General Introduction

1.1 Introduction

Proteomics research focuses on the study of proteins, their functioning and interactions with other biomolecules in the context of complex biological samples. Rather than focusing on a single protein, proteomics research takes on large numbers of distinct proteins at the same time, in an ideal setting the whole pool of proteins (called the proteome) expressed at a given time by an organism or cell type.\textsuperscript{1,2} In the overall study of biological systems proteomics is situated between genomics (study of the genome) and metabolomics (study of metabolites produced in cellular processes). The ultimate goal in proteomics is the complete understanding of each single protein in all physiological processes, which is of great importance in understanding pathological states. Since it is extremely difficult to study all proteins and their properties at once, proteomic research is usually performed by taking on one specific subset of proteins from two or more different biological systems (e.g. healthy cells and infected cells) in a comparative study. Traditionally, this is done by separation of the protein subset of choice, for example by 1D/2D gel-electrophoresis and immunoprecipitation, followed by determination of its level of abundance. An attractive alternative strategy developed for proteomics research is activity-based protein profiling (ABPP).\textsuperscript{3,4} This strategy is based on the determination of an enzyme’s activity rather than its levels of expression, which is important since an enzyme’s abundance is not necessarily linked to its activity in biological processes. A reason for this is that a large number of enzymes is translated as inactive pro-enzymes and therefore need to be transformed into their active form. In
addition, the enzyme’s state can be switched between active and inactive by post-translational modifications, such as phosphorylation, glycosylation, acetylation, ubiquitination and methylation. In the first part of this introduction the general idea of ABPP will be outlined. The major part of the introduction deals with an interesting aspect of this strategy in which a photophore is used to profile specific enzymes or enzyme families. The properties of the most commonly applied photophores and their use in chemical biology research are discussed. Finally, the aim and outline of this Thesis will be described at the end of the introduction.

1.2 Activity- and affinity-based protein profiling

Activity-based protein profiling makes use of relatively small organic molecules to label a specific enzyme (class) in their active state. These organic molecules are called ‘activity-based probes’ (ABPs) and comprise three major elements. The recognition unit directs the ABP to the target enzyme and is designed to resemble structural and functional motives of the natural substrate of the target enzyme. It can often be designed such that a number of related enzymes within a family can bind. Closely attached to the recognition part is a reactive group (also termed ‘warhead’), which reacts in the enzyme’s active site and thereby establishes a covalent bond between the ABP and enzyme. Depending on the type of enzyme and reactive group this interaction can be either reversible or irreversible. The recognition element and the warhead are linked, most often via a spacer, to a reporter group or ‘tag’, which allows visualization and/or purification of the bound enzyme. The most commonly used tags are fluorophores, radioactive labels and biotin, of which the latter can be used for both visualization and purification purposes (Figure 1, top).

![Figure 1. Schematic representation of an ABPP experiment. Both the 1-step (top) and the 2-step (bottom) labelling strategies are shown.](image-url)
An ingenious extension to ABPP has been developed for those cases in which the reporter group hampers the interaction between ABP and enzyme or drastically lowers the ability to cross the cell’s membrane, which is especially of interest for labelling in living cells. In this approach (referred to as two-step labelling) the ABP’s tag is replaced by a ligation handle, which can be connected to the reporter group after the enzyme has been captured (Figure 1, bottom). A requirement for this ligation handle is that it is unreactive towards all functionalities present in a biological sample (bioorthogonal) and the most popular ligation handles are the azide and the (terminal) alkyne. The ligation reactions used (shown in Figure 2) are the Staudinger-Bertozzi ligation (I) and the Huisgen 2,3-dipolar cycloaddition or ‘click’ reaction, which can be divided into copper-catalyzed (II) and strain-promoted (III). Recently Boons and co-workers developed a new reagent, which allows for a strain-promoted click reaction after irradiation with light and is therefore termed the ‘photoclick’ reaction (IV). In addition, the use of the Diels-Alder reaction as alternative ligation reaction is subject to growing attention and of the different Diels-Alder type reactions, the inverse-electron-demand variation appears the most promising for in vivo labelling.

The ABPP strategy as shown in Figure 1 is especially well suited for those enzymes that contain a nucleophilic amino acid side chain residue in their active site (e.g. serine, cysteine, threonine), which is responsible for the enzyme’s catalytic activity. The ABP reactive group which binds the target enzyme is designed such that it reacts with this nucleophilic residue to form a covalent bond and is therefore named ‘electrophilic trap’. Examples of enzymes targeted with this strategy are cysteine proteases, serine hydrolases and proteasome subunits. A difficulty arises for enzymes that do not rely on a nucleophilic residue in their active site, which precludes the use of an electrophilic trap. Among these are the metalloproteases, histone deacetylases (both of which employ a water molecule for their catalytic activity) and kinases. A good
alternative for the use of an electrophilic trap is the so-called photoaffinity labelling (PAL), in which the probe used is commonly referred to as ‘affinity-based probe’ (AfBP). The basic principle is shown in Figure 3. In this approach the AfBP binds the target enzyme in a reversible manner, either via non-covalent interactions (electrostatic, hydrophobic) or via a reversible covalent bond. Although these interactions can be relatively strong, they can not withstand harsh denaturing conditions often applied in biochemical protocols. An additional feature of the AfBP is the introduction of a photoreactive group (also termed photophore or photocrosslinker), which forms an irreversible covalent bond between probe and enzyme upon activation by light. In principle, photoaffinity labelling probes do not necessarily label active enzymes, however the probe can be designed as such that it has to enter the active site of an enzyme prior to photocrosslinking. Therefore, ABPP and PAL often go hand-in-hand.

Figure 3. Basic principle of photoaffinity labelling (PAL) using an affinity-based probe (AfBP).

1.3 Photoaffinity labelling

In order to use an AfBP for labelling of (active) enzymes in biological environments the photoreactive group must meet with certain criteria. First, the photophore has to be stable towards the various conditions a biological sample may be exposed to, as well as the intrinsic reactivity of the sample content, and must only be activated upon irradiation with light of a specific wavelength, which may not damage the biological system ($\lambda_{\text{act}} > 300 \text{ nm}$). Second, the generated reactive species needs to have a shorter lifetime compared to the lifetime of the studied enzyme-substrate complex in order to limit non-specific labelling. It is important that the activated species reacts with any chemical entity in close proximity, regardless of its nature (including relatively unreactive C-H bonds), and forms a stable covalent adduct. Finally, the photoreactive moiety must be relatively small, compared to the probe, so that it does not negatively influence the binding mode or activity of the AfBP towards the enzyme.

The use of PAL in enzyme modifications was first described in 1962, where Westheimer and co-workers reported on the use of a diazoacetyl group to inactivate chymotrypsin. Considerable research on the development of new PAL reagents has taken place ever since, but only a few number of photophores, which largely meet the above-mentioned requirements, are being used nowadays in AfBPs. These are aryl azides (first reported use in 1969), diazirines (1973) and benzophenones (1973). The chemistry of these three photoreactive groups after photolysis as well as their use in recently reported AfBPs will be discussed.
1.3.1 Aryl azide

Upon activation of an aryl azide (1, see Scheme 1) by irradiation with light of the appropriate wavelength, molecular nitrogen (N₂) is expelled and a singlet nitrene (2) is formed initially. This high energetic, highly reactive species has a short lifetime (~10⁻⁴ s) and is quickly converted into other intermediates. Intersystem crossing (ISC) leads to a triplet nitrene (3), which is about 20 kcal/mole lower in energy. A major difference between the two nitrene states is their nature of reactivity. Singlet nitrenes behave like electrophiles and can readily undergo an insertion reaction with C-H bonds, whereas the triplet state can be seen as a diradical, which first abstracts a hydrogen radical from a nearby C-H followed by coupling to the formed carbon radical. Although they react via two different mechanisms the product is the same (4). Singlet nitrenes can also undergo a rapid rearrangement into the corresponding benzazirine (5), which can further rearrange into dehydroazepine (7). Both these species are long-lived electrophiles and can react with a nearby nucleophile, which results in compounds 6 and 8 respectively. Two observed side-reactions that are not to be ignored when the aryl azide is applied in PAL are aerobic oxidation of the triplet nitrene to the corresponding nitro species 10 and reduction of the initial aryl azide to the amine 11 by dithiols, such as DTT.

Scheme 1. Possible reaction mechanisms of the reactive intermediates formed after photolysis of aryl azides.

Aryl azides can be easily prepared from their corresponding amines in one or two steps. Three examples are given in Scheme 2A. The most common method is the diazotization of the amine with sodium nitrite under acidic conditions, followed by addition of sodium azide in an aqueous medium (route a). In 2003 the synthetic method was improved by application of triflyl azide (TfN₃), which allowed a one-step conversion and higher yield (route b). Recently, the development of imidazole-1-sulfonyl azide 15 was reported, which proved to be a more stable reagent and allowed a conversion under mild conditions (route c).
A major drawback of phenyl azides that their maximum absorption wavelength being below 300 nm, since electromagnetic irradiation at these wavelengths can substantially damage the biological system. Consequently, a large number of substituted aryl azides have been made and evaluated for their absorption properties. In general most substituents ortho to the azide are to be avoided, since they can lead to undesired cyclizations after photolysis. It has been found that introduction of electron withdrawing substituents (e.g. nitro, hydroxyl and acyl groups, for example see compound 16 in Scheme 2B) has the dual effect of increasing the molar absorptivity and red-shifting the maximum absorption wavelength, both of which positively influence the photoactivatable properties. In addition, it has been found that (per)fluorinated aryl azides (such as 17) rearrange more slowly from the singlet nitrene species to the benzazirine and dehydroazepine, which leads to more efficient insertion reactions. The main advantage of aryl azides is their relatively small size and the possibility to incorporate them into natural biological compounds, such as phenylalanine 18 and adenosines 19 and 20, without significant alteration of the original structure.

Due to the many possible reaction pathways after irradiation (including capturing of the reactive intermediates by the solvent) cross-linking yields are often low (<30%). Arguably, the popularity of the aryl azide moiety in PAL studies is based on its relative ease of preparation and incorporation rather than on its photochemical properties.

Scheme 2. Preparation and examples of aryl azides.

(A) Three possible routes for the conversion of an aryl amine into its aryl azide: (a) via diazotization, (b) by the use of triflyl azide and (c) with imidazole-1-sulfonyl azide 15. (B) Some examples of substituted aryl azides.

An extensive study towards matrix metalloproteinases (MMPs) with the use of an aryl azide modified AβP was recently reported by Dive and co-workers. MMPs are metallo-proteases which reside in the extracellular matrix and are responsible for degradation of extracellular matrix material. Their mode of action depends on a Zn$^{2+}$ ion...
in the active site which coordinates the scissile bond carbonyl of the substrate and a water molecule. As a result, the carbonyl becomes more electrophilic and is subsequently hydrolysed. The fact that there is no formation of a covalent bond between enzyme and substrate during the proteolysis makes this class of enzymes an interesting target for photo-affinity labelling. The authors describe the use of radio-labelled compound 21 (Scheme 3A), a potent, subnanomolar MMP inhibitor, to label and visualize purified human MMPs. The aryl azide photoreactive group is located at the P1’ pocket, which leads to a tight interaction with the enzyme’s cavity. A big difference in terms of labelling efficiency and sensitivity between several MMPs was found, with MMP-12 giving the best results. The estimated crosslinking yield was ~42% after two minutes of irradiation, based on silver staining and as little as 2.5 fmol MMP-12 could be detected. In a second study the specifics of photocrosslinking were further explored, using compound 21 in combination with mass spectrometry and site-directed mutagenesis. Interestingly, the ε-amine side chain substituent of Lys241 in MMP-12 appeared to play a crucial role in the photocrosslinking. Two possible covalent constructs were proposed (Scheme 3B), but due to the fact that they have the same molecular weight it was impossible to distinguish between these using mass spectrometry. In theory, some other constructs are possible (see Scheme 1), however

**Scheme 3.** Affinity-based probes targeting MMPs from studies by Dive et al.

(A) Photoreactive A\(\beta\)P containing a tritium label. (B) Possible constructs formed between hMMP-12 and 21 after photolysis proposed by Dive et al. (C) Structures of A\(\beta\)Ps with or without a photophore for pull-down of active MMPs.
they were not mentioned by the authors. The lysine at position 241 is not conserved throughout the MMP family and photocrosslinking of 21 to other MMPs was therefore further explored.\textsuperscript{55} MMP-3 (containing His in position 241) and MMP-9 (Arg in position 241) could also be labelled, although with a lower overall efficiency. In addition, labelling performed at different pH values indicated that a more basic environment resulted in more efficient crosslinking. These results led to the conclusion that the nucleophilicity of the residue at position 241 plays a key role in the photoaffinity labelling. This conclusion was further substantiated by the finding that mutants of hMMP-12 (Ala\textsuperscript{241} and Thr\textsuperscript{241}) gave no labelling whatsoever.

In addition to the attempts of unravelling the modification site, Dive and co-workers also constructed two biotinylated AfBPs, 22 and 23 (Scheme 3C) and used these to study the difference in affinity- and photo-affinity MMP enrichment from a complex proteome.\textsuperscript{56} For this, tumor extracts were spiked with hMMP-12 and hMMP-8, after which compounds 22 and 23 were applied followed by streptavidin-coated magnetic beads for MMP pull-down. Affinity-based labelling with 23 appeared superior to photo-affinity-based labelling with 22 in terms of quantity of captured MMPs, although it should be noted that the compounds are structurally different and 23 is a 100 fold more potent MMP-8 and MMP-12 inhibitor.

1.3.2 Diazirine

One of the greatest advantages of the photolabile diazirine group over aryl azides is that all its members absorb most efficiently at a wavelength of 350-380 nm. This is well above the 300 nm limit (vide supra) and therefore no significant damage to the biological system will occur. The most important reactions that occur after photolysis of 3-aryl-3\textit{H}-diazirines are shown in Scheme 4. When a diazirine (such as 24 or 25) is irradiated molecular nitrogen is expelled and a singlet carbene is formed (26). Competitively, a substantial amount (>30%) of the diazirine is converted into diazoisomer 27. This diazo compound can be converted into the singlet carbene under the influence of light, however at the wavelengths normally used (360 nm) this process is relatively slow. For this reason the diazo species is relatively long-lived and thus has time to diffuse resulting in either aspecific labelling or hydrolysis. This problem was largely tackled when Brunner and co-workers reported the development of 3-aryl-3-(trifluoromethyl)-3\textit{H}-diazirine 25.\textsuperscript{57} The strong electron-withdrawing properties of the trifluoromethyl group stabilize the diazoisomer, which makes it almost completely resistant towards undesired ‘dark’ reactions. Singlet carbene 26 is a very short-lived species (t\textsubscript{1/2}~1 ns) and can be transformed into triplet carbene 34 via intersystem crossing (ISC). Singlet and triplet carbenes display a similar behaviour compared to their corresponding nitrenes. A singlet carbene can react as an electrophile, nucleophile or ambiphile, depending on the nature of its substituents, whereas triplet carbenes behave like diradicals. The formed singlet carbenes can give fast insertion reactions, in which they do not discriminate much between different reaction sites. Insertions into hydroxyl groups (giving 28) usually do give a higher yield compared to C-H insertions (29).\textsuperscript{58} Insertion into a primary or secondary N-H bond (30) can lead to an undesired side
reaction. The formed construct easily expels HF, thereby giving enamine 31, which is in equilibrium with imine 32. In aqueous environments, such as a physiological sample, these species are subsequently hydrolysed into the corresponding ketone, with loss of the captured substrate as the result.\(^5\) The triplet carbene can react with C-H bonds analogues to triplet nitrenes. Initial hydrogen abstraction leads to radical intermediate 35, which either reacts with the formed carbon radical to give a netto C-H insertion (29) or abstracts a second hydrogen from another C-H bond, resulting in a reduction (36). Another undesired side reaction occurs when the triplet carbene is oxidized by molecular oxygen (a ‘notorious scavenger of triplet states’) to the corresponding ketone 37.\(^3\) In general, unsubstituted 3-alkyl-3\(H\)-diazirines should be avoided since their corresponding carbenes are prone to hydride shift, which results in an olefin (see the insert in Scheme 4).\(^3\)

Scheme 4. Possible reactions of the intermediates formed after photolysis of 3-aryl-3\(H\)-diazirines.

Although the diazirine group itself is relatively small, aryl diazirines are quite bulky, but they can be incorporated into molecules with a structure similar to naturally occurring compounds. Furthermore, they are quite stable towards a wide variety of conditions, including strongly acidic, strongly basic, oxidative and several reducing agents, which is a big advantage of diazirines compared to aryl azides. Drawbacks of diazirines are the formation of substantial amounts (>30%) of the diazo species after photolysis and the intrinsic efficient reactivity of the singlet carbene with O-H bonds, which often leads to scavenging of the reactive species by water. Also, the synthesis of diazirines is somewhat complicated compared to aryl azides. The synthetic scheme often applied for the preparation of 3-aryl-3-(trifluoromethyl)-3\(H\)-diazirine nowadays is shown in Scheme 5A.\(^5\) It starts by lithiation of an aryl bromide (38), which subsequently reacts with N-(trifluoroacetyl)piperidine 39 (easily prepared from trifluoroacetic anhydride and piperidine) under the formation of trifluoroacetophenone 40. Next, the ketone is converted into the corresponding oxime 41, after which the hydroxyl group is
converted into its tosylate (42). Reaction with liquid ammonia (usually under pressure) allows the instalment of the diaziridine group (43). Subsequent oxidation with iodine finally results in the diazirine (44). This five-step reaction sequence is especially well compatible with acid labile protective groups, which are often used to protect and/or install functionalities at the R position.

Some interesting examples of diazirines used in biologically relevant studies are shown in Scheme 5B. Tritium functionalized adamantyl diazirine 45 was used for selective labelling of intrinsic membrane proteins in human erythrocytes. Despite the presence of α-hydrogen atoms, the formed carbene is not prone to hydride shift (see Scheme 4) due to the constraints of this caged ring system. However, photolabelling of species is reported to be quite inefficient, probably due to its propensity to intramolecular C–H insertion reactions and reaction with water.30,34 Among the diazirine functionalized amino acids developed, modified L-phenylalanine (Phe(Tmd), 46) is the most popular one. Its first stereoselective preparation was reported by Nassal in 198460 and it has been used extensively ever since.33,36 Recently, the synthesis of its D-phenylalanine analogue was reported and this compound was used to probe the sweet taste receptor.61 In 2005 Thiele and co-workers reported the chemo-enzymatic synthesis of diazirinized leucine (47) and methionine (48), which were abbreviated as ‘photo-Leu’ and ‘photo-Met’. It was shown that these unnatural amino acids could be incorporated into proteins by a eukaryotic cell with genetically unmodified translational machinery and this methodology was applied in the identification of protein-protein interactions in living cells.62 The synthesis of photo-Met was optimized by Muir and co-workers in 2007, who circumvented the enzymatic resolution step and incorporated the unnatural amino acid into a protein using solid phase peptide synthesis (SPPS) and expressed protein ligation (EPL) strategies.63

Scheme 5. Preparation and examples of diazirines.

(A) Synthetic scheme for preparation of 3-aryl-3-(trifluoromethyl)-3H-diazirines. (B) Some examples of diazirine functionalized compounds.
Some interesting examples of AFBPs containing the diazirine moiety, which were used to target active metalloenzymes, are shown in Figure 4. Yao and co-workers reported a library of hydroxamate oligopeptides 49, with varying types of amino acids at the P1 position. The hydroxamate moiety is a potent zinc binding group (ZBG). The oligopeptides were modified with an N-terminal aryl diazirine for covalent modification of the target enzyme and a fluorescent label (Cy3) for visualization. They were able to selectively label and visualize thermolysin (a Zn$^{2+}$ dependent metalloprotease found in gram-positive bacteria) spiked in a crude yeast extract after covalent modification by irradiation for twenty minutes. In addition, the library of compounds was incorporated in a large-scale profiling study, in which the 'fingerprint' labelling of twelve yeast metalloproteases towards probe library 49 was determined.

In two other studies photoaffinity labelling of metalloenzymes was combined with two-step modification and visualization using the Cu(I) catalyzed click reaction. Qiu et al. reported the use of succinylhydroxamate oligopeptide 50 containing an azide functionality in labelling MMP-2 (a secreted Zn$^{2+}$ dependent matrix metalloproteinase) both as a purified enzyme and in a mouse melanoma B16-F10 cell culture medium. Visualization of the photocaptured construct was achieved by a click reaction to biotin-propargylamide and subsequent streptavidin-HRP Western-blotting. The same group reported the development of one-step and two-step AFBPs 51a and 51b, the design of which was based on parent compound L288. The latter is a potent inhibitor (IC$_{50}$ 0.13 µM) of type I methionine aminopeptidase (MetAP1), a cobalt dependent metalloenzyme expressed by both prokaryotic and eukaryotic cells, and which removes N-terminal (initiator) methionine from polypeptides. The modifications made led to a slight decrease in inhibitory potency compared to L288. Incubation of overexpressed E. coli MetAP1 in E. coli cell lysate with compound 51b followed by UV irradiation, click reaction and Western blotting revealed labelling of the target enzyme, which could be competed away by L288. Interestingly, incubation with one-step probe 51a resulted in substantial non-selective labelling, which overwhelmed the MetAP1 signal. Apparently, the two-step probe is much more selective in this case.

Figure 4. Examples of AFBPs containing the diazirine moiety used to study active metalloenzymes.
1.3.3 Benzophenone

A major advantage of benzophenones is that they can be excited at wavelengths of 350-360 nm, just like diazirines. The possible reaction pathways of benzophenones after photolysis are shown in Scheme 6. Absorption of a photon of the proper wavelength by a benzophenone (52) initially results in the formation of a triplet state benzhydrid diradical (53). The formation of the triplet diradical is reversible and it can exist as long as 120 µs before relaxing back to its ground state in the case that it is unable to find a reaction partner. The first reaction step of the formed reactive species is abstraction of a hydrogen and the reaction rate is therefore dependent on the nature of a nearby X-H bond. In general, the diradical is more reactive towards C-H bonds than O-H bonds. Especially those C-H bonds that form relatively stable carbon radicals are prone to react and these include benzylic positions, amino acid α-positions, hydrogen atoms adjacent to heteroatoms and tertiary carbon centres. Reactions with aromatic and vinylic C-H bonds have not been reported. All amino acids can react, although it has been shown that there is a preference for the ε-H in methionine when the benzophenone moiety is mobile enough to choose. Abstraction of a hydrogen from an amino acid α-centre (54) by 53 results in the formation of a ketyl (55) and an alkyl radical (56), which recombine fast to form a benzhydrol (57). In the case of glycine, there is a possibility of elimination of water under the formation of olefin 58. A big advantage of the benzophenone group is that its photoactivated counterpart is more reactive towards C-H bonds compared to nitrenes and is less prone to intramolecular rearrangements than carbenes. Also, when the diradical inserts into water the corresponding hydrate (59) is formed. This species quickly dehydrates to form the ketone again, which can be recycled under irradiation to the diradical species. This ability of benzophenones to ‘search’ for a good reaction centre is a big advantage in terms of crosslink efficiency, however when the reactive species is not quenched in time there is a big chance of aspecific labelling, especially when a corresponding A:B:P is not interacting with the target enzyme, but moves around freely in the medium. A possible

Scheme 6. Chemistry of benzophenones after photolysis.
side reaction which can take place is the homodimerization of ketyl 55 to form benzopinacol 60, however due to the relatively big difference in reaction rates of hydrogen abstraction and recombination, normally only a very small amount of this is formed.

In contrast to aryl azides and diazirines, the most commonly used benzophenone building blocks are commercially available. Benzophenone substituted amino acid analogues were also created, similar to what was previously discussed for aryl azides and diazirines. Not surprisingly, the most studied amino acid derivative is the one derived from phenylalanine, commonly abbreviated as L-Bpa (61 in Scheme 6). Although the benzophenone group may seem like the ideal photocrosslinking reagent, it also suffers from some drawbacks. It is relatively bulky, which can negatively influence the interaction between enzyme and substrate. Also, the resulting steric hindrance can give rise to a discrimination between reaction sites and, as a result of that, can lead to a non-specific labelling. Finally, irradiation for prolonged times (>30 minutes) is often needed in order to obtain a reasonable crosslinking efficiency.

In recent literature many examples of AβPs containing benzophenones can be found. A first example concerns the study of histone deacetylases (HDACs). These enzymes catalyze the hydrolysis of acetylated lysine amine side chains in histones and are thus involved in the regulation of gene expression. There are approximately twenty human HDACs which are divided into three classes (I, II and III). Class I and II HDACs are zinc-dependent metallohydrolases that do not form a covalent bond with their substrates during their catalytic process, which is similar to MMPs. It has been found that hydroxamate 62 (SAHA, see Figure 5) is a potent reversible inhibitor of class I and II HDACs. In 2007 Cravatt and co-workers reported the transformation of SAHA into an AβP, by instalment of a benzophenone and an alkyne moiety, which resulted in SAHA-BPyn O (63). They showed that the probe can be used for the covalent modification and enrichment of several class I and class II HDACs from complex proteomes in an activity-dependent manner. In addition, they identified several HDAC associated proteins,

Figure 5. Examples of benzophenone modified two-step labelling probes to study HDACs.
possibly arising from the tight interaction with HDACs. Also, the probe was used to measure differences in HDAC content in human disease models. Later they reported the construction of a library of related probes and studied the differences in HDAC labelling. Their most interesting finding was that the labelling efficiency is not directly linked to a compound’s inhibitory potency. For example, compounds 64a, 64b and 65, containing Bpa (Figure 5), showed a higher potency (in terms of inhibition of HDAC activities) than 63, however the latter compound proved to be superior in HDAC labelling. A similar disparity in inhibitory potential and photo-affinity labelling will be addressed in Chapter 4. A similar approach was reported by Xu et al. in 2009, in which potent HDAC class I (comprising HDACs 1, 2, 3 and 8) inhibitor 66 was modified with a benzophenone-spacer-alkyne moiety (67) to study its binding affinities in more detail. Incubation of FRDA lymphoblast derived nuclear extract with 67 followed by photoaffinity crosslinking and a click reaction with either rhodamine azide or biotin azide, identified HDAC-3 as the single target of this inhibitor.

In addition to the aryl azide and diazirine groups, the benzophenone moiety was also incorporated in probes that target matrix metalloproteinases (see Figure 6). Potent, broad-spectrum succinyl-hydroxamate MMP inhibitor GM6001 (68) was converted into photoaffinity probe 69 by incorporation of Bpa (61) and a fluorophore, as reported by Cravatt and co-workers in 2004. It was shown that this probe can be used to covalently label (through photocrosslinking) and visualize several active MMPs in complex proteomes. In addition, the authors were able to identify a number of other metalloproteases targeted by GM6001, which do not belong to the MMP family. In order to address this metalloenzyme’s lack of selectivity towards a single inhibitor, this same group developed an alternative profiling strategy. A two-step photoaffinity labelling MMP probe library was constructed, in which the Bpa moiety was incorporated at either the P3’ (70a,b) or P2’ (70c) position and the P1’ and P2’ substituents were varied (see Figure 6). The library compounds were applied as a ‘cocktail’ to proteomes and instead

![Figure 6](image_url)

Figure 6. Examples of benzophenone modified probes to target matrix metalloproteinases. The amino acid three letter abbreviations in the R (70a,b) and R (70c) substituents refer to their corresponding side chains.
of identifying the affinity of each metalloprotease towards a single probe the labelling profiles towards the entire library were analyzed collectively. This method proved to be powerful in that more than twenty metalloproteases (MMPs and others) could be identified from a complex biological mixture. In a later study they showed the use of some of the library compounds and related photoaffinity probes as competitive AβBPs to study the affinity of four MMP-13 inhibitors, for a large number of other metalloproteases.79

A powerful application of the use of benzophenone containing AβBPs to locate the active site within a protease complex has been reported for γ-secretase.76 This multi-subunit, integral membrane protein complex is responsible for the proteolysis of transmembrane proteins. Together with β-secretase, it generates the amyloid β-protein, which is known as the central pathogenic feature in Alzheimer’s disease. Although the nature of the γ-secretase catalytic activity was determined to be of aspartate protease type, the exact location of its active site within the complex was unknown. A potent inhibitor of γ-secretase activity is L-685,485 (71, Figure 7), containing a hydroxyethylene dipeptide-isostere. This compound mimics the transition state of the aspartate protease catalytic process and hence, forms a reversible, non-covalent adduct with the active enzyme (the binding of the inhibitor in the enzyme’s active site is shown in Figure 7). Li and co-workers reported the modification of the inhibitor’s P2 and P3’ substituents with a benzophenone moiety, which led to AβBPs 72a and 72b.76 The modifications did not result in a decrease in inhibitory potency towards γ-secretase activity inhibition compared to parent compound 71. The interaction of both probes with γ-secretase was studied by addition of the probes to HeLa cell membranes containing solubilised γ-secretase and subsequent photocrosslinking. From the obtained results the authors were able to identify membrane-spanning protein presenilin 1 (PS1) as the γ-secretase active site bearing subunit. Four years later the same group reported the preparation of the synthetically more challenging P1’ Bpa containing biotinylated analogue of 71 (structure not shown).77 With this probe, they were able to label active γ-secretase in

![Figure 7. Examples of AβBPs modified with the benzophenone photophore that target γ-secretase.](image-url)
living cells and their results demonstrated, for the first time, that active γ-secretase is presented on the cell surface.

The use of click chemistry has been applied to the synthesis of benzophenone modified γ-secretase probes as well. The group of Yao reported the preparation of a compound library built up from Bpa containing alkyne 73 and azide 74 (Figure 7). The azide part contains a racemic hydroxyethylene moiety and variations were made in its aryl sulfonamide domain. The compound library was screened for its potency against γ-secretase inhibition and the most potent compounds were used to label active PS1 in a cell lysate. In addition, Fuwa and co-workers reported a divergent synthesis of γ-secretase AfBPs by means of click chemistry with alkyne 75 and azide 76. Variations were made in the aryl part of the alkyne (dibenzoazepine or benzodiazepine) and in the type of spacer between the benzophenone moiety and biotin in the azide. Photoaffinity labelling using these probes provided the authors with evidence that the molecular targets of this type of probes are the N-terminal fragment of PS1.

Another class of enzymes that is especially well suited for PAL is the kinase family. These enzymes catalyze the ATP dependent phosphorylation of several substrates, but do not form a covalent linkage with either reaction partner and can therefore not be caught by means of a suicide trap. Some examples of photoreactive AfBPs targeting kinases are shown in Figure 8. In 2003 the group of Sewald reported the preparation of fluorescently tagged AfBP 78, which was obtained after modification of the potent kinase inhibitor H-9 (77), an isooquinolinesulfonamide containing, competitive inhibitor, targeting a broad range of kinases by occupying their ATP binding site. The probe proved to be able of labelling several kinases in a concentration dependent manner and could be eliminated by preincubation of the kinase with competing ligands. Furthermore, the authors were able to label, although not very selective, creatine kinase added to a mixture of isolated thylakoid proteins.

In 2006 Kawamura and co-workers reported a study towards the photoactivated labelling of kinases with compound 79, with Aa being glycine (Figure 8). In an initial screen towards six different kinases it was shown that this probe selectively labelled one

![Figure 8. Examples of benzophenone containing AfBPs for labelling of kinases. D-‘pip’ = (R)-piperidine-2-carboxylic acid; GABA = γ-aminobutyric acid.](image-url)
kinase, namely leukocyte-specific protein tyrosine kinase (Lck) and that labelling could be blocked by adenine. In addition, the probe selectively labelled Lck in an extract of Jurkat cells. With the aid of LC-MS/MS after tryptic digest, it was possible to identify the labelled fragment within the kinase. Photocrosslinking had taken place in the Ile379-Arg386 tryptic fragment and, more precisely, to either Gly383 or Leu384. In a later study, the influence of target-binding affinity and conformational flexibility on the photocrosslinking efficiency was assessed. For this, a small library was prepared, in which the glycine moiety in 79 was replaced by other (non-proteinogenic) amino acids. Variations in length, flexibility and hydrogen bonding capability were made. The results corroborated the earlier mentioned finding (see the examples of HDAC inhibitors and Chapter 4) that the inhibitor potency does not necessarily correlate with the photocrosslinking efficiency. Interestingly, it was found that higher crosslinking yields were obtained for the more flexible compounds, while in other cases (for other enzymes) crosslinking efficiencies increased upon a more tightly bound probe (see for example Chapter 4), however this probably depends on the type of photocrosslinker used as well as the class of enzymes studied.

1.3.4 Comparing photocrosslinkers

As has been pointed out, different photophores can give rise to different labelling products. Therefore, the outcome of a photolabelling experiment substantially depends on the type of photocrosslinker applied. In order to select the ideal photophore for a specific experiment, one should carefully examine the different possibilities. Interestingly, there are only a few reports on the comparison of different photophores under otherwise identical conditions, especially where all three are taken into account or when they are applied in an AxBP. Some representative examples will be given here.

Weber and Beck-Sickinger reported a study towards the photochemical behaviour of the different photophores. For this, the synthetic pentapeptide thymopentin, or TP5 (Arg-Lys-Asp-Val-Thr) was modified with each of the photocrosslinkers (C-terminal Tyr was replaced by Phe(Tmd) and Bpa or 4-azidobenzoic acid was coupled to the N-terminus), thereby obtaining three different photoreactive peptides. The conversion and product formation were followed by LC-MS analysis after photolysis in water, n-propanol and a water/n-propanol mixture (1:1, v/v). It was found that both the diazirine and aryl azide were quickly converted into a reactive species, but only the diazirine gave a relatively pure product formation resulting from insertion into the solvent, predominantly in the O-H bond. The aryl azide also gave a clean O-H insertion reaction in water, however when n-propanol was present numerous unidentified products were formed, probably arising from intramolecular rearrangements and insertion reactions. The benzophenone containing peptide gave clean C-H insertions in the presence of n-propanol, but the conversion yield was low after fifteen minutes of irradiation (68% and 82% starting material left in water/n-propanol and n-propanol respectively). In pure water the conversion was much higher (94% after fifteen minutes), however due to the poor reactivity towards O-H bonds many unidentified products were generated. This led to the conclusion that the use of diazirines and aryl azides is preferred over
benzophenones, however the latter should be used in case the crosslinking site contains many water molecules.

Another direct comparison between the photophores, however not in an activity-based setting, was reported by Tate and co-workers. They synthesized dUTP analogues containing four different photoreactive moieties (Figure 9A) and incorporated them enzymatically into DNA constructs. With these, the DNA photoaffinity labelling of yeast RNA polymerase III transcription complexes was studied. It was shown that photolabelling with the diazirine construct rendered many protein-DNA contacts, whereas labelling with the other three photoreactive moieties proved only marginal.

In a recently reported study by Dalhoff and co-workers the aryl azide and diazirine photophores were directly compared in an activity-based labelling of methyltransferases (MTases). This class of enzymes catalyzes the transfer of the activated methyl group in S-adenosyl-L-methionine (SAM) to a nucleophilic position in various substrates (DNA, RNA, proteins) under the formation of S-adenosyl-L-homocysteine (SAH) as shown in the insert in Figure 9B. SAH itself is a general competitive product inhibitor of MTases and because of this it was transformed into four AβPs by attachment of a spacer moiety containing biotin and either an aryl azide or aryl diazirine (Figure 9B). Since the optimal modification site was unknown, the linker moiety was attached to either the N6 or C8 position, as indicated in Figure 9B. Initially, the compounds were used to label and capture (by means of streptavidin-coated magnetic beads) purified DNA adenine N6 MTase from Thermus aquaticus (M.Taql). All four compounds could be used for M.Taql pull-down in a yield of 5-10%, however the capture efficiency of the diazirine-based probes was slightly higher compared to the aryl azides, as determined from the Western blots. Labelling proved to be activity-based since it could be competed with SAH and, in addition, was shown to be light dependent. With these results in hand, the authors also tested the N6-aryl azide-derived probe for pull-down of three other NTases in purified form and two other SAH binding proteins in E. coli lysate, which were identified by LC-MS/MS analysis. Unfortunately, the diazirine derived probes were not further examined, because the authors argued that ‘the small improvement in capture yield might not balance the much higher cost of preparation’.

Up to here it seems that the diazirine-group is the most effective photocrosslinker. Interestingly, there are two reports showing that the potency of the applied photoreactive group largely depends on the system it is used in. In a study towards photoreactive metalloprotease probes by Yao et al., an additional, benzophenone containing, probe was synthesized. In order to compare the diazirine and benzophenone moieties, the authors replaced the diazirine moiety in compound (Figure 9C, this is the same compound as with P1 = Leu (see Figure 4)) with the corresponding benzophenone counterpart (82b). This compound proved to label purified thermolysin as well, however no signs of labelling were observed when applying the compound to crude yeast extract, spiked with thermolysin. In this case the diazirine-based probe appeared superior compared to its benzophenone counterpart. In contrast, Fuwa and co-workers obtained the opposite results, although in a completely different system. In their study towards the
development of benzophenone derived photoreactive γ-secretase probes (see paragraph 1.3.3 and Figure 7), they also incorporated the diazirine photophore into their probes (83, Figure 9D). These diazirine containing analogues of the compounds resulting from a click reaction between 75 and 76 (Figure 7), proved to be completely ineffective in the labelling of γ-secretase active subunits, despite the fact that they showed the same inhibitory potential compared to their benzophenone counterparts.

![Structures of compounds](image)

Figure 9. Structures of compounds used to compare the photoaffinity labelling of different photoreactive groups. (A) Photoreactive dUTP analogues 80a-d reported by Tate et al.84 (B) Photoreactive SAH analogues 81a,b reported by Dalhoff et al.85 The insert shows the transfer of the activated methyl group from SAM to nucleophilic positions (Nu) in various substrates by methyltransferases (MTase). A = adenosine. (C) Photoreactive metalloprotease probes 82a,b reported by Yao et al.64 (D) Diazirine containing analogue of clicked compound 75-76 for labelling of γ-secretase reported by Fuwa et al.79

These results show clearly that there is no such thing as the ‘ideal photophore’ (yet). Due to the differences in reaction site preferences and the nature of the reactive intermediates of each individual moiety, the optimal photophore type needed differs from system to system. In most cases it can be worthwhile to spend some time on the
optimization of the used photophore in a certain ABP, in order to obtain the best results (an issue that will also be addressed in Chapter 6). In addition, interesting information can be obtained when the results derived from different photophores are combined.\(^\text{34}\)

### 1.3.5 Identification of the photocrosslinked sites

As outlined in paragraph 1.2, the tag incorporated in an ABP or AFBP is used for visualization, both \textit{in vivo} and as read-out after SDS-PAGE, of the bound enzyme. In addition, a bifunctional tag, such as biotin, can be used as a pull-down handle for purification of the covalently modified complex. Subsequent (tryptic) digestion and LC-MS/MS analysis allows the identification of the tryptic peptides and hence, the labelled enzyme(s). In the case a warhead containing ABP is used, the modified peptide fragments can in general be found easily in the LC-MS/MS analysis, since the chemistry of modification and hence, the nature of the final construct is known. In photoaffinity labelling, the identification of the photolabelled fragment is usually much more difficult. This is largely due to the fact that the site of modification is unknown (usually there are multiple modification sites) and the photocrosslinking reaction itself can lead to many possible products, because of the multiple reactive species involved. In some cases, the identification of the photolabelled site is possible (see for instance the labelling of hMMP-12 reported by Dive \textit{et al.} (paragraph 1.3.1) or the labelling of Lck reported by Kawamura and colleagues (paragraph 1.3.3)), but this is generally difficult, because of complicated mixtures. A possible solution to this problem is the mixed isotope photoaffinity labelling strategy, which makes use of so-called ‘MS-friendly’ photoprobes, or target identification probes (TIPs).\(^\text{86}\) The general idea of this strategy is schematically shown in Figure 10. A proteome is treated with a 1:1 mixture of two structurally identical AFBPs, which only differ in their absolute mass by incorporation of stable isotopes (e.g. deuterium, \(^{13}\)C, \(^{18}\)O or \(^{15}\)N). Subsequent photocrosslinking, purification (using the tag) and tryptic digestion leads to a mixture of labelled and unlabelled tryptic peptides. It is

![Figure 10](image)

**Figure 10.** Mixed isotope photoaffinity labelling strategy for determination of the labelling site. After binding and photocrosslinking the target enzyme with a 1:1 mixture of the ‘light’ (D\(_0\)) and ‘heavy’ (D\(_7\)) isotopically labelled AFBPs, the construct is purified and tryptically digested. LC-MS/MS analysis of the peptide pool allows discrimination between the labelled and unlabelled peptides. The modified fragment can easily be retrieved and identified by searching for the isotope signature.
now possible to discriminate between these two types of peptides in a LC-MS/MS analysis because of the double peaks separated by the mass difference between the 'light' and 'heavy' isotopically labelled species. Thus, the modified peptide(s) can be identified by searching for the unique isotopic pattern, or 'isotope signature'. Although this concept is relatively new in the field of photoaffinity labelling, some examples have been recently reported. Lamos and colleagues reported the modification of cyclophilin A (CypA) binding immunosuppressive CsA with a benzophenone-D$_{11}$ and a biotin moiety (compound 84, Figure 11). As a proof of principle, they used a 1:1 mixture of this TIP and its non-deuterated isoster for the selective photoaffinity labelling and pull-down of CypA among three other proteins. Subsequent tryptic digestion of the elutes and LC-MS/MS analysis allowed for the identification of eleven CypA characteristic peptides, two of which were modified with the probe as evidenced from the double, 11 Da separated, peaks in the mass spectra. The large 11-Da mass difference allowed easy visual recognition of labelled peptides in the mass spectra, which makes this a powerful method in determination of the modification site after photoaffinity labelling, however application in more complex systems still remains to be done.

![Figure 11](https://via.placeholder.com/150)

**Figure 11.** Examples of 'MS-friendly' photoreactive compounds used in mixed photoaffinity labelling.

A second, more extensive, study was recently reported by the group of Heck. These authors describe the application of peptides containing stable isotopes and a photoreactive moiety in the localization of cGMP-dependent protein kinase (PKG) substrate binding sites. For this, they focussed on oligopeptide DT-2 (Figure 11), a potent and selective inhibitor of PKG, which is a construct of a PKG tight binding sequence (W45) and a membrane translocating sequence (DT-6). Both of these peptides were modified by incorporation of a benzophenone moiety (as Bpa) and an isotopically labelled amino acid Leu$^{13}$C$_6$/$^{14}$N$_1$ or Arg$^{13}$C$_6$/$^{15}$N$_4$ (peptides 85-87, notice the difference of the Bpa position between the W45 derived peptides). LC-MS/MS analysis was performed on photocrosslinked complexes of PKG with each of the peptides. The power of the strategy was reflected in their results, which led to the identification of each peptide’s binding site. For example, the binding site of compound 87 was located in tryptic peptide Gln$^{195}$-Arg$^{203}$ and fragmentation of this sequence identified Met$^{201}$ as the
single cross-linked amino acid. Interestingly, the binding site of the highly similar peptide 86 was located on a completely differently situated peptide in the PKG sequence, namely Phe$^{359}$-Glu$^{374}$. In this case the fragmentation did not result in the determination of the exact crosslinked amino acid, however it could be pinpointed to the residues Thr$^{364}$-Glu$^{374}$. With all the obtained results the authors were unfortunately unable to determine the exact binding site of DT-2, however they obtained the interesting results that both the DT-6 and W45 peptides are targeted to the same pocket in the PKG’s catalytic domain and that therefore DT-2 is preferentially bound to dimeric PKG.

Altogether, the mixed isotope photoaffinity labelling strategy is still in development, but is believed to become a powerful tool in photoaffinity-based protein profiling strategies in the near future. With its aid the photocrosslinking site(s) of an AfBP can more easily be identified, which will lead to a better understanding of protein structure and function in general.

1.4 Aim and outline of this Thesis

The research described in this Thesis aims at the development of chemical biology research tools to study proteolytic activities, with the main focus on metalloproteases (MMPs and ADAMS) and proteasomes. The first part of Chapter 2 describes the synthesis of the enantiomerically pure N,O-diprotected succinyl hydroxamate building block 88 (Figure 12) and its use in the construction of peptide hydroxamate-based MMP/ADAM inhibitors in a highly efficient, linear SPPS protocol. In addition, photoreactive probe 89 was synthesized and used for covalent modification of ADAM-10. The second part of this chapter deals with the preparation of a library containing 96 enantiopure peptide hydroxamates with general structure 90 by SPPS using building block 88. This compound library was used to study the inhibitor preference of three metalloproteases (MMP-9, MMP-12 and ADAM-17) with respect to the substituents at the P2’ and P3’ positions (R$_2$ and R$_1$ respectively).

![Figure 12. Structure of building block 88, used for the construction of photoreactive metalloprotease probe 89 and a library containing 96 inhibitors with general structure 90.](image)
Although the covalent modification of ADAM-10 with compound 89 was shown to be possible, the efficiency proved rather modest. Therefore, the effect of moving the photophore to the more firmly binding P1’ pocket was studied and is outlined in Chapter 3. The synthesis of photoreactive building block 91 (Figure 13) and its application in the preparation of activity-based probes 92a-c is described. It is further demonstrated that 92a is indeed the more efficient MMP/ADAM AβP, compared to 89, in a head to head comparison towards a range of recombinant, purified metalloproteases.

Figure 13. Structure of photoreactive building block 91, used for the synthesis of MMP/ADAM AβPs 92a-c.

The research described in Chapter 4 entails a study towards the effect of fluorine incorporation into proteasome inhibitors on their specificity towards the different proteasome subunits. It was found that fluorine incorporation led to a large decrease in potency towards the β1 and β2 subunits and, hence resulted in selectivity for β5. Tripeptide epoxyketone 93 (Figure 14) was identified as one of the most β5 selective inhibitors known to date.

Figure 14. Structures of β5 selective (93) and β2 selective (94) proteasome inhibitors.

The development of selective inhibitors for the proteasome’s trypsin-like sites (β2) is discussed in Chapter 5. Selectivity for β2 was achieved by instalment of basic side chain residues in the P1 position of tripeptide vinylsulfones. Incorporation of a 4-aminomethylene benzylamine substituent at this position resulted in compound 94 (Figure 14), which showed high β2 selectivity both in cell extracts and in living cells.

Chapter 6 describes the study towards the orientation of extended peptide vinylsulfones inside the proteasome’s catalytically active cavity. This study combined activity-based modification by means of a warhead with photocrosslinking. A panel of
bifunctional two-step probes, comprising of an extended peptide vinylsulfone equipped with an N-terminal photophore, was prepared. Of these probes, compound 95 (Figure 15) was able to establish a crosslink between the active β5 subunit and a neighbouring β6 subunit which could be identified by subsequent SDS-PAGE and LC-MS/MS analysis.

Figure 15. Structures of photoactivatable proteasome probe 95 and levulinoyl-based cleavable linker containing proteasome probe 96.

The design and preparation of a novel levulinoyl ester-based cleavable linker is described in Chapter 7. The cleavable linker was incorporated into potent proteasome inhibitor epoxomicin and attachment of biotin resulted in ABP 96 (Figure 15). The optimal characteristics, in terms of stability and cleavability, of the cleavable linker were reflected by the use of 96 for pull-down of proteasome active subunits from a cell lysate. Subsequent treatment with hydrazine allowed a chemoselective release of 96-derived bound proteins.

Finally, the research described in this Thesis is summarised in Chapter 8, and some future prospects based on the presented results are discussed.

References

(30) H. Bayley, Photogenerated reagents in biochemistry and molecular biology, Elsevier, Amsterdam, 1983.
Chapter 1


