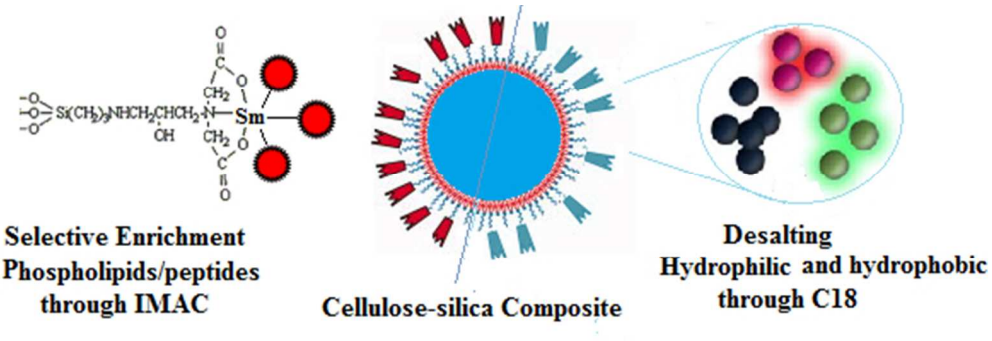
RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances



155x52mm (96 x 96 DPI)

New Cellulose-Silica Composite IMAC/C18 for Selective Enrichment of Phosphorylated Molecules and Improved Recovery of Hydrophilic Species

Muhammad Najam-ul-Haq[1]*, Fahmida Jabeen[1], Faiza Shafiq[1], Salman Sajid[1], Ambreen Saba[1]

[1]: Division of Analytical Chemistry, Institute of Chemical Sciences, Bahauddin Zakariya University, Multan 60800, Pakistan.

* Corresponding Author
Dr. M. Najam-ul-Haq
Institute of Chemical Sciences
Bahauddin Zakariya University
Multan 60800 Pakistan
Tel.: +92 306 7552653
Email: najamulhaq@bzu.edu.pk

Keywords: Cellulose-silica composite, Phosphopeptides, Phospholipids, Desalting, Prostate cancer, MALDI-MS.

Abstract

Cellulose and silica are the traditional quality sorbents with range of reported applications, particularly for the hydrophilic biomolecules. Considering those characteristics, cellulose and silica are brought into a composite form to take the benefits on a single platform. Choice of cellulose and silica is made to obtain complete hydrophilic composite, followed by its derivatization for the selective enrichment of phosphopeptides and as desalting material prior to the mass spectrometric analysis. The composite is characterized with FT-IR, EDX and SEM. As immobilized metal ion affinity chromatographic (IMAC) material, the cellulose-silica enriches phosphopeptides from complex mixture in which β-casein is spiked in de-phosphorylated HeLa cell extract. The inherent hydrophilic nature of composite gives higher selectivity up to 2000 times complexity level and sensitivity down to 1 femtomole for phosphopeptides. Other than the immobilized metal ions onto the IMAC composite, the role of base materials in the composite, i.e. cellulose and silica is also tested in the enrichment which helps thereof in optimizing the best sample preparation protocol. With the introduction of successive elution conditions, phospholipids and phosphopeptides are enriched and identified from egg yolk digest using single batch extraction method. The phosphopeptides are analyzed with MALDI-MS whereas newly designed gold nanoparticles and carbon based materials are used to analyze low molecular weight phospholipids through LDI-MS. Phosphopeptides are also enriched from complex serum digest and discussed in relevance to the prostate cancer. As RP material, it provides combination of hydrophilic composite and hydrophobic C18 chain. Efficient desalting of higher salt concentrations, improved recovery of hydrophilic peptides and the high sequence coverage in comparison to

commercial materials makes cellulose-silica C18 a better desalting tool with in-house economical synthesis.

Introduction

Affinity chromatography is developing along with the growing diversity in "Omics" technologies, such as genomics, proteomics and metabolomics[1,2]. The breakthrough development of selective affinity chromatography has enabled researchers to explore fields such as protein-protein interactions, post-translational modifications[3] and protein degradation that were not possible to be examined previously[4]. The coupling of reversed phase affinity chromatography with mass spectrometry has aided in the discovery of protein biomarkers[5,6]. Since past few decades polymer-inorganic composite materials have attracted interest in affinity chromatography, due to their synergistic and hybrid properties derived from individual components[7,8]. The properties of composites are not only the sum of individual contributions of both materials, but the role of inner interfaces can be predominant and provide interaction sites for various analytes[9].

The specificity is related to the immobilized surface functionalities, however the properties of support material must be chosen to limit the non-specific bindings. Supports which have little or no non-specific binding mimic the properties of aqueous mobile phase. Therefore, chemically inert support materials are hydrophilic in nature[10]. Chemical inertness of cellulose and silica enables selective binding of molecules of interest with little or no non-specific bindings. In addition, cellulose and silica are chemically stable under the normal operating conditions[11] and have been utilized individually in phosphoproteomics study[12,13]. Therefore they have enjoyed the attention of separation chemists for the development of selective enrichment tools in phosphoproteomics, glycoproteomics etc.

At clinical level and proteomic research centers, there is high demand of desalting materials. Commercial materials like ZipTip (Millipore), graphite spin columns (GC Tip, GL Science) or Eppendorf GELoader has been in use in recent times. However, loss of small, hydrophilic and some post-translationally modified peptides is still a limitation of commercially available desalting materials[14] and hence new materials are always in demand.

Knowing the uniqueness of cellulose and silica in proteomics, the present research is focused to synthesize cellulose-silica composite and to evaluate the benefits of this combination in comparison to cellulose and silica individually. Through functionalization as IMAC and RP, their adsorptions are modified as desired. In addition to the immobilized metal ions onto the IMAC composite, the role of base materials is tested to optimize the best enrichment strategy. Complete enrichment methodology is developed for phosphopeptides and phospholipids using standards and real samples. Cellulose-silica C18 is explored as unique hydrophobic/hydrophilic desalting tool to overcome the limitations of commercial materials that account for the loss of hydrophilic species.

Results and Discussion

Characterization of Cellulose-Silica

Cellulose-silica composite previously has been synthesized using different strategies and overview of their synthetic procedure and applications is given in Table S1. The current composite is synthesized using a new method, derivatized as IMAC and RP and applied in proteomics for the first time. The synthesis of composite involve oxidation of cellulose using periodate, amino modification of silica and finally composite fabrication is illustrated in Figure S1 whereas further derivatization of composite is described in Figure S2. The surface morphology is tested by Scanning Electron Microscopy (SEM) whereas the purity and elemental composition is determined by Energy Dispersive X-Ray Spectroscopy (EDX). The composite reveals the presence of micro-scale cellulose fibers and silica particles with heterogeneous morphology (Figure S3A). The elemental composition shows no impurities with carbon content 19.93 %, oxygen 52.72 % and silicon 27.34% (Figure S3 B).

Chemical functionalization like oxidation of cellulose, aminopropyl bonded silica formation and the final composite is confirmed by Fourier transform infrared spectroscopy (FT-IR). During oxidation of cellulose, there is change in the region of carbonyl group stretching vibration. In case of composite, the peak at 1630 cm^{-1} belongs to carbonyl in the case of amides which shows the synthesis of cellulose-silica composite. Peak at 3370 cm^{-1} represents the N-H stretch which also confirms the attachment of cellulose with amino-propyl bonded silica. The peak is broad due to the presence of OH groups. Peaks of Si-O and Si-C are observed in the region of 1100 cm^{-1} and 800 cm^{-1} (Figure S4). For cellulose-silica-IMAC, the C-N stretch is at 1056 cm^{-1} which indicates

the attachment of IDA to composite (Figure S5). Similarly, the functionalization as RP is confirmed by the signal at 3370 cm^{-1} for N-H stretch and strong peaks for C18 C-H stretch around 2900 cm^{-1} (Figure S6).

The commercially available silica used in this study has particle size 200-425 mesh, pore size (60 Å) , pore volume 0.75 cm^3/g and surface area 480 m^2/g. The pore size and surface chemistry (silanol groups) affect the mass transfer and surface reactivity[15] and physical characteristics of silica with respect to separation can be determined by near infrared spectroscopy, as reported earlier[16]. The higher surface area allows better derivatization of silica with cellulose to form the composite. The enrichment efficiency of composite also depends on the hydrophilicity in addition to the above-mentioned physical characteristics like porosity, particle size and surface area.

Cellulose Silica-IMAC - Metal Ion Immobilization

Four metal ions are immobilized onto cellulose-silica IMAC for phosphopeptides enrichment from β-casein digest. Choice of Fe^{3+} is made because of its traditional use in phosphoproteomics, however there is problem of acidic peptides enriched along with the phosphopeptides which hamper the selectivity and hence compromise the binding capacity (Figure S7a). Zr^{4+} shows better selectivity with five phosphopeptides (one tetra phosphorylated and 4 mono phosphorylated peptides) enriched (Figure S7b). However, focus in this study is placed on lanthanide ions due to their performance in phosphopeptides enrichment in comparison to the transition metal ions. In case of cellulose-silica IMAC-La^{3+} (Figure S7c) and cellulose-silica IMAC-Sm^{3+} (Figure S7d), there is a difference of one mono-phosphopeptide at 3054.892. Sm^{3+} is enriching the

highest number of phosphopeptides and hence immobilized onto the cellulose-silica IMAC for further studies.

Selectivity of Cellulose-Silica IMAC-Sm^{3+} for Phosphopeptides

Different complexity levels are used to determine the selectivity of cellulose-silica-IMAC-Sm^{3+} composite, i.e. the ability to bind the target molecules present in complex mixtures like serum or HeLa cell extract. The complexity levels are attained by mixing varying ratios of protein/peptide backgrounds. Selectivity depends on the nature of affinity material as well as on the nature of competing molecules. Phosphopeptides from β-casein digest are spiked in de-phosphorylated HeLa cell extract. Figure 1 represents the selectivity measurements done for spiked mixtures in ratio of 1:100; (b) 1:500; (c) 1:1000; (d) 1:1500; and (e) 1:2000. Five phosphopeptides are detected with amount of background increased up to 1500 times. Washing steps remove the non-phosphopeptides. For the sample 1:2000 (β-casein: de-HeLa cell), two phosphopeptides at m/z 2061 and 3122 are enriched. Both of these phosphopeptides are hydrophilic in nature; 2061 (FQSEEQQQTEDELQDK, 33-48, 1P) and 3122.9 (RELEELNVPGEIVESLSSSEESITRI, 26-41, 4P) with Krokhin_NET hydrophobicity values of 0.09920053 and 0.5137824 respectively. There is no methyl esterification carried out as it does not derivatize 100% of carboxylic acid groups, and that it can increase the complexity of the analysis due to signals originating from peptides with different degrees of O-methylation. Furthermore, the buffer used for O-methyl esterification can result into partial de-amidation forming glutamic acid from glutamine as well as aspartic acid from asparagine. These amino acid residues are subsequently O-methylated, adding further complexity to the analysis. In addition, O-methyl

esterification requires lyophilisation of the sample, which causes adsorptive losses of phosphopeptides to the surfaces.

Sensitivity of Cellulose-Silica IMAC-Sm^{3+}

The ability of composite to enrich the lower concentrations of phospho content is determined by applying varying concentrations. The complete study is designed with the starting concentration of 100 femtomole of β-casein and going down to 1 femtomole (Figure S8). Three phosphopeptides with one tetra phosphorylated peptide are evident in the case of 1 femtomole sample. Usually multi-phosphopeptides are difficult to enrich at lower concentrations due to their easy loss during washing steps[17], however, Sm^{3+} shows high affinity which is enhanced by the hydrophilic base materials (cellulose and silica) and overall there is improved enrichment of multi-phosphopeptides.

Role of Base Materials (Cellulose and Silica)

Base materials play important role in enrichments as reported earlier in the MELDI technique[18]. The consideration of only immobilized metal ions and ignoring the base materials leads to non-specific bindings. To observe the impact, β-casein digest is applied to cellulose, silica, cellulose-silica composite and end capped silica-lanthanum oxide[19] (Figure 2). In case of cellulose and silica, mostly the non-phosphopeptides are detected which are labelled as NPP (Figure 2A). The binding of these peptides is linked to the hydrophilic nature and surface hydroxyl groups of cellulose and silica; depicted through the common bound peptides (NPP) because of the structural similarities. Phosphopeptides interact through their phosphorylated serine site whereas non-phosphopeptides (NPP) interact through the carboxylic group of acidic amino acids as

glutamic acid (Glu) and aspartic acid (Asp). In case of cellulose, hydrophobic sites because of CH-O[20] allow binding of peptides with relatively higher degree of hydrophobicity, contributed to peptide backbone by alanine, valine, leucine, isoleucine and proline. For silica, silanol groups (Si-OH) are abundant on the surface and available for interaction with acidic and phosphorylated peptides. As a combination of both materials, cellulose-silica-IMAC-Sm^{3+} has three interaction sites namely cellulose, silica and Sm^{3+} ions. Nine phosphopeptides are detected in MS spectrum of cellulose-silica composite (Figure 2A-c). The presence of weak signals of non-phosphopeptides at m/z 1069 and 2106 are because of cellulose and silica. Under similar washing steps, it can be seen that for cellulose-silica composite, non-phosphopeptide interactions are weakened whereas these washings may not be enough for individual cellulose and silica.

A second comparison is made among silica, cellulose-silica IMAC-Sm^{3+} and end-capped silica-lanthanum oxide composite (Figure 2B). The presence of non-phosphopeptide at m/z 1069 because of silica is present in both MS spectra of individual silica (Figure 2B-a) and cellulose-silica-IMAC- Sm^{3+} (Figure 2B-b). The absence of non-phosphopeptides in silica-lanthanum oxide composite is because of the end-capping which blocks the un-derivatized silanol groups on silica (Figure 2B-c). The detail regarding peptides detected during the comparison is given in Table 1. In conclusion, not only the metal ions immobilized onto the IMAC material are taken into account while designing any material for enrichment, rather base material also contributes in enhancing the enrichment efficiency. Proper consideration towards the role of base material actually helps to optimize the sample preparation methodology.

Reproducibility

Development of enrichment strategy relies on the reproducibility of material and method independent of time and space. Figure S9 is representative of the reproducible results using serum digest. The intensity difference can be attributed to sample preparation, co-crystallization procedure and instrumental factors as laser power and laser intensity. For cellulose-silica IMAC-Sm^{3+}, statistical analysis is applied on the reproducibility study. The highest standard deviation (SD) is found 0.379 and the lowest 0.150 (Table S2).

Cellulose-Silica-IMAC-Sm^{3+} for Phosphoproteins/Phosphopeptides

The developed enrichment strategy for cellulose-silica IMAC-Sm^{3+} is tested on protein mixture containing phosphoproteins like α-, and β-casein with non-phosphoproteins as lysozyme, myoglobin, cytochrome c and BSA. The non-phosphoproteins are selected considering the range from hydrophobic to hydrophilic. Phosphoproteins α-, and β-casein having mass in the range of 24 kDa to 26 kDa are successfully enriched from the protein mixture with no trace of non-phosphoproteins (Figure S10). Both α-casein variants, αS1 and αS2 including the β-casein, are detected.

The complexity level is further enhanced and a fraction of this protein mixture (200 μL) is digested to provide high complexity background in the peptide range. Figure 3a shows MS spectrum of the direct analysis of peptide mixture after digestion. Only four phosphopeptides can be detected with high number of non-phosphopeptides. However, after the enrichment, 22 phosphopeptides are detected with 13 mono-phosphorylated, three di-phosphorylated, one tri-phosphorylated and five tetra-phosphorylated peptides derived from α-, and β-casein.

Cellulose Silica-IMAC-Sm^{3+} for Phospholipids Analysis (from standard)

The separation strategy for phospholipids is based on Folich's extraction method which is time consuming. IMAC based strategy is being reported for the first time to enrich phospholipids without any pre-fractionation and extraction. The composite- IMAC-Sm^{3+} and the optimized protocol simultaneously enriches both phospho- based classes of compounds, i.e. phospholipids and phosphopeptides in a single step. They are successively eluted, with no cross contamination. Phospholipids, due to their low mass/size, bind to IMAC materials and occupy/compete for the binding sites otherwise to be used for phosphopeptides. Therefore cellulose-silica IMAC-Sm^{3+} is also tried for the analysis of phospholipids with little modification in the enrichment protocol. As phospholipid has lipid as its counterpart, the bound phospholipids can be eluted using 5% ammonium hydroxide in methanol/dichloromethane mixture (1:9, v/v). It is important to consider that the phosphopeptides do not get eluted because of the hydrophobic nature of optimized elution buffer. Phospholipids mixture (Sigma Aldrich) is applied to the cellulose-silica IMAC-Sm^{3+}. As they are low mass species, gold nanoparticles in 0.1% TFA and carbon based LDI matrix are used for their MS analysis to avoid the matrix interferences (Figure S11a-b). All phospholipids present in the mixture are enriched with no background from materials used as LDI matrix.

Enrichment of Phospholipids/Phosphopeptides from Egg Yolk Digest

Hen egg yolk consists of phosphatidylcholine (PC, 73%), lysophosphatidylcholine (LPC, 5.8%), sphingomyelin (SM, 2.5%), phosphatidylethanolamine (PE, 15.0%), lysophosphatidylethanolamine (LPE, 2.1%) and phosphatidylinositol (PI, 0.6%). This composition can only be used as a rough guide, because the phospholipids and acyl composition of egg are influenced by many parameters, including the season, foods eaten

by the hen, etc. However, at least six different PLs are present in the egg yolk extract. Among egg yolk proteins, phosvitin (PV, 7%) and lipovitellin (LP, 69%) are phosphorylated. Single batch extraction is applied onto the cellulose-silica IMAC-Sm^{3+} for egg yolk digest. The stepwise elution of phospholipids and phosphopeptides is carried out. In first elution, enriched phospholipids are recovered, followed by an extra methanol and water wash step to remove any residual trace. They are analysed by LDI-MS. In the second eluted fraction, with 2% aqueous ammonium hydroxide, phosphopeptides are released, spotted with DHB buffer and recorded by MALDI-MS. Figure 4 shows MS spectra of tryptic digest of egg yolk before and after the enrichment. Only one phosphopeptide is identified prior to enrichment (Figure 4a) whereas after the enrichment by cellulose-silica IMAC-Sm^{3+}, 15 phosphopeptides are detected (Figure 4b). Figure 5 (3 expanded mass ranges) shows phospholipids detected in the eluted fraction belonging to different phospholipid classes. Phosphopeptides identified from phosvitin belong to Domain-III, position 1301-1322 SGHLEDDSSSSSSSSVLSKIWG (PV 190-211) with 6-10 phosphorylations. Highly phosphorylated peptide from phosvitin with 10P is detected at m/z 1894.98 ***S***GHLEDD***SSSSSSSS***VL***S***K (10P). For lipovitellin, Domain IV, position 1056-1076 IITEVNPE***S***EEEDE***SS***PYEDI, three phosphopeptides are detected.

Phosphorylation and Prostate Cancer

Phosphoproteins of clinical value are reported for prostate cancer. Their phosphorylated peptides in the serum digest are enriched by cellulose-silica IMAC-Sm^{3+} (Table S3). The same phosphorylation sites can also be used to enrich the whole phosphoprotein (as shown in Figure S10). Thus the developed enrichment material and protocol can enrich the phosphopeptides/phosphoproteins of clinical significance. Phosphopeptides

associated with the prostate cancer are highlighted in Table S3. To understand the importance, some of the phosphorylated proteins are discussed here in relevance to prostate cancer. CCDC124 (TIEDAIAVL**S**VAEEAADR, peak 26, inset Figure S12) is a novel factor operating both for proper progression of late cytokinetic stages in eukaryotes and for the establishment of Rap2 signaling dependent cellular functions proximal to the abscission site[21]. Down-regulation of phosphodiesterase 4B (PDE4B, ALDPQ**SS**PGLGRIMQAPVPHSQR, peak 36, inset Figure S12) activates protein kinase A and contributes to the progression of prostate cancer[22]. In clinical prostate cancer specimens, the expression of uncoupling protein 1 (UCP1, **S**RQ**T**MDCAT, peak 18, NNILADDVPCHLV**S**ALIAGFCA**T**AM**SS**PVDVVK, peak 45, inset Figure S12), a brown fat-specific marker, is enhanced with the level of expression correlated to disease progression from primary to bone metastatic cancers[23]. Aminopeptidase N (GA**S**VLRMLSSFLSEDVFK, peak 29, inset Figure S12) has been associated with the growth of prostate cancer and is suggested as a suitable target for anti-cancerous therapy[24]. Ang-III and related converting enzymes contribute to cell proliferation of prostate cancer, and may be implicated in cancer progression[25]. Functional PDC can form in mitochondria outside of the matrix in prostate cancer cells and PDHK1 (SF***SS***DSGSSPASER, peak 27, LFN**Y***M**Y***STAPRPRVETSR, peak 35, inset Figure S12) is commonly tyrosine phosphorylated by diverse oncogenic tyrosine kinases localized to different mitochondrial compartments. Expression of phosphorylation-deficient, catalytic hypomorph PDHK1 mutants in cancer cells leads to decreased cell proliferation under hypoxia and increased oxidative phosphorylation with enhanced mitochondrial utilization of pyruvate and reduced tumor growth in xenograft nude mice.

Together, tyrosine phosphorylation activates PDHK1 to promote the Warburg effect and tumor growth[26]. The expression patterns of the HOX network in human prostate cell phenotypes, representing different stages of prostate physiology and prostate cancer progression, make it possible to discriminate between different human prostate cell lines and to identify loci and paralogous groups harboring the HOX mostly involved in prostate organogenesis and cancerogenesis[27]. The amount of cross-linked fibrinogen gamma-chain dimer in plasma may correlate with tumor-associated fibrin deposition. The tumor-biological relevance of this potential marker protein needs to be explored[28]. Mitochondrial ribosomal protein L54 has been shown to result in prostate cellular differentiation and cell growth arrest. Up-regulation is identified in malignant prostate cancer cell lines and in prostate cancer tissue samples[29]. The information can be used along with PSA test for prostate cancer to diagnose with high confidence level.

According to the information provided by National Cancer Institute, screening tests for prostate cancer need method with high sensitivity and specificity. After optimization of protocol for the synthesized composite in different formats with commercial perspective like packed in-tip or in column, it can be used for screening of prostate cancer.

Cellulose-Silica C18 - Desalting and Recovery of Hydrophilic Molecules

Pre-concentration of peptide mixtures by desalting materials is preferred prior to mass spectrometric analysis for increased signal-to-noise ratio, sensitivity and consequently the sequence coverage. A large proportion of peptides, however, remain undetected by MS presumably because of their loss during the sample preparation, or suppression in the ionization process. Cellulose-silica C18 is synthesized as an alternative to the conventional desalting materials which face problem of loss of hydrophilic species. As

RP material, it provides combination of hydrophilic composite and hydrophobic C18 chain. This contrasting combination of cellulose-silica-C18 helps to retain moieties of both hydrophilic and hydrophobic origin and hence better recovery of proteins/peptides from range of samples. Cellulose-silica C18 is applied to casein mixture (α- and β-) having different salt concentrations (Figure S13). The high salt concentration of 8 M urea produces no adducts with all the peptides enriched and identified. The issue of loss of hydrophilic species by the most desalting materials is also verified. During the protein digestion procedures, the peptides containing methionine get oxidized which make the peptides more hydrophilic. For tryptic digest of α-casein, the styrene divinyl benzene (SDB) tips (GL Sciences®) and ZipTip C18 (Millipore®) show loss of the peptides (Figure S14a and S14b). For graphite tips (GL Sciences®), there is high number of peptides in mass range between 1500 to 3200 Da (Figure S14c). Cellulose-silica C18 shows distinctive peptide mass range with total of 32 peptides identified. The highlighted region shows the major difference between GL tips and cellulose-silica C18 (Figure S14d). The difference in relative ion intensity of peptides observed also indicates that hydrophilic peptides are favoured by cellulose-silica C18.

Another study with non-fat milk digest is carried out for the comparison and improvement of the sequence coverage. Peptides enriched by cellulose-silica C18 are shown in Figure S15. Peptides from milk caseins are identified using SwissProt database and sequence coverage analysis is also carried out for these milk proteins (Table S4). For α-casein, sequence coverage of 26%, 38%, 65.8% and 84.8% is shown by SDB Tip, ZipTip C18, GC Tip and cellulose-silica C18 respectively. For β-casein, comparison is distinctive with the highest sequence coverage shown by cellulose-silica C18 (98.4%)

and the lowest by ZipTip C18 (30.2%) and SDB Tips (32.9%). The efficiency difference shows that only hydrophobicity due to C18 chains is not enough for the desalting, rather contribution also comes from the hydrophilic base of cellulose and silica.

Experimental

Materials

The chemicals and reagents used in this study are given in the Supporting information.

Synthesis of Cellulose-Silica Composite

The derivatization of silica with aminopropyl silane, oxidation of cellulose and the chlorination is given in Supporting Information. Aminopropyl bonded silica was refluxed with cellulose acetyl chloride in 10 mL of anhydrous tetrahydrofuran (THF) for 24 h at 66 °C. The resulting composite was purified by washing with THF followed by methanol-water (1:1, v/v) to hydrolyze the unreacted acetyl chloride.

Functionalization of Cellulose-Silica Composite as IMAC

Non-selective oxidation of composite was carried out in suspension with 300 mL of 0.1 M aqueous solution of perchlorate ($HClO_4$) for 24 h at 25 °C. The subsequent chlorination involved reflux reaction in 10 mL anhydrous thionyl chloride for 24 h at 76 °C. After the removal of thionyl chloride and vacuum drying, cellulose-silica composite with acid chloride functionalities was suspended in 20 mL of anhydrous THF, refluxed with 1 g of iminodiacetic acid (IDA) and 1 mL of triethylamine (TEA) for 72 h at 66 °C. The product was washed with dichloromethane followed by ethanol and water.

Metal Immobilization

Different metal ions (Fe^{3+}, Zr^{4+}, La^{3+}, Sm^{3+}) were immobilized using their respective salt solutions. In general, 1 mg of cellulose-silica IMAC was incubated with 100 μL of 0.1 M salt solution for 45 min at 28 °C. The material was washed twice with water followed by

the activation buffer (100 μL, 80 % acetonitrile (ACN) in 0.1 % (v/v) trifluoroacetic acid (TFA).

Functionalization as Desalting Material

The synthesis of cellulose-silica C18 was carried out by suspending cellulose-silica acid chloride in 25 mL ACN and mixed with 2 g octadecylamine (ODA). The mixture was refluxed for 72 h at 80 °C. The product was washed with ethanol.

Sample Preparation for Enrichment/Desalting

The digestion of standards and biological fluids (non-fat milk, egg yolk and serum for prostate cancer), preparation of dephosphorylated HeLa cell sample, their use in selectivity, sensitivity and as desalting material is provided in Supporting Information.

Enrichment Protocol for Phosphopeptides/Phospholipids on Cellulose-Silica IMAC-Sm^{3+}

Phosphopeptides enrichment was carried out by incubating 20 μL of tryptic digest containing 0.1% TFA with cellulose-silica IMAC-Sm^{3+} at room temperature under gentle shaking for 30 min. The material was washed twice with 100 μL of 80% ACN in 0.1 % TFA followed by a wash with 110 mg DHB[30] dissolved in 10% ACN: 0.2% TFA. After a wash with water, 20 μL of 2% ammonium hydroxide (pH 10) was used for the elution of bound phosphopeptides. DHB could serve as displacer for the non-phosphorylated peptides which thus significantly prevented the binding of non-phosphorylated peptides. The idea of using DHB in washing solvent was supported by Larsen et al[31]. During the current study, it was optimized that 20 mg/mL DHB in 2 % ACN/0.1 % TFA was efficient in removing the non-phosphorylated compounds.

For phospholipids, phospholipid mixture from Sigma Aldrich (40 μL) was used as standard. The sample was loaded on activated cellulose-silica IMAC-Sm^{3+} and incubated for half an hour at room temperature. The material was washed with 800 μL of 1% CH_3COOH in $CHCl_3$ and CH_3OH (10:30) followed by 80% ACN in 0.2% TFA. For elution of phospholipids, 20 μL of 2% ammonia in CH_3OH (70:30) was used.

In case of egg yolk and serum, successive elution of phospholipids followed by phosphopeptides was carried out. First phospholipids were eluted using 2% ammonia in CH_3OH (70:30) followed by the elution of phosphopeptides using 1.5 % ammonium hydroxide solution.

Desalting by Cellulose-Silica C18

Desalting was carried out for different samples like β-casein in different salt concentrations (1 M urea, 4 M urea and 8 M urea) for desalting efficiency, α-casein for hydrophilic recovery and milk digest for improved sequence coverage. The comparison is made to the commercial desalting materials. Cellulose-silica C18 (1 mg) was activated by 100 μL of 80% ACN in 0.1% TFA. The sample (20 μL) was incubated for 10 min and the supernatant was removed. The material was washed with 100 μL of 5 % ACN in 0.1 % TFA followed by a wash with water. Phosphopeptides were eluted by using 10 μL of 50 % ACN in 0.1% TFA.

Mass Spectrometric Analysis

The MS analysis details are provided in the Supporting information.

Conclusion

A new methodology has been developed to synthesize cellulose-silica composite and its functionalization as IMAC and RP. FT-IR spectral study confirms the synthesis by showing the characteristic peak of carbonyl amide at 1630 cm^{-1} and C-N stretch at 1056 cm^{-1} for IMAC formation. Cellulose-silica IMAC-Sm^{3+} composite have been developed to enrich phosphopeptides/phospholipids in a single step followed by their two-step successive elution. Phosphopeptides are recorded with MALDI-MS whereas the low molecular weight phospholipids are analyzed by newly designed LDI-MS methodology. The role of metal ions on the designed IMAC composite is also compared to the contribution coming from the base materials, i.e. cellulose and silica. High selectivity of 1:2000 and sensitivity down to 1 fmol allows detection of low concentrated phosphopeptides from protein mixture digest and human serum. The material and the sample preparation protocol is optimized to the extent that the low concentrated phosphopeptides biomarkers relevant to prostate cancer are enriched from serum digest. Efficient enrichment from high salt concentrations, recovery of hydrophilic species and the improved sequence coverage makes cellulose-silica C18 a competent desalting tool as compared to SDB tip, GC Tip from GL Sciences and ZipTip C18.

Acknowledgments

This work is supported by the Higher Education Commission (HEC) of Pakistan. The authors are thankful to the Institute of Analytical Chemistry and Radiochemistry, Leopold-Franzens University, Innsbruck, Austria for providing access to laboratories and MALDI-MS instrument. Furthermore, the authors declare that they have no conflict of interest.

Table 1: Comparison of Cellulose-Silica-IMAC with Cellulose (powder, Sigma Aldrich), Silica gel (Sigma Aldrich mesh size 75-230), and end-capped Silica-Lanthanum oxide composite (reported) for β-casein tryptic digest. Phosphorylation at serine is shown with asterisk as S*. Glutamic acid (Glu) and aspartic acid (Asp) in bold as **E** and **D** for acidic amino acids. Hydrophobic amino acids alanine, valine, isoleucine, leucine and proline are represented as underlined A, V, I, L and P.

m/z	Sequences	Amino acid position	Nature of Peptide			Enriched Peptides			
			Phospho-rylation	Acidic peptide (Glu/Asp)	Degree of hydrophobicity (Krokhin_NET)	Cellulose	Silica	Cellulose -silica IMAC	Silica-La_2O_3
975	KFQS*EEQQQ	47-48	1P	2 Glu	0.0445766	-	-	*	*
1013	HK**E**MPFPK	106-113		1 Glu	0.140542	*	*	-	-
1094	PVLGPVRGPFPIIV	215-224			0.4724251	*	*	-	-
1104	KFQS*EEQQQT	47-49	1P		0.04997731	-	*	*	*
1994	LLYQ**E**PVLGPVRGPFPIIV	207-224		Glu	0.5902297	*	*	-	-
2061	FQS***EE**QQQT**EDE**LQ**D**K	33-48	1P		0.09920053	*	*	*	*
2106	FLLQ**E**PVLGPVRGPFPIIV	205-224		Glu	0.6408923	*	-	*	-
2187	**D**MPIQAFLLYQEPVLGPVR	184-202		Asp	0.5987214	*	*	*	-
2352	NVPG**E**IV**E**S*LS*S*S***EE**SITR	32-40	4P		0.300066	-	-	*	*
2556	FQS***EE**QQQ**TEDE**LQ**D**KIHPF	33-52	1P		0.2779962	-	-	*	*
2467	WMHQPHQPLPPTVMFPPQSVL	158-178			0.5060722	*	*	-	-
2909	**D**MPIQAFLLQ**E**PVLGPVRGPFPIIV	200-224		Asp/Glu	0.7687585	*	*	-	-
2965	R**E**L**EE**LNVPG**E**IV**E**S*LS*S*S***EE**SITR	1-24	4P		0.476857	-	-	*	*
3054	KIEK FQS***EE**QQQT**E**DE**E**LQ**D**KIHPF	29-52	1P		0.04226933	-	-	*	*
3124	R**E**L**EE**LNVPG**E**IV**E**S*LS*S*S***EE**SITR I	1-25	4P		0.5137824	*	*	*	*
3180	R**E**L**EE**LNVPG**E**IV**E**S*LS*S*S***EE**SIT	26-39	4P		0.4447257	*	*	*	*

Figure and Table Captions

Figure 1: MALDI-MS spectra of tryptic β-casein digest spiked in dephosphorylated HeLa cell extract using cellulose-silica-IMAC-Sm^{3+} in different ratios as: (a) 1:100; (b) 1:500; (c) 1:1000; (d) 1:1500; (e) 1:2000. The symbol **β** represents phosphopeptides derived from β-casein.

Figure 2: Comparison (A) MALDI-MS spectra recorded for (a) Cellulose (powder, Sigma Aldrich); (b) Silica gel (Sigma Aldrich mesh size 70-230); (c) Cellulose-Silica-IMAC-Sm^{3+}. Comparison (B) (a) Silica gel (Sigma Aldrich mesh size 75-230); (b) Cellulose-Silica-IMAC-Sm^{3+} (d) End-capped Silica-Lanthanum oxide composite. Tryptic digest of β-casein is applied (20 μL diluted in 0.1% TFA) following the optimized buffer conditions. Enriched phosphopeptides are labeled as β, whereas non-phosphopeptides are labeled as β-NPP.

Figure 3: Enrichment of α- and β-casein phosphopeptides from protein mixture digest (containing α-, β-casein, myoglobin, lysozyme, cytochrome c, BSA, 1 mg/mL) (a) before enrichment (b) after enrichment. Symbol α and β represent phosphopeptides with number of phosphorylated serine and position.

Figure 4: MALDI-MS analysis of phosphopeptides derived from phosvitin (P02845) and lipovitellin (P87498) in egg yolk digest (a) before enrichment; (b) after applying cellulose-silica-IMAC-Sm^{3+}. Phosphopeptides are labeled as PV for phosvitin and LP for lipovitellin with position numbers.

Figure 5: MALDI-MS spectra of phospholipids in egg yolk digest employing two batches of cellulose-silica-IMAC-Sm^{3+} using (a) gold nanoparticles in 0.1% TFA as LDI matrix (b) carbon based LDI matrix. Abbreviations used: PC, phosphatidylcholine; PE, phosphatidylethanolamine; PI, phosphatidylinositol; SM, sphingomyelin; LPC, lyso-phosphatidylcholine; LPE, lyso-phosphatidylethanolamine.

Table 1: Comparison of Cellulose-Silica-IMAC with Cellulose (powder, Sigma Aldrich), Silica gel (Sigma Aldrich mesh size 75-230), and end-capped Silica-Lanthanum oxide composite (reported) for β-casein tryptic digest. Phosphorylation at serine is shown with asterisk as S*. Glutamic acid (Glu) and aspartic acid (Asp) in bold as **E** and **D** for acidic amino acids. Hydrophobic amino acids alanine, valine, isoleucine, leucine and proline are represented as underlined A, V, I, L and P.

References

1 H. Ma, J. R. McLean, L. F. Chao, S. Mana-Capelli, M. Paramasivam, K. A. Hagstrom, K. L. Gould and D. McCollum, *Mol. Cell. Proteomics,* 2012, **11**, 501-511.

2 J. Zhu and G. Sun, *ACS Appl. Mater. Interfaces*, 2014, **6**, 925-932.

3 A. Fleitz, E. Nieves, C. Madrid-Aliste, S. J. Fentress, L. D. Sibley, L. M. Weiss, R. H. Angeletti and F. Y Che, *Anal. Chem.*, 2013, **85**, 8566-8576.

4 M. Ndao, *Methods Mol. Biol.*, 2012, **818**, 67-79.

5 A. Venugopal, R. Chaerkady and A. Pandey, *Ann. Indian Acad. Neurol.*, 2009, **12**, 3-11.

6 V. Guryča, D. Roeder, P. Piraino, J. Lamerz, A. Ducret, H. Langen and P. Cutler, *Biology*, 2014, **3**, 205-219.

7 A. Khan, A. M. Asiri, M. A. Rub, N. Azum, A. A. P. Khan, I. Khan and P. J. Mondal, *Int. J. Electrochem. Sci.*, 2012, **7**, 3854-3902.

8 S. Kango, S. Kalia, A. Celli, J. Njuguna, Y. Habibi and R. Kumar, *Prog. Polymer Sci.*, 2013, **38**, 1232-1261.

9 A. A. Taha, Y. Wu, H. Wang and F. Li, *J. Environ. Manag.*, 2012, **112**, 10-16.

10 A. Tolonen, W. Haas, A. C. Chilaka, J. Aach, S. P. Gygi, and G. M. Church, *Mol. Sys. Bio.*, 2011, **7**, 461-472.

11 T. Kimuraa, *Phys. Chem. Chem. Phys.*, 2013, **15**, 15056-15061.

12 I. Feuerstein, S. Morandell, G. Stecher, C. W. Huck, T. Stasyk, H. L. Huang, D. Teis, L. A. Huber and G. K. Bonn, *Proteomics*, 2005, **5**, 46-54.

13 L. Trojer, G. Stecher, I. Feuerstein, S. Lubbad and G. K. Bonn, *J. Chromatogr. A*, 2005, **1079**, 197-207.

14 J. Duan, H. Wang and Q. Cheng, *Anal. Chem.*, 2010, 82, 9211-9220.

15 T. Witoon and M. Chareonpanich, *J. Sci. Technol.*, 2012, **34**, 403-407.

16 C. W. Huck, R. Ohmacht, Z. Szabo and G. K. Bonn, *J. Near Infrared Spec.*, 2006, **14**, 51-57.

17 J. Fíla and J. Honys, *Amino Acids*, 2012, **43**, 1025-1047.

18 M. Najam-ul-Haq, M. Rainer, L. Trojer, I. Feuerstein, R. M. Vallant, C. W. Huck, R. Bakry and G. K. Bonn, *Expert Rev. Proteomics*, 2007, **4**, 447-452.

19 F. Jabeen, D. Hussain, B. Fatima, S. G. Musharraf, C. W. Huck, G. K. Bonn and M. Najam-ul-Haq, *Anal. Chem.*, 2012, **84**, 10180-10185.

20 Q. Li and S. Renneckar, *Biomacromolecules*, 2011, **12**, 650-659.

21 P. Telkoparan, S. Erkek, E. Yaman, H. Alotaibi, D. Bayık and U. H. Tazebay, *PLoS One*, 2013, **8**, e69289.

22 E. Kashiwagi, M. Shiota, A. Yokomizo, M. Itsumi, J. Inokuchi, T. Uchiumi and S. Naito, *Prostate*, 2012, **72**, 741-751.

23 H. E. Zhau, H. He, C. Y Wang, M. Zayzafoon, C. Morrissey, R. L. Vessella, F. F. Marshall, L. W. Chung and R. Wang, *Clin. Cancer. Res.*, 2011, **17**, 2159-2169.

24 M. Wickström, R. Larsson, P. Nygren and J. Gullbo, *Cancer Sci.*, 2011, **102**, 501-508.

25 J. I. Teranishi, H. Ishiguro, K. Hoshino, K. Noguchi, Y. Kubota and H. Uemura, *Prostate*, 2008, **68**, 1666-1673.

26 T. Hitosugi, J. Fan, T. W. Chung, K. Lythgoe, X. Wang, J. Xie, Q. Ge, T. L. Gu, R. D. Polakiewicz, J. L. Roesel, G. Z. Chen, T. J. Boggon, S. Lonial, H. Fu, F. R. Khuri, S. Kang, J. Chen, *Mol. Cell*, 2011, **44**, 864-877.

27 M. Cantile, A. Kisslinger, L. Cindolo, G. Schiavo, V. D'antò, R. Franco, V. Altieri, A. Gallo, A. Villacci, D. Tramontano and C. Cillo, *J. Cell. Physiol.*, 2005, **205**, 202-210.

28 C. Gerner, W. Steinkellner, K. Holzmann, A. Gsur, R. Grimm, C. Ensinger, P. Obrist and G. Sauermann, *Thromb Haemost.*, 2001, **85**, 494-501.

29 J. L. Goodin and C. L. Rutherford, *Mol. Biol. Rep.*, 2002, **29**, 301-316.

30 M. R. Mirza, M. Rainer, Y. Güzel, M. I. Choudhary and G. K Bonn, *Anal. Bioanal. Chem.*, 2012, **404**, 853-862.

31 M. R. Larsen, T. E. Thingholm, O. N. Jensen, P. Roepstorff and T. J. D. Jørgensen, *Mol. Cell. Proteomics*, 2005, **4**, 873-886.

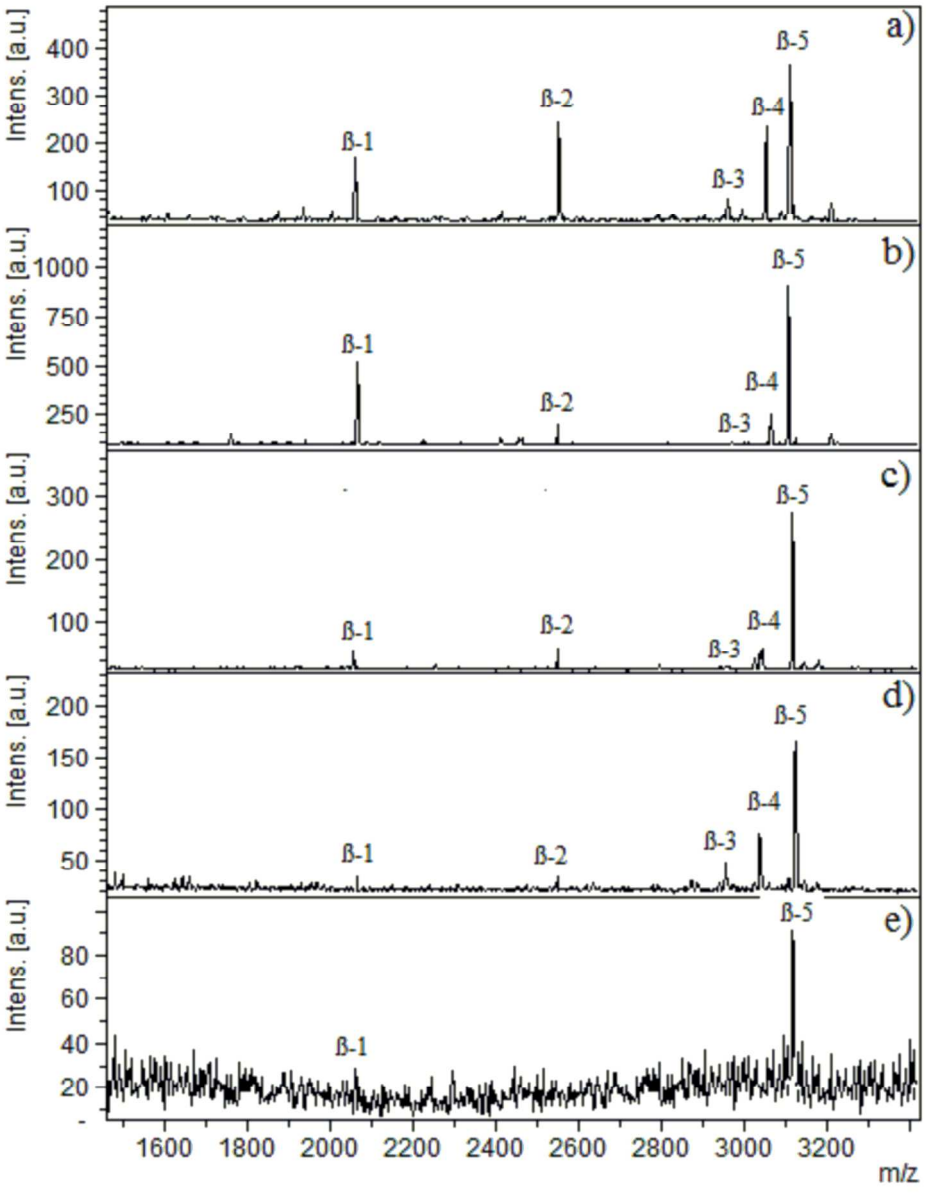


Figure 1: MALDI-MS spectra of tryptic β-casein digest spiked in dephosphorylated HeLa cell extract using cellulose-silica-IMAC-Sm3+ in different ratios as: (a) 1:100; (b) 1:500; (c) 1:1000; (d) 1:1500; (e) 1:2000. The symbol β represents phosphopeptides derived from β-casein.
153x197mm (300 x 300 DPI)

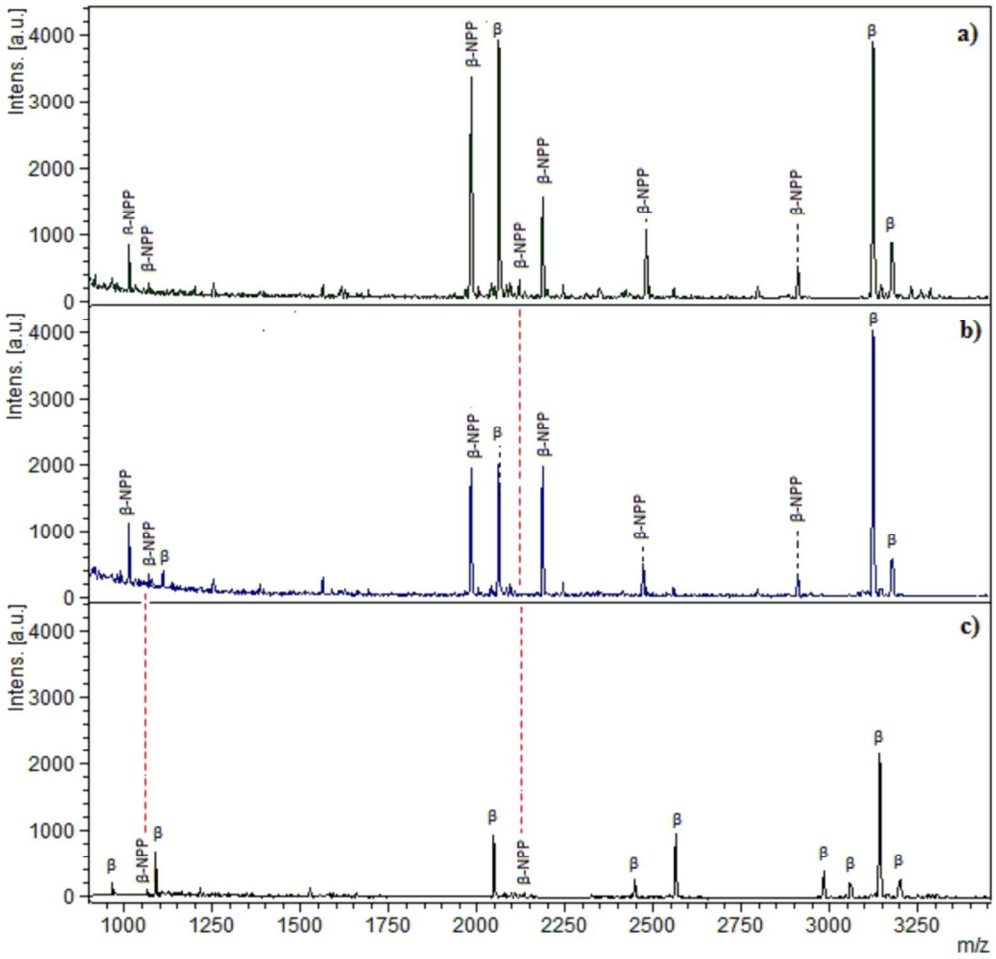


Figure 2: Comparison (A) MALDI-MS spectra recorded for (a) Cellulose (powder, Sigma Aldrich); (b) Silica gel (Sigma Aldrich mesh size 70-230); (c) Cellulose-Silica-IMAC-Sm3+. Comparison (B) (a) Silica gel (Sigma Aldrich mesh size 75-230); (b) Cellulose-Silica-IMAC-Sm3+ (d) End-capped Silica-Lanthanum oxide composite. Tryptic digest of β-casein is applied (20 μL diluted in 0.1% TFA) following the optimized buffer conditions. Enriched phosphopeptides are labeled as β, whereas non-phosphopeptides are labeled as β-NPP.
153x148mm (300 x 300 DPI)

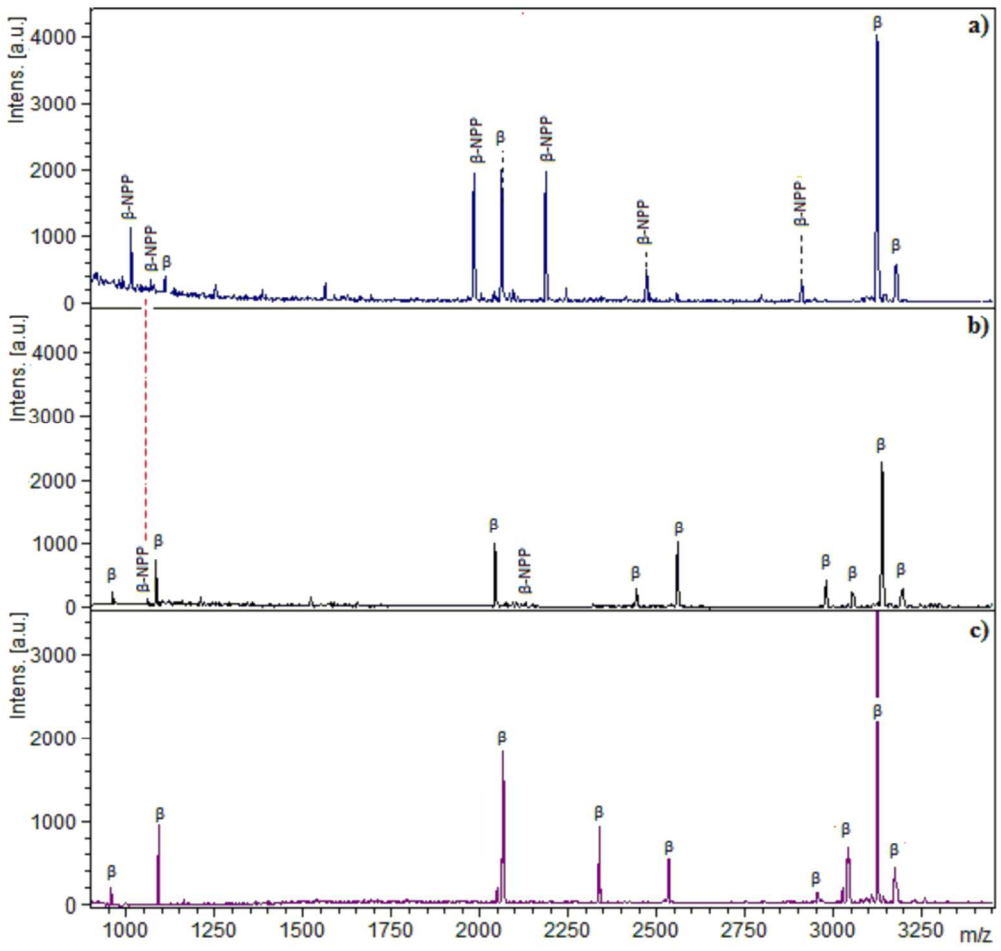


Figure 2: Comparison (A) MALDI-MS spectra recorded for (a) Cellulose (powder, Sigma Aldrich); (b) Silica gel (Sigma Aldrich mesh size 70-230); (c) Cellulose-Silica-IMAC-Sm3+. Comparison (B) (a) Silica gel (Sigma Aldrich mesh size 75-230); (b) Cellulose-Silica-IMAC-Sm3+ (d) End-capped Silica-Lanthanum oxide composite. Tryptic digest of β-casein is applied (20 μL diluted in 0.1% TFA) following the optimized buffer conditions. Enriched phosphopeptides are labeled as β, whereas non-phosphopeptides are labeled as β-NPP.
150x141mm (300 x 300 DPI)

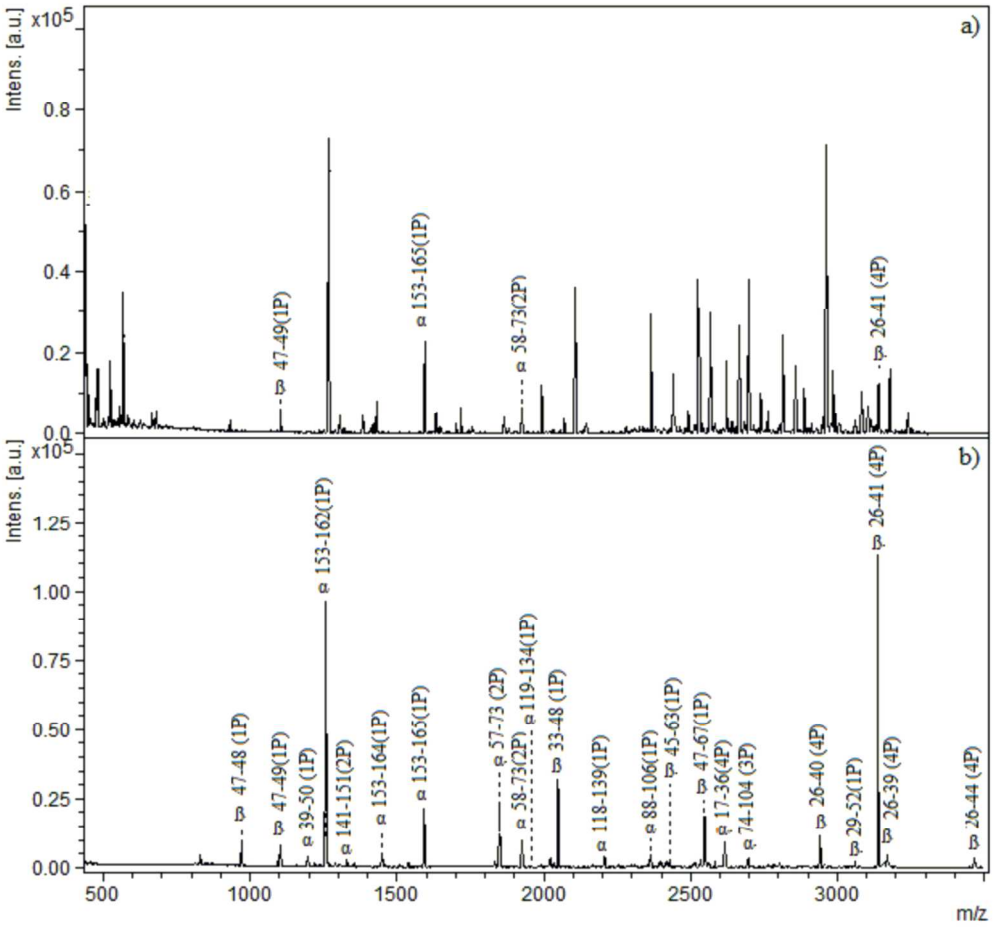


Figure 3: Enrichment of α- and β-casein phosphopeptides from protein mixture digest (containing α-, β-casein, myoglobin, lysozyme, cytochrome c, BSA, 1 mg/mL) (a) before enrichment (b) after enrichment. Symbol α and β represent phosphopeptides with number of phosphorylated serine and position.
153x142mm (300 x 300 DPI)

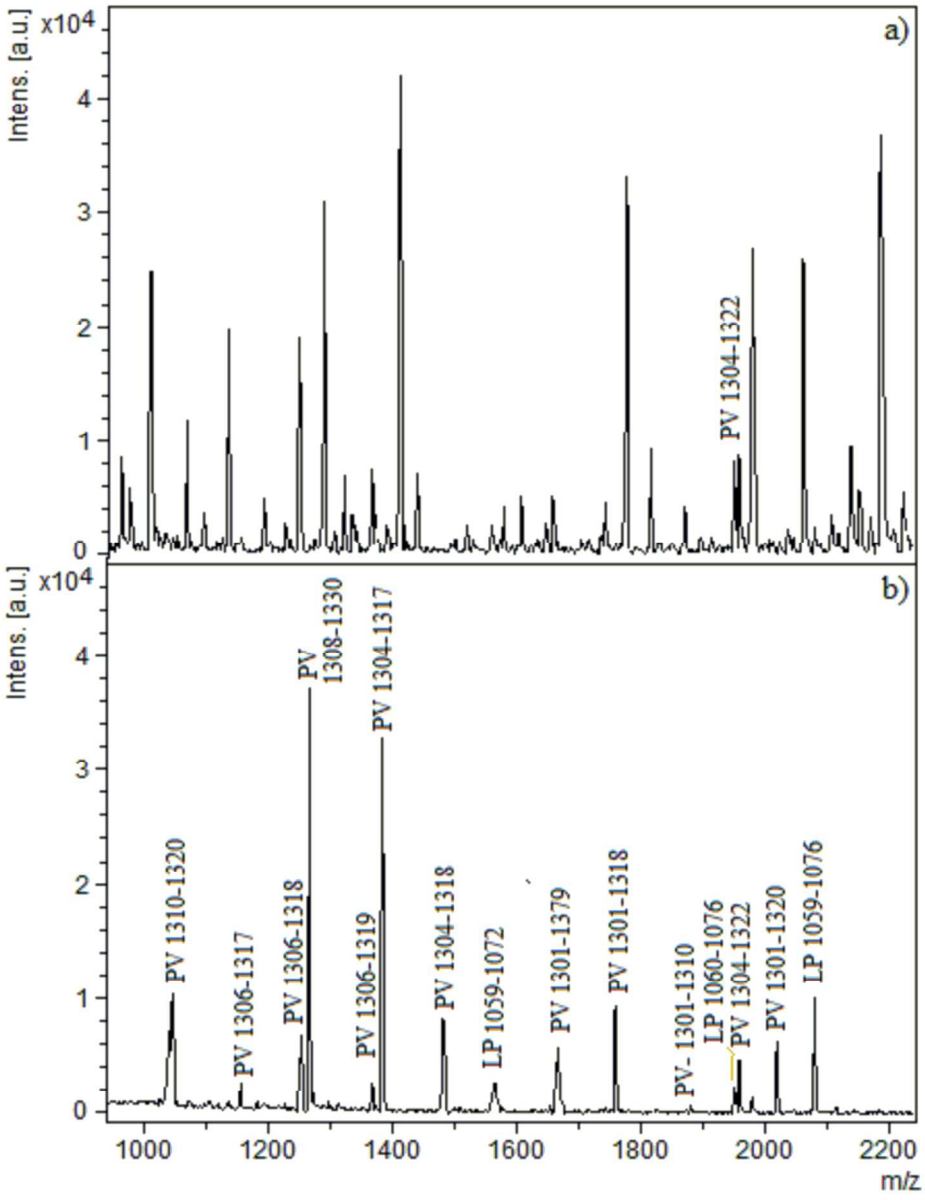


Figure 4: MALDI-MS analysis of phosphopeptides derived from phosvitin (P02845) and lipovitellin (P87498) in egg yolk digest (a) before enrichment; (b) after applying cellulose-silica-IMAC-Sm3+. Phosphopeptides are labeled as PV for phosvitin and LP for lipovitellin with position numbers.
153x198mm (300 x 300 DPI)

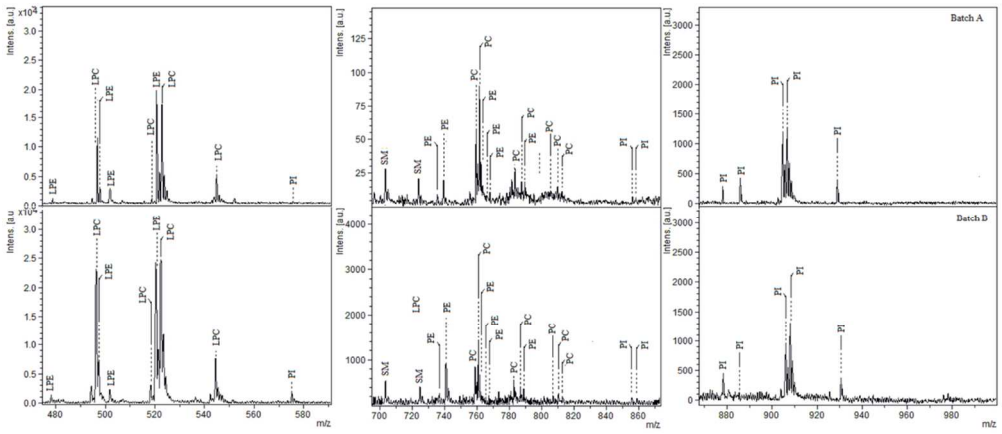


Figure 5: MALDI-MS spectra of phospholipids in egg yolk digest employing two batches of cellulose-silica-IMAC-Sm3+ using (a) gold nanoparticles in 0.1% TFA as LDI matrix (b) carbon based LDI matrix. Abbreviations used: PC, phosphatidylcholine; PE, phosphatidylethanolamine; PI, phosphatidylinositol; SM, sphingomyelin; LPC, lyso-phosphatidylcholine; LPE, lyso-phosphatidylethanolamine.
109x46mm (300 x 300 DPI)