Insights into chitosan-cleaving enzymes by simultaneous analysis of polymers and oligomers via SEC-RI-ESI-MS

Margareta J. Hellmann
University of Münster

Bruno M. Moerschbacher (moersch@uni-muenster.de)
University of Münster

Stefan Cord-Landwehr
University of Münster

Article

Keywords:

Posted Date: June 27th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3063686/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

The detailed characterization of chitosan-cleaving enzymes is of great importance to unveil structure-function relationships of this promising class of biomolecules, both for enzymatic fingerprinting analyses and to use the enzymes as biotechnological tools for the production of tailor-made chitosans for diverse applications. Analyzing polymeric substrates as well as oligomeric products has proven a suitable way to understand the enzymes’ actions but currently, this requires separate, rather laborious methods to obtain the full picture. Here, we describe size exclusion chromatography coupled to refractive index and electrospray ionization mass spectrometry detection (SEC-RI-ESI-MS) as a simple and fast way to semi-quantitatively analyze chitosan oligomers and to concomitantly determine the average molecular weight and its distribution for chitosan polymers. By sampling live from an ongoing enzymatic reaction, SEC-RI-ESI-MS offers the unique opportunity to analyze polymers and oligomers simultaneously – i.e. to follow the reduction in molecular weight of the polymeric substrate over the course of the digestion, while at the same time analyzing the emerging oligomeric products in a quantitative manner. In this way, a single simple analysis yields detailed insights into an enzyme’s action on a given substrate.

Introduction

Chitosans are a family of polysaccharides that attract increasing interest for diverse applications. They are industrially produced by the partial chemical deacetylation of chitin, a linear polymer consisting of β-1,4-linked N-acetyl-d-glucosamine units (GlcNAc, A-units) which can be found, e.g., in the exoskeletons of crustaceans and insects, or in the cell walls of fungi. The partial removal of N-acetyl groups turns the water insoluble chitin into soluble chitosans composed of β-1,4-linked GlcNAc and d-glucosamine units (GlcN, D-units). These linear copolymers have diverse structures which can be characterized by three key factors: The degree of polymerization (DP) describes the number of monosaccharide moieties to form a chitosan molecule; the proportion of GlcNAc units and their distribution along the sugar’s chain are termed fraction of acetylation (FA) and pattern of acetylation (PA), respectively. Fields for applications of these glycans range from agriculture over drug or gene delivery in nanoparticles to water purification, wound healing, food preservation and cosmetics. This high diversity in applications and activities is linked to the structural diversity of chitosans, which in turn is based on the natural variety of chitosan-synthesizing, -cleaving and modifying enzymes that influence DP, FA and PA.

A wide spectrum of chitosan hydrolyzing enzymes has been described, each preferring substrates and releasing products of certain DP, FA, and PA. Crucial factors for these differences are the shape of the catalytic site as well as the preferences for A- and D-units at the enzyme’s individual subsites, each binding one sugar unit. There exists a common terminology for glycan-cleaving enzymes: Subsites which bind the substrate’s subunits towards its non-reducing end from the hydrolyzed bond are subsites − 1, −2, et cetera; subsites binding towards the reducing end are subsites + 1, +2, et cetera. Chitinases prefer A-units at most subsites – their activity increases with increasing FA. In contrast, the activity of chitosanases decreases with increasing FA, because they favor D-units over A-units at most subsites.
On the one hand, the characterization of chitosan hydrolyzing enzymes is crucial to unravel structure-function relationships of the chitosans in vivo which in turn gives insights into the enzymes’ biological role. On the other hand, well-characterized enzymes can be exploited as biotechnological tools to produce tailor-made chitosans that show optimum performance in a given application. But the characterization of a chitinase, chitosanase or chitosanase is a laborious process that involves a thorough analysis of the chitosan substrates and the enzymatic products, at least concerning DP and FA. For polymeric substrates, the gold standards are nuclear magnetic resonance (NMR) spectroscopy to determine the average FA, and size exclusion chromatography coupled to refractive index and multi-angle laser light scattering detection (SEC-RI-MALLS) to analyze the molecular weight (MW) which allows the calculation of the DP if the FA is known. Apart from the MW distribution plot, the latter method gives the weight average MW \( M_w \), the number average MW \( M_n \) and the dispersity \( D_M \) – for details about the calculation and meaning of those three quantities, and about the difference between \( M_w \) and \( M_n \), see the supplementary information.

Both NMR spectroscopy and SEC-MALLS-RI require individual sample preparation as well as substantial amounts of chitosan polymer in the milligram range, and the obtained results are typically evaluated using manual integration, introducing a certain subjectivity into the generated values. Recently, enzymatic fingerprinting using detection by electrospray ionization mass spectrometry (ESI-MS) has been proposed as a new method to determine the average FA of polymers that uses only microgram amounts of sample and gives robust, objective results. In contrast, no optimized alternatives to SEC-MALLS-RI are available for MW analysis.

Of even greater importance for enzyme characterization is the analysis of small enzyme products, the released non-, partially, or fully acetylated chitooligosaccharides (COS). Whereas these can easily be annotated in terms of DP and FA based on their m/z values obtained by MS, different COS possess different signal response factors, making quantification via MS challenging. Currently, quantitative MS sequencing is the most sophisticated method that allows even the absolute quantification of oligomers of a certain DP, FA and PA with high accuracy. But it requires derivatization of the samples in form of complete chemical N-acetylation using isotope-labeled acetic anhydride, followed by reducing end \(^{18}\)O-labeling of MS samples, and addition of defined amounts of double isotope-labeled standards to MS samples. Moreover, quantitative MS sequencing is limited to oligomers up to DP 6 which need to be separated from remaining polymeric substrates by filtration before derivatization.

In the following, we present a new SEC-RI-ESI-MS setup to characterize chitosan-cleaving enzymes which combines the strengths of SEC coupled to RI to get insights into the polymeric substrate, and of RI coupled to MS to analyze oligomeric products. On the one hand, we established an efficient separation of COS with protein SEC columns as reported before for oligomers of pectin or (oxidized) cellulose, followed by detection using RI and ESI-MS (section “Oligomer analytics”). Furthermore, challenges in COS quantification are addressed, making it possible to quantitatively analyze COS of different DP and FA without the need for filtration and derivatization. On the other hand, we demonstrate that SEC-RI-ESI-MS
can be used for the MW analysis of chitosan polymers (section “Polymer analytics”). This is already of importance by itself, because it offers an alternative to SEC-RI-MALLS with considerably lower amounts of sample and shorter run times. But the crucial point is that oligomers and polymers can be analyzed using the same setup, resulting in the unique opportunity to simultaneously monitor the behavior of polymers during degradation and the consequent formation of oligomeric products. By taking consecutive live samples from a single ongoing reaction without additional sample preparation, this offers detailed insights into the course of an enzymatic reaction (results section “Live digestion”): How fast is a polymer degraded? How much of which COS are released? How does the product profile change over the course of the reaction?

While even live sampling of a single enzymatic reaction already gives helpful insights, analyzing a handful of samples with different enzymes or substrates results in a detailed comparison of the performance of various enzymes on the same substrate or of one enzyme on different substrates, respectively. Although the latter would not involve considerably more experimental effort, we decided to present just the data of a single enzymatic reaction using the very well-characterized chitosanase from Bacillus sp. MN (CsnMN)\textsuperscript{18,26,27}. In this way, we want to emphasize how much information can be derived from a single run of our SEC-RI-ESI-MS setup, and how closely the results match the findings of previous, more laborious analyses of this enzyme.

**Materials and methods**

**Polymers and oligomers**

The chitosan polymers used to determine the relationship between $M_w$ or $M_n$ and retention time (RT) in the SEC-RI-ESI-MS setup or as substrate for the live digestion were kindly provided by Heppe Medical Chitosan (HMC; Halle, Germany) or Gillet Chitosan (Plumaudan, France). All chitosans were produced by heterogeneous deacetylation of crustacean chitins. The FA was determined via enzymatic MS fingerprinting\textsuperscript{22}; $M_w$, $M_n$ and $\bar{D}_M$ were analyzed using SEC-RI-MALLS\textsuperscript{21} (for details see Supplementary Table S1). Additionally, polymeric and oligomeric pullulan standards (Polymer Standards Service, Mainz, Germany) as well as a chitosan oligomers mixture derived from the acid hydrolysis of polymeric chitosan (Kitostim, kindly provided by Gillet Chitosan) were measured using SEC-RI-ESI-MS. Their $M_w$ was given by the supplier for large pullulan standards or derived from the m/z values from MS spectra for chitosan or pullulan oligomers. To check their ionization efficiency and RI response, chitin oligomers of DP 2-6 (A2-6) were purchased from Megazyme (Wicklow, Ireland) and chitosan tetramers of different FA and PA were produced as described by Hembach et al.\textsuperscript{28}.

**SEC-RI-ESI-MS**

The setup proposed here consists of three parts: Initially, analytes were injected by an autosampler and separated using a Dionex Ultimate 3000RS UHPLC system (Thermo Fisher Scientific, Milford, USA)
equipped with an ACQUITY UPLC Protein BEH SEC column (either 125 or 200 Å pore size, 1.7 µm, 4.6 mm x 300 mm, Waters Corporation, Milford, USA). The isocratic elution was performed at 40°C with a flow rate of 0.4 mL/min\(^2\) with a runtime of 14 min using ammonium acetate (150 mM) and acetic acid (200 mM) dissolved in MilliQ water (final pH of 4.5) as solvent\(^\text{21}\). In general, polymer samples were dissolved overnight in the solvent whereas oligomer samples were dissolved in MilliQ water, both in a final concentration of 1 g/L. For polymer and oligomer samples, 3 µL and 1 µL were injected, respectively. Following separation, the flow was split using a 1:1 splitter (Accurate, Dionex Corporation, Sunnyvale, USA). One half is used to measure RI signals with an ERC RefractoMax 520 (Thermo Fisher Scientific) at 10 Hz and 40°C with recorder and integrator ranges of 512 µRIU and 125 µRIU/V, respectively. The other half flows into an ESI-MS\(^\text{n}\) detector (amaZon speed; Bruker, Bremen, Germany) in positive mode with a capillary voltage of 4.5 kV, an end plate offset voltage of 500 V, a nebulizer pressure of 15 psi, a dry gas flow rate of 8 L/min and a dry temperature of 180°C. For spectra acquisition, the device was set to enhanced resolution scan mode with a target mass of m/z 700 covering a scan range of m/z 50-2000. The ICC target was set to 200,000 and the maximum accumulation time to 10 ms. Whereas the analysis of MS data was performed using Data Analysis 4.1 (Bruker) and an in-house Python script based on the module pymzML\(^\text{29}\), the RI data were exported using Data Analysis 4.1 and evaluated using OriginPro 2023 (OriginLab, Northampton, USA). The latter is explained in more detail in the upcoming section.

**Live digestion**

A single sample is composed of substrate, a suitable buffer, and a suitable amount of enzyme. Pretests might be necessary to find a buffer type and concentration for appropriate enzyme activity as well as to determine the amount of enzyme that fits to the desired timeframe of the live digestion. Ideally, the first datapoints represent early timepoints of the reaction whereas the latest timepoint depicts (at least nearly) the end point. In the example shown in this work, chitosan 651 (Gillet Chitosan) with an FA of 0.22 and an \(M\text{w}\) of around 130 kDa was dissolved under slightly acidic conditions and served as substrate (1 g/L in sample) in sodium acetate buffer (40 mM, pH 6 in sample). The recombinantly expressed chitosanase CsnMN\(^\text{18,26,27,30}\) was used for enzymatic hydrolytic cleavage (12.5 nM in sample).

The workflow to perform a single live digestion experiment is summarized in Figure 1. First, all components of the sample except for the enzyme are prepared in an MS vial and moved to the autosampler which is heated or cooled to the intended reaction temperature. As soon as the samples reach this temperature, the reaction starts by adding the enzyme via manual pipetting or by using the autosampler. Immediately afterwards, the first sample representing \(t_0\) is drawn from the reaction vial and injected into the SEC-RI-ESI-MS setup with the 125 Å column as described above. If the time between enzyme addition and injection is too long, it is reasonable to use a control without enzyme as \(t_0\). Now the autosampler can perform an automated live sampling over the whole desired reaction time, directly resulting in raw MS and RI data for all chosen timepoints \(t_{1-x}\). In the example shown, we chose 300 min total reaction time at 37°C and injected samples of 3 µL from the 50 µL reaction. Because the SEC
runtime was set to 14 min and the operations of the autosampler take one minute between runs, we sampled every 15 min. If certain timepoints are not desired, MilliQ water, other samples, or controls can be injected instead to avoid wasting sample. In case the solvent used for the SEC differs from the reaction buffer inside the reaction vial in composition or pH, one should not contaminate the reaction vials with the solvent during injection. This can be achieved by changing the washing solution of the autosampler to water instead of connecting it with the solvent and/or by using air bubbles within the autosampler needle to stop the sample from mixing with the liquid in the sample loop.

Acquired data are analyzed in a partially automated way: The MS data are converted to mzML format and processed with an in-house Python script based on the module pymzML to give an output csv file with the arbitrary intensity of the MS signal per oligomer and timepoint. In contrast, RI data are transferred to OriginPro 2023, plotted, and the areas of the peaks of each DP per timepoint determined with the Peak Analyzer tool are copied to a second csv file. Both csv files are combined to a final output csv file with another Python script (ri_ms_combination.py) that is available in the supplementary information; additionally, examples for the three csv files and a user guideline for the Python script can be found there.

Results and discussion

Polymer analytics: SEC-RI-ESI-MS as an alternative to SEC-RI-MALLS

The first step towards a simultaneous analysis of chitosan polymers and oligomers via SEC-RI-ESI-MS was to establish our setup as an alternative to SEC-RI-MALLS to determine the average MW of polymers based on newly established regressions. Figure 2a shows the relationship between the $M_w$ determined via SEC-RI-MALLS and the RT of chitosan polymers and oligomers as well as pullulan standards in the SEC-RI-ESI-MS setup for two SEC columns of different pore size. During their flow through the column of porous beads, smaller oligomers can enter the pores and take a longer route, hence eluting later than larger polymers. When comparing the performance of the two columns, the separation is most efficient around 1 kDa or 80 kDa for the column with 125 Å or 200 Å pore size, respectively, which fits to the optimal MW ranges indicated by the supplier. As a consequence, resolving larger chitosan oligomers (e.g. the Kitostim sample) as well as the pullulans of $M_w$ 0.5-2 kDa into individual peaks was only possible using the 125 Å column, resulting in more data points. Moreover, analytes elute in general faster from the 125 Å column, as expected from its smaller pore size. When comparing different analytes, the pullulan standards elute always later than corresponding chitosans of a similar $M_w$. This can be explained by the more globular structure of pullulans with their mix of $\alpha$-1,4 and $\alpha$-1,6 linkages in contrast to the linear and rather stiff chitosans, making the former appear smaller during SEC. Consequently, the connection between $M_w$ and RT is different for different polysaccharides and the calculation of $M_w$ values always needs to be performed using regressions that are based on measurements of the same polysaccharide species. With the y-axis in logarithmic scale, exponential relationships between $M_w$ and RT of chitosans
appear as straight lines in the plot. It is not possible to find a suitable exponential equation to connect all \( M_w \) values to a RT, but for certain \( M_w \) ranges, a common straight line is reasonable: the low \( M_w \) range from 0.1-1 kDa, the medium \( M_w \) range from 1-10 kDa, and the high \( M_w \) range from 10-1000 kDa. The set of polymeric chitosans (HMC samples) in the high \( M_w \) range from Figure 2a is depicted as a close-up in Figure 2b. Whereas samples of higher \( M_w \) exhibit low standard deviations and fit well to the regression, samples of low \( M_w \) deviate more strongly. It needs to be mentioned that due to the low \( M_w \), the MALLS signals of these samples are weak which leads to inaccuracy and high deviations in the SEC-RI-MALLS results. Based on the \( R^2 \) values of the exponential regressions, the 200 Å column is more suitable for the analysis of chitosan polymers, although both columns give reliable results. An exponential relationship was also established between \( M_n \) and RT for both SEC columns with \( R^2 \) values above 0.96 (Supplementary Fig. S1).

The obtained exponential equations are used to convert signals from measurements of chitosan polymers in the SEC-RI-ESI-MS setup to the corresponding \( M_w \). Figure 2c shows the \( M_w \) distribution obtained by SEC-RI-MALLS of five selected samples; these were also measured via SEC-RI-ESI-MS (200 Å column) and based on the corresponding exponential equation, a second x-axis with the \( M_w \) was added to the chromatogram (Fig. 2d). As intended, both methods give similar results, but some differences are visible: On the high \( M_w \) end above 500 kDa, the SEC-RI-ESI-MS setup is not suitable for efficient separation as samples 75/1000 (\( M_w \): 404 kDa) and 75/3000 (\( M_w \): 535 kDa) elute simultaneously even on the 200 Å column. Moreover, the \( M_w \) range above 500 kDa was not covered when establishing the regressions. In contrast, SEC-RI-MALLS shows different results for the two samples as expected from viscosity data of the supplier, indicating a better performance of this method when analyzing very large polymers. However, for small polymers like 75/5 (\( M_w \): 66 kDa), it is difficult to get meaningful results on \( M_w \) and \( M_n \) via SEC-RI-MALLS because the intensity of the MALLS signal decreases with decreasing MW. On the contrary, SEC-RI-ESI-MS separates small polymers well and is just dependent on the RI signal, making it the better solution on the lower \( M_w \) end. Not unexpectedly, it becomes apparent that considering only the single value for the average MW is a strong simplification when characterizing a chitosan sample – it is clearly important to always consider the \( M_w \) or \( M_n \) distributions from SEC-RI-MALLS or SEC-RI-ESI-MS as well.

As mentioned above, SEC-RI-MALLS should still be the method of choice when analyzing very large polymers, at least when compared to the SEC columns used in the SEC-RI-ESI-MS setup used here, and it is obviously required to initially generate the regression describing the relationship between RT in SEC-RI-ESI-MS and MW of one type of analyte. Once this is established, using SEC-RI-ESI-MS has certain advantages over SEC-RI-MALLS: First, the amount of sample is multiple times lower and the separation times are much shorter, in our case 3 µg sample and 14 min for SEC-RI-ESI-MS compared to 100 µg sample and 90 min for SEC-RI-MALLS. Second, contaminations or degradation products can directly be identified via MS. And third, the data analysis is less subjective, because the RI peak maximum is used to calculate the sample's \( M_w \), whereas in SEC-RI-MALLS, the calculation is based on the manual definition of
the RI and MALLS peaks’ integrals. Therefore, SEC-RI-ESI-MS should be included into the general toolbox for chitosan polymer analytics as an alternative to SEC-RI-MALLS. But more importantly here, the setup forms an integral part of the live digestion method described below (results section “Live digestion”), addressing the polymer part of the sample.

**Oligomer analytics: Prerequisites for semi-quantitative analysis**

The second building block of this live digestion method is a simplified, quantitative oligomer analysis. Crucial for this is an adequate separation of COS with the same SEC columns used above for polymer MW determination. As discussed above, the results shown in Figure 2a for COS (Kitostim) and pullulan oligomers indicate a better performance in oligomer separation for the SEC column of pore size 125 Å compared to 200 Å, while maintaining reliable results in polymer analytics. Therefore, the 125 Å column was used in all further experiments. We were able to show that the efficiency of oligomer separation is not impaired by increased flow rates during chromatography (Supplementary Fig. S2). Hence, separation times can be reduced by increasing the flow rate as far as the pressure limit of the SEC column allows without loss of performance.

As mentioned in the introduction, the current main problem in the quantification of different COS are their different signal response factors in MS which in turn derive from their different ionization efficiencies during ESI. To work around this in the currently used quantitative MS sequencing method\(^\text{23}\), all D-units of COS are N-acetylated with isotope-labeled acetic anhydride before quantitative analysis. This derivatization turns non- or partially acetylated COS into chitin oligomers while still being able to distinguish between original and reacetylated GlcNAc units by MS due to the isotope label. Combined with the addition of defined amounts of double isotope-labeled standards of different DP, this allows for an absolute quantification of each COS of a certain DP and FA via MS. But the derivatization and the use of standards introduce limits into the analysis of oligomers: Because double isotope-labeled standards are only easily available up to DP 6 and even insoluble above DP 8, larger COS cannot be quantitatively analyzed using this method. Moreover, the full N-acetylation requires prior removal of remaining polymeric substrate and enzyme, making a simultaneous analysis of polymers and oligomers within one sample, or sampling live from an ongoing reaction impossible.

In the following, we describe how the combination of MS and RI signals paves the way towards a semi-quantitative analysis of chitosan oligomers without filtration and derivatization. This is only possible if three prerequisites are met: First, different oligomers of the same DP elute simultaneously, second, the RI signal intensity is only dependent on mass concentration, and third, the ionization efficiency of oligomers of one DP in the ESI-MS is similar. The first prerequisite is ensured by a high ammonium acetate concentration in the SEC solvent as shown in the supplementary information (Supplementary Fig. S3). In the following, the other two prerequisites are addressed.
At a constant temperature, the RI shows a linear correlation with a compound's mass concentration up to high concentrations\textsuperscript{32}. This is commonly used in SEC-RI systems with constant temperature, solvent composition, and flow rate for one analyte species to derive the relative mass concentration of the analyte from the RI peak integral. Indeed, with constant instrument parameters, the RI signal response is mainly dependent on the mass concentration, but also the MW and response factor of the analyte (dn/dc value) play a role\textsuperscript{33}, resulting in an influence of the chitosans’ DP and FA. Even though this is neglected in most SEC-RI applications on chitosans anyway, we confirmed that neither the DP (Fig. 3a) nor the FA or PA (Fig. 3b) have a systematic or considerable influence on the RI signal response of oligomers. Hence, a dependency of the RI signal intensity on the chitosan oligomer mass concentration alone can be assumed.

Like RI, also the signal response of chitosan oligomers in ESI-MS can differ dependent on the factors DP, FA, or PA because of the different ionization efficiencies during ESI. On the one hand, the ionization efficiency is vastly different for oligomers of different DP values (Fig. 3c): oligomers of DP 3 ionize best and are therefore overrepresented in the MS signal intensities when compared with other DP values. On the other hand, oligomers of the same DP but with different FA and PA values feature similar ionization efficiencies, as shown in Figure 3d.

The RI response factors allow the relative quantification of oligomers of one DP eluting in a single RI peak compared to oligomers of other DPs based on the integrals of the RI signal peaks. The MS response factors enable the relative quantification of oligomers of one FA compared to oligomers of other FAs based on the MS signal intensities, but only within one DP. A distinction between oligomers with the same DP and FA but differing PA is not possible here. Combining both signals makes it possible to quantify each oligomer of a certain DP and FA relative to all others. In comparison to quantitative MS sequencing\textsuperscript{23}, the combination of RI and MS proposed here indeed results in a less accurate quantification because it is based on the integration of RI peaks, because it does not feature absolute quantifications with internal standards, and most importantly, because the ionization of oligomers of the same DP with different FA is similar, but not identical — that is why we call it “semi-quantitative analysis”. However, our developed method is a unique way to easily quantify oligomers of defined DP and FA up to DPs with sufficient signal strength (typically at least up to DP 10) while simultaneously analyzing polymers within the same SEC-RI-ESI-MS run. Considering in addition that no further sample preparation is needed, this allows an easy and fast live analysis of chitosan-cleaving enzymes, as described in the following.

**Live digestion: Fast and easy analysis of an enzyme’s action**

The newly established methods for the analysis of polymers and oligomers via SEC-RI-ESI-MS were finally combined into a live digestion method for the fast and easy characterization of chitosan-cleaving enzymes. A single injected sample measured with the SEC-RI-ESI-MS setup already gives information on the amount, composition, and average FA of oligomer products as well as on the amount and estimated
$M_w$ of the remaining polymer substrate. But to create a full picture of the enzyme's activity on the substrate, it is necessary to monitor all these parameters over time. Because our method requires no sample preparation at all, the samples containing substrate, enzyme and buffer can be injected directly. That makes it possible to sample automatically over time from a single vial by using the autosampler. As shown in Figure 1, the enzymatic digestion is started by adding the enzyme and from this timepoint on, the autosampler takes live samples from the enzymatic digestion and injects them directly into the SEC-RI-ESI-MS system.

All data generated from a single vial with an enzymatic digestion are summarized in Figure 4, representative for the digestion of a chitosan polymer (651, FA: 0.22, $M_w$: 130 kDa) with the chitosanase CsnMN, sampled live over 300 min. We used DP 10 as the cutoff between small oligomers on the one hand, and polymers and large oligomers on the other hand, because even at low concentrations of products of DP $\leq$ 10, distinct RI peaks are formed.

First, one should consider the RI chromatograms (Fig. 4a, Fig. 5) which allow to directly follow the cleavage of the polymers into oligomers. Whereas the substrate contains just polymers as expected, a considerable part was already hydrolyzed after 45 min to smaller polymers (shift of the peak towards later RT), large oligomers (tailing peak) and small oligomers (distinct peaks). Because an exo-enzyme would produce only small oligomers directly from the large polymer substrate, this indicates an endo-cleaving activity which has been reported before for CsnMN$^{27}$. After 120 min, the majority of the polymers was hydrolyzed and substantial amounts of small oligomers of DP 2-6 were produced. When comparing the chromatograms at timepoints 120 min and 300 min, it becomes apparent that the enzyme degrades already small oligomers of DP 5 and 6 even further. The reaction was far progressed after 300 min, but considering the remaining polymers and large oligomers, probably not yet at its end point. To find the end point, one needs to just continue the ongoing live digestion until the RI chromatograms do not show changes anymore. Then it is useful to add fresh enzyme and take another sample after additional incubation time, to check if the reaction before had stopped due to activity loss of the enzyme, or if the real end point was reached. There, only substrates remain that do not fit to the enzyme's subsite preferences or required length, or the reaction has come to a standstill due to product inhibition.

Based on the relationship between $M_w$ and RT established by the regression in section “Polymer analytics” (Fig. 2b), the $M_w$ range of the polymers can be estimated as shown in Figure 4b by introducing a second x-axis. To plot the cleavage of the polymers over the full reaction time, the maximum of the polymer RI peak is converted to the corresponding $M_w$ (Fig. 4c). Both figures show the hydrolytic cleavage of the polymers over time: Whereas the non-digested substrate features a calculated $M_w$ of about 120 kDa (SEC-MALLS-RI of the substrate 651 had given 134 kDa, Supplementary Table 1), the $M_w$ of the RI maximum decreases over the course of the reaction – at first fast, then more slowly until it barely changes anymore. At the same time, the polymer RI peak integral is reduced drastically between start of the reaction and 300 min, indicating an efficient cleavage of the polymers.
The mass concentration-dependent integrals of RI signals can further be used to quantify the enzyme products of distinct size. Figure 4d shows the proportion of polymers and large oligomers (DP > 10) compared to small oligomers, with the latter fraction increasing and the former fraction decreasing linearly during the first 120 min. At this timepoint, half of the substrate was cleaved into small oligomers of DP ≤ 10. Afterwards, the curves level out, either because the enzyme lost its activity over time or because the reaction came close to its end point. Naturally, the same trend is visible when only plotting the amount of the small oligomers in micrograms (Fig. 4e), obtained by multiplying the proportion of small oligomers per timepoint with the total amount of chitosan per injection, in this case 3 µg.

In a second data analysis step, the MS and RI data are combined. The obtained dimensionless amounts of each oligomer of a certain DP and FA can be plotted for each timepoint as oligomer product profiles (Fig. 4f, Fig. 6a) which give insights how the product composition changes over time. At the earliest timepoint of 15 min, all products are fully deacetylated (D3-D9). Over the course of the reaction, two effects can be derived from the product profiles: First, the large fully deacetylated oligomers D6-D10 were further cleaved into even smaller oligomers, resulting in a high share of DP 2-5 products at late timepoints. Second, the proportion of products above DP 5 with one or two acetyl groups increases, for example A1D5, A1D6 or A2D7. This fits exactly to previous reports on CsnMN; the enzyme preferentially cleaves fully deacetylated sites of the substrate due to its high preference for D-units at subsites -3 to +3\textsuperscript{27}, resulting in fully deacetylated products. After these sites are hydrolyzed in the initial reaction phase, CsnMN binds increasingly to less acetylated sites, because some of the enzyme's subsites also accept A-units, but with lower affinity\textsuperscript{18}. The resulting partially acetylated products are again no optimal substrates and, thus, accumulate. This increasing acceptance of acetylated units is also visible in Figure 4g where the average FA of the small oligomer products calculated from the product profiles is plotted over time. Based on the weight percentages of each oligomer shown in Figures 4f and 6a, and the absolute amounts of oligomer products shown in Figure 4e, the total amount of individual oligomers can be followed over time (Fig. 4h, Fig. 6b). For example, D8 and D9 both increased in the beginning, but were further cleaved into smaller oligomers over the course of the reaction. The longer the fully deacetylated substrate, the more likely the enzyme cleaves; hence, the amount of D9 decreased earlier than that of D8. As for the product profiles and the average FA of oligomer products, the time courses of individual oligomers show an increasing acceptance of A-units. For DP 8, A1D7 started to accumulate after 60 min and even A2D6 was produced after 180 min. The longer the oligomer, the more likely is the presence of A-units in the product. Therefore, it is not surprising that both A1D8 and A2D7 were formed already after 60 min. A1D8 was even further degraded as indicated by its decrease after 240 min.

Overall, our results support multiple conclusions: CsnMN is an endo-acting chitosanase with a strong preference for D-units, as described before\textsuperscript{27}. Nevertheless, the enzyme tolerates A-units, especially once the strongly deacetylated parts of the substrate are already cleaved. This fits to previous findings of Weikert et al.\textsuperscript{18} that show a nearly absolute specificity for D-units at subsites -2 to +2 for initial timepoints of the reaction, whereas A-units are accepted at subsites +1 and +2 only towards the end point of the reaction for more highly acetylated substrates. In contrast, the strong specificity for D-units at subsites -1
and -2 is maintained throughout different substrates and reaction times\textsuperscript{18,26,27,30}. Because no partially acetylated oligomers of DP < 5 are produced, most A-units are accepted at subsites like -3 or +3 rather than close to the catalytic site. Even though after 45 min still less than 20\% of the sample was converted to smaller oligomers of DP \(\leq 10\), the \(M_w\) of the polymer peak was reduced drastically. All in all, our proposed live digestion method gives detailed insight into the enzyme based on a single enzymatic digestion.

**Conclusion and outlook**

In this study, we established SEC-RI-ESI-MS as a suitable tool for the MW analysis of chitosan polymers while simultaneously performing a semi-quantitative analysis of chitosan oligomers. Taken together in the proposed live digestion method to characterize chitosan-cleaving enzymes, we can i.a. follow the polymer cleavage over time, quantify the produced COS and plot the changing product profile over the course of the reaction. Our goal was not to replace currently used methods for enzyme characterization – especially quantitative MS sequencing\textsuperscript{23} is still the only tool to analyze quantitative subsite preferences of enzymes. But our SEC-RI-ESI-MS setup offers a faster and easier analysis of many aspects than currently used methods. For example, many previously reported insights into CsnMN\textsuperscript{18,26,27} derived from various laborious experiments were gained in this study by just a single live digestion experiment. This could enable mid throughput screenings of different enzymes on the same substrate and/or the same enzyme on different substrates, each for multiple timepoints along the reaction. Apart from making analyses faster and easier, SEC-RI-ESI-MS is the first method to analyze the degradation of the polymeric substrate along with the formation of oligomer products, an aspect that has often been neglected before.

In addition to live measurements, the setup offers the possibility of live sampling and stopping the reaction for downstream processing. One would need to prepare vials with a suitable amount of acidic solution and the autosampler could transfer sample into these vials in addition to the injection into the SEC-RI-ESI-MS system. The acidic conditions and volumes should be chosen such that the enzymatic reaction is stopped – the resulting samples could be \(N\)-acetylated and labeled for quantitative MS sequencing\textsuperscript{23} or used in bioactivity assays. Stopping under alkaline conditions is not suitable for chitosans, as chitosan polymers precipitate at high pH, unless this precipitation would be intended in order to harvest only oligomeric products. Apart from chitosans, the SEC-RI-ESI-MS setup might be suitable for the analysis of other polysaccharides as well. An interesting candidate is pectin, or more precisely homogalacturonan, a linear polysaccharide consisting of non-methylesterified and methylesterified galacturonic acid units. But for the RI- and MS-based quantification to work, it needs to be confirmed that the ionization efficiency of pectin oligomers of the same DP is (nearly) independent of their degree and pattern of methylesterification (section “Oligomer analytics”). But even if the quantification via integration of RI and MS signals is not feasible for certain polysaccharides, the SEC-RI-ESI-MS live digestion makes it possible to analyze any depolymerizing enzyme acting on soluble glycans over time – considering the declining polymer and the formed oligomers simultaneously.
Declarations

Data availability

The full datasets generated and analyzed during this study are available from the corresponding author upon request.

Data availability

The full datasets generated and analyzed during this study are available from the corresponding author upon request.

Acknowledgments

Margareta J. Hellmann is grateful for the financial support of the German Academic Scholarship Foundation within the framework of a doctoral scholarship. The authors acknowledge the financial support of the Open Access Publication Fund of the University of Münster (WWU) and thank Dr. Martin Bonin for critically reading the manuscript. Generous gifts of chitosans by Dr. Dominique Gillet (Gillet Chitosan) and Dr. Katja Richter (Heppe Medical Chitosan) are gratefully acknowledged.

Author contributions

All authors designed the experiments and discussed the results. M.J.H. performed the experiments and wrote the main manuscript. All authors reviewed and revised the manuscript.

Additional information

Supplementary information

The online version contains supplementary information in zip format available at DOI. Apart from a pdf file including i.a. all supplementary tables and figures, these also include exemplary csv files and the Python script for the combined analysis of RI and MS data.

Competing interests

The authors declare that they have no competing interests.
References


**Figures**

all components except enzyme in MS vial

Heated LC autosampler

Add enzyme

$t_0$

Automated live sampling over whole reaction time

Raw data of $t_{1-x}$

MS

RI

Python

Origin

CSV

Detailed insights into chitosan cleaving enzyme

**Figure 1**

Workflow of a single live digestion experiment.
Figure 2

SEC-RI-ESI-MS for the analysis of the $M_w$ of chitosans. **a.** Relationship between $M_w$ determined by SEC-RI-MALLS and RT of the RI signal maximum during SEC-RI-ESI-MS. Chitosan polymers from the company HMC and other suppliers, and oligomers (Kitostim from the company Gillet Chitosan) as well as pullulan standards were measured in the SEC-RI-ESI-MS setup with two SEC columns of different pore size, 125 Å (1-80 kDa optimum, yellow) or 200 Å (10-450 kDa optimum, blue). For smaller analytes, the MW was derived from the m/z values measured by MS; otherwise, the $M_w$ was either given by the supplier for pullulan standards or determined by SEC-RI-MALLS for chitosans. **b.** Exponential relationship between $M_w$ and RT of the RI signal maximum for chitosan polymers from HMC measured with SEC column of pore size 125 Å (yellow) or 200 Å (blue). **c.** $M_w$ distributions of selected HMC samples measured via SEC-RI-MALLS, or **d.** calculated from the RI chromatograms of the SEC-RI-ESI-MS setup (200 Å column) using the established relationship between $M_w$ and RT.

Figure 3

Response factors of different oligomers in RI and MS. All bars are averages of three replicates with 1 µg oligomer injected into the SEC-RI-ESI-MS system. Ax describes chitin oligomers of DP x, if the PA of the tetramers is indicated, the monomer sequence is written-out from non-reducing to reducing end. **a.** Integrated RI peaks of COS of different DP normalized to A3. **b.** Integrated RI peaks of tetramers of
different FA and PA normalized to AADA. c. Arbitrary intensities in ESI-MS of COS of different DP normalized to A3. d. Arbitrary intensities in MS of tetramers of different FA and PA normalized to AADA.

Figure 4

Data output of the SEC-RI-ESI-MS setup from a single enzymatic live digestion. A chitosan polymer of FA 0.22 was incubated with chitosanase CsnMN over 300 min, separation was performed with the 125 Å column. a. Selected raw RI chromatograms (more detailed in Fig. 5). b. Polymer peak in RI chromatogram with $M_w$ x-axis based on relationship of $M_w$ and RT established in section “Polymer analytics”. c. $M_w$ of polymer peak maximum.
polymer peak maximum in RI chromatogram over time calculated based on relationship of $M_w$ and RT established in section “Polymer analytics”. d. Proportion of small oligomers (DP 2-10) vs. polymers and large oligomers (DP > 10) over time based on RI signal integrals. e. Amount of small oligomers (DP 2-10) over time based on RI signal integrals and injected amount of chitosan (3 µg). f. Profiles of weight percentages of small oligomer products (DP 2-10) over time (more detailed with legend in Fig. 6a). g. Average FA of small oligomers (DP 2-10) over time based on weight percentages. h. Selected time courses of oligomers as amount of certain oligomer product over time (more detailed with legend in Fig. 6b).

![Figure 5](image)

**Figure 5**

*Selected RI chromatograms of the live digestion of FA 0.22 chitosan polymer with CsnMN.* Shown are the raw RI data of four timepoints from the start until the end of the reaction.
Figure 6

Amounts of oligomer products of specific DP and FA over time. The data are presented as a. weight percentages of each oligomer product in stacked bars per selected timepoint, or b. amount of selected oligomer products plotted over time. In both cases, the amounts are derived from the combination of RI signal integrals and signal intensities in the MS.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryinformation.zip