BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Peptide Backbone Conformation Affects the Substrate Preference of Protein Arginine Methyltransferase I

Knut Kölbel,[†,#] Christian Ihling,[‡] Uwe Kühn,[†] Ines Neundorf,[§,⊥] Silke Otto,[†] Jan Stichel,[§] Dina Robaa,[‡] Annette G. Beck-Sickinger,[§] Andrea Sinz,*[,‡] and Elmar Wahle*[,†]

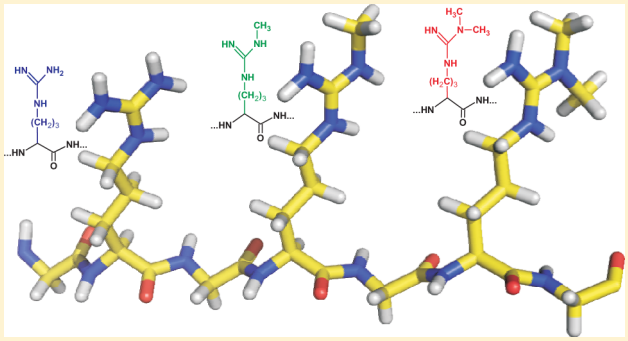
[†]Institute of Biochemistry and Biotechnology, Martin Luther University Halle-Wittenberg, Kurt-Mothes-Strasse 3, 06120 Halle, Germany

[‡]Institute of Pharmacy, Martin Luther University Halle-Wittenberg, Wolfgang-Langenbeck-Strasse 4, 06120 Halle, Germany

[§]Institute of Biochemistry, University of Leipzig, Brüderstrasse 34, 04103 Leipzig, Germany

Supporting Information

ABSTRACT: Asymmetric dimethylation of arginine side chains is a common post-translational modification of eukaryotic proteins, which serves mostly to regulate protein−protein interactions. The modification is catalyzed by type I protein arginine methyltransferases, PRMT1 being the predominant member of the family. Determinants of substrate specificity of these enzymes are poorly understood. The Nuclear poly(A) binding protein 1 (PABPN1) is methylated by PRMT1 at 13 arginine residues located in RXR sequences in the protein's C-terminal domain. We have identified a preferred site for PRMT1-catalyzed methylation in PABPN1 and in a corresponding synthetic peptide. Variants of these substrates were analyzed by steady-state kinetic analysis and mass spectrometry. The data indicate that initial methylation is directed toward the preferred arginine residue by an N-terminally adjacent proline. Enhanced methylation upon peptide cyclization suggests that induction of a reverse turn structure is the basis for the ability of the respective proline residue to enable preferred methylation of the neighboring arginine residue, and this notion is supported by far-UV circular dichroism spectroscopy. We suggest that the formation of a reverse turn facilitates the access of arginine side chains to the active sites of PRMT1, which are located in the central cavity of a doughnut-shaped PRMT1 homodimer.



Protein arginine methyl transferases (PRMTs) catalyze the dimethylation of arginine residues, a common posttranslational modification of eukaryotic proteins, which is important in modulating protein−protein interactions.[1−5] PRMTs are classified according to their products: Type I PRMTs form asymmetric dimethylarginine (ω-N^G,N^G-dimethylarginine), whereas type II PRMTs produce symmetric dimethylarginine (ω-N^G,N'^G-dimethylarginine). The only confirmed mammalian type II enzyme is PRMT5. Type I PRMTs, forming the apparently more common product, are PRMT1, -2, -3, -4, -6, and -8.[5−8]

Both types of PRMTs generate their products via the monomethylated intermediate, and both use *S*-adenosylmethionine (AdoMet) as the methyl donor. Type I enzymes obey a distributive reaction mechanism, in which both the sequential transfer of two methyl groups to a single arginine side chain and the modification of multiple side chains within one protein proceed through independent substrate binding events.[6,9−11] The catalytic core of the PRMTs is a conserved two-domain structure: The methyl donor is bound by the N-terminal Rossman fold domain, whereas the methyl-accepting arginine side chain is bound between this domain and the β-barrel forming the C-terminal domain.[12−17] Several PRMTs also present acidic grooves on their surfaces surrounding the active sites. These grooves are likely to bind additional positively charged amino acid residues in the vicinity of the methyl-accepting arginine and thus contribute to substrate selection.[14,16−18]

Substrate specificities of the PRMTs remain poorly characterized. The issue is complicated for a number of reasons: First, the PRMTs may associate with additional polypeptides in vivo that affect their substrate selection by additional interactions.[19,20] Second, the substrate spectra of different PRMTs are distinct but overlapping, so comparisons of methylated amino acid sequences are of little value unless the responsible methyl transferases are known. Furthermore, the conservation of methylated sequences is likely to reflect not only the determinants required for recognition by one or more methyl transferases but also the adaptation to a specific function, for example, binding to RNA and the nuclear import receptor in the case of the methylated C-terminal domain of

Received: March 22, 2012
Revised: June 6, 2012
Published: June 14, 2012

PABPN1 (see below). Finally, it is conceivable that PRMTs can bind their substrates in different orientations, each of which may have different sequence preferences.[11,14] Comparisons of synthetic peptides with full-length proteins from which they were derived showed that, at least in some cases, both substrates were methylated with similar efficiencies.[21,22] Thus, in these cases substrate recognition involves only local sequences surrounding the methyl accepting arginine residue. Nevertheless, more subtle contextual effects on substrate preferences and selection of methylated residues have been noted,[10,11,23,24] and accessibility of the peptide is likely to play an important role.[22] Even before the complex inventory of mammalian PRMTs was appreciated, a consensus sequence for arginine methylation, F/GGGRGGG/F, the "RGG box", was suggested on the basis of sequence comparisons.[25] However, even though a synthetic RGG box peptide is methylated by multiple PRMTs in vitro, it is not a particularly good substrate for the predominant enzyme, PRMT1.[21,22] Other substrates, like the nuclear poly(A) binding protein 1[26] or histones[27−30] do not conform to the RGG consensus. Studies with synthetic peptides have confirmed that PRMTs can modify a much wider range of substrates than indicated by the RGG box consensus.[11,22]

All type I PRMTs that have been crystallized show a characteristic homodimeric structure, in which the two monomers associate in a doughnut-like shape with the active sites accessible from the relatively small cavity in the middle.[12−17] Since the dimeric structure is important for enzyme function,[14,16−18] this arrangement of the active sites is likely to restrict the range of potential substrates through steric exclusion of larger folded structures. In fact, many observations support the notion that lack of a stable tertiary fold is an important determinant of substrate specificity (see Discussion).

The metazoan nuclear poly(A) binding protein 1 (PABPN1) is involved in mRNA polyadenylation.[31−36] All 13 arginine side chains within its 49 residue long C-terminal domain are quantitatively asymmetrically dimethylated in vivo.[26] The modification can be performed by PRMT1, -3 and -6, PRMT1 being the dominant enzyme.[22] Neither binding to poly(A) nor stimulation of poly(A) polymerase is affected by arginine methylation, even though the C-terminal domain is required for both.[32,37] Instead, methylation modulates the interaction of PABPN1 with its import receptor, transportin.[38] In contrast to many other PRMT substrates, PABPN1 contains 12 out of its 13 methylated arginines arranged in RXR motifs. The C-terminal domain harboring these motifs is likely to be unstructured: It carries a very high positive net charge, being entirely devoid of acidic residues, and it lacks amino acids with aliphatic side chains. Moreover, it is the binding site for transportin,[38,39] which is known to bind unstructured peptides.[40−42]

This study examines the proposition that backbone conformation is an important determinant of the substrate specificity of PRMTs. By comparing the PRMT1-mediated methylation of wild type full-length PABPN1, PAPBN1-derived peptides and a number of mutant variants, we identify a proline residue as a determinant for the preferred methylation of a C-terminally adjacent arginine. We present evidence suggesting that this effect is due to a reverse turn conformation induced by the proline residue.

EXPERIMENTAL PROCEDURES

Plasmids and Proteins. The plasmid pUK-CP encoding a fusion of CspB with the C-terminal domain of PABPN1 (CspRXR) was assembled as follows: Digestion of pUK-H4C[32] with NdeI and *Kpn*I left the open reading frame encoding the C-terminal domain of PABPN1 in the pUK vector. The open reading frame coding for *Bacillus subtilis* CspB[43] was amplified by PCR from pET11a-CspB (a gift of Franz-Xaver Schmid, University of Bayreuth, via Mirco Sackewitz, University of Halle) with the T7-promoter primer as forward primer and CGGGGTACCCGCTTCTTTAGTAACGGTAGC as reverse primer introducing a *Kpn*I restriction site. The PCR product was digested with NdeI and *Kpn*I and ligated into the digested pUK-H4C-plasmid in frame with the C-terminal domain of PABPN1.

PABPN1 variants R259A, R287A, and R294A were created by introduction of the respective point mutations into the plasmid pGM-synPABPN1[37] using the Gene Editor Site Directed Mutagenesis system (Promega). The mutagenesis primers were CAGCACAACAGACGCTGGCTTCCCAC (R259A); CAGTGGTTTTAACAGCGCGCCGCGGGGTCGCGTC (R287A); and GGGTCGCG-TCTATGCGGGCCGGGCTAG (R294A). PABPN1 variants P288A and R289A were created by site directed mutagenesis as described[32] using pGM-synPABPN1 as template and the oligonucleotides CAGTGGTTTTAACAGCAGGGCACGGGGTCGCGTCTACAG and CTGTAGACGCGACCCCGTGCCCTGCTGTTAAAACCACTG (PABPN1-P288A) as well as CAGTGGTTTTAACAGCAGGCCTGCAGGTCGCGTCTACAGGGGC and GCCCCTGTAGACGCGACCTGCAGGCCTGCTGTTAAAACCACTG (PABPN1-R289A) as primers. For expression of mutant PABPN1 in *Escherichia coli*, DNA fragments carrying the mutations were excised from pGM-synPABPN1 by XhoI and *Bam*HI and subcloned into pUK-synPABPN1ΔC49[37] opened with the same enzymes.

Expression of PABPN1 and its variants in *E. coli* BL21DE3 pUBS520 and purification by IMAC and ion exchange chromatography were conducted as described earlier.[37] CspRXR was obtained by the same procedure. His-tagged rat PRMT1v1 with a deletion of the first 10 amino acids[14] was expressed and purified as described.[22] Protein concentrations were determined spectrophotometrically using the theoretical absorption coefficients (PRMT1) or by quantification of band intensities from SDS−PAGE using a calibration curve from a BSA standard (PABPN1 variants).

Peptide Synthesis. All peptides were synthesized with free N- and amidated C-termini, except peptides T1 (free termini), T2 (acetylated N-terminus), and T3 (cyclic peptide). Peptides were synthesized according to the Fmoc/tBu solid-phase strategy (Syro peptide synthesizer; MultiSynTech, Witten, Germany). To obtain a free C-terminus or the amide, Wang or Rink amide resin, respectively, was used. The cyclic peptide T3 was synthesized as follows: Fmoc-Asp(OH)-ODmab (5 equiv in dimethylformamide) was coupled to the Rink amide resin (0.5 mmol of peptide per gram resin) by HOBt/DIC (5 equiv in dimethylformamide) activation, and the peptide chain was elongated. The Dmab-protecting group was cleaved by three treatments of the resin with hydrazine (2% in dimethylformamide) for 3 min. Cyclization was achieved by two treatments of the resin for 24 h with a solution of HOBt/DIC (3 equiv in dimethylformamide). After cleavage from the resin with TFA and precipitation from ice cold diethyl ether, the crude peptides

were purified by preparative HPLC and analyzed by reversed-phase analytical HPLC. Identity of the peptides was confirmed by MALDI-TOF mass spectrometry (Ultraflex III; Bruker Daltonik) or by electrospray ionization (ESI) mass spectrometry (Bruker Esquire HCT, peptides T1-3). Peptide RXR-1 was also obtained as a TFA-free preparation (ThermoFisher Scientific; Ulm, Germany), which was used for the CD measurements in the presence of TFE.

For methylation assays, lyophilized peptides were dissolved in double distilled water. Concentrations (typically ≈ 2 mM) were determined spectrophotometrically using the theoretical absorption coefficients. 280 nm was used for peptides containing tyrosines or tryptophans, 255 nm for the RGG-peptides. Concentrations of peptides lacking aromatic residues were determined by the absorption of the peptide bond (205, 210, 215, and 220 nm).

In Vitro Methylation Assays. All PRMT1 assays were done at pH 8.0 and 30 °C as described.[10] Steady state kinetic analysis with *S*-adenosyl-L-[*methyl*-^{14}C]methionine (GE Healthcare) as the methyl donor has been described.[10,22] Bovine serum albumin was omitted from the reaction buffer when protein substrates were to be analyzed by mass spectrometry. Concentrations of protein substrates were varied up to 2 μM and concentrations of peptides up to 100 μM, with actual concentrations for each particular substrate chosen on the basis of a preliminary first titration. Enzyme concentrations were at least 50-fold lower than substrate concentrations. At each substrate concentration, incorporation was determined at three to seven time points to ensure linearity of the progress curve.

Mass Spectrometry. For mass spectrometric measurements, peptides (final concentration 40 μM) were incubated with 0.2 μM PRMT1 and 100 μM AdoMet. After different time intervals, the reaction was stopped by addition of the inhibitor sinefungin (Sigma; final concentration 0.6 mM) and freezing. Samples were thawed, purified using C18 ZipTips (Millipore, Eschborn, Germany), and analyzed with an Ultraflex III MALDI-TOF/TOF mass spectrometer (Bruker Daltonik, Bremen, Germany). 0.5 μL of each sample was spotted, with the dried droplet method, onto an MTP 384 stainless steel target plate (Bruker Daltonik) with 0.5 μL matrix solution (50% (v/v) ACN (Merck), 0.1% (v/v) TFA (Sigma) saturated with α-cyanohydroxycinnamic acid (Bruker Daltonik)). MALDI-TOF mass spectra were recorded in the positive ionization and reflectron mode in the *m*/*z* range 800−5000. External calibration of the mass spectra was performed with the Peptide Calibration Standard II (Bruker Daltonik). Signals corresponding to methylated peptide species were selected for tandem MS analysis. FlexControl 3.0 software (Bruker Daltonik) was used for MS and MS/MS data acquisition. Data were processed with the FlexAnalysis 3.0 software (Bruker Daltonik). All arginine residues in a peptide were considered as potential methylation or dimethylation sites. In order to identify the methylation sites, fragment ions (a, b- and y-type ions) including all theoretical methylation patterns were calculated, compared to the MS/MS spectra and scored using the Biotools 3.1 software package (Bruker Daltonik). For peak assignment in MS/MS, a maximum mass deviation of 500 ppm was allowed.

Alternatively, samples were analyzed by ESI-MS using an LTQ-OrbitrapXL mass spectrometer (ThermoFisher Scientific, Bremen) equipped with a nano-ESI source using metal coated borosilicate emitters (Proxeon, Odense, Denmark). Mass spectra of peptides were acquired *offline* in the positive ionization mode in the *m*/*z* range 200−2000. The resolving power was 30000 at *m*/*z* 400. Fragments created by either collision-induced dissociation (CID, for RGG-peptides) or higher energy CID (for RXR-peptides) were detected in the orbitrap mass analyzer.

For mass spectrometric analysis of PABPN1 methylation, 3 μM protein was incubated with 0.1 μM PRMT1 and 1 mM AdoMet. Samples were processed as above and digested overnight with chymotrypsin (Roche Diagnostics) at 37 °C at an enzyme/substrate ratio of 1:80 (w/w). The digested samples were purified using ZipTips C18 (Millipore) and MALDI-TOF mass spectra were acquired as described above. Data were analyzed manually by comparison to predicted chymotryptic fragments and methylated derivatives.

Alternatively, the samples were analyzed by off-line nano-HPLC/MALDI-TOF/TOF-MS consisting of a nano-HPLC (Ultimate 3000, Dionex, Idstein, Germany) and an Ultraflex III MALDI-TOF/TOF mass spectrometer. 15 μL of the peptide mixture was loaded onto a trapping column (Acclaim PepMap100 C18, 5 μm, 100 Å, 300 μm I.D. × 5 mm, Dionex) and washed for 15 min with 0.1% TFA at a flow rate of 30 μL min^{-1}. Trapped peptides were separated over the separation column (Acclaim PepMap100 C18, 3 μm, 100 Å, 75 μm I.D. × 150 mm, Dionex), which had been equilibrated with 95% (v/v) solvent A (5% (v/v) ACN/95% (v/v) H_2O containing 0.05% (v/v) TFA), 5% (v/v) solvent B (80% (v/v) ACN, 0.04% (v/v) TFA). Peptides were eluted with a 30-min gradient from 5 to 50% (v/v) solvent B, followed by a 2-min gradient to 100% (v/v) B and 100% B (v/v) for 3 min. Separations were conducted at a flow rate of 300 nL/min and a temperature of 45 °C. A fraction collector Proteineer fc (Bruker Daltonik) was used to spot 25-s nano-HPLC fractions onto a 384 MTP 800 μm AnchorChip target and mix them with 1.1 μL matrix solution (0.71 mg/mL α-cyano-4-hydroxy cinnamic acid in 90% (v/v) ACN, 0.1% (v/v) TFA, 1 mM $NH_4H_2PO_4$) per spot. Mass spectra in the *m*/*z* range 740−5000 were acquired in the positive ionization and reflectron mode by accumulating data from 1800 laser shots per spot. Ion signals with a signal/noise ratio higher than 9 were automatically subjected to MALDI-LIFT-TOF/TOF-MS/MS. For determining methylation sites, mass spectra were searched against the amino acid sequence of His-tagged bovine PABPN1 using Biotools 3.1 and Mascot-server 2.2.

Far-UV CD. Lyophilized TFA-free peptide RXR-1 was dissolved to a final concentration of 0.87 mM in water with up to 80% 2,2,2-trifluoroethanol. 15−20 single scans were accumulated with a scan rate of 20 nm/min, 4 s response time, and a bandwidth of 1 nm at 4 °C. The optical path length was 0.5 mm. All other peptides were diluted into 100 mM Na_2HPO_4/NaH_2PO_4, pH 7.5, to final concentrations between 0.1 mM and 0.5 mM. Thirty single far-UV CD spectra were accumulated from 260 to 185 nm with a scan rate of 50 nm/min and 1 s response time, all other parameters being as above.

Structural Modeling. Available crystal structures of rat PRMT1 (PDB code: 1OR8, 1ORI, and 1ORH) are inadequate for use in modeling studies, as the N-terminal α-helix (residues 1−40), which constitutes a part of the substrate binding pocket, is missing. Complete forms of rat PRMT3 (50% sequence identity with PRMT1) and human PRMT4 (= CARM1; 35% sequence identity) were used to generate an accurate and complete homology model of the dimeric form of human PRMT1. Structures of rat PRMT1 (PDB code: 1OR8), rat PRMT3 (PDB code: 1F3L), and human PRMT4 (PDB code:

2Y1W) were obtained from the Protein Data Bank, and structural alignment with the protein sequence of human PRMT1 was performed using MOE 2011.10 (Chemical Computing Group, Montreal, Canada, 2011). A homology model was subsequently generated using the PRMT4 structure as a template to add the missing N-terminal α-helix. The dimeric structure of PRMT1 was developed by superimposition with the dimeric form of PRMT4 (chains A and C, 2Y1W) in MOE. The X-ray structure of rat PRMT1 reveals the substrate arginine situated in the deep cavity of the binding pocket, the guanidino group forming key H-bond interactions with the backbone carbonyl oxygen of E153 and the side chain of E144. Consequently, the substrate arginine could easily be placed into the binding pocket of the structural model of human PRMT1. A shell script was developed to generate spheres of different diameters which could be hosted by the substrate binding cavity of PRMT1.

RESULTS

In PABPN1, methylated arginines are restricted to the C-terminal domain. Although peptides of ≈20 amino acids derived from it are good substrates for PRMT1, they are not quite as good as the entire 49 amino acid domain in the context of the full-length protein.[22] The entire C-terminal domain could not be tested by itself in methylation assays due to its poor solubility. Therefore, it was fused to the cold shock protein from *B. subtilis*, Csp B, a small and highly soluble protein,[43] which, as a prokaryotic protein, is not a natural PRMT substrate. The chimeric protein, Csp-RXR, was expressed in *E. coli*, purified to homogeneity, and used for steady state kinetic methylation experiments. Its modification by PRMT1 was indistinguishable from PABPN1 methylation (Figure 1, Table 1). CspB without the PABPN1 appendix was

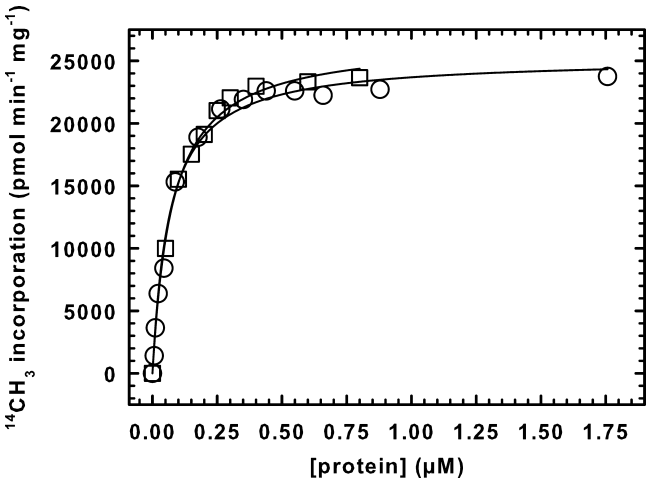


Figure 1. PRMT1 recognizes PABPN1 exclusively by its C-terminal domain. Steady state kinetic measurements of PRMT1-mediated methylation of wt-PABPN1 (○) and Csp-RXR (□) were performed as described under Experimental Procedures. Data for the wild-type protein are those published previously.[22]

not methylated to any detectable extent. Thus, PRMT1 recognizes exclusively the C-terminal domain of PABPN1, and contacts with the remainder of the protein are not necessary for substrate recognition.

Preferred methylation sites within the C-terminus of PABPN1 were first identified by peptide mass fingerprint analysis using MALDI-TOF-MS. PABPN1 was incubated with PRMT1 and AdoMet, and samples were taken at different time points and digested with chymotrypsin. Precursor ion mass spectra were searched for both unmodified and methylated chymotryptic peptides. A few additional cleavages were also observed. Proteolytic peptides covering nearly the whole length of the C-terminal domain of PABPN1 were identified, including all arginine residues (Figure 2). Substantial amounts of methylation were found only in two fragments, both covering the region N285/S286–Y293 containing arginine residues 287, 289, and 291 (Figure 2A,C). In order to obtain more detailed information, chymotryptic peptides were separated by reversed phase HPLC and analyzed by MALDI-MS/MS. The analysis of a sample taken after 10 min reaction time identified specifically Arg289 to be methylated. Both monomethylated and dimethylated Arg289 was found. Traces of methylated Arg251, Arg259, and Arg267 were detected in samples taken after longer incubation times (90 and 120 min). In the precursor ion spectra, however, only negligible amounts of these products were detected compared to the unmethylated counterpart (Figure 2B). Thus, in vitro methylation of PABPN1 by PRMT1 strongly prefers arginine 289.

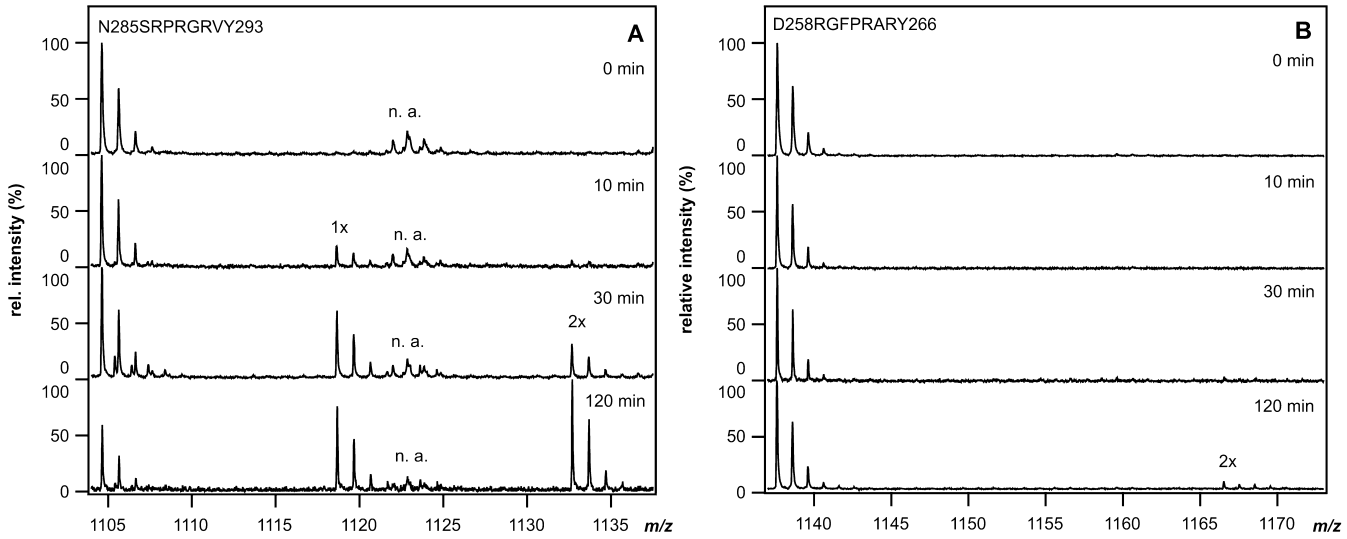
We have previously described the peptide RXR-1, which corresponds to amino acids 280–303 of PABPN1 with an internal deletion of amino acids 294–298 thus covering the critical region containing Arg289[10,22] (Table 2). In agreement with the preference for Arg289 in full-length PABPN1, RXR-1 is exclusively methylated at the corresponding position in the peptide, Arg10 [10] (Figure S1). When this residue was substituted either by lysine (peptide RXR-9) or by alanine (peptide RXR-10), the peptides were methylated by PRMT1 with a 100-fold lower efficiency (Figure 3 and Table 2). For comparison, the neighboring arginines, Arg12 and -8 (corresponding to Arg291 and Arg287 in full-length PABPN1) were individually substituted by either lysine or alanine. The alanine substitution at Arg8 (peptide RXR-11a) and the lysine mutation at Arg12 (peptide RXR-12b) were tolerated by PRMT1, whereas the other two mutations (R8K, peptide RXR-11b; and R12A, peptide RXR-12a) decreased the methylation efficiency (Table 2). Thus, while not every amino acid is tolerated, arginine is not essential at these positions. These results are in agreement with those obtained with a

Table 1. Kinetic Parameters of the PRMT1-Catalyzed Methylation Reaction of PABPN1 Variants[a]

protein	V_{max} (pmol min^{-1} mg^{-1})	k_{cat} (10^{-2} s^{-1})	K_M (μM)	k_{cat}/K_M (s^{-1} M^{-1})
wt-PABPN1[b]	25200	1.8	0.07	270000
Csp-RXR	26700	1.9	0.08	250000
PABPN1-R259A	25800	1.8	0.15	110000
PABPN1-R294A	24400	1.7	0.26	64000
PABPN1-R287A	12900	0.90	0.12	78000
PABPN1-R289A	11700	0.82	0.29	29000
PABPN1-P288A	17500	1.2	0.44	28000

[a]Point mutations were introduced into the full-length bovine protein, and the numbering refers to its predicted sequence,[33] including the N-terminal methionine but not the His-tag. [b]Data taken from ref 22. The standard errors of V_{max} and K_M did not exceed 2% and 9%, respectively.

dx.doi.org/10.1021/bi300373b | *Biochemistry* 2012, 51, 5463–5475



C

```
                  247 250          260          270          280          290          300
FRGR QIKVIPKRTN RPGISTTDRG FPRARYRART TNYNSSRSRF YSGFNSRPRG RVYRGRARAT SWYSPY      mma   dma

 RGR QIKVIPKR                                                                      -     -
 RGR QIKVIPKRT                                                                     -     -
 RGR QIKVIPKRTN                                                                    -     -
 RGR QIKVIPKRTN RPGISTT                                                            -     -
       VIPKRTN RPGISTT                                                             -     -
             N RPGISTTDRG FPRARY                                                   -     -
               RPGISTTDRG FPR                                                      2[a]  -
                    TTDRG FPRARY                                                   -     -
                     TDRG FPRARY                                                   -     -
                      DRG FPRARY                                                   1[a]  1[a]
                       RG FPRARY                                                   -     -
                                RART TNY                                           -     -
                                RART TNYNSSRSRF                                    -     -
                                     NSSRSRF                                       -     -
                                     NSSRSRF Y                                     -     -
                                                NSRPRG RVY                         1     1
                                                 SRPRG RVY                         -     1
                                                          RGRARAT SW               -     -
```

Figure 2. In vitro arginine methylation of PABPN1 by PRMT1 is largely restricted to arginine 289. PABPN1 was methylated by PRMT1 for 0, 10, 30, 90, and 120 min (spectra of samples taken after 90 min have been omitted), digested with chymotrypsin, and analyzed by MALDI-TOF-MS. (A) Mass spectra of the peptide N^{285}SRPRGRVY^{293}. (B) Mass spectra of the peptide D^{258}RGFPRARY^{266}. In both figures, the number of methylations (1x, 2x) is indicated. n.a., not assigned. (C) Chymotryptic peptides were analyzed by nano-HPLC/MALDI-TOF/TOF-MS/MS as described in Experimental Procedures. The sequence of the C-terminal domain of PABPN1 is shown (upper lane) in comparison to all identified peptides representing this region. Identified methylated residues are underlined. The amino acid numbering refers to the sequence of the bovine protein,[33] including the N-terminal methionine but not the His-tag of the recombinant protein. K247 marks the start of the C-terminal domain. Numbers on the right show the maximum number of monomethyl (mma)- and dimethylarginine (dma) residues found in the respective peptides. [a]: methylation was detected only in samples taken after at least 30 min and was present in negligible amounts.

similar series of RXR1 peptide derivatives, in which arginine residues were substituted by dimethylarginine: These substitutions had little impact on PRMT1 activity as long as Arg10 was not affected.[10] In summary, Arg10 in the peptide RXR-1, corresponding to Arg289 in full-length PABPN1, is the preferred site of methylation.

Arg10 is preceded by a proline residue. For simplicity, we will refer to the position at the N-terminal side of the methylated arginine as −1 and to the position on the C-terminal side as +1. In order to examine the importance of the proline residue at −1, several substitution variants were tested. When Pro9 was substituted by glycine (peptide RXR-13), no effect on the methylation efficiency was seen (Figure 3 and Table 2). Unexpectedly, however, the double mutant P9G R10K (peptide RXR-14) remained a good substrate for PRMT1 (Figure 3 and Table 2), whereas the R10K substitution by itself strongly reduced the ability of the peptide to be methylated (peptide RXR-9, see above). In other words, the substrate activity lost due to the R10K substitution was restored upon the additional substitution of the proline at position −1. As a control, Pro9 was also substituted by alanine (peptide RXR-15). The single exchange impaired methylation by PRMT1. Nevertheless, the combination of the P9A mutation with the R10K mutation (peptide RXR-16) generated a 10-fold better substrate compared with the R10K mutation alone (peptide RXR-9) (Table 2).

These data strongly indicate that the exclusive methylation of Arg10 in the 'wild-type' RXR-1 peptide depends on Pro9 and predict that the substitution of Pro9 should promote methylation of the two other arginines. Therefore, methylation

dx.doi.org/10.1021/bi300373b | *Biochemistry* 2012, 51, 5463−5475

Table 2. Kinetic Parameters of the PRMT1-Catalyzed Methylation of PABPN1-Derived Peptides[a]

peptide	sequence	V_{max} (pmol min^{-1} mg^{-1})	k_{cat} (10^{-2} s^{-1})	K_M (μM)	k_{cat}/K_M (s^{-1} M^{-1})
RXR-1[b]	FYSGFNSRPRGRVYATSWY	36000	2.5	1.3	20,000
RXR-9	FYSGFNSRP**K**GRVYATSWY	2830	0.18	13	140
RXR-10	FYSGFNSRP**A**GRVYATSWY	2380	0.16	11	150
RXR-11a	FYSGFNS**A**PRGRVYATSWY	23800	1.7	2.3	7400
RXR-11b	FYSGFNS**K**PRGRVYATSWY	1200	0.08	4.8	170
RXR-12a	FYSGFNSRPRG**A**VYATSVY	7500	0.53	9.8	540
RXR-12b	FYSGFNSRPRG**K**VYATSVY	18600	1.3	1.1	12000
RXR-13	FYSGFNSR**G**RGRVYATSWY	30,700	2.2	1.1	19000
RXR-14	FYSGFNSR**GK**GRVYATSWY	27000	1.9	1.6	12000
RXR-15	FYSGFNSR**A**RGRVYATSWY	9130	0.64	1.8	3600
RXR-16	FYSGFNSR**AK**GRVYATSWY	2140	0.15	0.70	2200

[a]Peptide RXR-1 corresponds to amino acids 280−303 of PABPN1 with an internal deletion of amino acids 294−298. In peptide RXR-1, Pro9 and Arg10, corresponding to Pro288 and Arg289 in the PABPN1 sequence, are underlined. Bold letters indicate amino acid substitutions. The standard errors of V_{max} and K_M typically did not exceed 4% and 14%, respectively. [b]Data taken from ref 22.

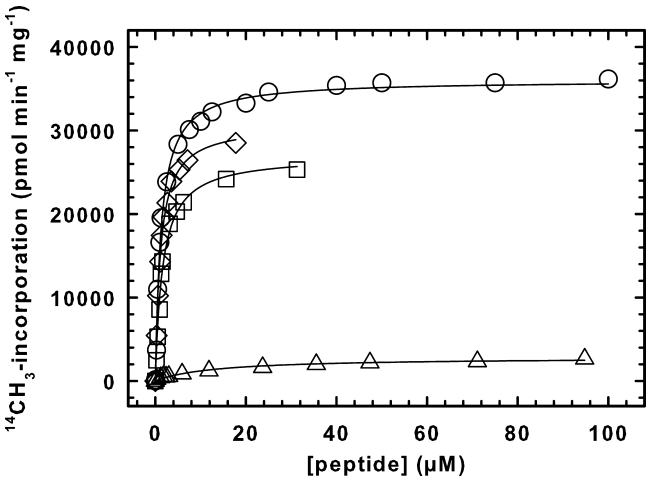


Figure 3. Methylation of RXR-peptides is determined by Pro9. v/S characteristics of PRMT1 with peptides RXR-9 (△), RXR-13 (◇), and RXR-14 (□) were determined as described under Experimental Procedures. Amino acid sequences and kinetic constants are listed in Table 2. The data of RXR-1 (○) have been published[22] and are included for comparison.

sites in peptide RXR-13 carrying the P9G mutation were analyzed by mass spectrometry. Products with one or two methyl groups were readily observed (Figure S2), and a triply methylated product, which was never found for RXR-1, was also present. This already implies that in RXR-13, in contrast to RXR-1, at least two arginine residues were methylated. Upon MALDI-MS/MS analysis of peptides, few methylated fragment ions were unambiguously identified, indicating a heterogeneous mixture. Distinct monomethylations at both Arg8 and Arg12 were seen (Figure 4, compare to Figure S1). Methylations at these residues were also found with RXR-15 (P9A variant) as a substrate (data not shown).

Analysis of peptides RXR-1, -13, and -15 was repeated using ESI-MS/MS. Analysis of RXR-1 confirmed exclusive methylation at Arg10 (data not shown). The fragment ion mass spectrum of the monomethylated precursor ion of RXR-13 revealed methylation at Arg8 and Arg12 (Figure 5A and Figure S3). In RXR-15, Arg10 and Arg12 were found to be methylated (Figure 5B and Figure S4). Thus, MS analysis confirms that the exclusive methylation of Arg10 in the wild-type sequence is abolished by substitution of Pro9.

In order to determine whether the proline effect is specific to the RXR-1 sequence or can be generalized, we introduced substitutions into the RGG-1 peptide, a generic PRMT1 substrate originally lacking any proline [22] (Table 3). As previous MALDI-MS/MS analysis had indicated that PRMT1 methylates RGG-1 at Arg3 and Arg15,[10] substitutions were generated in the −1 position relative to Arg9. The RGG-1 peptide is methylated by PRMT1 with modest efficiency.[22] When the glycine at position −1 was substituted by proline, the resulting peptide, RGG-2, was methylated with a similar k_{cat}, but significantly higher apparent affinity, resulting in a 10-fold improved catalytic efficiency (Table 3). RGG-3, as a control, contained an alanine substitution at −1. In this case, the K_M for methylation by PRMT1 was also decreased, but less so than with the G8P substitution (Table 3). In-depth ESI-MS/MS analysis (including MS^3 experiments) showed that even the original RGG-1 peptide was methylated at all three arginine residues (data not shown). Introduction of proline (Figure S5) or alanine (data not shown) at position 8 did not induce changes in the methylation pattern that would have been detectable by mass spectrometry. Thus, the proline effect in the RGG context is not nearly as pronounced as in the RXR sequence.

In order to determine whether the behavior of the various RXR peptides was representative of full-length PABPN1, we introduced corresponding mutations into the protein. Substitution of the preferred methylation acceptor, Arg289, by alanine reduced the efficiency of methylation 10-fold (Table 1). For comparison, three additional arginines were individually substituted by alanine: The substitution R259A had a weak influence, whereas R287A and R294A each had a 4-fold effect on the catalytic efficiency of PRMT1 (Table 1). In summary, among the arginine substitutions tested, that of Arg289 had the strongest effect on the methylation efficiency, in agreement with it being the preferred methylation site.

Pro288 was also substituted by alanine in the full-length protein. Methylation efficiency of the variant PABPN1-P288A by PRMT1 was reduced approximately 10-fold, similar to the R289A variant (Table 1). Preferred methylation sites of the P288A variant were analyzed by MALDI-MS/MS as described above for the wild type protein. Portions of the precursor ion mass spectra obtained after methylation for 0 and 120 min are shown in Figure 6. Whereas significant methylation of wt-PABPN1 had been found only in the region N285/S286−Y293 (see above), substantial methylation in the region R258-Y267

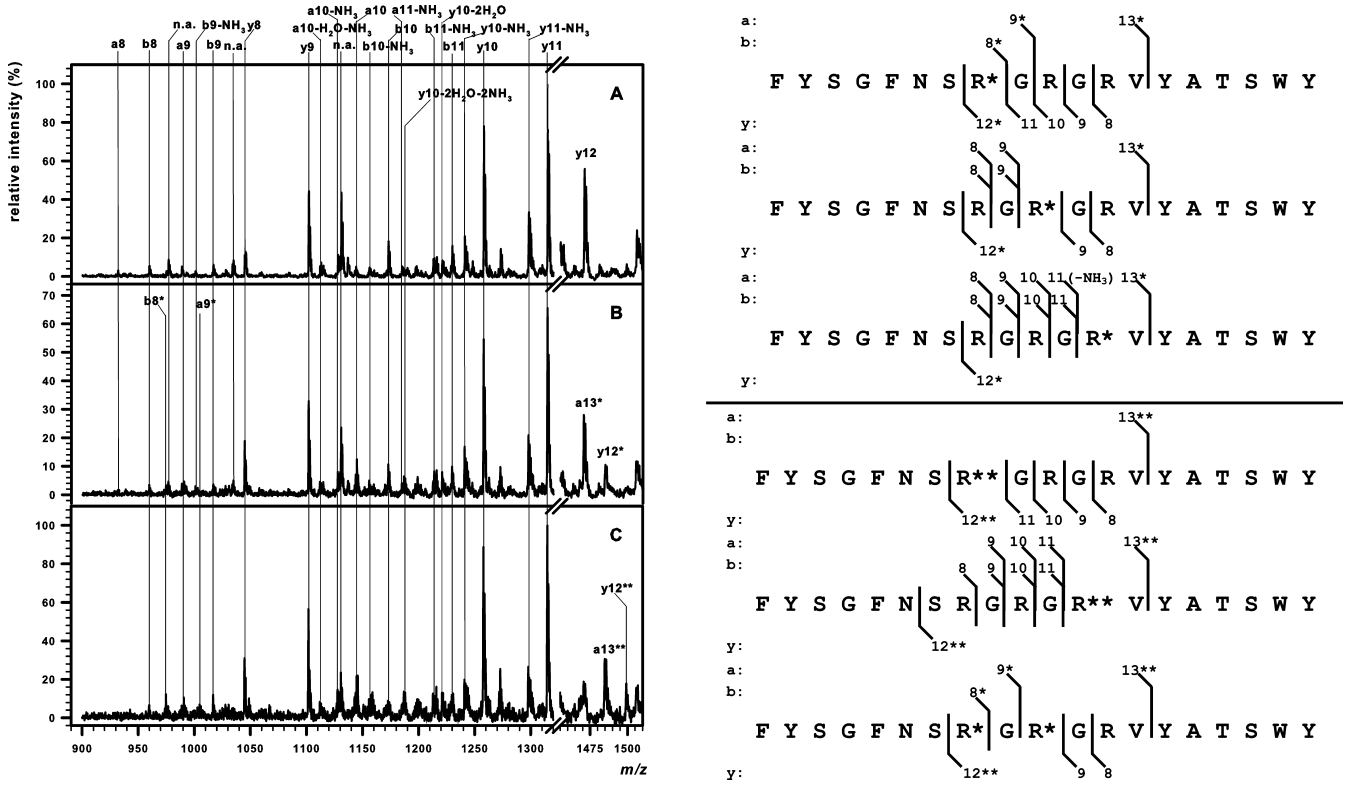


Figure 4. Methylation of peptide RXR-13 by PRMT1 analyzed by MALDI-TOF/TOF-MS/MS. Left: RXR-13 was methylated by PRMT1 for 90 min, and the sample was treated and analyzed as described under Experimental Procedures. The signals of the unmethylated (A), the mono- (B) and the dimethylated (C) precursor ions (marked with ●, ■, and *, respectively, in Supplemental Figure 2) were selected for fragmentation. Only relevant regions of fragment ion mass spectra are shown. Methylated species are marked by asterisk. Whereas the methylated b8 fragment proved Arg8 methylation, the unmethylated b11 fragment, derived from a monomethylated precursor, demonstrated methylation of Arg12 (B). Fragmentation of the dimethylated precursor (C) again produced the monomethylated b8 fragment, confirming monomethylation of Arg8. Dimethylation of Arg8 was proven by unmethylated y11, whereas the unmethylated b11 fragment indicated dimethylation of Arg12. Note that, due to similar masses of the theoretically possible fragment ions, signals at $m/z \sim 1484$ and 1498 were not unambiguously assigned and could correspond to y12 or a methylated a13 ion. These signals were therefore assigned relying on the methylation state of the respective precursor ion with the assumption that no fragmentation of the N–CH_3 bond occurs under MALDI-MS/MS conditions. Right: The amino acid sequences of the three possible variants of monomethylated (upper panel) or dimethylated (lower panel) RXR-13 are shown along with the identified fragment ions that are consistent with the respective methylation pattern. Only three of the theoretically possible variants are shown.

was easily detectable for the P288A variant (compare Figure 2 with Figure 6). Fragments Y273/N274–Y281 and R294-W302 were found to be methylated as well. The fragments from the C-terminal domain that were identified by LC/MALDI-MS/MS are summarized in Figure 6C. By MS/MS, residues 259 and 263 were shown to carry methyl groups. Clearly, the restriction of methylation to Arg289 was relieved upon the loss of Pro288. Thus, in full-length PABPN1, like in the peptides, preferred methylation of Arg289 depends on the proline residue in the −1 position.

How is arginine methylation favored by a preceding proline residue? The most likely explanation is an effect on the backbone conformation of the peptide. Proline is the most strongly favored amino acid at positions i and i+1 of various types of reverse turns.[44] We also note that the specifically methylated arginine in RXR-1 and full-length PABPN1 is followed by glycine, and this is the preferred amino acid in positions i+2 or i+3 of turns.[44] Together, proline and glycine would put the intervening arginine in either position i+1 or i+2 of the reverse turn. The arginine side chain in either position is predicted to point away from the peptide backbone forming the hairpin-like turn.[44] It is conceivable that this conformation would facilitate the insertion of the side chain into the active site pocket of PRMT1, which does not appear to be easily accessible.

Cyclization of short peptides is commonly used to stabilize them in a reverse turn conformation, and structural analysis of proline-containing cyclic peptides has confirmed the preference of this amino acid for position i+1 of the turn.[45−48] Thus, three different version of a pentapeptide, dAla-Pro-Arg-Gly-Asn, were synthesized (Table 4). Peptide T-1 contained this sequence in a linear form with free N- and C-termini. Peptide T-2 was also linear, but had an acetylated N-terminus and an amidated C-terminus. Peptide T-3, finally, was in the cyclic form with a peptide bond between dAla and Asn. According to its far-UV CD spectrum, peptide T-1 was in random coil conformation, lacking any ordered structure (Figure 7A). It was practically no PRMT1 substrate, its methylation being at the lower limit of detection (Figure 7B). The spectrum of peptide T-2 showed a slightly red-shifted minimum (Figure 7A), indicating a mixture of ordered and disordered structure. It was a better substrate with a significantly increased k_{cat} (Figure 7B, Table 4). The cyclic peptide T-3 was still a relatively weak PRMT1 substrate, but it was methylated about one hundred times faster than peptide T-1 (free termini) and fifteen times faster than peptide T-2 (blocked termini) (Figure 7B, Table 4). Although the CD spectrum of T3 lacks the characteristic maximum at ca. 190 nm,

dx.doi.org/10.1021/bi300373b | Biochemistry 2012, 51, 5463−5475

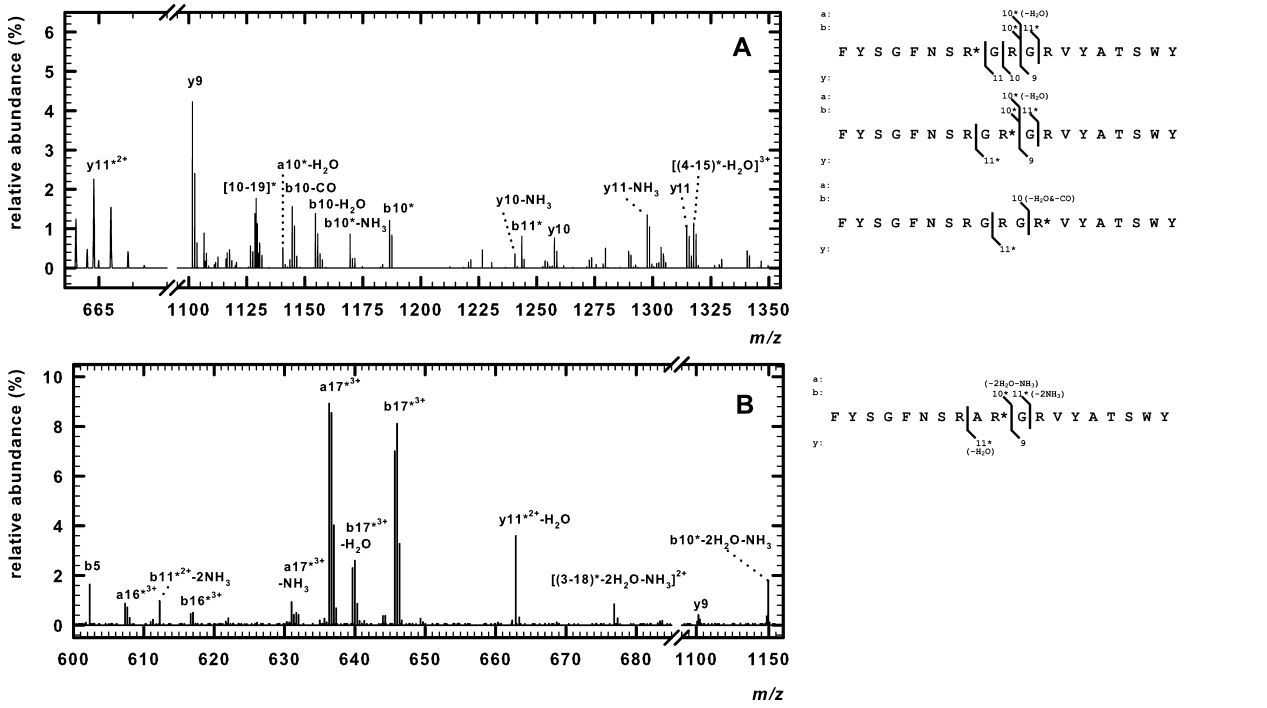


Figure 5. Methylation of peptides RXR-13 and -15 by PRMT1 analyzed by nano-ESI-LTQ-Orbitrap-MS/MS. Peptides RXR-13 (A) and RXR-15 (B) were methylated by PRMT1 for 90 min, and the sample was treated as described in Experimental Procedures. The triply charged monomethylated precursors were selected, fragmented by higher energy CID, and fragment ions were analyzed in the orbitrap. Only relevant regions of the fragment ion spectra are shown; (for the complete spectra see Figures S3 and S4). Methylation of Arg8 was proven by unmethylated fragment ions y8−11, whereas the unmethylated b10 fragment - although only detected with neutral losses of CO and water - implied methylation of Arg12 (A). The monomethylated b11 fragment of the dimethylated precursor also proves monomethylation of Arg12, assuming that the N−CH_3 bond is not cleaved under the conditions applied (B). Internal fragments are indicated by the amino acid numbers, for example, 3−5 represents the sequence SGF. On the right-hand side, the amino acid sequences of the possible variants of monomethylated RXR-13 are presented along with identified fragment ions that are consistent with the respective methylation pattern. Methylated arginine residues and fragment ions are indicated by asterisks. Only the most likely variant of monomethylated RXR-15 is shown, and identified fragments consistent with all three theoretically possible forms are omitted.

Table 3. Kinetic Parameters of the PRMT1-Catalyzed Methylation of RGG Peptides[a]

peptide	sequence	V_{max} (pmol min^{-1} mg^{-1})	k_{cat} (10^{-2} s^{-1})	K_M (μM)	k_{cat}/K_M (s^{-1} M^{-1})
RGG-1[b]	GGRGGFGGRGGFGGRGGFG	18200	1.3	3.0	4300
RGG-2	GGRGGFGPRGGFGGRGGFG	11200	0.8	0.16	49000
RGG-3	GGRGGFGARGGFGGRGGFG	14900	1.0	0.50	21000

[a]The standard errors of V_{max} and K_M typically did not exceed 4% and 17%, respectively. [b]Data taken from ref 22.

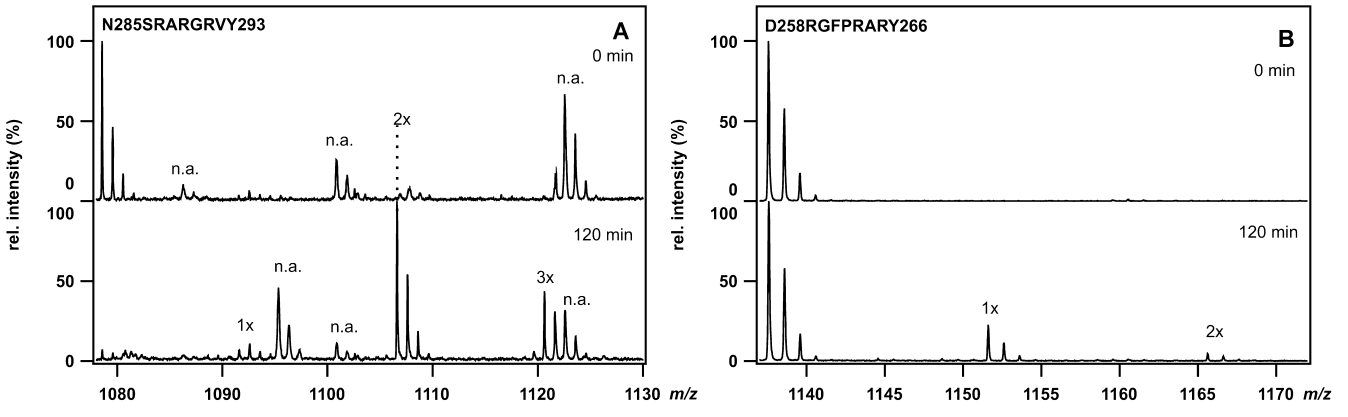
the minimum at 215 nm supports either β-strand or type I β-turn conformation, with the cyclic character of this short chain strongly favoring the latter one (Figure 7A). In this case, Arg3 should be in the i+2 position. The results are consistent with the notion that a reverse turn conformation displaying arginine in one of the apical positions may facilitate methylation by PRMT1.

The structures of several RXR peptides were also evaluated by far-UV CD spectroscopy. A low signal intensity was common to all of them, suggesting equilibria of two or more conformations. Nevertheless, the spectra obviously deviated from those expected of an unstructured peptide (Figure 8A,B, compare to that of T1, Figure 7A). The CD spectrum of RXR-1 showed a slight minimum at 210 nm, but the most prominent feature was a maximum at 225−230 nm. As the latter was not present in peptides lacking Arg8 or Tyr14 (data not shown), it probably results from a cation-π-interaction between the guanidine and aromatic moieties of these two amino acids. This peak may partially obscure the signal expected for a reverse turn. Upon addition of the secondary structure stabilizing reagent 2,2,2-trifluoroethanol, the peak at 225 nm was strongly reduced in favor of the minimum at 210 nm, which also shifted toward 215 nm (Figure 8A). Such a minimum is usually associated with β-strands or with type 1 β-turns.[49,50] This minimum was strongly reduced by substitution of Pro9 (peptide RXR13; Figure 8B), showing that Pro9 affects the conformation of the peptide and supporting the interpretation that this amino acid favors a reverse turn structure (data not shown). The spectra of RXR-9 and RXR-10 were indistinguishable from the spectrum of RXR-1 (data not shown); thus, the substitution of Arg10 does not have a major influence on the structure of the peptide.

DISCUSSION

The determinants by which protein arginine methyltransferases recognize their substrates are poorly understood. Complexes containing the type II enzyme PRMT5 bind their Sm protein substrates via an interaction between the noncatalytic subunit

dx.doi.org/10.1021/bi300373b | Biochemistry 2012, 51, 5463−5475



C

```
                247 250          260          270          280          290          300
FRGR QIKVIPKRTN RPGISTTDRG FPRARYRART TNYNSSRSRF YSGFNSRARG RVYRGRARAT SWYSPY      mma   dma

 RGR QIKVIPK                                                                       -     -
 RGR QIKVIPKR                                                                      -     -
 RGR QIKVIPKRT                                                                     -     -
 RGR QIKVIPKRTN                                                                    -     -
 RGR QIKVIPKRTN RPGISTT                                                            -     -
  GR QIKVIPKRT                                                                     -     -
     QIKVIPKRTN RPGISTT                                                            -     -
        VIPKRTN RPGISTT                                                            -     -
              N RPGISTTDRG FPRARY                                                  -     -
                  TTDRG FPRARY                                                     1     -
                    DRG FPR                                                        -     -
                    DRG FPRARY                                                     2     1
                     RG FPRARY                                                     -     -
                           RART TNY                                                -     -
                               YNSSRSRF Y                                          1     -   *
                                NSSRSRF Y                                          1     1   *
                                        SGFNSRARG RVY                              -     -
                                           NSRARG RVY                              1     2
                                            SRARG RVY                              -     2
                                                    RGRARAT SW                     1     -   *
```

Figure 6. The P288A substitution relieves the local restriction of methylation. PABPN1-P288A was methylated by PRMT1 for 0 or 120 min as indicated, digested with chymotrypsin and analyzed by MALDI-TOF-MS as described in Experimental Procedures. (A) Mass spectra of peptide N^{285}SRARGRVY^{293} with one, two, and three methyl groups. (B) Mass spectra of peptide D^{258}RGFPRARY^{266} with one or two methyl groups. n.a., not assigned. (C) Peptides derived from the C-terminal domain were analyzed by nano-HPLC/MALDI-TOF/TOF-MS/MS. They are aligned with the P288A mutant protein sequence (see legend to Figure 2). Methylated arginine residues confirmed by MS/MS experiments are underlined. Numbers on the right show the maximum number of monomethyl (mma) and dimethylarginine (dma) residues found in the respective peptides. Peptides marked with asterisks were found to be methylated in MS, but not in MS/MS.

Table 4. Kinetic Parameters of the PRMT1-Catalyzed Methylation of Linear and Cyclic Pentapeptides[a]

peptide	sequence	V_{max} (pmol min^{-1} mg^{-1})	k_{cat} (10^{-2} s^{-1})	K_M (μM)	k_{cat}/K_M (s^{-1} M^{-1})
T-1	ZPRGN	32	0.002	440	0.05
T-2	ZPRGN	241	0.017	620	0.27
T-3	ZPRGN	3574	0.25	1500	1.6

[a]Z denotes D-alanine. T-1, linear peptide with free N- and C-termini. T-2, linear peptide with acetylated N- and amidated C-terminus. T-3, cyclic peptide via peptide bond. The standard errors of V_{max} and K_M did not exceed 5% and 13%, respectively.

pICln and the Sm fold outside the methylated domain.[19,20] Likewise, arginine methylation of histone substrates does not rely exclusively on interactions between a particular PRMT and the N-terminal histone tail to be modified; instead, the modification is directed to particular sites in chromatin through an association of PRMTs with transcription factors.[2] However, recognition of local amino acid sequences does play a role in determining the substrate specificity of type I PRMTs[11,21,22] (Figure 1). Although machine learning methods have been reported to predict arginine methylation sites with some accuracy,[51−53] the rules for substrate recognition are not clear. It is also impossible to predict which member of the family of PRMTs will catalyze methylation of a particular arginine residue. In this paper we present evidence that the substrate specificity of PRMT1 is determined not only by specific contacts between the enzyme and particular amino acid side chains in the vicinity of the substrate arginines but also by the conformation of the peptide backbone in the substrate. We suggest that an arginine residue in one of the apical positions of a reverse turn is favorable for methylation.

The evidence can be summarized as follows: Out of 13 arginine residues that are modified in PABPN1 in vivo, PRMT1 selects one particular position, Arg289, that is almost exclusively methylated in vitro. The same preference is maintained in a peptide corresponding to the amino acid sequence surrounding Arg289. Both in the protein and in the peptide, selection of Arg289 depends on a proline residue as its N-terminal neighbor. This proline, together with the glycine

dx.doi.org/10.1021/bi300373b | Biochemistry 2012, 51, 5463−5475

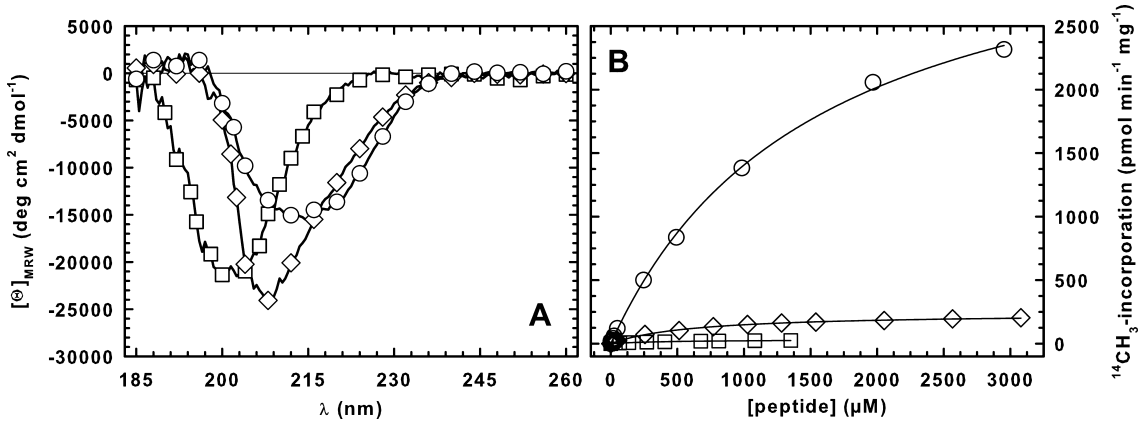


Figure 7. Cyclization of a pentapeptide facilitates its methylation by PRMT1. (A) Far-UV CD spectra of peptides T-1 (□), T-2 (◇), and T-3 (○) were recorded as described in Experimental Procedures. Intensities were corrected for the number of peptide bonds, four, four and five for the two linear and the cyclic peptide, respectively. (B) The same peptides were titrated into a methylation assay, and initial rates of methyl transfer were determined. Symbols as in (A). Kinetic constants derived from the data are shown in Table 4.

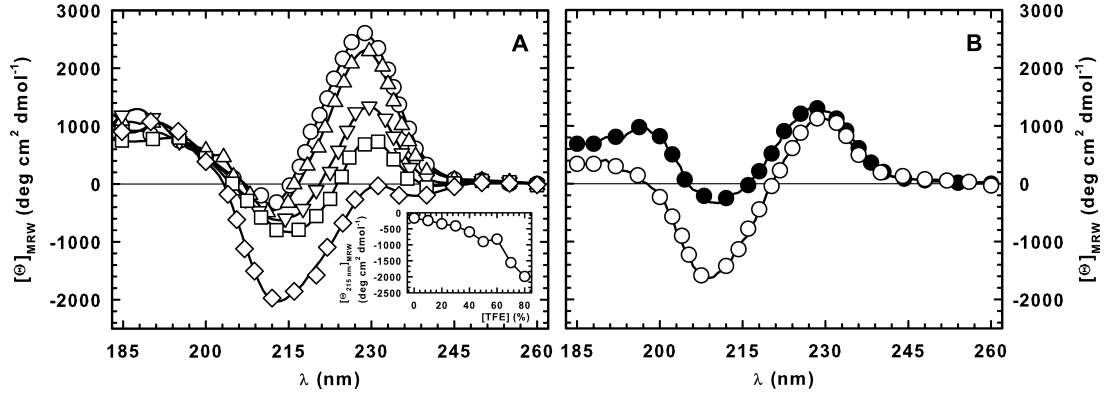


Figure 8. Secondary structure analysis of RXR family peptides. (A) Far-UV CD spectra of TFA-free RXR-1 were recorded in water containing 0% (○), 20% (△), 40% (▽), 60% (□), and 80% trifluoroethanol (◇). Inset: The signal at 215 nm is plotted versus the TFE concentration. (B) Far-UV CD spectra of RXR-1 (○) and RXR-13 (●) were measured in phosphate buffer as described under Experimental Procedures.

residue at position 290, may promote a reverse turn structure, putting the arginine in either the i+1 or the i+2 position of the turn and making its side chain point away from the tight turn of the backbone. In support of this hypothesis, cyclization of a pentapeptide containing a PRG sequence accelerated the rate of methylation 15-fold and improved catalytic efficiency 6-fold. Cyclization of such proline-containing peptides is known to favor reverse turn structures with proline in the i+1 position, thus putting the arginine residue in this particular peptide in the i+2 position.

We suggest that the proposed structure with the arginine side chain at the tip of a reverse turn facilitates access to the PRMT1 active site, which is facing the central hole of the doughnut-like homodimer. All PRMTs that have been crystallized are homodimers arranged in this manner,[12−17] so structural restraints on the range of possible substrates are probably common among all these enzymes. In all cases examined, catalytic activity of the PRMTs depends on homodimer formation.[14,16−18] This may serve to maintain their substrate specificity, restricting methylation activity to substrates that can access the central cavity of the homodimer and excluding larger folded structures. In order to test this idea, we simulated, by computation, an arginine side chain on the surface of spheres of different diameters and asked whether the arginine would be able to reach the active sites of a PRMT1 dimer. The results (Figure S7) showed that any sphere with a diameter larger than 8.9 Å prevented access to the active site, as expected from the dimensions of the central cavities of PRMT1, -3, and -4, approximately 8 × 12 Å.[17] As even the smallest protein fold (e.g., the Trp cage miniprotein with a diameter of ∼12−15 Å)[54] is larger than this, the absence of a stable tertiary fold is almost certainly a precondition for methylation by mammalian PRMT1, -3, and -4. Likely exceptions are arginine residues placed in loops emanating from folded regions, provided they are of sufficient length and dimensions. In contrast to the mammalian enzymes, the active sites of *Arabidopsis* PRMT 10 appear more easily accessible.[17]

Structural restraints on PRMT substrates due to poor accessibility of the active sites are reflected in several properties of known substrates: Many methylated arginine residues are embedded in glycine-rich sequences, as exemplified by the RGG consensus of methylation sites. In fact, although overall sequence conservation of arginine methylation sites is very low, glycine is preferred at all positions between −7 and +7 relative to a methylated arginine,[51,52] and most substitutions of glycine residues directly flanking the arginine decrease the efficiency of methylation.[11,23] The preference for glycine is likely to be due to its larger conformational freedom in comparison to other amino acids and the resulting flexibility of the peptide chain. More generally, arginine methylation appears to occur

preferentially in natively unfolded protein domains. This is known for example for the C-terminal domain of the Sm protein D_1, which is symmetrically dimethylated in RG dipeptide repeats[55] and disordered in a crystal structure[56] or for the N-terminal tails of histones, which are also disordered in crystals.[57] Lack of tertiary structure is also predicted for the methylated C-terminal domain of PABPN1, as discussed above.

Arginine methylation is frequently found in N- or C-terminal protein domains, as in the Sm proteins, histones, and PABPN1 discussed above or in nucleolin,[58,59] hnRNP A1,[25] or fibrillarin.[60] This arrangement may again facilitate access to the active site. However, some of these domains are fairly long. For example, in PABPN1 the most N-terminally located methylated arginine is 47 amino acids away from the C-terminus. Threading of this entire polypeptide chain into the central hole of the PRMT dimer for Arg259 to reach the active site does not seem very efficient. Threading is impossible in those cases where methylated arginines occur between two folded domains, for example, in hnRNP K.[61] Thus, the PRMT active site may be reached more easily by the peptide chain forming a transient loop or reverse turn. Alternatively, the PRMT dimer may open and close over the substrate sequence.

As the formation of a reverse turn depends on several features of the peptide sequence, it is not surprising that the proline effect that we describe here is context-dependent. For example, Arg263 in PABPN1 also has a proline in the −1 position, yet it is not a preferred methylation site. Likewise, the effects of introducing a proline residue into an RGG peptide were minor. Nevertheless, proline is the second or third preferred amino acid in almost all positions surrounding a methylation site,[51] presumably reflecting the disorder-promoting properties of this amino acid.[62]

The proposed turn structure (Figure S8), which we envision as a favored state in the conformational equilibrium of the C-terminal domain of PABPN1 or the RXR-1 peptide, is supported by the prediction of a folding program (http://www.scfbio-iitd.res.in/bhageerath/index.jsp); the same program predicts loss of the structure when the proline in the −1 position is substituted by glycine. It is possible that hydrophobic interactions between the tyrosine and phenylalanine residues scattered around R289 help to stabilize the turn conformation by hydrophobic interactions. These two aromatic amino acids are also hallmarks of the RGG domain.[25] While a reverse turn does not have one unique CD spectrum, the spectrum of RXR-1 is consistent with the presence of a reverse turn, and the change of the spectrum caused by substitution of P9 confirms a destabilization of this conformation. Interestingly, the methylated domain of nucleolin has previously been postulated, on the basis of spectroscopic data, to form a series of β-turns.[63]

The discovery of a reverse turn structure as an element favoring arginine methylation has been facilitated by the use of relatively short synthetic peptides, which are free to adopt preferred conformations based on local amino acid sequence. In the context of a larger protein, it is likely that contacts with other parts of the protein will affect the conformational equilibrium even of natively unfolded regions. This may be one of the reasons why the properties of isolated peptides as substrates for arginine methyl transferases do not always correspond to those of the native protein.[10,11,22−24]

■ ASSOCIATED CONTENT

Supporting Information

Additional mass spectrometric data and structural models of the PRMT1 active site accessibility are available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*(A.S.) Phone: (+49)-345-5525170. Fax: (+49)-345-5527026. E-mail: andrea.sinz@pharmazie.uni-halle.de. (E.W.) Phone: (+49)-345-5524920. Fax: (+49)-345-5527014. E-mail: ewahle@biochemtech.uni-halle.de.

Present Addresses

[#]Institute of Pharmacy, Martin Luther University Halle-Wittenberg, Wolfgang-Langenbeck-Strasse 4, 06120 Halle, Germany.

[⊥]Department of Chemistry, Institute of Biochemistry, University of Cologne, Zülpicher Strasse 47, 50674 Cologne, Germany.

Funding

This work was supported by grants from the EU (LSHM-CT-2005-018675) and the DFG (SFB 610) to E.W.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful to Ralph Golbik for helpful discussions, advice regarding CD spectroscopy, and critical reading of the manuscript; to Milton Stubbs for pointing out the Trp cage miniprotein; to Wolfgang Sippl for help with computations; to Kathrin Brunk for generating some of the PABPN1 variants; to Gudrun Scholz for help with protein purification; and to Franz-Xaver Schmid and Mirko Sackewitz for the CspB clone.

■ ABBREVIATIONS USED

AdoMet, *S*-adenosyl-L-methionine; DIC, diisopropylcarbodiimide; Dmab, 4-{*N*-[1-(4,4-dimethyl-2,6-dioxocyclohexylidene)-3-methylbutyl]-amino} benzyl; Fmoc/tBu, fluorenylmethoxycarbonyl/tertbutyl; Fmoc-Asp(OH)-ODmab, *N*-alpha-fluorenylmethoxycarbonyl-L-aspartic acid alpha-4-{*N*-[1-(4,4-dimethyl-2,6-dioxocyclohexylidene)-3-methylbutyl]-amino} benzyl ester); HOBt, 1-hydroxybenzotriazole; PABPN1, poly(A) binding protein nuclear 1; PRMT, protein arginine methyltransferase

■ REFERENCES

(1) Bachand, F. (2007) Protein arginine methyltransferases: from unicellular eukaryotes to humans. *Euk. Cell 6*, 889−898.

(2) Bedford, M. T., and Richard, S. (2005) Arginine methylation: An emerging regulator of protein function. *Mol. Cell 18*, 263−272.

(3) Krause, C. D., Yang, Z.-H., Kim, Y.-S., Lee, J.-H., Cook, J. R., and Pestka, S. (2007) Protein arginine methyltransferases: Evolution and assessment of their pharmacological and therapeutic potential. *Pharmacol. Therapeut. 113*, 50−87.

(4) Pahlich, S., Zakaryan, R. P., and Gehring, H. (2006) Protein arginine methylation: Cellular functions and methods of analysis. *Biochim. Biophys. Acta 1764*, 1890−1903.

(5) Bedford, M. T., and Clarke, S. G. (2009) Protein arginine methylation in mammals: Who, what, and why. *Mol. Cell 33*, 1−13.

(6) Lakowski, T. M., and Frankel, A. (2009) Kinetic analysis of human protein arginine N-methyltransferase 2: formation of monomethyl- and asymmetric dimethyl-arginine residues on histone H4. *Biochem. J. 421*, 253−261.

(7) Di Lorenzo, A., and Bedford, M. T. (2011) Histone arginine methylation. *FEBS Lett. 585*, 2024−2031.
(8) Zurita-Lopez, C. I., Sandber, T., Kelly, R., and Clarke, S. (2012) Human protein arginine methyltransferase 7 (PRMT7) is a type III enzyme forming omega-N^G-monomethylated arginine residues. *J. Biol. Chem. 287*, 7859−7870.
(9) Lakowski, T. M., and Frankel, A. (2008) A kinetic study of human protein arginine N-methyltransferase 6 reveals a distributive mechanism. *J. Biol. Chem. 283*, 10015−10025.
(10) Kölbel, K., Ihling, C., Bellmann-Sickert, K., Neundorf, I., Beck-Sickinger, A. G., Sinz, A., Kühn, U., and Wahle, E. (2009) Type I arginine methyltransferases PRMT1 and PRMT3 act distributively. *J. Biol. Chem. 284*, 8274−8282.
(11) Wooderchak, W. L., Zang, T., Zho, Z. S., Acuna, M., Tahara, S. M., and Hevel, J. M. (2008) Substrate profiling of PRMT1 reveals amino acid sequences that extend beyond the 'RGG' paradigm. *Biochemistry 47*, 9456−9466.
(12) Troffer-Charlier, N., Cura, V., Hassenboehler, P., Moras, D., and Cavarelli, J. (2007) Functional insights from structures of coactivator-associated arginine methyltransferase 1 domains. *EMBO J. 26*, 4391−4401.
(13) Yue, W. W., Hassler, M., Roe, S. M., Thompson-Vale, V., and Pearl, L. H. (2007) Insights into histone code syntax from structural and biochemical studies of CARM1 methyltransferase. *EMBO J. 26*, 4402−4412.
(14) Zhang, X., and Cheng, X. (2003) Structure of the predominant protein arginine methyltransferase PRMT1 and analysis of its binding to substrate peptides. *Structure 11*, 509−520.
(15) Zhang, X., Zhou, L., and Cheng, X. (2000) Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. *EMBO J. 19*, 3509−3519.
(16) Weiss, V. H., McBride, A. E., Soriano, M. A., Filman, D. J., Silver, P. A., and Hogle, J. M. (2000) The structure and oligomerization of the yeast arginine methyltransferase, Hmt1. *Nat. Struct. Biol. 7*, 1165−1171.
(17) Cheng, Y., Frazier, M., Lu, F., Cao, X., and Redinbo, M. R. (2011) Crystal structure of the plant epigenetic protein arginine methyltransferase 10. *J. Mol. Biol. 414*, 106−122.
(18) Lee, D. Y., Ianculescu, I., Purcell, D., Zhang, X., Cheng, X., and Stallcup, M. R. (2007) Surface-scanning mutational analysis of protein arginine methyltransferase 1: roles of specific amino acids in methyltransferase substrate specificity, oligomerization, and coactivator function. *Mol. Endocrinol. 21*, 1381−1393.
(19) Meister, G., Eggert, C., and Fischer, U. (2002) SMN-mediated assembly of RNPs: a complex story. *Trends Cell Biol. 12*, 472−478.
(20) Friesen, W. J., Paushkin, S., Wyce, A., Massenet, S., Pesiridis, G. S., van Duyne, G., Rappsilber, J., Mann, M., and Dreyfuss, G. (2001) The methylosome, a 20S complex containing JBP1 and pICln, produces dimethylarginine-modified Sm proteins. *Mol. Cell. Biol. 21*, 8289−8300.
(21) Osborne, T. C., Obianyo, O., Zhang, X., Cheng, X., and Thompson, P. R. (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.
(22) Fronz, K., Otto, S., Kölbel, K., Kühn, U., Friedrich, H., Schierhorn, A., Beck-Sickinger, A. G., Ostareck-Lederer, A., and Wahle, E. (2008) Promiscuous modification of the nuclear poly(A) binding protein by multiple protein arginin methyl transferases does not affect the aggregation behavior. *J. Biol. Chem. 283*, 20408−20420.
(23) Rawal, N., Rajpurohit, R., Lischwe, M. A., Williams, K. R., Paik, W. K., and Kim, S. (1995) Structural specificity of substrate for S-adenosylmethionine:protein arginine N-methyltransferases. *Biochim. Biophys. Acta 1248*, 11−18.
(24) Pahlich, S., Bschir, K., Chiavi, C., Belyanskaya, L., and Gehring, H. (2005) Different methylation characteristics of protein arginine methyltransferase 1 and 3 toward the Ewing sarcoma protein and a peptide. *Proteins 61*, 164−175.
(25) Kim, S., Merrill, B. M., Rajpurohit, R., Kumar, A., Stone, K. L., Papov, V. V., Schneiders, J. M., Szer, W., Wilson, S. H., Paik, W. K., and Williams, K. R. (1997) Identification of NG-methylarginine residues in human heterogeneous RNP protein A1: Phe/Gly-Gly-Gly-Arg-Gly-Gly-Gly/Phe is a preferred recognition motif. *Biochemistry 36*, 5185−5192.
(26) Smith, J. J., Rücknagel, K. P., Schierhorn, A., Tang, J., Nemeth, A., Linder, M., Herschman, H. R., and Wahle, E. (1999) Unusual sites of arginine methylation in poly(A)-binding protein II and in vitro methylation by protein arginine methyltransferases PRMT1 and PRMT3. *J. Biol. Chem. 274*, 13229−13234.
(27) Schurter, B. T., Koh, S. S., Chen, D., Bunick, G. J., Harp, J. M., Hanson, B. L., Henschen-Edman, A., Mackay, D. R., Stallcup, M. R., and Aswad, D. W. (2001) Methylation of histone H3 by coactivator-associated arginine methyltransferase 1. *Biochemistry 40*, 5747−5756.
(28) Wang, H., Huang, Z.-Q., Xia, L., Feng, Q., Erdjument-Bromage, H., Strahl, B. D., Briggs, S. D., Allis, C. D., Wong, J., Tempst, P., and Zhang, Y. (2001) Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. *Science 293*, 853−857.
(29) Guccione, E., Bassi, C., Casadio, F., Martinato, F., Cesaroni, M., Schuchlautz, H., Lüscher, B., and Amati, B. (2007) Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive. *Nature 449*, 933−937.
(30) Hyllus, D., Stein, C., Schnabel, K., Schiltz, E., Imhof, A., Dou, Y., Hsieh, J., and Bauer, U.-M. (2007) PRMT6-mediated methylation of R2 in histone H3 antagonizes H3 K4 trimethylation. *Genes Dev. 21*, 3369−3380.
(31) Bienroth, S., Keller, W., and Wahle, E. (1993) Assembly of a processive messenger RNA polyadenylation complex. *EMBO J. 12*, 585−594.
(32) Kerwitz, Y., Kühn, U., Lilie, H., Knoth, A., Scheuermann, T., Friedrich, H., Schwarz, E., and Wahle, E. (2003) Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. *EMBO J. 22*, 3705−3714.
(33) Nemeth, A., Krause, S., Blank, D., Jenny, A., Jenö, P., Lustig, A., and Wahle, E. (1995) Isolation of genomic and cDNA clones encoding bovine poly(A) binding protein II. *Nucleic Acids Res. 23*, 4034−4041.
(34) Wahle, E. (1991) A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. *Cell 66*, 759−768.
(35) Kühn, U., Gündel, M., Knoth, A., Kerwitz, Y., Rüdel, S., and Wahle, E. (2009) Poly(A) tail length is controlled by the nuclear poly(A) binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. *J. Biol. Chem. 284*, 22803−22814.
(36) Benoit, B., Mitou, G., Chartier, A., Temme, C., Zaessinger, S., Wahle, E., Busseau, I., and Simonelig, M. (2005) An essential cytoplasmic function for the nuclear poly(A) binding protein, PABP2, in poly(A) tail length control and early development in Drosophila. *Dev. Cell 9*, 511−522.
(37) Kühn, U., Nemeth, A., Meyer, S., and Wahle, E. (2003) The RNA binding domains of the nuclear poly(A)-binding protein. *J. Biol. Chem. 278*, 16916−16925.
(38) Fronz, K., Güttinger, S., Burkert, K., Kühn, U., Stöhr, N., Schierhorn, A., and Wahle, E. (2011) Arginine methylation of the nuclear poly(A) binding protein weakens the interaction with its nuclear import receptor, transportin. *J. Biol. Chem. 286*, 32986−32994.
(39) Calado, A., Kutay, U., Kühn, U., Wahle, E., and Carmo-Fonseca, M. (2000) Deciphering the cellular pathway for transport of poly(A) binding protein II. *RNA 6*, 245−256.
(40) Cansizoglu, A. E., Lee, B. J., Zhang, Z. C., Fontoura, B. M. A., and Chook, Y. M. (2007) Structure-based design of a pathway-specific nuclear import inhibitor. *Nat. Struct. Mol. Biol. 14*, 452−454.
(41) Imasaki, T., Shimizu, T., Hashimoto, H., Hidaka, Y., Kose, S., Imamoto, N., Yamada, M., and Sato, M. (2007) Structural basis for substrate recognition and dissociation by human transportin 1. *Mol. Cell 28*, 57−67.

(42) Lee, B. J., Cansizoglu, A. E., Süel, K. E., Louis, T. H., Zhang, Z., and Chook, Y. M. (2006) Rules for nuclear localization sequence recognition by karyopherinbeta2. *Cell 126*, 543−558.
(43) Schindelin, H., Marahiel, M. A., and Heinemann, U. (1993) Universal nucleic acid-binding domain revealed by crystal structure of the B. subtilis major cold-shock protein. *Nature 364*, 164−168.
(44) Hutchinson, G. E., and Thornton, J. M. (1994) A revised set of potentials for beta-turn formation in proteins. *Protein Sci. 3*, 2207−2216.
(45) Pease, G., and Watson, C. (1978) Conformational and ion binding studies of a cyclic pentapeptide. Evidence for beta and gamma turns in solution. *J. Am. Chem. Soc. 100*, 1279−1286.
(46) Kumaki, Y., Matshushima, N., Yoshida, H., Nitta, K., and Hikichi, K. (2001) Structure of the YSPTSPS repeat containing two SPXX motifs in the CTD of RNA polymerase II: NMR studies of cyclic model peptides reveal that the SPTS turn is more stable than SPSY in water. *Biochim. Biophys. Acta 1548*, 81−93.
(47) Stradley, S. J., Rizo, J., Bruch, M. D., Stroup, A. N., and Gierasch, L. M. (1990) Cyclic pentapeptides as models for reverse turns: Determination of the euqilibrium distribution between type I and type II conformations of Pro-Asn and Pro-Ala beta-turns. *Biopolymers 29*, 263−287.
(48) Hollósi, M., Kövèr, K. E., Holly, S., Radics, L., and Fasman, G. D. (1987) Beta-turns in bridged proline-containing cyclic peptide models. *Biopolymers 26*, 1555−1572.
(49) Hollosi, M., Kawai, M., and Fasman, G. D. (1985) Studies on proline-containing tetrapeptide models of beta-turns. *Biopolymers 24*, 211−242.
(50) Perczel, A., Hollosi, M. (1996) Turns, In *Circular Dichroism and the Conformational Analysis of Biomolecules* (Fasman, G. D., Ed.), pp 285−380, Plenum Press, New York, London.
(51) Chen, H., Xue, Y., Huang, N., Yao, X., and Sun, Z. (2006) MeMo: a web tool for prediction of protein methylation modifications. *Nucleic Acids Res. 34*, W249−W253.
(52) Shien, D.-M., Lee, T.-Y., Chang, W.-C., Hsu, J. B.-K., Horng, J.-T., Hsu, P.-C., Wang, T.-Y., and Huang, H.-D. (2009) Incorporating structural characteristics for identification of protein methylation sites. *J. Comput. Chem. 30*, 1532−1543.
(53) Shao, J., Xu, D., Tsai, S.-N., Wang, Y., and Ngai, S.-M. (2009) Computational identification of protein methylation sites through bi-profile Bayes feature extraction. *PLoS One 4*, e4920.
(54) Barua, B., Lin, J. C., Williams, V. D., Kummler, P., Neidigh, J. W., and Andersen, N. H. (2008) The Trp-cage: optimizing the stability of a globular miniprotein. *Protein Eng. Des. Sel. 21*, 171−185.
(55) Brahms, H., Raymackers, J., Union, A., de Keyser, F., Meheus, L., and Lührmann, R. (2000) The C-terminal RG dipeptide repeats of the spliceosomal Sm proteins D1 and D3 contain symmetrical dimethylarginines, which form a major B-cell epitope for anti-Sm autoantibodies. *J. Biol. Chem. 275*, 17122−17129.
(56) Kambach, C., Walke, S., Young, R., Avis, J. M., de la Fortelle, E., Raker, V. A., Lührmann, R., Li, J., and Nagai, K. (1999) Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell 96*, 375−387.
(57) Luger, K., and Richmond, T. J. (1998) The histone tails of the nucleosome. *Curr. Opin. Genet. Dev. 8*, 140−146.
(58) Lapeyre, B., Bourbon, H., and Amalric, F. (1987) Nucleolin, the major nucleolar protein of growing eukaryotic cells: An unusual protein structure revealed by the nucleotide sequence. *Proc. Natl. Acad. Sci. U.S.A. 84*, 1472−1476.
(59) Lapeyre, B., Amalric, F., Ghaffari, S. H., Rao, S. V. V, Dumbar, T. S., and Olson, M. O. J. (1986) Protein and cDNA sequence of a glycine-rich, dimethylarginine-containing region located near the carboxyl-terminal end of nucleolin (C23 and 100 kDa). *J. Biol. Chem. 261*, 9167−9173.
(60) Lischwe, M. A., Ochs, R. L., Reddy, R., Cook, R. G., Yeoman, L. C., Tan, E. M., Reichlin, M., and Busch, H. (1985) Purification and partial characterization of a nucleolar scleroderma antigen (Mr = 34 000; pI, 8.5) rich in NG, NG-dimethylarginine. *J. Biol. Chem. 260*, 14304−14310.
(61) Ostareck-Lederer, A., Ostareck, D. H., Rucknagel, K. P., Schierhorn, A., Moritz, B., Huttelmaier, S., Flach, N., Handoko, L., and Wahle, E. (2006) Asymmetric arginine dimethylation of heterogeneous nuclear ribonucleoprotein K by protein-arginine methyltransferase 1 inhibits its interaction with c-Src. *J. Biol. Chem. 281*, 11115−11125.
(62) Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. M., Hipps, K. W., Ausio, J., Nissen, M. S., Reeves, R., Kang, C., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, W., Garner, E. C., and Obradovic, Z. (2001) Intrinsically disordered protein. *J. Mol. Graph. Model. 19*, 26−59.
(63) Ghisolfi, L., Joseph, G., Amalric, F., and Erard, M. (1992) The glycine-rich domain of nucleolin has an unusual supersecondary structure responsible for its RNA-helix-destabilizing properties. *J. Biol. Chem. 267*, 2955−2959.