Inhibitor selectivity: profiling and prediction

Janssen, A.P.A.

Cover Page

The following handle holds various files of this Leiden University dissertation:
http://hdl.handle.net/1887/71808

Author: Janssen, A.P.A.
Title: Inhibitor selectivity: profiling and prediction
Issue Date: 2019-05-01

Inhibitor Selectivity:
Profiling and Prediction

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit van Leiden,
op gezag van Rector Magnificus Prof. Dr. C.J.J.M. Stolker,
volgens het besluit van het College voor Promoties
te verdedigen op woensdag 1 mei 2019
klokke 15:00 uur

door

Antonius Petrus Arnoldus Janssen
Geboren te Nootdorp op 7 december 1989

Promotiecommissie

Promotor: Prof. Dr. M. van der Stelt

Co-promotor: Dr. G.J.P. van Westen

Overige leden: Prof. Dr. H.S. Overkleeft

Prof. Dr. C.A.A. van Boeckel

Prof. Dr. H. van Vlijmen

Prof. Dr. A.P. IJzerman

Prof. Dr. M. Kaiser
Universität Duisburg-Essen, Essen, Germany

Dr. I. Dijkgraaf
Maastricht University, Maastricht, The Netherlands

Cover image: *Sharpest ever view of the Andromeda Galaxy* - NASA, ESA,
J. Dalcanton (University of Washington, USA), B. F. Williams (University of Washington,
USA), L. C. Johnson (University of Washington, USA), the PHAT team, and R. Gendler.
Constellations overlaid by the author.

You're only given a little spark of madness, if you lose that...
you're nothing.
Robin Williams.

Table of Contents

General introduction

The drug discovery process

The discovery and development of new small molecule medicines is a long and expensive process, which requires numerous fields of research to come together. The general timeline for the discovery and development of new drugs is depicted in Figure 1.1.[1,2] Initially, in the drug discovery phase, a biological target has to be found and validated. Next, the stages of hit finding and optimization are aimed at the discovery of molecules to effectively modulate the target. The best molecule, a so-called lead, is then taken into the lead optimization phase, where typically animal models are used to optimize the pharmacokinetic, efficacy
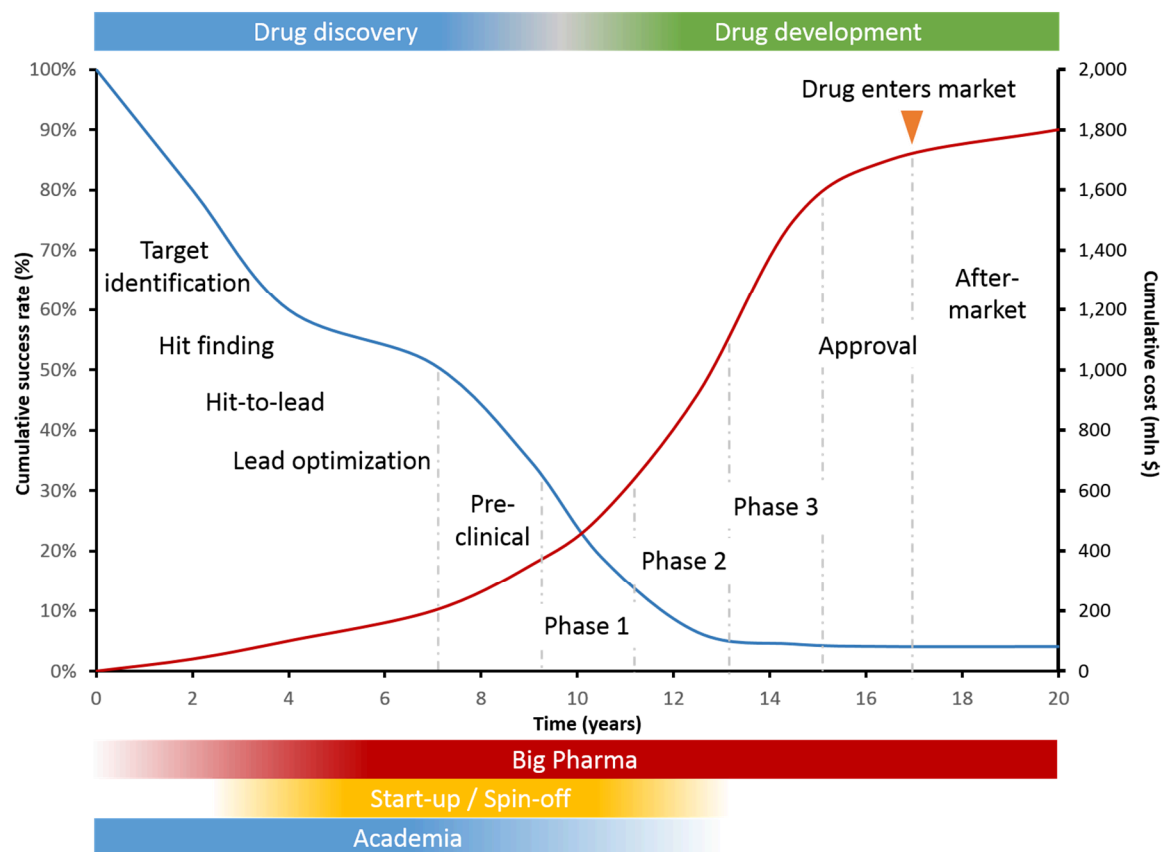
Figure 1.1 | Overview of the general stages of drug discovery and development, the timeline, and cumulative cost and success rate. Figure constructed based on Ref 1 with data from Ref 2.[1,2]

and safety profile of the leads, through iterative rounds of synthesis and testing. Once the drug candidate has been selected for drug development, extensive optimisation of the pharmaceutical formulation and pre-clinical toxicological profiling is performed, before the compound is tested in humans. The next 3 stages are referred to as the clinical trials. Phase 1 typically utilizes healthy volunteers mainly to assess the pharmacokinetic and safety profile of the drug candidate. Phase 2 has an emphasis on finding an efficacious dose in a small cohort of patients. If successful, phase 3 enrols a large patient cohort to determine the efficacy and safety of the experimental drug. It can then be submitted for approval by the authorities. Even after the drug enters the market, data is gathered to assess safety in the larger population, a process often referred to as phase 4. This stringent process of clinical trials is associated with high attrition rates, meaning that less than 1 in 10 devised therapies actually make it to the patient. This is exemplified in Figure 1.1 where the cumulative cost and success rate of drug discovery and development is tracked over time.[1] Failure in the clinical stage is mostly due to toxicity or to lack of efficacy. The high cost and slow progress in drug development is a major concern, as this inevitably pushes up drug prices.

Target-based drug discovery

As discussed above, modern drug discovery projects typically start from either a well-established biological target, or by validating a novel target. The biological target can be

anything from a co-factor to a protein-complex and generally performs a function which is, in the given condition, unwanted. Currently, four major classes of drug targets can be discerned: transporters, ion channels, receptors (nuclear or G-protein coupled) and enzymes, together accounting for just under 90% of the FDA approved drugs.[3] This thesis will focus primarily on enzymes. Pharmaceutical intervention in the functioning of enzymes, barring a few exceptions, always aims at inhibiting the reaction they catalyze. This is typically achieved by compounds that block the binding site of the natural substrate of the enzyme, called competitive inhibition.[4] To find the molecules that have the required interactions to efficiently block the binding site of a given protein, a plethora of techniques have been developed which found their way to the drug industry's toolbox.

The most important approach to find new inhibitors is, although conceptually simple, technically very challenging. So-called high-throughput screening campaigns typically use some biochemical or cellular assay capable of measuring the protein activity in a highly controlled sample and run this assay for thousands or even millions of compounds. This kind of screening is performed by specialized, highly expensive robots and are mostly restricted to the pharmaceutical industry. Smaller screens can be performed with less sophisticated robots or even by hand.

With the dawn of computers, it did not take long before they were put to use in the drug discovery world.[5] This field of research often utilizes virtual screenings, which are enabled by the availability of many 3D structures of proteins, either from crystal structures, NMR-structures, or homology models.[6] In virtual screenings, large numbers of molecules are tried, or docked, inside the (proposed) binding site of a protein, making it essentially the digital mimic of above-mentioned high-throughput screenings. The field of computational drug discovery comprises a broad set of other applications, which are not necessarily structure-based. Quantitative structure activity relations (QSAR) or machine learning approaches also strongly contribute, and the field, together with the number of developed techniques, is still growing. With the advent of machine learning and artificial intelligence this growth is likely to persist.

Off-target activity

The key of target-based drug discovery is that a specific enzyme is targeted to be inhibited. This inhibition is known or predicted to have a designated effect on physiology, which is supposed to be beneficial for the therapeutic indication at hand. With an estimated number of around 20,000 translated genes, it seems inevitable that the binding site of some proteins will be highly similar.[7] This is especially true for enzymes one step preceding or following the targeted enzyme in an enzyme cascade, as the product and substrate of these are identical. It is also the case for protein families within (large) protein families, with high overall similarity, such as the kinases or serine hydrolases. The high similarity of binding sites between proteins can lead to small molecules unintendedly inhibiting other proteins. These are referred to as off-targets.

Interfering with a number of proteins simultaneously can have additive beneficial effects and is referred to as polypharmacology.[8] Some drugs are actively tuned to inhibit multiple targets, which can be a challenging undertaking.[9] Usually however, small molecule

inhibitors are carefully optimized to minimize the cross-reactivity with off-targets, to minimize the chance of unwanted side effects or toxicity.[10]

The endocannabinoid system

The endocannabinoid system (ECS) influences many physiological processes in the human body, including food intake, energy balance, motor coordination, pain sensation, memory formation and anxiety.[11,12] The ECS has, therefore, been under active investigation for therapeutic exploitation in which its receptors and several metabolic enzymes serve as possible drug targets.[13,14] There are two main cannabinoid receptors, CB_1R and CB_2R, which belong to the family of GPCRs. They are activated by two endogenous ligands, *i.e.* anandamide (AEA) and 2-arachidonoyl glycerol (2-AG).[15,16] The production and degradation of these endocannabinoids is mainly performed by serine hydrolases (Figure 1.2). Diacylglycerol lipase-α and -β (DAGL-α and -β) are the main enzymes responsible for the biosynthesis of 2-AG through the *sn-1*-specific hydrolysis of diacylglycerol (DAG).[17–19] The DAGL-α isoform is expressed mainly in the brain, whereas the DAGL-β isoform is predominantly found in the periphery, and is highly abundant in macrophages.[20,21] Monoacylglycerol lipase (MAGL) and α/β-hydrolase-domain containing protein 6 and 12 (ABHD6 and ABHD12) together account for 99% of the 2-AG hydrolysis to arachidonic acid (AA) and glycerol in the brain.[22,23] The Ca^{2+}-dependent biosynthesis of endogenous AEA is mediated by the subsequent actions of PLA2G4E[24] and *N*-acylphosphatidylethanolamine-phospholipase D (NAPE-PLD) or ABHD4[25], although other biosynthetic pathways have also been uncovered.[13,14,26] Fatty acid amide hydrolase (FAAH) is the key enzyme for the hydrolysis of AEA to AA.[27,28] Several drugs targeting the ECS have already entered the market: Δ^9-tetrahydrocannabinol (THC, marketed as Marinol®), cannabidiol (CBD, marketed as Epidiolex®), a combination of THC and CBD (marketed as Sativex®) and Rimonabant®. The latter was withdrawn from the market after the discovery of psychological side effects. Several FAAH and MAGL inhibitors have entered clinical trials, but they have not (yet) reached the market.[29–34]

The exact function and tissue specific roles of the ECS are still poorly understood.[14] Inhibitors of the metabolic enzymes are thus crucial to investigate the biological role of the hydrolases and may serve as drug candidates to modulate the endocannabinoid levels in human disease. With its central role in the production of the main ECS signaling lipid 2-AG, modulation of DAGL activity holds large therapeutic promise. Specifically, DAGL modulation might aid in the alleviation of symptoms in neuroinflammatory conditions, such as observed in Parkinson's and Alzheimer's disease.[35,36]

Activity-based protein profiling

The DAGLs, like most of the synthetic and degradative enzymes in the ECS, belong to the superfamily of the serine hydrolases (SH).[37] This enzyme family has a conserved mechanism of action where a nucleophilic serine is used to hydrolyze ester or amide bonds. In the process, the serine forms a covalent bond with the substrate. This covalent intermediate is exploited in mechanism-based, covalent inhibitors, as well as in activity-based protein profiling (ABPP).[38] The enzymes of the endocannabinoid system have been
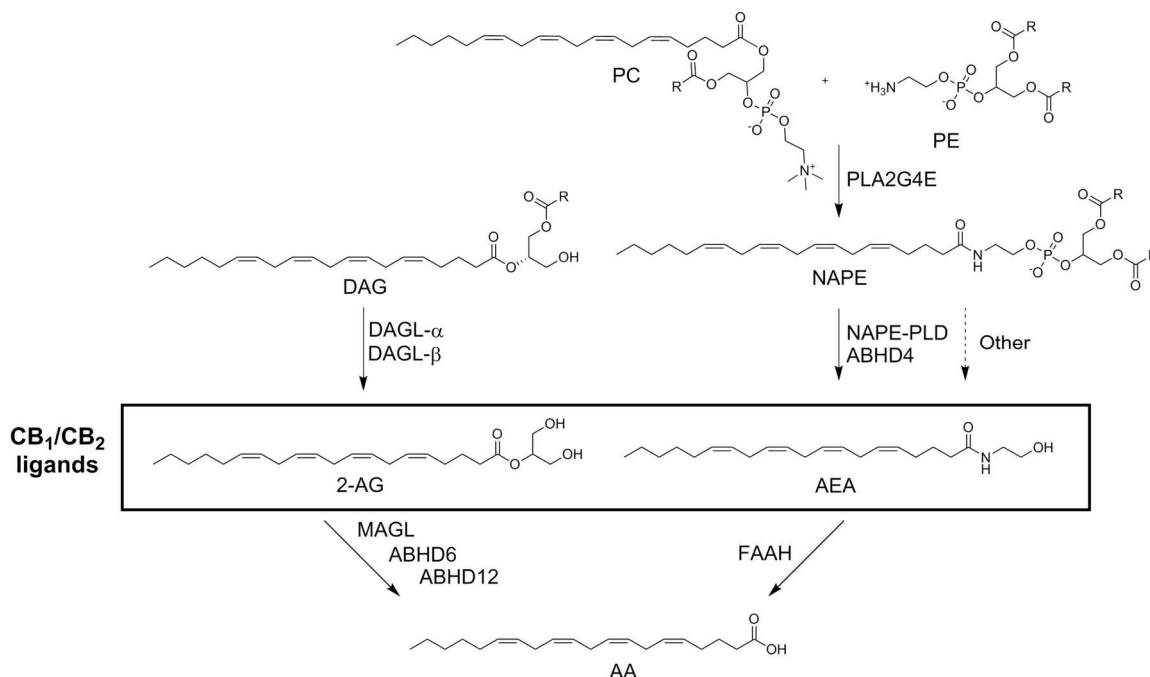
Figure 1.2 | Schematic overview of the main biosynthetic pathways within the endocannabinoid system. All enzymes except NAPE-PLD belong to the serine hydrolase protein family. PC: phosphatidylcholine; PE: phosphatidylethanolamine; DAG: diacylglycerol; NAPE: *N*-acylphosphatidylethanolamine; AA: arachidonic acid; PLA2G4E: phospholipase A2 group IV E; DAGL: diacylglycerol lipase; NAPE-PLD: *N*-acylphosphatidylethanol-amine phospholipase D; MAGL: monoacylglycerol lipase; ABHD: α/β-hydrolase-domain containing protein; FAAH: fatty acid amide hydrolase.

extensively investigated using ABPP.[24,39–45] Specifically the development of the tailored activity-based probe (ABP) MB064 to study the DAGLs was instrumental in the further development of several inhibitor classes.[39,42,43]

Activity-based protein profiling (ABPP) for serine hydrolases was introduced in the late ´90s (Figure 1.3).[46] In ABPP a chemical probe, typically consisting of a reactive 'warhead' and a reporter tag, reacts with the catalytically active nucleophilic serine of a serine hydrolase. The reporter tag can be either a fluorophore to visualize the probe-protein adduct by SDS-PAGE and fluorescence scanning,[47] or a biotin-group to enrich proteins from proteomes for identification by high resolution LC-MS/MS[48] or visualization by western blotting.[46] As the labeling of enzymes is typically activity dependent, inhibitor function can be studied using so-called competitive ABPP. This technique is used in drug discovery to efficiently profile activity of inhibitors in a wide variety of proteomes. Importantly, the selectivity of inhibitors over a protein family can also be investigated in native biological samples. Other advances in chemical biology, such as photoaffinity labeling, efficient bio-orthogonal chemistry, and improvements in analysis techniques such as proteomics, have broadened the scope of applications of this protein profiling technique significantly. The flexibility in potential protein sources and robust application to more complex samples make ABPP a powerful technique in all stages of drug discovery.
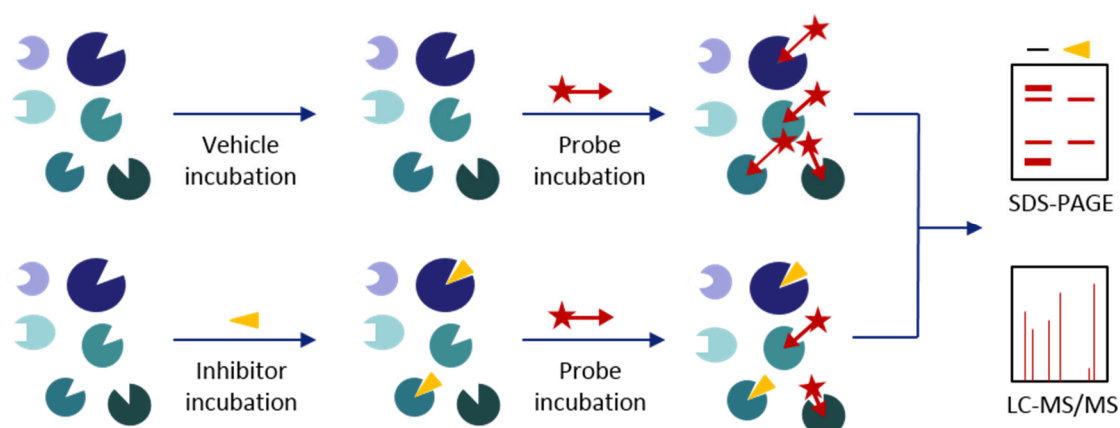
Figure 1.3 | General scheme of activity-based protein profiling. Probe binds specifically to active enzymes, which enables competitive ABPP (bottom). Probe labeling can be visualized using SDS-PAGE or LC-MS/MS analysis.

Aim and outline of this thesis

The aim of the research presented in this thesis is to develop methods to assess or predict the selectivity of (endocannabinoid related) inhibitors, and to use those methods in the discovery of better inhibitors of diacylglycerol lipase. The outline of this thesis is as follows:

Chapter 2 discusses the activity- and selectivity-driven optimization of a new scaffold of DAGL inhibitors found in a previously reported high-throughput screening campaign.[49] The structure activity relations are studied in detail and the optimized inhibitor is fully profiled biochemically. This is then further tested in cultures of murine neuroblastoma cells, before finally being administered to mice in a proof of target engagement study.

Chapter 3 combines the structural insights of Chapter 2 with those from literature[42] to enhance the physicochemical properties of the inhibitors from Chapter 2. The dual aim is to generate more drug-like compounds by enhancing pharmacokinetic properties, combined with increasing the topological polar surface area to restrict brain access. In this way, the more peripherally expressed DAGL-β isoform could be targeted selectively, reducing the risk of side effects mediated centrally. A small library of piperazine derivatives is synthesized and profiled biochemically, before testing the most promising compounds *in situ*.

Chapter 4 focusses more on the fundamental aspects of binding of covalent inhibitors. DH376 derivatives, wherein the number and positioning of the nitrogen atoms in the heterocyclic ring vary, are synthesized.[50] These are tested in an adapted surrogate substrate assay to determine the kinetics of enzyme inhibition and to study the role of affinity (K_i) and reactivity (k_{inact}). Surprisingly, the heterocycle is found to be more important in the former than in the latter. Insight in the specific binding kinetics is expected to aid in the design of inhibitors that are more potent and more selective, by increasing K_i and reducing k_{inact}, respectively.

Chapter 5 presents the work on the selectivity profiling of the experimental drug BIA 10-2474, which was designed as a FAAH inhibitor. This drug made headlines worldwide when, during a phase 1 first-in-human clinical trial in January 2016, a healthy volunteer passed away and four others were hospitalized with severe neurological symptoms.[51,52] Off-

target activity was quickly hypothesized as a possible cause of the observed toxicity. This hypothesis was investigated using activity-based protein profiling techniques, the results of which are disclosed in this chapter. BIA 10-2474, an important metabolite and three alkynylated derivatives are synthesized and extensively tested. It is shown that BIA 10-2474 is an a-specific and covalent inhibitor that severely disrupts neural lipid metabolism *in situ* through the inhibition of several serine hydrolase lipases.

Chapter 6 extends the chemical toolbox to study the endocannabinoid serine hydrolases. A new fluorophosphonate (FP) activity-based probe is synthesized and characterized. The profound influence on enzyme labeling efficiency due to the change in fluorophore is investigated in detail. The FP probe is found to label FAAH and MAGL at low-nanomolar concentrations, allowing it to be combined with MB064 to create an efficient probe cocktail capable of labeling most of the ECS-related serine hydrolases in one experiment. This cocktail is validated and used to profile the inhibition of two covalent MAGL inhibitors.

Chapter 7 attempts to take a prospective approach to target selectivity. By employing a relatively new machine learning algorithm, t-distributed Stochastic Neighbour Embedding (t-SNE), molecular similarity is shown for a large set of clinically relevant substances.[53] The same approach is able to visualize similarity in the protein sequences of the serine hydrolase superfamily, recapitulating phylogenetic information. A workflow is envisioned wherein bio-activity data spanning large amounts of compounds and targets are used to predict interaction profiles for serine hydrolase inhibitors.

Chapter 8 builds on the concepts explored in Chapter 7 for serine hydrolases and applies it to the more extensively studied kinase family. Using the Published Kinase Inhibitor Set, a model is trained and validated capable of predicting interaction profiles across the kinome.[54,55] The validated model is used to find new leads for the oncogene FLT3[56], which are validated in parallel using high-throughput methods. Two hits are resynthesized and profiled *in vitro* and *in situ* against acute myeloid leukaemia derived cells. The presented model is completely open source, and released as a readily usable executable.

Chapter 9 summarizes the work presented in this thesis, and shows future directions for the disclosed research.

References

1. Blass, B. E. Drug discovery and development. in *Basic principles of drug discovery and development* 1–27 (Academic Press, 2015).
2. Scannell, J. W., Blanckley, A., Boldon, H. & Warrington, B. Diagnosing the decline in pharmaceutical R&D efficiency. *Nat. Rev. Drug Discov.* **11**, 191–200 (2012).
3. Gaulton, A. *et al.* ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **40**, D1100–D1107 (2012).
4. Kenakin, T. P. Enzymes as Drug Targets. in *Pharmacology in Drug Discovery* 105–124 (Elsevier, 2012). doi:10.1016/B978-0-12-384856-7.00006-9
5. Hol, W. G. J. Protein Crystallography and Computer Graphics—toward Rational Drug Design. *Angew. Chemie Int. Ed. English* **25**, 767–778 (1986).
6. Lavecchia, A. & Di Giovanni, C. Virtual screening strategies in drug discovery: a critical review. *Curr. Med. Chem.* **20**, 2839–60 (2013).
7. Ezkurdia, I. *et al.* Multiple evidence strands suggest that there may be as few as 19 000 human protein-coding genes. *Hum. Mol. Genet.* **23**, 5866–5878 (2014).
8. Reddy, A. S. & Zhang, S. Polypharmacology: drug discovery for the future. *Expert Rev. Clin. Pharmacol.* **6**, 41–47 (2013).
9. Rodig, S. J. & Shapiro, G. I. Crizotinib, a small-molecule dual inhibitor of the c-Met and ALK receptor tyrosine kinases. *Curr. Opin. Investig. Drugs* **11**, 1477–90 (2010).
10. Bowes, J. *et al.* Reducing safety-related drug attrition: the use of in vitro pharmacological profiling. *Nat. Rev. Drug Discov.* **11**, 909–922 (2012).
11. Mechoulam, R. & Parker, L. A. The Endocannabinoid System and the Brain. *Annu. Rev. Psychol.* **64**, 21–47 (2013).
12. 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).
13. Donvito, G. *et al.* The Endogenous Cannabinoid System: A Budding Source of Targets for Treating Inflammatory and Neuropathic Pain. *Neuropsychopharmacology* **43**, 52–79 (2018).
14. Di Marzo, V. Targeting the endocannabinoid system: to enhance or reduce? *Nat. Rev. Drug Discov.* **7**, 438–55 (2008).
15. Devane, W. A. *et al.* Isolation and structure of a brain constituent that binds to the cannabinoid receptor. *Science* **258**, 1946–1949 (1992).
16. Mechoulam, R. *et al.* Identification of an endogenous 2-monoglyceride, present in canine gut, that binds to cannabinoid receptors. *Biochem. Pharmacol.* **50**, 83–90 (1995).
17. Gao, Y. *et al.* Loss of Retrograde Endocannabinoid Signaling and Reduced Adult Neurogenesis in Diacylglycerol Lipase Knock-out Mice. *J. Neurosci.* **30**, 2017–2024 (2010).
18. Tanimura, A. *et al.* The Endocannabinoid 2-Arachidonoylglycerol Produced by Diacylglycerol Lipase α Mediates Retrograde Suppression of Synaptic Transmission. *Neuron* **65**, 320–327 (2010).
19. Reisenberg, M., Singh, P. K., Williams, G. & Doherty, P. The diacylglycerol lipases: structure, regulation and roles in and beyond endocannabinoid signalling. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 3264–3275 (2012).
20. Bisogno, T. *et al.* Cloning of the first sn1-DAG lipases points to the spatial and temporal regulation of endocannabinoid signaling in the brain. *J. Cell Biol.* **163**, 463–468 (2003).
21. Murataeva, N., Straiker, A. & Mackie, K. Parsing the players: 2-arachidonoylglycerol synthesis and degradation in the CNS. *Br. J. Pharmacol.* **171**, 1379–1391 (2014).
22. Savinainen, J. R., Saario, S. M. & Laitinen, J. T. The serine hydrolases MAGL, ABHD6 and ABHD12 as guardians of 2-arachidonoylglycerol signalling through cannabinoid receptors. *Acta Physiol.* **204**, 267–276 (2012).
23. Long, J. Z. *et al.* Dual blockade of FAAH and MAGL identifies behavioral processes regulated by endocannabinoid crosstalk in vivo. *Proc. Natl. Acad. Sci.* **106**, 20270–20275 (2009).
24. Ogura, Y., Parsons, W. H., Kamat, S. S. & Cravatt, B. F. A calcium-dependent acyltransferase that produces N-acyl phosphatidylethanolamines. *Nat. Chem. Biol.* **12**, 669–671 (2016).
25. Simon, G. M. & Cravatt, B. F. Endocannabinoid biosynthesis proceeding through glycerophospho-N-acyl ethanolamine and a role for alpha/beta-hydrolase 4 in this pathway. *J. Biol. Chem.* **281**, 26465–72 (2006).
26. Liu, J. *et al.* Multiple pathways involved in the biosynthesis of anandamide. *Neuropharmacology* **54**, 1–7 (2008).
27. Cravatt, B. F. *et al.* Molecular characterization of an enzyme that degrades neuromodulatory fatty-acid amides. *Nature* **384**, 83–87 (1996).
28. Patricelli, M. P., Lovato, M. A. & Cravatt, B. F. Chemical and mutagenic investigations of fatty acid amide hydrolase: Evidence for a family of serine hydrolases with distinct catalytic properties. *Biochemistry* **38**, 9804–9812 (1999).

29. Huggins, J. P., Smart, T. S., Langman, S., Taylor, L. & Young, T. An efficient randomised, placebo-controlled clinical trial with the irreversible fatty acid amide hydrolase-1 inhibitor PF-04457845, which modulates endocannabinoids but fails to induce effective analgesia in patients with pain due to osteoarthritis of th. *Pain* **153**, 1837–1846 (2012).

30. Fraser, I. *et al.* Preclinical characterization and first-in-human administration of a selective monoacylglycerol lipase inhibitor, ABX-1431. in *Front. Pharmacol. Conference Abstract: EUFEMED 2017* (2017). doi:10.3389/conf.fphar.2017.62.00011

31. Cisar, J. S. *et al.* Identification of ABX-1431, a Selective Inhibitor of Monoacylglycerol Lipase and Clinical Candidate for Treatment of Neurological Disorders. *J. Med. Chem.* **61**, 9062–9084 (2018).

32. Li, G. L. *et al.* Assessment of the pharmacology and tolerability of PF-04457845, an irreversible inhibitor of fatty acid amide hydrolase-1, in healthy subjects. *Br. J. Clin. Pharmacol.* **73**, 706–716 (2012).

33. Postnov, A. *et al.* Fatty Acid Amide Hydrolase Inhibition by JNJ-42165279: A Multiple-Ascending Dose and a Positron Emission Tomography Study in Healthy Volunteers. *Clin. Transl. Sci.* **11**, 397–404 (2018).

34. Kiss, L. E. *et al.* Discovery of a Potent, Long-Acting, and CNS-Active Inhibitor (BIA 10-2474) of Fatty Acid Amide Hydrolase. *ChemMedChem* **13**, 2177–2188 (2018).

35. Janssen, F. J. & van der Stelt, M. Inhibitors of diacylglycerol lipases in neurodegenerative and metabolic disorders. *Bioorg. Med. Chem. Lett.* **26**, 3831–3837 (2016).

36. Baggelaar, M. P., Maccarrone, M. & van der Stelt, M. 2-Arachidonoylglycerol: A signaling lipid with manifold actions in the brain. *Prog. Lipid Res.* **71**, 1–17 (2018).

37. Long, J. Z. & Cravatt, B. F. The Metabolic Serine Hydrolases and Their Functions in Mammalian Physiology and Disease. *Chem. Rev.* **111**, 6022–6063 (2011).

38. Kidd, D., Liu, Y. & Cravatt, B. F. Profiling serine hydrolase activities in complex proteomes. *Biochemistry* **40**, 4005–4015 (2001).

39. Baggelaar, M. P. *et al.* Highly Selective, Reversible Inhibitor Identified by Comparative Chemoproteomics Modulates Diacylglycerol Lipase Activity in Neurons. *J. Am. Chem. Soc.* **137**, 8851–8857 (2015).

40. 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).

41. Baggelaar, M. P. *et al.* Chemical Proteomics Maps Brain Region Specific Activity of Endocannabinoid Hydrolases. *ACS Chem. Biol.* **12**, 852–861 (2017).

42. 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).

43. 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).

44. Johnson, D. S. *et al.* Discovery of PF-04457845: A Highly Potent, Orally Bioavailable, and Selective Urea FAAH Inhibitor. *ACS Med. Chem. Lett.* **2**, 91–96 (2011).

45. Adibekian, A. *et al. Optimization and characterization of a triazole urea inhibitor for alpha/beta hydrolase domain-containing protein 11 (ABHD11): anti-probe for LYPLA1/LYPLA2 dual inhibitor ML211. Probe Reports from the NIH Molecular Libraries Program* (National Center for Biotechnology Information (US), 2010).

46. Liu, Y., Patricelli, M. P. & Cravatt, B. F. Activity-based protein profiling: The serine hydrolases. *Proc. Natl. Acad. Sci.* **96**, 14694–14699 (1999).

47. Patricelli, M. P., Giang, D. K., Stamp, L. M. & Burbaum, J. J. Direct visualization of serine hydrolase activities in complex proteomes using fluorescent active site-directed probes. *Proteomics* **1**, 1067–1071 (2001).

48. Jessani, N. *et al.* A streamlined platform for high-content functional proteomics of primary human specimens. *Nat. Methods* **2**, 691–697 (2005).

49. Janssen, F. J. Discovery of sulfonyl-1,2,4-triazole ureas as DAGLα inhibitors by HTS-ABPP. in *Discovery of novel inhibitors to investigate diacylglycerol lipases and α/β-hydrolase domain 16A* 109–139 (2016).

50. Deng, H. *et al.* Triazole Ureas Act as Diacylglycerol Lipase Inhibitors and Prevent Fasting-Induced Refeeding. *J. Med. Chem.* **60**, 428–440 (2017).

51. Butler, D. & Callaway, E. Scientists in the dark after French clinical trial proves fatal. *Nature* **529**, 263–264 (2016).

52. Kerbrat, A. *et al.* Acute Neurologic Disorder from an Inhibitor of Fatty Acid Amide Hydrolase. *N. Engl. J. Med.* **375**, 1717–1725 (2016).

53. Van Der Maaten, L. & Hinton, G. Visualizing Data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).

54. Drewry, D. H., Willson, T. M. & Zuercher, W. J. Seeding collaborations to advance kinase science with the GSK Published Kinase Inhibitor Set (PKIS). *Curr. Top. Med. Chem.* **14**, 340–2 (2014).

55. Elkins, J. M. *et al.* Comprehensive characterization of the Published Kinase Inhibitor Set. *Nat. Biotechnol.* **34**, 95–103 (2015).

56. Larrosa-Garcia, M. & Baer, M. R. FLT3 Inhibitors in Acute Myeloid Leukemia: Current Status and Future Directions. *Mol. Cancer Ther.* **16**, 991–1001 (2017).

Hit-to-lead optimization of triazole sulfonamide DAGL-α inhibitors

Introduction

Diacylglycerol lipase (DAGL)-α and DAGL-β inhibitors have been under active investigation as potential drugs for the treatment of metabolic and neurodegenerative disorders for nearly a decade.[1,2] The recent generations of optimized inhibitors can be divided in three chemotypes: the α-ketoheterocycles (LEI104 and LEI105[3,4]), the glycine sulfonamides (*e.g.* LEI106)[5,6] and the triazole ureas (KT109, DO34 and DH376[7,8]). These inhibitors have been used successfully in a variety of cellular and animal models to study the effect of DAGL inhibition in detail.[1,4,7–9] However, as potential drugs, each of these

inhibitor classes has its own set of shortcomings. The α-ketoheterocycles are the most selective inhibitors published to date, with no known off-targets reported for LEI105.[4] However, LEI105 has a very low oral bioavailability due to its low solubility and high metabolic clearance (unpublished data). The glycine sulfonamides are dual DAGL and α/β-hydrolase domain containing protein 6 (ABHD6) inhibitors, and have a few further unidentified off-targets *in vitro*.[5] The free acid of the glycine moiety was found to be crucial for activity, but likely caused the observed significantly reduced cellular potency and low brain penetration.[9] Although the triazole urea derivatives DO34 and DH376 are less lipophilic and more selective than the original hit KT109[1], they still have several other serine hydrolases as off-targets. In all, the current state-of-the-art for DAGL inhibitors leaves room for improvement and warrants the search for new chemical scaffolds.

To find new hits for DAGL-α, a high throughput screening campaign was performed in collaboration with the Pivot Park Screening Centre and the European Lead Factory.[10] Initial screening at 10 μM in a previously reported surrogate substrate assay yielded 1932 hits (Figure 2.1).[3] These were confirmed at a lower concentration (1.25 μM) and verified in an orthogonal activity-based protein profiling (ABPP) assay. After final selections based on purity and legal restrictions, 46 qualified hits were disclosed. This set of compounds yielded a series of 1,2,4-triazole-3-sulfon(amide) urea inhibitors (**1-3**, Figure 2.1) that was selected for further optimization, because of its promising physicochemical properties (**1**: CLogP = 2.07, tPSA = 85.6, MW = 337). Here, the results of the structure activity relationship (SAR) investigations are described, leading to the most polar *in vivo* active triazole urea DAGL inhibitor reported to date.
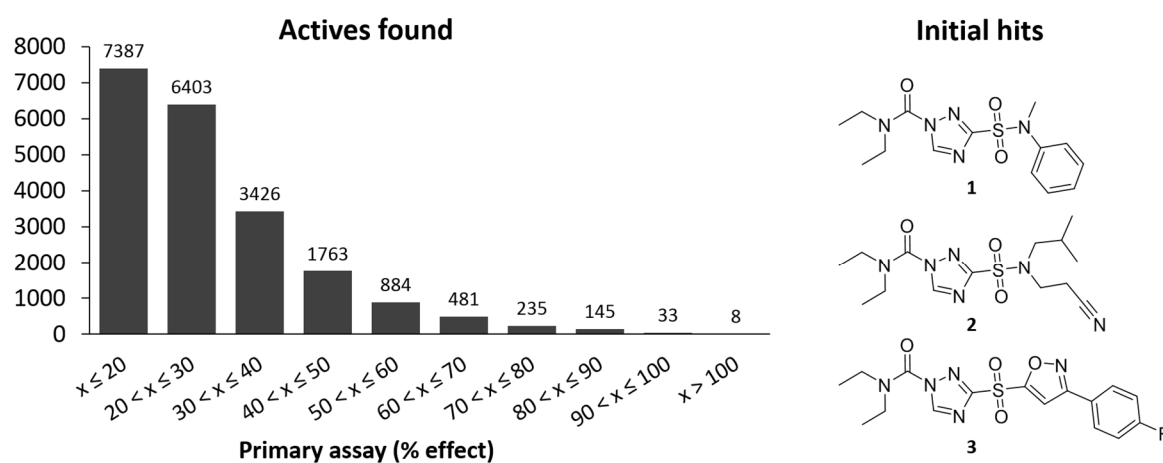
	Compounds		% Total
Total screened	302.655	-	100
Primary assay hits	1.932	> 50% eff. (10 μM)	0.64
Active confirmation hits	263	> 70% eff. (10 μM) > 50% eff. (1.25 μM)	0.09
Triaging and orthogonal assay	46	Chemical eye, clustering, legal, purity	0.02

Figure 2.1 | Summary of the high throughput screening campaign for DAGL-α. The triazole sulfonamide series was selected from the initial hits to be further optimized.

Results and discussion

To investigate the SAR of the triazole-sulfonamide-ureas a library of 65 compounds was synthesized at the European Lead Factory and screened using a previously reported colorimetric DAGL-α surrogate substrate assay.[3] The SAR is described in a topological fashion below. The initial variations focussed on finding the minimal requirements of the scaffold to inhibit DAGL-α (Table 2.1). The desired polar core of the scaffold was maintained, while varying the substituents on the sulfonamide (**1-8**) and on the urea moiety (**9-25**). Removal of one (**4**) or both (**5**) sulfonamide substituents significantly reduced DAGL-α inhibitory activity. Piperidine- (**6**) and morpholine- (**7**) substituted sulfonamides lost potency, while an inhibitor substituted with two isobutyl groups (**8**) was tolerated. This suggested that conformational flexibility and/or a lipophilic substituent on the sulfonamide is beneficial for the inhibitory activity.

Changing the diethyl substitution at the urea to a monomethyl (**9**), monoethyl (**10**) or di-methoxyethyl (**11**) group abolished activity against DAGL-α, but the dimethylamine (**12**) was active. Cyclization of the diethylamine substituents to pyrrolidine (**13**) or piperidine (**14**), but not morpholine (**15**) or 4-methylpiperazine (**16**), was tolerated. Even when decorated with a benzyl, mono-substituted amines failed to inhibit DAGL-α (**17** and **18**). Decorating the cyclic amines with a 3-phenyl or 2-benzyl substituent (**19-21**), but not a 2-phenyl (**22** and **23**), greatly improved binding to DAGL-α yielding a single digit nanomolar potent inhibitor (**21**). This is in line with the observations made by Hsu *et al.* in the optimization towards KT109[11] and suggests that 2-benzyl piperidine is a preferred motif for DAGL and the binding mode of this 3-substituted 1,2,4-triazole urea scaffold is comparable to that of the 4-substituted 1,2,3-triazole ureas. The loss of activity for compounds **24** and **25** further illustrates the lipophilic nature of the binding pocket.

Next the sulfonamide substitution pattern was further explored using the 3-phenyl-pyrrolidine, the 2-benzylpyrrolidine and the 2-benzylpiperidine as substituents on the amine urea for matched-molecular pair analysis (Table 2.2, **26-51**). In general, it was found that the 2-benzylpyrrolidine series (**19, 26-33**) was the least active, while the 2-benzylpiperidine series (**21, 45-51**) delivered the most active inhibitors. The SAR trend within each series was generally consistent across the three chemotypes. It was found that the increase in the size of the cycloalkyl substituent on the sulfonamide provided more active compounds. Virtually all methyl substituted variants (**19-21, 26-28, 34-37** and **45-46**) were more potent than their hydrogen counterparts (**29-33, 38-44** and **47-51,** respectively). None of the sulfonamide variants, however, provided an improvement over the originally found methyl aniline in **19-21**.

Finally, to investigate the necessity of the sulfonamide moiety a small selection of sulfide, sulfoxide and sulfone variants (**52-65**, Table 2.3) was tested. The sulfonamide was not required for potent inhibitory activity (e.g. **58, 60, 64** and **65**) and oxidation of the sulfur significantly increased the potency of the inhibitors (e.g. **52, 53** and **54** or **55** vs. **56, 57** vs. **58, 59** vs. **60** and **61** vs. **62**).

Table 2.1 | Initial variations of the sulfonamide substitution and urea bound amine groups. pIC$_{50}$s were determined by a surrogate substrate assay using membrane fractions of hDAGL-α overexpressing HEK293T cells, and are reported as mean of N=2, n=2 experiments.
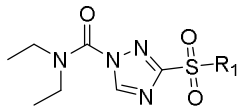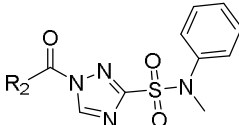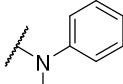

#	R$_1$	pIC$_{50}$ DAGL-α
1		5.9
2		5.5
4		5.5
5	NH$_2$	< 5
6		5.2
7		< 5
8		5.9
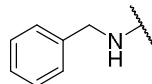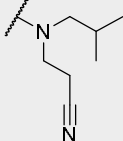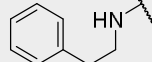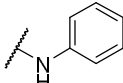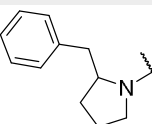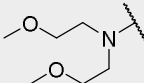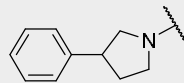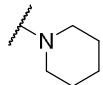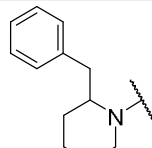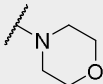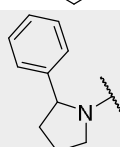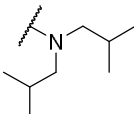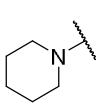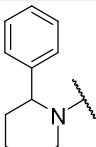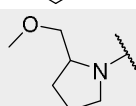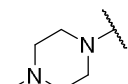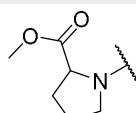
#	R$_2$	pIC$_{50}$ DAGL-α	#	R$_2$	pIC$_{50}$ DAGL-α
1		5.9	17		< 5
9		< 5	18		< 5
10		< 5	19		7.5
11		5.8	20		7.5
12		6.2	21		8.6
13		6.2	22		5.4
14		6.1	23		6.2
15		5.0	24		5.8
16		< 5	25		< 5

Table 2.2 | Sulfonamide variations for the three most promising retaining groups. pIC$_{50}$s were determined by a surrogate substrate assay using membrane fractions of hDAGL-α overexpressing HEK293T cells, and are reported as mean of N=2, n=2 experiments.
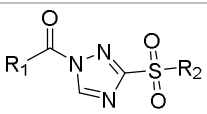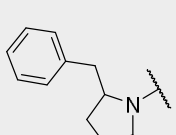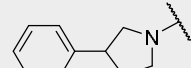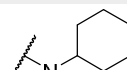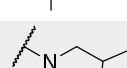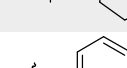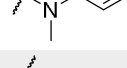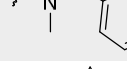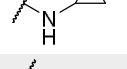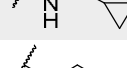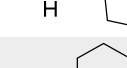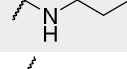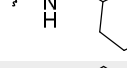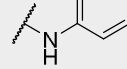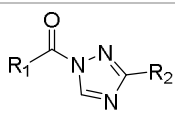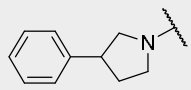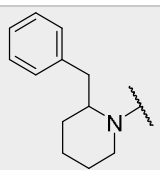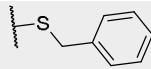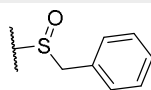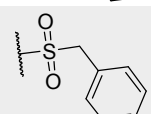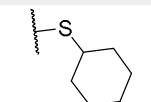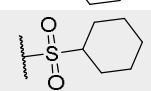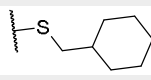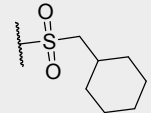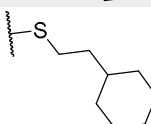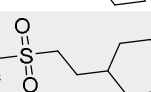
R$_2$	#	pIC$_{50}$ DAGL-α	#	pIC$_{50}$ DAGL-α	#	pIC$_{50}$ DAGL-α
N-methyl-N-cyclopropyl	26	6.0	34	6.9		-
N-methyl-N-cyclohexyl	27	5.9	35	7.9	45	7.6
N-methyl-N-(cyclohexylmethyl)	28	7.0	36	8.0	46	8.1
N-methyl-N-phenyl	19	7.5	20	7.5	21	8.6
N-methyl-N-benzyl		-	37	7.2		-
NH-cyclopropyl		-	38	6.5		-
NH-(cyclopropylmethyl)	29	6.0	39	7.1	47	7.2
NH-(cyclopentylmethyl)	30	6.7	40	7.4	48	7.8
NH-cyclohexyl	31	6.9	41	7.6	49	7.6
NH-(cyclohexylmethyl)	32	6.8	42	7.6	50	7.5
NH-phenyl	33	6.9	43	6.9	51	7.5
NH-benzyl		-	44	7.2		-

Table 2.3 | Sulfide, sulfone and sulfoxide variations for the two most promising urea substitutions. pIC_{50}s were determined by a surrogate substrate assay using membrane fractions of hDAGL-α overexpressing HEK293T cells, and are reported as mean of N=2, n=2 experiments.

R₁				
R₂	#	pIC_{50} DAGL-α	#	pIC_{50} DAGL-α
	52	6.1	61	6.6
	53	6.5		-
	54	7.7	62	7.8
	55	5.8		-
	56	7.3	63	7.3
	57	6.0		-
	58	8.0	64	8.0
	59	5.8		-
	60	8.0	65	8.0

Figure 2.2 | Correlation analysis of calculated lipophilicity (CLogP) and measured pIC_{50} for DAGL-α. Grey markers denote sulfide compounds **52**, **55**, **57**, **59** and **61**. Regression line with 99% confidence interval is shown.

For all compounds the CLogP was calculated (Chemdraw 16.0) and plotted against the potency (pIC_{50}) for a correlation analysis (Figure 2.2). Although in general lipophilic groups tended to increase potency, only a poor correlation with an R^2 of 0.25 was found. This indicates that the potency of the inhibitors might not only be dependent on an increase in entropy due to the de-solvation effect, but may also reflect enthalpic interactions and steric clashes with the enzyme. The most notable example is the sulfide series (**52**, **55**, **57**, **59**, **61**; grey markers in Figure 2.2), which indicated that the sulfone-oxygen atoms present in the sulfonamide and sulfone variants and not present in the sulfides have a positive contribution to the overall potency, potentially due to an increased electron withdrawing character. Compound **21** truly stands out in this plot, showcasing its large lipophilic efficiency (LipE = 4.4), which is only surpassed by a few more promiscuous molecules like **12** (Figure 2.3).

Off-target profiling

Next all compounds were screened for off-targets using activity-based protein profiling (ABPP). Two broad spectrum activity-based probes (ABPs) were used: the commercially available FP-TAMRA (ActiveX®) and the tailored ABP MB064, which together label most of the endocannabinoid serine hydrolases. This approach uniquely enables the medicinal chemistry program to focus not only on potency in the primary assay, but also to be steered towards the more selective compounds. Initial screening was performed at 10 μM inhibitor concentration, but for the more active compounds (pIC_{50} > 7.5) the screening was repeated at 100 nM. A typical screening result is displayed in Figure S2.1. A direct comparison of the most promising compounds is shown in Figure 2.3. The ABPP-data indicate that the 3-phenylpyrrolidine series is more promiscuous than the 2-benzylpyrrolidines and especially the 2-benzylpiperidines. The latter series shows a 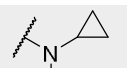dramatic reduction in inhibitory activity against several enzymes, among which ABHD16A, ABHD12 and FAAH. In all, compound **21** was selected for further biological profiling, as it displayed the highest overall activity (pIC_{50} = 8.6), large lipophilic efficiency (LipE = 4.4) and relatively good *in vitro* selectivity.
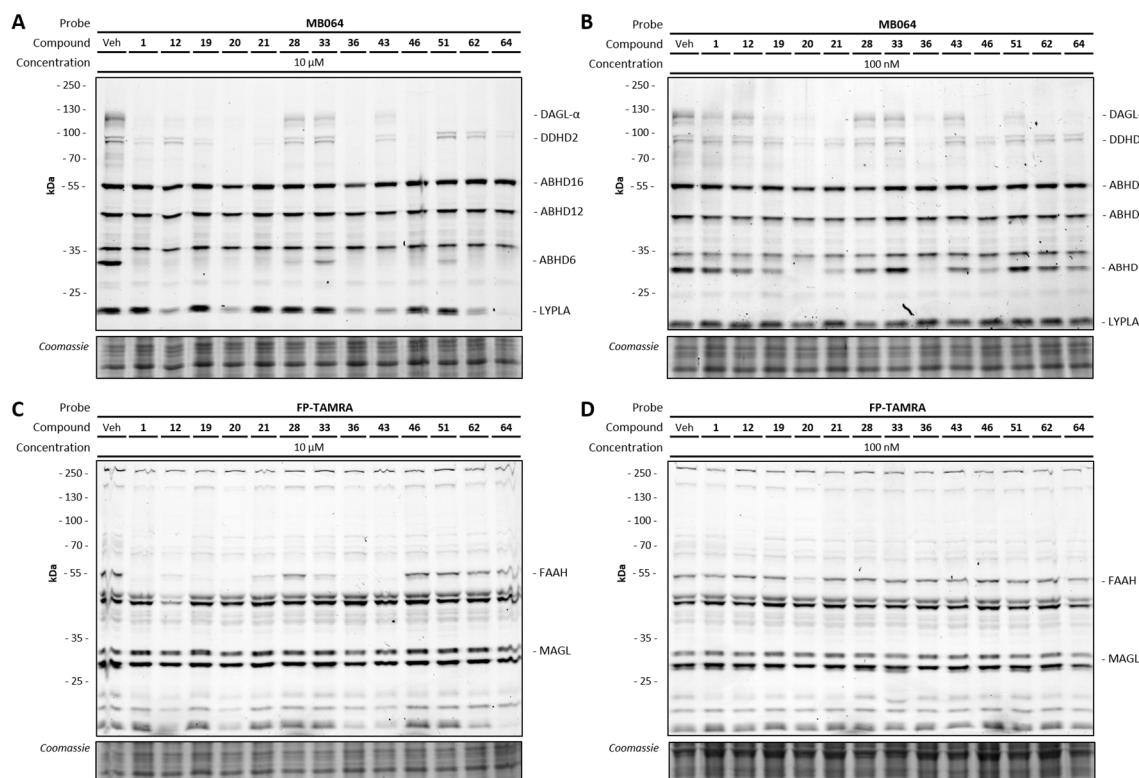
Figure 2.3 | Inhibition profiling of a selection of potent inhibitors in Table 2.1, 2.2 and 2.3 in *in vitro* treated mouse brain membrane proteome. Profiling was performed at compound concentrations of 10 μM (A, C) and 100 nM (B, D), using MB064 (A, B) and FP-TAMRA (C, D) for protein labeling.

Resynthesis of **21** and its activity-based probe (**77**)

Resynthesis of **21** was performed to enable further biological studies and was performed analogous to the synthetic methods used for the library synthesis (Scheme 2.1). In brief, 2-benzylpyridine (**66**) was hydrogenated using platinum oxide to yield the 2-benzyl-piperidine (**67**). In parallel, 1,2,4-triazole-3-thiol (**68**) was benzylated using benzylbromide. These were coupled in a triphosgene mediated coupling reaction to yield the triazole urea **70**. In a one-pot reaction the thioether was oxidized by *in situ* generated Cl_2 (from HCl and NaOCl) to the sulfonyl chloride, which was trapped by the addition of a large excess of *N*-methylaniline. This furnished **21** in moderate overall yield.

To use the covalent mode of action of these inhibitors to visualize target engagement, a two-step activity-based probe was designed (**77**, Scheme 2.1). This was done by fusing the previously described 5-propargylether substituted (S)-2-benzylpiperidine (**76**) of DH376 to the 1,2,4-triazole sulfonamide leaving group (**74**).[8] To this end, a new synthetic procedure was developed wherein the triphosgene coupling featured as the last step, to avoid putting the valuable chiral 2-benzyl piperidine through the harsh and low yielding chlorination step. The benzylated 1,2,4-triazole-3-thiol (**69**) was coupled to commercially available diphenylcarbamic chloride (**71**), before being reacted with chlorine and *N*-methylaniline. The urea moiety was then hydrolyzed using strong base and the resulting 1,2,4-triazole sulfonamide was coupled to the deprotected 5-substituted 2-benzylpiperidine (**76**), yielding **77**.
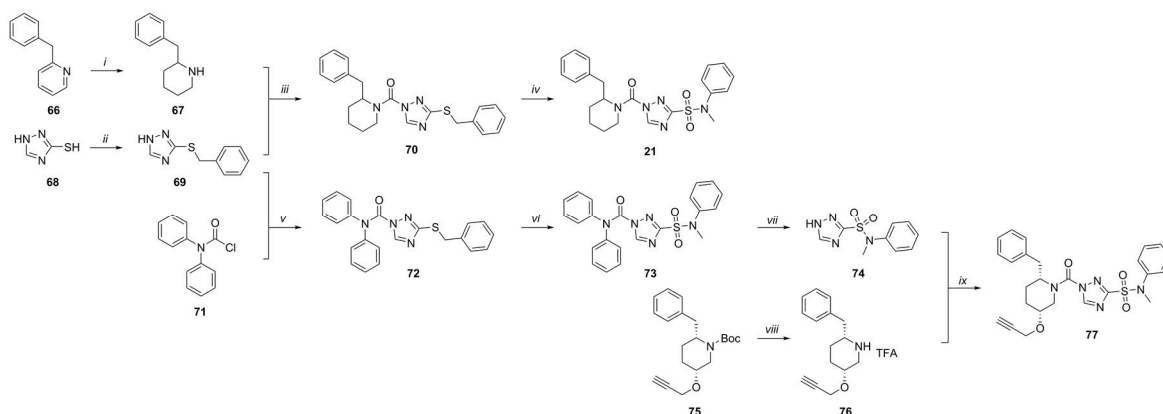
Scheme 2.1 | Synthesis of **21** and **77**. Synthesis of **75** was described in Deng *et al.*[8] Reagents and conditions: *i*) H_2, HCl, PtO_2, EtOH, 2 bar, 18 h, 64%; *ii*) benzylbromide, DMF, 18 h, 98%; *iii*) **67**, triphosgene, TEA, DCM, -5 °C → RT, 2 h, then **69**, K_2CO_3, DMF, 18 h, 32%; *iv*) HCl, NaOCl, DCM, -10 °C, 30 min, then *N*-methylaniline, -10 °C → RT, 2h, 52%; *v*) **69**, **71**, DIPEA, DMAP, THF, 65 °C, 5 h, 88%; *vi*) HCl, NaOCl, DCM, -10 °C, 20 min, then *N*-methylaniline, -10 °C → RT, 2h, 69%; *vii*) KOH, H_2O/THF (1:1), 18 h, 53%; *viii*) TFA, DCM, 5 h; *ix*) **76**, triphosgene, DIPEA, DCM, 0 °C → RT, 1 h, then **74**, DMAP, DIPEA, THF, 65 °C, 5 h, 67%.

Full *in vitro* and *in situ* profiling of **21** and **77**

Next, two orthogonal assays were performed to verify the activity of **21** and **77**. First, **21** and **77** were profiled in a competitive ABPP setting in a dose response fashion. Second, the compounds were tested in a previously developed natural substrate assay.[12] The resulting inhibition data are displayed in Table 2.4. Both compounds showed a slight reduction of potency in the competitive ABPP assay, compared to the surrogate substrate (PNP) assay, which could be due to a species difference as the surrogate substrate assay utilized human DAGL-α, whereas the ABPP assay was performed on mouse brain proteome. Both compounds showed a moderate selectivity *in vitro* over the major off-target ABHD6 of approximately 100-fold. In the natural substrate assay the activity of (human) DAGL-α was less potently inhibited, in line with previously reported observations.

The cellular activity of both compounds was tested in mouse Neuro-2a cells. Cells were treated with the compounds for one hour before the cells were harvested and analyzed by gel-based ABPP. The resulting gels (Figure 2.4) and quantification (Table 2.4) showed a similar potency for mouse DAGL-β *in situ* as for DAGL-α in the natural substrate assay, with a pIC_{50} of 6.5 and 7.2 for **21** and **77**, respectively. The potency for ABHD6 is remarkably increased to the single-digit nanomolar range, which is evidence of efficient cellular penetration of the compounds.

Table 2.4 | Activity data for **21** and **77** in the used biochemical and ABPP assays. Values are reported as mean ± standard deviation of N=3 (for ABPP) or N=2, n=2 (for PNP- and SAG-assay) replicates.

Compound	pIC_{50} DAGL-α PNP-assay	pIC_{50} DAGL-α ABPP-assay	pIC_{50} ABHD6 ABPP-assay	pIC_{50} DAGL-α SAG-assay	pIC_{50} DAGL-β *in situ*-assay
21	8.6 ± 0.3	8.1 ± 0.1	6.1 ± 0.1	6.4 ± 0.2	6.5 ± 0.1
77	8.6 ± 0.2	8.5 ± 0.1	6.8 ± 0.1	7.4 ± 0.1	7.2 ± 0.1
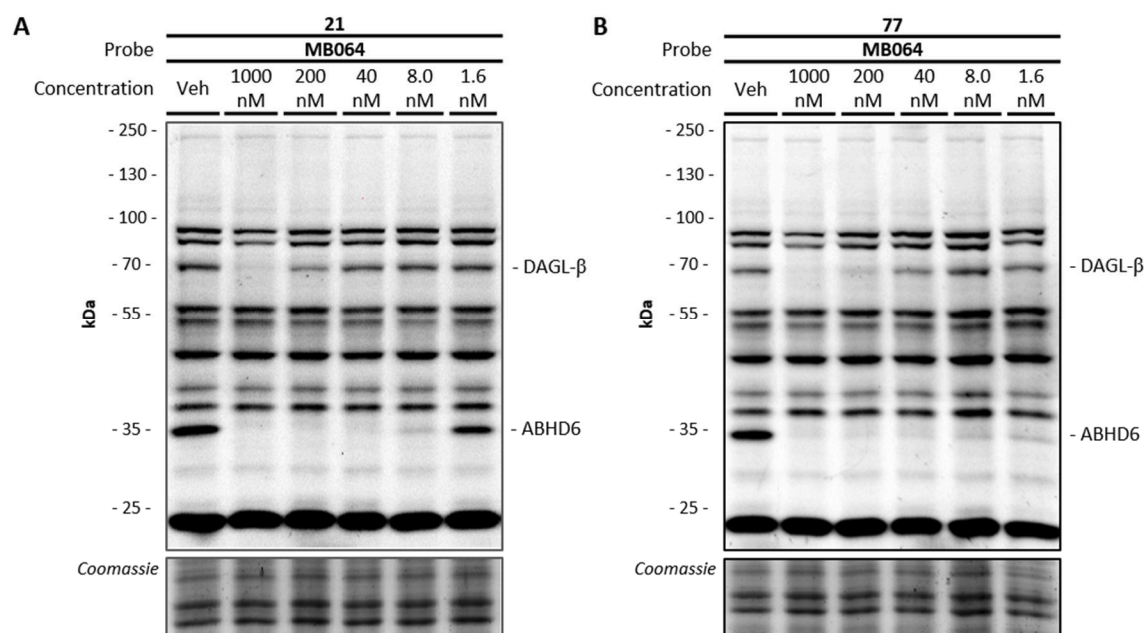
Figure 2.4 | Neuro-2a cells treated *in situ* with **21** (A) and **77** (B), labeled post-lysis with MB064 (2 µM, 20 min, RT). Coomassie was used for protein loading control.

21 and 77 are *in vivo* active

Considering the favourable physicochemical properties, *in vitro* and *in situ* potency and relatively high selectivity, both compounds were tested in mice to assess their *in vivo* target engagement. To this end, the compounds were administered intraperitoneal, using a high dose of 50 mg/kg of compound or vehicle (18:1:1 (v/v/v) saline:ethanol:PEG-40). The mice were sacrificed after two hours and the organs were harvested. Both the brain and the spleen were selected to assess target engagement of DAGL-α/β and ABHD6 in the central nervous system and peripheral organs. To visualize residual protein activity, the membrane proteomes were treated with both MB064 and FP-TAMRA (Figure 2.5). Compounds **21** and **77** abolished or reduced the fluorescent labeling of DAGL-α and ABHD6. This indicated that the compounds are *in vivo* active and are able to cross the blood-brain barrier. **77** was less active than **21**, as judged by the only partial inhibition of DAGL-α (Figure 2.5A). Both compounds seem to be quite selective, but **21** at this high dose also inhibited FAAH (band at 60 kDa labeled by FP-TAMRA). In the spleen the number of off-targets observed was somewhat higher. The strong band disappearing at 60 kDa is likely a group of carboxylic esterases, which are well-known to bind these types of inhibitors.[7,13] The labeling of DAGL-β in the spleen was faint, but inhibited by **21**.
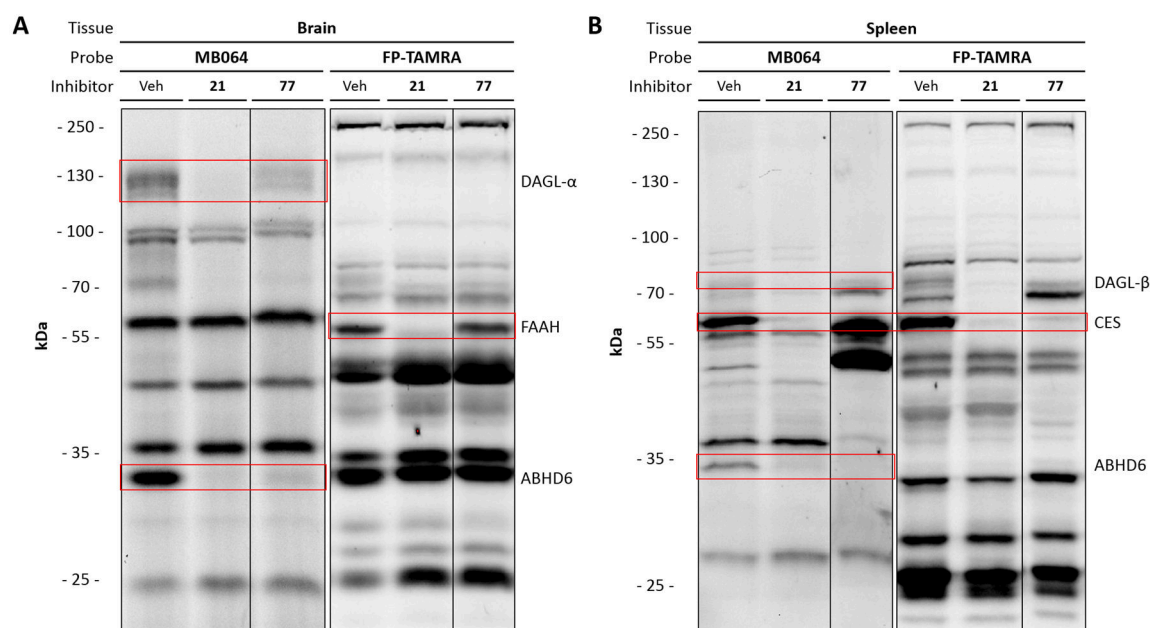
Figure 2.5 | Membrane fractions of mouse brain (A) and spleen (B) treated *in vivo* i.p. with 50 mg/kg of **21** and **77**, post-lysis labeled with MB064 or FP-TAMRA. Red boxes denote assigned bands (labels on the right of the gels).

Conclusion

All together, the discovery and optimization of a new series of DAGL inhibitors, based on a 1,2,4-triazole urea sulfonamide scaffold was reported in this Chapter. The optimized inhibitors had favourable physicochemical properties leading to a high lipophilic efficiency. Activity-based protein profiling was used to guide the optimization of potency and selectivity in parallel on endogenously expressed enzymes in brain lysates. Compound **21** proved cellular active and demonstrated *in vivo* activity in mice. Further dose-response studies have to be performed to establish the minimum dose required for complete DAGL-α inhibition in the brain.

Acknowledgements

Methods

Chemical Biology Methods

Cell Culture

HEK293T (human embryonic kidney) and Neuro-2a (mouse neuroblastoma) cells were cultured at 37 °C under 7% CO_2 in DMEM containing phenol red, stable glutamine, 10% (v/v) New Born Calf Serum (Thermo Fisher), and penicillin and streptomycin (200 µg/mL each; Duchefa). Medium was refreshed every 2-3 days and cells were passaged twice a week at 80-90% confluence by resuspension in fresh medium.
Cells lines were purchased from ATCC and were regularly tested for mycoplasma contamination. Cultures were discarded after 2-3 months of use.

Transient transfection

One day prior to transfection HEK293T cells were seeded to 15-cm dishes (~62.500 cells/cm²). Prior to transfection, culture medium was aspirated and a minimal amount of medium was added. A 3:1 (m/m) mixture of polyethyleneimine (PEI) (60 µg/dish) and human DAGL-α plasmid DNA (20 µg/dish) was prepared in serum-free culture medium and incubated for 15 min at RT. Transfection was performed by dropwise addition of the PEI/DNA mixture to the cells. Transfection with the empty pcDNA3.1 vector was used to generate control samples. After 24 h, medium was refreshed. Medium was aspirated 48 or 72 h post-transfection and cells were harvested by resuspension in PBS. Cells were pelleted by centrifugation (5 min, 1,000 g) and the pellet was washed with PBS. Supernatant was discarded and cell pellets were frozen in liquid nitrogen and stored at -80 °C until sample preparation.

In situ treatment of Neuro-2a cells

Cells were seeded in 6-well plates. *In situ* treatment was initiated 24 h later. Medium was aspirated and serum-free medium containing inhibitor or DMSO as vehicle was added (1.0 % v/v DMSO). After 1 h exposure to treatment medium, the medium was aspirated and cells were harvested and stored as described above until sample preparation.

Whole cell lysate

Cell pellets were thawed on ice, resuspended in cold lysis buffer (20 mM HEPES pH 7.2, 2 mM DTT, 250 mM sucrose, 1 mM $MgCl_2$, 2.5 U/mL benzonase) and incubated on ice (15-30 min). The cell lysate was used for membrane preparation (below) or was diluted to 2.0 mg/mL concentration in cold storage buffer (20 mM Hepes, pH 7.2, 2 mM DTT) for use as whole lysate. Protein concentrations were determined by a Quick Start™ Bradford Protein Assay and diluted samples were flash frozen in liquid nitrogen and stored at -80 °C until further use.

Membrane preparation from overexpression lysate

The membrane and cytosolic fractions of cell or tissue lysates were separated by ultracentrifugation (93,000 g, 45 min, 4 °C). The supernatant was collected (cytosolic fraction) and the membrane pellet was resuspended in cold storage buffer by thorough pipetting and passage through an insulin needle. Protein concentrations were determined by a Quick Start™ Bradford Protein Assay and samples were diluted to 2.0 mg/mL with cold storage buffer, flash frozen in liquid nitrogen and stored at -80 °C until further use.

Tissue preparation

Organs were isolated from C57BL/6 mice following standard guidelines as approved by the ethical committee of Leiden University (DEC#13191). Isolated organs were frozen in liquid nitrogen and stored at -80 °C. Organs were thawed on ice and homogenized by a glass stick douncing homogenizer in cold lysis buffer (20 mM HEPES, pH 7.2, 2 mM DTT, 1 mM $MgCl_2$, 2.5 U/mL benzonase). The resulting suspension was centrifuged at 1000 g for 5 minutes at 4 °C to get rid of residual solid tissue. The supernatant was centrifuged 45 minutes at 93,000 g at 4 °C to separate soluble (cytosol) and insoluble (membrane) fractions. The pellet (membrane) was resuspended in storage buffer (20 mM HEPES, pH 7.2, 2 mM DTT) using an insulin syringe. The protein concentration was measured using a Qubit™ protein assay and was adjusted to 2 mg/mL using storage buffer. The resulting lysates were frozen in liquid nitrogen and stored at -80 °C for later use.

Natural substrate-based fluorescence assay DAGL-α

The diacylglycerol substrate assay was performed as reported previously.[12] Standard assay conditions: 0.2 U/mL glycerol kinase (GK), glycerol-3-phosphate oxidase (GPO) and horseradish peroxidase (HRP), 0.125 mM ATP, 10 µM Amplifu™Red, 5% DMSO in a total volume of 200 µL. The assay additionally contained 5 µg/mL MAGL-overexpressing membranes, 100 µM SAG and 0.0075% (w/v) Triton X-100, with a final protein concentration of 50 µg/mL.

Surrogate substrate assay

The biochemical DAGL-α activity assay is based on the method previously described.[3] 200 µL reactions were performed in flat bottom Greiner 96-wells plates in a 50 mM pH 7.2 Hepes buffer. Membrane protein fractions from HEK293T cells transiently transfected with hDAGL-α (0.05 µg/µL final concentration) were used as hDAGL-α source. Inhibitors were introduced in 5.0 µL DMSO. The mixtures were incubated for 20 minutes before 10.0 µL 6 mM PNP-butyrate (final concentration 0.3 mM) in 50% DMSO was added (final DMSO concentration 5.0%). Reactions were allowed to progress for 30 minutes at 20 °C before OD (420 nm) was measured using a TECAN GENios plate reader. All experiments were performed at N=2, n=2 for experimental measurements and N=2, n=4 for controls.

Z'-factor of each plate was determined for the validation of each experiment, using the following formula: $Z' = 1-3(\sigma_{pc}+ \sigma_{nc})/(\mu_{pc} - \mu_{nc})$. The OD from the positive control (pc: DAGL DMSO), and the negative control (nc: 10 µM THL) was used. Plates were accepted for further analysis when Z' > 0.6. Measurements were corrected for the average absorption of the negative control. The average, standard deviation (SD) and standard error of mean (SEM) were calculated and normalized to the corrected positive control. Data was exported to Graphpad Prism 7.0 for the calculation of the pIC_{50} using a non-linear dose-response analysis with variable slope.

Activity-based protein profiling

For *in vitro* inhibition, lysates (19 µL per sample, 2 µg/µL) were thawed on ice. 0.5 µL of the inhibitor (40x stock in DMSO) or pure DMSO (as vehicle) was added to the sample, vortexed briefly and incubated for 20 minutes at RT. Subsequently, 0.5 µL probe (40x stock in DMSO, final concentration 250 nM for MB064, 500 nM for FP-TAMRA) was added to the proteome sample, vortexed briefly and incubated for 15 minutes at RT. For *in situ* inhibition, the *in situ*-treated cells (19.5 µL whole lysate) were directly incubated with the activity-based probe (40x stock in DMSO, final concentration 2 µM for MB064, 500 nM for FP-TAMRA, 20 min, RT). For *in vivo* inhibition the procedure is identical to *in vitro* inhibition, but the inhibitor incubation is skipped. Final volume in all cases was 20 µL (max. 5% DMSO). The reaction was quenched by the addition of 7.5 µL of 4*Laemmli-buffer (final concentrations: 60 mM Tris (pH 6.8), 2% (w/v) SDS, 10% (v/v) glycerol, 1.25% (v/v) β-mercaptoethanol, 0.01% (v/v) bromophenol blue). 10 µL (14 µg protein) of quenched reaction mixture was resolved on 10% acrylamide SDS-PAGE gels (180 V, 75 min). Fluorescence was measured using a Biorad ChemiDoc MP system (fluorescence channels Cy2 (460-490 nm), Cy3 (520-545 nm), Cy5 (625-650 nm) filters). Gels were then stained using coomassie staining and imaged for protein loading control.

Labeling quantification

Fluorescence quantification was performed using Imagelab 6.0 (Biorad). Intensities were normalized to the DMSO control and corrected for protein loading by coomassie staining. pIC_{50} values were calculated using GraphPad Prism 7.0. For all gel-based pIC_{50} determinations three replicates of each condition were used.

In vivo treatment

The animal experiments were conducted in accordance with the guidelines of the ethical committee of Leiden University (DEC#14137). *In vivo* studies were conducted in C57BL/6 mice. Mice were injected i.p. with the inhibitor dissolved in 18:1:1 (v/v/v) mixture of saline/ethanol/PEG-40 (ethoxylated castor oil, 100 µL injections). For the target engagement study, mice were treated with the compounds for 2 h and euthanized by cervical dislocation. Organs were harvested, flash frozen in liquid nitrogen and stored at -80 °C until sample preparation.

Triazole urea library

The triazole urea library was synthesized and characterized by the European Screening Centre in Dundee, Schotland. The general synthetic procedure is identical to that described for **21** (*vide infra*). Compounds were

purified by preparative HPLC/MS and analysed by H-NMR and LC/MS and only submitted to the library if the purity was > 90%.

Synthetic Methods

General remarks

All reactions were performed using oven- or flame-dried glassware and dry (molecular sieves) solvents. Reagents were purchased from Alfa Aesar, Sigma-Aldrich, Acros, and Merck and used without further purification unless noted otherwise. All moisture sensitive reactions were performed under an argon or nitrogen atmosphere.

^{1}H and ^{13}C NMR spectra were recorded on a Bruker DPX-300 (300 MHz), AV-400 (400 MHz) or DRX-500 (500 MHz). Used software for interpretation of NMR-data was Bruker TopSpin 1.3 and MestreNova 11.0. Chemical shift values are reported in ppm with tetramethylsilane or solvent resonance as the internal standard (CDCl$_3$: δ 7.26 for ^{1}H, δ 77.16 for ^{13}C; ACN-d3: δ 1.94 for ^{1}H, δ 1.32 for ^{13}C; MeOD: δ 3.31 for ^{1}H, δ 49.00 for ^{13}C).[14] Data are reported as follows: chemical shifts (δ), multiplicity (s = singlet, d = doublet, dd = double doublet, td = triple doublet, t = triplet, bs = broad singlet, m = multiplet), coupling constants J (Hz), and integration.

Liquid chromatography analysis was performed on a Finnigan Surveyor LC/MS system, equipped with a C18 column. Flash chromatography was performed using SiliCycle silica gel type SiliaFlash P60 (230−400 mesh). TLC analysis was performed on Merck silica gel 60/Kieselguhr F254, 0.25 mm. Compounds were visualized using KMnO$_4$ stain (K$_2$CO$_3$ (40 g), KMnO$_4$ (6 g), and water (600 mL)) or CAM stain (Ce(NH$_4$)$_4$(SO$_4$)$_4$·2H$_2$O (ceric ammonium sulfate: 10.0 g); ammonium molybdate (25 g); conc. H$_2$SO$_4$ (100 mL); H$_2$O (900 mL)). Preparative HPLC (Waters, 515 HPLC pump M; Waters, 515 HPLC pump L; Waters, 2767 sample manager; Waters SFO System Fluidics Organizer; Waters Acquity Ultra Performance LC, SQ Detector; Waters Binary Gradient Module) was performed on a Waters XBridgeTM column (5 μM C18, 150 x 19 mm). Diode detection was done between 210 and 600 nm. Gradient: ACN in (H$_2$O + 0.2% TFA). High resolution mass spectra (HRMS) were recorded by direct injection on a q-TOF mass spectrometer (Synapt G2-Si) equipped with an electrospray ion source in positive mode with Leu-enkephalin (m/z = 556.2771) as an internal lock mass. The instrument was calibrated prior to measurement using the MS/MS spectrum of Glu-1-fibrinopeptide B.

(R,S)-2-Benzylpiperidine (**67**)
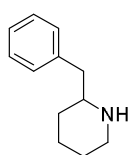
2-Benzylpyridine (5 mL, 31.1 mmol) was dissolved in absolute ethanol (100 mL) and concentrated HCl (10 mL) was added. Then PtO$_2$ (112 mg, 0.49 mmol) was added and the mixture was shaken under a hydrogen atmosphere of 2 bar at RT. After overnight shaking, solids were filtered off over celite. The solvent was removed under reduced pressure and the residue was purified using silica gel chromatography (10% methanol in DCM) to yield the title compound (4.2 g, 20 mmol, 64%). ^{1}H NMR (300 MHz, CDCl$_3$) δ 8.41 (s, 1H), 7.42 – 7.13 (m, 5H), 3.58 – 3.38 (m, 2H), 3.25 – 3.02 (m, 1H), 3.02 – 2.73 (m, 2H), 2.14 – 1.47 (m, 5H), 1.47 – 1.10 (m, 1H). ^{13}C NMR (75 MHz, CDCl$_3$) δ 136.16, 129.51, 128.76, 127.04, 58.59, 45.11, 40.07, 27.97, 22.61.

3-(Benzylthio)-1H-1,2,4-triazole (**69**)
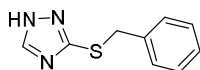
1H-1,2,4-Triazole-3-thiol (1.01 g, 10.0 mmol) and benzylbromide (1.09 mL, 10.0 mmol) were dissolved in DMF (10 mL) and stirred for 18 h. The reaction mixture was then diluted with EtOAc (50 mL) and washed with saturated aqueous NaHCO$_3$ (50 mL). The organic phase was separated and the aqueous phase was extracted with EtOAc (50 mL). The combined organic phases were washed with water (50 mL) and brine (50 mL). The volatiles were removed *in vacuo* and the product was obtained as white solid (1.86 g, 9.75 mmol, 98%). ^{1}H NMR (400 MHz, MeOD) δ 8.30 (s, 1H), 7.35 – 7.14 (m, 5H), 4.33 (s, 2H). ^{13}C NMR (101 MHz, MeOD) δ 147.99, 138.64, 129.91, 129.52, 128.48, 37.80.

(2-Benzylpiperidin-1-yl)(3-(benzylthio)-1H-1,2,4-triazol-1-yl)methanone (**70**)
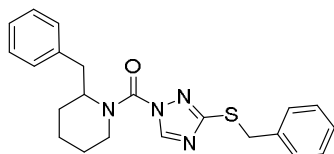
A solution of triphosgene (229 mg, 0.77 mmol) in DCM (5 mL) was cooled to -5 °C, and a solution of **67** (90 mg, 0.51 mmol) in DCM (3 mL) was added dropwise over a period of 5 minutes. To this mixture was added dropwise over a period of 15-30 minutes a solution of TEA (0.21 mL, 1.54 mmol) in DCM (3 mL). The cold bath was removed and stirring at RT was continued for 2 h. The mixture was diluted with DCM (10 mL) and washed with aqueous 1 M HCl (10 mL). The organic phase was separated, dried (MgSO$_4$), filtered and evaporated to dryness. The residue was taken up in DMF (2 mL) to which **69** (32 mg, 0.17 mmol) and K$_2$CO$_3$ (116 mg, 0.84 mmol) were added while stirring at RT. After stirring

overnight, the reaction was quenched with water and the product was extracted into EtOAc. The organic phase was evaporated to dryness and the residual material was purified on preparative HPLC to yield the title compound as a yellow gum (22 mg, 54 μmol, 32%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.31 (bs, 1H), 7.71 – 6.37 (m, 10H), 4.63 (bs, 1H), 4.38 (d, J = 2.5 Hz, 2H), 3.98 (d, J = 13.1 Hz, 1H), 3.62 – 3.20 (m, 2H), 3.09 (d, J = 3.5 Hz, 1H), 2.95 – 2.38 (m, 2H), 2.02 – 1.33 (m, 4H).

1-(2-Benzylpiperidine-1-carbonyl)-*N*-methyl-*N*-phenyl-1*H*-1,2,4-triazole-3-sulfonamide (21)
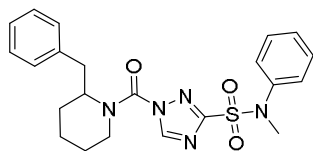
4 M HCl in dioxane (2 mL) was added to DCM (10 mL) and cooled to -10 °C. 15%$_{wt}$ NaOCl solution (aq., 2 mL) was slowly added, a yellow-green colour was observed. After 30 minutes of stirring, **70** (100 mg, 0.255 mmol) in DCM (1 mL) was added slowly. After 30 minutes *N*-methylaniline (2.0 mL, 18 mmol) was added in one batch and stirring was continued for 1 h on ice and 1 h at RT. The DCM was evaporated under reduced pressure and residual solvent was diluted with EtOAc (25 mL) and washed with 0.1 M aqueous HCl (25 mL). The aqueous layer was extracted with EtOAc (2x 25 mL). The combined organic layers were dried (MgSO$_4$), filtered and concentrated. The residue was purified by preparative HPLC to yield the title compound as a green gum (58 mg, 0.13 mmol, 52%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.91 (d, J = 18.9 Hz, 1H), 7.40 – 6.78 (m, 10H), 4.70 (s, 1H), 4.29 – 4.13 (m, 1H), 3.48 (s, 3H), 3.32 – 3.20 (m, 1H), 3.13 (s, 1H), 2.61 (d, J = 18.2 Hz, 1H), 1.74 (m, 6H). HRMS: Calculated for [C$_{22}$H$_{25}$N$_5$O$_3$S + Na]$^+$ = 462.1570, found = 462.1573.

3-(Benzylthio)-*N,N*-diphenyl-1*H*-1,2,4-triazole-1-carboxamide (72)
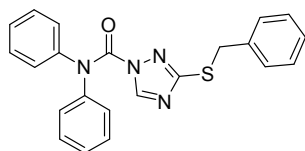
69 (2.137 g, 11.17 mmol) was dissolved in dry THF (40 mL) and diphenylcarbamic chloride (2.85 g, 12.3 mmol), DIPEA (2.34 mL, 13.4 mmol) and a catalytic amount of DMAP were added. The mixture was refluxed for 5 h and stirred for 18 h at RT. The reaction was quenched by the addition of sat. aq. Na$_2$CO$_3$ (40 mL). The organic layer was separated and the aqueous phase extracted with EtOAc (2x 40 mL). The combined organic layers were washed with brine (40 mL), dried (MgSO$_4$), filtered and concentrated. Column chromatography afforded the title compound as a white solid (3.81 g, 9.86 mmol, 88%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.70 (s, 1H), 7.46 – 7.10 (m, 15H), 3.86 (s, 2H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 162.81, 148.53, 147.18, 142.64, 136.61, 129.58, 129.04, 128.65, 127.60, 127.42, 126.80, 35.89.

3-(*N*-Methyl-*N*-phenylsulfamoyl)-*N,N*-diphenyl-1*H*-1,2,4-triazole-1-carboxamide (73)

4 N HCl in dioxane (2 mL) was added to DCM (5 mL) and cooled to -10 °C. 15%$_{wt}$ NaOCl (aq., 2 mL) was added slowly, forming a yellow-green Cl$_2$ solution. After 30 min **72** (100 mg, 0.259 mmol) was added dropwise as a solution in DCM (1 mL). After 20 min *N*-methylaniline (1 mL, 9 mmol) was added. The mixture was stirred for 2 h at RT. The reaction was quenched by the addition of 0.1 M HCl (10 mL). The organic layer was separated and the water layer was extracted with EtOAc (2x, 10 mL). The combined organic layers were dried (MgSO$_4$), filtered and concentrated. The residue was purified by column chromatography to yield the title compound as a brown solid (78 mg, 0.18 mmol, 69%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.80 (s, 1H), 7.47 – 6.99 (m, 15H), 3.07 (s, 3H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 160.95, 147.73, 147.28, 141.75, 140.35, 129.45, 128.87, 127.58, 127.47, 126.76, 126.40, 39.02.

N-Methyl-*N*-phenyl-1*H*-1,2,4-triazole-3-sulfonamide (74)

73 (1.05 g, 2.42 mmol) was dissolved in 2 N KOH in a 1:1 (v/v) H$_2$O/THF mixture (50 mL). The mixture was stirred for 18 h at RT. The mixture was concentrated and the residue was purified by column chromatography to yield the title compound (308 mg, 1.29 mmol, 53%). ^{1}H NMR (400 MHz, MeOD) δ 8.63 (s, 1H), 7.46 – 7.09 (m, 5H), 3.42 (s, 3H).

(2*R*,5*R*)-2-Benzyl-5-(prop-2-yn-1-yloxy)piperidine 2,2,2-trifluoroacetate (**76**)
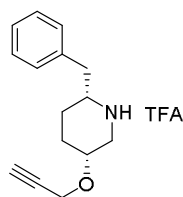
Previously prepared *tert*-butyl (2*R*,5*R*)-2-benzyl-5-(prop-2-yn-1-yloxy)piperidine-1-carboxylate (30 mg, 91 μmol) was dissolved in DCM (6 mL) to which trifluoroacetic acid (2 mL, 26 mmol) was added. The mixture was stirred for 5 h at RT. The volatiles were evaporated and the crude residue was coevaporated with toluene (3x 5 mL) and used without further purification for the next reaction.

1-((2*R*,5*R*)-2-Benzyl-5-(prop-2-yn-1-yloxy)piperidine-1-carbonyl)-*N*-methyl-*N*-phenyl-1*H*-1,2,4-triazole-3-sulfonamide (**77**)
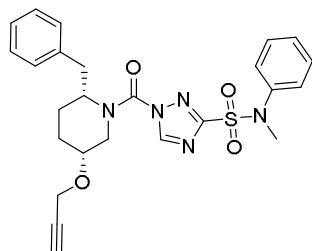
76 (31.3 mg, 0.091 mmol) was dissolved in DCM (5 mL) and cooled to 0 °C. Triphosgene (13.5 mg, 0.046 mmol) and DIPEA (40 μL, 0.30 mmol) were added and the mixture was stirred for 1 h warming to RT. The reaction was quenched with ice water (5 mL) and the organic layer was separated. The aqueous layer was extracted with EtOAc (2x 5 mL). The combined organic layers were dried (MgSO4), filtered and concentrated under reduced pressure. The crude was taken up in dry THF (5 mL) and **74** (24 mg, 0.10 mmol), DMAP (11 mg, 0.091 mmol) and DIPEA (40 μL, 0.30 mmol) were added. The mixture was refluxed for 5 h. The reaction was quenched with saturated aqueous NH4Cl (5 mL). The organic layer was diluted with EtOAc (5 mL) and separated. The aqueous layer was extracted with EtOAc (2x 10 mL). The combined organic layers were washed with brine, dried (MgSO4), filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography to yield the title compound as a white solid (30.2 mg, 61 μmol, 67%). ^{1}H NMR (400 MHz, CDCl3) δ 7.99 (s, 1H), 7.42 – 6.78 (m, 10H), 4.71 (s, 1H), 4.48 (s, 1H), 4.30 (s, 2H), 3.79 – 3.62 (m, 1H), 3.52 (s, 3H), 3.21 – 2.99 (m, 2H), 2.67 (s, 1H), 2.49 (s, 1H), 2.10 (d, *J* = 19.3 Hz, 1H), 1.80 (d, *J* = 11.8 Hz, 3H). HRMS: Calculated for $[C_{25}H_{27}N_5O_4S + Na]^+$ = 516.1676, found = 516.1680.

Supplementary Figures
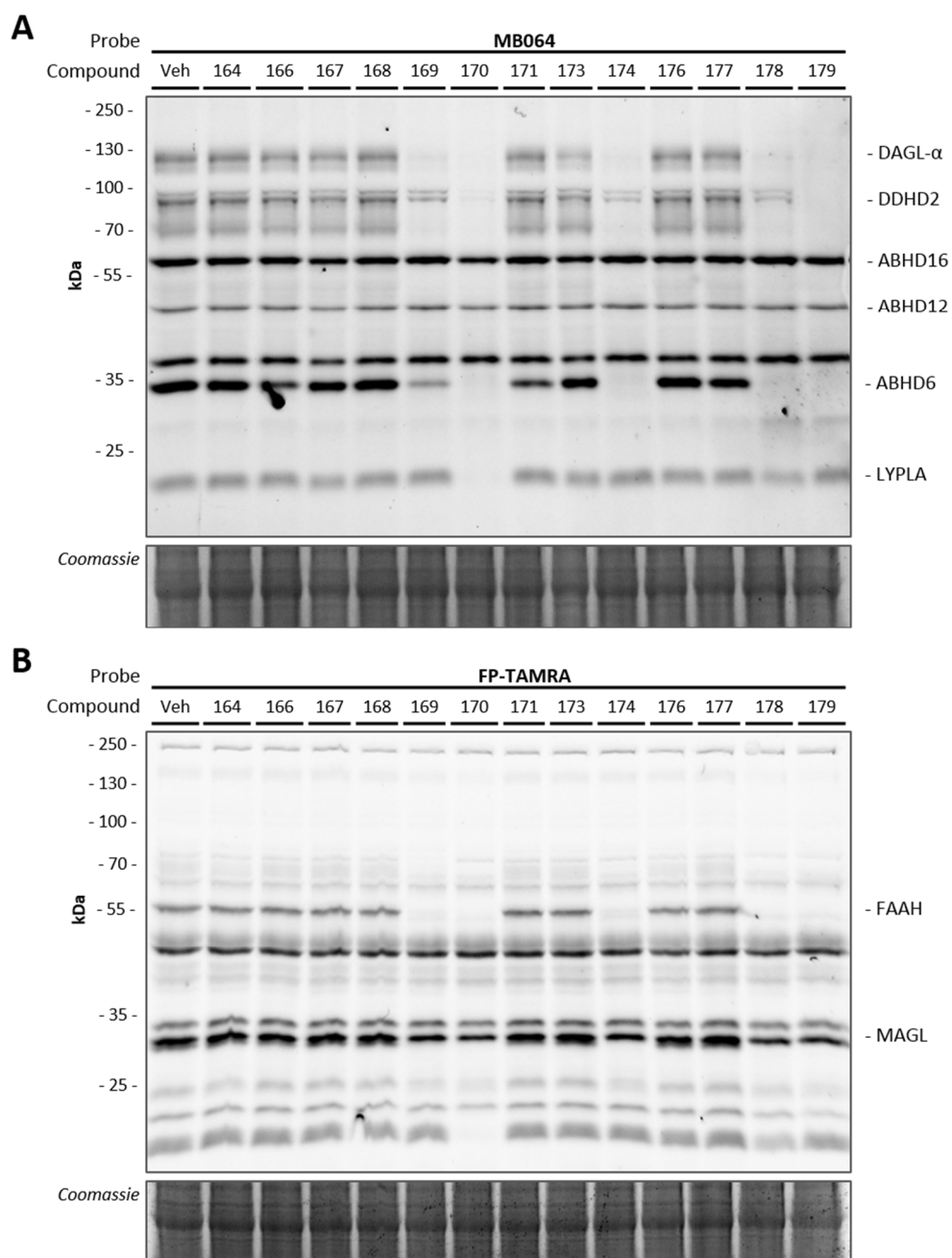
Figure S2.1 | Typical off-target profiling screen using MB064 (A) and FP-TAMRA (B). Compound numbers correspond to compound ESC-library codes. Compounds were screened at a concentration of 10 μM.

References

1.   Hsu, K. *et al.* DAGLβ inhibition perturbs a lipid network involved in macrophage inflammatory responses. *Nat. Chem. Biol.* **8**, 999–1007 (2012).
2.   Janssen, F. J. & van der Stelt, M. Inhibitors of diacylglycerol lipases in neurodegenerative and metabolic disorders. *Bioorg. Med. Chem. Lett.* **26**, 3831–3837 (2016).
3.   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).
4.   Baggelaar, M. P. *et al.* Highly Selective, Reversible Inhibitor Identified by Comparative Chemoproteomics Modulates Diacylglycerol Lipase Activity in Neurons. *J. Am. Chem. Soc.* **137**, 8851–8857 (2015).
5.   Janssen, F. J. *et al.* Discovery of Glycine Sulfonamides as Dual Inhibitors of sn -1-Diacylglycerol Lipase α and α/β-Hydrolase Domain 6. *J. Med. Chem.* **57**, 6610–6622 (2014).
6.   Hu, S. *et al.* Glycine chroman-6-sulfonamides for use as inhibitors of diacylglycerol lipase. US-patent US 8354548 B2. (2011).
7.   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).
8.   Deng, H. *et al.* Triazole Ureas Act as Diacylglycerol Lipase Inhibitors and Prevent Fasting-Induced Refeeding. *J. Med. Chem.* **60**, 428–440 (2017).
9.   Chupak, L. S. *et al.* Structure activity relationship studies on chemically non-reactive glycine sulfonamide inhibitors of diacylglycerol lipase. *Bioorg. Med. Chem.* **24**, 1455–1468 (2016).
10. Janssen, F. J., Discovery of sulfonyl-1,2,4-triazole ureas as DAGLα inhibitors by HTS-ABPP. in *Discovery of novel inhibitors to investigate diacylglycerol lipases and α/β-hydrolase domain 16A* 109–139 (2016).
11. Hsu, K.-L. *et al.* Development and Optimization of Piperidyl-1,2,3-Triazole Ureas as Selective Chemical Probes of Endocannabinoid Biosynthesis. *J. Med. Chem.* **56**, 8257–8269 (2013).
12. van der Wel, T. *et al.* A natural substrate-based fluorescence assay for inhibitor screening on diacylglycerol lipase α. *J. Lipid Res.* **56**, 927–935 (2015).
13. 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).
14. Gottlieb, H. E., Kotlyar, V. & Nudelman, A. NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. *J. Org. Chem.* **62**, 7512–7515 (1997).

Towards drug-like peripherally restricted DAGL-inhibitors

Introduction

The signalling of 2-arachidonoylglycerol (2-AG) via the endocannabinoid system is implicated in a wide range of physiological effects, such as neurotransmission, inflammation and energy metabolism.[1,2,3] The diacylglycerol lipase (DAGL) inhibitors, introduced in Chapter 2, have reasonable physicochemical properties and showed promising activity to study the role of DAGL in 2-AG signalling *in vivo*. They lacked, however, subtype selectivity over DAGL-α and DAGL-β, and also inhibited α/β-hydrolase domain containing protein 6 (ABHD6).[4] Despite studies using genetic knock-out mice, the role and importance of the two distinct DAGL isoforms is still poorly understood, and should ideally

be studied with isoform-selective inhibitors.[1,5,6] Since there is no structural data of these enzymes available, the search for such selective inhibitors is challenging. The spatial distribution of the isoforms in the body does point towards an alternative way to obtain largely selective modulation of DAGL activity: DAGL-α is predominantly expressed in the central nervous system (CNS), whereas DAGL-β is more highly expressed in the periphery.[7] By employing the drug design principles that do not allow drugs to pass the blood-brain barrier, compounds could in principle be restricted to the periphery, making them spatially selective towards DAGL-β. By limiting the brain penetration, unwanted CNS-side effects such as those observed for the retracted drug Rimonabant could be reduced or even prevented completely.[8]

It is generally accepted that the topological polar surface area (TPSA), together with the lipophilicity, expressed as (C)LogP, are the best predictors of brain penetration.[9] Hitchcock and Pennington described an optimal range for the lipophilicity where the LogP is between 2 and 4 and the limit for the polar surface area is below 90 Å^2. Further requirements include a maximum of three hydrogen bond donors. Compounds that do not meet these criteria are less likely to penetrate the brain.

The work presented here builds on the findings of Chapter 2 to work towards more drug-like and peripherally restricted compounds. Ultimately, these inhibitors could aid in the elucidation of the different physiological roles the DAG lipases play in human health and disease.

Structure activity relationship study on 4-substituted piperazine derivatives

During the hit optimization, which led to the discovery of *in vivo* active compound **1** (Figure 3.1, Chapter 2), a similarity in the structure activity relationships (SAR) with other triazole urea inhibitors was observed. The activity of DAGL inhibitor DO34 (Figure 3.1) demonstrated that there is ample room for variation in the 2-benzylpiperidine moiety[10], therefore it was envisioned that optimization of this part of **1** could be exploited to increase the TPSA. First, the substituted piperazine featured in DO34 was resynthesized in a 7-step procedure from commercially available *N*-Boc-phenylalanine and *N*-benzylglycine (Scheme S3.1), and was subsequently coupled to the two most promising scaffolds of Chapter 2, i.e. the 1,2,4-triazoles substituted with either an *N*-methylaniline-sulfonamide (**2**) or a benzyl-sulfone (**3**) (Table 3.1).

The sulfone analogue **3** showed comparable activity to **2** for DAGL-α and -β in the previously published surrogate substrate assay employing para-nitrophenol butyrate (PNPB) (Table 3.1).[11] In view of its synthetic accessibility, the sulfone scaffold was selected
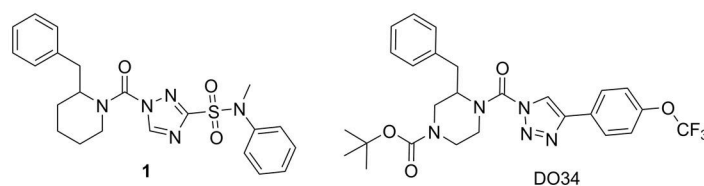
Figure 3.1 | Triazole urea compound **1**, presented in Chapter 2, and DO34, a published DAGL inhibitor featuring a 4-substituted piperazine.

Table 3.1 | Matched molecular pair analysis to validate the sulfone as synthon for a DAGL inhibitor library.

#	Structure	PNPB pIC$_{50}$ DAGL-α	ABPP pIC$_{50}$ DAGL-α	PNPB pIC$_{50}$ DAGL-β	CLogP	TPSA (Å^2)	LipE DAGL-α
2		8.97 ± 0.08	8.62 ± 0.13	8.25 ± 0.07	4.63	115.2	4.34
3		8.71 ± 0.06	7.88 ± 0.14	8.40 ± 0.07	4.97	112.0	3.74

to quickly generate SAR by varying the 4-piperazine substitutions in a matched molecular pair analysis. A small library was synthesized (Scheme S3.1) starting from the two 1-protected piperazine intermediates, which were synthesized according to Scheme S3.2. The 4-position was modified by 1-step reactions to generate the desired substitution. The resulting piperazines were deprotected and coupled to 3-(benzylsulfonyl)-1*H*-1,2,4-triazole, which is available in two high-yielding steps from commercial building blocks. The resulting library was tested in both the PNPB-assay for DAGL-α and DAGL-β activity and in a competitive activity-based protein profiling (ABPP) assay using MB064 as the activity-based probe (Table 3.2).

First, several amide variations were made with substituents increasing in size and lipophilicity from an acetyl (**4**), via cyclic and bulky alkyl substituents (**5** and **6**) to a benzoyl substituent (**7**). The potencies of the amides were all lower than *tert*-butyl carbamate containing compound **3**. The largest potency reduction (200-fold) was observed for the acetyl substitution in **4**. The more lipophilic substituents (**5-7**) retained most of their activity, with about a 10-fold reduced potency. The keto-piperazine analogue **8** showed a 10-fold drop as compared to the highly similar benzoyl substituted compound **7**. Three related urea moieties were introduced: unsubstituted (**9**), ethyl substituted (**10**) and a Boc-like *tert*-butyl urea (**11**). These followed the same trend in activity with the more substituted and lipophilic substituents showing higher inhibitory activity.

The introduction of a small alkyl substituent, generating a basic tertiary amine, did not dramatically reduce potency for the DAGLs (**12-13**). The calculated LipE for **13** (3.97) was the highest observed thus far, and was slightly better when using the (calculated) LogD for lipophilicity. As with the amide and urea substitutions, increasing the bulk and lipophilicity of the substituent increased the potency, reaching single digit nanomolar IC$_{50}$s for all variations (**14-18**). However, the TPSA of the most potent compounds (**13-18**) was actually reduced compared to the Boc-substituted **3**, and was below the 90 Å^2. Therefore, sulfonamide substitutions **19-22** were designed to incorporate another highly polar functionality in the scaffold. Again, a small substituent showed a large decrease in activity (**19**), but potency could be recovered by the introduction of an isopropyl (**20**) or phenyl substituent (**21, 22**). The sulfonamide substitution increased the TPSA to 120 Å^2, making it unlikely these compounds will be centrally active. The low CLogP combined with rather high potency for DAGL-α resulted in a LipE for **20** of 4.13, slightly higher than amine **13**.

Table 3.2 | Piperazine variations of the 1,2,4-triazole-sulfone scaffold.
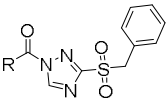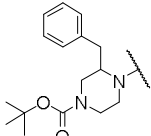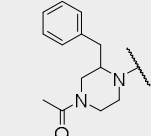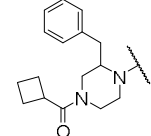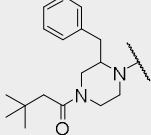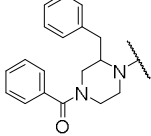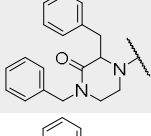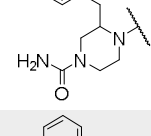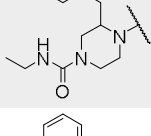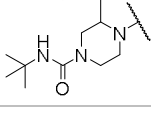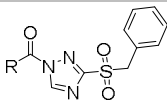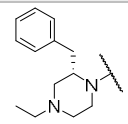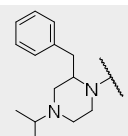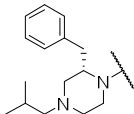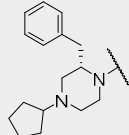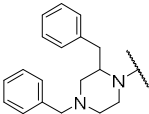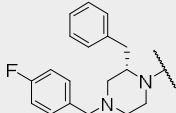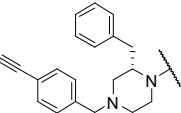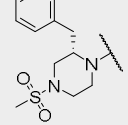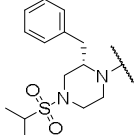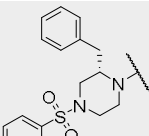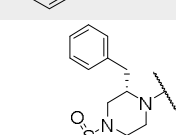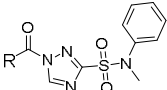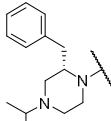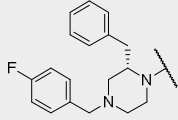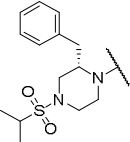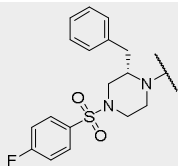
#	R	PNPB pIC$_{50}$ DAGL-α	ABPP pIC$_{50}$ DAGL-α	PNPB pIC$_{50}$ DAGL-β	CLogP	TPSA (Å^2)	LipE DAGL-α
3		8.71 ± 0.06	7.88 ± 0.14	8.40 ± 0.07	4.97	112.0	3.74
4		6.25 ± 0.08	5.61 ± 0.22	5.74 ± 0.22	2.58	102.7	3.67
5		7.41 ± 0.07	7.57 ± 0.45	7.25 ± 0.09	3.49	102.7	3.92
6		7.74 ± 0.07	7.91 ± 0.15	7.24 ± 0.07	4.44	102.7	3.30
7		7.71 ± 0.06	6.87 ± 0.25	7.24 ± 0.06	4.16	102.7	3.55
8		6.66 ± 0.20	6.14 ± 0.30	5.69 ± 0.14	4.58	102.7	2.08
9		< 5	< 5	< 5	1.77	128.7	-
10		6.52 ± 0.19	5.95 ± 0.19	6.27 ± 0.18	3.59	114.8	2.93
11		6.77 ± 0.05	6.54 ± 0.13	6.68 ± 0.07	4.14	114.8	2.63

Table 3.2 (continued)| Piperazine variations of the 1,2,4-triazole-sulfone scaffold.

#	R	PNPB pIC$_{50}$ DAGL-α	ABPP pIC$_{50}$ DAGL-α	PNPB pIC$_{50}$ DAGL-β	CLogP	TPSA (Å^2)	LipE DAGL-α
12		7.46 ± 0.05	N.D.	7.54 ± 0.23	4.09	85.7	3.37
13		8.37 ± 0.08	7.95 ± 0.25	7.99 ± 0.08	4.40	85.7	3.97
14		8.92 ± 0.03	8.67 ± 0.09	8.35 ± 0.11	5.02	85.7	3.90
15		8.56 ± 0.10	N.D.	8.81 ± 0.18	4.82	85.7	3.74
16		9.05 ± 0.06	8.53 ± 0.15	8.98 ± 0.10	5.28	85.7	3.77
17		8.12 ± 0.11	8.26 ± 0.14	N.D.	5.42	85.7	2.70
18		8.58 ± 0.11	7.83 ± 0.15	9.28 ± 0.12	5.55	85.7	3.03
19		6.75 ± 0.07	N.D.	6.73 ± 0.20	3.00	119.8	3.75
20		7.97 ± 0.12	7.72 ± 0.20	6.50 ± 0.31	3.84	119.8	4.13
21		8.50 ± 0.06	8.45 ± 0.10	7.88 ± 0.08	4.68	119.8	3.82
22		7.83 ± 0.11	N.D.	6.39 ± 0.30	4.82	119.8	3.01

Four piperazine substituents of the compounds with highest potency, LipE and drug-like character (Table 3.2) were combined with the *N*-methylaniline substituted 1,2,4-triazole sulfonamide scaffold of **1** to form compounds **23-26** (Table 2.2). The potency of the sulfonamide derivatives increased 5-10 fold compared to their sulfone analogues. Both amine compounds **23** and **24** showed sub-nanomolar potencies in the surrogate substrate assay and maintained high potency in the gel-based ABPP assays. Compounds **25** and **26** were slightly less active, but still showed IC$_{50}$s in the single nanomolar range. All four compounds were slightly less active against DAGL-β, consistent with the observations for compounds **3-22**. The lipophilic efficiency was further increased. The TPSA of the amines is just below 90 Å^2, which could allow these compounds to be centrally active. **25** and **26** have a TPSA > 120 Å^2, making it highly unlikely that they are able to cross the blood-brain barrier. Together, they could be employed to elucidate peripheral effects using **25** or **26**, or central effects, by comparing the effects of **23** or **24** with **25** or **26**.
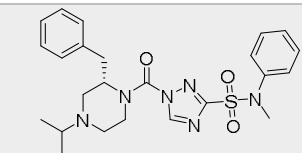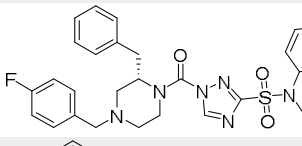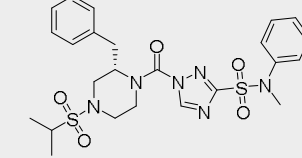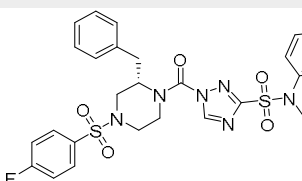
Table 3.3 | *In vitro* profiling of the four most promising piperazine groups combined with the sulfonamide leaving group. Activity-based protein profiling of DAGL-α is performed using mouse brain membrane proteome. For DAGL-β membrane fractions of HEK293T cells overexpressing mDAGL-β were used.

#	R	PNPB pIC$_{50}$ DAGL-α	ABPP pIC$_{50}$ DAGL-α	PNPB pIC$_{50}$ DAGL-β	ABPP pIC$_{50}$ DAGL-β	CLogP	TPSA (Å^2)	LipE DAGL-α
23		9.06 ± 0.07	8.93 ± 0.08	7.84 ± 0.25	8.54 ± 0.06	4.06	88.9	5.00
24		9.91 ± 0.11	9.45 ± 0.04	8.47 ± 0.19	8.44 ± 0.14	5.08	88.9	4.83
25		8.55 ± 0.10	7.97 ± 0.10	8.52 ± 0.11	7.70 ± 0.09	3.50	123.0	5.05
26		8.72 ± 0.12	9.44 ± 0.12	8.12 ± 0.08	8.26 ± 0.25	4.48	123.0	4.24

Cellular activity and selectivity of **23**-**26**

Cellular DAGL-β activity of compounds **23**-**26** was assessed by the treatment of murine neuroblastoma (Neuro-2a) cells for 1 hour with medium containing inhibitor or vehicle (DMSO) followed by gel-based ABPP. Post-lysis treatment with activity-based probes MB064, to visualize DAGL-β and other lipases such as Phospholipase DDHD2 (DDHD2) and ABHD6, showed clear inhibition of DAGL-β (Figure 3.2, Table 3.4). FP-BODIPY (Chapter 6) was used to label the endocannabinoid related hydrolase Fatty Acid Amide Hydrolase (FAAH) and to further assess general selectivity. Overall potency for DAGL-β of the inhibitors was reduced 10-50 fold in this setting. Quite surprisingly, the selectivity observed in the gel-based assays *in vitro* was attenuated *in situ*. The potency of the inhibitors against ABHD6 was equal or greater than that for DAGL-β, and further off-target activity was observed for DDHD2 and FAAH.

Table 3.4 | *In situ* potency of the optimized piperazine derivatives **23**-**26** tested in Neuro-2a cells determined by post-lysis ABPP and SDS-PAGE analysis.

#	Structure	pIC_{50} DAGL-β	pIC_{50} DDHD2	pIC_{50} ABHD6	pIC_{50} FAAH
23		6.68 ± 0.13	5.57 ± 0.07	8.00 ± 0.05	6.16 ± 0.08
24		7.58 ± 0.10	6.76 ± 0.11	7.84 ± 0.05	6.01 ± 0.22
25		6.09 ± 0.12	5.53 ± 0.21	7.06 ± 0.12	5.48 ± 0.13
26		6.70 ± 0.08	6.31 ± 0.11	7.38 ± 0.07	5.35 ± 0.41
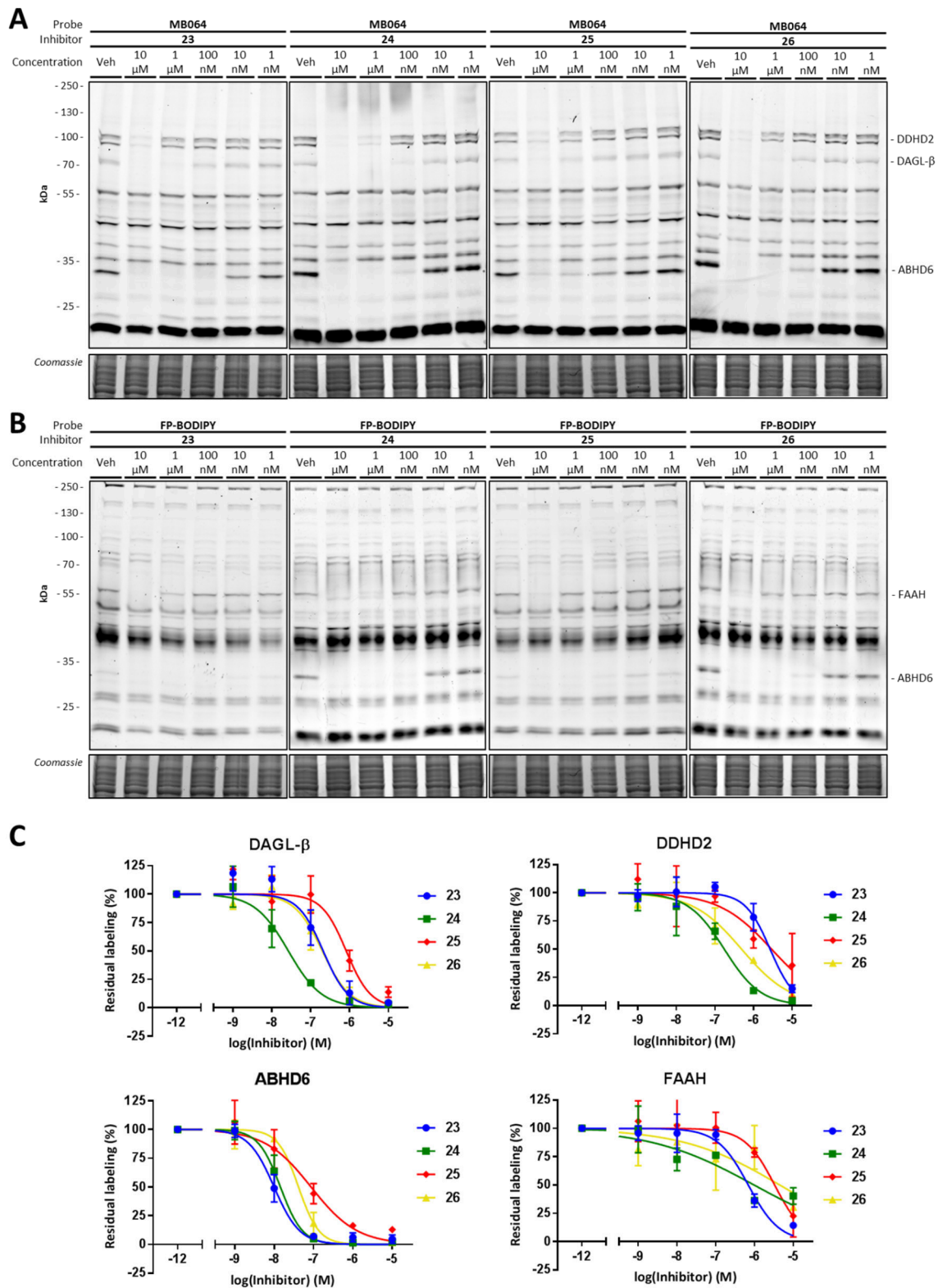
Figure 3.2 | Dose response *in situ* treatment of Neuro-2a cells with piperazine inhibitors **23-26**. A) Representative gels of lysates labeled with MB064 (250 nM). B) Representative gels of lysates labeled with FP-BODIPY (500 nM). C) Dose response curves for DAGL-β, DDHD2, ABHD6 and FAAH as based on A and B (N=3).

Discussion and conclusion

The optimization of the physicochemical properties of the DAG lipase inhibitors introduced in Chapter 2 was initiated by incorporating a previously published piperazine as the amine for the triazole urea scaffold.[10] Successful implementation of a matched molecular pair strategy significantly sped up the generation of a library of 2-benzyl-4-substituted piperazine triazole ureas. Similar trends were observed for acyl, alkyl, urea and sulfonyl substituents, with a strong preference for lipophilic and bulky sidechains. Intriguingly, the Boc-group (**3**), 3,3-dimethylbutyric acid (**6**) and *tert*-butylamine urea (**11**) variants show a nearly perfect trend, each decreasing a factor 10 in potency. The CLogP (ChemDraw 16.0) also follows a nearly perfect linear decrease, albeit with a slightly smaller slope: from 4.97 for **3**, 4.44 for **6** to 4.14 for **11**. This likely explains the observed drop in affinity.

The optimized inhibitors **23**-**26** showed (sub)nanomolar potency and good selectivity *in vitro*, but both were attenuated *in situ*. For compounds **25** and **26** the TPSA is probably too high for efficient cell penetration. The decrease in selectivity has not been investigated in detail, but might be attributed to different physical conditions, such as local pH, and differences in enzyme activity *in situ* as compared to *in vitro*. Especially the small selectivity window with respect to FAAH should be noted, as this is the main metabolizing enzyme for anandamide, the other major endocannabinoid.[12] Further profiling of the generated library *in situ* is warranted to possibly find more selective members.

Overall potency on both DAGL isoforms was slightly improved, but more importantly the lipophilic efficiency was increased from 4.34 for **2** to 5.05 for **25**, the highest LipE reported to date for a DAGL inhibitor. The topological polar surface area of **25** and **26** is > 120 Å^2, which should be sufficient to restrict these compounds to the periphery. However, their high TPSA likely hampers their cell penetration and may affect their oral bioavailability. Future studies should test whether these compounds can be used in animal models to inhibit DAGL-β in the periphery without inhibiting DAGL-α in the brain.

Acknowledgements

Methods

Chemical Biology Methods

Cell Culture

HEK293T (human embryonic kidney) and Neuro-2a (mouse neuroblastoma) cells were cultured at 37 °C under 7% CO_2 in DMEM containing phenol red, stable glutamine, 10% (v/v) New Born Calf Serum (Thermo Fisher), and penicillin and streptomycin (200 µg/mL each; Duchefa). Medium was refreshed every 2-3 days and cells were passaged twice a week at 80-90% confluence by resuspension in fresh medium.
Cells lines were purchased from ATCC and were regularly tested for mycoplasma contamination. Cultures were discarded after 2-3 months of use.

Transient transfection

One day prior to transfection HEK293T cells were seeded to 15-cm dishes (~62.500 cells/cm²). Prior to transfection, culture medium was aspirated and a minimal amount of medium was added. A 3:1 (m/m) mixture of polyethyleneimine (PEI) (60 µg/dish) and plasmid DNA (20 µg/dish, either mDAGL-α-FLAG or mDAGL-β-FLAG) was prepared in serum-free culture medium and incubated for 15 min at RT. Transfection was performed by dropwise addition of the PEI/DNA mixture to the cells. Transfection with the empty pcDNA3.1 vector was used to generate control samples. After 24 h, medium was refreshed. Medium was aspirated 48 h post-transfection and cells were harvested by resuspension in PBS. Cells were pelleted by centrifugation (5 min, 1,000 g) and the pellet was washed with PBS. Supernatant was discarded and cell pellets were frozen in liquid nitrogen and stored at -80 °C until sample preparation.

In situ treatment of Neuro-2a cells

Cells were seeded in 12-well plates. *In situ* treatment was initiated 24 h later. Medium was aspirated and medium (with serum) containing inhibitor or DMSO as vehicle was added (0.1% v/v DMSO). After 1 h exposure to the treatment medium, the medium was aspirated and cells were harvested and stored as described above until sample preparation.

Whole cell lysate

Cell pellets were thawed on ice, resuspended in cold lysis buffer (20 mM HEPES pH 7.2, 2 mM DTT, 250 mM sucrose, 1 mM $MgCl_2$, 2.5 U/mL benzonase) and incubated on ice (15-30 min). The cell lysate was used for membrane preparation (below) or was diluted to 2.0 mg/mL concentration in cold storage buffer (20 mM Hepes, pH 7.2, 2 mM DTT) for use as whole lysate. Protein concentrations were determined by a Quick Start™ Bradford Protein Assay and diluted samples were flash frozen in liquid nitrogen and stored at -80 °C until further use.

Membrane preparation from overexpression lysate

The membrane and cytosolic fractions of cell lysates were separated by ultracentrifugation (93,000 g, 45 min, 4 °C). The supernatant was collected (cytosolic fraction) and the membrane pellet was resuspended in cold storage buffer (20 mM HEPES, pH 7.2, 2 mM DTT) by thorough pipetting and passage through an insulin needle. Protein concentrations were determined by a Quick Start™ Bradford Protein Assay and samples were diluted to 2.0 mg/mL with cold storage buffer, flash frozen in liquid nitrogen and stored at -80 °C until further use.

Tissue preparation

Organs were isolated from C57BL/6 mice following standard guidelines as approved by the ethical committee of Leiden University (DEC#13191). Isolated organs were frozen in liquid nitrogen and stored at -80 °C. Organs were thawed on ice and homogenized by a glass stick douncing homogenizer in cold lysis buffer (20 mM HEPES, pH 7.2, 2 mM DTT, 1 mM $MgCl_2$, 2.5 U/mL benzonase). The resulting suspension was centrifuged at 1000 g for 5 minutes at 4 °C to get rid of residual solid tissue. The supernatant was centrifuged 45 minutes at 93,000 g at 4 °C to separate soluble (cytosol) and insoluble (membrane) fractions. The pellet (membrane) was resuspended in storage buffer (20 mM HEPES, pH 7.2, 2 mM DTT) using an insulin syringe. The protein concentration was measured using a Qubit™ protein assay and was adjusted to 2 mg/mL using storage buffer. The resulting lysates were frozen in liquid nitrogen and stored at -80 °C for later use.

Surrogate substrate assay

The biochemical DAGL-α or -β activity assay is based on the method previously described.[11] 200 μL reactions were performed in flat bottom Greiner 96-wells plates in a 50 mM pH 7.2 HEPES buffer, for DAGL-β this buffer was supplemented with 5 mM $CaCl_2$. Membrane protein fractions from HEK293T cells transiently transfected with mDAGL-α or -β (0.05 μg/μL final concentration) were used as mDAGL source. Inhibitors were introduced in 5.0 μL DMSO. The mixtures were incubated for 20 minutes before 10.0 μL 6 mM (DAGL-α) or 12 mM (DAGL-β) PNP-butyrate (final concentration 0.3 mM) in 50% DMSO was added (final DMSO concentration 5.0%). Reactions were allowed to progress for 30 minutes at 20 °C before OD (420 nm) was measured using a TECAN GENios plate reader. All experiments were performed at N=2, n=2 for experimental measurements and N=2, n=4 for controls.

Z'-factor of each plate was determined for the validation of each experiment, using the following formula: $Z' = 1 - 3(\sigma_{pc} + \sigma_{nc})/(\mu_{pc} - \mu_{nc})$. The OD from the positive control (pc: DAGL DMSO), and the negative control (nc: 10 μM THL) were used. Plates were accepted for further analysis when Z' > 0.6. Measurements were corrected for the average absorption of the negative control. The average, standard deviation (SD) and standard error of mean (SEM) were calculated and normalized to the corrected positive control. Data was exported to Graphpad Prism 7.0 for the calculation of the pIC_{50} using a non-linear dose-response analysis with variable slope.

Activity-based protein profiling

For *in vitro* inhibition, lysates (19 μL per sample, 2 μg/μL) were thawed on ice. 0.5 μL of the inhibitor (40x stock in DMSO) or pure DMSO (as vehicle) was added to the sample, vortexed briefly and incubated for 20 minutes at RT. Subsequently, 0.5 μL probe (40x stock in DMSO, final concentration 250 nM for MB064[11], 500 nM for FP-TAMRA[13], 1 μM for DH379[10]) was added to the proteome sample, vortexed briefly and incubated for 15 minutes at RT. For *in situ* inhibition, the *in situ*-treated cells (19.5 μL whole lysate) were directly incubated with the activity based probe (40x stock in DMSO, final concentration 2 μM MB064 or 500 nM FP-BODIPY, 20 min, RT). Final volume in all cases was 20 μL (max. 5% DMSO). The reaction was quenched by the addition of 7.5 μL of 4*Laemmli-buffer (final concentrations: 60 mM Tris (pH 6.8), 2% (w/v) SDS, 10% (v/v) glycerol, 1.25% (v/v) β-mercaptoethanol, 0.01% (v/v) bromophenol blue). 10 μL (14 μg protein) of quenched reaction mixture was resolved on 10% acrylamide SDS-PAGE (180 V, 75 min). Fluorescence was measured using a Biorad ChemiDoc MP system (fluorescence channels Cy2 (460-490 nm), Cy3 (520-545 nm), Cy5 (625-650 nm) filters). Gels were then stained using coomassie staining and imaged for protein loading control.

Labeling quantification

Fluorescence quantification was performed using Imagelab 6.0 (Biorad). Intensities were normalized to the DMSO control and corrected for protein loading by coomassie staining. pIC_{50} values were calculated with GraphPad Prism 7.0 using a non-linear dose-response analysis with variable slope. For all pIC_{50} determinations three replicates of each condition were used.

Synthetic Methods

General remarks

All reactions were performed using oven- or flame-dried glassware and dry (molecular sieves) solvents. Reagents were purchased from Alfa Aesar, Sigma-Aldrich, Acros, and Merck and used without further purification unless noted otherwise. All moisture sensitive reactions were performed under an argon or nitrogen atmosphere. ^{1}H and ^{13}C NMR spectra were recorded on a Bruker DPX-300 (300 MHz), AV-400 (400 MHz) or DRX-500 (500 MHz). Used software for interpretation of NMR-data was Bruker TopSpin 1.3 and MestreNova 11.0. Chemical shift values are reported in ppm with tetramethylsilane or solvent resonance as the internal standard (CDCl₃: δ 7.26 for ^{1}H, δ 77.16 for ^{13}C; ACN-*d3*: δ 1.94 for ^{1}H, δ 1.32 for ^{13}C; MeOD: δ 3.31 for ^{1}H, δ 49.00 for ^{13}C).[14] Data are reported as follows: chemical shifts (δ), multiplicity (s = singlet, d = doublet, dd = double doublet, td = triple doublet, t = triplet, q = quartet, bs = broad singlet, m = multiplet), coupling constants *J* (Hz), and integration.

Liquid chromatography analysis was performed on a Finnigan Surveyor LC/MS system, equipped with a C18 column. Flash chromatography was performed using SiliCycle silica gel type SiliaFlash P60 (230−400 mesh). TLC analysis was performed on Merck silica gel 60/Kieselguhr F254, 0.25 mm. Compounds were visualized using $KMnO_4$ stain (K_2CO_3 (40 g), $KMnO_4$ (6 g), and water (600 mL)) or CAM stain (Ce(NH₄)₄(SO₄)₄·2H₂O (ceric ammonium sulfate: 10 g); ammonium molybdate (25 g); conc. H_2SO_4 (100 mL); H_2O (900 mL)). Preparative HPLC (Waters, 515 HPLC pump M; Waters, 515 HPLC pump L; Waters, 2767 sample manager; Waters SFO

System Fluidics Organizer; Waters Acquity Ultra Performance LC, SQ Detector; Waters Binary Gradient Module) was performed on a Waters XBridgeTM column (5 μM C18, 150 x 19 mm). Diode detection was done between 210 and 600 nm. Gradient: ACN in (H_2O + 0.2% TFA). High resolution mass spectra (HRMS) were recorded by direct injection on a q-TOF mass spectrometer (Synapt G2-Si) equipped with an electrospray ion source in positive mode with Leu-enkephalin (m/z = 556.2771) as an internal lock mass. The instrument was calibrated prior to measurement using the MS/MS spectrum of Glu-1-fibrinopeptide B.

Scheme S3.1 | General scheme for the synthesis of the triazole urea inhibitors. *i*) introduction or substituent R; *ii*) deprotection of amine, General procedure 1 or 2; *iii*) triphosgene coupling to triazoles, General procedure 3.

General Procedure 1: Debocylation of piperazine intermediate

To a 1 M HCl solution in MeOH *tert*-butyl 4-substituted-2-benzylpiperazine-1-carboxylate was added (± 0.1 M). The mixture was stirred for 4-18 h until LC/MS analysis showed full conversion of starting material. Volatiles were removed under reduced pressure to yield the HCl salt of the deprotected piperazine.

General Procedure 2: Removal of trifluoroacetyl protecting group from piperazine intermediate

1-(2-Benzyl-4-substituted-piperazin-1-yl)-2,2,2-trifluoroethan-1-one was dissolved in 20% H_2O in MeOH (± 0.1 M). 3 Equivalents of K_2CO_3 were added. The mixture was stirred for 20-40 h until LC/MS analysis showed full conversion of starting material. The solids were filtered off and the filtrate concentrated *in vacuo* to obtain the deprotected piperazine.

General Procedure 3: Triphosgene coupling of piperazine derivative to 3-substituted-1*H*-1,2,4-triazoles

1-Substituted-3-benzylpiperazine was dissolved in dry DCM (± 0.1 M) and 3 equivalents of Na_2CO_3 were added. The mixture was cooled to 0 °C and 0.75 equivalent of triphosgene was added. The mixture was stirred at 0 °C for 1.5 h. The solids were filtered off and the solvent was evaporated under reduced pressure. The residue was redissolved in dry THF (± 0.1 M), brought under N_2-atmosphere and 0.95 equivalent of 3-substituted-1*H*-1,2,4-triazole, 0.1 equivalent of DMAP and 2 equivalents of DIPEA were added. The mixture was refluxed for 6-18 h. The reaction mixture was concentrated *in vacuo*.

3-(Benzylthio)-1*H*-1,2,4-triazole (**27**)
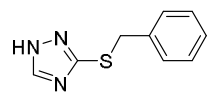
 1*H*-1,2,4-triazole-3-thiol (1.01 g, 10.0 mmol) and benzylbromide (1.09 mL, 10.0 mmol) were dissolved in DMF (10 mL) and stirred for 18 h. The reaction mixture was then diluted with EtOAc (50 mL) and washed with aqueous saturated $NaHCO_3$ (50 mL). The organic phase was separated and the aqueous phase was extracted with EtOAc (50 mL). The combined organic phases were washed with water (50 mL) and brine (50 mL). The volatiles are removed *in vacuo* and the product was obtained as white solid (1.86 g, 9.75 mmol, 98%). ^{1}H NMR (400 MHz, MeOD) δ 8.30 (s, 1H), 7.35 – 7.14 (m, 5H), 4.33 (s, 2H). ^{13}C NMR (101 MHz, MeOD) δ 147.99, 138.64, 129.91, 129.52, 128.48, 37.80.

3-(Benzylsulfonyl)-1*H*-1,2,4-triazole (**28**)
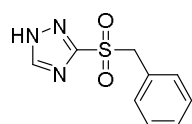
 27 (2.91 g, 15.2 mmol) was dissolved in DCM (40 mL) and peracetic acid (8.9 mL, 47 mmol) was slowly added. The reaction was stirred for 48 h at RT during which a white precipitate formed. The precipitate was filtered off and taken up in EtOAc (50 mL). This was washed with 1:1 saturated aqueous $Na_2S_2O_3$ and Na_2CO_3 (50 mL). The aqueous phase was extracted with EtOAc (2x 50 mL). Combined organic layers were dried ($MgSO_4$), filtered and concentrated to afford the title compound (2.60 g, 11.6 mmol, 76%). ^{1}H NMR (400 MHz, MeOD) δ 8.67 (s, 1H), 7.48 – 7.11 (m, 5H), 4.70 (s, 2H). ^{13}C NMR (101 MHz, MeOD) δ 161.51, 146.94, 132.20, 129.90, 129.58, 128.47, 61.59.
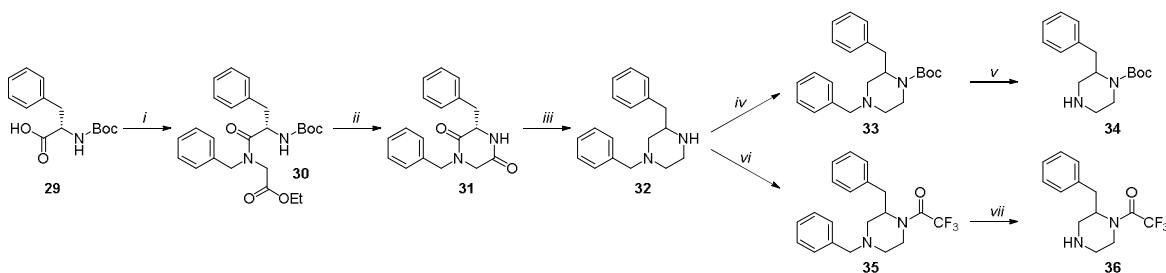
Scheme S3.2 | Reagents and conditions: *i*) HBTU, DIPEA, ethyl benzylglycinate, DCM, RT, 99%; *ii*) a: TFA, DCM, b: DIPEA, MeOH, RT, 82%; *iii*) LiAlH₄, THF, 65 °C, 90%; *iv*) Boc₂O, DCM, RT, 95%; *v*) ammonium formate, Pd/C, EtOH, 75 °C, 100%; *vi*) (CF₃CO)₂O, NEt₃, Et₂O, 100%; *vii*) ammonium formate, Pd/C, EtOH, 75 °C, 95%.

Ethyl *N*-benzyl-*N*-((*tert*-butoxycarbonyl)-L-phenylalanyl)glycinate (**30**)
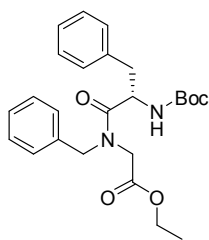
(*tert*-Butoxycarbonyl)-L-phenylalanine **29** (4.38 g, 16.5 mmol) and DIPEA (3.41 mL, 19.5 mmol) were dissolved in DCM (200 mL) on ice. HBTU (6.26 g, 16.5 mmol) was added and the mixture was stirred for 10 minutes. Ethyl benzylglycinate (2.76 mL, 15.0 mmol) was added and the mixture was stirred for 18 h at RT. The reaction was quenched with sat. aq. NH₄Cl (100 mL) and the organic phase was separated. The aqueous layer was extracted with EtOAc (100 mL). The combined organic layers were dried (MgSO₄), filtered and concentrated. The residue was purified by column chromatography to yield the title compound as a yellow oil (6.61 g, 15.0 mmol, 99%). ¹H NMR (400 MHz, CDCl₃) δ 7.47 – 6.97 (m, 10H), 5.37 (dd, *J* = 22.3, 8.9 Hz, 1H), 5.03 – 4.90 (m, 1H), 4.78 – 4.43 (m, 2H), 4.23 – 4.07 (m, 2H), 4.02 (dd, *J* = 20.2, 17.8 Hz, 1H), 3.78 (dd, *J* = 40.3, 17.8 Hz, 1H), 3.19 – 3.02 (m, 1H), 3.02 – 2.87 (m, 1H), 1.39 (s, 9H), 1.30 – 1.20 (m, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 172.89, 168.83, 155.03, 136.65, 136.46, 135.98, 135.41, 129.73, 129.58, 129.54, 128.95, 128.66, 128.52, 128.42, 128.01, 127.72, 127.35, 126.89, 126.86, 79.83, 61.73, 61.24, 51.98, 51.63, 50.11, 48.32, 47.12, 39.61, 28.34, 14.20 (carbon spectrum shows a mixture of rotamers).

(S)-1,3-Dibenzylpiperazine-2,5-dione (**31**)
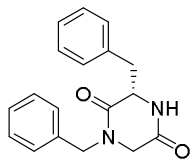
30 (6.61 g, 15.0 mmol) was dissolved in DCM (100 mL) and TFA (20 mL) was added. The mixture was stirred for 3 h at RT. Volatiles were removed under reduced pressure and the residue was taken up in methanol (50 mL). DIPEA (26.2 mL, 150 mmol) was added and the mixture stirred for 18 h. The mixture was concentrated *in vacuo* and purified by column chromatography. Product was isolated as a white gum (3.62 g, 12.3 mmol, 82%). ¹H NMR (400 MHz, CDCl₃) δ 7.38 – 7.26 (m, 3H), 7.26 – 7.18 (m, 1H), 7.18 – 7.06 (m, 7H), 4.41 (t, *J* = 3.7 Hz, 1H), 4.36 (t, *J* = 4.3 Hz, 1H), 4.22 (s, 1H), 3.38 (dd, *J* = 17.6, 2.8 Hz, 1H), 3.28 (dt, *J* = 13.7, 3.6 Hz, 1H), 3.02 (dt, *J* = 13.7, 3.7 Hz, 1H), 2.61 (dd, *J* = 17.6, 2.9 Hz, 1H). ¹³C NMR (101 MHz, CDCl₃) δ 169.56, 169.05, 137.52, 137.50, 133.19, 131.83, 131.70, 131.58, 131.17, 130.47, 59.39, 52.73, 51.19, 43.44.

1,3-Dibenzylpiperazine (**32**)
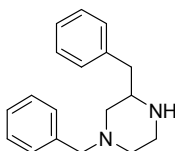
To a 0 °C suspension of **31** (3.61 g, 12.26 mmol) in dry THF (250 mL) was added in batches 2 M LiAlH₄ in THF (30.7 mL, 61.3 mmol). The mixture was stirred on ice for 15 min after which the reaction was heated and refluxed for 4 h. The solution was cooled and the excess LiAlH₄ was quenched by the careful addition of 2 M NaOH (8 mL). The cloudy mixture was stirred ON. Solids that formed were filtered off and the filtrate concentrated to provide 1,3-dibenzylpiperazine (2.95 g, 11.1 mmol, 90%) as a pale coloured oil. ¹H NMR (400 MHz, MeOD) δ 7.41 – 7.09 (m, 10H), 3.56 – 3.38 (m, 2H), 3.02 – 2.90 (m, 1H), 2.86 (dt, *J* = 12.7, 3.0 Hz, 1H), 2.81 – 2.68 (m, 3H), 2.65 – 2.59 (m, 2H), 2.07 (td, *J* = 11.5, 3.2 Hz, 1H), 1.84 (t, *J* = 10.6 Hz, 1H). ¹³C NMR (101 MHz, MeOD) δ 139.36, 138.11, 130.70, 130.21, 129.56, 129.26, 128.39, 127.51, 64.22, 59.67, 57.28, 53.79, 45.98, 41.36.

tert-Butyl 2,4-dibenzylpiperazine-1-carboxylate (**33**)
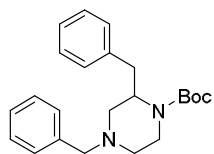
32 (0.70 g, 2.6 mmol) was dissolved in DCM (30 mL) and Boc-anhydride (0.64 mL, 2.8 mmol) was added. The mixture was stirred for 4 h. Volatiles were evaporated under reduced pressure and the resulting residue was purified by column chromatography yielding the title compound as a colourless oil (0.915 g, 2.50 mmol, 95%). ^{1}H NMR (300 MHz, CDCl$_3$) δ 7.44 – 6.95 (m, 10H), 4.16 (s, 1H), 3.94 (s, 1H), 3.55 (d, J = 12.9 Hz, 1H), 3.38 (d, J = 12.9 Hz, 1H), 3.22 (td, J = 12.7, 3.4 Hz, 1H), 3.08 (t, J = 11.0 Hz, 1H), 2.86 (d, J = 11.1 Hz, 2H), 2.65 (dt, J = 11.4, 1.9 Hz, 1H), 2.09 (td, J = 11.7, 3.5 Hz, 1H), 1.99 (dd, J = 11.5, 3.9 Hz, 1H), 1.40 (s, 9H). ^{13}C NMR (75 MHz, CDCl$_3$) δ 138.46, 129.36, 128.40, 127.29, 126.13, 79.63, 63.10, 53.44, 28.49.

tert-Butyl 2-benzylpiperazine-1-carboxylate (**34**)
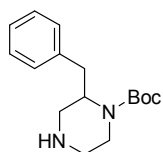
33 (964 mg, 2.63 mmol) was dissolved in absolute EtOH and ammonium formate (829 mg, 13.2 mmol) was added. The mixture was brought under N$_2$-atmosphere and 10% Pd on carbon (50 mg) was added. The mixture was refluxed for 48 h. The catalyst was filtered off and the volatiles were evaporated. Coevaporation with toluene afforded the pure product in quantitative yield (727 mg, 2.63 mmol, 100%). ^{1}H NMR (300 MHz, MeOD) δ 7.47 – 7.12 (m, 5H), 4.46 – 4.12 (m, 1H), 4.12 – 3.72 (m, 1H), 3.27 – 2.53 (m, 7H), 1.34 (s, 9H). ^{13}C NMR (75 MHz, CDCl$_3$) δ 156.01, 140.16, 130.37, 129.37, 127.22, 80.68, 46.25, 35.86, 28.82, 28.54.

1-(2,4-Dibenzylpiperazin-1-yl)-2,2,2-trifluoroethan-1-one (**35**)
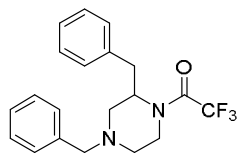
32 (2.01 g, 7.51 mmol) was dissolved in diethyl ether (30 mL) and triethylamine (1.15 mL, 8.26 mmol) was added. The mixture was cooled to 0 °C and 2,2,2-trifluoroacetic anhydride (1.15 mL, 8.26 mmol) was added slowly. The mixture was stirred for 3 h at RT. The reaction was quenched by the addition of saturated aqueous Na$_2$CO$_3$ (20 mL). The organic phase was separated and the water layer extracted with Et$_2$O (2x 25 mL). The combined organic layers were dried (MgSO$_4$), filtered and concentrated to yield **35** (2.72 g, 7.51 mmol, 100%) as a yellow oil, which was used without further purification for the next reaction. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.47 – 6.81 (m, 10H), 4.68 – 4.29 (m, 1H), 4.13 – 3.90 (m, 1H), 3.81 – 3.11 (m, 4H), 3.11 – 2.61 (m, 3H), 2.19 – 1.84 (m, 2H) (proton spectrum shows a mixture of rotamers). ^{13}C NMR (101 MHz, CDCl$_3$) δ 155.55 (q, J = 35.5 Hz), 155.35 (q, J = 35.5 Hz), 137.59, 137.43, 129.42, 129.32, 129.22, 128.55, 128.35, 127.45, 127.42, 126.65, 126.47, 116.69 (q, J = 288.1 Hz), 116.45 (q, J = 288.0 Hz), 62.51 (q, J = 7.0 Hz), 56.08 (q, J = 3.1 Hz), 52.74, 45.60, 41.90 (q, J = 3.7 Hz), 39.14, 36.67, 35.51, 29.69 (carbon spectrum shows a mixture of rotamers).

1-(2-Benzylpiperazin-1-yl)-2,2,2-trifluoroethan-1-one (**36**)
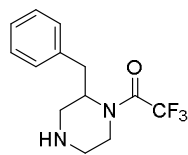
35 (1.09 g, 3.00 mmol) was dissolved in absolute ethanol (30 mL) and ammonium formate (0.95 g, 15 mmol) was added. The mixture was brought under N$_2$-atmosphere and 10% Pd/C catalyst (50 mg) was added. The mixture was refluxed for 70 h. Catalyst was filtered off and the solvent removed *in vacuo*. The residue was taken up in EtOAc (25 mL) and washed with saturated aqueous Na$_2$CO$_3$. The layers were separated and the aqueous phase extracted twice with EtOAc (25 mL). The combined organic layers were dried (MgSO$_4$), filtered and concentrated to yield the title compound (0.773 g, 2.84 mmol, 95%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.45 – 7.05 (m, 5H), 4.74 – 3.96 (m, 1H), 3.79 – 3.58 (m, 1H), 3.58 – 3.39 (m, 1H), 3.28 – 2.47 (m, 7H) (proton spectrum shows a mixture of rotamers). ^{13}C NMR (101 MHz, CDCl$_3$) δ 155.80 (d, J = 35.4 Hz), 155.61 (d, J = 35.4 Hz), 137.50, 137.44, 129.31, 129.28, 128.75, 128.52, 126.89, 126.68, 116.70 (q, J = 288.1 Hz), 116.45 (q, J = 288.2 Hz), 55.92, 52.16, 46.31 (d, J = 45.7 Hz), 46.14 (d, J = 55.0 Hz), 42.47 (q, J = 3.6 Hz), 39.50, 35.66, 34.64 (carbon spectrum shows a mixture of rotamers).

tert-Butyl 3-benzyl-4-(2,2,2-trifluoroacetyl)piperazine-1-carboxylate (**37**)
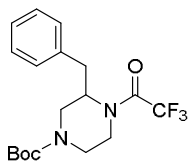
36 (192 mg, 0.705 mmol) was dissolved in DCM (10 mL) and Boc-anhydride (0.180 mL, 0.776 mmol) was added. The mixture was stirred for 18 h at RT. The solvents were removed under reduced pressure. The crude residue (263 mg, 0.706 mmol, quant.) was used without further purification in the next reaction.

tert-Butyl 3-benzylpiperazine-1-carboxylate (**38**)
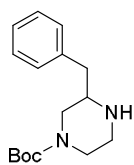
tert-Butyl 3-benzylpiperazine-1-carboxylate was synthesized from **37** (263 mg, 0.706 mmol) according to General Procedure 2. The crude residue (195 mg, 0.706 mmol, quant.) was of sufficient purity to be used without further purification in the next reaction. ¹H NMR (400 MHz, CDCl₃) δ 7.41 – 7.13 (m, 5H), 4.27 – 3.77 (m, 2H), 3.42 (bs, 1H), 3.04 – 2.73 (m, 4H), 2.71 – 2.47 (m, 3H), 1.45 (s, 9H). ¹³C NMR (101 MHz, CDCl₃) δ 154.65, 137.49, 129.22, 128.70, 126.73, 79.84, 56.16, 45.36, 39.78, 28.41.

tert-Butyl 4-acetyl-2-benzylpiperazine-1-carboxylate (**39**)
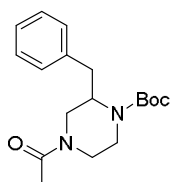
Acetylchloride (0.20 mL, 2.9 mmol), Net₃ (0.5 mL, 3.4 mmol) and **34** (720 mg, 2.6 mmol) were dissolved in DCM (20 mL) and stirred at 0 °C for 2 hours. The solution was washed with water (20 mL) and the water layer was extracted with EtOAc (2x 20 mL). The combined organic layers were dried (MgSO₄), filtered and concentrated *in vacuo*. The product was obtained as a yellow oil in quantitative yield (0.83 g, 2.6 mmol, 100%). ¹H NMR (300 MHz, CDCl₃) δ 7.21 – 6.86 (m, 5H), 4.41 – 3.99 (m, 2H), 3.92 – 3.47 (m, 2H), 3.09 – 2.82 (m, 2H), 2.81 – 2.40 (m, 3H), 1.83 (s, 3H), 1.18 (s, 9H). ¹³C NMR (75 MHz, CDCl₃) δ 169.19, 168.85, 153.78, 137.65, 137.26, 129.04, 128.63, 128.13, 127.86, 126.18, 125.85, 79.41, 59.66, 47.12, 45.49, 42.22, 40.67, 35.48, 34.93, 27.69, 20.85, 20.60 (carbon spectrum shows a mixture of rotamers).

1-(3-Benzylpiperazin-1-yl)ethan-1-one hydrochloride (**40**)

1-(3-Benzylpiperazin-1-yl)ethan-1-one was synthesized from **39** (490 mg, 1.54 mmol) according to General Procedure 1. The crude residue (370 mg, 1.45 mmol, 94%) was of sufficient purity to be used without further purification in the next reaction, as judged by LC/MS analysis.

tert-Butyl (S)-2-benzyl-4-(cyclobutanecarbonyl)piperazine-1-carboxylate (**41**)
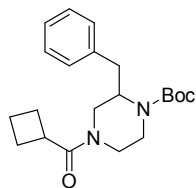
Cyclobutanecarboxylic acid (38.8 mg, 0.388 mmol), DIPEA (77 μL, 0.44 mmol) and HBTU (154 mg, 0.406 mmol) were dissolved in DCM (5 mL) and stirred for 20 minutes at RT. **34** (102 mg, 0.369 mmol) was then added as a solution in DCM (1 mL). The mixture was stirred for 20 h. Volatiles were removed *in vacuo* and the residue was purified by column chromatograpy. The title compound was obtained as a white solid (106 mg, 0.296 mmol, 80%). ¹H NMR (400 MHz, CDCl₃) δ 7.38 – 7.10 (m, 5H), 4.55 – 4.20 (m, 2H), 3.99 (s, 1H), 3.71 – 3.24 (m, 1H), 3.23 – 2.82 (m, 3H), 2.82 – 2.60 (m, 2H), 2.57 – 2.32 (m, 2H), 2.32 – 2.12 (m, 2H), 2.12 – 1.97 (m, 1H), 1.97 – 1.75 (m, 2H), 1.33 (s, 9H). ¹³C NMR (101 MHz, CDCl₃) δ 174.26, 173.81, 154.49, 154.33, 138.20, 137.74, 129.60, 129.19, 128.65, 128.50, 126.76, 126.46, 80.23, 44.82, 43.27, 41.65, 37.33, 36.88, 36.04, 29.76, 28.24, 25.61, 24.95, 18.11 (carbon spectrum shows a mixture of rotamers).

(3-Benzylpiperazin-1-yl)(cyclobutyl)methanone hydrochloride (**42**)

(3-Benzylpiperazin-1-yl)(cyclobutyl)methanone was synthesized from **41** (106 mg, 0.30 mmol) according to General Procedure 1. The crude residue (87 mg, 0.30 mmol, quant.) was of sufficient purity to be used without further purification in the next reaction, as judged by LC/MS analysis.

tert-Butyl 2-benzyl-4-(3,3-dimethylbutanoyl)piperazine-1-carboxylate (**43**)
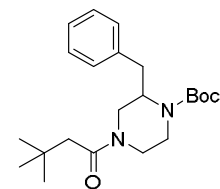
3,3-Dimethylbutanoic acid (45.0 mg, 0.388 mmol), DIPEA (77 μL, 0.44 mmol) and HBTU (154 mg, 0.406 mmol) were dissolved in DCM (5 mL) and stirred for 20 minutes at RT. **34** (102 mg, 0.369 mmol) was then added as a solution in DCM (1 mL). The mixture is stirred for 20 h. Volatiles were removed *in vacuo* and the residue was purified by column chromatograpy. The title compound was obtained as a white solid (67 mg, 0.18 mmol, 49%). ¹H NMR (400 MHz, DMSO, 353 K) δ 7.37 – 7.03 (m, 5H), 4.24 (s, 2H), 4.08 – 3.59 (m, 2H), 3.37 – 3.11 (m, 2H), 2.93 – 2.54 (m, 3H), 2.30 (s, 2H), 1.31 (s, 9H), 1.03 (s, 9H). ¹³C NMR (101 MHz, DMSO, 293 K) δ 169.88, 153.28, 137.94, 128.79, 127.77, 125.74, 78.59, 43.43, 30.44, 29.34, 27.54.

1-(3-Benzylpiperazin-1-yl)-3,3-dimethylbutan-1-one hydrochloride (**44**)

1-(3-Benzylpiperazin-1-yl)-3,3-dimethylbutan-1-one was synthesized from **43** (67 mg, 0.18 mmol) according to General Procedure 1. The crude residue (56 mg, 0.18 mmol, quant.) was of sufficient purity to be used without further purification in the next reaction, as judged by LC/MS analysis.

tert-Butyl 4-benzoyl-2-benzylpiperazine-1-carboxylate (**45**)
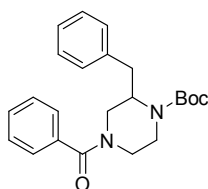
34 (120 mg, 0.434 mmol) was dissolved in DCM (5 mL) to which triethylamine (73 μL, 0.52 mmol) was added. Benzoyl chloride (55 μL, 0.48 mmol) was added and the mixture was stirred for 18 h. The reaction was quenched by the addition of 0.1 M aqueous HCl (5 mL). The organic layer was separated and the aqueous phase was extracted with DCM (2x 10 mL). The combined organic layers were dried (MgSO$_4$), filtered and concentrated. The residue was purified by column chromatography to yield the title compound (106 mg, 0.279 mmol, 64%) as white solid. ^{1}H NMR (400 MHz, CDCl$_3$, 323 K) δ 7.42 (m, 5H), 7.17 (m, 5H), 4.33 (s, 2H), 4.05 – 3.85 (m, 1H), 3.10 (m, 3H), 2.91 – 2.54 (m, 2H), 1.35 (s, 9H). ^{13}C NMR (101 MHz, CDCl$_3$, 323 K) δ 171.10, 154.42, 137.91, 135.67, 129.99, 129.40, 128.62, 128.50, 127.33, 126.53, 80.25, 53.48, 39.07, 36.10, 28.33 (carbon spectrum shows a mixture of rotamers).

(3-Benzylpiperazin-1-yl)(phenyl)methanone hydrochloride (**46**)

(3-Benzylpiperazin-1-yl)(phenyl)methanone was synthesized from **45** (106 mg, 0.279 mmol) according to General Procedure 1. The crude residue (88 mg, 0.28 mmol, quant.) was of sufficient purity to be used without further purification in the next reaction, as judged by LC/MS analysis.

tert-Butyl (*S*)-(1-(benzyl(2-hydroxyethyl)amino)-1-oxo-3-phenylpropan-2-yl)carbamate (**47**)
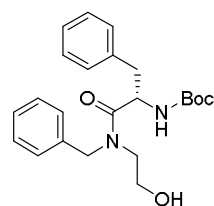
DIPEA (1.55 mL, 8.9 mmol) was added to a solution of Boc-L-phenylalanine (1.99 g, 7.5 mmol) in DCM (100 mL) on ice. HBTU (2.85 g, 7.5 mmol) was added and the mixture was stirred for 10 minutes. 2-Benzylaminoethanol (0.97 mL, 6.8 mmol) was added to the reaction mixture. The reaction mixture was stirred at RT for 24 h. The mixture was washed with 0.1 M HCl (aq., 100 mL) The water layer was extracted with EtOAc (3x 50 mL). The organic layers were combined and washed with 1 M NaOH (aq., 50 mL) and brine (50 mL). The organic layer was dried (MgSO$_4$), filtered and concentrated under reduced pressure. The residue was purified by column chromatography affording the title compound as a yellow oil (1.61 g, 4.0 mmol, 59%).^{1}H NMR (300 MHz, CDCl$_3$) δ 7.41 – 6.89 (m, 11H), 5.36 (dd, *J* = 40.4, 8.1 Hz, 1H), 5.07 – 4.74 (m, 1H), 4.34 (dd, *J* = 48.6, 16.0 Hz, 1H), 3.73 – 3.23 (m, 2H), 3.19 – 2.70 (m, 4H), 1.64 (s, 1H), 1.41 (s, 9H) (proton spectrum shows a mixture of rotamers).

tert-Butyl (*S*)-(1-(benzyl(2-chloroethyl)amino)-1-oxo-3-phenylpropan-2-yl)carbamate (**48**)
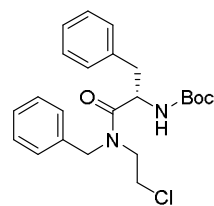
Triethylamine (0.63 mL, 5.1 mmol) was added to a solution of **47** (0.88 g, 2.2 mmol) in DCM (20 mL) on ice. Methanesulfonylchloride (0.18 mL, 2.4 mmol) was added and the mixture was refluxed for 5 h. The mixture was washed with saturated aqueous NaHCO$_3$ (20 mL). The water layer was extracted with EtOAc (20 mL). The combined organic layers were washed with brine (20 mL), dried (MgSO$_4$), filtered and concentrated *in vacuo*. The product was purified by column chromatography yielding the title compound (0.64 g, 1.5 mmol, 70%). ^{1}H NMR (300 MHz, CDCl$_3$) δ 7.36 – 6.93 (m, 10H), 5.62 (t, *J* = 9.9 Hz, 1H), 4.99 – 4.74 (m, 1H), 4.74 – 4.29 (m, 2H), 3.79 – 3.23 (m, 4H), 3.21 – 2.86 (m, 2H), 1.46 (s, 9H). ^{13}C NMR (75 MHz, CDCl$_3$) δ 172.75, 172.64, 155.36, 155.07, 136.72, 136.58, 136.11, 129.65, 129.60, 128.99, 128.70, 128.64, 128.02, 127.93, 127.63, 127.07, 127.00, 79.91, 79.75, 52.44, 51.82, 49.44, 48.39, 48.26, 40.89, 39.97, 28.40 (carbon spectrum shows a mixture of rotamers).

1,3-Dibenzylpiperazin-2-one (**49**)
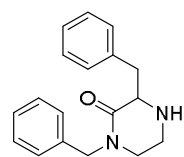
48 (640 mg, 1.53 mmol) was dissolved in 1 M HCl in MeOH (20 mL) and stirred for 19 h. The volatiles were removed *in vacuo* and the residue was taken up in dry DMF (20 mL) and brought under argon atmosphere. K$_2$CO$_3$ (300 mg, 2.17 mmol) and a catalytic amount of tetrabutylammonium iodide were added and the temperature was raised to 80 °C. The mixture was stirred for 24 h. The mixture was diluted with water (20 mL). The aqueous phase was extracted with EtOAc (3x 40 mL), the combined organic layers were washed with brine (40 mL), dried (MgSO$_4$), filtered and concentrated *in vacuo*. The residue was purified by column chromatography

affording the title compound as a white solid (150 mg, 0.54 mmol, 35%). ^{1}H NMR (300 MHz, MeOD) δ 7.67 – 7.02 (m, 10H), 4.62 (dt, *J* = 31.0, 7.0 Hz, 1H), 4.37 – 4.16 (m, 1H), 3.99 – 3.75 (m, 1H), 3.75 – 3.49 (m, 1H), 3.49 – 3.27 (m, 3H), 3.24 – 3.03 (m, 2H).

3-Benzyl-4-(2,2,2-trifluoroacetyl)piperazine-1-carboxamide (**50**)
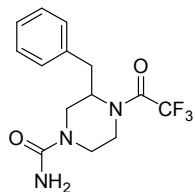
36 (156 mg, 0.573 mmol) and potassium cyanate (93 mg, 1.15 mmol) were dissolved in H_2O/AcOH (1:1, 6 mL) and stirred for 18 h at RT. The reaction was quenched by the addition of sat. aq. Na_2CO_3. The mixture was extracted with EtOAc (3x 20 mL). The combined organic layers were dried ($MgSO_4$), filtered and concentrated. The residue was purified by column chromatography to yield the title compound as a white solid (91 mg, 0.29 mmol, 50%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.42 – 7.13 (m, 5H), 5.24 (s, 2H), 4.81 – 4.66 (m, 1H), 4.51 – 4.01 (m, 2H), 3.80 (t, *J* = 13.4 Hz, 1H), 3.66 (dt, *J* = 13.6, 2.1 Hz, 1H), 3.50 (ddd, *J* = 14.3, 11.8, 3.5 Hz, 1H), 3.27 – 3.13 (m, 1H), 3.10 (dd, *J* = 13.7, 4.0 Hz, 1H), 3.04 – 2.80 (m, 4H) (proton spectrum shows a mixture of rotamers). ^{13}C NMR (101 MHz, $CDCl_3$) δ 158.67, 156.00 (q, *J* = 35.8 Hz), 155.71 (q, *J* = 36.0 Hz), 136.46 (d, *J* = 3.2 Hz), 129.39, 129.07, 128.86, 127.42, 127.20, 116.52 (q, *J* = 287.9 Hz), 116.25 (q, *J* = 288.0 Hz), 55.68 (d, *J* = 3.5 Hz), 53.53, 52.52, 44.66, 44.43, 43.16 (d, *J* = 4.1 Hz), 41.23 (q, *J* = 3.3 Hz), 38.37, 36.23, 35.20 (carbon spectrum shows a mixture of rotamers).

3-Benzylpiperazine-1-carboxamide (**51**)

3-Benzylpiperazine-1-carboxamide was synthesized from **50** (91 mg, 0.29 mmol) according to General Procedure 2. The crude residue (63 mg, 0.29 mmol, quant.) was of sufficient purity to be used without further purification in the next reaction, as judged by LC/MS analysis.

3-Benzyl-*N*-ethyl-4-(2,2,2-trifluoroacetyl)piperazine-1-carboxamide (**52**)
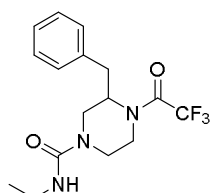
36 (112 mg, 0.411 mmol) was dissolved in DCM (5 mL) and isocyanatoethane (39 μL, 0.49 mmol) was added. The mixture was stirred for 18 h at RT. The reaction was quenched by the addition of 0.1 M HCl (aq., 10 mL). The organic layer was separated and the aqueous phase was extracted with DCM (2x 10 mL). The combined organic layers were dried ($MgSO_4$), filtered and concentrated to yield the title compound (141 mg, 0.411 mmol, 100%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.44 – 7.15 (m, 5H), 4.81 – 4.40 (m, 2H), 4.24 – 4.07 (m, 1H), 3.98 – 3.66 (m, 1H), 3.64 – 3.42 (m, 1H), 3.42 – 3.12 (m, 3H), 3.08 – 2.78 (m, 4H), 1.16 (td, *J* = 7.2, 3.1 Hz, 3H) (proton spectrum shows a mixture of rotamers). ^{13}C NMR (101 MHz, $CDCl_3$) δ 157.47, 156.00 (q, *J* = 35.8 Hz), 136.70, 129.40, 129.06, 128.86, 127.41, 127.20, 116.57 (q, *J* = 288.1 Hz), 116.30 (q, *J* = 288.0 Hz), 55.70 (q, *J* = 3.0 Hz), 53.53, 52.54, 44.56, 44.19, 43.11, 43.00, 41.42 (q, *J* = 3.3 Hz), 38.54, 36.21, 35.91, 35.88, 35.09, 30.34, 29.74, 15.48 (carbon spectrum shows a mixture of rotamers).

3-Benzyl-*N*-ethylpiperazine-1-carboxamide (**53**)

3-Benzyl-*N*-ethylpiperazine-1-carboxamide was synthesized from **52** (141 mg, 0.41 mmol) according to General Procedure 2. The crude residue (102 mg, 0.41 mmol, quant.) was of sufficient purity to be used without further purification in the next reaction, as judged by LC/MS analysis.

3-Benzyl-*N*-(*tert*-butyl)-4-(2,2,2-trifluoroacetyl)piperazine-1-carboxamide (**54**)
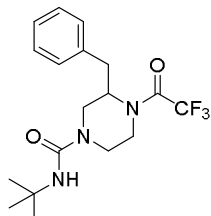
36 (114 mg, 0.419 mmol) was dissolved in DCM (6 mL) and isocyanatoethane (70 μL, 0.61 mmol) was added. The mixture was stirred for 18 h at RT. The reaction was quenched by the addition of 0.1 M HCl (aq., 10 mL). The organic layer was separated and the aqueous phase was extracted with DCM (2x 10 mL). The combined organic layers were dried ($MgSO_4$), filtered and concentrated to yield the title compound (156 mg, 0.419 mmol, 100%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.47 – 7.12 (m, 5H), 4.75 – 4.64 (m, 1H), 4.47 – 4.09 (m, 2H), 3.82 (d, *J* = 14.0 Hz, 1H), 3.61 – 3.37 (m, 2H), 3.26 – 3.15 (m, 1H), 3.08 – 2.94 (m, 1H), 2.92 – 2.70 (m, 2H), 1.35 (s, 9H) (proton spectrum shows a mixture of rotamers). ^{13}C NMR (101 MHz, $CDCl_3$) δ 156.52, 156.48, 155.91 (q, *J* = 35.9 Hz), 136.75, 136.68, 129.32, 129.26, 129.07, 128.85, 127.38, 127.14, 116.53 (q, *J* = 287.8 Hz), 116.25 (q, *J* = 288.0 Hz), 55.50 (q, *J* = 2.6 Hz), 52.52, 51.02, 44.88, 44.36, 42.74, 42.60, 41.41 (q, *J* = 3.0 Hz), 38.56, 36.06, 34.91, 31.83, 30.26, 29.28 (carbon spectrum shows a mixture of rotamers).

3-Benzyl-*N*-(*tert*-butyl)piperazine-1-carboxamide (**55**)

3-Benzyl-*N*-(*tert*-butyl)piperazine-1-carboxamide was synthesized from **54** (156 mg, 0.419 mmol) according to General Procedure 2. The crude residue (119 mg, 0.419 mmol, quant.) was of sufficient purity to be used without further purification in the next reaction, as judged by LC/MS analysis.

1-(2-Benzyl-4-isopropylpiperazin-1-yl)-2,2,2-trifluoroethan-1-one (**56**)
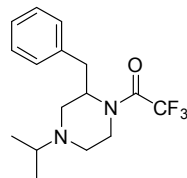
36 (194 mg, 0.71 mmol) and acetone (78 µL, 1.1 mmol) were dissolved in DCM (5 mL). The mixture was cooled to 0 °C and sodium triacetoxyborohydride (166 mg, 0.78 mmol) and acetic acid (49 µL, 0.86 mmol) were added. The mixture was stirred for 48 h at RT. The reaction was quenched by the addition of sat. aq. Na_2CO_3 (10 mL). The organic layer was separated and the aqueous phase was extracted with DCM (2x 10 mL). The combined organic layers were dried ($MgSO_4$), filtered and concentrated. The residue was purified by column chromatography to yield the title compound (120 mg, 0.382 mmol, 54%).^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36 – 7.12 (m, 5H), 4.74 – 4.28 (m, 1H), 4.14 – 3.60 (m, 1H), 3.59 – 3.40 (m, 1H), 3.31 – 3.13 (m, 1H), 2.94 – 2.57 (m, 4H), 2.38 – 2.22 (m, 1H), 2.22 – 2.05 (m, 1H), 1.06 – 0.91 (m, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 155.66 (q, *J* = 35.4 Hz), 155.45 (q, *J* = 35.4 Hz), 138.06, 137.92, 129.55, 129.44, 128.74, 128.54, 126.89, 126.65, 116.84 (q, *J* = 288.1 Hz), 116.58 (q, *J* = 288.2 Hz), 56.20 (d, *J* = 3.0 Hz), 54.54, 54.49, 52.83, 49.20, 48.77, 48.56, 48.32, 42.46 (d, *J* = 3.8 Hz), 39.62, 36.54, 35.40, 18.76, 18.70, 17.67, 17.58 (carbon spectrum shows a mixture of rotamers).

3-Benzyl-1-isopropylpiperazine (**57**)

3-Benzyl-1-isopropylpiperazine was synthesized from **56** (120 mg, 0.38 mmol) according to General Procedure 2. The crude residue (83 mg, 0.38 mmol, quant.) was of sufficient purity to be used without further purification in the next reaction, as judged by LC/MS analysis.

### *tert*-Butyl 2-benzyl-4-isobutylpiperazine-1-carboxylate (**58**)
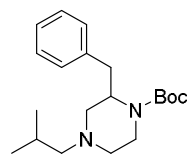
34 (53 mg, 0.19 mmol) and isobutyraldehyde (53 µL, 0.58 mmol) were dissolved in DCM (5 mL). The mixture was cooled to 0 °C and sodium triacetoxyborohydride (81 mg, 0.38 mmol) and acetic acid (24 µL, 0.42 mmol) were added. The mixture was stirred for 18 h at RT. The reaction was quenched by the addition of sat. aq. NaHCO₃ (10 mL). The organic layer was separated and the aqueous phase was extracted with DCM (2x 10 mL). The combined organic layers were dried ($MgSO_4$), filtered and concentrated. The residue was purified by column chromatography to yield the title compound (58 mg, 0.17 mmol, 91%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.40 – 7.10 (m, 5H), 4.17 (s, 1H), 3.91 (s, 1H), 3.18 (td, *J* = 12.7, 3.5 Hz, 1H), 3.07 (d, *J* = 11.4 Hz, 1H), 2.97 – 2.69 (m, 2H), 2.60 (dt, *J* = 11.6, 1.9 Hz, 1H), 2.04 (dd, *J* = 12.0, 7.6 Hz, 1H), 2.00 – 1.87 (m, 3H), 1.72 (m, 1H), 1.39 (s, 9H), 0.94 (m, 6H) (proton spectrum shows a mixture of rotamers). ^{13}C NMR (101 MHz, $CDCl_3$) δ 154.83, 139.66, 129.52, 128.44, 126.18, 79.50, 66.99, 55.11, 53.74, 36.42, 28.46, 25.58, 21.01.

3-Benzyl-1-isobutylpiperazine hydrochloride (**59**)

3-Benzyl-1-isobutylpiperazine was synthesized from **58** (58 mg, 0.17 mmol) according to General Procedure 1. The crude residue (47 mg, 0.17 mmol, quant.) was of sufficient purity to be used without further purification in the next reaction, as judged by LC/MS analysis.

tert-Butyl 2-benzyl-4-cyclopentylpiperazine-1-carboxylate (**60**)
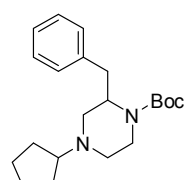
34 (50 mg, 0.18 mmol) and cyclopentanone (80 µL, 0.91 mmol) were dissolved in DCM (5 mL). The mixture was cooled to 0 °C and sodium triacetoxyborohydride (77 mg, 0.36 mmol) and acetic acid (23 µL, 0.40 mmol) were added. The mixture was stirred for 18 h at RT. The reaction was quenched by the addition of sat. aq. NaHCO₃ (10 mL). The organic layer was separated and the aqueous phase was extracted with DCM (2x 10 mL). The combined organic layers were dried ($MgSO_4$), filtered and concentrated. The residue was purified by column chromatography to yield the title compound (54 mg, 0.16 mmol, 87%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37 – 7.12 (m, 5H), 4.16 (s, 1H), 3.91 (s, 1H), 3.24 – 3.01 (m, 2H), 3.00 – 2.80 (m, 2H), 2.74 (d, *J* = 11.5 Hz, 1H), 2.52 – 2.33 (m, 1H), 1.95 (qd, *J* = 12.8, 12.2, 3.6 Hz, 2H), 1.83 (dtd, *J* = 11.5, 7.2, 2.8 Hz, 1H), 1.77 – 1.60 (m, 2H), 1.61 – 1.46 (m, 2H), 1.39 (s, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 154.78, 139.65, 129.57, 128.45, 126.22, 79.55, 67.36, 53.18, 52.38, 36.42, 30.54, 30.29, 28.48, 24.15.

3-Benzyl-1-cyclopentylpiperazine hydrochloride (**61**)

3-Benzyl-1-cyclopentylpiperazine was synthesized from **60** (54 mg, 0.16 mmol) according to General Procedure 1. The crude residue (44 mg, 0.16 mmol, quant.) was of sufficient purity to be used without further purification in the next reaction, as judged by LC/MS analysis.

tert-Butyl (*S*)-2-benzyl-4-(4-fluorobenzyl)piperazine-1-carboxylate (**62**)
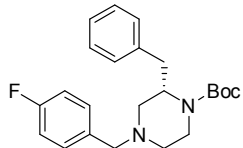
Commercial *tert*-butyl (S)-3-benzylpiperazine-1-carboxylate (50 mg, 0.18 mmol) and 4-fluorobenzaldehyde (21.4 µL, 0.199 mmol) were dissolved in DCM (5 mL). The mixture was cooled to 0 °C and sodium triacetoxyborohydride (46 mg, 0.22 mmol) and acetic acid (16 µL, 0.22 mmol) were added. The mixture was stirred for 18 h at RT. The reaction was quenched by the addition of sat. aq. Na_2CO_3 (10 mL). The organic layer was separated and the aqueous phase was extracted with DCM (2x 10 mL). The combined organic layers were dried ($MgSO_4$), filtered and concentrated. The residue was purified by column chromatography to yield the title compound (30 mg, 0.08 mmol, 43%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.49 – 6.91 (m, 9H), 4.13 (bs, 1H), 3.82 (bs, 1H), 3.51 (d, *J* = 13.0 Hz, 1H), 3.31 (d, *J* = 12.9 Hz, 1H), 3.20 (td, *J* = 12.8, 3.4 Hz, 1H), 3.12 – 2.94 (m, 1H), 2.93 – 2.74 (m, 2H), 2.60 (dt, *J* = 11.4, 1.9 Hz, 1H), 2.12 – 2.01 (m, 1H), 1.95 (dd, *J* = 11.5, 3.9 Hz, 1H), 1.38 (s, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 162.13 (d, *J* = 245.0 Hz), 154.72, 139.20, 133.91, 130.76 (d, *J* = 7.9 Hz), 129.37, 128.31, 126.11, 115.06 (d, *J* = 21.3 Hz), 79.62, 62.13, 53.27, 28.35.

(*S*)-3-Benzyl-1-(4-fluorobenzyl)piperazine 2,2,2-trifluoroacetate (**63**)

62 (30 mg, 0.08 mmol) was dissolved in DCM (5 mL) and cooled to 0 °C. 2,2,2-trifluoroacetic acid (1 mL, 13 mmol) was added and the mixture was stirred for 18 h at RT. The mixture was concentrated *in vacuo*. The residue was coevaporated with toluene (2x 5 mL). The residue was used without further purification in the next reactions.

tert-Butyl (*S*)-2-benzyl-4-(4-ethynylbenzyl)piperazine-1-carboxylate (**64**)
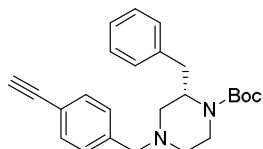
Commercial *tert*-butyl (S)-3-benzylpiperazine-1-carboxylate (50 mg, 0.18 mmol) and 4-ethynylbenzaldehyde (25.9 mg, 0.199 mmol) were dissolved in DCM (5 mL). The mixture was cooled to 0 °C and sodium triacetoxyborohydride (46 mg, 0.22 mmol) and acetic acid (16 µL, 0.22 mmol) were added. The mixture was stirred for 18 h at RT. The reaction was quenched by the addition of sat. aq. Na_2CO_3 (10 mL). The organic layer was separated and the aqueous phase was extracted with DCM (2x 10 mL). The combined organic layers were dried ($MgSO_4$), filtered and concentrated. The residue was purified by column chromatography to yield the title compound (43 mg, 0.11 mmol, 61%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.54 – 7.42 (m, 2H), 7.31 (d, *J* = 7.8 Hz, 2H), 7.24 – 6.92 (m, 5H), 4.37 – 3.67 (m, 2H), 3.52 (d, *J* = 13.1 Hz, 1H), 3.34 (d, *J* = 13.1 Hz, 1H), 3.20 (td, *J* = 12.5, 3.4 Hz, 1H), 3.08 (s, 1H), 3.01 (d, *J* = 12.4 Hz, 1H), 2.94 – 2.69 (m, 2H), 2.63 – 2.54 (m, 1H), 2.15 – 2.02 (m, 1H), 1.97 (dd, *J* = 11.6, 3.8 Hz, 1H), 1.38 (s, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.42, 132.18, 129.47, 129.27, 128.43, 126.21, 121.03, 83.72, 79.67, 62.73, 53.47, 29.83, 28.45.

(*S*)-3-Benzyl-1-(4-ethynylbenzyl)piperazine 2,2,2-trifluoroacetate (**65**)

64 (43 mg, 0.11 mmol) was dissolved in DCM (5 mL) and cooled to 0 °C. 2,2,2-trifluoroacetic acid (1 mL, 13 mmol) was added and the mixture was stirred for 18 h at RT. The mixture was concentrated *in vacuo*. The residue was coevaporated with toluene (2x 5 mL). The residue was used without further purification in the next reactions.

tert-Butyl (*S*)-3-benzyl-4-(3-(benzylsulfonyl)-1*H*-1,2,4-triazole-1-carbonyl)piperazine-1-carboxylate (**66**)
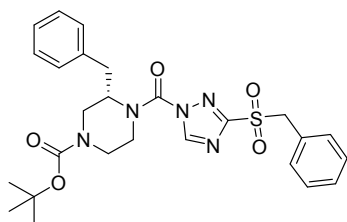
28 (178 mg, 0.796 mmol) was reacted with commercial *tert*-butyl (*S*)-3-benzylpiperazine-1-carboxylate (200 mg, 0.724 mmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (363 mg, 0.691 mmol, 95%). Spectral data in accordance with **3** (*vide infra*).

(S)-(2-Benzylpiperazin-1-yl)(3-(benzylsulfonyl)-1H-1,2,4-triazol-1-yl)methanone hydrochloride (67)
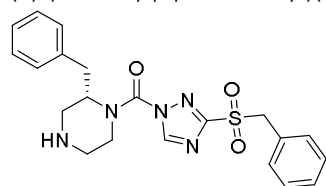
66 (363 mg, 0.691 mmol) was dissolved in diethylether (10 mL) and cooled to 0 °C. Hydrochloric acid (2 M in Et₂O, 1.7 mL, 3.4 mmol) was added and the mixture was stirred for 24 h at RT. The mixture was concentrated *in vacuo*. The residue was used without further purification in the next reactions.

3-(Benzylthio)-N,N-diphenyl-1H-1,2,4-triazole-1-carboxamide (68)
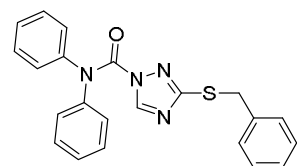
27 (2.137 g, 11.17 mmol) was dissolved in dry THF (40 mL) and diphenylcarbamic chloride (2.85 g, 12.3 mmol), DIPEA (2.34 mL, 13.4 mmol) and a catalytic amount of DMAP were added. The mixture was refluxed for 5 h and stirred 18 h at RT. The reaction was quenched by the addition of sat. aq. Na₂CO₃ (40 mL). The organic layer was separated and the aqueous phase extracted with EtOAc (2x 40 mL). The combined organic layers were washed with brine (40 mL), dried (MgSO₄), filtered and concentrated. Column chromatography afforded the title compound as a white solid (3.81 g, 9.86 mmol, 88%). ^{1}H NMR (400 MHz, CDCl₃) δ 8.70 (s, 1H), 7.46 – 7.10 (m, 15H), 3.86 (s, 2H). ^{13}C NMR (101 MHz, CDCl₃) δ 162.81, 148.53, 147.18, 142.64, 136.61, 129.58, 129.04, 128.65, 127.60, 127.42, 126.80, 35.89.

3-(N-Methyl-N-phenylsulfamoyl)-N,N-diphenyl-1H-1,2,4-triazole-1-carboxamide (69)

4 N HCl in dioxane (2 mL) was added to DCM (5 mL) and cooled to -10 °C. 15%$_{wt}$ NaOCl (aq., 2 mL) was added slowly, forming a yellow-green Cl₂ solution. After 30 min **68** (100 mg, 0.259 mmol) was added dropwise as a solution in DCM (1 mL). After 20 min N-methylaniline (1 mL, 9 mmol) was added. The mixture was stirred for 2 h at RT. The reaction was quenched by the addition of 0.1 M HCl (10 mL). The organic layer was separated and the water layer was extracted with EtOAc (2x 10 mL). The combined organic layers were dried (MgSO₄), filtered and concentrated. The residue was purified by column chromatography to yield the title compound as a brown solid (78 mg, 0.18 mmol, 69%). ^{1}H NMR (400 MHz, CDCl₃) δ 8.80 (s, 1H), 7.47 – 6.99 (m, 15H), 3.07 (s, 3H). ^{13}C NMR (101 MHz, CDCl₃) δ 160.95, 147.73, 147.28, 141.75, 140.35, 129.45, 128.87, 127.58, 127.47, 126.76, 126.40, 39.02.

N-Methyl-N-phenyl-1H-1,2,4-triazole-3-sulfonamide (70)

69 (1.05 g, 2.42 mmol) was dissolved in 2 N KOH in a 1:1 (v/v) H₂O/THF mixture (50 mL). The mixture was stirred for 18 h at RT. The mixture was concentrated and the residue was purified by column chromatography to yield the title compound (308 mg, 1.29 mmol, 53%). ^{1}H NMR (400 MHz, MeOD) δ 8.63 (s, 1H), 7.46 – 7.09 (m, 5H), 3.42 (s, 3H).

tert-Butyl (S)-3-benzyl-4-(3-(N-methyl-N-phenylsulfamoyl)-1H-1,2,4-triazole-1-carbonyl)piperazine-1-carboxylate (71)
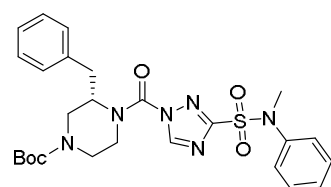
70 (30 mg, 0.13 mmol) was reacted with commercial *tert*-butyl (S)-3-benzylpiperazine-1-carboxylate (38.3 mg, 0.139 mmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (54 mg, 0.10 mmol, 79%). ^{1}H NMR (400 MHz, CDCl₃) δ 8.09 (bs, 1H), 7.38 – 7.08 (m, 9H), 6.98 (d, *J* = 8.0 Hz, 1H), 4.97 – 4.47 (m, 1H), 4.44 – 3.83 (m, 3H), 3.49 (s, 3H), 3.42 (td, *J* = 13.1, 3.6 Hz, 1H), 3.28 – 2.59 (m, 4H), 1.50 (s, 9H). ^{13}C NMR (101 MHz, CDCl₃) δ 154.79, 147.38, 140.53, 136.59, 129.31, 129.21, 128.91, 127.96, 127.22, 126.95, 80.87, 56.58, 40.60, 39.69, 35.88, 28.45.

(S)-1-(2-Benzylpiperazine-1-carbonyl)-N-methyl-N-phenyl-1H-1,2,4-triazole-3-sulfonamide 2,2,2-trifluoroacetate (72)
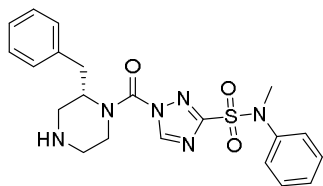
71 (54 mg, 0.10 mmol) was dissolved in DCM (5 mL) and cooled to 0 °C. 2,2,2-trifluoroacetic acid (1 mL, 13 mmol) was added and the mixture was stirred for 5 h at RT. The mixture was concentrated *in vacuo*. The residue was coevaporated with toluene (2x 5 mL). The residue was used without further purification in the next reactions.

tert-Butyl (S)-2-benzyl-4-isopropylpiperazine-1-carboxylate (73)
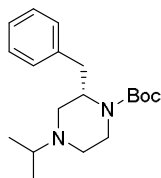
Commercial *tert*-butyl (S)-3-benzylpiperazine-1-carboxylate (100 mg, 0.362 mmol) and acetone (268 µL, 3.62 mmol) were dissolved in DCM (10 mL). The mixture was cooled to 0 °C and sodium triacetoxyborohydride (153 mg, 0.724 mmol) and acetic acid (46 µL, 0.80 mmol) were added. The mixture was stirred for 18 h at RT. The reaction was quenched by the addition of sat. aq. Na_2CO_3 (10 mL). The organic layer was separated and the aqueous phase was extracted with DCM (2x 10 mL). The combined organic layers were dried ($MgSO_4$), filtered and concentrated. The residue was purified by column chromatography to yield the title compound as a white solid (88 mg, 0.28 mmol, 76%). HPLC/MS: Calculated for $[C_{19}H_{30}N_2O_2 + H]^+ = 319.24$, found = 319.07.

(S)-3-Benzyl-1-isopropylpiperazine 2,2,2-trifluoroacetate (74)

73 was dissolved in DCM (5 mL) and TFA (1 mL) was added. The mixture was stirred for 6 h at RT, when TLC showed full conversion of the starting material. The reaction mixture was concentrated *in vacuo* and used without further purification for the next reaction.

tert-Butyl 3-benzyl-4-(3-(N-methyl-N-phenylsulfamoyl)-1H-1,2,4-triazole-1-carbonyl)piperazine-1-carboxylate (2)
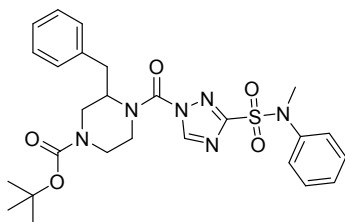
70 (50 mg, 0.21 mmol) was reacted with **38** (124 mg, 0.449 mmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (41 mg, 0.076 mmol, 36%). ^{1}H NMR (400 MHz, CDCl₃, 328K) δ 8.19 (bs, 1H), 7.49 – 6.81 (m, 10H), 4.74 (s, 1H), 4.09 (m, 3H), 3.48 (m, 3H), 3.46 – 3.34 (m, 1H), 3.19 – 2.68 (m, 4H), 1.50 (s, 9H). ^{13}C NMR (101 MHz, CDCl₃, 333K) δ 161.53, 154.89, 147.62, 140.92, 136.80, 129.40, 129.34, 128.93, 127.89, 127.24, 127.01, 80.88, 56.52, 53.46, 43.13, 39.69, 36.06, 28.54. HRMS: Calculated for $[C_{26}H_{32}N_6O_5S + Na]^+ = 563.2047$, found = 563.2053.

tert-Butyl 3-benzyl-4-(3-(benzylsulfonyl)-1H-1,2,4-triazole-1-carbonyl)piperazine-1-carboxylate (3)
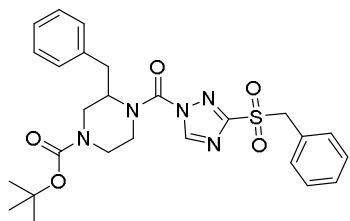
28 (150 mg, 0.67 mmol) was reacted with **38** (195 mg, 0.706 mmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (254 mg, 0.483 mmol, 69%). ^{1}H NMR (400 MHz, CDCl₃) δ 8.09 (s, 1H), 7.45 – 7.13 (m, 9H), 7.00 (s, 1H), 4.77 – 4.46 (m, 3H), 4.39 – 3.73 (m, 3H), 3.48 – 3.28 (m, 1H), 3.18 – 2.60 (m, 4H), 1.51 (s, 9H). HRMS: Calculated for $[C_{26}H_{31}N_5O_5S + Na]^+ = 548.1938$, found = 548.1945.

1-(3-Benzyl-4-(3-(benzylsulfonyl)-1H-1,2,4-triazole-1-carbonyl)piperazin-1-yl)ethan-1-one (4)
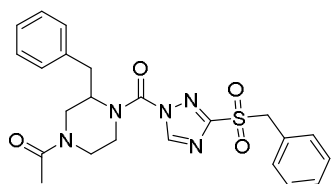
28 (49 mg, 0.22 mmol) was reacted with **40** (51 mg, 0.20 mmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (67 mg, 0.143 mmol, 72%). ^{1}H NMR (400 MHz, DMSO-d₆) δ 9.49 – 8.47 (m, 1H), 7.51 – 6.90 (m, 10H), 4.91 (s, 2H), 4.69 – 4.21 (m, 2H), 4.03 (q, *J* = 7.1 Hz, 1H), 3.74 (d, *J* = 13.9 Hz, 1H), 3.60 – 3.36 (m, 1H), 3.24 (t, *J* = 12.6 Hz, 1H), 2.98 – 2.66 (m, 3H), 2.14 (s, 1H), 1.99 (s, 2H) (spectrum shows a mixture of rotamers). HRMS: Calculated for $[C_{23}H_{25}N_5O_4S + H]^+ = 468.1699$, found = 468.1699.

(3-Benzyl-4-(3-(benzylsulfonyl)-1*H*-1,2,4-triazole-1-carbonyl)piperazin-1-yl)(cyclobutyl)methanone (**5**)
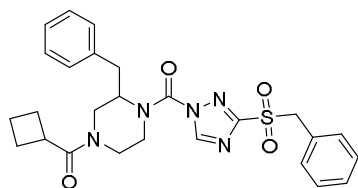
28 (62 mg, 0.28 mmol) was reacted with **42** (76 mg, 0.29 mmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (46 mg, 0.091 mmol, 31%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.13 – 7.79 (m, 1H), 7.53 – 6.72 (m, 10H), 4.97 – 4.68 (m, 1H), 4.70 – 4.51 (m, 3H), 4.22 – 4.04 (m, 1H), 3.94 – 3.53 (m, 1H), 3.54 – 3.09 (m, 3H), 3.07 – 2.51 (m, 3H), 2.51 – 1.79 (m, 6H) (spectrum shows a mixture of rotamers). HRMS: Calculated for [C$_{27}$H$_{33}$N$_5$O$_4$S + H]$^+$ = 524.232, found = 524.2328.

1-(3-Benzyl-4-(3-(benzylsulfonyl)-1*H*-1,2,4-triazole-1-carbonyl)piperazin-1-yl)-3,3-dimethylbutan-1-one (**6**)

28 (60 mg, 0.27 mmol) was reacted with **44** (78 mg, 0.28 mmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (88 mg, 0.17 mmol, 59%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.01 (s, 1H), 7.45 – 6.84 (m, 10H), 5.04 – 4.74 (m, 1H), 4.73 – 4.53 (m, 3H), 4.24 – 3.99 (m, 1H), 3.99 – 3.72 (m, 1H), 3.59 – 3.18 (m, 2H), 3.17 – 2.98 (m, 1H), 2.98 – 2.60 (m, 2H), 2.44 – 2.07 (m, 2H), 1.18 – 0.96 (m, 9H) (spectrum shows a mixture of rotamers). HRMS: Calculated for [C$_{27}$H$_{33}$N$_5$O$_4$S + H]$^+$ = 524.2325, found = 524.2328.
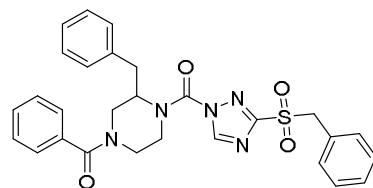
(4-Benzoyl-2-benzylpiperazin-1-yl)(3-(benzylsulfonyl)-1*H*-1,2,4-triazol-1-yl)methanone (**7**)

28 (53 mg, 0.24 mmol) was reacted with **46** (70 mg, 0.25 mmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (26.2 mg, 0.049 mmol, 20%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.91 (s, 1H), 7.61 – 6.57 (m, 15H), 4.71 (bs, 1H), 4.65 – 4.51 (m, 2H), 4.31 – 4.03 (m, 1H), 4.01 – 3.66 (m, 1H), 3.62 – 3.29 (m, 2H), 3.29 – 2.84 (m, 3H) (spectrum shows a mixture of rotamers). HRMS: Calculated for [C$_{28}$H$_{27}$N$_5$O$_4$S + H]$^+$ = 530.1856, found = 530.1857.

1,3-Dibenzyl-4-(3-(benzylsulfonyl)-1*H*-1,2,4-triazole-1-carbonyl)piperazin-2-one (**8**)
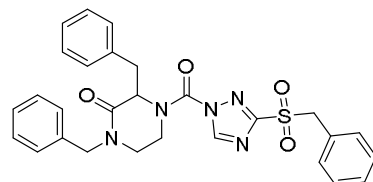
28 (53 mg, 0.24 mmol) was reacted with **49** (70 mg, 0.25 mmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (118 mg, 0.223 mmol, 89%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.97 (s, 1H), 7.52 – 7.02 (m, 14H), 6.92 (s, 1H), 5.61 (s, 1H), 5.10 (s, 1H), 4.81 – 4.51 (m, 2H), 4.45 – 4.04 (m, 2H), 3.41 (m, 2H), 3.31 – 3.04 (m, 2H), 2.92 – 2.66 (m, 1H). HRMS: Calculated for [C$_{28}$H$_{27}$N$_5$O$_4$S + H]$^+$ = 530.1856, found = 530.1859.

3-Benzyl-4-(3-(benzylsulfonyl)-1*H*-1,2,4-triazole-1-carbonyl)piperazine-1-carboxamide (**9**)

28 (61 mg, 0.27 mmol) was reacted with **51** (63 mg, 0.29 mmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (81.3 mg, 0.174 mmol, 60%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.13 (s, 1H), 7.43 – 7.11 (m, 9H), 6.99 (s, 1H), 5.01 (s, 2H), 4.78 – 4.48 (m, 3H), 4.26 – 4.05 (m, 1H), 3.91 (m, 1H), 3.75 (d, *J* = 13.7 Hz, 1H), 3.43 (t, *J* = 12.5 Hz, 1H), 3.15 (dd, *J* = 13.7, 3.8 Hz, 1H), 3.09 – 2.89 (m, 2H), 2.86 – 2.71 (m, 1H). HRMS: Calculated for [C$_{22}$H$_{24}$N$_6$O$_4$S + H]$^+$ = 469.1652, found = 469.1658.

3-Benzyl-4-(3-(benzylsulfonyl)-1*H*-1,2,4-triazole-1-carbonyl)-*N*-ethylpiperazine-1-carboxamide (**10**)
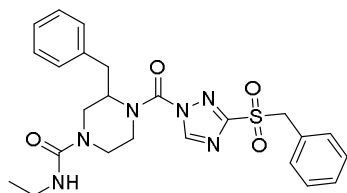
28 (87 mg, 0.39 mmol) was reacted with **53** (102 mg, 0.41 mmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (140 mg, 0.283 mmol, 69%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.18 (s, 1H), 7.42 – 7.11 (m, 9H), 7.00 (s, 1H), 4.84 – 4.52 (m, 4H), 4.32 – 4.04 (m, 1H), 3.93 (d, *J* = 18.4 Hz, 1H), 3.72 (d, *J* = 13.6 Hz, 1H), 3.51 – 3.36 (m, 1H), 3.34 – 3.16 (m, 2H), 3.17 – 2.87 (m, 3H), 2.87 – 2.75 (m, 1H), 1.15 (t, *J* = 7.2 Hz, 3H). HRMS: Calculated for [C$_{24}$H$_{28}$N$_6$O$_4$S + H]$^+$ = 497.1965, found = 497.1969.

3-Benzyl-4-(3-(benzylsulfonyl)-1*H*-1,2,4-triazole-1-carbonyl)-*N*-(*tert*-butyl)piperazine-1-carboxamide (**11**)
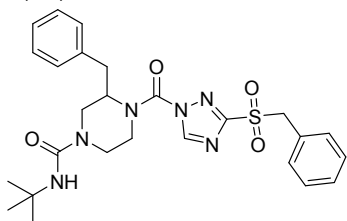
28 (92 mg, 0.41 mmol) was reacted with **55** (119 mg, 0.43 mmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (184 mg, 0.351 mmol, 81%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.28 (s, 1H), 7.44 – 7.15 (m, 9H), 7.05 (s, 1H), 4.73 – 4.49 (m, 3H), 4.40 – 4.01 (m, 2H), 3.97 – 3.80 (m, 1H), 3.56 (d, *J* = 13.2 Hz, 1H), 3.41 (d, *J* = 12.3 Hz, 1H), 3.14 – 2.95 (m, 2H), 2.95 – 2.77 (m, 2H), 1.35 (s, 9H). HRMS: Calculated for [C$_{26}$H$_{32}$N$_6$O$_4$S + H]$^+$ = 525.2278, found = 525.2286.

(2-Benzyl-4-ethylpiperazin-1-yl)(3-(benzylsulfonyl)-1*H*-1,2,4-triazol-1-yl)methanone (**12**)
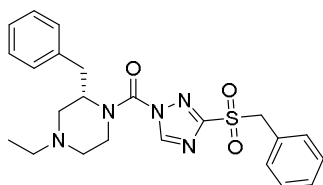
67 (60 mg, 0.13 mmol) was dissolved in DCM (5 mL) and cooled to 0 °C. Acetaldehyde (8.0 μL, 0.14 mmol), acetic acid (9.1 μL, 0.16 mmol) and sodium triacetoxyhydroborate (33.0 mg, 0.156 mmol) were added and the mixture was stirred for 18 h warming up to RT. The reaction was quenched with saturated aqueous Na$_2$CO$_3$ and the organic layer was separated. The water layer was extracted twice with DCM. The combined organic layers were dried (MgSO$_4$), filtered and concentrated. Column chromatography yielded the title compound as a white solid (20 mg, 44 μmol, 34%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.04 (s, 1H), 7.54 – 6.82 (m, 10H), 4.73 – 4.49 (m, 3H), 4.21 – 3.82 (m, 1H), 3.52 (t, *J* = 12.9 Hz, 1H), 3.38 – 3.16 (m, 1H), 3.03 (s, 1H), 2.93 – 2.68 (m, 2H), 2.60 – 2.00 (m, 4H), 1.10 (t, *J* = 7.3 Hz, 3H). HRMS: Calculated for [C$_{23}$H$_{27}$N$_5$O$_3$S + H]$^+$ = 454.1907, found = 454.1909.

(2-Benzyl-4-isopropylpiperazin-1-yl)(3-(benzylsulfonyl)-1*H*-1,2,4-triazol-1-yl)methanone (**13**)
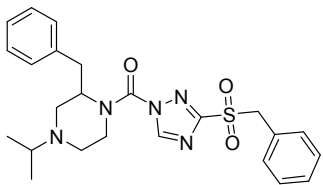
28 (81 mg, 0.36 mmol) was reacted with **57** (83 mg, 0.38 mmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (66.9 mg, 0.143 mmol, 38%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.02 (s, 1H), 7.25 (m, 9H), 7.07 – 6.91 (m, 1H), 4.66 – 4.49 (m, 3H), 4.19 – 3.82 (m, 1H), 3.47 (t, *J* = 12.5 Hz, 1H), 3.28 (t, *J* = 11.4 Hz, 1H), 3.02 – 2.61 (m, 4H), 2.53 – 2.18 (m, 2H), 1.08 – 0.92 (m, 6H). HRMS: Calculated for [C$_{24}$H$_{29}$N$_5$O$_3$S + H]$^+$ = 468.2063, found = 468.2066.

(2-Benzyl-4-isobutylpiperazin-1-yl)(3-(benzylsulfonyl)-1*H*-1,2,4-triazol-1-yl)methanone (**14**)
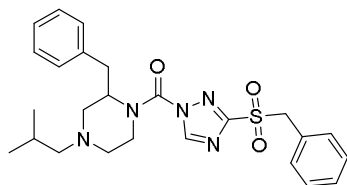
28 (43 mg, 0.19 mmol) was reacted with **59** (40 mg, 0.17 mmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (63.6 mg, 0.132 mmol, 76%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.04 (s, 1H), 7.44 – 6.83 (m, 10H), 4.71 – 4.41 (m, 3H), 4.21 – 3.69 (m, 1H), 3.51 (td, *J* = 12.9, 3.5 Hz, 1H), 3.36 – 3.20 (m, 1H), 3.08 – 2.66 (m, 3H), 2.33 – 1.85 (m, 4H), 1.85 – 1.54 (m, 1H), 0.95 (d, *J* = 6.5 Hz, 6H). HRMS: Calculated for [C$_{25}$H$_{31}$N$_5$O$_3$S + H]$^+$ = 482.2220, found = 482.2225.

(2-Benzyl-4-cyclopentylpiperazin-1-yl)(3-(benzylsulfonyl)-1*H*-1,2,4-triazol-1-yl)methanone (15)
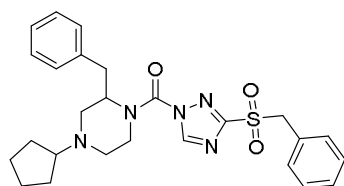
28 (39 mg, 0.17 mmol) was reacted with **61** (38 mg, 0.16 mmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (20.5 mg, 0.042 mmol, 27%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.01 (s, 1H), 7.43 – 6.90 (m, 10H), 4.69 – 4.47 (m, 3H), 4.34 – 3.84 (m, 1H), 3.71 – 2.65 (m, 6H), 2.62 – 1.28 (m, 10H) (spectrum shows a mixture of rotamers). HRMS: Calculated for [C$_{26}$H$_{31}$N$_5$O$_3$S + H]$^+$ = 494.2220, found = 494.2223.

(3-(Benzylsulfonyl)-1*H*-1,2,4-triazol-1-yl)(2,4-dibenzylpiperazin-1-yl)methanone (16)
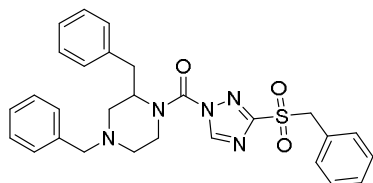
28 (46.1 mg, 0.206 mmol) was reacted with **32** (50 mg, 0.19 mmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (85 mg, 0.17 mmol, 88%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.03 (s, 1H), 7.45 – 7.19 (m, 11H), 7.19 – 6.73 (m, 5H), 4.60 (m, 3H), 4.12 (q, *J* = 7.0 Hz, 1H), 3.64 – 3.37 (m, 3H), 3.29 (t, *J* = 11.3 Hz, 1H), 3.01 (d, *J* = 11.9 Hz, 1H), 2.79 (dd, *J* = 28.0, 9.1 Hz, 2H), 2.34 – 2.07 (m, 2H). HRMS: Calculated for [C$_{28}$H$_{29}$N$_5$O$_3$S + H]$^+$ = 516.2063, found = 516.2070.

(*S*)-(2-Benzyl-4-(4-fluorobenzyl)piperazin-1-yl)(3-(benzylsulfonyl)-1*H*-1,2,4-triazol-1-yl)methanone (17)
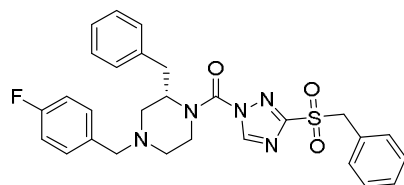
28 (19 mg, 85 µmol) was reacted with **63** (22 mg, 77 µmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (33 mg, 62 µmol, 80%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.04 (s, 1H), 7.42 – 6.81 (m, 14H), 4.60 (m, 3H), 4.23 – 3.83 (m, 1H), 3.62 – 3.12 (m, 4H), 2.99 (d, *J* = 11.8 Hz, 1H), 2.76 (dd, *J* = 26.6, 9.8 Hz, 2H), 2.37 – 2.04 (m, 2H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 162.33 (d, *J* = 245.7 Hz), 160.66, 147.88, 147.61, 137.47, 133.35 (d, *J* = 3.2 Hz), 131.14, 130.74, 128.97, 128.87, 127.00, 126.37, 115.37 (d, *J* = 21.3 Hz), 77.36, 61.88, 60.65, 56.96, 55.45, 55.06, 53.02, 52.52, 43.66, 41.47, 36.74, 35.95. HRMS: Calculated for [C$_{28}$H$_{28}$FN$_5$O$_3$S + H]$^+$ = 534.1969, found = 534.1975.

(*S*)-(2-Benzyl-4-(4-ethynylbenzyl)piperazin-1-yl)(3-(benzylsulfonyl)-1*H*-1,2,4-triazol-1-yl)methanone (18)
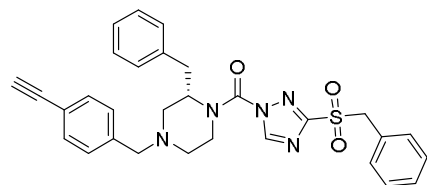
28 (27 mg, 0.12 mmol) was reacted with **65** (32 mg, 0.11 mmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (37 mg, 69 µmol, 62%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.04 (s, 1H), 7.56 – 6.83 (m, 14H), 4.60 (m, 3H), 4.05 (m, 1H), 3.71 – 3.14 (m, 4H), 3.10 (s, 1H), 2.99 (s, 1H), 2.76 (dd, *J* = 26.1, 11.7 Hz, 2H), 2.27 (s, 2H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 160.62, 148.16, 147.86, 147.58, 138.57, 137.39, 132.30, 131.10, 129.22, 129.09, 128.94, 128.85, 126.98, 126.31, 121.38, 83.47, 77.32, 62.27, 60.61, 56.91, 55.14, 53.06, 52.54, 43.63, 41.42, 36.69, 35.78, 30.39. HRMS: Calculated for [C$_{30}$H$_{29}$N$_5$O$_3$S + H]$^+$ = 540.2063, found = 540.2062.

(*S*)-(2-Benzyl-4-(methylsulfonyl)piperazin-1-yl)(3-(benzylsulfonyl)-1*H*-1,2,4-triazol-1-yl)methanone (19)
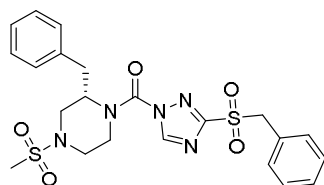
67 (60 mg, 0.13 mmol) was dissolved in DCM (5 mL). Mesyl chloride (12 µL, 0.16 mmol) and Hünig's base (68 µL, 0.39 mmol) were added. The mixture was stirred for 20 h at RT. The reaction mixture was concentrated under reduced pressure and purified by column chromatography to afford the title compound as a white solid (30 mg, 59 µmol, 45%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.08 (s, 1H), 7.44 – 6.85 (m, 10H), 4.63 (s, 3H), 4.35 – 3.83 (m, 2H), 3.80 – 3.63 (m, 1H), 3.61 – 3.32 (m, 1H), 3.21 (s, 1H), 2.83 (m, 6H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 160.92, 148.01, 147.84, 136.24, 131.14, 129.34, 129.03, 127.46, 126.34, 77.36, 60.58, 60.51, 56.18, 53.58, 48.15, 45.31, 40.53, 35.80, 34.85, 29.80. HRMS: Calculated for [C$_{22}$H$_{25}$N$_5$O$_5$S$_2$ + Na]$^+$ = 526.1189, found = 526.1191.

(S)-(2-Benzyl-4-(isopropylsulfonyl)piperazin-1-yl)(3-(benzylsulfonyl)-1*H*-1,2,4-triazol-1-yl)methanone (**20**)
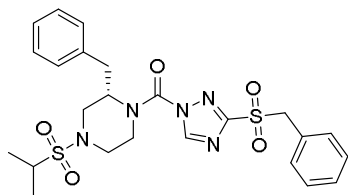
67 (68 mg, 0.15 mmol) was dissolved in DCM (5 mL) to which propane-2-sulfonyl chloride (36 μL, 0.18 mmol) and DIPEA (77 μL, 0.44 mmol) were added. The mixture was stirred for 20 h at RT. The reaction mixture was concentrated under reduced pressure and purified by column chromatography to afford the title compound as a white solid (31.8 mg, 60 μmol, 41%). ^{1}H NMR (400 MHz, CDCl3) δ 7.95 (s, 1H), 7.43 – 6.86 (m, 10H), 4.72 – 4.52 (m, 3H), 4.15 (d, *J* = 13.7 Hz, 1H), 4.01 (d, *J* = 13.0 Hz, 1H), 3.79 (d, *J* = 12.9 Hz, 1H), 3.58 – 3.40 (m, 1H), 3.37 – 2.89 (m, 4H), 2.79 (d, *J* = 14.2 Hz, 1H), 1.40 – 1.33 (m, 6H). ^{13}C NMR (101 MHz, CDCl3) δ 160.88, 148.29, 147.65, 136.45, 131.15, 129.37, 129.28, 129.05, 127.43, 126.44, 77.36, 60.57, 56.54, 53.82, 48.88, 45.68, 41.10, 35.80, 16.95, 16.88. HRMS: Calculated for [$C_{24}H_{29}N_5O_5S_2$ + Na]$^+$ = 554.1502, found = 554.1511.

(S)-(2-Benzyl-4-(phenylsulfonyl)piperazin-1-yl)(3-(benzylsulfonyl)-1*H*-1,2,4-triazol-1-yl)methanone (**21**)
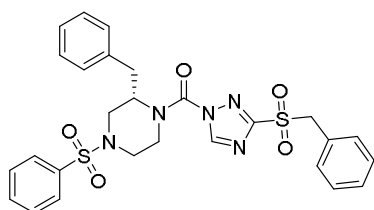
67 (60 mg, 0.13 mmol) was dissolved in DCM (5 mL). Benzenesulfonyl chloride (20 μL, 0.16 mmol) and DIPEA (68 μL, 0.39 mmol) were added. The mixture was stirred for 20 h at RT. The reaction mixture was concentrated under reduced pressure and purified by column chromatography to afford the title compound as a white solid (21.2 mg, 37 μmol, 29%). ^{1}H NMR (400 MHz, CDCl3) δ 8.01 (s, 1H), 7.86 – 6.87 (m, 15H), 4.80 – 4.51 (m, 3H), 4.26 – 4.08 (m, 1H), 4.08 – 3.86 (m, 1H), 3.86 – 3.40 (m, 2H), 3.38 – 3.15 (m, 1H), 3.05 – 2.79 (m, 1H), 2.60 – 2.24 (m, 2H). ^{13}C NMR (101 MHz, CDCl3) δ 160.89, 147.91, 147.74, 136.33, 135.02, 133.61, 131.13, 129.81, 129.56, 129.40, 129.33, 129.13, 129.01, 127.83, 127.46, 77.36, 60.54, 56.10, 54.81, 48.59, 46.59, 45.57, 42.58, 40.39, 36.06. HRMS: Calculated for [$C_{27}H_{27}N_5O_5S_2$ + Na]$^+$ = 588.1346, found = 588.1355.

(S)-(2-Benzyl-4-((4-fluorophenyl)sulfonyl)piperazin-1-yl)(3-(benzylsulfonyl)-1*H*-1,2,4-triazol-1-yl)methanone (**22**)
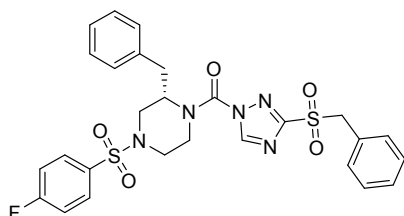
67 (68 mg, 0.15 mmol) was dissolved in DCM (5 mL). 4-fluorobenzenesulfonyl chloride (43 μL, 0.18 mmol) and DIPEA (77 μL, 0.44 mmol) were added. The mixture was stirred for 20 h at RT. The reaction mixture was concentrated under reduced pressure and purified by column chromatography to afford the title compound as a white solid (52 mg, 89 μmol, 61%). ^{1}H NMR (400 MHz, CDCl3) δ 8.04 (s, 1H), 7.82 – 6.80 (m, 14H), 4.79 – 4.51 (m, 3H), 4.20 (d, *J* = 13.7 Hz, 1H), 3.94 (d, *J* = 11.7 Hz, 1H), 3.78 – 3.41 (m, 2H), 3.39 – 3.16 (m, 1H), 3.03 – 2.84 (m, 1H), 2.55 – 2.31 (m, 2H). ^{13}C NMR (101 MHz, CDCl3) δ 165.58 (d, *J* = 256.4 Hz), 160.86, 148.32, 147.70, 136.15, 131.05, 130.49 (d, *J* = 9.4 Hz), 129.72, 129.29, 129.23, 129.02, 128.92, 127.38, 126.21, 116.80 (d, *J* = 22.6 Hz), 77.27, 60.44, 55.97, 48.42, 46.41, 45.48, 40.22, 35.93. HRMS: Calculated for [$C_{27}H_{26}FN_5O_5S_2$ + Na]$^+$ = 606.1251, found = 606.1258.

(S)-1-(2-Benzyl-4-isopropylpiperazine-1-carbonyl)-*N*-methyl-*N*-phenyl-1*H*-1,2,4-triazole-3-sulfonamide (**23**)
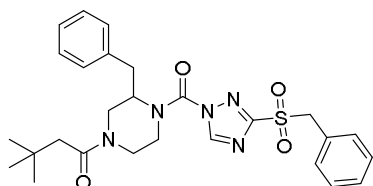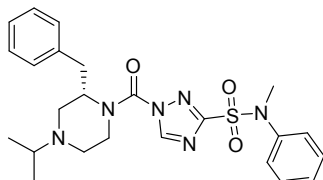
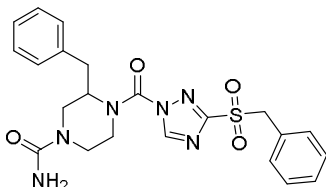**70** (60 mg, 0.25 mmol) was reacted with **72** (92 mg, 0.28 mmol) according to General Procedure 3. Column chromatography afforded the title compound as a white solid (56 mg, 0.12 mmol, 46%). ^{1}H NMR (500 MHz, CDCl3) δ 8.03 (s, 1H), 7.37 – 6.88 (m, 10H), 4.81 – 4.47 (m, 1H), 4.14 (d, *J* = 13.1 Hz, 1H), 3.56 – 3.44 (m, 4H), 3.34 – 3.18 (m, 1H), 3.07 – 2.87 (m, 1H), 2.87 – 2.64 (m, 3H), 2.55 – 2.14 (m, 2H), 1.13 – 0.87 (m, 6H). ^{13}C NMR (126 MHz, CDCl3) δ 160.77, 148.18, 147.86, 147.27, 140.64, 137.85, 129.59, 129.32, 128.82, 127.93, 127.02, 126.89, 57.01, 55.42, 54.52, 50.70, 49.20, 48.60, 44.10, 41.84, 39.73, 36.71, 35.74, 18.77, 17.76.

(S)-1-(2-Benzyl-4-(4-fluorobenzyl)piperazine-1-carbonyl)-N-methyl-N-phenyl-1H-1,2,4-triazole-3-sulfonamide (**24**)
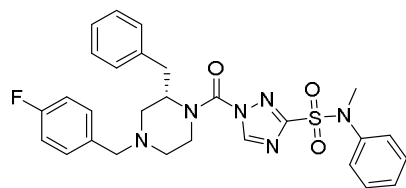
70 (98 mg, 0.18 mmol) was dissolved in DCM (5 mL) and cooled to 0 °C. 4-Fluorobenzaldehyde (27 µL, 0.25 mmol), acetic acid (12 µL, 0.21 mmol) and sodium triacetoxyhydroborate (45 mg, 0.21 mmol) were added and the mixture was stirred for 18 h warming up to RT. The reaction was quenched with saturated aqueous Na_2CO_3 and the organic layer was separated. The water layer was extracted twice with DCM. The combined organic layers were dried ($MgSO_4$) and concentrated. Column chromatography yielded the title compound as a white solid (9.4 mg, 17 µmol, 10%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.03 (s, 1H), 7.47 – 6.77 (m, 14H), 4.83 – 4.48 (m, 1H), 4.32 – 4.06 (m, 1H), 3.64 – 3.12 (m, 6H), 3.12 – 1.86 (m, 6H). HRMS: Calculated for $[C_{28}H_{29}FN_6O_3S + H]^+$ = 549.2078, found = 549.2090.

(S)-1-(2-Benzyl-4-(isopropylsulfonyl)piperazine-1-carbonyl)-N-methyl-N-phenyl-1H-1,2,4-triazole-3-sulfonamide (**25**)
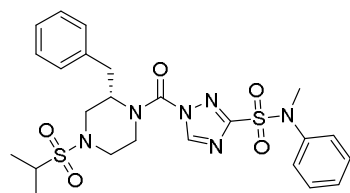
70 (205 mg, 0.37 mmol) was dissolved in DCM (10 mL). Propane-2-sulfonyl chloride (50 µL, 0.44 mmol) and Hünig's base (194 µL, 1.11 mmol) were added and the reaction mixture was stirred for 18 h at RT. The mixture was concentrated under reduced pressure and purified by column chromatography to yield the title compound as a white solid (127 mg, 232 µmol, 63%). ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.97 (s, 1H), 7.44 – 6.80 (m, 10H), 4.91 – 4.54 (m, 1H), 4.37 – 4.10 (m, 1H), 4.05 – 3.71 (m, 2H), 3.49 (s, 4H), 3.37 – 2.90 (m, 4H), 2.87 – 2.75 (m, 1H), 1.42 – 1.29 (m, 6H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 160.93, 148.39, 147.27, 140.41, 136.43, 129.62, 129.27, 129.14, 128.90, 127.95, 127.23, 126.87, 77.33, 61.85, 56.42, 54.95, 53.61, 48.64, 47.00, 45.57, 43.56, 40.93, 39.67, 35.73, 34.96, 16.82, 16.76.

(S)-1-(2-Benzyl-4-((4-fluorophenyl)sulfonyl)piperazine-1-carbonyl)-N-methyl-N-phenyl-1H-1,2,4-triazole-3-sulfonamide (**26**)
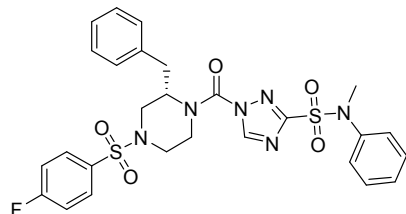
**70** (55 mg, 0.10 mmol) was dissolved in DCM (5 mL). 4-Fluorobenzene-sulfonyl chloride (23 mg, 0.12 mmol) and Hünig's base (52 µL, 0.30 mmol) were added and the reaction mixture was stirred for 18 h at RT. The mixture was concentrated under reduced pressure and purified by column chromatography to yield the title compound as a white solid (44.5 mg, 74 µmol, 74%). ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.04 (s, 1H), 7.81 – 6.87 (m, 14H), 5.03 – 4.52 (m, 1H), 4.50 – 4.16 (m, 1H), 3.94 (d, J = 11.6 Hz, 1H), 3.82 – 3.53 (m, 2H), 3.48 (s, 3H), 3.40 – 3.16 (m, 1H), 3.08 – 2.80 (m, 1H), 2.63 – 2.27 (m, 2H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 165.65 (d, J = 256.2 Hz), 161.21, 148.18, 147.48, 140.44, 136.29, 131.13, 130.57 (d, J = 9.4 Hz), 129.37, 129.07, 128.08, 127.42, 127.07, 116.88 (d, J = 22.7 Hz), 56.03, 54.60, 48.54, 46.52, 45.57, 42.53, 40.25, 39.85, 36.07, 35.05. HRMS: Calculated for $[C_{27}H_{27}FN_6O_5S_2 + Na]^+$ = 621.1360, found = 621.1368.

References

1. Baggelaar, M. P., Maccarrone, M. & van der Stelt, M. 2-Arachidonoylglycerol: A signaling lipid with manifold actions in the brain. *Prog. Lipid Res.* **71,** 1–17 (2018).
2. Baggelaar, M. P. *et al.* Highly Selective, Reversible Inhibitor Identified by Comparative Chemoproteomics Modulates Diacylglycerol Lipase Activity in Neurons. *J. Am. Chem. Soc.* **137,** 8851–8857 (2015).
3. Hsu, K. *et al.* DAGLβ inhibition perturbs a lipid network involved in macrophage inflammatory responses. *Nat. Chem. Biol.* **8,** 999–1007 (2012).
4. Blankman, J. L., Simon, G. M. & Cravatt, B. F. A Comprehensive Profile of Brain Enzymes that Hydrolyze the Endocannabinoid 2-Arachidonoylglycerol. *Chem. Biol.* **14,** 1347–1356 (2007).
5. Gao, Y. *et al.* Loss of Retrograde Endocannabinoid Signaling and Reduced Adult Neurogenesis in Diacylglycerol Lipase Knock-out Mice. *J. Neurosci.* **30,** 2017–2024 (2010).
6. Janssen, F. J. & van der Stelt, M. Inhibitors of diacylglycerol lipases in neurodegenerative and metabolic disorders. *Bioorg. Med. Chem. Lett.* **26,** 3831–3837 (2016).
7. Bisogno, T. *et al.* Cloning of the first sn1-DAG lipases points to the spatial and temporal regulation of endocannabinoid signaling in the brain. *J. Cell Biol.* **163,** 463–468 (2003).
8. European Medicines Agency. The European Medicines Agency recommends suspension of the marketing authorisation of Acomplia. https://www.ema.europa.eu/en/news/european-medicines-agency-recommends-suspension-marketing-authorisation-acomplia (2008).
9. Hitchcock, S. A. & Pennington, L. D. Structure−Brain Exposure Relationships. *J. Med. Chem.* **49,** 7559–7583 (2006).
10. 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).
11. 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).
12. Patricelli, M. P., Lovato, M. A. & Cravatt, B. F. Chemical and mutagenic investigations of fatty acid amide hydrolase: Evidence for a family of serine hydrolases with distinct catalytic properties. *Biochemistry* **38,** 9804–9812 (1999).
13. Janssen, A. P. A. *et al.* Development of a Multiplexed Activity-Based Protein Profiling Assay to Evaluate Activity of Endocannabinoid Hydrolase Inhibitors. *ACS Chem. Biol.* **13,** 2406–2413 (2018).
14. Gottlieb, H. E., Kotlyar, V. & Nudelman, A. NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. *J. Org. Chem.* **62,** 7512–7515 (1997).

Structure Kinetics Relationships for irreversible DAGL inhibitors

Introduction

The last decade has seen a renewed interest in the development of covalent inhibitors for several classes of drug targets, including the serine hydrolase and kinase superfamily.[1,2] Serine hydrolases perform a broad array of physiological functions and have diverse substrate preferences, but they all share a conserved catalytic nucleophilic serine residue. Serine hydrolases form a covalent acyl-enzyme intermediate, which can be exploited by irreversible, mechanism-based inhibitors. For example, PF-04457845 and ABX-1431 have entered clinical trials as fatty acid amide hydrolase (FAAH) and monoacylglycerol lipase

(MAGL) inhibitors for the treatment of neurological diseases, respectively.[3-8] Covalent, irreversible inhibitors initially bind in a reversible fashion to the protein (i.e. the Michealis-Menten complex) followed by a time-dependent chemical reaction that inactivates the enzyme (Figure 4.1A and 4.1B).[9] The half maximum inhibitory concentration (IC_{50}) for covalent inhibitors is dependent on a combination of binding affinity (K_i) and reactivity (k_{inact}). This fundamental dual aspect of covalent inhibition is often not taken into account during the optimization of covalent irreversible inhibitors, which is usually based on IC_{50} values.[2,10] This may lead to the prioritization of highly reactive molecules (large k_{inact}) based on their high potency.[11] Intrinsic high reactivity may, however, lead to a-specific binding to other members of the same enzyme family and unwanted adverse side effects as recently witnessed for BIA 10-2474 (Chapter 5).[12,13] Of note, the specificity constant $\left(\frac{k_{inact}}{K_i}\right)$ is sometimes employed to guide inhibitor optimization to avoid IC_{50}-values that are assay- and time-dependent.[14-16] The specificity constant is determined by measuring the observed rate constants (k_{obs}) using various inhibitor pre-incubation times, but does not allow the independent optimization of the affinity K_i, while minimizing the reactivity k_{inact}. Thus, alternative methods are required to determine K_i and k_{inact} in an independent manner.

Diacylglycerol lipases (DAGL) are serine hydrolases responsible for the synthesis of the endocannabinoid 2-arachidonoylglycerol (2-AG).[17] Modulation of DAGL activity holds therapeutic promise for the treatment of metabolic and neuroinflammatory diseases.[18-20] Several DAGL inhibitors, including KT109, DH376 and LEI105, have been developed (Figure 4.1C).[20-22] KT109 and DH376 belong to the class of triazole urea inhibitors, which are capable of irreversibly binding the catalytic serine through the formation of a stable carbamate adduct, thereby expelling a triazole-moiety as a leaving group (Figure 4.1A). This class of compounds has also shown merit as inhibitors for other serine hydrolases, such as α/β-hydrolase domain containing protein (ABHD) 6[21], ABHD11[23], DDHD domain-containing protein 2 (DDHD2)[24] and MAGL[11,25]. Structure activity studies of the DAGL (KT109 and DH376) and MAGL (JJKK-048) inhibitors have shown that the heterocyclic leaving group is crucially important for inhibitor activity.[11,26,27] The pK_a of the leaving group as determinant for the reactivity of the urea[28] was postulated to determine the activity of inhibitors.[11] The exact kinetic parameters of binding for these inhibitors have, however, thus far not been experimentally measured, thus the precise role of the triazole heterocycle in the inhibitor activity is unknown.

Here, we studied in detail the influence of the heterocycle in DAGL-α inhibitor DH376 on the binding and reactivity. To this end, we synthesized a coherent set of azole derivatives of DH376 to systematically investigate the role of the heterocycle in the activity of the inhibitor. Furthermore, we adapted a surrogate substrate assay of DAGL-α, which allowed us to independently measure K_i and k_{inact} of the new inhibitors. Surprisingly, we found that the azole has a crucial role in the formation of the Michaelis-Menten complex, rather than in modulating the reactivity.
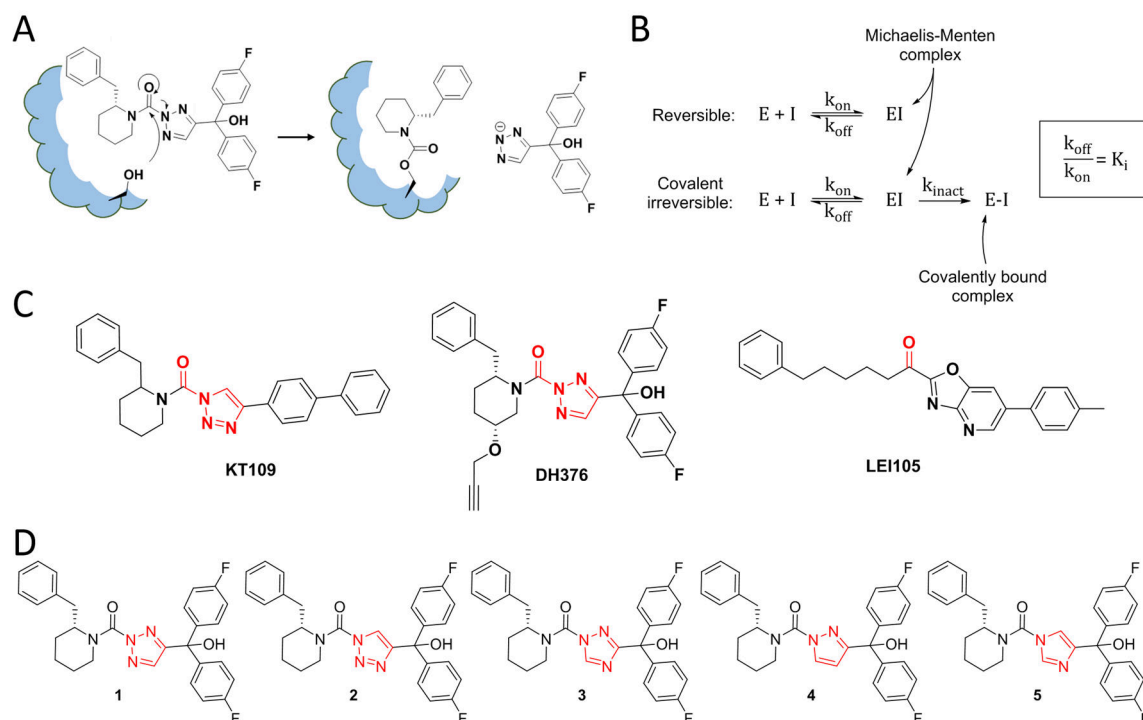
Figure 4.1 | A) Mechanism of action of mechanism based triazole urea inhibitors for serine hydrolases. B) The inhibition reactions for reversible and irreversible inhibition. E: enzyme, I: inhibitor, EI: Michaelis-Menten complex of E and I, E-I: covalently bound inhibitor-enzyme complex. C) Recently published potent DAGL inhibitors. D) Set of 5 inhibitors based on DH376 synthesized and characterized in this study.

Results

To study the role of the azole heterocycle in the activity of the DAGL inhibitor DH376, a focused set of DH376 analogues was synthesized (**1**-**5**) (Figure 4.1D). Four different heterocycles (1,2,3-triazole, 1,2,4-triazole, pyrazole and imidazole) and a regio-isomer were selected, because they differ five orders of magnitude in pK_a.[29–31] The compounds were synthesized according to Scheme 4.1. The enantioselective synthetic route towards (*R*)-2-benzylpiperidine was adapted from Deng *et al.*[27] We replaced the low yielding transamination step, used to introduce an alkene at the free amine of **10**, by a simple alkylation with 4-bromobut-1-ene after nosyl protection of the amine (Scheme 4.1). This protected diene was subjected to ring-closing metathesis and, after deprotection, yielded **15** in a three-fold higher overall yield than previously reported.[27] To synthesize the leaving group azole derivatives, a general synthetic route was devised featuring a Grignard reaction as core transformation to yield the di-(*p*-fluorophenyl)methanol moiety in a single step from accessible azole esters (Scheme 4.1). For all but the imidazole derivative this worked without the introduction of any protecting group. Tritylation of the imidazole ethyl ester was necessary to avoid degradation during the Grignard reaction. Finally, the secondary amine **15** was first transformed in a carbamoyl chloride using triphosgene and subsequently reacted with the diphenyl azoles (**18, 21, 25** and **29**), which furnished the inhibitors **1-5**.
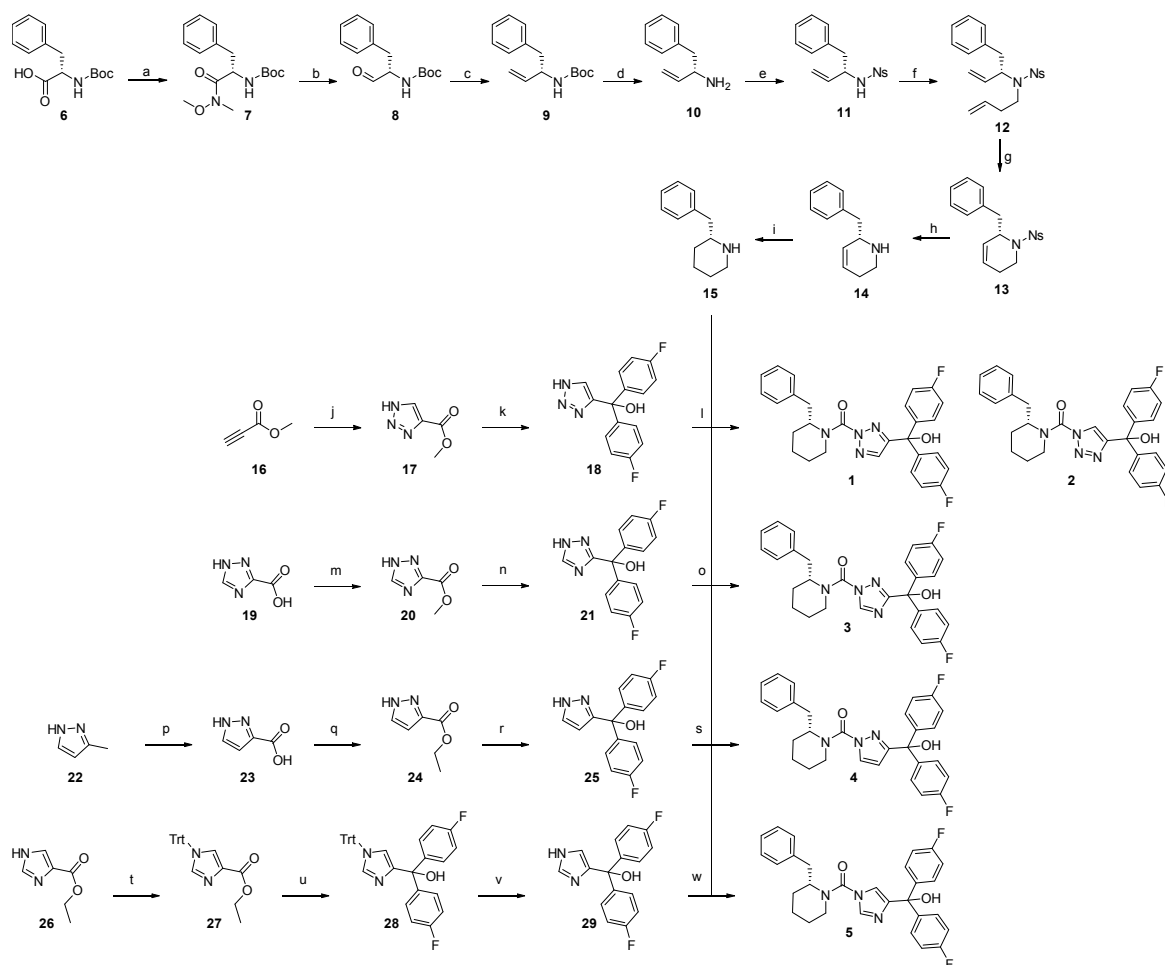
Scheme 4.1 | Synthesis of **1-5** implementing the optimized 9-step procedure for the synthesis of (*R*)-2-benzylpiperidine **15** from commercially available *N*-Boc-L-phenylalanine (**6**), followed by the coupling to the biphenyl-azole leaving groups synthesized using a general Grignard reaction. a) N,O-di-Me-hydroxyl-amine·HCl, EDCl·HCl, DCM, 0 °C → RT, 92%; b) LiAlH₄, THF, -20 °C, 96%; c) MeP(Ph)₃·Br, KHMDS, THF, -78°C → RT, 56%; d) HCl, MeOH/H₂O, quant.; e) NsCl, NEt₃, DMAP, DCM, 85%; f) 4-bromobut-1-ene, K₂CO₃, DMF, 70 °C, 80%; g) Grubbs' 1st gen., DCM, 40 °C, 62%; h) PhSH, NaOH, ACN/H₂O, 50 °C, 99%; i) RuCl₃(H₂O)₃, NaBH₄, DCE/MeOH, 87%; j) TMS-N₃, 90 °C, 68%; k) 4-F-PhMgBr, THF, 0 °C, 94%; l) *i*) **15**, triphosgene, Na₂CO₃, DCM; *ii*) **18**, DIPEA, DMAP, THF, 66 °C, 1.5% (**1**) 1.2% (**2**); m) MeOH, SOCl₂, 65 °C, 97%; n) 4-F-PhMgBr, THF, 0 °C, 89%; o) *i*) **15**, triphosgene, Na₂CO₃, DCM; *ii*) **21**, DIPEA, DMAP, THF, 66 °C, 35%; p) KMnO₄, H₂O, 100 °C; q) H₂SO₄, EtOH, 78 °C, 47% (two steps); r) 4-F-PhMgBr, THF, 0 °C, 83%; s) *i*) **15**, triphosgene, Na₂CO₃, DCM; *ii*) **25**, DIPEA, DMAP, THF, 66 °C, 21%; t) TrtCl, TEA, DCM, 0 °C → RT, 97%; u) 4-F-PhMgBr, THF, 0 °C, 86%; v) TFA, H₂O, DCM, 56%; w) *i*) **15**, triphosgene, Na₂CO₃, DCM; *ii*) **29**, DIPEA, DMAP, THF, 66 °C, 27%.

To determine the binding kinetics of the inhibitors we adapted the previously reported para-nitrophenolbutyrate (PNPB) assay.[32] The typical pre-incubation step with the inhibitors was omitted and enzyme activity was continuously measured from the start (t = 0). To obtain a higher specific signal at early time points, the substrate concentration was increased to 600 μM. Furthermore, the enzyme was pre-mixed with the assay buffer before addition to a 96-well plate, which contained concentrated inhibitor and substrate, to minimize initial mixing effects. This yielded reproducible substrate conversion curves (Figure 4.2).

Most available literature models, including the standard observed rate approximation (k_{obs}), assume that the enzyme concentration will not change during the incubation

($K_i \gg [E]$).[14,33-35] These models can, however, not be applied to potent inhibitors, such as KT109 and DH376, that will decrease the enzyme concentration. Therefore, the kinetic model of Schwartz *et al.* was selected to fit to the substrate conversion curves.[10] In this model, DynaFit software is used for the numerical fitting of the full set of differential equations governing the substrate conversion curves without making the $K_i \gg [E]$ assumption.[10,36,37] The kinetic model from Schwartz *et al.* was slightly adapted to incorporate the spontaneous enzyme inactivation observed for blank measurements, where substrate depletion cannot explain the decrease in substrate conversion rate. As the substrate concentration was well below the predicted K_M, the one-step substrate conversion proposed by Schwartz *et al.* was maintained. Initial values for the required rate constants were derived from several preliminary experiments, and were mostly left to be optimized by the algorithm (Figure S4.1).

The assay was validated using irreversible DAGL inhibitors (KT109 and DH376) and a reversible inhibitor (LEI105) (Figure 4.1C). All three inhibitors were previously found to be highly active with (sub)nanomolar potency (pIC_{50} 8.6 to 9).[20,22,38] Using seven inhibitor concentrations around their reported IC_{50}-values, a set of substrate conversion curves was generated. These curves were fitted with DynaFit (Figure S4.1). The resulting fits and values for K_i and k_{inact} are shown in Figure 4.2. As was expected, the model does not find a fit for the k_{inact} value for the covalent but reversible inhibitor LEI105. The found K_i-values (all between 0.2 and 0.4 nM) were generally in line with the high potency described in literature, although LEI105 was somewhat more active than previously reported.[22] The inactivation rates for KT-109 and DH376 were similar (k_{inact} = ± 0.07 min^{-1}).
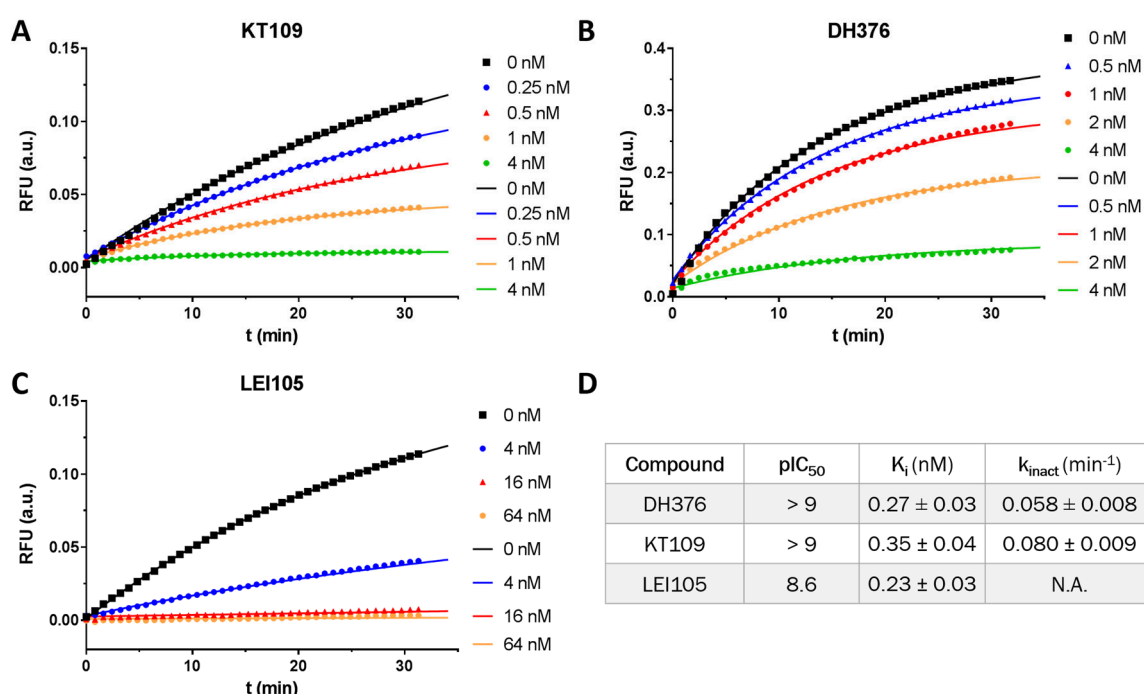
Figure 4.2 | Kinetics of binding of KT109 (A), DH376 (B) and LEI105 (C), data fits are summarized in (D). For clarity not all substrate conversion curves used for the curve fitting are shown. Markers denote measured absorbance, lines denote fitted model.

Next, we tested inhibitors **1-5** to determine the influence of the leaving group on their potency in the standard surrogate substrate assay. A large range in IC_{50}-values was observed (Figure 4.3A, Table 4.1). Both regioisomers **1** and **2** were low nanomolar inhibitors, whereas the 1,2,4-triazole (**3**) had a reduced potency (IC_{50} = 3.4 µM) and the imidazole (**5**) was inactive (IC_{50} > 10 µM). Pyrazole (**4**) had an intermediate potency with IC_{50} of 0.21 µM. The IC_{50}-values were used to guide the selection of inhibitor concentrations for the kinetic assay (Figure 4.3B, Table 4.1). For all but the imidazole compound **5** we were able to determine the kinetic parameters. Intriguingly, the inactivation rates for **1** and **2** (k_{inact} = 0.22 and 0.32 min^{-1}, respectively), were 2-3 times higher than for KT109 and DH376, but they had a lower binding affinity (Ki = 10 and 339 nM, respectively). Unexpectedly, the inactivation rates for **3** and **4** were comparable to DH376 and KT109, whereas there is 10,000-fold difference in pK_a between these heterocycles (Table 4.1). Yet, the binding affinity of **3** and **4** was substantially reduced (Ki > 10 µM), thereby explaining their higher IC_{50}-values.
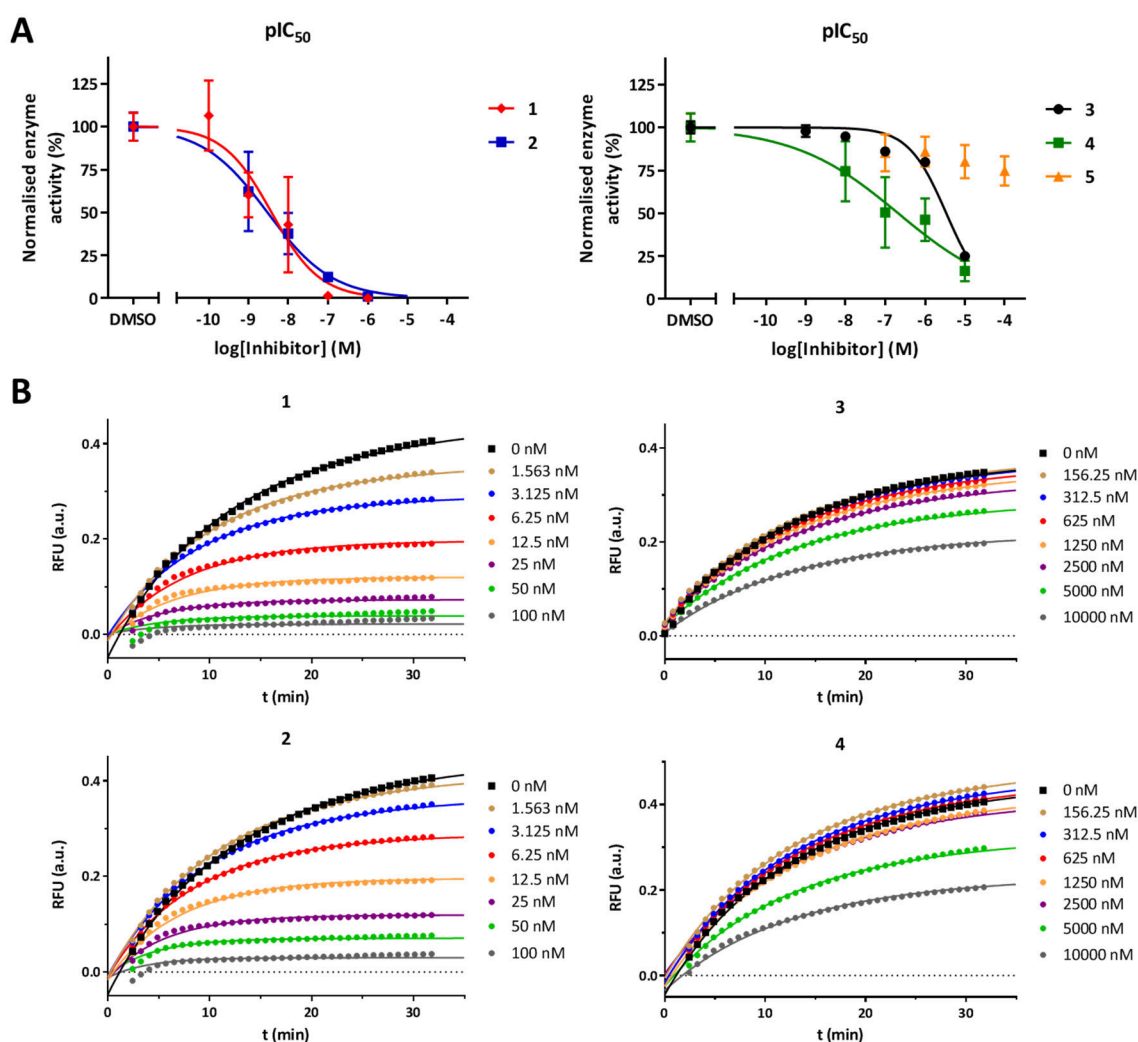
Figure 4.3 | pIC_{50} determination of compounds **1-5** (A) and kinetic fits (B). All data points are measured as N = 4, positive controls (DMSO) as N = 8. Markers denote mean values, error bars denote the SEM, lines are fitted data models.

Table 4.1 | Potency and kinetic parameters of the focused set of DH376 derivatives 1-5.

Compound	pIC_{50}	K_i (nM)	k_{inact} (min^{-1})	pK_a
1	8.52 ± 0.27	10.4 ± 2.0	0.22 ± 0.03	9.3[31]
2	8.42 ± 0.28	339 ± 55	0.32 ± 0.04	9.3[31]
3	5.47 ± 0.07	13770 ± 910	0.075 ± 0.005	10.0[29]
4	6.68 ± 0.41	10800 ± 910	0.080 ± 0.005	14.2[30]
5	< 4.5	N.D.	N.D.	14.4[30]
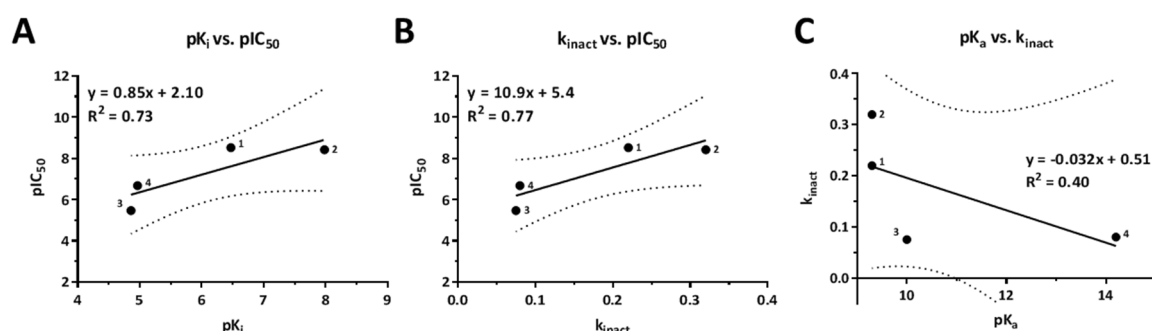
Figure 4.4 | pIC_{50} correlates with both pK_i (A) and k_{inact} (B) but pK_a (C) does not correlate with reactivity for the inhibitors 1-4. Dashed lines denote 90% confidence interval.

Discussion

We developed a surrogate substrate based assay to determine the kinetic parameters of binding and reactivity of triazole urea inhibitors of the serine hydrolase DAGL-α. Having the ability to discern the kinetics of binding should enable the optimization of the affinity of the inhibitors for the enzyme, the K_i, while controlling the reactivity k_{inact}, to minimize the off-target reactivity. The assay was validated using three well-characterized published DAGL inhibitors. Five DH376 derivatives were synthesized to study the role of the leaving group in the affinity and reactivity with DAGL-α. The IC_{50}-values correlated with the K_i and k_{inact} with an R^2 of 0.73 and 0.77, respectively (Figure 4.4A, B). The main reason for the large differences observed in IC_{50}-values for the five inhibitors was the strong reduction in binding affinity for the pyrazole and 1,2,4-triazole compounds compared to the 1,2,3-triazole inhibitors. Unexpectedly, it was shown that the 1,2,4-triazole 3 and pyrazole 4 inhibitors were as reactive as DH376 and KT109, which is in stark contrast to the five orders of magnitude difference in pK_a. This showed that the leaving group acidity does not correlate with the rate of inactivation k_{inact} (Figure 4.4C).

The triazole heterocycles of DH376 and KT109 could be involved in hydrogen bonds with active site amino acids, as has been previously shown for α-positioned heterocycles for related serine hydrolases like FAAH and others.[39,40] From our kinetic investigation we indeed conclude that the position of the nitrogen atoms in the ring is of great importance for the formation and strength of these bonds. Despite the lack of knowledge of the structure of DAGL-α, we could speculate that a major hydrogen bonding interaction has to occur from the direction in which His650 is drawn in Figure 4.5 (directionality relative to the

heterocycle). His650 is part of the catalytic triad and has been postulated to bind in this manner before, based on the covalent docking of α-ketoheterocycles in a homology model of DAGL-α.[41] The 1,2,3-triazoles would both be capable of forming this bond. Compounds 3-5 could not form this interaction, which would explain the strong drop in binding affinity. A second weaker hydrogen bonding interaction on the other side of the urea functionality could explain the further differences observed. A candidate hydrogen bond donor would be His471 (drawn in Figure 4.5).[41] Compound 1 would be optimally positioned to bind both, leading to the low K_i observed. The *N1*-isomer could only bind His650, leading to a slight drop in potency. Compounds 3 and 4 could both bind only through the weaker second interaction, leading to a poor overall binding. Imidazole 5 would be unable to pick up either bond. The proposed hydrogen bonding interactions are unable to directly explain the slight difference in binding between 3 and 4, though the specific electronic properties of the azoles (i.e. the efficiency of the hydrogen bond formation) could be a cause.
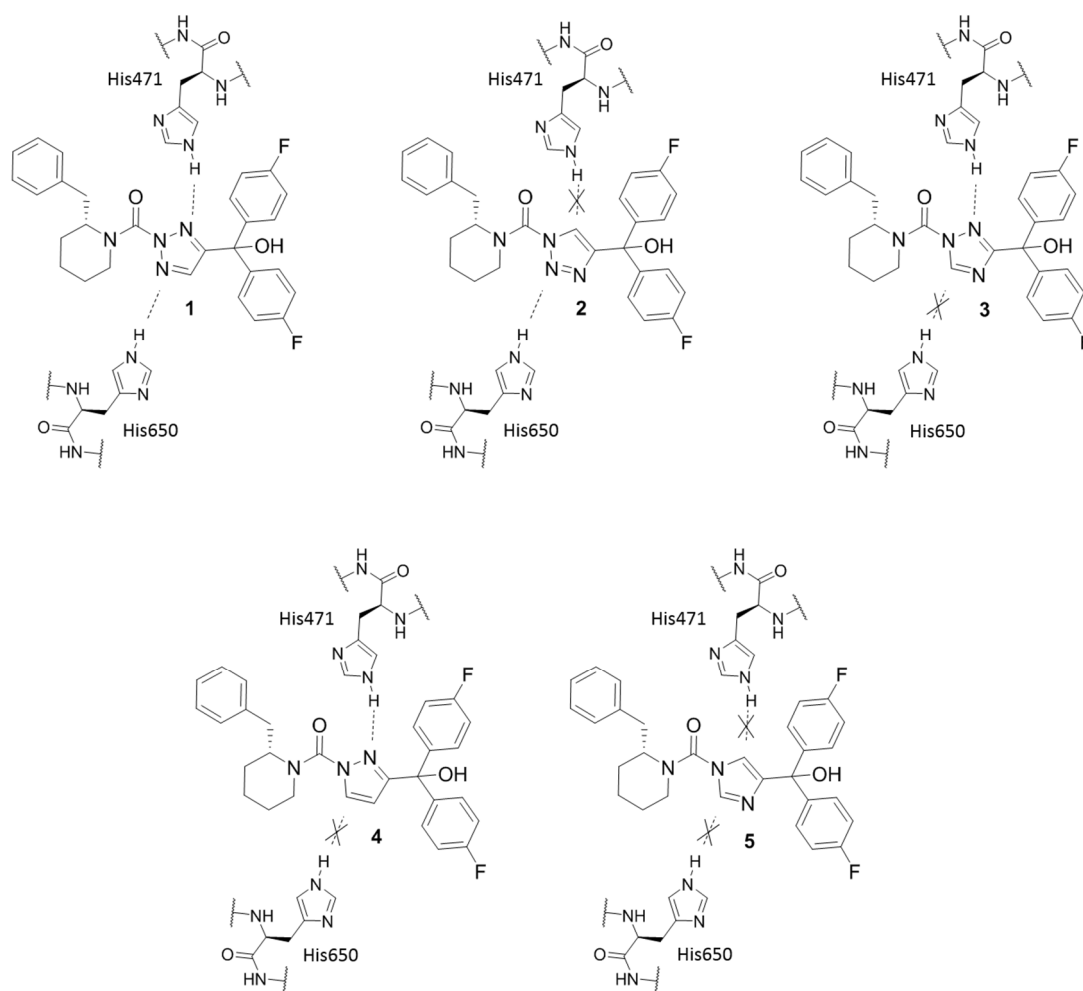
Figure 4.5 | Illustration of the hypotheses using potential hydrogen bonding to explain the difference in K_i. His650 is hypothesized to form the major hydrogen bond; His471 might form a secondary hydrogen bonding interaction.

The assay presented here, in combination with the data analysis through numerical fitting, could be translated to work on a multitude of serine hydrolases. As long as a sensitive and robust (surrogate) substrate assay is available that can be interrogated in a time-dependent manner it should in principle be possible to derive structure kinetics relationships. These relationships provide important insights into the mode of action and can aid in the optimization of covalent serine hydrolase inhibitors in an affinity directed manner. For the DAG lipase inhibitors, this may lead to more selective inhibitors more suitable for further *in vivo* target validation studies.

Conclusion

To conclude, we have developed a kinetic assay to study the influence of the heterocyclic core of triazole ureas as covalent, mechanism-based inhibitors of diacylglycerol lipase-α. We found that the pK_a of the leaving group did not correlate with the reactivity of the inhibitors, but that the position of the nitrogen atom in the heterocycle is of importance in its binding affinity. Detailed knowledge of structure kinetic relationships is expected to guide the optimization of more selective and well-balanced irreversible inhibitors of serine hydrolases.

Acknowledgements

Methods

Chemical biology methods

Cell culture and membrane preparation

HEK293T cells were grown in DMEM with stable glutamine and phenolred (PAA) with 10% New Born Calf serum, penicillin and streptomycin. Cells were passaged every 2-3 days by resuspending in medium and seeding them to appropriate confluence. Membranes were prepared from transiently transfected HEK293T cells. One day prior to transfection 10^7 cells were seeded in a 15 cm petri dish. Cells were transfected by the addition of a 3:1 mixture of polyethyleneimine (60 μg) and plasmid DNA (20 μg) in 2 mL serum free medium. The medium was refreshed after 24 hours, and after 72 h the cells were harvested by suspending them in 20 mL medium. The suspension was centrifuged for 10 min at 1000 g, and the supernatant was removed. The cell pellet was stored at -80 °C until use.

Cell pellets were thawed on ice and suspended in lysis buffer (20 mM Hepes, 2 mM DTT, 0.25 M sucrose, 1 mM MgCl$_2$, 25 U/mL benzonase). The suspension was homogenized by polytrone (3x 7 sec) and incubated for 30 min on ice. The suspension was subjected to ultracentrifugation (93,000 g, 30 min, 4 °C, Beckman Coulter, Type Ti70 rotor) to yield the cytosolic fraction in the supernatant and the membrane fraction as a pellet. The pellet was resuspended in storage buffer (20 mM Hepes, 2 mM DTT). The protein concentration was determined with Quick Start Bradford assay (Biorad). The protein fractions were diluted to a total protein concentration of 1 mg/mL and stored in small aliquots at -80 °C until use.

Surrogate substrate assay

The biochemical mDAGL-α activity assay is based on the method previously described.[32] 100 μL reactions were performed in flat bottom Greiner 96-wells plates in a 50 mM pH 7.2 Hepes buffer. Membrane protein fractions from HEK293T cells transiently transfected with mDAGL-α (0.05 μg/μL final concentration) were used as mDAGL-α source. Inhibitors were introduced in 2.5 μL DMSO. The mixtures were incubated for 20 minutes before 5.0 μL 6 mM PNP-butyrate (final concentration 0.3 mM) in 50% DMSO was added (final DMSO concentration 5.0%). Reactions were allowed to progress for 30 minutes at 20 °C before OD (420 nm) was measured using a TECAN GENios plate reader. All experiments were performed at N=2, n=2 for experimental measurements and N=2, n=4 for controls.

Data analysis: Z'-factor of each plate was determined for the validation of each experiment, using the following formula Z'=1-3(σ_{pc}+ σ_{nc})/(μ_{pc} - μ_{nc}). The OD from the positive control (pc: DAGL DMSO), and the negative control (nc: 10 μM THL) was used. Plates were accepted for further analysis when Z'>0.6. Measurements were corrected for the average absorption of the negative control (10 μM THL). The average, standard deviation (SD) and standard error of mean (SEM) were calculated and normalized to the corrected positive control. Data was exported to Graphpad Prism 7.0 for the calculation of the pIC$_{50}$ using a non-linear dose-response analysis.

DynaFit setup

DynaFit version 4 was used with an academic license. For batch processing the command line interface was used. The raw data were pre-processed using Microsoft Excel 2016 and exported to tab delimited text files for use with DynaFit. Scripts were generated manually or using a purpose-made python script. An example DynaFit script is shown in Figure S4.1.

The contents of the header [task] follow directly from the DynaFit manual and simply state that the program should fit progress curves using the data supplied.

The section [mechanism] was built based on standard enzyme kinetics. The simplified hit and run (`E + S ---> E + P`) was used for the substrate conversion as the K$_M$ of the surrogate substrate is too high to be determined reliably experimentally.[10,32] The additional enzyme degradation step (`E ---> E*`) was included as the progress rate curves for the DMSO blanks decreased more than could be explained by the reduction in substrate concentration (which is accounted for in the set of differential equations fitted).

The rate constants defined in [constants] were set empirically but are all left to be optimized. The exception is k$_{on}$, which is fixed to 100,000 μM^{-1}·min^{-1}, but the variable itself is dependent on k$_{off}$ (and vice versa), so only the ratio of the two (K$_i$) is physically meaningful in this experimental setup.

As the enzyme is obtained by overexpression in HEK293T cells the exact concentration is unknown. Data from the previously published PNPB-based assay used for pIC$_{50}$ determination indicate that the assay limit lies around 9, which puts the enzyme concentration at ± 1 nM. It is left to be optimised by DynaFit.

The value for P given in the [responses] section is essentially the absorption coefficient of the converted surrogate substrate in AU·μM^{-1}, which was determined experimentally (Figure S4.2).

Chemistry

General remarks

All reactions were performed using oven- or flame-dried glassware and dry (molecular sieves) solvents. Reagents were purchased from Alfa Aesar, Sigma-Aldrich, Acros, and Merck and used without further purification unless noted otherwise. All moisture sensitive reactions were performed under an argon or nitrogen atmosphere.

^{1}H and ^{13}C NMR spectra were recorded on a Bruker DPX-300 (300 MHz), AV-400 (400 MHz) or DRX-500 (500 MHz). Used software for interpretation of NMR-data was Bruker TopSpin 1.3 and MestreNova 11.0. Chemical shift values are reported in ppm with tetramethylsilane or solvent resonance as the internal standard (CDCl$_3$: δ 7.26 for ^{1}H, δ 77.16 for ^{13}C; ACN-d3: δ 1.94 for ^{1}H, δ 1.32 for ^{13}C; MeOD: δ 3.31 for ^{1}H, δ 49.00 for ^{13}C).[42] Data are reported as follows: chemical shifts (δ), multiplicity (s = singlet, d = doublet, dd = double doublet, td = triple doublet, t = triplet, q = quartet, bs = broad singlet, m = multiplet), coupling constants J (Hz), and integration.

Liquid chromatography was performed on a Finnigan Surveyor LC/MS system, equipped with a C18 column. Flash chromatography was performed using SiliCycle silica gel type SiliaFlash P60 (230–400 mesh). TLC analysis was performed on Merck silica gel 60/Kieselguhr F254, 0.25 mm. Compounds were visualized using KMnO$_4$ stain (K$_2$CO$_3$ (40 g), KMnO$_4$ (6 g), and water (600 mL)) or CAM stain (Ce(NH$_4$)$_4$(SO$_4$)$_4$·2H$_2$O (ceric ammonium sulfate: 10 g); ammonium molybdate (25 g); conc. H$_2$SO$_4$ (100 mL); H$_2$O (900 mL)). Preparative HPLC (Waters, 515 HPLC pump M; Waters, 515 HPLC pump L; Waters, 2767 sample manager; Waters SFO System Fluidics Organizer; Waters Acquity Ultra Performance LC, SQ Detector; Waters Binary Gradient Module) was performed on a Waters XBridgeTM column (5 μM C18, 150 x 19 mm). Diode detection was done between 210 and 600 nm. Gradient: ACN in (H$_2$O + 0.2% TFA). Chiral HPLC analysis was performed after benzoylation of the free amine on a Daicel Chiralpak AD column (250 x 4.5 mm, 10 μm particle size) using 10 % isopropyl alcohol in hexane as eluent (1.0 mL/min, UV-detection at 254 nm) Rt= 15.1 min (R-enantiomer 12.8 min). High resolution mass spectra (HRMS) were recorded by direct injection on a q-TOF mass spectrometer (Synapt G2-Si) equipped with an electrospray ion source in positive mode with Leu-enkephalin (m/z = 556.2771) as an internal lock mass. The instrument was calibrated prior to measurement using the MS/MS spectrum of Glu-1-fibrinopeptide B.
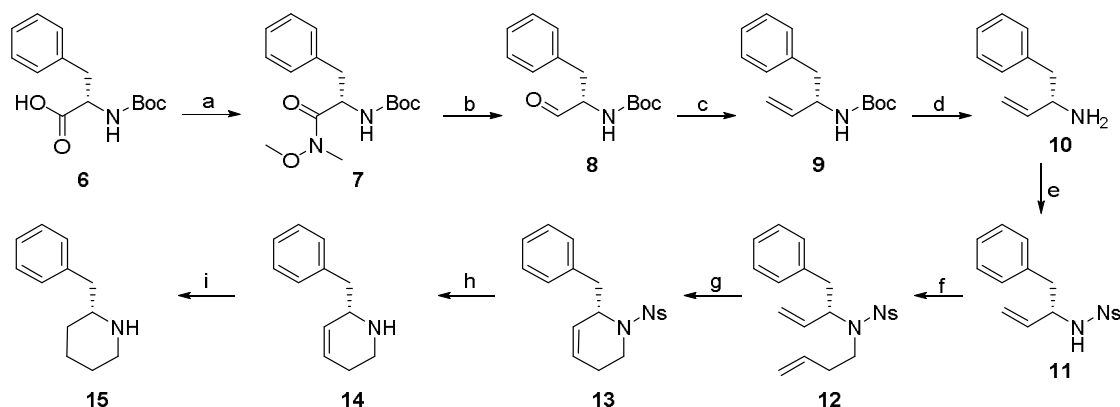
Scheme S4.1 | Synthesis of (*R*)-2-benzylpiperidine. Reagents and conditions: a) *N,O*-di-Me-hydroxylamine·HCl, EDCl·HCl, DCM, 0°C → RT, 92%; b) LiAlH$_4$, THF, -20°C, 96%; c) MeP(Ph)$_3$.Br, KHMDS, THF, -78 °C → RT, 56%; d) HCl, MeOH/H$_2$O, quant.; e) NsCl, NEt$_3$, DMAP, DCM 85%; f) 4-bromobut-1-ene, K$_2$CO$_3$, DMF; 70°C, 80%; g) Grubbs' 1st gen., DCM, 40°C, 62%; h) PhSH, NaOH, ACN/H$_2$O, 50°C, 99%; i) RuCl$_3$(H$_2$O)$_3$, NaBH$_4$, DCE/MeOH, 87%.

tert-Butyl (*S*)-(1-(methoxy(methyl)amino)-1-oxo-3-phenylpropan-2-yl)carbamic acid (**7**)
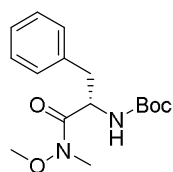
N-Boc-L-phenylalanine **6** (24.16 g, 91 mmol), *N,O*-dimethylhydroxylamine hydrochloride (9.95 g, 102 mmol) and 4-methylmorpholine (11.25 mL, 102 mmol) were dissolved in dichloromethane and cooled to 0 °C. EDCl-HCl (18.75 g, 98 mmol) was added in three portions with a 15 min interval. After consumption of the starting material, the reaction mixture was washed with sat. aq. NH$_4$Cl, sat. aq. NaHCO$_3$ and brine. The organic phase was dried (MgSO$_4$), filtered and concentrated under reduced pressure to obtain the pure title compound as a honey like oil (25.9 g, 84 mmol, 92%).^{1}H NMR (400 MHz, CDCl$_3$) δ 7.36 – 7.06 (m, 7H), 5.04 – 4.85 (m, 1H), 3.65 (s, 3H), 3.17 (s, 3H), 3.09 – 2.96 (m, 1H), 2.94 – 2.80 (m, 1H), 1.37 (s, 11H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 155.27, 136.70, 129.56, 128.45, 126.86, 79.67, 61.65, 51.62, 43.93, 38.97, 28.42.

tert-Butyl (*S*)-(1-oxo-3-phenylpropan-2-yl)carbamate (**8**)
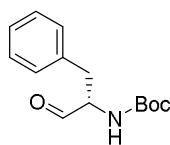
7 (25.9 g, 84 mmol) was dissolved in THF and cooled to -15 °C. Subsequently a 1 M solution of LiAlH₄ in THF (42 mL, 42 mmol) was added slowly. Upon completion, the reaction was quenched with 25% aq. KHSO₄, allowed to warm up to RT and stirred vigorously. Ethyl acetate was added, and the organic phase was separated, washed with sat. aq. NaHCO₃ and brine, dried (MgSO₄), filtered and concentrated under reduced pressure to give the title compound as a white solid. (20.2 g, 81 mmol, 96%). ^{1}H NMR (400 MHz, CDCl₃) δ 9.63 (s, 1H), 7.43 – 7.04 (m, 5H), 5.18 – 4.99 (m, 1H), 4.43 (q, *J* = 6.8 Hz, 1H), 3.12 (d, *J* = 6.7 Hz, 2H), 1.41 (s, 9H). ^{13}C NMR (101 MHz, CDCl₃) δ 199.56, 135.90, 129.46, 128.88, 127.19, 80.33, 60.91, 35.58, 28.39.

tert-Butyl (*S*)-(1-phenylbut-3-en-2-yl)carbamate (**9**)
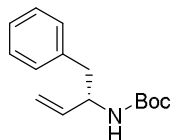
To a solution of potassium bis(trimethylsilyl)amide (27 g, 135 mmol) in dry THF was added methyltriphenylphosphonium bromide (54.8 g, 153 mmol). The resulting mixture was stirred for 3 hr and subsequently cooled to -78 °C, after which a solution of **8** (22.5 g, 90 mmol) in THF was added. The reaction was allowed to warm up to RT overnight and quenched with sat. aq. NH₄Cl. EtOAc was added and the organic phase was separated, washed with sat. aq. NaHCO₃ and brine, dried (MgSO₄), filtered and concentrated under reduced pressure. The residue was purified with silica chromatography (8% ethyl acetate in petroleum ether) to give the desired product as a white, waxy solid (12.6 g, 50.9 mmol, 56%). ^{1}H NMR (400 MHz, CDCl₃) δ 7.34 – 7.13 (m, 5H), 5.88 – 5.71 (m, 1H), 5.16 – 5.03 (m, 2H), 4.45 (d, *J* = 27.0 Hz, 2H), 2.83 (d, *J* = 6.5 Hz, 2H), 1.40 (s, 9H). ^{13}C NMR (101 MHz, CDCl₃) δ 155.33, 138.18, 137.53, 129.68, 128.44, 126.59, 114.84, 79.53, 53.58, 41.60, 28.47.

(*S*)-1-Phenylbut-3-en-2-amine (**10**)
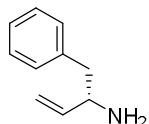
9 (9.7 g, 39.2 mmol) was dissolved in absolute methanol and 4 M aq. HCl was added. After TLC showed full consumption of the starting material, the mixture was concentrated under reduced pressure and diluted with water. The solution was washed with diethyl ether and aq. NaOH was used to make the aqueous layer basic (pH > 12). The aqueous phase was extracted 3x with chloroform. The organic layers were combined, dried (MgSO₄), filtered and concentrated under reduced pressure to give the title compound in quantitative yield. Epimerised product that was formed during Wittig olefination was removed by recrystallization form toluene/n-propyl alcohol with *N*-acetyl-ʟ-leucine according to literature procedure,[43] yielding the amine as a yellow oil with an enantiomeric ratio of 97:3 as determined by chiral HPLC (*vide infra*). ^{1}H NMR (400 MHz, CDCl₃) δ 7.36 – 7.12 (m, 5H), 5.89 (ddd, *J* = 16.8, 10.3, 6.2 Hz, 1H), 5.14 (dt, *J* = 17.2, 1.5 Hz, 1H), 5.04 (dt, *J* = 10.3, 1.4 Hz, 1H), 3.64 – 3.52 (m, 1H), 2.83 (dd, *J* = 13.3, 5.4 Hz, 1H), 2.62 (dd, *J* = 13.3, 8.3 Hz, 1H), 1.31 (s, 3H). ^{13}C NMR (101 MHz, CDCl₃) δ 142.49, 138.86, 129.51, 128.50, 126.45, 113.77, 55.57, 44.43.

	Before recrystallization			After recrystallization	
	3.315			2.833 4.063	
	9.8			12.802	
	12.827 13.679			14.998	
	15.142			23.425	

CHROMATOGRAM 1 MEMORIZED

C-R5A CHROMATOPAC
CHANNEL NO 1
SAMPLE NO 0
REPORT NO 95

FILE
METHOD

PKNO	TIME	AREA	MK	IDNO	CONC
1	12.827	539966 S			12.8909
		33.9sec			
2	15.142	3648758			87.109
		44.3sec			
	TOTAL	4188724			100

CHROMATOGRAM 1 MEMORIZED

C-R5A CHROMATOPAC
CHANNEL NO 1
SAMPLE NO 0
REPORT NO 96

FILE
METHOD

PKNO	TIME	AREA	MK	IDNO	CONC
1	12.802	334130			2.6044
		32.2sec			
2	14.998	12495560			97.3956
		48.2sec			
	TOTAL	12829690			100

(S)-2-Nitro-N-(1-phenylbut-3-en-2-yl)benzenesulfonamide (11)

10 (3.5 g, 23.8 mmol), triethylamine (4.88 mL, 35.7 mmol), *N,N*-dimethylaminopyridine (1.45 g, 11.9 mmol) and 2-nitrobenzenesulphonyl chloride (6.85 g, 30.9 mmol) were dissolved in dry DCM and stirred overnight, after which the reaction mixture was concentrated under reduced pressure. Flash column chromatography (20% ethyl acetate in petroleum ether) yielded the desired product with trace impurities as an orange oil (6.7 g, 20 mmol, 85%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.01 – 7.90 (m, 1H), 7.83 – 7.71 (m, 1H), 7.69 – 7.58 (m, 2H), 7.18 – 7.01 (m, 5H), 5.81 – 5.64 (m, 1H), 5.40 (d, *J* = 7.9 Hz, 1H), 5.13 (dt, *J* = 17.1, 1.2 Hz, 1H), 5.03 (dt, *J* = 10.3, 1.1 Hz, 1H), 4.32 – 4.17 (m, 1H), 2.92 (dd, *J* = 13.8, 6.1 Hz, 1H), 2.79 (dd, *J* = 13.8, 7.8 Hz, 1H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 137.21, 136.22, 134.84, 133.22, 132.92, 130.79, 129.46, 128.54, 127.03, 125.53, 116.57, 58.68, 42.16.

(S)-N-(But-3-en-1-yl)-2-nitro-N-(1-phenylbut-3-en-2-yl)benzenesulfonamide (12)
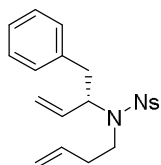
11 (6.7g, 20.2 mmol) was dissolved in DMF and K$_2$CO$_3$ (11.1 g, 81 mmol) and 4-bromo-1-butene (2.46 mL, 24.2 mmol) were added. The mixture was heated to 70 °C and stirred vigorously. After 24 hr, an extra portion of 4-bromo-1-butene (2.46 mL, 24.2 mmol) was added and the resulting mixture was stirred for 72 h. The reaction was allowed to cool down to RT, diluted with brine and extracted 3x with EtOAc. Combination, drying (MgSO$_4$), filtration and concentration of the organic phases yielded the desired product as a brown oil (6.24 g, 16.2 mmol, 80%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.92 – 7.84 (m, 1H), 7.69 – 7.52 (m, 3H), 7.25 – 7.12 (m, 5H), 5.89 – 5.64 (m, 2H), 5.17 – 5.02 (m, 4H), 4.72 – 4.62 (m, 1H), 3.50 – 3.31 (m, 2H), 3.08 (dd, *J* = 13.5, 5.9 Hz, 1H), 2.95 (dd, *J* = 13.5, 9.2 Hz, 1H), 2.50 – 2.31 (m, 2H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 137.45, 135.28, 134.61, 134.09, 133.45, 131.70, 130.88, 129.39, 128.55, 126.75, 124.29, 118.98, 117.34, 61.66, 44.69, 39.60, 35.63.

(S)-6-Benzyl-1-((2-nitrophenyl)sulfonyl)-1,2,3,6-tetrahydropyridine (13)
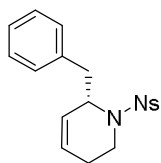
A solution of **12** (6.24 g, 16.2 mmol) in DCM was purged with argon and 1st generation Grubbs catalyst (400 mg, 3 mol%) was added. The mixture was heated to 40 °C and stirred overnight. Volatiles were removed under reduced pressure and flash column chromatography (40% diethyl ether in petroleum ether) yielded the desired product as brown powder (3.57 g, 9.96 mmol, 62%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.85 (d, *J* = 7.7 Hz, 1H), 7.69 – 7.48 (m, 3H), 7.25 – 7.09 (m, 5H), 5.87 – 5.74 (m, 1H), 5.62 (d, *J* = 10.6 Hz, 1H), 4.56 (s, 1H), 3.93 (dd, *J* = 14.4, 6.2 Hz, 1H), 3.15 (ddd, *J* = 15.2, 11.7, 4.1 Hz, 1H), 3.00 (dd, *J* = 13.1, 6.0 Hz, 1H), 2.89 (dd, *J* = 13.1, 8.3 Hz, 1H), 2.21 – 2.04 (m, 1H), 1.93 (dt, *J* = 18.1, 5.0 Hz, 1H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 137.28, 134.48, 133.30, 131.84, 130.44, 129.69, 128.50, 127.27, 126.79, 125.61, 124.27, 55.78, 41.71, 39.11, 24.33.

(S)-6-Benzyl-1,2,3,6-tetrahydropyridine (14)
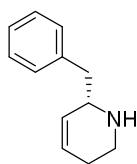
To a solution of thiophenol (2.56 mL, 24.9 mmol) in acetonitrile cooled with ice was added a 2 M aq. solution of NaOH (12.45 mL, 24.9 mmol). After stirring for 10 minutes, the ice bath was removed and a solution of **13** (3.57 g, 9.96 mmol) in acetonitrile was added slowly. The resulting mixture was heated to 50 °C. When TLC showed full conversion of the starting material, the reaction was cooled to RT and diluted with aq. HCl so that the pH was below 2. The aqueous layer was washed with Et$_2$O and diluted with aq. NaOH until the pH was above 12. It was then extracted 3x with ethyl acetate. The organic layers were combined, dried (MgSO$_4$), filtered and concentrated to afford the desired product as yellow oil (1.71 g, 9.86 mmol, 99%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.37 – 7.15 (m, 5H), 5.84 – 5.75 (m, 1H), 5.64 (dq, *J* = 10.1, 2.0 Hz, 1H), 3.61 – 3.51 (m, 1H), 3.17 – 2.99 (m, 1H), 2.85 – 2.75 (m, 2H), 2.70 (dd, *J* = 13.2, 8.8 Hz, 1H), 2.26 – 2.12 (m, 1H), 2.02 – 1.91 (m, 1H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 138.97, 130.25, 129.39, 128.57, 126.43, 126.22, 55.47, 42.56, 42.12, 25.86.

(R)-2-Benzylpiperidine (15)
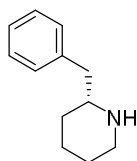
In a three-neck flask containing two stoppers and one septum with empty balloons, **14** (1.70 g, 9.81 mmol) was dissolved in dichloroethane/methanol 10:3. This solution was purged with argon and cooled to 0 °C, after which RuCl$_3$(H$_2$O)$_3$ (257 mg, 0.98 mmol) was added. NaBH$_4$ (1.86 g, 49 mmol) was added quickly while capturing the formed H$_2$ gas in the empty balloons, thus keeping the reaction under hydrogen atmosphere. The reaction was allowed to warm up to RT and stirred overnight. Aqueous HCl was then added, so that the water layer had a pH of

below 2. The aqueous layer was washed with Et$_2$O and diluted with aq. NaOH until the pH was above 12. The water layer was extracted thrice with ethyl acetate. The organic layers were combined, dried (MgSO$_4$), filtered and concentrated to yield the title compound as a yellow waxy solid (1.5 g, 8.6 mmol, 87%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.38 – 7.03 (m, 5H), 3.08 – 2.89 (m, 1H), 2.77 – 2.64 (m, 2H), 2.64 – 2.46 (m, 2H), 1.84 – 1.73 (m, 1H), 1.73 – 1.63 (m, 1H), 1.63 – 1.52 (m, 1H), 1.52 – 1.38 (m, 1H), 1.37 – 1.15 (m, 3H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 139.12, 129.23, 128.39, 126.20, 58.26, 47.08, 43.81, 32.77, 26.08, 24.80.

Methyl 1*H*-1,2,3-triazole-4-carboxylate (**17**)
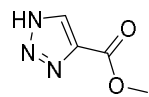
This protocol was based on literature procedure.[44] A mixture of azidotrimethylsilane (2.6 mL, 20 mmol) and methyl propiolate **16** (1.8 mL, 20 mmol) was heated for 4 h at 90 °C, concentrated and coevaporated with MeOH to yield the title compound as a white solid (1.74 g, 14 mmol, 68%). ^{1}H NMR (400 MHz, MeOD) δ 8.35 (s, 1H), 3.92 (s, 3H). ^{13}C NMR (101 MHz, MeOD) δ 162.61, 139.63, 131.92, 52.53.

Bis(4-fluorophenyl)(1*H*-1,2,3-triazol-4-yl)methanol (**18**)
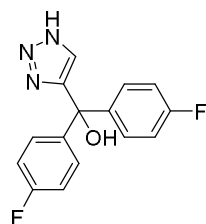
17 (100 mg, 0.787 mmol) was dissolved in THF and cooled to 0 °C. Under vigorous stirring, a 2M solution of 4-fluorophenylmagnesium bromide in Et$_2$O (1.38 mL, 2.75 mmol) was added dropwise. The reaction mixture was allowed to warm up to RT and stirred overnight. The reaction was quenched with sat. aq. NH$_4$Cl. The aqueous phase was extracted with DCM (3x). The combined organic layers were dried (MgSO$_4$), filtered and concentrated under reduced pressure. The residue was purified using flash column chromatography (40% to 60% ethyl acetate in pentane) in order to obtain the title compound as a white solid (210 mg, 0.731 mmol, 94%). ^{1}H NMR (400 MHz, MeOD) δ 7.60 (s, 1H), 7.45 – 7.33 (m, 4H), 7.15 – 6.98 (m, 4H).

Methyl 1*H*-1,2,4-triazole-3-carboxylate (**20**)
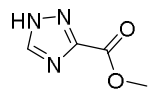
1*H*-1,2,4-triazole-3-carboxylic acid **19** (250 mg, 2.21 mmol) was dissolved in MeOH (50 mL) and cooled to 0 °C. Thionyl chloride (0.48 ml, 6.6 mmol) was slowly added to the solution. The mixture was then heated to reflux for 3 h after which it was cooled to RT and concentrated *in vacuo* to yield the title compound (271 mg, 2.14 mmol, 97%) as a white solid. ^{1}H NMR (400 MHz, MeOD) δ 9.24 (s, 1H), 4.05 (s, 3H).

Bis(4-fluorophenyl)(1*H*-1,2,4-triazol-3-yl)methanol (**21**)
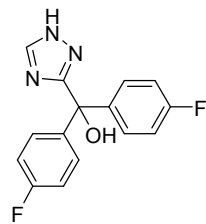
20 (100 mg, 0.787 mmol) was dissolved in THF and cooled to 0 °C. Under vigorous stirring, a 2 M solution of 4-fluorophenylmagnesium bromide in Et$_2$O (1.37 mL, 2.75 mmol) was added dropwise. The reaction mixture was allowed to warm up to RT and stirred overnight. The reaction was quenched with sat. aq. NH$_4$Cl. The aqueous phase was extracted with ethyl acetate (3x). The combined organic layers were dried (MgSO$_4$), filtered and concentrated under reduced pressure, yielding the title compound as an off-white solid (201 mg, 0.700 mmol, 89%). ^{1}H NMR (400 MHz, CD$_3$CN) δ 8.21 (s, 1H), 7.40 (dd, *J* = 8.7, 5.4 Hz, 4H), 7.08 (t, *J* = 8.7 Hz, 4H). ^{13}C NMR (101 MHz, MeOD) δ 163.23, 160.78, 140.15, 133.42, 128.89 (d, *J* = 8.2 Hz), 115.17, 114.42 (d, *J* = 21.7 Hz), 76.63.

1*H*-Pyrazole-3-carboxylic acid (**23**)
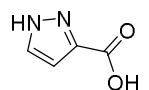
Synthesis based on published procedure.[45] 3-Methyl-1*H*-pyrazole **22** (750 mg, 9.13 mmol) was dissolved in water, KMnO$_4$ (3.18 g, 20.1 mmol) was added and the mixture was refluxed overnight. The reaction was cooled to room temperature, solids were filtered off, and the solvent was removed under reduced pressure. The resulting white powder was used directly without further purification.

Ethyl 1*H*-pyrazole-3-carboxylate (**24**)
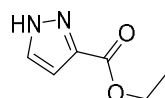
Crude **23** was dissolved in anhydrous ethanol with a catalytic amount of concentrated H$_2$SO$_4$ and refluxed overnight. The reaction was allowed to cool to RT, the solvent was partially removed and the residue was neutralized using sat. aq. NaHCO$_3$. The aqueous phase was extracted with ethyl acetate (3x), and the combined organic layers were dried (MgSO$_4$),

filtered and concentrated under reduced pressure, yielding the title compound as a white power (607 mg, 4.33 mmol, 47% over 2 steps). ^{1}H NMR (400 MHz, MeOD) δ 7.73 (d, J = 2.3 Hz, 1H), 6.85 (d, J = 2.3 Hz, 1H), 4.39 (q, J = 7.1 Hz, 2H), 1.40 (t, J = 7.1 Hz, 3H). ^{13}C NMR (101 MHz, MeOD) δ 108.84, 105.15, 61.94, 14.58.

Bis(4-fluorophenyl)(1H-pyrazol-3-yl)methanol (25)
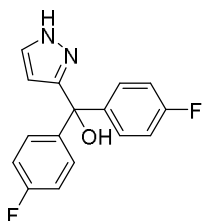
24 (103 mg, 0.735 mmol) was dissolved in THF and cooled to 0 °C. Under vigorous stirring, a 2 M solution of 4-fluorophenylmagnesium bromide in Et$_2$O (1.29 mL, 2.57 mmol) was added dropwise. The reaction mixture was allowed to warm to RT and stirred overnight. The reaction was quenched with sat. aq. NH$_4$Cl. The aqueous phase was extracted with DCM (3x). Combined organic layers were dried (MgSO$_4$), filtered and concentrated under reduced pressure. The residue was purified using silica flash chromatography (30% to 60% EtOAc in pentane) yielding the title compound as yellowish solid (175 mg, 0.611 mmol, 83%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.08 (s, 1H), 7.24 – 7.10 (m, 5H), 7.03 – 6.85 (m, 4H), 5.83 (d, J = 2.2 Hz, 1H). ^{13}C NMR (101 MHz, MeOH) δ 162.18 (d, J = 246.8 Hz), 141.90 (d, J = 3.1 Hz), 131.77, 129.21 (d, J = 8.1 Hz), 114.90 (d, J = 21.4 Hz), 110.10, 105.43, 77.61.

Ethyl 1-trityl-1H-imidazole-4-carboxylate (27)
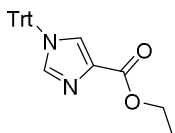
Ethyl 1H-imidazole-4-carboxylate **26** (100 mg, 0.714 mmol) was dissolved in DCM at 0 °C, after which trityl chloride (199 mg, 0.713 mmol) and triethylamine (0.117 mL, 0.856 mmol) were added. The reaction mixture was allowed to warm up to RT overnight, after which it was quenched with water. The organic phase was separated, dried (MgSO$_4$), filtered and concentrated under reduced pressure, giving the title compound as a white powder (265 mg, 0.693 mmol, 97%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.59 (d, J = 1.4 Hz, 1H), 7.45 (d, J = 1.4 Hz, 1H), 7.40 – 7.32 (m, 9H), 7.16 – 7.06 (m, 6H), 4.35 (q, J = 7.1 Hz, 2H), 1.37 (t, J = 7.1 Hz, 3H).

Bis(4-fluorophenyl)(1-trityl-1H-imidazol-4-yl)methanol (28)
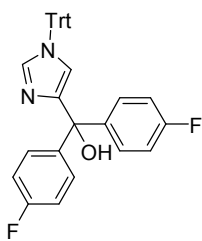
27 (265 mg, 0.693 mmol) was dissolved in THF and cooled to 0 °C. Under vigorous stirring, a 2 M solution of 4-fluorophenylmagnesium bromide in Et$_2$O (1.38 mL, 2.75 mmol) was added dropwise. The reaction mixture was allowed to warm up to RT and further stirred overnight. The reaction was quenched with sat. aq. NH$_4$Cl. The aqueous phase was extracted 3x with ethyl acetate. The combined organic layers were dried (MgSO$_4$), filtered and concentrated under reduced pressure. The residue was purified over silica column (50 % EtOAc in pentane) in order to obtain the product as a yellowish powder (315 mg, 0.596 mmol, 86%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.85 (s, 1H), 7.47 – 7.27 (m, 10H), 7.23 – 6.83 (m, 13H), 6.20 (s, 1H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 162.18 (d, J = 246.9 Hz), 144.16, 140.79, 140.65, 129.62, 129.10 (d, J = 8.1 Hz), 128.92, 128.68, 121.17, 115.33, 115.28 – 114.61 (m), 76.24.

Bis(4-fluorophenyl)(1H-imidazol-4-yl)methanol (29)
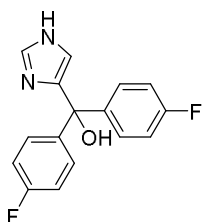
28 (275 mg, 0.520 mmol) was dissolved in 50% TFA/DCM with a few mL of water and stirred overnight. Solvents were removed under reduced pressure and the crude product was dissolved in diethyl ether and extracted with a 1 M aq. HCl solution. The aqueous phase was made basic (pH > 12) with NaOH and extracted with ethyl acetate (3x). Combination, drying (MgSO$_4$), filtered and concentration of the organic phases afforded crude product that was of sufficient purity to use in subsequent reactions as judged by LC/MS (83 mg, 0.29 mmol, 56%).

4-([1,1'-Biphenyl]-4-yl)-1H-1,2,3-triazole (30)
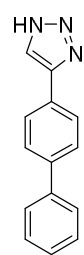
A mixture of formaldehyde (12.5 mL, 168 mmol), acetic acid (1.44 ml, 25.2 mmol) and 1,4-dioxane (125 mL) was stirred for 15 min. Sodium azide (1.64 g, 25.2 mmol) was added, followed by 4-ethynyl-1,1'-biphenyl (3.00 g, 16.8 mmol). After 10 min, sodium ascorbate (0.667 g, 3.37 mmol), and $CuSO_4$·$5H_2O$ (0.210 g, 0.842 mmol) in 1 mL of water were added. The resulting mixture was stirred for 18 h at RT. It was diluted with H_2O (60 mL) and extracted with chloroform (3x 30 mL). The combined organic layers were dried (MgSO4), filtered and concentrated. The residue was suspended in 2 M NaOH (6 mL) and stirred for 20 h at RT. The reaction was acidified with 4 M HCl (aq.) and the white precipitate was filtered off, yielding the desired product as a white solid (2.31 g, 10.4 mmol, 62%). ^1H NMR (400 MHz, DMSO) δ 15.18 (s, 1H), 8.40 (s, 1H), 7.99 – 7.93 (m, 2H), 7.81 – 7.75 (m, 2H), 7.75 – 7.69 (m, 2H), 7.53 – 7.45 (m, 2H), 7.42 – 7.34 (m, 1H). ^{13}C NMR (101 MHz, DMSO) δ 156.90, 145.31, 139.56, 128.99, 127.60, 127.17, 126.56, 126.09.

(R,S)-2-Benzylpiperidine (31)
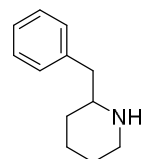
2-Benzylpyridine (5.0 mL, 31 mmol) was dissolved in ethanol (100 mL) and concentrated aqueous HCl (10 mL) was added. Then PtO_2 (112 mg, 0.49 mmol) was added and the mixture was shaken under a hydrogen atmosphere of 2 bar at RT. After overnight shaking, solids were filtered off over celite. The solvent was removed under reduced pressure and the residue was purified using flash column chromatography (10% methanol in DCM) to yield the title compound (4.2 g, 20 mmol, 64%). ^1H NMR (300 MHz, CDCl3) δ 8.41 (s, 1H), 7.42 – 7.13 (m, 5H), 3.58 – 3.38 (m, 2H), 3.25 – 3.02 (m, 1H), 3.02 – 2.73 (m, 2H), 2.14 – 1.47 (m, 5H), 1.47 – 1.10 (m, 1H). ^{13}C NMR (75 MHz, CDCl3) δ 136.16, 129.51, 128.76, 127.04, 58.59, 45.11, 40.07, 27.97, 22.61.

General procedure 1: triphosgene coupling of 2-benzylpiperidine and bis(4-fluorophenyl) heterocycle

Triphosgene (0.7 eq.) was dissolved in dry DCM (0.1 M) and to this solution 2-benzylpiperidine (1 eq.) and Na_2CO_3 (1 eq.) were added at 0 °C. The mixture was stirred for 1 hour, warming to RT. The mixture was then filtered and concentrated in vacuo. The residue was taken up in dry THF, followed by addition of a bis(4-fluoro-phenyl)heterocycle (1 eq.), DMAP (0.1 eq.) and DIPEA (1.1 eq.). The reaction mixture was refluxed to completion. The reaction was quenched with saturated NH_4Cl (aq.), after which the water layer was extracted three times with EtOAc. The combined organic layers were washed with brine, dried with MgSO4, filtered and concentrated under reduced pressure.

(R)-(2-Benzylpiperidin-1-yl)(4-(bis(4-fluorophenyl)(hydroxy)methyl-2H-1,2,3-triazol-2-yl)methanone (1) and (R)-(2-benzylpiperidin-1-yl)(4-(bis(4-fluorophenyl)(hydroxy)methyl)-1H-1,2,3-triazol-1-yl)methanone (2)
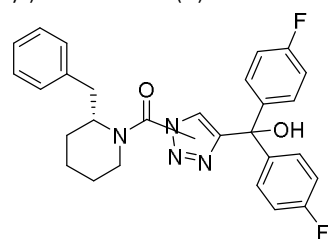
Synthesized according to General Procedure 1 from **18** (123 mg, 0.428 mmol). The N1-isomer was isolated as first eluting isomer (1.63 mg, 3.3 μmol, 1.2%). ^1H NMR (400 MHz, CDCl3) δ 7.41 – 6.57 (m, 14H), 4.76 (s, 1H), 4.40 – 4.18 (m, 1H), 3.56 (d, J = 66.9 Hz, 1H), 3.41 – 2.81 (m, 2H), 2.65 (s, 1H), 2.08 – 1.39 (m, 6H). HRMS: Calculated for $[C_{28}H_{26}F_2N_4O_2+H]^+$ = 488.2079, found = 488.2090. N2-isomer (2.14 mg, 4.4 μmol, 1.5%): ^1H NMR (400 MHz, CDCl3) δ 7.48 (s, 1H), 7.40 – 6.70 (m, 13H), 5.04 – 3.66 (m, 2H), 3.47 (s, 1H), 3.26 (t, J = 13.3 Hz, 1H), 3.16 – 2.81 (m, 2H), 1.94 – 1.43 (m, 6H). HRMS: Calculated for $[C_{28}H_{26}F_2N_4O_2+H]^+$ = 488.2079, found = 488.2091.

(R)-(2-Benzylpiperidin-1-yl)(3-(bis(4-fluorophenyl)(hydroxy)methyl)-1H-1,2,4-triazol-1-yl)methanone (3)
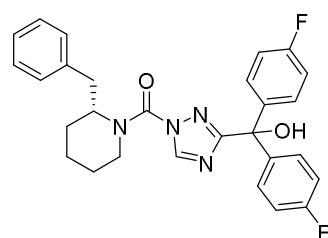
Synthesized according to General Procedure 1 from **21** (65 mg, 0.23 mmol). The title compound was obtained as a white solid (34.7 mg, 0.071 mmol, 35%). ^1H NMR (400 MHz, CD3CN) δ 8.04 (bs, 1H), 7.35 (d, J = 6.6 Hz, 4H), 7.12 (s, 3H), 7.07 – 6.70 (m, 7H), 4.88 (s, 1H), 4.72 (s, 1H), 4.03 (d, J = 15.5 Hz, 1H), 3.23 (t, J = 13.1 Hz, 1H), 2.70 (s, 1H), 1.83 – 1.32 (m, 6H). HRMS: Calculated for $[C_{28}H_{26}F_2N_4O_2+H]^+$ = 488.2097, found = 488.2096.

(*R*)-(2-Benzylpiperidin-1-yl)(3-(bis(4-fluorophenyl)(hydroxy)methyl)-1*H*-pyrazol-1-yl)methanone (**4**)
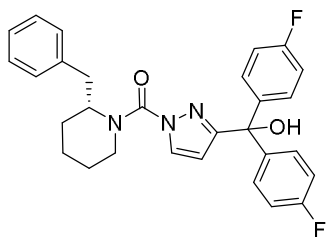
Synthesized according to General Procedure 1 from **25** (65 mg, 0.23 mmol). The title compound was obtained as a white solid (21 mg, 0.043 mmol, 21%). ^{1}H NMR (400 MHz, CD$_3$CN) δ 7.66 – 7.49 (m, 1H), 7.42 – 7.29 (m, 4H), 7.17 (dd, *J* = 5.1, 1.9 Hz, 3H), 7.13 – 6.92 (m, 6H), 6.22 (d, *J* = 2.6 Hz, 1H), 4.77 (s, 1H), 4.59 (s, 1H), 4.17 – 3.99 (m, 1H), 3.25 (td, *J* = 13.4, 3.0 Hz, 1H), 3.10 (dd, *J* = 13.5, 9.0 Hz, 1H), 2.83 – 2.68 (m, 1H), 1.88 – 1.21 (m, 6H). HRMS: Calculated for [C$_{29}$H$_{27}$F$_2$N$_3$O$_2$+H]$^+$ = 488.2144, found = 488.2140.

(*R*)-(2-Benzylpiperidin-1-yl)(4-(bis(4-fluorophenyl)(hydroxy)methyl)-1*H*-imidazol-1-yl)methanone (**5**)
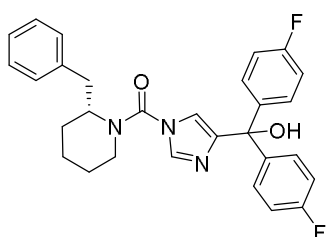
Synthesized according to General Procedure 1 from **29** (83 mg, 0.29 mmol). The title compound was obtained as a white solid (38 mg, 0.078 mmol, 27%). ^{1}H NMR (400 MHz, CD$_3$CN) δ 7.42 (s, 1H), 7.38 – 6.99 (m, 13H), 6.27 (s, 1H), 4.47 (s, 1H), 4.30 (qt, *J* = 7.1, 3.8 Hz, 1H), 3.81 (d, *J* = 13.7 Hz, 1H), 3.32 (td, *J* = 13.3, 2.9 Hz, 1H), 3.19 (dd, *J* = 13.7, 10.0 Hz, 1H), 2.78 (dd, *J* = 13.7, 5.5 Hz, 1H), 1.92 – 1.63 (m, 6H). HRMS: Calculated for [C$_{29}$H$_{27}$F$_2$N$_3$O$_2$+H]$^+$ = 488.2144, found = 488.2142.

(4-([1,1'-Biphenyl]-4-yl)-1*H*-1,2,3-triazol-1-yl)(2-benzylpiperidin-1-yl)methanone (KT109)
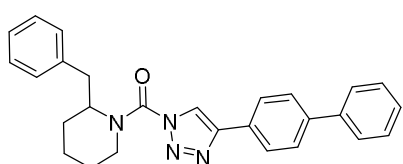
Synthesized according to General Procedure 1 from **30** (2.31 g, 10.4 mmol). The N1-isomer was isolated as first eluting isomer (621 mg, 1.47 mmol, 14%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.87 (s, 2H), 7.75 – 7.60 (m, 4H), 7.53 – 7.35 (m, 4H), 7.24 (d, *J* = 15.7 Hz, 6H), 4.86 (s, 1H), 4.37 (d, *J* = 13.5 Hz, 1H), 3.31 (d, *J* = 52.1 Hz, 2H), 2.70 (s, 1H), 2.11 – 1.62 (m, 6H).

HRMS: Calculated for [C$_{27}$H$_{26}$N$_4$O + H]$^+$ = 423.2179, found = 423.2183.

Supplementary Figures

```
[task]
data = progress | task = fit

[mechanism]
E + S ---> E + P : kprod
                          E + I <==> E.I : kon* koff
E.I ---> E-I : kinact
E ---> E* : kdeg

[constants] | kprod = 4.7 ?, kon* = 100000, koff = 400 ?, kinact = 1 ?, kdeg = 0.027 ?

[concentrations] | E = 0.001 ? (0.0005 .. 0.005), S = 600

[responses] | P = 0.001735

[data]
directory ../Input
sheet 2018-07-10_filename_001.txt
column 2 | offset = auto ? (-0.2 .. +0.2) | conc I = 0
column 7 | offset = auto ? (-0.2 .. +0.2) | conc I = 0.008
column 8 | offset = auto ? (-0.2 .. +0.2) | conc I = 0.004
column 9 | offset = auto ? (-0.2 .. +0.2) | conc I = 0.002
column 10 | offset = auto ? (-0.2 .. +0.2) | conc I = 0.001
column 11 | offset = auto ? (-0.2 .. +0.2) | conc I = 0.0005
column 12 | offset = auto ? (-0.2 .. +0.2) | conc I = 0.00025
column 13 | offset = auto ? (-0.2 .. +0.2) | conc I = 0.000125

[output]
directory ../Output/2018-07-10_filename_001/2018-07-10_filename_001_compound/

[settings] | {Constraints} |

[end]
```

Figure S4.1 | Example DynaFit script. All concentrations are in μM units, rate constants are given in min^{-1} (k_{off}, k_{inact} and k_{deg}) or M^{-1}min^{-1} (k_{prod} and k_{on})
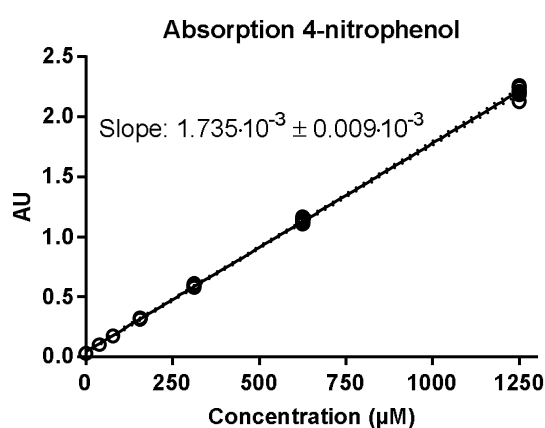
Figure S4.2 | Absorption of 4-nitrophenol. Linear curve fit of absorbance measured for a concentration of 4-nitrophenol in the assay buffer. 6 repetitions are given for each condition. Dashed lines denote 95% confidence interval. Slope value is given ± standard deviation.

References

1. Zhao, Z. & Bourne, P. E. Progress with covalent small-molecule kinase inhibitors. *Drug Discov. Today* **23**, 727–735 (2018).
2. Singh, J., Petter, R. C., Baillie, T. a & Whitty, A. The resurgence of covalent drugs. *Nat. Rev. Drug Discov.* **10**, 307–317 (2011).
3. Keith, J. M. *et al.* Preclinical Characterization of the FAAH Inhibitor JNJ-42165279. *ACS Med. Chem. Lett.* **6**, 1204–1208 (2015).
4. Postnov, A. *et al.* Fatty Acid Amide Hydrolase Inhibition by JNJ-42165279: A Multiple-Ascending Dose and a Positron Emission Tomography Study in Healthy Volunteers. *Clin. Transl. Sci.* **11**, 397–404 (2018).
5. Johnson, D. S. *et al.* Discovery of PF-04457845: A Highly Potent, Orally Bioavailable, and Selective Urea FAAH Inhibitor. *ACS Med. Chem. Lett.* **2**, 91–96 (2011).
6. Huggins, J. P., Smart, T. S., Langman, S., Taylor, L. & Young, T. An efficient randomised, placebo-controlled clinical trial with the irreversible fatty acid amide hydrolase-1 inhibitor PF-04457845, which modulates endocannabinoids but fails to induce effective analgesia in patients with pain due to osteoarthritis of th. *Pain* **153**, 1837–1846 (2012).
7. Cisar, J. S. *et al.* Identification of ABX-1431, a Selective Inhibitor of Monoacylglycerol Lipase and Clinical Candidate for Treatment of Neurological Disorders. *J. Med. Chem.* **61**, 9062–9084 (2018).
8. Fraser, I. *et al.* Preclinical characterization and first-in-human administration of a selective monoacylglycerol lipase inhibitor, ABX-1431. in *Front. Pharmacol. Conference Abstract: EUFEMED 2017* (2017). doi:10.3389/conf.fphar.2017.62.00011
9. Copeland, R. A. *Evaluation of enzyme inhibitors in drug discovery : a guide for medicinal chemists and pharmacologists.* (Wiley, 2013).
10. Schwartz, P. A. *et al.* Covalent EGFR inhibitor analysis reveals importance of reversible interactions to potency and mechanisms of drug resistance. *Proc. Natl. Acad. Sci.* **111**, 173–178 (2014).
11. Aaltonen, N. *et al.* Piperazine and Piperidine Triazole Ureas as Ultrapotent and Highly Selective Inhibitors of Monoacylglycerol Lipase. *Chem. Biol.* **20**, 379–390 (2013).
12. 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).
13. Kerbrat, A. *et al.* Acute Neurologic Disorder from an Inhibitor of Fatty Acid Amide Hydrolase. *N. Engl. J. Med.* **375**, 1717–1725 (2016).
14. Miyahisa, I., Sameshima, T. & Hixon, M. S. Rapid Determination of the Specificity Constant of Irreversible Inhibitors (k inact / K I) by Means of an Endpoint Competition Assay. *Angew. Chemie Int. Ed.* **54**, 14099–14102 (2015).
15. Ahn, K. *et al.* Mechanistic and Pharmacological Characterization of PF-04457845: A Highly Potent and Selective Fatty Acid Amide Hydrolase Inhibitor That Reduces Inflammatory and Noninflammatory Pain. *J. Pharmacol. Exp. Ther.* **338**, 114–124 (2011).
16. Strelow, J. M. A Perspective on the Kinetics of Covalent and Irreversible Inhibition. *SLAS Discov. Adv. Life Sci. R&D* **22**, 3–20 (2017).
17. Bisogno, T. *et al.* Cloning of the first sn1-DAG lipases points to the spatial and temporal regulation of endocannabinoid signaling in the brain. *J. Cell Biol.* **163**, 463–468 (2003).
18. Janssen, F. J. & van der Stelt, M. Inhibitors of diacylglycerol lipases in neurodegenerative and metabolic disorders. *Bioorg. Med. Chem. Lett.* **26**, 3831–3837 (2016).
19. Baggelaar, M. P., Maccarrone, M. & van der Stelt, M. 2-Arachidonoylglycerol: A signaling lipid with manifold actions in the brain. *Prog. Lipid Res.* **71**, 1–17 (2018).
20. 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).
21. Hsu, K.-L. *et al.* Discovery and Optimization of Piperidyl-1,2,3-Triazole Ureas as Potent, Selective, and in Vivo-Active Inhibitors of α/β-Hydrolase Domain Containing 6 (ABHD6). *J. Med. Chem.* **56**, 8270–8279 (2013).
22. Baggelaar, M. P. *et al.* Highly Selective, Reversible Inhibitor Identified by Comparative Chemoproteomics Modulates Diacylglycerol Lipase Activity in Neurons. *J. Am. Chem. Soc.* **137**, 8851–8857 (2015).
23. Adibekian, A. *et al. Optimization and characterization of a triazole urea inhibitor for alpha/beta hydrolase domain-containing protein 11 (ABHD11): anti-probe for LYPLA1/LYPLA2 dual inhibitor ML211. Probe Reports from the NIH Molecular Libraries Program* (National Center for Biotechnology Information (US), 2010).
24. Inloes, J. M. *et al.* The hereditary spastic paraplegia-related enzyme DDHD2 is a principal brain triglyceride lipase. *Proc. Natl. Acad. Sci.* **111**, 14924–14929 (2014).
25. Brindisi, M. *et al.* Development and Pharmacological Characterization of Selective Blockers of 2-Arachidonoyl Glycerol Degradation with Efficacy in Rodent Models of Multiple Sclerosis and Pain. *J. Med. Chem.* **59**, 2612–2632 (2016).
26. Hsu, K.-L. *et al.* Development and Optimization of Piperidyl-1,2,3-Triazole Ureas as Selective Chemical Probes of Endocannabinoid Biosynthesis. *J. Med. Chem.* **56**, 8257–8269 (2013).

27. Deng, H. *et al.* Triazole Ureas Act as Diacylglycerol Lipase Inhibitors and Prevent Fasting-Induced Refeeding. *J. Med. Chem.* **60**, 428–440 (2017).
28. Clayden, J., Greeves, N. & Warren, S. Nucleophilic substitution at the carbonyl group. in *Organic Chemistry* 197–221 (2012).
29. Blais, M.-J., Enea, O. & Berthon, G. Relations structure-réactivité grandeurs thermodynamiques de protonation d'h'et'erocycles saturés et non saturés. *Thermochim. Acta* **20**, 335–345 (1977).
30. Catalan, J. *et al.* Basicity and acidity of azoles: the annelation effect in azoles. *J. Am. Chem. Soc.* **110**, 4105–4111 (1988).
31. Hansen, L. D., West, B. D., Baca, E. J. & Blank, C. L. Thermodynamics of proton ionization from some substituted 1,2,3-triazoles in dilute aqueous solution. *J. Am. Chem. Soc.* **90**, 6588–6592 (1968).
32. 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).
33. Zhao, K.-Y. & Tsou, C.-L. Kinetics of substrate reaction during irreversible modification of enzyme activity where the modifier is not in great excess of the enzyme. *J. Theor. Biol.* **157**, 505–521 (1992).
34. Maurer, T. S. & Fung, H.-L. Comparison of methods for analyzing kinetic data from mechanism-based enzyme inactivation: Application to nitric oxide synthase. *AAPS PharmSci* **2**, 68–77 (2000).
35. Krippendorff, B.-F., Neuhaus, R., Lienau, P., Reichel, A. & Huisinga, W. Mechanism-based inhibition: deriving K(I) and k(inact) directly from time-dependent IC(50) values. *J. Biomol. Screen.* **14**, 913–23 (2009).
36. Kuzmič, P. Program DYNAFIT for the Analysis of Enzyme Kinetic Data: Application to HIV Proteinase. *Anal. Biochem.* **237**, 260–273 (1996).
37. Kuzmič, P. DynaFit--a software package for enzymology. *Methods Enzymol.* **467**, 247–80 (2009).
38. Hsu, K. *et al.* DAGLβ inhibition perturbs a lipid network involved in macrophage inflammatory responses. *Nat. Chem. Biol.* **8**, 999–1007 (2012).
39. Mileni, M. *et al.* Structure-guided inhibitor design for human FAAH by interspecies active site conversion. *Proc. Natl. Acad. Sci.* **105**, 12820–12824 (2008).
40. Edwards, P. D. *et al.* Design, synthesis, and kinetic evaluation of a unique class of elastase inhibitors, the peptidyl α-ketobenzoxazoles, and the x-ray crystal structure of the covalent complex between porcine pancreatic elastase and Ac-Ala-Pro-Val-2-benzoxazole. *J. Am. Chem. Soc.* **114**, 1854–1863 (1992).
41. Janssen, F. J. *et al.* Comprehensive Analysis of Structure–Activity Relationships of α-Ketoheterocycles as sn-1-Diacylglycerol Lipase α Inhibitors. *J. Med. Chem.* **58**, 9742–9753 (2015).
42. Gottlieb, H. E., Kotlyar, V. & Nudelman, A. NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. *J. Org. Chem.* **62**, 7512–7515 (1997).
43. Blacker, A. J. *et al.* Convenient Method for Synthesis of N-Protected α-Amino Epoxides: Key Intermediates for HIV Protease Inhibitors. *Org. Process Res. Dev.* **15**, 331–338 (2011).
44. Taherpour, A. A. & Kheradmand, K. One-pot microwave-assisted solvent free synthesis of simple alkyl 1,2,3-triazole-4-carboxylates by using trimethylsilyl azide. *J. Heterocycl. Chem.* **46**, 131–133 (2009).
45. Fatin-Rouge, N. *et al.* Lanthanide Podates with Programmed Intermolecular Interactions: Luminescence Enhancement through Association with Cyclodextrins and Unusually Large Relaxivity of the Gadolinium Self-Aggregates. *J. Am. Chem. Soc.* **122**, 10810–10820 (2000).

BIA 10-2474 is a non-selective FAAH inhibitor that disrupts lipid metabolism

Introduction

In January 2016, a first-in-human study of the fatty acid amide hydrolase (FAAH) inhibitor BIA 10-2474 led to the death of one volunteer and the hospitalization of four others.[1–4] All patients manifested mild-to-severe neurological symptoms.[3] FAAH is a membrane-bound serine hydrolase that degrades the endocannabinoid anandamide and related amidated lipids.[5–8] Three explanations for the clinical neurotoxicity of BIA 10-2474 have been proposed: (i) errors may have occurred in the clinical trial itself, either in the manufacturing or handling of the compound or in the conduct of the trial; (ii) through its inhibitory effects on FAAH, BIA 10-2474 may have produced high levels of long-chain fatty acid amides (e.g.,

anandamide) and their oxygenated metabolites, which could potentially overstimulate cannabinoid CB_1[8], TRPV1[9], and/or NMDA receptors[10]; or (iii) BIA 10-2474 and/or its metabolites might have off-target activities. The first hypothesis was dismissed by the French authorities.[4] The second hypothesis is considered unlikely because other FAAH inhibitors, such as PF-04457845, have exhibited favourable safety profiles in phase 1 and 2 clinical trials.[11,12] No information has previously been made available regarding the protein interaction profile of BIA 10-2474 that could help in defining the cause of its clinical toxicity, and in particular to directly evaluate the third hypothesis – the possibility that the observed clinical neurotoxicity might have resulted from off-target activity.[1] Therefore, aided by chemical proteomic methods for mapping the interaction landscapes of (ir)reversible serine hydrolase inhibitors[13–17], the serine hydrolase target selectivity of BIA 10-2474 versus other widely used FAAH inhibitors was desired. To this end a comparative study was set up to determine and compare the interaction profiles of BIA 10-2474, its major metabolite BIA 10-2639, the clinically safe PF-04457845 and URB597.

Results

BIA 10-2474 (Figure 5.1A) contains an electrophilic imidazole urea that may react with the nucleophilic serine of FAAH and other serine hydrolases to form covalent and irreversible adducts. On the basis of previously developed chemical proteomic methods to map the interaction landscapes of (ir)reversible serine hydrolase inhibitors,[13–17] it was anticipated that the serine hydrolase targets of BIA 10-2474 could be identified and compared to the selectivity profiles of other widely used FAAH inhibitors, such as URB597[6] and the clinical candidate PF-04457845.[18] PF-04457845 progressed to phase 2 trials without serious adverse events.[18] Therefore, BIA 10-2474 along with BIA 10-2639, a confirmed metabolite in which the *N*-oxide is reduced to a pyridine, were synthesized.[4] These compounds were tested for inhibition of FAAH-catalysed conversion of [^{14}C]-anandamide to arachidonic acid and [^{14}C]-ethanolamine.[19] Surprisingly, BIA 10-2474 showed very weak inhibitory activity ($IC_{50} > 1$ μM) against the FAAH species tested (human, mouse) compared to PF-04457845 and URB597, which exhibited IC_{50} values of ~0.01 and 0.13 μM, respectively (Figure 5.1B), that generally matched previously reported results for these inhibitors.[6,20]
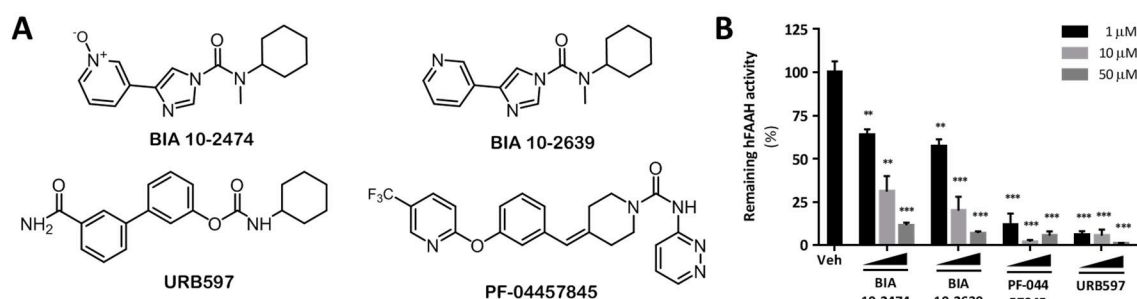
Figure 5.1 | Chemical structure and potency of human FAAH inhibitors. A) Structures of BIA 10-2474, metabolite BIA 10-2639, PF-04457845 and URB597. B) hFAAH activity relative to control after incubation with 1, 10 and 50 μM of indicated inhibitors as measured in a radiometric [^{14}C]-anandamide assay on purified recombinant hFAAH enzyme. Activity is expressed as mean ± SEM (n=3) with *P<0.05, ** P<0.01, *** P<0.001 (two-tailed *t*-test).

Activity-based protein profiling (ABPP) is used for rapid and efficient visualization of endogenous serine hydrolase activities in native biological samples.[13,15] Comparative and competitive gel-based ABPP studies were performed with two different activity-based probes: the broad-spectrum serine hydrolase-directed probe fluorophosphonate-rhodamine (FP-TAMRA)[15] and the tailored probe MB064 that preferentially reacts with endocannabinoid hydrolases diacylglycerol lipase-α (DAGL-α), α/β-hydrolase domain containing protein (ABHD) 6, and ABHD12, along with a handful of other enzymes.[14,17] Together, FP-TAMRA and MB064 provide target engagement assays for FAAH and a broad array of other brain serine hydrolases. As a first screen, a gel-based ABPP assay using mouse brain proteomes was performed. In total, 50 fluorescent bands were identified corresponding to putative serine hydrolases. Representative gel-based ABPP data are shown (Figure 5.2). The clinical trial subjects who developed neurological symptoms were exposed to a concentration of BIA 10-2474 that was 20 to 50 times higher than required for full blockade of FAAH activity.[4] Therefore, inhibitor activities against FAAH and other serine hydrolases were initially evaluated at two high concentrations (10 and 50 μM) in brain soluble and membrane proteomes. All tested compounds inhibited FAAH, with BIA 10-2474 showing the weakest activity, while also exhibiting distinctive off-target activities. In particular, BIA 10-2474 reduced the intensity of an additional fluorescent band in the membrane proteome (red box in panel A), which was identified as ABHD6 based on previous research.[16] PF-04457845 and URB597 reduced the fluorescent labeling of 2 other proteins, but ABHD6 was notably not among them. Moreover, BIA 10-2474 did not prevent labeling of the endocannabinoid hydrolases monoacylglycerol lipase (MAGL) or DAGL-α (Figure 5.2).

The brain target engagement profiles were confirmed and extended by performing ABPP coupled to high-resolution quantitative mass spectrometry (MS). This methodology allows for a more accurate quantification by avoiding band overlap (as observed with the gel-based assay) and enables screening over a broader range of specified serine hydrolases.[13] Mouse brain proteomes treated with inhibitor or vehicle were incubated with the serine hydrolase-directed activity-based probes FP-biotin and MB108, a biotinylated version of MB064. Probe-labeled enzymes were enriched by avidin chromatography, digested with trypsin, and the resulting tryptic peptides modified by reductive dimethylation (ReDiMe) of the NH_2-groups of N-termini and lysine residues using isotopically heavy and light formaldehyde. In these experiments, inhibited serine hydrolases were identified as enzymes exhibiting low heavy/light ratios. Quantitative MS confirmed complete inhibition of FAAH and validated ABHD6 as a major off-target of BIA 10-2474 and its metabolite, but not PF-04457845 (Figure 5.3). In addition to ABHD6, CES1c and ABHD11 were also identified as partial BIA 10-2474 off-targets in murine brain. PPME1 was identified as a potential, partial off-target for PF-04457845 with a heavy/light ratio < 0.5 (Figure 5.3).

Figure 5.2| Identification of serine hydrolase targets of BIA 10-2474, PF-04457845, URB597 and BIA 10-2639 by competitive ABPP of mouse brain proteome. Mouse brain membrane (A, B) or cytosol (C, D) proteome was incubated with inhibitors (10 and 50 μM, 30 min, 37 ˚C) or DMSO as vehicle. Samples were incubated with MB064 (A,C) or FP-TAMRA (B,D). Red boxes highlight ABHD6 (A) and FAAH (B).
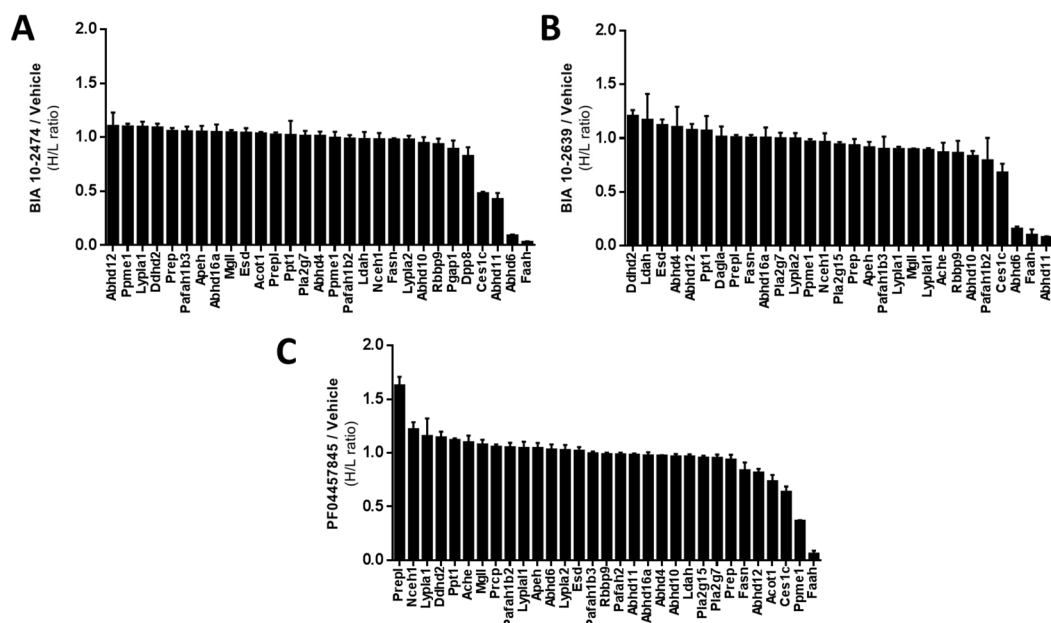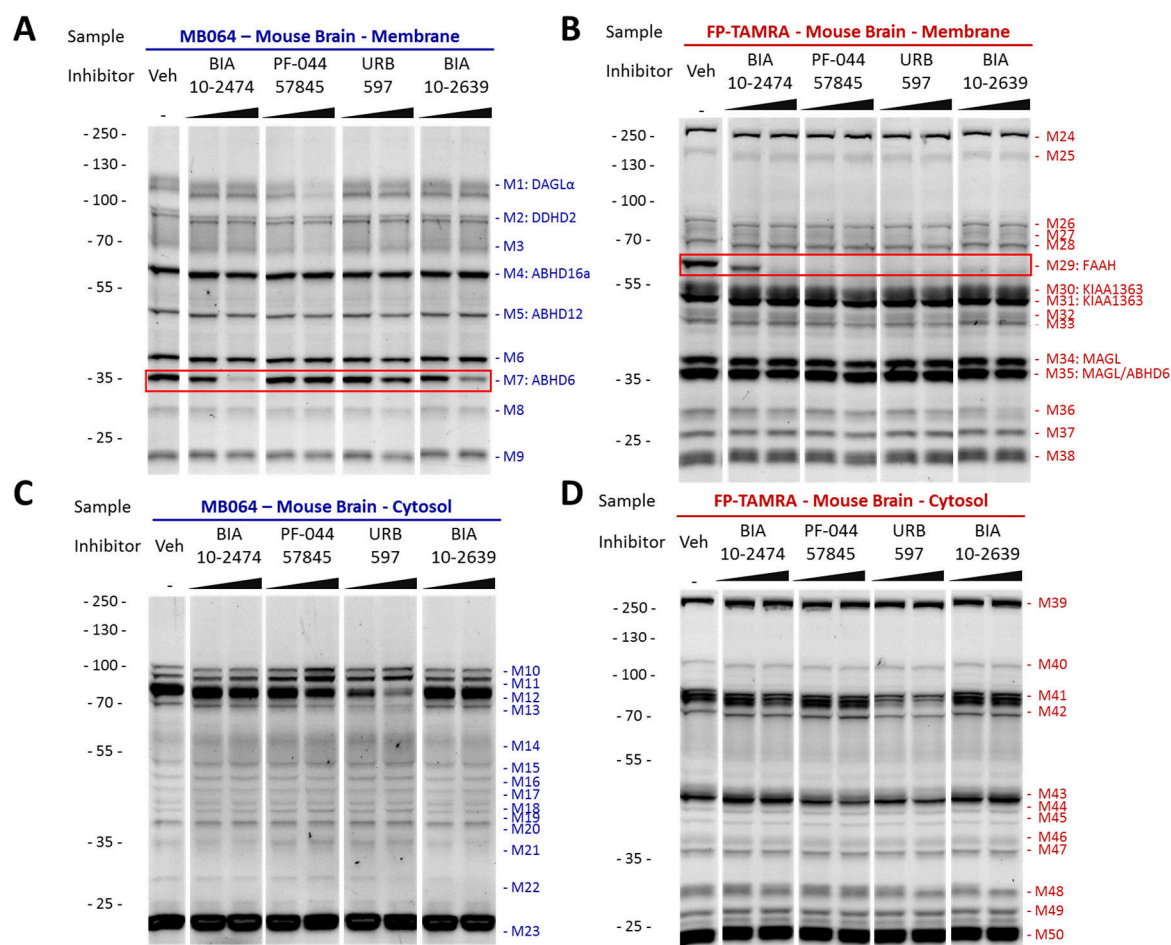
Figure 5.3 | MB108 and FP-biotin based chemoproteomic analysis of serine hydrolase activities in the mouse brain proteome treated with inhibitors BIA 10-2474 (A), BIA 10-2639 (B) and PF-04457845 (C) (50 μM, 30 min, 37 °C). Data is expressed as mean ± SEM (n=3).

Using dedicated activity and binding assays it was determined that BIA 10-2474 and PF-04457845 did not cross-react with other proteins of the endocannabinoid system, including cannabinoid CB$_1$ and CB$_2$ receptors, DAGL-α, DAGL-β, MAGL and *N*-acyl-phosphatidylethanolamine phospholipase D (NAPE-PLD), nor with the endocannabinoid-binding TRP ion channels (TRPV1-4, TRPM8 and TRPA1) (Supplementary Tables 1 and 2).

Finally, activity-based probes based on the structure of BIA 10-2474 were designed to identify potential non-serine hydrolase proteins that are directly modified by the inhibitor. To this end, compounds AJ167, AJ179 and AJ198 were synthesized (Figure 5.4), in which an alkyne functionality was introduced at different positions in BIA 10-2474. In all three probes, the alkyne group serves as a ligation handle to introduce fluorescent reporter groups via copper(I)-catalysed azide-alkyne cycloaddition ("click") chemistry.[21] Competitive ABPP revealed that compounds AJ179 and AJ198 are effective FAAH inhibitors and that ABHD6 can be partially inhibited by compounds AJ167 and AJ198 (Figure 5.5). Reaction of mouse brain proteomes treated with compound AJ167, AJ179, or AJ198 to a fluorophore using click chemistry revealed labeling of a band at the molecular weight of FAAH for probe AJ179 and AJ198. Additionally, compounds AJ179 and AJ198 also labeled a protein at the molecular weight of ABHD6. All FAAH and ABHD6 bands could be competed with BIA 10-2474. Importantly, these experiments definitively prove covalent binding of AJ179 and AJ198 to FAAH, which was maintained even under SDS-PAGE denaturing conditions. This observation, together with ~3-fold increased potency towards rat brain FAAH following a 20 min pre-incubation[22], supports an irreversible inhibition mechanism for BIA 10-2474. Few additional labeled proteins were detected, indicating that BIA 10-2474 has limited cross-reactivity with non-serine hydrolases in the brain proteome.
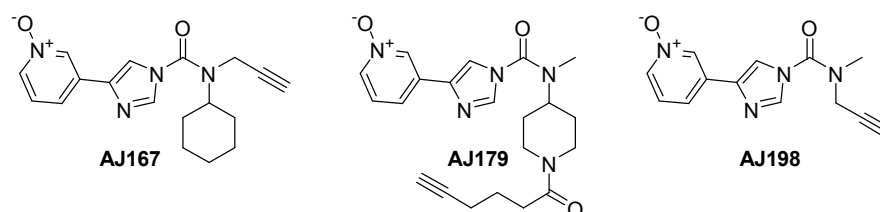
Figure 5.4 | Chemical structures of alkynylated BIA 10-2474-based probes AJ167, AJ179, and AJ198.
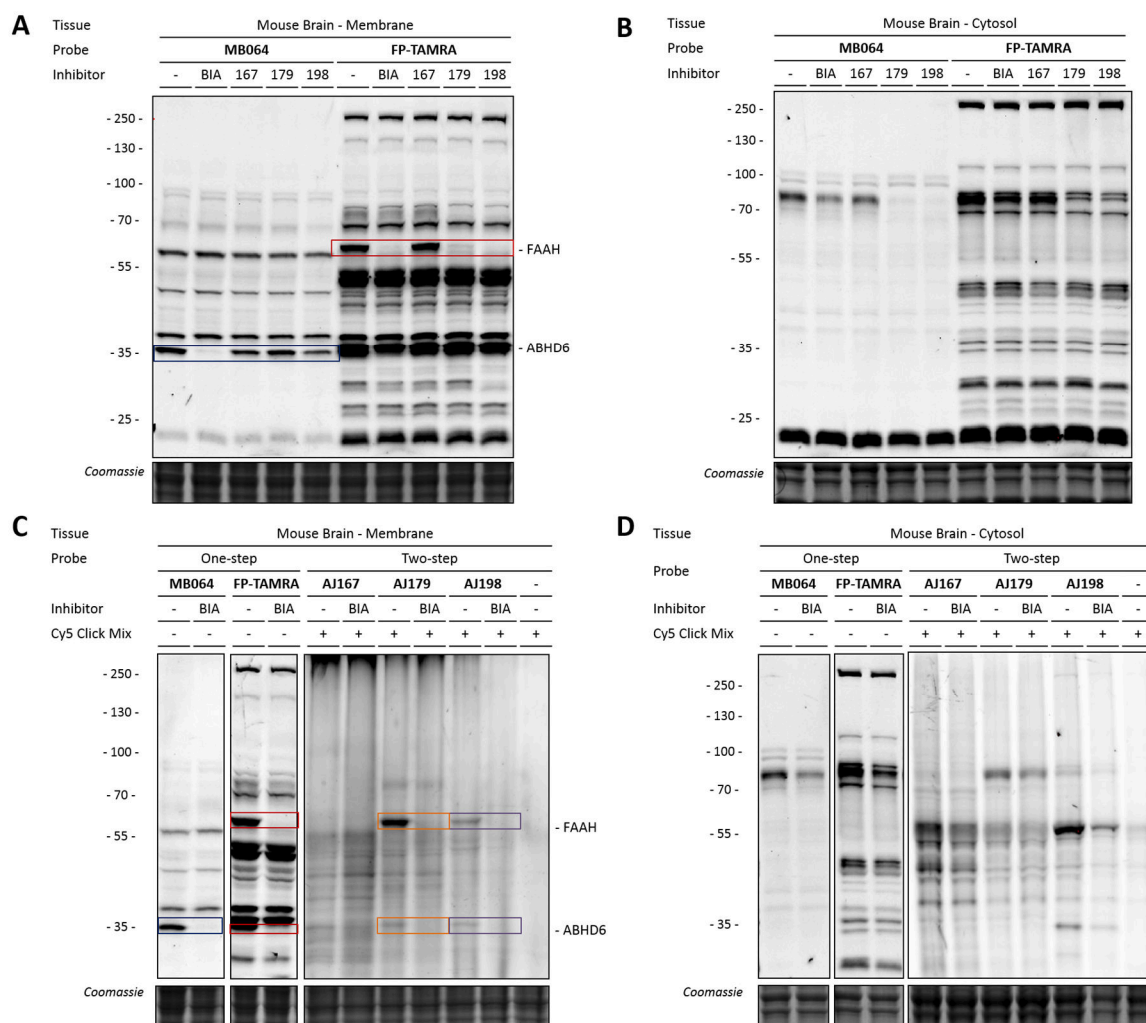
Figure 5.5 | Mouse brain membrane (A) or cytosol (B) proteomes were incubated with BIA 10-2474, the alkyne probes (50 µM, 30 min, 37 °C) or vehicle (DMSO). Residual activity was labeled with MB064 or FP-TAMRA. C-D) Mouse brain membrane (C) or cytosol (D) proteomes were incubated with BIA 10-2474 (50 µM, 30 min, 37 °C) or vehicle (DMSO). Samples were then incubated with alkyne probes (AJ167, AJ179 or AJ198; 50 µM, 30 min, 37 °C) or vehicle (DMSO), followed by a copper-catalyzed azide-alkyne cycloaddition to Cy5-azide. Competitive ABPP of BIA 10-2474 with MB064 or FP-TAMRA (as described for A, B) is shown as a reference.

As noted above, BIA 10-2474 exhibited weaker *in vitro* potency for FAAH compared to PF-04457845 and other advanced FAAH inhibitors. Remarkably, however, this difference was attenuated *in situ*, as gel-based ABPP experiments revealed that BIA 10-2474 inhibited FAAH, as well as other serine hydrolases (e.g., FAAH2, ABHD6) with substantially increased potency in human cells (Figure 5.6). The reason for the increased cellular activity of BIA 10-2474 is at present unclear, but it is not specific to one protein, which may suggest that cellular accumulation of the compound provides sufficiently high intracellular concentrations to inhibit FAAH and other serine hydrolases (*vide infra*).

In light of these data, the structural analogues AJ167, AJ176 and AJ198 were also expected to label more efficiently *in situ*. To test this hypothesis murine neuroblastoma (Neuro-2a) cells were treated with medium containing either BIA 10-2474, or either one of the three alkynated derivatives (Figure 5.7). From Figure 5.7A, it is clear that the two-step probes are much more efficient *in situ* (IS) than *in vitro* (IV). Especially AJ198 showed a

remarkable increase in labeling. Figure 5.7B shows a competition experiment between BIA 10-2474 and all three two-step probes. This demonstrated that the two bands shared by all probes were in fact FAAH and ABHD6 (red boxes) and these could be outcompeted by BIA 10-2474. All other clearly visible bands failed to be outcompeted. Therefore, it was decided not to pursue MS-based ABPP with these probes, because it is unlikely that this would yield additional off-targets.
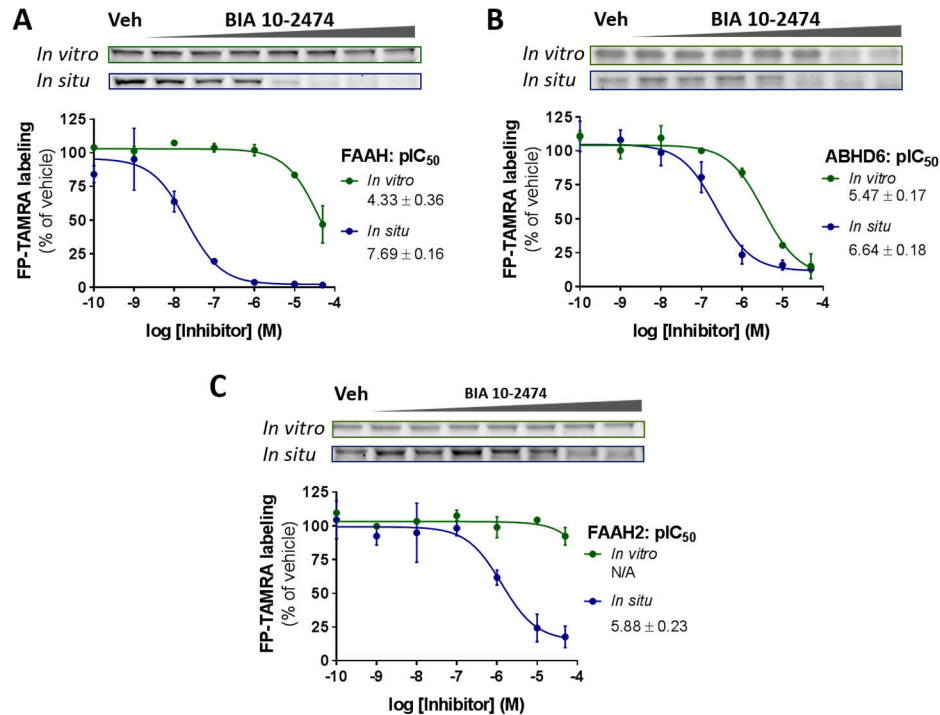
Figure 5.6 | HEK293T cells transiently overexpressing FAAH (A), endogenously expressing ABHD6 (B) and overexpressing FAAH2 (C) were treated with BIA 10-2474 (2 h, 37 °C) or DMSO as vehicle (n=3). Membrane fractions of untreated overexpressing cells were incubated with BIA 10-2474 (30 min, 37 °C) or DMSO as vehicle (*in vitro*). All samples were labeled with FP-TAMRA. Coomassie staining was used as loading control.
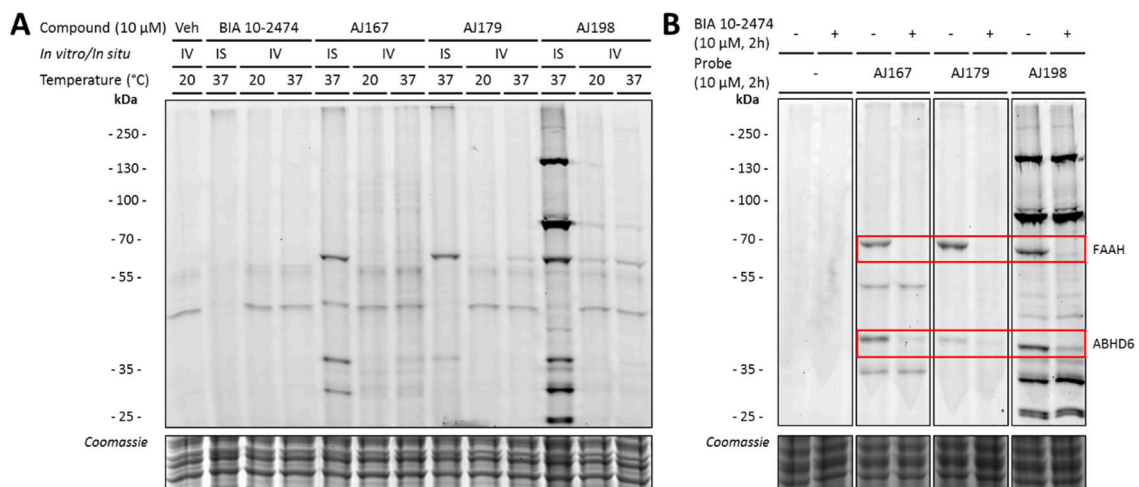
Figure 5.7 | *In situ* and *in vitro* activity profile of BIA 10-2474, AJ167, AJ179 and AJ198 by two-step competitive ABPP. Neuro-2a cells were treated *in situ* (IS) with BIA 10-2474, AJ167, AJ179 or AJ198 (10 µM, 2 h, 37 °C) or *in vitro* (IV) with BIA 10-2474, AJ167, AJ179, AJ198 (10 µM, 30 min, 20/37 °C) or vehicle (DMSO). Probe labeled targets were ligated to Cy5. Coomassie staining was used as a protein loading control.

The adverse effects observed during the clinical trial of BIA 10-2474 were not present in the preclinical toxicity profiling.[3] It was thus hypothesized that the off-target profile of BIA 10-2474 might also differ between species. Taking into account the potency difference between *in situ* versus *in vitro*, BIA 10-2474-treated cultured human neurons were profiled in ABPP studies. Several additional off-targets were identified: carboxyl esterase 2 (CES2), phospholipase 2 group XV (PLA2G15, also known as lysosomal phospholipase A2 (LPLA2)) and patatin-like containing phospholipase domain protein 6 (PNPLA6) (Figure 5.8) by means of chemical proteomics.
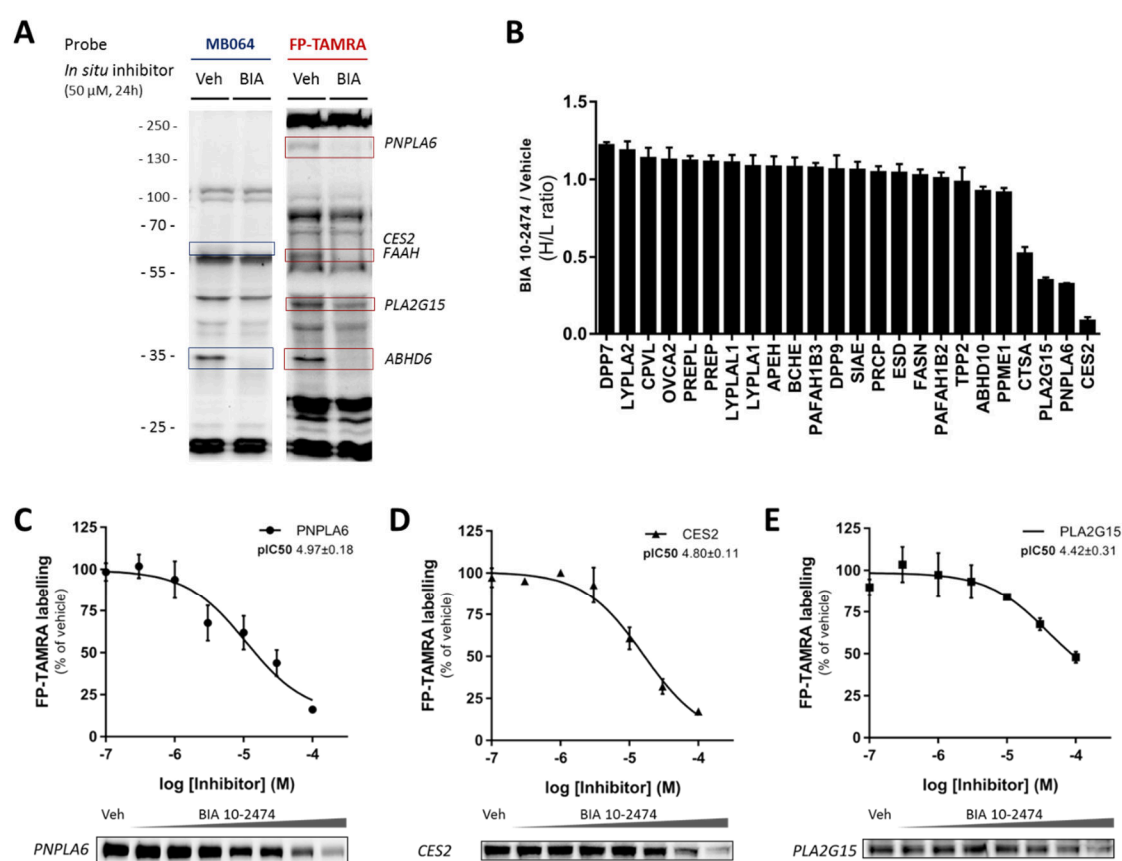
Figure 5.8 | *In situ* treatment of human neuronal cultures reveals additional BIA 10-2474 off-targets. A) Human neuronal cultures were treated *in situ* with BIA 10-2474 (50 μM, 24 h, 37 °C) or DMSO as vehicle (n=3). Lysates were incubated with FP-TAMRA or MB064. B) FP-biotin based chemoproteomic analysis of serine hydrolase activities in human neuronal cultures treated *in situ* with BIA 10-2474 or DMSO as vehicle (50 μM, 24 h, 37 °C). C-E) *In situ* dose response of BIA 10-2474 (0.1 - 100 μM, 24 h, 37 °C) and DMSO as vehicle on cells transiently overexpressing PNPLA6 (C), CES2 (D), and PLA2G15 (E). Data is expressed as mean ± SEM (n=3).

Most of the interaction partners of BIA 10-2474 identified in this study are involved in cellular lipid metabolism.[23–26] To confirm the interaction of BIA 10-2474 with lipolytic serine hydrolase off-targets, these human enzymes were transiently overexpressed in HEK293T cells. Concentration-dependent inhibition by BIA 10-2474 was assessed using gel-based ABPP (Figure 5.8C-E). BIA 10-2474 was found to inhibit PNPLA6, CES2 and PLA2G15 with IC_{50} values of 11, 16 and 38 μM, respectively. It is also noted that human CES2, as well as human ABHD6, were inhibited more potently by BIA 10-2474 and BIA 10-2639 than the mouse orthologues of these enzymes (data not shown). The found off-target profile led to the hypothesis that prolonged exposure to BIA 10-2474 might result in alterations of lipid metabolism in human cells. To test this hypothesis, targeted lipidomics analysis of human neuronal cultures treated with vehicle or BIA 10-2474 was performed. In total, 161 lipid species were quantified from which significant changes in several lipid classes were observed. Levels of *N*-acylethanolamines, triglycerides, monoacylglycerols and (lyso)phosphatidylcholines were increased in BIA 10-2474-treated cells, while those of free fatty acids and plasmalogens were reduced (Figure 5.9A). Notably, these alterations in cellular lipid metabolism are consistent with the inhibition of FAAH, FAAH2, ABHD6, CES2, PLA2G15 and PNPLA6 activities. Neuronal cells treated with PF-04457845 (1 μM) only showed significant increase in the *N*-acylethanolamine levels, consistent with selective FAAH inhibition (Figure 5.9B).
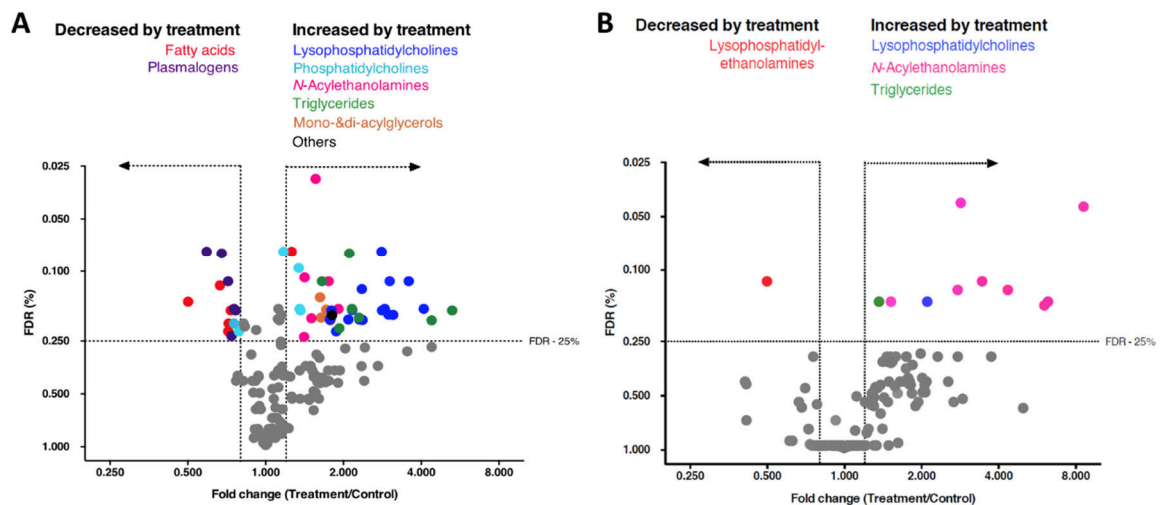
Figure 5.9 | Volcano plot of lipidomic analysis of human neurons treated *in situ* with BIA 10-2474 (50 μM, 48 h, 37 °C) (A) and PF-04457845 (1 μM, 48 h, 37 °C) (B) compared to cells treated with DMSO as vehicle. Lipids with a fold change (FC) threshold of ≥1.20 or ≤0.80 and Benjamini–Hochberg false discovery rate (FDR) ≥ 25% are represented by coloured circles distinguished by lipid class. Based on average lipid levels (n=3)

Discussion

Severe adverse effects of drug candidates are rarely observed in phase 1 clinical trials, due to extensive preclinical toxicological profiling in animals and precautions taken into account in the design of first-in-human studies.[1] No obvious toxicological results in rodents were found that could predict the observed human clinical neurotoxicity. In a study on dogs treated for 13 weeks with BIA 10-2474, a dose-dependent pulmonary toxicity was observed and two dogs from the subgroup receiving the highest dose were sacrificed. An initial toxicology study in primates showed that the highest administered dose led to axonal dystrophy in the spinal bulb. A follow-up primate study led to the death of one animal and the sacrifice of several others for undisclosed ethical reasons. These findings were not, however, considered to be sufficiently concerning to abandon the first-in-human studies due to the large therapeutic window in preclinical studies of BIA 10-2474.[4]

The main remaining hypothesis to explain the human clinical neurotoxicity of BIA 10-2474 as put forward by the French authorities[4] (i.e. off-target activity of BIA 10-2474 and/or its metabolites) was the basis of the here described study. BIA 10-2474 was shown to be an irreversible, potent FAAH and FAAH2 inhibitor that increases cellular levels of long-chain fatty acid ethanolamides. Using ABPP with mouse brain proteomes and human cortical neuron proteomes, BIA 10-2474, but not PF-04457845, was found to inhibit multiple lipases, including ABHD6, CES2, PLA2G15 and PNPLA6. BIA 10-2474 also disrupted neural lipid metabolism as witnessed by increased levels of N-acylethanolamines, monoacylglycerols, triglycerides phosphatidylcholine, (lyso)phosphatidylcholine, and reductions in free fatty acids and plasmalogens. When considering the basis for BIA 10-2474's broader interaction profile with serine hydrolases compared to PF-04457845, it is suspected that the greater intrinsic reactivity of BIA 10-2474 may be a contributing factor, though this warrants further investigation.

FAAH2 is a human-specific orthologue of FAAH, which also degrades long-chain fatty acid amides.[27] However, little is known about the neurobiological function of FAAH2. CES2 is a serine esterase involved in the hydrolysis of ester and amide bonds of xenobiotics and prodrugs, but also of endogenous lipids. It is highly expressed in the liver and in endothelial cells. Hepatic CES2 is required for triglyceride homeostasis by modulating lipolysis, ER stress and lipogenesis.[28] PLA2G15 is a lysosomal phospholipase highly expressed in alveolar macrophages and microglial cells. Inhibition of PLA2G15 leads to accumulation of phosphatidylcholines and phosphatidylethanolamines, and has previously been implicated in drug-induced phospholipidosis.[25]

PNPLA6 (also known as neuropathy target esterase, NTE) has previously been linked to neurodegeneration in humans resulting from exposure to organophosphate-based pesticides.[24,29] PNPLA6 is a (lyso)phospholipase that uses (lyso)phosphatidylcholine as a substrate and is highly expressed in endothelial cells, astrocytes and neurons. Proper neuronal phosphatidylcholine homeostasis mediated by PNPLA6 is required for axonal maintenance.[30] PNPLA6 inhibition leads to neurodegeneration following an age-dependent neuropathy of increasing severity in older age. Notably, PNPLA6-mediated neurotoxicity following acute organophosphate exposure is species dependent, with less pronounced effects in rodents than humans. Mice with a homozygous germline deletion of PNPLA6 have embryonic lethality due to placental and vasculature defects, however, brain-specific

deletion of PNPLA6 results in prominent neuronal pathology.[31] A threshold of >70% PNPLA6 inhibition has been determined as responsible for developing neurodegeneration upon organophosphate exposure in chicken.[24,29] In the light of this data, it is remarkable that many clinical symptoms arising from PNPLA6-mediated neurological sequelae following organophosphate exposure resemble the neuropathology observed in the clinical trial participants who received a cumulative dose of 250-300 mg of BIA 10-2474, including precedent for human neurotoxicity, brain region sensitivity, age dependency, species selectivity, dosing threshold and time course of neuropathology. However, while our data provide information about the selectivity of BIA 10-2474, they do not allow us to conclude that inhibition of one or more of the identified off-target proteins is responsible for the clinical neurotoxicity caused by this drug. Nor can we exclude the possibility that non-covalent interactions of BIA 10-2474 or its metabolites with other proteins might have contributed to the reported clinical effects.[32]

Regardless, our study highlights the general utility of ABPP as a versatile chemical proteomic method to assess on-target engagement and off-target activity of covalent drugs to guide therapeutic development.

Acknowledgements

Methods

Probes and inhibitors

Activity based probe TAMRA-fluorophosphonate (FP-TAMRA) was purchased from Thermo Fisher, MB064 was synthesized in-house as previously described.[14] PF-04457845 was synthesized in house and purchased from Sigma Aldrich. URB597 was purchased from Sigma Aldrich. Inhibitor BIA 10-2474, BIA 10-2639 and activity based probes AJ167, AJ179 and AJ198 were synthesized as described in the synthetic methods *(vide infra)*. All synthesized compounds were at least 95% pure and were analyzed by LC/MS, NMR and HRMS. Other chemicals, reagents, and primers were purchased from Sigma Aldrich unless indicated otherwise.

Cloning

Full-length human cDNA of ABHD6, FAAH, ABHD11, PNPLA6, CES2 (Source Bioscience) and FAAH2 (BioCat) was cloned into mammalian expression vector pcDNA3.1, containing genes for ampicillin and neomycin resistance. The inserts were cloned in frame with a C-terminal FLAG-tag and site-directed mutagenesis was used to remove restriction by silent point mutations. Plasmids were isolated from transformed XL-10 Z-competent cells (Maxi Prep kit: Qiagen) and sequenced at the Leiden Genome Technology Center. Sequences were analyzed and verified (CLC Main Workbench).

Cell culture

General

HEK293T (human embryonic kidney) and Neuro-2a (mouse neuroblastoma) cells were cultured at 37 °C under 7% CO_2 in DMEM containing phenol red, stable glutamine, 10% (v/v) New Born Calf Serum (Thermo Fisher), and penicillin and streptomycin (200 µg/mL each; Duchefa). Medium was refreshed every 2-3 days and cells were passaged twice a week at 80-90% confluence by resuspension in fresh medium.
Cells lines were purchased from ATCC and were regularly tested for mycoplasma contamination. Cultures were discarded after 2-3 months of use.

Human neural cell culture

Human iPSC-derived neural progenitor cells (NPCs) (Axol Biosciences, line ax0015) were plated on sterile coverslips in 12-well plates, coated with poly-L-ornithin/laminin (Sigma-Aldrich), in neural differentiation medium (Neurobasal medium, 1% N2 supplement, 2% B27-RA supplement, 1% MEM-NEAA, 20 ng/mL BDNF (ProSpec Bio)), 20 ng/mL GDNF (ProSpec Bio), 1 µM db-cAMP (Sigma-Aldrich), 200 µM ascorbic acid (Sigma-Aldrich), 2 µg/mL laminin, and 1% P/S), resulting in neural networks composed of neurons and glia. Cells were refreshed with neural differentiation medium 3 times per week. During weeks 1-4, medium was fully refreshed. After 4 weeks of neural differentiation, only half of the volume of medium per well was refreshed.

Transient transfection

One day prior to transfection HEK293T cells were seeded to 15-cm dishes or 12-well plates (~62.500 cells/cm^2). Prior to transfection, culture medium was aspirated and a minimal amount of medium was added. A 3:1 (m/m) mixture of polyethyleneimine (PEI) (60 µg/dish or 1.875 µg/well) and plasmid DNA (20 µg/dish or 0.625 µg/well) was prepared in serum-free culture medium and incubated for 15 min at RT. Transfection was performed by dropwise addition of the PEI/DNA mixture to the cells. Transfection with the empty pcDNA3.1 vector was used to generate control samples. After 24 h, medium was refreshed. Medium was aspirated 48 or 72 h post-transfection and cells were harvested by resuspension in PBS. Cells were pelleted by centrifugation (5 min, 1,000 *g*) and the pellet was washed with PBS. Supernatant was discarded and cell pellets were frozen in liquid nitrogen and stored at -80 °C until sample preparation.

In situ treatment of HEK293T or Neuro-2a cells

In situ treatment was initiated 24 h (for 24 – 48 h treatment) or 48 h (for 2 h treatment) post-transfection in 12-well plates. Medium was aspirated and medium containing inhibitor or DMSO as vehicle was added (0.1-1.0% v/v DMSO, DMSO concentration constant within each experiment). Final concentrations for inhibitors are indicated in the main text and figure legends. After 2, 4, 24, or 48 h exposure to treatment medium, medium was aspirated and cells were harvested and stored as described above until sample preparation (whole cell lysate).

In situ treatment of human neural cell culture

Cells were treated with inhibitor *in situ* 7-8 weeks after plating of NPCs. Culture slides were then collected from 6-well culture plates 24 h or 48 h after initiation of *in situ* treatment with BIA 10-2474 (50 µM, 0.25% DMSO) or vehicle (0.25% DMSO). The *in situ* treated cells were harvested in PBS as described above and stored at -80 °C until further use.

Sample preparation

Whole cell lysate

Cell pellets were thawed on ice, resuspended in cold lysis buffer (20 mM HEPES, pH 7.2, 2 mM DTT, 250 mM sucrose, 1 mM $MgCl_2$, 2.5 U/mL benzonase) and incubated on ice (15-30 min). The cell lysate was used for membrane preparation (below) or diluted to appropriate concentration in cold storage buffer (20 mM Hepes, pH 7.2, 2 mM DTT) for use as whole lysate (HEK293T: 2.0 mg/mL, human cortical neurons: 0.75 mg/mL, human neural cells: 1.5 mg/mL). Protein concentrations were determined by a Quick Start™ Bradford Protein Assay and diluted samples were flash frozen in liquid nitrogen and stored at -80 °C until further use.

Tissue lysate

Mouse brains (C57Bl/6) were isolated according to guidelines approved by the ethical committee of Leiden University (DEC#13191), frozen in liquid nitrogen, and stored at -80 °C until use. Tissues were thawed on ice, dounce homogenized in cold lysis buffer and incubated on ice (15 min), followed by two low-speed spins (3 min, 1,400−2,500 *g*, 4 °C) to remove debris. The supernatant fraction was collected for further use.

Membrane preparation from lysate

The membrane and cytosolic fractions of cell or tissue lysates were separated by ultracentrifugation (93,000 *g*, 45 min, 4°C). The supernatant was collected (cytosolic fraction) and the membrane pellet was resuspended in cold storage buffer by thorough pipetting and passage through an insulin needle. Protein concentrations were determined by a Quick Start™ Bradford Protein Assay and samples were diluted to 2.0 mg/mL with cold storage buffer, flash frozen in liquid nitrogen and stored at -80 °C until further use. FAAH and FAAH2 overexpression membranes were mixed in an appropriate ratio for ABPP (0.1 mg/mL: 1.0 mg/mL).

Activity based protein profiling

Gel based: Direct activity based probes

Gel-based ABPP was performed and analyzed with minor adaptations on previously reported procedures.[14] In brief, for *in vitro* inhibition, or mouse brain proteome or cell lysate (15 µL, 0.75, 1.5 or 2.0 mg/mL, lysate, cytosol or membrane fraction) was pre-incubated with vehicle or inhibitor (0.375 µL 40* inhibitor stock, 30 min, 37 °C) followed by an incubation with the activity-based probe (0.375 µL 40* stock, final concentrations: 250 nM for MB064 or 500 nM for FP-TAMRA, 20 min, RT). Final concentrations for the inhibitors are indicated in the main text and figure legends. For *in situ* inhibition, the *in situ*-treated cells (15 µL whole lysate) were directly incubated with the activity-based probe (0.375 µL 40* stock, final concentrations: 250 nM for MB064 or 500 nM for FP-TAMRA, 20 min, RT). Reactions were quenched with 4* Laemmli buffer (5 µL, 240 mM Tris (pH 6.8), 8% (w/v) SDS, 40% (v/v) glycerol, 5% (v/v) β-mercaptoethanol, 0.04% (v/v) bromophenol blue). 7.5, 15 or 20 µg per reaction was resolved on a 10% acrylamide SDS-PAGE gel (180 V, 75 min). Gels were scanned using Cy3 and Cy5 multichannel settings (605/50 and 695/55, filters respectively) and stained with Coomassie after scanning. Fluorescence was normalized to Coomassie staining and quantified with Image Lab (Bio-Rad). IC_{50} curves were fitted with Graphpad Prism® 7 (Graphpad Software Inc.).

Gel based: Two-step activity based probes

The two-step labeling protocol was adapted from previously developed methods.[33] In brief, human or mouse proteome (40 µL, 2 mg/mL, cytosol or membrane fraction) was pre-incubated with vehicle (DMSO) or inhibitor (50 µM, 30 min, 37 °C) followed by an incubation with the alkyne probe (50 µM, 30 min, 37 °C). Click reagent was freshly prepared by mixing copper sulfate (2 µL/reaction, 100 mM in water), THPTA (0.4 µL/reaction, 100 mM in water), sodium ascorbate (1.2 µL/reaction, 1 M in water), and Cyanine 5-Azide (Cy5−N_3, 1.60 µL/reaction, 1 mM in DMSO). Click reagent (5.0 µL) was added to each proteome, mixed by brief vortexing and incubated (60 min, RT). Proteins were precipitated by adding methanol (50 µL), chloroform (15 µL) and water (15 µL), mixed by brief vortexing and pelleted by centrifugation (10 min, 4000 rpm). The supernatant was removed and the pellet was washed twice with methanol (50 µL). The pellet was then redissolved in 40 µL

of 2* Laemmli buffer (120 mM (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 2.5% (v/v) β-mercaptoethanol, 0.02% (v/v) bromophenol blue) by vortexing. Samples were boiled (95 °C, 5 min) and 20 µg proteome per reaction was resolved on a 10% acrylamide SDS-PAGE gel (180 V, 75 min). Gels were scanned using Cy3 and Cy5 multichannel settings (605/50 and 695/55, filters respectively) and stained with Coomassie after scanning. Coomassie staining was performed as a protein loading control.

Activity-based proteomics

Activity-based proteomics was based on previously described procedures.[17] In summary, mouse whole brain proteome (250 µL cytosolic or membrane fraction at 1.0 mg/mL) was incubated with vehicle (2% DMSO) or inhibitor (BIA 10-2474, BIA 10-2639, or PF-04457845, 50 µM, 30 min, 37 °C). The proteome of *in situ* treated human neural cells (250 µL whole lysate at 0.35 mg/mL, 24 h treatment) was used without additional inhibitor incubation. The proteomes were then labeled with MB108 or FP-Biotin (10 µM, 60 min, RT). The labeling reaction was quenched and excess probe was removed by chloroform methanol precipitation. Precipitated proteome was resuspended in 6 M Urea/25 mM ammonium bicarbonate (250 µL) and incubated (15 min, RT). Subsequently DTT (2.5 µL, 1 mM) was added and the mixture was incubated (15 min, 65 °C). The sample cooled to RT and iodoacetamide (0.5 M, 20 µL) was added to alkylate the sample (30 min, RT, dark). SDS (70 µL, 10% (v/v)) was added and the proteome was heated (5 min, 65 °C). The sample was diluted with PBS (3 mL). 50 µL of a 50% slurry of Avidin−Agarose from egg white (Sigma-Aldrich) was washed with PBS and added to the proteome sample. The beads were incubated with the proteome (3 h, RT, shaking). The beads were isolated by centrifugation (2500 g, 2 min) and washed with SDS in PBS (10 mL, 0.5% (w/v)), followed by 3 washes with PBS. The beads were transferred to low-binding Eppendorf tubes and proteins were digested with sequencing grade trypsin (Promega) (500 ng per sample) in 250 µL buffer (100 mM Tris, 100 mM NaCl, 1 mM $CaCl_2$, 2 % acetonitrile) (37 °C, O/N, vigorous shaking). The pH was adjusted with formic acid to pH 3 and the beads were removed by filtration.

The peptides were isotopically labeled by on stage tip dimethyl labeling according to literature procedures[17], with the following modification. Dimethyl labeling was performed by the subsequent addition of 20, 20, 30, 30 and 40 µL of Light (vehicle) or Medium (inhibitor) reagent to the stage tips. The centrifugation speed during labeling was adjusted to have a flow through time of approximately 5 min (400-1000 g) per labeling step.

Targeted lipidomics

Sample extraction

Lipids were extracted from *in situ* treated human neural cells (48 h, 50 µM BIA 10-2474 or vehicle (0.25% DMSO)). The sample extraction was performed on ice. In brief, cell pellets with 1 million cells were transferred to 1.5 mL Eppendorf tubes, spiked with 10 µL each of deuterated labeled internal standard mix for endocannabinoids (N-arachidonoyl ethanolamine (AEA)-d8, N-arachidonoyldopamine (NADA)-d8, N-docosahexaenoylethanolamide (DHEA)-d4, 2-arachidonoylglycerol (2-AG)-d8, N-stearoylethanolamine (SEA)-d3, N-palmitoylethanolamine(PEA)-d4, N-linoleoylethanolamine (LEA)-d3 and N-oleoylethanolamine (OEA)-d4), positive apolar lipids (lysophosphatidylcholines (LPC)17:0, phosphatidylethanolamines (PE)17:0/17:0, phosphatidylcholines (PC)17:0/17:0, sphingomyelins (SM) d18:1/17:0, triglycerides (TG) 17:0/17:0/17:0, ceramides (Cer) d18:1/17:0) and negative polar lipids (fatty acid (FA)17:0-d33), followed by the addition of ammonium acetate buffer (100 µL, 0.1 M, pH 4). After extraction with methyl *tert*-butyl ether (1 mL), the tubes were thoroughly mixed for 4 min using a bullet blender at medium speed (Next Advance, Inc., Averill park, NY, USA), followed by a centrifugation step (5000 g, 12 min, 4 °C). Then 925 µL of the upper layer methyl *tert*-butyl ether was transferred into clean 1.5 mL Eppendorf tubes. Samples were dried in a speed-vac followed by reconstitution in acetonitrile:water (50 µL, 90:10, v/v). The reconstituted samples were centrifuged (14,000 g, 3 min, 4 °C) before transferring into LC-MS vials. Each sample was injected on three different lipidomics platforms: endocannabinoids (5 µL), positive apolar lipids (2 µL) and for negative polar lipids (8 µL).

LC-MS/MS Analysis for endocannabinoids

A targeted analysis of 21 endocannabinoids and related NAEs (N-acylethanolamines) were measured using an Acquity UPLC I class Binary solvent manager pump (Waters, Milford, USA) in conjugation with AB SCIEX 6500 quadrupole-ion trap (QTRAP) (AB Sciex, Massachusetts, USA). Separation was performed with Acquity HSS T3 column (1.2 x 100 mm, 1.8 µm) maintained at 45 °C. The aqueous mobile phase A consisted of 2 mM ammonium formate and 10 mM formic acid, and the organic mobile phase B was acetonitrile. The flow rate was set to 0.4 mL/min; initial gradient conditions were 55% B held for 2 min and linearly ramped to 100% B over 6 minutes and held for 2 minutes; after 10 s the system returned to initial conditions and held 2 min

before next injection. Electrospray ionization-MS was operated in positive mode for measurement of 21 endocannabinoids and NAEs, and a selective Multiple Reaction Mode (sMRM) was used for quantification.

LC-MS/MS analysis for positive apolar and negative polar lipids

Both lipidomics methods are adapted and modified from previously published work.[34] Briefly, these methods are measured on an Acquity UPLC Binary solvent manager pump (Waters) coupled to an Agilent 6530 electrospray ionization quadrupole time-of-flight (ESI-Q-TOF, Agilent, Jose, CA, USA) high resolution mass spectrometer using reference mass correction. The chromatographic separation was achieved on an Acquity HSS T3 column (1.2 x 100 mm, 1.8 µm) maintained at 40 °C for both methods. The positive polar lipids that include targets from different lipid classes including (lyso)phosphatidylcholines, triglycerides, ceramides. (lyso)phosphatidylethanolamines and sphingomyelins were separated using a flow of 0.4 mL/min over a 16 min gradient. In positive mode, the aqueous mobile phase A consisted of 60:40 (v/v) acetonitrile:H_2O with 10 mM ammonium formate, and the organic mobile phase B consisted of 10:90 (v/v) acetonitrile:isopropanol with 10 mM ammonium formate. The negative apolar lipids that constitute mainly free fatty acids and (lyso)phosphatidylcholines were separated with a flow of 0.4 mL/min over 15 min gradient. In negative mode, the aqueous mobile phase A consisted of 5:95 (v/v) acetonitrile:H_2O with 10 mM ammonium formate, and the organic mobile phase B consisted of 99% (v/v) methanol with 10 mM ammonium formate. The targets in both lipid methods were detected full scan (100-1000 m/z) in their respective ion charge mode.

Activity assays

Radiolabeled natural substrate assay hFAAH and mFAAH

The radiolabeled natural substrate based assay for human and mouse FAAH was performed as reported previously.[35] In brief, chemicals were of the purest analytical grade. Anandamide (AEA) was purchased from Sigma Chemical Co. (St. Louis, MO, USA). URB597 and purified hFAAH were obtained from Cayman Chemical (Ann Arbor, MI, USA). [14C-ethanolamine]-anandamide (60 Ci/mmol) was purchased from ARC (St. Louis, MO). Mouse brain membranes (50 µg per test), prepared as reported[35], or hFAAH (2.5 µg per test) were pre-incubated for 15 min at room temperature with each compound, then they were incubated with 10 µM [14C-ethanolamine]anandamide (15 min, 37 °C, in 500 µL of 50 mM Tris–HCl buffer (pH = 9)). The reaction was stopped by the addition of 0.6 mL of ice-cold methanol/chloroform (2:1, v/v). The mixture was centrifuged (3000 g, 5 min), the upper aqueous layer was put in a vial containing liquid scintillation cocktail (Ultima Gold XR, Perkin Elmer Life Sciences), and radioactivity was quantified in a β-counter.

Natural substrate-based fluorescence assay DAGL-α/β

The natural DAG substrate assay was performed as reported previously.[36] Standard assay conditions: 0.2 U/mL glycerol kinase (GK), glycerol-3-phosphate oxidase (GPO) and horseradish peroxidase (HRP), 0.125 mM ATP, 10 µM Amplifu™Red, 5% DMSO in a total volume of 200 µL. The assay additionally contained 5 µg/mL MAGL-overexpressing membranes, 100 µM SAG and 0.0075% (w/v) Triton X-100, with a final protein concentration of 50 µg/mL. The mDAGL-β assay was performed as the hDAGL-α assay, but assay buffer was supplemented with 5 mM $CaCl_2$ and the SAG concentration was 75 µM.

Natural substrate-based fluorescence assay MAGL

The natural substrate MAGL assay was performed as previously published.[36]

Surrogate substrate assay NAPE-PLD

The surrogate substrate assay was based on a previously reported method.[37] The membrane protein fraction HEK293T cells transiently overexpressing NAPE-PLD was diluted to 0.4 mg/mL in assay buffer (50 mM Tris-HCl (pH 7.5), 0.02% Triton X-100, 150 mM NaCl). The substrate PED6 (Invitrogen) stock (10 mM) was consecutively diluted in DMSO (25x) and in assay buffer (10x). Relevant concentrations of compounds were prepared in DMSO. The assay was performed in a dark Greiner 96 wells plate, end volume 100 µL. The membrane protein lysate (final concentration 0.04 mg/mL) was incubated with inhibitor or vehicle (30 min, 37 °C). A sample without membrane protein lysate was incorporated for background subtraction. Subsequently, substrate was added (PED6, final concentration: 1 µM) and the measurement was started immediately on a TECAN infinite M1000 pro (37 °C, scanning at 2 min intervals for 1 h: excitation 485 nm, emission 535 nm).

Radioligand displacement assays CB_1/ CB_2-receptor

$[^3H]CP55940$ displacement assays to determine the affinity for the cannabinoid CB_1 and CB_2 were performed as previously described[16,38], with the following changes: ligands of interest were incubated (25 °C, 2 h) with membrane aliquots containing 1.5 μg (CHOK1hCB2_bgal) membrane protein in 100 μL assay buffer (50 mM Tris–HCl, 5 mM $MgCl_2$, 0.1 % BSA, pH 7.4) with ~1.5 nM $[^3H]CP55940$ per assay point. Non-specific binding was determined in the presence of 10 μM AM630. Filtration was performed on 96-well GF/C filters, each well presoaked for 30 min with 25 μL 0.25 % PEI, using a 96-wells Filtermate harvester (PerkinElmer). Filter-bound radioactivity was determined by scintillation spectrometry using a Microbeta® 2450 microplate counter (PerkinElmer).

Fluorescent Ca^{2+} assays for TRP ion channels

HEK293 (human embryonic kidney) cells stably over-expressing recombinant human TRPV1 or rat TRPA1, TRPV2, TRPV3, TRPV4, or TRPM8 were grown on 100 mm diameter Petri dishes as mono-layers in minimum essential medium (MEM) supplemented with non-essential amino acids, 10 % fetal bovine serum, 2 mM glutamine, and maintained at 5 % CO_2 at 37 °C. Quantitative real time analysis was carried out to measure TRP gene over-expression in transfected-cells (data not shown). On the day of the experiment, cells were loaded with the methyl ester Fluo-4 AM in MEM (4 μM in DMSO containing 0.02 % Pluronic F-127, Invitrogen), kept in the dark at room temperature for 1 h, washed twice with Tyrode's buffer (145 mM NaCl, 2.5 mM KCl, 1.5 mM $CaCl_2$, 1.2 mM $MgCl_2$, 10 mM D-glucose, and 10 mM HEPES, pH 7.4), resuspended in the same buffer and transferred (about 100,000 cells) to the quartz cuvette of the spectrofluorimeter (PerkinElmer LS50B equipped with PTP-1 Fluorescence Peltier System; PerkinElmer Life and Analytical Sciences, Waltham, MA, USA) under continuous stirring. The effects on intracellular Ca^{2+} concentration ($[Ca^{2+}]_i$) before and after the addition of various concentrations of test compounds was measured by cell fluorescence (λ_{EX} = 488 nm, λ_{EM} = 516 nm) at 25 °C. The effects of compounds were normalized against the response to ionomycin (4 μM) in each experiment. The increases in fluorescence in wild-type HEK293 cells (i.e. not transfected with any construct) were used as baseline and subtracted from the values obtained from transfected cells. Efficacy was defined as the maximum response elicited by the compounds tested and was determined by comparing their effect with the analogous effect observed with 4 μM ionomycin (Cayman), while the potency of the compounds (EC_{50}) was determined as the concentration required to produce half-maximal increases in $[Ca^{2+}]_i$. Curve fitting (sigmoidal dose–response variable slope) and parameter estimation were performed with GraphPad Prism® (GraphPad Software Inc., San Diego, CA).

Antagonist/desensitizing behavior was evaluated by adding the test compounds in the quartz cuvette 5 min before stimulation of cells with agonists. In the case of human TRPV1-expressing HEK293 cells the agonist used was capsaicin (0.1 μM, in the case of SR141716A 10 nM was also used), which was able of elevating intracellular Ca^{2+} with a potency of EC_{50} = 5.3 ± 0.4 nM and efficacy = 78.6 ± 0.6 %.

For TRPV2, the rat TRPV2-HEK293 cells exhibited a sharp increase in $[Ca^{2+}]_i$ upon application of lysophosphatidylcholine (LPC) 3 μM. The concentration for half-maximal activation was 3.40 ± 0.02 μM and efficacy was 91.7 ± 0.5 %.

In the case of TRPV3, rat TRPV3-expressing HEK-293 cells were first sensitized with the non-selective agonist 2-aminoethoxydiphenyl borate (100 μM). Antagonist/desensitizing behavior was evaluated against thymol (100 μM), which showed an efficacy of 34.7 ± 0.2 % and a potency of EC_{50} = 84.1 ± 1.6 μM. In the case of rat TRPV4-expressing HEK293 cells the agonist used was GSK1016790A, (10 nM), which was able of elevating intracellular Ca^{2+} with a potency of EC_{50} = 0.46 ± 0.07 μM, and an efficacy of 51.9 ± 1.7 %. In the case of rat TRPM8-expressing HEK293 cells, antagonist/desensitizing behavior was evaluated against icilin at 0.25 μM and 0.10 μM. For icilin, efficacy was 75.1 ± 1.1 and potency EC_{50} = 0.11 ± 0.01 μM. In the case of HEK293 cells stably over-expressing recombinant rat TRPA1, the effects of TRPA1 agonists are expressed as a percentage of the effect obtained with 100 μM allyl isothiocyanate (AITC), which showed a potency of EC_{50} = 1.41 ± 0.04 μM and an efficacy of 65.9 ± 0.5.

The effect on $[Ca^{2+}]_i$ exerted by agonist alone was taken as 100%. Data are expressed as the concentration exerting a half-maximal inhibition of agonist-induced $[Ca^{2+}]_i$ elevation (IC_{50}), which was calculated again using GraphPad. All determinations were performed at least in triplicate. Statistical analysis of the data was performed by analysis of variance at each point using ANOVA followed by the Bonferroni's test.

Statistical methods

All statistical measures and methods are included in the respective Figure or Table captions. In brief: all data are shown as the mean ± SEM where applicable. A Student's *t*-test (unpaired, two-tailed) was used to determine differences between two groups. Data that included more than two groups or two factors were

analyzed by two-way ANOVA, with post-hoc Tukey HSD test. All statistical analyses were conducted using Excel or GraphPad Prism version 7, and a p-value less than 0.05 was considered significant throughout unless indicated otherwise. For the lipid profiling study (Figure 5.9), a Benjamini Hochberg correction (25 % false discovery rate) was applied.

A sample size of n=3 was sufficient to detect $\geq$ 50% inhibition of protein labeling with a 20% standard deviation and a power of 80% at a $p < 0.05$. Routinely, a protein is considered to be an off-target, if 50% inhibition of activity is reached at 10 µM. Since, BIA 10-2474 was weakly active on FAAH in in vitro assays (IC50 > 1 µM), the window of the ABPP assays was increased to a concentration of 50 µM detect off-targets.

Synthetic methods

General remarks

All reactions were performed using oven- or flame-dried glassware and dry solvents. Reagents were purchased from Sigma-Aldrich, Acros, and Merck and used without further purification unless noted otherwise. All moisture sensitive reactions were performed under an argon atmosphere.

^1H and ^{13}C NMR spectra were recorded on a Bruker AV 400 MHz spectrometer at 400.2 (^1H) and 100.6 (^{13}C) MHz using $CDCl_3$ as solvent, unless stated otherwise. Chemical shift values are reported in ppm with tetramethylsilane or solvent resonance as the internal standard ($CDCl_3$, δ 7.26 for ^1H, δ 77.16 for ^{13}C). Data are reported as follows: chemical shifts (δ), multiplicity (s = singlet, d = doublet, dd = double doublet, td = triple doublet, t = triplet, q = quartet, quintet = quint, br = broad, m = multiplet), coupling constants J (Hz), and integration. High-resolution mass spectra were recorded on a Thermo Scientific LTQ Orbitrap XL. Liquid chromatography was performed on a Finnigan Surveyor LC/MS system, equipped with a C18 column. Flash chromatography was performed using SiliCycle silica gel type SiliaFlash P60 (230−400 mesh). TLC analysis was performed on Merck silica gel 60/Kieselguhr F254, 0.25 mm. Compounds were visualized using $KMnO_4$ stain (K_2CO_3 (40 g), $KMnO_4$ (6 g), and water (600 mL)).

Synthesis of BIA 10-2639 (**6**) and BIA 10-2474 (**7**)
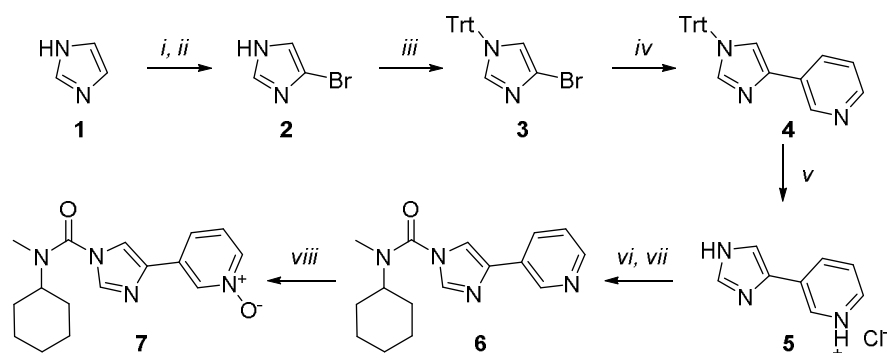
Scheme S5.1 | Reagents and conditions: *i)* NBS, DMF, RT; *ii)* 20% Na_2SO_3 (aq), 100°C, 20%; *iii)* TrtCl, DIPEA, DMF, RT, 93%; *iv)* Cs_2CO_3, Pd(PPh₃)₄, pyridin-3-ylboronic acid, DMF:H_2O (8:1), 95°C, 75%; *v)* 4N HCl in dioxane, MeOH, RT, 69%; *vi)* Na_2CO_3, triphosgene, *N*-methylcyclohexanamine, DCM, 0-20°C; *vii)* cyclohexyl(methyl)carbamic chloride, 3-(1*H*-imidazol-4-yl)pyridine hydrochloride, DMAP, DIPEA, THF, reflux, 74%; *viii)* peracetic acid, DCM, 0-20°C, 97%.

4-Bromo-1*H*-imidazole (**2**)
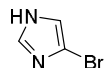
To a solution of 1*H*-imidazole **1** (4 g, 58.8 mmol) in DMF (100 mL) a solution of 1-bromopyrrolidine-2,5-dione (11.50 g, 64.6 mmol) in DMF (100 mL) was added over 1.5 h. Reaction left to stir for 110 h and concentrated to dryness. The residue obtained was taken up in 20% sodium sulfite solution in water and refluxed for 8 h. Upon cooling a precipitate formed which was filtered and dried to yield **2** (1.735 g, 11.81 mmol, 20%) as an off-white solid. ^1H NMR (400 MHz, MeOD) δ 7.61 (s, 1H), 7.11 (s, 1H). Spectroscopic data are in accordance with literature values.[39]

4-Bromo-1-trityl-1*H*-imidazole (**3**)
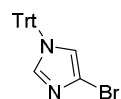
DIPEA (4.06 mL, 23.3 mmol) was added to a solution of trityl chloride (4.87 g, 17.5 mmol) and 4-bromo-1*H*-imidazole (1.71 g, 11.6 mmol) in DMF (20 mL). The resulting mixture was left to stir for 140 h. The resulting suspension was filtered and the yellowish solid was washed with methanol to yield a fine white solid which was purified by column chromatography to yield **3** (4.2 g, 11 mmol, 93%) as a white solid. ^{1}H NMR (400 MHz, CDCl₃) δ 7.43 – 7.30 (m, 10H), 7.19 – 7.04 (m, 6H), 6.80 (s, 1H). ^{13}C NMR (101 MHz, CDCl₃) δ 141.86, 138.71, 129.86, 128.45, 128.33, 120.90. Spectroscopic data are in accordance with literature values.[40]

3-(1-Trityl-1*H*-imidazol-4-yl)pyridine (**4**)
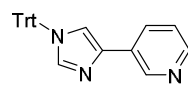
The title compound was synthesised under Suuzuki coupling conditions. Cesium carbonate (6.70 g, 20.6 mmol), **3** (2.00 g, 5.14 mmol) and pyridin-3-ylboronic acid (0.631 g, 5.14 mmol) were suspended in a mixture of DMF (40 mL) and water (5 mL). The resulting mixture was degassed by sonication for 30 min under a flow of argon. Tetrakis(triphenyl-phosphine)palladium(0) (148 mg, 0.128 mmol) was added and the mixture was heated to 90 °C for 16 h. The reaction mixture was diluted with EtOAc (100 mL) and washed with water. The organic layer was separated and the water layer extracted with EtOAc (2x 100 mL). The combined organic layers were washed with brine, dried (MgSO₄), filtered and concentrated. The oily residue was purified by column chromatography to yield **4** (1.50 g, 3.87 mmol, 75%) as a white solid. ^{1}H NMR (400 MHz, CDCl₃) δ 8.95 (m, 1H), 8.42 (m, 1H), 8.03 (m, 1H), 7.99 (s, 1H), 7.68 (m, 1H), 7.56 (d, *J* = 1.4 Hz, 1H), 7.47 – 7.40 (m, 1H), 7.34 (m, 9H), 7.24 (m, 1H), 7.24 – 7.16 (m, 7H). ^{13}C NMR (101 MHz, CDCl₃) δ 147.70, 146.42, 142.06, 139.66, 137.83, 132.04, 131.90, 129.97, 129.68, 128.15, 123.40, 117.92, 75.60.

3-(1*H*-Imidazol-4-yl)pyridine hydrochloride (**5**)
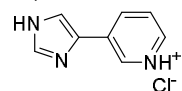
To a solution of **4** (1.24 g, 3.20 mmol) in MeOH HCl in dioxane (4 M, 4.0 mL) was added and the mixture was stirred for 1 h at RT. A white precipitate formed which was filtered off. The filtrate was vigorously stirred and pentane was slowly added to the mixture. More precipitate formed which was filtered off. The resulting solids were air dried to yield **5** (401.3 mg, 2.210 mmol, 69%). ^{1}H NMR (400 MHz, D₂O) δ 9.08 (d, *J* = 2.0 Hz, 1H), 8.84 (d, *J* = 1.3 Hz, 1H), 8.80 – 8.71 (m, 2H), 8.09 (dd, *J* = 8.3, 5.8 Hz, 1H), 8.00 (d, *J* = 1.3 Hz, 1H). ^{13}C NMR (101 MHz, D₂O) δ 143.10, 141.56, 138.96, 136.02, 127.80, 127.33, 127.14, 118.56.

N-Cyclohexyl-*N*-methyl-4-(pyridin-3-yl)-1*H*-imidazole-1-carboxamide (**6**)
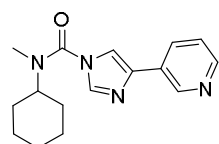
Triphosgene (0.831 g, 2.80 mmol) was added portionwise to a solution of *N*-methylcyclohexanamine (0.73 mL, 5.6 mmol) in DCM at 0 °C. After addition was complete, dried Na₂CO₃ (1.19 g, 11.2 mmol) was added in one batch and the mixture was allowed to warm to RT and stirred for 2 h. Na₂CO₃ was filtered off and the filtrate was concentrated to yield crude cyclohexyl(methyl)carbamic chloride which was used without further purification. To a stirred solution of **5** (203 mg, 1.12 mmol), DMAP (68.4 mg, 0.560 mmol) and DIPEA (0.587 mL, 3.36 mmol) in dry THF (40 mL) at 0 °C was added a solution of crude cyclo-hexyl(methyl)carbamic chloride in 5 mL THF. The resulting mixture was refluxed for 6 h under inert atmosphere. After the reaction was completed the cooled mixture was poured into saturated aqueous NH₄Cl solution and extracted three times with EtOAc. The combined organic layers were washed with brine, dried (MgSO₄), filtered and concentrated *in vacuo* to yield a white solid. Purification by column chromatography yielded an analytically pure sample of **6** (234 mg, 0.823 mmol, 74%). ^{1}H NMR (400 MHz, CDCl₃) δ 9.03 (dd, *J* = 2.3, 0.9 Hz, 1H), 8.54 (dd, *J* = 4.9, 1.6 Hz, 1H), 8.20 (dt, *J* = 8.0, 1.9 Hz, 1H), 7.94 (d, *J* = 1.3 Hz, 1H), 7.60 (d, *J* = 1.3 Hz, 1H), 7.41 (m, 1H), 3.95 (t, *J* = 12.0 Hz, 1H), 3.01 (s, 3H), 1.97 – 1.77 (m, 5H), 1.71 (d, *J* = 13.3 Hz, 1H), 1.59 (qd, *J* = 12.7, 12.2, 4.0 Hz, 2H), 1.46 – 1.31 (m, 2H), 1.14 (qt, *J* = 13.1, 3.5 Hz, 1H). ^{13}C NMR (101 MHz, CDCl₃) δ 150.98, 147.35, 145.67, 138.79, 137.47, 133.39, 129.63, 123.99, 114.32, 57.72, 31.42, 30.03, 25.44, 25.24. Spectroscopic data are in accordance with literature values.[41] HRMS (ESI+) m/z: calculated for $C_{16}H_{21}N_4O$ ([M+H]): 285.1710; found: 285.1710.

3-(1-(Cyclohexyl(methyl)carbamoyl)-1H-imidazol-4-yl)pyridine 1-oxide (7)
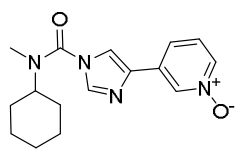
6 (100 mg, 0.352 mmol) was dissolved in dry DCM (25 mL) and cooled to 0 °C. Peracetic acid (0.119 mL, 0.703 mmol) as a 39% solution in acetic acid was added in one batch and the resulting mixture was stirred for 16 h. When TLC showed full conversion the reaction mixture was concentrated and purified by column chromatography to yield 7 (102.6 mg, 0.342 mmol, 97%) as a white solid. ^{1}H NMR (400 MHz, CDCl3) δ 8.72 (t, J = 1.6 Hz, 1H), 8.18 (ddd, J = 6.4, 1.8, 1.0 Hz, 1H), 7.93 (d, J = 1.3 Hz, 1H), 7.77 (dt, J = 8.1, 1.2 Hz, 1H), 7.60 (d, J = 1.3 Hz, 1H), 7.34 (dd, J = 8.0, 6.4 Hz, 1H), 3.99 – 3.84 (m, 1H), 3.00 (s, 3H), 1.87 (t, J = 13.5 Hz, 4H), 1.71 (d, J = 13.4 Hz, 1H), 1.59 (qd, J = 12.8, 12.2, 3.9 Hz, 2H), 1.45 – 1.29 (m, 2H), 1.13 (qt, J = 13.1, 3.5 Hz, 1H). ^{13}C NMR (101 MHz, CDCl3) δ 150.72, 137.84, 137.66, 136.75, 136.18, 133.22, 126.25, 124.17, 115.75, 57.95, 31.60, 30.13, 25.56, 25.33. Spectroscopic data are in accordance with literature values.[42] HRMS (ESI+) m/z: calculated for $C_{16}H_{21}N_4O_2$ ([M+H]): 301.1659; found: 301.1659.

Synthesis of three 2-step probes AJ167 (8), AJ179 (9) and AJ198 (10):
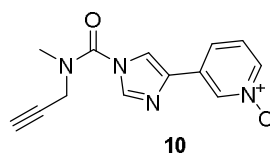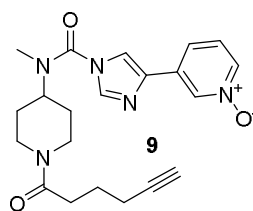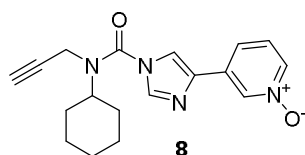
N-(Prop-2-yn-1-yl)cyclohexanamine (11)
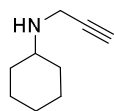
To a stirred solution of prop-2-yn-1-amine (0.50 mL, 7.8 mmol), cyclohexanone (0.404 mL, 3.90 mmol) and sodium triacetoxyborohydride (1.655 g, 7.81 mmol) in dry THF (40 mL) was slowly added one equivalent of acetic acid (0.223 mL, 3.90 mmol). The resulting mixture was stirred for 5 hours at room temperature after which the solvent was evaporated off. The crude was purified by column chromatography to yield 11 (0.521 g, 3.80 mmol, 97%) as a colourless oil. ^{1}H NMR (400 MHz, CDCl3) δ 3.46 (d, J = 2.5 Hz, 2H), 2.74 – 2.59 (m, 1H), 2.21 (t, J = 2.4 Hz, 1H), 1.92 – 1.80 (m, 2H), 1.80 – 1.67 (m, 2H), 1.68 – 1.53 (m, 2H), 1.31 – 0.98 (m, 5H). ^{13}C NMR (101 MHz, CDCl3) δ 82.51, 71.04, 54.95, 35.08, 33.01, 26.11, 24.83.

N-Cyclohexyl-N-(prop-2-yn-1-yl)-4-(pyridin-3-yl)-1H-imidazole-1-carboxamide (12)
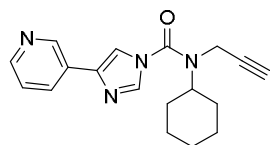
A solution of 11 (227 mg, 1.65 mmol) and triphosgene (163 mg, 0.551 mmol) in dry DCM (30 mL) was cooled to 0 °C and sodium carbonate (175 mg, 1.65 mmol) was added in one batch. The resulting suspension was stirred for 2 hours at room temperature. Na2CO3 was filtered off and the filtrate concentrated in vacuo. To a stirred solution of 5 (100 mg, 0.551 mmol), DMAP (67.3 mg, 0.551 mmol) and DIPEA (385 µL, 2.20 mmol) in dry THF (30 mL) at 0 °C was added a solution of crude carbamic chloride in 5 mL THF. The resulting mixture was refluxed for 4 hours under inert atmosphere. After the reaction was completed the cooled mixture was poured into saturated aqueous NH4Cl solution and extracted three times with EtOAc. The combined organic layers were washed with brine, dried (MgSO4), filtered and concentrated in vacuo to yield a white solid. This was purified by column chromatography to yield an analytically pure sample of 12 (110 mg, 0.357 mmol, 65%). ^{1}H NMR (500 MHz, CDCl3) δ 9.02 (d, J = 2.2 Hz, 1H), 8.53 (dd, J = 1.6, 4.9 Hz, 1H), 8.17 (d, J = 1.3 Hz, 1H), 8.13 (dt, J = 1.9, 7.9 Hz, 1H), 7.82 (d, J = 1.3 Hz, 1H), 7.35 (dd, J = 4.8, 7.9 Hz, 1H), 4.09 (d, J = 2.4 Hz, 2H), 4.00 (tt, J = 3.7, 12.1 Hz, 1H), 2.49 (t, J = 2.4 Hz, 1H), 1.93 (ddt, J = 2.8, 13.7, 41.4 Hz, 4H), 1.70 (qd, J = 3.5, 12.4 Hz, 2H), 1.47 – 1.09 (m, 4H). ^{13}C NMR (126 MHz, CDCl3) δ 150.70, 148.50, 146.69, 139.41, 132.62, 129.11, 123.74, 114.03, 79.91, 73.59, 58.82, 35.02, 30.55, 25.70, 25.26.

3-(1-(Cyclohexyl(prop-2-yn-1-yl)carbamoyl)-1H-imidazol-4-yl)pyridine 1-oxide AJ167 (8)

12 (20 mg, 0.065 mmol) is dissolved in dry DCM (10 mL) and cooled to 0 °C. Peracetic acid (20 µL, 0.118 mmol) as a 39% solution in acetic acid was added in one batch and the resulting mixture was stirred for 16 hours, slowly warming up to RT. The reaction mixture was concentrated and purified by column chromatography to yield 8 (11.5 mg, 0.035 mmol, 55%) as a white solid. ^{1}H NMR

(400 MHz, CDCl₃) δ 8.68 (s, 1H), 8.16 (d, J = 1.1 Hz, 2H), 7.86 – 7.80 (m, 1H), 7.75 (d, J = 8.0 Hz, 1H), 7.34 (t, J = 7.1 Hz, 1H), 4.07 (d, J = 2.4 Hz, 2H), 3.98 (tt, J = 12.0, 3.6 Hz, 1H), 2.49 (t, J = 2.4 Hz, 1H), 2.01 – 1.85 (m, 4H), 1.70 (qd, J = 12.2, 3.5 Hz, 2H), 1.46 – 1.10 (m, 4H). ¹³C NMR (101 MHz, CDCl₃) δ 150.24, 137.87, 137.69, 137.04, 132.80, 126.02, 123.17, 115.26, 79.65, 73.80, 58.89, 35.01, 30.47, 25.64, 25.19. HRMS (ESI+) m/z: calculated for $C_{18}H_{20}N_4O_2$ ([M+H]): 325.1659; found: 325.1658.

1-(Hex-5-ynoyl)piperidin-4-one (13)
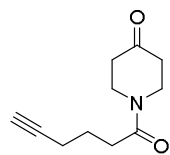
To a stirred solution of hex-5-ynoic acid (0.396 mL, 3.58 mmol) in a 1:1 mixture (v/v) of THF:DCM DIPEA (1.42 mL, 8.14 mmol) and HBTU (2.72 g, 7.16 mmol) were added in one batch. After the solution became homogenous piperidin-4-one hydrochloride hydrate (0.50 g, 3.3 mmol) was added. The mixture was left to stir for 215 h at room temperature. The reaction was quenched with saturated NaHCO₃ (aq.) and extracted three times with EtOAc. The combined organic layers were washed with brine, dried (MgSO₄), filtered and concentrated under reduced pressure. The residue was purified by column chromatography to yield **13** (572 mg, 2.96 mmol, 91%) as a colourless oil. ¹H NMR (400 MHz, CDCl₃) δ 3.90 (t, J = 6.4 Hz, 2H), 3.80 (t, J = 6.3 Hz, 2H), 2.58 (t, J = 7.3 Hz, 2H), 2.53 – 2.46 (m, 4H), 2.33 (td, J = 6.7, 2.7 Hz, 2H), 2.00 (t, J = 2.6 Hz, 1H), 1.91 (p, J = 7.0 Hz, 2H). ¹³C NMR (101 MHz, CDCl₃) δ 206.91, 171.11, 83.64, 69.32, 44.06, 41.29, 40.90, 38.66, 31.43, 23.68, 17.92.

1-(4-(Methylamino)piperidin-1-yl)hex-5-yn-1-one (14)
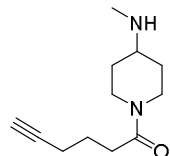
13 (572 mg, 2.96 mmol) was dissolved in DCM and methanamine hydrochloride (400 mg, 5.92 mmol), acetic acid (0.373 mL, 6.51 mmol) and DIPEA (1.09 mL, 6.22 mmol) were added. When cooled to 0 °C sodium triacetoxyborohydride (1.25 g, 5.92 mmol) was added in one portion. The resulting suspension was left to stir for 18 h warming up to RT. The reaction was quenched with a saturated aqueous Na₂CO₃ solution and extracted three times with EtOAc. The combined organic layers were washed with brine, dried (MgSO₄), filtered and concentrated *in vacuo*. The residue was purified by column chromatography to yield **14** (549 mg, 2.64 mmol, 89%) as a colorless oil. ¹H NMR (400 MHz, CDCl₃) δ 4.44 (m, 1H), 3.86 (m, 1H), 3.09 (m, 1H), 2.75 (m, 1H), 2.60 (m, 1H), 2.46 (t, J = 7.3 Hz, 2H), 2.44 (s, 3H), 2.27 (td, J = 6.8, 2.7 Hz, 2H), 2.01 (t, J = 2.6 Hz, 1H), 1.98 – 1.78 (m, 4H), 1.33 – 1.15 (m, 2H). ¹³C NMR (101 MHz, CDCl₃) δ 170.30, 83.68, 69.02, 56.32, 43.98, 40.19, 33.42, 32.55, 31.72, 31.49, 23.84, 17.88.

N-(1-(Hex-5-ynoyl)piperidin-4-yl)-*N*-methyl-4-(pyridin-3-yl)-1*H*-imidazole-1-carboxamide (15)
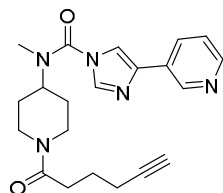
14 (274 mg, 1.32 mmol) was dissolved in DCM (25 mL) and cooled to 0 °C. Sodium carbonate (209 mg, 1.97 mmol) and triphosgene (390 mg, 1.32 mmol) were added and the mixture was stirred for 2 hours at RT. The solids were filtered off and the filtrate was concentrated under reduced pressure and redissolved in THF (25 mL). DIPEA (0.459 mL, 2.63 mmol), **5** (59.7 mg, 0.329 mmol) and DMAP (161 mg, 1.32 mmol) were added and the resulting mixture was refluxed for 4 h. The reaction was quenched with saturated aqueous NaHCO₃ solution and extracted three times with EtOAc. The organic layers were combined, washed with brine, dried (MgSO₄) and filtered. The solvent was removed and the crude purified by column chromatography to yield **15** (54.0 mg, 0.142 mmol, 43%) as a white solid. ¹H NMR (400 MHz, CDCl₃) δ 9.07 – 8.98 (m, 1H), 8.58 – 8.51 (m, 1H), 8.14 (m, 1H), 7.97 (d, J = 1.3 Hz, 1H), 7.61 (d, J = 1.3 Hz, 1H), 7.37 (dd, J = 7.9, 4.8 Hz, 1H), 4.85 (m, 1H), 4.29 (tt, J = 12.2, 4.2 Hz, 1H), 4.13 – 4.00 (m, 1H), 3.22 – 3.09 (m, 1H), 3.02 (s, 3H), 2.63 (m, 1H), 2.51 (t, J = 7.4 Hz, 2H), 2.31 (m, 2H), 1.99 (t, J = 2.6 Hz, 1H), 2.01 – 1.83 (m, 4H), 1.76 (qt, J = 12.2, 3.9 Hz, 2H). ¹³C NMR (101 MHz, CDCl₃) δ 170.67, 151.29, 148.44, 146.52, 139.42, 137.50, 132.79, 129.05, 123.80, 114.03, 83.78, 69.20, 55.57, 44.62, 40.94, 32.15, 31.53, 29.23, 28.53, 23.80, 18.00.

3-(1-((1-(Hex-5-ynoyl)piperidin-4-yl)(methyl)carbamoyl)-1*H*-imidazol-4-yl)pyridine 1-oxide AJ179 (9)
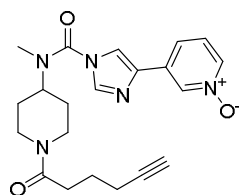
A solution of **15** (25 mg, 0.066 mmol) in DCM (25 mL) was cooled to 0 °C and peracetic acid (50 μL, 0.30 mmol) was added in one batch. The resulting mixture was left to stir for 16 h, after which the solvent was removed *in vacuo*. The resulting crude was purified by column chromatography to yield an analytically pure sample of **9** (4.5 mg, 0.011 mmol, 17%) as a white solid. ¹H NMR (400 MHz, MeOD) δ 8.82 (d, J = 1.7 Hz, 1H), 8.27 (d, J = 6.0 Hz, 1H), 8.22 (d, J = 1.2 Hz, 1H), 8.17 (d, J = 1.2 Hz, 1H),

8.05 (d, J = 8.3 Hz, 1H), 7.61 (dd, J = 8.1, 6.3 Hz, 1H), 4.73 (d, J = 13.3 Hz, 1H), 4.33 – 4.19 (m, 1H), 4.19 – 4.10 (m, 1H), 3.27 – 3.17 (m, 1H), 3.04 (s, 3H), 2.71 (dd, J = 14.0, 11.2 Hz, 1H), 2.61 – 2.55 (m, 2H), 2.32 – 2.25 (m, 2H), 2.06 (s, 1H), 2.02 – 1.74 (m, 6H). ^{13}C NMR (101 MHz, MeOD) δ 171.79, 151.05, 138.53, 137.24, 135.78, 135.46, 133.45, 126.77, 125.67, 116.99, 82.89, 68.90, 55.81, 44.57, 40.74, 31.16, 29.34, 28.56, 27.88, 24.03, 17.18.

N-Methyl-*N*-(prop-2-yn-1-yl)-4-(pyridin-3-yl)-1*H*-imidazole-1-carboxamide (**16**)
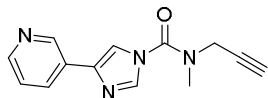
To a stirred solution of *N*-methylprop-2-yn-1-amine (0.084 ml, 1.0 mmol) in DCM (10 mL) at 0 °C sodium carbonate (106 mg, 1.00 mmol) was added. Triphosgene (223 mg, 0.750 mmol) was added and the mixture was stirred at RT for 1 hour. The suspension was then filtered and the filtrate was concentrated *in vacuo* to a yellow oil. This was redissolved in dry THF (10 mL, molecular sieves) and **5** (182 mg, 1.00 mmol) and DIPEA (0.437 mL, 2.50 mmol) were added. The mixture was heated and refluxed for 16 hours. The reaction was quenched by the addition of saturated aqueous NaHCO$_3$ and was extracted three times with EtOAc. The combined organic layers were washed with brine, dried (MgSO$_4$), filtered and concentrated. Purification using column chromatography yielded **16** (68 mg, 0.28 mmol, 28%). ^{1}H NMR (300 MHz, CDCl$_3$) δ 8.99 (s, 1H), 8.45 (d, J = 4.9 Hz, 1H), 8.30 – 8.13 (m, 2H), 8.02 (s, 1H), 7.47 (dd, J = 4.9, 8.1 Hz, 1H), 4.29 (d, J = 1.3 Hz, 1H), 3.22 (s, 4H), 2.93 (q, J = 2.1 Hz, 1H). ^{13}C NMR (75 MHz, CDCl$_3$) δ 148.78, 146.94, 134.57, 125.41, 116.53, 75.23, 40.81, 36.52, 30.74.

3-(1-(Methyl(prop-2-yn-1-yl)carbamoyl)-1*H*-imidazol-4-yl)pyridine 1-oxide AJ198 (**10**)

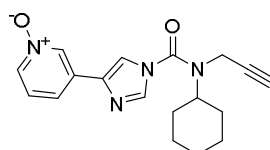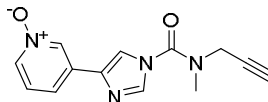**16** (20 mg, 0.083 mmol) was dissolved in DCM (10 mL) and cooled to 0 °C. Peracetic acid (92 μL, 0.42 mmol) as 30% solution in acetic acid was added in one batch. The mixture was stirred for 72 hours, slowly warming up to RT. The crude was concentrated under reduced pressure and purified by column chromatography to yield **10** (12 mg, 0.047 mmol, 56%) as a white solid. ^{1}H NMR (400 MHz, MeOD) δ 8.80 (t, J = 1.7 Hz, 1H), 8.25 (dt, J = 1.3, 6.5 Hz, 1H), 8.22 (d, J = 1.3 Hz, 1H), 8.15 (d, J = 1.3 Hz, 1H), 8.02 (dt, J = 1.2, 8.1 Hz, 1H), 7.58 (dd, J = 6.4, 8.1 Hz, 1H), 4.29 (d, J = 2.4 Hz, 2H), 3.22 (s, 3H), 2.92 (t, J = 2.5 Hz, 1H). HRMS (ESI+) m/z: calculated for C$_{13}$H$_{12}$N$_4$O$_2$ ([M+H]): 257.1033; found: 257.1037.

Supplementary Tables

Table S5.1 | Evaluating FAAH inhibitors against other endocannabinoid-related proteins. DAGL-α, DAGL-β, MAGL, ABHD6 and NAPE-PLD relative to DMSO as vehicle and [^{3}H]CP55940 displacement at overexpressing hCB1- and hCB2-receptor membranes.

Compound		hDAGL-α	mDAGL-β	hMAGL	hABHD6	hNAPE-PLD
		Remaining activity (%) ± SD				
BIA 10-2474	10 µM	102.7 ± 9.3	94.1 ± 4.3	117.9 ± 3.0	83.9 ± 8.2	96.9 ± 11.9
	30 µM	113.4 ± 4.0	102.9 ± 1.9	-	31.2 ± 13.2	-
	50 µM	-	-	-	-	95.9 ± 10.1
PF-04457845	10 µM	99.2 ± 8.7	105.5 ± 0.5	-	108.3 ± 21.6	98.6 ± 6.7
	30 µM	83.0 ± 8.4	106.2 ± 1.5	-	97.5 ± 26.7	-
	50 µM	-	-	-	-	70.2 ± 6.8
BIA 10-2639	10 µM	107.6 ± 2.6	94.5 ± 1.7	-	111.2 ± 8.3	89.0 ± 7.2
	30 µM	110.2 ± 3.2	102.1 ± 4.2	-	90.3 ± 11.2	-
	50 µM	-	-	-	-	89.9 ± 2.6

Table S5.1 (continued)

Compound		hCB$_1$	hCB$_2$
		Radioligand displacement (%) ± SEM	
BIA 10-2474	10 µM	29.8 ± 1.5	6.0 ± 5.4
	30 µM	-	-
	50 µM	38.4 ± 2.3	16.7 ± 8.9
PF-04457845	10 µM	59.9 ± 2.5	51.5 ± 4.6
	30 µM	-	-
	50 µM	103.8 ± 4.6	86.4 ± 1.3
BIA 10-2639	10 µM	36.8 ± 4.1	13.4 ± 8.0
	30 µM	-	-
	50 µM	44.0 ± 5.6	19.9 ± 5.4

Table S5.2 | Evaluating FAAH inhibitors for efficacy and affinity with TRPV-, TRPA1- and TRPM8-channels.

Compound	TRPV1		TRPV2	
	Efficacy (% ionomycin 4 µM)	IC_{50} (capsaicin 0.1 µM)	Efficacy (% ionomycin 4 µM)	IC_{50} (LPC 3 µM)
BIA 10-2474	< 10	> 20 µM	< 10	> 20 µM
PF-04457845	< 10	> 20 µM	< 10	> 20 µM
BIA 10-2639	< 10	> 20 µM	< 10	> 20 µM

Compound	TRPV3		TRPV4	
	Efficacy (% ionomycin 4 µM)	IC_{50} (thymol 100 µM)	Efficacy (% ionomycin 4 µM)	IC_{50} (GSK 10 nM)
BIA 10-2474	< 10	> 20 µM	< 10	> 20 µM
PF-04457845	< 10	> 20 µM	< 10	> 20 µM
BIA 10-2639	< 10	> 20 µM	< 10	> 20 µM

Compound	TRPA1		TRPM8
	Efficacy (%AITC 100 µM)	IC_{50} (AITC 100 µM)	IC_{50} (icilin 0.25 µM)
BIA 10-2474	< 10	> 20 µM	> 20 µM
PF-04457845	< 10	> 20 µM	> 20 µM
BIA 10-2639	< 10	> 20 µM	> 20 µM

References

1. Eddleston, M., Cohen, A. F. & Webb, D. J. Implications of the BIA-102474-101 study for review of first-into-human clinical trials. *Br. J. Clin. Pharmacol.* **81**, 582–586 (2016).
2. Butler, D. & Callaway, E. Scientists in the dark after French clinical trial proves fatal. *Nature* **529**, 263–264 (2016).
3. Kerbrat, A. *et al.* Acute Neurologic Disorder from an Inhibitor of Fatty Acid Amide Hydrolase. *N. Engl. J. Med.* **375**, 1717–1725 (2016).
4. Bégaud, B. *et al.* Report by the Temporary Specialist Scientific Committee (TSSC), 'FAAH (Fatty Acid Amide Hydrolase)', on the causes of the accident during a Phase 1 clinical trial. 1–28 (2016).
5. Cravatt, B. F. *et al.* Molecular characterization of an enzyme that degrades neuromodulatory fatty-acid amides. *Nature* **384**, 83–87 (1996).
6. Kathuria, S. *et al.* Modulation of anxiety through blockade of anandamide hydrolysis. *Nat. Med.* **9**, 76–81 (2002).
7. Devane, W. A. *et al.* Isolation and structure of a brain constituent that binds to the cannabinoid receptor. *Science* **258**, 1946–1949 (1992).
8. Long, J. Z. *et al.* Dual blockade of FAAH and MAGL identifies behavioral processes regulated by endocannabinoid crosstalk in vivo. *Proc. Natl. Acad. Sci.* **106**, 20270–20275 (2009).
9. van der Stelt, M. *et al.* Anandamide acts as an intracellular messenger amplifying Ca2+ influx via TRPV1 channels. *EMBO J.* **24**, 3026–3037 (2005).
10. Hampson, A. J. *et al.* Dual effects of anandamide on NMDA receptor-mediated responses and neurotransmission. *J. Neurochem.* **70**, 671–676 (1998).
11. Huggins, J. P., Smart, T. S., Langman, S., Taylor, L. & Young, T. An efficient randomised, placebo-controlled clinical trial with the irreversible fatty acid amide hydrolase-1 inhibitor PF-04457845, which modulates endocannabinoids but fails to induce effective analgesia in patients with pain due to osteoarthritis of th. *Pain* **153**, 1837–1846 (2012).
12. Li, G. L. *et al.* Assessment of the pharmacology and tolerability of PF-04457845, an irreversible inhibitor of fatty acid amide hydrolase-1, in healthy subjects. *Br. J. Clin. Pharmacol.* **73**, 706–716 (2012).
13. Niphakis, M. J. & Cravatt, B. F. Enzyme Inhibitor Discovery by Activity-Based Protein Profiling. *Annu. Rev. Biochem.* **83**, 341–377 (2014).
14. 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).
15. Liu, Y., Patricelli, M. P. & Cravatt, B. F. Activity-based protein profiling: The serine hydrolases. *Proc. Natl. Acad. Sci.* **96**, 14694–14699 (1999).
16. 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).
17. Baggelaar, M. P. *et al.* Highly Selective, Reversible Inhibitor Identified by Comparative Chemoproteomics Modulates Diacylglycerol Lipase Activity in Neurons. *J. Am. Chem. Soc.* **137**, 8851–8857 (2015).
18. Ahn, K. *et al.* Mechanistic and Pharmacological Characterization of PF-04457845: A Highly Potent and Selective Fatty Acid Amide Hydrolase Inhibitor That Reduces Inflammatory and Noninflammatory Pain. *J. Pharmacol. Exp. Ther.* **338**, 114–124 (2011).
19. Maccarone, M. *et al.* Anandamide Hydrolysis by Human Cells in Culture and Brain. *J. Biol. Chem.* **273**, 32332–32339 (1998).
20. Johnson, D. S. *et al.* Discovery of PF-04457845: A Highly Potent, Orally Bioavailable, and Selective Urea FAAH Inhibitor. *ACS Med. Chem. Lett.* **2**, 91–96 (2011).
21. 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).
22. 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).
23. Thomas, G. *et al.* The Serine Hydrolase ABHD6 Is a Critical Regulator of the Metabolic Syndrome. *Cell Rep.* **5**, 508–520 (2013).
24. Chang, P.-A. & Wu, Y.-J. Neuropathy target esterase: An essential enzyme for neural development and axonal maintenance. *Int. J. Biochem. Cell Biol.* **42**, 573–575 (2010).
25. Shayman, J. A. & Abe, A. Drug induced phospholipidosis: an acquired lysosomal storage disorder. *Biochim. Biophys. Acta* **1831**, 602–11 (2013).
26. Ross, M. K., Streit, T. M. & Herring, K. L. Carboxylesterases: Dual roles in lipid and pesticide metabolism. *J. Pestic. Sci.* **35**, 257–264 (2010).
27. Wei, B. Q., Mikkelsen, T. S., McKinney, M. K., Lander, E. S. & Cravatt, B. F. A second fatty acid amide hydrolase with variable distribution among placental mammals. *J. Biol. Chem.* **281**, 36569–78 (2006).
28. Li, Y. *et al.* Carboxylesterase 2 prevents liver steatosis by modulating lipolysis, endoplasmic reticulum stress, and lipogenesis and is regulated by hepatocyte nuclear factor 4 alpha in mice. *Hepatology* **63**, 1860–1874 (2016).

29. Richardson, R. J., Hein, N. D., Wijeyesakere, S. J., Fink, J. K. & Makhaeva, G. F. Neuropathy target esterase (NTE): overview and future. *Chem. Biol. Interact.* **203**, 238–44 (2013).
30. Read, D. J., Li, Y., Chao, M. V, Cavanagh, J. B. & Glynn, P. Neuropathy target esterase is required for adult vertebrate axon maintenance. *J. Neurosci.* **29**, 11594–600 (2009).
31. Moser, M. *et al.* Cloning and expression of the murine sws/NTE gene. *Mech. Dev.* **90**, 279–82 (2000).
32. Bertrand, D., Bertrand, S., Neveu, E. & Fernandes, P. Molecular characterization of off-target activities of telithromycin: a potential role for nicotinic acetylcholine receptors. *Antimicrob. Agents Chemother.* **54**, 5399–402 (2010).
33. Bachovchin, D. A. *et al.* Superfamily-wide portrait of serine hydrolase inhibition achieved by library-versus-library screening. *Proc. Natl. Acad. Sci.* **107**, 20941–20946 (2010).
34. Hu, C. *et al.* RPLC-ion-trap-FTMS method for lipid profiling of plasma: method validation and application to p53 mutant mouse model. *J. Proteome Res.* **7**, 4982–91 (2008).
35. Gattinoni, S. *et al.* Enol carbamates as inhibitors of fatty acid amide hydrolase (FAAH) endowed with high selectivity for FAAH over the other targets of the endocannabinoid system. *ChemMedChem* **5**, 357–360 (2010).
36. van der Wel, T. *et al.* A natural substrate-based fluorescence assay for inhibitor screening on diacylglycerol lipase α. *J. Lipid Res.* **56**, 927–935 (2015).
37. Peppard, J. V, Mehdi, S., Li, Z. & Duguid, M. S. Assay methods for identifying agents that modify the activity of NAPE-PLD or ABHD4 (WO2008150832A1). (2008).
38. Mukhopadhyay, P. *et al.* The novel, orally available and peripherally restricted selective cannabinoid CB 2 receptor agonist LEI-101 prevents cisplatin-induced nephrotoxicity. *Br. J. Pharmacol.* **173**, 446–458 (2016).
39. Sandtorv, A. H. & Bjørsvik, H.-R. Fast Halogenation of Some N - Heterocycles by Means of N , N′ -Dihalo-5,5-dimethylhydantoin. *Adv. Synth. Catal.* **355**, 499–507 (2013).
40. Roumen, L. *et al.* Synthesis, Biological Evaluation, and Molecular Modeling of 1-Benzyl-1 H -imidazoles as Selective Inhibitors of Aldosterone Synthase (CYP11B2). *J. Med. Chem.* **53**, 1712–1725 (2010).
41. Bial - Portela and Ca, S. A., Russo, D., Wahnon, J. B. R., Maton, W. & Eszenyi, T. Process for the synthesis of substituted urea compounds (WO2014017938A2). (2014).
42. Kiss, L. E. *et al.* Pharmaceutical Compounds (WO2010074588A2). (2010).

Development of a multiplexed ABPP assay to evaluate activity of endocannabinoid hydrolase inhibitors

Introduction

All endocannabinoid hydrolases except *N*-acetylphosphatidylethanolamine phospho-lipase D (NAPE-PLD) belong to the family of serine hydrolases. This family consists of over 200 proteins that use a nucleophilic serine to hydrolyze ester-, amide-, or thioesterbonds in small molecules and proteins via a covalent acyl-protein intermediate.[1,2] This mode of action is exploited in activity-based protein profiling (ABPP).[3,4] The archetypical activity-based probe (ABP) for serine hydrolases is the fluorophosphonate (FP) probe (FP-TAMRA (**1**), Figure 6.1), which was introduced by Liu *et al.* almost 20 years ago.[5] This probe is widely

used to study serine hydrolases in complex proteomes.[6,7] Although the FP-based probes are known for their broad reactivity, they do not react with all serine hydrolases.[4] Most notably, diacylglycerol lipase α (DAGL-α) is amongst the enzymes which cannot be visualized by FP-based ABPs.[8] To study DAGL-α, MB064 ((**2**), Figure 6.1), a tailored chemical probe with a BODIPY-TMR as fluorophore, was developed.[9] In terms of experimental efficiency with respect to time, cost-of-reagents, and use of valuable biological samples, it would be optimal to combine the commercially available FP-TAMRA (ActivX™) and MB064 in the same experiment. However, MB064 cannot be applied in conjunction with FP-TAMRA (**1**), because the excitation/emission spectra of their fluorophores overlap. Therefore, the aim of the current study was to synthesize, characterize and apply a new FP-based probe (**3**) with a different reporter tag (BODIPY-FL) that is compatible with MB064. Such a multiplexed assay, using different activity-based probes, has been shown to work for other enzyme classes in the past.[10-12] Here, a multiplexed ABPP assay with ABP (**3**) and MB064 was developed and used to study endocannabinoid hydrolase activity and to profile inhibitors on activity and selectivity in mouse brain proteomes.
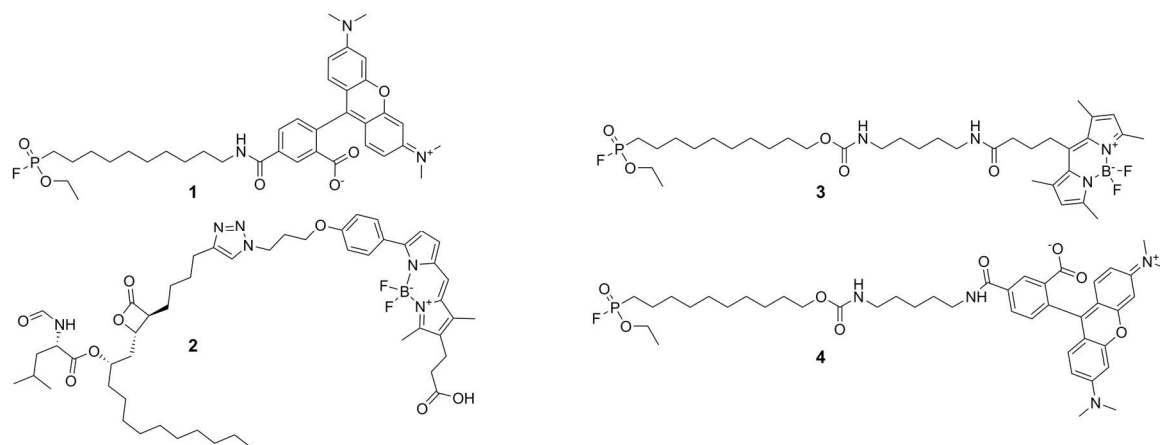
Figure 6.1 | Chemical structures of the four activity-based probes used in this study: FP-TAMRA (**1**), MB064 (**2**), FP-BODIPY (**3**) and a control probe (**4**)

Results and discussion

FP-BODIPY probe (**3**) was synthesized using a previously described method (Scheme S6.1).[13] In addition, commercially available FP-TAMRA (probe **1**) and control compound FP-TAMRA (**4**), containing the same linker as (**3**), were synthesized using similar procedures (Scheme S6.2 and S6.3, respectively).

To obtain a broad view of serine hydrolase labeling by the FP-probes in various tissues, first FP-TAMRA (**1**) and FP-BODIPY (**3**) were incubated with membrane and cytosol fractions of a panel of seven mouse tissues (brain, testes, kidney, spleen, heart, liver and pancreas) at a concentration of 500 nM (Figure 6.2). The proteins were resolved using SDS-PAGE and probe-labeled proteins were visualized by fluorescent scanning of the gel. The overall labeling profile in the various proteomes was comparable between probe **1** and **3**, but several differences were observed, denoted with boxes. In the brain, for example, membrane proteome FP-BODIPY (**3**) labeled additional targets, including in the top left box DAGL-α, the identity of which was confirmed by competition with LEI104 (Figure S6.1).
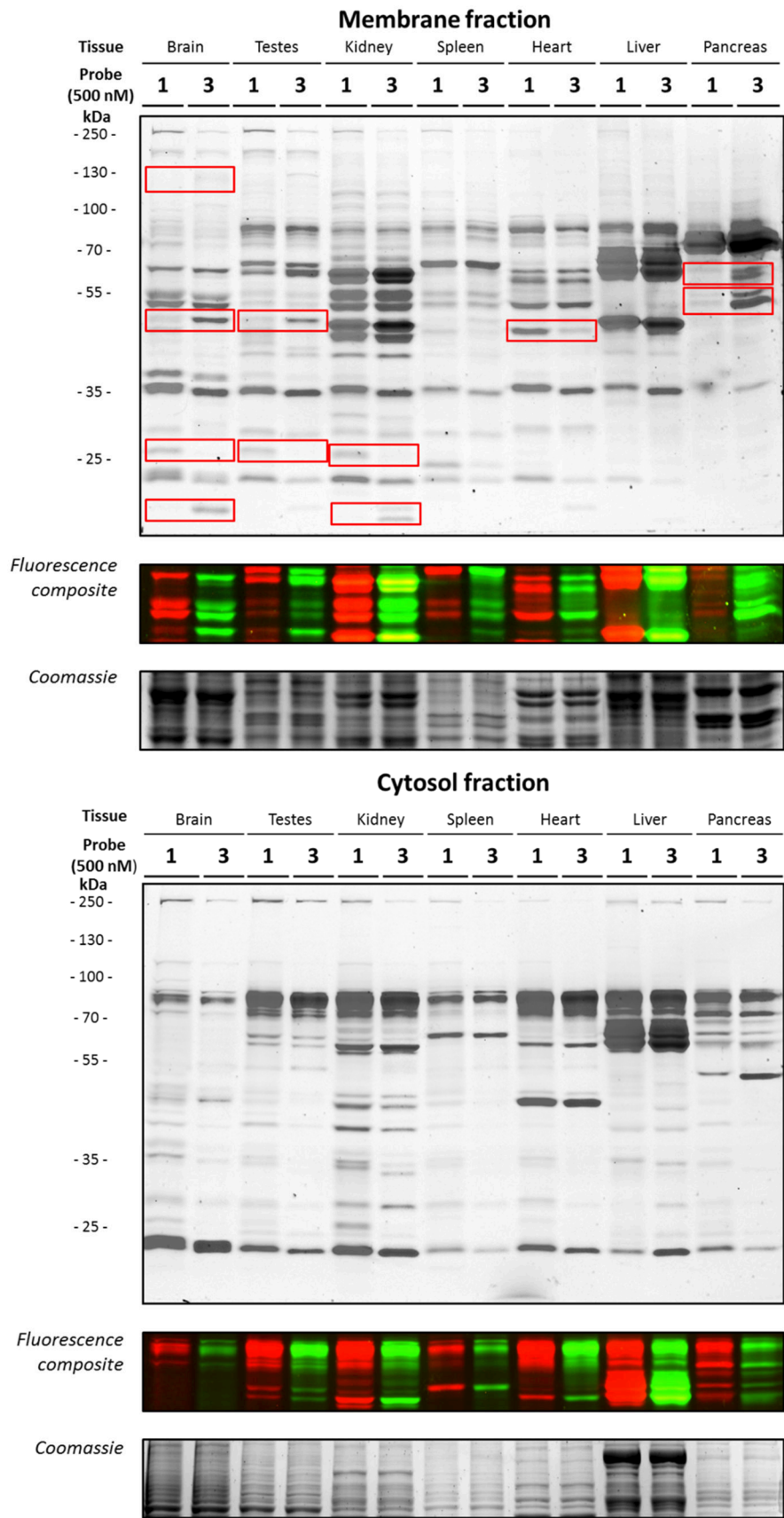
Figure 6.2 | Direct comparison of FP-TAMRA (**1**) and FP-BODIPY (**3**) labeling patterns
of seven mouse tissue lysates. Red boxes highlight strong differences between the
two fluorescent probes.

Detailed profiling of the fluorophosphonate probes **1**, **3** and **4**

Brain lysates were selected for further profiling of probe **1** and **3**, as well as control probe **4**, because the brain is the most studied target organ for ECS modulation. In an initial screen, the three ABPs were incubated with both brain membrane and soluble proteomes (Figure 6.3A, Figure S6.2). While the labeling profile in the soluble proteome was not significantly different between the three probes, FP-BODIPY (**3**) labeled various proteins at lower concentrations in the membrane proteome. To determine the half maximum effect (EC_{50}) values of serine hydrolase labeling in the membrane proteome, the probes were dosed at a wide range of concentrations (10 pM - 10 µM) and the fluorescent signal of 18 distinct bands was quantified and corrected for protein loading by coomassie staining (Figure S6.2 and S6.3). To study the effect of fluorophore and linker length on serine hydrolase labeling by FP-probes in mouse brain membrane proteome, the change in pEC_{50} of ABP **3** and **4** relative to ABP **1** was calculated (Figure 6.3B and Table S6.1). The increased linker length did not significantly alter the labeling efficiency for FP-TAMRA for most proteins, except for fatty acid amide hydrolase (FAAH) (left plot, Figure 6.3C), whereas the change in fluorophore led to a 10-fold increased potency in labeling for several proteins (3, 8 and 18). Of note, FAAH labeling was already visualized at 10 nM FP-BODIPY **3** and DAGL-α (band 3) at 500 nM (Figure 6.3C, A, respectively). The third plot in Figure 6.3B, comparing probe **3** and **4**, which only differ in reporter tag, shows that almost all the difference between the commercial FP-TAMRA probe **1** and the newly synthesized FP-BODIPY probe **3** observed in the central plot is due to the change in fluorophore. The most likely explanation for the observed potency increase when changing from TAMRA to BODIPY-FP is the strong increase in lipophilicity. The CLogP of BODIPY-FL is 3.7 points higher than that of TAMRA, which would make it more favourable to stick to proteins and membranes, possibly causing a higher local concentration and thus better labeling. This explanation is in line with the observation that the strongest differences are observed in the membrane fractions, and, between organs, in the brain. Finally, the impact of the addition of a reporter tag was visualized by pre-incubation with 'dark' alkyne-FP (**5**) (Figure 6.4). This competitive labeling shows that alkyne-FP only completely prevented labeling by the fluorescent probes at 5-10 µM, demonstrating the significantly reduced affinity of the fluorophosphonate inhibitor when lacking the reporter tag. All together, these data demonstrate that the choice of fluorophore influences the labeling efficiency of FP-based probes.
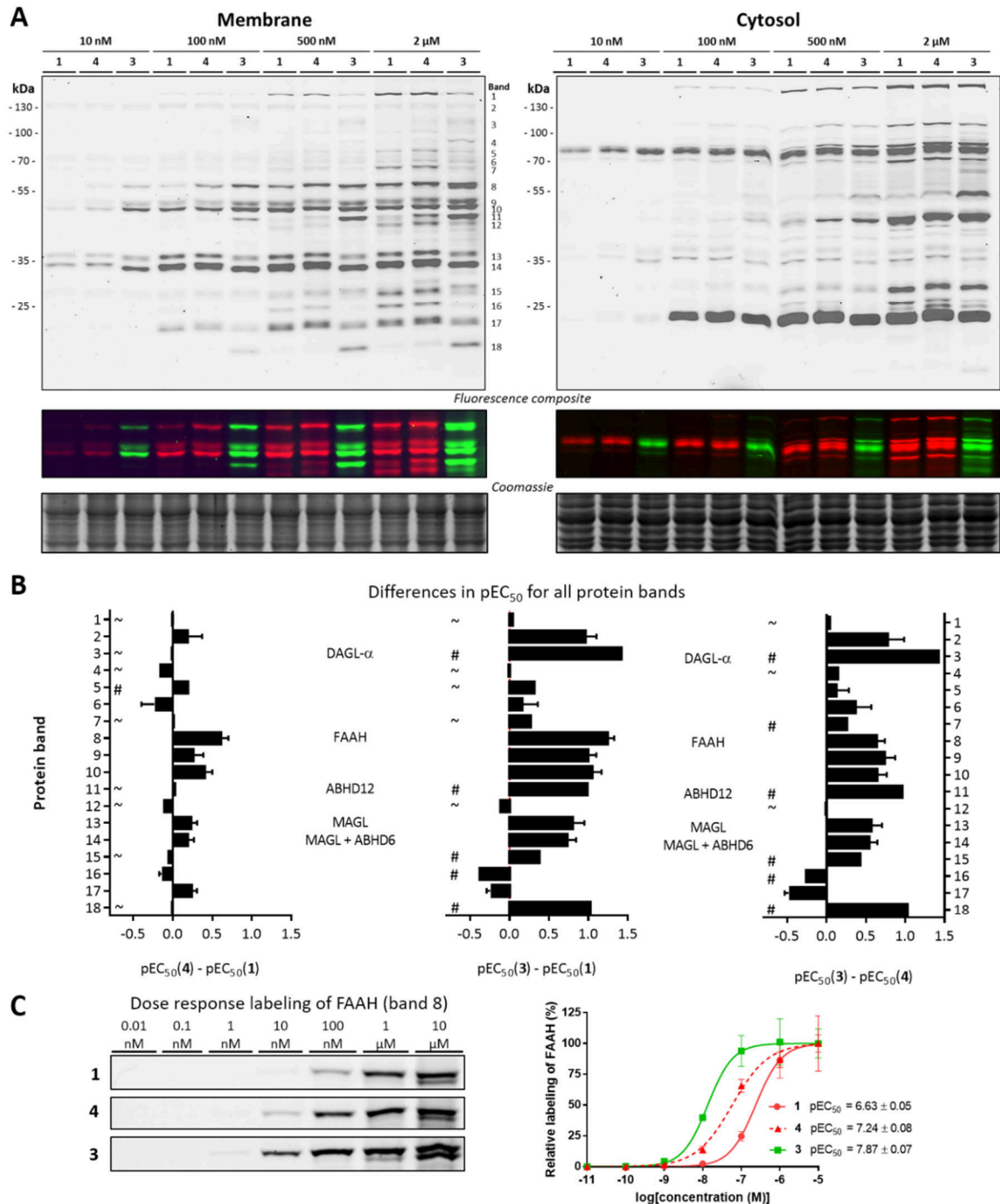
Figure 6.3 | Concentration dependent labeling of probes **1**, **3**, and **4**. A) Four doses of the FP probes label mouse brain proteome in a distinct pattern with different affinities. B) Quantified affinity-differences among the evaluated FP-probes for the 18 bands denoted in A. # indicates a $pEC_{50} \leq 6$ for one of the two probes, meaning that the difference is most likely greater than the given value. ~ indicates a $pEC_{50} \leq 6$ for both probes, meaning that both probes label these proteins only at high concentrations. All quantifications assumed 100% labeling of protein at 10 µM probe. C) Example of the labeling pattern of band 8 (FAAH) and corresponding pEC_{50} curves and values. DAGL-α: diacylglycerol lipase α, ABHD12, ABHD6: α/β hydrolase domain containing protein 12, 6, MAGL: monoacylglycerol lipase, FAAH: fatty acid amide hydrolase.
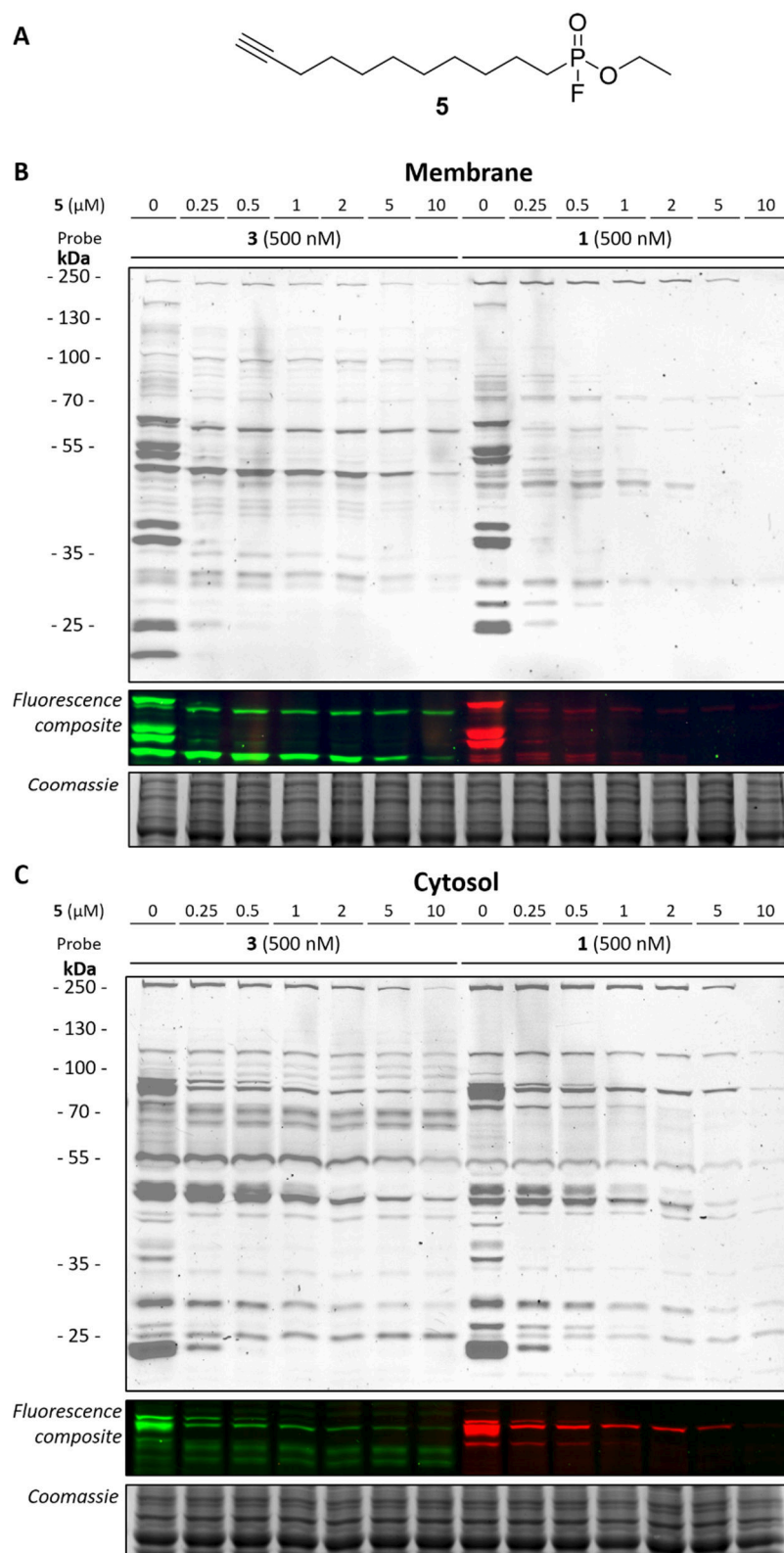
Figure 6.4 | Dose response inhibition of probe **1** and **3** by 'dark' alkyne-FP (**5**). A) Structure of alkyne-FP (**5**) B) Mouse brain membrane proteome was incubated with **5** in increasing dose to prevent labelling by probe **1** and **3**. C) As B) but utilizing mouse brain cytosolic proteome. The Figure is a grey scale composite of two channels. Fluorescence composite is included, green for probe **3** and red for probe **1**. Coomassie is shown as protein loading control.

Next, it was tested whether the activity and selectivity profile of serine hydrolase inhibitors would be dependent on the reporter group of the activity-based probe. To this end, a covalent irreversible FAAH inhibitor, PF-04457845[14], and a reversible inhibitor, LEI104[9] (Figure 6.5), were tested in a competitive ABPP setting using probe **1** (500 nM) and probe **3** (500 and 10 nM) (Figure 6.6A). Importantly, the pIC$_{50}$ values of both inhibitors were not dependent on the fluorescent reporter group of the probe, nor the probe concentration. This indicated that FP-BODIPY **3** can be used in a drug discovery setting to profile inhibitor activity using ABPP.

Development of a multiplexed assay

Having developed two complementary probes (FP-BODIPY (**3**) and MB064) with different reporter groups and distinct labeling patterns, it was tested whether they can be used in a multiplexed ABPP assay to profile the activity and selectivity of compounds inhibiting biosynthetic or catabolic enzymes of the ECS.[6,7,15] To this end, a cocktail of FP-BODIPY **3** (10 nM) and MB064 **2** (250 nM) was used to label mouse brain membrane proteomes. This enabled the simultaneous visualization and quantification of DAGL-α, DDHD2, α/β hydrolase domain containing protein 16a (ABHD16a), FAAH, monoacylglycerol lipase (MAGL), ABHD6 and ABHD12 activities in a single experiment (Figure 6.6B). Bands were identified based on previous studies.[9,16] Phopholipase 2 group IV E (PLA2G4E) and ABHD4 can be labeled by FP-BODIPY and MB064, respectively, but their endogenous expression in brain is too low to be visualized.[17] A panel of inhibitors consisting of JZL184 (MAGL)[18], DH376 (DAGL-α, ABHD6)[15], THL (DAGL-α, ABHD6, ABHD12, ABHD16a, DDHD2)[19], PF-04457845 (FAAH)[14], LEI104 (DAGL-α, FAAH)[9] and LEI105 (DAGL-α)[16] (Figure 6.5) was used to confirm the identity of each fluorescent band (Figure 6.6B). As a final validation, the inhibitory activities of LEI104 on DAGL-α and FAAH were confirmed in this new multiplexed ABPP assay and were found to be in line with previously reported data (Figure 6.6C).[9]
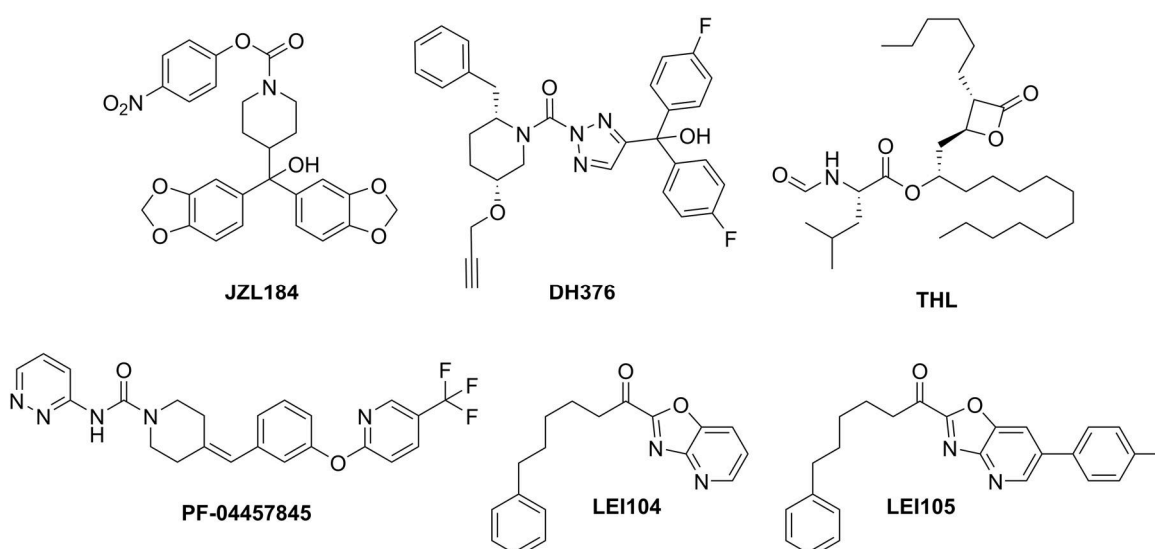
Figure 6.5 | Structures of endocannabinoid serine hydrolase inhibitors used in Figure 6.6.
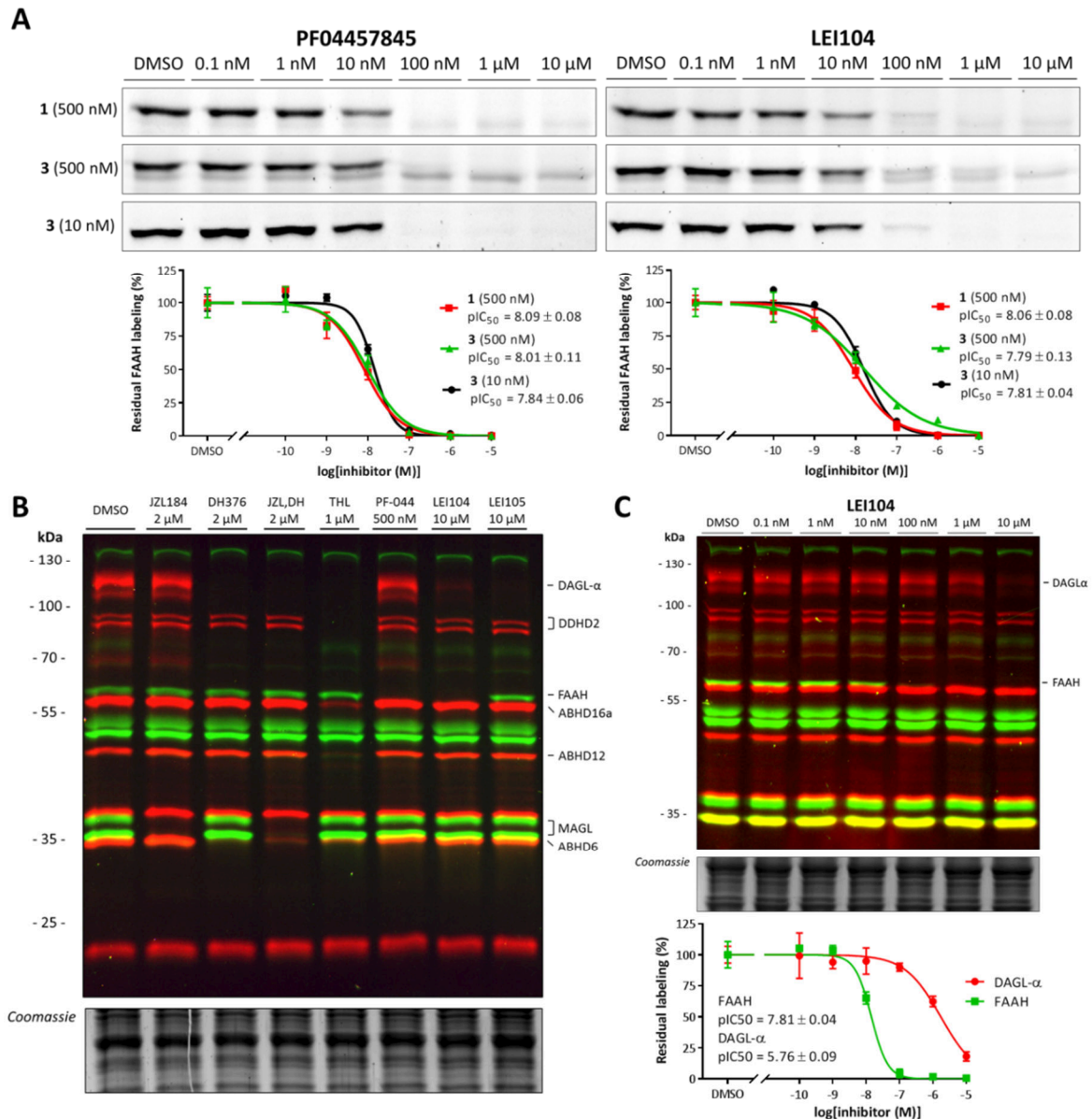
Figure 6.6 | Illustration of the applicability of the prepared probe cocktail. A) Dose response inhibition of FAAH using PF-04457845, covalent irreversible, and LEI104, reversible, to test the dependency of the pIC_{50} determination on probe affinity and concentration. No statistical significant differences have been found between the probe pairs ($p > 0.05$, Two-sided Student's t-test). B) Six inhibitors targeted for different endocannabinoid serine hydrolases were shown to inhibit their specific targets using the probe cocktail. C) Dose response inhibition with LEI104 shows, in one gel, the inhibition of DAGL-α and FAAH. Quantification shows agreement of the pIC_{50} with literature values.[9]

The validated assay was employed to study the selectivity and activity of two MAGL inhibitors. First the recently published β-lactam based MAGL inhibitor NF1819 (**6**) was tested (Figure 6.7A), which was active in several animal models of multiple sclerosis, pain and predator stress-induced long-term anxiety.[20,21] The target-interaction profile of NF1819 (**6**) was compared to the experimental drug ABX-1431 (**7**), currently in phase 2 clinical trials for the treatment of Tourette Syndrome.[22–24] To this end, they were incubated at various concentrations with mouse brain membrane proteome (Figure 6.7B, 6.7C and Figure S6.4). Inhibition of MAGL was confirmed with a pIC_{50} of 8.1 ± 0.1 for **6** and 6.7 ± 0.1 for **7** (Figure 6.7D, 6.7F), in accordance with previously published data.[20,22] Of note, for **6** various off-targets were observed, including ABHD6, LYPLA and an unidentified protein (Figure 6.7E). ABHD6 was inhibited at equal potency, whereas LYPLA demonstrated a 50-fold lower potency. FAAH labeling was only slightly reduced at concentrations > 10 μM. The target-interaction landscape of **7** is clean, even at 10 μM no clear off-targets were observed. The relatively small selectivity window of **6** over ABHD6 should be taken into account during the biological evaluations of this inhibitor as it may contribute to the rise of 2-AG levels.
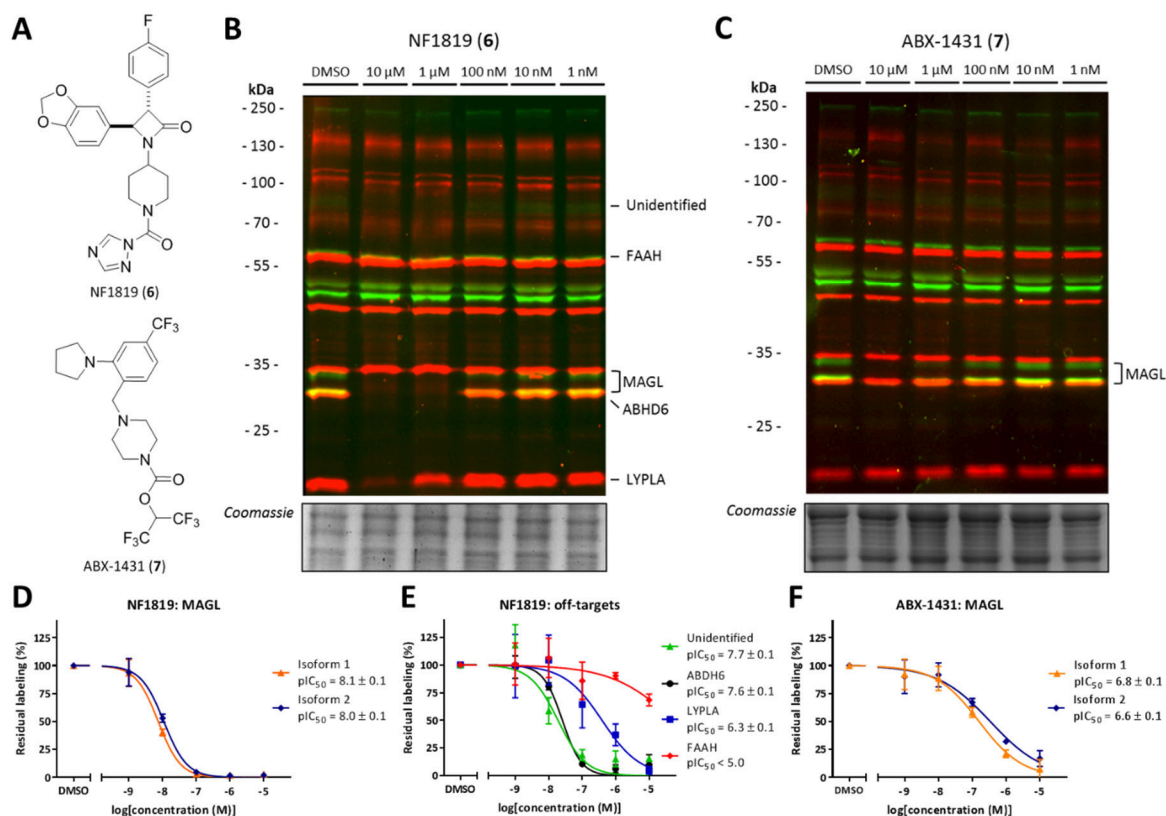
Figure 6.7 | Off-target profiling of β-lactam based MAGL inhibitor (**6**) and clinical candidate ABX-1431 (**7**) in mouse brain membrane proteome. A) Chemical structure of (**6**) and (**7**). B) Dose response inhibition with (**6**) shows several off-targets in mouse brain membrane. C) Dose response inhibition with (**7**) shows selective MAGL inhibition in mouse brain membrane. D) pIC_{50} curves and values of (**6**) against MAGL. E) pIC_{50} curves and values of (**6**) against its off-targets. F) pIC_{50} curves and values of (**7**) against MAGL.

Conclusion

In conclusion, FP-BODIPY (**3**) was synthesized and characterized as a new ABP, extending the chemical toolbox to study serine hydrolase activity in native biological samples. The choice of fluorophore was found to be of great influence on labeling patterns, even for broadly reactive probes such as fluorophosphonates. This finding should be taken into account when designing ABPs. FP-BODIPY (**3**) in conjunction with MB064 (**2**) was used to develop a multiplexed ABPP assay, which was validated by profiling inhibitor activity and selectivity on a broad range of endocannabinoid hydrolases in mouse brain tissue in a single experiment. This multiplexed ABPP assay was then applied to investigate the specificity of a recently published *in vivo* active MAGL inhibitor and an experimental drug currently going through clinical trials.

Methods

Materials

The inhibitors JZL184, PF-04457845 and tetrahydrolipstatin were commercially available (Sigma Aldrich). The syntheses of DH376, LEI104, LEI105, NF1819 and ABX-1431 were previously described.[9,15,16,20,23]

Tissue preparation

Organs were isolated from mice following standard guidelines as approved by the ethical committee of Leiden University (DEC#13191). Isolated organs were frozen in liquid nitrogen and stored at -80 °C. Organs were thawed on ice and homogenized by a glass stick douncing homogenizer in cold lysis buffer (20 mM HEPES, pH 7.2, 2 mM DTT, 1 mM $MgCl_2$, 2.5 U/mL benzonase). The resulting suspension was centrifuged at 1500 rpm for 5 minutes at 4 °C to get rid of residual solid tissue. The supernatant was centrifuged 45 minutes at 93,000 g at 4 °C to separate soluble (cytosol) and insoluble (membrane) fractions. The pellet (membrane) was resuspended in storage buffer (20 mM HEPES, pH 7.2, 2 mM DTT) using an insulin syringe. The protein concentration was measured using a Qubit™ protein assay and was adjusted to 2 mg/mL using storage buffer. The resulting lysates were frozen in liquid nitrogen and stored at -80 °C for later use.

Activity-based protein profiling

Lysate (19 µL, 2 µg/µL) was thawed on ice. For comparative ABPP (one probe and no inhibitor), 1 µL of probe (20x stock in DMSO) was added to the sample, vortexed briefly and incubated for 15 minutes at RT. For competitive ABPP (one inhibitor and one probe), 0.5 µL of the inhibitor (40x stock in DMSO) or pure DMSO (as vehicle) was added to the sample, vortexed briefly and incubated for 20 minutes at RT (37 °C for Figure 6.6A and C to compare with literature values). Subsequently, 0.5 µL probe (40x stock in DMSO) was added to the proteome sample, vortexed briefly and incubated for 15 minutes at RT. Final volume was 20 µL (5% DMSO). The reaction was quenched by the addition of 7.5 µL of 4*Laemmli-buffer (final concentrations: 60 mM Tris (pH 6.8), 2% (w/v) SDS, 10% (v/v) glycerol, 1.25% (v/v) β-mercaptoethanol, 0.01% (v/v) bromophenol blue). 9 µL (12 µg protein) of quenched reaction mixture was resolved on 10% acrylamide SDS-PAGE (180 V, 75 min). Fluorescence was measured using a Biorad ChemiDoc MP system (fluorescence channels Cy2, Cy3, Cy5). Gels were then stained using coomassie staining and imaged for protein loading control.

Labeling quantification

Fluorescence quantification was performed using Imagelab 6.0 (Biorad). Intensities were normalized to 10 µM (for Figure 6.3) or DMSO control (Figure 6.6 and 6.7) and corrected for protein loading by coomassie staining. pEC_{50} values were calculated using GraphPad Prism 7. For all pEC_{50} determinations three replicates of each condition were used.

Statistical methods

Two-sided Student's t-tests were performed using GraphPad Prism 7.0 with the standard Holm-Šídák method to correct for multiple comparisons.

Synthetic methods

General remarks

Reagents were purchased from Sigma Aldrich, Alfa Aesar, AnaSpec or ACROS organics at reagent grade and used without further purification. All moisture sensitive reactions were performed under a nitrogen atmosphere. All solvents were dried using molecular sieves. Ethyl acetate was distilled before use. Glassware was oven dried prior to use. ^{1}H and ^{13}C NMR spectra were recorded on a Bruker DPX-300 (300 MHz), AV-400 (400 MHz) or Bruker DRX-500 (500 MHz). Used software for interpretation of NMR-data was Bruker TopSpin 1.3 and MestreNova 11.0. Chemical shift values are reported in ppm with tetramethylsilane or solvent resonance as the internal standard ($CDCl_3$: δ 7.26 for ^{1}H, δ 77.16 for ^{13}C). ^{31}P-NMR shifts are given in ppm relative to phosphoric acid. ^{19}F-NMR shifts are given in ppm relative to $CFCl_3$. Data are reported as follows: chemical shifts (δ), multiplicity (s = singlet, d = doublet, dd = double doublet, td = triple doublet, t = triplet, q = quartet, bs = broad singlet, m = multiplet), coupling constants J (Hz), and integration. Liquid chromatography analysis was performed on a Finnigan Surveyor LC/MS system, equipped with a C18 column. Flash chromatography was performed using SiliCycle silica gel type SiliaFlash P60 (230−400 mesh). TLC analysis was performed on Merck silica gel 60/Kieselguhr F254, 0.25 mm. Compounds were visualized using $KMnO_4$

stain (K$_2$CO$_3$ (40 g), KMnO$_4$ (6 g), and water (600 mL)) or CAM stain (Ce(NH$_4$)$_4$(SO$_4$)$_4$·2H$_2$O (ceric ammonium sulfate: 10.0 g); ammonium molybdate (25 g); conc. H$_2$SO$_4$ (100 mL); H$_2$O (900 mL)). High resolution mass spectra (HRMS) were recorded by direct injection on a q-TOF mass spectrometer (Synapt G2-Si) equipped with an electrospray ion source in positive mode with Leu-enkephalin (m/z = 556.2771) as an internal lock mass. The instrument was calibrated prior to measurement using the MS/MS spectrum of Glu-1-fibrinopeptide B.
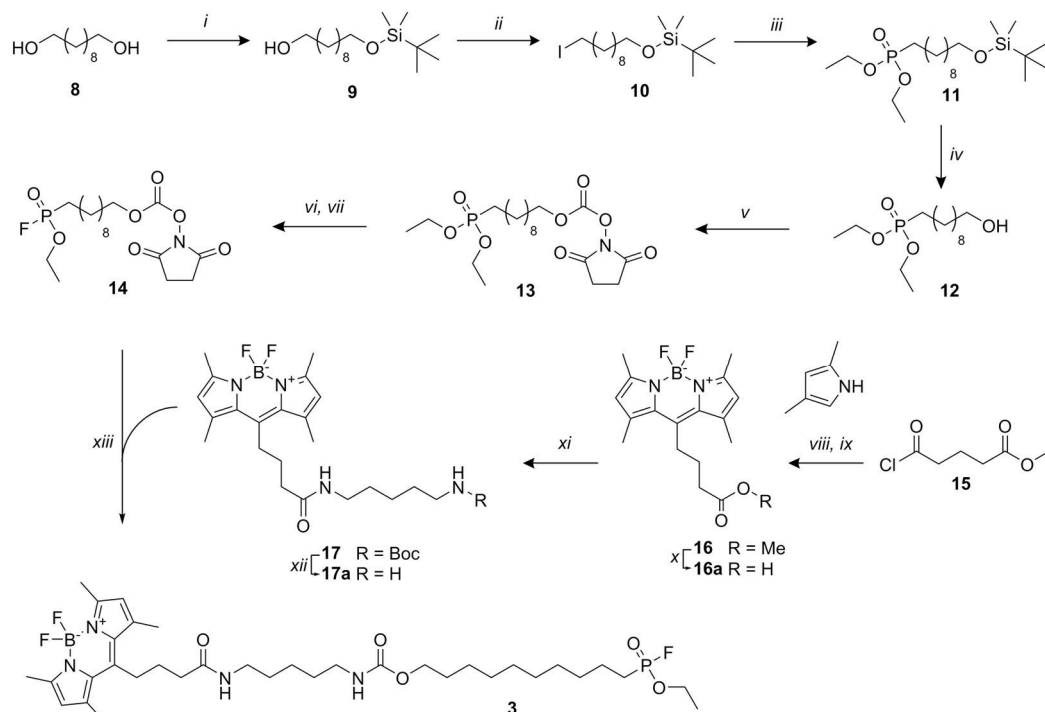
Scheme S6.1 | Synthesis of probe **3**. Reagents and conditions: *i*) TBDMS-Cl, DMAP, imidazole, DMF, RT, 42%; *ii*) I$_2$, triphenylphosphine, THF, RT, 93%; *iii*) POEt$_3$, reflux, 99%; *iv*) TBAF, THF, 0 – 20 °C, 86%; *v*) disuccinylcarbonate, triethylamine, DMF, RT, 91%; *vi*) oxalyl chloride, DCM, RT; *vii*) DAST, DCM, -78 °C, 46%; *viii*) DCM, reflux; *ix*) triethylamine, BF$_3$·OEt$_2$, DCM, toluene, 50 °C, 46%; *x*) NaOH, H$_2$O, THF, RT, 82%; *xi*) *N*-Boc-cadaverine, EDC·HCl, HOBt, triethylamine, DCM, RT, 56%; *xii*) TFA, DCM, RT, 41%; *xiii*) triethylamine, DCM, DMF, RT, 26%.

10-*tert*-Butyldimethylsilane-1-decanol (9)

8 (10.3 g, 57.4 mmol), imidazole (3.91 g, 57.4 mmol) and 4-(dimethylamino)-pyridine (80 mg, 0.65 mmol) were dissolved in dry DMF at RT. Subsequently, *tert*-butyl dimethyl silyl chloride (4.45 g, 28.8 mmol) was added and the mixture was stirred at RT for 5 hours. The mixture was quenched with saturated NaHCO$_3$ (200 mL, aq.) and H$_2$O (200 mL). The mixture was extracted with EtOAc (3x 200 mL) and the organic layers were combined and dried (MgSO$_4$), filtered and evaporated to afford the crude product as a white solid. Purification was performed using column chromatography (5% to 20% EtOAc in pentane). Pure fractions were combined and evaporated to dryness to afford the title compound (3.52 g, 12.1 mmol, 42%). ^{1}H NMR (300 MHz, CDCl$_3$) δ 3.69 – 3.49 (m, 4H), 1.64 – 1.42 (m, 5H), 1.37 – 1.21 (m, 12H), 0.88 (s, 9H), 0.03 (s, 6H). ^{13}C NMR (75 MHz, CDCl$_3$) δ 63.47, 63.14, 32.99, 32.92, 29.68, 29.55, 26.11, 25.91, 25.87, 18.51, -5.13.

10-Iododecyloxy-1-(*tert*-butyl)dimethylsilane (10)

Triphenylphosphine (3.5 g, 13 mmol) and imidazole (1.8 g, 26 mmol) were dissolved in THF (62 mL). Subsequently, iodine (3.5 g, 14 mmol) was slowly added to the solution. The mixture was stirred for 20 minutes at RT under inert atmosphere. Then, **9** (3.5 g, 12 mmol) dissolved in THF (14 mL) was slowly added to the stirred solution. The solution turned brown/red. During the reaction, a yellow precipitate formed. The reaction was allowed to stir at RT for 1 hour. The volatiles were evaporated and the residue was taken up in an emulsion of Et$_2$O (100 mL) and sodium thiosulfate (100 mL, 10%, aq.). The yellow colour disappeared after the addition of the thiosulfate. The organic layer was washed with water and with brine. The organic layer was dried (MgSO$_4$), filtered, and the solvent evaporated to afford white crystals. Purification was done using silica column chromatography (1% Et$_2$O in pentane), affording the

title compound as a colourless oil (4.31 g, 11.3 mmol, 93%). ^{1}H NMR (300 MHz, CDCl$_3$) δ 3.60 (t, J = 6.6 Hz, 2H), 3.19 (t, J = 7.0 Hz, 2H), 1.82 (p, J = 7.0 Hz, 2H), 1.56 – 1.44 (m, 2H), 1.44 – 1.22 (m, 12H), 0.89 (s, 9H), 0.06 (s, 6H). ^{13}C NMR (75 MHz, CDCl$_3$) δ 63.46, 33.72, 33.02, 30.65, 29.66, 29.51, 28.68, 26.14, 25.93, 7.49, -5.09.

Diethyl 10-(*tert*-butyldimethylsiloxy)decylphosphonate (11)
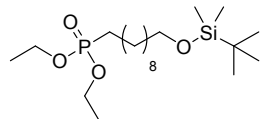
10 (4.31 g, 11.3 mmol) was dissolved in triethylphosphite (10 mL, 58 mmol). The solution was then heated to 155 °C under reflux conditions for 5 hours under inert atmosphere. The resulting solution was then concentrated under reduced pressure to afford the title compound as a light yellow oil (4.4 g, 11 mmol, 99%). ^{1}H NMR (300 MHz, CDCl$_3$) δ 4.21 – 3.98 (m, 4H), 3.59 (t, J = 6.6 Hz, 2H), 1.84 – 1.18 (m, 24H), 0.89 (s, 9H), 0.06 (s, 6H).

Diethyl 10-hydroxydecylphosphonate (12)
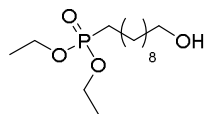
To a cooled (0 °C) solution of **11** (4.4 g, 11 mmol) in dry THF (40 mL), TBAF (12.3 mL, 1 M in THF) was slowly added under inert atmosphere. The reaction was allowed to stand overnight while warming up to RT. The reaction mixture was concentrated and dissolved in 100 mL DCM and washed twice with water (50 mL) and with brine (50 mL). The organic layer was dried (MgSO$_4$), filtered, and concentrated. The resulting yellow oil was purified using column chromatography (5% MeOH in DCM) to afford the title compound as a light yellow oil (2.84 g, 9.65 mmol, 86%). ^{1}H NMR (300 MHz, CDCl$_3$) δ 4.23 – 3.95 (m, 4H), 3.64 (t, J = 6.5 Hz, 2H), 1.88 – 1.46 (m, 9H), 1.46 – 1.18 (m, 16H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 62.98, 61.50 (d, J = 6.6 Hz), 32.87, 30.64 (d, J = 16.8 Hz), 29.56, 29.45, 29.34, 29.12, 25.82, 25.74 (d, J = 140.4 Hz), 22.44 (d, J = 5.2 Hz), 16.56 (d, J = 6.0 Hz). ^{31}P NMR (202 MHz, CDCl$_3$) δ 33.26.

10-(Diethoxyphophoryl) decyl-1-succinimidyl carbonate (13)
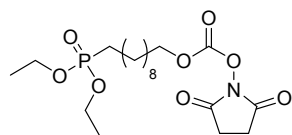
12 (189 mg, 0.64 mmol) was dissolved in 5 mL dry DMF and TEA (0.20 mL, 1.4 mmol) was added. To the stirred solution *N,N'*-disuccinimidyl carbonate (328 mg, 1.28 mmol) was added. The solution was stirred for 40 hours under argon atmosphere at RT. The resulting mixture was taken up in an emulsion of water (20 mL) and EtOAc (20 mL) and stirred for 10 minutes. Then, H$_2$O (50 mL) was added. The layers were separated and the aqueous layer was extracted using EtOAc (3x 20 mL). The combined organic layers were washed with brine, dried (MgSO$_4$), filtered and concentrated. Purification was performed using silica chromatography (5% MeOH in DCM) to afford **13** as a light yellow oil (254 mg, 0.58 mmol, 91%). ^{1}H NMR (300 MHz, CDCl$_3$) δ 4.32 (t, J = 6.7 Hz, 2H), 4.18 – 4.01 (m, 4H), 2.84 (s, 4H), 1.81 – 1.65 (m, 4H), 1.65 – 1.51 (m, 2H), 1.46 – 1.21 (m, 18H). ^{13}C NMR (75 MHz, CDCl$_3$) δ 168.82, 71.74, 61.49 (d, J = 6.5 Hz), 30.68 (d, J = 17.1 Hz), 29.40, 29.33, 29.13, 28.44, 25.77 (d, J = 140.1 Hz), 25.57, 25.50, 22.50 (d, J = 5.2 Hz), 16.59 (d, J = 6.0 Hz).

10-(Ethoxyfluorophophoryl) decyl-1-succinimidyl carbonate (14)
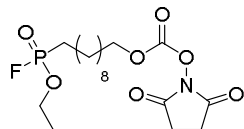
13 (254 mg, 0.58 mmol) was dissolved in 3 mL DCM. Subsequently, oxalyl chloride (0.54 mL, 6.2 mmol) was added and the reaction was stirred for 16 hours at room temperature under N$_2$-atmosphere. The reaction mixture was diluted with DCM (10 mL), then quenched with water (10 mL). The mixture was stirred for 15 minutes. The layers were separated and the aqueous layer was extracted using DCM and EtOAc (2x). The combined organic layers were dried (MgSO$_4$), filtered and concentrated. The residue was dissolved in DCM (3 mL), placed under N$_2$-atmosphere and cooled down to -78 °C. Slowly, diethylaminosulfur trifluoride (0.19 mL, 1.4 mmol) was added. The reaction was stirred for 30 minutes at -78 °C. The reaction was quenched with water (5 mL) and EtOAc (5 mL) and warmed up to room temperature. The layers were separated and the aqueous layer was extracted using EtOAc. The combined organic layers were dried (MgSO$_4$), filtered and concentrated *in vacuo* to afford the crude product. Purification was performed with column chromatography (50% EtOAc in pentane) to afford the title compound as a slightly yellow oil (109 mg, 0.27 mmol, 46%). R$_f$ (50% EtOAc in pentane) = 0.4. ^{1}H NMR (500 MHz, CDCl$_3$) δ 4.32 (t, J = 6.7 Hz, 2H), 4.29 – 4.18 (m, 2H), 2.84 (s, 4H), 1.95 – 1.83 (m, 2H), 1.81 – 1.71 (m, 2H), 1.71 – 1.58 (m, 2H), 1.45 – 1.35 (m, 7H), 1.35 – 1.21 (m, 8H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 168.83, 151.73, 71.75, 63.15 (d, J = 7.3 Hz), 30.37 (d, J = 16.9 Hz), 29.37, 29.24, 29.10, 29.01, 28.45, 25.59, 25.50, 24.41 (dd, J = 143.0, 22.2 Hz), 22.01 (d, J = 5.6 Hz), 16.47 (d, J

= 5.6 Hz). ^{31}P NMR (202 MHz, CDCl3) δ 32.52 (d, *J* = 1070 Hz). ^{19}F NMR (471 MHz, CDCl3) δ -65.03 (d, *J* = 1070 Hz).

4-(4,4-Difluoro-1,3,5,7-tetramethyl-4-bora-3a,4a-diaza-s-indacene-8-yl)-butyric methyl ester (BODIPY-butyric methyl ester) (16)
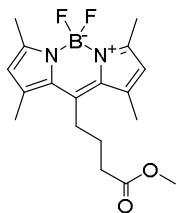
2,4-Dimethylpyrrole (716 mg, 7.3 mmol) was dissolved in DCM (30 mL) and placed under argon atmosphere. Then, methyl 5-chloro-5-oxopentanoate 15 (0.68 mL, 4.9 mmol) was added to the solution. The mixture was refluxed for 2 hours. TEA (3.5 mL, 25 mmol) was dissolved in toluene (20 mL) and was added to the solution. Subsequently, borontrifluoride etherate (4.1 mL, 33 mmol) was added and the solution was heated to 50 °C for 1.5 hours. The mixture was diluted with DCM. The solution was quenched with HCl (1 M, aq.) and water. The layers were separated and the organic layer was washed with HCl (1 M, aq.). The combined aqueous layers were extracted three times with DCM. The combined organic layers were dried (MgSO4), filtered and concentrated. Purification was performed using silica column chromatography (gradient of 10% – 20 % EtOAc in pentane). The product was obtained as a red/brown solid (777 mg). NMR showed presence of TEA·HCl, making up for 25%wt, as calculated by NMR-integrals. The resulting crude was used in subsequent reactions without further purification (584 mg, 1.68 mmol, 46%). ^{1}H NMR (300 MHz, CDCl3) δ 6.06 (s, 2H), 3.70 (s, 3H), 3.10 – 2.91 (m, 2H), 2.57 – 2.34 (m, 14H), 2.03 – 1.86 (m, 2H).

4-(4,4-Difluoro-1,3,5,7-tetramethyl-4-bora-3a,4a-diaza-s-indacene-8-yl)-butyric acid (16a)
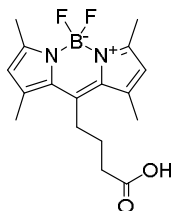
16 (285 mg, 0.82 mmol) was dissolved in THF (5 mL). NaOH (4 mL, 2M, aq.) and water (3 mL) were added to the solution. The reaction mixture was stirred 16 hours at room temperature under inert atmosphere. The mixture was acidified using HCl (6 mL of 2 M aq.) and HCl (1 mL 37%wt aq.). The resulting mixture was extracted twice with EtOAc. The combined organic layers were dried (MgSO4), filtered and concentrated. Purification was performed using column chromatography (5% AcOH and 15% EtOAc in pentane) to afford the title compound as a red solid (225 mg, 0.67 mmol, 82%). Rf (15:4:1 pentane:EtOAc:AcOH) = 0.3. ^{1}H NMR (400 MHz, CDCl3) δ 9.10 (s, 1H), 6.05 (s, 2H), 3.07 – 2.95 (m, 2H), 2.55 (t, *J* = 7.2 Hz, 2H), 2.51 (s, 6H), 2.41 (s, 6H), 2.01 – 1.90 (m, 2H). ^{13}C NMR (101 MHz, CDCl3) δ 178.81, 154.36, 144.85, 140.52, 131.55, 121.96, 34.27, 27.50, 26.63, 16.43, 14.60.

tert-Butyl (5-(4-BODIPY-butanamido)pentyl)carbamate (17)
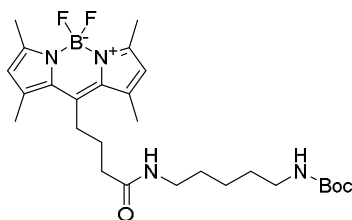
16a (225 mg, 0.67 mmol) was dissolved in DCM (30 mL), to which EDC·HCl (227 mg, 1.29 mmol), TEA (0.30 mL, 2.1 mmol), *N*-Boc cadaverine (0.25 mL, 1.22 mmol) and HOBt (106 mg, 0.78 mmol) were added. The reaction was stirred for 56 hours at room temperature. The reaction was quenched with HCl (20 mL, 0.1 M, aq.) and extracted using DCM. The combined organic layers were dried (MgSO4), filtered and concentrated. Purification using column chromatography (70% EtOAc in pentane) afforded the title compound as an orange solid (206 mg, 0.397 mmol, 56%). ^{1}H NMR (400 MHz, CDCl3) δ 6.05 (s, 2H), 5.77 (s, 1H), 4.60 (s, 1H), 3.27 – 3.19 (m, 2H), 3.11 (q, *J* = 6.6 Hz, 2H), 3.04 – 2.96 (m, 2H), 2.51 (s, 6H), 2.42 (s, 6H), 2.33 (t, *J* = 7.0 Hz, 2H), 2.00 – 1.90 (m, 2H), 1.57 – 1.46 (m, 4H), 1.44 (s, 9H), 1.39 – 1.30 (m, 2H). ^{13}C NMR (101 MHz, CDCl3) δ 156.30, 154.13, 152.80, 145.52, 140.67, 131.59, 121.86, 40.27, 39.57, 36.41, 29.93, 29.18, 28.54, 27.69, 27.58, 23.44, 16.55, 14.60.

N-(5-Aminopentyl)-4-BODIPY-butanamide (17a)
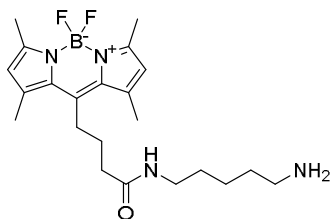
To a solution of 17 (206 mg, 0.40 mmol) in DCM (15 mL), TFA (1.0 mL, 13 mmol) was added under inert atmosphere and the reaction mixture was stirred for 16 hours at RT. The resulting mixture was concentrated to yield the title compound as a dark red solid. LC/MS showed partial removal of BF2-group by TFA. The impure product (68 mg) was used without further purification for subsequent reactions. ^{1}H NMR (400 MHz, MeOD) δ 6.15 (d, *J* = 8.1 Hz, 2H), 3.19 (t, *J* = 7.1 Hz, 2H), 3.09 – 2.97 (m, 2H), 2.71 (t, 2H), 2.45 (d, *J* = 4.9 Hz, 12H), 2.41 – 2.34 (m, 2H), 2.01 – 1.87 (m, 2H), 1.61 – 1.44 (m, 4H), 1.44 – 1.33 (m, 2H).

10-(Ethoxyfluorophosphoryl)decyl (5-(4-BODIPY-butanamido)pentyl)carbamate (3, BODIPY-FP)
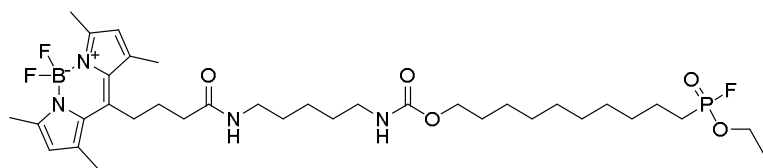
17a (68 mg) was placed under inert atmosphere and dissolved in DCM (3 mL). Dry DMF (2 mL) was added and **14** (42 mg, 0.10 mmol) was added to the solution. The reaction mixture was stirred for 4 hours at RT. The reaction was quenched with water (5 mL) and EtOAc (5 mL) was added. The aqueous layer was extracted with EtOAc (2x) and DCM (2x). The combined organic layers were dried (MgSO$_4$), filtered and concentrated. Purification was performed using silica column chromatography (100% EtOAc) to yield 39 mg of crude product. Further purification was performed on preparative HPLC (H$_2$O/acetonitrile/TFA). Pure fractions were combined and lyophilized to obtain the title compound as a red solid (19 mg, 27 µmol, 26%). ^{1}H NMR (500 MHz, CDCl$_3$) δ 6.05 (s, 2H), 5.79 (s, 1H), 4.72 (s, 1H), 4.33 – 4.19 (m, 2H), 4.07 – 3.96 (m, 2H), 3.29 – 3.20 (m, 2H), 3.20 – 3.11 (m, 2H), 3.05 – 2.95 (m, 2H), 2.51 (s, 6H), 2.43 (s, 6H), 2.39 – 2.29 (m, 2H), 2.01 – 1.93 (m, 2H), 1.93 – 1.83 (m, 2H), 1.71 – 1.47 (m, 8H), 1.43 – 1.24 (m, 17H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 171.77, 157.02, 154.04, 145.32, 140.49, 131.48, 121.74, 64.96, 63.09 (d, J = 7.2 Hz), 40.45, 39.47, 36.27, 30.23 (d, J = 16.8 Hz), 29.71, 29.37, 29.16, 29.13, 29.01, 28.92 – 28.77 (m), 27.56, 27.49, 25.82, 24.94 – 23.29 (m), 21.85 (d, J = 5.2 Hz), 16.45, 16.35 (d, J = 5.6 Hz), 14.47. ^{19}F NMR (471 MHz, CDCl$_3$) δ -64.70 (d, J = 1070 Hz), -144.63 – -150.06 (m). ^{31}P NMR (202 MHz, CDCl$_3$) δ 31.96 (d, J = 1070 Hz). HRMS: [M+H]$^+$ calculated for C$_{35}$H$_{58}$BF$_3$N$_4$O$_5$P: 713.4184; found: 713.4194
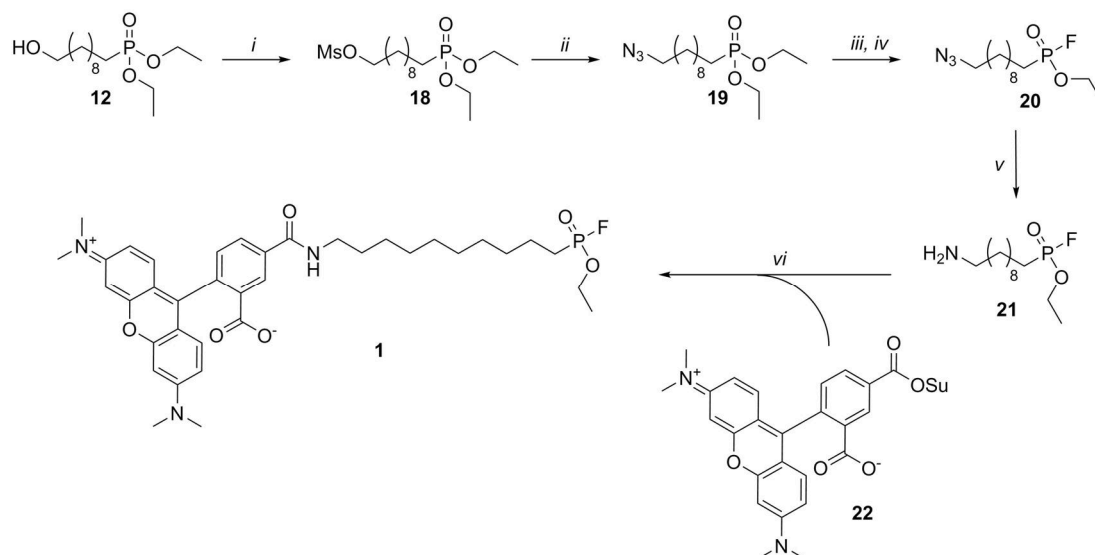
Scheme S6.2 | Synthesis of commercially available TAMRA-FP probe **1.** Reagents and conditions: *i*) MsCl, DIPEA, DCM, 0-20 °C; *ii*) NaN$_3$, DCM, DMF, 50 °C, 45%; *iii*) oxalyl chloride, DCM, 0-20 °C; *iv*) DAST, DCM, -78 °C, 83%; *v*) trimethylphospine, THF, 0-20 °C; *vi*) 5-TAMRA SE (**22**), NEt$_3$, DCM, RT, 54%.

10-(Diethoxyphosphoryl)decyl methanesulfonate (18)
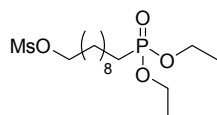
Diethyl (10-hydroxydecyl)phosphonate **12** (147 mg, 0.50 mmol) was dissolved in dry DCM (10 mL) and cooled to 0 °C. Methanesulfonyl chloride (43 µL, 63 mg, 0.55 mmol) and DIPEA (96 µL, 71 mg, 0.55 mmol) were added and the mixture is left to stir for 20 hours, warming to RT. The reaction was quenched by the addition of aqueous HCl (2 N) solution. The aqueous layer was extracted three times with DCM. The combined organic layers were dried (MgSO$_4$), filtered and concentrated yielding crude mesylated product which was used directly in the next reaction.

Diethyl (10-azidodecyl)phosphonate (19)
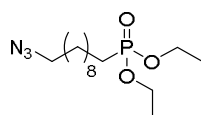
10-(diethoxyphosphoryl)decyl methanesulfonate **18** (186 mg, 0.50 mmol) was dissolved in dry DCM (5 mL). Dry DMF (5 mL) and sodium azide (98 mg, 1.50 mmol) were added and the mixture was stirred at 50 °C for 72 h. The crude mixture was poured into saturated aqueous NaHCO$_3$ and after separation of the layers the water layer was

extracted twice with DCM. The combined organic layers were dried (MgSO$_4$), filtered, and concentrated. Purification by column chromatography (50-90% EtOAc in pentane) afforded the product as a colourless oil (71 mg, 0.22 mmol, 45% yield over two steps). ^{1}H NMR (400 MHz, CDCl$_3$) δ 4.19 – 4.01 (m, 4H), 3.26 (t, J = 7.0 Hz, 2H), 1.80 – 1.66 (m, 2H), 1.66 – 1.52 (m, 4H), 1.42 – 1.20 (m, 18H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 61.43 (d, J = 6.5 Hz), 51.48, 30.69, 30.52, 29.42, 29.27, 29.13, 29.07, 28.85, 26.72, 25.67 (d, J = 140.3 Hz), 22.41 (d, J = 5.2 Hz), 16.51 (d, J = 6.0 Hz).

Ethyl (10-azidodecyl)phosphonofluoridate (20)
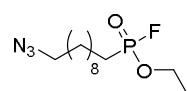
19 (71 mg, 0.22 mmol) was dissolved in dry DCM (5 mL) and stirred on ice. Oxalyl chloride (1.11 mL, 2.22 mmol) was added and stirred overnight, warming up to RT. The reaction was quenched by the addition of water (5 mL), diluted with DCM (5 mL) and stirred for 10 minutes. A cloudy suspension formed. The water layer was separated and extracted twice with DCM (water layer still cloudy) and twice with EtOAc (water layer now clear). The organic layers were combined, dried (MgSO$_4$), filtered and concentrated. The crude ethyl (10-azidodecyl)phosphonate was brought under nitrogen atmosphere, dissolved in dry DCM (3 mL) and cooled to -78 °C. Diethylamino sulfurtrifluoride (88 µL, 0.67 mmol) was added and the resulting mixture stirred at -78 °C for 30 minutes. The reaction was quenched by the addition of a 1:1 mixture (v/v) of water:EtOAc (10 mL). The water layer was separated and extracted twice with EtOAc. The organic layers were combined, dried (MgSO$_4$), filtered, and concentrated. Purification by column chromatography (50-100 % EtOAc in pentane) yielded the title compound (54 mg, 0.18 mmol, 83%). ^{1}H NMR (500 MHz, CDCl$_3$) δ 4.32 – 4.21 (m, 2H), 3.26 (t, J = 7.0 Hz, 2H), 1.95 – 1.83 (m, 2H), 1.72 – 1.54 (m, 4H), 1.46 – 1.15 (m, 15H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 63.11 (d, J = 7.2 Hz), 51.55, 30.40, 30.27, 29.43, 29.24, 29.16, 29.00, 28.91, 26.76, 24.38 (dd, J = 22.3, 143.1 Hz), 21.97 (d, J = 5.5 Hz), 16.43 (d, J = 5.6 Hz). ^{31}P NMR (202 MHz, CDCl$_3$) δ 32.44 (d, J = 1070 Hz). ^{19}F NMR (471 MHz, CDCl$_3$) δ -64.99 (d, J = 1070 Hz).

Ethyl (10-aminodecyl)phosphonofluoridate (21)
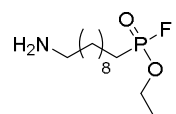
Ethyl (10-azidodecyl)phosphonofluoridate **20** (18 mg, 0.061 mmol) was coevaporated with dry toluene, dissolved in dry THF (5 mL) and cooled on ice. Trimethylphosphine (0.123 mL, 0.123 mmol) was added as a solution in THF (1 M). The resulting mixture was stirred for 3 h at 0°C after which water (0.2 mL) was added. The crude was concentrated *in vacuo* and coevaporated twice with dry toluene. The resulting crude product was used without further purification for the next reaction.

2-(3,6-Bis(dimethylamino)xanthylium-9-yl)-5-((10-(ethoxyfluorophosphoryl)decyl)carbamoyl)benzoate (1, FP-TAMRA)
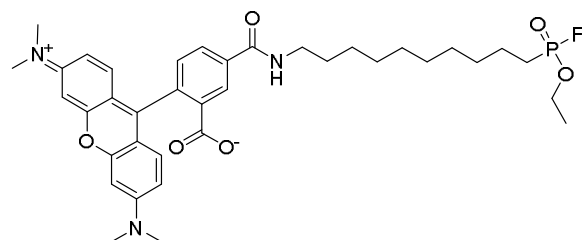
Crude **21** was dissolved in dry DCM and 5-TAMRA succinimidyl ester **22** (5.0 mg, 9.5 µmol) and triethylamine (7.9 µL, 0.057 mmol) were added. The mixture was stirred at RT overnight. TLC showed complete conversion (10% MeOH/EtOAc). The reaction mixture was concentrated *in vacuo* and purified by preparative HPLC. After lyophilization a purple powder was obtained of sufficient (>90 %) purity (3.47 mg, 5.10 µmol, 54% yield over two steps). HRMS: [M+H]$^+$ calculated for C$_{37}$H$_{48}$FN$_3$O$_6$P: 680.3259; found: 680.3250.
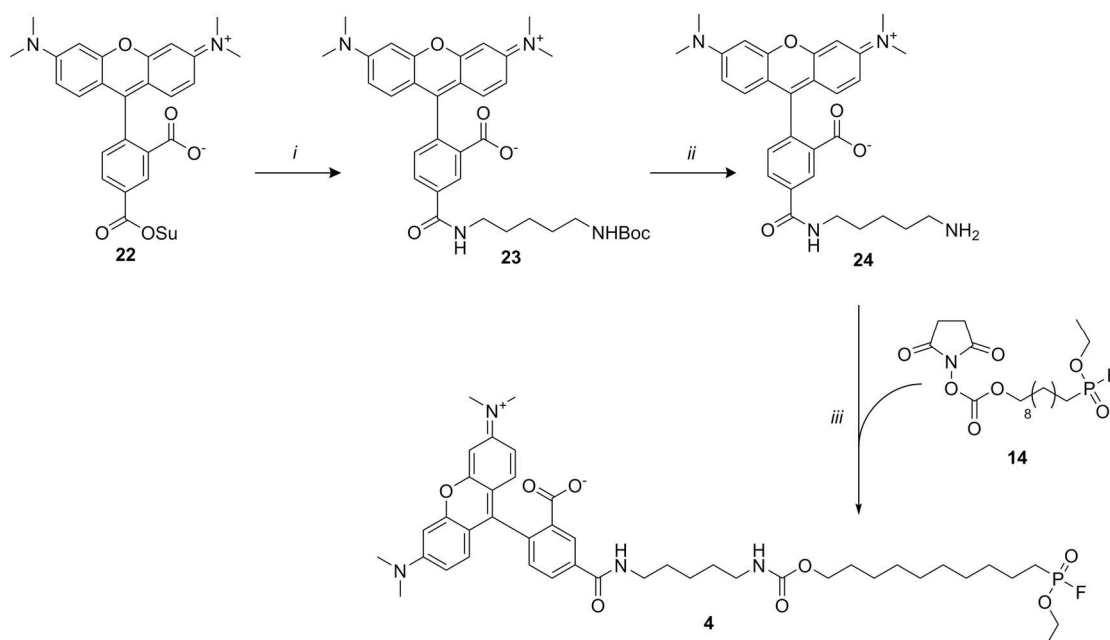
Scheme S6.3 | Synthesis of probe **4**. Reagents and conditions: *i*) *N*-Boc cadaverine, DCM, DMF, RT; *ii*) TFA, DCM, RT; *iii*) **14**, NEt₃, DCM, DMF, RT, 70%.

tert-Butyl (5-(5-TAMRA-amino)pentyl)carbamate (**23**)
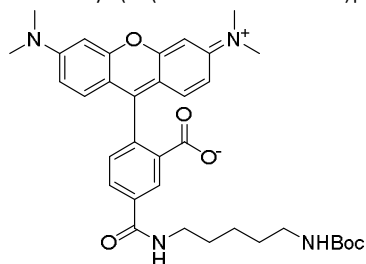
5-TAMRA-succinimidyl ester **22** (5.0 mg, 9.5 µmol) was dissolved in DCM (1.5 mL) and DMF (1 mL). To this solution *N*-Boc cadaverine (2.6 µL, 12 µmol) and TEA (2.4 µL, 17 µmol) were added. The reaction mixture was stirred for 4 hours at room temperature. LC/MS showed full conversion of **22**. The solvents were removed under low pressure to afford the crude product as a pink solid. The resulting crude product was used in subsequent reactions without further purification.

5-(5-TAMRA-amino)pentan-1-amine (**24**)
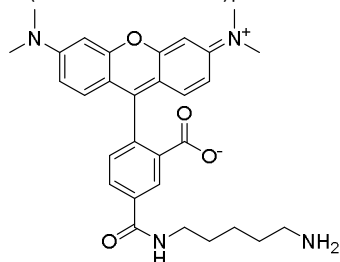
23 was dissolved in DCM (1.4 mL) and TFA (0.1 mL). The reaction mixture was stirred for 3 hours at room temperature, after which LC/MS showed full conversion of **23**. The solvent was removed under low pressure to afford the crude product as a red solid. The resulting crude product was used in subsequent reactions without further purification.

10-(Ethoxyfluorophosphoryl)decyl (5-(4-(3,6-bis(dimethylamino)xanthylium-9-yl)-3-carboxybenzamido)) pentylcarbamate (**4**)
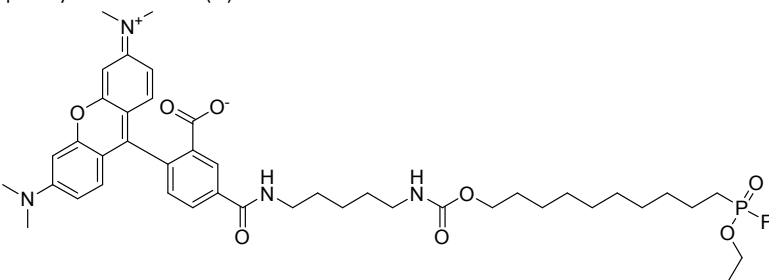
24 was dissolved in DCM (1.5 mL) and DMF (1 mL) and placed under inert atmosphere. **14** (19 mg, 46 µmol) and TEA (7.0 µL, 50 µmol) were added to the solution. The reaction mixture was stirred for 4 hours at room temperature. The solvent was removed under low pressure and purification was performed with preparative HPLC (50-55% acetonitrile in water, 0.1% TFA). Pure fractions were combined, extracted to 1,4-dioxane and lyophilized to yield the title compound as a purple solid (5.4 mg,
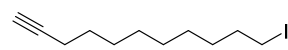
5.8 µmol, 61% over three steps). ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.75 (s, 1H), 8.29 (d, J = 7.8 Hz, 1H), 8.17 (s, 1H), 7.30 (d, J = 7.8 Hz, 1H), 7.19 (d, J = 9.4 Hz, 2H), 6.86 (dd, J = 9.5, 2.4 Hz, 2H), 6.74 (d, J = 2.4 Hz, 2H), 4.98 (s, 1H), 4.30 – 4.20 (m, 2H), 4.13 – 3.96 (m, 2H), 3.50 (q, J = 6.7 Hz, 2H), 3.28 (s, 12H), 3.16 (d, J = 16.4 Hz, 2H), 1.96 – 1.80 (m, 2H), 1.78 – 1.51 (m, 8H), 1.50 – 1.22 (m, 17H). ^{31}P NMR (202 MHz, CDCl$_3$) δ 32.15 (d, J = 1070 Hz). ^{19}F NMR (471 MHz, CDCl$_3$) δ -64.99 (d, J = 1070 Hz). HRMS: [M+H]$^+$ calculated for C$_{43}$H$_{59}$FN$_4$O$_8$P: 809.4049; found: 809.4059.
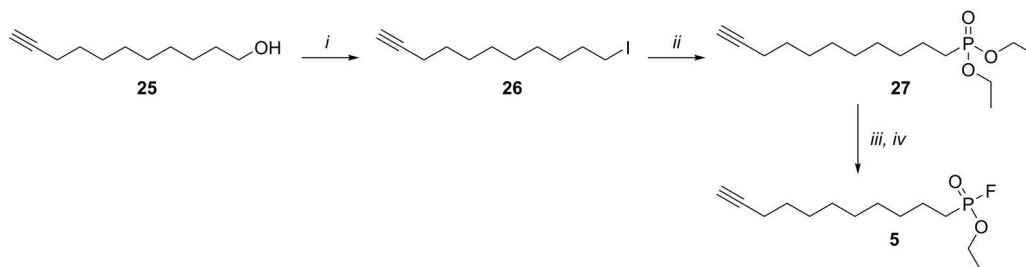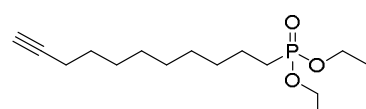
Scheme S6.4 | Synthesis of 'dark' alkyne-FP (**5**). Reagents and conditions: *i*) PPh$_3$, imidazole, I$_2$, DCM, RT, 20%; *ii*) P(OEt)$_3$, 160 °C, quantitative; *iii*) oxalyl chloride, DCM; *iv*) DAST, DCM, -78 °C, 41%.
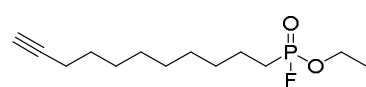
11-Iodoundec-1-yne (26)

Imidazole (102 mg, 1.4 mmol) and triphenylphosphine (355 mg, 1.35 mmol) were dissolved in DCM (10 mL) and iodine (330 mg, 1.3 mmol) was added gradually. After stirring for 20 min at RT, a solution of **25** (172 mg, 1.0 mmol) in DCM (5 mL) was added dropwise to the reaction mixture. After stirring for 16 h a solution of additional imidazole (102 mg, 1.5 mmol), PPh$_3$ (393 mg, 1.5 mmol) and I$_2$ (381 mg, 1.5 mmol) in THF (15 mL) was added to the reaction mixture. After 2 h an aqueous solution of Na$_2$S$_2$O$_3$ (10 mL, 1 M) was added to quench the reaction. The layers were separated and the organic layer was dried (MgSO$_4$), filtered and concentrated *in vacuo*. Purification by column chromatography (0-5% EtOAc in pentane) yielded the title compound as a yellow oil (58 mg, 0.20 mmol, 20%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 3.19 (t, J = 7.0 Hz, 2H), 2.18 (td, J = 2.6, 7.1 Hz, 2H), 1.94 (t, J = 2.6 Hz, 1H), 1.90 – 1.76 (m, 2H), 1.58 – 1.47 (m, 2H), 1.40 – 1.33 (m, 4H), 1.33 – 1.19 (m, 6H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 84.73, 68.13, 33.54, 30.49, 30.33, 29.28, 29.01, 28.69, 28.49, 18.41, 7.35.

Diethyl undec-10-yn-1-ylphosphonate (27)

Compound **28** (58 mg, 0.20 mmol) was dissolved in triethyl phosphite (500 µL) and heated to reflux. After stirring for 4 h the reaction mixture was cooled to RT and concentrated *in vacuo* at 75 °C yielding compound **27** (60 mg, 0.20 mmol, quantitative yield) as a yellow oil that was used without further purification. ^{1}H NMR (400 MHz, CDCl$_3$) δ 4.20 – 4.02 (m, 4H), 2.18 (td, J = 7.0, 2.7 Hz, 2H), 1.95 (t, J = 2.7 Hz, 1H), 1.81 – 1.66 (m, 2H), 1.65 – 1.56 (m, 2H), 1.56 – 1.47 (m, 2H), 1.46 – 1.22 (m, 16H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 84.77, 68.17, 61.65 – 61.32 (m), 30.72, 30.55, 30.37, 29.77, 29.27, 29.08, 28.74, 28.50, 25.72 (d, J = 140.4 Hz), 22.44 (d, J = 5.2 Hz), 18.44, 16.55 (d, J = 6.1 Hz).

Ethyl undec-10-yn-1-ylphosphonofluoridate (5)

A solution of oxalyl chloride in DCM (2.0 M, 1.0 mL, 2.0 mmol) was added to **27** (60 mg, 0.20 mmol) and the reaction mixture was stirred overnight at RT. DCM and H$_2$O were added and after 10 min of stirring the layers were separated and the aqueous layer was extracted two times with DCM. The combined organic layers were dried (MgSO$_4$), filtered and concentrated *in vacuo*. The crude product was coevaporated twice with toluene and redissolved in DCM. The reaction mixture was cooled to -78 °C, after which DAST (110 µL, 0.83 mmol) was added. After stirring for 30 min the reaction mixture was quenched with H$_2$O and aqueous layer was extracted three times with DCM. The combined organic layers were dried (MgSO$_4$), filtered and concentrated *in vacuo*. Purification by column chromatography yielded the title compound (17 mg, 64 µmol, 32%) as a yellow oil. ^{1}H NMR (400 MHz, CDCl$_3$) δ 4.32 – 4.20 (m, 2H), 2.18 (td, J = 2.6, 7.1 Hz, 2H), 1.94 (t, J = 2.7 Hz, 1H), 1.93 – 1.82 (m, 2H), 1.73 – 1.58 (m, 2H), 1.58 – 1.47 (m, 2H), 1.44 – 1.26 (m, 13H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 84.85, 68.24, 63.16 (d, J = 7.3 Hz), 30.48, 30.31, 29.26, 29.11, 29.03, 28.79, 28.56, 24.43 (dd, J = 22.3, 143.0 Hz), 22.02 (d, J = 5.6 Hz), 18.51, 16.50 (d, J = 5.8 Hz). ^{31}P NMR (162 MHz, CDCl$_3$) δ 32.37 (d, J = 1070 Hz). ^{19}F NMR (471 MHz, CDCl$_3$) δ -64.98 (d, J = 1070 Hz). HRMS: [M+Na]$^+$ calculated for C$_{13}$H$_{24}$FNaO$_2$P: 285.1390; found: 285.1395.

Supplementary Tables and Figures
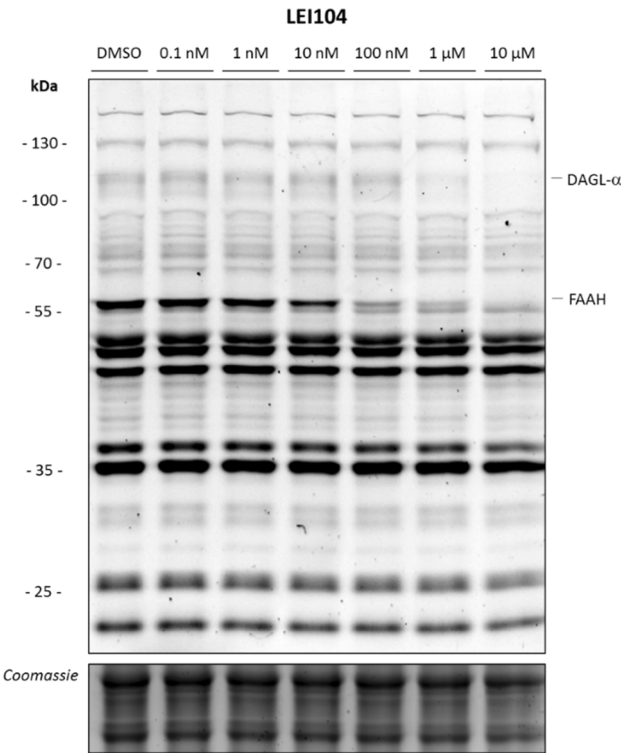
Figure S6.1 | Identification of DAGL-α labeling by probe **3** (500 nM) in mouse brain membrane proteome. LEI104 shows identical inhibition profile of the band at 120 kDa as reported previously when labeled with MB064 (**2**).[9]
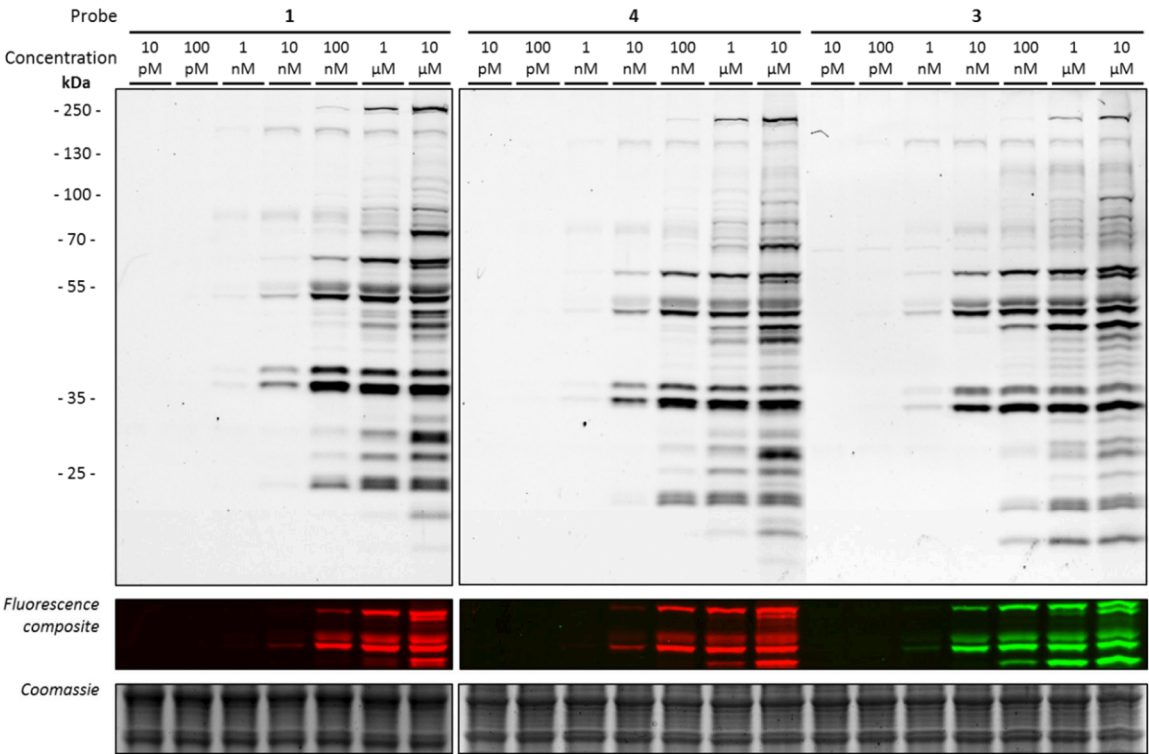
Figure S6.2 | Representative gels of the dose response analysis of the three probes of Figure 6.3. Final quantification was based on 3 gels.
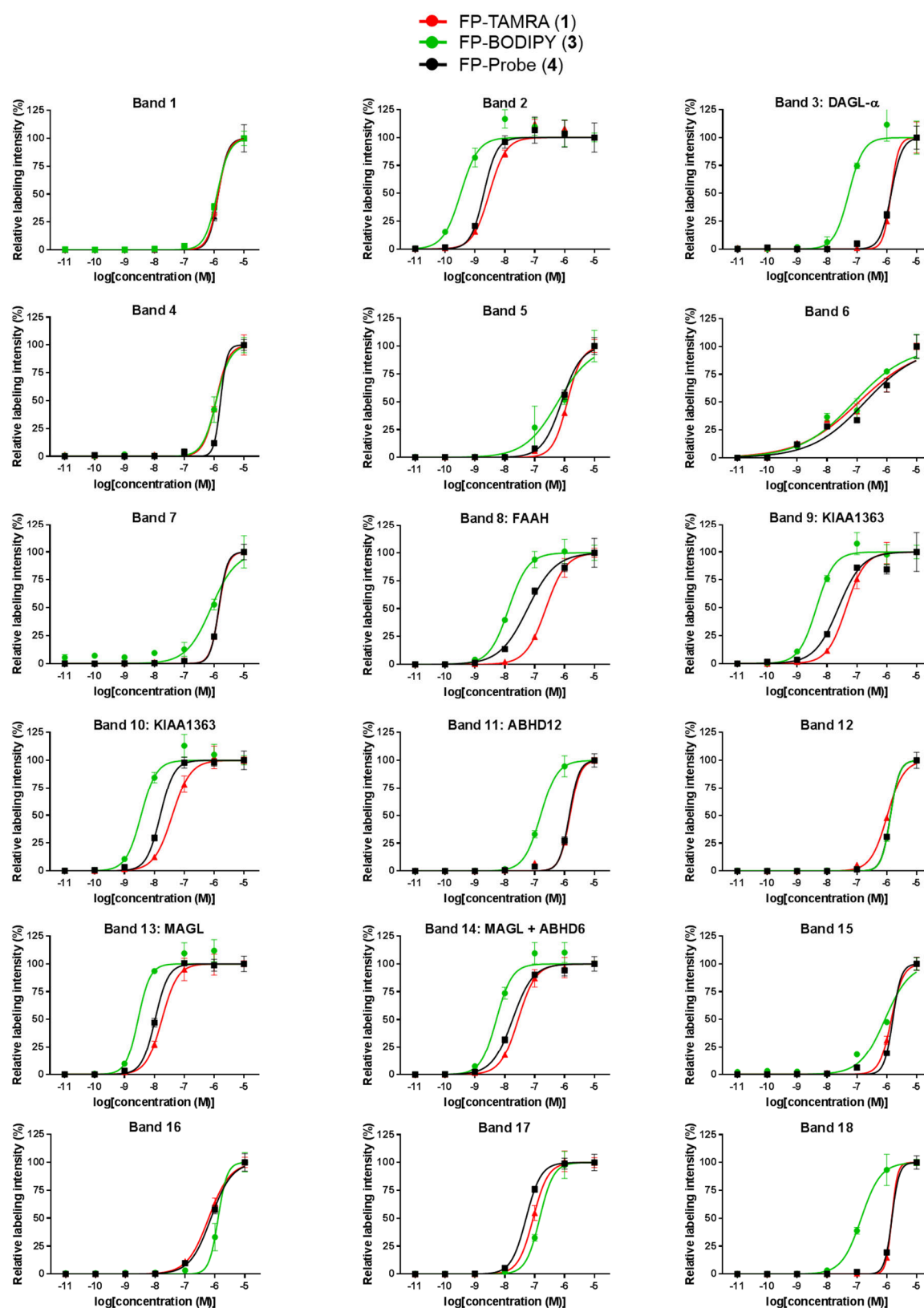
Figure S6.3 | EC₅₀ curves of the 18 quantified bands of Figure 6.3.

Table S6.1 | Overview of pEC$_{50}$ values of FP probes **1**, **3** and **4**. Band numbers correspond to Figure 6.3.

Band	Probe 1 pEC$_{50}$	Probe 1 ± SEM	Probe 3 pEC$_{50}$	Probe 3 ± SEM	Probe 4 pEC$_{50}$	Probe 4 ± SEM
1	≤ 6	-	≤ 6	-	≤ 6	-
2	8.52	0.07	9.48	0.12	8.71	0.17
3	≤ 6	-	7.28	0.09	≤ 6	-
4	≤ 6	0.04	≤ 6	-	≤ 6	-
5	≤ 6	-	≤ 6	-	≤ 6	-
6	6.94	0.13	7.10	0.15	6.73	0.14
7	≤ 6	-	≤ 6	-	≤ 6	-
8	6.63	0.05	7.87	0.07	7.24	0.08
9	7.36	0.07	8.35	0.08	7.62	0.11
10	7.4	0.07	8.45	0.10	7.80	0.07
11	≤ 6	-	6.80	0.05	≤ 6	-
12	≤ 6	-	5.87	0.09	≤ 6	-
13	7.75	0.07	8.55	0.13	7.98	0.03
14	7.55	0.07	8.28	0.09	7.74	0.05
15	≤ 6	-	6.22	0.09	≤ 6	-
16	6.21	0.04	≤ 6	-	6.09	0.04
17	7.05	0.04	6.83	0.06	7.29	0.05
18	≤ 6	-	6.85	0.05	≤ 6	-
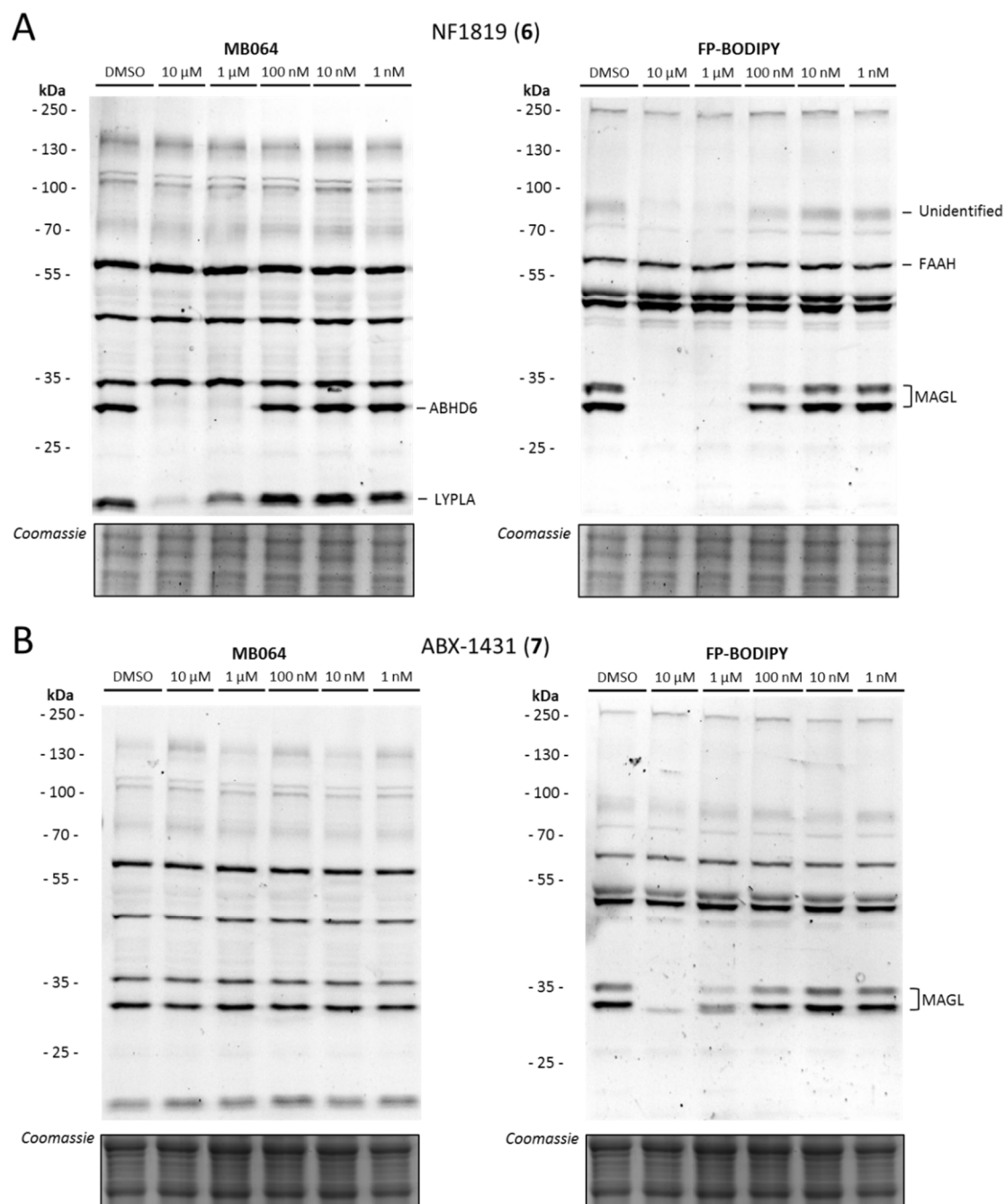
Figure S6.4 | Dose response inhibition of the probe cocktail by NF1819 (**6**) (A) and ABX-1431 (**7**) (B) in mouse brain membrane proteome, split by fluorescence channel to visualize overlapping bands of MAGL and ABHD6.

References

1. Long, J. Z. & Cravatt, B. F. The Metabolic Serine Hydrolases and Their Functions in Mammalian Physiology and Disease. *Chem. Rev.* **111**, 6022–6063 (2011).
2. Shahiduzzaman, M. & Coombs, K. M. Activity based protein profiling to detect serine hydrolase alterations in virus infected cells. *Front. Microbiol.* **3**, 1–5 (2012).
3. Wang, S. *et al.* Advanced Activity-Based Protein Profiling Application Strategies for Drug Development. *Front. Pharmacol.* **9**, (2018).
4. Simon, G. M. & Cravatt, B. F. Activity-based Proteomics of Enzyme Superfamilies: Serine Hydrolases as a Case Study. *J. Biol. Chem.* **285**, 11051–11055 (2010).
5. Liu, Y., Patricelli, M. P. & Cravatt, B. F. Activity-based protein profiling: The serine hydrolases. *Proc. Natl. Acad. Sci.* **96**, 14694–14699 (1999).
6. 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).
7. 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).
8. Hoover, H. S., Blankman, J. L., Niessen, S. & Cravatt, B. F. Selectivity of inhibitors of endocannabinoid biosynthesis evaluated by activity-based protein profiling. *Bioorg. Med. Chem. Lett.* **18**, 5838–41 (2008).
9. 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).
10. Adam, G. C., Sorensen, E. J. & Cravatt, B. F. Proteomic profiling of mechanistically distinct enzyme classes using a common chemotype. *Nat. Biotechnol.* **20**, 805–809 (2002).
11. 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).
12. Patricelli, M. P., Giang, D. K., Stamp, L. M. & Burbaum, J. J. Direct visualization of serine hydrolase activities in complex proteomes using fluorescent active site-directed probes. *Proteomics* **1**, 1067–1071 (2001).
13. Tully, S. E. & Cravatt, B. F. Activity-Based Probes That Target Functional Subclasses of Phospholipases in Proteomes. *J. Am. Chem. Soc.* **132**, 3264–3265 (2010).
14. Johnson, D. S. *et al.* Discovery of PF-04457845: A Highly Potent, Orally Bioavailable, and Selective Urea FAAH Inhibitor. *ACS Med. Chem. Lett.* **2**, 91–96 (2011).
15. 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).
16. Baggelaar, M. P. *et al.* Highly Selective, Reversible Inhibitor Identified by Comparative Chemoproteomics Modulates Diacylglycerol Lipase Activity in Neurons. *J. Am. Chem. Soc.* **137**, 8851–8857 (2015).
17. Ogura, Y., Parsons, W. H., Kamat, S. S. & Cravatt, B. F. A calcium-dependent acyltransferase that produces N-acyl phosphatidylethanolamines. *Nat. Chem. Biol.* **12**, 669–671 (2016).
18. Pan, B. *et al.* Blockade of 2-arachidonoylglycerol hydrolysis by selective monoacylglycerol lipase inhibitor 4-nitrophenyl 4-(dibenzo[d][1,3]dioxol-5-yl(hydroxy)methyl)piperidine-1-carboxylate (JZL184) Enhances retrograde endocannabinoid signaling. *J. Pharmacol. Exp. Ther.* **331**, 591–7 (2009).
19. Bisogno, T. *et al.* Cloning of the first sn1-DAG lipases points to the spatial and temporal regulation of endocannabinoid signaling in the brain. *J. Cell Biol.* **163**, 463–468 (2003).
20. Brindisi, M. *et al.* Development and Pharmacological Characterization of Selective Blockers of 2-Arachidonoyl Glycerol Degradation with Efficacy in Rodent Models of Multiple Sclerosis and Pain. *J. Med. Chem.* **59**, 2612–2632 (2016).
21. Lim, J. *et al.* Endocannabinoid Modulation of Predator Stress-Induced Long-Term Anxiety in Rats. *Neuropsychopharmacology* **41**, 1329–1339 (2015).
22. Blankman, J. L. *et al.* Methods of treating inflammation or neuropathic pain (WO2016183097A1). (2015).
23. Justin S. Cisar *et al.* Carbamate compounds and of making and using same (US20150018335A1). (2012).
24. Cisar, J. S. *et al.* Identification of ABX-1431, a Selective Inhibitor of Monoacylglycerol Lipase and Clinical Candidate for Treatment of Neurological Disorders. *J. Med. Chem.* **61**, 9062–9084 (2018).

t-Distributed Stochastic Neighbour Embeddings applied to drug discovery

Introduction

Chemical space is vast and can only be explored to a small extent by experimental methods to find suitable hits for drug discovery programs.[1,2] The search for new chemical starting points to modulate therapeutic targets is essential for the development of novel drugs. Due to the difficulties in finding suitable hits, it has been postulated that the best way to find a new drug is to start with an old drug.[3] This is in line with the central paradigm in medicinal chemistry that similar structures exert similar biological activities.[4] This paradigm can be mirrored to state that similar biological entities will be subject to modulation by similar molecular structures. It is this latter formulation that explains many

of the off-target activities observed when studying the inhibitory activity of a hit molecule against a subset of enzymes performing a similar function, which are part of the same cascade, or belong to a closely related family. An example from the endocannabinoid field is LEI104, which was found to be active against fatty acid amide hydrolase (FAAH) and diacylglycerol lipase (DAGL) α and DAGL-β. FAAH is responsible for the hydrolysis of *N*-arachidonoylethanolamine (anandamide), whereas the DAGLs produce 2-arachidonoyl-glycerol (2-AG).[5] The high similarity in structure of the substrate and product, respectively, of these enzymes implies that the binding pockets of the enzymes are similar, allowing LEI104 to bind to both.

Since most serine hydrolase inhibitors interact with the catalytically active amino acid, often in a covalent manner, off-target activity across this family is commonly observed for these compounds.[6] Such off-target activity is usually unwanted, and may sometimes lead to profound toxic effects as discussed in Chapter 5.[7] Despite the efficiency of activity-based protein profiling and other techniques to determine the selectivity of lead inhibitors, extensive screening campaigns are still expensive and time consuming. In the last decade, several large datasets with structure-activity relationships (SAR) of serine hydrolase inhibitors spanning a number of enzymes have been published.[6,8] These empirical datasets may serve as guides to explore chemical space around this drug target family and predict (off-)target activity using advanced computational chemistry methods, such as quantitative SAR models (QSAR), similarity ensemble approach (SEA), support vector machines, k-nearest neighbour, random forest, naïve bayes, (deep learning) neural networks (NN) and principal component analysis (PCA).[9–12]

Advanced machine learning models promise to revolutionize the field of drug discovery. Employing large high-dimensional datasets, these models are used to predict a wider range of biological activities for a compound, compared to traditional drug design methods (e.g. molecular modelling, docking and early QSAR-models, such as Hansch and Free-Wilson analysis[13]). Yet, many machine learning models are hampered in their applicability by lack of a clear interpretation and tendency to overfit on the high-dimensional data. Many of the best performing machine learning models are a black box in which it is unclear in what way the data are used to generate novel hypotheses. They also require in-depth knowledge of advanced cheminformatics and highly specialised or purpose-build software. These technical requirements slow down the implementation of the tools in the daily practice of drug discovery and consequently prevent the research community to take full advantage of the wealth of data becoming available. Therefore, there is a clear need for better tools to interpret and visualize complex, high-dimensional SAR datasets in an easy and intuitive manner and to predict the biological activity profile of novel hits for drug discovery programs.

t-Distributed Stochastic Neighbour Embedding

t-SNE was first published by van Maaten and Hinton in 2008 as the third variant of the Stochastic Neighbour Embedding concept.[14,15] The aim of the algorithm is to generate visualizations of high dimensional data by generating a two or three dimensional embedding thereof (see Box 7.1 for more detail). This makes t-SNE foremost a

dimensionality reduction algorithm like the more well-known principal component analysis (PCA)[16] and multidimensional scaling (MDS)[17]. The main difference, conceptually, between t-SNE, PCA and MDS is that the former is non-linear whereas the latter two are linear algorithms. This means that in general PCA and MDS predominantly preserve the large distances (differences) between points far apart in the high dimensional space, because these obtain a larger weight in the optimization algorithm than small differences. The result is that global structure is preserved well, but small variations in local structure are poorly mapped to the low dimensional embedding. t-SNE, like other algorithms before, aims to overcome these shortcomings by utilizing a non-linear cost-function, which weighs both the local and global structures. The success and general applicability of the t-SNE algorithm in dimensionality reduction is shown by the spur of publications utilizing it to visualize complex high-dimensional data sets in diverse experimental settings.[18-22] Here, the applicability of t-SNE to drug discovery is investigated. The applicability as a similarity metric is studied, with the ultimate goal to apply it in a general workflow to predict target-ligand interactions naïvely based on the similarity dogma.

Box 7.1 | A brief introduction to the t-SNE algorithm

t-Distributed Stochastic Neighbour Embedding is a non-linear dimensionality reduction algorithm, or 'manifold learning' algorithm. This branch of (unsupervised) machine learning is based on the premise that many real-world high-dimensional datasets are intrinsically low-dimensional. To be interpretable by humans, a low dimensional view of data is required. The process of dimensionality reduction is therefore aimed at removing redundant dimensions and deflating it to two or three dimensions, primarily to be visually interpretable by the analyst.

Mathematical framework

The original high-dimensional data-space is defined as $\mathbb{X}^D$ and the mapped space $\mathbb{Y}^2$, where D is the dimensionality. Original data points are denoted as x_i, mapped points y_i. The similarity of data point x_j to datapoint x_i was defined by Maaten and Hinton as the conditional probability, $p_{j|i}$, that x_j would be picked as its neighbour by x_i. If neighbours are picked in proportion to their probability density under a Gaussian centred at x_i the mathematical definition for $p_{j|i}$ is given by:

$$p_{j|i} = \frac{\exp\left(-\frac{|x_i - x_j|^2}{2\sigma^2}\right)}{\sum_{k \neq l} \exp\left(-\frac{|x_k - x_l|^2}{2\sigma^2}\right)} \tag{7.1}$$

The *perplexity* variable, which is set by the user, is directly coupled to the σ in (7.1), and provides a measure of the smoothness of the embedding.

In the original SNE, a very similar definition was used for the probability condition for the embedding, $q_{j|i}$, where for convenience the σ_i is set to $\frac{1}{\sqrt{2}}$ such that:

$$q_{j|i} = \frac{\exp\left(-|y_i - y_j|^2\right)}{\sum_{k \neq l} \exp\left(-|y_k - y_l|^2\right)} \tag{7.2}$$

This definition led to various problems in the optimization procedure and subsequent implementations, so in t-SNE a Student t's distribution is used as the probability condition for the low-dimensional map. This leads to a definition of q_{ij} as follows:

$$q_{ij} = \frac{\left(1 + |y_i - y_j|^2\right)^{-1}}{\sum_{k \neq l} \left(1 + |y_k - y_l|^2\right)^{-1}} \tag{7.3}$$

Where p_{ij} is the symmetrized version of $p_{j|i}$ according to $p_{ij} = \frac{p_{j|i} + p_{i|j}}{2}$, and likewise for q_{ij}. In a perfect embedding, p_{ij} and q_{ij} will be equivalent for all possibilities of ij, so the difference between these two quantities needs to be minimized in the embedding process. To achieve this, the Kullback-Leibler divergence is employed by t-SNE as a measure of the faithfulness of the embedding. The function to minimize, the cost function, is then given by:

$$C = KL(P||Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}} \tag{7.4}$$

in which P and Q represent the joint conditional probability distribution summed over all points where $j \neq i$. The values of p_{ii} and q_{ii} are set to zero. The cost function is optimized by a gradient descent which is given by:

$$\frac{\partial C}{\partial y_i} = 4 \sum_j (p_{ij} - p_{ij})(y_i - y_j)\left(1 + |y_i - y_j|^2\right)^{-1} \tag{7.5}$$

Physical interpretation

The optimization generating the low-dimensional embedding can be interpreted as a mass-spring system reaching its equilibrium. The magnitude of the Kullback-Leibler divergence for a given j will scale linearly with $(y_i - y_j)$ (7.5), analogous to Hooke's law, which scales linearly with the displacement of points. Intuitively the stiffness of the spring connecting i and j is determined by their distance in high-dimensional space, and also follows for t-SNE, where p_{ij} is strongly influenced by $(x_i - x_j)$.

t-SNE maps molecular similarity of drugs in drug-like chemical space

Based on the principle that the chemical structure of a compound determines its biological and chemical properties, a machine learning algorithm that predicts target-ligand interaction landscapes should be able to recognize molecular similarity between different molecules. Traditionally, chemical similarity is measured by the Tanimoto coefficient (Tc).[23] To calculate the distance between compounds the Tc uses a molecular fingerprint, which is a high-dimensional bit vector that captures the presence or absence of chemical groups in a molecule. As a similarity metric the Tc has its limitations, predominantly because it averages differences over all bits, thereby losing information.[24] Thus, we envisioned that the data contained in the molecular fingerprint should be used more efficiently by a machine learning algorithm to determine molecular similarity. Moreover, the Tanimoto coefficient is principally a distance metric between two fingerprints and does not allow

visualization of similarity within a set of molecules, making interpretation of this similarity metric nontrivial. t-SNE can be readily applied to bit vectors of any length and as such is easily applicable to chemical structures represented by molecular fingerprints. The algorithm could thus be used to find and cluster the most similar molecules in a large dataset and visualize that similarity clustering in a two- or three-dimensional space.

The applicability of the t-SNE algorithm to this problem was initially tested by visualizing the molecular similarity of molecules from the Drug Repurposing Hub, an online repository containing compounds that have been clinically tested or used in humans.[25] Only the launched drugs (2774) were selected and partially attributed to 27 chemotypes by hand. Morgan fingerprints (RD-Kit, 4096-bits, radius = 2, equivalent to ECFP4[26]) were generated for each of these 2774 clinical compounds using KNIME, an open source software package.[27,28] The fingerprints were fed into the Python implementation of the Barnes-Hut t-SNE algorithm to generate a map of the drug-like chemical space.[29] The resulting map (Figure 7.1) shows remarkable co-localization of most of the chemotypes. As an example, the family of penicillin-like structures at the far right of the plot (cyan) is completely separated from all other chemical matter. Some unannotated molecules (in grey) are visible in this cluster, but upon detailed inspection they all constitute β-lactams in which the sulfur is either substituted or omitted. In addition, many other highly dense unannotated clusters are visible at the boundaries of the map, corresponding to highly defined chemotypes, such as the rapamycin, conazole and oxytocin analogues. Of note, even in the apparently less defined centre of the map, clear co-localization of similar molecules can be observed, for example a cluster of aspirin-like molecules (orange, near origin).

Quantifying the quality of the generated embedding is not trivial. Despite the broad interest in machine learning algorithms and their application in numerous fields, few objective quality measures for embeddings are available.[30] Often co-ranking histograms are constructed, comparing the original high-dimensional space (H-space) to the constructed low-dimensional space (here: t-SNE, PCA or MDS space). Briefly, co-ranking histograms are constructed by counting the occurrences of co-ranking for each sample against all other samples. As this generates a two dimensional plot, they are often visualized as a heat-map. For the current purposes, such co-ranking histograms were constructed for t-SNE, PCA and MDS (Euclidian distances calculated in 2D embedding) and that of the Tanimoto distance (Figure 7.2A-D). Generally, the co-ranking histogram of t-SNE fanned out somewhat quicker than the Tanimoto plots, but for the lower ranked neighbours it corresponded well to the H-space. Compared to the two linear dimensionality reduction algorithms, t-SNE looks to be considerably better based on the heat-maps. There are several ways to summarize the information contained in a co-ranking histogram, of which the Q-score is conceptually one of the easiest to interpret.[30] It assesses the overall quality of the dimensionality reduction where 1 is a perfect score, and 0 is the lowest. The Q-score neglects information regarding the kind of ranking error (too high or too low) but summarizes the quality up until a certain value of the rank, k. The higher rate of co-ranking for low ranks that was visible in the histograms is immediately evident from the Q-score graphs (Figure 7.2E and F), which start

Legend:
- 2-aminopyrimidine-containing
- Artemisinin-like
- Aspirin-like
- Atropine-like
- Calcidol-like
- Chromone-containing
- Cyclizine-like
- Ephedrine-like
- Ergolide-like
- Estradiol-like
- Floxacin-like
- Indole-containing
- Lovastatin-like
- Opioid-like
- Penicillin-like
- Profenamine-like
- Purine-containing
- Quinazoline-containing
- Ribouridin-like
- Riopine-like
- Rubicin-like
- Steroid-like
- Sulfanilamide-like
- Tetracyclin-like
- Thromycin-like
- Unsaturated-lipid-like
- Vindesine-like
- Unannotated

Figure 7.1 | t-SNE visualization of chemical space. t-SNE embedding of the 'launched' drugs in the Drug Repurposing Hub.[25] Embedding is based on the 4096-bit Morgan fingerprint. t-SNE settings: perplexity = 25, learning rate = 50, iterations = 10,000. Markers are coloured according to 27 manually attributed chemotypes.

out high but quickly fall upon increasing k. This behaviour is observed for all three dimensionality reductions and the Tanimoto distance. t-SNE however performs much better than the two linear algorithms, even for small values of k. When comparing t-SNE for the more direct neighbours in the histogram and plot (Figure 7.2A and E) it is clear that t-SNE follows the Tanimoto distance quite closely. Of note, this comparison is technically a bit skewed, as the Tc is only a different distance measure derived from the same H-space, not the result of a dimensionality reduction as is the case for t-SNE. A benefit of the three dimensionality reduction algorithms is that the 2D embedding allows for interrogation of the directionality of similarity, whereas when using the Tc, it is impossible to assess the similarity between two molecules when their Tc distance to a third is known. It can be concluded that t-SNE is able to map the local chemical space of approved drugs well, much better than linear dimensionality reduction approaches, whilst following a chemist's intuition. It recognizes molecular similarity in a broad set of diverse drug-like molecules and is able to present the results in a visually attractive and interpretable manner.
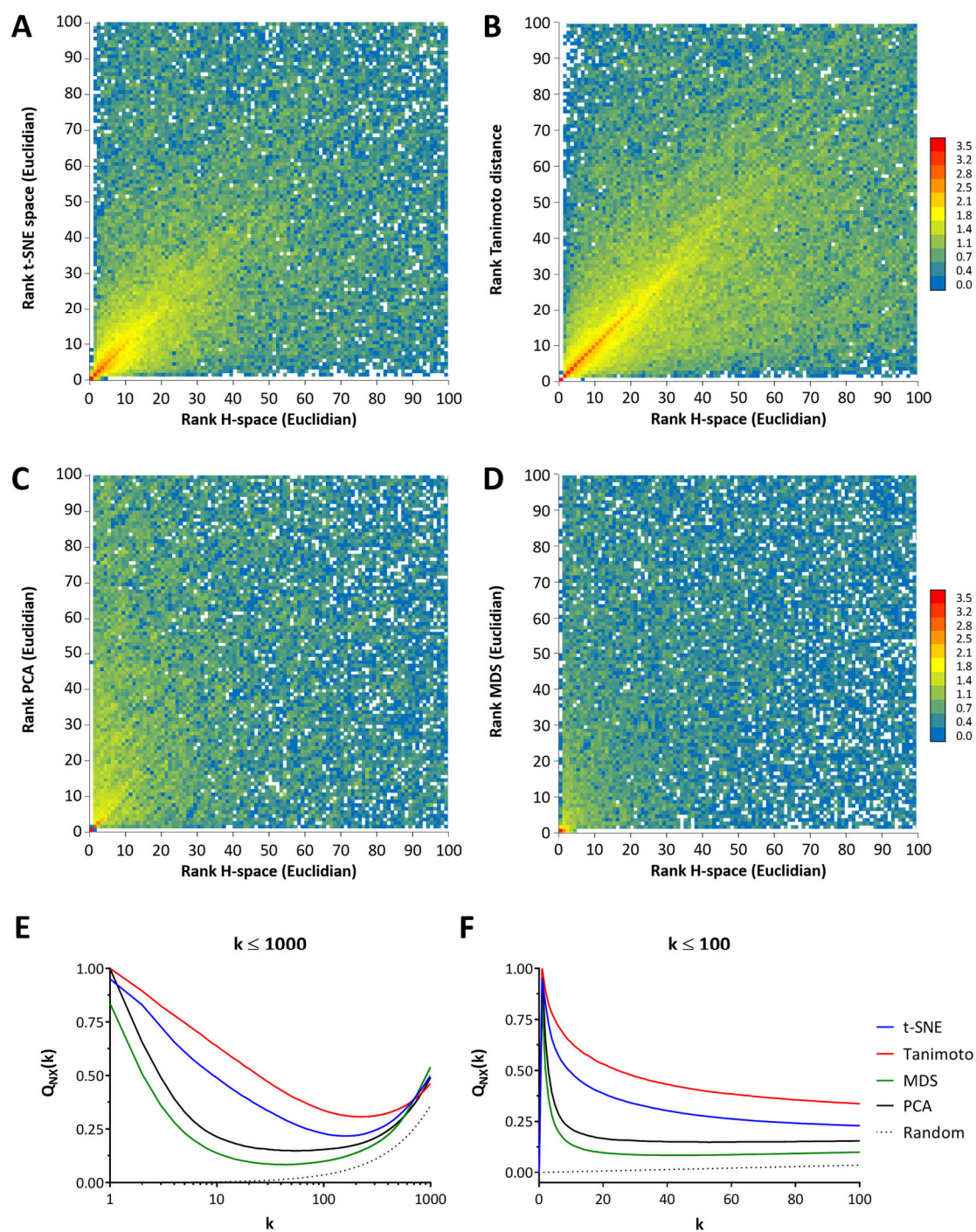
Figure 7.2 | Assessment of the quality of the dimensionality reduction as performed by t-SNE, PCA and MDS on the Drug Repurposing Hub molecules, based on the t-SNE embedding in Figure 7.1, the PCA plot in Supplementary Figure 7.1A and the MDS plot in Supplementary Figure 7.1B. Panels A-D show a $\log_{10}$-transformed co-ranking histogram for the top 100 ranked molecules for t-SNE (A), the Tanimoto coefficient (B), PCA (C) and MDS (D). The bottom panels show Q-score[30] graphs corresponding to the histograms above for the top 1000 (E) and top 100 (F) ranked molecules. H-space: the original high-dimensional space.

t-SNE map of serine hydrolases recapitulates phylogenetic information

Since closely related proteins typically bind similar endogenous molecules and (experimental) drugs, it was investigated whether the t-SNE algorithm is also capable of clustering proteins based on the similarity of their amino acids. Conceptually, this approach is analogous to proteochemometric modelling.[31] To quantify the similarity of serine hydrolases, first a manually curated list containing all serine endopeptidases and metabolic serine hydrolases (mSHs) was constructed. This was done in a similar manner as reported previously.[6] The amino acid sequences of the enzymes were then aligned using Clustal Omega.[32] This alignment was transformed to a fingerprint based on the physicochemical properties of the amino acids.[33] These fingerprints were used to create a two dimensional map of the target space by the t-SNE algorithm. The resulting map (Figure 7.3A) is capable of clustering most of the peptidases (triangles) and the mSHs (squares) in separate groups.

The embedding was compared with a more traditional phylogenetic tree (Figure 7.3B). As an illustration, several groups of proteins have been highlighted and annotated in the embedding. In purple all phospholipase A2 group IV proteins, except PLA2G4C, are clustered together. This separation is also observed in the phylogenetic tree. The small group of hepatocyte growth factor proteins (MST1, MST1L and HGF, in red) are related to the F2 protein, according to the phylogenetic tree, and contain kringle domains. In the t-SNE embedding however, F2 is quite far away, placed directly next to PRSS51 (not annotated), to which it is also quite similar. Intriguingly, the HGF related proteins are all expected to be catalytically inactive as they lack one or more of the typical catalytic triad residues, whereas F2 is confirmed to be a functional peptidase. This could indicate that t-SNE is able to separate similar proteins (global properties) on the absence or presence of key amino acids (local structure). It might also be one of the reasons why the red group is isolated in the plot. Other well defined protein clusters include the carboxylic esterases (CES, blue) and the products of genes *LIPA*, *LIPF* and *LIPM* (yellow). The centre of the embedding where less clusters are observed, still holds large co-localized families of proteins, e.g. the kallikreins (KLK, light green) or the acyl-coenzyme A thioesterases (ACOT, not annotated). It appears that t-SNE, when applied to binarized sequence alignments, is able to analyze and cluster these entities, maintaining the separation of large groups (peptidases and mSHs), whilst also keeping subclasses of proteins closely associated.

Figure 7.3 | Comparison of t-SNE embedding (A) and phylogeny tree (B) of the serine hydrolase (SH) family. t-SNE settings: perplexity = 30, learning rate = 50, iterations = 10,000.
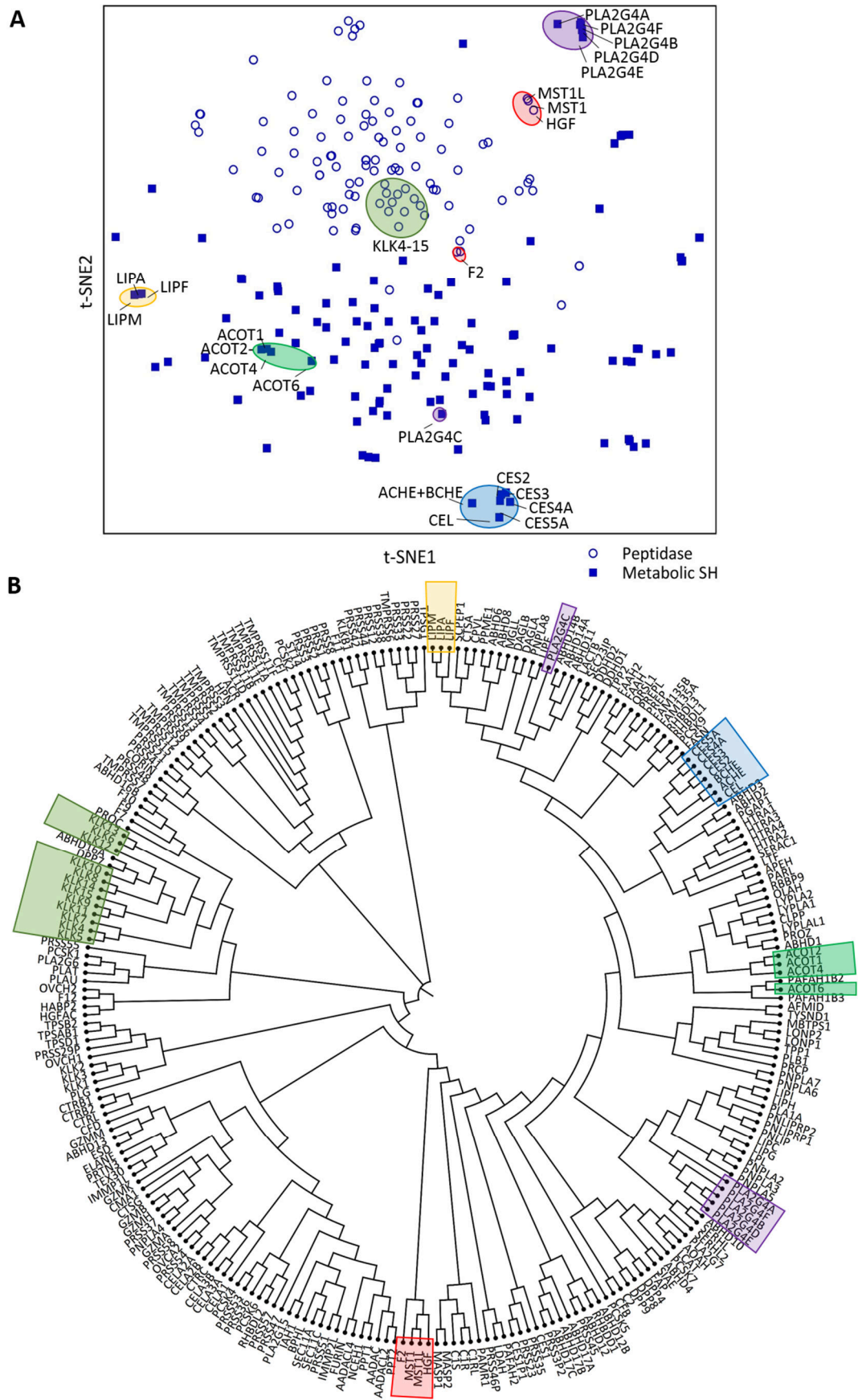
t-SNE analysis of serine hydrolase inhibition

t-SNE thus presents itself as a similarity visualization technique able to find and quantify similarities between molecules and proteins. With this technique a system could be envisioned where a newly designed molecule is compared to a large set of available molecules of which the bio-activity data is known. If one now finds the most similar known compounds utilizing t-SNE, the bio-activity data of those molecules could be used as a prediction for the bio-activity of the newly designed inhibitor. The t-SNE embedding of the protein space could then be used to append enzymes, for which no bio-activity data is available, as potential targets to the prediction based on the fact that they are closely related to known targets. Through this intuitive process, targets, or off-targets, of novel molecules may be predicted in an efficient way. This approach is comparable to PASS and various other techniques, reviewed by Bender *et al.*[34,35]

Currently, the largest dataset reported for serine hydrolases is that of Bachovchin *et al.*[6], which screens a library of 140 inhibitors against 72 SHs (10,080 data points). This dataset contains bio-activity at one concentration (50 µM) and all inhibitors have a carbamate warhead. The predictive utility of this dataset is, therefore, restricted to similar carbamate-based inhibitors. Another comprehensive serine hydrolase inhibitor dataset was generated using the EnPlex assay.[8] This dataset has bio-activity data for 55 quite diverse inhibitors (Figure 7.4) and 94 serine hydrolases (Figure 7.5). This dataset was processed using t-SNE to visualize the inhibition profiles projected on the embedding of molecules and proteins and thereby assess the co-localization.

When the 55 inhibitors were transformed in their respective fingerprints and clustered by t-SNE a quite disperse embedding was obtained (Figure 7.4A), in line with the diversity in the molecular set. Some more defined groups were observed, which upon inspection indeed corresponded to more similar molecules. When the bio-activity data was projected on this visualization by colouring the markers according to the pIC_{50}s for specific targets (Figure 7.4B-D) very distinct patterns could be observed. All DPP4 inhibitors were positioned in the far right of the embedding, with the exception of methoxyarachidonoyl fluorophosphonate (MAFP) which is in the centre of the plot and, by its reactive nature, inhibits virtually all serine hydrolases tested in this assay. For the endocannabinoid system related enzyme FAAH a similar observation can be made, with all inhibitors co-localized to the left of the embedding, barring some reactive molecules in the centre. Most of the tested FAAH-inhibitors are carbamate based, and this chemotype typically shows off-target activity against carboxylic esterases (CES). This is exemplified by Figure 7.4D where the potency for CES1 is annotated. Only the more chemically distinct urea based PF-3845 (topmost marker) is inactive for CES1.
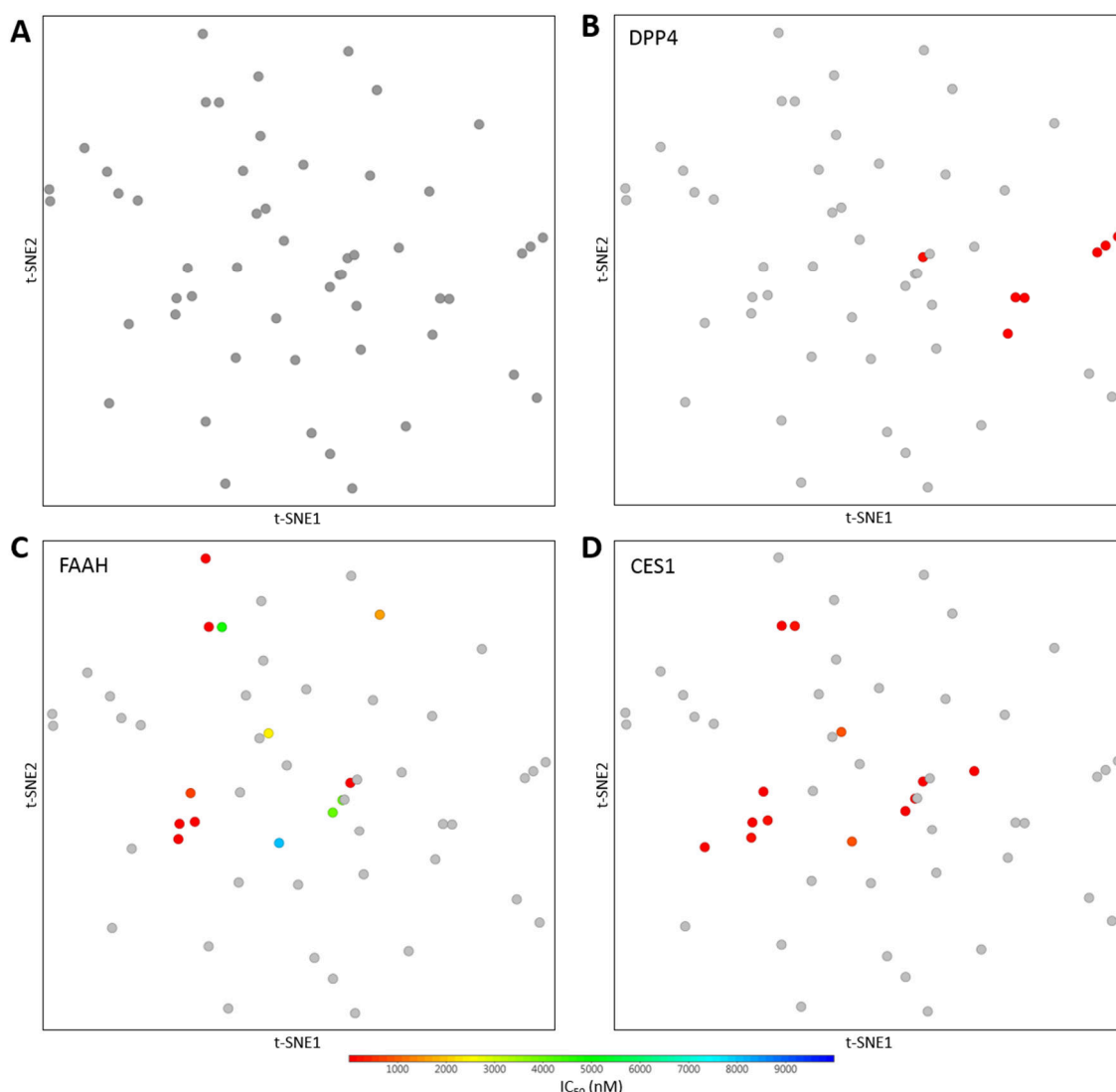
Figure 7.4 | t-SNE embeddings of the serine hydrolase inhibitors screened in the EnPlex assay.[8] A) Unannotated t-SNE clustering. B-D) All reported inhibition interactions of title enzymes are denoted in a gradient colouring. Only measured interactions < 10.000 nM are coloured in the embedding.

Next, the protein-target space was analysed. To aid in this analysis, the serine hydrolases for which no bio-activity data is available were filtered out (Figure 7.5A), and the markers were coloured according to the activity of three dissimilar inhibitors (Figure 7.5B-E). These inhibitors vary not only in structure, but also in reactive group (warhead), targeted subfamily, and selectivity. The localization of the inhibited SHs in the plots shows clear co-localization in pairs or groups. For example, JP104 (Figure 7.5B) has six measured IC_{50}s, which are arranged in the embedding as three duos. The active metabolite of lactacystin is, as might be expected from its small molecular structure containing a reactive β-lactone, less selective. Several clear groups of targets can be seen, most notably the seemingly unrelated cluster of retinoid inducible serine carboxypeptidase (SCPEP1), probable serine carboxypeptidase CPVL (CPVL), lysosomal protective protein (CTSA) and α/β-hydrolase domain containing protein 11 (ABHD11). Of note, the only enzyme classified as peptidase targeted by lactacystin is located close to several targeted mSHs. Finally, chymostatin is

quite selective, with only 2 observed interactions in this enzyme panel. The two peptidases targeted, chymase and chymotrypsin-C, are indeed closely related according to the t-SNE embedding, which is not directly evident from the phylogeny tree.
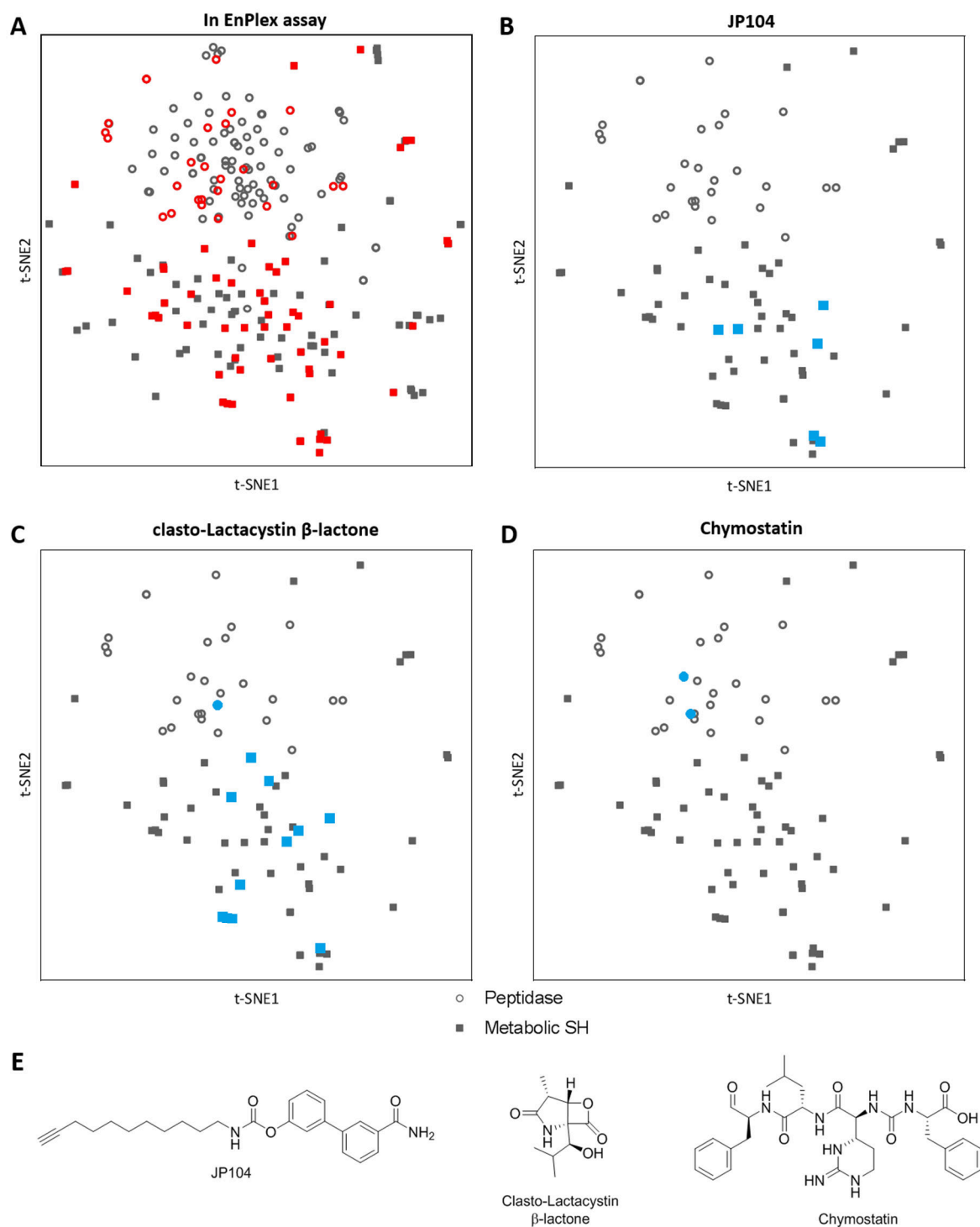
Figure 7.5 | t-SNE embeddings of the serine hydrolase superfamily. A) Red colouring denotes the inclusion in the EnPlex assay. B-D) All reported inhibition interactions of title inhibitor are denoted in blue. Only measured serine hydrolases are included in the embedding (those coloured red in A). E) Structures of inhibitors in B-D.

Conclusion

Taken together, it was shown that the dimensionality reduction algorithm t-Distributed Stochastic Neighbour Embedding can successfully be applied to molecular fingerprints, and that this generates a new, visually interpretable and quantitative similarity metric. Annotating a similarity visualization of a small inhibitor set with serine hydrolase inhibition data showed very distinct co-localization of activities in the embedding. t-SNE was also able to generate similarity maps for the serine hydrolase enzyme family when applied to a binarised multiple sequence alignment. This map revealed similarities between proteins not necessarily directly related in phylogeny. However, there is significant overlap between the traditionally assigned subfamilies and the co-localization in the t-SNE embeddings. By superimposing the available bio-activity data as a third dimension in colour on top of the t-SNE embeddings, definitive hotspots were shown where co-localized proteins were inhibited by the same inhibitor, which not always corresponded to the phylogenetic information. Given more (biological) data, it is expected that these mappings combined, in principle, could be applied to predict (off-)targets that are in these hotspots, or even predict (off-)targets for new molecules based on their molecular similarity to known compounds. It is expected that the target space visualization could be enhanced by improving the alignment of the enzymes. One option would be to centre the alignment around the catalytic serines, to improve the definition of the active site similarities, or truncate the sequences such that the alignment is more biased towards the actual binding site of the enzyme.

The concept presented here can easily be adapted to work with any dataset available. Because all data, algorithms, and data processing tools used are in the public domain or open source, it is highly adaptable and extensible. Concrete examples include different druggable protein classes, such as kinases (Chapter 8), G-protein coupled receptors, ion channels or nuclear hormones. It could also be trained on a different molecular set altogether, e.g. solubility, membrane permeability, metabolic stability, pharmacokinetics or toxicological data.

Methods

Cheminformatics tools

All molecular descriptors, molecular representations (SMILES, InChIKey) and fingerprints were generated using the RDKit software, either using the KNIME extensions or as the Python implementation.[28] Morgan fingerprints (4096-bits, radius 2) were used for all molecular fingerprints.

t-SNE algorithm

All t-SNE embeddings were generated with the Python Scikit-learn (v. 0.19) implementation of the Barnes-Hut t-SNE algorithm, either implemented in a 'Python for KNIME' node or as part of a Python script.[29]

Principal component analysis

Principal component analysis was performed using the Python Scikit-learn decomposition module (v. 0.19) based on the same 4096-bit fingerprints used for the other analyses.

Multidimensional scaling

The multidimensional scaling was performed using the Python Scikit-learn manifold module (v. 0.19), based on the 4096-bit fingerprints used for other analyses.

Co-ranking histogram and Q-scoring

The co-ranking histograms were constructed in Python, using SciPy's Rankdata functionality, using ordinal ranking.[36] Generated histograms were exported to Microsoft Excel 2016 where they were $\log_{10}$ transformed and heatmaps were generated.

The Q-score was calculated as described previously.[37] In brief, Q-scores were calculated by summarizing mild k intrusions and mild k extrusions for each k and normalizing to the sample size according to the following formula:

$$Q_{NX}(k) = \frac{1}{kN} \sum_{k,l} q_{kl}$$
(7.6)

where *N* is the total number of identities (2774, the number of molecules) and

$$q_{kl} = |\{(i,j) : \rho_{ij} = k,\ r_{ij} = l\}|$$
(7.7)

where ρ_{ij} denotes the rank in the low-dimensional manifold or the rank according to Tanimoto distance, and r_{ij} the rank in high dimensional space.

Serine hydrolase sequence information and bitstring

Sequence information of the serine hydrolases was retrieved from Uniprot.[38] The proteins were aligned using the online Clustal Omega tool provided by the EMBL-EBI using the default settings.[32] The standard "Clustal w/o numbers" output generated was transformed to a bitstring using the amino acid fingerprints as provided in Heil *et al.*[33] with the following additions: alignment dashes (-), stops (*) and blanks (X) were all considered empty, and were represented by 23 0's.

Bioactivity used

Activity data for the serine hydrolase superfamily was used without adaptations from Bachovchin *et al.*[8].

Supplementary Figures
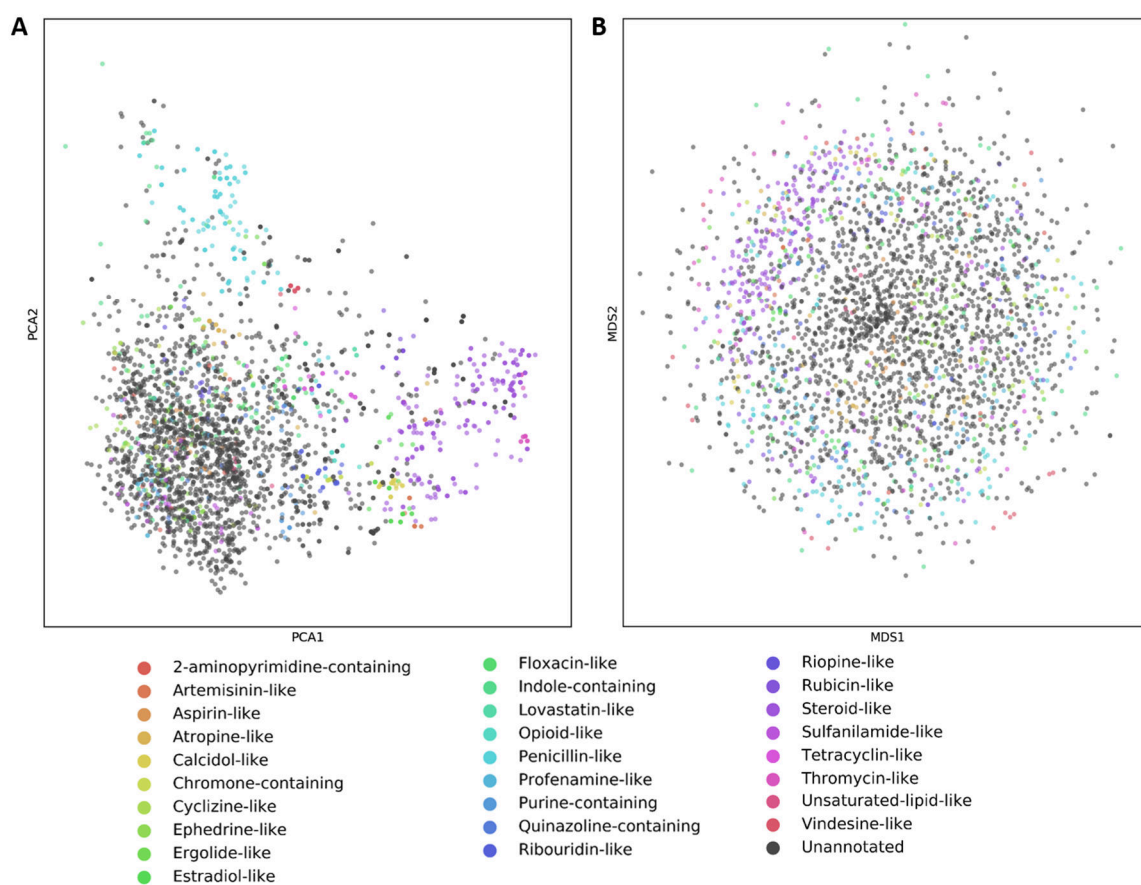
Figure S7.1 | Principal component analysis (A) and multidimensional scaling analysis (B) of the drug repurposing hub molecules with the status "Launched". Embeddings are based on the 4096-bit Morgan fingerprint (RD-Kit). Markers are coloured according to 27 manually attributed chemotypes.

References

1. Reymond, J.-L. The Chemical Space Project. *Acc. Chem. Res.* **48**, 722–730 (2015).
2. Kirkpatrick, P. & Ellis, C. Chemical space. *Nature* **432**, 823–823 (2004).
3. Imming, P. Medicinal Chemistry: Definitions and Objectives, Drug Activity Phases, Drug Classification Systems. in *The Practice of Medicinal Chemistry* (ed. Wermuth, C. G.) 3–13 (2011).
4. Bender, A. & Glen, R. C. Molecular similarity: a key technique in molecular informatics. *Org. Biomol. Chem.* **2**, 3204 (2004).
5. 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).
6. Bachovchin, D. A. *et al.* Superfamily-wide portrait of serine hydrolase inhibition achieved by library-versus-library screening. *Proc. Natl. Acad. Sci.* **107**, 20941–20946 (2010).
7. 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).
8. Bachovchin, D. A. *et al.* A high-throughput, multiplexed assay for superfamily-wide profiling of enzyme activity. *Nat. Chem. Biol.* **10**, 656–663 (2014).
9. Christmann-Franck, S. *et al.* Unprecedently Large-Scale Kinase Inhibitor Set Enabling the Accurate Prediction of Compound–Kinase Activities: A Way toward Selective Promiscuity by Design? *J. Chem. Inf. Model.* **56**, 1654–1675 (2016).
10. Sorgenfrei, F. A., Fulle, S. & Merget, B. Kinome-Wide Profiling Prediction of Small Molecules. *ChemMedChem* **13**, 495–499 (2018).
11. Merget, B., Turk, S., Eid, S., Rippmann, F. & Fulle, S. Profiling Prediction of Kinase Inhibitors: Toward the Virtual Assay. *J. Med. Chem.* **60**, 474–485 (2017).
12. Cichonska, A. *et al.* Computational-experimental approach to drug-target interaction mapping: A case study on kinase inhibitors. *PLOS Comput. Biol.* **13**, e1005678 (2017).
13. Hansch, C. & Fujita, T. ρ-σ-π Analysis. A Method for the Correlation of Biological Activity and Chemical Structure. *J. Am. Chem. Soc.* **86**, 1616–1626 (1964).
14. Van Der Maaten, L. & Hinton, G. Visualizing Data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
15. Hinton, G. & Roweis, S. Stochastic Neighbor Embedding. in *Proceedings of the 15th International Conference on Neural Information Processing Systems* 857–864 (MIT Press, 2002).
16. Hotelling, H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **24**, 417–441 (1933).
17. Torgerson, W. S. Multidimensional scaling: I. Theory and method. *Psychometrika* **17**, 401–419 (1952).
18. Mahfouz, A. *et al.* Visualizing the spatial gene expression organization in the brain through non-linear similarity embeddings. *Methods* **73**, 79–89 (2015).
19. Amir, E. D. *et al.* viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. *Nat. Biotechnol.* **31**, 545–52 (2013).
20. Abdelmoula, W. M. *et al.* Data-driven identification of prognostic tumor subpopulations using spatially mapped t-SNE of mass spectrometry imaging data. *Proc. Natl. Acad. Sci.* **113**, 12244–12249 (2016).
21. Bushati, N., Smith, J., Briscoe, J. & Watkins, C. An intuitive graphical visualization technique for the interrogation of transcriptome data. *Nucleic Acids Res.* **39**, 7380–7389 (2011).
22. Taskesen, E. & Reinders, M. J. T. 2D Representation of Transcriptomes by t-SNE Exposes Relatedness between Human Tissues. *PLoS One* **11**, e0149853 (2016).
23. Willett, P. & Winterman, V. A Comparison of Some Measures for the Determination of Inter-Molecular Structural Similarity Measures of Inter-Molecular Structural Similarity. *Quant. Struct. Relationships* **5**, 18–25 (1986).
24. Baldi, P. & Nasr, R. When is Chemical Similarity Significant? The Statistical Distribution of Chemical Similarity Scores and Its Extreme Values. *J. Chem. Inf. Model.* **50**, 1205–1222 (2010).
25. Corsello, S. M. *et al.* The Drug Repurposing Hub: a next-generation drug library and information resource. *Nat. Med.* **23**, 405–408 (2017).
26. Rogers, D. & Hahn, M. Extended-Connectivity Fingerprints. *J. Chem. Inf. Model.* **50**, 742–754 (2010).
27. Berthold, M. R. *et al.* KNIME: The Konstanz Information Miner. in *Data Analysis, Machine Learning and Applications* (ed. Preisach C., Burkhardt H., Schmidt-Thieme L., D. R.) 319–326 (Springer, Berlin, Heidelberg, 2008).
28. Landrum, G. RDKit: Open-source cheminformatics; http://www.rdkit.org.
29. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
30. Lee, J. A. & Verleysen, M. Scale-independent quality criteria for dimensionality reduction. *Pattern Recognit. Lett.* **31**, 2248–2257 (2010).
31. van Westen, G. J. P., Wegner, J. K., IJzerman, A. P., van Vlijmen, H. W. T. & Bender, A. Proteochemometric modeling as a tool to design selective compounds and for extrapolating to novel targets. *Med. Chem. Commun.* **2**, 16–30 (2011).

32. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
33. Heil, B., Ludwig, J., Lichtenberg-Frate, H. & Lengauer, T. Computational recognition of potassium channel sequences. *Bioinformatics* **22**, 1562–1568 (2006).
34. Lagunin, A., Stepanchikova, A., Filimonov, D. & Poroikov, V. PASS: prediction of activity spectra for biologically active substances. *Bioinformatics* **16**, 747–8 (2000).
35. Bender, A. *et al.* Chemogenomic data analysis: prediction of small-molecule targets and the advent of biological fingerprint. *Comb. Chem. High Throughput Screen.* **10**, 719–31 (2007).
36. Jones, E., Oliphant, T., Peterson, P. & others. SciPy: Open source scientific tools for Python.
37. Lee, J. A. & Verleysen, M. *Quality assessment of nonlinear dimensionality reduction based on K-ary neighborhoods.* **4**,
38. Consortium, T. U. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **45**, D158–D169 (2017).

Drug Discovery Maps visualizes and predicts kinome-inhibitor interaction landscapes

Introduction

Protein kinases are an important class of drug targets due to their key role in intracellular signal transduction processes involved in cancer, auto-immune diseases, and (neuro)inflammation.[1,2] The therapeutic value of the protein kinase family is demonstrated by the 38 kinase inhibitors currently approved by the FDA and the plethora of molecules being tested in clinical trials for this enzyme family.[3] It is anticipated that these clinically approved kinase inhibitors (KI) can serve as starting points to identify novel drug candidates for other kinases.

Most KIs interact with the structurally and functionally conserved ATP-binding site, present in all >500 human protein kinases. It is well established that KIs bind multiple members of the kinase family, and that this may affect their efficacy and toxicity.[4] Detailed investigation of the target-interaction landscape of KIs is, therefore, important to understand their molecular mode of action and offers the opportunity to identify new starting points for other therapeutically interesting kinases. Many complex, high-dimensional datasets with structure-activity relationships (SAR) of KIs over a broad selection of kinases have become available (Table 8.1).[5-10] The wealth of data present in literature has spurred the development of several computational platforms to predict kinase activities of inhibitors.[11,12] These reported highly specialised computational models which were able to predict, to some extent, the inhibition profiles of inhibitors against the kinome.

Given the large amount of data available and the apparent success of previous attempts to use these datasets in a predictive manner it was expected that the strategies introduced in Chapter 7 could be applied to the field of kinase inhibitors. This approach was dubbed Drug Discovery Maps (DDM), a machine learning tool that allows the visualization and prediction of target-ligand interaction landscapes.

Table 8.1 | Summary of large experimental (partially) public kinome screens for sets of kinase inhibitors. Numbers of inhibitors and kinases are as reported in original publications and include mutants, splicing variants and separate kinase domains from the same kinase. Coverage represents the percentage of possible kinase-inhibitor pairs for which a data point is reported.

Set	No. of inhibitors	No. of kinases	Dosage	No. of data points	Coverage
Karaman *et al.*	38	317	Dose response**	12046	100%
Metz *et al.*	3858	172	Dose response	258094	38.9%
Anastassiadis *et al.*	178	300	Single dose	52834	98.9%
Davis *et al.*	72	442	Dose response**	31824	100%
Elkins *et al.*	367*	224	Two doses	81940	99.9%

* contains duplicate molecules
** only if activity is observed at initial 10 µM single dose screen

t-SNE reliably reproduces human chemotype assignments

In Chapter 7 it was demonstrated that t-SNE is able to visualize molecular similarity in a drug-like set of 2774 clinically applied molecules. For the current purposes, it had to be verified that t-SNE is still able to recognize molecular similarity within a smaller set of drug-like molecules that is more homogenous and has higher molecular similarity. To this end, a t-SNE-mediated clustering on the molecules from the Published Kinase Inhibitor Set (PKIS) was performed.[13] The PKIS is a 364-member library of molecules assembled by GlaxoSmithKline that are all classified as inhibitors of protein kinases. The PKIS set represents 31 manually annotated chemotypes and their activity has been measured and published for 200 kinases (Table 8.1, last entry).[9] The resulting map of chemical space representing the KIs (Figure 8.1A) was coloured according to the manual chemotype attribution and showed clear co-localisation of chemotypes. A more in-depth analysis by the unsupervised cluster assignment algorithm DBSCAN confirmed the initial visual inspection

and shows high statistical correlation between the autonomously derived clustering and the human annotation.[14] Of the 31 chemotypes annotated, 23 were fully comprised in one computationally assigned cluster. For example, the orange and gold clusters on the left of the map are completely isolated and comprise all compounds of those chemotypes (Figure 8.1B). An enlargement of the cluster dominated by 3-amino-pyrazolopyridines (Figure 8.1B) showed that in fact three different chemotypes were included. Compounds **1-3** are examples of each, and they showed that indeed these compounds are highly similar, and perhaps arbitrarily assigned to different chemotypes. The 'stray' sky blue compound **4** is part of a chemotype that is less well defined in t-SNE space, but the similarities with compounds **1-3** are obvious. This analysis showed that t-SNE is capable to recognize and cluster molecular entities in a highly specific manner and that it allows the visual inspection of high-dimensional chemical structural data, or chemical space, in an easy and intuitive way.
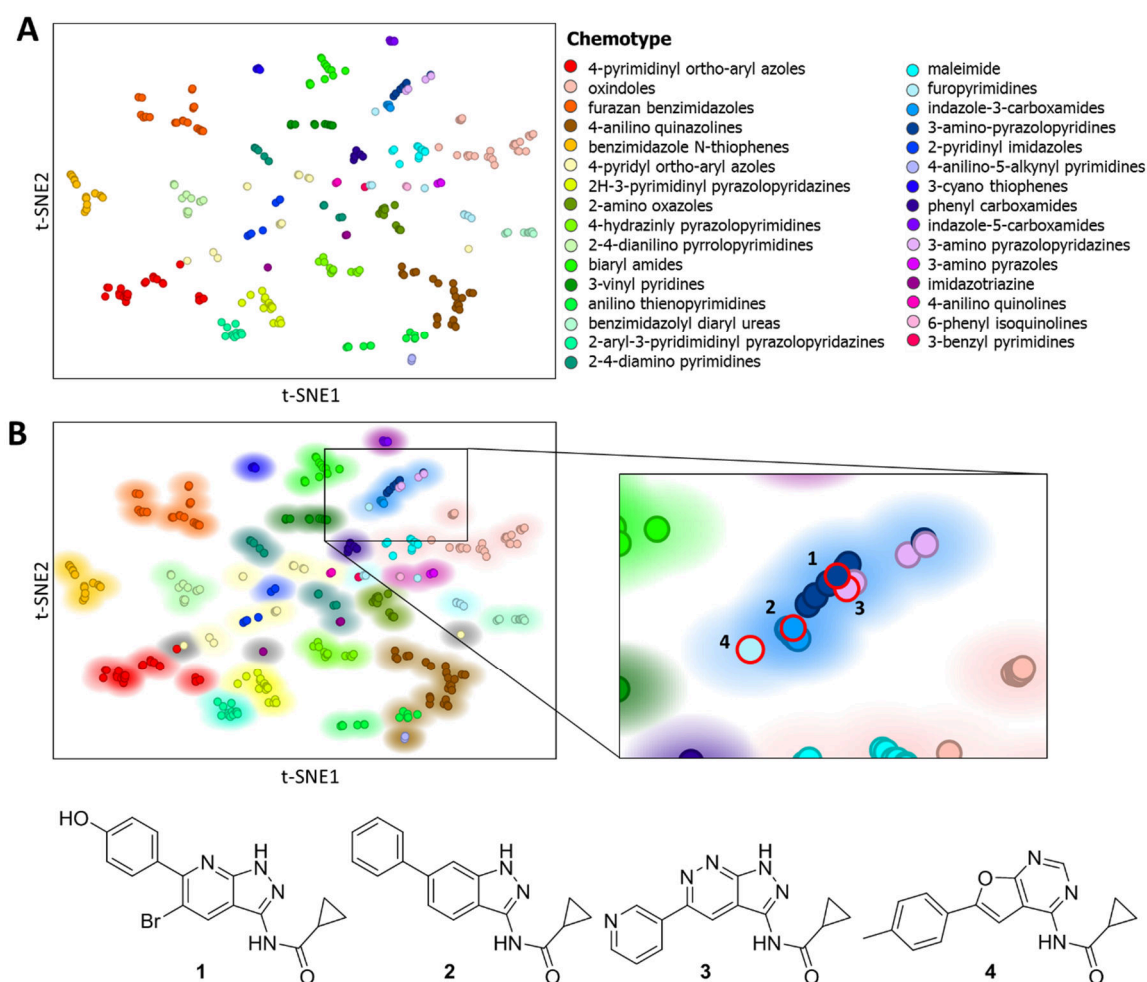
Figure 8.1 | A) t-SNE embedding of the Published Kinase Inhibitor Set. Embedding is based on the 4096-bit Morgan fingerprint. t-SNE settings: perplexity = 50, learning rate = 50, iterations = 10,000. Markers are coloured according to 31 manually attributed chemotypes. B) t-SNE embedding as in A) but including a background of the DBSCAN generated clustering, coloured by the dominant chemotype in that cluster (grey are singletons or duos). Markers are coloured according to the 31 chemotypes defined by Elkins *et al.* as in A).

t-SNE map of the kinase target space

To generate a visualization of the target space, and to quantify the similarity of kinases, first an alignment of the amino acid sequences of the whole kinase domains containing the ATP-binding pocket was created. This was translated to a fingerprint based on physico-chemical properties of the amino acids, analogous to the methods used in Chapter 7.[15] These fingerprints were used to create a two-dimensional map of the target space by the t-SNE algorithm. The resulting embedding (Figure 8.2A) is remarkable, as it almost seamlessly recreates the phylogenetic tree published by Manning *et al*. in 2002.[16] To assign the kinases to clusters, the coordinates of the t-SNE embedding were fed into the DBSCAN algorithm. All ten assigned clusters were significantly (P < 0.0001, Hypergeometric test) enriched for a specific kinase group as assigned by Manning *et al*. (Figure 8.2A). Closer inspection of some of the kinases unassigned by DBSCAN reveals that they belong to distinct branches of the phylogenetic tree, corresponding with their separation from the main clusters. As an example, the four tyrosine kinases (TK) in the far right of the embedding (burgundy) all belong to the JAK family (JAK1, 2, 3 and Tyk2), but only represent their second kinase domain. The first kinase domain is more closely associated with the rest of the TK group, and lies just outside the DBSCAN assigned cluster. The close association of the second kinase domains with the RGC cluster (coloured brown) is especially striking, as these domains, just as the RGC kinases, are considered pseudokinases. The same holds true for MLKL, IRAK2 and IRAK3. Intriguingly, the IRAK family of TKL kinases has four members, of which IRAK1 and IRAK4 are catalytically active, whereas IRAK2 and IRAK3 are not.[17] In the t-SNE embedding, the former are located in the major TKL cluster (orange) whereas the latter are actually assigned to the RGC-dominated cluster. MLKL has also been shown to indeed lack catalytic activity in at least one report.[18]

Another interesting feature is the separation of a group (left of plot) of TKL kinases from the major cluster. This subset features all but one of the STKR family of cell surface bound receptor kinases. Upon closer inspection, even the subfamilies of STRK1 and -2 are discernible. Strikingly, the MISR2 (AMHR2) kinase receptor is located with kinases categorised as 'Other'. This receptor kinase has an atypical DFG motif (DLG), and as such can indeed be classified as a pseudokinase, although phosphorylating activity has experimentally been shown.[19] The other members of the STKR family do all share the conserved DFG motif. Finally, on the lower side of the t-SNE plot several AGC coloured kinases have been clustered with the CAMK kinases. These actually represent the second kinase domains of the RSK family, which were also attributed to the CAMK group by Manning *et al*.[16]

In summary, this analysis of target space of the binding site of protein kinase domains indicated that this embedding is able to recognize overall similarity but also detect subtle differences between the different binding domains of most kinase inhibitors.
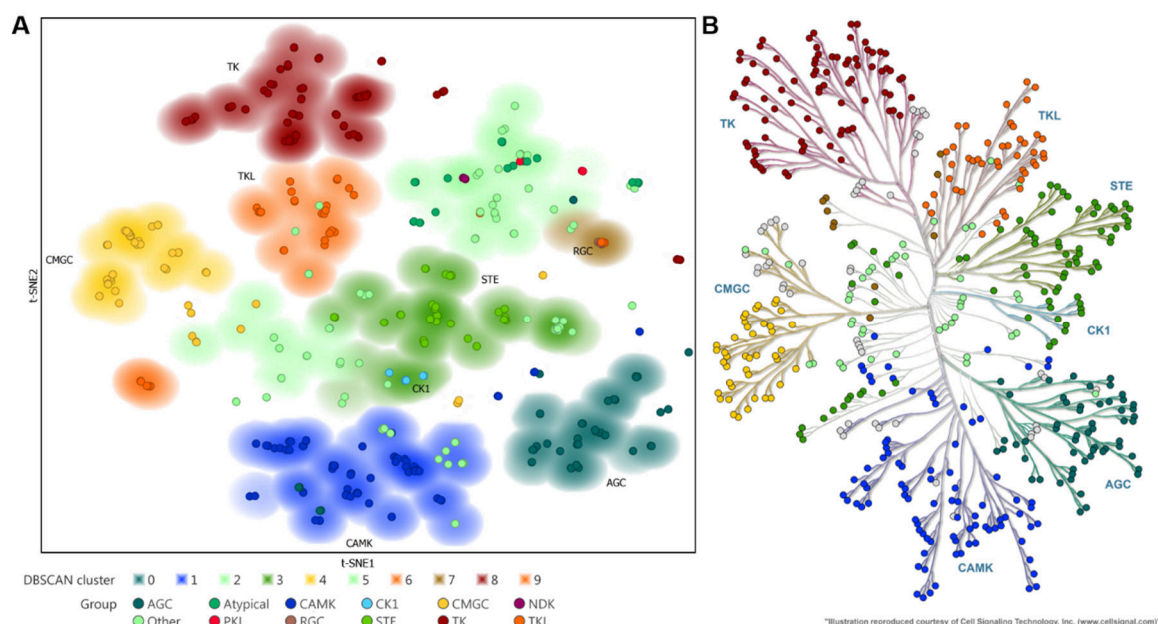
Figure 8.2 | t-SNE visualization of kinase domains reveals phylogenetic information: A) t-SNE embedding of physicochemical fingerprint of the kinase domains of 535 human kinase domains. t-SNE settings: perplexity = 50, learning rate = 50, iterations = 25,000. Arbitrary t-SNE coordinates are rotated to match dendrogram orientation of Manning *et al*. Markers are coloured according to the 12 groups defined by Manning *et al*., background is coloured based on the DBSCAN generated clustering, coloured by the dominant kinase group in that cluster (blanks are unclustered kinases). B) Manning *et al*. manually curated kinome dendrogram, overlaid with circles coloured according to the background colouring from the t-SNE map in A), based on the unsupervised DBSCAN clustering.[20]

DDM can predict target-ligand interaction landscapes

Based on the successful analysis of the chemical and target space of kinases and their inhibitors, a workflow was set up to predict the activity of novel compounds for the entire kinome. The bioactivity data measured by Elkins *et al*. for the Published Kinase Inhibitor Set was used as training set, as it contains the most unique interactions of all open datasets (Table 8.1).[9] The optimisation of the workflow with all parameters is described in more detail in the Methods. The final architecture of the algorithm is depicted in Figure 8.3, illustrated for the EGFR inhibitor erlotinib. At first, a t-SNE embedding is generated where erlotinib is mapped onto the chemical space of the PKIS (top left). This information is used to find the 9 most similar molecules (top right). Of these, the inhibition data measured by Elkins *et al*. is averaged, and all kinases above a threshold value *C* are considered targets (lower right). These kinases are then looked up in the target space map (Figure 8.2A), and the most similar kinases are appended (lower left) to yield the final prediction (centre). As the molecular t-SNE embedding is slightly stochastic, the described process is repeated several times (*R*) and the number of times a kinase is predicted is tracked.
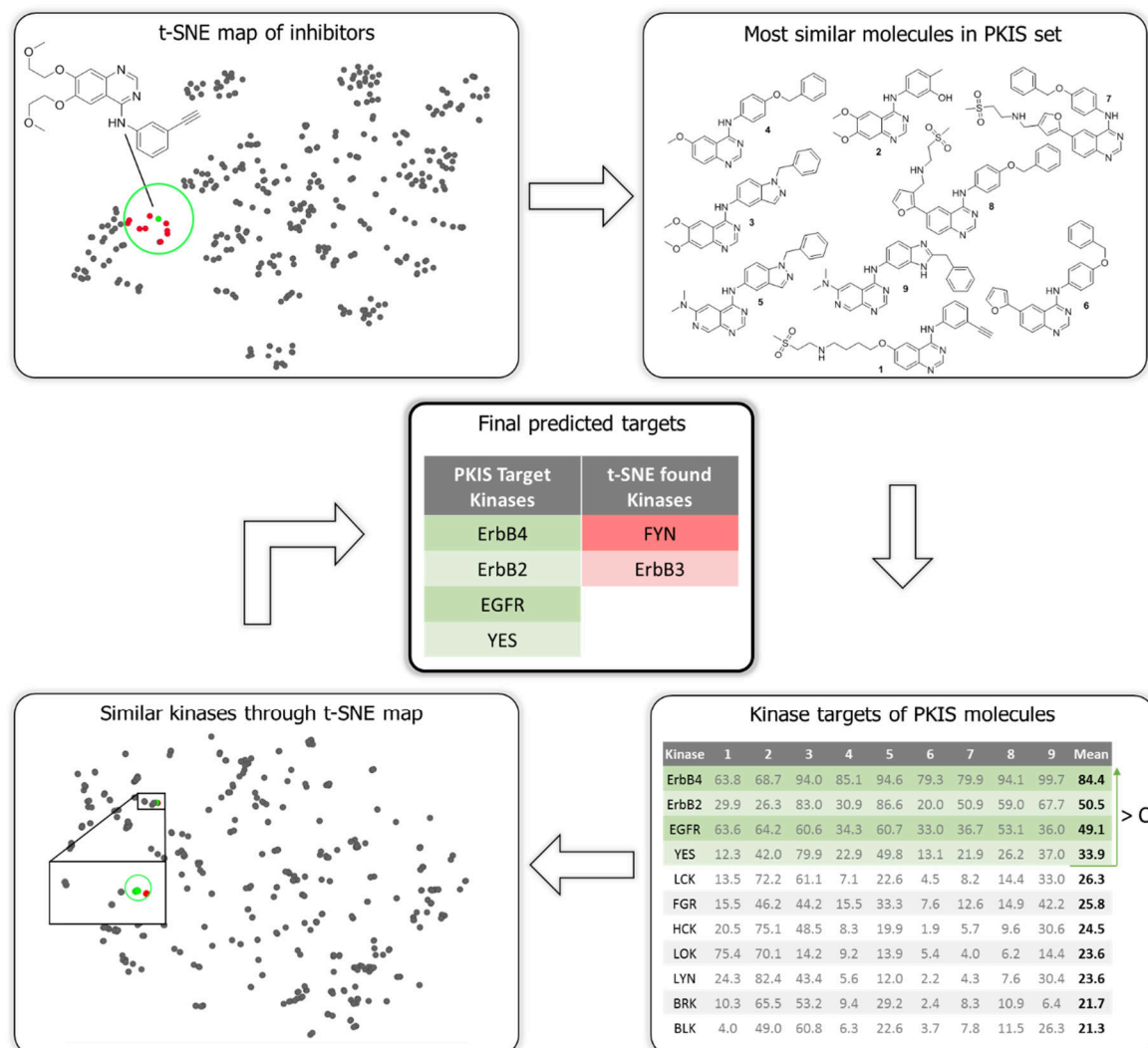
Figure 8.3 | Schematic overview of the DDM workflow. In this example, the targets of erlotinib are predicted. Based on a t-SNE embedding (top left) the PKIS inhibitors nearest to erlotinib are found, depicted in the top right panel. For these, the inhibition data as measured by Elkins *et al.* is averaged and used as initial prediction. These targeted kinases are then looked up in the t-SNE embedding in the lower left panel, where the most similar kinases are added to yield the final prediction (centre).

The DDM model was validated using an independent dataset generated by Karaman *et al.*[5] The resulting prediction statistics for each of the 38 compounds in this test set are summarised in Table S8.1. The average positive prediction value (PPV) was 40%, with a Matthews correlation coefficient (MCC) of 0.21. These statistics were compared with previously published methods, and it was found that DDM performed better than QSAR-models and was equal in performance to random forest based proteochemometric models (Figure 8.4). A receiver operating characteristic (ROC) analysis of the performance of DDM on this test set shows an area under the curve (AUC) of 0.76 (Figure 8.5). Taken all together, a novel machine learning model to predict kinome-inhibitor landscapes was developed and validated.
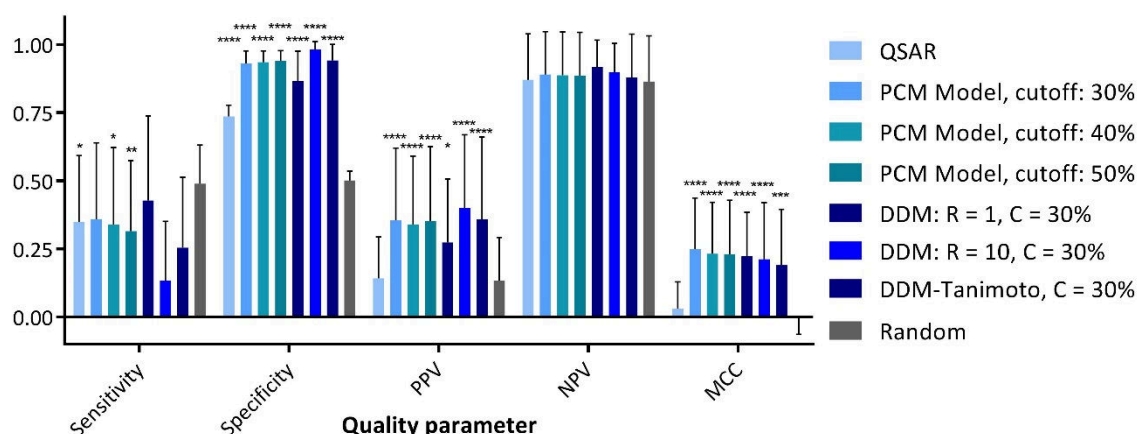
Figure 8.4 | Statistical comparison of the new t-SNE based model with four models based on published procedures. Bars denote mean ± SD of the 38 inhibitors. Significant improvements over random have been highlighted, * $p < 0.0332$, ** $p < 0.0021$, *** $p < 0.0002$, **** $p < 0.0001$ (Two-way ANOVA, Tukey's multiple comparison test). The relatively low PPV and high NPV for the random predictions are caused by the activity bias in the data set (13% active vs 87% inactive).
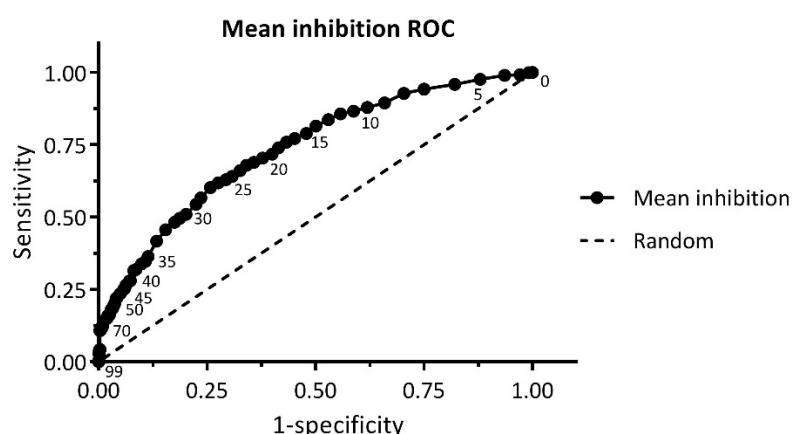
Figure 8.5 | Receiver operating characteristic curve of the test set predictions with varying cut-off C values. Marker labels represent *C* values. The area under the curve is 0.76, with the maximum deviation from random at C values between 25 and 35%. Axes have been normalized to take into account the inability to predict all kinases based on the PKIS coverage.

Discovery of novel FLT3 inhibitors using DDM

To investigate the utility of the model in early drug development, it was applied to the identification of new inhibitors for FMS-like Tyrosine kinase 3 (FLT3). FLT3 is implicated in advanced myeloid leukaemia, where approximately 30% of patients carry an internal tandem duplication (ITD) in their FLT3 gene that activates the kinase and acts as a driver mutation.[21] Recently, midostaurin has been FDA approved for the treatment of AML patients and several other inhibitors are currently being tested in clinical trials. However,

fast adaptive mutations in the FLT3 gene quickly result in drug-induced resistance of AML, warranting the search for novel chemotypes to inhibit this kinase. To this end, the DDM model was used to predict the kinome-ligand interaction landscape of a small kinase-focused library of 1152 molecules. They were analyzed using various values for the activity cut-off C, and were ultimately filtered using $C = 40\%$, and a prediction count of at least 9 out of 10 runs, to have a balanced number of molecules to be tested. These stringent cut-offs yielded a workable set of 44 compounds predicted to be active against FLT3.

To validate our virtual DDM screen, a time-resolved fluorescence-resonance energy transfer (FRET)-based biochemical assay was performed with all 1152 compounds against FLT3, at an initial concentration of 10 µM. This screen yielded 184 actives, with >50% loss of activity (16% of all compounds). Of these compounds, the pIC_{50} was measured, resulting in 135 compounds with a $pIC_{50} > 5$, with a mean of 6.7 ± 0.9. 18 of the 184 compounds were part of the 44 compounds identified by the DDM screen, which results in a positive predicted value (PPV; or hit rate) of 41% (Figure 8.6A, $P < 0.0001$ (Hypergeometric test)), which is almost three-fold higher compared to the hit rate of the biochemical assay. Of note, 15 of the predicted compounds demonstrated an IC_{50} of less than 2 µM (34%, $P < 0.0001$ (Hypergeometric test)) with an average pIC_{50} 7.3 ± 1.1, and also included the most active compound found in the screen (crenolanib, pIC_{50} of 9.0). The hit rate was nearly identical to the validation statistics for the test set (Figure 8.4), where an overall PPV of 40% was achieved. The same holds for the negative predictive value (89%) and the sensitivity (11%). The successful application of the DDM model for the FLT3 screen may partially be attributed to the high coverage for the TK family of kinases. Of note, the relatively low sensitivity (11%) is a balanced choice between minimising the number of compounds to screen and finding more actual hits. This can easily be tuned by varying the cut-off parameter.

Two of the predicted compounds, **5** and **6** (Figure 8.6B), were selected based on their chemical properties, novelty regarding FLT3 inhibition, and their predicted interaction profile (*vide infra*). These compounds were resynthesized using established methods (see Methods). The activity of the compounds was confirmed in a FRET assay using recombinant human FLT3 (Figure 8.6C). Compounds **5** and **6** showed a concentration-dependent activity with pIC_{50} values of 7.3 ± 0.1 and 8.8 ± 0.1, respectively. To determine the cellular activity of both compounds, a cell proliferation assay using the FLT3-dependent AML cell line MV4:11 was performed. Both compound **5** and **6** showed clear cellular activity with pEC_{50} values of 6.3 ± 0.1 and 8.5 ± 0.1, respectively. In summary, the experimental validation of the hits illustrated the power of the DDM workflow for compound selection in the lab.

Finally, to explain the potential binding mode of compound **5** and **6** both compounds were docked using a DFG-in model (for **5**) and the DFG-out structure 4RT7 (for **6**) (Figure 8.6E). Compound **5** binds to the hinge region with the aminopyrimidine moiety in a fashion typical for Type 1 kinase inhibitors. Compound **6** binds FLT3 in the DFG-out conformation analogous to RIPK2 (5AR7) by forming hydrogen bonds to the DFG-motif using the urea functionality and to the hinge region with the pyridine nitrogen.[22]
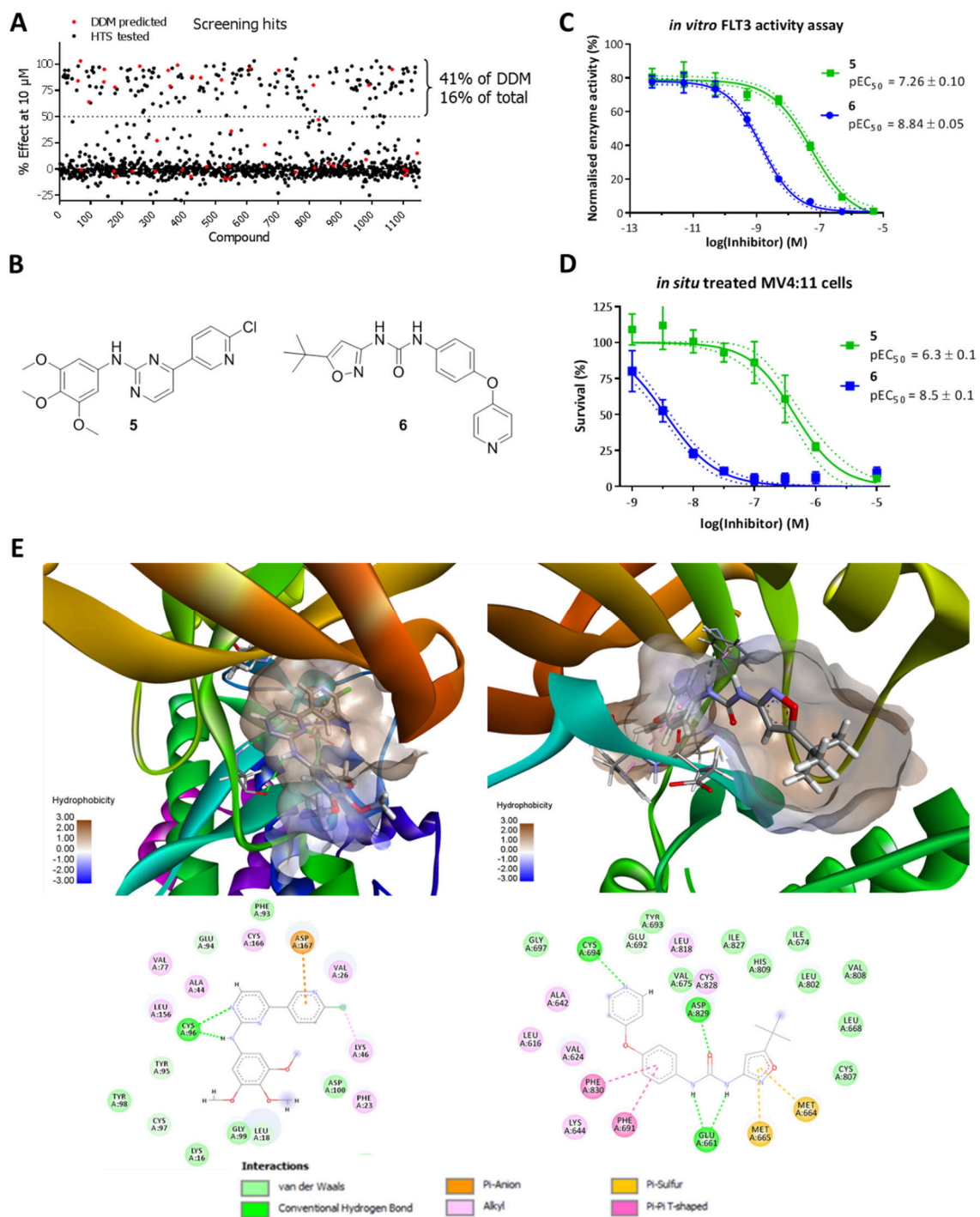
Figure 8.6 | Discovery of novel FLT3 inhibitors using DDM. A) Scatter plot of all compounds and their inhibitory effect at 10 µM as measured in the high-througput screen. DDM predicted molecules are marked red. B) Structures of the two compounds re-synthesized and tested *in situ* against MV4:11 cells. C) Dose response curves of compounds **5** and **6** against recombinant FLT3 in a FRET-based activity assay. Markers denote mean ± SD (N=4). Dotted lines denote 95% confidence interval of EC_{50} fit. D) Dose response curves of compounds **5** and **6** against MV4:11 leukaemia cells. Markers denote mean ± SD (N=3). Dotted lines denote 95% confidence interval of EC_{50} fit. E) Docking poses of **5** and **6** in the 3D models of FLT3 and the corresponding 2D interaction plots.

Kinome activity spectrum prediction using DDM

To reduce potential toxic side effects, kinase cross reactivity is ideally minimized. DDM enables rapid assessment of the predicted cross reactivity, because by default, DDM predicts the interactions with the entire kinome. Thus far, however, only the FLT3 prediction has been taken into account. As final validation, the activity of the two inhibitors on the predicted off-targets was tested in biochemical assays. In addition to FLT3, compounds **5** and **6** were predicted to be active against 35 and 33 kinases, respectively (C = 40%, R > 0.5). The off-targets were validated using KinaseProfiler™ by Eurofins at 10 µM. The inhibition data per compound are shown in Table S8.2. For compound **5** the predictions were 69% accurate (24 of the 35 off-targets confirmed (<50% remaining activity) with two additional off-targets in the low 50% residual activity range). For compound **6** the prediction was exceedingly accurate as 26 of the 33 targets (79%) were indeed inhibited >50%. To conclude, DDM was able to predict the kinome-inhibitor interaction landscape at a relatively high accuracy.

Predicting the activity profile of a large commercially available hinge binder set

The numerous recent approvals of small molecule kinase inhibitors by the FDA (6 in 2017, 13 since 2014) conceals the fact that none of the last 13 approvals targeted novel kinases or mechanisms. A lot of therapeutic potential may still lie in currently untargeted pathways. To aid in the elucidation of signalling pathways and validation thereof, easy access to KIs is crucial to allow biochemical evaluation by acute inhibition of regulating kinases. To facilitate this, and in the spirit of the Published Kinase Inhibitor Set goals, the model was used to predict the interaction landscape of a large commercially available hinge binder set. To generate a large dataset like this, the Enamine Kinase Hinge Region Directed Library of 18,020 molecules was downloaded and all targets were predicted for these molecules (Figure 8.7A).[23] After filtering for unpredicted kinases and inert molecules all 18,020 inhibitors remain, with 290 kinases as targets (confidence ≥ 0.1, C = 20%). The distribution of the targets across the kinome tree is quite homogenous, as shown in Figure 8.7B. The only major group that is underrepresented are the TKL kinases, but this follows directly from the t-SNE map depicted in Figure S8.1, as the biochemical data generated for the PKIS set did not include many TKL kinases. In total 1,888,418 interactions are predicted, with FLT3 as most inhibited kinase. For all compounds a value similar to the Selectivity Entropy (S_{sel}) could also be calculated, when the confidence parameter is assumed as the association constant ($1/K_d$).[24] This parameter can be used to judge the overall promiscuity of a selected inhibitor, to directly avoid pursuing the least selective scaffold.

As an illustration of the use of this large dataset, two kinase targets were chosen based on recent literature, for which few molecular modulators are known and which were not included in the PKIS dataset. Homeodomain-interacting protein kinase 3 (HIPK3, also known as FIST) is a protein kinase involved in transcription regulation, and is thought to negatively regulate apoptosis in certain triple negative breast cancer cell lines. Inhibition of this kinase may thus prove useful in treatment of these breast cancers.[25] p21-activated
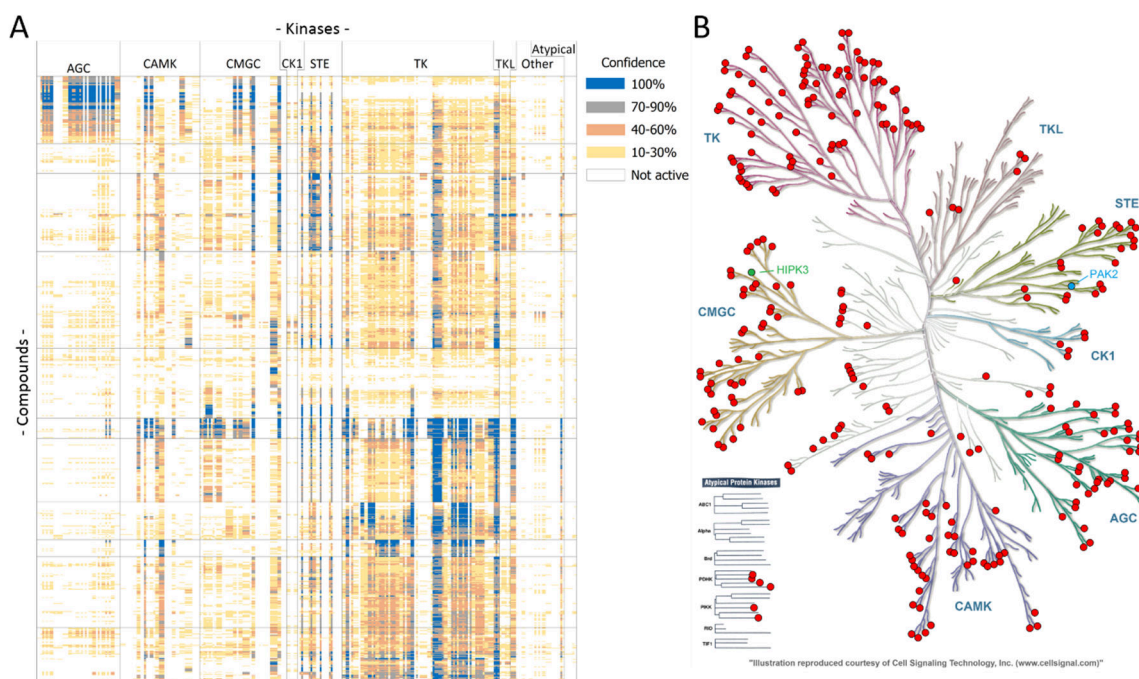
Figure 8.7 | A) Heatmap of the enamine inhibitor set. Compounds are hierarchically sorted on the predicted activity profile, kinases are sorted by group, then family. Cut-off value of 20% is used to increase sensitivity. B) Traditional kinase dendrogram showing the predicted kinases in the Enamine heatmap in red, PAK2 and HIPK3 are highlighted in blue/green. Picture was generated using KinMap.[20]

kinase 2 (PAK2) is activated by CDC42 and RAC1. When active it stimulates cell survival and growth. Recently it was implicated in oestrogen receptor positive (ER⁺) breast cancer as tamoxifen resistance inducer.[26] These two kinases belong to two distinct groups (CMGC and STE, respectively), also distinct from FLT3 (TK) treated above. HIPK3 was predicted as target quite frequently, leading to a set of 172 potential inhibitors ($C = 30\%$, confidence = 1). PAK2 on the other hand is only rarely predicted, and only 57 molecules were found, with poor confidence ($C = 20\%$, confidence = 0.1). The chosen targets are thus also at the far ends of the potentially usable spectrum of the dataset. For both targets 10 molecules were selected based on availability, price, and selectivity entropy. These 20 compounds were purchased and screened in the KinaseProfiler™ platform against both PAK2 and HIPK3. As there was no overlap between the two sets of predicted molecules, the molecules chosen for HIPK3 were used as negative controls for PAK2 and *vice versa* to check for specific enrichment by the model. The activity data, together with the molecular structures, is summarized in Table S8.3. These data showed no identified actives for HIPK3 or PAK2. The high positive prediction value reached for FLT3 above is thus shown not to hold for all targets.

Discussion

Drug discovery is still largely an empirical process, which is challenging, time consuming and hard.[27] The multi-parameter optimization of chemical structures, which is needed to balance the activity and selectivity of a drug candidate, requires the understanding of high-dimensional datasets. Machine learning algorithms have been employed to analyse and

predict compound activity using large datasets with varying success.[11,12,28] Some of the major drawbacks of most computational models are the complexity of the algorithm and the 'black box' nature of the systems. Implementation and interpretation of such systems is not trivial and, consequently, has not been widely adopted by the drug discovery community.

Drug Discovery Maps provides an intuitive, data-driven (bio)molecule similarity clustering procedure using state-of-the-art machine learning techniques. The model is based on the t-Distributed Stochastic Neighbour Embedding algorithm to generate a visualization of molecular similarity in two dimensions.[29,30] Colour is used as a third dimension to inter-actively visualize the biological activity or compound class (chemotype). DDM combines two different maps. The first map depicts the chemical space in which compounds are clustered based on their molecular similarity, whereas in the second map protein targets are clustered based on the chemical similarity of the amino acids making up the kinase domain. By combining both maps, DDM was able to predict bio-activities of small molecules across a protein family. DDM was applied to visualize the chemical space of the Published Kinase Inhibitor Set (PKIS) and the target space of the protein kinase family (kinome). It was able to predict the kinome activity profile of another independent set of kinase inhibitors with comparable or better scores than the currently available machine learning techniques. DDM was then used to identify new hits for the oncogene FMS-like tyrosine kinase 3 (FLT3), a validated therapeutic target for the treatment of acute myeloid leukaemia.[31] The hits were resynthesised and their biological activity was validated in biochemical and cellular assays. Finally, the off-target profile of the hits as predicted by DDM was validated in a panel of kinase assays. The method was then applied to a large commercially available screening collection to predict the interaction landscape of over 18,000 inhibitors. 20 inhibitors, predicted to be active for either HIPK3 or PAK2, were purchased and tested *in vitro*. Unfortunately, none proved to be active, which was an illustration of the limits of our approach. The average positive predictive value of 40% found in the validation and the FLT3 screen is apparently not universally achieved. This could be attributed to the sparse data density around the two selected targets. Another factor to take into account is plain luck of the draw (or lack thereof). If the PPV was 20%, there is an 11% chance of choosing 10 inactive inhibitors at random from a set of predicted actives.

Although the model performs equally well or better than the current computational drug discovery tools, it is envisioned that it can be further improved when more comprehensive datasets become available in the public domain. In the PKIS training set, 364 inhibitors were only tested at two concentrations on approximately 200 unique wild-type kinases. A more expansive dataset of a broader set of more diverse compounds tested on a larger number of kinases in a concentration-response fashion would inherently improve the predictions generated over the entire kinome.

The fact that a complete interaction profile is predicted uniquely enables the ranking of predicted hits by selectivity, for example by calculation of the predicted selectivity entropy.[24] This allows medicinal chemists in theory to rank scaffolds based on promiscuity and also to select for accepted off-targets, that depend on the biological questions or medical indication. This was demonstrated by the KinaseProfiler™ screen of predicted off-targets for the two FLT3 inhibitors.

To aid in the implementation of the tool as it is presented here, a Python based executable including a Graphical User Interface has been made available online[32] (Figure 8.8). The unpackaged Python script with a list of dependencies is also available. Also included is a fully annotated KNIME workflow, to allow step by step execution and analysis. This set of tools should enable the integration of this data-driven approach into any project without any need of investments *a priori*.

To conclude, the machine learning algorithm Barnes-Hut t-SNE was successfully implemented in a drug discovery setting to predict ligand-protein interaction landscapes. The concept of DDM is applicable to a multitude of drug discovery challenges, which, given the proper dataset, can be used to design a small molecule with a balanced set of physico-chemical and biological properties as required for drug candidates. It is envisioned that DDM may make the drug discovery process more efficient.
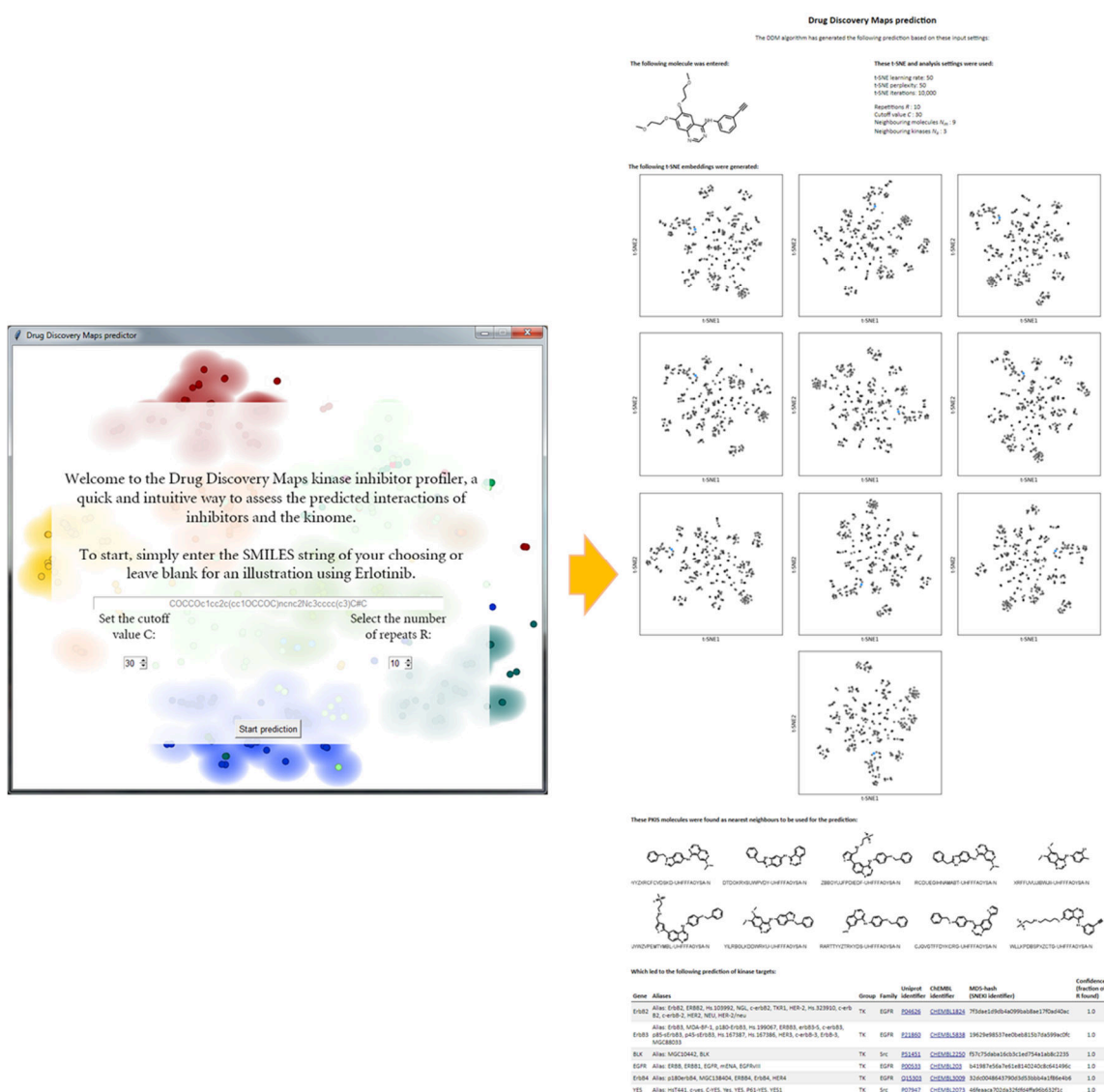
Figure 8.8 | Graphical User Interface (left) and generated output (right) of the Python implementation of the DDM algorithm presented here. Only a SMILES string is required as input, output is provided as depicted on the right. The packaged executable as well as the original Python script have been made available online.

Methods

Training dataset selection

Several large datasets screening dozens of small molecule inhibitors against hundreds of kinases have appeared in literature. Five datasets stand out when judged by size, these are summarized in Table 8.1. The first large dataset (>10,000 data points) published was that of Karaman *et al.* in 2008.[5] 38 commercially available kinase inhibitors were screened *in vitro* against 287 distinct kinases. In 2011, three large datasets were reported. Metz *et al.*[7] published a brief communication describing their analysis of more than 250,000 data points. However, a large portion of these interactions have not been made public. Later that same year, two papers were simultaneously published by Anastassiadis *et al.* and Davis *et al.*[6,8] Both studies comprise known kinase inhibitors and are screened against large portions of the kinome. The latest addition to the field is that by Elkins *et al.*, who screened the Published Kinase Inhibitor Set (PKIS) by GlaxoSmithKline in an *in vitro* assay against 224 kinases, including a number of mutants.[9,13] This dataset is particularly interesting, as the PKIS molecules are available free of charge for academic research, and the set contains the most drug like molecules at the largest coverage. For these reasons, the dataset by Elkins *et al.* was chosen as a starting point for our data-driven approach.

t-SNE for the Published Kinase Inhibitor Set

The curated dataset from ChEMBL 23[33] was used, as the original publication contained some duplicate molecules, and using the open source software package KNIME, the Morgan fingerprints (RD-Kit, 4096-bits, radius = 2) were generated.[33-35] These fingerprints were then clustered using the Python implementation of Barnes-Hut t-SNE.[36] Visual inspection of the embedding shows striking co-localization of the pre-defined chemotypes when the chemotype annotation by Elkins *et al.* is used post-hoc to colour the markers (Figure 8.1). To quantify the clustering quality, the embedding was analyzed using the DBSCAN algorithm, where the eps-parameter was optimized by maximizing the Silhouette Coefficient.[14] The best clustering was found for eps 0.9, for which the Adjusted Rand Index was 0.774. This unbiased clustering produced 33 clusters (Figure 8.1), 29 of which were significantly ($P < 0.0001$, hypergeometric test) enriched for a manually attributed chemotype. Of the 31 chemotypes annotated, 23 were fully comprised in one computationally assigned cluster. To inspect where it diverges from the human annotation, we inspected the cluster dominated by the 3-amino-pyrazolopyridines closer (Figure 8.1B). Inspection of the co-clustering of chemotypes shows that this cluster also contains all 3-amino-pyrazolopyridazines and the indazole-3-carboxamides, structurally very similar classes, as compounds **1**, **2** and **3** illustrate.

Nearest neighbour selection of molecules using t-SNE and initial target predictions

By appending a molecule to the PKIS dataset and regenerating the t-SNE mapping, simple Euclidian distances could be used to find its nearest neighbours. Selecting a set of neighbouring molecules should account for most chemical variation, leading to an 'average molecule' highly similar to our (new) molecule of interest.

For this set of neighbouring molecules, the interaction landscapes against 200 unique non-mutant kinases have been experimentally measured, which can be averaged to yield a predicted inhibition value for the new molecule for all these kinases. For the PKIS dataset inhibition was measured at 100 nM and 1 μM. The inhibition values measured at 1 μM were used as these have the highest information density and in many clinically relevant experiments low micromolar concentrations can (locally) be reached, thus warranting a screen for off-targets at this concentration.

Expanding target prediction using t-SNE to find most similar kinases

To be able to extrapolate from the kinases measured in the PKIS set to incorporate more of the kinome, again a t-SNE based approach was used. Of the entire kinome, the amino acid sequences of the kinase domains containing the ATP-binding pocket, were aligned and expressed using a fingerprint based on physicochemical properties, derived by Heil *et al.*[15] The resulting t-SNE embedding is shown in Figure 8.2 of the main text, and reproduced with annotation of the inclusion in the PKIS set (Figure S8.3). With this similarity mapping in hand, the prediction based on the training dataset could be expanded to theoretically include the whole human kinome, by considering neighbouring kinases as plausible additional targets. The distribution of the measured kinases in the PKIS set is visualized in Figure S8.3 and is rather well, but certainly not homogeneous. This means that for fair parts of the kinome no truly reliable prediction can be made based only on this set, as there are no close neighbouring kinases measured. In the workflow all kinases are however still included, as this will allow any dataset to be loaded as training set, without large adaptations.

The DDM workflow

The workflow envisaged was briefly described in Figure 8.3, and is depicted in more detail in Figure S8.1: a molecule of interest is converted, via its SMILES representation, into a 4096-bit fingerprint. This fingerprint is appended to the PKIS molecules and a 2D t-SNE embedding is generated, as in Figure 8.1A. Based on this embedding, the closest molecules are selected and their measured bioactivity is averaged and considered as the activity prediction. Then, using the t-SNE embedding of the kinase domains (Figure 8.2), the kinases most similar to the predicted kinases are appended to this prediction, which is then the output of the model. Since the t-SNE algorithm is inherently stochastic, this whole process is repeated R times, after which the initial outputs are weighed and finally returned as overall output. The model thus accepts any molecular string representation and returns a list of predicted kinase targets, with a confidence parameter based on the number of repetitions in which a specific target has been found. The PKIS molecules the prediction is based upon can also be viewed, to assess with a chemical eye whether the prediction is to be trusted.

Several optimizable parameters naturally arise in this workflow: the number of considered neighbouring molecules N_m, the number of considered neighbouring kinases N_k, the number of repetitions of the whole process R, and the cut-off value C of the mean inhibition above which an inhibitor is deemed active against that kinase.
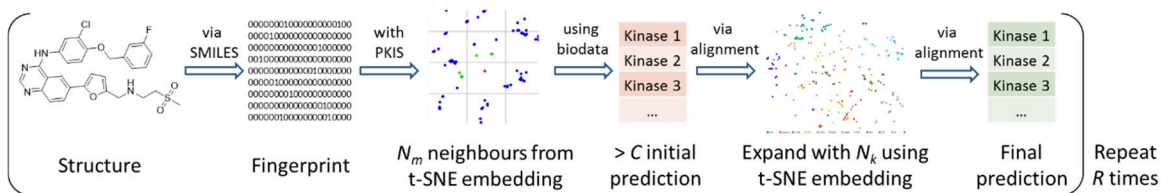
Figure S8.1 | Schematic overview of the DDM workflow. A chemical structure is converted into a Morgan fingerprint using the SMILES-string, and appended to the PKIS molecules set. After t-SNE embedding, the closest neighbour molecules (N_m) are identified and their biodata is averaged. A minimal average inhibition (C) is used as criterion to yield the initial prediction. The N_k kinases with the most similar kinase domain are then found in the kinase t-SNE embedding and appended to form the final prediction. This is repeated R times, weighing the final predictions to return a weighed final prediction.

Optimization of the model parameters

To optimize the model parameters and validate our model, the dataset generated by Karaman *et al*. was used as test set as this provided a comprehensive but diverse set of molecules, biochemically tested against a large set of kinases. For this test set, all K_d values below 1 µM were considered as actives, to mirror the training set, also measured at 1 µM. A multidimensional optimization was performed, initially varying the cut-off C and the two neighbouring values, N_m and N_k (Figure S8.2A and S8.2B). The number of repetitions was also optimized, as depicted in Figure S8.2C. The best predictions were found for 9 neighbouring molecules, 3 neighbouring kinases, and 10 repetitions. The cut-off percentage was found to be a valuable tuning parameter to either have a high sensitivity (low C) or rather a high positive predictive value (high C), which can be chosen depending on the projects specific demands. This is further illustrated by the ROC-curve in Figure 8.5. The mid-way value of C = 30% is recommended. The area under the ROC-curve is 0.76, scoring it as 'fair' according to standard criteria.

Comparison with state of the art methods

With this optimization of the model completed, the next step was to compare the new t-SNE based model with the state of the art in literature. To this end, three proteochemometric (PCM) models and a QSAR model were trained according to published procedures, and the DDM method was compared with those and a random model.[12,37] The result of this comparison is summarized in Figure 8.4, which shows that the DDM model is significantly better than both the random model and the QSAR, and performs similarly when compared to the PCM variants. A variant of the DDM model where the standard Tanimoto distances were used to find the nearest molecular neighbours in the first step of the workflow was also included in the comparison. t-SNE was used in the kinase lookup step. This DDM-Tanimoto variant performed rather well, and not significantly worse than our t-SNE based approach, for this test set.
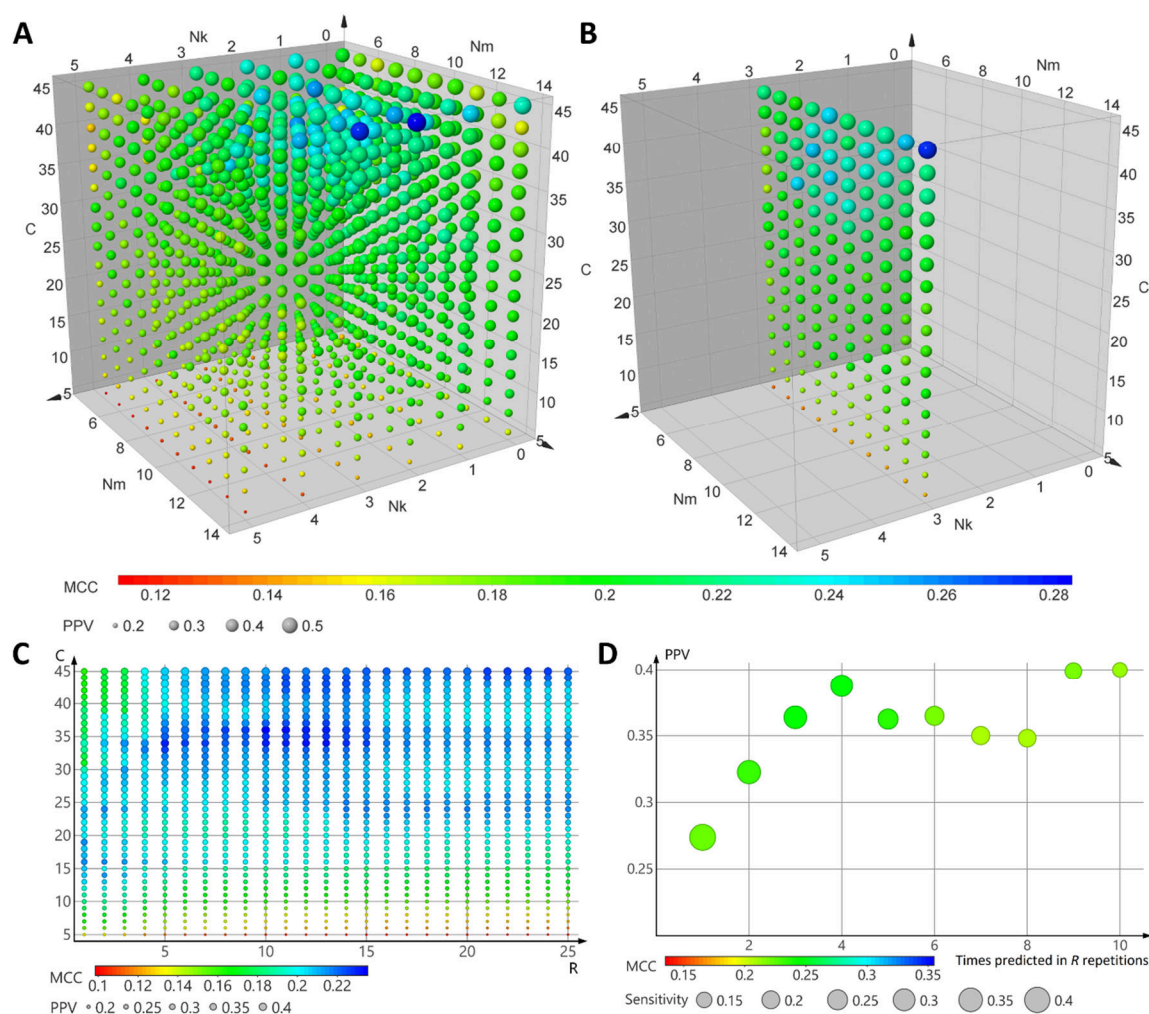
Figure S8.2 | Optimization of the model parameters using the Karaman *et al.* dataset as test set. Optimized parameters are the number of neighbouring molecules used (N_m), the number of neighbouring kinases used, N_k, and the cut-off inhibition percentage deemed active (*C*). Markers are sized based on the positive predictive value (PPV) and coloured based on the Matthews correlation coefficient (MCC). In A), all possible combinations are shown. In B), only the series for $N_k = 3$ is shown for clarity.

Bioactivity datasets used

Activity data for Karaman *et al.* and the Published Kinase Inhibitor Set by Elkins *et al.* was retrieved from the ChEMBL database version 23.[33,38] Data was retrieved from a local SQL install, directly from the website or through the KNIME extensions provided by the EMBL-EBI.

Kinase sequence information and bitstring

Sequence information of the kinase domains was retrieved from the KinBase situated on kinase.com, based on the paper by Manning *et al.*[16] Missing (pseudo)kinases were appended using data from the ChEMBL database version 23 and Uniprot.[39] The mapping of ChEMBL, Uniprot and KinBase identifiers was performed based on the KinBase website and extensive manual curation. The kinase domains were aligned using the online Clustal Omega tool provided by the EMBL-EBI.[40] The standard "Clustal w/o numbers" output generated was transformed to a bitstring using the amino acid fingerprints as provided in Heil *et al.*[15] with the following additions: alignment dashes (-), stops (*) and blanks (X) were all considered empty, represented by 23 0's.

t-SNE algorithm

All t-SNE embeddings were generated with the Python Scikit-learn (v. 0.19) implementation of the Barnes-Hut t-SNE algorithm, either implemented in a 'Python for KNIME' node or as part of a Python script.[36]

DBSCAN algorithm

The Python implementation of DBSCAN, available through the Scikit-learn module Cluster (v. 0.19) was used.[36] The clustering quality metrics Silhouette Coefficient and Adjusted Rand Index were calculated with the Metrics module and the former was optimised by tuning the eps-parameter in steps of 0.05. For the PKIS inhibitors a minimal cluster size of 1 was chosen, for the more disperse plotted kinases the minimal cluster size was set to 10. Comparisons to manual attributions was done manually. Statistical evaluation was performed using Microsoft Excel 2013.

Cheminformatics tools

All molecular descriptors, molecular representations (SMILES, InChIKey) and fingerprints were generated using the RDKit software, either using the KNIME extensions or as the Python implementation.[34]

QSAR and PCM models

QSAR and PCM models were trained as has been described before but for PCM using the fully aligned sequences rather than the binding site and using classification rather than regression.[12,37] PCM models were trained at an activity cut-off of 30%, 40%, and 50%. Pipeline Pilot 2016 (version 16.2.0.58 by BioVia) was used to process the data and random forests were trained in R (version 3.3, package randomforest) using 500 trees, equalizing class sizes, and randomly sampling the square root of the total present descriptors at individual splits.[41] Protein descriptors used were the first three z-scales with a mean value for each sequence for each z-scale.[42] Chemical descriptors used were circular fingerprints with a diameter of 6 bonds (FCFP_6) and physicochemical descriptors as was done previously.[37,43]

PCM models were trained on the full set of kinases and PKIs. QSAR models were trained per kinase but only if at least 5 active and 10 inactive PKIs per kinase where present (at a cut off of 30%). If no QSAR was trained missing predictions were completed to avoid bias and have the ability to compare identical prediction counts. This was obtained using a random number generator in the range of 0-1 where > 0.5 was deemed 'active' as was done previously.[37]

Out of Bag error estimates of PCM models trained on the full set were 8.47% (30% cut off), 8.15 % (40% cut off), and 7.83 % (50% cut off).

Docking of 5 and 6 in FLT3 crystal structure

Docking and structure based modeling was performed in the Schrödinger suite.[44] For 5 a DFG-in model was constructed on the basis of 4RT7 and 3LCD, in a similar fashion as has been done before,[45] using the knowledge-based potential in Prime.[46,47] Subsequently 5 was docked into this model using induced fit docking.[48] For the induced fit docking a hydrogen bond constraint on the backbone of Cys96 was used. Since 6 shares the same substructure as quizartinib, the crystal structure of FLT3 co-crystalized with quizartinib (4RT7)[49] was used as a starting point. 6 was docked using Glide SP.[50] Ligand surfaces, structure renderings and 2D interaction plots were generated using Discovery Studio Visualizer v16.1.

Statistical methods

Clustering enrichment was analysed using a hypergeometric test and was calculated using Microsoft Excel 2013. Significance was attributed only if P < 0.0001.

In the comparison of the quality of prediction of the various prediction models a regular 2-way ANOVA with Tukey's multiple comparison test was performed using GraphPad Prism 7. Significance is attributed according to the standard GraphPad style: * $p < 0.0332$, ** $p < 0.0021$, *** $p < 0.0002$, **** $p < 0.0001$.

High Troughput Screening FLT3

20 nL of 2 mM inhibitor solution in DMSO (row 9-48) or control (DMSO, row 1-8) was dispensed in a 1536-well plate. 2 µL assay buffer (50 mM HEPES (pH 7.5), 1 mM EGTA, 10 mM $MgCl_2$, 0.01% Tween-20, 2 mM DTT) without protein was loaded in row 5-6 as negative controls. Rows 1-4 and 7-48 were charged with 2 µL 0.75 ng/µL FLT3 in assay buffer. The plates were spun down for 30 seconds at 187.5x g and incubated for 30 minutes in the dark. Subsequently, 2 µL of substrate solution was added to all wells (50 mM HEPES (pH 7.5), 1 mM EGTA, 10 mM $MgCl_2$, 0.01% Tween-20, 2 mM DTT, 600 µM ATP, 12.5 nM Lance TK-peptide, 4 nM Lance anti-phosphotyrosine). The plates were spun down for 30 seconds at 187.5x g and incubated for 90 minutes in the dark at RT. Plates were then read on the Envision plate reader (Excitation 337 nm (laser), Emission first filter 615 nm, second filter 665 nm). Data was analysed using ActivityBase. Final assay concentrations: 10 µM

inhibitor, 300 μM ATP, 0.365 ng/μL FLT3, 6.25 nM (0.5 K_M) Lance TK-peptide, 2 nM Lance anti-phosphotyrosine, 0.5% DMSO.

In vitro FRET based FLT3 assay

In a 384-wells plate, 5 μL kinase+petide mix (0.06 ng/μL FLT3, 200 nM Lance TK-peptide) in assay buffer (50 mM HEPES (pH 7.5), 1 mM EGTA, 10 mM $MgCl_2$, 0.01% Tween-20, 2 mM DTT) was dispensed. Seperately inhibitor solutions (10 μM – 0.1 pM) were prepared in assay buffer containing 400 μM ATP and 1% DMSO. 5 μL of these solutions was dispensed and the plate was incubated for 90 minutes in the dark. After 90 minutes the reaction was quenched by the addition of 10 μL of 20 mM EDTA containing 4 nM Lance anti-phosphotyrosine. After thorough mixing it was incubated for 60 minutes in the dark. The FRET fluorescence was measured on a Tecan Infinite M1000 Pro plate reader (excitation 320 nm, emission first filter 615 nm, second filter 665 nm). Data was processed using Microsoft Excel 2013, pIC_{50} values were fitted using GraphPad Prism 7.0. Final assay concentrations during phosphorylation: 200 μM ATP, 0.03 ng/μL FLT3, 100 nM Lance TK-peptide, 0.5% DMSO)

In situ testing of kinase inhibitors

MV4:11 cells were grown in IMDM with 10% fetal bovine serum at 37 °C under 5% CO_2. For viability assays, 10,000 cells were seeded per well in a 96-wells plate and inhibitors were added at the indicated concentration. Three days later, cell viability was measured using the Cell Titer Blue viability assay (Promega), fluorescence was measured using the Clariostar (BMG Labtech). Relative survival was normalized to the untreated control and corrected for background signal.

Synthesis of *in situ* tested kinase inhibitors

General remarks

All reactions were performed using oven- or flame-dried glassware and dry solvents. Reagents were purchased from Sigma-Aldrich, Acros, and Merck and used without further purification unless noted otherwise. All moisture sensitive reactions were performed under a nitrogen atmosphere.
^1H and ^{13}C NMR spectra were recorded on a Bruker AV-400 (400 MHz). Used software for interpretation of NMR-data was Bruker TopSpin 1.3 and MestreNova 11.0. Chemical shift values are reported in ppm with tetramethylsilane or solvent resonance as the internal standard ($CDCl_3$: δ 7.26 for ^1H, δ 77.16 for ^{13}C; DMSO-*d6*: δ 2.50 for ^1H, δ 39.52 for ^{13}C). Data are reported as follows: chemical shifts (δ), multiplicity (s = singlet, d = doublet, dd = double doublet, td = triple doublet, t = triplet, q = quartet, bs = broad singlet, m = multiplet), coupling constants *J* (Hz), and integration. Liquid chromatography was performed on a Finnigan Surveyor LC/MS system, equipped with a C18 column. Flash chromatography was performed using SiliCycle silica gel type SiliaFlash P60 (230–400 mesh). TLC analysis was performed on Merck silica gel 60/Kieselguhr F254, 0.25 mm. Compounds were visualized using $KMnO_4$ stain (K_2CO_3 (40 g), $KMnO_4$ (6 g), and water (600 mL)) or CAM stain ($Ce(NH_4)_4(SO_4)_4 \cdot 2H_2O$ (ceric ammonium sulfate: 10.0 g); ammonium molybdate (25 g); conc. H_2SO_4 (100 mL); H_2O (900 mL)). Preparative HPLC (Waters, 515 HPLC pump M; Waters, 515 HPLC pump L; Waters, 2767 sample manager; Waters SFO System Fluidics Organizer; Waters Acquity Ultra Performance LC, SQ Detector; Waters Binary Gradient Module) was performed on a Waters XBridgeTM column (5 μM C18, 150 x 19 mm). Diode detection was done between 210 and 600 nm. Gradient: ACN in (H_2O + 0.2% TFA).

1-(3,4,5-Trimethoxyphenyl)guanidine (**7**)
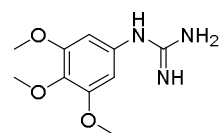
3,4,5-trimethoxyaniline (500 mg, 2.73 mmol) and cyanamide (574 mg, 13.6 mmol) were dissolved in ethanol (15 mL) before nitric acid (69%wt, 0.20 mL, 3.0 mmol) was added. The mixture was refluxed for 40 h and concentrated *in vacuo*. The resulting residue was suspended in diethylether (10 mL) and kept at 4 °C overnight. It was then filtered and air dried. The nitrate salt of the title compound was obtained as a dark purple solid (0.60 g, 2.1 mmol, 76%). ^1H NMR (400 MHz, DMSO) δ 9.51 (bs, 1H), 8.73 – 8.12 (m, 2H), 6.55 (s, 2H), 5.44 (s, 2H), 3.77 (s, 6H), 3.65 (s, 3H). ^{13}C NMR (101 MHz, DMSO) δ 161.53, 155.86, 153.38, 130.67, 102.91, 60.06, 56.01.

1-(6-Chloropyridin-3-yl)-3-(dimethylamino)prop-2-en-1-one (8)
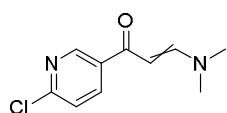
1-(6-chloropyridin-3-yl)ethan-1-one (1.00 g, 6.43 mmol) was dissolved in toluene (20 mL) and dimethylformamide diethylacetal (1.65 mL, 9.64 mmol) was added. The mixture was stirred for 16 h at 80 °C. TLC showed near complete conversion. The mixture was cooled to 50 °C and slowly concentrated at reduced pressure. The resulting crude residue was dissolved in warm toluene and pentane was slowly added. A yellow precipitate formed which was filtered off and rinsed with pentane. After drying this yielded the title compound (1.08 g, 5.12 mmol, 80%) as a yellow solid. ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.85 (dd, J = 0.7, 2.5 Hz, 1H), 8.17 (dd, J = 2.5, 8.3 Hz, 1H), 7.86 (d, J = 12.2 Hz, 1H), 7.38 (dd, J = 0.8, 8.2 Hz, 1H), 5.63 (d, J = 12.1 Hz, 1H), 3.20 (s, 3H), 2.97 (s, 3H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 184.90, 155.07, 153.38, 149.08, 138.10, 134.63, 124.10, 91.54, 45.46, 37.60.

4-(6-Chloropyridin-3-yl)-N-(3,4,5-trimethoxyphenyl)pyrimidin-2-amine (5)
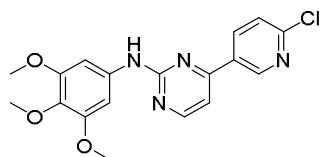
To a solution of 7 (376 mg, 1.30 mmol) and 8 (250 mg, 1.19 mmol) in ethanol (20 mL) K$_2$CO$_3$ (492 mg, 3.56 mmol) was added. The resulting mixture was refluxed for 20 h. The reaction was quenched by the addition of saturated aqueous Na$_2$CO$_3$ (15 mL) and the water layer was extracted with EtOAc (3x 25 mL). The combined organic layers were washed with brine, dried (MgSO$_4$), filtered and concentrated *in vacuo*. The resulting crude was purified using silica column chromatography (50-70% EtOAc/pentane). The resulting yellow solid was crystallised from warm diethylether to yield the title compound (137 mg, 0.37 mmol, 31%) as light yellow crystals. ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.07 (dd, J = 0.7, 2.5 Hz, 1H), 8.52 (d, J = 5.2 Hz, 1H), 8.32 (dd, J = 2.5, 8.3 Hz, 1H), 7.46 (dd, J = 0.7, 8.4 Hz, 1H), 7.38 (s, 1H), 7.14 (d, J = 5.2 Hz, 1H), 6.99 (s, 2H), 3.90 (s, 6H), 3.85 (s, 3H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 161.53, 160.28, 159.11, 153.74, 153.47, 148.62, 137.11, 135.30, 131.78, 124.52, 108.15, 97.42, 61.17, 56.26.

4-(Pyridin-4-yloxy)aniline (9)
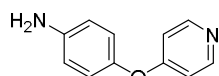
4-aminophenol (2.18 g, 20 mmol), 4-chloropyridine (3.15 g, 21 mmol) and NaOH (2.0 g, 50 mmol) were dissolved in DMSO (50 mL) and heated to 100 °C for 17 h. The mixture was cooled to RT and poured into ice water (300 mL). This was extracted with 10% MeOH in chloroform (4x 150 mL). The combined organic layers were washed with brine (2x 150 mL), dried (Na$_2$SO$_4$), filtered and concentrated *in vacuo*. The residue was flushed over a silica pad with pure EtOAc and concentrated to yield the title compound as a white solid (2.5 g, 13 mmol, 65%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.51 – 8.33 (m, 2H), 6.97 – 6.85 (m, 2H), 6.85 – 6.77 (m, 2H), 6.77 – 6.62 (m, 2H), 3.50 (s, 2H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 165.89, 151.25, 145.80, 144.25, 122.12, 116.36, 111.70.

1-(5-(tert-Butyl)isoxazol-3-yl)-3-(4-(pyridin-4-yloxy)phenyl)urea (6)
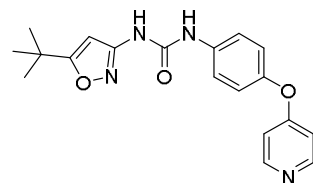
3-amino-5-tert-butylisooxazole (0.20 g, 1.4 mmol) was dissolved in DCM (14 mL) and DIPEA (0.50 mL, 2.9 mmol) was added. The mixture was cooled to 0 °C and triphosgene (0.42 g, 1.4 mmol) was added. After stirring the mixture for 19 h at RT it was refluxed for 1 h before the reaction was quenched by the addition of saturated aqueous NaHCO$_3$ (10 mL). The water layer was separated and extracted with DCM (3x 10 mL). The combined organic layers were washed with brine, dried (Na$_2$SO$_4$) and filtered. The solvent was removed under reduced pressure and the crude isocyanate was dissolved in 1,4-dioxane (14 mL). 9 (0.29 g, 1.6 mmol) was added to the solution and the mixture was heated to 110 °C for 2.5 h. After cooling to RT the mixture was diluted with DCM (30 mL) and saturated aqueous NaHCO$_3$ (30 mL) was added. The water layer was separated and extracted with DCM (3x 30 mL). The combined organic layers were washed with brine, dried (Na$_2$SO$_4$), filtered and concentrated *in vacuo*. The resulting yellow oil was purified by preparative HPLC to yield the title compound as a slightly yellow oil (0.22 g, 0.62 mmol, 44%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.41 (s, 1H), 8.67 (s, 1H), 8.47 (dd, J = 1.5, 5.1 Hz, 2H), 7.63 – 7.54 (m, 2H), 7.10 – 7.03 (m, 2H), 6.89 – 6.81 (m, 2H), 5.95 (s, 1H), 1.36 (s, 9H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 181.60, 165.38, 158.45, 152.58, 151.22, 149.95, 135.58, 122.15, 121.60, 112.17, 91.94, 33.08, 28.75.

Supplementary Tables and Figures

Table S8.1 | Overview of prediction qualities split per compound. Data corresponds to R = 10 and C = 30%. PPV = positive predictive value, NPV = negative predictive value, MCC = Matthew's correlation coefficient.

CHEMBL ID	True positives	False positives	True negatives	False negatives	Sensi-tivity	Speci-ficity	PPV	NPV	MCC
CHEMBL10	4	8	254	13	0.24	0.97	0.33	0.95	0.24
CHEMBL101253	7	41	230	1	0.88	0.85	0.15	1.00	0.32
CHEMBL103667	5	21	228	25	0.17	0.92	0.19	0.90	0.09
CHEMBL119385	2	11	265	1	0.67	0.96	0.15	1.00	0.31
CHEMBL124660	9	66	202	2	0.82	0.75	0.12	0.99	0.25
CHEMBL1336	20	28	217	14	0.59	0.89	0.42	0.94	0.41
CHEMBL1421	19	5	207	48	0.28	0.98	0.79	0.81	0.40
CHEMBL14762	0	49	227	3	0.00	0.82	0.00	0.99	-0.05
CHEMBL1721885	35	29	147	69	0.34	0.84	0.55	0.68	0.20
CHEMBL191003	20	43	174	42	0.32	0.80	0.32	0.81	0.12
CHEMBL215152	1	5	258	15	0.06	0.98	0.17	0.95	0.07
CHEMBL221959	1	46	226	6	0.14	0.83	0.02	0.97	-0.01
CHEMBL223360	26	62	187	4	0.87	0.75	0.30	0.98	0.41
CHEMBL24828	20	16	217	26	0.43	0.93	0.56	0.89	0.41
CHEMBL259084	20	65	191	3	0.87	0.75	0.24	0.98	0.37
CHEMBL261849	4	74	201	0	1.00	0.73	0.05	1.00	0.19
CHEMBL278041	2	14	249	14	0.13	0.95	0.13	0.95	0.07
CHEMBL296468	2	3	250	24	0.08	0.99	0.40	0.91	0.14
CHEMBL31965	3	3	261	12	0.20	0.99	0.50	0.96	0.29
CHEMBL388978	57	3	42	179	0.24	0.93	0.95	0.19	0.16
CHEMBL428690	2	19	228	30	0.06	0.92	0.10	0.88	-0.02
CHEMBL440084	0	12	263	4	0.00	0.96	0.00	0.99	-0.03
CHEMBL477772	22	56	185	16	0.58	0.77	0.28	0.92	0.26
CHEMBL483321	2	4	272	1	0.67	0.99	0.33	1.00	0.46
CHEMBL522892	24	31	183	41	0.37	0.86	0.44	0.82	0.24
CHEMBL535	44	30	117	89	0.33	0.80	0.59	0.57	0.14
CHEMBL553	3	3	253	20	0.13	0.99	0.50	0.93	0.23
CHEMBL554	2	0	277	0	1.00	1.00	1.00	1.00	1.00
CHEMBL558752	10	72	185	12	0.45	0.72	0.12	0.94	0.10
CHEMBL572878	2	11	199	67	0.03	0.95	0.15	0.75	-0.05
CHEMBL572881	6	16	242	15	0.29	0.94	0.27	0.94	0.22
CHEMBL573339	0	58	219	2	0.00	0.79	0.00	0.99	-0.04
CHEMBL574738	48	12	165	54	0.47	0.93	0.80	0.75	0.47
CHEMBL607707	11	45	203	20	0.35	0.82	0.20	0.91	0.14
CHEMBL608533	11	10	169	89	0.11	0.94	0.52	0.66	0.10
CHEMBL91829	1	12	240	26	0.04	0.95	0.08	0.90	-0.01
CHEMBL939	3	15	254	7	0.30	0.94	0.17	0.97	0.18
CHEMBL941	10	115	153	1	0.91	0.57	0.08	0.99	0.19
Average	12.05	29.29	211.58	26.18	0.38	0.88	0.32	0.89	0.21

Table S8.2 | Summary of the off-target screen performed for the *in situ* tested compounds **5** and **6**. Data are averages of two duplicates, and are percentages activity compared to vehicle control.

	5				6		
Gene Name	Uniprot accession	CHEMBL ID	Remaining activity (%)	Gene Name	Uniprot accession	CHEMBL ID	Remaining activity (%)
ABL1	P00519	CHEMBL1862	26	ABL1	P00519	CHEMBL1862	23
ABL2	P42684	CHEMBL4014	12	ABL2	P42684	CHEMBL4014	28
AurA	O14965	CHEMBL4722	2	AurA	O14965	CHEMBL4722	60
AurB	Q96GD4	CHEMBL2185	2	AurB	Q96GD4	CHEMBL2185	60
AurC	Q9UQB9	CHEMBL3935	8	AurC	Q9UQB9	CHEMBL3935	91
BLK	P51451	CHEMBL2250	34	DDR1	Q08345	CHEMBL5319	7
BRK	Q13882	CHEMBL4601	61	DDR2	Q16832	CHEMBL5122	1
EGFR	P00533	CHEMBL203	96	EphA3	P29320	CHEMBL4954	9
EphA4	P54764	CHEMBL3988	72	EphA4	P54764	CHEMBL3988	28
EphA5	P54756	CHEMBL3987	52	EphA5	P54756	CHEMBL3987	7
EphB2	P29323	CHEMBL3290	14	EphA7	Q15375	CHEMBL4602	-2
EphB3	P54753	CHEMBL4901	100	EphB1	P54762	CHEMBL5072	3
ErbB2	P04626	CHEMBL1824	94	EphB2	P29323	CHEMBL3290	5
ErbB4	Q15303	CHEMBL3009	108	EphB3	P54753	CHEMBL4901	55
FER	P16591	CHEMBL3982	52	FER	P16591	CHEMBL3982	86
FES	P07332	CHEMBL5455	19	FES	P07332	CHEMBL5455	86
FGFR1	P11362	CHEMBL3650	58	FGFR1	P11362	CHEMBL3650	6
FGFR3	P22607	CHEMBL2742	48	FGFR3	P22607	CHEMBL2742	29
FGFR4	P22455	CHEMBL3973	91	FGFR4	P22455	CHEMBL3973	70
FGR	P09769	CHEMBL4454	30	FLT1	P17948	CHEMBL1868	1
FLT1	P17948	CHEMBL1868	1	FLT3	P36888	CHEMBL1974	1
FLT3	P36888	CHEMBL1974	3	FLT4	P35916	CHEMBL1955	-1
FLT4	P35916	CHEMBL1955	1	FMS	P07333	CHEMBL1844	-1
FMS	P07333	CHEMBL1844	-1	KDR	P35968	CHEMBL279	7
FYN	P06241	CHEMBL1841	9	KIT	P10721	CHEMBL1936	7
HCK	P08631	CHEMBL3234	17	MUSK	O15146	CHEMBL5684	13
KDR	P35968	CHEMBL279	5	PDGFRa	P16234	CHEMBL2007	15
KIT	P10721	CHEMBL1936	2	PDGFRb	P09619	CHEMBL1913	9
LCK	P06239	CHEMBL258	15	RET	P07949	CHEMBL2041	-2
LYN	P07948	CHEMBL3905	26	TIE2	Q02763	CHEMBL4128	1
PDGFRa	P16234	CHEMBL2007	47	TRKA	P04629	CHEMBL2815	5
PDGFRb	P09619	CHEMBL1913	46	TRKB	Q16620	CHEMBL4898	0
RET	P07949	CHEMBL2041	53	TRKC	Q16288	CHEMBL5608	-2
SRC	P12931	CHEMBL267	4				
TNK1*	Q13470	CHEMBL5334	-				
YES	P07947	CHEMBL2073	10				

*TNK1 was not available in the KinomeProfiler™

Table S8.3 | Overview of the molecular structures and biological activity for the 20 purchased Enamine small molecules. Biological data are averages of two duplicates, and are percentages activity remaining compared to vehicle control.

	Predicted for PAK2		
Enamine ID	Structure	HIPK3 inhibition (%)	PAK2 inhibition (%)
Z1272746053		-4	-6
Z1537171258		2	-5
Z1557057955		3	1
Z1695667186		-2	8
Z226653560		-2	5
Z228788294		-3	2
Z237652740		2	3
Z316373096		7	5
Z367648718		1	10
Z785216724		-1	8

Table S8.3 (continued) | Overview of the molecular structures and biological activity for the 20 purchased Enamine small molecules. Biological data are averages of two duplicates, and are percentages activity remaining compared to vehicle control.

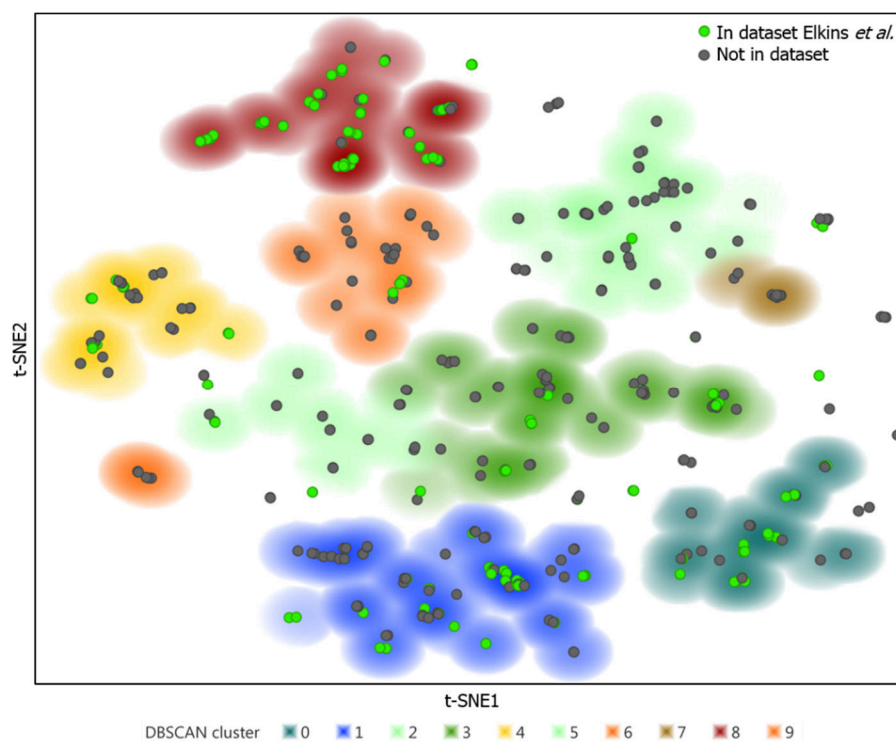| Predicted for HIPK3			
Enamine ID	Structure	HIPK3 inhibition (%)	PAK2 inhibition (%)
Z1041114058		-6	11
Z1102997288		3	-1
Z1272153794		-4	-8
Z1281728869		-1	2
Z1502786962		-6	9
Z1625107708		1	9
Z17559373		-1	-11
Z31057533		-2	6
Z644918986		2	13
Z909646456		0	-2

Figure S8.3 | t-SNE embedding of physicochemical fingerprint of the kinase domains of 535 human kinases as in Figure 8.2. Markers are coloured based on the presence (green) or absence (grey) in the Elkins *et al.* PKIS dataset.

References

1. Johnson, L. N. & Lewis, R. J. Structural Basis for Control by Phosphorylation. *Chem. Rev.* **101**, 2209–2242 (2001).
2. Adams, J. A. Kinetic and Catalytic Mechanisms of Protein Kinases. *Chem. Rev.* **101**, 2271–2290 (2001).
3. Klaeger, S. *et al.* The target landscape of clinical kinase drugs. *Science* **358**, (2017).
4. Müller, S., Chaikuad, A., Gray, N. S. & Knapp, S. The ins and outs of selective kinase inhibitor development. *Nat. Chem. Biol.* **11**, 818–821 (2015).
5. Karaman, M. W. *et al.* A quantitative analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **26**, 127–132 (2008).
6. Davis, M. I. *et al.* Comprehensive analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **29**, 1046–1051 (2011).
7. Metz, J. T. *et al.* Navigating the kinome. *Nat. Chem. Biol.* **7**, 200–202 (2011).
8. Anastassiadis, T., Deacon, S. W., Devarajan, K., Ma, H. & Peterson, J. R. Comprehensive assay of kinase catalytic activity reveals features of kinase inhibitor selectivity. *Nat. Biotechnol.* **29**, 1039–1045 (2011).
9. Elkins, J. M. *et al.* Comprehensive characterization of the Published Kinase Inhibitor Set. *Nat. Biotechnol.* **34**, 95–103 (2015).
10. Miduturu, C. V. *et al.* High-Throughput Kinase Profiling: A More Efficient Approach toward the Discovery of New Kinase Inhibitors. *Chem. Biol.* **18**, 868–879 (2011).
11. Merget, B., Turk, S., Eid, S., Rippmann, F. & Fulle, S. Profiling Prediction of Kinase Inhibitors: Toward the Virtual Assay. *J. Med. Chem.* **60**, 474–485 (2017).
12. Christmann-Franck, S. *et al.* Unprecedentedly Large-Scale Kinase Inhibitor Set Enabling the Accurate Prediction of Compound–Kinase Activities: A Way toward Selective Promiscuity by Design? *J. Chem. Inf. Model.* **56**, 1654–1675 (2016).
13. Drewry, D. H., Willson, T. M. & Zuercher, W. J. Seeding collaborations to advance kinase science with the GSK Published Kinase Inhibitor Set (PKIS). *Curr. Top. Med. Chem.* **14**, 340–2 (2014).
14. Ester, M., Kriegel, H.-P., Sander, J. & Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. in *KDD-96 Proceedings* 226–231 (1996).
15. Heil, B., Ludwig, J., Lichtenberg-Frate, H. & Lengauer, T. Computational recognition of potassium channel sequences. *Bioinformatics* **22**, 1562–1568 (2006).
16. Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The Protein Kinase Complement of the Human Genome. *Science* **298**, 1912–1934 (2002).
17. Murphy, J. M. *et al.* A robust methodology to subclassify pseudokinases based on their nucleotide-binding properties. *Biochem. J.* **457**, 323–334 (2014).
18. Murphy, J. M. *et al.* The Pseudokinase MLKL Mediates Necroptosis via a Molecular Switch Mechanism. *Immunity* **39**, 443–453 (2013).
19. Josso, N. & Clemente, N. di. Transduction pathway of anti-Müllerian hormone, a sex-specific member of the TGF-β family. *Trends Endocrinol. Metab.* **14**, 91–97 (2003).
20. Eid, S., Turk, S., Volkamer, A., Rippmann, F. & Fulle, S. KinMap: a web-based tool for interactive navigation through human kinome data. *BMC Bioinformatics* **18**, 16 (2017).
21. Larrosa-Garcia, M. & Baer, M. R. FLT3 Inhibitors in Acute Myeloid Leukemia: Current Status and Future Directions. *Mol. Cancer Ther.* **16**, 991–1001 (2017).
22. Charnley, A. K. *et al.* Crystal structures of human RIP2 kinase catalytic domain complexed with ATP-competitive inhibitors: Foundations for understanding inhibitor selectivity. *Bioorg. Med. Chem.* **23**, 7000–7006 (2015).
23. Enamine Ltd. Enamine. Available at: http://www.enamine.net/.
24. Uitdehaag, J. C. & Zaman, G. J. A theoretical entropy score as a single value to express inhibitor selectivity. *BMC Bioinformatics* **12**, 94 (2011).
25. Fokkelman, M. *et al.* Cellular adhesome screen identifies critical modulators of focal adhesion dynamics, cellular traction forces and cell migration behaviour. *Sci. Rep.* **6**, 31707 (2016).
26. Zhang, Y. *et al.* IGF1R signaling drives antiestrogen resistance through PAK2/PIX activation in luminal breast cancer. *Oncogene* **37**, 1869–1884 (2018).
27. Scannell, J. W. & Bosley, J. When Quality Beats Quantity: Decision Theory, Drug Discovery, and the Reproducibility Crisis. *PLoS One* **11**, e0147215 (2016).
28. Sorgenfrei, F. A., Fulle, S. & Merget, B. Kinome-Wide Profiling Prediction of Small Molecules. *ChemMedChem* **13**, 495–499 (2018).
29. Van Der Maaten, L. & Hinton, G. Visualizing Data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
30. Reutlinger, M. & Schneider, G. Nonlinear dimensionality reduction and mapping of compound libraries for drug discovery. *J. Mol. Graph. Model.* **34**, 108–117 (2012).
31. Zarrinkar, P. P. *et al.* AC220 is a uniquely potent and selective inhibitor of FLT3 for the treatment of acute myeloid leukemia (AML). *Blood* **114**, 2984–92 (2009).
32. Janssen, A. P. A. Drug Discovery Maps on GitHub. (2018). Available at: https://github.com/APAJanssen/DrugDiscoveryMaps/.
33. EMBL-EBI. CHEMBL database release 23. (2017). doi:10.6019/CHEMBL.database.23

34. Landrum, G. RDKit: Open-source cheminformatics; http://www.rdkit.org.
35. Berthold, M. R. *et al.* KNIME: The Konstanz Information Miner. in *Data Analysis, Machine Learning and Applications* (ed. Preisach C., Burkhardt H., Schmidt-Thieme L., D. R.) 319–326 (Springer, Berlin, Heidelberg, 2008).
36. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
37. Lenselink, E. B. *et al.* Beyond the hype: deep neural networks outperform established methods using a ChEMBL bioactivity benchmark set. *J. Cheminform.* **9**, 45 (2017).
38. Gaulton, A. *et al.* ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **40**, D1100–D1107 (2012).
39. Consortium, T. U. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **45**, D158–D169 (2017).
40. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
41. Breiman, L. Random Forests. *Mach. Learn.* **45**, 5–32 (2001).
42. van Westen, G. J. *et al.* Benchmarking of protein descriptor sets in proteochemometric modeling (part 2): modeling performance of 13 amino acid descriptor sets. *J. Cheminform.* **5**, 42 (2013).
43. Rogers, D. & Hahn, M. Extended-Connectivity Fingerprints. *J. Chem. Inf. Model.* **50**, 742–754 (2010).
44. Schrödinger Suite: Maestro. (2017).
45. Ke, Y.-Y. *et al.* Homology modeling of DFG-in FMS-like tyrosine kinase 3 (FLT3) and structure-based virtual screening for inhibitor identification. *Sci. Rep.* **5**, 11702 (2015).
46. Jacobson, M. P. *et al.* A hierarchical approach to all-atom protein loop prediction. *Proteins Struct. Funct. Bioinforma.* **55**, 351–367 (2004).
47. Jacobson, M. P., Friesner, R. A., Xiang, Z. & Honig, B. On the role of the crystal environment in determining protein side-chain conformations. *J. Mol. Biol.* **320**, 597–608 (2002).
48. Sherman, W., Day, T., Jacobson, M. P., Friesner, R. A. & Farid, R. Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects. *J. Med. Chem.* **49**, 534–553 (2006).
49. Smith, C. C. *et al.* Characterizing and Overriding the Structural Mechanism of the Quizartinib-Resistant FLT3 'Gatekeeper' F691L Mutation with PLX3397. *Cancer Discov.* **5**, 668–679 (2015).
50. Friesner, R. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **47**, 1739–1749 (2004).

I believe that scientific knowledge has fractal properties, that no matter how much we learn, whatever is left, however small it may seem, is just as infinitely complex as the whole was to start with.

Isaac Asimov

Summary and Future Prospects

General summary

The research presented in this thesis addressed one of the main challenges in modern drug development: selectivity. The whole area of *target-based* drug discovery is built on the notion that a certain protein target is a key regulator in a biological process. Although the use of this concept as the basis for drug discovery could be debated[1], the importance of chemical modulators capable of selectively altering the activity of the target protein is undisputed. The importance of selective inhibitors is especially great when they are used to prove the therapeutic value of a certain target. Generally, a therapeutic window needs to

be determined where the target protein is selectively modulated. At higher concentrations the effects of binding other targets could lead to side-effects and possible toxicity.

Because of the importance of selectivity, many techniques have been developed to assess the exact binding profiles of inhibitors, such as family-wide activity assays (e.g. for kinases and luciferases), and more specific assays like the 44-member CEREP-panel.[2] Other, more untargeted approaches have been developed, e.g. Cellular Thermal Shift Assay (CETSA) and a variety of uses of radiolabelled drug analogues. Another technique which has proven to be highly flexible and thus has found a large variety of applications in (early) drug discovery is activity-based protein profiling (ABPP). The concept of this technique was explained in **Chapter 1**. One field of research where ABPP has shown great use is that involving the endocannabinoid system (ECS), which is under active investigation for its role in anti-nociceptive and anti-inflammatory effects, amongst others.[3–6] This signalling system consists of two cannabinoid receptors (CB1 and CB2), and two main endogenous ligands: 2-arachidonoylglycerol (2-AG) and *N*-arachidonoylethanolamine (anandamide), and their biosynthetic and catabolic enzymes. 2-AG is the most abundant endocannabinoid and is mainly produced by diacylglycerol lipase (DAGL) α and β. The different roles of the two endocannabinoids and the specific functions of the two isoforms of DAGL are still poorly understood. This warrants the search for selective and *in vivo* active inhibitors to be able to study these processes in more detail.

Diacylglycerol lipase inhibitors: activity and specificity

Chapter 2 introduced the hit-to-lead optimization of a novel DAGL inhibitor scaffold: the 1,2,4-triazole-sulfonamides. This molecular series was found in a previously reported high-throughput screen.[7] In this Chapter the structure activity relations around the core scaffold were explored by screening 65 variants of the HTS-hit molecule. The potency for DAGL-α varied 400-fold from an IC_{50} of approximately 1 μM to 2.5 nM. By screening the library in mouse brain proteome using ABPP, the most selective inhibitor was identified and selected for further profiling. Compound **1** (Figure 9.1), a 2-benzylpiperidine substituted triazole urea, was active in mouse Neuro-2a cells. To test *in vivo* target engagement, compound **1** and an alkynylated analogue were administered intraperitoneal to mice. The serine hydrolase activities were analysed in the brain and spleen using gel-based ABPP after two hours. Both compounds passed the blood-brain barrier and potently engaged DAGL-α in the brain.

A single dose of 50 mg/kg was predominantly chosen to demonstrate target engagement *in vivo*. Due to the high dose, off-target activity was observed. A dose-response experiment is advised, which is ideally coupled to pharmacokinetic measurements, to determine the optimal dose of target engagement with minimal off-target activity. Considering the weak inhibition of DAGL-β in the spleen, it is possible that a dose can be selected where DAGL-α is selectively inhibited. This would enable for the first time detailed investigations of the separate DAGL isoforms by acute inhibition. In light of the similarities in the inhibition profile of **1** and DH376[5], pairing the inhibitors might lead to the identification of DAGL-β-specific effects in, for example, feeding behaviour.[8]
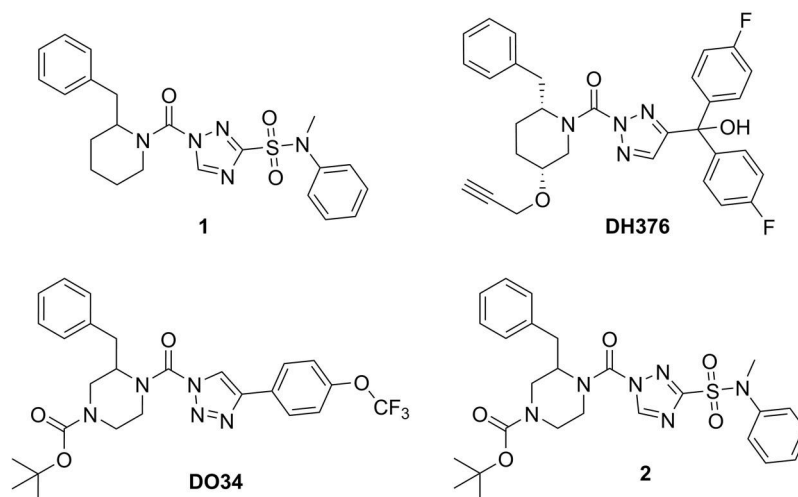
Figure 9.1 | Chemical structures of inhibitor **1**, reference inhibitors DH376 and DO34 and inhibitor **2**.

Chapter 3 was focussed on optimizing the physicochemical properties of compound **1**. In particular, the topological polar surface area (TPSA) was increased to restrict the inhibitors to the periphery to target predominantly DAGL-β. Hit optimization resulted in the discovery of highly potent compound **2**, in which the piperazine moiety of DO34 (Figure 9.1) was incorporated. Next, a small 24-membered library was synthesized in which the *tert*-butyl carbamate was replaced with various amides, ureas, tertiary amines and sulfonamides. This led to four highly potent and *in vitro* selective compounds, which were tested in Neuro-2a cells. A more than tenfold drop in IC_{50} values indicated that cell penetration was less efficient, possibly due to a reduced passive diffusion as a result of the increased polarity. Remarkably, the *in situ* the selectivity window was also reduced compared to the *in vitro* values.

Cellular profiling of the library is advised to assess potency and selectivity of different substitution patterns, ideally identifying more promising lead compounds. If more selective inhibitors are identified, an *in vivo* dose-response experiment should be performed to identify dosages that show full target engagement in peripheral tissue, but leave the brain unaffected. These compounds could then be instrumental in the elucidation of the workings and importance of the peripheral endocannabinoid signalling, and the role of DAGL-β in this process.

Chapter 4 showed the in depth investigation of the structure kinetic relationships of a small set of DH376 derivatives. A previously reported surrogate substrate assay was adapted to determine the affinity (K_i) and reactivity (k_{inact}) of mechanism-based covalent inhibitors.[3] DynaFit software[9] was used to fit a system of differential equations to the substrate conversion curves generated by this kinetics assay. Remarkably, the inactivation constants (k_{inact}) did not follow the trend in pK_a of the heterocycles, which is contradictory to standard theories of nucleofugality. In contrast, the change in number and position of the nitrogen atoms in the heterocycle strongly influenced the binding constant K_i. The 1,2,3-triazole based DH376 bound with picomolar affinity.

The approach to determine the binding affinity and reactivity presented in Chapter 4 is based on a traditional biochemical readout. Other novel techniques, such as Surface

Plasmon Resonance (SPR) and Isothermal Calorimetry (ITC), have recently been applied for in-depth evaluation of the formation of protein-inhibitor complexes.[10] The benefit of the here presented approach is its low cost and relatively high throughput: the current setup uses 96-well plates and allows for 6 compounds to be screened per (duplicate) plate and does not require purified enzymes. The original surrogate substrate assay was successfully miniaturized to 1536-well format, so there is room to improve the current throughput of the kinetics assay.[7] The assay could also be easily adopted by other enzymes capable of hydrolysing 4-nitrophenyl butyrate or alternative surrogate substrates. This could be especially relevant for major off-targets of the current series of inhibitors, like α/β-hydrolase domain containing protein 6 (ABHD6). This will guide the discovery of more selective DAGL inhibitors, including the inhibitors of Chapter 2 and 3.

Chapter 4 provided some key insights in the importance of the heterocycle of the DH376 variants to the binding affinity with DAGL-α. It should be investigated whether this SAR can be transferred to the triazole-sulfonamide series developed in Chapters 2 and 3. This can be achieved by synthesizing variants of **1** containing a 1,2,3-triazole (**3**) or pyrazole (**4**) (Scheme 9.1), in an analogous manner as presented in Chapters 2 and 3. Testing these compounds alongside **1** in the assay developed in Chapter 4 could provide novel insights in the binding-mode of these sulfonamide based inhibitors, and they may have interestingly different interaction profiles.
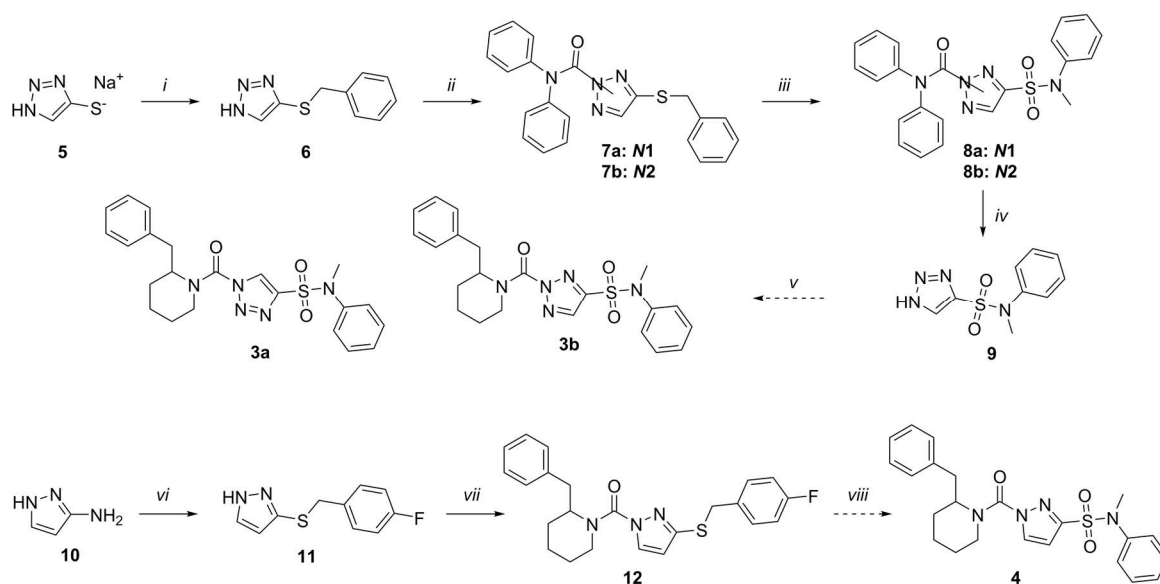
Scheme 9.1 | Synthesis of 1,2,3-triazoles (**3a** and **3b**) and pyrazole (**4**) derivatives of **1**. Reagents and conditions: *i*) benzylbromide, DMF, 18 h, 86%; *ii*) diphenylcarbamic chloride, DIPEA, DMAP, THF, 65 °C, 18 h, 71%; *iii*) HCl, NaOCl, DCM, -10 °C, 20 min, then *N*-methylaniline, -10 °C → RT, 2h, 65%; *iv*) KOH, H₂O/THF (1:1 v/v), 18 h, 57%; *v*) 2-benzylpiperidine, triphosgene, K₂CO₃, DCM, 0 °C → RT, 1 h, then **9**, DIPEA, THF, 6 h, 65 °C; *vi*) NaNO₂, H₂SO₄, H₂O, then NaOAc, NaOH, 4-fluorobenzylmercaptan, H₂O, 1 h, 0 °C, 42%; *vii*) 2-benzylpiperidine, triphosgene, DIPEA, THF, 0 °C, 1 h, then **11**, DIPEA, DMAP, THF, 65 °C, 6 h, 48%; *viii*) HCl, NaOCl, DCM, -10 °C, 20 min, then *N*-methylaniline, -10 °C → RT, 2h.

Activity-based protein profiling to study inhibitor selectivity

Chapter 5 showcased the utility of activity-based protein profiling in a highly relevant study to profile the interactions of the lethal drug-candidate BIA 10-2474, which was developed as a fatty acid amide hydrolase (FAAH) inhibitor. During the first-in-human phase 1 clinical trial held in France in early 2016, 5 members of the highest dosing group were hospitalised with severe neurological symptoms.[11] One of these volunteers succumbed to his symptoms several days after hospitalisation. The main hypothesis to explain the severe side effects was the possibility of off-target interactions. To investigate this, BIA 10-2474 was resynthesised, along with three alkyne-bearing analogues. The inhibitory capacity of BIA 10-2474 was assayed in a large panel of endocannabinoid related enzymes and both CB-receptors, and was found to only inhibit FAAH weakly *in vitro*. The interaction profile across nearly 50 serine hydrolases was assessed by ABPP using mouse brain proteomes. ABHD6 was found as an off-target at similar concentrations required for full FAAH inhibition. The covalent mechanism of action of the imidazole urea was shown by the fluorescent labeling of both FAAH and ABHD6 using 'Click'-chemistry on the alkyne analogues under denaturing SDS-PAGE conditions. *In situ* profiling of the inhibitors showed a large increase in potency for both FAAH and ABHD6, not observed for the safe FAAH inhibitor PF-04457845.[12] Using chemical proteomic profiling of *in situ* treated human neural cell cultures with BIA 10-2474 more off-targets were identified, which were subsequently verified by overexpression and competitive ABPP. Of note, these off-targets were not shared with PF-04457845, a safe FAAH inhibitor tested in phase 2 clinical trials, and were all involved in lipid metabolism. Finally, lipid analysis using targeted lipidomics showed a strong disruption of lipid networks in cultured neurons. This suggests that promiscuous lipase inhibitors have the potential to cause metabolic dysregulation in the nervous system. The importance of thorough off-target profiling, possible using ABPP, is emphasized and recommended.

None of the reported animal models used during pre-clinical characterization showed toxicity similar to that observed in humans.[13] Therefore, the assays performed here on mouse brain lysates and (neural) cell cultures should be repeated on actual patient-derived brain tissue samples. However, a call for access to the existing patient-material, such as blood samples from the trial, including those of unharmed volunteers, has not been followed up on by the sponsor of the trial, Bial Pharmaceuticals, thus making any further research in this direction infeasible.[14]

Another avenue of exploration could be the identification of protein interaction partners of BIA 10-2474 that were not targeted by activity-based probes used in the current study. To this end, chemical probes were generated based on the structure of BIA 10-2474, thereby creating an ABP closely related to the parent molecule. These probes, however, did not bind other targets than visualized in the competitive gel-based ABPP experiments with FP-TAMRA and MB064. The limitation of this approach is that the targets have to bind covalently to the urea electrophile to be visualized. Non-covalent interactions with proteins cannot be identified in this manner. Radioactively labelled BIA 10-2474 could aid in finding binding partners, but identification of specific proteins would be cumbersome. CETSA could be used, but will only reveal targets that are significantly thermally stabilized upon binding of BIA 10-2474. Photoaffinity labelling, combined with chemical biology methods, can be

used to identify non-covalent interaction partners.[15,16] This approach requires a photoreactive group to be present in the molecule. Typically, diazirines or benzophenones are used, but many other photoreactive groups exist.[15] An intriguing possibility arises in the case of BIA 10-2474, since the photoreactivity of heterocyclic *N*-oxides has been studied for decades for use in synthetic applications, but, to the best of my knowledge, has never been applied in a chemical biology setting.[17,18] Irradiation with ultraviolet light triggers a sequence of rearrangements, leading, in controlled conditions, to various reaction products.[17] No information is available regarding the reactivity of these species in close proximity to proteins. Experimental validation of this approach could consist of photolabelling experiments where the three alkynylated probes are irradiated. Incubation with or without BIA 10-2474 or the reduced pyridine metabolite BIA 10-2639, which should not be UV-active, can help to identify specific interactions. Conjugation to biotin using 'Click'-chemistry and consecutive enrichment for chemical proteomics should show the applicability of this photoreactive group in chemical biology. If successful, the abundance of pyridines and pyrimidines in drug-like molecules could open up a whole new field of target identification using heterocyclic *N*-oxides as photolabel. Finally, by applying computational target prediction algorithms on the BIA 10-2474 structure the amount of targets that need to be considered could be reduced, or at least the order of exploration could be optimized.
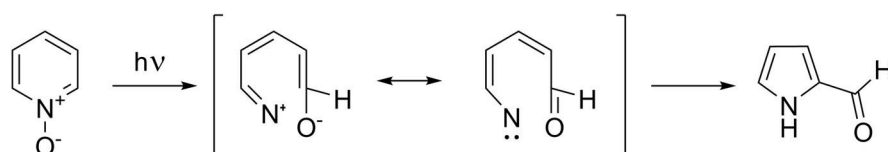
Scheme 9.2 | Proposed intermediates in the photoreaction of pyridine *N*-oxides.[17] The mechanism towards the bracketed transition state is not clear, and is believed to occur at a hyperplane.[18]

Chapter 6 described the development and in-depth characterisation of a green-fluorescent fluorophosphonate-probe, BODIPY-FP, an analogue of the widely used red-fluorescent TAMRA-fluorophosphonate. This new probe can be used in conjunction with the tailored DAG-lipase probe MB064, which carries a red-fluorescent BODIPY-TMR as reporter tag. The labeling efficiency of various enzymes with the new BODIPY-FP was found to be significantly different from TAMRA-FP. Especially the potency for FAAH and monoacyl-glycerol lipase (MAGL) was strongly increased. The probe was mixed at low concentration (10 nM) with MB064, thereby reliably visualising FAAH and MAGL without interfering with the labeling of MB064-targets. This probe cocktail allowed visualisation of most endocannabinoid related serine hydrolases in a single sample. This multiplex ABPP method was used to profile the selectivity of two MAGL inhibitors, NF1819 and ABX-1431.[19,20] The former was found to have several off-targets, including ABHD6. ABX-1431 showed no off-target activity up to 10 µM in mouse brain membrane proteome.

The combination of multiple probes with different reactive groups, also known as warheads, and orthogonal reporter tags, allows profiling of more proteins in a single sample. Especially in gel-based techniques broad reactivity of an ABP may lead to

overlapping protein bands, which hinders the quantification and identification. Since MB064 and BODIPY- or TAMRA-FP combined are not able to provide full coverage of the serine hydrolase family, the search for new activity-based probes is warranted. An interesting reactive group which has long been used in lipase inhibitors is the bromoenol lactone (BEL).[21] Previously, the introduction of a ligation handle in the form of an alkyne was proposed, but the synthesis proved cumbersome.[22] An alternative approach was envisioned based on a published synthetic procedure towards guanidine-containing protease inhibitors (Scheme 9.3).[23] This synthetic route furbished BEL probe **13** and a closely related inhibitor **14**.

Initial experiments using probe **13** *in situ* in a panel of breast cancer cell lines indicated that the BEL warhead is indeed able to covalently label several proteins in these cells in a dose-dependent manner (Figure 9.2). Most but not all bands could be outcompeted with pre-incubation of **14**, indicating specific targeting of several proteins. Chemical proteomics is required to identify the probe-targets, and to compare its interaction profile with fluorophosphonate-based probes.
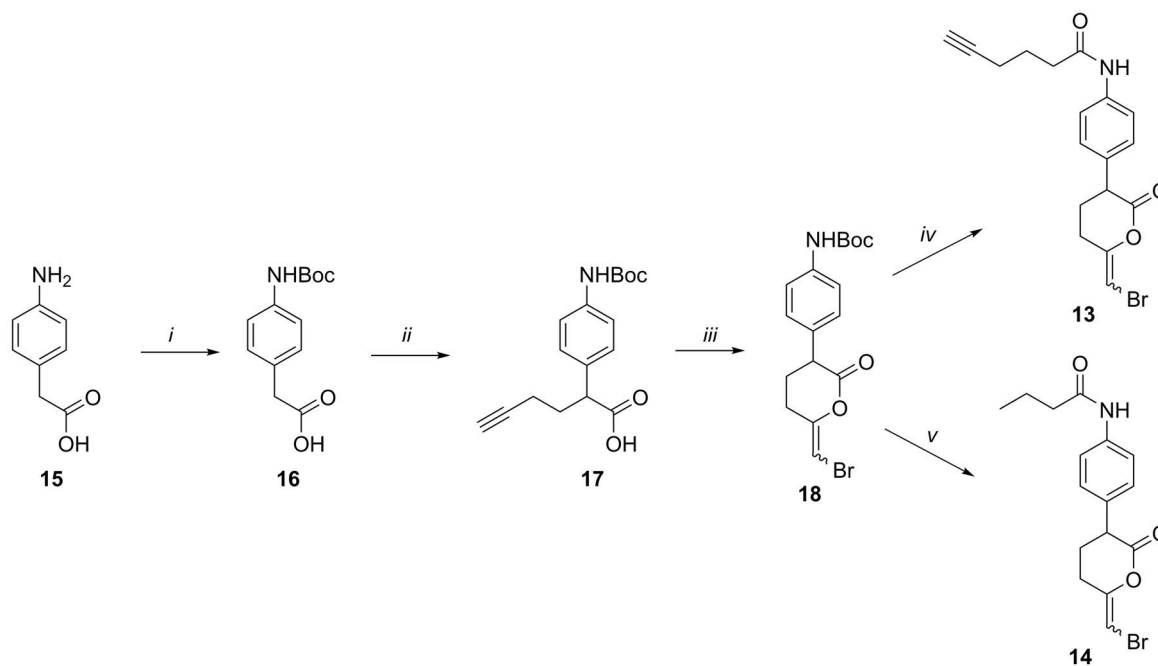
Scheme 9.3 | Synthesis of bromoenol lactone probe **13** and closely related inhibitor **14**. Reagents and conditions: *i*) Boc$_2$O, DCM, 18 h, 77%; *ii*) BuLi, THF, -30 °C, 30 min, then 4-bromobut-1-yne, -20 °C → 0 °C, 3 h, 40%; *iii*) K$_2$CO$_3$, DCM, then NBS, H$_2$O, 5 h, 22%; *iv*) TFA, DCM, 0 °C → RT, 1.5 h, then hex-6-ynoyl chloride, TEA, 0 °C 2 h, 53%; *v*) TFA, DCM, 0 °C → RT, 1.5 h, then butyryl chloride, TEA, 0 °C, 2 h, 63%.
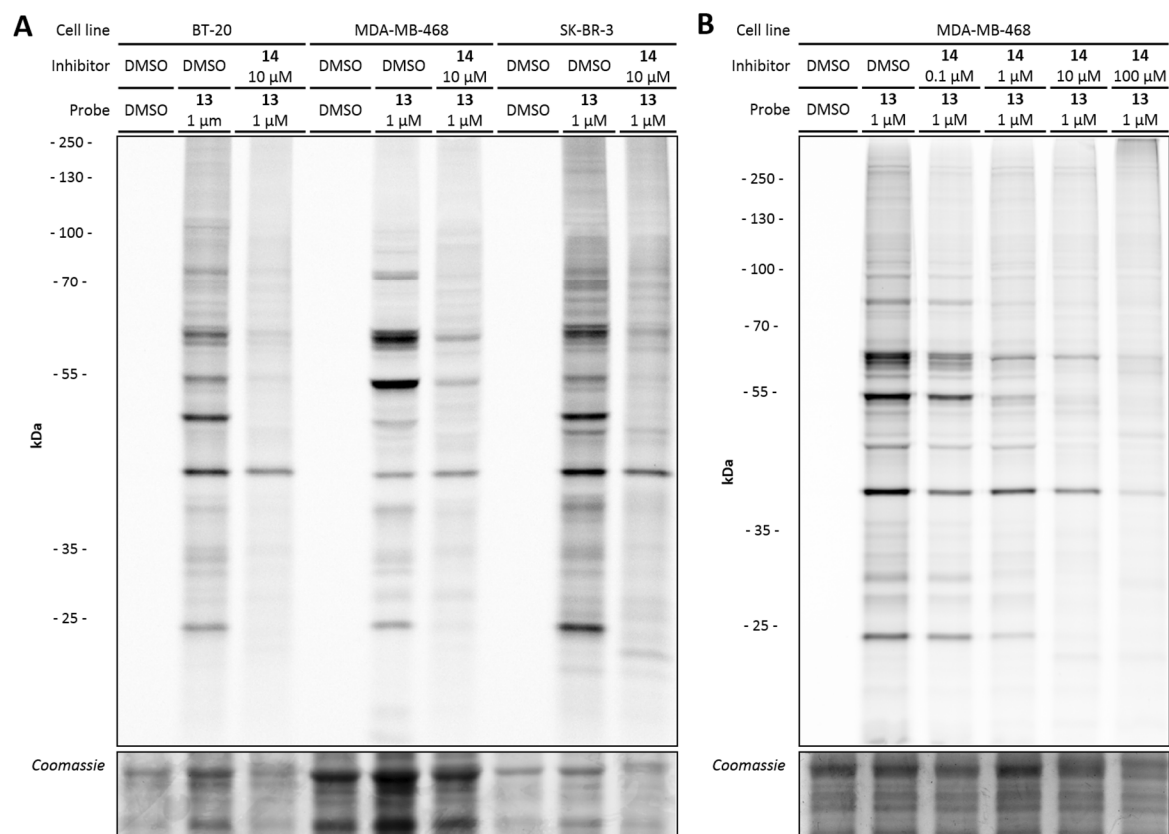
Figure 9.2 | Two-step labelling of breast cancer cell lines with BEL probe **13**, competed with a single dose (A) or a dose escalation (B) of inhibitor **14**. Coomassie staining is shown as protein loading control.

Machine learning as a tool to predict inhibition profiles

Chapter 7 introduced the application of the machine learning algorithm t-Distributed Stochastic Neighbour Embedding (t-SNE) as a similarity metric for (bio)molecules. In contrast to the commonly used Tanimoto coefficient (Tc), a bit string compatible distance metric, t-SNE generates a two dimensional visualisation of the high-dimensional bit string space, to some extent analogous to principal component analysis, which was shown to be outperformed by t-SNE in this context. t-SNE allows for a visual inspection of similar chemical matter, which is impossible when using the Tc. This similarity visualisation was applied to 2774 clinically used molecules as a demonstration of drug-like chemical space. This embedding showed strong co-localisation of chemically related compounds. t-SNE was further applied to a physicochemical bit string based on the alignment of the proteins of the serine hydrolase family. This revealed that t-SNE is capable of recapitulating phylogenetic information. The mapping of target-space was annotated with the inhibitory potencies of a diverse set of compounds, previously measured in an ABPP-based assay.[24] Despite the small datasets available for serine hydrolases, co-localisation of potently inhibited proteins could be observed, suggesting that this approach could be extended to predict affinities of inhibitors across an entire protein family.

Chapter 8 built upon the principles introduced in Chapter 7, but focussed on the large protein family of kinases. Very large screening sets are available for kinases, which facilitates machine learning approaches such as that described in Chapter 7. The protein

kinase space could be reliably mapped using t-SNE. This resulted in a very strong similarity between the previously published Kinase Family Tree[25] and the t-SNE map. The molecules of the Published Kinase Inhibitor Set, combined with the biological data available for ~200 kinases, provided a large training set to be used in the here developed model, which was named Drug Discovery Maps (DDM).[26,27] This model was validated using the data from Karaman *et al*. and shown to equal or outperform the state-of-the-art computational techniques, such as quantitative structure activity relations (QSAR) and proteochemometric modelling.[28,29] DDM was applied in a virtual high-throughput assay in search for new FMS-like tyrosine kinase 3 (FLT3) inhibitors. A positive predictive value of 41% was achieved, identifying 18 actives out of a predicted 44. Two of these hits were resynthesised and tested in MV4:11 cells. Both compounds were cellular active and the best compound showed single-digit nanomolar potency. The predicted off-targets were validated and around 70% of the predicted targets were found to be inhibited. The model was also used to predict the interaction profile of a commercially available set of compounds with a hinge-binding motif. 20 molecules were purchased and tested for their predicted activities against HIPK3 and PAK2. Unfortunately, no actives were identified amongst the purchased compounds, possibly due to a low coverage of the training data around these targets, or poor similarity between the molecules in the training set and the commercial library.

The workflow presented in Chapter 8 is flexible regarding the training data that is used in the t-SNE algorithm. Since the model is in essence retrained for every run, simply loading a different list of molecules with associated biological data will provide a new predictive model, which can then be optimised by tuning the hyperparameters. Implementing a larger training set, such as that described recently by Bembenek *et al*. could strongly improve the predictive capacities of the model, although perhaps hierarchical SNE should be used to avoid excessive computational times.[30,31] Alternatively, a completely different set of molecular data could be used, such as a binding dataset for the GPCRs, or different processes altogether, e.g. metabolic data, lipophilicity data or (cyto-)toxicity.

An interesting but challenging possibility of further exploration is the application of the concept of cross-modal, also known as multimodal, clustering.[32] This autonomous classification technique was designed to use two or more unrelated sensory inputs from the same observation, which can be interpreted in the broadest manner, to classify unannotated data points into an *a priori* unknown number of clusters. Cross-modal coupling could thus, theoretically, be implemented to use two or more t-SNE embeddings of the same set of entities but with different modalities. One set of modalities could be different settings in the t-SNE perplexity, which should translate to embeddings using a different 'resolution', shifting the focus of the embedding from more high-level structure (high perplexity) to subtler local differences (low perplexity), and combining those views in order to gain insight in similarities not immediately evident from either setting separately. Another promising direction would be to use the clustering of inhibitors based on (normalized) biological data in one embedding, and on a chemical fingerprint in another. Different molecular clusters, possibly far separated in chemical space, might show similar trends in inhibition profile. These effects could be picked up by cross-modal clustering, and could give a whole new aspect to the use of t-SNE embeddings in drug discovery.

Closing remarks

Less than 1 in 10 drug candidates that enter phase 1 clinical trials actually gets approved for human use. The high failure rate is in part due to unforeseen side effects or toxicity. A better understanding of the role of selectivity and a better insight in the off-target activities of drug candidates could greatly aid in preventing candidates to fail for these reasons. This thesis has tried to address some aspects in this challenging part of drug discovery. The use of activity-based protein profiling as presented in Chapters 2 and 3 in drug discovery and hit-to-lead optimization, and in Chapter 5 and 6 for the interaction profiling of a drug candidate, highlights the versatility and importance of this chemical biology technique. Combined with knowledge derived from biochemical assays, such as that developed in Chapter 4, ABPP can greatly aid the medicinal chemist. The recent surge in popularity of machine learning algorithms, backed by exponential growth of the amount of biological data available, holds great promise for drug discovery. Chapters 7 and 8 showed the applicability of one such algorithm, which was able to quite reliably predict interaction profiles. The challenges in finding, determining and predicting selectivity are far from solved, but, by incrementally expanding our understanding of the binding of small molecules to their (off-)targets, truly selective inhibitors might at some point become a reality or their necessity might be mitigated.

Acknowledgements

Methods

Chemical Biology Methods

Cell Culture

BT-20, MDA-MB-468 or SK-BR-3 cells were cultured at 37 °C under 5% CO_2 in RPMI medium containing 10% (v/v) Fetal Calf Serum (Thermo Fisher), penicillin and streptomycin (200 µg/mL each; Duchefa). Medium was refreshed every 2-3 days and cells were passaged twice a week at 80-90% confluence by trypsinisation and resuspension in fresh medium.

Cells lines were purchased from ATCC and were regularly tested for mycoplasma contamination. Cultures were discarded after 2-3 months of use.

In situ treatment of breast cancer cell lines

BT-20, MDA-MB-468 or SK-BR-3 cells were seeded in 6-well plates. *In situ* treatment was initiated 72 h later. Medium was aspirated and medium with serum containing inhibitor or DMSO as vehicle was added (0.1 % DMSO). After 1 h exposure to treatment medium at 37 °C, probe or DMSO as vehicle was added and incubated for 1 h at 37 °C. The medium was aspirated and cells were harvested, washed with PBS and flash frozen in liquid nitrogen and stored at -80 °C until further use.

Cell lysate preparation

Cell pellets were thawed on ice, resuspended in cold lysis buffer (2.5 U/mL benzonase and 1x protease inhibitor (Amresco) in Milli-Q water) and incubated on ice (15-30 min). Protein concentrations were determined by a Quick Start™ Bradford Protein Assay. The cell lysate was diluted to 1.0 mg/mL concentration in lysis buffer and diluted samples were flash frozen in liquid nitrogen and stored at -80 °C until further use.

Fluorescent ligation and gel-based analysis

The two-step labeling protocol was adapted from previously developed methods.[33] Click reagent was freshly prepared by mixing copper sulfate (25 µL, 18 mM in water), THPTA (5 µL, 18 mM in water), sodium ascorbate (15 µL, 150 mM in water), and AF-647 (5 µL, 90 µM in DMSO). Click reagent (5.0 µL) was added to each proteome (40 µL per condition), mixed by brief vortexing and incubated under continuous shaking (1 h, 37 °C). 15 µL of 4* Laemmli buffer was added. Samples were boiled (95 °C, 5 min) and 8 µg proteome per reaction (12 µL) was resolved on a 10% acrylamide SDS-PAGE gel (180 V, 75 min). Gels were scanned using Cy3 and Cy5 multichannel settings (605/50 and 695/55, filters respectively) and stained with Coomassie after scanning. Coomassie staining was performed as a protein loading control.

Synthetic Methods

General remarks

All reactions were performed using oven- or flame-dried glassware and dry solvents. Reagents were purchased from Alfa Aesar, Sigma-Aldrich, Acros, and Merck and used without further purification unless noted otherwise. All moisture sensitive reactions were performed under an argon or nitrogen atmosphere. ^1H and ^{13}C NMR spectra were recorded on a Bruker AV-300 (300 MHz), AV-400 (400 MHz) or DRX-500 (500 MHz). Used software for interpretation of NMR-data was Bruker TopSpin 1.3 and MestreNova 11.0. Chemical shift values are reported in ppm with tetramethylsilane or solvent resonance as the internal standard ($CDCl_3$: δ 7.26 for ^1H, δ 77.16 for ^{13}C; MeOD: δ 3.31 for ^1H, δ 49.00 for ^{13}C).[34] Data are reported as follows: chemical shifts (δ), multiplicity (s = singlet, d = doublet, dd = double doublet, td = triple doublet, t = triplet, q = quartet, quintet = quint, bs = broad singlet, m = multiplet), coupling constants J (Hz), and integration.

Liquid chromatography analysis was performed on a Finnigan Surveyor LC/MS system, equipped with a C18 column. Flash chromatography was performed using SiliCycle silica gel type SiliaFlash P60 (230–400 mesh). TLC analysis was performed on Merck silica gel 60/Kieselguhr F254, 0.25 mm. Compounds were visualized using $KMnO_4$ stain (K_2CO_3 (40 g), $KMnO_4$ (6 g), and water (600 mL)) or CAM stain ($Ce(NH_4)_4(SO_4)_4 \cdot 2H_2O$ (ceric ammonium sulfate: 10 g); ammonium molybdate (25 g); conc. H_2SO_4 (100 mL); H_2O (900 mL)). Preparative HPLC (Waters, 515 HPLC pump M; Waters, 515 HPLC pump L; Waters, 2767 sample manager; Waters SFO System Fluidics Organizer; Waters Acquity Ultra Performance LC, SQ Detector; Waters Binary Gradient Module) was performed on a Waters XBridgeTM column (5 µM C18, 150 x 19 mm). Diode detection was done between 210 and 600 nm. Gradient: ACN in (H_2O + 0.2% TFA). High resolution mass spectra (HRMS) were recorded by direct injection on a q-TOF mass spectrometer (Synapt G2-Si) equipped with an electrospray ion source in

positive mode with Leu-enkephalin (m/z = 556.2771) as an internal lock mass. The instrument was calibrated prior to measurement using the MS/MS spectrum of Glu-1-fibrinopeptide B.

4-(Benzylthio)-1*H*-1,2,3-triazole (**6**)
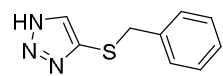
Sodium 1*H*-1,2,3-triazole-4-thiolate (1.23 g, 10.0 mmol) was dissolved in DMF (25 mL) and benzyl bromide (1.31 mL, 11.0 mmol) was added. The mixture was stirred for 18 h. The volatiles were removed under reduced pressure. The residue was purified by column chromatography to yield the title compound as an off-white solid (1.65 g, 8.64 mmol, 86%). ^{1}H NMR (300 MHz, CDCl$_3$) δ 14.54 (s, 1H), 7.48 (s, 1H), 7.21 – 7.05 (m, 5H), 4.04 (s, 2H). ^{13}C NMR (75 MHz, CDCl$_3$) δ 138.32, 137.00, 131.99, 128.73, 128.38, 127.26, 39.51.

4-(Benzylthio)-*N,N*-diphenyl-1,2,3-triazole-carboxamide (**7**)
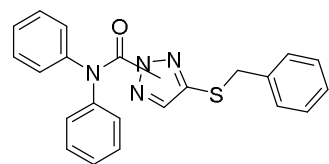
6 (1.248 g, 6.53 mmol) was dissolved in dry THF (30 mL) and diphenyl-carbamic chloride (1.66 g, 7.18 mmol), DIPEA (1.37 mL, 7.83 mmol) and a catalytic amount of DMAP were added. The mixture was refluxed for 18 h. The reaction was quenched by the addition of sat. aq. Na$_2$CO$_3$ (40 mL). The organic layer was separated and the aqueous phase extracted with EtOAc (2x 40 mL). The combined organic layers were washed with brine (40 mL), dried (MgSO$_4$), filtered and concentrated. Column chromatography afforded the title compound as an off-white solid, consisting of a mixture of 2 regioisomers (2:1) (1.79 g, 4.62 mmol, 71%). Regioisomer 1: ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.47 – 6.60 (m, 16H), 3.84 (s, 2H). Regioisomer 2: ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.47 – 6.59 (m, 16H), 3.69 (s, 2H) (spectrum shows a mixture of rotamers). ^{13}C NMR (101 MHz, CDCl$_3$) δ 151.18, 141.32, 137.43, 136.75, 133.21, 130.15, 129.68, 129.55, 129.00, 128.66, 128.57, 128.50, 127.32, 127.24, 127.21, 127.18, 126.95, 107.18, 106.05, 39.01, 37.26.

### 4-(*N*-Methyl-*N*-phenylsulfamoyl)-*N,N*-diphenyl-1,2,3-triazole-carboxamide (**8**)
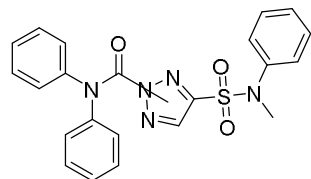
4 N HCl in dioxane (5 mL) was added to DCM (20 mL) and cooled to -10 °C. 15%$_{wt}$ NaOCl (aq., 5 mL) was added slowly, forming a yellow-green Cl$_2$ solution. After 30 min **7** (500 mg, 0.259 mmol) was added dropwise as a solution in DCM (5 mL). After 20 min *N*-methylaniline (5.0 mL, 46 mmol) was added. The mixture was stirred for 2 h at RT. The reaction was quenched by the addition of 0.1 M HCl (20 mL). The organic layer was separated and the water layer was extracted with EtOAc (2x 20 mL). The combined organic layers were dried (MgSO$_4$), filtered and concentrated. The residue was purified by column chromatography to yield the title compound as a brown solid (364 mg, 0.840 mmol, 65%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.58 (s, 1H), 7.45 – 6.79 (m, 15H), 3.01 (s, 3H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 147.92, 146.72, 141.68, 140.04, 136.10, 129.50, 129.05, 127.98, 127.61, 126.70, 126.28, 38.67.

N-Methyl-*N*-phenyl-1*H*-1,2,3-triazole-4-sulfonamide (**9**)
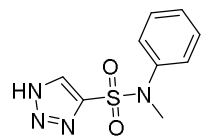
8 (364 g, 0.84 mmol) was dissolved in 1 N KOH in a 1:1 (v/v) H$_2$O/THF mixture (20 mL). The mixture was stirred for 18 h at RT. The mixture was concentrated and the residue was purified by column chromatography to yield the title compound (114 mg, 0.48 mmol, 57%). ^{1}H NMR (400 MHz, MeOD) δ 8.06 (s, 1H), 7.35 – 7.09 (m, 5H), 3.32 (s, 3H). ^{13}C NMR (101 MHz, MeOD) δ 145.60, 142.37, 130.68, 130.07, 128.78, 127.98, 39.25.

3-((4-Fluorobenzyl)thio)-1*H*-pyrazole (**11**)
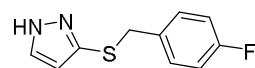
1*H*-Pyrazol-3-amine **10** (100 mg, 1.20 mmol) was dissolved in 40% aq. H$_2$SO$_4$ (5 mL). Sodium nitrite (100 mg, 1.44 mmol) in water (5 mL) was added slowly, keeping the temperature below 5 °C. After 15 minutes saturated aqueous NaOAc was added until the pH was ~5. The mixture was then added slowly to a cooled (0 °C) solution of (4-fluoro-phenyl)methanethiol (118 µL, 137 mg, 0.963 mmol) in 1 N aqueous NaOH (1 mL). A precipitate formed immediately upon addition. After addition was complete, the mixture was stirred for 1 h at 0 °C. The reaction mixture was extracted with EtOAc (3x 20 mL). The combined organic layers were washed with brine, dried (MgSO$_4$), filtered and concentrated. Column chromatography yielded the title compound (105.7 mg, 0.508 mmol, 42%) as a white solid. ^{1}H NMR (400 MHz, MeOD) δ 7.54 (d, *J* = 2.2 Hz, 1H), 7.25 – 7.15 (m, 2H), 6.97

– 6.89 (m, 2H), 6.20 (d, J = 2.2 Hz, 1H), 4.02 (s, 2H). ^{13}C NMR (101 MHz, MeOD) δ 163.31 (d, J = 244.1 Hz), 135.44, 131.65 (d, J = 8.1 Hz), 115.94 (d, J = 21.7 Hz), 109.97, 39.60.

(2-Benzylpiperidin-1-yl)(3-((4-fluorobenzyl)thio)-1*H*-pyrazol-1-yl)methanone (**12**)
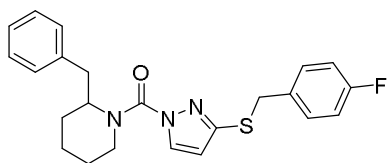
2-Benzylpiperidine hydrochloride (50.8 mg, 0.240 mmol) and triphosgene (71.2 mg, 0.240 mmol) were dissolved in dry THF (10 mL) at 0 °C. DIPEA (194 µL, 1.2 mmol) was added and the resulting mixture was stirred for 1 h. The mixture was poured into ice water (10 mL) and the organic layer was separated. The aqueous layer was extracted with EtOAc (3x 10 mL). The combined organic layers were dried (MgSO₄), filtered and concentrated under reduced pressure. The residue was taken up in dry THF (10 mL) to which **11** (50 mg, 0.24 mmol), DMAP (29.3 mg, 0.24 mmol) and DIPEA (194 µL, 1.2 mmol) were added. The mixture was heated to reflux for 6 h, after which it was poured into a saturated aqueous NH₄Cl solution (10 mL). The organic layer was separated and the aqueous layer was extracted with EtOAc (2x 10 mL). The combined organic layers were dried (MgSO₄), filtered and concentrated. Column chromatography afforded the title compound (47.3 mg, 0.116 mmol, 48%) as a white solid. ^{1}H NMR (400 MHz, CDCl₃) δ 7.64 (s, 1H), 7.38 – 6.89 (m, 9H), 6.11 (d, J = 2.8 Hz, 1H), 4.87 (s, 1H), 4.33 – 4.25 (m, 1H), 4.21 (t, J = 12.8 Hz, 2H), 3.28 – 2.97 (m, 2H), 2.84 (t, J = 10.7 Hz, 1H), 1.88 – 1.53 (m, 6H).
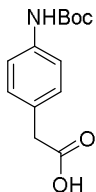
2-(4-((*tert*-Butoxycarbonyl)amino)phenyl)acetic acid (**16**)

2-(4-Aminophenyl)acetic acid **15** (1.0 g, 6.62 mmol) was dissolved in DCM (20 mL). Boc-anhydride (1.54 mL, 6.62 mmol) was added and the mixture was stirred for 18 h at RT. The mixture was concentrated under reduced pressure. The residue was purified by column chromatography. The title compound was obtained as an orange solid (1.28 g, 5.09 mmol, 77%). ^{1}H NMR (400 MHz, MeOD) δ 8.75 (s, 1H), 7.34 (d, J = 8.2 Hz, 2H), 7.20 – 7.11 (m, 2H), 3.52 (s, 2H), 1.50 (s, 9H). ^{13}C NMR (101 MHz, MeOD) δ 175.67, 155.16, 139.14, 129.89, 119.81, 80.70, 41.18, 28.69.

2-(4-((*tert*-Butoxycarbonyl)amino)phenyl)hex-5-ynoic acid (**17**)
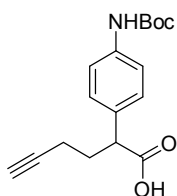
16 (620 mg, 2.47 mmol) was dissolved in THF (25 mL) and cooled to -30 °C, butyllithium (1.6 M in pentane, 5.09 mL, 8.14 mmol) was added dropwise. The mixture was stirred for 30 min. 4-Bromobut-1-yne (0.764 mL, 8.14 mmol) was slowly added and the mixture stirred for 2 h at -20 °C. The mixture was left to warm up to 0 °C over the course of 1 h. The reaction was quenched by the addition of ice. The organic layer was separated. The waterlayer was acidified to pH ~ 2 with 2 N HCl (aq.) and extracted with EtOAc (3x 50 mL). The combined organic layers were dried (MgSO₄), filtered and concentrated. The residue was purified using flash column chromatography. The desired product was obtained as a white powder (302 mg, 0.996 mmol, 40%). ^{1}H NMR (400 MHz, CDCl₃) δ 9.71 (bs, 1H), 7.41 – 7.06 (m, 4H), 6.74 (bs, 1H), 3.82 – 3.64 (m, 1H), 2.36 – 2.14 (m, 2H), 2.14 – 1.86 (m, 3H), 1.51 (s, 9H). ^{13}C NMR (101 MHz, CDCl₃) δ 179.44, 137.86, 132.24, 128.75, 119.26, 83.05, 69.62, 49.29, 31.36, 28.38, 16.28.

tert-Butyl (4-(6-(bromomethylene)-2-oxotetrahydro-2*H*-pyran-3-yl)phenyl)carbamate (**18**)
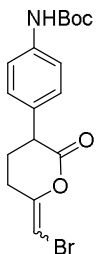
17 (150 mg, 0.494 mmol) was dissolved in dry DCM (7 mL). K₂CO₃ (41 mg, 0.30 mmol) was added to the solution. The resulting suspension was degassed for 15 min in a sonicator under continuous N₂-flow. *N*-Bromosuccinimide (97 mg, 0.54 mmol) and water (36 µL, 2 mmol) were added and the mixture was stirred for 5 h. The reaction mixture was quenched by the addition of Na₂S₂O₃ (5% aq., 10 mL). The organic layer was separated and diluted with DCM (10 mL), washed with H₂O (10 mL) and brine (5 mL). The combined organic layers were dried (MgSO₄), filtered and concentrated. The residue (yellow oil) was purified by column chromatography to yield the title compound as a white solid (42 mg, 0.11 mmol, 22%). ^{1}H NMR (400 MHz, CDCl₃) δ 7.44 – 7.30 (m, 2H), 7.18 – 7.07 (m, 2H), 6.60 (s, 1H), 6.12 – 6.03 (m, 1H), 3.74 (dd, J = 10.4, 5.4 Hz, 1H), 2.89 – 2.58 (m, 2H), 2.29 – 2.06 (m, 2H), 1.51 (s, 9H). ^{13}C NMR (101 MHz, CDCl₃) δ 168.76, 152.81, 151.62, 138.08, 131.42, 128.65, 118.98, 91.53, 80.79, 46.72, 28.42, 25.87, 23.84.

Hex-5-ynoyl chloride (19)
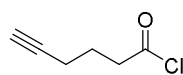
Hex-5-ynoic acid (0.197 mL, 1.78 mmol) was dissolved in DCM (10 mL) and oxalyl chloride (1.56 mL, 17.8 mmol) was slowly added. The mixture was refluxed for 2 h. The solvent and excess oxalyl chloride were removed *in vacuo* and the resulting oil was coevaporated twice with toluene. The crude hex-5-ynoyl chloride was used without further purification.

N-(4-(6-(bromomethylene)-2-oxotetrahydro-2*H*-pyran-3-yl)phenyl)hex-5-ynamide (13)
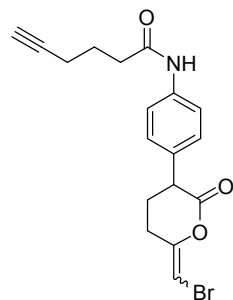
To a stirred solution of **18** (30 mg, 0.078 mmol) in DCM (3 mL) at 0 °C trifluoroacetic acid (0.302 mL, 3.92 mmol) was added. The mixture was stirred for 1.5 h at RT. **19** (102 mg, 0.785 mmol) was added and the mixture was cooled to 0 °C. TEA (0.60 mL, 4.3 mmol) was added and the mixture was stirred for 2 h at RT. The volatiles were evaporated under reduced pressure. Flash column chromatography yielded the title compound (15.6 mg, 0.041 mmol, 53%) as a white solid. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.58 – 7.45 (m, 2H), 7.40 (s, 1H), 7.21 – 7.11 (m, 2H), 6.11 (t, J = 1.8 Hz, 1H), 3.75 (dd, J = 10.6, 5.4 Hz, 1H), 2.96 – 2.81 (m, 1H), 2.76 – 2.64 (m, 1H), 2.51 (t, J = 7.3 Hz, 2H), 2.35 – 2.29 (m, 2H), 2.29 – 2.08 (m, 2H), 2.02 (t, J = 2.6 Hz, 1H), 1.94 (q, J = 7.3 Hz, 2H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 170.74, 168.78, 151.55, 137.53, 132.76, 128.71, 120.31, 91.71, 83.55, 69.60, 46.88, 36.04, 25.89, 24.00, 23.94, 17.90. HRMS: Calculated for [C$_{18}$H$_{18}$BrNO$_3$ + H]$^+$ = 376.0542, found = 376.0548.

N-(4-(6-(Bromomethylene)-2-oxotetrahydro-2*H*-pyran-3-yl)phenyl)butyramide (14)
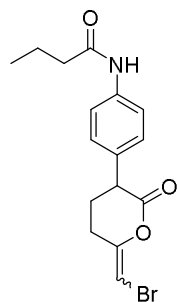
18 (10 mg, 0.026 mmol) was dissolved in DCM (1 mL) and trifluoroacetic acid (101 µL, 1.31 mmol) was added. The mixture was stirred for 1.5 h at RT, after which TLC showed full conversion of starting material. Butyryl chloride (27.1 µL, 0.262 mmol) was added and the mixture was cooled to 0 °C. Triethylamine (201 µL, 1.44 mmol) was added and the mixture stirred for 2 h at RT. The reaction mixture was concentrated and purified by column chromatography to yield the title compound (5.81 mg, 0.016 mmol, 63%) as a white solid. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.52 (d, J = 8.4 Hz, 2H), 7.21 – 7.11 (m, 3H), 6.11 (t, J = 1.8 Hz, 1H), 3.76 (dd, J = 10.5, 5.4 Hz, 1H), 2.93 – 2.81 (m, 1H), 2.77 – 2.63 (m, 1H), 2.34 (t, J = 7.4 Hz, 2H), 2.29 – 2.06 (m, 2H), 1.76 (h, J = 7.4 Hz, 2H), 1.01 (t, J = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 168.69, 151.57, 137.63, 132.69, 128.74, 120.26, 91.69, 46.87, 39.81, 25.91, 23.92, 19.18, 13.88.
HRMS: Calculated for [C$_{16}$H$_{18}$BrNO$_3$ + H]$^+$ = 352.0542, found = 352.0544.

References

1. Croston, G. E. The utility of target-based discovery. *Expert Opin. Drug Discov.* **12**, 427–429 (2017).
2. Bowes, J. *et al.* Reducing safety-related drug attrition: the use of in vitro pharmacological profiling. *Nat. Rev. Drug Discov.* **11**, 909–922 (2012).
3. 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).
4. Baggelaar, M. P. *et al.* Highly Selective, Reversible Inhibitor Identified by Comparative Chemoproteomics Modulates Diacylglycerol Lipase Activity in Neurons. *J. Am. Chem. Soc.* **137**, 8851–8857 (2015).
5. 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).
6. Baggelaar, M. P. *et al.* Chemical Proteomics Maps Brain Region Specific Activity of Endocannabinoid Hydrolases. *ACS Chem. Biol.* **12**, 852–861 (2017).
7. Janssen, F. J. Discovery of sulfonyl-1,2,4-triazole ureas as DAGLα inhibitors by HTS-ABPP. in *Discovery of novel inhibitors to investigate diacylglycerol lipases and α/β-hydrolase domain 16A* 109–139 (2016).
8. Deng, H. *et al.* Triazole Ureas Act as Diacylglycerol Lipase Inhibitors and Prevent Fasting-Induced Refeeding. *J. Med. Chem.* **60**, 428–440 (2017).
9. Kuzmič, P. DynaFit—a software package for enzymology. *Methods Enzymol.* **467**, 247–80 (2009).
10. Narayan, K. & Carroll, S. S. SPR Screening. in *Applied Biophysics for Drug Discovery* 93–105 (John Wiley & Sons, Ltd, 2017). doi:10.1002/9781119099512.ch6
11. Kerbrat, A. *et al.* Acute Neurologic Disorder from an Inhibitor of Fatty Acid Amide Hydrolase. *N. Engl. J. Med.* **375**, 1717–1725 (2016).
12. Johnson, D. S. *et al.* Discovery of PF-04457845: A Highly Potent, Orally Bioavailable, and Selective Urea FAAH Inhibitor. *ACS Med. Chem. Lett.* **2**, 91–96 (2011).
13. Bégaud, B. *et al.* Report by the Temporary Specialist Scientific Committee (TSSC), 'FAAH (Fatty Acid Amide Hydrolase)', on the causes of the accident during a Phase 1 clinical trial. 1–28 (2016).
14. Brøsen, K., Funck-Brentano, C., Kroemer, H. K., Pirmohamed, M. & Schwab, M. Open letter on access to the BIA 10-2474 clinical trial data. *Lancet* **375**, 1788–89 (2016).
15. Smith, E. & Collins, I. Photoaffinity labeling in target- and binding-site identification. *Future Med. Chem.* **7**, 159–183 (2015).
16. 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).
17. Lohse, C., Hagedorn, L., Albini, A. & Fasani, E. Photochemistry of pyridine-oxides. *Tetrahedron* **44**, 2591–2600 (1988).
18. Poole, J. S. Recent Advances in the Photochemistry of Heterocyclic N-Oxides and Their Derivatives. in *Heterocyclic N-Oxides* (ed. Oleg V. Larionov) 111–151 (Springer, 2017). doi:10.1007/7081_2017_4
19. Cisar, J. S. *et al.* Identification of ABX-1431, a Selective Inhibitor of Monoacylglycerol Lipase and Clinical Candidate for Treatment of Neurological Disorders. *J. Med. Chem.* **61**, 9062–9084 (2018).
20. Brindisi, M. *et al.* Development and Pharmacological Characterization of Selective Blockers of 2-Arachidonoyl Glycerol Degradation with Efficacy in Rodent Models of Multiple Sclerosis and Pain. *J. Med. Chem.* **59**, 2612–2632 (2016).
21. Daniels, S. B., Cooney, E., Sofia, M. J., Chakravarty, P. K. & Katzenellenbogen, J. A. Haloenol lactones. Potent enzyme-activated irreversible inhibitors for alpha-chymotrypsin. *J. Biol. Chem.* **258**, 15046–53 (1983).
22. Baggelaar, M. P. Summary and Future Prospects. in *Activity-based Protein Profiling of Diacylglycerol Lipases* 177–190 (2017).
23. Rai, R. & Katzenellenbogen, J. A. Guanidinophenyl-substituted enol lactones as selective, mechanism-based inhibitors of trypsin-like serine proteases. *J. Med. Chem.* **35**, 4150–9 (1992).
24. Bachovchin, D. A. *et al.* A high-throughput, multiplexed assay for superfamily-wide profiling of enzyme activity. *Nat. Chem. Biol.* **10**, 656–663 (2014).
25. Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The Protein Kinase Complement of the Human Genome. *Science* **298**, 1912–1934 (2002).
26. Drewry, D. H., Willson, T. M. & Zuercher, W. J. Seeding collaborations to advance kinase science with the GSK Published Kinase Inhibitor Set (PKIS). *Curr. Top. Med. Chem.* **14**, 340–2 (2014).
27. Elkins, J. M. *et al.* Comprehensive characterization of the Published Kinase Inhibitor Set. *Nat. Biotechnol.* **34**, 95–103 (2015).
28. Karaman, M. W. *et al.* A quantitative analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **26**, 127–132 (2008).
29. Christmann-Franck, S. *et al.* Unprecedently Large-Scale Kinase Inhibitor Set Enabling the Accurate Prediction of Compound–Kinase Activities: A Way toward Selective Promiscuity by Design? *J. Chem. Inf. Model.* **56**, 1654–1675 (2016).

30. Bembenek, S. D., Hirst, G. & Mirzadegan, T. Determination of a Focused Mini Kinase Panel for Early Identification of Selective Kinase Inhibitors. *J. Chem. Inf. Model.* **58**, 1434–1440 (2018).
31. Pezzotti, N., Höllt, T., Lelieveldt, B., Eisemann, E. & Vilanova, A. Hierarchical Stochastic Neighbor Embedding. *Comput. Graph. Forum* **35**, 21–30 (2016).
32. Coen, M. H. Cross-Modal Clustering. in *The Twentieth National Conference On Artificial Intelligence* 932–937 (2005).
33. Bachovchin, D. A. *et al.* Superfamily-wide portrait of serine hydrolase inhibition achieved by library-versus-library screening. *Proc. Natl. Acad. Sci.* **107**, 20941–20946 (2010).
34. Gottlieb, H. E., Kotlyar, V. & Nudelman, A. NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. *J. Org. Chem.* **62**, 7512–7515 (1997).

Ik geloof dat wetenschappelijke kennis fractale eigenschappen heeft, dat het niet uitmaakt hoeveel we leren, hetgeen overblijft, hoe klein het ook mag lijken, is nog steeds zo oneindig complex als het geheel was aan het begin.
Isaac Asimov (eigen vertaling)

Samenvatting

Algemene samenvatting

Het gepresenteerde onderzoek in dit proefschrift betreft één van de grootste uitdagingen in de moderne medicijnontwikkeling: selectiviteit. Het medicijnonderzoek dat gericht is op een moleculair doel, is gebaseerd op de notie dat een specifiek eiwit een sleutelrol speelt in een bepaald biologisch proces. Hoewel deze aanpak van medicijnontwikkeling kan worden bekritiseerd[1], staat het belang van selectieve aanpassing van doel-eiwitten met behulp van chemische modulatoren buiten kijf. Het belang van selectiviteit wordt nog versterkt wanneer de modulatoren worden gebruikt om het therapeutische potentieel van het doel-eiwit te bewijzen. In de regel dient in de praktijk een therapeutische marge te

worden bepaald, waarbinnen het doeleiwit selectief gemoduleerd wordt. Bij hogere concentraties kunnen de effecten van het binden van andere eiwitten tot bijwerkingen en zelfs tot toxiciteit leiden.

Vanwege het belang van selectiviteit zijn er veel technieken ontwikkeld om de selectiviteit van remmers te bepalen, zoals familie-brede activiteitstesten (bijvoorbeeld voor kinases en luciferases) en specifiekere tests zoals het 44 eiwitten tellende CEREP-panel.[2] Een andere veelgebruikte techniek is activiteit-gebaseerde eiwit profilering (in het Engels: *activity-based protein profiling* of ABPP). Deze techniek heeft haar flexibiliteit bewezen en heeft daardoor tot een grote variëteit aan toepassingen geleid binnen de medicijnontwikkeling. Het principe achter ABPP is geïntroduceerd in **Hoofdstuk 1**. Een specifiek veld waarbinnen ABPP haar waarde heeft bewezen is het onderzoek naar het endocannabinoïde systeem (ECS).[3–6] Het ECS staat in de belangstelling vanwege zijn rol in pijnperceptie en inflammatie. Het signaleringssysteem bestaat in hoofdzaak uit de twee cannabinoïde receptoren (CB1 en CB2) en de belangrijkste endogene liganden, 2-arachidonoylglycerol (2-AG) en *N*-arachidonoyl ethanolamine (anandamide). Daarnaast spelen de synthetische en katabolische enzymen een belangrijke rol. 2-AG is de meest voorkomende endocannabinoïde en wordt voornamelijk geproduceerd door diacylglycerol lipase (DAGL) α en β. De verschillende rollen van de twee ECS signaallipides en de precieze functies van de twee isovormen van DAG-lipase zijn nog verre van duidelijk. Om onderzoek naar de functies van DAG-lipase mogelijk te maken is het nodig om *in vivo* actieve remmers te vinden, en om die specifiek en in meer detail te kunnen bestuderen.

Diacylglycerol lipase remmers: activiteit en specificiteit

Hoofdstuk 2 introduceerde de *hit-to-lead* optimalisatie van een nieuw chemotype als DAGL-remmer, de 1,2,4-triazool-sulfonamides. Dit type moleculen werd gevonden in een eerder beschreven screening van ruim 300.000 stoffen.[7] In dit hoofdstuk zijn de structuur-activiteitsrelaties bestudeerd door het testen van 65 varianten. Tijdens deze optimalisatie werd de inhibitie van DAGL-α met een factor 500 verbeterd, van een half-maximale inhibitie concentratie (IC_{50}) van ongeveer 1 μM tot slechts 2.5 nM. Door met ABPP de hele set ook in eiwitpreparaten van muizenhersenen te testen, kon de meest selectieve remmer worden geïdentificeerd om deze vervolgens te profileren. Stof **1** (Figuur 1), een 2-gesubstitueerde benzylpiperidine triazool ureum, bleek actief te zijn in culturen van Neuro-2a cellen. Om ook de *in vivo* werking van deze stof te bepalen, werd zij, samen met een gealkynyleerd analogon, intra peritoneaal geïnjecteerd in muizen. Na een incubatie van twee uur werden de hersenen en de milt geprepareerd en geanalyseerd met gel-gebaseerde ABPP. Hierdoor kon de resterende activiteit van de serine hydrolases worden gevisualiseerd. Beide stoffen bleken de bloed-brein barrière te kunnen passeren en remden DAGL-α sterk in het hersenweefsel.

Hoofdstuk 3 draaide om de optimalisatie van de fysicochemische eigenschappen van stof **1**. De focus lag op het verhogen van het topologische polaire oppervlak om hierdoor de verspreiding van de remmers tot de perifere organen te beperken, om hierdoor voornamelijk de perifeer gelokaliseerde DAGL-β te remmen. De optimalisatie werd begonnen door de piperazine uit de eerder gepubliceerde stof DO34 op te nemen in de structuur, wat leidde tot de zeer potente remmer **2** (Figuur 1). Vervolgens werd een

coherente set van 24 varianten gesynthetiseerd waarin de *tert*-butyl carbamaat substitutie vervangen werd door een selectie van amides, ureum-verbindingen, tertiaire amines en sulfonamides. Dit leidde tot vier zeer potente en *in vitro* selectieve remmers welke in Neuro-2a culturen werden getest. Een ruime tienvoud afname in de inhibitieconstante toonde aan dat de cellulaire opname minder efficiënt verliep. Dit kan mogelijk worden verklaard door een verminderde passieve diffusie over het celmembraan als gevolg van de sterk verhoogde polariteit van de verbindingen. Opvallend genoeg bleek de selectiviteit die gevonden was *in vitro* zich niet goed te vertalen naar de cellulaire omstandigheden.
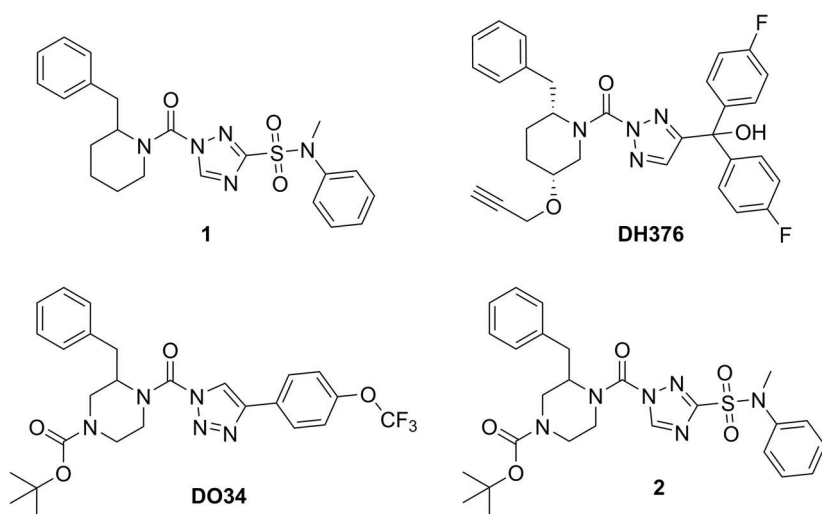
Figuur 1 | Chemische structuren van remmer **1**, referentiestof DH376 en DO34 en remmer **2**.

Hoofdstuk 4 beschreef de nauwgezette bestudering van de structuur-kinetiek relaties van een kleine set DH376 derivaten. Een eerder gepubliceerde biochemische surrogaat substraat analyse werd aangepast om de affiniteit (K_i) en reactiviteit (k_{inact}) van deze mechanisme-gebaseerde covalente remmers afzonderlijk te kunnen bepalen.[3] DynaFit programmatuur[8] werd gebruikt om een systeem van differentiaal vergelijkingen numeriek op te lossen, op basis van de door de kinetische analyse gegenereerde substraat omzettingscurves. Onverwachts volgden de inactivatie constantes (k_{inact}) totaal niet de trend in zuurtegraad (pK_a) van de heterocyclische verbindingen, wat op basis van de standaardtheorieën aangaande nucleofugiciteit werd voorspeld. Het bleek juist dat de positie en het aantal stikstofatomen in de heterocycle sterk de bindingsconstante K_i beïnvloedden. De 1,2,3-triazool van DH376 bleek te binden met een picomolaire affiniteit.

Activiteit-gebaseerde eiwit profilering om de selectiviteit van remmers te bestuderen

Hoofdstuk 5 demonstreerde de kracht van activiteit-gebaseerde eiwit profilering in een zeer relevante studie naar de interactieprofielen van het dodelijke kandidaat medicijn BIA 10-2474. Deze stof was ontworpen als fatty acid amide hydrolase (FAAH) remmer. Tijdens een fase 1 klinische studie in Frankrijk, begin 2016, moesten 5 deelnemers aan de

hoogste doseringsgroep worden opgenomen in het ziekenhuis met ernstige neurologische klachten. Een van de vrijwilligers overleed enkele dagen later. De belangrijkste hypothese om deze ernstige bijwerkingen te verklaren richtte zich op de mogelijkheid van zogenaamde bijvangsten, onbedoeld gebonden eiwitten. Om deze hypothese te kunnen testen werd BIA 10-2474 gesynthetiseerd, samen met drie alkyn-gesubstitueerde analoga. De inhibitie van BIA 10-2474 werd getest in een breed palet van ECS gerelateerde enzymen, alsmede beide CB-receptoren, en bleek *in vitro* enkel FAAH-activiteit (zwak) te remmen. Het inter-actieprofiel van BIA 10-2474 tegen 50 serine hydrolases werd getest met ABPP in lysaat van muizenhersenen. ABHD6 werd gevonden als bijvangst met een inhibitieconstante in dezelfde ordegrootte als nodig voor de volledige inhibitie van FAAH. Het covalente karakter van de inhibitie door de imidazool-ureum werd gedemonstreerd door de alkyn-analoga te ligeren aan fluoroforen met behulp van zogenaamde 'click'-chemie. Onder denaturerende SDS-PAGE-condities bleken FAAH en ABHD6 nog steeds meetbaar fluorescent gelabeld te worden. *In situ* profilering van de remmers toonde een dramatische toename in inhibitiecapaciteit voor zowel FAAH als ABHD6. Deze toename werd niet gevonden voor een veilige FAAH-remmer, PF-04457845.[9] Door *in situ* behandelde neurale celculturen te analyseren met hoge resolutie massaspectrometrische analyse technieken, werden additionele beiwitten geïdentificeerd. Deze werden gevalideerd door ze tot overexpressie te brengen en de *in situ* inhibitie vervolgens met ABPP te visualiseren. Een belangrijk punt is dat deze eiwitten niet werden geremd door PF-04457845, een FAAH-remmer die zich veilig heeft bewezen in fase 1 en 2 klinische studies. Alle eiwitten hadden met het lipidemetabolisme te maken, daarom werden de lipideniveaus geanalyseerd met gerichte lipodomics. Dit legde een ingrijpende verstoring van de lipidenetwerken in deze gecultuurde neuronen bloot. Dit suggereert dat niet-selectieve lipaseremmers metabolische deregulatie in het zenuwstelsel kunnen veroorzaken. Het belang van grondige interactieprofilering, mogelijk gemaakt door ABPP, werd nogmaals benadrukt en het gebruik ervan aangemoedigd.

Hoofdstuk 6 beschreef de ontwikkeling en diepgaande karakterisering van een groen-fluorescente fluorofosfonaat sensor, BODIPY-FP, een analogon van de veelgebruikte rood-fluorescente TAMRA-fluorofosfonaat. Deze nieuwe moleculaire sensor kan onafhankelijk van de DAG-lipase gerichtte, eveneens rood-fluorescente, sensor MB064 worden uitgelezen. De labelingsefficiëntie van verscheidene eiwitten door de nieuwe BODIPY-FP bleek significant af te wijken van TAMRA-FP. Met name de affiniteit voor FAAH en monoacyl-glycerol lipase (MAGL) bleek sterk verhoogd. Hierdoor kon BODIPY-FP op zeer lage concentratie (10 nM) gemengd worden met MB064 en detecteerbaar FAAH en MAGL labelen, zonder te interfereren met het labelingspatroon van MB064. Deze cocktail maakte het mogelijk om bijna alle endocannabinoïde gerelateerde serine hydrolases in een enkel monster te profileren. Deze multiplex ABPP-methode werd gebruikt om de selectiviteit van twee MAGL-remmers te testen, NF1819 en ABX-1431.[10,11] NF1819 bleek nog enkele andere eiwitten te remmen, waaronder ABHD6. ABX-1431 liet geen andere inhibitie activiteit zien bij 10 µM in de geteste muizenhersenpreparaten.

Automatisch leren om inhibitieprofielen te voorspellen

Hoofdstuk 7 introduceerde de toepassing van het automatisch lerende algoritme *t-Distributed Stochastic Neighbour Embedding* (t-SNE) als maat voor de vergelijkbaarheid van (bio)moleculen. In tegenstelling tot de veelgebruikte Tanimoto coëfficiënt (Tc), een afstandsmaat die met (moleculaire) bitstrings overweg kan, genereert t-SNE een tweedimensionale visualisatie van de hoogdimensionale bitstring ruimte. Dit maakt het mogelijk om de omringende chemische materie visueel te inspecteren, iets wat niet kan wanneer men de Tc gebruikt. De t-SNE vergelijkbaarheidsvisualisatie werd toegepast op 2774 klinisch toegepaste moleculen als demonstratie van de medicijnachtige chemische ruimte. Deze visualisatie toonde sterke colokalisering van chemisch gerelateerde stoffen. t-SNE werd ook toegepast op een fysicochemische bitstring gebaseerd op de uitgelijnde aminozuurvolgordes van eiwitten uit de serine hydrolase familie. Hiermee werd aangetoond dat t-SNE in staat is om hieruit fylogenetische informatie te herleiden. De in kaart gebrachte eiwitruimte werd geannoteerd met de inhibitiewaardes van een diverse set stoffen, eerder gemeten in een ABPP-gebaseerde analyse.[12] Ondanks het beperkte formaat van de beschikbare dataset kon enige groepering in de sterk geremde eiwitten worden opgemerkt. Dit duidt erop dat deze aanpak in theorie uitgebreid zou kunnen worden om de affiniteiten van remmers over de hele eiwitfamilie te voorspellen.

Hoofdstuk 8 bouwt voort op de principes die in Hoofdstuk 7 werden geïntroduceerd, maar concentreert zich op de grote familie van de kinases. Van deze eiwitfamilie zijn grote datasets beschikbaar in de wetenschappelijke literatuur die het mogelijk maken om automatisch lerende algoritmes zoals beschreven in het vorige hoofdstuk toe te passen. Ook de kinasefamilie liet zich uitstekend in kaart brengen met behulp van t-SNE, en toonde hierin sterke overeenkomsten met een eerder gepubliceerde kinasestamboom.[13] De stoffen die opgenomen zijn in de *Published Kinase Inhibitor Set* gecombineerd met de biologische data die beschikbaar is voor ± 200 kinases, vormden de basis voor de oefendata die gebruikt werd om het hier ontwikkelde *Drug Discovery Maps* model te onderwijzen.[14,15] Dit model werd gevalideerd aan de hand van de data van Karaman *et al.* en toonde zich evenwaardig of superieur aan geavanceerde computertechnieken als kwantitatieve structuur-activiteitsrelaties en proteochemometrische modellen.[16] Het ontwikkelde model werd toegepast in een virtuele zoektocht naar nieuwe FLT3-remmers. Een correct voorspelde ratio van 41% werd behaald, met 18 van de 44 voorspelde remmers inderdaad actief. Twee van deze stoffen werden gesynthetiseerd en getest in MV4:11 celculturen en beide bleken cellulair (zeer) actief. De andere kinases die voorspeld waren als doeleiwit van deze moleculen werden ook getest. Hieruit bleek dat ongeveer 70% van deze eiwitten inderdaad geremd werden. Tot slot werd het model gebruikt om de interactie-profielen te voorspellen van een commercieel verkrijgbare set stoffen die een motief bevatten, die de zogenaamde scharnierregio van kinases zou kunnen binden. Twintig stoffen hiervan werden gekocht en getest op hun activiteit tegen de kinases HIPK3 en PAK2. Helaas bleek geen van deze stoffen daadwerkelijk te werken, wat mogelijk te wijten is aan de lage dekkingsgraad van de oefendata rond deze doeleiwitten, of de matige vergelijkbaarheid tussen de moleculen in de oefendata en die in de commerciële bibliotheek.

Tot slot

Minder dan 1 op 10 kandidaat medicijnen die beginnen aan fase 1 klinische studies worden uiteindelijk goedgekeurd voor menselijk gebruik. Onverwachte bijwerkingen en toxiciteit zijn hier mede debet aan. Een beter begrip van de rol van selectiviteit en een beter inzicht in welke eiwitten daadwerkelijk door kandidaat medicijnen worden gebonden, kunnen bijdragen aan het voorkomen dat kandidaat medicijnen om deze redenen niet verder ontwikkeld worden. Dit proefschrift heeft enkele van deze uitdagende aspecten van de medicijnontwikkeling bestudeerd. Het gebruik van activiteit-gebaseerde eiwit profilering zoals in Hoofdstukken 2 en 3 tijdens medicijnontwikkeling en –optimalisatie, en in Hoofdstukken 5 en 6 om de interacties van een kandidaat medicijn in kaart te brengen, illustreert de veelzijdigheid en het belang van deze techniek uit de chemische biologie. Gecombineerd met de kennis verkregen uit biochemische analyses, zoals die ontwikkeld in Hoofdstuk 4, kan de medicinaal chemicus veel aan ABPP hebben. De recente enorme toename in de pop"ulariteit van zelflerende algoritmes, ondersteund door de exponentiële groei van beschikbare biologische data, biedt grote kansen voor de medicijnontwikkeling. Hoofdstukken 7 en 8 toonden de toepasbaarheid van één van die algoritmes aan, welke in staat bleek redelijk betrouwbare voorspellingen te doen over interactieprofielen van kinase-remmers. De moeilijkheden in het vinden, bepalen en voorspellen van selectiviteit zijn verre van opgelost, maar door ons begrip van de binding van remmers aan eiwitten, bedoeld of onbedoeld, te blijven uitbreiden, kunnen werkelijk selectieve remmers wellicht ooit werkelijkheid worden.

Citaties

1. Croston, G. E. The utility of target-based discovery. *Expert Opin. Drug Discov.* **12**, 427–429 (2017).
2. Bowes, J. *et al.* Reducing safety-related drug attrition: the use of in vitro pharmacological profiling. *Nat. Rev. Drug Discov.* **11**, 909–922 (2012).
3. 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).
4. Baggelaar, M. P. *et al.* Highly Selective, Reversible Inhibitor Identified by Comparative Chemoproteomics Modulates Diacylglycerol Lipase Activity in Neurons. *J. Am. Chem. Soc.* **137**, 8851–8857 (2015).
5. 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).
6. Baggelaar, M. P. *et al.* Chemical Proteomics Maps Brain Region Specific Activity of Endocannabinoid Hydrolases. *ACS Chem. Biol.* **12**, 852–861 (2017).
7. Janssen, F. J. Discovery of sulfonyl-1,2,4-triazole ureas as DAGLα inhibitors by HTS-ABPP. in *Discovery of novel inhibitors to investigate diacylglycerol lipases and α/β-hydrolase domain 16A* 109–139 (2016).
8. Kuzmič, P. DynaFit--a software package for enzymology. *Methods Enzymol.* **467**, 247–80 (2009).
9. Johnson, D. S. *et al.* Discovery of PF-04457845: A Highly Potent, Orally Bioavailable, and Selective Urea FAAH Inhibitor. *ACS Med. Chem. Lett.* **2**, 91–96 (2011).
10. Cisar, J. S. *et al.* Identification of ABX-1431, a Selective Inhibitor of Monoacylglycerol Lipase and Clinical Candidate for Treatment of Neurological Disorders. *J. Med. Chem.* **61**, 9062–9084 (2018).
11. Brindisi, M. *et al.* Development and Pharmacological Characterization of Selective Blockers of 2-Arachidonoyl Glycerol Degradation with Efficacy in Rodent Models of Multiple Sclerosis and Pain. *J. Med. Chem.* **59**, 2612–2632 (2016).
12. Bachovchin, D. A. *et al.* A high-throughput, multiplexed assay for superfamily-wide profiling of enzyme activity. *Nat. Chem. Biol.* **10**, 656–663 (2014).
13. Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The Protein Kinase Complement of the Human Genome. *Science* **298**, 1912–1934 (2002).
14. Drewry, D. H., Willson, T. M. & Zuercher, W. J. Seeding collaborations to advance kinase science with the GSK Published Kinase Inhibitor Set (PKIS). *Curr. Top. Med. Chem.* **14**, 340–2 (2014).
15. Elkins, J. M. *et al.* Comprehensive characterization of the Published Kinase Inhibitor Set. *Nat. Biotechnol.* **34**, 95–103 (2015).
16. Karaman, M. W. *et al.* A quantitative analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **26**, 127–132 (2008).

List of Publications

Drug Discovery Maps, a Machine Learning Model That Visualizes and Predicts Kinome–Inhibitor Interaction Landscapes
A.P.A. Janssen, S.H. Grimm, R.H.M. Wijdeven, E.B. Lenselink, J. Neefjes, C.A.A. van Boeckel, G.J.P. van Westen, M. van der Stelt.
Journal of Chemical Information and Modeling. **2018**, acs.jcim.8b00640.

Development of a Multiplexed Activity-Based Protein Profiling Assay to Evaluate Activity of Endocannabinoid Hydrolase Inhibitors
A.P.A. Janssen, D. van der Vliet, A.T. Bakker, M. Jiang, S.H. Grimm, G. Campiani, S. Butini, M. van der Stelt.
ACS Chemical Biology **2018**, *13*, 2406–2413.

Design and Synthesis of Quenched Activity-Based Probes for Diacylglycerol Lipase and α,β-Hydrolase Domain Containing Protein 6
E.J. van Rooden, M. Kohsiek, R. Kreekel, A.C.M. van Esbroeck, A.M.C.H. van den Nieuwendijk, A.P.A. Janssen, R.J.B.H.N. van den Berg, H.S. Overkleeft, M. van der Stelt.
Chemistry - An Asian Journal. **2018**.

Two-Step Activity-Based Protein Profiling of Diacylglycerol Lipase
E.J. van Rooden, R. Kreekel, T. Hansen, A.P.A. Janssen, A.C.M. van Esbroeck, H. den Dulk, R.J.B.H.N. van den Berg, J.D.C. Codée, M. van der Stelt.
Organic Biomolecular Chemistry **2018**, *16*, 5250–5253.

Activity-Based Protein Profiling Reveals off-target Proteins of the FAAH Inhibitor BIA 10-2474
A. C. M. van Esbroeck*, A.P.A. Janssen*, A.B. Cognetta*, D. Ogasawara*, G. Shpak, M. van der Kroeg, V. Kantae, M.P. Baggelaar, F.M.S. de Vrij, H. Deng, M. Allarà, F. Fezza, Z. Lin, T. van der Wel, M. Soethoudt, E.D. Mock, H. den Dulk, I.L. Baak, B.I. Florea, G. Hendriks, L. De Petrocellis, H.S. Overkleeft, T. Hankemeier, C.I. De Zeeuw, V. Di Marzo, M. Maccarrone, B.F. Cravatt, S.A. Kushner, M. van der Stelt.
Science, **2017**, *356*, 1084–1087.

Conformational Behaviour of Azasugars Based on Mannuronic Acid
E.R. van Rijssel, A.P.A. Janssen, A. Males, G.J. Davies, G.A. van der Marel, H.S. Overkleeft, J.D.C. Codée.
ChemBioChem **2017**, *18*, 1297–1304.

Discovery of an in vivo active NAPE-PLD inhibitor that reduces brain anandamide levels and pain behavior
E.D. Mock, M. Mustafa, R. Cinar, V. Kantae, X. Di, Z.V. Varga, J. Paloczi, G. Donvito, A.C.M. van Esbroeck, A.M.F. van der Gracht, I. Kotsogianni, J.K. Park, A. Martella, T. van der Wel, M. Soethoudt, M. Jiang, T.J. Wendel, <u>A.P.A. Janssen</u>, A. Bakker, B.I. Florea, J. Wat, H. van den Hurk, M. Wittwer, U. Grether, M.W. Buczynski, C.A.A. van Boeckel, T. Hankemeier, P. Pacher, A.H. Lichtman, M. van der Stelt.
Manuscript submitted

Discovery of *in vivo* active and selective *sn*-1-diacylglycerol lipase α inhibitors (*working title*)
<u>A.P.A. Janssen</u>*, F.J. Janssen*, M.P. Baggelaar, A.C.M. van Esbroeck, H. den Dulk, H. Deng, E. van Doornmalen, N. Smits, A. Morrison, E. Russell, J. Schultz, L. Brown, J. Hewitt, F. MacLeod,
J. Robinson, P.P. Geurink, H. Ovaa, B.I. Florea, H.S. Overkleeft, S.P.I. McElroy, C.A.A. van Boeckel, H. Rutjes, P.S. Jones, M. van der Stelt.
Manuscript in preparation

Structure Kinetics Relationships show crucial role for heterocycle leaving group in irreversible diacylglycerol lipase inhibitors (*working title*)
<u>A.P.A. Janssen</u>, J.M.A. van Hengst, O.J.M. Béquignon, H. Deng, G.J.P. van Westen, M. van der Stelt.
Manuscript in preparation

Discovery of 1,2,4-triazole sulfonamide ureas as *in vivo* active α/β-hydrolase domain 16A inhibitors (*working title*)
<u>A.P.A. Janssen</u>*, F.J. Janssen*, A. van der Ham, D. de Ridder, B.I. Florea, H.S. Overkleeft, M. van der Stelt.
Manuscript in preparation

*These authors contributed equally

Curriculum vitae

Anthe Janssen was born on December 7th, 1989 in Nootdorp, the Netherlands. He obtained his high school diploma (pre-university level, VWO) from the Huygenslyceum in Voorburg in 2008. He subsequently started a bachelor degree in Molecular Science and Technology, a shared program between the TU Delft and Leiden University. He obtained his bachelor's degree in 2011 *cum laude*. After a year as president of the study association Chemisch Dispuut Leiden, he started the master programme Chemistry with a specialisation in 'Design and Synthesis' in September 2012 at Leiden University. He obtained his master's degree *summa cum laude* in 2014.

During his master education he performed an industrial research internship at ASML in Veldhoven, the Netherlands, titled '*Directed Self-assembly: Revolutionary or complementary?*', which focussed on the physicochemical understanding of the phase separation of polymers in confined lithographic spaces, and potential applications thereof. He performed his major research internship in the Bio-organic Synthesis group with prof. dr. G. A. van der Marel titled '*An NMR-mediated study of pH-dependent conformational changes in 1-deoxymannojirimycin*'. This research studied the synthesis of small iminosugars and their conformational behaviour under varying acidity in aqueous and organic solvents. For this thesis he was awarded the *Unilever Research Prize 2014*.

In September 2014 he started his doctoral studies in the Leiden Institute of Chemistry, in the Bio-organic Synthesis group under supervision of prof. dr. H. S. Overkleeft and dr. M. van der Stelt. In his second year of research he transitioned to the newly founded group of Molecular Physiology, headed by (now) prof. dr. M. van der Stelt. The research described in Chapters 7 and 8 were performed in close collaboration with dr. G. J. P. van Westen, faculty member of the Leiden Academic Centre for Drug Research. The NWO is kindly acknowledged for the funding of this PhD research through ECHO-grant 711.014.009.

Posters on parts of the PhD research were presented at various (inter)national conferences: the International Cannabinoid Research Society's conference in Leiden, 2018; CHAINS in Veldhoven, in 2016 and 2017; and the FIGON Dutch Medicine Days in Ede in 2018. Oral presentations were held on parts of the PhD research at: CHAINS in Veldhoven, in 2016 and 2018, the FIGON Dutch Medicine Days in Ede in 2017, which received the *Best Presentation Award* in Medicinal Chemistry, and during the EFMC-YMCS in Ljubljana, Slovenia, in 2018, which received the *Public's Choice Prize* for the best presentation.

Curriculum vitae

Anthe Janssen werd geboren te Nootdorp, in Nederland, op 7 December 1989. Hij behaalde zijn middelbareschooldiploma (VWO-niveau) op het Huygenslyceum in Voorburg in 2008. Hierna begon hij zijn bacheloropleiding Molecular Science & Technology, een gedeeld opleidingsprogramma van de TU Delft en de Universiteit Leiden. Hij behaalde zijn BSc. graad in 2011 *cum laude*. Na een jaar als voorzitter van de studievereniging, het Chemisch Dispuut Leiden, begon hij zijn mastertraject in september 2012. Hij volgde de master Scheikunde met als specialisatie 'Ontwerp en Synthese' aan de Universiteit Leiden. Hij behaalde zijn MSc. graad in 2014 *summa cum laude*.

Tijdens zijn master heeft hij een industriële onderzoeksstage gevolgd bij ASML in Veldhoven, getiteld *'Gestuurde zelfassemblage: Revolutionair of Complementair?'*, waarin de fysicochemische fasescheidingen van polymeren in besloten lithografische ruimtes, en de mogelijke toepassingen daarvan, werden bestudeerd. Zijn hoofdstage heeft hij gedaan bij de Bio-organische Synthese groep bij prof. dr. G. A. van der Marel met als titel *'Een NMR-gemedieerde studie naar de pH-afhankelijke conformationele veranderingen van 1-deoxy-mannojirimycin'*. Dit onderzoek richtte zich op de synthese van kleine iminosuikers en hun conformationele gedragingen onder variërende zuurtegraad in waterige en organische oplosmiddelen. Voor het verslag hiervan ontving hij de *Unilever Onderzoeksprijs 2014*.

In september 2014 begon hij zijn doctoraalstudies in het Leids Instituut voor de Chemie, in de Bio-organische Synthese groep onder supervisie van prof. dr. H. S. Overkleeft en dr. M. van der Stelt. In het tweede jaar van zijn onderzoek stapte hij over naar de nieuw-opgerichte groep Moleculaire Fysiologie, geleid door (nu) prof. dr. M. van der Stelt. Het onderzoek beschreven in hoofdstukken 7 en 8 werd gedaan in nauwe samenwerking met dr. G. J. P. van Westen, werkzaam bij het Leids Academisch Centrum voor Medicijn-onderzoek. Dank gaat uit naar de Nederlandse Organisatie voor Wetenschappelijk Onderzoek voor de subsidie via ECHO-beurs 711.014.009 die dit doctoraal onderzoek mogelijk heeft gemaakt.

Posters over delen van het gepresenteerde doctoraal onderzoek werden gepresenteerd bij verscheidene (inter)nationale conferenties: de jaarlijkse conferentie van de International Cannabinoid Research Society in Leiden in 2018; CHAINS in Veldhoven, in 2016 en 2017; en de FIGON Dutch Medicine Days in Ede, in 2018. Mondelinge presentaties over delen van het onderzoek werden gegeven bij: CHAINS in Veldhoven, in 2016 en 2018; de FIGON Dutch Medicine Days in Ede in 2017, welke werd bekroond als beste presentatie in de medicinale chemie; en tijdens de EFMC-YMCS in Ljubljana, Slovenië, in 2018, welke door het publiek werd uitgekozen als beste presentatie van het congres.

Quod potui feci, faciant meliora potentes