Cell
PRESS

Arginylation and Methylation Double Up to Regulate Nuclear Proteins and Nuclear Architecture In Vivo

Sougata Saha,[1,5] Catherine C.L. Wong,[3,5] Tao Xu,[3,5] Suk Namgoong,[2] Henry Zebroski,[4] John R. Yates, III,[3] and Anna Kashina[1,*]
[1]Department of Animal Biology, School of Veterinary Medicine
[2]Department of Physiology, School of Medicine
University of Pennsylvania, Philadelphia, PA 19104, USA
[3]The Scripps Research Institute, San Diego, CA 92121, USA
[4]The Rockefeller University, New York, NY 10065, USA
[5]These authors contributed equally to this work
*Correspondence: akashina@vet.upenn.edu
DOI 10.1016/j.chembiol.2011.08.019

SUMMARY

Protein arginylation and arginine methylation are two posttranslational modifications of emerging importance that involve Arg residues and their modifications. To test a hypothesis that posttranslationally added arginines can be methylated, we used high-precision mass spectrometry and metabolic labeling to find whether posttranslationally added arginines can serve as methylation sites. We identified a number of proteins in vivo, on which posttranslationally added Arg have undergone mono- and dimethylation. This double modification predominantly affects the chromatin-containing nuclear fraction and likely plays an important regulatory role in chromatin-associated proteins. Moreover, inhibition of arginylation and Arg methylation results in a significant reduction of the nucleus size in cultured cells, suggesting changes in chromatin compaction and nuclear architecture. Our findings suggest a functional link between protein regulation by arginylation and methylation that affects nuclear structure in vivo.

INTRODUCTION

Posttranslational modifications regulate the majority of protein functions and physiological pathways and add an extra level of complexity to the structural and functional diversity of the proteome. Many modifications that have been discovered decades ago are now coming into the spotlight, as new studies show evidence of their major physiological role. One such modification is arginine methylation, recently demonstrated to regulate a variety of processes, including protein-protein interactions, RNA binding, transcription, signal transduction, and DNA repair (for review see Bedford and Richard, 2005). Another modification, whose role has emerged in recent studies, is arginylation—posttranslational addition of Arg onto proteins that has been shown to affect many proteins in vivo (Wong et al., 2007) and modulate critical physiological processes such as cell motility and cardiovascular development (Karakozova et al., 2006; Kwon et al., 2002).

Although no link between the two modifications has been previously suggested, a natural question arises of whether posttranslationally added Arg can serve as a site for the action of protein Arg-methyltransferases. Here, we addressed this question by examining posttranslational arginylation and methylation in subcellular fractions and found that posttranslationally arginylated proteins can be methylated on added arginines and that this double modification occurs with higher frequency in the nuclear compared to the cytosolic proteins, due likely to the higher nuclear activity of arginylation and methylation enzymes. The identity of the modified proteins and the location of arginylation and arginylation/methylation sites suggest that these modifications in the nucleus are involved in the global regulation of chromatin structure, alternative splicing, and gene expression. Furthermore, our studies in cultured cells showed that inhibition of arginylation and/or Arg methylation results in a significant reduction of the nuclear size, suggesting that these two modifications are essential for maintaining normal chromatin compaction and nuclear architecture. Our results are the first demonstration of posttranslational modifications doubled up on the same sites to regulate an important subset of proteins in vivo.

RESULTS

Posttranslationally Added Arginines Can Be Methylated In Vivo

To address the possible relationship between posttranslational arginylation and Arg methylation, we examined the samples isolated from subcellular fractions of mouse tissues and cultured cells by high-precision mass spectrometry to find out whether any of the posttranslationally added Arg are found in the methylated state (i.e., contain a mass addition of Arg [+156.1011] with an extra mass of either one [+14.0157] or two [+28.0314] methyl groups).

Analysis of the cytosolic fractions derived from cultured embryonic fibroblasts and mouse heart tissue revealed two peptides in different samples, in which an internal site within the molecule contained dimethylated Arg (mass addition +184.1325). One peptide belongs to α-cardiac actin (identified in a 2D gel spot corresponding to α-actin from the mouse heart Figure 1),
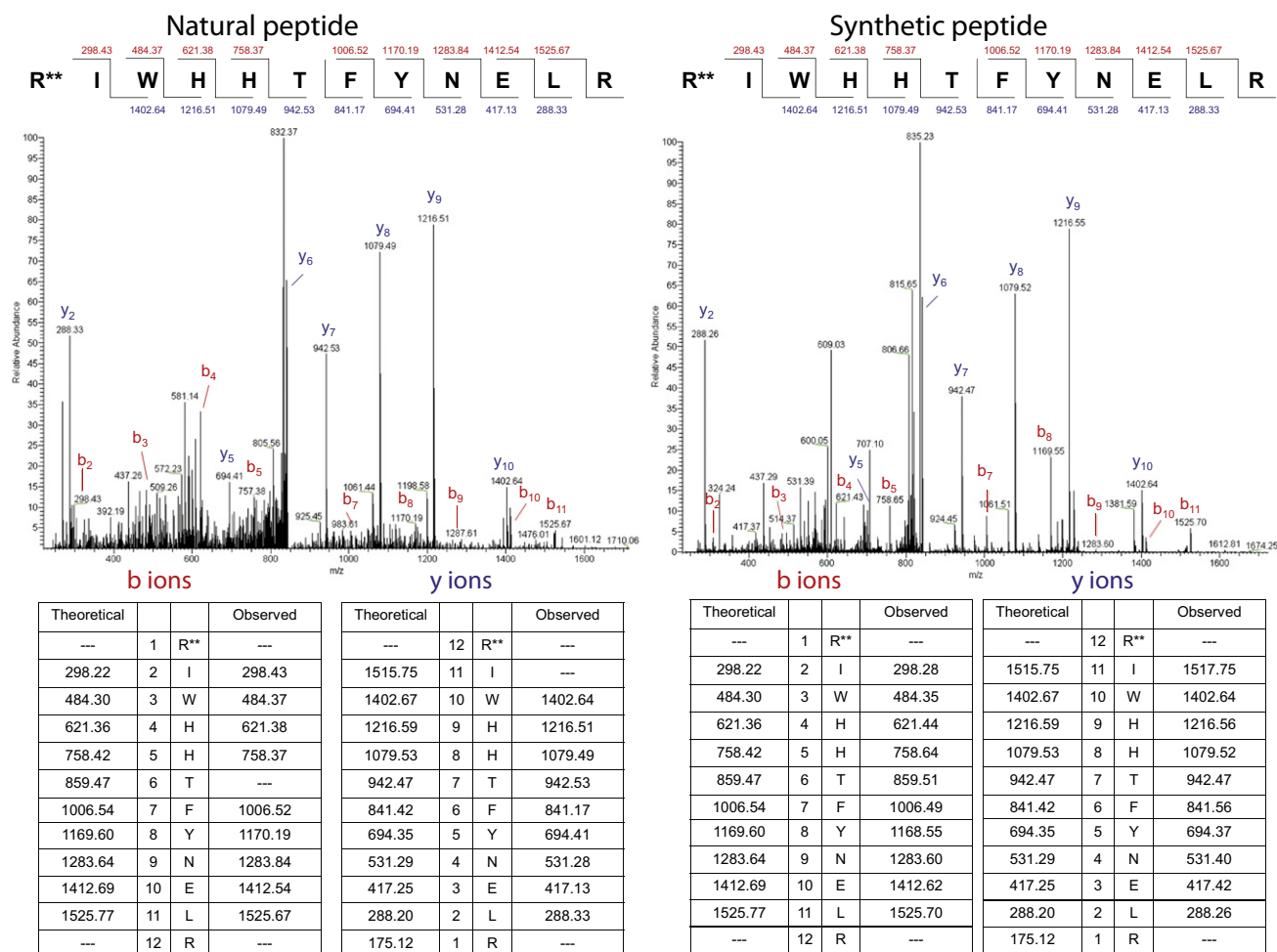
Figure 1. MS/MS Spectrum of the Experimentally Identified Arginylated Methylated Peptide Matches a Synthetic Peptide Standard of the Same Structure

MS/MS spectra of the experimentally identified dimethyl-arginylated actin peptide (NP-033736.1) (left, Natural peptide) and synthetic standard peptide containing N-terminal dimethylated Arg (right, dimethyl groups in both are denoted with double stars). Peptide sequences and b and y ion masses are shown on the top right for each spectrum and in the tables underneath.

See Table S1 for the search parameters. See also Figure S1.

and one is common between heat shock proteins 90 and 84b (identified by six different scans in a protein sample isolated by immunoaffinity chromatography with N-terminal Arg-specific antibody; Wong et al. [2007]; from mouse heart tissue) (Figure 1; see Figure S1 and Table S1 available online).

To confirm that the identified modification indeed represents Arg methylation and does not arise from an ambiguity that mimics a mass addition of +184.1325, we synthesized both peptides with a dimethyl-Arg residue in the N-terminal position, and compared the tandem mass spectra of these peptides with the tandem mass spectra of the arginylated/methylated peptides identified in the real sample. We found that the spectra between the standard peptides and the in vivo samples matched, both in the precursor masses and in the fragmentation patterns (see Figure 1 and Figure S1 for mass spectra and Table S1 for the precursor and ion fragment masses), confirming that the peptides identified by our analysis were truly arginylated/methylated.

Posttranslational Arginylation and Methylation Regulate Nuclear and Chromatin-Bound Proteins

Because Arg methylation is known to affect a large number of chromatin-associated proteins, it appears likely that if posttranslational arginylation and methylation indeed have a functional relationship, arginylation should also prominently affect chromatin-associated proteins, and such proteins could then be methylated on the added Arg to a higher extent than proteins in the cytosol. To test this hypothesis, we performed metabolic-labeling assays to compare the overall patterns of posttranslational Arg and methyl incorporation in different subcellular fractions (Figure 2A), by incubating cycloheximide and chloramphenicol-treated cells (to inhibit cotranslational incorporation of amino acids) with radioactively labeled ^{3}H-Arg or ^{3}H-Met (the source of methyl groups). As a negative control in these experiments, we used Arginyl-tRNA-protein transferase 1 knockout (*Ate1* KO) fibroblasts, in which arginylation does not occur, and
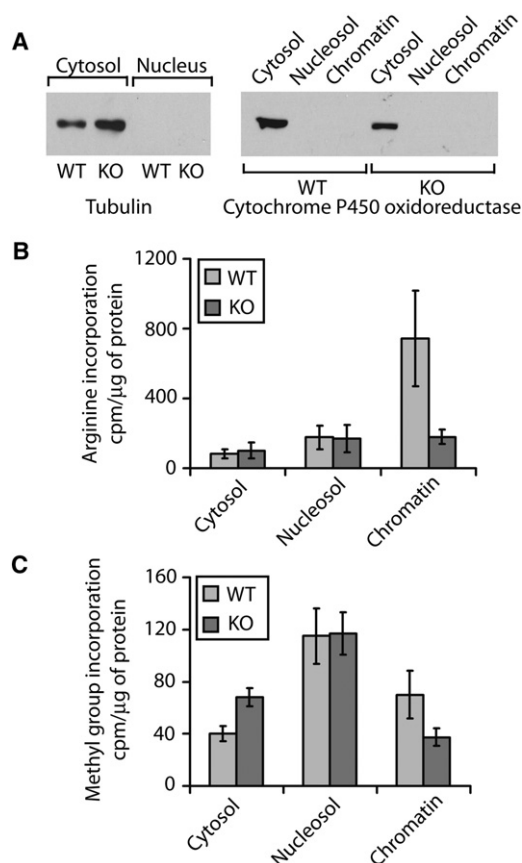
Figure 2. Arg and Methyl Group Incorporation into Different Subcellular Fractions

(A) Western blot of the cytosolic and nuclear fractions with the cytoplasmic marker tubulin (left) and the ER marker cytochrome P450 oxidoreductase confirms that purified nuclear fractions are devoid of cytosolic and ER contaminants. All the fractions for western blotting were normalized by total protein amount.

(B) In vivo incorporation of ^{3}H-Arg into cycloheximide and chloramphenicol-treated WT and *Ate1* KO cells presented as counts per minute (cpm)/microgram (μg) of protein. Error bars represent SEM (n = 4). Chromatin fraction p value is 0.0874.

(C) In vivo incorporation of ^{3}H-Met (the source of the methyl groups) into cycloheximide/chloramphenicol-treated WT and *Ate1* KO cells presented as cpm/μg of protein. Error bars represent SEM (n = 3). Cytosol and chromatin fraction p values are 0.016 and 0.1, respectively.

See also Figures S2 and S3.

protein methylation happens without the contribution of posttranslationally added Arg.

During both treatments, residual incorporation of radioactively labeled amino acids occurred due to incompletely inhibited protein synthesis, producing a background level of radioactive label incorporation (as seen from the *Ate1* KO bars in Figure 2B). Surprisingly, Arg incorporation into the cytosolic and nucleosolic protein fractions in wild-type (WT) cells did not exceed this background and did not significantly differ from that in the *Ate1* KO, suggesting that the overall stable level of posttranslational arginylation under these conditions is below the detection level for this assay. However, the amount of Arg incorporated into the chromatin-associated fraction in WT cells significantly exceeded

that in the *Ate1* KO (the rightmost set of bars in Figure 2B), indicating that posttranslational arginylation level is very significant in the chromatin fraction. Control experiments using cells without cycloheximide treatment (that incorporate Arg both during translation and posttranslationally) showed that such incorporation is also comparatively higher in the nucleosolic and chromatin fractions of WT cells, compared to the cytosolic protein fraction and to the *Ate1* KO cells (Figure S3A).

Comparison of methylation levels in different subcellular fractions from WT and arginylation KO cells revealed the same trend (Figure 2C). Although the nucleosol fraction from WT and *Ate1* KO cells had a similar amount of methyl group incorporation after 4 hr of metabolic labeling, the chromatin fraction in *Ate1* KO cells had significantly less methylation than WT. At the same time methylation levels in the cytosol in *Ate1* KO cells were higher than in WT cells, suggesting that absence of arginylation may lead to specific changes in protein Arg methylation in the cells. To test this and to evaluate the total methyl-Arg content in the protein-incorporated and free amino acid state, we performed amino acid analysis to detect methylated Arg derivatives in hydrolyzed total protein samples and in protein-free low molecular weight pools obtained from total cell lysates containing both the nuclear and the cytosolic fraction. We found that total methylation, as well as the amount of methylated Arg both in the total protein and in the free amino acid pool, was higher in *Ate1* KO than in WT (Figure S2), suggesting that in these cells the overall Arg methylation was increased via arginylation-independent mechanisms. Given these results, the increased arginylation and methylation in the chromatin fraction of WT appear even more striking and strongly suggest that arginylation-dependent methylation contributes significantly to the overall methyl incorporation specifically in the chromatin. Control experiments suggested that this methyl-arginylation was likely not mediated by arginyltransferase itself because purified ATE1 mixed in vitro with methylated Arg, tRNA, and Arg-tRNA-synthetase was unable to transfer methylated Arg onto the test substrate. Control experiments also suggested that inhibition of Arg methylation using the commercially available protein arginine methyltransferase inhibitor AMI-1 resulted in a reduction of methyl incorporation into the chromatin proteins in WT cells, but not in the *Ate1* KO cells, suggesting that only the Arg-methyltransferases specifically affected by this inhibitor are likely responsible for arginylation-dependent Arg methylation, whereas the overall increase of Arg methylation in *Ate1* KO is likely mediated by other types of Arg methyltransferases. In agreement with this, AMI-1 treatment had negligible and somewhat opposite effects on the cytosolic and nucleosolic Arg methylation in WT and *Ate1* KO cells (Figure S3B). Overall, these data suggest that posttranslational arginylation/methylation requires both the activity of arginyltransferase and the activity of specific type(s) of Arg-methyltransferase(s).

To test whether the higher incorporation of Arg into the nuclear proteins is related to the higher levels and/or activity of nuclear ATE1, we first detected the levels of ATE1 in the cytosol and the nucleus by immunoblotting and found that whereas ATE1 is present in both subcellular fractions, its level in the nucleus is significantly lower than that in the cytosol (inset in Figure 3A). We then used the cytosolic and the nuclear fractions in an in vitro arginylation assay and found that in the nuclear fraction,
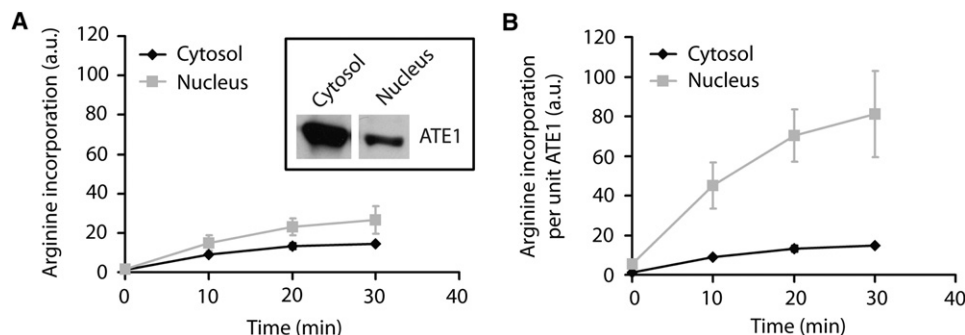
Figure 3. ATE1 Is Less Abundant but More Active in the Nucleus
(A) Inset is a western blot showing ATE1 levels in the cytosol and nucleus in the samples loaded at the equal protein level. The chart shows ^{3}H-Arg incorporation by endogenous ATE1 into cellular proteins in cytosolic and nuclear extracts at different time points. Error bars represent SEM (n = 2).
(B) The same curves as in (A) normalized to the ATE1 levels detected by the western blots shown in (A), inset. ATE1 activity per molecule is significantly higher in the nucleus than in the cytosol.
See also Figure S4.

arginylation activity was higher than in the cytosol (Figure 3A). When normalized to the comparative ATE1 levels in the two fractions, this activity was several times higher in the nucleus compared to the cytosol (Figure 3B). Control experiments showed that RNase A treatment abolished this Arg incorporation activity both in the nucleus and in the cytosol, confirming that the activity was RNA dependent (Figure S4). Therefore, increased arginylation of nuclear proteins is associated with a significantly higher arginyltransferase activity in the nucleus compared to the cytosol.

Identification of Arginylated and Arginylated/Methylated Proteins in the Nucleus

To identify proteins that are arginylated and arginylated/methylated in the nuclear subfractions, we analyzed total proteins found in the fractionated nuclear samples from mouse embryonic fibroblasts (MEFs) by mass spectrometry. This analysis revealed a large number of proteins that were arginylated in the nuclear subfractions. A significant subset of these proteins was found to contain posttranslationally added methyl-Arg, suggesting that these proteins have undergone methylation upon arginylation (see Tables 1 and 2 for the lists of proteins, and Datasets S1 and S2 for the mass spectra of arginylated/methylated and arginylated peptides, respectively). Some of these proteins contained dimethylated Arg, similarly to those previously identified in the cytosol. Others contained monomethylated Arg, suggesting that methylation of posttranslationally added Arg occurs in steps and is likely mediated by the same enzymes as regular Arg methylation. Remarkably, these proteins were identified without any special enrichment procedures in the total nuclear pool, suggesting that the extent of arginylation and arginylation/methylation on these proteins is relatively high.

Overall, the arginylated and arginylated/methylated proteins identified by this analysis fell into distinct functional groups. A prominent subset of these proteins included histones, which were arginylated and in some cases further methylated on the sites involved in DNA binding, ion interaction, and association with other histones in the nucleosome complex (Figure 4A). On these proteins, arginylated sites were identified in multiple scans

in different analyzed subsets, suggesting that the overall abundance of the arginylated form of these proteins was relatively high. Another subset of arginylated proteins has roles in gene expression and RNA splicing, including RNA-binding proteins and related factors. Although, similar to the previous studies (Wong et al., 2007), no arginylation consensus was found at or around these sites, a Lys residue was prevalent in the (−1) position to the arginylated site. This finding suggests that in the nucleus, unlike in the cytosol, recognition of proteins by the arginylation machinery may be mediated by Lys-containing primary sequence motifs. Structural analysis using the 3D structures available in the PDB database for these proteins or their close homologs revealed that in all cases arginylation sites were located on the protein surface (Figure 4; Figure S5) and in many cases fell onto functionally significant regulatory regions of these proteins. Structural modeling of the arginylated histone proteins and the nucleosome indicated that some of the added Arg and methyl-Arg have a distance of 5–7 Å from the nearest phosphate in the DNA backbone that, given the opposite charges in these groups, can potentially facilitate the interaction of the histones, with DNA (Figure 4A). These results confirm that arginylation and arginylation/methylation occur to a significant extent in the nucleus compared to the cytosol, and that these modifications in the nucleus likely play an important regulatory role.

Arginylation/Methylation Regulates Nuclear Size in Cultured Cells

To assess the biological role of arginylation/methylation of chromatin proteins, we analyzed the identity of these proteins (Tables 1 and 2) and noticed that a large fraction of these proteins had prominent roles in DNA compaction and architecture. To test the hypothesis that arginylation/methylation regulates nuclear architecture, we compared the nucleus size in cultured WT and *Ate1* KO MEFs and found that *Ate1* KO nuclei were ∼20% smaller compared to WT nuclei (Figures 5A and 5B). Strikingly, treatment of WT and *Ate1* KO cells with AMI-1 reduced the nuclear size of the WT MEFs to the same extent (by approximately 20%), without affecting the nuclear size in *Ate1* KO cells

Table 1. Posttranslationally Arginylated/Methylated Proteins

Accession Number	Name	Modified Residue	Modification	Cellular Fraction	Function
AAA37866	HSP90-α/HSP84b	S587	DMA	Cytosol	Heat shock
NP_033736.1	Actin, α skeletal muscle	I87	DMA	Cytosol	Actin cytoskeleton
NP_058557.2	PDZ and LIM domain protein 1	I30	MMA	Nucleosol	Actin cytoskeleton
NP_062752.2	Zinc finger and BTB domain containing 20 isoform L	I69	MMA	Nucleosol	Transcriptional repressor
NP_001104761.1	Cytoplasmic activation/ proliferation- associated protein 1 isoform c	L170	DMA	Nucleosol	Cell proliferation, translation regulation, RNA binding
NP_598432	Splicing factor U2AF 65 kDa subunit	G248	MMA	Nucleosol	Pre-mRNA splicing, 3′ end mRNA processing
NP_035561.1	116 kDa U5 sn RNP component isoform a	E689	DMA	Nucleosol	Pre-mRNA splicing
NP_598841.1	Filamin B	S1911	MMA	Chromatin	Actin cytoskeleton
CAM16973.1	Spectrin β-2	E741	MMA	Chromatin	Actin cytoskeleton
NP_835509.2	Histone H2B F	K47	MMA	Chromatin	Chromatin structure and organization
NP_033736.1	Actin, α skeletal muscle	N299	MMA	Chromatin	Actin cytoskeleton
NP_032017.2	Fibrillarin	Q160	MMA	Chromatin	snoRNP particle component, pre-rRNA processing, pre-rRNA methylation, ribosome assembly
NP_035897.2	Zinc finger RNA-binding protein	A444	DMA	Chromatin	RNA transport and localization

See Table S2 for the search parameters and peptide sequences, and see Figure 1, Figure S1, and Dataset S1 for the corresponding mass spectra.

(Figures 5A and 5B). Thus, inhibition of both arginylation and Arg methylation leads to exactly the same effect on the nuclear size, clearly indicating that arginylation combined with methylation regulates nuclear size in vivo. Because nuclear size can change due to alteration in chromatin compaction or other structural change in nucleus, the exact mechanism of arginylation/ methylation-mediated regulation of nuclear size remains to be investigated.

DISCUSSION

This work is the first demonstration that two posttranslational modifications can affect the same sites in a protein, and the first comprehensive analysis of arginylation in the nucleus that reveals the identity of a significant number of previously unknown protein arginylation targets and outlines a likely function for protein arginylation in chromatin structure and gene expression. Our result that ATE1 is less abundant and more active in the nucleus suggests that this enzyme undergoes an additional level of regulation that could increase or reduce its activity dependent on its intracellular localization. Such regulation could be achieved by modulating the interaction of ATE1 with different binding partners in the cytoplasm and the nucleus that may form either activatory or inhibitory complexes with this enzyme and/or regulate its activity by changing its conformation or posttranslationally modified state. Different chemical environments in these two intracellular compartments may also play a role, by creating favorable or unfavorable ionic conditions for arginylation.

It has been previously shown that some ATE1 isoforms showed preferential nuclear localization under certain condi-

tions, suggesting that this localization depends on the cell's physiological state (Kwon et al., 1999; Rai and Kashina, 2005; Wang et al., 2011). It is possible that higher arginylation activity in the nucleus could be necessary because of the existence of a higher number and/or amount of natural substrates of ATE1 in the nucleus. It was previously found that partially purified ATE1 preparation can arginylate nuclear proteins in vitro (Kaji, 1976). Our data demonstrate for the first time that ATE1 preferentially arginylates nuclear proteins in vivo and that the activity of this enzyme can be differentially regulated in different intracellular compartments.

Many of the proteins identified in the cytosolic and chromatin fractions have a demonstrated role in the corresponding compartments, such as HSP90 in the cytosol, and DNA- and RNA-binding proteins involved in chromatin structure, gene expression, and RNA processing in vivo. For such proteins, regulation by arginylation with or without subsequent methylation can conceivably facilitate their structural interactions and affect their in vivo functions. For nucleic acid-binding proteins, these modifications likely affect their ability to interact with the nucleic acids (which should be facilitated by the addition of the positively charged Arg), or other proteins, which could be repelled or attracted by the chemical groups on the matching surfaces of the interaction partners. However, a number of proteins found in the nucleosolic and chromatin fractions have been previously characterized as cytoplasmic proteins. For some of those, including actin, GAPDH, filamin, and spectrin, recent data demonstrate their involvement in gene expression and transcriptional regulation (Grummt, 2006; Loy et al., 2003; Sawa et al., 1997; Tang et al., 2003). It is likely that arginylation and arginylation/methylation of these proteins in the nucleus

Table 2. Proteins Arginylated in the Nucleus

Accession Number	Name	Modified Residue	Cellular Fraction	Function
NP_034928.1	Macrophage migration inhibitory factor	A35	Nucleosol	Proinflammatory cytokine
CAM24508.1	Far upstream element binding protein 3	I111	Nucleosol	ssDNA and RNA binding, transcriptional and posttranscriptional regulation
P11499.2	HSP 90β	L439	Nucleosol	Heat shock
NP_034368.2	Bromodomain-containing protein 2 isoform a	L10	Nucleosol	Transcriptional regulation, chromatin remodeling, signal transduction, nuclear kinase
EDL34349.1	Similar to glyceraldehyde-3-phoSphate dehydrogenase	V121	Nucleosol/ chromatin	Glycolysis, transcription activation, initiation of apoptosis, and ER to Golgi vesicle transport.
NP_035931.1	Craniofacial development protein 1	K104	Nucleosol	Transcriptional developmental signaling
CAI24328.1	RAB34	D75	Nucleosol	Lysosomal positioning, protein transport and secretion
NP_034063.1	Collagen α-1(VI) chain precursor	E576	Nucleosol	Extracellular matrix
NP_084206.1	Hypothetical protein LOC77574	G633	Nucleosol	Unknown
NP_663600.2	Eukaryotic translation initiation factor 4B	G138	Nucleosol	Translation
BAC36440.1	Histone H2B homolog	A59	Chromatin	Chromatin structure and organization
NP_783597.2	Histone H2B type 2-B	Q48; A59	Chromatin	Chromatin structure and organization
XP_981474.1	Histone H4	T136; V116	Chromatin	Chromatin structure and organization
AAA37763.1	Histone H2A.1	S41	Chromatin	Chromatin structure and organization
EDK99240.1	Ribophorin I, isoform CRA_a	V468	Chromatin	N-glycosylation, ER chaperone
NP_031419.1	Actin, β	D51	Chromatin	Actin cytoskeleton, transcriptional regulation
XP_001476183.1	Similar to ribosomal protein L23a	L140	Chromatin	Ribosomal organization
NP_035831.2	Vimentin	V224	Chromatin	Intermediate filament, cytoskeleton
NP_080447.1	Nuclear valosin-containing protein-like	T628	Chromatin	Nuclear ATPase

See Table S2 for the search parameters and peptide sequences and Dataset S2 for the corresponding mass spectra.

can facilitate this specific function. For others, such as collagen, no such function has been previously demonstrated. It is possible that their interaction with the chromatin is a nonspecific effect, resulting from the presence of the surface Arg residues that may induce their interaction with DNA after cell and nuclear lysis. However, it is also possible that their presence in the nuclear fractions found in this study reflects their previously unknown function in the chromatin and/or transcriptional control.

Structural and sequence analysis shows that arginylated sites for most of the identified proteins are located in the middle of their polypeptide chains, in conserved domains predicted to affect key molecular interactions (Figure 4; Figure S5). It has been previously found that for many arginylated proteins, sites for posttranslational addition of Arg are located in the middle of the polypeptides, far away from the N terminus, suggesting a novel regulatory mechanism that couples arginylation with partial regulatory proteolysis without disrupting the protein's quaternary structure (Wong et al., 2007). In agreement with this, all the available structural data presented in our current study (Figures 4; Figure S4) show that arginylated sites are exposed within the structurally important conserved domains on the protein surface, suggesting that arginylation on these sites affects fully synthesized and folded proteins and plays an important regulatory role. The measured distances of the added Arg or methyl-Arg residues in histones to the phosphate groups in the DNA backbone are particularly encouraging, because

these distances with the added Arg are short enough to significantly affect the interaction of histones with DNA, and thus, can affect DNA packaging and/or transcriptional regulation. Conceivably, addition of methyl groups onto Arg in these positions should stabilize the arginylation sites and ensure their longer-term action in the nucleosome. Further studies of the individual arginylation sites identified in this study will reveal the functional significance of their regulation by arginylation and advance the knowledge about this poorly understood posttranslational modification to the next level of understanding.

Our finding that posttranslationally added Arg can be methylated constitutes a proof of principle and the first demonstration that posttranslational modifications can double on the same sites within the proteins. It has been previously shown that a negative posttranslational regulation of methylation by phosphorylation can occur by modifying the neighboring residues to change the charge and conformation of the methylation sites (Bedford and Richard, 2005). We show for the first time a positive regulation by a posttranslational modification that can facilitate methylation by creating an acceptor site for methyl groups. Although many of the functions of such double modifications remain to be determined, our study shows a clear correlation between arginylation/methylation and the nuclear size. Because a smaller nucleus implies a higher compaction, which can affect nuclear functions significantly, it is reasonable to suggest that arginylation/methylation will have a significant effect on nuclear function. It also appears likely that arginylation
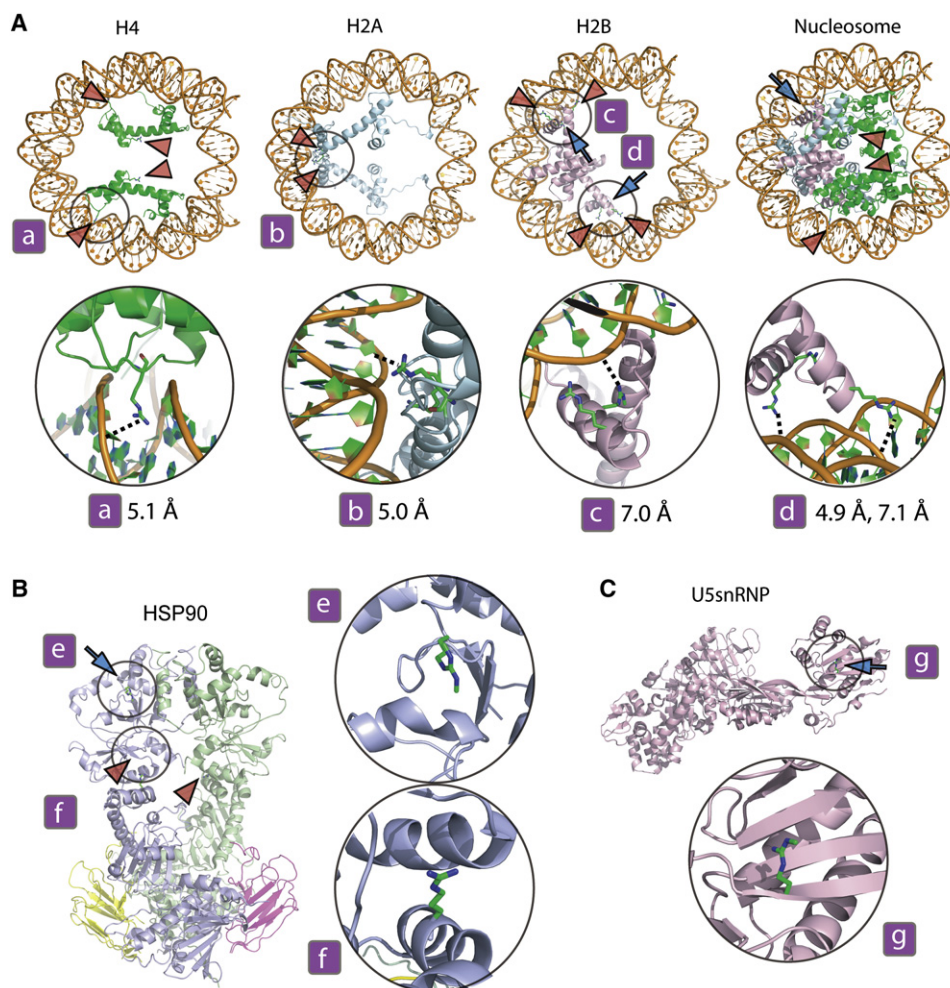
Figure 4. Structures of Some of the Arginylated and Arginylated/Methylated Proteins Identified in This Study

The position of the posttranslationally added Arg or methyl-Arg residues in the represented structures is modeled by incorporating Arg or methyl-Arg in the primary structure of the protein. Posttranslationally added arginines are marked with green, and the side-chain amino groups are in blue. The methyl groups in arginine are marked by green spots attached to the blue amino group. Arginylation sites are shown with red arrowheads and arginylation/methylation sites with blue arrows.

(A) Nucleosome (structures shown are selected from the complex between nucleosome core particle [H3, H4, H2a, H2b] and 146 bp long DNA fragment; PDB identifier: 1AOI). The position of individual histones (marked on top) and the entire nucleosome assembly (marked "Nucleosome" in the top rightmost panel) are shown in the topmost panels. The modified sites are located on the DNA binding and histone interaction surfaces likely to affect nucleosome assembly. The distances between the nitrogen atoms of added Arg or methyl-Arg and the nearest oxygen atoms of phosphate backbone of the DNA molecule are represented in the inset images a, b, c, and d and marked by black-dotted lines.

(B) Hsp90 dimer (structure shown is of *S. cerevisiae* Hsp90; PDB identifier: 2CG9). Modified sites are located on the protein surface and near subunit interaction sites. Inset images e and f show a detailed image of the modified regions.

(C) U5snRNP-specific protein (structure shown is of a homologous *S. cerevisiae* protein in complex with rRNA obtained by selecting the relevant structures from the *S. cerevisiae* ribosomal 80 s-Eef2-sordarin complex; PDB identifier: 1S1H). Inset image g shows a detailed image of the modified regions.

See also Figure S4 for additional structures. See also Figure S5, Datasets 1 and 2, and Table S2.

with subsequent methylation of the added Arg residues can serve as a second-order regulatory mechanism that could affect those major metabolic processes in which arginylation and methylation have been implicated. Methylation may also act to protect posttranslationally added Arg from chemical modifications in vivo that could destroy or inactivate them (for example, by methylglyoxal—a cytotoxic agent linked to disease), in line with a recently proposed model that methylation acts as a "guardian" of Arg (Fackelmayer, 2005). Finally, methylation

may be necessary to protect proteins against degradation by the N-end rule pathway, shown to affect some proteins with Arg in the N-terminal position (Varshavsky, 1997). It appears likely that the posttranslationally added Arg residues within a protein that further serve as targets of methylation should be especially important and/or long lived to allow their further modification by Arg methyltransferases, an additional level of regulation or protection of these proteins, and their important functional sites in vivo.
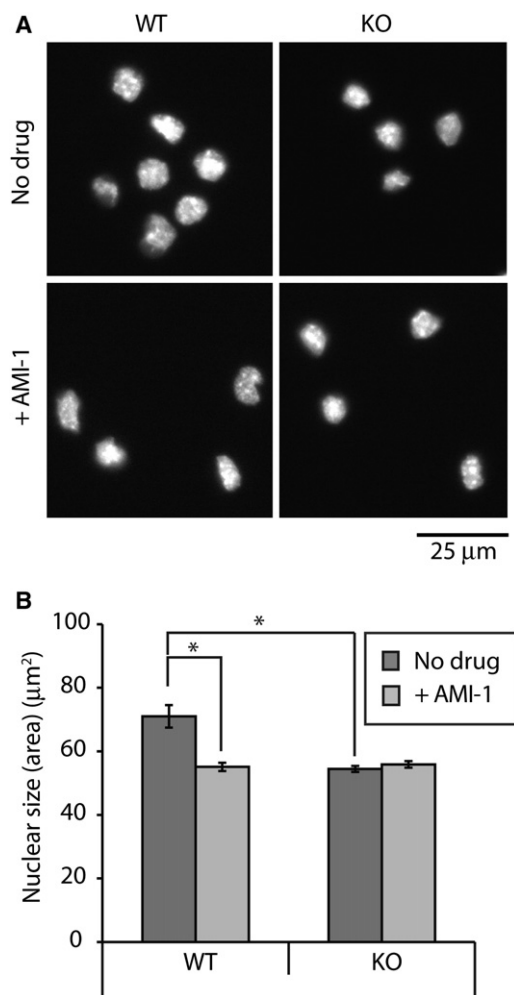
A

WT   KO

No drug

+ AMI-1

25 μm

B

Figure 5. Arginylation Methylation Regulates Nuclear Size
(A) DAPI stained untreated or AMI-1-treated WT and *Ate1* KO MEF nucleus. Scale bar, 25 μm.
(B) Quantification of nuclear size observed in the experiments presented in (A); the following numbers of nuclei were analyzed: WT, 41; *Ate1* KO, 96; WT + AMI-1, 84; and *Ate1* KO + AMI-1, 150. Error bars represent SEM. Star (*) indicates the p value <0.0001, as determined by Student's t test.

SIGNIFICANCE

The current study demonstrates that posttranslationally added arginines on proteins can be further modified by arginine methylation and that this double posttranslational modification can regulate DNA arrangement in the chromatin and nuclear architecture. This is the first evidence that posttranslationally modified sites of proteins can undergo subsequent posttranslational modifications, and the first functional link between protein arginylation and Arg methylation, which has been long known to regulate chromatin and nuclear proteins in vivo. It appears likely that such double modifications not only increase the complexity of the signaling through the arginylation and the methylation pathways but also exert other regulatory functions through a limited number of protein sites, and

even provide a unique type of protein code. Our study demonstrated that the interaction of the two important protein modifications, arginylation and methylation, affects the nuclear size, suggesting its possible role in nuclear architecture and function.

EXPERIMENTAL PROCEDURES

Nuclear Fractionation and Nuclear Protein Isolation for Mass Spectrometry

Cells were lysed in the lysis buffer (0.3 M sucrose, 4 mM Mg acetate, 12.5 mM KCl, 50 mM Tris [pH 8.0], 0.5 mM arginine, 1 mM DTT, 0.2 mM PMSF, protease inhibitor cocktail), and nuclear and cytosolic fractions were separated by centrifugation and further purification of nuclei by centrifugation through sucrose cushion. Pure nuclei were lysed in 400 mM NaCl-containing lysis buffer, and the soluble part was collected as nucleosol. The insoluble chromatin part was further extracted with 2 M NaCl-containing buffer, and finally by DNase I treatment (see Supplemental Experimental Procedures for the detailed procedure). Proteins from each fraction were precipitated by 20% TCA, and the pellets were analyzed by mass spectrometry.

Mass Spectrometry and Database Searches

Extracts from subcellular fractions of adult mouse heart (cytoplasmic fractions) and immortalized MEFs (cytoplasmic and nuclear fractions) were used throughout this study. For the initial identification of arginylation/methylation, samples obtained during the previous study on the global analysis of protein arginylation (Wong et al., 2007) were subjected to a new search to detect the N-terminal mass addition of Arg plus one or two methyl groups as described below and in Xu et al. (2009). In addition, one protein spot excised from 2D gel of mouse cardiac actin (Rai et al., 2008) was analyzed. For the analysis of nuclear proteins, samples were obtained as described above, and the TCA-precipitated protein pellets were dissolved in Invitrosol™ LC/MS Protein Solubilizer following the manufacturer's protocol (Invitrogen), reduced and alkylated with 10 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP; Roche Applied Science, Palo Alto, CA, USA) and 55 mM iodoacetamide (IAM; Sigma-Aldrich) in 100 mM ammonium bicarbonate. Digestion was performed in the presence of 50 mM ammonium bicarbonate and 5 mM calcium chloride (Sigma-Aldrich) using sequencing grade-soluble trypsin (Promega, Madison, WI, USA). The resulting peptides were extracted by 5% formic acid, redissolved into buffer A (5% acetonitrile with 0.1% formic acid), and were pressure loaded onto a reverse-phase (RP) or strong cation exchange (SCX)/RP MudPIT column and eluted with a linear gradient of acetonitrile or pulses of ammonium acetate salt (5%–100%). Synthetic peptides were dissolved in water and supplemented with 1% formic acid to the final concentration of 100 fmol μl^{-1} and analyzed under the same LC-MS/MS conditions as those used for the digested protein samples. Data-dependent tandem mass spectrometry (MS/MS) analysis was performed in a LTQ-Orbitrap mass spectrometer (Thermo Electron, San Jose, CA, USA). Full mass spectrometry spectra were acquired by the precursor ion scan using the Orbitrap analyzer with resolution set at 60,000, followed by nine MS/MS events in the linear ion trap (LTQ), sequentially generated on the first-, second-, and third-most intense ions selected from the full mass spectrometry spectrum. Mass spectrometry scan functions and HPLC solvent gradients were controlled by the Xcalibur data system (Thermo Electron). Tandem mass spectra were searched against EBI-IPI and NCBI mouse protein databases concatenated with reversed sequences to estimate false-positive rate, using the ProLuCID (Xu et al., 2006) protein database search algorithm with peptide mass tolerance of ±3 amu, fragment ion mass tolerance of 0.4 amu, and a static modification of 57.0215 on Cys due to carboxyamidomethylation. In order to identify N-terminal arginylated peptides, differential N-terminal modification searches for a mass addition of Arg (+156.1011) with an extra mass of either one (+14.0157) or two (+28.0314) methyl groups were performed with no enzymatic cleavage conditions imposed on the database search. ProLuCID search results were then filtered with DTASelect 2.0 (Tabb et al., 2002) using XCorr and DeltaCN at a false-positive rate of 0.1% (−fp 0.001) estimated by number of reverse hits. Half or fully tryptic (−y 1) peptides only with N-terminal arginylation and Arg methylation (−m 0) and

a DeltaMass ≤ 5 ppm ($-$DM 5) were accepted; the minimum number of peptides to identify a protein was set to 1 ($-$p 1). All spectra were verified by automated filtering and manual filtering against the mass ambiguities shown in Tables S3 and S4. The final proteins with arginylation modifications that passed these filtering criteria and manual validation are listed in Tables 1 and 2, and the corresponding mass spectra are shown in Figure 1 and Figure S1 and Datasets S1 and S2.

In Vivo Methylation Assay
A total of 1.2 million immortalized MEFs were seeded in 100 mm tissue culture dishes and grown for 15 hr. The plates were washed once with 5 ml of Met- and Cys-free DMEM without serum and incubated with 5 ml of the same media for 1 hr 30 min to induce amino acid starvation. After the initial starvation period, cycloheximide (100 μg/ml) and chloramphenicol (40 μg/ml) (as well as arginine methyltransferase inhibitor AMI-1 [100 μM] in case of AMI-1 treatment) were added to the medium, and the incubation was continued for an additional hour, followed by addition of 30 μCi of L-[methyl-^{3}H]-Methionine (GE Amersham) for a 4 hr incorporation. At the end of the incorporation period, cells were washed in PBS and harvested by scraping and centrifugation. Cell pellets were lysed by pipetting in buffer A (50 mM Tris [pH 8], 0.3 M Sucrose, 4 mM Mg-acetate, 12.5 mM KCl, 10 mM β-ME, 0.5% NP 40, 1 mM DTT, 0.2 mM PMSF, protease inhibitor cocktail). Nucleus and cytosol were separated using the protocol adapted from Kaji (1976) by spinning at 4°C for 5 min at 2000 × g. The supernatant (cytosol) was collected, and the nuclear pellet was resuspended in buffer B (50 mM Tris [pH 7.5], 25 mM KCl, 5 mM MgCl$_2$ 0.5% NP 40, 1 mM DTT, 0.2 mM PMSF, protease inhibitor cocktail) and layered onto an equal volume of 0.88 M sucrose in buffer B, followed by centrifugation at 4°C for 5 min at 2000 × g. The pellet containing pure nuclei was collected and lysed by pipetting in hypertonic buffer C (50 mM Tris [pH 8.0], 25 mM KCl, 400 mM NaCl, 1 mM DTT, 0.2 mM PMSF, protease inhibitor cocktail), followed by centrifugation at 4°C for 10 min at 16,000 × g. The supernatant was collected as nucleosol and the pellet as chromatin. Protein concentrations in the cytosol and nucleosol were estimated by the Bradford assay. Protein concentration in the chromatin fractions was estimated by running known amounts of nucleosolic proteins and 2× Laemmli buffer-solubilized chromatin pellets from similar preparations in SDS-PAGE, followed by densitometric analysis of the whole lane of Coomassie-stained gel. Equal amounts of protein from cytosol and nucleosol were precipitated by TCA, and incorporation into the pelleted protein was monitored by liquid scintillation. The chromatin pellet was washed three times with buffer B, and incorporation was monitored by liquid scintillation.

In Vivo Arginylation Assay
This assay was performed similar to the in vivo methylation assay, except for the following changes. Instead of Met- and Cys-free DMEM, custom-made Arg-free DMEM (Hyclone) was used. A total of 60 μCi of L-[2,3,4-^{3}H] Arginine (PerkinElmer) was added for 4 hr to each dish containing 5 ml of Arg-free DMEM without serum. For in vivo arginylation in the absence of translation inhibitors, the assay was performed in the same way except the cycloheximide/chloramphenicol treatment step was omitted.

Estimation of Intracellular Methyl-Arg Concentrations in the Protein and Free Metabolite Pools
Protein fractions prepared as described in the Supplemental Experimental Procedures were precipitated by 10% TCA, and the supernatant was analyzed by HPLC without acid hydrolysis. The pure methyl arginines (MMA, SDMA, and ADMA) were used as the standards in HPLC analysis (see Supplemental Experimental Procedures for standard concentrations). HPLC, amino acid hydrolysis, and Arg determination were performed by the W.M. Keck Foundation Biotechnology Resource Laboratory at Yale University. See Supplemental Experimental Procedures for the detailed procedure.

In Vitro Arginylation Assay to Estimate Endogenous ATE1 Activity in Cellular Fractions
This assay was performed using purified charged Arg-tRNA and protein fractions as the source of Arg-transferase as described in Wang et al. (2011) with some modifications. Cytosol and nuclei were purified in the same way as described in the nucleus fractionation section except the pure nuclei were

lysed by sonication followed by clarification by centrifugation at 16,000 × g and used for the assay. After the reaction, protein was precipitated with 10% cold TCA and counted in a liquid scintillation counter (see Supplemental Experimental Procedures for the detailed procedure).

Estimation of ATE1 Level in Cytosol and Nucleus Fractions
Cells were fractionated in the same way as described in the section "In Vivo Methylation Assay." Proteins in the cytosol and nucleosol fractions were estimated by Bradford assay. Equal amounts of proteins were analyzed by western blot using an ATE1 rat monoclonal antibody (Wang et al., 2011) to estimate the difference between the levels of ATE1 in these two fractions.

Estimation of the Nuclear Size
WT or Ate1 KO MEF cells, untreated or treated with 100 μM AMI-1 for 24 hr, were harvested by trypsinization. Cells were washed once with PBS and fixed for 20 min with 4% paraformaldehyde in PBS at room temperature. After fixation, cells were washed three times with PBS, stained with 1 μg/ml DAPI for 10 min, spotted on polylysine-coated coverslips, and mounted with Aqua-Poly/Mount (Polysciences, Inc). Five randomly chosen fields from each preparation were captured at 40× and analyzed using MetaMorph software. The pixel area covered by each nucleus was scored as nuclear size and converted to squared micrometers (μm^2) using the MetaMorph internal calibration tool.

Arginylated and methyl-arginylated protein structure models and the measurement of distances between arginyl and methylarginyl residues and the neighboring phosphate atoms of DNA in the histone complexes were performed using the program Coot (Emsley and Cowtan, 2004). Illustrations of structures were prepared with the program PyMOL (Schrödinger).

SUPPLEMENTAL INFORMATION

Supplemental Information includes four tables, five figures, two datasets, and Supplemental Experimental Procedures and can be found with this article online at doi:10.1016/j.chembiol.2011.08.019.

REFERENCES

Bedford, M.T., and Richard, S. (2005). Arginine methylation an emerging regulator of protein function. Mol. Cell 18, 263–272.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132.

Fackelmayer, F.O. (2005). Protein arginine methyltransferases: guardians of the Arg? Trends Biochem. Sci. 30, 666–671.

Grummt, I. (2006). Actin and myosin as transcription factors. Curr. Opin. Genet. Dev. 16, 191–196.

Kaji, H. (1976). Amino-terminal arginylation of chromosomal proteins by arginyl-tRNA. Biochemistry 15, 5121–5125.

Karakozova, M., Kozak, M., Wong, C.C., Bailey, A.O., Yates, J.R., 3rd, Mogilner, A., Zebroski, H., and Kashina, A. (2006). Arginylation of beta-actin regulates actin cytoskeleton and cell motility. Science 313, 192–196.

Kwon, Y.T., Kashina, A.S., and Varshavsky, A. (1999). Alternative splicing results in differential expression, activity, and localization of the two forms of

arginyl-tRNA-protein transferase, a component of the N-end rule pathway. Mol. Cell. Biol. 19, 182–193.

Kwon, Y.T., Kashina, A.S., Davydov, I.V., Hu, R.G., An, J.Y., Seo, J.W., Du, F., and Varshavsky, A. (2002). An essential role of N-terminal arginylation in cardiovascular development. Science 297, 96–99.

Loy, C.J., Sim, K.S., and Yong, E.L. (2003). Filamin-A fragment localizes to the nucleus to regulate androgen receptor and coactivator functions. Proc. Natl. Acad. Sci. USA 100, 4562–4567.

Rai, R., and Kashina, A. (2005). Identification of mammalian arginyltransferases that modify a specific subset of protein substrates. Proc. Natl. Acad. Sci. USA 102, 10123–10128.

Rai, R., Wong, C.C., Xu, T., Leu, N.A., Dong, D.W., Guo, C., McLaughlin, K.J., Yates, J.R., 3rd, and Kashina, A. (2008). Arginyltransferase regulates alpha cardiac actin function, myofibril formation and contractility during heart development. Development 135, 3881–3889.

Sawa, A., Khan, A.A., Hester, L.D., and Snyder, S.H. (1997). Glyceraldehyde-3-phosphate dehydrogenase: nuclear translocation participates in neuronal and nonneuronal cell death. Proc. Natl. Acad. Sci. USA 94, 11669–11674.

Tabb, D.L., McDonald, W.H., and Yates, J.R., 3rd. (2002). DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. J. Proteome Res. 1, 21–26.

Tang, Y., Katuri, V., Dillner, A., Mishra, B., Deng, C.X., and Mishra, L. (2003). Disruption of transforming growth factor-beta signaling in ELF beta-spectrin-deficient mice. Science 299, 574–577.

Varshavsky, A. (1997). The N-end rule pathway of protein degradation. Genes Cells 2, 13–28.

Wang, J., Han, X., Saha, S., Xu, T., Rai, R., Zhang, F., Wolf, Y.I., Wolfson, A., Yates, J.R., 3rd, and Kashina, A. (2011). Arginyltransferase is an ATP-independent self-regulating enzyme that forms distinct functional complexes in vivo. Chem. Biol. 18, 121–130.

Wong, C.C.L., Xu, T., Rai, R., Bailey, A.O., Yates, J.R., 3rd, Wolf, Y.I., Zebroski, H., and Kashina, A. (2007). Global analysis of posttranslational protein arginylation. PLoS Biol. 5, e258.

Xu, T., Wong, C.C., Kashina, A., and Yates, J.R., 3rd. (2009). Identification of N-terminally arginylated proteins and peptides by mass spectrometry. Nat. Protoc. 4, 325–332.

Xu, T., Venable, J.D., Park, S.K., Cociorva, D., Lu, B., Liao, L., Wohlschlegel, J., Hewel, J., Yates, J.R., and Ird, I.I. (2006). ProLuCID, a fast and sensitive tandem mass spectra-based protein identification program. Mol. Cell. Proteomics 5, S174.