Investigating Peptide-Coenzyme A Conjugates as Chemical Probes for Proteomic Profiling of N-Terminal and Lysine Acetyltransferases

Julia Sindlinger, Stefan Schön, Jürgen Eirich, Sören Kirchgäßer, Iris Finkemeier,* and Dirk Schwarzer*
Cell culture

HeLa S3 (DSMZ-No. ACC 161) cells were cultured in high glucose DMEM, supplemented with 10 % FCS and 1 % penicillin-streptomycin (all Sigma-Aldrich) at 37 °C in a humidified atmosphere with 5 % CO₂. Subcultivation was performed every 3-4 days. Therefore, cells were washed with phosphate-buffered saline (PBS), treated with 0.05 % trypsin-EDTA solution (Sigma Aldrich) and passaged after diluting with the aforementioned cell culture medium.

Preparation of cell extracts

HeLa cell extracts were prepared following a protocol adapted from Dignam et al. [S1]. All buffers were supplemented with 1 mM Na₃VO₄, 2 mM NaF and 0.5 mM 4-(2-aminoethyl)-benzensulfonyl fluoride (AEBSF) (instead of phenylmethylsulfonyl fluoride (PMSF)). EDTA was excluded. Cytosolic (CE) and nuclear extracts (NE) were either stored separately or combined in order to generate whole cell extract (WCE). Total protein concentration was determined by the BCA assay with the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific).

Cloning of bacterial NAA10 expression vector

The codon optimized (E. coli) synthetic gene encoding NAA10 (UniProt ID: P41227) was purchased from Thermo Fisher Scientific (GeneArt Gene Synthesis) as C-terminal TurboYFP fusion. The gene was inserted into the pET28a(+) expression vector (Novagen, Merck KGaA) at the Ndel/HindIII restriction sites, adding an N-terminal His6 tag. Restriction digests were carried out with FastDigest Nde I and FastDigest HindIII (Thermo Fisher Scientific) followed by ligation with T4 DNA ligase (Thermo Fisher Scientific). The sequence of the construct was confirmed by LIGHTRUN Sanger sequencing (Eurofins Genomics, Ebersberg, Germany).

Expression and purification of recombinant NAA10

E. coli BL21(DE3) cells (Agilent, Waldbronn, Germany) were transformed with the pET28a(+) NAA10-TurboYFP expression plasmid and cultured in LB media supplemented with kanamycin (30 μg/mL) at 37 °C while shaking (170 rpm). At an OD₆₀₀ of 0.5-0.6 expression was induced by adding isopropyl-β-D-thiogalactopyranosid (IPTG) (1 mM) and cells were cultivated at 25 °C for 18 h under shaking (170 rpm). In the following the cells were harvested by centrifugation (5000 × g, 10 min, 4 °C), resuspended in lysis buffer (20 mM Tris, 300 mM NaCl 0.1 % Triton X-100 (v/v), pH 8.0) and lysed with an EmulsiFlex-C5 homogenizer (Avestin, Ottawa,
Canada). The lysate was centrifuged (15,000 × g, 30 min, 4 °C) and NAA10-TurboYFP was purified on Ni-NTA Superflow resin (Qiagen, Venlo, Netherlands). Fractions were analyzed by SDS-PAGE and pure fractions were pooled. NAA10-TurboYFP was dialyzed (Spectra/Por Membrane, 6-8 kDa cutoff, Spectrum Laboratories, Rancho Dominguez, USA) against pull-down buffer (20 mM HEPES, 100 mM KCl, 20 % glycerol, pH 7.9) over night, aliquoted and stored at -20 °C. Protein concentration was determined on a NanoDrop ND-1000 device (VWR Peqlab, Radnor, USA).

**Pull-down assays**

An aliquot of each of the resin-bound peptide probes was transferred to a microcentrifuge filter unit (Ultrafree-MC-HV, Merck Millipore) with pull-down buffer (20 mM HEPES, pH 7.9, 100 mM KCl, 20 % glycerin). After centrifugation at 1700 × g for 2 min, the filtrate was discarded and the beads were equilibrated with pull-down buffer under agitation at 550 rpm for 2 min at room temperature (3 × 200 μL). The resin was drained by centrifugation at 1700 × g for 2 min after each step. Cell extracts were diluted to the desired concentration (Western blot for detection of NAA10: 1 mg/mL (WCE), NAA50: 1 mg/mL (CE), NAA25: 2 mg/mL (CE) and KAT7: 2 mg/mL (WCE); Label-free proteomics with WCE, CE and NE: 1 mg/mL) with pull-down buffer. The resin was incubated with 200 μL of the prepared solution by shaking at 550 rpm for 1 h at room temperature. After draining the supernatant, the resin was washed with pull-down wash buffer (20 mM HEPES, 300 mM KCl, 20 % glycerin, pH 7.9) (6 × 500 μL) by shaking at 550 rpm for 5 min at room temperature and subsequent centrifugation for 2 min at 1700 × g.

The peptide-bound proteins analyzed by Western Blot were eluted by incubation with 20 μL of 3 × SDS sample buffer (5 × SDS sample buffer: 250 mM Tris∙HCl, 10 % Sodium dodecyl sulfate (SDS), 30 % glycerin, 0.5 M dithiothreitol (DTT), 0.02 % bromophenol blue) for 10 min at 95 °C while shaking at 550 rpm, followed by centrifugation at 9400 × g for 5 min. An input sample was prepared by mixing 16 μL of the respective diluted HeLa cell extract solution with 4 μL of 5 × SDS sample buffer and heating for 10 min at 95 °C. All samples were loaded and separated by SDS-PAGE.

The elution for label-free proteomics investigations was carried out as described above with 20 μL of 3 × SDS sample buffer without bromophenol blue. An input sample was prepared by mixing 20 μL of the respective diluted HeLa cell extract with 30 μL of 5 × SDS sample buffer without bromophenol blue and heating for 10 min at 95 °C. Eluted proteins and the input sample were stored at 4 °C until subjected to filter aided sample preparation (FASP).

Pull-Downs with recombinant NAA10-TurboYFP were performed at an input concentration of 5 μM with 10 mg/mL BSA in pull-down wash buffer. The probe resins were incubated for 1 h at RT followed by 4 incubations with pull-down wash buffer and 2 incubations with 20 mM HEPES, 100 mM KCl, 20 % glycerin, pH 7.9 for 2 min each. Bound proteins were eluted by heating SDS sample buffer, separated on SDS-PAGE and stained with coomassie brilliant blue.
Western blotting

Protein samples, separated on SDS-PAGE, were transferred onto PVDF membranes in a wet tank blotting system. The membrane was blocked by incubation with blocking buffer (5% low fat powdered milk in PBS) for 30 min at room temperature.

After washing with PBST (0.1 % Tween 20 in PBS, pH 7.3) membranes were incubated with the respective primary antibodies in primary antibody buffer (2.5 % low fat powdered milk in PBST) at 4 °C over night. The following antibodies and dilutions were used: ARD1A (NAA10) rabbit polyclonal IgG (1:500, 14803-1-AP, proteintech, Rosemont, USA) NAP1 (NAA25) rabbit polyclonal IgG (1:200, ab79490, abcam, Cambridge, United Kingdom), NAT13 (NAA50) rabbit polyclonal IgG (1:1000, ab241487, abcam), HBO1 (KAT7) rabbit recombinant monoclonal IgG (1:1000, ab124993, abcam), lamin rabbit polyclonal IgG (1:500, #2032, Cell Signaling Technology, Danvers, USA), α-tubulin rabbit polyclonal IgG (1:500, #2144, Cell Signaling Technology).

The membranes were washed afterwards with PBST and incubated with the secondary antibody solution for 1 h at room temperature: IRDye 800 CW donkey anti-rabbit (1:10000, 925-32213, Li-Cor Bioscience, Lincoln, USA) in secondary antibody buffer (5% low fat powdered milk, 0.01 % SDS in PBST).

After washing with PBST and PBS IR-signals were detected using a ChemiDoc MP imaging system (Bio-Rad Laboratories, Munich, Germany).

Filter aided sample preparation (FASP)

Eluates from pull-down assays and the respective input samples were processed according to the FASP method as described previously [S2]. 450 μL urea buffer (8 M urea, 0.1 M Tris·HCl, pH 8.5) were added to each sample and incubated at 50 °C for 5 min under agitation. The mixtures were transferred to centrifuge filter units (Microcon YM-10, Merck Millipore). After centrifugation at 13900 × g for 20 min the filtrate was discarded. The samples were washed with urea buffer (3 × 450 μL) and drained by centrifugation (13900 × g, 20 min). 100 μL 55 mM chloroacetamide in urea buffer were added to the filter units and incubated in the dark for 20 min. After subsequent centrifugation at 13900 × g for 15 min, three washing steps were performed with 100 μL urea buffer each and centrifugation at 13900 × g for 10 min. The collection tubes were changed prior to enzymatic cleavage. To this end, 1 μg Lysyl endopeptidase (Lys-C, Wako Chemicals, Neuss, Germany) (0.5 μg / μL in 50 mM ammonium bicarbonate (ABC) buffer) in 40 μL urea buffer was added and the samples were incubated at room temperature for 1 h while shaking. Additional Lys-C (1 μg) was added and the mixture was incubated over night at room temperature under agitation. The solution was diluted with 300 μL ABC buffer and 0.2 μg trypsin (MS approved, SERVA, Heidelberg, Germany) (1 μg/ μL in 1 mM HCl) were added. The mixture was incubated at 37 °C for 3.5 h while shaking, afterwards trypsinization was repeated under the same conditions with 0.2 μg trypsin. After centrifugation at 13900 × g for 10 min, 50 μL ABC buffer were added to the filter units and centrifugation was repeated. The eluates were acidified with TFA to a final concentration of 1%, desalted and concentrated using C18-StageTips (pull-down samples) or SDB-StageTips (input samples) [S3].
Nano-LC-MS/MS

Peptides from tryptic digests were separated using an EASY-nLC1200 UHPLC system (Thermo Fisher Scientific) equipped with a 20 cm column, packed in-house with ReproSil-Pur C18-AQ 1.9 μm resin (Dr. Maisch). The column temperature was maintained at 50 °C, and the column was coupled to an Exploris 480 (E480) in case of samples related to the Pro and Glu probes or a Q Exactive HF (QEHF) mass spectrometer for all other samples (both Thermo Fisher Scientific) via a nano-electrospray source. Total peptides (0.5 μg) were loaded onto the column and separated over a segmented linear gradient from 3-80 % buffer B (0.5 % formic acid in ACN) in 120 min. The mass spectrometers were operated in data-dependent mode, survey scans were obtained in a mass range of 380 - 1500 (E480) and 300-1759 m / z (QEHF), at a resolution of 120 000 at 200 m / z and an AGC target value of 3 × 10^6. The 20 (E480) or 12 (QEHF) most intense ions were selected with an isolation width of 1.6 (E480) or 1.2 m/z (QEHF), fragmented in the HCD cell at a collision energy of 30 (E480) or 25 (QEHF) and the spectra recorded at a target value of 5 × 10^4 and a resolution of 15 000. Peptides with a charge of +1 or +6 and higher were excluded from fragmentation, the peptide match and exclude isotope features were enabled and selected precursors were dynamically excluded from repeated sampling for 30 s.

Data processing and quantification

Raw data were processed using the MaxQuant software package (Version 2.0.3.0 for samples related to Pro and Glu probes, Version 1.6.9.0 for others, https://www.maxquant.org/) [S4] and searched against the human reference proteome (UP000005640_9606, https://www.maxquant.org/), downloaded 2019-09-30) and an internal database containing standard contaminants. The search was performed with full trypsin specificity and a maximum of two missed cleavages and a protein false discovery rate of 1 %. Minimal peptide length was set to seven amino acids. Carbamidomethylation of cysteine residues was set as fixed, oxidation of methionine and N-terminal acetylation were set as variable modifications. All other search parameters were left at default. The MaxLFQ algorithm integral to MaxQuant was used for label-free quantification with match between runs enabled and the LFQ minimum ratio count was set to 1 [S5]. Input and pull-down samples were searched in two different parameter groups for separate LFQ normalization. Downstream data analysis was carried out in the Perseus software package (Version 1.6.1.1, https://www.maxquant.org/).[S4] All hits for contaminants and reversed sequences were removed and the LFQ values log2 transformed. Statistical analysis was performed with the Limma package [S6] for R (Version 3.6.0, https://www.r-project.org/) using the log2 ratios from three independent biological replicates. Volcano plots were generated with Microsoft Excel, plotting p-values from Limma analysis against the log2 ratio.
Data availability

Raw data of LC-MS/MS analysis is deposited on the jPOST server [S7] and available for download under the identifiers JPST001556 (https://repository.jpostdb.org/entry/JPST001556.3) for Pro and Glu probes and JPST000827 (https://repository.jpostdb.org/entry/JPST000827.0) for samples related to Ala and Lys related probes.

Supporting references

[S1] J. D. Dignam, P. L. Martin, B. S. Shastry, R. G. Roeder, Methods Enzymol. 1983, 101, 582-598.
[S2] J. Seidel, T. Meisinger, J. Sindlinger, P. Pieloch, I. Finkemeier, D. Schwarzer Chembiochem 2019, 20, 3001-3005.
[S3] J. Rappsilber, M. Mann, Y. Ishihama, Nat. Protoc. 2007, 2, 1896-1906.
[S4] J. Cox, M. Mann, Nat. Biotechnol. 2008, 26, 1367-1372.
[S5] J. Cox, M. Y. Hein, C. A. Luber, I. Paron, N. Nagaraj, M. Mann, Mol. Cell. Proteomics 2014, 13, 2513-2526.
[S6] M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, G. K. Smyth, Nucleic Acids Res. 2015, 43, e47.
[S7] S. Okuda, Y. Watanabe, Y. Moriya, S. Kawano, T. Yamamoto, M. Matsumoto, T. Takami, D. Kobayashi, N. Araki, A.C. Yoshizawa, T. Tabata, N. Sugiyama, S. Goto, Y. Ishihama Nucleic Acids Res. 2017, 45, D1107-D1111
Supporting figures

Supporting Figure S1: Synthesis scheme of probe peptides and immobilized probes. a) Synthesis of N-terminal peptide probe fragments and CoA conjugation. b) Linkage of N-terminal probe fragments and C-terminal linker fragment by Copper(I)-catalyzed Azide-Alkyne Cycloaddition. c) Immobilization of probe peptides. LC-MS of numbered intermediates 1 - 9 is illustrated in Supporting Figures 5 - 7. Abbreviations: Ahx: 6-Aminohexanoic acid, BAA-OSu: N-Succinimidyl bromoacetate, BAA₂: 2-bromoacetic anhydride, CoA-SH: Coenzyme A trilithium salt, DCM: Dichloromethane, HFIP: Hexafluoroisopropanol, TBTA: Tris((1-benzyl-4-triazolyl)methyl)amine, TFE: 2,2,2-Trifluoro-ethanol, TIPS: Triisopropyl silane
Supporting Figure S2: a) Western Blot analysis of WCE, CE, and NE extracts. Lamin and α-tubulin serve as marker proteins of the nuclear and cytosolic fraction, respectively. In total 20 µg or respective extracts were loaded per lane. b) Western blot validation results of pull-down experiments with α-Ala-CoA, ε-Lys-K-CoA and A-K-Ac with antibodies against NAA10, NAA25, NAA50 and KAT7. Pull-down input of Hela whole cell extract: 2 mg/mL for KAT7, 1 mg/mL for NAA10 and Hela cytosolic extract: 2 mg/mL for NAA25 and 1 mg/mL for NAA50. Gel input: 16 µg for NAA10 and NAA50, 32 µg for KAT7 and NAA25. All pull-down validation experiments were performed in triplicate (data shown here and Figure 3).
Supporting Figure S3: Volcano plots of α-Ala-CoA versus ε-Lys-CoA enriched proteins in nuclear extracts (NE) and cytosolic extracts (CE). Proteins enriched on α-Ala-CoA over ε-Lys-CoA are located in the upper right section of the plots. Proteins enriched on ε-Lys-CoA over α-Ala-CoA are located in the upper right section of the plots. (n=3; LIMMA test statistics).
Supporting Figure S4: Volcano plots of the a) cytosolic α-Ala-CoA over A-K-Ac enriched proteins, b) cytosolic ε-Lys-CoA over A-K-Ac enriched proteins, c) nuclear α-Ala-CoA over A-K-Ac enriched proteins and, d) nuclear ε-Lys-CoA over A-K-Ac enriched proteins. Proteins enriched on the respective CoA-conjugated probes over the A-K-Ac control are located in the upper right section of the plots. (n=3; LIMMA test statistics).
Supporting Figure S5: HPLC chromatograms and mass spectra of the purified bromoacetylated peptide precursors Ala-BAA (1a), Pro-BAA (1a), 3Glu-BAA (1c) and Lys-BAA (2).
Supporting Figure S6: HPLC chromatograms and mass spectra of N-terminal peptide fragments 3a-3c and 4 after CoA conjugation and purification before Click reaction. LC-MS analysis of N-terminal fragment 5 of N-K-Ac probe peptide and C-terminal peptide fragment 6.

Supporting Figure S7: HPLC chromatograms and mass spectra of probe peptide 7a-7c, 8, and 9 after purification. These peptides are subsequently coupled to the SulfoLink solid support and used as α-Ala-CoA, α-Pro-CoA, α-3Glu-CoA, ε-Lys-CoA, and A-K-Ac probes.