Delineating *Anopheles gambiae* coactivator associated arginine methyltransferase 1 automethylation using top–down high resolution tandem mass spectrometry

Peter Kuhn,[1] Qingge Xu,[2] Erika Cline,[1] Di Zhang,[1] Ying Ge,[2]* and Wei Xu[1]*

[1]McArdle Laboratory for Cancer Research, School of Medicine and Public Health, University of Wisconsin–Madison, Madison, Wisconsin 53706

[2]Human Proteomics Program and Department of Physiology, School of Medicine and Public Health, University of Wisconsin–Madison, Madison, Wisconsin 53706

Received 22 January 2009; Revised 8 April 2009; Accepted 9 April 2009
DOI: 10.1002/pro.139
Published online 16 April 2009 proteinscience.org

Abstract: Coactivator-associated arginine methyltransferase 1 (CARM1), originally defined as a coactivator for steroid receptors, is a member of the protein arginine methyltransferases. Here, we report the discovery and characterization of an automethylation event by *Ag*CARM1, a CARM1 homologue in the mosquito *Anopheles gambiae*, using top–down high resolution tandem mass spectrometry, which allows fine mapping of modifications in the intact protein accurately and quantitatively without priori knowledge. Unexpectedly, we found that *Ag*CARM1 has already been predominantly dimethylated during its expression in *Escherichia coli*. A single arginine methylation site, R485, was identified which is conserved among CARM1 in insects. No methylation was observed in the intact *Ag*CARM1^{R485K} mutant where R485 is mutated to lysine, which confirms that R485 is the only detectable methylation site. Using *Ag*CARM1 methyltransferase defective mutants, we confirmed that this is an automethylation event and show the automethylation of *Ag*CARM1 occurs intermolecularly. This study represents the first comprehensive characterization of an automethylation event by top–down mass spectrometry. The unexpected high percentage of automethylated recombinant *Ag*CARM1 expressed in *E. coli* may shed light on other bacterially expressed post-translational modifying enzymes, which could be modified but overlooked in biochemical and structural studies. Top–down high resolution tandem mass spectrometry thus provides unique opportunities for revealing unexpected protein modification, localizing specific modification to one amino acid, and delineating molecular mechanism of an enzyme.

Keywords: protein arginine methyltransferase; CARM1; automethylation; top–down mass spectrometry; Fourier transform mass spectrometry; post-translational modification; electron capture dissociation

Additional Supporting Information may be found in the online version of this article.

Abbreviations: AdoMet, S-adenosylmethionine; *Ag*CARM1, *A. gambiae* CARM1; CAD, collisionally activated dissociation; ECD, electron capture dissociation; ESI, electrospray ionization; FT, Fourier transform; LTQ, linear ion trap; MS, mass spectrometry; MS/MS, tandem mass spectrometry; PABP1, polyA binding protein 1; PRMT, protein arginine methyltransferase; PTM, post-translational modification; R, arginine.

Grant sponsors: Shaw Scientist Award from Greater Milwaukee Foundation; NIH; Grant number: R01CA125387; Grant sponsor: The Wisconsin Partnership Fund for a Healthy Future.

*Correspondence to: Ying Ge, 1300 University Ave, SMI 130, Madison, WI, 53706. E-mail: yge@physiology.wisc.edu or Wei Xu, 1400 University Ave, 421A McArdle, Madison, WI, 53706. E-mail: wxu@oncology.wisc.edu

Introduction

The methylation of arginine (R) residues is catalyzed by the protein arginine methyltransferase (PRMT) family of enzymes, which has been shown to play an important role in diverse cellular functions including transcription regulation, mRNA splicing, and DNA damage repair. The PRMTs are generally classified as either Type I or Type II enzymes: Type I PRMTs monomethylate and asymmetrically dimethylate the primary amine groups on the arginine side chain, whereas Type II PRMTs catalyze the formation of monomethyl arginine and symmetric dimethyl arginine.[1]

Coactivator-associated arginine methyltransferase 1 (CARM1), also known as PRMT4, is a Type I PRMT which was originally identified in a yeast two-hybrid screen as an associated protein for glucocorticoid receptor interacting protein 1, the p160 family steroid receptor coactivator.[2] Mutation and structural studies have identified a number of functional domains in CARM1.[3–5] CARM1's central catalytic core is characteristic of PRMT family proteins,[6] which contain an S-adenosylmethionine (AdoMet) binding domain and a substrate binding domain. Within this central core, the AdoMet binding domain contains a valine, leucine, aspartic acid (VLD, amino acids 189–191) sequence that is essential for enzymatic function.[2] This sequence is part of motif 1 (aliphatic-aliphatic-acidic-alphatic-G-X-G-X-G), the first of a set of shared motifs among methyltransferase enzymes.[7] The central dimerization motif within the substrate binding domain makes up an "arm" that facilitates dimerization. The C-terminal catalytic core is the substrate binding domain, which is necessary for recognition of the arginine to be methylated. CARM1 has been shown to methylate both histone and non-histone substrates.[8–13] In contrast to most PRMTs that methylate glycine-rich and arginine-rich patches (GAR motifs) within their substrates, CARM1 does not methylate the GAR motif but rather has a unique set of methylation targets,[1] which have no apparent recognition sequence.

The function of CARM1 has recently been shown to be regulated by post-translational modifications (PTMs). Phosphorylation of serine 228 blocks CARM1 dimerization, likely interrupting proper binding of the dimerization "arm" in phosphorylated CARM1 and thus inhibits its methylation activity.[14] Automethylation of CARM1 has been previously implicated *in vitro*[10] in analog with other PRMT family members including PRMT1, PRMT6, and PRMT8,[15,16] however, the automethylation of CARM1 and PRMT1 appeared to be significantly less than PRMT6 and PRMT8. The site(s) of CARM1 automethylation are unknown and the role of PRMT automethylation remains to be elucidated.

Mass spectrometry (MS) is the only technique that can universally provide information about protein PTMs without a priori knowledge of the modification.[17] In the conventional "bottom-up" MS approach proteins of interest are digested with an enzyme prior to MS analysis providing only partial coverage of the protein sequence.[18,19] In contrast, top–down MS analyzes intact protein directly, providing "a bird's eye view" of all possible protein modifications with full sequence coverage.[20–33] The top–down MS approach subsequently fragments the protein ions of interest in the mass spectrometer to locate the modification site(s), which allows fine mapping of modifications in the intact protein. Moreover, the top–down MS approach is especially attractive for estimating the relative abundance of protein species with specific modifications since the ionization efficiency of intact proteins is much less affected by the presence of modifying groups in comparison with peptides.[30,32,34,35] The newly developed MS/MS technique of electron capture dissociation (ECD)[36] greatly improves both efficiency and sequence coverage in top–down analyses. ECD cleaves NH-CHR bonds to produce mainly c and $z^{\cdot}$ ions, complementary to those from the well developed energetic dissociation methods such as collisionally activated dissociation (CAD).[21] More importantly, ECD generates far more fragmentation ions as well as unique cleavages due to its nonergodic dissociation of covalent protein backbone bonds causing local fast ($<10^{-12}$ s) cleavages of covalent bonds before energy randomization.[21,23,37,38]

In this study, we report the discovery and characterization of an unexpected automethylation event by *Ag*CARM1, a CARM1 homologue in the mosquito model *Anopheles gambiae*, using top–down high resolution Fourier transform (FT) tandem mass spectrometry (MS/MS). *Ag*CARM1 automethylation is shown to occur intermolecularly during expression in *Escherichia coli* as well as *in vitro*.

Results

Ag*CARM1 is the* A. gambiae *homologue of CARM1*

CARM1 contains a conserved PRMT catalytic core,[39] which is highly conserved across species even between mammals and insects. The *Drosophila* CARM1 homologue CARMER (or DART4) has been reported to have similar substrate specificity to mouse CARM1 (mCARM1) but is lacking the C-terminal activation domain found in other CARM1 sequences (Supporting Fig. S1).[40,41] To develop an animal model for CARM1 function, we chose to study CARM1 in the mosquito model *A. gambiae* since this model organism has a sequenced genome and a simplified hormone signaling pathway. We have cloned a sequence that is highly similar to known CARM1 genes from the *A. gambiae* genome (*Ag*CARM1, Genbank: FJ391182) and generated highly purified $AgCARM1^{1-509}$ protein in large quantities [Fig. 1(A)]. To verify $AgCARM1^{1-509}$ is the functional homologue of CARM1, we performed

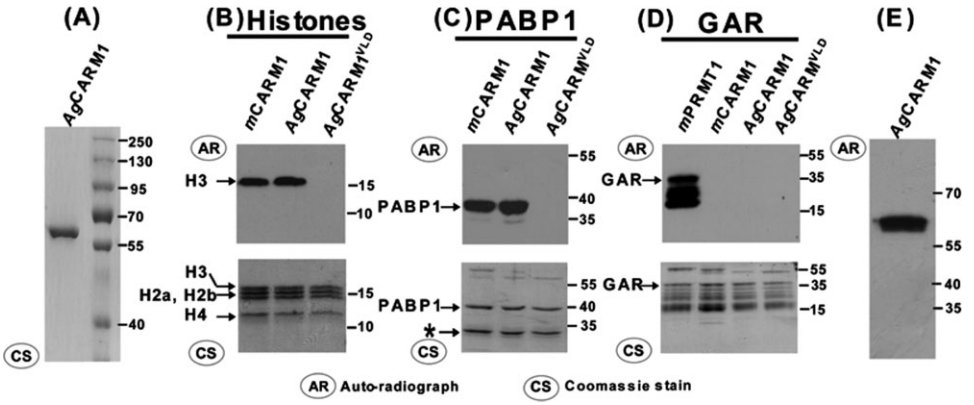

Figure 1. *In vitro* methylation studies of *Ag*CARM1$^{1–509}$ show similar substrate specificity to mCARM1. (A) Coomassie stain of Ni-NTA purified *Ag*CARM1.$^{1–509}$ (B,C) *In vitro* methylation of core histones (B) and PABP1 (C) by CARM1 variants in the presence of ^{3}H-AdoMet. Top panel is an autoradiograph showing radiolabel incorporation of the methyl group. Bottom panel is the commassie stained gel after exposure to film. (D) *In vitro* methylation of the fibronectin fragment GAR by various PRMTs in the presence of ^{3}H-AdoMet. (E) Autoradiograph of *Ag*CARM1$^{1–509}$ incubated without substrate in an otherwise identical *in vitro* methylation assay as seen in B–D. *Band that copurifies with PABP1, likely a degradation product. GST-GAR also exhibits significant degradation consistent with previous publications of this protein fragment. CS, Coomassie Stain; AR, Autoradiograph.

in vitro methylation assays using purified *Ag*CARM1$^{1–509}$ and *Ag*CARM1^{VLD} (catalytic inactive mutant where the amino acids in the AdoMet binding domain, VLD, were mutated to AAA) and compared their methylation specificity to mCARM1 and mPRMT1. These were incubated with known CARM1 substrates polyA binding protein 1 (PABP1), core histones from Hela cells, as well as GAR, a common substrate for PRMT family members including PRMT1 but not CARM1. *Ag*CARM1$^{1–509}$ was found to methylate the CARM1 specific targets exclusively [Fig. 1(B,C)], while exhibiting no activity towards GAR, which is methylated by PRMT1 [Fig. 1(D)]. Analogous to the mCARM1^{VLD} mutant,[2] the *Ag*CARM1^{VLD} mutant showed a complete loss of methylation activity [Fig. 1(B,C)]. *Ag*CARM1$^{1–509}$ also exhibits automethylation activity *in vitro* [Fig. 1(E)] as reported previously for mCARM1.[10] These results confirm that *Ag*CARM1$^{1–509}$ is the CARM1 homologue in *A. gambiae*.

Ag*CARM1*$^{1–509}$ *overexpressed in* E. coli *is dimethylated at R485* in vivo

We used top–down MS to characterize *Ag*CARM1$^{1–509}$ expressed in *E. coli,* later referred to as *Ag*CARM1^{WT}. After purification, the *Ag*CARM1^{WT} protein sample was directly analyzed by LTQ FT ultra high resolution mass spectrometer. Electrospray ionization (ESI)/FTMS spectrum revealed two molecular forms of *Ag*CARM1^{WT} (Fig. 2, top). The most abundant molecular weight of the minor form is 58549.36–58549.38, which is consistent with the DNA predicted mass value for *Ag*CARM1^{WT} with the removal of its N-terminal methionine (Calc'd M_r 58549.31–58549.38, 0.9 ppm). Surprisingly, the molecular weight of the major form, 58577.37–58577.38, is 28 Da more than the minor form suggesting a dimethylated form of *Ag*CARM1^{WT} (Fig. 2, bottom). Eighty-three percentage of the total *Ag*CARM1^{WT} was dimethylated and 17% was unmethylated, with no detectable amount of monomethylated *Ag*CARM1^{WT} based on peak height. This relative quantification is based on the assumption that the ionization efficiency of intact proteins is much less affected by a modifying group in comparison to peptides.[30] Since the analyzed sample was freshly purified after over-expression, it is conceivable that this methylation event occurred *in vivo* after expression in *E. coli*.

To locate the specific site of dimethylation, multiple charge states of *Ag*CARM1^{WT} were individually isolated and then fragmented by CAD and ECD. The six combined CAD spectra generated 39 *b* and 47 *y* ions whereas 3 ECD spectra generated 66 *c* and 12 $z^{\cdot}$ fragment ions representing 78 cleavages each. These covered 26% of the total 521 NH-CH available backbone bonds in *Ag*CARM1^{WT} and confirmed removal of methionine at the N-terminus. Fragmentation ions were consistent with dimethylation of R485, showing a +28 Da shift in those fragments containing the R485 residue. MS/MS spectra of a mixture of unmethylated and dimethylated *Ag*CARM1^{WT} show predominantly (~83%) dimethylated form of fragment ions containing R485 (Supporting Fig. S2) suggesting R485 is the only site for dimethylation. Large *b* and *y* ions were observed in CAD spectra including several complementary pairs although fragmentation occurred most frequently in the terminal regions of the protein (Fig. 2, bottom). The overlapping *b* and *y* ions span the entire sequence giving confidence that the assigned primary structure is correct. The Asn39-Pro40 bond was very sensitive to CAD dissociation, resulting in not only the complementary ion pair (y_{483} and b_{39}) but also a series of internal ions with Asn39 as the C-terminus (data not shown). This pattern was observed in all *Ag*CARM1 mutants analyzed hereafter.

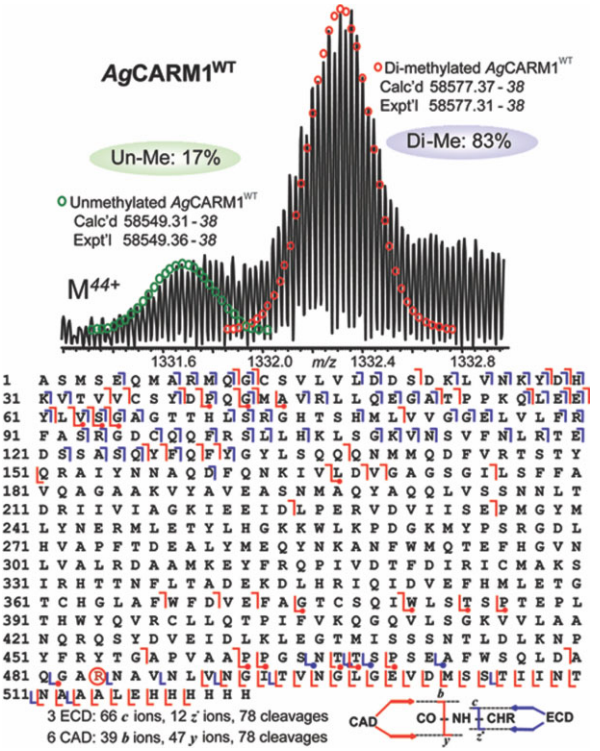

Figure 2. High resolution ESI/FTMS analysis of *Ag*CARM1^{WT} expressed in *E. coli*. Top, ESI/FTMS spectrum of *Ag*CARM1^{WT} intact protein ions (M^{44+}), suggesting *Ag*CARM1^{WT} is predominantly dimethylated. Bottom, fragmentation map from 3 ECD and 6 CAD spectra of both un- and di-methylated *Ag*CARM1^{WT} matched with assignments to the DNA-predicted sequence of *Ag*CARM1^{WT} with the removal of N-terminal methionine, localizing the dimethylation site to R485 (highlighted in circle). The fragmentation ions containing the dimethylated form were indicated in dots.

To validate that R485 is the sole methylation site on *Ag*CARM1^{WT}, R485 was mutated to lysine to maintain the positive charge while blocking methylation. The *Ag*CARM1^{R485K} mutant was expressed and purified as for wild-type protein. Only one molecular form was shown in ESI/FTMS spectrum, matching the mass of unmethylated *Ag*CARM1^{R485K} (Fig. 3, top). Similarly, CAD and ECD data confirmed the lack of detectable methylation in *Ag*CARM1^{R485K}, which is estimated to be <3% of the total protein ion population (Fig. 3, bottom). Two ECD spectra generated 63 *c* and 26 *z*· fragment ions representing 89 cleavages and two CAD spectra generated 24 *b* and 26 *y* ions representing 39 cleavages. This result confirms that R485 is the only significant methylation site on CARM1 during its expression in *E. coli*. Similar to the *Ag*CARM1^{WT} data, very large fragments were observed in broadband CAD spectra (see Fig. 4). The largest *b*/*y* fragmentation ions observed were b_{495} (Calc'd 55576.66–55576.37; Expt'l 55576.57–55576.37) and y_{483} (Calc'd 54169.99–54169.36; Expt'l 54169.94–54169.36), which is, to our knowledge, the largest fragmentation ions isotopically resolved in a MS/MS spectrum.

Ag*CARM1*^{R485K} *exhibits normal methylation activity towards other CARM1 substrates*

To examine whether methylation of R485 affects the methyltransferase activity of *Ag*CARM1, either *Ag*CARM1^{WT} or *Ag*CARM1^{R485K} was incubated with substrates PABP1, p300 and core histones in the presence of ^{3}H-AdoMet in an *in vitro* methylation assay. Methylation of substrates was determined to be at the same level between *Ag*CARM1^{WT} and *Ag*CARM1^{R485K} [Fig. 5(A)]. These results suggest that the methyltransferase activity of *Ag*CARM1 is independent of its methylation status at R485.

Ag*CARM1*^{WT} *is automethylated during its expression in* E. coli

There are no known arginine methyltransferases in *E. coli*. To ensure that methylation of *Ag*CARM1 in bacteria is due to automethylation activity and not a previously unidentified endogenous bacterial enzyme, the methyltransferase defective *Ag*CARM1^{VLD} mutant was bacterially expressed and analyzed by top–down MS. This mutation is not proximal to arginine 485, and so will likely maintain the sequence and structural

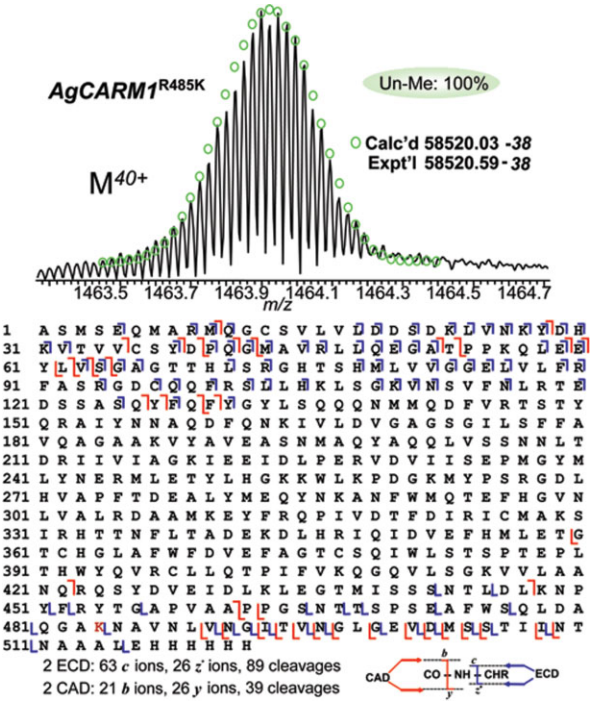

Figure 3. High resolution ESI/FTMS analysis of *Ag*CARM1^{R485K} mutant expressed in *E. coli*. Top, ESI/FTMS spectrum of *Ag*CARM1^{R485K} intact protein ions (M^{40+}), suggesting *Ag*CARM1^{R485K} is not methylated. Bottom, fragmentation map from 2 ECD and 2 CAD spectra of *Ag*CARM1^{R485K} matched with assignments to the DNA-predicted sequence of *Ag*CARM1^{R485K} without the N-terminal methionine.

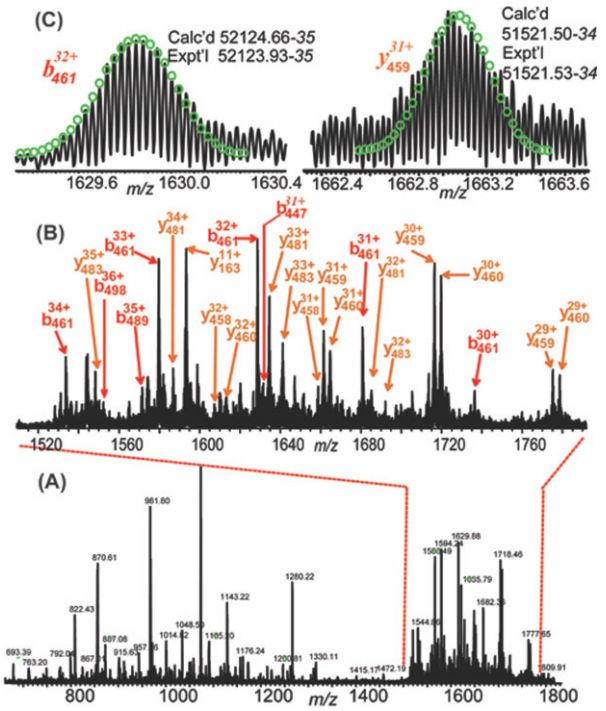

Figure 4. CAD spectrum of *Ag*CARM1^{R485K} intact protein ions. (A) CAD spectrum of isolated single charge state of *Ag*CARM1^{R485K} (M^{40+}). (B) Expanded spectrum of (A) showing large fragmentation ions observed at *m/z* 1500–1800. (C) Representative isotopically resolved large fragmentation ions.

integrity of the methylation site, thus serving as an effective control. A single molecular form was observed at 58433.83–58433.38 (Calc'd 58433.97–58433.38, 2 ppm), consistent with the unmethylated mass of *Ag*CARM1^{VLD} with the removal of N-terminal methionine. The overlapping *b* and *y* ions fragmentation ions from CAD of *Ag*CARM1^{VLD} span the entire sequence, which were consistent with an unmethylated parent ion (see Fig. 6). A single ECD spectrum generated 77 *c* and 38 *z*˙ fragment ions representing 115 cleavages; in contrast, one CAD spectrum generated 24 *b* and 24 *y* ions representing 47 cleavages. ECD fragmentation data showed a very rich fragmentation pattern mainly located on the two termini (Supporting Fig. S3). Lack of fragmentation of the central region of the protein was consistent among the *Ag*CARM1^{VLD}, *Ag*CARM1^{R485K}, and *Ag*CARM1^{WT} proteins. These results suggest that methylation is dependent on *Ag*CARM1 activity and *Ag*CARM1$^{1–509}$ is automethylated when expressed in bacteria.

Automethylation of Ag*CARM1 occurs intermolecularly*

In contrast to *Ag*CARM1^{WT}, whose automethylation indicates that it acts as both enzyme and substrate, no automethylation was observed *in vitro* when *Ag*CARM1^{R485K} or *Ag*CARM1^{VLD} mutants were incubated alone with ^{3}H-SAM [Fig. 5(B), compare lane 1 with lane 2 and 3]. *Ag*CARM1^{R485K} shows a lack of auto-methylation activity due to its inability to act as substrate in this reaction, while *Ag*CARM1^{VLD} has no enzymatic activity but still contains the methylation site arginine residue. To determine how the *Ag*CARM1 mutants and wild-type proteins interact, a number of mixed *in vitro* methylation reactions were performed.

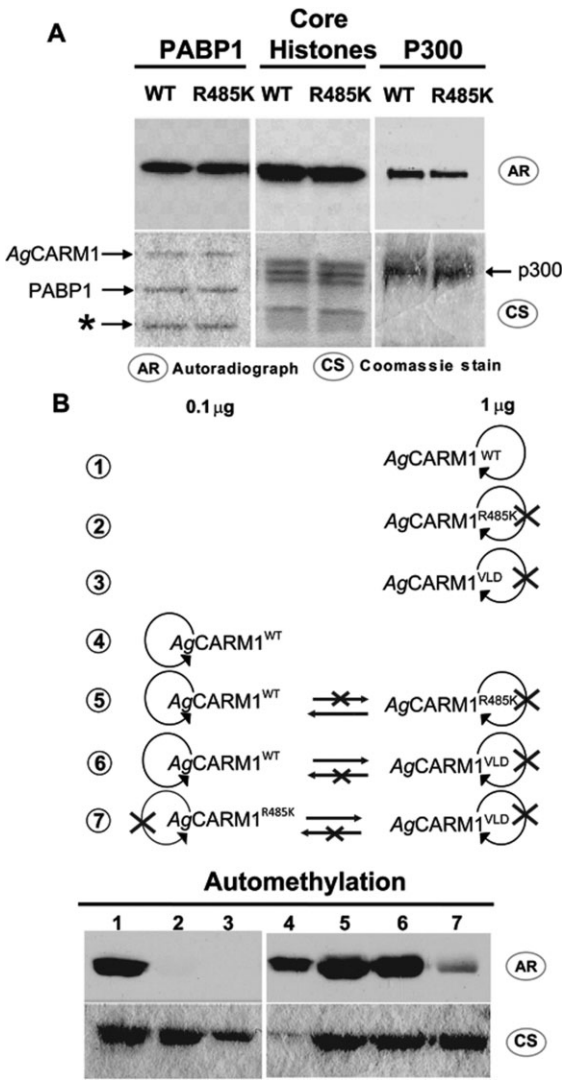

Figure 5. *In vitro* methylation assays characterizing *Ag*CARM1 automethylation and its affect on enzyme activity. ^{3}H-AdoMet was incubated with purified protein followed by SDS-PAGE, coomassie staining, and exposure to film. (A) *In vitro* methylation of known CARM1 substrates by *Ag*CARM1^{WT} and *Ag*CARM1^{R485K}. (B) Isolated and mixed automethylation reactions with 0.1 µg and/or 1 µg of *Ag*CARM1. Top shows a cartoon representation of reaction contents and possible interactions. Looping arrows indicate possible intramolecular and/or intermolecular automethylation. Straight arrows indicate possible intermolecular automethylation. Bottom shows *in vitro* methylation results of these reaction conditions, with lane numbers corresponding to the same numbered cartoon above. *Band copurified with PABP1, likely due to degradation. CS, Coomassie Stain; AR, Autoradiograph.

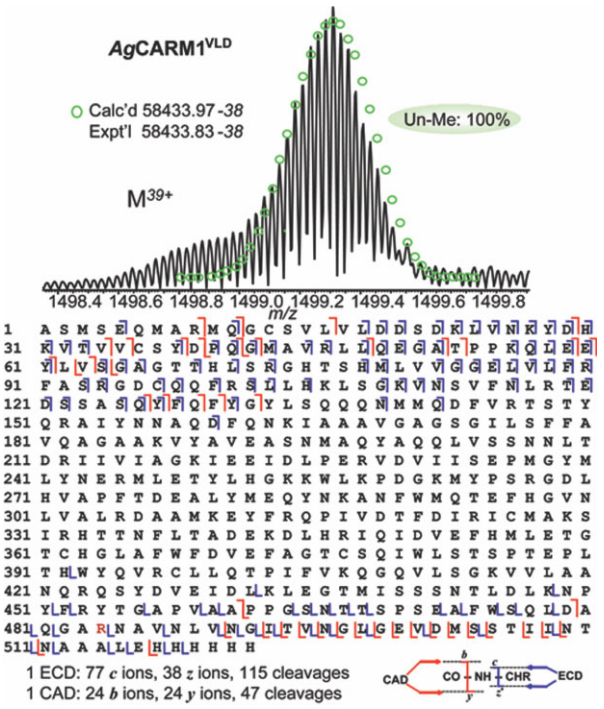

Figure 6. Top–down MS analysis of *Ag*CARM1^{VLD} mutant. Top, ESI/FTMS spectrum of *Ag*CARM1^{VLD} (M^{39+}), intact protein ions, suggesting *Ag*CARM1^{VLD} is not methylated. Bottom, Fragmentation map from both 1 CAD and 1 ECD spectra with assignments to the DNA-predicted sequence of *Ag*CARM1^{VLD} with the removal of N-terminal Met. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Although the *Ag*CARM1^{R485K} and *Ag*CARM1^{VLD} mutants do not automethylate in isolated reactions, mixing *Ag*CARM1^{WT} either with *Ag*CARM1^{R485K} or *Ag*CARM1^{VLD} at a 1:10 ratio *in vitro* results in increased methylation above that of *Ag*CARM1^{WT} alone [Fig. 5(B), compare lanes 5 and 6 with lane 4]. Mixing of *Ag*CARM1^{R485K} and *Ag*CARM1^{VLD} shows a weak methylation *in vitro* even though neither form shows automethylation activity by itself [Fig. 5(B), lane 7]. These data support an intermolecular mechanism for the automethylation of *Ag*CARM1 since these increases in methylation must be due to the interaction between different *Ag*CARM1 proteins.

Discussion

We have biochemically characterized a novel CARM1 homologue from the mosquito *A. gambiae*. *Ag*CARM1 shows conserved substrate specificity for CARM1 methylation targets [Fig. 1(B,C)] as well as exhibiting automethylation activity both *in vivo* during recombinant expression (see Fig. 2) and *in vitro* [Fig. 1(E)]. Unexpectedly, our top–down MS results indicate that *Ag*CARM1^{WT} has already been predominantly automethylated during its expression in *E. coli*. Recombinant proteins expressed in *E. coli* are generally thought to be lacking methylation and other PTMs; however, our study clearly shows protein arginine methyltransferase can utilize bacterial AdoMet and be automethylated when expressed in bacteria, suggesting that the same principle may apply to other post-translational modifying enzymes. Since PTMs can affect the activity or binding specificity of a given protein,[34,42,43] the possibility of PTM during recombinant expression may create structural and functional diversity that would affect characterization of these proteins.[44] This phenomenon may apply to other methyltransferases and specifically PRMT enzymes. Previous observation of weakly methylated CARM1 and PRMT1 *in vitro* has been logically attributed to weak automethylation activity.[16] Our results indicate that a weak *in vitro* automethylation signal can also be attributed to near saturated automethylation *in vivo* prior to purification of the enzyme. Thus, it may be inaccurate to compare the activity of these enzymes through *in vitro* methylation without prior characterization of their methylation states.

Mutation of R485 to lysine appears to have no effect on *Ag*CARM1's enzymatic activity, as CARM1 substrates are methylated equally by *Ag*CARM1^{WT} and *Ag*CARM1^{R485K}. This suggests that dimerization, AdoMet binding, and substrate specificity are not affected by loss of automethylation. R485 is located in the C-terminal domain adjacent to the substrate-binding site and, consistent with the above results, the corresponding region in mCARM1 is unstructured and not essential for the enzymatic activity of CARM1.[3] The *Ag*CARM1^{R485K} mutant shows a loss of automethylation both in bacteria (see Fig. 3) and when isolated *in vitro* [Fig. 5(A)], suggesting that R485 is the only detectable automethylation site on *Ag*CARM1. R485 is highly conserved in other sequenced insect species (Supporting Fig. S1).

The mechanism of automethylation has yet to be well characterized in PRMTs or other methyltransferases. An automethylation site on the lysine methyltransferase G9a has been determined,[45,46] and this work suggests that G9a automethylation may occur by an intramolecular mechanism. While we were not able to confirm or rule out intramolecular activity by *Ag*CARM1, we do show direct evidence of intermolecular methylation using a series of mixed *in vitro* methylation reactions [Fig. 5(B)]. Mixing *Ag*CARM1^{WT} with the catalytically inactive *Ag*CARM1^{VLD}, which can act only as a substrate in an automethylation reaction, increases radiolabeling above that of *Ag*CARM1^{WT} alone. Since *Ag*CARM1^{VLD} cannot contribute to the enzymatic activity of the reaction, this result indicates that *Ag*CARM1^{VLD} is being methylated by *Ag*CARM1^{WT}. Similar experiments show that *Ag*CARM1^{R485K} is able to methylate *Ag*CARM1^{VLD} in a similar manner. Mixing *Ag*CARM1^{WT} with *Ag*CARM1^{R485K} also shows an increase in the intensity of radiolabeling above *Ag*CARM1^{WT} alone, suggesting that *Ag*CARM1^{R485K} is methylating *Ag*CARM1^{WT}, since it does not appear to methylate itself. Altogether these data support an intermolecular mechanism for automethylation.

The ESI/FTMS analysis of *Ag*CARM1 expressed in bacteria showed predominant dimethylated protein, while the remaining *Ag*CARM1^{WT} was unmethylated. This observation was supported by fragmentation spectrum (Supporting Fig. S2), which showed a similar ratio of dimethylated and unmethylated forms with no detectable monomethylated form in fragmentation ions containing the dimethylation site, R485. We speculate that the unmethylated *Ag*CARM1^{WT} represents the freshly expressed protein, while the dimethylated *Ag*CARM1^{WT} may be expressed earlier and automethylated over time. This notion is supported by the top–down MS result that *Ag*CARM1^{WT} can be further dimethylated when incubated with cold AdoMet *in vitro* (data not shown). The specific lack of monomethylated *Ag*CARM1^{WT} suggests that *Ag*CARM1 may bind its substrate continuously during methylation, which is known as a processive mechanism. If validated, this would be in contrast to data from other Type I PRMTs such as PRMT1, PRMT3, and PRMT6. PRMT6 has been shown to methylate synthetic peptide substrates using a distributive mechanism, where a majority of substrate is monomethylated before dimethylation occurs.[47] PRMT1's mechanism is still under debate, as recent works suggest both a partially processive mechanism,[48] and a distributive mechanism.[49] These studies focused on the methylation of peptides instead of intact proteins. Thus, the role of extended substrate structure and the unique character of automethylation in defining the mechanism of PRMT activity remain poorly defined. Although not conclusive, our results indicate that automethylation may have a different mechanism than other substrates' methylation by PRMT family members.

The top–down MS used in our study allows us to unambiguously determine any and all detectable modifications of *Ag*CARM1. We were able to resolve the intact *Ag*CARM1 protein (>58 kDa) in a LTQ/FT mass spectrometer and achieved extensive fragmentation by CAD and ECD. Extremely large fragmentation ions (>55 kDa) were isotopically resolved in CAD spectra. The overlapping *b* and *y* ions span the entire sequence giving confidence that the assigned primary structure is correct. ECD provided more backbone cleavages than CAD; here one ECD spectrum of one isolated single charge state of a 58 kDa protein generated as many as 115 cleavages out of 521 bonds. Therefore, top–down MS shows great promise in the elucidation and quantitation of large intact proteins.[20–33] Recently, Webb *et al.* characterized methylation of lysine residues in yeast ribosomal protein Rpl42ab (12 kDa) via such a top–down MS approach. Despite its enormous promise, top–down MS has yet to be popularized mainly due to the high-resolution instrument requirement and technical difficulties in handling large intact proteins. Desalting is required before a protein sample can be analyzed on a mass spectrometer, however, proteins are more likely to precipitate in a solution free of salt. To overcome this dilemma, we found that it is critically important to analyze proteins freshly purified and desalted. Another limitation to top–down MS is the sensitivity of currently commercially available mass spectrometers. We need to acquire 1000–3000 transients per spectrum to ensure high quality MS and MS/MS, which required a relatively large amount of high purity proteins (>20 μg) and a highly stable electrospray. Hence, improvements in the instrument (i.e., sensitivity and resolution of a mass spectrometer) and software for automatic data processing[50] will greatly facilitate the use of top–down MS for large proteins.

Materials and Methods

Cloning, expression, and purification

*Ag*CARM1 was amplified from *A. gambiae* cDNA. The PCR product of the forward primer: (5′CGACTTG CTAGCATGAGTGAACAAATGGCTCG 3′) and reverse primer: (5′ATGGCAGCGGCCGCGTTGGTGTTGATGAT TGTCGACG 3′) was restriction digested and ligated into pET21a, adding seven N-terminal amino acids (MARMSEQ) and an C-terminal His tag *Ag*CARM1$^{1–509}$. *Ag*CARM1^{R485K}, *Ag*CARM1^{VLD}, GST-PABP1, and GST-GAR were expressed in BL21 cells and purified using tag affinity chromatography (Qiagen). mCARM1 and mPRMT1 were purified from insect cells as described previously.[10]

Site-directed mutagenesis

pET21a-*Ag*CARM1$^{1–509}$ was mutated by amplification with complementary primers containing mutations to induce codon changes. *Ag*CARM1^{VLD} forward primer (5′CGCAAGACTTCCAGAATAAGATCG***C***G***GCA***G***C***TGTG GGGGCCGGATCCG 3′) and *Ag*CARM1^{R485K} forward primer: (5′CAGCTGGACGCACAGGGCGCC***AAA***AATG CGGTCAACCTCGTCAAC3′) (bold = mutated).

In vitro *methylation assay*

Various enzymes and substrates were incubated in 15 μL of 5 m*M* $MgCl_2$, 20 m*M* Hepes pH 7.9, 1 m*M* EDTA, 1 m*M* DTT, 10% glycerol containing 1 μL of ^{3}H-S-adenosylmethionine (GE Healthcare) for 1 h. This was then separated by SDS-PAGE and fixed in methanol for 1 h. Fixed gels were incubated in "Amplify" (Amersham Biosciences) scintillation fluid for 30 min before exposure to film.

Top–down mass spectrometry

All the protein samples (20–100 μg) were desalted by ultrafiltration using microcon centrifugal filter devices (Millipore, Billerica, MA), dissolved in 10–50% ACN and 1% acetic acid in water at 0.2–1 μg/μL, and introduced to the mass spectrometer using an automated chip-based nanoESI source, the Triversa NanoMate (Advion BioSciences, Ithaca, NY) with a spray voltage of 1.25–1.6 kV versus the inlet of the mass spectrometer, resulting in a flow of 50–200 nL/min. Intact protein molecular ions were analyzed using a

linear trap/FT-ICR (LTQ FT Ultra) hybrid mass spectrometer (Thermo Scientific, Bremen, Germany). Capillary temperature was 200°C, capillary voltage was 36 V, and the tube lens voltage was 190 V. Ion transmission into the linear trap and further to the FT-ICR cell was automatically optimized for maximum ion signal with the target values (the approximate number of accumulated ions) for a full MS scan linear trap (LT) scan, FT-ICR cell (FT) scan, MS^n FT-ICR, ECD scan were 3×10^4, 10^6, 10^6, and 8×10^6, respectively. The resolving power of the FT-ICR mass analyzer was set at 100,000 or 400,000 $m/\delta m_{50\%}$ at m/z 400, resulting in one scan/s acquisition rate. Individual charge states of the protein molecular ions were first isolated and then dissociated by ECD using 2–3% "electron energy" and a 50–70 ms duration time with no delay. Isolated charge states were also dissociated by CAD using 15–20% collision energy. Up to 3000 transients were averaged per spectrum to ensure high quality ECD/CAD spectra. All FT-ICR spectra were processed with Xtract Software (FT programs 2.0.1.0.6.1.4, Xcallibur 2.0.5, Thermo Scientific, Bremen, Germany) using a signal to noise threshold of 1.5 and fit factor of 60% and validated manually. The resulting monoisotopic mass lists were further assigned using in-house "Ion Assignment" software (Version 1.0) based on the recombinant CARM1 protein sequence generated from DNA sequencing with consideration of the static (mono-, di-) methylation and 10 ppm monoisotopic precursor and fragment tolerance. The M_r value reported in the study are all most abundant masses and the mass difference (in units of 1.00235 Da) between the most abundant isotopic peak and the monoisotopic peak is denoted in italics after each M_r value. The top 10 most abundant isotopic peak heights were integrated to calculate the relative abundance of the intact protein.

Conclusions

We present, here, the discovery and comprehensive characterization of an automethylation event by AgCARM1 using top–down high resolution MS/MS. Unexpectedly, we found *Ag*CARM1^{WT} has already been predominantly automethylated on a single site during its expression in *E. coli* via this top–down MS approach which could easily have been missed in traditional bottom–up MS. Mutation of this site reveals automethylation occurs intermolecularly and is consistent with a processive model for *Ag*CARM1^{WT} automethylation. This study may shed light on other bacterially expressed post-translational modification enzymes, which could be modified but overlooked in biochemical and structural studies. Top–down high resolution MS/MS shows great promise for revealing unexpected protein modification and localizing such modification to one amino acid.

Acknowledgments

The authors thank Susan Paskewitz (UW-Madison) for the generous gift of *A. gambiae* cDNA, Mark Bedford (University of Texas MD Anderson Center) for the generous gift of GST-PABP1 and GST-GAR plasmids, Dick Burgess (UW-Madison) for helpful discussions, Yongna Xing (UW-Madison) and Tadhg Begley (Cornell University) for critical reading of the manuscript. Financial support was kindly provided by Shaw Scientist Award from Greater Milwaukee Foundation and NIH (W.X.) and the Wisconsin Partnership Fund for a Healthy Future (Y.G.).

References

1. Bedford MT, Richard S (2005) Arginine methylation: an emerging regulator of protein function. Mol Cell 18:263.
2. Chen D, Ma H, Hong H, Koh SS, Huang SM, Schurter BT, Aswad DW, Stallcup MR (1999) Regulation of transcription by a protein methyltransferase. Science 284: 2174–2177.
3. Teyssier C, Chen D, Stallcup MR (2002) Requirement for multiple domains of the protein arginine methyltransferase CARM1 in its transcriptional coactivator function. J Biol Chem 277:46066–46072.
4. Troffer-Charlier N, Cura V, Hassenboehler P, Moras D, Cavarelli J (2007) Functional insights from structures of coactivator-associated arginine methyltransferase 1 domains. EMBO J 26:4391–4401.
5. Yue WW, Hassler M, Roe SM, Vale VT, Pearl LH (2007) Insights into histone code syntax from structural and biochemical studies of CARM1 methyltransferase. EMBO J 26:4402–4412.
6. Martin JL, McMillan FM (2002) SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold. Curr Opin Struc Biol 12:783–793.
7. Katz JE, Dlakic M, Clarke S (2003) Automated identification of putative methyltransferases from genomic open reading frames. Mol Cell Proteomics 2:525–540.
8. Ma H, Baumann CT, Li H, Strahl BD, Rice R, Jelinek MA, Aswad DW, Allis CD, Hager GL, Stallcup MR (2001) Hormone-dependent, CARM1-directed, arginine-specific methylation of histone H3 on a steroid-regulated promoter. Curr Biol 11:1981.
9. Schurter RT, Koh SS, Chen D, Bunick GJ, Harp JM, Hanson BL, Henschen-Edman A, Mackay DR, Stallcup MR, Aswad DW (2001) Methylation of histone H3 by coactivator-associated arginine methyltransferase 1. Biochemistry 40:5747–5756.
10. Xu W, Chen H, Du K, Asahara H, Tini M, Emerson BM, Montminy M, Evans RM (2001) A transcriptional switch mediated by cofactor methylation. Science 294:2507–2511.
11. Lee J, Bedford MT (2002) PABP1 identified as an arginine methyltransferase substrate using high-density protein arrays. EMBO Rep 3:268–273.
12. Feng Q, Yi P, Wong J, O'Malley BW (2006) Signaling within a coactivator complex: methylation of SRC-3/AIB1 is a molecular switch for complex disassembly. Mol Cell Biol 26:7846–7857.
13. Naeem H, Cheng D, Zhao Q, Underhill C, Tini M, Bedford MT, Torchia J (2006) The activity and stability of the transcriptional coactivator p/CIP/SRC-3 is regulated by CARM1-dependent methylation. Mol Cell Biol 27: 120–134.
14. Higashimoto K, Kuhn P, Desai D, Cheng X, Xu W (2007) Phosphorylation-mediated inactivation of coactivator-associated arginine methyltransferase 1. Proc Natl Acad Sci USA 104:12318–12323.

15. Frankel A, Yadav N, Lee J, Branscombe TL, Clarke S, Bedford MT (2002) The novel human protein arginine N-methyltransferase PRMT6 is a nuclear enzyme displaying unique substrate specificity. J Biol Chem 277:3537–3543.
16. Sayegh J, Webb K, Cheng D, Bedford MT, Clarke SG (2007) Regulation of protein arginine methyltransferase 8 (PRMT8) activity by its N-terminal domain. J Biol Chem 282:36444–36453.
17. Nedelkov D, Kiernan UA, Niederkofler EE, Tubbs KA, Nelson RW (2006) Population proteomics—the concept, attributes, and potential for cancer biomarker research. Mol Cell Proteomics 5:1811–1818.
18. Chait BT (2006) Mass spectrometry: bottom-up or top–down? Science 314:65–66.
19. Kjeldsen F, Savitski MM, Nielsen ML, Shi L, Zubarev RA (2007) On studying protein phosphorylation patterns using bottom-up LC-MS/MS: the case of human alpha-casein. Analyst 132:768–776.
20. Kelleher NL, Lin HY, Valaskovic GA, Aaserud DJ, Fridriksson EK, McLafferty FW (1999) Top down versus bottom up protein characterization by tandem high-resolution mass spectrometry. J Am Chem Soc 121:806–812.
21. Ge Y, Lawhorn BG, ElNaggar M, Strauss E, Park JH, Begley TP, McLafferty FW (2002) Top down characterization of larger proteins (45 kDa) by electron capture dissociation mass spectrometry. J Am Chem Soc 124:672–678.
22. Sze SK, Ge Y, Oh H, McLafferty FW (2002) Top–down mass spectrometry of a 29-kDa protein for characterization of any posttranslational modification to within one residue. Proc Natl Acad Sci USA 99:1774–1779.
23. Ge Y, ElNaggar M, Sze SK, Bin Oh H, Begley TP, McLafferty FW, Boshoff H, Barry CE (2003) Top down characterization of secreted proteins from Mycobacterium tuberculosis by electron capture dissociation mass spectrometry. J Am Soc Mass Spectrom 14:253–261.
24. Ge Y, Lawhorn BG, Elnaggar M, Sze SK, Begley TP, McLafferty FW (2003) Detection of four oxidation sites in viral prolyl-4-hydroxylase by top–down mass spectrometry. Protein Sci 12:2320–2326.
25. Zabrouskov V, Giacomelli L, van Wijk KJ, McLafferty FW (2003) New approach for plant proteomics—characterization of chloroplast proteins of Arabidopsis thaliana by top–down mass spectrometry. Mol Cell Proteomics 2: 1253–1260.
26. Kelleher NL (2004) Top–down proteomics. Anal Chem 76:196A–203A.
27. Meng FY, Forbes AJ, Miller LM, Kelleher NL (2005) Detection and localization of protein modifications by high resolution tandem mass spectrometry. Mass Spec Rev 24:126–134.
28. Zhai HL, Dorrestein PC, Chatterjee A, Begley TP, McLafferty FW (2005) Simultaneous kinetic characterization of multiple protein forms by top down mass spectrometry. J Am Soc Mass Spectrom 16:1052–1059.
29. Han XM, Jin M, Breuker K, McLafferty FW (2006) Extending top–down mass spectrometry to proteins with masses greater than 200 kilodaltons. Science 314:109–112.
30. Pesavento JJ, Mizzen CA, Kelleher NL (2006) Quantitative analysis of modified proteins and their positional isomers by tandem mass spectrometry: human histone H4. Anal Chem 78:4271–4280.
31. Xie YM, Zhang J, Yin S, Loo JA (2006) Top–down ESI-ECD-FT-ICR mass spectrometry localizes noncovalent protein-ligand binding sites. J Am Chem Soc 128: 14432–14433.
32. Zabrouskov V, Han XM, Welker E, Zhai HL, Lin C, van Wijk KJ, Scheraga HA, McLafferty FW (2006) Stepwise deamidation of ribonuclease A at five sites determined by top down mass spectrometry. Biochemistry 45:987–992.
33. Siuti N, Kelleher NL (2007) Decoding protein modifications using top–down mass spectrometry. Nat Methods 4:817–821.
34. Zabrouskov V, Ge Y, Schwartz J, Walker JW (2008) Unraveling molecular complexity of phosphorylated human cardiac troponin I by top down electron capture dissociation/electron transfer dissociation mass spectrometry. Mol Cell Proteomics 7:1838–1849.
35. Webb KJ, Laganowsky A, Whitelegge JP, Clarke SG (2008) Identification of two SET domain proteins required for methylation of lysine residues in yeast ribosomal protein Rpl42ab. J Biol Chem 283:35561–35568.
36. Zubarev RA, Horn DM, Fridriksson EK, Kelleher NL, Kruger NA, Lewis MA, Carpenter BK, McLafferty FW (2000) Electron capture dissociation for structural characterization of multiply charged protein cations. Anal Chem 72:563–573.
37. Zubarev RA, Kelleher NL, McLafferty FW (1998) Electron capture dissociation of multiply charged protein cations. A nonergodic process. J Am Chem Soc 120:3265–3266.
38. Breuker K, Oh HB, Lin C, Carpenter BK, McLafferty FW (2004) Nonergodic and conformational control of the electron capture dissociation of protein cations. Proc Natl Acad Sci USA 101:14011–14016.
39. Bedford MT (2007) Arginine methylation at a glance. J Cell Sci 120:4243–4246.
40. Boulanger M, Miranda T, Clarke S, Di Fruscio M, Suter B, Lasko P, Richard S (2004) Characterization of the Drosophila protein arginine methyltransferases DART1 and DART4. Biochem J 379:283–289.
41. Cakouros D, Daish TJ, Mills K, Kumar S (2004) An arginine-histone methyltransferase, CARMER, coodinates Ecdysone-mediated apoptosis in Drosophila cells. J Biol Chem 279:18467–18471.
42. Mann M, Jensen ON (2003) Proteomic analysis of post-translational modifications. Nat Biotechnol 21:255–261.
43. Karanam B, Jiang L, Wang L, Kelleher NL, Cole PA (2006) Kinetic and mass spectrometric analysis of p300 histone acetyltransferase domain autoacetylation. J Biol Chem 281:40292–40301.
44. Olivares-Illana V, Meyer P, Bechet E, Gueguen-Chaignon V, Soulat D, Lazereg-Riquier S, Mijakovic I, Deutscher J, Cozzone AJ, Laprévote O, Morera S, Grangeasse S, Nessler S (2008) Structural basis for the regulation mechanism of the tyrosine kinase CapB from staphylococcus aureus. PLoS Biol 6:e143.
45. Chin HG, Esteve PO, Pradhan M, Benner J, Patnaik D, Carey MF, Pradhan S (2007) Automethylation of G9a and its implication in wider substrate specificity and HP1 binding. Nucleic Acids Res 35:7313–7323.
46. Sampath SC, Marazzi I, Yap KL, Sampath SC, Krutchinsky AN, Mecklenbräuker I, Viale A, Rudensky E, Zhou MM, Chait BT, Tarakhovsky T (2007) Methylation of a histone mimic within the histone methyltransferase G9a regulates protein complex assembly. Mol Cell 27:596–608.
47. Lakowski TM, Frankel A (2008) A kinetic study of human protein arginine N-methyltransferase 6 reveals a distributive mechanism. J Biol Chem 283:10015–10025.
48. Osborne TC, Obianyo O, Zhang X, Cheng X, Thompson PR (2007) Protein arginine methyltransferase 1: positively charged residues in substrate peptides distal to the site of methylation are important for substrate binding and catalysis. Biochemistry 46:13370–13381.
49. Kolbel K, Ihling C, Bellmann-Sickert K, Neundorf I, Beck-Sickinger AG, Sinz A, Kühn U, Wahle E (2009) Type I arginine methyltransferases PRMT1 and PRMT-3 act distributively. J Biol Chem 284:8274–8282.
50. Meng FY, Cargile BJ, Miller LM, Forbes AJ, Johnson JR, Kelleher NL (2001) Informatics and multiplexing of intact protein identification in bacteria and the archaea. Nat Biotechnol 19:952–957.