

**Poly (*N*-acryloxysuccinimide-co-ethylene glycol dimethacrylate) precursor monolith and its post polymerization modification with alkyl ligands, trypsin and lectins for reversed-phase chromatography, miniaturized enzyme reactors and lectin affinity chromatography, respectively**

Murthy Jonnada and Ziad El Rassi\*

Department of chemistry, Oklahoma State University, Stillwater, OK, 74078

\*Corresponding author: Regents Professor Ziad El Rassi, Department of Chemistry, Oklahoma State University, Stillwater, OK 74078-3071, USA. E-mail: [elrassi@okstate.edu](mailto:elrassi@okstate.edu). Fax: (405) 744-6007.

**Key words:** Nonpolar monolith; Immobilized enzyme monolith; Lectin monoliths; Microcolumns; Human serum

**Nonstandard abbreviations:** **ABs**, alkylbenzenes; **Con A**, concanavalin A; **EDMA**, ethylene glycol dimethacrylate; **GlcNAc**, N-acetyl glucosamine; **IMER**, immobilized enzyme reactor; **LAC**, lectin affinity chromatography; **LCA**, *Lens culinaris* agglutinin; **Me- $\alpha$ -Man**, methyl- $\alpha$ -D-mannopyranoside; **NAS**, *N*-acryloxysuccinimide; **NASM**, poly(NAS-co-EDMA) monolith; **PPM**, post polymerization modification; **RCA**, *Ricinus communis* agglutinin; **TFA**, trifluoroacetic acid.

### Abstract

This investigation was aimed at introducing a monolithic precursor that can be conveniently grafted with the desired chromatographic ligand *via* the process of post polymerization modification (PPM). The precursor was obtained by the *in-situ* polymerization of *N*-acryloxysuccinimide (NAS) and ethylene glycol dimethacrylate (EDMA) in a narrow bore stainless steel column of 1 mm i.d. yielding a poly(NAS-co-EDMA) monolithic column designated as the NASM column. In a first PPM, the NASM

Received: 05 30, 2017; Revised: 07 03, 2017; Accepted: 08 01, 2017

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1002/elps.201700221](https://doi.org/10.1002/elps.201700221).

This article is protected by copyright. All rights reserved.

column was bonded with octadecyl (OD) ligands yielding a nonpolar NASM-OD column that proved useful for RPC of proteins in gradient elution at increasing %ACN in the mobile phase. NASM-OD resulted from the reaction between the *N*-hydroxysuccinimide of NASM with octadecyl amine. In a second PPM, NASM was surface immobilized with trypsin generating a proteolytic narrow bore enzyme reactor called NASM-trypsin IMER that permitted the online digestion of proteins in a 20-min single pass through the IMER incorporated in a setup equipped with a short RPC column to achieve simultaneously a peptide tryptic map. This constituted a rapid turnover whereby ~ 95% of the protein was hydrolyzed by the immobilized trypsin. In a third PPM, the NASM column was surface immobilized with three different lectins (LCA, Con A and RCA) having complementary affinities toward serum glycoproteins thus permitting the capture of a wide range of glycoproteins/glycoforms. The three NASM-lectin columns when operated in a tandem format led to assessing the level of the various glycoforms in human serum *via* LC-MS/MS analysis of the captured protein fractions by each NASM-lectin column.

## 1 Introduction

Organic monolithic HPLC columns are becoming increasingly popular due to, among other things, (i) easy *in situ* preparation in columns or separation channels of different dimensions and (ii) tunable flow characteristics by adjusting the monomers to porogens ratio [1-4]. Monolithic columns with narrow inner diameter (i.d.) tubing represent a new trend in HPLC, which can be termed as "micro-scale HPLC" [5]. They have characteristic of low dead volume and separation efficiencies comparable to that of conventional HPLC columns within their typical low flow rates. The low flow rates and small column dimensions decrease the consumption of mobile phase and stationary phase, respectively, which are the main attractive feature of micro-scale HPLC. However, HPLC systems which use micro-bore columns with diameters of 200-500  $\mu\text{m}$  i.d. require specialized pumps and detector flow cells. To bypass the requirement of specialized instrumentation, an intermediate class was introduced, which can be termed as semi-micro columns with 1-2 mm i.d., and separations can be performed on a standard HPLC system with minor changes in the tubing (e.g., injection loop volume) to decrease the pre-column dead volume [5]. The main advantage of these semi-micro columns is that they consume far less solvent than the conventional HPLC columns having 4.6 mm i.d. thus leading to a reduction in the cost of analysis and yielding higher resolution by using longer columns [6].

The use of semi-micro columns with 1-2 mm i.d. dates back to the early history of HPLC in the mid-seventies [6, 7]. However, the lack of efficient packing methods has hindered the rise in using narrow bore columns which were soon substituted by  $\frac{1}{4}$ " columns (4.6 mm i.d.). But with the advent of polymeric based monolithic columns, the difficulties which arise from the packing of narrow bore columns with particles was eliminated, since the column material is formed by *in situ* polymerization.

In this research report, a narrow bore stainless steel HPLC column of 1 mm i.d. was filled with a monolithic precursor by the *in-situ* copolymerization of *N*-acryloxysuccinimide (NAS) and ethylene glycol dimethacrylate (EDMA). This NAS-based precursor monolith is referred to as NASM, and is characterized by its surface succinimide groups that can readily bind to various compounds with primary amines as well as to the lysine residues in proteins. Despite the fact that this type of (bio)-conjugation reaction between succinimide and amines is well known and has been exploited for a long time [8, 9], such an active functionality has found a rather scarce use in monolithic column technology in the so-called post polymerization modification (PPM) of monoliths with the desired chromatographic ligands. In fact, NASM prepared in fused silica capillaries of 75  $\mu\text{m}$  i.d. has been grafted with C4-C8 alkyl ligands by reacting the NASM with C4-C8 alkyl amines [10, 11], and the resulting columns were used in reversed-phase CEC of alkylbenzenes and other aromatic compounds. Also, NASM was converted to a zwitterionic surface by reacting the monolith with alkyl diamines and hydrolyzing the unreacted *N*-hydroxysuccinimide units to carboxylic acid groups yielding a HILIC/RP mixed mode column for CEC separations [12]. In another article, the NASM was subjected to PPM by a two steps click chemistry with propargyl amine first and then  $\beta$ -cyclodextrin to yield a chiral monolithic column [13] for the separation of flavanone enantiomers.

In this investigation, the NASM, which was produced from NAS and EDMA by a thermally initiated *in situ* polymerization process, was exploited in three different PPMs. In one PPM, the NASM was functionalized with octadecylamine yielding NASM with surface bound octadecyl (OD) ligands referred to as NASM-OD. The nonpolar NASM-OD was demonstrated in reversed phase chromatography (RPC) of nonpolar test solutes such as alkylbenzenes (ABs) and polyaromatic hydrocarbons (PAHs). Also, standard proteins were separated by a linear gradient elution at increasing %ACN (v/v) in the mobile phase.

To address the complexity of protein glycosylation, which is one of the most important posttranslational modifications of proteins [14], another PPM strategy was introduced whereby the NASM was surface immobilized with lectins, which are proteins of non-immune origin and are known to have affinity towards particular glycoconjugates thus serving as specific reagents for glycoprotein enrichment from biological fluids [15-17]. Three different lectins, namely concanavalin A (Con A), *Lens culinaris* agglutinin (LCA) and *Ricinus communis* I agglutinin (RCA) were immobilized on NASM yielding NASM-Con A, NASM-LCA and NASM-RCA columns, respectively. Since the glycans in glycoproteins are generally very complex and there exists not a single lectin that would cover all the glycoproteome in a given sample, a need for a multiple lectin column strategy emerges. So, in general an efficient approach to capture a wide range of glycoproteins has been the use of tandem lectin columns [16, 18] whereby a few lectin columns with different carbohydrate binding specificities are connected in series and are used to capture different glycoproteins from a given biological fluid (e.g., human serum). In this process, the serum is usually passed through the tandem lectin columns, and after washing the columns extensively with the binding mobile phase to remove all nonspecifically bound proteins, the tandem columns are disconnected from each other and each lectin column is eluted separately with a mobile phase containing a specific haptenic sugar for the

particular lectin. In brief, the three NASM-lectin columns were connected in the order NASM-LCA → NASM-Con A → NASM-RCA. LCA is known to interact with  $\alpha$ -mannosyl residues on *N*-glycans, but the *N*-acetyl glucosamine at the reducing terminal should be substituted with a fucose at its C-6 hydroxyl [19, 20]. High mannose and hybrid type glycans have affinity towards Con A [21, 22]. RCA generally binds to glycoproteins bearing branched *N*-glycans such as bi-, tri-, and tetra-antennary complex type glycans [21, 23, 24]. The haptenic sugar for LCA and Con A is the methyl- $\alpha$ -D-mannopyranoside (Me- $\alpha$ -Man) and that for RCA is lactose. Thus, it is expected that the tandem column order NASM-LCA → NASM-Con A → NASM-RCA with complimentary specificities will capture glycoproteins with the broadest glycosylation possible from human serum.

To investigate another important PPM of NASM, which would be beneficial for MS analysis of protein digest and may find use in proteomics profiling of biological samples, NASM with immobilized trypsin was introduced for the simultaneous online tryptic digestion and peptide mapping by RPC. This involved the development of NASM-trypsin immobilized enzyme reactor (IMER) referred to as NASM-trypsin IMER in narrow bore HPLC column, which can handle very small sample volumes and has the potential to be directly connected to MS to reduce the operator intervention and enhance throughput. IMER offers many distinct advantages compared to the in-solution digestion including the stability to harsh environments such as pH, temperature and mild organic solvents. Furthermore, IMER generally offers high turnover numbers compared to in-solution digestion. Currently, many publications have emerged which reported IMER made from beads, nanoparticles, membranes and other monolithic materials [25-27].

In summary, the originality of this investigation resides in 3 major developments including (i) extending the PPM of NASM with nonpolar ligands to long *n*-alkyl ligands (e.g., octadecyl ligands), (ii) introducing the PPM of NASM with a proteolytic enzyme (e.g., trypsin) to realize IMER in a narrow bore HPLC column and (iii) investigation of the strategy and the benefits of PPM of NASM with bio-ligands, e.g., lectins, for lectin affinity chromatography of glycoconjugates.

## 2 Experimental

See Supporting Information (SI) for Reagents and materials, Instrumentation, Fractionation of human serum glycoproteins by lectin affinity chromatography, Digestion of protein fractions by trypsin, LC-MS/MS methodology and LC-MS/MS data analysis.

### 2.1 Preparation of NASM in narrow bore columns

The monolithic precursor NASM was prepared by *in situ* polymerization in a stainless-steel column as follows: a 25-cm x 1 mm i.d. column bearing end fittings was filled with the polymerization reaction mixture of 2.0 g total weight consisting of 18.0% (w/w) NAS functional monomer, 15.8% (w/w) EDMA crosslinker, 66.2% (w/w) toluene as the porogen and 1% (w/w) of the initiator AIBN with respect to the monomers. The weight of AIBN was not taken into account when calculating the percentages of the reaction mixture. The polymerization mixture was then

transferred to the column with a syringe fitted with an appropriate connection to the column, which was kept vertical and the mixture was introduced into the column against gravity to ensure proper filling of the column and to avoid air bubble formation. After sealing the column at both ends, the polymerization was carried out at 50 °C in a water bath for 15 h. Thereafter, the column was flushed with ACN to remove the porogen and the unreacted monomers. This was followed by equilibrating the column with two column volumes of toluene before proceeding to PPM.

## 2.2 PPM of NASM with octadecyl ligands

The NASM was further modified with octadecylamine to form the NASM-OD (see Fig. 1, Path I). The concentration of octadecylamine solution prepared in toluene was 0.04 g/mL, which was pushed through the column for about two column volumes. The octadecylamine concentration was maintained in large excess in order to ensure high surface coverage of the NASM with OD ligands. Then, the column was washed with methanol to remove the unreacted octadecylamine and toluene. This was followed by scavenging the unreacted succinimide functions of the NASM by passing a 10% (v/v) methylamine solution in methanol through the column (Fig. 1, Path I). Typically, a 3.5 cm void was observed at the inlet of the column, which was cut yielding a column of 21.5 cm total length. After inserting a new ferrule at different position with a column end fitting, the shortened 21.5 cm column was then flushed with ACN before use.

## 2.3 Immobilization of proteins on NASM

All PPM reactions were done on column by injecting an adequately prepared solution of a given protein (e.g., trypsin or lectins) into the column by means of a syringe pump. The 1 mm i.d. narrow-bore HPLC column filled with NASM was washed with 5 column volumes each of water and 5 mM phosphate buffer, and then 1 mL of 10 mg/mL protein solution (trypsin/lectin) prepared in 20 mM phosphate buffer was injected for a period of 4 h. The lectin solutions had also their corresponding hapten sugars dissolved in the just-mentioned buffer in order to preserve the sugar-binding site of the given lectin. Con A and LCA were prepared in the above buffer containing 0.1 M Me- $\alpha$ -Man and 1mM of CaCl<sub>2</sub>, MgCl<sub>2</sub> and MnCl<sub>2</sub>, while the RCA solution was prepared with 0.1 M of its hapten sugar lactose. Proteins were immobilized onto the NASM column *via* the reaction of its *N*-hydroxysuccinimide surface groups with the primary amines of the side chain of lysine residues of the protein forming stable amide linkages. The residual succinimide groups of the column were then scavenged with 0.1 M ethanolamine/acetic acid buffer, pH 9.0, see Fig. 1, Path II. The entire protein immobilization setup was placed in a cold room maintained at 4 °C. While not in use, the columns thus prepared were stored at 4°C in the binding buffer containing 0.02% NaN<sub>3</sub>.

## 3 Results and discussion

### 3.1 NASM-OD

**Preparation.** The schematic of NASM-OD preparation is shown in Fig. 1, Path I. It should be noted that the NASM column was intended to have high mechanical stability to suit HPLC running conditions, which usually involve high backpressures at relatively high flow velocities. This was done

by increasing the percentage of EDMA crosslinker with respect to NAS functional monomer. The previously reported NASM was prepared for use in capillary electrochromatography (CEC) [11], and the ratio of NAS to EDMA was designed to suit the needs of CEC applications where the monolith does not encounter high backpressure. In the current work, the total percentage of monomers was 33.8% (w/w) with the ratio of NAS to EDMA being maintained at 1.33 and the single porogen toluene occupied the remaining 66.2% (w/w). Under these fabrication conditions, the NASM was mechanically stable for many consecutive runs.

For carrying out the PPM reactions, three different solvents namely toluene, dioxane and THF were examined in order to select the best solvent. The retention factors  $k$  of four ABs including toluene, ethylbenzene, propylbenzene and butylbenzene, were estimated with an NASM-OD column (21.5 cm x 1 mm i.d.) running a mobile phase of water:ACN 47.5:52.5 (v/v) at a flow rate of 0.1 mL/min. The NASM-OD prepared in toluene yielded  $k$  values of 1.27, 1.51, 2.05 and 2.80 for toluene, ethylbenzene, propylbenzene and butylbenzene, respectively, which were very similar to those obtained on the column modified in the presence of dioxane (compare to  $k$  values of 1.27, 1.50, 2.03 and 2.80 for the same ABs), but the column modified in toluene yielded a better peak shape. The column modified in the presence of THF gave lower retention and the peaks of toluene and ethylbenzene co-eluted. The  $k$  values for the 4 ABs were 1.24, 1.24, 1.70 and 2.40 for toluene, ethylbenzene, propylbenzene and butylbenzene, respectively. Thus, toluene was the chosen solvent for the rest of this investigation.

**Evaluation of NASM-OD column with small and large solutes.** This was done with hydro-organic mobile phases at various %ACN (v/v) in water using ABs as the test solutes. Figure 2A shows the retention behavior of ABs in terms of  $\log k$  vs. %ACN (v/v) in the % range of 35-52.5.  $\log k$  values decreased in a linear fashion ( $R^2$  values in the range 0.9972-0.9984) with increasing the %ACN in the mobile phase, which is indicative of RPC behavior. The magnitude of the slopes increased from 0.032 to 0.048 when going from toluene to butylbenzene indicating an increase in the extent of interaction with increased solute size. Linear curves were obtained when plotting  $\log k$  against the number of carbon in the ABs homologues series as shown in Fig. 2B, with slopes corresponding to the log of methylene group selectivity ( $\log \alpha_{\text{CH}_2}$ ).  $\alpha_{\text{CH}_2}$  values were calculated by extracting the antilog of the

slopes of the lines. As expected, the  $\alpha_{\text{CH}_2}$  values increased from 1.30 to 1.62 when the % (v/v) of ACN in the mobile phase decreased from 52.5% to 35% (v/v). A typical chromatogram for the separation of six alkyl benzenes is shown in Fig. 3A. It should be noted that the unmodified NASM column (i.e., blank column without OD ligands) did not show any retention for ABs, and all these test solutes eluted at the deadtime of the column.

For further characterization, the retention ability of NASM-OD column was tested with some polyaromatic hydrocarbons using a relatively elevated flow rate of 0.3 mL/min. Figure 3B shows the separation of 1-nitronaphthalene, fluorene, phenanthrene and pyrene. Pyrene being the most nonpolar with four benzene rings fused together eluted last.

As shown in Figs 3A and 3B, the NASM-OD column yielded rather broad peaks for small solutes. It was therefore very tempting to evaluate the column with large molecules such as proteins since polymer based monolithic columns are well suited for proteins separations owing to the absence of micropores in the monolithic structure [4]. In this regard, the NASM-OD column was tested with standard proteins using a 30-min linear gradient elution at increasing % ACN in the mobile phase at a flow rate of 0.1 mL/min. A mixture of six standard proteins consisting of ribonuclease A ( $M_r$ , 13.7 kDa; pI, 9.6), cytochrome C ( $M_r$ , 14.4 kDa; pI, 10.0), lysozyme ( $M_r$ , 14.3 kDa; pI, 9.3), conalbumin ( $M_r$ , 75.0 kDa; pI, 6.6),  $\beta$ -lactoglobulin A ( $M_r$ , 18.4 kDa; pI, 5.1) and human serum albumin ( $M_r$ , 66.0 kDa; pI, 4.7) was readily separated in 30 min, and the chromatogram is shown Fig. 3C.

*Reproducibility.* The reproducibility of analysis on the NASM-OD column was evaluated using the AB test solutes with a mobile phase of 50% (v/v) ACN at a flow rate of 0.1 mL/min. The  $k$  values were measured from run-to-run ( $n = 4$ ) for toluene, ethylbenzene, propyl benzene and butyl benzene, which yielded %RSD values of 1.27, 1.69, 1.80 and 2.52, respectively. Similarly, the column-to-column reproducibility was evaluated ( $n = 2$ ) using the same AB test solutes, which gave %RSD values of 0.86, 6.26, 6.76 and 5.91 for toluene, ethyl benzene, propyl benzene and butyl benzene, respectively.

### 3.2 Simultaneous on-line protein digestion by NASM-trypsin IMER and peptide mapping by RPC

The narrow bore NASM-trypsin IMER was evaluated in the proteolytic digestion of a typical standard protein, namely cytochrome C (Cyt C). In general monoliths, which are known to have interconnected networks (macropores) with small microglobules (mesopores) [1, 2] should in principle provide the perfect media for fast substrate (i.e., Cyt C) mass transfer and in turn the immobilized trypsin should result in an effective proteolytic digestion of the given protein. Cyt C of  $M_r$  12.4 kDa contains 104 amino acids with 20 trypsin cleavage sites that would yield 21 peptides of which 15 peptide fragments contain 2 or more amino acids ([http://web.expasy.org/peptide\\_cutter/](http://web.expasy.org/peptide_cutter/)).

Cytochrome C was first dissolved in an appropriate amount of the digestion mobile phase, and the whole process of the online digestion and RPC peptide mapping is explained in the legend of the instrumental set-up in Figs 4A and 4B, where also can be found details about the mobile phase composition, the direction of the mobile phase flow and the sequential operation of the valves in the set-up. During RPC gradient elution, the IMER was by-passed to prevent the immobilized trypsin to get in contact with ACN. The peptide map obtained for 20  $\mu$ g of injected Cyt C is shown in Fig. 4C using a two-step gradient. As shown in Fig. 4C, one can count 15 major peaks (indicated by arrows) and some 7 minor peaks indicated by asterisks. The 7 short peaks may be attributed to the non-specific cleavages, which seem to be very tiny peaks/fragments due to the relatively short single pass online digestion time of 20 min. The last peak on the map is that of undigested Cyt C, which

regardless of the amount digested of Cyt C in the range 2 – 20 µg represented less than 5% of the injected amount, thus indicating that the prepared narrow bore NASM-trypsin IMER provided high digestion efficiency at various injected concentrations of the protein under investigation. In addition, the online integration of NASM-trypsin IMER with the RPC column demonstrated the minimum operator intervention and a straightforward workflow.

### 3.3 Evaluation of NASM-lectin columns in the capturing of glycoproteins from disease free serum

First, the three NASM-lectin columns were tested for their nonspecific and specific binding towards standard glycoproteins and non-glycoproteins. The 3 columns showed no binding towards myoglobin, a known standard non-glycoprotein, which was eluted at  $t_0$  (chromatograms not shown). Then, the NASM-Con A column, which is known to have affinity towards high-mannose type glycans, hybrid type glycans and to a lesser extent bi-antennary complex type glycans (see Fig. 1S in SI) retained ribonuclease B (Fig. 5A) that is rich in high-mannose type glycans [28, 29]. On the other hand, a small fraction of ovalbumin was retained on the NASM-Con A column, as it eluted with 0.1 M Me- $\alpha$ -Man hapten sugar, indicating that this fraction may be attributed to ovalbumin glycoforms' with surface bound hybrid type glycans and to a lesser extent high mannose type glycans (Fig. 5B), which corroborate very well with an earlier finding by Harvey et al. on the composition of the various glycans in purified ovalbumin [30]. When human IgG was injected on the NASM-LCA column, a portion of IgG showed affinity toward immobilized LCA, which was eluted with 0.1 M Me- $\alpha$ -Man (Fig. 5C). This indicates the presence of IgG glycoforms with biantennary core fucosylated type glycans on their surface (for glycan structures, see Fig. 1S in SI), which is in agreement with the reported literature in a recent review article [31]. The NASM-RCA was tested with human transferrin (Fig. 5D) whereby some of the glycoforms were bound to the lectin column and eluted upon applying the eluting mobile phase containing of 0.1 M lactose. Most likely, the portion of human transferrin retained on the NASM-RCA column may be attributed to transferrin glycoforms with biantennary and triantennary complex type glycans on their surface [32, 33] toward which immobilized RCA has affinity (For the structures of glycans with affinity to RCA, see Fig. 1S in SI)

Thereafter, the 3 lectin columns were used in a tandem configuration in the order NASM-LCA  $\rightarrow$  NASM-Con A  $\rightarrow$  NASM-RCA to capture glycoproteins from disease-free serum. The order by which the 3 lectin columns were connected in series parallels the increasing order of their binding selectivity toward glycans; that is in the order of broader selectivity. While the NASM-LCA column, which is known for its narrowest selectivity and binds primarily to glycoproteins containing bound bi- and triantennary glycans with core fucosylation (see Fig. 1S in SI) was placed first in the series, NASM-Con A which has a wider selectivity as it binds to glycoproteins containing high mannose-type glycans and hybrid type glycans (see Fig. 1S in SI) was placed the second column in the series, and the NASM-RCA column with the widest selectivity as its binds to glycoproteins containing a variety of glycan structures (refer to Fig. 1S in SI) was placed as the last column in the tandem column series.

After injecting 20 µL of 1:3 diluted disease free serum, the unbound serum protein fraction was eluted first; Fig. 2SA in SI. Thereafter, the 3 lectin columns were disconnected from each other,



and each lectin column was eluted separately with its appropriate hapten sugar. The eluted glycoproteins which were captured by NASM-LCA, NASM-Con A and NASM-RCA columns from the injected serum are shown in Figs **2SB**, **2SC** and **2SD**, respectively, in **SI**. The pooled fractions collected from each lectin column after 7 consecutive serum injections were concentrated and the resulting protein pellets were subjected to trypsin digestion followed by LC-MS/MS analysis (see LC-MS/MS methodology and the data analysis sections in **SI**). Only proteins that exhibited protein and peptide identification probability of at least 99.9% and 95%, respectively, and contained at least three unique peptides were considered (see Table **1S** in **SI** for the list of identified proteins). A total of 64 non-redundant proteins were identified in the pooled protein fractions collected from the three lectin columns. The NASM-LCA column captured a total of 22 proteins while a total of 49 proteins were identified in the NASM-Con A column captured fraction and for the NASM-RCA column, a total of 57 proteins were identified in the captured fraction from disease-free serum, which may represent in major part glycoproteins with highly branched glycome. Human serum albumin, a nonglycosylated protein, was captured by the three lectin columns, yet the spectral counts observed were relatively low compared to its high abundance making up more than 50% of total protein mass in serum. This indicates some residual nonspecific interactions, which do not call for depleting such a high abundance protein when dealing with lectin capturing of glycoproteins given the fact that the amount of bound albumin is relatively low and will not cause any obscuring effect on other medium or low abundance glycoproteins.

In the 22 proteins captured by the NASM-LCA column (see Table **1S** in **SI**),  $\alpha$ -2-macroglobulin yielded the highest spectral count of 14 followed by the protein isoform 2 of Ig mu chain C region whose spectral count was 10, then Ig alpha-1 chain C region whose spectral count was 8 while Ig gamma-1 chain C region, Ig gamma-3 chain C region and trypsin have had each a spectral count of 6, and ficolin-3 registered a spectral count of 4. All remaining 15 proteins recorded 1 spectral count (7 proteins), 2 spectral count (6 proteins) and 3 spectral count (2 proteins). The spectral count represents the total number of spectra identified for a protein. It has gained acceptance as a practical, label-free, semi-quantitative measure of protein abundance in proteomic studies [34]. Therefore, for the proteins captured by NASM-LCA column, the recorded spectral count may indicate that the glycoproteins or their glycoforms that have core fucosylated glycans are rather of medium to low abundance in disease-free serum.

Sixteen proteins out of the 49 proteins captured by the NASM-Con A column recorded spectral count of double digits of 10 or more (see Table **1S** in **SI**), while 18 other proteins yielded spectral count in the range of 4 to 8, and the remaining 15 proteins had spectral count of 1 (4 proteins), 2 (9 proteins) and 3 (2 proteins). On this basis one can state safely that the majority of proteins captured by NASM-Con A column and presumably containing high mannose, hybrid and biantennary complex type glycans are of the high to medium abundance glycoproteins/glycoforms.

In the 57 proteins captured by the NASM-RCA column, 24 proteins registered spectral count of higher or equal to 10. Twenty-three proteins had spectral count in the range 4 to 9, while the remaining 10 proteins recorded spectral count of 1 (3 proteins), 2 (4 proteins) and 3 (3 proteins).

NASM-RCA column, which is the last column in the tandem column format, captured more high and medium abundance glycoproteins/glycoforms than the other two lectin columns because it has a broader specificity. This indicates that the majority of serum glycoproteins belong to those containing bisected hybrid type, tri- and tetra-antennary complex type glycans. NASM-RCA column bound exclusively 8 glycoproteins namely alpha-1acid glycoprotein 1, alpha-1acid glycoprotein 2, alpha-S1-casein, desmoglein-1, hemoglobin subunit alpha, hemoglobin subunit beta, leucine-rich alpha 2-glycoprotein and lysozyme C.

The tandem operation of lectin columns having different binding selectivities toward various glycans can assess *via* LC-MS/MS analysis the level of protein glycosylation with each type of glycans and the level of various glycoforms for a given glycoprotein. For instance,  $\alpha$ -2-macroglobulin binds to the 3 lectin columns *via* its glycoforms, and the spectral count registered on each lectin column expresses the level of the various glycoforms. It has been reported that  $\alpha$ -2-macroglobulin possesses four abundant glycans of the bi-antennary complex type with one or two sialic acids at the non-reducing ends and with or without core fucosylation [35]. Also, it has been reported earlier by Arnold et al. [36] that the Asn-846 of  $\alpha$ -2-macroglobulin is occupied by the high mannose type glycans (Man<sub>5-7</sub>). From these published findings, it is obvious that the portion of the protein retained on the NASM-LCA column reflects the level of glycoforms that have attached core fucosylated bi-antennary complex type glycans, while the larger portions retained on the NASM-Con A column (spectral count = 80) and NASM-RCA column (spectral count = 72) may be attributed to the predominance of glycoforms with attached bi-antennary complex type glycans without core fucosylation but with one or two sialic acids (strong affinity to RCA and to a lesser extent to Con A), and to the high mannose type glycans (primarily strong affinity to Con A), which may explain the close spectral count of 80 and 72 recorded on NASM-Con A and NASM-RCA, respectively (see Table 1S in SI). A similar observation can be brought up in the case of haptoglobin whereby much more glycoforms that bind to the NASM-RCA column existed

(spectral count = 61) than those having affinity to NASM-Con A column (spectral count = 32) with both are more preponderant than the glycoforms that have affinity to NASM-LCA column (spectral count = 3). This corroborate with the findings that haptoglobin has exclusively complex type glycans which are either biantennary (affinity to RCA and to lesser extent to Con A) or triantennary (affinity to RCA) with sialic acids at the non-reducing ends. These glycans may have fucose linked to the core *N*-acetyl glucosamine (GlcNAc) residue [37, 38], and may not be of high level which explain the lower spectral count observed for the glycoforms captured by the NASM-LCA column. Another case worth mentioning is that of ceruloplasmin that displays spectral counts of 24 and 2 on the NASM-RCA and NASM-Con A columns, respectively, reflecting the noticeable abundance of the glycoforms which have primarily the attached saccharides of the type that has strong affinity to RCA than to Con A by almost a factor of 12. The recently identified four glycosylation sites by Harazono et al. [39] in ceruloplasmin have been shown to be occupied by a sialylated biantennary or triantennary oligosaccharide with fucose residues (0, 1, or 2). A small amount of sialylated tetra-antennary oligosaccharide was also detected. Also, fucose residues were linked to the reducing end GlcNAc in biantennary oligosaccharides and to reducing end and/or  $\alpha$ 1–3 to outer arms GlcNAc in triantennary oligosaccharides and that roughly one of the antennas in triantennary oligosaccharides was  $\alpha$ 2–3 sialylated and occasionally  $\alpha$ 1–3 fucosylated at GlcNAc. The reversal can be mentioned for the case of complement C3 which possesses much more glycoforms of the NASM-Con A type (spectral count = 102) than the glycoforms of the NASM-RCA type (spectral count = 10), and that amount for almost a factor of 10. This is not surprising, since complement C3 is rich in high mannose type glycans [29] and has been found recently to have 5 high mannose glycans (e.g., Man5-9GlcNAc2) and one

monoglucosylated glycan Glc1Man9GlcNAc2 [40]. Other proteins that favors the binding to NASM-Con A than to NASM-LCA and NASM-RCA are apolipoprotein B-100, beta-2-glycoprotein 1, and complement component C6, to mention a few cases. Typical proteins that have glycoforms favoring the binding to NASM-RCA than NASM-Con A are antithrombin-III, desmoplakin, hornerin and isoform 2 of clusterin just to name a few (see Table 1S in SI).

#### 4 Concluding remarks

The NASM proved to be a useful precursor for PPM with octadecyl ligands, trypsin and lectins for use in RPC of proteins, in online tryptic digestion of proteins and capturing glycoproteins from human serum by LAC, respectively. The NASM-OD microbore column yielded high resolution separation of six standard proteins using a linear gradient elution at increasing ACN concentration in the mobile phase on a conventional HPLC system, yet consuming much less solvent than the regular analytical scale HPLC column. The NASM-trypsin IMER allowed the online digestion of the standard protein Cyt C in a single pass of 20 min with improved digestion efficiency compared to in-solution digestion protocols, which take several hours. Furthermore, the online setup allowed the simultaneous peptide mapping of the protein digest by RPC with virtually little or non-operator intervention. The narrow-bore HPLC columns filled with NASM-lectin columns were effective in LAC of glycoproteins and their capturing from human serum. When NASM-lectin columns were operated in tandem in the order of increasing broader selectivities; that is in the order NASM-LCA → NASM-Con A → NASM-RCA, they allowed not only the capturing of the widest range of glycoforms of glycoproteins from human serum but also to sort out the level of each type(s) of glycoforms *via* LC-MS/MS analysis. This strategy proved to be a convenient approach to capture the serum glycoproteins without the necessity of *a priori* depleting high abundance proteins from serum.

#### References

- [1] Guiochon, G., *J. Chromatogr. A* 2007, *1168*, 101-168.
- [2] Jonnada, M., Rathnasekara, R., El Rassi, Z., *Electrophoresis* 2015, *36*, 76-100.
- [3] Rathnasekara, R., Khadka, S., Jonnada, M., Rassi, Z. E., *Electrophoresis* 2017, *38*, 60-79.
- [4] Zhang, J., Wu, S.-L., Kim, J., Karger, B. L., *J. Chromatogr. A* 2007, *1154*, 295-307.
- [5] Takeuchi, T., *Fresenius' J. Anal. Chem.* 1990, *337*, 631-637.
- [6] Scott, R. P. W., Kucera, P., *J. Chromatogr. A* 1979, *169*, 51-72.
- [7] Ishii, D., Asai, K., Hibi, K., Jonokuchi, T., Nagaya, M., *J. Chromatogr.* 1977, *144*, 157-168.
- [8] Hermanson, G. T., *Bioconjugate techniques*, Elsevier, Amsterdam 2013.

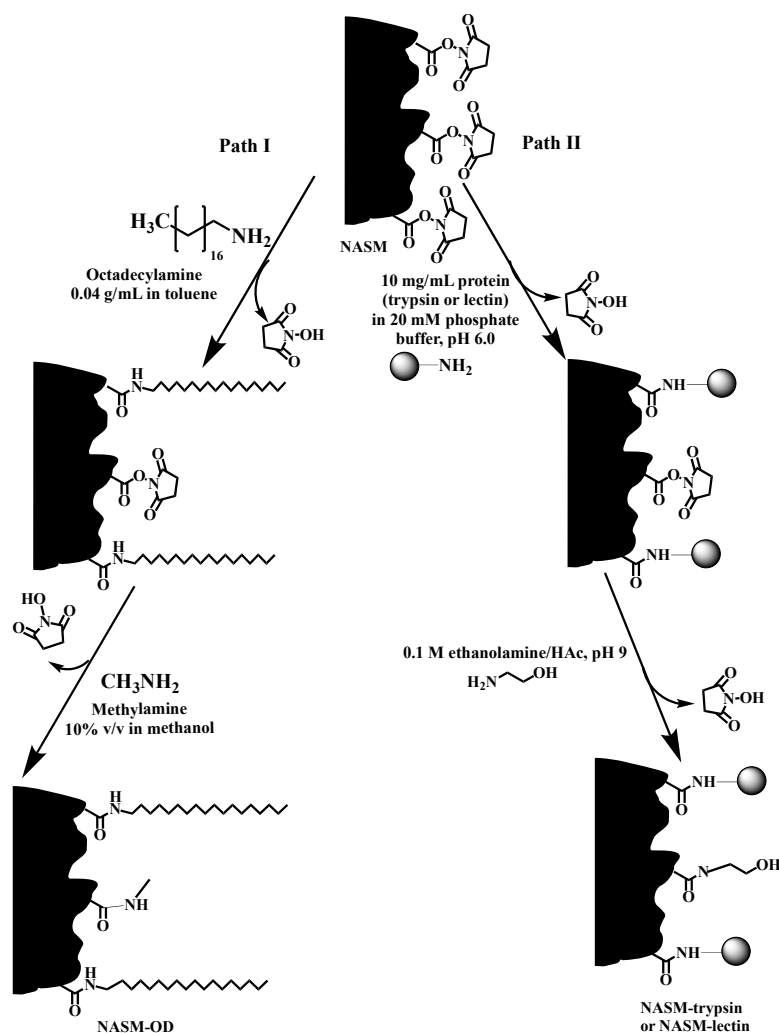
This article is protected by copyright. All rights reserved.

- [9] Lomant, A. J., Fairbanks, G., *J. Mol. Biol.* 1976, 104, 243-261.
- [10] Carbonnier, B., Guerrouache, M., Denoyel, R., Millot, M.-C., *J. Sep. Sci.* 2007, 30, 3000-3010.
- [11] Guerrouache, M., Carbonnier, B., Vidal-Madjar, C., Millot, M.-C., *J. Chromatogr. A* 2007, 1149, 368-376.
- [12] Guerrouache, M., Pantazaki, A., Millot, M.-C., Carbonnier, B., *J. Sep. Sci.* 2010, 33, 787-792.
- [13] Guerrouache, M., Millot, M.-C., Carbonnier, B., *Macromol. Rapid Commun.* 2009, 30, 109-113.
- [14] Spiro, R. G., *Adv. Protein Chem.* 1973, 27, 349-467.
- [15] Fanayan, S., Hincapie, M., Hancock, W. S., *Electrophoresis* 2012, 33, 1746-1754.
- [16] Selvaraju, S., El Rassi, Z., *J. Sep. Sci.* 2012, 35, 1785-1795.
- [17] Selvaraju, S., El Rassi, Z., *Proteomics* 2013, 13, 1701-1713.
- [18] Jung, K., Cho, W., *Anal. Chem.* 2013, 85, 7125-7132.
- [19] Debray, H., Decout, D., Strecker, G., Spik, G., Montreuil, J., *Eur. J. Biochem.* 1981, 117, 41-55.
- [20] Kornfeld, K., Reitman, M. L., Kornfeld, R., *J. Biol. Chem.* 1981, 256, 6633-6640.
- [21] Kobata, A., Yamashita, K., in: Fukuda, M., Kobata, A. (Eds.), *Glycobiology A Practical Approach*, IRL Press, Oxford 1993, pp. 103-125.
- [22] Yamamoto, K., Tsuji, T., Matsumoto, I., Osawa, T., *Biochemistry* 1981, 20, 5894-5899.
- [23] Baenziger, J. U., Fiete, D., *J. Biol. Chem.* 1979, 254, 9795-9799.
- [24] Lis, H., Sharon, N., in: Ginsburg, V., Robbins, P. W. (Eds.), *Biology of Carbohydrates*, John Wiley & Sons, New York 1984, pp. 9-22.
- [25] Marangoni, R., Chiarini, R., Iannone, G., Salerno, M., *Proteomics* 2008, 8, 2165-2167.
- [26] Li, Y., Xu, X., Deng, C., Yang, P., Zhang, X., *J. Proteome Res.* 2007, 6, 3849-3855.
- [27] Gao, J., Xu, J., Locascio, L. E., Lee, C. S., *Anal. Chem.* 2001, 73, 2648-2655.
- [28] Prien, J. M., Ashline, D. J., Lapadula, A. J., Zhang, H., Reinhold, V. N., *J. Am. Soc. Mass Spectrom.* 2009, 20, 539-556.
- [29] Solís, D., Feizi, T., Yuen, C.-T., Lawson, A. M., Harrison, R. A., Loveless, R. W., *J. Biol. Chem.* 1994, 269, 11555-11562.
- [30] Harvey, D., Wing, D., Küster, B., Wilson, I., *J. Am. Soc. Mass Spectrom.* 2000, 11, 564-571.
- [31] Huhn, C., Selman, M. H. J., Ruhaak, L.R., Deelder, A. M., Wuhrer, M., *Proteomics* 2009, 9, 882-913.
- [32] Marz, L., Hatton, M. W. C., Berry, L. R., Regoeczi, E., *Can. J. Biochem.* 1982, 60, 624-630.
- [33] G., d. J., van Eijk, H. G., *Electrophoresis* 1988, 9, 589-598.
- [34] D.H., L., Hwang, S. I., Wu, L., Han, D. K., *Expert Rev. Proteomics* 2010, 7, 39-53.
- [35] Lin, Z., Lo, A., Simeone, D. M., Ruffin, M. T., Lubman, D. M., *J. Proteomics Bioinform* 2012, 5, 127-134.
- [36] Arnold, J. A., Wallis, R., Willis, A. C., Harvey, D. J., Royle, L., Dwek, R. A., Rudd, P. M., Sim, R. B., *J. Biol. Chem.* 2006, 281, 6955-6963.
- [37] Dobryszczyka, W., in: Mackiewicz, A., Kushner, I., Baumann, H. (Eds.), *Acute Phase Proteins: Molecular Biology, Biochemistry, and Clinical Applications*, CRC Press, Boca Raton 1993, pp. 185-206.

- [38] Nilsson, B., Lowe, M., Osada, J., Ashwell, G., Zopf, D., in: Yamakawa, T., Osawa, T., Handa, S. (Eds.), *Glycoconjugates*, Japan Scientific Societies Press, Tokyo 1981, pp. 275-276.
- [39] Harazono, A., Kawasaki, N., Satsuki, I., Hashii, N., Ishii-Watabe, A., Kawanishi, T., Hayakawa, T., *Anal. Biochem* 2006, 348, 259-268.
- [40] Crispin, M. D., Ritchie, G. E., Critchley, A. J., Morgan, B. P., Wilson, I. A., Dwek, R. A., Sim, R. B., Rudd, P. M., *FEBS Lett.* 2004, 566, 270-274.

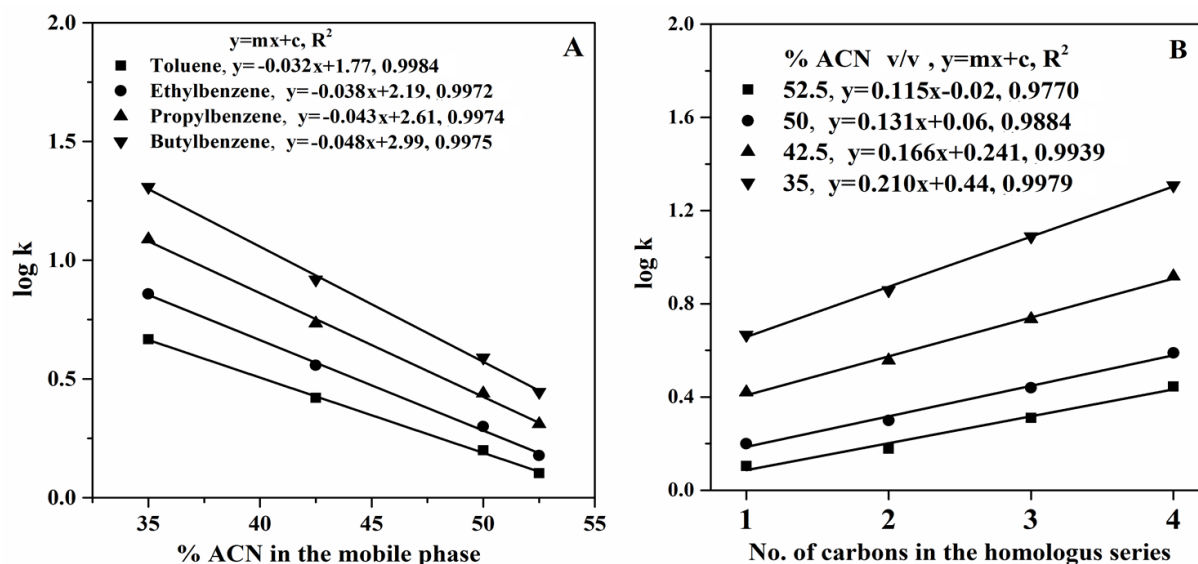
### Figure captions

**Figure 1.** Post polymerization modification of NASM with octadecyl and protein ligands. Path I: reaction of the surface succinimide groups of NASM with octadecylamine followed by scavenging the unreacted succinimide groups with methylamine; Path II: immobilization of protein (trypsin or lectin) on the surface succinimide groups of NASM followed by scavenging the unreacted succinimide groups of NASM with ethanolamine.

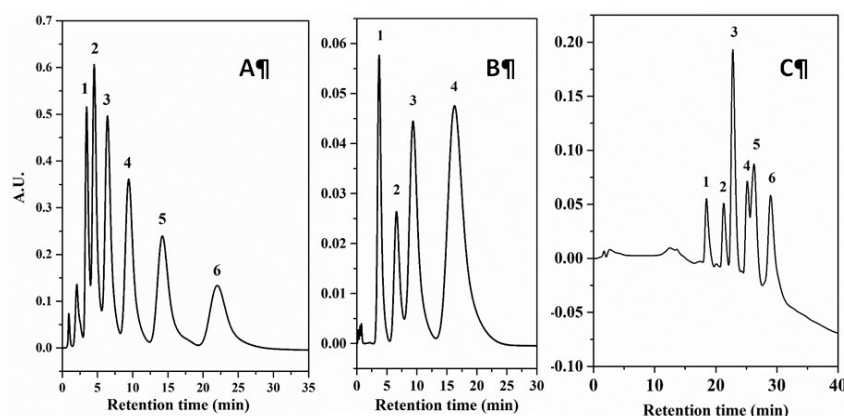


This article is protected by copyright. All rights reserved.

**Figure 2.** Plots of log *k* of the ABs homologous series vs. % ACN (v/v) in the mobile phase (A) and log *k* vs. carbon number of the homologous series at different ACN concentration in the mobile phase (B). Column, NASM-OD, 21.5 cm x 1 mm i.d.; flow rate, 0.1 mL/min; detection, 214 nm.

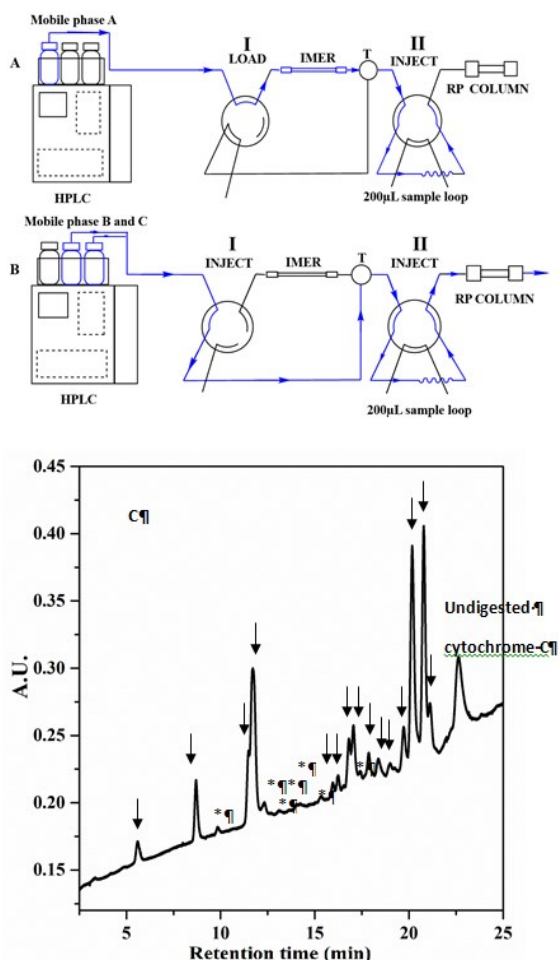


**Figure 3.** Chromatograms of ABs (A), PAHs (B) and standard proteins (C) obtained on the NASM-OD column. Column dimensions, 21.5 cm x 1 mm i.d.; flow rate, 0.1 mL/min; mobile phases in (A): water:ACN 45:55 (v/v); in (B): 60:40 (v/v) and in (C): linear gradient elution in 30 min from water:ACN, 95:5 (v/v) at 0.1% (v/v) TFA to water:ACN, 5:95 (v/v) at 0.1% (v/v) TFA. Peaks in (A): 1, toluene; 2, ethylbenzene; 3, propylbenzene; 4, butylbenzene; 5, pentylbenzene; 6, hexylbenzene. Peaks in (B): 1, 1-nitronaphthalene; 2, fluorine; 3, phenanthrene; 4, pyrene. Peaks in (C): 1, ribonuclease A; 2, cytochrome C; 3, lysozyme; 4, conalbumin; 5,  $\beta$ -lactoglobulin; 6, human serum albumin. UV detection at 214 nm in (A) and (C) and at 254 nm in (B).



This article is protected by copyright. All rights reserved.

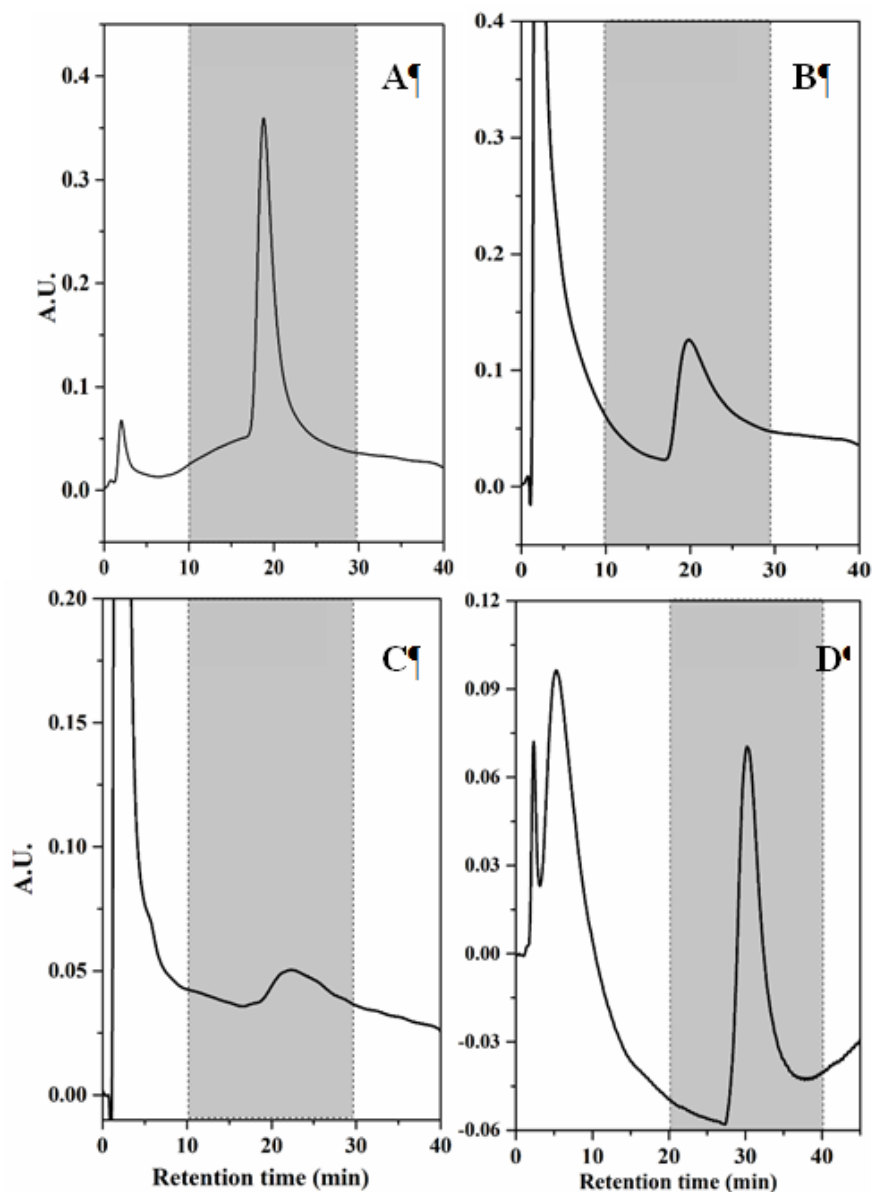
**Figure 4.** Chromatographic set-up for the online digestion of proteins by NASM-trypsin IMER (A and B) and RPC peptide mapping (C). (A): For allowing digestion mobile phase to go through the IMER, valve (I) is in the load position, while valve (II) is in the inject position to allow the digestion of Cyt C and the simultaneously collection of the digest in the 200  $\mu$ L loop of valve (II). (B): The valve before the IMER is changed to inject position to by-pass the mobile phase from the IMER and facilitate the subsequent injection of the collected digest in the 200  $\mu$ L loop into the RPC column to generate the peptide map. The arrows indicate the path of the mobile phase through the setup. Chromatographic conditions for online digestion: injection volume, 5  $\mu$ L; flow rate, 10  $\mu$ L/min; digestion mobile phase (Solvent A), 25mM  $\text{NH}_4(\text{HCO})_3$ , pH 8.0; digestion time, 20 min. Chromatographic conditions for peptide mapping in (C): column, octadecyl silica (ODS) UltraSphere™ column packed with 5- $\mu$ m average particle diameter, 4.5 cm x 4.6 mm i.d.; solvent B consists of 0.1% (v/v) TFA in 5% ACN; solvent C is made of 0.1% (v/v) TFA in 50% ACN; gradient conditions: two consecutive linear gradients from 0 to 60 %C in 20 min and then to 100% C in 5 min; flow rate, 1.0 mL/min; UV detection at 214 nm.



This article is protected by copyright. All rights reserved.



**Figure 5.** Chromatograms of ribonuclease B (A) and ovalbumin (B) obtained on the NASM-Con A column (21.5 cm x 1 mm i.d.). (C) *Chromatogram of IgG obtained on NASM-LCA (21.5 cm x 1 mm i.d.).* Binding mobile phase in (A), (B) and (C), 20 mM Tris HCl buffer, pH 7.4, containing 100 mM NaCl and 1mM  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$  and  $\text{Mn}^{2+}$ ; eluting mobile phase, 0.1 M Me- $\alpha$ -Man in the binding mobile phase; flow rate, 0.1 mL/min; wavelength, 214 nm. (D) *Chromatogram of human transferrin obtained on NASM-RCA (25.0 cm x 1 mm i.d.).* Binding mobile phase, 20 mM Tris HCl buffer, pH 7.4, containing 100 mM NaCl; eluting mobile phase, 0.1 M Lactose in the binding mobile phase; flow rate, 0.1 mL/min; wavelength, 214 nm. On all the chromatograms, elution mobile phase is indicated by the gray region.



This article is protected by copyright. All rights reserved.