Determination and Quantification of Collagen Types in Tissues Using HPLC-MS/MS

Statis Pataridis, Adam Eckhardt, Kateřina Mikulíková, Pavla Sedláková and Ivan Mikšík*

Institute of Physiology, Academy of Sciences of the Czech Republic, (v.v.i.), Vídeňská 1083, Prague 4, 142 20, Czech Republic

Abstract: A method for the determination and quantification of collagen types (I - V) using sample pretreatment has been developed. This work is a continuation of our previous work dealing with the determination of collagen types I and III in tissues [1].

The tissues (rat placenta and porcine cartilage) were firstly homogenized with a pestle in a grinding mortar with liquid nitrogen. Collagens were isolated from these tissues by cleavage with pepsin. The collagen types of interest were then precipitated successively by adding sodium chloride. For quantitation purposes, the sample preparation protocol has been simplified to the one-step precipitation of collagens from a solution containing 4.5 mol/L of sodium chloride. The fractions were fragmented by cyanogen bromide and digested with trypsin. After that, HPLC-MS/MS (high performance liquid chromatography coupled to an ion trap mass spectrometer) analyses of the resulting peptide mixtures (peptide maps) were performed. Based on these analyses, specific (marker) peptides for each of the collagen types were selected. The marker peptides were then synthesized and used to identify and quantify the above-mentioned collagen types in tissues using HPLC-MS/MS, and for determining the limits of detection and quantification.

The applicability of this method for collagen analysis was demonstrated.

Keywords: Collagen, Collagen types, HPLC-MS/MS, Proteomics.

1. INTRODUCTION

Collagens are a family of extracellular matrix proteins that play a dominant role in maintaining the structure of various tissues and also have many other important functions (for example adhesion, tissue remodeling). Collagens are the most abundant proteins in the human body, constituting approximately 30% of its protein mass. Vertebrates have at least 27 collagen types with 42 distinct polypeptide chains, and approximately 20 additional proteins with collagen-like domains and approximately 20 isoenzymes of various collagen-modifying enzymes [2, 3]. The most abundant collagens form extracellular fibrils or network-like structures, but the others fulfill a variety of biological functions. Fibril-forming collagens represent a set of at least nine different polypeptide chains, which constitute the molecular species of type (I, II, III, V, XI, XXIV and XXVII) collagens.

Collagen type I is predominantly present as a heterotrimeric molecule, composed of two α1(I) chains and one α2(I) chain, [α1(I)]2 α2(I), while the homotrimeric form [α1(I)]3 has been shown to occur at low levels in normal adult skin [4]. Type I collagen is frequently accompanied by several other collagen types such as type III and type V collagen [5]. Type II collagen is present in cartilage and the vitreous humor in association with type XI collagen [6]. Type IV collagen is only found in basement membranes, where it is the major structural component.

The determination of the molar ratios of particular collagen chains became of great importance due to its relationship with some deseases. For example, it has been shown that the more hydrophobic α2(I) chain stabilizes the heterotrimeric form of type I collagen [7]. On the other hand the homotrimeric collagen type I, which is composed of three α1(I) chains, [α1(I)]3, has been shown to be associated with certain forms of Ehlers-Danlos syndrome [8-10] and osteogenesis imperfecta [11, 12]. It has also been reported that the presence of the homotrimeric type I collagen isotype significantly weakens the aorta [13, 14].

Some of the collagen type II peptides can serve as potential biomarkers of the activity of matrix metalloproteinases [15], which leads to the destruction of cartilage, and therefore plays an important role in osteoarthritis.

Another example of clinical importance is the determination of the collagen I/III ratio. There is a lot of information indicating that their relative proportions change under conditions of inflammation, during wound healing [16, 17], and under some pathological conditions [4, 18].

Abnormalities in the basement membrane collagen structure and function are connected to both inherited and acquired diseases.

Above are just a few examples that illustrate the great importance of the determination of collagen types in tissues. In fact, over 400 mutations in just 6 different collagens cause a variety of human diseases that include osteogenesis imperfecta, chondrodysplasias, some forms of osteoporosis, some forms of osteoarthritis, and the renal disease known as Alport syndrome [2]. Considering all of the above-mentioned
examples, it is not surprising that there is a desire for a method which would allow us to determine and quantitate various collagen types or collagen chains.

While, there are many papers available dealing with the determination of water soluble proteins, the quantification of insoluble proteins such as collagens is usually accompanied by a lot of problems arising from the difficulties in sample preparation.

To date, several ways of quantifying collagen and/or collagen types or even particular collagen chains have been published such as the quantitation of hydroxyproline, which accounts for approximately 10% of the collagen molecule [19-21] or a method based on a spectrophotometric assay of collagen stained by a strong anionic dye (Sirius red F3BA) in picric acid solution [22-24]. However, the specificity of these methods is at least disputable.

Some of the more sophisticated methods include the radioactive labeling of proline and enzyme immunoassays using antibodies specific to each collagen type. These methods are very specific, but also very expensive and require many sample handling steps [25]. Other options which enable scientists to determine the molar ratios of particular collagen types employ the separation of the peptide mixture produced by enzymatic digestion by various separation methods, such as CE, SDS-PAGE or HPLC [26] and their detection by UV, mass spectrometry (MS) or tandem mass spectrometry (MS/MS) [1, 27].

In this study, we employed the HPLC-MS/MS analysis of specific marker peptides in the peptide mixture produced by cyanogen bromide / trypsin digestion. This approach was originally utilized by us for determining collagen types I and III in skin, tendon and aorta [1]. However, the usability of the above-mentioned method was limited to those tissues where a higher collagen content could be expected. Tissues with a lower collagen content require pre-concentration of the sample prior to its analysis. Therefore, an important step in the method presented here is the sample preparation (pre-concentration by salt), which pre-concentrates the sample, so that it is possible to determine collagen types in tissues even when direct cleavage by cyanogen bromide / trypsin is not applicable.

2. MATERIALS AND METHODS

2.1. Calibration Standards and Chemicals

Three of the peptides used for calibration (GSEGQPGVR, GDQGPVGR, GAGIFPGPEGGK, where P represents hydroxylated proline) were synthesized by Vidia (Jesenice u Prahy, Czech Republic), the other five (GPAIPQGPR, VGAIPGAPGAR, TGPAGAAAGAR, GPGGVPFGGSR, and GPEGPQQQR) were synthesized by Peptide 2.0 (Chantilly, VA, USA). The internal standard – tryptophanyl-glycine (WG), cyanogen bromide (CNBr), formic acid, and trypsin (lot 51K72501) were obtained from Sigma (St. Louis, MO, USA). The water used in the experiments was MilliQ, acetonitrile (HPLC gradient grade) was obtained from Merck (Darmstadt, Germany).

2.2. Tissues and Animals Used

Two collagen sources were used for both the isolation of collagens and their analyses. Rat placenta (two samples marked as A and B) obtained from Wistar strain rats bred in the Department of Physiology, 2nd Medical Faculty, Charles University (Prague, Czech Republic) was used for the identification of collagen types I, III, IV and V. Porcine cartilage served as a source of collagen II.

2.3. Sample Preparation

2.3.1. Identification of Marker Peptides

Placenta samples: The tissues were lyophilized and homogenized using a mortar and pestle. The resulting powder was then processed according to the procedure shown in Fig. (1), which uses collagen precipitation from sodium chloride solutions of various concentrations [28]. This procedure is often used to separate collagen types I-V. In our case, it was used to obtain fractions that were enriched in particular collagen types, so that it became possible to identify them. In order to minimize losses (for quantitation purposes), the isolation procedure was then simplified to include only one step (suspension in 4.5 mol/L NaCl + 50 mmol/L Tris-HCl, pH=7.5; (Fig. (2)). 10 µL of the internal standard stock solution (1 mg/mL) was added prior to analysis.

Cartilage samples: The cartilage was obtained from a porcine joint using a scalpel. The sample was lyophilized, frozen with liquid nitrogen and ground using a pestle and a mortar. The resulting powder was processed as shown in Fig. (2).

The fractions obtained by the processing described in Figs. (1) and (2) were cleaved with CNBr in 70% (v/v) formic acid under nitrogen. The samples were repeatedly dried (with a controlled warm air flow) and reconstituted in water (three times). After removing the CNBr, the samples were reconstituted in pH 7.8, 1 mL 0.05 mol/L ammonium bicarbonate buffer and treated with trypsin (1 : 50 enzyme/substrate ratio) at 37°C for 3 hours.

2.4. Total Protein Concentration

The total protein concentration was determined by amino acid analysis using the Pico Tag method (acid hydrolysis of protein followed by RP-HPLC determination of amino acids after pre-column derivatization by phenylisothiocyanate) [29].

2.5. Conditions for HPLC-MS/MS

Chromatographic separation was carried out in a Jupiter Proteo 90 A column, 250 x 2 mm (Phenomenex, Torrance, CA, USA). The HPLC apparatus used was a HP 1100 LC system (Agilent, Palo Alto, CA, USA) consisting of a degasser, a binary pump, an autosampler, a thermostated column compartment and a diode array detector. It was coupled to an ion-trap mass spectrometer (Agilent LC-MSD Trap XCT-Ultra).

The separation was achieved via a linear gradient between mobile phase A (water-formic acid, 100 : 0.03, v/v) and B (acetonitrile-formic acid, 100 : 0.025 v/v). Separation was started by running the system isocratically for two minutes with 2% mobile phase B, followed by a gradient elution to 35 % B at 40 minutes. Finally, the column was eluted with 100% B for 10 minutes. Equilibration before the next run
was achieved by washing with 2% mobile phase B for 10 minutes. The flow-rate was 0.25 mL/min, injection volume was 40 µL, and the column temperature was held at 25 °C.

Atmospheric pressure ionization-electrospray ionization (API-ESI) positive mode ion-trap mass spectrometry was used. Operating conditions: drying gas (N₂), 10 L/min; drying gas temperature, 350°C; nebulizer pressure, 172.37 kPa.

For peptide selection experiments, ions were observed over the mass range m/z 100-2200 (MS - standard mode, MS/MS - enhanced mode). Analysis was done in auto MS/MS mode (10 precursor ions, excluded after 2 spectra for 0.5 min). The analyses of selected peptides were performed in multiple reaction monitoring mode (precursor ions were selected with respect to the peptides studied - see Results). The fragmentation amplitude was set to 1.14 V.

Analysis of MS/MS data (peptide/protein identification and searching for possible post-translational modifications) was carried out using SpectrumMill software (v.3.02, Agilent). Searches were performed in the full protein databases SwissProt and NCBI nr and then on the data extracted from these databases.

2.6. Determination of Detection and Quantitation Limits

The limits of detection and quantitation for each marker peptide were determined from the calibration curve as 3N/m and 10N/m, respectively, where N is the noise and m is slope of the calibration curve (height of the peak vs. concentration). The values obtained were then recalculated for the whole chains using their molecular weights.

3. RESULTS

3.1. Marker Peptide Selection

The fractions obtained by processing placenta samples (see Fig. (1)) were analyzed for the presence of peptides unique to type I, III, IV, and V collagens. The cartilage sample was used for the identification of the characteristic peptides of type II collagen. The analyses showed that each collagen chain can be represented by at least one marker peptide.
peptide. The peptides selected as marker peptides had to meet all of the following criteria at the same time:

a) once detected in one sample, they had to be detected in all samples of a given type,
b) they had to be part of the collagenous domain,
c) they had to be CNBr/tryptic peptides, i.e. peptides obtained by cleavage with CNBr/trypsin at the specific sites,
d) their peak area in HPLC-MS/MS chromatograms of the standard samples had to be reproducible,
e) they had to be unique to a given collagen type, i.e. the peptide could not be part of any other protein structure in the SwissProt/NCBInr databases,
f) they had to be common to as many species as possible (preferably to human, bovine, and rat).

The presence of collagen marker peptides in particular fractions is illustrated in Figs. (1) and (2). The retention times and masses of both parent and product ions for α1 chains of type I - IV, and α2 chains of type I and V collagen marker peptides are summarized in Table 1.

The structures of all marker peptides were confirmed by HPLC-MS/MS (Figs. (3A) and (3B)). According to the Swiss-Prot database, each marker peptide was present only once per collagen chain for bovine, human and rat samples.

3.2. HPLC-MS/MS Analyses of Samples

Two types of tissue (placenta and cartilage) were analyzed for their content of collagen types I-V. Furthermore, the total protein concentration was determined for both the placenta and cartilage samples.

α1(I), α2(I), α1(III), and α1(IV) collagen chains were detected in the placenta when using the preparation procedure shown in Fig. (2). Types II and V (i.e. α1(II) and α2(V) chains) were not detected. The α2(V) collagen chain was detected only when analyzing pellet IV prepared as shown in Fig. (1). However, this is not surprising, because pellet IV (Fig. (1)) contains only collagen type V and a fraction of the total amount of collagen type I. Therefore, it is to be expected that the concentration of collagen type V in pellet IV will be higher when compared to the more complex mixture in pellet II (Fig. (2)), where all collagen chains are expected.
Table 1. List of Characteristic Marker Peptides for Collagen Types I - V

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Collagen Chain</th>
<th>HPLC Retention Time [min]</th>
<th>Charge State</th>
<th>Precursor MH⁺</th>
<th>Precursor Mass</th>
<th>Product Mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSEGPGGVR</td>
<td>α(I)</td>
<td>13.7</td>
<td>2+</td>
<td>886.3</td>
<td>443.7</td>
<td>613.3</td>
</tr>
<tr>
<td>GPGPQGPR</td>
<td>α(I)</td>
<td>13.7</td>
<td>2+</td>
<td>836.4</td>
<td>418.7</td>
<td>611.4</td>
</tr>
<tr>
<td>GDQGPVGR</td>
<td>α(I)</td>
<td>13.0</td>
<td>2+</td>
<td>785.3</td>
<td>393.2</td>
<td>485.3</td>
</tr>
<tr>
<td>VGPAPGAR</td>
<td>α(I)</td>
<td>13.9</td>
<td>2+</td>
<td>852.5</td>
<td>434.7</td>
<td>321.3</td>
</tr>
<tr>
<td>TGPAGAAGAR</td>
<td>α(I)</td>
<td>12.7</td>
<td>2+</td>
<td>828.3</td>
<td>414.7</td>
<td>335.8</td>
</tr>
<tr>
<td>GGAGPGPEGGK</td>
<td>α(III)</td>
<td>13.0</td>
<td>2+</td>
<td>980.5</td>
<td>498.7</td>
<td>377.8</td>
</tr>
<tr>
<td>GPPGGVGFGSRR</td>
<td>α(IV)</td>
<td>20.1</td>
<td>2+</td>
<td>1116.5</td>
<td>558.8</td>
<td>962.6</td>
</tr>
<tr>
<td>GPEGPGQQR</td>
<td>α(V)</td>
<td>12.3</td>
<td>2+</td>
<td>925.4</td>
<td>463.2</td>
<td>642.4</td>
</tr>
<tr>
<td>WG (internal std.)</td>
<td>-</td>
<td>17.4</td>
<td>2+</td>
<td>262.1</td>
<td>262.1</td>
<td>245.0</td>
</tr>
</tbody>
</table>

A)
The total protein concentration in pellet II (Fig. (2)) was found to be 94 and 96% (w/w) for placenta samples A and B, respectively. The cartilage samples had an average total protein concentration of 21%.

In order to determine the ratios of collagen chains in tissues, the calibration mixtures of the GSEGPQGVR, GPGPQGPR, GDQGPVGR, VGAPGAGAR, TGPAGAGAR, GAGPQPGEGGK, GPQGGVGFGSR, and GPEGPQGQR peptides were analyzed and the results obtained were recalculated for the whole chains using the molecular weights of the collagen chains. The results are summarized in Table 2. The molar concentrations and ratios obtained using the GSEGPQGVR and GPAGPQGPR peptides (both being part of the α1(I) chain) were identical. Similarly, the same results were obtained for GDQGPVGR and VGAPGAGAR, peptides that are part of the α2(I) chain. The detection and quantitation limits (also recalculated to collagen chains) are summarized in Table 3.

The Swiss-Prot search revealed the presence of other proteins in both the placenta and cartilage samples, such as the fibronectin precursor for placenta and keratin type I and II for placenta and cartilage.

4. DISCUSSION

HPLC-MS/MS is a powerful tool not only for the identification, but also quantification of peptides and proteins. While the use of marker peptides has been previously employed by us for the identification and quantification of collagen types I and III in aorta, skin and tendon [1], the quantification of other collagen types (particularly II and IV) remained a problem, even when applying this method to a tissue where a higher content of these collagen types was expected. In fact, the direct cleavage of placenta or cartilage with CNBr/trypsin didn’t provide satisfactory results. In this work, we have extended the use of HPLC-MS/MS to quantify other collagen types using marker peptides by modifying the sample preparation protocol.
It was found that using the CNBr/trypsin digestion of collagen types I-V leads to peptide mixtures containing, among others, peptides characteristic for a given collagen type. The versatility of the proposed method lies in the fact that each of these peptides is common to at least three species - human, bovine, and rat.

The usability of the proposed method was demonstrated by determining the content of collagen types in real tissues. The molar ratio of three main collagen chains found in placenta - α1(I) : α2(I) : α1(IV) was found to be 2 : 1 : 1.5 and 2.5 : 1 : 3.8 for the A and B placenta tissues, respectively. The content of α1(III) was very low (approx. 0.84% w/w of the total protein concentration). The molar ratio α1(I) : α2(I) = 2 shows that the type I collagen found in placenta A is predominantly composed of heterotrimeric molecules [α1(I)]2α2(I). For placenta B, the ratio of α1(I) : α2(I) = 2.5, indicates an increased homotrimer content. As expected, only the α1(II) and α1(III) collagens were found in cartilage. Similar results were obtained using CE-MS/MS (data not shown).

The method developed here represents a simple way of determining and quantifying the various collagen types in tissues using HPLC-MS/MS. Therefore, it retains the advantages of other highly specific methods of quantitation while avoiding the problems associated with them.

ACKNOWLEDGEMENTS
This work was supported by the Grant Agency of the Czech Republic, grants Nos. 203/09/0675, 203/08/1428, the Center for Cardiovascular Research 1M0510, and by Research Project AV0Z50110509.

ABBREVIATIONS
Dial = Dialysis
Lyo = Lyophilization

REFERENCEs
Determination and Quantification of Collagen Types


