Deep protein methylation profiling by combined chemical and immunoaffinity approaches reveals novel PRMT1 targets

Nicolas G. Hartel[1], Brandon Chew[1], Jian Qin[2,3], Jian Xu[2,4,5], Nicholas A. Graham[1,5]*

[1]Mork Family Department of Chemical Engineering and Materials Science, Viterbi School of Engineering; [2]Center for Craniofacial Molecular Biology, Herman Ostrow School of Dentistry; [3]Central Laboratory, Renmin Hospital, Wuhan University, Wuhan, Hubei, China; [4]Department of Biochemistry and Molecular Medicine, [5]Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, CA 90089

*To whom correspondence should be addressed: Nicholas A. Graham, University of Southern California, 3710 McClintock Ave., RTH 509, Los Angeles, CA 90089. Phone: 213-240-0449; E-mail: nagraham@usc.edu

Running title: Deep methylation profiling reveals novel PRMT1 targets

ABBREVIATIONS

ADMA, asymmetric dimethyl arginine; HILIC, hydrophilic interaction chromatography; IAP, immunoaffinity purification; Kme1, monomethyl lysine; Kme2, dimethyl lysine; Kme3, trimethyl lysine; KMT, lysine methyltransferase; LC, liquid chromatography; MS, mass spectrometry; MMA, monomethyl arginine; PTM, post-translational modification; PRMT, protein arginine methyltransferase; SCX, strong cation exchange; SDMA, symmetric dimethyl arginine;

ABSTRACT

Protein methylation has been implicated in many important biological contexts including signaling, metabolism, and transcriptional control. Despite the importance of this post-translational modification, the global analysis of protein methylation by mass spectrometry-based proteomics has not been extensively studied due to the lack of robust, well-characterized techniques for methyl peptide enrichment. Here, to better investigate protein methylation, we optimized and compared two methods for methyl peptide enrichment: immunoaffinity purification (IAP) and high pH strong cation exchange (SCX). Comparison of these methods revealed that they are largely orthogonal for monomethyl arginine (MMA), suggesting that the usage of both techniques is required to provide a global view of protein methylation. Using both IAP and SCX, we investigated changes in protein methylation downstream of protein arginine methyltransferase 1 (PRMT1) and quantified ~1,000 methylation sites on 407 proteins. Of these methylation sites, PRMT1 knockdown resulted in significant changes to 97 arginine methylation sites on 59 proteins. In contrast, zero lysine methylation sites were significantly changed upon PRMT1 knockdown. In PRMT1 knockdown cells, 84 MMA sites were either significantly downregulated or upregulated. PRMT1 knockdown also induced significant changes in both asymmetric dimethyl arginine (ADMA) and symmetric dimethyl arginine (SDMA), suggesting that loss of PRMT1 activity allows scavenging of PRMT1 substrates by other PRMTs. Using neutral loss fragmentation ions unique to ADMA and SDMA, we annotated dimethylarginines as either ADMA or SDMA. Through integrative analysis of methyl forms, we identified 12 high confidence PRMT1 substrates, 43 putative PRMT1 substrates, and 17 methylation sites that are scavenged by other non-PRMT1 arginine methyltransferases in the absence of PRMT1 activity. Taken together, our results suggest that deep protein methylation profiling by mass spectrometry requires orthogonal enrichment techniques to identify novel PRMT1 methylation targets and highlight the dynamic interplay between methyltransferases in mammalian cells.

INTRODUCTION

Protein post-translational modifications (PTMs) regulate diverse biological processes and provide additional complexity to proteins beyond their initial primary sequence [1]. Protein methylation was first identified over 50 years ago on both arginine [2] and lysine [3] residues, yet these PTMs are among the least studied compared to other modifications such as phosphorylation, acetylation, and ubiquitination [4]. Regardless, recent studies have identified protein arginine methylation as an important regulator of signal transduction [5]–[7], metabolism [8], [9, p. 1], cell cycle [10], and transcriptional control [11]–[13].

The mammalian genome encodes nine protein arginine methyltransferases (PRMT) and approximately 50 lysine methyltransferases (KMT). Both PRMTs and KMTs use S-adenosylmethionine (SAM) as a methyl donor to methylate either the guanidino nitrogens of arginine or the ε-amino group of lysine. The complexity of protein methylation is enhanced by the fact that both methyl-arginine and methyl-lysine occur in three distinct forms. Arginine exists in monomethyl (MMA), asymmetric dimethyl (ADMA), or symmetric dimethyl (SDMA) forms, whereas lysine exists in monomethyl (Kme1), dimethyl (Kme2), or trimethyl (Kme3) forms. PRMTs can be divided into two categories based on which type of arginine methylation they catalyze: Type I PRMTs catalyze MMA and ADMA (PRMT1, PRMT3, PRMT4, PRMT6, and PRMT8) [14], whereas Type II catalyze MMA and SDMA (PRMT5, PRMT7, and PRMT9) [15].

One reason why the study of protein methylation has lagged behind other PTMs is a lack of robust methyl peptide enrichment strategies [16]. Compared to strategies for enrichment of phospho-peptides with TiO_2 [17] or IMAC [18] or for enrichment of glycosylated peptides with hydrophilic interaction chromatography (HILIC) [19], techniques for enriching methyl peptides have not been as widely adopted. Enrichment of methyl peptides is more difficult than many other PTMs because methylation does not add significant steric bulk or change the charge of the amino

acids [20]. Despite this difficulty, advances in methyl peptide enrichment for mass spectrometry have been made using immunoaffinity enrichment (IAP) with antibodies that recognize various forms of protein methylation [21]–[25]. By combining IAP against MMA with sample fractionation, as many as 8,000 MMA sites have been identified in human cells [24]. Other enrichment strategies for methyl peptides include high pH strong cation exchange (SCX), chemical labeling, HILIC, and engineered MBT domains that bind methylated proteins [26]–[30]. High pH SCX relies on missed cleavage by trypsin of methylated arginine and lysine residues which results in methyl peptides with higher positive charge that can be enriched by SCX. HILIC, in contrast, enriches methyl peptides based on the highly hydrophilic nature of methyl peptides and can be enhanced by peptide deglycosylation to remove competing glycosylated peptides [28]. Together, these enrichment techniques have begun to shed insight into the global regulation of protein methylation, but there has been no extensive comparison of these methods to present a global picture of protein methylation.

Here, to better study protein methylation, we tested and optimized liquid-chromatography mass-spectrometry (LC-MS) gradients for different methyl peptide enrichment strategies including high pH SCX, IAP, and HILIC. To optimize high pH SCX, we shortened the LC gradient to exploit the hydrophilic nature of methyl peptides, thereby reducing instrument time without compromising detection of methyl-peptides. For IAP methyl peptide enrichment, we optimized the LC gradient to produce a 20% increase in unique methyl peptides. HILIC methods, by contrast, did not result in significant methyl peptide enrichment. Notably, comparison of high pH SCX and IAP revealed that these methods are largely orthogonal, suggesting that both methods are required for global analysis of protein methylation. We then used both optimized methyl proteomics methods in parallel to investigate the PRMT1 methylome. Knockdown of PRMT1 with shRNA led to global increases and decreases of MMA, general decreases in ADMA, and significant increases in SDMA, indicating potential competition and/or substrate scavenging from other PRMTs [31].

Additionally, examination of MS/MS fragments confirmed that ADMA and SDMA peptides can be distinguished by neutral ion loss from the methylated arginine [16], [32]–[34], even without the use of targeted MRM methods. Taken together, our results describe a general method for deep profiling of protein methylation and identify novel potential MMA and ADMA methylation targets of PRMT1.

EXPERIMENTAL PROCEDURES

Cell Culture — LN229 cells and HEK 293T cells expressing short hairpin RNA (shRNA) against PRMT1 or control were grown in DMEM media (Corning) supplemented with 10% FBS (Omega Scientific) and 100 U/mL penicillin/streptomycin (Thermo Scientific). Cells were cultured at 37 ^{o}C in humidified 5% CO_2 atmosphere. Generation of HEK 293T cells stably expressing shRNA against PRMT1 or control were previously described [5]. shPRMT1 and shControl cells were cultured with 4 ug/mL puromycin to maintain selection.

Cell Lysate Preparation — Cells were washed with PBS, scraped, and lysed in 50 mM Tris pH 7.5, 8M urea, 1 mM activated sodium vanadate, 2.5 mM sodium pyrophosphate, 1 mM β-glycerophosphate, and 100 mM sodium phosphate. Protein concentrations were measured by bicinchoninic assay. Lysates were sonicated and cleared by high speed centrifugation and then filtered through 0.22 um filter. Proteins were reduced, alkylated, and quenched with 5 mM dithiothreitol, 25 mM iodoacetamide, 10 mM dithiothreitol, respectively. Lysates were four-fold diluted in 100 mM Tris pH 8.0 and digested with trypsin at a 1:100 ratio and then quenched with addition of trifluoroacetic acid to pH 2. Peptides were purified using reverse-phase Sep-Pak C18 cartridges (Waters) and eluted with 30% acetonitrile, 0.1% TFA and then dried by vacuum. Dried peptides were subjected to high pH strong cation exchange or antibody immunoaffinity purification.

Immunoblot Analysis — Cells were lysed in modified RIPA buffer (50 mM Tris–HCl (pH 7.5), 150 NaCl, 50 mM β-glycerophosphate, 0.5 mM NP-40, 0.25% sodium deoxycholate, 10 mM sodium pyrophosphate, 30 mM sodium fluoride, 2 mM EDTA, 1 mM activated sodium vanadate, 20 μg/ml aprotinin, 10 μg/ml leupeptin, 1 mM DTT, and 1 mM phenylmethylsulfonyl fluoride). Whole-cell lysates were resolved by SDS–PAGE on 4–15% gradient gels and blotted onto nitrocellulose membranes (Bio-Rad). Membranes were blocked for 1 h in nonfat-milk, and then incubated with primary and secondary antibodies overnight and for 2 h, respectively. Blots were imaged using

the Odyssey Infrared Imaging System (Li-Cor). Primary antibodies used for Western blot analysis were: mono-methyl arginine (8015, Cell Signaling), asymmetric di-methyl arginine motif (13522, Cell Signaling), symmetric di-methyl arginine motif (13222, Cell Signaling), PRMT1 (2449, Cell Signaling), and anti-β-actin (10081-976, Proteintech).

High pH Strong Cation Exchange (SCX) — As described previously [27], in brief, 1 mg of digested protein was resuspended in loading buffer (60% acetonitrile, 40% BRUB (5 mM phosphoric acid, 5 mM boric acid, 5 mM acetic acid, pH 2.5) and incubated with high pH SCX beads (Sepax) for 30 minutes, washed with washing buffer (80% acetonitrile, 20% BRUB, pH 9), and eluted into five fractions using elution buffer 1 (60% acetonitrile, 40% BRUB, pH 9), elution buffer 2 (60% acetonitrile, 40% BRUB, pH 10), elution buffer 3 (60% acetonitrile, 40% BRUB, pH 11), elution buffer 4 (30% acetonitrile, 70% BRUB, pH 12), and elution buffer 5 (100% BRUB, 1M NaCl, pH 12). Eluates were dried, resuspended in 1% trifluoroacetic acid and desalted on STAGE tips [35] with 2 mg of HLB material (Waters) loaded onto 300 uL tip with a C8 plug (Empore, Sigma).

Immunoaffinity Purification (IAP) — 10 mg of digested proteins were dissolved in 1X immunoprecipitation buffer (50 mM MOPS, 10 mM Na_2HPO_4, 50 mM NaCl, pH 7.2, Cell Signaling). Modified symmetric dimethyl arginine peptides, asymmetric dimethyl arginine peptides, and monomethyl arginine peptides were immunoprecipitated by addition of 40 uL of PTMScan Symmetric Di-Methyl Arginine Motif Kit (13563, Cell Signaling), PTMScan Asymmetric Di-Methyl Arginine Motif Kit (13474, Cell Signaling), and PTMScan Mono-Methyl Arginine Motif Kit (12235, Cell Signaling), respectively. Modified methyl lysine peptides were enriched with PTMScan Pan-Methyl Lysine Kit (14809). Lysates were incubated for with PTMScan motif kits for 2 hours at 4 °C on a rotator. Beads were centrifuged and washed two times in 1X immunoprecipitation buffer followed by three washes in water, and modified peptides were eluted with 2 x 50 uL of 0.15% TFA and desalted on STAGE tips with C18 cores (Empore, Sigma). Enriched peptides were resuspended in 50 mM ammonium bicarbonate (Sigma) and digested

with trypsin for 2 hours, acidified with trifluoroacetic acid to pH 2 and desalted on STAGE tips.

Mass Spectrometric Analysis — All LC-MS experiments were performed on a nanoscale UHPLC system (EASY-nLC1200, Thermo Scientific) connected to an Q Exactive Plus hybrid quadrupole-Orbitrap mass spectrometer equipped with a nanoelectrospray source (Thermo Scientific). Peptides were separated by a reversed-phase analytical column (PepMap RSLC C18, 2 µm, 100 Å, 75 µm X 25 cm) (Thermo Scientific). For high pH SCX fractions a "Short" gradient was used where flow rate was set to 300 nl/min at a gradient starting with 0% buffer B (0.1% FA, 80% acetonitrile) to 29% B in 142 minutes, then washed by 90% B in 10 minutes, and held at 90% B for 3. The maximum pressure was set to 1,180 bar and column temperature was constant at 50 ^{o}C. For IAP samples a "Slow" gradient was used where flow rate was set to 300 nl/min at a gradient starting with 0% buffer B to 25% B in 132 minutes, then washed by 90% B in 10 minutes. The effluent from the HPLC was directly electrosprayed into the mass spectrometer. Peptides separated by the column were ionized at 2.0 kV in the positive ion mode. MS1 survey scans for DDA were acquired at resolution of 70k from 350 to 1800 m/z, with maximum injection time of 100 ms and AGC target of 1e6. MS/MS fragmentation of the 10 most abundant ions were analyzed at a resolution of 17.5k, AGC target 5e4, maximum injection time 120 ms for IAP samples, 240 ms for SCX samples, and normalized collision energy 26. Dynamic exclusion was set to 30 s and ions with charge 1 and >6 were excluded. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository with the dataset identifier PXD012357 (Username: reviewer93189@ebi.ac.uk, password: mnoFDcyY) [36].

Identification and Quantitation of Peptides — MS/MS fragmentation spectra were searched with Proteome Discoverer SEQUEST (version 2.2, Thermo Scientific) against the in-silico tryptic digested Uniprot *H. sapiens* database with all reviewed with isoforms (release Jun 2017, 42,140 entries). The maximum missed cleavage rate was set to 5 [27]. Trypsin was set to cleave at R

and K. Dynamic modifications were set to include mono-methylation of arginine or lysine (R/K, +14.01565), di-methylation of arginine or lysine (R/K, +28.0313), tri-methylation of lysine (K, +42.04695), oxidation on methionine (M, +15.995 Da, and acetylation on protein N-terminus (+42.011 Da). Fixed modification was set to carbamidomethylation on cysteine residues (C, +57.021 Da). The maximum parental mass error was set to 10 ppm and the MS/MS mass tolerance was set to 0.02 Da. Peptides with sequence of six to fifty amino acids were considered. Methylation site localization was determined by ptm-RS node in Proteome Discoverer, and only sites with localization probability greater or equal to 75% were considered. The False Discovery Rate threshold was set strictly to 0.01 using Percolator node validated by q-value. Relative abundances of parental peptides were calculated by integration of area-under-the-curve of the MS1 peaks using Minora LFQ node in Proteome Discoverer 2.2. Only spectra with an Xcorr > 1.5 were used for quantification of modified arginine peptides. The Proteome Discoverer export peptide groups abundance values were $\log_2$ transformed, normalized to the corresponding samples median values, and significance was determined using a permutation-based FDR approach in the Perseus environment [37] (release 1.6.2.3) with a q-value FDR of 0.05 and S_0 value of 0.5.

Motif Analysis — Motifs were analyzed by MotifX [38] and MOMO from MEME suite [39] to detect statistically significant patterns in methylation sequence data. Two sample motif analysis was performed using Two Sample Logo [40].

EXPERIMENTAL DESIGN AND STATISTICAL RATIONALE

Samples — Five LN229 samples were analyzed, four were enriched by high pH SCX enrichment and 1 enriched by mono-methyl arginine immunoaffinity purification. Two SCX samples were each run on the “Long” and “Short” SCX gradients, and the “Short” gradient SCX samples were

compared to the single mono-methyl arginine IAP sample. The single mono-methyl arginine IAP sample was injected in two equal amounts on a “Standard” and “Slow” gradient. Four shPRMT1 293T samples and four shControl 293T samples were analyzed and compared. Two of the four 293T samples were enriched by high pH SCX and another two were enriched by sequential immunoaffinity purification incubated with SDMA, ADMA, MMA, and Pan-K PTMScan Kits sequentially. 50ug of whole cell lysates from each sample were collected for immunoblots before SCX or IAP enrichment.

Replicates — LN229 samples had two biological replicates tested on each of the high pH SCX gradients (“Long” and “Short”). The LN229 mono-methyl arginine IAP sample had no replicates; however, equal volumes from the single sample were injected and run on the “Standard” and “Slow” gradients. 293T cells had two biological replicates each for high pH SCX enrichment and sequential IAP enrichment.

Controls and Randomization — 293T cells expressing shRNA against PRMT1 were controlled by using 293T expressing scrambled non-targeting short hairpin control RNA to account for biases introduced by stable transfection. LN229 and 293T samples were also randomized by sample prior to LC-MS injection.

Rationale — Using only two replicates for 293T shPRMT1 and 293T shControl is justified because of a low coefficient of variance for quantified peptides (85% of MMA IAP peptides below 60% CV for shControl and 78% of MMA IAP peptides below 60% CV for shPRMT1). A third of each dataset is also below 20% CV which indicates small variability between biological replicates. Two replicates for the 293T SCX replicates is also justified due to fractions E1, E2, and E3 having median CV’s of 21, 26, and 54 respectively for shControl samples and 32, 66, and 70 for the shPRMT1 samples. Fractions E1,2, and 3 contributed to 95% of quantified methyl peptides for high pH SCX. Sample abundances from Minora Label Free Quantitation through Proteome Discoverer 2.2 were $\log_2$ transformed, median normalized by sample, then subject to a

permutation-based FDR approach in Perseus software with a q-value FDR of 0.05 and S_0 of 0.5. All datasets showed a normal distribution of abundances after log2 transformation. Statistical significance was given to differential peptide pairs with q-value < 0.05 and required confident methyl site localization of 75% or more through ptm-RS node on Proteome Discoverer. An XCorr > 1.5 was required for quantified methyl peptides. Missing values across samples were excluded and only peptides confidently quantified in all replicates were considered for the permutation-based FDR calculation. For SCX methyl peptides identified in multiple fractions, the largest raw abundance was used to determine which fraction to use to calculate peptide abundance. Taken together, the low variability of biological replicates and rigorous statistical thresholds (peptide discovery FDR = 1% and permutation FDR = 5%) allow confident differential analysis of methyl peptides.

RESULTS

Optimization of LC gradients to improve methyl peptide identification

We first tested an antibody-free protocol that enriches methyl peptides using high-pH strong cation exchange (SCX) on LN229 glioblastoma cells. (Fig. 1A) [27]. High pH SCX exploits the tendency of methylation to induce missed trypsin cleavages, resulting in higher charge states for methyl peptides. After loading and washing tryptic peptides at high pH, five SCX fractions of increasing pH were collected for LC-MS analysis. Using the LC gradient originally employed for high-pH SCX methyl-proteomics, we observed that 70% of methyl peptides eluted early in the LC gradient when the percentage of ACN is below 12% (Fig. 1B, Supp. Table 1). In fact, 5% of methyl peptides eluted during the sample loading phase (2% ACN) when MS spectra are often not collected. The number of identified methyl peptides dropped off quickly after 150 minutes with very few identifications in the last 40 minutes of the gradient. In contrast, non-methyl peptides were more evenly distributed across the entire LC gradient. These results are consistent the known hydrophilic character of methyl-arginine and methyl-lysine peptides [26].

Because methyl peptides eluted early in the LC gradient, we sought to optimize the LC gradient to reduce instrument time without sacrificing the number of unique methyl peptide identifications. We thus tested a “short” LC gradient that i) starts at 0% ACN to prevent elution of methyl peptides during sample loading step; and ii) reduces the overall LC gradient time from 225 to 155 minutes by slightly increasing the ramp rate of ACN (Fig. 1C, Supp. Fig. 1). Using this short LC gradient, we found that the elution profile of methyl peptides was more evenly distributed across time compared to the long gradient (Fig. 1D, Supp. Fig. 2). The number of unique methyl peptides identified by each gradient was not significantly different (Fig. 1E). Taken together, these results demonstrate that the LC gradient can be shortened by 70 minutes while maintaining the sensitivity of the longer gradient for high-pH SCX methyl proteomics.

Next, we tested an antibody-based approach for enrichment of monomethyl arginine (MMA) peptides [21] (Fig. 2A). With a standard 130 minute LC gradient (3% B to 38% B over 110 minutes plus column washing at high ACN) [41], we found that MMA peptides eluted earlier than non-methyl peptides (Fig. 2B, Supp. Fig. 3, Supp. Table 2). These results are again consistent with the known hydrophilic character of methyl-arginine peptides [26]. We therefore reasoned that an LC gradient with a slower ACN ramp might improve the number of unique methyl peptides. We tested a "slow" gradient with an ACN ramp of 0.189 %B/min over 142 minutes against the standard 130-minute gradient which has a ramp of 0.32 %B/min (Fig. 2C). Indeed, the elution profile of MMA methyl peptides identified using the slow gradient was more normally distributed across the LC gradient than the standard gradient (Fig. 2D). In addition, the slow gradient increased the number of unique methyl peptides identified by 20% compared to the standard gradient (Fig. 2E). Taken together, these results demonstrate that optimization of the LC gradient can increase the number of methyl peptides identified by MMA IAP.

SCX and IAP target different subsets of protein methylome

Next, we tested whether SCX and IAP methyl peptide enrichment protocols identify the same or different components of the protein methylome. Lysates from LN229 glioblastoma cells were enriched for methyl peptides using either SCX or MMA IP, and the samples were run on the optimized LC gradients. Although SCX identifies all types of protein methylation (MMA, SDMA, ADMA, Kme1/2/3), we considered only mono-methyl arginine containing peptides to maintain consistency with the MMA IAP enrichment. Comparing the unique methyl peptides identified by each method, we found low overlap of mono-methyl arginine peptides identified by SCX and IAP (22%). 40% of MMA peptides were unique to SCX, and 38% of methyl peptides were unique to IAP (Fig. 3A). Gene ontology analysis of methyl peptides identified by SCX and IAP demonstrated that both techniques enriched for RNA and DNA binding proteins involved in transcription, translation, and alternative splicing in agreement with known distributions of methyl proteins

(Supp. Fig. 4) [22]. This data suggests that these enrichment techniques (SCX and IAP) target different MMA peptides and that the usage of both methods is required to achieve a more complete coverage of the protein arginine methylome.

To further characterize the differences between SCX and IAP-enriched methyl peptides, we compared the number of methyl arginine sites per peptide for each enrichment technique. This analysis revealed that 24% of methyl peptides identified by SCX featured only one methyl arginine site (Fig. 3B), whereas 41% of SCX methyl peptides had three or more methyl-arginine sites. In contrast, 59% of methyl peptides identified by IAP had one methyl site, and only 15% of identified peptides featured 3 or more methylation sites. Therefore, SCX selects preferentially for multi-methylated peptides.

Next, we conducted a motif analysis of SCX and IAP mono-methylarginine peptides [39]. Comparison against the human proteome revealed that IAP enriched peptides exhibit strong enrichment for the GRGG motif, a motif known as GAR (glycine-arginine rich) that is associated with arginine methylation (Fig. 3C) [42]. SCX peptides were also enriched for GAR motifs, however a notable difference was the presence of multiple non-central arginines +/- 2 residues from the central methyl arginine. SCX peptides were also enriched for the RP motif, known as proline-rich arginine motifs (PRAM) [26], while IAP peptides were enriched for RxxxxP where x is any amino acid.

Because SCX relies on missed trypsin cleavages to generate highly charged peptides and arginine methylation is reported to cause missed trypsin cleavage, we next calculated the frequency of MMA peptides identified by SCX and IAP that contained a mono-methyl arginine on the peptide C-terminus. Whereas nearly half of IAP methyl peptides contained a terminal mono-methyl R, only ~20% of SCX peptides terminated in MMA (Fig. 3D). This low percentage of

terminal mono-methyl R for SCX is consistent with successful trypsin cleavage of monomethyl arginine [24], thereby reducing the positive charge and affinity for the SCX media of the peptide. As a result, terminal mono-methyl R peptides are missed in SCX enrichment but not IAP.

Finally, we used a two-sample motif analysis to compare between MMA peptides enriched by SCX and IAP [40]. This analysis demonstrated that SCX enriched MMA peptides with prolines in the +1 position, recapitulating the PRAM sequence (Fig. 3E). In contrast, MMA IAP peptides were highly enriched for GAR motifs compared with SCX peptides. Taken together, these results demonstrate that SCX and IAP enrich for different subsets of protein arginine mono-methylation, suggesting that comprehensive analysis of protein methylation requires both techniques.

Identification of significantly changing MMA sites using combined SCX and IAP enrichment

PRMT1 has been reported to account for over 90% of ADMA methylation events in mammalian cells [43]. We thus sought to investigate the effects of PRMT1 depletion on the protein methylome using HEK 293T cells expressing a short hairpin RNA (shRNA) against PRMT1 or a control shRNA [5]. Western blotting confirmed PRMT1 knockdown and an upregulation of MMA levels in shPRMT1 cells (Fig. 4A), as reported previously [31]. To define PRMT1 methyl targets, cells expressing shPRMT1 or the shControl were subjected to both SCX and MMA IAP methyl peptide enrichment followed by mass spectrometry (Fig. 4B). As with LN229 cells, the overlap between MMA peptides identified by SCX and IAP was only 11%, justifying the usage of both enrichment techniques (Supp. Fig. 5A). Label-free quantitation revealed 87 significantly changing MMA peptides (q-value < 0.05), 24 of which were decreased and 63 of which were increased in MMA (Fig. 4C, Supp. Tables 3-4). Notably, both IAP and SCX identified significantly changing methylation sites, although IAP identified more significantly-changing sites than SCX (n = 80 for IAP, n = 7 for SCX, 84 unique sites). Peptides with decreased MMA levels represent putative

PRMT1 methylation targets. Because PRMT1 also catalyzes ADMA methylation, increased MMA levels could represent PRMT1 ADMA targets for which other PRMTs can catalyze MMA (ie, loss of PRMT1-mediated ADMA increases observed MMA levels). Indeed, significantly changing MMA sites were enriched for known PRMT1 interactors from the IntAct EBI database [44], regardless of the direction of change ($p=0.013$ by Fisher's Exact Test).

Two-sample motif analysis of the significantly upregulated MMA peptides revealed a strong preference for glycine in many positions, consistent with known GAR motifs that are associated with arginine methylation [42] (Fig. 4D). Additionally, significantly upregulated MMA peptides demonstrated a preference for serine in the -2 position and tyrosine in the +3 position. Interestingly, PRMT1 is known to recognize and methylate arginines followed by serine at +2 and tyrosine at +1 position [45]. Two-sample motif analysis of the significantly downregulated MMA peptides revealed a strong preference for glycine in the +1, +2, and +4 positions but not other positions. Similar to the upregulated MMA motif, the significantly downregulated MMA peptides were enriched for serine in the -2 position and tyrosine in the +3 position (Fig. 4D). Relative to the upregulated and downregulated MMA peptides, non-significantly changing MMA peptides were enriched for lysine at the +1 position and negatively-charged glutamic acid at the +4 position.

Notably, the most significantly upregulated MMA site was R455 of the RNA-binding protein EWSR1, which increased 690-fold (9.43 on a $\log_2$ scale, Fig. 4F). Two other EWSR1 methyl peptides, R464/R471 and R615 were also significantly upregulated in shPRMT1 cells (17- and 9-fold, respectively). This is consistent with observations that EWSR1 is highly methylated [46] and interacts with PRMT1 [47], although this data suggests that EWSR1 is also a target of other non-PRMT1 arginine methyltransferases. Additionally, EWSR1 R455 and R615 were observed in our dimethylarginine data sets (see below), but the dimethylation of these arginines was not significantly changed. Thus, our data suggests that PRMT1 knockdown induces large increases

in EWSR1 monomethylation but not dimethylation at these sites. Other MMA peptides that were significantly upregulated included OTUD4 R1061, HNRNPA1 R206, KHDRBS1 R304/R310, and BPTF R92. The arginine residues on these proteins have been previously reported to be both monomethylated [21], [22], [25], but none have been previously linked to PRMT1.

Examination of the significantly downregulated MMA sites revealed that a previously unreported methylation site on the cytoskeletal protein FLNA/FLNB R2242/R2228 was downregulated by 16-fold in PRMT1 knockdown cells (Fig. 4G). Other downregulated MMA sites included RBM33 R1028, DDX42 R12, SNRNP70 R201, and MAP3K20 R670. To our knowledge, none of these proteins have been previously reported to interact with PRMT1 and are thus putative novel PRMT1 targets. Interestingly, the identified arginine residues on DDX42, SNRNP70, and MAP3K20 have all been reported to be MMA modified, but to our knowledge, these residues have not been reported to be dimethylated. In our dimethylarginine analysis (see below), we failed to identify dimethylation at these sites, although we did identify decreased dimethylation of MAP3K20 on an adjacent methyl peptide containing R683/R693. Taken together, our data suggests that FLNA/FLNB, RBM33, DDX42, and SNRNP70 are PRMT1 MMA but not ADMA targets, whereas MAP3K20 may be a PRMT1 MMA and ADMA target.

Identification of significantly changing dimethylarginines using combined SCX and IAP enrichment

Next, we sought to analyze the effects of PRMT1 depletion on protein arginine dimethylation using HEK 293T cells expressing either PRMT1 or control shRNA. Western blotting with an antibody that recognizes ADMA showed a slight decrease in ADMA methylation upon knockdown of PRMT1 (Fig. 5A). In contrast, Western blotting with an antibody against SDMA showed increased SDMA levels at several molecular weights in PRMT1 knockdown cells, suggesting that Type II PRMTs are more active in cells with reduced PRMT1 activity [31]. To

identify proteins with changing dimethylarginines, cells expressing shPRMT1 or the shControl were subjected to SCX, ADMA IAP, and SDMA IAP methyl peptide enrichment followed by mass spectrometry (Fig. 5B). Like MMA, the overlap between dimethylarginine-containing peptides identified by SCX and ADMA IAP was low (24%), and the overlap between dimethylarginine-containing peptides identified by SCX and SDMA IAP was also low (19%), justifying the usage of both enrichment techniques (Supp. Fig. 5B-C). Because SCX enriches both ADMA and SDMA peptides, these numbers may overestimate the degree of overlap between the techniques.

Label-free quantitation of peptides enriched by ADMA IAP revealed 4 significantly downregulated methyl peptides (q-value < 0.05) including XRN2 R946, NEFL R23, LSG1 R14, and YBX3 R220 (Fig. 5C, Supp. Table S5). This downregulation suggests that these proteins are PRMT1 ADMA targets, but to our knowledge, none of these proteins have been previously linked to PRMT1. Known PRMT1 interactors ILF3 R609, WDR33 R1315, and YLPM1 R594/R599/R601 also exhibited downregulation of ADMA, although changes were not statistically significant. Interestingly, ADMA IAP also revealed that PRMT1 knockdown induced significant upregulation of 6 arginine methylation peptides, three of which were dimethylated at multiple arginines, on proteins including DHX9, HRNPA1, HNRNPA3, HNRNPU, and HNRNPUL1. The upregulation of these peptides may reflect increased activity of an alternative Type I PRMT or increased protein abundance in cells with reduced PRMT1 activity, although we cannot exclude the possibility that the ADMA antibody binds to and enriches SDMA methyl peptides.

Quantitation of methyl peptides enriched by SDMA IAP revealed 6 significantly upregulated (q-value < 0.05) methyl arginine peptides on the RNA binding proteins including DHX9, HNRNPA1, HNRNPA3, HNRNPH3, and KHDRBS1 (Fig. 5D, Supp. Table S6). Increased SDMA levels suggest that these methylation sites may become accessible to Type II PRMTs upon PRMT1 knockdown [31]. Notably, DHX9 R1249/R1253/R1265, HNRNPA1 R206, and HNRNPA3

R246 were also identified as significantly upregulated in our ADMA IAP data (Fig. 5C), suggesting that these proteins either undergo ADMA and SDMA methylation at the same residues or that the ADMA and SDMA antibodies cross-react with the unintended methyl-arginine form (see below for analysis of characteristic neutral ion loss). Amongst 84 methyl peptides identified by SDMA IAP, only one was significantly downregulated in PRMT1 knockdown cells, namely the novel dimethylation sites R281/R283 on the RNA-binding protein PRRC2C. Analysis of dimethylarginine peptides enriched by SCX revealed 3 significantly upregulated dimethyl peptides (Fig. 5E, Supp. Table S7) including the previously unreported sites HNRNPR R543/R546/R554. SCX also enriched peptides with both mono- and dimethyl marks, creating "mixed" methyl peptides, including three significantly upregulated methyl arginine peptides (Supp. Fig. 6, Supp. Table S8). Strikingly, methyl peptides that were significantly upregulated across all enrichment techniques were enriched for heterogeneous ribonuclear proteins (HNRNPs) (Fisher's exact p-value < 1 x 10^{-5}), which regulate many aspects of RNA metabolism including splicing, export, localization, stability and translation [48].

Two sample motif analysis using all identified ADMA to SDMA peptides from IAP also revealed interesting differences between the two groups. ADMA IAP peptides showed enrichment for the PRAM sequence (prolines in the -6, -5, -3, +2, and +4 positions) and also leucine at the +1 position which is a preferred PRMT1 substrate motif [45]. In contrast, SDMA peptides showed enrichment for the RXR motif that was also identified in SCX MMA sites along with general GAR motif enrichment (Fig. 5F).

Characteristic neutral losses of methylamine or dimethylamine allows discrimination of SDMA and ADMA

One limitation of methylarginine proteomics is the inability to independently confirm the dimethylarginine isomer (ie, whether a site is ADMA or SDMA) because these PTMs are identical

in mass. However, distinguishing ADMA and SDMA should be possible based on characteristic neutral ions losses from both dimethylarginine forms [32], [49]–[51]. Following the neutral loss of methylamine, SDMA produces an ion with 172.108 m/z, whereas the neutral loss of dimethylamine from ADMA results in an ion with 158.092 m/z. To test whether we could distinguish ADMA and SDMA based on these neutral losses, we calculated the ratio of intensities of (172.108 / 158.092) m/z for dimethylated peptides in all MS2 spectra acquired by ADMA and SDMA IAP. We did not include mixed methyl peptides which include MMA since neutral loss of methylamine from MMA produces the same ion as neutral loss of dimethylamine from ADMA. Comparing peptides identified using ADMA and SDMA antibodies revealed statistically significant differences in the 172/158 ratio with SDMA IAP peptides showing a higher 172/158 ratio than ADMA IAP peptides (Supp. Fig. 7). This suggests that the ADMA and SDMA antibodies are generally selective for their intended dimethylarginine form. In addition, our data suggests that we can discriminate methyl peptides as probable ADMA or SDMA sites with low or high 172/158 ratios, respectively (Supp. Tables 5-7).

We therefore used this approach to test whether significantly changing ($q < 0.05$) dimethylarginine peptides identified by IAP and SCX were likely to be either ADMA or SDMA. For significantly downregulated methyl peptides identified by ADMA IAP, three were conclusively ADMA (XRN2 R946, NEFL R23, and YBX3 R220), and one methyl peptide (LSG1 R14) was inconclusive (Supp. Fig. 8). For significantly upregulated methyl peptides identified by ADMA IAP, four of six were likely to be ADMA (HNRNPA3 R246, HNRNPU R707/709/715, HNRNPU R739, and HNRNPUL1 R645/656), and two were likely to be SDMA (HNRNPA1 R206 and DHX9 R1249/1253/1265). Note that we cannot exclude the possibility that multi-methyl peptides such as DHX9 R1249/1253/1265 contain both ADMA and SDMA. Examining methyl peptides enriched by SDMA IAP, five of six significantly upregulated methyl peptides were likely to contain SDMA. Only the site HNRNPA3 R246 (which was also identified in ADMA IAP) was likely to contain

ADMA. Notably, the methyl peptide PRRC2C R281/R283 which was significantly downregulated in PRMT1 knockdown cells was inconclusive (ie, did not have measurable 172 or 158 ions), but the singly methylated peptide PRRC2C R281 (identified in both IAP ADMA and IAP SDMA) was likely to contain ADMA based on its 172/158 ratio. For SCX, which identified only three significantly changing methyl peptides, all three were likely to contain SDMA based on 172/158 ratios. Thus, ADMA and SDMA IAP are broadly but not perfectly selective for their intended methyl forms.

Notably, these characteristic neutral losses enabled us to classify methyl peptides that had been identified in both ADMA and SDMA IAP data sets. Specifically, DHX9 R1249/R1253/R1265, HNRNPA1 R206, and HNRNPA3 R246 were identified as significantly upregulated in PRMT1 knockdown cells by both ADMA and SDMA IAP (Fig. 5C-D). Based on its large 172/158 ratio, DHX9 is likely to be SDMA but not ADMA modified at R1249/R1253/R1265 (Supp. Fig. 8). Although this peptide has been previously reported as an ADMA modification [21], the large 172/158 ratio suggests that this peptide is SDMA modified in 293T cells. Note that the presence of three dimethylation sites means that we cannot exclude mixed ADMA and SDMA methylation on R1249/R1253/R1265 of DHX9. Similarly, based on large 172/158 ratio, HNRNPA1 R206 likely is SDMA but not ADMA modified, whereas HNRNPA3 R246 is likely ADMA but not SDMA modified. Thus, these peptides may exist in only one form but exhibit cross-reactivity to both the ADMA and SDMA IAP antibodies. Together, this analysis demonstrates an approach to determine the identity of dimethylated arginines as SDMA or ADMA even though they are identical in mass.

Lysine methylation is not affected by PRMT1 knockdown

In addition to arginine methylation, both SCX and IAP can enrich peptides containing methyl lysine. To test whether PRMT1 depletion affected protein lysine methylation, 293T cells

expressing shPRMT1 or the shControl were subjected to both SCX and Pan-methyl-K IAP followed by mass spectrometry. Label-free quantitation of mono-, di-, and tri-methyl lysine by IAP showed no significant changes in protein lysine methylation (Supp. Fig. 9A, Supp. Table 9). Similarly, no significant changes in methyl-lysine-containing peptides were identified by SCX enrichment (Supp. Fig. 9B, Supp. Table 10). Together, these data demonstrate that PRMT1 depletion affects methylation of arginine but not lysine residues.

Integrated analysis of methyl-arginine forms reveals novel PRMT1 substrates and substrate scavenging by other PRMTs

Next, we sought to integrate data from our five types of methylarginine enrichment techniques (MMA IAP and SCX, ADMA and SDMA IAP, DMA SCX) to identify PRMT1 substrates (Supp. Table 11). Based on significant decreases in ADMA IAP levels and low 172/158 ratios, the methyl peptides YBX3 R220, NEFL R23, and XRN2 R946 appear to be *bona fide* PRMT1 substrates (Fig. 6A). XRN R946 was also reduced in MMA levels, suggesting that PRMT1 mediates both MMA and ADMA on this substrate. Nine other methyl peptides exhibited significant upregulation of MMA accompanied by downregulated dimethylation (eg, ADMA IAP or DMA SCX with ADMA-like 172/158 ratios). The fold change in MMA was consistent for peptides identified by both IAP and SCX. The combination of decreased dimethylation and increased monomethylation in cells with reduced PRMT1 activity suggests that other PRMTs catalyze MMA but not ADMA on these residues, such that MMA accumulates in PRMT1 knockdown cells. We suggest that these methyl peptides are high confidence PRMT1 substrates. In addition, we identified 24 methyl peptides with significantly decreasing MMA (Fig. 6B) and 19 with significantly increasing MMA (Fig. 6C) but without corresponding dimethylation data. We suggest that these methyl peptides are putative PRMT1 targets but that they do not meet the threshold for high confidence PRMT1 substrates because we cannot rule out that changes in methylation state are driven by changes in total protein levels.

In the absence of PRMT1 activity, other PRMTs may have access to substrates that were formerly blocked by PRMT1. Integrating mono- and dimethylation data revealed ten methyl peptides with increased levels of both MMA and SDMA (Fig. 6D). All these methylation sites were likely to SDMA based on their 172/158 ratio of characteristic neutral ion losses. In addition, we identified seven methyl peptides with increasing MMA and ADMA levels and 172/158 ratios indicative of ADMA (Fig. 6E). This data suggests that reduced PRMT1 activity enabled these substrates to be either SDMA or ADMA methylated by other, non-PRMT1 arginine methyltransferases, as reported elsewhere [31]. Interestingly, this analysis identified R128 of the known PRMT1-interactor protein chromatin target of PRMT1 (CHTOP) as an MMA and ADMA site that is upregulated upon PRMT1 knockdown [52], [53]. Taken together, we identified 12 high confidence PRMT1 substrates, 43 putative PRMT1 substrates, and 17 methyl peptides that are scavenged by other PRMTs in cells with reduced PRMT1 activity.

DISCUSSION

Despite its relevance for signal transduction, metabolism, transcription, and other cellular phenotypes, protein arginine and lysine methylation remains understudied. This is partly because of the inherent difficulty of enriching a small, neutral PTM and partly because methyl peptide enrichment strategies have been less comprehensively studied than other PTMs. In this study, we demonstrate the optimization and comparison of two methyl peptide enrichment techniques: high pH SCX and IAP. By exploiting the known hydrophilic characteristics of methyl peptides, we were able to reduce instrument time while maintaining sensitivity (high pH SCX, Fig. 1) and to increase the number of unique methyl peptides by ~20% (IAP, Fig. 2). We found that the overlap in methyl peptides identified by these two techniques is low (10-24%, Fig. 3A and Supp. Fig. 5), demonstrating that comprehensive measurement of the protein methylome requires both SCX and IAP enrichment strategies.

Using these orthogonal methyl peptide enrichment strategies, we investigated the PRMT1-dependent protein methylome. In total, we found that PRMT1 knockdown in 293T cells resulted in significant changes to 97 methylarginine peptides ($q < 0.05$) on 59 proteins (Fig. 4-5). Fewer than 10% of these significantly changing proteins are known to interact with PRMT1 (EBI database). Many of the significantly changing methyl proteins were annotated in the UniProt database as RNA binding proteins, consistent with the known function of protein arginine methylation [54]–[58]. Notably, we observed that PRMT1 knockdown did not affect protein lysine methylation (Supp. Fig. 9), although our data do support previous observations that lysine methylation is much less abundant *in vivo* than protein arginine methylation [21], [27]. Finally, through integration of mono- and dimethylation data with characteristic neutral losses from ADMA and SDMA, we identified high confidence and putative PRMT1 methylation substrates as well as PRMT substrate scavenging (Fig. 6).

Our findings validate the utility of published methods using SCX and IAP for enrichment of methyl peptides prior to LC-MS [21], [27]. We have extended these findings by demonstrating the usage of characteristic neutral losses to discriminate between ADMA and SDMA modifications. Differentiating between ADMA and SDMA is particularly important for SCX, which enriches methyl peptides based on missed trypsin cleavages. Although these characteristic ions have been used to discriminate free ADMA and SDMA in serum [32], here we used these characteristic ions to distinguish ADMA and SDMA in shotgun proteomic samples. Detection of 172.108 and 158.092 m/z ions for SDMA and ADMA, respectively, was required because detection of methylamine (31.042 Da) and dimethylamine (45.058 Da) is prohibited by the mass range of our MS system. Notably, the intensity of these ions is very low so that many dimethylarginine peptides could not be assigned as ADMA or SDMA in our data. In addition, we have extended the application of IAP by using commercially available anti-SDMA antibodies for methyl proteomics. Although the usage of this same suite of antibodies for MMA, ADMA, and Kme1/2/3 has been reported [21], this is to our knowledge the first reported usage of the anti-SDMA antibodies for MS-based proteomics. In addition, by using the 172/158 ratio, we were able to confirm the general specificity of these antibodies for their intended methyl arginine form, although we did note examples of likely SDMA peptides binding to anti-ADMA antibodies and vice versa.

Our data also highlight that PRMT1 is likely to regulate both RNA:protein interactions and the protein subcellular localization. Notably, we observed that significantly changing methyl peptides were significantly enriched for hnRNP, with 28 of 97 methyl peptides mapping to 12 different hnRNP family members. This is consistent with reports that arginine methylation affects the function of hnRNP [24], [54, p. 5], [56]–[59], including that PRMT1 knockdown increases the RNA binding function of HNRNPUL1 [24]. In addition, arginine methylation can affect protein subcellular localization, including the nucleo-cytoplasmic shuttling of RNA binding proteins

[60],[61]. In our data, PRMT1 knockdown decreased MMA of DHX9 R1160, a residue which is known to regulate the nuclear localization of DHX9 [62]. Interestingly, SDMA levels of DHX9 on the neighboring residues R1249/R1253/R1265 increased in PRMT1 knockdown cells, suggesting that DHX9 becomes more accessible to Type II PRMTs when localized to the cytoplasm.

Our results also highlight the value of combining analysis of both arginine monomethylation and dimethylation. Because PRMT1 can mediate both MMA and ADMA, putative PRMT1 substrates may exhibit either downregulation (Fig. 6B) or upregulation (Fig. 6C) of MMA upon PRMT1 knockdown. When available, measurement of both decreasing DMA and increasing MMA (Fig. 6A) provides increased confidence that the identified methyl peptide is a true substrate of PRMT1. In addition, switching between methyl forms can affect protein function, for example E2F-1 which is associated with apoptosis when ADMA modified by PRMT1 on R109 and associated with proliferation when SDMA modified by PRMT5 on R111 and R113 [12]. Here, through integrative analysis, we identified 10 SDMA and 7 ADMA sites that show increased methylation upon PRMT1 knockdown, suggesting that these methylation sites are normally competitively blocked by PRMT1. In our study, the increasing SDMA sites upon PRMT1 knockdown included R206 of the known PRMT5 target HNRNPA1 [54]. A crucial aspect of identifying ADMA/SDMA switching is the ability to discriminate these isobaric PTMs through neutral ion losses.

In summary, our results confirm that PRMT1 regulates a substantial amount of arginine methylation in mammalian cells. The fact that over 90% of significantly changing methyl arginine sites are not known interactors of PRMT1 demonstrates the need for continued comprehensive analysis of PRMTs and their substrates. This is especially relevant considering the growing body of evidence that dysregulation of arginine methylation may contribute to diseases affecting cancer [14]. The dynamic interplay between different methylation marks highlights the need for further

development of methods to quantify site occupancy across all methylation forms, as has been done for simpler PTMs including phosphorylation [65]. Towards this end, improved methods to distinguish ADMA and SDMA through fragmentation patterns will be valuable. Finally, given our demonstration that high pH SCX and IAP are largely orthogonal, the continued incorporation of fractionation techniques [24] and alternative methyl-peptide enrichment strategies [28]–[30] will enable deeper analysis of the protein methylome.

ACKNOWLEDGMENTS

This work was supported by the Viterbi School of Engineering (NAG), grant R01-DE026468 from the National Institute of Dental and Craniofacial Research (JX), and a Research Career Development Award from STOP CANCER Foundation (JX).

DATA AVAILABILITY

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository with the dataset identifier PXD012357 (Username: reviewer93189@ebi.ac.uk, password: mnoFDcyY).

REFERENCES

[1] P. Beltrao *et al.*, "Systematic Functional Prioritization of Protein Post-translational Modifications," *Cell*, vol. 150, no. 2, pp. 413–425, Jul. 2012.

[2] W. K. Paik and S. Kim, "Enzymatic methylation of protein fractions from calf thymus nuclei," *Biochem. Biophys. Res. Commun.*, vol. 29, no. 1, pp. 14–20, Oct. 1967.

[3] R. P. Ambler and M. W. Rees, "ε-N-Methyl-lysine in Bacterial Flagellar Protein," *Nature*, vol. 184, no. 4679, pp. 56–57, Jul. 1959.

[4] M. T. Bedford and S. Richard, "Arginine Methylation: An Emerging Regulatorof Protein Function," *Mol. Cell*, vol. 18, no. 3, pp. 263–272, Apr. 2005.

[5] J. Xu *et al.*, "Arginine methylation initiates BMP-induced Smad signaling," *Mol. Cell*, vol. 51, no. 1, pp. 5–19, Jul. 2013.

[6] J.-M. Hsu *et al.*, "Crosstalk between Arg 1175 methylation and Tyr 1173 phosphorylation negatively modulates EGFR-mediated ERK activation," *Nat. Cell Biol.*, vol. 13, no. 2, pp. 174–181, Feb. 2011.

[7] K. K. Biggar and S. S.-C. Li, "Non-histone protein methylation as a regulator of cellular signalling and function," *Nat. Rev. Mol. Cell Biol.*, vol. 16, no. 1, pp. 5–17, Jan. 2015.

[8] F. Liu *et al.*, "PKM2 methylation by CARM1 activates aerobic glycolysis to promote tumorigenesis," *Nat. Cell Biol.*, vol. 19, no. 11, pp. 1358–1370, Nov. 2017.

[9] Y.-P. Wang *et al.*, "Arginine Methylation of MDH1 by CARM1 Inhibits Glutamine Metabolism and Suppresses Pancreatic Cancer," *Mol. Cell*, vol. 64, no. 4, pp. 673–687, 17 2016.

[10] A. Scoumanne, J. Zhang, and X. Chen, "PRMT5 is required for cell-cycle progression and p53 tumor suppressor function," *Nucleic Acids Res.*, vol. 37, no. 15, pp. 4965–4976, Aug. 2009.

[11] K. B. Sylvestersen, H. Horn, S. Jungmichel, L. J. Jensen, and M. L. Nielsen, "Proteomic Analysis of Arginine Methylation Sites in Human Cells Reveals Dynamic Regulation During Transcriptional Arrest," *Mol. Cell. Proteomics*, vol. 13, no. 8, pp. 2072–2088, Aug. 2014.

[12] S. Zheng *et al.*, "Arginine methylation-dependent reader-writer interplay governs growth control by E2F-1," *Mol. Cell*, vol. 52, no. 1, pp. 37–51, Oct. 2013.

[13] G. Krapivinsky, L. Krapivinsky, N. E. Renthal, A. Santa-Cruz, Y. Manasian, and D. E. Clapham, "Histone phosphorylation by TRPM6's cleaved kinase attenuates adjacent arginine methylation to regulate gene expression," *Proc. Natl. Acad. Sci.*, vol. 114, no. 34, pp. E7092–E7100, Aug. 2017.

[14] Y. Yang and M. T. Bedford, "Protein arginine methyltransferases and cancer," *Nat. Rev. Cancer*, vol. 13, no. 1, pp. 37–50, Jan. 2013.

[15] M. T. Bedford and S. G. Clarke, "Protein Arginine Methylation in Mammals: Who, What, and Why," *Mol. Cell*, vol. 33, no. 1, pp. 1–13, Jan. 2009.

[16] Q. Wang, K. Wang, and M. Ye, "Strategies for large-scale analysis of non-histone protein methylation by LC-MS/MS," *Analyst*, vol. 142, no. 19, pp. 3536–3548, Sep. 2017.

[17] T. E. Thingholm, T. J. D. Jørgensen, O. N. Jensen, and M. R. Larsen, "Highly selective enrichment of phosphorylated peptides using titanium dioxide," *Nat. Protoc.*, vol. 1, no. 4, pp. 1929–1935, Nov. 2006.

[18] J. Villén and S. P. Gygi, "The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry," *Nat. Protoc.*, vol. 3, no. 10, pp. 1630–1638, Oct. 2008.

[19] P. Hägglund, J. Bunkenborg, F. Elortza, O. N. Jensen, and P. Roepstorff, "A New Strategy for Identification of N-Glycosylated Proteins and Unambiguous Assignment of Their Glycosylation Sites Using HILIC Enrichment and Partial Deglycosylation," *J. Proteome Res.*, vol. 3, no. 3, pp. 556–566, Jun. 2004.

[20] M. Evich, E. Stroeva, Y. G. Zheng, and M. W. Germann, "Effect of methylation on the side-chain pK a value of arginine," *Protein Sci. Publ. Protein Soc.*, vol. 25, no. 2, pp. 479–486, Feb. 2016.
[21] A. Guo *et al.*, "Immunoaffinity Enrichment and Mass Spectrometry Analysis of Protein Methylation," *Mol. Cell. Proteomics*, vol. 13, no. 1, pp. 372–387, Jan. 2014.
[22] M. Bremang, A. Cuomo, A. M. Agresta, M. Stugiewicz, V. Spadotto, and T. Bonaldi, "Mass spectrometry-based identification and characterisation of lysine and arginine methylation in the human proteome," *Mol. Biosyst.*, vol. 9, no. 9, pp. 2231–2247, Sep. 2013.
[23] X.-J. Cao, A. M. Arnaudo, and B. A. Garcia, "Large-scale global identification of protein lysine methylation in vivo," *Epigenetics*, vol. 8, no. 5, pp. 477–485, May 2013.
[24] S. C. Larsen *et al.*, "Proteome-wide analysis of arginine monomethylation reveals widespread occurrence in human cells," *Sci Signal*, vol. 9, no. 443, pp. rs9–rs9, Aug. 2016.
[25] V. Geoghegan, A. Guo, D. Trudgian, B. Thomas, and O. Acuto, "Comprehensive identification of arginine methylation in primary T cells reveals regulatory roles in cell signalling," *Nat. Commun.*, vol. 6, p. 6758, Apr. 2015.
[26] T. Uhlmann, V. L. Geoghegan, B. Thomas, G. Ridlova, D. C. Trudgian, and O. Acuto, "A Method for Large-scale Identification of Protein Arginine Methylation," *Mol. Cell. Proteomics MCP*, vol. 11, no. 11, pp. 1489–1499, Nov. 2012.
[27] K. Wang *et al.*, "Antibody-Free Approach for the Global Analysis of Protein Methylation," *Anal. Chem.*, vol. 88, no. 23, pp. 11319–11327, Dec. 2016.
[28] M. Ma *et al.*, "Strategy Based on Deglycosylation, Multiprotease, and Hydrophilic Interaction Chromatography for Large-Scale Profiling of Protein Methylation," *Anal. Chem.*, vol. 89, no. 23, pp. 12909–12917, Dec. 2017.
[29] K. E. Moore *et al.*, "A General Molecular Affinity Strategy for Global Detection and Proteomic Analysis of Lysine Methylation," *Mol. Cell*, vol. 50, no. 3, pp. 444–456, May 2013.
[30] Z. Wu *et al.*, "A chemical proteomics approach for global analysis of lysine monomethylome profiling," *Mol. Cell. Proteomics MCP*, vol. 14, no. 2, pp. 329–339, Feb. 2015.
[31] S. Dhar *et al.*, "Loss of the major Type I arginine methyltransferase PRMT1 causes substrate scavenging by other PRMTs," *Sci. Rep.*, vol. 3, Feb. 2013.
[32] M. B. Nabity *et al.*, "Symmetric Dimethylarginine Assay Validation, Stability, and Evaluation as a Marker for the Early Detection of Chronic Kidney Disease in Dogs," *J. Vet. Intern. Med.*, vol. 29, no. 4, pp. 1036–1044, 2015.
[33] J. Martens-Lobenhoffer and S. M. Bode-Böger, "Chromatographic-mass spectrometric methods for the quantification of l-arginine and its methylated metabolites in biological fluids," *J. Chromatogr. B*, vol. 851, no. 1, pp. 30–41, May 2007.
[34] K. Vishwanathan, R. L. Tackett, J. T. Stewart, and M. G. Bartlett, "Determination of arginine and methylated arginines in human plasma by liquid chromatography–tandem mass spectrometry," *J. Chromatogr. B. Biomed. Sci. App.*, vol. 748, no. 1, pp. 157–166, Oct. 2000.
[35] J. Rappsilber, Y. Ishihama, and M. Mann, "Stop and Go Extraction Tips for Matrix-Assisted Laser Desorption/Ionization, Nanoelectrospray, and LC/MS Sample Pretreatment in Proteomics," *Anal. Chem.*, vol. 75, no. 3, pp. 663–670, Feb. 2003.
[36] H. Hermjakob and R. Apweiler, "The Proteomics Identifications Database (PRIDE) and the ProteomExchange Consortium: making proteomics data accessible," *Expert Rev. Proteomics*, vol. 3, no. 1, pp. 1–3, Feb. 2006.
[37] S. Tyanova *et al.*, "The Perseus computational platform for comprehensive analysis of (prote)omics data," *Nat. Methods*, vol. 13, no. 9, pp. 731–740, Sep. 2016.

[38] D. Schwartz, M. F. Chou, and G. M. Church, "Predicting Protein Post-translational Modifications Using Meta-analysis of Proteome Scale Data Sets," *Mol. Cell. Proteomics MCP*, vol. 8, no. 2, pp. 365–379, Feb. 2009.
[39] A. Cheng, C. E. Grant, W. S. Noble, and T. L. Bailey, "MoMo: Discovery of statistically significant post-translational modification motifs," *Bioinforma. Oxf. Engl.*, Dec. 2018.
[40] V. Vacic, L. M. Iakoucheva, and P. Radivojac, "Two Sample Logo: a graphical representation of the differences between two sets of sequence alignments," *Bioinformatics*, vol. 22, no. 12, pp. 1536–1537, Jun. 2006.
[41] M. Pirmoradian, H. Budamgunta, K. Chingin, B. Zhang, J. Astorga-Wells, and R. A. Zubarev, "Rapid and deep human proteome analysis by single-dimension shotgun proteomics," *Mol. Cell. Proteomics*, p. mcp.O113.028787, Jan. 2013.
[42] P. Thandapani, T. R. O'Connor, T. L. Bailey, and S. Richard, "Defining the RGG/RG Motif," *Mol. Cell*, vol. 50, no. 5, pp. 613–623, Jun. 2013.
[43] J. Tang *et al.*, "PRMT1 Is the Predominant Type I Protein Arginine Methyltransferase in Mammalian Cells," *J. Biol. Chem.*, vol. 275, no. 11, pp. 7723–7730, Mar. 2000.
[44] S. Kerrien *et al.*, "IntAct—open source resource for molecular interaction data," *Nucleic Acids Res.*, vol. 35, no. suppl_1, pp. D561–D565, Jan. 2007.
[45] W. L. Wooderchak, T. Zang, Z. S. Zhou, M. Acuña, S. M. Tahara, and J. M. Hevel, "Substrate Profiling of PRMT1 Reveals Amino Acid Sequences That Extend Beyond the 'RGG' Paradigm," *Biochemistry*, vol. 47, no. 36, pp. 9456–9466, Sep. 2008.
[46] L. L. Belyanskaya, P. M. Gehrig, and H. Gehring, "Exposure on cell surface and extensive arginine methylation of ewing sarcoma (EWS) protein," *J. Biol. Chem.*, vol. 276, no. 22, pp. 18681–18687, Jun. 2001.
[47] D. O. Passos, G. C. Bressan, F. C. Nery, and J. Kobarg, "Ki-1/57 interacts with PRMT1 and is a substrate for arginine methylation," *FEBS J.*, vol. 273, no. 17, pp. 3946–3961, Sep. 2006.
[48] G. Dreyfuss, V. N. Kim, and N. Kataoka, "Messenger-RNA-binding proteins and the messages they carry," *Nat. Rev. Mol. Cell Biol.*, vol. 3, no. 3, pp. 195–205, Mar. 2002.
[49] D. P. Zolg *et al.*, "ProteomeTools: Systematic characterization of 21 post-translational protein modifications by LC-MS/MS using synthetic peptides," *Mol. Cell. Proteomics*, p. mcp.TIR118.000783, Jan. 2018.
[50] C. J. Brame, M. F. Moran, and L. D. B. McBroom-Cerajewski, "A mass spectrometry based method for distinguishing between symmetrically and asymmetrically dimethylated arginine residues," *Rapid Commun. Mass Spectrom.*, vol. 18, no. 8, pp. 877–881, Apr. 2004.
[51] J. Rappsilber, W. J. Friesen, S. Paushkin, G. Dreyfuss, and M. Mann, "Detection of Arginine Dimethylated Peptides by Parallel Precursor Ion Scanning Mass Spectrometry in Positive Ion Mode," *Anal. Chem.*, vol. 75, no. 13, pp. 3107–3114, Jul. 2003.
[52] H. Takai *et al.*, "5-Hydroxymethylcytosine plays a critical role in glioblastomagenesis by recruiting the CHTOP-methylosome complex," *Cell Rep.*, vol. 9, no. 1, pp. 48–60, Oct. 2014.
[53] T. B. van Dijk *et al.*, "Friend of Prmt1, a novel chromatin target of protein arginine methyltransferases," *Mol. Cell. Biol.*, vol. 30, no. 1, pp. 260–272, Jan. 2010.
[54] G. Gao, S. Dhar, and M. T. Bedford, "PRMT5 regulates IRES-dependent translation via methylation of hnRNP A1," *Nucleic Acids Res.*, vol. 45, no. 8, pp. 4359–4369, May 2017.
[55] Q. Liu and G. Dreyfuss, "In vivo and in vitro arginine methylation of RNA-binding proteins.," *Mol. Cell. Biol.*, vol. 15, no. 5, pp. 2800–2808, May 1995.
[56] E. C. Shen, M. F. Henry, V. H. Weiss, S. R. Valentini, P. A. Silver, and M. S. Lee, "Arginine methylation facilitates the nuclear export of hnRNP proteins," *Genes Dev.*, vol. 12, no. 5, pp. 679–691, Mar. 1998.

[57] G. Gurunathan, Z. Yu, Y. Coulombe, J.-Y. Masson, and S. Richard, “Arginine methylation of hnRNPUL1 regulates interaction with NBS1 and recruitment to sites of DNA damage,” *Sci. Rep.*, vol. 5, p. 10475, May 2015.

[58] J. Kzhyshkowska *et al.*, “Heterogeneous nuclear ribonucleoprotein E1B-AP5 is methylated in its Arg-Gly-Gly (RGG) box and interacts with human arginine methyltransferase HRMT1L1,” *Biochem. J.*, vol. 358, no. 2, pp. 305–314, Sep. 2001.

[59] V. H. Ryan *et al.*, “Mechanistic View of hnRNPA2 Low-Complexity Domain Structure, Interactions, and Phase Separation Altered by Mutation and Arginine Methylation,” *Mol. Cell*, vol. 69, no. 3, pp. 465-479.e7, Feb. 2018.

[60] F.-M. Boisvert, M. J. Hendzel, J.-Y. Masson, and S. Richard, “Methylation of MRE11 Regulates its Nuclear Compartmentalization,” *Cell Cycle*, vol. 4, no. 7, pp. 981–989, Jul. 2005.

[61] W. A. Smith, B. T. Schurter, F. Wong-Staal, and M. David, “Arginine Methylation of RNA Helicase A Determines Its Subcellular Localization,” *J. Biol. Chem.*, vol. 279, no. 22, pp. 22795–22798, May 2004.

[62] S. Aratani *et al.*, “The nuclear import of RNA helicase A is mediated by importin-α3,” *Biochem. Biophys. Res. Commun.*, vol. 340, no. 1, pp. 125–133, Feb. 2006.

[63] R. S. Blanc and S. Richard, “Arginine Methylation: The Coming of Age,” *Mol. Cell*, vol. 65, no. 1, pp. 8–24, Jan. 2017.

[64] J. J. Hamey, R. J. Separovich, and M. R. Wilkins, “MT-MAMS: Protein Methyltransferase Motif Analysis by Mass Spectrometry,” *J. Proteome Res.*, vol. 17, no. 10, pp. 3485–3491, Oct. 2018.

[65] J. V. Olsen *et al.*, “Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis,” *Sci Signal*, vol. 3, no. 104, pp. ra3–ra3, Jan. 2010.

FIGURES

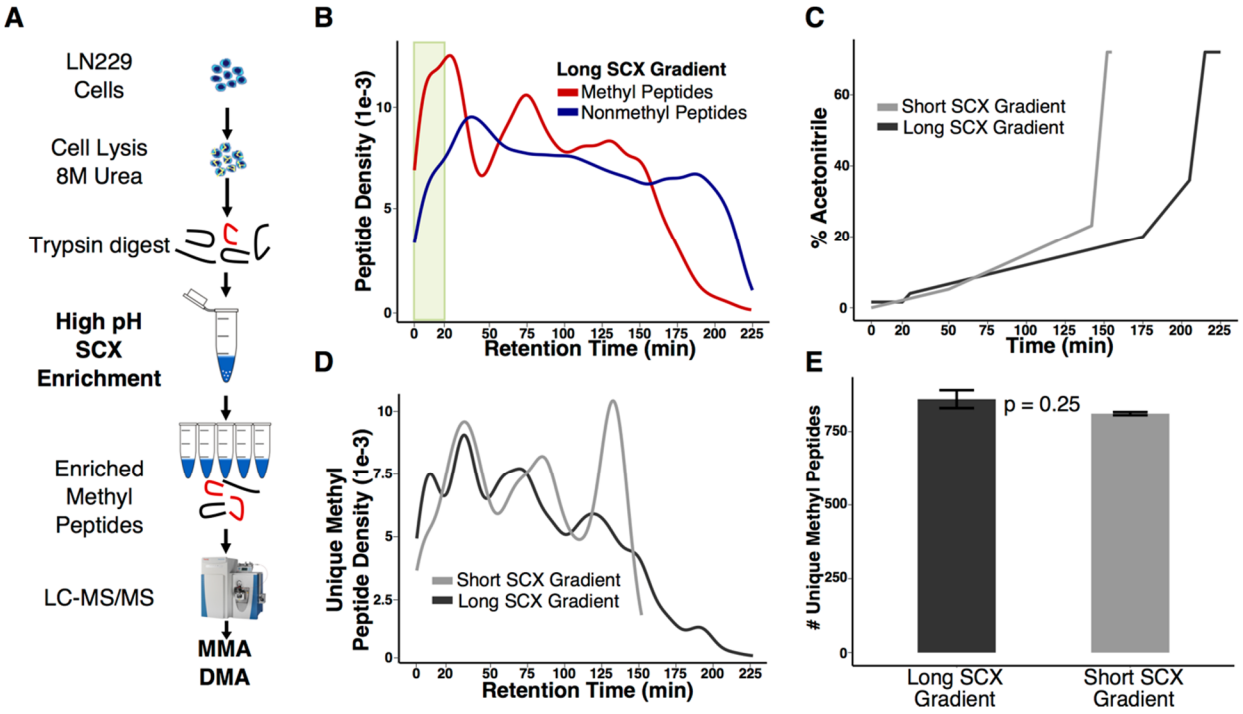


Figure 1: LC gradient optimization for high-pH SCX reduces instrument time while retaining sensitivity for methyl peptides

A) Schematic of the high-pH SCX enrichment protocol. Tryptic lysates from LN229 cells were incubated with SCX beads at high pH. Peptides were eluted off the SCX beads with high-pH buffers into five separate fractions. B) Density plot of methyl and non-methyl peptides versus time using the proposed LC gradient from Wang et al (1). The shaded bar indicates the sample loading phase from where the percentage of ACN is held constant (2%). Methyl peptides eluted earlier than non-methyl peptides. C) LC gradients showing the percentage of acetonitrile versus time for the short and long gradients. D) Density plot of methyl peptides identified versus time for the long and short LC gradients shown in C. E) Number of methyl peptides identified using the long and short gradients. Data are from 2 independent experiments. Error bars represent the standard deviation.

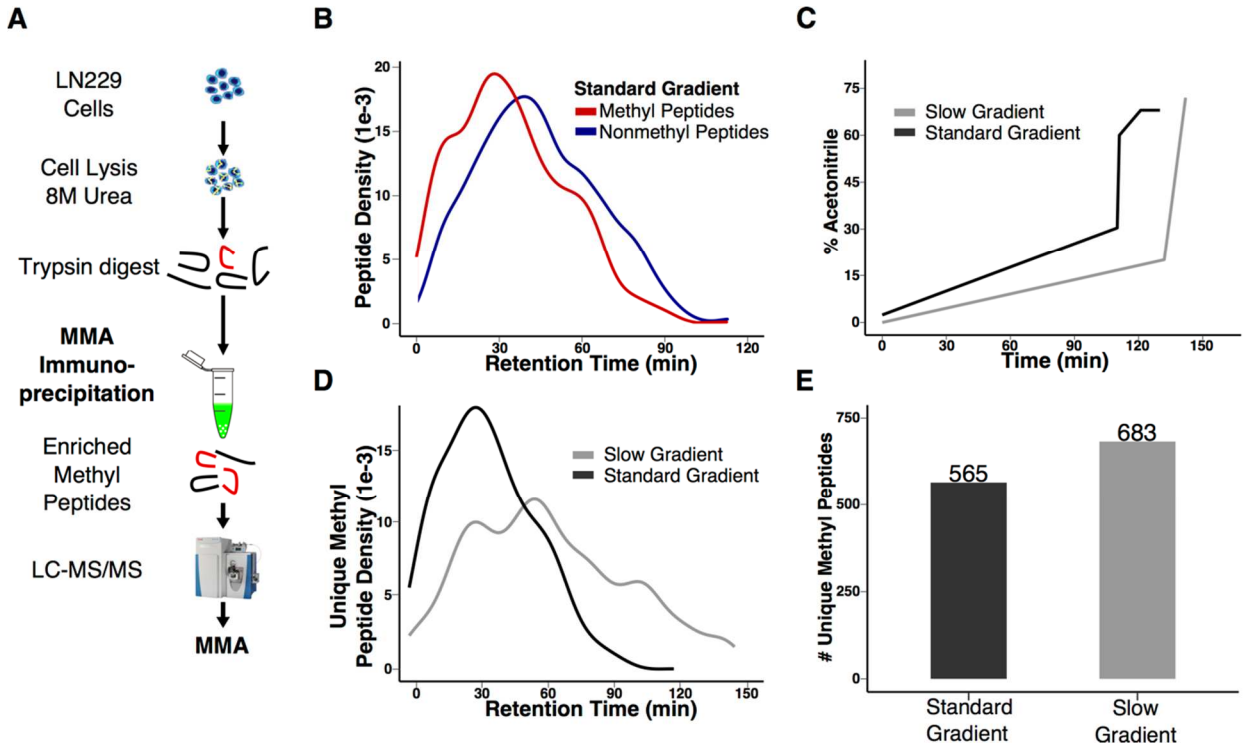


Figure 2: Slower gradient for Immunoaffinity Purification improves number of methyl peptides identified

A) Schematic of the IAP enrichment protocol. Tryptic lysates from LN229 cells were incubated with anti-MMA antibodies conjugated to agarose beads.

B) Density plot of methyl and non-methyl peptides versus time using a standard in-house proteomics gradient. Methyl peptides eluted earlier than non-methyl peptides.

C) LC gradients showing the percentage of Acetonitrile versus time for the standard and slow gradients.

D) Density plot of methyl peptides identified versus time for the standard and slow LC gradients shown in C.

E) Number of methyl peptides identified using the standard and slow gradients. Data are from 2 equal injections of single MMA IAP run on each gradient.

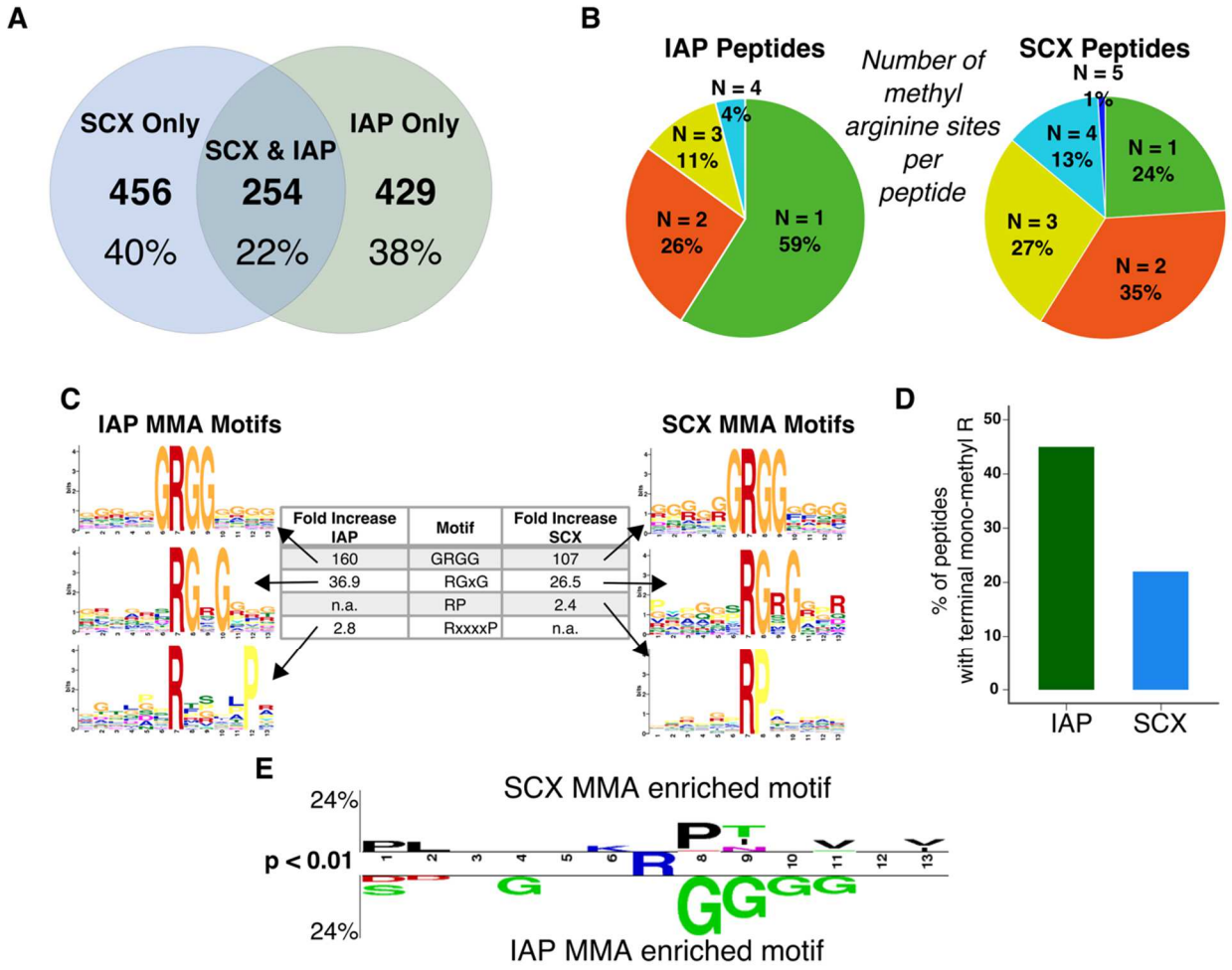


Figure 3: SCX and IAP enrich different subsets of methyl peptides

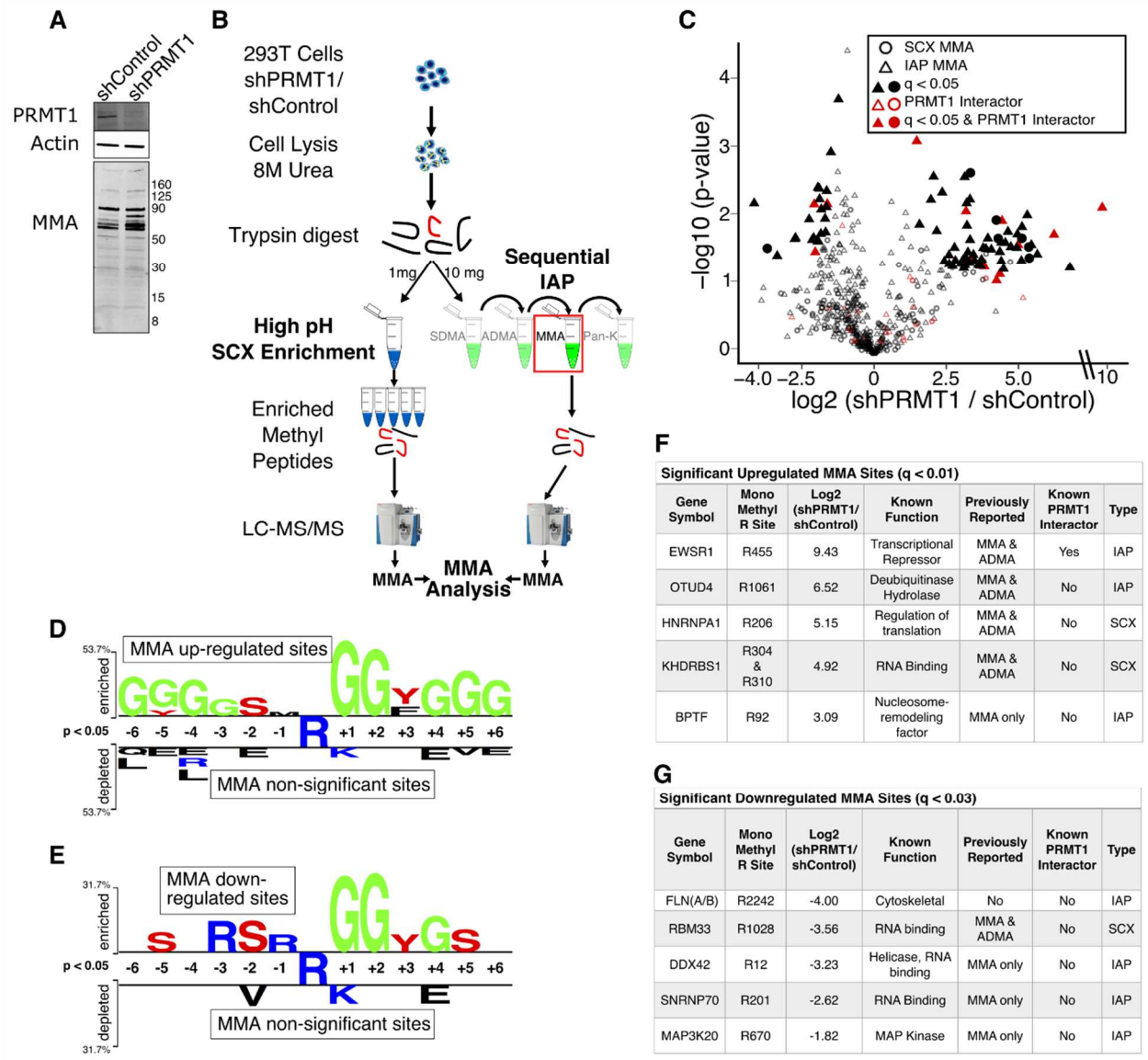
A) Overlap of MMA containing methyl peptides in SCX and IAP from LN229 cells.

B) Number of methyl R sites per peptide from A in SCX and IAP.

C) Motif table of enriched sequences from SCX MMA or IAP MMA sites in LN229 cells. All motifs shown were significant with adjusted p-value < 2e-8 [39].

D) Percent of peptides with terminal mono-methyl R sites in SCX and IAP in LN229 cells.

E) Two-sample motif analysis of MMA peptides enriched by SCX or IAP in LN229 cells.



F

Significant Upregulated MMA Sites (q < 0.01)

Gene Symbol	Mono Methyl R Site	Log2 (shPRMT1/ shControl)	Known Function	Previously Reported	Known PRMT1 Interactor	Type
EWSR1	R455	9.43	Transcriptional Repressor	MMA & ADMA	Yes	IAP
OTUD4	R1061	6.52	Deubiquitinase Hydrolase	MMA & ADMA	No	IAP
HNRNPA1	R206	5.15	Regulation of translation	MMA & ADMA	No	SCX
KHDRBS1	R304 & R310	4.92	RNA Binding	MMA & ADMA	No	SCX
BPTF	R92	3.09	Nucleosome-remodeling factor	MMA only	No	IAP

G

Significant Downregulated MMA Sites (q < 0.03)

Gene Symbol	Mono Methyl R Site	Log2 (shPRMT1/ shControl)	Known Function	Previously Reported	Known PRMT1 Interactor	Type
FLN(A/B)	R2242	-4.00	Cytoskeletal	No	No	IAP
RBM33	R1028	-3.56	RNA binding	MMA & ADMA	No	SCX
DDX42	R12	-3.23	Helicase, RNA binding	MMA only	No	IAP
SNRNP70	R201	-2.62	RNA Binding	MMA only	No	IAP
MAP3K20	R670	-1.82	MAP Kinase	MMA only	No	IAP

Figure 4: MMA analysis of PRMT1 depleted-293T cells reveals novel PRMT1 MMA and ADMA targets.

A) Western blot of PRMT1 and MMA proteins in 293T shPRMT1 or control cells. β-actin was used as an equal loading control.

B) Schematic of parallel SCX and IAP enrichment of 293T cells. Only MMA peptides were considered for analysis here. Analysis of di-methylarginine can be found in Fig. 5.

C) Volcano plot of quantified MMA peptides in 293T shPRMT1 or control cells. Enrichment method for each peptide is noted along with significance cutoff ($q < 0.05$) and whether the protein is a known PRMT1 interactor.

D) Two-sample motif analysis of MMA peptides enriched by SCX or IAP in 293T cells.

E) List of significantly changing MMA sites upon PRMT1 depletion.

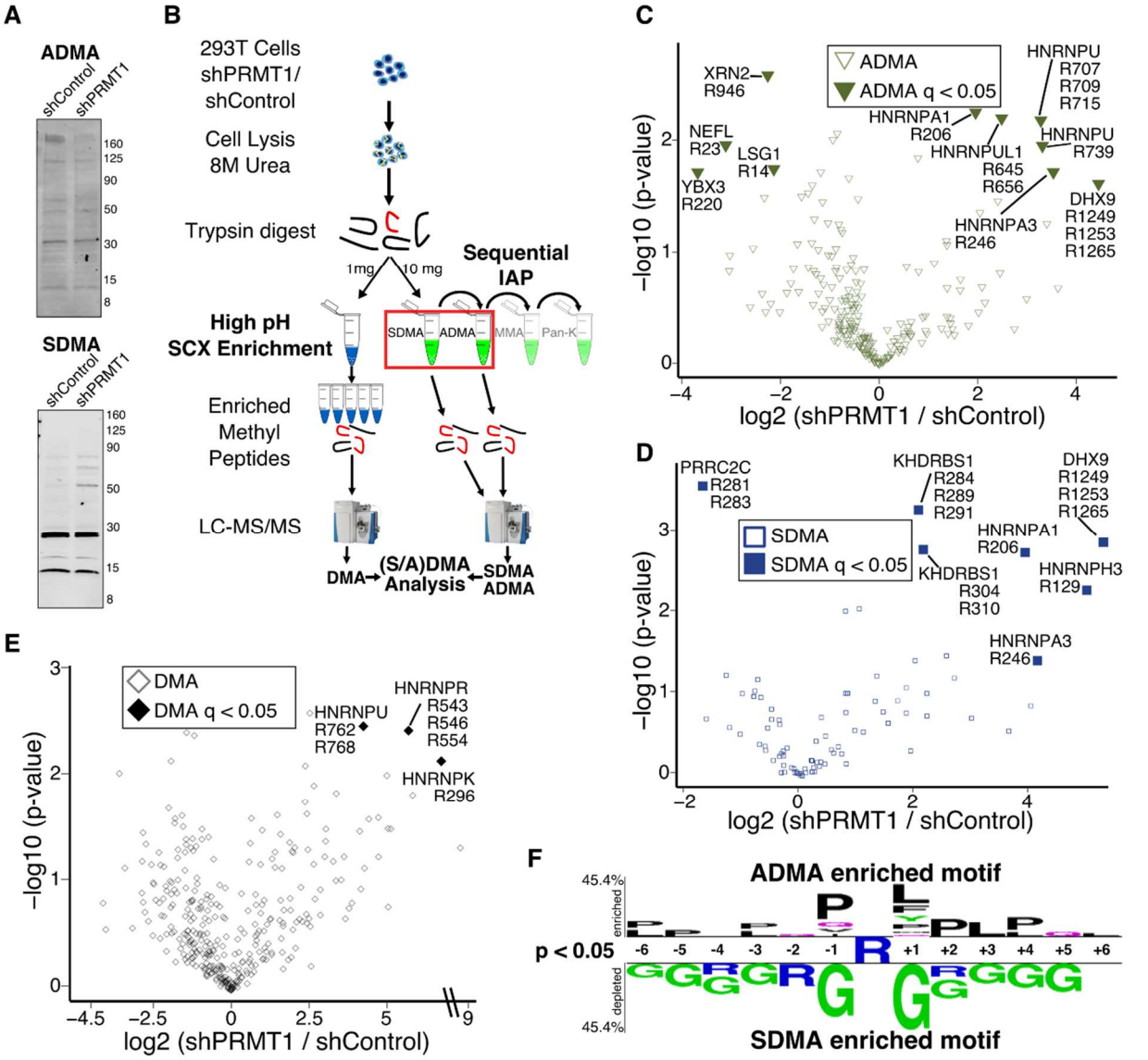


Figure 5: Analysis of dimethyl-arginine in PRMT1-depleted 293T cells.

A) Western blot of ADMA and SDMA proteins in 293T shPRMT1 or control cells. The equal loading control β-actin is shown in Fig. 4A.

B) Schematic of parallel SCX and IAP enrichment of 293T cells. Only (S/A)DMA peptides were considered for analysis here. Analysis of MMA can be found in Fig. 4.

C) Volcano plot of quantified ADMA peptides from IAP in 293T shPRMT1 or control cells. Enrichment method for each peptide is noted along with significance cutoff (q < 0.05).

D) Volcano plot of quantified SDMA peptides from IAP in 293T shPRMT1 or control cells ($q < 0.05$).

E) Volcano plot of quantified DMA peptides from SCX in 293T shPRMT1 or control cells ($q < 0.05$).

F) Two-sample motif analysis of ADMA peptides to SDMA peptides ($p < 0.05$)

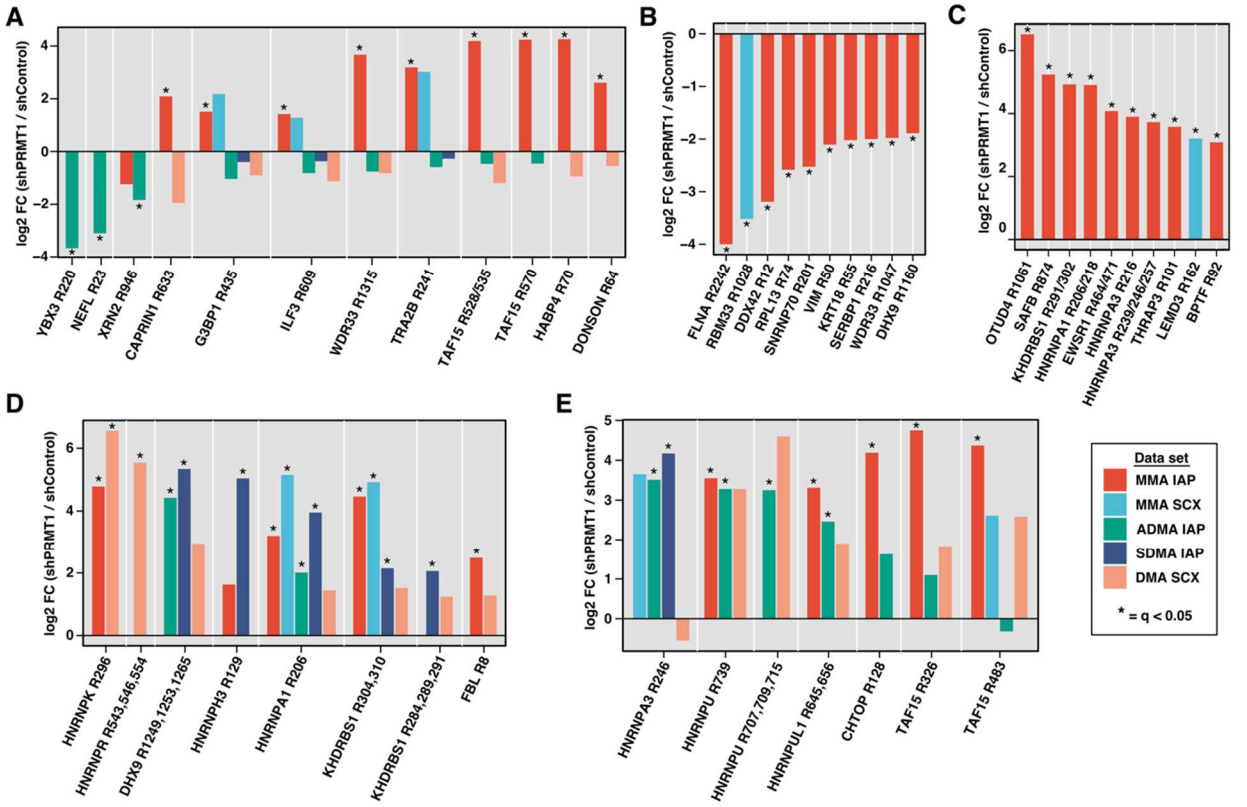


Figure 6: Integration of methyl arginine forms and characteristic neutral ion losses reveal high and medium confidence PRMT1 substrates and substrate scavenging by other PRMTs.

A) High confidence PRMT1 substrates that show a loss of ADMA upon knockdown of PRMT1 and low (172/158) characteristic ion ratios.

B) Putative PRMT1 MMA substrates that show loss of MMA upon knockdown of PRMT1. Of 24 identified methyl peptides, the 10 with the largest $\log_2$ fold-change are shown.

C) Putative PRMT1 ADMA substrates that show an accumulation of MMA upon knockdown of PRMT1. Of 19 identified methyl peptides, the 10 with the largest $\log_2$ fold change are shown.

D) Upregulated SDMA sites upon knockdown of PRMT1. Each methyl peptide was likely to contain SDMA (even when identified by ADMA IAP) based on the ratio of characteristic ions following neutral loss (et, 172/158 ratio).

E) Upregulated ADMA sites upon knockdown of PRMT1. Each methyl peptide was likely to contain ADMA (even when identified by SDMA IAP) based on the ratio of characteristic ions following neutral loss (et, 172/158 ratio).