SnO$_2$@Poly(HEMA-co-St-co-VPBA) Core-shell Nanoparticles Designed for Selectively Enriching Glycopeptides Followed by MALDI-MS Analysis

Wen-Wen Shen,^[a, b] Cheng-Ning Ma,^[a] Su-Feng Wang,[a] Huan-Ming Xiong,*^[a] Hao-Jie Lu,*^[a, b] and Peng-Yuan Yang^[a, b]

Abstract: The core-shell boronic-acid functionalized nanoparticles SnO$_2$@Poly(HEMA-co-St-co-VPBA) are designed for selectively enriching glycopeptides, followed by matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) analysis. Such 60 nm sized core-shell nanoparticles are prepared by means of copolymerization between 2-hydroxyethyl methacrylate (HEMA) grafted on SnO$_2$ nanoparticles, styrene, and 4-vinylphenylboronic acid (VPBA). All of the synthesis procedures are completed within 3 h. Cyclic boronate esters form between boronic-acid groups on the polymer chains and cis-diol groups on glycopeptides, and thus almost all intact glycopeptides from low-abundant horseradish peroxidase (HRP) and bovine asialofetuin (ASF) are enriched with high selectivity and efficiency. After enrichment, both intact N- and O-glycopeptides are characterized by multistage MS. Furthermore, we successfully apply this method to the human serum sample for characterizing the target glycoproteins haptoglobin and alpha-1-acid-glycoprotein. The present selective enriching method followed by multistage-MS analysis is proven to be a good choice for routine glycopeptide characterization.

Keywords: core-shell nanoparticles · enrichment · glycopeptides · mass spectrometry · polymerization

Introduction

Protein glycosylation, as one of the most important post-translational modifications (PTMs), regulates cellular mechanisms, including cell adhesion, receptor activation, signal transduction, molecular trafficking and clearance, and endocytosis.[1] Many diseases are caused by glycosylation or glycosidase deficiencies, such as autoimmune disorders[2] or infantile-onset symptomatic epilepsy.[3] The US Food and Drug Administration has agreed that over half of the recent cancer biomarkers are glycoproteins.[4] To identify glycosylation sites, as well as the primary structures of the glycans, mass spectrometric strategies have been developed.[5] However, as glycopeptides are always low in abundance when the glycoprotein is digested into peptides,[6] the MS signals of nonglycosylated peptides always heavily interfere with those of glycopeptides. Moreover, determination of glycan structures is usually difficult because of their vast diversity.[1] Therefore, selective enriching methods followed by multistage MS techniques are necessary for glycopeptide analysis.

A lot of strategies have been developed for enriching glycopeptides. Among them, the methods based on physical properties of the glycopeptides, for example, size exclusion[7] and hydrophilic interaction chromatography[8,9] are easily influenced by the complex components from practical samples. In contrast, the methods based on biospecificity or chemical binding have shown better performance on selective enriching. For example, lectin affinity chromatography is widely used for isolating glycopeptides from the peptide mixture based on the biospecificity interaction between lectin and the particular glycan structures.[10] However, this method is not suitable for the glycopeptides whose glycans cannot be recognized by the special lectin(s).[11] As another example, solid-phase micro-extraction through hydrazide chemistry depends on the covalent binding between the hydrazide resins and the oxidized cis-diol groups on gly-
Because this binding is irreversible, the following glycan cleavage by endoglycosidase, as required for MS analysis, always lost the glycan information. To obtain intact glycopeptides, it is optimal to utilize special chemical reactions to bond glycopeptides during enrichment and then recover them for analysis. Such reactions occur between polymers with oxylamino groups and aldehyde-attached glycopeptides[14,15] or between boronic acid groups and 1,2-cis-diol on glycans[16]. The latter boronate chemistry method is gentle and easy to handle. Several boronic-acid functionalized materials have been reported to date, such as m-amino-phenyl boronic-acid (APBA) functionalized agarose gels[16,17] APBA magnetic microparticles[18] APBA magnetic nanoparticles[19] and APBA mesoporous silica[20]. These materials isolate samples by means of solid phase extraction except the functionalized agarose gels, which are always prepared as a stationary phase for affinity chromatography. In general, the procedures by solid phase extraction were easier to deal with.

In this study, we prepared SnO$_2$@Poly(HEMA-co-St-co-VPBA) core-shell nanoparticles with boronic-acid groups and successfully applied them in the enrichment of glycopeptides. These nanoparticles were synthesized by means of a hydroxyl-exchange reaction and a polymerization route. The products were uniform, stable, monodispersed in water and their diameters were only about 60 nm. They had large surface areas and plenty of boronic-acid groups, so that glycopeptides with cis-diol groups could be enriched with high selectivity and efficiency. After enrichment and recovery treatment, the obtained glycopeptides were analyzed by MALDI-MS for identifying glycosylation sites and the corresponding glycan structures.

**Results and Discussion**

Illustrated in Scheme 1a are the synthetic routes of SnO$_2$@polymer nanoparticles. First, the freshly prepared SnO$_2$ nanoparticles were dispersed into 2-hydroxyethyl methacrylate (HEMA) to obtain the HEMA modified SnO$_2$. The exchange reaction mechanism has been reported in our previous work.[21,22] Secondly, the as-prepared SnO$_2$-HEMA sol was mixed with pure styrene (St) and 4-vinylphenylboronic acid (VPBA), and meanwhile, a solution containing styrene and K$_2$S$_2$O$_8$ was heated in water under stirring. After the two solutions were mixed together, copolymerization took place on the SnO$_2$-HEMA surfaces, resulting in nanoparticles with SnO$_2$-cores and hydrophilic boronic-acid surface groups. All of the procedures were completed within 3 h. It was rather simple and time-saving compared to the previous works, which needed over 16 h for the whole synthesis procedure.[18–20] The TEM images showed that the average diameter of the final produced uniform monodispersed nanoparticles was about 60 nm (Figure 1).

**Abstract in Chinese:**

研究发现，该富集材料能够从蔗糖化蛋白中高效分离富集出完整的N型和O型抗原，进而利用这些抗原对富集的抗原进行抗体结构解析。该方法不仅适用于标准体系化蛋白的抗原结构解析，还适用于复杂体系中新的抗原蛋白。我们成功富集并解析了人乳头状瘤病毒中的糖蛋白和酸性硫酸蛋白的抗原。该方法为后续研究提供了一种新的有效的工具。
The final produced nanoparticles were further characterized by IR, TGA, and ICP-AES measurements to investigate their structure and composition. IR spectra of SnO₂, SnO₂-HEMA, VPBA, and SnO₂@Poly(HEMA-co-St-co-VPBA) are shown in Figure S1 in the Supporting Information. The vibrations at $\tilde{\nu} = 538$ and $661 \text{ cm}^{-1}$ are ascribed to the Sn-O-Sn symmetric stretching inside SnO₂ nanoparticles and the Sn-O antisymmetric stretching on the SnO₂ surface, respectively.[22] Both vibrations are observed in the final nanoparticles, indicating that SnO₂ nanoparticles exist in the particle cores. The presence of HEMA in the final product is verified by the C=O stretching of ester groups at $\tilde{\nu} = 1717 \text{ cm}^{-1}$, whereas the presence of boronic-acid groups is verified by the B=O stretching at $\tilde{\nu} = 1365$ and $1340 \text{ cm}^{-1}$. Hence, the IR data confirms that the SnO₂ cores and boronic acid groups of the PHEMA related nanoparticle.

In Figure 2, thermogravimetric analysis (TGA) curves of SnO₂-HEMA and SnO₂@Poly(HEMA-co-St-co-VPBA) are shown. In Figure 2a, it is clear that there are three kinds of composition in the prepared SnO₂-HEMA sol. Physically adsorbed HEMA molecules (36.44 wt%) vaporized or decomposed before 150°C, followed by the decomposition of the HEMA molecules that were chemically bounded to the SnO₂ surface (39.28 wt%). When the temperature rose above 400°C, the reduction in weight stopped, indicating that only the inorganic component SnO₂ (24.28 wt%) remained. As for the SnO₂@Poly(HEMA-co-St-co-VPBA), four kinds of composition can be observed in Figure 2b. Between room temperature and 300°C, small molecules from physical adsorption (6.61 wt%) vaporized or decomposed.

Then, the polymers (73.05 wt%) decomposed within the range 300–400°C, followed by decomposition of the remaining phenylboronic acid species (8.29 wt%). Finally, above 500°C, SnO₂ and boron oxide remained (12.05 wt%). Analysis of the TGA data reveals that the content of B element is above 0.8 wt%. In addition, ICP-AES analysis indicated that there was 0.9 wt% B element in the nanoparticles. Thereby, the content of VPBA in the nanoparticles was calculated to be about 11 mg of VPBA segments in every 100 mg of powdered nanoparticles. For comparison, our previously reported APBA mesoporous silica contained 0.5 wt% B element.[23]

The affinity between the nanoparticle and glycopeptide was anticipated to be the typical boronate esterification since the boronic acid group could reversibly form cyclic-boronate esters with the 1,2-cis-diol of glycans,[14] as illustrated in Scheme 1b. To demonstrate that the newly prepared nanoparticles SnO₂@Poly(HEMA-co-St-co-VPBA) were able to selectively enrich glycopeptides from the mixture of glycosylated and non-glycosylated peptides, owing to the high contents of boronic-acid functional groups, mass spectra were acquired with both low and high mass ranges on the same spot by MALDI-MS. The final combined mass spectrum could have a wide range of mass from 500 to 6000 Da. In addition, SnO₂@Poly(HEMA-co-St) nanoparticles were used as a control material, to make sure that the enriching mechanism was based on the chemical reaction between VPBA of SnO₂@Poly(HEMA-co-St-co-VPBA) nanoparticles and glycopeptides. Two duplicate peptide mixtures of digested horseradish peroxidase (HRP) (1 mL and 100 ng/mL) were enriched by SnO₂@Poly(HEMA-co-St-co-VPBA) (Procedure I) or SnO₂@Poly(HEMA-co-St) (Procedure II). Both enriching procedures include exactly the same incubation–washing–elution sequence, as shown in Scheme 1b.

After the incubation step, no glycopeptides could be detected in the supernatant of Procedure I (Figure 3a), whereas many glycopeptide signals were found in the supernatant of Procedure II (Figure 3d), indicating that only the nanoparticles with boronic acid groups can selectively capture glycopeptides. After the washing step, the supernatant of Procedure I had only non-glycosylated peptides (Figure 3b), whereas the supernatant of Procedure II had both non-glycosylated peptides and glycopeptides (Figure 3e), indicating that non-glycosylated can interact with both nanoparticles through a non-specific manner. After the elution step, glycopeptides were expected to be released from the nanoparticles under the acidic condition if the adsorption interaction between the boronic-acid group and the cis-diol of glycans is specific. Indeed, glycopeptides were sensitively detected in the supernatant of Procedure I (Figure 3c), although no glycopeptides were seen in the supernatant of Procedure II (Figure 3f). Taking these results together, we can safely conclude that the interaction between the nanoparticles SnO₂@Poly(HEMA-co-St-co-VPBA) and glycopeptides is specific, arises from the boronic acid group on the surface of the nanoparticles.
The enriching efficiency of the SnO$_2$@Poly(HEMA-co-St-co-VPBA) for glycopeptides was studied. No glycopeptide signal could be found by MS for 2 mL digested HRP with the concentration of 100 ng mL$^{-1}$ without enriching procedure (Figure 3 g). After the enriching procedure, many signals assigned to glycopeptides were detected sensitively even 20% of the total eluent was analyzed (Figure 3 c). By contrast, 2 mL of digested HRP with the concentration of 10 ng mL$^{-1}$ was analyzed and the mass spectrum is shown in Figure 3 h. In Figure 3 h, the signals of glycopeptides are suppressed in the presence of these nonglycopeptides, because the nonglycopeptides would compete in desorption/ionization with the glycopeptides. For instance, the ratios of signal-to-noise (S/N) of the three glycopeptide peaks in the enriched sample ($m/z$: 2851, 3672, 3527) were 687 ($m/z$: 2851), 244 ($m/z$: 3672), 137 ($m/z$: 3527), whereas the S/N of the same peaks for the sample with 100-fold higher concentration was 304, 141, and 77, respectively. Therefore, a S/N enhancement factor could be estimated for the two samples. As a result, the S/N enhancement factor for glycopeptides achieved was nearly 3 orders of magnitude.

The enriching capacity of the SnO$_2$@Poly(HEMA-co-St-co-VPBA) nanoparticles for glycopeptides was also studied here. The 1 mL aliquot of 10–150 ng mL$^{-1}$ digested HRP was enriched by 100 ng nanoparticles. After enrichment, 10 mL elution solvent was added to release the glycopeptides and 0.5 mL eluent was spotted on the MALDI-plate for MS analysis. The glycopeptide ($m/z$: 3672) with the highest sensitivity here was used as the test peptide. If the concentration of the sample increased from 10 to 110 ng mL$^{-1}$, the S/N of this glycopeptide significantly increased. However, when the concentration increased over 110 ng mL$^{-1}$, the S/N of this glycopeptide stopped increasing, indicating that 100 ng nanoparticles reached enriching saturation for glycopeptides (Figure S2 in the Supporting Information). Therefore, we estimated the optimum enriching capacity of these nanoparticles for digested glycoprotein was about 1.1 gg$^{-1}$ of glycopeptide. In addition, the glycopeptide at $m/z$ 3672 was used for determination of the limit of detection (LOD) at a S/N of 3. As a result, LOD was calculated to be 0.5 ng mL$^{-1}$. By contrast, LOD of our most recent published material was about 2.5 ng mL$^{-1}$.[20]

To investigate the glycopeptides recovery of this enrichment method, the $^{18}$O/$^{16}$O labeling method[25] was used here. Glycopeptides of another digested glycoprotein ASF (1 ng mL$^{-1}$) with both N-glycosylation sites and O-glycosylation sites were enriched by these nanoparticles from 1 mL solution and the mass spectrum is shown in Figure 3 i. Equal amounts of digested ASF without enrichment were used as a control. Then the two sample parts were processed by Poroszyme immobilized trypsin (bulk media, Applied Biosystems) and labeled with H$_2$/$^{18}$O/H$_2$/$^{16}$O at the C-termini of the peptides from the enriched/control sample, respectively, as Wang et al. have previously described.[26] After $^{18}$O/$^{16}$O labeling, the two parts were mixed and analyzed by MS. Two glycopeptides ($m/z$ 3102 and $m/z$ 2738) were used for calculation. With the equation described by Yao et al.[26] and the MS results (Figure S3 in the Supporting Information), the average recovery was calculated to be 86.1%.
This enrichment method followed by multistage MS analysis provided a solution for determining both glycosylation sites and glycan structures of glycopeptides. It is acknowledged that mainly 6 types of monosaccharides are composed of glycan chains and the glycan chains are linked on peptides by the regular rules.\cite{5,7,27} Briefly speaking, there are mainly two types of glycopeptides, N- and O-glycopeptides. The N-glycopeptide requires the consensus sequence motif Asn-Xaa-Ser/Thr (Asn, asparagine; Ser, serine; Thr, threonine; Xaa is any amino acid except proline) and each glycan chain has a common core pentasaccharide Man\textsubscript{9}GlcNAc\textsubscript{2} and extends in the form of bi-, tri-, or tetra-antennary structure. As for O-glycopeptide, glycosylation happens on the Ser/Thr residues, and the glycan chain has eight kinds of cores, which start with GalNAc (N-acetylgalactosamine). Under these rules, the potential glycosylation sites can be predicted from the identified protein sequence, and glycan structures can be predicted from the mass spectrum in the first stage because glycan fragmentation occurs prior to peptide fragmentation.\cite{5} Then, the predictions of glycosylation sites and glycan structures are confirmed by multistage MS analysis. In this study, the intact glycopeptides from HRP and ASF were analyzed. From the mass spectra (Figures S4–S6 in the Supporting Information), their glycosylation sites and glycan structures were characterized. In agreement with other researchers,\cite{28–30} the detected glycopeptides from HRP are N-glycopeptides and those from ASF are both N- and O-glycopeptides. Table 1 shows 5 detected N-glycosylation sites from HRP and 4 detected glycosylation sites from ASF. Among these, the two N-glycan structures from ASF were characterized as the tri-antennary form and the composition of its four glycosylation sites was determined by analysis of the mass spectra (Figures S5 and S6 in the Supporting Information). As for the intact O-glycopeptides of ASF, it was difficult to individually ascribe the specified saccharide residues, as the two predicted O-glycosylation sites were differentiated by only one amino acid residue. Nevertheless, this was the first time that multistage-MS has been used for preliminary description of intact O-glycopeptides in ASF.

Besides the standard samples, we then detected target glycoproteins in real biological sample by using this method. Our purpose was to characterize two of the important human serum glycoproteins, haptoglobin\cite{31} and alpha-1-acid glycoprotein.\cite{32} Because of the wide dynamic range of protein abundances in human serum, 6 \( \mu \text{L} \) human serum (about 60 mg mL\textsuperscript{−1}) was loaded on the Agilent Human 14 Multiple Affinity Removal System column for collecting high-abundant proteins and low-abundant proteins per the manufacturer's instruction, respectively. All of the low-abundant proteins and 10\% of the high-abundant proteins were separated by sodium dodecyl sulfate/polyacrylamide gel (SDS-PAGE) electrophoresis (Figure 4). After separation, peptides were extracted from the protein bands in the high-abundant proteins lane and divided into 10 units for storage. 1-unit peptide solution from each band was used for protein identification per the previously reported method.\cite{33} The protein results are displayed in Table S1 in the Supporting Information. After that, another 1-unit peptide solution of alpha-1-acid glycoprotein and haptoglobin (detected as No.5 and 6 in Figure 4) was processed by SnO\textsubscript{2}@Poly(HEMA-co-St-co-VPBA) nanoparticles. Validated by the other report,\cite{34} glycopeptides with 4 N-glycosylation sites in both haptoglobin and alpha-1-acid glycoproteins were isolated and detected from the mass spectra (Figures S7 and S8 in the Supporting Information), as shown in Table 1. To our best knowledge, there are few studies on characterization of intact glycopeptides from real biological samples by using multistage-MS. Our method based on SnO\textsubscript{2}@Poly(HEMA-co-St-co-VPBA) nanoparticles might be widely applied in the future for detecting target glycopeptides.

### Table 1. List of the MS detected glycopeptides with their glycans at the glycosylation sites.

<table>
<thead>
<tr>
<th>Peptide sequence with glycosylation site(s) [a]</th>
<th>Glycan composition[b]</th>
</tr>
</thead>
<tbody>
<tr>
<td>QLTPTFYD\textsubscript{N\textsubscript{56}K}; VVHAVEVALATFNAES\textsubscript{N\textsubscript{176}GSYLQLVEISR}</td>
<td>(GlcNAc)\textsubscript{2}(Man)\textsubscript{3}(Fuc)(Xyl) [c]</td>
</tr>
<tr>
<td>VVHAVEVALATFNAES\textsubscript{N\textsubscript{126}GSYLQLVEISR}</td>
<td>(GlcNAc)\textsubscript{5}(Man)\textsubscript{3}(Gal) [c]</td>
</tr>
<tr>
<td>MVSH\textsubscript{N\textsubscript{60}LS} &amp; TTGATGATLINEO\textsubscript{N\textsubscript{70}TT}</td>
<td>(GlcNAc)\textsubscript{3}(Man)\textsubscript{5}(Gal) [c]</td>
</tr>
<tr>
<td>NLL\textsubscript{N\textsubscript{20}EHE\textsubscript{N\textsubscript{228}TTYLQTLR}</td>
<td>(GlcNAc)\textsubscript{3}(Man)\textsubscript{3}(Gal) [c]</td>
</tr>
<tr>
<td>VVLP\textsubscript{N\textsubscript{14}K} &amp; YSOVQDIKLK</td>
<td>(GlcNAc)\textsubscript{4}(Man)\textsubscript{3}(Gal) [c]</td>
</tr>
</tbody>
</table>

\[a\] Glycosylation sites are marked on the peptide sequence with a hash sign. \[b\] GlcNAc. N-acetylglucosamine; HexNAc. N-acetylgalactosamine; Man. mannose; Gal, galactose; Hex, mannose or galactose; Fuc, fucose; Xyl, xylose. \[c\] Two cysteine are linked by a disulfide bond. \[d\] Asterisk represents cysteine carboxymethylation by IAA.

### Conclusions

In summary, we have synthesized SnO\textsubscript{2}@Poly(HEMA-co-St-co-VPBA) core-shell nanoparticles for selectively enriching glycopeptides from the peptide mixtures, and developed a time-saving procedure to recover the
glycopeptides for MS analysis. The enrichment mechanisms are based on the reversible boronic esterification between the nanoparticle boronic surface groups and 1,2-cis-diol from glycopeptides, and thus, glycopeptides are recovered with high selectivity and efficiency. Simultaneous identification of the glycosylation site and glycan structure can be achieved by mass spectrometric analysis of the enriched intact glycopeptide from both standard and real biological samples.

**Experimental Section**

**Synthesis of SnO$_2$@Poly(HEMA-co-St-co-VPBA) Nanoparticles**

The freshly prepared SnO$_2$ nanoparticles, which were obtained by hydrolyzing a 0.5 m SnCl$_4$ aqueous solution in a poly(tetrafluoroethylene) (PTFE)-lined autoclave at 160°C for 2 h, were centrifuged, and then the white precipitates were dispersed in liquid 2-hydroxyethyl methacrylate (HEMA, Acros Organics) monomers. The solutions were precipitated and washed by absolute ether as a nonsolvent. The product was designated as SnO$_2$-HEMA in which HEMA groups were connected to the SnO$_2$ nanoparticles through a 0.5 m aqueous solution (10 mL) containing 1% trifluoroacetic acid (TFA) which was added into the washed deposit to release the glycopeptides from the nanoparticles for 1 h. In addition, the supernatant from centrifugation after incubation or washing was concentrated and re-dispersed by acidic aqueous solution (10 mL) containing 1% trifluoroacetic acid (TFA) which was added into the washed deposit to release the glycopeptides from the nanoparticles for 1 h. In addition, the supernatant from centrifugation after incubation or washing was concentrated and re-dispersed by acidic aqueous solution (10 mL) containing 1% TFA.

**MALDI MS Analysis**

For standard proteins, every 2 mL of eluate from concentrated supernatant was dried on the MALDI plate. For the real samples from human serum, all the eluate from each sample was spotted. The supernatants from centrifugation after incubation or washing was concentrated and re-dispersed by acidic aqueous solution (10 mL) containing 1% TFA.

**Acknowledgements**

This study was supported by the National Science and Technology Key Project of China (2007CB914100, 2008ZX10207, 2009CB825607, and 2006AA02A308), the National Natural Science Foundation of China (20735005, 20875016, 30672394, and 20873029), the Ph.D. Programs Foundation of the Ministry of Education of China (20080246001), NCET, and Shanghai Projects (08YZ2293601, Shu Guang, Eastern Scholar, 09QA1400400 and B109).