Immunoaffinity Enrichment and Mass Spectrometry Analysis of Protein Methylation

Ailan Guo*‡, Hongbo Gu*‡, Jing Zhou*, Daniel Mulhern*, Yi Wang*, Kimberly A. Lee*, Vicky Yang*, Mike Aguiar*, Jon Kornhauser*, Xiaoying Jia*, Jianmin Ren*, Sean A. Beausoleil*, Jeffrey C. Silva*, Vidyasiri Vemulapalli§, Mark T. Bedford§, and Michael J. Comb*¶

Protein methylation is a common posttranslational modification that mostly occurs on arginine and lysine residues. Arginine methylation has been reported to regulate RNA processing, gene transcription, DNA damage repair, protein translocation, and signal transduction. Lysine methylation is best known to regulate histone function and is involved in epigenetic regulation of gene transcription. To better study protein methylation, we have developed highly specific antibodies against monomethyl arginine; asymmetric dimethyl arginine; and monomethyl, dimethyl, and trimethyl lysine motifs. These antibodies were used to perform immunoaffinity purification of methyl peptides followed by LC-MS/MS analysis to identify and quantify arginine and lysine methylation sites in several model studies. Overall, we identified over 1000 arginine methylation sites in human cell line and mouse tissues, and ~160 lysine methylation sites in human cell line HCT116. The number of methylation sites identified in this study exceeds those found in the literature to date. Detailed analysis of arginine-methylated proteins observed in mouse brain compared with those found in mouse embryo shows a tissue-specific distribution of arginine methylation, and extends the types of proteins that are known to be arginine methylated to include many new protein types. Many arginine-methylated proteins that we identified from the brain, including receptors, ion channels, transporters, and vesicle proteins, are involved in synaptic transmission, whereas the most abundant methylated proteins identified from mouse embryo are transcriptional regulators and RNA processing proteins. ***Molecular & Cellular Proteomics 13: 10.1074/mcp.O113.027870, 372–387, 2014.***

Protein methylation is a posttranslational modification that occurs predominantly on arginine residues and lysine residues (1, 2). Arginine methylation is catalyzed by protein arginine methyltransferases (PRMTs)[1] (3–5), and lysine methylation is carried out by the protein lysine methyltransferase family of enzymes (6, 7). Human PRMTs are classified in two major groups, type I, including PRMT1, PRMT3, PRMT4 (CARM1), PRMT6, and PRMT8, and type II, including PRMT5 and PRMT7. Both groups catalyze the formation of monomethyl arginine (MMA). Type I PRMT can also add an additional methyl group to the same guanidino nitrogen atom of arginine to form asymmetric dimethyl arginine (ADMA), and type II enzymes further catalyze the formation of symmetric dimethyl arginine (SDMA) by adding the second methyl group to a different guanidine nitrogen atom of arginine. Proteins that are arginine methylated are involved in many different cellular processes, including RNA processing, transcriptional regulation, and DNA damage repair (3, 4, 8).

PRMTs have been shown to modify many different cytoplasmic and nuclear proteins. The majority of ADMA residues reside within glycine- and arginine-rich sequences called GAR motif (8, 9). However, coactivator-associated arginine methyltransferase I (CARM1) (PRMT4) does not modify the GAR motif and instead methylates arginine in the PGM motif (proline-, glycine-, and methionine-rich) (10).

Protein lysine methylation involves the addition of one, two, or three methyl groups to the epsilon amine of lysine by PKMTs to form monomethyl (Kme1), dimethyl (Kme2), or trimethyl (Kme3) lysine, respectively. Lysine methylation has been extensively described on many residues of histone proteins, providing a role in the regulation of chromatin compaction and gene transcription (6, 8).

Protein methylation was thought to be irreversible for many years, until the recent identification of protein lysine demethylases (11, 12). This group of enzymes removes methyl groups

From * Cell Signaling Technology Inc., 3 Trask Lane, Danvers, Massachusetts 01923; §Department of Molecular Carcinogenesis, The University of Texas MD Anderson Cancer Center, Smithville, Texas 78957

⌘ *Author's Choice*—Final version full access.

Received January 30, 2013, and in revised form, October 3, 2013

Published, MCP Papers in Press, October 15, 2013, DOI 10.1074/mcp.O113.027870

[1] The abbreviations used are: PRMT, protein arginine methyltransferase; MMA, monomethyl arginine; ADMA, asymmetric dimethyl arginine; SDMA, symmetric dimethyl arginine; PKMT, protein lysine methyltransferase; Kme1, monomethyl lysine; Kme2, dimethyl lysine; Kme3, trimethyl lysine; IAP, immunoaffinity purification; IAP-LC-MS/MS, Immunoaffinity purification coupled with liquid chromatography–tandem mass spectrometry; AdOx, adenosine-2′,3′-dialdehyde; CARM1, coactivator-associated arginine methyltransferase I.

from methylated proteins and further increases the level and complexity of the regulation of protein methylation. Protein arginine and lysine methyltransferases themselves are also subjected to other modifications such as phosphorylation (13, 14). When methyltransferases are knocked out in mice, it can result in embryonic lethality or early death (5, 15–17), indicating the potentially significant biological role this group of enzymes plays, which to this day remains largely unknown.

Protein methyltransferases and demethylases have been implicated in human health and disease (18, 19). CARM1 (PRMT4) is overexpressed in breast and prostate cancer (20, 21). PRMT1 aberrant expression was observed in breast and colon cancers (22, 23). Many protein lysine methyltransferases have been shown to be overexpressed in human tumors, including SUV39H1 (24) and EZH2 (25, 26). A gain-of-function mutation of EZH2 has also been reported to lead to tumorigenesis (27). Many PRMTs and PKMTs are pursued as therapeutic targets, and small molecules are screened as PRMT and PKMT inhibitors (28, 29).

Considering the important biological roles of protein methylation and its involvement in human disease mechanisms, there is a need for methods to identify methylated proteins *in vivo* at a depth comparable to that achieved by enrichment techniques that have been developed for protein phosphorylation and protein acetylation. Mass-spectrometry-based proteomics has been a great tool for studying protein posttranslational modification, and with it thousands of phosphorylation sites have been discovered through affinity enrichment via immobilized metal affinity chromatography (IMAC), titanium dioxide chromatography, and phosphorylation-specific antibodies (30–32). LC-MS/MS techniques have also been developed to study protein acetylation (33) and ubiquitination (34, 35), leading to the identification of thousands of sites via immunoaffinity enrichment and LC-MS/MS-based analysis of enriched peptides. There have been several proteomics studies to identify arginine- and lysine-methylated proteins. The first study utilized dimethyl arginine antibodies to purify arginine-methylated protein complexes and MS analysis of digested proteins (36), resulting in about 200 putative arginine-methylated proteins without precise information on sites. Another study took advantage of stable isotope labeling of amino acids in cell culture to label methyl groups using [$^{13}CD_3$]S-adenosyl methionine in the cell culture medium in order to increase the confidence of methylation site identification (37). Coupled with antibody enrichment of methyl proteins, the approach enabled the identification of around 60 protein methylation sites. Another recent study used bioinformatics tools to analyze 36,854 previously generated MS/MS spectra to identify 38 arginine and 45 lysine methylation sites from yeast (38). A very recent study (39) attempted large-scale proteomic identification of arginine methylation sites using different chromatography separation and enrichment of methylation peptides and identified 249 arginine methylation (mostly dimethyl arginine) sites in 131 proteins.

In this report we describe the development of several antibodies engineered against different specific arginine and lysine methylation states, and the utilization of these antibodies to enrich for methylated peptides that were subsequently characterized via LC-MS/MS. Using this approach, we were able to identify hundreds to thousands of methylated peptides from a single IAP-LC-MS/MS experiment using human cell line and mouse tissue samples. Most of the methylation sites that were identified in this study are novel, and the number of sites exceeded those reported in previous studies. We also found that many arginine-methylated proteins in brain are involved in synaptic transmission, whereas most abundantly arginine-methylated proteins in embryo are transcription regulators and RNA processing proteins.

EXPERIMENTAL PROCEDURES

Cell Treatment and Western Blot Analysis—HCT116 cells were cultured in DMEM with 10% FBS. Cells were treated with 20 μM AdOx for 24 h; then the medium was discarded, new medium with 20 μM AdOx was added, and the cells were incubated for another 24 h. Cells were harvested in 1x cell lysis buffer (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM Na_2EDTA, 1 mM EGTA, 1% Triton, 2.5 mM sodium pyrophosphate, 1 mM β-glycerophosphate, 1 mM Na_3VO_4, 1 μg/ml leupeptin) supplemented with Complete Mini EDTA-free protease inhibitor mixture (Roche). Lysates were sonicated and centrifuged at 14,000 rpm for 15 min. The protein concentration was measured using Coomassie protein assay reagent (Pierce Chemical Co., Rockford, IL). Equal amounts of total protein (25 to 30 μg) were resolved via SDS-PAGE gel and transferred to nitrocellulose membranes. Blots were incubated overnight at 4 °C with the appropriate antibodies. A total of 500 μg of protein lysate was used for protein immunoprecipitation. The cleared protein lysate was rocked with 2 μg of proper antibody and 15 μl of protein A agarose beads (Pierce) overnight at 4 °C. The beads were washed three times with 1x cell lysis buffer and boiled in 30 μl of 2x SDS-PAGE sample buffer for 5 min. The bound protein was then analyzed via Western blot.

Preparation of Lysates and Peptides—Subconfluent HCT116 cells were harvested with 8 M urea lysis buffer (20 mM HEPES, 8 M urea, 1 mM sodium vanadate, 1 mM β-glycerol phosphate, 2.5 mM sodium pyrophosphate). The lysates were sonicated two times for 30 s each at 15-W output power. Mouse tissue was homogenized using PolyTron for three bursts of 20 s each, then sonicated two times for 20 s each. The sonicated lysates were centrifuged for 15 min at 4 °C at 20,000*g*. The lysate concentration was measured using Coomassie protein assay reagent. 1 mg of lysate from mouse tissue was saved for Western blot analysis; the rest of the supernatants were reduced with 4.5 mM DTT for 30 min at 55 °C and then alkaylated with iodoacetamide (10 mM) for 15 min at room temperature in the dark. The samples were then diluted greater than 4-fold with 20 mM HEPES, pH 8.0, to make the lysate concentration 1 mg/ml and digested with trypsin (10 μg/ml) overnight at room temperature. The digests were acidified with 1% TFA, and the peptides were desalted and crudely purified from other cellular debris over Sep-Pak C18 columns (WAT051910, Waters, Milford, MA), eluted with 40% acetonitrile in 0.1% TFA, lyophilized, and stored at −80 °C.

Peptide Immunoaffinity Purification—Methylation specific motif antibodies (250 μg) were conjugated to protein A agarose beads (40 μl) by being rocked with the beads for at least 3 h and then were extensively washed with PBS (three times with 1 ml each time) and IAP buffer (50 mM MOPS, pH 7.2, 10 mM sodium phosphate, 50 mM NaCl) (twice with 1 ml each time). A total of 10 mg of peptides were

dissolved in 1.4 ml of IAP buffer and incubated with the antibody beads for 2 h at 4 °C. The beads were washed four times with MOPS IAP buffer and two times with 1 ml of dH_2O. Bound peptides were eluted from the beads with 50 μl of 0.15% TFA, followed by a second 50 μl. The eluates were combined and then desalted and concentrated over StageTips packed with Empore C18 (Sigma). Eluted peptides were analyzed via mass spectrometry.

LC-MS/MS Analysis—Methylation motif antibody-enriched peptides were dissolved in 0.1% TFA with 5% CH_3CN and separated on a 100 μm × 15 cm reversed-phase column packed with Magic C18 AQ (100 A × 3 μM, Michrom, Auburn, CA) into a pulled tip. Peptides were eluted using a 72-min linear gradient of 5%–30% acetonitrile in 0.125% formic acid delivered at 300 nl/min using an Easy nLC (Thermo Fisher). Tandem mass spectra were collected in a data-dependent manner with either an Orbitrap Elite or a Q Exactive mass spectrometer. The Orbitrap Elite utilized a top-20 method, collecting MS spectra in the Orbitrap mass analyzer (60,000 resolution) with an automatic gain control target of 1e6 (maximum ion time: 1000 ms) and collision-induced dissociation MS/MS spectra in the ion trap with an automatic gain control target of 4e3 (maximum ion time: 50 ms). The Q Exactive employed a top-10 method, collecting MS spectra at 70,000 resolution and 1e6 automatic gain control with a maximum ion time of 10 ms and higher-energy collisional dissociation MS/MS spectra at 17,500 resolution and 1e5 automatic gain control with maximum ion time of 120 ms and a normalized collision energy of 25. Both instruments scanned precursor masses from 300–1500 *m/z* and employed dynamic exclusion with a repeat count of 1 and a repeat duration of 30 s. Ions that had a charge of 1 or were unassigned were excluded from MS/MS analysis. The polydimethylsiloxane lock mass (*m/z* 371.10123) was used as an internal calibrant for all runs.

Following mass spectrometry data acquisition, RAW files were converted into mzXML format and processed using a suite of software tools developed in-house for the analysis of large-scale proteomics datasets. All precursors selected for MS/MS fragmentation were confirmed using algorithms to detect and correct errors in monoisotopic peak assignment and refine precursor ion mass measurement. All MS/MS spectra were then exported as individual DTA files and searched using SEQUEST (v. 28 (rev. 12), 1998–2007) against the human database (20110802_REV_20110627human.fasta, NCBI, June 27, 2011; contains 34,895 forward proteins) and the mouse database (20110802_REV_20110628mouse.fasta, NCBI, June 28, 2011; contains 30,120 forward proteins) and their reversed complements. A precursor mass tolerance of 50 ppm and a product ion tolerance of 1.0 Da (collision-induced dissociation spectra) or 0.02 Da (higher-energy collisional dissociation spectra) were allowed. One tryptic terminus was required, and four missed cleavages were allowed. Static carbamidomethylation of cysteine (+57.02146374) was required, and appropriate lysine or arginine modifications (+14.0156500642 for monomethyl and +28.0313001284 for dimethyl) and methionine oxidation (+15.9949146221) were dynamically allowed with a maximum of four modifications of one type per peptide. Peptide spectral matches were filtered to a 1% false discovery rate using linear discriminant analysis (40) in combination with the target-decoy strategy. Sites of lysine and arginine methylation were determined using a slight variation on the Ascore algorithm (41), with sites scoring >13 considered confidently assigned.

In quantification studies, the intensity of enriched methyl peptides was generated using Skyline (version 1.4) (42) in MS1 full-scan mode. The intensity of identified peptides was represented by the peak area of the extracted ion chromatogram of their monoisotopic peaks. The peptide with the highest intensity was picked to represent one methylation site of a protein if multiple peptides were identified for the same site. Extracted ion chromatograms of identified methylated peptides that showed abundance changes between mouse brain and embryo were manually verified to ensure the accuracy of the quantification. For statistical analysis, a two-tailed *t* test was performed for all methyl-peptides representing unique methylation sites.

Reductive Amination to Create Dimethyl Lysine Peptides—A total of 10 mg of HCT116 tryptic peptides were dissolved in 5 ml of 0.1 M sodium phosphate buffer, pH 6.6. 200 μl of 4% formaldehyde and 200 μl of 600 mM sodium cyanoborohydride were added to the reaction. The reaction was incubated for 1 h at room temperature and stopped by the addition of 800 μl of 10% TFA; labeled peptides were cleaned up using a Sep-Pak C18 column and lyophilized.

RESULTS

Development of Methylation-specific Motif Antibodies—We used an approach described previously (43) to generate antibodies against different forms of protein methylation (Fig. 1*A*). Two different peptide antigen libraries were used to immunize New Zealand White rabbits to obtain monomethyl arginine antibodies. The first library (XXXXXXXR*XXXXXX) contained a single monomethyl arginine (R*) in a peptide library with all the other positions containing the mixture of all natural amino acids except tryptophan (W), cysteine (C), and tyrosine (Y). The second library (XXXXXXXR*GGXXX) had a monomethyl arginine in the RGG motif to reflect that arginine methylation frequently occurs in RG-rich sequences. Rabbits with good reactivity were identified and used to generate two monoclonal antibodies, Me-R^4-100 (CST #8015, Cell Signaling Technology, Danvers, MA) and R*GG (D5A12) (CST #8711, Cell Signaling Technology, Danvers, MA). When generating dimethyl arginine antibodies we used a more complicated peptide library (XXXXR*XXR*XXXXR*XXXR*X) (R* = ADMA) to try to obtain more diverse antibodies in a single immunization; X represents a mix of all amino acids excluding W, C, and Y. Two rabbits, F8216 and BL8241, showing good reactivity against PRMT1 and CARM1 substrates (supplemental Fig. S1) were chosen to generate monoclonal antibody clones D6A8 and D4H5. Monomethyl lysine, dimethyl lysine, and trimethyl lysine antibodies were generated using a peptide library that had a single appropriately modified lysine in a peptide library containing degenerated amino acids at surrounding positions (XXXXXXK*XXXXXX). These antibodies were antigen affinity purified polyclonal antibodies (Kme1: BL10749; Kme2: BL10745; Kme3: D3978) at this stage.

All the antibodies generated were validated for Western blot (Figs. 1 and 4*A*). To better characterize these antibodies, we generated lysates from HCT116 cells that were untreated or treated with the commonly used methyltransferase inhibitor AdOx (adenosine-2′,3′-dialdehyde) (29). PRMTs and PKMTs belong to Ado-Met-dependent methyltransferase families that also include DNA methyltransferases. Ado-Met (*S*-adenosyl-l-methionine) is the methyl donor for methylation reactions. Ado-Met is metabolized to AdoHcy by methyltransferases, and AdoHcy is then converted into adenosine and homocysteine by AdoHcy hydrolase. AdOx inhibits AdoHcy hydrolase, causing an increase in AdoHcy levels. AdoHcy in turn inhibits the methyltransferases. Dimethyl arginine antibodies D6A8

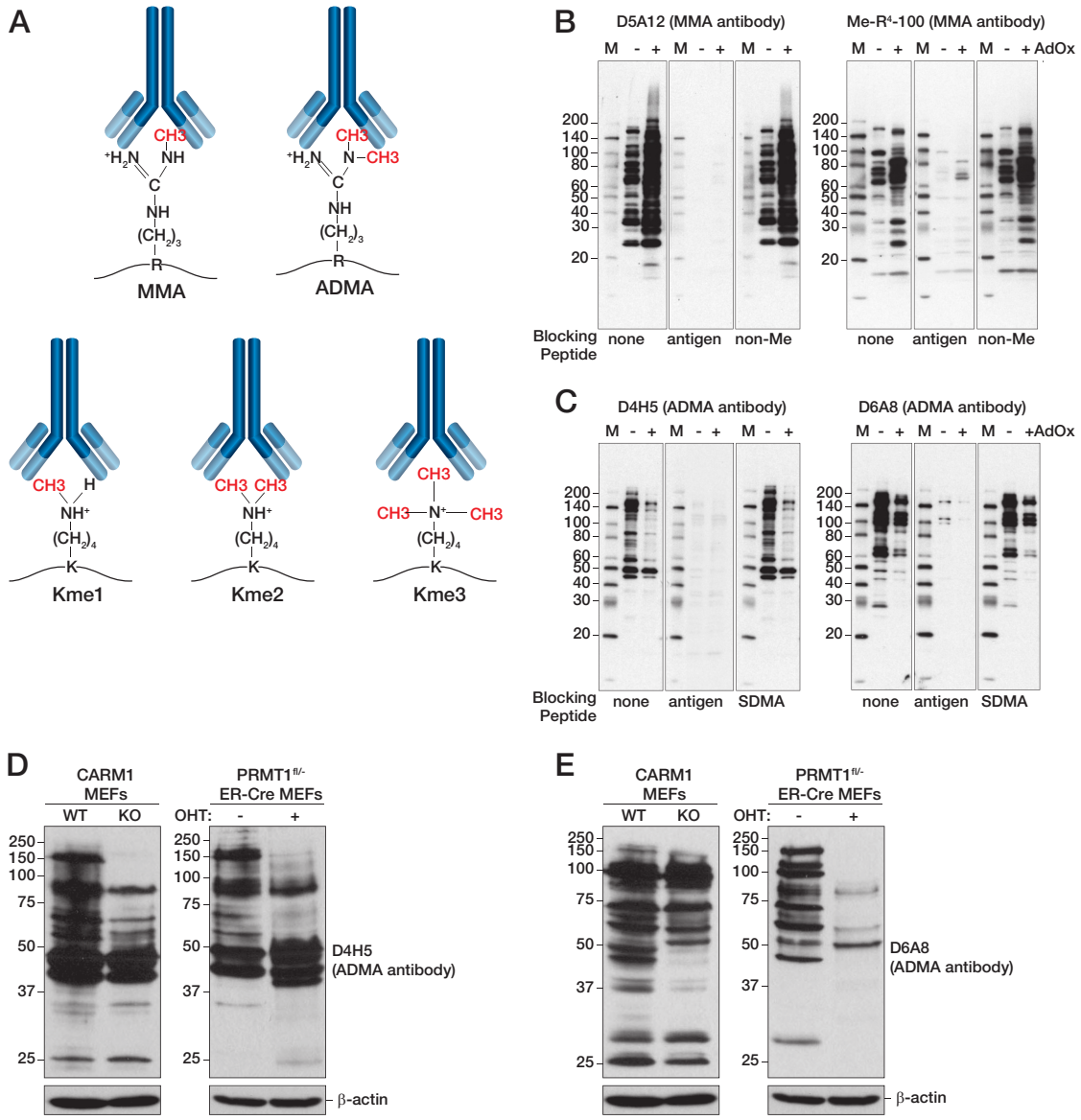
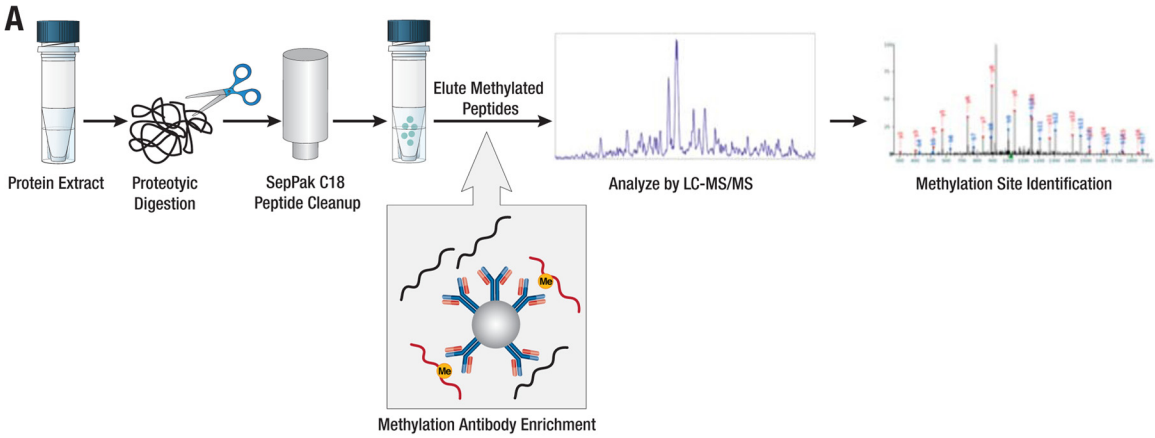


FIG. 1. **Characterization of methylation antibodies.** *A*, representation of different methylation antibodies generated using the peptide library approach. *B*, Western blot analysis of lysates from untreated or AdOx-treated (20 μM, 48 h) HCT116 cells using MMA antibodies D5A12 and Me-R^4-100 (left panels). The specific signal was blocked by monomethyl arginine peptides (middle panels), but not by non-methyl argnine control peptides (right panels). *C*, Western blot analysis of lysates from untreated or AdOx-treated (20 μM, 48 h) HCT116 cells using ADMA antibodies D4H5 and D6A8 (left panels). The specific signal was blocked by antigen ADMA peptides (middle panels), but not by corresponding SDMA control peptides (right panels). An equal amount of protein lysate (30 μg) was loaded into each lane in the gels in *B* and *C*, and Coommassie staining and β-actin signal were used to ensure equal loading of samples (shown in supplemental Fig. S3). M, protein marker lane. *D*, *E*, Western blot analysis of lysates from CARM1 wild-type (wt) and knockout (KO) mouse embryonic fibroblast cells and lysates from PRMT1$^{FL/-}$ER-Cre mouse embryonic fibroblasts with (+) and without (−) OHT (tamoxifen) treatment, using two asymmetric dimethyl arginine methylation (ADMA) antibodies, D4H5 and D6A8. D6A8 is more specific to PRMT1 substrates, showing decreased signal in PRMT1 KO cells (PRMT1$^{FL/-}$ER-Cre, + OHT). D4H5 showed CARM1- and PRMT1-specific bands, as well as some bands not blocked by CARM1 and PRMT1 KO. β-actin signal was used as a loading control.

and D4H5 showed clearly decreased signals in the AdOx-treated sample relative to the untreated sample, and the specific signals were blocked by antigen peptides, but not by the corresponding peptide that had symmetric dimethyl arginine (Fig. 1*C* and supplemental Fig. S2). Interestingly, when the same samples were probed with monomethyl arginine antibodies, the signal was shown to increase with AdOx treatment (Fig. 1*B* and supplemental Fig. S2). A recent paper reported that when PRMT1 is lost, an increased level of MMA is observed (44), which may partly explain our results. The monomethyl arginine antibody signals are specifically blocked only by monomethyl arginine peptides, and not by asymmet-



B

Methylation	Sites	Peptides	Proteins
MMA	1473	2053	716
ADMA	497	553	194
Kme1	132	234	111
Kme2	35	108	19
Kme3	31	61	23

FIG. 2. **IAP-LC-MS/MS method and experimental results in HCT116 cells.** *A*, schematic workflow of IAP-LC-MS/MS experiments. Protein extract from cell lines or tissue samples is proteolytically digested, and the resulting peptides are purified over a Sep-Pak C18 cartridge. Lyophilized, dried peptides are dissolved in IAP buffer and immunoaffinity purified using methylation-specific antibodies. Enriched methyl peptides are analyzed via LC-MS/MS to identify the methylation sites. *B*, methylation sites identified from IAP-LC-MS/MS experiments in HCT116 cells. 10 mg of tryptic peptides from HCT116 cells were immunoaffinity purified using 250 μg of methylation antibodies in each experiment. Reported MMA sites are the combined results from one IAP-LC-MS/MS experiment using D5A12 and one using the Me-R^4-100 MMA antibody. Reported ADMA sites are the combined results from one IAP-LC-MS/MS experiment using D4H5 and one using the D6A8 ADMA antibody. Kme1, Kme2, and Kme3 sites are from individual IAP-LC-MS/MS experiments using BL10749, BL10745, and D3978 antibodies, respectively.

ric or symmetric dimethyl arginine peptides. The same is true for asymmetric dimethyl arginine antibodies, which are not blocked by non-methyl arginine peptides, monomethyl arginine peptides, or symmetric dimethyl arginine peptides (Figd. 1*B* and 1*C*, supplemental Fig. S2). When ADMA antibodies D4H5 and D6A8 were tested with PRMT1 knockout and CARM1 knockout cells, clone D6A8 showed a more specific signal for PRMT1 substrates, whereas D4H5 recognized CARM1 as well as PRMT1 substrates and some other proteins that might be methylated by other PRMTs (Figs. 1*D* and 1*E*).

Identification of Methylation Sites in HCT116 Cells—Protein methylation site identification using methylation motif antibodies to enrich methylated peptides and analysis of the enriched peptides via LC-MS/MS were developed as an adaptation of a similar IAP-LC-MS/MS approach used to characterize phosphopeptides (30, 45) and ubiquitin-modified peptides (34). Fig. 2*A* shows a schematic workflow of the IAP-LC-MS/MS method. A typical experiment utilizes tryptic peptides from approximately 10 mg of protein extract derived from either cell lines or tissue samples and 100 to 250 μg of methylation-specific antibodies. Fig. 2*B* shows the number of methylation sites that were identified from colon cancer cell line HCT116 cells in single IAP-LC-MS/MS analysis experiments using different methylation-specific motif antibodies. The experimental details, the antibodies used in the study, and the numbers of peptides identified are listed in supplemental Table S1. Representative MS/MS spectra of several methyl peptides are shown in supplemental Fig. S4. Detailed peptide information with corresponding ModScore values used to score the site of methyl modification and a link to the corresponding MS/MS spectrum are provided for all identified methyl peptides in supplemental Table S2.

From a single experiment using Me-R^4-100 antibody along with tryptic peptides from HCT116 cells, we identified 1106 MMA sites from 1743 unique methylated peptides from 570 proteins. R*GG (D5A12) antibody identified 942 MMA sites from 1456 unique monomethyl arginine peptides from 489 proteins. With D4H5 ADMA antibody, we identified 226 ADMA sites from 287 unique peptides from 97 proteins. Using D6A8 ADMA antibodies with HCT116 tryptic peptides, 335 ADMA sites were identified from 440 peptides from 140 proteins (supplemental Tables S1 and S3). Using HCT116 tryptic peptides in lysine methylation profiling experiments with Kme1 antibody, we identified 132 monomethyl lysine sites from 111

proteins; we identified 35 Kme2 and 31 Kme3 sites from about 20 proteins (supplemental Tables S1 and S4).

We repeated the experiments many times, with good reproducibility of the corresponding total number of methylated peptides among the various experiments. Typical IAP-LC-MS/MS experiments with MMA antibodies generated ~1000 MMA sites in HCT116 cells, whereas the ADMA antibodies identified between 300 and 400 sites in HCT116 cells.

Motif and Protein Class Analysis of Arginine Methylation Sites—Motif analyses of peptides pulled down by Me-R^4-100 and R*GG (D5A12) antibodies are shown in Figs. 3*A* and 3*B*. In agreement with the existing literature, these identified methyl arginine residues resided in RG-rich sequences. 54.7% of sites that were enriched by Me-R^4-100 contained R*G sequences, and 70.4% of sites enriched by R*GG (D5A12) antibody contained R*G sequences, which reflects the antigen-sequence difference of the two antibodies.

Motif analyses for dimethyl arginine sites from D4H5 and D6A8 affinity enrichment are shown in Fig. 3*C* and supplemental Fig. S5. Again, RG-rich sequences are dominant (39.1%). Arginine methylation sites flanked by proline at −1 and +1 positions are also abundant, which is in agreement with Uhlmann *et al.*'s report (39), suggesting that our method is unbiased.

The overlap between the sites identified by four arginine methylation antibodies is shown in Venn diagrams in supplemental Fig. S6. These antibodies detected unique but also overlapping sites; this is reflected by their antibody specificity, which can be seen in the motif of sites identified by the antibodies (Fig. 3, supplemental Fig. S5). Using multiple antibodies to profile the same samples can increase the number of methylation sites identified.

The identified methylated proteins are abundant in several protein functional groups. Fig. 3*D* illustrates the protein class pie chart of monomethyl arginine sites, and Fig. 3*E* shows the protein class pie chart of asymmetric dimethyl arginine sites. Among the monomethyl arginine sites, close to 50% were found in proteins involved in RNA processing and transcriptional regulation, but of the dimethyl arginine sites, close to 70% were in proteins involved in transcriptional regulation and RNA processing. Figs. 3*F* and 3*G* show the statistically enriched protein classes of monomethyl and asymmetric dimethyl arginine-methylated proteins. The enriched terms for MMA and ADMA are very similar, but clearly, transcription factors and transcriptional regulator activity are more represented in ADMA proteins. Also, it is evident that most of the ADMA proteins are nuclear.

Protein Lysine Methylation Sites Identified in HCT116 Cells—Using Kme1 (BL10749), Kme2 (BL10745), and Kme3 (D3978) antibodies, in HCT116 cells, we identified in total 165 lysine methylation sites: 132 monomethyl K, 35 dimethyl K, and 31 trimethyl K (Fig. 2*B*, supplemental Table S4). Because lysine methylation datasets are quite small, we manually verified all spectra from these datasets to increase our confidence in the identification of these lysine methylation sites. Overall, there were far fewer lysine methylation sites than arginine methylation sites (Fig. 2*B*). Western blot analysis with lysine methylation antibodies did show fewer bands than arginine methylation blots, which might reflect fewer lysine methylation sites in the cells (Fig. 4*A*). To rule out the possibility that the lower number of lysine methylation site identifications was due to antibody affinity/specificity issues, we generated a diverse population of dimethylated peptides via reductive amination reaction (46) and used mix of dimethyl peptides and unlabeled peptides from HCT116 cells at a ratio of 1:9 (1 mg labeled:9 mg unlabeled) to do immunoaffinity enrichment using dimethyl lysine antibody (supplemental Fig. S7). From this experiment, we were able to identify 3857 dimethylated lysine peptides corresponding to 2763 unique sites. The fact that the antibody selectively enriched a diverse population of dimethylated sequences suggests that the small set of peptides identified *in vivo* from the HCT116 cells was not the result of antibody affinity/specificity.

In the lysine methylation site profiling experiment using tryptic peptides from HCT116 cells, we identified many known lysine methylation sites (Table I, supplemental Table S4), including histone H3 K28, K37, and K80 (corresponding to K27, K36, and K79 sites in the literature); Calm1 K116; and known sites in eEF1A. We also identified many related sites in different members of same-family proteins such as K663me of GRP94 and K607me of HSP90B, which was recently discovered and characterized (47), and K561me3 of HSP70 and K585me, -me2, and -me3 of GPR78. The identification of these known sites and highly related sites demonstrates the power of this immunoaffinity LC-MS/MS method in profiling lysine methylation sites. Indeed, we have identified many other novel lysine methylation sites (supplemental Table S4), including sites on components of the transcription complex such as transcription factors POLR2B and POLR3B. Interestingly, we also identified lysine methylation on several lysine methyl transferases including dimethylation of EZH2; mono-, di-, and trimethylation of SETDB1; and monomethylation of EZH1 (Table I, supplemental Table S4). The spectrum of EZH1 K736Me peptide is shown in Fig. 4*B*. Additional sample spectra of lysine-methylated peptides are shown in supplemental Fig. S4. To assess whether our method is useful for studying substrates of different lysine methyltransferases, we searched our identified sites against the consensus substrate motif of SET7/9 [KR][STA]K* (48, 49) and found many sites bearing this motif, among which were several transcription factors (Table I).

Motif analysis of Kme1 sites by Motif-X (50) showed that, with the exception of some selectivity of leucine at the −1 position, there are no other obvious motifs for lysine methylation sites (Fig. 4*C*).

Profiling Arginine Methylation of Mouse Brain and Mouse Embryo—We profiled protein arginine methylation in mouse brain (from 3-month-old mice) and mouse embryo (E16–17) via immunoaffinity enrichment using various arginine methyl-

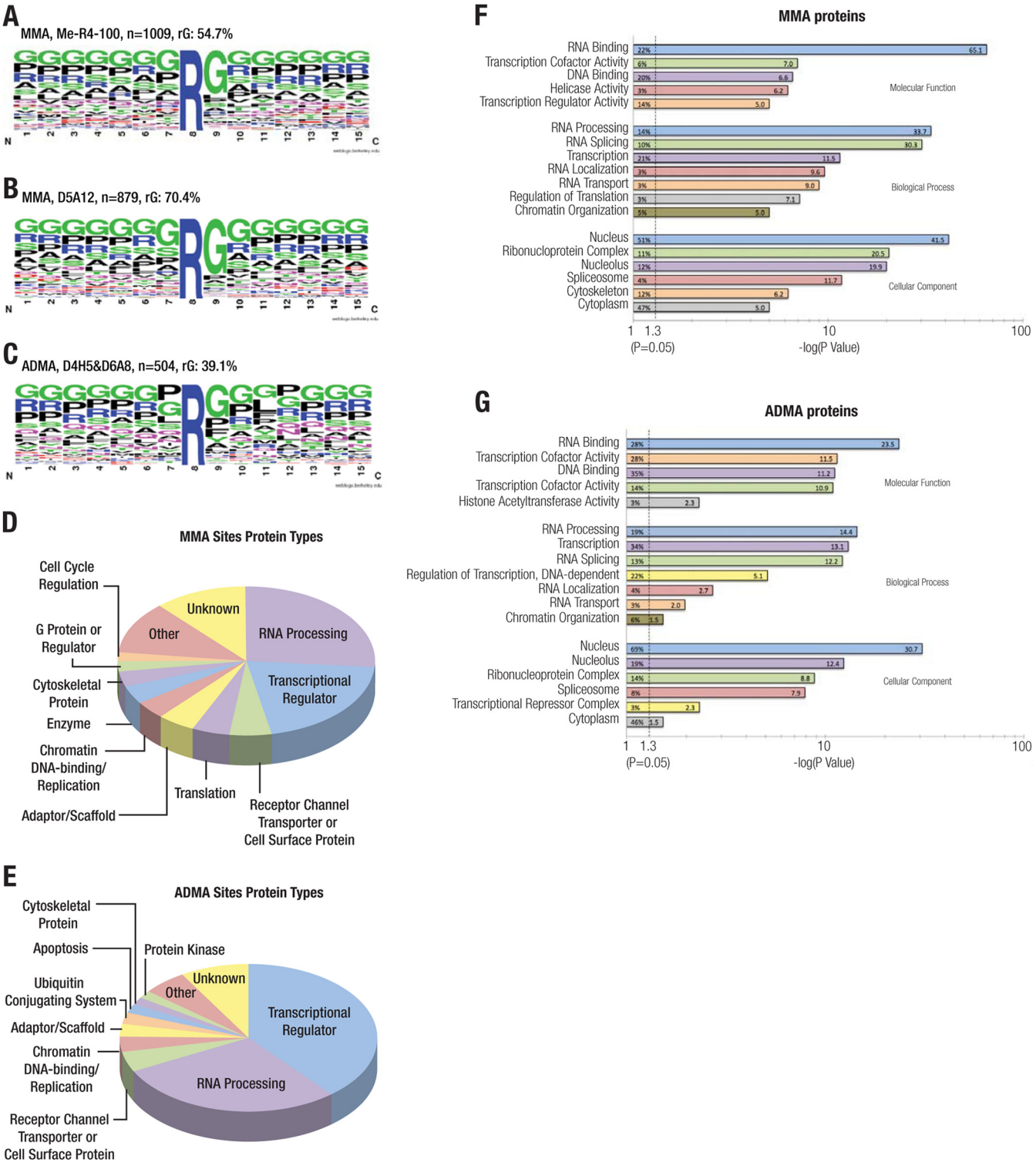


FIG. 3. **Motif and protein-type analysis of arginine methylation sites.** *A–C*, motif analysis of arginine methylation sites identified in HCT116 cells. The 15-mer sequences were generated by Motif X. Sites that were within seven residues of protein termini were not used in the motif logo. The frequency map was generated by Weblogo. *A*, of 1009 MMA sites identified using the Me-R^4-100 antibody, 54.7% had the rG motif. *B*, of 879 MMA sites identified with the D5A12 antibody, 70.4% had the rG motif. *C*, of the 466 ADMA sites identified by the D4H5 and D6A8 ADMA antibodies, 34.8% had the rG motif. *D*, *E*, pie charts of protein classes for MMA and ADMA sites identified in HCT116 cells based on the annotations in the PhosphoSite database. About half of the MMA sites are found in RNA processing proteins and transcriptional regulators. Transcriptional regulators account for the most ADMA sites (39%), followed by RNA processing proteins (29%). *F*, *G*, statistically enriched

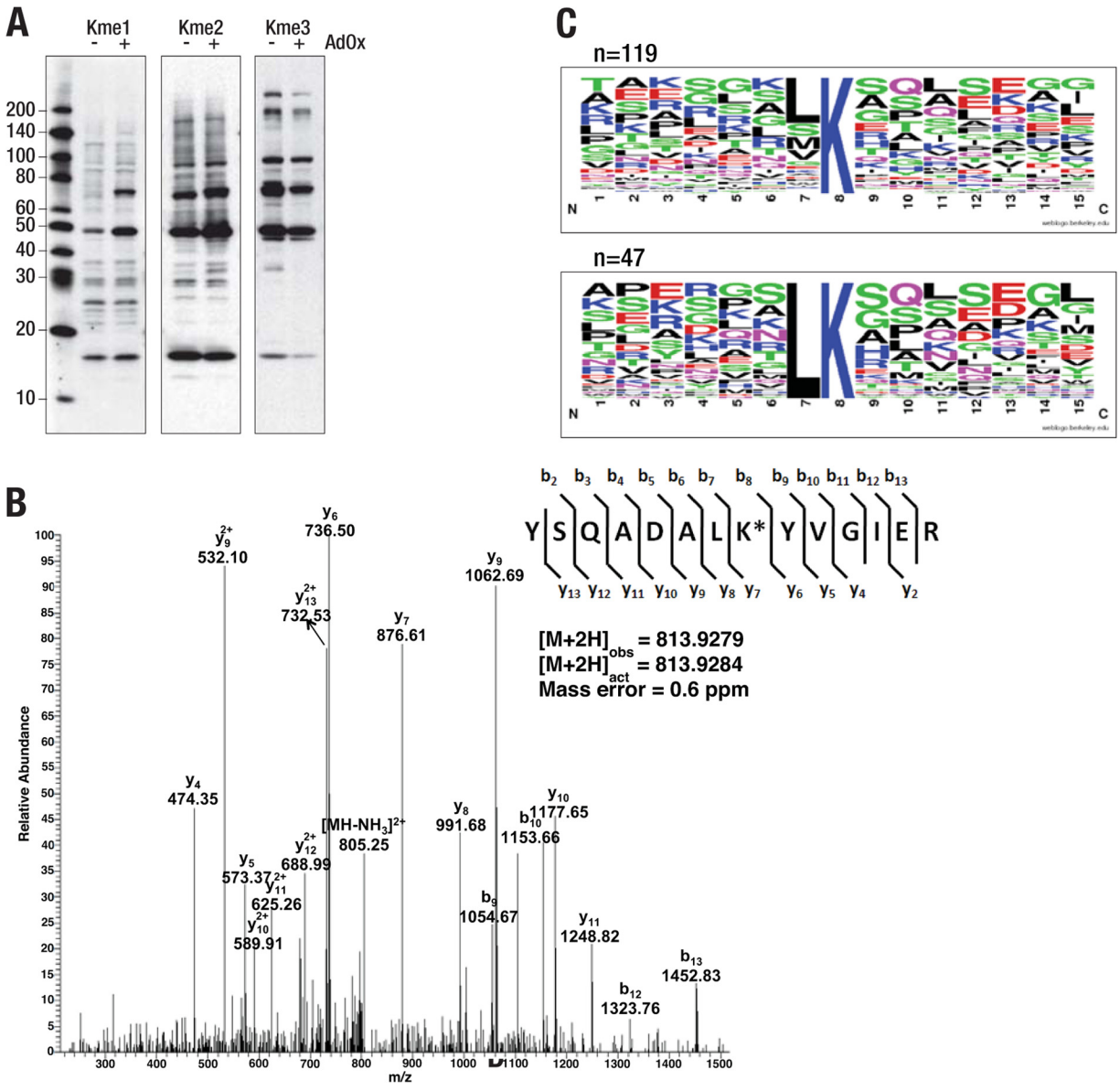


FIG. 4. **Profiling lysine methylation in HCT116 cells.** *A*, Western blot analysis of lysates from untreated or AdOx-treated (20 μM, 48 h) HCT116 cells by Kme1, Kme2, and Kme3 antibodies. *B*, MS/MS spectrum of monomethylated K736 of EZH1 (K735 of EZH2) identified from HCT116 cells using the Kme1 antibody. Site-determining ions y_6 and y_7 define the methylated residue. *C*, motif representation of 119 monomethyl lysine sites (upper panel) identified in HCT116 cells, of which 47 sites had leucine at the −1 position (lower panel).

ation antibodies and LC-MS/MS analysis. In total, we identified 807 MMA sites on 453 proteins from mouse brain and 598 MMA sites on 331 proteins from mouse embryo using a combination of D5A12 and Me-R[4]-1000 antibodies (Fig. 5*B*, supplemental Table S5), corresponding to a total of 1070 unique MMA sites from two tissues. Asymmetric dimethylated peptides were enriched with D4H5 and D6A8 antibodies or a combination of the two, and 697 ADMA sites on 321 proteins

monomethylated (MMA) protein classes and asymmetric dimethylated (ADMA) protein classes categorized by Gene Ontology terms. Significance is represented as the −log(*p* value); the significance threshold is 1.301 = −log(*p* = 0.05). Gene Ontology (GO) terms that were statistically overrepresented within each dataset were identified using DAVID. Selected GO terms that were significantly enriched in one or both datasets have been plotted. GO terms are listed on the vertical (*y*) axis; terms from each of the three major GO categories (molecular function, biological process, and cellular component) are grouped together. Within each major category, each GO term is depicted with the same bar color in both plots, to help in comparison between the two datasets. The length of each bar represents the statistical significance, plotted as the −log(*p* value); the number inside the end of each bar shows this value. The dashed line represents a significance threshold of 1.3 (*p* value = 0.05). The number inside the base of each bar indicates the percentage of genes within the dataset that have been annotated for the GO term.

TABLE I
A subset of representative lysine methylation sites identified in HCT116 cells

Protein	Sites and type	Peptide	Protein Type
	# = Known Sites		
Known sites and their homologs			
H3	# 28 (Me, 2Me, 3Me),	ATKAARKSAPATG	DNA binding protein
H3	# 37 (Me, 2Me, 3Me)	PSTGGVKKPHRYR	DNA binding protein
H3	# 80 (Me)	EIAQDFKTDLRFQ	DNA binding protein
Calm1	# 116 (Me, 2Me, 3Me)	MTNLGEKLTDEEV	Calcium-binding protein
eEF1A1	# 36 (Me, 3Me)	KCGGIDKRTIEKF	Translation elongation
eEF1A1	# 55 (Me, 2Me, 3Me)	MGKGSFKYAWVLD	Translation elongation
eEF1A1	# 79 (Me, 2Me, 3Me)	IDISLWKFETSKY	Translation elongation
eEF1A2	79 (Me, 2Me, 3Me)	IDISLWKFETTKY	Translation elongation
eEF1A1	# 165 (Me, 2Me, 3Me)	EPPYSQKRYEEIV	Translation elongation
eEF1A2	165 (Me, 2Me)	EPAYSEKRYDEIV	Translation elongation
EEF1A1	# 318 (2Me)	VKNVSVKDVRRGN	Translation elongation
eEF1A2	312 (2Me)	VKNVSVKDIRRGN	Translation elongation
Sites conserved in family members			
HSP90B	# 607 (Me)	NMERIMKAQALRD	chaperone
GRP94	663 (Me)	NMERIMKAQAYQT	chaperone
HSP70	561 (3Me)	DEGLKGKISEADK	chaperone
HSC70	561 (2Me)	DEKLQGKINDEDK	chaperone
GRP78	585 (Me, 2Me, 3Me)	KEKLGGKLSSEDK	chaperone
POLR2B	1052 (Me)	TYYQRLKHMVDDK	Transcription initiation complex
POLR3B	1013 (Me)	VYYQKLKHMVLDK	Transcription initiation complex
SLC25A5	52 (Me, 2Me)	TADKQYKGIIDCV	Membrane protein
SLC25A6	52 (3Me)	AADKQYKGIVDCI	Membrane protein
Sites from lysine methyltransferases			
EHMT1	153 (3Me)	LPGHAAKTLPGGA	lysine methyltransferase
EHMT2	189 (3Me)	ARKTMSKPGNGQP	lysine methyltransferase
EZH1	736 (Me)	SQADALKYVGIER	lysine methyltransferase
EZH2	510 (2Me)	WAAHCRKIQLKKD	lysine methyltransferase
EZH2	515 (2Me)	RKIQLKKDGSSNH	lysine methyltransferase
SETDB1	1170 (2Me, 3 Me)	TRGFALKSTHGIA	lysine methyltransferase
SETDB1	1178 (Me, 2Me, 3Me)	THGIAIKSTNMAS	lysine methyltransferase
SUZ12	403 (Me)	TQTIAVKESLTTD	lysine methyltransferase
Set7/9 substrate consensus site			
AKAP13	1670 (Me)	QICHRSKQQGFNY	Adaptor/scaffold
ZFP91	174 (Me)	TRSSRSKTGSLQL	C2H2-type zinc finger protein
STMN1	29 (Me)	ILSPRSKESVPEF	Cytoskeletal protein
ORC4L	7 (Me)	MSSRKSKSNSLIH	DNA binding
SH3PXD2B	619 (Me)	RPISKSKTDLPEE	Lipid binding protein
SLAIN2	553 (Me)	IPVPRSKLAQPVR	microtubule organization
cofilin 2	114 (Me)	SAPLKSKMIYASS	Transcriptional regulator
EDF1	25 (Me)	AAQAKSKQAILAA	Transcriptional regulator
GTF3C2	148 (Me)	KRGRKSKAELLLL	Transcriptional regulator
NCL	523 (Me)	NQNGKSKGYAFIE	Transcriptional regulator
ZAP3	808 (Me)	AEGTKSKWGMIPR	Transcriptional regulator
RPL29	5 (Me)	xxMAKSKNHTTHN	Translation

Note: Groups of shaded sites are homologous sites.
Literature-reported known site.

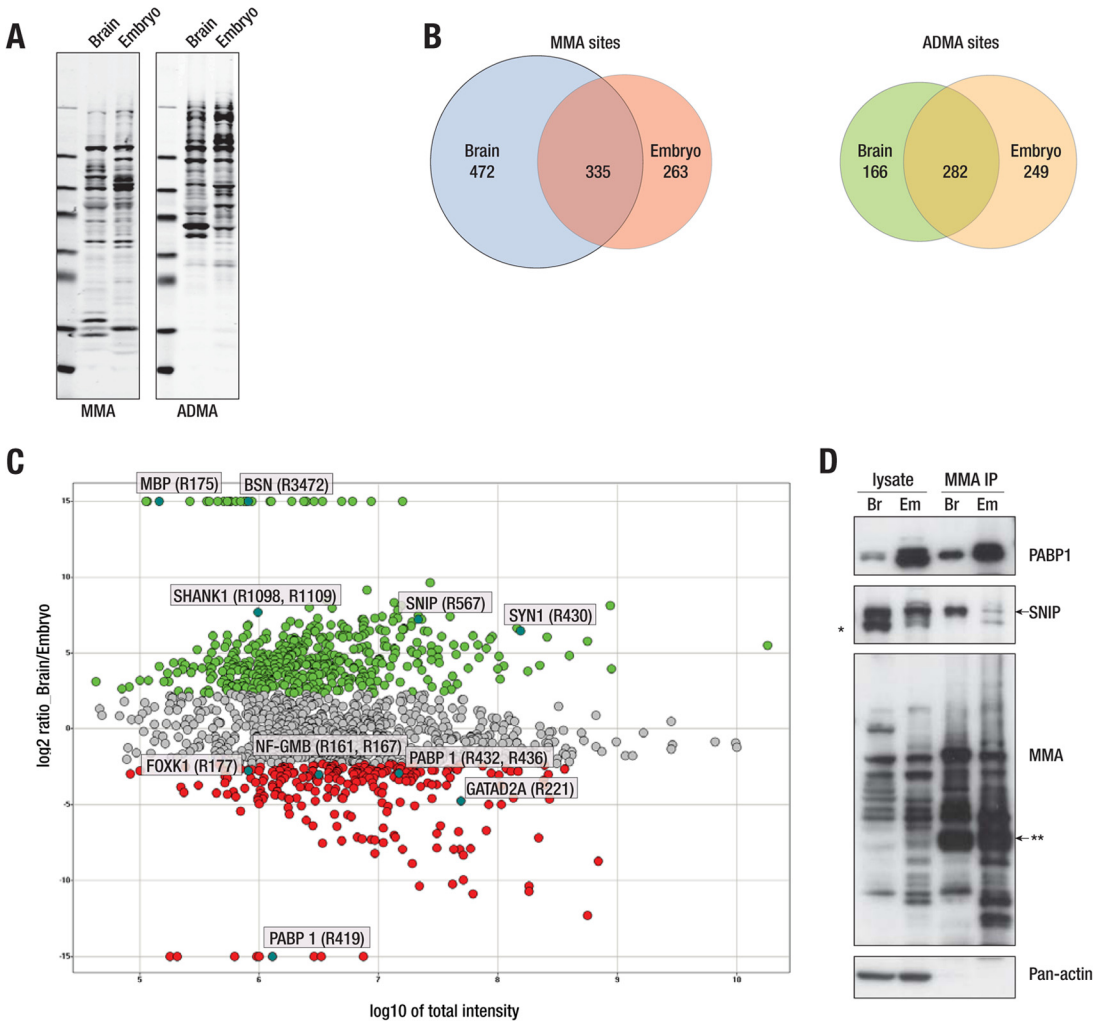


FIG. 5. **Quantitative analysis of arginine methylation in mouse brain and embryo.** *A*, Western blot analysis of protein extracts from mouse brain and mouse embryo by monomethyl arginine antibodies (MMA: mix of D5A12 and Me-R^4-100) and asymmetric dimethyl arginine antibodies (ADMA: mix of D4H5 and D6A8). *B*, Venn diagrams of the number of MMA and ADMA sites identified in mouse brain and mouse embryo. *C*, quantitative comparison of arginine monomethylation sites in two tissues. Each dot in the scatter plot represents a unique arginine monomethylation peptide identified in mouse brain and/or mouse embryo using the MMA antibody mixture. The *x*-axis is the log10 value of the total intensity of the representative peptide for a methylation site in mouse brain and embryo, and the *y*-axis shows the log2 ratio of the intensity of the peptide in mouse brain *versus* embryo. A cutoff of 5-fold was set to indicate increased arginine monomethylation peptide abundance in either brain (green dots) or embryo (red dots). For the methyl peptides that uniquely existed in a specific tissue, arbitrary log2 ratios of 15 (brain specific) and −15 (embryo specific) were assigned. Several representative brain (BSN, MBP, SYN1, SNIP, SHANK1) and embryo (GATAD2A, PABP1, FOXK1, NF-GMB) enriched methylated protein sites are highlighted and labeled on the graph. *D*, Western blot confirmation of methylation of SNIP and PABP1 in mouse brain and mouse embryo lysates. Total protein lysate from mouse brain and mouse embryo were immunoprecipitated by MMA antibody (mix of D5A12 and Me-R^4-100) and immunoblotted by SNIP, PABP1, MMA, and pan-actin antibodies. *Nonspecific band below SNIP in SNIP blot. **IgG band.

were identified (Fig. 5*B*, supplemental Table S6). MS/MS spectra of many proteins known to be brain-specific are identified in the brain samples but not in embryo as expected (supplemental Tables S5 and S6). Myelin basic protein (MBP) was among the first arginine-methylated proteins identified (51). Synapsin I (SYN1) is a neuronal phosphoprotein that associates with the cytoplasmic surface of synaptic vesicles and binds to the cytoskeleton. SHANK1 is a neuronal adaptor protein.

To demonstrate a quantitative assessment of the IAP-LC-MS/MS method for identifying and quantifying methyl proteins among samples, we first evaluated the technical variation of immunoprecipitation using monomethyl arginine antibodies in three independent IAPs. The experiments were performed using the same batch of mouse embryo peptides with a mixture of the two MMA antibodies Me-R^4-100 and D5A12. Each sample was processed separately, and the resulting enriched methylated peptides were analyzed via LC-

MS/MS in duplicate runs. In total, we identified 818 unique arginine monomethylation peptides, with over 66% of the sites identified in common among the three IAPs (supplemental Figs. S8*B* and S8*C*). Each site was quantified through label-free techniques using Skyline software. The median variations (percent coefficient of variation (%cv)) between replicate injections of matched methylated peptides for the three independent IAP samples were 14%, 13%, and 9%, respectively (supplemental Table S7). The variation associated with the IAP method was illustrated from the median %cv across three independent IAPs for the mouse embryo samples and was determined to be 13%. In addition, it is worth noting that 80% of the methylation sites that were quantified had %cv values of less than 20% (supplemental Table S7 and supplemental Fig. S8). The results from this experiment demonstrate that the IAP-LC-MS/MS method is quantitatively reproducible when experiments are performed in parallel and the LC-MS/MS analysis is collected in the same period of time.

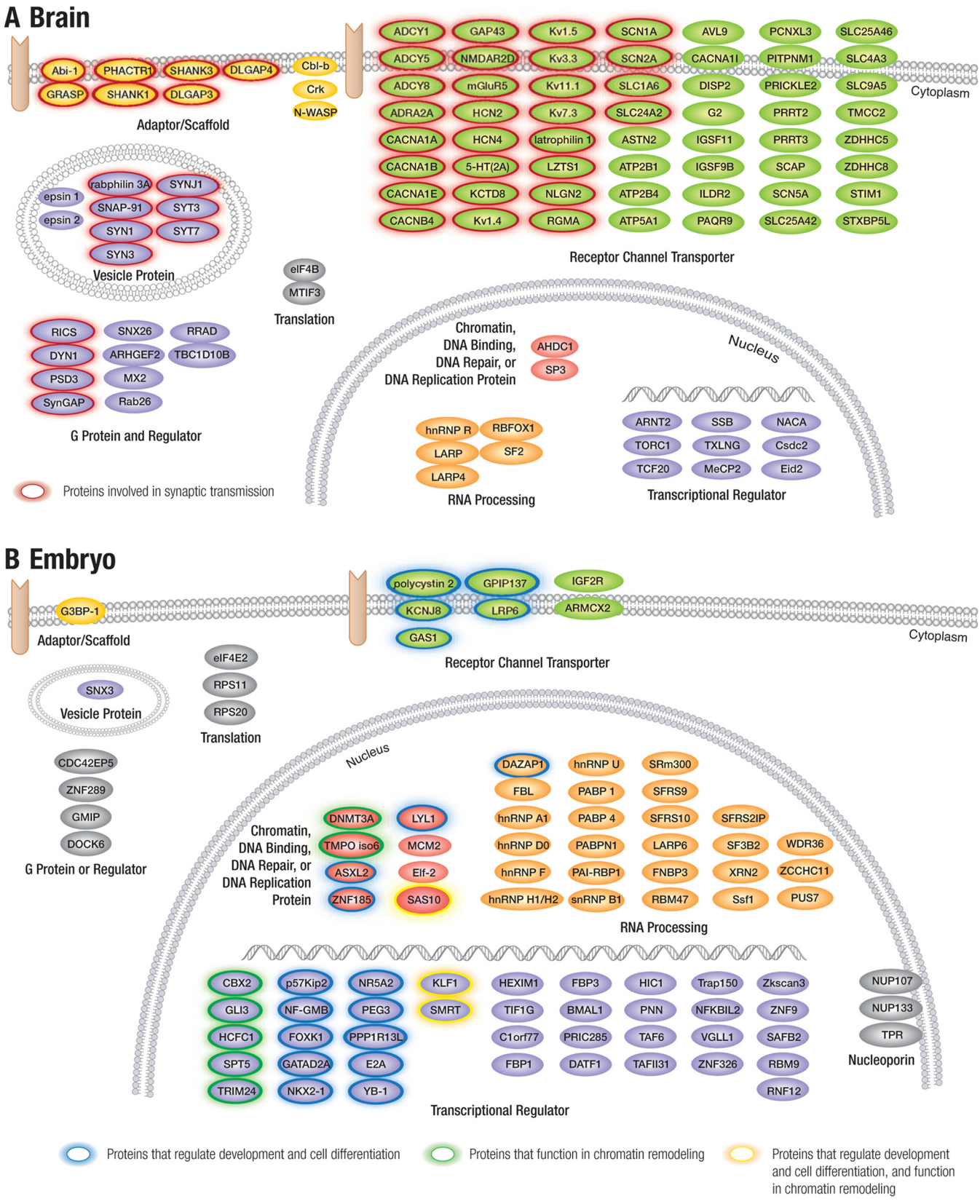
To gain a better understanding of the abundance of arginine-methylated proteins in different tissues, we analyzed the data from our MMA profiling of mouse brain and mouse embryo in biological triplicate experiments using Skyline software to compare the relative abundance of common methylation sites between the two tissues. In these experiments, a total of 10 mg of peptides from either mouse brain or mouse embryo were used in each IAP experiment, using the same amount of MMA antibodies. The IAPs were done in parallel at the same time, and each IAP sample was analyzed in duplicate LC-MS/MS runs. In total, there were six datasets from mouse brain and six datasets from mouse embryo (supplemental Tables S5 and S8). MS1 precursor intensities of methylated peptides were compared between corresponding MS runs from mouse brain and mouse embryo and averaged for each tissue to obtain the relative abundance of common methylation sites, which provided confident quantitative results of over 1000 unique arginine monomethylation sites from the tissues. We set an arbitrary fold change threshold of 5 as a quantitative cutoff to indicate whether a particular methylation peptide was more abundant in one tissue than the other. From this analysis, there were a total of 480 and 272 unique monomethylation peptides with high abundance in mouse brain and embryo, respectively, including 31 and 12 monomethylation peptides that existed in brain and embryo only, respectively (supplemental Table S8). A scatter plot of the log2 intensity ratio of brain/embryo of each methylation site *versus* the total peptide intensity was made to represent the peptides that were specific or rich in one particular mouse tissue (Fig. 5*C*). For statistic analysis, the *p* value was provided for each ratio by *t* test. We observed that because of the biological variation of mouse brain, many monomethylation peptides that were abundant in brain were associated with high *p* values. Three biological batches of mouse embryo tissues showed good consistency that made most of abundant monomethylation peptides in embryo with statistical significance (supplemental Table S8). The IAP experiment results are consistent with Western blot analyses of the lysates from the two tissues. The patterns of arginine methylation from the two tissues indicated that arginine methylation does vary in both directions for distinct protein bands (Figs. 5*A* and 5*D*).

To confirm some of our results, we did protein immunoprecipitation of mouse brain and embryo lysates using a mix of D5A12 and Me-R^4-100 MMA antibodies and used protein-specific antibodies to do Western blot analysis (Fig. 5*D*). PABP1 is more abundant in mouse embryo, and corresponding methylated PABP1 is also more abundant in mouse embryo. Interestingly, SNIP showed similar abundance in mouse brain and mouse embryo but was methylated only in mouse brain (Fig. 5*D*). This indicates some of the tissue-specific methylation we observed was due to different protein expression in different tissues, but some other tissue-specific methylation was due to the different levels of methylation on the relatively equally expressed proteins. Other types of validation studies such as Western blot analysis or total proteome profiling can be performed for important sites that have been identified by IAP-MS experiments. Nevertheless, our method could identify the high-level methylated proteins in given samples, and the methylation of these proteins could be important for protein function.

Arginine Methylated Proteins in Mouse Brain Versus Mouse Embryo—Further analysis of protein function/type was done for those arginine-methylated proteins specifically enriched in one tissue whose average peptide intensity was at least 5-fold abundant in the tissue as described in last section (Fig. 5*C*, supplemental Table S8). The results of the analysis showed distinct distribution patterns of protein function/type of enriched arginine-methylated proteins in brain and embryo. Of the brain enriched arginine-methylated proteins, many were receptor and membrane proteins, G proteins, and vesicle proteins, whereas transcription regulators and RNA processing proteins were overrepresented in mouse embryo. To visually show the difference, arginine-methylated proteins from eight protein classes that were specifically enriched in brain or embryo, including "receptor/channel/transporter," "adaptor/scaffold," "vesicle protein," "G protein and regulator," "translation," "chromatin/DNA binding," "DNA repair/replication proteins," "RNA processing," and "transcriptional regulator," are shown in Figs. 6*A* and 6*B* according to their type and localization. Over half of brain enriched arginine-methylated proteins were involved in synaptic transmission (Fig. 6*A*), whereas a significant portion of embryo enriched arginine-methylated proteins had functions in regulating tissue development and cell differentiation (Fig. 6*B*).

DISCUSSION

Methylation-specific Antibodies Enable Identification of in Vivo Methylation Sites in Proteomics Studies—Methylation of arginine and lysine protein residues has been known of for many years, yet because the field has lacked reliable tools


A Brain
Abi-1
PHACTR1
SHANK3
DLGAP4
GRASP
SHANK1
DLGAP3
Cbl-b
Crk
N-WASP
Adaptor/Scaffold
ADCY1
ADCY5
ADCY8
ADRA2A
CACNA1A
CACNA1B
CACNA1E
CACNB4
GAP43
NMDAR2D
mGluR5
HCN2
HCN4
5-HT(2A)
KCTD8
Kv1.4
Kv1.5
Kv3.3
Kv11.1
Kv7.3
Iatrophilin 1
LZTS1
NLGN2
RGMA
SCN1A
SCN2A
SLC1A6
SLC24A2
ASTN2
ATP2B1
ATP2B4
ATP5A1
AVL9
CACNA1I
DISP2
G2
IGSF11
IGSF9B
ILDR2
PAQR9
PCNXL3
PITPNM1
PRICKLE2
PRRT2
PRRT3
SCAP
SCN5A
SLC25A42
SLC25A46
SLC4A3
SLC9A5
TMCC2
ZDHHC5
ZDHHC8
STIM1
STXBP5L
Cytoplasm
Receptor Channel Transporter
epsin 1
epsin 2
rabphilin 3A
SNAP-91
SYN1
SYN3
SYNJ1
SYT3
SYT7
Vesicle Protein
eIF4B
MTIF3
Translation
RICS
DYN1
PSD3
SynGAP
SNX26
ARHGEF2
MX2
Rab26
RRAD
TBC1D10B
G Protein and Regulator
Chromatin,
DNA Binding,
DNA Repair, or
DNA Replication Protein
AHDC1
SP3
Nucleus
hnRNP R
LARP
LARP4
RBFOX1
SF2
RNA Processing
ARNT2
TORC1
TCF20
SSB
TXLNG
MeCP2
NACA
Csdc2
Eid2
Transcriptional Regulator
Proteins involved in synaptic transmission
B Embryo
G3BP-1
Adaptor/Scaffold
polycystin 2
KCNJ8
GAS1
GPIP137
LRP6
IGF2R
ARMCX2
Cytoplasm
Receptor Channel Transporter
SNX3
Vesicle Protein
eIF4E2
RPS11
RPS20
Translation
CDC42EP5
ZNF289
GMIP
DOCK6
G Protein or Regulator
Nucleus
Chromatin,
DNA Binding,
DNA Repair, or
DNA Replication
Protein
DNMT3A
TMPO iso6
ASXL2
ZNF185
LYL1
MCM2
Elf-2
SAS10
DAZAP1
FBL
hnRNP A1
hnRNP D0
hnRNP F
hnRNP H1/H2
hnRNP U
PABP 1
PABP 4
PABPN1
PAI-RBP1
snRNP B1
SRm300
SFRS9
SFRS10
LARP6
FNBP3
RBM47
SFRS2IP
SF3B2
XRN2
Ssf1
WDR36
ZCCHC11
PUS7
RNA Processing
CBX2
GLI3
HCFC1
SPT5
TRIM24
p57Kip2
NF-GMB
FOXK1
GATAD2A
NKX2-1
NR5A2
PEG3
PPP1R13L
E2A
YB-1
KLF1
SMRT
HEXIM1
TIF1G
C1orf77
FBP1
FBP3
BMAL1
PRIC285
DATF1
HIC1
PNN
TAF6
TAFII31
Trap150
NFKBIL2
VGLL1
ZNF326
Zkscan3
ZNF9
SAFB2
RBM9
RNF12
NUP107
NUP133
TPR
Nucleoporin
Transcriptional Regulator
Proteins that regulate development and cell differentiation
Proteins that function in chromatin remodeling
Proteins that regulate development and cell differentiation, and function in chromatin remodeling

and methods for studying these modifications, the role and function of protein methylation in cell biology is still poorly understood. The one exception is the epigenetic regulation of chromatin, in which histone lysine and arginine modifications play pivotal roles in the regulation of chromatin structure and gene expression. Although large families of both arginine and lysine methyltransferases and demethylase have been identified, we currently have very little knowledge of their substrates and overall function. Arginine methylation occurs in three different forms—monomethylation, asymmetric dimethylation, and symmetric dimethylation—and lysine methylation is found in mono-, di-, and trimethylated forms. Given the variety of lysine and arginine methylation substitutions, very little is currently known regarding the distribution and function of these modifications. We set out to develop new tools and high-throughput methods to identify and characterize the different forms and sites of arginine and lysine methylation at a proteomic scale.

With the antibodies and methods presented here, one can detect and quantify hundreds to thousands of methylated residues in a single experiment. This is a significant improvement over previous methods that are time consuming and identify few sites. A recent review summarizes some of these methods (52). A typical strategy usually begins by demonstrating that a recombinant PRMT or PKMT can methylate a recombinant protein substrate *in vitro* and then uses deletion mapping to identify the methylation regions. A recent paper reported 249 arginine methylations sites in 131 proteins (39), the largest list reported prior to our study. The study utilized two different proteases, three different chromatographic methods, and multiple MS runs for each sample to identify these methylation sites. This paper reported mostly DMA sites, and the analysis could not distinguish ADMA from SDMA modifications. An advantage of our approach is the development of antibodies that can distinguish between asymmetric and symmetric methylated arginine (Fig. 1 and supplemental Fig. S2) allowing enrichment for ADMA modifications. Our IAP-LC-MS/MS method is compatible with all standard quantitative proteomics techniques, including stable isotope labeling of amino acids in cell culture, reductive dimethyl labeling (excluding dimethyl lysine profiling), tandem mass tag/iTRAQ labeling, and label-free quantification methods. In this study, we demonstrated that the immunoaffinity enrichment method for methylated peptides is quantitative with a low degree of variation among replicate mouse embryo samples (13% cv; supplemental Fig. S8 and supplemental Table S7) and is suitable for large-scale quantitative analysis and discovery research.

Protein Lysine Methylation—Our current understanding of lysine methylation is largely restricted to histones, where it plays a critical role in epigenetic gene regulation. Lysine methyltransferases are components of large chromatin remodeling complexes such as the polycomb complex, and components such as EZH2 are frequently mutated in cancer. The large number of lysine methyltransferases suggests the importance of this modification in evolution, yet very few sites of lysine methylation have been identified. Using the antibody reagents described in this study for IAP-LC-MS/MS analysis, we identified thousands of arginine methylation sites but identified far fewer lysine methylation sites. Whereas our findings indicate that arginine methylation occurs on many different classes of proteins, we find that lysine methylation appears to be restricted to a much smaller group of proteins including histones, translation elongation factors eEF1A and eEF2A, and chaperone proteins such as HSP70. For the following reasons, we believe that this finding likely reflects what happens in cells, and is not an artifact of the restricted specificity of the antibodies used to identify methylated lysine residues. First, several different antibodies, each with broad selectivity against methylated lysine, gave similar results; and second, we have shown that the antibodies used in this study were capable of enriching thousands of different lysine-methylated sequences when they were produced via reductive dimethylation, which chemically methylates all lysine residues. In experiments involving both cell lines and mouse tissue, IAP-LC-MS/MS typically identified fewer than 100 lysine methylation sites, suggesting that this modification is restricted to a small subset of proteins. Our findings are in agreement with earlier work using chemical methods that detected far less lysine methylation than arginine methylation in PC12 cells (53). Taken together, these results suggest that the large families of lysine methyltransferases and demethylases selectively regulate a small number of important substrates, consistent with the critical role for lysine modifications in histone and chromatin function during development and cell differentiation (8, 54).

Surveying Arginine Methylation in Brain and Embryo—We performed an affinity enrichment and LC-MS/MS analysis and identified more than 1000 sites of arginine methylation from mouse brain and embryo to demonstrate that the method could be used for discovery-based proteomic studies. We further did some quantitative comparison of abundant groups of arginine-methylated proteins in mouse brain *versus* mouse embryo.

Proteins involved in RNA processing and transcriptional regulation represented the largest group of enriched arginine-

FIG. 6. **Schematic representation of brain (*A*) and embryo (*B*) enriched monomethyl arginine proteins according to protein type and localization.** The intensity ratio for a methylated protein was calculated by averaging the intensity ratios of all sites belonging to the same protein. A cutoff of 5-fold was set to indicate a significantly increased abundance of an arginine monomethylated protein in either brain (*A*) or embryo (*B*). Only proteins from "receptor/channel/transporter," "adaptor/scaffold," "vesicle protein," "G protein and regulator," "translation," "chromatin/DNA binding," "DNA repair/replication proteins," "RNA processing," and "transcriptional regulator" protein classes are represented in the figure. Circles are sized to accommodate the protein name and do not correlate with the fold change.

methylated proteins in mouse embryo. Many of the identified transcription factors function in regulating tissue development and cell differentiation and are involved in epigenetic regulation of gene expression (Fig. 6*B*). Arginine methyltransferases are known to play an essential role in embryo development, as PRMT1 and CARM1 null embryos are smaller in size and die early (15, 16). Arginine methylation's regulation of the cell fate and pluripotency of early mouse embryo was first noted in histones (55). Our large-scale proteomic results expanded the variety of sites regulated by arginine methylation to many non-histone proteins in embryo. We have identified several known non-histone arginine-methylated proteins responsible for transcriptional regulation, including CBX2 and 53BP1. In addition, we also identified many transcription factors that were not previously known to be arginine methylated, such as SMRT, NF-GMB, HCFC1, FOXK11, GATAD2A, etc. (Fig. 6*B*, supplemental Tables S5–S8). Our results expand arginine methylation to many non-histone proteins in embryo and provide many new nodes and clues to connect arginine methylation and embryo development.

In contrast to arginine methylation identified in mouse embryo, enriched arginine-methylated proteins in mouse brain showed a diverse distribution of protein functions. Interestingly, ion channels and transporter proteins represented the most abundantly methylated class of proteins in brain (Fig. 6*A*, supplemental Table S8). Among these, we identified several ion channels, including voltage-dependent channels (Ca^{2+}), hyperpolarization-activated cyclic nucleotide-gated channels (K^+/Na^+), and voltage-gated channels (K^+, Na^+). Although little is known regarding the role of arginine methylation in ion channel function, Beltran-Alvarez *et al.* recently reported arginine methylation at two sites on the cardiac sodium ion channel that are mutated in sodium channel disorders (56). Our findings indicate a larger role and expanded scope for PRMT regulation of the superfamily of ion channels. We also identified many sites in vesicle proteins suggesting a role in regulating protein trafficking or secretion in the brain. Overall, many of the molecules we identified as methylated in brain function in synaptic transmission, including receptors, channel proteins, adaptor/scaffold proteins, vesicle proteins, and G proteins (Fig. 6*A*). Our finding extends the known main function of protein arginine methylation (*i.e.* regulating RNA processing and transcription) to possible involvement in synaptic regulation, which resembles the scenario in which the identification of many O-GlcNAcylation sites in many synaptic proteins increased our knowledge of the new function of this posttranslational modification (57, 58). Among all the PRMTs, PRMT8 is solely expressed in brain and localizes to the plasma membrane via N-terminal myristoylation (59), suggesting that many of the proteins we have identified may be targets of PRMT8. It is tempting to speculate that PRMT8 methylates many of these receptor, channel, transporter, and adaptor proteins and that methylation is important for their function in synaptic transmission.

Application of Immunoaffinity Purification LC-MS/MS Analysis of Protein Methylation and Future Development—As demonstrated here, our IAP-LC-MS/MS method enables the large-scale identification and quantitation of protein methylation and will have broad applications in the study of the involvement of protein methylation in physiological and pathological processes. An essential component of this method is the availability of methylation-specific antibodies. Two of our monomethyl arginine antibodies identified overlapping and antibody-specific methylation peptides. The same is true for the two dimethyl arginine antibodies. This indicates that different antibodies can assess overlapping but distinct portions of the methylated proteome. We anticipate that the development of additional methyl arginine and methyl lysine antibodies with different specificities will help to broaden and functionally characterize our coverage of the methylome. Future studies include the development of antibodies against SDMA, which were not used in the current study, to help characterize differences between ADMA and SDMA modifications. One can also envision combining antibody enrichment with orthogonal fractionation techniques such as strong cation exchange or hydrophilic interaction liquid chromatography (HILIC) chromatography to allow a deeper analysis of the accessible methylome.

Acknowledgments—We thank Dr. Matthew Stokes for critical comments on the manuscript. We thank Rachel Matthews, Aileen Taylor, and David Comb for graphical assistance. We thank Dr. Bin Zhang and Dr. Peter Hornbeck for getting the publication status of all the methylation sites.

S This article contains supplemental material.

¶ To whom correspondence should be addressed: E-mail: mcomb@cellsignal.com.

‡ These authors contributed to this work equally.

REFERENCES

1. Aletta, J. M., Cimato, T. R., and Ettinger, M. J. (1998) Protein methylation: a signal event in post-translational modification. *Trends Biochem. Sci.* **23,** 89–91
2. Comb, D. G., Sarkar, N., and Pinzino, C. J. (1966) The methylation of lysine residues in protein. *J. Biol. Chem.* **241,** 1857–1862
3. Bedford, M. T., and Clarke, S. G. (2009) Protein arginine methylation in mammals: who, what, and why. *Mol. Cell* **33,** 1–13
4. Bedford, M. T., and Richard, S. (2005) Arginine methylation: an emerging regulator of protein function. *Mol. Cell* **18,** 263–272
5. Bedford, M. T. (2007) Arginine methylation at a glance. *J. Cell Sci.* **120,** 4243–4246
6. Dillon, S. C., Zhang, X., Trievel, R. C., and Cheng, X. (2005) The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol.* **6,** 227
7. Zhang, X., and Bruice, T. C. (2007) Histone lysine methyltransferase SET7/9: formation of a water channel precedes each methyl transfer. *Biochemistry* **46,** 14838–14844
8. Lee, D. Y., Teyssier, C., Strahl, B. D., and Stallcup, M. R. (2005) Role of protein methylation in regulation of transcription. *Endocrine Rev.* **26,** 147–170
9. Hyun, Y. L., Lew, D. B., Park, S. H., Kim, C. W., Paik, W. K., and Kim, S. (2000) Enzymic methylation of arginyl residues in -gly-arg-gly- peptides. *Biochem. J.* **348 Pt 3,** 573–578
10. Cheng, D., Cote, J., Shaaban, S., and Bedford, M. T. (2007) The arginine methyltransferase CARM1 regulates the coupling of transcription and

mRNA processing. *Mol. Cell* **25,** 71–83
11. Trojer, P., and Reinberg, D. (2006) Histone lysine demethylases and their impact on epigenetics. *Cell* **125,** 213–217
12. Anand, R., and Marmorstein, R. (2007) Structure and mechanism of lysine-specific demethylase enzymes. *J. Biol. Chem.* **282,** 35425–35429
13. Liu, F., Zhao, X., Perna, F., Wang, L., Koppikar, P., Abdel-Wahab, O., Harr, M. W., Levine, R. L., Xu, H., Tefferi, A., Deblasio, A., Hatlen, M., Menendez, S., and Nimer, S. D. (2011) JAK2V617F-mediated phosphorylation of PRMT5 downregulates its methyltransferase activity and promotes myeloproliferation. *Cancer Cell* **19,** 283–294
14. Higashimoto, K., Kuhn, P., Desai, D., Cheng, X., and Xu, W. (2007) Phosphorylation-mediated inactivation of coactivator-associated arginine methyltransferase 1. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 12318–12323
15. Pawlak, M. R., Scherer, C. A., Chen, J., Roshon, M. J., and Ruley, H. E. (2000) Arginine N-methyltransferase 1 is required for early postimplantation mouse development, but cells deficient in the enzyme are viable. *Mol. Cell. Biol.* **20,** 4859–4869
16. Yadav, N., Lee, J., Kim, J., Shen, J., Hu, M. C., Aldaz, C. M., and Bedford, M. T. (2003) Specific protein methylation defects and gene expression perturbations in coactivator-associated arginine methyltransferase 1-deficient mice. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 6464–6468
17. Dambacher, S., Hahn, M., and Schotta, G. (2010) Epigenetic regulation of development by histone lysine methylation. *Heredity* **105,** 24–37
18. Greer, E. L., and Shi, Y. (2012) Histone methylation: a dynamic mark in health, disease and inheritance. *Nat. Rev. Genetics* **13,** 343–357
19. Yang, Y., and Bedford, M. T. (2013) Protein arginine methyltransferases and cancer. *Nat. Rev. Cancer* **13,** 37–50
20. Majumder, S., Liu, Y., Ford, O. H., 3rd, Mohler, J. L., and Whang, Y. E. (2006) Involvement of arginine methyltransferase CARM1 in androgen receptor function and prostate cancer cell viability. *Prostate* **66,** 1292–1301
21. Frietze, S., Lupien, M., Silver, P. A., and Brown, M. (2008) CARM1 regulates estrogen-stimulated breast cancer growth through up-regulation of E2F1. *Cancer Res.* **68,** 301–306
22. Scorilas, A., Black, M. H., Talieri, M., and Diamandis, E. P. (2000) Genomic organization, physical mapping, and expression analysis of the human protein arginine methyltransferase 1 gene. *Biochem. Biophys. Res. Commun.* **278,** 349–359
23. Mathioudaki, K., Papadokostopoulou, A., Scorilas, A., Xynopoulos, D., Agnanti, N., and Talieri, M. (2008) The PRMT1 gene expression pattern in colon cancer. *Br. J. Cancer* **99,** 2094–2099
24. Kang, M. Y., Lee, B. B., Kim, Y. H., Chang, D. K., Kyu Park, S., Chun, H. K., Song, S. Y., Park, J., and Kim, D. H. (2007) Association of the SUV39H1 histone methyltransferase with the DNA methyltransferase 1 at mRNA expression level in primary colorectal cancer. *Int. J. Cancer* **121,** 2192–2197
25. Simon, J. A., and Lange, C. A. (2008) Roles of the EZH2 histone methyltransferase in cancer epigenetics. *Mutat. Res.* **647,** 21–29
26. Tsang, D. P., and Cheng, A. S. (2011) Epigenetic regulation of signaling pathways in cancer: role of the histone methyltransferase EZH2. *J. Gastroenterol. Hepatol.* **26,** 19–27
27. Yap, D. B., Chu, J., Berg, T., Schapira, M., Cheng, S. W., Moradian, A., Morin, R. D., Mungall, A. J., Meissner, B., Boyle, M., Marquez, V. E., Marra, M. A., Gascoyne, R. D., Humphries, R. K., Arrowsmith, C. H., Morin, G. B., and Aparicio, S. A. (2011) Somatic mutations at EZH2 Y641 act dominantly through a mechanism of selectively altered PRC2 catalytic activity, to increase H3K27 trimethylation. *Blood* **117,** 2451–2459
28. Yost, J. M., Korboukh, I., Liu, F., Gao, C., and Jin, J. (2011) Targets in epigenetics: inhibiting the methyl writers of the histone code. *Curr. Chem. Genomics* **5,** 72–84
29. Cheng, D., Yadav, N., King, R. W., Swanson, M. S., Weinstein, E. J., and Bedford, M. T. (2004) Small molecule regulators of protein arginine methyltransferases. *J. Biol. Chem.* **279,** 23892–23899
30. Rush, J., Moritz, A., Lee, K. A., Guo, A., Goss, V. L., Spek, E. J., Zhang, H., Zha, X. M., Polakiewicz, R. D., and Comb, M. J. (2005) Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat. Biotechnol.* **23,** 94–101
31. Villen, J., Beausoleil, S. A., Gerber, S. A., and Gygi, S. P. (2007) Large-scale phosphorylation analysis of mouse liver. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 1488–1493
32. Macek, B., Mann, M., and Olsen, J. V. (2009) Global and site-specific quantitative phosphoproteomics: principles and applications. *Annu. Rev. Pharmacol. Toxicol.* **49,** 199–221
33. Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V., and Mann, M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325,** 834–840
34. Kim, W., Bennett, E. J., Huttlin, E. L., Guo, A., Li, J., Possemato, A., Sowa, M. E., Rad, R., Rush, J., Comb, M. J., Harper, J. W., and Gygi, S. P. (2011) Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol. Cell* **44,** 325–340
35. Emanuele, M. J., Elia, A. E., Xu, Q., Thoma, C. R., Izhar, L., Leng, Y., Guo, A., Chen, Y. N., Rush, J., Hsu, P. W., Yen, H. C., and Elledge, S. J. (2011) Global identification of modular cullin-RING ligase substrates. *Cell* **147,** 459–474
36. Boisvert, F. M., Cote, J., Boulanger, M. C., and Richard, S. (2003) A proteomic analysis of arginine-methylated protein complexes. *Mol. Cell. Proteomics* **2,** 1319–1330
37. Ong, S. E., Mittler, G., and Mann, M. (2004) Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. *Nat. Methods* **1,** 119–126
38. Pang, C. N., Gasteiger, E., and Wilkins, M. R. (2010) Identification of arginine- and lysine-methylation in the proteome of *Saccharomyces cerevisiae* and its functional implications. *BMC Genomics* **11,** 92
39. Uhlmann, T., Geoghegan, V. L., Thomas, B., Ridlova, G., Trudgian, D. C., and Acuto, O. (2012) A method for large-scale identification of protein arginine methylation. *Mol. Cell. Proteomics* **11,** 1489–1499
40. Huttlin, E. L., Jedrychowski, M. P., Elias, J. E., Goswami, T., Rad, R., Beausoleil, S. A., Villen, J., Haas, W., Sowa, M. E., and Gygi, S. P. (2010) A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* **143,** 1174–1189
41. Beausoleil, S. A., Villen, J., Gerber, S. A., Rush, J., and Gygi, S. P. (2006) A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat. Biotechnol.* **24,** 1285–1292
42. Schilling, B., Rardin, M. J., MacLean, B. X., Zawadzka, A. M., Frewen, B. E., Cusack, M. P., Sorensen, D. J., Bereman, M. S., Jing, E., Wu, C. C., Verdin, E., Kahn, C. R., Maccoss, M. J., and Gibson, B. W. (2012) Platform-independent and label-free quantitation of proteomic data using MS1 extracted ion chromatograms in skyline: application to protein acetylation and phosphorylation. *Mol. Cell. Proteomics* **11,** 202–214
43. Zhang, H., Zha, X., Tan, Y., Hornbeck, P. V., Mastrangelo, A. J., Alessi, D. R., Polakiewicz, R. D., and Comb, M. J. (2002) Phosphoprotein analysis using antibodies broadly reactive against phosphorylated motifs. *J. Biol. Chem.* **277,** 39379–39387
44. Dhar, S., Vemulapalli, V., Patananan, A. N., Huang, G. L., Di Lorenzo, A., Richard, S., Comb, M. J., Guo, A., Clarke, S. G., and Bedford, M. T. (2013) Loss of the major Type I arginine methyltransferase PRMT1 causes substrate scavenging by other PRMTs. *Sci. Rep.* **3,** 1311
45. Moritz, A., Li, Y., Guo, A., Villen, J., Wang, Y., MacNeill, J., Kornhauser, J., Sprott, K., Zhou, J., Possemato, A., Ren, J. M., Hornbeck, P., Cantley, L. C., Gygi, S. P., Rush, J., and Comb, M. J. (2010) Akt-RSK-S6 kinase signaling networks activated by oncogenic receptor tyrosine kinases. *Sci. Signal.* **3,** ra64
46. Hsu, J. L., Huang, S. Y., Chow, N. H., and Chen, S. H. (2003) Stable-isotope dimethyl labeling for quantitative proteomics. *Anal. Chem.* **75,** 6843–6852
47. Donlin, L. T., Andresen, C., Just, S., Rudensky, E., Pappas, C. T., Kruger, M., Jacobs, E. Y., Unger, A., Zieseniss, A., Dobenecker, M. W., Voelkel, T., Chait, B. T., Gregorio, C. C., Rottbauer, W., Tarakhovsky, A., and Linke, W. A. (2012) Smyd2 controls cytoplasmic lysine methylation of Hsp90 and myofilament organization. *Genes Dev.* **26,** 114–119
48. Dhayalan, A., Kudithipudi, S., Rathert, P., and Jeltsch, A. (2011) Specificity analysis-based identification of new methylation targets of the SET7/9 protein lysine methyltransferase. *Chem. Biol.* **18,** 111–120
49. Couture, J. F., Collazo, E., Hauk, G., and Trievel, R. C. (2006) Structural basis for the methylation site specificity of SET7/9. *Nat. Struct. Mol. Biol.* **13,** 140–146
50. Schwartz, D., and Gygi, S. P. (2005) An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nat. Biotechnol.* **23,** 1391–1398
51. Brostoff, S., and Eylar, E. H. (1971) Localization of methylated arginine in the A1 protein from myelin. *Proc. Natl. Acad. Sci. U.S.A.* **68,** 765–769
52. Cheng, D., Vemulapalli, V., and Bedford, M. T. (2012) Methods applied to the study of protein arginine methylation. *Methods Enzymol.* **512,**

71–92
53. Najbauer, J., and Aswad, D. W. (1990) Diversity of methyl acceptor proteins in rat pheochromocytoma (PC12) cells revealed after treatment with adenosine dialdehyde. *J. Biol. Chem.* **265,** 12717–12721
54. Sims, R. J., 3rd, Nishioka, K., and Reinberg, D. (2003) Histone lysine methylation: a signature for chromatin function. *Trends Genet.* **19,** 629–639
55. Torres-Padilla, M. E., Parfitt, D. E., Kouzarides, T., and Zernicka-Goetz, M. (2007) Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature* **445,** 214–218
56. Beltran-Alvarez, P., Pagans, S., and Brugada, R. (2011) The cardiac sodium channel is post-translationally modified by arginine methylation. *J. Proteome Res.* **10,** 3712–3719
57. Vosseller, K., Trinidad, J. C., Chalkley, R. J., Specht, C. G., Thalhammer, A., Lynn, A. J., Snedecor, J. O., Guan, S., Medzihradszky, K. F., Maltby, D. A., Schoepfer, R., and Burlingame, A. L. (2006) O-linked N-acetylglucosamine proteomics of postsynaptic density preparations using lectin weak affinity chromatography and mass spectrometry. *Mol. Cell. Proteomics* **5,** 923–934
58. Tallent, M. K., Varghis, N., Skorobogatko, Y., Hernandez-Cuebas, L., Whelan, K., Vocadlo, D. J., and Vosseller, K. (2009) In vivo modulation of O-GlcNAc levels regulates hippocampal synaptic plasticity through interplay with phosphorylation. *J. Biol. Chem.* **284,** 174–181
59. Lee, J., Sayegh, J., Daniel, J., Clarke, S., and Bedford, M. T. (2005) PRMT8, a new membrane-bound tissue-specific member of the protein arginine methyltransferase family. *J. Biol. Chem.* **280,** 32890–32896