

Special Issue

STA-55, an Easily Accessible, Broad-Spectrum, Activity-Based Aldehyde Dehydrogenase Probe

Sebastiaan T. A. Koenders,^[a] Eva J. van Rooden,^[a] Hans van den Elst,^[b] Bogdan I. Florea,^[b] Herman S. Overkleef,^[b] and Mario van der Stelt^{*[a]}

Aldehyde dehydrogenases (ALDHs) convert aldehydes into carboxylic acids and are often upregulated in cancer. They have been linked to therapy resistance and are therefore potential therapeutic targets. However, only a few selective and potent inhibitors are currently available for this group of enzymes. Competitive activity-based protein profiling (ABPP) would aid

the development and validation of new selective inhibitors. Herein, a broad-spectrum activity-based probe that reports on several ALDHs is presented. This probe was used in a competitive ABPP protocol against three ALDH inhibitors in lung cancer cells to determine their selectivity profiles and establish their target engagement.

Introduction

Aldehyde dehydrogenases (ALDHs) perform important metabolic roles in eukaryotic cells because they convert endogenous and exogenous aldehydes into carboxylic acids.^[1] They are involved in many metabolic pathways and pathologies.^[1–7] ALDHs 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] Inhibitors have been developed for both these enzymes: NCT-505^[10] and CB7,^[11] respectively.

A recent study has shown that only a fraction of the commonly used and recently developed ALDH1A1 inhibitors showed cellular activity.^[12] The method described in this study is tailored towards ALDH1A1 activity, but, so far, no methods exist to study target engagement of all ALDHs present in a biological sample simultaneously. Activity-based protein profiling (ABPP) is a powerful technique capable of determining the selectivity profile of a drug candidate against an enzyme family in their native cellular environment.^[13] This technique relies on activity-based probes (ABPs), which are tailored to the enzyme family of interest and that react through its electrophilic warhead with a nucleophile within the active site of the

enzyme. We recently described the development of LEI-945, which is a first-in-class retinal-based probe for the ABPP of retinaldehyde dehydrogenases, ALDH1A1, ALDH1A2, and ALDH1A3.^[14] LEI-945 enabled the quantification of ALDH isozyme activities in a panel of cancer cell lines through both fluorescence and chemical proteomic approaches. The probe was superior to the widely used ALDEFLUOR assay in explaining the ability of breast cancer cells to produce *all-trans* retinoic acid. In addition, it revealed the cellular selectivity profile of NCT-505, an advanced ALDH1A1 inhibitor.^[14] However, the synthesis of LEI-945 is complex and challenging. We hypothesized that, by modifying the reported covalent pan-ALDH inhibitor Aldi-2,^[15] an easily accessible, broad-spectrum probe for the ALDH family could be made (Figure 1).

Herein, we describe the design, synthesis, biological validation, and application of STA-55 as a broad-spectrum probe for the family of ALDHs. Chemical proteomics showed that STA-55

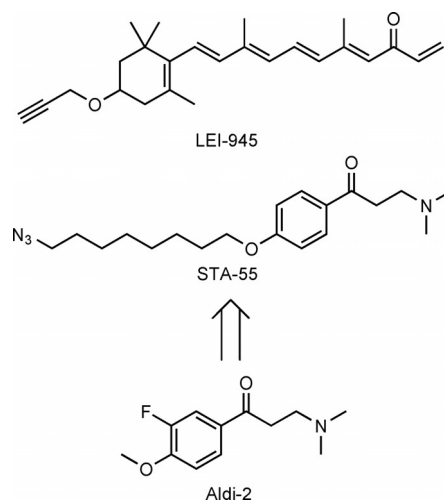


Figure 1. Chemical structures of LEI-945, STA-55, and Aldi-2.

[a] S. T. A. Koenders, Dr. E. J. van Rooden, Prof. Dr. M. van der Stelt
Department of Molecular Physiology
Leiden Institute of Chemistry, Leiden University
Einsteinweg 55, 2333 CC Leiden
(The Netherlands)
E-mail: m.van.der.stelt@chem.leidenuniv.nl

[b] H. van den Elst, Dr. B. I. Florea, Prof. Dr. H. S. Overkleef
Department of Bio-Organic Synthesis
Leiden Institute of Chemistry, Leiden University
Einsteinweg 55, 2333 CC Leiden
(The Netherlands)

Supporting information and the ORCID identification numbers for the authors of this article can be found under <https://doi.org/10.1002/cbic.201900771>.

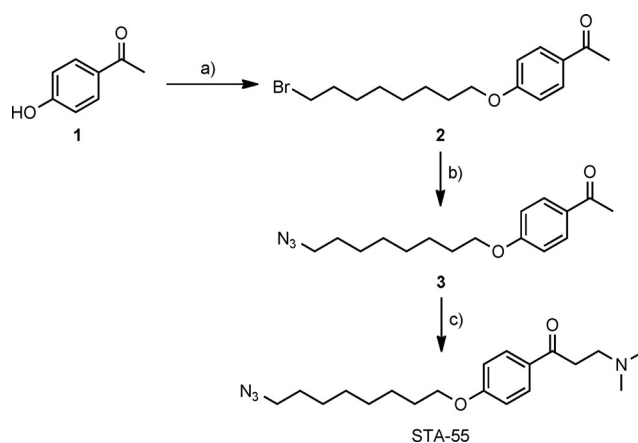
This article is part of a Special Collection on Chemical Proteomics and Metabolomics. To view the complete collection, visit our homepage.

could be used to enrich both ALDH1A1 and ALDH3A1 in A549 lung cancer cells. Competitive chemical proteomics with STA-55 was performed to determine the selectivity profiles of the known ALDH inhibitors DEAB,^[16] NCT-505,^[10] and CB7.^[11] Our results show 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

An ABP consists of a reactive group (termed “warhead” and often an electrophile), a recognition scaffold, and a ligation handle. Aldi-2 already incorporates a masked warhead, which, after liberation, reacts with the catalytic cysteine of the ALDH enzyme.^[15] Based on the reported structure–activity relationship studies,^[15] we developed a synthetic strategy for the introduction of an azide ligation handle at the position of the methoxy substituent (Scheme 1). The synthesis of probe STA-55 started from commercially available 4-hydroxyacetone (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 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.

To determine whether STA-55 interacted with the catalytic cysteine of an ALDH enzyme (Figure 2A), 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 copper(I)-catalyzed azide–alkyne [2+3] cycloaddition (CuAAC) ligation with a Cy5-alkyne resulted in a fluorescent band around 55 kDa in the wild-type sample after SDS-PAGE and fluorescent scanning, which suggested the ability of STA-55 to interact with ALDHs (Figure 2B).



Scheme 1. Synthesis of STA-55. a) 1,8-Dibromooctane, K₂CO₃, acetone, 56 °C, 18 h, 82%; b) NaN₃, DMF, 80 °C, 18 h, 93%; c) (CH₃)₂NH·HCl, paraformaldehyde, HCl, EtOH, 78 °C, 18 h, 21%.

The disappearance of this band in the catalytically inactive C314A mutant indicates that STA-55 has reacted covalently and irreversibly with the catalytic nucleophile of ALDH1A3.^[17]

To determine which members of the ALDH family were targeted by ABP STA-55, we performed a label-free quantitative proteomics experiment in the non-small-cell lung cancer cell line A549, which expressed 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-labeled proteins were subsequently enriched by using streptavidin beads and several washing steps to remove unbound proteins. On-bead tryptic digestion was followed by protein identification and quantification by means of mass spectrometry. In this way, 259 significantly enriched proteins were identified (fold-change > 2.0; *p* value < 0.05; Fig-



Figure 2. ALDH enzyme labeling with broad-spectrum probe STA-55. A) Proposed warhead deprotection mechanism and interaction of ABP with catalytic cysteine in the pocket of ALDH enzyme. B) Labeling of transiently transfected ALDH1A3^{WT} and ALDH1A3^{C314A} with STA-55 (1 μ M, 1 h) in U2OS cells.

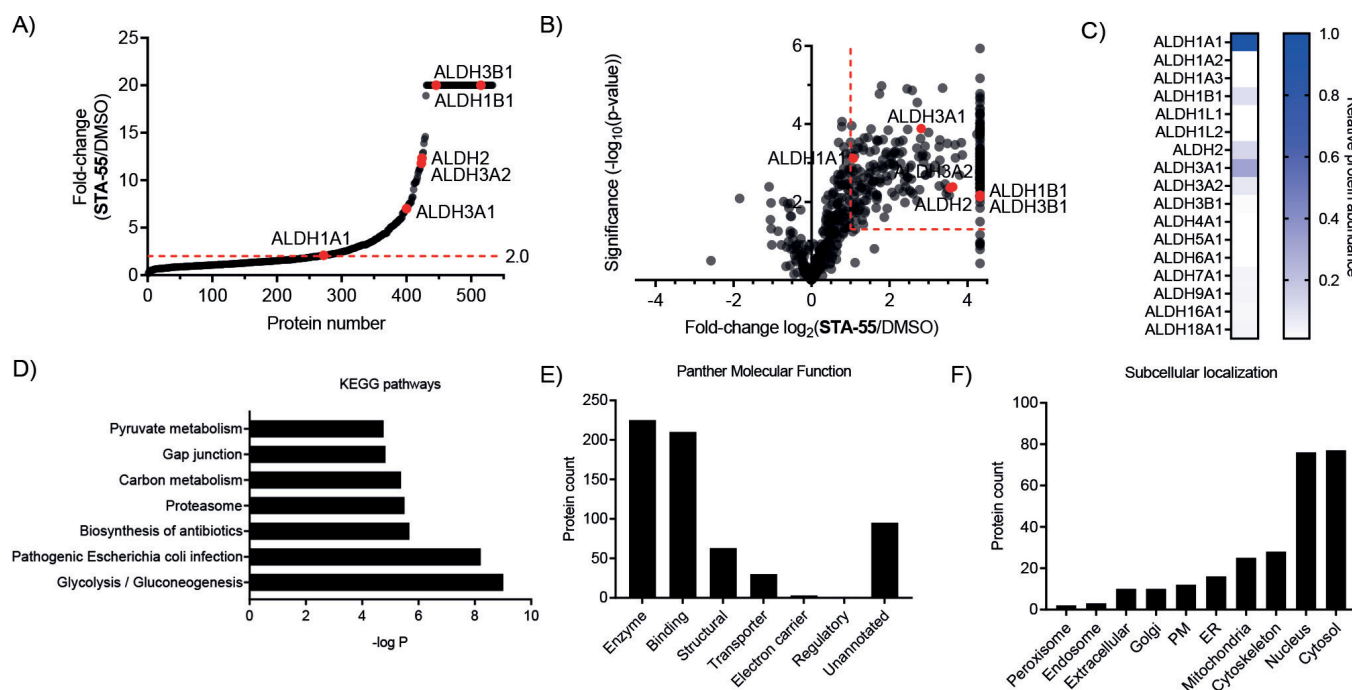


Figure 3. Chemical proteomics with broad-spectrum probe STA-55. A) Fold-change (STA-55/DMSO) plot for total proteins identified in a chemical proteomics experiment with probe STA-55 (10 μM). Red lines indicate the threshold fold-change of twofold enrichment and the maximum fold-change is set at 20. Red dots represent significantly enriched ALDH enzymes. B) Volcano plot for total proteins identified in a chemical proteomics experiment with probe STA-55 (10 μM). Red lines indicate threshold values marking significantly enriched proteins. Red dots represent significantly enriched ALDH enzymes. For A) and B), data are from $N=3$ experiments (biological replicates). C) Heat map showing relative ALDH protein levels in the non-small-cell lung cancer A549 cell line by using deep proteome data from the Expression Atlas. D) Top seven pathways enriched in the group of significantly enriched proteins, as determined by screening of the KEGG database. E) Biological functions attributed to significantly enriched proteins by the Panther database. F) Subcellular localization of significantly enriched proteins, as annotated by the UniProt database.

ure 3A, B). Six identified proteins belonged to the ALDH family: ALDH1A1, ALDH1B1, ALDH2, ALDH3A1, ALDH3A2, and ALDH3B1. Comparison with the Expression Atlas^[20,21] showed that all ALDH enzymes expressed in the A549 cell line were detected by the STA-55 broad-spectrum ALDH probe (Figure 3C), whereas the more specific retinal-based LEI-945 did not detect ALDH1B1 and ALDH3A1 in this cell line.^[14]

The proteins significantly enriched by STA-55 were further analyzed by using the KEGG,^[22] Panther,^[23] and UniProtKB databases.^[24] Proteins in glycolysis/gluconeogenesis, biosynthesis of antibiotics, carbon metabolism (ALDHs, lactate dehydrogenases, and phosphofructokinases), pathogenic *Escherichia coli* infection, gap junction (tubulins), and the proteasome were identified from the KEGG pathway database (Figure 3D). The majority of proteins identified possess enzyme activity or have specific protein interaction partners (Figure 3E). The subcellular localization showed that the majority of enriched proteins were derived from the cytosol and nucleus (Figure 3F). Taken together, our data show that STA-55 can be used as a broad-spectrum ABP for ALDHs, including the known cancer resistance biomarkers, ALDH1A1 and ALDH3A1.^[8,9]

Having established that STA-55 significantly enriches a broad range of ALDHs from A549 lung cancer cell extracts, we performed in situ selectivity profiling of three known ALDH inhibitors (DEAB, NCT-505, and CB7; Figure 4A). DEAB is a pan-ALDH inhibitor regularly used as a control compound in ALDH activity assays.^[12,16] NCT-505 is a recently reported ALDH1A1 inhibi-

tor,^[10] for which we showed, with LEI-945, that it inhibited ALDH1A3 to a similar extent as that of ALDH1A1 in living cells.^[14]

CB7 is reported as a selective ALDH3A1 inhibitor.^[11] A549 cells were preincubated for 30 min with inhibitor (10 μM), after which time STA-55 (1 μM) was added and incubated for 1 h. Subsequently, the chemical proteomics protocol was followed. The ALDH selectivity profiles obtained are shown in Figure 4B. DEAB inhibited ALDH1A1, ALDH2, ALDH3A1, and ALDH3A2, which was in agreement with previously reported results (Figure 4C).^[16] In addition to inhibition of ALDH1A1 and ALDH1A3,^[14] STA-55 revealed that NCT-505 inhibited ALDH3A1. CB7 also appeared to be more promiscuous than previously reported, inhibiting ALDH3A1, ALDH1A1, ALDH2, and ALDH3A2. These results challenge the reported selectivity for ALDH1A1^[10] and ALDH3A1,^[11] respectively. For both inhibitors, these selectivity claims are based on biochemical substrate assays with purified enzymes. However, the activity of isolated enzymes in a biochemical assay does not necessarily reflect their activity in a cellular environment. We argue that the selectivity profiles derived from 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^[11] 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 also has therapeutic

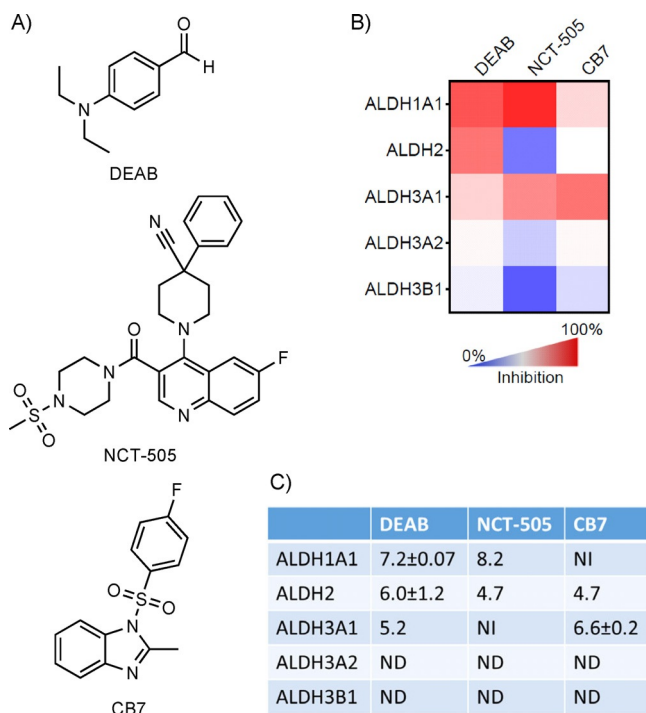


Figure 4. Competitive ABPP of ALDH inhibitors with broad-spectrum probe STA-55. A) Chemical structures of ALDH inhibitors used in this study. B) Heat map showing the selectivity profile of pan-ALDH inhibitor DEAB, ALDH1A1-selective inhibitor NCT-505, and ALDH3A1-selective inhibitor CB7, as determined by competitive ABPP with STA-55 (1 μ M). $N=4$ experiments (biological replicates). C) Reported pIC_{50} values of DEAB,^[12,16] NCT-505,^[10] and CB7^[11,12] for the ALDHs identified with STA-55. If divergent values are reported, data are represented as mean values \pm standard deviation (SD); NI: no inhibition, ND: no data.

tic potential as a dual inhibitor of these cancer therapy resistance biomarkers. Our results show the ability of STA-55 to be used for target identification and engagement studies with competitive ABPP of new drug candidates for ALDHs. STA-55 is easily accessible compared with synthetically challenging LEI-945.^[14] STA-55 is also capable of enriching all ALDH enzymes present in the A549 lung cancer cells, including the cancer biomarker ALDH3A1, whereas LEI-945 is more specifically targeted towards the retinaldehyde dehydrogenases. STA-55 is, therefore, a complementary broad-spectrum ALDH probe, which can conveniently be used for competitive ABPP applications.

Conclusions

We have described the design, development, and biological validation of STA-55, which is a broad-spectrum ABP for the family of ALDHs. Because certain ALDHs are often upregulated in cancer and confer therapy resistance, this probe enables the identification and quantification of these cancer biomarkers by using chemical proteomics. We furthermore show the ability of STA-55 to be used in cellular target identification and target engagement studies and propose that our probe may be used to thereby facilitate early drug-discovery studies aimed at the identification of selective and tissue-permeable ALDH inhibitors.

Experimental Section

General: All reactions were performed by 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.^[10,11] Tetrahydrofuran and *N,N*-dimethylformamide were stored over 4 Å molecular sieves before use. All moisture-sensitive reactions were performed under a nitrogen atmosphere. TLC analysis was performed by using Merck aluminum sheets (TLC silica gel 60/Kieselguhr F₂₅₄). Compounds were visualized by using a solution of KMnO₄ (7.5 g), K₂CO₃ (50 g), and 10% NaOH (6 mL) in H₂O (1 L). Column chromatography was performed by using Screening Device B.V. silica gel (particle size 40–63 μ m, pore diameter of 60 Å) with the indicated eluents. ¹H and ¹³C NMR spectra were recorded on Bruker AV-400 (400 and 101 MHz, respectively) or Bruker AV-500 MHz (500 and 125 MHz, respectively) spectrometers by using CDCl₃ as the solvent. Chemical shifts are reported relative to the residual solvent signal or tetramethylsilane. Coupling constants are given in Hz. 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⁻¹, 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) instrument equipped with an electrospray ion source in positive mode (ESI-TOF), injection through a NanoEquity system (Waters), with LeuEnk (m/z 556.2771) as the lock mass. Eluents used were 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).

Synthesis

1-(4-((8-Bromooctyl)oxy)phenyl)ethan-1-one (2): 4-Hydroxyacetophenone (1.0 g, 7.3 mmol) in acetone (20 mL) was added dropwise 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. The reaction mixture was stirred for 18 h at 56 °C. The reaction mixture was then allowed to cool, 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 means of column chromatography (CH₂Cl₂/pentane) afforded compound **2** as a white solid (2.0 g, 6.0 mmol, 82%). R_f (50% CH₂Cl₂ 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 ppm (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 ppm.

1-(4-((8-Azidoctyl)oxy)phenyl)ethan-1-one (3): Sodium azide (0.30 g, 4.6 mmol) was added under Ar to a solution of compound **2** (0.50 g, 1.5 mmol) in DMF (5 mL). The reaction mixture was stirred for 18 h at 80 °C and then allowed to cool. 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 means of column chromatography (pentane/CH₂Cl₂) afforded compound **3** (0.41 g, 1.4 mmol, 93%) as a colorless liquid. R_f (50% CH₂Cl₂ 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 ppm (m, 6H); ^{13}C NMR (100 MHz, CDCl_3): 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 ppm.

1-[4-[(8-Azidooctyl)oxy]phenyl]-3-(dimethylamino)propan-1-one (STA-55): Dimethylamine hydrochloride salt (92 mg, 1.1 mmol), paraformaldehyde (34 mg, 1.1 mmol), and one drop of concentrated HCl were added under Ar to a solution of compound **3** (164 mg, 0.57 mmol) in EtOH (2 mL). The reaction mixture was stirred at 90 °C for 18 h and then allowed to cool. Purification of the reaction mixture by means of column chromatography ($\text{CH}_2\text{Cl}_2/\text{MeOH}$) recovered compound **3** (127 mg, 0.43 mmol, 75%) and afforded STA-55 as a yellow solid (42.2 mg, 0.12 mmol, 21%). R_f (10% MeOH in CH_2Cl_2) = 0.5; ^1H (400 MHz, CDCl_3): 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 ppm (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 ppm; HRMS (ESI): m/z calcd for $\text{C}_{19}\text{H}_{30}\text{O}_2$ [$M+\text{H}$] $^+$: 347.24415; found: 347.24433.

In situ labeling procedure

ALDH plasmids: For the preparation of 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 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 means of sequence analysis at the Leiden Genome Technology Centre.

Cell culture: U2OS cells were grown in Dulbecco's modified eagle medium (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 six-well plate. Cells were transfected by the addition of a 3:1 mixture of polyethylenimine (6 μg) and plasmid DNA (2 μg) in 200 μL serum-free medium per well. The medium was refreshed after 24 h, and after 48 h the cells were used for subsequent assays.

In situ ABPP: Growth medium from cells grown to 70% confluence in a six-well plate was removed and 1 mL serum-free medium containing probe STA-55 (1 or 10 μM , 0.1% DMSO) was added. The cells were then incubated for 1 h. For competitive ABPP, cells were first incubated with vehicle or inhibitor (10 μM , 0.1% DMSO) for 30 min followed by STA-55 (1 μM final concentration) for 1 h. The medium was then removed, the cells were washed with phosphate-buffered saline (PBS; 2 mL) and then harvested in PBS (1 mL) using a cell scraper. The cells were moved to an Eppendorf tube and the suspension was centrifuged for 5 min at 135 g . PBS was removed and the samples were snap frozen in liquid nitrogen 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, 1 \times protease inhibitor cocktail (Roche cOmplete EDTA free)), and incu-

bated on ice (15–30 min). The protein concentration was determined by means of a Quick Start Bradford Protein assay (Bio-Rad). The protein fractions were diluted to a total protein concentration of 1 mg mL^{-1} . 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 per sample, 18 mM in H_2O), 45 mM sodium ascorbate (1.5 μL per sample, 150 mM in H_2O), 1.8 mM tris(3-hydroxypropyl-triazolylmethyl)amine (THPTA; 0.5 μL per sample, 18 mM in DMSO), and 9 μM Cy5-alkyne (0.5 μL per sample, 90 μM in DMSO from Thermo Fischer Scientific). The samples were incubated for 1 h at 37 °C and then SDS-PAGE sample buffer ($4 \times 15 \mu\text{L}$) was added. The samples were denatured at 100 °C for 5 min; 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) by using Cy3 and Cy5 multichannel settings (605/50 and 695/55, filters, respectively) and stained with Coomassie or transferred to 0.2 μm polyvinylidene difluoride membranes by using a Trans-Blot Turbo Transfer system (Bio-Rad) after scanning. Fluorescence was normalized to Coomassie staining or to the FLAG-tag signal and quantified with Image Lab software (Bio-Rad).

Western blotting: Proteins were transferred to 0.2 μm polyvinylidene difluoride membranes by using a 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 h 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 by using an imaging solution (10 mL Luminol, 100 μL ECL enhancer, and 3 μL 30% H_2O_2) and chemiluminescence was visualized by using a ChemiDoc XRS (Bio-Rad) system. The signal was normalized to Coomassie staining and quantified with Image Lab software (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: The protocol was adapted from a previously described procedure.^[26] Cells were treated in situ, harvested, lysed, and adjusted to 1 mg mL^{-1} protein concentration, as described above. An aliquot (250 μL) was taken from each sample and to this freshly prepared click mixture (25 μL), containing 1 mM CuSO_4 (2.5 μL per sample, 100 mM in H_2O), 5 mM sodium ascorbate (1.25 μL per sample, 1 M in H_2O), 0.4 mM THPTA (1 μL per sample, 100 mM in DMSO), 40 μM biotin-alkyne (2.5 μL per sample, 4 mM in DMSO), and MilliQ (17.75 μL per sample), was added. Samples were incubated for 1 h at 37 °C with 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 μL , 6 M urea and 25 mM ammonium bicarbonate), dithiothreitol (DTT; 2.5 μL , 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 to RT and then alkylated by the addition of iodoacetamide (20 μL , 0.5 M) for 30 min at RT in the dark. Addition of SDS (70 μL , 10%, v/v) was followed by heating at 65 °C for 5 min. For each sample, 50% slurry of Avidin-Agarose from egg white (50 μL Sigma-Aldrich) was washed with PBS (3 \times) and transferred in PBS (1 mL) to a 15 mL tube. To this, further PBS (2 mL) was added followed by the corresponding proteome sample. The beads were incubated with the proteome for 2 h at room temperature by 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 with PBS (3×). The beads were then transferred to low-binding Eppendorf tubes and proteins were digested overnight at 37 °C and 950 rpm in digestion buffer (250 µL; 100 mM Tris, 100 mM NaCl, 1 mM CaCl₂, 2% acetonitrile, and 0.5 µg sequencing grade trypsin (Promega)). Digestion was stopped by the addition of formic acid (12.5 µL) and the beads were filtered off by centrifugation (2 min, 600 g) by using a Bio-Spin column (Bio-Rad). Samples were then desalted by using stage tips, collected in low-binding Eppendorf tubes, concentrated by 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 measurements and analysis: Samples were reconstituted in LC-MS sample solution (50 µL; 3% acetonitrile, 0.1% formic acid, 20 fmol µL^{−1} enolase). Samples were then analyzed by 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 µL was loaded on a nanoEASE M/Z Symmetry C₁₈ trap column (particles 5 µm, 100 Å, 180 µm × 20 mm, Waters) with 0.1% formic acid and separated on an nanoEASE M/Z HSS C₁₈ T3 analytical column (particles 1.8 µm, 75 µm × 250 mm, Waters) heated at 80 °C. A multistep gradient running from 5 to 40% acetonitrile, containing 0.1% formic acid, during a 70 min method at 300 nL min^{−1} 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 s 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 by using drift-time specific collision energies. The lock mass was sampled every 30 s. MS raw files were analyzed with ProteinLynx Global SERVER (PLGS, v3.0.3; Waters). The MS^E identification was also performed with PLGS by 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, 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 peptides and quantified in at least three replicates of the positive control (probe treated). Proteins were designated as significantly enriched by the probe if they showed twofold enrichment in quantification value relative to the negative control (vehicle treated) with positive control (probe treated) samples and probability as determined by a Student's *t*-test (< 0.05).

Heat map competitive ABPP analysis: Only significantly enriched ALDH enzymes were selected for analysis. The mean raw label-free quantification (LFQ) intensities from quadruplicate measurements were normalized to DMSO (=0) and maximal LFQ STA-55 (=1) for each protein individually. The heat map was prepared by using Graphpad Prism 7 software (Graphpad Software Inc.).

Acknowledgements

This work was supported by the Institute for Chemical Immunology and the Oncode Institute.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: activity-based protein profiling • aldehyde dehydrogenases • cancer • proteomics • target engagement

- [1] V. Vasiliou, A. Pappa, D. R. Petersen, *Chem.-Biol. Interact.* **2000**, *129*, 1–19.
- [2] C. Kutzbach, E. L. R. Stokstad, *Biochim. Biophys. Acta Enzymol.* **1971**, *250*, 459–477.
- [3] M. K. Chern, R. Pietruszko, *Biochem. Biophys. Res. Commun.* **1995**, *213*, 561–568.
- [4] V. Vasiliou, A. Pappa, *Pharmacology* **2000**, *61*, 192–198.
- [5] V. De Laurenzi, G. R. Rogers, D. J. Hamrock, L. N. Marekov, P. M. Steinert, J. G. Compton, N. Markova, W. B. Rizzo, *Nat. Genet.* **1996**, *12*, 52–57.
- [6] M. T. Geraghty, D. Vaughn, A. J. Nicholson, W. W. Lin, G. Jimenez-Sanchez, C. Obie, M. P. Flynn, D. Valle, C. A. A. Hu, *Hum. Mol. Genet.* **1998**, *7*, 1411–1415.
- [7] K. L. Chambliss, D. D. Hinson, F. Trettel, P. Malaspina, A. Novelletto, C. Jakobs, K. M. Gibson, *Am. J. Hum. Genet.* **1998**, *63*, 399–408.
- [8] J. S. Moreb, D. Muhoczy, B. Ostmark, J. R. Zucali, *Cancer Chemother. Pharmacol.* **2007**, *59*, 127–136.
- [9] A. Emadi, R. J. Jones, R. A. Brodsky, *Nat. Rev. Clin. Oncol.* **2009**, *6*, 638–647.
- [10] S. M. Yang, N. J. Martinez, A. Yasgar, C. Danchik, C. Johansson, Y. Wang, B. Baljinnyam, A. Q. Wang, X. Xu, P. Shah, D. Cheff, X. S. Wang, J. Roth, M. Lal-Nag, J. E. Dunford, U. Oppermann, V. Vasiliou, A. Simeonov, A. Jadhav, D. J. Maloney, *J. Med. Chem.* **2018**, *61*, 4883–4903.
- [11] B. Parajuli, M. L. Fishel, T. D. Hurley, *J. Med. Chem.* **2014**, *57*, 449–461.
- [12] A. Yasgar, S. A. Titus, Y. Wang, C. Danchik, S. M. Yang, V. Vasiliou, A. Jadhav, D. J. Maloney, A. Simeonov, N. J. Martinez, *PLoS One* **2017**, *12*, e0170937.
- [13] R. Serwa, E. W. Tate, *Chem. Biol.* **2011**, *18*, 407–409.
- [14] S. T. A. Koenders, L. S. Wijaya, M. N. Erkelens, A. T. Bakker, V. E. van der Noord, E. J. van Rooden, L. Burggraaf, P. C. Putter, E. Botter, K. Wals, H. van den Elst, H. den Dulk, B. I. Florea, B. van de Westen, R. E. Mebius, H. S. Overkleeft, S. E. Le Dévédec, M. van der Stelt, *ACS Cent. Sci.* **2019**, *5*, 1965–1974.
- [15] M. Khanna, C. H. Chen, A. Kimble-Hill, B. Parajuli, S. Perez-Miller, S. Bas-karan, J. Kim, K. Dria, V. Vasiliou, D. Mochly-Rosen, T. D. Hurley, *J. Biol. Chem.* **2011**, *286*, 43486–43494.
- [16] C. A. Morgan, B. Parajuli, C. D. Buchman, K. Dria, T. D. Hurley, *Chem.-Biol. Interact.* **2015**, *234*, 18–28.
- [17] A. Moretti, J. Li, S. Donini, R. W. Sobol, M. Rizzi, S. Garavaglia, *Sci. Rep.* **2016**, *6*, 35710.
- [18] J. S. Moreb, J. R. Zucali, B. Ostmark, N. A. Benson, *Cytometry Part B* **2007**, *72*, 281–289.
- [19] J. H. Kang, S. H. Lee, D. Hong, J. S. Lee, H. S. Ahn, J. H. Ahn, T. W. Seong, C. H. Lee, H. Jang, K. M. Hong, C. Lee, J. H. Lee, S. Y. Kim, *Exp. Mol. Med.* **2016**, *48*, e272.
- [20] D. B. Bekker-Jensen, C. D. Kelstrup, T. S. Batth, S. C. Larsen, C. Haldrup, J. B. Bramsen, K. D. Sørensen, S. Høyer, T. F. Ørntoft, C. L. Andersen, M. L. Nielsen, J. V. Olsen, *Cell Syst.* **2017**, *4*, 587–599.
- [21] R. Petryszak, M. Keays, Y. A. Tang, N. A. Fonseca, E. Barrera, T. Burdett, A. Füllgrabe, A. M. Fuentes, S. Jupp, S. Koskinen, et al., *Nucleic Acids Res.* **2016**, *44*, 746–752.
- [22] H. Ogata, S. Goto, W. Fujibuchi, M. Kanehisa, *Biosystems* **1998**, *47*, 119–128.
- [23] P. D. Thomas, A. Kejariwal, M. J. Campbell, H. Mi, K. Diemer, N. Guo, I. Ladunga, B. Ulitsky-Lazareva, A. Muruganujan, S. Rabkin, J. A. Vandergriff, O. Doremieux, *Nucleic Acids Res.* **2003**, *31*, 334–341.
- [24] A. Bateman, *Nucleic Acids Res.* **2019**, *47*, 506–515.
- [25] A. C. M. Van Esbroeck, A. P. A. Janssen, A. B. Cognetta, D. Ogasawara, G. Shpak, M. Van Der Kroeg, V. Kantae, M. P. Baggelaar, F. M. S. De Vrij, H. Deng, et al., *Science* **2017**, *356*, 1084–1087.

- [26] E. J. Van Rooden, B. I. Florea, H. Deng, M. P. Baggelaar, A. C. M. Van Esbroeck, J. Zhou, H. S. Overkleeft, M. Van Der Stelt, *Nat. Protoc.* **2018**, *13*, 752–767.
- [27] J. Rappsilber, M. Mann, Y. Ishihama, *Nat. Protoc.* **2007**, *2*, 1896–1906.
- [28] U. Distler, J. Kuharev, P. Navarro, S. Tenzer, *Nat. Protoc.* **2016**, *11*, 795–812.
- [29] U. Distler, J. Kuharev, P. Navarro, Y. Levin, H. Schild, S. Tenzer, *Nat. Methods* **2014**, *11*, 167–170.
- [30] J. Kuharev, P. Navarro, U. Distler, O. Jahn, S. Tenzer, *Proteomics* **2015**, *15*, 3140–3151.

Manuscript received: December 19, 2019

Accepted manuscript online: January 27, 2020

Version of record online: February 27, 2020