Use of Raman spectroscopy and size-exclusion chromatography coupled with HDX-MS spectroscopy for studying conformational changes of small proteins in solution

Rui Wang, Ian Mangion, Alexey A. Makarov, Dmitry Kurouski

Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843, USA
Merck & Co., Inc., MBL Analytical Research & Development, 33 Avenue Louis Pasteur, Boston, MA 02115, USA
The Institute for Quantum Science and Engineering, Texas A&M University, College Station, Texas, 77843, USA

ARTICLE INFO
Article history:
Received 21 March 2020
Received in revised form 19 May 2020
Accepted 28 May 2020
Available online 10 June 2020

Keywords:
Raman spectroscopy
SEC-HDX-MS
Ubiquitin
Insulin
Protein melting

ABSTRACT
Protein-based drugs are a relatively new paradigm in modern therapeutics. These large molecule drugs often have much higher specificity and selectivity compared to small molecules that have been used in therapeutics for centuries. However, there are many analytical challenges associated with drug discovery and development of these new modalities. One of these analytical challenges concerns fast and robust assessment of peptides or small protein conformational structures in solution. In this study, we report a novel analytical approach that is based on Raman spectroscopy (RS) and size exclusion chromatography-hydrogen-deuterium exchange-mass spectrometry (SEC-HDX-MS) for probing conformational structures of proteins in solution. Specifically, we demonstrate that RS and SEC-HDX-MS can be used to probe temperature-induced changes in ubiquitin and insulin. We also show that a combination of these techniques provides a more comprehensive analysis and comparison of peptide or small protein conformational structures than by any one technique. Our results demonstrate that RS and SEC-HDX-MS allow for elucidation of sequential transformations in α-helix and β-sheet content of these proteins. These findings suggest that the proposed approach can be used for a fast investigation of changes in protein or peptide secondary structures under different solution conditions.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

In recent years, pharmaceutical industry efforts have included developing peptide and protein-based therapeutics [1]. Peptide and protein-based therapeutics have many advantages in terms of specificity and selectivity; however, there are a lot of analytical challenges that come with drug discovery and development of these new modalities [2,3]. One of these analytical challenges concerns fast and robust assessment of peptides or small protein conformational structures in solution. Nowadays, the synthesis of peptides for drug discovery efforts has evolved to include the simultaneous synthesis and purification of hundreds of compounds [4], which makes speed of analysis paramount for guiding both lead identification and optimization processes.

Raman spectroscopy (RS) is a modern analytical technique that provides information about molecular vibrations and consequently the structure of the analyzed specimen. The Raman effect is based on inelastic light scattering of photons by molecules that are being excited to higher vibrational or rotational states. RS has been broadly used in various research fields ranging from forensic analysis of bodily fluids and pesticides to food science and electrochemistry [5]. Our group recently demonstrated that Raman spectroscopy can be used to detect and identify plant diseases [6–8]. Specifically, we were able to identify whether maize kernels were healthy or infected by Aspergillus flavus, A. niger, Fusarium spp., or Diplodia spp. with 100% accuracy [6].

RS in general and deep UV resonance Raman (DUVRR) spectroscopy in particular can be used to monitor conformational changes in proteins [9]. For instance, using RS, Kurouski et al. demonstrated that conformations of disulfide bonds of insulin could be determined upon protein aggregation in amylloid fibrils [9]. The researchers demonstrated that aggregation of insulin occurred without scrambling of disulfides. The Lednev group pioneered the use of DUVRR to investigate structural changes in insulin
and other amyloid associated proteins that were taking place upon amyloid fibril formation [10–12]. Also, the Lednev group demonstrated that DUVRR could be used to probe structures of amyloid fibril polymorphs, aggregates that are formed from the same protein under the same or slightly different experimental conditions [13].

A typical Raman spectrum of a protein is composed of contributions from two major types of modes, vibrations of polypeptide backbone (amide bands) and amino acid residue side chains [14]. Amide modes include: the amide I vibration (1640–1680 cm−1), which primarily represents C=O stretching and a small amount of out-of-phase C–N stretching; the amide II vibration (~1550 cm−1), which consists of an out-of-phase combination of C–N stretching and NH– bending modes; and the amide III vibration (1200–1340 cm−1), a complex vibration mode which involves C–N stretching and NH– bending [15]. Analysis of amide I band is the most commonly used approach for elucidation of structural changes that is taking place in proteins. This is in part due to the overlay of the amide II and amide III bands with the vibrational frequencies of certain stretching modes, such as C–C, CN– and CH2, that substantially complicates their assignment and interpretation. The position of the amide I band depends on the conformation of the polypeptide backbone and intra- and intermolecular hydrogen bonds of the protein specimen. The amide I band (located in the 1665–1680 cm−1 range) corresponds to a β-sheet structure, while α-helical protein secondary structure corresponds to the amide I band (located in the 1640–1654 cm−1 range). The amide I vibration located in the 1654–1665 cm−1 range is typically assigned to unordered or disordered protein secondary structures [15].

Hydrogen deuterium exchange (HDX) is an established methodology for studying protein higher-order structure [16–18]. Labile protons can be exchanged to deuterons in amides, alcohols, carboxylic acids, or amines and the exchange is observable by MS and NMR [19]. Intrinsic H/D exchange rate depends on pH and temperature, for example in aqueous solution (at 25 °C and neutral pH) exchange occurs in a second-time scale [20]. There is an exchangeable NH proton in every peptide bond of a protein (except proline), which makes HDX a very useful technique for studying protein higher-order structure, dynamics and folding [21]. The utility of this technique has been recognized and workflows based on HDX methodology have been robustly implemented in the pharmaceutical and biopharmaceutical industries [22–24]. It was demonstrated that size-exclusion chromatography (SEC) coupled with differential HDX can be successfully used to compare global conformational changes of proteins under different solution conditions in many applications [24,25].

In this study we report an evaluation of the approach that combines two orthogonal techniques, Raman spectroscopy and SEC–HDX–MS, to investigate protein conformational structures in solution. Combination of these techniques aims to provide a more comprehensive analysis and comparison of peptide and protein conformational structures than by any one technique. The proposed approach can be applied for a fast investigation of protein higher-order structure under different solution conditions, as well as for the variant selection of peptide and small protein drug candidates based on conformational structure comparison during lead identification and optimization processes.

## 2. Materials and methods

### 2.1. Materials

Bovine recombinant insulin (bovine) and ubiquitin were purchased from Sigma-Aldrich Inc. (St. Louis, MO). Ultrapure water was obtained from a Milli-Q Gradient A10 from Millipore (Bedford, MA, USA). Trifluoroacetic acid (TFA), ammonium hydroxide and acetonitrile (ACN) HPLC grade were purchased from Fisher Scientific (Fair Lawn, NJ). The insulin and ubiquitin stock solution was prepared by dissolving the protein powder into the DI water with pH adjusted to pH 1.0 and 7.0 respectively. The final concentration of the insulin and ubiquitin stock solution was 140 and 100 mg/mL respectively. To perform the spectroscopic study, 1 mL of the insulin/ubiquitin stock solution was added to a capped glass vial, the vial was then placed in a self-made heater, whose temperature was controlled by circulating water from a thermostat. Besides, the temperature in the heater was also monitored by thermometer. The protein sample with the heater was placed on a piece of coverslip for Raman measuring.

### 2.2. Size exclusion chromatography-hydrogen-deuterium exchange-mass spectrometry (SEC–HDX–MS)

SEC–HDX–MS was performed using a Synapt G1 HDMS mass spectrometer with an Acquity UPLC™ system (Waters Corp., Milford, MA, USA), with Waters MassLynx V4.1 software for instrument control and data processing. An Acquity UPLC® Protein BEH SEC-125 4.6 × 150 mm, 1.7 μm column with 125 Å pore size from Waters was used in all experiments. The aqueous mobile phase (A) was 50 mM ammonium formate adjusted to pH 2.0 with TFA or to pH 5.5 with ammonium hydroxide and the organic mobile phase (B) was acetonitrile. Sample solutions of the proteins used for experiments were prepared at about 0.2 mg/mL in 0.1 v/v % TFA (insulin) or 50 mM ammonium formate adjusted to pH 5.5 (ubiquitin) in water or in deuterium oxide. All analytes were completely dissolved before injections. To label insulin and ubiquitin for SEC–HDX experiments, protein samples in deuterium oxide were incubated for 24 h at 40 °C. The SEC–HDX experiments were conducted in the similar manner as previously reported [24,26]. The SEC conditions consisted of a 5 % B isocratic run at 0.250 mL/min with total run time 10 min. The column conditions mimicked the same temperature unfolding as in Raman Spectroscopy experiments. The MaxEnt1 module was used for all mass spectra deconvolution. Note that in this study, we used an all-or-nothing approach and our SEC method was developed to eliminate exchange restrictions based on intrinsic rate of hydrogen–deuterium exchange for insulin and ubiquitin [24].

Protein solvent accessibility was estimated for the proteins used in this study with the computational chemistry software Biovia Discovery Studio v.17.1.0.16143 (Dassault Systemes Biovia) with input pdb (Protein Data Bank) files 2MJB for ubiquitin and 4IS2 for insulin (Table 1). The solvent inaccessible was defined as less than 10 % of solvent accessible surface (SAS) modeled by the double cubic lattice method (DCLM) [27]. The number of labile protons was calculated for the solvent-inaccessible part of the protein sequence (as determined by the software), based on the respective pdb files.

### Table 1

<table>
<thead>
<tr>
<th>Compound name</th>
<th>Number of solvent inaccessible labile protons*</th>
<th>Calculated total number of labile protons*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insulin</td>
<td>15</td>
<td>91</td>
</tr>
<tr>
<td>Ubiquitin</td>
<td>33</td>
<td>154</td>
</tr>
</tbody>
</table>

*Calculated based on pdb files 4IS2 and 2MJB respectively for insulin and ubiquitin.

** Solvent accessibility was calculated in Biovia Discovery Studio v.17.1.0.16143: solvent inaccessible was defined as less than 10 % of solvent accessible surface (SAS), based on approach by the DCLM method.
Table 2

<table>
<thead>
<tr>
<th>Frequency (cm⁻¹)</th>
<th>Assignment [9,14]</th>
</tr>
</thead>
<tbody>
<tr>
<td>646</td>
<td>Tyrosine</td>
</tr>
<tr>
<td>832</td>
<td>Tyrosine</td>
</tr>
<tr>
<td>859</td>
<td>Tyrosine</td>
</tr>
<tr>
<td>1007–1011</td>
<td>Phenylalanine</td>
</tr>
<tr>
<td>1355</td>
<td>CH deformation</td>
</tr>
<tr>
<td>1453</td>
<td>CH₂ deformation</td>
</tr>
<tr>
<td>1620</td>
<td>Tyrosine, phenylalanine</td>
</tr>
<tr>
<td>1667–1669</td>
<td>Amide I</td>
</tr>
</tbody>
</table>

2.3. Spectroscopy

Raman spectra were collected on an inverted microscope (Nikon TE-2000U) with 20X dry Nikon objective (NA = 0.45). A single longitudinal diode mode laser (Necsel, CA, USA) was used to generate a 785 nm excitation. Spectral acquisition time was 60 s with laser power at the sample of 15 mW. Three samples were tested in parallel for both insulin and ubiquitin. The signal was collected in a backscattering configuration and directed to a confocal IsoPlane SCT 320 Raman spectrometer (Princeton Instruments, NJ, USA) equipped with a 600 groove/mm grating blazed at 750 nm. Prior to entering the spectograph Rayleigh scattering was filtered with LP02–785RE-25 long-pass filter (Semrock, NY, USA). The dispersed light was then sent to PIXIS:400BR CCD (Princeton Instruments, NJ, USA). A motorized stage H117P2TE (Prior, MA, USA) controlled by Prior Proscan II was used to move the sample relative to the incident laser beam. All data was processed using GRAMS/AI 7.0 (Thermo Galactic, NH, USA). The reported spectra are baseline corrected.

3. Results and discussion

3.1. Temperature dependent Raman spectra of ubiquitin in solution

Solution of bovine ubiquitin (100 mg/mL, pH 7.0) was heated from 30 to 60 °C increment of 10 °C for each data point. Raman spectra collected from these solutions exhibited three clearly distinct vibrational bands at 1007, 1453 and 1667 cm⁻¹, which can be assigned to phenylalanine (Phe), CH₂ and amide I vibration respectively, Table 2, Fig. 1. A. For comparison of changes in intensities of vibrational bands, we normalized all collected spectra on 1453 cm⁻¹ peak (CH₂ vibration). This band has been chosen for normalization of Raman spectra because CH₂ groups are present in many amino acids located across the sequence of ubiquitin. Thus, such normalization is unbiased to the specific amino acid.

We have found that intensity of the amide I band gradually decreases with an increase in the solution temperature indicating deformation or melting of the peptide secondary structure, Fig. 1. B. We have also found that intensity of Phe remained nearly unchanged at low (30 °C) and high (60 °C) temperatures, Fig. 1. A, Fig. S1. This indicates that local environment of this amino acid did not change upon peptide melting [9].

Next, ubiquitin solution was gradually cooled down with an increment of 10 °C for each data point, Fig. 2.

We have found that intensity of amide I band gradually increases with an increase in the solution temperature reversible conformational changes in the peptide secondary structure, Fig. 2. B. These spectral changes suggest about re-folding of ubiquitin that is taking place upon depression of the peptide temperature. We have also found that intensity of Phe remained nearly unchanged at low (30 °C) and high (60 °C) temperatures, Fig. 2. A, Fig. S2.

We have plotted a temperature dependent change in the intensity of amide I band (1590–1730 cm⁻¹) to demonstrate that RS can be used for quantitative prediction of conformational changes in the protein secondary structure, Fig. 3. Our results demonstrate that RS can be used to predict a degree of temperature-induced melting of ubiquitin secondary structure. These data also show that RS can be used to probe the degree of ubiquitin re-folding that is taking place at a low temperature (50–30 °C) of the peptide solution.

RS offers a precision analysis of conformational changes that are taking place in different secondary structure elements of proteins [9]. As was discussed above, amide I band consists of individual contributions of β-sheet structure (1665–1680 cm⁻¹) and α-helix (1640–1654 cm⁻¹) [9,28]. Thus, precise analysis of amide I band can be used to investigate changes in the content of these secondary structures that are taking place upon temperature-induced melting and re-folding of proteins.

We fitted amide I in the spectra collected from ubiquitin solution at different temperatures to investigate conformational changes that are taking place in the α-helix (1643 cm⁻¹) and β-sheet (1670 cm⁻¹) of ubiquitin upon heating from 30 to 60 °C, Fig. 4.

We have found that a ratio of α-helix to β-sheet increased from 1.46 to 3.26 as the peptide solution was heated from 30 to 40 °C. These results suggest that α-helix melted faster than β-sheet in this temperature range. Deconvolution of Raman spectra of ubiquitin at 50 °C suggest even further degradation of α-helical component of this peptide, whereas the content of β-sheet does not experience similar changes. This quantitative analysis was limited to the spectrum of ubiquitin at 60 °C due to low signal-to-noise of the amide I region. This suggests significant loss of the peptide secondary structures at this temperature. Nevertheless, one can notice that the β-sheet component nearly disappeared in the Raman spectrum of ubiquitin collected at 60 °C suggesting that melting of β-sheet in ubiquitin is taking place between 50 and 60 °C.
Fig. 2. A. Raman spectra (λ = 785 nm; P = 15 mW; T = 60 s) collected from bovine ubiquitin (100 mg/mL; pH 7.0) at different temperatures (60-30 °C). Normalization was done based on 1453 cm⁻¹ peak (CH₂ vibration). B. Amide I region of these spectra.

SEC-HDX-MS was used as an orthogonal technique to screen for global conformational changes in protein (ubiquitin or insulin) structure upon temperature increase. In this study, deuterium-to-hydrogen exchange was performed on SEC column in the solution environments matching conditions in experiments using RS, and the overall degree of exchange was measured by MS detection on-line. We have previously reported on-line screening for global protein conformational changes in solution using combined SEC-HDX methodology [24-26, 29, 30]. The conformational structure changes were compared based on the number of labile protons which remained un-exchanged (solvent inaccessible) in the protein structure. The number of labile protons which remained un-exchanged (∆HDX) was measured based on the molecular mass difference between the deuterium-oxide-labeled protein after deuterium-to-hydrogen exchange (at specific conditions) and the molecular mass of the same protein not exposed.

Fig. 3. Temperature dependent change in the intensity of amide I band (1590-1730 cm⁻¹) of ubiquitin upon heating (black) and cooling (red). Corresponding changes in HDX profile of ubiquitin upon heating (blue). SEC-HDX-MS solvent: 95% aqueous 50 mM ammonium formate adjusted to pH 5.5/5% acetonitrile.

Fig. 4. Deconvolution of amide I band of ubiquitin in Raman spectra collected at four different heating temperatures (30, 40, 50 and 60 °C).
to deuterium oxide (under the same conditions). The %ΔHDX was calculated based on the total number of labile protons in the protein structure (Table 1). Fig. 3 demonstrates %ΔHDX for ubiquitin at different temperatures, which indicated the ubiquitin conformational structure unfolding upon temperature increase (%ΔHDX decreased upon temperature increase). Interestingly the slope of amide I band intensity upon heating was different than %ΔHDX slope upon heating (Fig. 3). This %ΔHDX slope indicates the speed of overall conformational changes in solution. This may corroborate observations based on RS data that a ratio of α-helix to β-sheet increased upon ubiquitin heating. Since %ΔHDX reflected the combined result of global conformational structure stability, the slower unfolding of β-sheet (demonstrated by RS) resulted in less steepness of the %ΔHDX slope for ubiquitin.

3.2. Temperature dependent Raman spectra of insulin in solution

Solution of bovine insulin (140 mg/mL; pH 1.0) was heated from 30 to 60 °C increment of 10 °C for each data point. Raman spectra collected from these solutions exhibited vibrational bands at 646, 832, 859, 1011, 1355, 1453, 1620 and 1669 cm⁻¹, Table 2. These vibrational bands can be assigned to tyrosine (Tyr) (646, 832 and 859 cm⁻¹), Phe (1011 cm⁻¹) and both of these amino acids (1620 cm⁻¹), as well as CH (1355 cm⁻¹) and CH₂ (1453 cm⁻¹) vibrations. We have also observed an amide I band centered at 1669 cm⁻¹. For comparison of changes in intensities of these vibrational bands, we normalized all collected spectra on 1453 cm⁻¹ peak (CH₂ vibration). This chemical moiety presents in nearly all classes of biological molecules. Therefore, normalization on CH₂ vibration makes spectral interpretation the least biased, Fig. 5.

We have found that the intensity of amide I band gradually decreased with an increase in the solution temperature indicating deformation or melting of insulin secondary structure, Fig. 5.

Fig. 5. A. Raman spectra (λ =785 nm; P =15 mW; T = 60 s) collected from bovine insulin (140 mg/mL; pH 1.0) at different temperatures (30-60 °C). Normalization was done based on 1454 cm⁻¹ peak (CH₂ vibration). B. Amide I region of these spectra.

B. We have also found that the intensity of Tyr and Phe remained nearly unchanged at low (30 °C) and high (60 °C) temperatures, Fig. 1, A, Fig. S3 and S4. This indicates that the local environment of this amino acid did not change upon peptide melting.

A decrease in temperature of insulin solution reflected in the increase in the intensity of amide I band. This suggests about reversible conformational changes (re-folding) in the insulin secondary structure. We have also found that intensity of Tyr and Phe remained nearly unchanged at low (30 °C) and high (60 °C) temperatures. For the clear visualization of these spectral changes, we have plotted a temperature dependent change in the intensity of amide I band (1590–1730 cm⁻¹), Fig. 6. These results demonstrate that RS can be used to predict a degree of temperature-induced melting of insulin secondary structure of insulin. These data also show that RS can be used to probe the degree of insulin re-folding that is taking place at low temperature of the peptide solution.

We performed detailed analysis of changes in amide I band of Raman spectra collected from insulin heated at 30, 40, 50 and 60 °C, Fig. 7. We have found that at 30 °C, presence of both α-helix (1631 cm⁻¹) and β-sheet (1668 cm⁻¹) with the ratio of 0.2 has been observed. At 40 °C, the ratio of α-helix to β-sheet became 0.85 suggesting that substantial amount of peptide β-sheet denatured at this temperature. We have found that at 50 °C, α-helical component of insulin nearly completely disappeared, whereas remaining β-sheet was preserved. We were not able to perform this quantitative analysis of changes in secondary structures of insulin that were taken place at 60 °C due to low signal-to-noise of the amide I region.

These results demonstrate that RS can be used to probe sequential changes in both α-helix and β-sheet content of insulin that are taking place upon temperature induced peptide melting.

Fig. 6. Temperature dependent change in the intensity of amide I band (1590–1730 cm⁻¹) of insulin upon heating (black) and cooling (red). Corresponding changes in HDX profile of ubiquitin upon heating (blue), SEC-HDX-MS solvent: 95 % aqueous 50 mM ammonium formate adjusted to pH 2.0/5 % acetonitrile.

The SEC-HDX-MS experiment for insulin was performed on column in solution environments matching those conditions in the experiment using Raman Spectroscopy (Fig. 6). Upon heating, the SEC-HDX conformational screening for insulin corroborated the observations by RS. Indeed, the slope of insulin’s %ΔHDX upon heating was almost identical to amide I intensity slope by RS upon temperature increase. This is an indication that overall insulin structure unfolding upon heating was mostly based on α-helix unfolding, which resulted in a direct correlation between global unfolding (by SEC-HDX) and specific α-helix unfolding (by RS).

4. Conclusions

The manuscript reports the experimental demonstration of combination of RS with SEC-HDX-MS for a fast elucidation of temperature-induced changes in two model proteins namely insulin and ubiquitin. Our results demonstrated complementar-
ity of RS and SEC-HDX-MS in terms of sensing of changes in protein secondary structure. We also show that in addition to this temperature-induced changes, RS allowed for elucidation of sequential transformations in α-helix and β-sheet content of these proteins. Combination of these techniques was able to provide a more comprehensive analysis and comparison of protein conformational structures in solution than by any single technique. The proposed approach can be applied for a fast investigation of protein higher-order structure in solution, as well as for the variant selection of peptide or small protein drug candidates based on conformational comparison during lead identification and optimization processes.

Compliance with ethical standards

None.

CRediT authorship contribution statement

Rui Wang: Conceptualization, Methodology, Investigation, Writing - original draft, Visualization, Writing - review & editing. Ian Mangion: Supervision, Project administration, Writing - review & editing. Alexey A. Makarov: Conceptualization, Methodology, Investigation, Writing - original draft, Visualization, Writing - review & editing. Dmitry Kurouski: Conceptualization, Methodology, Investigation, Writing - original draft, Visualization, Writing - review & editing. Supervision.

Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are grateful to Merck for provided financial support.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jpba.2020.113399.

References


