Twelve ways to confirm targets of activity-based probes in plants

Judit Kovács & Renier A. L. van der Hoorn

The Plant Chemetics Laboratory, Department of Plant Sciences, University of Oxford,
South Parks Lane, OX1 3RB Oxford, United Kingdom.

For correspondence: Tel.: +44 1865 275077; Fax: +44 1865 275074

E-mail address: renier.vanderhoorn@plants.ox.ac.uk

Running title: Target confirmation for ABPP

Key words: Activity-based protein profiling; mass spectrometry; Arabidopsis;
tomato; probe targets.

ABSTRACT

Activity-based probes are powerful tools to interrogate the functional proteome. Their covalent and often irreversible labeling of proteins facilitates the purification, identification and quantification of labeled proteins. However, the detection of labeled proteins often requires a confirmation, especially when unexpected proteins are identified, or to unravel fluorescent activity profiles. Here, we review twelve approaches towards target confirmation, grouped in approaches by direct target detection, target expression or target depletion. We discuss their proper use and limitations and illustrate these approaches with examples from plant science.

Introduction

Activity-based protein profiling (ABPP) is a powerful and robust functional proteomics tool that displays the active proteome in various biological systems (Cravatt et al. 2008; Serim et al., 2012; Willems et al., 2014; Nodwell & Sieber, 2012; Morimoto & Van der Hoorn, 2016). Chemical probes for ABPP react with active site of target proteins in a mechanism-dependent manner, resulting in a covalent and irreversible bond that facilitates the purification and detection of the labeled proteins.

There are four types of chemical probes (Cravatt et al., 2008; Morimoto & Van der Hoorn 2016). i) Mechanism-based probes are often inspired on a covalent,

mechanism-based inhibitor (e.g. DCG-04 (**1**) is biotinylated version of protease inhibitor E-64 (Greenbaum et al., 2000)). ii) Photoaffinity probes are reversible inhibitors equipped with a photoreactive group that facilitates crosslinking upon UV irradiation (e.g. JL01 (**2**) for metalloproteases (Lenger et al., 2012)). iii) Suicide probes become reactive upon activation by the enzyme and usually react when still in the substrate binding pocket (e.g. ethynylnaphthalene (**3**) for cytochrome P450s (Wright et al., 2007)). iv) Reactivity probes lack a binding group and preferentially react with hyper-reactive residues that are often active site residues (e.g. iodoacetamide probes (**4**) for hyper-reactive Cys residues (Weerapana et al., 2010)).

Probes carry a reporter tag for detection, which is most often a fluorophore or biotin. Some probes carry a chemical handle that facilitates the coupling to a reporter tag through bio-orthogonal chemistry (Kolb & Sharpless, 2003; Kaschani et al., 2009a; Willems et al., 2011).

Fluorescently labeled proteins are usually detected upon separation on protein gels by fluorescent scanning. The detection from fluorescently labeled proteins from gels is quantitative and relatively quick, but the identity of the labeled protein remains unknown unless candidate proteins are depleted from the proteome or produced and labeled. Another limitation is the resolving power of protein gels as similar proteins frequently overlap in molecular weight (MW). Occasionally also two-dimensional protein gels are used to increase the separation, but this approach is more elaborate and not more sensitive (Shabab et al., 2008).

Proteins labeled with biotinylated probes can be purified using avidin technology and stringent washing conditions to remove unlabeled proteins. These proteins are either eluted, separated on gel, digested ‘in-gel’ by trypsin, and analysed by mass spectrometry (MS), or digested ‘on-bead’ by trypsin and analyzed by MS. The in-gel approach is less sensitive but provides MW information which could help in the annotation of the signals detected with fluorescent probes. MS analysis of on-bead digests delivers more identities, but does not provide information of the MW of the protein (e.g. Kaschani et al., 2009b). Purification and MS analysis of labeled peptides purified from whole proteome ‘in-solution’ digests is a third detection method, which can only be followed for conjugates that can be detected by MS, and can deliver labeling sites of an even larger number of targets.

A single method for identifying labeled proteins are often not sufficient to confirm probe targets, especially when they are unexpected proteins. Detection by MS,

for example, is insufficient unless rigorous no-probe-controls have been included or the labeled peptide has been identified. Likewise, fluorescent signals can be caused by multiple labeled proteins and it is impossible to determine their relative contributions without additional assays.

This review summarizes twelve approaches to validate probe targets, and discusses their proper use and limitations, illustrated with examples from plant science. So far, 40 different probes representing 14 different probe chemotypes have been validated in plant science. These probes have the potential to label over 4500 different plant proteins, including proteases, glycosidases, Ser hydrolases, kinases, glutathione transferases and aldehyde dehydrogenases (Morimoto & Van der Hoorn, 2016). Although many probes hit their expected targets, several unexpected probe targets have also been identified. The twelve approaches to confirm probe targets can be grouped as follows: 1) direct detection of labeling; 2) (over)expression of putative target proteins; 3) depletion of the putative target from the proteome. Each of these approaches has its own advantages and limitations, and have different confidence levels (**Table I**), further explained below. A combination of different approaches is often required to validate the probe target.

1. Confirmation by robust detection of the target

1.1. Purify labeled proteins and identify by mass spectrometry

Purification of labeled proteins, followed by identification by mass spectrometry is the most frequently used method to identify labeled proteins. The biotin-avidin technology is almost exclusively used for affinity purification, although there are important variations in the type of avidin, the matrix, and the type of biotin (e.g. desthiobiotin). The risk in this approach is that some proteins co-purify on the matrix, either because they are extremely abundant, or sticky to avidin and/or the matrix, or because they are endogenously biotinylated. However, because the probe-protein conjugate is covalent and often irreversible, and because the avidin-biotin interaction is extremely strong, stringent washing conditions including 1% SDS and 6 M urea can be used to remove unlabeled proteins from the sample. In addition, endogenously biotinylated proteins can be removed by prepurification before labeling, or discarded because they are known. Nevertheless, a no-probe-control, where the solvent is added instead of the probe, is always useful to include. Purification of fluorescently labeled proteins using antibodies for the fluorophore has also been successful (Shaulov-Rotem et al., 2016), but this approach has not been taken in plant science, yet.

For detection by mass spectrometry, purified biotinylated proteins are treated in-gel or on-bead with trypsin (or other enzymes) and the released peptides analyzed by tandem mass spectrometry (MS). Peptide annotation delivers a list of proteins that were absent in the no-probe-control. In many cases, only proteins from expected classes appear in the substrate MS list. For example, 45 of the 45 FP-labeled proteins of Arabidopsis leaves, selected upon subtracting hits detected in the no-probe-control, are Ser hydrolases, consistent with the fact that FP (5) labels the active site of Ser hydrolases (Kaschani et al., 2009b). Similarly, pull downs and MS of proteomes labeled with DCG-04 (1), MVB070 (6), bAMS101 (7) and JJB111 (8), identified the expected Papain-like Cys Proteases (PLCPs), proteasome catalytic subunits, Vacuolar Processing Enzymes (VPEs), and retaining glycosidases, respectively (Van der Hoorn et al., 2004; Kolodziejek et al., 2011; Misas-Villamil et al., 2013a; Chandrasekar et al., 2014; **Figure 2A**). These examples illustrate that if the annotation of the labeled proteins is consistent with the expected target selectivity of the probe, then this is sufficient confirmation of labeling.

1.2. Fragmentation spectrum of labeled peptide

The annotation of a labeled peptide is possibly the strongest evidence to validate labeling, not only to confirm a probe-protein conjugate, but also to reveal the exact labeling site within the target protein. This evidence is, however, not always achievable. In some cases, the labeling site resides on a peptide that escapes detection by MS, e.g. because it is too large, too small, or its modification by the probe alters its properties such that the peptide escapes detection by MS, e.g. by fragmentation within the probe itself. In other cases, the probe-peptide conjugate is not stable enough to survive detection by MS. For instance, cyclophellitol aziridine probes (8) label retaining glycosidases via an ester linkage, which slowly hydrolyses in water, explaining why both unmodified and probe-modified active site peptides have been detected (Chandrasekar et al., 2014). Another challenge is the fact that labeled peptides often get a lower score in automated annotation programs because the fragmentation spectra contain peaks from the probe itself that are not assigned by the program. Manual annotation of these unassigned peaks to the probe-modified amino acid increases the confidence in the identification of the labeling site.

There are several good examples of annotated labeled peptides, e.g. for PsbP, GST-U20, HSP60 and TGG1. Protein P of photosystem II (PsbP) of Arabidopsis was unexpectedly labeled by IS4 (9), a probe containing a β -lactone warhead. Manual annotation of the fragmentation spectrum demonstrated that IS4 labels the N-terminus of mature PsbP, which was later explained by the discovery of a protease that can ligate IS4 and peptides to unmodified N-termini of proteins (Wang et al., 2008). Glutathione transferase U20 (GST-U20) was one of many Arabidopsis proteins labeled by sulfonyl fluoride probes (10). Manual

annotation of the MS2 spectrum of the labeled peptides revealed that labeling of GST-U20 occurs at a Tyr residue in the promiscuous substrate binding site, a pattern that also occurred with distantly related GSTs from mammals similar probes (Gu et al., 2013, **Figure 2B**). Arabidopsis heat shock protein 60 (HSP60) is labeled by acyl-ATP probes (**11**) in the ATP binding pocket, but the labeled peptides were only detected after the discovery that the biotin was oxidized in this experiment, explaining why these peptides initially escaped the detection (Villamor et al., 2013). This discovery indicates that the inclusion of biotin oxidation in searches may improve the identification of labeled peptides by MS. This work also includes the identification of a 242 additional labeling sites, often (but not always) in ATP binding pockets, including those of 24 different protein kinases (Villamor et al., 2013). Finally, TGG1 is a myrosinase of Arabidopsis that was labeled with JJB111 (**8**), a cyclophellitol aziridine-based probe originally designed for beta-glucosidases (Kallemeijn et al., 2012). Manual annotation of the MS2 spectrum of the labeled peptide confirms that labeling occurs at the active site glutamate/aspartate residue (Chandrasekar et al., 2014). However, also unlabeled active site peptides were detected, consistent with the hydrolysis of the unstable ester bond of this conjugate (Chandrasekar et al., 2014). Thus, detailed studies of labeling sites by interpreting MS2 spectra have confirmed labeling in many studies and have been crucial for understanding labeling mechanisms.

1.3. Probing pull-down with specific antibody

If a specific antibody recognizing the protein of interest is available, this can be used to detect if this protein is labeled by investigating if it is amongst the purified biotinylated proteins. A limitation of this approach is that larger quantities of biotinylated proteins should be purified, separated on protein gels, transferred to a membrane, and probed with the protein-specific antibodies. Also, because antibodies may also unspecifically bind to related proteins or other (abundant) proteins, a no-probe-control or other negative controls are required. Positive controls are also needed to show that the western blot display the anticipated signals.

A good example of using antibodies for the detection of labeled proteins is Rcr3, a secreted PLCP (Papain-like Cys Protease) of tomato (Rooney et al., 2005). Because of its low level in secreted proteomes, Rcr3 often remains undetected, especially because its closest paralog, Pip1, is more abundant and has a similar molecular weight, hampering the separation of these two proteases on protein gels. Detection using Rcr3-specific antibodies on purified biotinylated proteins, generated by large scale DCG-04 (**1**) labeling of secreted proteomes, was successful in the detection of Rcr3 among the labeled proteins. The signal was absent in labeled secreted proteomes of *rcr3-3* mutant tomato plants lacking Rcr3, and upon pre-incubation with the selective PLCP inhibitor E-64 (Rooney et al., 2005, **Figure 2C**). Likewise, western blots of purified labeled proteins using protein-specific antibodies were also used to

confirm labeling of the proteases RD21, XCP2 and AALP by DCG-04 (**1**) (Van der Hoorn et al., 2004), and to confirm labeling of protein P of photosystem II (PsbP) by the β -lactone probe IS4 (**9**) (Wang et al., 2008). Thus, if specific antibodies are available, sufficient labeled proteins can be purified, and negative controls are included, this western blotting approach can provide a robust confirmation of labeling.

1.4. Immuno-precipitation of labeled proteins

Protein-specific antibodies can also be used to test whether labeling occurred on the immunopurified putative target protein. This approach can be extended by fusion of the putative target protein with epitope tags (e.g. HA, cMyc or 6His) or protein tags (e.g. MBP, GST, GFP). Practically, specific proteins are immunoprecipitated using protein-specific antibodies from labeled proteomes and labeling is detected e.g. by fluorescence scanning (fluorescent probes) or streptavidin-HRP (biotinylated probe). An important limitation is that antibody-protein interactions tend to be weaker when compared to biotin-avidin interactions, so the washing conditions need to be less stringent. Consequently, IP experiments should include stringent controls to exclude co-purified, unspecific proteins, e.g. using other antibodies or the pre-immune serum. To our knowledge, no examples of this procedure are available from plant science.

1.5 Labeling characteristics

Labeling under different conditions can also be used to generate support for certain probe targets. Enzymatic activities are tightly regulated by the availability of cofactors and by pH, redox, salt concentrations and temperature. ABPP often reflects these relationships and labeling characteristics can therefore be used as support for certain probe targets. Obvious limitations of this strategy are that enzymatic activity parameters are only partially described and could be different in proteomes when compared to pure proteins.

Labeling characteristics were for example used to support that fluorescent signals caused by AMS101 (**7**) labeling in tomato are indeed from vacuolar processing enzymes (VPEs) (Sueldo et al., 2014). The VPEs were suspected as AMS101 targets in tomato because of the high selectivity of this probe in Arabidopsis proteomes and because the signals have a similar MW and intensity. However, the subcellular location was unexpected because these signals were observed in secreted proteomes (apoplast) whereas VPEs are vacuolar enzymes. Fortunately, VPEs have very specific activity characteristics. Being Cys proteases, VPEs require a reducing agent for activity and the activity is blocked by alkylating agents. Also, VPEs are only active at mildly acidic pH, and not at neutral pH. Labeling in the secreted proteomes occurs at pH5.5 and in the presence of DTT, and is absent at pH 7.5, or when DTT, is absent, or when iodoacetamide is added to alkylate Cys residues (**Figure 2D**, Sueldo et al.,

2014). These data, together with the suppression of labeling by Ac-YVAD-cmk but not E-64 (Figure 2D, section 3.2), suggest that the signals are derived from VPEs. The unexpected localization of this signal calls for further studies to confirm the identity of this signal and study its role in the apoplast, where its activity profile changes upon infection (Sueldo et al., 2014).

2. Confirmation by (over-)expression of putative target proteins

2.1. Labeling of transiently over-expressed target proteins

Transient over-expression by agroinfiltration is a frequently used method in plant science to confirm labeling of putative target proteins. Agroinfiltration is based on the use of disarmed *Agrobacterium tumefaciens* bacteria, which carries genes-of-interest on the transfer DNA (T-DNA) located on a binary plasmid. Infiltration of these bacteria into leaves results in the transfer of the T-DNA into the plant cell nucleus, from where it is expressed (Kapila et al., 1996). The wild tobacco plant *Nicotiana benthamiana* is often used for agroinfiltration because of its high transformation efficiency. Agroinfiltration results in transient expression of proteins of interest within a few days (Sparkes et al., 2006). As the gene-of-interest is often expressed to levels above that of endogenous homologous proteins, crude extracts labeled with activity-based probes usually display an additional strong signal when compared to agroinfiltrated control leaves that do not express the putative target protein.

Good examples of confirmation by labeling upon agroinfiltration are RD21 and SNG1. RD21 is a papain-like cysteine protease that was identified among proteins from Arabidopsis leaves labeled by DCG-04 (1) (Van der Hoorn et al., 2004). DCG-04 labeling of extracts from *Nicotiana benthamiana* transiently expressing RD21 resulted in several strong signals that were absent from leaves not expressing RD21 (Gu et al., 2012). Labeling was suppressed upon preincubation with inhibitor E-64 and absent for the C161A mutant RD21, which accumulates as protein but lacks the active site cysteine (Gu et al., 2012). SNG1 is a plant-specific acyltransferase involved in the biosynthesis of the secondary metabolite synapoylmalate (Lehfeldt et al., 2000). SNG1 was detected amongst proteins labeled with FP-biotin (5), which usually targets Ser hydrolases. Labeling of extracts from SNG1 agroinfiltrated leaves resulted in an additional signal, consistent with the presence of an active site Ser residue in SNG1 (Kaschani et al., 2009b). Likewise, labeling of Ser hydrolases FSH1, SH1, MES2, CXE12 and SCPL11 was confirmed by agroinfiltration (Kaschani et al., 2012, Figure 3A). Agroinfiltration was also used to study labeling of papain-like proteases Rcr3 (Rooney et al., 2005); Pip1 (Tian et al., 2007); CatB1 (Gilroy et al., 2007); C14, ALP and CYP3 of tomato (Shabab et al., 2008), and RD21B, RD21C, RDL2, XCP1, XCP2, SAG12, THI1, RD19A, RD19B, AALP and CTB3 of Arabidopsis (Richau et al., 2012). The relatively

high success rate of over-producing plant proteins by agroinfiltration makes this assay well suited to confirm labeling of proteins.

2.2. Labeling of purified (mutant) target proteins

In several cases, labeling has been confirmed on purified proteins produced by heterologous protein expression using *Escherichia coli* or *Pichia pastoris*. A risk with labeling purified proteins is that aspecific labeling can occur. DCG-04 (**1**), for example, can react with purified papain-like proteases, even if these do not carry the active site Cys residue (Kumari & Van der Hoorn, unpublished data). However, many probes remain selective on purified proteins. Mutant target proteins that lack the putative labeling site are very useful to include as negative control.

Specific labeling of purified proteins is illustrated by AvrPphB and SjGST. AvrPphB is a papain-like protease of *Pseudomonas syringae* that is injected into host cells as a type-III effector to cleave defence-related kinases (Shao et al., 2003). Incubation of purified AvrPphB produced in *E. coli* with a novel acyloxymethylketone probe (**12**) containing an AvrPphB targeting tripeptide Gly-Asp-Ala, resulted in strong labeling of wild-type AvrPphB, but not of the C98S active site mutant of AvrPphB (Lu et al., 2013). SjGST is a glutathione transferase from the human parasite *Schistosoma japonicum*, which has become the standard GST fusion protein. Purified SjGST produced in *E. coli* was used to confirm that sulfonyl fluoride probes (**10**) consistently label a Tyr residue in the promiscuous substrate binding site of GSTs. Labeling indeed occurs in wild-type SjGST but not of the Y111F mutant lacking this residue, consistent with this prediction (Gu et al., 2013, **Figure 3B**). An advantage of working with purified proteins is that these proteins can be tested for their activity in traditional enzymatic assays. SjGST, for instance, is an active enzyme, but its Y111F mutant not. Likewise, we have observed a strong correlation between active site labeling of aldehyde dehydrogenase ALDH3H1 by chloroacetamide probes (**4**) and its enzymatic activity (Stiti et al., 2016).

2.3. Labeling of spiked recombinant proteins

Problems caused by aspecific labeling on purified target proteins can be reduced by labeling a proteome that is mixed with a low amount of this recombinant target protein. This ‘spiking’ approach also facilitates the study of labeling under more natural conditions. This approach can be taken when the endogenous target protein is below detection level. A specific signal should only be detected when the recombinant protein is added. However, additional negative controls are required to exclude that the protein is aspecifically labeled, such as mutant proteins lacking the labeling site or the use of specific inhibitors to suppress labeling. There are no examples of this approach in plant science.

2.4. Crystallography

A very elegant way to confirm probe targets is to crystallize the purified labeled protein and elucidate its structure. This approach will also identify the labeling site and provide unique understanding of the specificity of the probe. A limitation is, however, that the production of an active, pure protein in a large quantity is challenging and that not all conjugates form crystals and that not all crystals diffract. At this stage, this approach has not been taken in plant science, but this may change in the future.

3. Confirmation by target depletion

3.1. Target depletion by genetics

Despite the above mentioned methods, it remains difficult to estimate the contributions of labeled proteins to the fluorescent signals in activity profiles when these signals are caused by multiple proteins. Reverse genetics is a robust method to determine the contribution of proteins to fluorescent signals as these signals will be reduced if the target protein is genetically depleted. However, this approach is limited to organisms that can be genetically manipulated and to genes that do not cause major phenotypes when depleted. In addition, genetic depletion may also indirectly suppress labeling, e.g. by removal of a positive regulator. The later issue requires labeling of the target protein for confirmation (sections 1 or 2). Also, the labeled signal should be consistent with the MW of the target proteins and labeling properties, e.g. with respect to dependency on pH, cofactors, and reducing agent.

There are at least seven examples for suppressed labeling upon genetic depletion. As first example, major signals in the DCG-04 (1) labeling profile of Arabidopsis leaves disappeared in the *rd21-1* and *aalp-1* null mutants, consistent with previous MS data, the expected MW of mature RD21 and AALP proteases and their relatively high abundance in leaves (Gu et al., 2012). Second, depletion of *Pip1* in tomato by RNAi removed a major signal from the DCG-04 labeling profile in secreted proteomes (Ilyas et al., 2015, **Figure 4A**). Third, Arabidopsis mutants lacking proteases CTB3 and AALP facilitated the discovery that these proteases are targeted by the probes JOGDA1 (12) and FY01 (13), respectively (Lu et al., 2015). The selectivity of these probes in the wild tobacco plant *Nicotiana benthamiana* was confirmed by depletion of the signals of the endogenous CTB and ALP proteases by RNAi (Lu et al., 2015). Fourth, signals from vacuolar processing enzymes (VPEs) are absent in *vpe* null mutants, both in leaves labeled with AMS101 (7) (Misas-Villamil et al., 2013a) and in seeds labeled with JOPD1 (14) (Lu et al., 2015). Arabidopsis has four VPEs and single mutant analysis deciphered the contributions of VPEs in leaves and seeds (Misas-Villamil et al., 2013a; Lu et al., 2015). Consistently, imaging with AMS101 (7) and JOPD1 (14) on living tissues shows vacuolar fluorescence that is absent in *vpe* mutants, both in Arabidopsis leaves and seeds, respectively (Misas-Villamil et al., 2013a; Ashnest et al., 2015). Fifth, the

major signal in activity profiles of Arabidopsis leaves labeled with fluorophosphonate probes (5) is absent in mutants lacking the CXE12 carboxypeptidase, consistent with its labeling (Kaschani et al., 2012). And finally, the two major signals in the glycosidase activity profile of Arabidopsis leaves are absent in mutants lacking both myrosinases TGG1 and TGG2, consistent with their labeling by JJB70 (8) (Chandrasekar et al., 2014). The latter example also illustrates that caution should be taken using mutants as sole method to confirm probe targets: solubility of TGG2 depends on the presence of TGG1, so the absence of the TGG2-derived signal in the *tgg1* mutant was indirect and the TGG2 signal was restored in *tgg1* mutants using a different protein extraction procedure (Chandrasekar et al., 2014). Conversely, mutants have been very useful to exclude suggested regulatory mechanisms: the presumed proteolytic activation of PLCPs by VPEs, for example, could be rejected because the DCG-04 (1) activity profile remains unaltered in the *vpe* quadrupole null mutant lacking all VPEs (Gu et al., 2012).

3.2. Target depletion using (independent) specific inhibitors

Putative targets can also be depleted chemically using known target-selective inhibitors. The known high selectivity of these inhibitors can then confirm labeling of the target by the suppression of labeling upon preincubation with the target inhibitor. A limitation of this approach is the data on the selectivity of inhibitors in the environment under study are often lacking. Also, while inhibitors tend to be selective between protein families, they are often not selective within protein families. However, multiple independent inhibitors can be used and this will increase the confidence in confirmation. A second limitation is that selective inhibitors are unknown for many targets of activity-based probes.

There are good examples of this approach in plant science for the proteasome, PLCPs, and VPEs. Labeling of leaf proteomes by a rhodamine-tagged syringolin A (RhSylA (15)), a known covalent proteasome inhibitor, but with unknown target selectivity, results in only two fluorescent signals that are suppressed upon preincubation with the selective proteasome inhibitor epoxomicin, both in Arabidopsis (Figure 4B) and *N. benthamiana* (Kolodziejek et al., 2011; Misas-Villamil et al., 2013b). Likewise, DCG-04 (1) labeling of papain-like cysteine proteases in Arabidopsis leaf extracts is suppressed not only by PLCP inhibitor E-64, on which DCG-04 has been based, but also upon preincubation with known PLCP inhibitors antipain, chymostatin, leupeptin and cystatin (Van der Hoorn et al., 2004). Labeling of VPEs by AMS101 (7) and JODP1 (14) is suppressed by the known VPE inhibitor Ac-YVAD-cmk (Misas-Villamil et al., 2013a; Lu et al., 2015, Figure 2D), and this last example also illustrates the limitation of this approach: Ac-YVAD-cmk was originally developed for caspases, which do not exist in plants, but this inhibitor also targets many other proteases, such as Cathepsin B (Lu et al., 2015). Likewise, the cathepsin B-selective inhibitor CA074

also blocks labeling of other PLCPs in plants (Gilroy et al., 2007). In fact, cross reactivity is common to many inhibitors. Proteasome inhibitors MG132 and MG115 also suppress labeling of PLCPs (Kaschani et al., 2009a; Lu et al., 2015), whilst PLCP inhibitor leupeptin also suppresses labeling of the proteasome catalytic subunits (Gu et al., 2010). In conclusion, suppression of labeling by chemical knock-out is indicative for confirmation of labeling but should be interpreted with caution and combined with another approaches for validation.

3.3. Target removal by immunodepletion

Another elegant method to confirm labeling through target depletion is by biochemical removal of the target using immobilized target-specific antibodies. This ‘immunodepletion’ results in a proteome that lacks the target protein and labeling this proteome should result in the absence of a signal. The practicality of this approach is limited by the availability of specific antibodies and the efficiency of target capture from native proteomes.

There is one example of immunodepletion used in plant science. The labeling of vacuolar processing enzymes of the wild tobacco plant *Nicotiana benthamiana* (NbVPE) by caspase-1 probe biotin-xVAD-fmk (**16**) was confirmed by preincubating leaf extracts with preimmune serum or anti-NbVPE serum to selectively deplete NbVPE from the leaf extract. Labeling of immunodepleted leaf extracts with Biotin-xVAD-fmk followed by detection of biotinylated proteins revealed a suppressed signal in the 38-40 kDa region, indicating that biotin-xVAD-fmk labels NbVPE (Hatsugai et al., 2004, **Figure 4C**).

Conclusion and prospects

The twelve different approaches to confirm probe targets differ in their confidence, research effort and pitfalls. The choice of approach very much depends on the question. Deciphering fluorescent labeling profiles require different approaches than confirming labeling of detected proteins or resolving a novel labeling mechanism. Available approaches also very much depend on the protein, proteome and organism under study. Some proteins are difficult to express heterologously, some proteomes have insufficient in quantity for large scale pull downs, and some organisms are challenging to manipulate genetically. Choices for approaches for confirmation are therefore mostly determined by the research question and biological system. Nevertheless, given the limitations of each approach, a combination of multiple validation methods will often be required to claim the identification of a probe target.

Acknowledgements

We would like to thank Friederike Grosse-Holz and Kyoko Morimoto for critically reading the manuscript. This work has been financially supported by the University of Oxford and ERC Consolidator Grant 616449 ‘GreenProteases’.

398

399 **References**

- 400 **Ashnest JR, Huynh DL, Dragwidge JM, Ford BA, Gendall AR** (2015) *Arabidopsis*
401 intracellular NHX-type sodium-proton antiporters are required for seed storage protein
402 processing. *Plant Cell Physiol* 56:2220-2233.
- 403 **Chandrasekar B, Colby T, Emran Khan Emon A, Jiang J, Hong TN, Villamor JG,**
404 **Harzen A, Overkleeft HS, van der Hoorn RAL** (2014) Broad-range glycosidase activity
405 profiling. *Mol Cell Proteomics* 13:2787-2800.
- 406 **Cravatt BF, Wright AT, Kozarich JW** (2008) Activity-based protein profiling: from
407 enzyme chemistry to proteomic chemistry. *Annu Rev Biochem* 77:383-414.
- 408 **Gilroy EM, Hein I, van der Hoorn RAL, Boevink PC, Venter E, McLellan H, Kaffarnik**
409 **F, Hrubikova K, Shaw J, Holeva M, López EC, Borrás-Hidalgo O, Pritchard L, Loake**
410 **GJ, Lacomme C, Birch PR** (2007) Involvement of cathepsin B in the plant disease resistance
411 hypersensitive response. *Plant J* 52:1-13.
- 412 **Greenbaum D, Medzihradszky KF, Burlingame A, Bogoy M.** (2000) Epoxide
413 electrophiles as activity-dependent cysteine protease profiling and discovery tools. *Chem Biol.*
414 7: 569-581.
- 415 **Gu C, Kolodziejek I, Misas-Villamil J, Shindo T, Colby T, Verdoes M, Richau KM,**
416 **Schmidt J, Overkleeft HS, van der Hoorn RAL** (2010) Proteasome activity profiling: a
417 simple, robust and versatile method revealing subunit-selective inhibitors and cytoplasmic,
418 defense induced proteasome activities. *Plant J* 62:160–170.
- 419 **Gu C, Shabab M, Strasser R, Wolters PJ, Shindo T, Niemer M, Kaschani F, Mach L,**
420 **van der Hoorn RAL** (2012) Post-translational regulation and trafficking of the granulin-
421 containing protease RD21 of *Arabidopsis thaliana*. *PLoS One* 7:e32422.
- 422 **Gu C, Shannon DA, Colby T, Wang Z, Shabab M, Kumari S, Villamor JG, McLaughlin**
423 **CJ, Weerapana E, Kaiser M, Cravatt BF, van der Hoorn RAL** (2013) Chemical
424 proteomics with sulfonyl fluoride probes reveals selective labeling of functional tyrosines in
425 glutathione transferases. *Chem Biol* 20:541-548.
- 426 **Hatsugai N, Kuroyanagi M, Yamada K, Meshi T, Tsuda S, Kondo M, Nishimura M,**
427 **Hara-Nishimura I** (2004) A plant vacuolar protease, VPE, mediates virus-induced
428 hypersensitive cell death. *Science* 305:855-858.
- 429 **Ilyas M, Hörger AC, Bozkurt TO, van den Burg HA, Kaschani F, Kaiser M, Belhaj K,**
430 **Smoker M, Joosten MH, Kamoun S, van der Hoorn RAL** (2015) Functional divergence of
431 two secreted immune proteases of tomato. *Curr Biol* 25:2300-2306.
- 432 **Kallemijn WW, Li KY, Witte MD, Marques AR, Aten J, Scheij S, Jiang J, Willems LI,**
433 **Voorn-Brouwer TM, van Roomen CP, Ottenhoff R, Boot RG, van den Elst H, Walvoort**
434 **MT, Florea BI, Codée JD, van der Marel GA, Aerts JM, Overkleeft HS** (2012) Novel

activity-based probes for broad-spectrum profiling of retaining β -exoglucosidases in situ and in vivo. *Angew Chem Int Ed Engl* 51:12529-12533.

Kapila J, De Rycke R, Van Montagu M, Angenon G (1996) An Agrobacterium-mediated transient gene expression system for intact leaves. *Plant Sci* 122:101-108.

Kaschani F, Verhelst SH, van Swieten PF, Verdoes M, Wong CS, Wang Z, Kaiser M, Overkleeft HS, Bogyo M, van der Hoorn RAL (2009a) Minitags for small molecules: detecting targets of reactive small molecules in living plant tissues using 'click chemistry'. *Plant J* 57:373-385.

Kaschani F, Gu C, Niessen S, Hoover H, Cravatt BF, van der Hoorn RAL (2009b) Diversity of serine hydrolase activities of unchallenged and botrytis-infected *Arabidopsis thaliana*. *Mol Cell Proteomics* 8:1082-1093.

Kaschani F, Shabab M, Bozkurt T, Shindo T, Schornack S, Gu C, Ilyas M, Win J, Kamoun S van der Hoorn RAL (2010) An effector-targeted protease contributes to defense against *Phytophthora Infestans* and is under diversifying selection in natural hosts. *Plant Physiol* 154: 1794-1804.

Kaschani F, Nickel S, Pandey B, Cravatt BF, Kaiser M, van der Hoorn RAL (2012) Selective inhibition of plant serine hydrolases by agrochemicals revealed by competitive ABPP. *Bioorg Med Chem* 20:597-600.

Kolb HC, Sharpless KB (2003) The growing impact of click chemistry on drug discovery. *Drug Discov Today* 8:1128-1137.

Kolodziejek I, Misas-Villamil JC, Kaschani F, Clerc J, Gu C, Krahn D, Niessen S, Verdoes M, Willems LI, Overkleeft HS, Kaiser M, van der Hoorn RAL (2011) Proteasome activity imaging and profiling characterizes bacterial effector syringolin A. *Plant Physiol* 155:477-489.

Lehfeldt C, Shirley AM, Meyer K, Ruegger MO, Cusumano JC, Viitanen PV, Strack D, Chapple C (2000) Cloning of the *SNG1* gene of *Arabidopsis* reveals a role for a serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism. *Plant Cell* 12:1295-1306.

Lenger J, Kaschani F, Lenz T, Dalhoff C, Villamor JG, Köster H, Sewald N, van der Hoorn RAL (2012) Labeling and enrichment of *Arabidopsis thaliana* matrix metalloproteases using an active-site directed, marimastat-based photoreactive probe. *Bioorgan Med Chem* 20:592-596.

Lu H, Wang Z, Shabab M, Oeljeklaus J, Verhelst SH, Kaschani F, Kaiser M, Bogyo M, van der Hoorn RAL (2013) A substrate-inspired probe monitors translocation, activation, and subcellular targeting of bacterial type III effector protease AvrPphB. *Chem Biol* 20:168-176.

471 **Lu H, Chandrasekar B, Oeljeklaus J, Misas-Villamil JC, Wang Z, Shindo T, Bogyo M,**
 472 **Kaiser M, van der Hoorn RAL** (2015) Subfamily-specific fluorescent probes for cysteine
 473 proteases display dynamic protease activities during seed germination. *Plant Physiol*
 474 168:1462-1475.
 475 **Misas-Villamil JC, Toenges G, Kolodziejek I, Sadaghiani AM, Kaschani F, Colby T,**
 476 **Bogyo M, van der Hoorn RAL** (2013a) Activity profiling of vacuolar processing enzymes
 477 reveals a role for VPE during oomycete infection. *Plant J* 73:689-700.
 478 **Misas-Villamil JC, Kolodziejek I, Crabill E, Kaschani F, Niessen S, Shindo T, Kaiser M,**
 479 **Alfano JR, van der Hoorn RAL** (2013b) *Pseudomonas syringae* pv. *syringae* uses
 480 proteasome inhibitor syringolin A to colonize from wound infection sites. *PLoS Pathog*
 481 9:e1003281.
 482 **Morimoto K, van der Hoorn RAL** (2016) The increasing impact of activity-based protein
 483 profiling in plant science. *Plant Cell Physiol in press*.
 484 **Nodwell MB, Sieber SA** (2012) ABPP methodology: introduction and overview. *Top Curr*
 485 *Chem* 324:1-41.
 486 **Richau KH, Kaschani F, Verdoes M, Pansuriya TC, Niessen S, Stüber K, Colby T,**
 487 **Overkleef HS, Bogyo M, Van der Hoorn RAL** (2012) Subclassification and biochemical
 488 analysis of plant papain-like cysteine proteases displays subfamily-specific characteristics.
 489 *Plant Physiol* 158:1583–1599.
 490 **Rooney HC, Van't Klooster JW, van der Hoorn RAL, Joosten MH, Jones JD, de Wit PJ**
 491 (2005) *Cladosporium Avr2* inhibits tomato *Rcr3* protease required for Cf-2-dependent disease
 492 resistance. *Science* 308:1783-1786.
 493 **Serim S, Haedke U, Verhelst SH** (2012) Activity-based probes for the study of proteases:
 494 recent advances and developments. *Chem Med Chem* 7:1146-1159.
 495 **Shabab M, Shindo T, Gu C, Kaschani F, Pansuriya T, Chintla R, Harzen A, Colby T,**
 496 **Kamoun S, van der Hoorn RAL** (2008) Fungal effector protein AVR2 targets diversifying
 497 defense-related Cys proteases of tomato. *Plant Cell* 20:1169–1183.
 498 **Shao F, Golstein C, Ade J, Stoutemyer M, Dixon JE, Innes RW** (2003) Cleavage of
 499 *Arabidopsis* PBS1 by a bacterial type III effector. *Science* 301:1230-1233.
 500 **Shaulov-Rotem Y, Merquiol E, Weiss-Sadan T, Moshel O, Salpeter S, Shabat D,**
 501 **Kaschani F, Kaiser M, Blum G** (2016) A novel quenched fluorescent activity-based probe
 502 reveals caspase-3 activity in the endoplasmic reticulum during apoptosis. *Chem Sci* 7: 1322-
 503 1337.
 504 **Sparkes IA, Runions J, Kearns A, Hawes C** (2006) Rapid, transient expression of
 505 fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat*
 506 *Protoc* 1: 2019-2025.

Stiti N, Chandrasekar B, Strubl L, Bartels D, Van der Hoorn RAL (2016) Nicotinamide cofactors suppress active-site labeling of aldehyde dehydrogenases. *ACS Chem Biol* *in press*.
Sueldo D, Ahmed A, Misas-Villamil J, Colby T, Tameling W, Joosten MH, van der Hoorn RAL (2014) Dynamic hydrolase activities precede hypersensitive tissue collapse in tomato seedlings. *New Phytol* 203:913-925.
Tian M, Win J, Song J, van der Hoorn RAL, van der Knaap E, Kamoun S (2007) A *Phytophthora infestans* cystatin-like protein targets a novel tomato papain-like apoplastic protease. *Plant Physiol* 143:364-377.
Van der Hoorn RAL, Leeuwenburgh MA, Bogyo M, Joosten MHAJ, Peck SC (2004) Activity profiling of papain-like cysteine proteases in plants. *Plant Physiol* 135:1170–1178.
Villamor JG, Kaschani F, Colby T, Oeljeklaus J, Zhao D, Kaiser M, Patricelli MP, van der Hoorn RAL (2013) Profiling protein kinases and other ATP binding proteins in *Arabidopsis* using Acyl-ATP probes. *Mol Cell Proteomics* 12:2481-2496.
Wang Z, Gu C, Colby T, Shindo T, Balamurugan R, Waldmann H, Kaiser M, van der Hoorn RAL (2008) Beta-lactone probes identify a papain-like peptide ligase in *Arabidopsis thaliana*. *Nat Chem Biol* 4:557-563.
Willems LI, van der Linden WA, Li N, Li KY, Liu N, Hoogendoorn S, van der Marel GA, Florea BI, Overkleeft HS (2011) Bioorthogonal chemistry: applications in activity-based protein profiling. *Acc Chem Res* 44:718-729.
Willems LI, Overkleeft HS, van Kasteren SI (2014) Current developments in activity-based protein profiling. *Bioconjug Chem* 25:1181-1191.
Wright AT, Cravatt BF (2007) Chemical proteomic probes for profiling cytochrome p450 activities and drug interactions in vivo. *Chem Biol* 14:1043-1051.
Weerapana E, Wang C, Simon GM, Richter F, Khare S, Dillon MBD, Bachovchin DA, Mowen K, Baker D, Cravatt BF (2010) Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* 468:790-795.

534 **Table I:** Twelve ways to confirm probe targets

Approach (rating)	Procedure	Examples
1.1 MS on labeled proteins (confidence: ***)	MS analysis and annotation on purified labeled proteome	SHs, glycosidases (Fig2A), PLCPs, proteasome, VPEs, ATP binding proteins
1.2 MS of labeled peptide (confidence: ****)	(Manual) MS2 analysis of labeled peptide	PsbP, GST-U20, HSP60 (Fig2B), TGG1.
1.3 Western on pull down (confidence: ****)	Purify labeled proteins and perform Western blot with target-specific antibody	Rcr3 (Fig2C), RD21, XCP2, AALP, PsbP
1.4 Immuno-precipitation (confidence: ****)	Immuno-precipitate putative target from labeled proteome and detect its labeling	None reported
1.5 Labeling characteristics (confidence: *)	Change labeling conditions to find enzymatic properties	VPEs (Fig2D)
2.1 Transient overexpression (confidence: ***)	Transiently (over)express putative target	Many SHs (Fig3A), PLCPs, VPEs
2.2 Purified proteins (confidence: **)	Label purified proteins that are expressed heterologously	AvrPphB, SjGST (Fig3B) ALDH3H1
2.3 Spiking (confidence: **)	Add recombinant protein to proteome lacking target	None reported
2.4 Crystallography (confidence: ***)	Crystallize purified labeled protein	None reported
3.1 Genetic target depletion (confidence: **)	Label proteome from organism lacking target protein (mutant/RNAi)	RD21, AALP, CTB3, Pip1 (Fig4A), VPEs, CXE12, TGG1, TGG2
3.2 Chemical target depletion (confidence: *)	Label proteome that is pre-incubated with known target inhibitor	Proteasome (Fig4B), PLCPs, VPEs
3.3 Immunodepletion (confidence: ***)	Label proteome from which target has been removed by immuno precipitation	VPE (Fig4C)

535 The number of stars (*) indicates the confidence level.

536

Epoxide (all PLCPs)
DCG-04 (Biotin)
MV201 (Bodipy-N3)
MV202 (Bodipy-Biotin)

Hydroxamate (Metalloproteases)
JL01 (Biotin)

Ethynylnaphthalene (Cytochrome P450s)
2EN-ABP (Alkyne)

Halo-acetamide (Cys residues)
CARh (Rhodamine)

Fluorophosphonate (Ser Hydrolases)
RhFP (Rhodamine)
FP-Biotin (Biotin)
TriFP (Rhodamine-Biotin)

Epoxyketone (proteasome)
MVB003 (Bodipy)
MVB070 (Bodipy-N3)
MVB072 (Bodipy-Biotin)

Aza-epoxide (VPEs)
AMS101 (Bodipy)
bAMS101 (Biotin)

Cyclophellitol aziridine (glycosidases)
JJB70 (Bodipy)
JJB111 (Biotin)

β-lactone (N-termini)
IS4 (Biotin)

Sulfonyl Fluoride (Tyr residues)
DS6B (Biotin)
DS6R (Rhodamine)

Acyl-ATP (ATP binding proteins)
BHAATP (Biotin)
DBAATP (Desthiobiotin)

AOMK (AvrPphB/CTBs)
FH11 (Rhodamine)
JOGDA1 (Bodipy)

Vinylsulfone (ALPs)
FY01 (Bodipy)

AOMK (VPEs)
JOPD1 (Bodipy)

Syrbactin (proteasome)
RhSylA (Rhodamine)

FMK (VPEs)
biotin-xVAD-fmk (Biotin)

Figure 1 Reactive groups of the chemical probes discussed in this review. Mentioned are the reactive group followed by main probe target in brackets (**bold**) and probe names with corresponding reporter tags between brackets. Binding groups that provide specificity to these probes are not shown because of limited space. Abbreviations: PLCPs, Papain-like Cys Proteases; VPEs (Vacuolar Processing Enzymes; CTBs, Cathepsin B-like proteases; ALPs, aleurain-like proteases; AOMK, acyloxymethylketone; FMK, fluoromethylketone.

Figure 2 Confirmation by target detection. A) MS analysis of purified labeled VPE proteins from Arabidopsis leaf extracts labeled with biotinylated AMS101 (7). Leaf extracts of wild-type Arabidopsis, a γVPE overexpressor (γOE) and null mutant (γKO) were labeled with and without biotinylated AMS101 (bAMS101). Labeled proteins were purified, separated on protein gel and stained by SyproRuby (left). Eight gel regions were excised, treated

with trypsin and eluted peptides analyzed by MS. Four γ VPE peptides (A-D) were identified, of which two were unique (u). Spectral counts are summarized for each excised gel region and each peptide. Adapted from (Misas-Villamil et al., 2013), with permission.

B) Annotation of the DS6B-labeled peptide of GST-U20. Arabidopsis leaf extracts were labeled with biotinylated sulfonyl fluoride probe DS6B (**10**) and labeled proteins were purified, separated on protein gel and stained by coomassie. Stained bands were excised, treated with trypsin and eluted peptides analyzed by MS. One of the detected peptides (IVAYAAEYR) carried the modification of DS6B on Y208 (underlined). Most major peaks are explained by fragmentation of the labeled peptide at amide bonds in both the peptide and the probe. Orange y-ions originate from the fragmentation of DS6B (orange). Adapted from (Gu et al., 2013), with permission.

C) Detection of Rcr3 proteases amongst purified DCG-04 labeled proteins. Secreted proteomes from wild-type and *rcr3-3* mutant tomato plants were isolated and labeled with DCG-04 (**1**). Biotinylated protein were purified, separated on protein gels and detected from a membrane using an Rcr3-specific antibody. A, excess Avr2; E, excess E-64. *, background protein, used as loading control. Adapted from (Rooney et al., 2005), with permission.

D) Labeling is consistent with enzyme characteristics. Secreted proteomes of tomato leaves containing DTT at pH 5.5 were preincubated with an excess E-64, Ac-YVAD-cmk, or iodoacetamide, or in a buffer that lacked DTT or had a pH of 7.5. These proteomes were labeled with AMS101 (**7**) and were separated on protein gels and detected by fluorescence scanning and coomassie staining. Adapted from (Sueldo et al., 2014), with permission.

Figure 3 Confirmation by target (over) expression.

A) Labeling of transiently expressed Ser hydrolases by RhFP (5). Arabidopsis Ser hydrolases FSH1 and SH1, methylesterase MES1, carboxylesterase CXE12 and Ser carboxypeptidase-like SCPL11 were transiently over expressed in *Nicotiana benthamiana* by agroinfiltration in the presence of the p19 silencing inhibitor. Leaf extracts were labeled with and without RhFP and labeled proteomes were separated on protein gels and detected by fluorescence scanning. *, endogenous Ser hydrolases, used as loading control. Adapted from (Kaschani et al., 2012), with permission.

B) SjGST but not its Y111F mutant derivative is labeled by the fluorescent sulfonyl fluoride probe DS6R (10). Wild-type (WT) and Y104F and Y111F mutants SjGST were produced in *E. coli* and treated with or without heating (ΔT , 10 min. 85 °C) and labeled with 2 μM DS6R for 15 minutes. Labeled proteomes were separated on protein gels and detected by fluorescence scanning and coomassie staining. Adapted from (Gu et al., 2013), with permission.

Figure 4 Confirmation by target depletion.

A) The bottom signal in the activity profile of secreted proteomes of tomato leaves is depleted upon suppression of the *Pip1* protease by RNAi. Secreted proteomes were isolated from the parental line (- asPip1) and two independent transformants (+ asPip1) and labeled with MV201 (**1**), a fluorescent probe for papain-like Cys proteases. Labeled proteomes were separated on protein gels and detected by fluorescence scanning. *, signal from unaffected C14 protease. Adapted from (Ilyas et al., 2015), with permission.

B) RhSylA-labeling is suppressed by proteasome inhibitor epoxomicin (Ep) but not by papain-like Cys protease inhibitor E-64d (E6d). Arabidopsis cell cultures were preincubated with or without 10 μ M E-64d or epoxomicin for 30 minutes and then labeled for 2 hours with and without 2 μ M RhSylA (**15**) for two hours. Labeled proteomes were separated on protein gels and detected by fluorescence scanning and coomassie staining. *, endogenous fluorescent signal. Adapted from (Kolodziejek et al., 2011), with permission.

C) Immunodepletion of NbVPE. Leaf extracts of *Nicotiana benthamiana* were incubated with preimmune serum or with anti-VPE serum and immunoprecipitates were removed. The immunodepleted samples were labeled with biotin-xVAD-fmk (**16**) and detected on protein blots using streptavidin HRP. *, endogenous biotinylated protein. Adapted from (Hatsugai et al., 2004), with permission.