Published in final edited form as:

J Proteome Res. 2012 October 5; 11(10): 4791–4802. doi:10.1021/pr201196e.

Myelin basic protein undergoes a broader range of modifications in mammals than in lower vertebrates

Chunchao Zhang¹, Angela K. Walker³, Robert Zand^{3,4}, Mario A Moscarello⁵, Jerry Mingtao Yan⁶, and Philip C Andrews^{1,2,3,*}

¹Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109

²Department of Chemistry, University of Michigan, Ann Arbor, MI 48109

³Department of Biological Chemistry, University of Michigan, Ann Arbor, MI 48109

⁴Department of Macromolecular Science and Engineering, University of Michigan, Ann Arbor, MI 48109

⁵Department of Molecular Structure, Hospital for Sick Children, Toronto, Canada

⁶Department of Radiology, University of Michigan, Ann Arbor, MI 48109

Abstract

Myelin basic protein (MBP) is an important component of the myelin sheath surrounding neurons and it is directly affected in demyelinating diseases. MBP contains a relatively large number of post-translational modifications (PTMs) which have been reported to play a role in multiple sclerosis while MBPs from lower vertebrates have been reported to be incapable of inducing multiple sclerosis or allergic encephalitis. This study reveals the extent of differences in PTM patterns for mammalian and non-mammalian MBPs. This included intact mass and *de novo* sequence analysis of approximately 85% of rattlesnake MBP, the first reptile MBP to be characterized and of bovine MBP. We identified 12 PTM sites in the five bovine MBP charge components which includes both previously reported and novel modifications. The most notable modification is an acetylation of lysine 121. Other modifications found in bovine MBP include N-terminal acetylation in components C1, C2, and C3; Oxidation of methionine 19 in all five components; All charge isomers had both a mono- and di-methylated (symmetric) arginine at position 106; Deimination in arginines 23 and 47 was found only in component C8b; Deimination of arginine 96 and deamidation in glutamine 102 was found in components C2, C3, C8a, and C8b; Phosphorylation in threonine 97 was restricted to charge components C2 and C3; Deimination in arginine 161 was only found in component C3; Deamidation of glutamine 120 was only observed in C1. All four deiminated arginines and one acetylated lysine were first experimentally revealed in this study for bovine MBP. Mascot database searching combined with *de novo* sequence analysis of rattlesnake MBP provided more than 85% sequence coverage. A few PTMs were also revealed in rattlesnake MBP: Mono- and dimethylated Arg, Protein N-terminal acetylation, and deiminated Arg. Overall, snake MBP was found to undergo less modification than bovine MBP based on the mass heterogeneity of the intact protein, the bottom-up structure analysis, and the limited complexity of rattlesnake MBP chromatography. The combined data from this study and information from previous studies extend the known MBP PTMs and PTMs unique to higher vertebrates are proposed.

*To whom correspondence should be addressed: Address: Rm. 1198, 300 N. Ingalls, Ann Arbor, MI 48109-0404., andrewsp@umich.edu; Phone: (734) 647-0951; Fax: (734) 615-4941.

Keywords

Myelin basic protein; Post-translational modification; Mass spectrometry; Multiple sclerosis

Introduction

Myelin basic protein, MBP, is a major component of the myelin membrane that envelops the nerve axons of mammalian and sub-mammalian species. It plays a crucial role in maintenance of the tight spiral wrappings of the myelin sheath required for proper functioning of nerves. The sequence of MBP is highly conserved in higher animals and it is a very basic protein (pI=11.65 for Bovine MBP, 18.3 kDa isoform) due to the large number of arginyl and lysyl residues present in the sequence. In bovine preparations, the 18.3 kDa form represents over 95% of the MBP. The primary structure of MBP undergoes extensive post-translational modifications (PTMs) which include phosphorylations, methylation of arginyl residues, citrullination of arginyl residues, deamidated glutaminy residues, and N-terminal acetylation¹⁻⁹. The number and extent of these modifications in mammals appears to be greater than in non-mammalian species^{5,6}.

MBP has been implicated in demyelinating diseases, in particular, multiple sclerosis¹⁰ and allergic encephalitis^{11,12}. Characterized as a human autoantigen, several studies have shown a role for autoantibodies against MBP in the pathogenesis of multiple sclerosis^{10,13}. Genomic studies have yet to provide direct insight into the mechanism underlying this phenomenon and interest in the relationship between MBP and multiple sclerosis has centered on the potential role of these MBP post-translational modifications. The purified mammalian protein containing extensive post-translational modifications differs significantly from the sub-mammalian proteins in its ability to elicit a mammalian demyelinating pathology in experimental allergic encephalomyelitis (EAE) which has been developed as an animal model of multiple sclerosis¹⁴. In this model, specific MBP forms can elicit an immunologic response in some species but fail to elicit a pathological response in other species^{15,16}.

The myriad PTMs in MBP such as methylation on arginyl residues and the addition of negatively charged phosphate groups on seryl and threonyl residues could trigger protein conformational changes, alter interactions between MBP and myelin membranes or between MBP and other myelin associated proteins. One hypothesis proposes that PTMs in MBP could be one of the key factors responsible for MBP antigenicity and the development of multiple sclerosis in mammals. The first inference comes from a quantitative study of the changes of PTMs in human MBP from normal and multiple sclerosis tissues⁸. This study demonstrated that MBP from multiple sclerosis patients was less phosphorylated but more methylated and more deaminated compared with the protein purified from normal tissues. Other studies also showed an aberrant pattern of PTMs that may contribute to demyelination, the protein antigenicity, and its association with a phospholipid membrane^{17,18}. These correlations are consistent with this hypothesis but a strict causal relationship has not yet been established. Which specific PTMs might be unique to higher animals, what differences in modification patterns exist between mammalian and non-mammalian species, and what types of PTMs are present in more evolutionarily distant animals such as reptiles, remain to be addressed and have a direct bearing on this hypothesis in light of the differences in antigenicity observed for lower vertebrates^{8,19-22}.

The extensive and variable PTMs of MBP lead to a complex mixture of isoforms which must be at least partially resolved before characterization. Most previous work was based on resolution of MBP charge isoforms by ion exchange chromatography prior to

characterization. To maintain the ability to correlate previous studies, this study compares the extent and nature of PTMs for these charge isoforms in bovine MBP with rattlesnake MBP and adds to the current limited information on structure, isoforms, and antigenic properties of MBP from mammalian and non-mammalian species. The objective of this study was to determine whether the less antigenic MBP from lower vertebrates exhibit fewer post-translational modifications and to further characterize the PTMs of the antigenic bovine MBP. We extend the characterization of the bovine MBP PTMs across different charge components to identify additional modification sites in each charge component and we identify the probable presence of acetyl-lysine in MBP for the first time. We also reveal the partial sequence of rattlesnake MBP for the first time, identify its PTMs, and document its relatively low level of post-translational modification relative to mammalian MBP.

Materials and Methods

Extraction, Isolation and Purification of MBP Charge Components

The protocol described here for bovine brain was adjusted appropriately for rattlesnake brains. Bovine MBP was isolated from the white matter of beef brains and the rattlesnake MBP were extracted from whole brains of Eastern Diamondback Rattlesnakes (*crotalus adamanteus*) from Georgia using a minor modification of the procedure of Martenson, *et al.*¹⁹. White matter (36–40 gm) from large brains or whole small brains were placed in a Waring blender and 750 ml of cold 2:1 (v/v) chloroform-methanol added plus 1 ml of 0.1 M phenylmethanesulfonyl fluoride and then homogenized at low speed for 1 minute, then at high speed for 3 minutes. The extract containing dispersed tissue was poured into a large beaker and stirred overnight at 4°C. The slurry was then filtered on Whatman No 1 filter paper in a Buchner funnel to form a pad of tissue. The tissue pad was macerated in acetone overnight, with stirring and filtered to give another filter cake. This second filter cake was stirred with 0.2 N H₂SO₄ overnight with phenylmethanesulphonyl fluoride and the extract filtered and transferred to dialysis bags. The extract was dialyzed against deionized and distilled water with at least 4 to 6 changes of water at 4 degrees until the pH of the dialysate is approximately 7. When the dialysis tested neutral for pH (pH 6.8–7.1) the solution was mixed with an equal volume of 95% ethanol and placed in a freezer at –20°C. The resultant precipitate was then collected by centrifugation at 10,000 rpm and the resulting precipitate was washed three times with cold ethanol, and allowed to dry at room temperature. The dried white powder was then stored at –80°C.

Isolation of Charge Components

Using the method of Martenson, *et al.*¹⁹, approximately 100mg (by O.D. 280) of the unfractionated charge component mixture was dissolved in a glycine-urea buffer at pH 9.5 room temperature and applied to a carboxymethyl cellulose column (Whatman CM 52, Pharmacia K9/30) equilibrated with the same buffer. The column was washed with glycine-urea buffer pH 10.5 and the charge component eluted with a linear gradient of 0 to 0.15M sodium chloride contained in 0.08 M glycine, 2M urea, pH 9.5 buffer. Fractions of 3 ml were collected. The charge components elute in order of increasing positive charge. Each fraction containing a major charge component was dialyzed at 4°C, against deionized water overnight, lyophilized and stored at –20°C. To remove minor contaminants in the rattlesnake MBP, the fraction was further separated by a Vydac C8 reverse phase column (Part No. 208TP54, 250mm × 4.6mm) with a linear gradient from 5% to 70% acetonitrile in 0.1% formic acid/water over 90 min. Each component was subjected to SDS gel electrophoresis under basic pH conditions to ascertain purity. Purity of the rattlesnake and bovine MBPs were estimated at 95% and 98% respectively by Coomassie Blue staining of SDS gels.

Protein Digestion, Mass spectrometer Analysis, and PTM Assignment

For each charge component, about 20 μ g of MBP was digested with sequencing-grade trypsin (Promega, Madison, WI) or Glu-C (Roche Applied Science) at a ratio of 1:30 enzyme:protein at 37 degrees, 6h in 100mM ammonium bicarbonate buffer. In addition to trypsin or Glu-C digestion, rattlesnake MBP was also digested with chymotrypsin (Roche Applied Science) at a ratio of 1:20 enzyme:protein at 37 degrees, 3h in the same digestion buffer. The digests were further vacuum dried and reconstituted in 0.1% formic acid for mass spectrometry analysis. For bottom up analysis of myelin basic protein, peptides after trypsin or Glu-C digestion were resolved with a 12cm, in-house made, C18 capillary column (5 μ m, 300Å, column ID: 99 μ m) and eluted into a Thermo Fisher Orbitrap XL mass spectrometer with a linear gradient from 5% to 90% acetonitrile in 0.1% formic acid/water at 200nl/min flow rate over 60 min. The mass spectrometer was run in data dependent mode with 5 CID or CID/HCD hybrid microscans for the most abundant peaks with dynamic exclusion set for 120 seconds. The full scan (MS1) covers the range of m/z from 400 to 2,000 m/z at resolution=60,000. The normalized collision energy was 27% for both CID and HCD. A neutral loss mass list with 32.67, 49, and 98 m/z was set to enhance the detection of phosphorylated peptides. Internal mass calibration was achieved by mass lock of 2 polydimethylcyclsiloxane (PCM) ions (m/z =445.12002 and 429.088735); For intact mass analysis, proteins were directly infused into or separated on a 15cm, in-house made, C8 capillary column (5 μ m, 300Å, column ID: 99 μ m) and eluted into a Thermo Fisher Orbitrap XL mass spectrometer under data dependent mode with 5 CID or HCD events. Digests of rattlesnake MBP were also analyzed on a Sciex model 4800 MALDI TOFTOF to extend the sequence coverage and to aid *de novo* analysis. Proteolytic digest were modified using sulfophenylisothiocyanate (SPITC) according to the method of Joss, *et al.*²³. Manual *de novo* analysis was performed by an experienced mass spectrometrists using MSExpedite, and in-house tool to aid *de novo* sequencing.

Spectra collected from bovine MBP were searched against the Uniprot bovine protein database using the Mascot search engine (Matrix Sciences, version 2.2.02) including its recently available top-down algorithm. Spectra from rattlesnake MBP were error-tolerant searched against NCBI protein database including all species. A false discovery rate was estimated and calculated from the bovine decoy database. Mass error tolerance for the precursor ion was 10ppm and was 0.8Da for the fragment ion. Up to two missed cleavages were allowed for enzyme digestion. The following variable modifications have been considered for both bottom-up and top-down methods: protein N-terminal acetylation, acetylation(K), deamidation(NQ), deimination(R), methylation(KR), phosphorylation(STY) and oxidation(M). Spectra from intact mass analysis were deconvoluted by the Xtract module of BioWorks software from ThermoFisher (Version 2.0.7), to get the monoisotopic mass and the isotope pattern of the intact proteins. *De novo* sequence analysis of rattlesnake MBP was achieved with PEAKS Studio 5.3. All significant spectra assigned with PTMs were considered candidates for further manual validation. Rules for those peptides assigned with PTMs based on *de novo* sequence analysis are: 1) Candidate must have a high-quality spectrum; 2) Most abundant ions should be assigned as b or y ions; 3) They should have more than 2 spectra observed; 4) There were at least 3 consecutive peaks covered; 5) Local confidence on a PTM should be at least more than 50%; and 6) Precursor has a mass error less than 5ppm. The rattlesnake *de novo* sequence analysis was performed manually from MALDI TOFTOF spectra using the MSExpedite software.

Bovine MBP isoforms or rattlesnake MBP resolved by SDS gel electrophoresis were verified from an in-gel trypsin digest of Coomassie Blue-stained gels after direct application to a MALDI plate in alpha-cyano-4-hydroxycinnamic acid matrix and analysis on a Sciex model 4800 MALDI TOFTOF. The spectra were identified by a MASCOT search against the protein database.

Western Blot Analysis of Protein Acetylation in Bovine MBP

Bovine MBP C1-C3 components with Tetrahymena histone H3 and bovine carbonic anhydrase II as positive and negative controls, respectively, were separated by SDS-PAGE under reducing conditions and transferred to nitrocellulose membranes. Membrane was washed and blocked with non-fat dry milk at room temperature and then incubated overnight with acetylated-lysine rabbit monoclonal antibody (Ac-K²-100, 1:500) purchased from Cell Signaling Technology. The membrane was washed and then incubated with peroxidase-labeled Anti-Rabbit IgG (1:1000) for 60 min and visualized using the ECL chemiluminescent reagent (GE Healthcare/Amersham) with a 10 min exposure. Finally, the film was imaged and analyzed with the Bio-Rad imaging system and associated software.

The data associated with this manuscript may be downloaded from ProteomeCommons.org Tranche using the following hash: jJT
+Xcym0KuG7DtnAI8+9Ngx859wDfcuYq0LSrKUtQsn/fsQGSMIB+njKvAwCseg5y
9sTq699B/TlfPmI5vrX9cOL8sAAAAAADTDg==

Results and Discussion

PTMs found in bovine MBP and those found in other species

Generally, the sequence coverage generated from trypsin digestion was more than 80% and it was more than 60% from Glu-C digestion for a total coverage of more than 90% for the combined methods. A typical false discovery rate for tryptic peptides at the identity threshold was around 1.5% and it was under 3% for those peptides cleaved from Glu-C digestion. At least three variants of bovine MBP appeared to be present in the preparation by SDS gel electrophoresis, a major form corresponding to approximately 18 kDa and minor forms corresponding to approximately 16 kDa, 23 kDa, and 37 kDa. Some of these minor forms were verified by MALDI peptide mass fingerprinting of the gel bands analysis of tryptic digests and altogether were estimated to correspond to less than 2% of the total MBP and thus did not interfere with the intact mass analyses. Rattlesnake MBP was estimated to be 19 kDa by SDS gel electrophoresis.

We identified a total of 12 PTMs in all 5 charge components from bovine MBP: C1–3, C8a and C8b based on LC-MS/MS of proteolytic digests (Figure 1, Table 1). The PTMs identified are: acetylated Ala at the protein N-terminus; oxidized M19; deimination at R23, R47, R96 and R161; deamidation at Q102 and Q120; phosphorylated T97; mono- and dimethylated R106; and acetylated K121.

Among these modifications of bovine MBP, the putative presence of acetylation of a lysyl residue has not been previously reported. The modified peptide identified in this study was consistent with the presence of acetyl-lysine (Figure 2) at position 121, adjacent to a deamidated glutamine at 120. The total mass shift of 43 Daltons is localized to the QKP tripeptide based on the observed fragment ions and was initially interpreted as a potential modification of the lysyl residue in the deamidated peptide. The observed mass (1842.8370 Da) at 2.4 ppm is consistent with the calculated mass for the acetyl-lysine peptide (1842.8326 Da) rather than trimethyllysine (1842.7966 Da) at 22 ppm but indistinguishable from the mass of the carbamylated and non-deamidated peptide (1842.8438) at 3.7 ppm. The average experimental mass accuracy for peptides identified in this study is 5 ppm.

Inspection of the MS/MS spectrum identified a weak fragment ion at 901.18 Da which might correspond to the neutral loss reported for carbamylated peptides^{24,25}, but the intensity is much lower than expected and the mass assignment is limited by mass accuracy in the MS/MS mode. To distinguish between the presence of acetylation or carbamylation on Lys 121, a Western blot of bovine MBP (Figure 3) using anti-acetyllysine antibody was

performed. The immunoblot is consistent with the presence of acetyllysine in MBP, with the highest level in component C3. The staining intensity observed for acetyl-lysine immunoreactivity in MBP is consistent with the low spectral count obtained for the putative acetylated form of this peptide (5 spectra were detected in C3, none were found in other components). This modified lysine appears to be a rare modification based on spectral counts of the modified versus non-modified peptide. Note that the observation of a probable acetyl-lysine only in the deamidated peptide may simply be a sampling issue and we cannot exclude the possibility that the parent peptide is not also acetylated.

We identified both mono- and dimethylated forms of R106 in all five MBP components (Figure 4a and b) which verified a previous study¹. The strong neutral losses of 31 Da and 49 Da observed in the MSMS spectra correspond respectively to monomethylamine (MMA, H₂N-CH₃) and MMA+H₂O²⁶, confirming that R106 was symmetrically dimethylated (Figure 4b). This is consistent with previous reports for other species which used non-mass spectrometric methods¹. The level of methylation observed varied in the different charge components. By intact mass analysis, the ratio of monomethylation vs dimethylation decreased in components C1 through C3 (Figure 5a–c) but it was not possible to verify this in components C8a and C8b because intact mass analysis failed to produce high-quality spectra due to impurities present in these two components. Brostoff, S. *et al.*, postulated that the biological function of the methylated arginine residue was to stabilize the double-chain structure for the protein induced by the triproline site¹. In human MBP, more highly methylated Arg was observed in multiple sclerosis patients compared with normal tissue and has been suggested to play an important role in the pathogenesis of multiple sclerosis⁸. It is noteworthy that this modification has been found in many species such as human, bovine, rabbit, and chicken but was not found in spiny dogfish. We also found clear evidence for this modification in rattlesnake based on intact mass analysis of rattlesnake MBP (Figure 5d). Spectra from MSMS also confirmed this modification in a peptide with the sequence GRGLSF_{SR} (Figure 6a and b) as this sequence matched well with known species and the fragmentation pattern of this peptide is very similar to that of the methylated peptide from bovine MBP (Figure 4a and b). Absence of this modification in some lower animals such as dogfish may be due to lack of an arginine in this position (Figure 7), suggesting methylated arginine is widely distributed from mammals to reptiles but absent in MBP from lower vertebrates which lack this residue.

Several sites were observed in bovine MBP for conversion of arginine to citrulline. Catalysed by peptidylarginine deiminases (PADs), deimination (citrullination) of arginine residues can have significant consequences for the structure and function of proteins, and has also been implicated in the pathogenesis of multiple sclerosis^{27–29}. Native MBP contains several non-deiminated arginines and forms tighter, more compact myelin sheaths. Deimination of MBP affects the stability of the myelin sheath by conversion of positively charged arginine into uncharged citrulline, which increases the hydrophobicity of the protein and reduces the interaction with the negatively charged phosphatidylserine in the membrane, leading to a more open structure susceptible to proteolysis by cathepsin D^{29,30}. The ratio of deiminated MBP/total MBP was found to be crucial in the physiological function of CNS and the degree of deimination of MBP correlated well with the severity of multiple sclerosis^{28,29}. MBP was highly deiminated in multiple sclerosis patients and in infants during normal CNS development and was less deiminated in healthy adults^{8,28,29}. More interestingly, possible crosstalk between deimination and methylation in MBP has been revealed by Pritzker *et al.*³¹ who observed that methylation of R106 in bovine MBP (R107 in human MBP) had a significant correlation to arginine deimination. In human MBP, increased deimination of arginyl residues accompanied with a decreased methylation of R107 has been observed. In this study, we found a total of 4 arginines deiminated at positions 23, 47, 96, and 161 respectively (Figure S1–4). Although deimination of arginine

in MBP has been found in other species, deimination has not previously been reported in bovine MBP. The residue corresponding to R23 in bovine MBP has been reported to be deiminated in human and chicken MBP^{6,8} and it was also observed in rattlesnake MBP in this study (Figure S5) indicating that deimination is highly conserved at this site. The residues corresponding to R47 and R161 in bovine MBP have only been previously observed to be deiminated in human MBP, while deiminated R96 found in this study has not been found in other species thus far. The conserved Arg in rattlesnake corresponding to bovine R161 was not found to be deiminated but the nearby R (in peptide sequence GYRYDGQGTLISK, Figure S6) was found to be partially deiminated (both forms observed by MS). There is no homologous Arg residue in mammalian MBPs. The degree of conservation of deimination in mammalian MBPs is still incomplete as deimination of Arg has not yet been examined in rabbit or other mammalian MBPs.

Acetylated alanine at the protein N-terminus was found in components C1, C2, and C3 in this study. This modification was not found in the previous study of bovine MBP PTMs by capillary electrophoresis-mass spectrometry². And it was also not found in the tryptic peptides in this study possibly because the peptide with N-acetylalanine (AAQK) was too short to be effectively detected by the mass spectrometry method or the peptide did not ionize well. When we digested MBP with Glu-C protease, the N-terminally acetylated peptide was clearly detected and the PTM was unambiguously assigned by tandem mass spectrometry (Figure S7). Intact protein mass analysis also agreed with our Glu-C bottom-up study (Figure 5a–c). Bovine MBP in C1–3 appears to be predominantly acetylated since no spectra corresponding to the unmodified peptide or protein were observed in both bottom-up and intact mass analyses. Acetylation of the N-terminal alpha-amino group of proteins is a common modification in eukaryotes and about 85% of eukaryotic proteins are N-terminally acetylated³², with N-terminal Ala being a common acetylation site. As a conserved and widespread modification, the biological role of this N-terminal acetylation is unclear, although it has been suggested that one biological function is as a signal for protein degradation³³. This modification was also found in other MBPs from mammals such as human, bovine, and rabbit, to non-mammals including chicken, dogfish, and rattlesnake in this study (Figure S8).

A wide range of MBP phosphorylation sites have been reported in previous studies^{1–9}, however, this study did not focus on phosphorylation so no effort was made to enrich for phosphopeptides. Despite the absence of enrichment methods, we identified a phosphorylation site at T97 which was present in bovine MBP components C2 and C3 (Figure S9).

Phosphorylation of T97 has been previously reported for human, bovine, rabbit, and chicken MBPs but not in spiny dogfish because the sequence (PRTPPP) associated with this PTM is highly conserved in higher animals but is missing in dogfish. Other conserved phosphorylated sites are: S7 found in human, bovine, rabbit and chicken; Phosphorylated Ser in sequence RGS^uGK corresponding to S54 in bovine, and S56 in human and rabbit; Phosphorylated Ser at protein C-terminus (SGSPXARR, X is M or V) in human, bovine, rabbit, and chicken. None of these phosphorylated sites were detected in spiny dogfish and appear to be absent also in rattlesnake as no clear evidence of phosphorylation was found from the intact mass analysis and *de novo* sequencing study (see below).

Deamidated Q102 was found to be present in components C2, C3, C8a, and C8b (Figure S10); A deamidated glutamine at position 120 is also presented in C3 (Figure 2). Deamidation of Q102 and Q120 was also observed in human MBP but has not been examined for rabbit MBP^{8,9}. Overall, the C1 fraction was least modified while the less basic C3 and C8b fractions were more highly modified. The analysis of proteolyzed MBP

fractions were also consistent with our intact mass analysis of MBP proteins that the deconvoluted spectra showed C3 had the most complicated spectra as compared with C1 and C2 fractions (Figure 5a–c). We have no unambiguous evidence that rattlesnake MBP is deamidated and note that at least some of the Gln and Asn residues reported to be deamidated in mammals are substituted in rattlesnake by other residues. However, the MALDI MS/MS sequence for the peptide isolated from a GluC digest having the sequence HKYAHQ/KGHQ/KGYRYDGE appears to end in a Glu residue despite strong evidence for a Gln residue in other peptides spanning this region.

The oxidized methionine in bovine MBP at position 19 has not been reported before and was found in all five components in this study (Figure S11). While it is quite likely that the methionine oxidation observed in this study resulted after the protein was extracted from the cell, we cannot rule out its presence *in vivo*. It is worth noting that several studies in other proteins have demonstrated that methionine oxidation has a significant effect on protein function, stability, aggregation and folding^{34–38}. More generally, oxidation of methionine can decrease protein stability. One explanation for the effect is that the extra oxygen atom introduced by methionine oxidation changes the protein hydrophobicity at that site³⁹.

Intact mass study of bovine MBP charge components and unfractionated rattlesnake MBP

The presence and distribution of several PTMs in three charge components of bovine MBP, C1–3, have been successfully characterized by intact mass analysis (Figure 5a–c). In C1, the first major component had a calculated monoisotopic mass of 18,354.44 Da while the theoretical monoisotopic mass calculated from the protein sequence was 18312.39 Da. The mass difference was +42.05 Da which is consistent with the N-terminal acetylation observed at the peptide level. The mass of the second major peak was further increased by 14.04 Da which is consistent with monomethylation of the N-terminally acetylated protein. The third major peak had a mass difference of +28.05 Da from the first peak suggesting that MBP in the C1 fraction was also modified with a dimethylation. The most abundant peak ($M_w=18,368.48$ Da) in the C1 fraction corresponds to MBP modified with both N-terminal acetylation and monomethylation. All the ions found in the C1 fraction were also present in the C2 fraction. However, the most abundant peak in C2 corresponds to the protein modified with N-terminal acetylation and dimethylation. An additional species mass shifted by +0.99 Da from 18,368.48 Da was found in C2, which is consistent with either deamidation or citrullination. Evidence consistent with the presence of an oxidized Met in C2 came from a peak with mass increased by +16 Da from 18,382.48 Da. The C3 component had the most complex spectrum arising from the combination of different PTMs. In addition to those PTMs found in C1 and C2, deamidated or citrullinated residues were suggested by mass increases of 1 or 2 Da. Although all three components were modified with a mono- and dimethylated Arg, the distribution of non-methylated, mono- and dimethylated forms changed in different charge isomers. For instance, the ratio of non-methylated: monomethylated: dimethylated was 1 : 1.18 : 0.80 in C1 while this ratio was changed to 1 : 0.45 : 1.14 in C2. Generally, the results from intact mass analysis were consistent with those from analysis of the proteolytic digests: C1 was least modified while C3 was most modified.

Unlike the spectra from bovine MBP fractions C1–C3, intact mass analysis of rattlesnake MBP revealed a much simpler spectrum (Figure 5d), suggesting rattlesnake MBP has considerably fewer PTMs than bovine MBP although the protein sequences are fairly well conserved in many species. Note that the complete protein sequence for rattlesnake MBP is unknown. The intact monoisotopic mass of rattlesnake MBP was approximately 19,564 Da. A form having a mass shift of 1 Da was observed for intact rattlesnake MBP suggesting rattlesnake MBP might be deamidated or citrullinated. This is consistent with the evidence with citrullination observed from the bottom-up analysis discussed below. No evidence was observed for deamidation in the bottom-up analysis. A mass shift of 14 Da is consistent with

the bottom-up analysis which identified methylation of Arg in the peptide, **GRGLSFR**. Spectra of intact rattlesnake MBP are generally less post-translationally modified compared with spectra of intact bovine MBP.

Partial protein sequence of rattlesnake MBP and its PTMs

We performed Mascot database searching against the entire protein sequence database with error tolerance as well as *de novo* sequence analysis of peptides from rattlesnake MBP to search for conserved sequences and possible PTMs present in the protein. The *de novo* sequencing software, PEAKS 5.3, was employed in this study for Orbitrap with manual confirmation. The error tolerance was 10ppm for precursor ions and 0.8 Da for fragment ions. Total Local Confidence (TLC) and Average Local Confidence (ALC%) scores have been used to determine the quality of the predicted sequence for particular MS/MS spectra⁴⁰. In addition to these procedures, independent manual *de novo* sequencing was performed on MALDI TOF/TOF MS/MS spectra by an expert (Walker). Mascot database searching together with the *de novo* interpretations identified more than 20 unique peptides (or a total of 150 AAs) and 4 PTMs (acetylation, monomethylation, dimethylation, deimination) at 4 sites in rattlesnake MBP (Table S1, Figure S12), which covers 85% of gekko and 89% of anole MBP sequences respectively (Figure 8). One additional potential modification was observed for deamidation of Gln to generate the C-terminal Glu residue in the peptide HKYAHKGHKG YRYDGE. The residue assignments for I/L and K/Q in Figure 8 were based on the Anole and Gekko sequences except when v and w fragment ions were able to distinguish I/L (See Figure S13 for an annotated example spectrum and Table S1 under MALDI TOF/TOF peptides). The sequence of rattlesnake MBP exhibited good sequence identity with most species but is closest to reptile MBP sequences (gekko and anole) (Figure 8). Mono- and dimethylated Arg were also identified in a peptide with the sequence **GRGLSFR** and this peptide exhibited a very similar fragmentation pattern to the homologous peptide from bovine MBP having only a single amino acid difference (Figures 6a and b). Other PTMs found in this species are: Protein N-terminal acetylation in a chymotryptic peptide (Figure S8) and a deiminated Arg in two tryptic peptides (LATASTIDHARHGSPR and GYRYDGQGTL SK) (Figures S5 and S6). Both of these sites are partially deiminated as evidenced by observation of peptides containing arginine as well as peptides containing citrulline. All MBPs from other species including rattlesnake, which have been previously investigated or addressed in this study, have been found to have acetylated alanine at the N-terminus indicating this modification is conserved across different species (Figure 7).

Conclusions

Similarly to histone proteins, myelin basic protein is a very basic, sequence-conserved protein exposed to various post-translational modifications. Extensive post-translational modifications including acetylation, deamidation, deimination, methylation, and phosphorylation are found in both proteins. Interestingly, the dynamic transitions between transcriptionally-active and transcriptionally-silent chromatin states is determined by a combinatorial set of these PTMs on one or more histones, a well-known ‘histone code’ hypothesis, which determines the compaction of eukaryotic chromatins^{41,42}. Some evidence has also emerged that the compaction of the myelin sheath in CNS could also be affected by dynamic protein post-translational modifications of MBP^{8,18} and the limited number or absence of these modifications in lower vertebrates may have implications for myelin function in these species.

The total number of conserved PTMs found in our study might be underestimated as the different methodologies employed to identify PTMs in different studies may have some

biases and low level modifications may be missed. Our intact MS analyses, however, strongly suggest that the bulk of rattlesnake MBP is much less modified than bovine MBP. This observation is consistent with the bottom-up analysis and with the single fraction observed on ion exchange chromatography.

Our study represents the most extensive MS/MS analysis of bovine myelin basic protein PTMs and leads to identification of new PTMs in bovine MBP and the relative levels of the PTM among isoforms has been revealed. This also represents the first analysis of PTMs in MBP from a reptile through tandem mass spectrometry and demonstration that citrullination and methylation of specific homologous Arg residues are conserved from snake to primates. We conclude that the results of this study are consistent with the general correlation between MBP PTMs and the antigenicity of MBP in demyelinating diseases such as multiple sclerosis and suggests specific modifications to evaluate.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank A.J. Molascon, University of Michigan Pathology Dept, for performing the western blot in this study. We also wish to thank Randy Campbell, Campbell Farms, Alpine, Tennessee 38543 for the rattlesnake brains. This work was funded in part by NIH Grant #1P41RR018627 (PCA).

References

1. Brostoff S, Eylar EH. Localization of methylated arginine in the A1 Protein from myelin. *Proc Natl Acad Sci U S A*. 1971; 68(4):765–9. [PubMed: 4994464]
2. Zand R, Li MX, Jin X, Lubman D. Determination of the sites of posttranslational modifications in the charge components of bovine myelin basic protein by capillary electrophoresis-mass spectroscopy. *Biochemistry*. 1998; 37(8):2441–9. [PubMed: 9485392]
3. Brostoff SW, Eylar EH. The proposed amino acid sequence of the P1 protein of rabbit sciatic nerve myelin. *Arch Biochem Biophys*. 1972; 153(2):590–8. [PubMed: 4662101]
4. Martenson RE, Law MJ, Deibler GE. Identification of multiple *in vivo* phosphorylation sites in rabbit myelin basic protein. *J Biol Chem*. 1983; 258(2):930–7. [PubMed: 6185481]
5. Zand R, Jin X, Kim J, Wall DB, Gould R, Lubman DM. Studies of posttranslational modifications in spiny dogfish myelin basic protein. *Neurochem Res*. 2001; 26(5):539–47. [PubMed: 11513482]
6. Kim J, Zhang R, Strittmatter EF, Smith RD, Zand R. Post-translational modifications of chicken myelin basic protein charge components. *Neurochem Res*. 2009; 34(2):360–72. [PubMed: 18618245]
7. Baldwin GS, Carnegie PR. Isolation and partial characterization of methylated arginines from the encephalitogenic basic protein of myelin. *Biochem J*. 1971; 123(1):69–74. [PubMed: 5128665]
8. Kim JK, Mastronardi FG, Wood DD, Lubman DM, Zand R, Moscarello MA. Multiple sclerosis: an important role for post-translational modifications of myelin basic protein in pathogenesis. *Mol Cell Proteomics*. 2003; 2(7):453–62. [PubMed: 12832457]
9. Kim J, Zand R, Lubman DM. Electrophoretic mobility for peptides with post-translational modifications in capillary electrophoresis. *Electrophoresis*. 2003; 24(5):782–93. [PubMed: 12627438]
10. Berger T, Rubner P, Schautzer F, Egg R, Ulmer H, Mayringer I, Dilitz E, Deisenhammer F, Reindl M. Antimyelin antibodies as a predictor of clinically definite multiple sclerosis after a first demyelinating event. *N Engl J Med*. 2003; 349(2):139–45. [PubMed: 12853586]
11. Mannie MD, Paterson PY, U'Prichard DC, Flouret G. Induction of experimental allergic encephalomyelitis in Lewis rats with purified synthetic peptides: delineation of antigenic

- determinants for encephalitogenicity, in vitro activation of cellular transfer, and proliferation of lymphocytes. *Proc Natl Acad Sci U S A*. 1985; 82(16):5515–9. [PubMed: 3875098]
12. Gupta MK. Myelin basic protein and demyelinating diseases. *Crit Rev Clin Lab Sci*. 1987; 24(4): 287–314. [PubMed: 2436854]
 13. Meinl E, Hohlfeld R. Immunopathogenesis of multiple sclerosis: MBP and beyond. *Clin Exp Immunol*. 2002; 128(3):395–397. [PubMed: 12067291]
 14. Steinman L. Assessment of animal models for MS and demyelinating disease in the design of rational therapy. *Neuron*. 1999; 24(3):511–4. [PubMed: 10595504]
 15. Driscoll BF, Kira J, Kies MW, Alvord EC. Mechanism of demyelination in the guinea pig. Separate sensitization with encephalitogenic myelin basic protein and non encephalitogenic brain components. *Neurochem Pathol*. 1988; 4(1):11–22. [PubMed: 2423932]
 16. Agrawal HC, Banik NL, Bone AH, Cuzner ML, Davion AN, Mitchell RF. The chemical composition of dogfish myelin. *Biochem J*. 1971; 124(5):70P.
 17. Harauz G, Libich DS. The classic basic protein of myelin-conserved structural motifs and the dynamic molecular barcode involved in membrane adhesion and protein-protein interactions. *Curr Protein Pept Sci*. 2009; 10(3):196–215. [PubMed: 19519451]
 18. Harauz G, Ladizhansky V, Boggs JM. Structural polymorphism and multifunctionality of myelin basic protein. *Biochemistry*. 2009; 48(34):8094–104. [PubMed: 19642704]
 19. Martenson RE, Deibler GE, Kies MW, Levine S, Alvord EC. Myelin basic proteins of mammalian and sub mammalian vertebrates: encephalitogenic activities in guinea pigs and rats. *J Immunol*. 1972; 109(2):262–70. [PubMed: 4114825]
 20. Agrawal HC, O'Connell K, Randle CL, Agrawal D. phosphorylation *in vivo* of four basic proteins of rat brain myelin. *Biochem J*. 1982; 201(1):39–47. [PubMed: 7082288]
 21. Patterson PY. A study of experimental encephalomyelitis employing mammalian and non mammalian nervous tissues. *J Immunol*. 1957; 78(6):472–5. [PubMed: 13463329]
 22. Zand, R. Myelin Basic Protein. Boggs, JM., editor. Nova Science Publishers, Inc; New York: 2008. p. 25-27.
 23. Joss JL, Molloy MP, Hinds LA, Deane EM. Evaluation of Chemical Derivatisation Methods for Protein Identification using MALDI MS/MS. *Int J Pept Res Ther*. 2006; 12(3):225–35.
 24. Park ZY, Sadygov R, Clark JM, Clark JJ. Assigning *in vivo* carbamylation and acetylation in human lens proteins using tandem mass spectrometry and database searching. *Int J Mass Spectrom*. 2007; 259(1–3):161–73.
 25. Lapko VN, Smith DL, Smith JB. Methylation and carbamylation of human γ -crystallins. *Protein Sci*. 2003; 12(8):1762–74. [PubMed: 12876325]
 26. Brame CJ, Moran MF, McBroom-Cerajewski LD. A mass spectrometry based method for distinguishing between symmetrically and asymmetrically dimethylated arginine residues. *Rapid Commun Mass Spectrom*. 2004; 18(8):877–81. [PubMed: 15095356]
 27. Ishigami A, Maruyama N. Importance of research on peptidylarginine deiminase and citrullinated proteins in age-related disease. *Geriatr Gerontol Int*. 2010; 10(Suppl 1):S53–8. [PubMed: 20590842]
 28. György B, Tóth E, Tarcsa E, Falus A, Buzás EI. Citrullination: a posttranslational modification in health and disease. *Int J Biochem Cell Biol*. 2006; 38(10):1662–77. [PubMed: 16730216]
 29. Harauz G, Musse AA. A tale of two citrullines--structural and functional aspects of myelin basic protein deimination in health and disease. *Neurochem Res*. 2007; 32(2):137–58. [PubMed: 16900293]
 30. Pritzker LB, Joshi S, Gowan JJ, Harauz G, Moscarello MA. Deimination of myelin basic protein. 1. Effect of deimination of arginyl residues of myelin basic protein on its structure and susceptibility to digestion by cathepsin D. *Biochemistry*. 2000; 39(18):5374–81. [PubMed: 10820008]
 31. Pritzker LB, Joshi S, Harauz G, Moscarello MA. Deimination of myelin basic protein. 2. Effect of methylation of MBP on its deimination by peptidylarginine deiminase. *Biochemistry*. 2000; 39(18):5382–8. [PubMed: 10820009]
 32. Glozak MA, Sengupta N, Zhang X, Seto E. Acetylation and deacetylation of non-histone proteins. *Gene*. 2005; 363:15–23. [PubMed: 16289629]

33. Hwang CS, Shemorry A, Varshavsky A. N-terminal acetylation of cellular proteins creates specific degradation signals. *Science*. 2010; 327(5968):973–7. [PubMed: 20110468]
34. Kim YH, Berry AH, Spencer DS, Stites WE. Comparing the effect on protein stability of methionine oxidation versus mutagenesis: steps toward engineering oxidative resistance in proteins. *Protein Eng*. 2001; 14(5):343–7. [PubMed: 11438757]
35. Wood MJ, Helena Prieto J, Komives EA. Structural and functional consequences of methionine oxidation in thrombomodulin. *Biochim Biophys Acta*. 2005; 1703(2):141–7. [PubMed: 15680222]
36. Wang ZY, Shimonaga M, Muraoka Y, Kobayashi M, Nozawa T. Methionine oxidation and its effect on the stability of a reconstituted subunit of the light-harvesting complex from *Rhodospirillum rubrum* *Eur J Biochem*. 2001; 268(12):3375–82.
37. Wong YQ, Binger KJ, Howlett GJ, Griffin MD. Methionine oxidation induces amyloid fibril formation by full-length apolipoprotein A-I. *Proc Natl Acad Sci U S A*. 2010; 107(5):1977–82. [PubMed: 20133843]
38. Sigalov AB, Stern LJ. Oxidation of methionine residues affects the structure and stability of apolipoprotein A-I in reconstituted high density lipoprotein particles. *Chem Phys Lipids*. 2001; 113(1–2):133–46. [PubMed: 11687233]
39. Hoshi T, Heinemann S. Regulation of cell function by methionine oxidation and reduction. *J Physiol*. 2001; 531(Pt 1):1–11. [PubMed: 11179387]
40. Ma B, Lajoie G. *De novo* interpretation of tandem mass spectra. *Curr Protoc Bioinformatics*. 2009; Chapt 25(Unit13.10.1–13.10.8)
41. Strahl BD, Allis CD. The language of covalent histone modifications. *Nature*. 2000; 403(6765):41–5. [PubMed: 10638745]
42. Jenuwein T, Allis CD. Translating the histone code. *Science*. 2001; 293(5532):1074–80. [PubMed: 11498575]

>sp|P02687|MBP_BOVIN; pI=11.65; Average Mass:18323.48; Monoisotopic Mass:18312.39

1-50 AAQKRPSQRS KYLASASTMD HARHGFLPRH RDTGILDSLG RFFGSDRGAP
 51-100 KRGSGKDGHH AARTTHYGSL PQKAQGHRPQ DENPVVHFFK NIVTPRTPPP
 101-150 SQGKGRLSL SRFSWGAEQ KPGFGYGGRA SDYKSAHKGL KGHDAQGTLS
 151-169 KIFKLGGRDS RSGSPMARR

Figure 1. All PTMs found in this study indicated in the bovine MBP primary sequence

A total of 11 PTMs were found in bovine MBP and their positions are underlined in the figure. Modifications include: acetylated Ala at the protein N-terminus; oxidized M19; deimination at R23, R47, R96 and R161; deamidation at Q102 and Q120; phosphorylated T97; mono- and dimethylated R106; and acetylated K121.

M/Z: 922.4236, Charge: +2, Deam (Q120), Acetyl(K121), Ion Score: 31, Expect:0.8

Figure 2. MS/MS Fragmentation of 113-FSWGAEQGKPGFGYGGR-129
Peptide was observed to be deamidated at Q120, and acetylated at K121 with M/
Z=922.4236, Charge: +2, Mass error: 6ppm, Ion Score: 31, Expect:0.8.

Figure 3. Confirmation of the acetylated-Lysine in MBP fraction C3 by Western Blot
 Acetylated-Lysine Rabbit monoclonal antibody: Ac-K²-100 (1:500); Anti_Rabbit IgG (1:1000); 10 minute exposure; Amount of protein loaded from Left to Right: histone H3 (tetrahymena thermophila): 3 μ g; carbonic anhydrase II (bovine): 30 μ g; MBP C1-C3 (bovine): 40 μ g. Immunoreactivity was detected in all 3 charge components (C1, C2, C3) and in the positive control (histone H3) while no band was found in the negative control (bovine carbonic anhydrase II). The strongest signal was observed for fraction C3 in which Acetyl-lysine was identified by MSMS.

a. Monomethylated R106 in peptide sequence 105-GRGLSLSR-112

M/Z: 430.2590, Charge: +2, Mass error: 1ppm, Ion Score: 39, Expect: 0.041

b. Dimethylated R106 at the same position

Bovine_MBP_C3_trypsin_1 #150-4624 RT: 1.79-58.48 AV: 25 NL: 1.68E3
T: Average spectrum MS2 437.27 (150-4624)

M/Z: 437.2669, Charge: +2, Mass error: 1ppm, Ion Score: 31, Expect: 0.17

Figure 4. Identification of arginine methylation at position 106 in all five components

a) Monomethylated R106 in peptide sequence 105-GRGLSLSR-112. M/Z: 430.2590, Charge: +2, Mass error: 1ppm, Ion Score: 39, Expect: 0.041; b) Dimethylated R106 at the same position. The strong neutral losses of 31 Da (monomethylamine) and 49 Da (monomethylamine + water) correspond to peaks labeled $[M-MMA + 2H]^{+2}$ and $[M-MMA-H_2O + 2H]^{+2}$ respectively, confirmed that R106 was a symmetrical dimethylation. M/Z: 437.2669, Charge: +2, Mass error: 1ppm, Ion Score: 31, Expect: 0.17. .

a) C1 component

b) C2 component

c) C3 component

d) Rattlesnake MBP

PTM pattern is much simpler in snake

Figure 5. Intact mass analysis of 3 bovine MBP components and unfractionated rattlesnake MBP

a) C1 component: The mass difference between the observed monoisotopic mass of 18,354.44 Da and the theoretical monoisotopic mass of 18312.39Da is +42.05 Da which is consistent with the N-terminal acetylation observed at the peptide level. The mass increases of 14.04 Da and 28.05 Da are also consistent with the masses of mono- and dimethylated forms. The most abundant peak ($M_w=18,368.48\text{Da}$) in the C1 fraction corresponds to MBP modified with both N-terminal acetylation and monomethylation. The ratio of unmethylated: monomethylated: dimethylated form was 1:1.18 :0.80 in C1. Me1 and Me2 indicate the monomethylated and dimethylated forms respectively. N-Ac, Deam, Citru, and Oxidation refer to N-terminal acetylation, deamidation of glutamine or asparagine, conversion of arginine to citrulline, and oxidation of methionine, respectively. b) C2 component: all the ions found in the C1 fraction were also present in C2 fraction. However, the most abundant peak in C2 corresponds to the protein modified with N-terminal acetylation and dimethylation. Deamidation or citrullination was found in C2 as +0.99 Da mass shift was observed. An oxidized Met in C2 came from a peak with mass increased by +16 Da from 18,382.48Da. The ratio of the unmethylated: monomethylated: dimethylated forms in C2 was 1:0.45:1.14. c) C3 component: C3 component had the most complex spectrum arising from the combination of different PTMs. More deamidated/citrullinated forms were observed by mass increases of 1 or 2 Da. d) Rattlesnake MBP: intact mass analysis of rattlesnake MBP revealed a much simpler spectrum as compared with spectra of bovine MBP charge components. Rattlesnake MBP has an intact monoisotopic mass roughly around 19,564 Da and forms having masses consistent with deamidation, citrullination, and methylation were observed.

a. Monomethylated Arg in peptide sequence GRGLSFSR

b. Dimethylated Arg in peptide sequence GRGLSFSR

Figure 6. Identification of arginine methylation in rattlesnake MBP

a) Monomethylated arginine in a peptide with sequence GRGLSFSR. M/Z: 447.2523, Charge: +2, Mass error: 2ppm, Ion Score: 30, Expect: 3.4; b) Dimethylated Arginine at the same position. M/Z: 454.2591, Charge: +2, Mass error: 0ppm, Ion Score: 23, Expect: 17.

Figure 7. Comparison of MBP PTMs in different species

PTM information collected from previous studies and including our present data (human[Ref:12–14], bovine[Ref:6, 7, and present data], rabbit[Ref:8, 9], chicken[Ref:11], dogfish[Ref:10]). Sequences of mammalian MBPs were taken from the major, 18.5kd isoforms. Sequence alignment is based on CLUSTAL 2.1 multiple sequence alignment. Asterisks below the sequences indicate highly conserved residues in the primary structures. For dogfish S120 and S121 (italic), it was not possible to distinguish which residue was phosphorylated. PTMs are indicated with symbols as ●: Phosphorylation, ♣: Acetylation, ♠: Methylation (Mono- or Dimethylation), ▲: deamidation, ◆: Oxidation, and ▼: Deimination. Conserved PTMs across species are indicated in bold type and boxed.

Bovine	AAQKRPSQR--SKYLASASTMDHARH--GFLPRHRDTGILDSLGRFFG--	44
Rabbit	ASQKRPSQRHGSKYLATASTMDHARH--GFLPRHRDTGILDSIGRFFS--	46
Human	ASQKRPSQRHGSKYLATASTMDHARH--GFLPRHRDTGILDSIGRFFG--	46
Snake	ASQKRSSFRHGSK-LATASTIDHARH--GSPR-YREPSLLDSLGRFFG--	
Anole	ASQKRSSFRHGSK-LASASIIDHARH--GSPRRYREPGLLDSLGRFFG--	45
Gekko	ASQKRSSFRHGSK-LATASTIDHARH--GSPRRYREPGLLDSLGRFFG--	45
Chicken	ASQKRSSFRHGSK-MASASTTDHARH--GSPR-HRDSGLLDSLGRFFG--	44
Dogfish	-----ASATTS DHAKQAGGAHSRQRDSGLLDQLGKLFGE	35
Bovine	SDRGAPKRGSGKDGHHAAARTTHYGSLPQKAQGH----RPQDENPVVHFFK	90
Rabbit	SDRGAPKRGSGKD--HAARTTHYGSLPQKS-GH----RPQDENPVVHFFK	89
Human	GDRGAPKRGSGKDSHHPARTAHYGSLPQKS-HG----RTQDENPVVHFFK	91
Snake	GDKHYPR---GKDYHHAARVSHVSSLPHPR-----YVDENPVVHFFR	
Anole	GDKHVPRR--GKDYHHAARASHVSSLPHQRSQR---GHYVDDNPVVHFFK	90
Gekko	GDKHVPRR--GKDYHHAARVSHVSSLPHPRSQHGRYGRYVDDNPVVHFFK	93
Chicken	GDRHVPRRGFGKDIH-AARASHVGSIPQ-RSQH---GRPDDNPVVHFFK	89
Dogfish	GSRKVPEK--GKE--PATRSVLMAPTTHKAHQG--ARRQTDDSPVVHFFK	79
Bovine	NIVTPRTPPPSQGKGRGLS--LSRFSWGAEGQKPGFGYGGGRASDYKSAHK	138
Rabbit	NIVTPRTPPPSQGKGRGTV--LSRFSWGAEGQKPGFGYGGRAADYKSAHK	137
Human	NIVTPRTPPPSQGKGRGLS--LSRFSWGAEGQRPFGFGYGGGRASDYKSAHK	139
Snake	NIVSPR-----AKGRGLS--FSRFSWGGETHKPGYGSG-KFYEHKYAHK	
Anole	SIVSPRSTPPMQAKGRGSP--FSRFSWGGETHKPGYGSG-KFFER---QK	134
Gekko	NIVSPRTPPPMQAKGRGLS--FSRFSWGGETPKHGYGSG-KFYEHKYAHK	140
Chicken	NIVSPRTPPPMQAKGRGLS--LTFRFSWGGEHKGPGYGSG-KFYEHKSAHK	136
Dogfish	NMMSPKKAPVQQAKSGASRAITKFIWGTGQRAHYGAAGSSKSK-----D	125
Bovine	GLKG--HDAQGTLSKIFKLGGDRDSR----SGSPMARR	169
Rabbit	GLKG--ADAQGTLSRLFKLGGDRDSR----SGSPMARR	168
Human	GFKG--VDAQGTLSKIFKLGGDRDSR----SGSPMARR	170
Snake	GHKGYRYDGQGTLSK---LGSSGSRPGSR-----	
Anole	GQKG--YDAQGTLSKIFKLGGSSGSRPGSRSGSPVARR	169
Gekko	GHKGYHYDGQGTLSKLFKLGGSSGSRPGSRHGSPIARR	177
Chicken	GHKGSYHEGQGTLSKIFKLGGSSGSRPGSRSGSPVARR	173
Dogfish	GFRG--RRDGSGLSSFFKMGKKGEG-----SPARR-	154

Sequence Coverage:
 150/177=85%(Gekko)
 150/169=89%(Anole)

Figure 8. Rattlesnake MBP partial sequence and PTMs matched to other species

The rattlesnake MBP sequence matches best with gekko or anole MBP sequence with only a few amino acid substitutions. A total of 150 AAs has been interpreted which covers 85% of gekko and 89% of anole MBP sequence respectively. PTMs identified in rattlesnake are shown in bold.

Table 1

Peptides with PTMs found in Bovine MBP isomers

Start	End	Sequence	PTM site information	Mascot ion score	Mass Error (ppm)	m/z ratio	Charge	Present in charge isomers
1	20	AAQKRPSQRSKYLASASTMD	Acetyl (N-term)	41	4	746.7175	+3	C1,C2,C3
10	23	SKYLASASTMDHAR	Oxidation (M19)	93	2	777.3734	+2	C1,C2,C3,C8a,C8b
10	23	SKYLASASTMDHAR	Oxidation (M19)	60	2	518.5853	+3	C1,C2,C3,C8a
12	23	YLASASTMDHAR	Oxidation (M19)	50	4	669.8116	+2	C1,C2,C3,C8b
12	23	YLASASTMDHAR	Oxidation (M19)	40	0	446.8751	+3	C1,C2,C3
12	29	YLASASTMDHARHGFLPR	Deimination (R23)	40	1	677.6674	+3	C8b
12	29	YLASASTMDHARHGFLPR	Oxidation(M19), Deimination(R23)	38	1	682.9996	+3	C8b
42	51	FFGSDRGAPK	Deimination (R47)	55	1	541.7674	+2	C8b
91	104	NIVTPRTPPPSQGGK	Deimination (R96)	56	3	746.9118	+2	C2,C3,C8a,C8b
91	104	NIVTPRTPPPSQGGK	Phosphorylation (T97)	32	3	786.4043	+2	C2,C3
91	104	NIVTPRTPPPSQGGK	Deamidation (Q102)	31	1	746.9102	+2	C2,C3,C8a,C8b
105	112	GRGLSLSR	Methylation (R106)	39	1	430.2590	+2	C1,C2,C3,C8a,C8b
105	112	GRGLSLSR	Dimethylation (R106)	31	1	437.2669	+2	C1,C2,C3,C8a,C8b
113	129	FSWGAEGQKPGFGYGGR	Deamidation(Q120), Acetyl(K121)	31	6	922.4236	+2	C3
159	168	DSRSGSPMAR	Deimination(R161)	24	-1	532.7426	+2	C3