SCISPACE
formerly Typeset

Open access • Posted Content • DOI:10.1101/2020.01.27.921650

Mapping physiological ADP-ribosylation using Activated Ion Electron Transfer Dissociation (AI-ETD) — Source link

Sara C. Buch-Larsen, Ivo A. Hendriks, Jean M. Lodge, Martin Rykær ...+5 more authors

Institutions: University of Copenhagen, University of Wisconsin-Madison

Published on: 28 Jan 2020 - bioRxiv (Cold Spring Harbor Laboratory)

Topics: Electron-transfer dissociation

Related papers:

- Mapping Physiological ADP-Ribosylation Using Activated Ion Electron Transfer Dissociation.
- Chemical Tools to Study Protein ADP-Ribosylation.
- Generation and Characterization of Recombinant Antibody-like ADP-Ribose Binding Proteins
- A General Approach Towards Triazole-Linked Adenosine Diphosphate Ribosylated Peptides and Proteins
- Mass spectrometry for serine ADP-ribosylation? Think o-glycosylation!

Share this paper:

View more about this paper here: https://typeset.io/papers/mapping-physiological-adp-ribosylation-using-activated-ion-567p4xsu8i

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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.

Mapping physiological ADP-ribosylation using Activated Ion Electron Transfer Dissociation (AI-ETD)

Sara C. Buch-Larsen[1,3], Ivo A. Hendriks[1,3], Jean M. Lodge[2], Martin Rykær[1], Benjamin Furtwängler[1], Evgenia Shishkova[2], Michael S. Westphall[2], Joshua J. Coon[2], and Michael L. Nielsen[1,*]

[1]Proteomics program, Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, 2200 Copenhagen, Denmark

[2]Genome Center of Wisconsin, University of Wisconsin-Madison, Madison, WI, 53706, USA

[3]These authors contributed equally

*Correspondence to: michael.lund.nielsen@cpr.ku.dk

Running title: Profiling the physiological ADP-ribosylome

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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.

SUMMARY (max. 150 words)

ADP-ribosylation (ADPr) is a post-translational modification that plays pivotal roles in a wide range of cellular processes. Mass spectrometry (MS)-based analysis of ADPr under physiological conditions, without relying on genetic or chemical perturbation, has been hindered by technical limitations. Here, we describe the applicability of Activated Ion Electron Transfer Dissociation (AI-ETD) for MS-based proteomics analysis of physiological ADPr using our unbiased Af1521 enrichment strategy. To benchmark AI-ETD, we profiled 9,000 ADPr peptides mapping to >5,000 unique ADPr sites from a limited number of cells exposed to oxidative stress, corresponding to 120% and 28% more ADPr peptides compared to contemporary strategies using ETD and EThcD, respectively. Under physiological conditions AI-ETD identified 450 ADPr sites on low-abundant proteins, including *in vivo* cysteine auto-modifications on PARP8 and tyrosine auto-modifications on PARP14, hinting at specialist enzymatic functions for these enzymes. Collectively, our data provides new insights into the physiological regulation of ADP-ribosylation.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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

ADP-ribosylation (ADPr) is an emerging post-translational modification (PTM) regulating a variety of biological processes, including cell signaling and DNA damage repair, by modifying proteins with either one ADPr moiety (mono-ADP-ribosylation; MARylation) or several ADPr moieties (poly-ADP-ribosylation; PARylation). ADP-ribosyltransferases (ARTs) catalyze the modification by transferring ADPr units from NAD+ onto target proteins. The group of responsible enzymes can be divided into two groups depending on the conserved structural features; ARTCs (cholera toxin-like) and ARTDs (diphtheria toxin-like, better known as poly(ADP-ribosyl)polymerases, PARPs), of which the latter is the largest and most characterized group consisting of 17 members in human (Lüscher et al., 2017). Within the ARTDs, four and twelve enzymes are reported to possess PARylation and MARylation activity, respectively, while PARP13 is reported to be catalytic inactive (Lüscher et al., 2017).

The most thoroughly studied PARP enzyme is PARP1, which is predominantly investigated in relation to DNA damage repair (De Vos et al., 2012). Under genotoxic stress, PARP1 mainly modifies serine residues in conjunction with the co-factor Histone PARylation factor 1 (HPF1) (Bonfiglio et al., 2017b), with this induction abrogated by PARP inhibitors (Larsen et al., 2018). ADPr is a reversible PTM, allowing it to rapidly and situationally adapt to cellular stimuli in a timely manner. Two groups of erasing enzymes exist; the macrodomain fold and the dinitrogenase reductase-activating glycohydrolase (DraG) (Crawford et al., 2018). Poly(ADP-ribose) glycohydrolase (PARG) is one of the erasers containing the macrodomain fold, and is the key enzyme accountable for degrading PAR chains into MAR (D'Amours et al., 1999). Recently, ARH3 was reported to be a serine hydrolase belonging to the DraG group of enzymes (Crawford et al., 2018; Fontana et al., 2017).

In recent years, mass spectrometry (MS)-based proteomics methods have advanced as the primary approach for studying ADPr in a global and unbiased manner. As is the case with many PTMs, ADPr is very low abundant, making enrichment methods necessary. MS-based analysis is further complicated by the dynamic nature of ADPr, with a rapid enzymatic turnover of the modification (Alvarez-Gonzalez and Althaus, 1989), numerous modifiable amino acid residues (Altmeyer et al., 2009; Bonfiglio et al., 2017b; McDonald and Moss, 1994; Moss and Vaughan, 1977; Ogata et al., 1980; Sekine et al., 1989; Van Ness et al., 1980), and the highly labile nature of the modification (Bonfiglio et al., 2017a).

Several methods have been established for enrichment of ADPr-modified peptides for mass spectrometric analysis (Daniels et al., 2014; Gagne et al., 2018; Hendriks et al., 2019; Larsen et al., 2018; Zhang et al., 2013). For large-scale studies of ADPr, the two most widely used strategies are based either on affinity enrichment using either the Af1521 macrodomain, or on a combination of boronate-affinity enrichment coupled with chemical conversion of the ADPr modification into an amide

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

using hydroxylamine. Whereas the Af1521-based enrichment strategy identifies the entire ADPr moiety on any amino acid residue (Larsen et al., 2018), the chemical-based strategy is only able to identify chemical marks on aspartic acid and glutamic acid residues (Li et al., 2019; Zhang et al., 2013). Moreover, while the Af1521 macrodomain is capable of identifying ADPr under endogenous conditions, knockdown of PARG is required to increase sensitivity for the chemical-based method, rendering investigation of ADPr under physiological conditions impossible (Li et al., 2019; Zhang et al., 2013).

We have previously demonstrated that MS acquisition relying on higher-energy collisional dissociation (HCD)-based fragmentation cannot be used for confidently localizing ADPr sites due to the labile nature of the bond between the modification and the amino acid residue (Larsen et al., 2018). Hence, the non-ergodic fragmentation propensity of electron transfer dissociation (ETD) fragmentation is necessary in order to confidently determine the exact amino acid residue being modified with ADPr. However, ETD is prone to non-dissociative electron transfer (ETnoD), resulting in fewer sequence-informative product ions (Ledvina et al., 2010), which we previously observed to especially affect ADPr-modified precursors (Larsen et al., 2018). To overcome the high degree of ETnoD, supplemental activation can be applied with electron-transfer higher-energy collisional dissociation (EThcD), which has been proven to be more effective than ETD for identification of ADPr sites (Bilan et al., 2017; Hendriks et al., 2019). However, EThcD did not have a major effect on the dissociation of charge-reduced ADPr peptide product ions (Hendriks et al., 2019).

By contrast, activated ion-ETD (AI-ETD) uses infrared photoactivation during the ETD reaction to overcome ETnoD, and has been shown to benefit glycoproteomics (Riley et al., 2019), top-down proteomics (McCool et al., 2019; Riley et al., 2018a; Riley et al., 2018b), as well as phosphoproteomics (Riley et al., 2017a). We therefore set out to investigate the utility of AI-ETD on an Orbitrap Fusion Lumos mass spectrometer for identification and confident localization of ADPr sites. Here, we explored different laser power settings for AI-ETD, and compared it to ETD and EThcD. Overall, AI-ETD proved very advantageous for analyzing ADPr-enriched samples, with AI-ETD mapping 79% and 23% more ADPr sites compared to ETD and EThcD, respectively. With the enhanced performance of AI-ETD, we could obtain a comparable depth of sequencing from markedly less input material. This increased sensitivity auspiciously enabled identification of 450 physiological ADPr sites, of which 60% were not previously identified under physiological conditions. Taken together, we demonstrate that AI-ETD is a superior dissociation technique for sensitive MS-based proteomics analysis of ADP-ribosylation, and facilitates analysis of physiological ADPr from limited amounts of starting material.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

RESULTS

Evaluation of AI-ETD

In order to investigate the ability of AI-ETD to increase the dissociation of ADP-ribosylated peptides, we utilized our established Af1521 method that allows for efficient and high-purity enrichment of peptides modified by ADP-ribosylation (Figure 1A). As the binding affinity of the macrodomain is towards the terminal moiety of ADP-ribose (Karras et al., 2005), the Af1521 strategy is capable of detecting modification of any amino acid residue type by ADP-ribosylation (Larsen et al., 2018). For the purpose of initial optimization and direct comparison of AI-ETD with other dissociation methodologies, we purified ADP-ribosylated peptides from H_2O_2-treated HeLa cells. This sample was used as a technical standard for quadruplicate comparison of different dissociation types, with each experiment corresponding to a comparatively low amount of ADP-ribosylated peptides purified from only two million cells. We compared standard ETD to EThcD, which we previously demonstrated to be superior for the identification of ADP-ribosylated peptides (Hendriks et al., 2019). Additionally, we used AI-ETD at five different laser power settings, spanning the range of routinely used power settings (Ledvina et al., 2010; Riley et al., 2017b).

We observed a higher number of identified and localized unique modified ADP-ribosylated peptides for all five AI-ETD laser power settings, compared to pure ETD (Figure 1B and Table S1). The optimum energy setting was 20%, with a gradual fall-off in identification rate for both lower and higher laser power settings. As we described previously, EThcD outperformed pure ETD, but overall AI-ETD performed significantly better. AI-ETD identified 120% more ADPr-modified peptides than ETD, and 28% more ADPr-modified peptides than EThcD. When considering the number of unique ADP-ribosylation sites on proteins, we found the same trend (Figure 1C and Table S2), with AI-ETD mapping 79% and 23% more sites compared to ETD and EThcD, respectively.

In terms of spectral quality, as assessed by Andromeda Score, AI-ETD demonstrated the highest scores when using relatively low laser power settings, with an optimum at 15% (Figure 1D). Higher laser power settings of 25% and above, as well as EThcD, resulted in a decline in spectral quality, suggesting excessive peptide fragmentation under these circumstances. When considering localization of the ADP-ribose to specific amino acid residues within the peptide, we found that on average AI-ETD at laser powers 15% and 20%, as well as EThcD, significantly outperformed pure ETD (Figure 1E). However, when filtering ADPr for a stringent localization cut-off of 90%, AI-ETD at 20% laser power was superior.

Previously, we demonstrated complementarity of trypsin and Lys-C digestions for identification of unique ADPr sites (Hendriks et al., 2019). Thus, we purified ADPr peptides as described above, but using Lys-C instead of trypsin digestion. We subjected these peptides to technical quadruplicate

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

analysis with a similar experimental setup, using the three optimal AI-ETD laser power settings. Overall, the aforementioned findings could be recapitulated, with the best performance of AI-ETD at laser power 20%, identifying 134% and 30% more ADPr peptides than ETD and EThcD, respectively (Figure S1A and Table S1). At the site-specific level, AI-ETD mapped 118% and 28% more ADPr sites than ETD and EThcD, respectively (Figure S1B and Table S2), and spectral quality was likewise improved by AI-ETD (Figure S1C).

Based on these observations, we wondered whether the increased performance of AI-ETD could be explained through increased dissociation of reduced charge-state precursors, i.e. decreased levels of ETnoD, and an overall more complete fragmentation of the ADP-ribosylated peptides. Intriguingly, AI-ETD considerably reduced the amount of ETnoD even at lower laser power settings, and dramatically reduced ETnoD to practically zero at the highest laser power settings (Figure 1F and S1D), whereas EThcD failed to resolve ETnoD product ions to the same extent. Concomitantly, the unmodified peptides within the same samples were only modestly affected by AI-ETD, as these were generally short peptides that did not display high levels of ETnoD to start with. Moreover, this highlights a striking susceptibility of ADPr peptides to AI-ETD. Correspondingly, when considering the overall level of precursor fragmentation, we found a notably increased fragmentation of ADPr peptides with increased levels of AI-ETD laser power, whereas unmodified peptides were not affected as much (Figure 1G and S1E). Contrarily, EThcD did not greatly increase fragmentation of ADPr peptides, but did affect unmodified peptides. Overall, AI-ETD reduced ETnoD and led to more complete fragmentation of specifically ADPr peptides.

ADPr peptides achieved a high level of fragmentation at the highest AI-ETD laser power settings, yet this did not result in the highest number of identifications. To investigate this, we compared the degree of peptide fragmentation to spectral score. For unmodified peptides, there was a global correlation with higher degrees of fragmentation corresponding to higher spectral quality (Figure S1F). For localized ADPr peptides, we found the highest quality spectra to frequently reside in the range from 30% to 45% fragmentation, and noted a strong decline in spectral quality when fragmentation exceeded 50% (Figure 1H). Indeed, when considering identified but non-localized ADPr peptides, a considerably higher number of these were >50% fragmented (Figure S1G), suggesting that very high degrees (>50%) of ADPr peptide fragmentation are detrimental to faithful identification and localization of this PTM. Collectively, we show that AI-ETD is superior to ETD and EThcD for confident profiling of ADPr sites.

Benchmarking AI-ETD for in-depth profiling of ADPr

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

To gauge the performance of AI-ETD in the context of deep analysis of the ADP-ribosylome, we employed our previously described high-pH fractionation on ADPr-modified peptides purified from H_2O_2-treated HeLa cells (Figure S2A) (Hendriks et al., 2019; Larsen et al., 2018). Since we established that AI-ETD is more effective than ETD and EThcD, we used AI-ETD at the optimal laser power setting of 20% for all further experiments. Additionally, to address the large amount of starting material usually required for proteomic profiling of ADPr (Abplanalp et al., 2017; Hendriks et al., 2019; Larsen et al., 2018; Li et al., 2019; Zhang et al., 2013), we performed mass spectrometric analysis on a considerably smaller amount of sample. In total, 8,939 unique ADPr-modified peptides were identified (Table S1), mapping to 5,206 confidently localized ADPr sites (Figure 2A and Table S2). An intermediate total number of ADPr sites was identified compared to previous studies, falling just short of the current single-study record of 7,000 (Hendriks et al., 2019). Nevertheless, the relative amount of input material analyzed using AI-ETD was only ~20%, yet still facilitated identification of ~75% of the total number of sites (Figure 2A). Additionally, our AI-ETD profiling of the ADP-ribosylome from limited input revealed ADPr modification of 1,800 proteins (Table S3), corresponding to 84% of the number of proteins identified in the largest single MS study to date (Hendriks et al., 2019). More than half of the ADPr target proteins were modified on two or more distinct ADPr sites, demonstrating deep coverage of this low stoichiometry PTM (Figure S2A). Collectively, our benchmarking data exemplify the great sensitivity of AI-ETD.

Overall, 1,798 (35%) and 2,045 (39%) sites were previously described by either or both our other studies, respectively (Figure 2B). Moreover, we identified 1,363 (26%) previously undescribed ADPr sites (Table S4). In terms of amino acid residue specificity, we observed the large majority of modification to occur on serine residues (83%), followed by modification of arginine (5.9%), histidine (4.5%), and tyrosine (2.8%) residues (Figure 2C, left). When only considering the subset of ADPr sites exclusively identified in this study, we still found the majority of these to reside on serine residues (58%), but with a relatively larger fraction of ADPr also targeting arginine (17%), histidine (8.7%), tyrosine (4.8%), lysine (4.7%), cysteine (3.2%), threonine (2.1%), aspartic acid (1.1%), and glutamic acid (0.7%) residues (Figure 2C, right).

The H_2O_2-induced ADP-ribosylome profiled using AI-ETD largely resided on proteins with functional annotations that are well-known in the context of ADP-ribosylation biology, including predominant modification of nuclear, nucleolar, and chromatin-centric proteins, and modification of proteins involved in RNA binding, DNA repair, chromatin regulation, and DNA replication (Figure 2D and Table S5). Taken together, AI-ETD is congruous with deep analysis of the ADP-ribosylome from a limited pool of starting material.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

Improved analysis of physiological ADPr using AI-ETD

ADP-ribosylation is frequently investigated in a non-physiological manner, e.g. by genetic perturbation of ADP eraser enzymes (Daniels et al., 2014; Zhang et al., 2013), or by exposure of cells to high levels of genotoxic or proteotoxic stress, which both serve to boost cellular levels of ADPr in order to facilitate detection. Levels of ADPr are much lower under physiological conditions, e.g. in unperturbed cells growing under standard conditions. This makes physiological ADPr technically challenging to study, despite the biomedical importance of elucidating how ADPr is regulated at baseline conditions, e.g. in clinical samples or patient biopsies. Therefore, we set out to evaluate the ability of AI-ETD to profile the physiological ADP-ribosylome.

Using our established strategy, we purified ADPr-modified peptides from unperturbed HeLa cells growing under standard conditions. To minimize the requirement for mass spectrometric machine time we did not perform high-pH fractionation, and analyzed the physiological ADPr samples as single shots using AI-ETD. In total, we identified 703 unique ADPr-modified peptides, corresponding to 450 ADPr sites, residing on 295 proteins (Figure 3A). Using AI-ETD, we identified 79% and 26% more ADPr sites compared to single shot and high-pH fractionated measurements from our previous study (Larsen et al., 2018), respectively. When considering the best replicate, we were able to map 387 sites from a single run spanning 50 minutes of analytical gradient time, while on average AI-ETD allowed us to identify and confidently localize up to 8 physiological ADPr sites per minute of machine time.

Out of the 450 ADPr physiological sites we identified, 286 (64%) were not previously mapped under physiological conditions (Figure 3B, right), and 70 were not found in proteomics screens under any previously investigated conditions (Table S4). Intriguingly, serine residues (55%) remained the most commonly ADP-ribosylated under physiological conditions (Figure 3B, left). Nonetheless, we also observed frequent modification of arginine (26%), histidine (12%), and cysteine (3.3%) residues. Physiological ADPr could be confidently localized (>90%) on cysteine, histidine, arginine, serine, threonine, and tyrosine residues (Figure 3C), while we did not observe consistent ADP-ribosylation of lysine, glutamate, and aspartate residues. We observed that under physiological conditions, >50% of serine ADPr resided in KS motifs (Figure S3A); an established phenomenon for serine ADPr (Larsen et al., 2018). Adherence to KS motifs was considerably lower in response to oxidative stress (Figure S3A), suggesting that under these conditions serine residues may be targeted more promiscuously. Physiological ADPr predominantly targeted nuclear proteins (71%; Figure S3B), although not to the same extent as H_2O_2-induced ADPr (85%; Figure S3C).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

Next, we investigated the expression levels of proteins which have previously been observed to be ADP-ribosylated (Figure 3D). With the routine use of H_2O_2 induction of ADP-ribosylation, which greatly boosts the amount of ADPr in cells, we found that protein copy-numbers of ADPr target proteins span six orders of magnitude. We observed that physiologically ADP-ribosylated target proteins covered a comparable range of expression levels, although overall the depth of sequencing was somewhat lower. Nonetheless, compared to physiological ADPr target proteins identified in our previous study (Larsen et al., 2018), we achieved a 57% greater depth of sequencing (Figure 3D).

Overall, the extent of ADP-ribosylation was vastly different between unperturbed cells and H_2O_2-treated cells. Combining previous proteomics-based knowledge on ADPr (Bilan et al., 2017; Bonfiglio et al., 2017b; Hendriks et al., 2019; Larsen et al., 2018), with the additional insight gained in this study, we set out to investigate systemic differences between physiological and H_2O_2-perturbed ADPr biology. Specifically, we overlapped 9,029 ADPr sites mapped in response to H_2O_2 with 643 ADPr sites detected under physiological conditions (Figure 3E). The vast majority of ADPr sites (93%) were exclusive to H2O2 treatment, with 6% of ADPr sites detected in both cases, and 1% of ADPr sites exclusively detected under physiological conditions. In terms of amino acid distribution, the H_2O_2-exclusive ADPr predominantly modified serine residues (84%), in agreement with previous reports (Hendriks et al., 2019; Larsen et al., 2018; Palazzo et al., 2018). ADPr detected under all conditions still targeted a majority of serine residues (62%), but moreover modified significant fractions of arginine (13%), histidine (13%), cysteine (5%), and lysine (4%) residues. Intriguingly, physiologically exclusive ADPr primarily modified arginine residues (66%), with further modification of histidine (15%) and cysteine (10%) residues, and only a small fraction of modification on serine residues (5.7%). Still, our data revealed a propensity for ADPr to be targeted to different amino acid residues dependent on cellular condition.

Next, we performed functional annotation enrichment analysis to assess relative differences between physiological and non-physiological ADPr (Figure 3E and Table S5). H_2O_2-specific ADPr was more frequently observed to target proteins involved in rRNA-binding, DNA replication, and cell cycle, and less frequently modified glycosyltransferases, nucleosome proteins, and membrane proteins. Physiological-specific ADPr preferentially targeted proteins with sulfotransferase activity, glycoproteins, and membrane proteins, and averted targeting nuclear proteins involved in transcription, cell cycle, and the DNA damage response. Finally, proteins ADP-ribosylated under all conditions represented terms canonically associated with ADP-ribosylation, including predominant modification of the ADP enzymatic machinery, histone proteins, and DNA repair proteins, and this subset of ADPr target proteins less frequently included transcription factors, chromatin regulators, and proteins involved in cell cycle regulation. Collectively, our comparison between physiological and H_2O_2-induced ADPr demonstrates a

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

considerable shift in the targeting of this PTM, not only at the substrate level, but also in terms of which amino acid residue types are modified.

Detection of PARP8 and PARP14 auto-modification in cells

It is widely acknowledged that PARP1 modifies itself in response to DNA binding (Bolderson et al., 2019), and we profiled the main auto-modification of PARP1 to target three serine residues (Larsen et al., 2018). Our physiological profiling of ADP-ribosylation using AI-ETD corroborated that the same three serine residues are also modified under standard cellular growth conditions, with the majority (99.4%) of modification targeting Serine-499, followed by Serine-507 and Serine-519 (Figure 4A), with further physiological ADPr sites on PARP1 predominantly on serine residues.

With the improved analytical sensitivity gained by AI-ETD, we were curious whether we could detect ADPr modification of other PARP enzymes, under physiological conditions or otherwise in response to H_2O_2 treatment. Based on deep proteomic profiling of HeLa cells, ten PARP family members are expressed at different levels in unperturbed HeLa cells (Bekker-Jensen et al., 2017). Using AI-ETD, we could detect ADPr modification of PARP1, PARP2, PARP8, PARP12, and PARP14 (Figure 4B and Tables S3 and S4). Strikingly, even though protein copy-numbers of PARP8 and PARP14 are 250 to 500 times lower than PARP1, we could still detect their ADP-ribosylation under physiological conditions.

On PARP2, we only observed modification of serine residues, similar to PARP1 (Figure 4A), corroborating the known inter-connected functions of the two enzymes (Menissier de Murcia et al., 2003; Schreiber et al., 2002). ADP-ribosylation on PARP12 was only found in response to H_2O_2 treatment, and at relatively low abundance. Intriguingly, PARP8 was ADP-ribosylated on three cysteine residues under physiological conditions, Cysteine-376, Cysteine-395, and Cysteine-332, whereas a fourth cysteine residue, Cysteine C-367, was modified in response to H_2O_2. This exclusive modification on cysteine residues could hint at PARP8 auto-modification, which may imply that PARP8 is capable of ADP-ribosylation of cysteine residues. Analogously, we observed physiological modification of three tyrosine residues in PARP14, Tyosine-895, Tyrosine-1088, and Tyrosine-59, which could likewise imply auto-modification, and a potential function of PARP14 as the catalyst for tyrosine ADP-ribosylation.

ADP-ribosylation of serine, arginine, tyrosine, and histidine residues, has previously been reported (Hendriks et al., 2019). We were intrigued by the ADPr modification of cysteine residues, with PARP8 modified on four cysteine residues, and knowledge regarding cysteine ADPr being relatively sparse. Altogether, we could map 123 cysteine ADPr sites to 112 proteins (Tables S3 and S4), with only one cysteine ADPr site detected for the majority of proteins. We generated an interconnected STRING network based on all proteins with cysteine ADPr (Figure 4C), with modification occurring under

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

physiological conditions or in response to oxidative stress. 85 out of 112 proteins (76%) were connected to the network, with significant enrichment for protein-protein interactions. Beyond PARP8, cysteine ADPr modification occurred on a considerable number of interconnected catalytic ADPr enzymes, including PARP1, PARP5A (TNKS), PARP7 (TIPARP), PARP12, PARP13 (ZC3HAV1), and PARP14. When only considering the 60 proteins predominantly ADP-ribosylated on cysteines, there was a significant presence of proteins involved in RNA metabolism, including DDX6, HNRNPH1, NSUN2, PAIP1, PARN, PCBP1, PPP2R1A, PSMB2, TNKS1BP1, and TRNT1. Taken together, our sensitive MS-based proteomics screening of ADPr provides further insight into modification of non-serine residue types, and may hint at enzymatic specificity of certain PARP enzymes towards specific amino acids.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

DISCUSSION

Here, we investigated the potential of AI-ETD for sensitive identification and confident localization of ADPr sites. To evaluate the ability of AI-ETD for profiling the ADP-ribosylome, we performed a comparison of pure ETD, EThcD, and AI-ETD at five different laser power settings.

ADP-ribosylated peptides are very labile when subjecting them to higher-energy collisional dissociation (HCD), with the ADP-ribose itself drawing much of the energy and being one of the first fragmentation events, ultimately prohibiting localization of the ADP-ribose to the correct amino acid, as no fragment ions remain with the ADP-ribose attached (Bonfiglio et al., 2017a; Larsen et al., 2018). Contrarily, ADP-ribosylated peptides are very stable when subjected to ETD, and here we observed ADPr peptides to be much more prone to non-dissociative electron transfer (ETnoD), compared to unmodified peptides within the same samples. Supplemental activation of the ETD reaction with HCD, i.e. EThcD, still largely results in an ETD-type fragmentation, while somewhat mitigating ETnoD and generating more comprehensive peptide bond coverage (Bilan et al., 2017; Hendriks et al., 2019). However, EThcD fails to completely overcome ETnoD, and results in decreased spectral quality at higher energies. Intriguingly, we observed that ADPr peptides are highly susceptible to AI-ETD, much more so than small unmodified peptides within the same sample. This susceptibility to AI-ETD resulted in a near-complete abolishing of ETnoD at laser power settings of 25% and higher, although these greater energies also resulted in over-fragmentation of the peptides and a sharp drop in spectral quality. At a modest level of energy, AI-ETD resulted in higher quality spectra, improved sensitivity, and more confident localization of ADP-ribosylation to acceptor amino acids, overall doubling the number of identifications compared to pure ETD.

Recent advances in mass spectrometry-based proteomics have largely focused on high-throughput analyses, essentially by minimizing time or input material required to achieve the same depth of sequencing (Bache et al., 2018; Batth et al., 2019). Compared to relatively abundant post-translational modifications such as phosphorylation and glycosylation, for which MS methodology is well-established (Hogrebe et al., 2018; Kelstrup et al., 2018; Riley et al., 2019), the study of ADP-ribosylation with MS-based proteomics remains daunting, requiring large amounts of input material and significant investments of mass spectrometric machine time (Hendriks et al., 2019; Larsen et al., 2018; Leutert et al., 2018; Li et al., 2019; Martello et al., 2016; Zhang et al., 2013). Using AI-ETD, we were able to identify >500 ADPr sites from single shot analyses of purified ADPr material originating from only ~5 million cells. Moreover, from fractionated samples originating from ~100 million cells, we were able to identify >5,000 ADPr sites, reaching a great depth of sequencing while using considerably less starting material

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

than other studies (Hendriks et al., 2019; Larsen et al., 2018; Leutert et al., 2018; Li et al., 2019; Martello et al., 2016; Zhang et al., 2013).

Compared to ADPr induced by oxidative stress, the identification of ADPr in unperturbed cells is exceptionally challenging, with the modification occurring at very low abundance and dynamically regulated (Daniels et al., 2015). We previously reported a few hundred ADPr sites under physiological conditions, without the need for genetic perturbation of eraser enzymes, and without exposing cells to high levels of genotoxic stress (Larsen et al., 2018). However, this previous identification of physiological ADPr sites required an extensive investment of mass spectrometric machine time and a larger amount of starting material, with a considerable portion of identified ADPr sites relying on pre-fractionation of samples prior to mass spectrometric analysis. With the increased analytical sensitivity of AI-ETD, we were now able to map 450 physiological ADPr sites from single shot analyses, while using limited starting material and a total of 200 minutes of analytical gradient time for four replicates. Taken together, we demonstrate that AI-ETD can assist the study of physiological ADPr under unperturbed conditions, ultimately reducing the amount of starting material required for sequencing the native ADP-ribosylome, and facilitating studying ADP-ribosylation in a biomedical context.

Auto-modification of PARPs is a known phenomenon (Vyas et al., 2014), with especially auto-modification of PARP1 being thoroughly studied (Chapman et al., 2013; Muthurajan et al., 2014). Here, we were able to identify ADPr occurring on PARP1, 2, 8, 12, and 14. For PARP1 and PARP2, serine was the main acceptor residue. This is consistent with PARP1 and PARP2 catalyzing ADPr of serine residues upon oxidative stress, in combination with HPF1 (Bonfiglio et al., 2017b; Larsen et al., 2018; Palazzo et al., 2018). While we observed a relatively high number of serine residues being ADP-ribosylated under physiological conditions, many of those were also detectable in response to oxidative stress. Contrarily, ADPr sites exclusively detected under physiological conditions were only rarely modifying serine residues. Taken together, this would suggest that serine ADP-ribosylation occurs predominantly in response to DNA damage, and that serine ADP-ribosylation observed under physiological conditions could be a result of baseline DNA damage, potentially as a consequence of culturing cells under atmospheric oxygen conditions.

Intriguingly, we only observed PARP8 to be ADP-ribosylated on cysteine residues. This exclusive modification of cysteine residues supports previous *in vitro* observations of the auto-modification of PARP8 (Vyas et al., 2014), and we confirmed all four hitherto reported cysteine residues, of which three were detectable under physiological conditions. In case of PARP14, physiological ADP-ribosylation only targeted tyrosine residues, similarly setting it apart from PARP1 and PARP2 which were primarily modified on serine residues. Auto-modification of PARP14 has been described previously

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

(Qin et al., 2019), and the interactome of PARP14 suggests a role in regulating RNA stability (Carter-O'Connell et al., 2018). We detected modification of PARP13 in cells on one cysteine residue, and PARP14 has been demonstrated to MARylate PARP13 *in vitro* (Carter-O'Connell et al., 2018), although there acidic amino acids were targeted for modification. Overall, the auto-modification of PARP8 and PARP14 could suggest that these enzymes may have the ability to specifically ADP-ribosylate cysteine and tyrosine residues, respectively. However, further investigation is required to elucidate whether, and how, such specificity would be mediated.

Taken together, the combination of the AI-ETD technology with our Af1521 strategy for unbiased purification of ADP-ribosylated peptides, has culminated in a highly sensitive MS-based proteomics approach that can confidently identify large numbers of ADP-ribosylation events from limited starting material. Moreover, this increased sensitivity makes it feasible to profile ADP-ribosylation in an entirely physiological context, taking a final step towards biomedical and clinical applicability.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

AUTHOR CONTRIBUTIONS

S.C.L. and I.A.H. prepared all samples. M.W., J.M.L, E.S., and J.J.C. established AI-ETD technology and infrastructure. S.C.L., I.A.H., and J.M.L. measured all samples on the mass spectrometer. I.A.H. and S.C.L. processed all MS raw data. I.A.H., S.C.L., and M.R. performed bioinformatics and statistical analyses. M.R. and B.F. wrote the Python script for determination of fragmentation efficiency. M.L.N. and J.J.C. supervised the project. I.A.H., S.C.L., and M.L.N. wrote the manuscript with input from all authors.

ACKNOWLEDGMENTS

The work carried out in this study was in part supported by the Novo Nordisk Foundation Center for Protein Research, the Novo Nordisk Foundation (grant agreement numbers NNF14CC0001 and NNF13OC0006477), The Danish Council of Independent Research (grant agreement numbers 4002-00051, 4183-00322A and 8020-00220B), and The Danish Cancer Society (grant agreement R146-A9159-16-S2). The proteomics technology applied were part of a project that has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement EPIC-XS-823839. We would like to acknowledge the lab of Michael O. Hottiger for the expression and purification of recombinant human PARG (University of Zurich), and we thank members of the NNF-CPR Mass Spectrometry Platform for instrument support and technical assistance.

DECLARATION OF INTERESTS

The authors declare no competing interests.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

REFERENCES

Abplanalp, J., Leutert, M., Frugier, E., Nowak, K., Feurer, R., Kato, J., Kistemaker, H.V.A., Filippov, D.V., Moss, J., Caflisch, A.*, et al.* (2017). Proteomic analyses identify ARH3 as a serine mono-ADP-ribosylhydrolase. Nature communications *8*, 2055.

Altmeyer, M., Messner, S., Hassa, P.O., Fey, M., and Hottiger, M.O. (2009). Molecular mechanism of poly(ADP-ribosyl)ation by PARP1 and identification of lysine residues as ADP-ribose acceptor sites. Nucleic acids research *37*, 3723-3738.

Alvarez-Gonzalez, R., and Althaus, F.R. (1989). Poly(ADP-ribose) catabolism in mammalian cells exposed to DNA-damaging agents. Mutation research *218*, 67-74.

Bache, N., Geyer, P.E., Bekker-Jensen, D.B., Hoerning, O., Falkenby, L., Treit, P.V., Doll, S., Paron, I., Muller, J.B., Meier, F.*, et al.* (2018). A Novel LC System Embeds Analytes in Pre-formed Gradients for Rapid, Ultra-robust Proteomics. Molecular & cellular proteomics : MCP *17*, 2284-2296.

Batth, T.S., Tollenaere, M.A.X., Ruther, P., Gonzalez-Franquesa, A., Prabhakar, B.S., Bekker-Jensen, S., Deshmukh, A.S., and Olsen, J.V. (2019). Protein Aggregation Capture on Microparticles Enables Multipurpose Proteomics Sample Preparation. Molecular & cellular proteomics : MCP *18*, 1027-1035.

Bekker-Jensen, D.B., Kelstrup, C.D., Batth, T.S., Larsen, S.C., Haldrup, C., Bramsen, J.B., Sorensen, K.D., Hoyer, S., Orntoft, T.F., Andersen, C.L.*, et al.* (2017). An Optimized Shotgun Strategy for the Rapid Generation of Comprehensive Human Proteomes. Cell systems *4*, 587-599.e584.

Bilan, V., Leutert, M., Nanni, P., Panse, C., and Hottiger, M.O. (2017). Combining Higher-Energy Collision Dissociation and Electron-Transfer/Higher-Energy Collision Dissociation Fragmentation in a Product-Dependent Manner Confidently Assigns Proteomewide ADP-Ribose Acceptor Sites. Analytical chemistry *89*, 1523-1530.

Bolderson, E., Burgess, J.T., Li, J., Gandhi, N.S., Boucher, D., Croft, L.V., Beard, S., Plowman, J.J., Suraweera, A., Adams, M.N.*, et al.* (2019). Barrier-to-autointegration factor 1 (Banf1) regulates poly [ADP-ribose] polymerase 1 (PARP1) activity following oxidative DNA damage. Nature communications *10*, 5501.

Bonfiglio, J.J., Colby, T., and Matic, I. (2017a). Mass spectrometry for serine ADP-ribosylation? Think o-glycosylation! Nucleic acids research *45*, 6259-6264.

Bonfiglio, J.J., Fontana, P., Zhang, Q., Colby, T., Gibbs-Seymour, I., Atanassov, I., Bartlett, E., Zaja, R., Ahel, I., and Matic, I. (2017b). Serine ADP-Ribosylation Depends on HPF1. Molecular cell *65*, 932-940.e936.

Carter-O'Connell, I., Vermehren-Schmaedick, A., Jin, H., Morgan, R.K., David, L.L., and Cohen, M.S. (2018). Combining Chemical Genetics with Proximity-Dependent Labeling Reveals Cellular Targets of Poly(ADP-ribose) Polymerase 14 (PARP14). ACS chemical biology *13*, 2841-2848.

Chapman, J.D., Gagne, J.P., Poirier, G.G., and Goodlett, D.R. (2013). Mapping PARP-1 auto-ADP-ribosylation sites by liquid chromatography-tandem mass spectrometry. Journal of proteome research *12*, 1868-1880.

Crawford, K., Bonfiglio, J.J., Mikoc, A., Matic, I., and Ahel, I. (2018). Specificity of reversible ADP-ribosylation and regulation of cellular processes. Critical reviews in biochemistry and molecular biology *53*, 64-82.

D'Amours, D., Desnoyers, S., D'Silva, I., and Poirier, G.G. (1999). Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions. The Biochemical journal *342 (Pt 2)*, 249-268.

Daniels, C.M., Ong, S.E., and Leung, A.K. (2014). Phosphoproteomic approach to characterize protein mono- and poly(ADP-ribosyl)ation sites from cells. Journal of proteome research *13*, 3510-3522.

Daniels, C.M., Ong, S.E., and Leung, A.K. (2015). The Promise of Proteomics for the Study of ADP-Ribosylation. Molecular cell *58*, 911-924.

De Vos, M., Schreiber, V., and Dantzer, F. (2012). The diverse roles and clinical relevance of PARPs in DNA damage repair: current state of the art. Biochemical pharmacology *84*, 137-146.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

Fontana, P., Bonfiglio, J.J., Palazzo, L., Bartlett, E., Matic, I., and Ahel, I. (2017). Serine ADP-ribosylation reversal by the hydrolase ARH3. Elife *6*.

Gagne, J.P., Langelier, M.F., Pascal, J.M., and Poirier, G.G. (2018). Hydrofluoric Acid-Based Derivatization Strategy To Profile PARP-1 ADP-Ribosylation by LC-MS/MS. Journal of proteome research *17*, 2542-2551.

Hendriks, I.A., Larsen, S.C., and Nielsen, M.L. (2019). An Advanced Strategy for Comprehensive Profiling of ADP-ribosylation Sites Using Mass Spectrometry-based Proteomics. Mol Cell Proteomics *18*, 1010-1026.

Hogrebe, A., von Stechow, L., Bekker-Jensen, D.B., Weinert, B.T., Kelstrup, C.D., and Olsen, J.V. (2018). Benchmarking common quantification strategies for large-scale phosphoproteomics. Nature communications *9*, 1045.

Jungmichel, S., Rosenthal, F., Altmeyer, M., Lukas, J., Hottiger, M.O., and Nielsen, M.L. (2013). Proteome-wide identification of poly(ADP-Ribosyl)ation targets in different genotoxic stress responses. Molecular cell *52*, 272-285.

Karras, G.I., Kustatscher, G., Buhecha, H.R., Allen, M.D., Pugieux, C., Sait, F., Bycroft, M., and Ladurner, A.G. (2005). The macro domain is an ADP-ribose binding module. The EMBO journal *24*, 1911-1920.

Kelstrup, C.D., Bekker-Jensen, D.B., Arrey, T.N., Hogrebe, A., Harder, A., and Olsen, J.V. (2018). Performance Evaluation of the Q Exactive HF-X for Shotgun Proteomics. Journal of proteome research *17*, 727-738.

Larsen, S.C., Hendriks, I.A., Lyon, D., Jensen, L.J., and Nielsen, M.L. (2018). Systems-wide Analysis of Serine ADP-Ribosylation Reveals Widespread Occurrence and Site-Specific Overlap with Phosphorylation. Cell Rep *24*, 2493-2505 e2494.

Ledvina, A.R., Beauchene, N.A., McAlister, G.C., Syka, J.E., Schwartz, J.C., Griep-Raming, J., Westphall, M.S., and Coon, J.J. (2010). Activated-ion electron transfer dissociation improves the ability of electron transfer dissociation to identify peptides in a complex mixture. Analytical chemistry *82*, 10068-10074.

Leslie Pedrioli, D.M., Leutert, M., Bilan, V., Nowak, K., Gunasekera, K., Ferrari, E., Imhof, R., Malmstrom, L., and Hottiger, M.O. (2018). Comprehensive ADP-ribosylome analysis identifies tyrosine as an ADP-ribose acceptor site. EMBO reports *19*.

Leutert, M., Menzel, S., Braren, R., Rissiek, B., Hopp, A.K., Nowak, K., Bisceglie, L., Gehrig, P., Li, H., Zolkiewska, A*., et al.* (2018). Proteomic Characterization of the Heart and Skeletal Muscle Reveals Widespread Arginine ADP-Ribosylation by the ARTC1 Ectoenzyme. Cell reports *24*, 1916-1929.e1915.

Li, P., Zhen, Y., and Yu, Y. (2019). Site-specific analysis of the Asp- and Glu-ADP-ribosylated proteome by quantitative mass spectrometry. Methods in enzymology *626*, 301-321.

Lüscher, B., Bütepage, M., Eckei, L., Krieg, S., Verheugd, P., and Shilton, B.H. (2017). ADP-Ribosylation, a Multifaceted Posttranslational Modification Involved in the Control of Cell Physiology in Health and Disease. Chemical reviews.

Martello, R., Leutert, M., Jungmichel, S., Bilan, V., Larsen, S.C., Young, C., Hottiger, M.O., and Nielsen, M.L. (2016). Proteome-wide identification of the endogenous ADP-ribosylome of mammalian cells and tissue. Nature communications *7*, 12917.

McCool, E.N., Lodge, J.M., Basharat, A.R., Liu, X., Coon, J.J., and Sun, L. (2019). Capillary Zone Electrophoresis-Tandem Mass Spectrometry with Activated Ion Electron Transfer Dissociation for Large-scale Top-down Proteomics. Journal of the American Society for Mass Spectrometry.

McDonald, L.J., and Moss, J. (1994). Enzymatic and nonenzymatic ADP-ribosylation of cysteine. Molecular and cellular biochemistry *138*, 221-226.

Menissier de Murcia, J., Ricoul, M., Tartier, L., Niedergang, C., Huber, A., Dantzer, F., Schreiber, V., Ame, J.C., Dierich, A., LeMeur, M*., et al.* (2003). Functional interaction between PARP-1 and PARP-2 in chromosome stability and embryonic development in mouse. EMBO J *22*, 2255-2263.

Moss, J., and Vaughan, M. (1977). Mechanism of action of choleragen. Evidence for ADP-ribosyltransferase activity with arginine as an acceptor. The Journal of biological chemistry *252*, 2455-2457.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

Muthurajan, U.M., Hepler, M.R., Hieb, A.R., Clark, N.J., Kramer, M., Yao, T., and Luger, K. (2014). Automodification switches PARP-1 function from chromatin architectural protein to histone chaperone. Proceedings of the National Academy of Sciences of the United States of America *111*, 12752-12757.

Ogata, N., Ueda, K., and Hayaishi, O. (1980). ADP-ribosylation of histone H2B. Identification of glutamic acid residue 2 as the modification site. The Journal of biological chemistry *255*, 7610-7615.

Palazzo, L., Leidecker, O., Prokhorova, E., Dauben, H., Matic, I., and Ahel, I. (2018). Serine is the major residue for ADP-ribosylation upon DNA damage. Elife *7*.

Qin, W., Wu, H.J., Cao, L.Q., Li, H.J., He, C.X., Zhao, D., Xing, L., Li, P.Q., Jin, X., and Cao, H.L. (2019). Research Progress on PARP14 as a Drug Target. Frontiers in pharmacology *10*, 172.

Riley, N.M., Hebert, A.S., Durnberger, G., Stanek, F., Mechtler, K., Westphall, M.S., and Coon, J.J. (2017a). Phosphoproteomics with Activated Ion Electron Transfer Dissociation. Analytical chemistry *89*, 6367-6376.

Riley, N.M., Hebert, A.S., Westphall, M.S., and Coon, J.J. (2019). Capturing site-specific heterogeneity with large-scale N-glycoproteome analysis. Nature communications *10*, 1311.

Riley, N.M., Sikora, J.W., Seckler, H.S., Greer, J.B., Fellers, R.T., LeDuc, R.D., Westphall, M.S., Thomas, P.M., Kelleher, N.L., and Coon, J.J. (2018a). The Value of Activated Ion Electron Transfer Dissociation for High-Throughput Top-Down Characterization of Intact Proteins. Analytical chemistry *90*, 8553-8560.

Riley, N.M., Westphall, M.S., and Coon, J.J. (2018b). Sequencing Larger Intact Proteins (30-70 kDa) with Activated Ion Electron Transfer Dissociation. Journal of the American Society for Mass Spectrometry *29*, 140-149.

Riley, N.M., Westphall, M.S., Hebert, A.S., and Coon, J.J. (2017b). Implementation of Activated Ion Electron Transfer Dissociation on a Quadrupole-Orbitrap-Linear Ion Trap Hybrid Mass Spectrometer. Analytical chemistry *89*, 6358-6366.

Rose, C.M., Rush, M.J., Riley, N.M., Merrill, A.E., Kwiecien, N.W., Holden, D.D., Mullen, C., Westphall, M.S., and Coon, J.J. (2015). A calibration routine for efficient ETD in large-scale proteomics. Journal of the American Society for Mass Spectrometry *26*, 1848-1857.

Schreiber, V., Ame, J.C., Dolle, P., Schultz, I., Rinaldi, B., Fraulob, V., Menissier-de Murcia, J., and de Murcia, G. (2002). Poly(ADP-ribose) polymerase-2 (PARP-2) is required for efficient base excision DNA repair in association with PARP-1 and XRCC1. J Biol Chem *277*, 23028-23036.

Sekine, A., Fujiwara, M., and Narumiya, S. (1989). Asparagine residue in the rho gene product is the modification site for botulinum ADP-ribosyltransferase. The Journal of biological chemistry *264*, 8602-8605.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research *13*, 2498-2504.

Shishkova, E., Hebert, A.S., Westphall, M.S., and Coon, J.J. (2018). Ultra-High Pressure (>30,000 psi) Packing of Capillary Columns Enhancing Depth of Shotgun Proteomic Analyses. Analytical chemistry *90*, 11503-11508.

Spitzer, M., Wildenhain, J., Rappsilber, J., and Tyers, M. (2014). BoxPlotR: a web tool for generation of box plots. Nature methods *11*, 121-122.

Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N.T., Roth, A., Bork, P.*, et al.* (2017). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic acids research *45*, D362-d368.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nature methods *13*, 731-740.

Van Ness, B.G., Howard, J.B., and Bodley, J.W. (1980). ADP-ribosylation of elongation factor 2 by diphtheria toxin. NMR spectra and proposed structures of ribosyl-diphthamide and its hydrolysis products. The Journal of biological chemistry *255*, 10710-10716.

Vyas, S., Matic, I., Uchima, L., Rood, J., Zaja, R., Hay, R.T., Ahel, I., and Chang, P. (2014). Family-wide analysis of poly(ADP-ribose) polymerase activity. Nature communications *5*, 4426.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

Zhang, Y., Wang, J., Ding, M., and Yu, Y. (2013). Site-specific characterization of the Asp- and Glu-ADP-ribosylated proteome. Nature methods *10*, 981-984.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

FIGURE LEGENDS

Figure 1. Comparison of AI-ETD vs. ETD and EThcD for mapping ADP-ribosylation. (A) Overview of the strategy employed to enrich ADP-ribosylated peptides. For technical comparison of dissociation methods, all samples were analyzed in technical quadruplicate; *n*=4. All dissociation methods were based on ETD to facilitate localization of the labile ADPr. (B) Overview of the number of ADPr peptide-spectrum-matches (PSMs) identified and localized (>90% probability) for each dissociation method and Supplemental Activation Energy (SAE). Error bars represent SD, *n*=4 technical replicates. (C) As **B**, but displaying the number of ADPr sites identified. (D) As **B**, but displaying the spectral quality (in Andromeda Score) of all identified ADPr-modified peptides. Distribution of data points is visualized, line limits; 1.5× interquartile range (IQR), box limits; 3rd and 1st quartiles, white dot; mean. Number of data points (*n*) is visualized below the distributions. (E) ADP-ribose localization probability plotted against the ranked fraction of all peptide-spectrum-matches (PSMs) resulting from the different dissociation methods and SAEs. Note that although all probabilities are displayed, only those over 0.9 were used for assignment of ADP-ribosylated peptides and sites. (F) Visualization of the average relative degree of non-dissociative electron transfer (ETnoD). Derived from all peptide-identified MS/MS spectra, and separately visualized for unmodified and ADP-ribosylated peptides. Error bars represent 5× SEM. (G) Visualization of the average degree of precursor fragmentation, calculated by dividing observed fragment ion peak intensity by the sum of non-ETD, ETnoD, and all fragment ion peak intensities. Derived from all peptide-identified MS/MS spectra, and separately visualized for unmodified, ADP-ribosylated, and localized ADP-ribosylated peptides. Error bars represent 5× SEM. (H) Spectral quality (in Andromeda Score) plotted against the average degree of precursor fragmentation, for localized ADP-ribosylated peptides. Coloring represents the relative density of dots in the plot, with higher values corresponding to higher density.

Figure 2. Benchmarking of AI-ETD. (A) Overview of the number of ADPr sites identified in this experiment, as compared to previous screens performed with the same ADPr-modified peptide enrichment strategy (Hendriks et al., 2019; Larsen et al., 2018). “Input” corresponds to the relative amount of the summed total ion current (observed at the MS1-level) across the analytical gradients of all experiments, and correlates with effective sample load on the column. (B) Scaled Venn diagram visualizing distribution of identified ADP-ribosylation sites in this experiment compared to two other ADP-ribosylation studies (Hendriks et al., 2019; Larsen et al., 2018). (C) Pie-chart overview of the amino acid residue distribution of all ADP-ribosylation sites identified in this experiment (left pie-chart), and ADP-

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

ribosylation sites exclusively identified in this study (right pie-chart) as compared to two other ADP-ribosylation studies (Hendriks et al., 2019; Larsen et al., 2018). (D) Term enrichment analysis using Gene Ontology annotations and Uniprot keywords, comparing proteins identified to be ADP-ribosylated in this study to the total proteome. Relative score is based on multiplication of logarithms derived from the enrichment ratio and the *q*-value. All displayed terms were significant with $q<0.02$, as determined through Fisher Exact Testing with Benjamini-Hochberg correction. The full term enrichment analysis is available in Table S5.

Figure 3. Analysis of physiological ADPr using AI-ETD. (A) Overview of the number of physiological ADPr sites identified in this study, compared to a previous study which also included physiological ADPr (Larsen et al., 2018). “Total” represents the number of unique ADPr sites identified across all four cell culture replicates, “Best rep.” represents the number of unique ADPr sites identified in the best replicate. “ADPr sites / minute” was calculated by dividing the best replicate number of ADPr sites by the effective analytical gradient time in minutes. “SS”; single-shot analysis, “All”; high-pH fractionated analysis. (B) Left; pie-chart overview of the amino acid residue distribution of all physiological ADP-ribosylation sites identified in this study. Right; scaled Venn diagram visualizing distribution of identified physiological ADP-ribosylation sites in this study compared to Larsen et al., 2018 (Larsen et al., 2018). (C) Average localization probability across all ADPr PSMs that were at least partially localized (>51%) to each specific amino acid residue type. Only amino acids with >10 partially localized PSMs were included. Error bars represent SEM, number of data points (*n*) is indicated. (D) Depth of sequencing analysis, plotting number of identified physiological ADPr target proteins versus their known copy-numbers. Proteins ADP-ribosylated in response to H_2O_2 were derived from this study, as well as various other studies (Bilan et al., 2017; Jungmichel et al., 2013; Larsen et al., 2018; Martello et al., 2016; Zhang et al., 2013). Protein copy-numbers were derived from a deep HeLa proteome study (Bekker-Jensen et al., 2017), which covers the vast majority (>99%) of all ADP-ribosylated proteins identified. Dotted lines with color indicator below represent the median protein copy-number for ADPr target proteins within each respective subset. (E) ‘Fidget spinner’ analysis, comprised of; scaled Venn diagram (center) visualizing distribution of H_2O_2-induced and physiological ADP-ribosylation sites; pie-chart overviews (around the center) of the amino acid residue distribution of each subset of ADPr sites, as indicated by background coloring; term enrichment analyses (outermost graphs) using Gene Ontology annotations and Uniprot keywords, comparing proteins identified to be ADP-ribosylated in each subset of ADPr sites as compared to the other subsets, indicated by background coloring. Relative score is based on multiplication of logarithms derived from the enrichment ratio and the *q*-value. All displayed terms were

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

significant with q<0.02, as determined through Fisher Exact Testing with Benjamini-Hochberg correction. The full term enrichment analysis is available in Table S5. Background and Venn coloring, red; H_2O_2-only, purple; both H_2O_2 and physiological, blue; physiological only.

Figure 4. ADP-ribosylation of different PARP family members. (A) Overview of the most abundantly ADPr-modified residues within PARP family members, as a fraction of total protein ADPr modification, either under physiological conditions ("Phys.") or in response to H_2O_2. All physiologically detected ADPr sites, or otherwise up to four of the most abundant total ADPr sites, are indicated on the graphical representation of the PARP family members. Blue lines; physiological ADPr sites, red lines; H_2O_2-induced ADPr sites. (B) Overview of the total amount of ADPr detected on each PARP family member, in relation to their known expression level ("IBAQ"). Both values are logarithmic, and graph scaling was normalized to PARP1 and PARP2 values. (C) STRING network visualizing functional interactions between proteins identified to be ADPr-modified on cysteine residues in this study or in Larsen et al., 2018 (Larsen et al., 2018). Default STRING clustering confidence was used (P>0.4), and disconnected proteins were omitted from the network. Proteins were significantly interconnected, with a protein-protein interaction enrichment q-value of $1.1*10^{-11}$. Nodes corresponding to proteins were annotated with visual properties, as highlighted in the figure legend. "Cys"; modification on cysteine residues, "Other"; modification on non-cysteine residues, "Phys."; physiological conditions. Distribution of color is relative to the number of sites within each category, sites detected both physiologically and in response to H_2O_2 were colored as physiological.

Figure S1. Comparison of AI-ETD vs. ETD and EThcD for mapping ADP-ribosylation in Lys-C digested peptides. (A) Overview of the number of ADPr peptide-spectrum-matches (PSMs) identified and localized (>90% probability) for each dissociation method and Supplemental Activation Energy (SAE). Error bars represent SD, n=4 technical replicates. (B) As **A**, but displaying the number of ADPr sites identified. (C) As **A**, but displaying the spectral quality (in Andromeda Score) of all identified ADPr-modified peptides. Distribution of data points is visualized, line limits; 1.5× interquartile range (IQR), box limits; 3rd and 1st quartiles, white dot; mean. Number of data points (n) is visualized below the distributions. (D) Visualization of the average relative degree of non-dissociative electron transfer (ETnoD). Derived from all peptide-identified MS/MS spectra, and separately visualized for unmodified and ADP-ribosylated peptides. Error bars represent 5× SEM. (E) Visualization of the average degree of precursor fragmentation, calculated by dividing observed fragment ion peak intensity by the sum of non-ETD, ETnoD, and all fragment ion peak intensities. Derived from all peptide-identified MS/MS spectra,

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

and separately visualized for unmodified, ADP-ribosylated, and localized ADP-ribosylated peptides. Error bars represent 5× SEM. (F) Spectral quality (in Andromeda Score) plotted against the average degree of precursor fragmentation, for unmodified peptides detected within ADPr samples. Coloring represents the relative density of dots in the plot, with higher values corresponding to higher density. (G) As **F**, but for non-localized ADP-ribosylated peptides.

Figure S2. Distribution of ADPr sites across proteins. (A) Visualization of the number of unique ADPr sites detected per protein.

Figure S3. KS motif adherence and subcellular localization. (A) Overview of KS motif adherence of all serine ADPr sites identified under physiological conditions or in response to H_2O_2 treatment, in this study or in two others (Hendriks et al., 2019; Larsen et al., 2018). KS motif modification is defined as an ADPr-modified serine with an N-terminal lysine residue. (B) Subcellular localization of ADPr target proteins identified under physiological conditions, with localization derived from Gene Ontology Cellular Compartments. (C) As **B**, but for ADPr target proteins identified in response to H_2O_2 treatment.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

METHODS

Cell culture

HeLa cells (CCL-2) were acquired via the American Type Culture Collection, and cultured at 37 °C and 5% CO2 in Dulbecco's Modified Eagle's Medium (Invitrogen) supplemented with 10% fetal bovine serum and a penicillin/streptomycin mixture (100 U/mL; Gibco). Cells were routinely tested for mycoplasma. Cells were not routinely authenticated. For initial optimization and benchmarking experiments, ADP-ribosylation was induced in HeLa by treatment of the cells with 1 mM H_2O_2 (Sigma Aldrich) for 10 min in PBS at 37°C. Five cell culture replicates were prepared, each corresponding to approximately 100 million HeLa cells, of which one was used for optimization, and four for benchmarking. For physiological experiments, cells were left untreated, and four cell culture replicates were prepared, each corresponding to approximately 100 million HeLa cells.

Cell lysis and protein digestion

The full procedure for enrichment of ADPr from cells was done as described previously (Hendriks et al., 2019; Larsen et al., 2018). Briefly, cells were washed twice with ice-cold PBS, and gently scraped at 4 °C in a minimal volume of PBS. Cells were pelleted by centrifugation at 500*g*, and lysed in 10 pellet volumes of Lysis Buffer (6 M guanidine-HCl, 50 mM TRIS, pH 8.5). Complete lysis was achieved by alternating vigorous shaking with vigorous vortexing, for 30 seconds, prior to snap freezing of the lysates using liquid nitrogen. Frozen lysates were stored at −80 °C until further processing. Lysates were thawed and sonicated at 30 W, for 1 second per 1 mL of lysate, spread across 3 separate pulses. Tris(2-carboxyethyl)phosphine and chloroacetamide were added to a final concentration of 5 mM, and the lysate was incubated for 1 hour at 30 °C. Proteins were digested using Lysyl Endopeptidase (Lys-C, 1:100 w/w; Wako Chemicals) for 3 hours, and diluted with three volumes of 50 mM ammonium bicarbonate. Half of the samples was further digested overnight using modified sequencing grade Trypsin (1:200 w/w; Sigma Aldrich), and the other half was re-digested overnight using Lys-C. Digested samples were acidified by addition of trifluoroacetic acid (TFA) to a final concentration of 0.5% (v/v), cleared by centrifugation, and purified using reversed-phase C18 cartridges (SepPak Classic, 360 mg sorbent, Waters) according to the manufacturer's instructions. Elution of peptides was performed with 30% or 40% ACN in 0.1% TFA, for peptides digested with either trypsin or Lys-C, respectively. Eluted peptides were frozen overnight at −80 °C, and lyophilized.

Purification of ADP-ribosylated peptides

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

GST-tagged Af1521 macrodomain was produced in-house and coupled to glutathione Sepharose 4B beads (Sigma-Aldrich), essentially as described previously (Hendriks et al., 2019; Larsen et al., 2018). Lyophilized peptides were dissolved in IP buffer (50 mM TRIS pH 8.0, 1 mM $MgCl_2$, 250 µM DTT, and 50 mM NaCl), after which long ADP-ribose polymers were reduced to monomers using recombinant PARG at a concentration of 1:10,000 (w/w), overnight and at room temperature. Subsequently, 100 µL of sepharose beads with GST-tagged Af1521 were added to the samples, and mixed at 4 °C for 3 h. Beads were sequentially washed four times with ice-cold IP Buffer, two times with ice-cold PBS, and two times with ice-cold MQ water. On the first wash, beads were transferred to 1.5 mL LoBind tubes (Eppendorf), and LoBind tubes were exclusively used from this point on to minimize loss of peptide. Additional tube changes were performed every second washing step to minimize carryover of background. ADP-ribosylated peptides were removed from the beads using two elution steps with 200 µL ice-cold 0.15% TFA, and the pooled elutions were cleared through 0.45 µm spin filters (Ultrafree-MC, Millipore) and subsequently through pre-washed 100 kDa cut-off filters (Vivacon 500, Sartorius). The filtered ADP-ribosylated peptides were purified using C18 StageTips at high pH (Hendriks et al., 2019), and eluted either as single shot samples (optimization and physiological experiments), or eluted as seven fractions (benchmarking experiments). Briefly, samples were basified by adding ammonium solution to a final concentration of 20 mM, and loaded onto StageTips carrying four layers of C18 disc material (punch-outs from 47mm C18 3M™ extraction discs, Empore). Elution was performed with 25% of acetonitrile (ACN) in 20 mM ammonium hydroxide for single shot samples, or sequentially performed with 2% (F1), 4% (F2), 7% (F3), 10% (F4), 15% (F5), and 25% (F6) of ACN in 20 mM ammonium hydroxide for fractionated samples. The seventh fraction (F0) was prepared by performing StageTip purification at low pH on the flowthrough fraction from sample loading at high pH. All samples were completely dried using a SpeedVac at 60 °C, and dissolved in a small volume of 0.1% formic acid. Final samples were frozen at −20 °C until measurement.

Mass spectrometric analysis

All MS samples were analyzed using a Fusion Lumos Orbitrap mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) modified for AI-ETD, coupled to a Dionex UltiMate 3000 nano-UHPLC system (Thermo Fisher, Sunnyvale, CA, USA). For AI-ETD implementation, a Firestar T-100 Synrad 60 Watt (W) CO_2 continuous wave laser (Mukilteo, WA, USA) was mounted at the back of the instrument. The IR photon beam was guided into the traps using focusing lenses and a multi-mode hollow-core fiber (Opto-Knowledge Systems, Torrance, CA, USA) prior to a zinc selenide window that was placed in the vacuum manifold. Chromatographic separation of peptides was performed on a 30-cm long analytical

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

column, with an internal diameter of 1.7 µm, packed in-house as described previously (Shishkova et al., 2018) using 130 Å pore size Bridged Ethylene Hybrid C18 particles (Waters, Milford, USA). The analytical column was heated to 50 °C using an in-house made column oven. Peptides were eluted from the column using a gradient of Buffer A (0.1% formic acid) and Buffer B (80% ACN in 0.1% formic acid). The primary gradient ranged from 3% buffer B to 24% buffer B over 50 minutes, preceded by sample loading time (5 minutes for evaluation experiments, 20 minutes for benchmarking and physiological experiments), and followed by a 20 minute washing block. Electrospray ionization (ESI) was achieved using a Nanospray Flex Ion Source (Thermo). Spray voltage was set to 2.25 kV, capillary temperature to 275°C, and RF level to 30%. Full scans were performed at a resolution of 60,000, with a scan range of 300 to 1,750 m/z, a maximum injection time of 60 ms, and an automatic gain control (AGC) target of 600,000 charges. Precursors were isolated with a width of 1.3 m/z, with an AGC target of 200,000 charges, and precursor fragmentation was accomplished using either electron transfer dissociation (ETD), electron transfer disassociation with supplemental higher-collisional disassociation (EThcD) at an NCE of 20, or activated-ion electron transfer dissociation (AI-ETD); all using calibrated charge-dependent ETD parameters (Rose et al., 2015). The laser powers were the percentage of power (W) from the Firestar T-100 Synrad 60-W CO_2, continuous wave laser. For AI-ETD evaluation, laser power settings of 10% (6 W), 15% (9 W), 20% (12 W), 25% (15 W), and 30% (18 W) were used, and for benchmarking and physiological experiments a laser power setting of 20% was used. Precursors with charge state 2-5 were isolated for MS/MS analysis, and prioritized from charge 3 (highest), to 4, to 5, to 2 (lowest), using the decision tree algorithm. For benchmarking and physiological experiments, only precursors with charge state 3-5 were isolated. Precursors were excluded from repeated sequencing by setting a dynamic exclusion time of 75 seconds for evaluation experiments, and 90 seconds for benchmarking and physiological experiments, with an exclusion mass tolerance of 10 ppm. MS/MS spectra were measured in the Orbitrap, with 5 data-dependent MS/MS scans per full MS scan, a maximum precursor injection time of 120 ms, and a scan resolution of 60,000. First mass for MS/MS scans was set to 100 for ETD and EThcD measurements, and variable for AI-ETD measurements, with the first mass defining the laser power applied. MS/MS first masses for AI-ETD were 141, 131, 142, 132, and 143, corresponding to application of laser powers 10%, 15%, 20%, 25%, and 30%, respectively.

Data analysis

Analysis of the mass spectrometry raw data was performed using MaxQuant software (version 1.5.3.30). MaxQuant default settings were used, with exceptions outlined below. Two separate computational

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

searches were performed, one for evaluation and physiological data, and the other for benchmarking data. For generation of the theoretical spectral library, a HUMAN.fasta database was extracted from UniProt on the 24th of May, 2019. N-terminal acetylation, methionine oxidation, cysteine carbamidomethylation, and ADP-ribosylation on all amino acid residues known to potentially be modified (C, D, E, H, K, R, S, T, and Y), were included as variable modifications. For analysis of trypsin samples, up to 6 missed cleavages were allowed. For analysis of Lys-C samples, up to 3 missed cleavages were allowed. A maximum allowance of 4 variable modifications per peptide was used. Second peptide search was enabled (default), and matching between runs was enabled with a match time window of 42 seconds and an alignment time window of 20 minutes. Mass tolerance for precursors was set to 20 ppm in the first MS/MS search and 4.5 ppm in the main MS/MS search after mass recalibration. For fragment ion masses, a tolerance of 20 ppm was used. Modified peptides were filtered to have an Andromeda score of >40 (default), and a delta score of >20. Data was automatically filtered by posterior error probability to achieve a false discovery rate of <1% (default), at the peptide-spectrum match, the protein assignment, and the site-specific levels.

Data filtering

Beyond automatic filtering and FDR control as applied by MaxQuant, the data was manually filtered in order to ensure proper identification and localization of ADP-ribose. PSMs modified by more than one ADP-ribose were omitted. PSMs corresponding to unique peptides were only used for ADP-ribosylation site assignment if localization probability was >0.90, with localization of >0.75 accepted only for purposes of intensity assignment of further evidences. Erroneous MaxQuant intensity assignments were manually corrected in the sites table, and based on localized PSMs only (>0.90 best-case, >0.75 for further evidences). For the ADP-ribosylation target proteins table, the proteinGroups.txt file generated by MaxQuant was filtered to only contain those proteins with at least one ADP-ribosylation site detected and localized post-filtering as outlined above, with cumulative ADP-ribosylation site intensities based only on localized evidences.

Determination of fragmentation efficiency

An in-house Python script was used to determine the intensities of non-reduced precursors (no ETD), and the corresponding non-dissociated charge-reduced precursor ions (ETnoD) for each MS/MS spectrum. For that, the centroided, de-isotoped, and charge-deconvoluted peak lists from the MaxQuant Andromeda Peak List (APL) files were used. Second peptide search APLs were not considered. “No ETD” precursor peaks were expected at:

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

$$m/z = Precursor_{m/z} * Charge - Charge + 1$$

"ETnoD" peaks were expected at:

$$ETnoD_{m/z} = Precursor_{m/z} * Charge - (0, 1, \dots, Charge - 2)$$

This compensates for MaxQuant charge deconvolution, which goes by the assumption that the charge state is purely dependent on added protons. The expected peaks were matched to peaks from the MaxQuant APL files, with a mass tolerance of 0.05 Da, and their intensities were stored. Intensity information was mapped to the msms.txt output file from MaxQuant using the scan number. Total fragmentation of peptides was calculated by division of fragment product ion intensities with the sum of all peak intensities.

Comparison to other studies

Data from other studies was retrieved from several publications and online databases. For ADP-ribosylation sites; Larsen et al. (Larsen et al., 2018), Hendriks et al. (Hendriks et al., 2019), Bonfiglio et al. (Bonfiglio et al., 2017b), Bilan et al. (Bilan et al., 2017), Leslie Pedrioli et al. (Leslie Pedrioli et al., 2018). For ADP-ribosylation proteins; Zhang et al. (Zhang et al., 2013), Jungmichel et al. (Jungmichel et al., 2013), Martello et al. (Martello et al., 2016), Bilan et al. (Bilan et al., 2017), Larsen et al. (Larsen et al., 2018), Hendriks et al. (Hendriks et al., 2019). For total proteome; Bekker-Jensen et al. (Bekker-Jensen et al., 2017). For comparison of proteins between studies, all protein identifiers were mapped to the human proteome as downloaded from Uniprot on the 24th of May, 2019. In case data sources did not include Uniprot IDs, ID mapping on Uniprot was used to convert other IDs to Uniprot IDs, and otherwise gene names were used. Non-existent, non-human, and redundant entries, were discarded from the analysis. For comparison of sites between studies, all parent proteins were mapped to Uniprot IDs as described above, and afterwards the reported positions of modification were used to extract 51 amino acid sequence windows (modified residue +/- 25 amino acids), with the sequence windows ultimately used to directly compare identified sites. Sites mapping to non-existent, non-human, or redundant proteins, were discarded. Sites not correctly aligning to the reported amino acid residues were discarded.

Statistical analysis

Details regarding the statistical analysis can be found in the respective figure legends. Statistical handling of the data was primarily performed using the freely available Perseus software (Tyanova et al., 2016), and includes term enrichment analysis through FDR-controlled Fisher Exact testing, and

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

density estimation for highly populated scatter plots. Protein Gene Ontology annotations and UniProt keywords used for term enrichment analysis were concomitantly downloaded from UniProt with the HUMAN.fasta file used for searching the RAW data. Boxplots and violin plots were generated using the BoxPlotR web tool (Spitzer et al., 2014). The online STRING database (version 11) was used for generation of protein interaction networks (Szklarczyk et al., 2017), and Cytoscape (version 3.7.1) was used for manual annotation and visualization of the STRING networks (Shannon et al., 2003).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

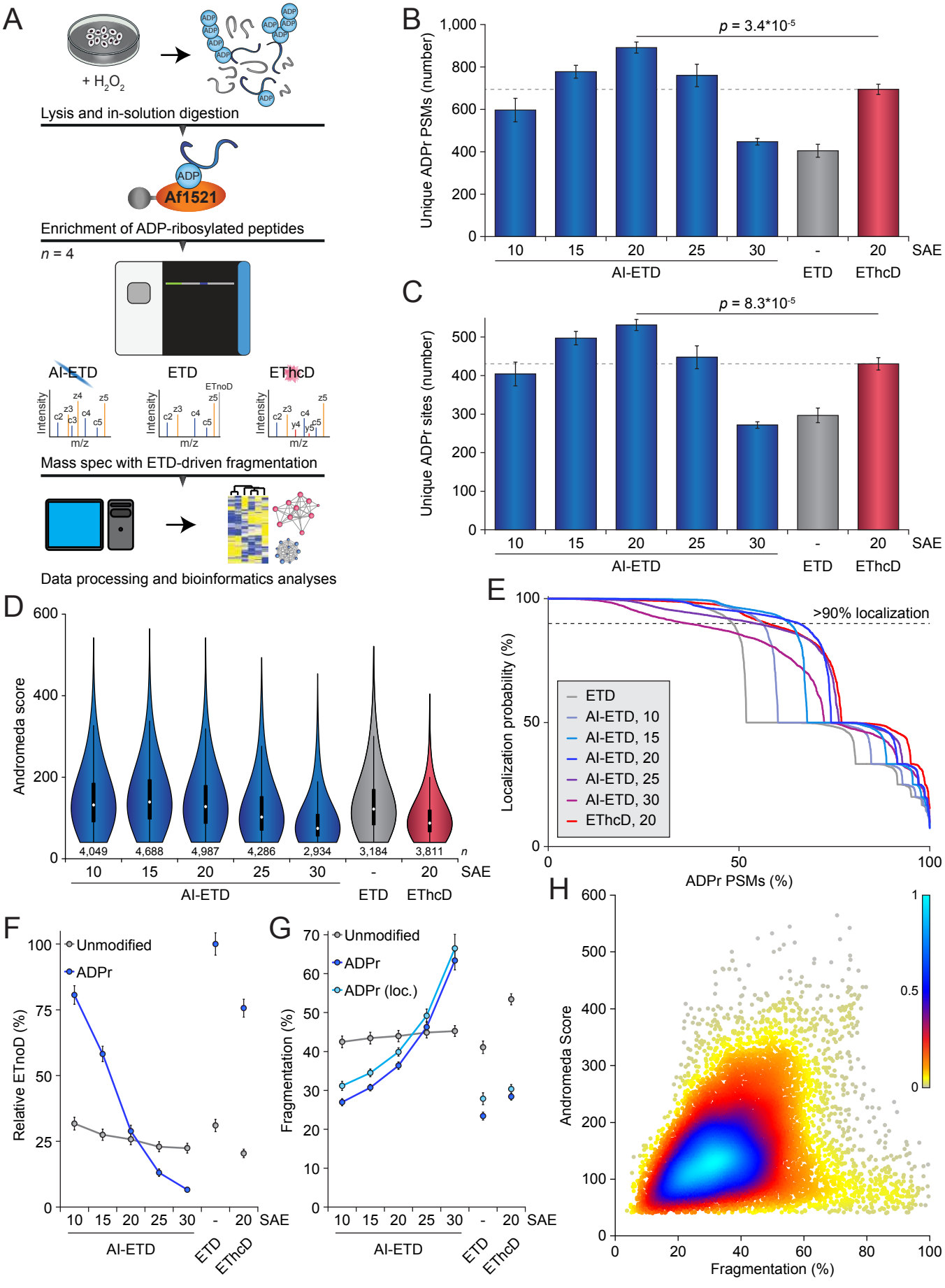


FIGURE 1

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

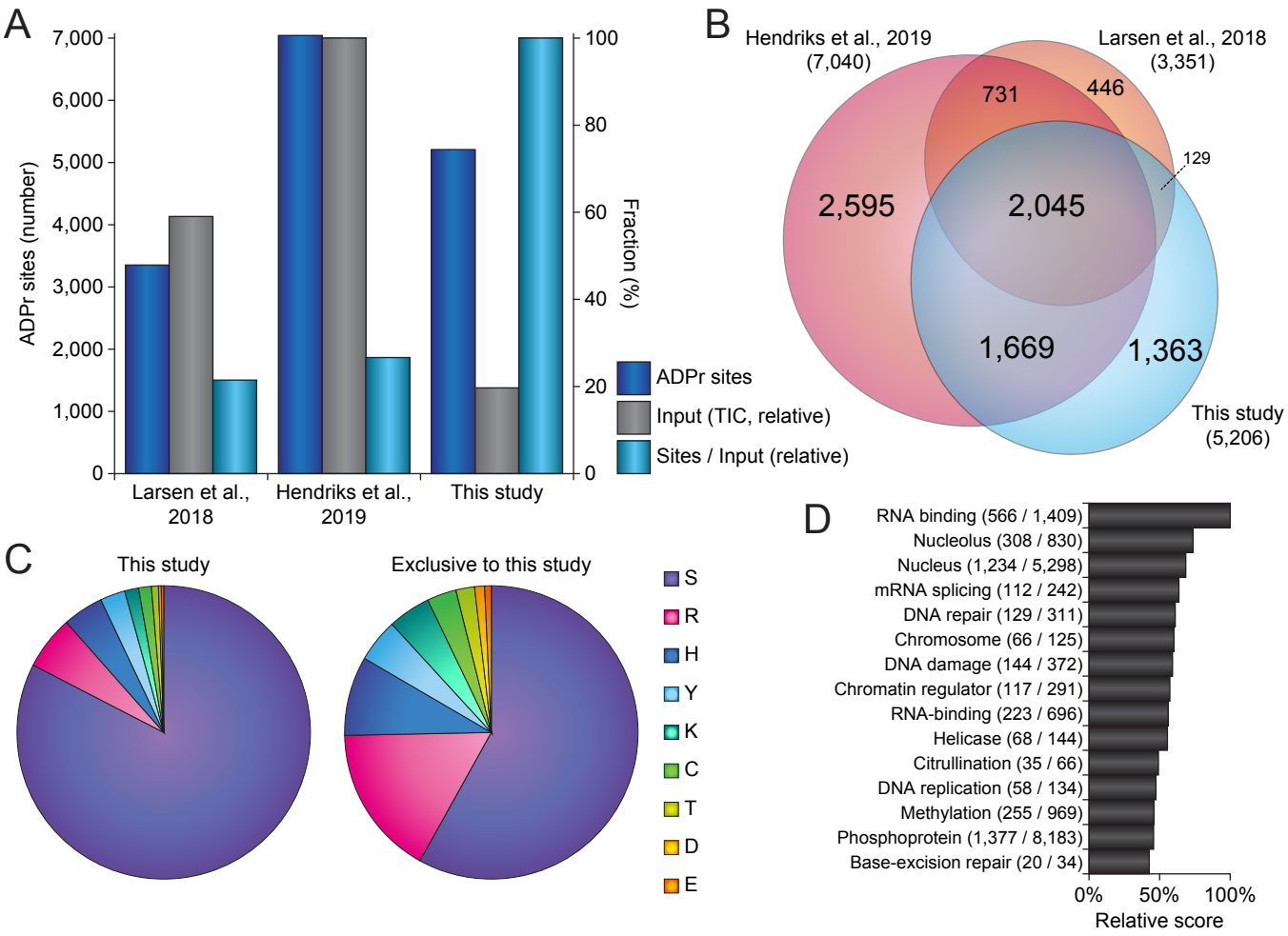


FIGURE 2

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

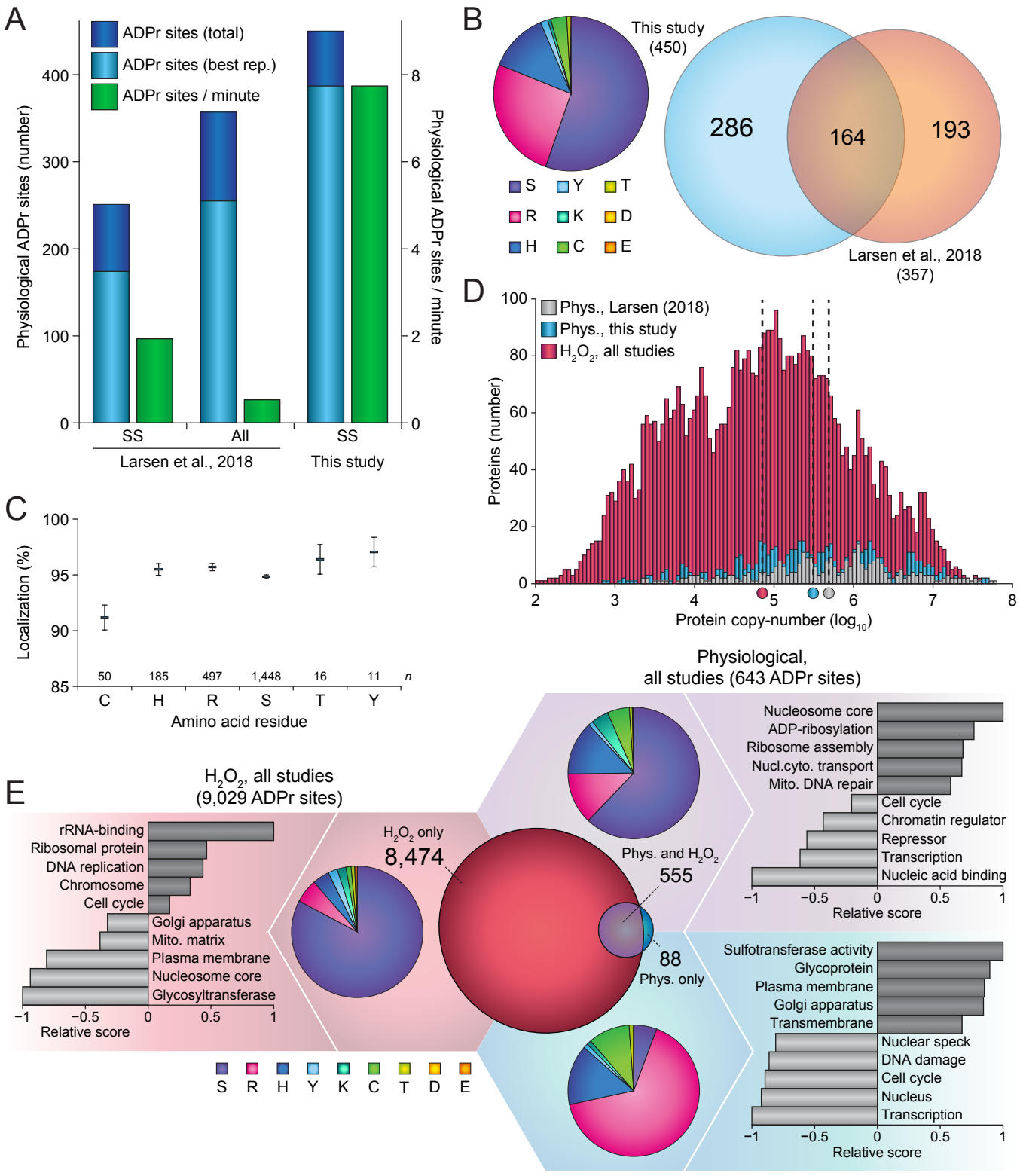
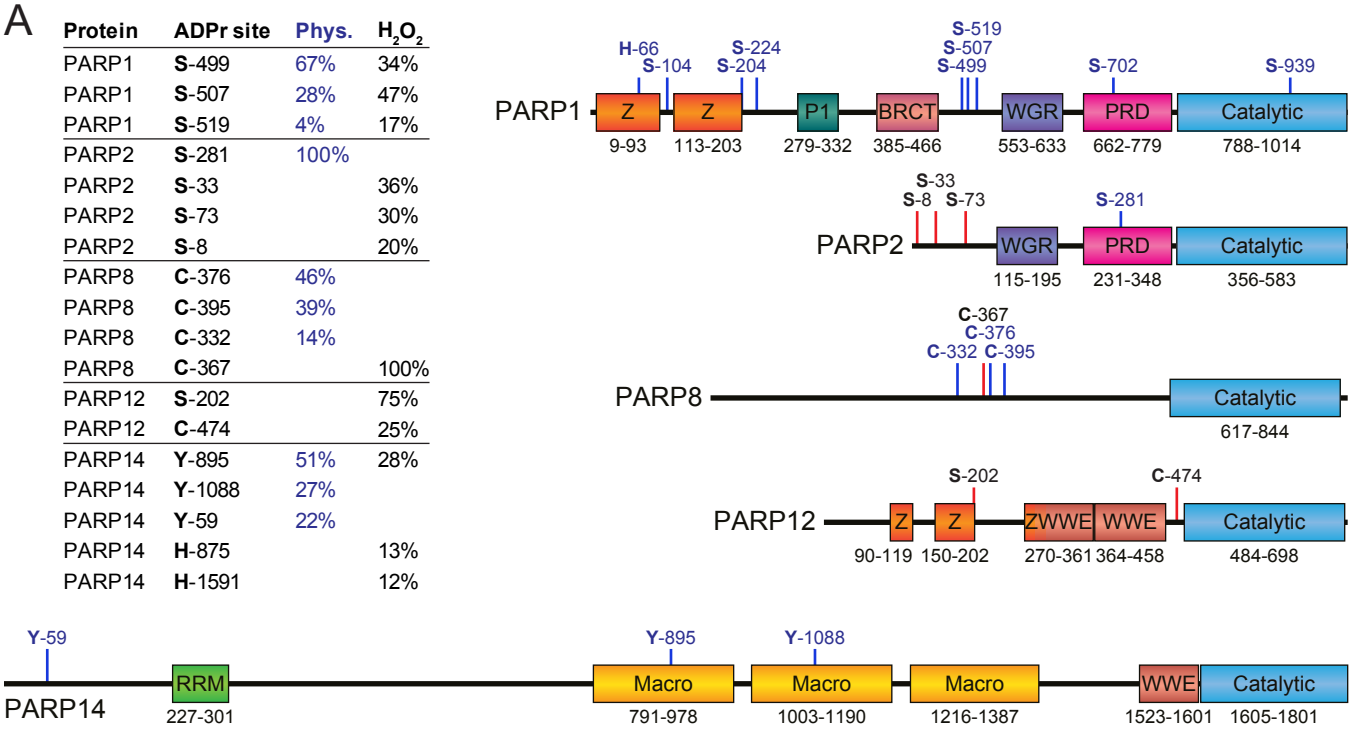


FIGURE 3

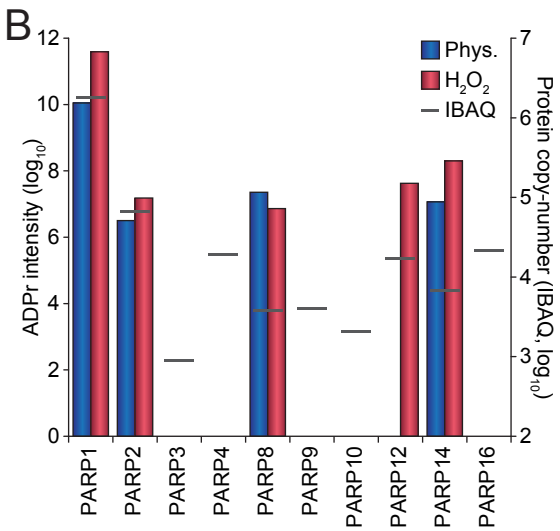
bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A

Protein	ADPr site	Phys.	H_2O_2
PARP1	**S**-499	67%	34%
PARP1	**S**-507	28%	47%
PARP1	**S**-519	4%	17%
PARP2	**S**-281	100%	
PARP2	**S**-33		36%
PARP2	**S**-73		30%
PARP2	**S**-8		20%
PARP8	**C**-376	46%	
PARP8	**C**-395	39%	
PARP8	**C**-332	14%	
PARP8	**C**-367		100%
PARP12	**S**-202		75%
PARP12	**C**-474		25%
PARP14	**Y**-895	51%	28%
PARP14	**Y**-1088	27%	
PARP14	**Y**-59	22%	
PARP14	**H**-875		13%
PARP14	**H**-1591		12%



B



C

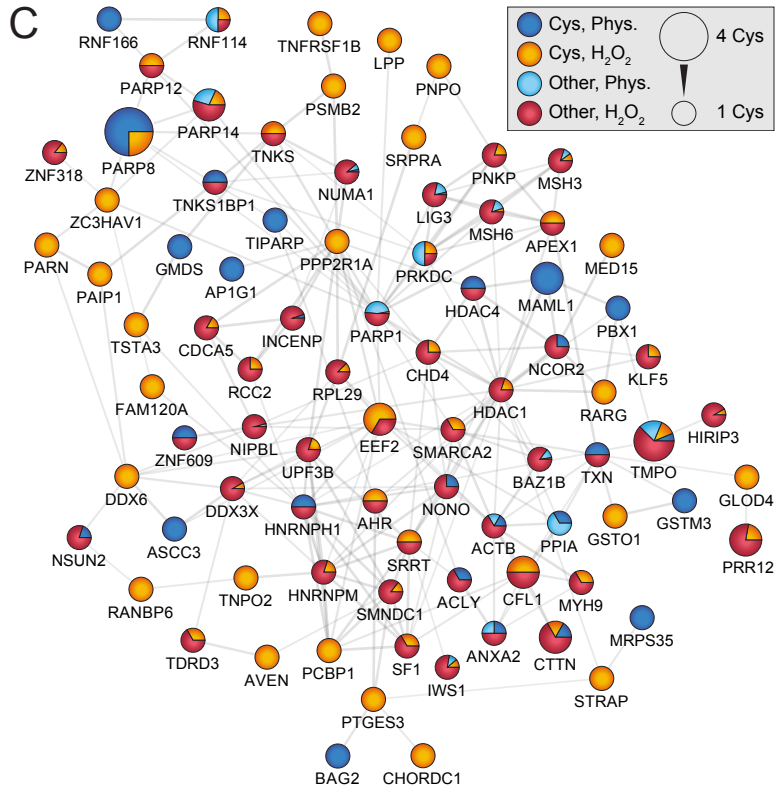


FIGURE 4

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

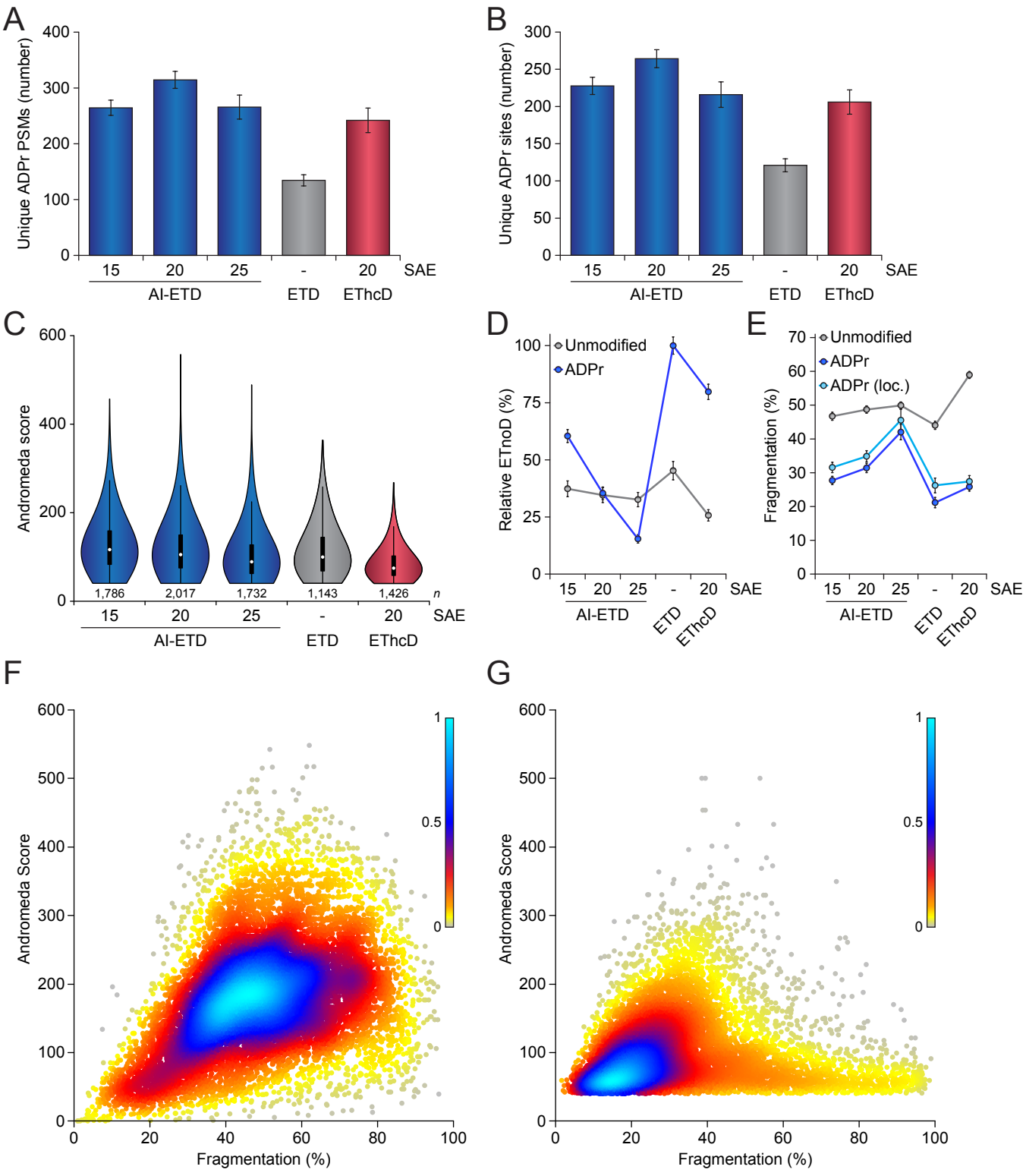


FIGURE S1

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

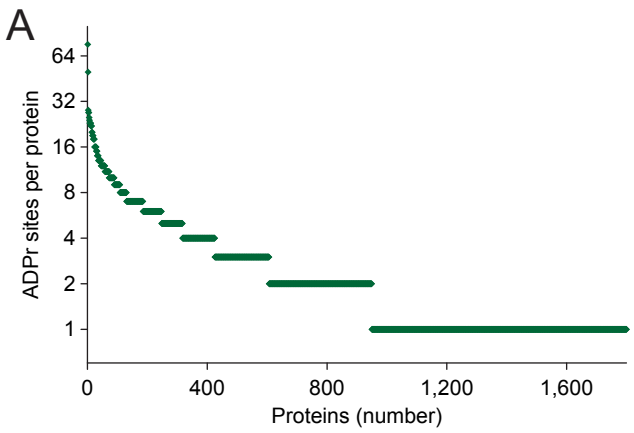


FIGURE S2

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921650; this version posted January 28, 2020. 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.

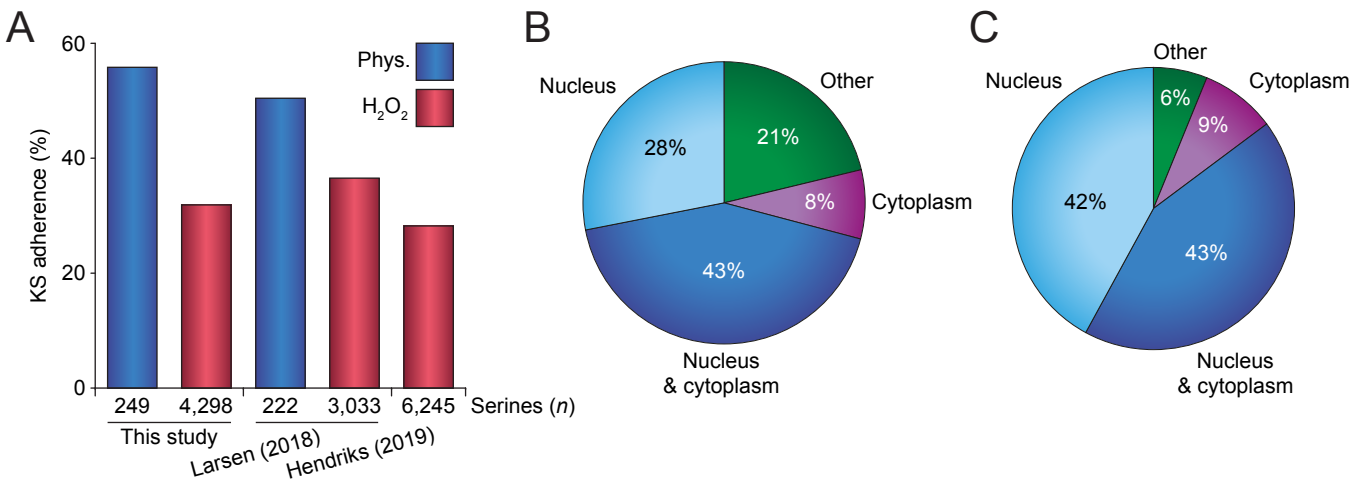


FIGURE S3