Chemical genetic approaches for target validation

Wel, T. van der

1

General introduction

The drug discovery process

In the last decades, the field of drug discovery and development has advanced tremendously. Improvements in synthetic chemistry, DNA sequencing, protein crystallography, high-throughput screening and computational drug design, among others, have contributed to a faster and more efficient drug discovery process.[1] Despite these efforts, it still takes 10-12 years for a drug candidate to reach the market and less than 10% of the drugs that enter clinical trials actually make it to the patient.[2,3]

Drug development is a time- and resource-consuming process and consists of multiple stages (Figure 1.1).[4] The search for a new drug usually starts at the discovery and validation of a (protein) target that contributes to the pathogenesis or progression of a disease.[5] After this, an assay that reports on the functional activity of the target protein has to be set up. The process of hit identification is then initiated, which aims to identify molecules that modulate the target's function, *e.g.* its catalytic activity in case of an enzyme. Nowadays, hit identification usually involves high-throughput screening campaigns in which large libraries of small molecules are screened against the target of interest in an automated setup.[6] Identified 'hit compounds' typically have low affinity, insufficient functional efficacy, and/or a poor selectivity profile and should be optimized in a hit-to-lead optimization program.[7] The most promising candidate, the 'lead compound', then enters the stage of lead optimization, where its pharmacokinetic and pharmacodynamic properties and toxicological profile are assessed, usually in multiple animal models. If a compound successfully passes this stage, it can enter clinical trials. At this point, it is first tested in healthy human subjects to investigate the drug's pharmacokinetic and safety profile. Next, the drug is administered to a small group of patients to evaluate its efficacious dose. The final phase of clinical trials aims to assess efficacy and safety of the compound on a larger patient group. If the experimental drug passes these stages and is approved by authorities, it can enter the market.[1]

Figure 1.1 – Overview of the different stages of drug discovery and development. Estimated duration is indicated for each stage.[3,5]

Chemical tools for target discovery and validation

The discovery and development of most new drugs starts with the target discovery. Nowadays, novel targets are generally identified using genetic methods, such as RNA interference (RNAi) or CRISPR/Cas9-based library screens that abrogate the expression of proteins, in combination with a functional assay that reports on a disease-relevant phenotypic response.[8,9] The identification of novel therapeutic targets can be a challenging endeavor in itself and greatly depends on the predictive value of the employed phenotypic assay.[10,11] Genetic knockdown or knockout of a target does not always match the effects of its acute pharmacological modulation.[12] Long-term genetic disruption may have different effects on cellular physiology in comparison to acute and dynamic modulation by small molecules. In addition, genetic models may be hampered by potential compensatory mechanisms that obscure the role of the target protein.[13]

It is thus essential to validate that pharmacological modulation (*e.g.* inhibition) of the target leads to the desired phenotype, a process which is collectively referred to as target validation.[14] Target validation heavily relies on the availability of suitable chemical tools to study engagement of the compound to the intended biological target, and the ability to connect these molecular interactions with proximal biomarkers or phenotypic effects.[15,16] These chemical tools, or 'probes', can guide in selecting the best compound from a panel of drug candidates and can help to determine the dose required for complete target occupancy without inducing off-target effects.[17,18] Depending on their application, probes can be diverse in chemical structure and characteristics, and include radioligands[19], fluorescent or biotinylated small molecules[20], positron emission tomography (PET) tracers[21], photoaffinity-based probes[22], and activity-based probes.[23,24]

Structurally, chemical probes typically consist of a binding element with affinity for the intended target, a reactive group that covalently links the probe to the target, and a reporter moiety that enables visualization (*e.g.* fluorescent group) or identification (*e.g.* biotin) of the probe-bound targets (Figure 1.2A).[25] A covalent mode of action renders these probes exceptionally useful for target engagement and target validation studies, since they can irreversibly react with one or multiple target proteins in a complex proteome mixture.[26,27] Fluorescent reporter groups can be used for visualization of these probe-bound targets using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) followed by in-gel fluorescence scanning. On the other hand, biotinylated probes allow streptavidin-based enrichment of probe-bound targets, followed by tryptic digestion and liquid chromatography coupled to mass spectrometry (LC-MS) analysis of target peptides.[16] Generally, the large size of a reporter group is a limiting factor, since it reduces probe solubility and cell-permeability. This limitation can be addressed by the use of 'two-step probes', which typically possess bioorthogonal ligation handles that enable conjugation of reporter groups in a later experimental stage.[28–30] Including an inhibitor

pre-incubation step allows a competitive experimental setup to profile a compound's proteome-wide target engagement (Figure 1.2B).

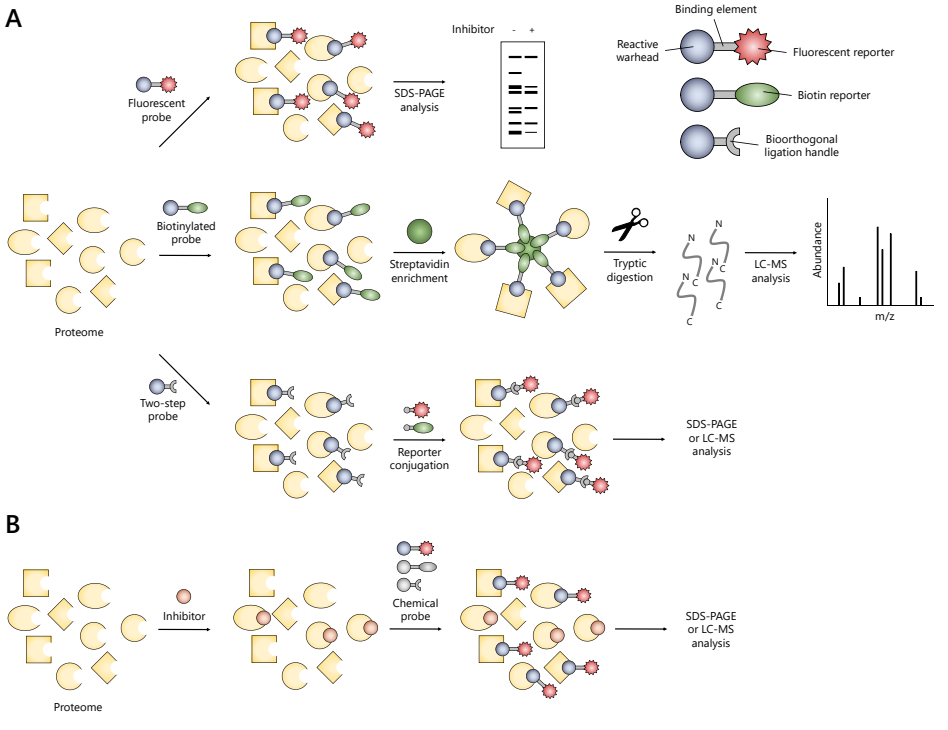

Figure 1.2 – Chemical probes as tools to study target engagement and target validation. (**A**) Chemical probes have an irreversible, covalent mode of action and can be used to label one or multiple proteins in a complex proteome sample. Fluorescent reporter groups can be used for visualization of probe-bound targets using SDS-PAGE and in-gel fluorescence scanning. Biotinylated probes enable streptavidin-based enrichment of probe-bound targets, followed by tryptic digestion and LC-MS analysis of target peptides. Two-step probes are versatile tools harboring bioorthogonal ligation handles, which enable conjugation to reporter tags in a later experimental stage. (**B**) Competitive experiment to study target engagement of inhibitors using chemical probes. The proteome is pre-incubated with inhibitor and subsequently labeled with a chemical probe of choice. Figures are modified from literature.[17]

A variety of chemical probes has been developed for many protein classes, in particular enzymes, including kinases[31], phosphatases[32], serine hydrolases[23,33], proteases[34,35], and glycosidases.[36] On the one hand, non-selective probes that bind to a broad spectrum of protein targets can serve as valuable tools for target discovery[37,38] or to investigate a compound's selectivity profile using chemical proteomics techniques.[23,31,39,40] On the other hand, selective probes are exceptionally suited for cellular target engagement studies[41] or for molecular imaging of the protein target.[35,42,43]

Chemical genetics

The development of a selective chemical probe that specifically targets a particular protein can be challenging to achieve due to the off-target activity towards structurally and/or functionally related homologs, *e.g.* other members within the same protein class. Selectivity can be improved in an iterative process of design, synthesis, testing, and refining the original scaffold, but this is a time-consuming and tedious effort. To overcome such limitations, various chemical genetic methods and technologies have been developed that allow for selective modulation of specific targets using small molecules (Figure 1.3).

Chemical genetics combines the specificity of genetics with benefits of acute, pharmacological modulation by a small molecule. It generally uses engineered, mutant proteins that can accommodate unnatural substrates or modified ligands that do not affect other, native proteins in the cellular environment. The most prominent example is the 'bump-hole' technology that has been successfully applied to a wide range of protein families, including kinases[44], GTPases[45], proteases[46], phosphatases[47], receptors[48] and various types of transferases.[49] This strategy is based on mutagenesis of bulky amino acid residues into smaller residues, creating an additional pocket ('hole') in the target protein's active site (Figure 1.3A). This engineered, mutant protein can accommodate bulkier ligands than its wild-type counterpart, making these ligands mutant-specific.[50–52] Since off-targets of the original ligand maintain their native active site structure, these proteins are typically not targeted by these bulky analogues.[53] An alternative to this steric complementation approach is the use of charge complementation, where active site residues are mutated to induce electronic repulsion to the original ligand while establishing novel electrostatic interactions (*e.g.* hydrogen bonds) with a modified ligand (Figure 1.3B).[54]

Although these concepts allow one to selectively modulate a target protein and study its function, the lack of a covalent binding mode limits the opportunities for use in target engagement studies. Covalent, irreversible ligands can have additional advantages over reversible ligands, such as sustained target occupancy, lower susceptibility to competition with high intracellular substrate concentrations and a pharmacodynamic profile that is dependent on the target's *de novo* protein synthesis rate. Most importantly, ligands with a covalent binding mode are powerful tools by serving as chemical probes. Shokat and co-workers introduced the chemical genetic strategy of 'covalent complementarity', which involves mutagenesis of active site residues into cysteines (Figure 1.3C).[55,56] The thiol group of cysteine can function as a nucleophile to covalently react with electrophilic ligands.[37,57] Although conceptually promising, mutant proteins may suffer from distorted protein folding, reduced catalytic activity or otherwise impeded protein function.[55] Another general limitation of all described chemical genetic strategies remains that they rely on overexpression of the mutant protein of interest, which in itself may

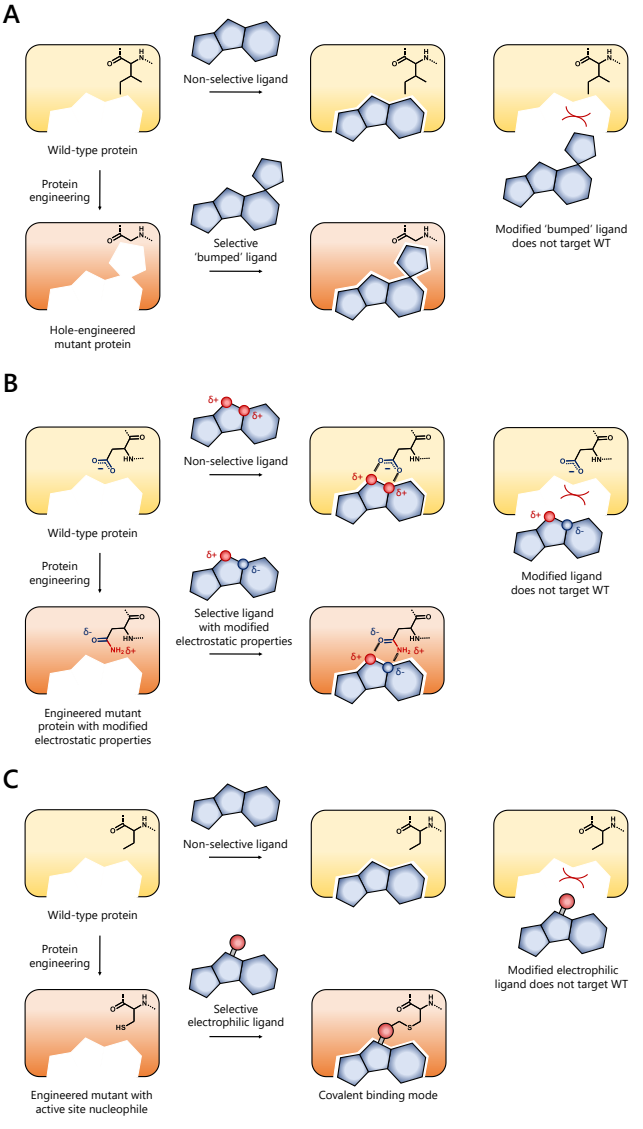

Figure 1.3 – Chemical genetic approaches to establish selective modulation of specific protein targets by small molecules. (**A**) Steric complementation or 'bump-hole' strategy. The target protein is engineered by mutation of a bulky amino acid residue in the active site into a smaller residue ('hole'). A non-selective ligand is modified by incorporation of a bulky substituent ('bump') to fill the mutant's pocket. This modified bumped ligand does not target the wild-type protein due to steric clash. (**B**) Charge complementation strategy. The target protein is engineered based on electrostatic interactions in the active site, *e.g.* by modification of hydrogen bonding patterns. A non-selective ligand is modified by changing its electronic properties to complement the mutated residue in the engineered protein pocket. This modified ligand does not target the wild-type protein due to electrostatic repulsion and/or loss of affinity. (**C**) Covalent complementarity strategy. The target protein is engineered by mutating an active site residue into a nucleophilic residue, *e.g.* cysteine. A non-selective ligand is modified by incorporation of an electrophile that covalently reacts with the nucleophilic residue. This modified electrophilic ligand does not target the wild-type protein due to a lack of nearby nucleophilic residues. Figures modified from literature.[58,59]

induce artefacts and disturbs normal cellular physiology.[60] Further improvements are thus required to facilitate the application of chemical genetics-based probes in target engagement and target validation studies on endogenously expressed proteins.

Aim and outline

The aim of the research described in this thesis is to develop a chemical genetic strategy that can be used for target engagement and target validation studies.

Chapter 2 describes the development of a chemical genetic toolbox for visualization of engineered kinases and their target engagement. Using the tyrosine kinase FES as exemplary target, various cysteine point mutants are generated and expressed, followed by comprehensive biochemical profiling. After identification of a suitable mutant, structure-based design is employed to synthesize mutant-specific probes that covalently react with the introduced cysteine. Next, the selectivity and covalent mode of action of the compounds are characterized in more detail. Cellular target engagement studies are performed to investigate the *in situ* potency. Lastly, broader application of the generated tools is examined on a panel of wild-type and cysteine mutant kinases.

Chapter 3 applies the tools developed in chapter 2 to investigate the role of FES activity in myeloid differentiation. The use of CRISPR/Cas9 gene editing allows the visualization of an endogenous mutant FES kinase in a relevant model system. Gel-based labeling experiments reveal the cellular target engagement profile of the probe, and chemical proteomics is used to study its proteome-wide selectivity. This chapter also demonstrates the power of the chemical genetic strategy to dissect on-target from off-target effects using mutant and wild-type cells, respectively.

Chapter 4 reports on the application of the chemical genetic tools from chapter 2 and the generated mutant cell line from chapter 3 to study the role of FES activity in neutrophil phagocytosis. A flow cytometry assay is used to measure phagocytic uptake of fluorescent *E. coli* by neutrophils, which reveals that FES plays a role in this process. Guided by a substrate profiling experiment, a novel FES substrate is identified and validated *in situ*. Phospho-specific immunoblot experiments are used to gain insight in the underlying molecular mechanism, which results in a model proposing a novel role of FES in neutrophil phagocytosis.

Chapter 5 extends the chemical genetic strategy described in chapter 2 to diacylglycerol lipase α (DAGLα), an enzyme belonging to the family of serine hydrolases. This chapter describes the first steps towards a strategy to subtype-selectively inhibit DAGLα without affecting its structurally related homolog DAGLβ. To this end, DAGLα cysteine mutants are designed based on a homology model, followed by biochemical profiling using activity-based protein profiling (ABPP) and substrate hydrolysis assays. The design and synthesis of mutant-selective inhibitors of DAGLα are described, followed by characterization of its irreversible, covalent binding mode. Competitive ABPP is used to

investigate the ability of the compounds to subtype-selectively target DAGLα in presence of DAGLβ.

Chapter 6 discusses the development and miniaturization of a biochemical activity assay for monoacylglycerol lipase (MAGL), a serine hydrolase that is currently considered as a therapeutic target for various diseases. Currently, nearly all MAGL inhibitors have an irreversible mode of action and the number of reversible compounds is limited. To identify novel reversible MAGL inhibitors, the assay was used in a high-throughput screening campaign on 233,820 unique compounds. Hit validation using orthogonal ABPP experiments results in a qualified list of hit compounds that constitute starting points for the development of novel, reversible MAGL inhibitors as well as chemical probes for use in chemical genetic strategies.

Chapter 7 provides a summary of the work described in this thesis and discusses future directions.

References

1. Scannell, J. W., Blanckley, A., Boldon, H. & Warrington, B. Diagnosing the decline in pharmaceutical R&D efficiency. *Nature Reviews Drug Discovery* **11**, 191–200 (2012).
2. Paul, S. M. *et al.* How to improve RD productivity: The pharmaceutical industry's grand challenge. *Nature Reviews Drug Discovery* **9**, 203–214 (2010).
3. Waring, M. J. *et al.* An analysis of the attrition of drug candidates from four major pharmaceutical companies. *Nature Reviews Drug Discovery* **14**, 475–486 (2015).
4. DiMasi, J. A., Grabowski, H. G. & Hansen, R. W. Innovation in the pharmaceutical industry: New estimates of R&D costs. *J. Health Econ.* **47**, 20–33 (2016).
5. Hughes, J. P., Rees, S. S., Kalindjian, S. B. & Philpott, K. L. Principles of early drug discovery. *British Journal of Pharmacology* **162**, 1239–1249 (2011).
6. Copeland, R. A. *Evaluation of Enzyme Inhibitors in Drug Discovery: A Guide for Medicinal Chemists and Pharmacologists: Second Edition*. (Wiley, 2013).
7. Keseru, G. M. & Makara, G. M. Hit discovery and hit-to-lead approaches. *Drug Discovery Today* **11**, 741–748 (2006).
8. Tiedemann, R. E. *et al.* Kinome-wide RNAi studies in human multiple myeloma identify vulnerable kinase targets, including a lymphoid-restricted kinase, GRK6. *Blood* **115**, 1594–1604 (2010).
9. Schuster, A. *et al.* RNAi/CRISPR Screens: from a Pool to a Valid Hit. *Trends in Biotechnology* **37**, 38–55 (2019).
10. Moffat, J. G., Rudolph, J. & Bailey, D. Phenotypic screening in cancer drug discovery-past, present and future. *Nat. Rev. Drug Discov.* **13**, 588–602 (2014).
11. Vincent, F. *et al.* Developing predictive assays: The phenotypic screening 'rule of 3'. *Sci. Transl. Med.* **7**, 293ps15 (2015).
12. Weiss, W. A., Taylor, S. S. & Shokat, K. M. Recognizing and exploiting differences between RNAi and small-molecule inhibitors. *Nat. Chem. Biol.* **3**, 739–744 (2007).
13. Knight, Z. A. & Shokat, K. M. Chemical Genetics: Where Genetics and Pharmacology Meet. *Cell* **128**, 425–430 (2007).
14. Nicolaou, K. C. The chemistry-biology-medicine continuum and the drug discovery and development process in academia. *Chemistry and Biology* **21**, 1039–1045 (2014).
15. Bunnage, M. E., Chekler, E. L. P. & Jones, L. H. Target validation using chemical probes. *Nat. Chem. Biol.* **9**, 195–199 (2013).
16. Drewes, G. & Knapp, S. Chemoproteomics and Chemical Probes for Target Discovery. *Trends in Biotechnology* **36**, 1275–1286 (2018).
17. Nomura, D. K., Dix, M. M. & Cravatt, B. F. Activity-based protein profiling for biochemical pathway discovery in cancer. *Nat. Rev. Cancer* **10**, 630–638 (2010).
18. Simon, G. M., Niphakis, M. J. & Cravatt, B. F. Determining target engagement in living systems. *Nat. Chem. Biol.* **9**, 200–205 (2013).
19. Maguire, J. J., Kuc, R. E. & Davenport, A. P. Radioligand binding assays and their analysis. *Methods Mol. Biol.* **897**, 31–77 (2012).
20. Dittus, L., Werner, T., Muelbaier, M. & Bantscheff, M. Differential Kinobeads Profiling for Target Identification of Irreversible Kinase Inhibitors. *ACS Chem. Biol.* **12**, 2515–2521 (2017).
21. Sharma, R. & Aboagye, E. Development of radiotracers for oncology - The interface with pharmacology. *British Journal of Pharmacology* **163**, 1565–1585 (2011).
22. Soethoudt, M. *et al.* Selective Photoaffinity Probe That Enables Assessment of Cannabinoid CB 2 Receptor Expression and Ligand Engagement in Human Cells. *J. Am. Chem. Soc.* **140**, 6067–6075 (2018).
23. Baggelaar, M. P. *et al.* Development of an activity-based probe and in silico design reveal highly selective inhibitors for diacylglycerol lipase-α in brain. *Angew. Chemie - Int. Ed.* **52**, 12081–12085 (2013).
24. Ogasawara, D. *et al.* Rapid and profound rewiring of brain lipid signaling networks by acute diacylglycerol lipase inhibition. *Proc. Natl. Acad. Sci.* **113**, 26–33 (2016).
25. Niphakis, M. J. & Cravatt, B. F. Enzyme Inhibitor Discovery by Activity-Based Protein Profiling. *Annu. Rev. Biochem.* **83**, 341–377 (2014).
26. Schenone, M., Dančík, V., Wagner, B. K. & Clemons, P. A. Target identification and mechanism of action in chemical biology and drug discovery. *Nature Chemical Biology* **9**, 232–240 (2013).
27. Arrowsmith, C. H. *et al.* The promise and peril of chemical probes. *Nat. Chem. Biol.* **11**, 536–541 (2015).
28. Rostovtsev, V. V., Green, L. G., Fokin, V. V. & Sharpless, K. B. A stepwise huisgen cycloaddition process: Copper(I)-catalyzed regioselective 'ligation' of azides and terminal alkynes. *Angew. Chemie - Int. Ed.* **41**, 2596–2599 (2002).
29. Van Rooden, E. J. *et al.* Mapping in vivo target interaction profiles of covalent inhibitors using chemical proteomics with label-free quantification. *Nat. Protoc.* **13**, 752–767 (2018).

30. van Rooden, E. J., Bakker, A. T., Overkleeft, H. S. & van der Stelt, M. Activity-Based Protein Profiling. in *eLS* 1–9 (John Wiley & Sons, Ltd, 2018).
31. Zhao, Q. *et al.* Broad-spectrum kinase profiling in live cells with lysine-targeted sulfonyl fluoride probes. *J. Am. Chem. Soc.* **139**, 680–685 (2017).
32. Kawaguchi, M., Hanaoka, K., Komatsu, T., Terai, T. & Nagano, T. Development of a highly selective fluorescence probe for alkaline phosphatase. *Bioorganic Med. Chem. Lett.* **21**, 5088–5091 (2011).
33. Liu, Y., Patricelli, M. P. & Cravatt, B. F. Activity-based protein profiling: The serine hydrolases. *Proc. Natl. Acad. Sci.* **96**, 14694–14699 (1999).
34. Sexton, K. B., Witte, M. D., Blum, G. & Bogyo, M. Design of cell-permeable, fluorescent activity-based probes for the lysosomal cysteine protease asparaginyl endopeptidase (AEP)/legumain. *Bioorganic Med. Chem. Lett.* **17**, 649–653 (2007).
35. de Bruin, G. *et al.* A Set of Activity-Based Probes to Visualize Human (Immuno)proteasome Activities. *Angew. Chemie - Int. Ed.* **55**, 4199–4203 (2016).
36. Kallemeijn, W. W. *et al.* Investigations on therapeutic glucocerebrosidases through paired detection with fluorescent activity-based probes. *PLoS One* **12**, e0170268 (2017).
37. Backus, K. M. *et al.* Proteome-wide covalent ligand discovery in native biological systems. *Nature* **534**, 570–574 (2016).
38. Hacker, S. M. *et al.* Global profiling of lysine reactivity and ligandability in the human proteome. *Nat. Chem.* **9**, 1181–1190 (2017).
39. Lanning, B. R. *et al.* A road map to evaluate the proteome-wide selectivity of covalent kinase inhibitors. *Nat. Chem. Biol.* **10**, 760–767 (2014).
40. van Esbroeck, A. C. M. *et al.* Activity-based protein profiling reveals off-target proteins of the FAAH inhibitor BIA 10-2474. *Science* **356**, 1084–1087 (2017).
41. Dubach, J. M. *et al.* Quantitating drug-Target engagement in single cells in vitro and in vivo. *Nat. Chem. Biol.* **13**, 168–173 (2017).
42. Oresic Bender, K. *et al.* Design of a Highly Selective Quenched Activity-Based Probe and Its Application in Dual Color Imaging Studies of Cathepsin S Activity Localization. *J. Am. Chem. Soc.* **137**, 4771–4777 (2015).
43. Kallemeijn, W. W. *et al.* Endo-β-Glucosidase Tag Allows Dual Detection of Fusion Proteins by Fluorescent Mechanism-Based Probes and Activity Measurement. *ChemBioChem* **17**, 1698–1704 (2016).
44. Bishop, A. C. *et al.* A chemical switch for inhibitor-sensitive alleles of any protein kinase. *Nature* **407**, 395–401 (2000).
45. Hwang, Y.-W. & Miller, D. L. A Mutation That Alters the Nucleotide Specificity of Elongation Factor Tu, a GTP Regulatory Protein. *J. Biol. Chem.* **262**, 13081–13085 (1987).
46. Morell, M. *et al.* Coupling protein engineering with probe design to inhibit and image matrix metalloproteinases with controlled specificity. *J. Am. Chem. Soc.* **135**, 9139–9148 (2013).
47. Hoffman, H. E., Blair, E. R., Johndrow, J. E. & Bishop, A. C. Allele-specific inhibitors of protein tyrosine phosphatases. *J. Am. Chem. Soc.* **127**, 2824–2825 (2005).
48. Tedesco, R., Thomas, J. A., Katzenellenbogen, B. S. & Katzenellenbogen, J. A. The estrogen receptor: A structure-based approach to the design of new specific hormone-receptor combinations. *Chem. Biol.* **8**, 277–287 (2001).
49. Luo, M. Current chemical biology approaches to interrogate protein methyltransferases. *ACS Chemical Biology* **7**, 443–463 (2012).
50. Belshaw, P. J., Schoepfer, J. G., Liu, K. -Q, Morrison, K. L. & Schreiber, S. L. Rational Design of Orthogonal Receptor–Ligand Combinations. *Angew. Chemie Int. Ed. English* **34**, 2129–2132 (1995).
51. Au-Yeung, B. B. *et al.* A genetically selective inhibitor demonstrates a function for the kinase Zap70 in regulatory T cells independent of its catalytic activity. *Nat. Immunol.* **11**, 1085–1092 (2010).
52. Carter-O'Connell, I. *et al.* Combining Chemical Genetics with Proximity-Dependent Labeling Reveals Cellular Targets of Poly(ADP-ribose) Polymerase 14 (PARP14). *ACS Chem. Biol.* **13**, 2841–2848 (2018).
53. Koh, J. T. Engineering selectivity and discrimination into ligand-receptor interfaces. *Chemistry and Biology* **9**, 17–23 (2002).
54. Tairis, N., Gabriel, J. L., Soprano, K. J. & Soprano, D. R. Alteration in the retinoid specificity of retinoic acid receptor-β by site-directed mutagenesis of Arg269 and Lys220. *J. Biol. Chem.* **270**, 18380–18387 (1995).
55. Garske, A. L., Peters, U., Cortesi, A. T., Perez, J. L. & Shokat, K. M. Chemical genetic strategy for targeting protein kinases based on covalent complementarity. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 15046–15052 (2011).
56. Xiao, J. *et al.* A coupled protein and probe engineering approach for selective inhibition and activity-based probe labeling of the caspases. *J. Am. Chem. Soc.* **135**, 9130–9138 (2013).
57. Barf, T. & Kaptein, A. Irreversible protein kinase inhibitors: Balancing the benefits and risks. *J. Med. Chem.* **55**, 6243–6262 (2012).
58. Islam, K. Allele-specific chemical genetics: Concept, strategies, and applications. *ACS Chem. Biol.* **10**, 343–363 (2015).

59. Islam, K. The Bump-and-Hole Tactic: Expanding the Scope of Chemical Genetics. *Cell Chem. Biol.* **25**, 1171–1184 (2018).
60. Prelich, G. Gene overexpression: Uses, mechanisms, and interpretation. *Genetics* **190**, 841–854 (2012).