacuo* to afford 200 mg of white solid as phthaloglycine methyl ester (**26**) in a yield of 6.7 %.

^{1}H-NMR ($CDCl_3$, 300MHz): δ 7.82(m, 2H), 7.70(m, 2H), 4.40(s, 2H), 3.71(s, 3H)
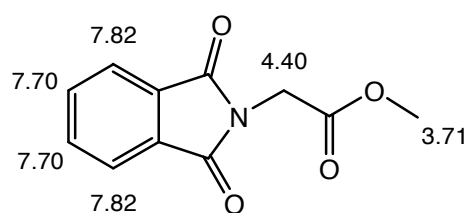
3.9. The Reaction Conditions in the Bromination of Phthaloglycine Methyl Ester (26) Using NBS/NaOH

The first reaction condition: into a 2-necked 25 ml RBF equipped with a condenser and a stir bar phthaloglycine methyl ester (**26**) (110 mg, 0.5 mmol), NBS (93.4 mg, 0.525 mmol, recrystallized), NaOH (120 mg, 3 mmol) and dry DMF (2.5 ml) were added in an Ar atmosphere to form a yellow-colored mixture. The reaction mixture was stirred at 80 $^{\circ}$C overnight and monitored by LC-ESIMS.

The second reaction condition: the temperature was subsequently increased to 150 $^{\circ}$C and stirred for 22 h. The reaction was monitored by LC-ESIMS.

3.10. The Reaction Condition in the Bromination of Phthaloglycine Methyl Ester (26) Using Br₂/NaH

Into a thoroughly dried RBF equipped with a stir bar NaH (60 % in mineral oil, 1.2 equiv.) and dry THF (1 ml) were added and the suspension was cooled to 0 $^{\circ}$C by an ice-water bath. A solution of phthaloglycine methyl ester (50 mg) in dry THF (0.5 ml) was injected dropwise to this reaction suspension and stirred at 0 $^{\circ}$C for 20 min. Then, a solution of Br$_2$ (15.8 μl) in dry THF (0.5 ml) was added dropwise into the reaction mixture and stirred at 0 $^{\circ}$C for 20 min. The reaction was monitored by LC-ESIMS as well as ^{1}H-NMR.

3.11. Biotin-LC-Maleimide (32)

Into a small Eppendorf tube 6-biotinamindohexanoic acid NHS ester (biotin-LC-NHS, 8.0 mg, 17.6 μmol), maleimidoethylamine trifluoroacetic acid salt (4.9 mg, 19.4 μmol), *N,N*-diisopropylethylamine (DIEA) (3.67 μl, 21.1 μmol) and dry DMF (176 μl, 0.2 M) were combined. The Ependorff tube was capped and fixed in a shaker via a conical flask and was shaken at RT for 6 h until the LC-ESIMS indicated that the full conversion of the starting

material to the product had been accomplished. The reaction mixture was dissolved in MeCN/DMF/H$_2$O to form a homogeneous solution and the target compound (**32**) was purified by PrepHPLC (λ_{max} 205 nm) to afford 8.2 mg of white solid product in a yield of 97 %.

^{1}H-NMR (DMSO-d^6, 300MHz): δ 7.85(t, *J*=5.9Hz, 1H), 7.72(t, *J*=5.4Hz, 1H), 7.00(s, 2H), 6.41(s, 1H), 6.38(s, 1H), 4.30(t, *J*=6.3Hz, 1H), 4.12(t, *J*=5.3Hz, 1H), 3.43(t, *J*=5.8Hz, 2H), 3.18(q, *J*=5.9Hz, 2H), 3.10(m, 1H), 2.98(m, 2H), 2.82(dd, *J$_1$*=12.3Hz, J$_2$=5.1Hz, 1H), 2.57(d, *J*=12.3Hz, 1H), 2.03(t, *J*=7.3Hz, 2H), 1.94(t, *J*=7.4Hz, 2H), 1.14~1.62(m, 12H)
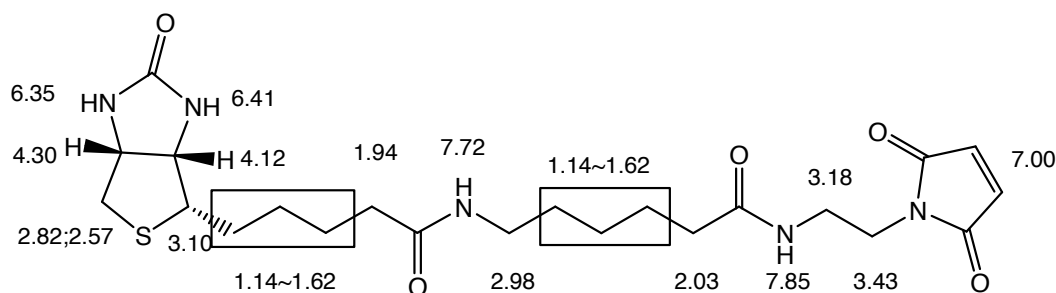
^{13}C-NMR (DMSO-d^6, 300MHz): δ 172.22, 171.72, 170.96, 162.62, 134.43, 60.99, 59.15, 55.35, 38.26~48.34(overlap with DMSO peaks), 37.20, 36.62, 35.72, 35.24, 35.17, 28.92, 28.15, 27.97, 26.04, 25.26, 24.76
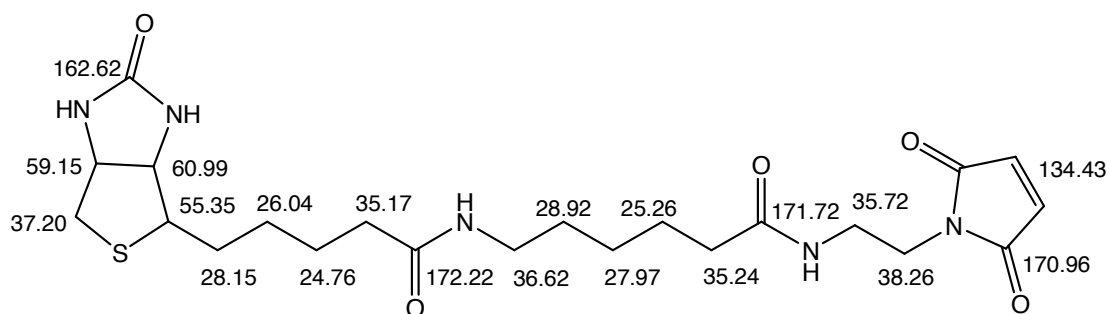
HR-ESIMS (+): C$_{22}$H$_{33}$N$_5$NaO$_5$S$^+$ calcd. 502.2095, found 502.2113

3.12. Monobiotinylation of HSA (**33**) via the Cysteine-Targeting Approach

To a degassed solution of HSA (68 mg, 1 μmol) in a phosphate buffer (60 ml, 50 mM, pH 6.5), a freshly prepared aqueous solution of TCEP (0.22 ml, 1 mM, 0.22 μmol) was added in an Ar atmosphere and the reaction mixture was stirred for ten minutes at RT. Then a solution

of biotin-LC-maleimide (1 mg) in DMSO (1 ml, 2 μmol, 200 % excess) was added and the reaction mixture which was stirred at RT for 1 h in an Ar atmosphere. The non-reacted biotin-LC-maleimide was removed after five times ultra-filtration with water and size-exclusion chromatography (Bio-Gel P 30, MQ-H$_2$O). Thereafter, unmodified HSA was removed via affinity chromatography by applying the Softlink Avidin Resin (Promega) according to its published protocol.

MALDI-ToF-MS (native HSA): m/e 66619 [M+H]$^+$

MALDI-ToF-MS (maleimide-HSA): m/e 67098 [M+H]$^+$

3.13. The Reaction Condition in the Transamination of RNase A (**18**) to Afford keto-RNase A (**35**)

The following reagents were prepared prior to the start of the experiment: phosphate buffer (50 mM, pH 6.5), 10 μM RNase A (1.35 mg/ml in phosphate buffer), and 20 mM of pyrodoxal-5-phosphate (4.94 mg/ml in phosphate buffer). These reagents were freshly prepared and stored at 4 $^{\circ}$C prior to use.

0.5 ml of RNase A solution was mixed with an equal volume of pyrodoxal-5-phosphate solution in a 1.5 ml Eppendorf tube and incubated at 37 $^{\circ}$C for 44 h. The resulting solution was dialyzed against MQ-H$_2$O to remove excess pyrodoxal-5-phophate. The purified keto-RNase A solution was kept at 4 $^{\circ}$C before proceeding to the next step. Theoretically, the concentration of the activated RNase A solution contained 50 μM of protein (0.675 mg/ml) since the solution volume doubled after the dialysis.

MALDI-ToF-MS (SA): calcd. 13.7 kD, found ~14 kD (br)

3.14. The Reaction Condition in the Hydrazone Formation of Keto-RNase A (**35**)

The freshly prepared keto-RNase A was treated with an equal volume of 4-hydrazinylbenzoic acid solution (100 mM in HOAc/H$_2$O (1:1)) at 37 ^{o}C for 24 h. The resulting solution was dialyzed against MQ-H$_2$O and the product was monitored by MALDI-ToF-MS (SA).

3.15. The Reaction Condition in the Aldol Reaction of Keto-RNase A (35)

The freshly prepared keto-RNase A was treated with 80 μl of DMSO solution and 4-nitrobenzaldeyde (10 mM) at 37 ^{o}C overnight. The removal of small molecules from the reaction mixture was achieved by gel filtration (Sephadex-G15 gel) before being monitored by MALDI-ToF-MS (SA).

3.16. The Reaction Condition in the Pictet-Spengler Reaction of Keto-RNase A (35)

25 μl of the freshly prepared keto-RNase A (18) was combined with 7.2 μl of 5-bromotryptophan methyl ester solution (68 μM in DMSO). The reaction mixture was incubated at 37 ^{o}C for 24 h and dialyzed against MQ-H$_2$O and monitored by MALDI-ToF-MS (SA).

3.17. 5-Bromotryptophan Methyl Ester (43)

Into a 25 ml 2-neck RBF 5-bromotryptophan (46.4 mg, 145 μmol) and 5 ml of MeOH were added under anhydrous conditions. The RBF was placed into an ice bath until the temperature of the reaction mixture was cooled to 0 ^{o}C. Thereafter, SOCl$_2$ (1 ml) was added dropwise and the ice water bath was detached. The reaction mixture was stirred at RT for 24 h until the TLC indicated the complete consumption of the starting material. SOCl$_2$ and MeOH were removed under high vacuum and dried overnight in *vacuo* to give 50.9 mg of solid as the product in a yield of 95 %.

^{1}H-NMR (DMSO-d^6, 300MHz): δ 11.45(s, 1H), 8.73(s, 3H), 7.71(s, 1H), 7.35(d, *J*=8.1Hz, 1H), 7.32(s, 1H), 7.18(d, *J*=8.1Hz, 1H), 4.20(s, 1H), 3.64(s, 3H), 3.44(s, 2H).

^{13}C-NMR (DMSO-d^6, 300MHz): δ 170.0, 135.3, 129.2, 127.1, 123.9, 120.8, 114.0, 111.7, 106.7, 53.0, 39.1~40.8(overlapped with DMSO-d^6), 26.2

ESI-MS (+): $C_{12}H_{14}BrN_2O_2^+$ calcd. 297.0, found 296.9

3.18. Biotin-LC-OH (71)

Biotin-LC-NHS (**31**) (3.6 mg, 0.009 mmol) was dissolved in DMSO-d^6 (0.5 ml) and transferred into a NMR tube. Subsequently, ethanolamine (30 mg, 0.49 mmol) was added to the solution and the solution was shaken vigorously. The NMR tube was capped, sealed by parafilm, fixed on a conical flask via a robber stopper and shaken at RT for 24 h. ^{1}H-NMR and ESI-MS indicated the complete consumption of the starting material which can be seen from the disappearance of the NHS ester signal.

Excess of ethanolamine and DMSO-d^6 were removed in *vacuo* overnight at RT. The residue was dissolved in 1 ml of MeOH/MeCN (1:1) and further purified by Prep-HPLC (t_R=10.1~12.2 min, λ=205 nm) to give 1.6 mg of white solid as the product (**71**) in a yield of 44 %.

^{1}H-NMR (DMSO-d^6, 300MHz): δ 7.78(t, *J*=5.6Hz, 1H), 7.73(t, *J*=5.6Hz, 1H), 6.41(s, 1H), 6.35(s, 1H), 4.68(sb, 1H), 4.30(td, J_1=5.1Hz, 1H), 4.12(td, J_1=4.6Hz, 1H), 3.33(m, 2H),

3.10(m, 2H), 3.08(m, 1H), 2.99(q, J=6.5Hz, 2H), 2.82(dd, J_1=12.3Hz, J_2=5.0Hz, 1H), 2.59(d, J=12.5Hz, 1H), 2.043(t, J=7.4Hz, 2H), 2.037(t, J=7.1Hz, 2H), 1.17-1.61(m, 12H)
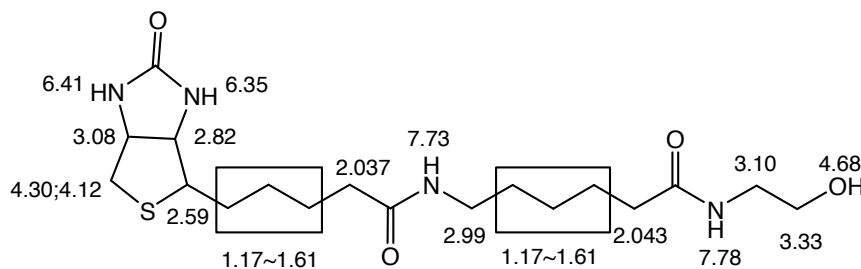
^{13}**C-NMR** (DMSO-d^6, 300MHz): δ 172.5, 172.1, 163.1, 61.4, 60.3, 59.6, 55.8, 41.8, 39.1~40.8(overlapped by DMSO peak), 38.7, 35.7, 35.6, 29.3, 28.6, 28.4, 26.5, 25.7, 25.4
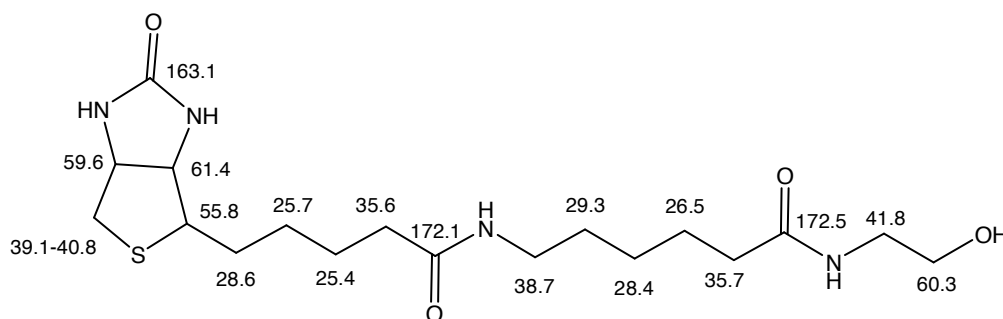
**ESI-MS** (+): $C_{18}H_{33}N_4O_4S^+$ calcd. 401.2, found 401.2[M+H]$^+$

3.19. Biotin-PEG-COO*t*Bu (47)

(+)-Biotin (93.4 mg, 0.393 mmol) and EDC·HCl (110 mg, 0.573 mmol) were introduced into an aluminum foil coated 5 ml RBF and were dried in *vacuo*. Then dry DMF (3.8 ml, 0.05 M) was injected at 0 °C followed by the dropwise injection of *tert*-butyl-12-amino-4,7,10-trioxadodecanoate (80 % purity, 53 μl, 0.191 mmol). The reaction mixture was stirred at 0 °C for 0.5 h and then allowed to warm to RT and stirred for 40 h. TLC indicated the complete consumption of the starting material and the LC-ESIMS spectrum revealed the formation of the target product (**47**).

The reaction mixture was concentrated at RT under vacuum and taken up in DCM. The organic phase was washed three times with DI-H$_2$O, once with 0.1 M HCl to remove the trace of basic impurities, once with DI-H$_2$O and finally, mixed with saturated Na$_2$CO$_3$ solution.

Later, the organic layer was separated and the aqueous phase was extracted twice by DCM. The Organic layers were combined, washed with brine, dried over anhydrous Na_2SO_4, filtrated and concentrated. Final purification via silica gel flash chromatography (CHCl$_3$:MeOH 5:1, I$_2$ stain) gave 51.4 mg of light yellow solid as the product (**47**) in a yield of 67 %. All steps were performed under the exclusion of light since biotin is considered to be a light sensitive species.

^{1}H-NMR (DMSO-d^6, 300MHz): δ 6.71(t, br, 1H), 5.85(s, 1H), 5.49(s, 1H), 4.41(m, 1H), 4.26(m, 1H), 3.63(t, J=6.2Hz, 2H), 3.53(m, 8H), 3.45(t, J=5.5Hz, 2H), 3.27(q, J=5.6Hz, 2H), 3.12-3.18(m, 1H), 2.87(dd, J_1=12.6, J_2=5.0, 1H), 2.63(d, J=12.8Hz, 1H), 2.40(t, J=6.2Hz, 2H), 2.13(t, J=7.2Hz), 1.48-1.73(m, 4H), 1.41(s, 9H), 1.37(t, J=7.2Hz, 2H)
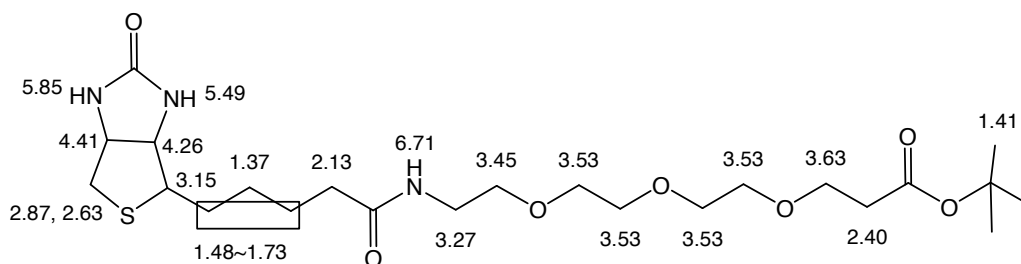
^{13}C-NMR (DMSO-d^6, 300MHz): δ 173.69, 171.64, 164.20, 80.76, 70.90, 70.85, 70.72, 70.18(two carbons), 67.36, 62.27, 60.71, 56.21, 40.96, 39.64, 36.89, 36.21, 28.90, 28.84, 28.14, 26.24
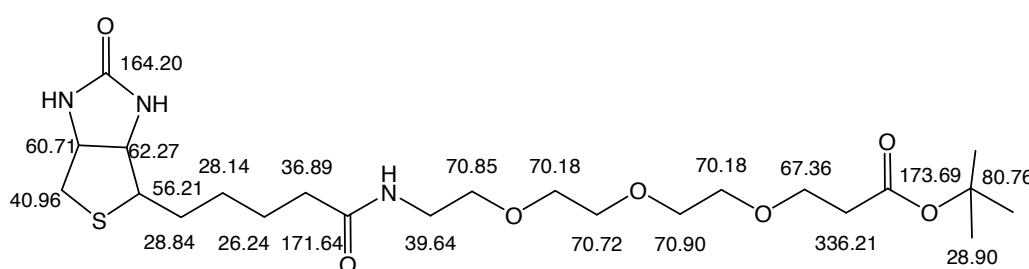
ESI-MS (+): $C_{23}H_{41}N_3NaO_7S^+$ calcd. 526, found 526

3.20. Biotin-PEG-COOH (**48**)

Biotin-PEG-COO*t*Bu (**47**) (75 mg, 0.147 mmol) was dissolved in DCM (1.73 ml) and transferred into a 25 ml RBF under Ar atmosphere. The solution was cooled in an ice-water

bath and TFA (0.87 ml) was added dropwise. The ice-water bath was removed and the RBF was wrapped with aluminum foil to exclude light. The reaction mixture was allowed to warm to RT and incubated overnight. DCM and TFA were removed under high vacuum leaving behind a white solid as the product **48** in a quantitative yield.

^{1}H-NMR (DMSO-d[6], 500MHz): δ 7.81(s, 1H), 6.49(s, br, 2H), 4.31(m, 1H), 4.13(m, 1H) 3.60(t, J=6.3Hz, 2H), 3.49(s, br, 8H, PEG H), 3.39(t, J=6.0Hz, 2H), 3.18(q, J=8.2Hz, 2H), 3.09(m, 1H), 2.82(dd, J_1=12.6Hz, J_2=5.05, 1H), 2.58(d, J=12.6Hz, 1H), 2.53(t, J=6.3Hz, 2H), 1.60(m, 1H), 1.49(m, 2H), 1.43(m, 1H), 1.29(m, 2H).

^{13}C-NMR (DMSO-d[6], 500MHz): 172.58, 172.10, 162.70, 69.69, 69.68, 69.60, 69.53, 69.14, 66.21, 61.04, 59.21, 55.37, 39.75, 38.44, 35.06, 34.72, 28.15, 28.00, 25.22

HR-ESIMS (+): $C_{19}H_{33}N_3NaO_7S^+$, cacld. 470.1931, found 470.1952

3.21. Biotin-PEG-NHS (**49**)

Into an aluminum foil coated 5 ml RBF equipped with a stir bar, Biotin-PEG-COOH (40.0 mg, 0.71 mmol), EDC·HCl (26 mg, 0.136 mmol) and *N*-hydroxysuccinimide (NHS) (13.6 mg, 0.12 mmol) were added. This mixture was dried under vacuum for 45 min before dry DMF (0.46 ml, 0.15 M) and DIEA (15.6 μl, 0.096 mmol) were introduced. The reaction solution was allowed to stir at RT for 23 h until TLC indicated the disappearance of the starting material and the formation of a new spot (CHCl$_3$: MeOH, R$_f$(SM) 0.4, R$_f$(P) 0.6, I$_2$ stain).

Thereafter, the solvent was removed at RT under vacuum and the viscous liquid was partitioned in CHCl$_3$/H$_2$O (3/1, V/V). The organic layer was collected by a dropper. The aqueous layer was extracted six times by CHCl$_3$. The organic layers were combined and washed with brine, dried under vacuum, purified by silica gel flash chromatography (CHCl$_3$: MeOH, 5:1, I$_2$ stain) to produce 24.1 mg of white solid **49** in a yield of 62 %.

^{1}H-NMR (MeOD, 500MHz): δ 4.49(t, J=6.2Hz, 1H), 4.31(t,m,1H), 3.84(t, J=5.95, 2H), 3.62(t,m, 8H),3.54(t, J=5.3Hz,2H), 3.35(t, J=5.3Hz,2H), 3.21(m, 1H), 2.9(m, 2H), 2.84(s,

4H), 2.69(m, 2H), 2.22(t, *J*=7.3Hz, 2H), 1.73(m, 1H), 1.65(m, 2H), 1.59(m, 1H), 1.44(m, 2H).

^{13}C-NMR (MeOD, 500MHz): δ 176.12, 171.69, 168.59, 166.08, 71.63, 71.61, 71,47, 71.26, 70.57, 66.86, 63.36, 61.62, 56.97, 41.04, 40.37, 36.74, 32.95, 29.48, 26.82, 26.52.

DEPT135 (MeOD, 500MHz): δ (+) 63.36, 61.62, 56.97; (-) 71.63, 71.61, 71.47, 71.26, 70.57, 41.04, 40.37, 36.74, 32.95, 29.74, 29.48, 26.83, 26.52.

ESI-MS (+): $C_{23}H_{36}N_4NaO_9S^+$ calcd. 567.2, found 567.3

3.22. Monobiotinylation of RNase A (18), Lysozyme C (50) and SST (51) via the Lysine-Targeting Approach

3.22.1. MATERIALS

Some of the materials which were required for this reaction were prepared as described below and have been numbered individually.

[1] Dissolve biotin-PEG-NHS (**49**) (2.66 mg) in MQ-H_2O (133 ml, 0.02 mg/ml) to prepare a 36.3 μM reagent solution. This should be freshly prepared in order to avoid the hydrolysis of this NHS ester in aqueous solution. Keep at 4 $^{\circ}$C.

[2] Dissolve RNase A (1.28 mg) in a 0.1 M PBS buffe (pH 7.2, 1.28 ml) to afford a 72.5 μM protein solution (1 mg/ml). Keep at 4 $^{\circ}$C.

[3] Dissolve lysozyme C (2.44 mg) in a 0.1 M PBS buffer (pH 7.2, 2.3 ml, 1.06 mg/ml) to afford a 72.5 μM protein solution. Keep at 4 $^{\circ}$C.

[4] Similarly, dissolve SST in a PBS buffer to afford a 0.12 mg/ml peptide solution whose molar concentration is also 72.5 μM. Keep at 4 $^{\circ}$C.

3.22.2. MODIFICATION PROTOCOL

The lysine targeting approach for the site-directed modification of proteins/peptide contains four major steps as detailed below.

3.22.2.1. Step I: Bioconjugation

In this step, mono-biotinylated proteins/peptide were formed. This is a protein mixture in buffer.

Into an autoclaved Falcon tube containing the protein/peptide solution (1 ml), biotin-PEG-NHS (**49**) solution (1 ml) was added portion-wise at a rate of 8.4 μl/min within 2 h. The reaction mixture was incubated at RT for an additional 1.5 h and the solution was subjected to MALDI-ToF-MS for analysis (Figure 13). According to the MALDI-ToF-MS spectrum, all three proteins/peptide were monofunctionalized, thus a mixture of native and modified proteins/peptide were formed in a PBS buffer solution.

3.22.2.2. Step II: Quenching

The reaction mixture contained unreacted biotin-PEG-NHS (**49**) which will interfere with the mono-modified protein after affinity chromatography. In order to deactivate any unreacted biotin-PEG-NHS (**49**), the following quenching step was applied.

Ethanolamine is a water soluble primary amine which can readily react with any NHS esters in aqueous media. Hence, an ethanolamine solution was employed in this step. The ethanolamine solution (2.67 mg/ml, 0.5 μl) in MQ-H$_2$O (1.33 μg, 1.2 equiv.) was added to each protein sample and incubated for 4 h. Ethanolamine will quench the unreacted biotin-PEG-NHS conjugation reagent.

3.22.2.3. Step III: Affinity Chromatography

Mono-biotinylated proteins/peptide were "fished" up from the reaction solution thus achieving the separation between modified and non-modified proteins/peptide. Since the presence of any biotin-containing reagent in the protein solution will block the active binding pockets of monomeric Avidin resin, they must be removed before affinity chromatography via e.g. gel filtration.

Procedure for Gel Filtration:

[1] Activate the Sephadex G15 gel (1.5 kD MWCO, 12 g) by dispersing it in MQ-H_2O (45 ml) in a 45 ml Falcon tube overnight.

[2] Divide the gel suspension into three equal portions and used it to pack three gel filtration columns.

[3] Elute each gel filtration column by MQ-H_2O until the eluent no longer displays any UV absorption.

[4] Load each reaction solution (1 ml) onto a separated column and eluted with MQ-H_2O. Collect around eight fractions of 1 ml each.

[5] Assay all fractions under UV absorption and combine the protein fractions.

Procedure for Affinity Chromatography:

Prior to affinity chromatography, the following materials/reagents were prepared:

  1. Resin: Promega Softlink Avidin Resin (cat. No. V2011, 1ml).

  2. Columns: 3 pieces of 1ml syringes without plungers and with a small piece of cotton at the bottom to hold the bed.

  3. Buffer A: fresh 0.1 M NaPO$_4$ buffer (pH 7.0), prepared by dissolving 156 mg of NaH$_2$PO$_4$·2H$_2$O and 222 mg of Na$_2$HPO$_4$ in 25.6 ml of MQ-H_2O.

  4. Buffer B: fresh PBS buffer (0.1 M) obtained by diluting a 1 M PBS buffer 10-times in MQ-H_2O.

  5. Regeneration Buffer: 10 % acetic acid in water (V/V).

  6. Elution Buffer: fresh 5 mM biotin/PBS solution prepared by dissolving 12.2 mg of biotin in a 10 ml Buffer B.

  7. Blocking Buffer: fresh 5 mM biotin/NaPO$_4$ (0.1 M) prepared by dissolving 12.2 mg of biotin in a 10 ml Buffer.

  8. Preservative Solution: 20 % EtOH in water, used for the storage of the column.

The whole affinity purification procedure consists of the following four major steps.

Step 1: block the irreversible binding pockets of the Avidin resin using a Blocking Buffer.

1. Pack the resin column with 0.25 ml of bead for each column.

2. Equilibrate the three columns with Buffer A.

3. Wash each column with Blocking Buffer at least two column volumes.

4. Stop flow for 15 min.

Step 2: regenerate the column

1. Wash each column with eight column volumes of Regeneration Buffer.

2. Wash each column with eight column volumes of Buffer A.

3. After the pH of eluent increased to pH 6.8, stop the flow for a minimum of 30 min to allow the Avidin to refold.

4. Equilibrate the column with Buffer B.

Step 3: purification

1. Load a 2 ml sample in each column.

2. Wash with 5 column volumes of Buffer B.

3. Elute each biotin-PEG-protein/peptide with Elution Buffer.

4. Collect fractions for each column with 0.4 ml per fraction.

5. Assay each fraction by applying a BioTEK microplate reader to identify the modified proteins/peptide solution.

Step 4: storage

1. Elute each column with Preservative Solution, and then stop the column.

2. Store the column at 4 ^{o}C.

3.22.2.4. Step IV: Gel Filtration

Since the isolated protein is dissolved in the Elution Buffer which contains biotin molecules and salt, they should be removed from the protein prior to storage. This can be achieved via gel filtration as described before in this section followed by freeze drying. Lyophilized protein powder was properly stored at -80 ^{o}C.

3.23. Rhodamine B NHS Ester (RB-NHS) (55)

In a typical reaction, rhodamine B (3.0 g, 6.28 mmol) and NHS (0.75 g, 6.52 mmol) were placed into a 250 ml 3-neck RBF equipped with a condenser and a 50 ml dropping funnel. Then dry MeCN (100 ml) was added into the flask and it was heated to 45 °C. A solution of DCC (1.5 g, 7.35 mmol) in dry MeCN (50 ml) was added dropwise. After stirring at 45 °C for 1 h, the reaction solution was cooled to RT and stirred for 21 h more. Thereafter, the reaction was cooled at 4 °C overnight to precipitate *N,N'*-dicyclohexylurea (DCU) from the reaction mixture. The reaction mixture was filtrated, removed of MeCN under reduced pressure and dried in *vacuo* to give 3.66 g of dark green crystals as the product in a quantitative yield.

¹H-NMR (CDCl$_3$, 300MHz): δ 8.41(dd, J_1=7.1Hz, J_2=1Hz, 1H), 7.97(td, J_1=7.1Hz, J_2=1Hz, 1H) 7.81(td, J_1=7.3Hz, J_2=0.8Hz, 1H), 7.47(dd, J_1= 7.1Hz, J_2=0.8Hz, 1H), 7.07(d, J=10.1Hz, 2H), 6.86(dd, J_1=10.1Hz, 2H), 6.86(d, 2H), 2.76(s, 4H)
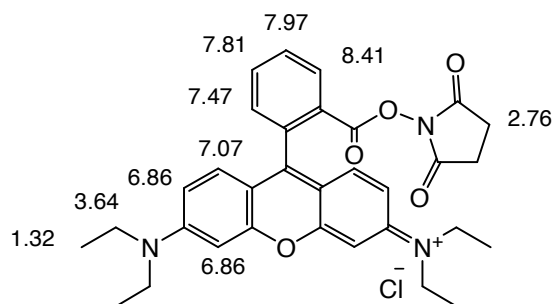
¹³C-NMR (CDCl$_3$, 300MHz): δ 168.7, 160.6, 157.6, 155.5, 155.5, 134.8, 134.2, 131.5, 130.9, 130.8, 130.6, 125.3, 114.4, 113.2, 96.4, 46.1, 25.5, 12.6
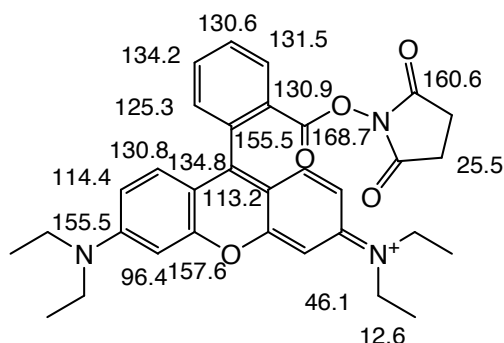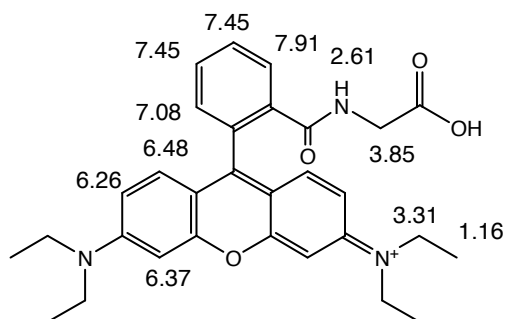

ESI-MS (+): C$_{32}$H$_{34}$N$_3$O$_5$$^+$ calcd. 540.3, found 540.3

3.24. Rhodamine B-Glycine Amide (72)

In a typical reaction, glycine (45.0 mg, 0.6 mmol) was dissolved in 1.8 ml of H_3BO_3 buffer (pH 8.5, 0.05 M) and the aqueous solution was stirred while a solution of RB-NHS (36 mg, 0.0626 mmol) in MeCN (1.8 ml) was added slowly to the buffer solution. After continuous stirring at RT for 4 h, MeCN was removed under reduced pressure. A violet precipitate was obtained and further recrystallized from $EtOH/H_2O$ to afford 3.8 mg of rhodamine B glycine amide (**72**) in the form of a bright violet crystal with a yield of 11 %.

^{1}H-NMR (CDCl$_3$, 300MHz): δ 7.91(dd, J_1=6.2Hz, 1H), 7.45(m, 2H), 7.08(m, 1H), 6.48(d, J=8.9Hz, 2H), 6.36(d, J=2.3Hz, 2H), 6.26(dd, J_1=8.9, J_2=2.3Hz, 2H), 3.85(s, 2H), 3.31(q, J=7.0Hz, 8H), 2.61(sb, 1H), 1.16(t, J=7.0Hz, 12H)
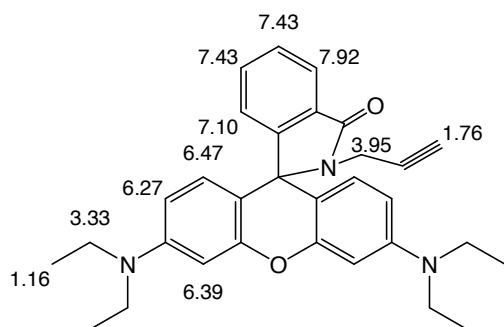

ESI-MS (+): $C_{30}H_{34}N_3O_4^+$ calcd. 500.25, found 500.34

3.25. Pseudo-Rhodamine B Propargyl Amide (**57**)

The synthetic route of pseudo-rhodamine B propargyl amide is similar to rhodamine glycine amide (section 3.2.1). The reaction scheme is shown below. A pink white powder was obtained as the product after recrystallization.
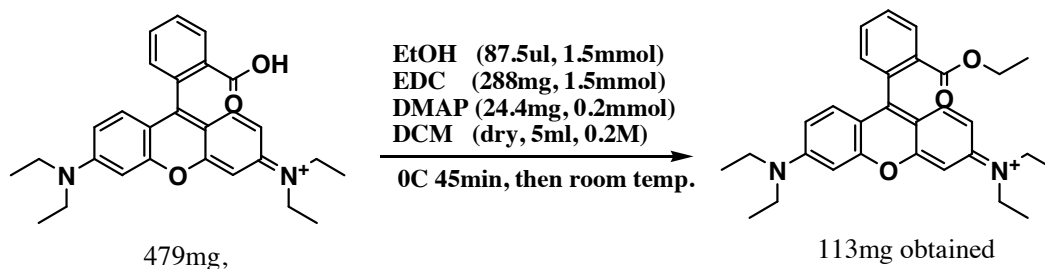

^{1}H-NMR (CDCl$_3$, 300MHz): δ 7.92(m, 1H), 7.43(m, 2H), 7.10 (m, 1H), 6.47 (d, J=8.9Hz, 2H), 6.39 (d, J=2.6Hz, 2H), 6.27 (dd, J_1=8.9Hz, J_2=2.6Hz, 2H), 3.95 (d, J=2.6Hz), 3.33 (q, J=7.1Hz, 8H), 1.76(d, J=2.5Hz, 1H), 1.16 (d, J=7.1Hz, 12H).
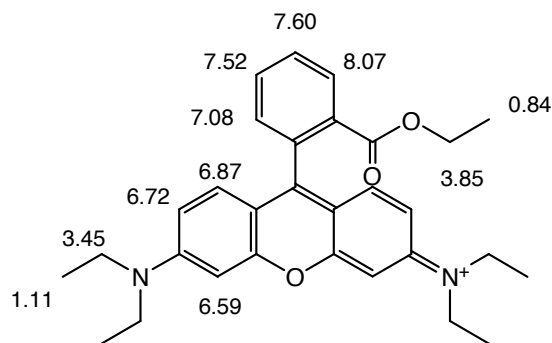
ESI-MS (+): $C_{31}H_{34}N_3O_2^+$ cacld. 480.3, found 480.3
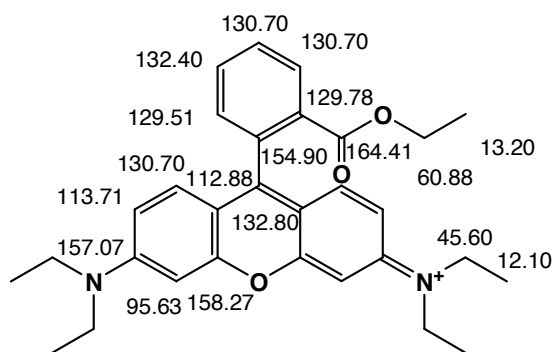
3.26. Rhodamine B Ethyl Ester (**60**)

The synthesis procedure of rhodamine B ethyl ester is similar to the synthesis of rhodamine B propargyl ester (section 3.2.4). The quantity and equivalent of each reagent is given by the reaction scheme below. Starting from 479 mg of rhodamine B (1 mmol), 112.9 mg of violet solid was obtained as the product with a 22 % yield.

H-NMR (CDCl₃, 300MHz): δ 8.07 (d, *J*=7.6Hz, 1H), 7.60 (t, *J*=7.3Hz, 1H), 7.52 (t, *J*=7.5Hz, 1H), 7.08 (d, *J*=7.0Hz, 2H), 6.87 (d, *J*=9.1Hz, 2H), 6.72(d, *J*=9.4Hz, 2H), 6.59 (s, 2H), 3.85 (q, *J*=7.0Hz, 2H), 3.45 (q, *J*=6.7Hz, 8H), 1.11 (t, *J*=6.3Hz, 12H), 0.84 (q, *J*=7.0Hz, 3H)

¹³C-NMR (CDCl₃, 300MHz): δ 164.41, 158.27, 157.07, 154.90, 132.80, 132.40, 130.70, 129.78, 129.51, 113.71, 112.88, 95.60, 60.88, 45.60, 13.20, 12.10
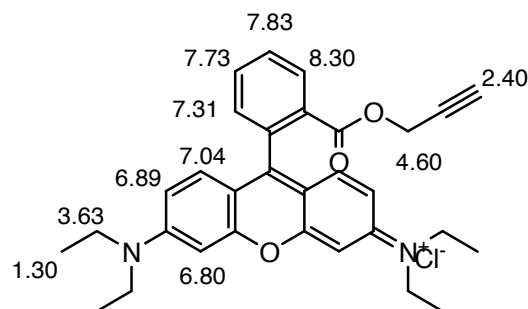
LC-ESIMS (10%~100% MeCN, 20min, 256nm): t_R 7.76min, 98.5% purity; $C_{30}H_{35}N_2O_3^+$ calcd. 471.26, found 471.15
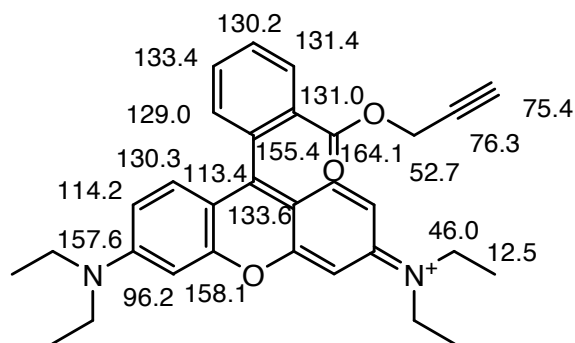
3.27. Rhodamine B Propargyl Ester (**73**)

Into a 25 ml 2-neck RBF equipped with a condenser, rhodamine B (1.91 g, 4 mmol), DCC (989 mg, 4.8 mmol), DMAP (97.6 mg, 0.8 mmol), DCM (121 ml) and 2-propynol (0.28 ml) were combined. The reaction mixture was stirred at RT for 2 h. Then the reaction mixture was filtrated to remove DCU, washed twice with 6 M HCl and brine, dried over anhydrous magnesium sulfate and filtrated. The solvent was removed under reduced pressure and the product was dried in *vacuo* to afford 9 mg of **73** in the form of dark violet crystals with a yield of 27 %.

¹H-NMR (CDCl₃, 300MHz): δ 8.30(dd, J_1=7.7Hz, J_2=1.1Hz, 1H), 7.83(td, J_1=7.5Hz, 1.2Hz, 1H), 7.73(td, J_1=7.7Hz, J_2=1.3Hz, 1H), 7.31(dd, J_1=7.6Hz, J_2=1.1Hz, 1H), 7.04(d, J=9.5Hz, 2H), 6.89(dd, J_1=9.5Hz, J_2=2.5Hz, 2H), 6.80(d, J=2.3Hz, 2H), 4.60(d, J=2.5Hz, 2H), 3.63(q, J=7.1Hz, 8H), 2.40(t, J=2.5Hz, 1H), 1.30(t, J=7.1Hz, 12H).

13C-NMR (CDCl$_3$, 300MHz): δ 164.1, 158.1, 157.6, 155.4, 133.6, 133.4, 131.4, 131.0, 130.3, 130.2, 129.0, 114.2, 113.4, 96.2, 76.3, 75.4, 52.7, 46.0, 12.5

HR-ESIMS (+):C$_{31}$H$_{33}$O$_3$N$_2$$^+$ cacld. 481.2486, found 481.2471

**EA**: C$_{31}$H$_{33}$O$_3$N$_2$Cl·2H$_2$O calcd. C 67.32%, H 6.74%, N 5.06%, Cl 6.41%; found C 67.46%, H 6.71%, N 5.14%, Cl 6.94%;

UV-Vis (H$_2$O): λ_{max} 559 nm, ε 1.43×10^5 M^{-1}·cm^{-1}

3.28. Azidoethanol (**62**)

Into a 5 ml RBF 2-bromoethanol (375 mg, 3.03 mmol), NaN$_3$ (257 mg, 6.1 mmol) and the phase transfer catalyst *n*-Bu$_4$NBr (25 mg, 0.075 mmol) were combined. The mixture was heated to 110 ^{o}C and stirred for 15 h. Afterwards the reaction mixture was cooled to RT and diethyl ether (1 ml) was added. The precipitate which contained NaBr, NaN$_3$ and *n*-Bu$_4$NBr was removed by filtration and the crude solid was washed three times with diethyl ether. The ether layers were combined and dried in *vacuo* to afford 36.2 mg of yellowish liquid as the product **62** in a yield of 13.6 %. During the vacuum drying, the product was also lost which resulted in a poor yield. This can be circumvented by using rotavap instead of high vacuum for the drying.

1H-NMR (CDCl$_3$, 300MHz): δ 3.78 (t, *J*=5.0, 2H), 3.34(t, *J*=5.0, 2H)
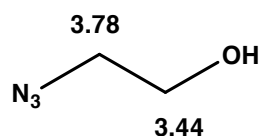
The ^{1}H-NMR of the starting material—2-bromoethanol—was also determined for the comparison with the product. The up-field shift clearly suggested the substitution of the bromo group by an azido group.

^{1}**H-NMR** (CDCl$_3$, 300MHz): δ 3.91 (t, 2H), 3.54 (t, 2H)
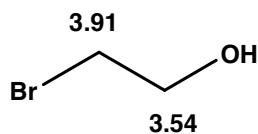
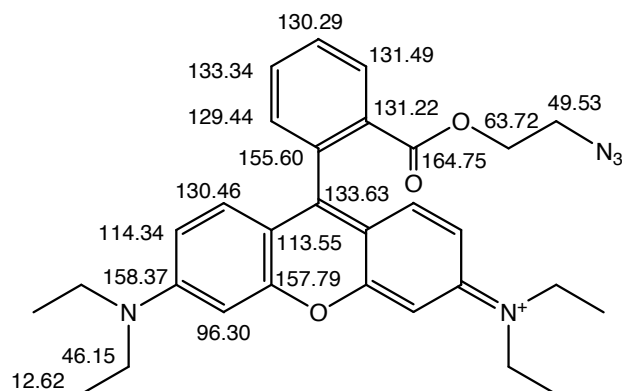## 3.29. Rhodamine B Azidoethanol Ester (**63**)

2-Azidoethanol (**62**, 26.2 mg, 0.30 mmol) was dissolved in anhydrous DCM to form a stock solution. Rhodamine B (141 mg, 0.30 mmol), DCC (97 mg, 0.45 mmol) and 4-dimethylaminopyridine (DMAP, 9.2 mg, 0.075 mmol) were placed into a 10 ml RBF which was wrapped by an aluminum foil and capped by a rubber stopper. The reaction mixture was further dried in vacuum and charged with Ar through a needle. Then 2-azidoethanol/DCM solution was injected by a syringe at 0 ^{o}C. Afterwards the reaction mixture was stirred, allowed to warm up slowly to RT and further stirred for 6 h. Subsequently, DCM was added and the reaction mixture was washed by water (2×), HCl (5×, 6 M), brine (1×), and concentrated to 3 ml. The organic layer was cooled at 4 ^{o}C to precipitate DCU from the solution. After filtration, the filtrate was evaporated and dissolved in 3.2 ml MeCN (HPLC grade). Half of the obtained sample was purified via PrepHPLC (λ=256 nm) to afford 24.3 mg of **63** in the form of a violet crystal with a yield of 43 %.

^{1}**H-NMR** (CDCl$_3$, 500MHz): δ 8.31 (dd, J_1=7.6Hz, J_2=too tiny, 1H), 7.81(dd, td, J_1=7.6Hz, J_2=too tiny, 1H), 7.75(td, J_1=8.2Hz, J_2=1.3Hz, 1H), 7.32(dd, J_1=7.6Hz, J_2=too tiny, 1H), 7.07(d, J=9.5Hz, 2H), 6.92 (dd, J_1=9.5Hz, J_2=2.6Hz, 2H), 6.82(d, J=2.5Hz, 2H)
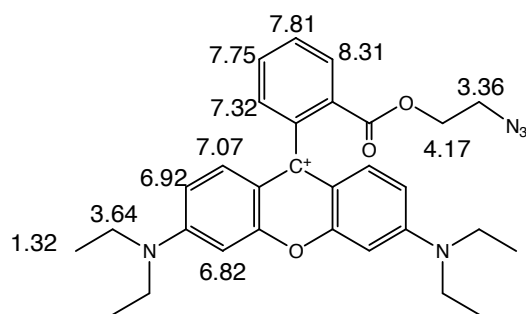
^{13}C-NMR (CDCl$_3$, 500MHz): δ 164.75, 158.37, 157.79, 155.60, 133.63, 133.34, 131.49, 131.22, 130.29, 129.44, 114.34, 113.55, 96.30, 63.72, 49.53, 46.15, 12.72

**ESI-HRMS(+):** C$_{30}$H$_{34}$N$_5$O$_3$$^+$ calcd. 512.2656, found 512.2639

UV-Vis (H$_2$O): λ_{max} 559 nm; ε 1.48×10^5 M^{-1}.cm^{-1}

3.30. 5-Ethynyltryptophan Rhodamine B Conjugate (65)

Rhodamine B azidoethanol ester (**63**, 1.8 mg, 6.9 μmol) was added to a test tube equipped with a rubber stopper in an Ar atmosphere. Thereafter, 45 μl of a solution of 5-ethynyltryptophan (**64**) /*t*BuOH (1.8 mg, 7.0 μmol) was injected, followed by the addition of 90 μl H$_2$O. Then, 0.7 μl of a freshly prepared solution containing CuSO$_4$ (18 μg, 0.07 μmol) and sodium ascorbate (69 μg, 0.35 μmol) were injected. After the test tube was detached from the Schlenk line, the rubber cap was sealed by parafilm and the test tube was coated with aluminum foil to exclude light. This tube was fixed in a suitable conical flask and shaken at RT overnight until LC-ESIMS indicated full conversion of the starting material (**63**) to the product (**65**). The reaction mixture was dissolved in 1.5 ml of MeCN/H$_2$O and purified by

PrepHPLC (λ 254 nm) to afford the pure product.

^{1}H-NMR (CDCl$_3$, 500MHz): δ 10.31(s, 1H), 8.37 (m, 1H), 8.03(s, 1H), 7.91(s, 1H), 7.78(m, 1H), 7.76(m, 1H), 7.49(m, 1H), 7.43(m, 1H). 7.15(m, 1H), 6.90(m, 2H), 6.62(m, 2H), 6.60(m, 2H), 4.55(s, br, 4H), 4.14(t, J=11.8Hz, 2H), 4.00(s, br, 2H), 3.69(s, br, 1H), 3.21(s, br, 2H), 1.29(m, 3H), 1.21(t, J=11.0Hz, 12H).
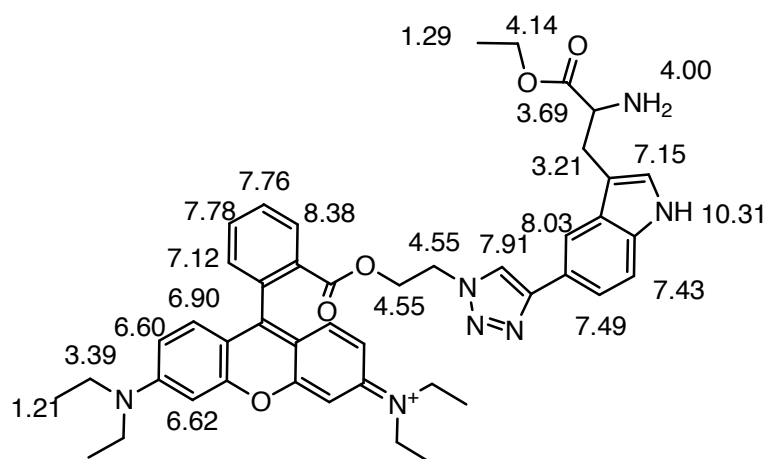
**HR-ESIMS** (+): C$_{45}$H$_{50}$N$_7$O$_5^+$ calcd. 768.3868, found 768.3877

UV-Vis (H$_2$O): λ_{max} 564 nm, ε 2.12$\times$10^5 M^{-1}.cm^{-1}

3.31. Bisulfone-OH (**68**) and Monosulfone-OH (**67**)

Into a small Ependorff tube 4-(3-tosyl-2-(tosylmethyl)propanoyl)benzoic acid *N*-hydroxysuccinimidyl ester (bisulfone-NHS, **66**) (20.2 mg, 34 μmol), DMF (340 μL, 0.1 M) and ethanolamine (30.5 μl, 51 μmol) were combined and the Ependorff tube was sealed and fixed on a conical flask with a rubber cap. The reaction mixture was shaken at 180 r/min at RT overnight (~18 h). Then, the reaction mixture was taken up in MeCN/H$_2$O/MeOH and purified by PrepHPLC to afford 4.2 mg of **68** in the form of a white solid with a yield of 23 %. De-toslated product—*N*-(2-hydroxyethyl)-4-(2-(tosylmethyl)acryloyl)benzamide (monosolfone-OH, **67**) was also isolated and characterized.

^{1}H-NMR (CDCl$_3$, 300MHz): δ 7.78(d, J=8.6Hz,2H), 7.70(d, J=8.0Hz, 2H), 7.69(d, J=8.2Hz, 4H), 7.36(d, J=7.9Hz, 4H), 4.37(m, 1H), 3.87(t, J=4.9Hz, 2H), 3.67(d, J=4.6Hz, 2H), 3.62(dd, J_1==14.3, J_2=6.6Hz, 2H), 3.49(dd, J_1=14.3Hz, J_2=5.9Hz, 2H), 2.48(s, 6H)
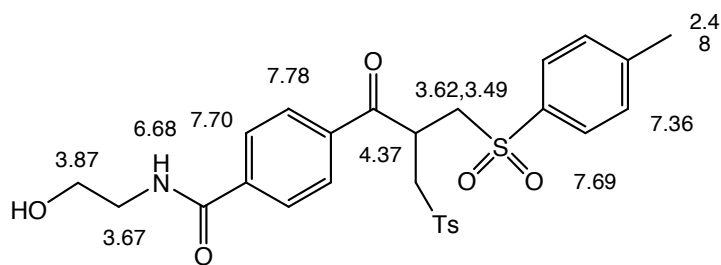
^{13}C-NMR (CDCl$_3$, 300MHz): δ 195.20, 167.02, 145.58, 139.07, 136.48, 130.22, 128.81, 127.55, 62.06, 55.68, 42.78, 35.60, 21.71
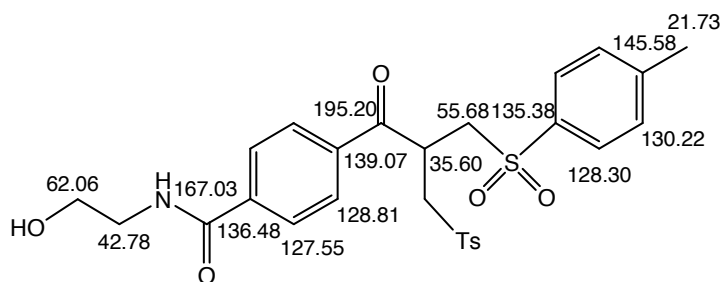
ESI-MS (+): C$_{27}$H$_{30}$NO$_7$S$_2^+$ calcd. 544, found 544

^{1}H-NMR (CDCl$_3$, 500MHz): δ 7.84(d, J=8.2Hz, 2H), 7.78(d, J=8.2Hz, 2H), 7.73(d, J=8.2Hz, 2H), 7.32(d, J=7.6Hz, 2H), 6.67(s, br, 1H), 6.29(s, 1H), 5.99(s, 1H), 3.87(t, J=2.2Hz, 2H), 3.66(m, 2H), 2.41(s, 3H)
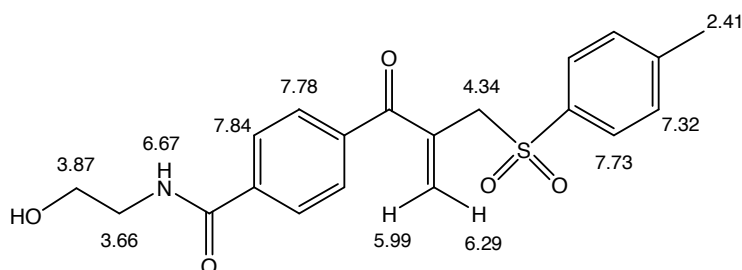
^{13}C-NMR (CDCl$_3$, 500MHz): δ 194.11, 167.33, 145.13, 138.89, 137.76, 135.93, 135.84, 134.36, 129.90, 129.79, 128.34, 126.98
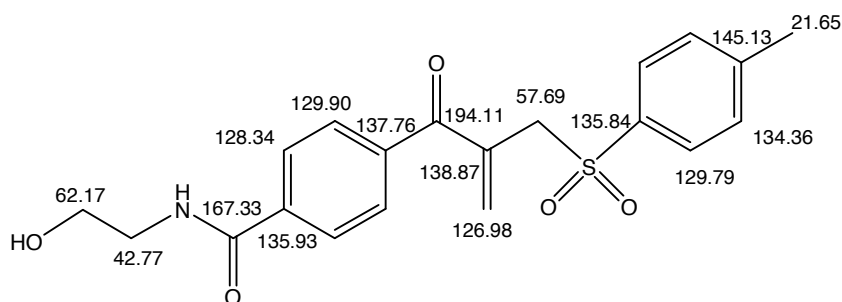
**HR-ESIMS** (-): $C_{20}H_{20}NO_5S^-$ cacld. 386.1068, found 386.1082

3.32. Monosulfone-OH Rhodamine B Ester (**69**)

Into a 2 ml test tube, monosulfone-OH (**67**) (1.6 mg, 4.13 μmol), rhodamine B (2.97 mg, 6.2 μmol), DCC (1.28 mg, 6.2 μmol), and DMAP (0.13 mg, 1.1 μmol) were combined and this mixture was dried in *vacuo* for half an hour. Later, dry DCM (100 μl, 0.04 M) was introduced via a syringe at 0 °C in an Ar atomsphere. Then the small test tube was coated with aluminum foil and stirred at 0 °C for half an hour. Later, this reaction mixture was shaken at RT overnight until the LC-ESI-MS indicated the complete conversion of the starting material to the product. The reaction mixture was mixed with DCM/MeCN and the DCU precipitate was removed by centrifugation. Finally purification via PrepHPLC (MeCN/H$_2$O) afforded a violet crystal as the product **69**.

LC-ESIMS (MeCN/H$_2$O)) of Monosulfone-OH Rhodamine B Ester

    a. HPLC spectrum
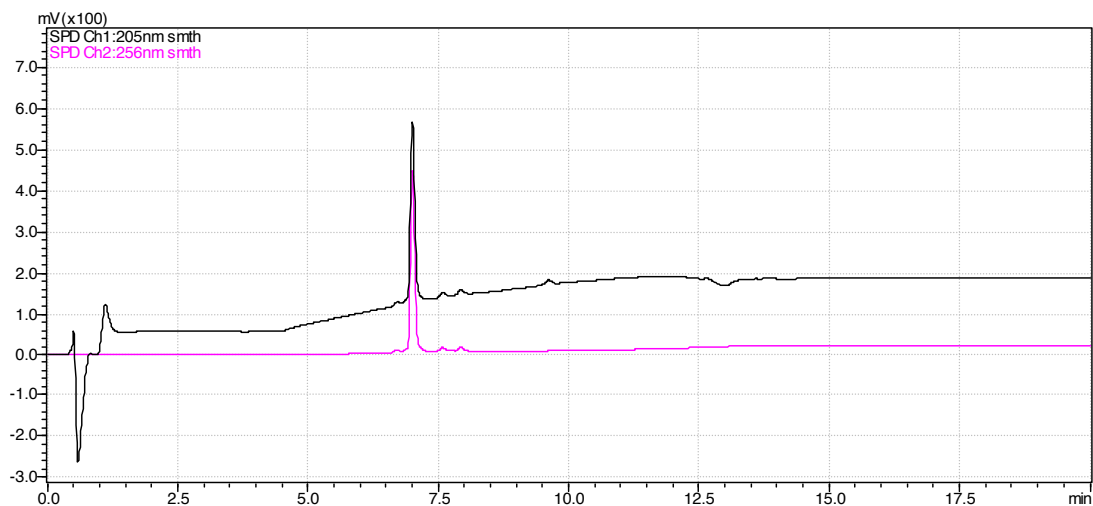
    b. ESI-MS (+): $C_{48}H_{50}N_3O_7S^+$ calcd. 812, found 812
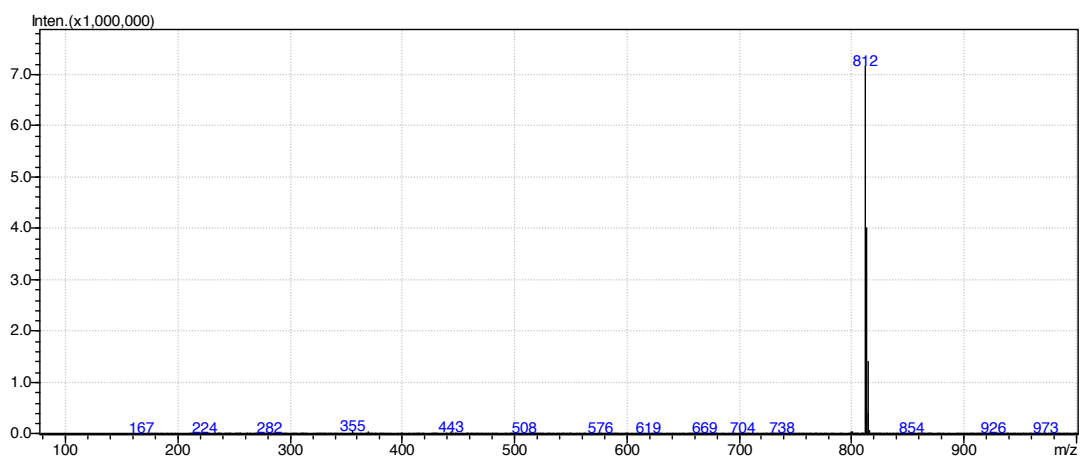
Chapter 4. CONCLUSIONS

In conclusion, one biology and two chemistry approaches toward site-directed protein modifications were developed. In conjunction with these protein modification strategies, three novel reactive rhodamine B dyes were prepared for straight forward modification of proteins and peptides.

The biology means, termed *selected pressure incorporation* (SPI) is considered a novel and practical approach towards site-directed protein functionalization. Two major steps were involved. The first step was based on the preparation of the unnatural amino acid analogue, called NCAA such as AHA. This stage involved chemistry efforts since a new NCAA needs to be rationally designed and synthesized since the molecular structure of the NCAA represents a vital factor determining its translational activity (whether it can be recognized by the cell's translational machinery) and whether it can be synthesized relatively easily. The second step is the incorporation of a NCAA into a protein sequence which requires the selection of an auxotrophic bacteria strain (bacteria which is unable to self-generate certain amino acids and must uptake them from environment in order to survive) and expression of the corresponding protein.

In this thesis, AHA, a surrogate of methionine was successfully prepared and incorporated into RNase A to yield bis-azido RNase A modified in a site-directed fashion. AHA was prepared with a reliably enhanced yield compared to recently reported protocols. The average yield of this compound was 77 % which is higher than the published protocol with variable yields between 34 % and 59 %.[14] Subsequently, this NCAA was successfully incorporated into RNase A whose methionine residues were all replaced by AHA. Two AHA were located on the protein surface, namely AHA 1 and AHA 41, as identified by nanoLC-HR(ESI)MS/MS. Compared to another biological means—*nonsense codon suppression*—this biological

approach is less time consuming as it does not require the preparation of a new transfer-RNA: aminoacyl-tRNA-synthetase pair. This approach could be valuable to produce a variety of other azido-functionalized proteins and apply these protein variants to achieve activity-controlled bio-macromolecules or other protein level machineries.

Two chemistry methods toward site-directed protein modifications were developed which allow fast and straight forward production of reasonable protein quantities compared with the SPI method and therefore, these approaches could be considered attractive alternatives. The first chemistry method targets a cysteine residue. Proteins that possess a single accessible cysteine residue on their surface were modified, such as serum albumin proteins, some antibody monomers etc. In this thesis, the bioconjugation reagent—biotin-PEG-maleimide—was developed and was successfully applied to achieve mono-biotinylated HSA via this cysteine-targeting approach. Since HSA bears a single cysteine residue at its surface and is considered suitable for this labeling strategy. However, not all proteins display an unpaired cysteine and therefore this approach still suffers from some inherent limitations.

A more attractive chemistry method developed in this thesis focuses on the modification of lysine residues which in principle displays a broader application scope than targeting cysteines since lysine is typically located at the protein surface. Based on the 3-dimentional structure of proteins, the solvent accessibility of each lysine residue is calculated via MOE and the results suggest that the most accessible lysine can be addressed in a site-directed manner. The interaction of lysine with other residues in close vicinity via e.g. H-H bonding is also expected to affect the site-directionality. The bioconjugation reagent—biotin-PEG-NHS (**49**)—was developed for this application featuring a biotin moiety as an affinity purification tag, a PEG linker which contributes to water-solubility and an amine-reactive NHS ester group. This reagent was successfully applied to yield mono-biotinylated RNase A, lysozyme C and SST as demonstrated by a MALDI-ToF-MS study.

Four major steps are involved in this lysine-targeting approach. First, biotin-PEG-NHS is applied to the protein/peptide solution as the limiting reagent in a portion-wise fashion which is called the *bioconjugation* step. Thus, mono-modified protein is generated but still containing the native protein as well. Second, this protein solution is treated with a quenching reagent to deactivate the unreacted conjugation reagent—the *quench* step. The remaining biotin-containing small molecules which interfere with affinity chromatography are facilely removed via subsequently gel filtration. In the third step, the mono-biotinylated protein is separated from the protein mixture using monomeric Avidin resin and the unreacted native protein is recovered as well—the *affinity chromatography* step. In the last step, the affinity purified protein solution is subjected to a second gel filtration in order to remove biotin molecules and salts which originate from affinity chromatography eluent—the *gel filtration* step. We believe that this approach has broad applicability toward achieving mono-modified proteins and we also envision that a variety of bioorthogonal groups can be introduced onto protein surfaces via this approach thus enabling further click reactions.

Finally, reactive rhodamine B dyes were prepared via one-step Steglich esterification and they were applied to decorate corresponding functionalized proteins/peptide. Three attractive rhodamine B derivatives were prepared including rhodamine B propargyl ester for labeling azido-proteins, rhodamine B azidoethyl ester for targeting ethynyl-proteins and monosulfone-OH rhodamine B ester as a disulfide-reactive chromophore. Their applicability for labeling corresponding type of proteins/peptide was demonstrated by experiments as well. Our efforts provide a new approach to prepare novel reactive rhodamine B as useful chromophores to decorate proteins to allow further photophysical investigations.

PUBLICATIONS

The work presented in this master thesis has been incorporated in the following published or submitted journal articles:

[1] <u>X. Chen</u>, K. Muthoosamy, A. Pfisterer, B. Neumann and T. Weil. Exploiting the intrinsic regioselectivity of lysine residues toward site-specific modification of native proteins and peptide: native chemical incorporation. *J. Am. Chem. Soc.* **2011** (submitted).

[2] <u>X. Chen</u>, Q Wu, L. Henschke, T. Weil. One-step practical synthesis of useful rhodamine probes with an ester bond linkage. *Dyes Pigments* **2011** (submitted).

[3] A. Pfisterer, K. Eisele, <u>X. Chen</u>, M. Wagner, K. Müllen and T. Weil. Bioactive unnatural somatostatin analogues through bioorthogonal iodo and ethynyl-disulfide intercalators. *Chem. Eur. J.* **2011** (online)

[4] M. Lillich, <u>X. Chen</u>, T. Weil. H. Barth and J. Fahrer. Streptavidin-conjugated C3 protein mediates the delivery of mono-biotinylated RNase A into macrophases. *J. Control. Release* **2011** (submitted).

BIBLIOGRAPHY

[1] G U Nienhaus, J Wiedenmann. Structure, dynamics and optical properties of fluorescent proteins: perspectives for marker development. *ChemPhysChem* **2009** (10) 1369

[2] C Bosisio, V Quercioli, M Collini, L D'Alfonso, G Baldini, S Bettati, B Campanini, S Raboni, G Chirico. *J. Phys. Chem. B* **2008** (112) 8806

[3] A E Miller, A J Fischer, T Laurence, C W Hollars, R J Saykally, J C Lagarias, T Huser. Single-molecule dynamics of phytochrome-bound fluorophores probed by fluorescence correlation spectroscopy. *P. Natl. Acad. Sci. USA* **2006** (103) 11136

[4] M Eisenhawer, S Cattarinussi, A Kuhn, H Vogel. Fluorescence resonance energy transfer shows a close helix-helix distance in the transmembrane M13 procoat protein. *Biochemistry* **2001** (40) 12321

[5] Y Ishii, T Yoshida, T Funatsu, T Wazawa, T Yanagida. Fluorescence resonance energy transfer between single fluorophores attached to a coiled-coil protein in aqueous solution. *Chem. Phys.* **1999** (247) 163

[6] X Michalet, S Weiss, M Jaoger. Single-molecule fluorescence studies of protein folding and conformational dynamics. *Chem. Rev.* **2006** (106), 1785

[7] M W Fried *et al*. PEGinterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N. Engl. J. Med.* **2002** (347) 975

[8] S J Hadziyannis *et al*. Peginterferon-alpha2a and ribavirin combination therapy in chronic hepatitis C: a randomized study of treatment duration and ribavirin dose. *Ann. Intern. Med.* **2004** (346)

[9] A H Talal et al. Pharmacodynamics of PEG-IFN alpha differentiate HIV/HCV coinfected sustained virological responders from nonresponders. *Hepatology* **2006** (43) 943

[10] J M Harris and R B Chess, Effect of PEGylation on pharmaceuticals. *Nat. Rev. Drug. Discov.* **2003** (2) 214

[11] D R Corey and P G Schultz. Generation of a hybrid sequence-specific single-stranded deoxyribonuclease. *Science* **1986** (238) 1401

[12] V W. Cornish, K M Hahn and P G Schultz. Site-specific protein modification using a ketone handle. *J. Am. Chem. Soc.* **1996** (118) 8150

[13] S Schoffelen, M H L Lambermon, M B van Eldijk, J C M van Hest, Site-specific modification of Candida antarctica8b lipase B *via* residue-specific incorporation of a non-canonical amino acid, *Bioconjugate Chem.* **2008** (19) 1127

[14] A J Link, M K S Vink, D A Tirrell, Preparation of the functionalizable methionine surrogate azidohomoalanine *via* copper-catalyzed diazo transfer. *Nat. Protoc.* **2007** (2)1879

[15] A J Link, M K S Vink, D A Tirrell. Synthesis of the functionalizable methione surrogate azidohomoalanine using Boc-homoserine as precursor. *Nat. Protoc.* **2007** (2) 1884

[16] E D Goddard-Borger and R B Stick. An Efficient, inexpensive, and shelf-stable diazotransfer reagent: imidazole-1-sulfonyl azide hydrochloride. *Org. Lett.* **2007** (9) 3797

[17] K Bernd, A Janina, B Wolfgang. Alkyne bridged-amino acids by palladium mediated coupling of alkynes with *N*-t-Boc-4-iodo-phenylalanine methyl ester. *Tetrahedron* **1997** (53) 2475

[18] V W Cornish, D R Benson, C A Altenbach, K Hideg, W L Hubbell, P G Schultz. Site-specific incorporation of biophysical probes into proteins. *P. Natl. Acad. Sci. USA* **1994** (91) 2910

[19] L Wang, A Brock, B Herberich, P G Schultz. Expanding the genetic codon of Escherichia coli. *Science* **2001** (292) 498

[20] J W. Chin, T A Croop, J C Anderson, M Mukherji, Z Zhang, P G Schultz. An expanded eukaryotic genetic code. *Science* **2003** (301) 964

[21] L Wang, P G Schultz, Expanding the genetic code. *Angew. Chem. Int. Ed.* **2005** (44) 34

[22] J M Xie, P G Schultz. A chemical toolkit for proteins—an expanded genetic code. *Nat. Rev.: Mol. Cell Bio.* **2006** (7) 775

[23] K L Kiick, J C M van Hest and DA Tirrell. Expanding the scope of protein biosynthesis by altering the methionyl-tRNA synthetase activity of a bacterial expression host. *Angew. Chem. Int. Ed.* **2000** (39) 2148

[24] J C M van Hest, K L Kiick, D A Tirrell. Efficient incorporation of unsaturated

methionine analogues into proteins in vivo. *J. Am. Chem. Soc.* **2000** (122) 1282

[25] K L Kiick, E Saxon, D A Tirrell, CR Bertozzi, Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *P. Natl. Acad. Sci. USA* **2002** (99)19

[26] C Tanrikulu, E Schmitt, Y Mechulam, W A Goddard III, D A Tirrell. Discovery of Escherichia coli methionyl-tRNA synthetase mutants for efficient labeling of proteins with azidonorleucine in vivo. *P. Natl. Acad. Sci. USA* **2009** (106) 15285

[27] K Kirshenbaum, I S Carrico and D A Tirrell. Biosynthesis of proteins incorporating a versatile set of phenylalanine analogues. *ChemBioChem* **2002** No. 02-03 235

[28] J Wilhelmi, X Chen, *et al*. Unpublished data **2009**

[29] J M Antos and M B Francis. Selective Tryptophan Modification with rhodium carbenoids in aqueous solution. *J. Am. Chem. Soc.* **2004** (126) 10256.

[30] S D Tilley and M B Francis. Tyrosine-selective protein alkylation using π-allylpalladium complexes. *J. Am. Chem. Soc.* **2006** (128) 1080

[31] S Brocchini, S Balan, A Godwin, J W Choi, M Zloh and S Shaunak. PEGylation of native disulfide bonds in proteins. *Nat. Protoc.* **2006** (1) 2241

[32] M Zloh, S Shaunak, S Balan and S Brocchini. Identification and insertion of 3-carbon bridges in protein disulfide bonds: a computational approach. *Nat. Protoc.* **2007** (2) 1070

[33] J M Gilmore, R A Scheck, A P Esser-Kahn, N S Joshi and M B Francis, *N*-Terminal protein modification through a biomimeric transamination reactiono. *Angew. Chem. Int. Ed.* **2006** (45) 5307

[34] S N P Murthy, T J Lukas, T S Jardetzky and L Lorand. Selectivity in the post-translational, transglutaminase-dependent acylation of lysine residues. *Biochemistry* **2009** (48) 2654.

[35] T Nguyen and M B Francis. Practical synthetic route to functionalized rhodamine dyes, *Org. Lett.* **2003** (5) 3245

[36] Q H Meng, M J Yu, H F Zhang, J C Ren and D Y Huang. Synthesis and application of *N*-hydroxysuccinimidyl rhodamine B ester as an amine-reactive fluorescent probe. *Dyes*

Pigments **2007** (73) 254

[37] J Futami, E Nukui, T Maeda, M Kosaka, H Tada, M Seno, H Yamada. Optimum modifcation for the highest cyttoxicity of cationized ribonuclease. *J. Biochem*. **2002** (132) 223

[38] J Futami, T Maeda, M Kitazoe, E Nukui, H Tada, M Seno, M Kosaka, H Yamada. Preparation of potent cytotoxic ribonucleases by cationization: enhanced cellular uptake and decreased interaction with ribonuclease inhibitor by chemical modification of carboxyl groups. *Biochemistry* **2001** (40) 7518

[39] N Koval'ov, I Kuzenetsova, E Burakova, V Sil'nikov, M Zenkova, V Vlassov. International conference on targeting RNA/Workshop on biology and biochemistry of extracellular nucleic acids. JUN 18-21, 2003 Novosibirsk, RUSSIA, *Nucleos. Nucleot. Nucleic Acids* **2004** (23) 977

[40] M Clarisse, P Hartmann, J P Obrecht. Titanium-mediated addition of silyl dienol ethers to electrophilic glycine: 4-ketopipecolic acid hydrochloride. *Org. Synth*. **1993** (71) 200 (**1998**, Coll. Vol9, p.527)

[41] T Sasaki, K Kodama, H Suzuki, S Fukuzawa, K Tachibana, *N*-terminal labeling of proteins by the Pictet-Spengler reaction. *Bioorg. Med. Chem. Lett*. **2008** (18) 4550

[42]M Zwolinska, P Smolewski, Onconase: a ribonuclease with antitumor activity. *Postepy Hig. Med. Dosw*. **2010** (64) 58

[43] A Kbeck, H I Pass, M Carbone, H Yang, Ranpirnase as a potential antitumor ribonuclease treatment for mesothelioma and other malignancies. *Future Onoco*. **2008** (4) 341

[44] L Strong, B Mei, J Kink, M Shaha, R Raines. Anti-cancer activity of human ribonuclease conjugates with enhanced pharmacokinetic profiles. *EJC Supplements* **2008** (6) 110

[45] S L Dong, L Merkel, L Moroder, N Budisa. Convenient syntheses of homopropargylglycine. *J Pept. Sci*. **2008** (14) 1148
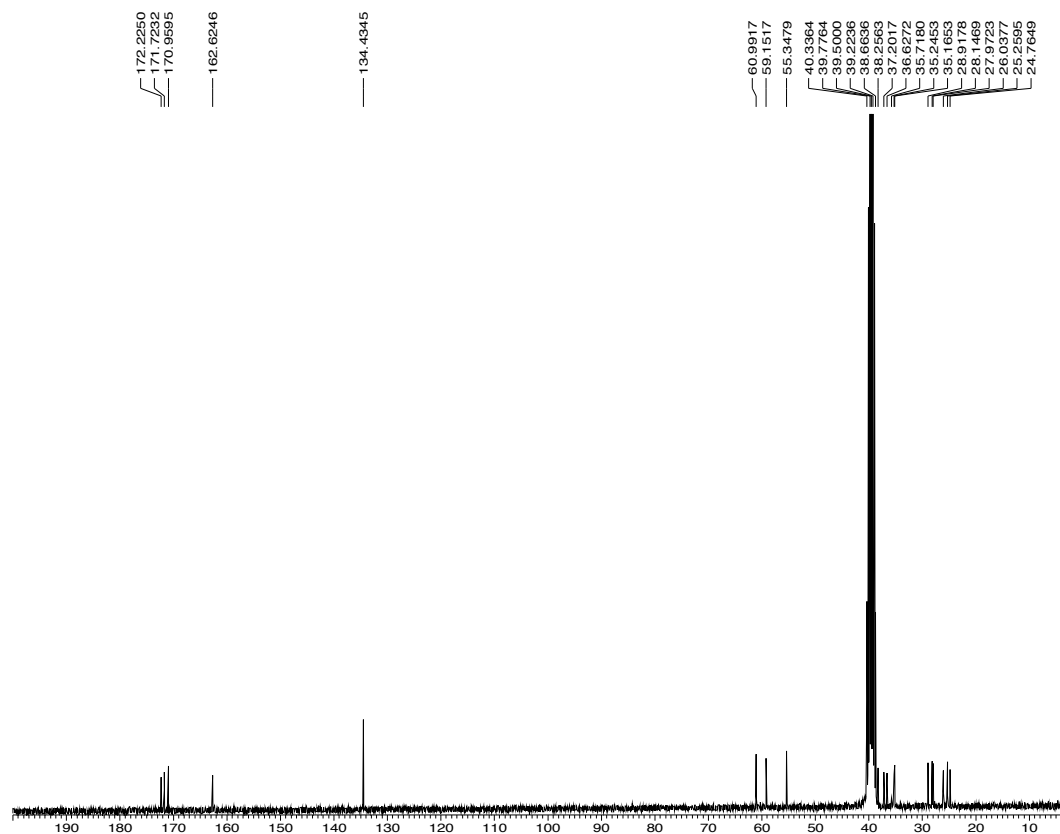
[46] M Otagiri and V T G Chuang. Pharmaceutically Important Pre- and Posttranslational Modifications on Human Serum Albumin. *Biol. Pharm. Bull*. **2009** (32) 527
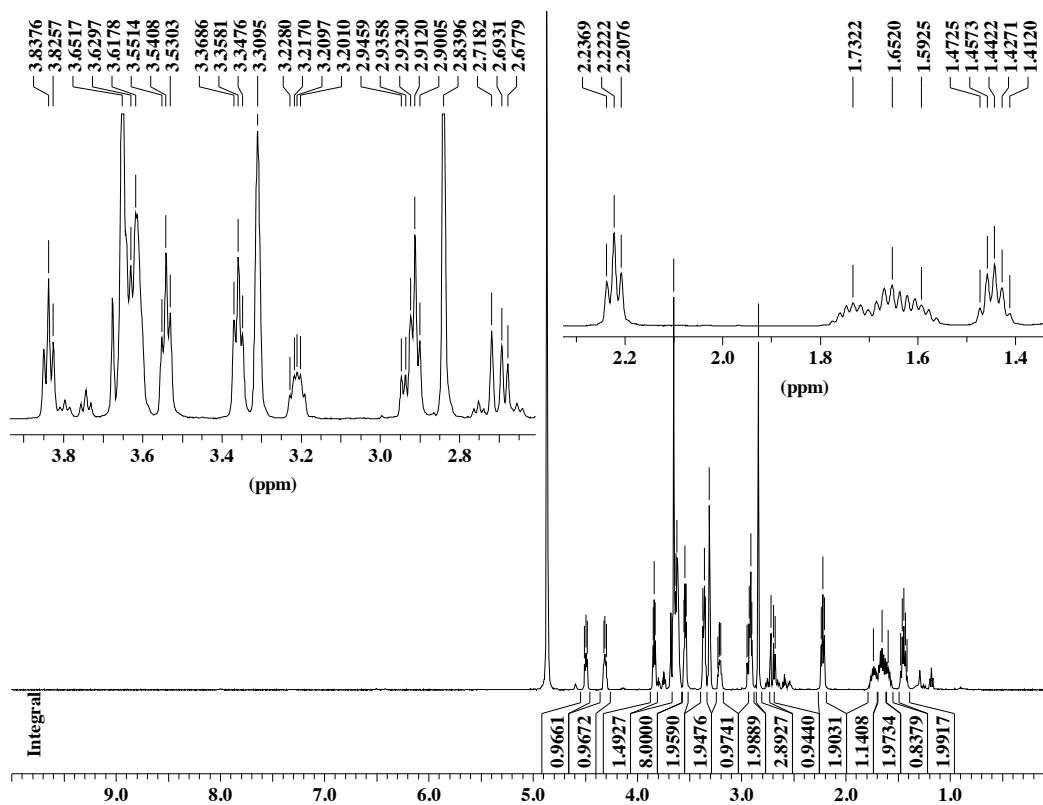
APPENDICES

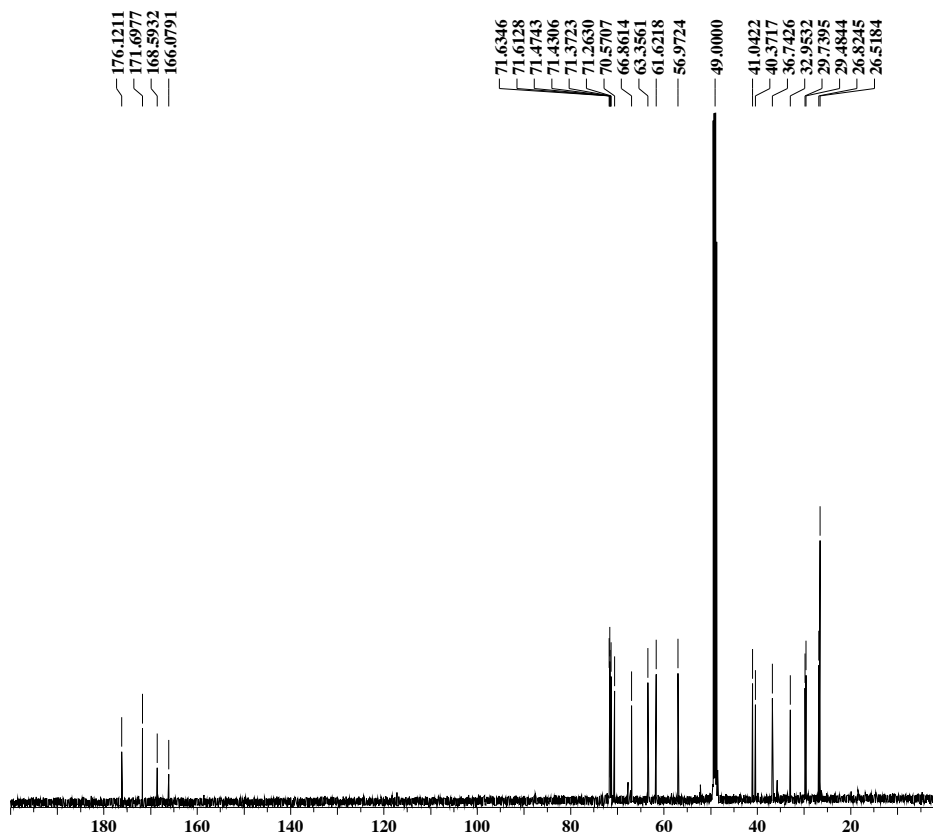1. **^{1}H-NMR** (CDCl$_3$) of Biotin-LC-Maleimide **32**
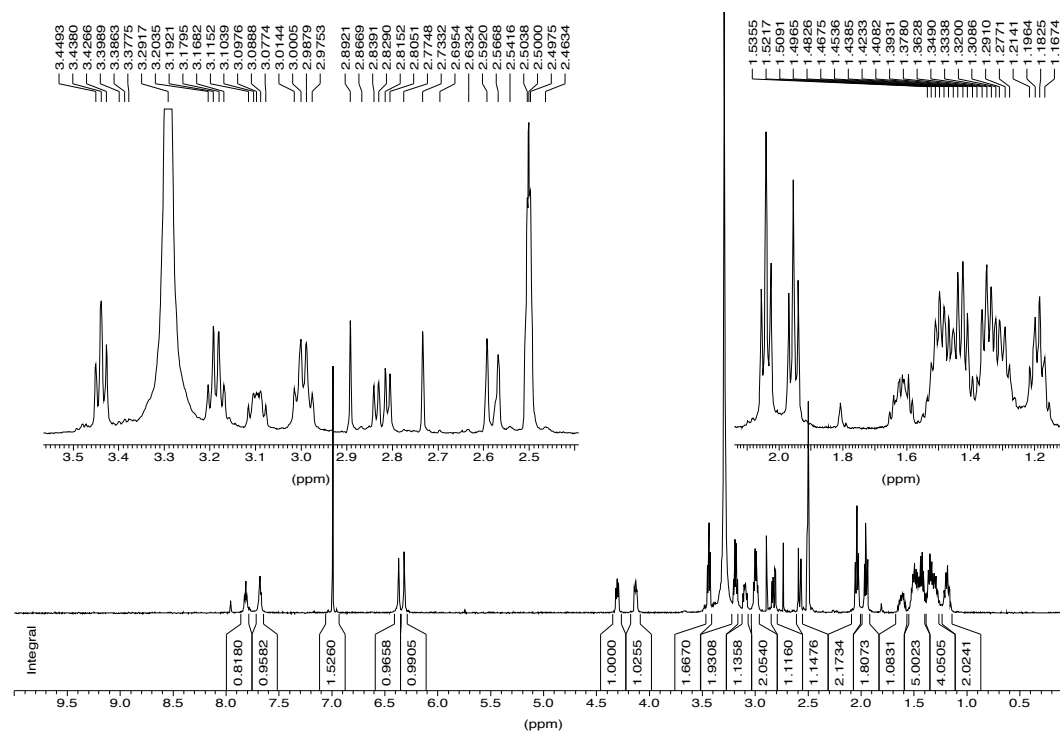
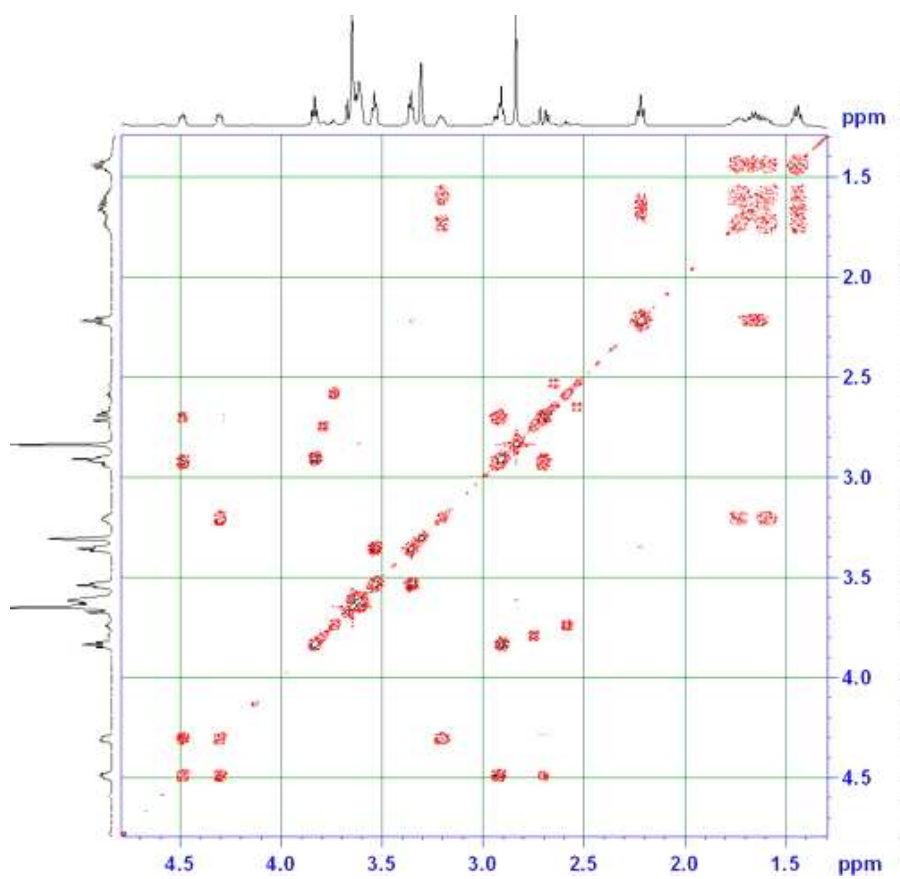2. **^{13}C-NMR** (CDCl$_3$) of Biotin-LC-Maleimide **32**
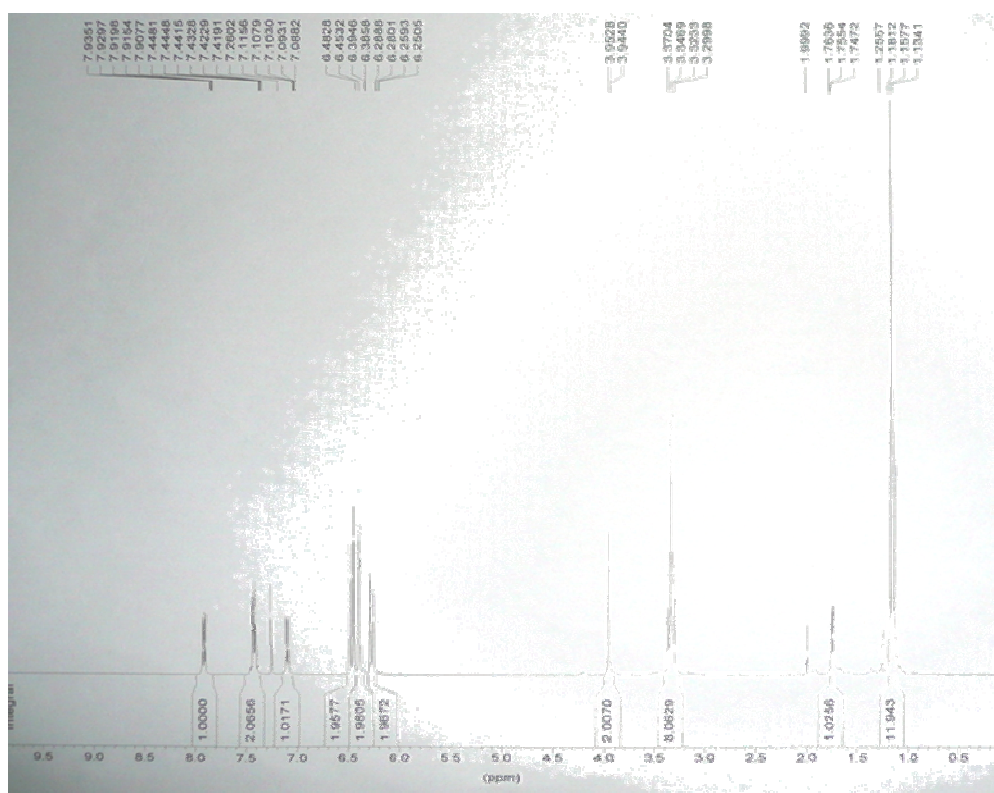
3. **^{1}H-NMR** (MeOD) of Biotin-PEG-NHS **49**

4. **^{13}C-NMR** (MeOD) of Biotin-PEG-NHS **49**

5. **^{1}H-^{1}H COSY** (MeOD) of Biotin-PEG-NHS **49**

6. **^{1}H-NMR** (CDCl$_3$) of Pseudo Rhodamine B Propargyl Amide **57**

7. ^{1}**H-NMR** (CDCl$_3$) of Rhodamine B Ethyl Ester **60**
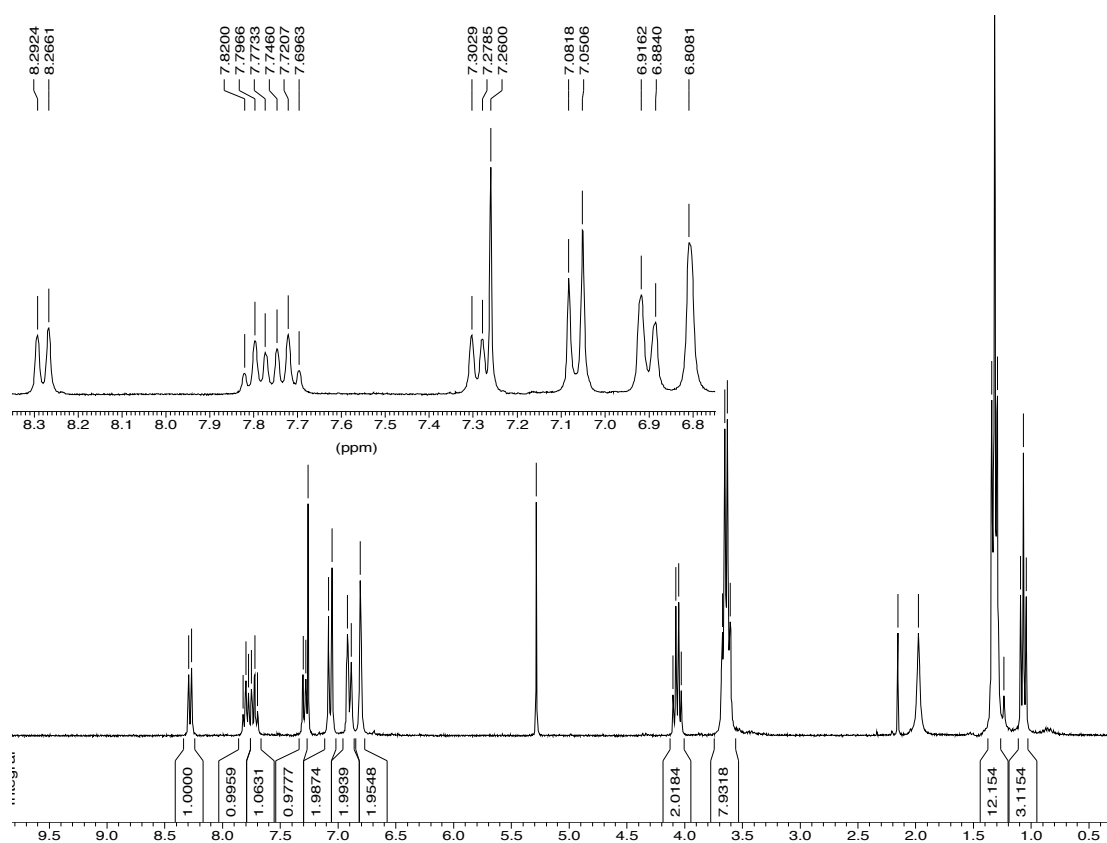
8. ^{13}**C-NMR** (CDCl$_3$) of Rhodamine B Ethyl Ester **60**
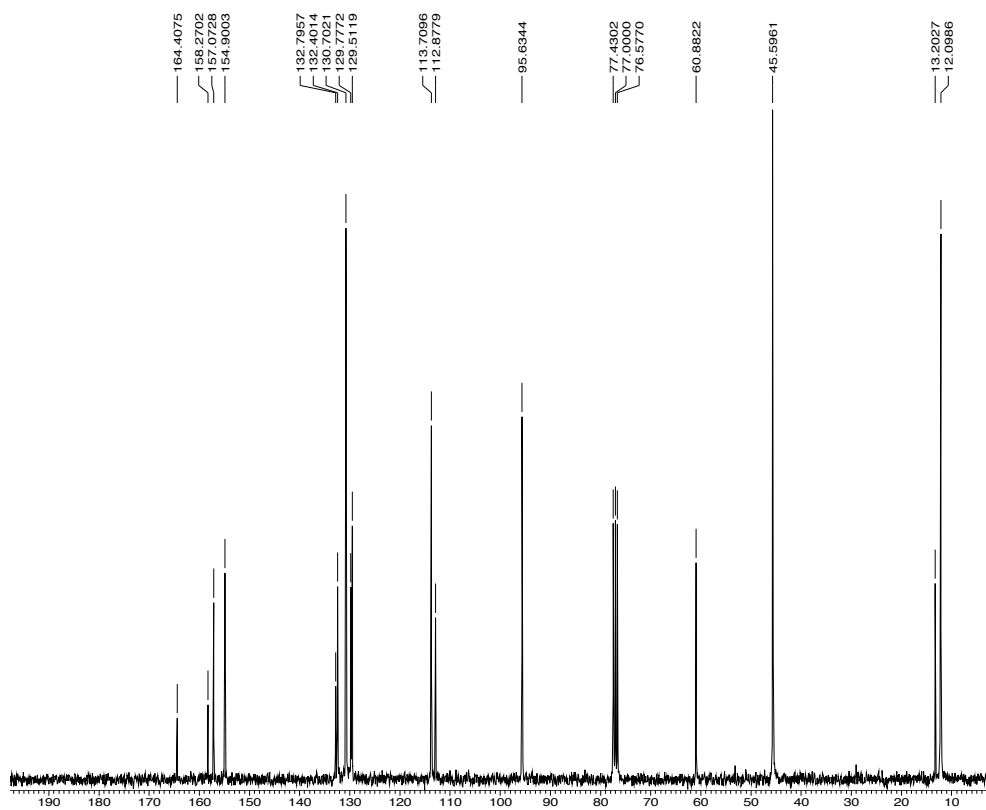
9. **^{1}H-NMR** (CDCl$_3$) of Rhodamine B Propargyl Ester **73**
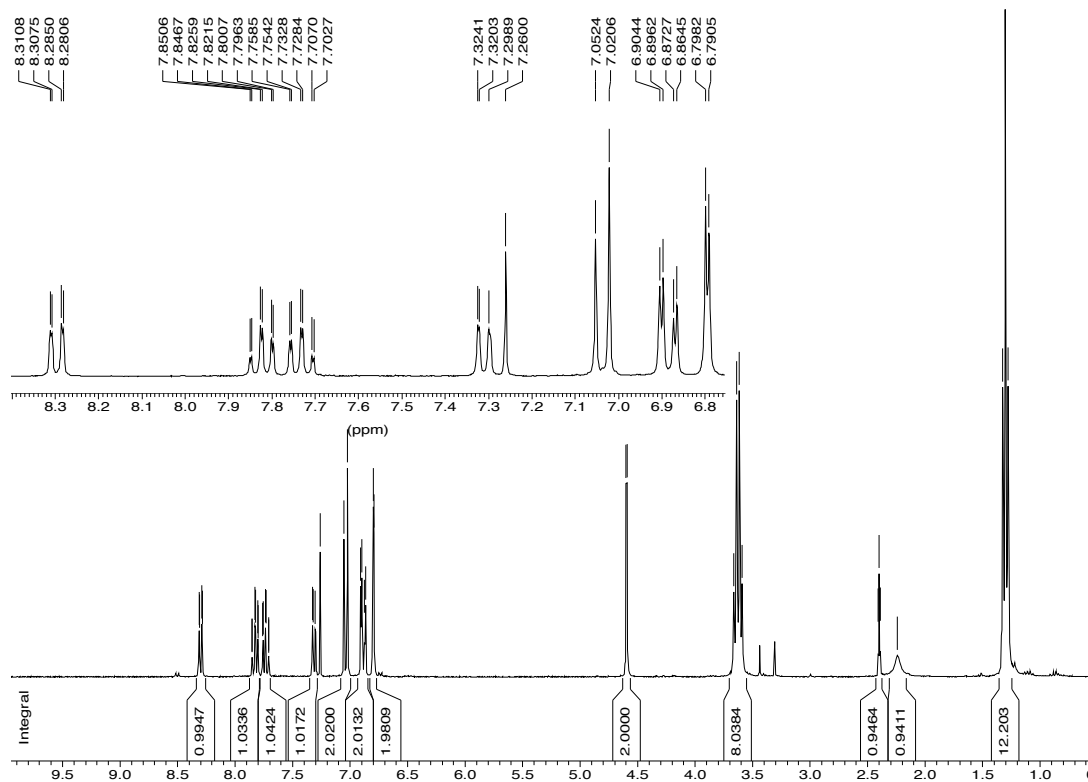
10. **^{13}C-NMR** (CDCl$_3$) of Rhodamine B Propargyl Ester **73**
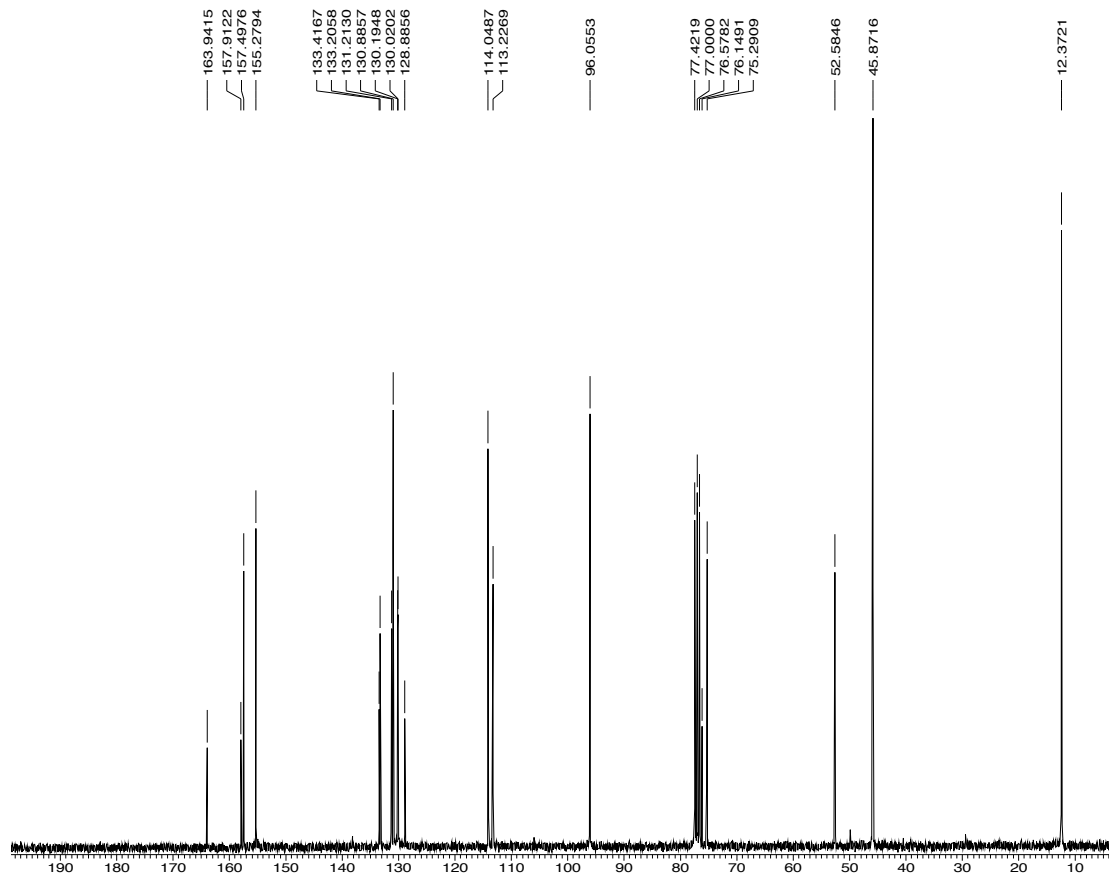
11. **^{1}H-NMR** (CDCl$_3$) of Rhodamine B Azidoethyl Ester **63**
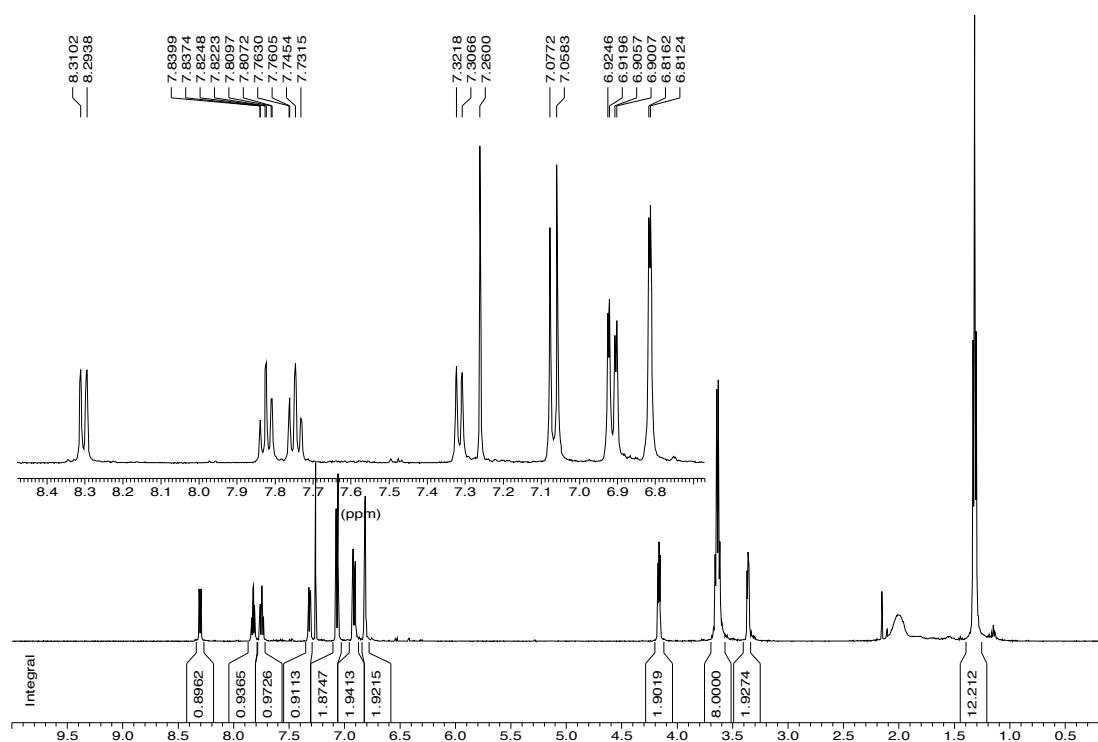
12. **^{13}C-NMR** (CDCl$_3$) of Rhodamine B Azidoethyl Ester **63**
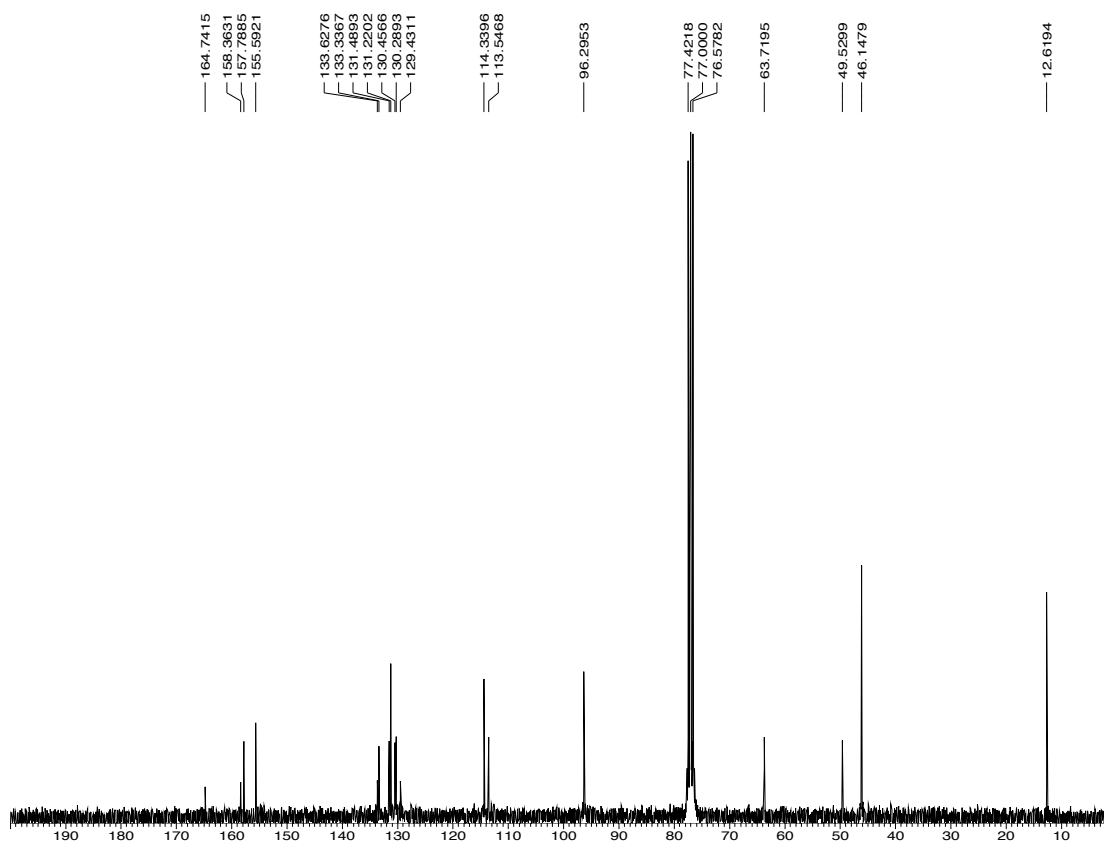
13. **^{1}H-NMR** (CDCl$_3$) of the Click Product **65**
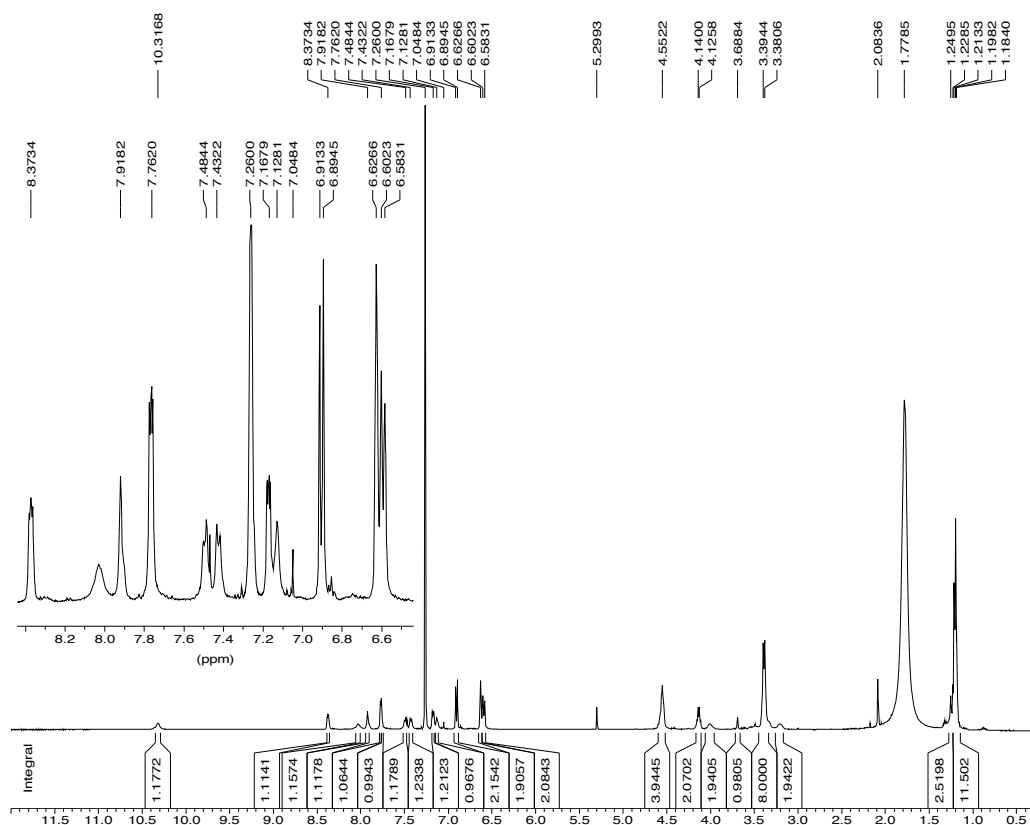
14. **^{1}H-NMR** (CDCl$_3$) of Monosulfone-OH **67**

15. **¹³C-NMR** (CDCl₃) of Monosulfone-OH **67**