
Design and synthesis of activity-based probes (ABPs),
adaptation of activity-based protein profiling (ABPP)
for plant proteomes studies

Zur Erlangung des akademischen Grades eines Doktors der
Naturwissenschaften von Fachbereich Chemie
der Technischen Universität Dortmund
angenommene

DISSERTATION

von

Diplom-Chemiker

Zhe Ming Wang

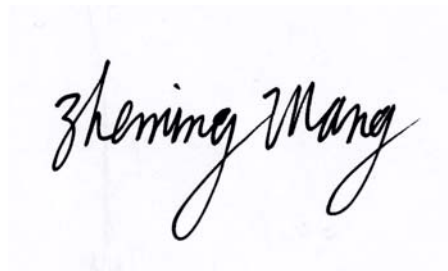
aus Hangzhou

1. Gutachter: Prof. Dr. H. Waldmann

2. Gutachter: Prof. Dr. R. Goody

Tag der mündlichen Prüfung: 28.09.2009

Die
vorliegende Arbeit
wurde unter der Betreuung von
Herrn Prof. Dr. Herbert Waldmann
in der Zeit von Dezember 2005 bis September 2009
am Fachbereich Chemie der Technischen Universität Dortmund,
sowie am Chemical Genomics Centre (CGC) der Max-Planck-Gesellschaft
in Dortmund angefertigt.

A handwritten signature in black ink that reads "Zheming Wang". The signature is written in a cursive style with a large, sweeping 'Z' and 'M'.

29.07.2009

釵頭鳳新曲

敬呈愛吾及吾愛之人

武陵人，躍龍門。

雪殘春曉鮮暇聞。

母呈食，父添茶。

影三相隨，夜狼孤嚎。

熬，熬，熬。

異鄉城，語新文。

客學西域結親朋。

雨千秋，聚川河。

舊時回首，醉貓獨笑。

妙，妙，妙。

Acknowledgements

Maybe this is the only opportunity to express my sincere gratitude to everyone at my working places (CGC, MPI-Dortmund and MPIZ) who offered any help or gave advices in the past.

Patrick, working at the fume hood beside me, offered so many help, big or small. I thank you for always standing beside me, in roomage and emotion.

Farnusch, Takayuki and Christian, my plant chemetics group colleagues in Cologne, did me great favours carrying out biological work. Your help is highly appreciated.

Prof. Waldmann, you gave me the great opportunity to explore the subtlety of chemical biology and arranged a nice position for me with two excellent group leaders taking good care of me. I thank you.

Markus, you spend me the lab, spend me the time, spend me the talks and spend me the ideas. How lucky to have met you. We shared happiness; we argued for the truth; we discussed for the problems and you cared for my private life. All things are engraved deeply on my memory. I say, thank you.

Renier, you bring me to your newly born group, bring me to so many beautiful places, bring me to the fantastic biological world, bring me to the nature. You gave me skills, your kind guiding and your useful advises triggered me to be mature. They are always on my mind. I say, thank you.

The last appreciation is well deserved for my most important home front, my papa, mama and Sumi as well as all my close friends from my hometown, from Bavaria and from other places. The time shared with you is wonderful.

Content

1	General Introduction	1
1.1	From genomics to functional genomics	1
1.2	Short historical background to proteins	3
1.3	From functional genomics to proteomics	5
1.4	From proteomics to functional proteomics	7
1.5	From functional proteomics to chemical proteomics	10
1.6	Chemical biology in chemical proteomics	13
1.7	Protease for life science	16
1.8	Background to activity-based protein profiling (ABPP)	19
1.9	ABPP for plant biology	23
2	Aim of the PhD mission	26
3	Results and discussions	29
3.1	Natural product-based ABP design of <i>anti</i> - β -lactones	29
3.1.1	Introduction	29
3.1.2	Syntheses	31
3.1.3	Bioassays	37
3.1.4	Discussion	43
3.2	Mechanism-based ABP design of <i>syn</i> - β -lactones	44
3.2.1	Introduction	44
3.2.2	Syntheses	45
3.2.3	Bioassays	47
3.2.4	Discussion	56

3.3	Non-directed ABP design of aziridines and azirines	58
3.3.1	Introduction	58
3.3.2	Syntheses	59
3.3.3	Bioassays	64
3.3.4	Discussion	66
3.4	Direct ABPs design of AEBSF-based probes	67
3.4.1	Introduction	67
3.4.2	Syntheses	68
3.4.3	Bioassays	69
3.4.4	Discussion	77
3.5	Mechanism-based ABP design of AOMK for AvrPphB	79
3.5.1	Introduction	79
3.5.2	Syntheses	80
3.5.3	Bioassays	84
3.5.4	Discussion	90
3.6	Mechanism-based ABP design of AOMK for VPE	92
3.6.1	Introduction	92
3.6.2	Syntheses	92
3.6.3	Bioassays	94
3.6.4	Discussion	95
3.7	Natural product syntheses of Gibbestatin	97
3.7.1	Introduction	97
3.7.2	Previous synthesis approaches for GNB	99
3.7.3	First synthesis approach for GNB	100
3.7.4	Second synthesis approach for GNB	106
3.7.5	Discussion	109

4	Summary (English)	111
4.1	Chemical part	111
4.2	Chemical biological part	113
4.3	Biological part	116
5	Zusammenfassung (German)	117
5.1	Chemischer Teil	117
5.2	Chemisch-biologischer Teil	119
5.3	Biologischer Teil	123
6	Experimental section	124
6.1	Chemical part	124
6.1.1	Instruments and Reagents	124
6.1.2	Synthetic procedures	128
6.2	Biological part	225
6.2.1	Materials of bioassays	225
6.2.2	Methods of bioassays	228
6.2.3	Procedures of ABPPs	232
7	References	242
8	NMR spectra	257

Abbreviations

1D: One-dimensional

2D: Two-dimensional

ABPs: Activity-based probes

ABPP: Activity-based protein profiling

ACN: Acetonitrile

AD: Activation domain

AEBSF: 4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride

AFP: Affinity-based profiling

AIBN: Azobis-*iso*-butyronitrile

AOMK: Acyloxymethyl ketone

BuLi: Butyl lithium

br: Broad

CDCl₃: Deuteriochloroform

cICAT: Cleavable isotope-code affinity tag

δ: Chemical shift

DCC: *N,N'*-Dicyclohexylcarbodiimide

DCM: Dichloromethane

DBD: DNA binding domain

DHP: 2,3-Dihydropyran

DIAD: Di-*iso*-propyl azodicarboxylate

DIC: *N,N'*-Di-*iso*-propylcarbodiimide

DIPEA: Di-*iso*-propylethylamine

DIPT: Di-*iso*-propyl tartarate

DMAP: 4-(Dimethylamino)pyridine

DMF: *N,N*-Dimethylformamide

DMP: Dess-Martin periodinane

DMSO: Dimethyl sulfoxide

DOS: Diversity oriented synthesis

DTT: Dithiothreitol

ELISA: Enzyme-linked immuno-sorbent assay

ESI: Electrospray ionization

EST: Expressed sequence tag

ETS: Effector-triggered susceptibility

ETI: Effector-triggered immunity

FCG: Forward chemical genetics

FP: Fluorophosphate

h: Hour

HBTU: *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate

HOBt: *N*-Hydroxybenzotriazole

HPLC: High performance liquid chromatography

HR: Hypersensitive cell death response

HRMS: High resolution mass spectroscopy

HTS: High-throughput screening

IBX: 2-Iodoxybenzoic acid

IMAC: Immobilized metal affinity chromatography

IR: Infrared-spectroscopy

J: Coupling constant in Hertz

LDA: Lithium diisopropylamide

LiHMDS: Lithium bis(trimethylsilyl)amide

m: Multiplet

MALDI: Matrix-assisted laser desorption ionization

MAMPs: Microbial-associated molecular patterns

MHz: Megahertz

min: Minute

MS: Mass spectrometry

MRM: Multiple reaction monitoring

NaOAc: Sodium acetate

NB-LRR: Nucleotide binding leucine rich repeat domain

NBS: *N*-Bromosuccinimide

ND: Natural product derivatives

NMM: *N*-Methylmorpholine

NMR: Nuclear magnetic resonance

NP: Natural products

ORF: Open reading frame

PAMPs: Pathogen-associated molecular patterns

PCD: Program cell death

PGs: Polygalacturonases

PMSF: Phenylmethanesulfonyl fluoride

PNBA: *Para*-nitrobenzoic acid

PPIs: Protein-protein interactions

PPTS: Pyridinium *para*-toluenesulfonate

PRRs: Pattern recognition receptors

PTI: PAMP-triggered immunity

PTM: Post-translational modification

PyBOP: Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate

q: Quartet

QIT: Quadrupole/linear ion trap

R protein: Disease resistance proteins

ROIs: Reactive oxygen intermediates

ROS: Reactive oxygen species
RGC: Reverse chemical genetics
s: Singlet
SA: Salicylic acid
SAGE: Serial analysis of gene expression
SDS-PAGE: Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SIL: Stable isotope labelling
SILAC: Stable isotope labelling of amino acids in cell culture
SNP: Single-nucleotide polymorphism
SRM: Selective reaction monitoring
SPPS: Solid phase peptide synthesis
t: Triplet
TBAF: Tetrabutylammonium fluoride
TBDPSCl: *tert*-Butyl diphenylchlorosilane
TBS: Tris-buffered saline
TEA: Triethylamine
TFA: Trifluoroacetic acid
TfOMe: Methyl trifluoromethanesulfonate
TGLA: Target guided ligand assembly
THF: Tetrahydrofuran
TLCK: *N*^α-Tosyl-lysyl-chloromethylketone
TMSCl: Chlorotrimethylsilane
TOF: Time-of-flight
TPCK: *N*-2-*p*-tosyl-l-phenylalanine-chloromethylketone
TPP: Triphenylphosphine
Tris buffer: *Tris*(hydroxymethyl)aminomethane buffer
t_R: Retention time

TTSS: Type three secretion system

VICAT: Visible isotope-coded affinity tag

VIGS: Virus-induced gene silencing

VS: Vinyl sulfone

Y2H: Yeast-two-hybrid system

Y3H: Yeast-three-hybrid system

1 General introduction

1.1 From genomics to functional genomics

Genomics is the large-scale study of the genomes of organisms by biochemical tools, which was established by Nobel laureate Fred Sanger with the complete sequencing of genomes of 5S ribosomal RNA from *Escherichia coli* in 1967 (Brownlee *et al.*, 1967) and bacteriophage phi X174 DNA in 1977 (Sanger *et al.*, 1977). However, the term “genomics” itself was coined by Tomas Roderick in 1986. Genomics studies based on DNA sequencing were the central issue in the Pre-Genomic era, many significant milestones were reached during that time, such as the first sequencing of a gene of bacteriophage MS2 coat protein in 1972 (Min Jou *et al.*, 1972), the first sequencing of a DNA-based genome of bacteriophage phi X174 in 1977 (Sanger *et al.*, 1977), the first sequencing of a free-living organism of *Haemophilus influenzae* in 1995 (Reischmann *et al.*, 1995) and a rough draft of the human genome by the Human Genome Project in early 2001 (International Human Genome Sequencing Consortium *et al.*, 2001; Venter *et al.*, 2001).

This draft of the human genome as a symbol marked the coming of the Post-Genomic era. However, at that time the intensive accumulation of complete gene codes in databases was still not sufficient to clarify their gene products. Several reasons result such difficulty: First, the existence of an open reading frame (ORF) in genomic data does not necessarily indicate the existence of the corresponding gene product, as was found in the case of alpha herpesviruses (Boldogkoei *et al.*, 1994); second, there is no strict linear relationship between genes and their corresponding gene products. This has been pointed out in by Pandey *et al.* in 2000 (Pandey *et al.*, 2000) and was indirectly supported with the study on the relationship between protein-fusion and gene co-expression by Gunter and

Gaasterland (Gunter *et al.*, 2001); third, until now and even in the near future, it is still difficult to predict genes accurately from genomic data, as pointed out from the sequencing of human chromosome 22 (Dunham *et al.*, 1999); fourth, as mentioned by Pandey *et al.*, “the sequencing of related organisms will ease the problem of gene prediction through comparative genomics, the success rate for correct prediction of the primary structure is still low” (Claverie 1997, Pandey *et al.*, 1999). The error rate was at least 8% in the annotations for 340 genes from the *Mycoplasma genitalium* genome (Brenner 1999) and even in the recent reports, the problem still exists in spite of the reduced error rate (Shampson *et al.*, 2009).

Furthermore, despite the difficult of identifying the gene products directly from their genome, the simple demonstrating the existence of the gene products is also not sufficient for understanding their function in a biological system. In order to understand the function of the genes, the genomes analyses have been improved. “Functional genomics” was initiated in the late eighties, which was conceived to be able to narrow the gap between traditional genome sequencing-based analysis and the required genome function analysis. The birth of functional genomics also marked the division of conventional genomics into two subtopics. Beside “functional genomics”, “structural genomics” was also invented (Burley *et al.*, 1999). During the early nineties, there were many different interpretations on functional genomics. In 1997, Philip Hieter and Mark Boguski pointed at a clearer task for functional genomics and enumerated several emerging methods (Hieter *et al.*, 1997). Functional genomics nowadays is interpreted as a field, which utilizes the genomic data to describe gene functions and interactions. Especially studies on gene transcription, translation, and protein-protein interactions are in the focus of functional genomics. To this end, many associated techniques were developed in a rapid way: DNA microarrays (Schena *et al.*, 1995, Shalon *et al.*, 1996) and serial analysis of gene expression (SAGE) (Velculescu *et al.*, 1995, Saha *et al.*, 2002,

Matsumura *et al.*, 2005) for mRNA, siRNA and cDNA transfection to study individual or libraries of genomic reagents in cell-based experiments (Mousses *et al.*, 2003, Pullmann *et al.*, 2007). A concept termed as “expressed sequence tag” (EST) (Adams *et al.*, 1991) emerged during that time. As an important resource, ESTs have been intensively studied, benefited from above advanced methods (Nagaraj *et al.*, 2006). Because most studies on ESTs are at the RNA level, the term “transcriptomics” was proposed. In analogy to genomics, transcriptomics is the study of the transcriptomes. A recent review summarized the achievements of transcriptomics studies until now (Wang *et al.*, 2009). The further that researchers go in addressing questions on functional genomics, the more they realize that studies only on gene function-related mutations and polymorphism, such as single-nucleotide polymorphism (SNP) (Vignal *et al.*, 2002) is insufficient, and that studies should also involve the measurement of final gene products (such as proteins and metabolites), which are opposite to transcripts, the prior gene product. Following the birth of transcriptomics, proteomics, the study of the proteomes and metabolomics, the study of the metabolomes (Oliver *et al.*, 1998) emerged from functional genomics.

1.2 Short historical background to proteins

Proteins as distinct biological polymers have a long history in chemistry. In the eighteenth century, they were recognized as chemical compounds coagulating or flocculating under treatments with heat or acid by the French chemist and clinician Antoine Fourcroy (1755-1809) and others. The first elemental analysis of common proteins was performed by the Dutch chemist Gerhardus Johannes Mulder (1802-1880), to reveal the astonishing fact that nearly all proteins have a same empirical formula $C_{400}H_{620}N_{100}O_{120}P_1S_1$ (Mulder 1838). Mulder believed the proteins consists of a fundamental material, named as *Grundstoff* and hydrolysis of

proteins delivered a substance, with both amino and carboxylic functionalities. Hence, the substance was called an amino acid. The first two amino acids leucine (in 1819) and glycine (in 1820) were isolated by the French chemist and pharmacist Henri Braconnet (1780–1855). Mulder continued to identify the molecular mass of leucine and chemically synthesized it as the first synthetic amino acid in 1855 (Perrett 2007). The term “protein”, derived from the greek word “prota” (of primary importance), however, was coined by Mulder’s associate in 1838, the Swedish chemist and clinician Joens Jakob Berzelius (1779-1848). Although later several other proteins such as albumins, fibrin and haemoglobin were isolated, the function of proteins in biological systems was not understood until the enzyme urease was demonstrated to be a protein (Sumner 1926). Before the phenylalanyl chain of insulin I/II in 1951 and the glycyll chain of insulin I/II in 1953 were sequenced by Frederick Sanger (Sanger *et al.*, 1951, Sanger *et al.*, 1953), proteins were presumed to be somewhat amorphous. The structures of myoglobin and haemoglobin were resolved by Sir John Cowdery Kendrew and Max Perutz, respectively (Kendrew *et al.*, 1958, Perutz *et al.*, 1960). Those were the first 3D structures determined by x-ray diffraction analysis.

While the properties and structures of proteins such as catalytic activities of enzymes were studied by chemists, the biological functions of proteins in living organisms were also increasingly studied by biologists. Initially, proteins were studied individually. However, when the time of studying protein cellular functions came - studies into actin or tubulin (Weisenberg 1981), enzyme cascades such as the coagulation cascade (Zimmermann *et al.*, 1978) or cell signalling such as Hedgehog signalling pathway (Nüsslein-Volhard *et al.*, 1981) - a systematic study of proteins was appreciated in biology.

1.3 From functional genomics to proteomics

Functional genomics made many achievements with mRNA studies. However when N. Leigh Anderson and Jeff Seilhamer presented the first multi-gene comparison plot of mRNA *vs* protein abundance for cellular gene products in 1997, they calculated a correlation coefficient between mRNA and protein of 0.48 (Anderson *et al.*, 1997). This correlation coefficient shows that the relationship between mRNA and proteins is not that strict. Studies of angiotensin (Pfeffer, 1993), 7-transmembrane G-protein coupled receptors (Wess, 1997) and Alzheimer amyloid precursor protein (Hooper *et al.*, 1997) showed that cellular control system can operate purely at the protein level. Transcriptomics cannot resolve the questions of cellular control systems. Furthermore, as proteins are more stable, they are much better samples than mRNA. Proteins are involved in nearly all biological activities, so they are the direct material to study functions of biological systems. “Functions: proteins 100 000 – mRNA 1” and “the current term functional genomics, which implies that function can be explored at the genomic level, is a bizarre...” were described in the review by Anderson (Anderson *et al.*, 1998). No matter whether that comment extreme is or not, the proteomics indeed completely succeeded the principle and mission of functional genomics in studying functions of biological systems.

Although the term “protein” originated in the field of chemistry as early as the nineteenth century, indeed much earlier than the term “gene”, “proteome” and “proteomics” were proposed later than “genome” and “genomics” in the field of molecular biology. The term proteome was coined by Marc Wilkins in 1994 in the symposium of “2D Electrophoresis: from protein maps to genomes” and later in his paper “from proteins to proteomes” to describe the proteome as “the PROTEin complement expressed by a genOME” (Wilkins *et al.*, 1996). Nowadays, a

proteome usually refers to the entire of proteins expressed by a genome, cell, tissue or organism. Hence, proteome is not as genome strictly defined. To describe a proteome in an accurate way, the time and defined conditions to express the referred proteome have to be noted. In reality, the term proteome sometimes refers more widely to a group of cellular proteins, or a collection of proteins in certain sub-cellular biological systems. (In certain cases, a mixture of proteins from different organisms is also called a mixed proteome.) Due to alternative splicing of genes and post-translational modifications (PTMs) of proteins, there are more proteins than genes. To study a proteome is therefore more complex than a genome. The proteomics cannot be limited to the accumulation of the protein sequences like the genomics is a list of DNA sequences, but also including information on protein structures and the functional interaction between proteins.

Although proteomics was first coined by James in 1997 (James, 1997), the origin of proteomics dates back to the late seventies when different high-resolution 2D-electrophoresis methods were developed by Klose (Klose, 1975), O'Farrell (O'Farrell, 1975) and Scheele (Scheele, 1975) to assemble proteins databases. Hence, proteomics stood originally for analysis of a large number of proteins from a given cell line or organism on 2D polyacrylamide gels. With further developed techniques, proteomics is nowadays more widely defined as large-scale studies of proteins in living organisms, particularly their structures and physiological functions.

Proteomics is complementary to genomics because it directly focuses on the proteins, which implements all biological processes. Verification of gene products by proteomics is a reliable way to interpret the genome. As a short summary, comparing to genomics or transcriptomics, the advantages of proteomics for studying the functions in biological systems are obvious: first, the existence of gene products can be confirmed by proteomics; second, the protein expression level can

be more accurately determined (Gygi *et al.*, 1999); third, the subcellular localization of gene products can be monitored experimentally; fourth, the interaction between protein and the cellular chemical structures such can be analysed only in the protein level; fifth, modifications of the proteins are not encoded directly in the genome, such as isoforms, PTM, and regulation of protein function by proteolysis, recycling and sequestration in cell compartments can be investigated only by proteomics.

1.4 From proteomics to functional proteomics

In the genomic era, proteomics is mainly dependent on annotation from 2D protein gel of cell lysates (Berndt *et al.*, 1999, Caron *et al.*, 2002) and visualization of protein expression levels on 2D protein gels (Greenspan *et al.*, 1995, Gygi *et al.*, 2000). The 2D gel approach dates back to the time when Frederick Sanger analysed the two polypeptide chains of Insulin (Sanger *et al.*, 1951, Sanger *et al.*, 1953). The resolution of the gel was always the bottleneck restricting the gel-based visualization of the proteome. The best 2D gel can resolve no more than 1000 proteins, usually only for the most abundant proteins. However when the technique of mass spectrometric identification of gel-separated proteins was combined with proteomics, protein annotation on gel was widely extended. Comparing to classical Edman degradation technique, mass spectrometry is more sensitive for complex proteomes and able to be combined with high-throughput screening (HTS) techniques (Marvin *et al.*, 2003).

Proteins can be identified by gel-based mass spectrometric methods, however the information of the corresponding protein such as 3D structure, skeleton dynamicity, surface conditions (e. i. charges, lipo- or hydrophilic properties), interaction partners, or cellular localization are all lost with this method. These lost information indeed are more important than the sequence information of a protein

for studying its biological functions. In the Post-genomic era, many other techniques besides mass spectrometry joined in the proteomics for the complete study of functions of the biological systems.

Mass spectrometric instrumentation for proteomics was quickly developed. Two soft ionization techniques, electrospray ionization (ESI) and matrix-assisted laser ionization (MALDI), enabled to ionize peptides or proteins (Fenn, *et al.*, 1989, Karas *et al.*, 1988). In order to analyse the complex proteome, four mass analyzers are routinely applied: quadrupole (Q), quadrupole/linear ion trap (QIT/LIT or LTQ), time-of-flight (TOF) and Fourier-transform ion cyclotron resonance (FTICR) mass analyzers. Nowadays, expensive hybrid analyzers are becoming popular for high sensitivity, such as Q-q-Q, Q-q-LIT, Q-TOF, TOF-TOF and LTQ-FTICR (Han *et al.*, 2008). LTQ-Orbitrap is a new analyzer type, invented in 1999 (Makarov 2000), which was applied for proteome analysis and led to a fourfold improvement of sensitivity (Hu *et al.*, 2005, Venable *et al.*, 2007). The fragmentation techniques for tandem mass spectrometry (MS/MS) were also developed besides conventional collision-induced dissociation (CID), such as electron-capture dissociation (ECD) (Zubarev *et al.*, 1998) or electron-transfer dissociation (Syka *et al.*, 2004). With these new fragmentation methods, PTM of proteins can be better analysed (Shi *et al.*, 2001, Chi *et al.*, 2007). Benefited from the advanced technologies, the strategies for proteome annotation and PTM characterization were also improved. Bottom-up and Top-down are two usual strategies to analyse proteins. The bottom-up approach refers to proteome separation before digestion followed by peptide mass fingerprinting-based acquisition or further peptide separation on-line coupled to tandem mass spectrometry to identify each protein. On the contrary, the top-down approach means that the proteome is fractionated and separated into pure single protein or less complex protein mixtures, followed by off-line static infusion of samples into the mass spectrometer for intact protein mass measurement and

intact protein fragmentation to identify each protein. The shotgun method is a complementary method for the bottom-up approach (McDonald *et al.*, 2003). However incorrect identification can lead to the loss of information from intact proteins and the limited dynamic range of mass spectrometric analysis only allows measurements of the most abundant peptides. The newer top-down approach provides direct analysis of modifications, but limits the ability to fragment intact proteins and the combination with throughout screening (Han *et al.*, 2006, Parks *et al.*, 2007).

Protein-protein interactions (PPIs) study is another major field in functional proteomics. Protein microarrays are the first major application of PPIs. Similar to DNA microarrays, protein microarrays are applied to identify the substrates of proteins such as kinases (Schutkowski *et al.*, 2004), transcription factors for gene-regulation (Sala *et al.*, 2009), or the targets of biologically active small molecules (Uttamchandani *et al.*, 2006). Antibody microarray are a well-established protein microarray, and it is well known as the enzyme-linked immuno-sorbent assay (ELISA), used as a diagnostic tool in medicine and co-purification tool in plant pathology. Purification of protein complexes is a second part of PPIs. To purify entire multi-protein complexes by affinity-based methods, fusion proteins, antibodies, peptides, oligonucleotides or small molecules, targeting to a cellular target can be chosen to immunoprecipitate the complex by an antibody or immobilize the complex on a solid support *via* an associated ‘bait’ (Strausberg *et al.*, 1999, Neubauer *et al.*, 1998). Yeast-two-hybrid system (Y2H) and phage display are two other methods to study PPIs. Y2H was originally devised by Stanley Fields to monitor PPIs in *Saccharomyces cerevisiae* system (Fields *et al.*, 1989). A bait protein is fused to the DNA binding domain (DBD) and a prey protein is fused to the activation domain (AD) of a transcriptional activator. The interaction between bait and prey reconstitutes the functional transcription factor

and results in the expression of the reporter gene for positive selection. The bait protein can be screened against genomic or cDNA prey libraries expressing all encoded or expressed proteins in the organism or tissue of interest. Nowadays, improved Y2H systems have been reported, such as YTH for bacterial PPIs, MYTH for membrane associated PPIs and MAPPIT for mammalian PPIs (Suter *et al.*, 2008). Like the Y2H, phage display is used for the HTS of PPIs, protein-peptide or protein-DNA interactions. It uses bacteriophages to express proteins on their coat protein (Smith 1985). Highly diverse libraries ($> 10^{10}$) can be represented as phage pools, and due to the link between genotype and phenotype, an antigen-specific clone can be selected *in vitro*. Commonly used bacteriophages are M13 and filamentous phage T7 (Kehoe *et al.*, 2005, Sidhu *et al.*, 2007).

There are also several other methods in functional proteomics, such as fluorescence protein fusion to study protein localization (Van Roessel *et al.*, 2002), proteomics analysis on large-scale mouse knockouts (Zambrowicz *et al.*, 1998), RNA interference, as exemplified by small virus-induced gene silencing (VIGS) (Colbère-Garapin *et al.*, 2005) or phenotypic analysis using deletion strains (Winzeler *et al.*, 1999).

1.5 From functional proteomics to chemical proteomics

Although proteomics like genomics belongs to the study field of molecular biology, it returned to the chemistry study field recently, due to the inherent relationship. As shown in the above section, protein studies started in chemistry and chemical modification of proteins has a long-standing history in this field (Lundbald, 2004). The chemical tags used in proteomics today were established during the sixties to seventies. Chemical proteomics sometimes is considered as “methods to measure the interaction between small molecule compounds and protein targets” (Kruse *et al.*, 2008). This view about chemical proteomics is

obviously too narrow; chemical proteomics is basically not different from conventional functional proteomics, but combines chemical tagging concepts (Leitner *et al.*, 2006) systematically with proteomics, and enables this new hybrid methodology to much better resolve the difficulties in functional proteomics.

Modifying functional groups in peptides to improve the sensitivity for mass spectrometry is one part of chemical proteomics. For example, the “tandem mass tag” (Thompson *et al.*, 2003) allows quantitation at the MS/MS level. Coumarin-based tags were used to enhance detection sensitivity and fragmentation efficiency in MALDI-MS and -MS/MS (Pashkova *et al.*, 2004).

Introduction of labels for relative or absolute quantification is the second part of chemical proteomics. In stable isotope labelling (SIL), heavy isotopes such as ^2H , ^{13}C , ^{15}N and ^{18}O are frequently used to determine protein levels. iTRAQ, an acronym for isobaric tag for relative and absolute quantification was introduced in 2004 (Ross *et al.*, 2004) and reductive dimethylation of amino groups (Hsu *et al.*, 2003) is a second method to label the proteome with heavy isotopes. These labelling methods enable mass spectrometry not only for determining the presence of a protein or PTM, but also for quantitative analysis of interiorly labelled proteomes. Stable isotope labelling of amino acids in cell culture (SILAC) (Ong *et al.*, 2002), selective reaction monitoring (SRM) (Mayya *et al.*, 2006) or multiple reaction monitoring (MRM) (Anderson *et al.*, 2004) are further emerging mass spectrometric strategies.

The introduction of affinity tags to enrich certain samples is the third important application in chemical proteomics. The cleavable isotope-code affinity tag (cICAT) originated from the isotope-code affinity tag (ICAT) technique, which is a method to quantify the proteome (Gygi *et al.*, 1999). Unlike ICAT, the tags in cICAT incorporate a cleavage site in the linker region to its solid support (Zhou *et al.*, 2002, Hansen *et al.*, 2003). For cleavage, photoreactions or acids are used and

cysteine containing peptides are employed for attachment to the solid support. Similar to cICAT, visible isotope-coded affinity tag (VICAT) reagents tag thiols from proteins, by introducing a biotin affinity handle and a photocleavable linker for removing a portion of the tag, and an isotope tag for distinguishing sample and internal standard peptides (Bottari *et al.*, 2004). Disulfide-exchange chromatography is used to enrich cysteine-containing peptides (Wang *et al.*, 2001). The guanidine group of arginine can be reacted with 2,3-butanedione and a solid based boronate reagent to form bicyclo ring, which is labile for acid. This method enables to selective capture arginine containing proteins (Foettinger *et al.*, 2005).

PTM enrichment and characterization is the fourth major and significant part in chemical proteomics. In this study field, chemical proteomic approaches show almost unique properties. Phosphorylation and glycosylation are currently two intense study areas for PTM. Terms like phosphoproteomics or glycoproteomics were coined to show their special status as sub-proteomics study fields. Quantitation of natural phosphorylation is a major issue for analyzing phosphorylated proteins, and selective enrichment of phosphorylated peptides after digestion is required. Immobilized metal affinity chromatography (IMAC) with different metal ions or oxides such as Fe^{3+} , Ga^{3+} , TiO_2 , ZrO_2 , Al_2O_3 were developed (Sun *et al.*, 2005, Schlosser *et al.*, 2005, Kweon *et al.*, 2006, Wolschin *et al.*, 2005). Alternatively, β -elimination of phosphate groups in pSer and pThr residues under alkaline conditions, followed by a Michael type addition of a thiol associated stable-isotope tags can also enrich or quantify the phosphorylated proteins (Adamczyk *et al.*, 2002, Goshe *et al.*, 2002, Amoresano *et al.*, 2004). The study of PTM for glycosylation also focuses for the enrichment of glycosylated proteins. Besides traditional bioaffinity methods such as lectins (Kaji *et al.*, 2003), chemical approaches were introduced, such as oxidation of the *cis*-diol to aldehyde, followed by a reaction with solid support hydrazine to form a hydrazone bond to immobilize

the glycosylated proteins (Zhang *et al.*, 2003). Similar to phosphorylation, the *N*-acetyl-glucosamine on serine or threonine residues can also be β -eliminated under basic conditions (Vosseller, *et al.*, 2006). Other PTMs, such as acetylation, methylation, S-nitrosylation and others were also studied by chemical proteomic approaches (Jaffrey *et al.* 2001, Hess *et al.* 2005).

Some other techniques were and still are appearing in chemical proteomics. Instead of fluorescent fusion proteins, small organic molecules imaging techniques show potential advantage (Fernández-Suárez *et al.*, 2008). PPIs can be influenced by small molecules, such as 14-3-3 proteins with fusicoccin (Malerba *et al.*, 2004). Chemical microarrays are used for drug screening and discovery (Ma *et al.*, 2006). Similiar to Y2H, Y3H is also a method to expand the Y2H principle (Schneider *et al.*, 2008) and is another milestone showing that chemical proteomics enlarges conventional functional proteomics.

Chemical proteomics is still a young family member in the field of life science, however due to its noble lineage from chemistry, it is destined to become one of the most prominent parts of proteomics. The publications achieved during only a short time period and earned high attentions, which is the best verification for its great potential.

1.6 Chemical biology in chemical proteomics

When scientists started to address biological questions based at the molecular level, the new interdisciplinary chemical biology was born. Chemical biology was already reported in the early twentieth century (Dale, 1934, Beard, 1952), and the first publication could date back to 1916 (Henderson, 1916). The term chemical biology however first appeared in publications from the sixties (Stirewalt, 1963).

One part of chemical biology is categorized as chemical genetics, which is defined as using chemical tools to study genetics (Mitchison, 1994, Schreiber,

1998). Forward chemical genetics (FCG) follows the way from effect to cause (“from phenotype to genotype” in forward genetics), first to screen the effect, that an inhibited or stimulated protein induces phenotype changes, and then to identify the small molecules causing the effects by regulation of the functions of certain proteins (Lokey 2003, Burdine *et al.*, 2004). In contrast, reverse chemical genetics (RCG) follows the way from cause to effect, (“from genotype to phenotype” in reverse genetics), first to search the cause, i. e. a ligand stimulating or inhibiting the important protein by screening compound libraries, and then to study the effect, that introduction of that specific ligand to a cell or organism to prove the resulting changes in phenotype (Blackwell *et al.*, 2003, Mayer, 2003). Chemical compounds libraries enable either forward or reverse chemical genetics. The required chemical compounds libraries consists for example of natural products (NPs), natural product derivatives (NDs), or collections derived from diversity oriented synthesis (DOS), tagged libraries, target guided ligand assembly (TGLA), dynamic combinatorial libraries (DCL) or annotated chemical libraries (ALC) (Walsh *et al.*, 2006). These established chemical tools were also applied for many other chemical biology investigations and not only in chemical genetics screening.

Another part of chemical biology focuses on medicine. The term “medicine” is derived from the Latin *ars medicina* (the art of healing). Considering this, medicine could be defined as the art and science of healing. It encompasses maintaining and restoring health and treating illnesses. To this end, chemical biology for medicine is also applied mainly for two sub-areas, pharmacology and physiology. For its application in pharmacology, clinicians initially argued that those “entering medicine should not waste their time on chemistry” (Rosenfeld, 2003), but after the “magic bullet” was introduced by Paul Ehrlich in the early twentieth century (Ehrlich, 1911), such opinions lost markets swiftly. Those magic bullets are small molecules, which target specific tissues or microbes. This is the

starting point for modern pharmacology. Nowadays the pharmacologic functions of small molecules are more carefully studied, implemented *via* interactions between small molecules with DNA, RNA and proteins (Schreiber, 2005). Unlike pharmacology, physiology focuses on basic questions in life science. The way in which chemical biology is studied in the context of physiology is also different from that for pharmacology. Since physiological study values more about the activity and function of enzymes, activity-based protein profiling (ABPP) and affinity-based profiling (AFP) as two chemical biology methods were specially developed (Fonović *et al.*, 2008, Beroza *et al.*, 2005). AFP is designed to target the non-catalytic residues of proteins or enzymes, whereas ABPP is designed to target the catalytic residue of enzymes and enzymes need to be active to facilitate the labelling. Although there are already many chemical biology techniques in chemical proteomics, they are mainly used for determining the PTM of proteins, measuring the changes of proteins abundance, studying the interacting partners and regulation mechanism of proteins, but not be able to study the activity of proteins. The changes in the phenotype may either be due to changes in the overall amount of proteins, or be caused by the changes in the activity of proteins. ABPP can directly measure the changes in the activity of proteins and is frequently used for studying the activity of proteases (Evans *et al.*, 2006).

The functions of a biological system consist of the function of all interior proteins. Thus, to study the function of a biological system need first to study the function of the individual protein. However to study the function of an individual protein individually is less sufficient and efficient than to study the protein in its natural proteome. That is why the *in vivo* results are more appreciated than the *in vitro* results. Since the studying object in life science already from protein becomes proteome, the principles in system biology are naturally involved in the modern chemical biology.

1.7 Protease for life science

Proteases, which hydrolyse the peptide bonds, form a subclass of the hydrolase enzyme family and are classified as EC 3.4 in the EC number classification of enzymes. Proteases can be classified into subgroups from different aspects. From the substrate aspect, proteases can be classified in two classes, which hydrolyse proteins or hydrolyse peptides. From the hydrolytic position aspect, proteases can be classified as endopeptidases (internal), aminopeptidases (N-terminal exopeptidase) or carboxypeptidases (C-terminal exopeptidase). From the catalytic mechanism aspect, proteases can be classified as serine, cysteine, aspartate, threonine, metallo or glutamate proteases. On the basis of evolutionary relationships, proteases are grouped into clans that share structural similarities (homology) and are then further subgrouped into families with similar sequences. In this thesis work, the proteases are mainly classified by their catalytic mechanisms.

From the protein structural aspect, all proteases can be roughly separated into a catalytic part and a structure maintaining part. The catalytic part sometimes is also called the active site pocket, which is the place for hydrolysis of the substrates. The oxyanion hole is a place inside the active site pocket, to polarize the carbonyl group of the amide bond, which will be hydrolyzed. The oxyanion hole contains Lewis acids such as an amide proton, acid proton or metal ion, which could interact with the oxygen atom of the carbonyl group. The activation energy of the hydrolysis of this amide bond will be reduced by the polarization of the carbonyl group in the oxyanion hole, so that the stable amide bond can be hydrolyzed under physiological conditions. The polarized carbonyl group will be attacked by a nucleophile. In case of serine, threonine or cysteine proteases, the nucleophile is the deprotonated hydroxyl or thiol group of the corresponding serine, threonine or cysteine amino acid. Whereas in case of aspartate, glutamate or metallo proteases,

the nucleophile is an activated water molecule. The substrate binding pocket stabilizes the peptide of the substrates and is responsible for selecting substrates. The selection of substrate pocket is achieved by the interaction between the amino acid of the peptide with each single substrate binding pocket, named as S1, S2, S3... or S1', S2', S3'.... "S" refers to subsite, "S or S'" refers to the direction (to N-terminal or C-terminal), and "1, 2, 3..." refers to the position (Van der Hoorn, 2008). The substrate position associated with the substrate binding pocket is noted as "P1, P2, P3... or P1', P2', P3'", relative to the substrate binding pocket (Nazif *et al.*, 2001).

Serine proteases are the most abundant proteases and their catalytic triad was the first well studied protease catalytic mechanism system (Damaschun *et al.*, 1968, Steitz *et al.*, 1969). The catalytic triad is a coordinated structure consisting of histidine, serine and aspartate as three essential amino acids. The hydroxyl group of the active site serine attacks the carbonyl group in the oxyanion hole, to cleave the peptide bond, release the amine-containing product fragment, and generate a peptidyl-*O*-protease covalent enzyme intermediate (acylation), which will be hydrolyzed in a second half-reaction (deacylation). Serine proteases are often divided into four major clans: chymotrypsin-like, subtilisin-like, alpha/beta hydrolase, and signal peptidase clans (Barrett *et al.*, 1995).

Cysteine proteases are another large group of proteases, which share a similar catalytic mechanism as serine proteases, sometimes lacking the aspartate residue of the catalytic triad. The thiol group of the active site cysteine attacks the carbonyl group to cleave the peptide bond and to generate a peptidyl-*S*-protease intermediate, which will be hydrolyzed in a second half-reaction. Cysteine proteases are often grouped into 7 major clans: CA, CD, CE, CF, CH, PA and PB (Barrett *et al.*, 2001).

Aspartate proteases utilize two aspartate residues to hydrolyse peptide substrates. The acid-base mechanism-based process includes the coordination of a water molecule between these two highly conserved aspartate residues, and one activates the water by abstracting a proton, enabling the water to attack the carbonyl group of the substrate (Brik *et al.*, 2003). Aspartate proteases are normally active at acidic pH values.

Metalloproteases hydrolysis of substrates *via* a metal ion in the active site pocket, which is coordinated by the imidazole groups of three conserved histidines. The fourth coordination position is left for a water molecule, which is activated by the chelated metal ion and deprotonated by the carboxylate group of one adjacent glutamate or aspartate residue to attack the carbonyl group of the substrate (Holz *et al.*, 2003). Most metalloproteases are zinc-dependent, but also cobalt.

The threonine proteases were described for the first time in 1995 in the proteasome (Orlowski *et al.*, 2000). The threonine hydroxyl group and its free amino group are coordinating a water molecule. Following a cyclic deprotonation process, the amino group abstracts a proton from the water molecule and the water molecule abstracts the proton of the hydroxyl group. The deprotonated hydroxyl group as a nucleophile attacks the carbonyl group of the substrate (Bochtler *et al.*, 1999).

Glutamate proteases are similar to aspartate protease and were reported recently (Meijers *et al.*, 2004). Most of these proteases have their zymogens and endogenous inhibitors.

Proteases exist in all organisms and act in a multitude of physiological reactions from simple degradation of proteins to highly regulated cascades. Hence, they are one of the most important enzyme classes in life science. They can either cleave a specific peptide residue *via* limited proteolysis, depending on the amino acid sequence of a protein or degrade a complete peptide into amino acids *via*

unlimited proteolysis. Through these proteolytic processes, proteases can abolish functions of a protein by destroying its essential domain or they can create an active protein from inactive precursors. Thereby proteases can work as biological switches and the regulatory mechanisms are achieved by controlling the lifetime of proteins, which play important physiological role like hormones, antibodies, or other enzymes. Through subsequent cooperative cascade reactions, proteases can effect a rapid and efficient amplification of a physiological response to a biological signal. For example, some acid aspartate proteases, such as pepsin, are secreted into the stomach (Johnson, 1985) and serine proteases, such as trypsin and chymotrypsin, are present in the duodenum (Goldberg, 2000) enable humans to digest the proteins in food; the blood-clotting cascades, the lyses of the clots as well as the correct action of the immune system are highly regulated by blood serum proteases such as thrombin, plasmin or the Hageman factor (Bouma *et al.*, 2006, Amara *et al.*, 2008, Borensztajn *et al.*, 2008, Krupiczajc *et al.*, 2008). Some proteases, such as elastase and cathepsin G present in leukocytes, play several roles in metabolic control (Korkmaz *et al.*, 2008). The cysteine proteases cathepsin K and caspases have vital roles in mammalian cells such as bone resorption, apoptosis, necrosis and inflammation (Le Gall *et al.*, 2008, Franchi *et al.*, 2009, Strasser *et al.*, 2009).

1.8 Background to activity-based protein profiling (ABPP)

Biological regulation consists of functions of numerous proteins, however the function of a protein is dependent on its activity. As discussed in the above sections, for a more detailed study on enzyme functions, we need to consider their activity. The variation of the active states enable enzymes to act as the biological switch for regulation of physiological processes. The activity of enzymes, for example a protease, is regulated by pH, co-enzymes, co-factors, temperature, ion

strength, etc. The ABPP approach is one of the best methods until now to directly study the activities of these enzymes.

The term ABPP was first coined by Benjamin Cravatt in 1999, appearing in the paper “Activity-based protein profiling: the serine hydrolases” (Liu *et al.*, 1999). The principle of ABPP is to utilize the catalytic feature of enzymes and label them with small organic molecules. Those small organic molecules are termed activity-based probes (ABPs), which mediate chemical reactions with active site residue of enzymes. The reactions are taking place only when the enzymes are active. For enzymes utilizing the catalytic amino acid as a nucleophile for the catalytic process, such as serine, cysteine, threonine proteases, glycosidases or phosphatases, ABPs usually react with the catalytic amino acid in the active site and form a covalent bond. For the enzymes utilizing an active water as a nucleophile for the catalytic process, such as aspartate, glutamate or metallo proteases, ABPs coordinate with the amino acids or metal ion, which activate the catalytic water, and, with the help of a photoreactive moiety, crosslink to the amino acid residue in the active pocket. For the non-hydrolytic enzymes such as lipid or protein kinases, ABPs (wortmannin-based probes) react with the amino acid (lysine) in the nucleotide binding site and form a covalent bond. All target enzymes, while covalently bound with ABPs, are permanently inactivated. Biotin or fluorophore tags in ABPs are used to visualize the labelled enzymes. Therefore ABPP is a multi-principles method, which is not only limited for screening or comparing the enzymes activities from different proteomes (Jessani *et al.*, 2002, Jessani *et al.*, 2004), but also can be applied for determining novel enzymes from an unknown proteome combined with different protein identification strategies (Borodovsky *et al.*, 2002, Jessani *et al.*, 2005), for inhibitor discovery (Greenbaum *et al.*, 2002, Leung *et al.*, 2003), for tissue based-imaging (Blum *et al.*, 2005) or the elucidation of novel enzymatic mechanisms (Wang *et al.*, 2008).

Different ABPs are needed to realize the plentiful ABPP strategies. Although the term ABPP was proposed till 1999 with the fluorophosphonate (FP) probe, the history of ABPs dates back to 1997, when Matthew Bogyo and his colleagues developed vinyl sulfone (VS) probes for the proteasome (Bogyo *et al.*, 1997). To enable ABPP, ABPs need three moieties, a warhead, a linker and a reporter tag. The warhead is the active moiety, as an electrophile, reacts with the enzyme to form a covalent bond. Therefore warhead is the most important moiety of ABPs, and usually responsible for selecting enzyme and determining its activity. The reporter tag is responsible for visualization of the labelled enzymes. Depending on the purposes of assays, the reporter tag could be an affinity or fluorescent tag. The linker is a spacer between warhead and reporter tag. A good linker can enhance the selectivity and labelling affinity of ABPs (Jeffery *et al.*, 2003).

According to the warhead design strategy, ABPs can be classified into four major subgroups: natural product-based, mechanism-based, direct and non-directed probes. Natural product-based probes are probes derived directly from a natural product, utilizing the probe to find out its related biological targets. The natural product has a reactive moiety and can form a covalent bond with its target. Successful examples are like E-64-based probes for the papain family of cysteine proteases, derived from E-64 a metabolite of *Aspergillus japonicus* (Hanada *et al.*, 1978, Bogyo *et al.*, 2000); wortmannin-based probes for lipid and protein kinases, derived from a furanosteroid metabolite of *Penicillium funiculosum* *Talaromyces (Penicillium) wortmannii*, a covalent inhibitor of phosphoinositide 3-kinases (PI3Ks) (Stoyanova *et al.*, 1997, Yee *et al.*, 2005); microcystin-based probes for serine or threonine phosphatases, derived from the cyclic nonribosomal peptide and cyanotoxin produced by cyanobacteria (Gupta *et al.*, 1997, Shreder *et al.*, 2004). Mechanism-based probes mimic the catalytic transition state of the substrate within an enzyme. Because the probe mimics the transition state, it can interact

with the enzyme and bind the enzyme *via* a covalent bond. Examples are like FP probes for serine hydrolases (Liu *et al.*, 1999); acyloxymethyl ketone (AOMK) probes for cysteine proteases (Kato *et al.*, 2005); VS probes for the proteasome and ubiquitin-specific cysteine proteases (Bogyo *et al.*, 1997, Borodovsky *et al.*, 2002); 2-Deoxy-2-fluoro glycoside probes for retaining *exo* and *endo*-glycosidases (Hekmat *et al.*, 2005) and α -Bromobenzylphosphonate for tyrosine phosphatases (Kumar *et al.*, 2004). Photoreactive probes are special mechanism-based probes for non-covalent substrate binding enzymes, such as metallo or aspartate protease. The photoreactive group, such as benzophenone, can crosslink enzymes covalently under UV irradiation. The reported examples are photoreactive hydroxamate probes for metalloproteases (Chan *et al.*, 2004) and photoreactive hydroxyl ethylene probes for aspartate proteases (Li *et al.*, 2000). A direct probe is a probe derived from a known inhibitor, which has usually defined targets. Whereas the non-directed probe is a probe based on a chemical reactive moiety and the targets of these probes are usually difficult to predicted before the experiment. The non-directed probes were first introduced by Benjamin Cravatt and a well known example is sulfonate ester (SE) based probes (Adam *et al.*, 2002), however many other probes have been developed in the meantime.

According to the reporter tag, ABPs can also be classified into three versions: affinity tagged, fluorophore or radioactive tagged and click chemistry support probes. Affinity tagged probes are widely used for purification or identification of labelled proteins. The most common affinity tag is biotin, which tightly binds avidin. The disadvantage of biotin is poor solubility and membrane permeability. Fluorophore tagged probes are very useful for profiling and imaging, comparable to conventional GFP or YFP. The fluorophore can be chosen depending on the assay conditions. Modified quencher fluorophore ABPs were introduced by Matthew Bogyo, revealed the unique advantage of activity-based imaging (Blum *et al.*,

2005). Click chemistry support probes combine the bio-orthogonal chemical labelling technique with ABPP methodology, enabling ABPs to behave more similarly to inhibitors *in vivo*, minimizing the influence of the tag on the warhead. It was introduced by Benjamin Cravatt (Speers *et al.*, 2003). The trifunctional click pattern enables click chemistry support probes to be more powerful (Speers *et al.*, 2004).

The last part of this section will shortly discuss the relationship between ABPs and inhibitors. ABPs and inhibitors are in principle very similar. However, ABPs must irreversibly bind the enzymes whereas inhibitors can either bind reversibly or irreversibly. Irreversible inhibitors can be directly converted into ABPs as directed probes, whereas reversible inhibitors, e. g. based on an aldehyde, boron acid, α -keto carbonyl, α -keto amide or α -keto aldehyde as active moieties (Bogyo *et al.*, 2002) can only be transformed into ABPs by adding a photoreactive group. ABPs need a reporter tag and linker whereas inhibitors do not require them. Therefore all APBs are synthetic products whereas inhibitors can also be natural products. ABPs are usually designed to label the enzymes in an one-to-one stoichiometry whereas inhibitors can be synthesized to label the enzymes in one-to-more stoichiometry (Loidl *et al.*, 1999a,b).

1.9 ABPP for plant biology

Organisms are divided into two domains: prokaryotes and eukaryotes. Eukaryotes can be further classified into the kingdoms of animal, plant, fungi and protist. Perhaps for the reason of cognate relationship, the animal kingdom is most well-studied. A search of the reports of chemical biology study in PubMed (www.pubmed.org) till 14.07.2009, the reports for animal science are 15294 pieces, almost five times as much as the reports for plant science (3219 pieces). This demonstrates that the research intensity in plant biology is much less strong than in

animal biology. Generally, animals are considered as a higher life form than plants, and therefore most advantage methodologies are applied priorly in animal biology. However, botanic system shares many similarities with animal system, so the study of plant biology could be of benefit for animal biology and vice versa. Plant biology even has the advantage for certain studies for juristic or ethic reasons. Hence, we are applying chemical biology tools to study plant biology.

Immunity is a common and essential physiological phenomenon, occurring in all organisms. Surprisingly, the plant immune system is similar to the animal innate immune system but features its own characters (Jones *et al.*, 2006). The plant immune system uses two large branches of disease resistance (R) proteins to defend microbes or pathogens. The first type, recognising microbial- or pathogen-associated molecular patterns (MAMPs or PAMPs), uses transmembrane pattern recognition receptors (PRRs). The second type, targeting pathogen effectors (proteins secreted by pathogen with virulent or avirulent effects) from diverse kingdoms uses polymorphic nucleotide binding leucine rich repeat domain (NB-LRR) motif proteins. The defense machinery works in a “zigzag” way, which consists of four phases. In phase one, PRRs result in a PAMP-triggered immunity (PTI) with recognising PAMPs (or MAMPs). In phase two, the PTI is interfered by pathogen virulent effectors, which leads to an effector-triggered susceptibility (ETS). In phase three, certain pathogen avirulent effectors (Avr) can be recognised by NB-LRR proteins by forming cognate Avr-R pairs according to the “gene-for-gene” theory (Flor, 1971). This causes an effector-triggered immunity (ETI). ETI accelerates and amplifies PTI, leading to disease resistance and usually a hypersensitive cell death response (HR). In phase four, due to natural selection of pathogens, ETI is suppressed by shedded or diversified pathogen effectors. Nitric oxide, reactive oxygen species (ROS), salicylic acid (SA) and jasmonates are such small molecules in plants, joining defense processes.

It is known that proteases play very important roles in the plant-pathogen interaction. Indeed, two recent reviews have discussed the function of plant proteases (Van der Hoorn *et al.*, 2008a,b). Plant cysteine proteases such as papain-like proteases Rcr3 and *NbCathB* are associated with HR (Rooney *et al.*, 2005, Gilroy *et al.*, 2007); caspase-like proteases mcII-Pa and VPEs are essential for program cell death (PCD) (Suarez *et al.*, 2004, Hatsugai *et al.*, 2006). Plant serine proteases such as carboxypeptidase-like proteases SNG1/2 are associated with UV-B protection (Lehfeldt *et al.*, 2000, Lorenzen *et al.*, 1996). Plant aspartate proteases such as pepsin-like proteases CDR1 and PCS1 induce high levels of reactive oxygen intermediates (ROIs) and SA or prevent cell death during gametogenesis (Xia *et al.*, 2004, Ge *et al.*, 2005). The other plant hydrolases like soybean *endo*- β -1,3-glucanase-A (EGaseA) are also defense-related enzymes (Rose *et al.*, 2002). The pathogen hydrolases such as *endo*- β -1,4-xylanases (GH10/11) and polygalacturonases (PGs) are important for pathogens to degrade the plant cell wall (Juge, 2006).

The application of ABPP for plant enzymes (e. g. hydrolases) studies can reveal novel aspects of the plant-pathogen interactions. ABPP of the plant proteome with known or novel ABPs can reveal novel enzymes, which might be responsible for the regulation of important defense processes in plants. Comparison of the ABPP results between diseased and normal plant proteomes can directly and quickly display variation of the activity for certain enzymes *in vitro* or *in vivo*. ABPs can also be used for inhibition of selected enzymes, or to observe phenotype changes. Fluorescent ABPs can visualize enzymes *in vivo* to study their localization, cyclization and lifetime.

2 Aim of the PhD mission

In chemical biology, chemical tools are used to elucidate complex biological systems. Chemical biology research usually consists of three different stages. First, an interesting biological question is formulated. To answer this question, suitable chemical probes are designed and synthesized in a second stage. These probes are then applied to the biological system to investigate the biological question.

Until now, the only technology to directly demonstrate the role of switching between the active/inactive states of enzymes during plant defence is ABPP. Since the discovery of ABPP ten years ago, several ABPs have been developed and used in pharmacological and physiological studies. However, a broader use of ABPP is still limited by the availability of suitable probes.

In an attempt to expand this limited applicability of ABPP, in this thesis, new types of ABPs were designed, synthesized and evaluated as potential probes for the study of plant-pathogen interactions.

To develop effective ABPs in an efficient way, different design strategies have been tested and discussed accompanying with five questions, to generate different types of novel ABPs:

Lipstatin is a natural product that can covalently bind to the active site of lipases. Its derivate, tetrahydrolipstatin (THL), is a well-known drug used for obesity. The THL-based probes with an *anti*- β -lactone moiety were generated with natural product-based design strategy. In order to target as many lipases as possible, only the reactive scaffold of THL was employed for probe design. This structural simplification, however, implies the danger of losing THL's biological function. How tolerant are the THL-based probes to chemical modification in respect to retaining intrinsic targeting selectivity? This is the first question to discuss ABPs

design strategy and the biological profiling with *anti*- β -lactones was studied consequently (chapter 3.1).

A mechanism-based probe design strategy was applied for two types of probes. *Syn*- β -lactone-based probes were chosen, as a *syn*- β -lactone moiety exists in cysteine protease inhibitors of picornavirus. Additionally, AOMK-based probes were chosen, as the AOMK moiety is a well-known warhead. The advantage of mechanism-based probes is the high possibility to label a certain class of enzymes, whilst the disadvantage is the low labelling selectivity. To overcome this disadvantage, mechanism-based probes with epoxide or AOMK warheads were studied to enhance the labelling selectivity. The second question addressed is that: is it possible to enhance the selectivity of *syn*- β -lactone-based probes with a similar strategy (chapter 3.2)?

In an attempt to develop structurally novel, non-directed probes, azirine and aziridine derivatives were also investigated. These derivatives were chosen as they are natural products with diverse biological activities featuring reactive moieties isolated from various species. With this project, the possibility of transforming azirines and aziridines into ABPs is discussed (chapter 3.3).

Moreover, the development of a direct probe based on the known serine protease inhibitor AEBSF was used to answer the fourth question: is it possible, as often proposed in literature, to directly transform AEBSF into a direct probe (chapter 3.4)?

The last question, being the most essential and challenging task in the thesis is: can we design specific probes for targets of interest in plant biology? The probes with the well-studied AOMK warhead were designed to investigate two interesting biological targets; the role of an avirulent protein AvrPphB in plant-pathogen interactions (chapter 3.5) and the functions of a plant caspase like-cysteine protease VPE (chapter 3.6).

Finally, the development of reactive natural products into ABPs was also investigated. The reactive natural product Gibbestatin B (GNB) featuring a reactive epoxide moiety was chosen for this purpose. As GNB is not commercially available, prior to probe design, a total synthesis of GNB is required. Consequently, synthetic approaches towards this interesting natural product were pursued in this thesis.

After the successful development of novel probes, they should be evaluated in ABPP assays to characterize their labelling properties. However, the ultimate goal is to use these newly developed ABPs for the biological study of plant-pathogen interactions. To this end, biological experiments employing host plants and plant pathogens were devised and performed.

3 Results and Discussions

3.1 Natural product-based ABP design of *anti*- β -lactones

3.1.1 Introduction

The natural product lipstatin is an inhibitor of pancreatic lipase and was originally isolated from the gram-positive bacterium *Streptomyces toxytricini* (Weibel *et al.*, 1987; Hochuli *et al.*, 1987). Its tetrahydro derivate, known as tetrahydrolipostatin (THL) (trade name Orlistat®), has been developed by Roche for clinic therapy of severe obesity and hyperlipidemia (Figure 1).

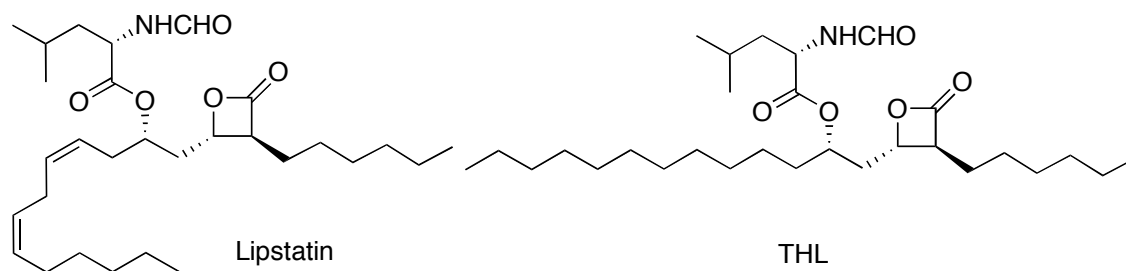


Figure 1. Structure of lipstatin and THL

Many biomedical studies on THL have been published during last two decades (Hauptman *et al.*, 1992; Zhi *et al.*, 1995; Ellrichmann *et al.*, 2007) demonstrating that THL inhibits lipases with high potency by a covalent mechanism, which makes it an ideal template for APBs. Although the direct structure of lipase inhibited by lipstatin or THL is still not reported, the crystal structure of thioesterase with THL showed that the hydroxyl group of serine attacked the carboxyl group of the β -lactone moiety, leading to ring opening and formation of a stable enzyme-acyl complex (Pemble *et al.*, 2007; Figure 2).

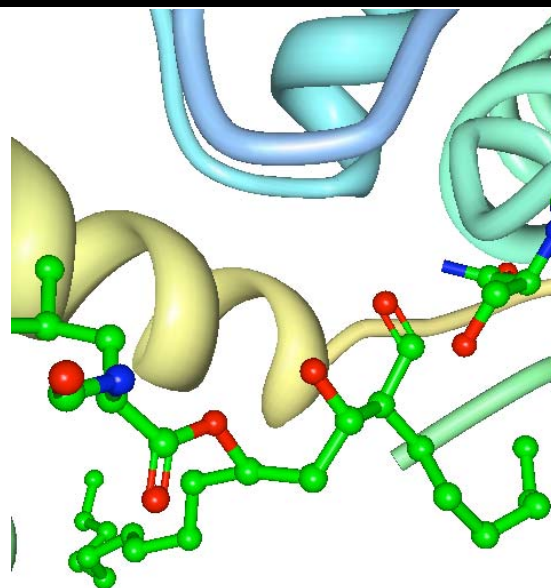
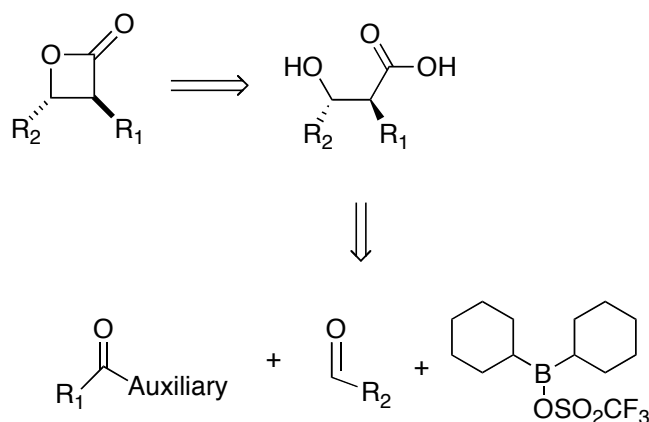


Figure 2. Crystal structure of the thioesterase domain of human fatty acid synthase (FAS) inhibited by THL: the serine (S-113) in the active site attacks the β -lactone moiety of THL to open the β -lactone ring and to form a new ester.

Several targets of THL in mammalian organisms are already known but no targets in plants have been described yet. To investigate the targets of THL in plants, ABPs based on THL were developed. To synthesize THL-based probes, several approaches can be followed. The first total synthesis of THL has been accomplished in 1993 (Hanessian *et al.*, 1993) and biosynthesis of lipstatin has also been investigated heavily (Eisenreich *et al.*, 1997; Goese *et al.*, 2000; Schuhr *et al.*, 2002). Key step of either total or biosynthesis is the formation of the 2-oxetanone moiety by an aldol reaction. Consequently, an aldol reaction also represents the key step of the synthesis of THL-based probes. As structure-activity relationship studies of THL-based lipase inhibition have revealed that the formylated leucine moiety in THL is not essential for inhibition, this moiety was not included in the probe design to simplify the required synthesis.

3.1.2 Syntheses

To generate THL-based probes, an asymmetric synthesis of *anti*- β -lactone is highly advantageous as they could be directly transformed into corresponding lactones by simple esterification of the required *anti*-(3*S*)-hydroxy-(2*S*)-alkyl-carboxylic acid derivatives, which can be obtained *via* an asymmetric aldol reaction between a corresponding aldehyde and a carbonyl compound attached to a chiral auxiliary by the Masamune methodology (Masamune *et al.*, 1996). Based on this strategy, the retro-synthetic analysis of THL-based probes requires two synthetic building blocks, i.e. a carboxyl auxiliary derivative, an aldehyde derivative and an organoborane-based reagent (Scheme 1).

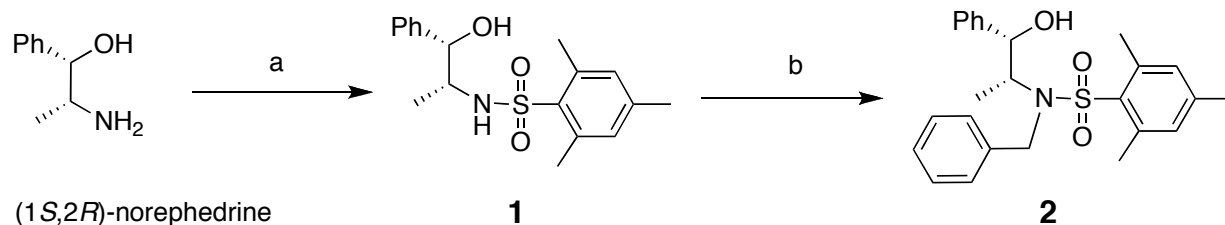


Scheme 1. Retro synthetic analysis of THL-based probes

In principle, the required *anti*-product could be obtained also with Evans-type chiral auxiliaries (Evans *et al.*, 1981; Evans *et al.*, 1982; Brown *et al.*, 1992), that deliver a *syn*-aldol product, which can then be converted to the *anti*-aldol product *via* Mitsunobu inversion of the 3-hydroxyl group (Mitsunobu *et al.*, 1967). Masamune-type auxiliaries however lead directly to the desired *anti*-aldol product and are therefore for our requirements more suitable. Two kinds of Masamune-types auxiliaries have been developed so far: The *S*-3-(3-ethyl)pentyl thioate-based auxiliaries deliver a very good *anti/syn*-selection ($\sim 30:1$) and ee % ($> 90\%$)

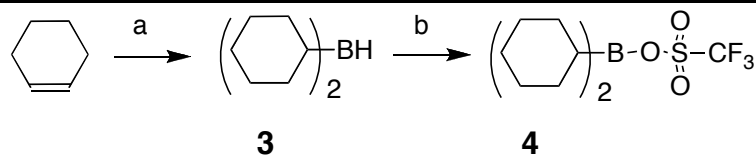
(Masamune *et al.*, 1986). However these thioates are not stable under basic conditions and the preparation of the required chiral organoborane is not convenient (Masamune *et al.*, 1985). The second type is based on chiral ester derivatives (Masamune *et al.*, 1996), which we chose for convenience for the synthesis of probes.

The auxiliary that we used is derived from the natural product norephedrine of which both enantiomers are commercially available. The synthesis was achieved following Masamune's protocol (Atsushi *et al.*, 1997), using (1*S*, 2*R*)-(+)-norephedrine to obtain the correct enantiomer for the THL-probes (Scheme 2). The amino group of (1*S*, 2*R*)-(+)-norephedrine was sulfonated with triMsCl to form **1**, using the organic base triethylamine. The sulfonamide group of **1** was subsequently benzylated with benzyl bromide and caesium carbonate to yield the final auxiliary **2**. All steps of the syntheses were convenient and occurred with high yields.



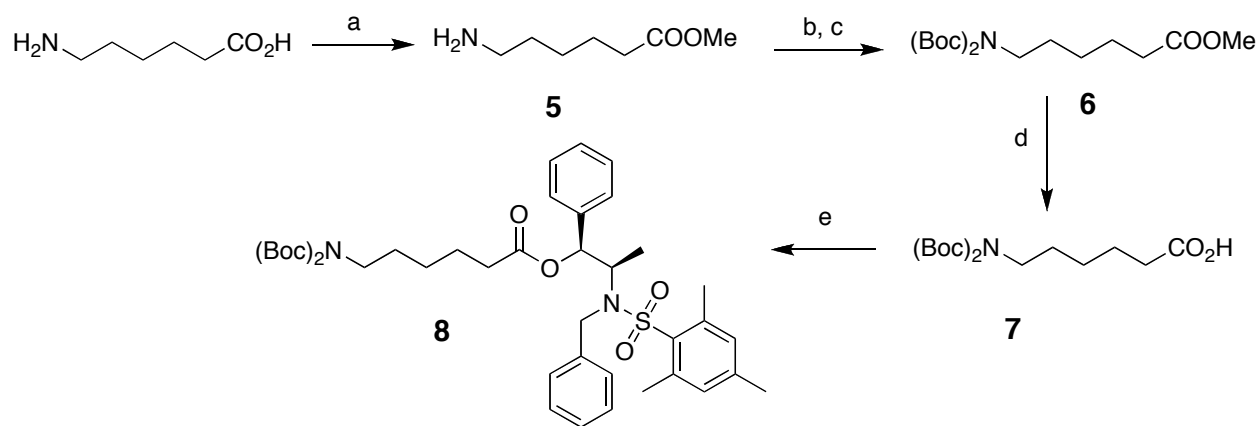
Scheme 2. Synthesis of the Masamune auxiliary: a. TEA, triMsCl, DCM, 0 °C – rt, 3 h, 99%; b. BzBr, Cs₂CO₃, MeCN, reflux, 0.5 h, 95%.

The preparation of the dicyclohexylboron triflate reagent was performed similar to the protocol of Atsushi (Atsushi, 2004) (Scheme 3), starting from the addition of borane to cyclohexene to obtain dicyclohexylboron **3**. **3** was then reacted with trifluoromethanesulfonic acid to generate the organoborane reagent **4**. The handling of trifluoromethanesulfonic acid proved as critical for the reaction, as it reacts heavily with traces of water including air moisture. Although the synthesized boron triflate is known as very sensitive, it could be stored under argon at 4 °C for half a month in a hexane solution.



Scheme 3. Synthesis of dicyclohexylboron triflate: a. $\text{BH}_3\text{-SMe}_2$, diethyl ether, $0\text{ }^\circ\text{C}$, 3 h, 99%; b. $\text{CF}_3\text{SO}_3\text{H}$, hexane, rt, 3 h, 89%.

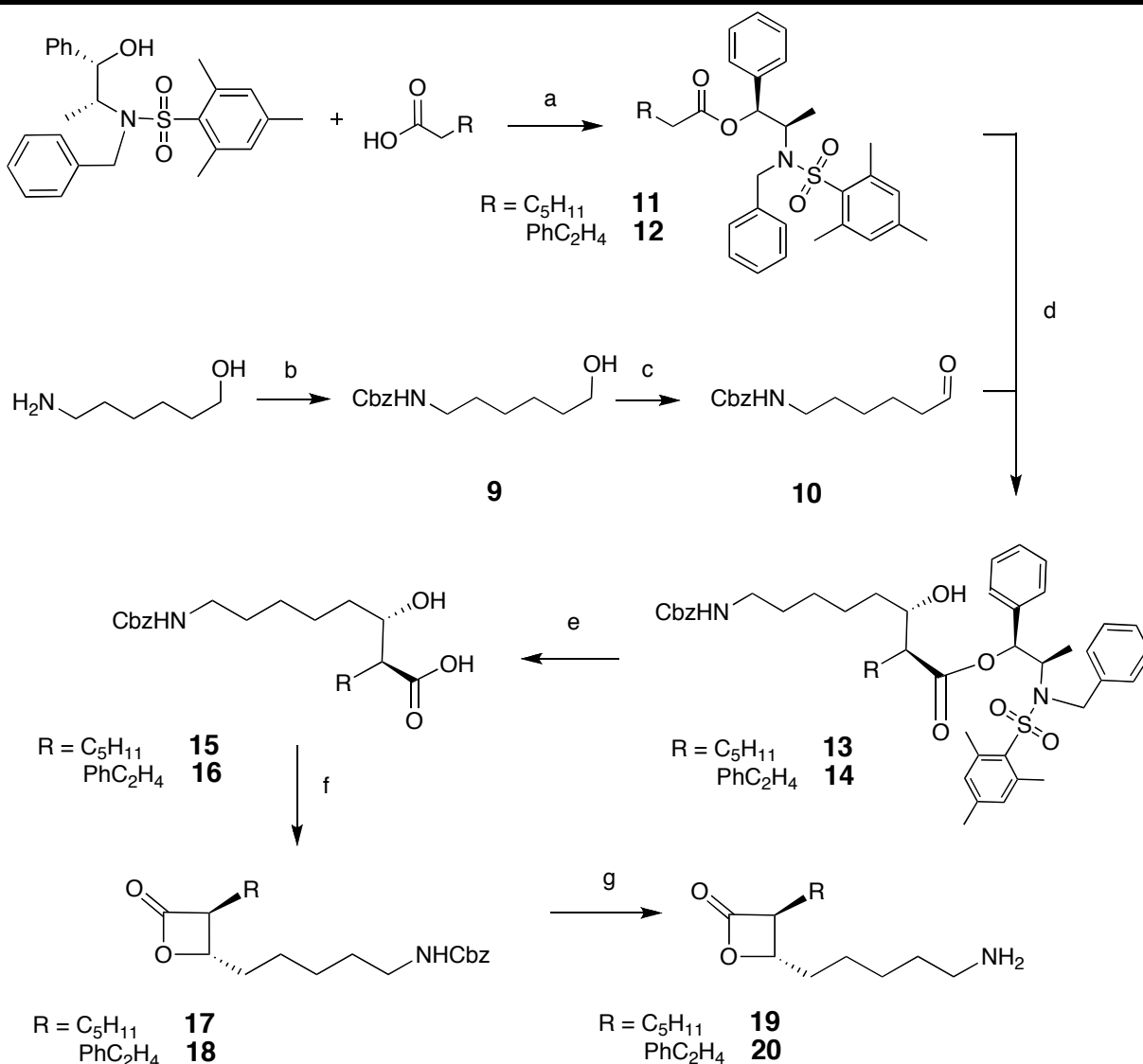
All ABPs require an appropriate reporter tag. For THL-based probes, two possibilities to link the tag to the molecules are reasonable, being either an attachment on the ester or on the aldehyde moiety. To link this tag to the reactive β -lactone warhead, an amide linker was chosen. To incorporate the required amine moiety into the molecule, we therefore employed an appropriately protected 6-amino hexanoic acid as a starting carboxyl building block for the aldol reaction. Consequently, derivate **8** was prepared (Scheme 4). Towards this purpose, 6-amino-hexanoic acid was transferred into its methyl ester **5**, and then its free amino group was *di*-Boc-protected in two steps to yield **6**. Hydrolysis of the methyl ester of **6** with the strong base *n*- Bu_4NOH led to **7**, which was then coupled with the chiral Masamune auxiliary **2** to obtain the desired building block **8**.



Scheme 4. Synthesis of building block **8**: a. SOCl_2 , MeOH, $0\text{ }^\circ\text{C}$, 8 h, 100%; b. $(\text{Boc})_2\text{O}$, Na_2CO_3 , dioxane/ H_2O (1:1), $0\text{ }^\circ\text{C}$ - rt, overnight, 100%; c. $(\text{Boc})_2\text{O}$, DMAP, THF, reflux, 21 h, 90%; d. *n*- Bu_4NOH , MeCN/ H_2O , rt, 30 min, 100%; e. **2**, DCC, DMAP, DCM, rt, overnight, 95%.

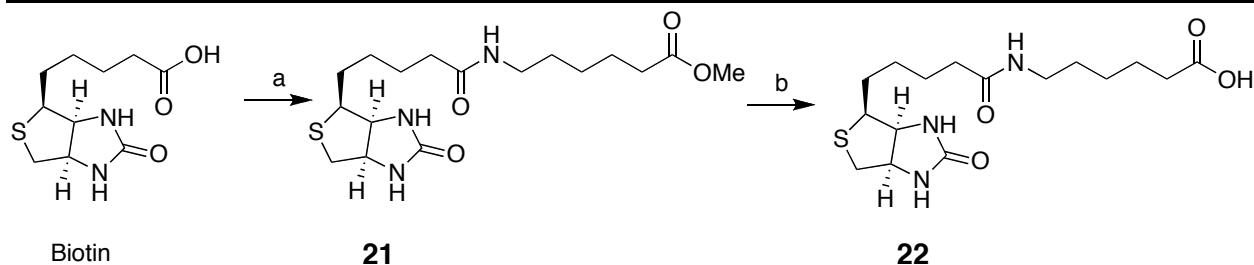
Unfortunately, **8** was instable under aldol reaction conditions. Most probably, the di-Boc protecting group was incompatible to the strongly electrophilic boron

reagent. The synthetic plan was therefore revised (Scheme 5) by employment of a benzyloxycarbonyl (Cbz) protecting group instead of the *di*-Boc protection. In addition, this amino group was introduced this time into the aldehyde building block **10**, which was derived from 6-amino-hexanol was *via* *N*-Cbz protection and oxidation with IBX in two steps. As a carboxylic acid building block, two lipophilic carboxylic acids, i.e. heptanoic acid or 4-phenyl-butanoic acid were coupled with the chiral auxiliary **2** to form the corresponding esters **11** and **12**. The Masamune *anti* aldol reaction of **10** and **11** or **12** formed the desired *anti*-(3*S*)-hydroxy-(2*S*)-alkyl-carboxylic acid esters **13** and **14**. The auxiliary of **13** and **14** was removed by hydrolysis with the mild base lithium hydroxide. Cyclization between the resulting free carboxyl group and the β -hydroxyl group with benzenesulfonyl chloride at 0 °C for two days and the subsequent Cbz deprotection by hydrogenation led to the molecules **19** and **20**.



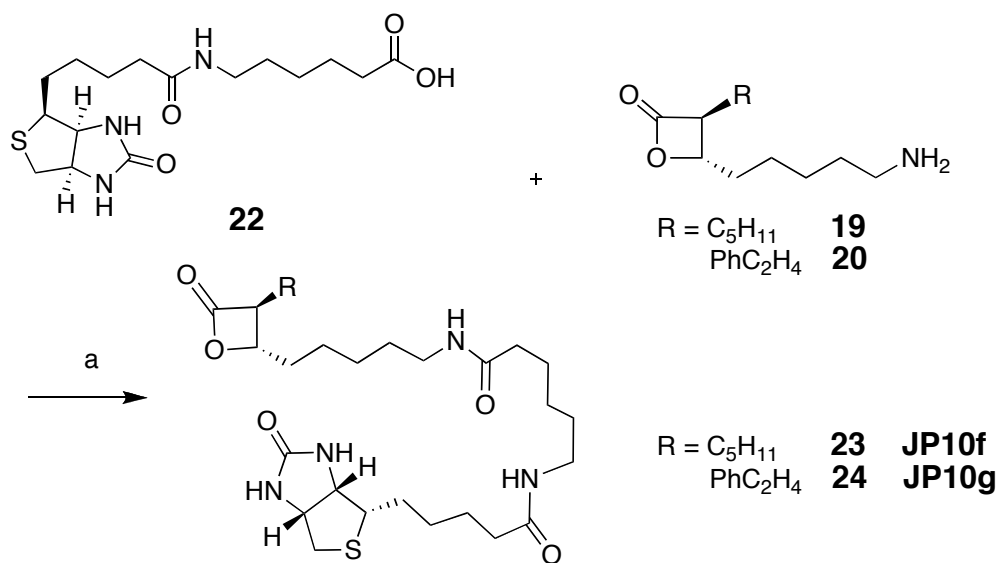
Scheme 5. Synthesis of β -lactones molecules: a. DCC, DMAP, DCM, rt, overnight, 84%; b. CbzCl, NaHCO_3 , dioxane/ H_2O (1:1), rt, 4 h, 87%; c. IBX, THF, 0 °C - rt, overnight, 95%; d. **4**, TEA, DCM, -78 – 0 °C, 4 h, H_2O_2 , pH=7, MeOH, rt, overnight, 88%; e. LiOH, THF/ H_2O (1:1), rt, 3 d, 75%; f. Py, PhSO_2Cl , 0 °C, 2 d, 70%; g. Pd/C, H_2 , EtOH, rt, overnight, 100%.

In the next step, an appropriate tag was added to these intermediates. For this type of probes, biotin as a suitable tag was chosen. To minimize the spatial influence of biotin on lipase inhibition, an additional spacer was introduced between the biotin and the β -lactone moiety (Scheme 6). Towards this purpose, biotin was coupled with previously prepared **5** (Scheme 4) to yield **6**, which was subsequently hydrolysed with *n*-Bu₄OH to obtain the biotinyl linker **22**.



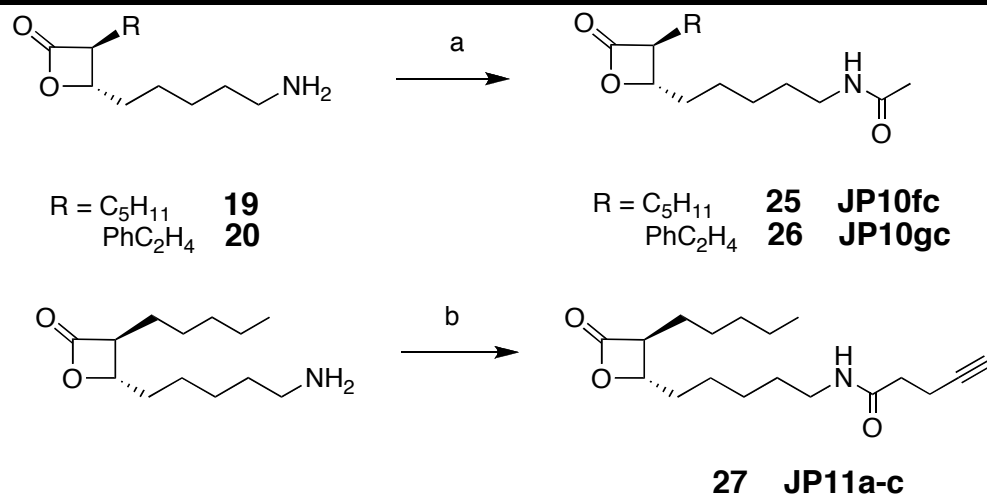
Scheme 6. Synthesis of the biotinyl linker: a. **5**, HBTU, TEA, MeCN, rt, 3 h, 78%; b. *n*-Bu₄NOH, MeCN/H₂O, rt, 30 min, 84%.

22 was then coupled with **19** and **20** to generate the probes **23** and **24** respectively (Scheme 7). Because of the low solubility of biotin in DCM, the coupling reaction was performed in a solution of DCM containing a minimal amount of DMSO to solubilize all reagents.



Scheme 7. Synthesis of *anti*- β -lactone probes: a. PyBOP, HOBt, TEA, DCM/DMSO, rt, overnight, 70%.

For biological competition experiments, also the non-tagged, acetylated derivatives **25** and **26** were synthesized by acetylation of **19** and **20**. In addition, for click-chemistry based profiling, also a probe **27** with an alkyne tag was generated by coupling of 4-pentynoic acid to **19** using DIC/HOBt as coupling reagents (Scheme 8).



Scheme 8. Synthesis of *anti*- β -lactone probes **25** – **27** from **19** and **20**: a. Ac_2O , DIPEA, DCM, rt, overnight, 57%; b. 4-pentynoic acid, DIC, DCM, rt, 4 h, 40%.

3.1.3 Bioassays

To test the reactivity of the designed probes, commercially available wheat germ lipase and *Pseudomonas fluorescense* lipase were tested in a labelling assay (Figure 3). Both probes **23** and **24** caused biotinylation of a 35 kDa and a 42 kDa proteins of wheat germ lipase, whereas no labelling was observed with the *Pseudomonas fluorescense* lipase. The reactivity of the probes to a plant lipase encouraged us to profile the probes in *Arabidopsis* proteomes.

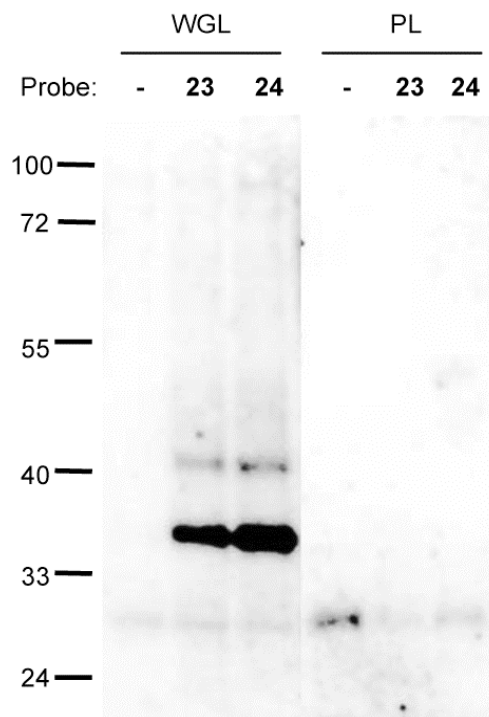


Figure 3. Probe **23** and **24** in vitro evaluation with commercially available lipases: wheat germ lipase Type I (WGL); *Pseudomonas fluorescense* lipase (PL); the commercial enzyme was prepared as a 0.5 mg/ml aqueous solution, and that solution was incubated with 4 μ M probe for 3 h. The proteins on the protein blots were detected by streptavidin-HRP.

Labelling with probes **23** and **24** was performed on leaf and cell culture extracts (Figure 4). In leaf extract, probe **23** caused two bands of 27 and 35 kDa, the lower band was stronger than the higher band. Probe **24** only gave one band of 27 kDa. In cell culture extract, both probes caused 27 and 35 kDa signals and signals at 42 and 56 kDa. There was also one strong 30 kDa background band that is also present in the no-probe-control. The labelling intensity from cell culture extract was higher because of a higher protein concentration. These data suggest that probes **23** and **24** share the same labelling targets but with different affinities. At high protein concentration (like in cell culture), both probes would cause a similar labelling pattern. The different affinity of both probes is caused by a different side chain. The longer aliphatic chain of **23** is more similar to the side

chain of THL, so it could fit the pocket of enzyme better than the aromatic side chain of **24**. Because probe **23** and **24** cause nearly identical labelling profiles, **23** was chosen for subsequent labelling assays.

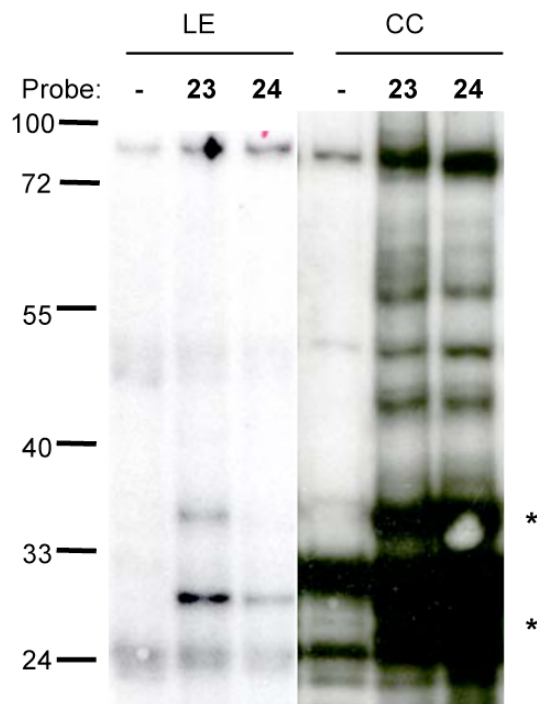


Figure 4. Labelling of *Arabidopsis* leaf extracts (LE) and cell cultures (CC) with probes **23** and **24**: Protein extracts were 5 mg/ml (LE) and 50 mg/ml (CC) and incubated at pH 6 with 4 μ M probes for 3 h. The proteins on the protein blots were detected by streptavidin-HRP.

To investigate the properties of probe **23**, different competitive labelling assays were performed (Figure 5). Since the probe **23** was analogous to THL from a chemical aspect, it was interesting to investigate if they showed the same labelling reactivity. The assay showed that in *Arabidopsis* leaf extracts, the labelling of the 27 and 35 kDa signals could not be competed by THL (Figure 5a). Labelling with probe **23** could also not be competed by E64, ebelactone B and **18**, but by **17** as well as **25** and **26** for the lower band (27 kDa) (Figure 5b). The results from THL, E64 and ebelactone B suggested that the targets of probe **23** were not lipases, cysteine proteases or membrane methyl esterases. **17** was better competing with **23**-labelling than **25** indicating that the benzyl oxy carbonyl residue (**17**) is closer to

biotinyl residue (**23**) than amino acetyl residue (**25**) from the aspect of labelling reactivity. **24** had lower labelling affinity than **23**, so its derivatives **18** and **25** could not compete **23** very well. The molecules with longer side chains (**17**, **23**) had a better affinity to the target, implying that the labelling pocket of target is quite deep.

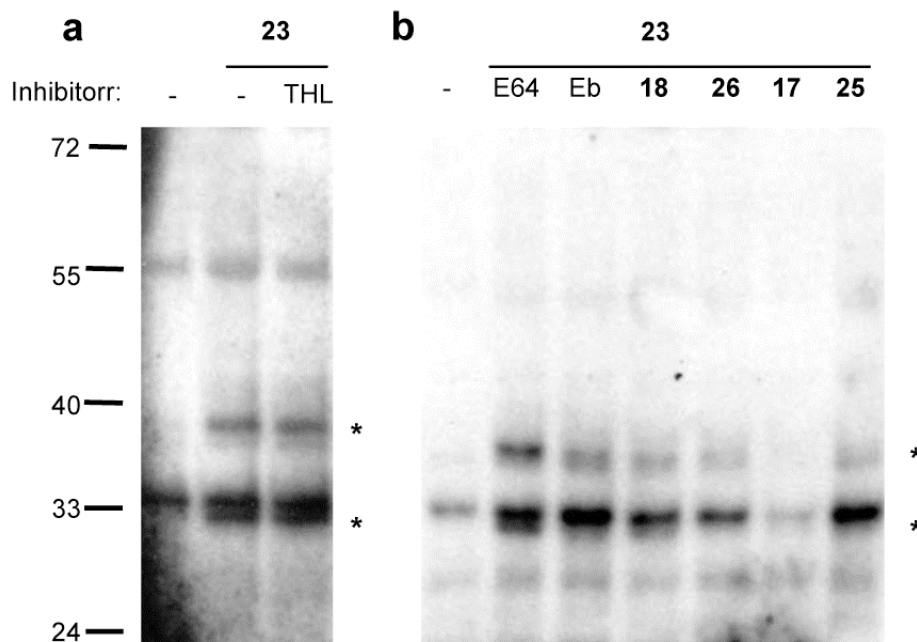


Figure 5. Competition assay of probe **23**: a. Competition with lipase inhibitor THL; b. Competition with cysteine protease inhibitor E64, esterase inhibitor ebelactone B (Eb) and derivatives **17**, **18**, **25** and **26**; Protein extracts (~5 mg/ml) were co-incubated at pH 6 with 40 μ M inhibitors and 4 μ M probe **23** for 3 h. The proteins on the protein blots were detected by streptavidin-HRP.

The competition assays suggest that the targets of the probe **23** were not the expected lipases or esterases. The direct way to identify the targets is to purify the labelled proteins, however all attempts were not successful due to the low abundance of the targets. Nevertheless several assays were performed to study the properties of the labelling. Labelling at various pH values showed that the optimal labelling condition was pH 6 (Figure 6a). At lower pH such as 4, labelling did not occur whereas labelling intensity also decreased with increasing pH value. There was no influence to the labelling by the absence or presence of the reduction reagent *L*-cysteine and labelling was also independent from calcium ions (Figure

6b). Labelling occurred in a very short time and was finished between 30 to 60 minutes (Figure 6c). The localization assay showed that the targets were mainly soluble proteins whereas the background bands were in the membrane (Figure 6d).

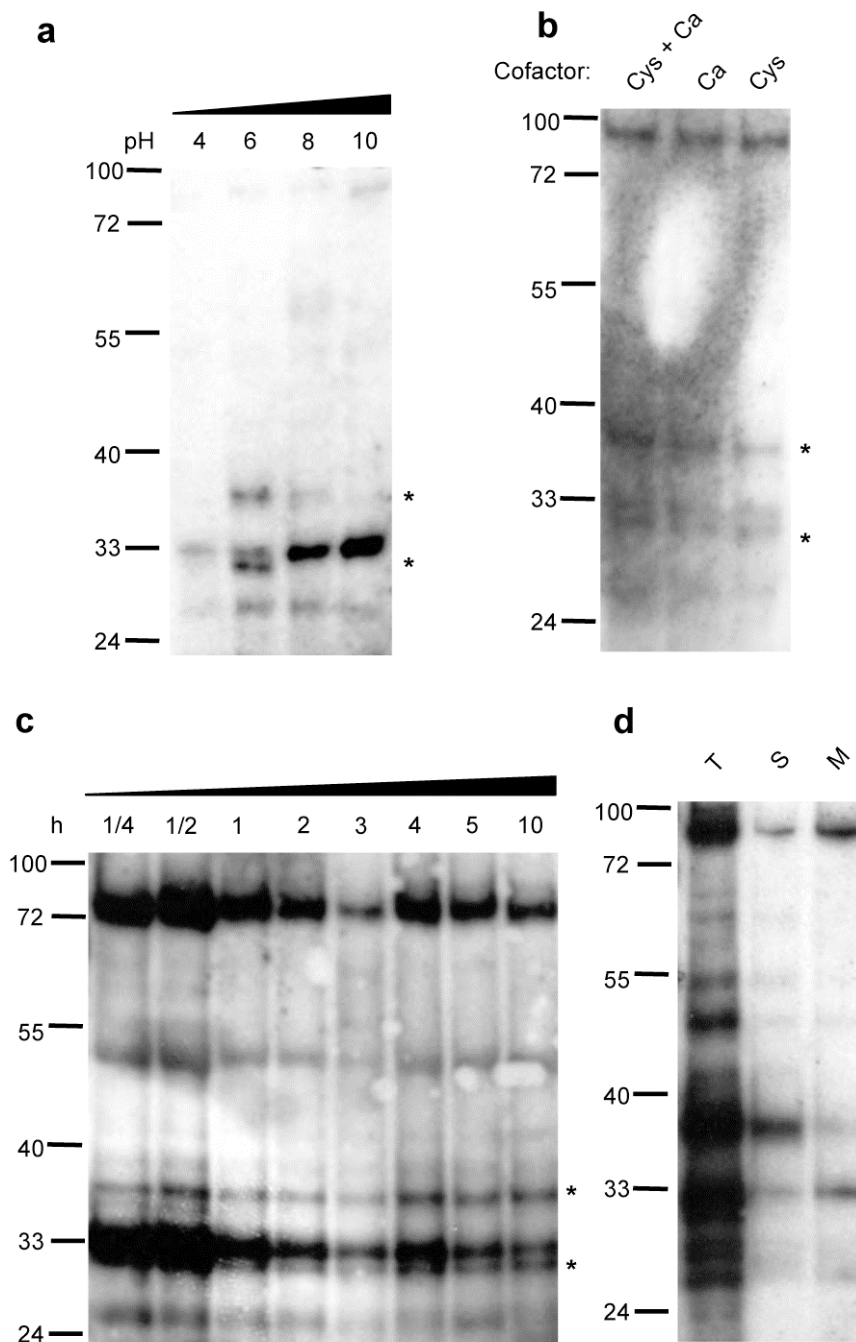


Figure 6. Characterisation of labelling with probe **23**: a. pH curve; b. Influence of cofactors: reduction reagent L-cysteine (Cys) and CaCl_2 (Ca); c. Time course of labelling; d. protein localization analysis: total cell culture (T), soluble protein (S) and membrane protein (P); protein extracts (~5 mg/ml) (except for d, 20 mg/ml), were incubated at pH 6 (except for pH curve) with 4 μM probe for 3 h. The proteins on the protein blots were detected by streptavidin-HRP.

Target identifications in bacterial systems were performed by Thomas Boettcher, a PhD student from the group of Dr. Stephan Sieber (LMU). Click probe **27** was used for *in vivo* labelling of five different bacterial proteomes (Figure 7). 18 labelled proteins were identified from the cytosol and from the membrane fractions. The probe caused many labelling signals in bacterial proteome.

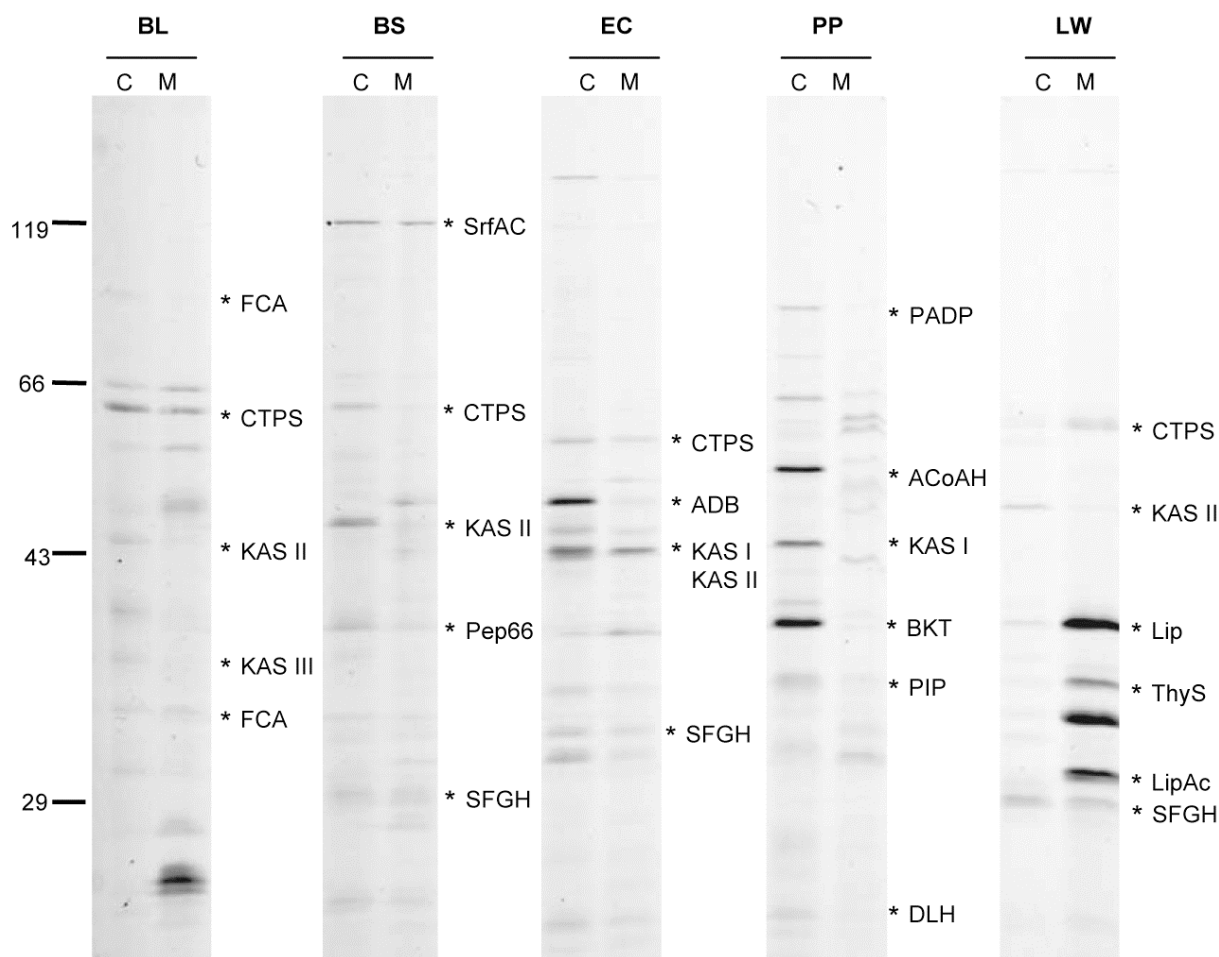


Figure 7. Target identification of probe **27** in bacterial proteome: Five different bacteria [*Bacillus licheniformis* (BL), *Bacillus subtilis* (BS), *Escherichia coli* (EC), *Pseudomonas putida* (PP), *Listeria welshimeri* (LW)] were labelled with probe **27** *in vivo*; the bacterial proteins were separated as cytosol proteins (C) and membrane proteins (M) in advance; The proteins were separated by SDS-PAGE and detected by an azide linked fluorescence tag.

3.1.4 Discussion

The synthesis of the probes was completed in 9 steps in generally good yields. As expected, the auxiliary delivered a good enantiomer selection for the natural product analogues. Unfortunately, the synthesized β -lactone probes proved as rather instable as hydrolysis of the probes in the DMSO stock solution even at -20 °C occurred after already half a year, which limits applicability of these probes. The labelling reactivity was normal in the *Arabidopsis* proteome, and it could label proteins from a proteome of a concentration of 5 mg/ml.

The idea of a natural product-based probe design did not work well in this project. The reason could be the strong modification of the molecule. The peptide moiety on the side chain of THL seems to have a more significant contribution as originally anticipated. Probes lacking this modified side chains seem to lose the original targeting ability of THL for lipases. However, this can still label other proteins, demonstrating the subtle conformational requisites for selective enzyme targeting with natural products.

3.2 Mechanism-based ABP design of *syn*- β -lactones

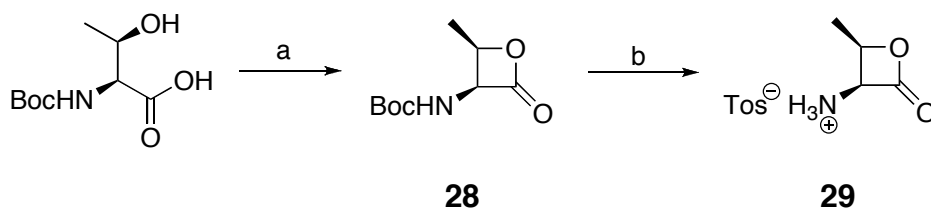
3.2.1 Introduction

The genome of the model plant *Arabidopsis thaliana* encodes for ~320 serine and ~140 cysteine proteases, which include large families of subtilase-like proteases (~60 SLPs), serine carboxypeptidase-like proteins (~60 SCPLs) and papain-like cysteine proteases (~30 PLCPs) (Beers *et al.*, 2004). Studies on some of these presumed plant proteases revealed that they can catalyze nonproteolytic reactions. Phytochelatin synthase, for example, acts as a glutathione transpeptidase, leading to phytochelatin which is required for heavy metal tolerance (Clemens *et al.*, 2006). Furthermore, a number of SCPLs act as acyltransferases in the production of sinapoyl secondary metabolites, which protect plants against UV radiation (Lehfeldt *et al.*, 2000). Some of the other proteases play key regulatory roles in defense and development, but the role, substrate and activation mechanisms of most of these enzymes are unknown (Van der Hoorn *et al.*, 2008).

So far, ABPP of plant proteases was mainly done using DCG-04, a biotinylated version E-64, which inhibits PLCPs (Greenbaum *et al.*, 2000). ABPP with DCG-04 on *Arabidopsis* leaf extracts revealed the activities of six different PLCPs, including RD21 (Responsive-to-dessication-21) and AALP (*Arabidopsis* Aleurain-like Protease) (Van der Hoorn *et al.*, 2004). To expand the range of serine and cysteine proteases that can be monitored by ABPP, a series of activity-based probes containing a 2-oxetanone reactive group were designed. This reactive group is found in covalent inhibitors of hepatitis A virus 3C (HAV-3C) proteinase and the proteasome (Dick *et al.*, 1997; Lall *et al.*, 1999). The inhibition mechanism of HAV-3C proteinase with β -lactones is very well understood, so the mechanism-based *syn*- β -lactone ABPs were developed.

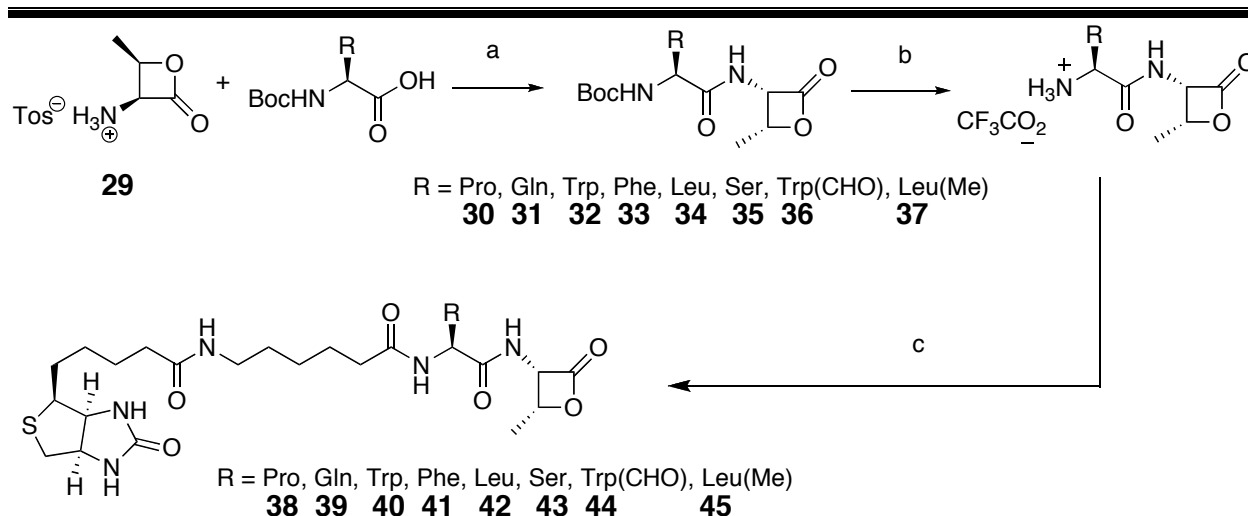
3.2.2 Syntheses

The employed synthetic strategy was straightforward, using the chiral amino acid pool to develop an enantioselective synthesis. Towards this purpose, threonine-based *syn*- β -lactones were generated *via* intramolecular threonine cyclization as established by Dr Rengarajan Balamurugan from Prof. Dr. H. Waldmann's group. Consequently, a *L*-threonine derived β -lactone was generated (Scheme 9). *N*-Boc-protected *L*-threonine as a starting material was cyclized *via* the coupling reagent PyBOP to form **28**. **28** was then deprotected with trifluoroacetic acid and *p*-toluenesulfonic acid to yield the β -lactone tosyl salt **29**, which proved as a good starting material for further coupling reactions.



Scheme 9. Synthesis of the threonine based β -lactone tosyl salt: a, PyBOP, TEA, DCM, 0 °C – rt, 6 h, 85%; b. TFA, PTSA, 0 °C. 15 min, 97%.

According to a previously published X-ray structure (Yin *et al.*, 2005), the β -lactone moiety should occupy the S1 pocket of the protease; therefore, an additional *L*-amino acid attached to its amino group should accommodate the S2 pocket, leading to more specific labelling. Consequently, eight different natural or non-natural amino acids were coupled with **29** to form *syn*- β -lactone derivatives **30** – **37**, which were then deprotected with trifluoroacetic acid and coupled with **22** to obtain the final threonine-based *syn*- β -lactone probes **38** – **45** (Scheme 10).



Scheme 10. Synthesis of threonine-based *syn*- β -lactone probes: a, ClCO_2Et , TEA, Py, DCM, -5°C - rt, overnight; b, TFA, DCM, rt, 1 h; c, **22**, PyBOP, TEA, DMSO, rt, overnight

Six natural amino acids were chosen to probe binding specificity: *L*-tryptophan (**40**) and *L*-phenylalanine (**41**) for evaluating the influence of aromatic residues on the P2 position; *L*-serine (**43**) and *L*-glutamine (**39**) as hydrophilic, non-charged amino acid residues at P2 position; *L*-leucine (**42**) for testing the lipophilic character of the P2 position and *L*-proline (**38**) as a turn-inducing moiety. Furthermore, two modified amino acids were also included for comparison (Figure 8), i.e. a tryptophan derivative in which the indole moiety was modified with a formaldehyde group (**44**) and a *N*-methyl leucine derivative (**45**). With such slight modifications, the sensitivity of recognition for the P2 position could be very well distinguished.

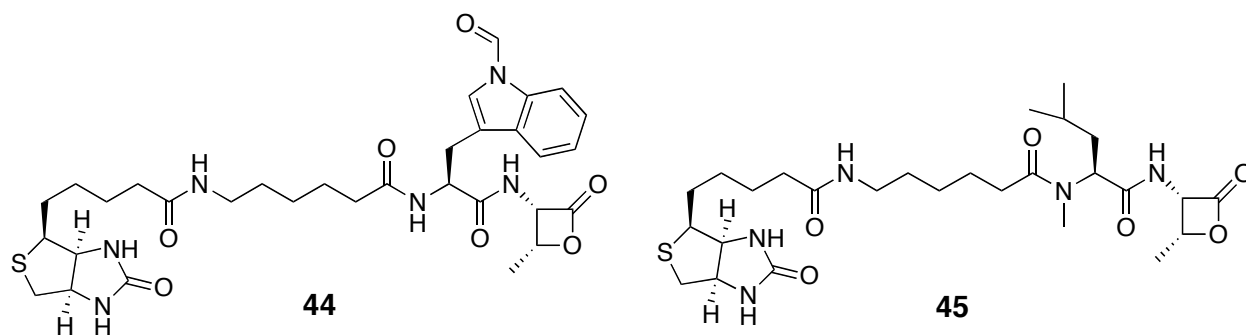
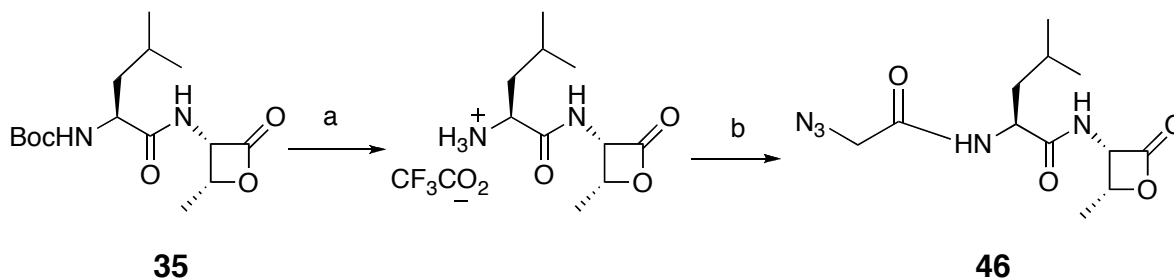


Figure 8. Structures of probes **44** and **45**

In order to obtain also a probe suitable for in vivo labelling, derivative **46** was also synthesized (Scheme 11). Thus, leucine β -lactone derivative **35** was deprotected with trifluoroacetic acid and coupled with 2-azido acetic acid to yield the probe **46**.



Scheme 11. Synthesis of click version probe **46**: a. TFA, DCM, rt, 1 h; b. $\text{N}_3\text{CH}_2\text{CO}_2\text{H}$, PyBOP, TEA, DCM, rt, overnight, 66%.

3.2.3 Bioassays

The probes were tested for labelling of *Arabidopsis* leaf proteomes of the first generation of probes (**39**, **44** and **45**). Only **39** caused labelling (Figure 9a). P2 positions of **44** and **45** have non-natural amino acids whereas **39** has a glutamine. This implied that modification of the natural amino acid prevents the probe for labelling. Probe **39** caused two weak signals of 50 and 60 kDa, indicating that glutamine is not the most suitable amino acid for the P2 position. The second-generation probes (**38**, **40**, **41**, **42** and **43**) carry proline, tryptophan, phenylalanine, leucine and serine. Probes **40**, **41** and **42** caused strong labelling whereas probes **38** and **43** did not cause labelling (Figure 9b). In case for the P2 position of probes with tryptophan and leucine, when they were natural amino acid, the probes (**40** and **42**) caused labelling while when they are chemical modified, the probes (**44** and **45**) did not cause labelling. This indicates the importance of the P2 position of the probes for labelling. However probes **41-42** caused almost identical labelling patterns, implying that tryptophan, phenylalanine and leucine do not influence the specificity of the probes. Probe **41** was chosen for further studies.

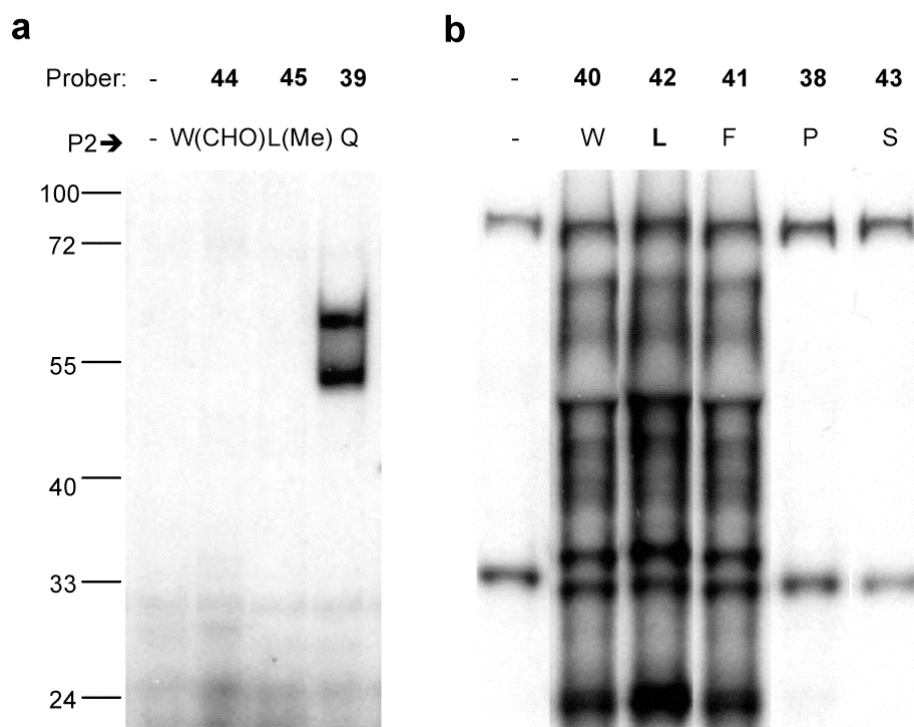


Figure 9. Labelling *Arabidopsis* leaf extracts with a series of threonine based β -lactone probes: a. Profiling with a first generation probes containing chemical modified tryptophan (**44**) or leucine (**45**) and natural glutamine (**39**) at P2 position; b. Profiling with second generation probes containing tryptophan (**40**), leucine (**42**), phenylalanine (**41**), proline (**38**) or serine (**43**) at P2 position; Protein extracts (~5 mg/ml (a) and 0.15 mg/ml (b)) were incubated at pH 8 with 2 μ M probes and 1 mM DTT for 3 h. The proteins on the protein blots were detected by streptavidin-HRP, 10 minutes long exposure (a) whereas 5 seconds short exposure (b).

Since β -lactone molecules can be inhibitors of cysteine proteases (Lall *et al.*, 1999), we compared labelling profiles of **41** with DCG-04, which labels papain-like cysteine protease (Figure 10). These assays were done in Caspase buffer at pH 7.4 and TBS buffer pH 7.5, and these buffers do not cause different labelling patterns. Probe **41** and DCG-04 have significantly different labelling profiles, indicating that they do not label the same targets. Labelling by probe **41** was competed by its non-biotinylated derivate **33** but not by **35**, the non-biotinylated derivate of probe **43**, which had no labelling reactivity. Surprisingly, labelling by probe **41** can also be competed by E-64, the non-biotinylated analogue of DCG-04. One hypothesis is

that E-64 has a broader target range than DCG-04, including the targets of probe **41**. A direct way to prove this hypothesis was to identify the targets of probe **41**.

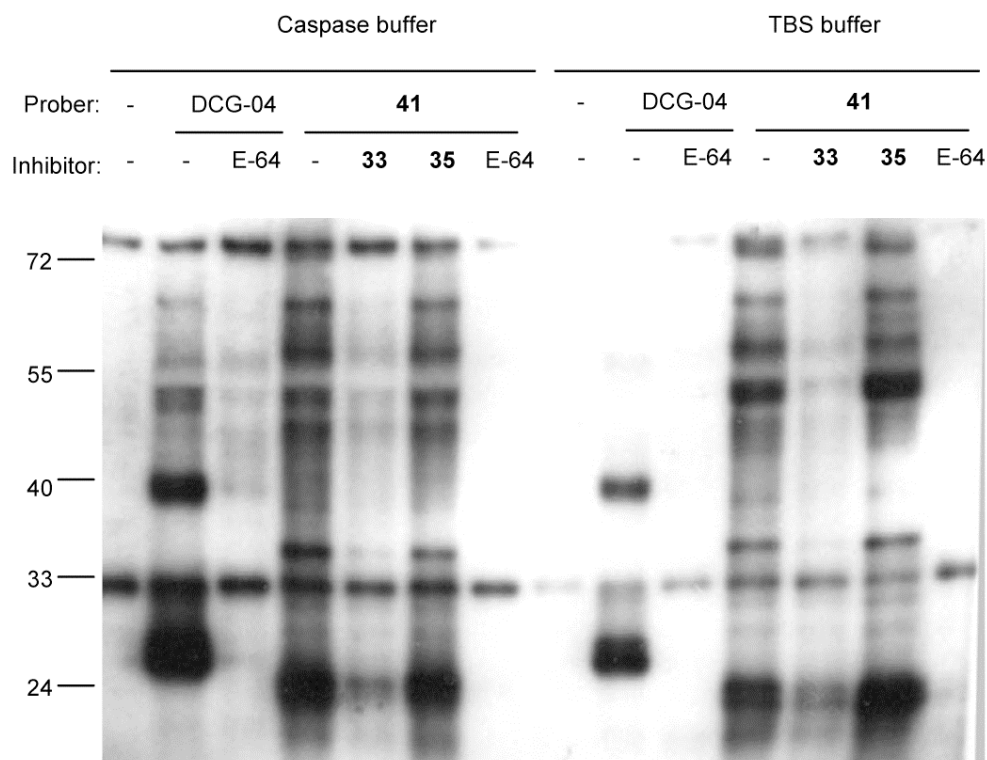


Figure 10. Comparison of labelling between DCG-04 and **41**: DCG-04 and **41** were profiled for a labelling assay as well as a competition assay with their own inhibitors in *Arabidopsis* leaf extracts with two buffer systems (Caspase pH=7.5, TBS pH=7.4) respectively. Protein extracts (~50 µg/ml) were co-incubated at pH 7.5 with 2 µM probe (DCG-04 and **41**) and 60 µM inhibitors (E-64, **33** and **35**) and 1 mM DTT for 3 h. The proteins on the protein blots were detected by streptavidin-HRP.

The purification of **41**-labelled proteins was performed by Christian Gu (Figure 11a) and protein mass spectrometry analysis was done by Dr. Tom Colby. The strongest band of 24 kDa was identified as the 23 kDa PsbP protein, which is a non-proteolytic protein but involving in photosynthesis. The MS data showed that **41** was bound to the N-terminus of PsbP (Figure 11d).

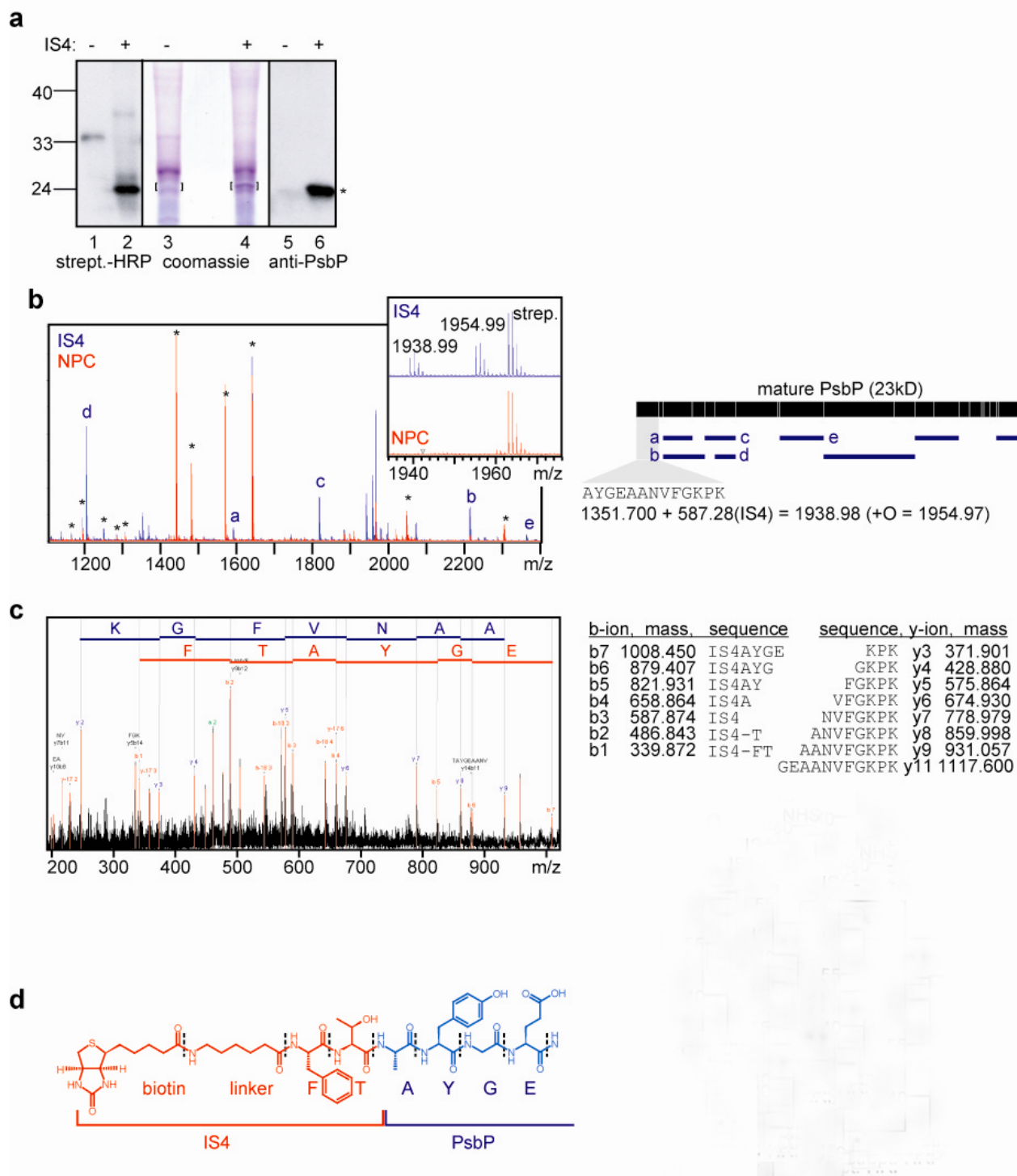


Figure 11. Identification of the major IS4 (41)-labelled protein: a. Purification of IS4 (41)-labelled proteins from Arabidopsis leaf extracts. Arabidopsis leaf extracts were labelled with and without IS4 (41) and biotinylated proteins were captured and purified on magnetic streptavidin beads. Proteins eluted from these beads were analyzed on protein blot, probed with streptavidin-HRP (left), and on coomassie-stained protein gel (middle). The differential protein band at 23 kDa was isolated (brackets), analyzed by tandem mass spectrometry and confirmed as being PsbP using specific PsbP antisera on the purified proteins (right); b. Peptide mass fingerprint (PMF) of

the 23 kDa protein band. Proteins isolated from the 23 kDa region of the IS4 (**41**)-labelled proteins (blue) and no-probe control (NPC, red) were analyzed by MS. Only part of the PMF is shown. Peptides from bovine serum albumine (BSA, *), streptavidin and trypsin were present in both IS4 and the NPC sample. The remaining IS4 (**41**)-specific peptides covered most of the 23 kDa mature PsbP protein (right). Peptides a-e are indicated in the PMF, the other matching peptides were outside the shown PMF. The section of the PMF with the IS4 (**41**)-modified N-terminal peptide is shown in the inset and explained on the right. Both the IS4 (**41**)-labeled peptide and its oxidized form have predicted masses that match the masses in the PMF inset; c. Fragmentation data of the IS4-labeled N-terminal peptide. The predicted y-ions (right) are found in the spectrum at the expected masses (left). The b-ions are also found in the spectrum, with an additional mass of IS4. Also IS4 itself, and fragments of IS4 (**41**) are found in the spectrum (b1, b2 and b3 ions); d. Proposed structure of the N-terminus of IS4-labeled PsbP, based on the peptide fragmentation data. IS4 is linked via a normal threonine through a peptide bond to the N-terminal Ala of PsbP.

To further study **41**-labelling, a series of assays were performed (Figure 12). Labelling occurred in mild basic conditions (Figure 12a), required reducing agent DTT (Figure 12b), completed rapidly within half an hour (Figure 12c) and increasing concentration of proteome resulted in enhanced labelling, indicating that the probe is over excess compared to its targets (Figure 12d).

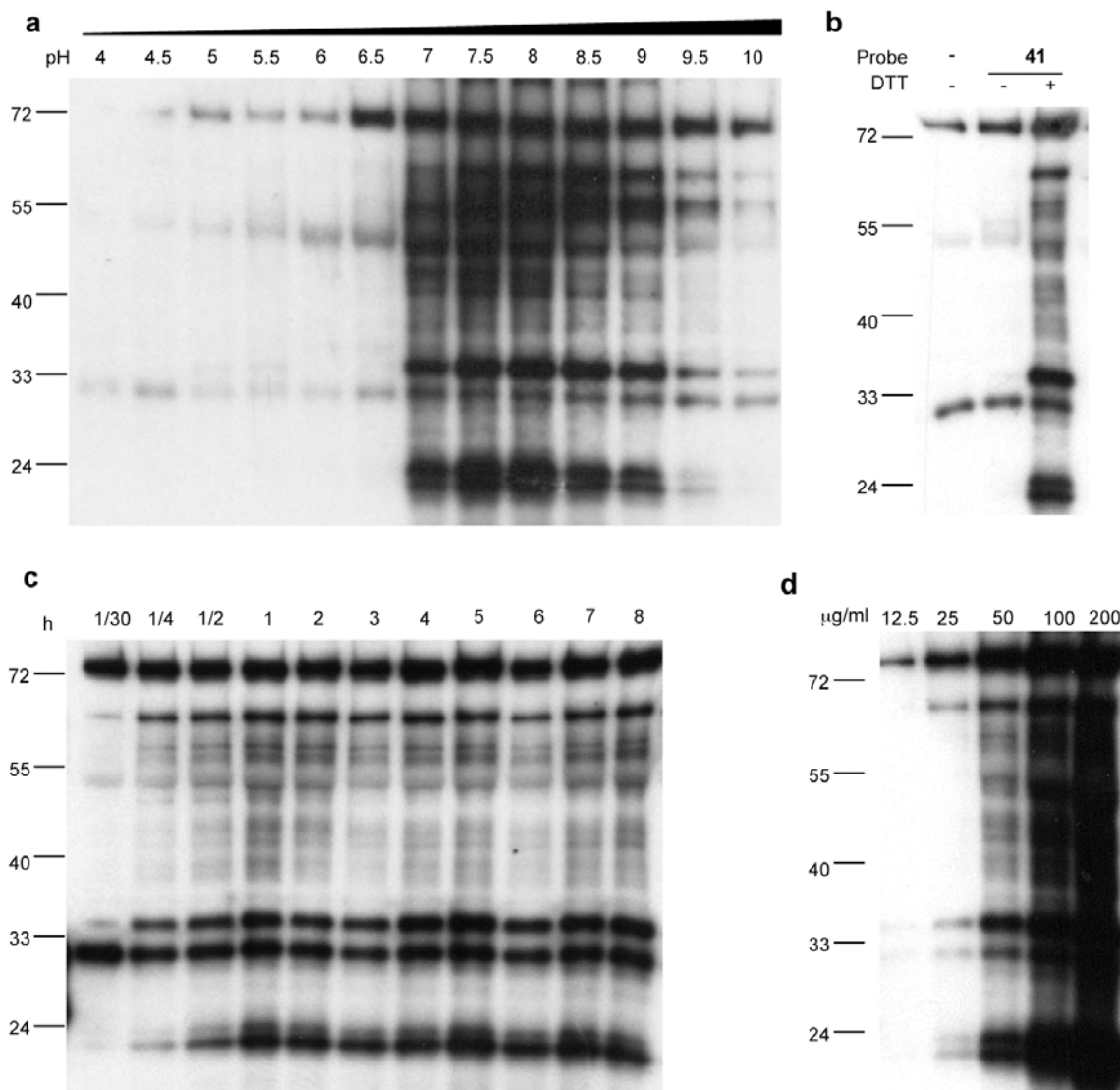


Figure 12. Labelling properties of probe **41**: a. pH curve; b. Cofactor DTT dependent labelling, DTT = 1 mM; c. Time course of **41** labelling; d. Protein concentration dependent labelling; Protein extracts (~50 µg/ml) (except for d) were incubated at pH 8 (except for a) with 2 µM probe **41** and 1 mM DTT (except for b) for 3 h (except for c). The proteins on the protein blots were detected by streptavidin-HRP.

Labelling by probe **41** can be prevented by adding E-64 (Figure 10). We next determined if E-64 derivatives and other proteases can also prevent **41**-labelling. Labelling by probe **41** could be competed by its non-biotinylated derivate **33** and all E-64 derivatives (E-64, E-64c and E-64d), whereas DCG-04 labelling could not be competed by **33**. When DCG-04 was mixed with probe **41**, DCG-04-signals were of the same labelling intensity but the intensity of **41**-signals were reduced.

The mixed probe labelling could be competed by E-64 (Figure 13a). All the cysteine protease inhibitors could compete with **41**-labelling whereas serine protease inhibitors such as PMSF, TLCK and TPCK could not compete (Figure 13b). This indicated that the labelling of **41** requires cysteine proteases.

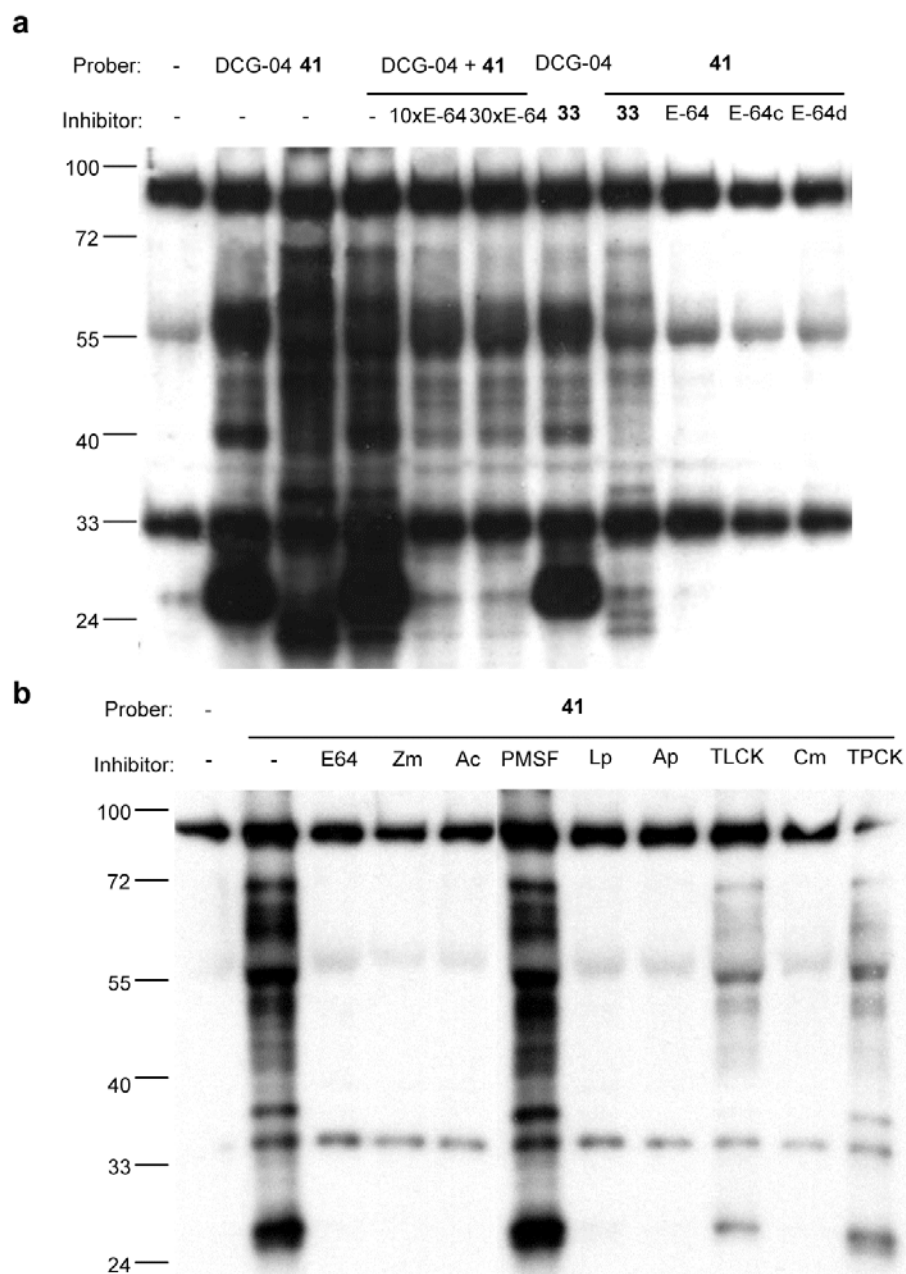


Figure 13. Competition assay of **41**: a. Investigation the labelling relationship between DCG-04 and **41**: DCG-04 vs **41**'s competitor **33**, **41** vs DCG-04's competitors (E-64, E-64c and E-64d) and the mixture (DCG-04 + **41**) vs E-64 in two doses (10 times- and 30 times-equivalent); b. Competition of **41** with other known serine or cysteine protease inhibitors: ZFAfmk (Zm), AcLvKCHO (Ac), leupeptin (Lp), Antipain (Ap), Chymostatin (Cm) as well as E64, PMSF,

TLCK and TPCK. Protein extracts (~50 µg/ml) were co-incubated at pH 8 with 2 µM probe (DCG-04 and **41**) and 30 µM of inhibitors (except for two marked E64 competition assays in a) and 1 mM DTT for 3 h. The proteins on the protein blots were detected by streptavidin-HRP.

If **41** had multiple targets, it probably would have different affinities to different targets. To monitor different affinities, several competitors were tested in a concentration series to compete with **41** labelling. The intensity of the strongest five signals were measured and plotted against the competitor concentration (Figure 14). These experiments showed that leupeptin was the strongest competitor (Figure 14c) and **33** was the weakest competitor (Figure 14d); E-64d was less potent than E-64 as a competitor (Figure 14a-b). However the intensities of all signals decreased similarly, indicating that **41** had almost the same affinity to all targets.

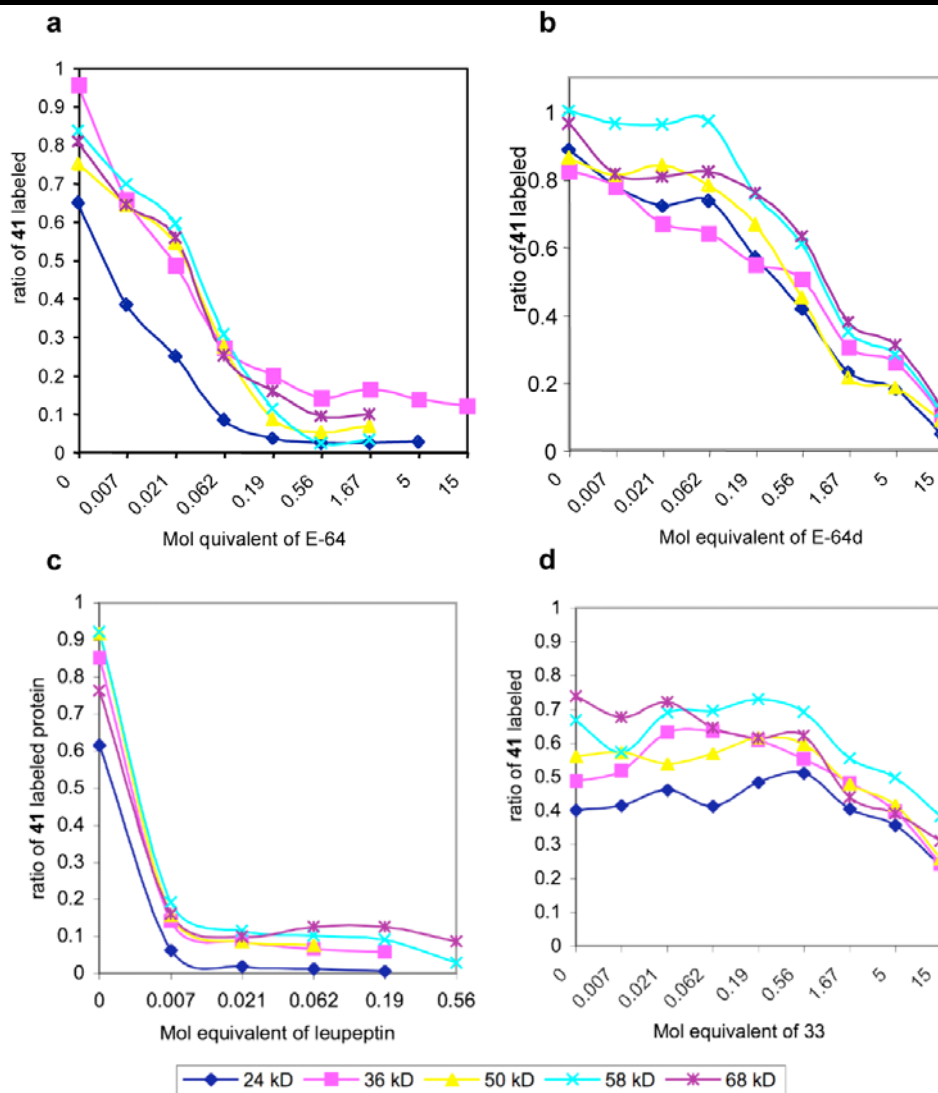


Figure 14. Monitoring the intensity variation of labelling bands of **41** with different competitors: E-64 (a), E-64d (b), leupeptin (c) and **33** (d) were diluted in a series of increased concentration to compete with **41**; The labelling intensity of the five strongest signals of **41** was plotted against the concentration of competitor.

Many evidences implied the labelling process of **41** is required for cysteine proteases, however the final labelled protein is a non-catalytic protein. **41**-labelling can be interfered by DCG-04 (Figure 13a) implied that the target of **41** will be one of targets of DCG-04. The *Arabidopsis* genome encodes for 30 PLCPs, of which at least ten are expressed in leaves (www.genevestigator.ethz.ch), and six were identified by DCG-04 in leaf extracts (Van der Hoorn *et al.* 2004). We reasoned that one of these leaf PLCPs could be responsible for IS4 labelling. We therefore

generated knock-out lines by selecting lines carrying a T-DNA insertion in the genes encoding leaf-expressed PLCPs. IS4 labelling occurs in leaf extracts of all mutants, except of the rd21-1 line. However by adding recombinant RD21 to proteomes of rd21-1 mutant plants can complement IS4 labelling again. This demonstrates that only RD21 is required for IS4 labelling in leaf extracts (Wang *et al.*, 2008). To understand the labelling mechanism, hypothesized that IS4 binds to RD21 and forms a thioester intermediate that can be transferred to the N-terminus of PsbP. Since a thioester bond is common to all intermediates of PLCPs with their substrates, we tested if thioesters formed from peptides could also be ligated to other proteins by RD21. To prove this hypothesis, we made a biotinylated peptide PepA (Bio-FTAYGE), which causes a labelling profile that is very similar to that of IS4. The unnormal labelling phenomenon finally revealed a new labelling mechanism of ABPP, proving mechanism of transligation (Figure 15).

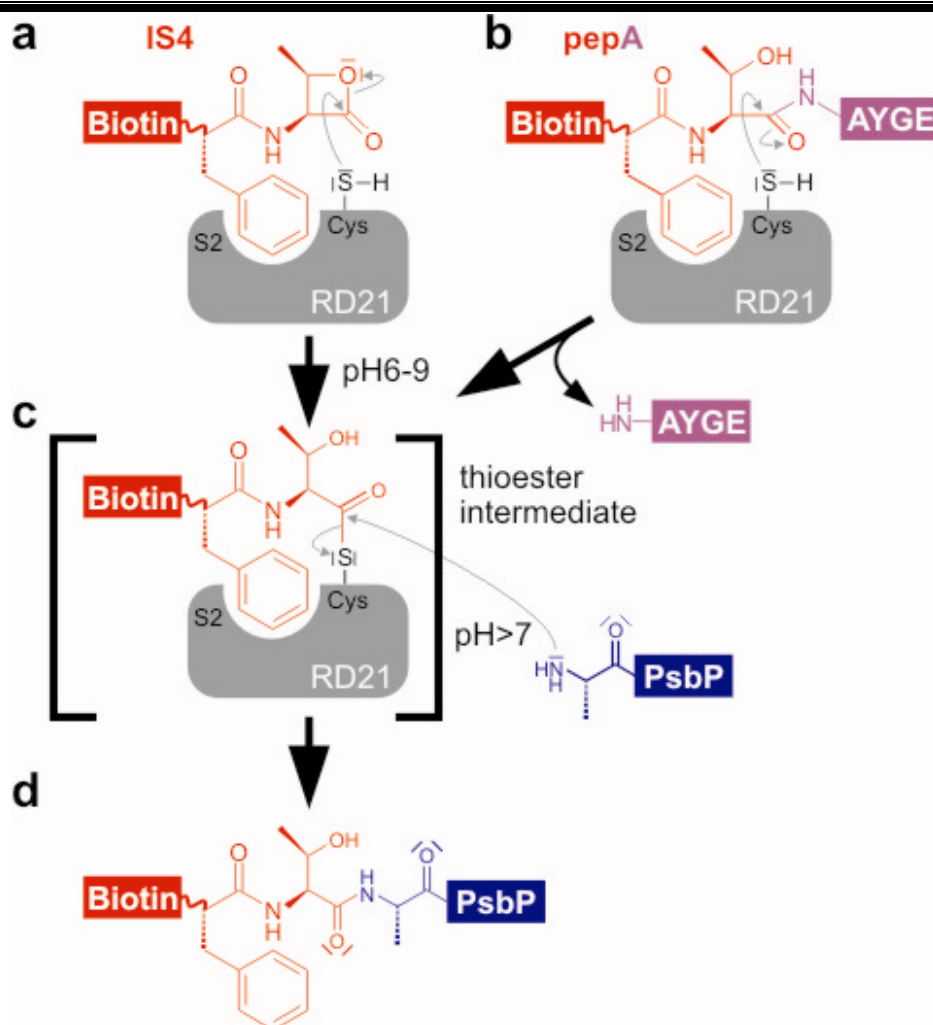


Figure 15. Model for IS4 and pepA labelling of PsbP by RD21. IS4 (**41**) (a) and pepA (b) bind to RD21. The Phe of these acceptor molecules is at the P2 position, making contact with the S2 substrate binding pocket of RD21. The active site Cys of RD21 acts as a nucleophile, resulting in a unstable thioester intermediate (c). At neutral to basic pH, the N-terminal amino group of PsbP acts as a nucleophile, resulting in IS4 (**41**) labelling of the N-terminus of PsbP through a peptide bond (d).

3.2.4 Discussion

Although we designed β -lactone derivatives as activity-based probes, the fate of these small molecules in plant extracts appears more complex than we predicted. IS4 labelling does not depend on the activity of the targeted proteins but is the result of an indirect labelling through a presumed protease. This illustrates that further investigation of unexpected labelling sites can lead to intriguing molecular

mechanisms. The discovery of the mechanism of transligation makes a complementary contribution to the ABPP proposal. Usually when we use designed ABPs to perform ABPP, we expect to find the targets in the gel. However this story tells that what is visualized does not need to be indeed the targets of the probe. The idea of mechanism-based probe design is reliable, however the enzyme-mediate labelling should also be considered. In this case, from a chemical aspect, it is known that the instability of the newly formed cysteine protease-probe conjugate, that can be imagined from the known labelling mechanism of cysteine proteases will use the thiol group of cysteine to attack the β -lactone ring and form a thioester bond. However we could not predict whether this instable thioester bond could be tolerant to the ABPP condition. The final result shows us that the instable thioester bond indeed was not tolerant to the labelling condition. However this instability of the thioester bond revealed an unusual protease labelling mechanism. Whether these reactions also occur in living cells will be as an exciting topic for further studies.

3.3 Non-directed ABPs design of aziridines and azirines

3.3.1 Introduction

The non-directed probe design strategy was introduced by Cravatt in 2002, employing a sulfonate ester moiety as a reactive warhead (Adam *et al.*, 2002). The non-directed strategy is based on the incorporation of a reactive functional group that does not represent a classical irreversible enzyme inhibitor as a warhead and to accommodate it with a linker and a reporter tag. Clearly, the reactivity of the warhead is critical for this approach as a too low reactivity will result in a weak or even no labelling while a too high reactivity will lead to false positive results resulting from unspecific labelling. As noted by Evans *et al.* in a review on this strategy, the “application of these probes significantly expanded the scope of enzymes addressable by ABPP, further underscoring the utility of combinatorial strategies for probe discovery.” was mentioned in the review (Evans *et al.*, 2006).

To date, several reactive functional groups have already been used for probes design, such as epoxides, β -lactones, or halogeno methyl ketones. Further examples are the α -chloroacetamide (Barglow *et al.*, 2004) or reactive spiroepoxide (Evans *et al.*, 2005) based probes. We rationalized that aziridines and azirines, in which a nitrogen residue replaces the oxygen of epoxide, might represent good candidates for non-directed probes. Besides the structural similarity of aziridines and azirines with epoxides, these also feature a similar reactivity, being an electrophilic reactive group that can react in nucleophilic ring opening reactions.

The aziridine moiety can be found in various natural products such as azinomycin or miraziridin A (Corre *et al.*, 2004 and Schaschke *et al.*, 2004). Also azirines are elements of several bioactive natural products such as Azirinomycin

isolated from *Streptomyces Aureus* (Miller *et al.*, 1971) or dysidazirine isolated from the marine sponge *Dysidea Fragilis* (Molinski *et al.*, 1988) (Figure 16).

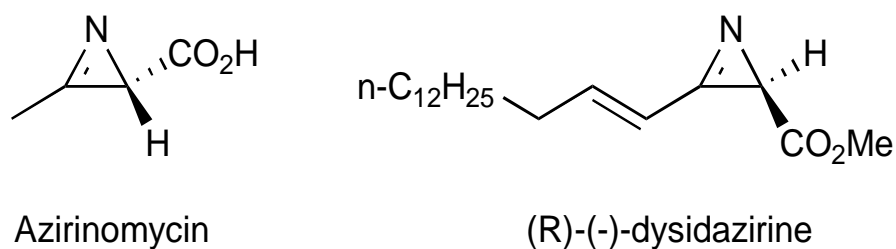


Figure 16. Structure of azirinomycin and R-(-)-dysidazirine

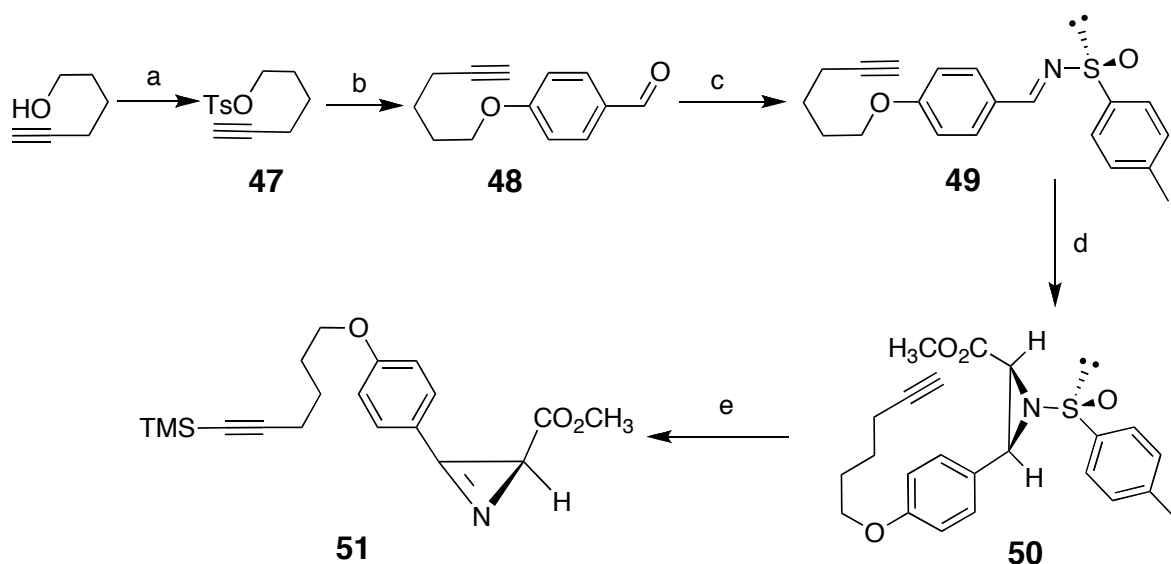
Their central feature is the chiral center on the C3 position of the cyclic ring, which is connected to a carboxylic acid moiety, either in a free acid form or as a methyl ester. Therefore, ABPs based on this natural product motif could employ the required tag moiety either on the C3 attached carboxyl group or to the C2 ring position.

3.3.2 Syntheses

The chiral 2H-azirines (Figure 16) are synthetically accessible from corresponding chiral aziridine precursors via elimination (Davis *et al.*, 1995). This intermediate can be prepared *via* an aza-Darzens asymmetric synthesis from *N*-sulfinyl imines. (Davis *et al.*, 1999) This synthetic route offers the advantage that it proceeds *via* an activated *N*-(*p*-toluenesulfinyl) aziridine intermediate which could also represent an interesting non-directed probe. In order to evaluate the impact of the attachment of the tag on labelling efficiency, two kinds of probes were initially synthesized.

The first probe **51** was prepared with the tag on the C-2 side chain (Scheme 12). To obtain this compound, the hydroxyl group of 5-hexyn-1-ol was converted into **47** with tosyl chloride in pyridine at 0 °C, followed by a substitution with 4-hydroxy-benzaldehyde in DMF at 100 °C to form **48**. **48** was then converted into

sulfinyl imine **49** with an *in situ* prepared lithium tolylsulfinyl trimethylsilyl amide in tetrahydrofuran. The sulfinyl imine can be isolated at room temperature and reacted with 2-bromo methyl acetate *via* an aza-Darzens reaction to generate the chiral aziridine **50** in good yields. Addition of lithium diisopropyl amide then induced elimination to the final product 2H-azirine **51**.

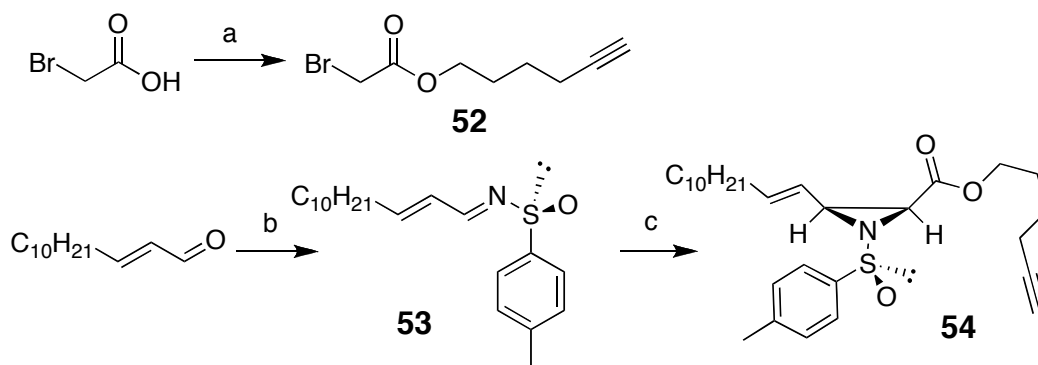


Scheme 12. Synthesis of **51**: a. TsCl, Py, 0 °C, overnight, 54%; b. 4-hydroxybenzaldehyde, Cs₂CO₃, DMF, 100 °C, 16 h, 87%; c. (1*R*, 2*S*, 5*R*)-Menthyl-(*S*)-*p*-tolylsulfinate, THF, LiHMDS, -78 °C, 7 h, 37%; d. BrCH₂CO₂Me, THF, LiHMDS, -78 °C, 20 min, 70.6%; e. LDA, TMSCl, THF, -78 °C, 15 min, 48%.

With this strategy, 2H-azirine **51** could be conveniently prepared in a short reaction sequence with modest yields. Originally, we aimed at the synthesis of an acetylene derivative of **51** (Scheme 12). However, due to a side reaction during the last reaction step, the TMS-acetylene derivative **51** was obtained because an excess of TMSCl was used. As removal of the TMS group risked the degradation of the probe and as a TMS-acetylene residue in principle should be compatible with a “click chemistry” protocol, we decided to employ **51** as a first test probe.

As this side reaction is an inherent problem of our synthesis strategy, we stopped our in parallel-performed synthesis (Scheme 13). **54** was obtained by following essentially the synthetic protocol as described for **50**. In brief, the

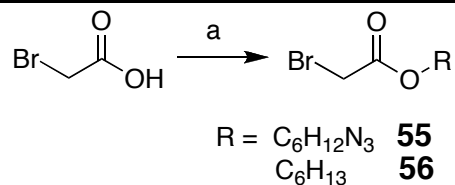
commercially available *trans*-2-tridecenal was converted into its sulfinyl imine **53** and the second building block 2-bromo 5-hexyne acetate **52** was prepared by esterification of 2-bromo acetic acid. Aza-Darzens reaction of **52** with **53** then led to the chiral aziridine intermediate **54**.



Scheme 13. Synthesis of **54**: a. HO(CH₂)₄CCH, DCC, DMAP, DCM, rt, overnight, 56%; b. (1*R*, 2*S*, 5*R*)-menthyl-(*S*)-*p*-tolylsulfinate, THF, LiHMDS, -78 °C, 7 h, 85%; c. **52**, THF, LiHMDS, -78 °C, 20 min, 87%.

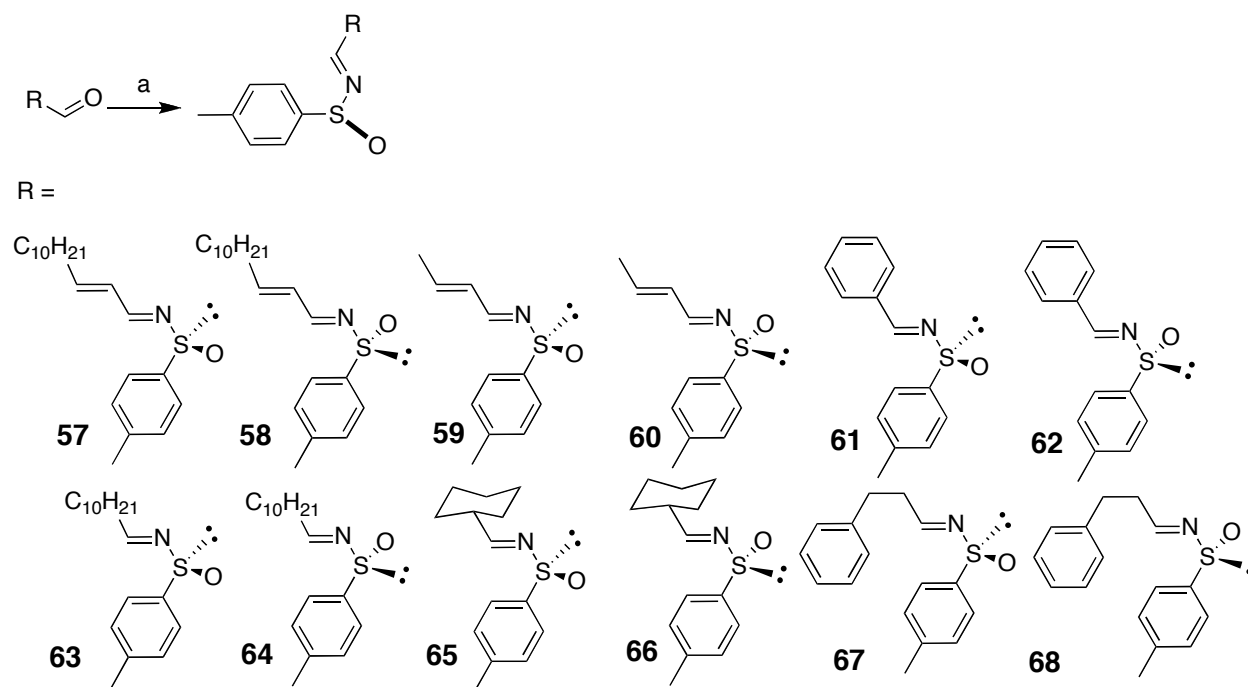
With these two probes in hand, some first labelling assays were performed. Unfortunately, no labelling with **51** could be detected. **54** however showed promising labelling characteristics (Figure 17). The efficiency of labelling with **54** demonstrated that the attachment of a tag to an ester group is compatible with labelling condition; secondly, the *N*-(*p*-toluenesulfinyl) aziridine is a suitable reactive warhead for labelling. These findings forced us to investigate these probe types more deeply. To prevent however any future problems with the acetylene moiety, the azide functionality was introduced as a tag for click chemistry.

As both (1*R*, 2*S*, 5*R*)-(-)-menthyl-(*S*)-*p*-toluenesulfinyl and its enantiomer (1*S*, 2*R*, 5*S*)-(+)-menthyl-(*R*)-*p*-toluenesulfinyl are commercially available, the syntheses of the next generation of probes was established in the following manner. First, the 2-bromo 6-azido-hexyl acetate **55** and 2-bromo hexyl acetate **56** were prepared once in a large amount (Scheme 14). Starting from 2-bromo acetic acid, esterifications with two different alcohols were performed in good yields.



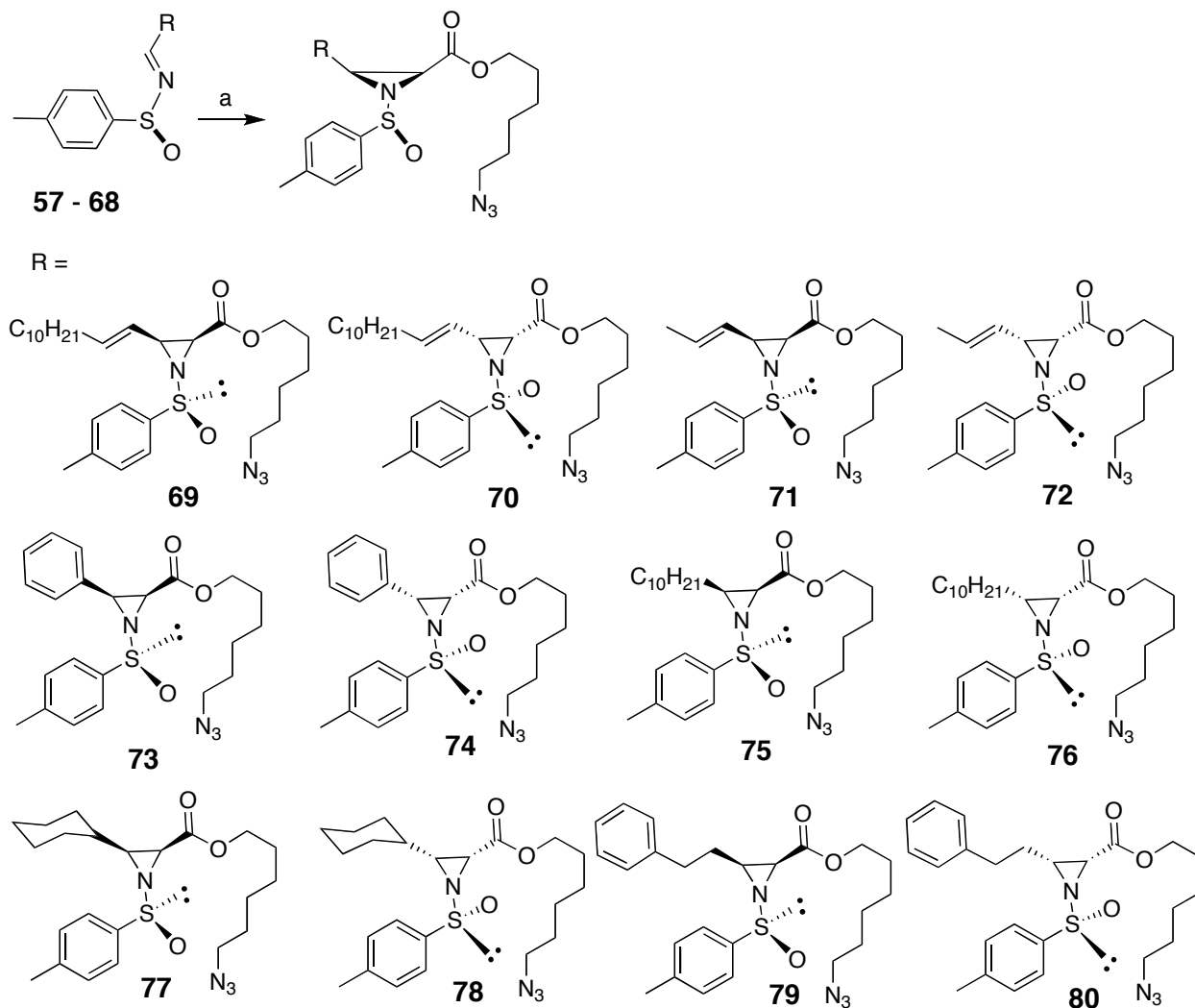
Scheme 14. Synthesis of **55** and **56**: a. alcohol, DCC, DMAP, DCM, rt, overnight, 53%.

To introduce diversity into the probes, the C2 *N*-sulfinyl imine reaction intermediates were prepared from different aldehydes with both chiral menthyl-*p*-toluenesulfonates (Scheme 15). The employed aldehydes were chemically different: The derivatives **57** and **58** were synthesized to incorporate the side chain of **54**, which led to successful labelling in the first generation probes. A shorter aliphatic version of **57** and **58** was also prepared (**59** and **60**). In addition, derivatives with aromatic side chains were synthesized (**61**, **62**, **67** and **68**). To investigate the role of the double bond on the side chain, compounds **63** and **64** were made.



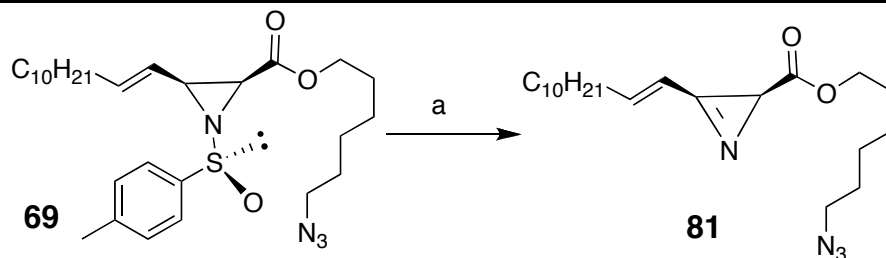
Scheme 15. Synthesis of the collection of *N*-sulfinyl imines: a. (1*R*, 2*S*, 5*R*)-(-)-menthyl-(*S*)/(1*S*, 2*R*, 5*S*)-(+)-menthyl-(*R*)-*p*-tolylsulfonate, THF, LiHMDS, -78 °C, 7 h.

The subsequent aza-Darzens reaction between these *N*-sulfinyl imines **57-68** and 2-bromo acetic acid esters **55** and **56**, then allowed the generation of a collection of *N*-(*p*-toluenesulfinyl) aziridines were obtained (Scheme 16).



Scheme 16. Synthesis of the collection of *N*-(*p*-toluenesulfinyl) aziridine probes: a. **55**, THF, LiHMDS, -78 °C, 20 min.

Additionally, to follow up the initial hit with the azirine derivative **54**, the 2H-azirine **81** this time carrying an azide instead of alkyne click tag was prepared from **69** with the same protocol as for **51** (Scheme 17).



Scheme 17. Synthesis of **81**: a. LDA, TMSCl, THF, -78 °C, 15 min, 20%.

3.3.3 Bioassays

Probes **51** and **54** were tested immediately after preparation. The labelling procedure was performed as described in Chapter 3.1 and 3.2, complemented by an additional “click chemistry” step (Kaschani *et al.*, 2009, see experimental part for the reaction condition). A representative example is shown in Figure 17.

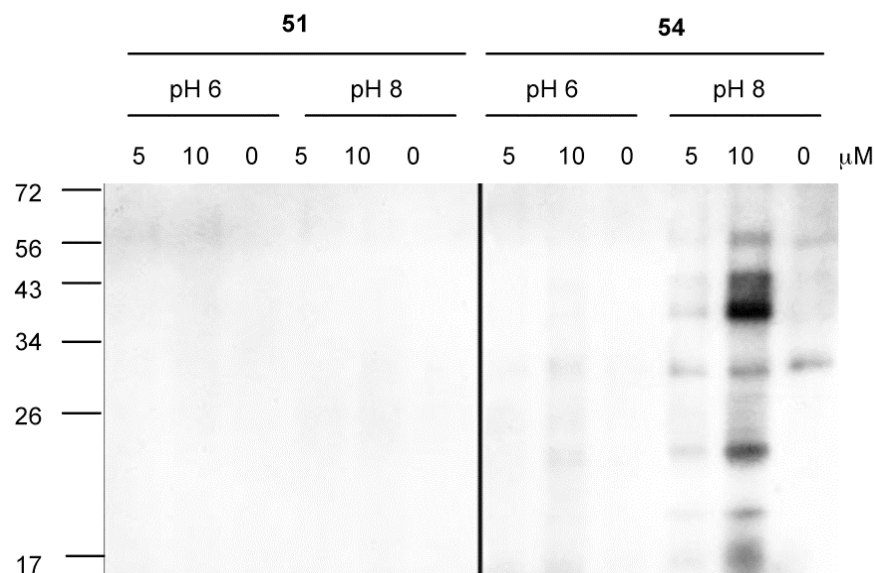


Figure 17. Labelling assays of probes **51** and **54**: Protein extracts (~50 μg/ml) were incubated with probe of 5 and 10 μM, in pH 6 and 8 buffer solution for 3 h and one step more of a “click chemistry” to add a biotin tag on the probe. The proteins on the protein blots were detected by streptavidin-HRP.

Unfortunately, the 2H-azirine probes **51** and **81** did not cause any labelling pattern neither at low pH or high pH. However, the *N*-(*p*-toluenesulfonyl) aziridine probe **54** caused a strong labelling band around 40 kDa at pH 8. Nevertheless, the

required probe labelling concentration of 10 μM was quite high and the experiment was hampered by a difficult reproducibility of labelling.

Profiling experiments with the collection of *N*-(*p*-toluenesulfinyl) aziridines **69-80** however resulted in a strong labelling pattern as depicted in Figure 18.

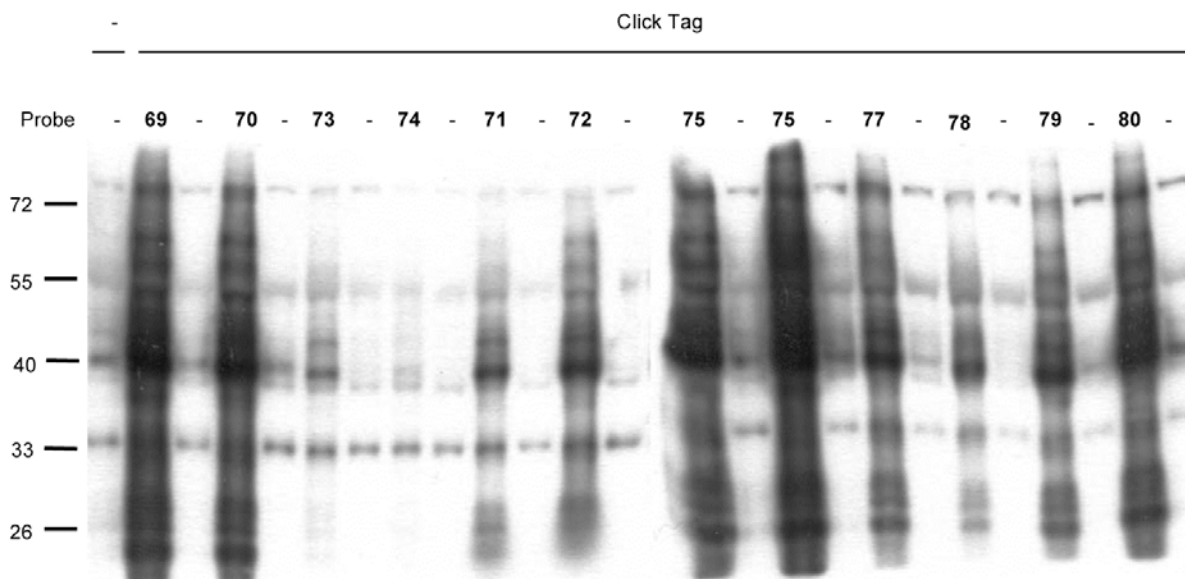


Figure 18. Labelling assays of *N*-(*p*-toluenesulfinyl) aziridine probes series: Protein extracts (~ 50 $\mu\text{g}/\text{ml}$) were incubated at pH 8 with 5 μM probe for 2 h and one step more of a “click chemistry” to add a biotin tag on the probe. The proteins on the protein blots were detected by streptavidin-HRP.

The different side chains of the probes had a distinct influence on labelling: **69** and **70** with long unsaturated aliphatic chains, **75** and **76** with long saturated aliphatic side chains, **79** and **80** with long aromatic side chains generally caused stronger labelling than the probes with short side chain such as unsaturated aliphatic, aromatic and cyclic **71 – 74** and **77 – 78**. Surprisingly, the different enantiomers did not cause different labelling patterns and although the probes featured different side chains and thus were chemically diverse, all labelling patterns seemed after visual inspection identical. Finally, the resolution of the gels was in all cases rather low, resulting in smeared bands. Unfortunately, also different labeling conditions did not improve the profiling quality. Despite these disadvantages, a pull down assay for identification of the targets was performed

analogously to the one described in chapter 3.2, but failed to deliver any positive results. After these negative results, a further follow up of this project was stopped.

3.3.4 Discussion

The non-directed 2H-azirine and *N*-(*p*-toluenesulfinyl) aziridines probes were synthesized by a straightforward synthesis route. A probe collection was generated by variation of the substitution pattern and the chiral centre on the three-membered ring. However, the adjacent biological assays were unsuccessful.

Despite that 2H-azirine is a very reactive organic molecule, it exists as a stable compound in nature. It was therefore reasonable to assume that it could be used for labelling. However, no reproducible 2H-azirine based labelling could be detected. This unpleasant finding might result from the instability of the reactive ring system in the employed labelling conditions.

The synthetic intermediate *N*-(*p*-toluenesulfinyl) aziridine is a less reactive organic molecule and showed some promising labelling patterns. This indicates a mild chemical reactivity of the warhead which is a prerequisite for the design of proper non-directed probes. The variation of the backbones from *N*-(*p*-toluenesulfinyl) aziridine did not result in distinct labelling patterns. Thus, it seems that also this reactive group is much less reactive than 2H-azirines, it still reacts unspecific with the proteome, leading to the conclusion that also these type of aziridines are unsuitable for ABPP applications.

3.4 Direct ABP design of AEBSF-based probes

3.4.1 Introduction

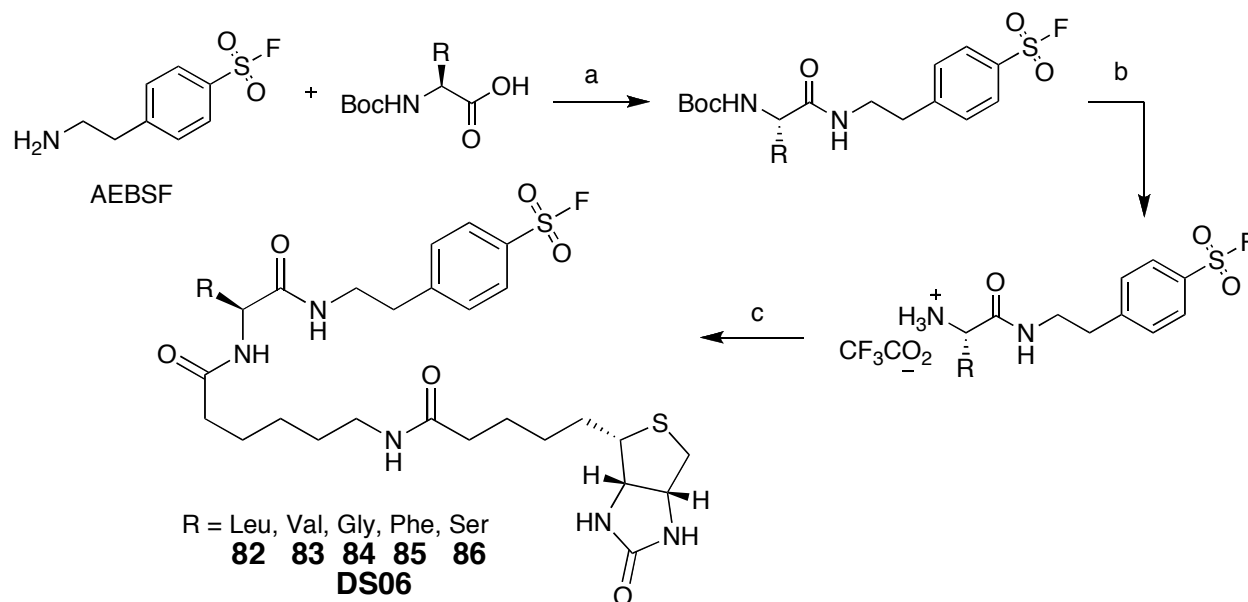
4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride (AEBSF) is a water soluble, irreversible serine protease inhibitor. Its specificity is similar to another serine protease inhibitor PMSF, nevertheless with higher stability at low pH values. It was discovered and applied as a pharmacological agent to inhibit chymotrypsin, kallikrein, plasmin, thrombin, and trypsin proteases, and many studies about the physiological functions of AEBSF from neurology, immunology and medicine were reported (Rideout *et al.*, 2001; Nakabo *et al.*, 1996; Okada *et al.*, 2003).

Sulfonate ester (SE) probes as nondirected probes were introduced by Cravatt, Sorensen and co-workers several years ago (Adam *et al.*, 2001; Adam *et al.*, 2002 and Adam *et al.*, 2004). The mild chemically reactive group of SE could endow the carbon electrophile with an ability to target a broad spectrum of labelling sites in enzymes. SE probes proved to be a versatile chemotype for the creation of ABPs that target a broad range of enzyme classes.

Compared to sulfonate esters, benzenesulfonyl fluorides have a much higher reactivity, and this high reactivity could disturb its labelling selectivity, similar to that of the azirine probes (Chapter 3.3). However AEBSF has been used for pharmacological studies, so directed probes i.e. AEBSF-based probes are still worth to generate and to test for monitoring serine proteases activities in plant proteomes. In addition to several known serine protease probes, such as Fluorophosphonate (FP) probes (Kidd *et al.*, 2001), AEBSF-based probes may target new classes of serine proteases.

3.4.2 Syntheses

AEBSF is commercially available and was used as a starting material. The design of the probes introduced the diversity on the amine function of AEBSF. To this end, peptide couplings that were performed in a parallel manner were used (Scheme 18). In brief, AEBSF was coupled with diverse *N*-Boc protected amino acids using PyBOP as the coupling reagent and TEA as a base. Subsequent deprotection with trifluoro acetic acid and coupling with the biotinyl linker building block **22** led after HPLC purification to the final products (**82** – **86**).

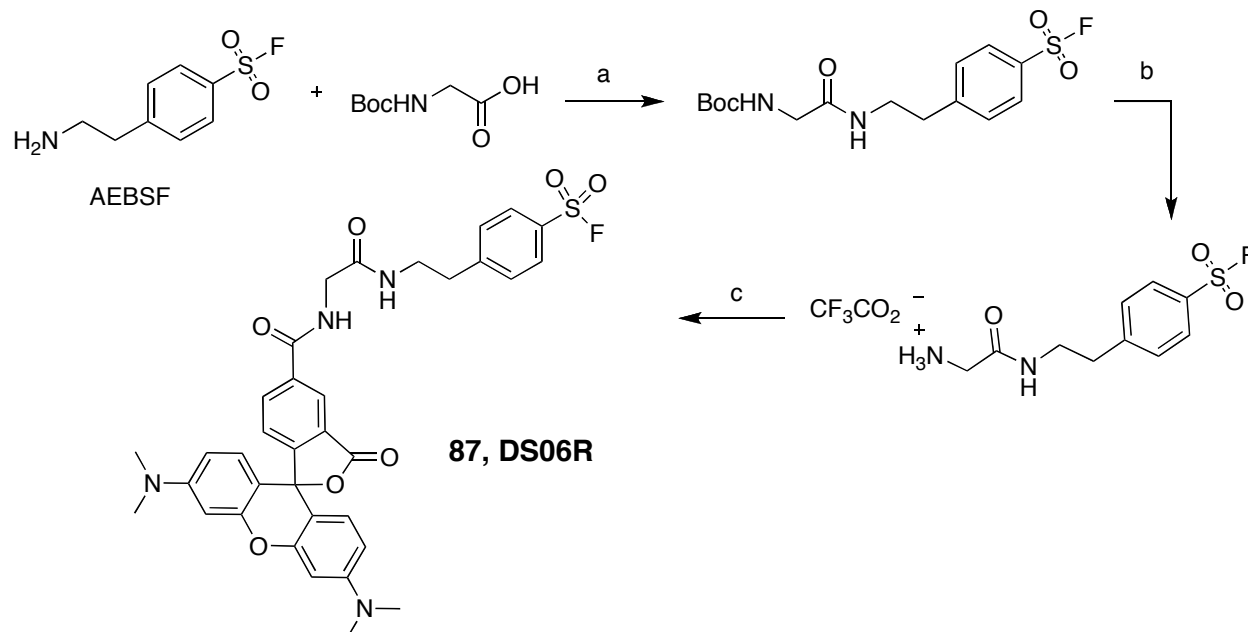


Scheme 18. Synthesis of a collection of sulfonyl fluoride probes: a. PyBOP, TEA, DCM, rt, 2 h; b. TFA, DCM, rt, 1 h; c. **22**, PyBOP, TEA, DMSO, rt, overnight.

The probes **82** to **86** were synthesized in a parallel reaction synthesizer without isolation of the reaction intermediates in a moderate yield of 30%. This might have resulted from the known instability of sulfonyl fluoride moiety, which could be improved by shortening the reaction time.

Subsequent biological evaluation of the synthesized probes (see next section) revealed **84** (later on named as DS06) as a representative example compound of all other probes. Therefore, a fluorescence derivative of DS06 (**84**) was synthesized,

resulting in the generation of probe **87** (later on named as DS06R, scheme 19). The synthetic route for DS06R (**87**) was similar to the biotinylated probes, but in the last step rhodamine *N*-succinimidyl ester was coupled instead of the biotinyl linker.



Scheme 19. Synthesis of the fluorescence probe **87**: a. PyBOP, TEA, DCM, rt, 2 h; b. TFA, DCM, rt, 1 h; c. 5(6)-Carboxytetramethylrhodamine *N*-succinimidyl ester, PyBOP, TEA, DMSO, rt, overnight, total 8.2%.

3.4.3 Bioassays

Arabidopsis leaf extracts were labelled with AEBSF-based-probes **82** to **87**. All probes caused strong labelling of multiple proteins (Figure 19). There was no difference in the labelling pattern between the different probes, which indicates that variation of the amino acid in the main chain of the probes did not affect labelling selectivity.

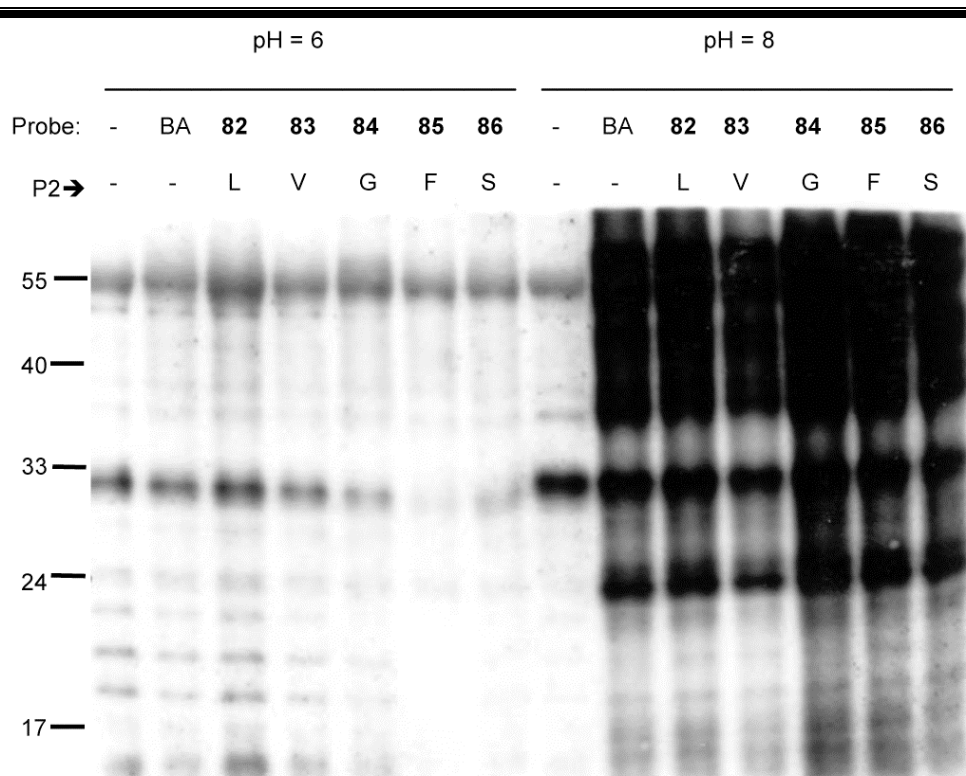


Figure 19. Labelling of *Arabidopsis* leaf extracts with sulfonamide probes: Biotinylated AEBSF (BA) and its amino acid conjugates (**82** – **86**) were incubated with *Arabidopsis* leaf extracts in two buffer systems (NaOAc pH 6 and Tris pH 8). The protein extracts (~0.2 mg/ml) were incubated at pH 6 (NaOAc) or pH 8 (Tris) with 20 μ M probe for 2 h. The proteins on the protein blots were detected by streptavidin-HRP.

A similar labelling profile was generated by labelling with the fluorescent probe **87** (Figure 20). Since this fluorescent probe represented the labelling profiles of all the other probes and since it is easy to handle, probe **87** was chosen for further studies.

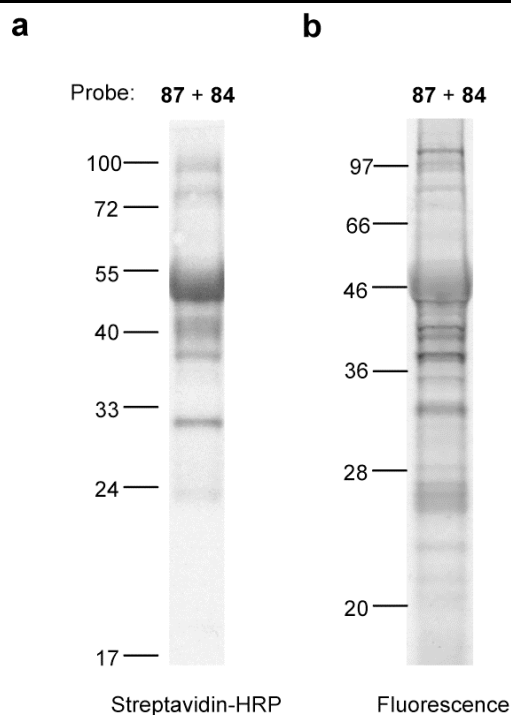


Figure 20. Co-labelling of **87** and **84**: equal concentration of fluorescent probe **87** and biotinylated probe **84** were mixed and incubated with an *Arabidopsis* leaf extract: The labelled proteins were detected by streptavidin-HRP (a) and by a fluorescence photometer (b). The protein extracts ($\sim 30 \mu\text{g/ml}$) were incubated at pH 8 with $2 \mu\text{M}$ probe for 2 h.

To further characterise labelling, probe concentration and pH value were varied (Figure 21). Low probe concentration ($\sim 60 \text{ nM}$) already caused labelling (Figure 21a) illustrated the labelling sensitivity of probe **87**. The signal intensities increased with increasing probe concentrations in the same ratio, which suggested that the probe **87** had almost the same affinity to all the targets. Furthermore, there was no saturation of the intensity of the signals, indicating that labelling cannot be saturated. Labelling at various pH revealed that the intensities of the signals increase with the rising pH value. This increased labelling is in contrast to the presumed probe stability since sulfonyl fluoride is less stable at higher pH.

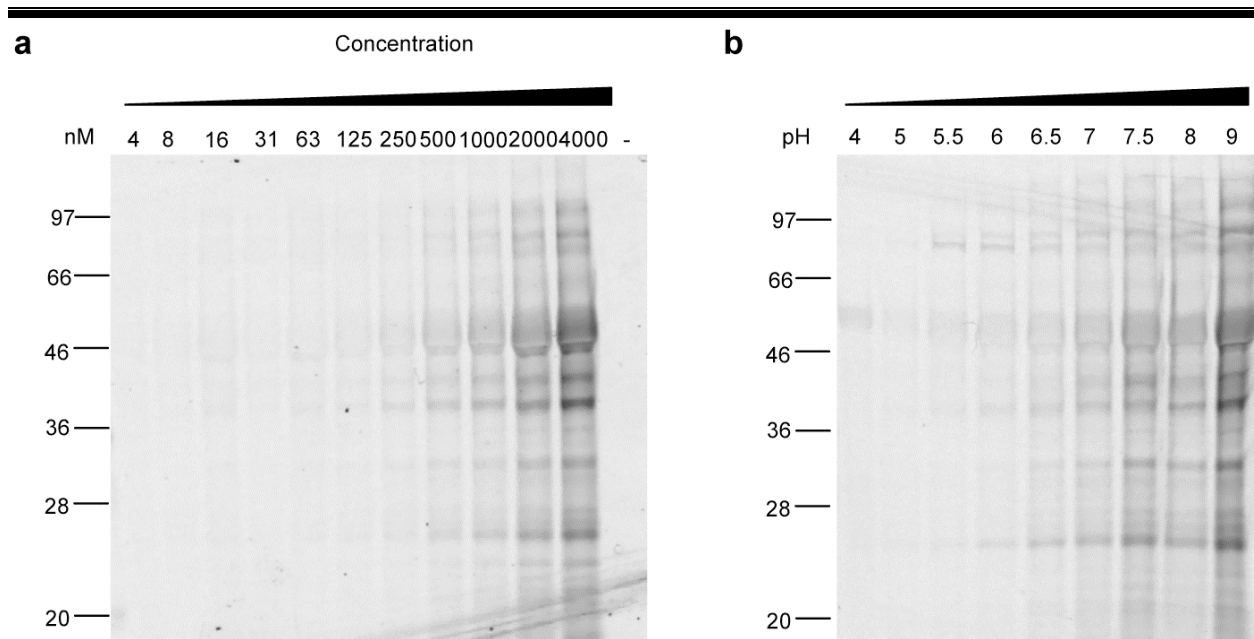
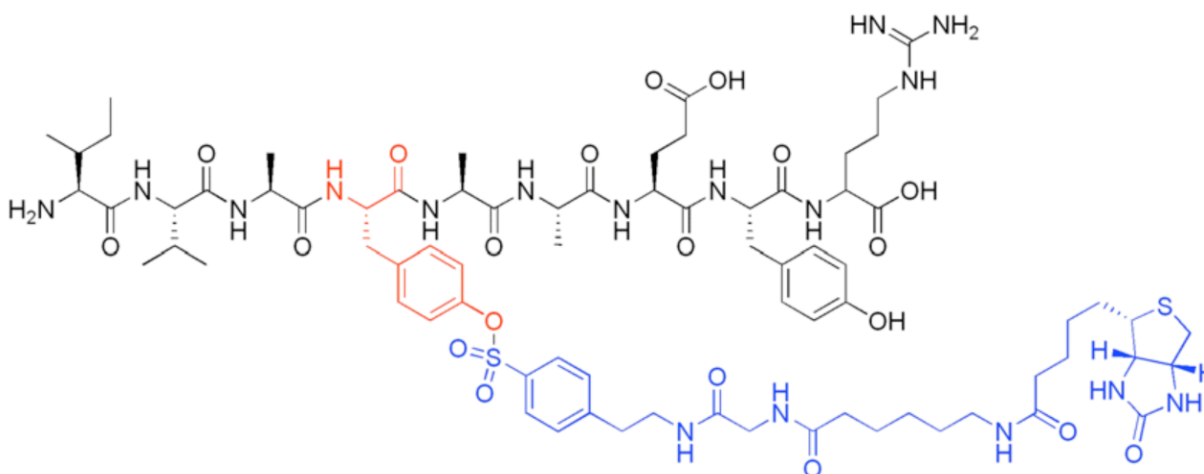


Figure 21. Labelling at various probe concentrations and various pH values: a. the probe concentration of **87** ranges from 4 nM to 4 μ M; b. the pH value ranges from 4 to 9; Protein extracts (\sim 20 μ g/ml) were incubated at pH 8 (a) with 2 μ M probe (b) for 2 h. The labelled proteins on protein blot were detected by fluorescence photometer.

Whether **87** is an activity-based probe was still not clear based on the above results. To understand the labelling mechanism, identification of the targets was required. Christian Gu performed a pull down assay with **84**-labelled proteins and Dr. Colby did the protein mass spectroscopy analysis. More than hundred proteins were identified but there were no serine proteases or other proteases. The identified proteins represent different classes providing no clue on the labelling mechanism of the probes (**84** and **87**). However one **84** labelled peptide from a glutathione S-transferase (*AtGSTU20*, class τ 20) was identified (Figure 22). The labelling position of probe **84** was not on the active centre residue but on a tyrosine, distant from the active site.

AtGSTU20 class Tau

MALDI-MS: 1634.7590 / Exact Mass:1633.7633

Figure 22. Mass spectrometry of **84**-labelled peptide: the sulfonyl group was bound with a phenol group from tyrosine side chain; the probe **84** (blue); the peptide from *AtGSTU20* (black) with the labelled tyrosine was (red) colour.

To illustrate the unusual labelling position, a structure-based model was required. There is no crystal structure of *Arabidopsis AtGSTU20* available, however the structure of its rice homologue *OsGSTU1* (Dixon *et al.*, 2003) with substrate – glutathione (GSH) was reported (Figure 23). GSH locates in the substrate-binding pocket of GST, the sulphur atom of GSH points to the adjacent tyrosine (Y-116), distant from the active site similar to the **84**-labelled tyrosine in *AtGSTU20*. If that GSH could be replaced by probe **84**, the tyrosine can attack the sulfonyl fluoride moiety of probe **84**. This illustration helps us to understand the possible mechanism of probe **84** labelling the peptide from *AtGSTU20* at the tyrosine residue.

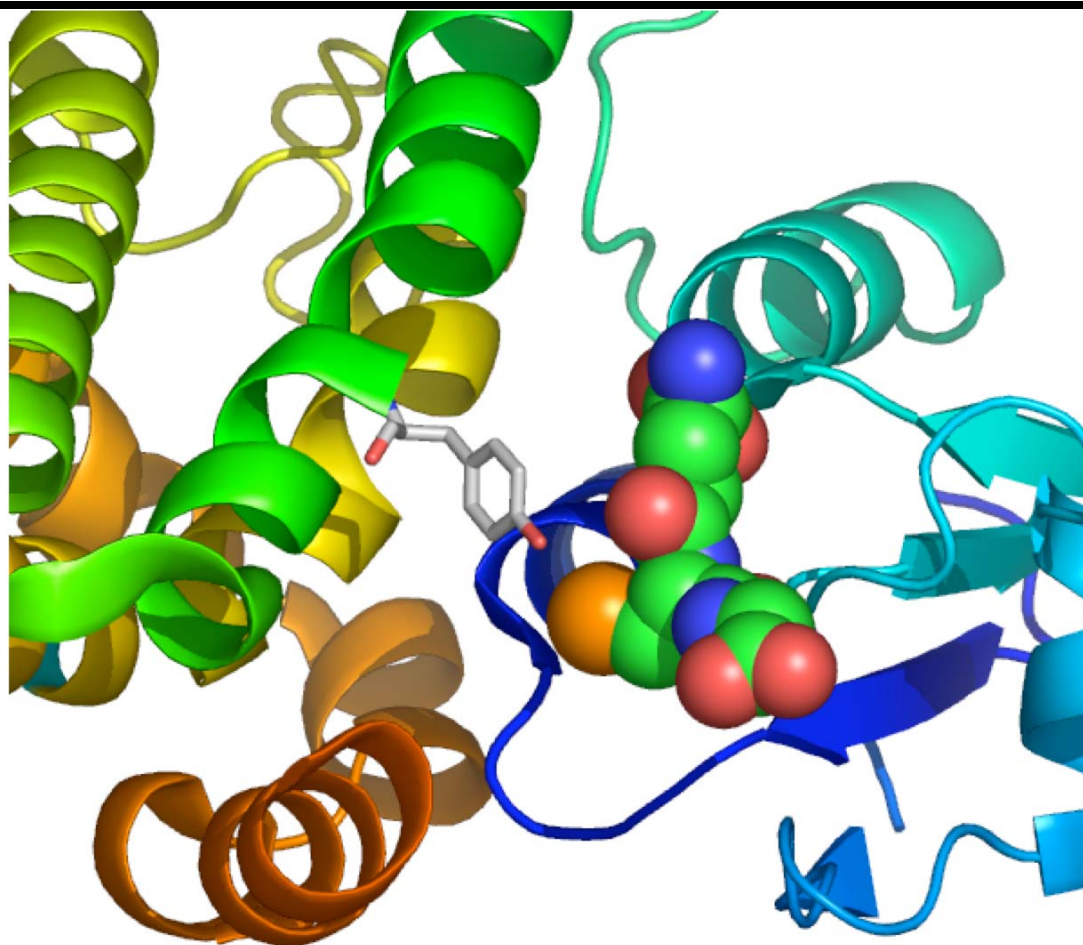


Figure 23. Crystal structure of rice GST1 (*OsGSTU1*) in complex with glutathione: In the C-domain GSH (space filling model) interacts with Y-116; the phenol group of Y-116 points to the cysteine side chain of GSH.

In the meanwhile, the group of Prof. Dr. Cravatt (Scripps Institute, La Jolla) also made the same observation. With the alkyne-labelled AEBSF probe (BC) (Figure 24), they identified also many unrelated proteins including GST from labelled mouse liver extracts.

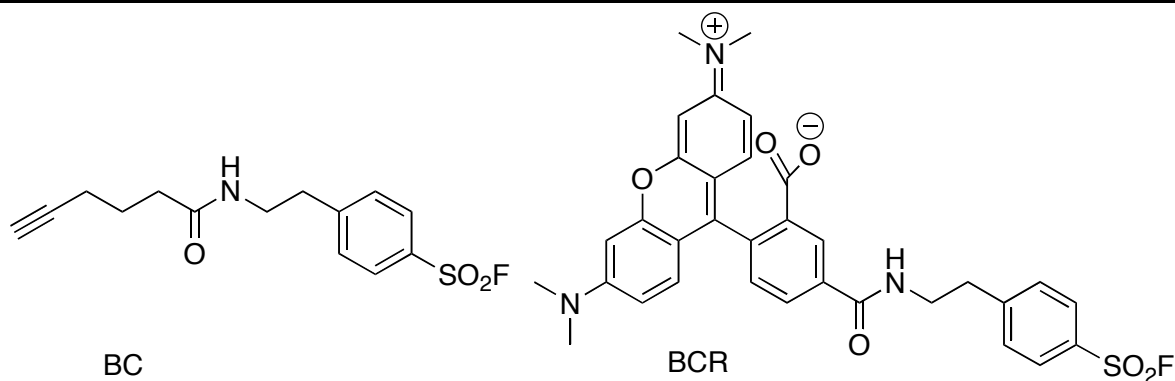


Figure 24. Structures of probes from the Cravatt lab: Click probe BC and fluorescence probe BCR

They identified 43 labelled peptides, 72% of those peptides were labelled on tyrosines and 28% were labelled on serines. Furthermore, it is found that probe BC labelled tyrosines at multiple positions of the same peptide indicating that the targets could be multi-labelled. Labelled proteins included three classes of GST (class μ 1, ψ 1 and ω 1) as well as other proteins. Labelling with the different probes on plant and animal proteomes are mostly in agreement, except for competition assays. The signals of BCR could not be competed by adding an excess of AEBSF (in mouse liver extract) whereas the signals of **87** were successfully competed by AEBSF (in *Arabidopsis* leaf extract) (Figure 25).

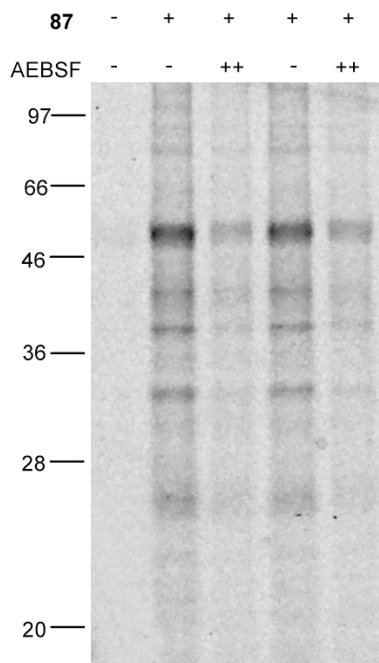


Figure 25. Competition assay between **87** and AEBSF: *Arabidopsis* protein extracts ($\sim 20 \mu\text{g/ml}$) were at pH 8 pre-incubated with $600 \mu\text{M}$ AEBSF for 30 min and then labelled with $2 \mu\text{M}$ probe **87** for 2 h. The labelled proteins on the protein blot were detected by a fluorescence photometer.

The opposite observations from two competitions can only be caused by two different factors. One is based on the biological difference between the two proteomes; the other is based on the chemical difference between BCR and **87**. As it unreasonable to assume that the proteome changes the labelling activity of probe, the only cause seems to be the difference between those two probes, and an assay for comparing **87** and BCR was hence designed (Figure 26).

The labelling pattern of **87** and BCR are different both in *Arabidopsis* leaf extract and mouse liver extract. The chemical structure difference between **87** and BCR is only one additional amino acid, however the side chain of that amino acid is not important, since probes **82** – **86** caused similar labelling patterns. Therefore the difference between **87** and BCR can only be explained by a spatial distance between warhead and reporter tag. The rhodamine moiety in **87** is further away from the warhead than in BCR. If the labelling pocket of the protein is too narrow, **87** would have better labelling affinity than BCR.

Consistent with distinct probe targets, **87** and BCR had also their own competitors. **87**-labelling can be competed by AEBSF but not by acetylated AEBSF (Ac-AEBSF). In contrast, BCR can be competed by Ac-AEBSF but not by AEBSF.

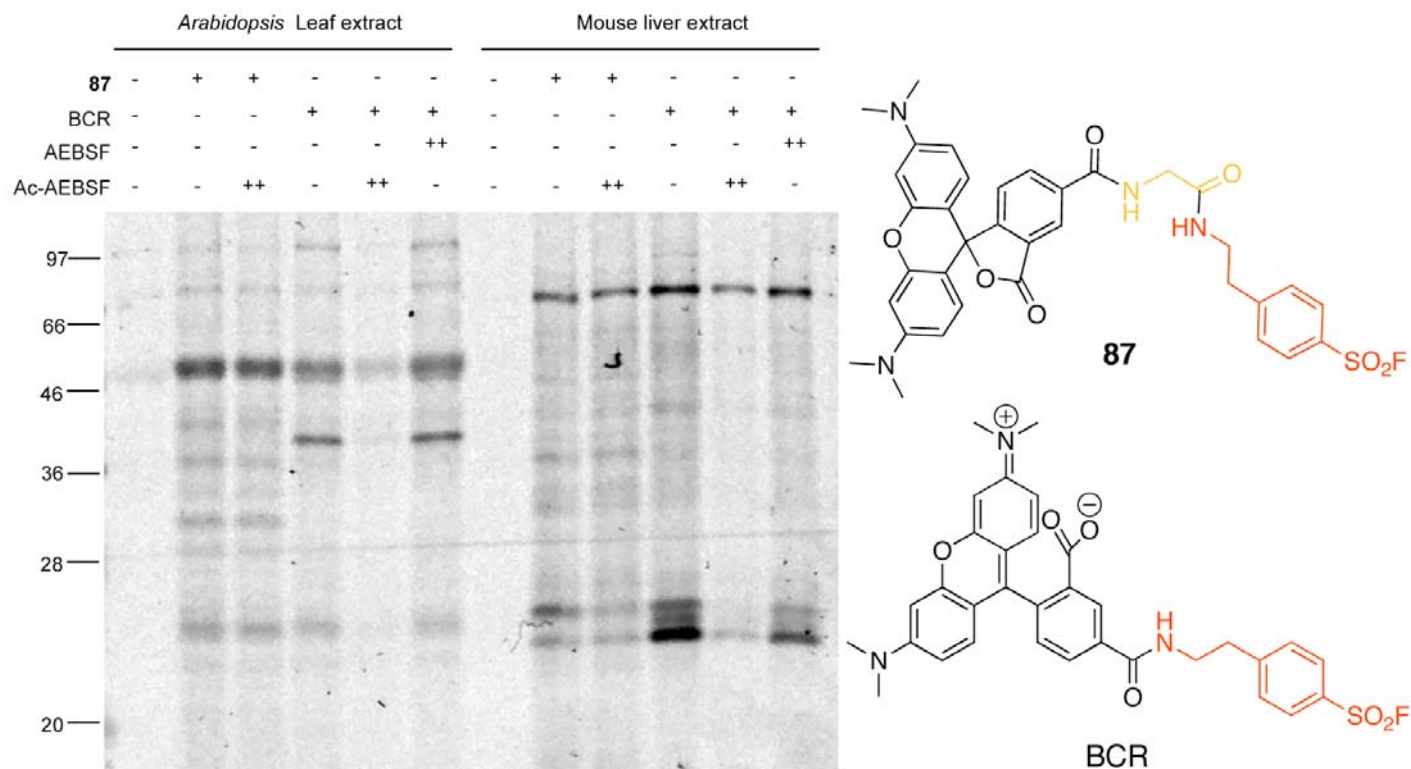


Figure 26. Comparison of the labelling profile between **87** and BCR: **87** and BCR have a same AEBSF warhead (red), a same fluorescence tag (black), the slight difference is that **87** has a glycine more (yellow). *Arabidopsis* leaf extracts or mouse liver extracts (~20 $\mu\text{g/ml}$) were pre-incubated at pH 8 with 600 μM AEBSF or Ac-AEBSF for 30 min and then labelled with 2 μM probe **87** or BCR at pH 8 for 2 h. The labelled proteins on protein blot were detected by fluorescence photometer.

3.4.4 Discussion

AEBSF-based probes did not label any serine proteases but many other enzymes such as GST, both in mammalian and plant proteomes. There are four possible explanations: One reason could be that AEBSF indeed does not target serine proteases, so the probe could not target serine proteases. Due to many clinic results showing that AEBSF cause phenol type, it could be that AEBSF does only

target serine proteases *in vivo* but not *in vitro*; an alternative explanation is that it targets serine proteases *in situ* but that label could be transferred to other enzymes under *in vitro* condition; the last possibility is that **87** is not equal to AEBSF itself, although it can be competed by AEBSF. Our studies could give previous pharmacological study of AEBSF a new idea. AEBSF can inhibit serine proteases *in vitro* however it does not mean that AEBSF must target serine proteases *in vivo* and it could have many other targets.

Except for the targets, a significant variation of labelling selectivity caused by different linker is another prominent finding in this study. The different amino acids in probes **82** – **86** do not lead to a variation of labelling selectivity. But removing this amino acid significantly changes the labelling properties of the probe. Variation of the amino acid (**82** – **86**) does not change the spacer distance between warhead and tag, and does not change the labelling selectivity, but removing this amino acid shortens the spacer, changing the labelling properties of the probe. A long spacer may reduce the mobility of the probe whereas a short spacer may cause spatial problems in the presence of a bulky reporter tag.

The reason of AEBSF-based probes preferably labelling at tyrosine could be explain that the benzene ring in AEBSF has an additional interaction with the phenol ring in tyrosine, and this interaction enhances the labelling affinity of AEBSF-based probes to tyrosine.

The difference between AEBSF and its acetylated form Ac-AEBSF is the free amino group in AEBSF, causing a positive charge at pH < 9 on AEBSF. But whether this additional dipole really induces the different competition ability of AEBSF and Ac-AEBSF to probe **87** is still unknown.

Based on our observations, AEBSF-based probes can be considered as multi-targets-reactivity probes, especially sensitive to GST, but they are not ABPs. Since

they do label many targets in one time, they could be used as an inhibitor instead of inhibitors cocktail in proteomics studies.

3.5 Mechanism-based ABPs design of AOMK for AvrPphB

3.5.1 Introduction

Plant–pathogen interactions are governed by specific interactions between a pathogen *avr* (avirulence) gene and the corresponding plant disease *R* (resistance) gene. When corresponding *R* and *avr* genes are present in both host and pathogen, disease will be resisted. However, if either of the two genes is inactive or absent, disease takes place (Flor *et al.*, 1971). This “gene-for-gene” hypothesis was proposed 40 years ago but the gene-for-gene interaction had been interpreted on the protein level in the nineties for the first time (Keen *et al.*, 1990). A model for the interaction between Avr and R proteins has been refined into “guard” and “decoy” models that involves an additional host protein (Van der Biezen *et al.*, 1998; Dangl *et al.*, 2001; Van der Hoorn *et al.*, 2008). To complete the model, more Avr and R protein pairs need to be studied.

The avirulence gene *avrPphB* was original discovered from *Pseudomonas syringae* pv. *phaseolicola*, which infects bean. Bean cultivars carrying the *R3* resistance gene introduce a hypersensitive reaction (HR) (Anastasia *et al.*, 2002). *AvrPphB* was later found to induce a HR in *Arabidopsis* plants carrying the NB-LRR resistance gene *RPS5* (Warren *et al.*, 1998 and Shao *et al.*, 2002). *AvrPphB* is secreted into host cells by the Type Three Secretion System (TTSS) (Tampakaki *et al.*, 2002), and belongs to a novel family of cysteine proteases. Its natural substrate is the *Arabidopsis* PBS1 protein kinase, which is specifically required for *AvrPphB*/*RPS5*-mediated resistance (Swiderski *et al.*, 2001). The crystal structure showed that *AvrPphB* has a papain-like fold with a distinct substrate-binding site (Zhu *et al.*, 2004). The information from the *AvrPphB*'s structure and amino acid

sequence of its natural substrate enable the design of its specific mechanism-based ABPs.

Acyloxymethyl ketones (AOMKs) were introduced first in the last century (Bromme *et al.*, 1989 and Pliura *et al.*, 1992) and were reported as inhibitors of asparaginyl endopeptidase (Loak *et al.*, 2004). ABPs with an AOMK warhead were generated by Boygo and his co-workers (Kato *et al.*, 2005) and showed exceptional class-wide reactivity with cysteine proteases in proteomes. AOMK has a high selectivity for two major clans of cysteine proteases, the CA clan (cathepsin B and L) and the CD clan (caspase-3, legumain, Arg- and Lys-gingipains) (Evans *et al.*, 2006). Since AvrPphB belongs to the CA clan, AOMK was chosen as warhead for the AvrPphB probes.

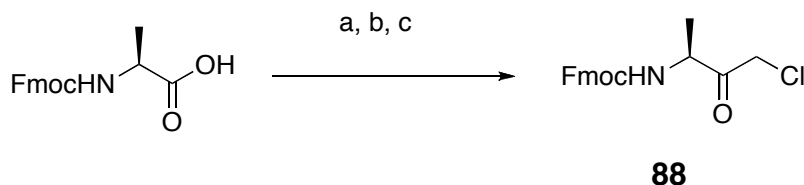
3.5.2 Synthesis

For the synthesis of these probes (Figure 27), a solid phase peptide synthesis strategy (SPPS) was chosen, following essentially the protocols developed by the Bogyo laboratory.

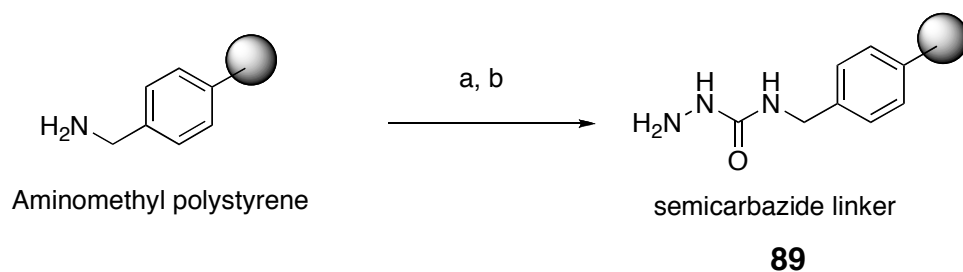
The AOMK moiety is introduced into the P1 position of the probe *via* a substitution reaction. To this end, the corresponding amino acid is transformed into a chloro or bromo methyl ketone (CMK or BMK) intermediate and then reacted with 2,6-dimethylbenzoic acid to obtain the AOMK-based amino acid (AS-AOMK). In order to prevent side reaction during the reaction of the CMK or BMK intermediate and 2,6-dimethylbenzoic acid, two different protocols have been used which take into consideration the different chemical properties of P1 inhibitor residue.

The designed probes for AvrPphB were based on the XXA-AOMK motif (thus alanine as a P1 residue), requiring the synthesis of an alanine AOMK (A-AOMK) building block. Its synthesis could be separated into two parts, i.e. the

generation of Fmoc-Ala-CMK in solution (Scheme 20) and preparation of a semicarbazide linker on aminomethylpolystyrene resin (Scheme 21). For the synthesis of Fmoc-Ala-CMK, Fmoc-protected alanine was first converted *in situ* into a mixed anhydride with isobutylchloroformate and *N*-methyl morpholine in tetrahydrofuran, followed by a reaction with freshly generated diazomethane. The newly formed organo diazomethane compounds were immediately reacted with a one to one mixture of HCl and acetic acid to obtain the final chloro methyl ketone **88**. The resin **89** with a semicarbazide linker was obtained from commercial aminomethyl polystyrene resin, to which first carbonyl diimidazole and then hydrazine were added.

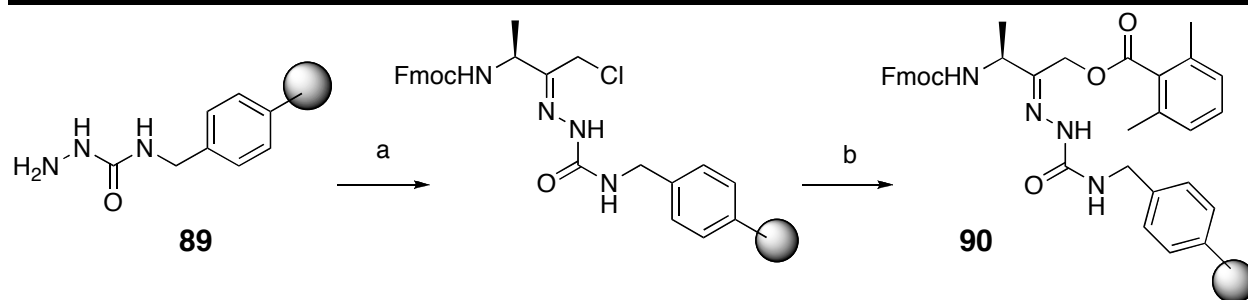


Scheme 20. Synthesis of **88**: a. isobutyl chloroformate, NMM, THF, -10°C , 25 min; b. diazomethane, 0°C – rt, 3h; c. HCl/AcOH, 0°C , 5 min, total 100%.



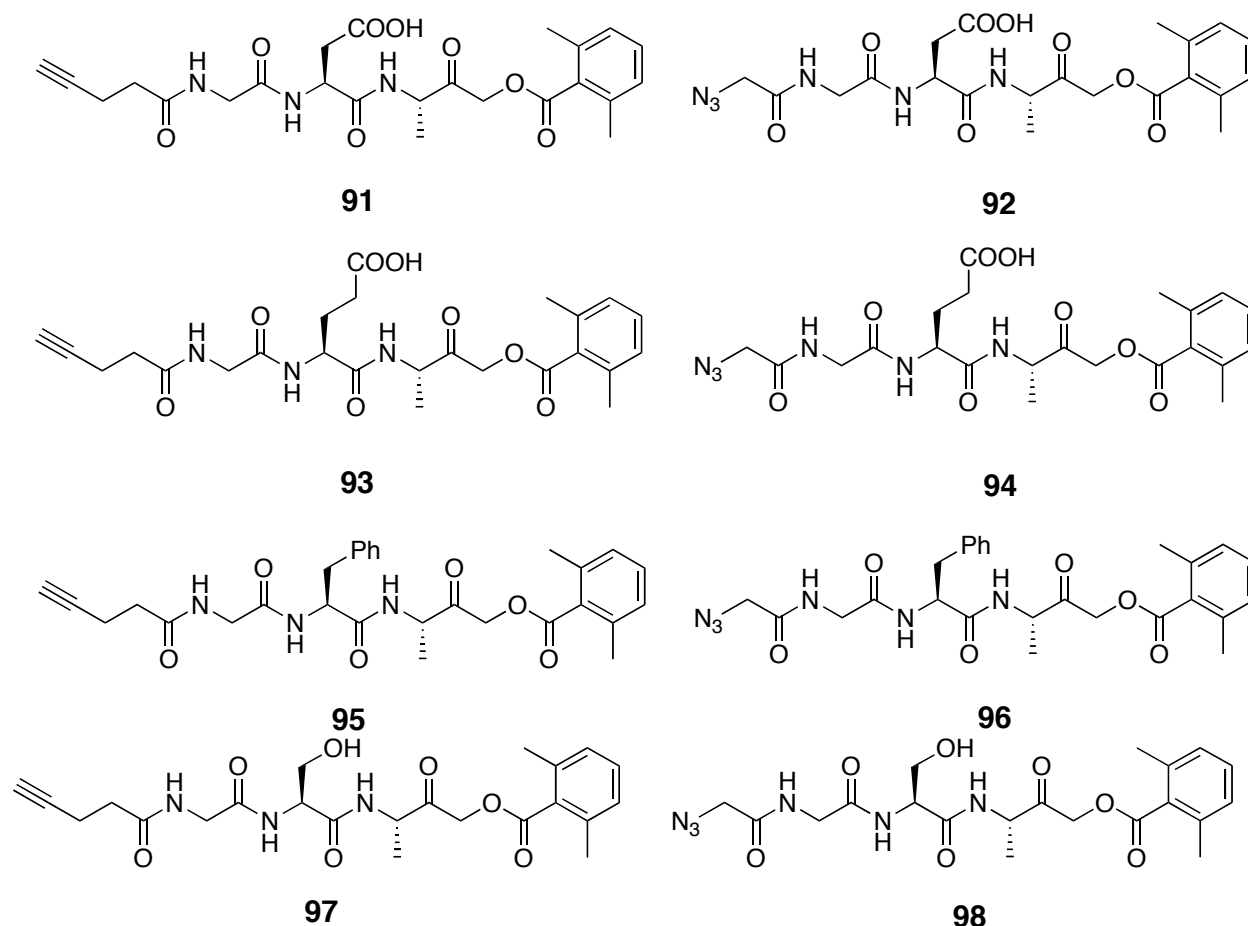
Scheme 21. Preparation of the semicarbazide linker **89**: a. carbonyldiimidazole, DMF, rt, 3h; b. hydrazine, DMF, rt, 1h.

After preparation of the semicarbazide linker, **88** was loaded on the resin and converted to resin-bound Fmoc-A-AOMK by reaction with dimethylbenzoic acid and KF in DMF to obtain the acyoxymethyl hydrazone **90** (Scheme 22).



Scheme 22. Synthesis of building block **90** for SPPS: a. **88**, DMF, 50°C, 3h; b. dimethylbenzoic acid, KF, DMF, rt, overnight.

After complete syntheses of intermediate **90**, standard SPPS was applied for the further synthesis of the probes (Figure 27), coupling first either aspartate, glutamate, phenylalanine, serine or arginine, then as a next amino acid glycine and finally 4-pentynoic acid or 2-azido acetic acid as the reporter tag of the probes.



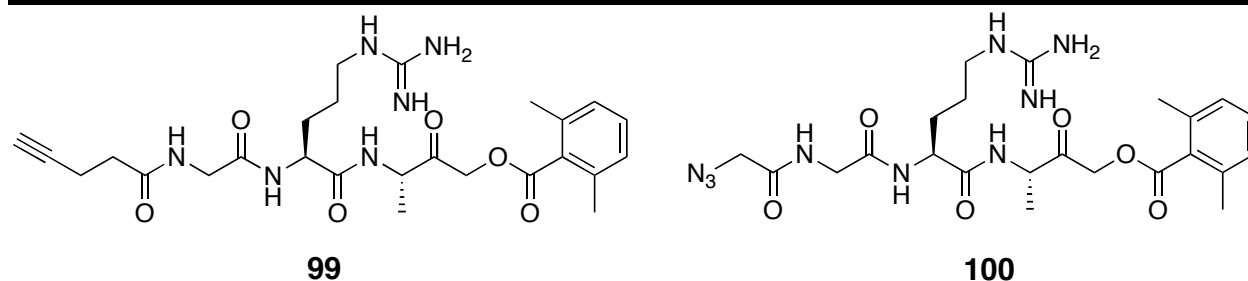


Figure 27. A collection of AOMK based probes for AvrPphB

As mentioned above, the natural sequence of the peptide is XXGDK, and the G241A, D242A, and K243A mutations reduce the peptide cleavage by about 90, 75, and 15%, respectively (Shao *et al.*, 2003). Consequently, variation of the P2 position of the probes was pursued to investigate the role of the unique acidic property in P2 position, which is highly unusual for ‘standard’ cysteine proteases of the papain-like super family. The former experience with probes such as IS4 or E64 suggested that papain-like cysteine proteases usually prefer hydrophobic groups such as a benzyl or *iso*-butyl group on the P2 position. Hence, the probes **95** and **96** were prepared. In order to check the impact of the negative charge on binding, probes **99** and **100** featuring a positive charged P2 position were also synthesized. Finally, to probe the general influence of side chain residues on binding, probes **91-94** and **97-98** were synthesized, which should allow a first glimpse on the spatial arrangement of the critical P2 position.

Later, after a first profiling assay revealed that **91** and **92** are most promising, a rhodamine labelled version **101** (later on named as FH11) was synthesized for more thorough screening (Figure 28).

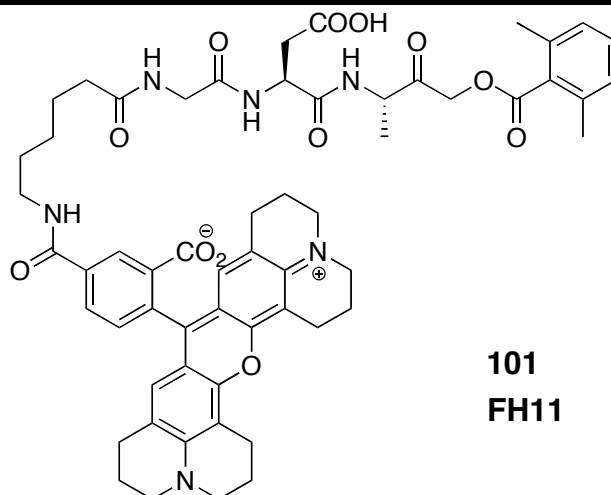


Figure 28. Structure of the fluorescently labelled probe **101**

3.5.3 Bioassays

AvrPphB is an avirulence protein from *Pseudomonas syringae* pv. *phaseolicola*. To obtain AvrPphB for labelling assays, we need first to clone the *avrPphB* gene and overexpress the protein. AvrPphB is injected into the host cytoplasm via the TTSS. When during the infection it is activated is unknown, and one hypothesis is that AvrPphB is only activated upon interacting with the plant proteome. Based on this hypothesis, we wanted to make sure to overexpress the active enzyme for labelling, so AvrPphB was overexpressed in *Nicotiana benthamiana* by infiltration of *Agrobacterium tumefaciens*. The *avrPphB* gene sequence was amplified from *Pseudomonas syringae* pv. *tomato* carrying an *avrPphB* encoding plasmid and cloned into a binary vector, behind a plant-specific constitutive 35S promoter (Figure 29a). The haemagglutinin (HA) epitope was fused to the C-terminus of AvrPphB to detect the protein with an anti-HA antibody. This binary vector (pZM05) was transformed into *Agrobacterium tumefaciens* via electroporation. The transformed *Agrobacterium tumefaciens* were infiltrated into leaves of *Nicotiana benthamiana*. Proteins were extracted after 2-5 days and analyzed by anti-HA western blot. However, AvrPphB was not detected with an

antibody on western blot. Instead, pZM05 triggered cell death within 2.5 days after agroinfiltration, which might explain the absence of AvrPphB-HA protein accumulation (Figure 29b).

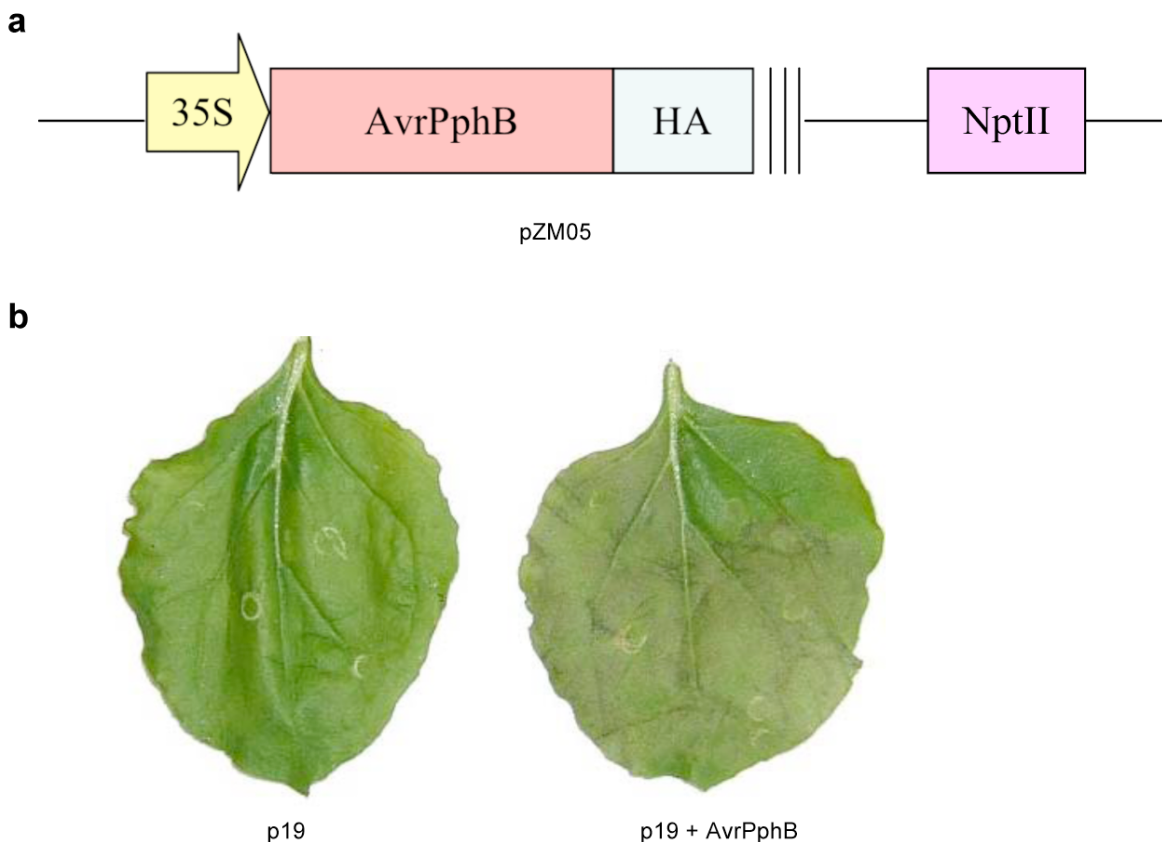


Figure 29. AvrPphB expression in *planta*: *Agrobacteria* carrying binary plasmid pZM05 were infiltrated into *N. benthamiana* leaves in the presence of silencing inhibitor p19; a. map of transfer DNA of pZM05 (35S, constitutive promoter; HA, haemagglutinin epitope; NptII, kanamycin resistance gene); b. AvrPphB expression triggers cell death. Pictures were taken at 2.5 days after agroinfiltration.

A second strategy for protein expression was to express AvrPphB in *Pseudomonas syringae* pv. *tomato* DC3000, which does not contain the *avrPphB* gene. Although AvrPphB would be expressed without interacting with plant proteome, it was still worth to test if AvrPphB is active. To perform this strategy, a new vector for bacterial over expression was constructed by Dr. Kaschani, using pZM05 as a template (Figure 30a).

Bacteria containing the vectors pFK141 (*AvrPphB-HA*) and pFK142 (*AvrPphB-his*) as well as the non-transformed bacteria were grown overnight at 28 °C in normal growth medium (NYG). To determine if *Pseudomonas syringae* secretes AvrPphB, the medium and bacterial pellet were separated and proteins from the medium were precipitated with acetone. Analysis of the proteins on western blot using anti-HA antibodies, revealed a 28 kDa signal in the bacterial pellet sample of bacteria expressing AvrPphB-HA, which matched the calculated molecular weight of AvrPphB-HA (Figure 30b). The protein expression level was low, and 10 minutes exposure not only caused a strong signal but also increased background signals of 34 kDa above the AvrPphB-HA signal.

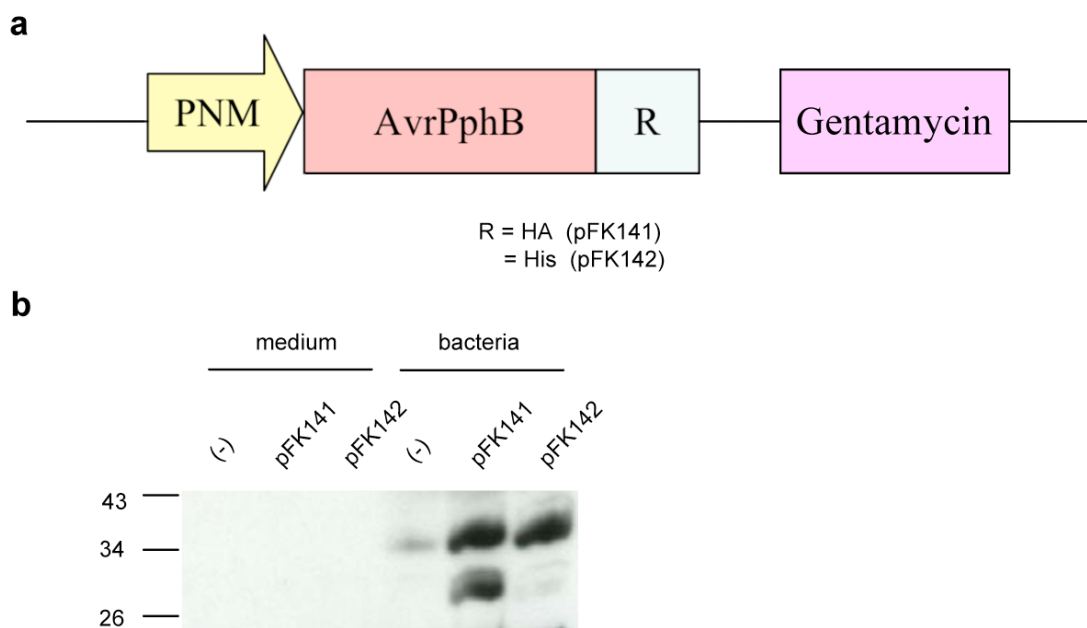


Figure 30. a, Structure of AvrPphB vectors: pFK141 contains a terminal HA epitope whereas pFK142 contains a terminal His tag; PNM, a strong constitutive promoter; b, AvrPphB expression in *Pseudomonas*: non-transformed *Pseudomonas* (-), *Pseudomonas* expressing HA-tagged (pFK141) or His-tagged (pFK142) AvrPphB were grown in rich medium; Secreted proteins (medium) and protein in bacterial pellet (bacteria) were analyzed on western blot, using anti-HA antibody.

These results demonstrated that AvrPphB was not secreted by *Pseudomonas* into the medium but it accumulated inside bacteria at low level. To increase protein

levels, *Pseudomonas* was grown in minimal medium (min-A), which mimics the natural habitat of *Pseudomonas* during infection. Western blot analysis with anti-HA antibody on proteins from bacteria revealed a strong 28 kDa signal for bacteria expressing AvrPphB-HA but only when the bacteria were grown in min-A (Figure 31a).

Protein extracts of *Pseudomonas* grown in min-A medium were used for a labelling experiment with probe **101**. Fluorescence scanning revealed clear 28 kDa signals from *Pseudomonas* carrying pFK141 and pFK142 suggesting that AvrPphB is covalently labelled with probe **101** (Figure 31b). The fluorescence bands correspond with the band on the anti-HA western blot and the coomassie gel.

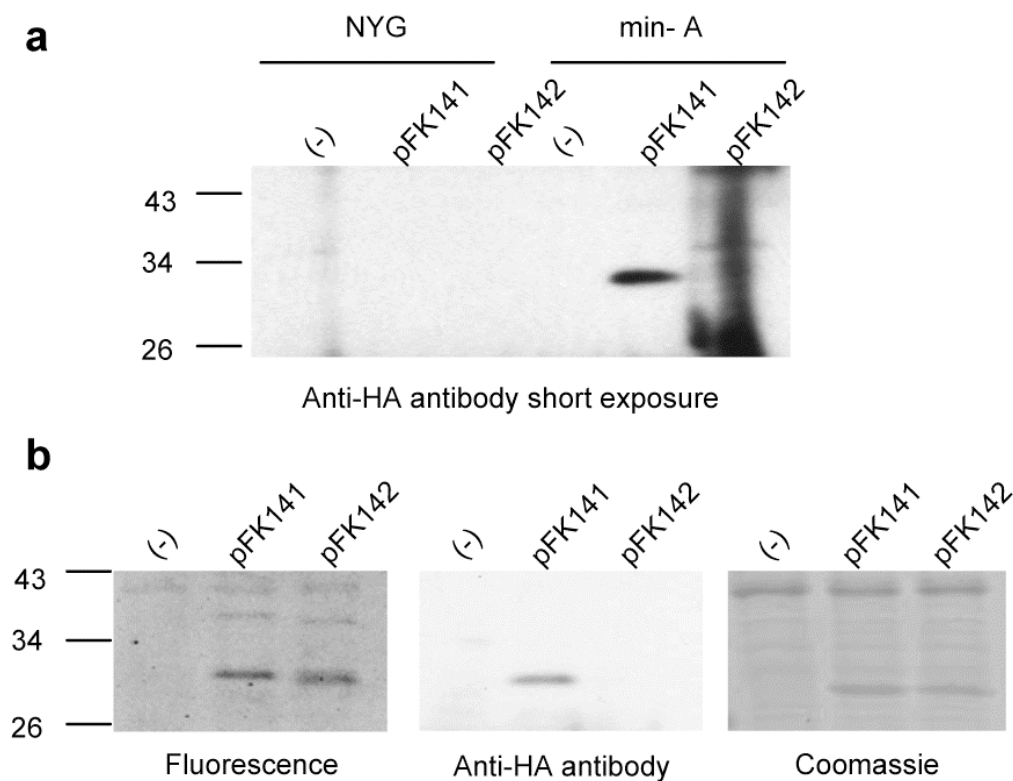


Figure 31. a. AvrPphB expression in two media: Different *Pseudomonas* strains were grown overnight in NYG or mim-A medium; The bacterial pellet was used for anti-HA antibody western blot analysis; b. Labelling of AvrPphB with probe **101**: Bacteria pellet proteins were incubated at pH 8 with 5 μ M **101** for 2 h. Fluorescent Proteins were detected by an in-gel fluorescence scanner, anti-HA antibody and coomassie staining in sequence.

To confirm that the labelled protein indeed represents AvrPphB, a pull-down assay was performed using Nickel beads, which bind the His tag of AvrPphB-his. The 28 kDa fluorescence signal only appeared in the enriched sample from *Pseudomonas* carrying vector pFK142 (Figure 32). This indicates that AvrPphB-His was captured on the Nickel beads in a high yield and the fluorescence band showed labelling of AvrPphB with **101**.

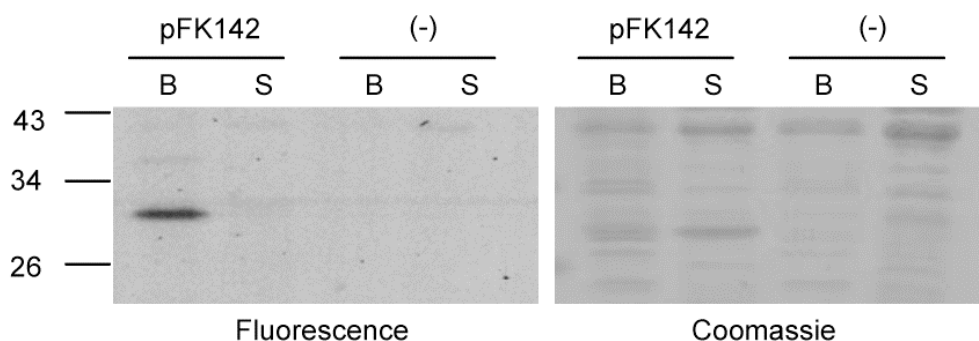


Figure 32. Pull down assay of **101**-labelled AvrPphB-His: Extracts from pFK142-containing bacteria were labelled with **101** for 2 hours and incubated Nickel beads for 2 h to immobilize His-tagged proteins; Proteins on Beads (B) or Supernatant (S) were detected in gel by fluorescence scanning and coomassie staining.

These results not only showed the activity of the probe but also the activity of AvrPphB. Based on these results, two properties of AvrPphB were revealed. First, AvrPphB is already active in the extract from *Pseudomonas*. Second, AvrPphB is not secreted by *Pseudomonas* unless there is a stimuli, which could happen during infection.

To further investigate the properties of AvrPphB, labelling was performed at various pH values. This revealed that AvrPphB is active at neutral pH value, coinciding with the pH value of the plant cytoplasm (Figure 33). In fact, pH 8 was not the good labelling pH for probe **101** but the neutral pH values.

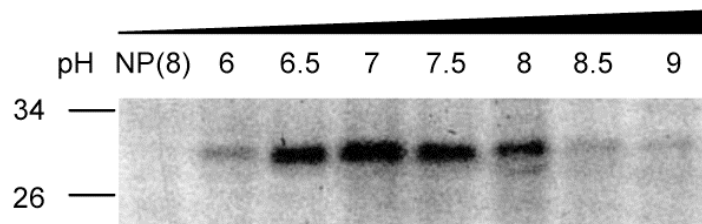


Figure 33. pH curve of labelling of AvrPphB with **101**: AvrPphB-containing extracts were incubated at various pH with 5 μ M **101** for 2 h, Fluorescent proteins were detected by in-gel fluorescence scanner.

PLCPs are usually active at acidic pH value, but AvrPphB is active at neutral pH value. Since the cytoplasm of *Pseudomonas syringae* is mildly acidic (Dawson *et al.*, 2009), AvrPphB is inactive inside bacteria. When the bacteria inject AvrPphB into the cytoplasm of host cell, having a neutral pH value, AvrPphB is activated. Thus in our pH-dependent ABPP assay, the AvrPphB shows the most labelling activity at pH 7, which matches the physiological reason in nature.

The last assay was the competition assay to test the specificity of probe **101** (Figure 34). Labelling by **101** (P2 = D) could be competed by pre-incubation with **92** (P2 = D), but not by the other inhibitors (P = E, F, S, R). This shows that **101** and **92** have a high labelling specificity. Especially in the case of **92** where the P2 position is aspartic acid whereas the P2 position of **94** was glutamic acid. The fact that inhibitors containing P2 = D are effective, but P2 = E not, is intriguing since the two inhibitors differ only in one methylene group, indicating that the S2 pocket of AvrPphB is too shallow to accommodate P2 = E.

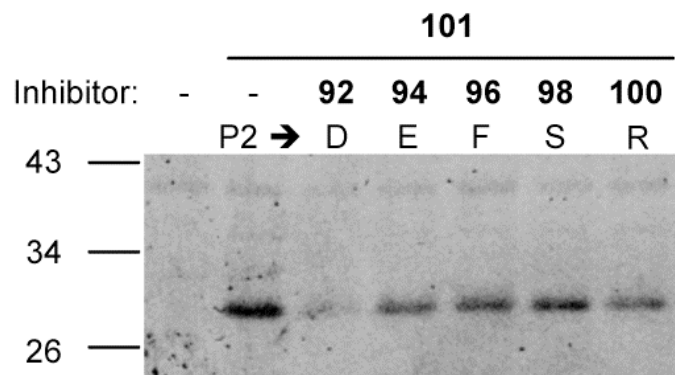


Figure 34. Competition assay of **101**: The AvrPphB containing extracts were pre-incubated with alkyne-tagged probes **92**, **94**, **96**, **98**, **100** at 100 μ M for 30 minutes, and remaining non-inhibited AvrPphB was labelled with fluorescence probe **101** of 5 μ M for another 1.5 hours. Fluorescent proteins were detected by in-gel fluorescence scanner.

3.5.4 Discussion

The design of probe **101** as well as its click version **91** and **92** were successful according to the biological results. Due to the highly selective AOMK moiety and using aspartate and glycine at P2 and P3 position, the probes are specific to AvrPphB. Variation of amino acid at the P2 position confirms that the AvrPphB active pocket is selective towards the P2 position of the substrate. This is a unique character among PLCPs, since most PLCPs prefer substrates with aliphatic or aromatic residues at the P2 position. However, the natural substrate of AvrPphB carries aspartate at P2 position. The competition experiment showed that only inhibitors with aspartic acid at P2 position prevent **101** labelling but even with glutamic acid did not. This implies that the S2 substrate-binding pocket in AvrPphB is too shallow to accommodate the extra chemical residue. This experimental observation is supported by the crystal structure, in which an arginine residue is located in the S2 pocket that could be responsible for the narrow space. The absent labelling activity of AvrPphB at low pH value could result the denaturation of AvrPphB. Recently another avr protein, AvrPto was reported that it is denaturated at low pH in order to be transportable via TTSS (Dawson *et al.*,

2009). The pH-sensitivity and that strict selectivity for P2 = D by AvrPphB could be related to the physiological function of this protease.

After generation of ABPs for AvrPphB, future biological work on AvrPphB could be benefited and carried on utilizing the ABPP. We could still investigate the time course of expression and secretion of AvrPphB by *Pseudomonas* into different hosts *in vivo*; the localisation of secreted AvrPphB in plant intra- or extracellular; or to monitor the proteolytic in- or active state of AvrPphB. To explore the role of AvrPphB in the plant-pathogen interactions, the AvrPphB probes could be utilized as a powerful tool.

3.6 Mechanism-based ABPs design of AOMK for VPE

3.6.1 Introduction

Vacuolar processing enzymes (VPEs) are cysteine proteases responsible for the maturation of various vacuolar proteins in higher plants (Hara-Nishimura *et al.*, 1991). VPE can be subdivided into two subfamilies, one is specific for protein storage vacuoles in seeds and another is specific for lytic vacuoles in vegetative organs (Kinoshita *et al.*, 1995). The VPE-mediated processing is similar in both situations and VPE cleaves at Asn-Gln bonds of natural substrates with a unique mechanism (Yamada *et al.*, 1999). Activation of VPE occurs through self-catalytic removal of an auto-inhibitory C-terminal propeptide (Kuroyanagi *et al.*, 2002). This self-activation process is similar to that of AvrPphB. Like caspases, VPE is involved in a mediate virus-induced hypersensitive cell death (Hatsugai *et al.*, 2004). There are four VPEs in *Arabidopsis*, namely α , β , γ and δ -VPE. The reversible Caspase-1 inhibitor Ac-YVAD-CHO and VPE inhibitor Ac-ESEN-CHO were introduced already. To monitor the change of activities for VPE *in vivo* and *in vitro*, specific ABPs need to be generated.

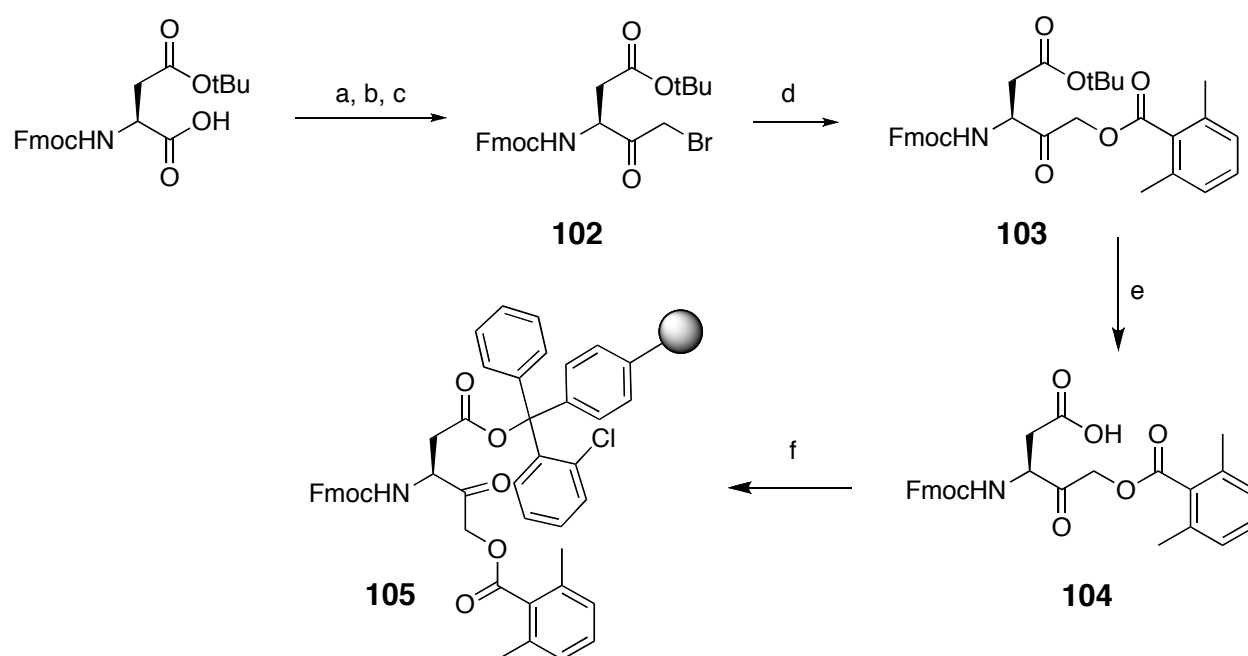
Based on the success of AvrPphB probes, a similar probe design strategy was applied. AOMK was chosen as the warhead, since it can carry a P1 amino acid residue that determines selectivity to clan CD proteases like caspase and VPE.

3.6.2 Syntheses

The AOMK-based probes for VPE were synthesized by the same strategy as the probes for AvrPphB. The VPE probes were based on a XXN-AOMK motif. Contrary to Fmoc-Ala-AOMK, which was linked to the resin *via* a semicarbazide linker, the corresponding Asn-AOMK derivative can be obtained *via* side chain

linking of a corresponding Asp-AOMK building block to a 2-chlorotrityl chloride resin (Scheme 23). Consequently, Fmoc-Asp-AOMK was required as a building block for solid phase synthesis.

To this end, Fmoc-Asp(O*t*Bu)-OH was used as a starting material and converted into Fmoc-Asp(O*t*Bu)-BMK, followed by solution transformation into the corresponding Fmoc-Asp(O*t*Bu)-AOMK **103** (Scheme 23). Deprotection of the side chain *t*-Bu-ester with TFA generated building block **104**, which was then coupled to 2-chlorotrityl chloride resin to deliver resin bound **105**.



Scheme 23. Synthesis of building block **105** for SPPS: a. isobutylchloroformate, NMM, THF, -10 °C, 25 min; b. diazomethane, 0 °C - rt, 3h; c. HBr/AcOH, 0 °C, total 100%; d. dimethylbenzoic acid, KF, DMF, rt, overnight, 60%; e. TFA, DCM, rt, 1h, 100%; f. 2-Cl-trityl resin, DIPEA, DCM, rt, 2h.

The BMK reaction condition was much harsher than for CMK and hydrogen bromide was applied very carefully. After the synthesis of **105**, standard SPPS was used to synthesize the biotinylated and fluorophore attached probes **106** and **107** (Figure 35).

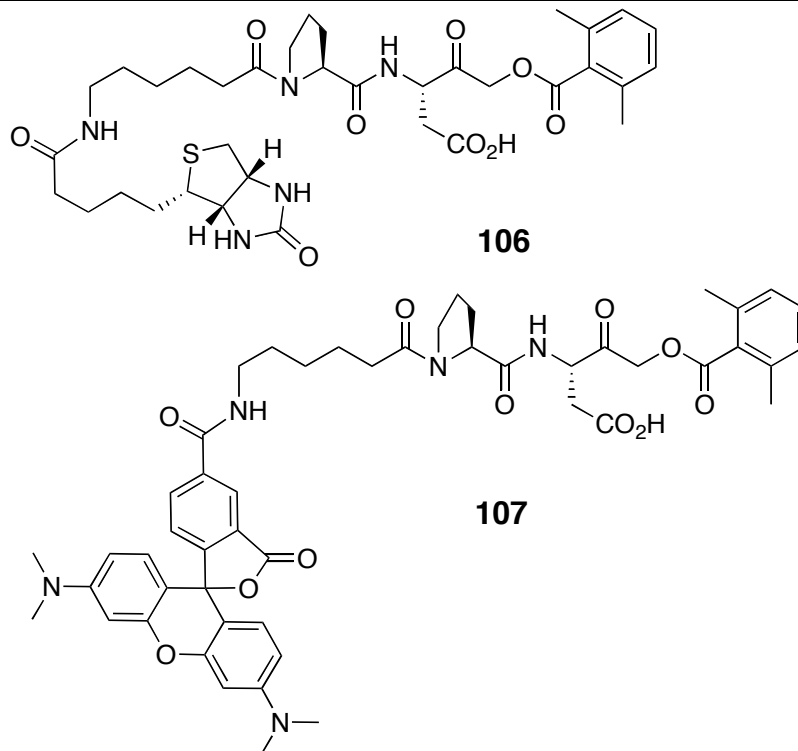


Figure 35. Structures of the biotinylated probe **106** and fluorescently labelled probe **107**

3.6.3 Bioassays

The probe **107** was directly labelled with *Arabidopsis* leaf extract of different lines: *Columbia* (*Col*) wildtype contains an endogenous amount of VPE; γ -VPE overexpressed (*O γ*) line has high level of γ -VPE protein; γ -VPE knockout (*K γ*) line and VPE quadruple knockout (*K*) lines lack γ -VPE or all VPEs respectively. The result showed that probe **107** caused two 43 kDa signals in *Col* and *O γ* lines. The size of the labelling coincides with the theoretical VPE protein size. The signal was absent in single and quadruple knockout lines but more intensive in the γ -VPE overexpressed line, confirming that the signal represents VPE (Figure 36a). To prove that the labelling was specific, a competition assay of probe **107** was carried out with its cognate biotin tagged probe **106**. The assay showed that the labelling of **107** can be competed by **106** (Figure 36b).

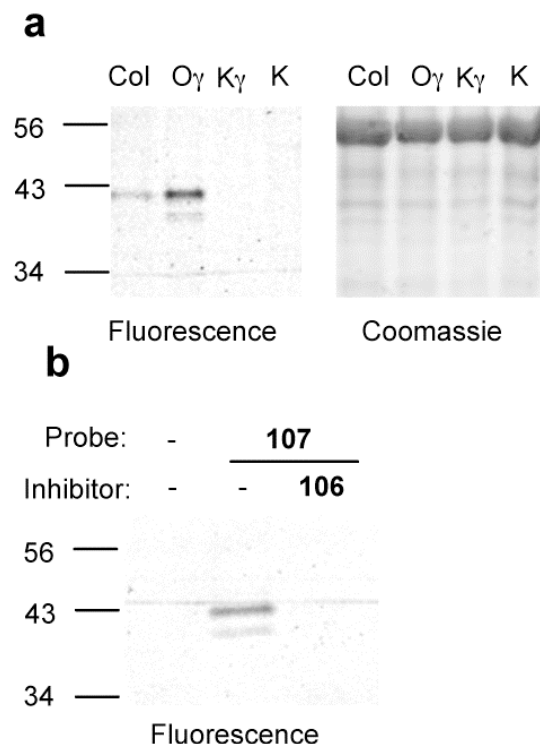


Figure 36. Evaluation of probe **107** as an ABP for VPE: a, Labelling of probe **107** with *Arabidopsis* leaf extract of different lines: *Columbia* wildtype (Col), γ -VPE overexpressed ($O\gamma$), γ -VPE knockout ($K\gamma$), α , β , γ , δ -VPEs quadruple knockout (K); b, biotin tagged probe **106** prevents the labelling of probe **107** in *Arabidopsis* leaf extract of γ -VPE overexpressed line; *Arabidopsis* leaf extracts of different lines (~ 100 $\mu\text{g/ml}$) was labelled with 5 μM probe **107** (a) or pre-incubated with 100 μM **106** for 30 min and then labelled with 5 μM probe **107** (b) at pH 5.5 for 2 h. The labelled proteins on protein blot were detected by fluorescence photometer and coomassie stain.

3.6.4 Discussion

The simple biological assays prove that **107** is an ABP for VPE of *Arabidopsis*. The probe is very specific and highly selective for VPE, so this project gave another good example of the selectivity of the AOMK warhead to cysteine proteases. Unlike AOMK-based probe **101** for AvrPphB (chapter 3.5), which is a CA clan cysteine protease, this AOMK-based probe **107** is for VPE, a CD clan cysteine protease. The difference between probes **101** and **107** is only two different amino acids in the P1 and P2 positions, illustrating that the subtle

selectivity of targeting cysteine protease is dependent on the properties of the substrate pocket of the enzyme.

3.7 Natural product syntheses of Gibbestatin

3.7.1 Introduction

To date, Gibbestatins A-C have been described (Hayashi *et al.*, 1999) of which Gibbestatin B (GNB) is the most important derivative. It was originally isolated in 1984 as the antibiotic 2-11-B from *Streptomyces* sp. 2-11 (Kinashi *et al.*, 1984), however without stereochemical assignments. In 1999, together with two other gibbestatins, the structure of GNB (Figure 37) was more thoroughly elucidated by spectroscopic analyses (Hayashi *et al.*, 1999). Despite these advances, two chiral centres at C16 and C17 could only be assigned in relative stereochemistry (*trans*). Later, the same group reported an inhibitory effect of GNB on both gibberellin A3 (GA)-induced α -amylase production in the cereal aleurone layers and the expression of various genes induced by GA, abscisic acid (ABA) and auxin on tobacco and *Arabidopsis* plants (Hayashi *et al.*, 2000). This report concluded with the notion that GNB acts in a similar manner as typical plant hormones.

To investigate the exact biological functional of GNB, an identification of its biological target is required. To this end, two strategies are in principle applicable: first, GNB can be transformed into an ABP for target pull-down experiments or second, a Y3H can be performed, necessitating a transformation of GNB into suitable hybrid ligands (Schneider *et al.*, 2008). As access to the natural product is limited, both strategies require a previous natural product synthesis.

The structure of GNB features one epoxide moiety and several double bonds. Its mode of action is still unknown, however several possibilities for a potential reaction of GNB with cellular proteins can be imagined from its functional groups. Thus, the epoxide group could be attacked by a nucleophile in an analogues manner

as in the natural product E-64, leading to ring opening either at C-16 as well as C-17; in addition, GNB could be attacked by a nucleophile in a Michael-type reaction. Importantly, both reactions would lead to stable, covalent protein-GNB conjugates, allowing efficient pull-down experiments with GNB in complex proteomes.

As GNB represents an interesting natural product due to its inherent biological activity, a total synthesis was pursued, aiming at the synthesis of sufficient amounts of GNB for further biological screening and target identification. To date, no synthesis of GNB or a structurally related natural product has yet been reported.

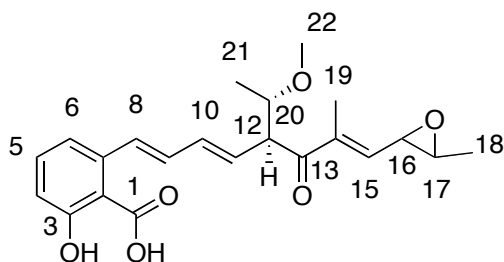
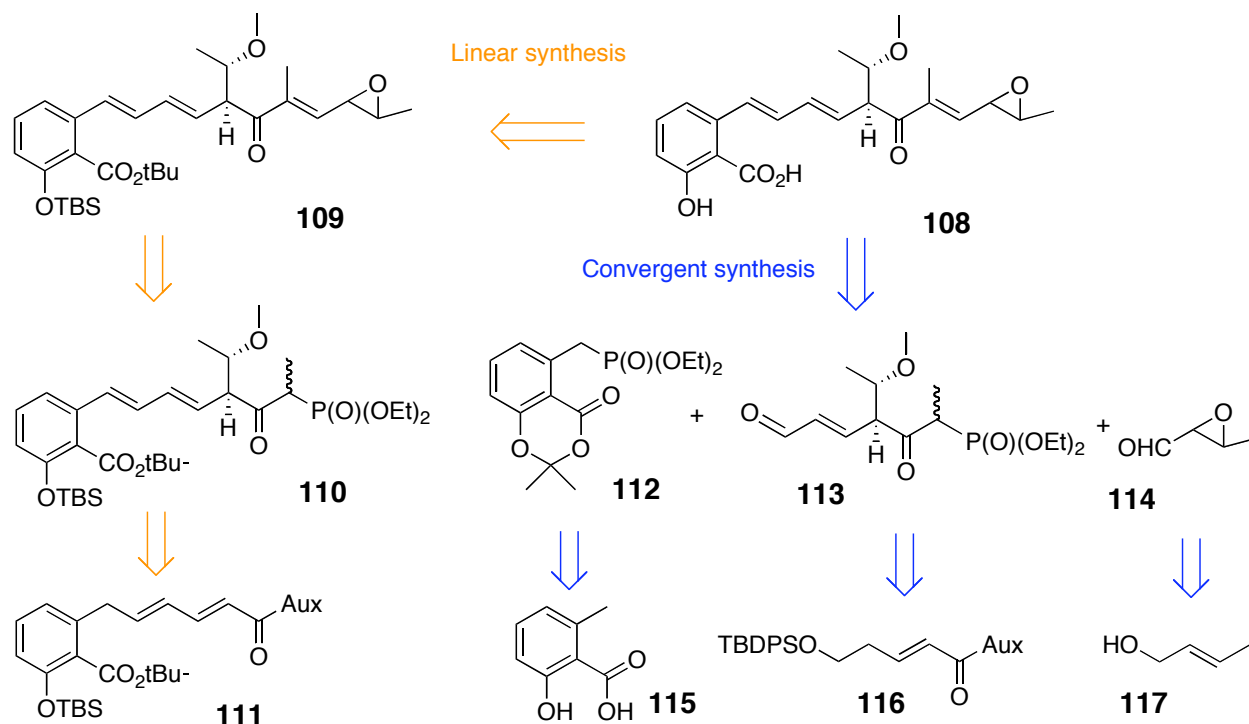


Figure 37. Structure of GNB

Although GNB is only a middle size natural product, the high density of functional groups still turns it into an interesting synthetic challenge. GNB features three double bonds, one benzene ring, one epoxide moiety and one carbonyl group in a only 22-carbon atoms molecule. Retro-synthetically, several strategies for synthesis of GNB can be differentiated.

3.7.2 Previous synthesis approaches for GNB



Scheme 24. Comparison of two retro-synthetic strategies towards GNB synthesis developed by Dr. Rama Narayana: a linear synthesis route (orange) and a more convergent synthesis route (blue) as an improved second route. However, both routes failed to deliver the desired product at a late stage of synthesis.

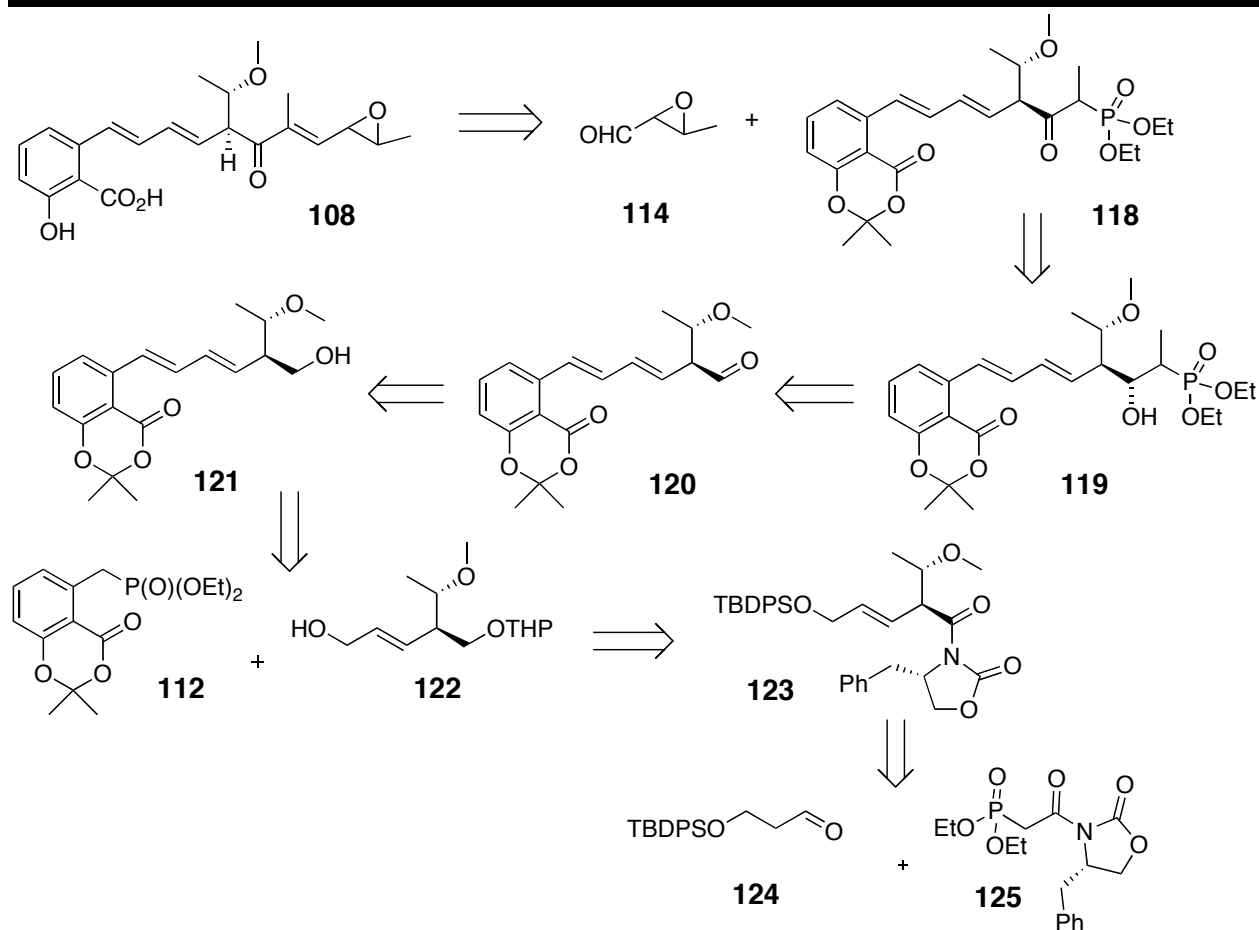
The previous synthetic trial was performed by Dr. Rama Narayana (A postdoctoral fellow in the Kaiser group), and he first used a linear synthetic strategy to synthesize GNB. Because only the chiral centres on C12 and C20 were assigned during previous spectroscopic studies, an enantioselective aldol reaction was adapted for the generation of the 12*S* and 20*S* chiral centres. A Sharpless asymmetric epoxidation was utilized to generate both derivatives (16*S*, 17*S* & 16*R*, 17*R*) on the epoxide moiety, because chiral information on C16 and C17 are still unknown (Scheme 24). Thus, first a deconjugated salicylic acid derivative **111** was synthesized (Scheme 24), which was then subsequently transformed into the final product by an *anti*-aldol and a HWE reaction as key steps. Starting from 2-methoxy benzoic acid, 11 steps are required to synthesize **111** (unpublished data).

It however turned out that the last step of this reaction sequence, consisting of a hydrolysis of methyl ester and adjacent coupling with the chiral auxiliary for the following aldol reaction could be achieved only in disappointing bad yields. Consequently, an improved synthetic route was sought.

To this end, a convergent synthetic strategy was devised, basing again on an auxiliary-mediated enantioselective aldol reaction and a Sharpless epoxidation. Due to the instability of the TBS and *t*-butyl protecting groups on the salicylic acid moiety during the first synthetic trial, an alternative protecting group, i.e. an acetone acetal was used during the convergent synthetic route. As a key idea, GBN was divided into three building blocks *via* a formal separation of C8-C9 and C14-C15 (Scheme 24). The synthesis of each building block can be achieved by straightforward chemistry. Indeed, the syntheses of the C8 benzyl phosphonate building block **112** (Scheme 24) and C15 epoxide building block **114** (Scheme 24) were optimized during the synthetic studies and could be achieved in good to high yields. The synthesis of the required building block **113** (Scheme 24) however could still not be achieved satisfyingly as the aldol reaction proved unexpectedly difficult (unpublished data). Although an assembly of **113** with the **114** could be achieved, the subsequent HWE reaction with **112** to complete the synthesis could not be performed under the investigated reaction conditions.

3.7.3 First synthesis approach for GNB

Retro-synthetic analysis of the first synthesis approach



Scheme 25. Second convergent retro-synthetic strategy

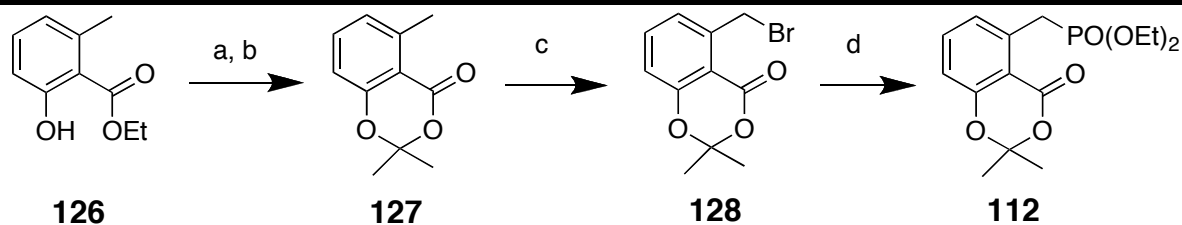
At this point in time, an improved synthetic route including the gained knowledge of the previous trials of Dr. Narayana was developed during this thesis (Scheme 25): As the last step, i.e. the HWE reaction between **112** and **113/114** building block failed in the previous synthetic route, the order of reactions was changed in the second convergent retro-synthesis. Thus, first **113** and **112** building blocks were connected by a HWE reaction and then as a last step, a HWE reaction between **114** and the other C14 phosphonate derivative **118** to form the C14-C15 double bond was envisaged. The synthesis of **114** was already established previously. Consequently, only the generation of **118** had to be investigated. To this end, the C14 phosphonate moiety **118** was envisioned to be accessible by an elongation of the C13 aldehyde moiety **120** by a nucleophilic addition with diethyl

ethylphosphonate to form the C13-C14 bond. **120** could be obtained from **121**, which in turn could be prepared by a HWE reaction between **112**, whose synthesis was previously established, and the C9 aldol product based moiety **122**. **122** could be generated with an auxiliary-based enantioselective aldol reaction as a key step within total 12 steps in total.

However, a critical view on this synthetic strategy reveals that the generation of the two chiral centers on C12 and C20 is rather elaborate. As a consequence, the total yield of the reaction sequence was rather low, turning the final assembly of the synthesized building block into a tedious task. Thus, an alternative strategy would be advantageous for a more efficient synthesis of GNB.

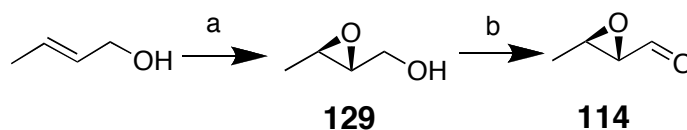
Synthesis of the first synthesis approach

Benefited from the former experience of the GNB synthesis by Dr. Rama Narayana and the already obtained intermediates from his former synthesis, the synthesis of building block **112** was started with **126**. **126** was refluxed in 10% sodium hydroxide ethanol solution for 2 hours to hydrolyze the ethyl ester. The obtained 6-methyl salicylic acid was isolated without further purification and reacted with acetone, thionyl chloride and a catalytic amount of DMAP in dimethoxyethane to protect the hydroxyl and carboxyl groups of salicylic acid as acetonide **127**. Benzylic bromination with NBS and AIBN yielded **128**, which then was refluxed in anhydrous toluene with triethylphosphite to generate the desired benzyl phosphonate building block **112** (Scheme 26).



Scheme 26. Synthesis of building block **112**: a. NaOH, EtOH, reflux, 2 h; b. DMAP, CH₃COCH₃, SOCl₂, Dimethoxyethane, < 30 °C, 3 h, 89%; c. NBS, AIBN, CCl₄, reflux, overnight, 99%; d. P(OEt)₃/Toluene, reflux, 3 h, 96%.

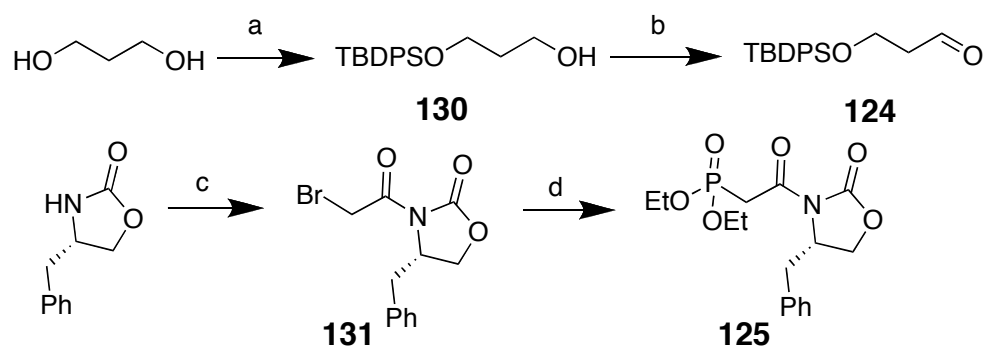
To simplify the synthetic work, the (16*S*, 17*S*) isomer on the epoxide moiety was first chosen as the target molecule. After successful establishment of the synthetic route, the corresponding (16*R*, 17*R*) isomer could be synthesized easily. The synthesis of the required building block **114** started with a Sharpless asymmetric epoxidation of (*E*)-2-buten-1-ol to obtain [(2*S*, 3*S*)-3-methyloxiran-2-yl]methanol (**129**), which was then oxidized with DMP to generate the volatile building block **114** (Scheme 27).



Scheme 27. Synthesis of building block **114**: a. DIPT, tBuOOH, Ti(OiPr)₄, DCM, - 20 °C, 2 h, 54%; b. DMP, DCM, 0 °C, 1 h, 90%.

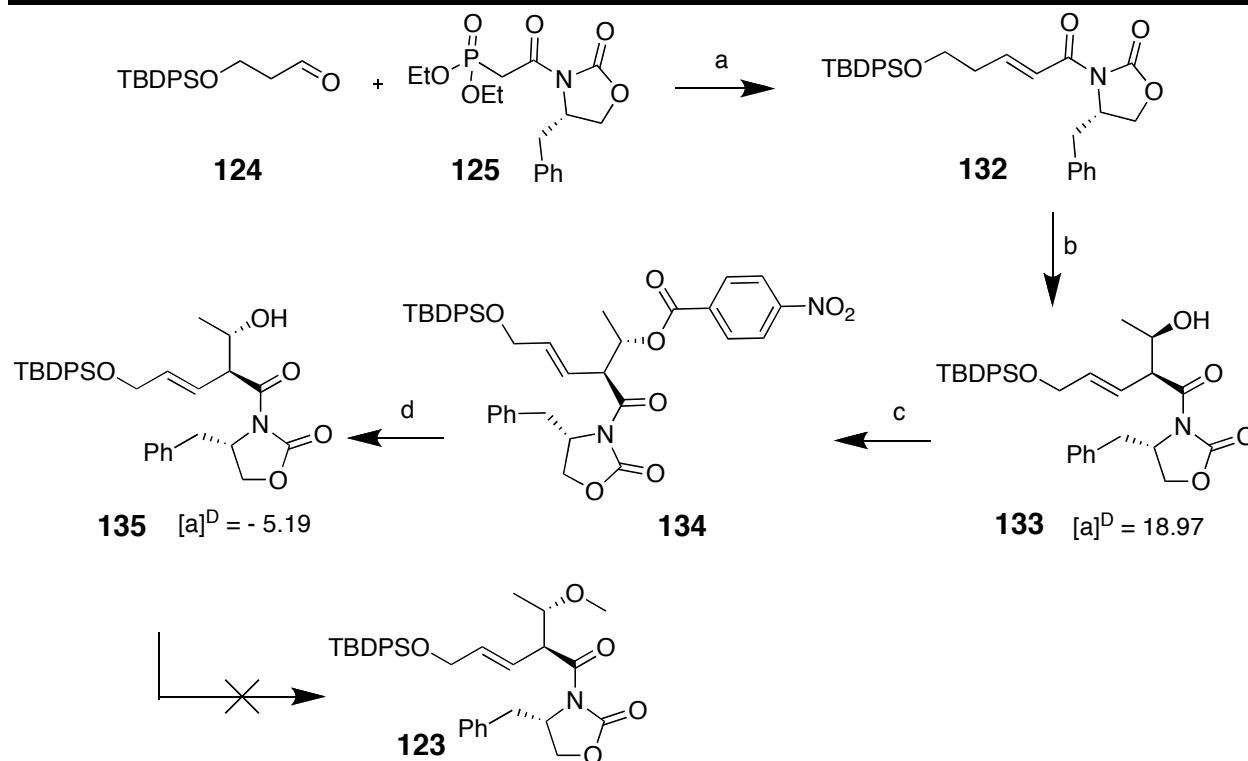
To perform the second convergent retro-synthetic strategy, preparation of the aldol-derived building block **122** is essential. The direct Masamune *anti*-aldol reaction was not compatible with the later reaction conditions, so a combination of a *syn*-aldol reaction and Mitsunobu inversion was adapted (unpublished data). To perform this strategy, two small fragments **124** and **125** were required. **124** was synthesized by mono-protection of 1,3-propanediol with TBDPSCl to generate **130** followed by oxidation to the aldehyde **124** with a Swern oxidation. **125** was prepared from a coupling reaction of bromoacetyl bromide with of (2*R*)-4-(phenyl

methyl)-2-oxazolidinone (Evans auxiliary) to form **131** followed by bromide displacement by refluxing with triethylphosphite (Scheme 28).



Scheme 28. Syntheses of building blocks **124** and **125**: a. TBDPSCl, DIPEA, DCM, rt, overnight, 90%; b. DMSO, (COCl)₂, DCM, TEA, -78 °C, 2 h, 70%; c. bromoacetyl bromide, *n*-BuLi, THF, -78°C, 1 h, 94%; d. P(OEt)₃, reflux, overnight, 80%.

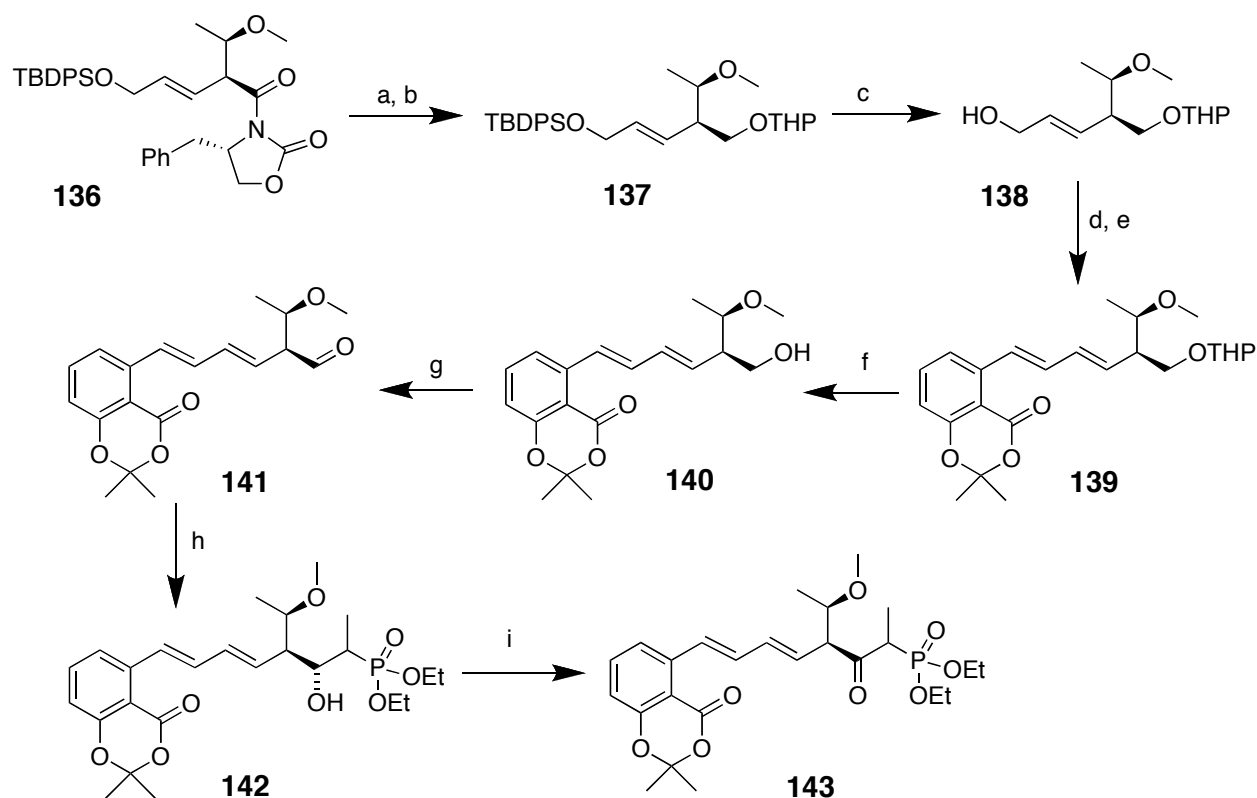
After accomplishing the preparation of **124** and **125**, a HWE reaction between both led to **132**, which then was subjected to a *syn*-aldol reaction with acetaldehyde to generate the *syn*-aldol product **133**. In order to achieve the required inversion of the hydroxyl group, **133** was first transformed into an *anti-para*-nitrobenzoic ester **134** by a Mitsunobu-type reaction, which was then hydrolyzed to generate the desired *anti*-aldol product **135**. Surprisingly, the subsequent methylation of **135** failed with all employed methylation reagents such as iodomethane, methyl triflate and diazomethane in combination with different bases such as 2,6-di-*tert*-butyl-4-methyl pyridine, pyridine or TEA. The reason for the unsuccessful methylation could be the unfavorable spatial conformation of the substrate. The auxiliary seemed to strongly shield the reaction center of **135** by other reagents. As a consequence, the required building block **123** could not be synthesized (Scheme 29).



Scheme 29. Synthesis of building block **123**: a. NaHMDS, THF, 0 °C – rt, 1 h, 85%; b. Bu₂BOTf, TEA, CH₃CHO, DCM, -78 °C – rt, 4 h, 91%; c. DIAD, TPP, PNBA, THF, 0 °C, 30 min, 91%; d. K₂CO₃, MeOH, rt, 1.5 h, 79%.

Although **135** could not be methylated directly, this problem could be circumvented by removing the shielding auxiliary, which only results in some additional steps. However, in order to test if the designed synthetic strategy is feasible at all, the methylated *syn*-aldol product **136** was used to perform the test reactions. **136** was reduced with LiBH₄ to remove the auxiliary and then protected with DHP to generate **137**. The TBDPS protecting group was released with TBAF to obtain **138**, which was then *in situ* oxidized to an aldehyde with DMP and transformed into **139** *via* a HWE reaction with building block **112**. Deprotection of THP by PPTS in methanol yielded **140**, which was then oxidized with DMP to **141**. The aldehyde **141** was reacted with deprotonated diethylethylphosphonate to generate **142**, which was then transformed into **143** by Dess-Martin oxidation (Scheme 30). The subsequent synthesis had to be stopped at this step because the

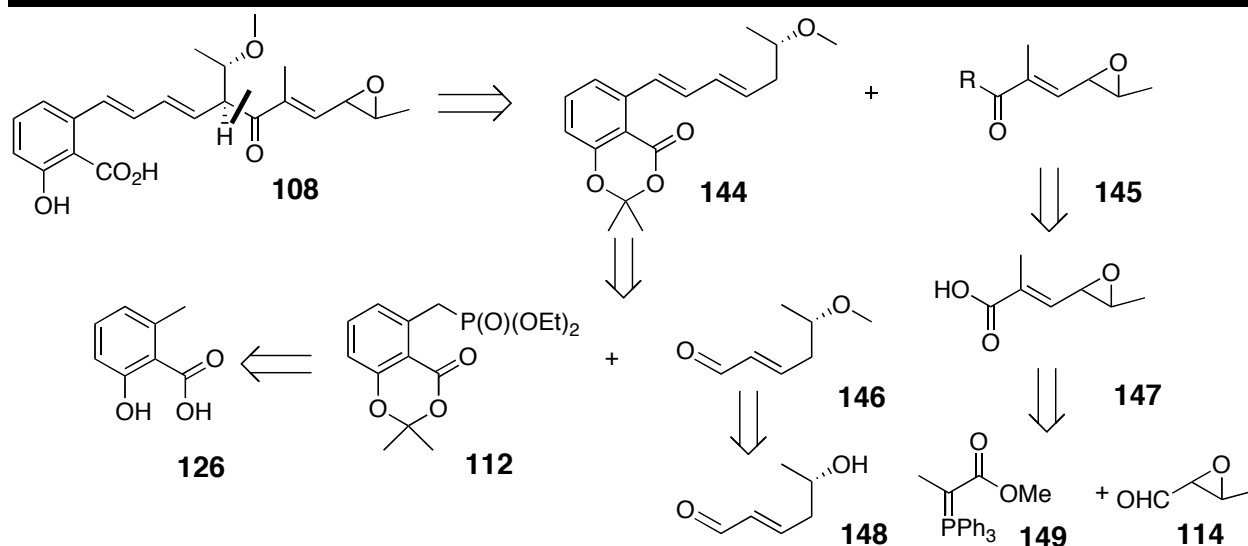
synthetic intermediate **143** could be obtained only in too low amounts for the further reactions overall steps. From **136** to **143**, 9 steps were required for synthesis. As a consequence, although a transformation of **135** into the corresponding isomer of **136** by an alternative strategy in principle would be possible, the overall reaction sequence seems very uneconomic and an alternative strategy for the synthesis of GNB was sought.



Scheme 30. Test reactions for final part of the synthesis: a. LiBH_4 , Ether/Methanol, $0\text{ }^\circ\text{C}$ – rt, 3 h, 93%; b. DHP, PPTS, DCM, rt, 5 h, 80%; c. TBAF, THF, rt, 3 h, 85%; d. DMP, DCM, $0\text{ }^\circ\text{C}$, 1 h, 90%; e. *t*-BuOK, **112**, THF, $0\text{ }^\circ\text{C}$ – rt, 1 h, 57%; f. PPTS, MeOH, rt, 2 d, 66%; g. DMP, DCM, $0\text{ }^\circ\text{C}$, 1 h, 86%; h. $\text{EtPO}(\text{OEt})_2$, *n*-BuLi, THF, $-78\text{ }^\circ\text{C}$, 1 h, 23%; i. DMP, DCM, $0\text{ }^\circ\text{C}$, 1 h, 50%.

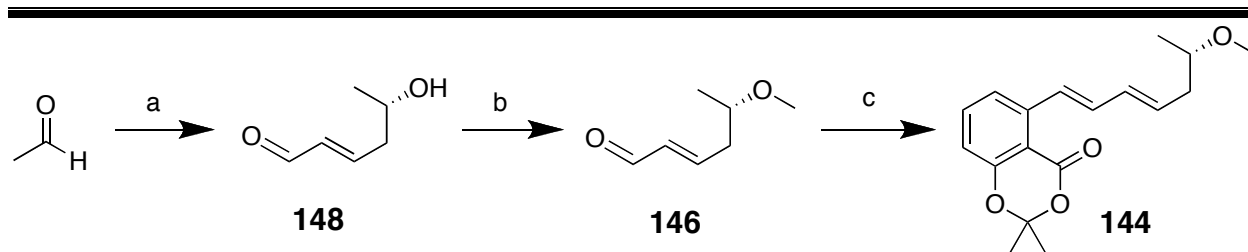
3.7.4 Second synthesis approach for GNB

Retro-synthetic analysis of second synthesis approach



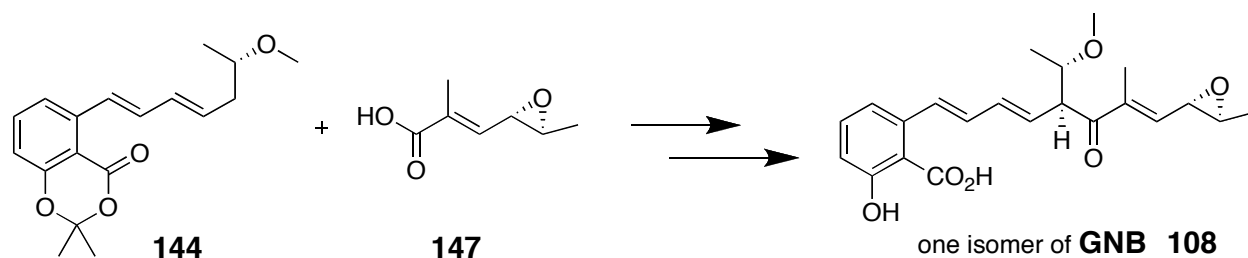
Scheme 31. Third convergent retro-synthetic route

The third convergent retro-synthesis developed within this thesis is based on a division of GNB into two parts *via* the C12-C13 bond (Scheme 31). Although the generation of a carbanion at C12 of **144** as a prerequisite for a nucleophilic substitution and a subsequent stereoselective substitution at an activated carbon acid moiety surely is a difficult task, the suggested synthesis route is much more straightforward than previous ones. Thus, this reaction requires careful optimization of reaction conditions. As the complementary reaction partner of the carbanion, a suitable leaving group at the C13 moiety **145** was carefully discussed. Several potential groups could react with the generated carbanion in the desired manner, e.g. an acyl halide or active anhydride, which all are available from a carboxylic acid derivative. Thus, a HWE reaction on the former **114** building block could generate such a carboxylic acid. As for **144**, it could also be obtained from the former **112** building block and a new C9 conjugated aldehyde building block **146** by a HWE reaction. **146** could then be prepared within 2 steps from acetaldehyde by a proline-catalyzed one pot asymmetric synthesis of 5-hydroxy-(2*E*)-hexenal (Cordova *et al.*, 2002). It is obviously that the third version of the



Scheme 33. Synthesis of building block **144**: a. *L*-Proline, THF, 4 °C, 14 h; b. TfOMe, 2,6-di-(*t*-butyl)-4-methylpyridine, DCM, rt, overnight, 90%; c. *t*-BuOK, **112**, THF, 0 °C – rt, 1 h, 80%.

After obtaining the building blocks of **144** and **147**, the completion of the synthesis of GNB is very close. Deprotonation studies on **144** revealed *n*-BuLi as most efficient. The further synthesis is anticipated as following: **147** will be transformed into an active anhydride, which is added to the carbanion of **144** to generate a protected GNB derivative. Complete deprotection then delivers GNB (Scheme 34). Although the critical coupling reaction might not occur with high ee ratio, this route is still more efficient due to the in comparison to previous trials highly shortend reaction sequence.



Scheme 34. Future synthetic plan for the synthesis of GNB

3.7.5 Discussion

We hope that with this fourth strategy, the total synthesis of GNB is very close to the end. This strategy for the first time does not rely on an auxiliary-based aldol reaction that strongly shortens the overall synthetic steps as no introduction and cleavage of the auxiliary is necessary. Instead, this method relies on an efficient organocatalytic transformation that introduces enantioselectively a

required first stereocenter that will be used to direct the outcome of subsequent reactions.

We did not accomplish the synthesis yet. However, what we have learned *via* performing the synthesis is almost more meaningful than the synthesis itself. To study the synthesis is equally important as to complete the synthesis. The total synthesis of GNB will finish in soon future within few steps.

4 Summary (English)

The present thesis includes the study and development of suitable small molecules for ABPP. Within this thesis, 4 different types of ABP design strategies were studied and 6 types of ABPs collections were generated. To this end, around 150 organic compounds were characterised as reaction intermediates or products; 4 proteomes from botanic, bacterial and mammalian species were profiled *in vitro* or *in vivo* respectively; around 200 1D-SDS-PAGE and manifold MS proteomics analyses were performed; several targets of the designed probes have been identified and the associated genetic work of cloning and *in planta* protein expression has also been completed.

The summary of the whole PhD work is separated in three parts – the first chemical, the second chemical biological, and the third biological – in order to facilitate a more detailed discussion.

4.1 Chemical part

3 probes and 2 inhibitors with an *anti*- β -lactone warhead (chapter 3.1), 9 probes and 8 inhibitors with a *syn*- β -lactone warhead (chapter 3.2), 2 probes with an azirine warhead (chapter 3.3), 13 probes and 1 inhibitor with an aziridine warhead (chapter 3.3), 6 probes and 1 inhibitor with an AEBSF warhead (chapter 3.4) and finally 18 probes with an AOMK warhead (chapter 3.5 and 3.6) were generated. Of these probes, 3 were rhodamine tagged fluorescent probes, 16 were biotin tagged affinity probes and 27 were azide or acetylene tagged chemical orthogonal probes.

For generation of the *anti*- β -lactone compounds, a Masamune auxiliary based synthetic strategy was used and delivered the desired β -lactones with a high enantioselection and yield (chapter 3.1). With this methodology, probes with variant chemical moieties were generated. However, in later synthetic approaches (chapter 3.7), this methodology proved incompatible for variant aldol reactions such as a 1,3 or 1,5-shifted aldol reaction, probably due to lower reactivity and stability of the ester-based Masamune aldol reaction partner than the amide-based Evans aldol reaction partner.

The cyclization of serine or threonine to form *syn*- β -lactone compounds followed by peptide couplings proved as a direct and convenient method to generate a *syn*- β -lactone-based probe collection in a short time (chapter 3.2).

The Aza-Darzens asymmetric synthesis was demonstrated to be a versatile method to generate aziridine or azirine compounds. The reaction proved as compatible with many chemical moieties and application of the commercially available chiral menthyl-*p*-tolylsulfinate auxiliary led to a good enantioselection of the sulfinyl imine intermediate. Consequently, this methodology enabled the generation of a complete probe library (chapter 3.3) with yields ranging from 30% to 90%.

Also AEBSF-derived probes featuring a sulfonyl fluoride moiety were synthetically straightforward available, although reaction yields were only modest due to hydrolysis under peptide coupling or purification conditions (chapter 3.4).

The AOMK derivatives were rapidly and conveniently accessible by SPPS, enabling the generation of several AOMK-based peptide probes (chapter 3.5 and 3.6).

Finally, the total synthesis of GNB was studied intensively. To date, four different synthetic strategies were evaluated, two of them within this thesis (chapter 3.7). To this end, around 40 reaction intermediates were prepared. Each re-

evaluation of the synthesis strategy improved the former one, resulting in a more efficient synthesis regarding protecting group strategy or the number of intermediates. The latest retro-synthesis involved an allyl anion nucleophilic substitution based strategy, replacing the previously used aldol reaction-based strategy, thereby shortening the required synthetic steps significantly. The proline-catalyzed one pot asymmetric synthesis of 5-hydroxy-(2*E*)-hexenal represents a short and efficient way to prepare the required building block. Although the total synthesis is not yet completed, the required building blocks were fully characterised, requiring now the final assembly. However, the employed strategy seems very promising and therefore finalisation of the synthesis should be possible in the near future.

4.2 Chemical biological part

Four ABP design strategies being direct, non-directed, natural product based and mechanism based probe design were studied within this thesis. The natural product based probe design strategy was used to obtain two THL-based ABPs. However, the targets of these two ABPs proved different from the expected lipase targets. One reason for this might be the modification of the peptide moiety on the side chain of THL, indicating that even subtle chemical modifications can change the target of a natural product. For future probe designs, it seems preferably to keep the synthesized probe as similar as possible to the natural product, although this burdens the required chemical synthesis.

With the mechanism-based probe design strategy, 9 *syn*- β -lactone-based probes and 18 AOMK-based probes were obtained. Interestingly, the 9 *syn*- β -lactone-based probes finally did not prove as “classical” ABPs. Instead, a transpeptidase activity of a previously known cysteine protease RD21 was discovered (chapter 3.2). The 18 AOMK-based probes for AvrPphB and VPE were

demonstrated to act as ABPs, labelling their target enzymes very specifically (chapter 3.5 and 3.6). These “opposite” results from the two types of probes evaluated during this thesis indicate that “classical” mechanism-based probes can be obtained only if the reacting warhead irreversibly binds the target enzyme while a “transiently” formed covalent bond can lead to unexpected results.

Following the direct probe design strategy, 6 AEBSF-based probes were generated. Unexpectedly, these probes proved not as ABPs (chapter 3.4). One reason for this strange finding could be that the additional linker and reporter tag change the properties of the serine protease inhibitor AEBSF. This indicates that a direct transformation of an inhibitor into an ABP is not trivial. An alternative explanation could be that AEBSF contrary to previous reports is not a selective active site inhibitor.

The non-directed probe design strategy was used to obtain 13 aziridine and 2 azirine-based probes. However, the evaluation of their labelling potential proved as difficult and it seems that they cannot be used as valuable probes (chapter 3.3). To date, there is only a limited number of non-directed probes available, probably as a result of difficult probe design. The results of this thesis suggest that indeed chemical reactivity of the synthesized probes is a key point in non-directed probe design as the generated probes have to be in a very narrow reactivity window, being reactive enough to bind to active site residues but being inactive enough to prevent unspecific labelling. Consequently, non-directed probe design strategies still need further improvement.

Regarding the specificity of the reactive warheads towards certain enzyme classes, AOMKs proved as most specific, targeting only cysteine proteases as reported in the literature (chapter 3.5 and 3.6). β -Lactones proved as a potential new warhead class, targeting many classes of hydrolases. Consequently, in order to

obtain specific β -lactone based probes, chemical modifications e.g. of the linker moiety (chapter 3.1 and 3.2) are necessary. Finally, AEBSF, aziridine and azirine-derived probes were demonstrated to be too active for specific labelling, thereby generally limiting their applicability for the development of ABPs (chapter 3.3 and 3.4).

In addition, 3 types of tags were studied. Biotin proved as a good affinity tag not only for detection but also specially for purification of the targets *in vitro*. However, the poor solubility and membrane permeability of biotin-tagged ABPs severely limited their use in ABPP applications.

Fluorescent tags proved as very versatile and easy to apply in many labelling studies, nevertheless application is limited due to the incapability to purify the labelled targets.

Click chemistry-based chemically orthogonal tags proved as an elegant way to balance the advantages and disadvantages of affinity and fluorescent tags. The targets of interest could be labelled first either *in vivo* or *in vitro*, and only then an affinity or fluorescent tag is coupled. The only disadvantage of this method is the required click chemistry-mediated reaction, in presence of copper, which could cause protein aggregation.

Studies of different linker system have been performed to investigate the influence of labelling selectivity. The results showed that small changes of linker composition such as the addition of one glycine moiety is able to cause a different labelling pattern, highlighting the often-neglected role of linkers during probe design (chapter 3.4). To date, there are no general rules to evaluate proper linker design and further attempts in this direction are required in future.

Profiling experiments in this thesis were performed exclusively with 1D protein gels, which proved as very convenient for evaluating profiling efficiency and to quickly display a chemical proteome map. However, the proteins of interest could be studied further more by 2D gel profiling. The target identification of probes could be carried out *via* gel-based or gel-free strategies.

4.3 Biological part

The biological part of this thesis focuses on the roles of enzymes for plant-pathogen interactions. Several literature reports highlight the roles of hydrolytic enzymes during plant defence processes. To monitor the changes of the activity states of these enzymes in normal, susceptible or diseased plant proteomes, several probe types were developed as described previously. While some probes failed to deliver suitable labelling properties, the *syn*- β -lactone protease probe could be used successfully to elucidate a transpeptidase mechanism of the *Arabidopsis thaliana* PLCP RD21 (chapter 3.2). The physiological meaning of this PTM will be studied further in future.

Cloning of the *Pseudomonas syringae* PLCP AvrPphB and labelling experiments of this avirulence protein with an AOMK-based probe revealed several important properties of this enzyme. Especially the pH dependent activity of AvrPphB seems related with its physiological function in TTSS (chapter 3.5).

Finally, the AOMK-based VPE probe has also been used to profile the activity of VPE in wild or mutant *Arabidopsis* strains (chapter 3.6). However, further studies on the implications of the observed different activity states are still required.

5 Zusammenfassung (German)

Die vorliegende Doktorarbeit beschäftigt sich mit dem Studium und der Entwicklung geeigneter kleiner Moleküle zum ABPPs. Hierzu wurden vier Typen von ABPs entwickelt und sechs Substanzsammlungen generiert. Es wurden 150 organische Verbindungen, entweder Zwischen- oder Endprodukte, charakterisiert; Ein Profiling von vier Proteomen, aus Pflanzen, Bakterien und Säugetieren, sowohl *in vitro*, als auch *in vivo*, erstellt; ungefähr 200 1D-SDS-PAGE und eine große Anzahl an MS-proteomischen Analysen wurden durchgeführt; einige der Zielproteine konnten durch die entwickelten Sonden identifiziert werden; zusätzlich konnte die dazugehörige genetische Arbeit der Klonierung und der *in planta* Proteinexpression abgeschlossen werden.

Die Zusammenfassung der gesamten Doktorarbeit ist in drei Abschnitte unterteilt, um eine detailliertere Diskussion zu ermöglichen. Diese sind ein chemischer, ein chemisch-biologischer und ein biologischen Teil.

5.1 Chemischer Teil

Es wurde eine Vielzahl an Sonden und Inhibitoren mit verschiedenen reaktiven Gruppen synthetisiert, d.h. es wurden mit *anti*- β -Lactonen 3 Sonden und 2 Inhibitoren (Kapitel 3.1), mit *syn*- β -Lactonen 9 Sonden und 8 Inhibitoren (Kapitel 3.2), mit Azirinen 2 Sonden (Kapitel 3.3), mit Aziridin 13 Sonden und 1 Inhibitor (Kapitel 3.3), mit AEBSF 6 Sonden und 1 Inhibitor (Kapitel 3.4), mit AOMK 18 Sonden (Kapitel 3.5 und 3.6) hergestellt. Von diesen Verbindungen sind 3 Fluoreszenzsonden mit Rhodamin, 16 Affinitätssonden mit Biotin und 27 weitere chemisch orthogonale Sonden mit Azid oder Acetylen versehen worden.

Eine auf das Masamune-Auxiliar-basierende synthetische Strategie wurde genutzt, um die *anti*- β -Lacton-Verbindungen in hoher Enantioselektivität und Ausbeute zu generieren (Kapitel 3.1). Mit dieser Methodologie konnten Sonden mit verschiedenen chemischen Motiven generiert werden. Dennoch hat sich diese Vorgehensweise in späteren synthetischen Versuchen (Kapitel 3.7) für variierte Aldol Reaktionen als inkompatibel erwiesen, beispielsweise bei der 1,3 oder 1,5-verschobenen Aldolreaktion. Dies liegt wahrscheinlich an der niedrigeren Reaktivität und Stabilität der auf Estern basierenden Masamune-Aldolreaktionspartner, im Vergleich zu den auf Amidien basierenden Evans-Aldolreaktionspartnern.

Um eine Sammlung aus *syn*- β -Lacton-Sonden in kurzer Zeit zu generieren, erwies sich die Zyklisierung der Serine und Threonine, gefolgt von Peptidkupplungen als überzeugende Methode (Kapitel 3.2).

Die asymmetrische Aza-Darzens-Synthese hat sich als vielseitige Methode erwiesen, um Aziridine und Azirine zu synthetisieren. Dieser Mechanismus führte zu Sulfinylimin-Intermediaten in guter Enantioselektivität, ist kompatibel mit vielen chemischen Resten und erlaubt die Verwendung des chiralen, kommerziell verfügbaren Menthyl-*p*-tolylsulfinat-Auxiliars. Daher erlaubte diese Methodologie die Erstellung einer kompletten Sondensammlung (Kapitel 3.3) mit Ausbeuten von 30 bis 90%.

Auch die auf AEBSF-basierenden-Sonden besitzen ein Sulfonylfluoridrest und sind synthetisch schnell verfügbar. Allerdings waren die Ausbeuten moderat, da sowohl unter den Bedingungen bei Peptidkupplungen, als auch bei der Aufreinigung eine Hydrolyse stattfand (Kapitel 3.4).

Die AOMK-Derivate konnten schnell und einfach durch SPPS hergestellt werden. Dies ermöglichte die Synthese verschiedener peptidischer Sonden, die auf AOMK-Motiv basierten (Kapitel 3.5 und 3.6).

Abschließend wurde die Synthese von GNB ausgiebig studiert. Bisher sind vier verschiedene synthetische Strategien zur Synthese dieses Naturstoffes evaluiert worden; zwei dieser Strategien wurden in dieser Arbeit behandelt (Kapitel 3.7). Innerhalb der Doktorarbeit wurden zu diesem Zweck fast 40 Intermediate hergestellt. Jede neuerliche Evaluierung der synthetische Strategie erzielte Verbesserungen im Vergleich zu vorhergehenden Versuchen; dies führte zu einer effizienten Synthese in Bezug auf die Schutzgruppen und die Anzahl der Intermediate. Die neueste Retrosynthese basiert auf einer allylisch-anionischen, nukleophilen Substitutionsstrategie, auf diese Weise konnte die auf der Aldolreaktion basierende Retrosynthese erheblich verkürzt werden. Dies hat die Anzahl der benötigten Reaktionsschritte signifikant reduziert. Die prolinkatalysierte, asymmetrische Eintopfsynthese des 5-Hydroxy-(2*E*)-hexenal ist ein kurzer und effizienter Weg, um den benötigten Baustein zu synthetisieren. Obwohl die Totalsynthese noch nicht abgeschlossen werden konnte, sind alle Bausteine voll charakterisiert und können nun miteinander verknüpft werden. Da die verwendete Strategie äußerst vielversprechend erscheint, sollte ein Abschluss der Synthese bald erreicht werden können.

5.2 Chemisch biologischer Teil

Innerhalb dieser Doktorarbeit wurden vier Ansätze zur Generierung von ABP studiert. Diese sind das Design von direkten-, nicht-direkten-, naturstoffbasierten- und Mechanismus-basierten Sonden.

Die Naturstoffstrategie wurde genutzt, um zwei THL-basierte ABPs zu erhalten. Allerdings waren die Zielproteine dieser ABPs andere, als die der erwarteten Lipasezielstrukturen. Ein Grund für diese unterschiedlichen Zielproteine könnte die Modifikation des peptidischen Restes der THL-Seitenkette sein. Dies ist ein Anzeichen, dass selbst geringste chemische Veränderungen die Zielproteine

von Naturstoffen ändern können. Bei zukünftigen Planungen von Sonden sollten daher die synthetisierten Verbindungen dem Naturstoff so ähnlich wie möglich sein, auch wenn dies die Synthese erschwert.

Mittels Mechanismus-basiertem Sondendesign konnten 9 *syn*- β -Lactone und 18 AOMK-basierte Sonden erhalten werden. Interessanterweise erwiesen sich die 9 *syn*- β -Lactone nicht als "klassische" ABPs, stattdessen erlaubten sie die Entdeckung der Transpeptidaseaktivität der bekannten Cysteinprotease RD21 (Kapitel 3.2). Es konnte gezeigt werden, dass die 18-AOMK basierten Sonden für AvrPphB und VPE als ABPs fungieren und ihre Zielenzyme sehr spezifisch markieren (Kapitel 3.5 und 3.6). Die gegensätzlichen Ergebnisse der zwei Sondentypen geben Hinweise, dass "klassische" auf mechanismus basierende Sonden nur erhalten werden können, wenn die reaktive Gruppe irreversibel am Zielenzym bindet, während eine "vorübergehende", kovalente Bindung zu unerwarteten Resultaten führen kann.

Dem direkten Sondendesign folgend wurden 6 AEBSF-basierte Sonden generiert. Diese Sonden erwiesen sich überraschender Weise nicht als ABPs (Kapitel 3.4). Dies könnte an dem zusätzlichen Linker und Reportertag liegen, der die Eigenschaften des Serinproteaseinhibitors AEBSF verändert. Dies deutet an, dass eine direkte Transformation eines Inhibitors in ein ABP nicht trivial ist. Außerdem wäre es möglich, dass entgegen bisherigen Erkenntnissen, AEBSF kein selektiver Inhibitor aktiver Zentren ist.

Das nicht-direkte Sondendesign wurde genutzt, um 13 auf Aziridine und 2 auf Azirine basierende Sonden zu synthetisieren. Allerdings hat sich die Evaluierung dieser Verbindungen als schwierig erwiesen und es scheint als könnten diese nicht als Sonden genutzt werden (Kapitel 3.3). Bis heute ist lediglich eine begrenzte Anzahl an nicht-direkten Sonden verfügbar, wahrscheinlich aufgrund der Schwierigkeiten beim Design der Sonden. Die Ergebnisse dieser Dissertation

weisen darauf hin, dass die chemische Reaktivität der synthetisierten Sonden absolut essentiell ist. Die Sonden haben ein enges Reaktivitätsfenster, sie müssen so reaktiv sein, dass sie am aktiven Zentrum binden, dürfen allerdings nicht so reaktiv sein, dass es zu unspezifischen Markierungen kommen kann. Nicht-direktes Sondendesign erfordert daher weitere Verbesserungen.

In Bezug auf die Spezifität der reaktiven Gruppen zu bestimmten Enzymklassen haben sich AOMKs als am geeignetsten erwiesen. Wie in der Literatur angegeben, sind diese spezifisch gegenüber Cysteinproteasen (Kapitel 3.5 und 3.6). β -Lactone haben sich als neue Klasse für reaktive Gruppen erwiesen, da sie mit vielen Klassen der Hydrolasen interagieren. Um jedoch spezifische Sonden auf Basis von β -Lactonen zu erhalten, sind chemische Modifikationen notwendig, Z.B. des Linkers (Kapitel 3.1 und 3.2). Abschließend hat es sich gezeigt, dass Sonden abgeleitet von AEBSF, Aziridinen und Azirinen zu reaktiv sind, um als spezifische Markierungen eingesetzt werden zu können; dies limitiert die Entwicklung und die Applikation dieser Verbindungen für ABPs generell (Kapitel 3.3 und 3.4).

Zusätzlich wurden drei Typen von Tags untersucht. Biotin hat sich als geeignetes Tag sowohl zur Detektion, als auch speziell zur Affinitätsaufreinigung der Zielproteine *in vitro* erwiesen. Die Nutzung von Biotin für ABPP-Anwendungen ist jedoch aufgrund der schlechten Löslichkeit und Membranpermeabilität begrenzt.

Fluoreszierende Sonden haben sich als vielseitig einsetzbar erwiesen und konnten in vielen Markierungsstudien verwendet werden. Dennoch ist die Anwendung eingeschränkt, da markierte Proteome nicht aufgereinigt werden können.

Auf Click-Chemie basierende, orthogonale Tags haben sich als vielseitige Alternative erwiesen. Sie besitzen eine gute Balance zwischen den Vor- und Nachteilen der Tags, die auf Affinität oder Fluoreszenz basieren. Interessante Zielstrukturen konnten zunächst entweder *in vivo* oder *in vitro* markiert werden und dann in Abhängigkeit der jeweiligen Umstände mit einem Affinitäts- oder Fluoreszenz Tag versehen werden. Der einzige Nachteil dieser Methode ist, dass die Click-Chemie-Reaktion Kupfer benötigt, welches Proteinaggregation verursachen kann.

Des Weiteren wurden Untersuchungen angestellt, um die Bedeutung verschiedener linker in Bezug auf die Selektivität der Markierungen aufzuklären. Die Ergebnisse haben gezeigt, dass kleine Änderungen im Aufbau des Linkers wie beispielsweise das Hinzufügen eines zusätzlichen Glycins, unterschiedliche Markierungsmuster zur Folge haben können. Dies zeigt, wie wichtig bei der Planung von Sonden die oft vernachlässigte Rolle des Linkers ist (Kapitel 3.4). Bis heute gibt es keine generellen Erkenntnisse zur Evaluierung des Designs von Linkern und künftig sind weitere Anstrengungen in diese Richtung notwendig.

Alle Profiling-Experimente in dieser Doktorarbeit wurden ausschließlich mit 1D-Proteingelen durchgeführt. Diese eignen sich sehr gut, um die Effizienz des Profiling zu evaluieren und zur schnellen Kartographierung von Proteomen. Ein 2D-Gelprofil könnte für ein gründlicheres Studium interessanter Proteine unter Umständen notwendig sein. Eine Identifikation der Zielstrukturen der Sonden könnte durch auf Gele basierende oder gelfreie Strategien durchgeführt werden.

5.3 Biologischer Teil

Der biologische Teil dieser Doktorarbeit fokussiert sich auf die Rolle von Enzymen bei der Pflanzen-Pathogen Interaktion. Verschiedene Literaturquellen stellen die Rolle von hydrolytischen Enzymen bei Abwehrprozessen der Pflanzen heraus; um die Veränderungen der Aktivität dieser Enzyme in normalen, empfänglichen oder erkrankten Pflanzenproteomen zu überwachen, sind die zuvor erwähnten, unterschiedlichen Sondentypen entwickelt worden. Während einige Sonden nicht die notwendigen Eigenschaften zur Markierung aufwiesen, war es jedoch möglich die *syn*- β -Lacton-Sonde erfolgreich zu nutzen, um den Mechanismus der Transpeptidase von *Arabidopsis thaliana* PLCP RD21 (Kapitel 3.2) zu studieren. Die physiologische Bedeutung dieser PTM wird in darauffolgenden Studien weiter untersucht werden.

Die Klonierung der *Pseudomonas syringae* PLCP AvrPphB, sowie Markierungsexperimente dieses avirulenten Proteins mit AOMK-basierten Sonden haben einige wichtige Eigenschaften dieses Proteins offenbart. Insbesondere die pH-abhängige Aktivität von AvrPphB, welche anscheinend mit ihrer physiologischen Funktion im TTSS zusammenhängt (Kapitel 3.5) ist zu erwähnen.

Außerdem konnte die AOMK basierte Sonde dazu genutzt werden, die Aktivität von VPE in wildtyp oder mutierten *Arabidopsis*-Stämmen zu untersuchen (Kapitel 3.6). Dennoch sind weitere Untersuchungen in Bezug auf die Implikationen der beobachteten verschiedenen Aktivitätsstadien notwendig.

6 Experimental section

6.1 Chemical part

6.1.1 Instruments and Reagents

NMR Spectroscopy

The NMR spectra were measured with the following machines

Varian Mercury VX-400: 400 MHz ^1H -NMR and 100.5 MHz ^{13}C -NMR

Bruker DRX 500: 500 MHz ^1H -NMR and 125.7 MHz ^{13}C -NMR

Chemical shifts are expressed in parts per million (ppm) from internal deuterated solvent standard (CDCl_3): $\delta_{\text{H}} = 7.27$ ppm, $\delta_{\text{C}} = 77.16$ ppm; CD_3OD : $\delta_{\text{H}} = 4.85$ ppm, $\delta_{\text{C}} = 49.00$ ppm; DMSO : $\delta_{\text{H}} = 2.48$ ppm, $\delta_{\text{C}} = 39.43$ ppm; CD_3CN : $\delta_{\text{H}} = 1.94$ ppm, $\delta_{\text{C}} = 1.24$ ppm; CD_3COCD_3 : $\delta_{\text{H}} = 2.04$ ppm, $\delta_{\text{C}} = 29.8$ ppm). Coupling constants (J) are given in Hertz (Hz) and the following notations indicate the multiplicity of the signals: s (singlet), d (doublet), t (triplet), dd (doublet of doublet), m (multiplet), br (broad signal).

Optical rotation

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

High resolution mass spectrometry

HRMS was measured by a HPLC Agilent 1100 Series with one mass spectrometer of LTQ Orbitrap and nano-electrospray for ionisation. A C18-Dionex column was applied.

Reversed phase liquid chromatography – electrospray ionisation mass spectrometry (LC-MS)

LC-MS measurements were carried out on a Agilent 1200 Series HPLC with a 150/4.6 Eclipse XDB-C18 5 μ M and a 125/4 Nucleodur C4 Gravity 5 μ M. The ESI-mass spectrometer was a LCQ Advantage MAX from Thermo.

The standard gradient is the following:

Solvent A: 0.1 % HCOOH in H₂O; solvent B: 0.1 % HCOOH in ACN; Flow rate: 1 ml/min, 1 min 10 % B, linearly increasing to 100 % B in 9 min, and 2 min with 100% B.

Melting point determination

Melting points were measured with a Buechi melting point B-540 apparatus with open capillary (uncorrected values were reported).

Fourier transform infrared spectroscopy (FT-IR)

IR spectra were measured in Bruker vector 22 with a diffuse reflectance head A527 from Spectra Tech (KBr as a matrix) and a Bruker tensor 27 spectrometer with transmission and attenuated total reflection (ATR) and coupled with infrared

microscope from Spectra Tech (neat). The following notations indicate the intensity of the absorption bands: s = strong, m = middle, w = weak, b = broad.

Thin layer chromatography (TLC)

TLCs were carried out on Silica gel 60 F₂₅₄ from Merck with ultraviolet light irradiation at 254/360 nm or the following staining agents:

Staining solution A: Molybdato-phosphoric acid (20g) in ethanol (100 ml);

Staining solution B: KMnO₄ (5 g), NaOH solution (100 mg) in H₂O (100 ml);

Staining solution C: Ninhydrin (100 mg) in ethanol (50 ml) and acetic acid (1 ml).

Preparative reverse phase high performance liquid chromatography (prep HPLC)

Purification of compounds was performed on an LC-8A from Shimadzu with VP 125/21 Nucleodur C18 Gravity 5 µm column from Macherey-Nagel.

The standard gradient is the following:

Solvent A: 0.1 % TFA in H₂O; solvent B: 0.1 % TFA in ACN; Flow rate: 20 ml/min.

Solvents and Reagents

The reagents and solvents were purchased from Acros Chimica, Aldrich, Fluka, Merck, Novabiochem, Riedel de Haen, Roth and were used without further purification.

Deionized water was obtained using a Millipore Q-plus System.

Anhydrous solvents and reagents such as DCM, THF, diethyl ether, toluene, ACN, methanol, DMF, DMSO and DIPEA were purchased from Aldrich, Fluka or Acros Chimica.

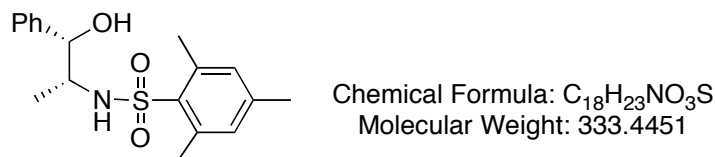
Triethylamine was distilled from CaH_2 under argon and stored with KOH.

NBS was recrystallized from water.

Phosphate buffer pH 7.0: $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ (2.9 g) and $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$ (5.15 g) were dissolved in water (50 ml).

6.1.2 Synthetic procedures

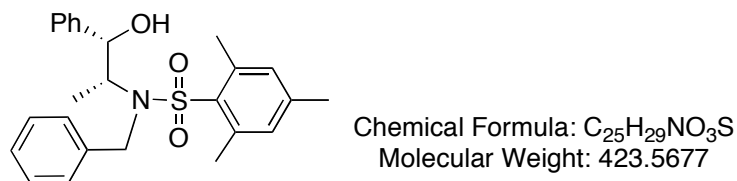
(1*S*, 2*R*)-2-(*N*-Mesitylenesulfonyl)amino-1-phenyl-1-propanol (**1**)



To a stirred solution of (1*S*,2*R*)-norephedrine (3.78 g, 25 mmol) and TEA (4.2 ml, 30 mmol) in anhydrous DCM (100 ml) was added mesitylenesulfonyl chloride (5.47g, 25 mmol) at 0 °C. The reaction mixture was stirred at 0 °C for 50 min and at room temperature for another 1.5 h. The mixture was washed with water (50 ml), 1N HCl (40 ml), water (30 ml), saturated NaHCO₃ solution (40 ml) in sequence. The organic layer was separated and concentrated in *vacuo* to afford the product **1** (yield: 8.29 g, 24.9 mmol, 99%) as a yellow oil, which was used in the next step without further purification (Recrystallization from DCM/hexane if crystal required).

TLC (MeOH/DCM = 1:9): R_f = 0.82; LC-MS (ESI): t_R = 8.95 min, calcd. for C₁₈H₂₃NO₃S [M]⁺, 333.45, found 333.78.

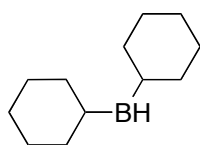
(1*S*, 2*R*)-2-(*N*-benzyl-*N*-Mesitylenesulfonyl)amino-1-phenyl-1-propanol (**2**)



A mixture of **1** (8.29 g, 24.9 mmol), benzyl bromide (3.5 ml, 30 mmol) and Cs₂CO₃ (8.25 g, 37.5 mmol) in acetonitrile (100 ml) was refluxed for 0.5 h. The cooled mixture is filtered and the salt is washed with diethyl ether (30 ml) four times. Combined organic layers were concentrated and purified by flash chromatography (ethyl acetate/cyclohexane = 1:5) to afford the product **2** (yield: 10.05 g, 23.8 mmol, 95 %) as a white powder and the side product *N,O*-dibenzyl compound (0.53 g).

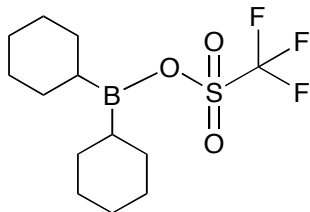
TLC (ethyl acetate/cyclohexane = 1:5): R_f = 0.63; ¹H NMR (CDCl₃) δ = 7.34-6.93 (m, 12H), 5.0 (s br, 1H), 4.77 (ABq, J = 16 Hz), 4.54 (ABq, J = 16 Hz), 3.83 (dd, J = 7.2 Hz), 2.65 (s, 6H), 2.29 (s, 3H), 2.12 (br s, 1H), 1.03 (d, J = 7.2 Hz, 3H); ¹³C NMR (CDCl₃) δ = 142.8, 142.3, 140.4, 138.8, 132.3, 128.8, 128.3, 127.9, 127.6, 127.5, 125.7, 59.9, 49.3, 48.0, 23.2, 21.1, 10.1; LC-MS (ESI): t_R = 8.95 min, calcd. for C₂₅H₂₉NO₃S [M]⁺, 423.57, found 423.86.

Dicyclohexylborane (**3**)



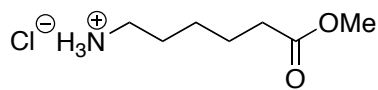
Chemical Formula: C₁₂H₂₃B
Molecular Weight: 178.1220

Cyclohexene (3.34 ml, 33 mmol) and anhydrous diethyl ether (10 ml) were charged in a 50 ml round flask and kept at 0 °C under argon. Borane-dimethyl sulphide complex (2.94 ml, 16 mmol) was added dropwise. 5 minutes after addition of the borane, a precipitate was formed. The reaction mixture was stirred for 3 h at 0 °C, and the diethyl ether was removed as much as possible by a syringe. The remained solid was dried under reduced pressure to give the product **3** (yield: 2.8 g, 15.8 mmol, 99%) as a white powder, which is used without further purification for the preparation of the triflate.

Dicyclohexylboron trifluoromethanesulfonate (4)

Chemical Formula: $C_{13}H_{22}BF_3O_3S$
Molecular Weight: 326.1832

To a suspension of **3** (2.8 g, 15.8 mmol) in anhydrous hexane (10 ml) was added dropwise *via* a glass drop funnel trifluoromethanesulfonic acid (1.42 ml, 16 mmol) under stirring, a gentle gas evolution occurred. The solid starting material gradually dissolved and after 5 min the solution turned clear and colourless. Stirring was continued at room temperature for 1.5 h. The reaction was completed when there was no gas evolution any more. A slight amount of crystalline powder formed on the bottom of the flask, which was the acid precipitated from the hexane. After filtration, the filtrate was collected. (If crystals were needed, then the solution was placed in a $-20^{\circ}C$ freezer for 24 h. Larger crystals formed and the mother liquor was transferred *via* cannula to another flask. The crystals were dried under reduced pressure.) Hexane was removed at reduced pressure to obtain the product (yield 4.6 g, 14.1 mmol, 89%) as colourless crystals. The isolated crystals could be preserved as an 1M hexane solution at $4^{\circ}C$.

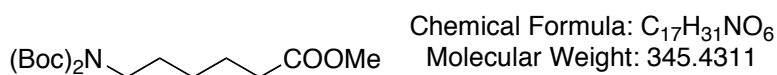
Methyl 6-aminocaproate hydrochloride (5)

Chemical Formula: $C_7H_{16}ClNO_2^+$
Molecular Weight: 181.6599

To a white suspension of 6-aminocaproic acid (2.62 g, 20 mmol) in ice-cooled anhydrous methanol (35 mL) was added dropwise thionyl chloride (2.04 mL, 28 mmol). The reaction mixture was allowed to warm to room temperature and was subsequently stirred overnight. Removal of the volatiles under reduced pressure gave methyl 6-aminocaproate hydrochloride **5** (yield: 3.63 g, 20 mmol, 100%) as a colourless crystalline solid and was used in the next step without further purification.

^1H NMR (D_2O) δ = 3.59 (s, 3H), 2.89 (t, J = 7.6 Hz, 2H), 2.32 (t, J = 7.6 Hz, 2H), 1.55 (m, 4H), 1.31 (m, 2H).

Methyl (6-di-Boc-amino)caproate (**6**)



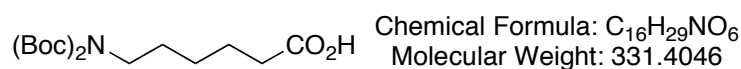
A solution of **5** (2.9 g, 20 mmol) in a mixture of dioxane (40 ml), water (20 ml) and Na₂CO₃ (2.12 g, 20 mmol) was stirred and cooled in an ice bath. Di-*tert*-butyl pyrocarbonate (4.8 g, 22 mmol) was added and stirred at room temperature for another 30 min. The solution was concentrated in *vacuo* to about 10 to 15 ml, cooled in an ice bath, covered with a layer of DCM (60 ml) and acidified with a dilute solution of KHSO₄. The organic layer was separated, washed with NaHCO₃ aqueous solution (20 ml) and brine (20 ml), dried over Na₂SO₄ and concentrated in *vacuo* to obtain the product (6-mono-bocamino)caproate (yield 4.9 g, 20 mmol, 100%) as a colourless oil.

^1H NMR (CDCl_3) δ = 4.5 (br s, 1H), 3.66 (s, 3H), 3.11 (q, J = 6.4 Hz, 2H), 2.31 (t, J = 7.2 Hz, 2H), 1.68-1.32 (m, 6H), 1.44 (s, 9H).

To a solution of above obtained methyl 6-(monobocamino)caproate (4.9 g, 20 mmol) and DMAP (220 mg, 1.8 mmol) in THF (120 ml) was added a solution of Di-*tert*-butyl pyrocarbonate (6.54 g, 30 mmol) in THF (50 ml) and the mixture was refluxed for 16 h. Additional Di-*tert*-butyl pyrocarbonate (4.36 g, 20 mmol) was added and the heating was continued for 5 h. The deep orange colour reaction mixture was cooled down, concentrated, purified by column chromatography (ethyl acetate/cyclohexane = 3:97) to get the pure product **6** (yield: 6.2 g, 18 mmol, 90%) as a light yellow oil.

TLC (ethyl acetate/cyclohexane = 3:97): R_f = 0.18; ¹H NMR (CDCl₃) δ = 3.64 (s, 3H), 3.53 (t, J = 7.5 Hz, 2H), 2.29 (t, J = 7.5 Hz, 2H), 1.66-1.53 (m, 4H), 1.48 (s, 18H), 1.35-1.28 (m, 2H); ¹³C NMR (CDCl₃) δ = 173.8, 152.6, 81.9, 51.3, 46.1, 33.9, 28.5, 27.9, 26.2, 24.5.

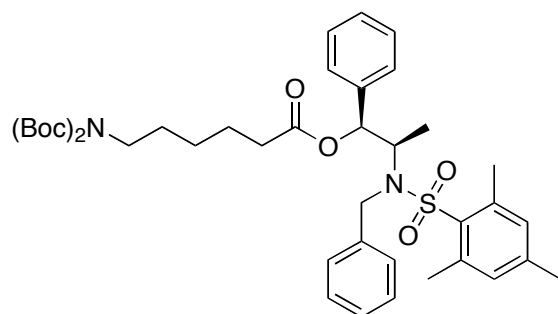
(6-Di-Boc-amino)caproic acid (**7**)



6 (1.38 g, 4 mmol) in acetonitrile (20 ml) was cooled to 0°C. *n*-Bu₄NOH × 30 H₂O (4.8 g, 6 mmol) was added under stirring (ACN/H₂O = 2.5:1 to 6:1). The reaction was complete in 15 to 30 min. The organic layer was washed with 1N HCl and water, concentrated in *vacuo* to obtain the product **7** (yield 1.32 g, 4 mmol, 100 %) as a colourless oil.

^1H NMR (CDCl_3) δ = 3.54 (t, J = 7.5 Hz, 2H), 2.35 (t, J = 7.5 Hz, 2H), 1.68-1.56 (m, 4H), 1.49 (s, 18H), 1.45-1.35 (m, 2H); ^{13}C NMR (CDCl_3) δ = 178.4, 152.6, 82.0, 46.1, 33.8, 28.5, 27.9, 26.1, 24.5.

(1*S*, 2*R*)-2-(*N*-benzyl-*N*-mesitylenesulfonyl)amino-1-phenyl-1-propanyl (6-di-Boc-amino)caproate (8**)**



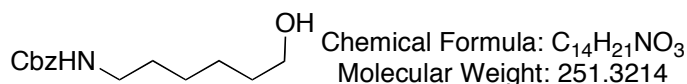
Chemical Formula: $\text{C}_{41}\text{H}_{56}\text{N}_2\text{O}_8\text{S}$
Molecular Weight: 736.9569

To a solution of **7** (662 mg, 2 mmol), DCC (412 mg, 2 mmol) and DMAP (0.01 g, 0.1 mmol) in DCM (10 ml), **2** (0.846 g, 2 mmol) was added. The reaction was stirred overnight, concentrated in *vacuo*, purified by column chromatography (ethyl acetate/cyclohexane = 1:9) to get the pure product **8** (yield: 1.4 g, 1.9 mmol, 95%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 1:9): R_f = 0.2; ^1H NMR (CDCl_3) δ = 7.33-7.29 (m, 3H), 7.23-7.18 (m, 5H), 6.93-6.87 (m, 4H), 5.82/5.00 (d/d, 1H, J = 4 Hz), 4.77/4.71 (d/d, J = 16.5 Hz, 1H), 4.59/4.54 (d/d, J = 16.5 Hz, 1H), 4.04/3.81 (q/q, J = 2.5 Hz, 1H), 3.55/3.50/3.19 (t/t/t, J = 7.5 Hz, 2H), 2.65/2.61/2.50 (s/s/s, 6H), 2.29/2.27 (s/s, 3H), 2.18-2.02 (m, 2H), 1.81-1.68 (m, 2H), 1.51/1.49 (s/s, 18H), 1.33-1.18 (m, 4H), 1.11/1.02/0.96 (d/d/d, J = 7 Hz, 3H); ^{13}C NMR (CDCl_3) δ = 171.9, 152.8, 142.6, 140.4, 138.8, 138.7, 133.5, 132.3, 128.5, 128.3, 127.9, 127.9, 127.5, 127.4, 127.2, 126.1, 125.7, 82.2, 78.1, 56.8, 48.3, 46.3, 35.1, 34.3, 28.8, 28.2,

26.4, 24.5, 23.1, 21.0, 13.0; LC-MS (ESI): $t_R = 11.95$ min, calcd. for $C_{41}H_{56}N_2O_8SNa [M+Na]^+$, 759.95, found 759.26.

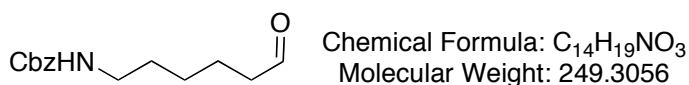
(6-Hydroxyhexyl)carbamic acid benzyl ester (**9**)



6-Amino-1-hexanol (585 mg, 5 mmol) was dissolved in dioxane (7 ml) and sodium bicarbonate saturated aqueous solution (10 ml) at room temperature. To this solution, carbobenzoxy chloride (1 g, 6 mmol) in dioxane (3 ml) was added. The reaction was complete in 3 h and extracted with DCM and washed with brine, dried over sodium sulfate, purified by column chromatography (ethyl acetate/cyclohexane = 2:1) to get the pure product **9** (yield: 1.09 g, 4.3 mmol, 87%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 2:1): $R_f = 0.44$; 1H NMR ($CDCl_3$) $\delta = 7.35-7.31$ (m, 5H), 5.09 (s, 2H), 4.74 (br s, 1H), 3.63 (t, $J = 6.4$ Hz, 3H), 3.19 (m, 2H), 1.58-1.49 (m, 4H), 1.40-1.34 (m, 4H); ^{13}C NMR ($CDCl_3$) $\delta = 156.5, 136.79, 128.66, 128.24, 66.76, 62.88, 41.05, 32.71, 30.11, 26.52, 25.44$; HPLC: $t_R = 12.18$ min.

(6-Formylhexyl)carbamic acid benzyl ester (**10**)

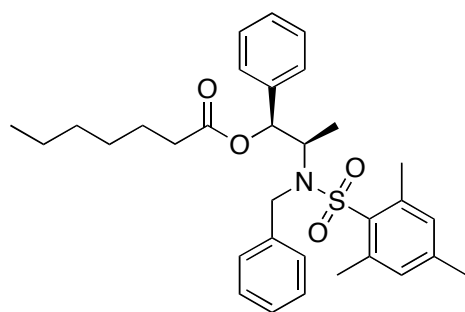


To a solution of **9** (0.502 g, 2 mmol) in THF (7.5 ml) was added a solution of IBX [0.09 g/ml (DMSO), 7.5 ml, 2.4 mmol] at 0 °C. The clear reaction solution was

then stirred at room temperature overnight. The reaction mixture was filtrated before extraction. The product was extracted with DEE, washed with brine, dried and purified by a short column chromatography (ethyl acetate) to get the pure product **10** (yield: 0.47 g, 1.9 mmol, 95%) as a colourless oil.

TLC (ethyl acetate): $R_f = 0.87$; $^1\text{H NMR}$ (CDCl_3) $\delta = 9.76$ (s, 1H), 7.37-7.32 (m, 5H), 5.10 (s, 2H), 4.77 (br s, 1H), 3.20 (m, 2H), 2.44 (t, $J = 7.24$ Hz, 3H), 1.66 (m, 2H), 1.53 (m, 2H), 1.35 (m, 2H); $^{13}\text{C NMR}$ (CDCl_3) $\delta = 202.49$, 156.5, 136.75, 128.66, 128.25, 66.77, 43.85, 40.93, 29.93, 26.34, 21.77; HPLC: $t_R = 14.04$ min.

(1S, 2R)-2-(N-benzyl-N-mesitylenesulfonyl)amino-1-phenyl-1-propanyl enanthate (11)



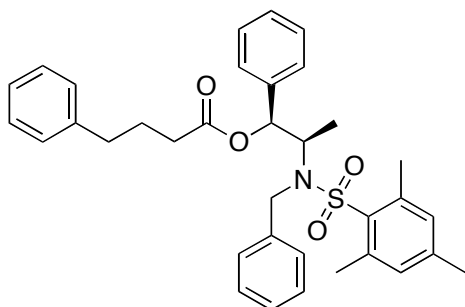
Chemical Formula: $\text{C}_{32}\text{H}_{41}\text{NO}_4\text{S}$
Molecular Weight: 535.7372

To a solution of enanthic acid (311 μl , 2.2 mmol), DCC (453 mg, 2.2 mmol) and DMAP (12 mg, 0.1 mmol) in DCM (10 ml), **2** (846 mg, 2 mmol) was added. The reaction mixture was stirred overnight, concentrated in *vacuo*, purified by column chromatography (ethyl acetate/cyclohexane = 1:9) to get the pure product **11** (yield: 0.92 g, 1.7 mmol, 86%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 1:9): $R_f = 0.33$; $^1\text{H NMR}$ (CDCl_3) $\delta = 7.34$ -7.33 (d, $J = 7$ Hz, 2H), 7.19-7.18 (m, 6H), 6.89-6.88 (m, 4H), 5.82 (d, $J = 4$ Hz, 1H), 4.74

(ABq, $J = 16.5$ Hz, 1H), 4.54 (ABq, $J = 16.5$ Hz, 1H), 4.04 (dd, $J = 2.5, 4$ Hz, 1H), 2.51 (s, 6H), 2.28 (s, 3H), 2.17-2.09 (m, 2H), 1.5-1.48 (m, 2H), 1.23-1.22 (m, 6H), 1.12 (d, $J = 7$ Hz, 3H), 0.86 (t, $J = 7$ Hz, 3H); ^{13}C NMR (CDCl_3) $\delta = 171.9, 142.4, 140.1, 138.6, 138.4, 133.3, 132.0, 128.3, 128.2, 127.7, 127.3, 127.0, 125.9, 77.8, 56.6, 48.1, 34.1, 31.3, 28.6, 24.5, 22.9, 22.3, 20.8, 13.9, 12.8$; LC-MS (ESI): $t_{\text{R}} = 11.75$ min, calcd. for $\text{C}_{32}\text{H}_{41}\text{NO}_4\text{SNa}$ $[\text{M}+\text{Na}]^+$, 558.72, found 558.26.

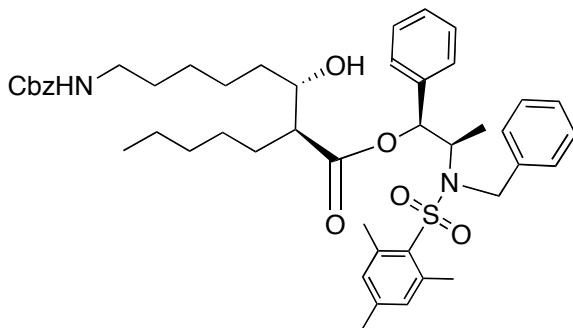
(1S, 2R)-2-(N-benzyl-N-mesitylenesulfonyl)amino-1-phenyl-1-propanyl (4-phenyl)butyrate (12)



Chemical Formula: $\text{C}_{35}\text{H}_{39}\text{NO}_4\text{S}$
Molecular Weight: 569.7535

Following the same protocol as above for **11** led to the pure product **12** (yield: 0.95 g, 1.66 mmol, 83%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 1:9): $R_f = 0.32$; ^1H NMR (CDCl_3) $\delta = 7.34-7.27$ (d, 4H), 7.24-7.18 (m, 7H), 7.11-7.10 (d, $J = 2.2$ Hz, 2H), 6.92-6.88 (m, 4H), 5.83 (d, $J = 4$ Hz, 1H), 4.74 (d, $J = 16.5$ Hz, 1H), 4.54 (d, $J = 16.5$ Hz, 1H), 4.04 (m, 1H), 2.54 (t, $J = 7.5$ Hz, 2H), 2.51 (s, 6H), 2.28 (s, 3H), 2.17-2.09 (m, 2H), 1.82 (m, 2H), 1.11 (d, $J = 7$ Hz, 3H), 0.86 (t, $J = 7$ Hz, 3H); ^{13}C NMR (CDCl_3) $\delta = 171.8, 142.6, 141.4, 140.4, 138.8, 138.7, 133.5, 132.3, 128.6, 128.5, 128.4, 127.5, 127.2, 126.2, 126.1, 78.1, 56.8, 48.3, 35.1, 33.5, 27.1, 26.2, 23.2, 21.0, 13.0$; LC-MS (ESI): $t_{\text{R}} = 11.43$ min, calcd. for $\text{C}_{35}\text{H}_{39}\text{NO}_4\text{SNa}$ $[\text{M}+\text{Na}]^+$, 592.74, found 592.23.

Anti- β -(S)-Hydroxyl- α -(S)-pentyl-(8-Cbz-amino)caprylic Masamune auxiliary ester (13)


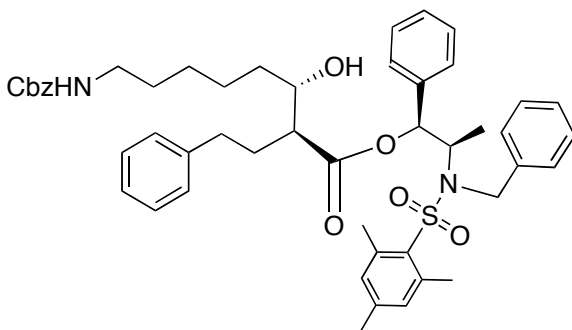
Chemical Formula: C₄₆H₆₀N₂O₇S
Molecular Weight: 785.0428

To a solution of **11** (474 mg, 884 μ mol) in anhydrous DCM (15 ml) at -78 °C was added TEA (525 μ l, 3.8 mmol) and **4** (1 M in hexane, 1.77 ml). After 2 h at -78 °C, **10** (264 mg, 1.06 mmol) in anhydrous DCM (3 ml) was added dropwise. The reaction was stirred at -78 °C for 1 h and at 0 °C for another 1 h before quenched by adding pH 7 buffer (10 ml), MeOH (10 ml) and 30% H₂O₂ (10 ml). The heterogeneous mixture was stirred vigorously for 12 h and extracted with DCM thoroughly. The combined organic layers were dried over Na₂SO₄, concentrated in *vacuo*, purified by column chromatography (ethyl acetate/cyclohexane = 1:2) to get the pure product **13** (yield: 0.614 g, 0.78 mmol, 88%) as a colourless oil.

$[\alpha]_D^{20} = -20.8$ (c = 4.38, CDCl₃); TLC (ethyl acetate/cyclohexane = 1:2): R_f = 0.5; ¹H NMR (CDCl₃) δ = 7.33-7.11, 6.84 (m, 17H, Ar), 5.82 (d, J = 5.64 Hz, 1H), 5.09 (s, 2H), 4.74 (ABq, J = 16.2 Hz, 1H), 4.70 (m, 1H), 4.49 (ABq, J = 16.2 Hz, 1H), 4.16 (m, 1H), 3.62 (m, 1H), 3.16 (m, 2H), 2.42 (s/s, 6H/1H), 2.27 (s, 3H), 1.90 (m, 2H), 1.73 (m, 2H), 1.49-1.02 (m, 25H), 1.13 (d, J = 2.16 Hz, 3H), 0.78 (t, J = 6.8 Hz, 3H); ¹³C NMR (CDCl₃) δ = 174.7, 156.5, 149.5, 142.6, 140.5, 138.5, 138.0, 136.8, 136.7, 133.3, 128.6, 128.5, 128.2, 128.0, 127.8, 127.4, 127.1, 126.8, 78.2, 72.1, 70.4, 66.7, 56.5, 51.5, 48.2, 41.0, 35.7, 35.2, 31.8, 30.1, 29.7, 26.9, 26.7, 25.6,

25.3, 24.3, 23.0, 22.5, 21.0, 14.6, 14.0; LC-MS (ESI): $t_R = 11.63$ min, calcd. for $C_{46}H_{60}N_2O_7SNa [M+Na]^+$, 807.04, found 807.21.

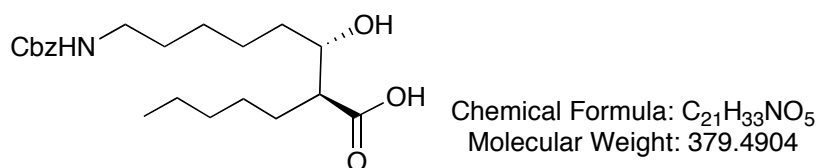
***Anti*- β -(*S*)-hydroxyl- α -(*S*)-(2-phenylethyl)-(8-Cbzamino)caprylic masamune auxiliary ester (14)**



Chemical Formula: $C_{49}H_{58}N_2O_7S$
Molecular Weight: 819.0590

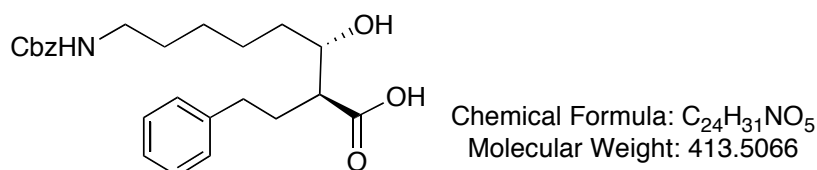
Following the same protocol as above for **13** led to the pure product **14** (yield: 0.475 g, 0.58 mmol, 70%) as a colourless oil.

$[\alpha]_D^{20} = -27.6$ ($c = 4.5$, $CDCl_3$); TLC (ethyl acetate/cyclohexane = 1:2): $R_f = 0.31$; 1H NMR ($CDCl_3$) $\delta = 7.35-7.11$, 6.84 (m, 22H), 5.82 (d, $J = 5.64$ Hz, 1H), 5.09 (s, 2H), 4.74 (ABq, $J = 16.2$ Hz, 1H), 4.70 (m, 1H), 4.49 (ABq, $J = 16.2$ Hz, 1H), 4.16 (m, 1H), 3.62 (m, 1H), 3.16 (m, 2H), 2.45 (m, 1H), 2.42 (s, 6H), 2.27 (s, 3H), 2.05 (s, 1H), 1.87 (m, 2H), 1.32-1.02 (m, 10H), 1.13 (d, $J = 2.16$ Hz, 3H); ^{13}C NMR ($CDCl_3$) $\delta = 174.2$, 156.6, 142.8, 142.7, 141.5, 141.3, 140.6, 138.4, 138.0, 136.9, 133.3, 132.3, 128.7, 128.68, 128.66, 128.62, 128.59, 128.56, 128.51, 128.46, 128.42, 128.31, 128.21, 127.99, 127.5, 127.1, 127.0, 126.9, 126.3, 78.4, 77.4, 72.1, 66.8, 56.5, 51.2, 50.9, 41.1, 35.2, 33.4, 31.4, 27.2, 26.7, 26.6, 23.1, 21.1, 15.1, 14.8; LC-MS (ESI): $t_R = 17.51$ min, calcd. for $C_{49}H_{58}N_2O_7SNa [M+Na]^+$, 842.04, found 841.34.

Anti- β -(S)-hydroxyl- α -(S)-pentyl-(8-Cbzamino)caprylic acid (15)

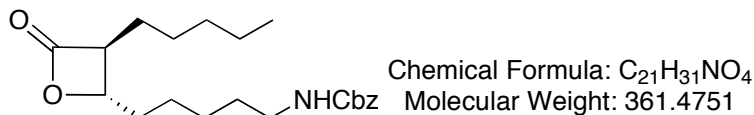
To a solution of **13** (395 mg, 0.5 mmol) in a THF/water mixture (2:1, 5 ml) was added LiOH × H₂O (420 mg, 10 mmol). The reaction mixture was stirred at room temperature for 3 days and poured into water (10 ml) and extracted with ether to recover the auxiliary. The aqueous layer was acidified (pH 3) with 1N HCl and extracted with ether. The organic layer was carefully washed with brine, dried, and the solvent was removed at reduced pressure. The product **15** was nearly pure without additional purification (yield: 0.141 g, 0.37 mmol, 74.6%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 1:2): R_f = 0.11; ¹H NMR (CDCl₃) δ = 7.36-7.31 (m, 5H), 5.09 (s, 2H), 4.78 (s br, 1H), 3.70 (m, 1H), 3.17 (m, 2H), 2.46 (m, 1H), 1.73-1.30 (m, 16H), 0.89 (t, J = 6.44 Hz, 3H); ¹³C NMR (CDCl₃) δ = 182.96, 156.87, 134.76, 131.41, 128.76, 128.75, 128.36, 72.00, 52.77, 51.00, 45.8, 43.44, 38.41, 32.51, 31.9, 29.7, 27.2, 26.4, 22.66, 22.65, 14.21.

Anti-β-(S)-hydroxyl-α-(S)-(2-phenylethyl)-(8-Cbzamino)caprylic acid (16)

Following the same protocol as above for **15** led to the pure product **16** (yield: 25 mg, 0.06 mmol, 71%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 1:2): R_f = 0.08; ¹H NMR (CDCl₃) δ = 7.36-7.13 (m, 10H), 5.09 (s, 2H), 4.82 (s br, 1H), 3.79/3.67 (m, 1H), 3.10 (m, 2H), 2.74-2.43 (m, 3H), 2.03-1.99 (m, 1H), 1.86-1.79 (m, 1H), 1.42-1.23 (m, 9H); ¹³C NMR (CDCl₃) δ = 174.61, 156.89, 148.64, 141.72, 141.53, 136.78, 128.76, 128.72, 128.66, 128.64, 128.35, 126.3, 126.27, 72.23, 72.09, 66.94, 51.04, 50.4, 41.01, 34.1, 34.0, 33.76, 31.5, 30.57, 30.02, 26.55, 26.48, 25.54.

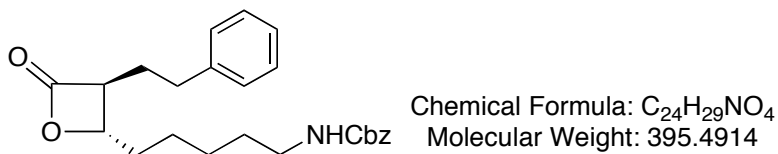
(3S, 4S)-3-pentyl-4-[(5-cbzamino)pentyl]-2-oxetanone (17)

To a cooled and stirred solution of **15** (54.2 mg, 143 μmol) in anhydrous pyridine (5.15 mL), 2-nitrobenzenesulfonyl chloride (63.2 mg, 286 μmol) in pyridine (1.2 mL) was added dropwise *via* a syringe at 0 °C. The light yellow solution was stirred for two days at 0 °C and diluted with ether (100 mL). Water (30 mL) was added, and the aqueous layer was extracted with ether (3 x 10 mL). The combined organic layers were washed once with water (20 mL) and dried over anhydrous Na₂SO₄,

concentrated in *vacuo*, purified by column chromatography (ethyl acetate/cyclohexane = 1:2) to get the pure product **17** (yield: 35.3 mg, 0.1 mmol, 70%) as a colourless oil.

$[\alpha]_D^{20} = -18.0$ ($c = 3.5$, CD_3CN); TLC (ethyl acetate/cyclohexane = 1:2): $R_f = 0.51$; $^1\text{H NMR}$ (CD_3CN) $\delta = 7.33\text{-}7.31$ (m, 5H), 5.56 (s br, 1H), 5.09 (s, 2H), 4.25 (m, 1H), 3.24 (m, 1H), 3.09 (m, 2H), 1.74-1.70 (m, 4H), 1.47-1.31 (m, 12H), 0.92 (t, $J = 6.84$ Hz, 3H); $^{13}\text{C NMR}$ (CD_3CN) $\delta = 172.90, 157.37, 138.62, 129.44, 129.11, 128.82, 128.67, 126.51, 79.00, 66.65, 56.75, 41.4, 34.8, 32.17, 30.4, 28.44, 27.3, 27.03, 25.41, 23.11, 14.27$; LC-MS (ESI): $t_R = 12.31$ min, calcd. for $\text{C}_{21}\text{H}_{31}\text{NO}_4\text{Na}$ $[\text{M}+\text{Na}]^+$, 384.46, found 384.13.

(3*S*, 4*S*)-3-(2-phenylethyl)-4-[(5-Cbz-amino)pentyl]-2-oxetanone (18)

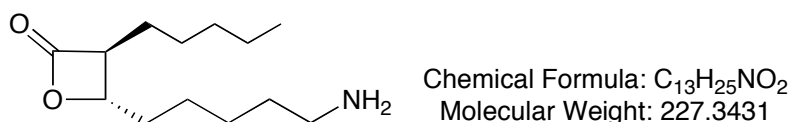


Following the same protocol as above for **17** led to the product **18** (yield: 24.6 mg, 0.06 mmol, 71%) as a colourless oil, however 40% of product was racemized.

$[\alpha]_D^{20} = -8.94$ ($c = 2.5$, CD_3CN); TLC (ethyl acetate/cyclohexane = 1:2): $R_f = 0.36$; $^1\text{H NMR}$ (CD_3CN) $\delta = 7.33\text{-}7.21$ (m, 10H), 5.58 (s br, 1H), 5.09 (s, 2H), 4.50(*syn*)/4.26(*anti*) (m/m, 1H), 3.66(*syn*)/3.24(*anti*) (m, 1H), 3.09 (m, 2H), 2.8~2.65 (m/m, *syn/anti*, 2H), 2.05 (m, 2H), 1.72-1.64 (m, 2H), 1.46-1.31 (m, 6H); $^{13}\text{C NMR}$ (CD_3CN) $\delta = 173.38, 172.60, 154.45, 143.14, 142.05, 129.45, 129.44, 129.41, 128.78, 128.63, 127.13, 127.09, 79.17$ (*anti*), 76.77(*syn*), 66.61, 56.19(*anti*),

52.66(*syn*), 41.35, 34.6, 34.1, 33.63, 30.55, 30.33, 30.27, 26.97, 26.93, 26.5, 25.94, 25.37; LC-MS (ESI): $t_R = 14.03$ min, calcd. for $C_{24}H_{29}NO_4Na$ $[M+Na]^+$, 418.48, found 418.05.

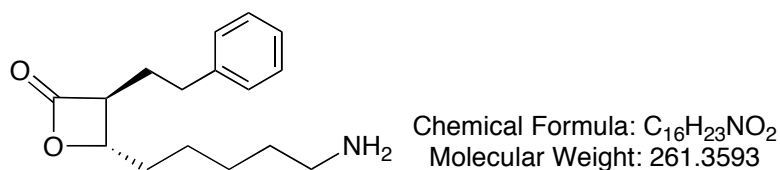
(3*S*, 4*S*)-3-pentyl-4-[(5-amino)pentyl]-2-oxetanone (19)



17 (80 μ mol, 28.9 mg) was weighted in a 10 ml flask with two necks equipped with gas catheters. Ethanol (2 ml), chloroform (25 μ l) and Pd (10%) on active carbon (2 mg, 2% mol) were added. The reaction mixture was heated by an oil bath to 40 $^{\circ}C$ and a hydrogen gas stream was applied. After 4 h, the suspension was filtrated and the filtrate was concentrated *in vacuo* to obtain the pure product **19** (yield: 18 mg, 80 μ mol, 100%) as a colourless oil. As free amino group could cause the degradation of the product, the product **19** was immediately transformed in a next step without further analyses.

TLC (ethyl acetate/cyclohexane = 2:1): $R_f = \sim 0$.

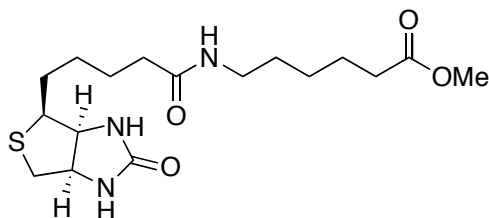
(3*S*, 4*S*)-3-(2-Phenylethyl)-4-[(5-amino)pentyl]-2-oxetanone (20)



Following the same protocol as above for **19** led to the racemized product **20** (yield: 15.6 mg, 60 μ mol, 100%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 2:1): $R_f = \sim 0$.

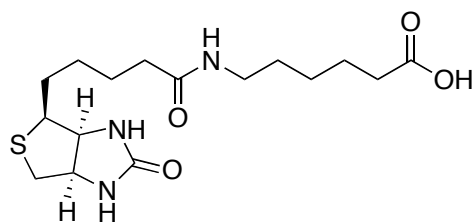
Methyl 6-(Biotinyl)-aminocaproate (**21**)



Chemical Formula: $C_{17}H_{29}N_3O_4S$
Molecular Weight: 371.4949

To a solution of **5** (284 mg, 1.56 mmol), biotin (366 mg, 1.5 mmol) and TEA (1.1 mL) in acetonitrile (4 mL) was added HBTU (591 mg, 1.56 mmol) and HOBt (210 mg, 1.56 mmol). After 30 min stirring at room temperature, a precipitate had formed. Water was added and the product was extracted with DCM. The organic phase was dried over anhydrous Na_2SO_4 , evaporated to dryness and the crude product was purified by silica gel chromatography (MeOH/DCM = 1:9) to obtain **21** (yield: 434 mg, 1.17 mmol, 78%) as a white powder.

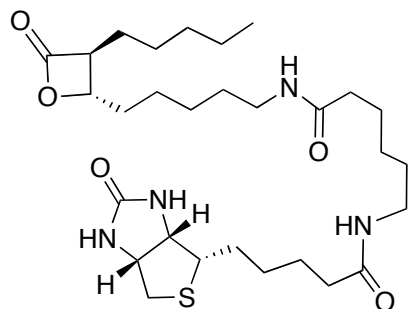
TLC (MeOH/DCM = 1:9): $R_f = 0.63$; 1H NMR ($CDCl_3$): $\delta = 6.46$ (s, 1H), 6.15 (br s, 1H), 5.56 (br s, 1H), 4.52 (m, 1H), 4.31 (m, 1H), 3.66 (s, 3H), 3.32 (q, $J = 6.8$ Hz, 2H), 3.15 (m, 1H), 2.91 (dd, $J = 4.9, 8$ Hz, 1H), 2.73 (d, $J = 12.7$ Hz, 1H), 2.32 (t, $J = 7.6$ Hz, 2H), 2.21 (t, $J = 7.6$ Hz, 2H), 1.6–1.3 (m, 13H); ^{13}C NMR ($CDCl_3$): $\delta = 174.4, 173.4, 164.2, 62.0, 60.4, 55.9, 51.8, 40.8, 39.4, 36.3, 34.1, 29.5, 28.4, 28.3, 26.6, 25.9, 24.7$; LC-MS (ESI): $t_R = 11.74$ min, calcd. for $C_{17}H_{30}N_3O_4S$ $[M+H]^+$, 372.19, found 372.18.

6-(Biotinyl)-aminocaproic acid (22)

Chemical Formula: C₁₆H₂₇N₃O₄S
Molecular Weight: 357.4683

21 (434 mg, 1.17 mmol) was dissolved in acetonitrile (7.5 mL) and cooled to 0 °C. (*n*-Bu₄N)OH × 30 H₂O (1.80 g, 2.25 mmol) was added under vigorous stirring. After 40 min, the reaction mixture was acidified by addition of 1 M HCl. The desired product **22** precipitated as a white solid. It was filtered off, washed with water and dried to obtain **22** (yield: 350 mg, 0.98 mmol, 84 %) as a white solid.

¹H NMR (DMSO-d₆): δ = 11.96 (br s, 1H), 7.71 (s, 1H), 6.40 (s, 1H), 6.34 (s, 1H), 4.28 (m, 1H), 4.11 (m, 1H), 3.08 (m, 1H), 2.97 (q, J = 6.8 Hz, 2H), 2.82 (dd, J = 4.9, 8 Hz, 1H), 2.73 (d, J = 12.7 Hz, 1H), 2.12 (t, J = 7.6 Hz, 2H), 2.01 (t, J = 7.6 Hz, 2H), 1.6 – 1.3 (m, 13H); ¹³C NMR (DMSO-d₆): δ = 174.4, 171.7, 162.7, 61.0, 59.2, 55.4, 38.2, 35.2, 33.6, 28.9, 28.2, 28.0, 26.0, 25.3, 24.2.

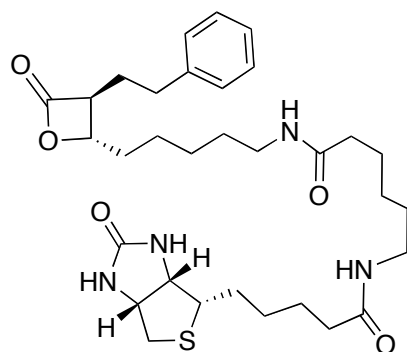
(3S, 4S)-3-Pentyl-4-[(5-(6-biotinylamino)caproylamino)pentyl]-2-oxetanone (23)

Chemical Formula: C₂₉H₅₀N₄O₅S
Molecular Weight: 566.7961

To a solution of **22** (13.4 mg, 37.6 μmol) and TEA (14.3 μl , 103 μmol) in DMSO (500 μl), PyBOP (19.8 mg, 38 μmol) and HOBt hydrate (5.1 mg, 38 μmol) were added at room temperature. After 15 min, **19** (7.76 mg, 34.2 μmol) in DMSO (200 μl) was added. The reaction was completed overnight. Because of the poor solubility in many organic solvents, the product was precipitated with ethyl acetate/diethyl ether 10 ml (1:1), washed with sodium bicarbonate solution, ethyl ether and dried over Na_2SO_4 (yield: 13.6 mg, 24 μmol , 70%) as a white solid.

^1H NMR (DMSO- d_6) δ = 6.40 (s, 1H), 6.34 (m, 1H), 4.33 (m, 2H), 4.12 (m, 1H), 3.09 (m, 1H), 3.01 (m, 4H), 2.83 (dd, J = 12.5 Hz, 4.95 Hz, 1H), 2.57 (d, J = 12.5 Hz, 1H), 2.2~2.0 (m, 4H), 1.74 ~1.27 (m, 28H), 0.86 (t, J = 6.6 Hz, 3H); ^{13}C NMR (DMSO- d_6) δ = 172.44, 166.94, 161.26, 157.12, 83.09, 78.29, 61.74, 59.89, 56.12, 50.93, 49.36, 41.14, 38.9, 35.9, 31.55, 29.67, 28.9, 28.72, 27.67, 26.63, 26.0, 24.86, 22.54, 17.56; LC-MS (ESI): t_{R} = 7.99 min, calcd. for $\text{C}_{29}\text{H}_{51}\text{N}_4\text{O}_5\text{S}$ $[\text{M}+\text{H}]^+$, 567.80, found 567.41.

(3S, 4S)-3-(2-Phenylethyl)-4-[(5-(6-biotinylamino)caproylamino)pentyl]-2-oxetanone (24)

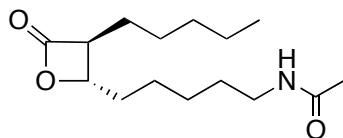


Chemical Formula: $\text{C}_{32}\text{H}_{48}\text{N}_4\text{O}_5\text{S}$
Molecular Weight: 600.8123

Following the same protocol as above for **23** led to the racemized product **24** (yield: 16.2 mg, 27 μ mol, 65%) as a colourless oil.

^1H NMR (DMSO- d_6) δ = 7.32-7.27 (m, 5H), 6.40 (s, 1H), 6.33 (s, 1H), 4.34 (m, 1H), 4.27 (m, 1H), 4.11 (m, 1H), 3.09 (m, 1H), 3.01 (m, 4H), 2.83 (dd, J = 12.5 Hz, 4.95 Hz, 1H), 2.57 (d, J = 12.5 Hz, 1H), 2.04 (m, 6H), 1.74-1.19 (m, 25H); ^{13}C NMR (DMSO- d_6) δ = 172.44, 172.24, 168.83, 168.41, 163.39, 163.37, 162.98, 141.44, 131.27, 129.08, 129.06, 128.87, 128.4, 127.93, 126.77, 76.96, 59.89, 56.11, 55.3, 49.43, 47.46, 46.53, 39.0, 35.9, 34.05, 33.03, 29.68, 28.9, 28.72, 26.83, 25.78; LC-MS (ESI): t_R = 7.93 min, calcd. for $\text{C}_{32}\text{H}_{48}\text{N}_4\text{O}_5\text{S}$ $[\text{M}+\text{H}]^+$, 601.81, found 601.30.

(3*S*, 4*S*)-3-Pentyl-4-[(5-acetylamino)pentyl]-2-oxetanone (**25**)



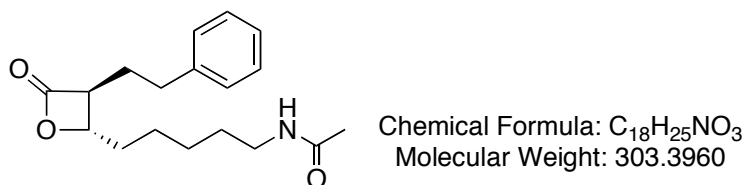
Chemical Formula: $\text{C}_{15}\text{H}_{27}\text{NO}_3$
Molecular Weight: 269.3798

Acetic anhydride (14.28 mg, 140 μ mol) and DIPEA (14.6 μ l, 84 μ mol) were dissolved in anhydrous DCM (280 μ l) at room temperature. To this solution, **19** (6.29 mg, 27.7 μ mol) in anhydrous DCM (100 μ l) was added. The reaction was stirred overnight and the mixture was purified directly by a short column chromatography (ethyl acetate/cyclohexane = 5:1) to get the pure product **25** (yield: 4 mg, 15 μ mol, 54%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 2:1): R_f = 0.1; ^1H NMR (CD_3CN) δ = 6.3 (s br, 1H), 4.26 (m, 1H), 3.26 (m, 1H), 3.09 (m, 2H), 1.82 (s, 3H), 1.74-1.70 (m, 4H),

1.47-1.31 (m, 12H), 0.92 (t, $J = 6.84$ Hz, 3H); ^{13}C NMR (CD_3CN) $\delta = 180.09$, 153.68, 79.00, 56.71, 39.68, 34.8, 32.14, 30.1, 28.4, 27.3, 27.2, 25.4, 23.1, 23.04, 14.22.

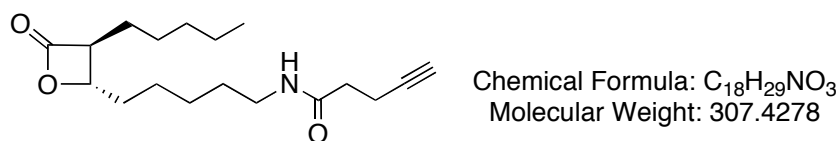
(3*S*, 4*S*)-3-(2-phenylethyl)-4-[(5-acetylamino)pentyl]-2-oxetanone (26)



Following the same protocol as above for **25** led to the racemized product **26** (yield: 3.7 mg, 12 μmol , 59%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 2:1): $R_f = 0.08$; ^1H NMR (CD_3CN) $\delta = 7.38\sim 7.23$ (m, 5H), 6.3 (s br, 1H), 4.55(*syn*)/4.29(*anti*) (m/m, 1H), 3.69(*syn*)/3.27(*anti*) (m, 1H), 3.09 (m, 2H), 2.8~2.65 (m/m, *syn/anti*, 2H), 2.05 (m, 2H), 1.84 (s, 3H), 1.72~1.66 (m, 2H), 1.46~1.31 (m, 6H); ^{13}C NMR (CD_3CN) $\delta = 183.3$, 147.67, 142.11, 136.64, 129.5, 129.4, 127.18, 85.97, 83.62, 70.61, 65.02, 66.61, 51.73, 39.73, 34.66, 33.67, 30.07, 27.21, 27.18, 26.56, 25.46, 23.08.

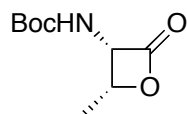
(3*S*, 4*S*)-3-Pentyl-4-[(5-pentynylamino)pentyl]-2-oxetanone (27)



To a solution of 4-pentynoic acid (4.6 μmol , 0.5 mg) in anhydrous DCM (100 μl) were added HOBt (4.6 μmol , 0.6 mg) and DIC (6 μmol , 1 μl). After a short time, **19** (0.9 mg, 4 μmol) in anhydrous DCM (100 μl) was added to this reaction mixture. The resulting reaction was stirred for 3 h and directly purified by HPLC to get the pure product **27** (yield: 0.5 mg, 1.6 μmol , 40%) as a yellow oil.

TLC (ethyl acetate/cyclohexane = 2:1): $R_f = 0.15$; $^1\text{H NMR}$ (CDCl_3) $\delta = 5.61$ (m, 1H), 4.20 (m 1H), 3.28 (t, $J = 6.44$ Hz, 2H), 3.16 (m, 1H), 2.53 (t, $J = 6.84$ Hz, 2H), 2.38 (t, $J = 6.84$ Hz, 2H), 2.01 (t, $J = 2.52$ Hz, 1H), 1.72-1.25 (m, 16H), 0.89 (t, $J = 6.8$ Hz, 3H); LC-MS (ESI): $t_R = 7.25$ min, calcd. for $\text{C}_{18}\text{H}_{30}\text{NO}_3$ $[\text{M}+\text{H}]^+$, 308.43, found 308.52.

(3*S*,4*R*)-3-[(*tert*-Butyloxycarbonyl)amino]-4-methyloxetan-2-one (**28**)



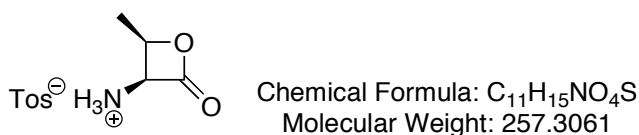
Chemical Formula: $\text{C}_9\text{H}_{15}\text{NO}_4$
Molecular Weight: 201.2197

To a mixture of Boc-*L*-Thr-OH (205 mg, 1 mmol), and PyBOP (0.624 g, 1.2 mmol) in anhydrous DCM (20 mL), TEA (0.42 mL, 3 mmol) was added dropwise at 0 $^\circ\text{C}$. The reaction mixture was stirred at 0 $^\circ\text{C}$ for 30 min and then warmed to room temperature. After stirring for 6 h, it was concentrated in *vacuo* and purified by column chromatography (ethyl acetate/cyclohexane = 1:6) to get the pure product **28** (yield: 0.159 g, 0.79 mmol, 85%) as a white solid.

M.p. = 141-142 $^\circ\text{C}$; $[\alpha]_D^{20} = +20.4$ ($c = 1.16$, CHCl_3); TLC (ethyl acetate/cyclohexane = 1:1): $R_f = 0.53$; $^1\text{H NMR}$ (CDCl_3): $\delta = 5.43$ -5.40 (m, 1H), 5.25 (br s, 1H), 4.84 (p, $J = 6.2$ Hz, 1H), 1.46 (s, 9 H), 1.41 (d, 6.3 Hz, 3 H); ^{13}C

NMR (CDCl₃): δ = 169.2, 154.4, 81.3, 74.8, 60.1, 28.1, 15.0; HRMS (ESI) calcd. for C₉H₁₆NO₄ [M+H]⁺, 202.1071, found 202.1079.

(3*S*,4*R*)-3-amino-4-methyloxetan-2-one p-toluenesulfonate salt (29**)**



To a mixture of *para*-toluenesulfonic acid (0.334 g, 1.94 mmol) and TFA (5 mL), **28** (0.368 g, 1.83 mmol) was added dropwise at 0 °C and the resulting mixture was stirred for 15 min. TFA was removed in *vacuo* and the remaining residue was co-evaporated twice with toluene. Dry ether was added to the residue, the white solid **29** was filtered, washed twice with dry ether and used in the subsequent reaction without further purification (yield: 0.457 g, 1.78 mmol, 97%).

General method for syntheses of 30 – 34

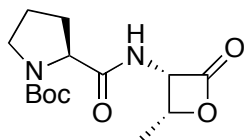
Method A

To a solution of *N*-Boc protected amino acid (0.5 mmol) in anhydrous DCM (4 mL) was added at -5 °C TEA (51 mg, 0.5 mmol) and ethyl chloroformate (54 mg, 0.5 mmol). After stirring for 20 min, **29** (0.128 g, 0.5 mmol) and pyridine (79 mg, 1 mmol) were added. After further 30 min stirring at -5 °C, the reaction mixture was warmed to room temperature and stirred overnight. The solvent was removed and the residue was dissolved in ethyl acetate, washed with water, dried over anhydrous Na₂SO₄ and concentrated to dryness. The residue was purified by silica gel column chromatography with ethyl acetate/cyclohexane.

Method B

To a solution of the *N*-Boc protected amino acid (1.05 mmol), PyBOP (1.1 mmol), HOBt (1.1 mmol) and DIEA (0.42 mL) in DMF/DCM (1:1, 5 mL), **29** (0.256 g, 1 mmol) was added at 0 °C. The reaction mixture was slowly warmed to room temperature and stirred for 4 h. Solvents were removed in *vacuo*, the residue was dissolved in ethyl acetate, washed with water, dried over anhydrous Na₂SO₄ and concentrated to dryness. The remaining residue was purified by silica gel column chromatography with ethyl acetate/cyclohexane.

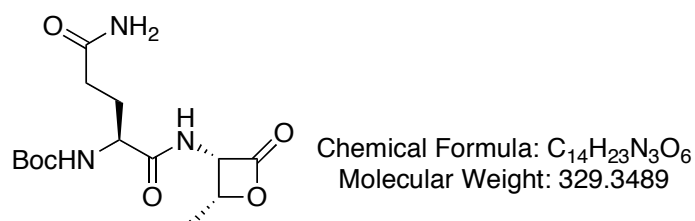
(3*S*,4*R*)-3-[(*N*-Boc prolinyl)amidyl]-4-methyloxetan-2-one (30)



Chemical Formula: C₁₄H₂₂N₂O₅
Molecular Weight: 298.3349

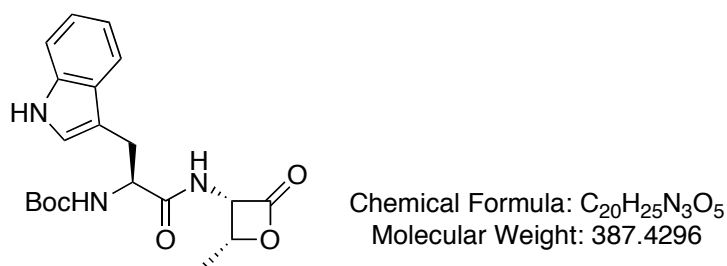
Method A was followed, yield: 0.128g, 0.43 mmol, 86%; white solid.

M.p. = 136-137 °C; $[\alpha]_D^{20} = -71.4$ (c = 1.55, CHCl₃); TLC (ethyl acetate/cyclohexane = 1:1): R_f = 0.28; ¹H NMR (CDCl₃): δ = 8.15 (br s, 1H), 5.58 (br s, 1H), 4.82 (p, J = 6.2 Hz, 1H), 4.32-4.21 (m, 1H), 3.53-3.25 (m, 2H), 2.35-1.85 (m, 4H), 1.42 (s, 9H), 1.36 (d, J = 6.2 Hz, 3H); ¹³C NMR (CDCl₃): δ = 171.8, 169.0, 156.1, 80.7, 74.3, 59.2, 58.6, 47.0, 28.2, 27.3, 24.6, 14.9; HRMS calcd. for C₁₄H₂₃N₂O₅ [M+H]⁺, 299.1607, found 299.1588.

(3S,4R)-3-[(N-Boc glutamyl)amidyl]-4-methyloxetan-2-one (31)

Method A was followed, yield: 0.11 g, 0.33 mmol, 67%; white solid.

M.p. = 144-145 °C; $[\alpha]_D^{20} = +8.6$ (c = 1.09, CH₃OH), TLC (ethyl acetate): R_f = 0.15; ¹H NMR (CD₃OD): δ = 5.51 (d, J = 5.8 Hz, 1H), 4.86 (p, J = 6.2 Hz, 1H), 4.07 (dd, J = 8.5, 5.4 Hz, 1H), 2.32 (t, J = 7.4 Hz, 2H), 2.07-1.99 (m, 1H), 1.92-1.85 (m, 1H), 1.43 (s, 9H), 1.40 (d, J = 6.4 Hz, 3H); ¹³C NMR (CD₃OD): δ = 178.5, 175.9, 171.4, 158.6, 81.6, 76.9, 60.8, 56.4, 33.4, 29.7, 29.5, 16.1; HRMS calcd. for C₁₄H₂₄N₃O₆ [M+H]⁺, 330.1665, found 330.1649.

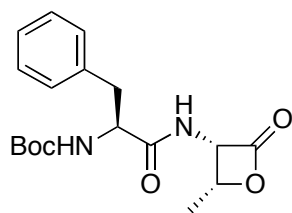
(3S,4R)-3-[(N-Boc tryptophanyl)amidyl]-4-methyloxetan-2-one (32)

Method B was followed, yield: 0.325 g, 0.84 mmol, 84%; white solid.

M.p. = 135-136 °C; $[\alpha]_D^{20} = +10.4$ (c = 1.03, CH₃OH), TLC (ethyl acetate/cyclohexane = 1:1): R_f = 0.26; ¹H NMR (CD₃OD) δ = 7.63 (d, J = 6.2 Hz,

1H), 7.36 (d, $J = 6.4$ Hz, 1H), 7.13-7.03 (m, 3H), 5.47 (d, $J = 4.6$ Hz, 1H), 4.80 (p, $J = 5.0$ Hz, 1H), 4.44 (t, $J = 5.5$ Hz, 1H), 3.26 (dd, $J = 11.5, 4.7$ Hz, 1H), 3.13 (dd, $J = 11.5, 6.1$ Hz, 1H), 1.40 (s, 9H), 1.30 (d, $J = 5.0$ Hz, 3H); ^{13}C NMR (CD_3OD): $\delta = 175.9, 171.2, 158.3, 138.8, 129.6, 125.4, 123.3, 120.7, 120.2, 113.1, 111.4, 81.5, 76.8, 60.8, 57.5, 29.9, 29.5, 15.9$; HRMS calcd. for $\text{C}_{20}\text{H}_{25}\text{N}_3\text{O}_5$ $[\text{M}]^+$, 387.1794, found 387.1786.

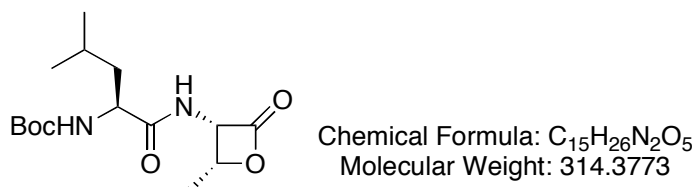
(3S,4R)-3-[(N-Boc phenylalaninyl)amidyl]-4-methyloxetan-2-one (33)



Chemical Formula: $\text{C}_{18}\text{H}_{24}\text{N}_2\text{O}_5$
Molecular Weight: 348.3936

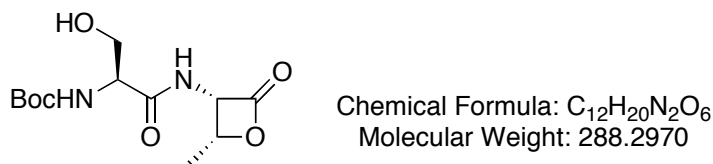
Method B was followed, yield: 0.24 g, 0.69 mmol, 69%; white solid.

M.p. = 141-143 °C; $[\alpha]_D^{20} = +10.3$ ($c = 1.01, \text{CHCl}_3$), TLC (ethyl acetate/cyclohexane = 1:1): $R_f = 0.36$; ^1H NMR (CD_3OD): $\delta = 7.32-7.18$ (m, 5H), 7.14 (d, $J = 7.7$ Hz, 1H), 5.52 (dd, $J = 8.0, 6.1$ Hz, 1H), 5.07 (d, $J = 7.8$ Hz, 1H), 4.84 (p, $J = 6.2$ Hz, 1H), 4.37 (q, $J = 7.2$ Hz, 1H), 3.11 (dd, $J = 13.9, 6.5$ Hz, 1H), 3.01 (dd, $J = 13.9, 7.6$ Hz, 1H), 1.40 (s, 9H), 1.32 (d, $J = 6.3$ Hz, 3H); ^{13}C NMR (CDCl_3): $\delta = 171.7, 168.4, 155.7, 136.1, 129.2, 128.8, 127.2, 81.4, 74.4, 58.8, 55.6, 37.4, 28.2, 14.9$; HRMS calcd. for $\text{C}_{18}\text{H}_{25}\text{N}_2\text{O}_5$ $[\text{M}+\text{H}]^+$, 349.1763, found 349.1732.

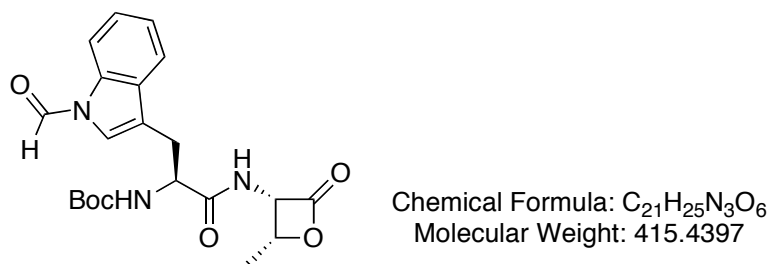
(3*S*,4*R*)-3-[(*N*-Boc leucanyl)amidyl]-4-methyloxetan-2-one (34)

Method B was followed, yield: 0.267 g, 0.85 mmol, 85%; white solid.

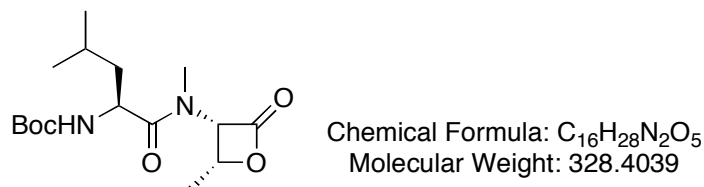
M.p. = 139-140 °C; $[\alpha]_D^{20} = -10.6$ ($c = 1.28$, CHCl₃); TLC (ethyl acetate/cyclohexane = 1:3): R_f = 0.18; ¹H NMR (CDCl₃): δ = 8.39 (d, J = 8.2 Hz, 1H), 5.90 (d, J = 8.2 Hz, 1H), 5.58 (dd, J = 8.5, 5.9 Hz, 1H), 4.90 (p, J = 6.0 Hz, 1H), 4.26-4.20 (m, 1H), 1.71-1.50 (m, 3H), 1.38 (s, 9H), 1.29 (d, J = 6.1 Hz, 3H), 0.89 (d, J = 6.6 Hz, 3H), 0.88 (d, J = 6.6 Hz, 3H); ¹³C NMR (CDCl₃): δ = 173.7, 169.0, 155.8, 79.8, 74.8, 58.8, 52.9, 40.1, 28.3, 24.6, 22.7, 21.9, 14.7; HRMS calcd. for C₁₅H₂₇N₂O₅ [M+H]⁺, 315.1920, found 315.1913.

(3*S*,4*R*)-3-[(*N*-Boc serinyl)amidyl]-4-methyloxetan-2-one (35)

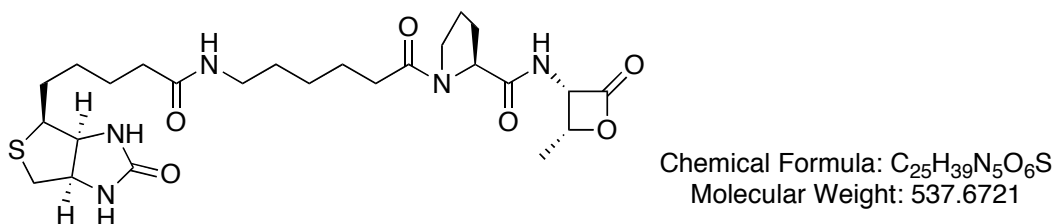
Method B was followed, yield: 0.225 g, 0.78 mmol, 78%; colourless oil.

(3S,4R)-3-[(N-Boc formyltryptophanyl)amidyl]-4-methyloxetan-2-one (36)

Method B was followed, yield: 0.341 g, 0.82 mmol, 82%; colourless oil.

(3S,4R)-3-[(N-Boc leuciny)methylamidyl]-4-methyloxetan-2-one (37)

Method B was followed, yield: 0.246 g, 0.75 mmol, 75%; white solid.

(3S,4R)-3-[(N-biotinyl prolinyl)amidyl]-4-methyloxetan-2-one (38)

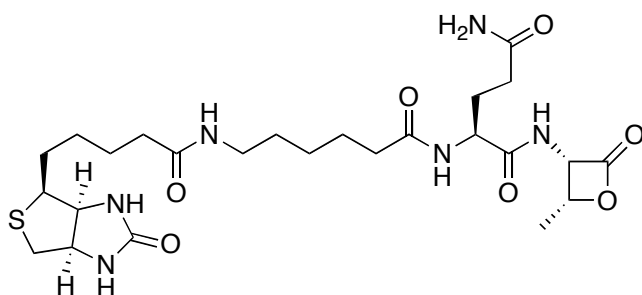
30 (44.7 mg, 150 μmol) was dissolved in DCM (2 mL). A mixture of triisopropylsilane (100 μL) and TFA (2 mL) was added and the resulting solution was stirred for 30 min. Evaporation to dryness, followed by co-evaporation in

presence of toluene delivered the deprotected intermediate which was used without further purification (yield: 46.8 mg, >98%).

22 (53.6 mg, 150 μmol) was dissolved in acetonitrile (3 mL). TEA (62.3 μL , 450 mmol) and few drops of DMSO were added. PyBOP (85.8 mg, 165 μmol) and HOBt hydrate (25.2 mg, 165 μmol) were then added and the resulting solution was stirred for further 15 min, after which deprotected **30** (46.8 mg, 150 μmol) dissolved in acetonitrile (1 mL) was added. The resulting reaction mixture was heated to 50 $^{\circ}\text{C}$ for 5 min and then stirred at room temperature overnight. The mixture was evaporated to dryness, the residue was taken up in a small amount of DCM and precipitated by addition of ethylacetate/diethyl ether (1:1, 20 mL) at -20 $^{\circ}\text{C}$ over two days. The precipitate was then purified by preparative HPLC (0 to 5 min, 10% aq. acetonitrile (0.1% TFA), 5 to 35 min, from 10% aq. acetonitrile to 70% aq. acetonitrile (0.1% TFA), 35 to 42 min, from 70% aq. acetonitrile to 100% acetonitrile (0.1% TFA)) to yield pure **38** (yield: 56.4 mg, 100 μmol , 67%) as a white powder.

^1H NMR (DMSO- d_6): δ = 6.42 (s, 1H), 6.36 (s, 1H), 4.30 (m, 1H), 4.13 (m, 1H), 3.09 (m, 4H), 3.00 (m, 4H), 2.80 (dd, J = 4, 10 Hz, 1H), 2.56 (d, J = 10 Hz, 1H), 2.18 (m, 2H), 2.04 (m, 2H), 1.89 – 1.25 (m, 21H); ^{13}C NMR (DMSO- d_6) δ = 173.4, 170.8, 169.9, 161.7, 73.6, 60.1, 58.2, 57.2, 54.4, 44.9, 37.3, 34.2, 32.6, 28.1, 27.2, 24.9, 24.3, 23.2, 13.7; LC-MS (ESI): t_{R} = 5.94 min, calcd. for $\text{C}_{25}\text{H}_{39}\text{N}_5\text{O}_6\text{S}$ $[\text{M}+\text{H}]^+$, 538.26, found 538.53.

(3*S*,4*R*)-3-[(*N*-biotinyl glutaminy)amidyl]-4-methyloxetan-2-one (39)



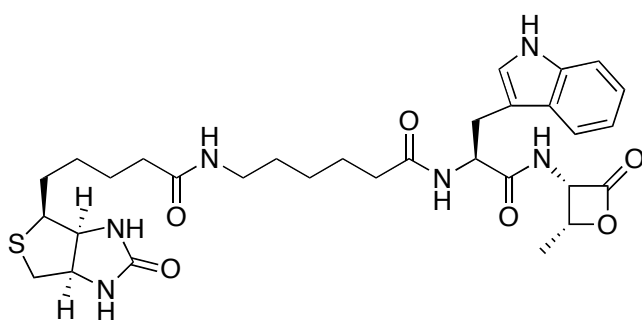
Chemical Formula: C₂₅H₄₀N₆O₇S
Molecular Weight: 568.6861

31 (20 mg, 60.8 μ mol) was dissolved in DCM (1 mL). A mixture of triisopropylsilane (50 μ L) and TFA (1 mL) was added and the resulting solution was stirred for 30 min. Evaporation to dryness, followed by co-evaporation in presence of toluene delivered the deprotected intermediate which was used without further purification (yield: 20.8 mg, >98%).

22 (21.8 mg, 61 μ mol) was dissolved in acetonitrile (1.2 mL). TEA (25.3 μ L, 183 mmol) and few drops of DMSO were then added. PyBOP (34.9 mg, 67.1 μ mol) and HOBt hydrate (10.3 mg, 67.1 μ mol) were added and the resulting solution was stirred for further 15 min, after which deprotected **31** (20.8 mg, 60.8 μ mol) dissolved in acetonitrile (400 μ L) was added. The resulting reaction mixture was heated to 50 $^{\circ}$ C for 5 min and then stirred at room temperature overnight. The mixture was evaporated to dryness, the residue was taken up in a small amount of DCM and precipitated by addition of ethylacetate/diethyl ether (1:1, 20 mL) at -20 $^{\circ}$ C over two days. The precipitate was then purified by preparative HPLC (0 to 5 min, 10% aq. acetonitrile (0.1% TFA), 5 to 40 min, from 10% aq. acetonitrile to 75% aq. acetonitrile (0.1% TFA), 40 to 46 min, from 75% aq. acetonitrile to 100% acetonitrile (0.1% TFA)) to yield pure **39** (yield: 23.8 mg, 42 μ mol, 69%) as a white powder.

LC-MS (ESI): $t_R = 9.95$ min, calcd. for $C_{25}H_{40}N_6O_7S$ $[M+H]^+$, 569.27, found 569.22.

(3*S*,4*R*)-3-[(*N*-biotinyl tryptophanyl)amidyl]-4-methyloxetan-2-one (40)



Chemical Formula: $C_{31}H_{42}N_6O_6S$
Molecular Weight: 626.7668

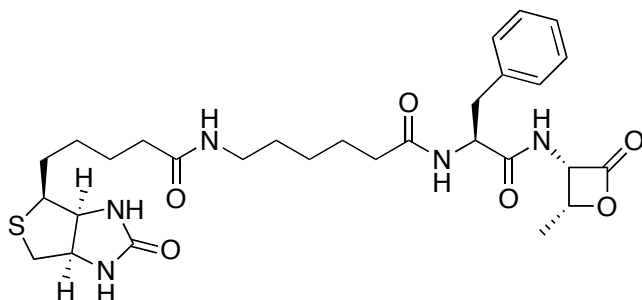
32 (15.5 mg, 40 μ mol) was dissolved in DCM (1 mL). A mixture of triisopropylsilane (50 μ L) and TFA (1 mL) was added and the resulting solution was stirred for 30 min. Evaporation to dryness, followed by co-evaporation in presence of toluene delivered the deprotected intermediate which was used without further purification (yield: 16.0 mg, >98%).

22 (14.3 mg, 40 μ mol) was dissolved in acetonitrile (1 mL). TEA (16.6 μ L, 120 mmol) and few drops of DMSO were then added. PyBOP (22.8 mg, 44 μ mol) and HOBt hydrate (6.7 mg, 44 μ mol) were added and the resulting solution was stirred for further 15 min, after which deprotected **32** (16.0 mg, 40 μ mol) dissolved in acetonitrile (400 μ L) was added. The resulting reaction mixture was heated to 50 $^{\circ}$ C for 5 min and then stirred at room temperature overnight. The mixture was evaporated to dryness, the residue was taken up in a small amount of DCM and precipitated by addition of ethylacetate/diethyl ether (1:1, 20 mL) at -20 $^{\circ}$ C over two days. The precipitate was then purified by preparative HPLC (0 to 5 min, 10% aq. acetonitrile (0.1% TFA), 5 to 40 min, from 10% aq. acetonitrile to 75% aq. acetonitrile (0.1% TFA), 40 to 46 min, from 75% aq. acetonitrile to 100%

acetonitrile (0.1% TFA)) to yield pure **40** (yield: 16.7 mg, 26.6 μmol , 67%) as a white powder.

^1H NMR (DMSO- d_6): δ = 7.97-7.73 (m, 5H), 6.40 (s, 1H), 6.34 (s, 1H), 5.52 (m, 1H), 4.85 (t, J = 6.2 Hz, 1H), 4.63 (m, 1H), 4.29 (m, 1H), 4.12 (m, 1H), 3.09 (m, 4H), 3.01 (m, 4H), 2.81 (dd, J = 5, 12.3 Hz, 1H), 2.57 (d, J = 12.3 Hz, 1H), 2.04 (m, 4H), 1.49 – 1.31 (m, 16H); ^{13}C NMR (DMSO- d_6) δ = 177.1, 172.5, 171.5, 162.4, 152.7, 145.9, 136.1, 118.8, 113.5, 112.6, 111.1, 95.7, 83.3, 81.7, 77.1, 74.2, 74.1, 63.9, 81.7, 77.1, 74.2, 63.9, 60.8, 58.9, 56.5, 55.1, 45.5, 34.9, 28.7, 27.9, 25.0, 24.5, 14.0; LC-MS (ESI): t_{R} = 6.93 min, calcd. for $\text{C}_{31}\text{H}_{42}\text{N}_6\text{O}_6\text{S}$ $[\text{M}+\text{H}]^+$, 627.29, found 627.47.

(3*S*,4*R*)-3-[(*N*-biotinyl phenylalaninyl)amidyl]-4-methyloxetan-2-one (41)



Chemical Formula: $\text{C}_{29}\text{H}_{41}\text{N}_5\text{O}_6\text{S}$
Molecular Weight: 587.7307

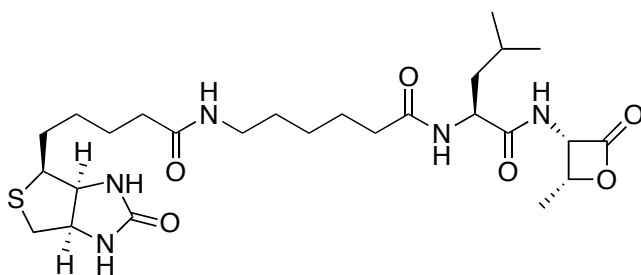
33 (20 mg, 57.5 μmol) was dissolved in DCM (1 mL). A mixture of triisopropylsilane (50 μL) and TFA (1 mL) was added and the resulting solution was stirred for 30 min. Evaporation to dryness, followed by co-evaporation in presence of toluene delivered the deprotected intermediate which was used without further purification (yield: 20.8 mg, >98%).

22 (20.7 mg, 58 μmol) was dissolved in acetonitrile (1.5 mL). TEA (24 μL , 174 mmol) and few drops of DMSO were then added. PyBOP (33.2 mg, 63.8 μmol)

and HOBt hydrate (9.7 mg, 63.8 μmol) were added and the resulting solution was stirred for further 15 min, after which deprotected **33** (20.8 mg, 57.5 μmol) dissolved in acetonitrile (400 μL) was added. The resulting reaction mixture was heated to 50 $^{\circ}\text{C}$ for 5 min and then stirred at room temperature overnight. The mixture was evaporated to dryness, the residue was taken up in a small amount of DCM and precipitated by addition of ethylacetate/diethyl ether (1:1, 20 mL) at -20 $^{\circ}\text{C}$ over two days. The precipitate was then purified by preparative HPLC (0 to 5 min, 20% aq. acetonitrile (0.1% TFA), 5 to 40 min, from 20% aq. acetonitrile to 80% aq. acetonitrile (0.1% TFA), 40 to 46 min, from 80% aq. acetonitrile to 100% acetonitrile (0.1% TFA)) to yield pure **41** (yield: 23.9 mg, 41 μmol , 71%) as a white powder.

^1H NMR (DMSO- d_6): δ = 7.26-7.18 (m, 5H), 6.42 (s, 1H), 6.35 (s, 1H), 5.48 (m, 1H), 4.84 (t, J = 5 Hz, 1H), 4.53 (m, 1H), 4.29 (m, 1H), 4.12 (m, 1H), 3.08 (m, 2H), 2.99 (m, 4H), 2.82 (m, 2H), 2.55 (d, J = 10 Hz, 1H), 2.02 (m, 4H), 1.74 – 1.45 (m, 16H); ^{13}C NMR (DMSO- d_6): δ = 172.0, 171.9, 169.0, 162.7, 137.7, 129.1, 128.0, 126.3, 65.0, 61.0, 59.2, 58.3, 55.4, 53.7, 45.9, 38.3, 37.4, 35.2, 28.9, 28.2, 25.9, 24.9, 14.6; LC-MS (ESI): t_{R} = 6.88 min, calcd. for $\text{C}_{29}\text{H}_{41}\text{N}_5\text{O}_6\text{S}$ $[\text{M}+\text{H}]^+$, 588.28, found 588.40.

(3S,4R)-3-[(N-biotinyl leuciny)amidyl]-4-methyloxetan-2-one (42)



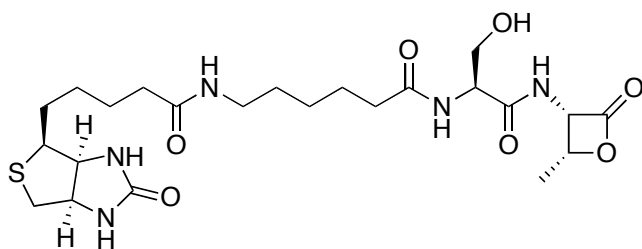
Chemical Formula: $\text{C}_{26}\text{H}_{43}\text{N}_5\text{O}_6\text{S}$
Molecular Weight: 553.7145

34 (12.6 mg, 40 μ mol) was dissolved in DCM (1 mL). A mixture of triisopropylsilane (50 μ L) and TFA (1 mL) was added and the resulting solution was stirred for 30 min. Evaporation to dryness, followed by co-evaporation in presence of toluene delivered the deprotected intermediate which was used without further purification (yield: 13.1 mg, >98%).

22 (14.3 mg, 40 μ mol) was dissolved in acetonitrile (1.5 mL). TEA (16.6 μ L, 120 mmol) and few drops of DMSO were then added. PyBOP (22.8 mg, 44 μ mol) and HOBt hydrate (6 mg, 44 μ mol) were added and the resulting solution was stirred for further 15 min, after which deprotected **34** (13.1 mg, 40 μ mol) dissolved in acetonitrile (400 μ L) was added. The resulting reaction mixture was heated to 50 $^{\circ}$ C for 5 min and then stirred at room temperature overnight. The mixture was evaporated to dryness, the residue was taken up in a small amount of DCM and precipitated by addition of ethylacetate/diethyl ether (1:1, 20 mL) at -20 $^{\circ}$ C over two days. The precipitate was then purified by preparative HPLC (0 to 5 min, 10% aq. acetonitrile (0.1% TFA), 5 to 35 min, from 10% aq. acetonitrile to 65% aq. acetonitrile (0.1% TFA), 35 to 40 min, from 65% aq. acetonitrile to 100% acetonitrile (0.1% TFA)) to yield pure **42** (yield: 13.7 mg, 24.8 μ mol, 62%) as a white powder.

^1H NMR (DMSO- d_6): δ = 6.38 (s, 1H), 6.32 (s, 1H), 5.52 (m, 1H), 4.82 (t, J = 6.2 Hz, 1H), 4.28 (m, 1H), 4.11 (m, 1H), 3.07 (m, 4H), 2.99 (m, 4H), 2.80 (dd, J = 5.1, 12.3 Hz, 1H), 2.54 (d, J = 12.3 Hz, 1H), 2.08 (m, 4H), 1.73 – 1.14 (m, 17H), 0.86 (d J = 6.4 Hz, 3H), 0.82 (d J = 6.4 Hz, 3H); ^{13}C NMR (DMSO- d_6) δ = 177.3, 172.2, 171.7, 169.7, 162.7, 76.1, 61.1, 59.2, 55.4, 54.9, 49.4, 45.8, 35.2, 33.4, 32.8, 28.2, 25.9, 25.3, 24.9, 21.5, 14.5; LC-MS (ESI): t_{R} = 5.61 min, calcd. for $\text{C}_{26}\text{H}_{43}\text{N}_5\text{O}_6\text{S}$ $[\text{M}+\text{H}]^+$, 554.29, found 554.47.

(3*S*,4*R*)-3-[(*N*-biotinyl serinyl)amidyl]-4-methyloxetan-2-one (43)



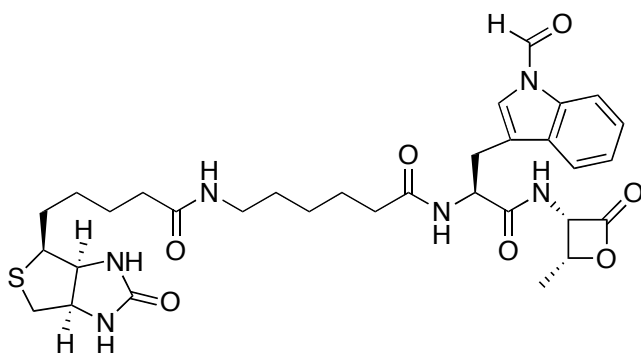
Chemical Formula: C₂₃H₃₇N₅O₇S
Molecular Weight: 527.6342

35 (43.2 mg, 150 μ mol) was dissolved in DCM (2 mL). A mixture of triisopropylsilane (100 μ L) and TFA (2 mL) was added and the resulting solution was stirred for 30 min. Evaporation to dryness, followed by co-evaporation in presence of toluene delivered the deprotected intermediate which was used without further purification (yield: 45.3 mg, >98%).

22 (53.6 mg, 150 μ mol) was dissolved in acetonitrile (3 mL). TEA (62.3 μ L, 450 μ mol) and few drops of DMSO were then added. PyBOP (85.8 mg, 165 μ mol) and HOBt hydrate (25.2 mg, 165 μ mol) were added and the resulting solution was stirred for further 15 min, after which deprotected **35** (45.3 mg, 150 μ mol) dissolved in acetonitrile (1000 μ L) was added. The resulting reaction mixture was heated to 50 $^{\circ}$ C for 5 min and then stirred at room temperature overnight. The mixture was evaporated to dryness, the residue was taken up in a small amount of DCM and precipitated by addition of ethylacetate/diethyl ether (1:1, 20 mL) at -20 $^{\circ}$ C over two days. The precipitate was then purified by preparative HPLC (0 to 5 min, 10% aq. acetonitrile (0.1% TFA), 5 to 35 min, from 10% aq. acetonitrile to 65% aq. acetonitrile (0.1% TFA), 35 to 45 min, from 65% aq. acetonitrile to 100% acetonitrile (0.1% TFA)) to yield pure **43** (yield: 43.4 mg, 82 μ mol, 55%) as a white powder.

^1H NMR (DMSO- d_6): δ = 6.42 (s, 1H), 6.35 (s, 1H), 4.28 (m, 1H), 4.12 (m, 1H), 3.08 (m, 4H), 3.01 (m, 4H), 2.81 (dd, J = 4.9, 9.9 Hz, 1H), 2.56 (d, J = 9.9 Hz, 1H), 2.17 (m, 2H), 2.03 (m, 2H), 1.72 – 1.16 (m, 21H); ^{13}C NMR (DMSO- d_6): δ = 171.8, 164.3, 162.7, 61.0, 59.2, 57.0, 55.4, 45.8, 38.2, 35.2, 33.6, 29.0, 28.2, 25.9, 25.3, 18.9; LC-MS (ESI): t_{R} = 6.03 min, calcd. for $\text{C}_{23}\text{H}_{37}\text{N}_5\text{O}_7\text{S}$ $[\text{M}+\text{H}]^+$, 528.24, found 528.40.

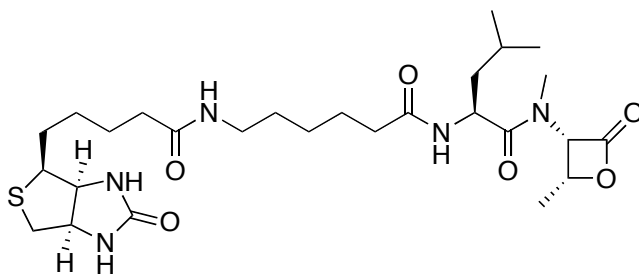
(3S,4R)-3-[(N-biotinyl formyltryptophanyl)amidyl]-4-methyloxetan-2-one (44)



Chemical Formula: $\text{C}_{32}\text{H}_{42}\text{N}_6\text{O}_7\text{S}$
Molecular Weight: 654.7769

Following the same protocol as above for **40** led to the product **44** (yield: 17.7 mg, 27 μmol , 66%) as a colourless oil.

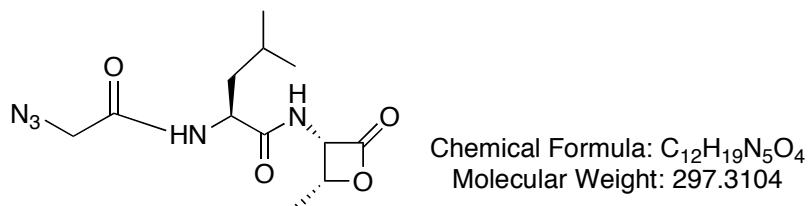
(3S,4R)-3-[(N-biotinyl leuciny)methylamidyl]-4-methyloxetan-2-one (45)



Chemical Formula: $\text{C}_{27}\text{H}_{45}\text{N}_5\text{O}_6\text{S}$
Molecular Weight: 567.7411

Following the same protocol as above for **42** led to the product **45** (yield: 15.9 mg, 28 μmol , 70%) as a colourless oil.

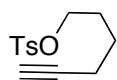
(3S,4R)-3-[(N-azidoacetyl leuciny)amidyl]-4-methyloxetan-2-one (46)



34 (24.3 mg, 77 μmol) was dissolved in DCM (2 mL). A mixture of triisopropylsilane (90 μL) and TFA (2 mL) was added and the resulting solution was stirred for 30 min. Evaporation to dryness, followed by co-evaporation in presence of toluene delivered the deprotected intermediate which was used without further purification (yield: 25.3 mg, >98%).

Azidoacetic acid (8 mg, 80 μmol) was dissolved in DCM (1 mL). TEA (38 μL , 280 mmol) and PyBOP (41.6 mg, 80 μmol) were added and the resulting solution was stirred for further 15 min, after which deprotected **34** (25.3 mg, 77 μmol) dissolved in DCM (200 μL) was added. The resulting reaction mixture was heated to 50 $^{\circ}\text{C}$ for 5 min and then stirred at room temperature overnight. The mixture was evaporated to dryness, the residue was then direct purified by a short column (ethyl acetate/cyclohexane = 1:1) to yield pure **46** (yield: 15.1 mg, 50.8 μmol , 66%) as a light yellow oil.

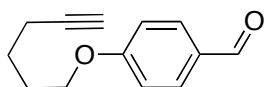
TLC (ethyl acetate/cyclohexane = 1:1): R_f = 0.38; LC-MS (ESI): t_R = 6.74 min, calcd. for $\text{C}_{12}\text{H}_{20}\text{N}_5\text{O}_4$ $[\text{M}+\text{H}]^+$, 298.31, found 298.52.

5-Hexyn-1-OTs (47)

Chemical Formula: C₁₃H₁₆O₃S
Molecular Weight: 252.3293

Tosyl chloride (380 mg, 2 mmol) was added to pyridine (5 ml) at 0 °C. After a short time, 5-Hexyn-1-ol (217 µl, 2 mmol) was added. The reaction was stirred overnight and the product was extracted by DCM and washed with brine. The organic phase was dried over anhydrous Na₂SO₄, evaporated to dryness and the crude product was purified by silica gel chromatography (ethyl acetate/cyclohexane = 2:1) to obtain **47** (yield: 271.6 mg, 1.08 mmol, 54%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 2:1): R_f = 0.76; ¹H NMR (CDCl₃) δ = 7.79 (d, J = 8 Hz), 7.35 (d, J = 8 Hz, 2H), 4.06 (t, J = 6.2 Hz, 2H), 2.45 (s, 3H), 2.17 (m, 2H), 1.92 (s, 1H), 1.78 (m, 2H), 1.56 (m, 2H); ¹³C NMR (CDCl₃) δ 144.86, 133.19, 129.95, 127.96, 83.48, 70.02, 69.06, 27.86, 24.32, 21.71, 17.82; LC-MS (ESI): t_R = 14.85 min, calcd. for C₂₀H₂₅N₅S₂ [M+Ts+H]⁺, 409.53, found 409.21.

4-(5-hexyn-1-oxy)-benzaldehyde (48)

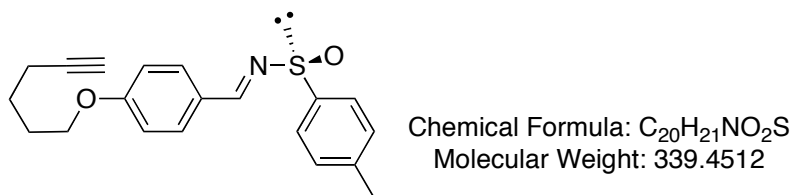
Chemical Formula: C₁₃H₁₄O₂
Molecular Weight: 202.2491

4-Hydroxybenzaldehyde (73.2 mg, 0.6 mmol) was added to **47** (151 mg, 0.6 mmol) and Cs₂CO₃ (292 mg, 0.9 mmol) in DMF (2 ml) under argon. The mixture was heated to 100 °C for 16 h, poured into water (5 ml) and extracted with diethyl ether. The crude product was purified by silica gel chromatography (ethyl

acetate/cyclohexane = 1:5) to obtain **48** (yield: 105.3 mg, 0.52 mmol, 87%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.23$; $^1\text{H NMR}$ (CDCl_3) $\delta = 9.88$ (s, 1H), 7.82 (d, $J = 8$ Hz, 2H), 6.68 (d, $J = 8$ Hz, 2H), 4.06 (t, $J = 6.2$ Hz, 2H), 2.28 (m, 2H), 1.96 (m, 3H), 1.71 (m, 2H); $^{13}\text{C NMR}$ (CDCl_3) $\delta = 190.96, 164.3, 132.2, 130.14, 114.97, 84.06, 69.04, 67.94, 28.28, 25.14, 18.34$; LC-MS (ESI): $t_R = 8.95$ min, calcd. for $\text{C}_{13}\text{H}_{15}\text{O}_2$ $[\text{M}+\text{H}]^+$, 202.25, found 203.30.

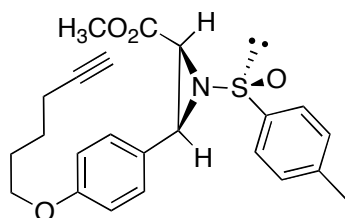
4-(5-hexyn-1-oxy)-phenyl sulfinate imine (**49**)



(1*R*, 2*S*, 5*R*)-(-)-menthyl (*S*)-*p*-toluenesulfinate (135.7 mg, 0.46 mmol) was dissolved in anhydrous THF (2.5 ml) at -78 °C. A solution of LiHMDS (0.6 mmol, 1.0 M in THF) (0.6 ml) was added dropwise and the reaction mixture is allowed to warm to rt with stirring. After 5 h, the reaction mixture was cooled to -78 °C again and **48** (103.2 mg, 0.51 mmol) was added and stirred for 2 h at -78 °C. The reaction mixture was quenched with water (0.6 ml), diluted with ethyl ether (5 ml) and warmed to room temperature. The organic layer was washed with water and brine, dried, and the crude product was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:5) to obtain pure product **49** (yield: 56.8 mg, 0.17 mmol, 37%) as a white solid. As the product is an instable imine, it was directly transferred into the next product and no detailed analysis was carried out.

TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.33$; $^1\text{H NMR}$ (CDCl_3) $\delta = 8.67$ (s, 1H), 7.78 (d, $J = 8$ Hz, 2H), 7.62 (d, $J = 8$ Hz, 2H), 7.30 (d, $J = 8$ Hz, 2H), 6.93 (d, $J = 8$ Hz, 2H), 4.03 (t, $J = 6.2$ Hz, 2H), 2.39 (s, 3H), 2.28 (m, 2H), 1.96 (m, 3H), 1.71 (m, 2H).

Aziridine (50)



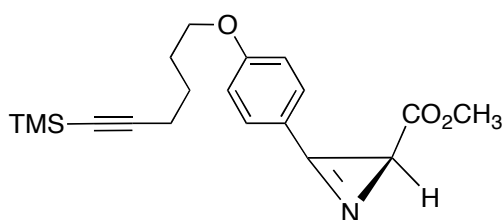
Chemical Formula: $\text{C}_{23}\text{H}_{25}\text{NO}_4\text{S}$
Molecular Weight: 411.5139

49 (56.8 mg, 0.17 mmol) in anhydrous THF (1 ml) was cooled to -78 °C and methyl α -bromoacetate (30.9 μl , 0.336 mmol) was added. After 3 min, LiHMDS (219 μl , 1M in THF) was added dropwise. The reaction was stirred at -78 °C for 20 min, quenched with water (1 ml) and diluted with ethyl acetate (3 ml). The organic layer was separated and the aqueous phase was washed with ethyl acetate. The combined organic layer was washed with brine, dried and the crude product was purified by silica gel chromatography (ethyl acetate/cyclohexane = 5:1) to obtain pure product **50** (yield: 49.3 mg, 0.12 mmol, 70.6%) as a white solid.

TLC (ethyl acetate/cyclohexane = 5:1): $R_f = 0.14$; $^1\text{H NMR}$ (CDCl_3) $\delta = 7.71$ (d, $J = 8$ Hz, 2H), 7.37 (d, $J = 8$ Hz, 2H), 7.32 (d, $J = 8$ Hz, 2H), 6.84 (d, $J = 8$ Hz, 2H), 3.97 (t, $J = 6.2$ Hz, 2H), 3.81 (d, $J = 7.32$ Hz, 1H), 3.45 (d, $J = 7.32$ Hz, 1H), 3.41 (s, 3H), 2.42 (s, 3H), 2.28 (m, 2H), 1.96 (m, 3H), 1.71 (m, 2H); $^{13}\text{C NMR}$ (CDCl_3) $\delta = 166.52, 159.31, 142.4, 141.15, 135.83, 135.82, 129.96, 129.24, 125.47, 124.73, 114.39, 84.28, 68.89, 67.5, 52.27, 52.26, 47.65, 42.22, 35.09, 28.48, 25.26, 21.72,$

18.38; LC-MS (ESI): $t_R = 5.45$ min, calcd. for $C_{23}H_{26}NO_4S$ $[M+H]^+$, 412.51, found 412.14.

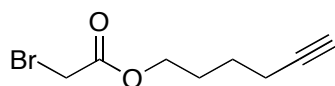
Azirine (51)



Chemical Formula: $C_{19}H_{25}NO_3Si$
Molecular Weight: 343.4922

To a solution of **50** (41.1 mg, 0.1 mmol) in anhydrous THF (2 ml), $TMSCl$ (77 μ l, 0.6 mmol) was added at -78 °C. 2 min later, LDA (0.3 ml, 0.3 mmol, 1.0 M in THF) was added dropwise. The reaction was stirred at -78 °C for 15 min, quenched with water (1 ml), and diluted with ethyl acetate (5 ml). The organic phase was separated and aqueous phase was washed with ethyl acetate. The combined organic layer was washed with brine, dried and the crude product was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:5) to obtain pure product **51** (yield: 16.5 mg, 48 μ mol, 48%) as a colourless oil.

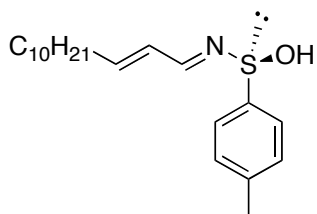
TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.46$; 1H NMR ($CDCl_3$) $\delta = 7.74$ (d, $J = 8.7$ Hz, 2H), 7.03 (d, $J = 8.7$ Hz, 2H), 4.07 (t, $J = 6.3$ Hz, 2H), 3.67 (s, 3H), 2.32 (t, $J = 7$ Hz, 3H), 1.94 (m, 2H), 1.72 (m, 2H), 0.14 (s, 9H); ^{13}C NMR ($CDCl_3$) $\delta = 174.52, 163.41, 132.47, 115.47, 106.79; 85.28, 67.93, 52.11, 31.27, 28.26, 25.18, 19.68, 0.28$; LC-MS (ESI): $t_R = 7.62$ min, calcd. for $C_{19}H_{26}NO_3Si$ $[M+H]^+$, 344.49, found 344.76.

5-hexynyl alpha-bromoacetate (52)

Chemical Formula: C₈H₁₁BrO₂
Molecular Weight: 219.0757

In a 25 ml flask, Bromoacetic acid (304 mg, 2.2 mmol), DCC (453 mg, 2.2 mmol) and DMAP (12.2 mg, 0.1 mmol) were dissolved in DCM (10 ml). To this mixture, 5-Hexyn-1-ol (217 μ l, 2 mmol) was added. The reaction was completed overnight. The organic layer was washed with brine, dried and the crude product was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:2) to obtain pure product **52** (yield: 246 mg, 1.12 mmol, 56%) as a yellow oil.

TLC (ethyl acetate/cyclohexane = 1:2): R_f = 0.6; ¹H NMR (CDCl₃) δ = 4.21 (t, J = 8.05 2H), 3.82 (s, 2H), 2.25 (m, 2H), 1.96 (t, J = 3.4 Hz, 1H), 1.81 (m, 2H), 1.64 (m, 2H); ¹³C NMR (CDCl₃) δ = 167.41, 83.8, 69.01, 65.9, 27.59, 25.94, 24.87, 18.17.

***trans*-2-tridecenyl sulfinic imine (53)**

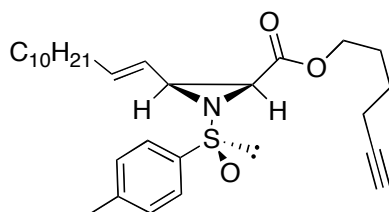
Chemical Formula: C₂₀H₃₁NOS
Molecular Weight: 333.5312

Following the same protocol as above for **49** led to the pure product **53** (yield: 130 mg, 0.39 mmol, 85%) as a white solid.

TLC (ethyl acetate/cyclohexane = 1:5): R_f = 0.61; ¹H NMR (CDCl₃) δ = 8.33 (d, J = 9.5 Hz, 1H), 7.56 (d, J = 8.25 Hz, 2H), 7.29 (d, J = 8.25 Hz, 2H), 6.55 (m, 1H),

6.38 (m, 1H), 2.39 (s, 3H), 2.25 (m, 2H), 1.45 (m, 2H), 1.26 (m, 16H), 0.87 (t, J = 6.7 Hz, 3H); ^{13}C NMR (CDCl_3) δ = 162.2, 152.7, 141.8, 130.0, 128.7, 124.8, 45.2, 33.2, 32.0, 29.7, 29.5, 29.3, 28.3, 22.8, 21.6, 14.3; LC-MS (ESI): t_{R} = 15.26 min, calcd. for $\text{C}_{20}\text{H}_{32}\text{NOS}$ $[\text{M}+\text{H}]^+$, 334.53, found 334.03.

Aziridine (54)

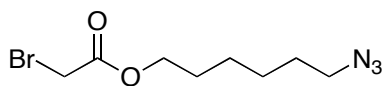


Chemical Formula: $\text{C}_{28}\text{H}_{41}\text{NO}_3\text{S}$
Molecular Weight: 471.6950

Following the same protocol as above for **50** led to the pure product **54** (yield: 99 mg, 0.21 mmol, 87%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 1:5): R_f = 0.52; ^1H NMR (CDCl_3) δ = 7.59 (m, 2H), 7.31 (m, 2H), 6.02 (m, 1H), 5.38 (m, 1H), 4.20 (m, 2H), 4.03 (t, J = 6.2 Hz, 1H), 3.59 (m, 1H), 3.27 (d, J = 3.75 Hz, 1H), 2.41 (s, 3H), 2.23 (m, 2H), 2.14 (m, 2H), 2.08 (m, 2H), 1.96 (m, 1H), 1.81 (m, 2H), 1.61 (m, 4H), 1.41-1.26 (m, 11H), 0.89 (t, J = 6.12 Hz, 3H); LC-MS (ESI): t_{R} = 11.16 min, calcd. for $\text{C}_{28}\text{H}_{42}\text{NO}_3\text{S}$ $[\text{M}+\text{H}]^+$, 472.70, found 472.24.

6-azido hexyl bromo acetate (55)

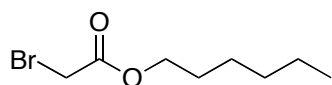


Chemical Formula: $\text{C}_8\text{H}_{14}\text{BrN}_3\text{O}_2$
Molecular Weight: 264.1197

Following the same protocol as above for **52** led to the pure product **55** (yield: 279 mg, 1.06 mmol, 53%) as a yellow oil.

TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.42$; $^1\text{H NMR}$ (CDCl_3) $\delta = 4.19$ (t, $J = 6.64$ Hz, 2H), 3.84 (s, 2H), 3.28 (t, $J = 6.84$ Hz, 2H), 1.68 (m, 2H), 1.61 (m, 2H), 1.42 (m, 4H); $^{13}\text{C NMR}$ (CDCl_3) $\delta = 167.4, 66.3, 51.5, 28.9, 28.4, 26.5, 26.0, 25.5$.

hexyl bromo acetate (**56**)

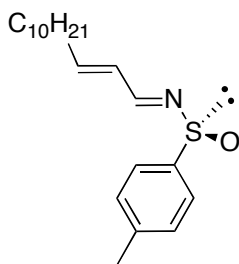


Chemical Formula: $\text{C}_8\text{H}_{15}\text{BrO}_2$
Molecular Weight: 223.1075

Following the same protocol as above for **52** led to the pure product **56** (yield: 227 mg, 1.02 mmol, 51%) as a light lemon oil.

TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.53$; $^1\text{H NMR}$ (CDCl_3) $\delta = 4.23$ (t, $J = 6.64$ Hz, 2H), 4.05 (s, 2H), 1.65 (m, 2H), 1.53 (m, 2H), 1.31 (m, 4H), 0.89 (t, $J = 6.8$ Hz, 3H); $^{13}\text{C NMR}$ (CDCl_3) $\delta = 167.1, 63.6, 30.5, 28.7, 26.2, 25.3, 22.4, 15.3$.

S-trans-2-tridecenyl sulfinat imine (**57**)

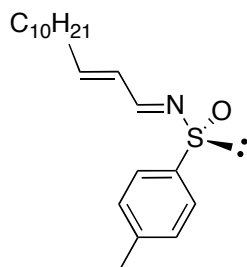


Chemical Formula: $\text{C}_{20}\text{H}_{31}\text{NOS}$
Molecular Weight: 333.5312

Following the same protocol as above for **49** led to the pure product **57** (yield: 130 mg, 0.39 mmol, 85%) as a white solid.

TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.61$; $^1\text{H NMR}$ (CDCl_3) $\delta = 8.33$ (d, $J = 9.5$ Hz, 1H), 7.56 (d, $J = 8.25$ Hz, 2H), 7.29 (d, $J = 8.25$ Hz, 2H), 6.55 (m, 1H), 6.38 (m, 1H), 2.39 (s, 3H), 2.25 (m, 2H), 1.45 (m, 2H), 1.26 (m, 16H), 0.87 (t, $J = 6.7$ Hz, 3H); $^{13}\text{C NMR}$ (CDCl_3) $\delta = 162.2, 152.7, 141.8, 130.0, 128.7, 124.8, 45.2, 33.2, 32.0, 29.7, 29.5, 29.3, 28.3, 22.8, 21.6, 14.3$; LC-MS (ESI): $t_R = 15.26$ min, calcd. for $\text{C}_{20}\text{H}_{32}\text{NOS}$ $[\text{M}+\text{H}]^+$, 334.53, found 334.03.

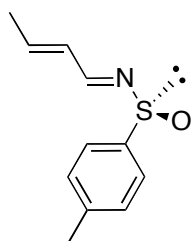
***R-trans*-2-tridecenyl sulfinate imine (**58**)**



Chemical Formula: $\text{C}_{20}\text{H}_{31}\text{NOS}$
Molecular Weight: 333.5312

Following the same protocol as above for **49**, but with (1*S*, 2*R*, 5*S*)-(+)-menthyl (*R*)-*p*-toluenesulfinate led to the pure enantiomer **58** (yield: 138 mg, 0.42 mmol, 90%) as a white solid.

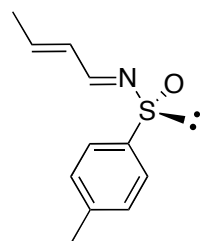
TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.61$; $^1\text{H NMR}$ (CDCl_3) $\delta = 8.34$ (d, $J = 9.16$ Hz, 1H), 7.57 (d, $J = 8.2$ Hz, 2H), 7.30 (d, $J = 8.2$ Hz, 2H), 6.55 (m, 1H), 6.42 (m, 1H), 2.40 (s, 3H), 2.27 (m, 2H), 1.45 (m, 2H), 1.26 (m, 16H), 0.88 (t, $J = 6.7$ Hz, 3H); LC-MS (ESI): $t_R = 15.34$ min, calcd. for $\text{C}_{20}\text{H}_{31}\text{NOS}$ $[\text{M}+\text{H}]^+$, 334.53, found 334.32.

***S-trans*-crotonyl sulfinate imine (59)**

Chemical Formula: C₁₁H₁₃NOS
Molecular Weight: 207.2920

Following the same protocol as above for **49** led to the product **59** (yield: 93.1 mg, 0.45 mmol, 45%) as a white solid.

TLC (ethyl acetate/cyclohexane = 1:5): R_f = 0.3; ¹H NMR (CDCl₃) δ = 8.34 (d, J = 9.16 Hz, 1H), 7.55 (d, J = 8 Hz, 2H), 7.28 (d, J = 8 Hz, 2H), 6.56 (m, 1H), 6.39 (m, 1H), 2.38 (s, 3H), 1.95 (d, J = 5.6 Hz, 3H); ¹³C NMR (CDCl₃) δ = 161.7, 147.1, 141.9, 141.6, 130.1, 129.8, 124.5, 34.5, 23.1, 20.9.

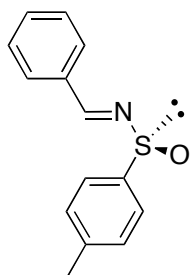
***R-trans*-crotonyl sulfinate imine (60)**

Chemical Formula: C₁₁H₁₃NOS
Molecular Weight: 207.2920

Following the same protocol as above for **49**, but with (1*S*, 2*R*, 5*S*)-(+)-menthyl (*R*)-*p*-toluenesulfinate led to the enantiomer **60** (yield: 107 mg, 0.52 mmol, 52%) as a white solid.

TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.3$; $^1\text{H NMR}$ (CDCl_3) $\delta = 8.34$ (d, $J = 9.16$ Hz, 1H), 7.55 (d, $J = 8$ Hz, 2H), 7.28 (d, $J = 8$ Hz, 2H), 6.56 (m, 1H), 6.39 (m, 1H), 2.38 (s, 3H), 1.95 (d, $J = 5.6$ Hz, 3H).

S-phenyl sulfinatone imine (61)

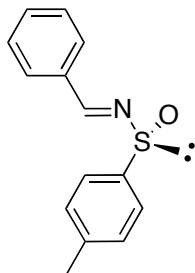


Chemical Formula: $\text{C}_{14}\text{H}_{13}\text{NOS}$
Molecular Weight: 243.3241

Following the same protocol as above for **49** led to the product **61** (yield: 211 mg, 0.87 mmol, 87%) as a light yellow solid.

TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.41$; $^1\text{H NMR}$ (CDCl_3) $\delta = 8.77$ (s, 1H), 7.85 (d, $J = 7.44$ Hz, 2H), 7.64 (d, $J = 8.2$ Hz, 2H), 7.49 (m, 3H), 7.32 (m, 2H), 2.41 (s, 3H); $^{13}\text{C NMR}$ (CDCl_3) $\delta = 160.8, 141.9, 141.9, 134.1, 132.7, 130.0, 129.7, 129.0, 124.9, 21.6$.

R-phenyl sulfinatone imine (62)

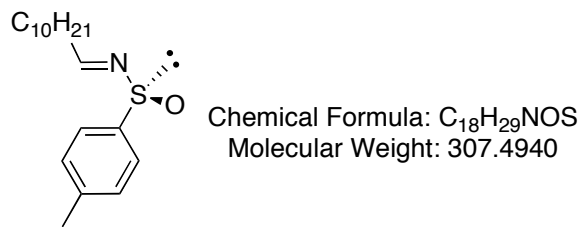


Chemical Formula: $\text{C}_{14}\text{H}_{13}\text{NOS}$
Molecular Weight: 243.3241

Following the same protocol as above for **49**, but with (1*S*, 2*R*, 5*S*)-(+)-menthyl (*R*)-*p*-toluenesulfinate led to the enantiomer **62** (yield: 218 mg, 0.9 mmol, 90%) as a light yellow solid.

TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.41$; $^1\text{H NMR}$ (CDCl_3) $\delta = 8.77$ (s, 1H), 7.84 (d, $J = 7.44$ Hz, 2H), 7.62 (d, $J = 8.2$ Hz, 2H), 7.48 (m, 3H), 7.32 (m, 2H), 2.41 (s, 3H);

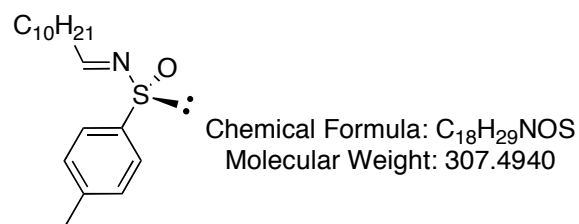
S-decanyl sulfinate imine (**63**)



Following the same protocol as above for **49** led to the product **63** (yield: 98 mg, 0.32 mmol, 32%) as a white solid.

TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.35$; $^1\text{H NMR}$ (CDCl_3) $\delta = 8.22$ (t, $J = 5.2$ Hz, 1H), 7.54 (d, $J = 7.44$ Hz, 2H), 7.28 (d, $J = 8.2$ Hz, 2H), 2.39 (s, 3H), 1.58-1.42 (m, 18H), 0.89 (t, $J = 6.5$ Hz, 3H).

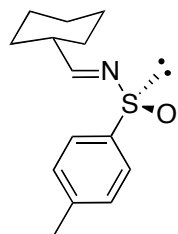
R-decanyl sulfinate imine (**64**)



Following the same protocol as above for **49**, but with (1*S*, 2*R*, 5*S*)-(+)-menthyl (*R*)-*p*-toluenesulfinate led to the enantiomer **64** (yield: 92 mg, 0.3 mmol, 30%) as a white solid.

TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.35$; $^1\text{H NMR}$ (CDCl_3) $\delta = 8.22$ (t, $J = 5.2$ Hz, 1H), 7.52 (d, $J = 7.44$ Hz, 2H), 7.26 (d, $J = 8.2$ Hz, 2H), 2.39 (s, 3H), 1.58-1.42 (m, 18H), 0.89 (t, $J = 6.5$ Hz, 3H).

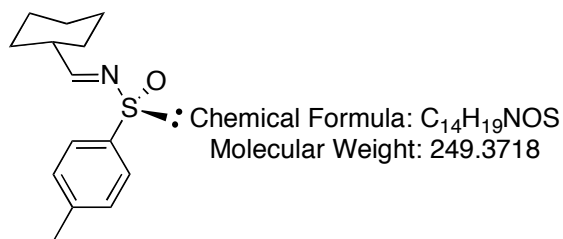
S-cyclohexyl sulfinate imine (65)



Chemical Formula: $\text{C}_{14}\text{H}_{19}\text{NOS}$
Molecular Weight: 249.3718

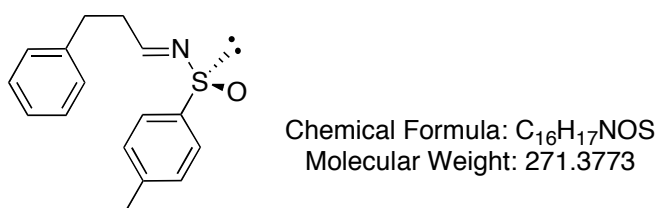
Following the same protocol as above for **49** led to the product **65** (yield: 204 mg, 0.82 mmol, 82%) as a white solid.

TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.34$; $^1\text{H NMR}$ (CDCl_3) $\delta = 8.09$ (d, $J = 4.8$ Hz, 1H), 7.52 (d, $J = 8$ Hz, 2H), 7.27 (d, $J = 8$ Hz, 2H), 2.37 (s, 3H), 1.95 (m, 1H), 1.93 (m, 2H), 1.89 (m, 2H), 1.83 (m, 2H), 1.74 (m, 4H); $^{13}\text{C NMR}$ (CDCl_3) $\delta = 170.6, 142.4, 141.8, 130.0, 124.9, 50.4, 45.3, 43.9, 29.4, 26.0, 25.5, 21.2$.

R-cyclohexyl sulfinate imine (66)

Following the same protocol as above for **49**, but with (1*S*, 2*R*, 5*S*)-(+)-menthyl (*R*)-*p*-toluenesulfinate led to the enantiomer **66** (yield: 212 mg, 0.85 mmol, 85%) as a white solid.

TLC (ethyl acetate/cyclohexane = 1:5): R_f = 0.34; ¹H NMR (CDCl₃) δ = 8.09 (d, J = 4.8 Hz, 1H), 7.52 (d, J = 8 Hz, 2H), 7.27 (d, J = 8 Hz, 2H), 2.37 (s, 3H), 1.95 (m, 1H), 1.93 (m, 2H), 1.89 (m, 2H), 1.83 (m, 2H), 1.74 (m, 4H).

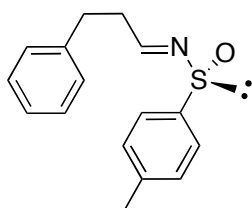
S-(3-phenyl propyl) sulfinate imine (67)

Following the same protocol as above for **49** led to the product **67** (yield: 67.3 mg, 0.33 mmol, 33%) as a white solid.

TLC (ethyl acetate/cyclohexane = 1:5): R_f = 0.22; ¹H NMR (CDCl₃) δ = 8.28 (t, J = 5.3 Hz, 1H), 7.42 (d, J = 8 Hz, 2H), 7.27 (d, J = 8 Hz, 2H), 7.13 (m, 3H), 7.04 (m,

2H), 2.37 (s, 3H), 1.95 (m, 2H), 1.93 (m, 2H); ^{13}C NMR (CDCl_3) $\delta = 165.9, 148.5, 142.1, 140.9, 139.0, 138.1, 129.3, 128.3, 127.5, 44.8, 34.3, 31.4, 21.9$.

S-(3-phenyl propyl) sulfinat imine (68)

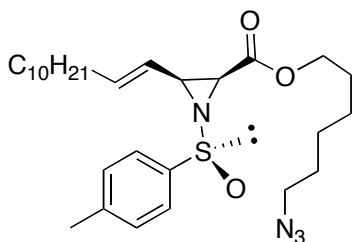


Chemical Formula: $\text{C}_{16}\text{H}_{17}\text{NOS}$
Molecular Weight: 271.3773

Following the same protocol as above for **49**, but with (1*S*, 2*R*, 5*S*)-(+)-menthyl (*R*)-*p*-toluenesulfinate led to the enantiomer **68** (yield: 92 mg, 0.34 mmol, 34%) as a white solid.

TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.22$; ^1H NMR (CDCl_3) $\delta = 8.21$ (t, $J = 5.3$ Hz, 1H), 7.42 (d, $J = 8$ Hz, 2H), 7.24 (d, $J = 8$ Hz, 2H), 7.13 (m, 3H), 7.04 (m, 2H), 2.31 (s, 3H), 1.95 (m, 2H), 1.93 (m, 2H); ^{13}C NMR (CDCl_3) $\delta = 165.9, 148.5, 142.1, 140.9, 139.0, 138.1, 129.3, 128.3, 127.5, 44.8, 34.3, 31.4, 21.9$.

Aziridine (69)

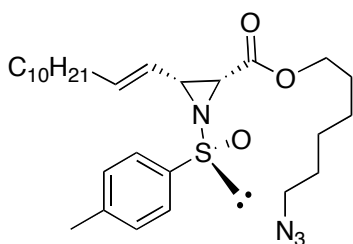


Chemical Formula: $\text{C}_{28}\text{H}_{44}\text{N}_4\text{O}_3\text{S}$
Molecular Weight: 516.7390

To a solution of **57** (186.5 mg, 0.56 mmol) in anhydrous THF (3 ml), **55** (297 mg, 1.13 mmol) in anhydrous THF (3 ml) was added at -78 °C. After 3 min, LiHMDS (0.73 ml, 0.73 mmol, 1.0 M in THF) was added dropwise. The reaction was stirred at -78 °C for 1 h and warmed up to -38°C for another 1 h, quenched with water (0.5 ml) and diluted with ethyl acetate (3 ml). The organic layer was separated and the aqueous phase was washed with ethyl acetate. The combined organic layer was washed with brine, dried and the crude product was purified by silica gel chromatography (ethyl acetate/cyclohexane = 5:1) to obtain pure product **69** (yield: 245 mg, 0.47 mmol, 85%) as a yellow oil.

TLC (ethyl acetate/cyclohexane = 1:5): R_f = 0.42; ¹H NMR (CDCl₃) δ = 7.59 (m, 2H), 7.30 (m, 2H), 6.00 (m, 1H), 5.50 (m, 1H), 4.17 (t, J = 6.44 Hz, 1H), 3.99 (m, 2H), 3.25 (m, 5H), 2.39 (s, 3H), 2.05 (m, 2H), 1.56-1.25 (m, 22H), 0.87 (t, J = 6.64 Hz, 3H); ¹³C NMR (CDCl₃) δ = 167.3, 142.3, 141.3, 139.5, 129.9, 125.4, 125.2, 122.2, 66.3, 65.2, 64.3, 51.5, 42.6, 32.9, 32.8, 29.8, 29.6, 29.1, 28.9, 28.5, 26.5, 25.5, 22.9, 21.7, 14.3.

Aziridine (**70**)

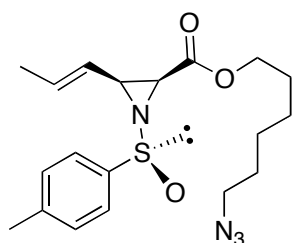


Chemical Formula: C₂₈H₄₄N₄O₃S
Molecular Weight: 516.7390

Following the same protocol as above for **69** led to the product **70** (yield: 239 mg, 0.46 mmol, 83%) as a yellow oil.

TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.42$; $^1\text{H NMR}$ (CDCl_3) $\delta = 7.59$ (m, 2H), 7.28 (m, 2H), 5.97 (m, 1H), 5.46 (m, 1H), 4.17 (t, $J = 6.64$ Hz, 1H), 3.99 (m, 2H), 3.26 (m, 5H), 2.39 (s, 3H), 2.05 (m, 2H), 1.56-1.25 (m, 22H), 0.87 (t, $J = 6.64$ Hz, 3H); $^{13}\text{C NMR}$ (CDCl_3) $\delta = 167.3, 142.3, 141.3, 139.5, 129.9, 125.4, 125.2, 122.2, 66.3, 65.2, 64.3, 51.5, 42.6, 32.9, 32.8, 29.8, 29.6, 29.1, 28.9, 28.5, 26.5, 25.5, 22.9, 21.7, 14.3$.

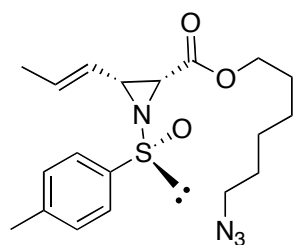
Aziridine (71)



Chemical Formula: $\text{C}_{19}\text{H}_{26}\text{N}_4\text{O}_3\text{S}$
Molecular Weight: 390.4997

Following the same protocol as above for **69** led to the product **71** (yield: 63 mg, 0.16 mmol, 63%) as a yellow oil.

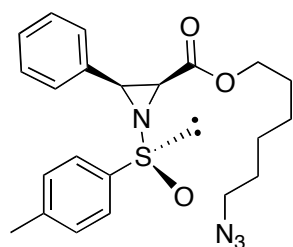
TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.44$; $^1\text{H NMR}$ (CDCl_3) $\delta = 7.57$ (m, 2H), 7.29 (m, 2H), 5.98 (m, 1H), 5.49 (m, 1H), 4.15 (m, 1H), 3.99 (m, 2H), 3.58 (m, 1H), 3.25 (m, 5H), 2.42 (s, 3H), 1.65-1.23 (m, 8H); $^{13}\text{C NMR}$ (CDCl_3) $\delta = 167.3, 142.9, 141.3, 134.0, 129.9, 125.4, 125.2, 123.6, 71.9, 68.5, 65.1, 64.3, 51.5, 42.5, 32.8, 28.9, 26.5, 25.5, 21.7, 18.3$.

Aziridine (72)

Chemical Formula: C₁₉H₂₆N₄O₃S
Molecular Weight: 390.4997

Following the same protocol as above for **69** led to the product **72** (yield: 70 mg, 0.18 mmol, 70%) as a yellow oil.

TLC (ethyl acetate/cyclohexane = 1:5): R_f = 0.44; ¹H NMR (CDCl₃) δ = 7.57 (m, 2H), 7.29 (m, 2H), 5.98 (m, 1H), 5.49 (m, 1H), 4.15 (m, 1H), 3.99 (m, 2H), 3.58 (m, 1H), 3.25 (m, 5H), 2.42 (s, 3H), 1.65-1.23 (m, 8H); ¹³C NMR (CDCl₃) δ = 167.3, 142.9, 141.3, 134.0, 129.9, 125.4, 125.2, 123.6, 71.9, 68.5, 65.1, 64.3, 51.5, 42.5, 32.8, 28.9, 26.5, 25.5, 21.7, 18.3.

Aziridine (73)

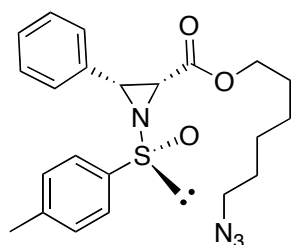
Chemical Formula: C₂₂H₂₆N₄O₃S
Molecular Weight: 426.5318

Following the same protocol as above for **69** led to the product **73** (yield: 215 mg, 0.51 mmol, 90%) as a yellow oil.

TLC (ethyl acetate/cyclohexane = 1:5): R_f = 0.27; ¹H NMR (CDCl₃) δ = 7.74 (m, 2H), 7.45 (m, 2H), 7.33 (m, 5H), 3.85 (m, 2H), 3.72 (m, 1H), 3.49 (d, J = 7.4 Hz,

1H), 3.19 (t, J = 7.04 Hz, 2H), 2.42 (s, 3H), 1.61-1.17 (m, 6H), 1.01 (m, 2H); ^{13}C NMR (CDCl_3) δ = 166.0, 142.4, 141.3, 133.0, 130.0, 128.4, 128.0, 125.5, 65.0, 51.5, 42.3, 35.3, 28.8, 28.4, 26.4, 25.3, 21.7.

Aziridine (74)

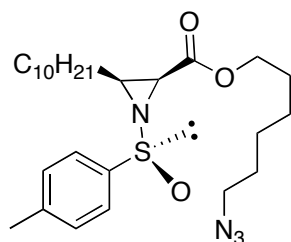


Chemical Formula: $\text{C}_{22}\text{H}_{26}\text{N}_4\text{O}_3\text{S}$
Molecular Weight: 426.5318

Following the same protocol as above for **69** led to the product **74** (yield: 210 mg, 0.49 mmol, 88%) as a yellow oil.

TLC (ethyl acetate/cyclohexane = 1:5): R_f = 0.27; ^1H NMR (CDCl_3) δ = 7.74 (m, 2H), 7.45 (m, 2H), 7.33 (m, 5H), 3.85 (m, 2H), 3.72 (m, 1H), 3.49 (d, J = 7.4 Hz, 1H), 3.19 (t, J = 7.04 Hz, 2H), 2.42 (s, 3H), 1.61-1.17 (m, 6H), 1.01 (m, 2H); ^{13}C NMR (CDCl_3) δ = 166.0, 142.4, 141.3, 133.0, 130.0, 128.4, 128.0, 125.5, 65.0, 51.5, 42.3, 35.3, 28.8, 28.4, 26.4, 25.3, 21.7.

Aziridine (75)

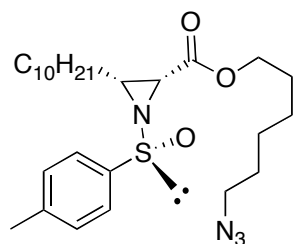


Chemical Formula: $\text{C}_{26}\text{H}_{42}\text{N}_4\text{O}_3\text{S}$
Molecular Weight: 490.7017

Following the same protocol as above for **69** led to the product **75** (yield: 84 mg, 0.17 mmol, 85%) as a light yellow oil.

TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.28$; $^1\text{H NMR}$ (CDCl_3) $\delta = 7.58$ (m, 2H), 7.32 (m, 2H), 4.18 (t, $J = 6.64$ Hz, 2H), 3.84 (m, 2H), 3.28 (m, 2H), 2.41 (m, 3H), 1.63-1.26 (m, 26H), 0.88 (t, $J = 6.64$ Hz, 3H); $^{13}\text{C NMR}$ (CDCl_3) $\delta = 167.5$, 129.9, 126.4, 125.5, 125.4, 66.3, 64.3, 51.5, 32.1, 29.8, 29.7, 29.6, 28.9, 28.5, 26.5, 26.0, 25.6, 22.9, 14.3.

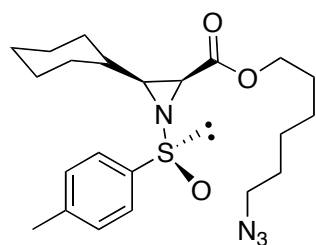
Aziridine (76)



Chemical Formula: $\text{C}_{26}\text{H}_{42}\text{N}_4\text{O}_3\text{S}$
Molecular Weight: 490.7017

Following the same protocol as above for **69** led to the product **75** (yield: 86 mg, 0.18 mmol, 87%) as a light yellow oil.

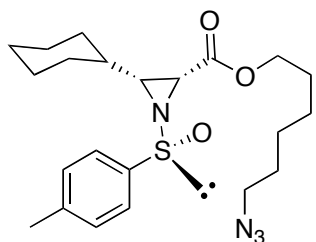
TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.28$; $^1\text{H NMR}$ (CDCl_3) $\delta = 7.58$ (m, 2H), 7.32 (m, 2H), 4.18 (t, $J = 6.64$ Hz, 2H), 3.84 (m, 2H), 3.28 (m, 2H), 2.41 (m, 3H), 1.63-1.26 (m, 26H), 0.88 (t, $J = 6.64$ Hz, 3H); $^{13}\text{C NMR}$ (CDCl_3) $\delta = 167.5$, 129.9, 126.4, 125.5, 125.4, 66.3, 64.3, 51.5, 32.1, 29.8, 29.7, 29.6, 28.9, 28.5, 26.5, 26.0, 25.6, 22.9, 14.3.

Aziridine (77)

Chemical Formula: C₂₂H₃₂N₄O₃S
Molecular Weight: 432.5795

Following the same protocol as above for **69** led to the product **77** (yield: 169 mg, 0.39 mmol, 70%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 1:5): R_f = 0.35; ¹H NMR (CDCl₃) δ = 7.59 (m, 2H), 7.29 (m, 2H), 4.18 (m, 1H), 3.99 (m, 2H), 3.58 (m, 1H), 3.25 (m, 2H), 2.42 (s, 3H), 1.62-1.24 (m, 19H); ¹³C NMR (CDCl₃) δ = 167.8, 142.9, 129.9, 129.8, 125.4, 125.2, 66.3, 65.2, 64.3, 51.5, 46.6, 36.2, 31.3, 30.1, 29.7, 28.9, 28.5, 26.5, 25.6, 21.7

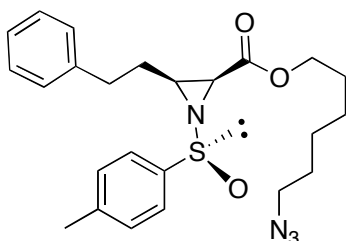
Aziridine (78)

Chemical Formula: C₂₂H₃₂N₄O₃S
Molecular Weight: 432.5795

Following the same protocol as above for **69** led to the product **78** (yield: 164 mg, 0.38 mmol, 68%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.35$; $^1\text{H NMR}$ (CDCl_3) $\delta = 7.59$ (m, 2H), 7.29 (m, 2H), 4.18 (m, 1H), 3.99 (m, 2H), 3.58 (m, 1H), 3.25 (m, 2H), 2.42 (s, 3H), 1.62-1.24 (m, 19H); $^{13}\text{C NMR}$ (CDCl_3) $\delta = 167.8, 142.9, 129.9, 129.8, 125.4, 125.2, 66.3, 65.2, 64.3, 51.5, 46.6, 36.2, 31.3, 30.1, 29.7, 28.9, 28.5, 26.5, 25.6, 21.7$

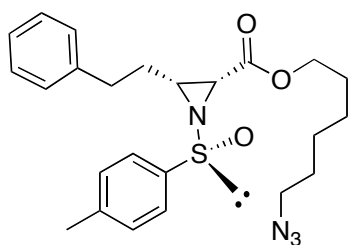
Aziridine (79)



Chemical Formula: $\text{C}_{24}\text{H}_{30}\text{N}_4\text{O}_3\text{S}$
Molecular Weight: 454.5850

Following the same protocol as above for **69** led to the product **79** (yield: 69 mg, 0.15 mmol, 50%) as a colourless oil.

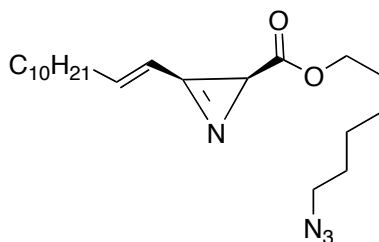
TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.29$; $^1\text{H NMR}$ (CDCl_3) $\delta = 7.59$ (m, 2H), 7.18 (m, 7H), 4.10 (m, 1H), 3.75 (m, 2H), 3.44 (m, 1H), 3.14 (m, 2H), 2.65 (m, 1H), 2.43 (m, 1H), 2.32 (s, 3H), 1.55-1.13 (m, 10H); $^{13}\text{C NMR}$ (CDCl_3) $\delta = 166.4, 142.2, 141.9, 138.9, 129.8, 128.8, 128.7, 128.6, 125.4, 66.3, 65.0, 64.4, 51.5, 43.4, 36.1, 35.5, 34.2, 30.4, 29.0, 28.5, 26.5, 25.5, 21.7$.

Aziridine (80)

Chemical Formula: C₂₄H₃₀N₄O₃S
Molecular Weight: 454.5850

Following the same protocol as above for **69** led to the product **80** (yield: 76 mg, 0.17 mmol, 55%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 1:5): R_f = 0.29; ¹H NMR (CDCl₃) δ = 7.59 (m, 2H), 7.18 (m, 7H), 4.10 (m, 1H), 3.75 (m, 2H), 3.44 (m, 1H), 3.14 (m, 2H), 2.65 (m, 1H), 2.43 (m, 1H), 2.32 (s, 3H), 1.55-1.13 (m, 10H); ¹³C NMR (CDCl₃) δ = 166.4, 142.2, 141.9, 138.9, 129.8, 128.8, 128.7, 128.6, 125.4, 66.3, 65.0, 64.4, 51.5, 43.4, 36.1, 35.5, 34.2, 30.4, 29.0, 28.5, 26.5, 25.5, 21.7.

Aziridine (81)

Chemical Formula: C₂₁H₃₆N₄O₂
Molecular Weight: 376.5361

To a solution of **69** (51.6 mg, 0.1 mmol) in anhydrous THF (1 ml), LDA (130 μl, 0.13 mmol, 1.0 M in THF) was added dropwise at -78 °C. The reaction was stirred at -78 °C for 15 min, at which time iodomethane (12.4 μl, 0.2 mmol) was added. After stirring for 20 min, the reaction was quenched with water (0.2 ml) and diluted

with ethyl acetate (5 ml). The organic phase was separated and aqueous phase was washed with ethyl acetate. The combined organic layer was washed with brine, dried and the crude product was purified by silica gel chromatography (ethyl acetate/cyclohexane = 5:1) to obtain product **81** (yield: 7 mg, 19 μ mol, 20%) as a brown oil.

TLC (ethyl acetate/cyclohexane = 1:5): R_f = 0.85.

General methods for the syntheses of compounds 82 – 86

Syntheses were carried out in a parallel reaction synthesizer without isolation of the reaction intermediates.

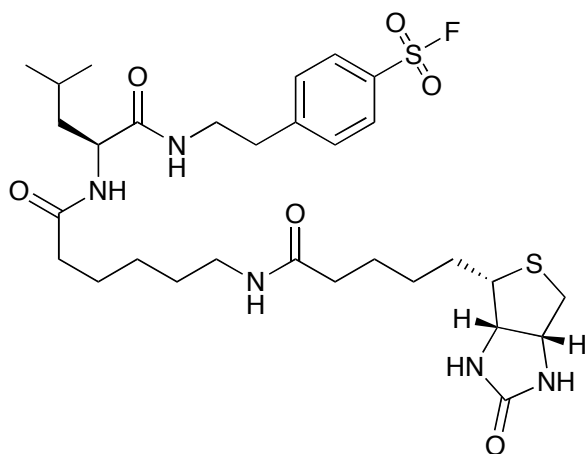
Amino acids (0.2 mmol) were weighted in 12 different reaction tubes. HBTU (1.1 g, 2.88 mmol), HOBT (440 mg, 2.88 mmol) and TEA (996 μ l, 7.2 mmol) were dissolved in DCM (final volume 12 ml) and aliquoted to each tube. The reaction mixture was stirred for 5 min and a solution of AEBSF (48 mg, 0.24 mmol) in DCM (0.25 ml) was added and stirred at room temperature for another 3 h (TLC control).

The solvent in each reaction tube was removed carefully at reduced pressure and 50% TFA in DCM (1 ml) was added. The resulting reaction mixtures were stirred for 3 h. Toluene (5 ml) was added and removed at reduced pressure.

22 (857 mg, 2.4 mmol), PyBOP (1.25 g, 2.4 mmol) and TEA (1.32 ml, 9.6 mmol) were dissolved in DMSO (final volume 6 ml) at room temperature. After stirring for 15 min, the DMSO solution (500 μ l) was aliquoted to each reaction tube. The reactions were completed overnight. The products were isolated after removing DMSO by precipitation with ethyl acetate/diethyl ether 5 ml (1:1) and washed with

sodium bicarbonate solution and ethyl ether, dried and purified by preparative HPLC.

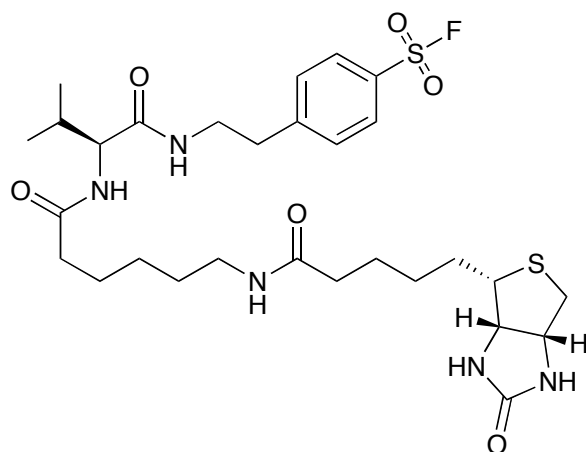
Biotinyl leucinyl AEBSF (**82**)



Chemical Formula: $C_{30}H_{46}FN_5O_6S_2$
Molecular Weight: 655.8445

Following the above protocol led to the product **82** (yield: 46 mg, 0.07 mmol, 35%) as a white solid.

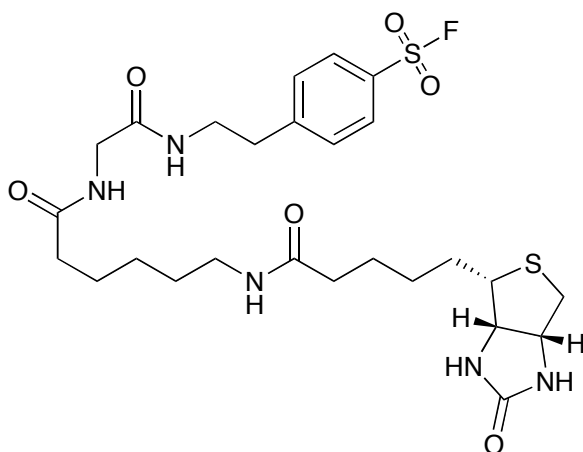
1H NMR (DMSO- d_6) δ = 8.04 (d, J = 8.2 Hz, 2H), 7.95 (t, J = 5.48 Hz, 1H), 7.79 (d, J = 8.2 Hz, 1H), 7.68 (t, J = 5.48 Hz, 1H), 7.58 (d, J = 8.2 Hz, 2H), 6.35 (s/s, 2H), 4.31 (m, 1H), 4.15 (m, 2H), 3.14 (t, J = 8.4 Hz, 1H), 3.08 (m, 1H), 2.96 (q, J = 6.84 Hz, 2H), 2.84 (t, J = 6.84 Hz, 2H), 2.79 (dd, J = 12.52, 5.08 Hz, 1H), 2.55 (d, J = 12.52 Hz, 1H), 2.04 (m, 4H), 1.61-1.26 (m, 14H), 0.92 (t, J = 7.24 Hz, 2H), 0.78 (d, J = 6.44 Hz, 3H), 0.74 (d, J = 6.44 Hz, 3H); ^{19}F NMR (DMSO- d_6) δ = -74.0, -199.2; MS (ESI): calcd. for $C_{30}H_{47}FN_5O_6S_2$ $[M+H]^+$, 656.84, found 656.20.

Biotinyl valinyl AEBSF (83)

Chemical Formula: C₂₉H₄₄FN₅O₆S₂
Molecular Weight: 641.8180

Following the above protocol led to the product **83** (yield: 36 mg, 0.06 mmol, 28%) as a white solid.

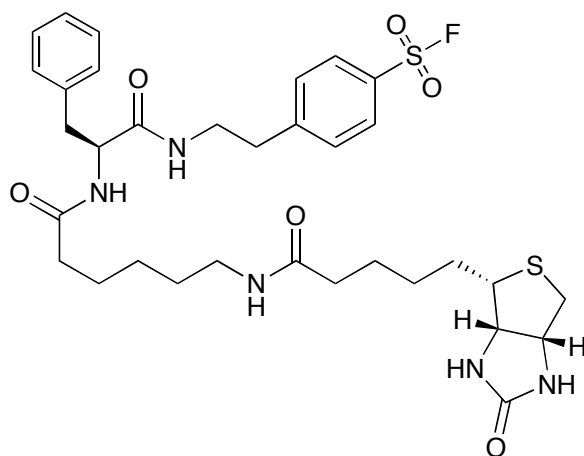
¹H NMR (DMSO-d₆) δ = 7.99 (d, J = 8.4 Hz, 3H), 7.70 (t, J = 3.32 Hz, 2H), 7.60 (d, J = 8.4 Hz, 2H), 6.35 (s/s, 2H), 4.27 (m, 1H), 4.10 (m, 1H), 4.00 (m, 1H), 3.06 (m, 1H), 2.96 (q, J = 6.84 Hz, 2H), 2.86 (t, J = 6.84 Hz, 2H), 2.79 (dd, J = 12.52, 5.08 Hz, 1H), 2.55 (d, J = 12.52 HZ, 1H), 2.08 (m, 4H), 1.61-1.26 (m, 15H), 0.70 (t, J = 6.64 Hz, 6H); ¹⁹F NMR (DMSO-d₆) δ = -74.1, -199.2; MS (ESI): calcd. for C₂₉H₄₅FN₅O₆S₂ [M+H]⁺, 642.82, found 642.27.

Biotinyl glycinyl AEBSF (84)

Chemical Formula: C₂₆H₃₈FN₅O₆S₂
Molecular Weight: 599.7382

Following the above protocol led to the product **84** (yield: 43 mg, 0.07 mmol, 36%) as a white solid.

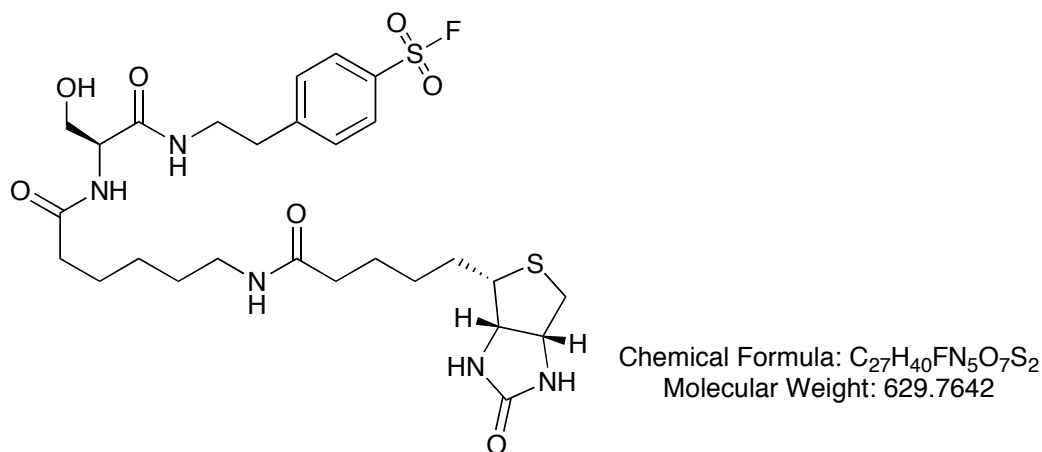
¹H NMR (DMSO-d₆) δ = 8.04 (d, J = 8.2 Hz, 2H), 7.95 (t, J = 5.48 Hz, 1H), 7.79 (d, J = 8.2 Hz, 1H), 7.68 (t, J = 5.48 Hz, 1H), 7.58 (d, J = 8.2 Hz, 2H), 6.35 (br s, 2H), 4.28 (m, 1H), 4.12 (m, 1H), 3.59 (d, J = 5.84 Hz, 2H), 3.33 (m, 3H), 3.09 (m, 2H), 2.99 (m, 1H), 2.86 (t, J = 6.84 Hz, 2H), 2.78 (dd, J = 12.52, 5.08 Hz, 1H), 2.55 (d, J = 12.52 Hz, 1H), 2.06 (m, 4H), 1.61-1.26 (m, 11H); ¹⁹F NMR (DMSO-d₆) δ = -75.0, -199.2; MS (ESI): calcd. for C₂₆H₃₉FN₅O₆S₂ [M+H]⁺, 600.74, found 600.20.

Biotinyl phenylalaninyl AEBSF (85)

Chemical Formula: C₃₃H₄₄FN₅O₆S₂
Molecular Weight: 689.8608

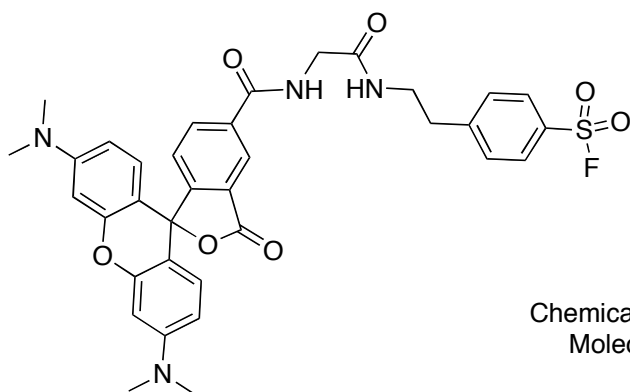
Following the above protocol to get the product **85** (yield: 45 mg, 0.07 mmol, 33%) as a white solid.

¹H NMR (DMSO-d₆) δ = 8.04 (m, 3H), 7.56 (d, J = 8.36 Hz, 2H), 7.16 (m, 7H), 6.35 (br s, 2H), 4.38 (m, 1H), 4.28 (m, 1H), 4.10 (m, 1H), 3.33 (m, 4H), 3.14 (t, J = 8.4 Hz, 2H), 3.08 (m, 1H), 2.96 (q, J = 6.84 Hz, 2H), 2.84 (t, J = 6.84 Hz, 2H), 2.66 (dd, J = 12.52, 5.08 Hz, 1H), 2.55 (d, J = 12.52 Hz, 1H), 2.04 (m, 4H), 1.61-1.26 (m, 8H), 0.92 (t, J = 7.24 Hz, 2H); ¹⁹F NMR (DMSO-d₆) δ = -74.8, -199.2; MS (ESI): calcd. for C₃₃H₄₅FN₅O₆S₂ [M+H]⁺, 690.86, found 690.27.

Biotinyl serinyl AEBSF (86)

Following the above protocol to get the product **86** (yield: 38 mg, 0.06 mmol, 30%) as a white solid.

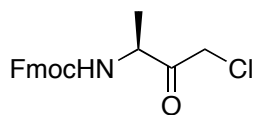
^1H NMR (DMSO- d_6) δ = 8.04 (d, J = 8.2 Hz, 2H), 7.95 (m, 1H), 7.71 (t, J = 5.48 Hz, 2H), 7.59 (d, J = 8.2 Hz, 2H), 6.35 (s/s, 2H), 4.30 (m, 1H), 4.26 (m, 1H), 4.12 (m, 1H), 3.34 (m, 2H), 3.14 (m, 4H), 2.96 (q, J = 6.84 Hz, 1H), 2.84 (t, J = 6.84 Hz, 2H), 2.79 (d/d, J = 12.52, 5.08 Hz, 1H), 2.55 (d, J = 12.52 Hz, 1H), 2.09 (m, 3H), 1.81 (m, 4H), 1.61-1.26 (m, 10H); ^{19}F NMR (DMSO- d_6) δ = -74.8, -199.1; MS (ESI): calcd. for $C_{27}H_{41}FN_5O_7S_2$ $[\text{M}+\text{H}]^+$, 630.76, found 630.20.

Rhodaminyl glycinyl AEBSF (87)

Chemical Formula: C₃₅H₃₃FN₄O₇S
Molecular Weight: 672.7225

Boc-Gly-OH (35 mg, 0.2 mmol), HOBt (27 mg, 0.2 mmol) and DIC (31 μ l, 0.2 mmol) were dissolved in DCM (2 ml). To this reaction mixture, AEBSF (48 mg, 0.2 mmol) was added and the reaction mixture was stirred for the following 3 h. After that, DCM was removed and 50% TFA in DCM (1 ml) was added to cleave the Boc protecting group within 1 h. The solvent was removed and TEA was added to neutralize excess TFA. 5(6)-Carboxytetramethylrhodamine N-succinimidyl ester (25 mg, 47.4 μ mol) in DCM (1 ml) was added, stirred overnight, evaporated and the product **87** was purified directly via HPLC (yield: 11 mg, 16 μ mol, 8.2%) as a violet solid.

LC-MS (ESI): $t_R = 7.32$ min, calcd. for C₃₅H₃₄FN₄O₇S [M+H]⁺, 673.72, found 673.24.

Fmoc-Ala-CMK (88)

Chemical Formula: C₁₉H₁₈ClNO₃
Molecular Weight: 343.8041

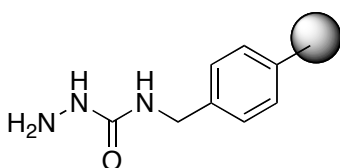
A 0.2 M solution of *N*^α-Fmoc alanine (1.65 g, 5 mmol) in anhydrous THF was stirred in an ice/ acetone bath at -10 °C. To this solution, *N*-methylmorpholine (686 μl, 6.25 mmol) and *iso*-butylchloroformate (752 μl, 5.75 mmol) were sequentially added. Immediately after the addition of the latter compound, a white precipitate formed. The reaction mixture was maintained at -10 °C for 25 min.

Diazomethane was generated *in situ* using the procedure described in the Aldrich Technical Bulletin (AL-180). Ethereal diazomethane (20 mmol) was transferred to the stirred solution of the mixed anhydride at 0 °C. The reaction mixture was warmed to room temperature over 3 h.

To obtain the corresponding chloromethyl ketones, a solution of concentrated hydrochloric acid and acetic acid 1:1 (15 mL) was added dropwise to the reaction mixture at 0 °C. Ethyl acetate was added and the organic layer was separated, washed with water, brine, saturated aqueous NaHCO₃, and dried over Na₂SO₄, concentrated under reduced pressure. The product **88** was obtained (yield: 1.82 g, 5 mmol, 100%) as a white solid without further purification.

LC-MS (ESI): *t*_R = 7.83 min, calcd. for C₁₉H₁₉ClNO₃ [M+H]⁺, 344.80, found 344.20.

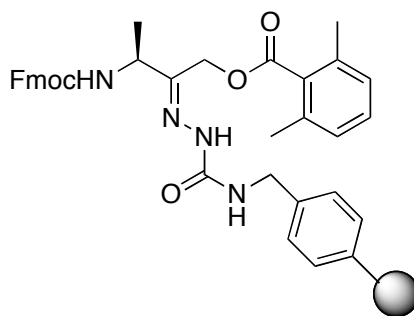
Semicarbazide linker (89)



Aminomethylpolystyrene resin (2.2 g, 1.1 mmol/g) was dried *in vacuo* overnight in a 12-mL polypropylene cartridge. The resin was presolvated with DMF for 30 min

and another 30 min with DCM. A 1 M solution of *N, N'*-Carbonyldiimidazole (1.95 g, 12 mmol) in DCM was added to the resin and the resin was shaken at room temperature for 3 h. The reagent was drained and the resin was washed with DCM followed by DMF. A 10 M solution of hydrazine (3.76 ml, 120 mmol) in DMF was added to the resin, and the resin was shaken at room temperature for 1 h. The resin was washed with DMF followed by DCM, dried *in vacuo*, and stored at -4 °C.

Acyoxymethyl hydrazone (90)



A 0.5 M solution of **88** (1 g, 3 mmol) in DMF was added to the resin. The cartridge was tightly sealed and shaken at 50 °C for 3 h. The resin was washed with DMF. A 0.5 M solution of 2,6-dimethylbenzoic acid (750 mg, 5 mmol) and potassium fluoride (580 mg, 10 mmol) were added to the resin. The resin was shaken at room temperature overnight. After the solution was removed, the resin was washed with DMF followed by DCM, and dried *in vacuo*. The resin load was estimated by UV absorption of cleaved Fmoc (loading: 0.565 mmol/g).

General protocol for syntheses of compounds 91-100

Protocol for deprotection

1. Add 20% piperidine in DMF and shake for 15 minutes.
2. Drain solution
3. Wash resin with DMF (3x)

Protocol for coupling

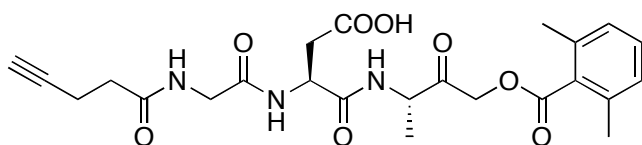
1. Add Fmoc-amino acid and HOBt to the cartridge (both 3 eq.).
2. Add DMF to the resin (so that final concentration will be 0.25 M of Fmoc-amino acid).
3. Add DIC (3 eq.) to the cartridge.
4. Shake for 2 h.
5. Drain solution
6. Wash with DMF (3x)
7. Wash with DCM (2x) to get rid of residual DMF
8. Kaiser test, if necessary.

Protocol for elongation

1. Wash resin with DMF.
2. Deprotect Fmoc following deprotection protocol.
3. Couple amino acid following coupling protocol.

Protocol for cleavage

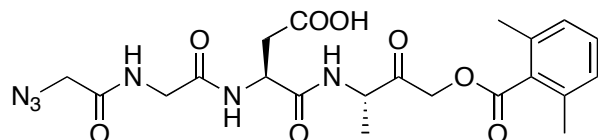
1. Make a stock of 95% TFA, 2.5% H₂O and 2.5% TIS.
2. Add TFA mixture to resin.
3. Cleave for 1h.
4. Collect the solution into a little flask.
5. Rinse the resin twice with 95% TFA to collect all the peptide.
6. Evaporate the TFA, dry peptide *in vacuo* and purify it with preparative HPLC.

Ak-GDA-AOMK (91)

Chemical Formula: C₂₄H₂₉N₃O₈
Molecular Weight: 487.5024

Following the above protocol led to the product **91** (yield: 2.4 mg, 4.9 μmol, 9.9%) as a brown solid.

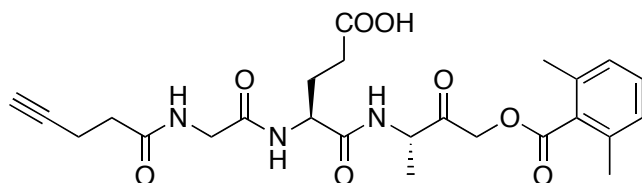
LC-MS (ESI): $t_R = 7.77$ min, calcd. for C₂₄H₃₀N₃O₈ [M+H]⁺, 488.50, found 488.07.

Az-GDA-AOMK (92)

Chemical Formula: C₂₁H₂₆N₆O₈
Molecular Weight: 490.4665

Following the above protocol led to get the product **92** (yield: 3 mg, 6 μmol, 12%) as a yellow solid.

LC-MS (ESI): $t_R = 7.72$ min, calcd. for C₂₁H₂₇N₆O₈ [M+H]⁺, 491.47, found 491.07.

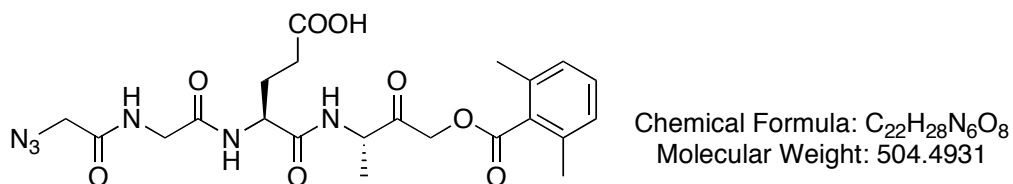
Ak-GEA-AOMK (93)

Chemical Formula: C₂₅H₃₁N₃O₈
Molecular Weight: 501.5289

Following the above protocol led to the product **93** (yield: 2.5 mg, 5 μ mol, 10%) as a brown solid.

LC-MS (ESI): $t_R = 7.75$ min, calcd. for $C_{25}H_{32}N_3O_8$ $[M+H]^+$, 502.53, found 502.13.

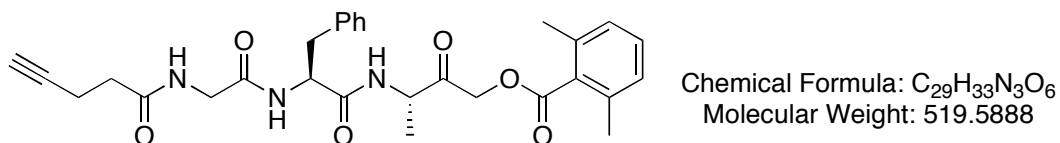
Az-GEA-AOMK (**94**)



Following the above protocol led to the product **94** (yield: 2 mg, 4 μ mol, 8%) as a brown solid.

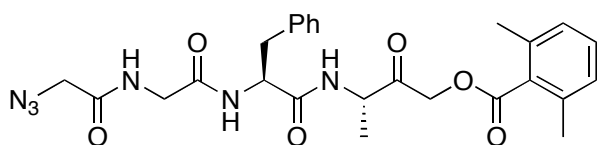
LC-MS (ESI): $t_R = 7.69$ min, calcd. for $C_{22}H_{29}N_6O_8$ $[M+H]^+$, 505.49, found 505.13.

Ak-GFA-AOMK (**95**)



Following the above protocol led to the product **95** (yield: 1.5 mg, 2.9 μ mol, 5.8%) as a white solid.

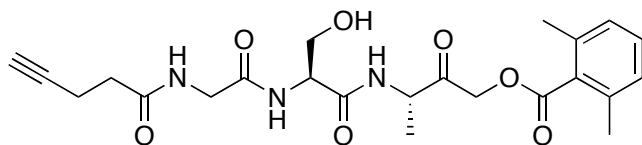
LC-MS (ESI): $t_R = 9.20$ min, calcd. for $C_{29}H_{34}N_3O_6$ $[M+H]^+$, 520.59, found 520.20.

Az-GFA-AOMK (96)

Chemical Formula: $C_{26}H_{30}N_6O_6$
Molecular Weight: 522.5530

Following the above protocol led to the product **96** (yield: 2 mg, 3.8 μ mol, 7.6%) as a yellow solid.

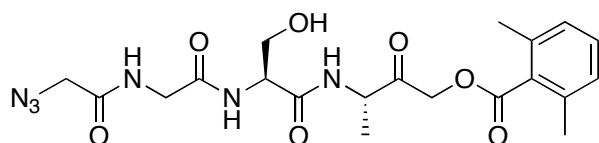
LC-MS (ESI): $t_R = 9.19$ min, calcd. for $C_{26}H_{33}N_6O_5$ $[M+H]^+$, 523.55, found 523.13.

Ak-GSA-AOMK (97)

Chemical Formula: $C_{23}H_{29}N_3O_7$
Molecular Weight: 459.4923

Following the above protocol led to the product **97** (yield: 1.8 mg, 3.9 μ mol, 7.8%) as a white solid.

LC-MS (ESI): $t_R = 7.61$ min, calcd. for $C_{23}H_{30}N_3O_7$ $[M+H]^+$, 460.49, found 460.13.

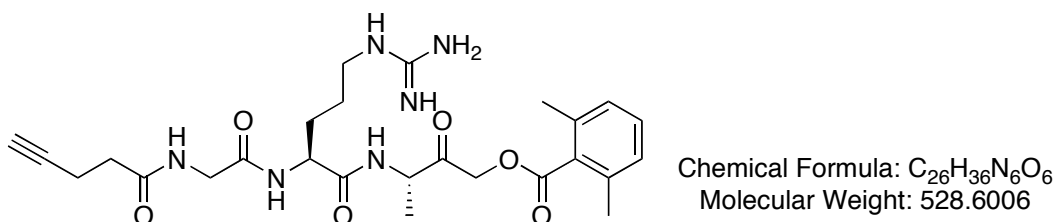
Az-GSA-AOMK (98)

Chemical Formula: $C_{20}H_{26}N_6O_7$
Molecular Weight: 462.4564

Following the above protocol led to the product **98** (yield: 2 mg, 4.3 μmol , 8.6%) as a white solid.

LC-MS (ESI): $t_R = 7.54$ min, calcd. for $\text{C}_{20}\text{H}_{27}\text{N}_6\text{O}_7$ $[\text{M}+\text{H}]^+$, 460.49, found 460.07.

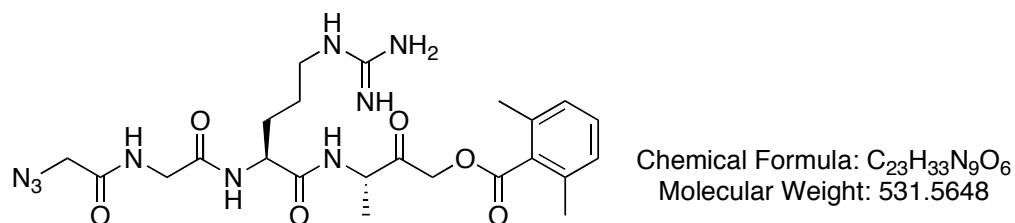
Ak-GRA-AOMK (**99**)



Following the above protocol led to the product **99** (yield: 4.1 mg, 7.7 μmol , 15.4%) as a brown solid.

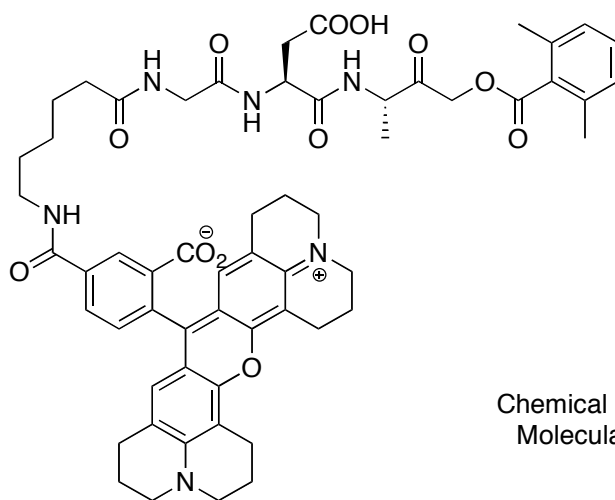
LC-MS (ESI): $t_R = 6.53$ min, calcd. for $\text{C}_{26}\text{H}_{37}\text{N}_6\text{O}_6$ $[\text{M}+\text{H}]^+$, 529.60, found 529.40.

Az-GRA-AOMK (**100**)



Following the above protocol led to the product **100** (yield: 4 mg, 7.5 μmol , 15%) as a brown solid.

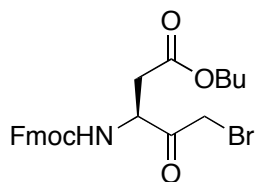
LC-MS (ESI): $t_R = 6.51$ min, calcd. for $\text{C}_{23}\text{H}_{34}\text{N}_9\text{O}_6$ $[\text{M}+\text{H}]^+$, 532.56, found 532.33.

Rho-GDA-AOMK (101)

Chemical Formula: $C_{58}H_{64}N_6O_{12}$
 Molecular Weight: 1037.1618

Following the above protocol led to the product **101** (yield: 6.2 mg, 6 μ mol, 12%) as a violet solid.

LC-MS (ESI): $t_R = 8.01$ min, calcd. for $C_{58}H_{65}N_6O_{12}$ $[M+H]^+$, 1038.16, found 1037.80.

Fmoc-Asp(OtBu)-BMK (102)

Chemical Formula: $C_{24}H_{26}BrNO_5$
 Molecular Weight: 488.3709

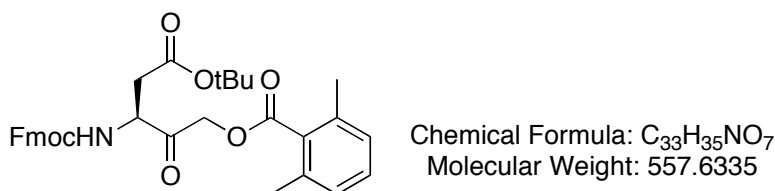
A 0.2 M solution of *N*-Fmoc (*O*tBu) aspartate (2 g, 5 mmol) in anhydrous THF was stirred in an ice/acetone bath at -10 °C. To this solution, *N*-methylmorpholine (686 μ l, 6.25 mmol) and *iso*-butylchloroformate (752 μ l, 5.75 mmol) were sequentially

added. Immediately after the addition of the latter compound, a white precipitate formed. The reaction mixture was maintained at -10°C for 25 min.

Diazomethane was generated *in situ* using the procedure described in the Aldrich Technical Bulletin (AL-180). Etheral diazomethane (20 mmol) was transferred to the stirred solution of the mixed anhydride at 0°C . The reaction mixture was warmed to room temperature over 3 h.

To obtain the corresponding bromomethyl ketones, 30% hydrogen bromide in acetic acid (10 ml) was added to the reaction mixture at 0°C . Workup was carried out as described for the chloromethyl ketone synthesis. The product **102** was obtained (yield: 2.4 g, 5 mmol, 100%) as a yellow oil without further purification.

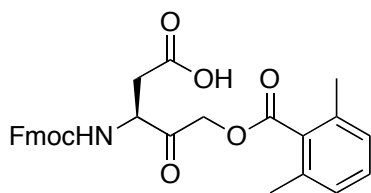
Fmoc-Asp(OtBu)-AOMK (**103**)



A 0.2 M solution of **102** (2.4 g, 5 mmol) in DMF was stirred at 0°C . To this solution, potassium fluoride (870 mg, 15 mmol) and 2,6-dimethylbenzoic acid (900 mg, 6 mmol) were added. The reaction mixture was warmed to room temperature and stirred overnight, diluted with ethyl acetate. The organic layer was separated and washed with water, brine, saturated aqueous NaHCO_3 and dried over MgSO_4 . The solvent was removed under reduced pressure. The crude product was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:5) to obtain pure product **103** (yield: 1.65 g, 2.96 mmol, 60%) as a white solid.

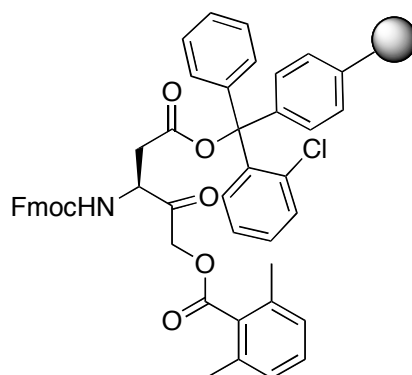
TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.2$; $^1\text{H NMR}$ (CDCl_3) $\delta = 7.82$ (d, $J = 8.2$ Hz, 2H), 7.63 (m, 2H), 7.41 (m, 2H), 7.33 (m, 2H), 7.20 (t, $J = 7.6$ Hz, 1H), 7.04 (d, $J = 7.64$ Hz, 2H), 5.89 (d, $J = 8.8$ Hz, 1H), 5.07 (q, $J = 16.8$ Hz, 3H), 4.65 (m, 2H), 4.24 (t, $J = 6.44$ Hz, 1H), 2.97 (dd, $J = 17.1, 4.88$ Hz, 1H), 2.91 (dd, $J = 17.1, 4.88$ Hz, 1H), 2.40 (s, 6H), 1.45 (s, 9H); $^{13}\text{C NMR}$ (CDCl_3) $\delta = 201.1, 169.0, 156.2, 143.7, 141.5, 141.5, 135.8, 132.7, 129.8, 127.9, 127.8, 125.2, 120.2, 120.2, 82.4, 67.3, 66.8, 54.9, 47.4, 36.7, 28.1, 27.1, 20.0$; LC-MS (ESI): $t_R = 11.86$ min, calcd. for $\text{C}_{33}\text{H}_{36}\text{NO}_7$ $[\text{M}+\text{H}]^+$, 558.63, found 558.32.

Fmoc-Asp(OH)-AOMK (104)

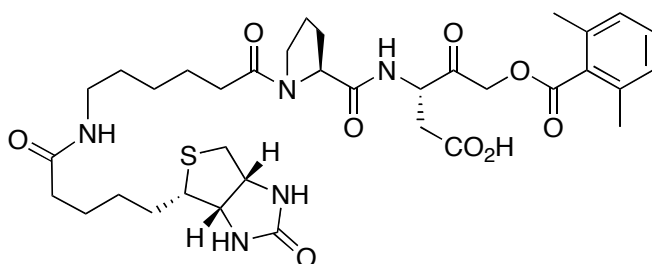


Chemical Formula: $\text{C}_{29}\text{H}_{27}\text{NO}_7$
Molecular Weight: 501.5272

A 0.2 M solution of **103** (1.65 g, 2.96 mmol) was dissolved in 25% TFA in DCM and allowed to stand for 30 min with occasional shaking. The reaction mixture was diluted with DCM. The cleavage solution was removed by coevaporation with toluene. The product was further dried *in vacuo*. The crude product **104** was used without further purification.

Fmoc-Asp(*O*-trityl resin)-AOMK (105)

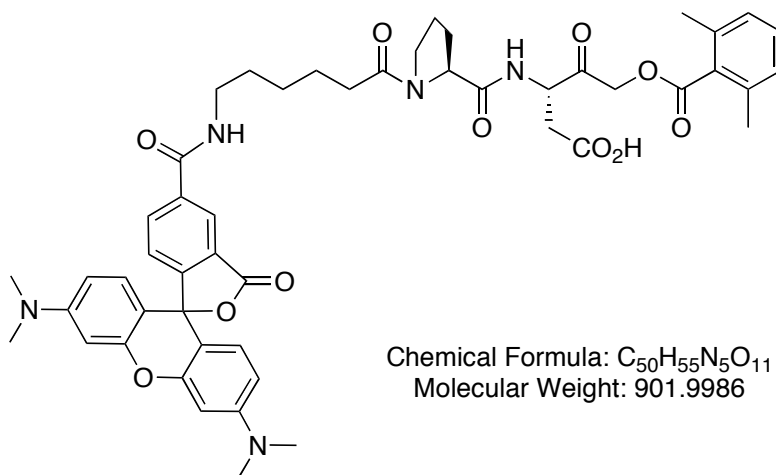
2-chlorotrityl chloride resin (1.5 g, 1.95 mmol) was washed with DCM. A 0.5 M solution of **104** (1 g, 2 mmol) and DIPEA (1.34 ml, 8 mmol) were added to the resin, shaking for 2 h. The resin was washed with DCM and DMF, yielding the loaded resin. Resin load was determined by UV absorption of cleaved Fmoc (loading: 0.77 mmol/g).

Bio-AHxPD-AOMK (106)

Chemical Formula: C₃₅H₄₉N₅O₉S
Molecular Weight: 715.8567

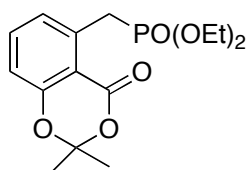
Following the above SPPS protocol led to the product **106** (yield: 75 mg, 0.1 mmol, 20%) as a white solid.

LC-MS (ESI): $t_R = 7.47$ min, calcd. for C₃₅H₅₀N₅O₉S [M+H]⁺, 716.86, found 716.47.

Rho-AHxPD-AOMK (107)

Following the above SPSS protocol led to the product **107** (yield: 5.2 mg, 6 μ mol, 12.2%) as a violet solid.

LC-MS (ESI): $t_R = 7.49$ min, calcd. for $C_{50}H_{56}N_5O_{11}$ $[M+H]^+$, 903.00, found 902.60.

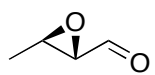
Diethyl (2, 2-dimethyl-4-oxo-4H-benzo[d][1,3]dioxin-5-yl)methyl phosphonate (112)

Chemical Formula: $C_{15}H_{21}O_6P$
Molecular Weight: 328.2974

To a stirred solution of **128** (1.1 g, 4 mmol) in anhydrous toluene (10 ml) at room temperature was added triethylphosphite (10 ml, 61.2 mmol) dropwise. The resulting mixture was refluxed for 2 h, allowed to cool to room temperature and directly transferred on top of a silica gel column. Flash chromatography (ethyl acetate/petroleum ether = 4:1) furnished the desired diethyl benzyl phosphonate **112** (yield: 1.3 g, 3.84 mmol, 96%) as a colourless oil.

TLC (ethyl acetate/petroleum ether = 4:1): $R_f = 0.1$; $^1\text{H NMR}$ (CDCl_3): $\delta = 7.42$ (t, $J = 7.8$ Hz, 1H), 7.06 (dd, $J = 3.0, 7.6$ Hz, 1H), 6.87 (ddd, $J = 1.1, 2.4, 8.2$ Hz, 1H), 4.08 (m, 4H), 3.99 (s, 1H); 3.93 (s, 1H); 1.71 (s, 6H); 1.24 (m, 6H); $^{13}\text{C NMR}$ (CDCl_3): $\delta = 160.82, 157.01, 136.53, 134.82, 126.36, 116.23, 113.04, 105.41, 62.07, 30.92, 25.58, 16.30$; IR (neat): 2987, 1730, 1016, 964 cm^{-1} ; HRMS: Calcd. for $\text{C}_{15}\text{H}_{22}\text{O}_6\text{P}$: $[\text{M}+1]^+$ 329.2974, found 329.1150.

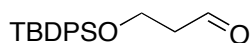
(2R, 3S)-3-methyloxirane-2-carbaldehyde (114)



Chemical Formula: $\text{C}_4\text{H}_6\text{O}_2$
Molecular Weight: 86.0892

To a solution of oxalylchloride (0.447 mL, 5.2 mmol) in DCM (15 ml) was added dropwise DMSO (0.695 mL, 10.6 mmol) at -78 °C. After ceasing of the gas evolution, a solution of **129** (0.3 g, 3.4 mmol) in DCM (2 mL) was added. After stirring for 90 min at -78 °C, TEA (2 mL, 14.3 mmol) was added dropwise. After 20 min the mixture was allowed to warm to room temperature. The precipitate was filtered off and the filtrate was concentrated *in vacuo*. The crude product was purified by column chromatography (petroleum ether/diethyl ether = 1:1) to obtain the product **114** (yield: 263 mg, 3 mmol, 90%) as a liquid. Because **114** has a very low boiling point, it can only be obtained in a diether ether solution. No NMR spectrum could be measured in good quality and the yield was only estimated.

3-(tert-Butyldiphenylsilyloxy)-propanal (124)



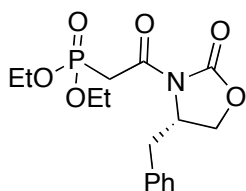
Chemical Formula: $\text{C}_{19}\text{H}_{24}\text{O}_2\text{Si}$
Molecular Weight: 312.4782

To a solution of oxalylchloride (0.54 ml, 6.36 mmol) in DCM (15 ml) was added DMSO (0.9 g, 12.7 mmol) dropwise at -78 °C. After ceasing of the gas evolution, a

solution of **130** (1 g, 3.18 mmol) in DCM (10 ml) was added. After 90 min of stirring at -78°C , TEA (2.65 ml 18.8 mmol) was added dropwise. After additional 20 min, the mixture was allowed to warm to room temperature. The reaction mixture was quenched with water, the organic layer was separated and the aqueous layer was extracted with DCM. The combined organic layers were successively washed with cold water, brine, dried over Na_2SO_4 and solvent was removed under reduced pressure. The crude product was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:2) to obtain pure product **124** (yield: 0.7 g, 2.2 mmol, 70%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 1:2): $R_f = 0.72$; ^1H NMR (CDCl_3): $\delta = 9.82$ (t, $J = 2.1$ Hz, 1H), 7.66 (d, $J = 1.56$ Hz, 4 H), 7.41 (m, 6H), 4.02 (t, $J = 5.9$ Hz, 2H), 2.61 (dt, $J = 2.1, 5.9$ Hz, 2H), 1.04 (s, 9H); ^{13}C NMR (CDCl_3): $\delta = 202.0, 135.7, 135.0, 133.4, 130.0, 127.9, 58.5, 46.5, 26.9, 19.3$; LC-MS (ESI): $t_R = 12.45$ min, calcd. for $\text{C}_{19}\text{H}_{25}\text{O}_2\text{Si}$ $[\text{M}+\text{H}]^+$, 313.48, found 312.86.

(S)-Diethyl 2-(4-benzyl-2-oxooxazolidin-3-yl)-2-oxoethylphosphonate (**125**)

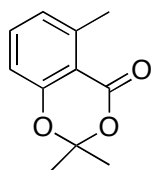


Chemical Formula: $\text{C}_{16}\text{H}_{22}\text{NO}_6\text{P}$
Molecular Weight: 355.3227

A stirred solution of **131** (2 g, 6.71 mmol) in triethylphosphite (10 ml) was refluxed for 12 h, allowed to cool to room temperature and directly transferred on top of a silica gel column. Flash chromatography (ethyl acetate/cyclohexane = 5:1) furnished the desired product **125** (yield: 1.91 g, 5.4 mmol, 80%) as a viscous colourless oil.

TLC (ethyl acetate/cyclohexane = 1:1): $R_f = 0.1$; $^1\text{H NMR}$ (CDCl_3): $\delta = 7.33\text{-}7.23$ (m, 5H), 4.70 (m, 1H), 4.18 (m, 6H), 3.80 (m, 2H), 3.35 (dd, $J = 3.3, 13.4$ Hz, 1H), 2.75 (dd, $J = 9.8, 13.4$ Hz, 1H), 1.35 (t, $J = 7.0$ Hz, 6H); $^{13}\text{C NMR}$ (CDCl_3): $\delta = 165.2, 164.9, 153.5, 135.3, 129.6, 129.1, 127.5, 66.2, 62.9, 55.6, 37.8, 35.1, 33.9, 16.5$; HRMS: Calcd. for $\text{C}_{16}\text{H}_{23}\text{NO}_6\text{P}$: $[\text{M}+1]^+$ 356.3227, found 356.1255.

2,2,5-trimethyl-4H-benzo[d]dioxin-4-one (127)



Chemical Formula: $\text{C}_{11}\text{H}_{12}\text{O}_3$
Molecular Weight: 192.2112

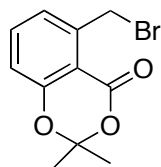
126 (1.93 g, 10.7 mmol) was dissolved in 10% NaOH/EtOH solution (20 ml) and refluxed for 2 h. The ethanol solution was acidified and extracted with acetyl acetate. The solvent was removed and the intermediate was used for the next step.

To a solution of 6-methylsalicylic acid (1.62 g, 10.7 mmol), DMAP (67 mg, 0.55 mmol) and acetone (0.96 ml, 13 mmol) in dimethoxyethane (3.6 ml) was added dropwise a solution of thionylchloride (1.1 ml, 15 mmol) in dimethoxyethane (0.5 ml). The temperature was maintained below 30 °C. The reaction mixture was stirred for 3 h, after which it was concentrated. The crude product was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:2) to obtain **127** (yield: 1.82 g, 9.5 mmol, 89%) as a colourless solid.

TLC (ethyl acetate/cyclohexane = 1:2): $R_f = 0.55$; $^1\text{H NMR}$ (CDCl_3): $\delta = 7.37$ (t, $J = 7.8$ Hz, 1H), 6.91 (d, $J = 7.6$ Hz, 1H), 6.79 (dq, $J = 0.5, 8.2$ Hz, 1H), 2.67 (s, 3H), 1.70 (s, 6H); $^{13}\text{C NMR}$ (CDCl_3): $\delta 160.60, 156.98, 143.53, 135.01, 125.70, 115.02,$

112.46, 105.13, 25.67, 21.94; LC-MS (ESI): $t_R = 11.04$ min, calcd. for $C_{11}H_{13}O_3$ $[M+H]^+$, 193.21, found 193.00.

5-(bromomethyl)-2,2-dimethyl-4H-benzo[d]dioxin-4-one (128)

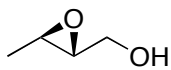


Chemical Formula: $C_{11}H_{11}BrO_3$
Molecular Weight: 271.1072

A solution of **127** (1.55 g, 8 mmol), *N*-bromosuccinimide (1.56 g, 8.8 mmol) and AIBN (197 mg, 0.8 mmol) in carbon tetrachloride (20 ml) was refluxed for 15 h. The reaction mixture was filtered, and the filtrate was concentrated *in vacuo*. The crude product was dissolved in ethyl acetate, and the organic phase was washed with water, brine and dried over Na_2SO_4 . The crude product was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:2) to obtain **128** (yield: 2.15 g, 7.9 mmol, 99%) as a white solid.

TLC (ethyl acetate/cyclohexane = 1:2): $R_f = 0.61$; 1H NMR ($CDCl_3$): $\delta = 7.47$ (dd, $J = 7.6, 8.2$ Hz, 1H), 7.16 (d, $J = 7.6$ Hz, 1H), 6.94 (dd, $J = 1.1, 8.2$ Hz, 1H), 5.05 (s, 2H), 1.72 (s, 6H); ^{13}C NMR ($CDCl_3$): $\delta = 159.70, 157.26, 141.76, 135.57, 125.75, 117.99, 111.45, 105.70, 30.85, 25.58$; LC-MS (ESI): $t_R = 10.05$ min, calcd. for $C_{11}H_{12}BrO_3$ $[M+H]^+$, 271.99, found 271.07.

[(2*S*, 3*S*)-3-methyloxiran-2-yl] methanol (129)

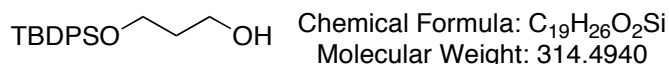


Chemical Formula: $C_4H_8O_2$
Molecular Weight: 88.1051

Molecular sieves (4 g, 4 Å) in a Schlenk flask were carefully dried under reduced pressure (ca. 0.1 mbar). In an argon atmosphere, DCM (150 ml) was cooled to -20 °C. *L*-(+) isopropyl tartrate (1 g, 4.3 mmol), (*E*)-2-buten-1-ol (5 g, 69.4 mmol) and titanium (IV)-isopropoxide (1 g, 10 mmol) were added. After stirring for 15 min at -20 °C, *tert*-butyl hydroperoxide solution (23.6 ml, 5.5 M) was added dropwise. The mixture was stirred for 2 h at -20 °C. Then, tributylphosphine (17.1 ml) was added to the reaction mixture until no peroxides were detected. The solvent was removed under reduced pressure; the residue was distilled using a Vigreux column (13 mbar, 64 °C) to give pure product **129** (yield: 3.3 g, 37.5 mmol, 54%) as colorless liquid.

B.p. = 64 °C at 13 mbar; ^1H NMR (CDCl_3): δ = 3.89 (dd, J = 12.6, 2.16 Hz, 1H), 3.61 (dd, J = 12.6, 3.92 Hz, 1H), 3.03 (q, J = 2.32 Hz, 1H), 2.88 (t, J = 2.16 Hz, 1H), 1.86 (br s, 1H), 1.33 (d, J = 5.28 Hz, 3H); ^{13}C NMR (CDCl_3): δ = 61.73, 59.52, 52.06, 17.29.

3-(*tert*-Butyldiphenylsilyloxy)-propanol (130)

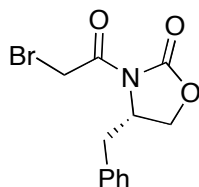


To a stirred solution of propanediol (5.8 g, 76.4 mmol) in DCM (50 ml) was added DIPEA (5.2 ml, 38.2 mmol) at 0 °C under an inert atmosphere. After 10 min, TBDPSCl (7.0 g, 25.5 mmol) in DCM was added at 0 °C and stirred for 12 h at room temperature. Water was added and the organic layer was separated. The aqueous layer was re-extracted with DCM. The combined organic layers were washed with brine, dried, evaporated and the residue was purified by silica gel

chromatography (ethyl acetate/cyclohexane = 1:2) to obtain **130** (yield: 7.2 g, 23 mmol, 90%) as a viscous liquid.

TLC (ethyl acetate/cyclohexane = 1:2): $R_f = 0.4$; $^1\text{H NMR}$ (CDCl_3): $\delta = 7.68$ (m, 4 H), 7.42 (m, 6H), 3.84 (m, 4H), 1.81 (m, 2H), 1.06 (s, 9H); $^{13}\text{C NMR}$ (CDCl_3): $\delta = 135.7, 133.4, 129.9, 127.9, 63.4, 62.1, 34.4, 27.0, 19.2$; LC-MS (ESI): $t_R = 12.04$ min, calcd. for $\text{C}_{19}\text{H}_{27}\text{O}_2\text{Si}$ $[\text{M}+\text{H}]^+$, 315.49, found 315.07.

(S)-4-Benzyl-3-(2-bromoacetyl)oxazolidin-2-one (131)



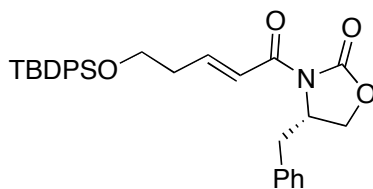
Chemical Formula: $\text{C}_{12}\text{H}_{12}\text{BrNO}_3$
Molecular Weight: 298.1326

To a solution of (2R)-4-(phenyl methyl)-2-oxazolidinone (4 g, 22.5 mmol) in anhydrous THF (50 ml) at $-78\text{ }^\circ\text{C}$ was added *n*-butyl lithium (15.5 ml, 31.5 mmol, 1.6 M in hexane), followed by bromoacetyl bromide (2.76 ml, 24.8 mmol). The solution was stirred at $-78\text{ }^\circ\text{C}$ for 10 min and then the cooling bath was removed. After 20 min, the reaction was quenched by addition of sat. aqueous NH_4Cl (50 ml) and the residue was extracted with DCM. The organic phases were dried over Na_2SO_4 and concentrated. The crude product was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:2) to obtain **131** (yield: 6.3 g, 21 mmol, 94%) as a yellow oil.

TLC (ethyl acetate/cyclohexane = 1:2): $R_f = 0.28$; $^1\text{H NMR}$ (CDCl_3 , 400 MHz): $\delta = 7.34$ (m, 3H), 7.21 (m, 2H), 4.70 (m, 1H), 4.55 (d, $J = 12.7$ Hz, 1H), 4.52 (d, $J = 12.7$ Hz, 1H), 4.25 (m, 2H), 3.33 (dd, $J = 3.3, 13.4$ Hz, 1H), 2.81 (dd, $J = 9.5, 13.4$

Hz, 1H); ^{13}C NMR (CDCl_3): $\delta = 166.1, 153.1, 134.9, 129.6, 129.2, 127.7, 66.8, 55.6, 37.7, 28.3$; LC-MS (ESI): $t_{\text{R}} = 9.46$ min, calcd. for $\text{C}_{12}\text{H}_{13}\text{BrNO}_3$ $[\text{M}+\text{H}]^+$, 298.56, found 298.27.

(S,E)-4-Benzyl-3-(5-(tert-butyldiphenylsilyloxy)pent-2-enoyl)oxazolidin-2-one (132)

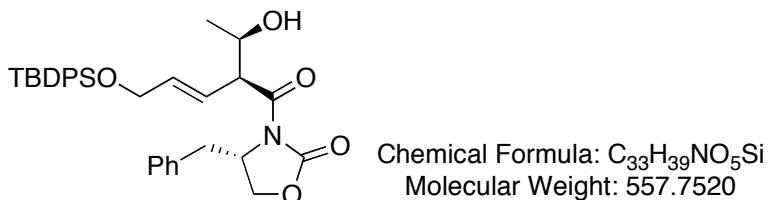


Chemical Formula: $\text{C}_{31}\text{H}_{35}\text{NO}_4\text{Si}$
Molecular Weight: 513.6994

To a solution of **125** (1 g, 3.2 mmol) in anhydrous THF (20 ml) at 0 °C was added NaHMDS (0.586 g, 3.2 mmol), leading to a hazy orange solution, to which was added a solution of **124** (1.13 g, 3.2 mmol) in anhydrous THF (14 ml). The reaction mixture was allowed to warm to room temperature and stirred for 1 h before it was poured into ethyl acetate and sat. aqueous NH_4Cl . The layers were separated and the aqueous layer was extracted with ethyl acetate. The combined organic extracts were dried with Na_2SO_4 , and concentrated *in vacuo*. The crude product was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:2) to obtain **132** (yield: 1.4 g, 2.7 mmol, 85%) as a colourless oil.

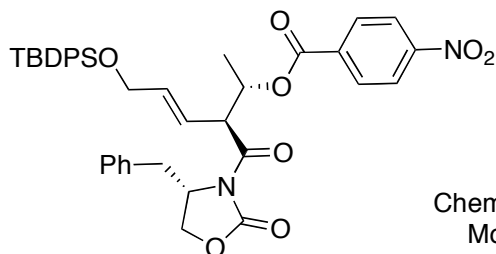
TLC (ethyl acetate/cyclohexane = 1:2): $R_f = 0.65$; ^1H NMR (CDCl_3): $\delta = 7.68$ (m, 4H), 7.46-7.19 (m, 13H), 4.77-4.69 (m, 1H), 4.23-4.14 (m, 2H), 3.81 (t, $J = 6.3$ Hz, 2H), 3.33 (dd, $J = 3.2, 13.3$ Hz, 1H), 2.80 (dd, $J = 9.5, 13.3$ Hz, 1H), 2.54 (q, $J = 6.3$ Hz, 2H), 1.06 (s, 9H); ^{13}C NMR (CDCl_3): $\delta = 164.8, 153.3, 148.4, 135.6, 135.4, 133.6, 133.6, 129.7, 129.4, 128.9, 127.7, 127.3, 122.0, 66.1, 62.4, 55.3, 37.9, 36.0, 26.8, 19.2$; HRMS: Calcd. for $\text{C}_{31}\text{H}_{36}\text{NO}_4\text{Si}$ $[\text{M}+\text{H}]^+$ 514.6994, found 514.2404.

(S)-4-Benzyl-3-((S)-5-(tert-butyldiphenylsilyloxy)-2-((R)-1-hydroxyethyl)pent-3-enoyl) oxazolidin-2-one (133)



To a solution of **132** (1.2 g, 2.34 mmol) in anhydrous DCM (20 ml), TEA (0.445 ml, 3.26 mmol) and Bu₂BOTf (2.8 mL, 2.8 mmol, 1.0 M in DCM) were added at -78 °C. The mixture was stirred for 1 h at -78 °C and 1 h at 0 °C. Acetaldehyde (0.262 ml, 4.6 mmol) in DCM (1 mL) was added dropwise at -78 °C. The mixture was stirred for 2 h at -78 °C and 2 h at 0 °C. The reaction was quenched with a mixture of phosphate buffer pH 7, MeOH and H₂O₂ (1:1:1, 15 ml) and the mixture was kept under vigorous stirring overnight. The product was then extracted with ethyl acetate and the combined extracts were washed with water and brine, dried over Na₂SO₄, and concentrated under reduced pressure. The crude product was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:2) to obtain **133** (yield: 1.2 g, 2.1 mmol, 91%) as a light yellow oil.

$[\alpha]_D^{20} = +18.97$ (c = 0.78, CDCl₃); TLC (ethyl acetate/cyclohexane = 1:2): R_f = 0.16; ¹H NMR (CDCl₃): δ = 7.72 (m, 4H), 7.46 (m, 6H), 7.29 (m, 3H), 7.18 (m, 2H), 5.94 (d, J = 6.08 Hz, 1H), 4.72 (m, 1H), 4.52 (m, 1H), 4.28-4.14 (m, 5H), 3.22 (dd, J = 9.36, 13.36 Hz, 1H), 2.70 (dd, J = 9.36, 13.36 Hz, 1H), 1.20 (d, J = 6.24 Hz, 3H), 1.07 (s, 9H); ¹³C NMR (CDCl₃): δ = 174.3, 153.1, 136.3, 135.7, 135.1, 133.7, 129.8, 129.60, 127.8, 123.9, 123.0, 66.1, 64.2, 60.5, 55.2, 52.2, 48.7, 37.7, 27.0, 20.0, 19.3; HRMS: Calcd. for C₃₃H₄₀NO₅Si [M+H]⁺ 558.7520, found 558.2667.

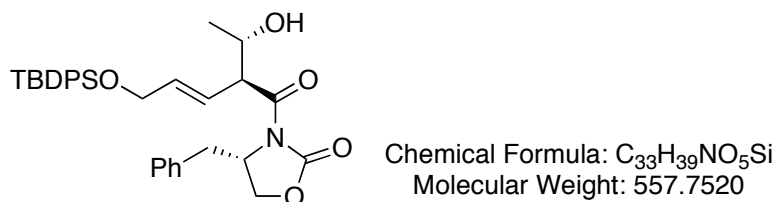
(S)-4-Benzyl-3-((S)-5-(tert-butyl-diphenylsilyloxy)-2-((S)-1-para-nitrobenzoicyclopent-3-enyl)oxazolidin-2-one (134)

Chemical Formula: C₄₀H₄₂N₂O₈Si
Molecular Weight: 706.8556

To an ice-cooled solution of **133** (0.84 g, 1.5 mmol) in THF (50 ml) were added *para*-nitrobenzoic acid (0.5 g, 3 mmol), triphenylphosphine (0.79 g, 3 mmol) and dropwise di-*iso*-propyl azodicarboxylate (627 μ l, 3 mmol). The reaction mixture was stirred for 30 min at 0 °C and 30 min at room temperature. After concentration *in vacuo*, the residue was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:2) to obtain **134** (yield: 0.96 g, 1.36 mmol, 91%) as a yellow oil.

TLC (ethyl acetate/cyclohexane = 1:2): R_f = 0.5; ¹H NMR (CDCl₃): δ = 7.65 (m, 5H), 7.46-7.32 (m, 14H), 6.69 (m, 1H), 5.85 (m, 1H), 4.69 (m, 1H), 4.29 (m, 2H), 4.14 (m, 4H), 3.42 (dd, J = 9.36, 13.36 Hz, 1H), 2.72 (dd, J = 9.36, 13.36 Hz, 1H), 1.86 (d, J = 6.24 Hz, 3H), 1.07 (s, 9H); ¹³C NMR (CDCl₃): δ = 181.4, 169.5, 135.3, 134.9, 134.2, 133.2, 130.5, 129.4, 129.1, 128.7, 127.4, 121.8, 121.4, 65.8, 63.8, 55.0, 37.6, 26.6, 21.3, 18.9, 13.3; LC-MS (ESI): t_R = 13.10 min, calcd. for C₄₀H₄₃N₂O₈Si [M+H]⁺, 707.86, found 707.32.

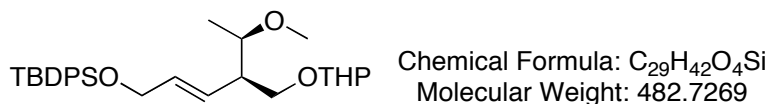
(S)-4-benzyl-3-((S)-5-(tert-butyl-diphenylsilyloxy)-2-((S)-1-hydroxyethyl)pent-3-enoyl) oxazolidin-2-one (135)



To a solution of **134** (0.96 g, 1.36 mmol) in MeOH (50 mL) was added K₂CO₃ (0.414 g, 3 mmol). After stirring for 2 h at room temperature, the reaction mixture was filtered and then evaporated. The residue was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:2) to obtain **135** (yield: 0.6 g, 1.07 mmol, 79%) as a colourless oil.

$[\alpha]_D^{20} = -5.19$ (c = 0.67, CDCl₃); TLC (ethyl acetate/cyclohexane = 1:2): R_f = 0.2; ¹H NMR (CDCl₃): δ = 7.72 (m, 4H), 7.46 (m, 6H), 7.29 (m, 3H), 7.18 (m, 2H), 6.52 (m, 2H), 5.68 (m, 2H), 4.24 (m, 3H), 4.14 (m, 1H), 3.66 (m, 2H), 2.86 (m, 2H), 2.45 (br s, 1H), 1.76 (d, J = 6.24 Hz, 3H), 1.08 (s, 9H); ¹³C NMR (CDCl₃): δ = 169.4, 137.7, 135.7, 135.6, 134.8, 133.6, 132.3, 129.9, 129.3, 128.8, 126.9, 122.7, 121.4, 69.8, 65.2, 64.0, 53.5, 37.0, 27.0, 19.4, 14.0; LC-MS (ESI): t_R = 12.43 min, calcd. for C₃₃H₄₀NO₅Si [M+H]⁺, 558.75, found 558.42.

tert-Butyl-((4R,5R)-5-methoxy-4-((tetrahydro-2H-pyran-2-yloxy)methyl)hex-2-enyloxy)diphenylsilane (137)

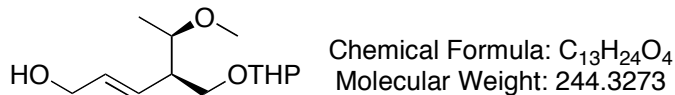


To a solution of **136** (0.8 g, 1.4 mmol) and methanol (115 μ l, 2.8 mmol) in diethyl ether (25 ml) was added LiBH₄ (62 mg, 2.8 mmol) at 0 °C. After stirring for 2 h, the resulting mixture was warmed to room temperature and stirred for 1 h. The reaction mixture was quenched with sat. NH₄Cl and extracted with ethyl acetate. The combined organic extracts were washed with brine, dried over Na₂SO₄ and concentrated *in vacuo*. The residue was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:2) to obtain the intermediate (yield: 0.52 g, 1.3 mmol, 93%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 1:2): R_f = 0.5;

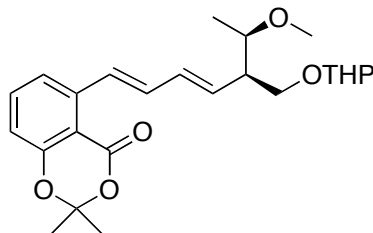
To a solution of the above intermediate (0.46 g, 1.14 mmol) in DCM (20 ml) were added DHP (0.32 mL, 3.5 mmol) and PPTS (28 mg, 0.11 mmol) at room temperature. After 5 h, the mixture was quenched by addition of NaHCO₃, the organic layer was separated and the aqueous layer was re-extracted. The combined organic layers were dried over Na₂SO₄, concentrated and the crude product was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:9) to obtain product **137** (yield: 0.43 g, 0.9 mmol, 80%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 1:9): R_f = 0.67; ¹H NMR (CDCl₃): δ = 7.68 (m, 4H), 7.38 (m, 6H), 5.82 (m, 1H), 5.68 (m, 1H), 4.55 (m, 1H), 4.27 (m, 1H), 4.18 (m, 1H), 3.73 (m, 2H), 3.54-3.33 (m, 3H), 3.31 (s, 3 H), 2.48 (m, 1H), 1.87-1.43 (m, 6H), 1.10 (d, J = 6.3 Hz, 3H), 1.04 (s, 9H); ¹³C NMR (CDCl₃): δ = 134.6, 133.9, 132.0, 129.5, 128.3, 127.6, 99.0, 98.8, 76.5, 67.9, 67.8, 64.7, 62.2, 56.7, 48.2, 47.9, 30.7, 26.9, 26.8, 25.5, 19.5, 19.2, 16.6.

(4*R*, 5*R*)-5-Methoxy-4-((tetrahydro-2*H*-pyran-2-yl)oxy)methyl)hex-2-en-1-ol (138)

To a solution of **137** (0.36 g, 0.74 mmol) in THF (30 ml), TBAF (0.26 g, 0.82 mmol) was added at room temperature. After 3 h, the reaction mixture was quenched by addition of sat. aqueous NH₄Cl and extracted with ethyl acetate. The organic layer was washed with water and brine, dried over Na₂SO₄, and concentrated in *vacuo*. The crude product was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:2) to obtain product **138** (yield: 0.15 g, 0.63 mmol, 85%) as a yellow oil.

TLC (ethyl acetate/cyclohexane = 1:2): R_f = 0.23; ¹H NMR (Acetone-d₆): δ = 5.68 (m, 1H), 5.58 (m, 1H), 4.56 (t, J = 8.4 Hz, 1H), 4.03 (m, 2H), 3.80 (m, 1H), 3.55 (m, 2H), 3.45 (m, 1H), 3.28 (s, 3H), 3.25 (m, 1H), 2.76 (t, J = 1.0 Hz, 1H); 2.35 (m, 1H), 1.85-1.40 (m, 6H), 1.07 (d, J = 8.2 Hz, 3H); ¹³C NMR (Acetone-D₆): δ = 135.3, 129.6, 100.5, 77.7, 69.7, 64.3, 63.3, 57.7, 50.3, 32.5, 27.4, 21.2, 17.9; LC-MS (ESI): t_R = 7.32 min, calcd. for C₁₃H₂₅O₄ [M+H]⁺, 245.33, found 245.01.

5-((1E,3E,5R,6R)-6-Methoxy-5-((tetrahydro-2H-pyran-2-yloxy)methyl)hepta-1,3-dienyl)-2,2-dimethyl-4H-benzo[d][1,3]dioxin-4-one (139)

Chemical Formula: C₂₄H₃₂O₆
Molecular Weight: 416.5073

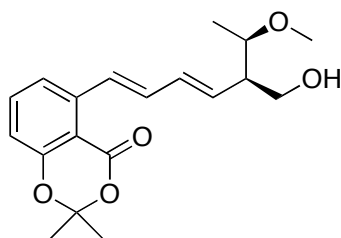
To a solution of **138** (120 mg, 0.48 mmol) in DCM (15 mL) was added NaHCO₃ (81 mg, 1 mmol) and DMP (312 mg, 0.73 mmol). The mixture was stirred at 0 °C for 45 min and quenched by addition of aqueous Na₂S₂O₃ and aqueous NaHCO₃. The layers were separated and the organic phase was extracted with diethyl ether. The combined organic layers were washed with brine, dried over Na₂SO₄ and concentrated *in vacuo* to give the crude aldehyde intermediate (105 mg) which was used immediately.

To a solution of **112** (141 mg, 0.43 mmol) and the above aldehyde intermediate (105 mg, 0.43 mmol) in anhydrous THF (10 ml) was added KO^tBu (50 mg, 0.45 mmol) at 0 °C and stirred for 1 h. The reaction mixture was quenched by addition of NH₄Cl and extracted with ethyl acetate. The organic layer was separated, dried, concentrated and the crude product was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:1) to obtain product **139** (yield: 100 mg, 0.24 mmol, 57%) as a colourless oil.

TLC (ethyl acetate/cyclohexane = 1:1): R_f = 0.62; ¹H NMR (CDCl₃): δ = 7.63 (d, J = 14.7 Hz, 1H), 7.42 (t, J = 7.9 Hz, 1H), 7.29 (d, J = 8 Hz, 1H), 6.80 (q, J = 8.2 Hz, 2H), 6.42 (m, 1H), 5.85 (m, 1H), 4.59 (t, J = 3.4 Hz, 1H), 3.89 (m, 2H), 3.75 (m,

3H), 3.34 (s, 3H), 2.48 (m, 1H), 1.88-1.74 (m, 2H), 1.70 (s, 6H), 1.64-1.46 (m, 4H), 1.14 (d, $J = 6.8$ Hz, 3H); ^{13}C NMR (CDCl_3): $\delta = 157.1, 135.1, 134.4, 134.3, 133.9, 128.6, 120.8, 115.9, 105.2, 99.3, 98.9, 68.1, 62.4, 56.9, 49.1, 30.8, 27.0, 25.8, 19.7, 17.0$.

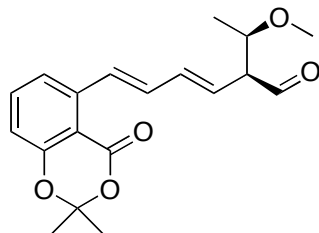
5-((1*E*,3*E*,5*R*,6*R*)-5-(Hydroxymethyl)-6-methoxyhepta-1,3-dienyl)-2,2-dimethyl-4*H*-benzo[*d*][1,3]dioxin-4-one (140)



Chemical Formula: $\text{C}_{19}\text{H}_{24}\text{O}_5$
Molecular Weight: 332.3909

To a stirred solution of **139** (90 mg, 0.22 mmol) in methanol (15 ml) were added catalytic amounts of PPTS (12 mg) at 0°C under an inert atmosphere. The reaction mixture was stirred for 2 days at room temperature. The solvent was removed under reduced pressure and the residue was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:1) to obtain product **140** (yield: 46 mg, 0.14 mmol, 66%) as a colourless oil.

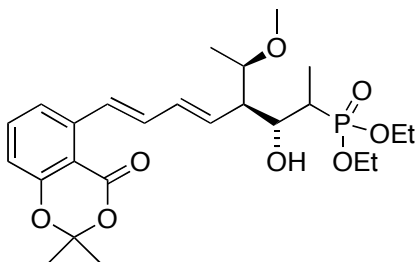
TLC (ethyl acetate/cyclohexane = 1:1): $R_f = 0.21$; ^1H NMR (CDCl_3): $\delta = 7.65$ (d, $J = 15.6$ Hz, 1H), 7.43 (t, $J = 7.9$ Hz, 1H), 7.29 (d, $J = 7.9$ Hz, 1H), 6.80 (m, 2H), 6.42 (m, 1H), 5.83 (dd, $J = 9.1, 15.2$ Hz, 1H), 3.82-3.61 (m, 3H), 3.36 (s, 3H), 2.52 (m, 1H), 1.70 (s, 6H), 1.16 (d, $J = 6.24$ Hz, 3H); ^{13}C NMR (CDCl_3): $\delta = 160.6, 157.1, 141.9, 135.2, 134.3, 133.2, 132.5, 129.4, 128.1, 120.8, 116.1, 105.3, 79.2, 64.6, 56.7, 50.1, 25.8, 16.2$; IR (neat): 3435, 2928, 1725, 1042 cm^{-1} ; LC-MS (ESI): $t_R = 8.24$ min, calcd. for $\text{C}_{19}\text{H}_{25}\text{O}_5$ $[\text{M}+\text{H}]^+$, 333.39, found 333.22.

5-((1*E*,3*E*,5*R*,6*R*)-5-Formyl-6-methoxyhepta-1,3-dienyl)-2,2-dimethyl-4*H*-benzo[*d*][1,3]dioxin-4-one (141)

Chemical Formula: C₁₉H₂₂O₅
Molecular Weight: 330.3750

To a solution of **140** (40 mg, 0.12 mmol) in DCM (5 ml) was added NaHCO₃ (21 mg, 0.25 mmol) and DMP (0.08 g, 0.18 mmol). The resulting mixture was stirred at 0 °C for 30 min and directly transferred on top of a silica gel column. Flash chromatography (ethyl acetate/cyclohexane = 1:1) furnished the desired **141** (yield: 33 mg, 0.1 mmol, 86%) as a colourless oil, which was used immediately.

TLC (ethyl acetate/cyclohexane = 1:1): R_f = 0.56.

5-((1*E*,3*E*,5*R*,6*R*)-5-(1-Hydroxyl-2-diethylphosphonyl-propyl)-6-methoxyhepta-1,3-dienyl)-2,2-dimethyl-4*H*-benzo[*d*][1,3]dioxin-4-one (142)

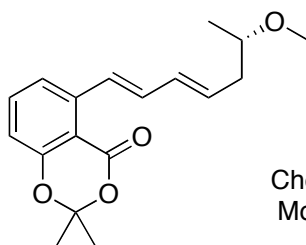
Chemical Formula: C₂₅H₃₇O₈P
Molecular Weight: 496.5302

To a solution of diethylethylphosphonate (50 mg, 0.3 mmol) in THF (1 ml) at -78 °C was added *n*-BuLi (125 μl, 0.2 mmol, 1.6 M in hexanes) dropwise. The reaction mixture was stirred for 15 min and then a solution of **141** (33 mg, 0.1 mmol) in

THF (1 ml) was added dropwise. The dark yellow solution was stirred at - 78 °C for 1h, and a sat. NH₄Cl solution (2 ml) was added. The mixture was diluted with water (2 ml) and the organic layer was separated, dried over Na₂SO₄, and concentrated *in vacuo*. The residue was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:1) to obtain product **142** (yield: 11 mg, 23 μmol, 23%) as a yellow oil, which was impure due to remaining impurity.

TLC (ethyl acetate/cyclohexane = 1:1): R_f = 0.47; LC-MS (ESI): t_R = 7.67 min, calcd. for C₂₅H₃₈O₈P [M+H]⁺, 497.53, found 497.26.

5-((1E,3E,6R)-6-Methoxyhepta-1,3-dienyl)-2,2-dimethyl-4H-benzo[d][1,3]dioxin-4-one (144)

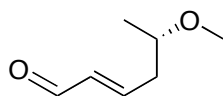


Chemical Formula: C₁₈H₂₂O₄
Molecular Weight: 302.3649

To a solution of **112** (435 mg, 1.33 mmol) in anhydrous THF (5 ml) was added KO^tBu (145 mg, 1.3 mmol) at 0 °C. After the reaction mixture turns red, **146** (170 mg, 1.33 mmol) in anhydrous THF (4 ml) was added. The reaction mixture was stirred at 0 °C for 1 h and at room temperature for 1 h, quenched by addition of sat. NH₄Cl and extracted with ethyl acetate. The organic layer was separated, dried over Na₂SO₄, and concentrated *in vacuo*. The residue was purified by silica gel chromatography (ethyl acetate/cyclohexane = 1:5) to obtain pure product **144** (yield: 321 mg, 1.06 mmol, 80%) as a yellow oil.

TLC (ethyl acetate/cyclohexane = 1:5): $R_f = 0.27$; $^1\text{H NMR}$ (CDCl_3): $\delta = 7.59$ (d, $J = 15.6$ Hz, 1H), 7.38 (t, $J = 8$ Hz, 1H), 7.29 (q, $J = 7.8$ Hz, 1H), 6.76 (m, 2H), 6.35 (m, 1H), 5.86 (m, 1H), 3.38 (q, $J = 6.08$ Hz, 1H), 3.32 (s, 3H), 2.31 (m, 2H), 1.68 (s, 6H), 1.14 (d, $J = 6.04$ Hz, 3H); $^{13}\text{C NMR}$ (CDCl_3): $\delta = 160.5, 157.0, 142.1, 135.1, 133.5, 133.2, 133.1, 128.4, 120.7, 115.8, 110.7, 105.2, 76.6, 56.2, 39.5, 25.7, 19.1$; LC-MS (ESI): $t_R = 11.20$ min, calcd. for $\text{C}_{18}\text{H}_{23}\text{O}_4$ $[\text{M}+\text{H}]^+$, 303.36, found 303.07.

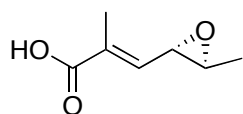
5-Methoxy-(2E)-hexenal (146)



Chemical Formula: $\text{C}_7\text{H}_{12}\text{O}_2$
Molecular Weight: 128.1690

To a solution of 5-(*S*)-Hydroxy-(2*E*)-hexenal (500 mg, 4.38 mmol) and 2,6-di-*tert*-butyl-4-methylpyridine (1.17 g, 5.7 mmol) in anhydrous DCM (8 ml) was added methyltriflate (645 μl , 5.7 mmol) at room temperature. After stirring overnight, the mixture was directly transferred on top of a silica gel column. Flash chromatography (ethyl acetate/petroleum ether = 1:2) furnished the desired **146** (yield: 504 mg, 3.9 mmol, 90%). However it could not be obtained as pure compound due to its low boiling point, and used directly for the next reaction.

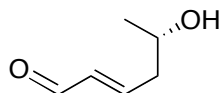
B.p. = 33 °C at 16 mbar; TLC (ethyl acetate/petroleum ether = 1:2): $R_f = 0.52$.

1-Methyl-(3S,4S)-oxirane-(2E)-hexenoic acid (147)Chemical Formula: C₇H₁₀O₃

Molecular Weight: 142.1525

To a 0.1 M NaOH solution (THF/H₂O 1:1, 10 ml) was added **150** (167 mg, 1.1 mmol) at room temperature for 3 h. The unreacted **150** was removed by addition of ethyl acetate and extraction. The aqueous phase was acidified by addition of 0.1 M HCl to pH 4 and the product was extracted by addition of ethyl acetate. The organic layer was dried over Na₂SO₄ and concentrated *in vacuo*. The residue was purified by silica gel chromatography (methanol/DCM = 1:9) to obtain pure product **147** (yield: 79 mg, 0.56 mmol, 51%) as a viscous colourless oil.

TLC (methanol/DCM = 1:9): R_f = 0.48; ¹H NMR (MeOD): δ = 6.65 (d, J = 9.1 Hz, 1H), 4.20 (dd, J = 6.24, 9.16 Hz, 1H), 3.68 (m, 1H), 1.88 (s, 3H), 1.12 (d, J = 6.44 Hz, 3H); ¹³C NMR (MeOD): δ = 171.3, 141.7, 131.4, 73.6, 71.5, 18.7, 13.4.

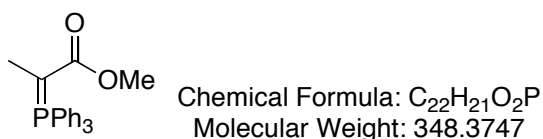
5-Hydroxy-(2E)-hexenal (148)Chemical Formula: C₆H₁₀O₂

Molecular Weight: 114.1424

A mixture of THF/acetaldehyde (4:1, 500 ml) and *L*-proline (1.2 g) was stirred for 14 h at 4 °C. The crude reaction mixture was filtered through silica gel and concentrated *in vacuo*. The residue was then purified by silica gel chromatography (ethyl acetate/petroleum ether = 1:1) to obtain pure product **148** (yield: 3.2 g, 28 mmol) as a colourless oil.

TLC (ethyl acetate/petroleum ether = 1:1): $R_f = 0.38$. The obtained NMR spectrums were in accordance to the literature values.

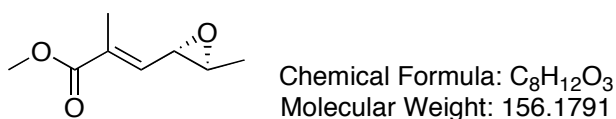
Methyl (triphenylphosphoranylidene)propionate (**149**)



A mixture of triphenylphosphine (2.11 g, 8.06 mmol) and methyl bromopropionate (1.00 ml, 8.96 mmol) in H_2O (10 mL) was stirred for 24 h at 70 °C. The mixture was cooled to room temperature, then a solution of NaOH (716 mg, 17.9 mmol) in H_2O (21 mL) was added. The resulting mixture was stirred for 5 min at room temperature, then DCM was added to redissolve the suspension formed overnight. The organic layer was separated, and the aqueous layer was extracted with DCM. The combined organic layers were dried over Na_2SO_4 and concentrated. The residue was triturated with hexane and filtered. The resulting residue was dried *in vacuo* to give **149** (2.66 g, 7.64 mmol, 95%) as a light yellow solid.

The obtained NMR spectrums were in accordance to the literature values.

1-methyl-(3S,4S)-oxirane-(2E)-hexenoic methyl ester (**150**)



To a solution of **129** (317 mg, 3.6 mmol) and **149** (1.5 g, 4.33 mmol) in DCM (90 ml) was added MnO_2 (2.5 g, 29 mmol) at room temperature. The resulting mixture

was stirred for 2 days and then filtrated through Celite and washed with DCM. The filtrate was concentrated. The residue was purified by silica gel chromatography (ethyl acetate/petroleum ether = 1:5) to obtain pure product **150** (yield: 546 mg, 3.5 mmol, 97%) as a colourless oil.

TLC (ethyl acetate/petroleum ether = 1:5): $R_f = 0.47$; $^1\text{H NMR}$ (CDCl_3): $\delta = 6.28$ (d, $J = 8.8$ Hz, 1H), 3.74 (s, 3H), 3.31 (dd, $J = 1.76, 8.8$ Hz, 1H), 3.03 (m, 1H), 2.00 (s, 3H), 1.39 (d, $J = 5.28$ Hz, 3H); $^{13}\text{C NMR}$ (CDCl_3): $\delta = 167.8, 138.4, 132.4, 56.4, 55.6, 52.2, 17.7, 13.1$; LC-MS (ESI): $t_R = 7.99$ min, calcd. for $\text{C}_8\text{H}_{12}\text{O}_3$ $[\text{M}]^+$, 156.18, found 156.93.

6.2 Biological part

6.2.1 Material of bioassays

Chemicals, antibiotics and inhibitors

All chemicals and antibiotics were purchased from Sigma, Roth, Merck and Duchefa. DCG-04 was provided by Dr. H. Overkleeft (Leiden University, Netherlands) and Dr. M. Bogyo (Stanford Medical School, US). “Click Chemistry” trifunctional tag was provided by Dr. B. F. Cravatt (Skaggs Institute for Chemical Biology, US). Inhibitors for competition assays were purchased from Sigma, Bachem, Fluka, Calbiochem and from labs of Dr. Dr. H. Overkleeft, Dr. M. Bogyo, Dr. B. F. Cravatt and Dr. S. Q. Yao (National University of Singapore).

Enzymes

Restriction enzymes were purchased from Fermentas and New England Biolabs. Taq polymerase for standard PCR was purchased from Promega and BioBudget. High-fidelity polymerase was purchased from Roche. Ligase was purchased from Promega and Fermentas. Wheat germ lipase Type I and *Pseudomonas fluorescence* lipase were purchased from Sigma-aldrich.

Vectors

Plasmid pFK26 was supplied by Dr. F. Kaschani at MPIZ (Colonge, Germany).

Kits and primers

Oligonucleotide primers were purchased from Invitrogen. Kits for isolating DNA were purchased from Qiagen (Hilden, Germany). Kits for isolating Plasmid were purchased from Peqlab (Erlangen, Germany) or Macherey-Nagel (Duren, Germany).

Pathogens

Pseudomonas syringae pv. *tomato* strain DC3000 (*Pst*) was obtained from Dr. Silke Robatzek and Dr. Jane Parker at the MPIZ (Cologne, Germany). *Pseudomonas syringae* pv. *tomato* DC3000 carrying *AvrPphB* was obtained from lab of Dr. R. A. van der Hoorn at MPIZ (Cologne, Germany).

Bacterial strains

Cloning was applied with *Escherichia coli* strain DH10B. Agrobacterium-infiltration and plant transformation was utilized with *Agrobacterium tumefaciens* strain GV3101.

Plant material

The *Arabidopsis thaliana* was carried out using ecotype Columbia (Col-0). γ -VPE overexpressor and knockout lines described in (Rojo *et al.*, 2003 and Rojo *et al.*, 2004) and quadrupole mutant of VPE described in (Gruis *et al.*, 2004) were used for labelling assays. *Arabidopsis thaliana* ecotype Landsberg cell suspension

culture was obtained from Sainsbury lab (John Innes centre, Norwich, UK). *N. benthamiana* (310A) were grown at the MPIZ (Cologne, Germany).

6.2.2 Methods of bioassays

Plant growth conditions

Arabidopsis used for labelling assays were grown in two different ways: long day (16:8 day/night regime) or short day condition (12:12 day/night regime). Four to five weeks old plants were applied for labelling assays. *N. benthamiana* were grown in a climate chamber at a 14 h light regime at 18°C (night) and 22°C (day). Four to six weeks old plants were applied for agroinfiltration.

All the plants were planted by labmates from plant chemetics lab and workers in MPIZ.

Preparation of leaf extracts

The leaves of *Arabidopsis* were usually 1 cm long and 0.5 cm broad. The well growing leaves, without speckles on the surface, were used for assays.

a. Extraction with 10% triton TX-100

Leaves in an Eppendorf tube (1.5 ml) with water (100 µl) were grinded with a blue stick (sigma C7353). After that, water (890 µl) was added with 10% triton TX-100 (10 µl). The mixture was centrifugated at 13.2 krpm for 1 min after a short time of vibration by a vortex. Clear green aqueous extract (900 µl) was removed out and used for labelling assays.

b. Extraction without 10% triton TX-100

The process is same as above without using triton.

General principle of ABPP**Basic Material:**

TBS: 150 mM NaCl and 50 mM pH 7.5 Tris-HCl Buffer

TBST: TBS + 0.1% Tween-20

TBS + 3% BSA: 50 ml TBS + 1.5 g BSA

Loading buffer (2 x SDS buffer): 5 ml total, 0.5 ml 1M pH 6.8 Tris, 3.5 ml 50% glycerol, 1.6 ml 10% SDS, 0.4 ml BME + BFB

Master solution: per each assay

25 μ l buffer (final 50 mM);

0.5 μ l 1M CaCl₂ (end 1 mM); if needed

0.5 mg *L*-Cys (final 1 mg/ml) or DTT (final 1 mM); if needed

375 μ l H₂O

Labelling

To one Eppendorf tube (1.5 ml) with master solution (400 μ l) was added fresh leaf extract (100 μ l) and ABP (n μ M). The mixture was rotated with a rotary shaker for n h and the labelling reaction was quenched with a -20°C acetone (1 ml). The proteins were centrifugated at 13 krpm for 2 min and supernatant was removed. 70 % acetone (500 μ l) of -20°C was added again. The proteins in acetone solution

were vortexed and centrifugated (1 min, 13 krpm) and the supernatant was removed. The protein pellets were dried at room temperature, resuspended in 50 µl SDS loading buffer, heated at 90°C for 10 min and stored at -20°C.

Western blot for detecting biotinylated proteins:

Samples were loaded onto polyacrylamide gel (12%) and proteins were transferred to a PVDF membrane (Millipore IVPH00010). The membrane was incubated with ultrasensitive streptavidin-HRP (Sigma S2438, dilution of 1:3000) and signals were detected using enhanced chemiluminescence (Pierce FEMTO ECL 34095)/(Pico ECL 34080) (Thermo Fisher Scientific, Bonn, Germany) on X-ray films (Kodak, Germany).

Fluorescence scan for detecting fluorescent probe labelled proteins

Samples were loaded onto polyacrylamide gel (12%). After a short time of washing the gel, proteins on the gel were directly scanned with a Typhoon 8600 variable mode imager.

No probe control

The protocol is same as above, only using same volume of DMSO instead of ABP solution.

Cloning for AvrPphB

AvrPphB was amplified from *Pseudomonas syringae* pv. *tomato* DC3000 carrying *AvrPphB* with the primers we designed. Cloning vector pFK26 was digested with restriction enzymes XhoI and PstI. PCR products were digested and ligated into pFK26 and the plasmids were transformed into *E. coli*. Successful clones, validated by nucleotide sequencing, were digested using the same restriction enzymes for shuttling into the PTP05 binary vector. Inserts in the generated pZM05 plasmid were confirmed by PCR using vector specific primers.

Agrobacterium infiltration of AvrPphB construct

pZM05 plasmids were transformed into *Agrobacterium tumefaciens* strain GV3101. Spin down the bacteria from the medium, add 0.5 to 1 ml MES Buffer (10 mM MgCl₂, 10 mM 2-(*N*-Morpholino)ethanesulfonic acid) and measure OD. Tune the concentration and dilute OD into 2, add acetosyringon (final 10 mM). Let the mixture stay in darkness for 2 hours and in the meanwhile put water on the plant soil under light to activate the plant. When plant is treated water under light, the stomas back side of leaves will be open. The bacteria were infiltrated into the leaves of *N. benthamiana* with a syringe.

6.2.3 Procedures of ABPP

All the procedures follow the general principle of ABPP, and the detail information of each assay is described according to the figure of the assay.

Figure 3

Wheat germ lipase Type I (WGL) and *Pseudomonas fluorescence* lipase (PL) were prepared as 0.5 mg/ml aqueous solutions, which were incubated with probes **23** and **24** at a final concentration (4 μ M) for 3 h at room temperature. Preparation of protein samples was followed the general method and the proteins on the protein blots were detected by streptavidin-HRP.

Figure 4

Arabidopsis leaf extracts (LE) and cell cultures (CC) were prepared as 5 mg/ml (LE) and 50 mg/ml (CC) pH 6 NaOAc buffer solutions respectively, which were incubated with probes **23** and **24** at a final concentration (4 μ M) for 3 h at room temperature. Preparation of protein samples was followed the general method and the proteins on the protein blots were detected by streptavidin-HRP.

Figure 5a/b

Arabidopsis leaf extracts were prepared as 5 mg/ml pH 6 NaOAc buffer solutions, which were incubated with probe **23** at a final concentration (4 μ M) and co-incubated with inhibitors at a final concentration (40 μ M) for competition labelling, for 3 h at room temperature. Preparation of protein samples was followed the

general method and the proteins on the protein blots were detected by streptavidin-HRP.

Figure 6

a

Arabidopsis leaf extracts were prepared as 5 mg/ml pH (4 or 6) NaOAc and (8 or 10) Tris buffer solutions respectively, which were incubated with probe **23** at a final concentration (4 μ M) for 3 h at room temperature.

b

Arabidopsis leaf extracts were prepared as 5 mg/ml pH 6 NaOAc buffer solutions with additional cofactors (Cysteine or CaCl₂) respectively, which were incubated with probe **23** at a final concentration (4 μ M) for 3 h at room temperature.

c

Arabidopsis leaf extracts were prepared as 5 mg/ml pH 6 NaOAc buffer solutions, which were incubated with probe **23** at a final concentration (4 μ M) for (1/4, 1/2, 1, 2, 3, 4, 5, 10) h at room temperature.

d

Arabidopsis leaf extract was separated as membrane proteome (M) and soluble proteome (S) with an ultra-centrifugation. Another equal leaf extract as total proteome (T) with M and S were prepared as 20 mg/ml (T) pH 6, NaOAc buffer solutions respectively, which were incubated with probe **23** at a final concentration (4 μ M) for 3 h at room temperature.

All preparation of protein samples was followed the general method and the proteins on the protein blots were detected by streptavidin-HRP.

Figure 9a/b

Arabidopsis leaf extracts were prepared as 5 mg/ml (a) and 0.15 mg/ml (b) pH 8 Tris buffer solutions with DTT (1 mM), which were incubated with probes **38 - 45** at a final concentration (2 μ M) for 3 h at room temperature. Preparation of protein samples was followed the general method and the proteins on the protein blots were detected by streptavidin-HRP.

Figure 10

Arabidopsis leaf extracts were prepared as 50 μ g/ml pH 7.4 TBS and pH 7.5 caspase buffer solutions with DTT (1 mM) respectively, which were incubated with probes **41** and DCG-04 at a final concentration (2 μ M) and co-incubated with inhibitors at a final concentration (60 μ M) for competition labelling, for 3 h at room temperature. Preparation of protein samples was followed the general method and the proteins on the protein blots were detected by streptavidin-HRP.

Figure 12**a**

Arabidopsis leaf extracts were prepared as 50 μ g/ml pH (4 – 6.5) NaOAC and pH (7 – 10) Tris buffer solutions with DTT (1 mM) respectively, which were incubated with probes **41** at a final concentration (2 μ M) for 3 h at room temperature.

b

Arabidopsis leaf extracts were prepared as 50 μ g/ml pH 8 Tris buffer solutions with or without DTT (1 mM) respectively, which were incubated with probes **41** at a final concentration (2 μ M) for 3 h at room temperature.

c

Arabidopsis leaf extracts were prepared as 50 µg/ml pH 8 Tris buffer solutions with DTT (1 mM), which were incubated with probes **41** at a final concentration (2 µM) for (1/30, 1/4, 1/2, 1, 2, 3, 4, 5, 6, 7, 8) h at room temperature.

d

Arabidopsis leaf extracts were prepared as (12.5, 25, 50, 100, 200) µg/ml pH 8 Tris buffer solutions with DTT (1 mM) respectively, which were incubated with probes **41** at a final concentration (2 µM) for 3 h at room temperature.

Preparation of protein samples was followed the general method and the proteins on the protein blots were detected by streptavidin-HRP.

Figure 13a/b

Arabidopsis leaf extracts were prepared as 50 µg/ml pH 8 Tris buffer solutions with DTT (1 mM), which were incubated with probes **41** and DCG-04 at a final concentration (2 µM) and co-incubated with inhibitors at a final concentration (30 µM) for competition labelling, for 3 h at room temperature. Preparation of protein samples was followed the general method and the proteins on the protein blots were detected by streptavidin-HRP.

Figure 18

Arabidopsis leaf extracts were prepared as 50 µg/ml pH 8 Tris buffer solutions, which were incubated with probes **69 - 80** at a final concentration (5 µM) for 2 h at room temperature. The next step was followed “click chemistry” protocol (Kaschani *et al.*, 2009). Preparation of protein samples was followed the general method and the proteins on the protein blots were detected by streptavidin-HRP.

Figure 19

Arabidopsis leaf extracts were prepared as 200 µg/ml pH 6 NaOAc and pH 8 Tris buffer solutions, which were incubated with probes **82** - **86** and biotinylated AEBSF at a final concentration (20 µM) for 2 h at room temperature. Preparation of protein samples was followed the general method and the proteins on the protein blots were detected by streptavidin-HRP.

Figure 20

Arabidopsis leaf extracts were prepared as 30 µg/ml pH 8 Tris buffer solutions, which were incubated with a mixed probe of **84** and **87** at a final concentration (2 µM) for 2 h at room temperature. Preparation of protein samples was followed the general method and the proteins labelled by fluorescent probe **87** on protein gel were detected by fluorescence photometer and proteins labelled by affinity probe **84** on the protein blots from the same protein gel were detected by streptavidin-HRP.

Figure 21**a**

Arabidopsis leaf extracts were prepared as 20 µg/ml pH 8 Tris buffer solutions, which were incubated with probe **87** at a final concentration (4 nM - 4 µM) for 2 h at room temperature.

b

Arabidopsis leaf extracts were prepared as 20 µg/ml pH (4 – 6.5) NaOAc and pH (7 – 9) Tris buffer solutions, which were incubated with probe **87** at a final concentration (2 µM) for 2 h at room temperature.

Preparation of protein samples was followed the general method and the proteins on protein gel were detected by fluorescence photometer.

Figure 25

Arabidopsis leaf extracts were prepared as 20 µg/ml pH 8 Tris buffer solutions, which were pre-incubated with AEBSF at a final concentration (600 µM) for 1/2 h at room temperature and then incubated with probe **87** at a final concentration (2 µM) for 2 h at room temperature. Preparation of protein samples was followed the general method and the proteins on protein gel were detected by fluorescence photometer.

Figure 26

Arabidopsis leaf extracts and mouse liver extracts were prepared as 20 µg/ml pH 8 Tris buffer solutions, which were pre-incubated with AEBSF or Ac-AEBS at a final concentration (600 µM) for 1/2 h at room temperature and then incubated with probe **87** and BCR at a final concentration (2 µM) for 2 h at room temperature. Preparation of protein samples was followed the general method and the proteins on protein gel were detected by fluorescence photometer.

Figure 30

Non-transformed *Pseudomonas* (-), *Pseudomonas* expressing HA-tagged (pFK141) or His-tagged (pFK142) AvrPphB were grown in rich medium (NYG) containing gentamicin (10 µg/ml) and rifampicin (30 µg/ml) at 28°C overnight. The bacteria were measured OD at 600 nm and calculated as 1.25 (-), 0.97 (pFK141) and 1.03 (pFK142). The bacterial medium was centrifuged at 5000 rpm for 10 min and the liquid and pellets were separated. Secreted proteins in medium and proteins in bacterial pellet were Samples were loaded onto polyacrylamide gel (12%) and proteins were transferred to a PVDF membrane (Millipore IVPH00010). The membrane was incubated with anti-HA antibody (plant chemetics lab, dilution of 1:3000) and followed with ultrasensitive streptavidin-HRP (Sigma S2438, dilution of 1:3000) and signals were detected using enhanced chemiluminescence (Pierce femto ECL 34095)/(Pierce pico ECL 34080) (Thermo Fisher Scientific, Bonn, Germany) on X-ray films (Kodak, Germany).

Figure 31**a**

Non-transformed *Pseudomonas* (-), *Pseudomonas* expressing HA-tagged (pFK141) or His-tagged (pFK142) AvrPphB were grown in rich medium (NYG) and min-A medium containing gentamicin (10 µg/ml) and rifampicin (30 µg/ml) at 28°C overnight. The bacteria were measured OD at 600 nm and calculated as 0.78 (-), 1.02 (pFK141) and 0.86 (pFK142) in NYG medium and 0.88 (-), 1.04 (pFK141) and 1.06 (pFK142) in min-A medium. The western blot was performed as same as above.

b

Pseudomonas pellets were frozen with liquid nitrogen and crashed by metal balls with a shaker. After one minute shaking, the solid bacterial powder were prepared as pH 8 Tris buffer solutions, which were incubated with probe **101** at a final concentration (5 μ M) for 2 h at room temperature. Preparation of protein samples was followed the general method and the proteins on protein gel were detected by fluorescence photometer and the proteins were transferred to a PVDF membrane (Millipore IVPH00010) to perform the western blot as above and the protein blot was finally stained by coomassie.

Figure 32

Non-transformed *Pseudomonas* (-) and *Pseudomonas* expressing His-tagged (pFK142) AvrPphB were grown in min-A medium containing gentamicin (10 μ g/ml) and rifampicin (30 μ g/ml) at 28°C overnight. The bacteria were measured OD at 600 nm and calculated as 0.47 (-) and 0.5 (pFK142). The bacterial proteome solutions were prepared following above procedure, which were incubated with probe **101** at a final concentration (5 μ M) for 2 h at room temperature. After that, the **101** labelled AvrPphB were incubated with Nickel beads (50 μ l, Qiagen Ni-NTA Agarose) in 8 ml pH 7.4 PBS buffer for 1 h. The supernatant and beads were separated by centrifugation at 3000 rpm for 5 min, and preparation of protein samples in both components was followed the general method. The proteins on protein gel were detected by fluorescence photometer and then stained by coomassie.

Figure 33

Pseudomonas expressing HA-tagged (pFK141) was grown in min-A medium containing gentamicin (10 µg/ml) and rifampicin (30 µg/ml) at 28°C overnight. The bacteria were measured OD at 600 nm and calculated as 0.5 (pFK141). The bacterial proteome solutions were prepared following above procedure but with variant pH (6 – 9), which were incubated with probe **101** at a final concentration (5 µM) for 2 h at room temperature. Preparation of protein samples was followed the general method. The proteins on protein gel were detected by fluorescence photometer.

Figure 34

Pseudomonas expressing HA-tagged (pFK141) was grown in min-A medium containing gentamicin (10 µg/ml) and rifampicin (30 µg/ml) at 28°C overnight. The bacteria were measured OD at 600 nm and calculated as 0.47 (pFK141). The bacterial proteome solutions were prepared following above procedure at pH 7, which were pre-incubated with inhibitors (**92, 94, 96, 98, 100**) at a final concentration (100 µM) for 1/2 h at room temperature and then incubated with probe **101** at a final concentration (5 µM) for 1.5 h at room temperature. Preparation of protein samples was followed the general method. The proteins on protein gel were detected by fluorescence photometer.

Figure 36**a**

Arabidopsis leaf extracts of different lines: *Columbia* wildtype (Col), γ -VPE over expressed ($O\gamma$), γ -VPE knockout ($K\gamma$), α , β , γ , δ -VPEs quadruple knockout (K) were prepared as 100 $\mu\text{g/ml}$ pH 5.5 NaOAc buffer solutions, which were incubated with probe **107** at a final concentration (5 μM) for 2 h at room temperature. Preparation of protein samples was followed the general method and the proteins on protein gel were detected by fluorescence photometer and stained by coomassie.

b

Arabidopsis leaf extracts of γ -VPE over expressed ($O\gamma$), were prepared as 100 $\mu\text{g/ml}$ pH 5.5 NaOAc buffer solutions, which were pre-incubated with **106** at a final concentration (100 μM) for 1/2 h at room temperature and then incubated with probe **107** at a final concentration (5 μM) for 2 h at room temperature. Preparation of protein samples was followed the general method and the proteins on protein gel were detected by fluorescence photometer.

7 References

1. Adam G.C.; Cravatt B.F.; Sorensen E.J. *J. Chem. Biol.* **2001**, *8*, 81-95.
2. Adam G.C.; Sorensen E.J.; Cravatt B.F. *Nat. Biotechnol.* **2002**, *20*, 805-809.
3. Adam G.C.; Burbaum J.; Kozarich J.W.; Patricelli M.P.; Cravatt B.F. *J. Am. Chem. Soc.* **2004**, *126*, 1363-1368.
4. Adams M. D.; Kelley J. M.; Gocayne J. D. *et al.*, *Science* **1991**, *252*, 1651-1656.
5. Adamczyk M.; Gebler J. C.; Wu J. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 999-1001.
6. Amara U.; Rittirsch D.; Flierl M.; Bruckner U.; Klos A.; Gebhard F.; Lambris J. D.; Huber-Lang M. *Adv. Exp. Med. Biol.* **2008**, *632*, 71-79.
7. Amoresano A.; Marino G.; Cirulli C.; Quemeneur E. *Eur. J. Mass Spectrom.* **2004**, *10*, 401-412.
8. Anderson L.; Seilhamer J. *Electrophoresis* **1997**, *18*, 533-537.
9. Anderson N. L.; Anderson N. G. *Electrophoresis* **1998**, *19*, 1853-1861.
10. Anderson N. L.; Anderson N. G.; Haines L. R.; Hardie D. B.; Olafson R. W.; Pearson T. W. *J. Proteome Res.* **2004**, *3*, 235-244.
11. Atsushi A. *Organic Syntheses* **2004**, *Coll. 10*, 273; **2002**, *79*, 103.
12. Atsushi A.; Liu J.F.; Masamune S. *J. Am. Chem. Soc.* **1997**, *119*, 2586-2587.
13. Barglow K.T.; Cravatt B.F. *Chem. Biol.* **2004**, *11*, 1523.
14. Barrett A. J.; Rawlings N. D. *Arch. Biochem. Biophys.* **1995**, *318*, 247-250.
15. Barrett A. J.; Rawlings N. D. *Biol. Chem.* **2001**, *382*, 727-733.
16. Berndt P.; Hobohm U.; Langen H. *Electrophoresis* **1999**, *20*, 3521-3526.
17. Beroza P.; Damodaran K.; Lum R. T. *Curr. Top. Med. Chem.* **2005**, *5*, 371-381.
18. Beard R. L, Heumann K. F. *Science* **1952**, *116*, 553-554.
19. Beers E. P.; Jones A. M.; Dickerman A. W. *Phytochemistry* **2004**, *65*, 43-58.
20. Blackwell H. E.; Zhao Y. D. *Plant Physiol.* **2003**, *133*, 448-455.

-
21. Blum G.; Mullins S. R.; Keren K.; Fonovic M.; Jedeszko C.; Rice M. J.; Sloane B. F.; Bogoy M. *Nat. Chem. Biol.* **2005**, *1*, 203-209.
 22. Bochtler M.; Ditzel L.; Groll M.; Hartmann C.; Huber R. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 295-317.
 23. Bogoy M.; McMaster J. S.; Gaczynska M.; Tortorella D.; Goldberg A. L.; Ploegh H. *Proc. Natl. Acad. Sci. USA.* **1997**, *94*, 6629-6634.
 24. Bogoy M.; Verhelst S.; Bellingard-Dubouchaud V.; Toba S.; Greenbaum D. *Chem. Biol.* **2000**, *7*, 27-38.
 25. Bogoy M.; Wang E. W. *Curr Top Microbiol. Immunol.* **2002**, *268*, 185-208.
 26. Boldogkoei Z.; Murvai J. *Virus Genes.* **1994**, *9*, 47-51.
 27. Borensztajn K.; Peppelenbosch M. P.; Spek C. A. *Trends Mol. Med.* **2008**, *14*, 429-440.
 28. Borodovsky A.; Ovaas H.; Kolli N.; Gan-Erdene T.; Wilkinson K. D.; Ploegh H. L.; Kessler B. M. *Chem. Biol.* **2002**, *9*, 1149-1159.
 29. Bottari P.; Aebbersold R.; Turecek F.; Gelb M. H. *Bioconj. Chem.* **2004**, *15*, 380-388.
 30. Bouma B. N.; Mosnier L. O. *Ann Med.* **2006**, *38*, 378-388.
 31. Brik A.; Wong C. H. *Org. Biomol. Chem.* **2003**, *1*, 5-14.
 32. Bromme D.; Schierhorn A.; Kirschke H.; Wiederander B.; Barth A.; Fittkau S.; Demuth H. U. *Biochem. J.* **1989**, *263*, 861.
 33. Brown H. C.; Dhar R.K.; Ganesan K.; Singaram B. *J. Org. Chem.* **1992**, *57*, 499.
 34. Brownlee G. G.; Sanger F.; Barrell B. G. *Nature* **1967**, *215*, 735-736.
 35. Burdine L.; Kodadek T. *Chem. Biol.* **2004**, *11*, 593-597.
 36. Burley S. K. *et al.*, *Nat. Genet.* **1999**, *23*, 151-157.
 37. Caron M.; Imam-Sghiouar N.; Poirier F.; Le Caër J. P.; Labas V.; Joubert-Caron R. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2002**, *771*, 197-209.

-
- 38.Chan E. W.; Chattopadhyaya S.; Panicker R. C.; Huang X.; Yao S. Q. *J. Am. Chem. Soc.* **2004**, *126*, 14435-14446.
- 39.Chi A.; Huttenhower C.; Geer L. Y.; Coon J. J.; Syka J. E. P.; Bai D. L.; Shabanowitz J.; *et al.* *Proc Natl Acad Sci U S A* **2007**, *104*, 2193-2198.
- 40.Claverie, J. M. *Hum. Mol. Genet.* **1997**, *6*, 1735–1744.
- 41.Clemens S. *J. Plant Physiol.* **2006**, *163*, 319-332.
- 42.Colbère-Garapin F.; Blondel B.; Saulnier A.; Pelletier I.; Labadie K. *Microbes Infect.* **2005**, *7*, 767-775.
- 43.Córdova A.; Notz W.; Barbas C. F. 3rd *J. Org. Chem.* **2002**, *67*, 301-303.
- 44.Corre C.; Lowden P. A. *Chem. Commun. (Camb).* **2004**, *21*, 990-991.
- 45.Dale H. *Science* **1934**, *80*, 343-349.
- 46.Damaschun G.; Fichtner P.; Pürschel H. V.; Reich J. G. *Acta Biol Med Ger.* **1968**, *21*, 309-316.
- 47.Dangl J.L.; Jones J.D. *Nature* **2001**, *411*, 826-833.
- 48.Davis F.A.; Reddy V.G.; Liu H. *J. Am. Chem. Soc.* **1995**, *117*, 3651-3652.
- 49.Davis F.A.; Liu H.; Zhou P.; Fang T.; Reddy G. V.; Zhang Y. *J. Org. Chem.* **1999**, *64*, 7559-7567.
- 50.Dawson J. E.; Seckute J.; De S.; Schueler S. A.; Oswald A. B.; Nicholson L. K. *Proc. Natl. Acad. Sci USA.* **2009**, *106*, 8543-8548.
- 51.Dick L. R.; Cruikshank A. A.; Destree A. T.; Grenier L.; McCormack T. A.; Melandri F. D.; Nunes S. L.; *et al.* *J. Biol. Chem.* **1997**, *272*, 182-188.
- 52.Dixon D. P.; McEwen A. G.; Laphorn A. J.; Edwards R. *J. Biol. Chem.* **2003**, *278*, 23930-23935.
- 53.Dunham, I. *et al.*, *Nature* **1999**, *402*, 489–495.
- 54.Ehrlich P. *Folia Serologica VII*, **1911**, 697-714.
- 55.Eisenreich W.; Kupfer E.; Weber W.; Bacher A. *J. Biol. Chem.* **1997**, *272*, 867-874
-

-
56. Ellrichmann M; Ritter P.R.; Otte J.; Schrader H.; Banasch M; Brunke G.; Herzig K.; *et al. Reg. Pept.* **2007**, *139*, 136-140.
57. Evans D.A.; Nelson J.V.; Vogel E.; Tabler T.R. *J. Am. Chem. Soc.* **1981**, *103*, 3099.
58. Evans D.A.; Ennis M.D.; Mathre D.J. *J. Am. Chem. Soc.* **1982**, *104*, 1737-1739.
59. Evans M.J.; Saghatelian A.; Sorensen E.J.; Cravatt B.F. *Nat. Biotechnol.* **2005**, *23*, 1303.
60. Evans M.J.; Cravatt B.F. *Chem. Rev.* **2006**, *106*, 3279-3301.
61. Fenn J. B.; Mann M.; Meng C. K.; Wong S. F.; Whitehouse C. M. *Science* **1989**, *246*, 64-71.
62. Fernández-Suárez M.; Ting A. Y. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 929-943.
63. Fields S.; Song O. *Nature* **1989**, *340*, 245-246.
64. Fleischmann R.; Adams M.; White O.; Clayton R.; Kirkness E.; Kerlavage A.; Bult C.; *et al. Science* **1995**, *269*, 496-512.
65. Flor H. H. *Annu. Rev. Phytopathol.* **1971**, *9*, 275-296.
66. Foettinger A.; Leitner A.; Lindner W. *J. Chromatogr. A* **2005**, *1079*, 187-196
67. Fonović M.; Bogoy M. *Expert Rev Proteomics.* **2008**, *5*, 721-730.
68. Franchi L.; Eigenbrod T.; Muñoz-Planillo R.; Nuñez G. *Nat. Immunol.* **2009**, *10*, 241-247.
69. Ge X.; Dietrich C.; Matsuno M.; Li G.; Berg G. *et al. EMBO Rep.* **2005**, *6*, 282-288.
70. Gilroy E.; Hein I.; van der Hoorn R. A. L.; Boevink P.; Venter E.; *et al. Plant J.* **2007**, *53*, 1-13.
71. Goese M.; Eisenreich W.; Kupfer E.; Weber W.; Bacher A. *J. Biol. Chem.* **2000**, *275*, 21192-21196.
72. Goldberg D. M. *Clin. Chim. Acta.* **2000**, *291*, 201-221.
-

-
73. Goshe M. B.; Veenstra T. D.; Panisko E. A.; Conrads T. P. *et al.*, *Anal. Chem.* **2002**, *74*, 607–616.
74. Greenbaum D.; Medzihradzky K. F.; Burlingame A.; Bogyo M. *Chem. Biol.* **2000**, *7*, 569-581.
75. Greenspan P.; Mao F. W.; Ryu B. H.; Gutman R. L. *J. Chromatogr A.* **1995**, *698*, 333-339.
76. Greenbaum D. C.; Arnold W. D.; Lu F.; Hayrapetian L.; Baruch A.; Krumrine J.; Toba S.; Chehade K.; *et al.* *Chem Biol.* **2002**, *9*, 1085-1094.
77. Gruis D.; Schulze J.; Jung R. *Plant Cell.* **2004**, *16*, 270-290.
78. Gunther C. S.; Gaasterland T. *Genome Inform.* **2001**, *12*, 34-43.
79. Gupta V.; Ogawa A. K.; Du X.; Houk K. N.; Armstrong R. W. *J. Med. Chem.* **1997**, *40*, 3199-3206.
80. Gygi S.; Rochon Y.; Franza B. R.; Aebersold, R. *Mol. Cell. Biol.* **1999**, *19*, 1720–1730.
81. Gygi S. P.; Rist B.; Aebersold R. *Curr. Opin. Biotechnol.* **2000**, *11*, 396-401.
82. Gygi, S. P.; Rist B.; Gerber S. A.; Turecek, F. *et al.*, *Nat. Biotechnol.* **1999**, *17*, 994–999.
83. Hanessian S.; Tehim A.; Chen P. *J. Org. Chem.* **1993**, *58*, 7768-7781.
84. Hara-Nishimura I.; Inoue K.; Nishimura M. *FEBS Lett.* **1991**, *294*, 89-93.
85. Hatsugai N.; Kuroyanagi M.; Yamada K.; Meshi T.; Tsuda S.; Kondo M.; Nishimura M.; Hara-Nishimura I. *Science* **2004**, *305*, 855-858.
86. Hauptman J.B.; Jeunet F.S.; Hartmann D. *Am. J. Clin. Nutr.* **1992**, *55*, 309S-313S.
87. Han X.; Aslanian A.; Yates J. R. 3rd *Curr. Opin. Chem. Biol.* **2008**, *12*, 483-490.
88. Han X.; Jin M.; Breuker K.; McLafferty F. W. *Science* **2006**, *314*, 109-112.
89. Hanada K.; Tamai M.; Yamagishi M.; Ohmura S.; Sawada J. Tanaka I. *Agric. Biol. Chem.* **1978**, *42*, 523.
-

-
90. Hansen K. C.; Schmitt-Ulms G.; Chalkley R. J.; Hirsch J.; Baldwin M. A.; Burlingame A. L. *Mol. Cell. Proteomics* **2003**, *2*, 299–314.
91. Hatsugai N.; Kuroyanagi M.; Nishimura M.; Hara-Nishimura I. *Apoptosis* **2006**, *11*, 905–911.
92. Hayashi K.; Nozaki H. *J. Antibiot.* **1999**, *52*, 917-920.
93. Hayashi K.; Inoguchi M.; Kondo H.; Nozaki H. *Phytochem.* **2000**, *55*, 1-9.
94. Hekmat O.; Kim Y. W.; Williams S. J.; He S.; Withers S. G. *J. Biol. Chem.* **2005**, *280*, 35126-35135.
95. Henderson L. J. *Proc. Natl. Acad. Sci. USA.* **1916**, *2*, 654-658.
96. Hess D. T.; Matsumoto A.; Kim S. O.; Marshall H. E.; Stamler J. S. *Nat. Rev. Mol. Cell. Biol.* **2005**, *6*, 150–166.
97. Hieter P.; Boguski M. *Science* **1997**, *278*, 601-602.
98. Hochuli E.; Kupfer E.; Maurer R.; Meister W.; Mercadal Y.; Schmidt K. *J. Antibiot. (Tokyo)* **1987**, *40*, 1086-1091.
99. Holz R. C.; Bzymek K. P.; Swierczek S. I. *Curr. Opin. Chem. Biol.* **2003**, *7*, 197-206.
100. Hooper N. M.; Karran E. H.; Turner A. J. *Biochem. J.* **1997**, *321*, 265-279.
101. Hsu J. L.; Huang S. Y.; Chow N. H.; Chen S. H. *Anal. Chem.* **2003**, *75*, 6843–6852.
102. Hu Q.; Noll R. J.; Li H.; Makarov A.; Hardman M.; Cooks R. G. *J. Mass. Spectrom.* **2005**, *40*, 430-443.
103. International Human Genome Sequencing Consortium *Nature* **2001**, *409*, 860–921.
104. James P. *Q Rev Biophys.* **1997**, *30*, 279-331.
105. Jaffrey S. R.; Erdjument-Bromage H.; Ferris C. D.; Tempst P.; Snyder S. H. *Nat. Cell. Biol.* **2001**, *3*, 193–197.
106. Jeffery D. A.; Bogyo M. *Curr Opin Biotechnol.* **2003**, *14*, 87-95.
-

-
107. Jessani N.; Liu Y.; Humphrey M.; Cravatt B. F. *Proc Natl Acad Sci U S A*. **2002**, *99*, 10335-10340.
 108. Jessani N.; Humphrey M.; McDonald W. H.; Niessen S.; Masuda K.; Gangadharan B.; Yates J. R. 3rd; *et al.* *Proc. Natl. Acad. Sci. USA*. **2004**, *101*, 13756-13761.
 109. Jessani N.; Niessen S.; Wei B. Q.; Nicolau M.; Humphrey M.; Ji Y.; Han W.; *et al.* *Nat. Methods*. **2005**, *2*, 691-697.
 110. Johnson L. R. *Annu. Rev. Physiol.* **1985**, *47*, 199-215.
 111. Jones J. D. G.; Dangl J. L. *Nature* **2006**, *444*, 323-329.
 112. Juge N. *Trends Plant Sci.* **2006**, *11*, 359-367.
 113. Kaji H.; Saito H.; Yamauchi Y.; Shinkawa T. *et al.*, *Nat. Biotechnol.* **2003**, *21*, 667-672.
 114. Kaschani F.; Verhelst S. H. L.; Van Swieten P. F.; Verdoes M.; Wong C. S.; Wang Z.M.; Kaiser M.; *et al.* *Plant J.* **2009**, *57*, 373-385.
 115. Karas M.; Hillenkamp F. *Anal. Chem.* **1988**, *60*, 2299-2301.
 116. Kato D.; Boatright K. M.; Berger A. B.; Nazif T.; Blum G.; Ryan C.; Chehade K. A. H.; *et al.* *Nat. Chem. Biol.* **2005**, *1*, 33-38.
 117. Keen N.T. *Annu. Rev. Genet.* **1990**, *24*, 447-463.
 118. Kehoe J. W.; Kay B. K. *Chem. Rev.* **2005**, *105*, 4056-4072.
 119. Kendrew J. C.; Bodo G.; Dintzis H. M.; Parrish R. G.; Wyckoff H. Phillips D. C. *Nature* **1958**, *181*, 662-6666.
 120. Kidd D.; Liu Y.; Cravatt B.F. *Biochemistry.* **2001**, *40*, 4005-4015.
 121. Kinashi H.; Sakaguchi K. *Agric. Biol. Chem.* **1984**, *48*, 245-247.
 122. Kinoshita T.; Nishimura M.; Hara-Nishimura I. *Plant Cell Physiol.* **1995**, *36*, 1555-1562.
 123. Klose J. *Humangenetik* **1975**, *26*, 231-243
 124. Korkmaz B.; Moreau T.; Gauthier F. *Biochimie.* **2008**, *90*, 227-242

-
125. Kruse U.; Bantscheff M.; Drewes G.; Hopf C. *Mol. Cell Proteomics*. **2008**, *7*, 1887-901.
 126. Krupiczkoj M. A.; Scotton C. J.; Chambers R. C. *Int. J. Biochem. Cell Biol.* **2008**, *40*, 1228-1237.
 127. Kuroyanagi M.; Nishimura M.; Hara-Nishimura I. *Plant Cell Physiol.* **2002**, *43*, 143-151.
 128. Kumar S.; Zhou B.; Liang F.; Wang W. Q.; Huang Z.; Zhang Z. Y. *Proc. Natl. Acad. Sci. USA.* **2004**, *101*, 7943-7948.
 129. Kweon H. K.; Håkansson K. *Anal. Chem.* **2006**, *78*, 1743-1749.
 130. Lall M. S.; Karvellas C.; Vederas J. C. *Org. Lett.* **1999**, *1*, 803-806.
 131. Lehfeldt C.; Shirley A.M.; Meyer K.; Ruegger M.O.; Cusumano J.C.; Viitanen P. V.; Strack D.; Chapple C. *Plant Cell* **2000**, *12*, 1295-1306.
 132. Leitner A.; Lindner W. *Proteomics* **2006**, *6*, 5418-5434.
 133. Le Gall C.; Bonnelye E.; Clézardin P. *Curr Opin Support Palliat Care.* **2008**, *2*, 218-222.
 134. Leung D.; Hardouin C.; Boger D. L.; Cravatt B. F. *Nat. Biotechnol.* **2003**, *21*, 687-691.
 135. Li Y. M.; Xu M.; Lai M. T.; Huang Q. *et al. Nature.* **2000**, *405*, 689-694.
 136. Liu Y.; Patricelli M. P.; Cravatt B. F. *Proc Natl Acad Sci U S A.* **1999**, *96*, 14694-14699.
 137. Loak K.; Ni-Li D.; Manoury B.; Billson J.; Morton F.; Hewitt E.; Watts C. *Biol. Chem.* **2003**, *384*, 1239-1246.
 138. Lokey R. S. *Curr. Opin. Chem. Biol.* **2003**, *7*, 91-96.
 139. Loidl G.; Groll M.; Musiol H. J.; Ditzel L.; Huber R.; Moroder L. *Chem. Biol.* **1999a**, *6*, 197-204.
 140. Loidl G.; Groll M.; Musiol H. J.; Huber R.; Moroder L. *Proc. Natl. Acad. Sci. USA.* **1999b**, *96*, 5418-5422.
-

-
141. Lorenzen M.; Racicot V.; Strack D.; Chapple C. *Plant Physiol.* **1996**, *112*, 1625–1630.
 142. Lundblad R. L. *Chemical Reagents for Protein Modification*, 3rd Edn., CRC Press, Boca Raton **2004**.
 143. Ma H.; Horiuchi K. Y. *Drug Discov. Today* **2006**, *11*, 661-668.
 144. Malerba M.; Crosti P.; Cerana R.; Bianchetti R. *Physiol. Plant.* **2004**, *120*, 386-394.
 145. Marvin L. F.; Roberts M. A.; Fay L. B. *Clin. Chim. Acta.* **2003**, *337*, 11-21.
 146. Masamune S.; Kim B.; Petersen J.S.; Sato T.; Veenstra S.J. *J. Am. Chem. Soc.* **1985**, *107*, 4549-4551.
 147. Masamune S.; Sato T.; Kim B.; Wollmann T.A. *J. Am. Chem. Soc.* **1986**, *108*, 8279-8281.
 148. Masamune S.; Atsushi A.; Liu J.F. *J. Org. Chem.* **1996**, *61*, 2590-2591.
 149. Matsumura H.; Ito A.; Saitoh H.; Winter P.; Kahl G.; Reuter M.; Krüger D. H.; Terauchi R. *Cell Microbiol.* **2005**, *7*, 11-18.
 150. Mayer T. U. *Trends Cell Biol.* **2003**, *13*, 270-277.
 151. Mayya V.; Rezual K.; Wu L.; Fong M. B.; Han D. K. *Mol Cell Proteomics* **2006**, *5*, 1146-1157.
 152. McDonald W. H.; Yates Iii J. R. *Curr. Opin. Mol. Therap.* **2003**, *5*, 302-309.
 153. Meijers R.; Blagova E. V.; Levnikov V. M.; Rudenskaya G. N.; Chestukhina G. G.; Akimkina T. V.; Kostrov S. V.; *et al.* *Biochemistry.* **2004**, *43*, 2784-2791.
 154. Miller T.W.; Tristram E.W.; Wolf F.J. *J. Antibiot.* **1971**, *24*, 48.
 155. Min Jou W.; Haegeman G.; Ysebaert M.; Fiers W. *Nature* **1972**, *237*, 82-88.
 156. Mitsunobu, O.; Yamada, Y. *Bull. Chem. Soc. Japan* **1967**, *40*, 2380-2382.
 157. Mitchison T. J. *Chem. Biol.* **1994**, *1*, 3-6.
 158. Molinski T.F.; Ireland C.M. *J. Org. Chem.* **1988**, *53*, 2103.
-

-
159. Mousses S.; Caplen N. J.; Cornelison R.; Weaver D.; Basik M.; Hautaniemi S.; Elkahouloun A. G.; *et al. Genome Res.* **2003**, *13*, 2341-2347.
 160. Mulder G. J. *Bulletin des Sciences Physiques et Naturelles en Neerlande* **1838**.
 161. Nagaraj S. H.; Gasser R. B.; Ranganathan S. *Brief Bioinform.* **2007**, *8*, 6-21.
 162. Nakabo Y.; Pabst M.J. *J. Leukoc. Biol.* **1996**, *60*, 328-336.
 163. Nazif T.; Bogoyo M. *Proc. Natl. Acad. Sci. USA.* **2001**, *98*, 2967-2972.
 164. Neubauer G.; King A.; Rappsilber J.; Calvio C.; Watson M.; Ajuh P.; Sleeman J.; *et al. Nat. Genet.* **1998**, *20*, 46-50.
 165. Nüsslein-Volhard C.; Wieschaus E. *Nature* **1980**, *287*, 795-801.
 166. O'Farrell P. H. *J. Biol. Chem.* **1975**, *250*, 4007-4021.
 167. Okada T.; Haze K.; Nadanaka S.; Yoshida H.; Seidah N.G.; Hirano Y.; Sato R.; *et al. J. Biol. Chem.* **2003**, *278*, 31024-31032.
 168. Oliver S. G.; Winson M. K.; Kell D. B.; Baganz F. *Trends Biotechnol.* **1998**, *16*, 373-378.
 169. Ong S. E.; Blagoev B.; Kratchmarova I.; Kristensen D. B.; Steen H.; Pandey A.; Mann M. *Mol. Cell Proteomics* **2002**, *1*, 376-386.
 170. Orłowski M.; Wilk S. *Arch. Biochem. Biophys.* **2000**, *383*, 1-16.
 171. Pandey A.; Mann M. *Nature* **2000**, *405*, 837-846.
 172. Pandey, A.; Lewitter, F. *Trends Biochem. Sci.* **1999**, *24*, 276-280.
 173. Parks B. A.; Jiang L.; Thomas P. M.; Wenger C. D.; Roth M. J.; Boyne Ii M. T.; *et al. Anal. Chem.* **2007**, *79*, 7984-7991.
 174. Pashkova A.; Moskovets E.; Karger B. L. *Anal. Chem.* **2004**, *76*, 4550-4557.
 175. Pemble C. W. 4th; Johnson L. C.; Kridel S. J.; Lowther W. T. *Nat. Struct. Mol. Biol.* **2007**, *14*, 704-709.
 176. Perrett D. *Proteomics Clin. Appl.* **2007**, *1*, 720-738.

-
177. Perutz M. F.; Rossmann M. G.; Cullis A. F.; Muirhead H.; Will G.; North A. C. *Nature* **1960**, *185*, 416-422.
178. Pfeffer M. A. *Am. Heart. J.* **1993**, *126*, 789-793.
179. Pullmann R. Jr.; Kim H. H.; Abdelmohsen K.; Lal A.; Martindale J. L.; Yang X.; Gorospe M. *Mol. Cell Biol.* **2007**, *27*, 6265-6278.
180. Rideout H. J.; Zang E.; Yeasmin M.; Gordon R.; Jabado O.; Park D.S.; Stefanis L. *Neuroscience* **2001**, *107*, 339-352.
181. Rojo E.; Zouhar J.; Carter C.; Kovaleva V.; Raikhel N. V. *Proc. Natl. Acad. Sci. USA.* **2003**, *100*, 7389-7394.
182. Rojo E.; Martín R.; Carter C.; Zouhar J.; Pan S.; Plotnikova J.; Jin H.; *et al.* *Curr. Biol.* **2004**, *14*, 1897-1906.
183. Rooney H. C. E.; van't Klooster J. W.; van der Hoorn R. A. L.; Joosten M. H. A. J.; Jones J. D. G.; de Wit P. J. G. M. *Science* **2005**, *308*, 1783-1786.
184. Rose J. K. C.; Ham K. S.; Darvill A. G.; Albersheim P. *Plant Cell* **2002**, *14*, 1329-1345.
185. Rosenfeld L. *Clin. Chem.* **2003**, *49*, 1696-1710.
186. Ross P. L.; Huang Y. N.; Marchese J. N.; Williamson B. *et al.* *Mol. Cell. Proteomics* **2004**, *3*, 1154-1169.
187. Pliura D. H.; Bonaventura B. J.; Smith R. A.; Coles P. J.; Krantz A. *Biochem. J.* **1992**, *288 (part 3)*, 759.
188. Saha S.; Sparks A. B.; Rago C.; Akmaev V.; Wang C. J.; Vogelstein B.; Kinzler K. W.; Velculescu V. E. *Nat. Biotechnol.* **2002**, *19*, 508-512.
189. Sala C.; Grainger D. C.; Cole S. T. *Cell Host Microbe.* **2009**, *5*, 430-437.
190. Sampson J. N.; Zhao H. *BMC Bioinformatics.* **2009**, *10*, 68-82.
191. Sanger F.; Air G. M.; Barrell B. G.; Brown N. L.; Coulson A. R.; Fiddes C. A.; Hutchison C. A.; *et al.* *Nature* **1977**, *265*, 687-695.
192. Sanger F.; Tuppy H. *Biochem J.* **1951**, *49*, 463-490.
-

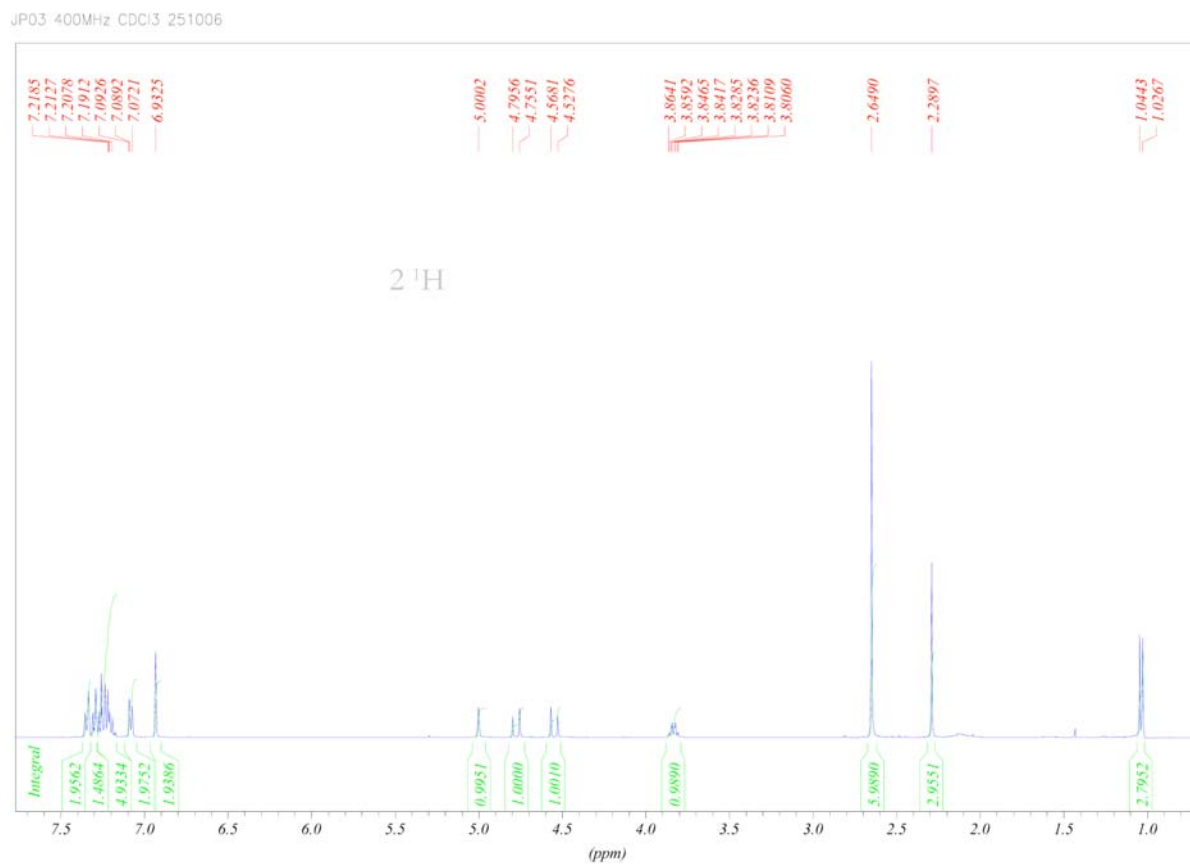
-
193. Sanger F.; Thompson E. O. *Biochem J.* **1953**, *53*, 353-374.
 194. Schaschke N. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 855-857.
 195. Scheele G. A. *J. Biol. Chem.* **1975**, *250*, 5375-5385.
 196. Schena M.; Shalon D.; Davis R. W.; Brown P. O. *Science* **1995**, *270*, 467–470.
 197. Schlosser A.; Vanselow J. T.; Kramer A. *Anal. Chem.* **2005**, *77*, 5243–5250.
 198. Schneider K.; Wang Z.M.; Kaiser M.; Kombrink E. *Curr. Topics in Phytochem.* **2008**, *9*, 1-16.
 199. Shi S. D. H.; Hemling M. E.; Carr S. A.; Horn D. M.; Lindh I.; McLafferty F. W. *Anal. Chem.* **2001**, *73*, 19-22.
 200. Shreder K. R.; Liu Y.; Nomanhboy T.; Fuller S. R.; Wong M. S.; Gai W. Z.; Wu J.; *et al.* *Bioconjug. Chem.* **2004**, *15*, 790-798.
 201. Schreiber S. L. *Bioorg. Med. Chem.* **1998**, *6*, 1127-1152.
 202. Schreiber S. L. *Nat. Chem. Biol.* **2005**, *1*, 64-66.
 203. Schuhr C.A.; Eisenreich W.; Goese M.; Stohler P.; Weber W.; Kupfer E.; Bacher A. *J. Org. Chem.* **2002**, *67*, 2257-2262.
 204. Schutkowski M.; Reineke U.; Reimer U. *Chembiochem* **2005**, *6*, 513-521.
 205. Shalon D.; Smith S. J.; Brown P. O. *Genome Res.* **1996**, *6*, 639-645.
 206. Shao F.; Merritt P. M.; Bao Z.; Innes R. W.; Dixon J. E. *Cell* **2002**, *109*, 575-588.
 207. Shao F.; Golstein C.; Ade J.; Stoutemyer M.; Dixon J.E.; Innes R.W. *Science* **2003**, *301*, 1230-1233.
 208. Sidhu S. S.; Koide S. *Curr. Opin. Struct. Biol.* **2007**, *17*, 481-487.
 209. Smith G. P. *Science* **1985**, *228*, 1315–1317.
 210. Speers A. E.; Adam G. C.; Cravatt B. F. *J. Am. Chem. Soc.* **2003**, *125*, 4686-4687.
 211. Speers A. E.; Cravatt B. F. *Chem. Biol.* **2004**, *11*, 535-546.

-
212. Steitz T. A.; Henderson R.; Blow D. M. *J. Mol. Biol.* **1969**, *46*, 337-348.
 213. Stirewalt M. A. *Ann N Y Acad Sci.* **1963**, *113*, 36-53.
 214. Stoyanova S.; Bulgarelli-Leva G.; Kirsch C.; Hanck T.; Klinger R.; Wetzker R.; Wymann M. P. *Biochem. J.* **1997**, *324* (Pt 2), 489-495.
 215. Strasser A.; Jost P. J.; Nagata S. *Immunity* **2009**, *30*, 180-192.
 216. Strausberg R. L.; Feingold E. A.; Klausner R. D.; Collins F. S. *Science* **1999**, *286*, 455-457.
 217. Suarez M. F.; Filonova L. H.; Smertenko A.; Savenkov E. I.; Clapham D. H. *et al. Curr. Biol.* **2004**, *14*, R338-340.
 218. Sumner J. B. *J. Biol. Chem.* **1926**, *69*, 435-441.
 219. Sun X.; Chiu J. F.; He Q. Y. *Expert Rev. Proteomics* **2005**, *2*, 649-657.
 220. Suter B.; Kittanakom S.; Stagljar I. *Curr. Opin. Biotechnol.* **2008**, *19*, 316-323.
 221. Swiderski M. R.; Innes R.W. *Plant J.* **2001**, *26*, 101-112.
 222. Syka J. E. P.; Coon J. J.; Schroeder M. J.; Shabanowitz J.; Hunt D. F. *Proc. Natl. Acad. Sci. USA.* **2004**, *101*, 9528-9533.
 223. Tampakaki A. P.; Bastaki M.; Mansfield J. W.; Panopoulos N. *J. Mol. Plant Microbe Interact.* **2002**, *15*, 292-300.
 224. Thompson A.; Schaefer J.; Kuhn K.; Kienle S. *et al., Anal. Chem.* **2003**, *75*, 1895-1904. Correction published in *Anal. Chem.* **2006**, *78*, 4235.
 225. Uttamchandani M.; Wang J.; Yao S. Q. *Mol. Biosyst.* **2006**, *2*, 58-68.
 226. Van der Biezen E.A.; Jones J.D. *Trends Biochem. Sci.* **1998**, *23*, 454-456.
 227. Van der Hoorn R. A. L. *Ann. Rev. Plant Biol.* **2008**, *59*, 191-223.
 228. Van der Hoorn R. A. L.; Kamoun S. *Plant Cell* **2008**, *20*, 2009-2017.
 229. Van der Hoorn R. A. L.; Leeuwenburgh M. A.; Bogyo M.; Joosten M. H. A. J.; Peck, S.C. *Plant Physiol.* **2004**, *135*, 1170-1178.
 230. Van Roessel P.; Brand A. H. *Nat. Cell Biol.* **2002**, *4*, E15-20.
-

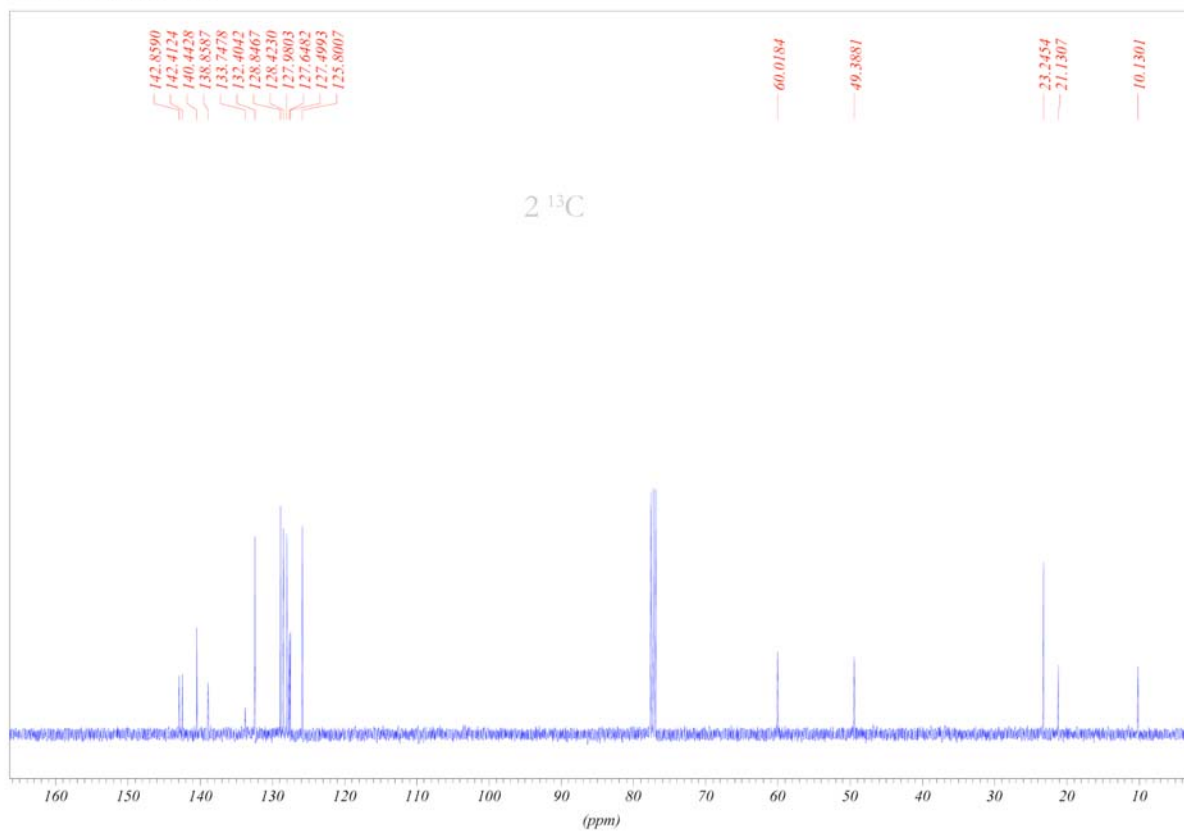
-
231. Velculescu V. E.; Zhang L; Vogelstein B.; Kinzler K. W. *Science* **1995**, *270*, 484–487.
232. Venable J. D.; Wohlschlegel J.; McClatchy D. B.; Sung K. P.; Yates Iii J. R. *Anal. Chem.* **2007**, *79*, 3056-3064.
233. Istrail S.; Sutton G. G.; Florea L.; Halpern A. L.; Mobarry C. M.; Lippert R.; Venter J. C. *et al.*, *Science* **2001**, *291*, 1304–1351.
234. Vignal A.; Milan D.; Sancristobal M.; Eggen A. *Genet. Sel. Evol.* **2002**, *34*, 275-305.
235. Vosseller K.; Trinidad J. C.; Chalkley R. J.; Specht C. G. *et al.*, *Mol. Cell. Proteomics* **2006**, *5*, 923–934.
236. Walsh D. P.; Chang Y. T. *Chem. Rev.* **2006**, *106*, 2476-2530.
237. Wang S.; Regnier F. E.; *J. Chromatogr. A* **2001**, *924*, 345–357.
238. Wang Z.M.; Gu C.; Colby T.; Shindo T.; Balamurugan R.; Waldmann H.; Kaiser M.; Van der Hoorn R. A. L. *Nat. Chem. Biol.* **2008**, *4*, 557-563.
239. Wang Z.; Gerstein M.; Snyder M. *Nat. Rev. Genet.* **2009**, *10*, 57-63.
240. Warren R.F.; Henk A.; Mowery P.; Holub E.; Innes R.W. *Plant Cell* **1998**, *10*,1439-1452.
241. Weibel E.K.; Hadvary P.; Hochuli E.; Kupfer E.; Lengsfeld H. *J. Antibiot. (Tokyo)* **1987**, *40*, 1081-1085.
242. Weisenberg R. C. *Cell Motil.* **1981**, *1*, 485-497.
243. Wess J. *FASEB J.* **1997**, *11*, 346-354.
244. Wilkins M. R.; Pasquali C.; Appel R. D.; Ou K.; Golaz O.; Sanchez J. C.; Yan J. X.; *et al.* *Nat. Biotechnol.* **1996**, *14*, 61-65.
245. Winzeler E. A.; Shoemaker D. D.; Astromoff A.; Liang H.; Anderson K.; Andre B.; Bangham R. *et al.* *Science* **1999**, *285*, 901–906.
246. Wolschin F.; Wienkoop S.; Weckwerth W. *Proteomics* **2005**, *5*, 4389–4397.

-
247. Xia Y.; Suzuki H.; Borevitz J.; Blount J.; Guo Z.; *et al.* *EMBO J.* **2004**, *23*, 980–988.
248. Yamada K.; Shimada T.; Kondo M.; Nishimura M.; Hara-Nishimura I. *J. Biol. Chem.* **1999**, *274*, 2563-2570.
249. Yee M. C.; Fas S. C.; Stohlmeyer M. M.; Wandless T. J.; Cimprich K. A. *J Biol Chem.* **2005**, *280*, 29053-29059.
250. Yin J.; Bergmann E. M.; Cherney M. M.; Lall M. S.; Jain R. P.; Vederas J. C.; James M. N. *J. Mol. Biol.* **2005**, *354*, 854-871.
251. Zambrowicz B. P.; Friedrich G. A.; Buxton E. C.; Lilleberg S. L.; Person C.; Sands A. T. *Nature* **1998**, *392*, 608–611.
252. Zhang H.; Li X. j.; Martin D. B.; Aebersold R. *Nat. Biotechnol.* **2003**, *21*, 660–666.
253. Zhi J.; Melia A.T.; Eggers H.; Joly R.; Patel I.H. *J. Clin. Pharmacol.* **1995**, *35*, 1103-1108.
254. Zhou H.; Ranish J. A.; Watts J. D.; Aebersold R. *Nat. Biotechnol.* **2002**, *19*, 512–515.
255. Zhu M.; Shao F.; Innes R.W.; Dixon J.E.; Xu Z. *Proc. Natl. Acad. Sci. USA.* **2004**, *101*, 302-307.
256. Zimmermann R. E.; Lubinus J. *Klin. Wochenschr.* **1978**, *56*, 781-788.
257. Zubarev R. A.; Kelleher N. L.; McLafferty F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.

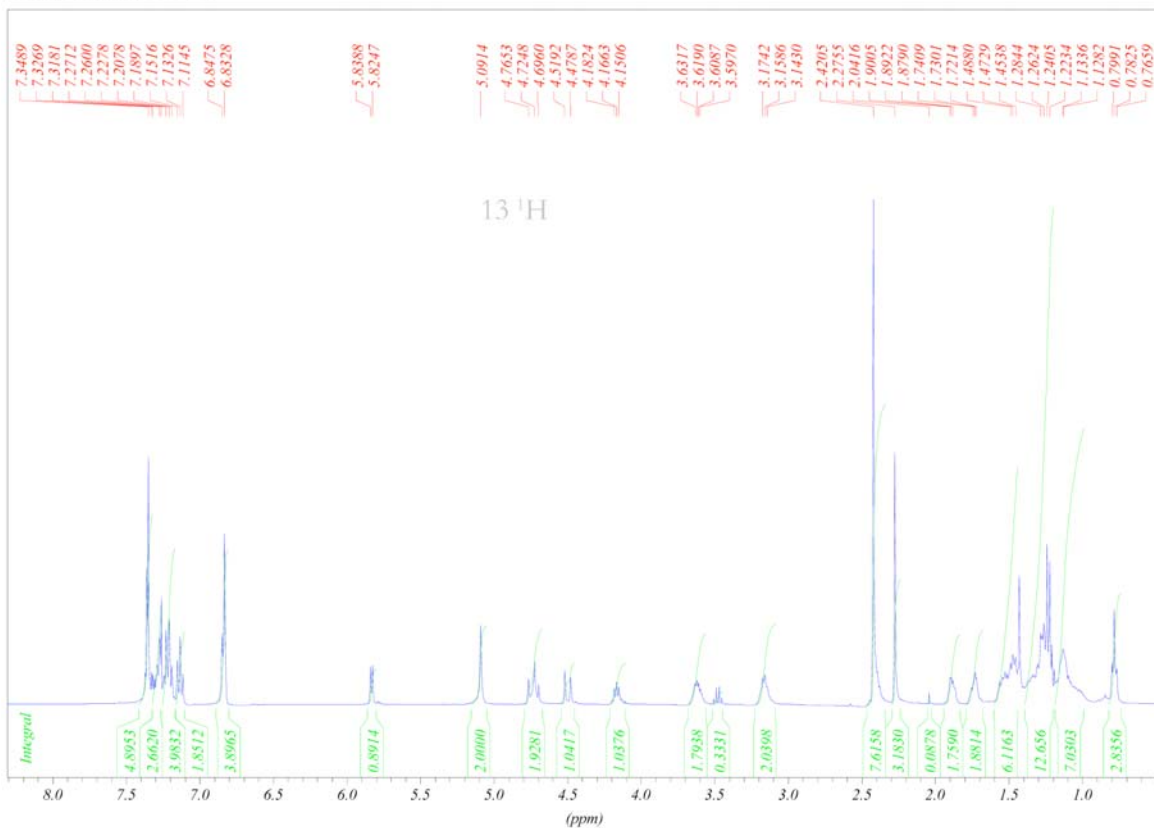
8 NMR Spectra

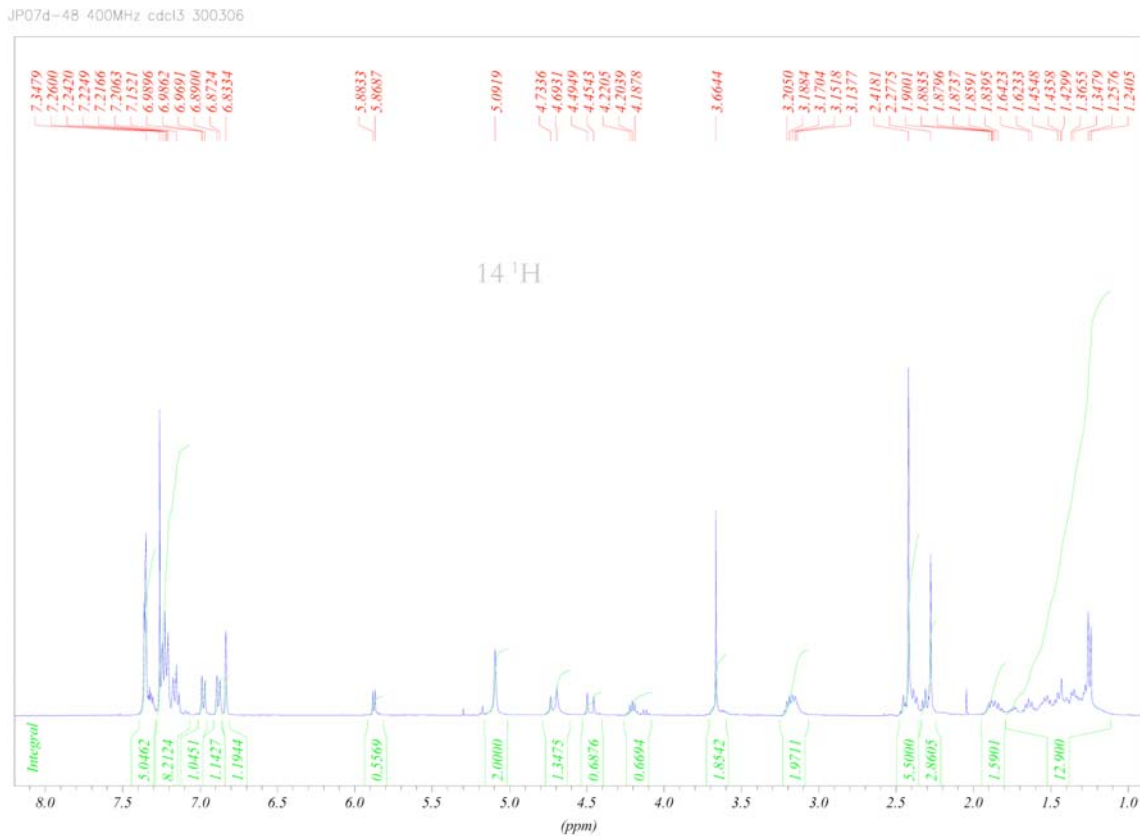
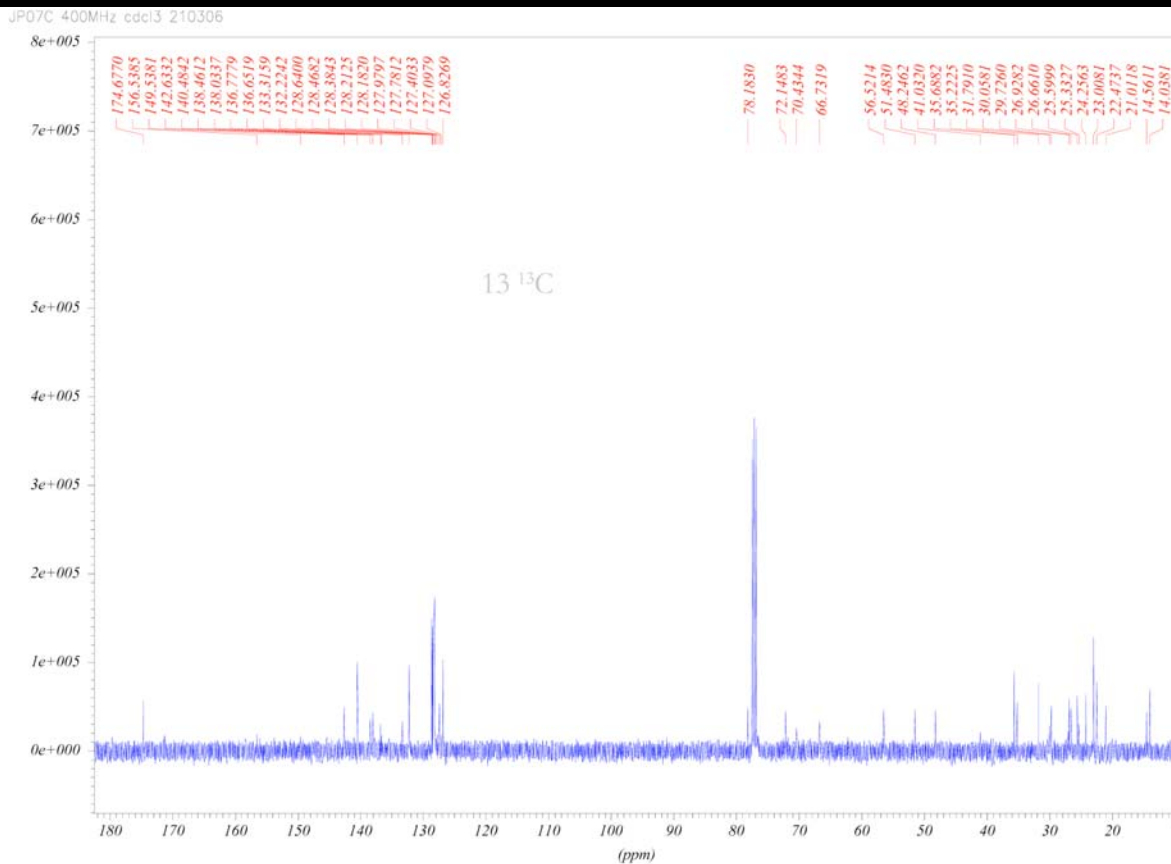


JP03 400MHz CDCl3 251006

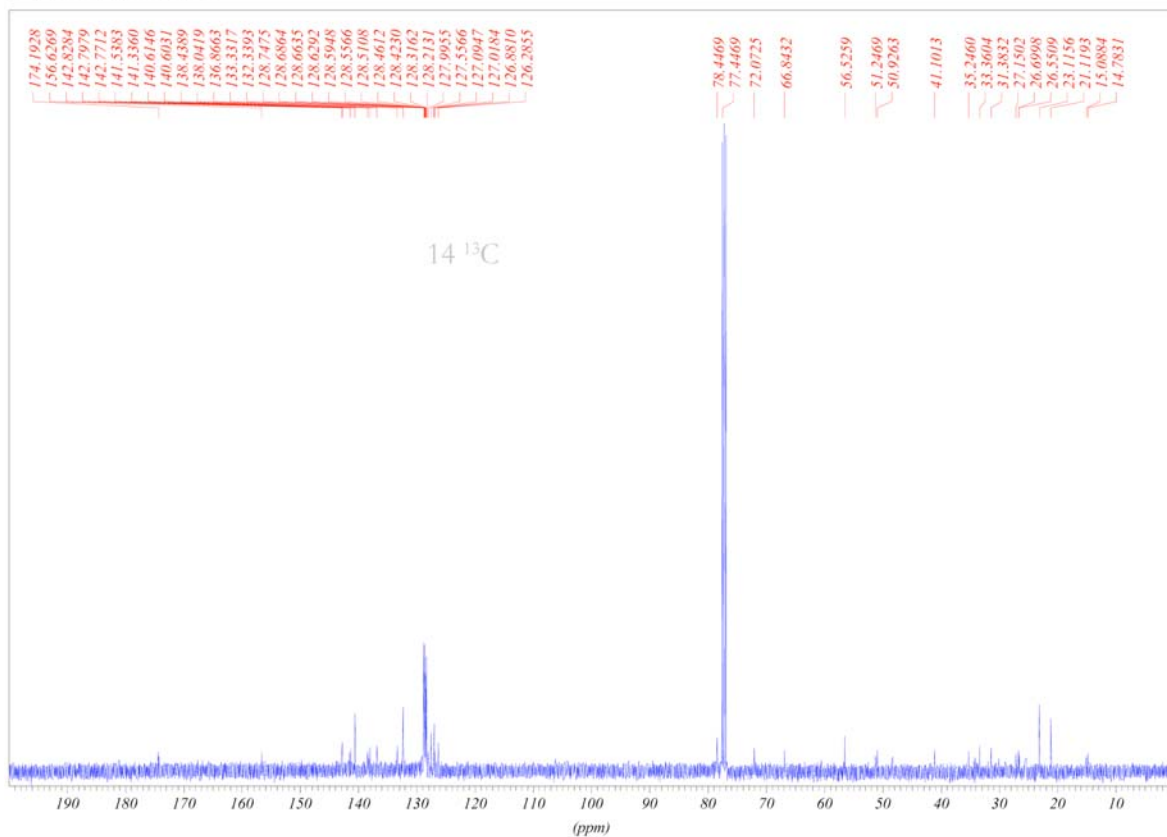


JP07C 400MHz cdcl3 210306

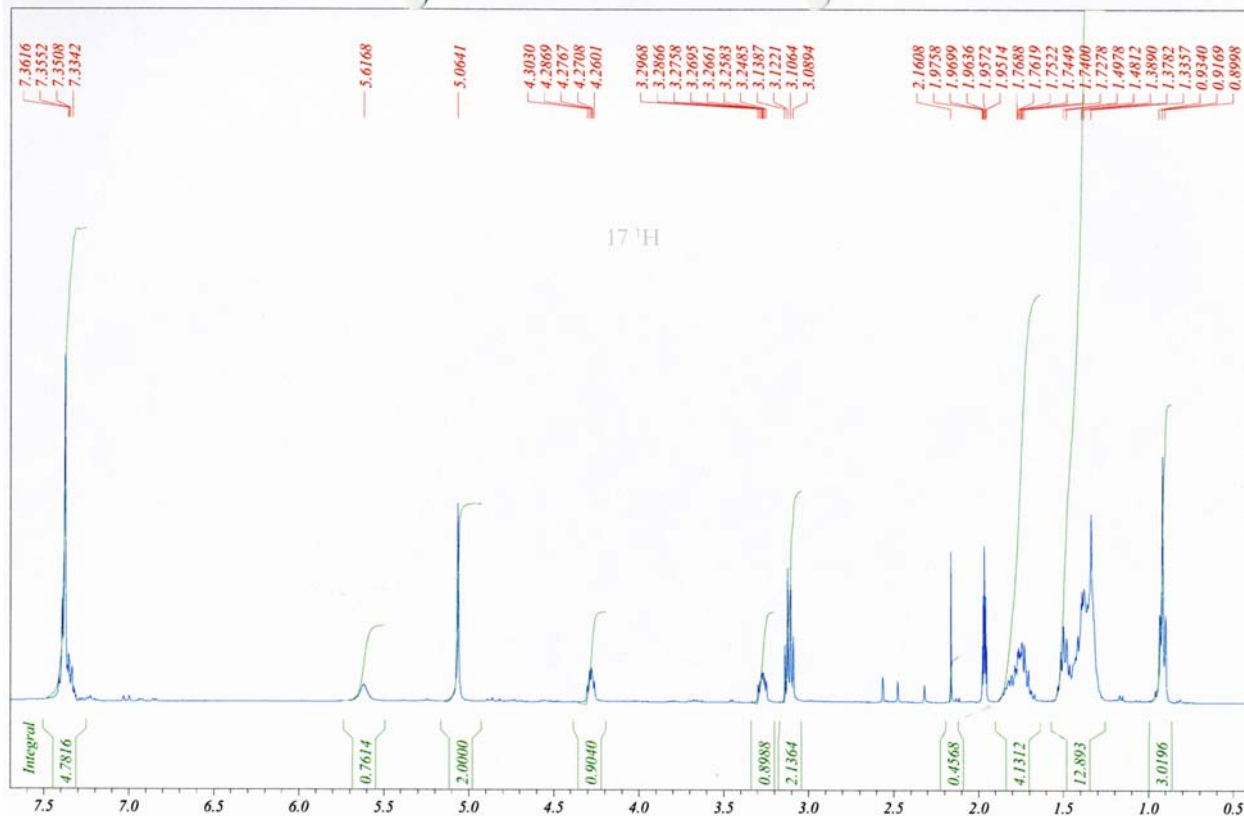


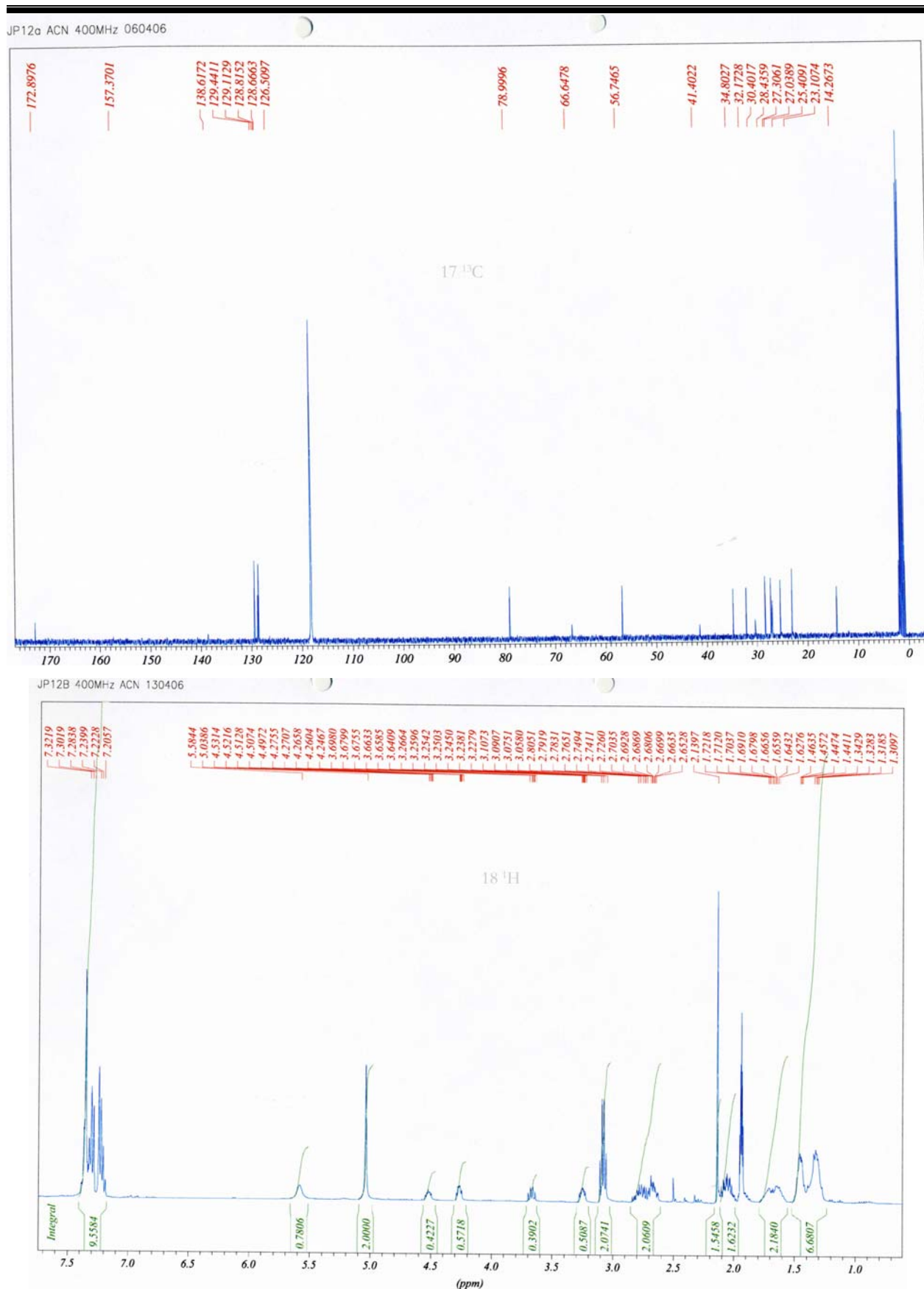


JP07d 400MHz cdcl3 040406

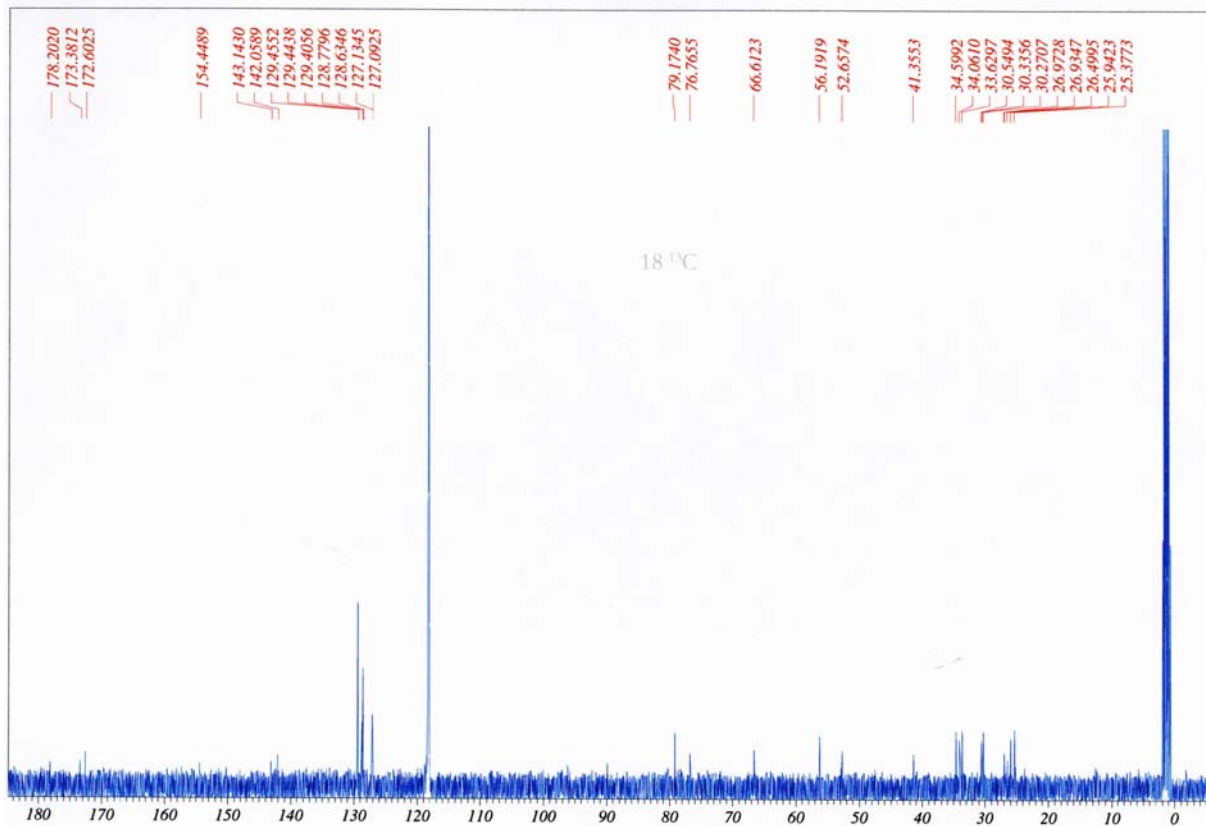


JP12a ACN 400MHz 060406

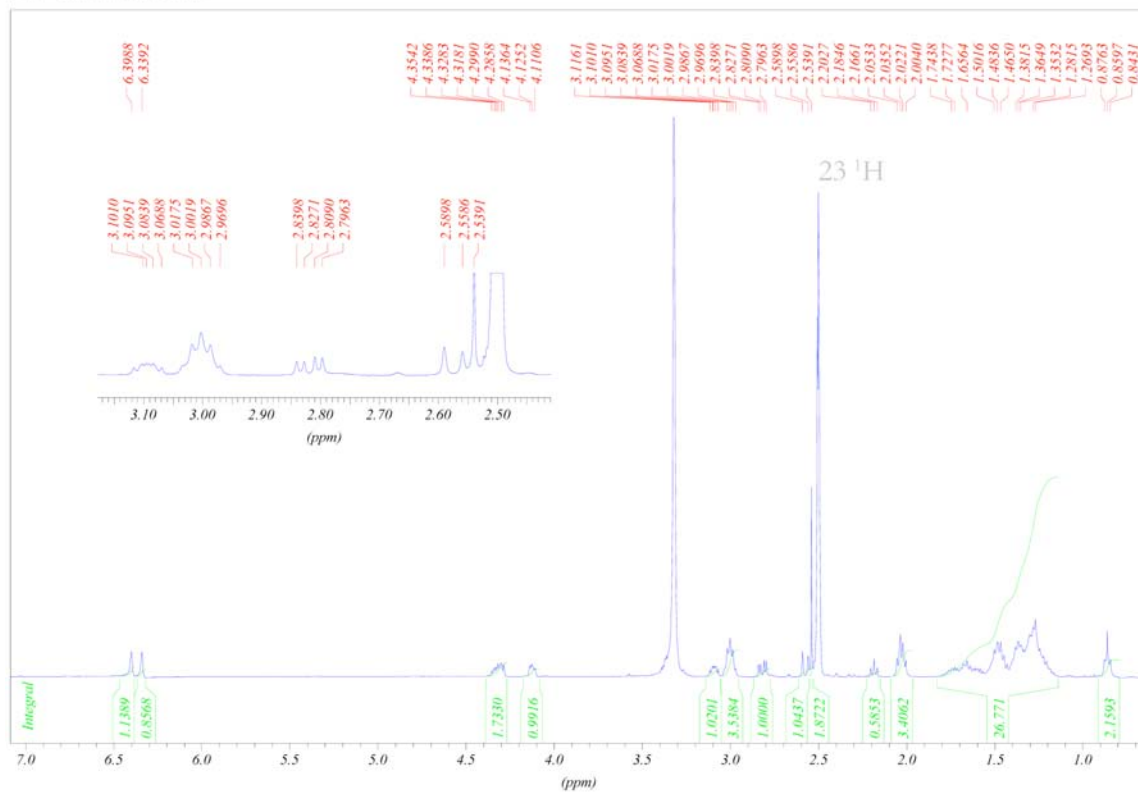




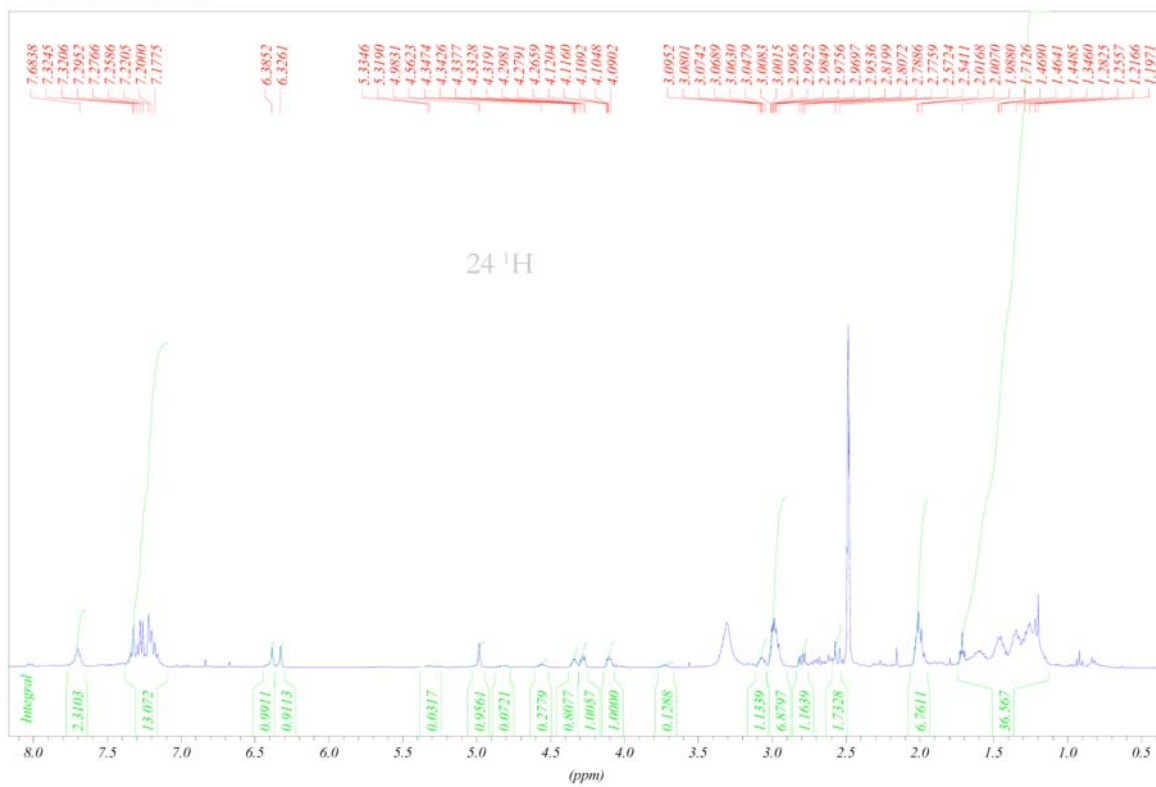
JP12B 400MHz ACN 130406



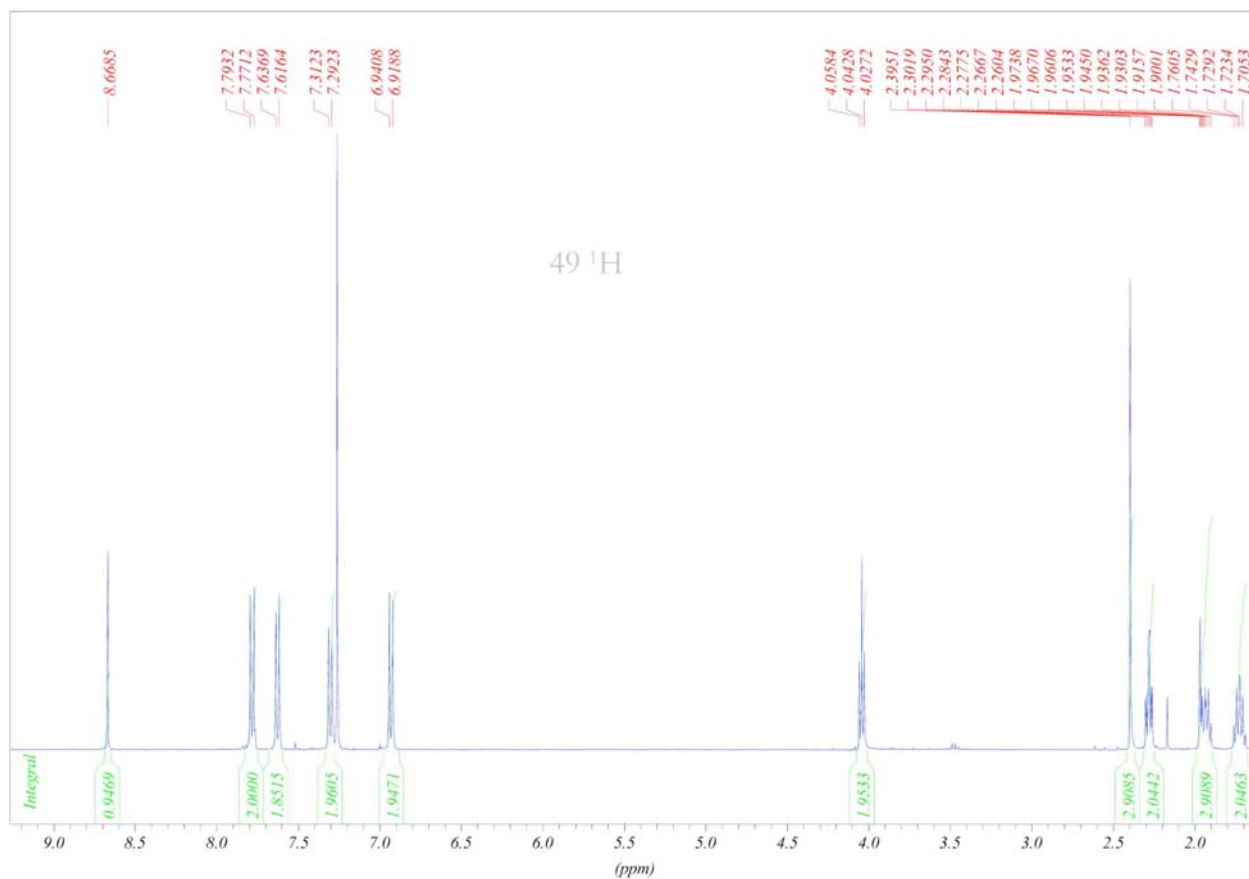
JP10F 400MHz DMSO 090409

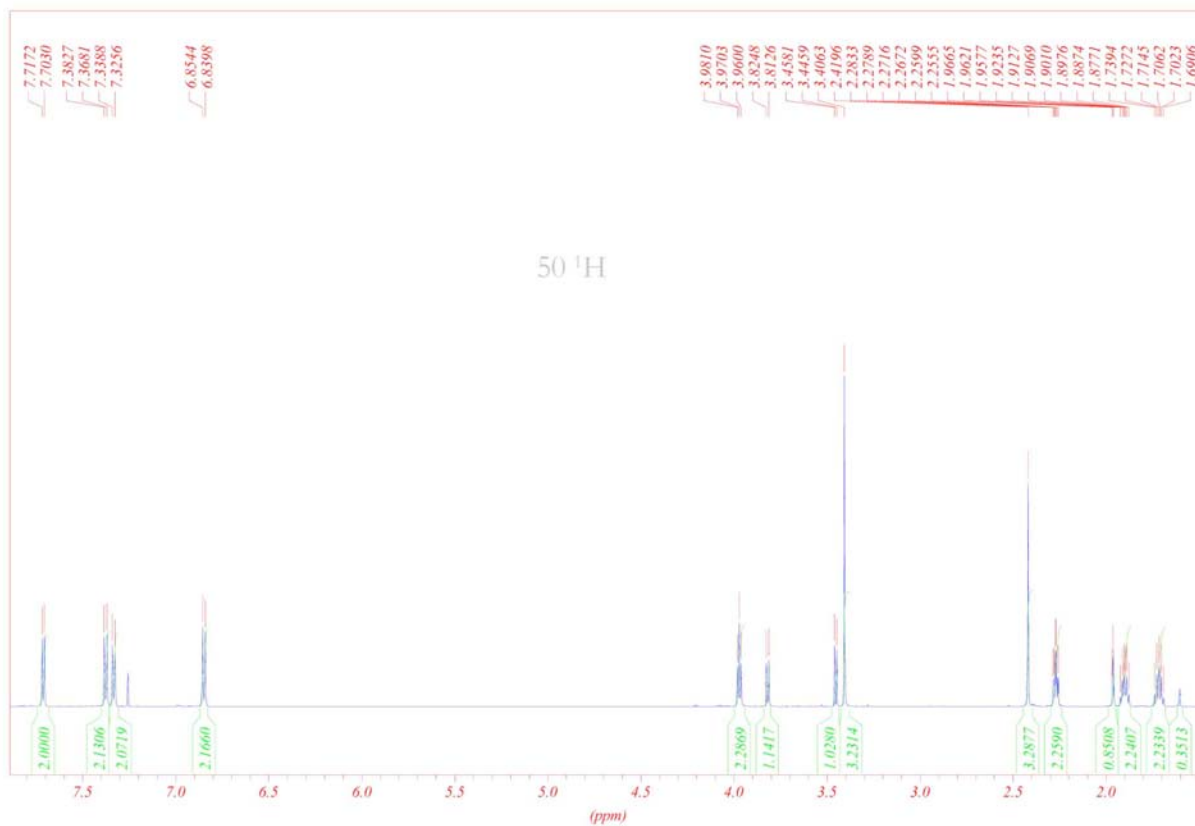
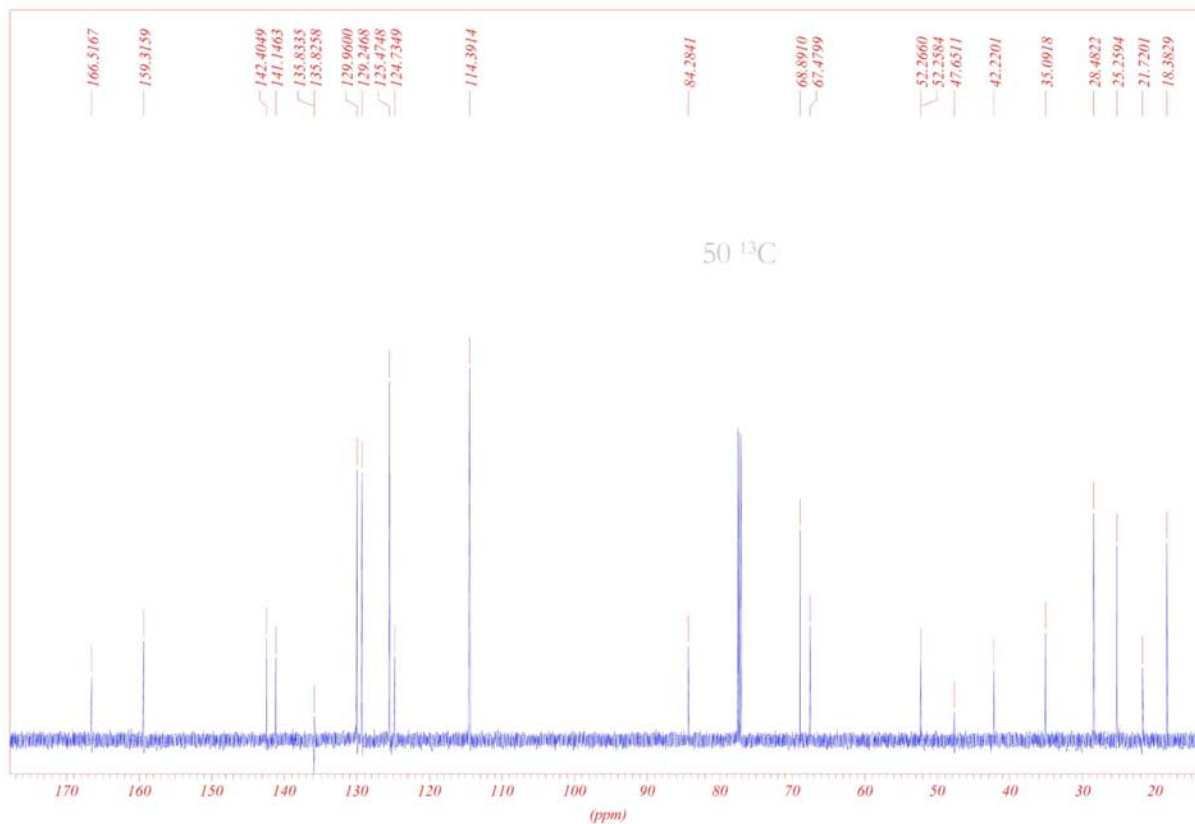


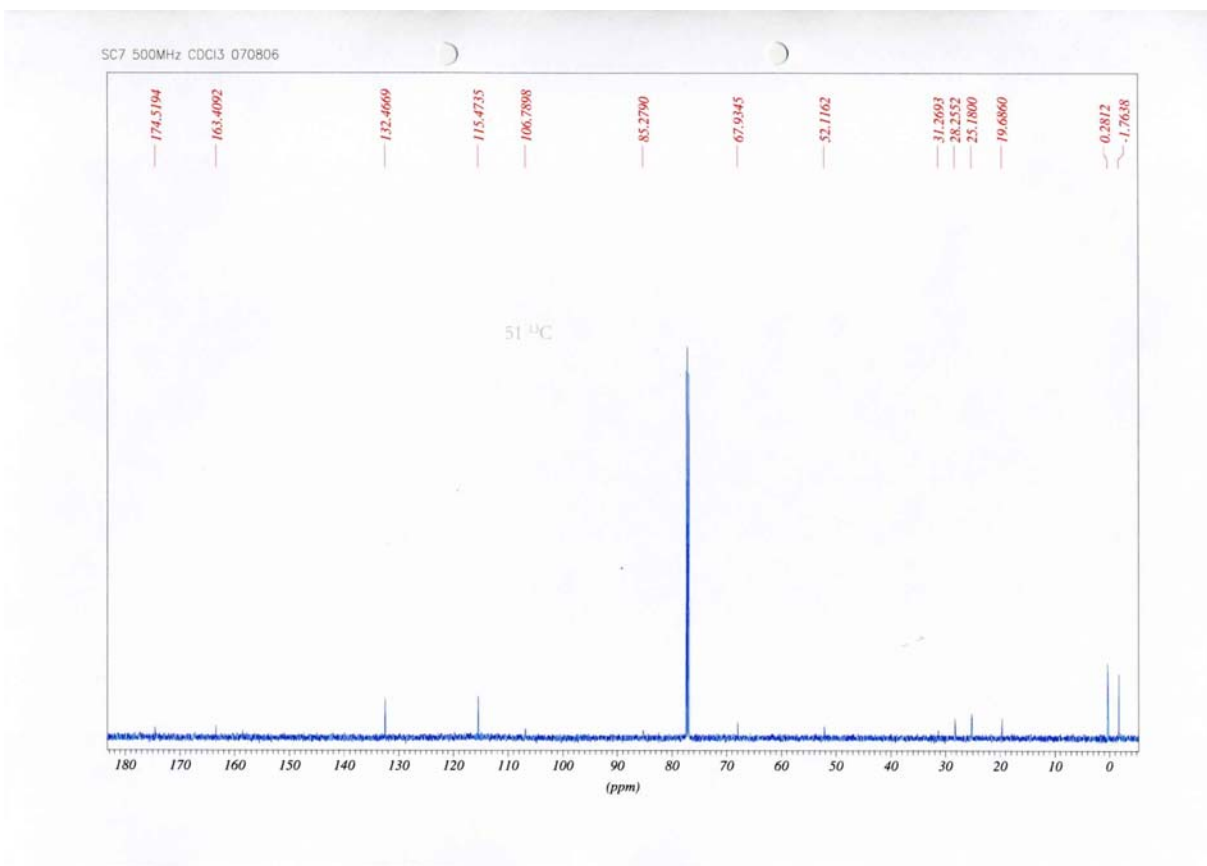
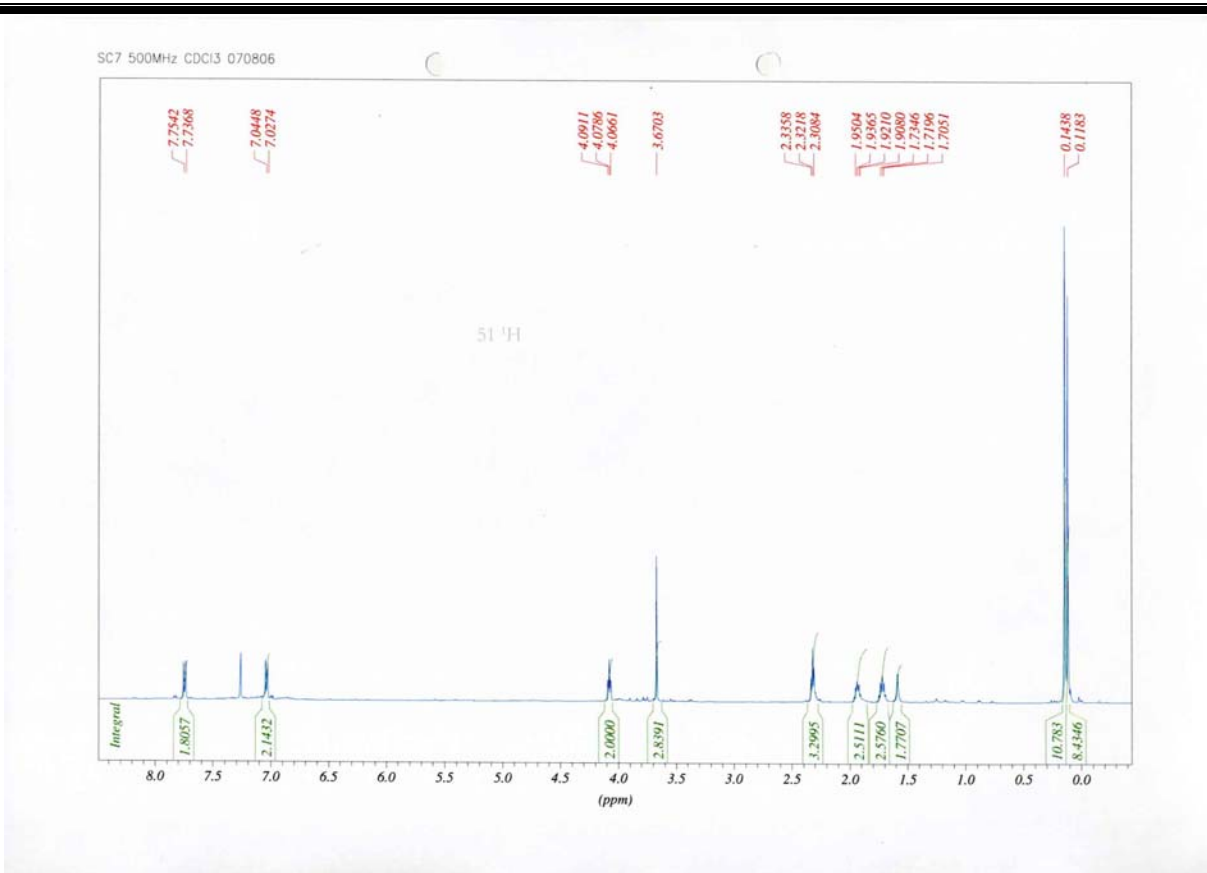
JP10g DMSO 400MHz 16040b



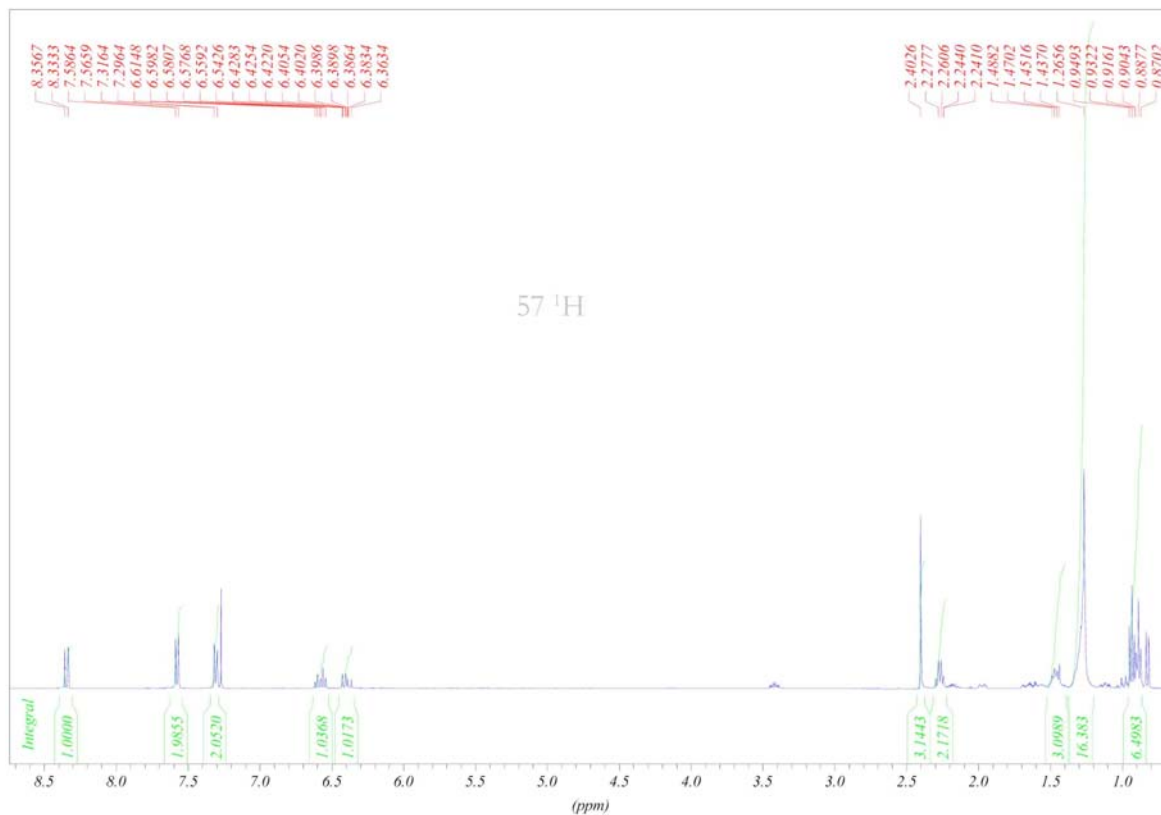
SC5 400MHz CDCl3 020806



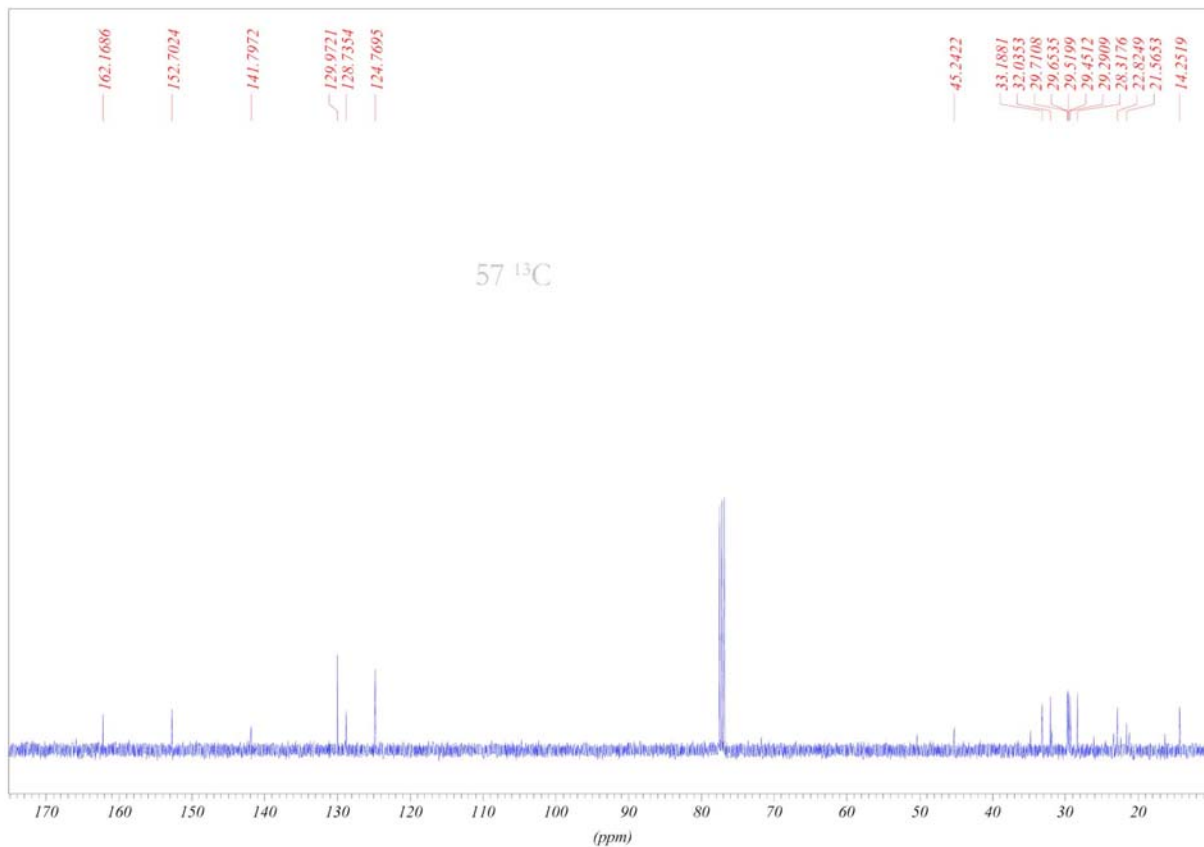
SC6 600MHz CDCl₃ 040806SC6 600MHz CDCl₃ 040806



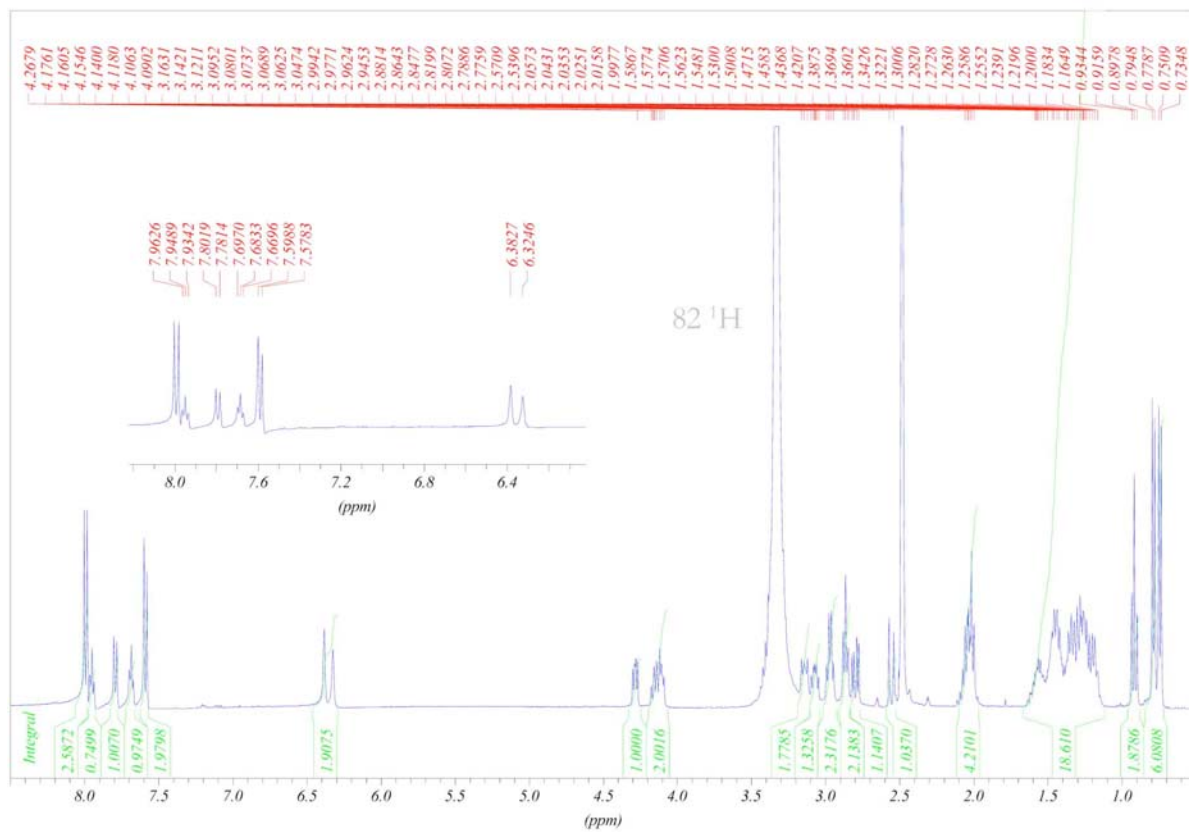
SC09S 400MHz CDCl3 100707



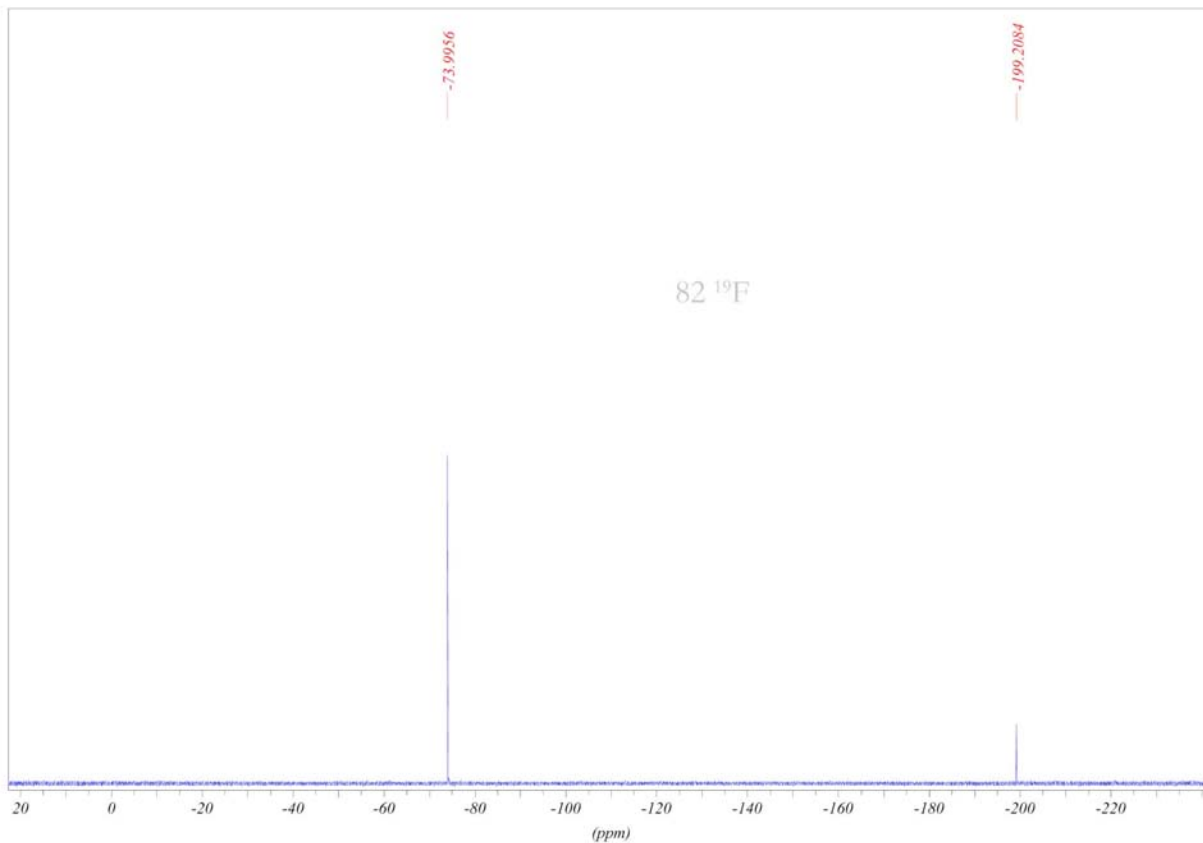
SC09S 400MHz CDCl3 100707



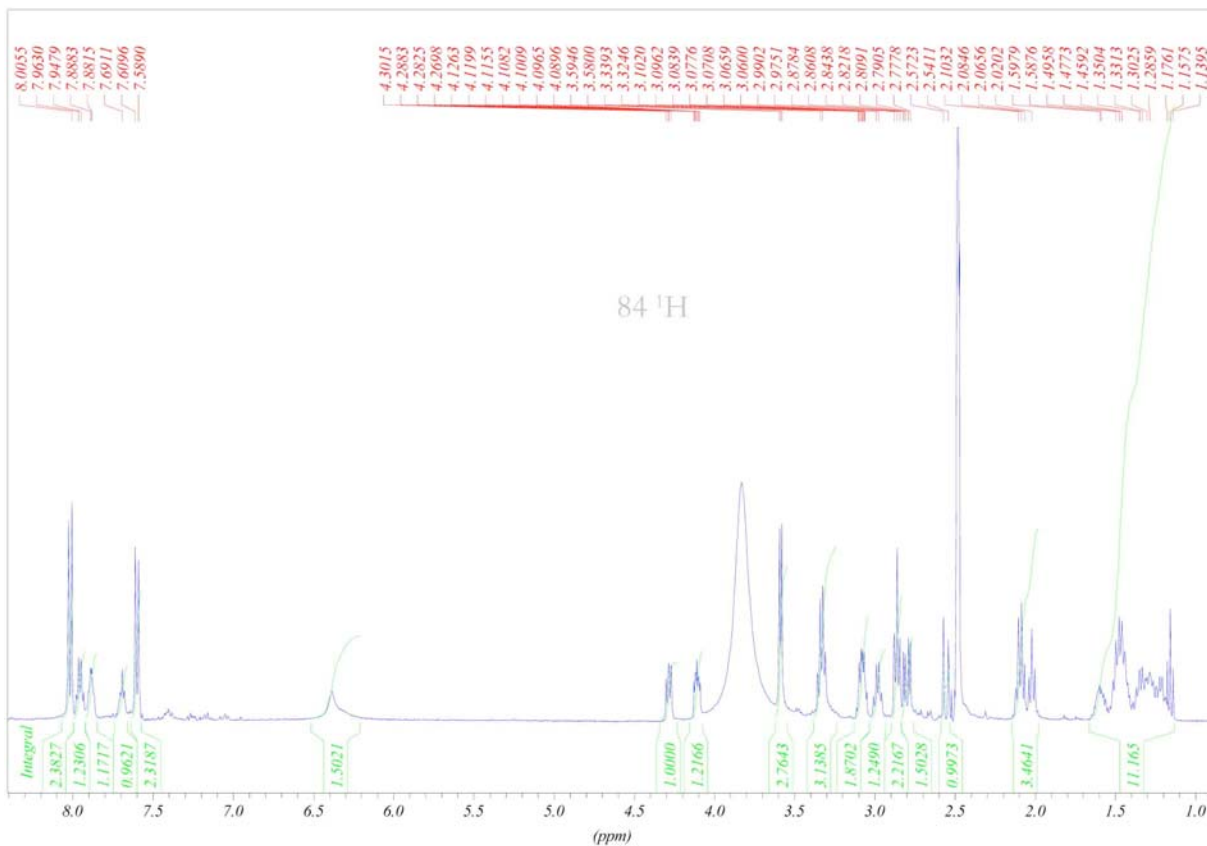
DS1 400MHz DMSO d90207



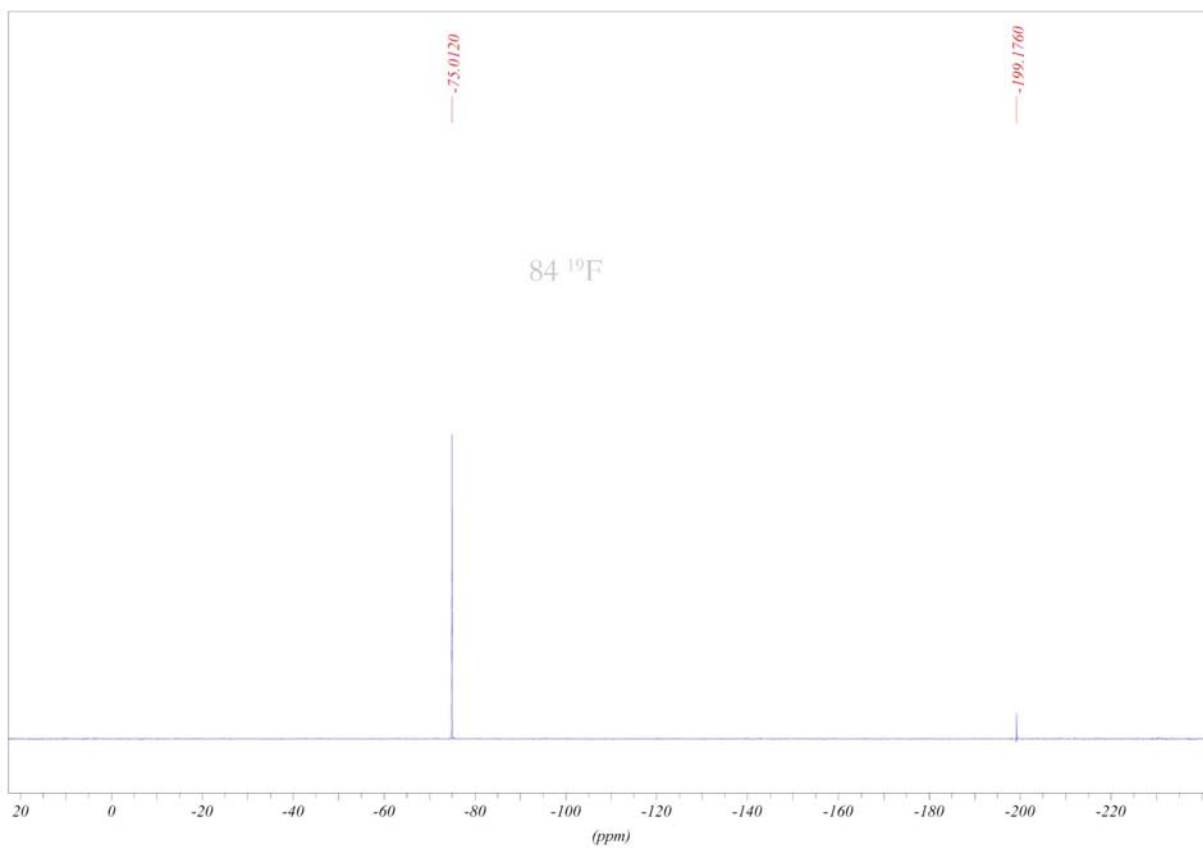
DS1 400MHz DMSO d90207

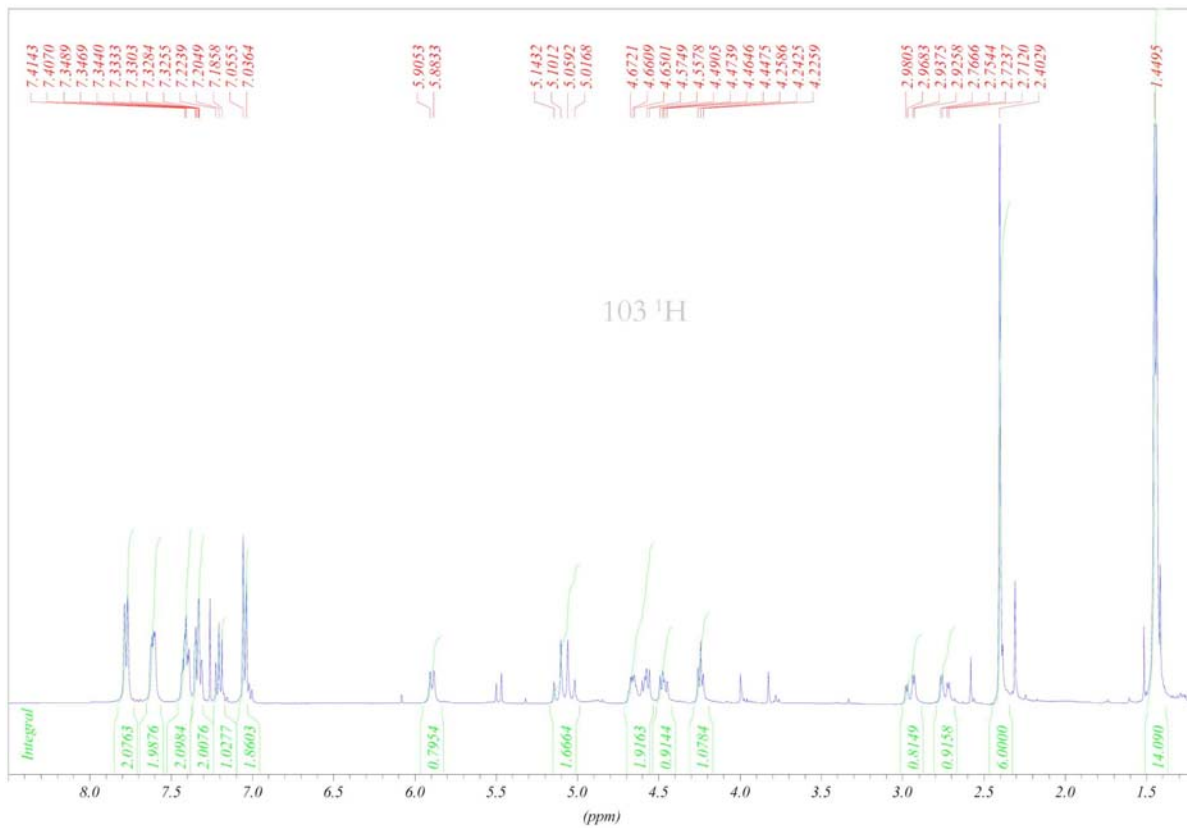
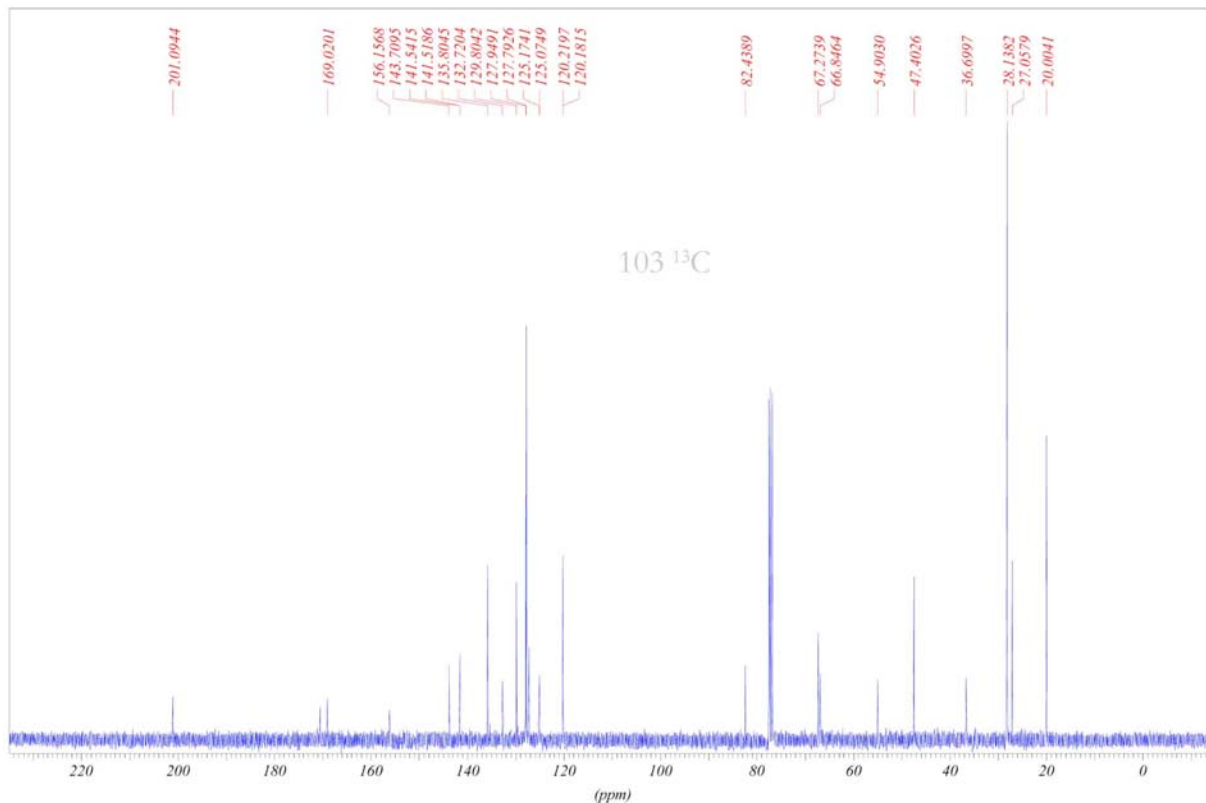


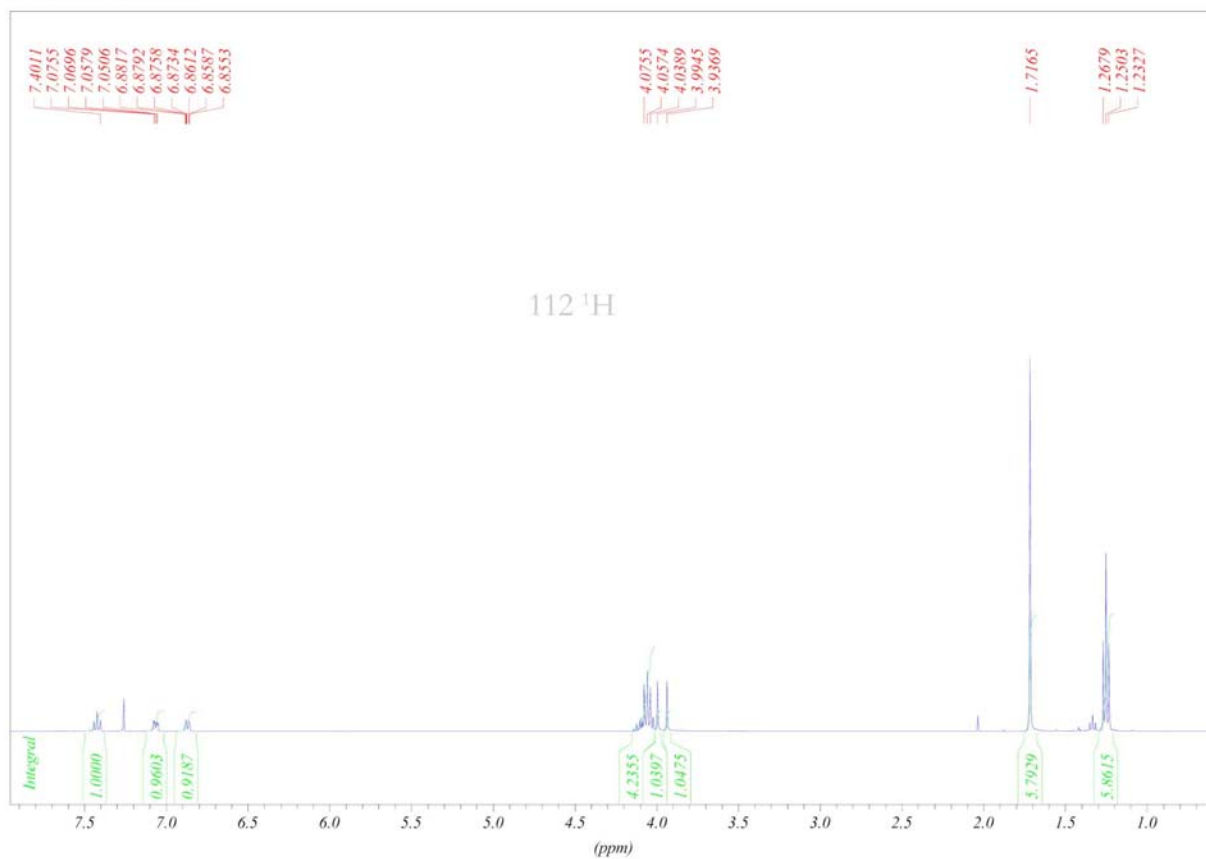
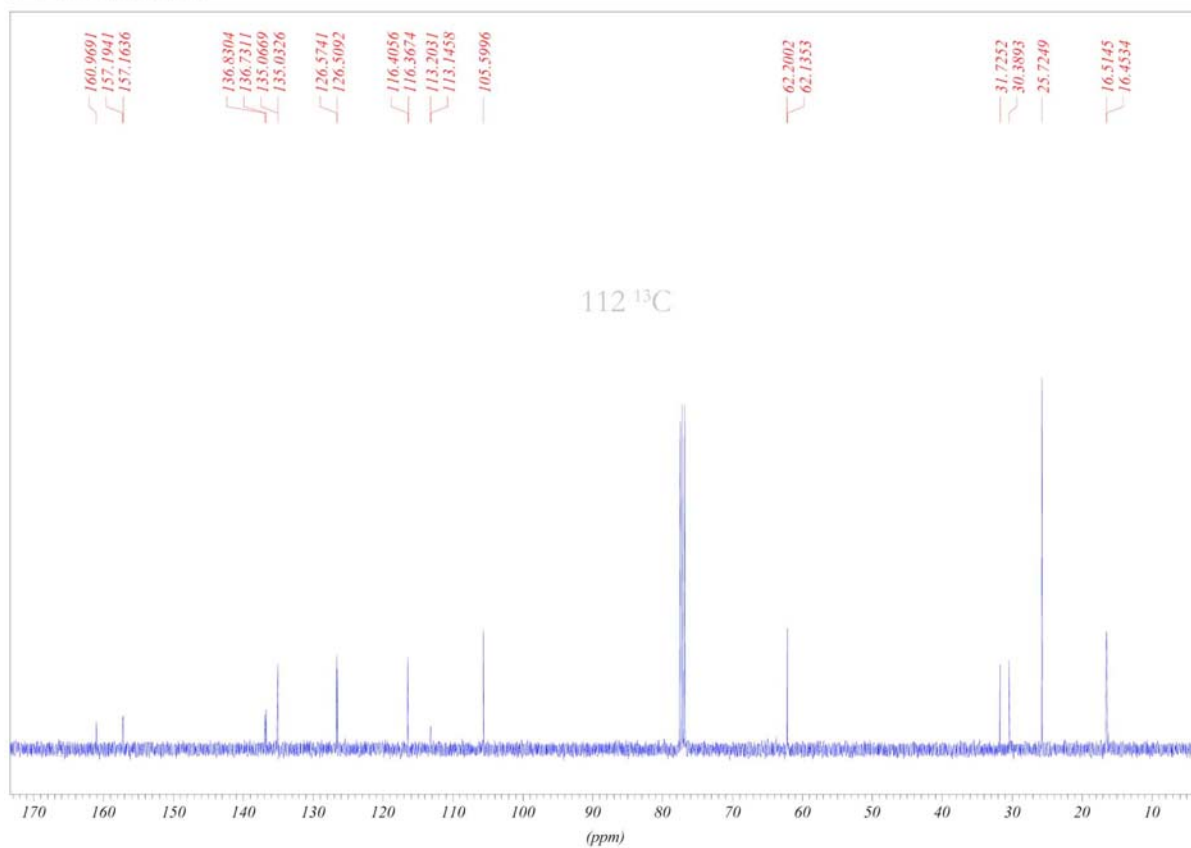
DS6 400MHz CDCl3 150207

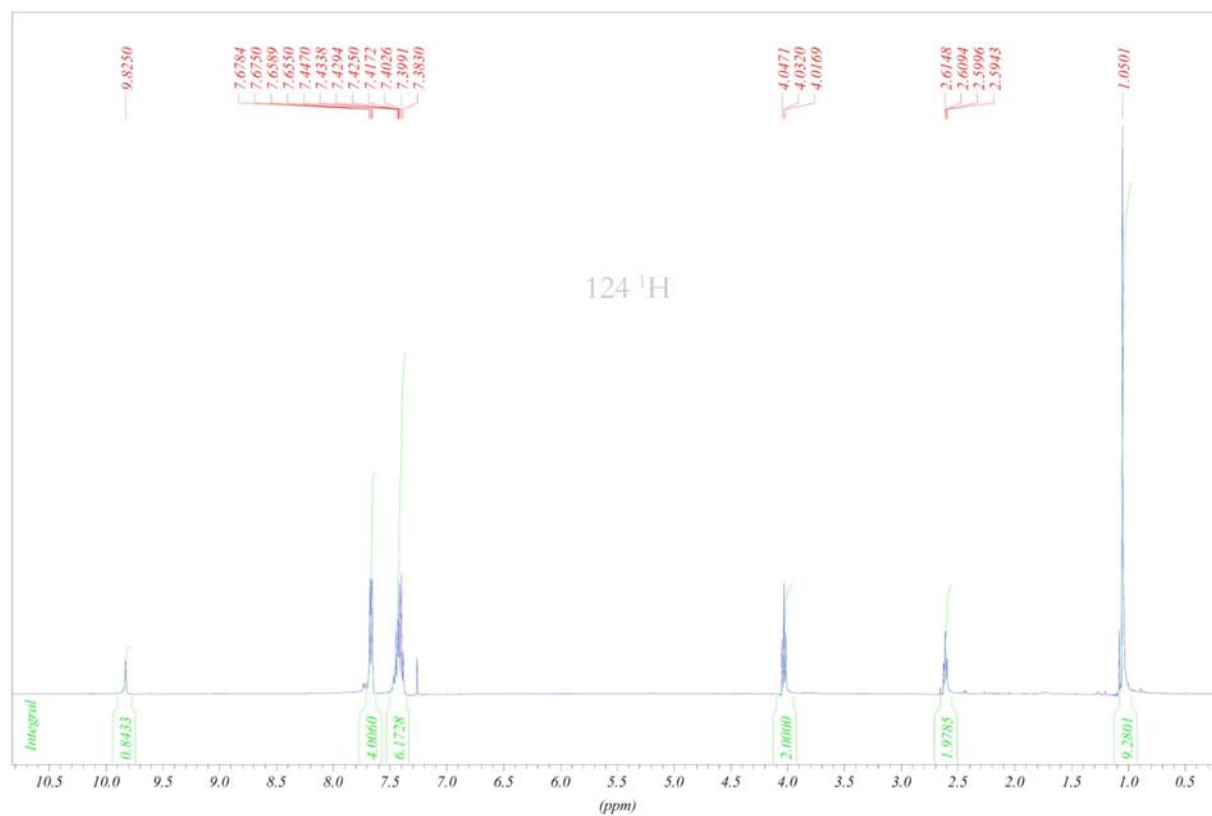
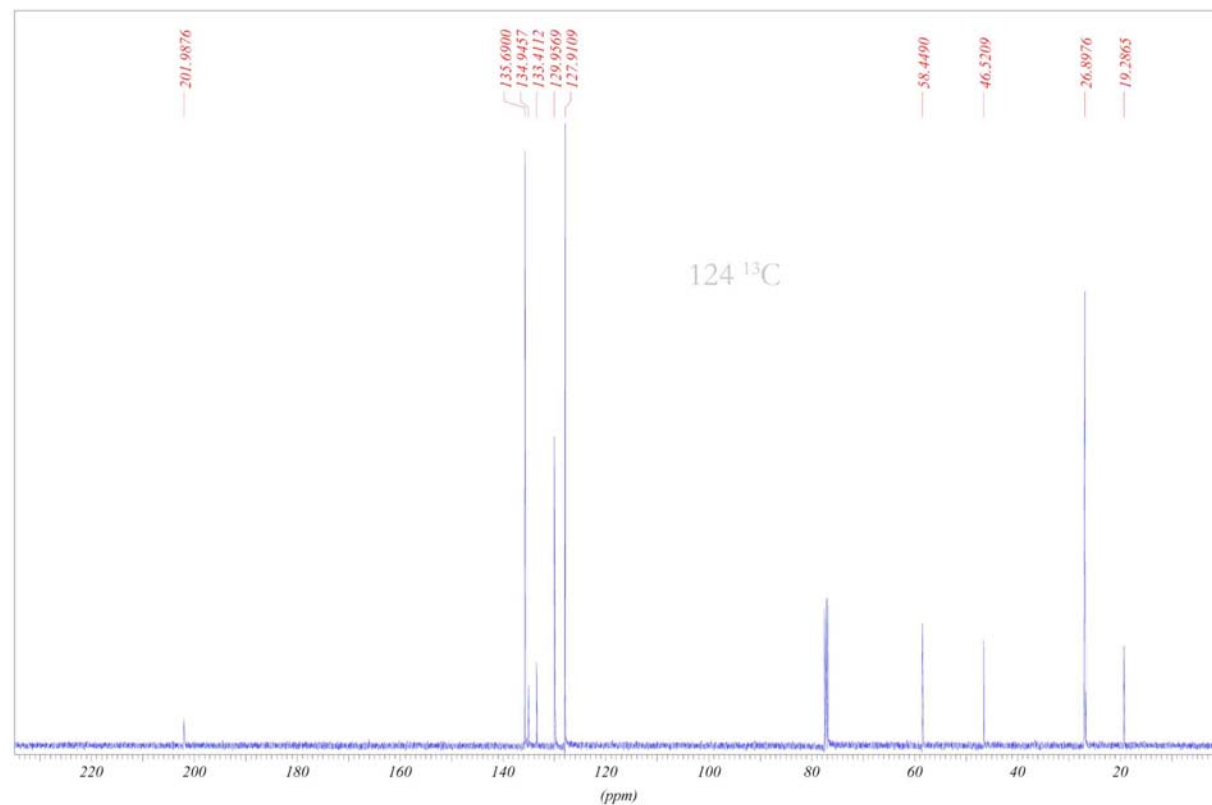


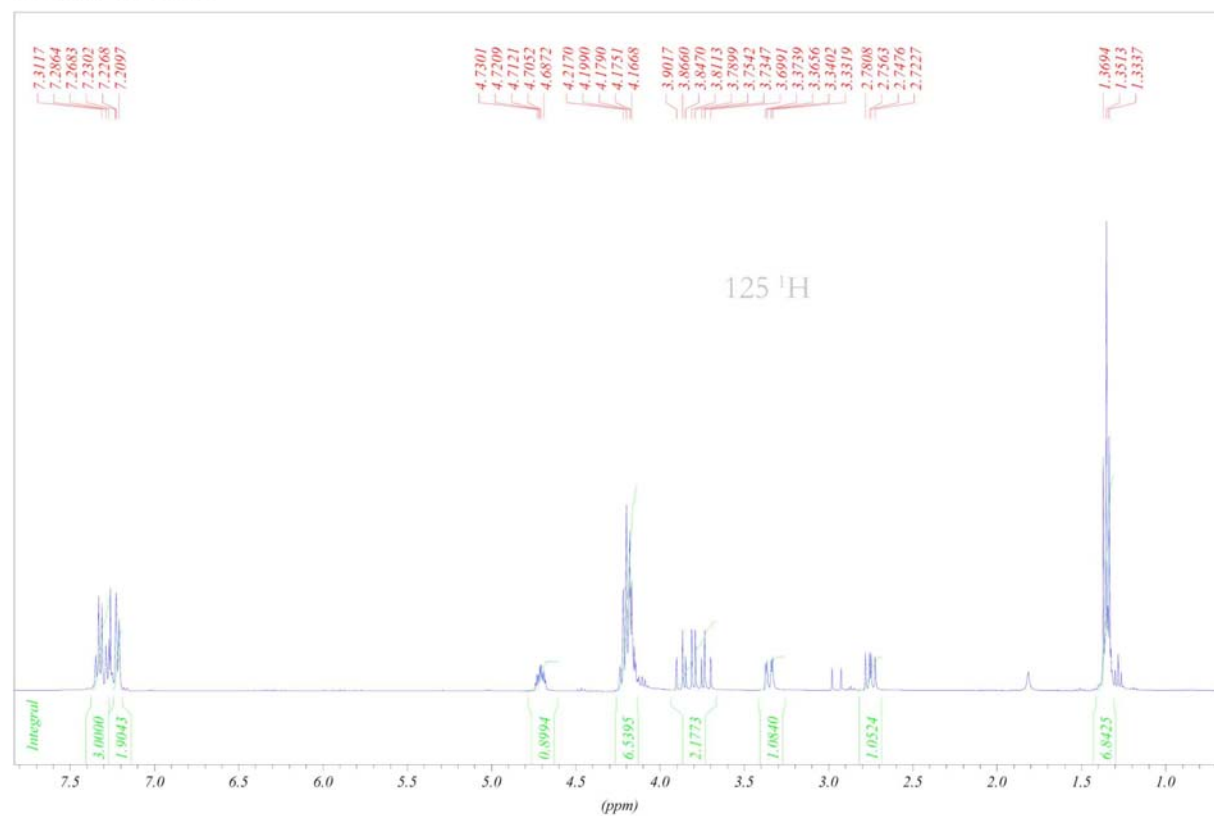
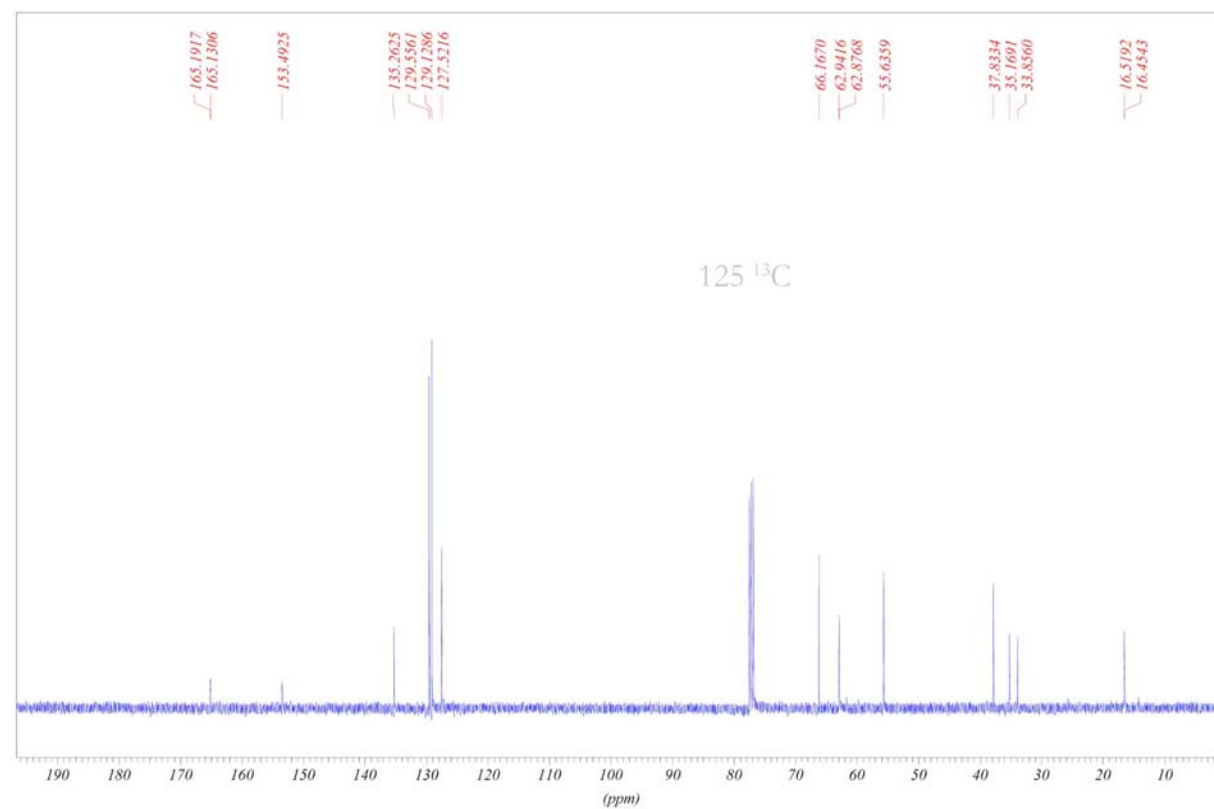
DS6 400MHz CDCl3 150207



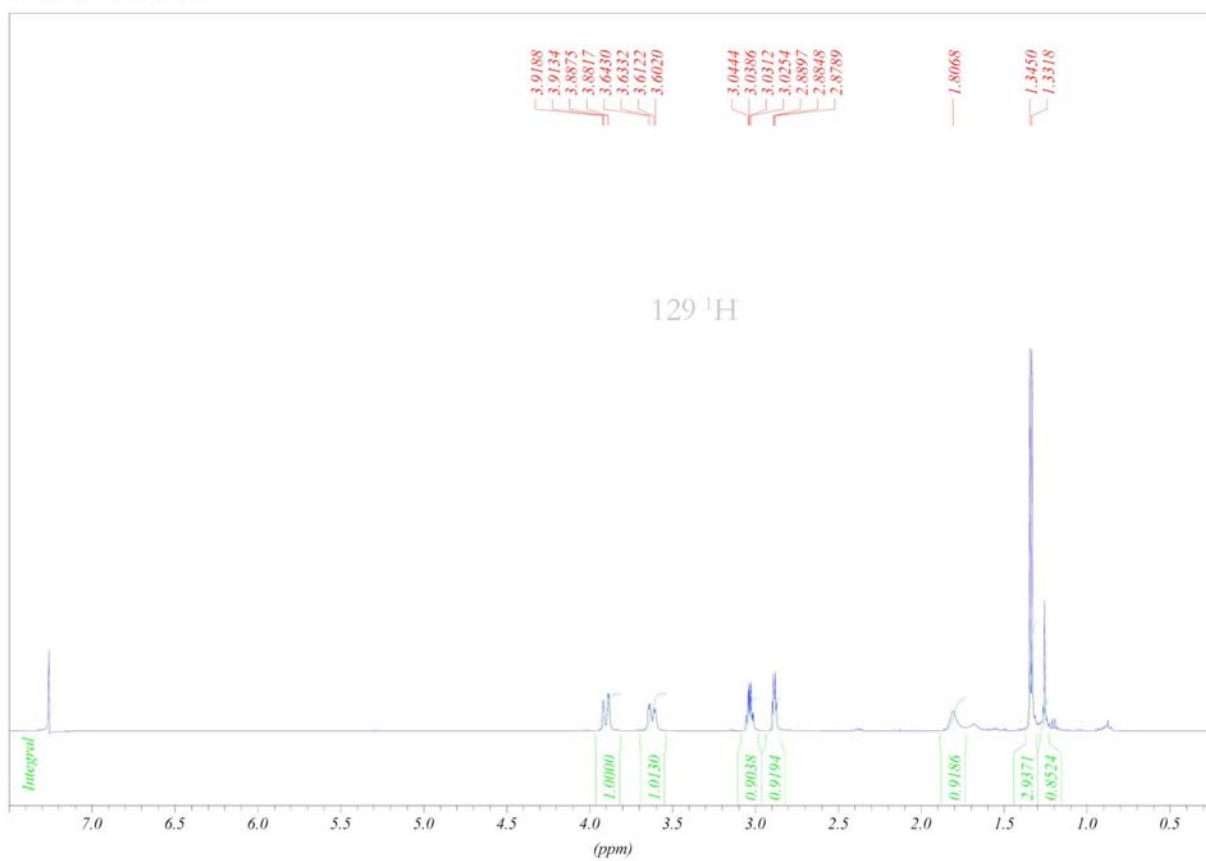
FmocNHD(OtBu)AOMK 400MHz CDCl₃ 271003FmocNHD(OtBu)AOMK 400MHz CDCl₃ 271003

G7 400MHz CDCl₃ 250208G7 400MHz CDCl₃ 250208

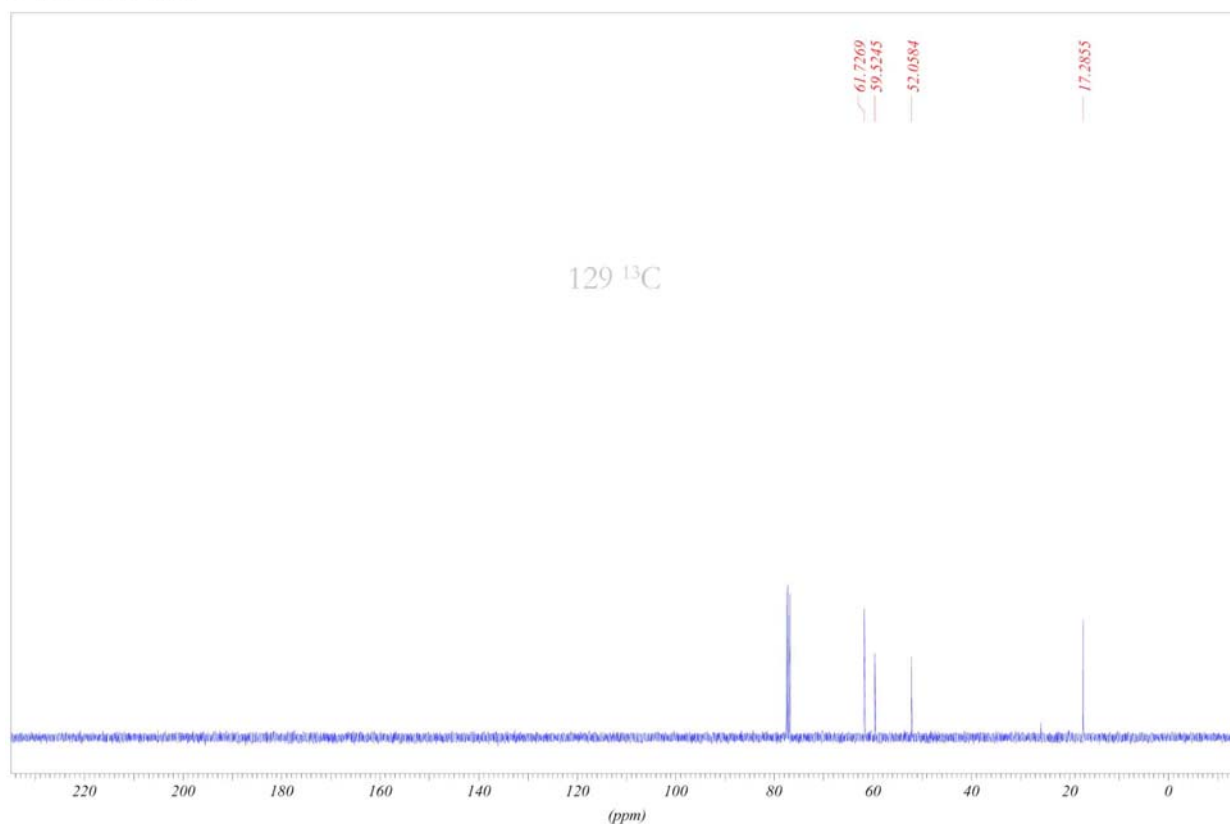
G12 CDCl₃ 400MHz 260109G12 400MHz CDCl₃ 260109

G10 CDCl₃ 400MHz 260109G10 400MHz CDCl₃ 260109

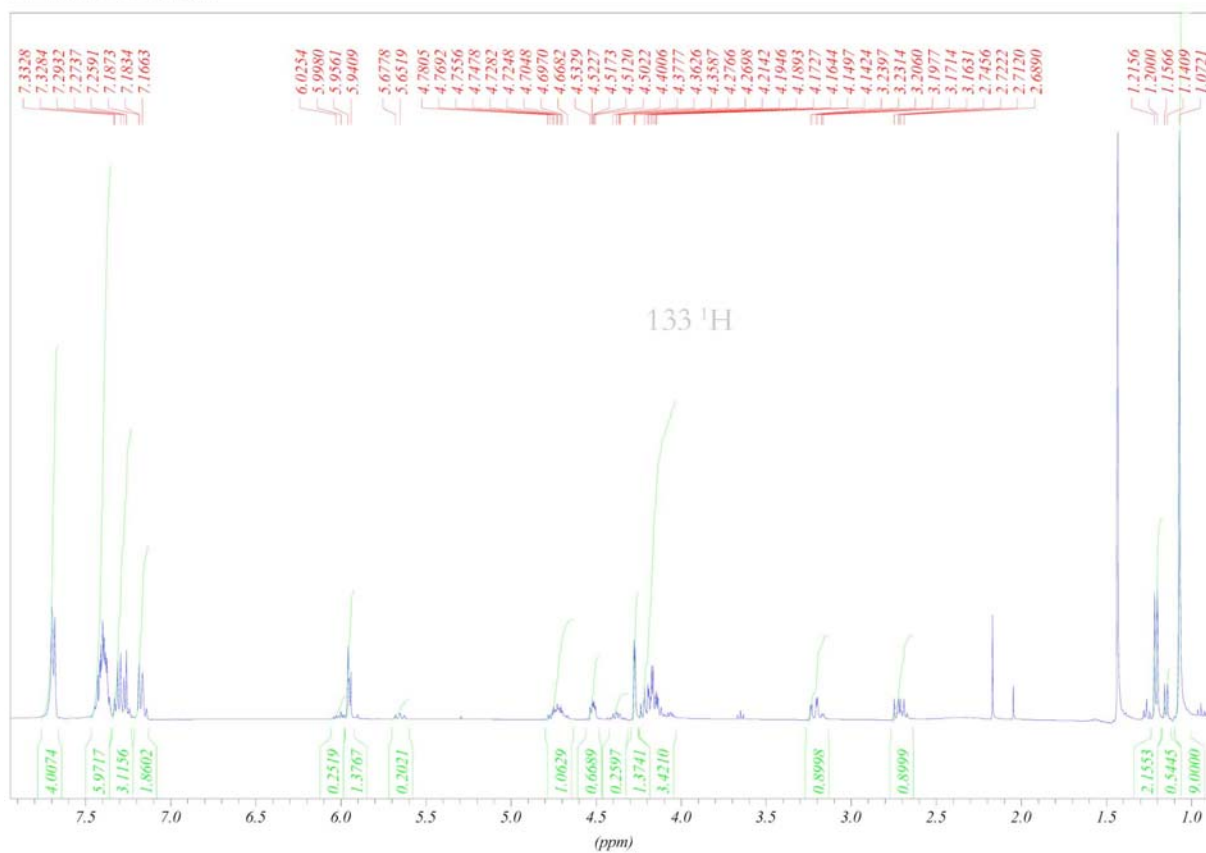
G2 400MHz CDCl3 020409



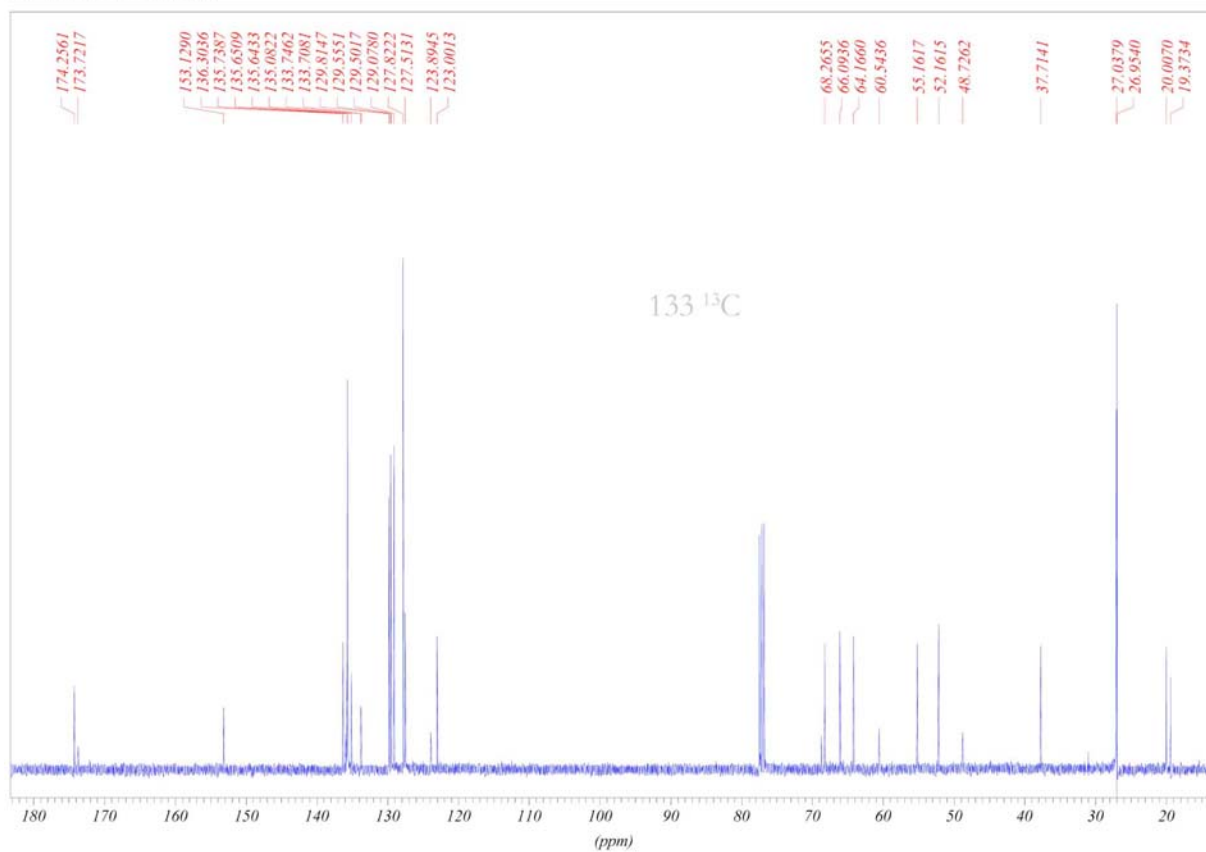
G2 400MHz CDCl3 020409



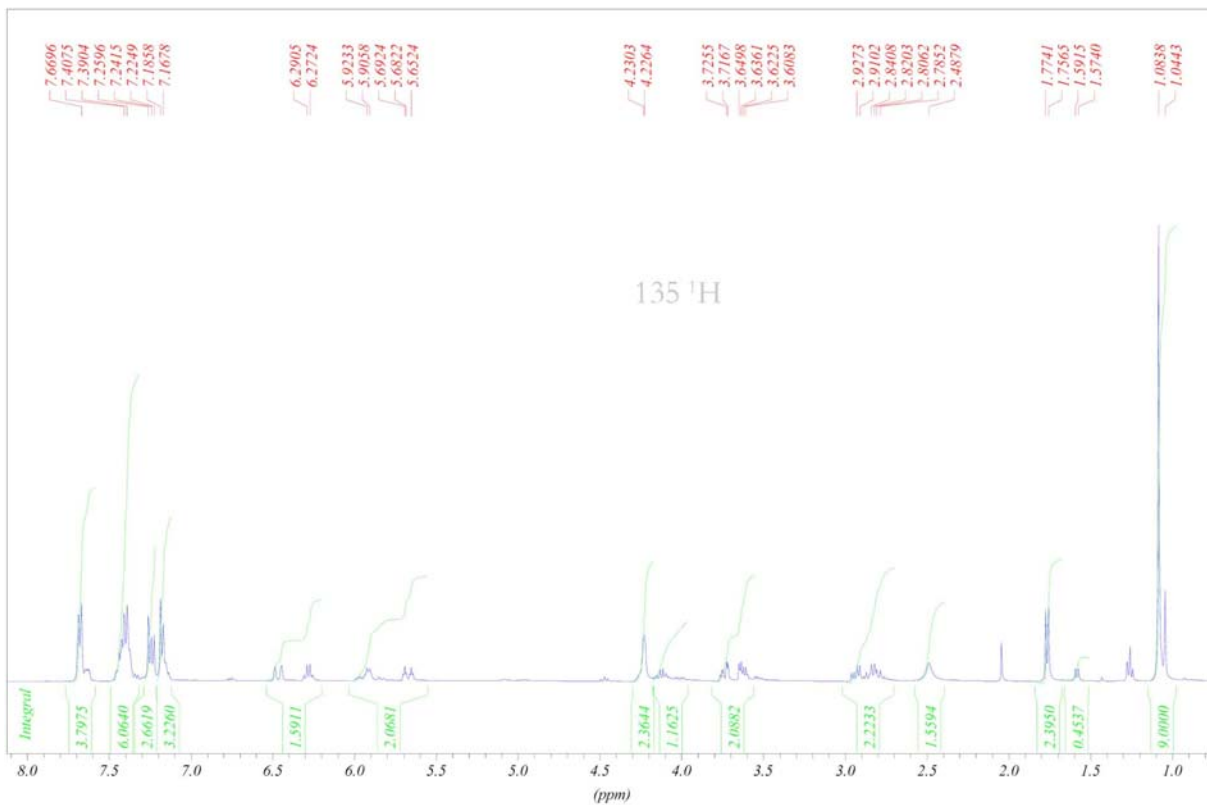
G14 400MHz CDCl3 010608



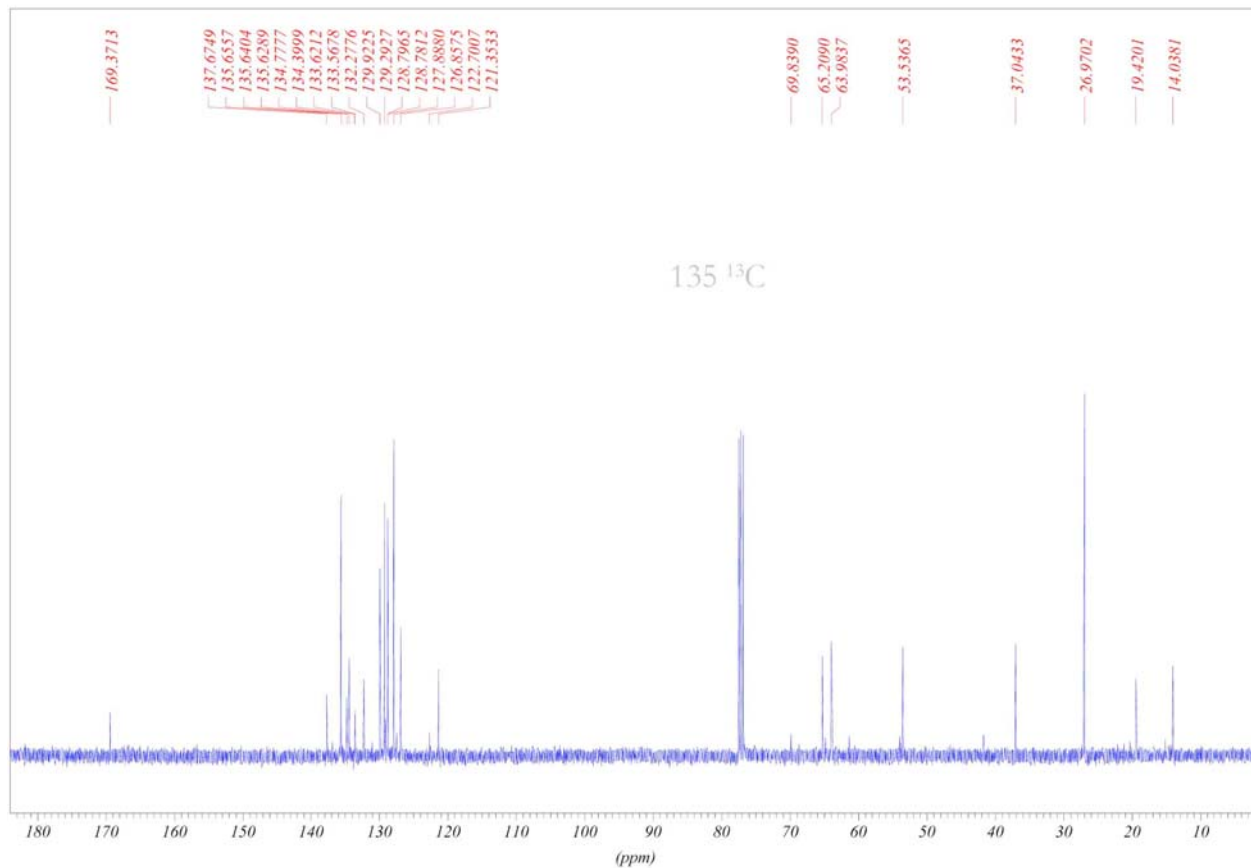
G14 400MHz CDCl3 010608



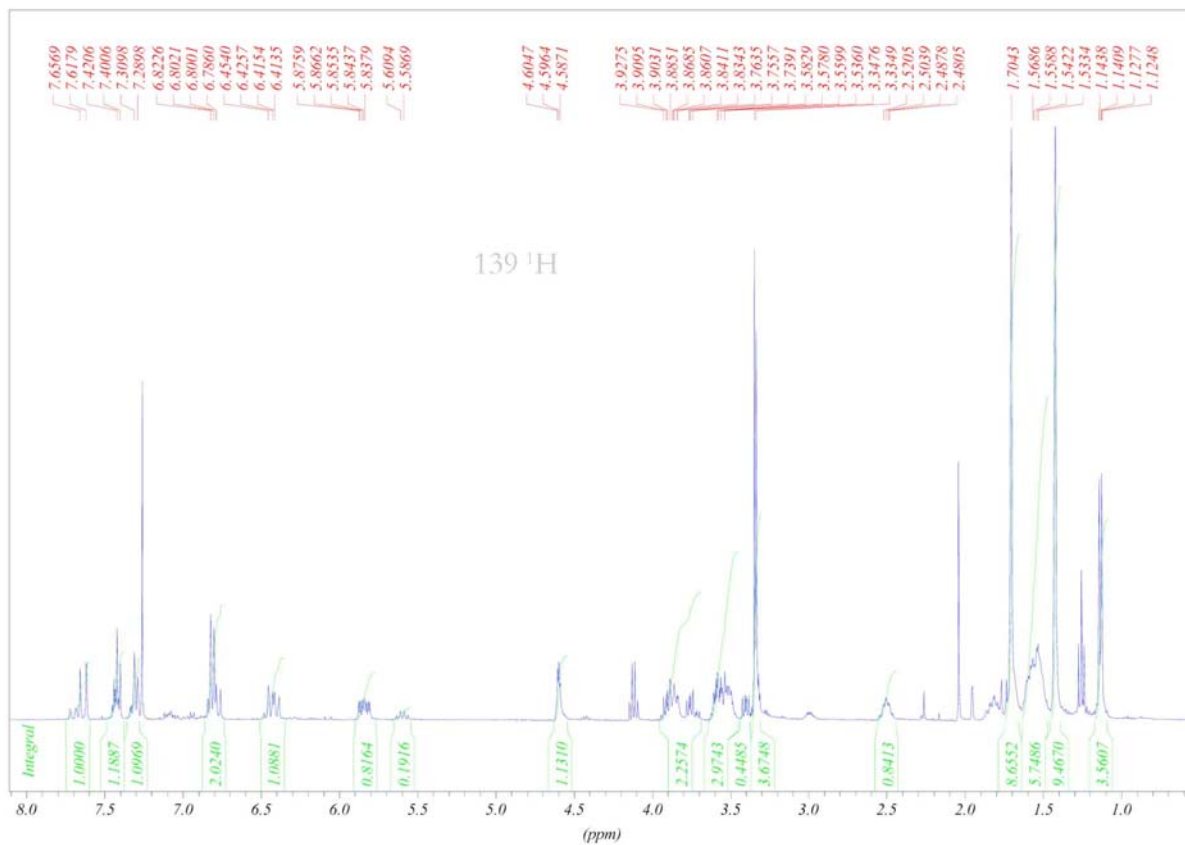
G34 400MHz CDCl3 020608



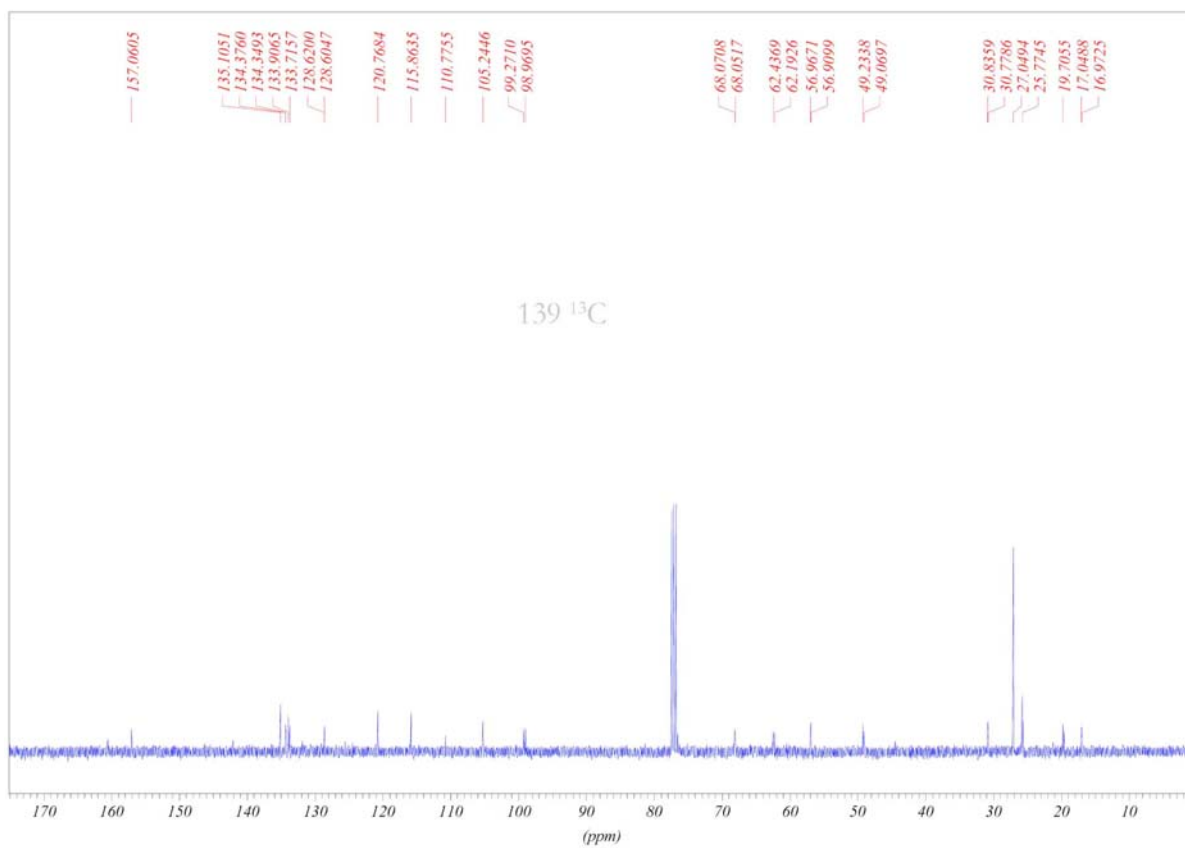
G34 400MHz CDCl3 020608



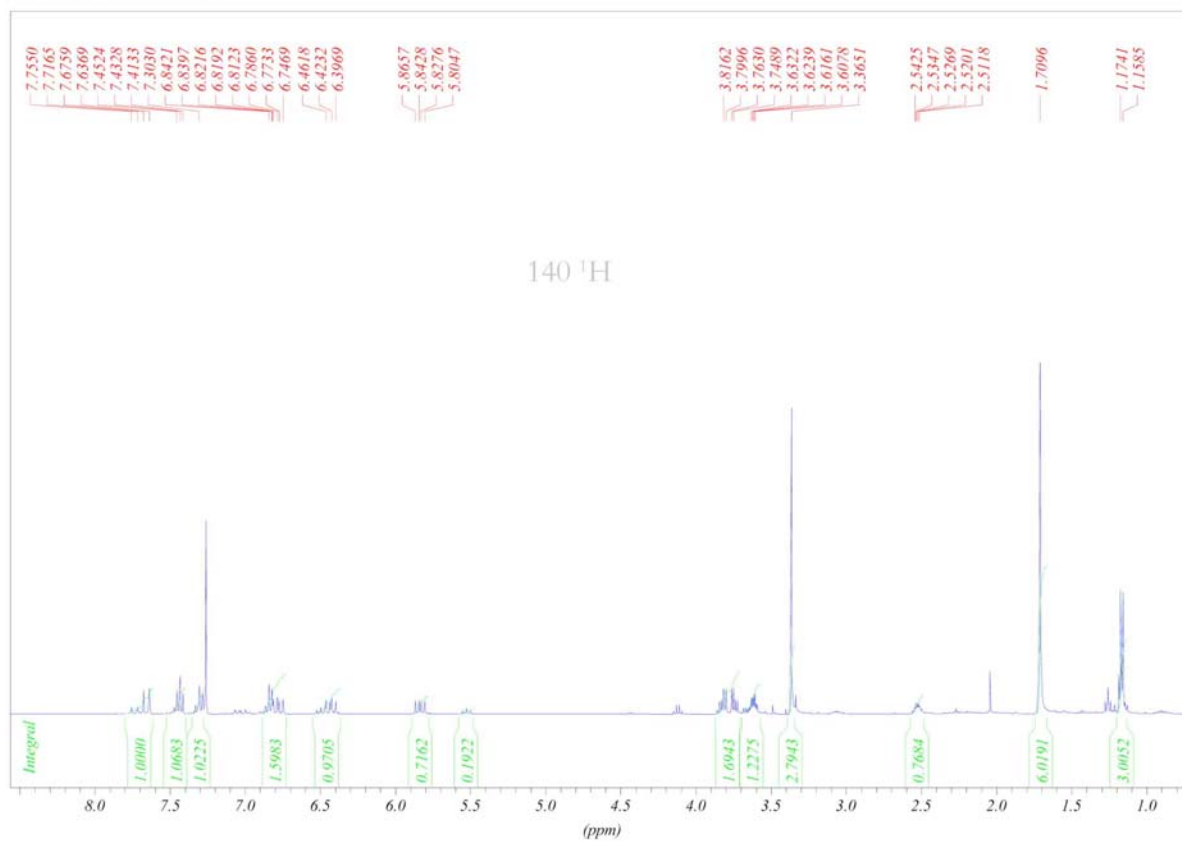
G24 400MHz CDCl3 260208



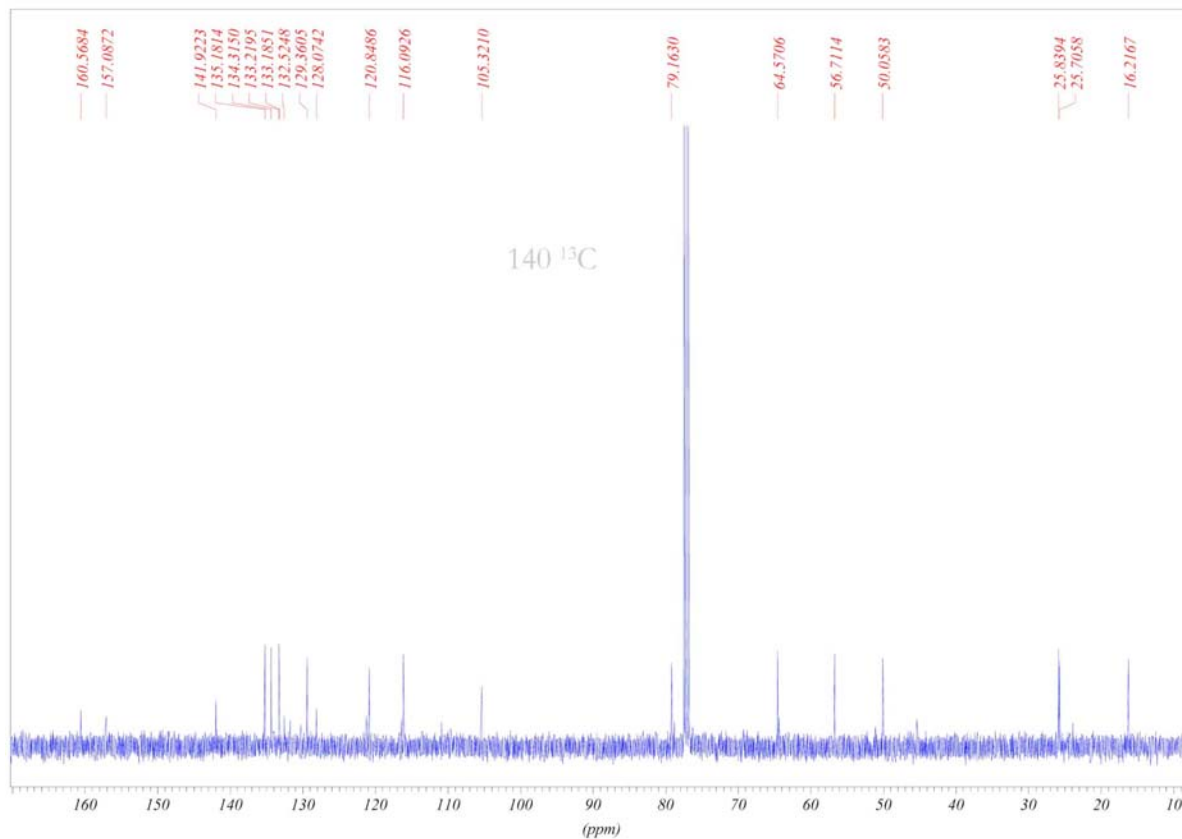
G24 400MHz CDCl3 260208



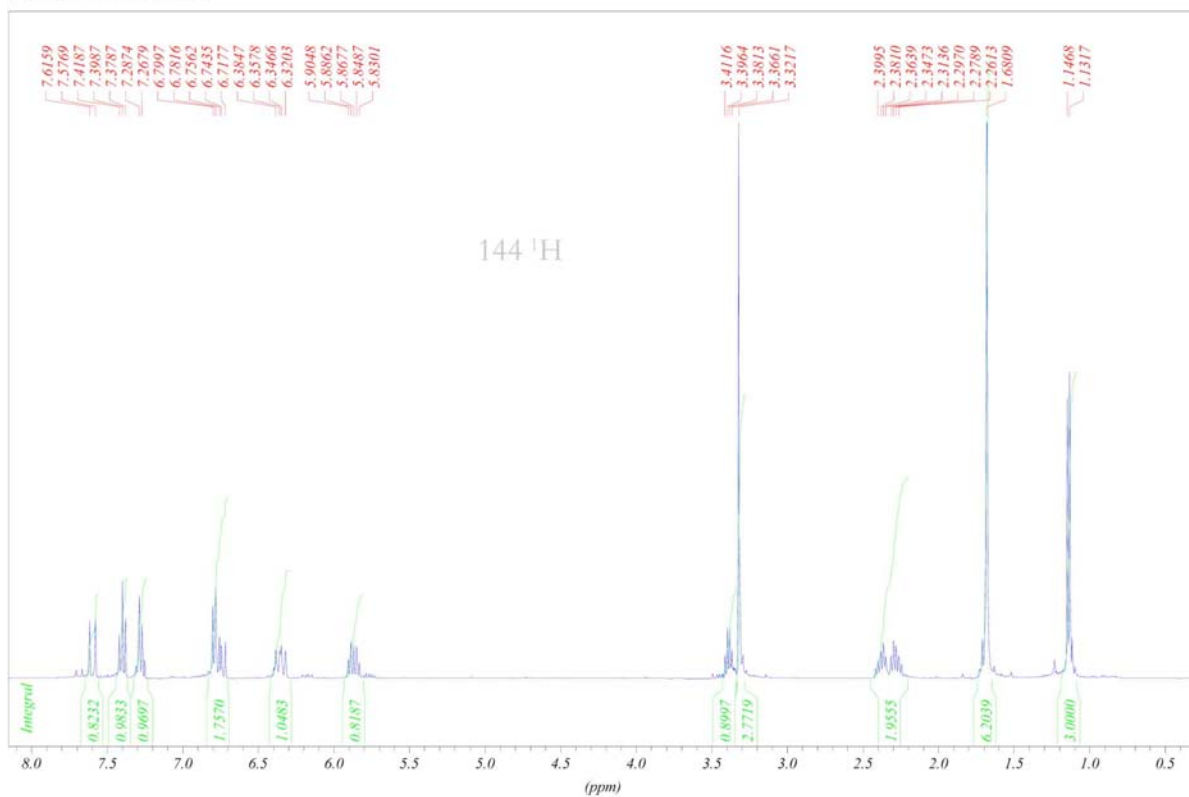
G25 400MHz CDCl3 030308



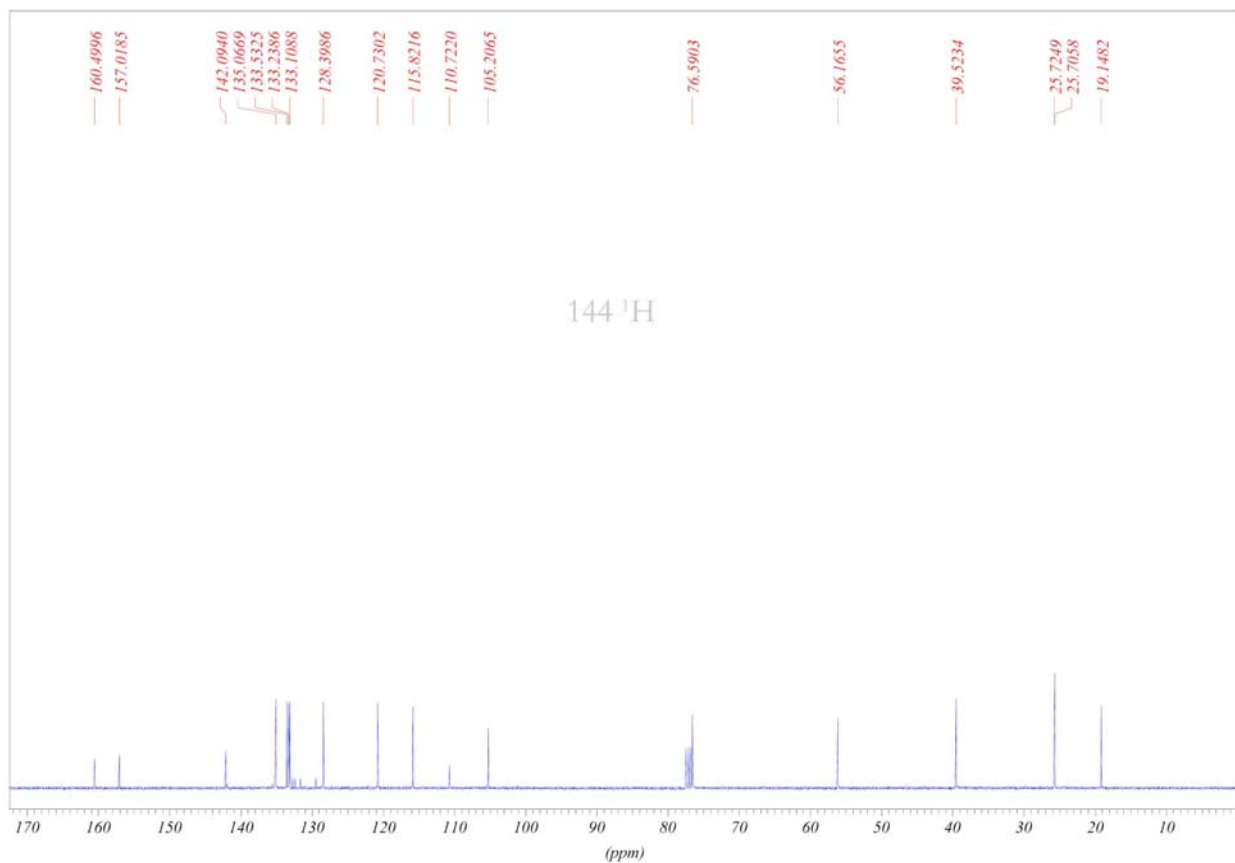
G25 400MHz CDCl3 030308



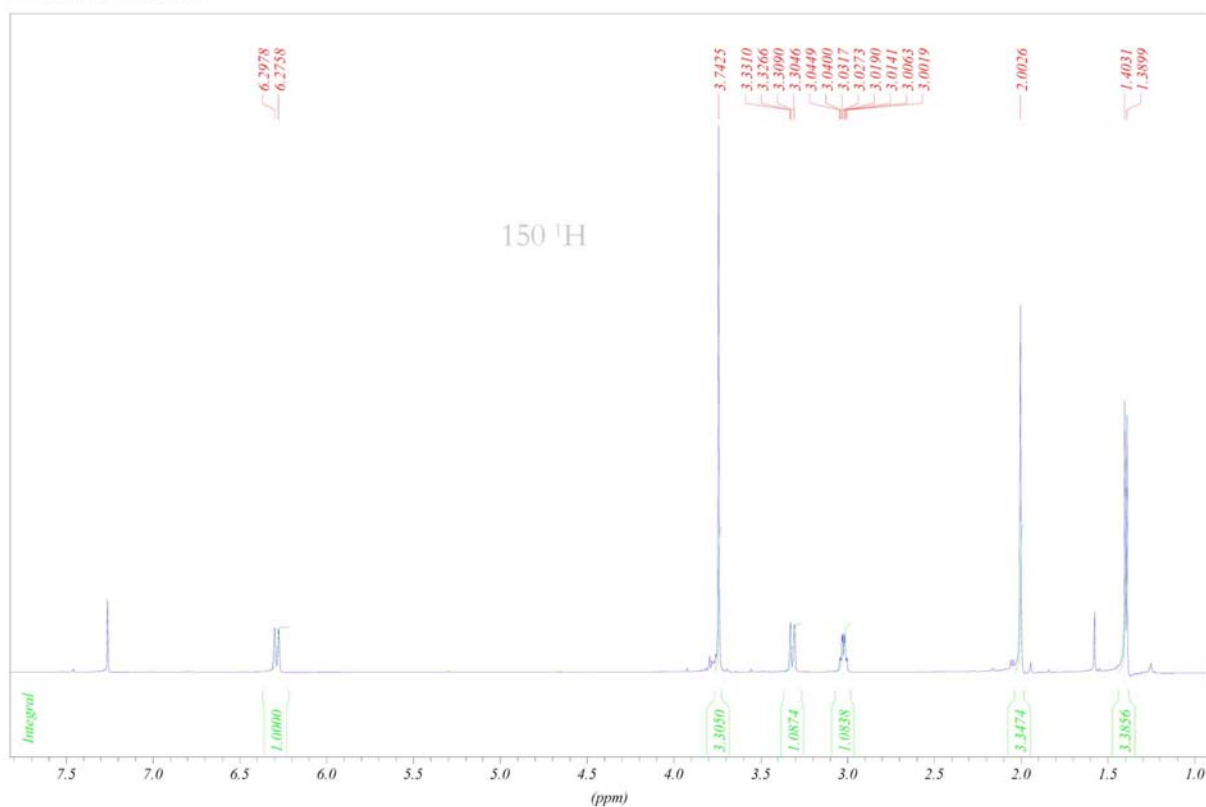
G49 400MHz CDCl3 040309



G49 400MHz CDCl3 040309



G48 400MHz CDCl3 060309



G48 400MHz CDCl3 060309

