SCIENTIFIC REPORTS

OPEN

An Integrated Chemical Proteomics Approach for Quantitative Profiling of Intracellular ADP-Ribosylation

Received: 14 November 2018
Accepted: 11 April 2019
Published online: 30 April 2019

Karunakaran Kalesh[1,2], Saulius Lukauskas[1,3], Aaron J. Borg[4], Ambrosius P. Snijders[4], Vinay Ayyappan[5], Anthony K. L. Leung[5], Dorian O. Haskard[6] & Peter A. DiMaggio[1]

ADP-ribosylation is integral to a diverse range of cellular processes such as DNA repair, chromatin regulation and RNA processing. However, proteome-wide investigation of its cellular functions has been limited due to numerous technical challenges including the complexity of the poly(ADP-ribose) (PAR) chains, low abundance of the modification and lack of sensitive enrichment methods. We herein show that an adenosine analogue with a terminal alkyne functionality at position 2 of the adenine (2-alkyne adenosine or 2YnAd) is suitable for selective enrichment, fluorescence detection and mass spectrometry proteomics analysis of the candidate ADP-ribosylome in mammalian cells. Although similar labelling profiles were observed via fluorescence imaging for 2YnAd and 6YnAd, a previously reported clickable NAD^+ precursor, quantitative mass spectrometry analysis of the two probes in MDA-MB-231 breast cancer cells revealed a significant increase in protein coverage of the 2YnAd probe. To facilitate global enrichment of ADP-ribosylated proteins, we developed a dual metabolic labelling approach that involves simultaneous treatment of live cells with both 2YnAd and 6YnAd. By combining this dual metabolic labelling strategy with highly sensitive tandem mass tag (TMT) isobaric mass spectrometry and hierarchical Bayesian analysis, we have quantified the responses of thousands of endogenous proteins to clinical PARP inhibitors Olaparib and Rucaparib.

Protein ADP-ribosylation is a post-translational modification (PTM) where members of a family of enzymes known as ADP-ribosyltransferases (or PARPs) covalently link ADP-ribose moieties derived from NAD^+ to the side chains of several possible amino acid residues (glutamate, aspartate, arginine, lysine, cysteine, serine) on target proteins. This PTM exists either as a monomer (mono-ADP-ribosylation), consisting of a single ADP-ribose moiety on the target residue, or as a polymer (poly-ADP-ribosylation) where a branched polymeric chain of ADP-ribose units originates from the target residue. Of the 17 PARPs encoded in the human genome, PARPs 1, 2 and 3 are known to have important roles in DNA regulation, particularly in the base excision repair pathway[1]. Consequently, PARP inhibitors have been developed as cancer therapeutics to exploit the concept of synthetic lethality in patients with germline mutations in *BRCA1* or *BRCA2* genes[2]. Beyond the established roles of these nuclear PARPs in DNA damage responses, the broad cellular functions of the majority of the other members of the PARP family remain elusive primarily due to the lack of analytical techniques for the large-scale profiling of intracellular ADP-ribosylation.

Mass spectrometry (MS) proteomics studies of ADP-ribosylation in particular have been limited due to numerous technical challenges. These include low specificity and/or affinity when using recombinant macro domains[3,4], *cis*-diol binding boronate beads[5], antibodies[6] or adapted phosphoproteome enrichment strategies[7] to enrich for this low abundance PTM against complex cellular backgrounds. Additionally, most methods reported to date require *a priori* knockdown of poly-ADP-ribose glycohydrolase (PARG) activity to sufficiently increase the baseline level of poly-ADP-ribosylation for MS detection, which is not ideal since PARG knockdown is known to induce physiological changes in cells[8]. A further limitation of many previous studies is that they have been

[1]Department of Chemical Engineering, Imperial College London, South Kensington Campus, London, SW7 2AZ, UK. [2]Present address: Department of Chemistry, Durham University, Stockton Road, Durham, DH1 3LE, UK. [3]Present address: Institute of Functional Epigenetics, Helmholtz Zentrum München, Neuherberg, Germany. [4]Mass Spectrometry Proteomics Platform, The Francis Crick Institute, London, NW1 1AT, UK. [5]Department of Biochemistry and Molecular Biology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, 21205, USA. [6]Faculty of Medicine, National Heart & Lung Institute, Vascular Science Section, Hammersmith Campus, London, UK. Correspondence and requests for materials should be addressed to P.A.D. (email: p.dimaggio@imperial.ac.uk)

performed under stress induction, which activates PARP1 and thus potentially masks the underlying activities of the other PARP enzymes[3–7].

Recently, it has been demonstrated that 6-alkyne adenosine (6YnAd), a compound that was previously demonstrated to be suitable for labelling poly(A) tails of mRNAs in mammalian cells[9], enables sensitive fluorescence profiling of ADP-ribosylated proteins in live cells[10]. In this work, we analysed the labelling efficiency of 6YnAd for the first time using tandem mass tag (TMT)-based quantitative proteomics in mammalian cells and found that 6YnAd alone limits substrate coverage, but that inclusion of a similar adenosine analogue, 2-alkyne adenosine (2YnAd), allows a more comprehensive assessment of the putative ADP-ribosylome. We also report an integrated chemical proteomics approach that involves simultaneous treatment of live cells with both 2YnAd and 6YnAd followed by robust enrichment of the labelled proteome and its sensitive profiling by TMT quantitative mass spectrometry.

Results

2-alkyne adenosine (2YnAd) treatment results in labelling of ADP-ribosylated proteins. We treated MDA-MB-231 breast cancer cells with equal concentrations of 2YnAd and 6YnAd in parallel. The cells were then lysed and the whole-cell proteome was clicked with a trifunctional capture reagent Azido-TAMRA-Biotin (Suppl. Fig. 1) and resolved on SDS-PAGE. In-gel fluorescence scan revealed qualitatively similar protein labelling profiles for both 2YnAd and 6YnAd treatments (Fig. 1A,B). Similarly, following affinity capture of the labelled proteins on NeutrAvidin-Agarose resin, the combination of in-gel fluorescence imaging (Fig. 1C) and Western blot analysis using an anti-pan-ADP-ribose antibody (Fig. 1D)[11] confirmed significant enrichment of ADP-ribosylated proteins in both 2YnAd and 6YnAd treated cells relative to the input control lanes.

2YnAd provides greater coverage of ADP-ribosylated targets than 6YnAd. After establishing the use of the 2YnAd analogue for live cell labelling of ADP-ribosylated proteins, we performed a proteomic comparison of the labelling between 2YnAd and 6YnAd using the highly sensitive tandem mass tags (TMT)-based isobaric tagging for quantitative mass spectrometry[12]. Using the MDA-MB-231 breast cancer cell line as a model system, we first developed a workflow (Fig. 2) that involves the following sequential steps: (i) parallel metabolic incorporation of 2YnAd and 6YnAd, (ii) bioorthogonal tagging of the ADP-ribosylated proteome using a copper (I) catalysed azide-alkyne cycloaddition (CuAAC) click reaction, (iii) affinity enrichment of the labelled proteome using the robust biotin-NeutrAvidin pull-down system, (iv) TMT6plex isobaric labelling and (v) liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis.

Our chemical proteomics workflow for profiling intracellular ADP-ribosylation events identified over 1600 protein targets (Supplementary Dataset File 1, in which proteins previously reported in ADPriboDB[13] are highlighted in bold), representing a significant improvement in proteomic coverage compared to other published studies despite requiring less protein input (500 µg total protein input before affinity enrichment per condition). Importantly, statistical analysis of the replicate data sets using the hierarchical Bayesian model implemented in LIMMA[14] revealed that specific subsets of protein targets were preferentially labelled with either 2YnAd or 6YnAd treatment (Fig. 3, Suppl. Fig. 2), despite the overall qualitative similarity between the two probes in the labelling observed in the in-gel fluorescence scans (Fig. 1A). The 2YnAd was found to be preferentially incorporated, labelling more ADP-ribosylated protein targets than the 6YnAd (Fig. 3). Additionally, analysis of the biological pathways of these targets (Suppl. Fig. 3, Supplementary Dataset File 2) specifically revealed a 2YnAd enrichment for proteins involved in proteasome regulation, including several 26S proteasomal subunits.

Dual metabolic labelling using both 2YnAd and 6YnAd elucidates the quantitative response of thousands of ADP-ribosylated proteins to clinical PARP inhibitors. We rationalised that the observed quantitative differences in protein targets labelled by 2YnAd and 6YnAd could be exploited for more comprehensive profiling of the cellular ADP-ribosylome by simultaneously treating cells with both compounds and combining this dual metabolic labelling approach with our quantitative proteomics workflow. Additionally, as several PARP inhibitors have recently emerged as promising cancer therapeutics, we decided to evaluate the dual metabolic labelling strategy for comprehensively and quantitatively profiling global changes to the cellular ADP-ribosylome upon PARP inhibitor treatment (Fig. 4). We treated MDA-MB-231 cells with equimolar concentrations of 2YnAd and 6YnAd to investigate differential labelling of protein ADP-ribosylation in the presence of five different concentrations (0, 0.2, 1, 5, 25 µM) of the clinical PARP inhibitors Olaparib and Rucaparib[15]. The samples were combined in a TMT10plex experiment and over 2300 statistically significant hits were identified following hierarchical Bayesian analysis (Supplementary Dataset File 3), which is the largest experimental candidate data set representing the cellular ADP-ribosylome to date (Fig. 5A). Statistical analysis of the data sets revealed a correlative response in the levels of ADP-ribosylation with respect to Olaparib and Rucaparib treatment, which indicates a congruent set of PARP substrates (Fig. 5B), highlighting the robustness and reproducibility of this integrated chemical biology approach. Protein targets found to decrease in ADP-ribosylation upon treatment of both inhibitors were captured in the lower left quadrant of the scatter plot (Fig. 5B). Interestingly, several mRNA splicing factors (highlighted by cyan rectangles) were found among the most responsive proteins. In Fig. 5C, the dose-dependent response in the levels of ADP-ribosylation to both inhibitors is shown for two of these mRNA splicing proteins, namely pre-mRNA splicing factor 38B (PRPF38B) and serine/arginine repetitive matrix protein 2 (SRRM2). It is shown that PRPF38B is more sensitive to Olaparib treatment, whereas SRRM2 exhibits a congruent response to both PARP inhibitors. Pathway analysis of all the protein targets found to decrease in ADP-ribosylation upon treatment with the inhibitors also identified a preferential enrichment for factors involved in mRNA splicing (Supplementary Dataset File 4, Suppl. Fig. 4).

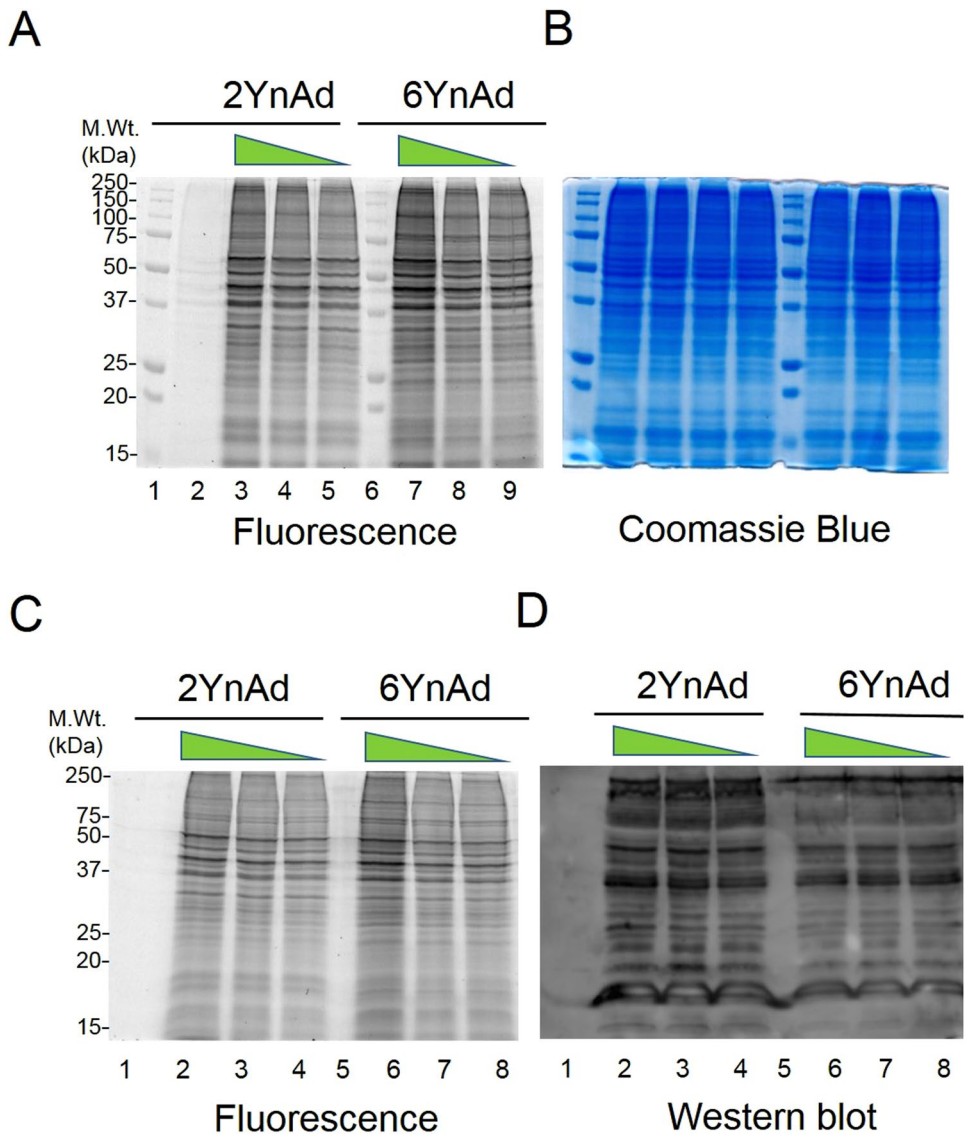

Figure 1. (**A**) Qualitative assessment of 2YnAd and 6YnAd labelling in MDA-MB-231 cells by in-gel fluorescence scanning. Lanes 1 and 6: molecular weight marker; lane 2: DMSO control; lanes 3, 4 and 5 (1 mM, 0.5 mM and 0.25 mM 2YnAd respectively) and lanes 7, 8 and 9 (1 mM, 0.5 mM and 0.25 mM 6YnAd respectively). (**B**) Coomassie blue staining of the same gel. (**C**) In-gel fluorescence scan following the metabolic incorporation, cell lysis, click chemistry and affinity enrichment. Lane 1: DMSO control; Lanes 2, 3 and 4: 1 mM, 0.5 mM and 0.25 mM respectively of 2YnAd was used for the metabolic labelling; lane 5: empty and lanes 6, 7 and 8: 1 mM, 0.5 mM and 0.25 mM respectively of 6YnAd was used for the metabolic labelling. (**D**) Western blot of the same gel after electrotransfer of proteins onto a nitrocellulose membrane probed using an anti-pan-ADP-ribose antibody (MABE1016, EMD Millipore).

Discussion

Functionalised full-length NAD^+ analogues have recently emerged as promising tools for fluorescence-based visualisation of ADP-ribosylation in cell-lysates[11,16–18] and in cells[19,20]. Due to the lack of cell-permeability of the full-length NAD^+ analogues, intracellular labelling has been performed by transient permeabilisation of the cell membrane using Triton X-100 surfactants. Although permeabilisation of cell membrane using the surfactant is a powerful approach for imaging applications, successful adaptation of the method to large-scale mass spectrometry (MS) proteomics studies of ADP-ribosylation has not been demonstrated to date. Nevertheless, a chemical genetic discovery method using clickable full-length NAD^+ analogues has recently enabled selective enrichment and MS profiling of ADP-ribosylated proteins, although the lack of cell permeability of the NAD^+ analogues has restricted these studies to cell lysates[11]. In a recent advancement in the field, Howard Hang and co-workers reported a chemical biology approach for live cell labelling of ADP-ribosylated proteins using a terminal alkyne functionalised and cell-permeable NAD^+-precursor compound, 6-alkyne adenosine (6YnAd)[10]. Although successful labelling of a few hundred ADP-ribosylated proteins was demonstrated using 6YnAd, previous studies have shown that the corresponding full-length clickable NAD^+ cofactor (6-alkyne-NAD) is a weaker substrate of

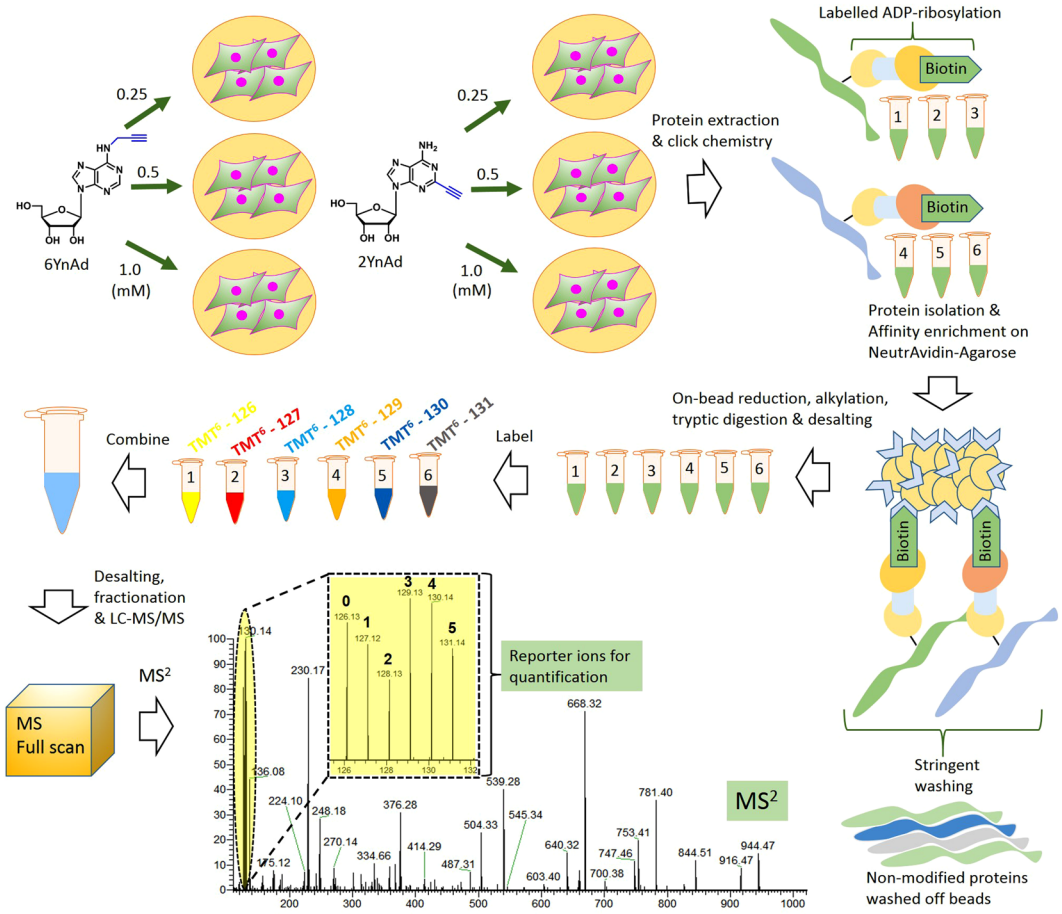

Figure 2. Schematic illustrating the experimental workflow for quantitative comparison of metabolic labelling between 2YnAd and 6YnAd. MDA-MB-231 cells were metabolically labelled with 3 different concentrations of 2YnAd and 6YnAd in parallel. After cell lysis and click reaction with a biotin-containing capture reagent, the labelled proteins are affinity enriched on NeutrAvidin-Agarose resins. Stringent washing at this stage removes unmodified proteins from the beads. After on-bead reduction, alkylation and tryptic digestion, the eluted peptides are desalted and each sample is labelled using a unique TMT6plex channel. After quenching the TMT reactions, the 6 samples are mixed together, fractionated and subjected to nanoflow LC-MS/MS analysis. Quantification of reporter ions from the MS/MS spectra provides relative quantification of labelling efficiencies of the two compounds at all three tested concentrations.

PARP1 than endogenous NAD^+ [16]. This raised the question as to whether the placement of the alkyne group on the NAD^+ cofactor analogue affects binding and utilisation by the different endogenous PARP enzymes.

We aimed to further develop the metabolic labelling strategy using clickable NAD^+ precursors, as this approach has the unique advantage of labelling of the ADP-ribosylome directly in the native cellular environment. Since 2-alkyne adenosine (2YnAd), the metabolic precursor for the corresponding full-length clickable NAD^+ analogue 2-alkyne-NAD^+, was also reported to efficiently label the poly(A) tails of mRNA transcripts in live cells[21] similar to 6YnAd and recent studies have demonstrated that PARP enzymes are more tolerant to substitutions in this position on NAD^+ [19,20], we investigated if 2YnAd can also be used for the metabolic labelling of ADP-ribosylome. Our results (Fig. 1) showed for the first time that 2YnAd is also useful for the in-cell labelling of ADP-ribosylated proteins similar to the 6YnAd probe.

As with most chemical proteomics approaches, we cannot eliminate the possibility that a fraction of the clickable adenosine analogues might be utilised for other types of protein modifications. Comparison of the fluorescence and Western-blot experiments (Fig. 1C,D) reveals similar band patterns, thus confirming the enrichment of ADP-ribosylated proteins, although differences in labelling were also observable between the two. However, these techniques are fundamentally different and have their own limitations in sensitivity and selectivity, which can vary with protein substrate, and neither are ideal for detecting protein-specific changes in complex mixtures. Therefore, we decided to apply quantitative mass spectrometry-based proteomics using the highly sensitive TMT isobaric tagging to conduct a head-to-head comparison of the labelling of protein targets by the two probes. The TMT labelling strategy offers sample multiplexing and importantly the isobaric nature of the tag results in the pooling of peptide signal across all experimental conditions, which increases the sensitivity in detecting lower abundance ADP-ribosylated proteins relative to any non-specific background. Hierarchical Bayesian analysis of

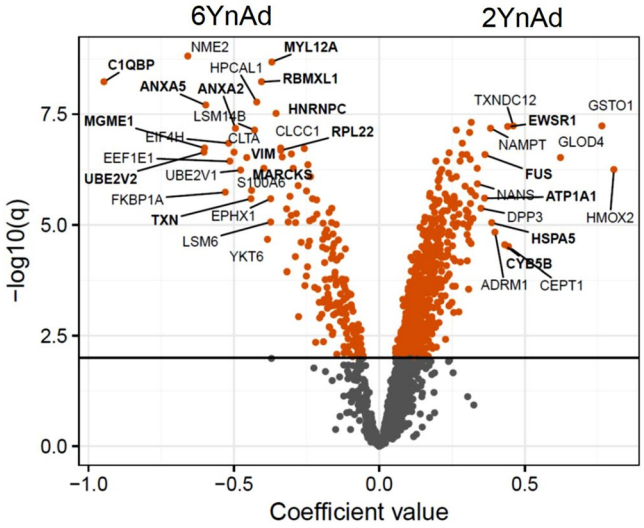

Figure 3. Volcano plot for the cofactor effect coefficients in the linear model, where positive values on the x-axis indicate preference for 2YnAd and negative values indicate preference for 6YnAd. The significance of the cofactor effect coefficients is computed using a moderated t-test based on the hierarchical Bayesian model as implemented in LIMMA, and adjusted p values (−log10 scale) are plotted on the y-axis with significantly enriched proteins coloured in red. Proteins reported to be ADP-ribosylated in other studies from ADPriboDB[22] are highlighted in bold.

our TMTsixplex datasets using LIMMA revealed that the 2YnAd was preferentially incorporated as it produced more ADP-ribosylated proteins than the 6YnAd probe (Fig. 3), which is only hinted at by the relatively insensitive Western-blot analysis (Fig. 1D). This suggests the possibility that the endogenous PARP enzymes elicit a differential structural preference for the two NAD^+ cofactor analogues.

Interestingly, biological pathway analysis of the identified targets revealed enrichment of several proteins involved in proteasome regulation specifically among the 2YnAd targets, including several 26S proteasomal subunits. 26S proteasomal activity has been previously reported to be promoted by ADP-ribosylation[22]. Similarly, a co-operative interaction between the proteasome subunit PSMD14 and BRCA1 has been implicated in cellular responses to DNA double strand breaks[23], a niche where nuclear PARPs are known to have key functional roles. It has also been suggested that the negatively charged PARylation may serve as a recognition site for the E3 ubiquitin protein ligases, which facilitates subsequent degradation of the protein in the 26S proteasome[24]. However, the precise biological function of ADP-ribosylation in proteasome regulation remains unknown.

Our dual metabolic labelling approach involving simultaneous treatment of cells with equimolar concentrations of 2YnAd and 6YnAd probes and its combination with the TMT10plex quantitative proteomics resulted in the largest candidate experimental dataset of cellular ADP-ribosylome with over 2300 significant proteins. The superior sensitivity of our proteomics technology enabled the systematic profiling of the cellular ADP-ribosylome in an unbiased manner. As an example application, we utilised this methodology to elucidate the first direct and quantitative comparison between the intracellular response in ADP-ribosylation to two clinical PARP inhibitors, Olaparib and Rucaparib, which are able to induce synthetic lethality in homologous recombination (HR)-deficient cancers due to germline and/or somatic mutations in *BRACA1* or *BRACA2* genes[2]. In particular, both inhibitors are FDA-approved for the treatment of patients with advanced ovarian cancer and are used in the maintenance treatment of recurrent ovarian epithelial, fallopian tube and primary peritoneal cancer. Olaparib is also approved for the treatment of *BRACA1/2* mutated metastatic breast cancer and it is expected that these PARP inhibitors will find further applications in treating other HR-deficient cancers.

In Fig. 5B it is shown that several mRNA splicing factors exhibited the largest response to the PARP inhibitor treatments, which is in agreement with recent reports indicating pivotal functional roles of PARylation in mRNA splicing and modulation of RNA[25–29]. This finding is further supported by a recent CRISPR screen that identified several high-confidence mRNA splicing factors, including DDX46 and SRSF11 (lower left quadrant in Fig. 5B), which resulted in increased sensitivity to the PARP inhibitor Olaparib across several cell lines when mutated[30].

Due to the heterogeneity of the target residues and the variations of the possible branched polymer structures for ADP-ribosylation, the development of robust mass spectrometry-based methods for comprehensive site localisation of this modification at the proteome-scale remains a challenge. Our approach circumvents the inherent bias arising from site promiscuity by enriching the target proteins via the modification and then elucidating the quantitative changes at the protein level across a variety of experimental conditions via multiplexed TMT labelling mass spectrometry and hierarchical Bayesian analysis. We have also developed a complementary methodology where we enrich for the modification at the peptide-level (unpublished) and then perform MS/MS site localisation after treatment with reagents such as hydroxylamine to reduce the size of the modification[11]. Using this approach, we have confidently identified several sites of ADP-ribosylation (a few representative MS/MS annotations are provided in Suppl. Fig. 6) and also eliminated the possibility of other modifications that could originate from adenosine cofactors, such as AMPylation.

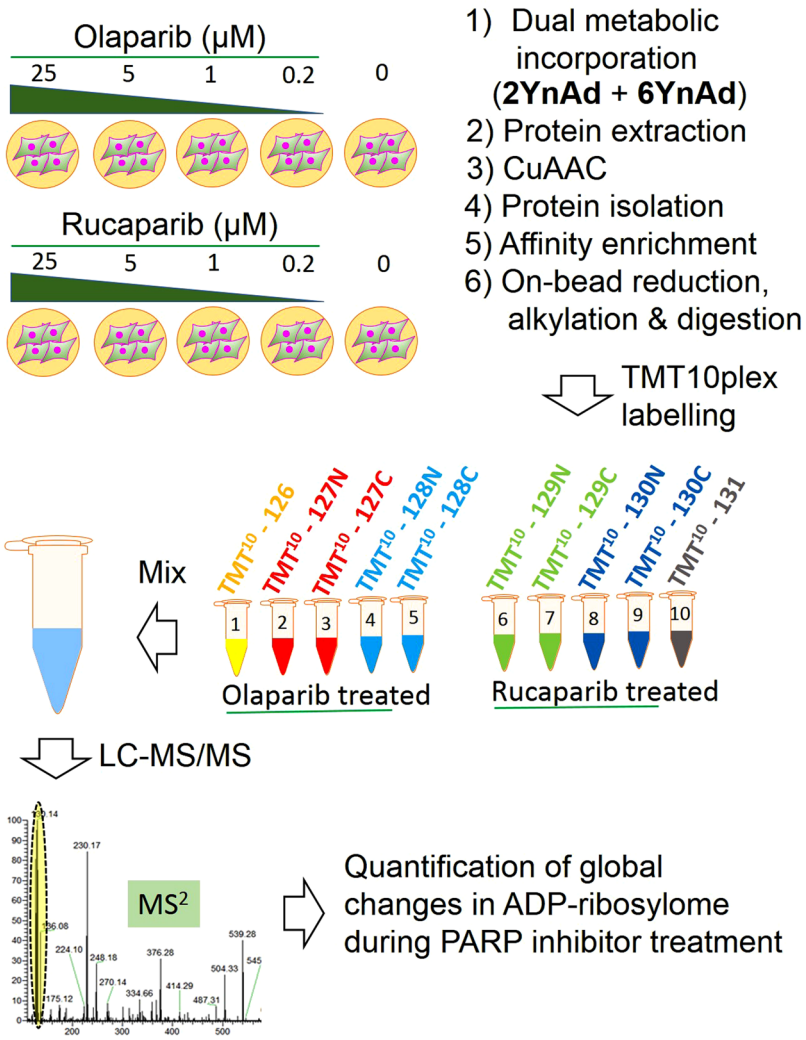

Figure 4. Schematic illustrating the workflow involving parallel treatment of PARP inhibitors Olaparib and Rucaparib followed by dual metabolic labelling of 6YnAd and 2YnAd, click chemistry, affinity enrichment of the labelled proteins on NeutrAvidin-Agarose beads and TMT10plex-based quantification of global changes in the cellular ADP-ribosylome.

In summary, we have presented a quantitative chemical proteomics approach that combines metabolic labelling using clickable adenosine analogues with TMT isobaric MS to measure ADP-ribosylation in live cells without experimental manipulation of PARG activity. We demonstrated that our method robustly profiled several thousands of intracellular protein targets, and yielded the largest quantitative proteomics data set measuring the ADP-ribosylome reported to date. The method is enabled by a dual metabolic labelling strategy using clickable NAD^+ precursors 6YnAd and 2YnAd, with the latter demonstrating superior protein coverage. Robust pull-down of the modified proteins is achieved using the selective NeutrAvidin-Biotin interaction and the pooling effect resulting from utilising TMT isobaric tagging to multiplex experimental conditions provides increased sensitivity. A head-to-head comparison between two clinical PARP inhibitors revealed a correlative response, which further highlights the sensitivity and robustness of this chemical proteomics platform for studying ADP-ribosylation in live cells. This approach has also been applied to profile global ADP-ribosylation in other immortalised cell lines (e.g. MCF-7 breast cancer and HEK-293T fluorescence labelling images are provided in Suppl. Fig. 7) as well as primary cell lines. We envision this enabling chemical proteomics technology will transform our ability to systematically characterise the intracellular functions of ADP-ribosylation in fundamental and disease biology, including cancer research, bacterial and viral infections and autoimmune diseases. In addition, the approach will be instrumental in determining differences in the targets of established PARP inhibitors and the effects of an emerging range of novel compounds designed to be selective for particular PARP enzymes[31].

Methods

Cell culture. MDA-MB-231 cells (ATCC) were grown in Dulbecco's Modified Eagle Medium (DMEM) (Gibco) supplemented with 10% foetal bovine serum (FBS) (Gibco) without antibiotics in a humidified incubator with 5% CO_2 at 37 °C and passaged at every 2–3 days at 80–90% confluency.

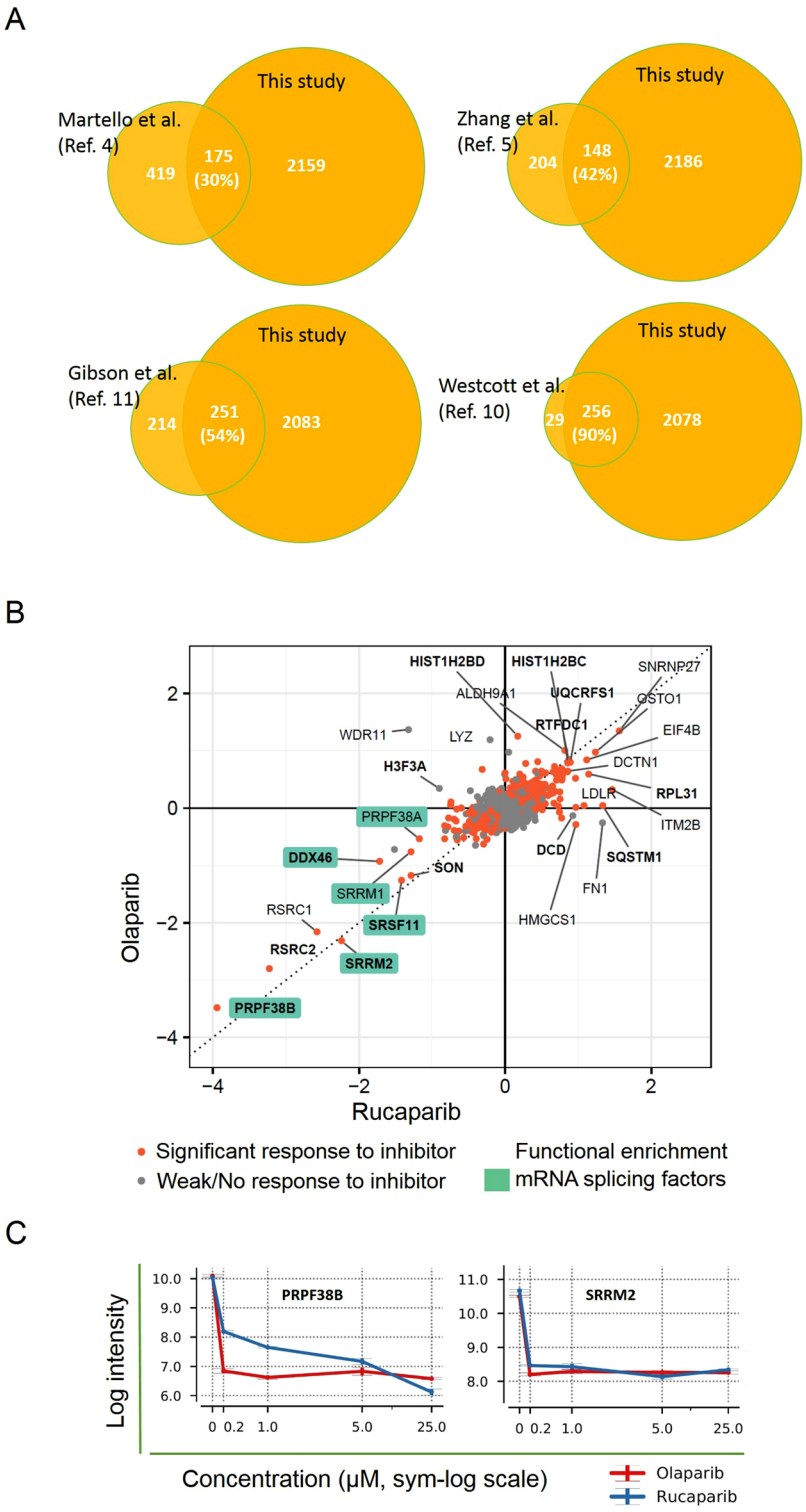

Figure 5. (**A**) Venn diagram assessment of the overlap between the proteins identified to be ADP-ribosylated in this study and previously published data sets. (**B**) Scatter plot showing mean log2 responses in the level of ADP-ribosylation of endogenous proteins to 25 μM of the inhibitors Olaparib and Rucaparib after removing the effects associated with technical replicate bias. Proteins whose response to inhibitors is significant (see LIMMA processing and statistical testing) are highlighted in orange. Proteins reported to be ADP-ribosylated in other studies from ADPriboDB[14] are highlighted in bold. Factors involved in mRNA splicing are highlighted with a cyan rectangle around their labels. (**C**) Quantitative changes in the reporter intensities (log2 scale) of mRNA splicing proteins pre-mRNA splicing factor 38B (PRPF38B) and serine/arginine repetitive matrix protein 2 (SRRM2) as a function of concentration of the inhibitors Olaparib (red) and Rucaparib (blue).

Metabolic labelling and preparation of cell extracts. In a 6-well plate, 4×10^5 cells were treated with 1 ml of fresh media (with 10% FBS) containing 2 µl of appropriate concentration of 6YnAd or 2YnAd (stock solutions in dimethyl sulfoxide [DMSO]) or DMSO control. For labelling experiments in 10 cm plates, 5 ml of fresh media with 10% FBS and 10 µl of the clickable adenosine analogue stock solutions were used. The cells were incubated at 37 °C and 5% CO_2 for 1 h. The cells were then washed with PBS three times and lysed using 4% SDS lysis buffer (50 mM HEPES pH 7.4, 150 mM NaCl, 4% SDS and 500 U Benzonase) and whole-cell lysates were collected using a cell scraper. Lysates were cleared at 16,000 g for 5 min and protein concentrations were estimated using Bio-Rad DC™ protein assay.

Click chemistry and in-gel fluorescence detection. The cell lysates at 1 mg/ml concentration were treated with freshly premixed click chemistry reaction cocktail (100 µM capture reagent, 1 mM $CuSO_4$ solution, 1 mM tris[2-carboxyethyl] phosphine hydrochloride [TCEP], 100 µM [tris(1-benzyl-1H-1,2,3-trazol-4-yl) methyl]amine [TBTA]) for 3 h at room temperature. The following stock solutions of the click chemistry reagents were used. Capture reagent (Azido-TAMRA-Biotin or Azide-PEG3-Biotin (Jena Bioscience)): 10 mM in DMSO; $CuSO_4$: 50 mM in MilliQ water; TCEP: 50 mM in MilliQ water; TBTA: 10 mM in DMSO. Proteins were precipitated by adding ice-cold methanol (4 volumes), chloroform (1.5 volumes) and water (3 volumes) and the precipitates were collected by centrifugation at 16,000 g for 5 min. The protein precipitates were then washed twice with ice-cold methanol (10 volumes) and the supernatants were discarded. The protein pellets were air dried for 20 min and suspended in resuspension buffer (4% SDS, 50 mM HEPES pH 7.4, 150 mM NaCl) to 3 mg/ml concentration. Equal volumes of 2X SDS loading buffer (Bio-Rad) containing 10% β-mercaptoethanol were added and the samples were boiled at 95 °C for 10 min and allowed to cool to room temperature. 30 µg of protein was loaded per gel lane (12% SDS Tris-HCl gels) and resolved by SDS-PAGE. The gels were scanned for fluorescence labelling using a GE typhoon 5400 gel scanner.

Affinity enrichment on NeutrAvidin-Agarose. The dried protein pellets obtained after click chemistry and protein precipitation were suspended in resuspension buffer (4% SDS, 50 mM HEPES pH 7.4, 150 mM NaCl) to 10 mg/ml concentration. In a typical experiment, 500 µg of whole-cell lysate after click chemistry and protein precipitation was resuspended in 50 µl of the resuspension buffer. The samples were subsequently diluted 20 fold with HEPES buffer (50 mM HEPES pH 7.5, 150 mM NaCl) so that the final SDS amount is 0.2%. 50 µl of NeutrAvidin-Agarose beads (Thermo scientific) were washed three times with 0.2% SDS buffer (0.2% SDS 50 mM HEPES pH 7.5, 150 mM NaCl) and added to the samples. The mixtures were rotated on an end-over-end rotating shaker for 1.5 h at room temperature. The beads were then washed 4 times with 1% SDS buffer (1% SDS, 50 mM HEPES pH 7.5, 150 mM NaCl), 4 times with 5 M urea buffer (5 M urea in 50 mM HEPES pH 7.5, 150 mM NaCl) and 4 times with 50 mM freshly prepared ammonium bicarbonate buffer. Each washing was performed with 20 volumes of the washing solutions and centrifugation of the beads between washing steps were carried out at 2,000 g for 1 min. For Western blot analysis, the beads, after one round of washing with 1% SDS buffer, were boiled in 4% SDS loading buffer and the proteins were collected by centrifugation. For LC-MS/MS analysis, on-bead reduction, alkylation and tryptic digestion were performed.

Western blotting. Proteins resolved by SDS-PAGE on 12% Tris-HCl gels were electrotransferred onto a nitrocellulose membrane. Western blotting for ADP-ribosylation was performed using anti-pan-ADP-ribose antibody (MABE1016, lot # 3074935, EMD Millipore) as described earlier with minor modifications[11]. Briefly, the nitrocellulose membrane after electrotransfer was blocked with freshly prepared 5% BSA in TBS-T for 1 hour at room temperature with constant shaking. The membrane was then incubated with the primary antibody (1:1000 dilution) in 5% BSA in TBS-T for 1 hour at room temperature with constant shaking. The membrane was then washed 4 times for 5 minutes with TBS-T and incubated with Goat-anti-rabbit IgG-HRP secondary antibody (1:5000 dilution) in 5% BSA in TBS-T for 1 hour at room temperature with constant shaking, washed again 4 times for 15 minutes with TBS-T and visualised with Pierce™ ECL Plus Western blotting substrate (Thermo Scientific, Prod # 32132, Lot # SD244941).

On-bead reduction, alkylation, tryptic digestion and desalting. The beads after affinity enrichment were resuspended in 100 µl of 50 mM ammonium bicarbonate buffer and treated with 3 mM dithiothreitol (DTT) (100 mM stock solution in water) solution for 30 min at room temperature. The beads were washed once with 50 mM ammonium bicarbonate buffer and resuspended in 100 µl of 50 mM ammonium bicarbonate buffer and treated with 10 mM iodoacetamide (IAA) (500 mM stock solution in water) in dark for 30 min at room temperature. The beads were again washed with 50 mM ammonium bicarbonate buffer and resuspended in 100 µl of fresh 50 mM ammonium bicarbonate buffer and treated with 2 µg of sequencing grade modified trypsin (Promega; Product code V5113) at 37 °C overnight. The samples were acidified to pH 3 using formic acid and allowed to stand for 5 min, centrifuged and collected the supernatant. The beads were washed with 0.1% formic acid solution in water, centrifuged and the supernatant were mixed with the previous supernatant. The collected tryptic peptides were desalted on C18 Empore disks on micropipette tip format and the desalted peptide mixture were dried on a speedvac.

TMTsixplex labeling. MDA-MB-231 cells were fed with three different concentrations (1, 0.5 and 0.25 mM) of 6YnAd and 2YnAd in parallel. After 1 h of incubation, whole-cell lysates were prepared and proteins were quantified. 500 µg of proteins from each condition were subjected to click chemistry using the capture reagent Azido-TAMRA-Biotin. Proteins were precipitated, re-solubilised and subjected to affinity enrichment on NeutrAvidin-Agarose beads. After extensive washings, the beads were subjected to DTT treatment, IAA treatment and overnight trypsinisation. The peptide digests were collected, desalted, dried and re-dissolved in 100 µl

of 100 mM triethylammonium bicarbonate (TEAB) buffer (pH 8.5). Each sample was subjected to TMT labelling with a unique TMTsixplex reagent for 1 h at room temperature. After quenching the TMT reactions with hydroxylamine (10 µL of 5% hydroxylamine per reaction for 15 min at room temperature) treatment, the six samples were mixed together, concentrated to complete dryness and re-dissolved in 0.1% formic acid. The sample was desalted on C18 Empore disks on micropipette tip format and the desalted sample was dried on a speedvac.

Mass-spectrometry analysis of TMT labelled samples. Peptides were reconstituted in 0.1% trifluoroacetic acid and chromatographically resolved using an Ultimate 3000 RSLCnano (Dionex) UHPLC. Peptides were first loaded onto an Acclaim PepMap 100 C18, 3 µm particle size, 100 Å pore size, 20 mm × 75 µm ID (Thermo Scientific) trap column heated to 40 °C, using a loading buffer (2% acetonitrile (MeCN) and 0.05% trifluoroacetic acid in 97.05% H_2O) with a flow rate of 7 µL/minute. Chromatographic separation was achieved using an EASY-Spray column, PepMap C18, 2 µm particles, 100 Å pore size, 500 mm × 75 µm ID (Thermo Scientific) heated to 40 °C. The gradient utilised a flow of 0.3 µL/minute, starting at 98% mobile A (0.1% formic acid, 5% DMSO in H_2O) and 2% mobile B (0.1% formic acid, 75% acetonitrile, 5% DMSO and 19.9% H_2O). After 5.5 minutes mobile B was increased to 8% over 0.5 minutes, to 30% over 27 minutes, further increased to 40% in 10 minutes, and finally up to 95% in 2 minutes and held for 4 minutes. Mobile B was reduced back to 5% over 1 minute for the rest of the acquisition.

Data was acquired in real time over 60 minutes using an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Scientific). Eluents were ionised using 2.1 kV spray voltage and a capillary temperature of 275 °C. For all samples MS1 spectra were measured with a resolution of 120,000 at 200 *m/z* and a mass range of 375 to 1500 *m/z*, with a 50 ms maximum injection time, and 4e5 AGC target.

High-resolution HCD MS2 spectra were generated in positive mode for precursor masses using a normalised collision energy of 38%, within a 0.7 *m/z* isolation window, using quadrupole isolation, in top speed mode over a 3 second cycle time, at 50,000 resolution in profile mode, 100 ms maximum injection time and 5e4 AGC target, with an instrument determined scan range beginning at 100 *m/z*.

To ensure quality peptide fragmentation, a number of filters were utilised, including; peptide mono isotopic precursor selection, minimum intensity exclusion of 5e4, and charge state inclusion of 2–5 + ions. To prevent repeat sampling a dynamic exclusion with exclusion count 1, exclusion duration of 30 seconds, +/−10 ppm tolerance window, with isotope exclusion and single charge state selection were used.

Mass-spectrometry data processing. All raw files were analysed using MaxQuant software (version 1.5.3.30) queried against the Uniprot complete human database (October 2017). The following search parameters were used: Reporter ion MS2 with multiplicity 6plex TMT for the TMTsixplex experiments and 10plex TMT for the TMT10plex experiments, trypsin digestion with maximum 2 missed cleavages, oxidation of methionine and acetylation of protein N-termini as variable modifications, carbamidomethylation of cysteine as a fixed modification, minimum peptide length of 6, protein FDR 0.01. Appropriate correction factors for the individual TMT channels for both peptide N-terminal labelling and lysine side-chain labelling as per the TMT kits used (Thermo Scientific) were configured into the integrated Andromeda search prior to the MaxQuant search.

PARP inhibitor treatment, dual metabolic incorporation and TMT10plex labeling. MDA-MB-231 cells grown to 80–90% confluency in 10 cm plates were treated with 25, 5, 1, 0.2 and 0 (DMSO control) µM concentrations of the PARP inhibitors Olaparib (10 mM stock solution in DMSO) and Rucaparib (10 mM stock solution in DMSO) in parallel. All treatments were performed in 10 ml DMEM with 10% FBS and the cells were placed in an incubator with 5% CO_2 and 37 °C for 1 h. After 1 h, the media were removed and the cells were supplemented with fresh complete media (5 ml) containing 0.25 mM concentrations of both 6YnAd (500 mM stock solution in DMSO) and 2YnAd (500 mM stock solution in DMSO) continued incubation for one more hour. The cells were washed with PBS three times and whole-cell lysates were prepared and proteins were quantified. 500 µg protein per each condition were subjected to click reactions using Azide-PEG3-Biotin capture reagent. Proteins were precipitated using the methanol/chloroform/water system and air-dried precipitates were resuspended in the pull-down buffer (50 mM HEPES, pH 7.4, 0.2% SDS, 150 mM NaCl) and subjected to affinity enrichment on NeutrAvidin-Agarose beads. After extensive washings, the beads were subjected to DTT treatment, IAA treatment and overnight trypsinisation. The peptide digests from the 10 conditions were collected, desalted, dried and re-dissolved in 100 µl of 100 mM triethylammonium bicarbonate (TEAB) buffer (pH 8.5). Each sample was subjected to TMT labelling with a unique TMT10plex reagent for 1 h at room temperature. After quenching the TMT reactions with hydroxylamine treatment, the 10 samples were mixed together, concentrated to complete dryness and re-dissolved in 0.1% formic acid solution in water. The sample was desalted on C18 Empore disks on micropipette tip format and the desalted sample was dried on a speedvac. The sample was dissolved in 300 µl of 0.1% TFA solution in water and subjected to fractionation using High pH Reverse-Phase Peptide Fractionation Kit (Thermo scientific) to 8 fractions and dried on a speedvac.

Pre-processing of MaxQuant output for data analysis. Columns corresponding to the corrected reporter intensity were extracted from the proteinGroups.txt files in MaxQuant output. Potential contaminants, proteins only identified by site, reverse proteins and proteins containing less than two peptides were discarded. Out of the remaining proteins, only proteins that have non-zero intensities across all reporters and replicates were kept (1604 proteins in 2YnAd-6YnAd dataset, 2334 in the inhibitor dataset). Intensities of the remaining proteins were normalised using TMM method[32] and converted to log2-per-million values. The resulting data was then compared with previously reported annotations of ADP-ribosylated proteins available in ADPriboDB (v1.0)[13], which was augmented to include annotations reported since 2015 (v2.0; unpublished), after prefiltering the database to contain only annotations derived from human datasets.

LIMMA linear model design. Both cofactor preference and inhibitor treatment datasets were analysed using LIMMA method. The linear methods for each of these datasets were designed separately and described below.

2YnAD/6YnAD. The normalised intensities are assumed to be modelled by the following effects in a format described below:

$$\begin{aligned}\log 2(\text{normalised intensity}) \sim\ & [\text{Cofactor effect, Sum encoding}] \\ & + [\log(\text{concentration}) \text{ effect}] \\ & + [\text{Replicate effect}] + \text{Intercept}.\end{aligned}$$

We used the following encoding for the effects.

1. **Cofactor effect** (i.e. the main measured effect). Encoded by +1 in reporters corresponding to 2YnAd and −1 in reporters corresponding to 6YnAd (e.g. sum-to-one encoding). Positive coefficients to this factor therefore indicate 2YnAd preference, and negative values indicate 6YnAd preference.
2. **log(concentration) effect**: encoded as log values of the concentrations in mM. Effect is assumed to be the same for both cofactors. Higher values indicate stronger response.
3. **Replicate**: 1-of-K encoding (minus one redundant term absorbed in intercept). This term accounts for technical-replicate associated bias in protein intensities.

The complete design matrix is available as Supplementary Table 1.

PARP Inhibitor dataset. For this dataset we have applied the linear modelling only on the data points corresponding to 25 µM of inhibitor and 0 µM (reporters 0, 4, 5 and 9), not taking into account the data for intermediate concentrations to avoid overfitting. The following linear model was assumed.

$$\begin{aligned}\text{Log(intensity)} \sim\ & [\text{Inhibitor preference effect}] + [\text{Inhibition effect}] \\ & + [\text{Replicate bias}] + [\text{Intercept}].\end{aligned}$$

The three factors have been encoded as follows:

1. **Inhibitor preference effect**. One of the two major effects in this experiment. This factor is designed to capture the effect of a particular inhibitor on the intensity of the protein. We encoded this using the following map: −1 = Olaparib, +1 = Rucaparib and 0 = No inhibitor (i.e. concentration = 0 µM). In this way, positive values correspond to protein having to higher intensity upon Rucaparib treatment and negative values corresponding to higher intensity upon Olaparib treatment.
2. **Inhibition effect**. This term corresponds to the second major effect in the experiment, which captures the general effect of inhibition to the protein that does not depend on the inhibitor. It was encoded as 1 (in cases where inhibitor is present) and 0 otherwise. This way values below than zero correspond to decrease in intensity upon inhibition.
3. **Replicate bias**: This term accounts for technical-replicate associated bias in protein intensities. This was encoded with one-in-K encoding with one redundant term absorbed to intercept.

The complete design matrix is available in Supplementary Table 2.

LIMMA processing and statistical testing. We have used a similar procedure to the one described by Kammers *et al.*[33]. The linear models were encoded as described above. LIMMA analysis was run with the trend parameter set to TRUE (commonly referred to as LIMMA-trend method). After the model has been fit, we have performed statistical testing on the effect coefficients in the model using the equations provided by Kammers *et al.* Specifically, we have tested the null hypothesis that cofactor effect coefficient is zero in 2YnAd/6YnAd dataset, and the null hypotheses that inhibitor preference effect and inhibition effect coefficients are not zero (separately) in inhibitor treatment dataset. The p values were adjusted with Benjamini-Hochberg method. Any adjusted p values below 0.01 were deemed significant and annotated in the plots and output. In order to generate debiased plots, we have subtracted the values for bias coefficients (e.g. replicate bias and intercept) from the normalised log2 intensities.

Pathway enrichment analysis. Biological pathway enrichment analysis was performed against Reactome database[34] using the R packages ReactomePA[35] and clusterProfiler[36] using the default settings, showing the most descriptive categories with the "by = rowPercentage" parameter set.

Data Availability

The tandem mass spectrometry datasets generated during the current study are available via the PRIDE database (project accession: PXD011631). All of the data analysed during this study are included in this published article (and its Supplementary Information files).

References

1. Hottiger, M. O. Nuclear ADP-Ribosylation and Its Role in Chromatin Plasticity, Cell Differentiation, and Epigenetics. *Annu. Rev. Biochem.* **84**, 227–263 (2015).
2. Lord, C. J. & Ashworth, A. PARP inhibitors: Synthetic lethality in the clinic. *Science* **355**, 1152–1158 (2017).
3. Jungmichel, S. *et al.* Proteome-wide identification of poly(ADP-Ribosyl)ation targets in different genotoxic stress responses. *Mol. Cell* **52**, 272–285 (2013).
4. Martello, R. *et al.* Proteome-wide identification of the endogenous ADP-ribosylome of mammalian cells and tissue. *Nat. Commun.* **7**, 12917 (2016).
5. Zhang, Y., Wang, J., Ding, M. & Yu, Y. Site-specific characterization of the Asp- and Glu-ADP-ribosylated proteome. *Nat. Methods* **10**, 981–984 (2013).
6. Gagne, J. P. *et al.* Proteome-wide identification of poly(ADP-ribose) binding proteins and poly(ADP-ribose)-associated protein complexes. *Nucleic Acids Res.* **36**, 6959–6976 (2008).
7. Daniels, C. M., Ong, S. E. & Leung, A. K. Phosphoproteomic approach to characterize protein mono- and poly(ADP-ribosyl)ation sites from cells. *J. Proteome Res.* **13**, 3510–3522 (2014).
8. Min, W., Cortes, U., Herceg, Z., Tong, W. M. & Wang, Z. Q. Deletion of the nuclear isoform of poly(ADP-ribose) glycohydrolase (PARG) reveals its function in DNA repair, genomic stability and tumorigenesis. *Carcinogenesis* **31**, 2058–2065 (2010).
9. Grammel, M., Hang, H. & Conrad, N. K. Chemical reporters for monitoring RNA synthesis and poly(A) tail dynamics. *ChemBioChem* **13**, 1112–1115 (2012).
10. Westcott, N. P., Fernandez, J. P., Molina, H. & Hang, H. C. Chemical proteomics reveals ADP-ribosylation of small GTPases during oxidative stress. *Nat. Chem. Biol.* **13**, 302–308 (2017).
11. Gibson, B. A. *et al.* Chemical genetic discovery of PARP targets reveals a role for PARP-1 in transcription elongation. *Science* **353**, 45–50 (2016).
12. Dayon, L. *et al.* Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags. *Anal. Chem.* **80**, 2921–2931 (2008).
13. Vivelo, C. A., Wat, R., Agrawal, C., Tee, H. Y. & Leung, A. K. ADPriboDB: The database of ADP-ribosylated proteins. *Nucleic Acids Res.* **45**, D204–D209 (2017).
14. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
15. Knezevic, C. E. *et al.* Proteome-wide Profiling of Clinical PARP Inhibitors Reveals Compound-Specific Secondary Targets. *Cell Chem. Biol.* **23**, 1490–1503 (2016).
16. Jiang, H., Kim, J. H., Frizzell, K. M., Kraus, W. L. & Lin, H. Clickable NAD analogues for labeling substrate proteins of poly(ADP-ribose) polymerases. *J. Am. Chem. Soc.* **132**, 9363–9372 (2010).
17. Wang, Y., Rösner, D., Grzywa, M. & Marx, A. Chain-terminating and clickable NAD(+) analogues for labeling the target proteins of ADP-ribosyltransferases. *Angew. Chem. Int. Ed.* **53**, 8159–8162 (2014).
18. Carter-O'Connell, I., Jin, H., Morgan, R. K., David, L. L. & Cohen, M. S. Engineering the substrate specificity of ADP-ribosyltransferases for identifying direct protein targets. *J. Am. Chem. Soc.* **136**, 5201–5204 (2014).
19. Wallrodt, S., Buntz, A., Wang, Y., Zumbusch, A. & Marx, A. Bioorthogonally Functionalized NAD(+) Analogues for In-Cell Visualization of Poly(ADP-Ribose) Formation. *Angew. Chem. Int. Ed.* **55**, 7660–7664 (2016).
20. Buntz, A. *et al.* Real-Time Cellular Imaging of Protein Poly(ADP-ribos)ylation. *Angew. Chem. Int. Ed.* **55**, 11256–11260 (2016).
21. Curanovic, D. *et al.* Global profiling of stimulus-induced polyadenylation in cells using a poly(A) trap. *Nat. Chem. Biol.* **9**, 671–673 (2013).
22. Cho-Park, P. F. & Steller, H. Proteasome regulation by ADP-ribosylation. *Cell* **153**, 614–627 (2013).
23. Kakarougkas, A. *et al.* Co-operation of BRCA1 and POH1 relieves the barriers posed by 53BP1 and RAP80 to resection. *Nucleic Acids Res.* **41**, 10298–10311 (2013).
24. Haikarainen, T., Krauss, S. & Lehtio, L. Tankyrases: structure, function and therapeutic implications in cancer. *Curr. Pharm. Des.* **20**, 6472–6488 (2014).
25. Matveeva, E. *et al.* Involvement of PARP1 in the regulation of alternative splicing. *Cell Discov* **2**, 15046 (2016).
26. Jubin, T. *et al.* Poly ADP-ribose polymerase-1: Beyond transcription and towards differentiation. *Semin. Cell Dev. Biol.* **63**, 167–179 (2017).
27. Ji, Y. & Tulin, A. V. Post-transcriptional regulation by poly(ADP-ribosyl)ation of the RNA-binding proteins. *Int. J. Mol. Sci.* **14**, 16168–16183 (2013).
28. Ji, Y. & Tulin, A. V. Poly(ADP-ribosyl)ation of heterogeneous nuclear ribonucleoproteins modulates splicing. *Nucleic Acids Res.* **37**, 3501–3513 (2009).
29. Melikishvili, M., Chariker, J. H., Rouchka, E. C. & Fondufe-Mittendorf, Y. N. Transcriptome-wide identification of the RNA-binding landscape of the chromatin-associated protein PARP1 reveals functions in RNA biogenesis. *Cell Discov.* **3**, 17043 (2017).
30. Zimmermann, M. *et al.* CRISPR screens identify genomic ribonucleotides as a source of PARP-trapping lesions. *Nature* **559**, 285–289 (2018).
31. Thorsell, A.-G. *et al.* Structural Basis for Potency and Promiscuity in Poly(ADP-ribose) Polymerase (PARP) and Tankyrase Inhibitors. *J Med Chem* **60**, 1262–1271 (2017).
32. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
33. Kammers, K., Cole, R. N., Tiengwe, C. & Ruczinski, I. Detecting Significant Changes in Protein Abundance. *EuPA Open Proteom.* **7**, 11–19 (2015).
34. Fabregat, A. *et al.* The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **46**, D649–D655 (2018).
35. Yu, G. & He, Q. Y. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol. BioSyst.* **12**, 477–479 (2016).
36. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* **16**, 284–287 (2012).

Acknowledgements

This work was supported by GlaxoSmithKline through the Imperial College London Engineering Medicines Laboratory Project (K.A.K, D.O.H and P.A.D) and the UK Biotechnology and Biological Sciences Research Council Doctoral Training Partnership award (S.L). A.P.S. and A.B. are supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001999), the UK Medical Research Council (FC001999) and the Wellcome Trust (FC001999).

Author Contributions

K.K. and P.A.D. conceived of the study; K.K. performed experiments; S.L. and P.A.D. performed statistical analysis of data; K.K., S.L., A.J.B., A.P.S., D.O.H., V.A., A.K.L.L. and P.A.D. contributed to experimental design, data analysis and data interpretation. K.K. and P.A.D. wrote the manuscript with input from all authors.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-43154-1.

Competing Interests: K.K., S.L. and P.A.D. have a patent application pending for the utilisation of the dual metabolic labelling using 2YnAd and 6YnAd for identification of intracellular ADP-ribosylation (Application number: 1810697.1). A.J.B., A.P.S., D.O.H., V.A. and A.K.L.L. declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019