Size-selective proteolysis on mesoporous silica-based trypsin nanoreactor for low-MW proteome analysis†

Qianhao Min,†a Ren’an Wu,†a Liang Zhao,†a Hongqiang Qin,†a Mingliang Ye,†a Jun-Jie Zhu*b and Hanfa Zou*a

Received 27th March 2010, Accepted 28th June 2010
DOI: 10.1039/c0cc00619j

In this study, the concept of size-selective proteolysis was first described by using the mesoporous silica-based trypsin nanoreactor. For analysis of a complex protein sample, low-MW proteins were preferentially digested for identification while high-MW proteins were excluded from digestion.

The low-molecular weight (MW) proteome has drawn great attention in understanding biological systems for its adequate information of probable biomarkers and signal molecules in biology and pathology.1,2 Due to the interference from high-MW proteins, the determination of low-MW proteins in biofluids and tissue extracts has become even more difficult. To date, advances in isolating low-MW proteins from complex biological samples followed by enzymatic digestion for the low-MW proteome analysis have been achieved by utilizing accurate MW ultrafiltration,3 acetonitrile precipitation,3 SDS-PAGE4 and silica material extraction.5 Whereas, the strategy in analyzing low-MW proteins undergoing a multistep procedure of isolation followed by proteolysis is considered to be time- and labor-consuming, and particularly not well-suited for minute biological samples because of the probable sample loss. Mesoporous silicas (MPSs) (typically SBA-156,7 have played a significant role in size-selective separation of biomolecules based on the size-exclusive interaction8–10 for their distinct mesopore structures. Besides benefiting from their high surface areas and large pore volumes, MPSs have been employed as promising supports for enzyme immobilization.11–14 Moreover, in proteomics application, increased digestion rate, improved sequence coverage and enhanced detection sensitivity have been realized by MPS-based enzyme reactors.15,16 Nevertheless, the size-selective enzymatic digestion specifically for the proteolysis of low-MW proteins in complex protein mixtures using MPSs, to the best of our knowledge, has never been investigated.

Herein, we establish an MPS-based “one-step” size-selective enzymatic digestion mode for achieving the highly efficient proteome analysis of low-MW proteins from a complex protein mixture by combination of size-exclusive separation and enzyme immobilization on a thiol-modified SBA-15 material (SBA-15-SH). The SBA-15 materials used in this study were characterized by TEM and N2 sorption. As shown in Fig. 1, the pore sizes of the SBA-15 materials before and after thiol modification are 6.5 nm and 5.7 nm, respectively, which are the pore sizes to prevent the entry of high-MW proteins (e.g., bovine serum albumin (BSA) with MW = 66 400 Da and size = 5.0 × 7.0 × 7.0 nm3).12 The size-exclusive effect of the MPS carrier on proteins with distinct sizes can be obtained owing to the mesopore structures.

To achieve the size-selective enzyme digestion, trypsin (3.8 × 3.8 × 3.8 nm3)12 was immobilized on calcinated SBA-15 via physical adsorption. However, the leaching of trypsin from SBA-15 material definitely interferes with the size-selective digestion of enzyme toward the substrate. Thus, thiol groups were introduced to simply minimize the leaching of immobilized trypsin (see Fig. S1 ESI† for comparison experiment). Using the thiol functionalized SBA-15 material (SBA-15-SH), four proteins, Cytochrome c (Cyt c) (MW = 12 384 Da, 2.6 × 3.2 × 3.3 nm3), lysozyme (MW = 14 388 Da, 1.9 × 2.5 × 4.3 nm3), myoglobin (MW = 17 000 Da, 2.1 × 3.5 × 4.4 nm3) and BSA (MW = 66 400 Da, 5.0 × 7.0 × 7.0 nm3),12 with different molecular sizes were chosen as the model compounds to examine the size-selective proteolysis of the SBA-15-SH enzyme reactor toward proteins. In the model illustrated in Scheme 1, the proteins of Cyt c, lysozyme and myoglobin with sizes less than 5.0 × 5.0 × 5.0 nm3 are supposed to be entrapped in the nanochannels of SBA-15-SH and digested; while BSA with a larger size is supposed to be unable to enter the mesopores of the support for subsequent proteolysis.

To verify the size-selective interaction of the above-mentioned proteins with MPS SBA-15-SH, the adsorption of

![Fig. 1 Pore size distributions calculated from the desorption branch of nitrogen adsorption/desorption isotherm (Inset: TEM images) of SBA-15 (A) before and (B) after thiol modification.](Image)
these proteins on SBA-15-SH was first carried out (Fig. S2, ESI†). The smallest proteins of Cyt c and lysozyme demonstrate the ultrafast adsorption rates (reach equilibrium in less than 10 min, adsorption curves A and B). Myoglobin, with a little larger dimension, reaches the adsorption plateau in a longer time (ca. 5 h). For the biggest protein, BSA, little adsorption was observed within a 24 h incubation period. Thereupon, it can be concluded that the SBA-15-SH did exhibit a size-dependent adsorption performance toward proteins with different MWs. With the size-selective effect, proteins with dimensional sizes less than the pore size of SBA-15-SH will be allowed to enter the nanochannels and consequently be degraded into peptide fragments by enzyme inside the mesopores.

To further confirm the expected size-selective proteolysis, SBA-15-SH with immobilized trypsin was applied to digest protein mixtures composed of BSA and either Cyt c, lysozyme or myoglobin with concentration of 5 nmol ml⁻¹ for each protein. For comparison, these protein pairs were also digested by free trypsin in solution. Tryptic digests of all protein pairs were analyzed by MALDI-TOF mass spectrometry (Fig. 2). With free trypsin, as demonstrated in Fig. 2A, peptides from BSA dominate the MALDI-TOF mass spectrum in the protein pair of Cyt c and BSA. In contrast, using SBA-15-SH immobilized trypsin, the detected peptides are most attributed to the small protein of Cyt c with only few BSA peptide fragments observed in the same spectrum. In a data summary (Table S1, ESI†), 10 Cyt c peptides were determined using SBA-15-SH immobilized trypsin, but instead only 5 Cyt c peptides in the case of free trypsin. In contrast, no more than 2 BSA peptides were resolved for using the SBA-15-SH trypsin reactor, which were 28 for free trypsin. Replacing Cyt c with lysozyme, similar results were observed (Fig. 2B) for both approaches. Namely, the SBA-15-SH immobilized trypsin demonstrates the proteolysis selectivity to the small protein of lysozyme rather than the big protein of BSA. Applied to the myoglobin/BSA protein pair, the SBA-15-SH trypsin reactor demonstrates the preferential digestion toward myoglobin rather than BSA (Fig. 2C). All these results indicate that proteolysis of low-MW proteins on the mesoporous SBA-15-SH trypsin reactor is size-dependent digestion which results from the size-exclusive interaction of the mesopores against the high-MW protein of BSA. In this manner, the peptide information from low-MW proteins will be selectively enhanced by the size-selective proteolysis. On the other hand, the proteolysis of high-MW proteins will be inhibited, yielding little peptide information to complicate the mass spectrum.

To confirm the improvement of protein identification (specifically for low-MW proteins) by the size-selective proteolysis, the proteolytic products of the three protein pairs were analyzed by nanoscale liquid chromatography tandem mass spectrometry (nano-LC-MS/MS) followed by database searching. The identified unique peptides and sequence coverages are presented in Table S4 ESI†. The apparently increased unique peptide numbers and the improved sequence coverages for these low-MW proteins of Cyt c, lysozyme and myoglobin indicate the superior proteolysis efficiency of the SBA-15-SH enzyme reactor to low-MW proteins. With even lower concentration of Cyt c in the mixture of Cyt c and BSA (mole ratio of Cyt c-to-BSA at 1:10000), Cyt c could still be confidently identified by the trypsin nanoreactor, while the mole ratio of Cyt c-to-BSA at 1:300 is the limited level for the

![Scheme 1 Flowchart of size-selective proteolysis of the protein pair of high-MW protein (BSA) and low-MW protein (Cyt c, lysozyme or myoglobin).](image-url)

![Fig. 2 MALDI-TOF spectra of tryptic digests of (A) Cyt c, (B) lysozyme and (C) myoglobin mixed with BSA after being treated by trypsin immobilized on SBA-15-SH and free trypsin at 37 °C for 12 h. Peptide peaks at relatively high S/N are labeled with “*” and “**”. “*” is assigned to peptides from digest of BSA; “**” is assigned to peptides derived from (A) Cyt c, (B) lysozyme and (C) myoglobin, respectively.](image-url)
identification of Cyt c by the conventional trypsin strategy (Table S5 and S6, ESI†). Nevertheless, it must be noticed that the absolute exclusion of high-MW protein by the MPS materials is hard to achieve as the possible adsorption of trypsin on the external surface and the pore opening. Ascribed to this defect, in this study, we observed that the prolonged incubation time would help the digestion of high-MW protein such as BSA in our case (Table S7 ESI†), which thus weakened the size-exclusive effect of the enzyme reactor to high-MW proteins. Hence, shortening the digestion time was an effective way to avoid the extra digestion of high-MW proteins. In this work, 1 h with approximately maximum identified unique peptides and sequence coverage was optimized for the size-selective digestion of low-MW proteins, which either ensures the sufficient proteolysis of low-MW proteins or maximally reduces the possible digestion of high-MW proteins (Table S7 ESI†).

Inspired by the size-selective digestion of the SBA-15-SH enzyme reactor toward low-MW proteins, we further applied this reactor to the selective proteolysis of low-MW proteins in a real complex biological sample (human serum, with proteins of MWs from $10^3$ to $10^5$Da, and protein concentrations ranging from 1 to $10^{10}$ pg ml$^{-1}$). After 1 h trypptic digestion, peptides originating from human serum proteins were analyzed by nano-LC-MS/MS, followed by spectral counting analysis. As a comparison, free trypptic digestion went along with the identical conditions. As shown in Fig. 3A, most identified proteins with the SBA-15-SH trypsin reactor are of low-MW proteins (less than 40 kDa, which is the MW cutoff in agreement with the literature) in human serum. Conversely, the proteins with MWs mainly located in the range 60–80 kDa dominate the spectral counting distribution after incubation with free trypsin (Fig. 3B). In particular, as listed in Table S8 ESI†, a spectral count of 1076 for human serum albumin (HSA), with overwhelming abundance, was obtained by using free trypsin, which decreased to 14 (ca. two orders of magnitude lower) after exposure to the SBA-15-SH trypsin reactor. In addition, other significant proteins in human serum (with MWs in the range of 40–80 kDa), such as transferrin, α1-antitrypsin and haptoglobin, did not produce their peptides. On the contrary, proteins with MWs less than 40 kDa produced their accordant peptides with certain identified spectral counts by immobilized trypsin (Table S8 ESI†).

By the SBA-15-SH trypsin reactor than free trypsin seem to be somewhat at odds with the size-selective digestion, the fragments of complement C3 generated by endogenous proteases in serum could explain this observation. This discriminating digestion of low-MW proteins by the mesoporous SBA-15-SH trypsin reactor is attributed to size exclusion of the nanochannels toward the high-MW proteins.

In conclusion, size-selective proteolysis of proteins with diverse MWs has been uncovered and described by constructing an enzyme reactor based on mesoporous materials, which incorporates the size-exclusive interaction of mesopore structure toward high-MW proteins with the in situ proteolysis by the immobilized enzyme within the pore area. Importantly, this in situ digestion approach integrates the separation and digestion into a one-step operation, thus realizing the rapid proteolysis of low-MW proteins from complex biological samples, particularly for the minute protein samples, without the conventional tedious pretreatment procedures.

This work was supported by the National Natural Science Foundation of China (No. 20735004, 20875089 and 20635020), the State Key Basic Research Development Program of China (973 Program) (No. 2007CB914102), the National High Technology Research, the Development Program of China (863 Program) (No. 2008AA02Z211), and the Knowledge Innovation Program of the Chinese Academy of Sciences (No. KSCX2-YW-G-046).

Notes and references