Published in final edited form as:
Anal Chem. 2007 March 1; 79(5): 2007–2014. doi:10.1021/ac061606u.

Selective Enrichment and Fractionation of Phosphopeptides from Peptide Mixtures by Isoelectric Focusing after Methyl Esterification

Chong-Feng Xu[1], **Huaibin Wang**[2], **Daming Li**[3], **Xiang-Peng Kong**[2], and **Thomas A. Neubert**[1,*]

1 *Department of Pharmacology and Skirball Institute of Biomolecular Medicine, NYU School of Medicine, New York, NY 10016*

2 *Department of Biochemistry, NYU School of Medicine, New York, NY 10016*

3 *LITEC Systems Corporation, New York, NY 10006*

Summary

We have developed a new strategy to enrich and fractionate phosphopeptides from peptide mixtures based on difference in their isoelectric points (pIs) after methyl esterification. After isoelectric focusing (IEF) of a methylated tryptic digest of a mixture of α-S-casein and β-casein, phosphopeptides were selectively enriched at acidic and neutral pHs while nonphosphopeptides left the focusing gel because their pIs are higher than the upper limit of the immobilized pH gradient (IPG). We wrote a web-based program, pIMethylation, to predict the pIs for peptides with and without methyl esterification. Theoretical calculations using pIMethylation indicated that methylated phosphopeptides and non-phosphopeptides can be grouped based on the number of phosphate groups and basic residues in each peptide. Our IEF results were consistent with theoretical pIs of methylated peptides calculated by pIMethylation. We also showed that 2, 6-dihydroxyacetophenone (DHAP) is superior to 2, 5-dihydroxybenzoic acid (DHB) as a matrix for MALDI Q-TOF MS of methylated phosphopeptides in both positive and negative ion modes.

Keywords

isoelectric focusing; methyl esterification; phosphorylation; phosphopeptide; enrichment; isoelectric point; mass spectrometry

Introduction

Reversible phosphorylation of proteins plays a key role in many biological processes, including cellular signal transduction, molecular recognition and interaction, and cell differential and proliferation. Mass spectrometry (MS) has become an effective tool for characterizing protein phosphorylation sites. Mass spectrometry is especially useful for high throughput analysis: Hundreds or thousands of MS/MS spectra of phosphopeptides can be acquired in a single LC-MS run. However, performing a comprehensive phosphoproteomics analysis is still challenging for the following reasons: a) proteins are present in cells over a wide range of abundances; b) the stoichiometry of phosphorylation at each site can be very low and dynamic; and c) the ionization efficiencies of phosphorylated and nonphosphorylated peptides in MALDI

* To whom correspondence should be addressed. Tel: (212) 263-7265. Fax (212) 263-8214. E-mail: neubert@saturn.med.nyu.edu.

or ESI can differ significantly, leading to potential ion suppression of phosphopeptides. As a result, an enrichment of phosphopeptides from peptide mixtures is often required for phosphorylation analysis of either single proteins or complex protein mixtures.

Within the past five years, many highly specific enrichment methods have been developed, such as immobilized metal affinity chromatography (IMAC) after methyl esterification of peptide mixtures[1] and improved TiO_2, ZrO_2 and $Al(OH)_3$ chromatography[2–4]. These methods have greatly reduced nonspecific binding of non-phosphorylated acidic peptides. However, phosphoproteomics research is still challenged by the complexity of biological samples and wide dynamic range of protein concentrations and phosphorylation stoichiometries. Therefore, methods that can effectively fractionate the enriched phosphorylated peptides prior to reverse phase liquid phase chromatography-mass spectrometry analysis can be useful. A method that could perform the enrichment and fractionation at the same time would be especially useful.

Recently, isoelectric focusing (IEF) on an immobilized pH gradient (IPG) gel has been demonstrated to be very effective for fractionating peptides mixture prior to μLC-ESI MS analysis, and can lead to more protein identifications than optimized off-line strong cation exchange (SCX) chromatography[5, 6]. Since IEF separates peptides based on their pIs, information about peptide pIs obtained during separation can be compared to predicted pIs based on amino acid sequences as a constraint to improve confidence in protein identifications in LC-MS/MS experiments[7]. IEF of peptides mixture therefore can be thought of as an "information added" pre-fractionation approach for shotgun proteomics[8].

Peptide pIs decrease after phosphorylation due to the replacement of neutral hydroxyl groups on serines, threonines and tyrosines by negatively charged phosphate groups. Previous work has attempted to enrich phosphopeptides from peptides mixture based on their lower pIs. For example, capillary electrophoresis has been used for the separation of phosphopeptides from protein digests dissolved in acidic buffer[9]. Strong cation exchange chromatography also has been used to enrich phosphopeptides which generally elute in early fractions[10]. IEF also has been used[11]. Christoph Turck and colleagues have used IPG gels to fractionate peptides from digests of AtT-20 cells and found that phosphopeptides were enriched in the pH range of 3.5–4.5 (personal communication). Not surprisingly, the selectivity of enrichments using the above methods suffer from nonspecific enrichment of non-phosphorylated acidic peptides that contain multiple Asp/Glu resides in their sequences.

Previously, we have used methyl esterification to increase specificity by converting carboxylic acids to methyl esters for the selective detection of phosphopeptides in negative ion mode by MALDI Q-TOF MS[12]. Here we use IEF for the selective enrichment and fractionation of phosphopeptides after methyl esterification because methylated phosphopeptides have much lower pIs than methylated non-phosphopeptides. We also wrote and used a web-based program, pIMethylation, to predict the pIs for peptides with and without methyl esterification to confirm the theoretical basis for this work.

Experimental section

Reagents

2, 6-dihydroxyacetophenone (DHAP), 2, 4, 6-trihydroxyacetophenone (THAP), 2, 5-dihydroxybenzoic acid (DHB), diammonium hydrogen citrate (DAHC), acetyl chloride, ammonium bicarbonate, formic acid (FA), trifluoroacetic acid (TFA), phosphoric acid (PA), n-octyl-glucoside, bovine serum albumin (BSA), α-S-casein, β-casein, mono-phosphopeptide and tetra-phosphopeptide from β-casein were purchased from Sigma-Aldrich (St. Louis, MO). Sequencing grade modified trypsin was purchased from Promega (Madison WI). HPLC grade water, acetonitrile (AcN), and anhydrous methanol were purchased from Fisher Scientific

(Hanover Park, IL). IPG strips and Bio-Lyte ampholytes were purchased from Bio-Rad (Hercules, CA). μ-C18 ZipTips were purchased from Millipore (Billerica, MA)

Preparation of methyl esterified peptide mixtures

Tryptic digests of proteins were prepared by dissolving each protein in 25 mM ammonium bicarbonate, denaturing at 95 °C for 10 min, and treating with trypsin at 40:1 (w/w) at 37 °C for 6h. We prepared two peptide mixtures from the above protein digests. Peptide mixture 1 was used for the IEF experiment and contained 8 pmol each of tryptic digests of α-S-casein and β-casein. Peptide mixture 2 was used for the optimization of MALDI matrix, which contained two standard phosphopeptides from β-casein, α-S-casein tryptic digests and BSA tryptic digests at 200 fmoles each.

Methyl esterification of tryptic peptides followed the same protocol described in our previous work[7]. This protocol slightly modified the protocol of Ficarro et al.[1] by using more methanolic HCl and a longer incubation time to achieve nearly complete methyl esterification. Methanolic HCl solution was prepared by the dropwise addition of 160 μl of acetyl chloride to 1 ml of dry methanol[1]. The protein tryptic digests were lyophilized and redissolved in 50 μl of 2 M methanolic HCl regent. Methyl esterification was allowed to proceed for 2–3 h at room temperature.

Isoelectric focusing and extraction of methylated tryptic digests of proteins

IEF of methylated digests were performed with a Bio-Rad Protean IEF cell using 7 cm IPG strips exhibiting linear pH gradients of 3 to 10, 3 to 8.5 and 6.3 to 8.3. The pH 3–8.5 strips were made by cutting 1 cm off the basic end of pH 3–10 strips. 8 pmol methylated digest from peptide mixture 1 were loaded onto the gels either during or after rehydration. To facilitate the direct MS analysis of peptides in preliminary experiments, HPLC grade water alone was used to rehydrate the IEF gels. However, better focusing was achieved using a 0.2% ampholyte solution as rehydration buffer. Detergents, reducing agents and denaturing reagents were not used in the experiment. IEFs were performed at 3500V or 4000V for 1 to 2 hrs with maximum current of 50 μA. Gel strips were cut into 5–7 pieces after IEF.

60 μL 50%AcN in 0.1% TFA was used to extract methylated peptides from each piece of gel. The gel immersed in the extraction buffer was vortexed using a thermo mixer at 1000 rpm for 15 min, and then sonicated in a water bath for another 15 min. After centrifugation, the supernatant fluids were collected and dried using a vacuum centrifuge, and redissolved in 0.1% TFA solution for clean-up using μ-C18 ZipTips. One-fourth of each extract was analyzed by MALDI Q-TOF MS using DHAP as matrix.

Matrix selection for the detection of phosphopeptides with and without methyl esterification

10–100 mM DHAP or 100–300 mg/mL DHB were prepared by dissolving the matrices in 50% AcN and 0.1% TFA. DAHC at final concentrations of 0–50 mM was added to the DHAP matrix solution. Samples were dissolved in peptide dissolving solution (PDS) consisting of 5% FA, 0.1% TFA, 2% PA and 0.1% n-octyl-glucoside[8]. The original PDS was further diluted with HPLC grade water to form a series of solutions with concentrations of 6.25%, 12.5%, 25%, 50% and 100% PDS. Peptide mixture 2, a mixture of 200 fmol/μl (final concentration) methylated or non-methylated tryptic peptides from BSA, α-s-casein, mono- and tetra-phosphopeptide from β-casein, was dissolved in different concentration of PDS solution, then mixed with matrix solution at a 1:1 ratio. Two μl of the mixtures containing 200 fmoles of the peptides were spotted onto a MALDI target plate and dried in air. Both positive and negative ion mode mass spectra were acquired using MALDI Q-TOF MS. In each spectrum, the ion intensities of five phosphopeptides, including 2 mono-phosphopeptides, 1 di-phosphopeptide,

1 tetra-phosphopeptide and 1 penta-phosphopeptide, were summed for comparing the ion intensities using different matrix compositions.

MALDI Q-TOF and MALDI TOF mass spectrometry

A Waters-Micromass (Manchester, UK) MALDI Q-TOF Ultima mass spectrometer operated in either positive or negative ion mode was used in the experiment. Ions were generated by irradiation of analyte/matrix deposits by a nitrogen laser at 337nm with fixed incident energy. Mass spectra were acquired using the same number of shots for each experiment. To avoid bias in spectral acquisitions, the laser automatically fired in either a “line” or “circle” pattern, and each acquisition in the matrix selection experiment was repeated 3 times and intensities were averaged. Masslynx 4.0 software was used to smooth and subtract background from the raw data.

A Waters-Micromass Tof Spec 2E MALDI TOF mass spectrometer operated in positive ion reflectron mode was used to acquire data for all two hour focusing experiments. A nitrogen laser (λ=337 nm, pulse time 39ns) was used. The incident laser energy was tuned to the minimum effective setting to avoid neutral loss of phosphopeptides and achieve better resolution for peptides. Three hundred laser shots were summed into a single mass spectrum. Masslynx 4.0 software was used to smooth and subtract background from the raw data.

Results and Discussion

IEF of methylated digests of proteins

Our goal was to develop a new strategy to enrich and fractionate phosphopeptides present in complex mixtures of phosphorylated and nonphosphorylated peptides. We did this by exploiting the acidic properties of phosphopeptides to resolve them from nonphosphopeptides in IEF experiments after methyl esterification to reduce contamination by acidic nonphosphopeptides. In preliminary experiments using pH 3–10 gel strips and a rehydration time of 12 h, we observed many hydrolysis products of nonphosphopeptide methyl esters in low pH fractions that contained the majority of the phosphopeptides. As shown in Supplementary Figure S1, phosphopeptides were enriched in the pH 7–8 fraction. However, we observed hydrolysis products of methyl esters of nonphosphopeptides in the pH 6–7 fraction. Later we found that hydrolysis of methyl esters was significantly reduced by loading the sample using the “well loading” method after rehydration of the gel in 0.2% pH 3–10 Bio-Lyte solution for 12 h. We also found that using a pH 3–8.5 gel greatly reduced methyl ester hydrolysis, which is favored at pH 9.0 and above. In the pH 3–8.5 gel most of the phosphopeptides were present in fractions with pH 6.5–7.5 and 7.5–8.5, with little contamination from nonphosphopeptides (data not shown). This observation agreed well with our calculations showing that methyl esterified mono-phosphorylated peptides have pIs close to 7.0, and that multiply phosphorylated peptides have even lower pIs.

Figure 1 shows positive ion MALDI Q-TOF MS spectra of a tryptic digest of methyl esterified α- and β-casein before and after IEF, using conditions that minimize hydrolysis of methyl esters. We observed that mono-phosphopeptides were predominantly in the pH 6.7–7.1 and pH 6.3–6.7 fractions. The masses of these protonated peptides after methyl esterification were 1538.7 (Methionine oxidation product of 1522.7), 1716.9, 1888.9 and 2159.9. We also detected one multiply phosphorylated peptide at mass 3234.4 (4 phosphates, calculated pI 1.68) in pH 7.5–7.9, pH 7.1–7.5, and pH 6.7–7.1 fractions, presumably due to anomalous migration through the strip or poor focusing of this peptide. In the pH 7.9–8.3 and 7.5–7.9 fractions, we detected one mono-phosphopeptide containing a missed tryptic cleavage at protonated mass 2009.0.

In a preliminary experiment where no ampholyte was used and focusing time was limited to one hour to minimize methyl ester hydrolysis (Fig, S2), in addition to the above phosphopeptides, we were able to observe several additional multiply phosphorylated peptides at mass 2025.9 (2 phosphates, calculated pI 3.84), 3216.3 (4 phosphates, calculated pI 5.75) in the lowest pH fraction (6.3–6.7). In the pH 7.9–8.3 and 7.5–7.9 fractions, we detected two mono-phosphopeptides containing one missed cleavage at protonated masses 1650.8 and 2009.0. In the pH 7.9–8.3 fraction, we also observed two peptides having lost phosphate groups at protonated masses 1570.8 and 1928.1. The observation of additional, low pI phosphopeptides here was most likely due to incomplete focusing. For the same reason, the peptide at protonated mass 2009.0 appeared in several different pH fractions.

All together, we observed 9 phosphopeptides with a total of 14 phosphorylation sites. The observed and predicted (using pIMethylation, see below) pIs of these phosphopeptides are shown in Table 1. Overall, the distribution of the phosphopeptides in the IEF gel after methyl esterification is in agreement with pIs calculated by pIMethylation. However, under incomplete focusing conditions several of these phosphopeptides were present in the fraction of nominal pH 6.3–6.7 even though their calculated pIs are lower than 6.3 because this fraction was the lowest pH fraction of the gel.

An important issue in our IEF experiments is the optimization of methyl ester formation. Peptides with Asn/Gln residues may undergo a side reaction of deamidation and methyl esterification[12, 13]. However, methylation of Asn/Gln does not affect the pIs of peptides to a significant degree. On the other hand, incomplete methylation did affect our results by allowing contamination of phosphopeptides by acidic peptides, so we optimized conditions to maximize methylation while allowing the Asn/Gln methylation side reaction to proceed to some extent. The efficiency of the methylation reaction was improved by using dehydrated methanol, increasing the amount of methanolic HCl, using longer incubation times, and performing the methylation reaction twice when necessary to achieve complete methylation of acidic residues.

Matrix selection for the detection of methylated phosphopeptides by MALDI-MS

DHB, DHAP and THAP have been shown to be very useful matrices for the analysis of phosphopeptides by MALDI-TOF MS[12, 14, 15]. In addition, phosphoric acid (PA) is a popular additive to improve the detection of phosphopeptides[14]. We investigated the use of DHB and DHAP for the detection of methylated and non-methylated phosphopeptides in both positive and negative ion modes of MALDI QTOF MS. THAP was not investigated because of its similarity to DHAP. Our optimized DHAP matrix solution consisted of 50mM DHAP, 20mM DAHC, 12.5% PDS, 50% AcN and 0.1% TFA. Our optimized DHB matrix solution was 100 mg/mL DHB, 12.5% PDS, 50% AcN and 0.1% TFA (data not shown). We found that 12.5% PDS (peptide dissolving solution, consisting of 5% FA, 0.1% TFA, 2% PA and 0.1% n-octyl-glucoside), which leads to a final matrix concentration of 0.25% PA, is a very effective additive for the detection of multiply phosphorylated peptides using MALDI Q-TOF MS. However, PA is not a necessary additive for the detection of monophosphopeptides. We found that for non-methylated peptides, DHB was approximately 1.5 times more sensitive than DHAP in positive ion mode MALDI Q-TOF MS, although their sensitivities in negative ion mode were similar. In Fig. 2d we observed several peaks with mass 154 Da greater than some of the known phosphopeptides. MS/MS experiments confirmed that these were DHB adducts of the methylated phosphopeptides. These adduct peaks were only observed for the methylated peptides in negative ion mode using DHB as the MALDI matrix.

Although DHB is effective for the detection of non-methylated phosphopeptides, Fig. 2 shows that DHAP is approximately 50% more sensitive in positive ion mode and over three times as sensitive in negative ion mode for methylated phosphopeptides compared to DHB. We think the difference in the matrix selectivity of peptides before and after methyl esterification is due

to the increased hydrophobicity of peptides after methyl esterification, as DHAP is less polar than DHB. The match of relative polarity between sample and matrix can be critical in MALDI MS analysis[16]. We have observed that the choice of matrix for MALDI analysis can also be influenced by other factors such as laser intensity and wavelength, and is often instrument-specific.

Prediction of peptide isoelectric points by pIMethylation

To optimize IEF conditions and better understand our results, we felt it necessary to calculate the theoretical isoelectric points of the methyl esterified peptides we analyzed based on their amino acid sequences. Various algorithms have been proposed to predict the pIs of peptides, and excellent agreement between the theoretical and experimental pIs has been achieved for unmodified peptides[17, 18]. A list of available programs that can predict the pIs of peptides is shown in Table 2. Three out of the six programs can predict the pIs of phosphorylated peptides. Unfortunately, none of these programs consider methyl esterification as a modification, so none of them can directly predict the pIs of methylated peptides. In addition, because different pKa values of phosphate group have been used for the predictions, significant variations in the predicted pIs of phosphopeptides have been obtained by these programs.

We wrote a web-based, open source software program, pIMethylation, to calculate the pIs of peptides with and without methyl esterification (http://saturn.med.nyu.edu/~xucf/cgi-bin/pi.cgi). pIMethylation considers the contributions of phosphates, N-terminal amino and C-terminal carboxyl groups, and the side chains of Arg, Lys, His, Tyr, Cys, Glu and Asp residues to the pIs of non-methylated peptides. The pKas (pKa of the side chain) of these chargeable functional groups (except phosphates) were taken from the paper of Bjellqvist et al.[17]. For calculating the pIs of methylated peptides, we removed the contributions to pIs from Asp and Glu residues and C-termini of the sequences. The source code for pIMethylation is freely available by contacting the corresponding author.

To evaluate the performance of pIMethylation, we compared it with Scansite, one of the online programs listed in Table 2. For several dozen random nonphosphorylated and nonmethylated peptide sequences we tested, pIMethylation gave the same prediction results as Scansite (data not shown). In Table 1 we compared the pI predictions using Scansite and pIMethylation for the major tryptic phosphopeptides from β-casein and α-S-casein. For mono-phosphopeptides, the predicted pIs given by the two programs were very close. However, for multiply phosphorylated peptides, significant differences in the prediction results were observed. These differences were due to the use of different pKa values for the phosphate group in the two programs. For Scansite, the pKa of phosphate is taken from that of phosphoric acid, which is 2.12 for pKa1 and 7.21 for pKa2. However, the phosphate groups on peptides are more acidic than phosphoric acid[2, 19]. pIMethylation allows users to define the pKa values of phosphate in the web-based interface, and we recommend using 1.2 and 6.5 as the pKas of peptide phosphates based on experimental values[19]. Because Scansite and other programs do not predict the pIs for methylated peptides, the predicted pIs of methylated phosphopeptides were only given by pIMethylation.

Phosphorylated and non-phosphorylated peptides can be separated based on calculated pI after methyl esterification

Figure 3 shows a typical output from the pIMethylation program, depicting the predicted pIs of the mono-phosphorylated tryptic peptide from β-casein (FQpSEEQQTEDELQDK, MW 2060) and its non-phosphorylated version. Without methylation, the pIs for the phosphopeptide and its nonphosphorylated version are 3.34 and 3.77, respectively, which is a subtle difference. However, after methyl esterification the pI for the phosphopeptide is 7.00, and the pI for the nonphosphorylated version is greater than 13. This dramatic difference in pIs suggests that

phosphopeptides and non-phosphopeptides can be efficiently separated based on pI differences after methyl esterification.

The pIs of peptides are mostly determined by their acidic and basic amino acid residues. For methylated peptides, presence of phosphorylation and basic residues (Lys, and Arg) are the major determinants of pIs. Cys (pKas=9.0), Tyr (pKas=10.0), and His (pKas=6.0) can also affect the pIs of methylated peptides, but the presence of other amino acid residues including Asp and Glu do not significantly affect the pIs of methylated peptides. Using pIMethylation, we predicted the pIs for methylated tryptic peptides with different amino acid compositions and numbers of phosphate groups, and the results are illustrated in Figure 4. Figure 4a shows the predictions of pIs for peptides with 0 missed tryptic cleavages (i.e. single lysine or arginine), 0 histidine residues, and one to five phosphorylated amino acid residues to be 7.00, 3.84, 1.50, 1.20 and 1.02, respectively. For mono-phosphorylated peptides, the presence of histidines does not cause significant increase in pI. However, for multiply phosphorylated peptides, one histidine residue in the sequence is equivalent to removal of one phosphorylation. Monophosphopeptides with one missed cleavage have slightly basic pIs; the pIs for peptides containing two lysines, one lysine and one arginine, or two arginines are 8.61, 8.77 and 9.62, respectively. For monophosphopeptides with one missed cleavage, the presence of histidine does not affect the pIs significantly. In summary, the pIs of methylated tryptic phosphopeptides can be classified based on the number of phosphorylations, and this classification can be affected by the presence of additional basic residues in the sequence.

Methylated nonphosphopeptides have very high pIs, greater than 13.0 if there are no tyrosines or free cysteines in the sequences. Nonphosphopeptides with 1–3 tyrosines have pIs between 9.70 and 11.0, and nonphosphopeptides with 1–3 cysteines have pIs in the range of 8.7 –10.5. The pI distribution of nonphosphorylated peptides with 1–3 Tyr or Cys residues are illustrated in Fig. 4b. Because the frequencies of cysteine and tyrosine in vertebrates are both 3.3%, there are few tryptic peptides, which normally contain less than 30 amino acids, with more than three Tyr/Cys present in their amino acid sequences.

As indicated in Fig. 4, phosphopeptides have acidic or nearly neutral pIs after methyl esterification, and non-phosphopeptides generally have very basic pIs. The threshold for differentiating phosphopeptides from non-phosphopeptides is approximately pH 8.7. In theory, if all free cysteines were removed by S-carbamidomethylation, the minimum pI of nonphosphorylated methylated peptides would be increased to pH 9.7 for even more efficient separation of phosphorylated and nonphosphorylated peptides. Non-phosphopeptides can be removed based on this threshold and phosphopeptides can be sorted according to the number of phosphorylations and basic residues. The unique pattern of pI distributions of peptides after methyl esterification enables the efficient separation of phosphopeptides from non-phosphopeptides by IEF.

In addition to gel-based IEF, several other techniques rely on differences in pIs to separate proteins or peptides, such as liquid phase IEF, capillary IEF, and free flow electrophoresis. More recently, chromatofocusing has been used as a column-based method for separating complex mixture of proteins according to pI. OFFGEL electrophoresis is another recent advance in separation technology that fractionates proteins or peptides based on their pIs[20]. It is very likely that any of these methods could be used effectively for the enrichment and fractionation of phosphopeptides after methyl esterification. Methyl esterification also has been used to increase the selectivity of IMAC for the enrichment of phosphorylated peptides[1] and successfully applied to the selective detection of phosphopeptides in negative ion mode MALDI-MS[12]. Indeed, calculations of pIs for methylated and nonmethylated phosphopeptides by pIMethylation help explain the increase in specificity for these methods after methylation of the peptides. Our method of phosphopeptide enrichment by IEF is compatible with these

techniques. We believe a combination of these methods will enable more comprehensive analyses of phosphoproteomes.

Summary

We have demonstrated the use of IEF for selective enrichment and fractionation of phosphopeptides from protein digests after methyl esterification of tryptic peptides. We have also written a freely available web-based program, pIMethylation, to calculate the isoelectric points of the methyl esterified phosphopeptides. Our theoretical pI calculations agreed with the isoelectric focusing results. In addition, we have shown that DHAP can be a more effective MALDI matrix than DHB for the detection of methyl esterified peptides by MALDI-TOF MS.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Dr. Guoan Zhang and Joshua Hart for helpful discussions. We gratefully acknowledge support from NIH Shared Instrumentation Grant S10 RR017990 and S10 RR14662 to T.A.N, NIH NCI core grant P30 CA016087 to Steven Burakoff, and NIH grants R01 DE013686-06 to Moosa Mohammadi and DK52206 to X.-P.K.

References

1. Ficarro SB, McCleland ML, Stukenberg PT, Burke DJ, Ross MM, Shabanowitz J, Hunt DF, White FM. Nature Biotechnology 2002;20:301–305.
2. Larsen MR, Thingholm TE, Jensen ON, Roepstorff P, Jorgensen TJD. Molecular & Cellular Proteomics 2005;4:873–886. [PubMed: 15858219]
3. Kweon HK, Hakansson K. Analytical Chemistry 2006;78:1743–1749. [PubMed: 16536406]
4. Wolschin F, Wienkoop S, Weckwerth W. Proteomics 2005;5:4389–4397. [PubMed: 16222723]
5. Essader AS, Cargile BJ, Bundy JL, Stephenson JL. Proteomics 2005;5:24–34. [PubMed: 15672457]
6. Gan CS, Reardon KF, Wright PC. Proteomics 2005;5:2468–2478. [PubMed: 15880631]
7. Cargile BJ, Bundy JL, Freeman TW, Stephenson JL. Journal of Proteome Research 2004;3:112–119. [PubMed: 14998171]
8. Xie HW, Bandhakavi S, Griffin TJ. Analytical Chemistry 2005;77:3198–3207. [PubMed: 15889909]
9. Zhang HX, Zhang CJ, Lajoie GA, Yeung KKC. Analytical Chemistry 2005;77:6078–6084. [PubMed: 16159144]
10. Beausoleil SA, Jedrychowski M, Schwartz D, Elias JE, Villen J, Li JX, Cohn MA, Cantley LC, Gygi SP. Proceedings of the National Academy of Sciences of the United States of America 2004;101:12130–12135. [PubMed: 15302935]
11. Cantin GT, Venable JD, Cociorva D, Yates JR. Journal of Proteome Research 2006;5:127–134. [PubMed: 16396503]
12. Xu CF, Lu Y, Ma JH, Mohammadi M, Neubert TA. Molecular & Cellular Proteomics 2005;4:809–818. [PubMed: 15753120]
13. He T, Alving K, Feild B, Norton J, Joseloff EG, Patterson SD, Domon B. Journal of the American Society for Mass Spectrometry 2004;15:363–373. [PubMed: 14998538]
14. Kjellstrom S, Jensen ON. Analytical Chemistry 2004;76:5109–5117. [PubMed: 15373450]
15. Yang XF, Wu HP, Kobayashi T, Solaro RJ, van Breemen RB. Analytical Chemistry 2004;76:1532–1536. [PubMed: 14987115]
16. Hoteling AJ, Erb WJ, Tyson RJ, Owens KG. Analytical Chemistry 2004;76:5157–5164. [PubMed: 15373456]
17. Bjellqvist B, Hughes GJ, Pasquali C, Paquet N, Ravier F, Sanchez JC, Frutiger S, Hochstrasser D. Electrophoresis 1993;14:1023–1031. [PubMed: 8125050]

18. Shimura K, Zhi W, Matsumoto H, Kasai K. Analytical Chemistry 2000;72:4747–4757. [PubMed: 11028642]
19. Halligan BD, Ruotti V, Jin W, Laffoon S, Twigger SN, Dratz EA. Nucleic Acids Research 2004;32:W638–W644. [PubMed: 15215467]
20. Horth P, Miller CA, Preckel T, Wenz C. Molecular & Cellular Proteomics 2006;5:1968–1974. [PubMed: 16849286]

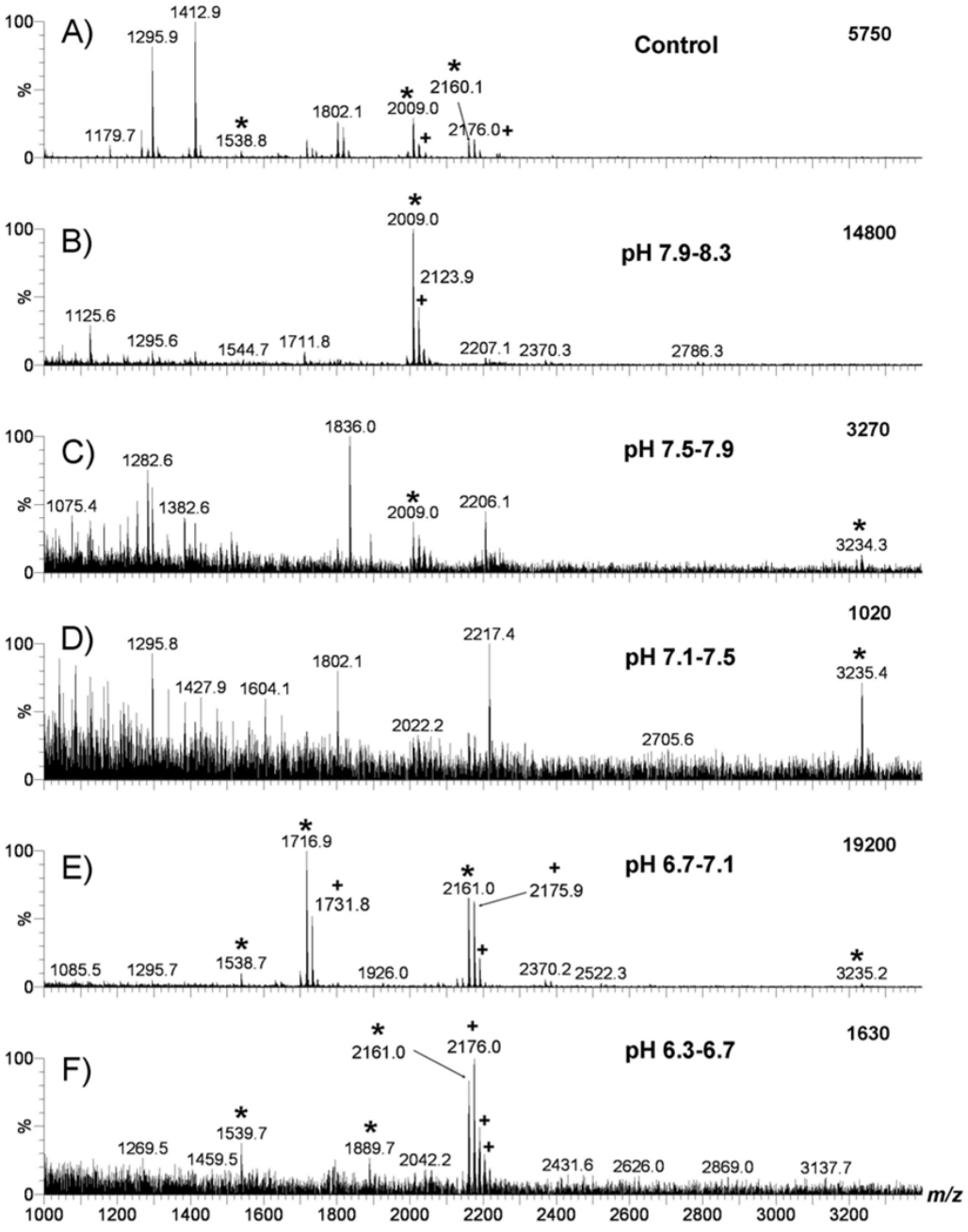


Figure 1.

IEF separation of methylated tryptic digests of α-casein and β-casein (1:1). Eight pmol of methylated digest was loaded onto the 7cm pH 6.3–8.3 linear pH gradient gel which had been rehydrated in 0.2% pH 6.3–8.3 Bio-Lyte buffer for 12 h. IEF was performed at 4000V for 120 min with a current limit of 50 μA. The gel strip was cut into 5 pieces after focusing, and 25% of each extract was analyzed by positive ion MALDI TOF MS. Fig. 1a shows the MS spectrum of the sample prior to IEF. Figs. 1b – 1f show the spectra of peptides extracted from each gel piece. Predominant phosphopeptide species are labeled (*), and methyl ester side products from deamidated Gln or Asn residues are labeled (+). Base peak (maximum) ion counts are shown in the upper right corner of each spectrum. In Fig 1e, the mass 1538.7 is the methionine

oxidation product of the peptide with protonated mass 1522.7. The sequences of the labeled peaks are shown in Table 1.

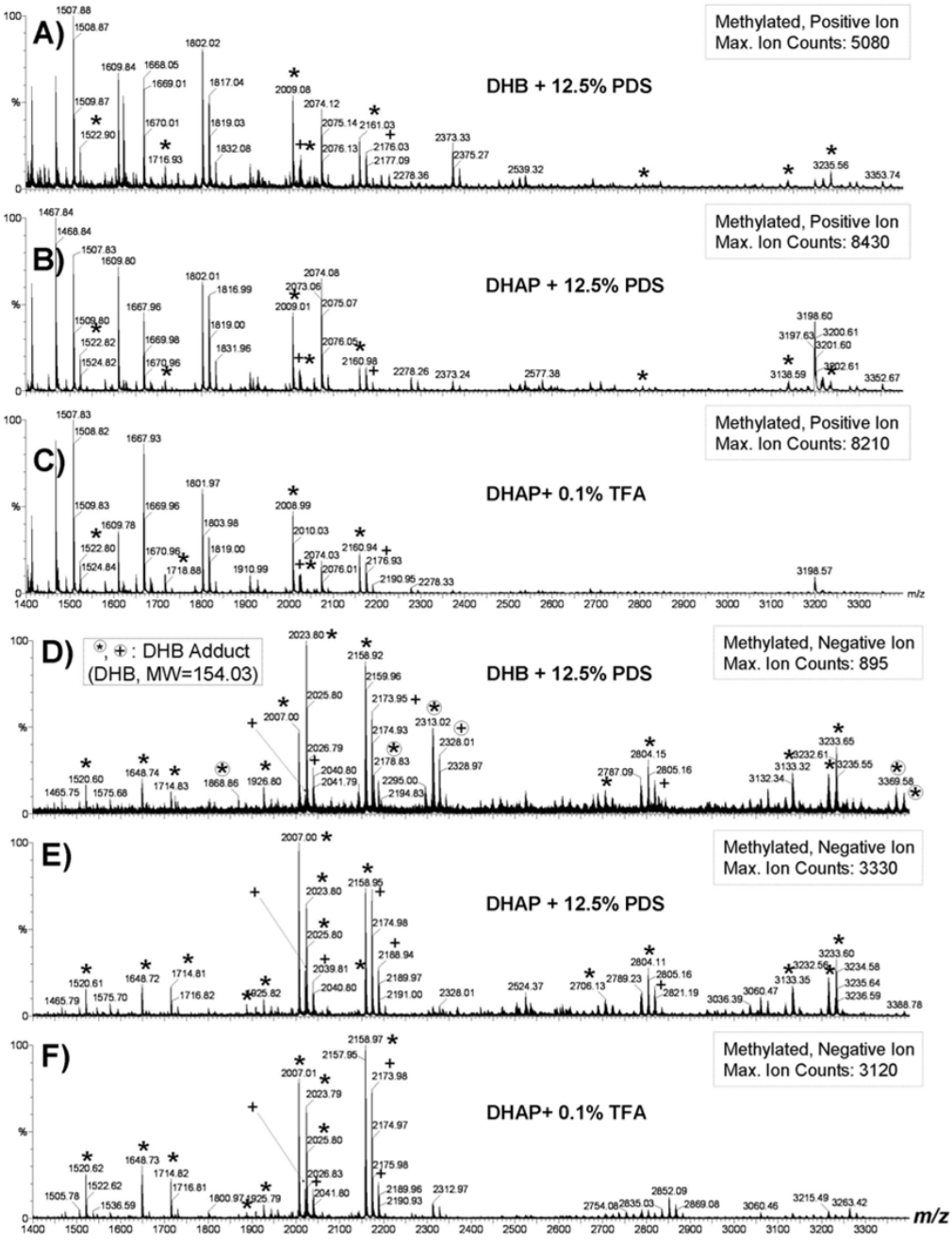
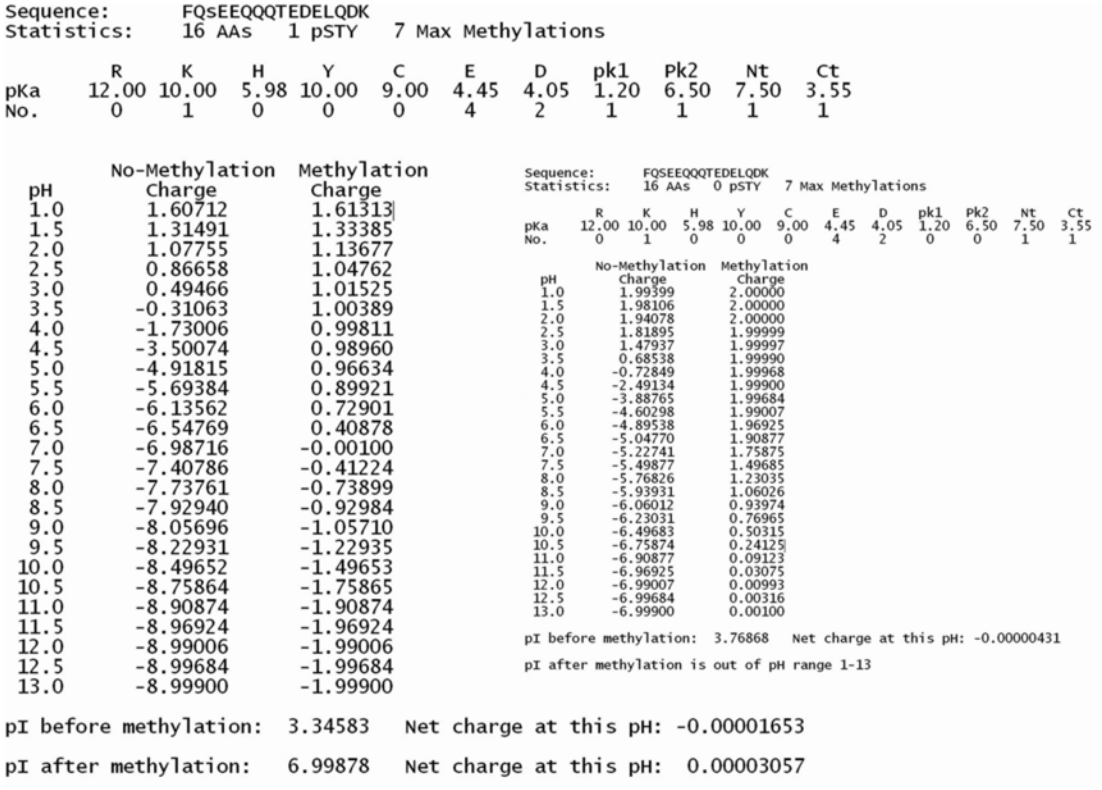


Figure 2.

Positive and negative ion MALDI Q-TOF MS of methylated peptides using different matrices and additives. The sample was a mixture of 200 fmol methylated tryptic peptides from BSA, α-S-casein, and the mono- and tetra- phosphopeptides from β-casein. Matrices were either DHAP (50mM DHAP+20mM DAHC in 50% AcN and 0.1%TFA) or DHB (100mg/ml DHB in 50% AcN and 0.1% TFA). The solvents for the methyl esters were either 0.1% TFA or 12.5% PDS (PDS, peptide dissolving solution: 5% FA + 0.1% TFA + 2% PA + 0.1% n-octyl-glucoside). Predominant phosphopeptide species are labeled (*), undermethylated (one or more carboxyl groups not methylated) are labeled (−), methyl ester side products of deamidated Q

or N residues are labeled (+). In Fig. 2d, DHB adducts of methylated phosphopeptides are circled. The amino acid sequences of labeled peptides are listed in Table 1.



```
Sequence:       FQsEEQQQTEDELQDK
Statistics:     16 AAs    1 pSTY    7 Max Methylations

          R     K     H     Y     C     E     D    pk1   Pk2    Nt    Ct
pKa   12.00 10.00  5.98 10.00  9.00  4.45  4.05  1.20  6.50  7.50  3.55
No.       0     1     0     0     0     4     2     1     1     1     1

          No-Methylation  Methylation
  pH          Charge          Charge
  1.0          1.60712         1.61313
  1.5          1.31491         1.33385
  2.0          1.07755         1.13677
  2.5          0.86658         1.04762
  3.0          0.49466         1.01525
  3.5         -0.31063         1.00389
  4.0         -1.73006         0.99811
  4.5         -3.50074         0.98960
  5.0         -4.91815         0.96634
  5.5         -5.69384         0.89921
  6.0         -6.13562         0.72901
  6.5         -6.54769         0.40878
  7.0         -6.98716        -0.00100
  7.5         -7.40786        -0.41224
  8.0         -7.73761        -0.73899
  8.5         -7.92940        -0.92984
  9.0         -8.05696        -1.05710
  9.5         -8.22931        -1.22935
 10.0         -8.49652        -1.49653
 10.5         -8.75864        -1.75865
 11.0         -8.90874        -1.90874
 11.5         -8.96924        -1.96924
 12.0         -8.99006        -1.99006
 12.5         -8.99684        -1.99684
 13.0         -8.99900        -1.99900

pI before methylation:  3.34583   Net charge at this pH: -0.00001653

pI after methylation:   6.99878   Net charge at this pH:  0.00003057
```

```
Sequence:       FQSEEQQQTEDELQDK
Statistics:     16 AAs    0 pSTY    7 Max Methylations

          R     K     H     Y     C     E     D    pk1   Pk2    Nt    Ct
pKa   12.00 10.00  5.98 10.00  9.00  4.45  4.05  1.20  6.50  7.50  3.55
No.       0     1     0     0     0     4     2     0     0     1     1

          No-Methylation  Methylation
  pH          Charge          Charge
  1.0          1.99399         2.00000
  1.5          1.98106         2.00000
  2.0          1.94078         2.00000
  2.5          1.81895         1.99999
  3.0          1.47937         1.99997
  3.5          0.68538         1.99990
  4.0         -0.72849         1.99968
  4.5         -2.49134         1.99900
  5.0         -3.88765         1.99684
  5.5         -4.60298         1.99007
  6.0         -4.89538         1.96925
  6.5         -5.04770         1.90877
  7.0         -5.22741         1.75875
  7.5         -5.49877         1.49685
  8.0         -5.76826         1.23035
  8.5         -5.93931         1.06026
  9.0         -6.06012         0.93974
  9.5         -6.23031         0.76965
 10.0         -6.49683         0.50315
 10.5         -6.75874         0.24125
 11.0         -6.90877         0.09123
 11.5         -6.96925         0.03075
 12.0         -6.99007         0.00993
 12.5         -6.99684         0.00316
 13.0         -6.99900         0.00100

pI before methylation:  3.76868   Net charge at this pH: -0.00000431

pI after methylation is out of pH range 1-13
```

Figure 3.

Prediction by pIMethylation of pIs for the mono-phosphopeptide in β-casein (FQpSEEQQTEDELQDK, MW=2060) and its non-phosphorylated version. The first line indicates the input sequence; the second line contains statistics about the number of amino acids, number of phosphorylations and maximum number of methyl esterifications; lines 4–6 list the chargeable functional groups in the sequence that were used to calculate the pI. Below line 8, the net charges of peptides with and without methylation were calculated for pH 1 to pH 13. The last two lines display the calculated pIs of the peptides with and without methylation. The inset shows the pI prediction for the non-phosphorylated version of this peptide.

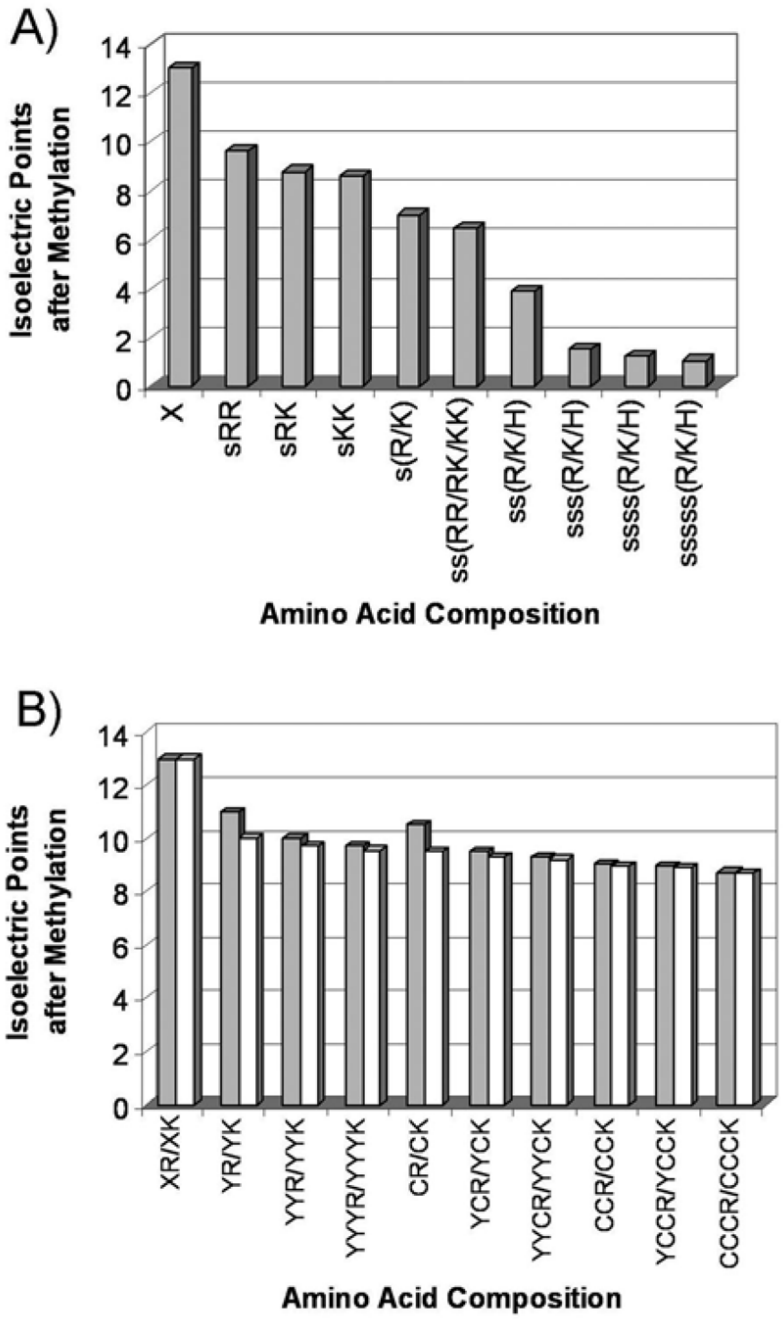


Figure 4.

Classification of phosphopeptides and non-phosphopeptides by pI after methyl esterification. Fig. 4a shows the pIs for methylated tryptic phosphopeptides as predicted by pIMethylation. X represents any amino acid sequence without Tyr, free Cys or phosphorylation; s represents pS, pT or pY. For the pI calculation of multiply phosphorylated peptides, adding an R/K/H residue is equivalent to the removal of one phosphorylation. Fig. 4b shows the prediction of the pIs for methylated tryptic non-phosphopeptides containing 0–3 Cys/Tyr residues and no missed tryptic cleavages (i.e. only one K or R). X, any amino acid sequence without Tyr, free Cys or phosphorylation.

Table 1

Prediction of the pIs for phosphopeptides using pIMethylation and Scansite

Protein	Seq. No.	Sequence	MW+H w/ o Methylation	Methylation	pI (w/o Methylation) Scansite	pI (w/o Methylation) pIMethylation	pI (Methylation) pIMethylation	[a] Observed pI (pH 6.3–8.3 gel)
β-casien	33–48	$FQ_pSEEQQQTEDELQDK$	2061.83	2159.94	3.37	3.34	7.00	6.7–7.1
	1–25	$RELEELNVPGEIVE_pSL_pS_pS_pSEESITR$	3122.27	3234.40	2.51	1.66	1.68	6.7–7.9
α-S1-casein	106–119	$VPQLEIVPN_pSAEER$	1660.80	1716.86	3.69	3.67	6.97	6.7–7.1
	104–119	$YLGEYLIVPN_pSAEER$	1832.83	1888.89	3.69	3.67	7.00	6.3–6.7
	43–58	$DIG_pSE_pSTEDQAMEDIK$	1927.69	2025.80	2.78	2.33	3.84	≤ 6.3–6.7[b]
	104–119	$YKVPQLEIVPN_pSAEER$	1951.95	2008.01	4.26	4.25	8.61	7.9–8.3
	59–79	$QMEAE_pSI_pS_pS_pSEEIVPN_pSVEQK$	2720.91	2805.00	1.93	1.02	1.02	N/A
α-S2-casein	138–149	$TVDME_pSTEVFTK$	1466.61	1522.67	3.60	3.58	6.66	6.7–7.1
	138–150	$TVDME_pSTEVFTKK$	1594.70	1650.76	4.14	4.13	8.14	7.5–7.9[b]
	1–24	$KNTMEHV_pS_pS_pSEESII_pSQETYKQEK$	3132.20	3216.29	3.59	3.52	5.75	≤ 6.3–6.7[b]

[a] Observed pI was the pH fraction that contained the majority of phosphopeptides in the 2h focusing experiment;

[b] phosphopeptides that were observed only in the 1 h focusing experiment.

Table 2

Publicly available tools for the prediction of pIs of proteins or peptides

pI Prediction Tools and their Links	pKa of Phosphates
1. Compute pI/Mw Tool http://us.expasy.org/tools/pi_tool.html	Not Considered
2. EMBL WWW Gateway to Isoelectric Point Service http://www.embl-heidelberg.de/cgi/pi-wrapper.pl	Not Considered
3. Protein Calculator v3.3 http://www.scripps.edu/~cdputnam/protcalc.html	Not Considered
4. Calculate Molecular Weight and Isoelectric Point http://scansite.mit.edu/calc_mw_pi.html	2.12, 7.21
5. ProMost (Protein Modification Screening Tool) http://proteomics.mcw.edu/promost/index.jsp	1.2, 6.5 (or 6.9)
6. ProteoMod (A New Tool to Quantitate Protein PTMs) *Proteomics 2004, 4, 1672–1683*	pS/pT: 2.12, 7.12 pY: 1.0, 7.0