Chemical tools for the study of proteolytic activities associated with antigen presentation

Swieten, Paul Franciscus van

CHEMICAL TOOLS FOR THE STUDY OF PROTEOLYTIC ACTIVITIES ASSOCIATED WITH ANTIGEN PRESENTATION

Cover design and printing:
Optima Grafische Communicatie, Rotterdam

CHEMICAL TOOLS FOR THE STUDY OF PROTEOLYTIC ACTIVITIES ASSOCIATED WITH ANTIGEN PRESENTATION

PROEFSCHRIFT

TER VERKRIJGING VAN

DE GRAAD VAN DOCTOR AAN DE UNIVERSITEIT LEIDEN,

OP GEZAG VAN RECTOR MAGNIFICUS DR. D. D. BREIMER,

HOOGLERAAR IN DE FACULTEIT DER WISKUNDE EN

NATUURWETENSCHAPPEN EN DIE DER GENEESKUNDE,

VOLGENS BESLUIT VAN HET COLLEGE VOOR PROMOTIES

TE VERDEDIGEN OP DONDERDAG 18 JANUARI 2007

KLOKKE 16.15 UUR

DOOR

PAUL FRANCISCUS VAN SWIETEN

GEBOREN TE DE LIER IN 1977

Samenstelling promotiecommissie

Promotores	:	Prof. dr. H. S. Overkleeft
		Prof. dr. G. A. van der Marel
Referent	:	Dr. J. W. Drijfhout
Overige leden	:	Prof. dr. J. J. C. Neefjes
		Dr. B. M. Kessler (University of Oxford)
		Dr. H. Ovaa
		Prof. dr. J. Brouwer

Table of contents

List of abbreviations[1]

ABP	activity based probe
Ac	acetyl
Ada	1-adamantyl acetyl
Ahx	6-amino-hexanoic acid
amc	7-amino-4-methyl coumarin
Aoh	7-amino-3-oxaheptanoic acid
aq.	aqueous
Bn	benzyl
Boc	tert-butyl-oxycarbonyl
br s	broad singlet
Cat	cathepsin
cv	column volume
d	doublet
DAPI	4',6-diamidino-2-phenylindole
DCG	Doron C. Greenbaum
dd	doublet of doublets
DIC	diisopropyl carbodiimide
DiPEA	*N*,*N*-diisopropyl ethylamine
DMAP	4-dimethylaminopyridine
DMF	*N*,*N*-dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDC	1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide
EDTA	ethylene diamine tetraacetic acid
ESI-MS	electrospray ionization mass spectrometry
Et	ethyl
eq.	molar equivalent
FALI	fluorophore-assisted laser inactivation
FCS	fetal calf serum
Fmoc	9-fluorenylmethyloxycarbonyl
FRET	fluorescence resonance energy transfer
GlcNAc	*N*-acetyl-D-glucosamine
GFP	green fluorescent protein
GM-CSF	granulocyte macrophage-colony stimulating factor
GST	glutathione S-transferase
hAGT	human *O*(6)-alkylguanine-DNA alkyltransferase
HATU	*N*-[(Dimethylamino)-1*H*-1,2,3-triazolo[4,5-b]pyridine-1-lmethylene]-*N*-methylmethanaminium hexafluorophosphate
HCTU	*N*-[(1*H*-6-Chlorobenzotriazolo-1-yl)(dimethylamino)methylene]-*N*-methylmethanaminium hexafluorphosphate *N*-oxide
HIV	human immunodeficiency virus
HOBt	1-hydroxybenzotriazole
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
HRP	horseradish peroxidase
ICAT	isotope coded affinity tag
JPM	Joseph P. Meara
LC/MS	liquid chromatography/mass spectrometry
LPS	lipopolysaccharide
m	multiplet

HMPB	4-(4-hydroxymethyl-3-methoxy-phenoxy)butyryl
Me	methyl
MHC	major histocompatibility complex
Mtt	4-methyl-triphenylmethyl
Nle, nL	norleucine
NMP	*N*-methyl-2-pyrrolidone
NMR	nuclear magnetic resonance
OVA	ovalbumin
PAGE	polyacrylamide gel electrophoresis
Pbf	2,2,4,6,7-pentamethyldihydrobenzo-furan-5-sulfonyl
PBMC	peripheral blood mononuclear cell
PBS	phosphate buffered saline
Pmc	2,2,5,7,8-pentamethylchroman-6-sulfonyl
ppm	parts per million
PVDF	polyvinyl difluoride
PyBOP	(benzotriazol-1-yloxy)-tripyrrolidino-phosphonium hexafluorophosphate
q	quartet
RNA	ribonucleic acid
RP	reverse phase
RPMI	Roswell Park Memorial Institute
rt	retention time
s	singlet
sat.	saturated
SDS	sodium dodecyl sulfate
SPPS	solid phase peptide synthesis
Strept	streptavidine
Su	succinimidyl
Suc	succinyl
t	triplet
tBu	tert butyl
TES	triethyl silane
Tf	trifluoromethylsulfonyl
TFA	trifluoro acetic acid
THF	tetrahydrofuran
TIS	triisopropyl silane
TLC	thin layer chromatography
TMEDA	*N,N,N',N'*-tetramethylene ethylenediamine
TMS	tetramethyl silane
Tr	triphenylmethyl
tRNA	transfer ribonucleic acid
Ts	4-tolylsulfonyl
UV	ultraviolet

1. The one and three letter codes for the amino acids follow the recommendations of IUPAC. *J. Biol. Chem.* 1968, **243**, 3557-3559 and *J. Biol. Chem.* 1972, **247**, 977-983.

General introduction

This thesis describes the development of strategies to study the proteasome and cysteine proteases of the cathepsin family. Key to all strategies is the use of an electrophilic trap, which allows covalent labeling of the proteolytic activities with a number of desirable properties, depending on the question at hand. The development of such labels, often based on irreversible inhibitors, is not unique to this thesis. Indeed, many research efforts on the development of irreversible protease inhibitors have appeared in the literature over the past decades.[1] These probes are widely regarded to be very useful in medicinal and biological research, both for the purpose of drug development and, perhaps more importantly, for the assessment of protease activities in complex biological samples in what is now called chemical, or functional, proteomics.[2]

An ideal protease label in the field of proteomics should have a number of functionalities embedded in its structure. These include comparable and high binding to all members of a given class of proteases, but no reactivity towards other components in the biological sample. Other useful attributes are cell permeability and the means to track (for instance via a fluorescent tag) and/or retrieve the modified proteases. A further attractive feature would be the ability to monitor the relative abundance or proteolytic activities stemming from different, but closely related biological samples. Obviously, such ideal tools

are not within reach for each class of proteolytic activities present in nature. The first requirement is the ability to covalently and irreversibly modify proteolytic activities. This can be achieved with relative ease in case a side chain heteroatom of the protease(s) at hand participates in proteolysis, exerting a nucleophilic attack on the peptide bond of the substrate. Fortunately, this is the case for many protease families,[1] for instance the threonine secondary alcohol (the proteasome), the serine primary alcohol (serine proteases such as trypsin) and the cysteine thiol (many cathepsins, caspases, ubiquitin isopeptidases). The highly reactive nucleophiles can be used as a site for labeling of the protease via selective alkylation. However, protease families that recruit a water molecule for the initial nucleophilic attack (matrix metalloproteases, aspartic proteases) have proven to be more resistant towards chemical proteomics strategies employing protease inhibitor tools.[3] A broad spectrum irreversible protease inhibitor is an appropriate starting point for designing protease labels. Some restraint is required, though, because covalent attachment of an identification or isolation tag to a broad spectrum irreversible protease inhibitor may abolish cell permeability or may result in a drop in activity.[4]

The aim of this thesis is to address some of the issues raised in the previous paragraph. In case it is not possible to include all desirable properties of broad spectrum activity, cell permeability and identification potential in one chemical probe, an alternative would be to modify proteases in a two-step procedure: first labeling with a cell-permeable probe and second selectively functionalizing the labeled protease with an identification moiety in a cell-free environment. One solution, presented in this thesis for the two-step labeling of the proteasome, relies on the use of the Bertozzi-Staudinger ligation,[5] one of the major recent breakthroughs in the development of selective reactions in complex mixtures of biological origin. A second advance in protease probe design is the inclusion of an isotope encoded spacer, described in this thesis as a set of 'light' (H8) and 'heavy' (D8) probes that inhibit a broad array of cathepsin cysteine proteases. This strategy extends the isotope coded affinity tagging (ICAT) strategy as developed by Aebersold and coworkers,[6] enabling the monitoring of relative levels of protease activity, rather than levels of expression. The ICAT strategy is based on the alkylation of cysteine residues in two samples with either a heavy or light reagent, the mass difference of the products stemming from isotopes incorporated. The underlying idea is that the incorporated stable isotopes neither influence the reactivity of the alkylating agent towards cysteine thiols nor influence its chromatography properties. The two samples are combined and prepared for mass

spectrometric analysis, where fragments containing alkylated cysteine residues serve as mutual internal standard. The relative abundance of proteins in the two samples becomes apparent through comparing the magnitudes of the signal pairs originating from cysteine containing fragments.

Although the probes described in the research chapters in this thesis can be used to study either the proteasome or the cathepsin family of cysteine proteases, the general strategy upon which their design is based should be easy to transfer to probes for studying other classes of biomolecules. As such, the contents of this thesis fall within the broad area of research that can be described loosely as 'Chemical Biology'.[7] Although not new, especially the last decade has witnessed an explosion of research activities that take place at the interface of chemistry and biology. Next to the two examples briefly outlined above (bioorthogonal labeling, ICAT based proteomics), diverse areas of research such as chemical genetics (disabling gene products through inhibition with a chemical compound, rather than silencing the gene),[8] but also engineering the actual genetic code[9] are considered to belong to the Chemical Biology area. A specific area of Chemical Biology research with high relevance for part of the contents of this thesis is highlighted in **Chapter 1** and encompasses the current status in the two-step labeling of selected biomolecules in live cells.

Chapters 2 and 3 describe a new strategy to decorate the proteolytic activities of the 26S proteasome with a biotin moiety in a two-step strategy. The 26S proteasome is the business end of the ubiquitin-proteasome machinery that is responsible for the degradation of the majority of cytosolic and nuclear proteins in eukaryotes.[10] The degradation and turnover of proteins is indispensable for living cells, both to remove malfunctioning or redundant proteins, and to allow reuse of amino acids in the synthesis of other proteins. Attachment of a polyubiquitin chain (ubiquitin is a 76 residue signal protein) to lysine residues in a given protein marks it for degradation. The proteasome processes ubiquitinylated polypeptides to produce oligopeptides of a length varying from three to about 20 amino acid residues. The bulk of these are then further processed through the action of downstream aminopeptidases. A small portion, however, will (after potential further trimming by aminopeptidases) escape to the endoplasmic reticulum and complex to major histocompatibility complex (MHC) class I assemblies for presentation to the immune system.[11] As such, the proteasome is indispensable to mammalian immunity in monitoring

the cytosolic and nuclear protein content which may, next to self proteins, also encompass viral encoded proteins.

The 26S proteasome is assembled from a 20S core proteolytic particle to which two 19S cap protein assemblies are associated. Substrate recognition takes place at the 19S caps, where also protein unfolding and ubiquitin removal activities are located. The actual protein degradation takes place within the 20S core particle, which consists of 28 protein subunits. Of these, 14 α-subunits make up the outer parts as two rings of 7 subunits each, and 14 β-subunits make up the inner parts, again as two rings of 7 subunits. The 7 α-subunits in each ring, although highly homologues, have an unique sequence and the same holds true for the 7 individual β-subunits in the inner rings. Three of the β-subunits in higher organisms possess catalytic activity, each with a distinct (although as yet poorly understood) substrate preference. To complicate matters further, higher organisms may express a second set of proteasome proteolytic activities, which also assemble into a 20S particle. This particle, induced by specific cytokines (for instance interferon gamma) that are the result of initial immune response processes, is therefore called the immunoproteasome,[12] and it also harbors three distinct catalytic activities. In **Chapter 2** a two step labeling strategy is described that enables the visualization of the combined six proteolytic activities of both the constitutively expressed proteasome and the immunoproteasome with almost equal efficiency. The strategy hinges on the use of two synthetic tools, one being a peptide vinyl sulfone broad spectrum proteasome inhibitor equipped with an azide, and the other a biotinylated phosphine. After addition of the vinyl sulfone compound to growing cells, labeling of all proteasome catalytic subunits with azido groups occurs. Cell lysis is followed by denaturation of the protein content. In the next step the azide functionalities are addressed through Staudinger ligation, resulting in the attachment of a biotin to each catalytic subunit. These are now visualized through SDS-PAGE gel electrophoresis followed by streptavidin blotting. **Chapter 3** describes the application of this two-step labeling strategy in the identification of a peptide vinyl sulfone based inhibitor that specifically inhibits one of the three proteasome proteolytic activities, namely the one with preference to cleave after acidic residues (β1 or the caspase-like activity). Further, this inhibitor was found to specifically inhibit the corresponding immunoproteasome subunit (β1i), making this compound potentially useful in the study of the involvement of the immunoproteasome in MHC class I antigen presentation.

Chapters 4, 5 and 6 describe the development of tools to modify and analyse lysosomal cysteine proteases of the cathepsin class.[13] Lysosomal proteases, of which the cathepsins constitute the major class, are responsible for the degradation of proteins that are collected into the lysosomal compartment through phagocytosis, pinocytosis or autophagy. Proteins targeted for lysosomal degradation can be either cell surface membrane associated proteins or proteins from the extracellular matrix, and may be either self proteins or proteins originating from host pathogens, typically bacteria. As such, the body of lysosomal proteases complement the ubiquitin-proteasome pathway, in that the latter degrades most intracellular proteins and the former takes care of the extracellular protein content. Another parallel is their role in the adaptive immune system: a small portion of the oligopeptides resulting from lysosomal proteolysis escape further degradation and are presented on the cell surface by MHC class II molecules.[14] In this way, also the protein content of the extracellular matrix, including that arising from bacterial infection, is monitored by the immune system. This thesis describes the development of a number of advanced chemical probes for the study of cathepsin activity in a number of different settings. All probes are based on the leucine epoxysuccinate motif as present in the parent compound known in the literature as DCG-04,[15] which was developed by Bogyo and coworkers as a broad-spectrum cysteine protease label. **Chapter 4** presents the development of a set of isotope-encoded activity based probes for the assessment of the relative abundance of cysteine proteases. The probes are based on the aforementioned DCG-04, but contain either eight protons or eight deuterium nuclei. The first inroads towards the use of these compounds in biological studies and the difficulties in these are discussed. In **Chapter 5** a strategy is presented through which the progress of an extracellular protein through the endocytic pathway and the cysteine proteases it encounters along that way can be monitored.[16] A broad-spectrum cysteine protease label equipped with a biotin moiety is linked to ovalbumin as a model protein through a photocleavable linker. Initial studies in which the conjugate is incubated with recombinant cathepsin S followed by photocleavage are presented. **Chapter 6** makes use of a related strategy. A cysteine protease probe is linked to nona-arginine, a peptide known for its ability to deliver cargo to the cytosol.[17] Endocytosis studies demonstrate that this conjugate labels at least some of the pool of cathepsin cysteine proteases, thereby indicating that cell penetrating peptides of this type do make use of endocytic processes in reaching the cytoplasm. Finally, some concluding remarks and future prospects are presented.

References

1 J. C. Powers, J. L. Asgian, O. D. Ekici and K. E. James, *Chem. Rev.* 2002, **102**, 4639-4750.

2 D. A. Jeffery and M. Bogyo, *Curr. Op. Chem. Biol.*, 2003, **14**, 87-95.

3 A. Saghatelian, N. Jessani, A. Joseph, M. Humphrey and B. F. Cravatt, *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 10000-10005.

4 B. M. Kessler, D. Tortorella, M. Altun, A. F. Kisselev, E. Fiebinger, B. G. Hekking, H. L. Ploegh and H. S. Overkleeft, *Chem. Biol.*, 2001, **8**, 913-929.

5 M. Kohn and R. Breinbauer, *Angew. Chem., Int. Ed.*, 2004, **43**, 2-12.

6 S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb and R. Aebersold, *Nat. Biotechnol.*, 1999, **17**, 994-999.

7 A perspective on the field of chemical biology is given: S. L. Schreiber, *Nat. Chem. Biol.*, 2005, **1**, 64-66.

8 D. R. Spring, *Chem. Soc. Rev.*, 2005, **34**, 472-482.

9 L. Wang and P. G. Schultz, *Angew. Chem., Int. Ed.*, 2005, **44**, 34-66.

10 K. L. Rock and A. L. Goldberg, *Annu. Rev. Immunol.*, 1999, **17**, 739-797.

11 J. W. Yewdell, E. Reits and J. Neefjes, *Nat. Rev. Immunol.*, 2003, **3**, 952-961.

12 B. J. van den Eynde and S. Morel, *Curr. Opin. Immunol.*, 2001, **13**, 147-153.

13 K. Honey and A. Y. Rudensky, *Nat. Rev. Immunol.*, 2003, **3**, 472-482.

14 C. Watts, *Nat. Immunol.*, 2004, **5**, 685-692.

15 D. Greenbaum, K. F. Medzihradszky, A, Burlingame and M. Bogyo, *Chem. Biol.*, 2000, **7**, 569-581.

16 A.-M. Lennon-Duménil, A. H. Bakker, R. Maehr, E. Fiebinger, H. S. Overkleeft, M. Rosemblatt, H. L. Ploegh and C. Lagaudrière-Gesbert, *J. Exp. Med.*, 2002, **196**, 529-539.

17 S. Futaki, *Int. J. Pharm.*, 2002, **245**, 1-7.

Chapter 1

Bioorthogonal chemistry in living cells: novel strategies for labeling biomolecules[1]

Introduction

Research in contemporary life sciences is increasingly directed towards the detailed analysis of biological processes at the molecular level. In order to understand and intervene in processes at the DNA/RNA level, protein-protein interactions and enzymatic action - systems which all involve large biopolymers - a detailed insight into structural and chemical mode of action is often a prerequisite. At the same time, scientists endeavour to generate such detailed molecular information from macroscopic systems, that is, in whole cells and organisms.

The ability to selectively tag a biomolecule in living cells or organisms allows researchers to a) monitor its involvement in a given biological event which progresses in its natural environment, and/or to b) retrieve and analyse in detail the tagged biomolecule or its metabolites. The biomolecular toolbox provides a set of useful tags which can be attached to proteins of interest through standard genetic engineering techniques. Widely used tags include short oligopeptide stretches that can be selectively isolated through

immunoprecipitation (such as the FLAG tag) or affinity chromatography (for instance the nickel-binding His tag). Of particular interest is the family of fluorescent proteins, the most prominent member of which is green fluorescent protein (GFP).[2]

The availability of fluorescent proteins emitting at different wavelengths has resulted in the development of fluorescence resonance energy transfer (FRET)[3] strategies for the visualisation of the spatial arrangement and interaction dynamics of proteins. In the last decade, a wealth of information on many biological systems has emerged from studies based on the use of fluorescent fusion proteins. An obvious extension of this technique would be to develop systems that express tagged non-proteinogenic entities, or those that report on enzymatic activities. The realisation that such can not be achieved by standard molecular biology techniques alone has led to a shift of attention towards organic chemistry-based approaches. In principle, the use of synthetic tags is not limited to a particular class of biomolecules and, moreover, the variety in the nature of the tag is only limited by the imagination of the researcher. In addition, proper design of the synthetic tag enables not only monitoring of the presence or absence of the tagged biomolecule in a specific system (as with GFP approaches), but will also provide information about enzyme activity or about the cellular fate of a metabolite of any given class of biomolecules.

Most biochemical studies are executed in cell lysates, since these are generally easier to exploit experimentally as compared to whole cells or animal models. In line with this, chemical tagging of a biomolecule in a cell lysate obviates the need for the tag to be cell permeable. Reporter entities, such as biotin, fluorescent labels and radiolabels, may therefore be incorporated into the molecular probes liberally, as long as they do not interfere with the process of interest. However, because of intrinsic disadvantages of working with cell lysates (dilution factor, pH change, loss of subcellular compartmentalisation), live cell systems are preferred where possible.

Tagging biomolecules in living cells does, however, impose some important restrictions on the design of the chemical tags. For instance, the cell-impermeability or the reduced profiling potency[4] of many biotinylated constructs essentially precludes the direct incorporation of biotin in a molecular probe. To a lesser extent, the same also holds true for many radiolabels and fluorescent tags. It is therefore not surprising that in recent years a variety of two-step labeling strategies have been developed.[5] These strategies are based on a) introduction of a bioorthogonal[6] attachment site to a biomolecule in a living cell system,

and b) addition of the reporter moiety, through a chemoselective ligation, to the attachment site at a desirable stage in the experiment.

Three different approaches for introducing the bioorthogonal entity into the biomolecule(s) of choice and attaining the chemoselective ligation with a tagging or crosslinking agent can be discerned (see Figure 1):

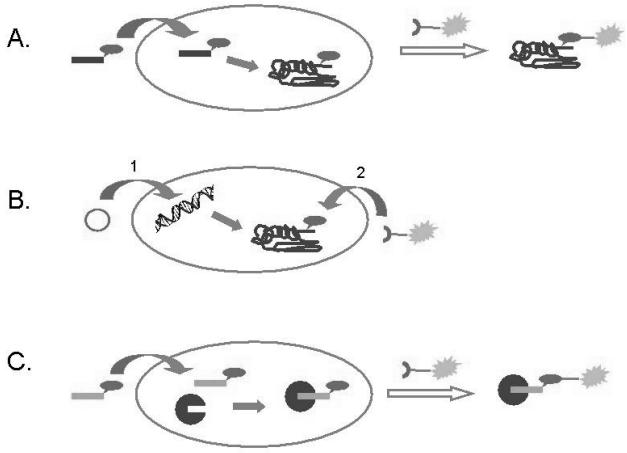


Figure 1. Three different strategies for the labeling of biomolecules in living cells: A) introduction of a chemically modified metabolic precursor, then addressing the modification; B) creation of fusion proteins with a bioorthogonal attachment site, followed by ligation *in vivo*; C) *in vivo* tagging of enzymes with activity-based probes, then ligation with a reporter molecule.

A. Treatment of tissue with a chemically modified metabolic precursor will lead to incorporation of a bioorthogonal attachment site in the target biomolecules. Subsequently, the attachment site is addressed in a chemoselective ligation. Examples of this strategy include the introduction of bioorthogonally functionalised amino acids into the cell's protein synthesis machinery[7] and the introduction of other modified precursors of biomolecules into post-translational modifications and oligosaccharide clusters.[8]

B. Creation of mutant cell lines expressing fusion proteins with attachment functionalities that specifically react with a complementary synthetic probe of choice.[9]

C. Tagging of enzymes with small functionalised active site directed probes in living cells, then addressing the functionality in a bioorthogonal chemical ligation step.[10]

A. Use of chemically modified metabolic precursors

Methionine tRNA-synthetase is able to accept unnatural methionine isosteres as substrates.[11] Using this feature, Van Hest *et al.* succeeded in introducing amino acids with unsaturated side chains (for instance, 2-amino-5-hexynoic acid) into proteins.[12] In a joint effort of the Tirrell and Bertozzi groups, this approach was used for the decoration of *E. coli* proteins with azide residues.[13] Addition of azidohomoalanine (**1**, see Figure 2) to a

methionine-depleted bacterial culture resulted in the incorporation of azide moieties instead of methionine residues. After cell lysis, the azide functionalities were addressed chemoselectively by a modified Staudinger reaction (also referred to as the Staudinger-Bertozzi ligation, see Scheme 1), thereby attaching the antigenic FLAG-peptide to the modified proteins.

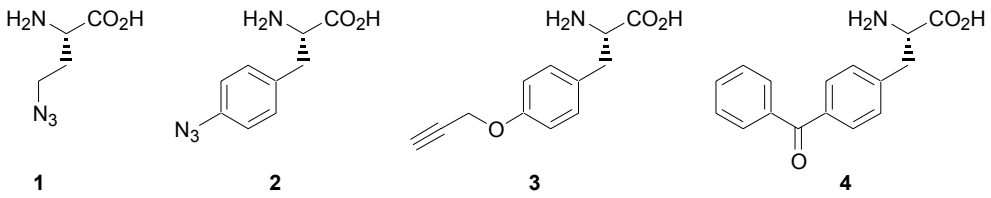
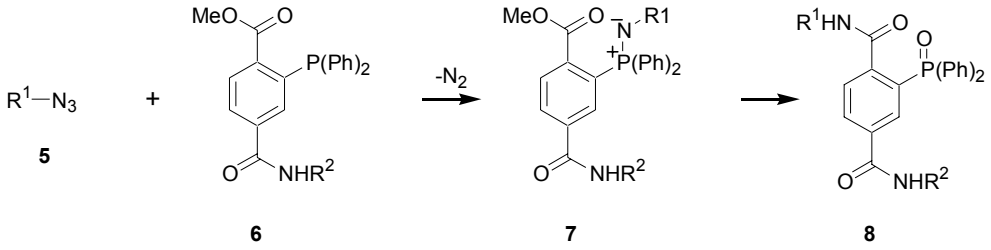
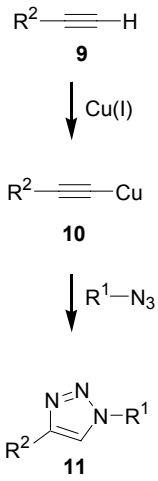


Figure 2. Some bioorthogonally functionalised amino acids that have been built into proteins *via* the protein synthesis machinery in living cells.



Scheme 1. Staudinger-Bertozzi ligation to connect R^1 and R^2.

Using the same technique for the construction of modified proteins, Tirrell and co-workers reported that azide-decorated cell surface proteins can be labeled using a copper catalyzed Huisgen-type [3+2] cycloaddition (see Scheme 2).[14] This reaction, that can be executed under aqueous conditions, and that features two reactive components that are largely inert in physiological environments, was developed independently by the groups of Meldal[15] and Sharpless.[16] A biotin propargyl amide was used to address the cell surface azides by use of this cycloaddition reaction. After staining with fluorescent avidin, the biotin-modified cells were easily distinguished from the non-modified cells.



Scheme 2. The copper-catalysed Huisgen-type cycloaddition.

Schultz and co-workers have embarked on the development of genetically engineered micro-organisms that encode for non-natural amino acids.[17] The end goal of this research program is to generate

altered micro-organisms that endogenously synthesise a 21st amino acid with any desirable functionality. This amino acid residue is recognised by a novel tRNA synthetase that subsequently transfers it to a tRNA equipped with a unique anticodon, allowing site-specific incorporation of the amino acid in selected proteins. The current state of the art dictates that bioorthogonally functionalised amino acids for incorporation still need to be added to the growth medium of altered organisms.[18] For example, proteins with either *p*-azidophenylalanine (**2**, Figure 2) or *O*-propargyltyrosine (**3**)[19] as unnatural amino acid, depending on the composition of the growth medium, were expressed in genetically modified yeast. After affinity purification of the target proteins, both modifications were addressed by their fluorescent counterparts (that is, azide- or alkyne-functionalised dansyl or fluorescein derivatives) in a Huisgen-type cycloaddition. This procedure allowed for in-gel visualisation of the modified target proteins. The authors are currently investigating an *in vivo* labeling strategy based on the same principles.

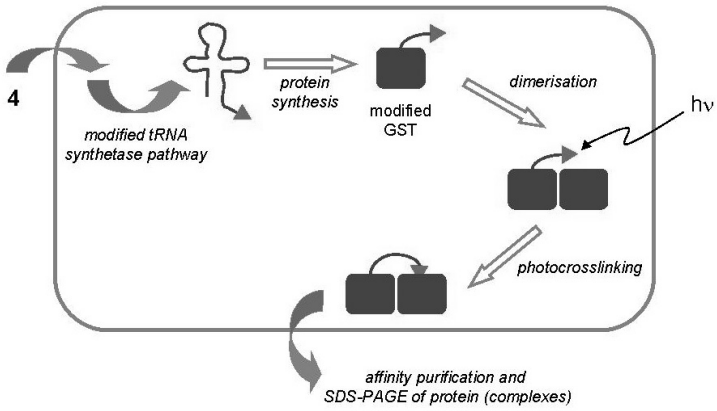


Figure 3. Introduction of a photocrosslinker amino acid (**4**) specifically into GST of *E. coli*, followed by irradiation with UV light leads to crosslinked GST molecules, thereby confirming that GST exists as dimers *in vivo*.

A major breakthrough in protein engineering is the development of techniques that allow the incorporation, at a predesigned site, of non-natural amino acids that are highly reactive under photolytic conditions.[20] After exposure to UV light, the resulting intermediates, normally being carbenes or nitrenes, will react with surrounding functionalities, including amide linkages. Crosslinking may occur with surrounding proteins, thereby providing information on protein-protein interactions. Until recently, however, the site-specific incorporation of desired amino acids could only be achieved in *in vitro* translation systems and with low efficiency. Further, the application of the modified proteins obtained through *in vitro* translation met with considerable technical difficulties. The *in vivo* translation system developed by the group of Schultz opens the way to many

exciting applications, in which protein-protein interactions can be monitored in the proper environment of the living cell.

A relevant example of this strategy entails the following set of experiments.[21] Azidophenylalanine was incorporated at the binding surface of glutathione S-transferase (GST) in the appropriate *E. coli* mutant, and the protein was purified through its fused His-tag. Irradiation with UV-light effected crosslinking to another GST molecule. This finding indicates that GST occurs in dimers *in vitro*. In addition, the crosslinking proved to be possible in living cells (see Figure 3),[22] using a different photocrosslinking amino acid, *p*-benzoyl-L-phenylalanine (**4**). When incorporated at the same site of GST as azidophenylalanine, similar results were obtained as in the *in vitro* crosslinking experiment. These results clearly demonstrate the value of the *in vivo* translation technique for probing protein-protein interactions in living cells.

Transforming micro-organisms such that they incorporate non-natural amino acids in selected proteins requires considerable efforts in altering their protein synthesis machinery. An alternative strategy for the selective incorporation of bioorthogonal attachment sites is represented by adaptation of specific protein post-translational modification events. In such a strategy, a key intermediate in a given post-translational modification pathway is replaced by a synthetic equivalent containing a ligation handle. Successful incorporation now rests on the ability of the involved enzymes to recognise the chemically altered metabolic precursor and process it like its natural counterpart, without interfering with the incorporated ligation functionality.

Seminal work by Bertozzi and co-workers on the latent labeling of cell surface glycoproteins comprises one of the hallmarks of contemporary chemical biology research.[23] It was demonstrated that a synthetic *N*-acetylmannosamine derivative equipped with an azide functionality is tolerated by both the sialic acid biosynthetic pathway and the protein glycosylation machinery, resulting in the decoration of the cell surface glycoproteins with azide-containing sialic acid residues. In short, Jurkat cells were grown on medium containing tetra-*O*-acetyl-*N*-(azidoacetyl)mannosamine (**12**, Figure 4) for three days. After washing, the cells were treated with phosphine-biotin reagent **14** and subsequently stained with fluorescent avidin. In this way, cells fed with **12** could be selectively visualised. This result highlights the metabolic stability of the azide group in a living cell, and thereby its suitability for use in bioorthogonal chemistry. Recently, the scope of the Staudinger ligation

was extended to performing the ligation in living animals. Mice were treated with **12** for seven days, and the cell surface sialic acid residues were addressed by injection of an analogue of **14** in which the biotin had been replaced with an FLAG-epitope.[24]

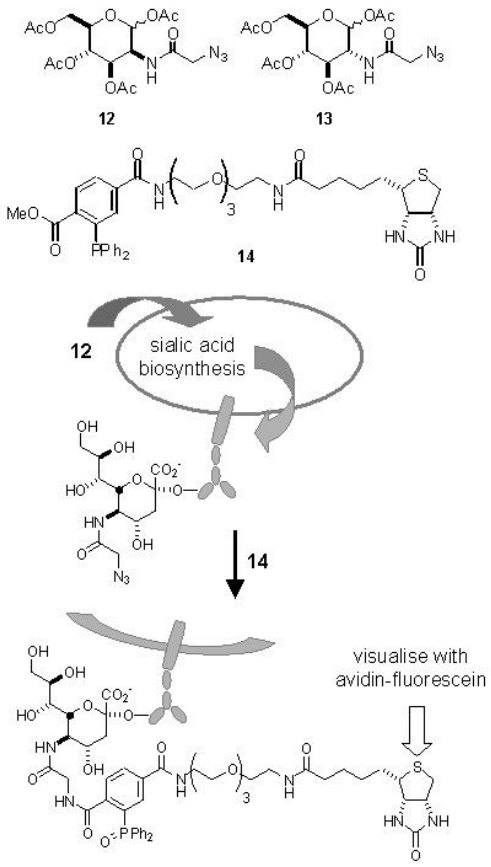


Figure 4. The structures of azide-modified carbohydrates and the use of modified ManNAc **12** in cell surface labeling.

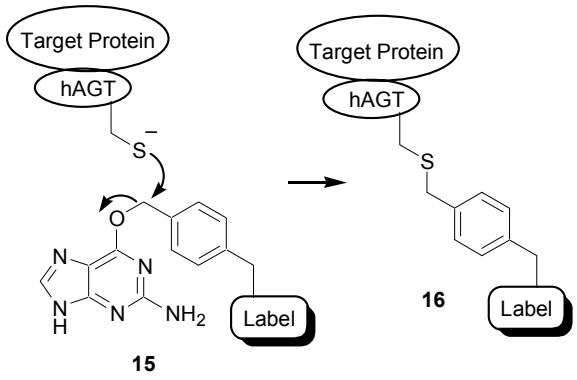
The modification of specific serine and threonine residues in a wide array of cytosolic and nuclear proteins with a single *N*-acetylglucosamine residue (Ser/Thr *O*-GlcNAc modification) has attracted considerable interest in recent years. In many cases, the modification appears to be reversible and of a highly dynamic nature. Further, the finding that this modification often occurs at sites also prone to phosphorylation has led to speculations that protein *O*-GlcNAcylation, as is the case with protein phosphorylation, is involved in the control of many biological events. These considerations prompted the Bertozzi group[25] to establish whether protein *O*-GlcNAcylation could be addressed through selective chemical ligation in a similar fashion as outlined above for sialic acid containing glycoproteins. A cell line having a defect in its *de novo* hexosamine synthesis pathway was cultured in the presence of azide-modified GlcNAc (compound **13**). After cell lysis and isolation of nuclear material, the modified GlcNAc moieties were equipped with an affinity label using the Staudinger-Bertozzi ligation. As a result, the authors detected a large number of nuclear proteins that bear the azido-GlcNAc modification.

B. Tagging of fusion proteins with synthetic probes

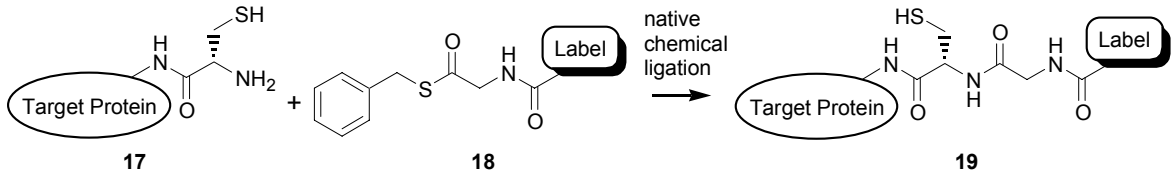
The ability to label a protein of choice site-selectively, in order to track its cellular localisation and fate, has been the objective of many research efforts in recent decades. Fusion proteins containing GFP or its analogues remain the tools of choice for this purpose.

However, the fact that these proteins are bulky and thereby potentially perturbative, has spurred the search for alternatives. Some examples of recent results are described in the following section.

Johnsson and co-workers have developed a labeling strategy that exploits the activity of the human DNA repair protein 6-*O*-alkylguanine-DNA-alkyltransferase (hAGT).[26] The biological function of hAGT resides in its ability to recognise 6-*O*-alkylated guanine and transfer the alkyl group to one of its cysteine thiols, thereby restoring the structural integrity of the DNA. It appeared that hAGT possesses broad substrate specificity with respect to the alkyl moiety and, especially, that functionalised benzyl moieties are well tolerated. This observation paved the way for the following strategy. A fusion is created between the protein of interest and hAGT. In the next step, ligation between the fusion protein and the label of choice (biotin, fluorescent or bifunctional tag[27]) is accomplished by nucleophilic attack of hAGT's activated sulfur on the benzylic position of modified guanine derivatives (see Scheme 3). It should be noted that special care needs to be taken with mammalian cells, since these also express AGT which can react with the 6-*O*-alkylated guanine, thereby causing high background labeling.[28] This problem can be avoided by the use of an AGT knock-out cell line. Furthermore, the higher activity of AGT-mutants relative to wild-type AGT allows selective substoichiometric labeling of AGT fusion proteins.[29] The same group developed an extension of this approach, in which a chemical inducer of dimerisation (methotrexate) was linked to a hAGT-fusion protein, thereby constructing heterodimers of proteins that regulate DNA transcription in yeast.[30] This method can be regarded as a more flexible alternatieve of the GFP method, in the sense that the tag that is transferred to the fusion protein of interest can be chosen liberally. Although hAGT is 31 amino acid residues smaller than GFP and is thus expected to cause less perturbation of the protein of interest, it is still rather bulky. In that sense, it suffers from the same main objection against the GFP method.



Scheme 3. The labeling of hAGT-fusion proteins using benzylguanine derivatives.

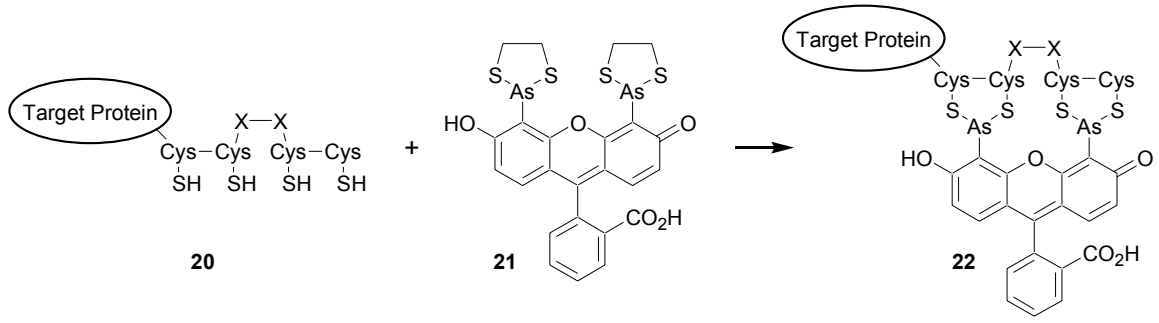


Scheme 4. The reaction of *N*-terminal cysteine engineered proteins with thioester-derivatised labels.

The problem of unwanted cross-reactivity with native proteins, as noted in the example above, is circumvented in an elegant approach by the group of Yao,[31] who make use of the highly selective and biocompatible native chemical ligation reaction of an N-terminal cysteine and a thioester (Scheme 4). Thus, cell permeable probes containing a reporter moiety (label) and a benzyl thioester (**18**) are reacted with an N-terminal cysteine modified target protein overexpressed in *E. coli* cells. In this way, a highly site-selective labeling strategy is attained, without the necessity for knock-out cell lines. It should be noted that labeling of internal cysteine residues or other thiols by transesterification is highly likely. However, this non-specific labeling will not be observed since under the conditions of sample preparation for gel electrophoresis all thioesters will be hydrolysed. The advantage of this technique, as compared to the above example, is that the necessary modification (only one cysteine residue) is less perturbing in terms of size. However, protein engineering still remains a necessity and side reactions on the N-terminal cysteine modified proteins may be introduced.

An exquisite combination of organic chemistry and biochemical engineering is represented by the fluorescent labeling strategy developed by Tsien and coworkers.[32] The toxicity of trivalent arsenic compounds, originating from strong binding to thiol pairs of adjacent cysteine residues, inspired the design of novel protein labels. Fluorescent biarsenides (exemplified by the so-called FlAsH derivative **21**, see Scheme 5) were prepared and studied with respect to their binding affinity with synthetic cysteine-containing oligopeptides. It was found that peptides containing the CCXXCC motive bind strongly and selectively to FlAsH. Fusion of the tetracysteine motif to target proteins now allows their labeling with FlAsH or its analogues[33] at the desired time. The validity of this approach was illustrated by the Tsien group in the following experiment.[34] Cells were generated that express recombinant connexin fused with the tetracysteine tag. Connexins are cell surface proteins that establish channels between two aligned cells for exchange of subcellular material. Treatment of the culture with the cell-permeable FlAsH compound

resulted in selective labeling of clusters of cell surface connexin molecules. At this point, an interesting feature of the two-step labeling presented itself. Whereas FlAsH molecules bound to recombinant connexins appeared to be relatively stable, excess non-bound FlAsH could be readily washed from the culture. Then, treatment with another biarsenic label, ReAsH (which colours red as opposed to the green FlAsH), rendered any newly synthesised connexin red, while the earlier synthesised connexins remained green. In this way, important information on the development of connexin patches was obtained. It is unlikely that similar results can be obtained by making use of GFP fusion type constructs.



Scheme 5. The reaction of a fusion protein containing a tetracystein motif with the FlAsH reagent (**21**).

The toxicity and the high background signals arising from FlAsH derivatives as well as the cross reactivity as observed in the hAGT-approach are addressed in the approach of Nolan.[35] The high-affinity interaction between a FKBP12 mutant protein and the synthetic ligand, SLF', is exploited to label and inactivate proteins in living cells by fluorophore assisted laser inactivation (FALI). Fusion proteins of β-galactosidase and mutant FKBP12 were expressed in NIH 3T3 cells. The cells were incubated with cell permeable fluorescein derivative of SLF', allowing detection of labeled fusion proteins with confocal microscopy. In a subsequent FALI-experiment, irradiation of the fluorescein group generated oxygen radicals near the galactosidase fusion protein, thereby inactivating the enzyme. This method can be used to create spatially and temporally defined enzyme inactivation.

C. Tagging of enzymes with activity-based probes

The examples described thus far entail the development of techniques that enable the labeling of protein families through the incorporation of selective attachment sites in either the polypeptide backbone or in post-translationally added entities. After chemoselective ligation of the reporter group, information on both expression levels and the

localisation of the targeted protein (family) is obtained. Families of enzymatic activities provide an alternative strategy for chemoselective attachment of reporter molecules. Chemical probes that are designed to bind specifically and irreversibly within the active site of a given enzyme, and that are equipped with a reporter molecule, may in principle report on both localisation and expression level. Moreover, the fact that binding is dependent on actual enzymatic activity introduces a third functional read-out, namely that of proper functioning of the targeted enzyme(s). The value of this strategy has been widely recognised in recent years, especially in the field of proteolysis.[36] In general, short oligopeptides that serve as recognition elements for a given protease (family) are equipped with both an electrophilic trap (a chemical entity with high reactivity towards the enzyme active site) and a reporter group (for instance biotin, fluorescent group, radioisotope). The resulting construct is referred to as an activity-based probe (ABP). Treatment of cell lysates with an ABP will result in irreversible labeling of the targeted proteases, which can be visualised and analysed using standard gel electrophoresis. A distinct advantage of the ABP strategy over related strategies that are based on the use of fluorogenic oligopeptide substrates is the covalent and irreversible nature. Modified proteases can thereby be visualised and, optionally, isolated and identified in a later stage of the experiment, facilitating profiling of the enzymatic activities in complex mixtures. ABP labeling of many protease families, including serine proteases,[37] cysteine proteases (cathepsins,[38] caspases[39]) and the proteasome,[4] has been reported in recent years. A drawback encountered in these procedures, however, is the finding that the presence of the reporter entity generally diminishes or even precludes cellular uptake[4] of the probes, necessitating cellular lysis prior to the labeling experiment. To overcome this shortcoming, several groups have endeavoured to introduce small latent ligation sites in ABPs, with the idea that such constructs would retain cell permeability. After labeling in living cells, attachment of the reporter entity may then be executed *ex vivo* through a chemoselective ligation reaction.

This strategy is well illustrated by the recent work of Cravatt and co-workers on the profiling of a range of enzyme activities in mammalian tissue cultures (for instance human breast cancer glutathione-*S*-transferase and mouse liver aldehyde dehydrogenase).[40] Making use of the finding that a diverse range of enzymes are irreversibly disabled by simple benzenesulfonyl derivatives, a set of azide containing ABPs were prepared (such as compound **23**, Figure 5). Cells were incubated with the sulfonate, washed, lysed and the azide moiety was addressed in the cell lysate by an alkyne-functionalysed rhodamine (**24**)

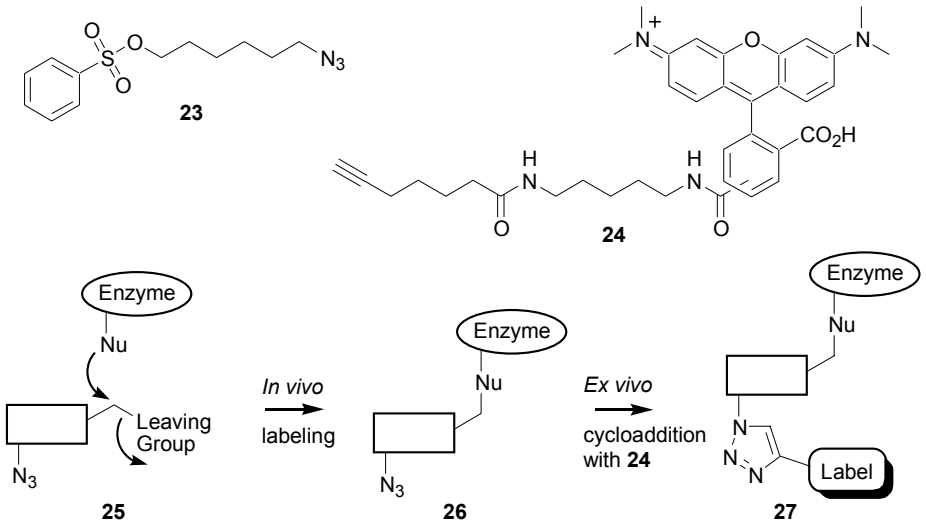


Figure 5. Two-step activity-based enzyme profiling using Huisgen-type cyclisation.

in a stepwise Huisgen reaction. In a more recent report,[41] it was shown that the alkyne moiety is not fully bioorthogonal. In a reversed experiment, the ABP was equipped with an acetylene moiety and the reporter entity with an azide. It appeared that application of the alkyne-functionalysed rhodamine (and therefore an azide-ABP) led to higher background labeling than an azide-functionalised rhodamine (with an alkyne ABP). Also, the Cravatt group demonstrated the first labeling of proteins in live organisms (mice), and subsequent in-gel visualisation of retrieved proteins.[40]

Conclusion and perspectives

The approaches presented all rely on the organic-chemical modification of biomolecules (in most cases proteins) in living cells, with the aim to visualise and/or retrieve the biomolecules or their metabolites. Application of organic chemistry allows the introduction of a multitude of modifications into selected biomolecules. The philosophy behind executing the labeling in living cells stems from the desire to trap the targets as much in their natural environment as possible, so that artefacts related with cell lysis are reduced to a minimum.

In this sense, it may seem counterintuitive to use molecular biological techniques to engineer proteins in such a way as to accommodate a modification. In the examples shown in category B (see above), fusion proteins are created, containing an attachment site (ranging from a single amino acid - N-terminal Cys - to an entire protein - hAGT). Although these modifications are all smaller than GFP, the target protein that is studied is

considerably altered compared to its natural counterpart. Further, in the case of the hAGT examples, an additional artefact is introduced by the need to knock out the native AGT mechanism (in the case of mammalian tissue) in order to diminish background labeling. Finally, the creation of fusion proteins for each separate process that one would like to study constitutes a tremendous amount of biotechnological effort. However, the collective strongpoint of the approaches described in this category is that these are, up to now, the only truly *in vivo* chemical labeling strategies, in the sense that both steps (the tagging and ligation) occur in live cells.

A less invasive strategy is to incorporate, at a designated site of a biomolecule, an inconspicuous bioorthogonal modification that can be addressed after the fact (as described under category A). Techniques required for altering the protein synthesis machinery in such a way that it accepts a non-natural amino acid with a bioorthogonal handle are by no means standardised yet. Furthermore, all potential target proteins have to be engineered separately. In this respect, the other approach described under A (to feed the cell certain modified metabolic precursors that are later expressed in sites of interest) seems less laborious, in the sense that non-engineered (wild type) tissues can be studied. Of course, this strategy only works if the modified precursors are accepted by the natural metabolic pathway. But with the groundbreaking work described in section A, and some careful considerations of the modified precursor to be introduced (in terms of site of modification, its size and thereby conspicuousness), it is possible to envisage applications in the area of, for instance, lipid metabolism and post-translational modifications other than glycanation or glycosidation. In fact, the metabolic incorporation of a synthetic azido-farnesyl analogue into farnesylated proteins has been shown recently.[42]

The third category, the activity-based enzyme profiling approach, is also independent of protein engineering. The intrinsic reactivity of native enzymes is exploited to a) visualise their activity and b) retrieve the biomolecules from a biological sample. It may seem that this strategy can only be applied to enzymatic activities that actually partake, through the involvement of an amino acid side chain functionality inherent to the enzyme, in the chemical transformation at hand. However, it may well be that, in the near future, hydrolytic activities that involve water as the nucleophilic species (rather than a serine- or cysteine residue), may be addressed in an ABP fashion, for instance through the application of well-designed suicide substrates.

In the field of activity based profiling, there are many areas still to be developed. Firstly, all studies done so far in the area of activity-based profiling have been on the qualitative level (whether a labeled protein can be detected or not). Not much effort has been devoted to efficacy studies of both the tagging and the subsequent ligation step, in terms of side reactions, molar yield and kinetics. Of course, this requires robust analytical methods for the characterisation of the tagging and ligation products. With detailed information in this area in hand, it will be possible to advance the *in vivo* labeling strategy to a truly quantitative approach, in which the levels of certain enzymatic activities can be compared between, for instance, diseased and control tissues. One can readily imagine a merge of activity based profiling of enzymatic activities with the isotope-encoded affinity probe techniques developed by Aebersold and coworkers.[43] In such a strategy, tissue from different sources would be treated with a pair of ABPs differing only in hydrogen/deuterium content.

There is much to be gained from studies on enzymatic activities other than proteases, which up to now have been the subject of *in vivo* ABP-based techniques. With the advances made in the design and synthesis of irreversible inhibitors for other enzymes (for instance glycosidases and phosphatases), it is now possible to embark on similar *in vivo* strategies with these enzymes. The first step toward *in vivo* activity based profiling of glycosidases has been made by the Bertozzi group. A known mechanism-based glycosidase inhibitor was equipped with an azide moiety and used to profile enzyme activity in cell lysates.[44]

A last, but certainly not least, perspective on strategy C entails the approach best described as forward chemical genomics.[45] This entails the incubation of a certain tissue with a label of choice, followed by the ligation reaction, retrieval and identification of the products. This may lead to the discovery of enzymatic activities or metabolites that may not have been envisaged or have been overlooked, because they are, for instance, not present in cell lysates.

It can be concluded from the examples described in this chapter that the labeling of proteins in living cells using bioorthogonal organic chemistry is an effective and reliable approach with high potential for further application. Although the described approaches which employ organic chemistry, rather than biotechnology, are only just past the stage of proof-of-principle, it is obvious that there is a considerable interest from both the organic

chemistry, biology and medical research communities. It is therefore only a matter of time before these techniques will find acceptance and applicability in fundamental biochemistry, pharmaceutics and medicine.

References and notes

1. Original paper: P. F. van Swieten, M. A. Leeuwenburgh, B. M. Kessler and H. S. Overkleeft, *Org. Biomol. Chem.*, 2005, **3**, 20-27.
2. R. Y. Tsien, Ann. Rev. Biochem., 1998, **67**, 509-544.
3. J. Neefjes and N. P. Dantuma, *Nat Rev. Drug Discov.*, 2004, **3**, 58-69; J. Zhang, R. E. Campbell, A. Y. Ting and R. Y. Tsien, *Nat. Rev. Mol. Cell Biol.*, 2002, **3**, 906-918.
4. B. M. Kessler, D. Tortorella, M. Altun, A. F. Kisselev, E. Fiebinger, B. G. Hekking, H. L. Ploegh and H. S. Overkleeft, *Chem. Biol.*, 2001, **8**, 913-929.
5. S. W. Michnick, *Drug Discov. Today*, 2004, **9**, 262-267.
6. Bioorthogonal is defined in this chapter as: not associated with - and chemically inert to - the conditions in the particular physiological system.
7. T. Hohsaka and M. Sisido, *Curr. Op. Chem. Biol.*, 2002, **6**, 809-815; A. Strømgaard, A. A. Jensen and K. Strømgaard, *ChemBioChem*, 2004, **5**, 909-916.
8. D. H. Dube and C. R. Bertozzi, *Curr. Op. Chem. Biol.*, 2003, **7**, 616-625.
9. N. Johnsson and K. Johnsson, *ChemBioChem*, 2003, **4**, 803-810.
10. N. Jessani and B. F. Cravatt, *Curr. Op. Chem. Biol.*, 2004, **8**, 54-59.
11. W. A. Hendrickson, J. R. Horton and D. M. Lemaster, *EMBO J.*, 1990, **9**, 1665-1672; N. Budisa, B. Steipe, P. Demange, C. Eckerskorn, J. Kellerman and R. Huber, *Eur. J. Biochem.*, 1995, **230**, 788-796; H. Duewel, E. Daub, V. Robinson and J. F. Honek, *Biochemistry*, 1997, **36**, 3404-3416.
12. J. C. M. van Hest, K. L. Kiick and D. A. Tirrell, *J. Am. Chem. Soc.*, 2000, **122**, 1282-1288; K. L. Kiick, J. C. M. van Hest, and D. A. Tirell, *Angew. Chem., Int. Ed.*, 2000, **39**, 2148-2152.
13. K. L. Kiick, E. Saxon, D. A. Tirell and C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 19-24.
14. A. J. Link and D. A. Tirell, *J. Am. Chem. Soc.*, 2003, **125**, 11164-11165.
15. C. W. Tornøe, C. Christensen and M. Meldal, *J. Org. Chem.*, 2002, **67**, 3057-3064.
16. V. V. Rostovsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chem., Int. Ed.*, 2002, **41**, 2596-2599.
17. J. W. Chin, T. A. Cropp, J. C. Anderson, M. Mukherji, Z. Zhang and P. G. Schultz, *Science*, 2003, **301**, 964-967.
18. R. A. Mehl, J. C. Anderson, S. W. Santoro, L. Wang, A. B. Martin, D. S. King, D. M. Horn and P. G. Schulz, *J. Am. Chem. Soc.*, 2003, **125**, 935-939.

19. A. Deiters, T. A. Cropp, M. Mukherji, J. W. Chin, J. C. Anderson and P. G. Schultz, *J. Am. Chem. Soc.*, 2003, **125**, 11782-11783.
20. J. Brunner, *Chem. Soc. Rev.*, 1993, **22**, 183-189.
21. J. W. Chin, S. W. Santoro, A. B. Martin, D. S. King, L. Wang and P. G. Schultz, *J. Am. Chem. Soc.*, 2002, **124**, 9026-9027.
22. J. W. Chin and P. G. Schultz, *ChemBioChem*, 2002, **3**, 1135-1137.
23. E. Saxon and C. R. Bertozzi, *Science*, 2000, **287**, 2007-2010.
24. J. A. Prescher, D. H. Dube and C. R. Bertozzi, *Nature*, 2004, **430**, 873-877.
25. D. J. Vocadlo, H. C. Hang, E.-J. Kim, J. H. Hanover and C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 9116-9121.
26. A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel and K. Johnsson, *Nat. Biotechnol.*, 2003, **21**, 86-89.
27. M. Kindermann, I. Sielaff and K. Johnsson, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 2725-2728.
28. A. Keppler, M. Kindermann, S. Gendreizig, H. Pick, H. Vogel and K. Johnsson, *Methods*, 2004, **32**, 437-444.
29. A. Keppler, H. Pick, C. Arrivoli , H. Vogel and K. Johnsson, *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 9955-9959.
30. S. Gendreizig, M. Kindermann and K. Johnsson, *J. Am. Chem. Soc.*, 2003, **125**, 14970-14971.
31. D. S. Y. Yeo, R. Srinivasan, M. Uttamchandani, G. Y. J. Chen, Q. Zhu and S. Q. Yao, *Chem. Commun.*, 2003, 2870-2871.
32. B. A. Griffin, S. R. Adams and R. Y. Tsien, *Science*, 1998, **281**, 269-272.
33. S. R. Adams, R. E. Campbell, L. A. Gross, B. R. Martin, G. K. Walkup, Y. Yao, J. Llopis and R. Y. Tsien, *J. Am. Chem. Soc.*, 2002, **124**, 6063-6076.
34. G. Gaietta, T. J. Deerinck, S. R. Adams, J. Bouwer, O. Tour, D. W. Laird, G. E. Sosinsky, R. Y. Tsien and M. H. Ellisman, *Science*, 2002, **296**, 503-507.
35. K. M. Marks, P. D. Braun and G. P. Nolan, *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 9982-9987.
36. A. E. Speers and B. F. Cravatt, *ChemBioChem*, 2004, **5**, 41-47.
37. D. Kidd, Y. Liu and B. F. Cravatt, *Biochemistry*, 2001, **40**, 4005-4015.
38. D. Greenbaum, K. F. Medzihradszky, A, Burlingame and M. Bogyo, *Chem. Biol.*, 2000, **7**, 569-581; J.-P. Falgueyret, W. C. Black, W. Cromlish, S. Desmarais, S. Lamontagne, C. Mellon, D. Riendeau, S. Rodan, P. Tawa, G. Wesolowski, K. E. Bass, S. Venkatraman and M. D. Percival, *Anal. Biochem.*, 2004, **335**, 218-227.
39. L. P. Tan, R. C. Panicker, G. Y. J. Chen and S. Q. Yao, *Chem. Commun.*, 2005, 596-598; M.-L. Liau, R. C. Panicker and S. Q. Yao, *Tetrahedron Lett.*, 2003, **44**, 1043-1046.
40. A. E. Speers, G. C. Adam and B. F. Cravatt, *J. Am. Chem. Soc.*, 2003, **125**, 4686-4687.
41. A. E. Speers and B. F. Cravatt, *Chem. Biol.*, 2004, **11**, 535-546.

42. Y. Kho, S. C. Kim, C. Jiang, D. Barma, S. W. Kwon, J. Cheng, J. Jaunbergs, C. Weinbaum, F. Tamanoi, J. Falck and Y. Zhao, *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 12479-12484.
43. S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb and R. Aebersold, *Nat. Biotechnol.*, 1999, **17**, 994-999.
44. D. J. Vocadlo and C. R. Bertozzi, *Angew. Chem., Int. Ed.*, 2004, **43**, 5338-5342.
45. S. M. Khersonsky and Y.-T. Chang, *ChemBioChem*, 2004, **5**, 903-908.

Chapter 2

Detection of active proteasomes by a two-step labeling strategy[1]

Introduction

Major efforts in unraveling of the human genome and the genetic material of most relevant human pathogens, enabled the focus in biomedical and biological sciences to shift towards the global assessment of expression levels and function of the gene products. The reason for the renewed interest in protein activity is obvious: biological processes are modulated at the protein level in health and disease. Approaches that report on transcription levels are not informative for the levels of activities of the products encoded by these transcripts. Equally important, the relevant activities are those in living cells and not what is measured in vitro. At the same time, the global assessment of highly complex and dynamic protein mixtures, as found in intact cells, is a much more arduous task than that of the relatively static genome. This holds true especially when aiming for insight in the activity of proteins rather than for their expression levels.

Chemistry-based functional proteomics approaches[2,3] have been developed based on the use of synthetic compounds that modify a selected subset of proteins covalently and

irreversibly. These methodologies combine the attractive features of simplifying the complex proteome by selecting protein families on the basis of their function.[4] For instance, broad-spectrum, irreversible protease inhibitors have been used in the profiling of serine proteases,[5] cysteine proteases,[6] and the catalytically active subunits of the proteasome.[7,8] These inhibitors are equipped with either a radioisotope, a biotin moiety or a fluorescent tag, to allow respectively visualization, isolation and quantification of the proteases. The cell-impermeability of the compounds used in these examples limits their use to in vitro applications and this limitation is an unavoidable consequence of the use of such probes.

Two-step labeling strategy

Here, a new functional proteomics strategy is described, that allows two-step labeling of the catalytically active subunits of the proteasome in living cells.[9] The proteasome is a multi-catalytic protease that accounts for the bulk of cytosolic and nuclear proteolysis.[10] Previously, a set of extended, peptide-based, irreversible proteasome inhibitors have been reported.[7] The most potent of these, AdaAhx$_3$L$_3$VS **1** (Figure 1), is unique in that it targets all catalytically active β-subunits of both the constitutive- and the interferon-γ-inducible immunoproteasome with approximately equal efficiency. Although AdaAhx$_3$L$_3$VS **1** proved to be cell permeable, it lacks a label that allows easy detection.

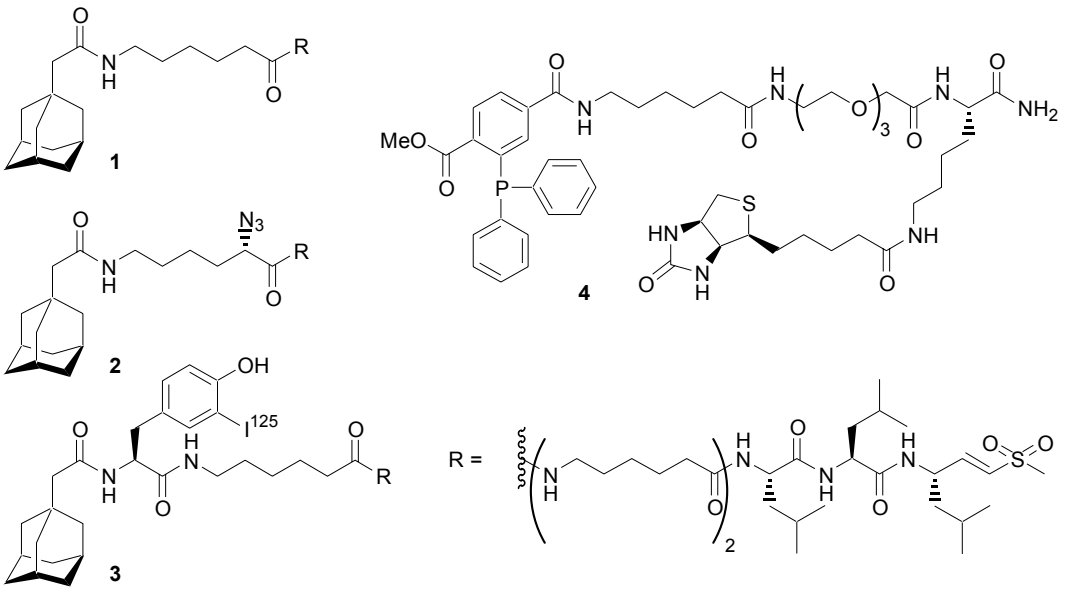


Figure 1. Structure of the broad-spectrum cell-permeable proteasome inhibitor **1**, the cell-impermeable radioiodinated derivative **3**, and the Staudinger ligation partners **2** and **4** described here.

Here, it is demonstrated that modification of **1** with an azide group (to give **2**; Figure 1) neither interferes with its inhibitory profile nor with its cell permeability.[11] Labeling of whole cells with **2** decorates the catalytically active β-subunits of the proteasome with an azide as a latent ligation handle. After cell lysis, retrieval and denaturation of cellular protein content, the azido groups can be addressed by the biotinylated phosphane reagent **4** in a modified Staudinger ligation, as developed by Bertozzi and coworkers (Figure 2).[12,13]

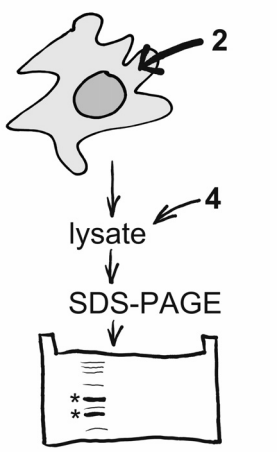
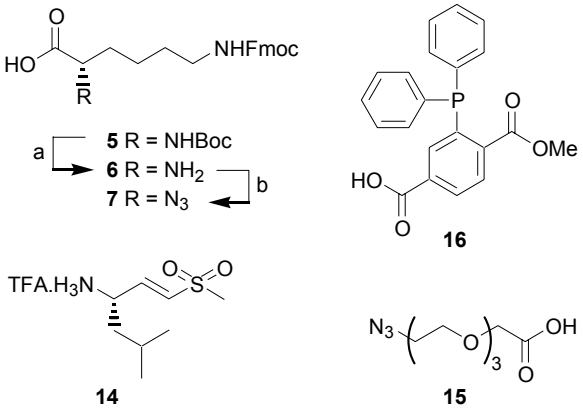


Figure 2. General strategy: irreversible proteasome inhibition and subsequent Staudinger ligation-mediated biotinylation to enable visualization of proteasomes in an activity-based manner.

Synthesis

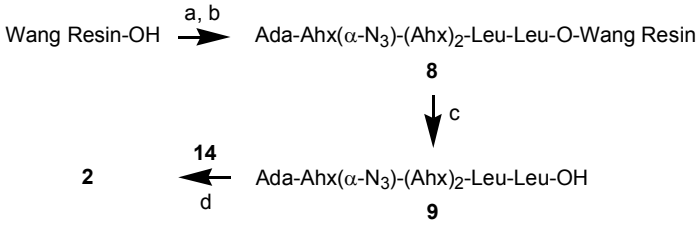
The synthesis of the azide-containing proteasome inhibitor **2** was accomplished as follows (Scheme 1 and 2). Treatment of Boc-Lys(Fmoc)-OH (**5**) with trifluoroacetic acid, followed by subjection to diazotransfer conditions (TfN_3, $CuSO_4$)[14] readily afforded (2*S*)-2-azido-6-(fluorenylmethyloxycarbonylamino)-hexanoic acid (**7**; Scheme 1). Standard Fmoc-based solid phase peptide synthesis (SPPS) afforded immobilized peptide **8** on acid-labile Wang resin (Scheme 2). Cleavage from the resin and solution-phase condensation of carboxylate **9** with leucine vinyl sulfone **14**[15] afforded target compound **2** that was purified by silica gel chromatography (58% overall yield).



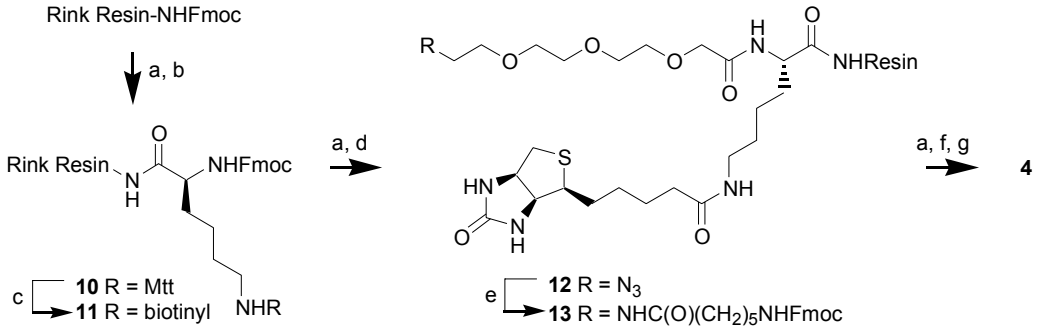
Scheme 1. Non-standard building blocks used in the synthesis of **2** and **4**. Reagents and conditions: a) 50% TFA/CH_2Cl_2, *92%*; b) TfN_3, $CuSO_4$, H_2O/MeOH, *89%*.

The synthesis of phosphane reagent **4** commenced with the condensation of Rink amide linker with Fmoc-Lys(Mtt)-OH (Scheme 3).[16] The side chain protective group in **10** was removed, followed by condensation with biotin to afford immobilized biocytin **11**. Standard solid phase peptide synthesis employing azido acid

15,[17] Fmoc-aminohexanoic acid and phosphane **16**, respectively, followed by acidic cleavage from the resin and HPLC purification gave the homogeneous target compound **4** in 16% yield.



Scheme 2. Synthesis of azide-containing proteasome inhibitor **2**. Reagents and conditions: a) Fmoc-Leu-OH, DIC, DMAP, CH_2Cl_2; b) Repeated cycles of SPPS: Fmoc cleavage: 20% piperidine in DMF; amino acid condensation: Fmoc protected amino acid, PyBOP, DiPEA, DMF. Fmoc protected amino acids building blocks employed in consecutive order: Fmoc-Leu-OH, Fmoc-Ahx-OH, Fmoc-Ahx-OH, **7**, adamantane acetic acid; c) 10% H_2O/TFA; d) HBTU, DiPEA, DMF, 58% overall yield.



Scheme 3. Synthesis of biotinylated phosphane reagent **4**. Reagents and conditions: a) 20% piperidine in DMF; b) Fmoc-Lys(Mtt)-OH, PyBOP, DiPEA, DMF; c) 1% TFA/CH_2Cl_2, then biotin, PyBOP, DiPEA, DMF; d) **15**, PyBOP, DiPEA, DMF; e) Me_3P, 20% H_2O/dioxane, then Fmoc-Ahx-OH, PyBOP, DiPEA, DMF; f) **16**, EDC, HOBt, CH_2Cl_2; g) 50% TFA/CH_2Cl_2, 16% overall yield.

Inhibitory results

In order to establish the inhibition profile of **2**, a set of competition experiments was performed. The cell line used was EL-4, derived from a murine thymoma. This cell line is advantageous because it expresses both the constitutive proteasome and the interferon-γ-inducible immunoproteasome, containing, in total, 6 distinct catalytically active β-subunits. The cell line HeLa, derived from a human cervical carcinoma, was also used. HeLa expresses only the constitutive proteasome (3 catalytically active β-subunits). When lysates of EL-4 cells were incubated with azide-containing proteasome inhibitor **2** at different

concentrations, prior to treatment with radio-iodinated peptide vinyl sulfone **3** (Figure 1), it was observed (Figure 3A) that labeling of the six individual subunits was abolished at final inhibitor concentrations of 10-30 μM, demonstrating **2** to be a proteasome inhibitor of equal potency as **1**.[7] The ability of peptide vinyl sulfone **2** to disable the proteasome in living cells

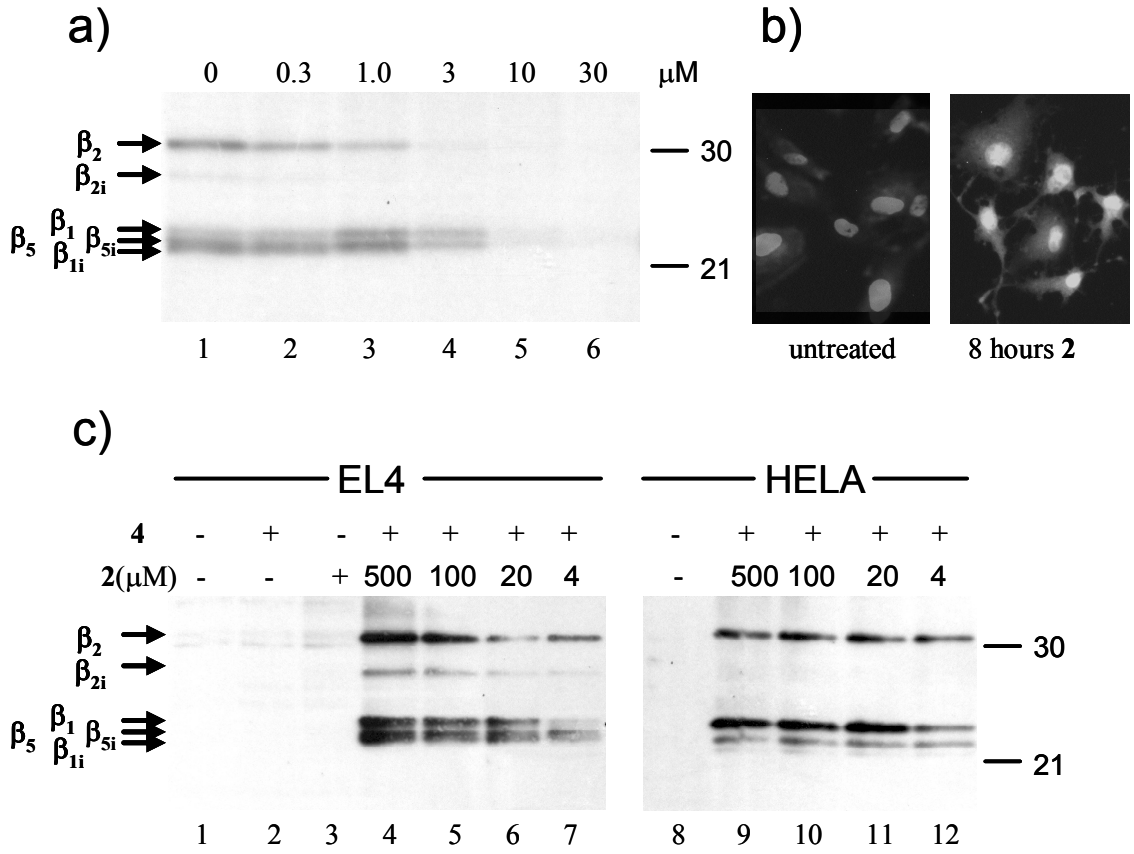


Figure 3. a) Cell lysate prepared from EL-4 cells was incubated with **2** at concentrations ranging from 0 to 30 μM. Residual unmodified subunits were labeled by subsequent incubation with radioiodinated inhibitor **3**. Labeled subunits were resolved by SDS-PAGE and visualized by autoradiography. b) Ub-R-GFP accumulates when proteasomal degradation is blocked. Cells were incubated either with a solvent control or **2** (50 μM final concentration) for 8 hours and fixed, followed by blue nuclear staining of the DNA with DAPI (blue). Confocal laser scanning microscopy revealed **2** to be a cell-permeable proteasome inhibitor. c) Lysates from EL-4 and HeLa were reacted with **2** at 37˚C for one hour and then boiled in the presence of SDS, to achieve protein denaturation and exposure of the azido moieties of conjugated **2**. The azido moieties were biotinylated *via* a Staudinger ligation by adding an aqueous solution of reagent **4** to the reaction mixture, followed by incubation for 1 h at 37˚C. Samples were separated by SDS-PAGE and transferred to polyvinylidene difluoride membrane. Incubation with streptavidin-horseradish peroxidase (strept-HRP) conjugates allowed the visualization of active proteasomal β-subunits by chemiluminescence.

was determined with the following procedure: U373 cells expressing a green fluorescent protein (GFP)-ubiquitin fusion protein[18] (Ub-R-GFP) were treated with compound **2** at 50 μM (final concentration) and compared with untreated cells for the presence of GFP fluorescence. Ub-R-GFP is rapidly degraded by the proteasome with little green fluorescence at steady state as a result (Figure 3B). However, in cells treated with **2**, a time-dependent accumulation of fluorescence was observed, demonstrating the capacity of **2** to inactivate the proteasome in living cells.

Labeling of active proteasomes

Encouraged by these results, the suitability of a Staudinger ligation was established for two-step visualization of catalytically active proteasome subunits in cell lysates as well as in living cells. In the first experiment, cell lysates from EL-4 and HELA cells were exposed to **2** at various concentrations, prior to denaturation of cellular protein. The resulting mixtures were incubated with biotinylated Staudinger ligation reagent **4** and separated by SDS-PAGE. Transfer of the separated protein mixture onto a polyvinylidene difluoride (PVDF) membrane, followed by chemiluminescence induced by horseradish peroxidase-streptavidin conjugate, resulted in a distinct labeling profile. Labeling intensity depended on the dose of **2**. The pattern of labeling conforms to that established for radio-iodinated probe **3**.[7] Importantly, proteasome derived polypeptides were detected only when both inhibitor **2** and Staudinger reagent **4** were used (Figure 3C, lanes 4-7 and 9-12). These results establish the selectivity of **4** in complex physiological mixtures to target only those proteins modified with an azide functionality.

Next, the possibility of covalent proteasome inhibition in living cells followed by post-lysis Staudinger ligation and immunoblotting was investigated. EL-4 cells were incubated overnight with **2** (Figure 4). Subsequent glass bead lysis, incubation with **4**, ensuing SDS-PAGE separation and Western blotting afforded a labeling pattern virtually indistinguishable from

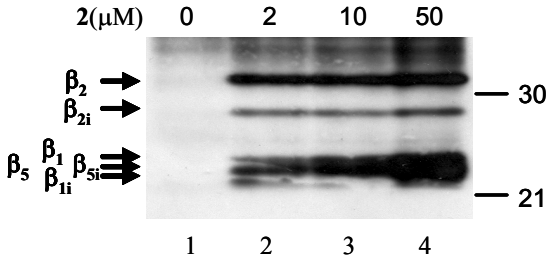


Figure 4. Proteasome labeling in living cells. Incubation of living cells (5 × 10^6 at 37°C) with **2** followed by post-lysis ligation and immunoblotting reveals the active proteasomal content and composition in living cells.

that obtained for labeling of cell lysates (Figure 3C). Importantly, *in vivo* labeling appeared to be more effective (compare Figure 4 with Figure 3C), indicating a more efficient targeting of all proteasomal subunits in living cells as compared to the labeling patterns obtained *in vitro*. The latter observation is possibly due to partial dissociation of the proteasome particle during cell lysis and storage. It was concluded that inhibitor **2** can be used in combination with biotinylation reagent **4** for the visualization of active proteasomes in living cells.

Conclusions

In summary, we have presented a novel strategy to visualize active enzymes in living cells. Compound **2** was identified as a powerful, cell-permeable inhibitor of all proteasomal activities, and **2** can subsequently be addressed by post-lysis labeling *via* a chemoselective Staudinger ligation. This protocol opens the way towards the screening, in living cells, of proteasomal activity in, for example, human tissue samples.

The measurement of proteasome activity in live cells remains an important goal, not only in the context of novel treatment strategies for cancer, but also in biological systems more generally. For instance, malfunctioning of the ubiquitin-proteasome system has been implicated in both cancer[19] and neurodegeneration.[20]

Importantly, the two-step methodology (this is, covalent, irreversible enzyme modification followed by chemoselective modification) may be extended towards the development of novel chemoselective ligation partners compatible with desired cellular environments. The application of this strategy for the assessment of the activity of other protease enzyme activities in living cells is further elaborated in chapter 3.

Experimental section

General methods and materials: Solvents used in the solid phase peptide synthesis, DiPEA and TFA were all of peptide synthesis grade (Biosolve) and used as received. The protected amino acids, Wang resin (0.86 mmol/g), Rink amide MBHA resin (0.78 mmol/g) and PyBOP were obtained from NovaBiochem. Fmoc-Lys(Mtt)-OH was from Senn Chemicals, EDC and DIC from Acros and anhydrous HOBt from Neosystem. Adamantane acetic acid was purchased from Aldrich. Leucine vinyl sulfone was prepared as reported.[15] SPPS was carried out using a 180° Variable Rate Flask Shaker (St. John Associates, Inc.). LC/MS analysis was performed on a Jasco HPLC system (detection simultaneously at 214 and 254 nm) coupled to a Perkin Elmer Sciex API 165 mass

spectrometer equipped with a custom-made Electrospray Interface (ESI). An analytical Alltima C_{18} column (Alltech, 4.6 × 250 mm, 5 μm particle size) was used. Buffers: A = H_2O; B = CH_3CN; C = 0.5% aq. TFA. For RP HPLC-purification of **2** and **4** a Biocad "Vision" automated HPLC system (PerSeptive Biosystems, Inc.) was used. ^{1}H-NMR and ^{13}C-NMR spectra were recorded with a Bruker AC200 instrument at 200 and 50.1 MHz, respectively.

Solid phase peptide synthesis: Where appropriate removal of the Fmoc protecting group was accomplished by treatment of the resin-bound peptide with 20% (v/v) piperidine in DMF for 20 min. Peptide coupling steps were performed by treatment of the resin with a premixed (5 min) solution of the of the appropriate acid (5 eq.), PyBOP (5 eq.) and DiPEA (6 eq.) in DMF for 1 h. Coupling efficiencies were monitored with the Kaiser test and couplings were repeated if necessary. After coupling and deprotecting steps the resin was washed with DMF (5×).

Synthesis of Fmoc-Ahx(α-N_3)-OH (7): Boc-Lys(Fmoc)-OH (2.0 g, 4.2 mmol) was treated with TFA/CH_2Cl_2 (20 ml, 1/1 v/v) for 40 minutes. Solvents were removed *in vacuo* yielding 1.9 g (3.9 mmol, 92%) of crude TFA-salt of H-Lys(Fmoc)-OH **6**. The crude product was treated with trifluoromethanesulfonyl azide as described[14] followed by a modified work-up procedure. After removal of the organic solvents, the aqueous slurry was acidified with 1 N HCl to pH 2 and extracted with EtOAc (4×). The combined organic phases were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Purification of the crude product over silica gel (hexanes/EtOAc 2/1 v/v) yielded 1.4 g (3.5 mmol, 89%) of a white foam. ^{1}H NMR (200 MHz, $CDCl_3$, 25°C, TMS): δ= 7.75 (d, ^3J (H,H)=7.3 Hz, 2H; ArH), 7.57 (d, ^3J (H,H)=7.3 Hz, 2H; ArH), 7.40-7.25 (m, 4H; ArH), 4.41 (d, ^3J (H,H)=6.6 Hz, 2H; CHCH_2O), 4.19 (t, ^3J (H,H)=6.6 Hz, 1H; CHCH_2O), 3.92 (t, ^3J (H,H)=6.7 Hz, 1H; αCH), 3.18 (m, 2H; εCH_2), 1.7-1.3 ppm (m, 6H; βCH_2, γCH_2, δCH_2); ^{13}C NMR (50.1 MHz, $CDCl_3$, 25°C, TMS): δ=174.1, 156.8, 143.5, 140.9, 127.1, 126.7, 124.6, 119.6, 66.4, 62.1, 60.4, 46.7, 30.6, 28.8, 22.4 ppm; ESI-MS: *m/z* (%) = 417.2 (100) [M+Na^+], 811.5 (20) [2M+Na^+].

Synthesis of 2: Wang resin (1.0 g, 0.86 mmol) was condensed with Fmoc-Leu-OH (1.2 g, 3.4 mmol) in CH_2Cl_2 (25 ml) under the agency of DIC (0.54 ml, 3.44 mmol) and a catalytic amount of 4-(dimethylamino)pyridine for 2 h. The resin was filtered off, washed (3× CH_2Cl_2-MeOH, CH_2Cl_2 and Et_2O), and air-dried. Loading was determined by quantification (UV) of Fmoc cleavage, and proved to be 0.66 mmol/g. The resin, 260 mg (0.17 mmol) was elongated using standard Fmoc-based SPPS to give resin-bound AdaAhx(α-N_3)Ahx_2Leu_2-OH. Treatment of the immobilized peptide with TFA/H_2O (95/5 v/v) for 1 h, filtration and removal of solvent *in vacuo* was followed by a solution phase block coupling with leucine vinyl sulfone TFA salt **14** (1 eq.) under the agency of HBTU (1 eq.) and DiPEA (2.2 eq.) in DMF (1.5 ml). After evaporation of the solvent, the residue was dissolved in EtOAc. The product was precipitated by sonication. The precipitate was filtered and washed with EtOAc, Et_2O and hexanes to yield 146 mg (0.15 mmol, 88%) of the title

compound in 90% purity as judged by LMCS. Silica gel purification (0-10% MeOH in EtOAc) of 21 mg (22 µmol) of the crude product yielded 14 mg (14 µmol) of the title compound (58% overall yield) ready to use in biological experiments. LC/MS: m/z = 974.9 [$M+H^+$].

Synthesis of 4: Fmoc Rink amide resin (128 mg, 100 µmol) was deprotected as described. Fmoc-Lys(Mtt)-OH (0.5 mmol, 312 mg) was coupled. The Mtt protecting group was removed by treatment of the resin with 1% TFA in CH_2Cl_2 for 0.5 min.[16] This treatment was repeated 9 times until no more yellow color was observed in the eluted solution. The resin was neutralized with 10% DiPEA in DMF. Biotin (0.5 mmol, 122 mg) was coupled under the agency of PyBOP (1 eq.) and DiPEA (1.2 eq.) in DMF. After removal of the Fmoc group, the resulting free amine was condensed with **15**[17] (0.5 mmol, 237 mg). After washing of the immobilized peptide with dioxane (3×), the azide moiety was reduced by treatment with Me_3P (0.6 ml of a 1 M solution in toluene, 0.6 mmol, 6 eq.) in dioxane/water (2 ml, 4/1 v/v) for 40 min[14] followed by dioxane washes (3×), and Fmoc-Ahx-OH (0.5 mmol, 180 mg) was coupled. Of the obtained resin, 50 µmol was treated with piperidine as described above to remove the Fmoc protecting group and the resin was then washed with CH_2Cl_2 (3×). Phosphane **16** (90 mg, 0.25 mmol) was activated with EDC (48 mg, 0.25 mmol) and HOBt (41 mg, 0.3 mmol) in CH_2Cl_2 (2 ml) under an argon atmosphere for 5 min, and subsequently added to the resin. The resin was agitated under argon atmosphere for 1 h, and the resin was washed (CH_2Cl_2, then DMF-MeOH alternating (3×), CH_2Cl_2-MeOH alternating (3×), and CH_2Cl_2), while keeping the resin under argon atmosphere. The immobilized peptide was liberated from the resin by treatment with TFA/CH_2Cl_2 (1/1 v/v) for 1 h. Evaporation of the solvents *in vacuo* followed by HPLC-purification of the crude product (linear gradient in B: 25-55% B in 3 column volumes) yielded 8.0 mg (8 µmol, 16%) of a white solid. LC/MS: calculated = 1020 [$M+H^+$], found = 1020 [$M+H^+$]. The product was kept as aliquots of a stock solution (1.6 mM) in degassed DMSO at –80°C.

Proteasome labeling in living cells: EL-4 cells, cultured in RPMI (Gibco, Invitrogen Corp.) supplemented with L-glutamine, fetal calf serum, penicillin and streptomycin were incubated overnight with concentrations of inhibitor **2** as indicated in Figure 3. Cells were harvested by centrifugation. After glass-bead lysis, 10 µg of protein in 33 µl of lysis buffer (50 mM Tris, 5 mM $MgCl_2$, 0.5 mM EDTA, 0.25 mM Sucrose, pH 7.4) was denatured by the addition of 2 µl of 20% SDS followed by brief boiling. The denatured sample was further incubated with 10 µl of **4** (100 µM) in DMSO/CH_3CN/H_2O (6/1/1 v/v/v) for 1 h at 37°C. Without further heating 4× sample buffer was added and samples were run on either a 12.5% or 15% SDS-PAGE gel followed by transfer to PVDF membrane. The membrane was blocked with milk and washed briefly with PBS containing 1% TWEEN-20 (Sigma) (3×) and incubated with streptavidin-HRP (Amersham Life Sciences,

dilution 1: 10.000) for 1 h at ambient temperature. The membrane was washed as before and soaked in Western Lightning Chemiluminescence (Perkin Elmer) reagent followed by developing a film.

References and notes

1. Original publication: H. Ovaa, P. F. van Swieten, B. M. Kessler, M. A. Leeuwenburgh, E. Fiebinger, A. M. C. H. van den Nieuwendijk, P. J. Galardy, G. A. van der Marel, H. L. Ploegh and H. S. Overkleeft, *Angew. Chem., Int. Ed.*, 2003, **42**, 3626-3629.
2. A. Borodovsky, H. Ovaa, N. Kolli, T. Gan-Erdene, K. D. Wilkinson, H. L. Ploegh and B. M. Kessler, *Chem. Biol.*, 2002, **9**, 1149-1159.
3. D. C. Greenbaum, W. D. Arnold, F. Lu, L. Hayrapetian, A. Baruch, J. Krumrine, S. Toba, K. Chehade, D. Brömme, I. D. Kuntz and M. Bogyo, *Chem. Biol.*, 2002, **9**, 1085-1094.
4. G. C. Adam, E. J. Sorensen and B. F. Cravatt, *Nature Biotech*., 2002, **20**, 805-809.
5. G. C. Adam, B. F. Cravatt and E. J. Sorensen, *Chem. Biol.*, 2001, **8**, 81-95.
6. A.-M. Lennon-Duménil, A. H. Bakker, R. Maehr, E. Fiebinger, H. S. Overkleeft, M. Rosemblatt, H. L. Ploegh and C. Lagaudrière-Gesbert, *J. Exp. Med.*, 2002, **196**, 529-539.
7. B. M. Kessler, D. Tortorella, M. Altun, A. F. Kisselev, E. Fiebinger, B. G. Hekking, H. L. Ploegh and H. S. Overkleeft, *Chem. Biol.*, 2001, **8**, 913-929.
8. T. Nazif and M. Bogyo, *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 2967-2972.
9. The latent *in vivo* labeling of serine proteases has also been reported: A. E. Speers, G. C. Adam and B. F. Cravatt, *J. Am. Chem. Soc.*, 2003, **125**, 4686-4687.
10. K. L. Rock and A. L. Goldberg, *Annu. Rev. Immunol.*, 1999, **17**, 739-797.
11. Large substituents such as biotin or fluorescence tags can also have a detrimental effect on the inhibitory profile. Incorporation of ^{125}I-Tyr or biocytin abolishes cell permeability and, in case of biocytin, also partially impairs its inhibition profile.[6]
12. E. Saxon and C. R. Bertozzi, *Science*, 2000, **287**, 2007-2010.
13. K. L. Kiick, E. Saxon, D. A. Tirrell and C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 19-24.
14. J. T. Lundquist and J. C. Pelletier, *Org. Lett*., 2001, **3**, 781-783.
15. M. Bogyo, J. S. McMaster, M. Gaczynska, D. Tortorella, A. L. Goldberg and H. L. Ploegh, *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 6629-6634.
16. L. Bourel, O. Carion, H. Gras-Masse and O. Melnyk, *J. Pept. Sci.*, 2000, **6**, 264-270.
17. S. W. Yeong and D. F. O'Brien, *J. Org. Chem.*, 2001, **66**, 4799-4802.
18. N. P. Dantuma, K. Lindsten, R. Glas, M. Jellne and M. G. Masucci, *Nature Biotech.*, 2000, **18**, 538-543.
19. T. Hideshima and K. C. Anderson, *Nature Rev. Cancer*, 2002, **2**, 927-937.
20. C . Soto, *Nature Rev. Neurosci.*, 2003, **4**, 49-60.

Chapter 3

A cell-permeable inhibitor and activity based probe for the caspase-like activity of the proteasome[1]

Introduction

The proteasome is a multi-catalytic protease that degrades the majority of cytosolic and nuclear proteins to oligopeptides.[2] This degradation is a highly regulated process, and a small portion of the oligopeptides produced escape further degradation and are loaded on MHC class I molecules for presentation to the immune system.[3] Proteasomal protein degradation takes place in a barrel shaped structure that consists of four rings of seven subunits each. Three of the β subunits in the middle two rings, β1, β2 and β5, display protease activity. Each catalytic subunit has a different substrate specificity. The β1 subunit cleaves preferentially after acidic residues and is therefore referred to as the caspase-like subunit. The β2 subunit has a preference for peptide bond cleavage after basic residues and is responsible for the trypsin-like activity. The β5 subunit displays a chymotrypsin-like activity and cleaves after large, hydrophobic residues. In the interferon-γ-inducible immunoproteasome, the catalytic β1, β2 and β5 subunits are replaced by the β1i, β2i and

β5i immunosubunits. The immunosubunits display a substrate specificity slightly different from to that of the respective constitutive subunits. As a result, the immunoproteasome and proteasome generate a different spectrum of oligopeptides from the same protein, which effects a different spectrum of antigens presented on MHC class I molecules.[4]

Inhibitor design

To gain more insight in the role of the different proteasomal subunits in the process of protein degradation and antigen presentation, it would be highly useful to have tools to switch off one of the proteasome activities selectively. Some inhibitors with selectivity for either the trypsin-like or the chymotrypsin-like subunits have been reported.[5] In contrast, no cell-permeable selective inhibitor or activity based probe for the caspase-like proteasomal subunits has been reported to date. In order to develop such tools, precedents from the literature were used to select an appropriate recognition sequence and electrophilic trap. In this respect, it appeared that the selective reversible inhibitor **1** (Figure 1) of the caspase-like subunit of the proteasome was a suitable starting point for the development of an irreversible inhibitor.[6] Furthermore, it has been shown before that vinyl sulfones are suitable electrophilic traps to alkylate the N-terminal threonine residue, the active site nucleophile of the proteasome. The ability of vinyl sulfones to alkylate proteasomal catalytic residues depends on the functional group downstream of the vinyl sulfone moiety. For example, Bogyo and coworkers showed that in case of GL_3-vinyl sulfone, phenolic vinyl sulfone (**2**; Figure 1) was a better inhibitor of the caspase-like site than the corresponding methyl vinyl sulfone (**3**; Figure 1).[7] Thus, a phenolic vinyl sulfone seemed to be the obvious electrophilic trap to be employed.

To avoid the elaborate synthesis of a suitably protected aspartic acid vinyl sulfone moiety that would be required

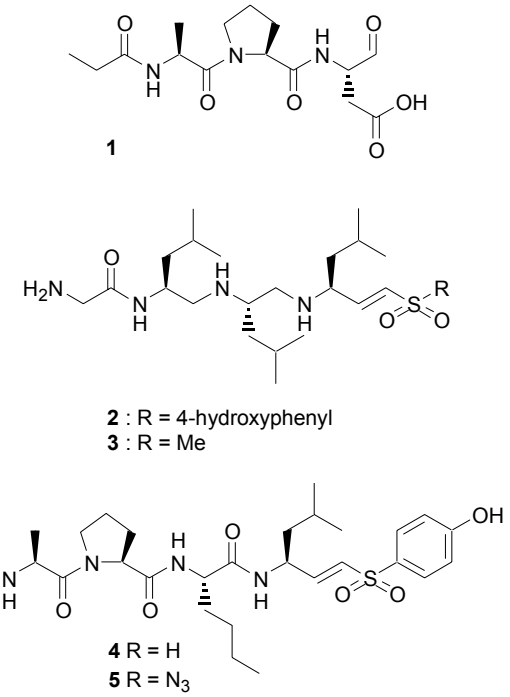


Figure 1. Relevant reversible and irreversible proteasome inhibitors.

for block coupling of the peptide vinyl sulfone building block to a peptide sequence, the P1 aspartic acid was replaced by leucine, thus generating inhibitor **4** (Figure 1). It is known that the caspase-like subunit also cleaves after hydrophobic branched chain amino acids,[6,8,9,10] and this concept has been used in irreversible inhibitor design. An epoxyketone based inhibitor with leucine in the P1 position can inhibit the caspase-like subunit with significant, albeit not absolute, specificity.[11] Furthermore, the leucine residue renders the inhibitor more apolar and thus possibly better cell-permeable.

In order to convert inhibitor **4** into an activity probe for the caspase-like subunit of the proteasome, an azide moiety was introduced on the N-terminal acetyl function to give **4** (Figure 1). In Chapter 2 of this thesis, an approach employing the Staudinger ligation to biotinylate proteasome modified with an azide-containing probe in a two-step labeling fashion is described. Based on literature evidence,[12] direct attachment of a biotin moiety to the probe was expected to influence both cell permeability and subunit specificity.

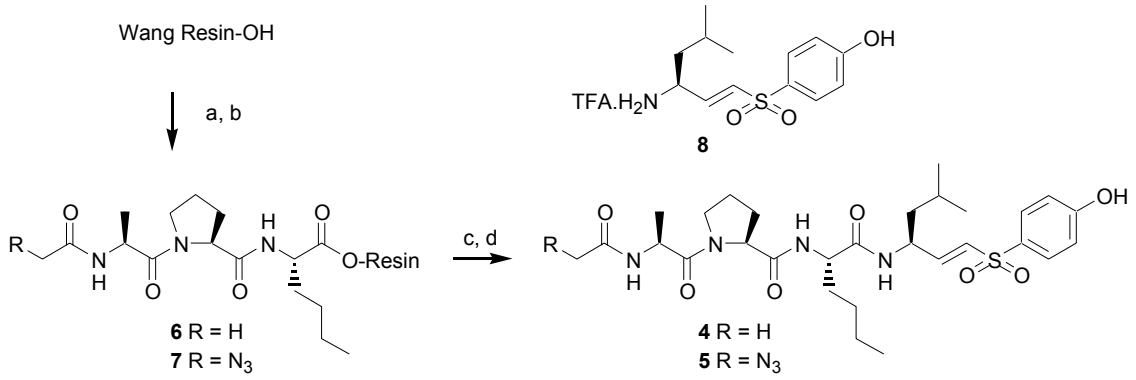
Here, the first highly specific, cell-permeable inhibitors of the caspase-like activities of the proteasome (**4** and **5**; Figure 1) are presented. At low concentrations, these compounds show selectivity towards the caspase-like subunit of the immunoproteasome in favour of the corresponding caspase-like activity of the constitutive proteasome. Activity probe **5** allows two step labeling of active caspase-like subunits of the immunoproteasome and, at higher concentrations, of the constitutive proteasome. Labeled subunits can then be visualized *via* post-lysis ligation and Western blotting.

Synthesis

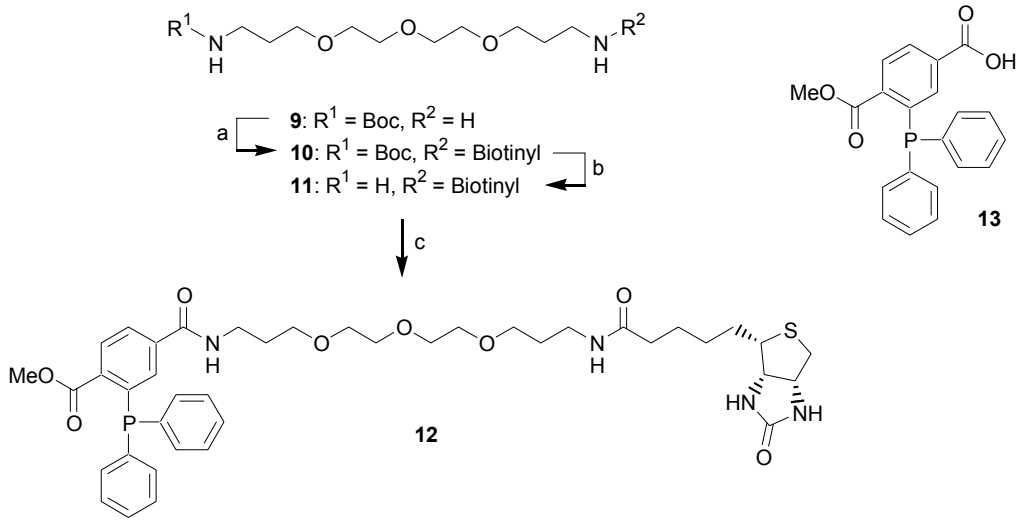
The synthesis of inhibitor **4** started with the synthesis of tripeptide **6** on acid-labile Wang resin (Scheme 1). Acidolysis of the peptide from the resin and subsequent solution phase condensation with vinyl sulfone **8**[7] afforded the target compound in 45% overall yield after RP-HPLC purification. The synthesis of activity probe **5** was accomplished in a similar fashion. Azide-functionalized peptide **7** was synthesized on Wang resin. Cleavage of the peptide from the solid support and coupling to **8** gave target compound **5** in 12% yield after RP-HPLC purification.

For the Staudinger ligation, new biotinylation agent **12** (Scheme 2) was designed. The relatively facile synthesis of **12**, as well as the possibility of purification using silica gel chromatography, allow for production of larger amounts than the comparable reagent

presented in Chapter 3. The synthesis starts with biotinylation of monoprotected bisamine **9** using a standard peptide coupling method. After removal of the Boc protecting group with 5% water in TFA, the liberated amine functionality was acylated with phosphane **13**.[13] Both the use of the EDC/HOBt system to activate the acid moiety and a protective argon atmosphere proved necessary to suppress premature oxidation of phosphane **12** to the corresponding phosphane oxide.



Scheme 1. Synthesis of inhibitor 4 and activity probe **5**. Reagents and conditions: a) Fmoc-Nle-OH, DIC, DMAP, CH_2Cl_2; b) Repeated cycles of SPPS: Fmoc cleavage: 20% piperidine in DMF; amino acid condensation: Fmoc protected amino acid, HCTU, DiPEA, NMP. Fmoc protected amino acid building blocks employed in consecutive order: Fmoc-Pro-OH, Fmoc-Ala-OH, acetic anhydride or azidoacetic acid; c) TFA/H_2O (95/5 v/v); d) **8**, HATU, DiPEA, DMF, overall yield 18% and 26% for **4** and **5** respectively.



Scheme 2. Synthesis of Staudinger-reagent **12**. Reagents and conditions: a) biotin, HCTU, DiPEA, DMF; b) TFA/H_2O (95/5 v/v), 56% yield over two steps; c) **13**, HOBt, EDC, DiPEA, CH_2Cl_2, argon atmosphere, 22% yield.

Evaluation of inhibitory activity

The subunit specificity of newly synthesized inhibitor **4** was determined as follows. Purified rabbit muscle 26S proteasome was incubated in the presence or absence of 50 μM inhibitor, and aliquots of the mixture were assayed with fluorogenic substrates for the respective active site. The effect of the inhibitor on the proteasomal activity was determined by relative fluorescence in the presence or absence of inhibitor.[14] Figure 2 shows that **4** selectively blocks the caspase-like site of the proteasome. Furthermore, an increased activity of the trypsin-like site is observed.

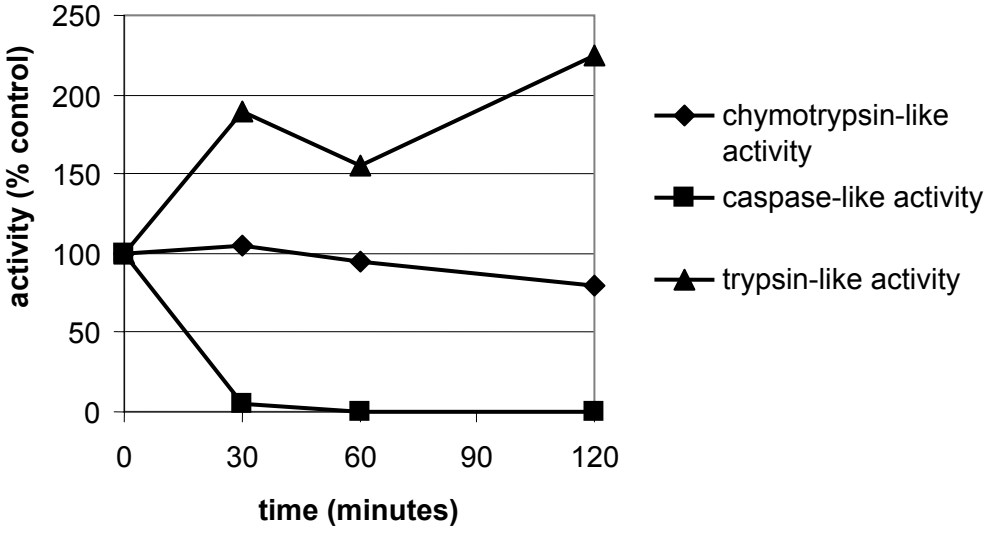


Figure 2. Protease assay with purified proteasome shows specificity of **4** for the caspase-like site of the proteasome. Proteasome was incubated with 50 μM inhibitor and protease activity was measured by the amount of fluorescence generated by degradation of 7-amido-4-methylcoumarin (amc) substrates for the three proteasomal subunits. Suc-LLVY-amc was used for the chymotrypsin-like site, Ac-nLPnLD-amc for the caspase-like site and Ac-RLR-amc for the trypsin-like site.

Next, the effects of compound **4** and **5** on proteasomes in cultured cells were measured. Mouse T-cell lymphoma EL-4 cells, human multiple myelome RMPI8226 cells and human immortalized B-lymphocytes (LG2 clone) were incubated for 16 h with different concentrations of inhibitors **4** and **5**. Cytosol was squeezed out of digitonin-permeabilized cells by centrifugation, and the activity of all three types of active sites was determined by a fluorogenic substrate assay (Figure 3).[15] Results of experiments with **4** in

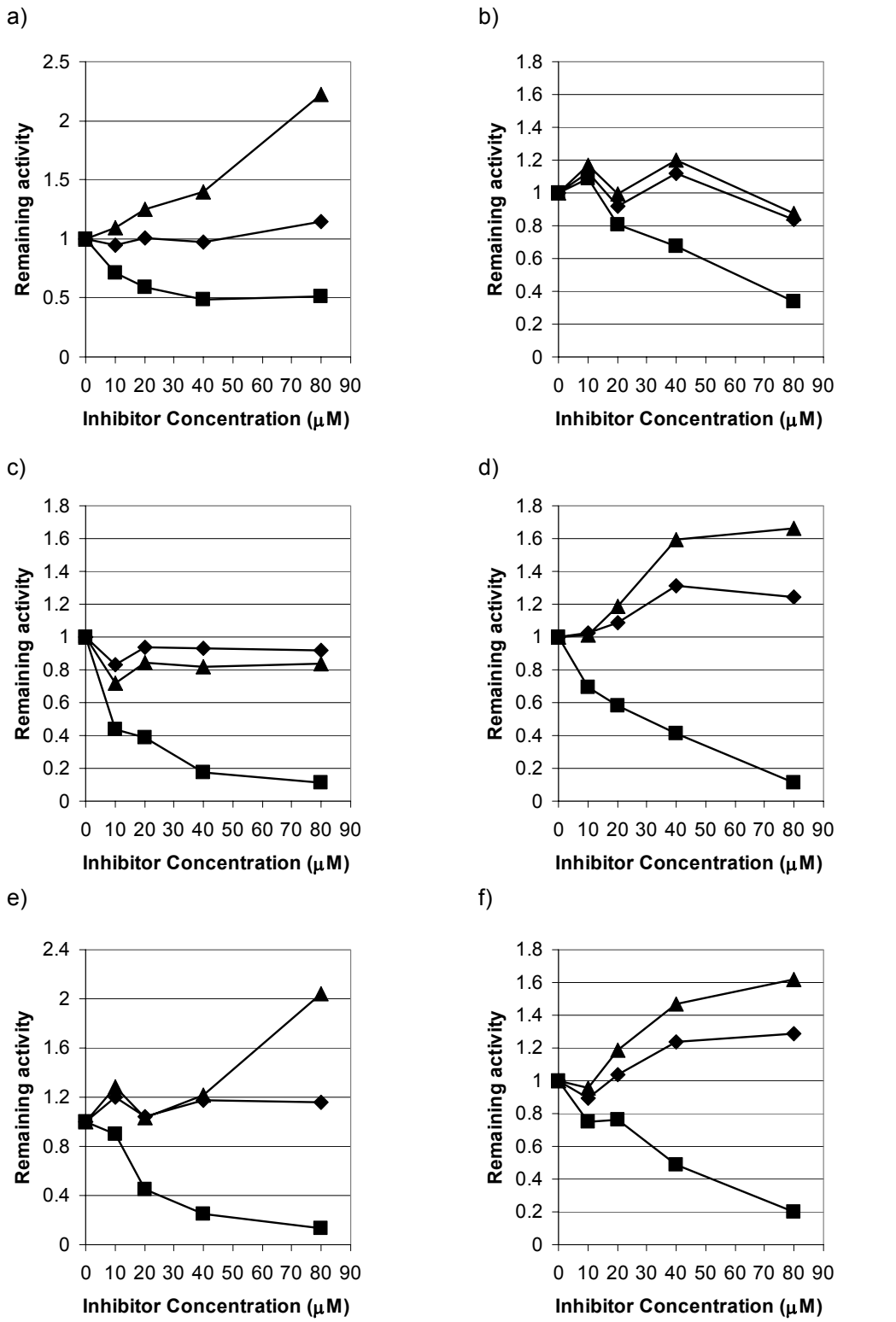


Figure 3. Protease activity measured in cell extracts after treatment with inhibitor **4** (a,c,e) or **5** (b,d,f). EL-4 cells (a,b), RPMI 8226 cells (c,d) and LG2 cells (e,f) were incubated 16-18 h with a concentration range of **4** or **5**. Cells were lysed, and remaining proteasomal activity was addressed by performing a protease assay.

EL-4 cells (Figure 3a) show a plateau at 50% inhibition of the caspase-like activity. As EL-4 cells express both proteasome and immunoproteasome in equal amounts, this could indicate that **4** targets only one of the caspase-like sites. Also, the trypsin-like activity is increased over two-fold, indicating allosteric activation or an increase in proteasome expression due to inhibition of the caspase-like site. In RMPI cells, a plateau in caspase-like activity occurs between 10 and 20 μM **4** (Figure 3c), and at higher concentrations all caspase-like subunits can be blocked. In this cell type, no activation of the trypsin-like site is observed. Inhibitor **5** is less potent in these cells and does not show a plateau (Figure 3d), but some activation of the other subunits is observed. In LG2 cells (Figure 3e and 3f), inhibitor **4** shows a plateau around 10 μM, and at higher concentrations activation of the trypsin-like activity occurs. Azide-containing inhibitor **5** shows a plateau at higher concentrations, and less activation of the other subunits.

Labeling of caspase-like subunits

In order to visualize the subunits that are modified by inhibitor **5**, the azide moieties of labeled proteasomal subunits were modified with a biotin moiety by Staudinger ligation with biotinylated phosphane **12**. In short, cells were incubated for 16 h with different concentrations of **5**, permeabilized with digitonin, and cytosol was squeezed out by centrifugation, followed by extraction of organellar proteins by sonication. Both extracts were treated with phosphane **12** to introduce a biotin moiety, separated on SDS-PAGE and transferred onto PVDF membranes. Biotinylated proteins were visualized using a streptavidin-IRdye 800CW conjugate (Figure 4). In all cell lines, at lower concentration only β1i is modified, whereas at higher concentration of **5**, both caspase-like subunits (β1 and β1i) are targeted. The lack of labeling of the other proteasomal subunits confirms the specificity for **5** for the caspase-like site as found in the fluorogenic peptidase assays described above.

Because the proteasomal subunits are not always resolved on SDS-PAGE, a 2D separation was performed, which is known to separate all catalytic active proteasomal subunits. Cells were incubated with 40 μM and 80 μM of inhibitor **5**. Cytosolic extracts were prepared as described above, incubated with phosphane **12** and separated on a 2D gel (Figure 5). At both concentrations, the caspase-like subunit of the immunoproteasome is labeled exclusively.

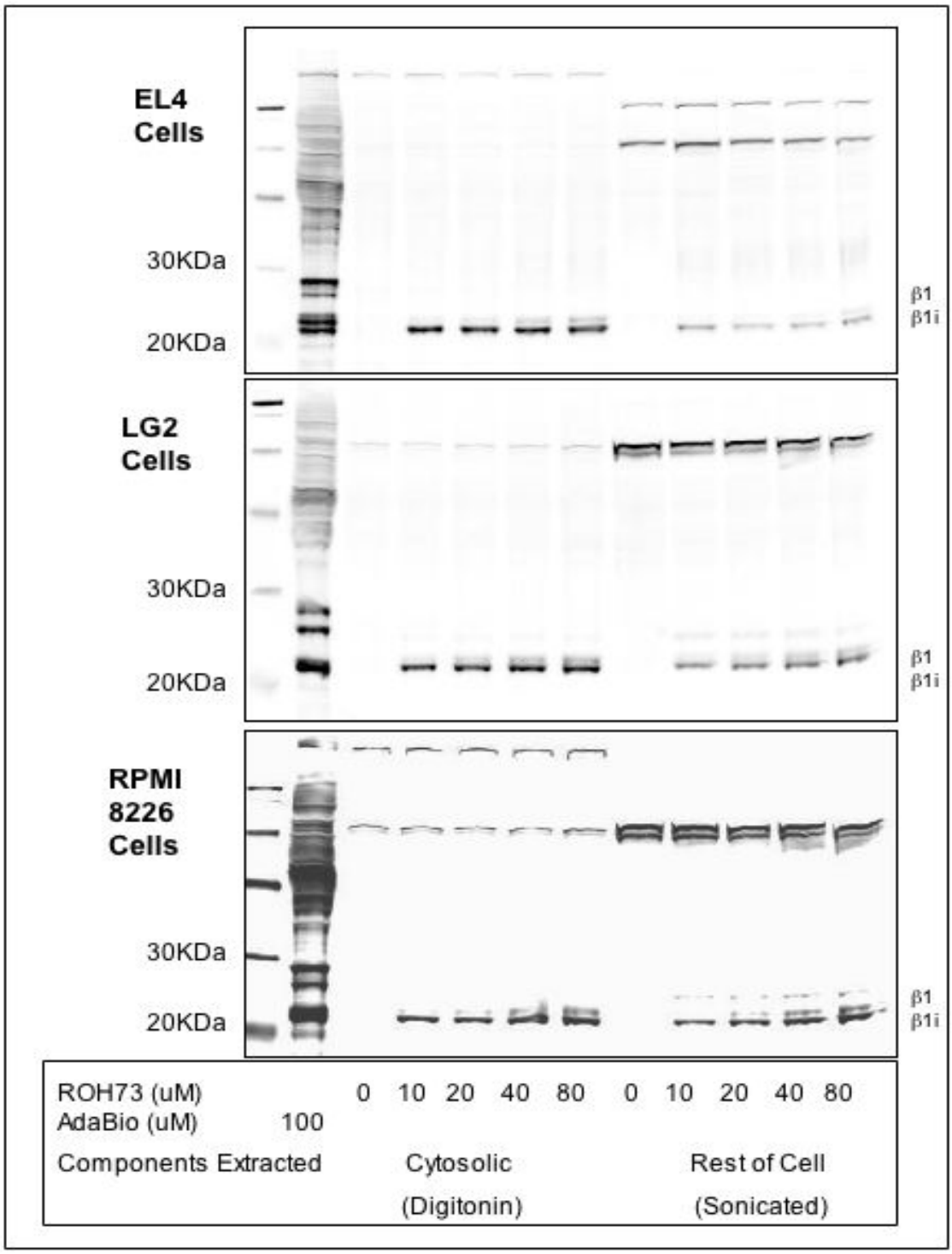


Figure 4. Two-step visualization of targets of compound **5**. EL-4 cells, RPMI 8226 cells and LG2 cells were incubated 16-18 h with different concentrations of **5**. Cells were permeabilized with digitonin, allowing separation of the cytosolic fraction from the organellar fraction. Both were treated with biotinylated phosphane **12**, followed by fractionation on SDS-PAGE and Western blotting. Labeled proteins were visualized by streptavidin-IRdye800CW.

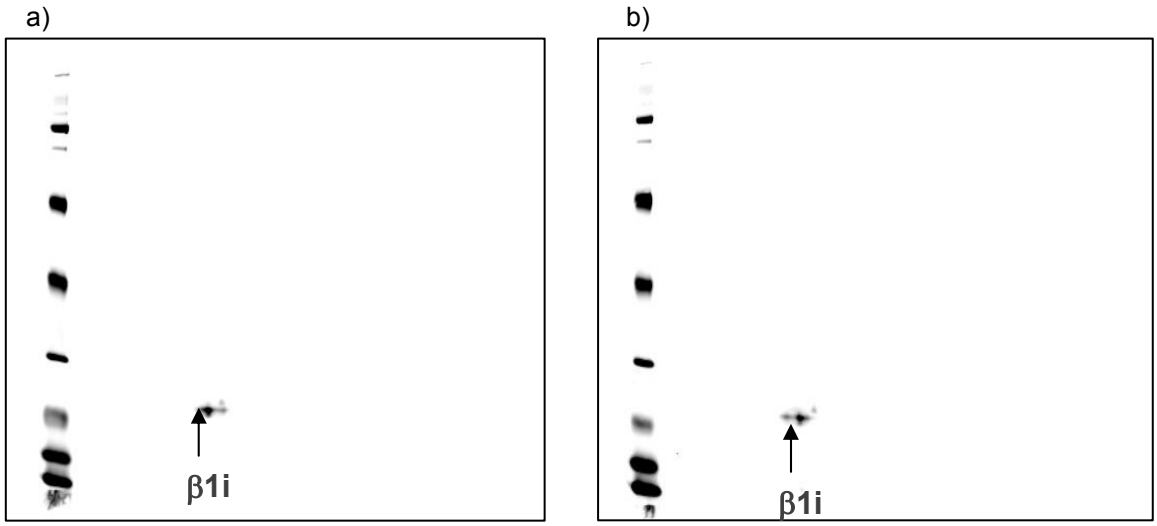


Figure 5. 2D SDS-PAGE of cytosilic extracts of labeled EL-4 cells treated with 10 μM (a) and 80 μM (b) **5**. Extracts were generated and treated as before, and separated by 2D gel electrophoresis. Proteins were stained with streptavidin-IRdye800CW indicating that only the caspase-like activities of the immunoporteasome (β1i) are labeled.

Conclusions

Two new vinyl sulfone based proteasome inhibitors are presented here that are selective inhibitors of the caspase-like subunits of the proteasome. The uniqueness of these reagents is that at lower concentrations they selectively inhibit the caspase-like sites of the immunoproteasomes. The introduction of an azide-moiety did not change the inhibition profile, and confirmed selective targeting of the caspase-like subunits by visualization *via* a Staudinger ligation followed by and SDS-PAGE and Western blotting.

Both inhibitors presented here, as well as the Staudinger ligation protocol applied, might be of value for future research aimed at the role of the caspase-like subunit in the processing of antigens. Furthermore, at lower concentration of inhibitor, the specific contribution of the immunoproteasomal caspase-like site in antigen processing can be studied.

Experimental section

General methods and materials: Solvents used in the solid phase peptide synthesis, DiPEA and TFA were all peptide synthesis grade (Biosolve) and used as received. The protected amino acids, Wang resin (0.86 mmol/g), and HCTU were obtained from NovaBiochem. SPPS was carried out using a 180° Variable Rate Flask Shaker (St. John Associates, Inc.). LC/MS analysis was performed on a Jasco HPLC system (detection simultaneously at 214 and 254 nm) coupled to a Perkin Elmer

Sciex API 165 mass spectrometer equipped with a custom-made Electrospray Interface (ESI). High resolution mass spectra were recorded on a Finnigan LTQ-FT (Thermo electron).An analytical Alltima C_{18} column (Alltech, 4.6 × 250 mm, 5 μm particle size) was used. Buffers: A = H_2O; B = CH_3CN; C = 0.5% aq. TFA. For RP HPLC-purifications a Biocad "Vision" automated HPLC system (PerSeptive Biosystems, Inc.) was used. The applied buffers were A, B and C. ^{1}H-NMR and ^{13}C-NMR spectra were recorded with a Bruker AC200 instrument at 200 and 50.1 MHz, respectively. 600 MHz ^{1}H-NMR spectra were recorded with a Bruker DMX 600 instrument with chemical shifts (δ) relative to tetramethylsilane. TLC analysis was performed on Merck 25DC Plastikfolien Kieselgel 60 F_{254}, with detection by UV absorption (254 nm) and spraying with $KMnO_4$ (10 g/l in 2% aq Na_2CO_3). Fluka silica gel (230-400 mesh) was used for column chromatography.

(*S*)-4-(3-Amino-5-methyl-hex-1-ene-1-sulfonyl)-phenol (8): To a solution of KOtBu (1 g, 8.94 mmol) and (4-hydroxy-benzenesulfonylmethyl) phosphonic acid diethyl ester (1.5 g, 4.87 mmol) in THF was added known Boc-protected leucinal[16] (0.92 g, 4.1 mmol). The mixture was stirred for 16 h, the mixture was diluted with Et_2O, washed with sat. aq. $NaHCO_3$ (3×) and sat. aq. NaCl (3×), dried ($MgSO_4$), filtered and concentrated *in vacuo* yielding 1.3 g of crude Boc protected title compound. The Boc protecting group was subsequently removed by treatment with TFA/H_2O (1 ml, 95/5 v/v) for 1 h. The reaction mixture was poured into Et_2O (50 ml), and the precipitate collected yielding 0.45 g, (1.2 mmol, 29% over two steps) of the title compound, which was used without further purification.

N_3AcAlaProNle-OH: Wang Resin (2.0 g, 1.7 mmol) was condensed with Fmoc-Nle-OH (4.2 g, 12 mmol) in CH_2Cl_2 (30 ml) under the agency of DIC (1.8 ml, 12 mmol) and about a spatula tip of 4-(dimethylamino)pyridine for 16 h. The resin was filtered off and washed (DMF 3×, CH_2Cl_2 3×, Et_2O 3×) and air-dried. Loading was determined by quantification (UV) of Fmoc cleavage, and proved to be 0.77 mmol/g. The resin (130 mg, 100 μmol) was elongated using standard Fmoc-based SPPS to give resin-bound N_3AcAlaProNle-OH. Where appropriate removal of the Fmoc protecting group was accomplished by treatment of the resin-bound peptide with piperidine/DMF (1/4 v/v) for 20 min. Peptide coupling steps were performed by treatment of the resin with a premixed (5 min) solution of the appropriate acid (5 eq.), HCTU (5 eq.) and DiPEA (6 eq.) in NMP (5 ml) for 1 h unless stated otherwise. Coupling efficiencies were monitored with the Kaiser test and couplings were repeated if necessary. After coupling and deprotection steps the resin was washed with DMF (5×). After the last coupling step, the resin was washed extensively (alternating CH_2Cl_2-MeOH 3×, alternating CH_2Cl_2-Et_2O 3×). The resin was treated with TFA/H_2O (2 ml, 95/5 v/v) for 2 h, the resin was filtered into toluene, the mixture was filtered and the resin washed with TFA (2× 1 ml). The

filtrate was diluted (toluene, 10 ml) and concentrated *in vacuo*. The crude product was coevaporated (toluene, 3×).

AcAlaProNleLeuVSOH (4): Crude AcAlaProNle-OH (100 µmol, prepared *via* in the same method as has been described above for N_3AcAlaProNle-OH) was coupled with the TFA-salt of phenolic leucine vinyl sulfone and purified as described for N_3AcAlaProNleLeuVSOH to give 13.6 mg (23 µmol, 45%) of the title compound. ^{1}H NMR (DMSO, 300 K, DMX 600): δ, 10.58 (s, 1H), 8.39 (d, 1H, *J* = 7.4 Hz), 8.36 (d, 1H, *J* = 7.3 Hz), 7.92 (d, 1H, *J* = 5.2 Hz), 7.83-7.80 (m, 2H), 7.60 (d, 2H, *J* = 8.7 Hz), 6.92 (d, 2H, *J* = 8.7 Hz), 6.73-6.48 (m, 2H), 4.59-4.47 (m, 2H), 4.29-4.26 (m, 1H), 4.07-4.02 (m, 1H), 2.03-1.95 (m, 1H), 1.90-1.33 (m, 9H), 1.23-1.10 (m, 14H), 0.85-0.79 (m, 7H). HRMS: $C_{29}H_{44}N_4O_7S + H^+$ requires 593.3004, found 593.2995.

N_3AcAlaProNleLeuVSOH (5): Crude N_3AcAlaProNle-OH (100 µmol), HATU (38 mg, 0.1 mmol) and DiPEA (0.2 ml, 0.12 mmol) were stirred for 5 min in DMF. The TFA-salt of phenolic leucine vinyl sulfone (46 mg, 0.12 mmol) and DiPEA (0.3 ml, 0.18 mmol) were added and the mixture was stirred for 16 h. The reaction mixture was concentrated in vacuo, the residue was dissolved in EtOAc and was washed with 1N HCl, sat. aq. $NaHCO_3$ and sat. aq. NaCl. The organic layer was dried with $MgSO_4$ and concentrated *in vacuo*. The product was purified to homogeneity by RP-HPLC (linear gradient of 40-50% B in 3 column volumes) to yield 7.43 mg (12 µmol, 12%) of the title compound. ^{1}H NMR (DMSO, 300 K, DMX 600): δ, 10.58 (s, 1H), 8.13 (d, 1H, *J* = 7.5 Hz), 7.79 (d, 1H, *J* = 7.6 Hz), 7.77 (d, 1H, *J* = 8.7 Hz), 7.60 (d, 2H, *J* = 8.7 Hz), 6.92 (d, 2H, *J* = 8.7 Hz), 6.69 (dd, 1H, *J* = 15.0 Hz, *J* = 4.9 Hz), 6.55 (dd, 1H, *J* = 15.0 Hz, *J* = 1.3 Hz), 4.56-4.48 (m, 1H), 4.27-4.22 (m, 1H), 4.07-4.02 (m, 1H), 3.61-3.50 (m, 2H), 2.04-1.94 (m, 1H), 1.88-1.72 (m, 4H), 1.63-1.30 (m, 5H), 1.24-1.09 (m, 14H), 0.86-0.77 (m, 7H). HRMS: $C_{29}H_{43}N_7O_7S + H^+$ requires 633.2928, found 633.2935.

***Tert*-butyl (4,7,10-trioxa-13-(biotinylamino)tridecyl) carbamate (10)**: Biotin (488 mg, 2.2 mmol) and HCTU (826 mg, 2 mmol) were dissolved in DMF (5 ml) and DiPEA (0.40 ml, 2.4 mmol) was added, upon which the reaction mixture turned brown. (1-(tert-butyloxycarbonyl-amino)-4,7,10-trioxa-13-tridecanamine) **9** (704 mg, 2.2 mmol) was added slowly and the reaction mixture was stirred and turned red. After 2 h, TLC-analysis ($CHCl_3$/CH_3OH 5/1 v/v) indicated a completed reaction. The mixture was concentrated *in vacuo*, the residue was taken up in CH_2Cl_2 and washed (1N HCl 2×, saturated $NaHCO_3$), dried ($MgSO_4$) and filtered. The filtrate was concentrated *in vacuo* to yield 1.3 g of a brown oil, which was used without further purification.

4,7,10-Trioxa-13-(biotinylamino)tridecylammonium trifluoroacetate (11): Crude **10** (1.3 g, 2.0 mmol) was stirred in TFA (2 ml) with water (1 drop). After 1 h, TLC-analysis ($CHCl_3$/CH_3OH 5/1 v/v) indicated a completed reaction. The mixture was concentrated *in vacuo*, and coevaporated (3×

toluene and 3× methanol). The residue was taken up in 3 ml methanol and poured into Et_2O (200 ml). Slowly, a syrup formed on the surface of the flask. The mother liquor was decanted and the oil (750 mg, 56% over two steps) was collected.

Methyl-2-(diphenylphosphino)-4-(4,7,10-trioxa-13-(biotinylamino)tridecylamino)-carbonyl)benzoic acid (12): Crude **11** (375 mg, 0.56 mmol) was treated with a preactivated solution of 2-(diphenylphosphino)-4-(methoxycarbonyl)benzoic acid (236 mg, 0.65 mmol), HOBt (88 mg, 0.65 mmol), EDC (128 mg, 0.67 mmol) in CH_2Cl_2 under argon atmosphere in the presence of DiPEA (0.1 ml, 0.67 mmol). The mixture was stirred for 16 hour under argon atmosphere and LC/MS analysis indicated a completed reaction. The reaction mixture was diluted (CH_2Cl_2) and washed (1N HCl 2×, saturated $NaHCO_3$ 2×), dried ($MgSO_4$) and filtered. The filtrate was concentrated *in vacuo* and the residue was purified (SiO_2; CH_2Cl_2/CH_3OH 16/1 v/v) to yield a yellow syrup (112 mg, 0.14 mmol, 22%). The product was taken up in CH_3CN/CH_3OH (1/1 v/v) and lyophilised into 5.4 mg aliquots that were stored at -20°C under argon atmosphere. LC/MS-analysis (10-90% B, rt=13.79 min) indicated that the purity of the product was 88%. $[M+H]^+$ = 793.5, $[M+2H]^{2+}$ = 397.3. ^{1}H NMR (200 MHz, CD_3OD): δ 7.95 (dd, *J*=3.3, 8.0, 1H), 7.67 (dd, *J*=1.7, 8.2, 1H), 7.32 (dd, 1.5, 4.0, 1H), 7.3-7.1 (m, 10H), 4.36 (dd, *J*=4.0, 8.0, 1H), 4.16 (dd, 4.4, 7.7, 1H), 3.58 (s, 3H), 3.5-3.0 (m, 18H), 2.80 (dd, *J*=4.7, 12.8, 1H), 2.58 (d, 12.8 Hz, 1H), 2.07 (t, *J*=7.1, 2H), 1.7-1.3 (m, 10H). ^{13}C NMR (50.1 MHz, CD_3OD): δ, 175.8, 169.0, 168.1, 165.9, 142.6, 142.0, 139.0, 138.6, 138.4, 137.8, 135.2, 134.8, 134.4, 131.6, 130.1, 129.8, 129.6, 127.7, 71.4, 71.2, 69.9, 69.8, 63.2, 61.5, 56.9, 52.7, 41.1, 38.7, 36.8, 30.4, 30.2, 29.7, 29.4, 26.8. ^{31}P NMR (50.1 MHz, CD_3OD): δ, -2.72. HRMS: $C_{41}H_{53}N_4O_8PS + H^+$ requires 793.3395, found 793.3384.

Continuous assay of peptidase activities on a fluorescent plate reader: An assay of proteasome activity was performed on a 96-well plate. The ssay mixture was prepared by mixing 100 µl of 2× substrate, 50 µl of buffer (or 4× inhibitor), and 50 µl of 4× enzyme (or extract) to give a final assay volume of 200 µl. Substrates used were Suc-LLVY-amc for chymotrypsin-like sites, Ac-nLPnLD-amc for caspase-like sites and Ac-RLR-amc for trypsin-like sites, all at a final concentration of 100 µM. All substrates were initially prepared as 100× stocks in DMSO. All peptidase assays of the 26S proteasome use the same assay buffer, which consists of 50 mM Tris-HCl, pH 7.5, 40 mM KCl, 5 mM $MgCl_2$, 0.5 mM ATP and 1 mM dithiothreitol (DTT). ATP is needed to prevent dissociation of the 26S proteasome into its components, to ensure maximal activity, and in a way that 26S and not 20S activity is measured. The plate was then placed in a fluorescent plate reader, and the fluorescence of 4-amino-7-methyl-coumarin reaction product was monitored continuously (excitation 380 nm and emission 460 nm were used). The reaction rates were determined from the slopes of the reaction progress curves. In order to account for the contribution of other proteases that may be present in extracts to the cleavage of proteasomal substrates, extracts were preincubated

in the presence or absence of 20 μM epoxomicin for 30 min at 37°C. For each sample, the proteasome activity was calculated by subtracting activity in the epoxomicin-treated extract from the activity in the mock-treated extracts.[14]

Determination of inhibition of proteasomes in cells: To determine the effect of the inhibitors in living cells, cells were incubated at 37°C overnight with varying concentrations of inhibitor. Cells were then washed three times in ice cold 1× PBS by centrifugation at 4°C. The cells were frozen briefly at -80°C. After thawing, the cells were re-suspended in 4 volumes of homogenization buffer (50 mM Tris-HCl, pH 7.5, 250 mM sucrose, 5 mM $MgCl_2$, 2 mM ATP, 1 mM DTT, 0.5 mM EDTA) containing 0.025% digitonin. Cells were then incubated for at least 5 minutes on ice to allow permeabilization by digitonin. The cytosol was then "squeezed out" by centrifugation at 20.000×g for 15 min (4°C). Protein concentration of the extract was determined by Bradford assay. The extract was then assayed as above or treated with 0.5 mM phosphane, and analyzed on a SDS-PAGE (12% Novex NuPAGE Bis-Tris gel with MOPS running buffer) or 2D-gels (using Invitrogen ZOOM system with pH 3-10 isoelectric focusing strips, and a 4-12% gradient NuPAGE Bis-Tris gel with MOPS buffer). In some cases, the remaining cell pellet was re-suspended in the same homogenization buffer and then sonicated to extract the remaining components. The extracts were then analyzed by SDS-PAGE exactly as described for cytosolic extracts.

References and notes

1. The work presented in this chapter was performed in close collaboration with A. Kisselev and E. Samuel at Norris Cotton Cancer Center, Dartmouth Medical School, Lebanon, USA.
2. K. L. Rock and A. L. Goldberg, *Annu. Rev. Immunol.*, 1999, **17**, 739-797.
3. J. W. Yewdell, E. Reits and J. Neefjes, *Nat. Rev. Immunol.*, 2003, **3**, 952-961.
4. B. J. van den Eynde and S. Morel, *Curr. Opin. Immunol.*, 2001, **13**, 147-153.
5. C. Garcia-Echeverria, *Mini Rev. Med. Chem.*, 2002, **2**, 247-259.
6. A. F. Kisselev, M. Garcia-Calvo, H. S. Overkleeft, E. Peterson., M. W. Pennington, H. L. Ploegh, N. A. Thornberry and A. L. Goldberg, *J. Biol. Chem.*, 2003, **278**, 35869-35877.
7. M. Bogyo, S. Shin, J. S. McMaster and H. L. Ploegh, *Chem. Biol.*, 1998, **5**, 307-320.
8. T. P. Dick, A. K. Nussbaum, M. Deeg, W. Heinemeyer, M. Groll, M. Schirle, W. Keilholz, S. Stevanovic, D. H. Wolf, R. Huber, H.-G. Rammensee and H. Schild, *J. Biol. Chem.*, 1998, **273**, 25637-25646.
9. T. A. McCormack, A. A. Cruikshank, L. Grenier, F. D. Melandri, S. L. Nunes, L. Plamondon, R. L. Stein and L. R. Dick, *Biochemistry*, 1998, **37**, 7792-7800.
10. C. Cardozo, C. Michaud and M. Orlowski, *Biochemistry*, 1999, **38**, 97768-9777.

11. B. M. Kessler, D. Tortorella, M. Altun, A. F. Kisselev, E. Fiebinger, B. G. Hekking, H. L. Ploegh and H. S. Overkleeft, *Chem. Biol.*, 2001, **8**, 913-929.
12. J. Myung, K. B. Kim, K. Lindsten. N. P. Dantuma and C. M. Crews, *Mol. Cell*, 2001, **7**, 411-420.
13. E. Saxon and C. R. Bertozzi, *Science*, 2000, **287**, 2007-2010.
14. A. F. Kisselev and A. L Goldberg, *Meth. Enzymol.*, 2005, **398**, 364-378.
15. The use of digitonin for cell lysis has been shown to avoid release of non-proteasomal enzymes that also cleave proteasomal substrates and thus interfere with the assay.[14]
16. D. H. Rich, E. T. Sun and A. S. Boparai, *J. Org. Chem.*, 1978, **43**, 3624-2626.

Chapter 4

Development of an isotope-coded activity based probe for the quantitative profiling of cysteine proteases[1]

Introduction

Proteomics research aims at the study of the expression levels and functioning of (subsets of) the proteins present in a biological sample. Proteomics research presents a number of challenges to the researchers. The protein content in a cell is dynamic and cannot be amplified easily. Further, there is a difference of several orders of magnitude between the most and the least abundant proteins. Traditionally, 2D-gel electrophoresis has been used to separate mixtures of proteins, allowing identification of the proteins, originally using Edman degradation. Nowadays, protein sequencing is normally performed using mass spectrometry techniques, as introduced by Watanabe and coworkers in their groundbreaking report.[2] This mass spectrometry based approach also opened the way to quantitative analyses on protein fragments.

The use of isotope labeled amino acids for relative quantification of protein was introduced by the group of Chait.[3] Isotopic labeling was accomplished by growing cells on ^{15}N enriched growth media. Two cell pools, one isotope labeled and one unlabeled, are mixed, lysed and digested with trypsin. The two isotope forms of each peptide have very similar chromatographic properties. In practical applications, they coelute and therefore can serve as mutual internal standards in the mass spectrometer. The relative intensity of both peptides reflects the relative abundance of the protein in the two cell pools. The amino acid sequence of the peptides, and thereby the identity of the proteins they originated from, was established by tandem ms.

Aebersold and coworkers used an isotope coded affinity tag (ICAT) to introduce an affinity tag, biotin, and an isotope label.[4] The ICAT strategy is based on the presence of eight D (heavy) or H (light) atoms in a biotin-containing cysteine-reactive probe. The light and heavy probe share almost all physical and chemical properties, and only in a mass spectrometer the mass difference is revealed. After labeling a denatured protein sample from two different sources with either the light or heavy ICAT probe, the combined fragments are digested using trypsin. The biotinylated fragments are enriched for, and both the relative quantity and sequence identity of the proteins from which the biotinylated peptides originated are determined by automated multistage mass spectrometry.[5] By selecting only cysteine containing tryptic peptides the sample complexity is reduced while protein quantification and identification are still achieved. ICAT experiments have been used by others to study the influence of perturbations on protein expression levels, to gain new insights in specific cellular processes, and to study protein complexes and post-translational modifications.[6]

Quantitative activity based profiling

In chemical proteomics approaches, a complex biological mixture of proteins is simplified before analysis by labeling a specific set of related proteins with an affinity- or fluorescence tag.[7] For instance, broad-spectrum, irreversible protease inhibitors have been used in the profiling of serine proteases,[8] cysteine proteases,[9] and the catalytically active subunits of the proteasome.[10, 11] A relevant example of a chemical proteomics probe is represented by **1a** (DCG-04, Figure 1), developed by Bogyo and coworkers as an irreversible cysteine protease inhibitor, and applied by Lennon-Duménil *et al.* to monitor the proteolytic activity of maturing phagosomes in live antigen-presenting cells.[12,13]

Compound **1a** consists of three functionalities: 1) an electrophilic epoxysuccinate, that alkylates the active site cysteine residue, 2) a short peptide sequence that allows recognition of the probe by the cathepsin family of cysteine proteases and 3) biotin, for the detection and isolation of the modified enzymes (Figure 1). The biotin is connected to the peptide epoxysuccinate by an aminohexanoic acid residue. It was reasoned that incorporation of an isotopic encoded entity, in analogy to the ICAT reagent,[4] would allow for both quantitative and functional assessment of the cathepsin family of cysteine proteases from complex biological samples.

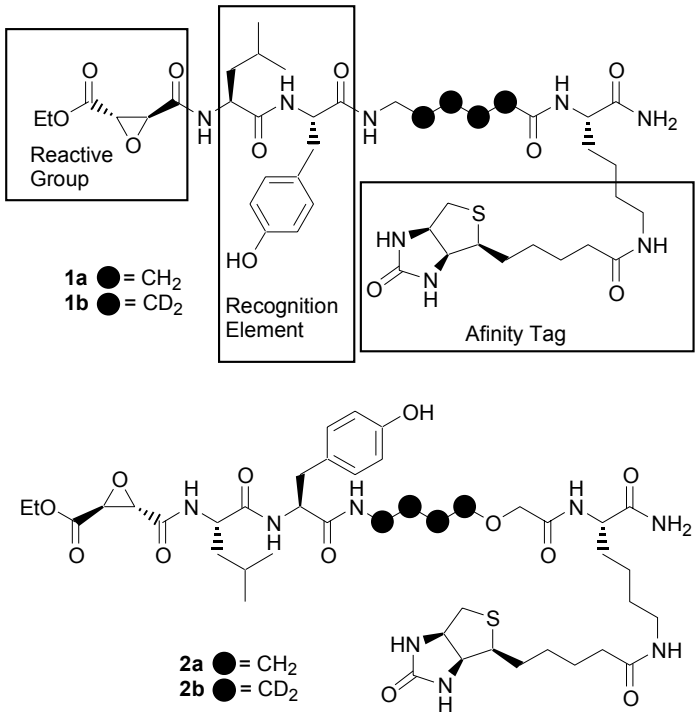
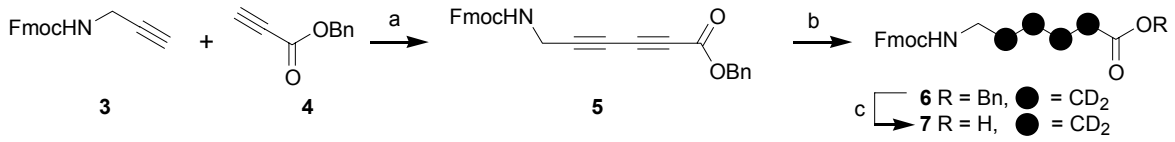


Figure 1. Target Structures

Synthesis

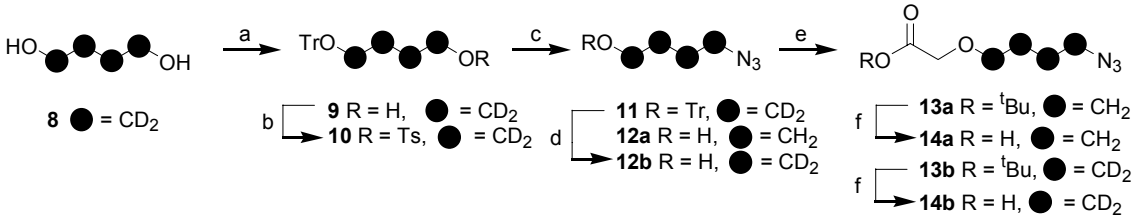
Here, the syntheses of a new azide-protected amino acid spacer, 7-azido-3-oxaheptanoic acid (N_3-Aoh-OH, **14a**) and two d8-enriched amino acids, 2,2,3,3,4,4,5,5 octadeutero-6-(Fmoc-amino)hexanoic acid (Fmoc-Ahx-OH-D8, **7**) and 4,4,5,5,6,6,7,7-octadeutero-7-azido-3-oxaheptanoic acid (N_3-Aoh-OH-D8, **14b**) are described. Their use in the synthesis of two sets of DCG-04-based isotopic encoded activity based probes, namely compound **1b** (in combination with DCG-04 **1a**) and compounds **2a** and **2b** is demonstrated. It is also shown here that neither incorporation of an isotopic label (compounds **1b** and **2b**, Figure 1) nor substitution of the aminohexanoic acid moiety by an 7-amino-3-oxa-heptanoic acid residue (compounds **2ab**) affects the broad spectrum affinity for cysteine proteases of the cathepsin family.



3 4 5

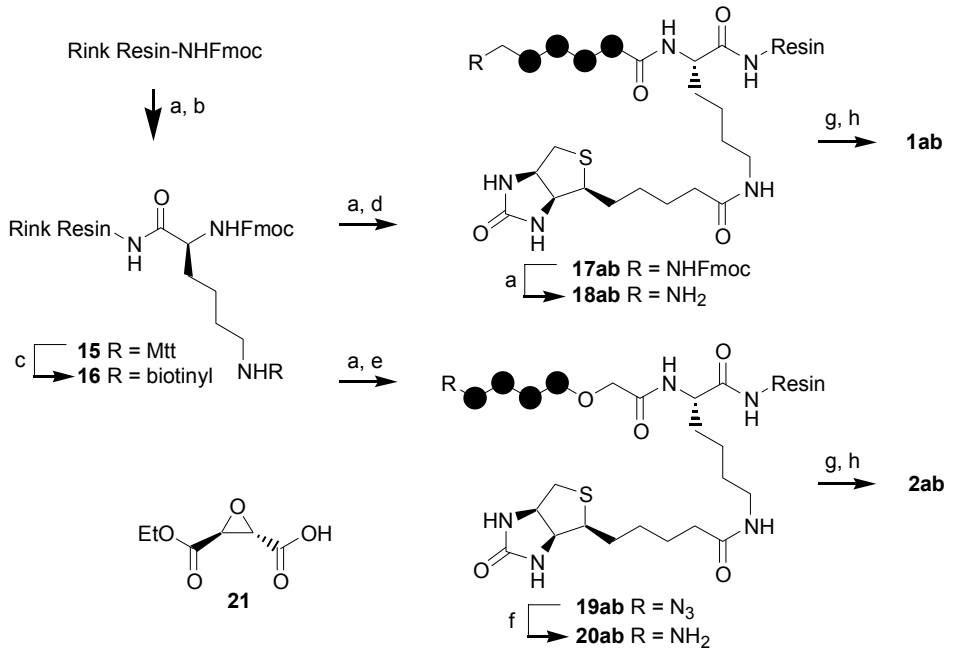
Scheme 1. Synthesis of deuterated Fmoc-Ahx-OH **7**. Reagents and conditions: a) CuCl, *N,N,N′,N′*-tetramethylene ethylenediamine (TMEDA), O_2, THF, 45°C, 36% yield; b) D_2, EtOAc, Pd/C, 84% yield; c) DCl in D_2O, dioxane, 4 h, 90°C, 85% yield.

The synthesis of Fmoc-Ahx-OH-D8 (**7**; Scheme 1) starts with a copper(I) catalyzed oxidative Glaser coupling[14] of Fmoc-protected propargylamine[15] (**3**) and benzyl propiolate[16] (**4**) to furnish diyne **5**. Upon reduction of diyne **5** with deuterium gas and palladium on carbon, the benzyl ester remained intact, yielding deuterated **6**. Contrary to the results of reduction of several analogs of **5**,[17] no isotopic scrambling had taken place as judged from the isotope distribution in the products. Finally, the benzyl ester was hydrolyzed[18] using DCl in deuterium oxide and dioxane to give isotopically labeled Fmoc-Ahx-OH-D8 (**7**).



Scheme 2. Synthesis of azidooxaheptanioc acids **14a** and **14b**. Reagents and conditions: a) TrCl, Et_3N, DMAP, 23% yield, 68% recovered **8**; b) TsCl, Et_3N, 95% yield; c) NaN_3, DMF, 50°C, quantitative; d) TFA, TES, CH_2Cl_2, 83% yield; e) tert-butyl bromoacetate, Bu_4NBr, NaOH, H_2O, toluene, 61% (**13a** and **13b**) yield; f) TFA/CH_2Cl_2 (1/1 v/v), 81% (**14a**) and 87% (**14b**) yield.

The syntheses of N_3-Aoh-OH-D0 (**14a**) and N_3-Aoh-OH-D8 (**14b**) were accomplished as follows (Scheme 2). Upon tritylation of deuterated **8**, monotrityl ether **9** was obtained in 23% yield, while the unreacted deuterated butanediol was recovered easily (68%). Tosylation of the primary alcohol followed by the replacement of the tosylate by azide gave protected azidoalcohol **11**. Detritylation (TFA/TES in CH_2Cl_2) afforded azidoalcohol **12b**. Both known 4-azido-1-butanol[19] **12a** and deuterated **12b** were alkylated under phase transfer conditions with *tert*-butyl bromoacetate to furnish the corresponding *tert*-butyl ester **13a** and **13b**. Subsequent acidolysis of the *tert*-butyl esters employing TFA/CH_2Cl_2 (1/1 v/v) yielded azido acid **14a** in 49% over two steps and **14b** in an overall yield of 42% over the last five steps.



Scheme 3. Synthesis of activity based probes **1ab** and **2ab**. Reagents and conditions: a) 20% piperidine in NMP; b) Fmoc-Lys(Mtt)-OH, HCTU, DiPEA, NMP; c) 1% TFA in CH_2Cl_2, then biotin, HCTU, DiPEA, NMP; d) Fmoc-Ahx-OH or **7**, HCTU, DiPEA, NMP; e) **14a** or **14b**, HCTU, DiPEA, NMP; f) Me_3P, 20% H_2O in dioxane; g) Repeated cycles of SPPS: Fmoc cleavage: 20% piperidine in NMP; amino acid condensation: Fmoc- protected amino acid, HCTU, DiPEA, NMP; Fmoc-protected building blocks were used in the following order: Fmoc-Tyr(tBu)-OH, Fmoc-Leu-OH, **21**; h) 5% H_2O in TFA.

The incorporation of spacers **7** and **14ab** into the respective cysteine protease inhibitors **1b** and **2ab** is shown in Scheme 3. Immobilization of biocytin on a Rink linker was accomplished as described in chapter 3. Standard solid phase peptide synthesis (SPPS) with the sequential addition of a spacer (**7**, **14a** and **14b,** respectively), Fmoc-Tyr(tBu)-OH, Fmoc-Leu-OH and ethyl (2S,3S)oxirane-2,3-dicarboxylate **21**,[20] acidic cleavage from the resin and purification by repeated precipitation afforded target compounds **1b**, and **2a** and **2b** in 34%, 29% and 40%, respectively in 80-90% purity as judged by LC/MS. A small portion of each product was purified to homogeneity by RP-HPLC.

Inhibitory results

To establish the inhibition profile of the newly synthesized probes, a set of labeling experiments with cell lysates of the mouse macrophage cell line J774 was performed. Cell lysates were incubated with DCG-04 (**1a**) as a control and with the new probes **1b**, **2a** and **2b** for 60 minutes at 37°C. The resulting mixtures were separated by SDS-PAGE. The separated proteins were transferred onto a polyvinylidene difluoride (PVDF) membrane,

followed by chemiluminescence induced by horseradish peroxidase–streptavidin conjugate (Figure 2). Probes **2ab** label the cysteine proteases Cat B, L, S, and Z in a cell lysate with the same efficiency as DCG-04, which has been shown previously to effectively target these proteolytic enzymes.[12] This suggests that both sets of isotopic coded activity based probes **1ab** and **2ab** are viable quantitative functional proteomics tools for the cathepsin family of cysteine proteases.

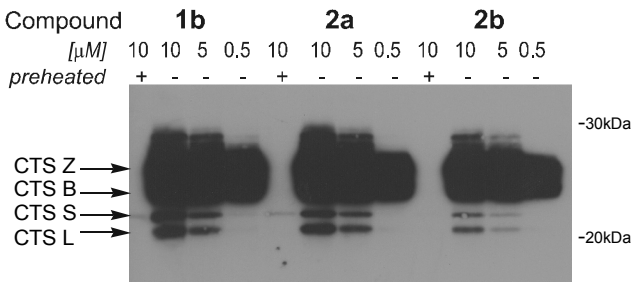


Figure 2. Derivatization of the DCG-04 molecule does not alter targeting of active protease species in J774 cells. Crude extracts prepared from J774 cells were labeled for one hour at 37°C with different concentrations of compounds **1b**, **2a** and **2b**. As a control, extracts were pre-heated for 5 minutes at 100°C prior to labeling (lanes 1, 5 and 9). Proteins were then separated by SDS-PAGE on 12.5% gels and labeled polypeptides visualized by streptavidin blotting. Polypeptide species corresponding to Cat Z, B, S and L are indicated based on previous studies.[10]

Attempts to synthesize labeled Cat B active-site fragment

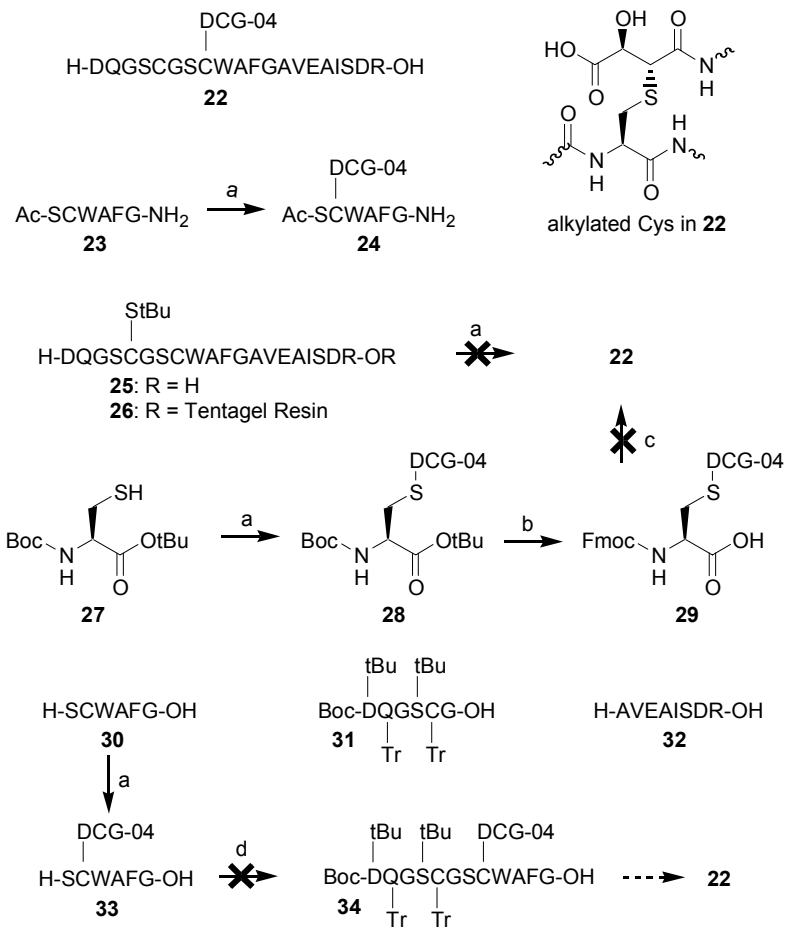
Next an experiment was set up to identify the labeled active site of a cysteine protease. Recombinant Cat B was labeled with epoxide **1a** and treated with trypsin. The labeled catalytic cysteine residue, Cys^{105}, is expected to be in the DQGSCGSC^{105}WAGGAVEAISDR sequence, originating from proteolytic cleavage after Arg^{100} and Arg^{120}. The tryptic digest was analyzed by mass spectrometry, but the labeled active-site fragment could not be found. In order to gain as much information on the expected labeled tryptic fragment, it was decided to embark on the synthesis of this fragment **22** (Scheme 4).

The key step in the synthesis of **22** would be the alkylation of Cys^{105} with epoxide **1a**. To investigate the feasibility of such an alkylation, a series of base catalyzed alkylations was performed on pentapeptide **23**. The use of solvents such as ethanol, DMF, THF and mixtures thereof with DMSO, as well as the use of inorganic bases such as KOH or Na_2CO_3 effected hydrolysis of the ethyl ester present in **1a**. The reaction proved to proceed with the best result using DMSO as solvent and DiPEA as base to afford **24** in moderate efficiency as judged by LC/MS analysis. As only a small amount of **22** was needed for mass spectrometry experiments, it was decided to use these conditions to react peptide **25**, in

which Cys[102] was protected as an tert butyl disulfide, with probe **1a**. However, no product formation could be detected by LC/MS analysis. Subjecting the resin-bound and protected peptide **26**, with only Cys[105] liberated, to the same conditions did not afford target compound **22** either.

As alkylation of Cys[105] within the context of the tryptic fragment proved to be unsuccessful, an attempt was made to incorporate S-alkylated cysteine into the growing peptide chain on the solid support. Reaction of protected cysteine **27** with epoxide **1a** afforded **28**. Removal of the acid labile protecting groups and installation of an Fmoc protecting group yielded **29**, which was purified by recrystallization. This building block proved unsuitable for coupling to the growing tryptic digest peptide chain. Where coupling of a trityl protected cysteine building block to the N-terminal Trp on the resin was difficult and needed double couplings, **29** could not be coupled to the Trp residue at all.

A third route that might lead to **22** was found in a double block coupling strategy. As alkylation of **22** went in moderate yield, it was expected that reaction of **30** with **1a** would afford some **33**, which, after HPLC purification, could be reacted with the succinimidyl ester of protected N-terminal fragment **31**. Activation of **31** *via* the same method and coupling to unprotected C-terminal



Scheme 4. Synthetic routes toward alkylated tryptic digest fragment of Cat B. Reagents and conditions: a) **1a**, DMSO, DiPEA; b) TFA/H_2O (95/5 v/v), then Fmoc-OSu, DiPEA, DMSO; c) Fmoc based SPPS; d) **31**, HOSu, DIC, DMF, then **33**.

fragment **32** would eventually lead to target **22**. However, alkylation of **30** gave product **33** in a marginal 2% yield after HPLC purification. Transformation of **31** to the corresponding succinimidyl ester went uneventful, but subsequent reaction with **33** did not yield any product.

Conclusions

In summary, the efficient synthesis of two pairs of isotopic coded spacers is presented and it is shown that their incorporation into a known cysteine protease inhibitor does not alter the inhibitory profile of the label. This opens the way to quantitative functional proteomics studies on a functional subset of the proteome, namely the cathepsin family of cysteine proteases. One hurdle that remains to be taken is the identification of the labeled tryptic digest fragment of the Cat B active site in a mass spectrometric experiment. Synthesis of this fragment has not been succesfull, but the shorter peptide fragment that was S-alkylated with the cysteine protease probe could be of use as well. Importantly, this concept may be extended towards other isotopic coded spacers (^{13}C, ^{15}N) and activity based probes targeting other proteins.[7]

Experimental section

General methods and materials: All solvents used in synthesis were of p.a. grade and stored over molecular sieves. Pyridine, CH_2Cl_2, toluene, triethylamine, DiPEA, DMF, NMP, DMSO and TFA were bought at Biosolve. Tritylchloride, tosylchloride, 4-(dimethylamino)pyridine, deuterium chloride (20 wt% solution in D_2O) and triethylsilane were purchased from Acros. 1,1,2,2,3,3,4,4-Octadeutero-1,4-butanediol, deuterium (99.8% D) and *tert*-butyl bromoacetate were purchased from Aldrich. Sodium azide was bought at Merck. Fmoc-Leu-OH, Fmoc-Tyr(tBu)-OH and Fmoc-Lys(Mtt)-OH were bought at Senn Chemicals.

TLC analysis was performed on Merck 25DC Plastikfolien Kieselgel 60 F_{254}, with detection by UV absorption (254 nm) and spraying with one of the following solutions: *a* 20% H_2SO_4 in EtOH followed by charring, *b* ammonium molybdate (25 g/l) in 10% H_2SO_4 followed by charring, *c* $KMnO_4$ (10 g/l in 2% aq. Na_2CO_3). Fluka silica gel (230-400 mesh) was used for column chromatography. ^1H-NMR and ^{13}C-NMR spectra were recorded with a Bruker AC200 instrument at 200 and 50.1 MHz, respectively. Chemical shifts (δ) are given in ppm, relative to tetramethylsilane as an internal standard. Mass spectra were recorded with a Perkin Elmer Sciex API 165 mass instrument equipped with a custom-made Electrospray Interface (ESI). LC/MS analysis was performed on a Jasco HPLC system (detection simultaneously at 214 and 254 nm) coupled to a

Perkin Elmer Sciex API 165 mass spectrometer equipped with a custom-made Electrospray Interface (ESI). An analytical Alltima C_{18} column (Alltech, 4.6 × 250 mm, 5 μm particle size) was used. Buffers: A: H_2O; B: CH_3CN; C: 0.5% aq TFA. For RP HPLC-purifications a Biocad "Vision" automated HPLC system (PerSeptive Biosystems, inc.) was used. The applied buffers were A, B and C. High resolution mass spectra were recorded on a Finnigan LTQ-FT (Thermo electron).

(N-Fluorenylmethyloxycarbonyl)-propargylamine (3): The title compound was prepared according to the literature procedure.[15] ^1H and ^{13}C NMR analyses were in agreement with the reported values.

Benzyl propiolate (4): The title compound was prepared according to the literature procedure.[16] ^1H and ^{13}C NMR analyses were in agreement with the reported values.

6-(Fluorenylmethyloxycarbonylamino)-hexa-2,4-diynoic acid benzyl ester (5): In a 100 ml three necked reaction flask were placed, Cu(I)Cl (180 mg, 1.79 mmol), TMEDA (0.260 ml, 1.79 mmol), THF (20 ml) and *N*-Fmoc propargyl amine (1.03 g, 3.72 mmol). Oxygen was bubbled through the suspension and at 30°C a solution of benzyl propiolate (2.58 g, 16.1 mmol) in THF (20 ml) was added dropwise over a 1.5 h period. The reaction was stirred for another 2 h at 40°C while bubbling oxygen. The resulting black mixture was diluted with EtOAc (100 ml) and washed subsequently with 0.1 M HCl (60 ml) and brine (60 ml). After drying ($MgSO_4$), evaporation of the solvent and column chromatography of the black residue (hexanes:EtOAc 95:5 → 9:1 → 3:1), the title compound was isolated as an orange/red solid (590 mg, 36 %). 0.60 g *N*-Fmoc-propargyl amine was recovered. ^1H-NMR ($CDCl_3$); δ 4.12 (d, 2H, *J*=5.1 Hz, CH_2N); 4.21 (t, 1H, *J*=6.6 Hz, CH); 4.45 (d, 2H, *J*=6.6 Hz, CH_2O); 4.98 (broad s, 1H, NH); 5.22 (s, 2H, CH_2O); 7.25-7.44 (m, 9H, arom); 7.56 (d, 2H, *J*=7.3 Hz, arom); 7.76 (d, 2H, *J*=7.3 Hz, arom). ^{13}C-NMR ($CDCl_3$); δ 31.2, 46.8, 65.8, 66.9, 67.3, 67.9, 68.1, 70.7, 119.8, 124.8, 126.9, 127.6, 128.3, 128.5, 134.2, 141.0, 143.4, 152.2, 155.7.

6-(Fluorenylmethyloxycarbonylamino)–2,2,3,3,4,4,5,5-octadeuterohexanoic acid benzyl ester (6): Benzyl ester **5** (307 mg, 0.705 mmol) was dissolved in a mixture of dry 1,4-dioxane (50 ml) and D_2O (2 ml). The solvents were evaporated *in vacuo* and the procedure was repeated once. Then the residue was dissolved in EtOAc (12 ml) and Pd/C-10% (14 mg) was added. The mixture was put in a Parr apparatus and shaken under D_2 (2.5 bar) for 4 h. The solution was filtered, concentrated and purified by column chromatography (hexanes:EtOAc 9:1 → 3:1), to afford the title compound as a pale yellow solid (267 mg, 84 %). ^1H-NMR ($CDCl_3$); δ 3.15 (d, 2H, *J*=5.8 Hz, CH_2N); 4.21 (t, 1H, *J*=6.6 Hz, CH); 4.40 (d, 2H, *J*=6.6 Hz, CH_2O); 4.74 (m, 1H, NH); 5.11 (s, 2H, CH_2O); 7.25-7.39 (m, 9H, arom); 7.58 (d, 2H, *J*=7.3 Hz, arom); 7.76 (d, 2H, *J*=7.3 Hz, arom). ^{13}C-NMR ($CDCl_3$); δ 20.1-21.7 (m), 22.0-23.1 (m), 25.3-26.4 (m), 30.0-31.4 (m), 38.1, 44.8, 63.6, 63.9,

117.4, 122.5, 124.5, 125.1, 125.7, 126.0, 133.5, 138.8, 141.5, 154.0, 170.9. ESI-MS *m/z* 452.1 $[M+H]^+$, 474.2 $[M+Na]^+$, 903.5 $[2M+H]^+$, 925.6 $[2M+Na]^+$.

6-(Fluorenylmethyloxycarbonylamino)–2,2,3,3,4,4,5,5-octadeuterohexanoic acid (7): Under argon benzyl ester **6** (205 mg, 0.455 mmol) was dissolved in dry 1,4-dioxane (8 ml) and DCl/D_2O (0.30 ml, 1/4 w/w)) was added. The reaction mixture was stirred at 90°C and after 4 hours TLC showed complete conversion of the benzyl ester. The mixture was cooled and the solvents were evaporated *in vacuo*. Purification by column chromatography (hexanes:EtOAc 9:1 → 3:1 → 1:1 +0.2% HOAc), afforded the target compound as a white solid (140 mg, 85 %). ^{1}H-NMR (CD_3OD); δ 2.99 (s, 2H, CH_2N); 4.10 (t, 1H, *J*=6.7 Hz, CH); 4.26 (d, 2H, *J*=6.7 Hz, CH_2O); 7.17-7.33 (m, 4H, arom); 7.55 (d, 2H, *J*=7.3 Hz, arom); 7.70 (d, 2H, *J*=7.3 Hz, arom). ^{13}C-NMR (CD_3OD); δ 23.7-24.9 (m), 25.3-26.7 (m), 28.7-30.1 (m), 33.2-34.7 (m), 41.3, 48.3, 67.4, 120.8, 126.0, 128.0, 128.6, 142.4, 145.2, 158.7, 177.6. ESI-MS *m/z* 362.2 $[M+H]^+$, 723.6 $[2M+H]^+$. HRMS $[M+H]^+$ calculated: 362.220, found: 362.220.

1,1,2,2,3,3,4,4-Octadeutero-4-trityloxybutan-1-ol (9): A mixture of 1,1,2,2,3,3,4,4-octadeutero-1,4-butanediol (2.0 g, 20.4 mmol), trityl chloride (1.9 g, 6.8 mmol), triethylamine (0.95 ml, 6.8 mmol) and about a spatula tip of 4-(dimethylamino)pyridine was stirred until TLC-analysis indicated consumption of tritylchloride. EtOAc (10 ml) was added and the mixture was filtered and concentrated in vacuo. Purification over silica gel (hexanes:EtOAc 6:1→1:1, then EtOAc:MeOH 9:1) yielded 1.57 g (4.6 mmol, 68% yield to TrCl, 23% conversion of diol) of the title compound. ^{1}H NMR ($CDCl_3$): δ, 7.35-7.15 (m, 15H, H_{arom}), 1.8 (br. s. 1H OH); ^{13}C NMR ($CDCl_3$): δ, 143.9 ($C_{q,arom}$) 128.1, 127.2, 126.3 (C_{arom}), 86.0 (C_q Tr), 63-59 (m CD_2O 2x) 29-27 (m, CD_2), 26-24 (m, CD_2). HRMS $[M+Na]^+$ calculated: 363.217, found: 363.218. 1.36 g (13.8 mmol, 68%) of the deuterated diol was recovered. Overall, 91% of the deuterated starting material was converted or recovered.

1,1,2,2,3,3,4,4-Octadeutero-4-trityloxybutanol tosylate (10): To a solution of **9** (1.57 g, 4.6 mmol) in CH_2Cl_2 were added triethylamine (1.28 ml, 9.2 mmol, 2 eq.), a catalytic amount of 4-(dimethylamino)pyridine and tosyl chloride (1.31 g, 6.9 mmol, 1.5 eq.) and the mixture was stirred for 3 h, after which TLC-analysis indicated a completed reaction. Excess tosyl chloride was quenched by addition of MeOH (5 ml) and the mixture was concentrated *in vacuo*. The residue was taken up in EtOAc, filtered and the filtrate was concentrated *in vacuo*. The crude product was purified over silica gel (hexanes:EtOAc 9:1) to yield 2.16 g (4.4 mmol, 95%) of a white solid. ^{1}H NMR ($CDCl_3$): δ, 7.74 (d, 2H, *J*=9.7 Hz, H_{arom}), 7.4-7.1 (m, 17H, H_{arom}), 2.42 (s, 3H, Me); ^{13}C NMR ($CDCl_3$): δ, 144.4, 143.9, 132.9 ($C_{q,arom}$), 129.6, 128.3, 127.8, 127.5, 126.7 (C_{arom}), 86.1 (C_q Tr), 21.3 (Me). HRMS $[M+Na]^+$ calculated: 517.229, found: 517.227.

1,1,2,2,3,3,4,4-Octadeutero-4-trityloxybutane azide (11): To a solution of **10** (2.2 g, 4.4 mmol) in DMF was added sodium azide (426 mg, 6.55 mmol, 1.5 eq.) and the mixture was stirred at 50°C until TLC analysis indicated consumption of tosylate **2**. The reaction mixture was concentrated *in vacuo*, taken up in Et_2O and washed (H_2O, 3×). The combined water layers were extracted with Et_2O. The combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo* to yield 1.68 g (quant.) of a colourless oil. ^{1}H NMR ($CDCl_3$): δ, 7.5-7.1 (m, 15H, H_{arom}); ^{13}C NMR ($CDCl_3$): δ, 143.8 ($C_{q,arom}$) 128.0, 127.2, 126.4 (C_{arom}), 85.9 (C_q Tr), 62-60 (m, CD_2O), 51-49 (m, CD_2N_3), 26-24 (m, CD_2 2×).

1,1,2,2,3,3,4,4-Octadeutero-4-azidobutan-1-ol (12b): To a stirred solution of **11** (1.6 g, 4.37 mmol) and triethylsilane (1 ml, 6.5 mmol, 1.5 eq.) in CH_2Cl_2 (2.5 ml) was added TFA dropwise, until no more yellow colour of trityl cations was observed, and the mixture was stirred for 15 min. The mixture was concentrated *in vacuo*, and purified over silica gel (hexanes:EtOAc 9:1, 4cv, then hexanes:EtOAc 1:1, 4cv) to yield 445 mg (3.6 mmol, 83%) of a yellow oil. ^{1}H NMR (200 MHz, $CDCl_3$): δ, 5.11 (br s, 1H, OH); ^{13}C NMR ($CDCl_3$): δ, 61-59 (m, CD_2OH), 51-49 (m, CD_2N_3), 29-27 (m, CD_2), 25-23 (m, CD_2).

7-Azido-3-oxoheptanoic acid tert-butyl ester (13a): To a solution of known[21] 4-azido-butane-1-ol (1.7 g, 15 mmol), *tert*-butyl bromoacetate (4.2 ml, 30 mmol, 2 eq.), tetrabutylammonium bromide (490 mg, 3 mmol, 10 mol%) and a catalytic amount of tetrabutylammonium iodide in toluene (60 ml) was added aq. NaOH (60 ml, 50% by weight) and the mixture was stirred vigorously at ambient temperature until TLC-analysis indicated consumption of **12a**. Water (600 ml) was added to the mixture, and stirring was continued for 1 h. The organic layer was separated and the aq. layer extracted with toluene (3×). The combined organic layers were washed (sat. aq. NH_4Cl, sat. aq. $NaHCO_3$, brine), dried over $MgSO_4$ and concentrated *in vacuo*. Purification of the crude product over silica gel (hexanes:EtOAc 9:1) yielded 2.1 g (9.1 mmol, 61%) of a yellow liquid. ^{1}H NMR ($CDCl_3$): δ, 3.95 (s 2H, CH_2C=O), 3.6-3.5 (m, 2H, CH_2O), 3.25-3.15 (m, 2H, CH_2N_3), 1.8-1.65 (m, 4H, CH_2 2x), 1.48 (s, 9H, tBu); ^{13}C NMR ($CDCl_3$): δ, 168.7 (C=O), 80.3 (C_q tBu), 69.8 (CH_2O), 67.8 (CH_2C=O), 50.3 (CH_2N_3), 26.0, 24.9 (m, CH_2). HRMS $[M+H]^+$ calculated: 230.150, found: 230.150.

4,4,5,5,6,6,7,7-octadeutero-7-azido-3-oxaheptanoic acid tert-butyl ester (13b): To a solution of **12b** (445 mg, 3.6 mmol), *tert*-butyl bromoacetate (1.06 ml, 7.2 mmol, 2 eq.), tetrabutylammonium bromide (119 mg, 0.36 mmol, 10 mol%) and a catalytic amount of tetrabutylammonium iodide in toluene (15 ml) was added aq. NaOH (15 ml, 50% by weight) and the mixture was stirred vigorously at ambient temperature until TLC-analysis indicated consumption of **12b**. Water (150 ml) was added to the mixture, and stirring was continued for 1 h. The organic layer was separated

and the aq. layer extracted with toluene (3×). The combined organic layers were washed (sat. aq. NH_4Cl, sat. aq. $NaHCO_3$, brine), dried over $MgSO_4$ and concentrated *in vacuo*. Purification of the crude product over silica gel (hexanes:EtOAc 9:1) yielded 530 mg (2.2 mmol, 61%) as a colourless oil. ^1H NMR ($CDCl_3$): δ, 3.94 (s 2H, CH_2), 1.48 (s, 9H, tBu); ^{13}C NMR ($CDCl_3$): δ, 169.1 (C=O), 80.7 (C_q tBu), 70-68 (m CD_2O), 68.0 (CH_2), 50-48 (m CD_2N_3), 26-23 (m, CD_2 2×).

7-azido-3-oxoheptanoic acid (14a): *tert*-Butyl ester **13a** (1 g, 4.3 mmol) was treated with TFA/CH_2Cl_2 (20 ml, 1/1 v/v) for 2 h at ambient temperature. Toluene (20 ml) was added, and the mixture was concentrated to half its volume. Toluene (20 ml) was added again and the mixture was coevaporated with toluene 3× to remove residual TFA. The resulting yellow oil was dissolved in EtOAc and extracted with sat. aq. $NaHCO_3$ (3×). The organic layer was discarded, the aqueous layer acidified with conc. HCl to pH 1 and extracted with Et_2O (6×). The combined Et_2O layers were dried ($MgSO_4$), filtered and concentrated *in vacuo* to yield 610 mg (3.5 mmol, 81%) of a yellowish oil. ^1H NMR ($CDCl_3$): δ, 10.68 (br s, 1H, COOH), 4.15 (s 2H, CH_2C=O), 3.65-3.55 (m, 2H, CH_2O), 3.35-3.25 (m, 2H, CH_2N_3), 1.8-1.6 (m, 4H, CH_2 2x); ^{13}C NMR ($CDCl_3$): δ, 175.1 (C=O), 70.8 (CH_2O), 67.1 (CH_2C=O), 50.6 (CH_2N_3), 26.1, 25.0 (m, CH_2). HRMS $[M+H]^+$ calculated: 174.087 found: 174.088.

4,4,5,5,6,6,7,7-octadeutero-7-azido-3-oxaheptanoic acid (14b): *tert*-Butyl ester **13b** (530 mg, 2.2 mmol) was treated with TFA/CH_2Cl_2 (10 ml, 1/1 v/v) for 2 h at ambient temperature. Toluene (10 ml) was added, and the mixture was concentrated to half its volume. Toluene (10 ml) was added again and the mixture was coevaporated with toluene 3× to remove residual TFA. The resulting yellow oil was dissolved in EtOAc and extracted with sat. aq. $NaHCO_3$ (3×). The organic layer was discarded, the aqueous layer acidified with conc. HCl to pH 1 and extracted with Et_2O (6×). The combined Et_2O layers were dried ($MgSO_4$), filtered and concentrated *in vacuo* to yield 346 mg (1.9 mmol, 87%) of a yellowish oil. ^1H NMR ($CDCl_3$): δ, 9.86 (br s, 1H, COOH), 4.14 (s 2H, CH_2); ^{13}C NMR ($CDCl_3$): δ, 175.0 (C=O), 71-69 (m CD_2O), 66.9 (CH_2), 51-49 (m CD_2N_3), 26-23 (m, CD_2 2×). ESI-MS:[M+Na]=204.0.

Ethyl (2S,3S)oxirane-2,3-dicarboxylate (21): The title compound was prepared according to the literature procedure.[20] ^1H and ^{13}C analysis were in agreement with the reported value. α_D^{22} +105 (*c* 1.2, Et_2O). As a control, Ethyl (2R, 3R)oxirane-2,3-dicarboxylate was prepared following the same procedure. α_D^{22} -114 (*c* 1.2, Et_2O).

Synthesis of the novel DCG-04 analogues: Fmoc Rink amide resin (78 mg, 50 μmol) was elongated by standard Fmoc-based SPPS. Where appropriate removal of the Fmoc protecting group was accomplished by treatment of the resin-bound peptide with 20% (v/v) piperidine in NMP for 20 minutes. Peptide coupling steps were performed by treatment of the resin with a premixed (5 min)

solution of the of the appropriate acid (5 eq.), HCTU (5 eq.) and DiPEA (6 eq.) in NMP (0.5 ml) for 1 h unless stated otherwise. Coupling efficiencies were monitored with the Kaiser test and couplings were repeated if necessary. After coupling and deprotecting steps the resin was washed with DMF (5×). Lysine side chain Mtt protecting groups were removed by treatment of the resin with 1% TFA in CH_2Cl_2 (4×10 min) The resin was washed (CH_2Cl_2 5×, then NMP 2×) and neutralized with 10% DiPEA in NMP. Azides were reduced to amines on solid phase by treatment of the resin, after washing (dioxane 5×), with Me_3P (0.3 ml 1M in toluene, 0.3 mmol, 6 eq.) in dioxane/H_2O (4/1 v/v, 1 ml) for 40 min, followed by extensive washing of the resin (dioxane 3×, alternating CH_2Cl_2-MeOH 3×, and NMP 3×). Deuterated building blocks **7** and **14b** were coupled using 2 eq. of the appropriate acid. After coupling of known **21** to the immobilised peptide, the resin was washed extensively (alternating CH_2Cl_2-MeOH 3×, alternating CH_2Cl_2-Et_2O 3×. The resin was transferred into a clean vial, washed with CH_2Cl_2 and treated with TFA/H_2O (1 ml, 95/5 v/v) for 2 h. The mixture was filtered and the resin washed with TFA (2× 1 ml). The filtrate was diluted (toluene, 10 ml) and concentrated *in vacuo*. To the residue, toluene was added and the concentration was repeated. The residue was taken up in MeOH pored into Et_2O. The precipitate was filtered off and washed with Et_2O. This precipitation was repeated to give the DCG-04 analogues in 29-40% yield. A portion of every analogue was purified to homogeneity by RP-HPLC, applying a linear gradient (20-40% B in 3 column volumes). **1b** was isolated in 34% yield. HRMS $[M+H]^+$ calculated: 933.497, found: 933.505. **2a** was isolated in 29% yield. HRMS $[M+H]^+$ calculated: 941.441, found: 941.449. **2b** was isolated in 40% yield. HRMS $[M+H]^+$ calculated: 949.492, found: 949.499.

Peptide 24: Fmoc Rink amide resin (78 mg, 50 µmol) was elongated by manual solid phase peptide synthesis to give Ac-SCWAFG-NH_2 after cleavage from the resin. **23** was obtained as indicated above. 9 mg (2.5 µmol) of the crude product and **1a** (6 mg, 6 µmol) were dissolved in DMSO saturated with argon and DiPEA (5 µl) was added. The reaction mixture was stirred under argon atmosphere, and after 3 h, LC/MS analysis indicated formation of target **24**, as well as saponified **1a**, saponified **24** and the disulfide of **24**. Longer reaction times resulted in more saponifiaction, whereas shorter reaction times resulted in less product formation as well. As judged by LC/MS, the yield of the reaction was estimated around 30%.

Peptide 25: Tentagel HMB resin preloaded with Fmoc-Arg (50 µmol) was elongated using automated standard Fmoc-based SPPS, except for the two difficult couplings after the Trp residue, where double couplings of 2 h were applied. The peptide was cleaved from the resin with TFA/H_2O/TIS (95/2.5/2.5 v/v/v), purified by RP-HPLC and reacted with **1a** as described for compound **24**. LC/MS analysis gave no indication for formation of **22**. Also, removal of the Mmt protecting group of the target cysteine residue on the solid support followed by on-resin alkylation with **1a** did not result in productive alkylation.

Boc-Cys(DCG-04)-OtBu (28): To a solution of Boc-Cys-OtBu (**26**) (28 mg, 0.1 mmol), **1a** (90 mg, 0.1 mmol) in DMSO (2 ml) saturated with argon, DiPEA (68 µl, 0.4 µmol) was added and the mixture was stirred for 3 h. The reaction mixture was poured into Et_2O/EtOAc (30 ml, 1/1 v/v), cooled (0°C) and agitated to aid precipitation. The white precipitate was collected by centrifugation and washed (Et_2O) to yield 111 mg (94 µmol, 94%) of a white powder, with a purity of 90% as judged from LC/MS. $[M+H]^+$ calculated for $C_{55}H_{89}N_9O_{15}S_2$: 1180.5, found: 1180.9.

Fmoc-Cys(DCG-04)-OH (29): Crude **28** (94 µmol) was treated with TFA/H_2O (1 ml, 95/5 v/v) for 3 h. The product was precipitated with cold Et_2O (35), aided by vortexing. The white precipitate was collected by centrifugation and washed (Et_2O), and redissolved in DMSO (1 ml) with Fmoc-OSu (95 mg) and DiPEA (124 µl). After 3 h of stirring, the mixture was poured into Et_2O (30 ml) and left for precipitation for 16 h at 4°C. The precipitate was collected by centrifugation and washed (Et_2O), to yield 78 mg (47 µmol, 50%) of the target compound. LC/MS $[M+H]^+$ calculated for $C_{61}H_{83}N_9O_{15}S_2$: 1245.6, found: 1245.5. Attempts to couple this alkylated cysteine to the Trp residue of growing peptide chain WAFGAVEAISDR remained fruitless.

Boc-D(tBu)Q(Tr)GS(tBu)C(Tr)G-OH (31): Peptide **31** was synthesized on hyper acid labile Fmoc-Gly-HMPB tentagel resin. The protected peptide was cleaved from the resin by batchwise treatment with 1% TFA in CH_2Cl_2 (20 × 1 ml) for 0.5 min and immediate neutralization of the acidic batches in MeOH/pyridine (10 ml, 9/1 v/v). The solution was concentrated *in vacuo* and **31** was purified to homogeneity by RP-HPLC on a semi-preperative Alltima CN column (Alltech, 10 × 250 mm, 5 µm particle size) and the A-buffer replaced by H_2O/MeOH (1/1 v/v) to yield 5 mg (4 µmol, 8%) of the target compound. LC/MS $[M+H]^+$ calculated for $C_{70}H_{83}N_7O_{13}S$: 1262.6, found: 1262.7.

H-SC(DCG-04)WAFG-OH (33): Peptide **30** was synthesized on 50 µmol resin using automated standard Fmoc-based SPPS and cleavage of the product from resin by treatment with TFA/H_2O/TIS (95/2.5/2.5 v/v/v). Peptide **30** was precipitated (Et_2O) and used without any further purification. LC/MS $[M+H]^+$ calculated for $C_{31}H_{39}N_7O_8S$: 670.2, found: 670.4. **30** (50 µmol) was dissolved in DMSO (1.5 ml) and **1a** (54 mg, 60 µmol) was added. The mixture was saturated with argon before DiPEA (250 µmol, 40 µl) was added, and the reaction mixture was stirred under an argon atmosphere for 16 h. LC/MS analysis of the reaction mixture indicated formation of **33**, as well as peptide disulfide formation. **33** was purified to homogeneity by RP-HPLC to yield 1.5 mg (1 µmol, 2%) of the target compound. LC/MS $[M+H]^+$ calculated for $C_{74}H_{105}N_{15}O_{19}S_2$: 1572.7, found: 1572.9.

34: Activation of **31** (4 µmol) with DIC (1 µl, 4 µmol) and *N*-hydroxysuccinimide (1.2 mg, 10 µmol) in DMF (0.5 ml) was confirmed by LC/MS analysis after 2 h. LC/MS $[M+H]^+$ calculated for

$C_{74}H_{86}N_8O_{15}S$: 1359.6, found: 1359.9. The activated ester solution and DiPEA (4 μmol) were added to **33** (1 μmol) and the mixture was stirred for 16 h. LC/MS analysis did not reveal the presence for target **34**.

Active site labeling of cysteine proteases in J774 cell lysates and detection by streptavidin-blotting: The mouse macrophage cell line J774 was cultured in Dulbecco's Modified Eagle's medium containing 10% fetal calf serum (FCS), 1% penicillin/streptomycine and 1% glutamine. Cells were harvested by centrifugation at 4°C, washed in phosphate buffered saline (PBS) pH 6.8 and cell pellets corresponding to 5×10^7 cells were frozen at –80°C. Cell pellets were thawed on ice and lysed in 100 μl lysis buffer pH 5 (50 mM NaOAc pH 5, 5 mM $MgCl_2$, 0.5% Nonidet-P40), incubated for 30 minutes and centrifuged for 15 min at 13.000 × g to remove nuclei. Protein concentration of the cell extract was measured using the BioRad Protein Assay with bovine serum albumin as standard. Cell lysate (25 μg protein) was incubated with DCG-04, **1b**, **2a** and **2b** respectively for 60 min at 37°C. The reaction was stopped by addition of double concentrated, reducing Laemmli sample buffer and boiling for 10 min. Samples were analyzed by SDS-PAGE and transferred to a polyvinylidene membrane (Immobilon P; Millipore). After blocking over-night in phosphate-buffered saline (PBS) pH 7.2 containing 10% non-fat dry milk, the membrane was incubated with strept-HRP (1:2500; Amersham) in PBS containing 0.2% Tween 20 for 60 min at room temperature followed by extensive washing in PBS-Tween. Polypeptide species reactive to the compounds were detected by Enhanced Chemiluminescence (ECL NEN; Perkin Elmer).

References and notes

1. Part of this chapter was published: P. F. van Swieten, R. Maehr, A. M. C. H. van den Nieuwendijk, B. M. Kessler, M. Reich, C.-S. Wong, H. Kalbacher, M. A. Leeuwenburgh, C. Driessen, G. A. van der Marel, H. L. Ploegh and H. S. Overkleeft, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 3131-3134.
2. W. J. Henzel, T. M. Billeci, J. T. Stults, S. C. Wong, C. Grimley and C. Watanabe, *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 5011-5015.
3. Y. Oda, K. Huang, F. R. Cross, D. Cowburn and B. T. Chait, *Proc. Natl. Acad. Sci. USA*, 1999, **96**, 6591-6596.
4. S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb and R. Aebersold, *Nat. Biotechnol.*, 1999, **17**, 994-999.
5. S. P. Gygi, B. Rist, T. J. Griffin, J. Eng, and R. Aebersold, *J. Prot. Res.*, 2002, **1**, 47-54.
6. W. A. Tao and R. Aebersold, *Curr. Opin. Biotechnol.*, 2003, **14**, 110-118; H. Zhang, W. Yan and R. Aebersold, *Curr. Op. Chem. Biol.*, 2004, **8**, 66-75.

7. A. E. Speers and B. F. Cravatt, *ChemBioChem*, 2004, **5**, 41-47; D. A. Jeffery and M. Bogyo, *Curr. Opin. Biotechnol.*, 2003, **14**, 87-95; D. A. Campbell and A. K. Szardenings, *Curr. Opin. Chem. Biol.*, 2003, **7**, 296-303; J. W. Kozarich, *Curr. Opin. Chem. Biol.*, 2003, **7**, 78-83.
8. Y. Liu, M. P. Patricelli and B. F. Cravatt, *Proc. Natl. Acad. Sci. USA*, 1999, **96**, 14694-14699.
9. D. Greenbaum, K. F. Medzihradszky, A, Burlingame and M. Bogyo, *Chem. Biol.*, 2000, **7**, 569-581.
10. M. Bogyo, J. S. McMaster, M. Gaczynska, D. Tortorella, A. L. Goldberg and H. L. Ploegh, *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 6629-6634.
11. B. M. Kessler, D. Tortorella, M. Altun, A. F. Kisselev, E. Fiebinger, B. G. Hekking, H. L. Ploegh and H. S. Overkleeft, *Chem. Biol.*, 2001, **8**, 913-929.
12. A.-M. Lennon-Duménil, A. H. Bakker, R. Maehr, E. Fiebinger, H. S. Overkleeft, M. Rosemblatt, H. L. Ploegh and C. Lagaudrière-Gesbert, *J. Exp. Med.*, 2002, **196**, 529-539.
13. C. Kocks, R. Maehr, H. S. Overkleeft, E. W. Wang, L. K. Iyer, A. M. Lennon-Duménil, H. L. Ploegh and B. M. Kessler, *Mol. Cell. Proteomics*, 2003, **2**, 1188-1197.
14. A. L. Grunidin, I. M. Koshika and I. N. Domnin, *Russ. J. Org. Chem.*, 1993, **29**, 340-341.
15. G. Tong, J. M. Lawlor, G. W. Tregear and J. Haralambidis, *J. Org. Chem.*, 1993, **58**, 2223-2231.
16. J. H. Bowie and D. H. Williams, *Tetrahedron*, 1967, **23**, 305-320.
17. The reduction of 6-(Fmoc-amino)-hexa-2,4-diyn-1-ol, 6-(Fmoc-amino)-hexa-2,4-diyn-1-tetrahydropyranyl ether and 6-(Fmoc-amino)-hexa-2,4-diynoic acid methyl ester, prepared *via* Glaser oxidative couplings of Fmoc-protected propargylamine with propargyl alcohol, 2-(prop-2-yn-1-yloxy)tetrahydro-2*H*-pyran and methyl propiolate respectively, under the same conditions resulted in substantial amounts of hexa-, hepta- and nonadeuterated products.
18. M. Lange and K. Undheim, *Tetrahedron*, 1998, **54**, 5337-5344.
19. P. F. Alewood, M. Benn and R. Reinfried, *Can. J. Chem.*, 1974, **52**, 4083-4089.
20. K. Mori and H. Iwasawa, *Tetrahedron*, 1980, **36**, 87-90.
21. R. Hirschmann, K. C. Nicolaou et al. *J. Am. Chem. Soc.* 1993, **115**, 12550-12568.

Chapter 5

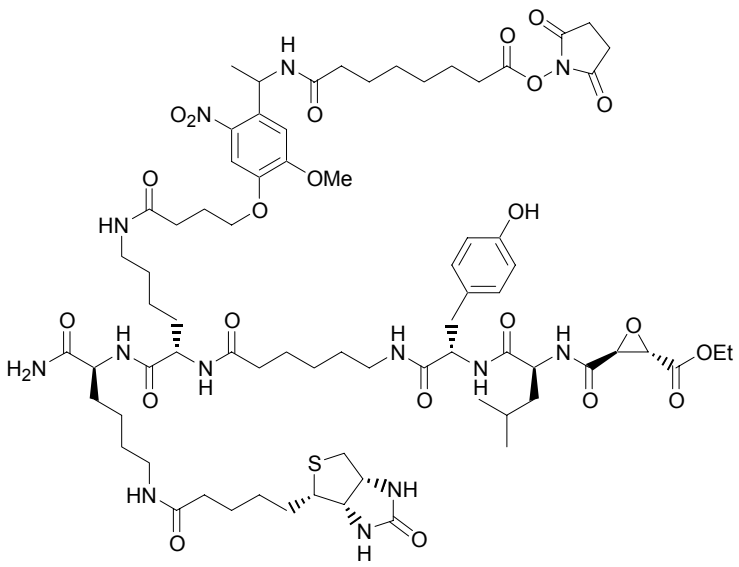
A cysteine protease probe with an additional amine-reactive group[1]

Introduction

Phagocytosis is, in many cases, the first step towards the presentation of exogenic antigenic proteins on MHC class II molecules of professional antigen presenting cells.[2] The ingested antigens travel to the lysosomes and are degraded by a set of different hydrolases. Some of the peptide fragments generated by these proteases escape further degradation by binding to MHC class II molecules. The loaded MHC class II molecules are transported to the cell surface for presentation to CD4+ T-cells (helper T cells). Depending on other factors, like the conditions of the antigen presentation, the CD4+ T-cells can initiate a B-lymphocyte or CD8+ cytotoxic T-lymphocyte mediated immune response.

The cathepsin family of lysosomal cysteine proteases plays an important role in the process of antigen processing: these proteases are able to cleave the majority of peptide bonds in antigens.[3] The activity of the different proteases in the lysosomal pathway is not constant.[4] Different protease activities are added to the maturing lysosome sequentially and

the activity of all proteases is also dependent on the pH. As a result, the processing of epitopes is governed by gradients of tightly regulated protease activities.



1

Figure 1. Target structure

Here, a cysteine protease probe is described for use in the monitoring of the activity of the lysosomal cysteine proteases that a model antigen encounters after internalisation. The new probe contains an amine reactive group to allow conjugation to a model antigen, as well as a photolabile linker to release the probe from the model antigen before analysis of the target proteases. A known cysteine protease activity based probe, DCG-04,[5] was equipped with an activated ester to facilitate conjugation of this probe to a lysine side chain amine of an antigen of choice (**1**, Figure 1). In a typical experiment, probe **1** could be conjugated to ovalbumin (**2**) to furnish construct **3** (Figure 2). This ovalbuminated probe captures an active cathepsin of the cysteine protease family to give **4**. This way, the biotin-bearing cysteine protease probe has become covalently bound to the targeted cysteine protease. An photolabile ortho-nitrobenzyl linker would allow for photorelease of the ovalbumin carrier protein, and biotinylated cathepsin **5** could be visualised by standard avidin based blotting techniques.

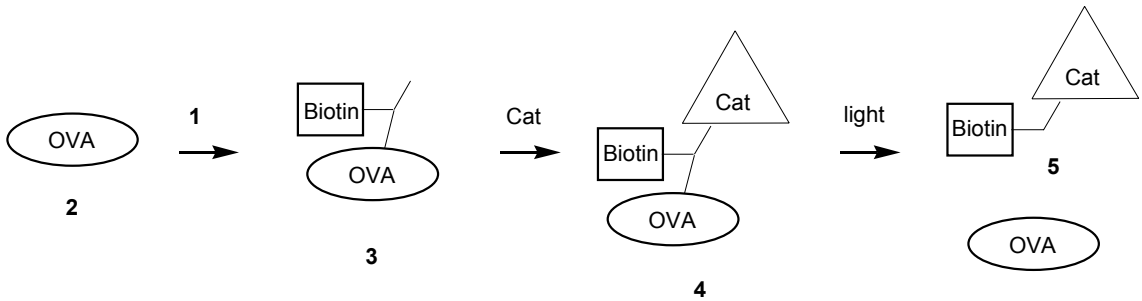
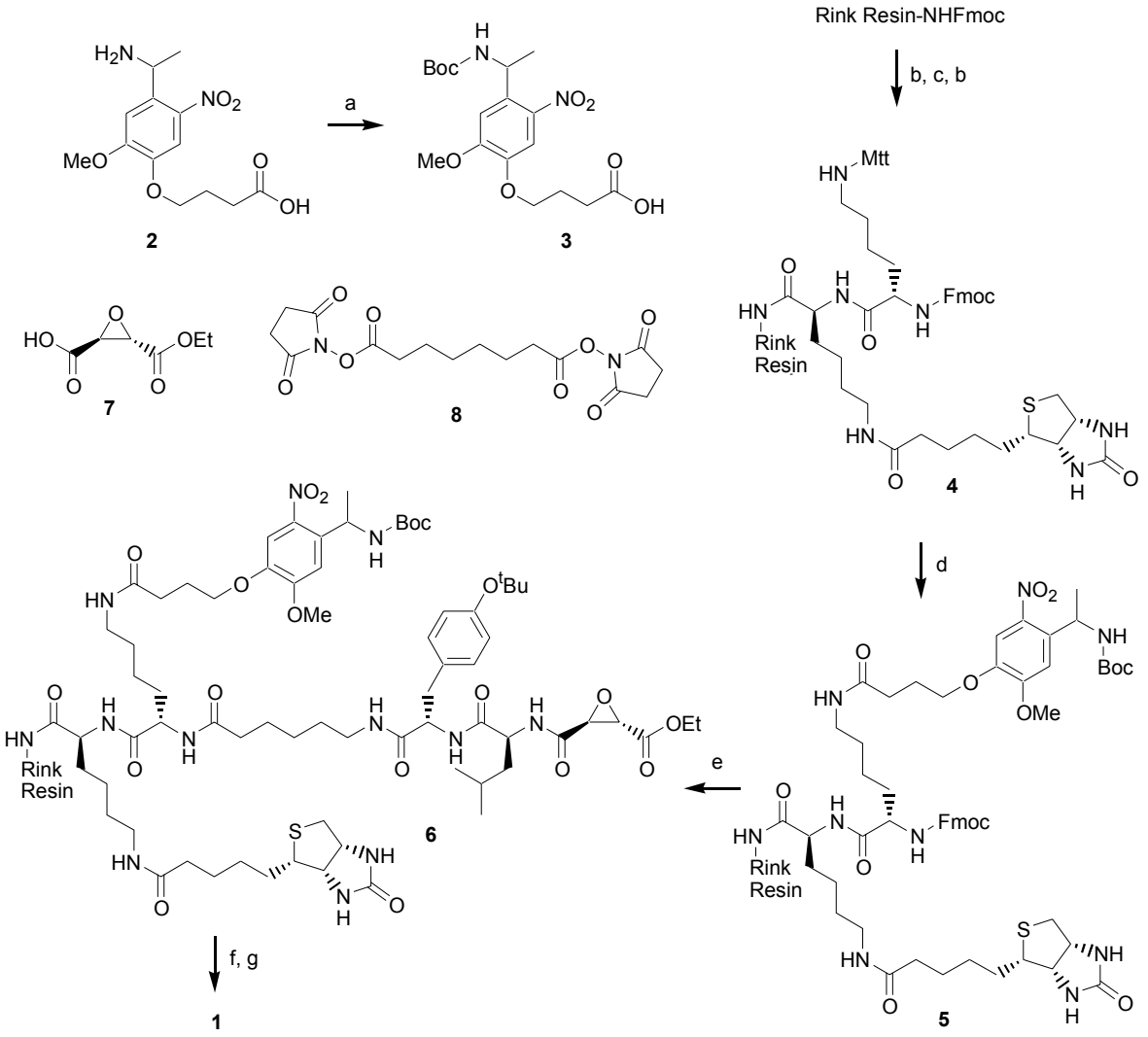


Figure 2. Schematic representation of the cathepsin labeling approach. Ovalbumin (OVA) is conjugated to probe **1**, a cathepsin (Cat) reacts with the probe and OVA is photoreleased to furnish labeled cathepsin **5**.

Synthesis

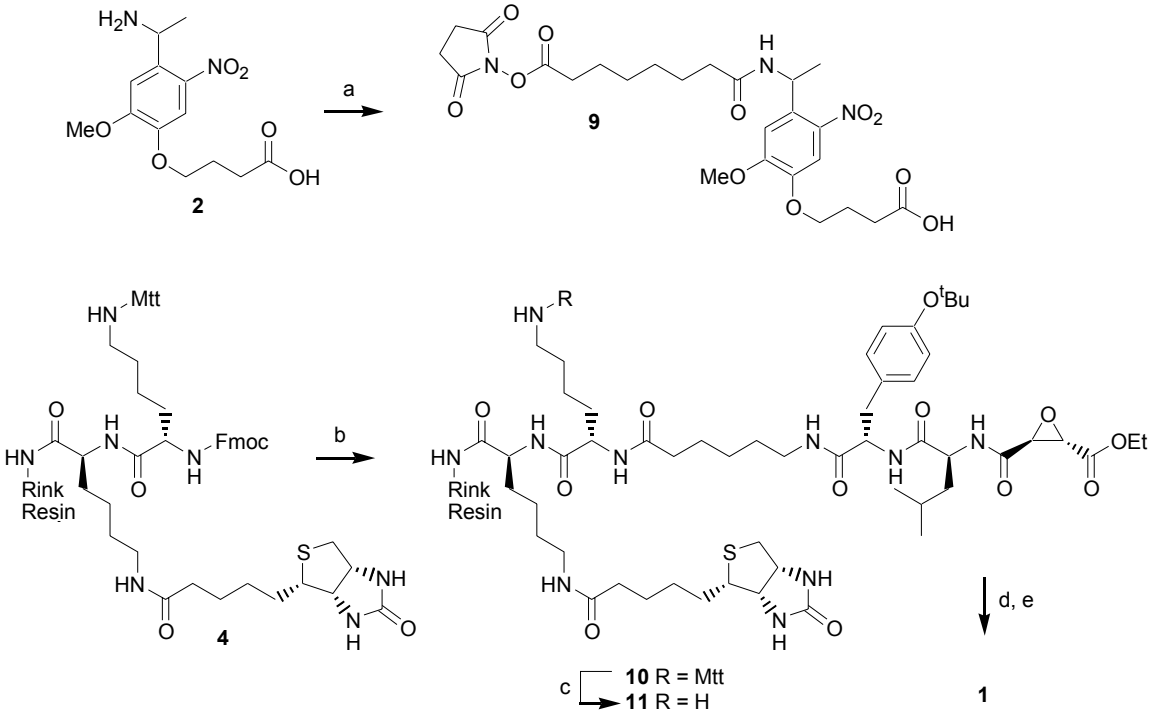
The first route of synthesis that was explored is depicted in Scheme 1. Suitable protected photocleavable linker **3** (Scheme 1) was obtained by Boc-protection of known photolabile amino acid **2**.[6] Immobilised biocytin was created as described in Chapter 3 and elongated with Fmoc-Lys(Mtt)-OH to afford **4**. The lysine side chain amine was liberated by treatment with 1% TFA in CH_2Cl_2, followed by acylation with Boc-protected nitrovaleryl acid **3** to give immobilised **5**. The peptide was elongated by standard Fmoc-based SPPS to



Scheme 1. Synthesis of photocleavable activated probe **1**. Reagents and conditions: a) Boc_2O, dioxane, H_2O, NaOH; b) 20% piperidine in NMP, then Fmoc-Lys(Mtt)-OH, HCTU, DiPEA, NMP; c) 1% TFA in CH_2Cl_2, then biotin, HCTU, DiPEA, NMP; d) 1% TFA in CH_2Cl_2 then **3**, HCTU, DiPEA, NMP; e) Repeated cycles of SPPS: Fmoc cleavage: 20% piperidine in NMP; amino acid condensation: Fmoc-protected amino acid, HCTU, DiPEA, NMP; Fmoc-protected building blocks were used in the following order: Fmoc-Ahx-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Leu-OH, **7**; f) TFA/H_2O (95/5 v/v); g) **8**, DiPEA, DMF.

give protected and immobilized **6**. The peptide was cleaved from the resin by treatment with TFA/H_2O/TIS, removing also the acid labile side chain protecting groups. The crude intermediate was purified to homogeneity by RP-HPLC and reacted with excess bisactivated ester **8**. LC/MS analysis confirmed the formation of product **1** but HPLC purifications did not effect separation of the target compound from the excess of **8**.

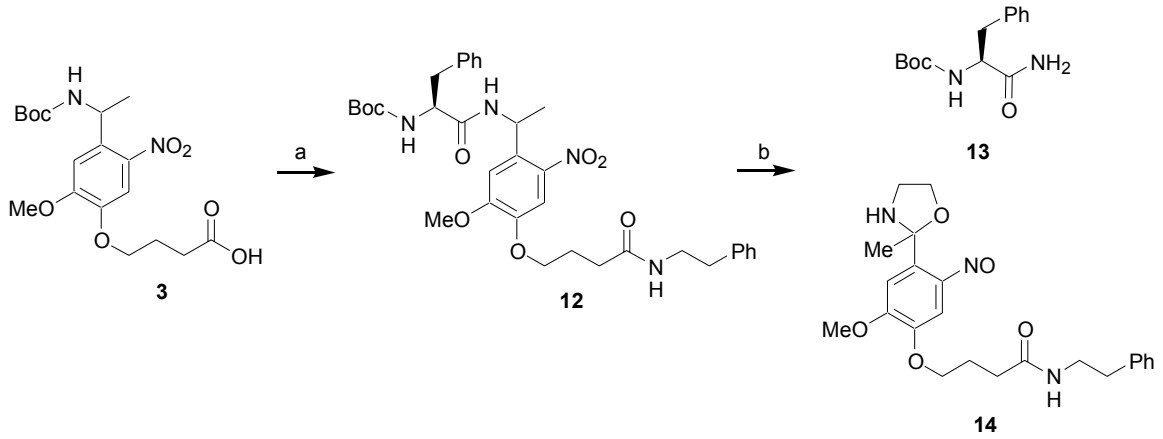
To circumvent this purification problem, it was decided to perform all coupling steps on the solid support. For this purpose, activated ester **9** (Scheme 2) was prepared by reaction of photolabile linker **5** with disuccinimidyl suberate **8**. Elongation of **4** by standard Fmoc-based SPPS afforded Mtt-protected immobilised peptide **10**. After Mtt cleavage, the lysine side chain was elaborated using acid **9** at low temperature. Nucleophilic attack on the less reactive OSu ester was suppressed by the reaction conditions. Liberation of the peptide from the solid support followed by RP-HPLC purification afforded activated ester **1** in 6% overall yield.



Scheme 2. Synthesis of photocleavable activated probe **1.** Reagents and conditions: a) **8**, Et_3N, DMF; b) Repeated cycles of SPPS: Fmoc cleavage: 20% piperidine in NMP; amino acid condensation: Fmoc-protected amino acid, HCTU, DiPEA, NMP; Fmoc-protected building blocks were used in the following order: Fmoc-Ahx-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Leu-OH, **7**; c) 1% TFA in CH_2Cl_2; d) **9**, HATU, DiPEA, NMP, -20°C; e) TIS/H_2O/TFA (2.5/2.5/95 v/v/v).

Inhibitory results and photocleavage characteristics

Next, the photocleavage characteristics of the linker were investigated. Model compound **12**, conveniently synthesised from **3** in three steps (Scheme 3), was chosen for initial studies. Photocleavage proceeded smooth, and after 4 h all starting material had been converted to amide **12** and aminal **13** as judged from LC/MS analysis of the reaction mixture.



Scheme 3. Initial photocleavage studies. Reagents and conditions: a) 2-phenylethylamine, DIC, DMAP, CH_2Cl_2, then TFA/CH_2Cl_2 (1/1 v/v) then Boc-Phe-OH, DIC, DMAP, CH_2Cl_2; b) **12** was dissolved in DMSO/MeOH/MeCN (10/45/45 v/v/v) and two drops of ethanolamine were added. The reaction mixture was irradiated in a quartz cuvet using a Rayonet reactor equipped with 7 × 3500 Å (15 W) lamps. Samples were taken at t=0, 1 h, 2 h, 4 h, 8 h and analyzed by LC/MS.

With the photocleavage conditions in hand, an ovalbumin-**1** conjugate was synthesized. Ovalbumin was incubated with 3 equivalent of OSu ester **1** for 2 h at 4°C, and unreacted and hydrolyzed **1** were removed by dialysis. The resulting construct was exposed to the action of Cat B at 37°C for 30 min. Western blot analysis indicated the presence of biotinylated Cat B, but no heavy band for the Cat B-**1**-Ova was observed. This observation forced us to exclude the presence of **1** after dialysis.

As a model for the dialysis of **1**, ovalbumin was incubated with DCG-04 (compound **1a,** Chapter **4**) under the same conditions as **1**. Unexpectedly, not only cathepsin labeling was found upon incubation with Cat B, but also labeling of ovalbumin was detected by Western blotting with strept-HRP. This would indicate that ovalbumin could react with the epoxide moiety in DCG-04. To test which nucleophile was responsible for this labeling, ovalbumin was treated with maleimide to block all cysteine residues. DCG-04 was added

and western blotting indicated no alkylation with DCG-04. However, labeling returned if a dialysis step was included after maleimide treatment, albeit with lower efficiency as seen before. As a third control, DCG-04 and ovalbumin were mixed in blotting buffer, boiled briefly to denaturate and separated on SDS-PAGE. Again, biotinylation was confirmed with western blotting. From these experiments, it was concluded that 1) the dialysis protocol is insufficient for removal of DCG-04 and possibly also for **1** and 2) that boiling in Laemmli sample buffer is sufficient to effect alkylation of cysteine residues within ovalbumin. These results demand a proper purification protocol for ovalbumin after the reaction with **1** and a method to prove that **1** is attached to ovalbumin *via* the photocleavable side chain and not through the electrophilic site of the epoxide moiety.

Conclusions

The synthesis and purification of probe **1** required a strategy that allowed all coupling steps to be perfomed on a solid support. The target compound was obtained in 6% overall yield. The photocleavage characteristics of the nitrovaleryl linker were established. The labeling of ovalbumin with **1**, in order to study the inhibition characteristics of the construct, needs optimization. The purity and homogeneity of ovalbumin labeled with probe **1** could not be confirmed, due to possible alkylation of ovalbumin *via* a cysteine thiol moiety. This would indicate that the labeling reaction of lysine with hydroxysuccimidyl ester is unsuitable to be performed in the presence of an epoxide moiety, and possibly an bioorthogonal labeling approach using the Staudinger ligation or Huisgen cyclization could produce efficient labeling of ovalbumin.

Experimental part

General methods and materials: Solvents used in the solid phase peptide synthesis, DiPEA and TFA were all peptide synthesis grade (Biosolve) and used as received. The protected amino acids, Rink amide MBHA resin (0.78 mmol g^{-1}) and HCTU were obtained from NovaBiochem. Ethyl (2S,3S)oxirane-2,3-dicarboxylate was prepared as described in Chapter 4. Fmoc-Lys(Mtt)-OH was from Senn Chemicals, SPPS was carried out using a 180° Variable Rate Flask Shaker (St. John Associates, Inc.). LC/MS analysis was performed on a Jasco HPLC system (detection simultaneously at 214 and 254 nm) coupled to a Perkin Elmer Sciex API 165 mass spectrometer equipped with a custom-made Electrospray Interface (ESI). An analytical Alltima C_{18} column (Alltech, 4.6 × 250 mm, 5μm particle size) was used. Buffers: A: H_2O; B: CH_3CN; C: 0.5% aq. TFA. For RP HPLC-purifications a Biocad "Vision" automated HPLC system (PerSeptive

Biosystems, Inc.) was used. The applied buffers were A, B and C. ^{1}H-NMR spectra were recorded with a Bruker DMX 600 instrument at 600 MHz with chemical shifts (δ) relative to tetramethylsilane. When appropriate, reactions were performed under the exclusion of light.

Synthesis of 4-(4-(1-(succinimidylsuberoylamino)ethyl)-2-methoxy-5-nitrophenoxy)butanoic acid (9): A solution of known 4-(4-(1-(amino)ethyl)-2-methoxy-5-nitrophenoxy)butanoic acid[6] (298 mg, 1 mmol) and Et_3N (0.14 ml, 1 mmol) in DMF (5 ml) was added dropwise to a solution of disuccinimidyl suberate (Alldrich, 386 mg, 1.3 mmol) in DMF (5 ml) and the mixture was stirred for 1.5 h until TLC-analysis (ethyl acetate) indicated a completed reaction. The mixture was concentrated *in vacuo* and the residue was taken up in ethyl acetate. The white precipitate was filtered off and the remainder solution was purified by silica gel chromatography (0-1% acetic acid in ethyl acetate) to yield 250 mg (0.45 mmol, 45%) of a faint yellow solid. ^{1}H NMR ($CDCl_3$): δ, 7.56 (s, 1H), 7.19 (s, 1H), 6.52 (d, *J*=7.3 Hz, 1H), 5.6-5.4 (m, 1H), 4.2-4.0 (m, 2H), 3.92 (s, 3H), 2.84 (s, 4H), 2.59 (t, 4H, *J*=7.3 Hz), 2.3-2.0 (m, 4H), 1.8-1.2 (m, 11H). ESI-MS $C_{25}H_{33}N_3O_{11}$ + H^+ requires 552.2, found 551.9.

Synthesis of 1: Fmoc-protected biocytinylated Rink amide resin (132 mg, 50 μmol) was elongated by standard Fmoc-based SPPS to give resin-bound Lys(Bio)Lys(Mtt)AhxTyr(tBu)Leu-Epoxirane. In brief, where appropriate removal of the Fmoc protecting group was accomplished by treatment of the resin-bound peptide with 20% (v/v) piperidine in NMP for 20 min. Peptide coupling steps were performed by treatment of the resin with a premixed (5 min) solution of the of the appropriate acid (5 eq.), HCTU (5 eq.) and DiPEA (6 eq.) in NMP (0.5 ml) for 1 hour unless stated otherwise. Coupling efficiencies were monitored with the Kaiser test and couplings were repeated if necessary. After coupling and deprotecting steps the resin was washed with NMP (5×). After the last coupling step, the resin was washed with CH_2Cl_2 (3×). Of 25 μmol of resin-bound Lys(Bio)Lys(Mtt)AhxTyr(tBu)LeuEpoxirane, the lysine side chain Mtt protecting group was removed by treatment with 1% TFA in CH_2Cl_2 (4×10 min). The resin was washed (CH_2Cl_2 5×, then NMP 2×) and neutralised with 10% DiPEA in NMP. To cooled (-20°C) resin was added a premixed (5 min) and cooled (-20°C) solution of active ester **8** (55 mg, 0.10 mmol), HATU (38 mg, 0.10 mmol) and DiPEA (34 μl, 0.20 mmol) in NMP and the resin was agitated for 1 h, and allowed to warm to ambient temperature. The resin was washed extensively (alternating CH_2Cl_2-MeOH 3×, alternating CH_2Cl_2-Et_2O 3×), transferred into a clean vial and treated with TFA/H_2O/TIS (1 ml, 95/2.5/2.5 v/v/v) for 1 h. The mixture was filtered and the resin washed with TFA (2× 1 ml). The filtrate was diluted (toluene, 10 ml) and concentrated in vacuo. The crude product was coevaporated (toluene, 3×) and purified to homogeneity by RP-HPLC, applying a linear gradient (20-55% B in 3 column volumes) to yield 2.38 mg (1.5 μmol, 6.1%) of product. ^{1}H NMR (DMSO, 300 K, DMX 600): δ, 9.12 (s, 1H), 8.50 (d, 1H, *J* = 7.6 Hz), 8.44 (d, 1H, *J* 7.6 Hz), 8.07 (d, 1H, *J* = 8.2 Hz), 7.89

(d, 2H, *J* = 7.7 Hz), 7.83-7.77 (m, 2H), 7.77-7.65 (m, 6H), 7.46 (s, 2H), 7.14 (s, 1H), 6.99-6.91 (m, 5H), 6.60 (d, 4H, *J* = 8.3 Hz), 6.41 (s, 1H), 6.34 (s, 1H), 5.38-5.31 (m, 1H), 4.37-4.26 (m, 5H), 4.21-4.08 (m, 8H), 4.04-3.94 (m, 3H), 3.86 (s, 3H), 3.69 (d, 1H, *J* = 1.7 Hz), 3.57 (d, 1H, *J* = 1.7 Hz), 3.11-3.04 (m, 3H), 2.83-2.74 (m, 6H), 2.73-2.50 (m, 8H), 2.23-1.88 (m, 12H), 1.67-1.09 (m, 28H), 0.87-0.74 (m, 7H). ESI-MS: $C_{74}H_{109}N_{13}O_{22}S + H^+$ requires 1564.8, found 1565.0;

Conjugation of 1 to ovalbumin: Ovalbumin (22 nmol) was incubated with **1** (66 nmol, 5 µl of a 13 mM stock in DMSO) in phosphate buffer (0.5 ml, 50 mM, pH 7.8) for 2 h at 4°C. To purify the labeled ovalbumin, the reaction mixture was filtered over a 10 kD cutoff filter and the resulting solution was lyophilized. A sample of the ovalbumin conjugate was boiled for 3 min in 4× Laemmli sample buffer (5% SDS, 46% glycerol, 0.156 mM Tris-HCl pH 6.8, 0.25% bromophenol blue, 1.4 M β-mercaptoethanol in H_2O) and resolved by 12.5% SDS-PAGE. Proteins were visualized by silver staining (Amersham). Western blotting and incubation with streptavidin-horseradish peroxidase (strept-HRP) conjugates allowed the visualization of biotinylated ovalbumin by chemiluminescence.

Cathepsin activity profiling: 2 pmol Cat B (R&D Systems) was pre-incubated at 10 µg/ml in 25 mM 2-(*N*-morpholino)ethanesulfonic acid (MES), 5 mM dithiothreitiol, pH 5.0 for 15 min at room temperature. Ovalbumin-**1** conjungate (2 pmol) was incubated with the activated Cat B for 1 h at ambient temperature. The mixture was neutralized to pH 7 by addition of aq. NaOH (0.4 µl of a 1M solution) and resolved by 10% SDS-PAGE followed by Western blotting. Biotinylated proteins were visualized by chemiluminescence after incubation with strept-HRP.

References and notes

1. The work presented in this chapter was performed in close collaboration with Bobby Florea and Micha Slegt.
2. I. Mellman, S. J. Turley and R. M. Steinman, *Tr. Cell Biol.*, 1998, **8**, 231-237.
3. K. Honey and A. Y. Rudensky, *Nature Rev. Immunol.*, 2003, **3**, 472-482.
4. A.-M. Lennon-Duménil, A. H. Bakker, R. Maehr, E. Fiebinger, H. S. Overkleeft, M. Rosemblatt, H. L. Ploegh and C. Lagaudrière-Gesbert, *J. Exp. Med.*, 2002, **196**, 529-539.
5. D. Greenbaum, K. F. Medzihradszky, A, Burlingame and M. Bogyo, *Chem. Biol.*, 2000, **7**, 569-581.
6. C. P. Holmes and D. G. Jones, *J. Org. Chem.*, 1995, **60**, 2318-2319.

Chapter 6

A cell penetrating cysteine protease probe[1]

Introduction

Around a decade ago, it was found that specific oligopeptides can traverse lipid bilayers that make up the cell membrane of mammalian cells, and that they can do so when covalently bound to proteins or other relevant biomolecules. This property, which was first identified for peptides derived from the HIV-1 Tat[2] and *Drosophila* Antennapedia[3] (Ant) proteins, has attracted much attention. As a result, a wide variety of natural and synthetic peptides displaying transferring properties have been identified and used to deliver cargo into cells.[4-6] Common to most of these sequences is the presence of numerous basic amino acid residues, as is evident in the Tat (GRKKRRQRRRPPQ) and Ant (RQIKIWFQNRRMKWKK) peptides. It is thought that these cationic residues are responsible for the translocation activity.

To investigate the role of arginine residues further, oligoarginine peptides built from seven to eleven arginine moieties were used in uptake experiments. The group of Futaki showed that oligoarginine translocated both fluoresceine and carbonic anhydrase into RAW cells.[7] The optimal number of arginine residues for translocation was found to lie between six and eight. Wender *et al.* synthesised both all-D and all-L oligoarginine peptides, and

demonstrated that both are capable of translocation of fluoresceine.[8] They found that nine arginine residues were optimal to induce translocation, and that the all-D peptide is a better translocator than the corresponding all-L peptide. In another study,[9] the uptake of fluoresceine conjugates with oligomers of arginine, lysine or histidine was compared, and it was found that only the oligoarginine conjugates effected fluorescence within cells. Octamers and nonamers of arginine exhibited comparable uptake efficiency, and again the all-D peptides residues performed better than the all-L peptides.

The mechanism of translocation mediated by cell penetrating peptides still has to be revealed. From the observation that the all-D peptides are at least as efficient in translocation as the corresponding all-L peptides, it can be concluded that there is no specific receptor involved. There is, however, no consensus whether the translocation is caused by spontaneous permeabilisation of the cell membrane or by active uptake of the cargo *via* the endosomal pathway. Evidence supporting the latter option is accumulating slowly,[10] but the possibility of multiple pathways can not be excluded yet. The endosomal pathway is a feature of the cell that is used to regulate uptake of nutritients and communication signals from the environment of the cell. There are different uptake pathways for molecules to enter the cell.[11] Generally, the first step involves a membrane rearrangement resulting in the formation of a small, membrane enclosed vesicle in the cytosol. These vesicles can travel to the endoplasmatic reticulum or golgi apparatus, and deliver their cargo there, or enter the lysosomal pathway, resulting in the degradation of the contents of the vesicle through the action of lysosomal cysteine proteases.[12]

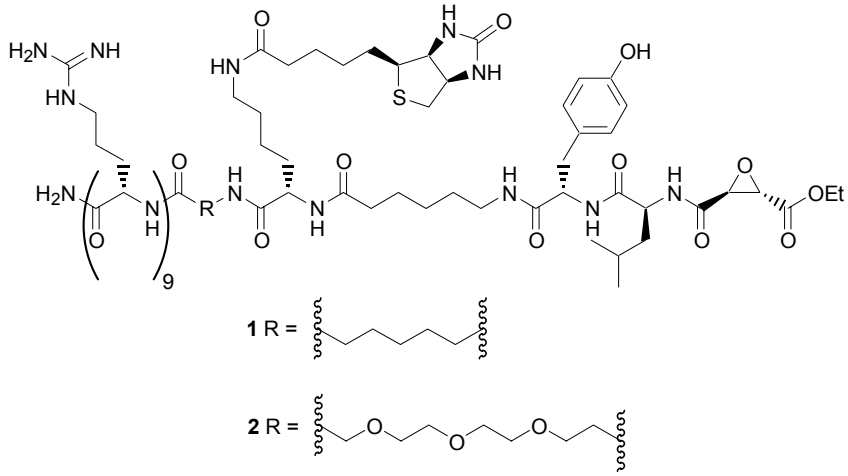


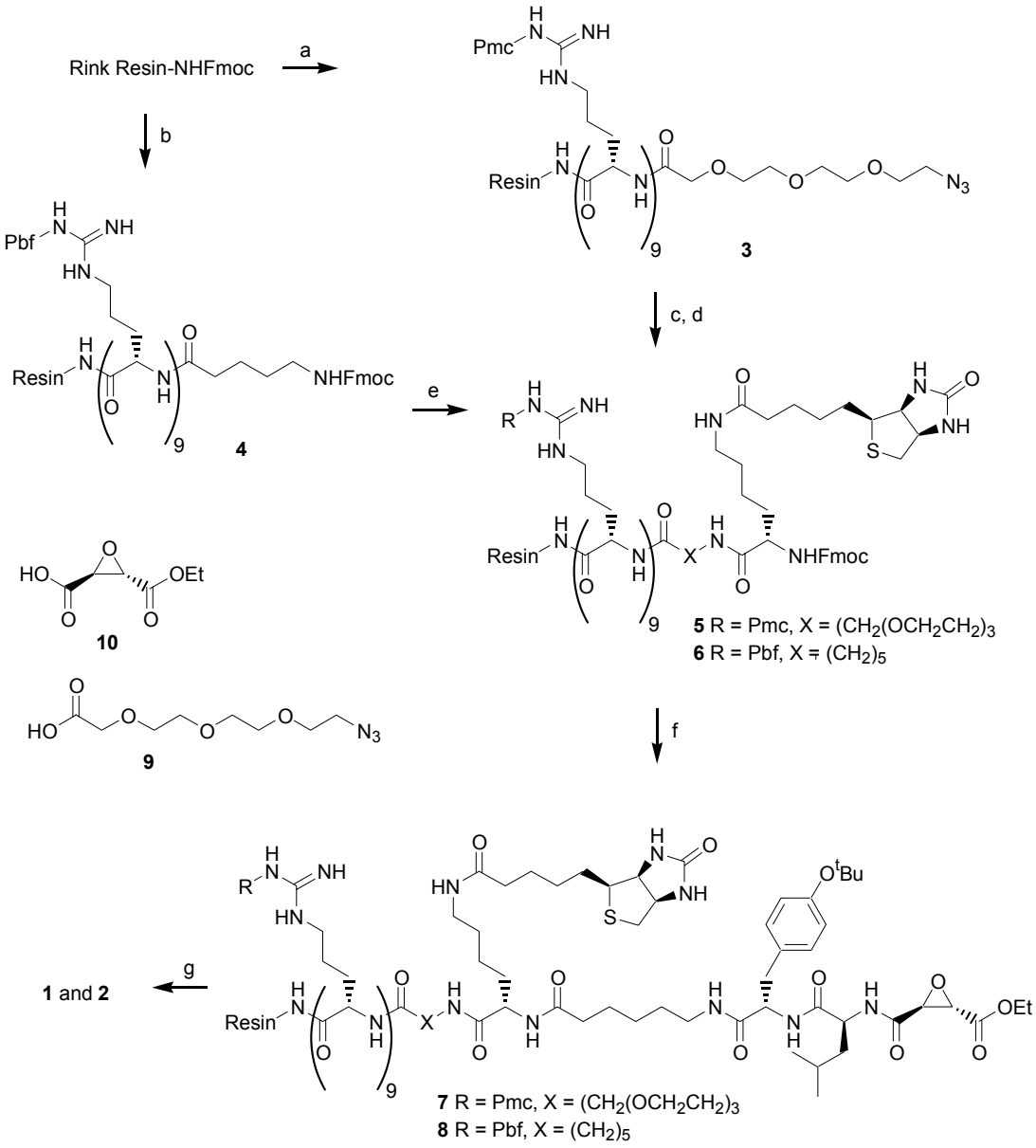
Figure 1. Target structures.

It was envisaged that involvement of the endosomal pathway could be demonstrated by the activity of lysosomal cysteine proteases, which may be visualised *via* a suitable activity based probe. Fusion of an activity based cysteine protease probe to the nonaarginine cell penetrating peptide would yield a probe that enters the cell in the fashion dictated by the nonaarginine sequence. Any cysteine protease that is encountered in the uptake pathway can be visualised, thus generating insight in the actual path that the probe, and thus the nonaarginine peptide, follows. For this purpose, two conjugates were synthesized of nonaarginine with DCG-04,[13] a broad class cysteine protease activity based probe. The cell penetrating peptide is linked to DCG-04 either by a short hydrophobic or by a longer hydrophilic linker (compounds **1** and **2**, figure 1).

Synthesis

The synthesis of probe **1** (Scheme 1) starts with the conjugation of nine Pbf-protected arginine residues to the acid labile rink resin *via* standard Fmoc-based SPPS. Elongation of the growing peptide chain with an aminohexanoic acid spacer yields protected and immobilised **4**. The biotin moiety was introduced by direct coupling of a side chain biotinylated lysine building block and the synthesis was continued by Fmoc-based SPPS furnishing peptide **8** on the solid support. Cleavage of the peptide from the solid phase with concomitant removal of the arginine en tyrosine side chain protecting groups gave endocytosis probe **1** in 11% overall yield after RP-HPLC purification.

The synthesis of probe **2** was performed in a slightly different way. Rink resin was elongated with nine Pmc-protected arginine residues and spacer **9**.[14] The azide group in compound **3** was transformed to an amine by a Staudinger reduction. Elongation with Fmoc-Lys(Mtt)-OH and removal of the hyper acid-labile Mtt protecting group was followed by biotinylation of the lysine side chain to give immobilised **5**. From that point, the synthesis was continued by Fmoc-based SPPS furnishing peptide **7** on the solid support. Cleavage of the peptide from the solid phase and RP-HPLC purification afforded probe **2** in 13% yield.



Scheme 1. Synthesis of polyargininated cysteine protease probes **1** and **2**. Reagents and conditions: General reagents for SPPS: Fmoc cleavage: 20% (v/v) piperidine in NMP; amino acid condensation: Fmoc protected amino acid, HCTU, DiPEA, NMP. a) Repeated cycles of SPPS: Building blocks employed: Fmoc-Arg(Pmc)-OH (9×), **9**; b) Repeated cycles of SPPS: Building blocks employed: Fmoc-Arg(Pbf)-OH (9×), Fmoc-Ahx-OH; c) Me_3P, 20% water in dioxane, then Fmoc-Lys(Mtt)-OH, HCTU, DiPEA, NMP; d) 1% TFA in CH_2Cl_2, then biotin, HCTU, DiPEA, NMP; e) 20% piperidine in NMP, then Fmoc-Lys(Biotin)-OH; f) Repeated cycles of SPPS: Building blocks employed: Fmoc-Ahx-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Leu-OH, epoxirane **10**; g) TFA/TIS/H_2O (95/2.5/2.5 v/v/v), 3 h. Overall yield of **1** and **2**: 11% and 13%, respectively.

Biological activity of the probes

Next, the labeling efficiency of the newly synthesised probes was investigated. Cell lysates and endocytic extracts[15] from different origin were incubated with **1**, **2** or a broad class cysteine protease probe developed in Chapter 5 (**11**; Figure 2). The labeling with **1** can be abolished by pre-treatment of the lysate with E-64,[16] a broad-class cysteine protease inhibitor (Figure 3, lane 1), indicating that **1** is a label for active cysteine proteases. For the other lanes, it is of interest to note that cathepsins labeled with either **1** or **2** are modified with a probe that is 1.5 kD heavier than **11**, and thus all cathepsin band are shifted 1.5 kD upwards. In U937 extracts (lane 2 and 3), probe **1** and **2** label only Cat S and Cat B, whereas **11** also hits Cat X (lane 4). In DC cell lysates, the new probes clearly target Cat B, which may appear as a double band (lane 5 and 6). Cathepsin probe **11** (lane 7) also hits Cat X and Cat S. Monocyte extract labeling with probe **1** and **2** (lane 8 and 9) shows Cat B, Cat H and Cat S. Again, **11** also labels Cat X. Overall, this experiment demonstrates that Cat B, Cat H and Cat S are targeted by **1** and **2** with an efficiency comparable to broad class cysteine protease inhibitor **11**.

Encouraged by these results, the labeling efficiency in living cells was addressed. In a pulse experiment, immature dendritic cells (DC) were incubated with control **11** or probe **1**. Immature DC are known for their high endocytic activity, whereas in mature DC endocytosis is greatly reduced.[17] The cells were washed and lysed in the presence of JPM-565[18] to block remaining cysteine protease activity. The labeling of cysteine proteases was visualised *via* strept-HRP blotting (Figure 4). As a control, immature and mature dendritic cell lysates were labeled with **11** (lane 1 and 2), and it shows that maturity of the dendritic cells does not affect cathepsin labeling. In lane 3, 4 and 5, it is seen that labeling of Cat X, Cat B and Cat S with **11** is almost completely abolished at 4°C, and that it is independent of

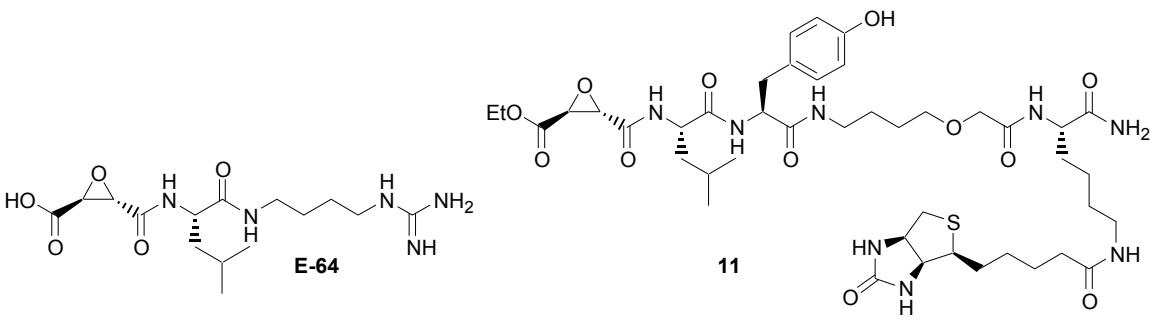


Figure 2. Broad class epoxide based cysteine protease active site alkylating agents used as reference compounds.

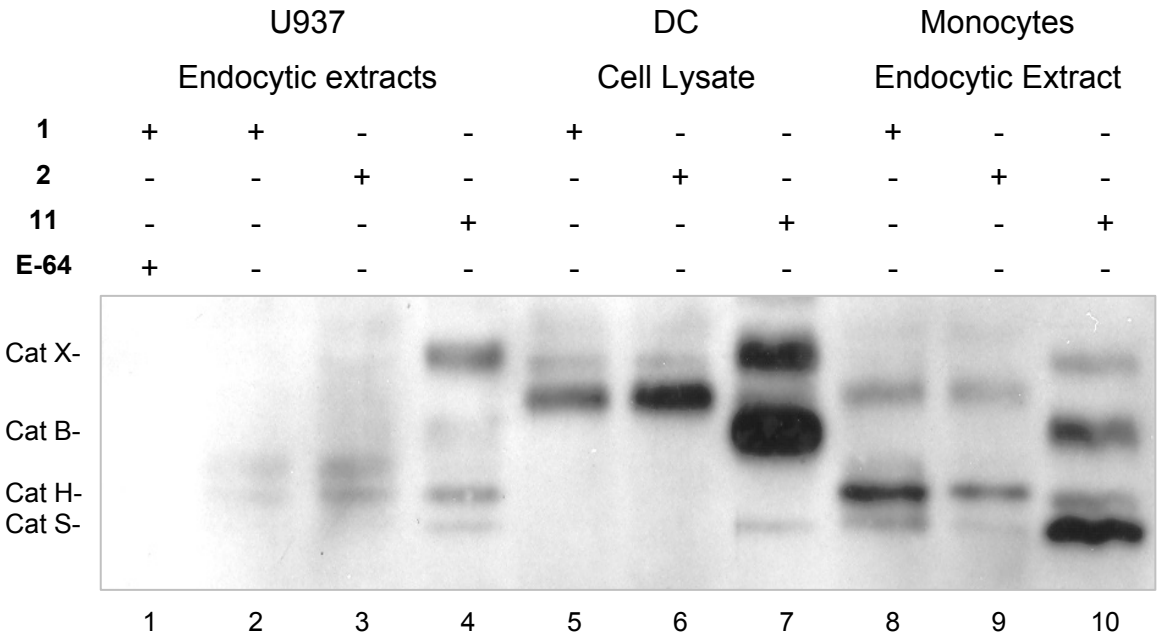


Figure 3. Labeling behaviour of **1** and **2** in cell lysates and endocytic extracts. Lysates and extracts were incubated with **1**, **2** and **11** at 10 μM and ambient temperature for 30 minutes. In lane 1, the extract was pretreated with 40 μM E-64 for 1 hour. Samples were separated by SDS-PAGE and transferred to polyvinyliden difluoride membrane. Incubation with strept-HRP conjugate allowed the visualisation of active cathepsins.

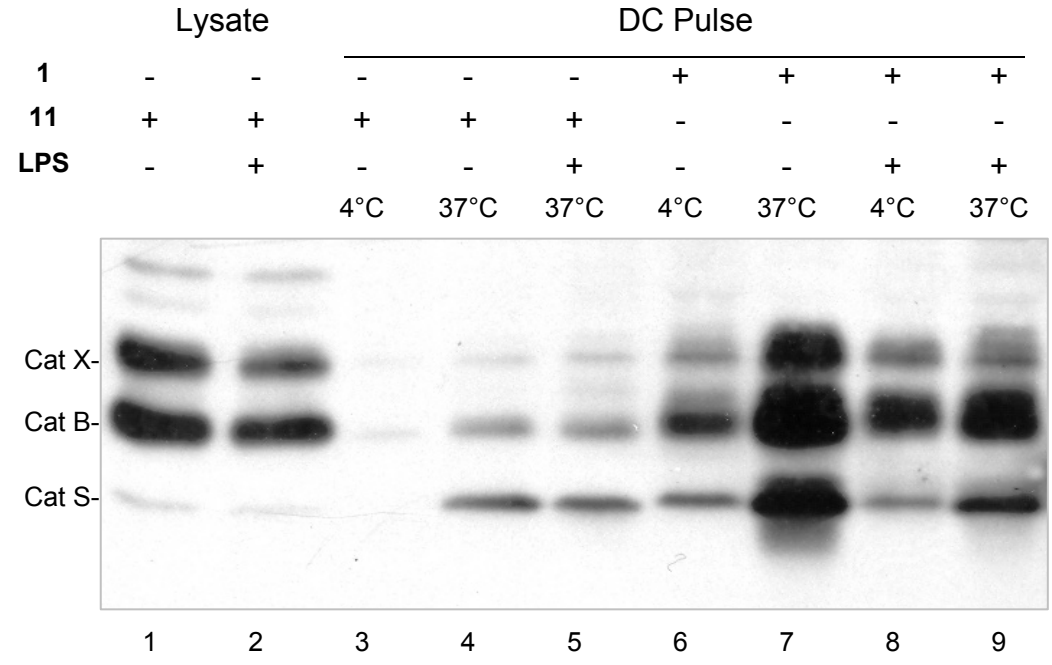


Figure 4. Labeling of dendritic cells with probe **1** or **11**. Cells were incubated for 1 hour with 10 μM of either **11** or **1**. Cells were lysed in the precence of 100 μM JPM-565 and the labeling was addressed by developing a blot.

maturity. However, in comparable experiments (data not shown), labeling generally decreased upon maturation. This indicates that labeling with **11** is at least partly independent of endocytosis. In lane 6-9, cathepsin labeling is greatly reduced at 4°C relative to 37°C, but not blocked. This indicates contributions both of energy dependent and energy independent pathways. In mature cells (lane 9), the uptake is lower than in immature dendritic cells (lane 7), which indicates that endocytosis is a large component of the uptake in immature cells at 37°C. Interestingly, in mature cells the difference in uptake between 4°C and 37°C (lane 8 and 9) is not so pronounced, which suggests that in mature cells no energy dependent pathway is employed for the uptake of the probe. This suggests that the energy-dependent component of uptake is abolished both at 4°C and by maturation, and endocytosis could be this component in both cases.

Conclusion

In this chapter, it is shown that a broad class cysteine protease probe that is hardly cell permeable can be made cell penetrating *via* the attachment of a nonaargine cell penetrating vector. The labeling of cysteine proteases with the new probes in living cells seems to be largely energy-dependent, because it is greatly reduced at 4°C.

These probes may play a role in the elucidation of the pathway that is employed to take up nonaarginine vectors. Further experiments addressing the role of individual steps of the uptake pathway could help to gain insight in details of the uptake. Furthermore, it will be interesting to investigate the uptake efficiency and the targeting of cysteine proteases in other cell types or tissues.

Experimental section

General methods and materials: Solvents used in the solid phase peptide synthesis, DiPEA and TFA were all of peptide synthesis grade (Biosolve) and used as received. The protected amino acids, Rink amide MBHA resin (0.78 mmol/g) and HCTU were obtained from NovaBiochem. Ethyl (2S,3S)oxirane-2,3-dicarboxylate was prepared as described in Chapter 4. Fmoc-Lys(Mtt)-OH was from Senn Chemicals. SPPS was carried out using a 180° Variable Rate Flask Shaker (St. John Associates, Inc.) or on a 443A Peptide Synthesiser (Applied Biosystems). LC/MS analysis was performed on a Jasco HPLC system (detection simultaneously at 214 and 254 nm) coupled to a Perkin Elmer Sciex API 165 mass spectrometer equipped with a custom-made Electrospray Interface (ESI). High resolution mass spectra were recorded on a Finnigan LTQ-FT (Thermo electron). An analytical Alltima C_{18} column (Alltech, 4.6 × 250 mm, 5 μm particle size) was used.

Buffers: A: H_2O; B: CH_3CN; C: 0.5% aq TFA. For RP HPLC-purifications a Biocad "Vision" automated HPLC system (PerSeptive Biosystems, Inc.) was used. The applied buffers were A, B and C. ^{1}H-NMR spectra were recorded with a Bruker DMX 600 instrument at 600 MHz with chemical shifts (δ) relative to tetramethylsilane. U937 cells were cultured on complete medium (RPMI 1640 medium supplemented with 2 mM L-glutamine, 10% fetal calf serum, and antibiotics).

Solid phase peptide synthesis: Where appropriate removal of the Fmoc protecting group was accomplished by treatment of the resin-bound peptide with 20% (v/v) piperidine in NMP for 20 min. Peptide coupling steps were performed by treatment of the resin with a premixed (5 min) solution of the appropriate acid (5 eq.), HCTU (5 eq.) and DiPEA (6 eq.) in NMP for 1 h. Coupling efficiencies were monitored with the Kaiser test and couplings were repeated if necessary. After coupling and deprotecting steps the resin was washed with NMP (5×).

Synthesis of 1: Fmoc Rink amide resin (78 mg, 50 μmol) was elongated using automated standard Fmoc-based SPPS to give resin-bound $(Arg(Pbf))_9$. The synthesis was continued by manual Fmoc-based SPPS to give protected and resin-bound **8**. The resin was washed extensively (alternating CH_2Cl_2-MeOH 3×, alternating CH_2Cl_2-Et_2O 3×. An aliquot of resin (100 mg, 16 μmol) was transferred into a clean vial, washed with CH_2Cl_2 and treated with TFA/H_2O/TIS (0.7 ml, 95/2.5/2.5 v/v/v) for 2 h. The mixture was filtered into cold Et_2O and the white precipitate was collected by centrifugation and decantation. The precipitate was washed (Et_2O), followed by RP-HPLC purification of the crude product (linear gradient in B: 20-30% B in 3 column volumes) to yield 4.3 mg (1.8 μmol, 11%) of product. ^{1}H NMR (D_2O, 295 K, DMX 600): δ, 7.08 (d, 2H, J = 7.9 Hz), 6.80 (d, 2H, J = 8.2 Hz), 4.58-4.55 (m, 1H), 4.45-4.41 (m, 1H), 4.38-4.35 (m, 1H), 4.34-4.18 (m, 12H), 4.13-4.09 (m, 1H), 3.68 (br s, 1H), 3.48-3.45 (m, 1H), 3.30-3.25 (m, 1H), 3.24-3.09 (m, 24H), 3.03-2.86 (m, 4H), 2.79 (s, 1H), 2.69 (s, 1H), 2.28-2.18 (m, 6H), 1.80-1.20 (m, 61H), 0.89-0.80 (m, 6H). ESI-MS: $C_{105}H_{189}N_{45}O_{24}S$ + 4 H^+ requires 625.5, found 625.4; $C_{105}H_{189}N_{45}O_{24}S$ + 3 H^+ requires 833.7, found 833.6.

Synthesis of 2: Fmoc Rink amide resin (78 mg, 50 μmol) was elongated using automated standard Fmoc-based SPPS to give resin-bound **3**. The azide moiety was reduced to an amine on solid phase by treatment of the resin, after washing (dioxane 5×), with Me_3P (0.3 ml 1M in toluene, 0.3 mmol, 6 eq.) in dioxane/H_2O (4/1 v/v, 1 ml) for 40 min, followed by extensive washing of the resin (dioxane 3×, alternating CH_2Cl_2-MeOH 3×, and NMP 3×). The peptide chain was elongated manually with Fmoc-Lys(Mtt)-OH and the lysine side chain Mtt protecting groups was removed by treatment with 1% TFA in CH_2Cl_2 (4×10 min) The resin was washed (CH_2Cl_2 5×, then NMP 2×) and neutralized with 10% DiPEA in NMP. The lysine side chain amine was acylated with biotin to give resin-bound **5**. The synthesis was continued by manual Fmoc-based SPPS to give protected and resin-bound **7**.

The resin was washed extensively (alternating CH_2Cl_2-MeOH 3×, alternating CH_2Cl_2-Et_2O 3×. An aliquot of resin (55 mg, 10 µmol) was transferred into a clean vial, washed with CH_2Cl_2 and treated with TFA/H_2O/TIS (0.7 ml, 95/2.5/2.5 v/v/v) for 3 h. The mixture was filtered into cold Et_2O and the white precipitate was collected by centrifugation and decantation. The precipitate was washed (Et_2O), followed by RP-HPLC purification of the crude product (linear gradient in B: 20-30% B in 3 column volumes) to yield 3.3 mg (1.3 µmol, 13%) of product. ^{1}H NMR (D_2O, 295 K, DMX 600) δ, 7.09 (d, 2H, *J* = 8.3 Hz), 6.80 (d, 2H, *J* = 8.2 Hz), 4.58-4.55 (m, 1H), 4.47-4.43 (m, 1H), 4.38-4.23 (m, 13H), 4.20-4.16 (m, 1H), 4.11 (s, 2H), 3.76-3.61 (m 20H), 3.59-3.55 (m, 2H), 3.54-3.50 (m, 1H), 3.40-3.36 (m, 1H), 3.35-3.31 (m, 2H), 3.28-3.24 (m, 1H), 3.22-3.12 (m, 24H), 2.99-2.90 (m, 5H), 2.69 (s, 2H), 2.24-2.20 (m, 4H), 1.82-1.20 (m, 60H), 0.88-0.80 (m, 6H). ESI-MS: $C_{103}H_{185}N_{45}O_{21}S$ + 4 H^+ requires 606.5, found 606.5; $C_{103}H_{185}N_{45}O_{21}S$ + 3 H^+ requires 808.3, found 808.3.

DC: Isolation and cultures: PBMC (peripheral blood mononuclear cells) were isolated by Ficoll/Paque (Life Technologies, Inc., Grand Island, NY) density gradient centrifugation of heparinized blood obtained from buffy coat of healthy volunteers from the blood bank of the University of Tübingen. Isolated PBMC were plated (1×10^8 cells/8 ml per tissue culture flask) into tissue culture flasks (cellstar 75 cm^2, Greiner Bio-One GmbH, Frickenhausen, Germany) in complete medium. After 1.5 h of incubation at 37°C, nonadherent cells were removed, and the adherent cells (12–19% of the incubated cells) were cultured in complete medium supplemented with GM-CSF and IL-4 for 6 days.[19, 20] For maturation, medium was supplemented with LPS (Sigma) at day 6 for additional 24 h.

Isolation of monocytes: Human PBMC were freshly isolated from buffy coats of donor blood by density gradient centrifugation (Ficoll), followed by several washing steps with PBS and a Percoll gradient (Percoll™, Amersham Biosciences, Uppsala, Sweden) consisting of 12 ml 2× PBS and 14 ml Percoll. After centrifugation four distinct layers become discernable, where (from top to bottom) the first contains debris and few thrombocytes, the second enriched monocytes with little lymphocytes and the remaining thrombocytes, the next layer is enriched for lymphocytes with little thrombocytes while the bottom phase contains remaining granulocytes and red blood cells. The second phase (containing the bulk of monocytes) was carefully collected. The majority of the remaining contaminating cells were T cells (15-20%) with small amounts of B cells present (<10%). Cells were kept on ice for all successive procedures to prevent non-specific stimulation. Generation of postnuclear supernatants immediately followed monocyte enrichment.

Active site labeling of cysteine proteases and detection by streptavidin-blotting: Endocytic extracts of at least 1×10^8 U937 cells or at least 5×10^7 monocytes were prepared following the method described.[15] DC lysates were prepared as follows: 5×10^5 immature DC were washed with

PBS, centrifuged and resolved with 100 µl 2× lysis buffer (100 mM citrate/phosphate, 2 mM EDTA, 1% NP40, pH 5) and kept to lysis for 30 min at 4°C. After pelleting membrane fragments by centrifugation for 10 min at 13.000 rpm in an eppendorf-centrifuge, the supernatant contained crude cell lysate. 3 µg total endocytic protein of U937, 1.5 µg endocytic protein of primary monocytes or 3 µg DC cell lysate protein was incubated with 10 µM of **1**, **2** or **11** at ambient temperature for 30 min. Reactions were terminated by addition of SDS reducing sample buffer and immediate boiling. Samples were resolved by 12.5% SDS-PAGE gel, and then blotted on a PVDF-membrane and visualised using strept-HRP and the ECL-detection kit. In lane 1, E-64 was competed against **1**. The same amount of total endocytic protein was incubated with 40 µM of E-64 for one hour at ambient temperature, and **1** was added to a concentration of 10 µM. After 30 min, the reaction was terminated and samples were treated as described above.

Uptake and labeling in living cells *via* a pulse experiment: At least 1×10^6 DC per time point were pulsed in complete culture medium (300 µl per time point) for 1 hour with 10 µM **1** or **11** at 37°C. Additionally, at least 1×10^6 DC were incubated for the same time with 10 µM of either **1** or **11** at 4°C. After the pulse cells were washed to remove excess **1** or **11** at 4°C for four times by centrifugation, removing of supernatant and resolving the pellet in PBS. Cells were lysed with 50 µl of 2× lysis buffer (100 mM citrate/phosphate, 2 mM EDTA, 1% NP40, pH 7), supplemented with 100 µM free JPM-565. SDS reducing sample buffer was added to 50 µg total protein of each time point and immediately boiled.

References and notes

1 The work presented in this chapter was performed in close collaboration with M. Reich, C. Driessen and R. Brock at the University of Tübingen.

2 S. Fawell, J. Seery, Y. Daikh, C. Moore, L. L. Chen, B. Pepinsky and J. Barsoum, *Proc. Natl. Acad. Sci. USA*, 1994, **91**, 664-668.

3 D. Derossi, A. H. Joliot, G. Chassaing and A. Prochiantz, *J. Biol. Chem.*, 1994, **269**, 10444-10450.

4 H. Brooks, B. Lebleu and E. Vives, *Adv. Drug Delivery Rev.*, 2005, **57**, 559-577.

5 B. Gupta, T. S. Levchenko and V. P. Torchilin, . *Drug Delivery Rev.*, 2005, **57**, 637-651.

6 J. Temsamani and P. Vidal, *Drug Discov. Today*, 2004, **9**, 1012-1019.

7 S. Futaki, T. Suzuki, W. Ohashi, T. Yagami, S. Tanaka, K. Ueda and Y. Sugiura, *J. Biol. Chem.*, 2001, **276**, 5836-5840.

8 P. A. Wender, D. J. Mitchell, K. Pattabiraman, E. T. Pelkey, L. Steinman and J. B. Rothbard, *Proc. Natl. Acad. Sci. USA*, 2000, **97**, 13003-13008.

9 D. J. Mitchell, D. T. Kim, L. Steinman, C. G. Fathman and J. B. Rothbard, *J. Peptide Res.*, 2000, **56**, 318-325.

10 R. Fischer, K. Köhler, M. Fotin-Mleczek and R. Brock, *J. Biol. Chem.*, 2004, **279**, 12625-12635.

11 S. D. Conner and S. L. Schmid, *Nature*, 2003, **422**, 37-44.

12 K. Honey and A. Y. Rudensky, *Nature Rev. Immunol.*, 2003, **3**, 472-482.

13 D. Greenbaum, K. F. Medzihradszky, A, Burlingame and M. Bogyo, *Chem. Biol.*, 2000, **7**, 569-581.

14 S. W. Yeong and D. F. O'Brien, *J. Org. Chem.*, 2001, **66**, 4799-4802.

15 C. J. Schröter, M. Braun, J. Englert, H. Beck, H. Schmid and H. Kalbacher, *J. Immunol. Methods,* 1999, **227**, 161-168.

16 A. J. Barrett, A. A. Kembhavi, M. A. Brown, H. Kirschke, C. G. Knight, M. Tamai and K. Hanada, *Biochem. J.*, 1982, **201**, 189-198.

17 F. Sallusto, M. Cella, C. Danieli and A. Lanzavecchia, *J. Exp. Med.*, 1995, **182**, 389-400.

18 G.-P. Shi, J. S. Munger, J. P. Meara, D. H. Rich and H. A. Chapman, *J. Biol. Chem.*, 1992, **267**, 7258-7262.

19 A. Lautwein, T. Burster, A.M. Lennon-Duménil, H. S. Overkleeft, E. Weber, H. Kalbacher and C. Driessen, *Eur. J. Immunol.*, 2002, **32**, 3348-3357.

20 M. Cella, A. Engering, V. Pinet, J. Pieters, and A. Lanzavecchia, *Nature*, 1997, **388**, 782-787.

Summary and future prospects

Organic synthesis has provided tools to study the proteasome and cysteine proteases of the cathepsin family. In this thesis, new tools are presented, extending the possibilities to study these proteolytic activities. In **Chapter 1** of this thesis, an overview is presented of chemical labeling of biomolecules in living cells. Incorporation of unnatural reactive handles into native proteins, followed by employing these reactivities for the installation of a label, allowed selective tagging of a single protein in a cellular system. Another strategy of labeling a single protein was found in the creation of fusion proteins between the protein of interest and either a special peptide sequence or an active protein. The fusion proteins are labeled *via* the unnatural part with a fluorescent dye reacting specifically with the special peptide sequence. The third strategy explored, uses cell permeable activity based probes to introduce a latent bioorthogonal reactivity on active proteins, and after cell lysis, the reporter group of choice can be attached to the latent reactivity.

In **Chapter 2**, an approach is described to target active proteasomes in living cells and detect the proteasomes after cell lysis *via* two-step labeling. After modifying active proteasome with an irreversible inhibitor that was equipped with an azide moiety, the Staudinger-Bertozzi ligation was employed to biotinylate the proteasome selectively. This is the first example in which all proteolytic activities of the proteasome in a cellular system

are visualized. In this way, labeling artifacts as associated with cell lysis and measuring enzymatic activities in cell lysates are precluded. The scope of this method could be extended by the use of different irreversible proteasome activity probes and different ligation methodology, such as the Huisgen cyclisation as applied by Cravatt and coworkers.[1]

In **Chapter 3**, the two-step labeling approach presented in Chapter 2 was used to assess the specificity of two subunit-specific cell permeable inhibitors of the caspase-like activity of the proteasome. Based on studies towards reversible subunit specific proteasome inhibitors, an irreversible and cell permeable inhibitor was designed. Modification of the newly synthesized inhibitor with an azide moiety yielded a probe for the caspase-like activity of the proteasome. The selectivity of the newly synthesized probes was assessed both by peptidase assays and Western blots. At a lower concentration of the inhibitor, selectivity towards the immunoproteasome was observed. To investigate the role of the caspase like subunit of immunoproteasome in antigen processing, the epoxomicin[2] (**1**, Figure 1) warhead and the peptide boronate known from bortezomib[3] (**2**) combined with the used peptide sequence, might lead to even more potent inhibitors **3** and **4**.

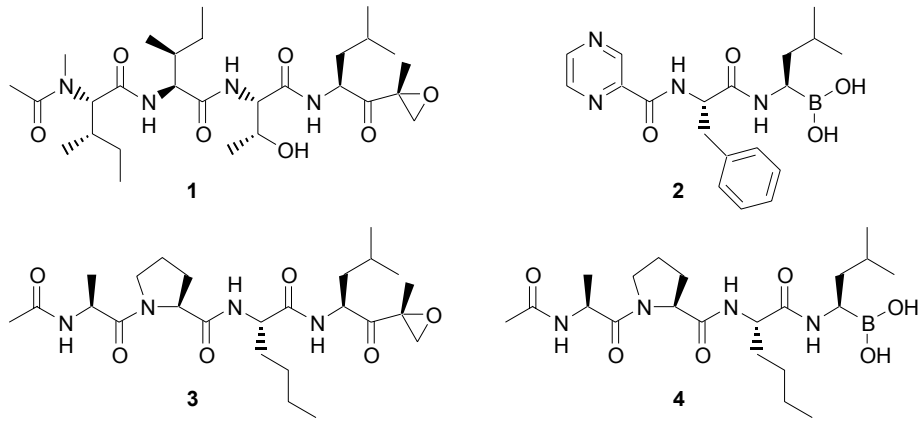
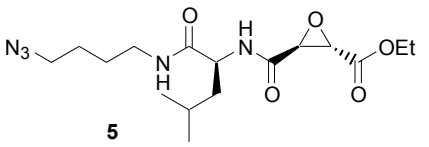


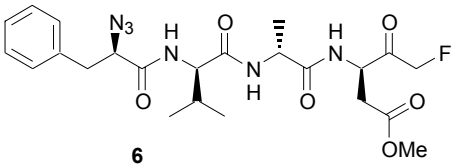
Figure 1. Possible immunoproteasome inhibitors based on other warheads

Furthermore, the technique of finding inhibitor targets by incorporation of a latent chemical group, like the azide moiety, in other irreversible inhibitors may contribute to elucidating the selectivity of the inhibitors and the identification of their minor targets. This could be of importance in recognizing side effects of drugs. In Figure 2, some azide-modified cysteine protease inhibitors are depicted, that might be used for two-step labeling. Compound **5** is based on E-64[4], a natural product that is an inhibitor of cysteine proteases

of the cathepsin family, whereas inhibitor **6** is an azide modification of ZVAD-FMK[5], an irreversible caspases inhibitor.



5



6

Figure 2. Azide-modified irreversible protease inhibitors

Chapter 4 shows the synthesis of an isotope-coded activity based probe for cysteine proteases of the cathepsin class. A set of deuterium encoded spacer molecules suitable for incorporation by solid phase peptide synthesis were developed. Modification of a known broad class cysteine protease inhibitor with these isotope coded spacers did not impair the inhibitory profile. It was expected that by labeling two cysteine protease samples by an isotope coded inhibitor pair, the relative activity of cysteine proteases between two samples can be established by mass spectrometry. However, the detection of the expected isotope-labeled active site fragment of cathepsin B proved troublesome, and therefore an attempt was made to synthesise this fragment. The isotope coded linker developed in this chapter could also find their way into the proteasome inhibitor or biotinylation phosphine in presented in **Chapter 2** to allow quantitation of proteasomal activity in living cells. Furthermore, the methodology could be extended by the use of other isotope labels, such as ^{13}C, which is reported to have less effect of chromatographic separations.[6]

Chapter 5 deals with the synthesis of a new cysteine protease probe. This probe is based on a known probe, extended both with an activated ester functionality that is reactive towards amines and an photolabile linker. The active ester was employed to conjugate the cysteine protease probe covalently to a model antigen. Presentation of this modified antigen to a professional antigen presenting cell could result in activity profiling of cysteine proteases involved in the proteolytic processing of the model antigen. The photolabile linker could be of importance to remove the remained of the model antigen from the labeled cysteine proteases, thereby simplifying the identification of the labeled cysteine proteases. Initial experiments involving the cleavage of the photolabile linker, the conjugation to ovalbumin as a model antigen and labeling of cathepsin B proved encouraging. However, the purity and homogeneity of the ovalbumin-probe conjugate could not be established yet.

Some peptide sequences are known for their capacity to translocate through cell membranes. In **Chapter 6**, a cysteine protease probe was fused to such a peptide sequence,

rendering a construct that was actively transported into cells. Meanwhile, the cysteine protease probe sampled the pathway for active cysteine proteases. In a set of uptake experiments, it was shown that the uptake pathway has at least an energy-independent component. In future research, different cell penetrating peptides could be fused to a cysteine protease probe, in order to see if all peptide sequences dictate the same uptake mechanism.

References

1. A. E. Speers, G. C. Adam and B. F. Cravatt, *J. Am. Chem. Soc.*, 2003, **125**, 4686-4687.
2. M. Hanada, K. Sugawara, K. Kaneta, S. Toda, Y. Nishiyama, K. Tomita, H. Yamamoto, M. Konishi and T. Oki, *J. Antibiot.*, 1992, **45**, 1746-1752.
3. J. Adams, M. Behnke, S. Chen, A. A. Cruickshank, L. R. Dick, L. Grenier, J. M. Klunder, Y.-T. Ma, L. Plamondon, and R. L. Stein. *Bioorg. Med. Chem. Lett.*, 1998, **8**, 333-338.
4. K. Hanada, M. Y. Tamai, M. Yamagishi, S. Ohmura, J. Samada and I. Tanaka, *Agric. Biol. Chem.*, 1978, **42**, 523–528.
5. L. Revesz, C. Briswalter, R. Heng, A. Leutwiler, R. Mueller and H.-J. Wuethrich, *Tetrahedron Lett.*, 1994, **35**, 9693-9696.
6. C. N. Filer, *J. Labelled Compd. Radiopharm*., 1999, **42**, 169-197.

Samenvatting

Het bestuderen van functionele eiwitten in de context van de levende cel is een belangrijk veld van onderzoek. Dit proefschrift richt zich op het ontwerp en de synthese van probes waarmee de enzymactiviteiten van het proteasoom en van cysteineproteases van de cathepsine klasse kunnen worden gedetecteerd en gekwantificeerd. In **Hoofdstuk 1** van dit proefschrift wordt een overzicht gegeven van de manieren waarop biomoleculen in levende cellen van een label kunnen worden voorzien. In natieve eiwitten kan een reactieve groep ingebouwd worden, die vervolgens wordt gebruikt om het eiwit selectief van een label te voorzien. Een andere strategie omvat de biosynthese van fusie-eiwitten die bestaan uit het te bestuderen eiwit en een speciaal peptide of actief enzym. Het toegevoegde deel van het fusie-eiwit maakt het aanbrengen van een label mogelijk. De derde strategie maakt gebruik van speciaal ontworpen synthetische probes, die celpermeabel zijn. Deze probes binden covalent aan het te onderzoeken eiwit en zijn bovendien voorzien van een bio-orthogonale groep die inert is in de cellulaire omgeving, maar na cellysis wordt gebruikt voor het aanbrengen van een label.

In **Hoofdstuk 2** worden de synthese en toepassing van een nieuw type probe beschreven. Een bekende cel permeabele en irreversibele remmer van het proteasoom werd uitgerust met een azide als bio-orthogonale groep. Na incubatie van cellen met deze probe

en cellysis, kon het proteasoom middels de Staudinger-Bertozzi ligatie selectief worden gebiotinyleerd. Met deze tweestapsprocedure konden voor het eerst alle proteolytische activiteiten van het proteasoom in de levende cel zichtbaar worden gemaakt en werden mogelijke artefacten, die kunnen optreden bij het lyseren van cellen en het meten van enzymactiviteiten in cellysaten, voorkomen.

Het proteasoom vertoont verschillende soorten enzymatische activiteit, die gelokaliseerd zijn in afzonderlijke subunits. In **Hoofdstuk 3** worden twee remmers gepresenteerd voor één van deze activiteiten van het proteasoom, namelijk degene die wat substraatherkenning betreft op caspases lijkt. Een van deze nieuwe remmers bevat een azidegroep, en de specificiteit van deze remmer voor de caspase-achtige activiteit van het proteasoom werd aangetoond middels de tweetrapsprocedure uit Hoofdstuk 2. Bij een lagere concentratie van de remmer werd bovendien selectiviteit voor het immunoproteasoom waargenomen.

Hoofdstuk 4 beschrijft de vaste drager synthese van twee paar probes voor cysteine proteases uit de cathepsineklasse. Kenmerkend voor deze probes zijn de speciale linkers, namelijk één in een lichte vorm, en de ander in een zware vorm, waarin een aantal waterstofatomen is vervangen door deuterium. Het inbouwen van de nieuwe linker in een bekende cathepsineremmer had geen invloed op het remmingsprofiel. Met behulp van een lichte en een zware probe zou de relatieve cysteine proteaseactiviteit tussen twee samples middels massaspectrometrie kunnen worden vastgesteld. Voorwaarde voor succes van deze methode is de massaspectrometrische detectie van het door de remmer gealkyleerde active site fragment van het enzym.

Hoofdstuk 5 behandelt de synthese van een nieuwe cysteine protease probe. Deze probe omvat een bekende remmer voor cysteine proteases die via een fotolabiele linker werd uitgebreid met een geactiveerde ester. De actieve ester werd gebruikt om een covalente binding tot stand te brengen tussen de probe en de aminogroepen van een modelantigeen. Door dit gemodificeerde antigen aan te bieden aan een professionele antigen presenterende cel zouden de cysteineproteases, die een rol spelen bij de verwerking van dit antigen voor presentatie, in kaart gebracht kunnen worden. De identificatie van deze proteases zou vereenvoudigd kunnen worden door het verwijderen van het modelantigeen door de invloed van ultraviolet licht op de fotolabiele linker. De eerste

labelingsexperimenten verliepen echter niet eenduidig, wat mogelijk een gevolg is van een inhomogeniteit van het gemodificeerde antigen.

Sommige peptidesequenties zijn bekend om hun eigenschap andere moleculen door celmembranen heen te transporteren. Het mechanisme van dit transport is nog niet volledig opgehelderd. In **Hoofdstuk 6** wordt een benadering gepresenteerd die het inzicht in het opnamemechanisme kan vergroten. Een cysteine protease probe werd geconjugeerd aan een celpenetrerende peptidesequentie, opdat dit construct actief cellen in wordt getransporteerd. Tijdens het transport bindt de probe aan de actieve cysteine proteases die het onderweg tegenkomt. In een reeks experimenten werd aangetoond dat het opnamepad minimaal een energieonafhankelijke component kent. In toekomstig onderzoek zouden verschillende celpenetrerende peptiden aan een cysteine protease probe kunnen worden geconjugeerd om te zien of alle peptide hetzelfde opnamepad volgen.

List of publications

Montmorillonite K-10 clay assisted transformation of vinylketoses into spirochromans and arylketoses
P. A. V. van Hooft, P. F. van Swieten, G. A. van der Marel, C. A. A. van Boeckel and J. H. van Boom, *Synlett*, 2001, 269-271.

Chemistry in living cells: detection of active proteasomes by a two-step labeling strategy
H. Ovaa, P. F. van Swieten, B. M. Kessler, M. A. Leeuwenburgh, E. Fiebinger, A. M. C. H. van den Nieuwendijk, P. J. Galardy, G. A. van der Marel, H. L. Ploegh and H. S. Overkleeft, *Angew. Chem., Int. Ed.*, 2003, **42**, 3626-3629.

Analysis of the proteome of a living cell by labelling the proteins in the intact cell
H. S. Overkleeft, M. A. Leeuwenburgh, P. F. van Swieten, H. Ovaa, B. M. Kessler and H. L. Ploegh, *International Patent*, 2004.

Development of an isotope-coded activity-based probe for the quantitative profiling of cysteine proteases
P. F. van Swieten, R. Maehr, A. M. C. H. van den Nieuwendijk, B. M. Kessler, M. Reich, C.-S. Wong, H. Kalbacher, M. A. Leeuwenburgh, C. Driessen, G. A. van der Marel, H. L. Ploegh, H. S. Overkleeft, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 3131-3134.

Bioorthogonal organic chemistry in living cells: novel strategies for labeling biomolecules
P. F. van Swieten, M. A. Leeuwenburgh, B. M. Kessler and H. S. Overkleeft, *Org. Biomol. Chem.*, 2005, **3**, 20-27.

Chemical proteomics profiling of proteasome activity
M. Verdoes, C. R. Berkers, B. I. Florea, P. F. van Swieten, H. S. Overkleeft and H. Ovaa, *Methods Mol. Biol.*, 2006, **328**, 51-69.

Curriculum vitae

Paul Franciscus van Swieten werd op 23 oktober 1977 geboren te De Lier. Na het behalen van het VWO-diploma aan het Sint Stanislascollege te Delft in 1996 werd in september van dat jaar begonnen met de studie scheikunde aan de Universiteit Leiden. In 1997 werd de propedeuse behaald, en van september 1999 tot juli 2000 werd er in het kader van de hoofdvakstage onderzoek verricht binnen de vakgroep Bio-organische Synthese onder leiding van prof. dr. J. H. van Boom. Dit onderzoek omvatte de synthese van een kleine bibliotheek van geglycosyleerde aromaten met schimmelwerende werking. Deze hoofdvakstage werd bekroond met de Unilever Research Prijs in 2000. Daarnaast werd van september 2000 tot maart 2001 een onderzoeksstage volbracht bij de research and development afdeling van Organon te Newhouse, Schotland, onder leiding van dr. J. Cai en dr. D. C. Rees. Het doctoraal diploma werd behaald in mei 2001. In mei en juni van dat jaar werd er bij TNO Rijswijk in de groep van dr. D. Noort onderzoek gedaan naar derivatisering van histidine en gealkyleerd histidine met fluorescente agentia.

Van juli 2001 tot september 2005 werd als assistent in opleiding gewerkt aan het in dit proefschrift beschreven onderzoek binnen de vakgroep Biosyn onder supervisie van prof. dr. H. S. Overkleeft en prof. G. A. van der Marel. In januari 2003 werd een lezing gegeven tijdens de Holland Research School for Molecular Chemistry Conference te Amsterdam. Het werk, beschreven in hoofdstuk 3, werd in oktober 2003 tijdens de driedaagse bijeenkomst van de Studiegroepen Ontwerp en Synthese, Structuur en Reactiviteit en Biomoleculaire Chemie, te Lunteren bekroond met een Shell Posterprijs. In januari 2004 werd deelgenomen aan de Ilab Initiative workshop in Wiesbaden, Duitsland.

Vanaf september 2005 wordt als onderzoeker binnen een samenwerkingsverband tussen ISA Pharmaceuticals, de sectie IHB van het LUMC en de vakgroep Biosyn gewerkt aan de synthese van immuun stimulerende agentia.

Nawoord

Het interdisciplinaire werk beschreven in dit proefschrift kon slechts tot stand komen door intensieve samenwerking tussen wetenschappers. Na de synthese van elk van de diverse probes uit dit proefschrift, werden deze probes in verschillende laboratoria gebruikt in een lopend biochemisch onderzoek. Soms bleek het ontwerp nog niet optimaal, en werd dit na doornemen van de resultaten aangepast.

I appreciated the collaboration with Michael Reich, during his stay at the Biosyn lab. This collaboration got a valuable extension during the discussions and the practical work performed on the cell penetrating cysteine protease probe. I am indebted to Ana-Maria Lennon-Duménil, Benedikt Kessler, René Maehr, Hidde Ploegh, Alexei Kisselev, Emlyn Samuel, Christoph Driessen, and Huib Ovaa for their biochemical efforts. Micha Slegt en Bobby Florea hebben bijgedragen aan het biochemische en fotochemische werk met betrekking tot de fotolabiele cysteineproteaseprobe.

De wetenschappelijke discussies met collegae van de vakgroep Bioorganische Synthese heb ik als stimulerend ervaren. It was a pleasure to me to work with Nacho Vaya, Chung-Sing Wong, Rosa Orient Hernández, Jimmy van Rijn, Carol de Dicastillo Bergamo and Nerea Picazas, as they joined the Biosyn group to make a contribution to the work described in this thesis. Binnen de proteomics studiegroep waren de kennis en ervaring van Michiel Leeuwenburgh en Rian van den Nieuwendijk van grote waarde.

Hans van den Elst en Nico Meeuwenoord hebben wezenlijke bijdragen geleverd aan dit proefschrift middels het meten, synthetiseren en zuiveren van intermediairen en eindproducten. Op Fons Lefeber en Cees Erkelens kon ik rekenen voor het meten van NMR.