bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

System-wide Identification of Enzyme Substrates by Thermal Analysis (SIESTA)

Authors: Amir Ata Saei[1], Juan Astorga Wells[1], Pierre Sabatier[1], Christian Beusch[1], Alexey Chernobrovkin[1,2], Sergey Rodin[1,3], Katja Näreoja[4], Ann-Gerd Thorsell[4], Tobias Karlberg[4], Qing Cheng[5], Ákos Végvári[1], Elias S.J. Arnér[5], Herwig Schüler[4], Roman A. Zubarev[1,6*]

Affiliations:

[1]Division of Physiological Chemistry I, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, SE-17 177 Stockholm, Sweden

[2]Pelago Bioscience AB, 171 48, Solna, Sweden

[3]Department of Surgical Sciences, Uppsala University, Sweden

[4]Department of Biosciences and Nutrition, Karolinska Institutet, SE-14183 Huddinge, Sweden

[5]Division of Biochemistry, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, SE-171 77 Stockholm, Sweden

[6]Department of Pharmacological & Technological Chemistry, I.M. Sechenov First Moscow State Medical University, Moscow, Russia

*Correspondence and requests for materials should be addressed to R.A.Z. (email: Roman.Zubarev@ki.se)

One sentence summary: Monitoring protein stability can reveal changes due to specific enzymatic post-translational modifications at the proteome level.

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Despite the immense importance of enzyme-substrate reactions, there is a lack of generic and unbiased tools for identifying the molecular components participating in these reactions on a cellular level. Here we developed a universal method called System-wide Identification of Enzyme Substrates by Thermal Analysis (SIESTA). The approach assumes that enzymatic post-translational modification of substrate proteins changes their thermal stability, and applies the concept of specificity to reveal potential substrates. For selenoprotein thioredoxin reductase 1, SIESTA confirmed several known protein substrates and suggested novel candidates. For poly-(ADP-ribose) polymerase-10, SIESTA revealed a number of putative substrates, which were confirmed by targeted mass spectrometry and functional assays. Wider application of SIESTA can enhance our understanding of the role of enzymes in homeostasis and disease, and facilitate drug discovery.

Keywords: ARTD10; mono-ADP-ribosylation, PARP10; post-translational modification; proteomics; target; thermal stability; TrxR1; TXNRD1

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

At least a third of all proteins possess enzymatic activity. One of the most comprehensive enzyme databases BRENDA (*1*) comprises >9 million protein sequences and encompasses >7000 classes of enzyme-catalyzed reactions (http://genexplain.com/brenda/). Many of these enzymes catalyze the modifications of protein substrates. Transient modulation of protein post-translational modifications (PTMs) controls numerous cellular processes, with PTM occupancies often acting as independent regulators in respect to protein expression levels (*2*). PTMs induce a host of downstream effects, such as changes in protein function, stability, hemostasis, localization and cellular diversification (*3*). Not surprisingly, mechanisms and kinetics of protein modifications have become a vibrant research area (*4*). An important aspect of PTM research is the characterization of enzyme-substrate associations, which is essential for our understanding of cell biology and disease mechanisms. Moreover, many high-throughput screening assays rely upon modified substrates as a readout (*5*). The lack of information on the physiological substrates of enzymes hampers the development of effective therapeutics, e.g. in Parkinson's disease (*6*) and cancer (*7*).

For thorough understanding of homeostasis and disease, all substrates need to be determined for every PTM-catalyzing enzyme. However, conventional techniques used for identifying specific substrates are enzyme-specific, labor-intensive and often not straightforward. Such experiments include the use of genetic and pharmacologic perturbations (*8*), substrate-trapping mutants (*9*), affinity purification-mass spectrometry (*10*), utilizing peptide (*11*) or protein arrays (*12*), tagging the client proteins by substrate analogues using engineered enzymes (*13*), peptide immunoprecipitation (*14*) or the use of sophisticated computational tools (*15*). Most of these techniques are specifically designed for a certain enzyme or enzyme class, which limits their applicability. Engineering mutant enzymes can alter the biology of the system, potentially introducing a bias and increasing the risk of false positive discoveries. Therefore, designing an unbiased, universal and system-wide method not involving modification of the enzymes or

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

their co-factors can prove to be a significant methodological advancement useful in a wide range of studies.

Any interaction with a small molecule, metabolite, protein or nucleic acid, as well as a genetic mutation, can affect protein stability (*16-19*). PTMs can also alter thermal stability of the substrate protein (*20-22*). A cell lysate contains most potential substrates for any cellular enzyme. Therefore, proteome-wide monitoring of the thermal stability changes (*19*) in the cell lysate upon addition of a recombinant enzyme and an eventual cofactor has the potential to reveal enzyme substrates. However, the concomitant protein-enzyme and protein-cofactor interactions will interfere with and can mask modification-specific thermal stability changes of the substrates. This problem is addressed in our method of System-wide Identification of Enzyme Substrates by Thermal Analysis (SIESTA). SIESTA identifies *specific* thermal stability changes induced in substrate proteins by a combination of enzyme and co-factor as compared to the changes induced by either enzyme or co-factor alone. The idea of specific response is borrowed from our method of Functional Identification of Target by Expression Proteomics (FITeXP) (*23*). FITExP reveals specific changes in protein abundances, by contrasting proteome responses to a given drug with those to all other drugs and controls. We have recently shown that excellent specificity can be achieved in some cases with a few contrasting treatments, while in more general cases >10 contrasting treatments are required (*24*). Since in our experience, thermal stability changes are more specific than the proteome abundance responses, the number of contrasting treatments in SIESTA can be as low as three: "control" (cell lysate incubated with vehicle), "enzyme" (cell lysate incubated with an added purified enzyme, recombinant or cell-isolated), and "co-factor" (cell lysate incubated with an added cofactor). These treatments are contrasted with "enzyme + co-factor" to reveal specifically shifting proteins (Fig. 1).

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

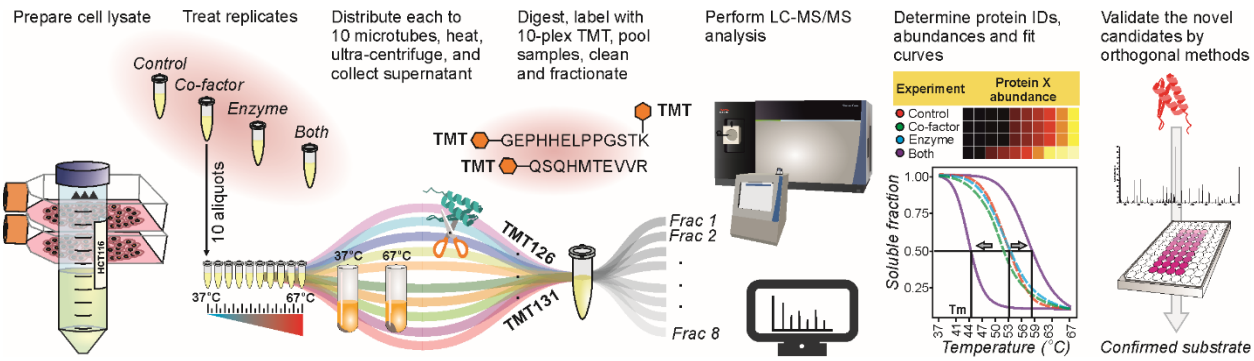

Fig. 1. SIESTA workflow for unbiased proteome-wide identification of enzyme protein substrates. A master cell lysate is prepared by multiple freeze-thawing in a non-denaturing buffer. The cell lysate aliquots are treated with vehicle (control), co-factor, enzyme or combination of enzyme with co-factor. After treatment, each aliquot is split into 10 tubes, with each tube heated to a temperature point in the range from 37 °C to 67 °C. After removing unfolded proteins by ultracentrifugation, identical volumes of supernatants are digested with trypsin. The samples are then serially labeled with 10-plex TMT reagents, pooled, cleaned and fractionated by reversed-phase chromatography. After LC-MS/MS analyses of each fraction, protein IDs and abundances are determined, and sigmoid curves are fitted to determine the melting temperature T_m for each protein. For each non-vehicle treatment, the read-out is the protein's ΔT_m shifts (of both signs) compared to control. Any protein shifting significantly more upon addition of enzyme and co-factor compared to when they are added alone, are putative substrates of the enzyme under study. Such candidate protein substrates are subsequently confirmed by orthogonal verification methods.

SIESTA identified known and putative TXNRD1 substrates

As a proof of principle of the SIESTA approach, we first selected an enzymatic reaction involving disulfide bond reduction. Since such a reaction should destabilize substrate proteins and lead to negative ΔT_m, the asymmetry between positive and negative values will be easy to verify. For this reaction we employed human selenoprotein thioredoxin reductase 1 (TXNRD1), a key oxidoreductase

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

that catalyzes the reduction of several substrate proteins using NADPH as a co-factor (*25*). HCT116 cell lysate was therefore treated with vehicle, NADPH (1 mM), TXNRD1 (1 µM), or both, all experiments being done in duplicates. In LC-MS/MS analysis, 5864 proteins were quantified, of which 5637 proteins with at least two peptides (Table S1).

Changes in T_m after NADPH treatment revealed stabilization of several known NADPH-interacting proteins, as expected (Fig. 2A-B; listed in Table S2). The analysis of specific ΔT_m shifts in the TXNRD1+NADPH treatment revealed that in the presence of NADPH, TXNRD1 destabilized 13 known substrate proteins (Fig. 2C). Several novel candidate substrate proteins were found. In general, the expected asymmetry in melting temperature shifts in favor of destabilization was well pronounced (Fig. 2C). A partial least square discriminant analysis (OPLS-DA) model (*26*) contrasting TXNRD1+NADPH with all other treatments was also used to reveal the specifically destabilized proteins (Fig. 2D). Such modeling showed its utility in FITExP (*23, 24*) and its main advantage is in the ability to use arbitrary number of contrasting treatments.

Examples of melting curves for proteins destabilized by TXNRD1 are shown in Fig. 2E. The 28 identified substrates (Table S3) mapped to the following INTERPRO Protein Domains and Features pathways: “thioredoxin-like fold” (8 proteins, $p < 3e-11$) and “peroxiredoxin, C-terminal” (4 proteins, $p < 3e-08$). GO and KEGG terms included “oxidoreductase activity” (13 proteins, $p < 5e-10$) and “glutathione metabolism” (4 proteins, $p < 0.0002$), in very good agreement with the known physiological roles of TXNRD1 (*25*).

GPX1 was the protein showing the strongest destabilization (Fig. 2C-E). GPX1 is a cytosolic selenoprotein usually considered to be glutathione-dependent (*27*), but at least some GPX isoenzymes are known to be directly reduced by TXNRD1 (*28*). Several peroxiredoxins (PRDXs), the highly abundant thioredoxin-dependent peroxidases, were also destabilized by the addition of TXNRD1 and NADPH, presumably through the action of thioredoxin present in the cell lysate (*29*). TXNL1 (or TRP32) (*30*), NXN (*31*), COPS5 (Or Jab-1) (*32*) and NFKB1 (*33*) are also well-known substrates of TXNRD1 or

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TXNRD1-dependent enzymes. GSTO1 and GSTO2 have thioredoxin-like domains with glutaredoxin-like activities (*34*) and are also likely to be TXNRD1-dependent. RNASET2 is an extracellular ribonuclease, dependent upon correct disulfide formation in the endoplasmic reticulum, which in turn is dependent upon TXNRD1 (*35*). There are also links between ETHE1 and thioredoxin system, as this protein affects polysulfide status (*36*). Other identified proteins are of interest as potential TXNRD1 substrates; their validation was outside the scope of this work.

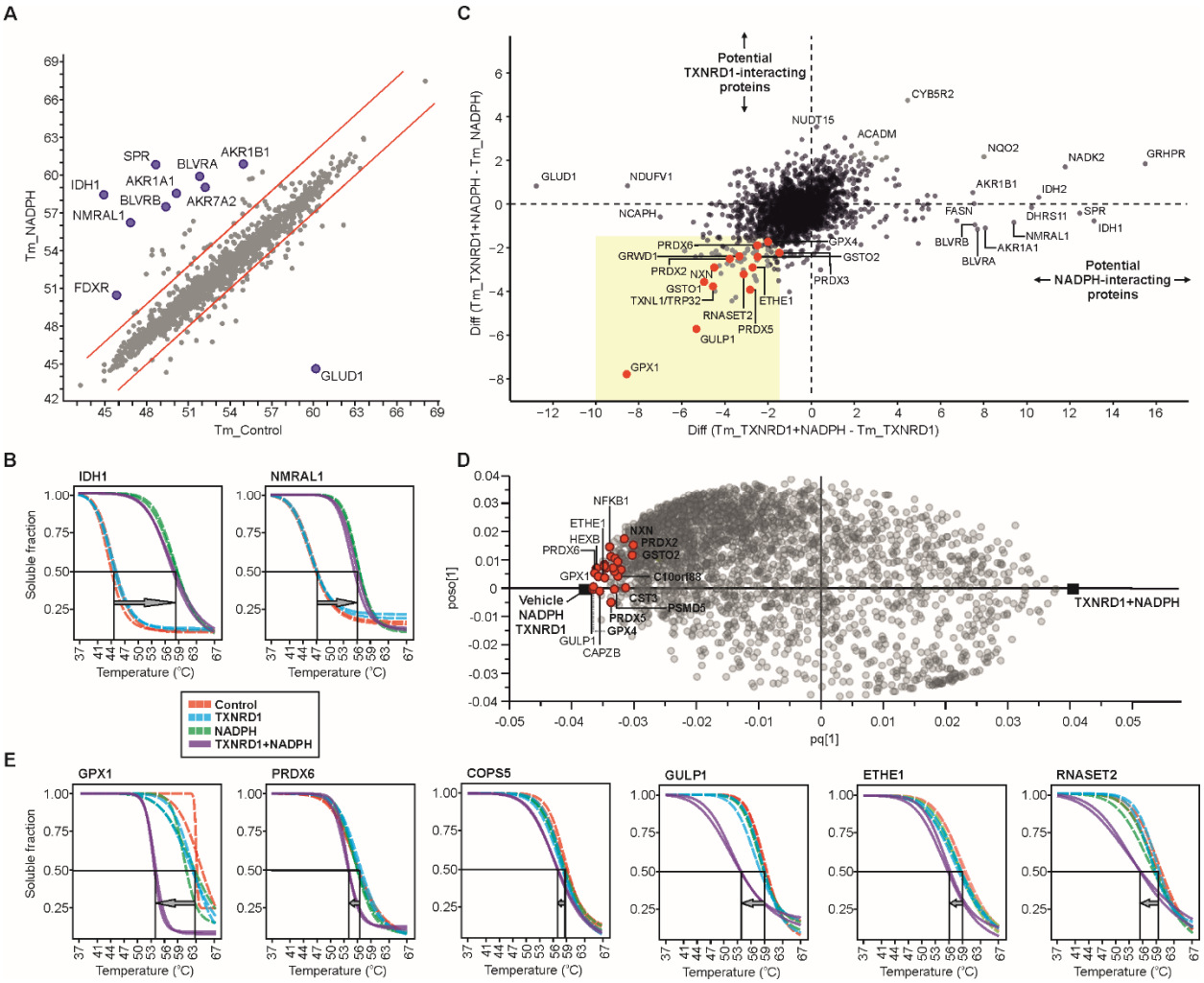

Fig. 2. Proof-of-principle SIESTA experiment revealed known TXNRD1 substrates and suggested novel candidates. A) SIESTA data for control and NADPH were filtered for proteins with a T_m difference between the two replicates of ≤1°C. A linear regression was then performed with a 95%-

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

prediction interval as a cutoff. The outliers are known and putative NADPH interacting proteins. Purple circles correspond to representative proteins known to interact with NADPH. **B**) Representative stabilization of NADPH binding proteins IDH1 and NMRAL1. **C**) Scatterplot of T_m differences reveals the melting temperature shifts occurring only after simultaneous TXNRD1 and NADPH addition; these shifts are thus likely due to enzymatic modifications (shaded yellow, putative and known substrates are shown as red circles). **D**) Potential substrates (red circles) are located close to the negative reference point (black square) in an OPLS-DA model contrasting the "TXNRD1+NADPH" T_m against all other treatments. **E**) Representative melting curves of known TXNRD1 targets (GPX1, PRDX6 and COPS5) as well as putative substrates (GULP1, ETHE1 and RNASET2).

SIESTA identified many novel putative substrates for PARP10

We next selected a more challenging system involving poly-(ADP-ribose) polymerase-10 (PARP10), a member of the PARP family of proteins that performs mono-ADP ribosylation of proteins (*37*). ADP-ribosylation is involved in cell signaling, DNA repair, gene regulation and apoptosis (*38*). Identification of PARP family substrates by mass spectrometry has generally proven challenging, as ADP-ribosylation is a glycosidic modification that can be easily lost during protein extraction or sample processing. It is also highly labile in the gas phase, which hampers its detection by MS/MS. Different strategies have thus been used to enrich the modified peptides for mass spectrometric analysis and use "gentle" MS/MS methods (*39*, *40*). Unlike poly-ADP-ribosyltransferases, the mono-ADP-ribosyltransferases such as PARP10 have not been studied to great extent and this enzyme is in need of further functional characterization.

In contrast to the TXNRD1 system where protein substrates were mainly destabilized, we expected both destabilization and stabilization of substrate proteins with PARP10 system in SIESTA. Overall, 5194 proteins were quantified, of which 4979 with at least two peptides (Table S4). As a

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

verification of experiment quality, several proteins known to interact with NAD were identified by comparing the ΔT_m shifts in the NAD vs vehicle treatments (Fig. S1A-B, Table S5).

In total, 28 proteins changed their stability only when both PARP10 and NAD were present (12 proteins were destabilized and 16 stabilized), i.e. were identified by SIESTA as potential PARP10 substrates (Fig. 3A, Table S6). Melting curves for some of these proteins are shown in Fig. 3B. An OPLS-DA model contrasting "PARP10+NAD" T_m vs. those from all other treatments is given in Fig. S1C.

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

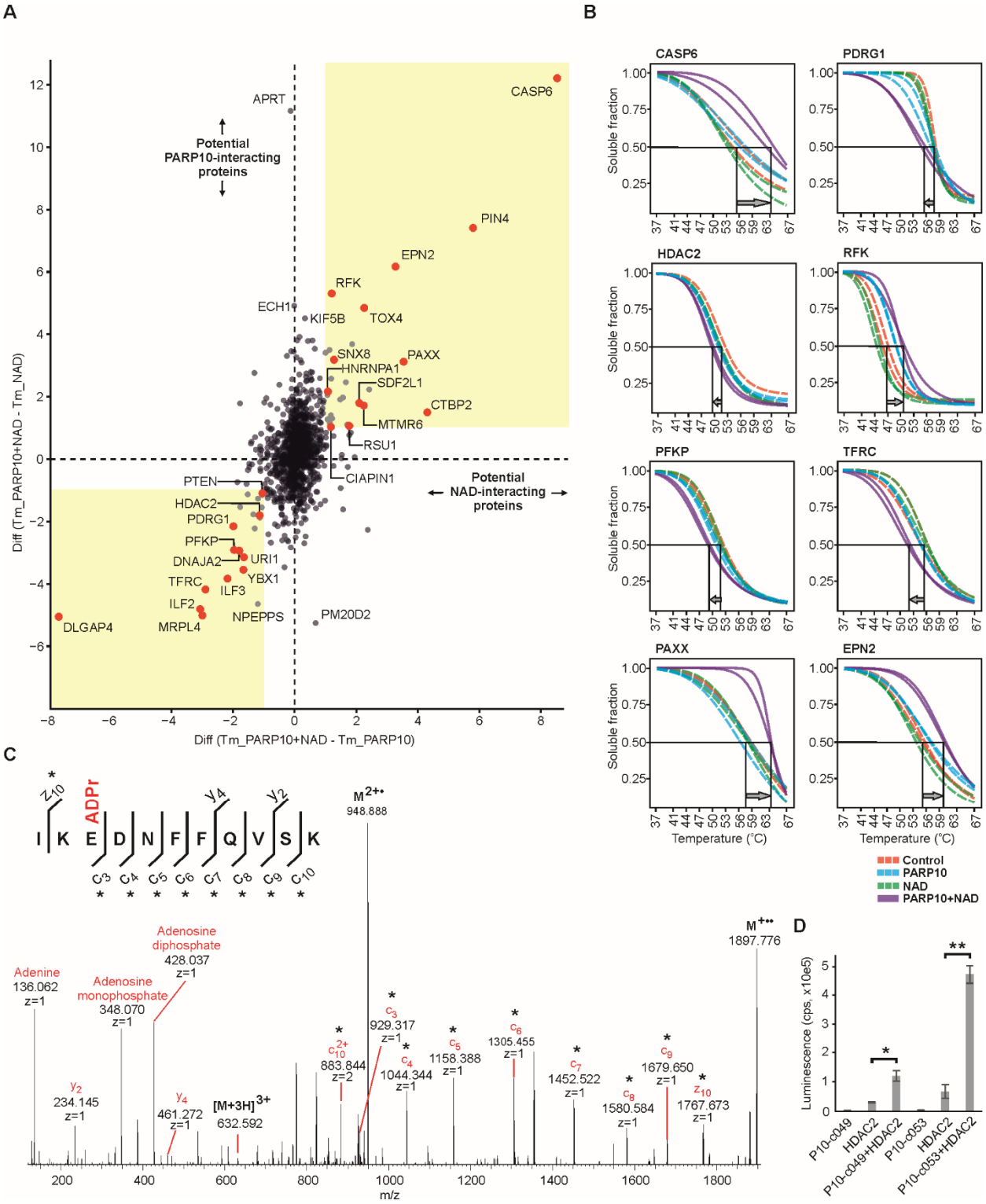

Fig. 3. SIESTA identified novel PARP10 substrates. A) Scatterplot of T_m differences reveals the shifts occurring when both NAD and PARP10 are incubated with cell lysate. Red circles denote candidate

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

substrates with reproducible shifts. **B**) Representative melting curves of putative PARP10 substrates. **C**) Targeted ETD MS/MS of a RFK peptide revealed mono-ADP-ribosylation on glutamic acid residue (the site with the highest sequence-fitting score). The fragments carrying the modification are marked with an asterisk. **D**) The mono-ADP-ribosylation of HDAC2 was confirmed in two independent experiments using a purified recombinant protein and two PARP10 catalytic subunit constructs (data are represented as mean±SD; *$p < 0.02$ and **$p < 0.005$).

In GO terms analysis, the identified potential protein substrates of PARP10 mapped to "ribonucleoprotein complex" (7 proteins, $p < 0.02$), "nuclear part" (15 proteins, $p < 0.02$) and "nucleus" (19 proteins, $p < 0.05$), which is in line with known PARP10 functions and its cellular location (*41*). The stabilized proteins did not map to any common pathway, while three of the 12 destabilized proteins mapped to "double-stranded RNA binding" ($p < 0.02$) and six to "poly(A) RNA binding" ($p < 0.04$). Several proteins, such as ILF2 and ILF3, have already been reported as PARP10 substrates (*42*), which serves as an additional confirmation of the method.

The majority of the identified PARP10 substrates were novel, and to rank them for validation priority, their melting curves were manually examined and OPLS-DA loadings were sorted. Based on the availability of high purity recombinant proteins, the destabilized PDRG1 and HDAC2 as well as the stabilized CASP6 and RFK putative substrates were chosen for verification of the presence of mono-ADP-ribosylation as a result of PARP10-catalysis. After incubation with recombinant PARP10 and NAD, the above proteins were digested and analyzed with LC-MS/MS. Every higher energy collision dissociation (HCD) MS/MS event triggered in data-dependent acquisition was investigated in real time for the presence of signature ions of adenine (m/z 136.0623), adenosine-18 (m/z 250.094) and adenosine monophosphate (AMP, m/z 348.0709). The presence of these triggers would then initiate a second MS/MS event using electron-transfer dissociation (ETD) with a supplementary HCD activation. The obtained RFK sequence coverage was 94%, and ADP ribose moieties were found in four positions: on

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Glu140, Glu131, Glu113 and Arg14, ordered from highest to lowest peptide score (Table S7). The ETD MS/MS spectrum of a peptide with Glu140 is shown in Fig. 3C. The PDRG1 sequence coverage was 74%, and the protein was found modified with ADP in three locations: on Glu110, Glu75 and Asp32 (Table S7, Fig. S1D).

Since the HDAC2 sequence coverage with trypsin digestion was low despite supplementary LysC digestion, we used an *in vitro* chemiluminescence assay to validate the mono-ADP-ribosylation of HDAC2. Two different PARP10 catalytic domain constructs were used for that analysis, and both of them indeed significantly modified HDAC2 (Fig. 3D). These results verified RFK, PDRG1 and HDAC2 as novel PARP10 substrates.

Caspase-6 showed the largest specific stabilization (10.4 ^{o}C, Fig. 3A-B), but its modification was not verified in either of the two *in vitro* assays. It should also be noted that PARP10 was suggested to be a substrate for caspase-6 during apoptosis (*43*). PARP10 has a major cleavage site at D406 that is preferentially recognized by caspase-6 (*43*). The large specific thermal stabilization might therefore indicate that PARP10 induces a conformational change in caspase-6 and thus an increase in its stability by binding, as has been reported for other caspase-6 substrates (*44*). The reason why caspase-6 stabilization was not observed upon PARP10 addition in the absence of NAD may be that auto-modified PARP10 (in presence of NAD) is required for effective caspase-6 binding.

In summary, we demonstrated SIESTA to be a general approach for unbiased identification of protein substrates for specific enzymes in a proteome-wide manner. Besides confirming several known specific substrates for both TXNRD1 and PARP10, we uncovered a number of interesting potential novel candidates for these proteins, implicating them in important cellular processes.

SIESTA will likely be able to identify most types of protein substrates for a wide range of enzymes. The spatial resolution of the method can be increased by sub-cellular fractionation of the lysate prior to analysis. Furthermore, cell- or tissue-specific substrates should be possible to discover by

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

comparing lysates from different sources. The ease of identifying enzyme-specific substrates offered by this technology can enhance our understanding of enzyme systems and disease, accelerate constructing high-throughput assays and thus facilitate drug discovery.

Materials and methods

Experimental Design

A SIESTA layout with four parallel treatments was designed (Fig. 1), with vehicle, enzyme, co-factor or both being added to a cell lysate of choice and incubated. Subsequently, the thermal stability of the incubated mixture was monitored at the proteome level. The specific melting temperature shifts resulting from combined action of enzyme and co-factor were identified by filtering out the shifts due to enzyme alone or cofactor alone. The top candidate substrate proteins were validated by targeted tandem mass spectrometry and *in vitro* chemiluminescence.

Cell culture

Human colorectal carcinoma HCT116 (ATCC, USA) cells were grown at 37 °C in 5% CO_2 using McCoy's 5A modified medium (Sigma-Aldrich, USA) supplemented with 10% FBS superior (Biochrom, Berlin, Germany), 2 mM L-glutamine (Lonza, Wakersville, MD, USA) and 100 units/mL penicillin/streptomycin (Gibco, Invitrogen). Low-number passages (<10) were used for the experiments.

Recombinant proteins

Human TXNRD1 was expressed recombinantly in *E.coli* and purified as described earlier (*45*). PARP10 full length protein and catalytic domain construct were produced as detailed before (*46*). RFK (ab89009) and PDRG1 (PRO-007) were purchased from Abcam and Prospec, respectively, while Caspase-6 (ALX-201-060-U100) and HDAC2 (BML-SE533-0050) were obtained from Enzo.

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

SIESTA experiment

Cells were cultured in 175 cm^2 flasks, and were then trypsinized, washed twice with PBS, counted, resuspended in 50 mM HEPES pH 7.5, 2 mM EDTA (for TXNRD1) or in 50 mM HEPES pH 7.5, 100 mM NaCl and 4 mM MgCl2 (for PARP10), both with complete protease inhibitor cocktail (Roche). The cells were then freeze-thawed 5 times. The cell lysates were centrifuged at 21,000 g for 20 min and the soluble fraction was collected. The protein concentration in the lysate was measured using Pierce BCA assay (Thermo) and equally distributed into 8 aliquots (1 mL each). For TXNRD1, each pair of samples were incubated with vehicle, 1 mM NADPH, 1 µM TXNRD1, or with TXNRD1+NADPH at 37 °C for 30 min. For PARP10, each pair of samples were incubated with vehicle, 100 µM NAD, 400 nM PARP10, or with PARP10+NAD at 37 °C for 1 h. Each replicate was then aliquoted into 10 PCR microtubes and incubated for 3 min in SimpliAmp Thermal Cycler (Thermo) at temperature points of 37, 41, 44, 47, 50, 53, 56, 59, 63, and 67 °C. Samples were cooled for 3 min at room temperature and afterwards kept on ice. Samples were then transferred into polycarbonate thickwall tubes and centrifuged at 100,000 g and 4 °C for 20 min.

The soluble protein fraction was carefully transferred to new Eppendorf tubes. Protein concentration was measured in the samples treated at lowest temperature points (37 and 41° C) using Pierce BCA Protein Assay Kit (Thermo), the same volume corresponding to 50 µg of protein at lowest temperature points was transferred from each sample to new tubes and urea was added to a final concentration of 4 M. Dithiothreitol (DTT) was added to a final concentration of 10 mM and samples were incubated for 1 h at room temperature. Subsequently, iodoacetamide (IAA) was added to a final concentration of 50 mM and samples were incubated in room temperature for 1 h in the dark. The reaction was quenched by adding an additional 10 mM of DTT. Proteins were precipitated using methanol/chloroform. The dry protein pellet was dissolved in 8 M urea, 20 mM EPPS (pH=8.5) and diluted to 4 M urea. LysC was added at a 1 : 100 w/w ratio at room temperature overnight. Samples were diluted with 20mM EPPS to the final urea concentration of 1 M, and trypsin was added at a 1 : 100 w/w

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ratio, followed by incubation for 6 h at room temperature. Acetonitrile (ACN) was added to a final concentration of 20% and TMT reagents were added 4x by weight to each sample, followed by incubation for 2 h at room temperature. The reaction was quenched by addition of 0.5% hydroxylamine. Samples were combined, acidified by TFA, cleaned using Sep-PaK cartridges (Waters) and dried using DNA 120 SpeedVac™ Concentrator (Thermo). Samples were then resuspended in 0.1% TFA and fractionated into 8 fractions using Pierce™ High pH Reversed-Phase Peptide Fractionation Kit (Thermo).

Samples dissolved in buffer A (0.1% formic acid and 2% ACN in water) were loaded onto a 50 cm EASY-Spray column (75 µm internal diameter, packed with PepMap C18, 2 µm beads, 100 Å pore size) connected to the EASY-nLC 1000 (Thermo) and eluted with a buffer B (98% ACN, 0.1% FA, 2% H_2O) gradient from 5% to 38% of at a flow rate of 250 nL/min. The eluent was ionized by electrospray, with molecular ions entering an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific). The survey mass spectrum was acquired at a nominal resolution of 120,000, with the m/z range from 400 to 1600. Precursors were isolated in 3s cycle time using 0.7 Da isolation width and fragmented via HCD with excitation set at 40%. MS2 spectra were acquired at 60,000 resolution, 105 ms maximum injection time and AGC target of 1e5. Dynamic exclusion was set to 60s.

Data processing

The raw LC-MS data (SIESTA) were analyzed by MaxQuant, version 1.5.6.5 (*47*). The Andromeda search engine matched MS/MS data against the Uniprot complete proteome database (human, version UP000005640_9606, 92957 entries). TMT10-plex on the MS/MS level was used for quantification of protein abundances. Cysteine carbamidomethylation was used as a fixed modification, while methionine oxidation was selected as a variable modification. Trypsin/P was selected as enzyme specificity. No more than two missed cleavages were allowed. A 1% false discovery rate was used as a filter at both protein and peptide levels. For all other parameters, the default settings were used. After removing all the contaminants, only proteins with at least two peptides were included in the final dataset.

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Network mapping

For GO term and pathway analyses, STRING version 10.5 (http://string-db.org) protein network analysis tool was used (*48*). Medium confidence threshold (0.4) was used to define protein-protein interactions. In-built gene set enrichment analysis with the whole genome as a background was used to identify enriched gene ontology terms and pathways.

Validation of mono-ADP-ribosylation by targeted tandem mass spectrometry

Recombinant RFK (5 µg) and PDRG1 (5 µg) were diluted with 50 Mm HEPES (pH = 7.5), 0.5 mM TCEP, 100 mM NaCl, 100 µM NAD, 4 mM $MgCl_2$ and incubated with 400 nM of PARP10 for 1 h. Proteins were reduced with 10 mM DTT for 30 min and alkylated with 50 mM IAA for 30 min in the dark. Afterwards, 1 M urea was added to the samples and LysC (overnight) and Trypsin (6 h) were added sequentially at 1 : 100 w/w to protein. After acidification, samples were cleaned using StageTips. Samples were dissolved in 0.1% FA and 1 µg of each samples was analyzed with LC-MS using a 1 h gradient.

The chromatographic separation of peptides was achieved using a 50 cm Easy C18 column connected to an Easy1000 LC system (Thermo Fisher Scientific). The peptides were loaded onto the column at a flow rate of 1000 nL/min, and then eluted at 300 nL/min for 50 min with a linear gradient from 4% to 26% ACN/0.1% formic acid. The eluted peptides were ionized with electrospray ionization and analyzed by an Orbitrap Fusion mass spectrometer. The survey mass spectrum was acquired at the resolution of 120,000 in the m/z range of 300-1750. The first MS/MS event data were obtained with a HCD at 32% excitation for ions isolated in the quadrupole with a m/z width of 1.6 at a resolution of 30,000. Mass trigger filters targeting adenine, adenosine and AMP ions were used to initiate a second MS/MS event using ETD MS/MS with HCD supplementary activation at 30% collision energy and with a 30,000 resolution. Samples treated with NAD but no PARP10 were used as negative controls.

Spectra were converted to Mascot generic format (MGF) using in-house written RAWtoMGF v. 2.1.3. The MGFs files were then searched against the UniProtKB human database (v. 201806), which

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

included 71434 sequences. Mascot 2.5.1 (Matrix Science) was used for peptide sequence identification. Enzyme specificity was set to trypsin, allowing up to two missed cleavages. C, D, E, K, N, R and S residues were set as variable ADP-ribose acceptor sites. Carbamidomethylation was set as a fixed modification on C and oxidation as a variable modification on M.

***In vitro* mono-ADP-ribosylation assay**

Hexahistidine-tagged PARP10 catalytic domain (auto-modification) or protein substrate (substrate protein modification) was immobilized on Ni^{2+}-chelating microplates (5-PRIME). TEV-cleaved PARP10 catalytic domain was used for evaluation of substrate protein modification. Mono-ADP-ribosylation was assessed after incubation with 100 µM NAD^+ (including 2% biotinylated NAD^+, Trevigen) prior to chemiluminescence detection of biotinyl-ADP-ribose in a Clariostar microplate reader (BMG Labtech) as described in detail before (*46*).

Statistical Analysis

Curve normalization and fitting was done by an in-house R package, available in GitHub (https://github.com/RZlab/SIESTA). Briefly, after removing the contaminant proteins and those quantified with less than two peptides, protein abundances in temperature points 41-67 °C were normalized to the total proteome melting curve similar to Franken at al. (*49*). Individual protein abundances were scaled so that the lowest temperature intensity was set to 1. For each protein in each replicate, a sigmoid curve was fitted using non-linear least squares method according the formula:

$$I \sim (1-Pl) / [1 + \exp((T-T_m)/bT)] + Pl,$$

where Pl – high-temperature plateau of the melting curve, T_m – melting temperature, b – slope of the curve.

Difference in T_m between samples was assessed using modified t-test. T-statistics was calculated assuming normal distribution of T_m estimation errors with non-equal variances between replicates. T-test

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

derived p-values were adjusted for multiple comparison using Benjamini–Hochberg method. The melting curves for all putative candidates were also manually inspected to exclude false positives.

Acknowledgements

We are grateful to Marie Ståhlberg and Carina Palmberg for their assistance in different proteomic experiments. **Funding:** The research is funded by grants from the Knut and Alice Wallenbergs Foundation (grant KAW 2015.0063) and VINNOVA (grant Oxidocurin) awarded to R.Z.; E.A. is supported by grants from Karolinska Institutet, Swedish Cancer Society and Swedish Research Council, and H.S. by Swedish Research Council (grant 2015-4603). K.N. was supported by a stipend from the Wenner-Gren foundation. **Authors contributions:** Conceptualization, R.A.Z. and A.A.S.; methodology and experiment design, A.A.S., R.A.Z., J.A.W., H.S., E.A., S.R. and T.K.; project organization, training, resources and funding acquisition, R.A.Z., H.S. and E.A; SIESTA experiments, A.A.S. and P.S.; targeted mass spectrometry for ADP-ribose confirmation, A.A.S. and A.V.; data analysis and visualization, A.A.S., C.B. and A.C.; production and testing of enzymes, A.G.T. and Q.C.; *in vitro* mono-ADP-ribosylation assays, K.N.; writing—original draft, A.A.S. and R.A.Z.; writing—review & editing, A.A.S., R.A.Z., H.S., E.A and J.A.W. **Competing interests:** Authors declare no conflicts of interest. **Data and materials availability:** Excel files containing the analyzed data are provided in Supplementary Materials. The mass spectrometry data were deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (*50*) with the dataset identifier PXD010554.

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

List of Supplementary Materials

Materials and Methods

Tables S1 to S7

Fig. S1

References and Notes

1. A. Chang, I. Schomburg, S. Placzek, L. Jeske, M. Ulbrich, M. Xiao, C. W. Sensen, D. Schomburg, BRENDA in 2015: exciting developments in its 25th year of existence. *Nucleic Acids Res*. **43**, D439-D346 (2014).
2. M. Mann, O. N. Jensen, Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* **21**, 255-261 (2003).
3. D. Houde, Y. Peng, S. A. Berkowitz, J. R. Engen, Post-translational modifications differentially affect IgG1 conformation and receptor binding. *Mol. Cell. Proteomics* **9**, 1716-1728 (2010).
4. B. T. Weinert, T. Narita, S. Satpathy, B. Srinivasan, B. K. Hansen, C. Schölz, W. B. Hamilton, B. E. Zucconi, W. W. Wang, W. R. Liu, J. M. Brickman, E. A. Kesicki, A. Lai, K. D. Bromberg, P. A. Cole, C. Choudhary, Time-Resolved Analysis Reveals Rapid Dynamics and Broad Scope of the CBP/p300 Acetylome. *Cell* **174**, 231-244 (2018).
5. O. Von Ahsen, U. Bömer, High-throughput screening for kinase inhibitors. *ChemBioChem* 6, 481-490 (2005).
6. M. Steger, F. Tonelli, G. Ito, P. Davies, M. Trost, M. Vetter, S. Wachter, E. Lorentzen, G. Duddy, S. Wilson, M. A. Baptista, B. K. Fiske, M. J. Fell, J. A. Morrow, A. D. Reith, D. R. Alessi, M. Mann, Phosphoproteomics reveals that Parkinson's disease kinase LRRK2 regulates a subset of Rab GTPases. *eLife* **5**, e12813 (2016).
7. K. M. Byrd, B. S. Blagg, Chaperone substrate provides missing link for cancer drug discovery. *J. Biol. Chem.* **293**, 2381-2382 (2018).
8. H.-C. S. Yen, S. J. Elledge, Identification of SCF ubiquitin ligase substrates by global protein stability profiling. *Science* **322**, 923-929 (2008).
9. A. J. Flint, T. Tiganis, D. Barford, N. K. Tonks, Development of “substrate-trapping” mutants to identify physiological substrates of protein tyrosine phosphatases. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 1680-1685 (1997).
10. T. Y. Low, M. Peng, R. Magliozzi, S. Mohammed, D. Guardavaccaro, A. J. Heck, A systems-wide screen identifies substrates of the SCFβTrCP ubiquitin ligase. *Sci. Signal.* **7**, rs8 (2014).
11. M. Köhn, M. Gutierrez-Rodriguez, P. Jonkheijm, S. Wetzel, R. Wacker, H. Schroeder, H. Prinz, C. M. Niemeyer, R. Breinbauer, S. E. Szedlacsek, H. Waldmann, A microarray

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

strategy for mapping the substrate specificity of protein tyrosine phosphatase. *Angew. Chem. Int. Ed.* **46**, 7700-7703 (2007).

12. T. Feilner, C. Hultschig, J. Lee, S. Meyer, R. G. Immink, A. Koenig, A. Possling, H. Seitz, A. Beveridge, D Scheel, D. J. Cahill, H. Lehrach, J. Kreutzberger, B. Kersten, High throughput identification of potential Arabidopsis mitogen-activated protein kinases substrates. *Mol. Cell. Proteomics* **4**, 1558-1568 (2005).

13. J. A. Ubersax, E. L. Woodbury, P. N. Quang, M. Paraz, J. D. Blethrow, K. Shah, K. M. Shokat, D. O. Morgan, Targets of the cyclin-dependent kinase Cdk1. *Nature* **425**, 859-864 (2003).

14. S. Matsuoka, B. A. Ballif, A. Smogorzewska, E. R. McDonald 3rd, K. E. Hurov, J. Luo, C. E. Bakalarski, Z. Zhao, N. Solimini, Y. Lerenthal, Y. Shiloh, S. P. Gygi, S. J. Elledge, ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **316**, 1160-1166 (2007).

15. L. Anders, N. Ke, P. Hydbring, Y. J. Choi, H. R. Widlund, J. M. Chick, H. Zhai, M. Vidal, S. P. Gygi, P. Braun, P. Sicinski, A systematic screen for CDK4/6 substrates links FOXM1 phosphorylation to senescence suppression in cancer cells. *Cancer Cell* **20**, 620-634 (2011).

16. C. Park, S. Marqusee, Pulse proteolysis: a simple method for quantitative determination of protein stability and ligand binding. *Nat. Methods* **2**, 207-212 (2005).

17. F. H. Niesen, H. Berglund, M. Vedadi, The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2**, 2212-2221 (2007).

18. D. Martinez Molina, R. Jafari, M. Ignatushchenko, T. Seki, E. A. Larsson, C. Dan, L. Sreekumar, Y. Cao, P. Nordlund, Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science* **341**, 84-87 (2013).

19. M. M. Savitski, F. B. Reinhard, H. Franken, T. Werner, M. F. Savitski, D. Eberhard, D. Martinez Molina, R. Jafari, R. B. Dovega, S. Klaeger, B. Kuster, P. Nordlund, M. Bantscheff, G. Drewes, Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science* **346**, 1255784 (2014).

20. W. R. Drake, C.-W. Hou, N. E. Zachara, C. L. Grimes, New use for CETSA: monitoring innate immune receptor stability via post-translational modification by OGT. *J. Bioenerg. Biomembr.* **50**, 231-240 (2018).

21. I. Becher, A. Andrés-Pons, N. Romanov, F. Stein, M. Schramm, F. Baudin, D. Helm, N. Kurzawa, A. Mateus, M. T. Mackmull, A. Typas, C. W. Müller, P. Bork, M. Beck, M. M. Savitski, Pervasive Protein Thermal Stability Variation during the Cell Cycle. *Cell* **173**, 1495-1507 (2018).

22. L. Dai, T. Zhao, X. Bisteau, W. Sun, N. Prabhu, Y. T. Lim, R. M. Sobota, P. Kaldis, P. Nordlund, Modulation of Protein-Interaction States through the Cell Cycle. *Cell* **173**, 1481-1494 (2018).

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

23. A. Chernobrovkin, C. Marin-Vicente, N. Visa, R. A. Zubarev, Functional Identification of Target by Expression Proteomics (FITExP) reveals protein targets and highlights mechanisms of action of small molecule drugs. *Sci. Rep.* **5**, 11176 (2015).

24. A. A. Saei, A. Chernobrovkin, P. Sabatier, B. Zhang, C. Beusch, U. Guler Tokat, M. Gaetani, A. Vegvari, R. Zubarev, https://www.biorxiv.org/content/early/2018/09/18/421115. (2018).

25. E. S. Arnér, A. Holmgren, Physiological functions of thioredoxin and thioredoxin reductase. *FEBS J.* **267**, 6102-6109 (2000).

26. M. Bylesjö, M. Rantalainen, O. Cloarec, J. K. Nicholson, E. Holmes, J. Trygg, OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *J. Chemom.* **20**, 341-351 (2006).

27. R. Brigelius-Flohé, M. Maiorino, Glutathione peroxidases. *Biochim. Biophys. Acta Gen. Subj.* **1830**, 3289-3303 (2013).

28. M. Björnstedt, J. Xue, W. Huang, B. Akesson, A. Holmgren, The thioredoxin and glutaredoxin systems are efficient electron donors to human plasma glutathione peroxidase. *J. Biol. Chem.* **269**, 29382-29384 (1994).

29. H. Z. Chae, H. J. Kim, S. W. Kang, S. G. Rhee, Characterization of three isoforms of mammalian peroxiredoxin that reduce peroxides in the presence of thioredoxin. *Diabetes Res. Clin. Pract.* **45**, 101-112 (1999).

30. A. Jiménez, M. Pelto-Huikko, J. A. Gustafsson, A. Miranda-Vizuete, Characterization of human thioredoxin-like-1: Potential involvement in the cellular response against glucose deprivation. *FEBS Lett.* **580**, 960-967 (2006).

31. Y. Funato, T. Michiue, M. Asashima, H. Miki, The thioredoxin-related redox-regulating protein nucleoredoxin inhibits Wnt–β-catenin signalling through Dishevelled. *Nat. Cell Biol.* **8**, 501-508 (2006).

32. C. Y. Hwang, Y. S. Ryu, M. S. Chung, K. D. Kim, S. S. Park, S. K. Chae, H. Z. Chae, K. S. Kwon, Thioredoxin modulates activator protein 1 (AP-1) activity and p27Kip1 degradation through direct interaction with Jab1. *Oncogene* **23**, 8868-8875 (2004).

33. W. Jeong, Y. Jung, H. Kim, S. J. Park, S. G. Rhee, Thioredoxin-related protein 14, a new member of the thioredoxin family with disulfide reductase activity: Implication in the redox regulation of TNF-α signaling. *Free Radic. Biol. Med.* **47**, 1294-1303 (2009).

34. P. G. Board, M. Coggan, G. Chelvanayagam, S. Easteal, L. S. Jermiin, G. K. Schulte, D. E. Danley, L. R. Hoth, M. C. Griffor, A. V. Kamath, M. H. Rosner, B. A. Chrunyk, D. E. Perregaux, C. A. Gabel, K. F. Geoghegan, J. Pandit, Identification, characterization, and crystal structure of the Omega class glutathione transferases. *J. Biol. Chem.* **275**, 24798-24806 (2000).

35. J. Lundström, A. Holmgren, Protein disulfide-isomerase is a substrate for thioredoxin reductase and has thioredoxin-like activity. *J. Biol. Chem.* **265**, 9114-9120 (1990).

36. É. Dóka, I. Pader, A. Bíró, K. Johansson, Q. Cheng, K. Ballagó, J. R. Prigge, D. Pastor-Flores, T. P. Dick, E. E. Schmidt, E. S. Arnér, P. Nagy, A novel persulfide detection

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

method reveals protein persulfide-and polysulfide-reducing functions of thioredoxin and glutathione systems. *Sci. Adv.* **2**, e1500968 (2016).

37. H Kleine, E Poreba, K Lesniewicz, P. O. Hassa, M. O. Hottiger, D. W. Litchfield, B. H. Shilton, B. Lüscher, Substrate-assisted catalysis by PARP10 limits its activity to mono-ADP-ribosylation. *Mol. Cell* **32**, 57-69 (2008).

38. R. Gupte, Z. Liu, W. L. Kraus, PARPs and ADP-ribosylation: recent advances linking molecular functions to biological outcomes. *Genes Dev.* **31**, 101-126 (2017).

39. R Martello, M Leutert, S Jungmichel, V Bilan, SC Larsen, C Young, MO Hottiger, ML Nielsen, Proteome-wide identification of the endogenous ADP-ribosylome of mammalian cells and tissue. *Nat. Commun.* **7**, 12917 (2016).

40. I. Carter-O'Connell, H. Jin, R. K. Morgan, L. L. David, M. S. Cohen, Engineering the substrate specificity of ADP-ribosyltransferases for identifying direct protein targets. *J. Am. Chem. Soc.* **136**, 5201–5204 (2014).

41. F. J. Bock, T. T. Todorova, P. Chang, RNA regulation by poly (ADP-ribose) polymerases. *Mol. Cell* **58**, 959-969 (2015).

42. I Carter-O'Connell, H. Jin, R. K. Morgan, R. Zaja, L. L. David, I. Ahel, M. S. Cohen, Identifying family-member-specific targets of mono-ARTDs by using a chemical genetics approach. *Cell Rep.* **14**, 621-631 (2016).

43. N. Herzog, J. D. Hartkamp, P. Verheugd, F. Treude, A. H. Forst, K. L. Feijs, B. E. Lippok, E. Kremmer, H. Kleine, B. Lüscher, Caspase-dependent cleavage of the mono-ADP-ribosyltransferase ARTD10 interferes with its pro-apoptotic function. *FEBS J.* **280**, 1330-1343 (2013).

44. S. Vaidya, E. M. Velázquez-Delgado, G. Abbruzzese, J. A. Hardy, Substrate-induced conformational changes occur in all cleaved forms of caspase-6. *J. Mol. Biol.* **406**, 75-91 (2011).

45. Q. Cheng, E. S. Arnér, Selenocysteine insertion at a predefined UAG codon in a release factor 1 (RF1) depleted Escherichia coli host strain bypasses species barriers in recombinant selenoprotein translation. *J. Biol. Chem.* **292**, 5476-5487 (2017).

46. A. G. Thorsell, T. Ekblad, T. Karlberg, M. Löw, A. F. Pinto, L. Trésaugues, M. Moche, M. S. Cohen, H. Schüler, Structural basis for potency and promiscuity in poly (ADP-ribose) polymerase (PARP) and tankyrase inhibitors. *J. Med. Chem.* **60**, 1262-1271 (2016).

47. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367-1372 (2008).

48. D. Szklarczyk, J. H. Morris, H. Cook, M. Kuhn, S. Wyder, M. Simonovic, A. Santos, N. T. Doncheva, A. Roth, P. Bork, L. J. Jensen, C. von Mering, The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**, D362-D368 (2016).

49. H. Franken, T. Mathieson, D. Childs, G.M. Sweetman, T. Werner, I. Tögel, C. Doce, S. Gade, M. Bantscheff, G. Drewes, FB. Reinhard, W. Huber, MM. Savitski, Thermal

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. *Nat. Protoc.* **10**, 1567-1593 (2015).

50. J. A. Vizcaíno, E. W. Deutsch, R. Wang, A. Csordas, F. Reisinger, D. Ríos, J. A. Dianes, Z. Sun, T. Farrah, N. Bandeira, P. A. Binz, I. Xenarios, M. Eisenacher, G. Mayer, L. Gatto, A. Campos, R. J. Chalkley, H. J. Kraus, J. P. Albar, S. Martinez-Bartolomé, R. Apweiler, G. S. Omenn, L. Martens, A. R. Jones, H. Hermjakob, ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **32**, 223-226 (2014).

Table S7. The PARP10-mediated mono-ADP-ribosylation sites of RFK and PDRG1*

Protein	Start	End	Mascot Score	Peptide (site of modification in bold)
RFK	10	18	1	R.GQVV**R**GFGR.G + ADP-Ribosyl (R)
	108	127	61	K.NFDSL**E**SLISAIQGDIEEAK.K + Acetyl (N-term); ADP-Ribosyl (E)
	130	137	14	R.L**E**LPEHLK.I + ADP-Ribosyl (E)
	130	148	14	R.L**E**LPEHLKIKEDNFFQVSK.S + Acetyl (N-term); ADP-Ribosyl (E)
	138	148	39	K.IK**E**DNFFQVSK.S + ADP-Ribosyl (E)
PDRG1	29	37	1	R.QIV**D**LDTKR.N + ADP-Ribosyl (D)
	75	90	12	K.**E**MIEKDQDHLDKEIEK.L + ADP-Ribosyl (E)
	102	112	12	R.LFEAQGKP**E**LK.G + ADP-Ribosyl (E)

*The presence of fingerprint ions was manually verified in each spectrum.

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

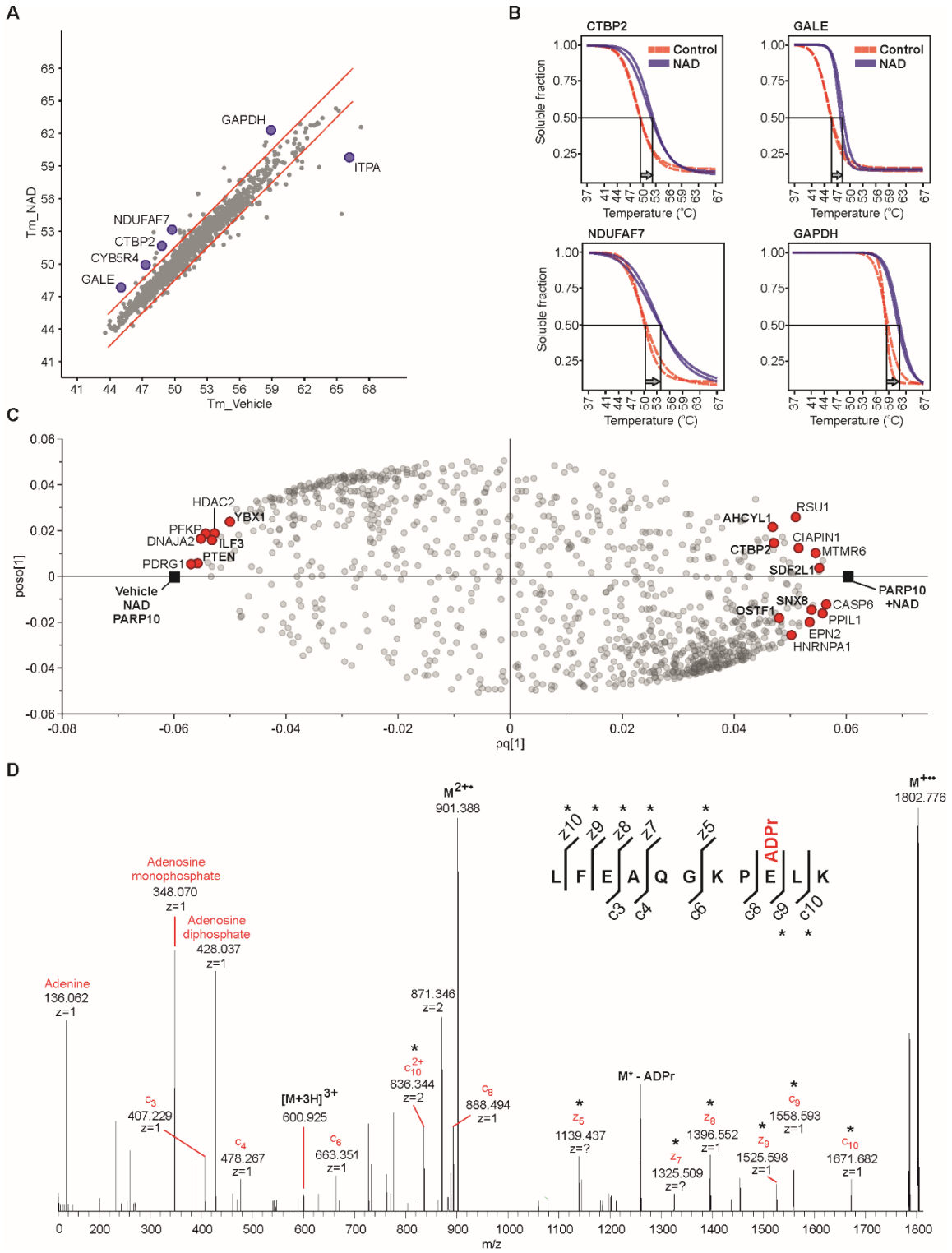

Fig. S1. Identification and validation of mono-ADP-ribosylated proteins. A) SIESTA data for vehicle and NAD were filtered. The proteins with a T_m difference between the replicates of ≤1 °C were used to

bioRxiv preprint doi: https://doi.org/10.1101/423418; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

perform a linear regression with a 95%-prediction cutoff. The outliers are the proteins presumably interacting with NAD; several proteins are known NAD interaction partners (purple circles). **B**) Representative melting curves of NAD sensor CTBP2 and other known NAD interactors. **C**) Loadings of OPLS-DA model contrasting the "PARP10+NAD" T_m vs. those in all other treatments singled out potential substrates (red circles indicate proteins with melting curve shifts verified by manual inspection). **D**) Mono-ADP-ribosylation on a glutamic acid residue Glu110 in the PDRG1 peptide with the highest sequence-fitting score revealed by ETD MS/MS of 3+ molecular ions.