



Universiteit
Leiden
The Netherlands

Visualization of vitamin A metabolism

Koenders, S.T.A.

Citation

Koenders, S. T. A. (2020, September 17). *Visualization of vitamin A metabolism*. Retrieved from <https://hdl.handle.net/1887/136528>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/136528>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/136528> holds various files of this Leiden University dissertation.

Author: Koenders, S.T.A.

Title: Visualization of vitamin A metabolism

Issue date: 2020-09-17

Visualization of Vitamin A Metabolism

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus Prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 17 september 2020
klokke 13.45 uur

door

Sebastiaan Theodorus Antonius Koenders
geboren te Nijmegen in 1991

Promotiecommissie

Promotoren: Prof.dr. M. Van der Stelt

Prof.dr. H.S. Overkleeft

Overige leden: Prof.dr. J. Brouwer

Prof.dr. G. Van der Marel

Prof.dr. R. Mebius

Vrije Universiteit, Amsterdam

Prof.dr. S. Sieber

Technische Universität München, Munich, Germany

Dr. S. Le Dévédec

Prof.dr. J. Neefjes

Cover designed by Frouke van Dijk

ISBN: 978-94-6402-509-5

Chance favors the prepared mind

Louis Pasteur

Table of Contents

Chapter 1	
General Introduction	7
Chapter 2	
Opportunities for Lipid-based Probes in the Field of Immunology	19
Chapter 3	
Design and Synthesis of an Activity-based Retinoid Probe	35
Chapter 4	
Biological Evaluation of LEI-945	49
Chapter 5	
Comparative and Competitive ABPP of ALDHs in Breast Cancer Cells using LEI-945	67
Chapter 6	
Visualizing Vitamin A Metabolism in the Small Intestines	85
Chapter 7	
Synthesis and Biological Evaluation of Clickable Vitamin A	105
Chapter 8	
Design, Synthesis and Biological Evaluation of a Broadspectrum ALDH Probe	135
Chapter 9	
Summary and Future Prospects	151
Nederlandse Samenvatting	167
List of Publications	176
Curriculum Vitae	177

Chapter 1

General Introduction

Aldehyde dehydrogenases

The aldehyde dehydrogenases (ALDHs) belong to a family of 19 oxidoreductases based on sequence homology.^{1,2} They perform important metabolic roles in our cells as they convert endogenous and exogenous aldehydes into carboxylic acids.³ ALDHs are nicotinamide adenine dinucleotide (NAD^+) dependent enzymes and have a conserved catalytic cysteine and glutamic acid. The mechanism by which they convert their aldehyde substrates into a carboxylic acid is shown in **Fig. 1.1**.¹ The catalytic cysteine is activated by the neighboring glutamic acid residue via deprotonation. Nucleophilic attack of the catalytic cysteine on the aldehyde results in a hemithioacetal. Subsequent hydride abstraction by NAD^+ results in a thioester. The thioester bond is hydrolyzed by a water molecule, which is deprotonated by the neighboring glutamate, thereby releasing the product.⁴ The conversion of aldehydes into carboxylic acids is vital as aldehydes are reactive functional groups, which may form adducts with both proteins and DNA leading to cellular damage.⁵

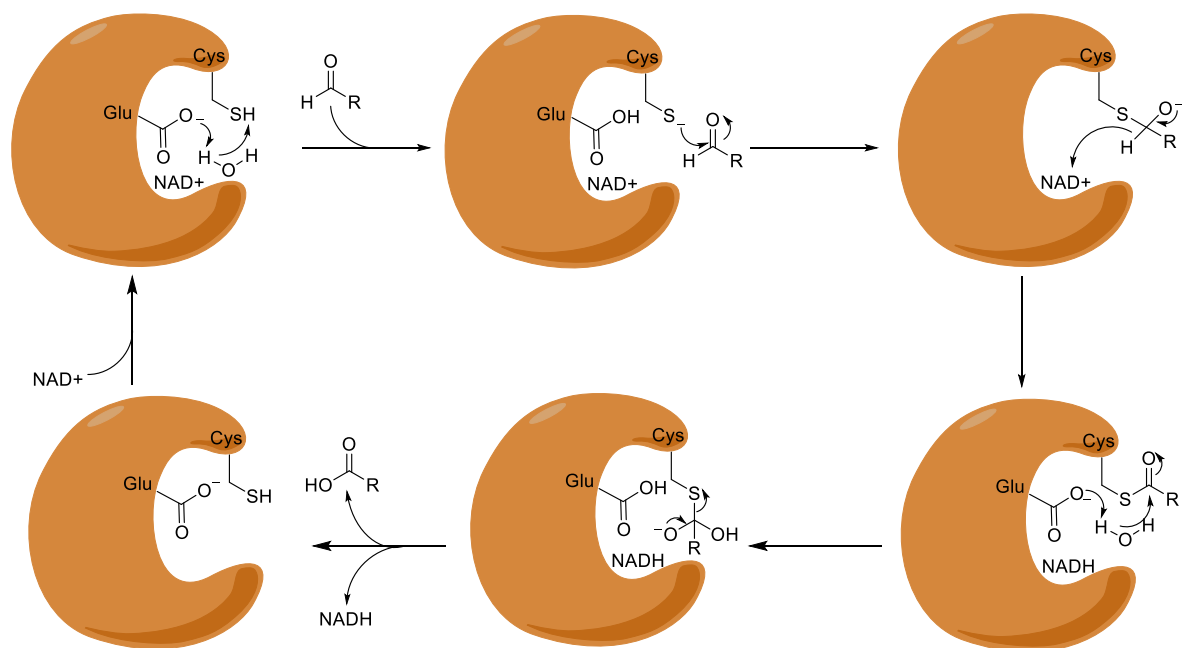


Fig. 1.1 | General enzymatic mechanism of ALDH enzymes. Starting from the upper left: the cofactor NAD⁺ binds and the glutamate activates the catalytic cysteine via deprotonation. The substrate then binds to the enzyme, followed by nucleophilic attack of the catalytic cysteine on the aldehyde. Hydride abstraction by NAD⁺ results in a thioester. Glutamate assisted hydrolysis of the thioester bond results in the release of the product and NADH.

ALDH activity is not only essential for the detoxification of aldehydes, but also for the production of several functional biomolecules, such as retinoic acid², γ -aminobutyric acid (GABA)⁶, folate⁷ and betaine⁸, that regulate important cellular processes.⁹ Not surprisingly, mutations in certain ALDH enzymes that compromise their enzymatic activity are associated with metabolic diseases¹⁰, such as Sjögren-Larsson syndrome¹¹, type II hyperprolinemia¹² and 4-hydroxybutyric aciduria.¹³ These are autosomal recessive disorders linked to mutations in ALDH3A2, ALDH4A1 and ALDH5A1, respectively.^{11–13}

Upregulation of ALDH expression levels is commonly observed in cancer cells. ALDH activity in breast cancer has been linked with chemo- and radiotherapy resistance.¹⁴ For example, ALDH1A1 overexpression is predictive of chemoresistance towards cyclophosphamide and the expression of ALDH1A3 is associated with poor clinical outcome in breast cancer.^{15–19} Consequently, ALDH inhibition is a potential therapeutic strategy to overcome therapy resistance in cancer.

To date, the only ALDH inhibitor used in the clinic is disulfiram (Antabuse, **Fig. 1.2**), which is prescribed to support chronic alcoholics to maintain sobriety. The compound was originally used in industry as an accelerator of the rubber vulcanization process. It was first linked to ethanol metabolism in 1937 by

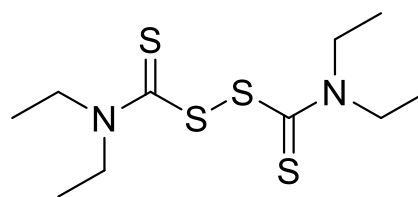


Fig. 1.2 | Chemical structure of disulfiram

E. Williams, who reported on the inability to consume alcohol of workers handling disulfiram.²⁰ Hald and Jacobsen showed that the sensitization to ethanol by disulfiram is accompanied by increased acetaldehyde concentrations in the blood.²¹ Although disulfiram has been shown to be a pan-ALDH inhibitor, its effect on ethanol metabolism has mainly be ascribed to its inhibition of mitochondrial ALDH2. ALDH2 is the main ALDH responsible for converting acetaldehyde, formed after the dehydrogenation of ethanol in the first metabolic step, into acetate in the liver.^{22,23}

Vitamin A: biological effects and metabolism

Vitamin A or retinol is taken up from the diet as the vitamin A precursors, β -carotene or retinyl esters.^{24,25} Retinyl esters are hydrolyzed into retinol in the intestinal lumen by pancreatic enzymes after which retinol is taken up by the intestinal epithelial cells.²⁶ Inside these cells, retinol can be re-esterified into retinyl esters and stored in blood circulating lipoprotein particles, called chylomicrons, which are transported to the liver, the main storage depot of vitamin A in the human body.²⁷ Alternatively, Retinol can be converted by the intestinal epithelial cells into retinal by retinol dehydrogenases (RDHs), which are members of the short chain dehydrogenase/reductase (SDR) superfamily²⁸, and alcohol dehydrogenases (ADHs).²⁹ The contrary reaction, reduction of retinal to retinol, can be performed by aldo-keto reductases (AKRs)³⁰ and RDHs.²⁸

β -carotene is taken up directly by intestinal epithelial cells and cleaved into two retinal molecules.³¹ Retinal is subsequently oxidized to retinoic acid by ALDHs in an irreversible step. A schematic overview of the enzymes involved in vitamin A metabolism in the small intestines is shown in **Fig. 1.3**.

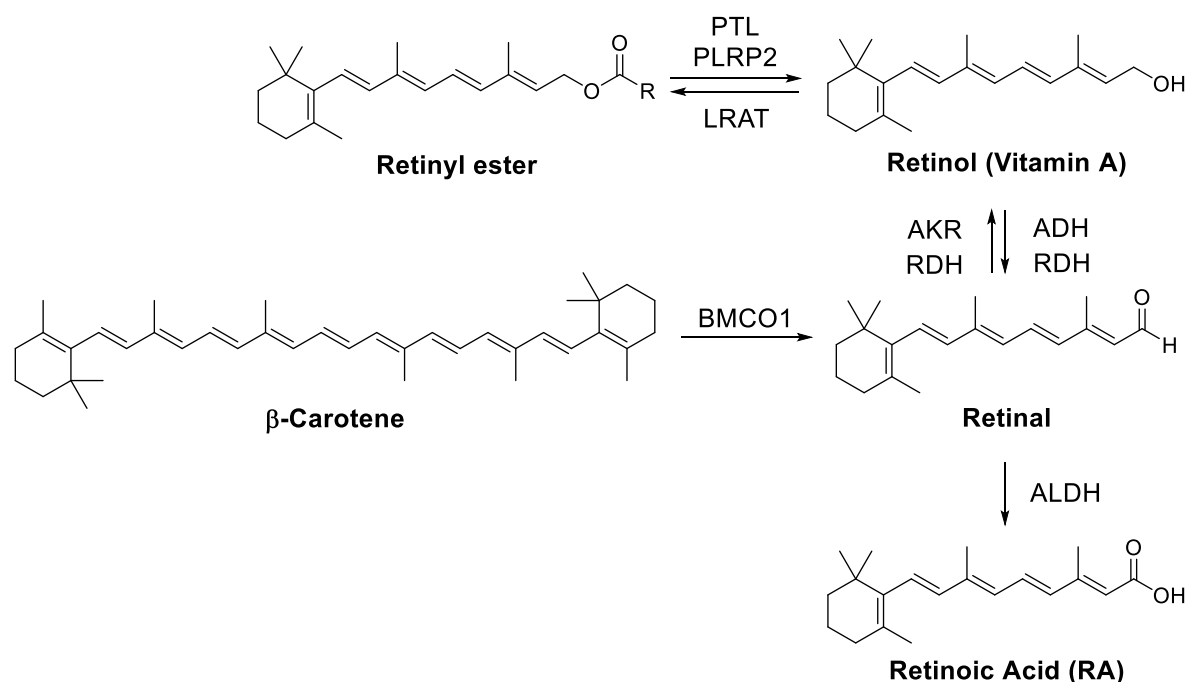


Fig. 1.3 | Schematic overview of vitamin A metabolism. Retinyl esters are converted in the lumen by pancreatic triglyceride lipase (PTL) and pancreatic lipase-related protein 2 (PLRP2) into retinol and then taken up by the intestinal epithelial cells (IECs). The IECs can convert retinol either back into retinyl esters intended for storage in the liver via lecithin retinol acyltransferase (LRAT) or into retinal by alcohol dehydrogenases (ADHs) and retinol dehydrogenases (RDHs). Retinal can be transformed back into retinol by aldo-keto reductases (AKRs) and RDHs. β-Carotene is cleaved by β-carotene 15,15'-dioxygenase (BMCO1) forming two molecules of retinal in the IECs. Retinal can subsequently be metabolized into retinoic acid by aldehyde dehydrogenases (ALDHs).

Retinoic acid regulates many cellular and physiological functions, including embryonic development, immunomodulation, neuronal differentiation and (cancer) stem cell proliferation.^{32–35} Most of these functions of retinoic acid are mediated via the retinoic acid receptor (RAR), which forms heterodimers with the retinoid X receptor (RXR). Binding of retinoic acid to the RAR/RXR heterodimer complex modulates gene transcription by recruiting different cofactors to the DNA-bound complex in a cell specific manner.^{36,37} Retinoic acid is essential in chordate animals as disruption of retinoic acid signaling can lead to severe (neural) developmental defects, autoimmunity disorders and cancer.^{32,38–40} The key function of retinoic acid in biological signaling implies that its cellular levels are tightly regulated.

Three retinaldehyde dehydrogenases (ALDH1A1, ALDH1A2 and ALDH1A3) have been identified that produce retinoic acid.^{1,2} These ALDHs have a variable and inducible cellular expression pattern. Their activity is regulated by post-translational modifications.^{41,42} Immunoblotting and quantitative real-time polymerase chain reaction (RT-PCR) are currently used to determine retinaldehyde dehydrogenase expression levels in cells, but these assays report solely on protein expression levels and not on activity.^{16,43}

The ALDEFLUOR assay does report on global ALDH activity levels in (cancer) stem cells. This assay uses a fluorescent aldehyde that upon oxidation to a charged fluorescent carboxylate is trapped within the cell. Therefore, cells with high ALDH activity become fluorescent and can be detected using fluorescence activated cell sorting (FACS) (**Fig. 1.4**).

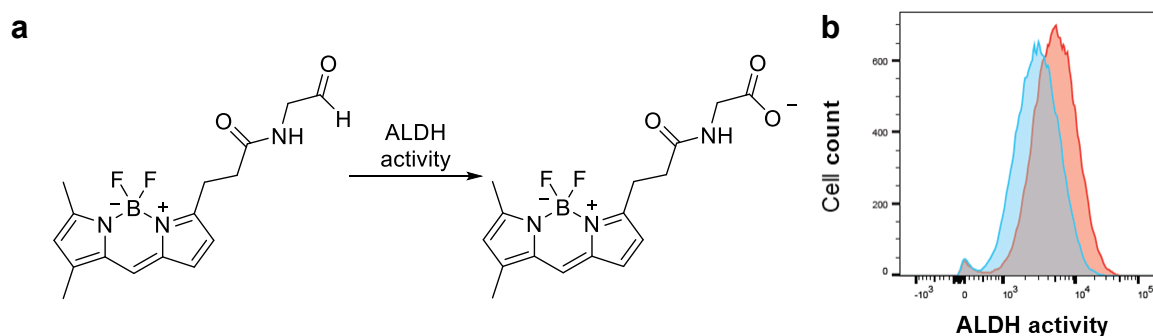


Fig. 1.4 | Determination of ALDH activity using the ALDEFLUOR assay. **a**, Chemical representation of the conversion of the ALDEFLUOR substrate by ALDHs. **b**, Representative FACS plot of cells treated with ALDEFLUOR (red) and a negative control in which cells were pretreated with an ALDH inhibitor (blue).

However, the ALDEFLUOR assay does not discriminate between individual ALDHs.⁴⁴ Recently developed selective fluorescent substrates report on the activity of a single enzyme, but do not provide an overview of the global ALDH activity present in a biological system.^{45,46} The development of chemical tools and methods to profile the levels of active retinaldehyde dehydrogenases is, therefore, important to study retinoic acid formation in health and disease, and for the discovery of effective ALDH inhibitors for therapeutic strategies.

Activity-based protein profiling

Activity-based protein profiling (ABPP) is a powerful chemical biological method that is ideally suited to report on the abundance and identity of active enzymes in complex biological systems.⁴⁷ It employs activity-based probes (ABP) that covalently and irreversibly react with conserved catalytic amino acid residues in the active site of proteins. ABPs constitute an electrophilic warhead to covalently react with the catalytic nucleophile, a recognition element for enzyme affinity and in the case of a two-step probe a ligation handle for the introduction of reporter molecules, such as fluorophores or biotin. These reporter groups enable target visualization by resolving samples on sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) followed by in-gel fluorescent scanning or target identification by affinity enrichment and chemical proteomics, respectively (**Fig. 1.5**).

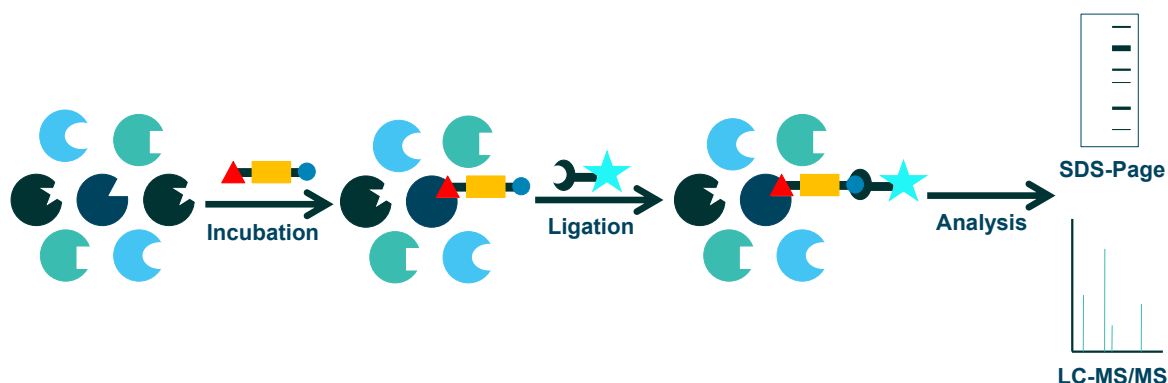


Fig. 1.5 | Activity-based Protein Profiling. Schematic representation of an ABPP workflow using a two-step probe.

Commonly used ligation handles are alkynes and azides which are coupled to the reporter groups using a copper(I)-catalyzed azide-alkyne [2+3] cycloaddition.⁴⁸ The introduction of the reporter group after covalent interaction of the probe with its target offers the advantage of avoiding potential issues with potency, selectivity and cell permeability.⁴⁹

There are two main applications for which ABPs are commonly used: comparative and competitive ABPP. Comparative ABPP studies the abundance of active enzymes in biological samples derived from different biological states, such as healthy versus disease states^{50,51}, different differentiation states, developmental stages or from various tissues.^{52,53} Competitive ABPP reports on the potency and selectivity of inhibitors. Preincubation of biological samples with an inhibitor blocks the active site of interacting enzymes, which can

therefore, not be labeled by the ABP. This results in a concentration-dependent decrease in labeled proteins, which can be quantified using chemical proteomics. ABPP can be applied in a wide variety of biological samples ranging from cell lysate and living cells to *in vivo* animal models.⁵⁴

ABPP has been successfully and widely applied in the field of serine hydrolases^{52,55}, proteases⁵⁶, kinases⁵⁷ and glycosidases.⁵⁸ However, only a few ABPs for oxidoreductases, which target P450⁵⁹, MAO⁶⁰ and 15-LOX⁶¹, have been published. There are currently no ABPs that report on ALDH activity. It is envisioned that chemical tools and methods that allow the identification and quantitation of the levels of active ALDHs in human cells is of utility to study the role of these enzymes in health and disease.

Aim and outline

The aim of this thesis is to design, synthesize and apply activity-based probes for the profiling of aldehyde dehydrogenase activity in complex biological samples. The outline of this thesis is as follows:

Chapter 2 provides an overview of the currently available lipid-based ABPs and discusses the (future) applications of lipid-based probes in immunology.

Chapter 3 describes the design and synthesis of the first-in-class retinal-based ABP **LEI-945** for the profiling of ALDHs.

Chapter 4 focusses on the biological validation of **LEI-945** as an ABP for ALDHs. The detection of various ALDH isozymes by **LEI-945** was demonstrated in living cells using fluorescence and chemical proteomic approaches.

Chapter 5 describes the utility of **LEI-945** in quantitating ALDH isozyme activity in a panel of breastcancer subtypes. **LEI-945** was superior to the widely used ALDEFLUOR assay in explaining the ability of breast cancer (stem) cells to produce retinoic acid. Furthermore, the probe revealed the cellular selectivity profile of an advanced ALDH1A1 inhibitor, thereby prompting the investigation of its cytotoxic nature.

Chapter 6 elaborates on the role of vitamin A metabolism in intestinal epithelial cells in immune homeostasis. Comparative ABPP using **LEI-945** of intestinal cells isolated from wild type and ALDH1A1 knockout mice pointed to ALDH1B1 as a potential alternative retinoic acid producing enzyme.

Chapter 7 describes the synthesis of alkyne functionalized retinoid analogues without an electrophilic warhead. Clickable retinoic acid differentiated HL60 cells into granulocytes in a similar fashion as retinoic acid. The probes were subsequently used in photoaffinity based chemical proteomics to define the retinoid interactome.

Chapter 8 describes the design, synthesis and application of an easily accessible broadspectrum ABP with a non-retinoid scaffold in the profiling of ALDH activity in cancer cells.

Chapter 9 summarizes the work presented in this thesis and reflects on future opportunities for the application of chemical probes profiling aldehyde dehydrogenases and retinoid interacting partners.

References

1. Koppaka, V. *et al.* Aldehyde Dehydrogenase Inhibitors: a Comprehensive Review of the Pharmacology, Mechanism of Action, Substrate Specificity, and Clinical Application. *Pharmacol. Rev.* **64**, 520–539 (2012).
2. Duester, G., Mic, F. A. & Molotkov, A. Cytosolic retinoid dehydrogenases govern ubiquitous metabolism of retinol to retinaldehyde followed by tissue-specific metabolism to retinoic acid. in *Chemico-Biological Interactions* **143–144**, 201–210 (Elsevier, 2003).
3. Vasilou, V., Pappa, A. & Petersen, D. R. Role of aldehyde dehydrogenases in endogenous and xenobiotic metabolism. *Chem. Biol. Interact.* **129**, 1–19 (2000).
4. Wang, X. & Weiner, H. Involvement of Glutamate 268 in the Active Site of Human Liver Mitochondrial (Class 2) Aldehyde Dehydrogenase as Probed by Site-Directed Mutagenesis. *Biochemistry* **34**, 237–243 (1995).
5. Setshedi, M., Wands, J. R. & De La Monte, S. M. Acetaldehyde adducts in alcoholic liver disease. *Oxidative Medicine and Cellular Longevity* **3**, 178–185 (2010).
6. Kurys, G., Ambroziak, W. & Pietruszkos, R. Human Aldehyde Dehydrogenase. *J. Biol. Chem.* **264**, 4715–4721 (1989).
7. Kutzbach, C. & Stokstad, E. L. R. Mammalian methylenetetrahydrofolate reductase Partial purification, properties, and inhibition by S-adenosylmethionine. *BBA - Enzymol.* **250**, 459–477 (1971).
8. Chern, M. K. & Pietruszko, R. Human aldehyde dehydrogenase E3 isozyme is a betaine aldehyde dehydrogenase. *Biochem. Biophys. Res. Commun.* **213**, 561–568 (1995).
9. Jackson, B. *et al.* Update on the aldehyde dehydrogenase gene (ALDH) superfamily. *Hum. Genomics* **5**, 283–303 (2011).
10. Vasilou, V. & Pappa, A. Polymorphisms of Human Aldehyde Dehydrogenases. *Pharmacology* **61**, 192–198 (2000).
11. Laurenzi, V. De *et al.* Sjögren–Larsson syndrome is caused by mutations in the fatty aldehyde dehydrogenase gene. *Nat. Genet.* **12**, 52–57 (1996).
12. Geraghty, M. T. *et al.* Mutations in the Δ^1 -pyrroline 5-carboxylate dehydrogenase gene cause type II hyperprolinemia. *Hum. Mol. Genet.* **7**, 1411–1415 (1998).
13. Chambliss, K. L. *et al.* Two Exon-Skipping Mutations as the Molecular Basis of Succinic Semialdehyde Dehydrogenase Deficiency (4-Hydroxybutyric Aciduria). *Am. J. Hum. Genet.* **63**, 399–408 (1998).
14. Croker, A. K. & Allan, A. L. Inhibition of aldehyde dehydrogenase (ALDH) activity reduces chemotherapy and radiation resistance of stem-like ALDH hiCD44 + human breast cancer cells. *Breast Cancer Res. Treat.* **133**, 75–87 (2012).
15. Qiu, Y. *et al.* The expression of aldehyde dehydrogenase family in breast cancer. *J. Breast Cancer* **17**, 54–60 (2014).
16. Marcato, P. *et al.* Aldehyde Dehydrogenase Activity of Breast Cancer Stem Cells Is Primarily Due To Isoform ALDH1A3 and Its Expression Is Predictive of Metastasis. *Stem Cells* **29**, 32–45 (2011).
17. Sládek, N. E., Kollander, R., Sreerama, L. & Kiang, D. T. Cellular levels of aldehyde dehydrogenases (ALDH1A1 and ALDH3A1) as predictors of therapeutic responses to cyclophosphamide-based chemotherapy of breast cancer: A retrospective study. *Cancer Chemother. Pharmacol.* **49**, 309–321 (2002).
18. Tomita, H., Tanaka, K., Tanaka, T. & Hara, A. Aldehyde dehydrogenase 1A1 in stem cells and cancer. *Oncotarget* **7**, 11018–32 (2016).
19. De Beca, F. F. *et al.* Cancer stem cells markers CD44, CD24 and ALDH1 in breast cancer special histological types. *J. Clin. Pathol.* **66**, 187–191 (2013).
20. Williams, E. E. Effects of Alcohol on Workers with Carbon Disulfide. *Journal of the American Medical Association* **109**, 1472–1473 (1937).
21. Hald, J. & Jacobsen, E. The Formation of Acetaldehyde in the Organism after Ingestion of Antabuse (Tetraethylthiuramdisulphide) and Alcohol. *Acta Pharmacol. Toxicol. (Copenh.)* **4**, 305–310 (1948).
22. Yasgar, A. *et al.* A High-Content assay enables the automated screening and identification of small molecules with specific ALDH1A1-Inhibitory activity. *PLoS One* **12**, 1–19 (2017).
23. Chen, C.-H., Ferreira, J. C. B., Gross, E. R. & Mochly-Rosen, D. Targeting Aldehyde Dehydrogenase 2: New Therapeutic Opportunities. *Physiol. Rev.* **94**, 1–34 (2014).
24. D'Ambrosio, D. N., Clugston, R. D. & Blaner, W. S. Vitamin A metabolism: An update. *Nutrients* **3**, 63–103 (2011).
25. Woollard, G. Retinol, retinoic acid, carotenes and carotenoids: Vitamin a structure and terminology. *Food Nutr. Components Focus* **1**, 3–22 (2012).
26. Harrison, E. H. Mechanisms of digestion and absorption of dietary vitamin A. *Annu. Rev. Nutr.* **25**, 87–103 (2005).

27. Blomhoff, R., Green, M. H., Green, J. B., Berg, T. & Norum, K. R. Vitamin A metabolism: new perspectives on absorption, transport, and storage. *Physiol.Rev.* **71**, 951–990 (1991).
28. Lidén, M. & Eriksson, U. Understanding retinol metabolism: Structure and function of retinol dehydrogenases. *J. Biol. Chem.* **281**, 13001–13004 (2006).
29. Molotkov, A. *et al.* Stimulation of retinoic acid production and growth by ubiquitously expressed alcohol dehydrogenase Adh3. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 5337–5342 (2002).
30. Ruiz, F. X., Porté, S., Parés, X. & Farrés, J. Biological role of aldo-keto reductases in retinoic acid biosynthesis and signaling. *Front. Pharmacol.* **3**, 58 (2012).
31. D'Ambrosio, D. N., Clugston, R. D. & Blaner, W. S. Vitamin A Metabolism: An Update. *Nutrients* **3**, 63–103 (2011).
32. Zile, M. H. Vitamin A and Embryonic Development: An Overview. *J. Nutr.* **128**, 455–458 (1998).
33. Wiseman, E. M., Bar-El Dadon, S. & Reifen, R. The vicious cycle of vitamin a deficiency: A review. *Crit. Rev. Food Sci. Nutr.* **57**, 3703–3714 (2017).
34. Mora, J. R., Iwata, M. & Andrian, U. H. Von. Vitamin effects on the immune system. *Nat. Rev. Immunol.* **8**, 685–698 (2008).
35. Blomhoff, H. K. *et al.* Vitamin A is a key regulator for cell growth, cytokine production, and differentiation in normal B cells. *J. Biol. Chem.* **267**, 23988–23992 (1992).
36. Carlberg, C. Lipid soluble vitamins in gene regulation. *BioFactors* **10**, 91–97 (1999).
37. Balmer, J. E. & Blomhoff, R. Gene expression regulation by retinoic acid. *J. Lipid Res.* **43**, 1773–808 (2002).
38. Noy, N. Between Death and Survival: Retinoic Acid in Regulation of Apoptosis. *Annu. Rev. Nutr.* **30**, 201–17 (2010).
39. Schambach, F., Schupp, M., Lazar, M. A. & Reiner, S. L. Activation of retinoic acid receptor- α favours regulatory T cell induction at the expense of IL-17-secreting T helper cell differentiation. *Eur. J. Immunol.* **37**, 2396–2399 (2007).
40. Erkelens, M. N. & Mebius, R. E. Retinoic Acid and Immune Homeostasis: A Balancing Act. *Trends in Immunology* **38**, 168–180 (2017).
41. Wang, J. *et al.* Phosphorylation-dependent regulation of ALDH1A1 by Aurora kinase A: Insights on their synergistic relationship in pancreatic cancer. *BMC Biol.* **15**, 1–22 (2017).
42. Zhao, D. *et al.* NOTCH-induced aldehyde dehydrogenase 1A1 deacetylation promotes breast cancer stem cells. *J. Clin. Invest.* **124**, 5453–5465 (2014).
43. Moreb, J. S., Zucali, J. R., Ostmark, B. & Benson, N. A. Heterogeneity of aldehyde dehydrogenase expression in lung cancer cell lines is revealed by Aldefluor flow cytometry-based assay. *Cytom. Part B Clin. Cytom.* **72B**, 281–289 (2007).
44. Zhou, L. *et al.* Identification of cancer-type specific expression patterns for active aldehyde dehydrogenase (ALDH) isoforms in ALDEFLUOR assay. *Cell Biol. Toxicol.* **35**, 161–177 (2019).
45. Anorma, C. *et al.* Surveillance of Cancer Stem Cell Plasticity Using an Isoform-Selective Fluorescent Probe for Aldehyde Dehydrogenase 1A1. *ACS Cent. Sci.* **4**, 1045–1055 (2018).
46. Yagishita, A. *et al.* Development of Highly Selective Fluorescent Probe Enabling Flow-Cytometric Isolation of ALDH3A1-Positive Viable Cells. *Bioconjug. Chem.* **28**, 302–306 (2017).
47. Cravatt, B. F., Wright, A. T. & Kozarich, J. W. Activity-Based Protein Profiling: From Enzyme Chemistry to Proteomic Chemistry. *Annu. Rev. Biochem.* **77**, 383–414 (2008).
48. Hein, J. E. & Fokin, V. V. Copper-catalyzed azide-alkyne cycloaddition (CuAAC) and beyond: New reactivity of copper(I) acetylides. *Chemical Society Reviews* **39**, 1302–1315 (2010).
49. 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).
50. van Rooden, E. J. *et al.* Chemical proteomic analysis of serine hydrolase activity in Niemann-Pick Type C mouse brain. *Front. Neurosci.* **12**, 440 (2018).
51. Nomura, D. K., Dix, M. M. & Cravatt, B. F. Activity-based protein profiling for biochemical pathway discovery in cancer. *Nature Reviews Cancer* **10**, 630–638 (2010).
52. Liu, Y., Patricelli, M. P. & Cravatt, B. F. Activity-based protein profiling: The serine hydrolases. *Proc. Natl. Acad. Sci.* **96**, 14694–14699 (1999).
53. Baggelaar, M. P. *et al.* Chemical Proteomics Maps Brain Region Specific Activity of Endocannabinoid Hydrolases. *ACS Chem. Biol.* **12**, 852–861 (2017).
54. 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).

55. 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).
56. Serim, S., Haedke, U. & Verhelst, S. H. L. Activity-Based Probes for the Study of Proteases: Recent Advances and Developments. *ChemMedChem* **7**, 1146–1159 (2012).
57. Zhao, Q. *et al.* Broad-spectrum kinase profiling in live cells with lysine-targeted sulfonyl fluoride probes. *J. Am. Chem. Soc.* **139**, 680–685 (2017).
58. Kuo, C. L. *et al.* Activity-Based Probes for Glycosidases: Profiling and Other Applications. in *Methods in Enzymology* **598**, 217–235 (Academic Press, 2018).
59. Wright, A. T. & Cravatt, B. F. Chemical Proteomic Probes for Profiling Cytochrome P450 Activities and Drug Interactions In Vivo. *Chem. Biol.* **14**, 1043–1051 (2007).
60. Krysiak, J. M. *et al.* Activity-based probes for studying the activity of flavin-dependent oxidases and for the protein target profiling of monoamine oxidase inhibitors. *Angew. Chemie Int. Ed.* **51**, 7035–7040 (2012).
61. Eleftheriadis, N., Thee, S. A., Zwinderman, M. R. H., Leus, N. G. J. & Dekker, F. J. Activity-Based Probes for 15-Lipoxygenase-1. *Angew. Chemie Int. Ed.* **55**, 12300–12305 (2016).

Chapter 2

Opportunities for Lipid-based Probes in the Field of Immunology

Published as part of S.T.A. Koenders, B. Gagestein & M. Van der Stelt., *Current Topics in Microbiology and Immunology*, **420**, 283–319 (2018).

Introduction

Lipids are defined as hydrophobic biomolecules that dissolve in organic solvents, but not in water. They perform a wide range of functions inside the cell, ranging from structural building blocks of membranes and energy storage to cell signaling. The discovery of the inhibitory effect of aspirin on prostaglandin biosynthesis showed that lipids can modulate the immune system and that the enzymes involved in their metabolism constitute potential drug targets.¹ Since this initial discovery, many connections have been made between the immune system and signaling lipids in the field of endocannabinoids, resolvins, steroid hormones and vitamins A and D.^{2–5} However, due to the low abundance, high lipophilicity and inherent instability of many lipid signaling molecules, their lipid-protein interaction profile and mode of action have remained largely elusive.

In recent years, several technical advances in mass spectrometry and innovative chemical biology strategies have been developed to shed light on signaling lipids and their protein interaction landscapes as well as their biological function. Since its inception the field of lipodomics, which is the analysis of lipids and their interacting partners within a biological system, has made great strides forward.⁶ Standardization of protocols, increased availability of deuterated lipids and the high mass accuracy and resolution of modern mass spectrometers have made it possible to identify and quantify many lipids in complex biological samples.⁷

In an effort to systematically classify the rapidly expanding database of characterized lipids, the Lipid Metabolites and Pathways Strategy (LIPID MAPS) consortium has come up with a more concise and generally accepted definition of a lipid, which in their view comprises a small hydrophobic or amphipathic molecule that is formed at least partially by the condensation of ketoacyl thioesters and/or isoprene units.⁸ Based on these two building blocks eight major lipid classes are defined within this classification: glycerophospholipids, glycerolipids, fatty acyls, sphingolipids, sterol lipids, prenol lipids, saccharolipids and polyketides (**Fig. 2.1**).⁹

The variability in chemical and physical properties between the various lipid groups makes it a challenge to measure all lipid species in one experiment. Different sample preparation methods and measuring techniques are required to quantify different classes of lipids.¹⁰ Adding to the challenge is the wide range of alterations that lipids can undergo to form still non-identified lipid derivatives. By combining the fields of chemistry and biology some of these challenges can be overcome.

The field of chemical biology has developed two main approaches to study lipids: a) chemically modified lipids to study their biological role and b) chemical tools to visualize and modulate the proteins involved in lipid metabolism. To track lipids in a biological system they can be functionalized with alkyne or azide tags.^{11,12} Using bioorthogonal ligation chemistry these functionalized lipids can be visualized and identified.¹³ This method has been used for the visualization of lipids in membranes, modification of proteins by lipids and to study lipid metabolism.^{14–16}

Although the introduction of a ligation handle enables affinity purification, this will only reveal protein interaction partners that are covalently bound to the lipid. To visualize lipid interacting partners, such as binding proteins or metabolizing proteins, photoaffinity or activity-based labeling can be used. These methods require bifunctional lipids and probes. These molecules contain a photoreactive group or an electrophilic warhead and an alkyne or azide, which are employed in affinity-based protein profiling (A/BPP) and activity-based protein profiling (ABPP).¹⁷ Both methods allow for the visualization and identification of protein-lipid interactions.¹⁸

In this chapter, the recent developments concerning activity-based protein probes based on lipids are summarized. The probes are grouped based on the enzyme class they target. As well, some opportunities for future research are presented.

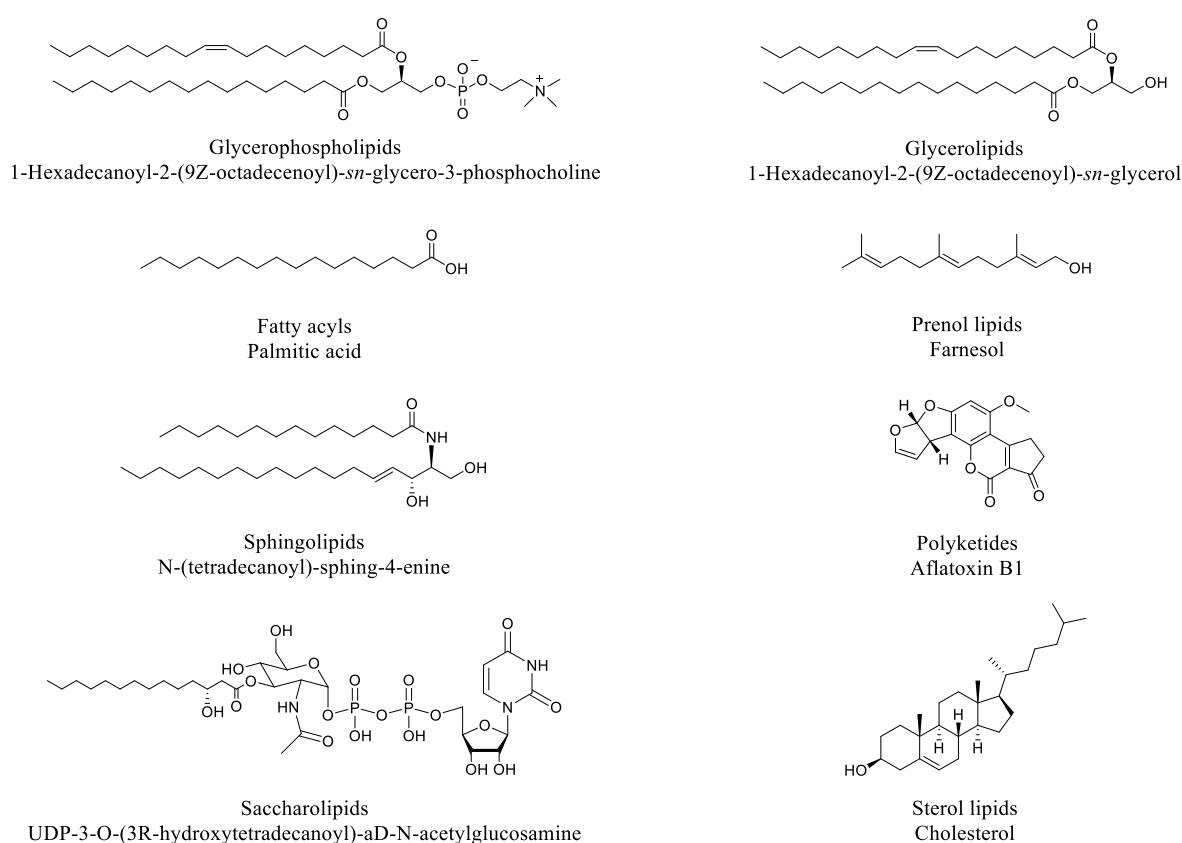


Fig. 2.1 | Representative structures of the eight lipid categories as defined by the LIPID MAPS consortium.^{8,9}

Lipid-based activity-based probes

Activity-based protein profiling (ABPP) uses a chemical probe to covalently label and identify an enzyme or class of enzymes in a biological sample. Activity-based probes (ABPs) consist of a reactive group or warhead and a reporter tag. ABPs can further be classified in one-step and two-step probes based on the presence of a reporter tag during protein labeling or ligation of the tag after the labeling event using click chemistry, respectively. The chemical structure of the probe combined with the warhead determines the affinity, reactivity and selectivity of the ABP.

The advantage of using ABPP over transcriptomics or whole-cell proteomics is the ability to quantify the amount of active proteins, whereas alternative methods do not discriminate between active or inactive protein forms. If a protein is inactive due to a post-translational modification (PTM) or blocked by a non-allosteric inhibitor it will not react with the ABP, because the reactive group interacts with the active site of the enzyme.

Most of the foundational work in the field of ABPP has been done in the field of serine hydrolases. This class of enzymes consists of more than 200 enzymes in humans of which about half are serine proteases and the other half metabolic enzymes with an active site serine.¹⁹ The group of metabolic serine hydrolases can be divided into (thio)esterases, lipases, and peptidases and thus can hydrolyze a wide variety of substrates.²⁰ Lipases hydrolyze triglycerides feeding liberated fatty acids back into the β -oxidative pathway. Next to their role in the energy household of cells, these lipases also play a role in cellular signaling as some of their products, such as 2-arachidonoylglycerol (2-AG) and anandamide, are endogenous ligands of the cannabinoid receptors.

The first ABP developed for serine hydrolases is fluorophosphonate (FP)-biotin, which is a one-step probe with a FP-warhead.²¹ Together with the FP-TAMRA probe **1**, they are routinely used for the mapping of activity profiles of serine hydrolases in different animals or tissues as well as for the determination of the selectivity profiles of inhibitors.^{22–26}

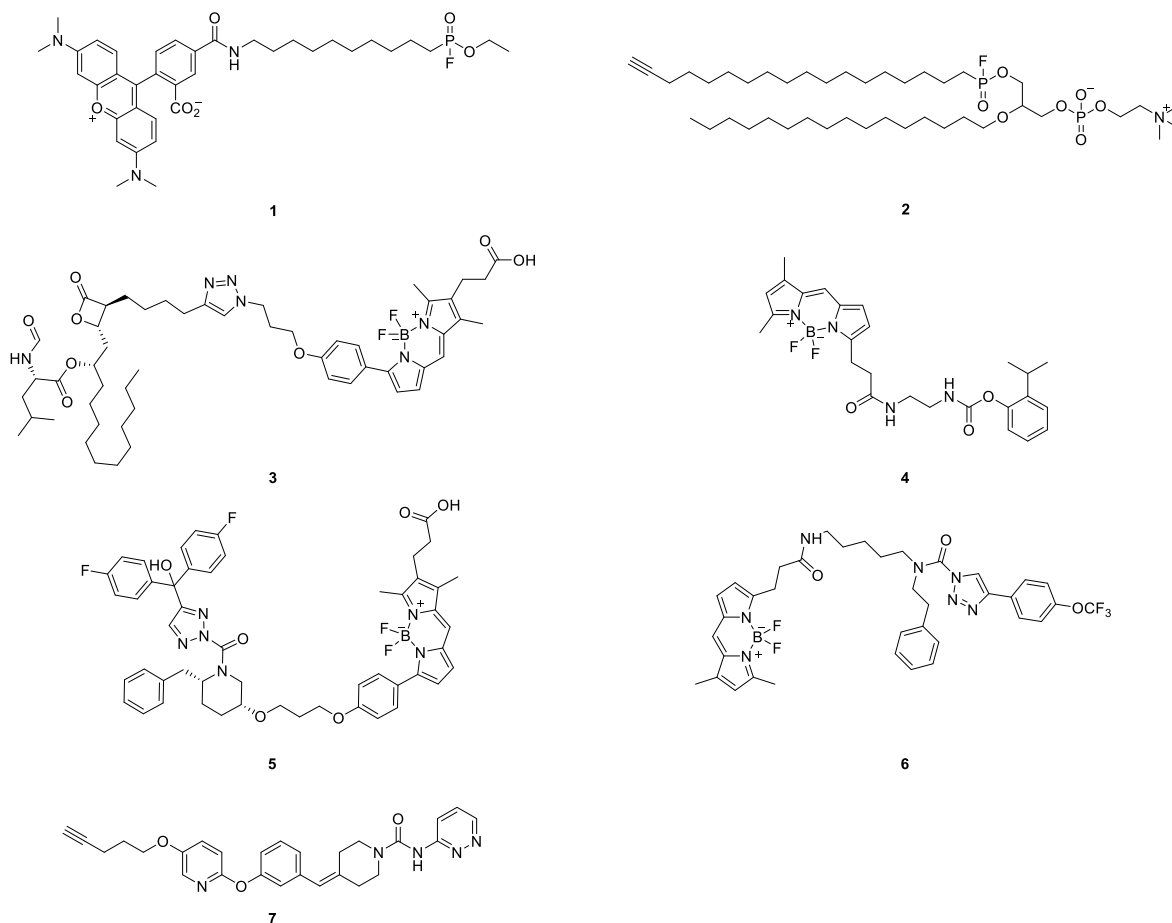


Fig. 2.2 | Structures of some activity-based probes for serine hydrolases. Broadspectrum serine hydrolase FP probes **1-2**, THL based probe MB064 **3**, KIAA1363 selective probe JW576 **4**, DAGL-α and DAGL-β selective probes DH379 **5** and HT-01 **6**, FAAH selective probe PF-04457845yne **7**.

By tweaking the scaffold of traditional FP probes, subclass-selective probe **2** for serine phospholipases was made.²⁷ This demonstrates that the scaffold and reactivity of the warhead are key for affinity and selectivity of a probe (**Fig. 2.2**). The FP-probes enable measurement of the activity of a wide range of serine hydrolases, but do not cover the entire family. Several complementary ABPs have therefore been developed over the past decade.

For example, MB064 (**3**) is based on the scaffold of tetrahydrolipstatin (THL), a promiscuous lipase inhibitor that also labelled diacylglycerol lipase alpha (DAGL-α), which is not targeted by FP-TAMRA (**1**).²⁸ Compared to FP-TAMRA (**1**), MB064 (**3**) covered a smaller subset of serine hydrolases, but labelled enzymes not targeted by FP-TAMRA (**1**).²² Another example is JW576 (**4**), an ABP selective for KIAA1363, which can also be used as a biomarker in imaging (**Fig. 2.2**).²⁹

Another group of selective serine hydrolase probes are the triazole urea probes represented by DH379 (**5**) and HT-01 (**6**).^{30,31} These probes show selectivity for DAGL- α and DAGL- β . HT-01 has been used to study the regulatory role of DAGL- β in the inflammatory response of macrophages.³¹ The last example in this group is the fatty-acid amide hydrolase 1 (FAAH)-selective probe PF-04457845yne (**7**) (**Fig. 2.2**).³² The amidase FAAH hydrolyses a variety of endocannabinoid lipid amides, thereby inactivating these lipid messengers. Probe **7** was synthesized to study the off-targets of the covalent FAAH inhibitor PF-04457845 directly. This ABP proved to be selective for FAAH in mouse brain and liver tissue.

More recently, two probes (**8** and **9**) for lipid modifying lysosomal cysteine hydrolases, have been published (**Fig. 2.3**).^{33,34} This group of hydrolases is involved in the hydrolysis of lipids, but performs its activity via an active site cysteine instead of a serine. Both probes are activity-based and react with the catalytic cysteine. However, they use different warheads for this purpose. The carmofur-based probe **8** uses a urea group and ARN14686 (**9**) uses a β -lactam. Probe **8** labeled acid ceramidase (AC) and ARN14686 (**9**) was shown to label both AC and *N*-acyl ethanolamide acid amidase (NAAA). Changing the ligation tag of **9** led to the development of norbornene-ABP **10** and an BODIPY-ABP **11** based on a *N,O*-substituted threonine- β -lactam.³⁵ Both probes react with NAAA and could be used to label NAAA directly (**Fig. 2.3**).

Another interesting group of ABPs are the probes **12** and **13** for β -glucosidases, such as glucosylceramidase (GBA) that hydrolyze glucosylceramide (**Fig. 2.3**). ABP **12** was shown to be selective for GBA1, whereas ABP **13** acted as a broadspectrum probe for GBA1, GBA2, GBA3 and lactase/phlorizin hydrolase (LPH).^{36,37} These probes are currently used to study Gaucher disease in which deficiency of GBA leads to accumulation of its substrate.

ABPP has also made its entry into the field of the lipid vitamins. RPE65 or retinoid isomerohydrolase is essential in the visual cycle and converts all-*trans*-retinyl esters to 11-*cis*-retinol.³⁸ LRAT transfers acyl groups from lecithin to all-*trans*-retinol yielding retinyl esters, which provides the substrate for RPE65 and is therefore also part of the visual cycle.³⁹

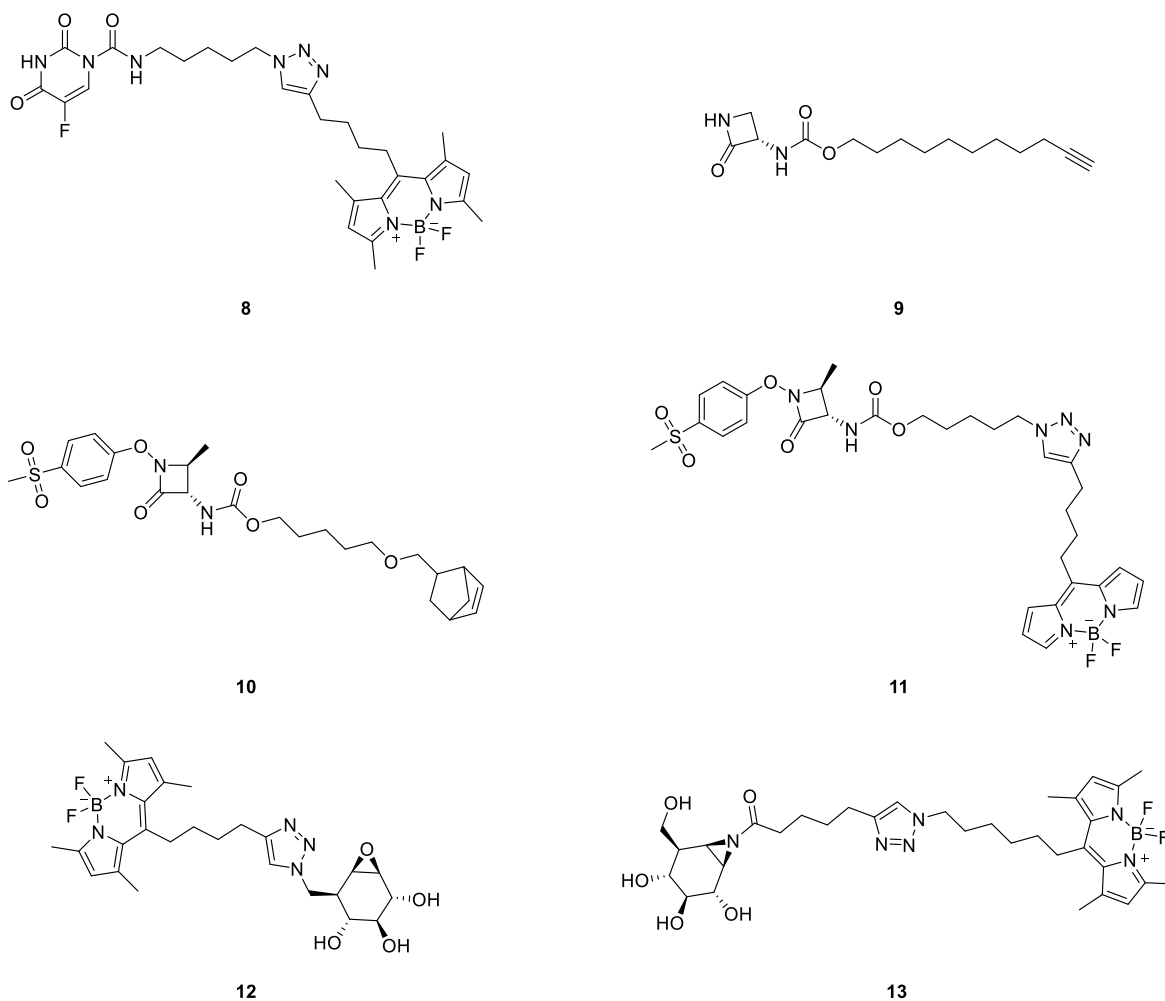


Fig. 2.3 | Structures of activity-based probes for cysteine hydrolases 8-11 and β -glucosidases 12-13.

The one-step probe, all-*trans*-retinyl chloroacetate (RCA) (**14**), mimicking retinyl acetate was synthesized in 2002 and contains a chloroacetate warhead and a cleavable biotin linker (**Fig. 2.4**).⁴⁰ This ABP **14** was shown to label RPE65 and lecithin retinol acyltransferase (LRAT) in retinal pigment epithelial membrane.⁴¹

Another lipid soluble vitamin, vitamin D₃, was used as a scaffold for probe **15** (**Fig. 2.4**). Probe **15** was shown to bind the vitamin D-binding protein (DBP) in a similar fashion as the natural vitamin D₃ ligand, positioning the bromoacetate electrophile for S_N2 displacement of the bromide by a neighboring tyrosine phenol.⁴² The resulting covalent complex could be visualized via the carbon-14 (¹⁴C) of probe **15** using phosphor imaging. After separation of the proteins using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) the gel is placed on a photostimulable phosphor plate.

Exposure to the ionizing radiation from the ^{14}C radioisotope excites electrons in the phosphor plates and traps them in this excited state. This latent image can be released using a scanning laser, which lets the excited electrons return to their base state, while emitting photons. The luminescence is captured and used to visualize the gel image. A similar affinity probe based on the 1,25-dihydroxyvitamin D_3 equipped with a bromoacetate was shown to label the vitamin D nuclear receptor (VDR).⁴³

Lipoxygenases oxygenate polyunsaturated fatty acids (PUFAs) and are essential metabolic enzymes in the formation of resolvins. N144 (**16**) is a recently developed ABP for 15-lipoxygenase-1 (15-LOX-1).⁴⁴ 15-LOX-1 oxidizes its natural substrate linoleic acid through a radical mechanism. N144 (**16**) mimics linoleic acid and interacts via its bis(alkyne)core with the active site of the enzyme (**Fig. 2.4**). At the active site a hydrogen is abstracted from the bis(alkyne)core resulting in a radical, which can insert itself into one of the neighboring amino acid bonds. As a result the enzyme is irreversibly inhibited and covalently labelled by N144 (**16**). Probe **16** possesses a terminal alkene, which can be utilized as a bioorthogonal handle in an oxidative Heck reaction with biotinylated phenylboronic acid to ligate a biotin reporter tag.⁴⁵

Arylfluorosulfate probe **17** has recently been shown to covalently label intracellular lipid binding proteins (**Fig. 2.4**). Probe **17** and similar arylfluorosulfate probes, have been shown to react with a reactive tyrosine inside the binding pocket of cellular retinoic acid binding proteins (CRABPs) and fatty acid binding proteins (FABPs).⁴⁶ These arylfluorosulfates enable the visualization of these lipid trafficking proteins and their ligands without the need of probes with a photoaffinity group.

Opportunities

Activity-based probes are relatively underrepresented in the category of lipid-based chemical probes.⁴⁷ The development of these lipid-based ABPs provides opportunities in the field of chemical biology. Even more so in chemical immunology, as lipid signaling molecules are increasingly recognized as important regulators of the immune response.

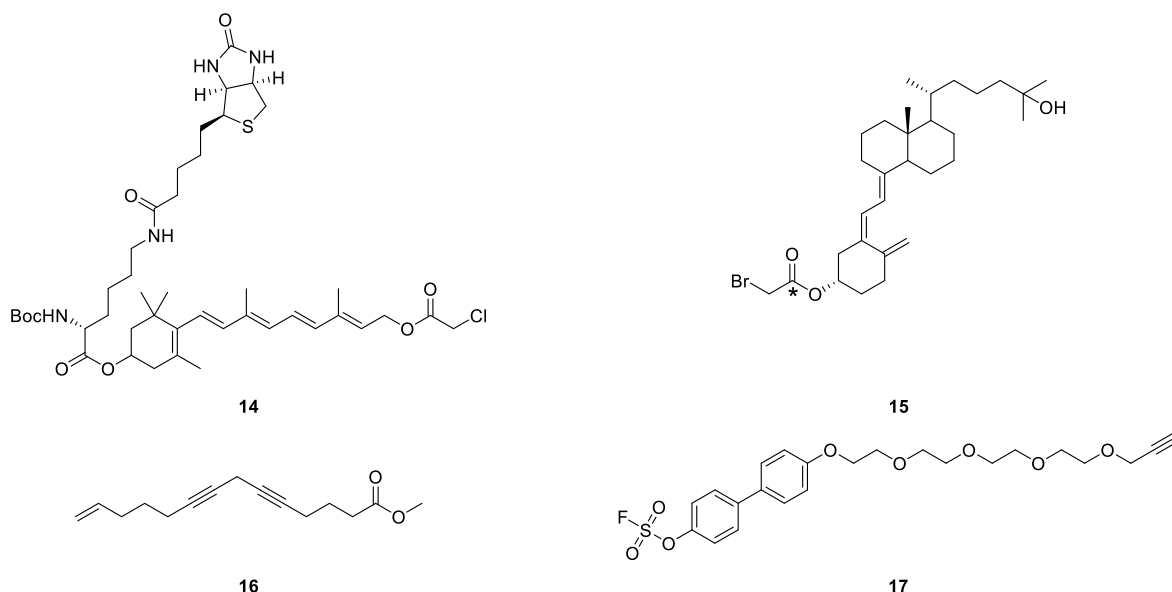


Fig. 2.4 | Structures of activity-based probes based on retinyl acetate 14, vitamin D 15 (* indicates ^{14}C), linoleic acid 16 and arylfluorosulfonate probe 17.

Endocannabinoids

The serine hydrolase ABPs discussed previously have been widely used to study the endocannabinoid system (ECS) and the effect of its lipid messengers on the central nervous system (CNS). The main receptors for these lipids are the cannabinoid receptor type 1 and type 2 (CB_1 and CB_2 receptor respectively). While CB_1 is mainly expressed within the central nervous system, CB_2 is mainly expressed in immune cells and tissue.⁴⁸

Since the discovery of the CB_2 receptor, determining its expression level in different immune cell types is done by either quantifying its mRNA levels or by Western blotting.⁴⁹ Both of these methods however have drawbacks. While mRNA levels do not always correlate with protein levels, they also provide no information on the activity of the protein. Protein degradation, PTMs and protein-protein interactions are important factors in regulating the amount of active enzyme. The presence of protein mRNA does therefore not predictably correlate with the amount and activity of an enzyme. While correlating with the amount of protein, antibodies on the other hand can have cross-reactivity problems.⁵⁰ Without the necessary quality controls and with different groups using different antibodies, it is a difficult task to determine which results are significant and reproducible. Recently it has been shown that photoaffinity probes can be used as an alternative to antibodies.⁵¹

With endocannabinoids being implicated in several immune diseases, the field of immunology presents a clear opportunity for lipid-based chemical probes.^{52–55} These ABPs would enable the visualization of levels of active enzyme in immune cells as has been shown with the HT-01 probe **6** for DAGL- β in macrophages.³¹ Coupled with cell sorting techniques, the effect of different kinds of stimulation on the ECS can be analyzed. This technique could also be used to compare healthy and diseased tissues.

Steroid hormones

The steroid hormones can be divided into two categories: sex steroids and corticosteroids. They are derived from cholesterol and are fat-soluble. These hormones act as lipid messengers and interact with nuclear steroid receptors influencing gene expression. More recently they have been implicated to also act via a different, faster signaling mechanism within the plasmamembrane.⁵⁶ Due to the sexual dimorphism in the immune system, sex steroids have long been implicated to affect the immune system. Testosterone seems to suppress the immune system leading to a lowered immune response to infections and vaccines in men.^{57,58} Women display a more robust immune response, but are more susceptible to autoimmune diseases possibly due to the modulatory effects of estrogen.^{59,60} Most research in this field has focused on administering steroid hormones, removing hormone producing organs or protein knockouts.^{61,62} It is therefore envisioned that steroid hormone-based probes have potential in the field of immunology. For example, no steroid hormone-based activity-based probes have been published so far. These probes would enable investigation of steroid hormone modifying enzymes in a biological system.

Lipid soluble vitamins

Vitamin A and D regulate gene expression by binding to their nuclear receptors the retinoic acid receptor (RAR) and vitamin D receptor (VDR), respectively.⁶³ Vitamin D, synthesized in the skin under influence of light, is a secosteroid and therefore also structurally resembles the steroid hormones. Both lipid vitamins have been implicated to affect the immune system.⁵ A radioactive photoaffinity probe and a radioactive affinity-based probe based on vitamin D have been synthesized and applied to characterize the binding pocket of the vitamin D binding protein.^{42,64}

These probes could be replaced by new ABPs with a diazirine as a less bulky photoactive group and addition of a click-handle to enable proteomics studies. A vitamin D-based ABP would enable the study of enzymes involved in the metabolism of vitamin D, which could be used to study its involvement in immunological processes.

In the case of vitamin A and its metabolites, direct photoaffinity labeling with radioactive retinoic acid was enabled by its inherent photoreactive characteristics.^{65,66} A radioactive photoaffinity probe based on retinoic acid has been synthesized, but A/BPP has not been performed due to lack of a ligation handle for application in chemical proteomics studies.⁶⁷ The ABP for retinyl ester processing enzymes has not yet been tested outside of the retinal pigment epithelial membrane.⁴¹ The application of this ABP in the field of immunology would be interesting as vitamin A is involved in the immunohomeostasis of the gut.⁶⁸ As these retinyl ester processing enzymes are involved in the metabolism of vitamin A, comparing their activity levels in intestinal epithelial cells in different disease models could indicate whether vitamin A is processed for storage in the liver or used to regulate the local immune response.^{69–71} Furthermore, ABPs based on retinol, retinal and the immunologically active metabolite, retinoic acid (RA), would be valuable additions to this field as RA is involved in the differentiation and gut-homing of lymphocytes.^{68,72} Probes based on these lipids would therefore enable research into the factors regulating their metabolism.

Other lipid soluble vitamins, such as vitamin E and K, have not been studied using ABPs. Vitamin E has been associated with T cell differentiation and decreased cellular immunity in aging.^{73,74} Next to its antioxidant activity, it has also been implicated as a potential orphan nuclear receptor ligand.^{63,75} The synthesis of a vitamin E derivative containing a diazirine and an alkyne click-handle would therefore be a potential chemical tool to unravel its protein-interactions and mechanism of action. Such a photoaffinity probe could help to find the proposed nuclear receptor or discover other cellular pathways involved in vitamin E biology.

Conclusion

Chemoproteomics using lipid-based probes is an invaluable strategy to study the biology of lipid messengers. A/BPP using photoaffinity probes can provide insight into the target interaction landscape of lipid messengers with previously unknown proteins, such as transporters and receptors, while ABPP using activity-based probes could identify the enzymes that control the metabolism of these important messengers in health and disease. The combination of A/BPP and ABPP is a powerful approach to obtain a global and detailed view of the biological processes mediated by lipid signaling molecules. The current set of lipid-based probes, however, does not cover many lipid classes yet. A potential reason could be that most long-chain, polyunsaturated lipids and their metabolites are challenging to synthesize and are inherently instabile. Although chemists are still working on the total synthesis of these low abundant biologically active lipids and their probes, there is still a need for further elucidation of important cellular processes performed by lipids, especially in the field of immunology.^{76,77}

References

1. Vane, J. R. Inhibition of prostaglandin synthesis as a mechanism of action for aspirin-like drugs. *Nat. New Biol.* **231**, 232–235 (1971).
2. Pandey, R., Mousawy, K., Nagarkatti, M. & Nagarkatti, P. Endocannabinoids and immune regulation. *Pharmacol. Res.* **60**, 85–92 (2009).
3. Serhan, C. N. & Petasis, N. A. Resolvins and Protectins in Inflammation Resolution. *Chem. Rev.* **111**, 5922–5943 (2011).
4. Marshall-Gradisnik, S., Green, R., Brenu, E. & Weatherby, R. Anabolic androgenic steroids effects on the immune system: a review. *Open Life Sci.* **4**, 19–33 (2009).
5. Mora, J. R., Iwata, M. & Andrian, U. H. Von. Vitamin effects on the immune system. *Nat. Rev. Immunol.* **8**, 685–698 (2008).
6. Wenk, M. R. The emerging field of lipidomics. *Nat. Rev. Drug Discov.* **4**, 594–610 (2005).
7. Wenk, M. R. Lipidomics: New Tools and Applications. *Cell* **143**, 888–895 (2010).
8. Fahy, E., Cotter, D., Sud, M. & Subramaniam, S. Lipid classification, structures and tools. *Biochim. Biophys. Acta* **1811**, 637–47 (2011).
9. Fahy, E. *et al.* Update of the LIPID MAPS comprehensive classification system for lipids. *J. Lipid Res.* **50**, 9–14 (2009).
10. Yang, K. & Han, X. Lipidomics: Techniques, Applications, and Outcomes Related to Biomedical Sciences. *Trends in Biochemical Sciences* **41**, 954–969 (2016).
11. Gaebler, A. *et al.* Alkyne lipids as substrates for click chemistry-based in vitro enzymatic assays. *J. Lipid Res.* **54**, 2282–2290 (2013).
12. Robichaud, P. P. *et al.* On the cellular metabolism of the click chemistry probe 19-alkyne arachidonic acid. *J. Lipid Res.* **57**, 1821–1830 (2016).
13. Hein, J. E. & Fokin, V. V. Copper-catalyzed azide-alkyne cycloaddition (CuAAC) and beyond: New reactivity of copper(I) acetylides. *Chemical Society Reviews* **39**, 1302–1315 (2010).
14. Gaebler, A., Penno, A., Kuerschner, L. & Thiele, C. A highly sensitive protocol for microscopy of alkyne lipids and fluorescently tagged or immunostained proteins. *J. Lipid Res.* **57**, 1934–1947 (2016).
15. Tate, E. W., Kalesh, K. A., Lanyon-Hogg, T., Storck, E. M. & Thion, E. Global profiling of protein lipidation using chemical proteomic technologies. *Curr. Opin. Chem. Biol.* **24**, 48–57 (2015).
16. Hofmann, K. *et al.* A novel alkyne cholesterol to trace cellular cholesterol metabolism and localization. *J. Lipid Res.* **55**, 583–591 (2014).
17. Cravatt, B. F., Wright, A. T. & Kozarich, J. W. Activity-Based Protein Profiling: From Enzyme Chemistry to Proteomic Chemistry. *Annu. Rev. Biochem.* **77**, 383–414 (2008).
18. Wright, M. H. & Sieber, S. A. Chemical proteomics approaches for identifying the cellular targets of natural products. *Nat. Prod. Rep.* **33**, 681–708 (2016).
19. Long, J. Z. & Cravatt, B. F. The metabolic serine hydrolases and their functions in mammalian physiology and disease. *Chemical Reviews* **111**, 6022–6063 (2011).
20. Simon, G. M. & Cravatt, B. F. Activity-based proteomics of enzyme superfamilies: Serine hydrolases as a case study. *Journal of Biological Chemistry* **285**, 11051–11055 (2010).
21. Liu, Y., Patricelli, M. P. & Cravatt, B. F. Activity-based protein profiling: The serine hydrolases. *Proc. Natl. Acad. Sci.* **96**, 14694–14699 (1999).
22. Baggelaar, M. P. *et al.* Chemical Proteomics Maps Brain Region Specific Activity of Endocannabinoid Hydrolases. *ACS Chem. Biol.* **12**, 852–861 (2017).
23. van Rooden, E. J. *et al.* Chemical proteomic analysis of serine hydrolase activity in Niemann-Pick Type C mouse brain. *Front. Neurosci.* **12**, 440 (2018).
24. 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).
25. Lentz, C. S. *et al.* Identification of a *S. aureus* virulence factor by activity-based protein profiling (ABPP) article. *Nat. Chem. Biol.* **14**, 609–617 (2018).
26. Zweerink, S. *et al.* Activity-based protein profiling as a robust method for enzyme identification and screening in extremophilic Archaea. *Nat. Commun.* **8**, 15352 (2017).
27. 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).
28. 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).
 29. Chang, J. W., Moellering, R. E. & Cravatt, B. F. An Activity-Based Imaging Probe for the Integral Membrane Hydrolase KIAA1363. *Angew. Chemie Int. Ed.* **51**, 966–970 (2012).
 30. 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).
 31. Hsu, K. L. *et al.* DAGL β inhibition perturbs a lipid network involved in macrophage inflammatory responses. *Nat. Chem. Biol.* **8**, 999–1007 (2012).
 32. 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).
 33. Romeo, E. *et al.* Activity-Based Probe for *N*-Acylethanolamine Acid Amidase. *ACS Chem. Biol.* **10**, 2057–2064 (2015).
 34. Ouairy, C. M. J. *et al.* Development of an acid ceramidase activity-based probe. *Chem. Commun.* **51**, 6161–6163 (2015).
 35. Petracca, R. *et al.* Novel activity-based probes for *N*-acylethanolamine acid amidase. *Chem. Commun.* **53**, 11810–11813 (2017).
 36. Witte, M. D. *et al.* Ultrasensitive in situ visualization of active glucocerebrosidase molecules. *Nat. Chem. Biol.* **6**, 907–913 (2010).
 37. Kallemijn, W. W. *et al.* Novel activity-based probes for broad-spectrum profiling of retaining β -exoglucosidases in situ and in vivo. *Angew. Chemie - Int. Ed.* **51**, 12529–12533 (2012).
 38. Cai, X., Conley, S. M. & Naash, M. I. RPE65: Role in the visual cycle, human retinal disease, and gene therapy. *Ophthalmic Genetics* **30**, 57–62 (2009).
 39. Jahng, W. J., Xue, L. & Rando, R. R. Lecithin Retinol Acyltransferase Is a Founder Member of a Novel Family of Enzymes. *Biochemistry* **42**, 12805–12812 (2003).
 40. Nesnas, N., Rando, R. R. & Nakanishi, K. Synthesis of biotinylated retinoids for cross-linking and isolation of retinol binding proteins. *Tetrahedron* **58**, 6577–6584 (2002).
 41. Jahng, W. J., David, C., Nesnas, N., Nakanishi, K. & Rando, R. R. A cleavable affinity biotinylating agent reveals a retinoid binding role for RPE65. *Biochemistry* **42**, 6159–6168 (2003).
 42. Swamy, N., Addo, J., Vskokovic, M. R. & Ray, R. Probing the Vitamin D Sterol-Binding Pocket of Human Vitamin D-Binding Protein with Bromoacetate Affinity Labeling Reagents Containing the Affinity Probe at C-3, C-6, C-11, and C-19 Positions of Parent Vitamin D Sterols. *Arch. Biochem. Biophys.* **373**, 471–478 (2000).
 43. Swamy, N. *et al.* Molecular modeling, affinity labeling, and site-directed mutagenesis define the key points of interaction between the ligand-binding domain of the vitamin D nuclear receptor and 1 α ,25-dihydroxyvitamin D₃. *Biochemistry* **39**, 12162–12171 (2000).
 44. Eleftheriadis, N., Thee, S. A., Zwiderman, M. R. H., Leus, N. G. J. & Dekker, F. J. Activity-Based Probes for 15-Lipoxygenase-1. *Angew. Chemie Int. Ed.* **55**, 12300–12305 (2016).
 45. Ourailidou, M. E. *et al.* Aqueous Oxidative Heck Reaction as a Protein-Labeling Strategy. *ChemBioChem* **15**, 209–212 (2014).
 46. Chen, W. *et al.* Arylfluorosulfates Inactivate Intracellular Lipid Binding Protein(s) through Chemoselective SuFEx Reaction with a Binding Site Tyr Residue. *J. Am. Chem. Soc.* **138**, 7353–7364 (2016).
 47. Koenders, S. T. A., Gagestein, B. & van der Stelt, M. Opportunities for lipid-based probes in the field of immunology. in *Current Topics in Microbiology and Immunology* **420**, 283–319 (Springer, Cham, 2018).
 48. Cabral, G. A. & Griffin-Thomas, L. T. Emerging role of the cannabinoid receptor CB₂ in immune regulation: Therapeutic prospects for neuroinflammation. *Expert Reviews in Molecular Medicine* **11**, 1–25 (2009).
 49. Turcotte, C., Blanchet, M. R., Laviolette, M. & Flamand, N. The CB₂ receptor and its role as a regulator of inflammation. *Cellular and Molecular Life Sciences* **73**, 4449–4470 (2016).
 50. Weller, M. G. Quality Issues of Research Antibodies. *Anal. Chem. Insights* **2016**, 21–27 (2016).
 51. Soethoudt, M. *et al.* Selective Photoaffinity Probe That Enables Assessment of Cannabinoid CB₂ Receptor Expression and Ligand Engagement in Human Cells. *J. Am. Chem. Soc.* **140**, 6067–6075 (2018).
 52. Chiurchiù, V., Battistini, L. & Maccarrone, M. Endocannabinoid signalling in innate and adaptive immunity. *Immunology* **144**, 352–364 (2015).
 53. Chiurchiù, V., van der Stelt, M., Centonze, D. & Maccarrone, M. The endocannabinoid system and its therapeutic exploitation in multiple sclerosis: Clues for other neuroinflammatory diseases. *Prog. Neurobiol.* **160**, 82–100 (2018).

54. Basu, S. & Dittel, B. N. Unraveling the complexities of cannabinoid receptor 2 (CB2) immune regulation in health and disease. *Immunol. Res.* **51**, 26–38 (2011).
55. Pacher, P. & Mechoulam, R. Is lipid signaling through cannabinoid 2 receptors part of a protective system? *Prog. Lipid Res.* **50**, 193–211 (2011).
56. Norman, A. W., Mizwicki, M. T. & Norman, D. P. G. Steroid-hormone rapid actions, membrane receptors and a conformational ensemble model. *Nature Reviews Drug Discovery* **3**, 27–41 (2004).
57. Trigunaite, A. & Dimo, J. Suppressing effects of androgens on the immune system. *Cell. Immunol.* **294**, 87–94 (2015).
58. Furman, D. Sexual dimorphism in immunity: Improving our understanding of vaccine immune responses in men. *Expert Review of Vaccines* **14**, 461–471 (2015).
59. Cunningham, M. & Gilkeson, G. Estrogen receptors in immunity and autoimmunity. *Clin. Rev. Allergy Immunol.* **40**, 66–73 (2011).
60. Waldmann, H., Ahmed, S. A. & Khan, D. The immune System is a natural Target for estrogen Action: Opposing effects of estrogen in Two Prototypical Autoimmune Diseases. *Front. Immunol* **6**, 6353389–635 (2016).
61. Homo-Delarche, F. *et al.* Sex steroids, glucocorticoids, stress and autoimmunity. *J. Steroid Biochem. Mol. Biol.* **40**, 619–637 (1991).
62. Lai, J. J. *et al.* Androgen receptor influences on body defense system via modulation of innate and adaptive immune systems: Lessons from conditional AR knockout mice. *American Journal of Pathology* **181**, 1504–1512 (2012).
63. Carlberg, C. Lipid soluble vitamins in gene regulation. *BioFactors* **10**, 91–97 (1999).
64. Ray, R., Bouillon, R., Van Baelen, H. & Holick, M. F. Synthesis of 25-hydroxyvitamin D3 3 β -3'-[N-(4-azido-2-nitrophenyl)amino]propyl ether, a second-generation photoaffinity analogue of 25-hydroxyvitamin D3: Photoaffinity labeling of rat serum vitamin D binding protein. *Biochemistry* **30**, 4809–4813 (1991).
65. Bernstein, P. S., Choi, S. Y., Ho, Y. C. & Rando, R. R. Photoaffinity labeling of retinoic acid-binding proteins. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 654–658 (1995).
66. Chen, G. & Radominska-Pandya, A. Direct photoaffinity labeling of cellular retinoic acid-binding protein I (CRABP-I) with all-trans-retinoic acid: Identification of amino acids in the ligand binding site. *Biochemistry* **39**, 12568–12574 (2000).
67. Shimazawa, R. *et al.* Fluorescent and photoaffinity labeling probes for retinoic acid receptors. *Biochem. Biophys. Res. Commun.* **179**, 259–265 (1991).
68. Erkelens, M. N. & Mebius, R. E. Retinoic Acid and Immune Homeostasis: A Balancing Act. *Trends in Immunology* **38**, 168–180 (2017).
69. McDonald, K. G. *et al.* Epithelial expression of the cytosolic retinoid chaperone cellular retinol binding protein II is essential for in vivo imprinting of local gut dendritic cells by luminal retinoids. *Am. J. Pathol.* **180**, 984–997 (2012).
70. Goverse, G. *et al.* Diet-Derived Short Chain Fatty Acids Stimulate Intestinal Epithelial Cells To Induce Mucosal Tolerogenic Dendritic Cells. *J. Immunol.* **198**, 2172–2181 (2017).
71. Grizotte-Lake, M. *et al.* Commensals Suppress Intestinal Epithelial Cell Retinoic Acid Synthesis to Regulate Interleukin-22 Activity and Prevent Microbial Dysbiosis. *Immunity* **49**, 1103–1115 (2018).
72. Iwata, M. Retinoic acid production by intestinal dendritic cells and its role in T-cell trafficking. *Seminars in Immunology* **21**, 8–13 (2009).
73. Moriguchi, S. The role of vitamin E in T-cell differentiation and the decrease of cellular immunity with aging. *BioFactors* **7**, 77–86 (1998).
74. Moriguchi, S. & Kaneyasu, M. Role of Vitamin E in Immune System. *Ser. Rev. J.Clin. Biochem.* **34**, 97–109 (2003).
75. Soontjens, C. D., Rafter, J. J. & Gustafsson, J. Å. Ligands for orphan receptors? *J. Endocrinol.* **150**, 241–257 (1996).
76. Ogawa, N., Sugiyama, T., Morita, M., Suganuma, Y. & Kobayashi, Y. Total Synthesis of Resolvin D5. *J. Org. Chem.* **82**, 2032–2039 (2017).
77. Rodriguez, A. R. & Spur, B. W. First total synthesis of pro-resolving and tissue-regenerative resolvin sulfido-conjugates. *Tetrahedron Lett.* **58**, 1662–1668 (2017).

Chapter 3

Design and Synthesis of an Activity-based Retinoid Probe

Published as part of S.T.A. Koenders *et al.*, *ACS Cent. Sci.*, **5**, 1965-1974 (2019).

Introduction

Retinoic acid was synthesized for the first time in 1946 by the Dutch chemists Arens and Van Dorp while working for Organon International.¹ In the year thereafter the first commercially viable synthesis was developed by Otto Isler while in employment of Hoffman-La Roche.² Progress in the field of organic synthesis opened up new synthesis routes towards retinoids with the development of reactions such as the Wittig reaction and the Julia olefination.³ The synthesis of retinoids on an industrial scale is still performed using these transformations, while research into new routes using cross-couplings or cross-metathesis reactions is still ongoing.⁴

In cells retinoic acid is formed via the oxidation of retinal. The enzymes catalyzing this reaction are called retinaldehyde dehydrogenases. The human genome expresses three of these: ALDH1A1, ALDH1A2 and ALDH1A3.^{5,6} Nucleophilic attack of their catalytic cysteine on the aldehyde of retinal results in a hemithioacetal. Hydride abstraction by nicotinamide adenine dinucleotide (NAD⁺) with concomitant deprotonation results in net dehydrogenation and formation of a thioester adduct.⁵ Hydrolysis of this thioester leads to the formation of retinoic acid.

ALDH1A1 and ALDH1A3 have been reported as cancer stem cell biomarkers⁷ and ALDH1A1 activity confers resistance against chemo- and radiation therapy.^{8–10} The ability to discern the contribution of specific retinaldehyde dehydrogenases to the global aldehyde dehydrogenase (ALDH) activity would be helpful to understand the underlying biology and to determine which ALDHs are potential therapeutic targets in cancer. Retinaldehyde dehydrogenases have a variable and inducible cellular expression pattern. Their activity is regulated by post-translational modifications, such as the phosphorylation of threonine residues by Aurora kinase A¹¹ and acetylation of a lysine residue.¹²

Activity-based protein profiling (ABPP) has become one of the key methodologies to map enzyme activities on a global scale in living systems, such as cells and animals.^{13,14} ABPP is a technology that relies on activity-based chemical probes that covalently and irreversibly react with the catalytic nucleophile in the active site of an enzyme in their native biological context.¹⁵ Since this process requires a catalytically active protein, these chemical probes report on the abundance of active enzymes. ABPP enables the determination of target engagement and selectivity profiling of drug candidates in a physiologically relevant environment, which enhances the therapeutic relevance of the observed interaction profile.^{19–21} An activity-based probe (ABP) generally consists of an electrophilic warhead, a scaffold that is recognised by the protein (family) of interest and a fluorophore or biotin as reporter group. To date, no ABP for retinaldehyde dehydrogenases has been developed, despite the fact that the catalytic mechanism employed by these enzymes during substrate processing involves a covalent intermediate.

The development of an activity-based probe to profile cellular retinaldehyde dehydrogenase activity would enable the study of retinoic acid production in cells and the selectivity profiling of ALDH inhibitors. This chapter describes the design and synthesis of **LEI-945**, a first-in-class activity-based probe based on the natural substrate retinal to profile retinaldehyde dehydrogenases.

Results and discussion

Design of the retinal-based probe LEI-945

Taking advantage of the fact that covalent enzyme-substrate intermediates emerge during catalysis an ABP based on the structure of retinal was designed. It was hypothesized that the aldehyde in retinal could be exchanged with a vinyl ketone to trap the nucleophilic catalytic cysteine through conjugate addition (**Fig. 3.1**).

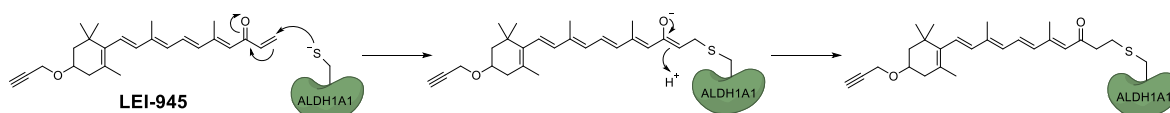


Fig. 3.1 | Covalent interaction of LEI-945 with ALDH1A1. The vinyl ketone warhead of **LEI-945** undergoes nucleophilic attack by the catalytic cysteine of ALDH1A1. The resulting thioether cannot be hydrolysed and results in an irreversible covalent interaction between **LEI-945** and ALDH1A1.

Docking of the retinal-derived vinyl ketone into the crystal structure of ALDH1A1 (PDB: 4WP7)¹⁹ suggested that the nucleophile (the catalytic cysteine thiol) and the electrophile (the unsaturated ketone) would be in close proximity. It pointed as well to a suitable position, which is a solvent exposed position where probe contact with the enzyme active site is negligible, for installation of a bioorthogonal ligation handle (**Fig. 3.2**). An alkyne ligation handle was selected as it is expected to minimally affect the lipophilicity and physico-chemical properties of the probe.^{20,21} This led to the design of **LEI-945** as a potential two-step retinaldehyde dehydrogenase ABP (**Scheme 3.1**).

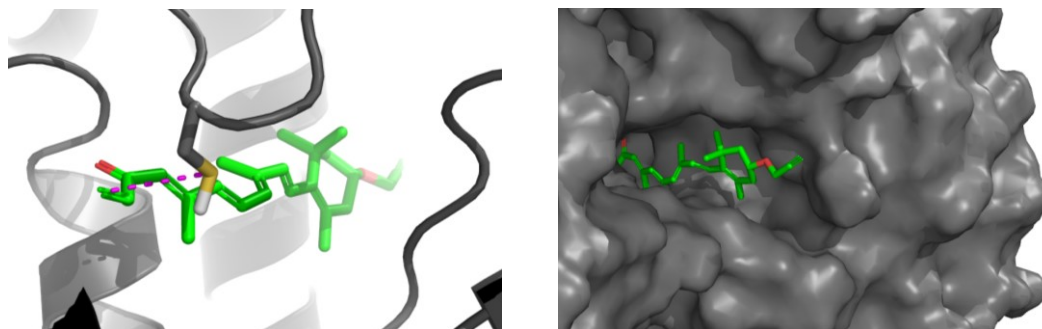
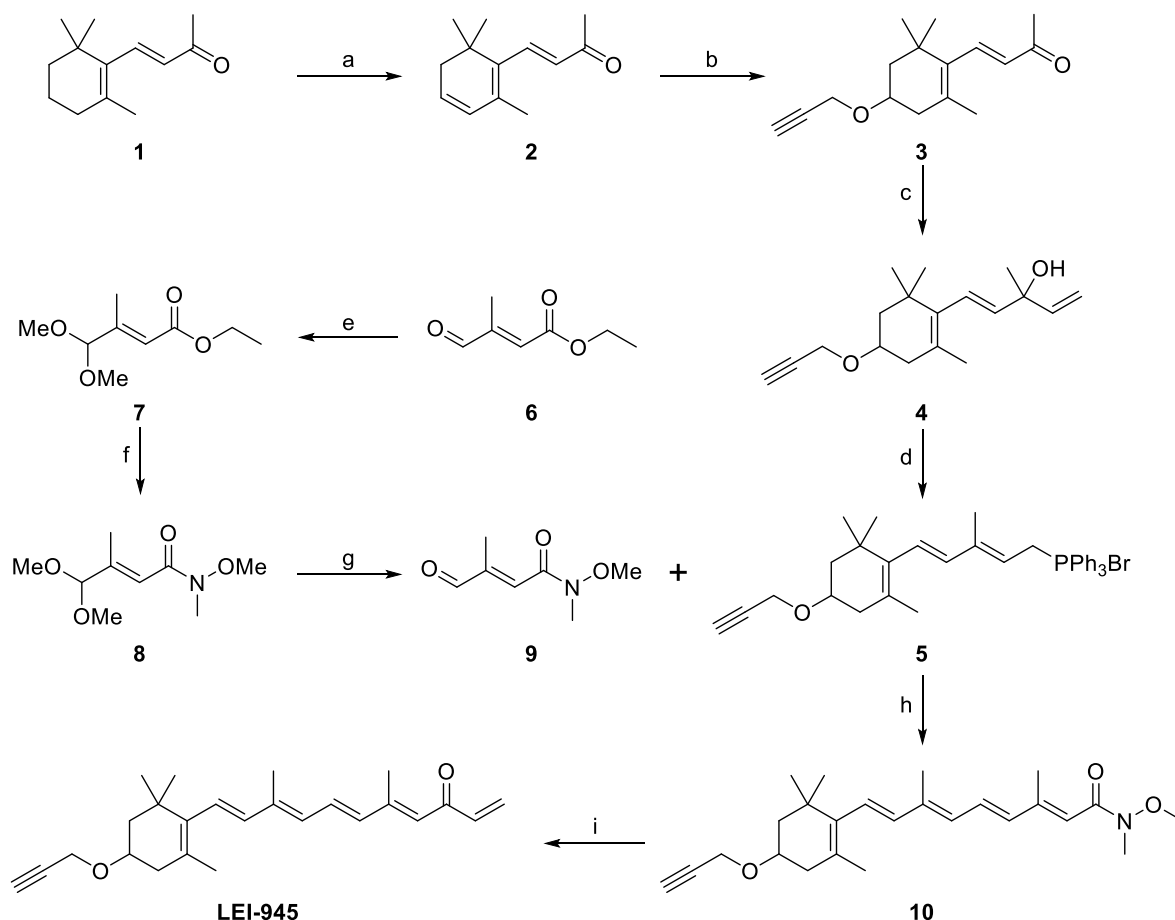


Fig. 3.2 | Docking pose of LEI-945 in the crystal structure of ALDH1A1. The left panel shows the vinyl ketone warhead in close proximity to the catalytic cysteine (yellow; C303) of ALDH1A1. The right panel shows the binding pocket of ALDH1A1 with the alkyne ligation handle solvent exposed.

Synthesis of the retinal-based probe LEI-945

LEI-945 was synthesized using a convergent synthesis starting from commercially available β -ionone **1** and ethyl 3-methyl-4-oxocrotonate **6** (**Scheme 3.1**). Bromination of **1** using recrystallized *N*-bromosuccinimide (NBS), followed by dehydrobromination of 3-bromo- β -ionone with Na_2CO_3 afforded 3-dehydro- β -ionone **2** in 81% yield. As previously described by Henbest *et al.* for this substrate the amount of NBS in this reaction is critical.²² A slight surplus of NBS is necessary to fully consume the starting material as some NBS will be lost in side reactions such as the bromination of the methyl ketone. Too much NBS, however, will lead to more side reactions consuming the product and substrate. Purification of NBS through recrystallization is, therefore, required to permit accurate determination of the amounts of NBS added.

Conjugate addition of propargyl alcohol to 3-dehydro- β -ionone **2** following a modification of the literature procedure²³ provided alkyne **3** in 23% yield. It should be noted that rapid degradation of 3-dehydro- β -ionone **2** was observed even when stored at $-20\text{ }^\circ\text{C}$ and the intermediate should, therefore, be used straightaway. Grignard reaction of **3** and vinylmagnesium bromide quantitatively afforded tertiary alcohol **4**. Addition of triphenylphosphine hydrobromide in methanol provided phosphonium salt **5** in 78% yield.²⁴ Intermediate **9** required for the key Wittig reaction towards the retinal scaffold was synthesised from commercially available ethyl 3-methyl-4-oxocrotonate **6** as follows. Treatment of **6** with trimethyl orthoformate and catalytic acid, yielded acetal **7** in quantitative yield. The ester in **7** was converted into Weinreb amide **8** using *N,O*-dimethylhydroxylamine hydrochloride also in quantitative yield.



Scheme 3.1 | Synthesis of retinal-based probe LEI-945. Reagents and conditions: a) NBS, AIBN, CCl₄, 80 °C, 2 h, then Na₂CO₃, DMF, 81% over two steps; b) H₂SO₄, propargyl alcohol, 4 °C, 18 h, 23%; c) vinylmagnesium bromide, THF, 18 h, quant.; d) PPh₃HBr, 18 h, 78%; e) trimethyl orthoformate, *p*TsOH, MeOH, 2 h, quant.; f) *N,O*-dimethylhydroxylamine hydrochloride, *i*-PrMgCl, THF, 3 h, quant.; g) TFA, DCM/H₂O, 2 h, quant.; h) *n*-BuLi, THF, -78 °C to rt, 2 h, 46% (1:1 E/Z); i) vinyl magnesium bromide, THF, 2 h, 43% (7:3 E/Z).

The acetal in **8** was quantitatively deprotected using aqueous TFA, yielding aldehyde **9**. Phosphonium salt **5** was deprotonated at -78 °C using *n*-BuLi after which addition of aldehyde **9** and stirring at room temperature afforded compound **10** as a 1:1 E/Z mixture. Finally, treatment of Weinreb amide **10** with vinylmagnesium bromide furnished **LEI-945** in 20% yield over the last two steps. The E/Z mixture was inseparable and the product prone to isomerization under the influence of light. As retinoids are rapidly converted *in vivo* into their biological equilibrium of stereoisomers, **LEI-945** was used as the reported E/Z mixture.^{25,26}

The retinal-based probe **LEI-945** was stored in aliquots as DMSO stocks at -80 °C under nitrogen. When stored under these conditions stocks can be kept for at least up to two years. DMSO was chosen instead of ethanol, because ethanol is converted into acetaldehyde *in situ*. As acetaldehyde is also a substrate of ALDH1A1 this might interfere with efficient probe labelling.

Conclusion

40

Acknowledgements

Lindsey Burggraaff is kindly acknowledged for performing the docking studies.

Experimental procedures

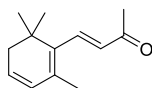
Docking

All calculations were performed using the Schrödinger Suite.²⁷ The X-ray structure of ALDH1A1 was extracted from the PDB (PDB ID: 4WP7).^{19,28} The centroid of the co-crystallized inhibitor CM026 was used to define the binding pocket coordinates. The protein was prepared for docking using the protein preparation tool: metals and ions were removed, and hydrogens and missing side chains were added. Retinoic acid and **LEI-945** were docked into the CM026 binding pocket of ALDH1A1. A constraint was used in docking of retinoic acid for hydrogen bond formation with either the backbone of Cys303 or the side chain of Asn170.

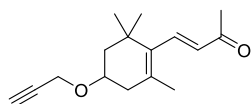
Synthetic methods

General remarks. All reactions were performed using oven or flame-dried glassware and dry solvents. Reagents were purchased from Sigma Aldrich, Acros, Biosolve, VWR, Fluka, Fischer Scientific and Merck and used as received unless stated otherwise. Tetrahydrofuran (THF) and *N,N*-dimethylformamide (DMF) were stored over 4 Å molecular sieves before use. All moisture sensitive reactions were performed under a nitrogen atmosphere. TLC analysis was performed using Merck aluminium sheets (TLC silica gel 60/Kieselguhr F₂₅₄). Compounds were visualized using a solution of KMnO₄ (7.5 g), K₂CO₃ (50 g), 10% NaOH (6 mL) in H₂O (1 L). Column chromatography was performed using Screening Device B.V. silica gel (particle size 40 – 63 µm, pore diameter of 60 Å) with the indicated eluents. ¹H- and ¹³C-NMR spectra were recorded on Bruker AV-400 (400 MHz and 101 MHz, respectively) or Bruker AV-500 MHz (500 MHz and 150 MHz, respectively) using CDCl₃ as solvent. Chemical shifts are reported in ppm (δ) relative to the residual solvent peak or tetramethylsilane. Coupling constants are given in Hz. High-resolution mass spectrometry (HRMS) analysis was performed with a LTQ Orbitrap mass spectrometer (Thermo Finnigan), equipped with an electrospray ion source in positive mode (source voltage 3.5 kV, sheath gas flow 10 mL/min, capillary temperature 250 °C) with resolution R = 60000 at m/z 400 (mass range m/z = 150 – 2000) and dioctyl phthalate (m/z = 391.28428) as a “lock mass”, or with a Synapt G2-Si (Waters), equipped with an electrospray ion source in positive mode (ESI-TOF), injection via NanoEquity system (Waters), with LeuEnk (m/z = 556.2771) as “lock mass”. Eluents used: MeCN:H₂O (1:1 v/v) supplemented with 0.1% formic acid. The high-resolution mass spectrometers were calibrated prior to measurements with a calibration mixture (Thermo Finnigan). The retinoids were handled under dark conditions using amber coloured flasks or aluminium foil when containing more than 3 conjugated double bonds. Retinoid intermediates were stored in the dark under nitrogen at -30 °C and final compounds as powder or as DMSO stock at -80 °C under nitrogen atmosphere. For further information on retinoid handling and storage we refer to the review from Barua and Furr.²⁹

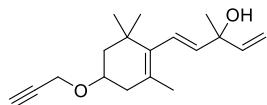
(*E*)-4-(2,6,6-Trimethylcyclohexa-1,3-dien-1-yl)but-3-en-2-one (2):



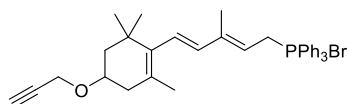
A solution of purified β-ionone (1.0 g, 5.2 mmol), recrystallized NBS (1.2 g, 6.8 mmol) and AIBN (43 mg, 0.26 mmol) in tetrachloromethane (50 mL) was refluxed at 80 °C for 2 hours and was then allowed to cool down.²² The solution was filtered and then DMF (12.5 mL) and Na₂CO₃ (1.3 g) were added. The tetrachloromethane was distilled off and the mixture cooled to room temperature. Et₂O (100 mL) was added and the solution was filtered. The organic layer was washed with 1 M HCl and brine, dried over MgSO₄, filtered and concentrated under reduced pressure. Purification of the residue by column chromatography (Et₂O/pentane) afforded the title compound **2** (0.80 g, 4.2 mmol, 81%) as an orange oil. *R*_f (10% Et₂O in pentane) = 0.45. ¹H NMR (500 MHz, CDCl₃) δ 7.27 (d, *J* = 16.4 Hz, 1H), 6.20 (d, *J* = 16.4 Hz, 1H), 5.88 (s, 2H), 2.31 (s, 3H), 2.12 (d, *J* = 2.2 Hz, 2H), 1.91 (s, 3H), 1.08 (s, 6H). ¹³C NMR (126 MHz, CDCl₃) δ 198.3, 135.9, 133.2, 132.7, 130.3, 129.6, 128.2, 50.4, 39.9, 32.8, 26.5, 20.3.

(E)-4-(2,6,6-Trimethyl-4-(prop-2-yn-1-yloxy)cyclohex-1-en-1-yl)but-3-en-2-one (3):

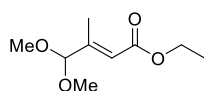
Compound **2** (11.6 g, 61 mmol) was dissolved in propargyl alcohol (180 mL, 3.2 mol) and stirred at 4 °C under a nitrogen atmosphere.²³ Then concentrated sulfuric acid (2.0 mL, 38 mmol) was added and the reaction was stirred overnight at 4 °C. The reaction mixture was then poured onto ice water and then 50% NaOH aq. was added. After stirring for several minutes the mixture was extracted with Et₂O and the organic layer washed with brine, dried over MgSO₄, filtered and concentrated under reduced pressure. Purification of the residue by column chromatography (EtOAc/pentane) afforded the title compound **3** (3.5 g, 14 mmol, 23%; 32% yield based on 29% starting material recovered) as an orange oil and recovered compound **2** (3.3 g, 17 mmol, 29%). *R_f* (10% EtOAc in pentane) = 0.45. ¹H NMR (400 MHz, CDCl₃) δ 7.21 (d, *J* = 16.4 Hz, 1H), 6.11 (d, *J* = 16.4 Hz, 1H), 4.22 (d, *J* = 2.4 Hz, 2H), 3.92 – 3.82 (m, 1H), 2.47 – 2.41 (m, 2H), 2.30 (s, 3H), 2.13 – 2.10 (m, 1H), 1.90 – 1.84 (m, 1H), 1.78 (s, 3H), 1.50 – 1.41 (m, 1H), 1.13 (s, 3H), 1.11 (s, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 198.5, 142.2, 135.7, 132.4, 132.1, 80.2, 74.1, 71.0, 55.3, 44.5, 39.6, 36.6, 30.0, 28.5, 27.3, 21.6. HRMS (ESI) *m/z*: [M + H]⁺ calculated for C₁₆H₂₂O₂: 247.16926, found 247.16918.

(E)-3-Methyl-1-(2,6,6-trimethyl-4-(prop-2-yn-1-yloxy)cyclohex-1-en-1-yl)Penta-1,4-dien-3-ol (4):

To a stirred solution of **3** (3.0 g, 12 mmol) in dry THF (120 mL) at 0 °C under nitrogen atmosphere was added vinylmagnesium bromide (1 M solution in THF, 18 mL, 18 mmol). The mixture was stirred overnight at room temperature and was then cooled to 0 °C. Sat. NH₄Cl aq. was added and the mixture was stirred for 30 minutes. The mixture was extracted with Et₂O. The organic layer washed with brine and concentrated under reduced pressure affording the title compound **4** (3.5 g, 12 mmol, quant.) as an orange oil. *R_f* (10% EtOAc in pentane) = 0.4. ¹H NMR (400 MHz, CDCl₃) δ 6.09 – 5.95 (m, 2H), 5.55 (dd, *J* = 16.2, 1.2 Hz, 1H), 5.26 (dd, *J* = 17.3, 1.2 Hz, 1H), 5.08 (dd, *J* = 10.6, 1.2 Hz, 1H), 4.22 (d, *J* = 2.4 Hz, 2H), 3.90 – 3.80 (m, 1H), 2.43 – 2.41 (m, 1H), 2.41 – 2.34 (m, 1H), 2.07 – 1.98 (m, 1H), 1.87 – 1.80 (m, 1H), 1.68 (s, 3H), 1.42 (s, 4H), 1.05 (s, 3H), 1.03 (s, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 144.0, 139.7, 136.8, 125.3, 124.8, 112.2, 80.3, 73.9, 73.5, 71.6, 55.1, 44.2, 38.8, 36.6, 30.0, 28.3, 28.0, 21.2. HRMS (ESI) *m/z*: [M + H]⁺ calculated for C₁₈H₂₆O₂·H₂O: 257.18999, found: 257.18994.

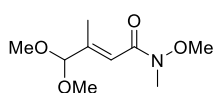
Bromo((2E,4E)-3-methyl-5-(2,6,6-trimethyl-4-(prop-2-yn-1-yloxy)cyclohex-1-en-1-yl)Penta-2,4-dien-1-yl)triphenyl-λ⁵-phosphane (5):

To a stirred mixture of **4** (3.0 g, 11 mmol) in MeOH (60 mL) was added triphenylphosphine hydrobromide (4.1 g, 12 mmol).²⁴ The reaction mixture was stirred overnight at room temperature and then concentrated under reduced pressure. H₂O was added and the product extracted with DCM. The organic layer was washed with brine, dried over Na₂SO₄, filtered, concentrated and purified by column chromatography (EtOAc followed by MeOH) affording title compound **5** (5.1 g, 8.5 mmol, 78%) as a yellow/orange foam. *R_f* (10% MeOH in DCM) = 0.6. ¹H NMR (400 MHz, CDCl₃) δ 7.92 – 7.86 (m, 6H), 7.80 – 7.77 (m, 3H), 7.70 – 7.67 (m, 6H), 5.93 (s, 2H), 5.41 – 5.32 (m, 1H), 5.07 – 4.86 (m, 2H), 4.20 (d, *J* = 2.4 Hz, 2H), 3.82 (dddd, *J* = 11.9, 9.2, 5.5, 3.4 Hz, 1H), 2.41 (t, *J* = 2.4 Hz, 1H), 2.40 – 2.33 (m, 1H), 2.05 – 1.97 (m, 1H), 1.84 – 1.78 (m, 1H), 1.64 (s, 3H), 1.44 – 1.39 (m, 1H), 1.37 (d, *J* = 3.5 Hz, 3H), 1.01 (s, 3H), 0.99 (s, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 143.5 (d, *J* = 13.8 Hz), 137.0, 136.3, 134.8, 133.9 (d, *J* = 9.7 Hz), 130.2 (d, *J* = 12.4 Hz), 127.5, 126.3, 118.3 (d, *J* = 85.2 Hz), 113.4 (d, *J* = 11.4 Hz), 73.9, 71.4, 55.1, 44.1, 39.0, 36.6, 30.0, 28.5, 24.9 (d, *J* = 48.5 Hz), 21.5, 12.8. HRMS (ESI) *m/z*: [M]⁺ calculated for C₃₆H₄₀OP⁺: 519.28113, found: 519.28065.

Ethyl (E)-4,4-dimethoxy-3-methylbut-2-enoate (7):

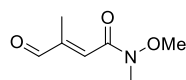
To a stirred solution of trimethyl orthoformate (8.6 mL, 77 mmol) and ethyl 3-methyl-4-oxocrotonate **6** (10 g, 70 mmol) in MeOH (40 mL) was added *p*TsOH (268 mg, 1.4 mmol). The mixture was stirred for 2 hours and then H₂O (100 mL) was added. The product was extracted with Et₂O (3x200 mL). The combined organic layers were washed with NaHCO₃ aq., brine and dried over Na₂SO₄, filtered and carefully concentrated affording title compound **7** (13.1 g, 69 mmol, quant.) as a colourless oil. *R_f* (10% Et₂O in pentane) = 0.9. ¹H NMR (400 MHz, CDCl₃) δ 6.03 (s, 1H), 4.61 (s, 1H), 4.18 (q, *J* = 7.1 Hz, 2H), 3.31 (s, 6H), 2.11 (s, 3H), 1.29 (t, *J* = 7.2 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 166.2, 152.2, 118.3, 104.8, 59.7, 53.1, 14.1, 13.8.

(E)-N,4,4-Trimethoxy-N,3-dimethylbut-2-enamide (8):



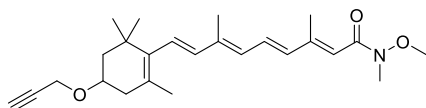
To a stirred solution of **7** (2.0 g, 11 mmol) in dry THF (30 mL) under N₂ at 0 °C was added *N,O*-dimethylhydroxylamine hydrochloride (1.6 g, 16 mmol) followed by *i*-PrMgCl (16 mL, 32 mmol, 2 M in THF). The reaction was stirred for 3 hours at room temperature and was then quenched with sat. NH₄Cl aq. The mixture was extracted with Et₂O and the organic layer washed with brine, dried over Na₂SO₄, filtered and concentrated affording title compound **8** (2.2 g, 11 mmol, quant.) as a yellowish oil. *R*_f (10% Et₂O in pentane) = 0.05; ¹H NMR (400 MHz, CDCl₃) δ 6.47 (s, 1H), 4.64 (s, 1H), 3.69 (s, 3H), 3.32 (s, 6H), 3.23 (s, 3H), 2.06 (s, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 117.1, 105.2, 67.9, 61.5, 53.2, 14.1. HRMS (ESI) *m/z*: [M + H]⁺ calculated for C₉H₁₇NO₄: 204.12303, found: 204.12299.

(E)-N-methoxy-N,3-dimethyl-4-oxobut-2-enamide (9):



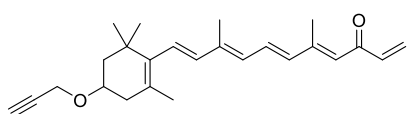
To a stirred solution of **8** (0.50 g, 2.5 mmol) in DCM (20 mL) at room temperature was added a solution of TFA (2.5 mL) in H₂O (2.5 mL). The reaction mixture was stirred for 2 hours and then K₂CO₃ was added until the pH was 8. The water layer was then extracted with DCM. The combined organic layers washed with brine, dried over MgSO₄, filtered and concentrated affording title compound **9** (378 mg, 2.5 mmol, quant.) as a yellow oil. *R*_f (50% EtOAc in pentane) = 0.5. ¹H NMR (400 MHz, CDCl₃) δ 9.58 (s, 1H), 7.01 (s, 1H), 3.74 (s, 3H), 3.30 (s, 3H), 2.11 (s, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 194.6, 165.8, 147.9, 135.3, 61.9, 31.9, 10.7. HRMS (ESI) *m/z*: [M + H]⁺ calculated for C₇H₁₁NO₃: 158.08117, found 158.08116.

(2E,4E/Z,6E,8E)-N-Methoxy-N,3,7-trimethyl-9-(2,6,6-trimethyl-4-(prop-2-yn-1-yloxy)cyclohex-1-en-1-yl)nona-2,4,6,8-tetraenamide (10):

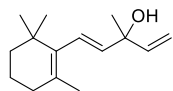


To a stirred solution of **5** (0.50 g, 0.83 mmol) in dry THF (4 mL) at -78 °C under N₂ was added *n*-BuLi (0.5 mL, 1.6 M in hexane, 0.8 mmol). The mixture was stirred for 30 minutes at -78 °C and then **9** (0.12 g, 0.76 mmol) dissolved in dry THF (1 mL) was added. The reaction was stirred for 1 hour at -78 °C and was then allowed to reach room temperature in 2 hours. Then the reaction was quenched with sat. NH₄Cl aq. and extracted with Et₂O. The combined organic layers were washed with H₂O and brine, dried over MgSO₄, filtered and concentrated. Purification of the residue by column chromatography (EtOAc/pentane) afforded title compound **10** (139 mg, 0.35 mmol, 46%; 1:1 4E/Z) as a yellow oil. *R*_f (50% Et₂O in pentane) = 0.8. NMR spectra are obtained from the mixture of stereoisomers. ¹H NMR (400 MHz, CDCl₃) δ 6.99 – 5.92 (m, 6H), 4.22 (s, 2H), 3.92 – 3.81 (m, 1H), 3.68 (d, *J* = 8.3 Hz, 3H), 3.23 (s, 3H), 2.47 – 2.38 (m, 2H), 2.35 – 2.28 (m, 3H), 2.13 – 2.06 (m, 1H), 2.01 – 1.95 (m, 3H), 1.88 – 1.81 (m, 1H), 1.74 – 1.70 (m, 3H), 1.50 – 1.40 (m, 1H), 1.07 (s, 6H). ¹³C NMR (101 MHz, CDCl₃) δ 168.1, 138.7, 138.4, 138.3, 138.1, 137.6, 132.7, 130.1, 129.7, 127.9, 127.1, 127.0, 126.6, 126.4, 216.3, 118.0, 117.5, 80.3, 73.9, 71.5, 61.5, 55.1, 44.4, 39.2, 36.8, 30.2, 28.6, 21.6, 19.1, 13.9, 12.8, 12.4. HRMS (ESI) *m/z*: [M + H]⁺ calculated for C₂₅H₃₅NO₃: 398.26897, found 398.26875.

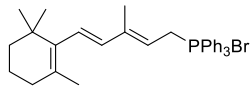
(4E,6E/Z,8E,10E)-5,9-Dimethyl-11-(2,6,6-trimethyl-4-(prop-2-yn-1-yloxy)cyclohex-1-en-1-yl)undeca-1,4,6,8,10-Pentaen-3-one (LEI-945):



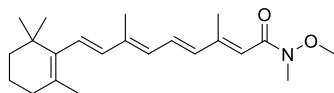
To a solution of **10** (0.10 g, 0.25 mmol) in THF (1 mL) at 0 °C under N₂ was added vinyl magnesium bromide (0.27 mmol, 0.30 mL) dropwise. The solution was then stirred for 2 hours at room temperature and then quenched with sat. NH₄Cl aq. The mixture was then extracted with Et₂O and the organic layer washed with brine, dried over Na₂SO₄, filtered and concentrated. Purification of the residue by column chromatography (EtOAc/pentane) afforded title compound **LEI-945** (39 mg, 0.11 mmol, 43%; 7:3 6E/Z) as a yellow oil. The E/Z mixture was inseparable and prone to further isomerization under the influence of light. As retinoids are rapidly converted *in vivo* into their biological equilibrium of stereoisomers, **LEI-945** was used as the reported E/Z mixture.^{25,26} *R*_f (10% EtOAc in pentane) = 0.9. NMR spectra are obtained from the mixture of stereoisomers. ¹H NMR (400 MHz, CDCl₃) δ 7.19 – 6.96 (m, 1H), 6.61 – 5.96 (m, 7H), 5.80 – 5.70 (m, 1H), 4.22 (s, 2H), 3.92 – 3.82 (m, 1H), 2.47 – 2.41 (m, 2H), 2.40 – 2.35 (m, 3H), 2.14 – 2.05 (m, 1H), 2.02 – 1.97 (m, 3H), 1.89 – 1.82 (m, 1H), 1.74 (s, 3H), 1.49 – 1.41 (m, 1H), 1.08 (s, 6H). ¹³C NMR (101 MHz, CDCl₃) δ 190.4, 153.1, 139.8, 138.6, 138.5, 138.3, 138.0, 137.6, 136.1, 132.4, 132.3, 130.1, 129.8, 127.9, 127.8, 127.1, 126.8, 126.5, 125.0, 124.7, 86.8, 73.9, 71.5, 55.2, 44.4, 39.3, 36.8, 30.2, 28.6, 21.7, 14.5, 12.9. HRMS (ESI) *m/z*: [M + H]⁺ calculated for C₂₅H₃₂O₂: 365.24751, found 365.24749.

(*E*)-3-Methyl-1-(2,6,6-trimethylcyclohex-1-en-1-yl)Penta-1,4-dien-3-ol (11):

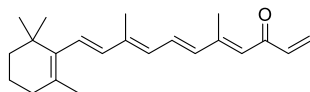
The title compound **11** was synthesized from β -ionone **1** (0.80 g, 4.2 mmol) according to the procedure described for compound **4**. This yielded **11** (0.88 g, 4.0 mmol, quant.) as a yellow oil. R_f (10% Et₂O in pentane) = 0.3. ¹H NMR (400 MHz, CDCl₃) δ 6.12 – 5.94 (m, 2H), 5.57 – 5.49 (m, 1H), 5.32 – 5.21 (m, 2H), 5.12 – 5.03 (m, 1H), 1.97 (t, J = 5.2 Hz, 2H), 1.66 (s, 3H), 1.62 – 1.57 (m, 2H), 1.47 – 1.43 (m, 2H), 1.42 (s, 3H), 0.98 (s, 6H). ¹³C NMR (101 MHz, CDCl₃) δ 144.2, 138.9, 136.8, 128.4, 125.7, 112.0, 73.6, 39.3, 34.0, 32.6, 28.7, 28.0, 21.3, 19.2.

Bromo((2*E*,4*E*)-3-methyl-5-(2,6,6-trimethylcyclohex-1-en-1-yl)Penta-2,4-dien-1-yl)triphenyl- λ^5 -phosphane (12):

The title compound **12** was synthesized from **11** (0.88 g, 4.0 mmol) according to the procedure described for compound **5**. This yielded **12** (1.95 g, 3.6 mmol, 90%) as a yellow foam. R_f (10% MeOH in DCM) = 0.5. ¹H NMR (400 MHz, CDCl₃) δ 7.92 – 7.70 (m, 15H), 6.07 – 5.89 (m, 2H), 5.43 – 5.35 (m, 1H), 4.89 – 4.75 (m, 2H), 1.97 (s, 2H), 1.63 (s, 3H), 1.58 (s, 2H), 1.46 – 1.40 (m, 2H), 1.37 (d, J = 3.2 Hz, 3H), 0.96 (s, 6H). ¹³C NMR (101 MHz, CDCl₃) δ 143.1 (d, J = 13.7 Hz), 136.4, 134.4, 133.2 (d, J = 9.7 Hz), 129.7 (d, J = 12.4 Hz), 128.9, 127.9, 117.4 (d, J = 85.2 Hz), 112.1 (d, J = 11.6 Hz), 38.6, 33.4, 32.1, 28.2, 24.1 (d, J = 48.9 Hz), 21.0, 18.5, 12.1.

(2*E*,4*E*/*Z*,6*E*,8*E*)-*N*-Methoxy-*N*,3,7-trimethyl-9-(2,6,6-trimethylcyclohex-1-en-1-yl)nona-2,4,6,8-tetraenamide (13):

The title compound **13** was synthesized from **9** (0.13 g, 0.83 mmol) and **12** (0.50 g, 0.92 mmol) as described for compound **10**. This yielded **13** (0.12 g, 0.36 mmol, 43%; 1:1 4*E*/*Z* mixture) as a yellow oil. R_f (50% EtOAc in pentane) = 0.8. NMR spectra are obtained from the mixture of stereoisomers. ¹H NMR (500 MHz, CDCl₃) δ 6.99 – 5.89 (m, 6H), 3.67 (s, 3H), 3.23 (s, 3H), 2.35 – 2.29 (m, 3H), 2.04 – 2.01 (m, 2H), 2.00 – 1.96 (m, 3H), 1.72 – 1.69 (m, 3H), 1.65 – 1.58 (m, 2H), 1.50 – 1.45 (m, 2H), 1.04 – 1.01 (m, 6H). ¹³C NMR (126 MHz, CDCl₃) δ 168.2, 139.0, 138.8, 137.7, 137.6, 137.4, 136.1, 132.3, 129.9, 129.8, 129.6, 128.2, 128.0, 126.1, 117.9, 117.3, 61.6, 60.4, 39.5, 34.2, 33.0, 28.9, 21.7, 19.2, 19.2, 13.9, 12.8, 12.4. HRMS (ESI) m/z : $[M + H]^+$ calculated for C₂₂H₃₃NO₂: 344.25841, found 344.25830.

(4*E*,6*E*/*Z*,8*E*,10*E*)-5,9-Dimethyl-11-(2,6,6-trimethylcyclohex-1-en-1-yl)undeca-1,4,6,8,10-Pentaen-3-one (STA-186):

The title compound **STA-186** was synthesized from **13** (100 mg, 0.29 mmol) as described for compound **LEI-945**. This yielded **STA-186** (25.7 mg, 83 μ mol, 29%; 2:1 6*E*/*Z* mixture). R_f (20% EtOAc in pentane) = 0.9. NMR spectra are obtained from the mixture of stereoisomers. ¹H NMR (400 MHz, CDCl₃) δ 7.14 – 7.04 (m, 1H), 6.53 – 6.12 (m, 7H), 5.75 – 5.70 (m, 1H), 2.39 – 2.10 (m, 3H), 2.06 – 1.99 (m, 5H), 1.77 – 1.71 (m, 3H), 1.66 – 1.59 (m, 2H), 1.51 – 1.44 (m, 2H), 1.03 (s, 6H). ¹³C NMR (101 MHz, CDCl₃) δ 190.1, 152.9, 139.9, 138.2, 137.4, 136.9, 135.4, 132.2, 129.3, 128.8, 126.4, 124.2, 122.4, 39.3, 33.9, 32.8, 28.6, 21.4, 18.9, 14.2, 12.6. HRMS (ESI) m/z : $[M + H]^+$ calculated for C₂₂H₃₀O: 311.23694, found 311.23667.

References

1. van Dorp, D. A. & Arens, J. F. The synthesis of “vitamin A acid”, a biologically active substance. *Recl. des Trav. Chim. des Pays-Bas* **65**, 338–345 (1946).
2. Isler, O., Huber, W., Ronco, A. & Kofler, M. Synthese des Vitamin A. *Helv. Chim. Acta* **30**, 1911–1927 (1947).
3. Parker, G. L., Smith, L. K. & Baxendale, I. R. Development of the industrial synthesis of vitamin A. *Tetrahedron* **72**, 1645–1652 (2016).
4. Álvarez, R., Vaz, B., Gronemeyer, H. & De Lera, R. A. Functions, therapeutic applications, and synthesis of retinoids and carotenoids. *Chemical Reviews* **114**, 1–125 (2014).
5. Koppaka, V. *et al.* Aldehyde Dehydrogenase Inhibitors: a Comprehensive Review of the Pharmacology, Mechanism of Action, Substrate Specificity, and Clinical Application. *Pharmacol. Rev.* **64**, 520–539 (2012).
6. Duester, G., Mic, F. A. & Molotkov, A. Cytosolic retinoid dehydrogenases govern ubiquitous metabolism of retinol to retinaldehyde followed by tissue-specific metabolism to retinoic acid. in *Chemico-Biological Interactions* **143–144**, 201–210 (Elsevier, 2003).
7. Luo, Y. *et al.* ALDH1A isozymes are markers of human melanoma stem cells and potential therapeutic targets. *Stem Cells* **30**, 2100–2113 (2012).
8. Qiu, Y. *et al.* The expression of aldehyde dehydrogenase family in breast cancer. *J. Breast Cancer* **17**, 54–60 (2014).
9. Sládek, N. E., Kollander, R., Sreerama, L. & Kiang, D. T. Cellular levels of aldehyde dehydrogenases (ALDH1A1 and ALDH3A1) as predictors of therapeutic responses to cyclophosphamide-based chemotherapy of breast cancer: A retrospective study. *Cancer Chemother. Pharmacol.* **49**, 309–321 (2002).
10. Croker, A. K. & Allan, A. L. Inhibition of aldehyde dehydrogenase (ALDH) activity reduces chemotherapy and radiation resistance of stem-like ALDH hiCD44 + human breast cancer cells. *Breast Cancer Res. Treat.* **133**, 75–87 (2012).
11. Wang, J. *et al.* Phosphorylation-dependent regulation of ALDH1A1 by Aurora kinase A: Insights on their synergistic relationship in pancreatic cancer. *BMC Biol.* **15**, 1–22 (2017).
12. Zhao, D. *et al.* NOTCH-induced aldehyde dehydrogenase 1A1 deacetylation promotes breast cancer stem cells. *J. Clin. Invest.* **124**, 5453–5465 (2014).
13. Serwa, R. & Tate, E. W. Activity-based profiling for drug discovery. *Chemistry and Biology* **18**, 407–409 (2011).
14. Cravatt, B. F., Wright, A. T. & Kozarich, J. W. Activity-Based Protein Profiling: From Enzyme Chemistry to Proteomic Chemistry. *Annu. Rev. Biochem.* **77**, 383–414 (2008).
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. 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).
17. 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).
18. Ahn, K. *et al.* Discovery and Characterization of a Highly Selective FAAH Inhibitor that Reduces Inflammatory Pain. *Chem. Biol.* **16**, 411–420 (2009).
19. Morgan, C. A. & Hurley, T. D. Characterization of two distinct structural classes of selective aldehyde dehydrogenase 1A1 inhibitors. *J. Med. Chem.* **58**, 1964–1975 (2015).
20. Thiele, C. *et al.* Tracing fatty acid metabolism by click chemistry. *ACS Chem. Biol.* **7**, 2004–2011 (2012).
21. Gaebler, A. *et al.* Alkyne lipids as substrates for click chemistry-based in vitro enzymatic assays. *J. Lipid Res.* **54**, 2282–2290 (2013).
22. Henbest, H. B. 237. Studies in the polyene series. Part XXXVII. Preparation of 3-dehydro- β -ionone and some 3-substituted β -ionones. *J. Chem. Soc.* **0**, 1074–1078 (1951).
23. Stjrmatis, J. D. & Thommen, R. A Total Synthesis of Astaxanthin Dimethyl Ether. *J. Org. Chem.* **32**, 180–184 (1967).
24. Haugan, J. A. *et al.* Total Synthesis of C31-Methyl Ketone Apocarotenoids: Sintaxanthin and (3R)-3-Hydroxysintaxanthin. *Acta Chem. Scand.* **48**, 657–664 (2008).
25. McBee, J. K., Van Hooser, J. P., Jang, G. F. & Palczewski, K. Isomerization of 11-cis-Retinoids to All-trans-retinoids in Vitro and in Vivo. *J. Biol. Chem.* **276**, 48483–48493 (2001).

CHAPTER 3

26. Kojima, R. *et al.* In vivo isomerization of retinoic acids. Rapid isomer exchange and gene expression. *J. Biol. Chem.* **269**, 32700–32707 (1994).
27. Schrödinger Release 2017-4: LigPrep, Schrödinger, LLC, New York, NY, 2017.
28. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–42 (2000).
29. Barua, A. B. & Furr, H. C. Properties of retinoids. Structure, handling, and preparation. *Applied Biochemistry and Biotechnology - Part B Molecular Biotechnology* **10**, 167–182 (1998).

Chapter 4

Biological Evaluation of LEI-945

Published as part of S.T.A. Koenders *et al.*, *ACS Cent. Sci.*, **5**, 1965-1974 (2019).

Introduction

Aldehyde dehydrogenases (ALDHs) perform an important role in cellular metabolism by converting aldehydes into carboxylic acids, some of which perform important biological roles, such as retinoic acid.¹ Aldehydes are intrinsically reactive and may bind to DNA or proteins, resulting in cellular damage. A buildup of these metabolites in the cell is toxic.² Fast conversion of such compounds into less toxic and easily excreted carboxylic acid metabolites is therefore essential for cell survival and proliferation. The human ALDH family consists of 19 enzymes, each with its own preferred type of substrates and covering a wide range of endogenous and exogenous aldehydes.^{3,4}

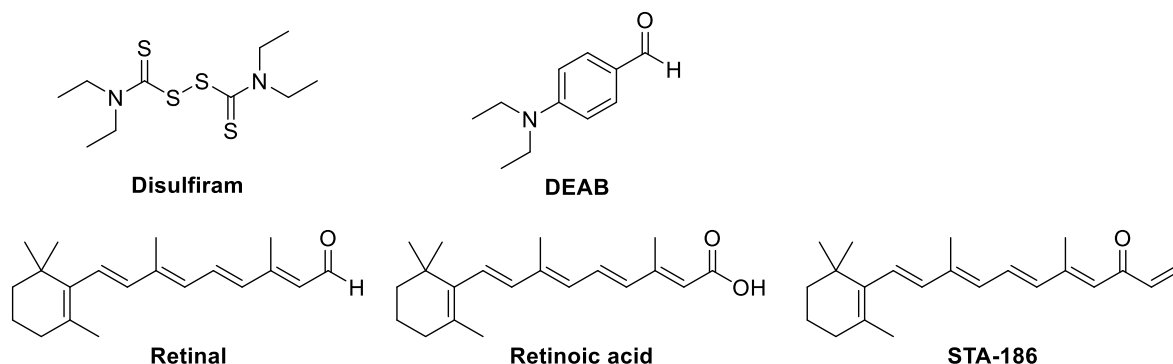


Fig. 4.1 | Chemical structures of ALDH inhibitors and retinoids used in this chapter

Three ALDHs (ALDH1A1, ALDH1A2, ALDH1A3) have been shown to convert retinal into retinoic acid and are therefore classified as retinaldehyde dehydrogenases.^{3,4} Retinoic acid regulates many cellular and physiological functions. Enabling a specific readout of the activity of these enzymes would allow researchers to study the individual contribution of these enzymes to processes such as embryonic development, immunomodulation, neuronal differentiation and (cancer) stem cell proliferation.⁵⁻⁸

Chapter 3 described the design and synthesis of the first-in-class retinal based probe **LEI-945** for the profiling of retinaldehyde dehydrogenases. In this chapter the biological evaluation of activity-based probe (ABP) **LEI-945** is reported.

Results and discussion

LC-UV/MS based assay for in situ retinal conversion

Most biochemical ALDH assays described in literature rely on the change in absorbance of nicotinamide adenine dinucleotide (NAD^+) upon conversion into NADH, which is produced during enzymatic aldehyde dehydrogenation. However, the catalytic cysteine of ALDHs can be easily oxidized in non-cellular environments. To overcome this problem these *in vitro* assays rely on the addition of reducing agents, such as dithiothreitol or tris(2-carboxyethyl)phosphine, to keep the cysteine in its reduced and catalytically active state. Often *in vitro* assays are also performed in buffers with a pH of ≥ 8.0 , further increasing the reactivity of the cysteine. As these assay conditions do not resemble physiologically relevant conditions, there is a need for a cellular assay that reports directly on the formation of retinoic acid by ALDH enzymes.

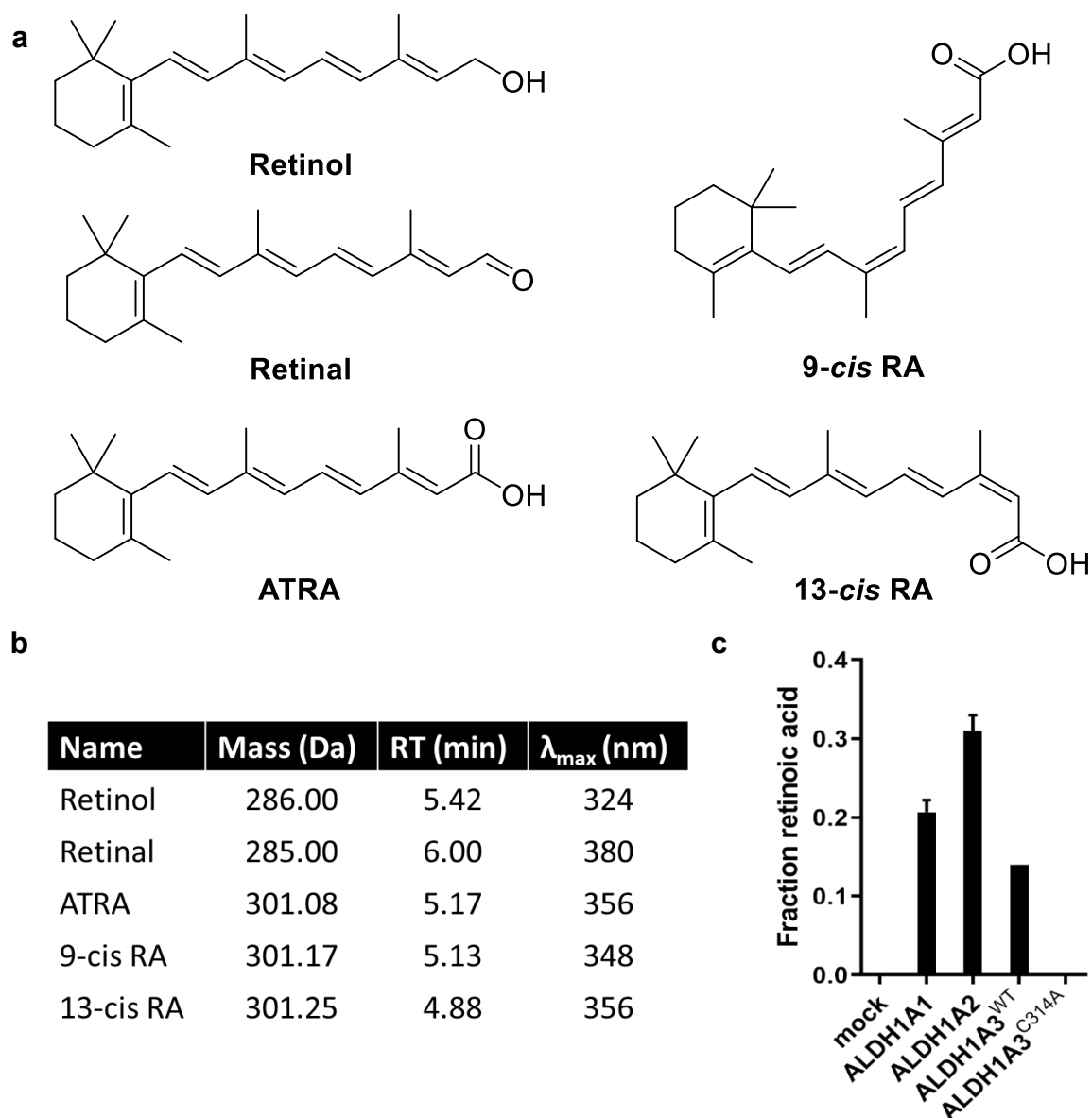


Fig. 4.2 | Chemical structures of retinoid standards and data from *in situ* retinal conversion assay. **a**, Chemical structures of retinoid standards. **b**, Detected values for each of the retinoid standards, detected mass, retention time (RT) and wavelength maximum ($\lambda_{\max}$). **c**, Retinoic acid production by U2OS cells overexpressing ALDHs. Data represent mean values $\pm$ SD; $N = 3$ biological replicates measured twice.

To determine whether recombinant ALDH enzymes are active and capable of converting retinal, an *in situ* substrate assay was developed. It was envisioned that a LC-UV/MS-based method could be established to measure the cellular conversion of retinal into retinoic acid. First, the retention time and maximum wavelength ($\lambda_{\max}$) of retinoid standards were measured on a LC-UV/MS using a gradient of 10% acetonitrile for 1.5 minutes followed by 90% acetonitrile for 7.5 minutes (**Supplementary Fig. 4.1**).

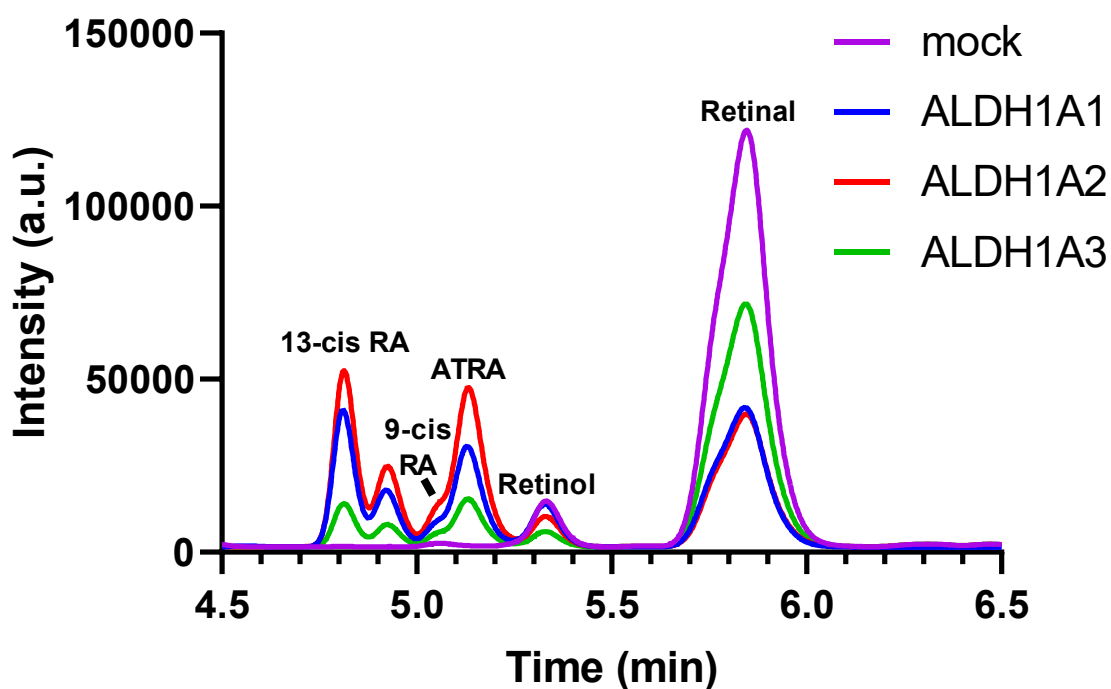


Fig. 4.3 | *In situ* retinal conversion assay using LC-UV/MS. Overlay of UV traces for mock, ALDH1A1, ALDH1A2 and ALDH1A3 overexpressed in U2OS cells.

The λ_{max} values were 324, 380, 356, 348 and 356 nm for retinol, retinal, all-*trans*-retinoic acid (ATRA), 9-*cis* RA and 13-*cis* RA, respectively (**Fig. 4.2b**). This is consistent with reported values in ethanol.⁹ A calibration curve was made using a mixture of the five standards which was measured at 324, 356 and 380 nm. The retention times of the retinoid standards were 5.42, 6.00, 5.17, 5.13 and 4.88 minutes, respectively (**Fig. 4.2b**). The MS data were used to confirm the correct assignment of the peaks (**Fig. 4.2b**).

ALDH1A1, 1A2 and 1A3 were transiently overexpressed in U2OS human osteosarcoma cells. This cell line was used as it showed no endogenous conversion of retinal. Two days after transfection, cells were treated with retinal (30 μM , 4 h) in medium with serum. Cells were then lysed with acetonitrile. Insoluble compounds were removed by centrifugation. The samples were concentrated, dissolved in LC-MS sample solution and measured using the protocol previously described. An overlay of the UV traces for ALDH1A1, 1A2, 1A3 is shown in **Fig. 4.3**.

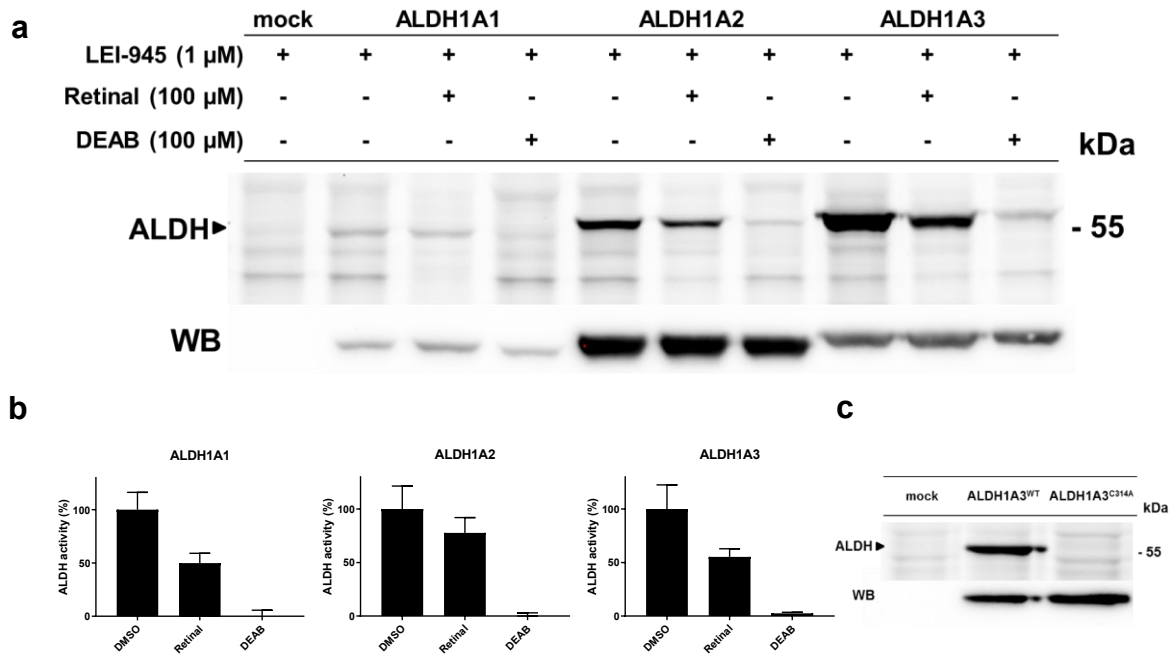


Fig. 4.4 | *In situ* labeling of ALDH by LEI-945 and competition with the natural substrate and pan-ALDH inhibitor DEAB. **a**, *In situ* labeling of ALDH1A1, ALDH1A2, ALDH1A3 transiently transfected in U2OS cells using **LEI-945** (1 μ M) for 1 h at 37 °C. Competition was performed with natural substrate retinal (100 μ M) or general ALDH inhibitor DEAB (100 μ M). **b**, Competition of **LEI-945** labeling of overexpressed ALDH enzymes in U2OS cells. Data represent mean values $\pm$ SD; $N = 3$ experiments per group (biological replicates). **c**, *In situ* labeling of ALDH1A3^{WT} and mutant ALDH1A3^{C314A} with **LEI-945** (1 μ M) for 1 h at 37 °C.

Formation of ATRA and its *cis*-isomers is clearly catalyzed by all three ALDH isozymes. An additional peak is detected at 4.92 minutes with the same mass as retinoic acid, probably corresponding with 11-*cis* retinoic acid. Overexpression of ALDH1A3^{C314} in U2OS confirmed that Cys314 was required for the enzymatic activity of ALDH1A3 (**Fig. 4.2c**). These data show that recombinant ALDH1A1, ALDH1A2 and ALDH1A3 are active and can be used to establish whether **LEI-945** acts as an activity-based probe for retinaldehyde dehydrogenases.

LEI-945 reacts with overexpressed ALDHs in situ to form a covalent adduct

To assess whether **LEI-945** acts as a retinaldehyde dehydrogenase ABP, pcDNA3.1 plasmids containing either recombinant human ALDH1A1, ALDH1A2 or ALDH1A3 fused to a FLAG-tag were transiently transfected in U2OS cells. Living cells were treated with **LEI-945** (1 μ M) in serum free medium for 1 hour, lysed and the probe labelled proteins ligated with Alexa Fluor 647 dye under copper(I)-catalysed azide-alkyne [2+3] cycloaddition conditions.

The protein samples were resolved by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by in-gel fluorescent scanning. A fluorescent band at 55 kDa was apparent in samples harvested from all three transfected cell cultures, but not in the mock-treated cells (**Fig. 4.4a**). All fluorescent bands were at the same height as the signal generated by an α -FLAG antibody in a separate Western blotting experiment. This molecular weight matches with that of the transfected retinaldehyde dehydrogenases, suggesting that **LEI-945** indeed reacted with these proteins in a covalent manner. Fluorescent labeling intensity was reduced by pretreatment with retinal (100 μ M) as well as with the general ALDH inhibitor, DEAB (100 μ M) (**Fig. 4.4a,b**).¹⁰

Site-directed mutagenesis of the catalytic Cys314 into alanine in ALDH1A3, which was chosen as a representative retinaldehyde dehydrogenase, abolished labeling of the enzyme with **LEI-945**, whereas protein expression was not affected as determined by an α -FLAG antibody (**Fig. 4.4c**). Altogether, these data demonstrate that **LEI-945** can efficiently label ALDH1A1, ALDH1A2 and ALDH1A3 in an activity-dependent manner by forming a covalent and irreversible bond with the catalytic cysteine thiol.

Mapping of endogenous ALDHs in A549 lung cancer cells using LEI-945

To establish the potential of **LEI-945** to detect endogenous retinaldehyde dehydrogenases, non-small-cell lung cancer cell line A549, known to express ALDH1A1^{11,12}, was incubated with **LEI-945**. This yielded a fluorescent band at around 55 kDa, which could be significantly and dose-dependently reduced by preincubation with either DEAB, retinal, disulfiram or **STA-186**, with pIC₅₀ values in the range of 5.0-5.9 (**Fig. 4.5**).

To find out which proteins are irreversibly and covalently labelled by **LEI-945**, a label-free chemical proteomics experiment was performed. **LEI-945** (1 μ M, 1 h)-treated A549 cells were lysed and reacted with biotin-N₃, after which probe-labelled proteins were enriched for by streptavidin bead pull-down. Subsequent tryptic digestion, protein identification and quantification by mass spectrometry resulted in the identification 34 significantly enriched proteins, including ALDH1A1, ALDH2, ALDH3A2 and ALDH3B1 (**Fig. 4.6a,b**, **Supplementary Table 1**).

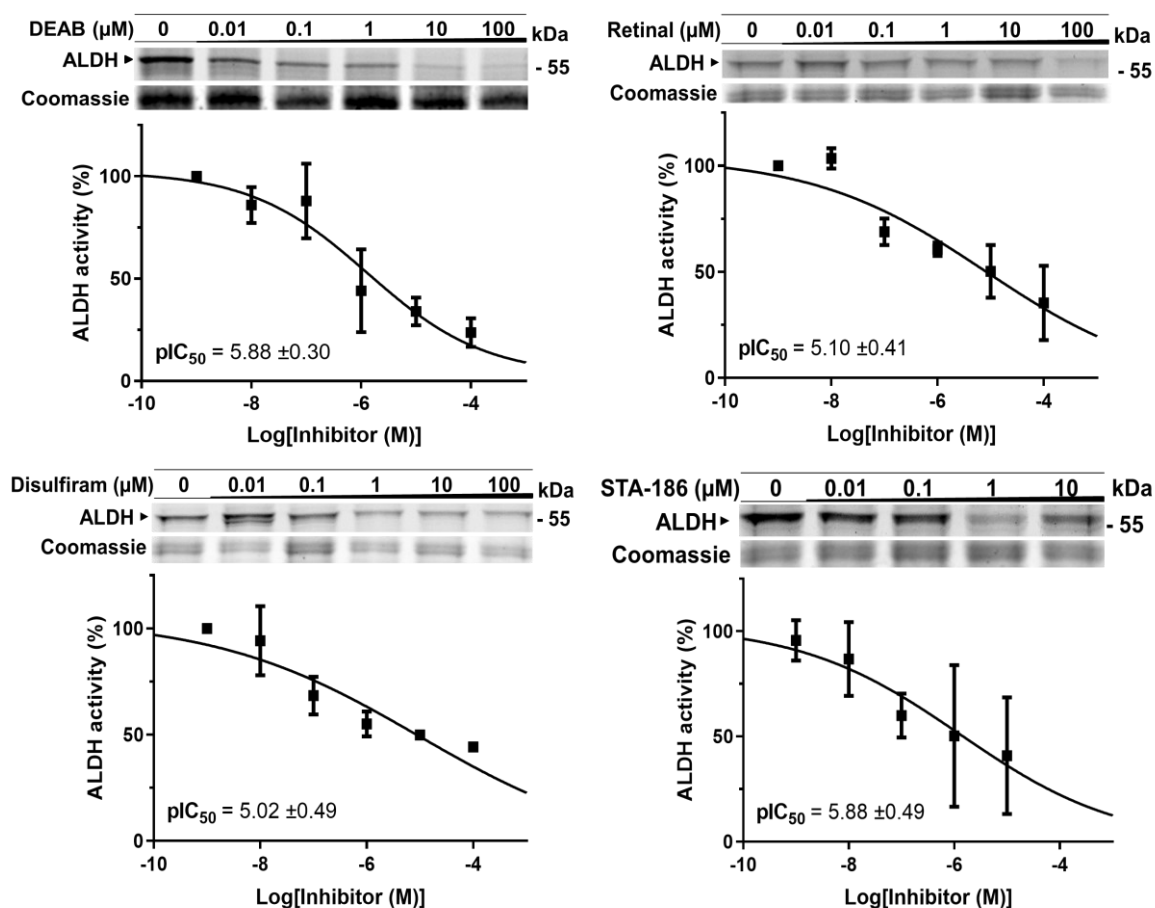


Fig. 4.5 | *In situ* labeling of endogenous ALDHs by LEI-945 and competition with natural substrate and inhibitors. *In situ* labeling of endogenous ALDH enzymes in A549 lung cancer cells with **LEI-945** (1 μM) and competition with ALDH inhibitors or natural substrate. Graphs showing the pIC_{50} values $\pm$ SD of several ALDH inhibitors as determined by competitive ABPP with probe **LEI-945** (1 μM). Data represent mean values $\pm$ SD; $N = 3$ experiments per group (biological replicates).

To exclude the possibility that the warhead covalently binds to reactive cysteines irrespective of recognition based on the retinal scaffold, the list of 34 proteins significantly enriched by **LEI-945** was compared with a list of the 150 most reactive cysteines. This list was compiled by Cravatt *et al.*¹³ using ABPP with the reactive nucleophile iodoacetamide in MDA-MB-231, MCF7 and Jurkat cancer cell lines. Only two proteins (ALDH2 and RTN3) from this list were identified as interacting partners of **LEI-945**. This suggested that the scaffold of **LEI-945** confers selectivity to its binding partners and does not randomly label proteins with hyper reactive cysteines.

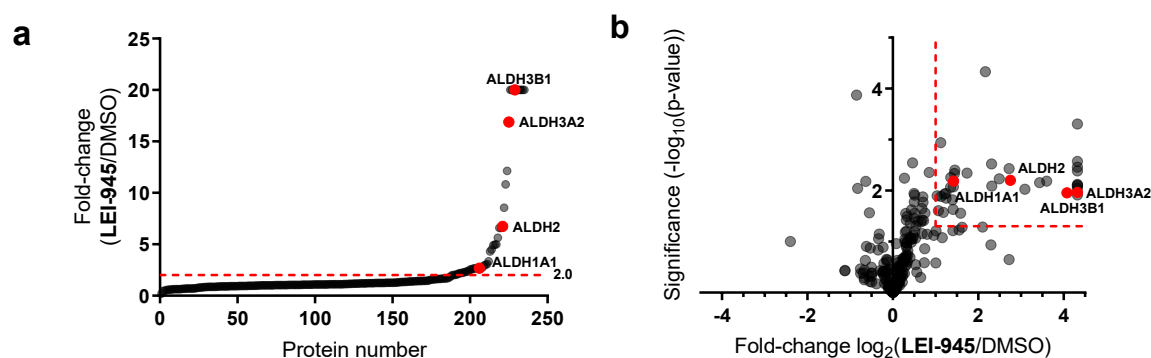


Fig. 4.6 | ABPP of ALDHs in A549 cells. **a**, Fold-change (LEI-945/DMSO) plot for total proteins identified in chemical proteomics experiment with probe LEI-945 (1 μ M). Red lines indicate the threshold fold-change of 2-fold enrichment and the maximum fold-change is set at 20. Red dots represent significantly enriched ALDH enzymes. **b**, Volcano plot for total proteins identified in chemical proteomics experiment with probe LEI-945 (1 μ M). Red lines indicate threshold values (fold-change > 2; p-value < 0.05) marking significantly enriched proteins. Red dots represent significantly enriched ALDH enzymes. For parts **a** and **b**, data are from $N = 4$ experiments (biological replicates).

Some of the proteins found, have been previously reported to interact with retinoids and related lipids, such as sarcoplasmic/endoplasmic reticulum calcium ATPase 2 (ATP2A2), ADP/ATP translocase 2 (SLC25A5), ADP/ATP translocase 3 (SLC25A6), voltage-dependent anion-selective channel protein 2 (VDAC2) and voltage-dependent anion-selective channel protein 3 (VDAC3).^{14–17}

Pathway analysis using the KEGG¹⁸, UniProt¹⁹ and PANTHER²⁰ databases and DAVID²¹ analytic tools, revealed that most enzymes and transporters are located in the endoplasmic reticulum and mitochondria (**Fig. 4.7a**). Many proteins involved in fatty acid degradation, β -alanine and histidine metabolism were enriched, which is in line with the proposed role of ALDH enzymes in these cellular pathways (**Fig. 4.7b**).³

A significant number of unannotated proteins were found, which may be involved in retinal/retinoic acid biochemistry and biology (**Fig. 4.7c**). To investigate whether these proteins are also targeted by retinal and DEAB, a competitive chemical proteomics experiment was performed (**Fig. 4.7d**). Retinal inhibited over 85% of the targets of LEI-945, including all identified ALDH enzymes apart from ALDH2, whereas DEAB was much more selective and reduced the labeling of only three proteins (ALDH1A1, ALDH2 and PCNA-interacting partner (PARPBP)).

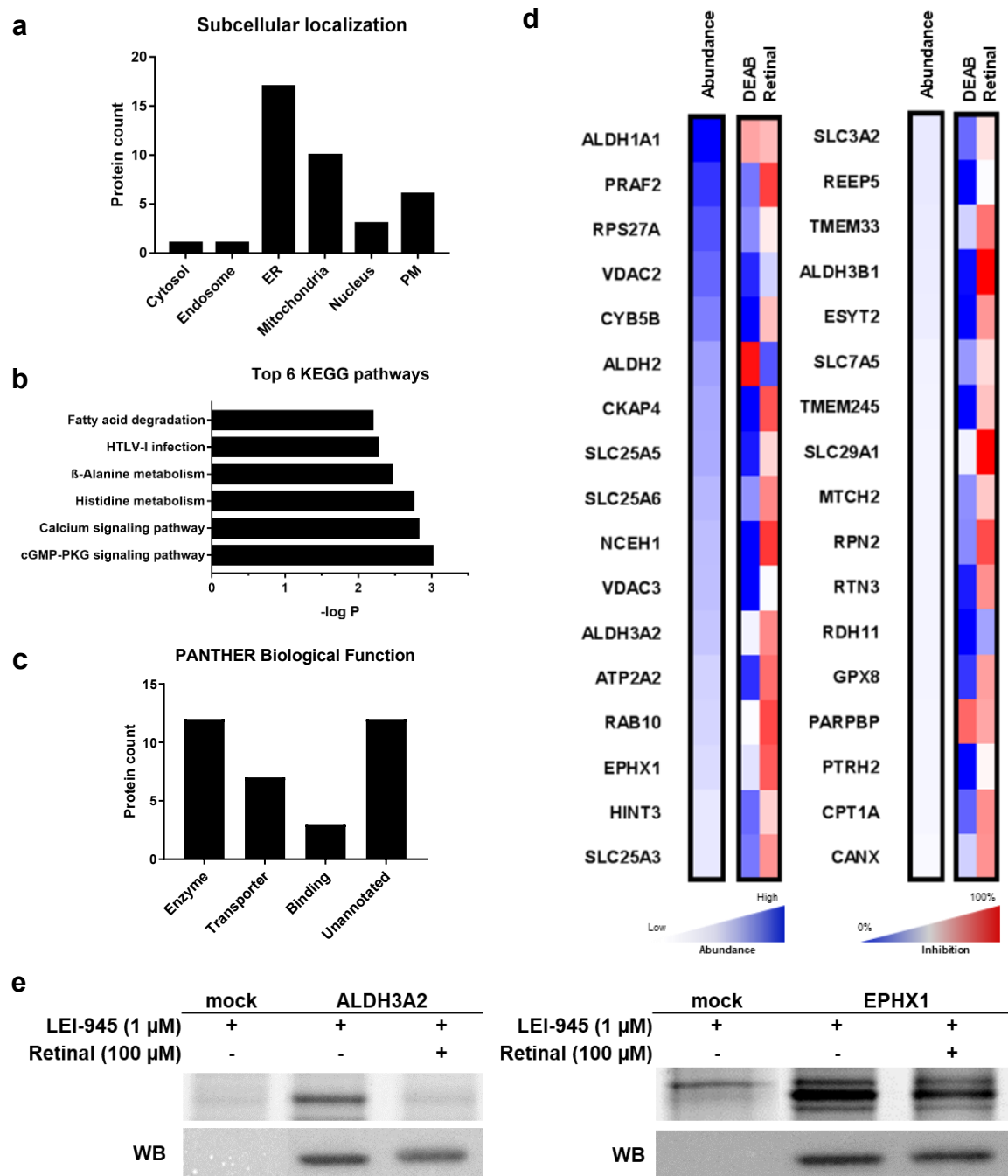


Fig. 4.7 | Proteomics data of retinal interacting proteins in A549 cells. **a**, Subcellular localization of significantly enriched proteins as annotated by the UniProt database. **b**, Top 6 pathways enriched in the group of significantly enriched proteins as determined by screening on the KEGG database. **c**, Biological functions annotated to significantly enriched proteins by the PANTHER database. **d**, Maximum abundance (white = low; blue = high) and inhibition (blue = 0%; red = 100%) by pan-ALDH inhibitor DEAB (100 μ M) and natural substrate retinal (100 μ M) in competitive ABPP experiments with **LEI-945**. Data are from $N = 4$ experiments (biological replicates). **e**, *In situ* labeling of recombinant ALDH3A2 and EPHX1 with **LEI-945** (1 μ M) for 1 h at 37 $^{\circ}$ C. Competition was performed with natural substrate retinal (100 μ M).

To validate the identified probe targets as retinoid-interacting proteins, we overexpressed recombinant ALDH3A2 and epoxide hydrolase 1 (EPHX1) as representative examples in U2OS cells and fluorescently labelled them with **LEI-945** (1 μ M). Competition with retinal (100 μ M) confirmed their ability to bind retinoids (**Fig. 4.7e**). Taken together, these data show that **LEI-945** enables the detection of various retinaldehyde dehydrogenases and other retinoid-interacting proteins in living human cells.

Conclusion

In conclusion, activity-based probe **LEI-945** was validated as the first-in-class ABP for ALDH using fluorescence and chemical proteomic approaches. **LEI-945** labelled overexpressed and endogenous retinaldehyde dehydrogenases in an activity-based fashion. Using chemical proteomics 34 proteins, including ALDH1A1, ALDH2, ALDH3A2 and ALDH3B1, have been significantly enriched from the A549 lung cancer cell line.

Acknowledgements

Hans van den Elst is kindly acknowledged for his contributions to the optimization of the LC/UV-MS assay, Hans den Dulk and Kim Wals for technical assistance with the cell culture and providing the plasmids, Pasquale Putter for determination of the *in situ* pIC₅₀ values of ALDH inhibitors and Eva van Rooden for her advice on chemoproteomics and mass spectrometry analysis.

Experimental procedures

In situ labeling experiments

ALDH plasmids. For the preparation of the different constructs, full length human cDNA was purchased from Source Bioscience and cloned into mammalian expression vector pcDNA3.1, containing genes for ampicillin and neomycin resistance. ALDH1A1, ALDH1A2, ALDH1A3 and ALDH3A2 were cloned into pcDNA3.1. A FLAG-linker was cloned into the vector at the C-terminus of ALDH1A1, ALDH1A2 or ALDH1A3. EPHX1 (OHu14852D) was purchased directly from GenScript. Two step PCR mutagenesis was performed to substitute the active site cysteine for an alanine in the ALDH1A3-FLAG to obtain ALDH1A3-C314A-FLAG. All plasmids were grown in XL-10 Z-competent cells and prepped (Maxi Prep, Qiagen). The sequences were confirmed by sequence analysis at the Leiden Genome Technology Centre.

Cell culture. U2OS cells were grown in DMEM with stable glutamine and phenol red with 10% New Born Calf serum, penicillin and streptomycin at 37 °C and 7% CO₂. A549 cells were grown in DMEM with stable glutamine and phenol red with 10% New Born Calf serum, penicillin and streptomycin at 37°C and 5% CO₂. Medium was refreshed every 2-3 days and cells were passaged twice a week. Cell lines were purchased from ATCC and were regularly tested for mycoplasma contamination. Cultures were discarded after 2-3 months of use.

Transient transfection in U2OS. One day prior to transfection 4*10⁵ U2OS were seeded in a 6-wells plate. Cells were transfected by addition of a 3:1 mixture of polyethyleneimine (6 µg) and plasmid DNA (2 µg) in 200 µL serum free medium per well. The medium was refreshed after 24 hours and after 48 hours the cells were ready to be used.

***In situ* activity-based protein profiling.** Growth medium from cells grown in 6-wells plate was removed and 1 mL serum free medium containing probe **LEI-945** (1 µM, 0.1% DMSO) was added. The cells were then incubated for 1 hour. The medium was then removed, the cells were washed with 2 mL PBS and then harvested in 1 mL PBS using a cell scraper. The cells were moved to an Eppendorf tube and the suspension was centrifuged for 5 minutes at 1200 rpm. The PBS was removed and the samples snap frozen and stored at -80 °C until use.

***In situ* competitive activity-based protein profiling.** Growth medium from cells grown in 6-wells plate was removed and 1 mL serum free medium containing inhibitor or DMSO as vehicle was added (0.1% DMSO). After 1 hour of incubation 1 mL of serum free medium containing probe **LEI-945** (2 µM, 0.1% DMSO) was added. The cells were then incubated for another hour. The medium was then removed, the cells were washed with 2 mL PBS and then harvested in 1 mL PBS using a cell scraper. The cells were moved to an Eppendorf tube and the suspension was centrifuged for 5 minutes at 1200 rpm. The PBS was removed and the samples snap frozen and stored at -80 °C until further use.

CuAAC reaction and in-gel fluorescence analysis. Cell pellets were thawed on ice, lysed by addition of ice-cold lysis buffer (MilliQ, 1x protease inhibitor cocktail (Roche cOmplete EDTA free)) and incubated on ice (15-30 min). The protein concentration was determined by a Quick Start™ Bradford Protein assay (Bio-Rad). The protein fractions were diluted to a total protein concentration of 1 mg/mL. From each sample 40 µL was taken and treated with 5 µL from a freshly prepared “click” mixture containing 9 mM CuSO₄ (2.5 µL/sample, 18 mM in H₂O), 45 mM NaAsc (1.5 µL/sample, 150 mM in H₂O), 1.8 mM THPTA (0.5 µL/sample, 18 mM in DMSO) and 9 µM Alexa Fluor 647 azide (0.5 µL/sample, 90 µM in DMSO from Thermo Fischer Scientific). The samples were incubated for 1 hour at 37 °C and then 15 µL 4x SDS page sample buffer was added. The samples were denatured at 100 °C for 5 minutes. 8 µg per sample was resolved on a SDS-PAGE gel (10% acrylamide, 180 V, 75 min). Gels were visualized with a ChemiDoc XRS (Bio-Rad) using Cy3 and Cy5 multichannel settings (605/50 and 695/55, filters respectively) and stained with Coomassie or transferred to a 0.2 µm polyvinylidene difluoride membranes by Trans-Blot Turbo™ Transfer system (Bio-Rad) after scanning. Fluorescence was normalized to Coomassie staining or to FLAG-tag signal and quantified with Image Lab (Bio-Rad). IC₅₀ curves were fitted with Graphpad Prism® 7 (Graphpad Software Inc.).

CHAPTER 4

Western Blotting. Proteins were transferred to a 0.2 μm polyvinylidene difluoride membranes by Trans-Blot TurboTM Transfer system (Bio-Rad). Membranes were washed with TBS (50 mM Tris, 150 mM NaCl), washed with TBS-T (50 mM Tris, 150 mM NaCl, 0.05% Tween 20) and then blocked with 5% w/v milk powder in TBS-T for 1 hour at room temperature. Membranes were then incubated with primary antibody in 5% w/v milk powder TBS-T (α -FLAG: 1 h, RT), washed three times with TBS-T, incubated with matching secondary antibody in 5% w/v milk powder TBS-T (1 h, RT) and washed with TBS-T and TBS. The blot was developed in the dark using an imaging solution (10 mL Luminol, 100 μL ECL enhancer and 3 μL 30% H_2O_2) and chemiluminescence was visualized using a ChemiDoc XRS (Bio-Rad). The signal was normalized to Coomassie staining and quantified with Image Lab (Bio-Rad). Primary antibodies: monoclonal mouse anti-FLAG (1:5000, Sigma-Aldrich, F3165). Secondary antibodies: HRP-coupled goat-anti-mouse (1:5000, Santa Cruz, sc2005).

In situ retinal conversion LC-UV/MS assay

Sample preparation. U2OS cells were grown in 6-wells plates and transiently transfected ($N = 3$). After 48 hours growth medium was replaced with medium containing retinal (30 μM) with serum. Cells were incubated for 4 hours, then washed with PBS and then PBS (1250 μL) was added. Cells were harvested using a cell scraper and then suspended in PBS. 1 mL of the suspension was moved to low-binding Eppendorf tube and 250 μL to a separate Eppendorf tube used to check for successful overexpression. The low-binding Eppendorf tube was spun down at 1000 g for 5 min and then PBS was removed. Samples were snap frozen and stored at -80°C until needed. A live/dead count was performed on the cells in the remaining Eppendorf tube after which cells were lysed using a probe sonicator (5 sec, 30%). Samples were then denatured and loaded on gel, transferred to a membrane and the overexpression visualized using anti-FLAG antibody. When required low-binding Eppendorf tubes were thawed on ice after which acetonitrile (600 μL , LC-MS grade) was added. The cells were lysed using a probe sonicator (5 sec, 30%) and spun down (14000 g, 5 min). The lysate was transferred to a new low-binding Eppendorf tube. Lysates were collected in low-binding Eppendorf tubes, concentrated using a SpeedVac (Eppendorf) and reconstituted in LC-MS sample solution (110 μL , 90% acetonitrile). Samples were vortexed, spun down (14000 g, 5 min) and transferred (100 μL) to a LC-MS vial after which they were measured.

LC-UV/MS measurement and analysis. Samples were injected onto a C18 column (50 x 4.6 mm, 3 μm ; Nucleodur Gravity, Macherey-Nagel) connected to a Vanquish UHPLC system (Thermo Scientific) with a Vanquish Diode Array detector (Thermo Scientific) coupled to a LCQTM Fleet (Thermo Scientific) via electrospray ionisation (ESI). Acetonitrile and water containing TFA (0.1%) were used for chromatographic separation of the retinoids. The solvent gradient was run from 10% acetonitrile for 1.5 min and then increased to 90% for 7.5 min. UV spectra were recorded between 200 and 600 nm. A calibration curve was made from a mix of retinoid standards (retinol, retinal, all-*trans* retinoic acid, 9-*cis* retinoic acid and 13-*cis* retinoic acid) measured at increasing concentrations (10 nM, 25 nM, 50 nM, 100 nM, 250 nM, 500 nM, 1 μM and 2 μM ; $N = 2$, $n = 4$). Samples were then measured in duplo with a washing step after each sample. Quantification was performed using XcaliburTM software (Thermo Scientific) after which the ratio between the UV spectra selected at 324, 356 and 380 nm were calculated for each peak to determine their purity. Concentrations were calculated using the calibration curve and averaged over the two measurements. The percentages of each retinoid as part of the total amount of retinoids detected in the samples were calculated.

In situ activity-based proteomics

Sample preparation. Protocol adapted from previously described procedure.²² Cells were treated *in situ*, harvested, lysed and adjusted to 1 mg/mL protein concentration as described above. 250 μ L was taken from each sample and to this 25 μ L freshly prepared “click” mixture containing 1 mM CuSO_4 (2.5 μ L/sample, 100 mM in H_2O), 5 mM NaAsc (1.25 μ L/sample, 1 M in H_2O), 0.4 mM THPTA (1 μ L/sample, 100 mM in DMSO), 40 μ M biotin- N_3 (2.5 μ L/sample, 4 mM in DMSO) and MilliQ (17.75 μ L/sample) was added. Samples were incubated for 1 hour at 37 °C while shaking (300 rpm). Excess click reagents were then removed by chloroform/methanol precipitation followed by a wash with methanol. Precipitated proteomes were then suspended in urea buffer (250 μ L, 6 M urea and 25 mM ammonium bicarbonate), DTT (2.5 μ L, 1M) was added and the mixture was then incubated for 15 min at 65 °C while shaking (600 rpm). The samples were then allowed to cool down to RT and then alkylated by addition of iodoacetamide (20 μ L, 0.5 M) for 30 minutes at RT in the dark. Addition of SDS (70 μ L, 10% (v/v)) was followed by heating at 65 °C for 5 minutes. For each sample 50 μ L 50% slurry of Avidin-Agarose from egg white (Sigma-Aldrich) was washed three times with PBS and transferred in PBS (1 mL) to a 15 mL tube. To this another 2 mL of PBS was added followed by the corresponding proteome sample. The beads were incubated with the proteome for 2 hours at room temperature using an overhead shaker. The beads were then isolated by centrifugation (2 min, 2500 g), washed with SDS in PBS (0.5% (w/v)) and washed three times with PBS. The beads were then transferred to low-binding Eppendorf tubes and proteins were digested overnight at 37 °C and 950 rpm shaking in 250 μ L digestion buffer (100 mM Tris, 100 mM NaCl, 1 mM CaCl_2 , 2% acetonitrile and 0.5 μ g sequencing grade trypsin (Promega)). Digestion was stopped by addition of formic acid (12.5 μ L) and the beads filtered off by centrifugation (2 min, 600g) using a Bio-Spin column (Bio-Rad). Samples were then desalted using stage tips, collected in low-binding Eppendorf tubes, concentrated using a SpeedVac (Eppendorf) and stored at -20 °C until reconstitution before measurement.²³ All samples were prepared in at least three biological replicates.

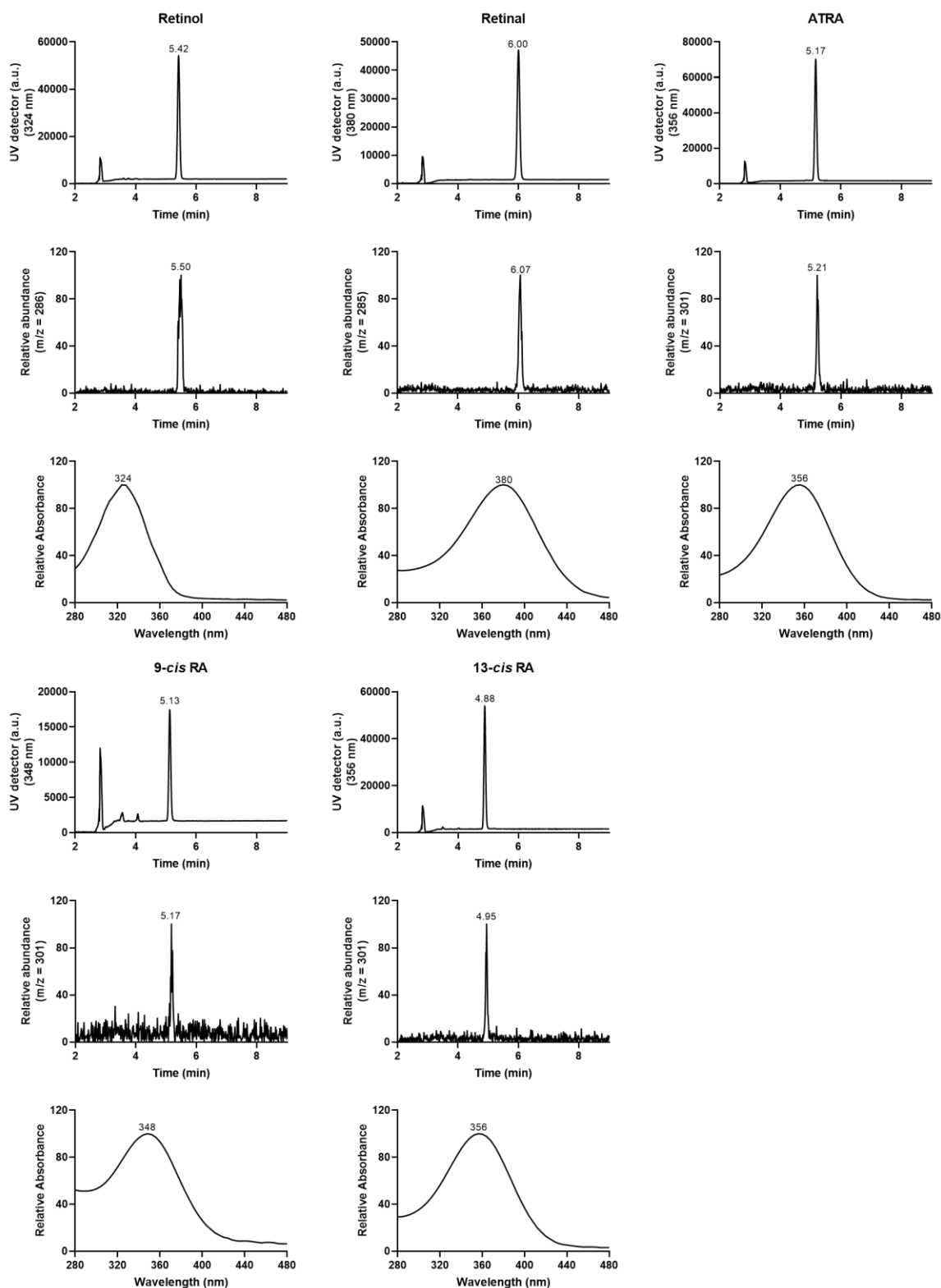
LC-MS/MS measurement and analysis. Samples were reconstituted in LC-MS sample solution (50 μ L, MilliQ, 3% acetonitrile/0.1% formic acid/20 fmol/ μ L enolase). Samples were then analysed using a NanoACQUITY UPLC System (Waters) coupled to a SYNAPT G2-Si high-definition mass spectrometer (Waters) as previously described.^{22,24} Of each sample 5 μ L was loaded on a nanoEASE™ M/Z Symmetry C18 trap column (particles 5 μ m, 100 Å, 180 μ m x 20 mm, Waters) with 0.1% formic acid and separated on an nanoEASE™ M/Z HSS C18 T3 analytical column (particles 1.8 μ m, 75 μ m x 250 mm, Waters) heated at 80 °C. A multistep gradient running from 5-40% acetonitrile containing 0.1% formic acid during a 70 minute method at 300 nL/min was used to achieve peptide separation. Survey scans (m/z 50-2000 Da) were acquired in the Synapt with a scan time of 0.6 seconds in positive, resolution mode. The collision energy is set to 4 V in the trap cell for low-energy MS mode. For the elevated energy scan, the transfer cell collision energy is ramped using drift-time specific collision energies. The lock mass is sampled every 30 seconds. MS raw files were analysed with ProteinLynx Global SERVER (PLGS, v3.0.3, Waters). The MS^E identification was also performed with PLGS using the human proteome from Uniprot (Uniprot-homo-sapiens-trypsin-reviewed-2016-08-29.fasta). The following parameter settings were used: low energy threshold 150 counts, elevated energy threshold 30, peptide and protein FDR 1%, enzyme specificity trypsin, max missed cleavages max 2, variable modification methionine oxidation, fixed modification carbamidomethylation cysteine, at least: fragments/peptide 2, fragments/protein 5, peptides/protein 1 and number of peptides to measure per protein 3. For label-free quantification ISOQuant (v1.5) was used.^{25,26} Data were filtered to retain only proteins with two or more reported unique peptides and quantified in at least 3 replicates of the positive control (probe-treated). Proteins were designated as significantly enriched by the probe when they showed 2-fold enrichment in quantification value when comparing negative control (vehicle-treated) with positive control (probe-treated) samples, probability as determined by a Student's *t* test (<0.05) and Benjamini-Hochberg correction with an FDR of 10%. The mass spectrometry proteomics data (raw data and IsoQuant output tables for proteins groups and peptides) have been deposited in the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository with the dataset identifier PDX015495.^{27,28}

Heatmap Competitive ABPP analysis. Only proteins significantly enriched were selected for analysis. Proteins were sorted on their maximal abundance. The mean raw LFQ intensities from quadruplicate measurements were normalized to DMSO (= 0) and maximal LFQ LEI-945 (= 1) for each protein individually. The heatmap was prepared using Graphpad Prism® 7 (Graphpad Software Inc.).

Supplementary Data

Supplementary Table 4.1 | Proteins significantly enriched using LEI-945 (1 μ M) in A549 cells corresponding to Fig. 4.6b.

Name	Gene name	Accession	Uniprot ID	Fold change	Significance (-log ₁₀)
Retinol dehydrogenase 11	RDH11	Q8TC12	RDH11_HUMAN	20	2,1
Receptor expression-enhancing protein 5	REEP5	Q00765	REEP5_HUMAN	20	3,3
Peptidyl-tRNA hydrolase 2 mitochondrial	PTRH2	Q9Y3E5	PTH2_HUMAN	20	2,1
Histidine triad nucleotide-binding protein 3	HINT3	Q9NQE9	HINT3_HUMAN	20	2,1
Cytochrome b5 type B	CYB5B	O43169	CYB5B_HUMAN	20	2,0
Transmembrane protein 245	TMEM245	Q9H330	TM245_HUMAN	20	2,6
Probable glutathione peroxidase 8	GPX8	Q8TED1	GPX8_HUMAN	20	1,9
Extended synaptotagmin-2	ESYT2	A0FGR8	ESYT2_HUMAN	20	2,5
Reticulon-3	RTN3	O95197	RTN3_HUMAN	20	2,4
Aldehyde dehydrogenase family 3 member B1	ALDH3B1	P43353	AL3B1_HUMAN	20	2,0
Fatty aldehyde dehydrogenase	ALDH3A2	P51648	AL3A2_HUMAN	16,9	2,0
Voltage-dependent anion-selective channel protein 3	VDAC3	Q9Y277	VDAC3_HUMAN	12,1	2,2
Neutral cholesterol ester hydrolase 1	NCEH1	Q6PIU2	NCEH1_HUMAN	10,8	2,2
Voltage-dependent anion-selective channel protein 2	VDAC2	P45880	VDAC2_HUMAN	8,5	2,0
Aldehyde dehydrogenase mitochondrial	ALDH2	P05091	ALDH2_HUMAN	6,7	2,2
Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	ATP2A2	P16615	AT2A2_HUMAN	6,6	2,4
Cytoskeleton-associated protein 4	CKAP4	Q07065	CKAP4_HUMAN	5,6	2,2
PRA1 family protein 2	PRAF2	O60831	PRAF2_HUMAN	5,0	2,1
Carnitine O-palmitoyltransferase 1 liver isoform	CPT1A	P50416	CPT1A_HUMAN	5,0	2,5
Mitochondrial carrier homolog 2	MTCH2	Q9Y6C9	MTCH2_HUMAN	4,5	4,3
Large neutral amino acids transporter small subunit 1	SLC7A5	Q01650	LAT1_HUMAN	3,3	2,3
Microsomal glutathione S-transferase 3	MGST3	O14880	MGST3_HUMAN	3,0	1,6
PCNA-interacting partner	PARBP	Q9NWS1	PARI_HUMAN	2,7	2,4
4F2 cell-surface antigen heavy chain	SLC3A2	P08195	4F2_HUMAN	2,7	2,3
Retinal dehydrogenase 1	ALDH1A1	P00352	AL1A1_HUMAN	2,7	2,2
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 2	RPN2	P04844	RPN2_HUMAN	2,7	2,1
ADP/ATP translocase 2	SLC25A5	P05141	ADT2_HUMAN	2,6	1,8
Ubiquitin-40S ribosomal protein S27a	RPS27A	P62979	RS27A_HUMAN	2,6	2,3
Transmembrane protein 33	TMEM33	P57088	TMM33_HUMAN	2,6	2,1
ADP/ATP translocase 3	SLC25A6	P12236	ADT3_HUMAN	2,4	1,9
Epoxide hydrolase 1	EPHX1	P07099	HYEP_HUMAN	2,4	1,9
Phosphate carrier protein mitochondrial	SLC25A3	Q00325	MPCP_HUMAN	2,3	1,9
Calnexin	CANX	P27824	CALX_HUMAN	2,2	1,9
Equilibrative nucleoside transporter 1	SLC29A1	Q99808	S29A1_HUMAN	2,2	2,9
Ras-related protein Rab-10	RAB10	P61026	RAB10_HUMAN	2,1	1,9



Supplementary Fig. 4.1 | LC-UV/MS spectra for retinoid standards. LC-MS traces and absorption spectra of retinoid standards used in this study. Retention time and maximum wavelength (λ_{max}) of each retinoid were determined using a gradient of 10% acetonitrile for 1.5 minutes followed by 90% acetonitrile for 7.5 minutes.

References

1. Vasiliou, V., Pappa, A. & Petersen, D. R. Role of aldehyde dehydrogenases in endogenous and xenobiotic metabolism. *Chem. Biol. Interact.* **129**, 1–19 (2000).
2. Setshedi, M., Wands, J. R. & De La Monte, S. M. Acetaldehyde adducts in alcoholic liver disease. *Oxidative Medicine and Cellular Longevity* **3**, 178–185 (2010).
3. Koppaka, V. *et al.* Aldehyde Dehydrogenase Inhibitors: a Comprehensive Review of the Pharmacology, Mechanism of Action, Substrate Specificity, and Clinical Application. *Pharmacol. Rev.* **64**, 520–539 (2012).
4. Duester, G., Mic, F. A. & Molotkov, A. Cytosolic retinoid dehydrogenases govern ubiquitous metabolism of retinol to retinaldehyde followed by tissue-specific metabolism to retinoic acid. in *Chemico-Biological Interactions* **143–144**, 201–210 (Elsevier, 2003).
5. Zile, M. H. Vitamin A and Embryonic Development: An Overview. *J. Nutr.* **128**, 455–458 (1998).
6. Wiseman, E. M., Bar-El Dadon, S. & Reifen, R. The vicious cycle of vitamin a deficiency: A review. *Crit. Rev. Food Sci. Nutr.* **57**, 3703–3714 (2017).
7. Mora, J. R., Iwata, M. & Andrian, U. H. Von. Vitamin effects on the immune system. *Nat. Rev. Immunol.* **8**, 685–698 (2008).
8. Blomhoff, H. K. *et al.* Vitamin A is a key regulator for cell growth, cytokine production, and differentiation in normal B cells. *J. Biol. Chem.* **267**, 23988–23992 (1992).
9. Kane, M. A. & Napoli, J. L. Quantification of endogenous retinoids. *Methods Mol. Biol.* **652**, 1–54 (2010).
10. Morgan, C. A., Parajuli, B., Buchman, C. D., Dria, K. & Hurley, T. D. N,N-diethylaminobenzaldehyde (DEAB) as a substrate and mechanism-based inhibitor for human ALDH isoenzymes. *Chem. Biol. Interact.* **234**, 18–28 (2015).
11. Moreb, J. S., Zucali, J. R., Ostmark, B. & Benson, N. A. Heterogeneity of aldehyde dehydrogenase expression in lung cancer cell lines is revealed by Aldefluor flow cytometry-based assay. *Cytom. Part B Clin. Cytom.* **72B**, 281–289 (2007).
12. Kang, J. H. *et al.* Aldehyde dehydrogenase is used by cancer cells for energy metabolism. *Exp. Mol. Med.* **48**, 1–13 (2016).
13. Weerapana, E. *et al.* Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **468**, 790–797 (2010).
14. Notario, B. All-trans-Retinoic Acid Binds to and Inhibits Adenine Nucleotide Translocase and induces Mitochondrial Permeability Transition. *Mol. Pharmacol.* **63**, 224–231 (2003).
15. Dadsena, S. *et al.* Ceramides bind VDAC2 to trigger mitochondrial apoptosis. *Nat. Commun.* **10**, (2019).
16. Budelier, M. M. *et al.* Photoaffinity labeling with cholesterol analogues precisely maps a cholesterol-binding site in voltage-dependent anion channel-1. *J. Biol. Chem.* **292**, 9294–9304 (2017).
17. Rostovtseva, T. K. & Bezrukov, S. M. VDAC regulation: Role of cytosolic proteins and mitochondrial lipids. *J. Bioenerg. Biomembr.* **40**, 163–170 (2008).
18. Ogata, H., Goto, S., Fujibuchi, W. & Kanehisa, M. Computation with the KEGG pathway database. *Biosystems* **47**, 119–128 (1998).
19. Bateman, A. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, 506–515 (2019).
20. Thomas, P. D. *et al.* PANTHER: A browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Research* **31**, 334–341 (2003).
21. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
22. 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).
23. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
24. Distler, U., Kuharev, J., Navarro, P. & Tenzer, S. Label-free quantification in ion mobility-enhanced data-independent acquisition proteomics. *Nat. Protoc.* **11**, 795–812 (2016).
25. Distler, U. *et al.* Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics. *Nat. Methods* **11**, 167–170 (2014).
26. Kuharev, J., Navarro, P., Distler, U., Jahn, O. & Tenzer, S. In-depth evaluation of software tools for data-independent acquisition based label-free quantification. *Proteomics* **15**, 3140–3151 (2015).
27. Vizcaino, J. A. *et al.* ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nature*

- Biotechnology* **32**, 223–226 (2014).
28. Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* **47**, 442–450 (2019).

Chapter 5

Comparative and Competitive ABPP of ALDHs in Breast Cancer Cells using LEI-945

Published as part of S.T.A. Koenders *et al.*, *ACS Cent. Sci.*, **5**, 1965-1974 (2019).

Introduction

Aldehyde dehydrogenases (ALDHs) play an important role in the detoxification of the cellular environment. It is therefore not surprising that an upregulation of ALDH activity in breast cancer cells has been linked with chemo- and radiotherapy resistance.¹ For example, the presence of ALDH1A1 has been found to correlate with resistance in cancer tissue towards cyclophosphamide and the expression of ALDH1A3 is associated with poor clinical outcome in breast cancer.²⁻⁶ Noteworthy, ALDH1A1 and ALDH1A3 have also been reported as biomarkers for cancer stem cells.⁷

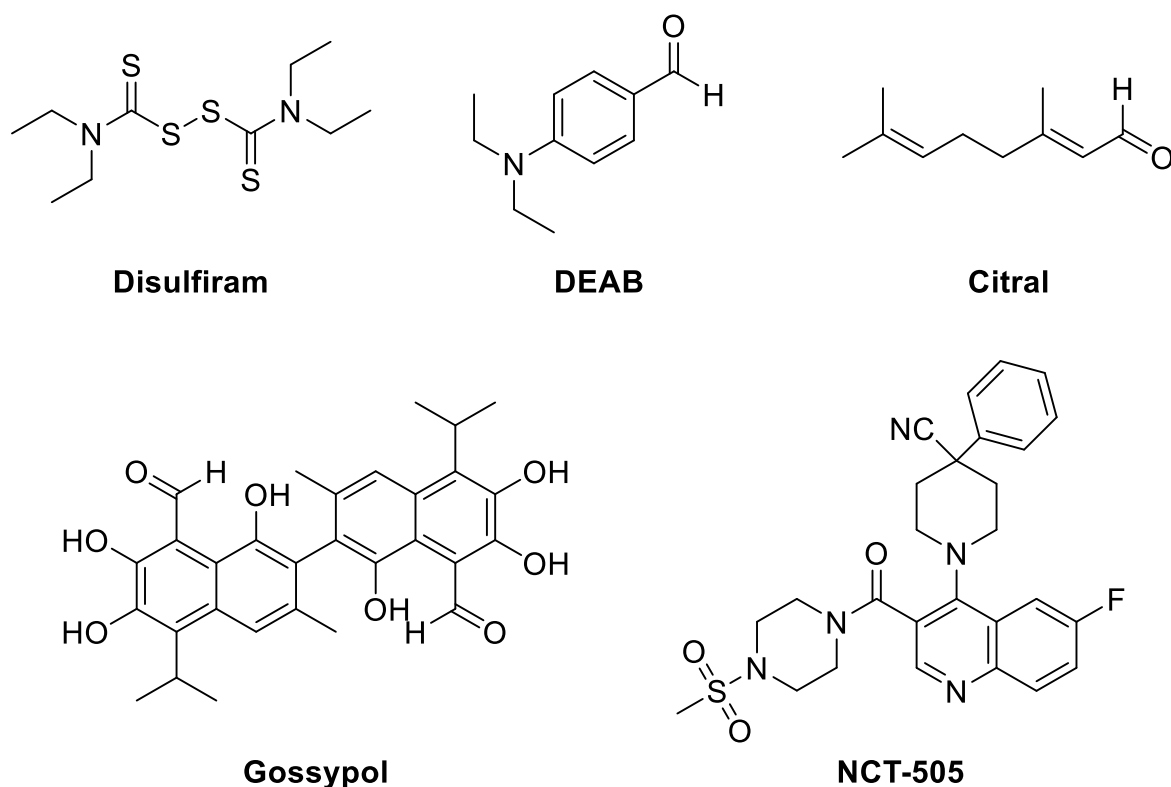


Fig. 5.1 | Chemical structures of ALDH inhibitors disulfiram, DEAB, citral, gossypol and NCT-505.

Inhibition of these ALDHs is therefore a potential therapeutic strategy to overcome drug resistance in cancer. The activity of these enzymes is currently measured using the fluorescence-activated cell sorting (FACS) based ALDEFLUOR assay, which reports on global ALDH activity.⁸ It uses a fluorescent aldehyde that upon oxidation to a charged carboxylate is trapped within the cell. The ALDEFLUOR assay does not, however, discriminate between individual ALDHs. Selective ALDH inhibitors are required to study the physiological role of ALDHs in cancer cells in an acute and dynamic matter. Such inhibitors may serve as potential drug candidates. Reported ALDH inhibitors, such as disulfiram, 4-diethylaminobenzaldehyde (DEAB), citral and gossypol, however, are weakly active and/or demonstrate promiscuous behaviour, which complicates the interpretation of their biological effects.^{9,10} Recently, NCT-505 was developed as the first potent ALDH1A1 inhibitor with a >1000-fold selectivity over ALDH1A3 as determined in a biochemical assay.¹¹ NCT-505 was cytotoxic to ovarian cancer cells and sensitized them to paclitaxel. The cellular selectivity profile and mode-of-action of NCT-505 has not been reported yet, and knowledge hereof would be of importance to guide its therapeutic development. As well, the determination of target protein engagement and off-target activities of small molecules is an essential step in drug discovery.

This chapter describes mapping of ALDH activity in breast cancer cells using comparative ABPP using the first-in-class retinal-based **LEI-945**.¹² Competitive activity-based protein profiling (ABPP) using **LEI-945** was employed to determine the cellular target engagement and selectivity profile of NCT-505 in MDA-MB-468 breast cancer cells. Both ALDH1A1 and ALDH1A3 were inhibited by NCT-505 resulting in an overall reduced cell viability. Following establishment of the selectivity profile of NCT-505, live cell imaging was performed to study the physiological effects of ALDH inhibition in cancer.

Results

LEI-945 reveals distinct ALDH activities in cells of different breast cancer subtypes

To determine whether **LEI-945** can be used to profile the activities of these ALDHs in breast cancer cell lines, a panel of seven breast cancer cells was selected ranging from aggressive basal origin (MDA-MB-231, BT-20, MDA-MB-468 and HCC38) to less aggressive luminal origin (SK-BR-7, MCF7 and SK-BR-3).¹³ Their general ALDH activity was measured using the ALDEFLUOR assay.¹⁴

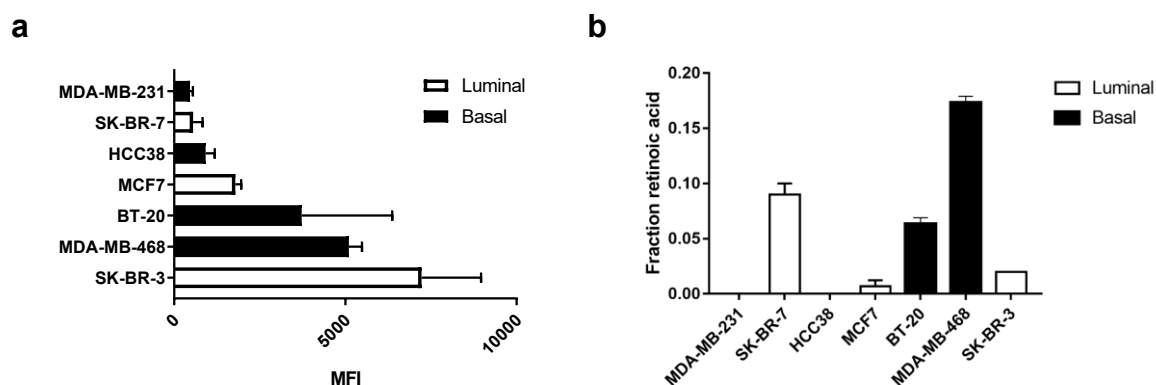


Fig. 5.2 | ALDH activity in breast cancer cells. **a**, Mean fluorescence intensity (MFI) of ALDH activity measured in breast cancer cell lines using the ALDEFLUOR assay. Data are represented as mean values $\pm$ SD; $N = 3$ independent experiments with each $n = 3$ experiments per group (biological replicates). **b**, Retinoic acid production by breast cancer cell lines. Data represent mean values $\pm$ SD; $N = 3$ experiments (biological replicates) measured twice.

Based on this assay the breast cancer cell lines were divided in two groups: ALDH^{low} (MDA-MB-231, SK-BR-7, HCC38 and MCF7) and ALDH^{high} (BT-20, MDA-MB-468 and SK-BR-3) (**Fig. 5.2b**, **Supplementary Fig. 5.1**). These two groups did not correlate with the breast cancer subtype.

Next, the capability of the breast cancer cell lines to convert retinal into retinoic acid was determined by liquid chromatography/mass spectrometry. Retinal (10 μ M, 1 h)-treated cells were lysed and the lipids extracted after which the unique UV absorption of the retinoids relative to other metabolites was used to determine to which extent retinal was converted (**Fig. 5.2b**). The results of the ALDEFLUOR assay (**Fig. 5.2a**) did not correlate with the ability of cells to produce retinoic acid. For example, while MDA-MB-468 and SK-BR-3 were both classified as ALDH^{high}, only MDA-MB-468 cells produced significant amounts of retinoic acid.

To determine whether ABPP using **LEI-945** would be able to explain these apparent contradictory results by determining the levels of active aldehyde dehydrogenases in each breast cancer cell line, chemical proteomics experiments were performed with samples from all seven breast cancer cell lines using **LEI-945** (1 μ M, 1 h) (**Fig. 5.3a**). ALDH1A1 was found in SK-BR-7 and MDA-MB-468. ALDH1A3 was detected in BT-20, MDA-MB-468 and SK-BR-3. ALDH2 was significantly enriched in MDA-MB-231, HCC38, MDA-MB-468 and SK-BR-3. ALDH3A2 was identified in MDA-MB-231, SK-BR-7, BT-20, MDA-MB-468 and SK-BR-3.

Comparison of the chemical proteomics data with data from transcriptomics and global proteomics experiments of these breast cancer cell lines revealed in general a high correlation with r of 0.80 and 0.74 between enzyme activity and mRNA and protein levels, respectively (**Fig. 5.3b,c**). Levels of active ALDH1A1 and ALDH1A3 enzymes were higher than expected on the basis of the mRNA and protein levels, indicating that the activity of these enzymes is potentially regulated post-translationally, as was previously reported.^{15,16}

Active ALDH1A1 was most abundant in the MDA-MB-468 cells, whereas SK-BR-3 had high levels of active ALDH2 and no discernible levels of active ALDH1A1. Since human ALDH2 does not convert retinal into retinoic acid¹⁷, but is an efficient oxidizer of numerous aldehydes, these data may explain the discrepancy between the ALDEFLUOR assay and the retinal conversion assay. Taken together, this data shows that **LEI-945** has the ability to identify individual ALDH isozyme activities in cancer cells and report on the well-known cancer biomarkers ALDH1A1 and ALDH1A3.⁷

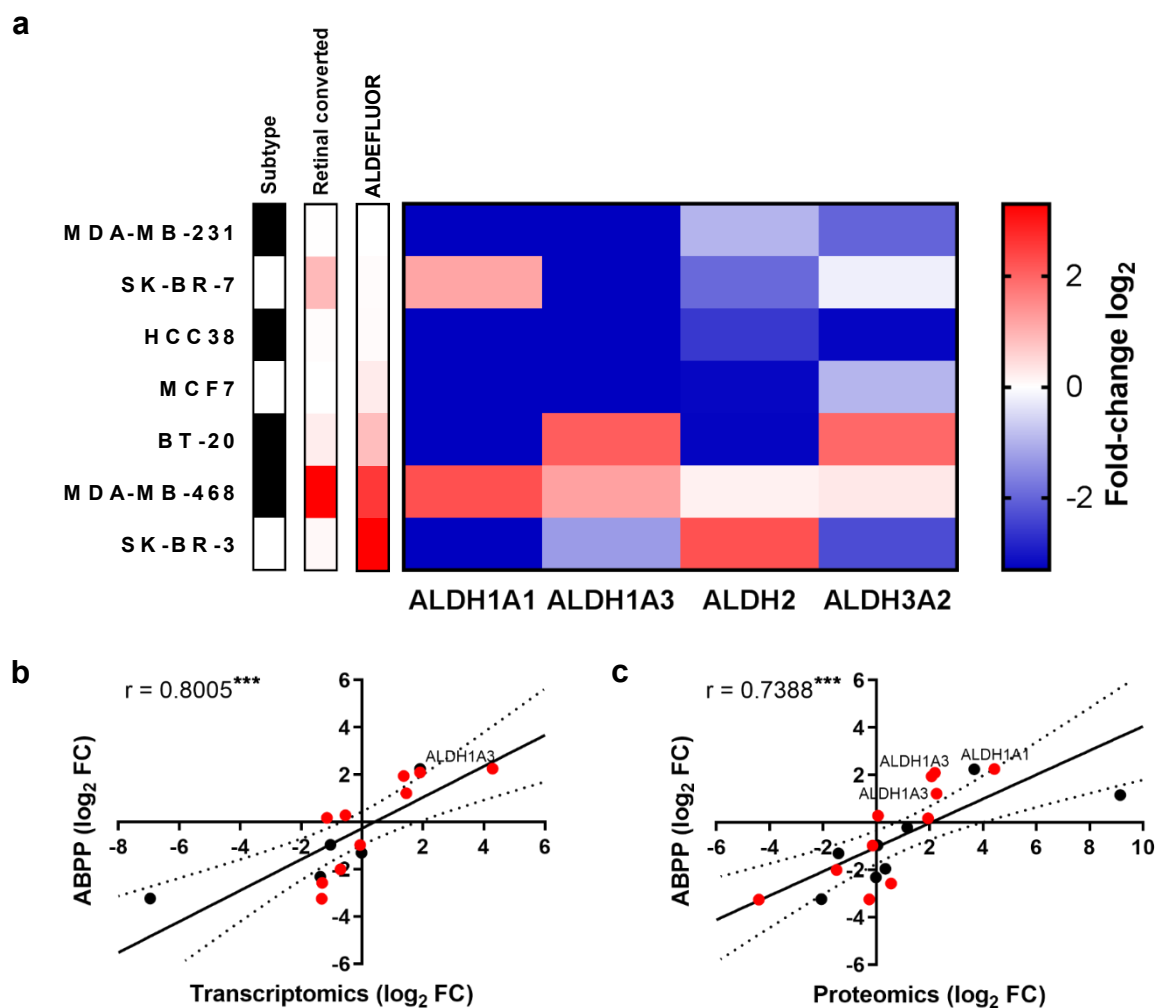


Fig. 5.3 | Profiling levels of active ALDH enzymes in breast cancer cell lines. **a**, ALDH profiling of breast cancer cell lines using chemical proteomics. The subtype column indicates if the cell line belongs to the luminal subtype (white) or basal subtype (black). The retinal converted column shows the amount of retinal converted to retinoic acid over 4 h in a gradient from 0% (white) to 100% (red). The ALDEFLUOR column shows the ALDH activity as determined by the ALDEFLUOR assay in a gradient from low (white) to high (red). The heatmap shows the fold-change in LFQ value for each ALDH enzyme compared to the average for each ALDH enzyme. $N = 2$ independent experiments with each at least $n = 3$ experiments per group (biological replicates). **b**, Correlation of transcriptomics data with ABPP data. ALDHs measured in cell lines with a basal subtype are shown as red dots and the ones with a luminal subtype are shown as black dots and the Pearson's correlation reported for the correlation between ABPP and transcriptomics, ($r = 0.8005$; $p = 0.0003$). **c**, Correlation of proteomics data with ABPP data. ALDHs measured in cell lines with a basal subtype are shown as red dots and the ones with a luminal subtype are shown as black dots and the Pearson's correlation reported for the correlation between ABPP and proteomics, ($r = 0.7388$; $p = 0.0003$).

NCT-505 inhibits ALDH1A1 and ALDH1A3 in MDA-MB-468 cells

For successful drug discovery it is advantageous to understand the molecular and cellular mode-of-action of a drug candidate. So far, studies on the physiological effects of ALDH inhibition in cancer have been mostly performed with poorly active and/or promiscuous inhibitors, which have not been properly characterized in biological systems.^{1,10,18–20} This complicates the interpretation of the physiological function ascribed to a specific ALDH. Here, **LEI-945** was used to study the cellular ALDH interaction profile of the ALDH1A1 inhibitor NCT-505. First, it was confirmed that NCT-505 was able to significantly reduce cell viability and proliferation of MDA-MB-468 cells, which expressed the highest levels of ALDH1A1 (**Fig. 5.4a,b**).

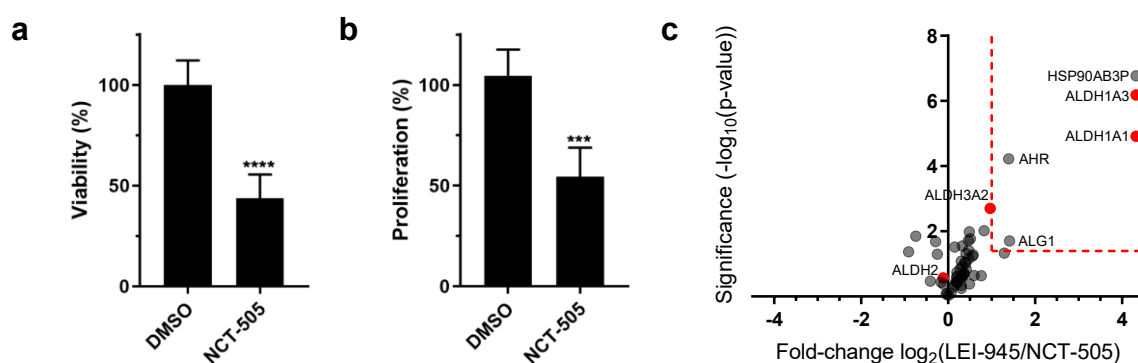


Fig. 5.4 | Target engagement studies of NCT-505 in MDA-MB-468 cells. **a**, Cell viability of MDA-MB-468 cells after treatment with NCT-505 (30 μM) for 72 h. **b**, Cell proliferation of breast cancer cell lines after treatment with NCT-505 (30 μM) for 72 h. For parts **a** and **b**, data represent mean values ± SD; $N = 3$ biological replicates with each $n = 3$ experiments per group. *** $P < 0.001$; **** $P < 0.0001$; t test, two sided. **c**, Volcano plot of *in situ* competitive ABPP experiment in MDA-MB-468 to determine off-targets of ALDH inhibitor NCT-505 (30 μM). $N = 4$ experiments per group (biological replicates).

To determine the cellular target engagement and off-target activities, a chemical proteomics experiment with **LEI-945** was performed using MDA-MB-468 cells *in situ* treated with NCT-505. This competitive ABPP experiment showed that both ALDH1A1 and ALDH1A3 were inhibited by NCT-505 at the concentration used in the cell viability and proliferation assay (**Fig. 5.4c**). Putative heat shock protein HSP 90-β-3 (HSP90AB3P), the aryl hydrocarbon receptor (AHR) and chitobiosyldiphosphodolichol β-mannosyltransferase (ALG1) were also identified as off-targets of NCT-505.

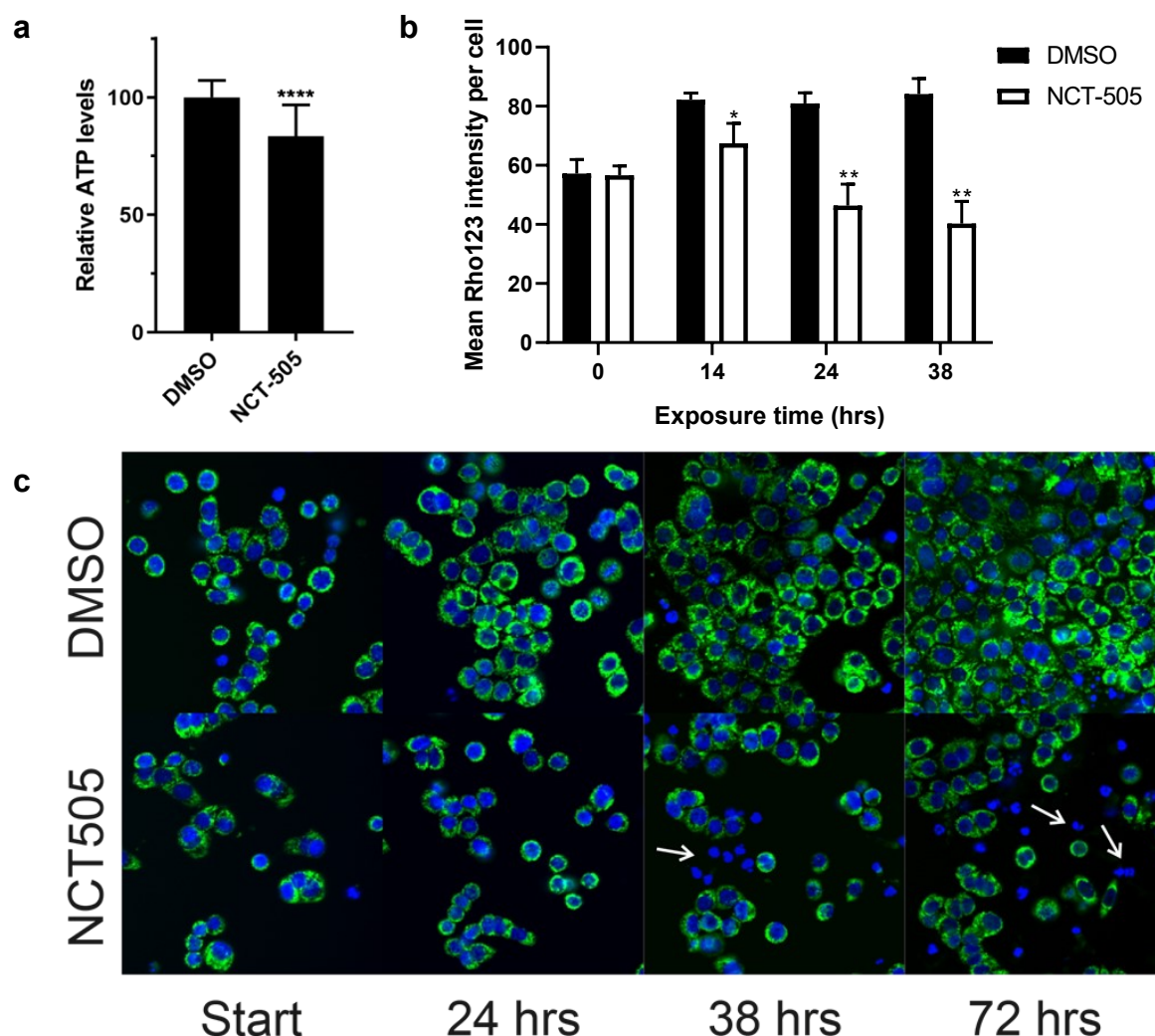


Fig. 5.5 | Impact of NCT-505 on mitochondrial function. **a**, Relative ATP levels of vehicle and NCT-505 (30 μ M)-treated MDA-MB-468 cells as determined by CellTiter-Glo assay. Data represent mean values $\pm$ SD; $N = 3$ biological replicates with each $n = 3$ experiments per group. **** $P < 0.0001$. **b**, Relative mitochondrial membrane potential ($\Delta\Psi_m$) of vehicle and NCT-505 (30 μ M)-treated MDA-MB-468 cells visualized by rhodamine 123 and measured over time. Data represent mean values $\pm$ SD; $N = 3$ biological replicates with each 5 experiments per group. ** $P < 0.01$; t test, two sided. **c**, Representative images of time-lapses performed by confocal microscopy of mitochondrial potential. Mitochondrial membrane potential was visualized using rhodamine 123 dye (green) while the nuclei were detected with live Hoechst (blue). Arrows indicate dying cells without rhodamine signal.

Having established the cellular aldehyde dehydrogenase interaction profile of NCT-505 by ABPP using **LEI-945**, the biological effect of this dual ALDH1A1/ALDH1A3 inhibitor in MDA-MB-468 cells was further characterized by studying its effects on mitochondrial function, cell cycle and cell death. NCT-505 produced a limited, but significant, reduction in adenosine triphosphate (ATP) levels (**Fig. 5.5a**).

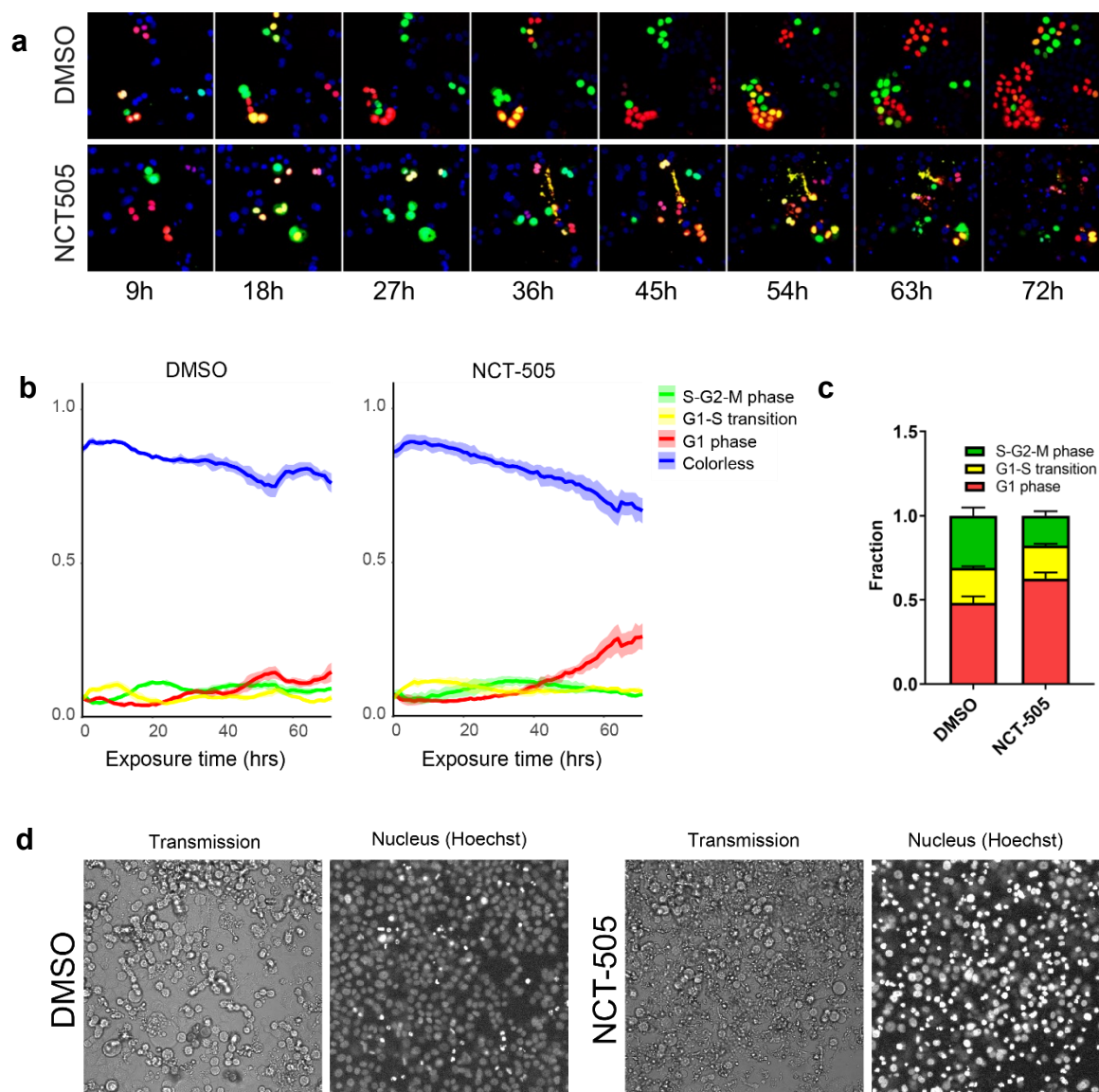


Fig. 5.6 | NCT-505 arrests cells in the G1 phase. **a**, Representative images of FUCCI-expressing MDA-MB-468 cells. Cell cycle phases were visualized by FUCCI expression: G1 phase (red), G1-S transition (yellow) and S-G2-M phase (green). The nuclei were detected with live Hoechst (blue). **b**, FUCCI analysis of the cell cycle using time-lapse microscopy. MDA-MB-468 FUCCI cells were treated with NCT-505 (30 μ M) for 72 h and imaged every hour. **c**, Distribution over the different cell cycle phases of vehicle and NCT-505 (30 μ M)-treated FUCCI-expressing MDA-MB-468 cells after 72 h. Data represent mean values $\pm$ SD; $N = 3$ biological replicates with each 5 experiments per group. **d**, Representative images of nuclei condensation. MDA-MB-468 cells were treated with NCT-505 (30 μ M) for 72 h. Hoechst staining showed highly condensed nuclei in the NCT-505-treated cells compared to vehicle.

Mitochondrial membrane potential ($\Delta\Psi_m$) upon preincubation of MDA-MB-468 cells was also reduced as assessed with rhodamine 123, a fluorescent marker for mitochondrial activity, suggesting that mitochondrial function was altered (**Fig. 5.5b,c**).

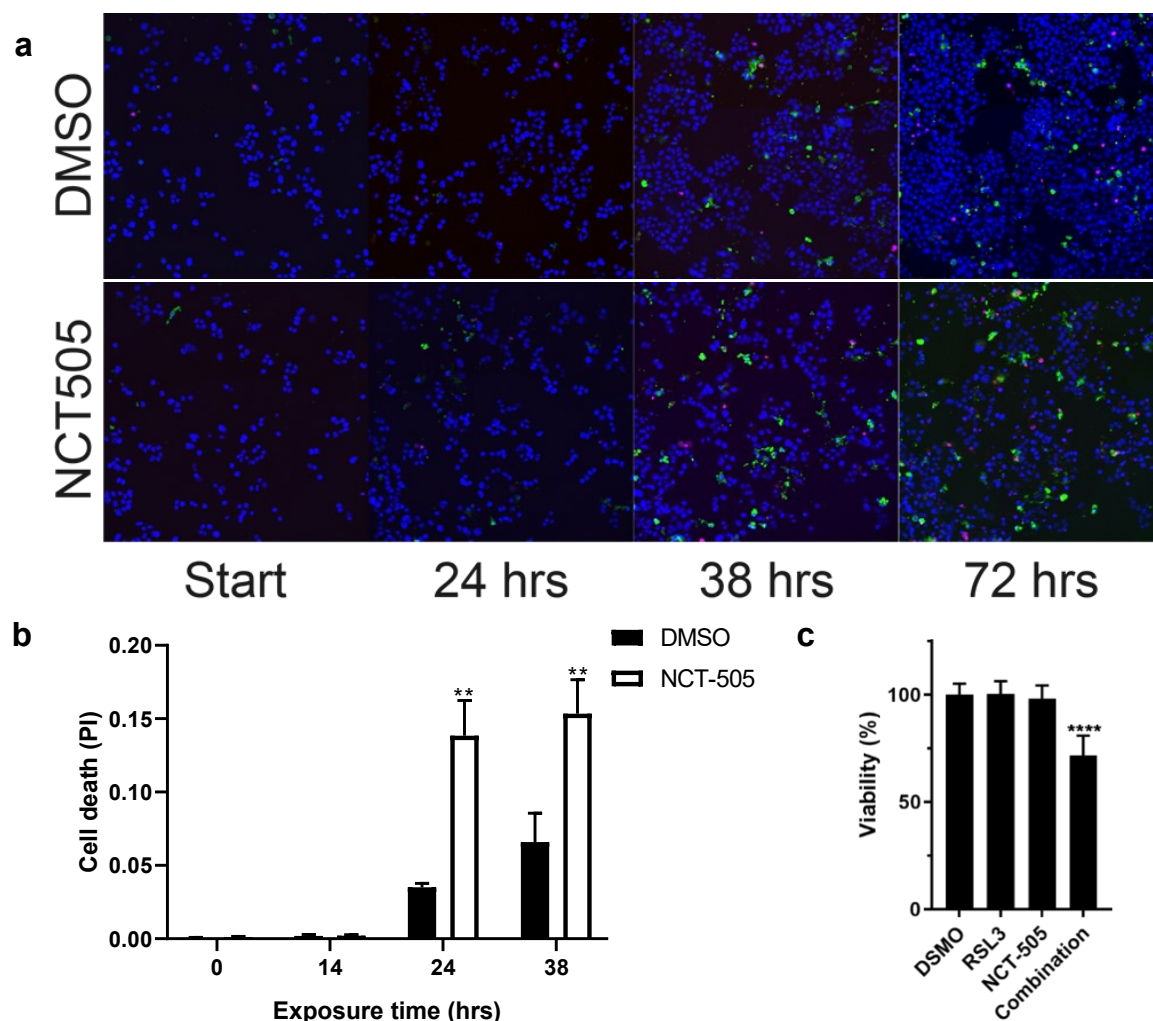


Fig. 5.7 | ALDH1A1/1A3 inhibition cells undergo non-apoptotic cell death. **a**, Representative images of time-lapses performed by confocal microscopy of cell death. MDA-MB-468 cells were treated with NCT-505 (30 μ M) for 72 h. Cell death was visualized using propidium iodide (PI; red), Annexin V (AnV; green) and the nuclei were detected with live Hoechst (blue). **b**, Cell death profile of vehicle and NCT-505 (30 μ M)-treated MDA-MB-468 cells as determined by propidium iodide (PI) staining and measured over time. Data represent mean values $\pm$ SD; $N = 3$ biological replicates with each 5 experiments per group. ** $P < 0.01$; t test, two-sided. **c**, Synergy between RSL3 (30 nM) and NCT-505 (3 μ M) measured using a viability assay after 72 h of treatment. Data represent mean values $\pm$ SD; $N = 3$ biological replicates with each 6 experiments per group. **** $P < 0.0001$.

In addition, exposure to NCT-505 arrested cells in the G1 cycle as determined by fluorescent ubiquitination-based cell cycle indicators (FUCCI) (Fig. 5.6a,b,c).²¹ The nuclei of the NCT-505-treated cells were highly condensed compared to vehicle-treated cells indicating dying cells, which had no rhodamine 123 signal as well (Fig. 5.6d, see arrows in Fig. 5.5c). A significant increase in the fraction of propidium iodide (PI) positive cells was observed in NCT-505-treated cells, which suggested that the cells underwent necrosis (Fig. 5.7a,b).

As changes in mitochondrial morphology are a distinctive feature of ferroptosis²², the observed non-apoptotic cell death in combination with alterations in mitochondrial function might indicate that the dying cells undergo ferroptosis. To determine whether incubation with RSL3, a compound that induces ferroptosis via inhibition of GPX4^{23,24} and inhibition of ALDH1A1/ALDH1A3 could synergistically induce cell death, MDA-MB-468 cells were treated with ineffective, low concentrations of GPX4 inhibitor RSL3 (30 nM) or NCT-505 (3 μ M). At these concentrations, the inhibitors did not show any reductions in cell viability, whereas the combination led to a significant decrease in cell viability with a combination index value of 0.87 indicating synergy (**Fig. 5.7c**).

Discussion

LEI-945 was able to detect and compare individual active ALDH isozymes in breast cancer cell lines. The ALDH profiles generated using **LEI-945** could be used to explain the ability of certain cell lines to produce retinoic acid, where the ALDEFLUOR assay could not. It also detected levels of active ALDH1A1 enzyme in the SK-BR-7 cell line, while the ALDEFLUOR assay reported no ALDH activity. This method can also be used in combination with the ALDEFLUOR assay by sorting cells based on global ALDH activity and subsequently subjecting the sorted cell populations to comparative ABPP analysis. Thus providing both qualitative and quantitative information on the levels of active ALDH enzymes in therapy-resistant cancer (stem) cells enabling a better mechanistic understanding of the underlying biology, which is ALDH isozyme dependent.²⁵

Furthermore, **LEI-945** enabled target engagement and cellular selectivity studies of the recently reported ALDH1A1 inhibitor NCT-505. This ABPP method revealed that NCT-505 does not only inhibit ALDH1A1, but also ALDH1A3, at concentrations that significantly reduced cell viability and proliferation of MDA-MB-468 cells via inhibition of mitochondrial function, cell cycle arrest in the G1 phase and necrosis-induced cell death. The concentration used in this study is in accordance with the reported EC₅₀ value derived from OV-90 ovarian cancer cells grown in 2D cell culture.¹¹ Maloney *et al.* showed an increase in sensitivity towards NCT-505 when cells are grown in 3D cell culture, possibly due to the accompanying elevated levels of ALDH1A1 expression.¹¹ Elevated ALDH expression in 3D cell culture compared with 2D has also been described by others.^{26,27} This phenomenon could be studied in more depth by looking at changes in the levels of active enzymes using **LEI-945**.

Ferroptosis induced via GPX4 inhibition was synergistic with NCT-505-mediated toxicity. This combined therapy is of interest as both therapeutic targets, GPX4 and ALDH1A1, have been linked to therapy resistance in cancer.^{20,23,24} Since both enzymes are involved in the detoxification of lipid peroxidation products, this might explain the observed synergistic effect.^{28–30} Since NCT-505 also inhibits ALDH1A3, this data does not rule out the possibility that ALDH1A3 may be an interesting drug target to tackle drug resistance in cancer.

Conclusion

To conclude, **LEI-945** can be used for comparative and competitive ABPP of cancer cells and ALDH inhibitors, thereby providing guidance in the target validation and discovery of cancer (stem) cell-based therapies. Thus, these results showcase the application of substrate-based probes in interrogating pathologically relevant enzyme activities. They also highlight the general power of chemical proteomics in driving the discovery of new biological insights and guiding drug discovery efforts.

Acknowledgements

Martje Erkelens is kindly acknowledged for performing the ALDEFLUOR assay, Alexander Bakker and Bogdan Florea for mass spectrometry analysis, Vera van der Noord for her advice and technical assistance with the cellular assays, Else Botter for her collaboration on performing the cellular assays, Lukas Wijaya and Sylvia Le Dévédec for performing and analysing the live cell imaging.

Experimental procedures

Biological methods

Cell culture. The breast cancer cell lines BT-20, HCC38, MCF7, MDA-MB-231, MDA-MB-468, SK-BR-3 and SK-BR-7 were grown in RPMI-1640 with stable glutamine and phenol red with 10% Fetal Calf serum, penicillin and streptomycin at 37 °C and 5% CO₂. Medium was refreshed every 2-3 days and cells were passaged twice a week. Cell lines were purchased from ATCC and were regularly tested for mycoplasma contamination. Cultures were discarded after 2-3 months of use.

Flow cytometry. Cells from the different cell lines were used in the logarithmic growth phase for analysis with flow cytometry. The ALDEFUOR assay (Stemcell, Cologne, Germany) was performed according to manufacturer's protocol. Subsequently, cells were stained with CD24 APC (Clone ML5, 1:100), CD44 PE Cy7 (Clone BJ18, 1:800), CD133 PE/Dazzle 594 (Clone 7, 1:100, all Biolegend, London, United Kingdom). Just before acquiring, DAPI (Invitrogen, Bleiswijk, the Netherlands) was added as a live death stain (1:40.000). Samples were acquired on LSR-Fortessa X20 (BD Bioscience, Mountain View, CA, USA) and data was analysed with Flowjo software (Tree Star, San Carlos, CA, USA).

Cell viability and proliferation. Breast cancer cells were grown to 80% confluence in 10 cm plates. Cells were then seeded in a 96-wells plate depending on their rate of growth (BT-20: 12 k/well; MDA-MB-468 & SK-BR-3: 10 k/well; SK-BR-7 & HCC38: 8 k/well; MDA-MB-231: 4 k/well). For the SRB proliferation assay cells were also seeded on an extra plate to be able to determine the starting point protein content. Cells were allowed to adhere overnight and were then treated with vehicle (0.1% DMSO) or inhibitor NCT-505 (30 µM, 0.1% DMSO) in 100 µL RPMI medium. Cells were then incubated for 72 hours.

For the MTT viability assay another 100 µL of medium containing MTT (1 mg/mL) was added and the cells were incubated for another 4 hours. The medium was then removed and the formazan crystals were dissolved in 100 µL DMSO during 1 hour in the stove. The absorbance was then measured in a CLARIOstar Plus plate reader at 570 nm.

For the SRB assay cells were fixated by incubation with 30 µL 50% TCA for 1 hour at 4 °C. The plates were washed 5 times with water and dried. 60 µL 0.4% SRB was added and incubated for 30 minutes at room temperature. The plates were then washed another four times with acetic acid (0.1%) and dried. The SRB was dissolved by adding 150 µL Tris (10 mM) and incubating for 30 minutes at room temperature. Absorbance was then measured at 540 nm on a CLARIOstar Plus plate reader. Graphpad Prism® 7 (Graphpad Software Inc.) was used to plot data and calculation of mean values and standard deviation.

Determination of ATP levels. MDA-MB-468 cells were seeded in 96-wells plates at 10 k/well. Cells were allowed to adhere overnight and were then treated with vehicle (0.1% DMSO) or inhibitor NCT-505 (30 µM, 0.1% DMSO) for 4 hour. The plate was then allowed to equilibrate for 30 minutes at room temperature. CellTiter-Glo assay (Promega) was then performed according to manufacturer's protocol. Measured on a CLARIOstar Plus plate reader and Graphpad Prism® 7 (Graphpad Software Inc.) was used to plot data and calculation of mean values and standard deviation.

Synergy of NCT-505 and RSL3. MDA-MB-468 cells were seeded in 96-wells plates at 10k/well. Cells were allowed to adhere overnight and were then treated with vehicle (0.1% DMSO), NCT-505 (3 µM, 0.1% DMSO), RSL3 (30 nM, 0.1% DMSO, Selleck Chemicals) or a combination of NCT-505 (3 µM) and RSL3 (30 nM). Cells were then incubated for 72 hours and a MTT assay was performed as described above. Graphpad Prism® 7 (Graphpad Software Inc.) was used to plot data and calculation of mean values and standard deviation. The IC28 values for both NCT-505 and RSL3 were calculated and used to determine the combination index value.

In situ activity-based proteomics

Sample preparation. Protocol adapted from previously described procedure.³¹ Cells were treated *in situ*, harvested, lysed and adjusted to 1 mg/mL protein concentration as described above. 250 μ L was taken from each sample and to this 25 μ L freshly prepared “click” mixture containing 1 mM CuSO₄ (2.5 μ L/sample, 100 mM in H₂O), 5 mM NaAsc (1.25 μ L/sample, 1 M in H₂O), 0.4 mM THPTA (1 μ L/sample, 100 mM in DMSO), 40 μ M biotin-N₃ (2.5 μ L/sample, 4 mM in DMSO) and MilliQ (17.75 μ L/sample) was added. Samples were incubated for 1 hour at 37 °C while shaking (300 rpm). Excess click reagents were then removed by chloroform/methanol precipitation followed by another wash with methanol. Precipitated proteomes were then suspended in urea buffer (250 μ L, 6 M urea and 25 mM ammonium bicarbonate), DTT (2.5 μ L, 1 M) was added and the mixture was then incubated for 15 min at 65 °C while shaking (600 rpm). The samples were then allowed to cool down to RT and then alkylated by addition of iodoacetamide (20 μ L, 0.5M) for 30 minutes at RT in the dark. Addition of SDS (70 μ L, 10% (v/v)) was followed by heating at 65 °C for 5 minutes. For each sample 50 μ L 50% slurry of Avidin-Agarose from egg white (Sigma-Aldrich) was washed three times with PBS and transferred in PBS (1 mL) to a 15 mL tube. To this another 2 mL of PBS was added followed by the corresponding proteome sample. The beads were incubated with the proteome for 2 hours at room temperature using an overhead shaker. The beads were then isolated by centrifugation (2 min, 2500 g), washed with SDS in PBS (0.5% (w/v)) and washed three times with PBS by centrifugation. The beads were then transferred to low-binding Eppendorf tubes and proteins were digested overnight at 37 °C and 950 rpm shaking in 250 μ L digestion buffer (100 mM Tris, 100 mM NaCl, 1 mM CaCl₂, 2% acetonitrile and 0.5 μ g sequencing grade trypsin (Promega)). Digestion was stopped by addition of formic acid (12.5 μ L) and the beads filtered off by centrifugation (2 min, 600 g) using a Bio-Spin column (Bio-Rad). Samples were then desalted using stage tips, collected in low-binding Eppendorf tubes, concentrated using a SpeedVac (Eppendorf) and stored at -20 °C until reconstitution before measurement.³² All samples were prepared in at least three biological replicates.

LC-MS/MS measurement and analysis. Samples were reconstituted in LC-MS sample solution (50 μ L, MilliQ, 3% acetonitrile/0.1% formic acid/20 fmol/ μ L enolase). Samples were then analysed using a NanoACQUITY UPLC System (Waters) coupled to a SYNAPT G2-Si high-definition mass spectrometer (Waters) as previously described.^{31,33} Of each sample 5 μ L was loaded on a nanoEASE™ M/Z Symmetry C18 trap column (particles 5 μ m, 100 Å, 180 μ m x 20 mm, Waters) with 0.1% formic acid and separated on an nanoEASE™ M/Z HSS C18 T3 analytical column (particles 1.8 μ m, 75 μ m x 250 mm, Waters) heated at 80 °C. A multistep gradient running from 5-40% acetonitrile containing 0.1% formic acid during a 70 minute method at 300 nL/min was used to achieve peptide separation. Survey scans (m/z 50-2000 Da) were acquired in the Synapt with a scan time of 0.6 seconds in positive, resolution mode. The collision energy is set to 4 V in the trap cell for low-energy MS mode. For the elevated energy scan, the transfer cell collision energy is ramped using drift-time specific collision energies. The lock mass is sampled every 30 seconds. MS raw files were analysed with ProteinLynx Global SERVER (PLGS, v3.0.3, Waters). The MS^E identification was also performed with PLGS using the human proteome from Uniprot (uniprot-homo-sapiens-trypsin-reviewed-2016-08-29.fasta). The following parameter settings were used: low energy threshold 150 counts, elevated energy threshold 30, peptide and protein FDR 1%, enzyme specificity trypsin, max missed cleavages max 2, variable modification methionine oxidation, fixed modification carbamidomethylation cysteine, at least fragments/peptide 2, fragments/protein 5, peptides/protein 1 and number of peptides to measure per protein 3. For label-free quantification ISOQuant (v1.5) was used.^{34,35} Data were filtered to retain only proteins with two or more reported unique peptides and quantified in at least 3 replicates of the positive control (probe-treated). Proteins were designated as significantly enriched by the probe when they showed 2-fold enrichment in quantification value when comparing negative control (vehicle-treated) with positive control (probe-treated) samples, probability as determined by a Student's *t* test (<0.05) and Benjamini-Hochberg correction with an FDR of 10%. The mass spectrometry proteomics data (raw data and IsoQuant output tables for proteins groups and peptides) have been deposited in the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository with the dataset identifier PDX015495.^{36,37}

Heatmap ALDH profiles breastcancer analysis and correlation with transcriptomics and proteomics. Only ALDH enzymes significantly enriched in at least one breastcancer cell line (ALDH1A1, ALDH1A3, ALDH2, ALDH3A2) were selected for analysis. The mean raw LFQ intensities of triplicate measurements from two independent experiments were normalized using the LFQ mean raw LFQ intensity of the enolase internal standard and averaged. Average LFQ values for each ALDH enzymes were then calculated and used to determine the fold-change log₂. The heatmap was prepared using Graphpad Prism® 7 (Graphpad Software Inc.). The log₂ FC was plotted against transcriptomics and proteomics log₂ FC data.

CHAPTER 5

Breast cancer cell line mRNA microarray data used were previously established and are available in the Gene Expression Omnibus data repository (GEO) (accession number: GSE41313).³⁸ Protein abundance in the breast cancer cell lines was previously analysed using proteomics (manuscript in preparation, data available upon request). Transcriptomics data did not include the SK-BR-7 cell line. Linear regression and Pearson's correlation were calculated and plotted using Graphpad Prism[®] 7 (Graphpad Software Inc.).

In situ retinal conversion LC-UV/MS assay

Sample preparation. U2OS cells were grown in 6-wells plates and transiently transfected ($N = 3$). After 48 hours growth medium was replaced with medium containing retinal (30 μ M) with serum. Cells were incubated for 4 hours, then washed with PBS and then PBS (1250 μ L) was added. Cells were harvested using a cell scraper and then suspended in PBS. 1 mL of the suspension was moved to low-binding Eppendorf tube and 250 μ L to a separate Eppendorf tube used to check for successful overexpression. The low-binding Eppendorf tube was spun down at 1000g for 5 min and then PBS was removed. Samples were snap frozen and stored at -80 °C until needed. A live/dead count was performed on the cells in the remaining Eppendorf tube after which cells were lysed using a probe sonicator (5 sec, 30%). Samples were then denatured and loaded on gel, transferred to a membrane and the overexpression visualized using anti-FLAG antibody. When required low-binding Eppendorf tubes were thawed on ice after which acetonitrile (600 μ L, LC-MS grade) was added. The cells were lysed using a probe sonicator (5 sec, 30%) and spun down (14000 g, 5 min). The lysate was transferred to a new low-binding Eppendorf tube. Lysates were collected in low-binding Eppendorf tubes, concentrated using a SpeedVac (Eppendorf) and reconstituted in LC-MS sample solution (110 μ L, 90% acetonitrile). Samples were vortexed, spun down (14000 g, 5 min) and transferred (100 μ L) to a LC-MS vial after which they were measured.

LC-UV/MS measurement and analysis. Samples were injected onto a C18 column (50 x 4.6 mm, 3 μ m; Nucleodur Gravity, Macherey-Nagel) connected to a Vanquish UHPLC system (Thermo Scientific) with a Vanquish Diode Array detector (Thermo Scientific) coupled to a LCQ[™] Fleet (Thermo Scientific) via electrospray ionisation (ESI). Acetonitrile and water containing TFA (0.1%) were used for chromatographic separation of the retinoids. The solvent gradient was run from 10% acetonitrile for 1.5 min and then increased to 90% for 7.5 min. UV spectra were recorded between 200 and 600 nm. A calibration curve was made from a mix of retinoid standards (retinol, retinal, all-*trans* retinoic acid, 9-*cis* retinoic acid and 13-*cis* retinoic acid) measured at increasing concentrations (10 nM, 25 nM, 50 nM, 100 nM, 250 nM, 500 nM, 1 μ M and 2 μ M; $N = 2$, $n = 4$). Samples were then measured in duplo with a washing step after each sample. Quantification was performed using Xcalibur[™] software (Thermo Scientific) after which the ratio between the UV spectra selected at 324, 356 and 380 nm were calculated for each peak to determine their purity. Concentrations were calculated using the calibration curve and averaged over the two measurements. The percentages of each retinoid as part of the total amount of retinoids detected in the samples were calculated.

Live cell imaging

Cell culture and lentivirus production. HEK293T cells were maintained in high-glucose Dulbecco's Modified Eagle Medium (DMEM, HyClone) with 5% fetal bovine serum (FBS, Gemini Bio) and 2 mM glutamine (Life Technologies) at 37 °C with 5% CO₂. To package lentivirus, HEK293T cells at ~70% confluence were transfected with pMDLg-RRE (gag/pol), pCMV-VSVG, pRSV-REV, and FUCCI plasmids (addgene#83841) using polyethylenimine (PEI). Two days post-transfection, viral supernatant was filtered with a 0.45 μ m polyethersulfone filter before using to infect the MDA-MB-468 cells.

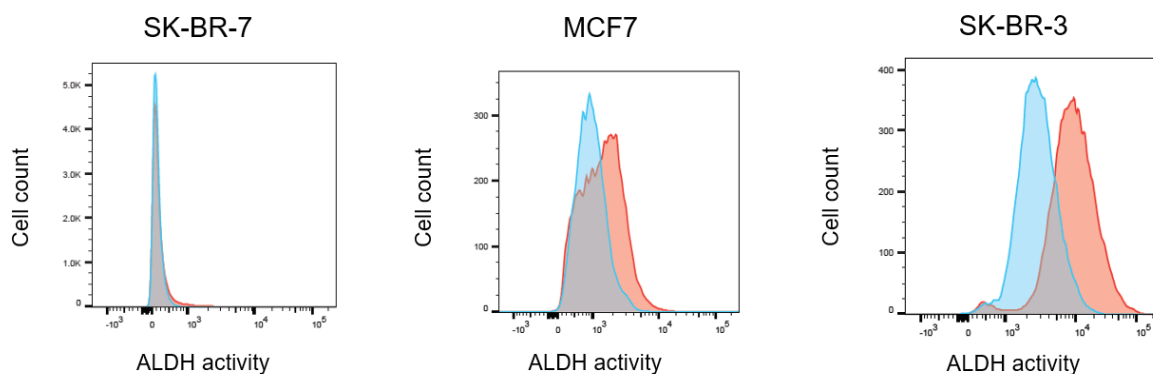
Exposure and live cell imaging. Cells were seeded in Greiner black μ -CLEAR 96 well plates at 10k/well. Prior to the NCT-505 exposure, the cells were incubated with 100 ng/ml live Hoechst 33342 in complete RPMI for 2 hours. Thereafter, the medium was refreshed with complete medium. For cell death experiment, 0.2 μ M propidium iodide (PI) and Annexin-V-Alexa633 (AnV) were added to the complete medium. For the live mitochondrial membrane potential imaging, the cells were incubated with rhodamine123 dye (0.5 μ M; Sigma Aldrich) in combination with live Hoechst 33342 for 2 hours, thereafter we refreshed with complete medium containing rhodamine 123 only (0.075 μ M) for time-lapse imaging. Then complete medium containing vehicle or NCT-505 (30 μ M final concentration) was added and the plates were directly imaged onto the microscope stage for live cell imaging.

For the cell death experiment and mitochondrial membrane potential imaging, the plates were imaged at 0, 14, 24, 40, and 72 hours after the compound exposure. The cell cycle was imaged every hour after the exposure. The imaging was performed using a Nikon TiE2000 confocal laser microscope (lasers : 647 nm, 540 nm, 488 nm, and 408 nm), equipped with automated stage and perfect focus system. During the imaging, the plates were maintained in humidified atmosphere at 37 °C and 5% CO₂. The experiment was conducted with 3 biological replicates with 4-5 technical replicates.

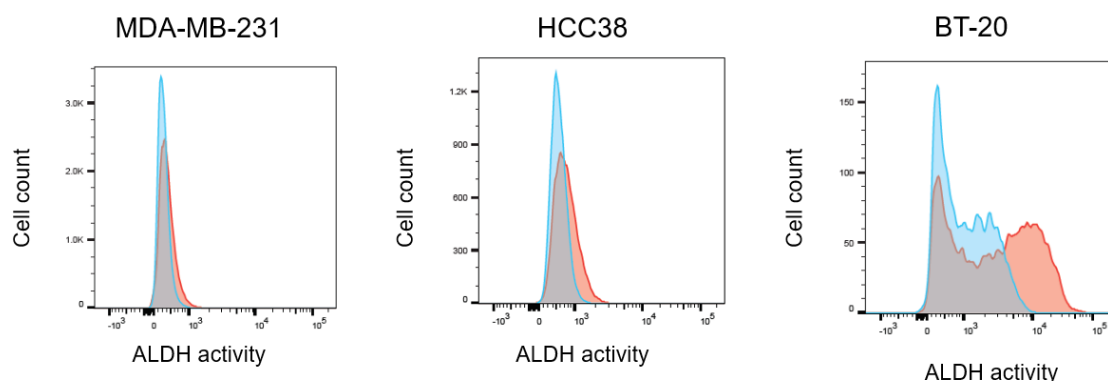
Image analysis. The quantitative image analysis was performed with ImageJ version 1.52p, CellProfiler version 2.2.0 and Ilastik 1.3.2. Firstly, the nuclei per image were segmented using the watershed masked algorithm on ImageJ. For images obtained from the cell death and mitochondrial membrane potential experiment, the nuclear segmentation was performed with Ilastik 1.3.2. The images were processed with an in house developed CellProfiler module.^{39,40} The images from the FUCCI cell line were analyzed and identified the object nuclei, geminin, and cdt with the IdentifyPrimaryObject module. The object geminin and cdt were masked with the segmented nuclei with the MaskImage module to eliminate the background. The cells expressing both proteins were identified as the yellow object by masking the images from the red channel with the geminin objects. The number of the nuclear object, geminin object, cdt object, and yellow object were counted as the outcome. For the cell death and mitochondrial membrane potential experiment , the nuclear Hoechst 33342 intensity levels, rhodamine 123 integrated intensity, PI area, and Annexin V area were measured at the single cell level. The results were stored as HDF5 files. Data analysis, quality control, and graphics were performed using the in house developed R package h5CellProfiler (manuscript in preparation).

Supplementary data

Luminal



Basal



Supplementary Fig. 5.1 | Additional graphs for Figure 5.2a. Representative FACS plots of ALDEFLUOR assay in breast cancer cell lines. Cells treated with ALDEFLUOR (red) and a negative control in which cells were pre-treated with DEAB (blue).

References

1. Croker, A. K. & Allan, A. L. Inhibition of aldehyde dehydrogenase (ALDH) activity reduces chemotherapy and radiation resistance of stem-like ALDH hiCD44 + human breast cancer cells. *Breast Cancer Res. Treat.* **133**, 75–87 (2012).
2. Qiu, Y. *et al.* The expression of aldehyde dehydrogenase family in breast cancer. *J. Breast Cancer* **17**, 54–60 (2014).
3. Marcato, P. *et al.* Aldehyde Dehydrogenase Activity of Breast Cancer Stem Cells Is Primarily Due To Isoform ALDH1A3 and Its Expression Is Predictive of Metastasis. *Stem Cells* **29**, 32–45 (2011).
4. Sládek, N. E., Kollander, R., Sreerama, L. & Kiang, D. T. Cellular levels of aldehyde dehydrogenases (ALDH1A1 and ALDH3A1) as predictors of therapeutic responses to cyclophosphamide-based chemotherapy of breast cancer: A retrospective study. *Cancer Chemother. Pharmacol.* **49**, 309–321 (2002).
5. Tomita, H., Tanaka, K., Tanaka, T. & Hara, A. Aldehyde dehydrogenase 1A1 in stem cells and cancer. *Oncotarget* **7**, 11018–32 (2016).
6. De Beca, F. F. *et al.* Cancer stem cells markers CD44, CD24 and ALDH1 in breast cancer special histological types. *J. Clin. Pathol.* **66**, 187–191 (2013).
7. Luo, Y. *et al.* ALDH1A isozymes are markers of human melanoma stem cells and potential therapeutic targets. *Stem Cells* **30**, 2100–2113 (2012).
8. Zhou, L. *et al.* Identification of cancer-type specific expression patterns for active aldehyde dehydrogenase (ALDH) isoforms in ALDEFLUOR assay. *Cell Biol. Toxicol.* **35**, 161–177 (2019).
9. Koppaka, V. *et al.* Aldehyde Dehydrogenase Inhibitors: a Comprehensive Review of the Pharmacology, Mechanism of Action, Substrate Specificity, and Clinical Application. *Pharmacol. Rev.* **64**, 520–539 (2012).
10. Yasgar, A. *et al.* A High-Content assay enables the automated screening and identification of small molecules with specific ALDH1A1-Inhibitory activity. *PLoS One* **12**, 1–19 (2017).
11. Yang, S. M. *et al.* Discovery of Orally Bioavailable, Quinoline-Based Aldehyde Dehydrogenase 1A1 (ALDH1A1) Inhibitors with Potent Cellular Activity. *J. Med. Chem.* **61**, 4883–4903 (2018).
12. Koenders, S. T. A. *et al.* Development of a Retinal-Based Probe for the Profiling of Retinaldehyde Dehydrogenases in Cancer Cells. *ACS Cent. Sci.* **5**, 1965–1974 (2019).
13. Hollestelle, A. *et al.* Distinct gene mutation profiles among luminal-type and basal-type breast cancer cell lines. *Breast Cancer Res. Treat.* **121**, 53–64 (2010).
14. Storms, R. W. *et al.* Isolation of primitive human hematopoietic progenitors on the basis of aldehyde dehydrogenase activity. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 9118–9123 (1999).
15. Wang, J. *et al.* Phosphorylation-dependent regulation of ALDH1A1 by Aurora kinase A: Insights on their synergistic relationship in pancreatic cancer. *BMC Biol.* **15**, 1–22 (2017).
16. Zhao, D. *et al.* NOTCH-induced aldehyde dehydrogenase 1A1 deacetylation promotes breast cancer stem cells. *J. Clin. Invest.* **124**, 5453–5465 (2014).
17. Yoshida, A., Hsu, L. C. & Dave, V. Retinal oxidation activity and biological role of human cytosolic aldehyde dehydrogenase. *Enzyme* **46**, 239–244 (1992).
18. Kang, J. H. *et al.* Aldehyde dehydrogenase is used by cancer cells for energy metabolism. *Exp. Mol. Med.* **48**, 1–13 (2016).
19. Burgos, C., de Burgos, N. M. G., Rovai, L. E. & Blanco, A. In vitro inhibition by gossypol of oxidoreductases from human tissues. *Biochem. Pharmacol.* **35**, 801–804 (1986).
20. Raha, D. *et al.* The cancer stem cell marker aldehyde dehydrogenase is required to maintain a drug-tolerant tumor cell subpopulation. *Cancer Res.* **74**, 3579–3590 (2014).
21. Sakaue-Sawano, A. *et al.* Visualizing Spatiotemporal Dynamics of Multicellular Cell-Cycle Progression. *Cell* **132**, 487–498 (2008).
22. Dixon, S. J. *et al.* Ferroptosis: An iron-dependent form of nonapoptotic cell death. *Cell* **149**, 1060–1072 (2012).
23. Hangauer, M. J. *et al.* Drug-tolerant persister cancer cells are vulnerable to GPX4 inhibition. *Nature* **551**, 247–250 (2017).
24. Viswanathan, V. S. *et al.* Dependency of a therapy-resistant state of cancer cells on a lipid peroxidase pathway. *Nature* **547**, 453–457 (2017).
25. Marcato, P., Dean, C. A., Giacomantonio, C. A. & Lee, P. W. K. Aldehyde dehydrogenase its role as a cancer stem cell marker comes down to the specific isoform. *Cell Cycle* **10**, 1378–1384 (2011).
26. Reynolds, D. S. *et al.* Breast Cancer Spheroids Reveal a Differential Cancer Stem Cell Response to Chemotherapeutic Treatment.

- Sci. Rep.* **7**, 1–12 (2017).
27. Fujiwara, D., Kato, K., Nohara, S., Iwanuma, Y. & Kajiyama, Y. The usefulness of three-dimensional cell culture in induction of cancer stem cells from esophageal squamous cell carcinoma cell lines. *Biochem. Biophys. Res. Commun.* **434**, 773–778 (2013).
 28. Yang, W. S. *et al.* Regulation of ferroptotic cancer cell death by GPX4. *Cell* **156**, 317–331 (2014).
 29. Schneider, C., Tallman, K. A., Porter, N. A. & Brash, A. R. Two distinct pathways of formation of 4-hydroxynonenal. Mechanisms of nonenzymatic transformation of the 9- and 13-hydroperoxides of linoleic acid to 4-hydroxyalkenals. *J. Biol. Chem.* **276**, 20831–20838 (2001).
 30. Yoval-Sánchez, B. & Rodríguez-Zavala, J. S. Differences in susceptibility to inactivation of human aldehyde dehydrogenases by lipid peroxidation byproducts. *Chem. Res. Toxicol.* **25**, 722–729 (2012).
 31. 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).
 32. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
 33. Distler, U., Kuharev, J., Navarro, P. & Tenzer, S. Label-free quantification in ion mobility-enhanced data-independent acquisition proteomics. *Nat. Protoc.* **11**, 795–812 (2016).
 34. Distler, U. *et al.* Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics. *Nat. Methods* **11**, 167–170 (2014).
 35. Kuharev, J., Navarro, P., Distler, U., Jahn, O. & Tenzer, S. In-depth evaluation of software tools for data-independent acquisition based label-free quantification. *Proteomics* **15**, 3140–3151 (2015).
 36. Vizcaino, J. A. *et al.* ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nature Biotechnology* **32**, 223–226 (2014).
 37. Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* **47**, 442–450 (2019).
 38. Riaz, M. *et al.* MiRNA expression profiling of 51 human breast cancer cell lines reveals subtype and driver mutation-specific miRNAs. *Breast Cancer Res.* **15**, R33 (2013).
 39. Wink, S., Hiemstra, S., Herpers, B. & van de Water, B. High-content imaging-based BAC-GFP toxicity pathway reporters to assess chemical adversity liabilities. *Arch. Toxicol.* **91**, 1367–1383 (2017).
 40. Yan, K. & Verbeek, F. J. Segmentation for high-throughput image analysis: Watershed masked clustering. in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **7610**, 25–41 (Springer, Berlin, Heidelberg, 2012).

Chapter 6

Visualizing Vitamin A Metabolism in the Small Intestines

Introduction

Retinoic acid (RA) is essential for embryonic development, cell differentiation and immunomodulation.^{1,2} Its importance in immunology has already been recognized for 100 years.³ In recent years the role of RA in the regulation of the immune system has been revisited and was shown to be important in the gut-homing of lymphocytes and function of dendritic cells (DCs).⁴⁻⁶

The precursors of RA, retinol (vitamin A) and β -carotene, are taken up from the diet by the epithelial cells of the small intestines.⁷ These cells convert RA precursors into retinal. The next step, the conversion of retinal, which can be oxidized into RA by retinaldehyde dehydrogenases, such as ALDH1A1, ALDH1A2 and ALDH1A3 (**Fig. 6.1**).^{8,9}

RA is required for the activation of DCs in the tissue of the small intestines.¹⁰ RA binds to retinoic acid receptors (RAR), which modulate gene transcription and induce the expression of ALDH1A2 in DCs, thereby activating these cells.^{10–12} As a result, the DCs are capable of producing RA by themselves. The activated DCs travel from the intestinal tissue to the gut-draining mesenteric lymph nodes, where they provide RA to lymphocytes (T and B cells).¹⁰ RA induces the expression of the integrin $\alpha 4\beta 7$ and the chemokine receptor CCR9 in the lymphocytes, which may bind to the mucosal vascular addressin cell adhesion molecule-1 (MadCAM-1) and the chemokine CCL25, respectively.^{13,14} This process is termed the “imprinting of lymphocytes”. Once the RA-imprinted lymphocytes leave the mesenteric lymph nodes, they travel through the bloodstream to find their interaction partners, which are expressed by the endothelial cells lining the vasculature of the small intestines. The RA-imprinted lymphocytes leave the bloodstream and enter the small intestines. The RA imprinting of lymphocytes by DCs therefore induces a gut-homing phenotype (**Fig. 6.1**).^{5,6}

RA imprinted T and B cells undergo several cell-type specific changes. RA differentiates B cells into IgA plasma cells, which secrete IgA into the intestinal lumen. These cells are required for an effective control of the gut microbiome.¹⁵ Activated T cells express Foxp3, which leads in the presence of TGF- β to the differentiation of T cells into immunosuppressive regulatory T (Treg) cells instead of the proinflammatory T helper 17 (Th17) cells.^{16–19} RA controls, therefore, the balance between Treg and Th17 cells in the gut associated lymphoid tissue, which is crucial for immune homeostasis. An imbalance between Treg and Th17 cells may lead to autoimmune pathologies, such as inflammatory bowel disease, rheumatoid arthritis, multiple sclerosis and asthma, which can be characterized by an excess of Th17 cells.⁴

Currently, it is unknown which ALDH subtype is responsible for the production of RA required for the activation of DCs and the imprinting of the lymphocytes. Genetic deletion or pharmacological inhibition of these enzymes could shed light on this process. There are, however, no truly selective and potent ALDH inhibitors available²⁰ and constitutive genetic deletion of ALDH1A2 or ALDH1A3 is embryonically lethal.²¹

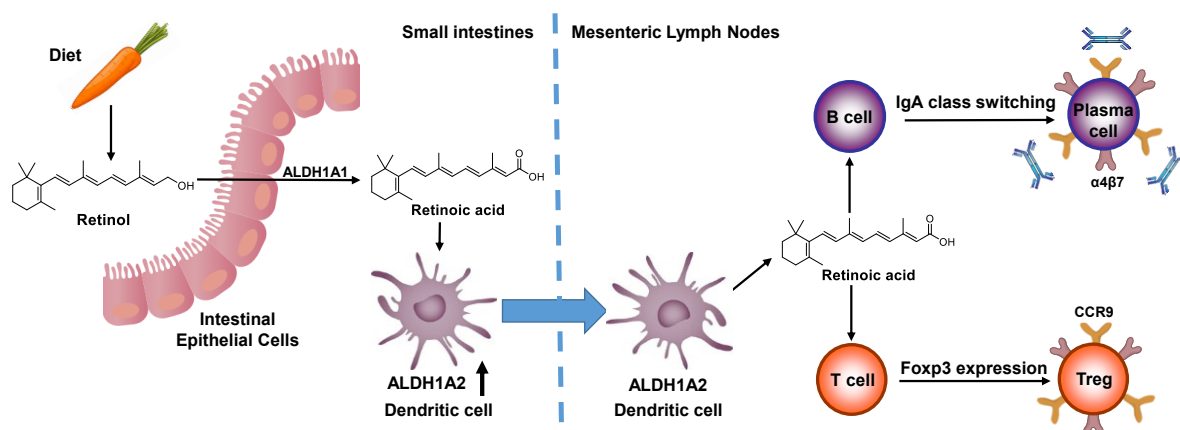


Fig. 6.1 | Schematic representation of the imprinting of DCs and lymphocytes by RA. On the left the uptake and conversion of retinol into RA by intestinal epithelial cells expressing ALDH1A1 is depicted. This results in a gradient of RA capable of imprinting DCs present in the underlying tissue of the small intestines. The retinoic acid binds to the nuclear retinoic acid receptor of the DCs and induces the expression of retinaldehyde dehydrogenase ALDH1A2. The activated DCs then travel to the mesenteric lymph nodes where they can provide T and B cells with the RA stimulus required for the expression of the gut-homing integrin $\alpha 4 \beta 7$ and chemokine receptor CCR9. Furthermore, the DCs stimulate T and B cells with RA to differentiate into IgA producing plasma cells and Foxp3 expressing Treg cells, respectively.

ALDH1A1 knock-out mice were viable, but did not display any significant phenotype, such as differences in Treg cells and fecal IgA levels (**Supplementary Fig. 6.1**).²² This suggested that ALDH1A1 is not essential for the production of RA to activate DCs and to imprint lymphocytes. Of note, increased ALDH1A3 expression levels in the intestinal epithelial cells of vitamin A deficient mice have been previously described.¹⁰ This may suggest that other aldehyde dehydrogenases (ALDHs) could be upregulated in the intestinal epithelial cells to compensate for the loss of ALDH1A1 in the knockout (KO) mice.

Activity-based protein profiling (ABPP) is a powerful chemical biological method that is ideally suited to report on the abundance and identity of active enzymes in complex biological systems.²³ It is a technology that relies on mechanism-based chemical probes that covalently and irreversibly react with the catalytic nucleophile in the active site of an enzyme in their native biological context.²⁴ Using the retinal-based probe, **LEI-945**²⁵, introduced and characterized in **chapters 3-5**, it was envisioned that retinaldehyde dehydrogenase activity could be identified and quantified in cell extracts derived from mouse intestines.

This chapter describes the development and application of a comparative ABPP method to profile the retinaldehyde dehydrogenase activities in intestinal cells derived from *Aldh1a1*^{WT} and *Aldh1a1*^{-/-} mice. Chemical proteomics indicated that ALDH1B1 activity may contribute to the conversion of retinal into RA in the epithelial cells of the small intestines. Analysis of the ALDH1A1/ALDH1B1 double knock-out animals revealed a proinflammatory phenotype in which the gut-homing of Treg cells was significantly decreased. These data suggest that ALDH1B1 has retinaldehyde dehydrogenase activity, and is partly responsible for the production of RA in intestinal epithelial cells to induce Treg cells.

Results

*Comparative ABPP of retinaldehyde dehydrogenases in intestinal epithelial cells derived from *Aldh1a1*^{WT} mice and *Aldh1a1*^{-/-} mice*

Since ALDH activity relies on a catalytic cysteine, which can be easily oxidized in non-cellular environments, chemical proteomics experiments were performed in intact cells to ensure a physiologically relevant readout. Freshly harvested epithelial cells from the small intestines and cells from the mesenteric lymph nodes of wild type mice were treated with **LEI-945** (1 μ M, 1 h), washed, pelleted and snap frozen until the proteomics experiment was performed. The same procedure for label-free quantification (LFQ) proteomic analysis was followed as described in **Chapters 4 and 5**. In short, the cells were thawed, lysed in MilliQ and subjected to Copper(I)-catalyzed alkyne-azide [2+3] cycloaddition with a biotin-azide. Incubation with avidin beads, followed by several washing steps, tryptic digestion, desalting and analysis using LC-MS/MS. This chemical proteomics experiment resulted in the detection of over 300 proteins.

ALDH1A1, ALDH1A7, ALDH1B1, ALDH2 and ALDH3A2 were significantly enriched in the intestinal epithelial cells (**Fig. 6.2a**), whereas ALDH1A2, ALDH2 and ALDH3A2 were identified in cells from the mesenteric lymph nodes (**Fig. 6.2b**). Based on the previously reported tissue distribution of ALDHs in mice, this subset covers all the ALDHs expressed in the intestinal cells of mice.²⁶ An overview of all significantly enriched enzymes in the wild type intestinal cells is shown in **Supplementary Tables 6.1 and 6.2**.

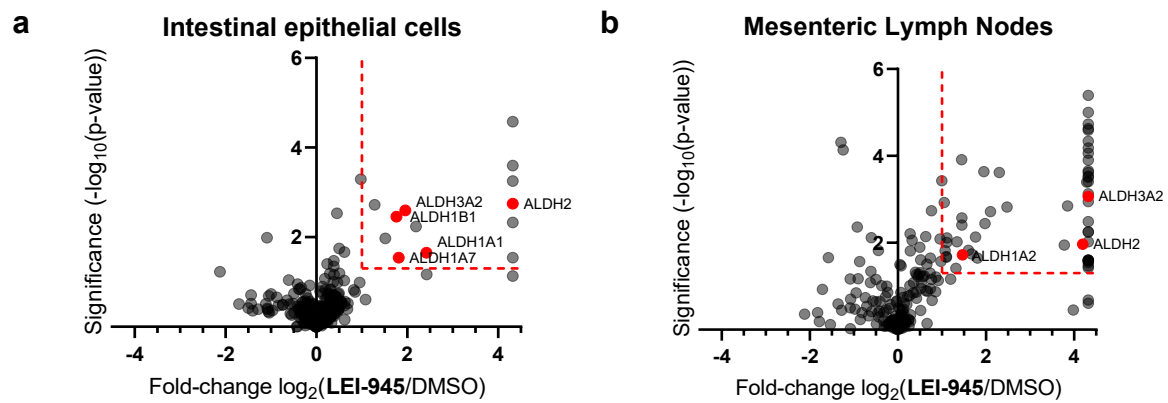


Fig. 6.2 | Chemical proteomics data of ALDH enzymes intestinal cells derived from wild type mice.

a, Volcano plot for total proteins identified in a chemical proteomics experiment with probe **LEI-945** (1 μ M) in intestinal epithelial cells (IEC). Red lines indicate threshold values marking significantly enriched (fold-change > 2; p-value < 0.05) proteins. Red dots represent significantly enriched ALDH enzymes. **b**, Volcano plot for total proteins identified in a chemical proteomics experiment with probe **LEI-945** (1 μ M) in mesenteric lymph node (MLN) cells. Red lines indicate threshold values marking significantly enriched (fold-change > 2; p-value < 0.05) proteins. Red dots represent significantly enriched ALDH enzymes. For parts **a** and **b**, data are from $N = 4$ experiments (biological replicates).

To determine which retinaldehyde dehydrogenase may compensate for the loss of ALDH1A1 activity in the small intestines of *Aldh1a1*^{-/-} mice, a comparative chemical proteomics experiment using **LEI-945** was performed using epithelial cells from *Aldh1a1*^{WT} and *Aldh1a1*^{-/-} mice. To this end, the LFQ values of ALDHs from *Aldh1a1*^{WT} and *Aldh1a1*^{-/-} mice were determined (**Fig. 6.3a**). Chemical proteomics confirmed that the knockout mice did not show any ALDH1A1 activity. Of note, the levels of active ALDH1A7, a mouse specific isozyme, which cannot convert retinal into retinoic acid^{26,27}, were also decreased in the intestinal epithelial cells from the *Aldh1a1*^{-/-} mice. This is in line with a previous report in which the loss of ALDH1A7 activity was attributed to their adjacent genetic localization on chromosome 19.²⁸ The levels of ALDH1B1 activity had notably increased in the *Aldh1a1*^{-/-} mice, whereas the levels of ALDH2 and ALDH3A2 were unchanged. ALDH1B1 and ALDH2 were the most abundant ALDH activities in intestinal epithelial cells.

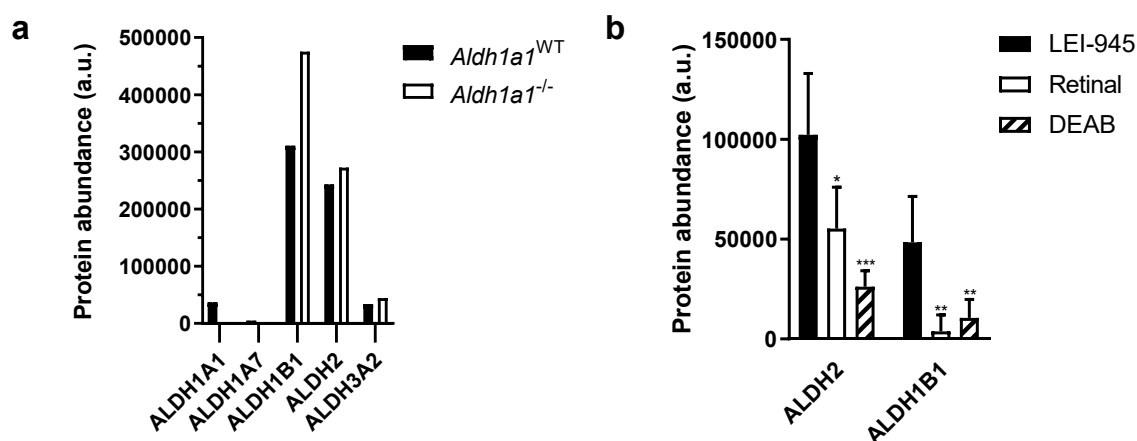


Fig. 6.3 | Comparative and competitive ABPP of ALDH enzymes in intestinal epithelial cells. **a**, Data represent mean LFQ values for ALDHs significantly enriched in intestinal epithelial cells derived from *Aldh1a1*^{WT} and *Aldh1a1*^{-/-} mice. Data are from $N \geq 1$ experiments (biological replicates). **b**, Competitive ABPP data from competition of retinal (30 μ M) or DEAB (30 μ M) with **LEI-945** (1 μ M) in mouse intestinal epithelial cells. Data represent mean LFQ values $\pm$ SD; $N = 4$ experiments (biological replicates). $P^* < 0.05$, $P^{**} < 0.01$ and $P^{***} < 0.001$; t test, two-sided.

Competitive ABPP experiments with general ALDH inhibitor 4-diethylaminobenzaldehyde (DEAB) and the natural substrate retinal were performed to exclude non-specific labeling (**Fig. 6.3b**). Competitive ABPP with DEAB showed a significant decrease in labeling for ALDH1B1 and ALDH2, whereas competition with retinal showed a significant, but small (less than 50%) inhibition of ALDH2, while ALDH1B1 showed almost complete inhibition by the natural substrate (**Fig. 6.3b**).

Since ALDH2 is unable to convert retinal^{29,30}, it is considered an unlikely candidate to compensate for the loss of ALDH1A1 activity. Although acetaldehyde was previously described as the main substrate of ALDH1B1³¹, retinal can also serve as a substrate for recombinant human ALDH1B1 in *in vitro* assays.³⁰ Of note, ALDH1B1 is primarily expressed in the small intestines²⁶ and upregulated in mice fed a high fiber diet in a similar fashion as ALDH1A1 (data not shown).⁸ Of the ALDHs identified using ABPP, ALDH1B1 has one of the highest levels of active enzyme in mouse epithelial cells. Together, these data suggest a possible tissue specific role of ALDH1B1 in RA synthesis and immune homeostasis and that it could compensate for the loss of ALDH1A1.

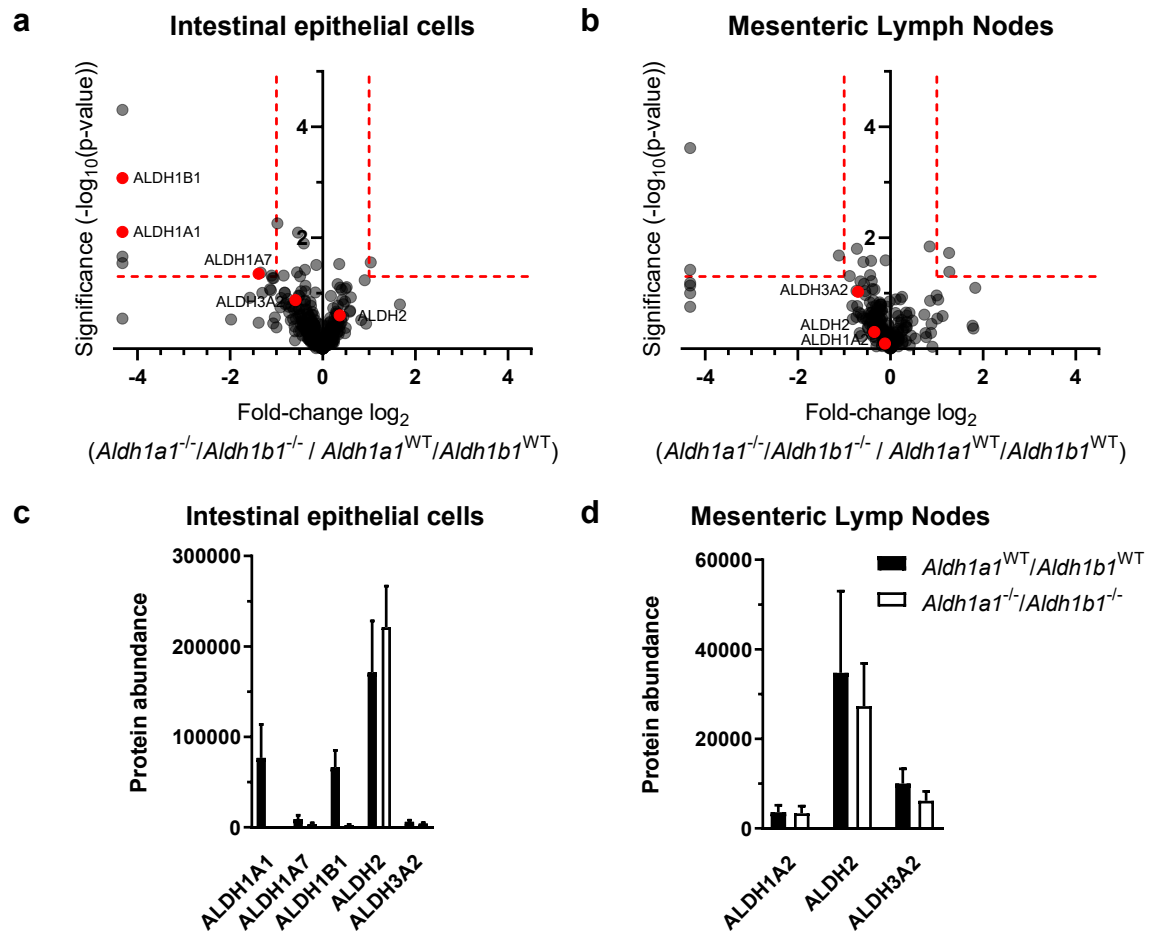


Fig. 6.4 | Comparative ABPP of ALDHs in intestinal cells derived from $Aldh1a1^{WT}/Aldh1b1^{WT}$ and $Aldh1a1^{-/-}/Aldh1b1^{-/-}$ mice. **a**, Volcano plot for total proteins identified in chemical proteomics experiment with probe **LEI-945** (1 μM) comparing intestinal epithelial cells derived from $Aldh1a1^{WT}/Aldh1b1^{WT}$ with $Aldh1a1^{-/-}/Aldh1b1^{-/-}$ mice. Red lines indicate threshold values marking significantly enriched (fold-change > 2 ; p-value < 0.05) proteins. Red dots represent significantly enriched ALDH enzymes. **b**, Volcano plot for total proteins identified in chemical proteomics experiment with probe **LEI-945** (1 μM) comparing cells derived from the mesenteric lymph nodes of $Aldh1a1^{WT}/Aldh1b1^{WT}$ and $Aldh1a1^{-/-}/Aldh1b1^{-/-}$ mice. Red lines indicate threshold values marking significantly enriched proteins. Red dots represent significantly enriched ALDH enzymes. **c**, Protein abundance of ALDHs. Data represent mean LFQ values $\pm$ SD for significantly enriched ALDHs in intestinal epithelial cells derived from $Aldh1a1^{WT}/Aldh1b1^{WT}$ or $Aldh1a1^{-/-}/Aldh1b1^{-/-}$ mice. **d**, Protein abundance of ALDHs. Data represent mean LFQ values $\pm$ SD for significantly enriched ALDHs in cells derived from the mesenteric lymph nodes of $Aldh1a1^{WT}/Aldh1b1^{WT}$ and $Aldh1a1^{-/-}/Aldh1b1^{-/-}$ mice. For parts **a-d**, $N = 4$ experiments (biological replicates).

Aldh1a1^{-/-}/Aldh1b1^{-/-} mice display a more proinflammatory phenotype

ALDH1B1 has not yet been studied in the context of gut-homing and immune homeostasis. To determine whether ALDH1B1 plays a role in the regulation of the immune system within the small intestines, mice deficient for both *Aldh1a1* and *Aldh1b1* genes were investigated. Chemical proteomics using **LEI-945** confirmed the epithelial cells derived from the small intestines of these transgenic mice were indeed lacking ALDH1A1 and ALDH1B1 activity (**Fig. 6.4a,c**).

Furthermore, the ABPP analysis indicated that the activity of ALDH2 and ALDH3A2 did not change significantly, while ALDH1A7 activity was decreased (**Fig 6.4a,c**). ALDH1A2, ALDH2 and ALDH3A2 were detected in cells derived from the mesenteric lymph nodes (**Fig. 6.4b,c**). The ALDHs detected in the mesenteric lymph nodes, did not show any significant change in activity in the double-knockout compared with the wild type.

To determine whether *Aldh1a1^{-/-}/Aldh1b1^{-/-}* mice did display an altered phenotype, the amount of Treg cells, activated DCs and IgA levels were measured. The presence of Treg cells in *Aldh1a1^{-/-}/Aldh1b1^{-/-}* mice was determined within the tissue of the intestines using fluorescence-activated cell sorting (FACS). A significant decrease in the frequency of Foxp3⁺CD25⁺ Treg cells was detected (**Fig. 6.5a**). This decrease in immunosuppressive Treg cells was accompanied by an increase of CD45⁺ leukocytes within the epithelium (**Fig. 6.5b**). There was no change in the amount of activated DCs present (**Fig. 6.5c,d**). Of note, the amount of IgA in the lumen was also significantly increased (**Fig. 6.5e**). These data suggest that the combined absence of ALDH1A1 and ALDH1B1 proteins leads to a shift in the immune homeostasis of the small intestines towards a proinflammatory phenotype.

Discussion

The hypothesis that production of RA by the intestinal epithelial cells required for the effective imprinting of dendritic cells depends on ALDH1A1 activity, was previously contested by a lack of disruption in the immune homeostasis in *Aldh1a1^{-/-}* mice. To investigate which ALDH enzyme could compensate for the loss of ALDH1A1 activity, comparative ABPP using **LEI-945** in intestinal epithelial cells derived from wild type and ALDH1A1 knockout mice was performed. High levels of ALDH1B1 and ALDH2 activity were identified in the intestinal epithelial cells.

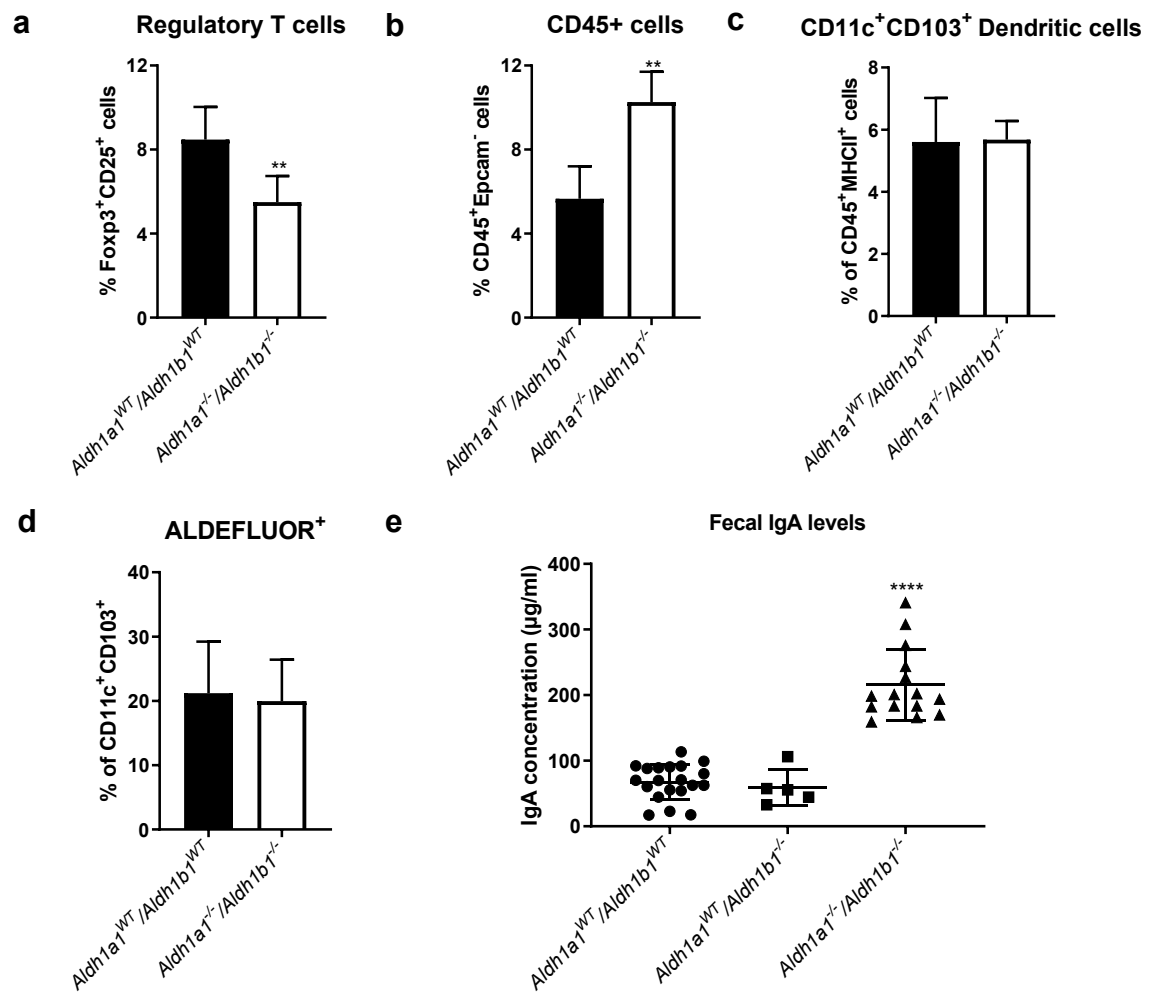


Fig. 6.5 | *Aldh1a1*^{-/-}/*Aldh1b1*^{-/-} mice display a more proinflammatory phenotype. **a**, Graph showing the proportion of regulatory T cells (Foxp3⁺CD25⁺) in CD45⁺CD3⁺CD4⁺ T cells. Data represent mean values $\pm$ SD; *N* = 6 experiments (biological replicates). ***P* < 0.01; Mann-Whitney test; two-tailed. **b**, Graph showing the proportion of leukocytes (CD45⁺) in the small intestinal epithelial layer. Data represent mean values $\pm$ SD; *N* = 5 experiments (biological replicates). ***P* < 0.01; Mann-Whitney test, two-tailed. **c**, Graph showing the proportion of CD11c⁺CD103⁺ dendritic cells in CD45⁺MHCII⁺ cells. **d**, Graph showing the proportion of ALDEFLUOR⁺ cells in CD11c⁺CD103⁺ dendritic cells. For parts **c** and **d**, data represent mean values $\pm$ SD; *N* = 4 experiments (biological replicates). **e**, Graph showing the amount of IgA in the lumen of *Aldh1a1*^{WT}/*Aldh1b1*^{WT}, *Aldh1a1*^{WT}/*Aldh1b1*^{-/-} and *Aldh1a1*^{-/-}/*Aldh1b1*^{-/-} mice. Data represent mean values $\pm$ SD; *N* = 14. *****P* < 0.0001; Mann-Whitney test, two-tailed.

However, the probe labelling of ALDH1B1 could be efficiently competed with retinal, whereas labelling of ALDH2 only decreased modestly. This suggested that ALDH1B1 could be an alternative enzyme capable of producing RA in the small intestines.

To date, the role of ALDH1B1 in immune homeostasis has not been investigated, therefore the gut-homing of lymphocytes in transgenic mice lacking both *Aldh1a1* and *Aldh1b1* genes was studied. FACS analysis of the immune cell populations in the small intestines of these mice showed a lymphocyte composition corresponding with a more proinflammatory phenotype.

Chemical proteomics experiments on intestinal cells derived from the *Aldh1a1*^{-/-}/*Aldh1b1*^{-/-} mice showed no changes in the levels of other ALDHs to compensate for the loss of ALDH activity. There was also no alteration in the amount of ALDH1A2 activity in the mesenteric lymph nodes, corresponding with the observation of unchanged levels of DCs with high ALDH activity or ALDEFLUOR⁺ DCs. Therefore, only the ALDH activity in the epithelial layer seems to be directly affected by the *Aldh1a1*^{-/-}/*Aldh1b1*^{-/-} knockout.

Although the *Aldh1a1*^{-/-}/*Aldh1b1*^{-/-} knockout mice did show a proinflammatory phenotype, only the expected decrease in the amount of Treg cells was observed. The amount of activated DCs was not affected and IgA levels increased, while a decrease of both was expected as a result of lower RA levels. Several factors might be considered in order to explain these seemingly contradictory results: retinoid levels, composition of the gut microbiome and short-chain fatty acid (SCFA) levels.

In a recent study it has been shown that the gut microbiome influences vitamin A metabolism by subduing retinol dehydrogenase 7 (*Rdh7*) expression in intestinal epithelial cells.³² Suppressing the levels of *Rdh7*, which converts retinol into retinal, lowered the formation of RA locally in the epithelial layer. As a result, the conversion of retinol shifted to the production of retinyl esters for storage in the liver, while retinaldehyde dehydrogenase levels remained unchanged.³² This shift is favourable to commensal bacteria as RA enhances the production of IL-22, which supports epithelial integrity and the antimicrobial response via the production of antimicrobial peptides, such as Reg3β and Reg3γ.³³ A knockout of *Rdh7* led to 3-fold lower levels of RA and a reduction of IL-22, Reg3β and Reg3γ.³² However, in the *Aldh1a1*^{-/-}/*Aldh1b1*^{-/-} mice no significant alteration in the levels of these antimicrobial markers was detected (data not shown).

This beckons the question to what extent the retinoid levels in the intestinal tissue of the knockout mice have changed and whether the observed changes are due to a change in RA levels or affect the intestinal environment in other ways. To answer these questions the actual levels of RA in the intestinal tissues of these mice will have to be measured. Additionally, the LC-UV/MS assay introduced in **Chapter 4** could be used to determine the ability of intestinal epithelial cells derived from wild type and knockout mice to convert retinal into RA.

Nonetheless, there was a significant reduction in the proportion of Treg cells in the intestinal tissue of *Aldh1a1^{-/-}/Aldh1b1^{-/-}* mice. The decrease in Treg cells was also accompanied by an increase of leukocytes in the epithelium. This phenotype resembles that of pathologies such as coeliac disease.³⁴ However, the previously proposed model of disrupted activation of DCs resulting in less Treg cells, cannot explain this observed reduction as no difference in DC subsets in the mesenteric lymph nodes were observed. A more detailed characterization of the leukocytes in the epithelial layer of *Aldh1a1^{-/-}/Aldh1b1^{-/-}* mice might provide insight into the factors influencing the observed changes.

Remarkably, an increase in fecal IgA levels was detected instead of the expected decrease due to reduced lymphocyte recruitment. A possible explanation might be found in the composition of the gut microbiome. A balanced gut microbiome is essential for immune homeostasis as germ free mice are ineffective at sustaining Tregs.³⁵ This has been linked to the short-chain fatty acids (SCFAs), such as acetate, propionate and butyrate, produced by commensal bacteria from dietary fiber.^{8,36–38} SCFAs can effect gut homeostasis via two pathways: histone deacetylase inhibition and binding to G-protein-coupled receptors, such as GPR43 and GPR109a.³⁹ Propionate and butyrate can inhibit histone deacetylases, inducing ALDH1A2 expression in dendritic cells and Foxp3 expression in T cells^{8,38}, whereas binding of acetate to GPR43 induces ALDH1A2 expression in dendritic cells and an increase of IgA production in B cells.⁴⁰

Changes in the composition of the gut microbiome or an aberrant reaction to the microbiome, are therefore other factors to be considered. However, SCFA levels and the composition of the bacteria making up the gut microbiome have not been determined in this study. These could provide an answer for the discrepancy between the expected decrease in IgA producing plasma cells and the detected increase in fecal IgA.

Conclusion

ABPP with **LEI-945** was capable of determining the levels of active ALDH enzymes in intestinal cells derived from mice. This shows that **LEI-945** can not only be used for the profiling of ALDHs in cultured (cancer) cell lines, but also in primary material *ex vivo*. Using both comparative and competitive ABPP, ALDH1B1 was identified as a potential enzyme contributing to the production of RA in the small intestines. *Aldh1a1^{-/-}/Aldh1b1^{-/-}* mice showed a proinflammatory phenotype with an influx of leukocytes in the epithelium, a decrease in Treg cells and higher levels of fecal IgA.

Acknowledgements

Martje Erkelens is kindly acknowledged for harvesting the intestinal cells from mice, performing the FACS assays, her collaboration on preparing the proteomics samples and helpful discussions. Bogdan Florea is acknowledged for mass spectrometry analysis.

Experimental procedures

Biological methods

Materials, probe and inhibitors. Activity-based probe **LEI-945** was synthesized in-house as previously described in **Chapter 3**. DEAB and retinal were purchased from Sigma Aldrich.

Mice. *Aldh1a1* deficient mice¹⁴ were obtained from The Jackson Laboratory, USA (stock #012247), and were bred heterozygous x heterozygous so both wild type and knock out animals were available from the same litter. Both male and female mice were used aged 6-20 weeks. Mice were housed under specific pathogen free conditions at the VU University Animal Facility. All experiments have been approved by the VU University Scientific and Animal Ethics Committees according to Dutch law. *Aldh1a1/1b1* deficient mice were created at Yale University and experiments were performed with age matched wild type animals. Mice were housed under specific pathogen free conditions at the Yale University Animal Facility, New Haven, CT USA. At the time of the experiments all mice were aged 8-12 weeks and experiments were performed in accordance with local ethical regulations.

Ex vivo single cell suspensions. Small intestines and mesenteric lymph nodes were dissected from mice and fatty tissue was removed as much as possible. Luminal contents were collected in 1 ml PBS and stored at -20 °C. Then, intestines were opened longitudinally and washed in phosphate-buffered saline without calcium (D-PBS) (Gibco, Life Technologies, Paisley, United Kingdom) to remove all fecal contents. To isolate intestinal epithelial cells, the intestines were transferred to 1x Hank's Balanced Salt Solution (HBSS) (Invitrogen, Breda, The Netherlands) containing 5% Fetal Calf Serum (FCS) (HyClone Laboratories/Greiner Bio-One, Alphen aan den Rijn, The Netherlands) and 2.5 mM EDTA (Sigma Aldrich, Zwijndrecht, The Netherlands) and incubated on ice for 30 minutes. Tissues were vortexed vigorously for the removal of epithelial cells. The remaining tissue and the mesenteric lymph nodes were minced and incubated separately while constantly stirring for 30 minutes in 37 °C in HBSS + 5% FCS + 0,5 mM collagenase IV (Roche, Penzberg, Germany) in order to obtain a single cell suspension of both the lamina propria and the mesenteric lymph nodes. The lamina propria suspension was filtered through a 70 µm cell-strainer (BD Biosciences, Breda, The Netherlands) and washed once with HBSS containing 5% FCS. Suspensions were either used for FACS analysis or ABPP.

Flow cytometry. The ALDEFLUOR assay kit (StemCell Technologies, Grenoble, France) was used to measure ALDH activity using flow cytometry according to the manufacturer's protocol. In some cases cells were incubated with an increasing series of retinal (Sigma Aldrich, Zwijndrecht, the Netherlands) as a competitive substrate while the ALDEFLUOR assay was performed. Subsequently, cells were stained with anti-Epcam-PE, anti-CD45-PE-Cy7, anti-CD11c e450 (eBioscience, Amsterdam, the Netherlands), anti-MHCII 647, anti-CD103-biotin (BD Bioscience, Breda, The Netherlands), and Streptavidin eFluor 605NC (both eBioscience, Amsterdam, the Netherlands). anti-podoplanin (clone GP38, made in house), anti-CD45-viogreen (Miltenyi Biotec, Leiden, The Netherlands), anti-CD31-PE-Cy7 (eBioscience, Amsterdam, the Netherlands) and goat-anti-hamster 647 (Invitrogen, Landsmeer, the Netherlands). 7-AAD (Thermo Fisher Scientific, Breda, The Netherlands), and DAPI (Invitrogen, Landsmeer, the Netherlands) were used to discriminate between live and death cells. Cells were analyzed with LSR Fortessa X-20 (BD care, Erembodegem, Belgium). Data was analyzed using FlowJo software (Tree Star, Ashland, OR, USA).

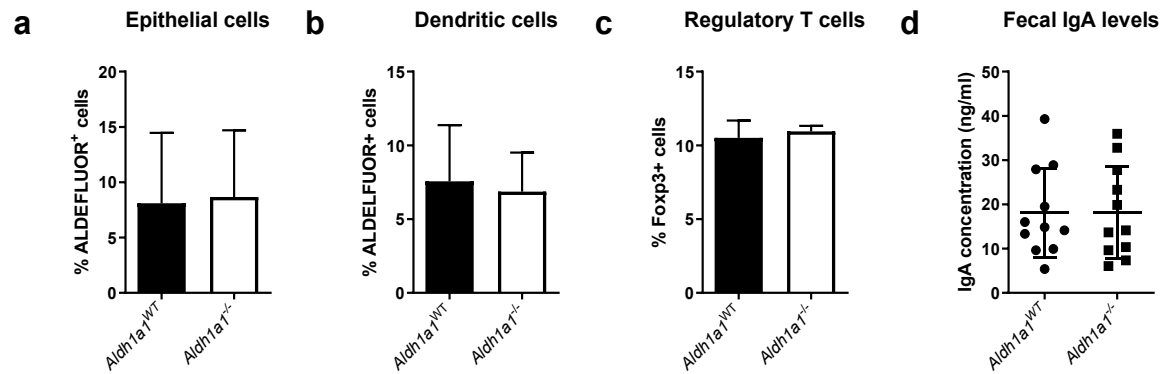
Enzyme-linked immunosorbent assay (ELISA) for secretory IgA. Collected luminal contents were thawed and shaken at 4 °C for 60 minutes to homogenize the samples. Then, samples were centrifuged at maximum speed for 5 minutes to remove the debris and supernatant was collected for analysis by ELISA. Plates were coated with anti-mouse-IgA antibody to capture secretory IgA or mouse-IgA used as a standard (clone s107), followed by anti-mouse-IgA-biotin (all Southern Biotech, Birmingham, AL, USA) antibody, and subsequently streptavidin labeled with HRP (Thermo Fischer Scientific, Landsmeer, the Netherlands). Samples were analyzed with a Fluostar Optima microplate reader (BMG Labtech, Isogen Lifescience, De Meern, the Netherlands).

Ex vivo activity-based proteomics

Sample preparation for chemical proteomics. Protocol adapted from previously described procedure.⁴¹ Cells were treated with vehicle or **LEI-945** (1 μ M) in PBS. In the case of competitive ABPP vehicle, DEAB (100 μ M) or retinal (100 μ M) was added. After 1 hour cells were washed, snap frozen and stored at -20 °C. When required cells were thawed, lysed and adjusted to 1 mg/mL protein concentration as determined by a Quick Start™ Bradford Protein assay (Bio-Rad). 250 μ L was taken from each sample and to this 25 μ L freshly prepared “click” mixture containing 1 mM CuSO₄ (2.5 μ L/sample, 100 mM in H₂O), 5 mM NaAsc (1.25 μ L/sample, 1 M in H₂O), 0.4 mM THPTA (1 μ L/sample, 100 mM in DMSO), 40 μ M biotin-N₃ (2.5 μ L/sample, 4 mM in DMSO) and MilliQ (17.75 μ L/sample) was added. Samples were incubated for 1 hour at 37 °C while shaking (300 rpm). Excess click reagents were then removed by chloroform/methanol precipitation followed by another wash with methanol. Precipitated proteomes were then suspended in urea buffer (250 μ L, 6 M urea and 25 mM ammonium bicarbonate), DTT (2.5 μ L, 1 M) was added and the mixture was then incubated for 15 min at 65 °C while shaking (600 rpm). The samples were then allowed to cool down to RT and then alkylated by addition of iodoacetamide (20 μ L, 0.5M) for 30 minutes at RT in the dark. Addition of SDS (70 μ L, 10% (v/v)) was followed by heating at 65 °C for 5 minutes. For each sample 50 μ L 50% slurry of Avidin-Agarose from egg white (Sigma-Aldrich) was washed three times with PBS and transferred in PBS (1 L) to a 15 mL tube. To this another 2 mL of PBS was added followed by the corresponding proteome sample. The beads were incubated with the proteome for 2 hours at room temperature using an overhead shaker. The beads were then isolated by centrifugation (2 min, 2500 g), washed with SDS in PBS (0.5% (w/v)) and washed three times with PBS. The beads were then transferred to low-binding Eppendorf tubes and proteins were digested overnight at 37 °C and 950 rpm in 250 μ L digestion buffer (100 mM Tris, 100 mM NaCl, 1 mM CaCl₂, 2% acetonitrile and 0.5 μ g sequencing grade trypsin (Promega)). Digestion was stopped by addition of formic acid (12.5 μ L) and the beads filtered off by centrifugation (2 min, 600 g) using a Bio-Spin column (Bio-Rad). Samples were then desalted using stage tips, collected in low-binding Eppendorf tubes, concentrated using a SpeedVac (Eppendorf) and stored at -20 °C until reconstitution before measurement.⁴²

LC-MS/MS measurement and analysis. Samples were reconstituted in LC-MS sample solution (50 μ L, MilliQ, 3% acetonitrile/0.1% formic acid/20 fmol/ μ L enolase). Samples were then analysed using a NanoACQUITY UPLC System (Waters) coupled to a SYNAPT G2-Si high-definition mass spectrometer (Waters) as previously described.^{41,43} Of each sample 5 μ L was loaded on a nanoEASE™ M/Z Symmetry C18 trap column (particles 5 μ m, 100 Å, 180 μ m x 20 mm, Waters) with 0.1% formic acid and separated on an nanoEASE™ M/Z HSS C18 T3 analytical column (particles 1.8 μ m, 75 μ m x 250 mm, Waters) heated at 80 °C. A multistep gradient running from 5-40% acetonitrile containing 0.1% formic acid during a 70 minute method at 300 nL/min was used to achieve peptide separation. Survey scans (m/z 50-2000 Da) were acquired in the Synapt with a scan time of 0.6 seconds in positive, resolution mode. The collision energy is set to 4 V in the trap cell for low-energy MS mode. For the elevated energy scan, the transfer cell collision energy is ramped using drift-time specific collision energies. The lock mass is sampled every 30 seconds. MS raw files were analysed with ProteinLynx Global SERVER (PLGS, v3.0.3, Waters). The MS^E identification was also performed with PLGS using the human proteome from Uniprot (Uniprot-homo-sapiens-trypsin-reviewed-2016-08-29.fasta). The following parameter settings were used: low energy threshold 150 counts, elevated energy threshold 30, peptide and protein FDR 1%, enzyme specificity trypsin, max missed cleavages max 2, variable modification methionine oxidation, fixed modification carbamidomethylation cysteine, at least: fragments/peptide 2, fragments/protein 5, peptides/protein 1 and number of peptides to measure per protein 3. For label-free quantification ISOQuant (v1.5) was used.^{44,45} Data were filtered to retain only proteins with two or more reported unique peptides and quantified in at least 3 replicates of the positive control (probe-treated). Proteins were designated as significantly enriched by the probe when they showed 2-fold enrichment in quantification value when comparing negative control (vehicle-treated) with positive control (probe-treated) samples and probability as determined by a Student's *t* test (<0.05).

Supplementary Data



Supplementary Fig. 6.1 | *Aldh1a1*^{-/-} mice show no change in immune homeostasis. **a**, Graph showing the proportion of ALDEFLUOR⁺ cells in the small intestine epithelial layer. **b**, Graph showing the proportion of ALDEFLUOR⁺ cells in dendritic (CD45⁺MHCII⁺CD11c⁺CD103⁺) cells. For parts **a** and **b**, data represent mean values $\pm$ SD; $N = 7$ experiments (biological replicates). **c**, Graph showing the proportion of Foxp3⁺ cells in T cells. Data represent mean values $\pm$ SD; $N = 4$ experiments (biological replicates). **d**, Graph showing the amount of IgA in the lumen of *Aldh1a1*^{WT} and *Aldh1a1*^{-/-} mice. Data represent mean values $\pm$ SD; $N = 11$ experiments (biological replicates).

Supplementary Table 6.1 | Proteins significantly enriched using LEI-945 (1 μ M) in intestinal epithelial cells derived from *Aldh1a1*^{WT} mice corresponding to Fig. 6.2a.

Name	Gene name	Accession	Uniprot ID	Fold-change	Significance ($-\log_{10}$)
Heat shock 70 kDa protein 1A	Hspa1a	Q61696	HS71A_MOUSE	20.0	3.6
Aldehyde dehydrogenase_ mitochondrial	Aldh2	P47738	ALDH2_MOUSE	20.0	2.7
Thioredoxin-related transmembrane protein 1	Tmx1	Q8VBT0	TMX1_MOUSE	20.0	4.6
Voltage-dependent anion-selective channel protein 1	Vdac1	Q60932	VDAC1_MOUSE	20.0	3.3
Voltage-dependent anion-selective channel protein 3	Vdac3	Q60931	VDAC3_MOUSE	20.0	2.3
Keratin_type II cytoskeletal 72	Krt72 PE	Q6IME9	K2C72_MOUSE	20.0	1.5
Retinal dehydrogenase 1	Aldh1a1	P24549	AL1A1_MOUSE	5.4	1.7
Protein disulfide-isomerase	P4hb	P09103	PDIA1_MOUSE	4.6	2.2
Fatty aldehyde dehydrogenase	Aldh3a2	P47740	AL3A2_MOUSE	3.9	2.6
Aldehyde dehydrogenase_ cytosolic 1	Aldh1a7	O35945	AL1A7_MOUSE	3.5	1.5
Aldehyde dehydrogenase X_ mitochondrial	Aldh1b1	Q9CZS1	AL1B1_MOUSE	3.4	2.5
Mitochondrial carrier homolog 2	Mtch2	Q791V5	MTCH2_MOUSE	2.9	2.0
Cytochrome b5 type B	Cyb5b	Q9CQX2	CYB5B_MOUSE	2.4	2.7

Supplementary Table 6.2 | Proteins significantly enriched using LEI-945 (1 μ M) in mesenteric lymph nodes derived from *Aldh1a1*^{WT} mice corresponding to Fig. 6.2b.

Name	Gene name	Accession	Uniprot ID	Fold-change	Significance (-log ₁₀)
Stimulator of interferon genes protein	Tmem173	Q3TBT3	STING_MOUSE	20.0	1.5
Fructose-bisphosphate aldolase A	Aldoa	P05064	ALDOA_MOUSE	20.0	4.1
Tubulin beta-2A chain	Tubb2a	Q7TMM9	TBB2A_MOUSE	20.0	4.3
14-3-3 protein theta	Ywhaq	P68254	I433T_MOUSE	20.0	1.6
40S ribosomal protein S18	Rps18	P62270	RS18_MOUSE	20.0	1.6
Acetyl-CoA acetyltransferase_mitochondrial	Acat1	Q8QZT1	THIL_MOUSE	20.0	1.6
Estradiol 17-beta-dehydrogenase 11	Hsd17b11	Q9EQ06	DHB11_MOUSE	20.0	3.1
Fatty aldehyde dehydrogenase	Aldh3a2	P47740	AL3A2_MOUSE	20.0	3.1
Glutamate dehydrogenase 1_mitochondrial	Glud1	P26443	DHE3_MOUSE	20.0	4.6
Proteasome activator complex subunit 2	Psme2	P97372	PSME2_MOUSE	20.0	3.9
Extended synaptotagmin-2	Esyt2	Q3TZZ7	ESYT2_MOUSE	20.0	4.7
40S ribosomal protein SA	Rpsa	P14206	RSSA_MOUSE	20.0	1.6
Sideroflexin-3	Sfxn3	Q91V61	SFXN3_MOUSE	20.0	3.5
Immunity-related GTPase family M protein 1	Irgm1	Q60766	IRGM1_MOUSE	20.0	3.5
Elongation factor 1-gamma	Eef1g	Q9D8N0	EF1G_MOUSE	20.0	1.5
60S ribosomal protein L5	Rpl5	P47962	RL5_MOUSE	20.0	1.4
Voltage-dependent anion-selective channel protein 2	Vdac2	Q60930	VDAC2_MOUSE	20.0	3.4
Voltage-dependent anion-selective channel protein 3	Vdac3	Q60931	VDAC3_MOUSE	20.0	2.9
Inactive hydroxysteroid dehydrogenase-like protein 1	Hsd11	Q8BTX9	HSDL1_MOUSE	20.0	2.3
H-2 class II histocompatibility antigen gamma chain	Cd74	P04441	HG2A_MOUSE	20.0	2.0
Cytochrome b5 type B	Cyb5b	Q9CQX2	CYB5B_MOUSE	20.0	2.2
Thioredoxin-related transmembrane protein 1	Tmx1	Q8VBT0	TMX1_MOUSE	20.0	2.5
Voltage-dependent anion-selective channel protein 1	Vdac1	Q60932	VDAC1_MOUSE	20.0	3.7
40S ribosomal protein S2	Rps2	P25444	RS2_MOUSE	20.0	5.4
60S ribosomal protein L17	Rpl17	Q9CPR4	RL17_MOUSE	20.0	4.6
40S ribosomal protein S5	Rps5	P97461	RS5_MOUSE	20.0	5.0
Triosephosphate isomerase	Tpi1	P17751	TPIS_MOUSE	20.0	4.2
Aconitate hydratase_mitochondrial	Aco2	Q99KI0	ACON_MOUSE	20.0	1.6
Hemoglobin subunit beta-2	Hbb-b2	P02089	HBB2_MOUSE	20.0	3.6
Delta(14)-sterol reductase	Lbr	Q3U9G9	LBR_MOUSE	19.5	3.4
Aldehyde dehydrogenase_mitochondrial	Aldh2	P47738	ALDH2_MOUSE	18.3	2.0
ADP/ATP translocase 2	Slc25a5	P51881	ADT2_MOUSE	14.4	2.8
Mitochondrial carrier homolog 2	Mtch2	Q791V5	MTCH2_MOUSE	13.7	1.9
Phosphate carrier protein_mitochondrial	Slc25a3	Q8VEM8	MPCP_MOUSE	5.6	2.8
Guanine nucleotide-binding protein G(i) subunit alpha-2	Gnai2	P08752	GNAI2_MOUSE	4.9	3.6
ADP/ATP translocase 1	Slc25a4	P48962	ADT1_MOUSE	4.3	2.7
Tubulin alpha-1C chain	Tuba1c	P68373	TBA1C_MOUSE	3.9	2.4
Ubiquitin-40S ribosomal protein S27a	Rps27a	P62983	RS27A_MOUSE	3.9	3.6
Glyceraldehyde-3-phosphate dehydrogenase	Gapdh	P16858	G3P_MOUSE	3.5	1.6
Tubulin alpha-4A chain	Tuba4a	P68368	TBA4A_MOUSE	3.4	2.1
L-lactate dehydrogenase A chain	Ldha	P06151	LDHA_MOUSE	3.2	1.7
Plastin-2	Lcp1	Q61233	PLSL_MOUSE	3.0	1.8
Retinal dehydrogenase 2	Aldh1a2	Q62148	AL1A2_MOUSE	2.8	1.7
Pyruvate kinase PKM	Pkm	P52480	KPYM_MOUSE	2.7	2.6
Endoplasmic reticulum chaperone BiP	Hspa5	P20029	BIP_MOUSE	2.7	2.4
Elongation factor 1-alpha 1	Eef1a1	P10126	EF1A1_MOUSE	2.7	3.9
Neutral cholesterol ester hydrolase 1	Nceh1	Q8BLF1	NCEH1_MOUSE	2.5	2.0
Actin_cytoplasmic 2	Actg1	P63260	ACTG_MOUSE	2.5	1.4
ATP synthase subunit alpha_mitochondrial	Atp5f1a	Q03265	ATPA_MOUSE	2.2	2.1
Antigen peptide transporter 1	Tap1	P21958	TAP1_MOUSE	2.2	2.0
Tubulin beta-5 chain	Tubb5	P99024	TBB5_MOUSE	2.1	1.7
Profilin-1	Pfn1	P62962	PROF1_MOUSE	2.1	1.7
Adenylyl cyclase-associated protein 1	Cap1	P40124	CAP1_MOUSE	2.1	1.8
Antigen peptide transporter 2	Tap2	P36371	TAP2_MOUSE	2.1	2.9

References

1. Zile, M. H. Vitamin A and Embryonic Development: An Overview. *J. Nutr.* **128**, 455–458 (1998).
2. Kiser, P. D., Golczak, M. & Palczewski, K. Chemistry of the retinoid (visual) cycle. *Chemical Reviews* **114**, 194–232 (2014).
3. Green, H. N. & Mellanby, E. Vitamin a as an anti-infective agent. *Br. Med. J.* **2**, 691–696 (1928).
4. Erkelens, M. N. & Mebius, R. E. Retinoic Acid and Immune Homeostasis: A Balancing Act. *Trends in Immunology* **38**, 168–180 (2017).
5. Iwata, M. *et al.* Retinoic acid imprints gut-homing specificity on T cells. *Immunity* **21**, 527–538 (2004).
6. Zeng, R. *et al.* Retinoic acid regulates the development of a gut-homing precursor for intestinal dendritic cells. *Mucosal Immunol.* **6**, 847–856 (2013).
7. Harrison, E. H. Mechanisms of digestion and absorption of dietary vitamin A. *Annu. Rev. Nutr.* **25**, 87–103 (2005).
8. Goverse, G. *et al.* Diet-Derived Short Chain Fatty Acids Stimulate Intestinal Epithelial Cells To Induce Mucosal Tolerogenic Dendritic Cells. *J. Immunol.* **198**, 2172–2181 (2017).
9. Lampen, A., Meyer, S., Arnhold, T. & Nau, H. Metabolism of vitamin A and its active metabolite all-trans-retinoic acid in small intestinal enterocytes. *J. Pharmacol. Exp. Ther.* **295**, 979–985 (2000).
10. Molenaar, R. *et al.* Expression of Retinaldehyde Dehydrogenase Enzymes in Mucosal Dendritic Cells and Gut-Draining Lymph Node Stromal Cells Is Controlled by Dietary Vitamin A. *J. Immunol.* **186**, 1934–1942 (2011).
11. Ohoka, Y., Yokota-Nakatsuma, A., Maeda, N., Takeuchi, H. & Iwata, M. Retinoic acid and GM-CSF coordinately induce retinal dehydrogenase 2 (RALDH2) expression through cooperation between the RAR/RXR complex and Sp1 in dendritic cells. *PLoS One* **9**, 1–11 (2014).
12. Coombes, J. L. *et al.* A functionally specialized population of mucosal CD103⁺ DCs induces Foxp3⁺ regulatory T cells via a TGF- β - and retinoic acid-dependent mechanism. *J. Exp. Med.* **204**, 1757–1764 (2007).
13. Svensson, M. *et al.* CCL25 mediates the localization of recently activated CD8 $\alpha\beta$ -lymphocytes to the small-intestinal mucosa. *J. Clin. Invest.* **110**, 1113–1121 (2002).
14. Hamann, A., Andrew, D. P., Jablonski-Westrich, D., Holzmann, B. & Butcher, E. C. Role of α 4-integrins in lymphocyte homing to mucosal tissues in vivo. *J. Immunol.* **152**, 3282–93 (1994).
15. Pantazi, E. *et al.* Cutting Edge: Retinoic Acid Signaling in B Cells Is Essential for Oral Immunization and Microflora Composition. *J. Immunol.* **195**, 1368–1371 (2015).
16. Schambach, F., Schupp, M., Lazar, M. A. & Reiner, S. L. Activation of retinoic acid receptor- α favours regulatory T cell induction at the expense of IL-17-secreting T helper cell differentiation. *Eur. J. Immunol.* **37**, 2396–2399 (2007).
17. Benson, M. J., Pino-Lagos, K., Roseblatt, M. & Noelle, R. J. All-trans retinoic acid mediates enhanced T reg cell growth, differentiation, and gut homing in the face of high levels of co-stimulation. *J. Exp. Med.* **204**, 1765–1774 (2007).
18. Mucida, D. *et al.* Reciprocal T H 17 and regulatory T cell differentiation mediated by retinoic acid. *Science* **317**, 256–260 (2007).
19. Sun, C.-M. *et al.* Small intestine lamina propria dendritic cells promote de novo generation of Foxp3 T reg cells via retinoic acid. *J. Exp. Med.* **204**, 1775–1785 (2007).
20. Yasgar, A. *et al.* A High-Content assay enables the automated screening and identification of small molecules with specific ALDH1A1-Inhibitory activity. *PLoS One* **12**, 1–19 (2017).
21. Niederreither, K. & Dollé, P. Retinoic acid in development: Towards an integrated view. *Nature Reviews Genetics* **9**, 541–553 (2008).
22. Fan, X. *et al.* Targeted Disruption of Aldh1a1 (Raldh1) Provides Evidence for a Complex Mechanism of Retinoic Acid Synthesis in the Developing Retina. *Mol. Cell. Biol.* **23**, 4637–4648 (2003).
23. Cravatt, B. F., Wright, A. T. & Kozarich, J. W. Activity-Based Protein Profiling: From Enzyme Chemistry to Proteomic Chemistry. *Annu. Rev. Biochem.* **77**, 383–414 (2008).
24. Liu, Y., Patricelli, M. P. & Cravatt, B. F. Activity-based protein profiling: The serine hydrolases. *Proc. Natl. Acad. Sci.* **96**, 14694–14699 (1999).
25. Koenders, S. T. A. *et al.* Development of a Retinal-Based Probe for the Profiling of Retinaldehyde Dehydrogenases in Cancer Cells. *ACS Cent. Sci.* **5**, 1965–1974 (2019).
26. Alnouti, Y. & Klaassen, C. D. Tissue distribution, ontogeny, and regulation of aldehyde dehydrogenase (Aldh) enzymes mRNA by prototypical microsomal enzyme inducers in mice. *Toxicol. Sci.* **101**, 51–64 (2008).
27. Hsu, L. C., Chang, W.-C., Hoffmann, I. & Duester, G. Molecular analysis of two closely related mouse aldehyde dehydrogenase

- genes: identification of a role for Aldh1, but not Aldh-pb, in the biosynthesis of retinoic acid. *Biochem. J.* **339**, 387–395 (1999).
28. Levi, B. P., Yilmaz, Ö. H., Duester, G. & Morrison, S. J. Aldehyde dehydrogenase 1a1 is dispensable for stem cell function in the mouse hematopoietic and nervous systems. *Blood* **113**, 1670–1680 (2009).
 29. Yoshida, A., Hsu, L. C. & Dave, V. Retinal oxidation activity and biological role of human cytosolic aldehyde dehydrogenase. *Enzyme* **46**, 239–244 (1992).
 30. Moore, S. A. *et al.* Sheep liver cytosolic aldehyde dehydrogenase: The structure reveals the basis for the retinal specificity of class I aldehyde dehydrogenases. *Structure* **6**, 1541–1551 (1998).
 31. Stagos, D. *et al.* Aldehyde dehydrogenase 1B1: Molecular cloning and characterization of a novel mitochondrial acetaldehyde-metabolizing enzyme. *Drug Metab. Dispos.* **38**, 1679–1687 (2010).
 32. Grizotte-Lake, M. *et al.* Commensals Suppress Intestinal Epithelial Cell Retinoic Acid Synthesis to Regulate Interleukin-22 Activity and Prevent Microbial Dysbiosis. *Immunity* **49**, 1103–1115 (2018).
 33. Mielke, L. A. *et al.* Retinoic acid expression associates with enhanced IL-22 production by $\gamma\delta$ T cells and innate lymphoid cells and attenuation of intestinal inflammation. *J. Exp. Med.* **210**, 1117–1124 (2013).
 34. Chang, F., Mahadeva, U. & Deere, H. Pathological and clinical significance of increased intraepithelial lymphocytes (IELs) in small bowel mucosa. *APMIS* **113**, 385–399 (2005).
 35. Furusawa, Y. *et al.* Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* **504**, 446–450 (2013).
 36. Smith, P. M. *et al.* The microbial metabolites, short-chain fatty acids, regulate colonic T reg cell homeostasis. *Science* **341**, 569–573 (2013).
 37. Tan, J. *et al.* Dietary Fiber and Bacterial SCFA Enhance Oral Tolerance and Protect against Food Allergy through Diverse Cellular Pathways. *Cell Rep.* **15**, 2809–2824 (2016).
 38. Arpaia, N. *et al.* Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. *Nature* **504**, 451–455 (2013).
 39. Skelly, A. N., Sato, Y., Kearney, S. & Honda, K. Mining the microbiota for microbial and metabolite-based immunotherapies. *Nature Reviews Immunology* **19**, 305–323 (2019).
 40. Wu, W. *et al.* Microbiota metabolite short-chain fatty acid acetate promotes intestinal IgA response to microbiota which is mediated by GPR43. *Mucosal Immunol.* **10**, 946–956 (2017).
 41. 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).
 42. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
 43. Distler, U., Kuharev, J., Navarro, P. & Tenzer, S. Label-free quantification in ion mobility-enhanced data-independent acquisition proteomics. *Nat. Protoc.* **11**, 795–812 (2016).
 44. Distler, U. *et al.* Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics. *Nat. Methods* **11**, 167–170 (2014).
 45. Kuharev, J., Navarro, P., Distler, U., Jahn, O. & Tenzer, S. In-depth evaluation of software tools for data-independent acquisition based label-free quantification. *Proteomics* **15**, 3140–3151 (2015).

Chapter 7

Synthesis and Biological Evaluation of Clickable Vitamin A

Introduction

The functionalization of endogenous lipids with a ligation handle is commonly used for the visualization of lipid metabolism and its localization within the cell.¹ Examples that can be found in the literature include fatty acids and cholesterol.^{2–7} So far this approach has not been applied to vitamin A, possibly due to a lack of efficient synthetic approaches for the facile introduction of a ligation handle.

The chemical method introduced in **Chapter 3** was used to make retinal-based probe **LEI-945**, suitable for activity-based protein profiling (ABPP).⁸ This method also enables the synthesis of retinoid analogues containing an alkyne click handle starting from the key Wittig salt **1**. Such analogues can be used to study the intracellular localization of vitamin A or the exchange of retinoic acid (RA) between different cells of the immune system, which is thought to be essential for immune homeostasis.^{9,10}

Clickable vitamin A analogues can also be used in affinity-based protein profiling (AfBPP) to identify its interaction partners. Photoaffinity labelling normally requires the introduction of a photocrosslinking group, such as a diazirine¹¹, but retinoids are inherently photoreactive. Under the influence of visible light photoisomerization can occur¹², which is an essential reaction in the visual cycle.¹³ UV light is capable of exciting non-binding electrons within the conjugated retinoid system, which can lead to radical formation.^{14,15} These radicals can react with a protein interaction partner to form a covalent adduct.¹⁶ Previously, radiolabelled retinoids have been used for the UV-induced labelling of retinoid-binding proteins, such as albumin, the cellular retinoic acid-binding protein (CRABP) and the retinoid X receptor β (RXR β).^{17–19}

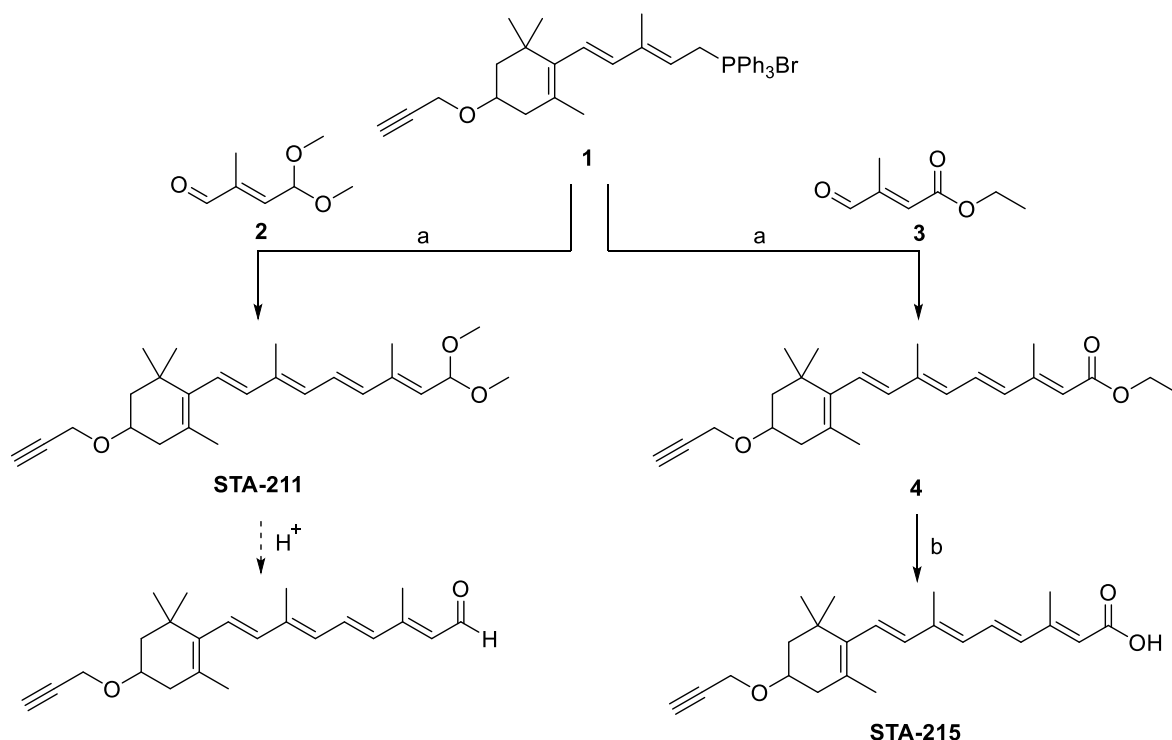
In this chapter, the inherent photocrosslinking ability of clickable retinoids was employed in chemical proteomics approaches to circumvent the use of radioactive compounds.

Results and discussion

Synthesis of clickable retinoid analogues

The clickable retinoid analogues **STA-211** and **STA-215** were synthesized using the convergent synthesis route introduced in **Chapter 3**. Wittig salt **1** was coupled with aldehyde **2** to make the protected retinal analogue **STA-211** in 8% yield. Aldehyde **2** was synthesized following a literature procedure, which can be found in **Supplementary scheme 7.1**.²⁰ The final product was obtained in poor yield due to several purification steps, which were required to separate the product from the unreacted aldehyde **2**.

RA analogue **STA-215** was made in a similar fashion starting from Wittig salt **1** and commercially available aldehyde **3**. Compound **4** was treated with 2 M NaOH in 24% yield. **STA-211** and **STA-215** were stored in ethanol at -80 °C. Exposure of **STA-211** to slightly acidic conditions using a buffer or MilliQ instantly liberates the aldehyde after which the compound can be used for biochemical experiments.



Scheme 7.1 | Synthesis of clickable vitamins. Reagents and conditions: a) *n*-BuLi, THF, -78 °C to RT, 3 h, 8% for **STA-211** and 39% for **4** (3.3:1 E/Z); b) 2 M NaOH, MeOH, 50 °C, 4 h, 24% (7:3 E/Z).

Clickable retinoic acid analogue STA-215 is biologically equivalent

To determine whether the clickable retinoic acid was biologically equivalent, its ability to differentiate human leukaemia HL60 cells was compared. HL60 cells are commonly used as a model for human myeloid cell differentiation.²¹ When treated with DMSO and retinoic acid (RA) these cells differentiate into granulocytes expressing CD11b, which can be used as a biomarker for successful differentiation (**Fig. 7.1a**).²² HL60 cells were cultured in normal RPMI growth medium or medium containing **STA-215** (1 μ M), RA (1 μ M), DMSO (1.25% v/v), DMSO (1.25% v/v) with **STA-215** (1 μ M) and DMSO (1.25% v/v) with RA (1 μ M). Cells were cultured for 72 hours, stained for CD11b and measured using fluorescence activated cell sorting (FACS). Both RA and **STA-215** on their own did not induce differentiation (**Fig. 7.1b**).

Treatment with DMSO (1.25% v/v) on its own induced granulocyte differentiation, but when combined with **STA-215** (1 μ M) or RA (1 μ M) the percentage of differentiated cells increased significantly with 1.5 and 2-fold, respectively. **STA-215** was able to differentiate HL60 cells in combination with DMSO into granulocytes, although to a lesser extent than RA.

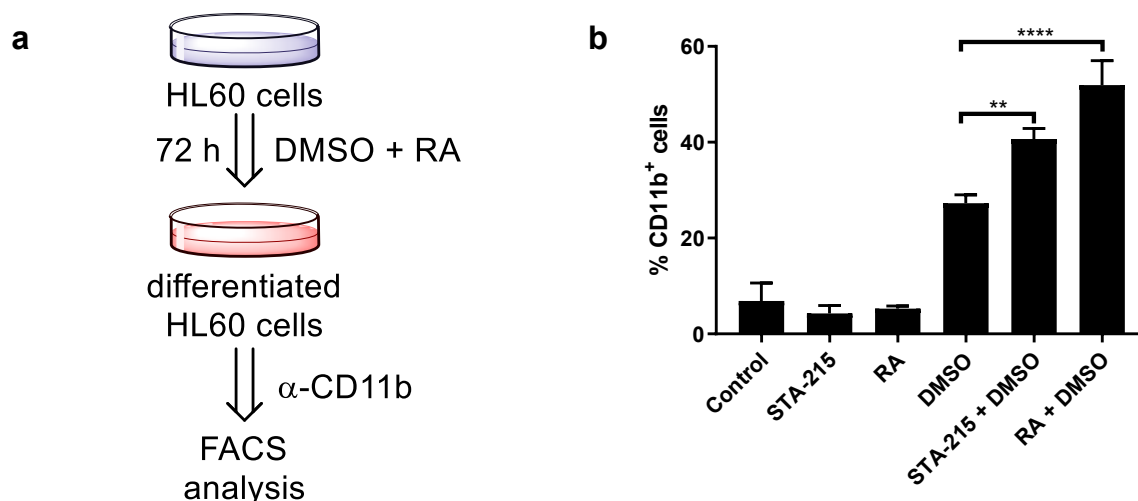


Fig. 7.1 | Validation of biological equivalence of clickable retinoic acid. **a**, Schematic representation of HL60 differentiation experiment. HL60 cells are treated with DMSO and retinoic acid, incubated for 72 h, stained for CD11b and analysed using FACS. **b**, Graph showing the proportion of CD11b⁺ cells indicating the differentiation of HL60 cells into granulocytes.

To further validate whether the clickable retinoid analogues are biologically equivalent, dendritic cells were treated with **STA-211** (10 μ M) or **STA-215** (10 μ M) for 6 days. RA differentiated these dendritic cells, increasing aldehyde dehydrogenase (ALDH) expression in these cells as measured by the ALDEFLUOR assay (**Supplementary Fig. 7.1a**). The differentiation marker CD103 was also increased upon RA treatment (**Supplementary Fig. 7.1b**). Retinoic acid analogue **STA-215** was able to induce the same increase in ALDH activity and CD103 expression with no significant difference compared to the endogenous compound. Also between retinal and **STA-211**, no significant differences could be detected. Taken together, these data show that the clickable RA **STA-215** is capable of inducing differentiation.

Photoaffinity labelling with clickable vitamins

Having validated the biological equivalence of clickable retinoic acid, its application in a photoaffinity proteomics approach was studied. HL60 cells were treated with either **STA-211** (10 μ M, 1 h) or **STA-215** (10 μ M, 1 h) and irradiated for 10 minutes with UV-light (350 nm). The samples were washed, lysed, clicked and resolved by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by in-gel fluorescent scanning.

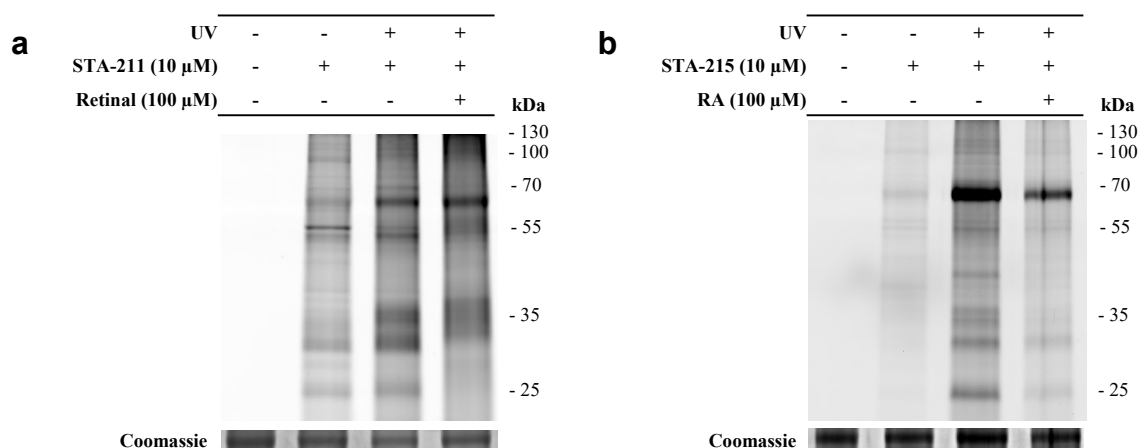


Fig. 7.2 | Validation of biological equivalence of clickable retinoic acid and photoaffinity labelling of clickable retinoids. **a**, Fluorescence labelling of HL60 cells treated with **STA-211** (10 µM, 1 h) and irradiated with UV-light (350 nm, 10 min). **b**, Fluorescence labelling of HL60 cells treated with **STA-215** (10 µM, 1 h) and irradiated with UV-light (350 nm, 10 min).

Cells treated with **STA-211** showed a relatively high amount of background labelling, probably due to interactions of the reactive aldehyde group with enzymes (**Fig. 7.2a**). One clear fluorescent band is visible around 55 kDa in the non-UV induced samples. Upon UV irradiation the intensity of the labelling increases, which could be partially outcompeted by retinal (100 µM).

In HL60 cells treated with **STA-215** (10 µM, 1 h) there is an increase in labelling upon UV irradiation (**Fig. 7.2b**). Around 70 kDa a prominent fluorescent band is visible. The UV-induced labelling by **STA-215** can also be competed with RA (100 µM). These data indicate that the clickable retinoid analogues can be used as photoaffinity probes. To identify the labelled proteins, **STA-215** was subjected to a chemical proteomics experiment. HL60 cells were treated as described above. Samples were ligated with biotin azide and the probe-labelled proteins subsequently enriched using avidin-beads. After several washing steps the proteins were digested on-bead using trypsin and then measured and identified using mass spectrometry. 34 Proteins were enriched by **STA-215** without UV-irradiation, including apolipoprotein A-1 (APOA1) (**Fig. 7.3a**). UV-induced labelling resulted in 192 UV-enriched proteins of which 36 could be competed away by RA (**Fig. 7.3b,c**, **Supplementary Table 7.1**). Interestingly, fatty acid binding protein 5 (FABP5)²³, a known RA binding protein, was not enriched by **STA-215**, while two apolipoproteins (APOA1 and APOM) and serum albumin were significantly enriched by the RA analogue.

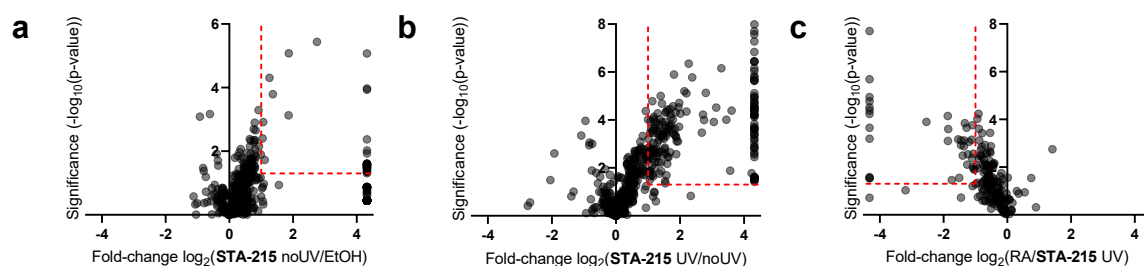


Fig. 7.3 | Photoaffinity labelling of HL60 cells with clickable retinoic acid. **a**, Volcano plot of total proteins identified in a chemical proteomics experiment with probe **STA-215** (10 μ M). Red lines indicate threshold values (fold-change > 2; p-value < 0.05) marking significantly probe-enriched proteins. **b**, Volcano plot of total proteins identified in a chemical proteomics experiment with probe **STA-215** (10 μ M). Red lines indicate threshold values (fold-change > 2; p-value < 0.05) marking significantly UV-enriched proteins. **c**, Volcano plot of UV-enriched proteins in a chemical proteomics experiment with probe **STA-215** (10 μ M). Red lines indicate threshold values (fold-change < 0.5; p-value < 0.05) marking significantly competed proteins by retinoic acid (RA). For parts **a-c**, data are from $N = 4$ experiments (biological replicates).

Analysis of the UV-enriched proteins with DAVID²⁴ analytic tools using the KEGG²⁵ database showed a high representation of enzymes involved in protein processing. Furthermore, enzymes involved in Alzheimer's, Huntington's and Parkinson's disease were found in this group of UV-enriched enzymes (**Fig. 7.4a**). The enriched proteins annotated to this group, include the NADH:ubiquinone oxidoreductase subunits, cytochrome c oxidase subunits, solute carriers and voltage dependent anion channels. Analysis using the PANTHER²⁶ database indicated that the majority of UV-enriched proteins are annotated as catalytic enzymes or binding/transporter proteins. These data show that the clickable retinoic acid analogue **STA-215** can be used for the A_fBPP of retinoid-interacting proteins.

An activity-based chemical proteomics approach with the activity-based probe **LEI-945** in HL60 cells identified 231 significantly enriched proteins (**Supplementary Table 7.2**). Of these proteins, 142 were also UV-enriched by **STA-215**. 24 Proteins of this group were competed by RA in a competitive A_fBPP experiment (**Fig. 7.4c**). Only one ALDH was found in these cells, ALDH3A2. This aldehyde dehydrogenase was also UV-enriched by **STA-215** and competed by RA.

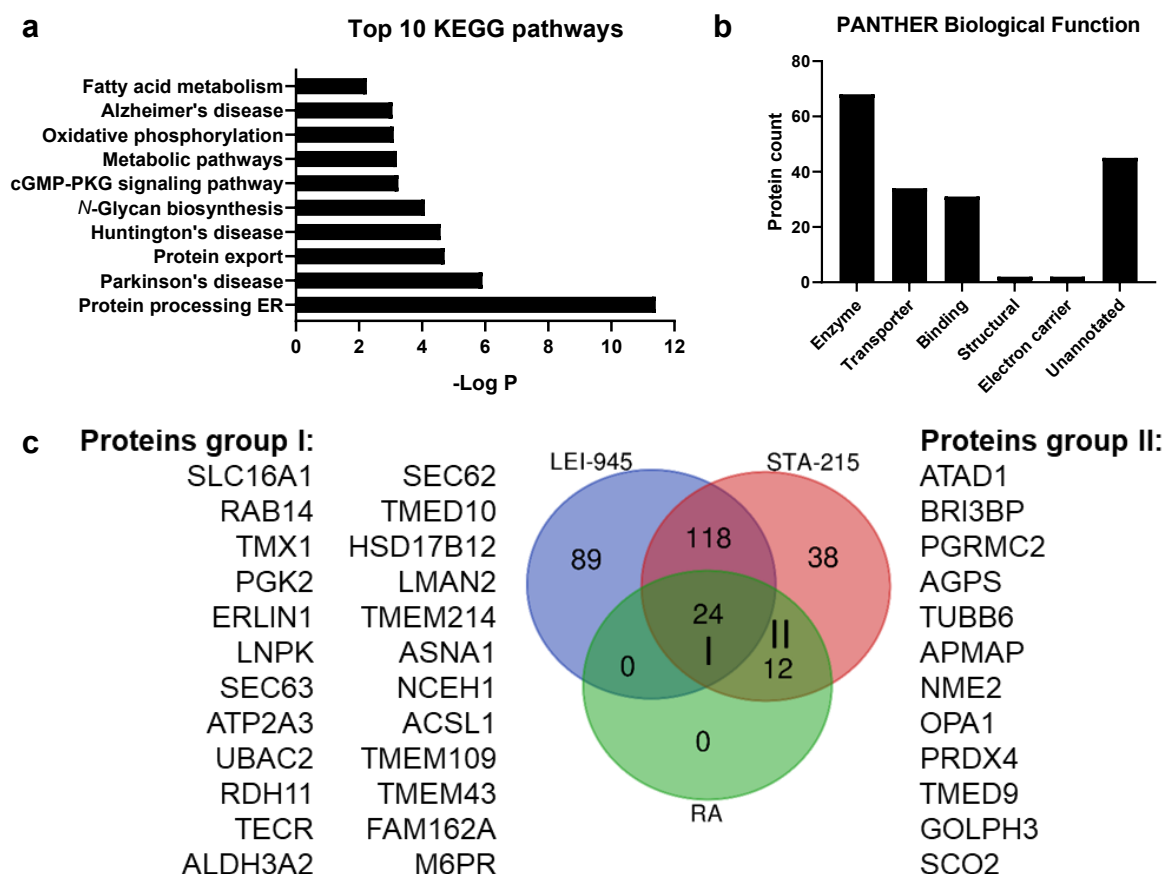


Fig. 7.4 | Proteomics data of UV-enriched proteins by STA-215 in HL60 cells and comparison with LEI-945. **a**, Top 10 pathways enriched in the group of significantly UV-enriched proteins as determined by screening on the KEGG database. **b**, Molecular functions attributed to significantly UV-enriched proteins by the PANTHER database. **c**, Venn diagram showing the overlap in proteins enriched by **LEI-945**, UV-enriched by **STA-215** and UV-enriched by **STA-215** and competed by retinoid acid.

Both **STA-211** and **STA-215** were then used for A_fBPP in the adherent non-small cell lung cancer cell line A549. As this cell line is capable of producing retinoic acid, it was expected that more retinoid-interacting proteins would be expressed in A549 cells. Samples from cells treated with **STA-211** (10 μ M, 30 min), showed a fluorescent band around 55 kDa without UV irradiation. Upon UV-induced labelling a fluorescent band around 40 kDa was visible, which could be competed for by both retinal and retinoic acid. Cells treated with **STA-215** (10 μ M, 30 min) showed a similar labelling pattern to **STA-211**, but with an extra fluorescent band appearing around 30 kDa and less prominent labelling of the proposed ALDH1A1 band around 55 kDa. Overall there was less background labelling in the samples treated with **STA-215** compared to **STA-211**, similar to the samples derived from the HL60 cells.

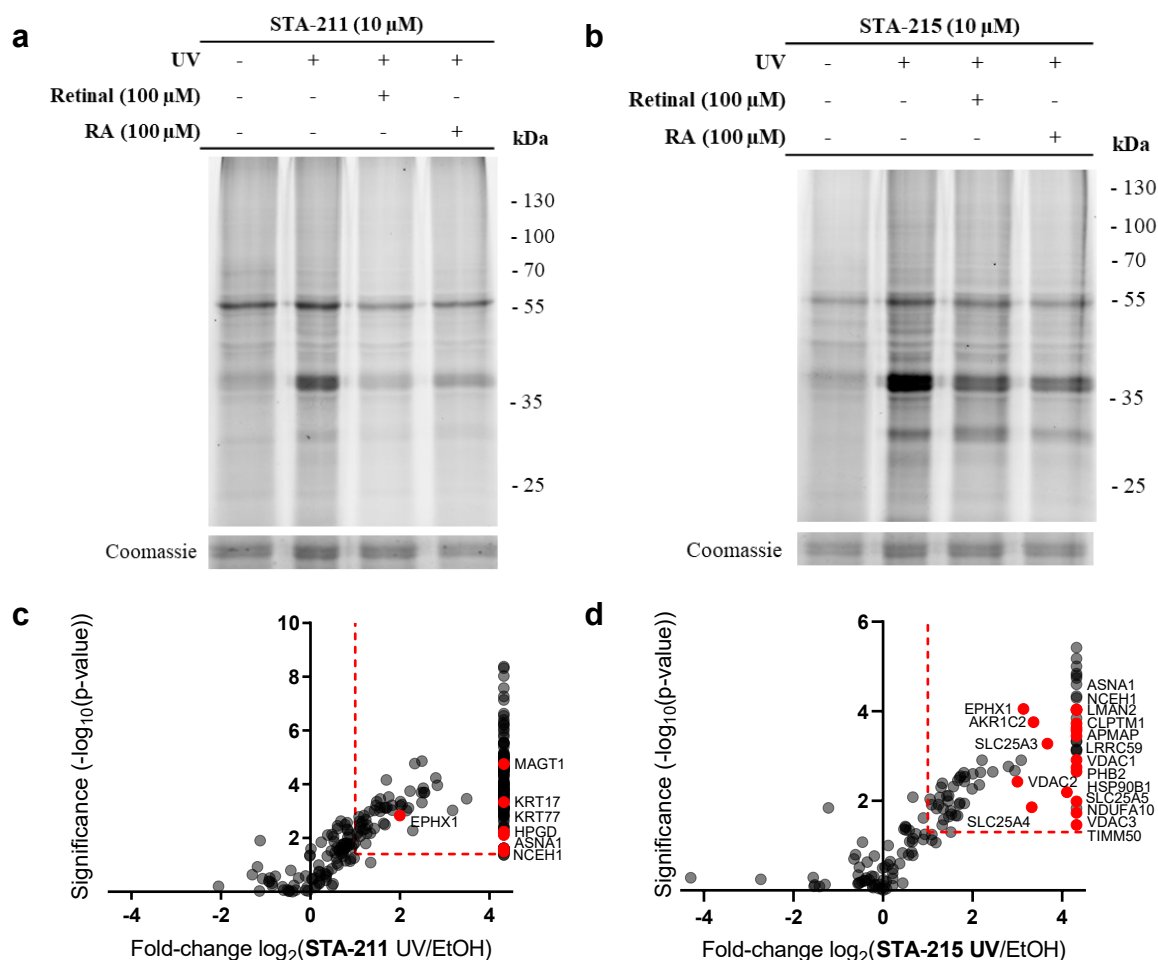


Fig. 7.5 | Photoaffinity labelling of A549 cells with clickable retinoids. **a**, A549 cells treated with **STA-211** (10 μ M, 30 min) and irradiated on ice with UV-light (350 nm, 10 min). **b**, A549 cells treated with **STA-215** (10 μ M, 30 min) and irradiated on ice with UV-light (350 nm 10 min). **c**, Volcano plot of total proteins identified in chemical proteomics experiment with probe **STA-211** (10 μ M). Red lines indicate threshold values (fold-change > 2; p-value < 0.05) marking significantly enriched proteins. Red dots indicate UV-enriched proteins. Data are from $N = 4$ experiments (biological replicates). **d**, Volcano plot of total proteins identified in chemical proteomics experiment with probe **STA-215** (10 μ M). Red lines indicate threshold values (fold-change > 2; p-value < 0.05) marking significantly enriched proteins. Red dots indicate UV-enriched proteins. Data are from $N = 2$ experiments (biological replicates).

The samples were then subjected to a chemical proteomics ArBPP experiment as described previously. In correspondence with the results from the fluorescent labelling, **STA-211** significantly enriched 71 proteins without UV irradiation, whereas **STA-215** only enriched 14 proteins. Upon UV irradiation only 7 proteins were enriched by **STA-211**, however, 5 of these were only identified in the UV-enriched samples (**Fig. 7.5c**).

Of the UV-enriched proteins, 5 could be significantly competed by pretreatment with retinal (100 μ M): epoxide hydrolase 1 (EPHX1), keratin (KRT17), neutral cholesterol ester hydrolase (NCEH1), ATPase (ASNA1) and magnesium transporter protein 1 (MAGT1). Although **STA-211** has high background labelling of itself, probably due to the labelling of proteins via its aldehyde, UV irradiation does expand the coverage of the retinal interacting proteins detected.

STA-215 on the other hand, UV-enriched 21 proteins (**Fig. 7.5d**), 7 of which could be competed by pretreatment with retinoic acid (100 μ M), which include epoxide hydrolase 1 (EPHX1), neutral cholesterol ester hydrolase (NCEH1), ADP/ATP translocase 1 (SLC25A4), vesicular integral-membrane protein VIP36 (LMAN2), aldo-keto reductase family 1 member C2 (AKR1C2), mitochondrial phosphate carrier protein (SLC25A3) and prohibitin-2 (PHB2). An overview of the proteins enriched in these experiments can be found in **Supplementary table 7.3** and **7.4**.

In comparison with the proteins enriched by **LEI-945**, neutral cholesterol ester hydrolase 1 (NCEH1) and epoxide hydrolase 1 (EPHX1) are shared with **STA-211** and **STA-215**. Four are shared with **STA-215** alone, which are the voltage-dependent anion-selective channel protein 2 (VDAC2), voltage-dependent anion-selective channel protein 3 (VDAC3), ADP/ATP translocase 2 (ADT2) and mitochondrial phosphate carrier protein (SLC25A3).

Cellular localization of clickable vitamins using imaging flow cytometry

To determine the cellular localization of the clickable retinoids, imaging flow cytometry was used. This technique combines flow cytometry with fluorescence cell imaging, which provides spatial and morphological information in a quantitative manner.²⁷ A549 cells were treated with vehicle, **STA-211** (10 μ M) or **STA-215** (10 μ M) for 1 hour. Cells were then suspended into a single cell suspension, fixed and clicked with Alexa Fluor 488. After staining of the nucleus, cells were analysed using imaging flow cytometry. The treatment of vehicle-treated cells with the Alexa Fluor 488 “click mix” led to some background labelling, however a clear increase in fluorescence intensity could be seen in cells treated with a clickable vitamin (**Supplementary Fig. 7.2**).

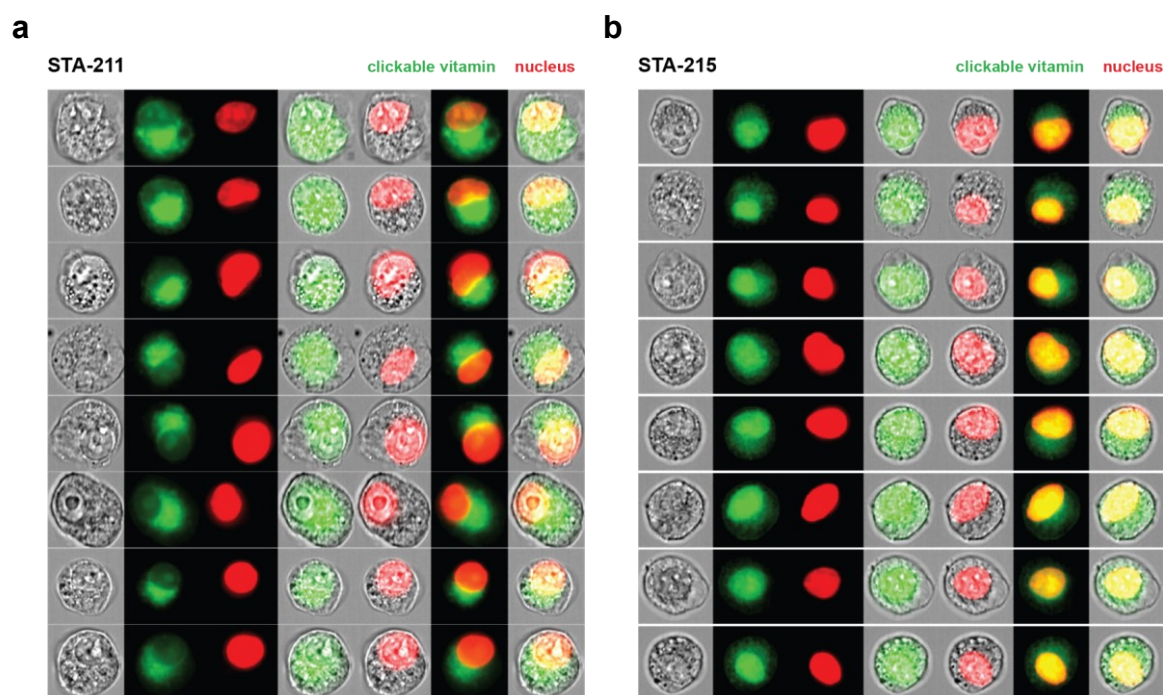


Fig. 7.6 | Localization of clickable vitamins in A549 cells using imaging flow cytometry. **a**, Representative images of individual cells treated with **STA-211** (10 μ M) showing low colocalization. Alexa Fluor 488 fluorescence is shown in green and DAPI in red. Colocalization is visualized in yellow. **b**, Representative images of individual cells treated with **STA-215** (10 μ M) showing high colocalization. Alexa Fluor 488 fluorescence is shown in green and DAPI in red. Colocalization is visualized in yellow.

Preliminary data showed localization of **STA-211** in the cytoplasm (**Fig. 7.6a**), whereas **STA-215** staining localized in the nucleus (**Fig. 7.6b**). As retinoic acid exerts its effect on gene regulation in the nucleus, this difference in localization might be explained by active transport of the clickable retinoic acid to the nucleus. These data also indicate that (imaging) flow cytometry can be used to track the exchange of retinoids between cells, which is of interest in the field of mucosal immunology as discussed in **Chapter 6**.

Conclusion

Using the chemical methods introduced in **Chapter 3**, retinal and retinoic acid analogues functionalized with a ligation handle were synthesized. The clickable retinoic acid analogue was shown to be biologically equivalent to its endogenous counterpart. Both probes were used in a chemical proteomics experiment and showed protein enrichment upon UV irradiation, validating their use as photoaffinity probes.

However, the retinal analogue **STA-211** showed high background labelling, probably due to covalent interactions its aldehyde with proteins. The retinoic acid analogue **STA-215**, on the other hand, showed low background labelling and enriched more proteins upon UV irradiation. Further validation and characterization of the enriched probe targets is required to unravel their involvement in retinoid biology.

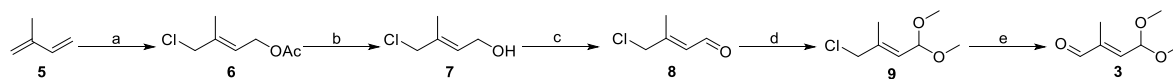
Acknowledgements

Tom van der Wel is kindly acknowledged for his technical advice and collaboration on performing the HL60 differentiation experiment, Bogdan Florea for mass spectrometry analysis, Pasquale Putter for his collaboration on the A_fBPP in A549 cells and Martje Erkelens for performing and analysing the imaging flow cytometry.

Experimental procedures

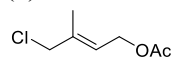
Synthetic methods

General remarks. All reactions were performed using oven or flame-dried glassware and dry solvents. Reagents were purchased from Sigma Aldrich, Acros, Biosolve, VWR, Fluka, Fischer Scientific and Merck and used as received unless stated otherwise. Tetrahydrofuran (THF) and *N,N*-dimethylformamide (DMF) were stored over 4 Å molecular sieves before use. All moisture sensitive reactions were performed under a nitrogen atmosphere. TLC analysis was performed using Merck aluminum sheets (TLC silica gel 60/Kieselguhr F₂₅₄). Compounds were visualized using a solution of KMnO₄ (7.5 g), K₂CO₃ (50 g), 10% NaOH (6 mL) in H₂O (1L). Column chromatography was performed using Screening Device B.V. silica gel (particle size of 40 – 63 µm, pore diameter of 60 Å) with the indicated eluents. ¹H- and ¹³C-NMR spectra were recorded on Bruker AV-400 (400 MHz and 101 MHz, respectively) or Bruker AV-500 MHz (500 MHz and 150 MHz, respectively) using CDCl₃ as solvent. Chemical shifts are reported in ppm (δ) relative to the residual solvent peak or tetramethylsilane. Coupling constants are given in Hz. High-resolution mass spectrometry (HRMS) analysis was performed with a LTQ Orbitrap mass spectrometer (Thermo Finnigan), equipped with an electrospray ion source in positive mode (source voltage 3.5 kV, sheath gas flow 10 mL/min, capillary temperature 250 °C) with resolution R = 60000 at m/z 400 (mass range m/z = 150 – 2000) and dioctyl phthalate (m/z = 391.28428) as a “lock mass”, or with a Synapt G2-Si (Waters), equipped with an electrospray ion source in positive mode (ESI-TOF), injection via NanoEquity system (Waters), with LeuEnk (m/z = 556.2771) as “lock mass”. Eluents used: MeCN:H₂O (1:1 v/v) supplemented with 0.1% formic acid. The high-resolution mass spectrometers were calibrated prior to measurements with a calibration mixture (Thermo Finnigan). The retinoids were handled under dark conditions using amber colored flasks or aluminum foil when containing more than 3 conjugated double bonds. Retinoid intermediates were stored in the dark under nitrogen at -30 °C and final compounds as powder or as EtOH stock at -80 °C under nitrogen atmosphere. For further information on retinoid handling and storage we refer to the review from Barua and Furr²⁸.



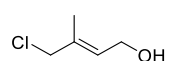
Supplementary Scheme 7.1 | Synthesis of protected aldehyde intermediate. Reagents and conditions: a) *t*BuOCl, AcOH, 1 h and then CuSO₄, H₂SO₄, 96 h, 30%; b) Na₂CO₃, MeOH/H₂O, 5 h, 0 °C, 33%; c) PCC, DCM, 90 min, 0 °C, 51%; d) trimethyl orthoformate, *p*TsOH, MeOH, 24 h, 67%; e) K₂HPO₄/KH₂PO₄, NaBr, DMSO, 24 h, 80 °C, 55%.

(E)-4-Chloro-3-methylbut-2-en-1-yl acetate (6):

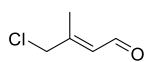


To a stirred solution of isoprene **5** (20 mL, 0.20 mol) in acetic acid (60 mL) at 0 °C under N₂ was added *tert*-butyl hypochlorite (18 mL, 0.16 mol). The reaction mixture was stirred for 1 hour and then quenched with H₂O and extracted with Et₂O. The organic layer was washed with sat. aq. NaHCO₃ and brine, dried with MgSO₄, filtered and concentrated under reduced pressure. The residue was then dissolved in acetic acid (60 mL) and to this was added CuSO₄ (320 mg, 2.0 mmol) and H₂SO₄ (0.2 mL, 4 mmol). The reaction mixture was stirred for 96 hours at room temperature and was then quenched with H₂O and extracted with Et₂O. The organic layer was washed with sat. aq. NaHCO₃ and brine, dried with MgSO₄, filtered and concentrated under reduced pressure. Purification of the residue by column chromatography (pentane/Et₂O) afforded the title compound **6** (7.8 g, 48 mmol, 30%). ¹H NMR (400 MHz, CDCl₃) δ 5.70 (t, *J* = 6.8 Hz, 1H), 4.62 (d, *J* = 6.9 Hz, 2H), 4.02 (s, 2H), 2.06 (s, 3H), 1.82 (s, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 170.4, 136.6, 123.5, 60.4, 50.5, 20.5, 14.2. Spectroscopic data for H-NMR are in agreement with those reported.²⁹

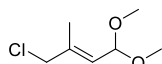
(E)-4-Chloro-3-methylbut-2-en-1-ol (7):



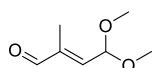
Compound **6** (7.8 g, 48 mmol) was dissolved in MeOH (120 mL). To this was added a solution of Na₂CO₃ (7.6 g, 72 mmol) in H₂O (40 mL). The reaction mixture was stirred for 5 hours at 0 °C, filtered and concentrated under reduced pressure. The remaining water layer was then extracted with CHCl₃. The organic layer was washed with brine, dried with MgSO₄, filtered and concentrated under reduced pressure. Purification of the residue by column chromatography (pentane/Et₂O) afforded the title compound **7** (1.9 g, 16 mmol, 33%). *R*_f (50% Et₂O in pentane) = 0.5. ¹H NMR (400 MHz, CDCl₃) δ 5.70 (s, 1H), 4.15 (s, 2H), 4.02 (s, 2H), 3.43 (s, 1H), 1.76 (s, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 128.8, 58.5, 51.2, 42.7, 14.1.

(E)-4-Chloro-3-methylbut-2-enal (8):

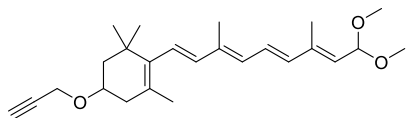
A solution of **7** (1.9 g, 16 mmol) in dry DCM (20 mL) was added to a stirred suspension of PCC (5.1 g, 24 mmol) in dry DCM (40 mL) at 0 °C under N₂ containing molecular sieves. The reaction mixture was stirred for 90 minutes at 0 °C and was then filtered over celite and concentrated under reduced pressure. Purification of the residue by column chromatography (Et₂O/pentane) afforded the title compound **8** (0.94 g, 8.0 mmol, 51%). *R*_f (50% Et₂O in pentane) = 0.8. ¹H NMR (400 MHz, CDCl₃) δ 10.04 (d, *J* = 7.7 Hz, 1H), 6.12 (d, *J* = 7.7 Hz, 1H), 4.12 (s, 2H), 2.28 (s, 3H); ¹³C NMR (101 MHz, CDCl₃) δ 191.0, 155.2, 128.6, 49.3, 15.5.

(E)-4-Chloro-1,1-dimethoxy-3-methylbut-2-ene (9):

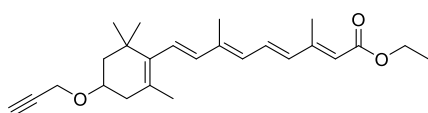
To a solution of **8** (0.94 g, 7.9 mmol) in MeOH (7 mL) was added trimethyl orthoformate (7.1 mL, 40 mmol) and *p*TsOH (0.80 mmol, 84 mg). The reaction mixture was stirred overnight at room temperature and was then quenched with sat. aq. NaHCO₃ and extracted with Et₂O. The organic layer was washed with brine, dried with MgSO₄, filtered and concentrated under reduced pressure. Purification of the residue by column chromatography (Et₂O/pentane) afforded the title compound **9** (0.88 g, 5.3 mmol, 67%). *R*_f (10% Et₂O in pentane) = 0.5. ¹H NMR (400 MHz, CDCl₃) δ 5.51 (d, 6.0 Hz, 1H), 5.04 (d, *J* = 6.2 Hz, 1H), 4.01 (s, 2H), 3.32 (s, 6H), 1.85 (s, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 137.4, 126.6, 99.6, 52.2, 50.6, 14.9.

(E)-4,4-Dimethoxy-2-methylbut-2-enal (2):

To a stirred solution of **9** (0.78 g, 4.7 mmol) in DMSO (25 mL) was added K₂HPO₄ (1.1 g, 6.2 mmol), KH₂PO₄ (0.23 g, 1.7 mmol) and NaBr (73 mg, 0.71 mmol). The mixture was stirred for 18 hours at 80 °C. Then H₂O was added and the mixture extracted with CHCl₃. The organic layer was washed with brine, dried with MgSO₄, filtered and concentrated under reduced pressure. Purification of the residue by column chromatography (Et₂O/pentane) afforded the title compound **2** (0.39 g, 2.7 mmol, 57%). *R*_f (10% Et₂O in pentane) = 0.3. ¹H NMR (400 MHz, CDCl₃) δ 9.49 (s, 1H), 6.39 (d, *J* = 5.9 Hz, 1H), 5.27 (d, *J* = 5.9 Hz, 1H), 3.38 (s, 6H), 1.85 (s, 3H); ¹³C NMR (101 MHz, CDCl₃) δ 194.4, 146.7, 141.5, 99.1, 52.5, 9.5. Spectroscopic data are in agreement with those reported.²⁰

2-((1E,3E,5E,7E)-9,9-Dimethoxy-3,7-dimethylnona-1,3,5,7-tetraen-1-yl)-1,3,3-trimethyl-5-(prop-2-yn-1-yloxy)cyclohex-1-ene (STA-211):

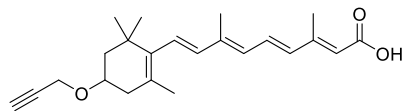
To a stirred solution of **1** (420 mg, 0.70 mmol) in dry THF (5 mL) at -78 °C under N₂ was added *n*-BuLi (0.3 mL, 2.5 M in hexane, 0.76 mmol). The mixture was stirred for 30 minutes at -78 °C and then a solution of **2** (100 mg, 0.70 mmol) in dry THF (5 mL) was added. The reaction mixture was stirred for 1 hour at -78 °C and then for 3 hours at room temperature. The reaction was quenched with sat. aq. NH₄Cl and extracted with Et₂O. The combined organic layers were washed with H₂O and brine, dried with MgSO₄, filtered and concentrated under reduced pressure. Purification of the residue by column chromatography (Et₂O/pentane) with neutral silica afforded the title compound **STA-211** (21 mg, 55 μmol, 7.9%). ¹H NMR (500 MHz, CDCl₃) δ 6.70 – 6.07 (m, 5H), 5.56 – 5.13 (m, 2H), 4.22 (d, *J* = 2.4 Hz, 2H), 3.86 (dt, *J* = 12.0, 5.7, 3.3 Hz, 1H), 3.33 (d, *J* = 10.1 Hz, 6H), 2.46 – 2.38 (m, 2H), 2.07 (dd, *J* = 16.8, 9.6 Hz, 1H), 1.96 (s, 3H), 1.91 (s, 3H), 1.86 – 1.82 (m, 1H), 1.73 (s, 3H), 1.44 (t, *J* = 12.0 Hz, 1H), 1.07 (s, 6H). ¹³C NMR (126 MHz, CDCl₃) δ 154.6, 138.9, 138.3, 137.7, 136.3, 136.1, 130.4, 127.8, 126.1, 125.9, 125.7, 100.2, 80.4, 73.8, 71.6, 55.1, 52.2, 44.5, 39.2, 36.8, 30.2, 28.6, 21.6, 13.2, 12.7. HRMS (ESI) *m/z*: [M + H]⁺ calculated for deprotected **STA-211** C₂₃H₃₀O₂: 339.23186, found: 339.23185.

Ethyl (2E,4E/Z,6E,8E)-3,7-dimethyl-9-(2,6,6-trimethyl-4-(prop-2-yn-1-yloxy)cyclohex-1-en-1-yl)nona-2,4,6,8-tetraenoate (4):

To a stirred solution of **1** (0.50 g, 0.83 mmol) in dry THF (4 mL) at -78 °C under N₂ was added *n*-BuLi (0.50 mL, 1.6 M in hexane, 0.83 mmol). The mixture was stirred for 30 minutes at -78 °C and then ethyl 3-methyl-4-oxocrotonate **3** (0.10 mL, 0.76 mmol) was added. The reaction was stirred for 1 hour at -78 °C and then for 2 hours at room temperature. The reaction was quenched with sat. aq. NH₄Cl and extracted with Et₂O. The combined organic layers were washed with H₂O and brine, dried with MgSO₄ and concentrated under reduced pressure. Purification of the residue by column chromatography (EtOAc/pentane) afforded the title compound **4** (0.11 g, 0.30 mmol, 39%) as a yellow oil as an E/Z mixture (1:0.3). *R*_f (10% EtOAc in pentane) = 0.95. NMR spectra are obtained from the mixture of stereoisomers. ¹H NMR (400 MHz, CDCl₃) δ 6.98 (dd, *J* = 15.1, 11.4 Hz, 1H), 6.55 – 6.09 (m, 4H), 5.93 – 5.76 (m, 1H), 4.25 – 4.21 (m, 2H), 4.21 – 4.13 (m, 2H), 3.92 – 3.81 (m, 1H), 2.48 – 2.38 (m, 2H), 2.37 – 2.32 (m, 3H), 2.12 – 2.03 (m, 1H), 2.01 – 1.94 (m, 3H), 1.88 – 1.82 (m, 1H), 1.73 (s, 3H), 1.45 (t, *J* = 12.0 Hz, 1H), 1.29 (t, *J* = 7.1 Hz, 3H), 1.11 – 1.05 (m, 6H). ¹³C NMR (101 MHz, CDCl₃) δ 167.13, 152.54, 139.10, 137.98,

137.59, 135.51, 130.70, 129.95, 127.48, 126.60, 118.79, 80.30, 73.88, 71.48, 59.64, 55.13, 44.39, 39.23, 36.80, 30.16, 28.58, 21.62, 14.32, 13.78, 12.87. HRMS (ESI) m/z : $[M + H]^+$ calculated for $C_{25}H_{34}O_3$: 383.25807, found 383.25801.

(2E,4E/Z,6E,8E)-3,7-Dimethyl-9-(2,6,6-trimethyl-4-(prop-2-yn-1-yloxy)cyclohex-1-en-1-yl)nona-2,4,6,8-tetraenoic acid (STA-215):



To a stirred solution of **4** (64 mg, 0.167 mmol) in MeOH (2.5 mL) was added 2 M NaOH (aq) (2.5 mL). The reaction was heated at 50 °C for 3 hours and then another 5 mmol of NaOH (200 mg) was added. The reaction was then stirred for 1 hour at 50 °C and then made acidic with 1 M HCl. The product was then extracted with EtOAc, washed with

brine, dried over $MgSO_4$, filtered and concentrated. Purification by column chromatography (Et_2O /pentane) afforded the title compound **STA-215** (14 mg, 39 μ mol, 24%; 7:3 4E/Z). The E/Z mixture was inseparable and prone to further isomerization under the influence of light. As retinoids are rapidly converted *in vivo* into their biological equilibrium of stereoisomers. **STA-215** was used as the reported E/Z mixture.^{30,31} NMR spectra are obtained from the mixture of stereoisomers. 1H NMR (500 MHz, $CDCl_3$) δ 7.04 (dd, J = 15.0, 11.4 Hz, 1H), 6.62 – 6.50 (m, 1H), 6.35 – 6.11 (m, 3H), 5.94 – 5.79 (m, 1H), 4.22 (s, 2H), 3.91 – 3.83 (m, 1H), 2.47 – 2.40 (m, 2H), 2.39 – 2.33 (m, 3H), 2.13 – 2.04 (m, 1H), 2.02 – 1.96 (m, 3H), 1.88 – 1.83 (m, 1H), 1.74 (s, 3H), 1.48 – 1.42 (m, 1H), 1.08 (s, 6H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 172.1, 155.1, 139.8, 137.9, 137.6, 135.2, 131.7, 129.9, 127.9, 126.8, 117.8, 80.3, 73.9, 71.5, 55.2, 44.4, 39.2, 36.8, 30.2, 28.6, 21.6, 14.0, 12.9. LC-MS (ESI) m/z : $[M + H]^+$ calculated for $C_{23}H_{30}O_3$: 355.22677, found: 355.16667.

Biological methods

Cell culture. HL60 cells were grown in RPMI with stable glutamine and phenol red with 10% Fetal Calf serum, penicillin and streptomycin at 37 °C and 5% CO_2 . Cells were cultured between 0.2×10^6 and 2.0×10^6 cells/mL confluence. A549 cells were grown in DMEM with stable glutamine and phenol red with 10% New Born Calf serum, penicillin and streptomycin at 37 °C and 5% CO_2 . Medium was refreshed every 2-3 days and cells were passaged twice a week. Cell lines were purchased from ATCC and were regularly tested for mycoplasma contamination. Cultures were discarded after 2-3 months of use.

HL60 differentiation and FACS protocol. HL60 cells were cultured as described above. 2.4×10^6 cells per condition were collected in a 15 mL tube, the growth medium was removed and replaced with 6 mL medium containing DMSO (1.25% v/v), **STA-215** (1 μ M), RA (1 μ M) or a combination hereof. Cells were divided over 3 wells in a 6-wells plate and cultured for 72 hours. Cells were counted using a TC20™ automated cell counter and a maximum of 1.0×10^6 cells was transferred to an Eppendorf tube and centrifuged (500 g, 3 min). Cells were suspended in FACS buffer (100 μ L; 1% FCS, 1% BSA, 0.1% NaN_3 and 2 mM EDTA in PBS) containing human FcR blocking reagent (4 μ L; Miltenyi Biotec), transferred to a V-bottom 96-wells plate and split providing the isotype control samples. Cells were incubated for 10 minutes in the dark at 4 °C. Then staining mixture (50 μ L; FACS buffer (4.65 μ L), antibody (1 μ L) and 7-AAD (2 μ g/mL)) was added containing either the CD11b-APC antibody (Miltenyi Biotec) or rat IgG2b-APC (Miltenyi Biotec). Samples were mixed and incubated for 30 minutes in the dark at 4 °C. PBS (200 μ L) was added and the plate was centrifuged (500 g, 3 min). Cells were suspended in PBS (100 μ L) and fixated by adding 1% PFA in PBS (100 μ L). After incubating for 15 minutes in the dark at 4 °C PBS was added (200 μ L) and the plate was centrifuged. This was repeated once and cells were suspended in FACS buffer (200 μ L). Samples were diluted 10x in a new plate to ensure a cell density of <1000 cells/ μ L. Samples were measured on a Guava® easyCyte™ system. Samples were gated on living single cells and the isotype control was used to gate for background signal for differentiation. Using these settings the percentage of differentiated cells was determined.

Human dendritic cell culture. Peripheral blood mononuclear cells (PBMCs) were isolated by gradient centrifugation of peripheral blood obtained from healthy donors on lymphoprep ($d=1.077$, Fresenius Kabi, Bad Homburg, Germany). Blood was drawn after receiving signed informed consent according to the guidelines of the Medical Ethical Committee of the VU University Medical Center (Amsterdam, the Netherlands) in accordance with the Declaration of Helsinki. $CD14^+$ monocytes were isolated from the PBMC fraction with positive selection using $CD14$ magnetic microbeads (Miltenyi Biotec, Leiden, the Netherlands). $CD14^+$ cells were cultured in RPMI supplemented with 10% FCS, 1% PSG, 20 ng/ml GM-CSF and 20 ng/ml IL-4 (both Immunotools, Friesoythe, Germany) for 6 days in order to induce the development of dendritic cells in the presence of 10 μ M retinoic acid, retinal, (both Sigma Aldrich, Zwijndrecht, the Netherlands) **STA-211**, **STA-215** or a vehicle control.

At day 6, cells were harvested and an ALDEFUOR assay (Stemcell, Cologne, Germany) was performed according to manufacturer's protocol. Cells were subsequently stained with CD1c BV421 (1:20, clone L161, Biolegend, London, United Kingdom), CD14 BV510 (1:25, clone MφP9, BD Bioscience, San Jose, CA, USA), CD86 PE (1:100, clone 2331, BD Bioscience, San Jose, CA, USA) and CD103 APC (1:20, clone Ber-ACT8, BD Bioscience, San Jose, CA, USA). Before acquisition 7-AAD was added as a live death marker. Samples were acquired on LSR-Fortessa X20 (BD Bioscience, Mountain View, CA, USA) and data was analyzed with Flowjo software (Tree Star, San Carlos, CA, USA).

***In situ* affinity-based protein profiling.** STA-211 was deprotected by dilution in MilliQ or acidic buffer. Growth medium from cells grown in 6-wells plate was removed and 1 mL serum free medium containing probe STA-211 (10 μ M, 0.1% EtOH) or STA-215 (10 μ M, 0.1% EtOH) was added. The cells were incubated for 1 hour. For competitive A₁BPP cells were first incubated with vehicle or competitor (100 μ M, 0.1% EtOH) for 30 minutes followed by STA-211 (10 μ M final concentration) or STA-215 (10 μ M final concentration) for 1 hour. The medium was then removed, the cells were washed with 2 mL PBS and then irradiated (350 nm, 10 min) in 1 mL ice-cold PBS. The cells were then harvested using a cell scraper and transferred to an Eppendorf tube and the suspension was centrifuged for 5 minutes at 1200 rpm. The PBS was removed and the samples snap frozen and stored at -80 °C until further use.

Copper-catalysed azide-alkyne cycloaddition reaction and in-gel fluorescence analysis. Cell pellets were thawed on ice, lysed by addition of ice-cold lysis buffer (MilliQ, 1x protease inhibitor cocktail (Roche cOmplete EDTA free)) and incubated on ice (15-30 min). The protein concentration was determined by a Quick Start™ Bradford Protein assay (Bio-Rad). The protein fractions were diluted to a total protein concentration of 1 mg/mL. From each sample 40 μ L was taken and treated with 5 μ L from a freshly prepared "click" mixture containing 9 mM CuSO₄ (2.5 μ L/sample, 18 mM in H₂O), 45 mM NaAsc (1.5 μ L/sample, 150 mM in H₂O), 1.8 mM THPTA (0.5 μ L/sample, 18 mM in DMSO) and 9 μ M Alexa Fluor 647 azide (0.5 μ L/sample, 90 μ M in DMSO from Thermo Fischer Scientific). The samples were incubated for 1 hour at 37 °C and then 15 μ L 4x SDS page sample buffer was added. The samples were denatured at 100 °C for 5 minutes. 8 μ g per sample was resolved on a SDS-PAGE gel (10% acrylamide, 180V, 75 min). Gels were visualized with a ChemiDoc XRS (Bio-Rad) using Cy3 and Cy5 multichannel settings (605/50 and 695/55, filters respectively) and stained with Coomassie. Fluorescence was normalized to Coomassie staining.

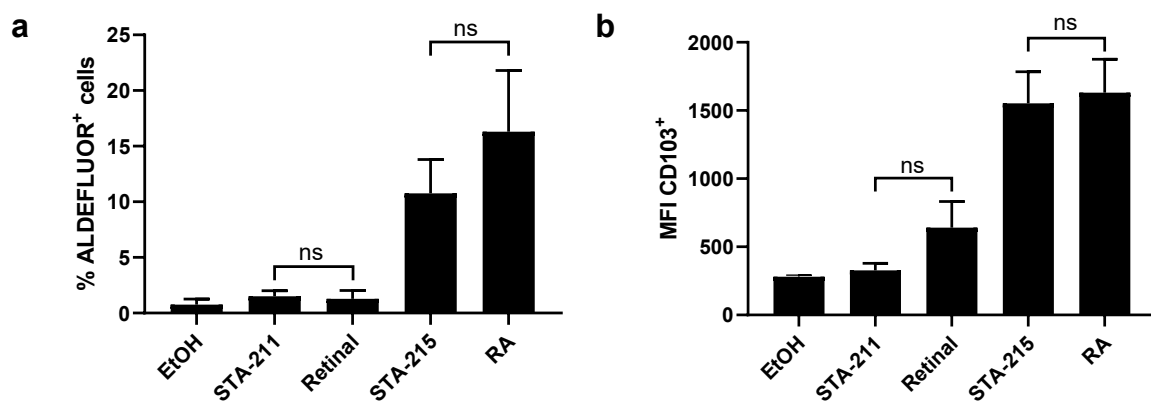
Imaging flow cytometry. A549 cells were incubated with STA-211 (10 μ M), STA-215 (10 μ M) or a vehicle control (ethanol) for 1 hour under serum-free conditions. Cells were then fixated using 1% PFA. Click chemistry mixture (click mix) was prepared (1 mM CuSO₄, 5 mM NaAsc, 0.4 mM THPTA (all Sigma Aldrich, Zwijndrecht, the Netherlands), 40 μ M biotin-alkyne, Alexa Fluor 488-azide (all Click Chemistry Tools, Scottsdale, AZ, United States) and PBS was added. Samples were incubated for 1 hour at 37 °C. DAPI was added as a nuclear stain just before acquisition. Samples were analyzed using the ImageStream. The ImageStream is an instrument that combines flow cytometry and microscopy in one platform and is thereby able to generate population statistics on subcellular morphological measurements made per cell across a sample of thousands. The system has a magnification lens in the light path after the flow cell and in essence replaces the PMT's of a conventional flow cytometer with a camera. It is thus able to image thousands of cells per second with microscopic resolution. Sample were acquired using the INSPIRE acquisition software on a two camera, 4 laser (125mW 405nm, 100mW 488 nm, 200 mW 561 nm, 150mW 642 nm) ImageStream ISX from Amnis (a Luminex subsidiary). With the 488nm laser and 405nm laser set at 100 mW and 125 mW respectively, the Alexa Fluor 488 signal was measured in Channel 2 of Camera 1 and the DAPI signal was measured in Channel 7 of Camera 2. The magnification was set at 60x with a sample core velocity of 40mm/min. The acquired images and data were analyzed using the image analysis software IDEAS. IDEAS is flexible and powerful enough to identify and measure any discernible object in an image. To this end, a threshold mask of 70% of the default DAPI mask was created for the nucleus and the extent of nuclear translocation for the clickable vitamins was quantified using this nuclear mask as the region of interest for both the Bright Detail Similarity and Internalization features in the IDEAS software.

In situ activity-based proteomics

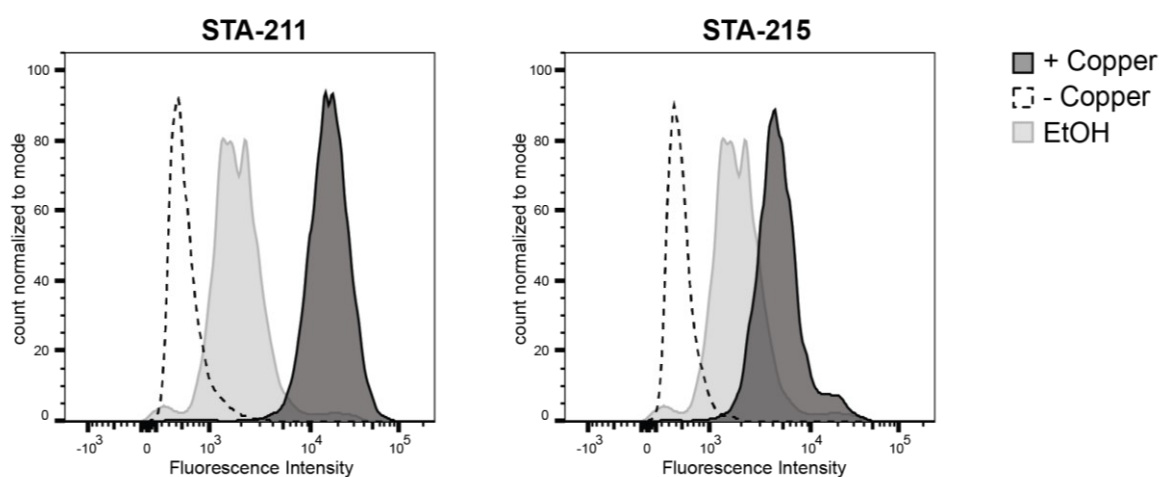
Sample preparation. Protocol adapted from previously described procedure.³² Cells were treated *in situ*, irradiated, harvested, lysed and adjusted to 1 mg/mL protein concentration as described above. 250 μ L was taken from each sample and to this 25 μ L freshly prepared “click” mixture containing 1 mM CuSO₄ (2.5 μ L/sample, 100 mM in H₂O), 5 mM NaAsc (1.25 μ L/sample, 1M in H₂O), 0.4 mM THPTA (1 μ L/sample, 100 mM in DMSO), 40 μ M biotin-alkyne (2.5 μ L/sample, 4 mM in DMSO) and MilliQ (17.75 μ L/sample) was added. Samples were incubated for 1 hour at 37 °C while shaking (300 rpm). Excess click reagents were then removed by chloroform/methanol precipitation followed by another wash with methanol. Precipitated proteomes were then suspended in urea buffer (250 μ L, 6 M urea and 25 mM ammonium bicarbonate), DTT (2.5 μ L, 1M) was added and the mixture was then incubated for 15 min at 65 °C while shaking (600 rpm). The samples were then allowed to cool down to RT and then alkylated by addition of iodoacetamide (20 μ L, 0.5M) for 30 minutes at RT in the dark. Addition of SDS (70 μ L, 10% (v/v)) was followed by heating at 65 °C for 5 minutes. For each sample 50 μ L 50% slurry of avidin-agarose beads from egg white (Sigma-Aldrich) was washed three times with PBS and transferred in PBS (1 mL) to a 15 mL tube. To this another 2 mL of PBS was added followed by the corresponding proteome sample. The beads were incubated with the proteome for 2 hours at room temperature using an overhead shaker. The beads were then isolated by centrifugation (2 min, 2500 g), washed with SDS in PBS (0.5% (w/v)) and washed three times with PBS. The beads were then transferred to low-binding Eppendorf tubes and proteins were digested overnight at 37 °C and 950 rpm shaking in 250 μ L digestion buffer (100 mM Tris, 100 mM NaCl, 1 mM CaCl₂, 2% acetonitrile and 0.5 μ g sequencing grade trypsin (Promega)). Digestion was stopped by addition of formic acid (12.5 μ L) and the beads were filtered off by centrifugation (2 min, 600 g) using a Bio-Spin column (Bio-Rad). Samples were then desalted using stage tips, collected in low-binding Eppendorf tubes, concentrated using a SpeedVac (Eppendorf) and stored at -20 °C until reconstitution before measurement.³³ All samples were prepared in at least three biological replicates.

LC-MS/MS measurement and analysis. Samples were reconstituted in LC-MS sample solution (50 μ L, MilliQ, 3% acetonitrile/0.1% formic acid/20 fmol/ μ L enolase). Samples were then analysed using a NanoACQUITY UPLC System (Waters) coupled to a SYNAPT G2-Si high-definition mass spectrometer (Waters) as previously described.^{32,34} Of each sample 5 μ L was loaded on a nanoEASE™ M/Z Symmetry C18 trap column (particles 5 μ m, 100 Å, 180 μ m x 20 mm, Waters) with 0.1% formic acid and separated on an nanoEASE™ M/Z HSS C18 T3 analytical column (particles 1.8 μ m, 75 μ m x 250 mm, Waters) heated at 80 °C. A multistep gradient running from 5-40% acetonitrile containing 0.1% formic acid during a 70 minute method at 300 nL/min was used to achieve peptide separation. Survey scans (m/z 50-2000 Da) were acquired in the Synapt with a scan time of 0.6 seconds in positive, resolution mode. The collision energy is set to 4 V in the trap cell for low-energy MS mode. For the elevated energy scan, the transfer cell collision energy is ramped using drift-time specific collision energies. The lock mass is sampled every 30 seconds. MS raw files were analysed with ProteinLynx Global SERVER (PLGS, v3.0.3, Waters). The MS^E identification was also performed with PLGS using the human proteome from Uniprot (Uniprot-homo-sapiens-trypsin-reviewed-2016-08-29.fasta). The following parameter settings were used: low energy threshold 150 counts, elevated energy threshold 30, peptide and protein FDR 1%, enzyme specificity trypsin, max missed cleavages max 2, variable modification methionine oxidation, fixed modification carbamidomethylation cysteine, at least fragments/peptide 2, fragments/protein 5, peptides/protein 1 and number of peptides to measure per protein 3. For label-free quantification ISOQuant (v1.5) was used.^{35,36} Data were filtered to retain only proteins with two or more reported unique peptides and quantified in at least 3 replicates of the positive control (probe-treated). Proteins were designated as significantly enriched by the probe when they showed 2-fold enrichment in quantification value when comparing negative control (vehicle-treated) with positive control (probe-treated) samples and probability as determined by a Student's *t* test (<0.05).

Supplementary Data



Supplementary Fig. 7.1 | Clickable retinoid (10 μ M) induced differentiation of dendritic cells. **a**, Graph showing the general ALDH activity of activated dendritic cells as determined by the ALDEFLUOR assay. **b**, Graph showing the differentiation of dendritic cells as measured by the expression of the CD103 marker. For parts **a** and **b**, data represent mean values $\pm$ SD. $N = 3$ experiments per group (biological replicates).



Supplementary Fig. 7.2 | Flow cytometry approach to visualize clickable vitamins in A549 cells. Fluorescence intensity of Alexa Fluor 488 conjugated to clickable vitamins of cells treated with vehicle (EtOH), STA-211 (10 μ M) or STA-215 (10 μ M).

Supplementary Table 7.1 | Proteins significantly enriched by STA-215 (10 μ M) in HL60 cells corresponding to Fig. 7.3b,c.

Name	Accession	Uniprot ID	Fold-change (UV/noUV)	Significance (-log ₁₀)	RA competed
Phosphoglycerate kinase 2	P07205	PGK2_HUMAN	20.0	7.7	Yes
Endoplasmic reticulum junction formation protein lunapark	Q9C0E8	LNP_HUMAN	20.0	4.9	Yes
Peroxisiredoxin-4	Q13162	PRDX4_HUMAN	20.0	4.8	Yes
Dynamin-like 120 kDa protein_mitochondrial	O60313	OPA1_HUMAN	20.0	4.5	Yes
ATPase family AAA domain-containing protein 1	Q8NBU5	ATAD1_HUMAN	20.0	4.4	Yes
Tubulin beta-6 chain	Q9BUF5	TBB6_HUMAN	20.0	4.2	Yes
Protein SCO2 homolog_mitochondrial	O43819	SCO2_HUMAN	20.0	3.6	Yes
Membrane-associated progesterone receptor component 2	O15173	PGRC2_HUMAN	2.0	2.3	Yes
Retinol dehydrogenase 11	Q8TC12	RDH11_HUMAN	20.0	1.6	Yes
Fatty aldehyde dehydrogenase	P51648	AL3A2_HUMAN	20.0	1.5	Yes
Ubiquitin-associated domain-containing protein 2	Q8NBM4	UBAC2_HUMAN	20.0	1.5	Yes
Nucleoside diphosphate kinase B	P22392	NDKB_HUMAN	20.0	5.7	Yes
Keratin_type I cytoskeletal 19	P08727	K1C19_HUMAN	2.9	1.5	No
Alpha-2-HS-glycoprotein	P02765	PETUA_HUMAN	2.3	2.2	No
Golgi phosphoprotein 3	Q9H4A6	GOLP3_HUMAN	20.0	4.4	Yes
Neutral cholesterol ester hydrolase 1	Q6PIU2	NCEH1_HUMAN	20.0	4.9	Yes
Transmembrane emp24 domain-containing protein 10	P49755	TMEDA_HUMAN	8.3	3.9	Yes
ATPase ASNA1	O43681	ASNA_HUMAN	20.0	4.5	Yes
Translocation protein SEC63 homolog	Q9UGP8	SEC63_HUMAN	2.3	2.6	Yes
Transmembrane protein 214	Q6NUQ4	TM214_HUMAN	3.0	3.6	Yes
Protein FAM162A	Q96A26	F162A_HUMAN	20.0	2.4	Yes
Sarcoplasmic/endoplasmic reticulum calcium ATPase 3	Q93084	AT2A3_HUMAN	3.7	3.3	Yes
ADP-ribosylation factor 4	P18085	ARF4_HUMAN	3.0	2.2	No
Transmembrane protein 43	Q9BTV4	TMM43_HUMAN	3.4	3.5	Yes
Very-long-chain 3-oxoacyl-CoA reductase	Q53GQ0	DHB12_HUMAN	6.4	3.5	Yes
Alkylidihydroxyacetonephosphate synthase_peroxisomal	O00116	ADAS_HUMAN	2.3	3.6	Yes
Thioredoxin-related transmembrane protein 1	Q9H3N1	TMX1_HUMAN	3.3	4.1	Yes
Transmembrane protein 109	Q9BVC6	TM109_HUMAN	4.6	5.1	Yes
Adipocyte plasma membrane-associated protein	Q9HDC9	APMAP_HUMAN	2.7	3.3	Yes
Ras-related protein Rab-14	P61106	RAB14_HUMAN	20.0	3.4	Yes
Monocarboxylate transporter 1	P53985	MOT1_HUMAN	2.4	3.6	Yes
Erlin-1	O75477	ERLN1_HUMAN	20.0	8.0	Yes
Translocation protein SEC62	Q99442	SEC62_HUMAN	2.6	4.0	Yes
BR13-binding protein	Q8WY22	BR13B_HUMAN	3.4	4.0	Yes
Vesicular integral-membrane protein VIP36	Q12907	LMAN2_HUMAN	3.9	4.7	Yes
Lysosome-associated membrane glycoprotein 1	P11279	LAMP1_HUMAN	20.0	1.6	No
Very-long-chain enoyl-CoA reductase	Q9NZ01	TECR_HUMAN	2.1	2.9	Yes
Long-chain-fatty-acid--CoA ligase 1	P33121	ACSL1_HUMAN	20.0	4.5	Yes
Cation-dependent mannose-6-phosphate receptor	P20645	MPRD_HUMAN	3.6	3.3	Yes
Transmembrane emp24 domain-containing protein 9	Q9BVK6	TMED9_HUMAN	3.0	3.3	Yes
Down syndrome cell adhesion molecule-like protein 1	Q8TD84	DSCL1_HUMAN	2.2	2.1	No
Very-long-chain (3R)-3-hydroxyacyl-CoA dehydratase 3	Q9P035	HACD3_HUMAN	2.7	3.3	No
Cytochrome b-c1 complex subunit 7	P14927	QCR7_HUMAN	20.0	3.6	No
Derlin-1	Q9BUN8	DERL1_HUMAN	20.0	7.6	No
Major facilitator superfamily domain-containing protein 10	Q14728	MFS10_HUMAN	2.4	2.5	No
Nuclear pore membrane glycoprotein 210	Q8TEM1	PO210_HUMAN	2.1	1.5	No
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit	P39656	OST48_HUMAN	2.3	4.0	No
Endoplasmic reticulum resident protein 29	P30040	ERP29_HUMAN	3.3	4.2	No
Voltage-dependent anion-selective channel protein 1	P21796	VDAC1_HUMAN	12.2	4.4	No
Probable ergosterol biosynthetic protein 28	Q9UKR5	ERG28_HUMAN	3.1	4.4	No
Reticulon-4	Q9NQC3	RTN4_HUMAN	2.9	1.4	No
Endoplasmic reticulum chaperone BiP	P11021	BIP_HUMAN	2.2	3.5	No
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 2	P04844	RPN2_HUMAN	2.5	4.3	No
Signal recognition particle receptor subunit beta	Q9Y5M8	SRPRB_HUMAN	3.6	4.3	No
Lanosterol synthase	P48449	ERG7_HUMAN	2.2	3.4	No
3-beta-hydroxysteroid-Delta(8)-Delta(7)-isomerase	Q15125	EBP_HUMAN	3.1	2.5	No
Receptor expression-enhancing protein 4	Q9H6H4	REEP4_HUMAN	3.9	4.5	No
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3B	Q8TCJ2	STT3B_HUMAN	2.4	3.4	No
Receptor expression-enhancing protein 5	Q00765	REEP5_HUMAN	2.1	1.9	No
Apolipoprotein A-I	P02647	APOA1_HUMAN	18.9	1.8	No
Monocarboxylate transporter 4	O15427	MOT4_HUMAN	2.9	2.8	No
Cullin-associated NEDD8-dissociated protein 1	Q86VP6	CAND1_HUMAN	3.7	2.6	No
Extended synaptotagmin-1	Q9BSJ8	ESYT1_HUMAN	10.9	4.0	No
Cleft lip and palate transmembrane protein 1-like protein	Q96KA5	CLP1L_HUMAN	3.8	3.1	No
AFG3-like protein 2	Q9Y4W6	AFG32_HUMAN	20.0	5.6	No
Sigma intracellular receptor 2	Q5BJF2	SGMR2_HUMAN	20.0	3.0	No
Transmembrane protein 33	P57088	TMM33_HUMAN	2.6	3.8	No
Carnitine O-palmitoyltransferase 1_liver isoform	P50416	CPT1A_HUMAN	3.6	3.6	No
NADH-cytochrome b5 reductase 3	P00387	NBSR3_HUMAN	2.1	1.7	No
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1	P04843	RPN1_HUMAN	2.2	3.6	No
Large neutral amino acids transporter small subunit 1	Q01650	LAT1_HUMAN	2.8	3.2	No
Voltage-dependent anion-selective channel protein 2	P45880	VDAC2_HUMAN	8.8	4.5	No
B-cell receptor-associated protein 31	P51572	BAP31_HUMAN	3.0	1.7	No
Keratin_type II cytoskeletal 1b	Q7Z794	K2C1B_HUMAN	20.0	3.2	No
CAAX prenyl protease 1 homolog	O75844	FACE1_HUMAN	2.5	2.6	No
Protein kish-A	Q8TBQ9	KISHA_HUMAN	2.4	3.0	No
Protein transport protein Sec61 subunit alpha isoform 1	P61619	S61A1_HUMAN	2.8	3.4	No
Atlastin-3	Q6DD88	ATLA3_HUMAN	20.0	4.1	No
Protein disulfide-isomerase TMX3	Q96J77	TMX3_HUMAN	2.0	3.2	No
Basigin	P35613	BAS1_HUMAN	2.8	3.5	No
Protein disulfide-isomerase	P07237	PDI1A_HUMAN	2.7	4.7	No
Translocon-associated protein subunit alpha	P43307	SSRA_HUMAN	3.2	3.7	No
Inactive hydroxysteroid dehydrogenase-like protein 1	Q3SXM5	HSDL1_HUMAN	20.0	4.4	No
Lysophosphatidylcholine acyltransferase 2	Q7L5N7	PCAT2_HUMAN	2.3	4.0	No
Protein YIPF5	Q969M3	YIPF5_HUMAN	2.1	1.8	No
ADP-ribosylation factor-like protein 6-interacting protein 1	Q15041	AR6P1_HUMAN	20.0	2.9	No
Acylglycerol kinase_mitochondrial	Q53H12	AGK_HUMAN	20.0	4.8	No
Mitochondrial fission process protein 1	Q9UDX5	MTFP1_HUMAN	20.0	3.4	No

SYNTHESIS AND BIOLOGICAL EVALUATION OF CLICKABLE VITAMIN A

Name	Accession	Uniprot ID	Fold-change (UV/noUV)	Significance (-log ₁₀)	RA competed
Heat shock protein beta-1	P04792	HSPB1_HUMAN	2.1	2.8	No
Reticulon-3	O95197	RTN3_HUMAN	2.8	2.6	No
Translocator protein	P30536	TSPO_HUMAN	3.3	3.6	No
Translocon-associated protein subunit delta	P51571	SSRD_HUMAN	3.9	3.7	No
4F2 cell-surface antigen heavy chain	P08195	4F2_HUMAN	20.0	4.7	No
Leukocyte surface antigen CD47	Q08722	CD47_HUMAN	2.1	2.8	No
Voltage-dependent anion-selective channel protein 3	Q9Y277	VDAC3_HUMAN	6.7	4.0	No
Prohibitin-2	Q99623	PHB2_HUMAN	7.0	5.1	No
Calcium-binding mitochondrial carrier protein Aralar1	O75746	CMC1_HUMAN	20.0	4.6	No
Serum albumin	P02768	ALBU_HUMAN	3.7	3.1	No
Calcium-binding mitochondrial carrier protein Aralar2	Q9UJS0	CMC2_HUMAN	20.0	6.2	No
Keratin_type II cytoskeletal 6B	P04259	K2C6B_HUMAN	20.0	2.6	No
Prohibitin	P35232	PHB_HUMAN	4.8	6.3	No
Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	P16615	AT2A2_HUMAN	2.1	3.9	No
NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 10_mitochondrial	O95299	NDUAA_HUMAN	20.0	4.2	No
Ras-related protein Rab-10	P61026	RAB10_HUMAN	2.7	3.4	No
Neutral amino acid transporter B(0)	Q15758	AAAT_HUMAN	3.2	4.5	No
Myeloid-associated differentiation marker	Q96S97	MYADM_HUMAN	3.2	3.7	No
Ras-related protein Rab-39A	Q14964	RB39A_HUMAN	20.0	4.6	No
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit DAD1	P61803	DAD1_HUMAN	2.3	2.7	No
Sodium/potassium-transporting ATPase subunit beta-3	P54709	AT1B3_HUMAN	4.0	4.3	No
Ras-related C3 botulinum toxin substrate 2	P15153	RAC2_HUMAN	2.0	2.8	No
Mitochondrial import inner membrane translocase subunit TIM50	Q3ZCQ8	TIM50_HUMAN	5.2	5.8	No
Calcium load-activated calcium channel	Q9UM00	TMCO1_HUMAN	2.4	3.9	No
Sodium/potassium-transporting ATPase subunit alpha-2	P50993	AT1A2_HUMAN	11.8	1.9	No
NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 8_mitochondrial	O95169	NDUB8_HUMAN	20.0	3.7	No
Arachidonate 5-lipoxygenase-activating protein	P20292	AL5AP_HUMAN	2.2	1.7	No
Ras-related protein Rap-1b	P61224	RAP1B_HUMAN	2.1	2.6	No
Ubiquitin-60S ribosomal protein L40	P62987	RL40_HUMAN	2.6	5.0	No
Mitochondrial import receptor subunit TOM40 homolog	O96008	TOM40_HUMAN	3.0	3.5	No
ADP-dependent glucokinase	Q9BRR6	ADPGK_HUMAN	20.0	2.8	No
Sorting and assembly machinery component 50 homolog	Q9Y512	SAM50_HUMAN	20.0	6.4	No
Dol-P-Man:Man(5)GlcNAc(2)-PP-Dol alpha-1_3-mannosyltransferase	Q92685	ALG3_HUMAN	2.6	1.8	No
ATPase family AAA domain-containing protein 3A	Q9NV17	ATD3A_HUMAN	2.3	3.7	No
GDP-Man:Man(3)GlcNAc(2)-PP-Dol alpha-1_2-mannosyltransferase	Q2TAA5	ALG11_HUMAN	2.9	3.9	No
Protein disulfide-isomerase A6	Q15084	PDIA6_HUMAN	2.3	2.7	No
Sigma non-opioid intracellular receptor 1	Q99720	SGMR1_HUMAN	6.6	4.2	No
ATPase family AAA domain-containing protein 3B	Q5T9A4	ATD3B_HUMAN	2.0	4.0	No
GPI-anchor transamidase	Q92643	GPI8_HUMAN	20.0	5.1	No
Estradiol 17-beta-dehydrogenase 11	Q8NBQ5	DHBT1_HUMAN	2.4	2.8	No
CDGSH iron-sulfur domain-containing protein 2	Q8N5K1	CISD2_HUMAN	20.0	1.4	No
Phosphatidate cytidyltransferase_mitochondrial	Q96BW9	TAM41_HUMAN	20.0	3.6	No
Protein disulfide-isomerase A3	P30101	PDIA3_HUMAN	2.3	4.1	No
Equilibrative nucleoside transporter 1	Q99808	S29A1_HUMAN	3.0	3.4	No
NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 11_mitochondrial	Q9NX14	NDUBB_HUMAN	20.0	4.2	No
Myosin-11	P35749	MYH11_HUMAN	2.5	2.6	No
MICOS complex subunit MIC60	Q16891	MIC60_HUMAN	3.4	4.6	No
Lysocardiolipin acyltransferase 1	Q6UWP7	LCLT1_HUMAN	2.9	2.8	No
Cytochrome c-type heme lyase	P53701	CHCL_HUMAN	20.0	3.8	No
Sideroflexin-1	Q9H9B4	SFXN1_HUMAN	2.9	4.3	No
Mitochondrial inner membrane protein OXA1L	Q15070	OXA1L_HUMAN	4.6	4.4	No
Guanine nucleotide-binding protein G(o) subunit alpha	P09471	GNAO_HUMAN	2.4	2.5	No
Signal peptidase complex subunit 3	P61009	SPCS3_HUMAN	20.0	1.5	No
DnaJ homolog subfamily C member 11	Q9NVH1	DJC11_HUMAN	20.0	5.8	No
Tricarboxylate transport protein_mitochondrial	P53007	TXTPT_HUMAN	2.2	4.8	No
ADP/ATP translocase 3	P12236	ADT3_HUMAN	3.3	3.8	No
Microsomal glutathione S-transferase 3	O14880	MGST3_HUMAN	3.2	2.6	No
GTP-binding protein SAR1a	Q9NR31	SAR1A_HUMAN	2.2	1.7	No
Tetrapeptide repeat protein 39C	Q8N584	TT39C_HUMAN	20.0	5.0	No
Lamina-associated polypeptide 2_isoforms beta/gamma	P42167	LAP2B_HUMAN	2.1	2.5	No
Phosphate carrier protein_mitochondrial	Q00325	MPCP_HUMAN	4.4	5.2	No
Ras-related protein Rab-8A	P61006	RAB8A_HUMAN	2.1	1.8	No
Mitochondrial import receptor subunit TOM70	O94826	TOM70_HUMAN	20.0	6.4	No
Apolipoprotein M	O95445	APOM_HUMAN	20.0	3.2	No
Cytochrome c oxidase subunit 2	P00403	COX2_HUMAN	2.2	2.7	No
Squalene synthase	P37268	FDFT_HUMAN	2.3	3.0	No
Ras-related protein Rab-5C	P51148	RAB5C_HUMAN	20.0	4.2	No
Vitamin K epoxide reductase complex subunit 1-like protein 1	Q8N0U8	VKORL_HUMAN	20.0	3.5	No
ADP/ATP translocase 2	P05141	ADT2_HUMAN	9.8	6.2	No
Mitochondrial dicarboxylate carrier	Q9UBX3	DIC_HUMAN	20.0	6.0	No
ATP synthase subunit g_mitochondrial	O75964	ATP5L_HUMAN	2.9	3.8	No
Thioredoxin domain-containing protein 5	Q8NBS9	TXND5_HUMAN	2.1	3.5	No
ATP-binding cassette sub-family D member 3	P28288	ABCD3_HUMAN	20.0	6.8	No
MICOS complex subunit MIC13	Q5XKP0	MIC13_HUMAN	20.0	2.6	No
Vesicle-associated membrane protein 8	Q9BV40	VAMP8_HUMAN	2.5	2.2	No
Protein ERGIC-53	P49257	LMAN1_HUMAN	3.0	4.4	No
Membrane-associated progesterone receptor component 1	O00264	PGRC1_HUMAN	2.9	2.6	No
Synaptic vesicle membrane protein VAT-1 homolog	Q99536	VAT1_HUMAN	2.7	3.9	No
Solute carrier family 43 member 3	Q8NB15	S43A3_HUMAN	20.0	3.6	No
Hemoglobin subunit alpha	P69905	HBA_HUMAN	2.7	2.2	No
NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 9_mitochondrial	Q16795	NDUA9_HUMAN	2.2	2.4	No
Mitochondrial import receptor subunit TOM20 homolog	Q15388	TOM20_HUMAN	2.3	2.0	No
Mitochondrial proton/calcium exchanger protein	O95202	LETM1_HUMAN	2.5	3.1	No
ATP-binding cassette sub-family B member 10_mitochondrial	Q9NRK6	ABCB10_HUMAN	2.5	3.7	No
Guanine nucleotide-binding protein subunit alpha-13	Q14344	GNA13_HUMAN	3.5	5.9	No
Ran-specific GTPase-activating protein	P43487	RANG_HUMAN	20.0	1.6	No
14-3-3 protein epsilon	P62258	I433E_HUMAN	2.7	1.9	No

CHAPTER 7

Name	Accession	Uniprot ID	Fold-change (UV/noUV)	Significance (-log ₁₀)	RA competed
Phosphatidylinositol glycan anchor biosynthesis class U protein	Q9H490	PIGU_HUMAN	2.1	1.9	No
NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 4	O95168	NDUB4_HUMAN	20.0	8.4	No
Emerin	P50402	EMD_HUMAN	20.0	5.6	No
Fatty acyl-CoA reductase 1	Q8WVX9	FACR1_HUMAN	20.0	1.6	No
Minor histocompatibility antigen H13	Q8TCT9	HM13_HUMAN	20.0	1.6	No
Calcium-binding mitochondrial carrier protein SCaMC-1	Q6NUK1	SCMC1_HUMAN	20.0	1.6	No
Ras-related protein Rab-27A	P51159	RB27A_HUMAN	20.0	1.6	No
Mitochondrial carrier homolog 2	Q9Y6C9	MTCH2_HUMAN	3.1	2.4	No
ADP-ribosylation factor 1	P84077	ARF1_HUMAN	3.6	3.4	No
Catechol O-methyltransferase	P21964	COMT_HUMAN	20.0	1.5	No
Mitochondrial import inner membrane translocase subunit TIM44	O43615	TIM44_HUMAN	20.0	7.3	No
Peptidyl-tRNA hydrolase 2_mitochondrial	Q9Y3E5	PTH2_HUMAN	20.0	2.8	No
OCIA domain-containing protein 1	Q9NX40	OCAD1_HUMAN	2.6	3.3	No
Spliceosome RNA helicase DDX39B	Q13838	DX39B_HUMAN	20.0	5.8	No
Cytochrome c oxidase subunit 6C	P09669	COX6C_HUMAN	20.0	1.5	No
Ras-related protein Rab-1A	P62820	RAB1A_HUMAN	2.7	2.7	No
POTE ankyrin domain family member E	Q6S8J3	POTEE_HUMAN	20.0	6.4	No

Supplementary Table 7.2 | Proteins significantly enriched by LEI-945 (1 μ M) in HL60 cells.

Name	Accession	Uniprot ID	Fold-change (LEI-945/DMSO)	Significance (-log ₁₀)
Voltage-dependent anion-selective channel protein 1	P21796	VDAC1_HUMAN	3.2	5.0
Ubiquitin-60S ribosomal protein L40	P62987	RL40_HUMAN	4.6	4.7
Voltage-dependent anion-selective channel protein 2	P45880	VDAC2_HUMAN	20.0	8.9
ADP/ATP translocase 2	P05141	ADT2_HUMAN	4.7	6.6
L-lactate dehydrogenase B chain	P07195	LDHB_HUMAN	2.0	3.7
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit DAD1	P61803	DAD1_HUMAN	3.2	6.3
Prohibitin-2	Q99623	PHB2_HUMAN	2.3	3.4
Microsomal glutathione S-transferase 3	O14880	MGST3_HUMAN	20.0	4.6
Tubulin alpha-1C chain	Q9BQE3	TBA1C_HUMAN	2.0	4.2
Voltage-dependent anion-selective channel protein 3	Q9Y277	VDAC3_HUMAN	20.0	5.6
ADP/ATP translocase 3	P12236	ADT3_HUMAN	3.2	5.7
Up-regulated during skeletal muscle growth protein 5	Q96IX5	USMG5_HUMAN	3.4	4.5
Protein disulfide-isomerase	P07237	PDIA1_HUMAN	8.1	7.9
Ras-related protein Rab-1A	P62820	RAB1A_HUMAN	20.0	5.7
Ras-related protein Rab-10	P61026	RAB10_HUMAN	20.0	5.4
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1	P04843	RPN1_HUMAN	3.4	5.5
Tubulin alpha-4A chain	P68366	TBA4A_HUMAN	20.0	4.2
Very-long-chain 3-oxoacyl-CoA reductase	Q53GQ0	DHB12_HUMAN	3.5	4.2
Ras-related protein Rab-8A	P61006	RAB8A_HUMAN	20.0	5.2
Phosphate carrier protein_mitochondrial	Q00325	MPCP_HUMAN	3.7	5.3
Transmembrane protein 33	P57088	TMM33_HUMAN	3.4	4.9
ADP-ribosylation factor 1	P84077	ARF1_HUMAN	4.4	5.4
Calnexin	P27824	CALX_HUMAN	2.1	7.2
Fructose-bisphosphate aldolase A	P04075	ALDOA_HUMAN	2.1	3.3
Apolipoprotein A-I	P02647	APOA1_HUMAN	20.0	6.5
Delta(14)-sterol reductase	Q14739	LBR_HUMAN	3.5	7.1
Sideroflexin-1	Q9H9B4	SFXN1_HUMAN	2.4	4.4
Protein kish-A	Q8TBQ9	KISHA_HUMAN	2.6	3.0
ATPase ASNA1	O43681	ASNA_HUMAN	20.0	7.5
Chitobiosylidiphosphodolichol beta-mannosyltransferase	Q9BT22	ALG1_HUMAN	3.3	4.8
Glucose-6-phosphate isomerase	P06744	G6PI_HUMAN	2.1	4.8
Plastin-2	P13796	PLSL_HUMAN	2.6	5.1
Elongation factor 1-gamma	P26641	EF1G_HUMAN	2.1	4.5
Thioredoxin-dependent peroxide reductase_mitochondrial	P30048	PRDX3_HUMAN	20.0	4.8
Mitochondrial carrier homolog 2	Q9Y6C9	MTCH2_HUMAN	20.0	4.3
Gamma-enolase	P09104	ENOG_HUMAN	20.0	3.6
Eukaryotic initiation factor 4A-I	P60842	IF4A1_HUMAN	2.5	4.7
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit	P39656	OST48_HUMAN	3.7	4.7
Translocon-associated protein subunit delta	P51571	SSRD_HUMAN	20.0	6.0
Mitochondrial import receptor subunit TOM40 homolog	O96008	TOM40_HUMAN	20.0	1.4
ATPase family AAA domain-containing protein 3A	Q9NV17	ATD3A_HUMAN	20.0	7.9
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3A	P46977	STT3A_HUMAN	3.8	4.8
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 2	P04844	RPN2_HUMAN	3.4	5.3
Keratin_type II cytoskeletal 6B	P04259	K2C6B_HUMAN	20.0	2.6
Signal peptidase complex catalytic subunit SEC11A	P67812	SC11A_HUMAN	20.0	3.5
Phosphoglycerate kinase 2	P07205	PGK2_HUMAN	20.0	5.7
Vitamin K epoxide reductase complex subunit 1-like protein 1	Q8N0U8	VKORL_HUMAN	20.0	3.2
Arachidonate 5-lipoxygenase-activating protein	P20292	AL5AP_HUMAN	3.5	4.6
Inactive hydroxysteroid dehydrogenase-like protein 1	Q3SXM5	HSDL1_HUMAN	20.0	6.2
Ras-related protein Rab-14	P61106	RAB14_HUMAN	20.0	6.1
Signal recognition particle receptor subunit beta	Q9Y5M8	SRPRB_HUMAN	7.4	6.9
Estradiol 17-beta-dehydrogenase 11	Q8NBQ5	DHB11_HUMAN	4.0	4.5
14-3-3 protein beta/alpha	P31946	I433B_HUMAN	20.0	6.2
Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	P16615	AT2A2_HUMAN	7.2	7.6
MICOS complex subunit MIC60	P16891	MIC60_HUMAN	20.0	6.0
Ras-related protein Rab-30	Q15771	RAB30_HUMAN	20.0	6.4
Neutral cholesterol ester hydrolase 1	Q6PIU2	NCEH1_HUMAN	20.0	5.3
Basigin	P35613	BASI_HUMAN	2.2	5.0
Vesicular integral-membrane protein VIP36	Q12907	LMAN2_HUMAN	3.3	5.3
Lamina-associated polypeptide 2_ isoforms beta/gamma	P42167	LAP2B_HUMAN	20.0	3.9
14-3-3 protein gamma	P61981	I433G_HUMAN	2.5	1.7
Major facilitator superfamily domain-containing protein 10	Q14728	MFS10_HUMAN	2.5	3.1
Ras-related protein Rab-39A	Q14964	RB39A_HUMAN	20.0	2.4
Heterogeneous nuclear ribonucleoprotein F	P52597	HNRPF_HUMAN	2.2	3.3
NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 4	O95168	NDUB4_HUMAN	20.0	6.1
Keratin_type II cytoskeletal 1b	Q7Z794	K2C1B_HUMAN	20.0	2.7
Ras-related protein Rab-5C	P51148	RAB5C_HUMAN	20.0	6.6
14-3-3 protein eta	Q04917	I433F_HUMAN	20.0	2.1
OCIA domain-containing protein 1	Q9NX40	OCAD1_HUMAN	20.0	6.8
Very-long-chain enoyl-CoA reductase	Q9NZ01	TECR_HUMAN	3.8	5.4
ATP synthase subunit g_mitochondrial	O75964	ATP5L_HUMAN	20.0	6.6
Lanosterol synthase	P48449	ERG7_HUMAN	2.4	6.3
ATP synthase subunit O_mitochondrial	P48047	ATPO_HUMAN	20.0	1.6
Thioredoxin-related transmembrane protein 1	Q9H3N1	TMX1_HUMAN	4.0	5.1
ATP-dependent RNA helicase DDX39A	O00148	DX39A_HUMAN	20.0	7.2
Guanine nucleotide-binding protein G(i) subunit alpha-2	P04899	GNAI2_HUMAN	2.1	2.8
Histidine triad nucleotide-binding protein 3	Q9NQE9	HINT3_HUMAN	20.0	4.4
Mitochondrial import inner membrane translocase subunit TIM50	Q3ZCQ8	TIM50_HUMAN	20.0	3.8
Transmembrane emp24 domain-containing protein 10	P49755	TMEDA_HUMAN	2.1	3.1
Acyl-CoA 6-desaturase	O95864	FADS2_HUMAN	2.2	4.8
Ras-related C3 botulinum toxin substrate 2	P15153	RAC2_HUMAN	2.7	4.0
Secretory carrier-associated membrane protein 3	O14828	SCAM3_HUMAN	20.0	4.4
Extended synaptotagmin-2	A0FGR8	ESYT2_HUMAN	3.8	4.4
Sorting and assembly machinery component 50 homolog	Q9Y512	SAM50_HUMAN	20.0	4.6
Alpha-actinin-1	P12814	ACTN1_HUMAN	2.1	2.5
Equilibrative nucleoside transporter 1	Q99808	S29A1_HUMAN	20.0	4.8
Extended synaptotagmin-1	Q9BSJ8	ESYT1_HUMAN	20.0	4.9

CHAPTER 7

Name	Accession	Uniprot ID	Fold-change (LEI-945/DMSO)	Significance (-log ₁₀)
Protein YIPF5	Q969M3	YIPF5_HUMAN	20.0	6.3
14-3-3 protein epsilon	P62258	I433E_HUMAN	3.1	3.1
Protein FAM162A	Q96A26	F162A_HUMAN	20.0	1.5
Reticulon-4	Q9NQC3	RTN4_HUMAN	20.0	3.4
NADH-cytochrome b5 reductase 3	P00387	NBSR3_HUMAN	2.1	2.3
Vesicle-associated membrane protein 8	Q9BV40	VAMP8_HUMAN	20.0	3.4
MICOS complex subunit MIC27	Q6UXV4	MIC27_HUMAN	20.0	6.2
Surfeit locus protein 4	O15260	SURF4_HUMAN	3.0	1.8
CAAX prenyl protease 1 homolog	O75844	FACE1_HUMAN	3.3	4.0
Very-long-chain (3R)-3-hydroxyacyl-CoA dehydratase 3	Q9P035	HACD3_HUMAN	3.2	4.0
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3B	Q8TCJ2	STT3B_HUMAN	3.9	6.0
Transmembrane protein 214	Q6NUQ4	TM214_HUMAN	2.6	4.0
Ras-related protein Rab-11A	P62491	RB11A_HUMAN	20.0	2.6
Synaptogyrin-2	O43760	SN2G2_HUMAN	20.0	4.9
Lysophosphatidylcholine acyltransferase 2	Q7L5N7	PCAT2_HUMAN	6.5	5.0
Protein disulfide-isomerase A6	Q15084	PDIA6_HUMAN	20.0	7.0
Translocon-associated protein subunit alpha	P43307	SSRA_HUMAN	20.0	3.8
Transmembrane protein 43	Q9BTV4	TMM43_HUMAN	3.0	3.8
Sigma non-opioid intracellular receptor 1	Q99720	SGMR1_HUMAN	20.0	4.2
Mitochondrial dicarboxylate carrier	Q9UBX3	DIC_HUMAN	20.0	1.6
Glutathione S-transferase omega-1	P78417	GSTO1_HUMAN	2.7	4.0
ADP-ribosylation factor-like protein 6-interacting protein 1	Q15041	AR6P1_HUMAN	20.0	3.5
Phosphatidylserine synthase 1	P48651	PTSS1_HUMAN	20.0	3.3
Malectin	Q14165	MLEC_HUMAN	2.4	2.9
Mitochondrial inner membrane protein OXA1L	Q15070	OXA1L_HUMAN	20.0	5.3
3-beta-hydroxysteroid-Delta(8)_Delta(7)-isomerase	Q15125	EBP_HUMAN	4.2	3.1
Translocator protein	P30536	TSPO_HUMAN	3.5	3.8
Long-chain-fatty-acid--CoA ligase 3	O95573	ACSL3_HUMAN	20.0	4.9
Catechol O-methyltransferase	P21964	COMT_HUMAN	20.0	3.8
T-complex protein 1 subunit beta	P78371	TCPB_HUMAN	2.1	3.0
Tropomyosin alpha-3 chain	P06753	TPM3_HUMAN	20.0	1.6
Carnitine O-palmitoyltransferase 1_liver isoform	P50416	CPT1A_HUMAN	20.0	6.6
ATP-binding cassette sub-family D member 3	P28288	ABCD3_HUMAN	20.0	5.5
Cleft lip and palate transmembrane protein 1-like protein	Q96KA5	CLP1L_HUMAN	20.0	5.9
Protein transport protein Sec61 subunit alpha isoform 1	P61619	S61A1_HUMAN	4.7	4.1
B-cell receptor-associated protein 31	P51572	BAP31_HUMAN	5.5	3.2
ATP synthase F(0) complex subunit B1_mitochondrial	P24539	AT5F1_HUMAN	2.3	4.7
Calcium load-activated calcium channel	Q9UM00	TMCO1_HUMAN	2.9	3.2
Peroxisomal biogenesis factor 3	P56589	PEX3_HUMAN	2.7	3.2
Proteasome subunit beta type-2	P49721	PSB2_HUMAN	20.0	4.9
Monocarboxylate transporter 4	O15427	MOT4_HUMAN	6.7	6.0
Calcium-binding mitochondrial carrier protein Aralar2	Q9UJS0	CMC2_HUMAN	20.0	1.5
Mitochondrial proton/calcium exchanger protein	O95202	LETM1_HUMAN	20.0	5.0
Receptor expression-enhancing protein 5	Q00765	REEP5_HUMAN	20.0	7.7
Cytochrome c oxidase subunit 2	P00403	COX2_HUMAN	2.2	2.8
RUS1 family protein C16orf58	Q96GQ5	RUS1_HUMAN	20.0	6.9
Guanine nucleotide-binding protein G(o) subunit alpha	P09471	GNAO1_HUMAN	20.0	3.1
Lysocardiolipin acyltransferase 1	Q6UWP7	LCLT1_HUMAN	20.0	5.8
Protein jagunal homolog 1	Q8NSM9	JAGN1_HUMAN	20.0	6.2
Guanine nucleotide-binding protein subunit alpha-13	Q14344	GNA13_HUMAN	2.4	3.1
Sodium/potassium-transporting ATPase subunit beta-3	P54709	AT1B3_HUMAN	20.0	4.9
Myosin-10	P35580	MYH10_HUMAN	2.4	2.6
Lysosome-associated membrane glycoprotein 1	P11279	LAMP1_HUMAN	20.0	1.6
NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 11_mitochondrial	Q9NX14	NDUBB_HUMAN	20.0	1.6
Myosin-11	P35749	MYH11_HUMAN	8.3	4.3
Peptidyl-tRNA hydrolase 2_mitochondrial	Q9Y3E5	PTH2_HUMAN	20.0	3.7
Thymidylate synthase	P04818	TYSY_HUMAN	20.0	7.5
Ataxin-10	Q9UBB4	ATX10_HUMAN	20.0	8.2
ATP-binding cassette sub-family B member 10_mitochondrial	Q9NRK6	ABCB10_HUMAN	20.0	7.2
Translocation protein SEC63 homolog	Q9UGP8	SEC63_HUMAN	2.1	3.5
Myeloblastin	P24158	PRTN3_HUMAN	2.2	3.2
Vesicle-associated membrane protein-associated protein A	Q9POL0	VAPA_HUMAN	2.9	3.0
Methylsterol monooxygenase 1	Q15800	MSMO1_HUMAN	3.4	4.8
Reticulon-3	O95197	RTN3_HUMAN	20.0	2.6
Asparagine synthetase [glutamine-hydrolyzing]	P08243	ASNS_HUMAN	5.4	5.1
Membrane-associated progesterone receptor component 1	O00264	PGRC1_HUMAN	20.0	1.6
Dol-P-Man:Man(5)GlcNAc(2)-PP-Dol alpha-1_3-mannosyltransferase	Q92685	ALG3_HUMAN	20.0	4.5
MICOS complex subunit MIC13	Q5XKP0	MIC13_HUMAN	20.0	4.1
Calcium-binding mitochondrial carrier protein Aralar1	O75746	CMC1_HUMAN	20.0	1.5
Leukocyte surface antigen CD47	Q08722	CD47_HUMAN	20.0	4.0
Receptor expression-enhancing protein 4	Q9H6H4	REEP4_HUMAN	20.0	5.9
Very long-chain acyl-CoA synthetase	O14975	S27A2_HUMAN	20.0	5.4
Importin-5	O00410	IPO5_HUMAN	2.4	3.2
Transmembrane protein 109	Q9BVC6	TM109_HUMAN	2.9	3.3
Translocation protein SEC62	Q99442	SEC62_HUMAN	3.3	3.3
Signal peptidase complex subunit 3	P61009	SPCS3_HUMAN	20.0	5.4
DNA replication licensing factor MCM6	Q14566	MCM6_HUMAN	2.4	4.4
Protein disulfide-isomerase TMX3	Q96J77	TMX3_HUMAN	20.0	8.2
Signal recognition particle receptor subunit alpha	P08240	SRPRA_HUMAN	2.6	3.2
4F2 cell-surface antigen heavy chain	P08195	4F2_HUMAN	20.0	6.6
Saccharopine dehydrogenase-like oxidoreductase	Q8NBX0	SCPDL_HUMAN	20.0	7.1
Myeloid-associated differentiation marker	Q96S97	MYADM_HUMAN	4.8	2.6
Tyrosine-protein kinase Lyn	P07948	LYN_HUMAN	20.0	1.6
Derlin-1	Q9BUN8	DERL1_HUMAN	20.0	5.9
Endoplasmic reticulum junction formation protein lunapark	Q9C0E8	LNP_HUMAN	20.0	1.6
Monocarboxylate transporter 1	P53985	MOT1_HUMAN	3.3	3.8
Sterol-4-alpha-carboxylate 3-dehydrogenase decarboxylating	Q15738	NSDHL_HUMAN	20.0	6.9

SYNTHESIS AND BIOLOGICAL EVALUATION OF CLICKABLE VITAMIN A

Name	Accession	Uniprot ID	Fold-change (LEI-945/DMSO)	Significance (-log ₁₀)
Atlastin-3	Q6DD88	ATLA3_HUMAN	20.0	5.4
Scavenger receptor class B member 1	Q8WTV0	SCRBI_HUMAN	5.2	5.4
Small ubiquitin-related modifier 1	P63165	SUMO1_HUMAN	20.0	3.4
26S proteasome non-ATPase regulatory subunit 11	O00231	PSD11_HUMAN	20.0	3.9
Neutral amino acid transporter B(0)	Q15758	AAAT_HUMAN	20.0	7.5
Long-chain-fatty-acid--CoA ligase 4	O60488	ACSL4_HUMAN	20.0	7.3
Heterogeneous nuclear ribonucleoprotein M	P52272	HNRPM_HUMAN	20.0	8.3
Minor histocompatibility antigen H13	Q8TCT9	HM13_HUMAN	20.0	1.6
Threonylcarbamoyladenosine tRNA methyltransferase	Q5VV42	CDKAL_HUMAN	20.0	6.3
Ras-related protein Rab-27A	P51159	RB27A_HUMAN	20.0	4.9
Endoplasmic reticulum resident protein 29	P30040	ERP29_HUMAN	20.0	1.6
Erlin-1	O75477	ERLN1_HUMAN	20.0	5.1
TraB domain-containing protein	Q9H4I3	TRABD_HUMAN	20.0	5.2
Keratin_type I cytoskeletal 12	Q99456	K1C12_HUMAN	20.0	1.4
Basic leucine zipper and W2 domain-containing protein 1	Q7L1Q6	BZW1_HUMAN	20.0	5.0
ADP-dependent glucokinase	Q9BRR6	ADPGK_HUMAN	20.0	4.0
NADH dehydrogenase [ubiquinone] iron-sulfur protein 2_ mitochondrial	O75306	NDUS2_HUMAN	20.0	3.8
tRNA (cytosine(34)-C(5))-methyltransferase	Q08J23	NSUN2_HUMAN	20.0	6.8
Protein RFT1 homolog	Q96AA3	RFT1_HUMAN	3.5	3.9
Mitochondrial import receptor subunit TOM70	O94826	TOM70_HUMAN	20.0	1.6
Ubiquitin-associated domain-containing protein 2	Q8NBM4	UBAC2_HUMAN	20.0	5.2
NADH-ubiquinone oxidoreductase chain 5	P03915	NU5M_HUMAN	20.0	4.8
Vesicle-associated membrane protein 7	P51809	VAMP7_HUMAN	20.0	4.6
Protein ERGIC-53	P49257	LMAN1_HUMAN	20.0	4.2
Nicalin	Q969V3	NCLN_HUMAN	20.0	5.1
Choline/ethanolaminephosphotransferase 1	Q9Y6K0	CEPT1_HUMAN	20.0	6.7
Peptidyl-prolyl cis-trans isomerase FKBP8	Q14318	FKBP8_HUMAN	20.0	5.9
DnaJ homolog subfamily C member 11	Q9NVH1	DJC11_HUMAN	20.0	6.2
Squalene synthase	P37268	FDFT_HUMAN	20.0	5.4
Sorbitol dehydrogenase	Q00796	DHSO_HUMAN	20.0	7.0
Thioredoxin domain-containing protein 5	Q8NBS9	TXND5_HUMAN	20.0	7.5
CDGSH iron-sulfur domain-containing protein 2	Q8NSK1	CISD2_HUMAN	20.0	4.5
Signal recognition particle subunit SRP68	Q9UHB9	SRP68_HUMAN	20.0	5.4
Retinol dehydrogenase 11	Q8TC12	RDH11_HUMAN	20.0	4.6
Cation-dependent mannose-6-phosphate receptor	P20645	MPRD_HUMAN	20.0	4.5
GPI-anchor transamidase	Q92643	GPI8_HUMAN	20.0	1.6
Heat shock 70 kDa protein 4	P34932	HSP74_HUMAN	2.1	3.4
Adenosine 3'-phospho 5'-phosphosulfate transporter 1	Q8TB61	S35B2_HUMAN	20.0	4.3
GDP-Man:Man(3)GlcNAc(2)-PP-Dol alpha-1_2-mannosyltransferase	Q2TAA5	ALG11_HUMAN	20.0	4.9
Phosphatidylinositol glycan anchor biosynthesis class U protein	Q9H490	PIGU_HUMAN	2.9	2.7
Spermidine synthase	P19623	SPEE_HUMAN	20.0	1.6
Heme oxygenase 2	P30519	HMOX2_HUMAN	20.0	5.5
Exportin-2	P55060	XPO2_HUMAN	2.0	4.5
Solute carrier family 43 member 3	Q8NBI5	S43A3_HUMAN	20.0	5.5
Cytochrome c-type heme lyase	P53701	CCHL_HUMAN	20.0	1.6
Fatty aldehyde dehydrogenase	P51648	AL3A2_HUMAN	20.0	6.4
ATP-dependent 6-phosphofructokinase_liver type	P17858	PFKAL_HUMAN	20.0	8.0
Sigma intracellular receptor 2	Q5BJF2	SGMR2_HUMAN	20.0	2.4
Sarcoplasmic/endoplasmic reticulum calcium ATPase 3	Q93084	AT2A3_HUMAN	20.0	7.0
Large neutral amino acids transporter small subunit 1	Q01650	LAT1_HUMAN	20.0	6.4
Kelch-like ECH-associated protein 1	Q14145	KEAP1_HUMAN	20.0	6.5
Long-chain-fatty-acid--CoA ligase 1	P33121	ACSL1_HUMAN	20.0	5.4
PCI domain-containing protein 2	Q5JVF3	PCID2_HUMAN	20.0	1.6
FAS-associated factor 2	Q96CS3	FAF2_HUMAN	20.0	5.1
Glycine--tRNA ligase	P41250	GARS_HUMAN	20.0	3.2
Sphingomyelin phosphodiesterase 4	Q9NXE4	NSMA3_HUMAN	20.0	6.9
Protein SET	Q01105	SET_HUMAN	20.0	6.1

Supplementary Table 7.3 | Proteins significantly enriched by STA-211 (10 μ M) in A549 cells corresponding to Fig. 7.5c.

Name	Accession	Uniprot ID	Fold change (UV/EtOH)	Significance (-log ₁₀)	UV enriched
15-hydroxyprostaglandin dehydrogenase [NAD(+)]	P15428	PGDH_HUMAN	20.0	1.5	Yes
Keratin type II cytoskeletal 1b	Q7Z794	K2C1B_HUMAN	20.0	2.2	Yes
Keratin type I cytoskeletal 17	Q04695	K1C17_HUMAN	20.0	3.3	Yes
Neutral cholesterol ester hydrolase 1	Q6PIU2	NCEH1_HUMAN	20.0	2.1	Yes
ATPase ASNA1	O43681	ASNA_HUMAN	20.0	1.6	Yes
Magnesium transporter protein 1	Q9H0U3	MAGT1_HUMAN	20	4.8	Yes
Epoxide hydrolase 1	P07099	HYEP_HUMAN	4.0	2.8	Yes
Ras-related protein Rab-10	P61026	RAB10_HUMAN	20.0	5	No
Fatty aldehyde dehydrogenase	P51648	AL3A2_HUMAN	20.0	4.3	No
Ras-related protein Rab-1A	P62820	RAB1A_HUMAN	20.0	4.6	No
Alpha-actinin-1	P12814	ACTN1_HUMAN	20.0	3.2	No
14-3-3 protein theta	P27348	1433T_HUMAN	20.0	4	No
Vimentin	P08670	VIME_HUMAN	20.0	5.1	No
Voltage-dependent anion-selective channel protein 3	Q9Y277	VDAC3_HUMAN	20.0	4.1	No
14-3-3 protein gamma	P61981	1433G_HUMAN	20.0	4.7	No
14-3-3 protein beta/alpha	P31946	1433B_HUMAN	20.0	3.5	No
Keratin type II cytoskeletal 7	P08729	K2C7_HUMAN	20.0	4.7	No
Tubulin beta-4B chain	P68371	TBB4B_HUMAN	20.0	5.5	No
Tubulin alpha-4A chain	P68366	TBA4A_HUMAN	20.0	5.5	No
Dolichyl-diphosphooligosaccharide--protein subunit DAD1	P61803	DAD1_HUMAN	20.0	3.7	No
14-3-3 protein epsilon	P62258	1433E_HUMAN	20.0	3.6	No
Microsomal glutathione S-transferase 3	O14880	MGST3_HUMAN	20.0	2.7	No
Poly(rC)-binding protein 1	Q15365	PCBP1_HUMAN	20.0	4.9	No
Poly(rC)-binding protein 2	Q15366	PCBP2_HUMAN	20.0	3.4	No
Cytochrome b5 type B	O43169	CYB5B_HUMAN	20.0	2.8	No
ATP synthase subunit alpha mitochondrial	P25705	ATPA_HUMAN	20.0	4.1	No
Keratin type I cytoskeletal 13	P13646	K1C13_HUMAN	20.0	3.1	No
Prohibitin-2	Q99623	PHB2_HUMAN	20.0	3.2	No
Heat shock 70 kDa protein 1A	P0DMV8	HS71A_HUMAN	20.0	5	No
Prohibitin	P35232	PHB_HUMAN	20.0	3.6	No
Keratin type I cytoskeletal 19	P08727	K1C19_HUMAN	20.0	5.9	No
Keratin type II cytoskeletal 75	O95678	K2C75_HUMAN	20.0	3.5	No
Sulfide:quinone oxidoreductase mitochondrial	Q9Y6N5	SQOR_HUMAN	20.0	5	No
Sideroflexin-1	Q9H9B4	SFXN1_HUMAN	20.0	3.6	No
40S ribosomal protein S3	P23396	RS3_HUMAN	20.0	4.5	No
Mitochondrial import inner membrane translocase subunit TIM50	Q3ZCQ8	TIM50_HUMAN	20.0	1.6	No
40S ribosomal protein S20	P60866	RS20_HUMAN	20.0	3.1	No
Putative heat shock protein HSP 90-beta 2	Q58F88	H90B2_HUMAN	20.0	3.5	No
Trifunctional enzyme subunit beta mitochondrial	P55084	ECHB_HUMAN	20.0	6	No
78 kDa glucose-regulated protein	P11021	BIP_HUMAN	20.0	5	No
ADP-ribosylation factor 1	P84077	ARF1_HUMAN	20.0	4.3	No
Ubiquitin carboxyl-terminal hydrolase isozyme L1	P09936	UCHL1_HUMAN	20.0	3.8	No
Mitochondrial import receptor subunit TOM40 homolog	O96008	TOM40_HUMAN	20.0	3.1	No
Dolichyl-diphosphooligosaccharide--protein subunit 1	P04843	RPN1_HUMAN	20.0	3.1	No
Very-long-chain (3R)-3-hydroxyacyl-CoA dehydratase	H3BPZ1	H3BPZ1_HUMAN	20.0	2.4	No
Thioredoxin-related transmembrane protein 1	Q9H3N1	TMX1_HUMAN	20.0	3.9	No
Protein transport protein Sec61 subunit alpha isoform 1	P61619	S61A1_HUMAN	20.0	4.3	No
Nucleophosmin	P06748	NPM_HUMAN	20.0	3.8	No
Aldehyde dehydrogenase mitochondrial	P05091	ALDH2_HUMAN	20.0	5.1	No
Glutathione S-transferase P	P09211	GSTP1_HUMAN	20.0	4.2	No
40S ribosomal protein S16	P62249	RS16_HUMAN	20.0	3.5	No
Rho-related GTP-binding protein RhoC	P08134	RHOC_HUMAN	20.0	4.8	No
60S ribosomal protein L11	P62913	RL11_HUMAN	20.0	3.7	No
Phosphoglycerate kinase 1	P00558	PGK1_HUMAN	20.0	6.5	No
ATP-dependent 6-phosphofructokinase platelet type	Q01813	PFKAP_HUMAN	20.0	3.4	No
60S acidic ribosomal protein P0	P05388	RLA0_HUMAN	20.0	3.3	No
Malate dehydrogenase mitochondrial	P40926	MDHM_HUMAN	20.0	5	No
PRA1 family protein 2	O60831	PRAF2_HUMAN	20.0	2.9	No
Keratin type I cuticular Ha1	Q15323	K1H1_HUMAN	20.0	8.3	No
Tetraspanin	P48509	CD151_HUMAN	20.0	2.3	No
Leucine-rich repeat-containing protein 59	Q96AG4	LRC59_HUMAN	20.0	5.2	No
Ezrin	P15311	EZRI_HUMAN	20.0	2.4	No
Estradiol 17-beta-dehydrogenase 11	Q8NBQ5	DHB11_HUMAN	20.0	3.6	No
Malate dehydrogenase cytoplasmic	P40925	MDHC_HUMAN	20.0	4.6	No
Trifunctional enzyme subunit alpha mitochondrial	P40939	ECHA_HUMAN	20.0	4.6	No
Carnitine O-palmitoyltransferase 1 liver isoform	P50416	CPT1A_HUMAN	20.0	1.6	No
Stress-70 protein mitochondrial	P38646	GRP75_HUMAN	20.0	8.4	No
Minor histocompatibility antigen H13	Q8TCT9	HM13_HUMAN	20.0	2.8	No
Synaptogyrin-2	O43760	SNG2_HUMAN	20.0	1.5	No
Ras-related C3 botulinum toxin substrate 2	P15153	RAC2_HUMAN	20.0	6.1	No
Filamin-A	P21333	FLNA_HUMAN	20.0	5	No
60S ribosomal protein L5	P46777	RL5_HUMAN	20.0	3.1	No
Isocitrate dehydrogenase [NADP] cytoplasmic	O75874	IDHC_HUMAN	20.0	6.6	No
Transgelin-2	P37802	TAGL2_HUMAN	20.0	3	No
Heterogeneous nuclear ribonucleoprotein K	P61978	HNRPK_HUMAN	20.0	4.6	No
Destrin	P60981	DEST_HUMAN	20.0	4.1	No
40S ribosomal protein SA	P08865	RSSA_HUMAN	20.0	5.4	No
Tetraspanin	P21926	CD9_HUMAN	20.0	5.1	No
Keratin-like protein KRT222	Q8N1A0	KT222_HUMAN	20.0	2.2	No
Heterogeneous nuclear ribonucleoprotein A1	P09651	ROA1_HUMAN	20.0	4.9	No
Nucleolin	P19338	NUCL_HUMAN	20.0	7.3	No
60S ribosomal protein L12	P30050	RL12_HUMAN	20.0	3.8	No
Importin subunit beta-1	Q14974	IMB1_HUMAN	20.0	3.9	No
Bifunctional purine biosynthesis protein PURH	P31939	PUR9_HUMAN	20.0	4.7	No
Thioredoxin-related transmembrane protein 2	Q9Y320	TMX2_HUMAN	20.0	4.5	No
Dehydrogenase/reductase SDR family member 7	Q9Y394	DHRS7_HUMAN	20.0	1.5	No
Thioredoxin	P10599	THIO_HUMAN	20.0	3.9	No
Glutathione reductase mitochondrial	P00390	GSHR_HUMAN	20.0	4.4	No
Ras-related protein Rab-11A	P62491	RB11A_HUMAN	20.0	5.1	No
Adenosylhomocysteinase	P23526	SAHH_HUMAN	20.0	2.9	No

SYNTHESIS AND BIOLOGICAL EVALUATION OF CLICKABLE VITAMIN A

Name	Accession	Uniprot ID	Fold change (UV/EtOH)	Significance (-log ₁₀)	UV enriched
B-cell receptor-associated protein 31	P51572	BAP31_HUMAN	20.0	7.6	No
ADP-ribosylation factor 4	P18085	ARF4_HUMAN	20.0	1.6	No
Isoform 4 of Filamin-B	O75369-4	FLNB_HUMAN	20.0	3.4	No
Glutathione S-transferase omega-1	P78417	GSTO1_HUMAN	20.0	6.2	No
Leucine-rich PPR motif-containing protein_mitochondrial	P42704	LPPRC_HUMAN	20.0	7.1	No
Stress-induced-phosphoprotein 1	P31948	STIP1_HUMAN	20.0	8	No
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 2	P04844	RPN2_HUMAN	20.0	4.5	No
Elongation factor 1-gamma	P26641	EF1G_HUMAN	20.0	4.8	No
Glycogen phosphorylase brain form	P11216	PYGB_HUMAN	20.0	4.5	No
Long-chain-fatty-acid--CoA ligase 3	O95573	ACSL3_HUMAN	20.0	1.4	No
Preylcysteine oxidase 1	Q9UHG3	PCYOX_HUMAN	20.0	2.6	No
Lamin-B receptor	Q14739	LBR_HUMAN	20.0	4	No
Sodium/potassium-transporting ATPase subunit alpha-1	P05023	AT1A1_HUMAN	20.0	1.5	No
60S ribosomal protein L26	P61254	RL26_HUMAN	20.0	4	No
Protein disulfide-isomerase A3	P30101	PDI3_HUMAN	20.0	4.6	No
Very-long-chain enoyl-CoA reductase	Q9NZ01	TECR_HUMAN	20.0	3.8	No
Myosin-9	P35579	MYH9_HUMAN	20.0	3	No
Alpha-1 3-glucosyltransferase	Q9BVK2	ALG8_HUMAN	20.0	1.5	No
Vesicular integral-membrane protein VIP36	Q12907	LMAN2_HUMAN	20.0	1.6	No
Lysophospholipid acyltransferase 7	Q96N66	MBOA7_HUMAN	20.0	4.5	No
ATP synthase membrane subunit g	O75964	ATP5L_HUMAN	20.0	1.6	No
Tricarboxylate transport protein_mitochondrial	P53007	TXTP_HUMAN	20.0	1.6	No
Mitochondrial inner membrane protein OXA1L	Q15070	OXA1L_HUMAN	20.0	1.5	No
T-complex protein 1 subunit theta	P50990	TCPO_HUMAN	20.0	1.6	No
Adipocyte plasma membrane-associated protein	Q9HDC9	APMAP_HUMAN	20.0	3.1	No
Protein disulfide-isomerase A4	P13667	PDI4A_HUMAN	20.0	5.6	No
Elongation factor Tu_mitochondrial	P49411	EFTU_HUMAN	20.0	5.8	No
ATP-binding cassette sub-family D member 3	P28288	ABCD3_HUMAN	20.0	6.4	No
Aspartate aminotransferase_mitochondrial	P00505	AATM_HUMAN	20.0	5.6	No
Chloride intracellular channel protein 1	O00299	CLIC1_HUMAN	20.0	1.6	No
Serum albumin	P02768	ALBU_HUMAN	20.0	1.4	No
Mitochondrial import receptor subunit TOM20 homolog	Q15388	TOM20_HUMAN	20.0	1.6	No
Splicing factor proline- and glutamine-rich	P23246	SFPQ_HUMAN	20.0	3.1	No
T-complex protein 1 subunit beta	P78371	TCPB_HUMAN	20	6.2	No
Endoplasmic reticulum metalloproteinase 1	Q7Z2K6	ERMP1_HUMAN	20	1.6	No
Cytoskeleton-associated protein 4	Q07065	CKAP4_HUMAN	20	1.6	No
Elongation of very long chain fatty acids protein 1	Q9BW60	ELOV1_HUMAN	20	1.6	No
Filamin-A	P21333	FLNA_HUMAN	20	6.3	No
X-ray repair cross-complementing protein 6	P12956	XRCC6_HUMAN	20	3.8	No
T-complex protein 1 subunit delta	P50991	TCPD_HUMAN	20	4.6	No
Spectrin beta chain non-erythrocytic 1	Q01082	SPTB2_HUMAN	20	1.5	No
C-1-tetrahydrofolate synthase_cytoplasmic	P11586	C1TC_HUMAN	20	2.3	No
ATP-dependent RNA helicase A	Q08211	DHX9_HUMAN	20	4	No
ADP/ATP translocase 2	P05141	ADT2_HUMAN	11.3	3.5	No
Voltage-dependent anion-selective channel protein 2	P45880	VDAC2_HUMAN	9.0	3	No
Aldo-keto reductase family 1 member C3	P42330	AK1C3_HUMAN	7.1	4	No
Keratin type II cuticular Hb1	Q14533	KRT81_HUMAN	7.0	4.2	No
Heat shock protein HSP 90-beta	P08238	HS90B_HUMAN	5.9	3.7	No
Keratin type II cytoskeletal 8	P05787	K2C8_HUMAN	5.8	3.7	No
Glyceraldehyde-3-phosphate dehydrogenase	P04406	G3P_HUMAN	5.7	3.9	No
Aldo-keto reductase family 1 member B10	O60218	AK1BA_HUMAN	5.6	4.9	No
Aldo-keto reductase family 1 member C2	P52895	AK1C2_HUMAN	5.6	3.7	No
Phosphate carrier protein_mitochondrial	Q00325	MPCP_HUMAN	5.6	3.3	No
Retinal dehydrogenase 1	P00352	AL1A1_HUMAN	5.0	4.8	No
Voltage-dependent anion-selective channel protein 1	P21796	VDAC1_HUMAN	4.9	2.3	No
Annexin A1	P04083	ANXA1_HUMAN	4.7	3.7	No
Large neutral amino acids transporter small subunit 1	Q01650	LAT1_HUMAN	4.6	3.1	No
Keratin type I cytoskeletal 18	P05783	K1C18_HUMAN	4.4	3	No
L-lactate dehydrogenase A chain	P00338	LDHA_HUMAN	4.1	3.2	No
L-lactate dehydrogenase B chain	P07195	LDHB_HUMAN	4.0	3.4	No
Elongation factor 1-alpha 1	P68104	EF1A1_HUMAN	4.0	3.2	No
Peroxisomal dehydrogenase [quinone] 1	Q06830	PRDX1_HUMAN	3.6	3.2	No
NAD(P)H dehydrogenase [quinone] 1	P15559	NQO1_HUMAN	3.5	3	No
Fructose-bisphosphate aldolase A	P04075	ALDOA_HUMAN	3.4	3.5	No
Tubulin alpha-1A chain	Q71U36	TBA1A_HUMAN	3.3	3.1	No
Pyruvate kinase PKM	P14618	KPYM_HUMAN	3.2	4.2	No
Alpha-actinin-4	O43707	ACTN4_HUMAN	3.2	2.5	No
Elongation factor 2	P13639	EF2_HUMAN	3.2	3.1	No
6-phosphogluconate dehydrogenase decarboxylating	P52209	6PGD_HUMAN	3.1	2.5	No
Tubulin beta chain	P04350	TBB4A_HUMAN	3.1	3.1	No
Aldehyde dehydrogenase dimeric NADP-preferring	P30838	AL3A1_HUMAN	3.0	3.5	No
Alpha-enolase	P06733	ENOA_HUMAN	2.9	3	No
Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	P16615	AT2A2_HUMAN	2.9	3.1	No
Profilin-1	P07737	PROF1_HUMAN	2.8	3.4	No
Endoplasmic reticulum chaperone	P14625	ENPL_HUMAN	2.8	3.2	No
ADP/ATP translocase 3	P12236	ADT3_HUMAN	2.7	2.4	No
Protein S100	P06703	S10A6_HUMAN	2.7	3	No
Heat shock protein HSP 90-alpha	P07900	HS90A_HUMAN	2.5	2.2	No
Histone H1.2	P16403	H12_HUMAN	2.5	3	No
Actin cytoplasmic 1	P60709	ACTB_HUMAN	2.4	2.2	No
Polyubiquitin-B	P0CG47	UBB_HUMAN	2.4	2.7	No
Transketolase	P29401	TKT_HUMAN	2.4	1.8	No
Putative heat shock protein HSP 90-beta-3	Q58FF7	H90B3_HUMAN	2.4	2.5	No
60S ribosomal protein L27a	P46776	RL27A_HUMAN	2.3	2.6	No
Cofilin 1 (Non-muscle) isoform CRA a	G3V1A4	G3V1A4_HUMAN	2.3	2.3	No
60 kDa heat shock protein_mitochondrial	P10809	CH60_HUMAN	2.3	2.8	No
60S ribosomal protein L6	Q02878	RL6_HUMAN	2.3	2.3	No
Glucose-6-phosphate isomerase	Q92643	GPI8_HUMAN	2.2	2.1	No
Heterogeneous nuclear ribonucleoproteins A2/B1	P22626	ROA2_HUMAN	2.2	2.5	No
60S ribosomal protein L7a	P62424	RL7A_HUMAN	2.2	2.4	No
14-3-3 protein zeta/delta	P63104	1433Z_HUMAN	2.1	2.2	No
Nicotinamide phosphoribosyltransferase	P43490	NAMPT_HUMAN	2.1	2.6	No
40S ribosomal protein S8	P62241	RS8_HUMAN	2.1	2.2	No
ATP-citrate synthase	P53396	ACLY_HUMAN	2.1	2.8	No

Supplementary Table 7.4 | Proteins significantly enriched by STA-215 (10 μ M) in A549 cells corresponding to Fig. 7.5d.

Name	Accession	Uniprot ID	Fold-change (UV/EtOH)	Significance (-log ₁₀)	UV enriched
Endoplasmic	P14625	ENPL_HUMAN	20.0	2.7	Yes
Voltage-dependent anion-selective channel protein 1	P21796	VDAC1_HUMAN	20.0	2.9	Yes
Vesicular integral-membrane protein VIP36	Q12907	LMAN2_HUMAN	20.0	3.7	Yes
Neutral cholesterol ester hydrolase 1	Q6PIU2	NCEH1_HUMAN	20.0	4.0	Yes
Prohibitin-2	Q99623	PHB2_HUMAN	20.0	2.7	Yes
Adipocyte plasma membrane-associated protein	Q9HDC9	APMAP_HUMAN	20.0	3.4	Yes
Mitochondrial import inner membrane translocase subunit TIM50	Q3ZCQ8	TIM50_HUMAN	20.0	1.5	Yes
Voltage-dependent anion-selective channel protein 3	Q9Y277	VDAC3_HUMAN	20.0	1.7	Yes
ATPase ASNA1	Q43681	ASNA_HUMAN	20.0	4.0	Yes
Cleft lip and palate transmembrane protein 1	O96005	CLPT1_HUMAN	20.0	3.6	Yes
Leucine-rich repeat-containing protein 59	Q96AG4	LRC59_HUMAN	20.0	3.6	Yes
NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 10_mitochondrial	O95299	NDUAA_HUMAN	20.0	2.0	Yes
ADP/ATP translocase 2	P05141	ADT2_HUMAN	17.3	2.2	Yes
Phosphate carrier protein_mitochondrial	Q00325	MPCP_HUMAN	12.7	3.3	Yes
Aldo-keto reductase family 1 member C2	P52895	AK1C2_HUMAN	10.3	3.8	Yes
ADP/ATP translocase 1	P12235	ADT1_HUMAN	10.0	1.9	Yes
Epoxide hydrolase 1	P07099	HYEP_HUMAN	8.8	4.1	Yes
Voltage-dependent anion-selective channel protein 2	P45880	VDAC2_HUMAN	8.0	2.4	Yes
Transgelin-2	P37802	TAGL2_HUMAN	20.0	11.2	No
Protein S100	P06703	S10A6_HUMAN	20.0	4.6	No
Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	P16615	AT2A2_HUMAN	20.0	3.1	No
Serum albumin	P02768	ALBU_HUMAN	20.0	4.3	No
Rab GDP dissociation inhibitor beta	P50395	GDIB_HUMAN	20.0	5.0	No
Isocitrate dehydrogenase [NADP] cytoplasmic	O75874	IDHC_HUMAN	20.0	3.8	No
Keratin_type II cytoskeletal 7	P08729	K2C7_HUMAN	20.0	4.0	No
Chloride intracellular channel protein 1	O00299	CLIC1_HUMAN	20.0	3.3	No
Elongation factor 1-gamma	P26641	EF1G_HUMAN	20.0	5.2	No
40S ribosomal protein SA	P08865	RSSA_HUMAN	20.0	4.8	No
Stress-70 protein_mitochondrial	P38646	GRP75_HUMAN	20.0	4.7	No
Phosphoglycerate kinase 1	P00558	PGK1_HUMAN	20.0	9.4	No
Isoform 4 of Filamin-B	O75369-4	FLNB_HUMAN	20.0	4.3	No
Transaldolase	P37837	TALDO_HUMAN	20.0	4.8	No
Protein disulfide-isomerase	P07237	PDIA1_HUMAN	20.0	5.4	No
Leucine-rich PPR motif-containing protein_mitochondrial	P42704	LPPRC_HUMAN	20.0	6.8	No
Ubiquitin carboxyl-terminal hydrolase	P09936	UCHL1_HUMAN	20.0	3.2	No
Sideroflexin-1	Q9H9B4	SFXN1_HUMAN	20.0	3.1	No
15-hydroxyprostaglandin dehydrogenase [NAD(+)]	P15428	PGDH_HUMAN	20.0	7.8	No
Trifunctional enzyme subunit alpha_mitochondrial	P40939	ECHA_HUMAN	20.0	1.8	No
4F2 cell-surface antigen heavy chain	P08195	4F2_HUMAN	20.0	7.2	No
Putative keratin-87 protein	A6NCN2	KR87P_HUMAN	20.0	3.3	No
Protein disulfide-isomerase A4	P13667	PDIA4_HUMAN	20.0	6.0	No
78 kDa glucose-regulated protein	P11021	BIP_HUMAN	20.0	7.3	No
Large neutral amino acids transporter small subunit 1	Q01650	LAT1_HUMAN	8.5	2.9	No
Aldo-keto reductase family 1 member C3	P42330	AK1C3_HUMAN	7.7	2.7	No
Aldo-keto reductase family 1 member B10	O60218	AK1BA_HUMAN	7.0	2.8	No
14-3-3 protein zeta/delta	P63104	1433Z_HUMAN	5.5	2.6	No
Heat shock protein HSP 90-alpha	P07900	HS90A_HUMAN	4.6	2.9	No
Retinal dehydrogenase 1	P00352	AL1A1_HUMAN	4.5	2.6	No
Annexin A1	P04083	ANXA1_HUMAN	4.4	2.4	No
Elongation factor 1-alpha 1	P68104	EF1A1_HUMAN	4.0	2.6	No
Polyubiquitin-B	P0CG47	UBB_HUMAN	4.0	2.3	No
Heat shock cognate 71 kDa protein	P11142	HSP7C_HUMAN	3.6	2.5	No
L-lactate dehydrogenase B chain	P07195	LDHB_HUMAN	3.5	1.9	No
UDP-glucose 6-dehydrogenase	O60701	UGDH_HUMAN	3.5	2.4	No
Annexin A2	P07355	ANXA2_HUMAN	3.5	2.6	No
Heat shock protein HSP 90-beta	P08238	HS90B_HUMAN	3.3	2.6	No
L-lactate dehydrogenase A chain	P00338	LDHA_HUMAN	3.3	2.0	No
Vimentin	P08670	VIME_HUMAN	3.3	2.5	No
Alpha-enolase	P06733	ENOA_HUMAN	3.3	1.9	No
14-3-3 protein theta	P27348	1433T_HUMAN	3.2	2.5	No
Keratin_type I cytoskeletal 18	P05783	K1C18_HUMAN	3.2	1.7	No
Aldose reductase	P15121	ALDR_HUMAN	3.1	1.7	No
Heat shock 70 kDa protein 1A	P0DMV8	HS71A_HUMAN	3.0	2.3	No
Pyruvate kinase PKM	P14618	KPYM_HUMAN	2.9	1.7	No
Glucose-6-phosphate 1-dehydrogenase	P11413	G6PD_HUMAN	2.9	1.8	No
Fructose-bisphosphate aldolase A	P04075	ALDOA_HUMAN	2.8	1.6	No
Elongation factor 2	P13639	EF2_HUMAN	2.8	2.1	No
Keratin_type II cytoskeletal 8	P05787	K2C8_HUMAN	2.7	2.0	No
Triosephosphate isomerase	P60174	TPIS_HUMAN	2.7	2.1	No
Actin_cytosolic 1	P60709	ACTB_HUMAN	2.6	1.6	No
Aldehyde dehydrogenase_dimeric NADP-preferring	P30838	AL3A1_HUMAN	2.5	1.8	No
Filamin-A	P21333	FLNA_HUMAN	2.5	2.4	No
Tubulin beta chain	P04350	TBB4A_HUMAN	2.5	1.5	No
Glutathione reductase_mitochondrial	P00390	GSHR_HUMAN	2.5	2.3	No
Transketolase	P29401	TKT_HUMAN	2.4	2.1	No
Nicotinamide phosphoribosyltransferase	P43490	NAMPT_HUMAN	2.4	1.9	No
Profilin-1	P07737	PROF1_HUMAN	2.3	1.8	No
Tubulin alpha-1C chain	Q9BQE3	TBA1C_HUMAN	2.1	1.9	No
Glutathione S-transferase P	P09211	GSTP1_HUMAN	2.1	2.0	No

References

1. Haberkant, P. & Holthuis, J. C. M. Fat & fabulous: Bifunctional lipids in the spotlight. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1841**, 1022–1030 (2014).
2. Thiele, C. *et al.* Tracing fatty acid metabolism by click chemistry. *ACS Chem. Biol.* **7**, 2004–2011 (2012).
3. Robichaud, P. P. *et al.* On the cellular metabolism of the click chemistry probe 19-alkyne arachidonic acid. *J. Lipid Res.* **57**, 1821–1830 (2016).
4. Gaebler, A. *et al.* Alkyne lipids as substrates for click chemistry-based in vitro enzymatic assays. *J. Lipid Res.* **54**, 2282–2290 (2013).
5. Hannoush, R. N. & Arenas-Ramirez, N. Imaging the lipidome: ω -alkynyl fatty acids for detection and cellular visualization of lipid-modified proteins. *ACS Chem. Biol.* **4**, 581–587 (2009).
6. Heal, W. P., Wickramasinghe, S. R., Leatherbarrow, R. J. & Tate, E. W. N-Myristoyl transferase-mediated protein labelling in vivo. *Org. Biomol. Chem.* **6**, 2308–2315 (2008).
7. Hang, H. C. *et al.* Chemical probes for the rapid detection of fatty-acylated proteins in mammalian cells. *J. Am. Chem. Soc.* **129**, 2744–2745 (2007).
8. Koenders, S. T. A. *et al.* Development of a Retinal-Based Probe for the Profiling of Retinaldehyde Dehydrogenases in Cancer Cells. *ACS Cent. Sci.* **5**, 1965–1974 (2019).
9. Budhu, A. S. & Noy, N. Direct channeling of retinoic acid between cellular retinoic acid-binding protein II and retinoic acid receptor sensitizes mammary carcinoma cells to retinoic acid-induced growth arrest. *Mol. Cell. Biol.* **22**, 2632–41 (2002).
10. Erkelens, M. N. & Mebius, R. E. Retinoic Acid and Immune Homeostasis: A Balancing Act. *Trends in Immunology* **38**, 168–180 (2017).
11. Kleiner, P., Heydenreuter, W., Stahl, M., Korotkov, V. S. & Sieber, S. A. A Whole Proteome Inventory of Background Photocrosslinker Binding. *Angew. Chemie Int. Ed.* **56**, 1396–1401 (2017).
12. Murayama, A., Suzuki, T. & Matsui, M. Photoisomerization of retinoic acids in ethanol under room light: A warning for cell biological study of geometrical isomers of retinoids. *J. Nutr. Sci. Vitaminol. (Tokyo)*. **43**, 167–176 (1997).
13. Kiser, P. D., Golczak, M. & Palczewski, K. Chemistry of the retinoid (visual) cycle. *Chem. Rev.* **114**, 194–232 (2014).
14. Dillon, J., Gaillard, E. R., Bilski, P., Chignell, C. F. & Reszka, K. J. The photochemistry of the retinoids as studied by steady-state and pulsed methods. *Photochem. Photobiol.* **63**, 680–685 (1996).
15. Lo, K. K. N., Land, E. J. & TRUSCOTT, T. G. PRIMARY INTERMEDIATES IN THE PULSED IRRADIATION OF RETINOIDS. *Photochem. Photobiol.* **36**, 139–145 (1982).
16. Bayley, H. & Knowles, J. R. Photoaffinity labeling. *Methods Enzymol.* **46**, 69–114 (1977).
17. Bernstein, P. S., Choi, S. Y., Ho, Y. C. & Rando, R. R. Photoaffinity labeling of retinoic acid-binding proteins. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 654–658 (1995).
18. Chen, G. & Radomska-Pandya, A. Direct photoaffinity labeling of cellular retinoic acid-binding protein I (CRABP-I) with all-trans-retinoic acid: Identification of amino acids in the ligand binding site. *Biochemistry* **39**, 12568–12574 (2000).
19. Radomska-Pandya, A. & Chen, G. Photoaffinity labeling of human retinoid X receptor β (RXR β) with 9-cis-retinoic acid: Identification of phytanic acid, docosahexaenoic acid, and lithocholic acid as ligands for RXR β . *Biochemistry* **41**, 4883–4890 (2002).
20. Lambertin, F., Wende, M., Quirin, M. J., Taran, M. & Delmond, B. New retinoid analogs from δ -pyronene, a natural synthon. *European J. Org. Chem.* **1999**, 1489–1494 (1999).
21. Birnie, G. D. The HL60 cell line: a model system for studying human myeloid cell differentiation. *Br. J. Cancer. Suppl.* **9**, 41–5 (1988).
22. Miller, L. J., Bainton, D. F., Borregaard, N. & Springer, T. A. Stimulated mobilization of monocyte Mac-1 and p150,95 adhesion proteins from an intracellular vesicular compartment to the cell surface. *J. Clin. Invest.* **80**, 535–544 (1987).
23. Napoli, J. L. Cellular retinoid binding-proteins, CRBP, CRABP, FABP5: Effects on retinoid metabolism, function and related diseases. *Pharmacology and Therapeutics* **173**, 19–33 (2017).
24. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
25. Ogata, H., Goto, S., Fujibuchi, W. & Kanehisa, M. Computation with the KEGG pathway database. *Biosystems* **47**, 119–128 (1998).

26. Thomas, P. D. *et al.* PANTHER: A browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Research* **31**, 334–341 (2003).
27. McGrath, K. E., Bushnell, T. P. & Palis, J. Multispectral imaging of hematopoietic cells: Where flow meets morphology. *J. Immunol. Methods* **336**, 91–97 (2008).
28. Barua, A. B. & Furr, H. C. Properties of retinoids. Structure, handling, and preparation. *Applied Biochemistry and Biotechnology - Part B Molecular Biotechnology* **10**, 167–182 (1998).
29. Babler, J. H., Coghlan, M. J., Feng, M. & Fries, P. Facile Synthesis of 4-Acetoxy-2-methyl-2-butenal, a Vitamin A Precursor, from Isoprene Chlorohydrin. *J. Org. Chem.* **44**, 1716–1717 (1979).
30. McBee, J. K., Van Hooser, J. P., Jang, G. F. & Palczewski, K. Isomerization of 11-cis-Retinoids to All-trans-retinoids in Vitro and in Vivo. *J. Biol. Chem.* **276**, 48483–48493 (2001).
31. Kojima, R. *et al.* In vivo isomerization of retinoic acids. Rapid isomer exchange and gene expression. *J. Biol. Chem.* **269**, 32700–32707 (1994).
32. 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).
33. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
34. Distler, U., Kuharev, J., Navarro, P. & Tenzer, S. Label-free quantification in ion mobility-enhanced data-independent acquisition proteomics. *Nat. Protoc.* **11**, 795–812 (2016).
35. Distler, U. *et al.* Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics. *Nat. Methods* **11**, 167–170 (2014).
36. Kuharev, J., Navarro, P., Distler, U., Jahn, O. & Tenzer, S. In-depth evaluation of software tools for data-independent acquisition based label-free quantification. *Proteomics* **15**, 3140–3151 (2015).

Chapter 8

Design, Synthesis and Biological Evaluation of a Broadspectrum ALDH Probe

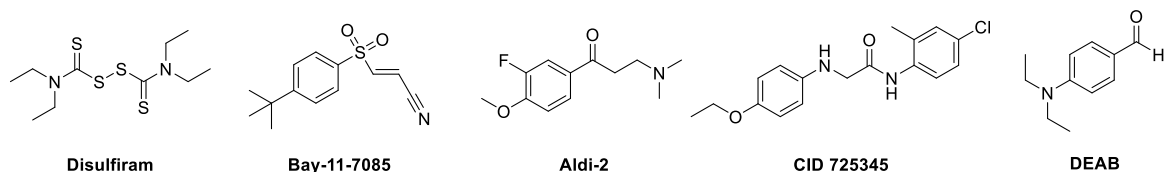
Published as S.T.A. Koenders *et al.*, *ChemBioChem*, **21**, 1911-1917 (2020).

Introduction

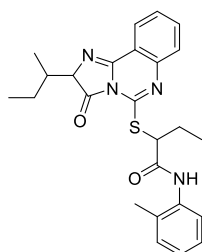
Aldehyde dehydrogenases (ALDHs) are involved in many metabolic pathways and pathologies.¹⁻⁷ They are often upregulated in cancer and have been linked to cancer therapy resistance. For example, both ALDH1A1 and ALDH3A1 induce resistance to the commonly used chemotherapeutic, cyclophosphamide.^{8,9}

A recent study has shown that only 7 out of 20 commonly used or recently developed ALDH1A1 inhibitors showed cellular activity (**Fig. 8.1**).¹⁰ Two compounds (compound 5 and NCT-501) were selective for ALDH1A1, whereas five were pan-ALDH inhibitors (Disulfiram, Bay-11-7085, Aldi-2, CID 725345 and DEAB).¹⁰ CB7 was claimed to be a selective ALDH3A1 inhibitor.¹⁰

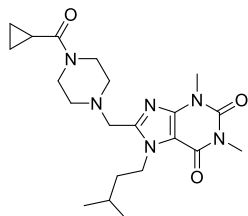
Pan-ALDH inhibitors



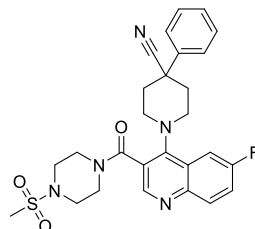
ALDH1A1 inhibitors



Compound 5

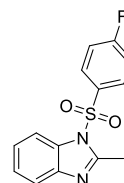


NCT-501



NCT-505

ALDH3A1 inhibitor



CB7

Fig. 8.1 | Chemical structures of *in situ* active ALDH inhibitors

Although *in vitro* assays are useful for the rapid screening of large compound libraries, the biochemical ALDH assays have two main drawbacks. First, the conditions in the biochemical assays differ from the cellular assays, which result in altered potency and selectivity profiles. For example, most biochemical assays are performed at pH > 7.4 and reducing agents, such as dithiothreitol (DTT) or tris(2-carboxyethyl)phosphine (TCEP), are added. This will change the reactivity of catalytic cysteine. Second, the biochemical assays do not take into account factors, such as cell permeability and metabolic stability.

So far, no methods exists to study the target-engagement of multiple ALDHs present in a biological sample simultaneously. Activity-based protein profiling (ABPP) is a powerful technique capable of determining the selectivity profile of drug candidate against an enzyme family in their native cellular environment.¹¹ This technique relies on activity-based probes, which are tailored to the enzyme family of interest and that react via its electrophilic warhead with a nucleophile within the active site of the enzyme. **Chapters 3-6** described the development of a first-in-class retinal-based probe **LEI-945** for the profiling of retinaldehyde dehydrogenases.¹² With **LEI-945**, six of the 19 enzymes in the ALDH family in humans, ALDH1A1, ALDH1A2, ALDH1A3, ALDH2, ALDH3A2 and ALDH3B1 can be detected from cell extract. The synthesis of **LEI-945**, which can be found in **Chapter 3**, is complex and challenging. It was envisioned that by modifying the reported covalent pan-ALDH inhibitor, Aldi-2²⁰ (**Fig. 8.1**), an easily accessible broadspectrum probe for the ALDH family could be made.

This chapter describes the design, synthesis, biological validation and application of **STA-55** as a broadspectrum probe for the family of aldehyde dehydrogenases. Chemical proteomics showed that **STA-55** could be used to enrich both ALDH1A1 and ALDH3A1 in A549 lung cancer cells. Competitive proteomics with **STA-55** was performed to determine the selectivity profiles of known ALDH inhibitors DEAB¹³, NCT-505¹⁴ and CB7.¹⁵ These results showed that **STA-55** can be used to identify therapy resistance biomarkers in cancer and to validate target-engagement of ALDH drug candidates.

Results and discussion

Synthesis and biological evaluation of a broadspectrum activity-based ALDH probe

An activity-based probe (ABP) consists of a reactive group (termed ‘warhead’ and often an electrophile), a recognition handle and a ligation handle. Aldi-2 already incorporates a masked warhead which after liberation reacts with the catalytic cysteine of the ALDH enzyme (**Fig. 8.2**).¹⁶ Based on the reported structure-activity relationship studies, a synthetic strategy for the introduction of an azide ligation handle at the position of the methoxy-substituent was developed.

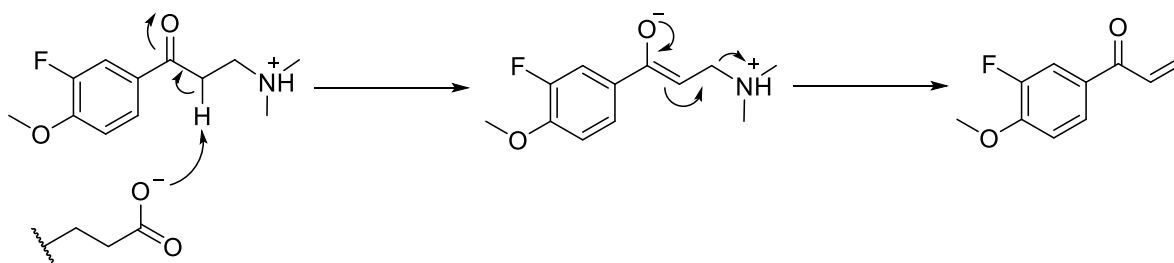
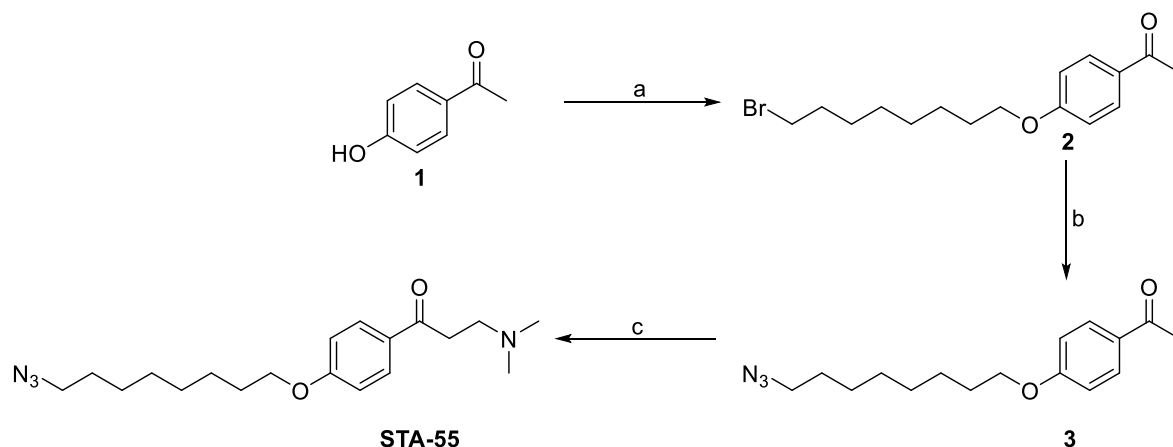


Fig. 8.2 | Proposed Aldi-2 warhead deprotection mechanism

The synthesis of probe **STA-55** started from commercially available 4-hydroxyacetophenone **1** (**Scheme 8.1**). Reaction of **1** with 1,8-dibromooctane and potassium carbonate provided **2** in 82% yield. Treatment of **2** with sodium azide led to the substitution of the bromine for an azide, yielding compound **3** in 93% yield. Finally, Mannich reaction of **3** with dimethylamine and paraformaldehyde gave tertiary amine **STA-55** in 21% yield.



Scheme 8.1 | Synthesis of probe STA-55. Reagent and conditions: a) 1,8-dibromooctane, K_2CO_3 , acetone, 56 °C, 18 h, 82%; b) NaN_3 , DMF, 80 °C, 18 h, 93%; c) $(CH_3)_2NH \cdot HCl$, paraformaldehyde, HCl, EtOH, 90 °C, 18 h, 21%.

To determine whether **STA-55** interacts with the catalytic cysteine of an ALDH enzyme, recombinant ALDH1A3^{WT} and catalytically inactive ALDH1A3^{C314A} were overexpressed in human osteosarcoma U2OS cells. Treatment with **STA-55** (1 μ M, 1 h), lysis and Cu(I)-catalysed azide-alkyne [2+3] cycloaddition with a Cy5-alkyne resulted in a fluorescent band around 55 kDa after resolving the samples by SDS-PAGE and fluorescent scanning, suggesting the ability of **STA-55** to interact with ALDHs. The disappearance of this band in the catalytically inactive C314A mutant indicates that **STA-55** has reacted covalently with the catalytic nucleophile of ALDH1A3 (**Supplementary Fig. 8.1**).¹⁷

To determine which members of the ALDH family are targeted by ABP **STA-55**, a label-free quantitative proteomics experiment was performed in the non-small-cell lung cancer cell line A549, which expresses high levels of ALDH activity.^{18,19} **STA-55**-treated A549 cells (10 μ M, 1 h) were harvested, lysed and the covalently bound enzymes conjugated with a biotin-alkyne. The probe labelled proteins were subsequently enriched using streptavidin-beads and several washing steps to remove unbound proteins. On-bead digestion was followed by protein identification and quantification by mass spectrometry. In this way, 259 were identified (fold-change > 2.0; p-value < 0.05) (**Fig. 8.3a, b**). Six identified proteins belong to the ALDH family: ALDH1A1, ALDH1B1, ALDH2, ALDH3A1, ALDH3A2 and ALDH3B1.

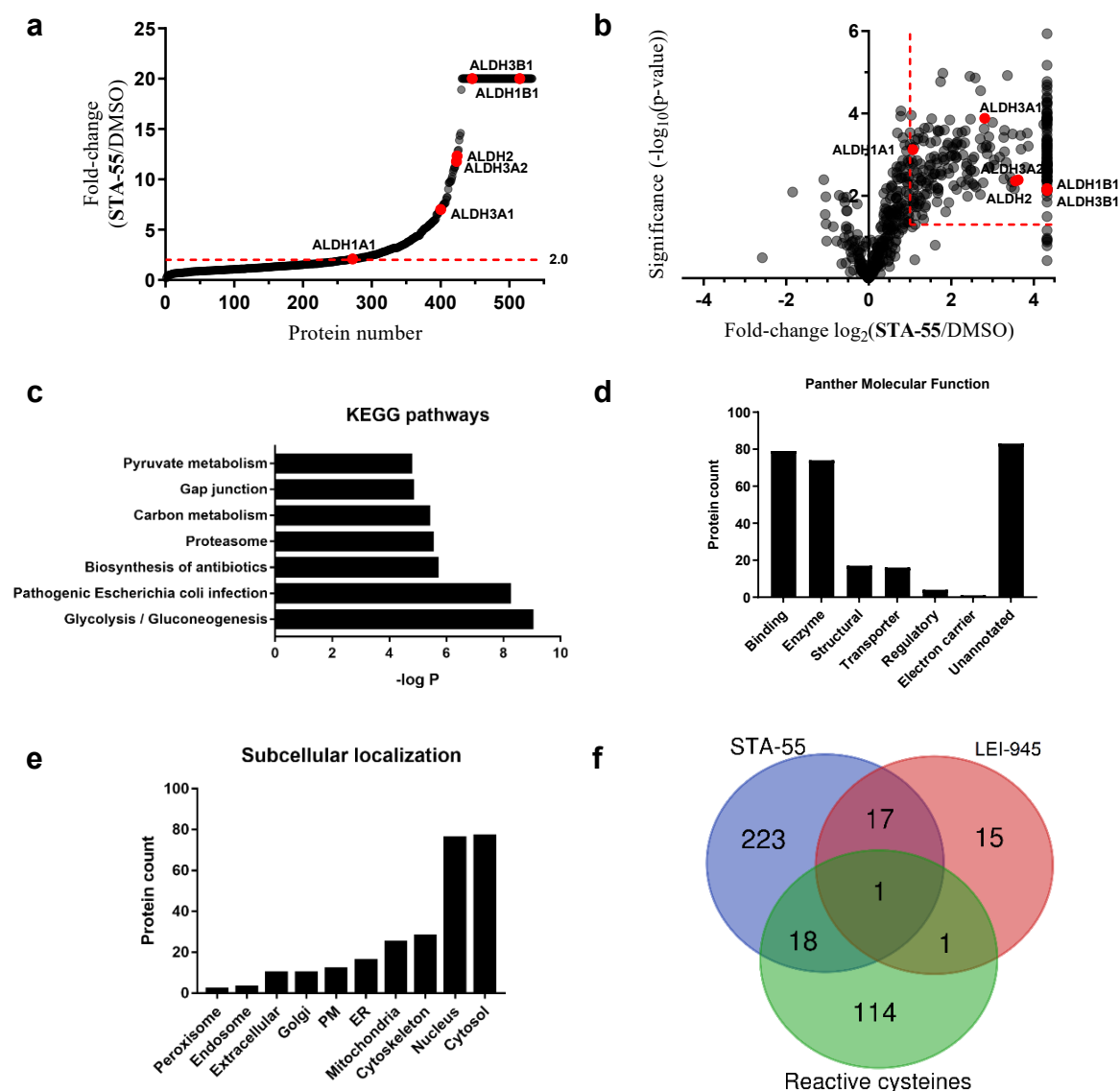


Fig. 8.3 | Chemical proteomics with broadspectrum ALDH probe STA-55. **a**, Fold-change (STA-55/DMSO) plot for total proteins identified in chemical proteomics experiment with probe STA-55 (10 μ M). Red lines indicate the threshold fold-change of 2-fold enrichment and the maximum fold-change is set at 20. Red dots represent significantly enriched ALDH enzymes. **b**, Volcano plot for total proteins identified in chemical proteomics experiment with probe STA-55 (10 μ M). Red lines indicate threshold values marking significantly enriched proteins. Red dots represent significantly enriched ALDH enzymes. For parts **a** and **b**, data are from $N = 3$ experiments (biological replicates). **c**, Top 7 pathways enriched in the group of significantly enriched proteins as determined by screening on the KEGG database. **d**, Biological functions annotated to significantly enriched proteins by the PANTHER database. **e**, Subcellular localization of significantly enriched proteins as annotated by the UniProt database. **f**, Venn diagram showing the distribution of proteins enriched by the broadspectrum STA-55 probe, retinoid probe LEI-945 and proteins containing highly reactive cysteines as determined by Weerapana *et al.*

Comparison with the Expression Atlas (**Supplementary Fig. 8.2**)^{20,21} showed that all the ALDH enzymes expressed in the A549 cell line were detected by the **STA-55** broadspectrum probe, whereas the more specific **LEI-945** probe did not detect ALDH1B1 and ALDH3A1 in this cell line. The proteins significantly enriched by **STA-55** were further analysed using the KEGG²², PANTHER²³ and UniProtKB databases.²⁴ Proteins in the glycolysis/gluconeogenesis, biosynthesis of antibiotics and carbon metabolism (ALDHs, lactate dehydrogenases and phosphofructokinases), pathogenic *E. Coli* infection and gap junction (tubulins) and the proteasome were identified via the KEGG pathway database (**Fig. 8.3c**).

The majority of proteins identified possess enzyme activity or have specific protein interaction partners (**Fig. 8.3d**). The subcellular localization showed that the bulk of the enriched proteins are derived from the cytosol and nucleus (**Fig. 8.3e**). Comparing the probe targets of **STA-55** and **LEI-945** shows that 18 proteins enriched by **LEI-945** are also detected by **STA-55** (**Fig. 8.3f**). Although Aldi-2 was claimed to be selective for ALDHs over other enzymes containing reactive cysteines, such as cysteine proteases, alcohol dehydrogenases and tyrosine phosphatases, **STA-55** did enrich 19 enzymes containing reactive cysteines, including ATPases, hydrolases, cysteine proteases, disulphide isomerases and kinases (**Fig. 8.3f**).¹⁶ Taken together, these data indicate that **STA-55** can be used as a broadspectrum activity-based probe for aldehyde dehydrogenases, including the known cancer resistance biomarkers, ALDH1A1 and ALDH3A1.^{8,9}

Competitive ABPP of ALDH inhibitors using STA-55

Having established that **STA-55** significantly enriches a broad range of ALDHs in the A549 lung cancer cell extracts, *in situ* selectivity profiling of three known ALDH inhibitors (DEAB^{10,13}, NCT-505¹⁴ and CB7¹⁵) was performed. DEAB is a pan-ALDH inhibitor regularly used as control compound in ALDH activity assays.^{10,13} NCT-505 is a recently reported selective ALDH1A1 inhibitor¹⁴ and CB7 is reported as a selective ALDH3A1 inhibitor.¹⁵

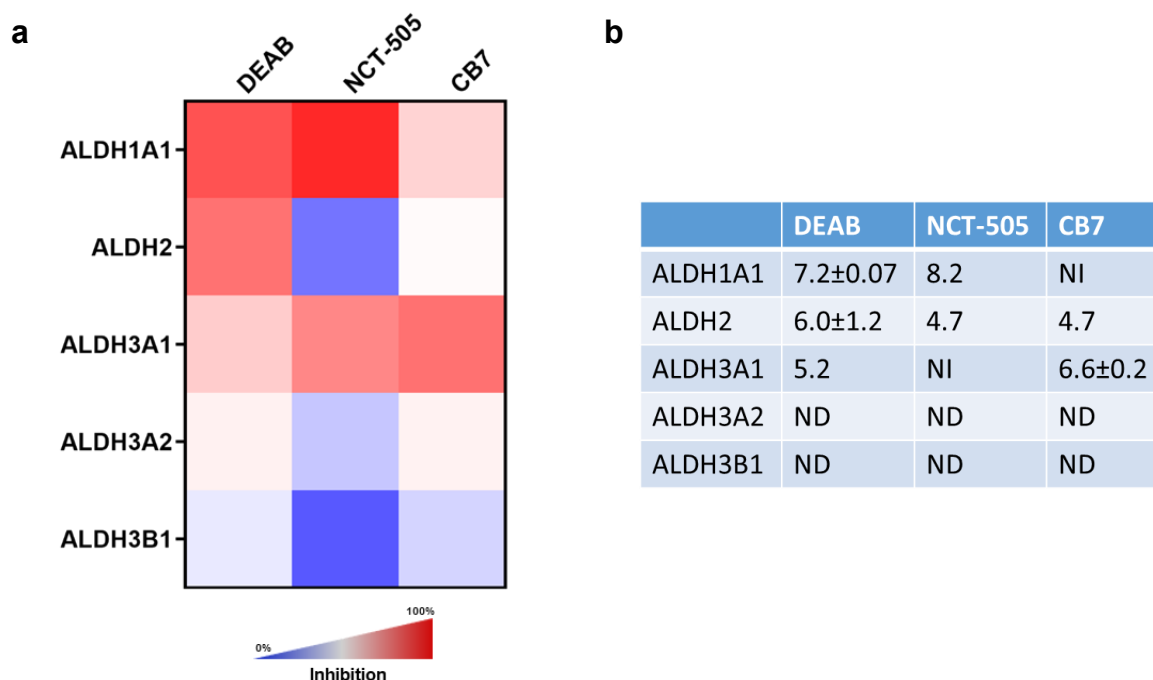


Fig. 8.4 | Competitive ABPP of STA-55 with ALDH inhibitors. **a**, Heatmap showing the selectivity profile of pan-ALDH inhibitor DEAB, ALDH1A1 selective inhibitor NCT-505 and ALDH3A1 selective inhibitor CB7 as determined by competitive ABPP with **STA-55** (1 μ M). $N = 4$ experiments (biological replicates). **b**, Reported pIC50 values of DEAB^{10,13}, NCT-505¹⁴ and CB7^{10,15} for the ALDHs identified using **STA-55**. When divergent values have been reported, data are represented as mean values $\pm$ SD. NI = no inhibition and ND = no data.

A549 cells were preincubated for 30 minutes with the inhibitor (10 μ M) after which **STA-55** (1 μ M) was added and incubated for one hour. After which the chemical proteomics protocol described before was followed. The ALDH selectivity profiles obtained are shown in **Fig. 8.4a**. Of note, ALDH1B1 was not detected in the competitive ABPP experiment, probably because a 10-fold lower concentration of **STA-55** was used and ALDH1B1, one of the lowest abundant ALDH enzymes in this cell line, did not reach the detection threshold (**Supplementary Fig. 8.3**).

DEAB inhibited ALDH1A1, ALDH2, ALDH3A1 and ALDH3A2, which is in agreement with previously reported results (**Fig. 8.4b**).¹³ NCT-505 inhibited ALDH1A1, ALDH1A3 and ALDH3A1. CB7 appeared to be more promiscuous than reported and inhibited ALDH3A1, ALDH1A1, ALDH2 and ALDH3A2 under these conditions. These results challenge the reported selectivity for ALDH1A1 and ALDH3A1 of NCT-505¹⁴ and CB7¹⁵, respectively.

For both inhibitors these selectivity claims are based on biochemical substrate assays using purified enzymes. However, the activity of isolated enzymes in a biochemical assay does not necessarily reflect their activity in a cellular environment. It can therefore be argued that the selectivity profiles derived via an *in situ* competitive ABPP method provide a more accurate representation of the cellular target-engagement of a drug candidate.²⁵ The ability of CB7 to sensitize lung cancer cells to cyclophosphamide treatment,¹⁵ can therefore not be solely attributed to its inhibition of ALDH3A1, but might actually require dual inhibition of ALDH1A1 and ALDH3A1. From this point of view, NCT-505 has therapeutic potential as a dual inhibitor of these cancer therapy resistance biomarkers.

These results show the ability of **STA-55** to be used for target identification and engagement studies using competitive ABPP of new drug candidates for ALDHs. **STA-55** is easily accessible compared with the synthetically challenging **LEI-945**. **STA-55** is also capable of enriching all ALDH enzymes present in the A549 lung cancer cells, including the cancer biomarker ALDH3A1, where **LEI-945** is more specifically targeted at the retinaldehyde dehydrogenases.

Conclusion

In summary, this chapter describes the design, development and biological validation of **STA-55**, a broadspectrum activity-based probe for the family of aldehyde dehydrogenases. Certain ALDHs are often upregulated in cancer and confer therapy resistance. This probe enables the identification and quantification of these cancer biomarkers using chemical proteomics. Furthermore, it shows the ability of **STA-55** to be used in cellular target identification and target engagement studies, and proposes that the probe may be used to facilitate early drug discovery studies aimed at the identification of selective and tissue-permeable ALDH inhibitors.

Acknowledgements

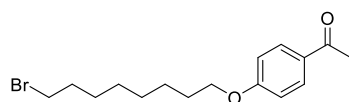
Eva van Rooden and Bogdan Florea are kindly acknowledged for the mass spectrometry analysis.

Experimental procedures

Synthetic methods

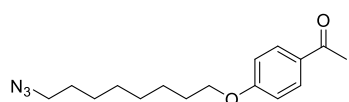
General remarks. All reactions were performed using oven or flame-dried glassware and dry solvents. Reagents were purchased from Sigma Aldrich, Acros, Biosolve, VWR, Fluka, Fischer Scientific and Merck and used as received unless stated otherwise. Inhibitors NCT-505 and CB7 were prepared as previously described.^{14,15} Tetrahydrofuran (THF) and *N,N*-dimethylformamide (DMF) were stored over 4 Å molecular sieves before use. All moisture sensitive reactions were performed under a nitrogen atmosphere. TLC analysis was performed using Merck aluminum sheets (TLC silica gel 60/Kieselguhr F₂₅₄). Compounds were visualized using a solution of KMnO₄ (7.5 g), K₂CO₃ (50 g), 10% NaOH (6 mL) in H₂O (1 L). Column chromatography was performed using Screening Device B.V. silica gel (particle size 40 – 63 µm, pore diameter of 60 Å) with the indicated eluents. ¹H- and ¹³C-NMR spectra were recorded on Bruker AV-400 (400 MHz and 101 MHz, respectively) or Bruker AV-500 MHz (500 MHz and 125 MHz, respectively) using CDCl₃ as solvent. Chemical shifts are reported in ppm (δ) relative to the residual solvent peak or tetramethylsilane. Coupling constants are given in Hz. High-resolution mass spectrometry (HRMS) analysis was performed with a LTQ Orbitrap mass spectrometer (Thermo Finnigan), equipped with an electrospray ion source in positive mode (source voltage 3.5 kV, sheath gas flow 10 mL/min, capillary temperature 250 °C) with resolution R = 60000 at m/z 400 (mass range m/z = 150 – 2000) and dioctyl phthalate (m/z = 391.28428) as a “lock mass”, or with a Synapt G2-Si (Waters), equipped with an electrospray ion source in positive mode (ESI-TOF), injection via NanoEquity system (Waters), with LeuEnk (m/z = 556.2771) as “lock mass”. Eluents used: MeCN:H₂O (1:1 v/v) supplemented with 0.1% formic acid. The high-resolution mass spectrometers were calibrated prior to measurements with a calibration mixture (Thermo Finnigan).

1-(4-((8-Bromooctyl)oxy)phenyl)ethan-1-one (2):



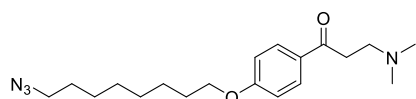
To a mixture of 1,8-dibromooctane (4.1 mL, 22 mmol) and K₂CO₃ (1.1 g, 8.1 mmol) in acetone (40 mL) at 56 °C was added dropwise 4-hydroxyacetophenone (1.0 g, 7.3 mmol) in acetone (20 mL). The reaction mixture was stirred for 18 hours at 56 °C. The reaction mixture was then allowed to cool down, filtered and concentrated under reduced pressure. The residue was then dissolved in EtOAc, the organic layer washed with H₂O and brine, dried with Na₂SO₄, filtered and concentrated under reduced pressure. Purification of the residue by column chromatography (DCM/pentane) afforded the title compound **2** (2.0 g, 6.0 mmol, 82%) as a white solid. *R*_f(50% DCM in pentane) = 0.5. ¹H NMR (400 MHz, CDCl₃): δ 7.93 (d, *J* = 4.8 Hz, 2H), 6.92 (d, *J* = 4.8 Hz, 2H), 4.02 (t, *J* = 6.8 Hz, 2H), 3.41 (t, 6.8 Hz, 2H), 2.56 (s, 3H), 1.83 (m, 4H), 1.47 (m, 4H), 1.39 (m, 4H). ¹³C NMR (100 MHz, CDCl₃): δ 196.8, 163.1, 130.6, 130.1, 114.1, 68.1, 34.0, 32.7, 29.1, 29.0, 28.6, 28.1, 26.3, 25.9.

1-(4-((8-Azidooctyl)oxy)phenyl)ethan-1-one (3):



To a solution of compound **2** (0.50 g, 1.5 mmol) in DMF (5 mL) under Ar was added sodium azide (0.30 g, 4.6 mmol). The reaction mixture was stirred for 18 hours at 80 °C and then allowed to cool down. H₂O was added and the aqueous layer was extracted with Et₂O. The combined organic layers were washed with brine, dried with MgSO₄, filtered and concentrated under reduced pressure. Purification of the residue by column chromatography (pentane/DCM) afforded the title compound **3** (0.41 g, 1.4 mmol, 93%) as a colorless liquid. *R*_f(50% DCM in pentane) = 0.45. ¹H NMR (400 MHz, CDCl₃): δ 7.93 (d, *J* = 4.8 Hz, 2H), 6.92 (d, *J* = 4.8 Hz, 2H), 4.02 (t, *J* = 6.8 Hz, 2H), 3.26 (t, 6.8 Hz, 2H), 2.55 (s, 3H), 1.81 (m, 2H), 1.61 (m, 2H), 1.47 (m, 2H), 1.39 (m, 6H). ¹³C NMR (100 MHz, CDCl₃): δ 196.8, 163.0, 130.5, 130.1, 114.1, 68.1, 51.4, 29.1, 29.0, 28.8, 26.6, 26.3, 25.8.

1-(4-((8-Azidooctyl)oxy)phenyl)-3-(dimethylamino)propan-1-one (STA-55):



To a solution of compound **3** (164 mg, 0.57 mmol) in EtOH (2 mL) under Ar was added dimethylamine hydrochloride salt (92 mg, 1.1 mmol), paraformaldehyde (34 mg, 1.1 mmol) and one drop of conc. HCl. The reaction mixture was stirred at 90 °C for 18 hours and then allowed to cool down. Purification of the reaction mixture by column chromatography (DCM/MeOH) recovered compound **3** (127 mg, 0.43 mmol, 75%) and afforded the title compound **STA-55** (42.2 mg, 0.12 mmol, 21%) as yellow solid. *R*_f(10% MeOH in DCM) = 0.5. ¹H (400 MHz, CDCl₃): δ 7.96 (d, *J* = 8.8 Hz, 2H), 6.92 (d, *J* = 8.8 Hz, 2H); 4.03 (t, *J* = 6.4 Hz, 2H), 3.68 (t, *J* = 6.4 Hz, 2H), 3.51 (t, *J* = 6.4 Hz, 2H), 3.27 (t, *J* = 7.2 Hz, 2H), 2.84 (s, 6H), 1.81 (m, 2H), 1.59 (m, 2H), 1.46 (m, 2H), 1.37 (m, 6H).

^{13}C NMR (100 MHz, CDCl_3): δ 194.1, 163.9, 130.6, 128.2, 114.4, 68.2, 52.8, 51.4, 43.3, 33.4, 29.1, 29.0, 28.9, 28.7, 26.6, 25.8. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ calculated for $\text{C}_{19}\text{H}_{30}\text{O}_2$: 347.24415, found 347.24433.

In situ labeling procedure

ALDH plasmids. For the preparation of the different constructs, full length human cDNA was purchased from Source Bioscience and cloned into mammalian expression vector pcDNA3.1, containing genes for ampicillin and neomycin resistance. ALDH1A3 was cloned into pcDNA3.1. A FLAG-linker was cloned into the vector at the C-terminus of ALDH1A3. Two step PCR mutagenesis was performed to substitute the active site cysteine for an alanine in the hALDH1A3-FLAG to obtain hALDH1A3-C314A-FLAG. All plasmids were grown in XL-10 Z-competent cells and prepped (Maxi Prep, Qiagen). The sequences were confirmed by sequence analysis at the Leiden Genome Technology Centre.

Cell culture. U2OS cells were grown in DMEM with stable glutamine and phenol red with 10% New Born Calf serum, penicillin and streptomycin at 37 °C and 7% CO_2 . A549 cells were grown in DMEM with stable glutamine and phenol red with 10% New Born Calf serum, penicillin and streptomycin at 37 °C and 5% CO_2 . Medium was refreshed every 2-3 days and cells were passaged twice a week. Cell lines were purchased from ATCC and were regularly tested for mycoplasma contamination. Cultures were discarded after 2-3 months of use.

Transient transfection of U2OS cells. One day prior to transfection 4×10^5 U2OS cells were seeded in a 6-wells plate. Cells were transfected by addition of a 3:1 mixture of polyethyleneimine (6 μg) and plasmid DNA (2 μg) in 200 μL serum free medium per well. The medium was refreshed after 24 hours and after 48 hours the cells were used for subsequent assays.

***In situ* activity-based protein profiling.** Growth medium from cells grown in 6-wells plate was removed and 1 mL serum free medium containing probe STA-55 (10 μM , 0.1% DMSO) was added. The cells were then incubated for 1 hour. For competitive ABPP cells were first incubated with vehicle or inhibitor (10 μM , 0.1% DMSO) for 30 minutes followed by STA-55 (1 μM final concentration) for 1 hour. The medium was then removed, the cells were washed with 2 mL PBS and then harvested in 1 mL PBS using a cell scraper. The cells were moved to an eppendorf tube and the suspension was centrifuged for 5 minutes at 1200 rpm. The PBS was removed and the samples snap frozen and stored at -80 °C until use.

CuAAC reaction and in-gel fluorescence analysis. Cell pellets were thawed on ice, lysed by addition of ice-cold lysis buffer (MilliQ, 1x protease inhibitor cocktail (Roche cOmplete EDTA free)) and incubated on ice (15-30 min). The protein concentration was determined by a Quick Start™ Bradford Protein assay (Bio-Rad). The protein fractions were diluted to a total protein concentration of 1 mg/mL. From each sample 40 μL was taken and treated with 5 μL from a freshly prepared “click” mixture containing 9 mM CuSO_4 (2.5 μL /sample, 18 mM in H_2O), 45 mM NaAsc (1.5 μL /sample, 150 mM in H_2O), 1.8 mM THPTA (0.5 μL /sample, 18 mM in DMSO) and 9 μM Cy5-alkyne (0.5 μL /sample, 90 μM in DMSO from Thermo Fischer Scientific). The samples were incubated for 1 hour at 37°C and then 15 μL 4x SDS page sample buffer was added. The samples were denatured at 100 °C for 5 minutes. 8 μg per sample was resolved on a SDS-PAGE gel (10% acrylamide, 180V, 75 min). Gels were visualized with a ChemiDoc XRS (Bio-Rad) using Cy3 and Cy5 multichannel settings (605/50 and 695/55, filters respectively) and stained with Coomassie or transferred to a 0.2 μm polyvinylidene difluoride membranes by Trans-Blot Turbo™ Transfer system (Bio-Rad) after scanning. Fluorescence was normalized to Coomassie staining or to FLAG-tag signal and quantified with Image Lab (Bio-Rad).

Western Blotting. Proteins were transferred to a 0.2 μm polyvinylidene difluoride membranes by Trans-Blot Turbo™ Transfer system (Bio-Rad). Membranes were washed with TBS (50 mM Tris, 150 mM NaCl), washed with TBS-T (50 mM Tris, 150 mM NaCl, 0.05% Tween 20) and then blocked with 5% w/v milk powder in TBS-T for 1 hour at room temperature. Membranes were then incubated with primary antibody in 5% milk TBS-T (α -FLAG: 1 h, RT), washed three times with TBS-T, incubated with matching secondary antibody in 5% milk TBS-T (1 h, RT) and washed with TBS-T and TBS. The blot was developed in the dark using an imaging solution (10 mL Luminol, 100 μL ECL enhancer and 3 μL 30% H_2O_2) and chemiluminescence was visualized using a ChemiDoc XRS (Bio-Rad). The signal was normalized to Coomassie staining and quantified with Image Lab (Bio-Rad). Primary antibody: monoclonal mouse anti-FLAG (1:5000, Sigma-Aldrich, F3165). Secondary antibody: HRP-coupled goat-anti-mouse (1:5000, Santa Cruz, sc2005).

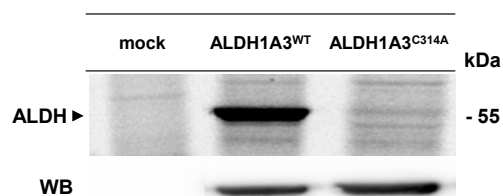
In situ activity-based proteomics

Sample preparation. Protocol adapted from previously described procedure.²⁶ Cells were treated *in situ*, harvested, lysed and adjusted to 1 mg/mL protein concentration as described above. 250 μ L was taken from each sample and to this 25 μ L freshly prepared “click” mixture containing 1 mM CuSO₄ (2.5 μ L/sample, 100 mM in H₂O), 5 mM NaAsc (1.25 μ L/sample, 1 M in H₂O), 0.4 mM THPTA (1 μ L/sample, 100 mM in DMSO), 40 μ M biotin-alkyne (2.5 μ L/sample, 4 mM in DMSO) and MilliQ (17.75 μ L/sample) was added. Samples were incubated for 1 hour at 37 °C while shaking (300 rpm). Excess click reagents were then removed by chloroform/methanol precipitation as described before. Precipitated proteomes were then suspended in urea buffer (250 μ L, 6 M urea and 25 mM ammonium bicarbonate), DTT (2.5 μ L, 1 M) was added and the mixture was then incubated for 15 min at 65 °C while shaking (600 rpm). The samples were then allowed to cool down to RT and then alkylated by addition of iodoacetamide (20 μ L, 0.5 M) for 30 minutes at RT in the dark. Addition of SDS (70 μ L, 10% (v/v)) was followed by heating at 65 °C for 5 minutes. For each sample 50 μ L 50% slurry of Avidin-Agarose from egg white (Sigma-Aldrich) was washed three times with PBS by centrifugation and transferred in PBS (1 mL) to a 15 mL tube. To this another 2 mL of PBS was added followed by the corresponding proteome sample. The beads were incubated with the proteome for 2 hours at room temperature using an overhead shaker. The beads were then isolated by centrifugation (2 min, 2500 g), washed with SDS in PBS (0.5% (w/v)) and washed three times with PBS. The beads were then transferred to low-binding Eppendorf tubes and proteins were digested overnight at 37 °C and 950 rpm shaking in 250 μ L digestion buffer (100 mM Tris, 100 mM NaCl, 1 mM CaCl₂, 2% acetonitrile and 0.5 μ g sequencing grade trypsin (Promega)). Digestion was stopped by addition of formic acid (12.5 μ L) and the beads filtered off by centrifugation (2 min, 600 g) using a Bio-Spin column (Bio-Rad). Samples were then desalted using stage tips, collected in low-binding Eppendorf tubes, concentrated using a SpeedVac (Eppendorf) and stored at -20 °C until reconstitution before measurement.²⁷ All samples were prepared in at least three biological replicates.

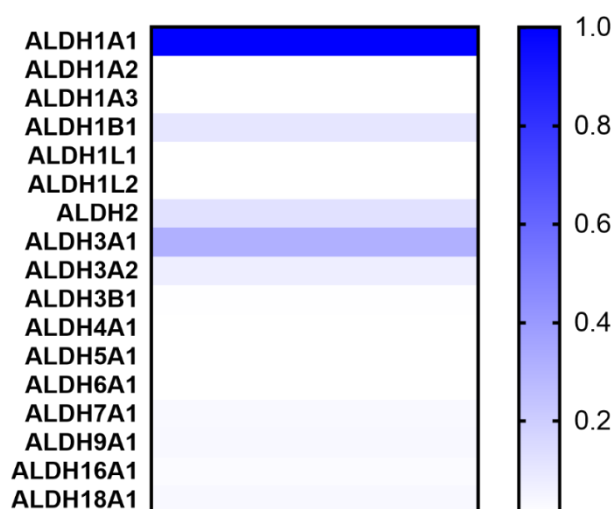
LC-MS/MS measurement and analysis. Samples were reconstituted in LC-MS sample solution (50 μ L, MilliQ, 3% acetonitrile/0.1% formic acid/20 fmol/ μ L enolase). Samples were then analysed using a NanoACQUITY UPLC System (Waters) coupled to a SYNAPT G2-Si high-definition mass spectrometer (Waters) as previously described.^{26,28} Of each sample 5 μ L was loaded on a nanoEASE™ M/Z Symmetry C18 trap column (particles 5 μ m, 100 Å, 180 μ m x 20 mm, Waters) with 0.1% formic acid and separated on an nanoEASE™ M/Z HSS C18 T3 analytical column (particles 1.8 μ m, 75 μ m x 250 mm, Waters) heated at 80 °C. A multistep gradient running from 5-40% acetonitrile containing 0.1% formic acid during a 70 minute method at 300 nL/min was used to achieve peptide separation. Survey scans (m/z 50-2000 Da) were acquired in the Synapt with a scan time of 0.6 seconds in positive resolution mode. The collision energy was set to 4 V in the trap cell for low-energy MS mode. For the elevated energy scan, the transfer cell collision energy was ramped using drift-time specific collision energies. The lock mass was sampled every 30 seconds. MS raw files were analysed with ProteinLynx Global SERVER (PLGS, v3.0.3, Waters). The MS^E identification was also performed with PLGS using the human proteome from Uniprot (uniprot-homo-sapiens-trypsin-reviewed-2016-08-29.fasta). The following parameter settings were used: low energy threshold 150 counts, elevated energy threshold 30, peptide and protein FDR 1%, enzyme specificity trypsin, max missed cleavages max 2, variable modification methionine oxidation, fixed modification carbamidomethylation cysteine, at least: fragments/peptide 2, fragments/protein 5, peptides/protein 1 and number of peptides to measure per protein 3. For label-free quantification ISOQuant (v1.5) was used.^{29,30} Data were filtered to retain only proteins with two or more reported unique peptides and quantified in at least 3 replicates of the positive control (probe-treated). Proteins were designated as significantly enriched by the probe when they showed 2-fold enrichment in quantification value when comparing negative control (vehicle-treated) with positive control (probe-treated) samples and probability as determined by a Student's *t* test (<0.05).

Heatmap Competitive ABPP analysis. Only significantly enriched ALDH enzymes were selected for analysis. The mean raw LFQ intensities from quadruplicate measurements were normalized to DMSO (= 0) and maximal LFQ STA-55 (= 1) for each protein individually. The heatmap was prepared using Graphpad Prism® 7 (Graphpad Software Inc.).

Supplementary Data

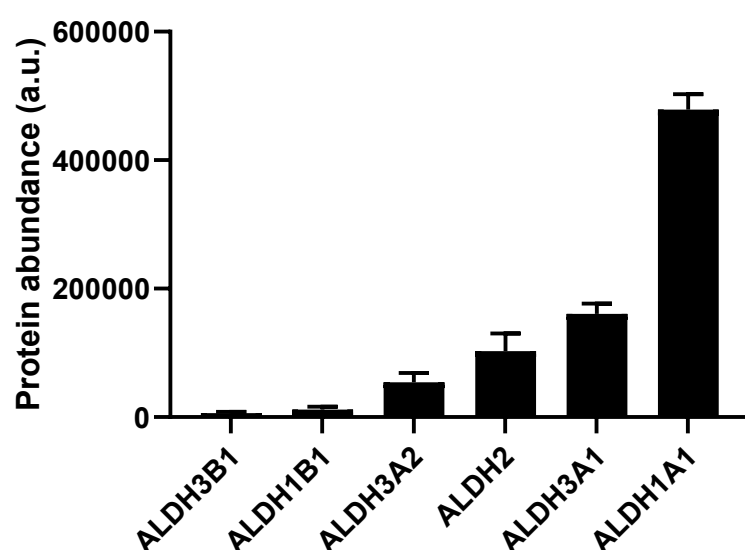


Supplementary Fig. 8.1 | *In situ* labeling of ALDH1A3^{WT} and mutant ALDH1A3^{C314A} with STA-55 (1 μ M) for 1 h at 37 °C and α -FLAG Western blot.



Supplementary Fig. 8.2 | Heatmap showing relative protein levels of ALDH enzymes in A549 cells using deep proteome data from Bekker-Jensen *et al.* analysed using the Expression Atlas.^{20,21}

Levels of active ALDHs in A549 cells



Supplementary Fig. 8.3 | Data represent mean LFQ levels $\pm$ SD of the ALDHs significantly enriched in the A549 lung cancer cell line. $N = 3$ experiments (biological replicates).

References

1. Vasiliou, V., Pappa, A. & Petersen, D. R. Role of aldehyde dehydrogenases in endogenous and xenobiotic metabolism. *Chem. Biol. Interact.* **129**, 1–19 (2000).
2. Kutzbach, C. & Stokstad, E. L. R. Mammalian methylenetetrahydrofolate reductase Partial purification, properties, and inhibition by S-adenosylmethionine. *BBA - Enzymol.* **250**, 459–477 (1971).
3. Chern, M. K. & Pietruszko, R. Human aldehyde dehydrogenase E3 isozyme is a betaine aldehyde dehydrogenase. *Biochem. Biophys. Res. Commun.* **213**, 561–568 (1995).
4. Vasiliou, V. & Pappa, A. Polymorphisms of Human Aldehyde Dehydrogenases. *Pharmacology* **61**, 192–198 (2000).
5. Laurenzi, V. De *et al.* Sjögren–Larsson syndrome is caused by mutations in the fatty aldehyde dehydrogenase gene. *Nat. Genet.* **12**, 52–57 (1996).
6. Geraghty, M. T. *et al.* Mutations in the Δ^1 -pyrroline 5-carboxylate dehydrogenase gene cause type II hyperprolinemia. *Hum. Mol. Genet.* **7**, 1411–1415 (1998).
7. Chambliss, K. L. *et al.* Two Exon-Skipping Mutations as the Molecular Basis of Succinic Semialdehyde Dehydrogenase Deficiency (4-Hydroxybutyric Aciduria). *Am. J. Hum. Genet.* **63**, 399–408 (1998).
8. Moreb, J. S., Muhoczy, D., Ostmark, B. & Zucali, J. R. RNAi-mediated knockdown of aldehyde dehydrogenase class-1A1 and class-3A1 is specific and reveals that each contributes equally to the resistance against 4-hydroperoxycyclophosphamide. *Cancer Chemother. Pharmacol.* **59**, 127–136 (2007).
9. Emadi, A., Jones, R. J. & Brodsky, R. A. Cyclophosphamide and cancer: Golden anniversary. *Nature Reviews Clinical Oncology* **6**, 638–647 (2009).
10. Yasgar, A. *et al.* A High-Content assay enables the automated screening and identification of small molecules with specific ALDH1A1-Inhibitory activity. *PLoS One* **12**, 1–19 (2017).
11. Serwa, R. & Tate, E. W. Activity-based profiling for drug discovery. *Chemistry and Biology* **18**, 407–409 (2011).
12. Koenders, S. T. A. *et al.* Development of a Retinal-Based Probe for the Profiling of Retinaldehyde Dehydrogenases in Cancer Cells. *ACS Cent. Sci.* **5**, 1965–1974 (2019).
13. Morgan, C. A., Parajuli, B., Buchman, C. D., Dria, K. & Hurley, T. D. N,N-diethylaminobenzaldehyde (DEAB) as a substrate and mechanism-based inhibitor for human ALDH isoenzymes. *Chem. Biol. Interact.* **234**, 18–28 (2015).
14. Yang, S. M. *et al.* Discovery of Orally Bioavailable, Quinoline-Based Aldehyde Dehydrogenase 1A1 (ALDH1A1) Inhibitors with Potent Cellular Activity. *J. Med. Chem.* **61**, 4883–4903 (2018).
15. Parajuli, B., Fishel, M. L. & Hurley, T. D. Selective ALDH3A1 inhibition by benzimidazole analogues increase mafosfamide sensitivity in cancer cells. *J. Med. Chem.* **57**, 449–461 (2014).
16. Khanna, M. *et al.* Discovery of a novel class of covalent inhibitor for aldehyde dehydrogenases. *J. Biol. Chem.* **286**, 43486–43494 (2011).
17. Moretti, A. *et al.* Crystal structure of human aldehyde dehydrogenase 1A3 complexed with NAD⁺ and retinoic acid. *Sci. Rep.* **6**, 35710 (2016).
18. Moreb, J. S., Zucali, J. R., Ostmark, B. & Benson, N. A. Heterogeneity of aldehyde dehydrogenase expression in lung cancer cell lines is revealed by Aldefluor flow cytometry-based assay. *Cytom. Part B Clin. Cytom.* **72B**, 281–289 (2007).
19. Kang, J. H. *et al.* Aldehyde dehydrogenase is used by cancer cells for energy metabolism. *Exp. Mol. Med.* **48**, 1–13 (2016).
20. Bekker-Jensen, D. B. *et al.* An Optimized Shotgun Strategy for the Rapid Generation of Comprehensive Human Proteomes. *Cell Syst.* **4**, 587–599 (2017).
21. Petryszak, R. *et al.* Expression Atlas update - An integrated database of gene and protein expression in humans,

- animals and plants. *Nucleic Acids Res.* **44**, 746–752 (2016).
22. Ogata, H., Goto, S., Fujibuchi, W. & Kanehisa, M. Computation with the KEGG pathway database. *Biosystems* **47**, 119–128 (1998).
 23. Thomas, P. D. *et al.* PANTHER: A browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Research* **31**, 334–341 (2003).
 24. Bateman, A. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, 506–515 (2019).
 25. 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).
 26. 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).
 27. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
 28. Distler, U., Kuharev, J., Navarro, P. & Tenzer, S. Label-free quantification in ion mobility-enhanced data-independent acquisition proteomics. *Nat. Protoc.* **11**, 795–812 (2016).
 29. Distler, U. *et al.* Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics. *Nat. Methods* **11**, 167–170 (2014).
 30. Kuharev, J., Navarro, P., Distler, U., Jahn, O. & Tenzer, S. In-depth evaluation of software tools for data-independent acquisition based label-free quantification. *Proteomics* **15**, 3140–3151 (2015).

Chapter 9

Summary and Future Prospects

General Summary

Aldehyde dehydrogenases (ALDHs) convert endogenous and exogenous reactive aldehydes into carboxylic acids to reduce or prevent cellular damage (see **Chapter 1**).¹⁻³ In addition, ALDHs perform an essential metabolic role in several cellular processes.⁴⁻⁷ For example, retinoic acid is an important signaling lipid, generated by ALDHs. Retinoic acid controls many cellular and physiological functions via gene transcription.⁸⁻¹³ Retinoic acid is produced by retinaldehyde dehydrogenases ALDH1A1, ALDH1A2 and ALDH1A3 from its precursor retinal.¹⁴ Retinaldehyde dehydrogenases are tightly regulated via an inducible cellular expression pattern and by post-translation modifications.^{15,16}

Current methods allow researchers to measure expression levels of these enzymes or general ALDH activity in cells or tissue, but not the activity of individual enzymes. New strategies are highly desired to visualize and quantify ALDH activity of specific enzymes in health and disease. The aim of this thesis was, therefore, to develop and apply new chemical tools to study the activity of ALDHs in complex biological samples.

In **Chapter 2**, the recent developments in the field of chemical probes to study lipid biology, especially in immunology, are summarized and potential avenues for future research are indicated.¹⁷ Lipids perform numerous functions inside the cell, ranging from structural building block of membranes and energy storage to cell signaling. The mode of action of many signaling lipids has remained elusive due to their low abundance, high lipophilicity, and inherent instability. Various chemical biology approaches, such as photoaffinity or activity-based protein profiling methods, have recently been developed to shed light on the biological role of lipids, lipid-protein interactions and lipid processing enzymes.

For example, activity-based protein profiling (ABPP) is a powerful methodology to map enzyme activities in complex biological samples.^{18,19} ABPP relies on activity-based probes containing an electrophilic warhead, which reacts with a conserved catalytic amino acid residue of an enzyme. Visualization of the probe targets via reporter groups, such as fluorophores or biotin, is performed by in-gel fluorescent scanning or chemical proteomics. Comparative and competitive ABPP are the two main applications used in this thesis. In comparative ABPP the abundance of active enzymes derived from different biological samples are studied, whereas competitive ABPP is used to determine the selectivity profile of an inhibitor.

Activity-based protein profiling of retinaldehyde dehydrogenases

Chapter 3 describes the design and synthesis of a first-in-class probe for ALDHs. Guided by the co-crystal structure of ALDH1A1, retinal was functionalized with an alkyne ligation handle and an electrophilic warhead. This led to the design of **LEI-945**.²⁰ A convergent synthesis route utilizing a key Wittig reaction was developed to synthesize **LEI-945**. This retinal-based probe was biologically validated in **Chapter 4**.

LEI-945 labelled ALDH1A1, ALDH1A2 and ALDH1A3 by a covalent interaction with their catalytic cysteines. In addition, the probe was able to detect endogenously expressed ALDHs in A549 lung cancer cells, including ALDH2, ALDH3A2 and ALDH3B1. **LEI-945** also revealed the selectivity profile of pan-ALDH inhibitor 4-diethylaminobenzaldehyde (DEAB) and retinal using competitive ABPP in A549 cells.

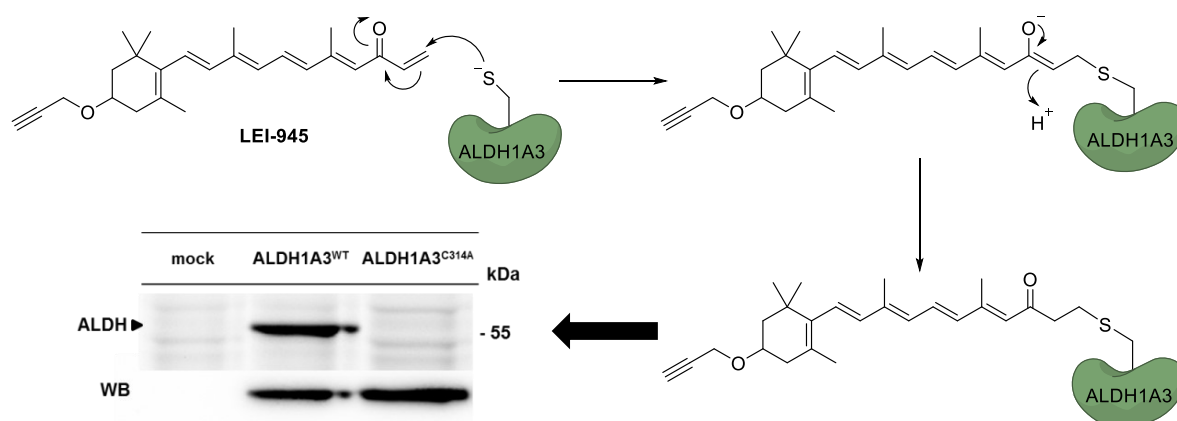


Fig. 9.1 | Covalent interaction of LEI-945 with and visualization of ALDH1A3. The vinyl ketone warhead of **LEI-945** undergoes nucleophilic attack by the catalytic cysteine of ALDH1A3. The resulting thioether cannot be hydrolyzed and results in an irreversible covalent interaction between ALDH1A3 and **LEI-945**. Ligation with AlexaFluor-647 and visualization by in-gel fluorescence showed a clear band around 55 kDa. Altering the catalytic cysteine into an alanine completely abolished labeling.

Profiling of ALDHs in breast cancer subtypes

Having established that **LEI-945** can act as an ABP for ALDHs, the probe was applied in **Chapter 5** for comparative ABPP using different breast cancer cell types. Upregulation of certain ALDHs has been previously linked to therapy resistance in cancer.²¹ For example, ALDH1A3 is associated with poor clinical outcome in breast cancer.^{22–26} The ALDEFLUOR assay, which is routinely used to measure ALDH activity, was not capable of predicting the ability of breast cancer cell lines to produce retinoic acid. Using **LEI-945** distinct ALDH activity profiles were made for each cell line. These profiles were used to explain the differences in retinal conversion based on the abundance of active ALDH1A1 and/or ALDH1A3 enzymes.

Remarkably, SK-BR-3, the breast cancer cell line showed the highest level of overall ALDH activity and displayed exceptionally high levels of active ALDH2, but converted a relatively low fraction of retinal. The low levels of ALDH1A3 activity in this cell line as determined by ABPP using **LEI-945** were sufficient to explain the conversion of retinal into retinoic acid. This suggested that human ALDH2 does not have a biologically relevant role in vitamin A metabolism.

Comparative ABPP

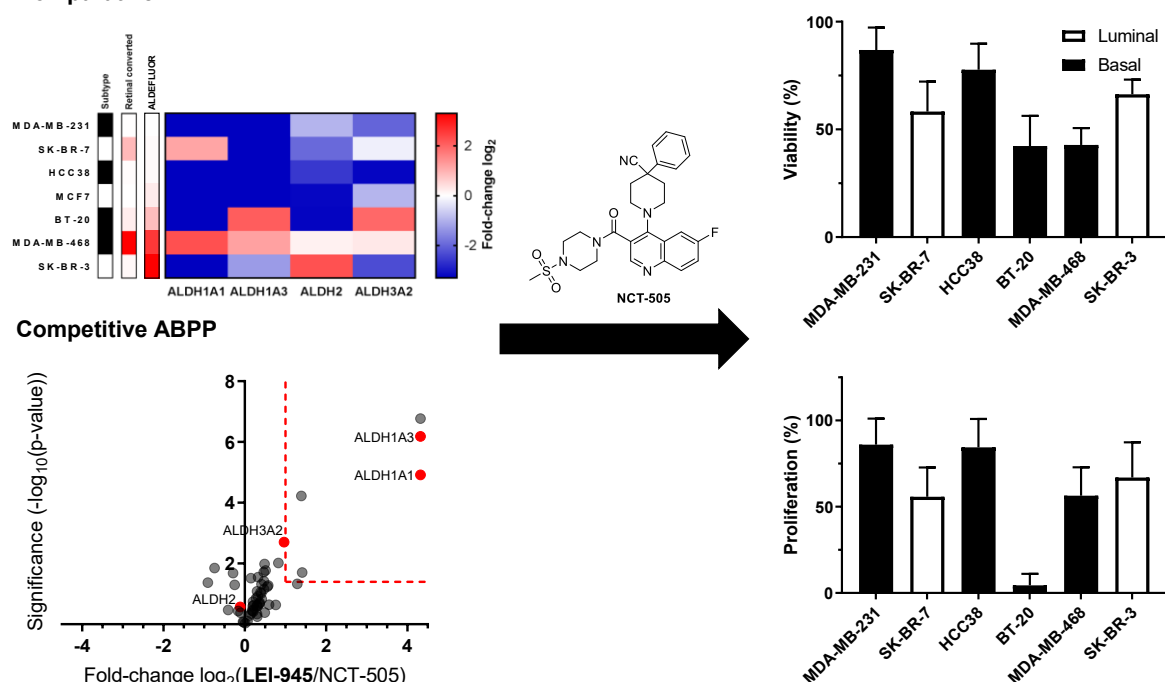


Fig. 9.2 | Comparative and competitive ABPP of ALDHs and subsequent cellular assays with NCT-505 in breast cancer cells. *Comparative ABPP*: ALDH profiling of breast cancer cell lines using chemical proteomics. The subtype column indicates if the cell line belongs to the luminal subtype (white) or basal subtype (black). The retinal converted column shows the amount of retinal converted to retinoic acid over 4 h in a gradient from 0% (white) to 100% (red). The ALDEFLUOR column shows the ALDH activity as determined by the ALDEFLUOR assay in a gradient from low (white) to high (red). The heatmap shows the fold-change in LFQ value for each ALDH enzyme compared to the average for each ALDH enzyme. $N = 2$ independent experiments with each at least $n = 3$ experiments per group (biological replicates). *Competitive ABPP*: Volcano plot of the *in situ* competitive ABPP experiment in MDA-MB-468 to determine off-targets of ALDH inhibitor NCT-505 (30 μM). $N = 4$ experiments per group (biological replicates). *Cellular assays*: cell viability of breast cancer cell lines after treatment with NCT-505 (30 μM) for 72 h. For viability and proliferation assay data represent mean values $\pm$ SD; $N = 3$ biological replicates with each $n = 3$ experiments per group.

LEI-945 was used for competitive ABPP in the MDA-MB-468 cell line to determine the selectivity profile of NCT-505, an advanced ALDH1A1 selective inhibitor.²⁷ Chemical proteomics analysis showed that NCT-505 inhibited both ALDH1A1 and ALDH1A3 and showcased the potential utility of **LEI-945** to guide drug discovery efforts toward selective ALDH inhibitors.

Combining the results from the comparative and competitive ABPP experiments, indicated that treatment of the breast cancer cell lines with the dual ALDH1A1/ALDH1A3 inhibitor would mostly affect the viability and proliferation of the SK-BR-7, BT-20 and MDA-MB-468 cell lines due to their high levels of ALDH1A1 or ALDH1A3 activity (**Fig. 9.2**). A slight decrease could also be detected for the SK-BR-3 cell line, which might be related to the relatively low amount of ALDH1A3 detected in these cells. Of note, the ALDEFLUOR assay did not show any ALDH activity in the SK-BR-7 cell line. This highlights the sensitivity of **LEI-945** to detect low amounts of ALDH activity, which might be used to guide the selection of (therapeutic) effective inhibitors.

Another way in which **LEI-945** could be applied in this field, is to study whether the reported increase in ALDH expression in 3D cell culture compared with 2D cell culture is accompanied by an increase in the activity levels of specific ALDH isozymes.^{28,29} **LEI-945** could also be used in combination with the ALDEFLUOR assay and fluorescence-activated cell sorting, which is generally used to detect and sort therapy-resistant cancer (stem) cells. The sorted cell populations can be subjected to comparative ABPP, providing a better understanding of the underlying biology.³⁰ Although the research in this thesis has focused on the role of ALDH isozymes in breast cancer, it is envisioned that **LEI-945** could also be used to study the specific ALDH isozymes involved in other cancer types that express ALDHs.²⁵

The role of ALDHs in lipid peroxidation

The physiological effects of ALDH1A1 and ALDH1A3 inhibition were studied in more depth by determining cellular ATP levels and using live cell imaging techniques. NCT-505 reduced ATP levels and the mitochondrial membrane potential. Analyzing cell states using FUCCI showed arrest of these cells in the G₁ phase of the cell cycle. A significant increase of propidium iodide staining indicated that the cells died via a necrosis-like pathway. Combination of NCT-505 with the glutathione peroxidase 4 (GPX4) inhibitor RSL3 had a therapeutically synergetic effect. RSL3 induces ferroptosis, which is an iron-dependent form of non-apoptotic cell-death.^{31,32} Polyunsaturated fatty acids (PUFAs) can be oxidized by lipoxygenases into hydroxyperoxides.³³ Under normal conditions these reactive species are converted by GPX4 into lipid alcohols (**Fig. 9.3**).

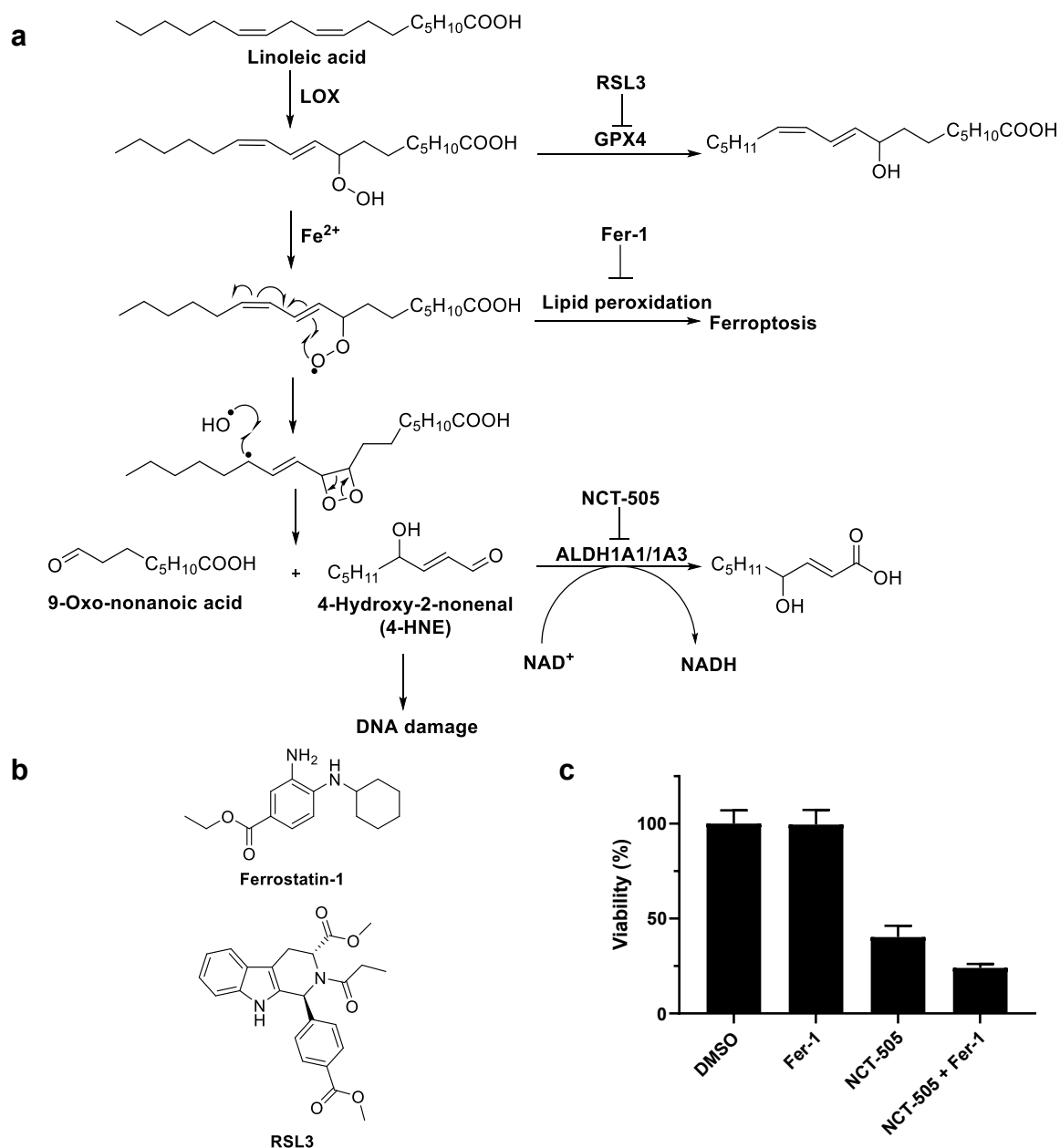


Fig. 9.3 | Overview of ferroptosis and lipid peroxidation. **a**, Schematic overview of ferroptosis and lipid peroxidation processes. Polyunsaturated fatty acid (PUFA), linoleic acid, is converted into a hydroperoxide by lipoxygenase (LOX). This hydroperoxide can then either be converted into a lipid alcohol by glutathione peroxidase 4 (GPX4) or a lipid radical under influence of iron. An accumulation of these lipid radicals induces ferroptosis, but this process can be inhibited by Fer-1. An alternative pathway is the formation of reactive aldehydes, such as 4-hydroxy-2-nonenal (4-HNE), which can bind to DNA. These reactive aldehydes are detoxified by aldehyde dehydrogenases (ALDHs) accompanied by the formation of NADH. ALDH1A1 and ALDH1A3 can be inhibited by NCT-505. **b**, Chemical structures of the radical trapping agent Ferrostatin-1 (Fer-1) and the GPX4 inhibitor RSL3. **c**, Graph showing the cell viability of MDA-MB-468 cells treated with DMSO, Fer-1 (10 μ M), NCT-505 (30 μ M) or Fer-1 (10 μ M) and NCT-505 (30 μ M). Data represent mean values $\pm$ SD; $n = 6$.

Inhibition of GPX4 by RSL3 results in an increase of lipid reactive oxygen species (ROS).³⁴ The propagation of these lipid radicals is catalyzed by iron and induces ferroptosis. Ferrostatin-1 (Fer-1; **Fig. 9.3b**), a radical trapping agent capable of blocking lipid radical propagation, was unable to rescue cells treated with NCT-505 and combinatorial treatment led to a significant ($p = 0.0002$) decrease in viability (**Fig. 9.3c**), suggesting that NCT-505 does not induce ferroptosis.

Effects of NCT-505 on MDA-MB-231 cells

Another interesting observation was made in NCT-505-treated MDA-MB-231 cells, which are derived from a triple-negative breast tumor. NCT-505 reduced the viability of the MDA-MB-231 cells to a small extent (**Fig. 9.2**), but it did induce a specific phenotype.

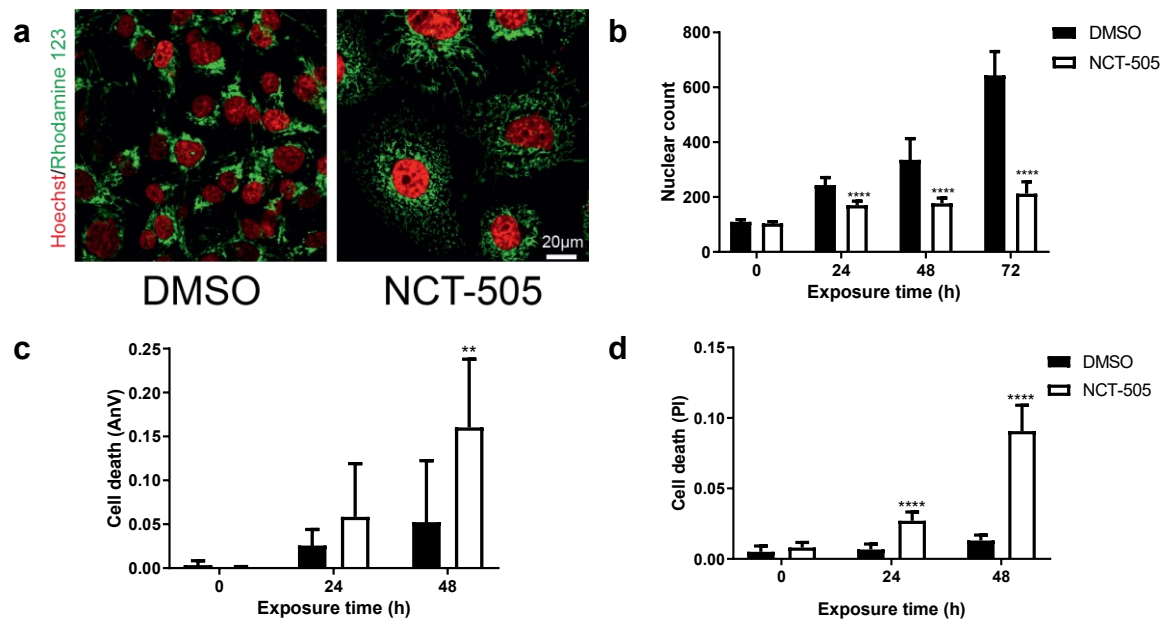


Fig. 9.4 | Effects of NCT-505 on MDA-MB-231 breast cancer cells. **a**, Representative images of live-cell confocal microscopy of mitochondrial potential after 72 h treatment with vehicle or NCT-505 (30 μ M). Mitochondrial membrane potential was visualized using rhodamine 123 dye (green) while the nuclei were detected with Hoechst 33342 (red). **b**, Nuclear count of vehicle and NCT-505 (30 μ M)-treated MDA-MB-231 cells measured over time. Data represent mean values $\pm$ SD; $N = 3$ biological replicates with each 3 experiments per group. **c**, Cell death profile of vehicle and NCT-505 (30 μ M)-treated MDA-MB-231 cells as determined by Annexin V (AnV) staining and measured over time. **d**, Cell death profile of vehicle and NCT-505 (30 μ M)-treated MDA-MB-231 cells as determined by propidium iodide (PI) staining and measured over time. For parts **c** and **d**, data represent mean values $\pm$ SD; $N = 3$ biological replicates with each 5 experiments per group. ** $P < 0.01$ and **** $P < 0.0001$; t test, two-sided.

The volume of the cells was increased along with an expansion of the nucleus (**Fig. 9.4a**). This was accompanied with a decrease in cell motility (data not shown). This phenotype seems to resemble a senescent state. Further experimentation using specific biomarkers is required to confirm this hypothesis.³⁵

Since, virtually no ALDH isozyme activity was detected in this cell line (**Fig. 9.2**), the observed phenotype is probably due to an off-target of NCT-505. Proteomics based methods, such as the cellular thermal shift assay (CETSA), can be applied to identify the target.^{36,37} Another option would be to synthesize a probe based on the scaffold of NCT-505, which includes a photo-reactive cross-linker, such as a diazirine, for affinity-based protein profiling (A_fBPP).

Of note, cell imaging showed that the nuclear count of vehicle-treated cells increased more than 3-fold after 72 hours, whereas the nuclear count of NCT-505-treated MDA-MB-231 cells increased only marginally (**Fig. 9.4b**). This inhibition of cell proliferation was not detected in the Sulforhodamine B assay, which measures total protein content. These results highlight a benefit of live cell imaging.

Visualizing the role of vitamin A metabolism in immune homeostasis

Chapter 6 describes the search for the specific retinaldehyde dehydrogenases responsible for the conversion of retinal into retinoic acid in the small intestines of mice. Retinoic acid plays an important role in immune homeostasis via the activation and recruitment of dendritic, T and B cells.^{38,39} Comparative ABPP of intestinal cells derived from *Aldh1a1*^{WT} and *Aldh1a1*^{-/-} mice with **LEI-945** was performed and identified ALDH1B1 as a possible additional source of retinoic acid in intestinal epithelial cells.

Aldh1a1^{-/-}/*Aldh1b1*^{-/-} mice showed a more proinflammatory phenotype, indicating involvement of both ALDH1A1 and ALDH1B1 in immune homeostasis. However, to determine whether the retinoid levels in the tissue of the small intestines from the knockout mice have actually changed, these levels will have to be quantified using liquid chromatography-mass spectrometry (LC-MS). The ability of *Aldh1a1*^{-/-}/*Aldh1b1*^{-/-} intestinal epithelial cells to convert retinal into retinoic acid has not been assessed, but could be measured using the LC-UV/MS assay developed in **Chapter 4**.

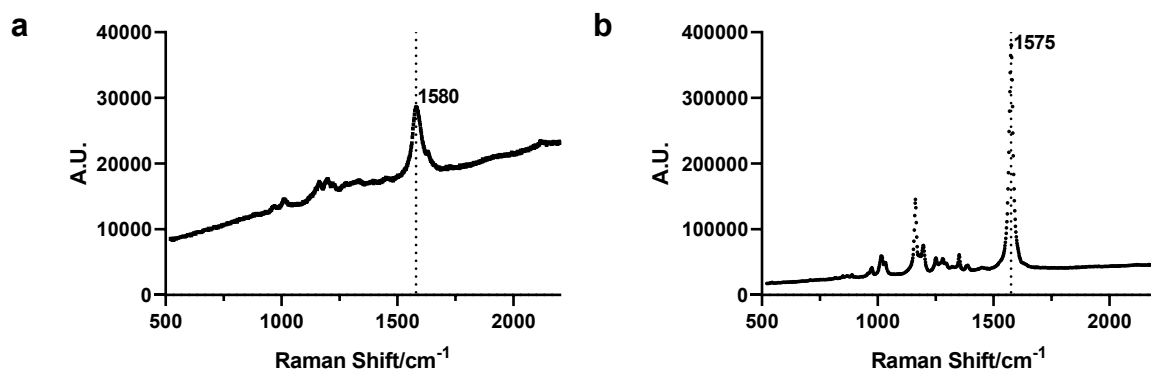


Fig. 9.5 | Raman spectra of clickable retinoids. **a**, Raman spectrum of **STA-211**. **b**, Raman spectrum of **STA-215**. For parts **a** and **b**, spectra were acquired using 532 nm excitation, ca. 0.15 mW laser power, 20 s acquisition time and a spectral range of 500 to 2200 cm^{-1} . Characteristic peaks were identified at 1580 cm^{-1} and 1575 cm^{-1} for **STA-211** and **STA-215**, respectively.

Other sources of retinoic acid should also be considered, such as bile retinoids excreted from the liver in the form of retinoyl glucuronides^{40,41} or the conversion of retinal into retinoic acid by the gut microbiome. ABPP of the gut microbiome using **LEI-945** could be used to identify bacterial enzymes and bacteria populations capable of this transformation.⁴²

New chemical tools to study vitamin A interacting proteins are introduced in **Chapter 7**. These clickable vitamins were shown to be biologically equivalent to their endogenous counterparts and were applied for ABPP without the introduction of a photo-reactive group. Another envisioned application of these probes is to study the trafficking of retinoids between cells. One interesting new technique in this field is the application of Raman microscopy to track alkyne functionalized lipids in living cells.⁴³ This method utilizes the characteristic peak of the alkyne at 2110 cm^{-1} in the biologically silent region of the Raman spectrum.⁴⁴ To determine the suitability of the clickable retinoids for this application, Raman spectra were acquired using 532 nm excitation, ca. 0.15 mW laser power and 20 seconds acquisition time in a spectral range from 500 to 2200 cm^{-1} (**Fig. 9.5a,b**).

The characteristic alkyne peak was not visible at these settings (which were 10-fold lower compared to similar experiments), but characteristic retinoid peaks of 1580 cm^{-1} and 1575 cm^{-1} were detected for the **STA-211** and **STA-215** probe, respectively.⁴⁵ The intrinsic characteristic signal of the retinoids is relatively intense and might be used to track retinoic exchange between cells depending on its signal-to-noise ratio.

Drug screening against aldehyde dehydrogenases using competitive ABPP

The development of an additional activity-based probe, **STA-55**, for aldehyde dehydrogenases is described in **Chapter 8**. This probe, derived from the reported ALDH inhibitor Aldi-2⁴⁶, was characterized as a broadspectrum probe for the ALDH family improving the coverage of the ALDH family by enriching ALDH1B1 and ALDH3A1 in addition to the ALDHs already enriched by **LEI-945** in A549 cells. It was subsequently used to determine the selectivity profile of three ALDH inhibitors by competitive ABPP.

This led to the discovery of another off-target of the ALDH1A1 inhibitor NCT-505²⁷ and revealed broadspectrum off-target activity for ALDH3A1 inhibitor CB7.⁴⁷ As **STA-55** enriched ALDH1A1 and ALDH3A1, both markers of cyclophosphamide resistance in cancer⁴⁸, it can potentially be used to predict therapy efficacy by identifying the activity of these enzymes in patient material. It may also be used in the development of new ALDH inhibitors by determining *in situ* target engagement and selectivity.

To determine whether **STA-55** truly covers the entire ALDH family the 13 remaining ALDH enzymes will have to be transiently overexpressed and subjected to labelling experiments. To enable efficient screening of selectivity profiles using ABPP, a cell line containing preferably all or at least the most clinically relevant ALDHs would be required. Thus far, the breast cancer MDA-MB-468 cell line has been used for screening, as it expresses both ALDH1A1 and ALDH1A3. Another candidate would be the lung cancer cell line A549, which proliferates rapidly and can be easily transfected.

Closing remarks

Several new chemical tools are presented in this thesis that enable the study of aldehyde dehydrogenases using activity- and affinity-based protein profiling techniques. Applications of the first-in-class retinal-based probe, **LEI-945**, include, but are not limited to, the profiling of ALDH activities in cancer in and clarifying vitamin A metabolism in the small intestines. In addition to this tailored retinal probe, the synthesis and use of clickable retinoids are discussed and a broad-spectrum probe for the ALDH family is developed. Together these chemical tools provide researchers with a comprehensive toolkit to study ALDH enzymes. The results obtained using these chemical tools showcase the application of substrate-based probes in interrogating pathologically relevant enzyme activities. They also highlight the general power of chemical proteomics in driving the discovery of new biological insights and its utility to guide drug discovery efforts.

Acknowledgements

Lukas Wijaya and Sylvia Le Dévédec are kindly acknowledged for performing and analyzing the live cell imaging, and Freek Ariese for recording the Raman spectra.

Experimental procedures

Biological methods

Cell culture. The breast cancer cell lines BT-20, HCC38, MCF7, MDA-MB-231, MDA-MB-468, SK-BR-3 and SK-BR-7 were grown in RPMI-1640 with stable glutamine and phenol red with 10% Fetal Calf serum, penicillin and streptomycin at 37 °C and 5% CO₂. Medium was refreshed every 2-3 days and cells were passaged twice a week. Cell lines were purchased from ATCC and were regularly tested for mycoplasma contamination. Cultures were discarded after 2-3 months of use.

Cell viability and proliferation. Breast cancer cells were grown to 80% confluency in 10 cm plates. Cells were then seeded in a 96-wells plate depending on their rate of growth (BT-20: 12k/well; MDA-MB-468 & SK-BR-3: 10k/well; SK-BR-7 & HCC38: 8k/well; MDA-MB-231: 4k/well). For the SRB proliferation assay cells were also seeded on an extra plate to be able to determine the starting point protein content. Cells were allowed to adhere overnight and were then treated with vehicle (0.1% DMSO) or inhibitor NCT-505 (30 µM, 0.1% DMSO) in 100 µL RPMI medium. Cells were then incubated for 72 hours. For the MTT viability assay another 100 µL of medium containing MTT (1 mg/mL) was added and the cells were incubated for another 4 hours. The medium was then removed and the formazan crystals were dissolved in 100 µL DMSO during 1 hour in the stove. The absorbance was then measured in a CLARIOstar Plus plate reader at 570 nm. For the SRB assay cells were fixated by incubation with 30 µL 50% TCA for 1 hour at 4 °C. The plates were washed 5 times with water and dried. 60 µL 0.4% SRB was added and incubated for 30 minutes at room temperature. The plates were then washed another four times with acetic acid (0.1%) and dried. The SRB was dissolved by adding 150 µL Tris (10 mM) and incubating for 30 minutes at room temperature. Absorbance was then measured at 540 nm on a CLARIOstar Plus plate reader. Graphpad Prism[®] 7 (Graphpad Software Inc.) was used to plot data and calculation of mean values and standard deviation.

Combination of NCT-505 and Fer-1. MDA-MB-468 cells were seeded in 96-wells plates at 10k/well. Cells were allowed to adhere overnight and were then treated with vehicle (0.1% DMSO), NCT-505 (30 µM, 0.1% DMSO), Fer-1 (10 µM, 0.1% DMSO, Sigma-Aldrich) or a combination of NCT-505 (30 µM) and RSL3 (10 µM). Cells were then incubated for 72 hours and a MTT assay was performed as described above. Graphpad Prism[®] 7 (Graphpad Software Inc.) was used to plot data and calculation of mean values and standard deviation.

Live cell imaging

Exposure and live cell imaging. Cells were seeded in Greiner black µ-CLEAR 96 well plates at 10k/well. Prior to the NCT-505 exposure, the cells were incubated with 100 ng/ml live Hoechst 33342 in complete RPMI for 2 hours. Thereafter, the medium was refreshed with complete medium. For cell death experiment, 0.2 µM propidium iodide (PI) and Annexin-V-Alexa633 (AnV) were added to the complete medium. For the live mitochondrial membrane potential imaging, the cells were incubated with rhodamine123 dye (0.5 µM; Sigma Aldrich) in combination with live Hoechst 33342 for 2 hours, thereafter we refreshed with complete medium containing rhodamine123 only (0.075 µM) for time-lapse imaging. Then complete medium containing vehicle or NCT-505 (30 µM final concentration) was added and the plates were directly imaged onto the microscope stage for live cell imaging.

For the cell death experiment and mitochondrial membrane potential imaging, the plates were imaged at 0, 14, 24, 40, and 72 hours after the compound exposure. The cell cycle was imaged every hour after the exposure. The imaging was performed using a Nikon TiE2000 confocal laser microscope (lasers : 647 nm, 540 nm, 488 nm, and 408 nm), equipped with automated stage and perfect focus system. During the imaging, the plates were maintained in humidified atmosphere at 37 °C and 5% CO₂. The experiment was conducted with 3 biological replicates with 4-5 technical replicates.

Image analysis. The quantitative image analysis was performed with ImageJ version 1.52p, CellProfiler version 2.2.0 and Ilastik 1.3.2. Firstly, the nuclei per image were segmented using the watershed masked algorithm on ImageJ. For images obtained from the cell death and mitochondrial membrane potential experiment, the nuclear segmentation was performed with Ilastik 1.3.2. The images were processed with an in house developed CellProfiler module^{49,50}. For the cell death and mitochondrial membrane potential experiment, the nuclear Hoechst 33342 intensity levels, rhodamine 123 integrated intensity, PI area, and Annexin V area were measured at the single cell level. The results were stored as HDF5 files. Data analysis, quality control, and graphics were performed using the in house developed R package h5CellProfiler (manuscript in preparation).

References

1. Koppaka, V. *et al.* Aldehyde Dehydrogenase Inhibitors: a Comprehensive Review of the Pharmacology, Mechanism of Action, Substrate Specificity, and Clinical Application. *Pharmacol. Rev.* **64**, 520–539 (2012).
2. Duester, G., Mic, F. A. & Molotkov, A. Cytosolic retinoid dehydrogenases govern ubiquitous metabolism of retinol to retinaldehyde followed by tissue-specific metabolism to retinoic acid. in *Chemico-Biological Interactions* **143–144**, 201–210 (Elsevier, 2003).
3. Vasilou, V., Pappa, A. & Petersen, D. R. Role of aldehyde dehydrogenases in endogenous and xenobiotic metabolism. *Chem. Biol. Interact.* **129**, 1–19 (2000).
4. Kurys, G., Ambroziak, W. & Pietruszkos, R. Human Aldehyde Dehydrogenase. *J. Biol. Chem.* **264**, 4715–4721 (1989).
5. Kutzbach, C. & Stokstad, E. L. R. Mammalian methylenetetrahydrofolate reductase Partial purification, properties, and inhibition by S-adenosylmethionine. *BBA - Enzymol.* **250**, 459–477 (1971).
6. Chern, M. K. & Pietruszko, R. Human aldehyde dehydrogenase E3 isozyme is a betaine aldehyde dehydrogenase. *Biochem. Biophys. Res. Commun.* **213**, 561–568 (1995).
7. Jackson, B. *et al.* Update on the aldehyde dehydrogenase gene (ALDH) superfamily. *Hum. Genomics* **5**, 283–303 (2011).
8. Zile, M. H. Vitamin A and Embryonic Development: An Overview. *J. Nutr.* **128**, 455–458 (1998).
9. Wiseman, E. M., Bar-El Dadon, S. & Reifen, R. The vicious cycle of vitamin a deficiency: A review. *Crit. Rev. Food Sci. Nutr.* **57**, 3703–3714 (2017).
10. Mora, J. R., Iwata, M. & Andrian, U. H. Von. Vitamin effects on the immune system. *Nat. Rev. Immunol.* **8**, 685–698 (2008).
11. Blomhoff, H. K. *et al.* Vitamin A is a key regulator for cell growth, cytokine production, and differentiation in normal B cells. *J. Biol. Chem.* **267**, 23988–23992 (1992).
12. Carlberg, C. Lipid soluble vitamins in gene regulation. *BioFactors* **10**, 91–97 (1999).
13. Balmer, J. E. & Blomhoff, R. Gene expression regulation by retinoic acid. *J. Lipid Res.* **43**, 1773–808 (2002).
14. D'Ambrosio, D. N., Clugston, R. D. & Blaner, W. S. Vitamin A metabolism: An update. *Nutrients* **3**, 63–103 (2011).
15. Wang, J. *et al.* Phosphorylation-dependent regulation of ALDH1A1 by Aurora kinase A: Insights on their synergistic relationship in pancreatic cancer. *BMC Biol.* **15**, 1–22 (2017).
16. Zhao, D. *et al.* NOTCH-induced aldehyde dehydrogenase 1A1 deacetylation promotes breast cancer stem cells. *J. Clin. Invest.* **124**, 5453–5465 (2014).
17. Koenders, S. T. A., Gagestein, B. & van der Stelt, M. Opportunities for lipid-based probes in the field of immunology. in *Current Topics in Microbiology and Immunology* **420**, 283–319 (Springer, Cham, 2018).
18. Serwa, R. & Tate, E. W. Activity-based profiling for drug discovery. *Chemistry and Biology* **18**, 407–409 (2011).
19. Cravatt, B. F., Wright, A. T. & Kozarich, J. W. Activity-Based Protein Profiling: From Enzyme Chemistry to Proteomic Chemistry. *Annu. Rev. Biochem.* **77**, 383–414 (2008).
20. Koenders, S. T. A. *et al.* Development of a Retinal-Based Probe for the Profiling of Retinaldehyde Dehydrogenases in Cancer Cells. *ACS Cent. Sci.* **5**, 1965–1974 (2019).
21. Croker, A. K. & Allan, A. L. Inhibition of aldehyde dehydrogenase (ALDH) activity reduces chemotherapy and radiation resistance of stem-like ALDH hiCD44 + human breast cancer cells. *Breast Cancer Res. Treat.* **133**, 75–87 (2012).
22. Qiu, Y. *et al.* The expression of aldehyde dehydrogenase family in breast cancer. *J. Breast Cancer* **17**, 54–60 (2014).
23. Marcato, P. *et al.* Aldehyde Dehydrogenase Activity of Breast Cancer Stem Cells Is Primarily Due To Isoform ALDH1A3 and Its Expression Is Predictive of Metastasis. *Stem Cells* **29**, 32–45 (2011).
24. Sládek, N. E., Kollander, R., Sreerama, L. & Kiang, D. T. Cellular levels of aldehyde dehydrogenases (ALDH1A1 and ALDH3A1) as predictors of therapeutic responses to cyclophosphamide-based chemotherapy of breast cancer: A retrospective study. *Cancer Chemother. Pharmacol.* **49**, 309–321 (2002).
25. Tomita, H., Tanaka, K., Tanaka, T. & Hara, A. Aldehyde dehydrogenase 1A1 in stem cells and cancer. *Oncotarget* **7**, 11018–32 (2016).
26. De Beca, F. F. *et al.* Cancer stem cells markers CD44, CD24 and ALDH1 in breast cancer special histological types. *J. Clin. Pathol.* **66**, 187–191 (2013).
27. Yang, S. M. *et al.* Discovery of Orally Bioavailable, Quinoline-Based Aldehyde Dehydrogenase 1A1 (ALDH1A1) Inhibitors with Potent Cellular Activity. *J. Med. Chem.* **61**, 4883–4903 (2018).
28. Reynolds, D. S. *et al.* Breast Cancer Spheroids Reveal a Differential Cancer Stem Cell Response to Chemotherapeutic Treatment.

- Sci. Rep.* **7**, 1–12 (2017).
29. Fujiwara, D., Kato, K., Nohara, S., Iwanuma, Y. & Kajiyama, Y. The usefulness of three-dimensional cell culture in induction of cancer stem cells from esophageal squamous cell carcinoma cell lines. *Biochem. Biophys. Res. Commun.* **434**, 773–778 (2013).
 30. Marcato, P., Dean, C. A., Giacomantonio, C. A. & Lee, P. W. K. Aldehyde dehydrogenase its role as a cancer stem cell marker comes down to the specific isoform. *Cell Cycle* **10**, 1378–1384 (2011).
 31. Dixon, S. J. *et al.* Ferroptosis: An iron-dependent form of nonapoptotic cell death. *Cell* **149**, 1060–1072 (2012).
 32. Yang, W. S. *et al.* Regulation of ferroptotic cancer cell death by GPX4. *Cell* **156**, 317–331 (2014).
 33. Brash, A. R. Lipoxygenases: Occurrence, functions, catalysis, and acquisition of substrate. *Journal of Biological Chemistry* **274**, 23679–23682 (1999).
 34. Viswanathan, V. S. *et al.* Dependency of a therapy-resistant state of cancer cells on a lipid peroxidase pathway. *Nature* **547**, 453–457 (2017).
 35. de Jesus, B. B. & Blasco, M. A. Assessing cell and organ senescence biomarkers. *Circ. Res.* **111**, 97–109 (2012).
 36. Martinez Molina, D. & Nordlund, P. The Cellular Thermal Shift Assay: A Novel Biophysical Assay for In Situ Drug Target Engagement and Mechanistic Biomarker Studies. *Annu. Rev. Pharmacol. Toxicol.* **56**, 141–161 (2016).
 37. Dziekan, J. M. *et al.* Identifying purine nucleoside phosphorylase as the target of quinine using cellular thermal shift assay. *Sci. Transl. Med.* **11**, (2019).
 38. Svensson, M. *et al.* CCL25 mediates the localization of recently activated CD8 α β +lymphocytes to the small-intestinal mucosa. *J. Clin. Invest.* **110**, 1113–1121 (2002).
 39. Hamann, A., Andrew, D. P., Jablonski-Westrich, D., Holzmann, B. & Butcher, E. C. Role of alpha 4-integrins in lymphocyte homing to mucosal tissues in vivo. *J. Immunol.* **152**, 3282–93 (1994).
 40. Jaensson-Gyllenbäck, E. *et al.* Bile retinoids imprint intestinal CD103+ dendritic cells with the ability to generate gut-tropic T cells. *Mucosal Immunol.* **4**, 438–447 (2011).
 41. Dunagin, P. E., Zachman, R. D. & Oslo, J. A. The identification of metabolites of retinal and retinoic acid in rat bile. *BBA* **124**, 71–85 (1966).
 42. Whidbey, C. *et al.* A Probe-Enabled Approach for the Selective Isolation and Characterization of Functionally Active Subpopulations in the Gut Microbiome. *J. Am. Chem. Soc.* **141**, 42–47 (2019).
 43. Jamieson, L. E. *et al.* Tracking intracellular uptake and localisation of alkyne tagged fatty acids using Raman spectroscopy. *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **197**, 30–36 (2018).
 44. Yamakoshi, H. *et al.* Alkyne-tag Raman imaging for visualization of mobile small molecules in live cells. *J. Am. Chem. Soc.* **134**, 20681–20689 (2012).
 45. Failloux, N., Bonnet, I., Baron, M. H. & Perrier, E. Quantitative Analysis of Vitamin A Degradation by Raman Spectroscopy. *Appl. Spectrosc.* **57**, 1117–1122 (2003).
 46. Khanna, M. *et al.* Discovery of a novel class of covalent inhibitor for aldehyde dehydrogenases. *J. Biol. Chem.* **286**, 43486–43494 (2011).
 47. Parajuli, B., Fishel, M. L. & Hurley, T. D. Selective ALDH3A1 inhibition by benzimidazole analogues increase mafosfamide sensitivity in cancer cells. *J. Med. Chem.* **57**, 449–461 (2014).
 48. Moreb, J. S., Muhoczy, D., Ostmark, B. & Zucali, J. R. RNAi-mediated knockdown of aldehyde dehydrogenase class-1A1 and class-3A1 is specific and reveals that each contributes equally to the resistance against 4-hydroperoxycyclophosphamide. *Cancer Chemother. Pharmacol.* **59**, 127–136 (2007).
 49. Wink, S., Hiemstra, S., Herpers, B. & van de Water, B. High-content imaging-based BAC-GFP toxicity pathway reporters to assess chemical adversity liabilities. *Arch. Toxicol.* **91**, 1367–1383 (2017).
 50. Yan, K. & Verbeek, F. J. Segmentation for high-throughput image analysis: Watershed masked clustering. in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **7610**, 25–41 (Springer, Berlin, Heidelberg, 2012).

Nederlandse Samenvatting

Algemene samenvatting

Aldehydedehydrogenases (ALDHs) zetten endogene en exogene aldehyden om in carbonzuren. Aangezien aldehyden reactief zijn en zich kunnen binden aan enzymen en DNA, verminderen ALDHs op deze manier mogelijke celschade (zie **Hoofdstuk 1**)¹⁻³. Daarnaast zijn ALDHs betrokken bij essentiële metabolische cellulaire processen⁴⁻⁷, zoals bijvoorbeeld de vorming van retinolzuur. Retinolzuur is een belangrijk signaallipide, dat verschillende cellulaire en fysiologische processen beïnvloedt door middel van genregulatie.⁸⁻¹³ Retinolzuur wordt gevormd door de omzetting van retinal door de retinaldehydrogenases ALDH1A1, ALDH1A2 en ALDH1A3. Deze eiwitten worden nauwgezet gereguleerd via induceerbare cellulaire expressie patronen en posttranslationale eiwit modificaties.^{14,15}

De huidige methodes stellen onderzoekers in staat om de expressie van specifieke enzymen of de algemene ALDH activiteit te meten. Het is echter niet mogelijk om de activiteit van individuele enzymen te bepalen. Nieuwe strategieën om de ALDH activiteit van specifieke enzymen te visualiseren en kwantificeren in modellen van ziekte en gezondheid zijn daarom gewenst. Het doel van het onderzoek beschreven in dit proefschrift was om nieuwe moleculaire sensoren (*probes*) te ontwikkelen, die gebruikt kunnen worden om de activiteit van ALDHs te bestuderen in gecompliceerde biologische milieus.

In **Hoofdstuk 2**, werden de recente ontwikkelingen samengevat op het gebied van chemische vetachtige *probes* om de lipide biologie te bestuderen. Tevens werden verschillende mogelijkheden voor toekomstig onderzoek besproken.¹⁶ Lipiden voeren talrijke functies uit in de cel. Deze variëren van structurele grondstof en energieopslag tot cellulaire communicatie. De werkingwijze van vele signallipiden is nog steeds ongrijpbaar door toedoen van hun geringe abundantie, hoge lipofiliciteit en intrinsieke instabiliteit. Verschillende chemisch biologische strategieën, zoals fotoaffiniteit- en activiteit-gebaseerde eiwitprofilering (*affinity- and activity-based protein profiling, ABPP and ABPP*), zijn recentelijk ontwikkeld om de rol van lipiden, lipide-eiwit interacties en lipide verwerkende enzymen te onderzoeken.

ABPP is bijvoorbeeld een krachtige techniek gebleken om enzymatische activiteit te profileren in complexe biologische monsters.^{17,18} *ABPP* maakt gebruik van activiteit-gebaseerde *probes*. Deze activiteit-gebaseerde *probes* beschikken over een reactieve groep, waarmee ze covalente kunnen reageren met actieve enzymen. Visualisatie van de doeleiwitten vindt plaats door het ligeren van een label, zoals een fluorofoor of een affiniteitslabel, aan het *probe*-eiwit complex door middel van ‘click’-chemie. De covalent gebonden enzymen kunnen na ligatie in kaart gebracht worden met behulp van respectievelijk gel-analyse (SDS-PAGE) of massaspectrometrie. Comparatieve en competitieve *ABPP* zijn de twee toepassingen die voornamelijk gebruikt worden in dit onderzoek. Comparatieve *ABPP* vergelijkt de activiteitsprofielen van verschillende biologische toestanden, terwijl competitieve *ABPP* gebruikt kan worden om de potentie en het selectiviteitsprofiel van een toekomstig geneesmiddel te bepalen door de activiteitsprofielen van behandelde en onbehandelde cellen te vergelijken.

Activiteit-gebaseerde eiwit profilering van retinaldehydedehydrogenases

Hoofdstuk 3 beschreef het ontwerp en de synthese van een retinal-gebaseerde activiteit-gebaseerde *probe* voor ALDHs. Met behulp van een eerder gerapporteerde kristalstructuur van ALDH1A1 werd de optimale locatie op het retinal molecuul bepaald voor de toevoeging van een alkyn groep, die in een later stadium gebruikt kan worden om visualisatie labels aan te bevestigen. Tevens werd de aldehyde groep veranderd in een vinyl keton.

Dit leidde tot het ontwerp van **LEI-945**, welke gesynthetiseerd werd door middel van een convergente synthese gebruikmakend van een Wittig reactie.¹⁹ Activiteit-gebaseerde *probe* **LEI-945** werd vervolgens biologisch gevalideerd in **Hoofdstuk 4**. **LEI-945** bleek in staat de enzymen ALDH1A1, ALDH1A2 en ALDH1A3 te visualiseren door middel van een covalente interactie met hun katalytische cysteïne. Vervolgens werd aangetoond dat de *probe* in staat was om ook endogeen tot expressie gebrachte ALDHs, waaronder ALDH2, ALDH3A2 en ALDH3B1, te visualiseren in levende cellen behorende tot de longkanker cellijn A549. **LEI-945** werd ook gebruikt om de potentie en selectiviteitsprofielen van de ALDH remmer, 4-diethylaminobenzaldehyde (DEAB), en het natuurlijke substraat, retinal, te bepalen door middel van de competitieve *ABPP* methode in levende A549 cellen.

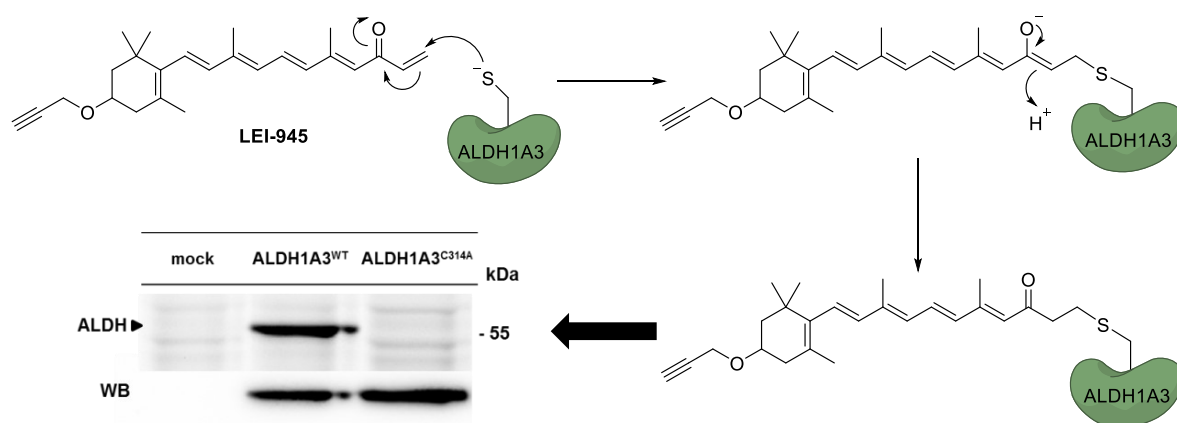


Fig. 9.1 | Visualisatie van de covalente interactie van LEI-945 met ALDH1A3. De reactieve groep van **LEI-945** reageert covalent met de katalytische cysteïne van ALDH1A3. De gevormde thio-ether kan niet gehydrolyseerd worden en vormt een onomkeerbare covalente interactie tussen het enzym ALDH1A3 en de activiteit-gebaseerde *probe* **LEI-945**. Visualisatie vindt plaats door middel van koppeling met de fluorofor AlexaFluor-647 en SDS-PAGE, dit laat een duidelijke band zien rond 55 kDa. Door de katalytische cysteïne te veranderen in een alanine verdwijnt de interactie tussen **LEI-945** en het enzym volledig.

Het in kaart brengen van ALDH activiteit in borstkanker subtypes

In **Hoofdstuk 5** werd **LEI-945** gebruikt om de activiteitsprofielen van ALDHs te vergelijken in verschillende borstkanker celtypes door middel van comparatieve *ABPP*. De verhoogde expressie van bepaalde ALDHs is eerder in verband gebracht met therapie bestendige kanker.²⁰ ALDH1A3 wordt bijvoorbeeld geassocieerd met een verminderd klinisch resultaat in borstkanker.^{21–25} De ALDEFLUOR techniek wordt routinematig gebruikt om de algemene cellulaire ALDH activiteit te bepalen, maar was niet in staat om het vermogen van borstkanker cellijnen om retinolzuur te produceren, te voorspellen.

Door gebruik te maken van **LEI-945** konden specifieke ALDH activiteitsprofielen gemaakt worden voor elke cellijn. Deze profielen konden vervolgens gebruikt worden om de verschillen in retinal omzettingsvermogen te verklaren gebaseerd op de mate waarin actieve ALDH1A1 en ALDH1A3 enzymen aanwezig waren in de verschillende cellijnen.

Interessant genoeg vertoonde de borstkanker cellijn SK-BR-3 de hoogste algemene cellulaire ALDH activiteit, maar was deze cellijn slechts in staat om een relatief kleine hoeveelheid retinal om te zetten. Op basis van het ALDH activiteitsprofiel kon de hoge ALDH activiteit verklaard worden door een exceptioneel hoge aanwezigheid van actieve ALDH2 enzymen, terwijl er slechts een lage mate van ALDH1A3 activiteit werd waargenomen. Deze bevinding suggereert dat de menselijke variant van ALDH2 geen biologisch relevante rol heeft in vitamine A metabolisme.

LEI-945 werd vervolgens gebruikt in een competitieve *ABPP* methode om het selectiviteitsprofiel van NCT-505²⁶, een vergevorderde selectieve ALDH1A1 remmer, te bepalen en interactie van de remmer van het doeleiwit (*target engagement*) aan te tonen in de MDA-MB-468 borstkanker cellijn. Vergelijking van de activiteitsprofielen van behandelde en onbehandelde cellen toonde aan dat NCT-505 niet alleen met ALDH1A1, maar ook met ALDH1A3 een interactie aan ging. Hierdoor werd de toegevoegde waarde van **LEI-945** in het ontwikkelingsproces voor selectieve ALDH remmers aangetoond.

De fysiologische effecten van ALDH1A1 en ALDH1A3 remming werden verder bestudeerd door de cellulaire ATP niveaus te bepalen en verschillende microscopische technieken waarbij levende cellen over tijd gevolgd kunnen worden (*live cell imaging*) te gebruiken. NCT-505 verminderde de cellulaire ATP niveaus en het mitochondriaal membraan potentieel.

Door middel van de FUCCI techniek werd bepaald dat behandelde cellen vastzaten in de G₁ fase van de celcyclus. Een significante toename van propidium iodide kleuring duidde erop dat behandelde cellen bezweken volgens de necrotische celdood route. Een combinatie therapie van NCT-505 met de glutathione peroxidase 4 (GPX4) remmer, RSL3^{27,28}, had een therapeutisch synergistisch effect.

Visualisatie van de rol van vitamine A metabolisme in immuun homeostase

Hoofdstuk 6 beschreef de zoektocht voor de specifieke retinaldehydedehydrogenases betrokken bij de omzetting van retinal in retinolzuur in de dunne darmen van muizen. Retinolzuur speelt een belangrijke rol in immuun homeostase. Het is verantwoordelijk voor de activatie en rekrutering van dendritische, T- en B-cellen.^{29,30} Comparatieve *ABPP* met **LEI-945** werd toegepast op darmcellen afkomstig van muizen die in het bezit waren van het ALDH1A1 enzym en muizen waarin het ALDH1A1 eiwit uit het genetisch materiaal verwijderd was (*knockout*). Op deze manier werd ALDH1B1 geïdentificeerd als een mogelijk additioneel ALDH enzym dat in staat is om retinolzuur te produceren in intestinale epitheliale cellen. *Aldh1a1^{-/-}/Aldh1b1^{-/-}* dubbele *knockout* muizen lieten een pro-inflammatoir fenotype zien. Dit duidde op de betrokkenheid van zowel ALDH1A1, als ALDH1B1 in de regulering van immuun homeostase.

Nieuw chemische gereedschappen om de interactie tussen vitamine A en eiwitten te bestuderen werden beschreven in **Hoofdstuk 7**. Deze klikbare vitamines bleken biologisch identiek aan hun endogene tegenhangers en werden gebruikt in *A₁BPP* toepassingen zonder de toevoeging van een fotoreactieve groep.

Ontwikkeling van een breedspectrum ALDH probe

Hoofdstuk 8 beschreef de ontwikkeling van een breedspectrum activiteit-gebaseerde *probe*, **STA-55**, voor de familie van ALDHs. Deze *probe* is gebaseerd op de eerder gerapporteerde ALDH remmer Aldi-2³¹ en is gekarakteriseerd als een breedspectrum *probe* voor de ALDH familie, die naast de door selectieve *probe* **LEI-945** gedetecteerde ALDHs ook ALDH1B1 en ALDH3A1 verrijkte in levende A549 cellen. De breedspectrum *probe* is vervolgens gebruikt om de potentie en selectiviteitsprofielen te bepalen drie ALDH remmers door middel van competitieve *ABPP*.

Dit leidde tot de ontdekking van verdere onbedoelde eiwit vinding van de ALDH1A1 remmer NCT-505²⁶ en onthulde ook breedspectrum remming van niet-doeleiwitten door de ALDH3A1 remmer CB7.³² Een verhoogde activiteit van ALDH1A1 of ALDH3A1 is een biologische indicatie van cyclofosfamide resistentie in kanker.³³ Aangezien **STA-55** in staat is zowel de activiteit van ALDH1A1, als ALDH3A1 te visualiseren, kan deze *probe* potentieel gebruikt worden om therapeutische effectiviteit te voorspellen door de activiteit van deze enzymen te kwantificeren in patiëntmateriaal. **STA-55** kan verder toegepast worden in het medicijnontwikkelingsproces door de selectiviteitsprofielen en *target engagement* van nieuwe ALDH remmers vast te stellen.

Tot slot

Dit proefschrift beschreef de ontwikkeling van diverse nieuwe chemische *tools*, die gebruikt kunnen worden om ALDHs te onderzoeken door middel van *ABPP* en *A_fBPP*. Toepassingen waarvoor de retinal-gebaseerde *probe*, **LEI-945**, gebruikt is, omvatten het in kaart brengen van de activiteitsprofielen van ALDHs in kanker en het visualiseren van vitamine A metabolisme in de dunne darmen. Naast deze op maat gemaakte retinal *probe* werd de synthese van klikbare vitamine A analoga en de ontwikkeling van een breedspectrum ALDH *probe* beschreven. Gezamenlijk geven deze chemische *tools* onderzoekers de beschikking over een uitgebreide chemische gereedschapskist om ALDH enzymen te kunnen bestuderen. De door het gebruik van deze *tools* behaalde resultaten laten de toegevoegde waarde zien van de toepassing van substraat-gebaseerde *probes* in het onderzoeken van pathologisch relevante enzymactiviteiten. Tevens belichten zij de algemene kracht van eiwit profilering in het ontdekken van nieuwe biologische inzichten en het nut hiervan in het medicijnontwikkelingsproces.

Referenties

1. Koppaka, V. *et al.* Aldehyde Dehydrogenase Inhibitors: a Comprehensive Review of the Pharmacology, Mechanism of Action, Substrate Specificity, and Clinical Application. *Pharmacol. Rev.* **64**, 520–539 (2012).
2. Duester, G., Mic, F. A. & Molotkov, A. Cytosolic retinoid dehydrogenases govern ubiquitous metabolism of retinol to retinaldehyde followed by tissue-specific metabolism to retinoic acid. in *Chemico-Biological Interactions* **143–144**, 201–210 (Elsevier, 2003).
3. Vasilioi, V., Pappa, A. & Petersen, D. R. Role of aldehyde dehydrogenases in endogenous and xenobiotic metabolism. *Chem. Biol. Interact.* **129**, 1–19 (2000).
4. Kurys, G., Ambroziak, W. & Pietruszkos, R. Human Aldehyde Dehydrogenase. *J. Biol. Chem.* **264**, 4715–4721 (1989).
5. Kutzbach, C. & Stokstad, E. L. R. Mammalian methylenetetrahydrofolate reductase Partial purification, properties, and inhibition by S-adenosylmethionine. *BBA - Enzymol.* **250**, 459–477 (1971).
6. Chern, M. K. & Pietruszko, R. Human aldehyde dehydrogenase E3 isozyme is a betaine aldehyde dehydrogenase. *Biochem. Biophys. Res. Commun.* **213**, 561–568 (1995).
7. Jackson, B. *et al.* Update on the aldehyde dehydrogenase gene (ALDH) superfamily. *Hum. Genomics* **5**, 283–303 (2011).
8. Zile, M. H. Vitamin A and Embryonic Development: An Overview. *J. Nutr.* **128**, 455–458 (1998).
9. Wiseman, E. M., Bar-El Dadon, S. & Reifen, R. The vicious cycle of vitamin a deficiency: A review. *Crit. Rev. Food Sci. Nutr.* **57**, 3703–3714 (2017).
10. Mora, J. R., Iwata, M. & Andrian, U. H. Von. Vitamin effects on the immune system. *Nat. Rev. Immunol.* **8**, 685–698 (2008).
11. Blomhoff, H. K. *et al.* Vitamin A is a key regulator for cell growth, cytokine production, and differentiation in normal B cells. *J. Biol. Chem.* **267**, 23988–23992 (1992).
12. Carlberg, C. Lipid soluble vitamins in gene regulation. *BioFactors* **10**, 91–97 (1999).
13. Balmer, J. E. & Blomhoff, R. Gene expression regulation by retinoic acid. *J. Lipid Res.* **43**, 1773–808 (2002).
14. Wang, J. *et al.* Phosphorylation-dependent regulation of ALDH1A1 by Aurora kinase A: Insights on their synergistic relationship in pancreatic cancer. *BMC Biol.* **15**, 1–22 (2017).
15. Zhao, D. *et al.* NOTCH-induced aldehyde dehydrogenase 1A1 deacetylation promotes breast cancer stem cells. *J. Clin. Invest.* **124**, 5453–5465 (2014).
16. Koenders, S. T. A., Gagestein, B. & van der Stelt, M. Opportunities for lipid-based probes in the field of immunology. in *Current Topics in Microbiology and Immunology* **420**, 283–319 (Springer, Cham, 2018).
17. Serwa, R. & Tate, E. W. Activity-based profiling for drug discovery. *Chemistry and Biology* **18**, 407–409 (2011).
18. Cravatt, B. F., Wright, A. T. & Kozarich, J. W. Activity-Based Protein Profiling: From Enzyme Chemistry to Proteomic Chemistry. *Annu. Rev. Biochem.* **77**, 383–414 (2008).
19. Koenders, S. T. A. *et al.* Development of a Retinal-Based Probe for the Profiling of Retinaldehyde Dehydrogenases in Cancer Cells. *ACS Cent. Sci.* **5**, 1965–1974 (2019).
20. Croker, A. K. & Allan, A. L. Inhibition of aldehyde dehydrogenase (ALDH) activity reduces chemotherapy and radiation resistance of stem-like ALDH hiCD44 + human breast cancer cells. *Breast Cancer Res. Treat.* **133**, 75–87 (2012).
21. Qiu, Y. *et al.* The expression of aldehyde dehydrogenase family in breast cancer. *J. Breast Cancer* **17**, 54–60 (2014).
22. Marcato, P. *et al.* Aldehyde Dehydrogenase Activity of Breast Cancer Stem Cells Is Primarily Due To Isoform ALDH1A3 and Its Expression Is Predictive of Metastasis. *Stem Cells* **29**, 32–45 (2011).
23. Sládek, N. E., Kollander, R., Sreerama, L. & Kiang, D. T. Cellular levels of aldehyde dehydrogenases (ALDH1A1 and ALDH3A1) as predictors of therapeutic responses to cyclophosphamide-based chemotherapy of breast cancer: A retrospective study. *Cancer Chemother. Pharmacol.* **49**, 309–321 (2002).
24. Tomita, H., Tanaka, K., Tanaka, T. & Hara, A. Aldehyde dehydrogenase 1A1 in stem cells and cancer. *Oncotarget* **7**, 11018–32 (2016).
25. De Beca, F. F. *et al.* Cancer stem cells markers CD44, CD24 and ALDH1 in breast cancer special histological types. *J. Clin. Pathol.* **66**, 187–191 (2013).
26. Yang, S. M. *et al.* Discovery of Orally Bioavailable, Quinoline-Based Aldehyde Dehydrogenase 1A1 (ALDH1A1) Inhibitors with Potent Cellular Activity. *J. Med. Chem.* **61**, 4883–4903 (2018).
27. Dixon, S. J. *et al.* Ferroptosis: An iron-dependent form of nonapoptotic cell death. *Cell* **149**, 1060–1072 (2012).
28. Yang, W. S. *et al.* Regulation of ferroptotic cancer cell death by GPX4. *Cell* **156**, 317–331 (2014).

NEDERLANDSE SAMENVATTING

29. Svensson, M. *et al.* CCL25 mediates the localization of recently activated CD8 $\alpha\beta$ +lymphocytes to the small-intestinal mucosa. *J. Clin. Invest.* **110**, 1113–1121 (2002).
30. Hamann, A., Andrew, D. P., Jablonski-Westrich, D., Holzmann, B. & Butcher, E. C. Role of alpha 4-integrins in lymphocyte homing to mucosal tissues in vivo. *J. Immunol.* **152**, 3282–93 (1994).
31. Khanna, M. *et al.* Discovery of a novel class of covalent inhibitor for aldehyde dehydrogenases. *J. Biol. Chem.* **286**, 43486–43494 (2011).
32. Parajuli, B., Fishel, M. L. & Hurley, T. D. Selective ALDH3A1 inhibition by benzimidazole analogues increase mafosfamide sensitivity in cancer cells. *J. Med. Chem.* **57**, 449–461 (2014).
33. Moreb, J. S., Muhoczy, D., Ostmark, B. & Zucali, J. R. RNAi-mediated knockdown of aldehyde dehydrogenase class-1A1 and class-3A1 is specific and reveals that each contributes equally to the resistance against 4-hydroperoxycyclophosphamide. *Cancer Chemother. Pharmacol.* **59**, 127–136 (2007).

List of Publications

Development of a Retinal-Based Probe for the Profiling of Retinaldehyde Dehydrogenases in Cancer Cells

S.T.A. Koenders, L.S. Wijaya, M.N. Erkelens, A.T. Bakker, V.E. van der Noord, E.J. van Rooden, L. Burggraaff, P.C. Putter, E. Botter, K. Wals, H. van den Elst, H. den Dulk, B.I. Florea, B. van de Water, G.J.P. van Westen, R.E. Mebius, H.S. Overkleeft, S.E. Le Dévédec, M. van der Stelt.

ACS Central Science, **2019**, 5, 1965-1974.

STA-55, An Easily Accessible Broadspectrum Activity-Based Aldehyde Dehydrogenase Probe

S.T.A. Koenders, E.J. van Rooden, H. van den Elst, B.I. Florea, H.S. Overkleeft, M. van der Stelt

ChemBioChem, **2020**, 21, 1911-1917.

Opportunities for Lipid-Based Probes in the Field of Immunology

S.T.A. Koenders, B. Gagestein, M. van der Stelt.

Current Topics in Microbiology and Immunology, **2018**, 420, 283-319.

Curriculum vitae

Sebastiaan Theodorus Antonius Koenders was born on April 23, 1991 in Nijmegen, the Netherlands. He obtained his high school diploma from the Gymnasium Apeldoorn in 2009 and won the national chemistry Olympiad in that same year. Based on his interest in both biology and chemistry he started the bachelor Biopharmaceutical Sciences at Leiden University. He obtained his bachelor's degree in 2013 *cum laude*. During his bachelor he competed in rowing at a national level as a member of the Royal Student Rowing Society Njord. After competing in rowing for three years he was elected as the president of the rowing club, presiding over the 139th board of Njord from September 2012 until August 2013. He continued his academic education in September 2013 with a research master in Chemistry at Leiden University and obtained his master's degree *cum laude* in 2015.

During his master program he did a research internship in the Bio-Organic Synthesis group under supervision of Prof. Dr. G.A. van der Marel and Dr. J.D.C. Codée titled “*Synthesis of glycerolphosphate probes for the characterization of pilin phosphotransferase B*”. This research focused on the synthesis of novel chemical tools based on glycerolphosphate to study the role of pilin phosphotransferase B in the pathogenesis of *Neisseria meningitidis*. His second research internship was performed in the lab of Prof. Dr. H. Waldmann at the Max-Planck Institute of Molecular Physiology in Dortmund supported by an Erasmus grant. The aim of this study was *The development of an asymmetric [4+2] annulation of ketimines and allenates to enable structure-activity relationship studies around the indoloquinolizine scaffold*.

In September 2015 he started his doctoral studies in the Molecular Physiology group under supervision of Prof. Dr. H. S. Overkleeft and Prof. Dr. M. van der Stelt. The research described in **Chapter 5** was performed in collaboration with Dr. S.E. le Dévédec, faculty member of the Leiden Academic Centre for Drug Research, and the work presented in **Chapter 6** was performed in collaboration with Prof. Dr. R. Mebius from the department of Molecular Cell Biology and Immunology at the Amsterdam UMC. The Institute for Chemical Immunology and the Oncode Institute are kindly acknowledged for the funding of this PhD research.

Posters on parts of the research were presented at various national and international conferences: the annual Chemical Immunology Congress (Amsterdam, the Netherlands, 2016, 2017 and 2019); CHAINS (Veldhoven, the Netherlands, 2017 and 2018); and the annual Activity-Based Protein Profiling meeting (Oxford, United Kingdom, 2018). Oral presentations were held at: the annual Chemical Immunology Congress (Amsterdam, the Netherlands, 2018); the EMBO workshop Chemical Biology (Heidelberg, Germany, 2018); the Figon Dutch Medicine Days (Ede, the Netherlands, 2018); the annual Activity-Based Protein Profiling meeting (Leuven, Belgium, 2019); and the Oncode Tumor Heterogeneity meeting (Veldhoven, the Netherlands, 2019).