ACS Partner Journal

Journal of the American Society for Mass Spectrometry

pubs.acs.org/jasms

Research Article

Label-Free Higher Order Structure and Dynamic Investigation Method of Proteins in Solution Using an Enzymatic Reactor Coupled to Electrospray High-Resolution Mass Spectrometry Detection

Elodie Grifnée, Christopher Kune, Cédric Delvaux, Loïc Quinton, Johann Far,*,† Gabriel Mazzucchelli,† and Edwin De Pauw†

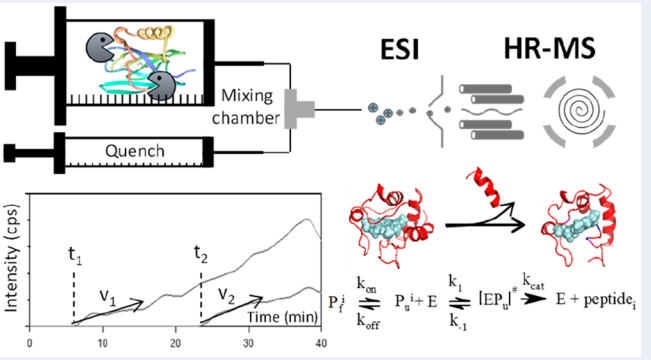
Cite This: *J. Am. Soc. Mass Spectrom.* 2022, 33, 284−295

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

Downloaded via UNIV DE LIEGE on September 2, 2022 at 14:04:29 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

ABSTRACT: For decades, structural analysis of proteins have received considerable attention, from their sequencing to the determination of their 3D structures either in the free state (e.g., no host−guest system, apoproteins) or (non)covalently bound complexes. The elucidation of the 3D structures and the mapping of intra- and intermolecular interactions are valuable sources of information to understand the physicochemical properties of such systems. X-ray crystallography and nuclear magnetic resonance are methods of choice for obtaining structures at the atomic level. Nonetheless, they still present drawbacks which limit their use to highly purified systems in a relatively high amount. On the contrary, mass spectrometry (MS) has become a powerful tool thanks to its selectivity, sensitivity, and the development of structural methods both at the global shape and the residue level. The combination of several MS-based methods is mandatory to fully assign a putative structure in combination with computational chemistry and bioinformatics. In that context, we propose a strategy which complements the existing methods of structural studies (e.g., circular dichroism, hydrogen/deuterium exchange and cross-links experiments, nuclear magnetic resonance). The workflow is based on the collection of structural information on proteins from the apparition rates and the time of appearance of released peptides generated by a protease in controlled experimental conditions with online detection by electrospray high-resolution mass spectrometry. Nondenaturing, partially or fully denatured proteins were digested by the enzymatic reactor, i.e., β-lactoglobulin, cytochrome *c*, and β-casein. The collected data are interpreted with regard to the kinetic schemes with time-dependent rates of the enzymatic digestion established beforehand, considering kinetics parameters in the Michaelis−Menten formalism including k_{cat} (the turnover number), k_1 (formation of the enzyme−substrate complex), k_{-1} (dissociation of the enzyme−substrate complex), k_{off} (local refolding of the protein around the cleavage site), and k_{on} (local unfolding of the protein around the cleavage site). Solvent-accessible surface analysis through digestion kinetics was also investigated. The initial apparition rates of released peptides varied according to the protein state (folded vs denatured) and informs the k_{off}/k_{on} ratio around the cleavage site. On the other hand, the time of appearance of a given peptide is related to its solvent accessibility and to the resilience of the residual protein structure in solution. Temperature-dependent digestion experiments allowed estimation of the type of secondary structures around the cleavage site.

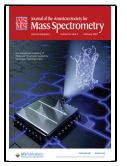


KEYWORDS: *Protein structures, enzymatic reactor, dynamics, kinetic scheme, mass spectrometry*

INTRODUCTION

During the last decades, mass spectrometry (MS) has significantly contributed to the high-throughput sequencing of proteins[1,2] (i.e., primary structures) as well as protein quantification based on label-free or protein-tagging methods of biological origin (although, not exclusively) samples.[3−5] The structure elucidation of proteins based on mass spectrometry usually relies on methods such as hydrogen/deuterium exchange (HDX-MS),[6] chemical cross-linking (XL-MS),[7] and/or ion mobility mass spectrometry (IM-MS).[8] To take full advantage of HDX methods, the primary sequence characterization is a prerequisite, while the knowledge of the PDB structures is highly desirable in the context of ligand binding studies (e.g., epitope mapping, supramolecular interactions...). In the best cases, the structural information

Received: September 4, 2021
Revised: November 5, 2021
Accepted: December 13, 2021
Published: December 31, 2021



ACS Publications

© 2021 American Society for Mass Spectrometry. Published by American Chemical Society. All rights reserved.

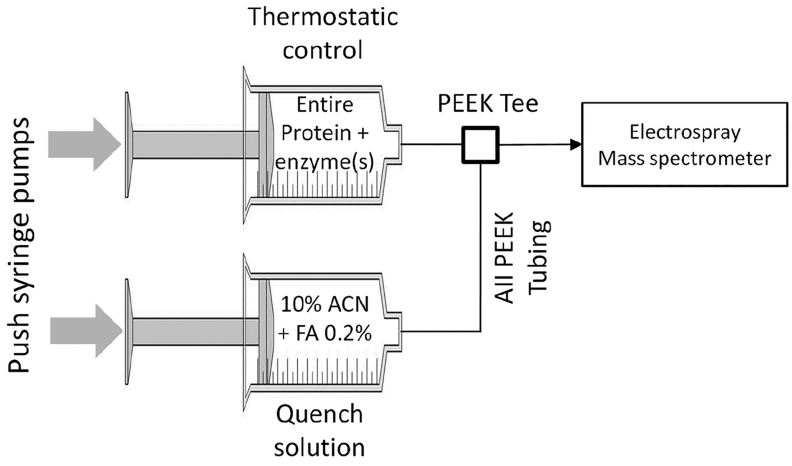


Figure 1. Schematic representation of the enzymatic reactor. Two push syringe pumps and two glass gastight Hamilton RN1725 syringes (250 μL, needle point style3) were used to deliver the protein and enzyme mixture and the quench solution. PEEK tubing (0.007 in. I.D. 1/16 in. O.D. 11 in. length), PEEK tee (3-ways PEEK tee 0.002 in. bore 10-32 thread, Supelco 57661), and PEEK syringe needle ports (Supelco 57639) were used for connecting the devices. The syringe containing the mixture of entire protein and enzyme was temperature-controlled using a Thermo Haake 56-1165-34 Amersham Biosciences (Germany) MultiTemp III (part code 18-1102-78) and a homemade jacketed syringe. See the Material and Methods for details and Figure S1.

that can be retrieved operates at the residue level (i.e., HDX-MS with the use of ECD or ETD). XL-MS provides information at the domain level and suffers from cross-linking between non-interacting regions due to the dynamics of the system (unspecific cross-linking),[9] while efficient software to assist the data processing has been actively developed. IM-MS has to be used with the mandatory support of computational chemistry to extract structural information at the atomic to domain level derived from the experimental collision cross section.[10,11]

Spectroscopic methods such as nuclear magnetic resonance (NMR)[12] and crystallography (XRC)[13] provide the most stable 3D structure at the atomic level and, to some extent, the distribution of the most floppy structures in space. On the other side, circular dichroism (CD)[14] is a very popular spectroscopic method focusing on the evolution of conformation but is rather limited in terms of structural information. Fluorescence resonance energy transfer (FRET) coupled to MS, CD, or NMR are emergent promising methods for structural elucidation purposes but strongly rely on advanced instrumentations. In addition, the capability to bring fluorescent active moieties on the protein structure without affecting the assumed native (local) structure of the model is still under investigation.[15] Moreover, the correlation of the data from these different techniques requires heavy computational chemistry skills.

Limited proteolysis strategies were introduced to improve the sequence coverage of proteins (i.e., production of overlapped peptides), especially in complex mixtures and/or presenting post translational modifications[16,17] In this context, limited (proteolytic) digestion of proteins was promisingly anticipated to be used to probe their surface accessibility and extract global structural information by Fontana et al.[18,19] Although these authors argued that the emergence of MS-based proteomics could lead to unexpected development, they anticipated that the limited digestion contribution in structural studies would be hindered. Indeed, the expected influence of the kinetic constants of the local protein refolding and protein unfolding (k_{off} and k_{on} respectively), should provide information about the access of the proteolytic cleavage sites of the enzyme instead of the surface topology and the surface exposed to the solvent of the cleavage site. This hypothesis is valid in the context of limited digestion performed in batch followed by subsequent analysis of the generated peptides. The direct coupling of an enzymatic reactor with mass spectrometry will cover the kinetics (rate of apparition) and temporal (time of appearance) aspects of the released peptides from the protease activity. Combining the kinetic scheme (i.e., with time dependent rates) in the Michaelis−Menten formalism with the contribution of k_{on} and k_{off}, leads to information about the local flexibility of the peptide chain around the cleavage site as well as the protein surface accessibility to the protease (and the solvent). To this end, the apparition rates and the times of appearance of the MS detected peptides are used in combination with the previously determined primary structure of the protein (straightforwardly assessed nowadays by MS) obtained from the bottom-up strategy.

In this work, an in-house low-volume enzymatic digestion reactor was coupled to a high-resolution mass spectrometry (HRMS) to assess the structural and dynamics information on three model proteins without any labeling strategy (e.g., isotope, fluorochrome, (non)-covalent labeling), i.e., bovine β-lactoglobulin,[20] equine cytochrome c,[21] and bovine β-casein.[22] The results obtained for bovine β-lactoglobulin and cytochrome c were favorably compared to those obtained by NMR (dynamic conformation) to evaluate the relevance and interest of enzymatic digestion kinetics as a structural tool.

Bovine β-casein was selected in the Uniprot data bank because no experimental 3D structure was available (suggesting an intrinsically unordered protein) and was digested by the reactor to challenge the developed method. Our generated data set complements circular dichroism and HDX-MS experiments to interpret the tridimensional structures of protein in the case where PDB files are lacking. The kinetics of the released peptides performed at different temperatures of digestion also provided an indication of the presence of the unfolding of secondary structures as well as the impact of proline residues around the cleavage sites.

https://doi.org/10.1021/jasms.1c00274
J. Am. Soc. Mass Spectrom. 2022, 33, 284−295

MATERIALS AND METHODS

Enzymatic Digestion Reactor Coupled to High-Resolution Mass Spectrometry. The designed enzymatic digestion reactor (see Figure 1) was composed of two pumps, a first one pushing from a thermostatized 250 μL syringe (4 μL.min^{-1}) the targeted protein (about 10 μM) and the enzyme (see details in online enzymatic digestion) and a second 250 μL syringe (4 μL.min^{-1}) filled with formic acid 0.2% (v/v) with 10% acetonitrile (ACN) used to quench the enzymatic digestion in the mixing chamber prior to the MS analysis. The enzyme was added to the sample directly before the injection to follow the digestion kinetics as a function of temperature, reaction time, and enzyme concentration. The reactor was coupled to a SolariX XR 9.4 T ESI-Q-FT-ICR MS from Bruker Daltonics, Bremen, Germany. Instrument conditions were optimized as follows: capillary voltage = 4.5 kV, nebulizer = 1.1 bar, dry gas = 4L.min^{-1}, dry temperature = 200 °C, average scans = 10, accumulation = 0.5, detection mode = broadband, est resolving power = 130000 at m/z 400 and chromatography mode 40 min. All other relevant parameters regarding these experiments are provided in Table S1. Mass spectra were acquired during the digestion time (here until 40 min), allowing identification of generated peptides and their apparition rate during digestion. Peptides were assigned manually. Apparition rates and slopes were determined after extraction of the absolute intensities of the target m/z over the time of the experiments. Dead volume was determined in a separate experiment based on the detection of the cluster ions of infused 0.1 mg·mL^{-1} NaTFA solution (in 50% ACN) and corresponds to an apparition time of around 5 min (depending on the solvent viscosity and the temperature fixed during the experiments which also affect the viscosity of the fluid). The dead time is defined in this work as the contributions of the dead volume of the enzymatic reactor (tubbing and mixing chamber) and the required time to prepare the protein and enzyme mixture in the syringe (note that this latter contribution is negligible compared the dead volume of the enzymatic reactor).

Chemicals. Bovine β-lactoglobulin (P02754), equine hearth cytochrome c (P00004), bovine β-casein (P02666), ammonium bicarbonate (NH_4HCO_3), formic acid (FA), acetonitrile (ACN), dithiothreitol (DTT), and iodoacetamide (IAA) were purchased from Sigma-Aldrich (Bornem, Belgium). MS-grade lyophilized trypsin protease was purchased from ThermoFisher Scientific (Waltham, MA). They were reconstituted on ice to a stock concentration of 1 mg·mL^{-1}, using respectively HCl 10 mM, Milli-Q ultrapure water, and HCl 1 mM before use. All reagents were used without further purification. A Milli-Q ultrapure water system (Millipore, Molsheim, France) was used throughout the study.

Protein Denaturation Reaction. Protein denaturation (disulfide bonds reduction followed by iodoacetamide alkylation) was performed as follows in a three-step reaction in buffered conditions (NH_4HCO_3 (50 mM)): (a) reduction by DTT (10 mM) at 56 °C for 40 min, (b) alkylation by IAA (20 mM) at room temperature for 30 min, and (c) quenching of the alkylation reaction by addition of DTT (11 mM). The denatured samples were then washed twice with NH_4HCO_3 500 mM and twice with NH_4HCO_3 50 mM on an Amicon Ultra centrifugal filters (3K, 0.5 mL, Merck Millipore, Ireland) at 15000 rpm (Eppendorf Centrifuge 5415 R) according to the manufacturer's protocol.

Online Enzymatic Digestion. Protein samples—native-like folding or fully denatured—were diluted in preheated NH_4HCO_3 (50 mM) to a final concentration of 10 μM in a 1.5 mL Eppendorf tube. For enzymatic digestion, on-ice trypsin was added in protease-to-protein ratios of 1/1,500, 1/1,000, 1/500, and 1/100 and then immediately mixed using vortex agitator and transferred to the enzymatic reactor (syringe) coupled to ESI-MS. Digestion temperatures of 10, 21, 40, and 60 °C were also investigated. For cytochrome c digestion, trypsin was also used in protease-to-protein ratio of 1/1000 at 40 °C.

Software. *MS Data Processing.* DataAnalysis 5.0 was used for data processing generated from the SolariX XR 9.4T (Bruker). Sequence editor v 3.2 (Bruker) was used to perform the in-silico digestion and the prediction of the potential peptides that could be released during the experiment. Pymol v 1.5.0.5 was used to visualize the PDB (NMR resolved PDB) structures of the investigated proteins (i.e., bovine pancreas trypsin (2PTN), bovine β-lactoglobulin (1CJ5), and equine hearth cytochrome c (1AKK) (www.rcsb.org)) as well as determining the position of the cleavage sites within the protein structure and identification of α-helices, β-sheets, and loops.

For HDX-MS experiments, 20 μM of β-casein was dissolved in a 5 mM phosphate buffer pH 7. A 3 μL portion of this solution was added in 7 μL of labeling buffer (5 mM phosphate buffer in D_2O pD 7) and then quenched after defined time at 0 °C in 50 mM phosphate quench buffer containing 4 M guanidine chloride adjusted at pH 2.3. The β-casein was then submitted to online immobilized pepsin digestion (Enzymate BEH pepsin column 2.1 mm × 30 mm, Waters) during 3 min using a back-pressure of 10000 psi (100 μL·min^{-1}) of formic acid 0.1% (pH 2). The trapped peptides generated by immobilized pepsin from the β-casein were submitted to UPLC separation at 0 °C (Acquity UPLC BEH C18 1.7 μm 1 mm × 100 mm, Waters) after trapping (Acquity UPLC BEH C18 1.7 μm VanGuard precolumn 2.1 mm × 5 mm, Waters) using the HDX manager of the HDX-MS system from Waters. The peptides were detected using a Synapt G2-Si with T-wave ion mobility separation (see Table S2 for details). Data processing was performed using ProteinLynx Global SERVER v3.0.3 and DynamX HDX data analysis software v3.0.0.139.

Solvent Accessible Surface Analysis (SASA) Computation. SASA were computed for each cleavable amino acid using an in-house Python 3.9 script. The open-source PyMOL library (https://github.com/schrodinger/pymol-open-source, release 2.4.0, open-source license: BSD-like) was used. In brief, the algorithm uses the hard sphere model applied for a solvent molecule using a water radius value of 1.4 Å to probe the surface of a given protein structure (from PDB file) accessible to this solvent. Relying on the PYMOL code, the script was used residue-per-residue to compute the solvent-exposed to the surface for each residue of interest.

RESULTS AND DISCUSSION

Method Formalism. *Choice of Trypsin.* The selected protease for the proof-of-concept of enzymatic reactor coupled online with electrospray ionization and high-resolution mass spectrometry (see Figure 1) is trypsin (from porcine pancreas *Sus scrofa*). This endoprotease belongs to the serine protease family. Natural and modified trypsins are relatively tolerant to thermal denaturation (above 60 °C and much more).[23]

https://doi.org/10.1021/jasms.1c00274
J. Am. Soc. Mass Spectrom. 2022, 33, 284−295

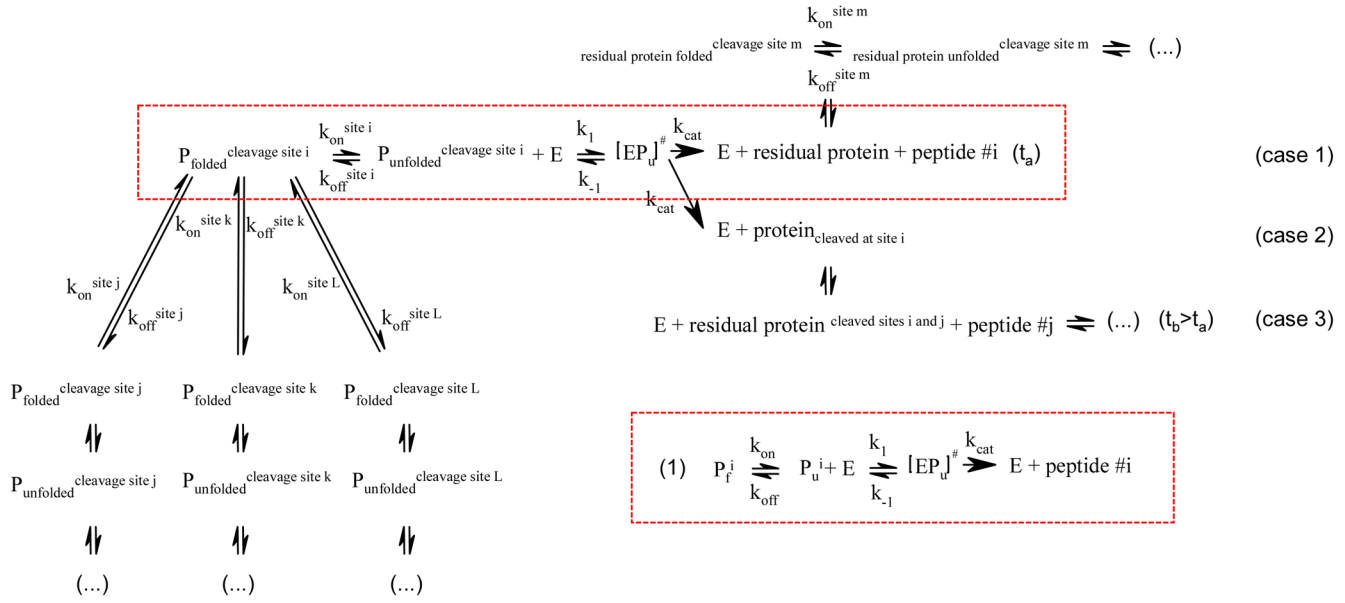


Figure 2. Full kinetic scheme showing all of the competitive reactions of trypsin digestion of the protein under investigation which occur in the enzymatic reactor. The reaction rate equations for the P_f^i to P_u^i protein folded state at the cleavage site *i* to protein unfolded state at the cleavage site *i*, determined by k_{on} and k_{off} rates, enzyme−protein transition state (k_1 and k_{-1} rates), and cleavage of the protein (k_{cat} rate at site *i*) of the reaction inside the red dashed rectangle were determined for one of this kinetic scheme (see eq 1 in Scheme 1). Case 1: Released peptide detected by ESI-MS with experimental time of appearance t_a. Case 2: Not yet released peptide requiring more than one cleavage event before being released. Case 3: Second cleavage event releasing the peptide. This peptide #j appears necessarily later than peptide #i and has an experimental time of appearance t_b with $t_b > t_a$. All of the competitive paths occur at different rates of reaction and with their own time of appearance due to their sequence of reaction.

Porcine pancreas trypsin is characterized by a catalytic triad composed of H_{57}, D_{102}, and S_{195} and cleaved after K and R.[24,25]

Kinetic Digestion Scheme with Time-Dependent Rates. Several competitive reactions occur inside the enzymatic reactor as shown in Figure 2. The kinetic scheme (Scheme 1, eqs 1 and 2) of the protease in the formalism of Michaelis−Menten including the contribution of k_{cat} (enzyme catalytic constant), k_1 (formation of the enzyme−substrate complex) and k_{-1} (dissociation of the enzyme−substrate complex), k_{off} (local refolding of the protein around the cleavage site), k_{on} (local unfolding of the protein around the cleavage site), and a modified Michaelis−Menten constant K'_m (K_m which explicitly includes the other kinetic constants as apparent K_m) are provided below. The kinetic parameters remain constant until the enzyme is saturated, meaning that the protein is in excess compared to the amount of enzyme. In this case, the peptide apparition rate will adopt a linear behavior as a function of time. The complete analysis of the kinetic scheme is provided in Supporting Information. K_m is also provided in terms of k_{-1}, k_{cat}, and k_1 according to the original Michaelis−Menten formalism (Scheme 1, eq 3) to discuss the influence of each constant regarding the contribution of k_{on} and k_{off}. Finally, eq 4 (Scheme 1) is a reminder about the dependence between K_m and the reaction rate.

Scheme 1. Kinetic Scheme

Eq.1 $P_f^i \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} P_u^i + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [EP_u]^{\#} \xrightarrow{k_{cat}} E + \text{peptide \#i}$

Eq.2 $K'_m = \frac{[E][P]}{[EP_u]^{\#}} = \frac{k_{off}(k_{cat} + k_{-1})}{k_{on}k_1}$

Eq.3 $K_m = \frac{k_{-1}+k_{cat}}{k_1}$

Eq.4 $v \propto 1/Km$

The affinity of trypsin (k_1/k_{-1}) for a specific sequence of the investigated protein will affect K_m and, consequently, also affects the catalytic efficiency k_{cat}/K_m. On the contrary, k_{-1} needs to be at least 1 order of magnitude larger than k_1 to significantly affect k_{cat}/K_m. A numerical simulation (see Figure S2) obtained from eqs 1−4 (Scheme 1) showed, as expected, that the reaction rate is optimal when k_{off}/k_{on} is low (i.e., mainly locally unfolded state) and k_{-1}/k_1 is low (i.e., enzyme is saturated). On the other hand, the reaction rate heavily decreases when both k_{off}/k_{on} and k_{-1}/k_1 are high (i.e., mainly locally refolded state in combination to a nonsaturated enzyme). The apparent catalytic efficiency k_{cat}/K'_M will therefore be affected depending on the respective values of k_{off} and k_{on} with regard to k_{-1} and k_1 values. As the targeted protein was always in large excess compared to trypsin (i.e., trypsin is saturated), the experimentally determined peptide apparition rates are mainly dependent on k_{off}/k_{on} (unfolding state around the cleavage site). k_{off}/k_{on} will be affected by the secondary structures around the cleavage site such as α-helices and β-sheets and the presence of proline near the cleavage site. The apparition rates will be also dependent on the trypsin affinity to the amino acid sequence of the peptide to be cleaved. The numerical simulation (see Figure S2) suggests that k_{off}/k_{on} more heavily affects the apparition rate than the affinity constant of trypsin for the cleavage site (K_m which depends on k_{cat}, k_1 and k_{-1}, see eq 3, Scheme 1). Because

https://doi.org/10.1021/jasms.1c00274
J. Am. Soc. Mass Spectrom. 2022, 33, 284−295

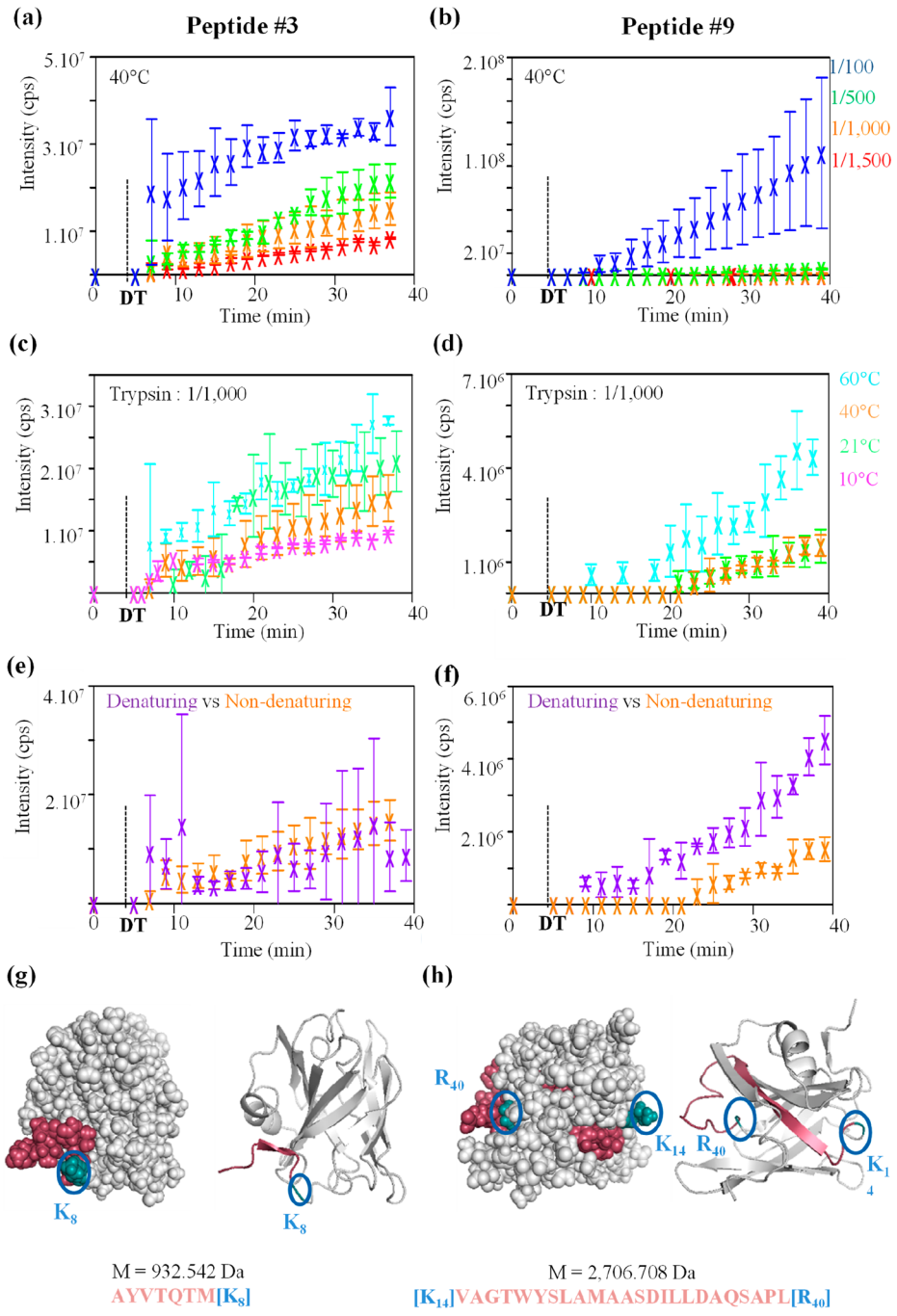


Figure 3. Intensity as a function of time of a β-lactoglobulin peptide accessible to digestion (Peptide #3, M = 932.542 Da) and a peptide partially accessible to digestion (Peptide #9, M = 2,706.708 Da) for (a,b) protease-to-protein ratios comparison at 40 °C (c,d) temperature of digestion comparison at protease-to-protein ratio of 1/1000, and (e,f) denaturing versus nondenaturing conditions comparison. (g,h) β-Lactoglobulin 3D representation (PDB: 1CJ5). Peptides #3 and #9 are colored in red and their cleavage sites surrounded by blue circles. Dashed lines represent the dead volume of the system (reactor and capillary). Each experiment was performed in triplicate, and error bars correspond to the standard deviation.

trypsin performs only 1 cleavage event at the time, the peptides requiring a single cleavage (those in N_{ter} or C_{ter}) event to be released will display lower times of appearance than those requiring several cleavage events. On the contrary, when the cleavage sites are buried in the 3D structure, the peptide will be characterized by a later time of appearance because trypsin needs to "peel" the protein before accessing the core cleavage sites. Moreover k_{off}/k_{on} will also be affected as more steric hindrance around the cleavage site leads to a lower probability to provide a locally unfolded state of the cleavage site.

Time of Appearance and Apparition Rate Determination as Tools for Structural Characterization. Peptide apparition rate slopes expressed in cps·min^{-1} were obtained by calculation of the initial velocity. Apparition rates were obtained by estimating the respective slope tangent curves at the time of appearance of each peptide (see Figure S4). The time of

https://doi.org/10.1021/jasms.1c00274
J. Am. Soc. Mass Spectrom. 2022, 33, 284−295

appearance is the difference between the dead time of the system and the detection time of the peptide (i.e., above the limit of detection of the experimental setup). The response factor of ionization was described as dependent on the hydrophobicity of the chain of the peptide as well as the content in basic residues.[26] In this work, this factor was assumed to be approximately equivalent for all investigated peptides. This was however verified based on the basicity (content of basic residues) and the hydropathic pattern (content of hydrophobic residues). In addition, the linear behavior of the signal as a function of time suggests the absence of a significant suppression effect. Nonetheless, a more in-depth characterization of the impact of these effects on the response factor should be performed on "outsider" peptides such as very acidic, basic, or hydrophobic peptides, which are not the most common cases. On the basis of the observations, recommendations about specific corrections could be considered based on the potentially troublesome peptide sequences.

Method Optimization: β-Lactoglobulin Analysis. Bovine β-lactoglobulin (Uniprot P02754, 162 residues without signal peptide) was selected as a model protein for the reactor optimization due to its moderately low molecular weight (18 kDa) and its peculiar structure containing both flexible and structuring domains such as α-helices (1) and β-sheets (9).[20] It also contains two disulfide bridges (C_{66}–C_{160} and C_{106}–C_{119} or C_{106}–C_{121}), and the thermal denaturation process of this protein is also well characterized in literature.[27] The sequence of β-lactoglobulin predicts 16 potential cleavable sites sensitive to trypsin which are spread all along the protein sequence (see Figure S3). Ten to 16 cleavage sites are located within the core of the protein, providing a perfect model to test the accessibility of these sites throughout the kinetics of digestion. The thermal denaturation of β-lactoglobulin can be described as a multistage process.[27] Between 40 and 55 °C, some side chains of the protein undergo minor conformational changes. Above 60 °C, the protein undergoes a reversible partial unfolding (molten globule state) up to 140 °C, the temperature at which the protein is fully thermally denatured (complete destruction of the secondary structures).

Online Digestion Optimization. The digestion optimization of the online enzymatic reactor was performed by investigating the protein-to-enzyme ratios and the temperature of the digestion. Other parameters were kept as trypsin optimum condition of operation (pH and MS-friendly buffers) and nondenaturing buffer for the investigated protein models.

In Figure 3, two peptides were selected based on the in-silico digestion prediction and the PDB structure of β-lactoglobulin to the influence of the protease-to-protein ratio. Peptide #3 is accessible to digestion (well exposed to the surface, sequence LIVTQTMK/K_8, Figure 3g) while the peptide #9 (sequence K_{14}/VAGTWYSLAMAASDISLLDAQSAPLR.R_{40}, Figure 3h) is partially accessible to digestion. As shown in Figure 3, these two different peptide traces (absolute intensities of the target m/z over the time of the experiments, see Figure 3a,b) increased when the trypsin-to-β-lactoglobulin ratio was raised from 1/1500 to 1/100 at constant temperature. For example, the partially accessible peptide #9 appeared earlier (3 times the dead time of the capillaries) when this ratio was raised from 1/1500 to 1/100 (Figure 3b). As expected, the enzymatic digestion was more effective when higher trypsin-to-β-lactoglobulin ratios were used. Indeed, more peptides were generated after 40 min upon an increase of the enzyme–protein ratio: 11 peptides for 1/1500, 14 peptides for 1/1000, 18 peptides for 1/500 and 24 peptides for 1/100. This increased trypsin concentration also enhanced the apparition rate of the released peptides (the amount of detected peptide during the first minutes of the experiment, expressed in cps·min^{-1}). In some cases, this also reduced the time of appearance of these peptides (the smallest time when the peptide appears in the mass spectrum, expressed in minute). Indeed, while K_{60}, K_{91}, K_{100}, K_{101}, and K_{135} were not detected at low ratios (1/1500 and 1/1000) after 40 min of digestion, these cleavage sites became accessible at higher ratios (1/500 and 1/100). Two distinct cleavage sites (K_{47} and K_{69}) appeared to be constantly not accessible to trypsin. As reported in the literature, K_{47} is followed by a P, preventing trypsin to efficiently cleave this peptide bond.[24] On the other hand, K_{69} is located on a peptide section constrained by a disulfide bridge. For further optimization, the protease-to-protein ratio of 1/1000 was selected as it provides the most differential apparition rates and times of appearance under 40 min of enzymatic reaction by providing access to the first step of the digestion.

The effect of temperature on the digestion efficiency was investigated between 10 and 60 °C at a constant protease-to-protein ratio of 1/1000 (Figure 3c,d). This range of temperature was investigated in regard to the thermal denaturation process of β-lactoglobulin.[27] Here, as well, peptides #3 and #9 have been monitored to investigate the temperature effect on digestion kinetics. As expected, both intensities and apparition rates were lower for the peptide #9 than for peptide #3, see Supporting Information Table S3. Interestingly, the times of appearance or the peptide intensities of these two peptides was not influenced by the increase of the temperature from 21 to 40 °C. At 60 °C, the times of appearance and intensities of peptide #3 but especially peptide #9 were increased. This effect was attributed to the thermal denaturation process of β-lactoglobulin, which is notably above 40 °C.[27] Peptide #9 presents an accessible cleavage site and a partially accessible cleavage site. Moreover, despite the second cleavage event, peptide #9 remains embedded within the cleaved β-lactoglobulin after the action of trypsin (see Figure 3h). As a result, the release of this peptide requires some degree of flexibility of the substrate. This is in excellent agreement with the reported thermal denaturation data by Tolkach and Kulozik.[27] In the selected experimental conditions, an increase of temperature was also directly related to the higher number of detected unique peptides: 5 peptides for 10 °C, 7 for 21 °C, 14 for 40 °C (as mentioned before) and 15 unique peptides detected for 60 °C. In the studied temperature range, peptides released from the cleavage sites K_{47}, K_{60}, K_{69}, K_{100}, and K_{135} were never detected. Again, K_{47} is followed by a proline and K_{69} is located on a peptide section constrained by a disulfide bridge, while K_{60}, K_{100}, and K_{135} are located in the protein core, leading to sterically hindered cleavage sites for trypsin digestion. Peptides cleaved from K_{77}-R_{124} appeared at 40 °C and K_{75}-K_{101} at 60 °C. These differences can be explained by the moderate thermal denaturation of β-lactoglobulin in this temperature range. In addition to the thermal effect on β-lactoglobulin denaturation, the trypsin activity could also be affected by temperature. Although its activity is approximately constant between 40 and 60 °C,[23] trypsin was reported to be three times less active at 10 °C. The surface accessibility of the trypsin cleavage sites as well as the thermal effect on a protein being investigated is then achievable from 40 to 60 °C, without introducing significant experimental bias due to drastic trypsin activity variation

https://doi.org/10.1021/jasms.1c00274
J. Am. Soc. Mass Spectrom. 2022, 33, 284−295

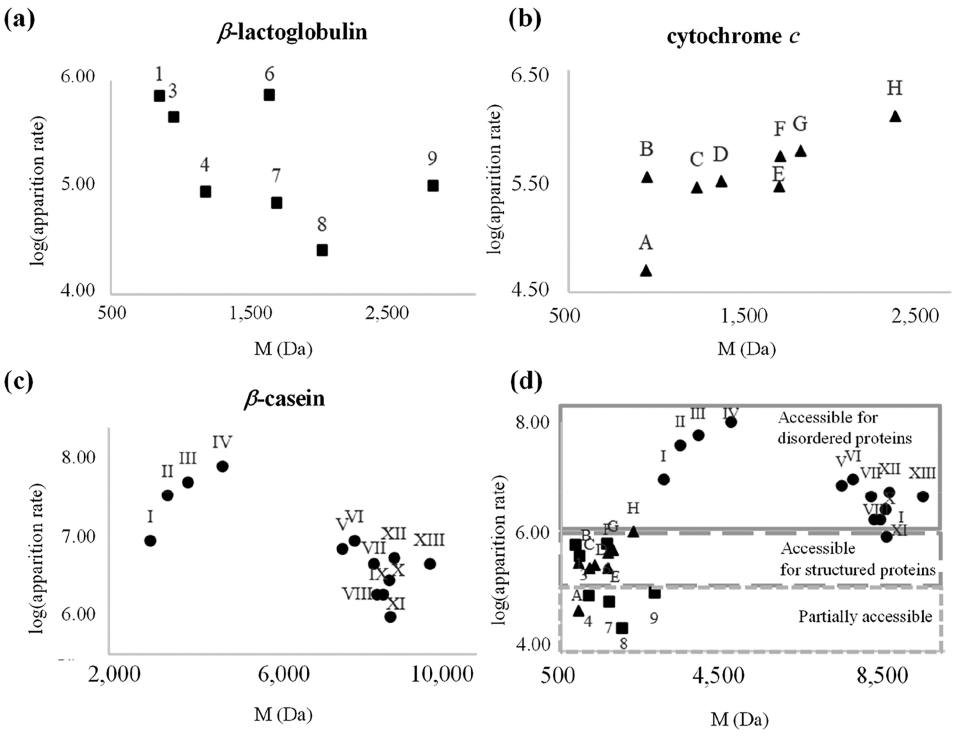


Figure 4. Graph plotting $\log_{10}$ (apparition rate) versus masses for peptides generated by trypsin at 40 °C with the online enzymatic reactor coupled to ESI-MS for (a) β-lactoglobulin doubly positively charged peptides (square ■) in nondenaturing conditions, (b) cytochrome *c* doubly positively charged peptides (triangle ▲) in nondenaturing conditions, and (c) β-casein peptides in nondenaturing conditions, all charge states considered (circle ●). (d) Combined graph plotting $\log_{10}$ (apparition rate) versus masses for peptides generated for the three proteins.

In general, the apparition rates of the peptides increased with the temperature of digestion. However, some peptides (e.g., peptide #7, and #8, Figure S5) were detected only if the temperature of digestion was above 40 °C. This different behavior was attributed to the presence of disulfide bridges close to the peptides #5 and peptides #7 and #8 where the dynamics of the peptide chains cannot be significantly affected by temperature due to the lack of flexibility of the structure. However, peptide #5, which is also located near a cysteine residue involved in a disulfide bond, was barely detected for each investigated temperature of digestion. One can expect that the presence of this secondary structure capped trypsin efficiency to cleave at this potential additional cleavage site when lower temperature of digestion is used.

The optimized digestion parameters were defined at a 1/1000 enzyme-to-protein ratio and at a temperature of 40 °C to investigate the kinetics of peptide released between denaturing (reduction and alkylation of disulfide bonds) and nondenaturing conditions of the investigated protein. The denatured β-lactoglobulin presents an altered tertiary structure, while the three-dimensional structure is theoretically preserved in the nondenaturing 50 mM NH_4HCO_3 buffer. During the digestion reaction, 26 peptides were generated under denaturing conditions while only 14 peptides were produced under nondenaturing conditions (due to the presence of the disulfide bridges), where 11 were common to both conditions. As an example, the peptide #3 exhibiting an accessible cleavage site and the peptide #9 having a partially accessible cleavage sites were compared in denaturing and nondenaturing conditions. As shown in Figure 3e,f, the peptide #3 (N_{ter} loop cleaved at K_8) presented similar signal over time profile whereas peptide #9 (cleaved at K_{14} and R_{40}) displayed a 3 times dead time shifted time of appearance from denaturing to nondenaturing conditions. This shifted time of appearance is assumed to be related to the required additional digestion steps (i.e., steric hindrance around the cleavage site needs to be removed) before the target peptide can be released. This is in total agreement with the kinetics scheme and equations described previously.

To assess the contribution of k_{off}/k_{on} in regard to the exposition of the cleavage site to the solvent, the experimental reaction rates were compared (see Figure 4). The computed solvent-accessible surface area (SASA)[28,29] values were determined from a PDB structure of β-lactoglobulin using our in-house Python3 script and the open source Pymol library (see the Material and Methods for details). For β-lactoglobulin, PDB 1CJ5[20] obtained by solution NMR was used as reference (note that the residue sequence of 1CJ5 is corresponding to an isoform of P02754 with no modification on the cleavage sites of trypsin). β-Lactoglobulin 3D representation in Figure 3g clearly shows that peptide #3 (N_{ter}-K_8 in blue) presenting only one cleavage site seems to be effectively accessible to trypsin since located at the protein surface. The SASA computed value for K_8 cleavage site was 198 Å^2 (total SASA of β-lactoglobulin is about 8600 Å^2), confirming the accessibility of the cleavage site to the enzyme. The high value of apparition rate (i.e., $4.57 \times 10^{+5}$ cps·s^{-1}) and the short time of appearance (i.e., 1 min 20 s) is in good agreement with these computed SASA values, pointing out a link between the surface location of the cleavage site and the accessibility for the trypsin digestion. K_8 is part of a loop, which is an unstructured region of the β-lactoglobulin that connects two consecutive secondary structures.[30] Due to their lack of defined pattern, loops are more flexible than α-helices[31]

https://doi.org/10.1021/jasms.1c00274
J. Am. Soc. Mass Spectrom. 2022, 33, 284−295

or β-sheets and are more likely to be accessible to digestion due to their dynamics (i.e., lower k_{off}/k_{on} expected). In contrast, peptide #2 (K_{75}-K_{83}) cleavage sites are both located on β-sheet secondary structures (i.e., larger k_{off}/k_{on} ratio expected) but their SASA values are in the same range as peptides released with the highest apparition rates. Nonetheless, the apparition rate of peptide #2 was weak (see Table S3). A good correlation between the time of appearance of the cleaved peptides and the solvent-accessible surface analysis was also observed. Peptide #9 (K_{14}-R_{40}) had an experimental apparition rate ($1.06 \times 10^{+5}$ cps.s^{-1}) above the median (i.e., $9.25 \times 10^{+4}$ cps·s^{-1}), while this peptide presented one of the highest times of appearance (detected after 15 min, see Table S3). This is also in good agreement with the SASA calculation (i.e, 129 and 24 Å^2 respectively). For this peptide, the cleavage sites are both located on flexible loops (K_{14} and R_{40}, assumed low k_{off}/k_{on} ratio). In regard to the primary structure, peptide #9 requires only a couple of cleavage events to be released with an expected time of appearance similar to peptide #4 (i.e., 1 min 55 s or close value). Nonetheless, peptide #9 (orange) is embedded inside the 3D structure of the β-lactoglobulin as shown in Figure S6 and as mentioned above. The cleavage of peptide #3 or #6 is a prerequisite before the release of peptide #9 based on the structural arrangement of the protein. A similar result was obtained for peptide #5 (the highest time of appearance as 19 min 30 s) which also required the prior cleavage of peptide #3 or #6 to be released. Peptide #3 is obtained from direct cleavage by trypsin.

The apparition rates of the peptides are well correlated to the presence of secondary structure near the cleavage site, affecting the k_{off}/k_{on} ratio as supported by the kinetic scheme provided in eqs 1 and 2 (Scheme 1). SASA values had some correlation with the apparition rates but actually reflect both the accessibility to the cleavage site of the protease and an indicator of the local steric hindrance around the cleavage site which also affects k_{off}/k_{on} near the cleavage sites. Finally, the times of appearance are well supported by the amount of cleavage events required to release the peptides. Embedded peptides in the tridimensional structure of the entire protein also need more digestion time before being released.

The times of appearance and the rates of apparition appeared to be constant (linear) during the experiments (see Table S3). This suggests that the conformation of the residual core structure(s) of the native β-lactoglobulin in solution was not drastically affected by the digestion. Otherwise, the access of the most buried cleavage sites to trypsin would be significantly facilitated and the observed times of appearance would be reduced while the rates of apparition would be increased. Moreover, detected residual structures seems to reach a stable intensity over time. A progressive loss of signal would be anticipated in the case of major structural change of the residual structures exposing the cleavage site to the trypsin located in the core of the protein. Note that it requires the residual structure to be efficiently detected by MS during the online digestion. A time-of-flight mass analyzer could be more suited for this purpose.

Cytochrome *c* Analysis. The optimized trypsin digestion (temperature of 40 °C using a protease-to-protein ratio of 1/1000) was also applied on another structured protein of low molecular weight, the heart equine cytochrome *c* (12 kDa, 104 amino acids). The structure of this protein in solution is characterized (PDB: 1AKK[21]). This second model was selected due to the presence of both flexible and structured regions. Indeed, cytochrome *c* contains 5 α-helices, representing approximately 39% of the protein structure, leaving the remaining 61% as poorly structured regions (loops). Cytochrome *c* also contains a heme covalently linked to the C_{14} and C_{17} (the numbering is excluding the removed first initiation methionine residue) of the protein and noncovalently bound to the heme iron with H_{18} and M_{80}. In addition to its role of electron transfer, the heme is known to have a role in the structural features of cytochrome *c*.[32] In this protein, all tryptic cleavable sites are located on the surface of the protein, assuming a total access of the potential cleavage sites to trypsin (see Figure S7). Consequently, identical or very similar times of appearance of the released peptides were expected. On the other hand, similar but larger apparition rates were expected due to a limited dependence of k_{off}/k_{on} as the protein mainly contains loop regions, which causes notable differences compared to α-helix and/or β-sheet regions.

During a 40 min experiment, the tryptic digest of cytochrome *c* led to the generation of 15 peptides. As seen in Figure 4b–d, the apparition rates of the released peptides of cytochrome *c* (doubly charged species) are located in the "accessible to digestion" region of the chart. They are of the same magnitude order than those observed for β-lactoglobulin. The difference in the apparition rates is indeed well explained by the presence of secondary structure elements located near the cleavage site. Among the measured apparition rates, the lowest value could be attributed to the peptide #A (M = 901.466 Da, see Tables S4 and S5) characterized by a single cleavage site (K_8) localized in the midst of an α-helix (which was considered as thermodynamically unfavorable according to Fontana and co-workers[19]). Most of the other peptides which shared apparition rates above $0.5 \times 10^{+5}$ cps.min^{-1} (i.e., above 4.5 in $\log_{10}$ scale) are characterized by trypsin cleavage sites located in loops, with the notable exception of peptides #E and #F. However, the cleavage sites for these peptides are positioned at the very edge of the α-helix which seems to poorly affect the digestion rate of trypsin when the digestion was performed at 40 °C. One could expect that k_{off}/k_{on} ratio is low enough to significantly boost the rate of reaction.

Surprisingly, the time of appearance of the peptide releases by trypsin, especially for the N_{ter} and C_{ter} peptides, appeared to be noticeably delayed compared to those observed for the β-lactoglobulin. This phenomenon was attributed to the structuring contribution of the heme due to the covalent and/or noncovalent binding interactions with the residue chain of the cytochrome *c*, affecting the probability of trypsin to perform a cleavage event disregarding the SASA values of the cleavage site computed from the PBD structure as predicted by Fontana and co-workers.[19] In the case of the peptides #B and #H (see Table S5), the noncovalent binding of these peptides both containing M_{80} and interacting with the heme affects the time of appearance (6 min 30 s and 4 min 40 s for peptides #B (M_{80}-K_{87}) and peptide #H (M_{80}-C_{ter}), respectively). Nonetheless, the chronological release sequence of the trypsin generated peptide (see Table S5) suggests that cytochrome *c* is more efficiently digested from its C_{ter} extremity than the N_{ter} extremity. This observation reflects the more pronounced presence of α-helices in the N_{ter} region than the C_{ter} region.

β-Casein Analysis. Finally, the developed approach was applied to a third protein: bovine β-casein. This protein, characterized by a mass of approximately 25 kDa (224 amino acids), does not present 3D structure available in the protein data bank, making it a good candidate to our structural

https://doi.org/10.1021/jasms.1c00274
J. Am. Soc. Mass Spectrom. 2022, 33, 284−295

Table 1. Effect of the Temperature on the Apparition Rate Determined with the Enzymatic Reactor with On-Line ESI-MS Detection of the Released Peptides by Trypsin of Bovine β-Casein*

	21°C → 40°C → 60°C apparition rate acceleration factor		Presence of Proline near the cleavage site [a]	Assumed secondary structures around the cleavage sites
Peptide #I	x1.5	x1.5	N_{ter}	loop / C_{ter}
Peptide #II	x5	x10		N_{ter} / loop
Peptide #III	**x8**	x12	N_{ter}	loop / C_{ter}
Peptide #IV	x2.5	x5	N_{ter}	loop / C_{ter}
Peptide #V	**x15**	x15	N_{ter} + C_{ter}	loop / loop
Peptide #VI	**x10**	x20	N_{ter} + C_{ter}	loop / loop
Peptide #VII	x5	x10		loop / loop
Peptide #VIII	x3	x3	N_{ter} + C_{ter}	loop / loop
Peptide #IX	x3	x5	C_{ter}	loop / loop
Peptide #X	**x5**	x10	N_{ter}+C_{ter}	loop / loop
Peptide #XI	x1.5	x4.5		loop / loop
Peptide #XII	**x7**	x6	C_{ter}	loop / loop
Peptide #XIII	**x6**	x7.5	C_{ter}	loop / loop

*Peptides sharing the same background color are related in terms of sequence and are due to missed cleavage events. 21°C was selected as reference temperature. The acceleration factors were determined based on the ratio between apparition rate at the temperature of interest and the apparition rate at the reference temperature. [a]Proline residue(s) is/are located within a three-residue range near the cleavage site of trypsin.

elucidation method based on our online enzymatic reactor with ESI MS. The secondary structure of this protein is described to contain approximately 31% of β-sheets, 28% of turns, and 21% combining α-helices and loops. The remaining 20% are unspecified as reported by Fourier transform infrared and circular dichroism spectroscopy.[33] In addition, the presence of many proline residues (17%) introduces turns into the structure and interrupts the secondary structure domains, resulting in a supposed open tertiary structure. Moreover, the lack of disulfide bridges combined to the previous observations support a "molten globule state" structural behavior for β-casein.[34]

In the case of β-casein, the apparition rates are 1 to 2 orders of magnitude higher than those obtained from β-lactoglobulin and cytochrome *c*, while the times of appearance for these peptides are significantly shorter. A large apparition rate means that the k_{off}/k_{on} ratios for all the detected peptides are low, assuming low conformational barriers between conformers. In other words, the structure around the cleavage site is highly flexible and poorly structured. The difference in terms of apparition rates could be attributed to the presence of proline near the cleavage site which negatively affects k_{off}/k_{on} and, by consequence, the reaction rate of the digestion by trypsin (see Figure 4). This is also supported by the digestion experiments performed at different temperatures as reported in Table 1. We assumed that the apparition rate acceleration factors of the peptides of poorly structured region were almost not affected by the temperature of digestion while the most structured region (i.e., α-helix and β-sheet) would experience higher acceleration factor of the apparition rates.

The most intense acceleration factor (in bold) of the apparition rate for temperature of digestion from 21 to 40 °C seems to be related to the presence of proline residue(s) near the cleavage site. The following acceleration factor are in each case very limited, suggesting that all the remaining regions are lacking secondary structures, in good agreement with the CD[33] and NMR data,[22] the predicted structure by computational chemistry,[35] and our reported HDX-MS experiment shown in Figure 5. The slight delay in the apparition rates obtained from our enzymatic reactor is attributed to the number of cleavage events required to release the detected peptide in regards of the k_{cat}/K_m of trypsin.

The data obtained from our enzymatic reactor online with ESI-MS were favorably compared to the hydrogen−deuterium exchange mass spectrometry (HDX-MS) experiments. The highest apparition rates (peptides II−IV) observed from the enzymatic reactor correspond to the C_{ter} region of β-casein (only 1 cleavage event), which also displayed the highest deuterium uptake, as reported in Figure 5, Table 1 and Table S6. The lowest apparition rates (peptides VIII, IX, and XI) were correlated to the lowest deuterium uptake regions, with notable exception for peptides VIII and IX (very similar peptide sequence and position). In these peptides, the C_{ter} is located in very high deuterium uptake while N_{ter} is located in a high deuterium uptake region. However, a proline is present very close to the N_{ter} cleavage site of trypsin. In other words, the enzymatic reactor online coupled to ESI-MS was able to probe a more rigid region (higher k_{off}/k_{on} ratio around the cleavage site) than the HDX-MS suggested, especially when monitoring several temperatures of digestion. Similar results

https://doi.org/10.1021/jasms.1c00274
J. Am. Soc. Mass Spectrom. 2022, 33, 284−295

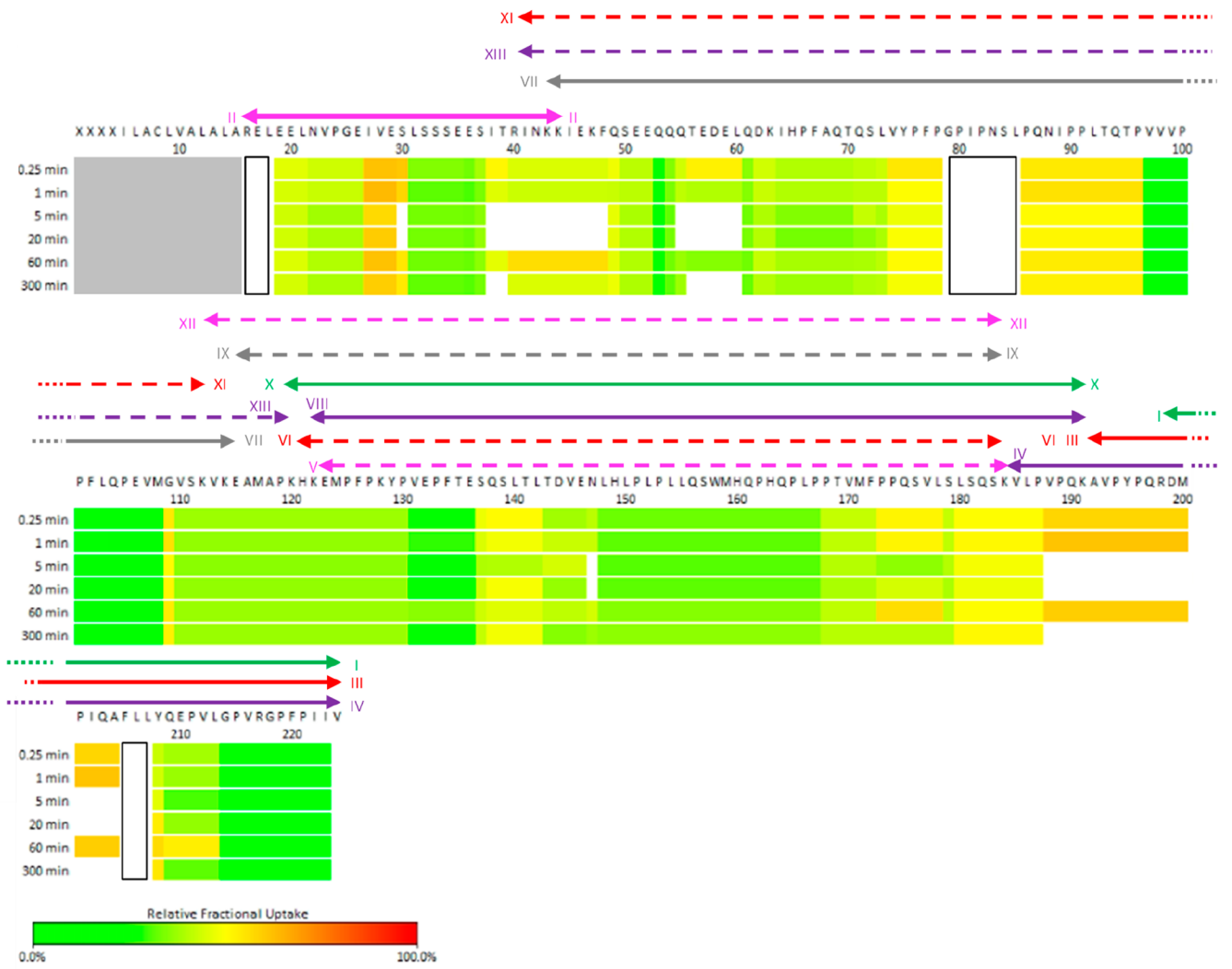


Figure 5. Deuterium uptake mapping for β-casein after 15 s, 1 min, 5 min, 20 min, 60 min, and 300 min of hydrogen/deuterium exchange. Peptides generated during online digestion with trypsin (1/1000 40 °C) are represented by a double arrow. The gray rectangle represents the signal peptide, and blank rectangles correspond to missing HDX data.

could be observed for the intermediate apparition rates where the presence of proline more heavily affects the apparition rate than suggested by the deuterium uptake obtained from the HDX-MS experiments. We attributed this observation to the real time measurement of the enzymatic reactor, while the HDX-MS experiments are performed in sequential batches at different times where H/D scrambling and back-exchange will prevent the observation of this local rigidification of the β-casein structure due to proline.

CONCLUSIONS

In this work, we introduced an easy label-free method using an enzymatic reactor coupled online with high-resolution mass spectrometry to extract structural information and dynamics of proteins around the cleavage sites of the selected protease(s). The workflow was designed to be easily implemented and not challenging in terms of MS setup or too exigent in terms of MS instrumentation. Time-of-flight instruments commercialized in this past decade can be used as well as Orbitrap or FT-ICR fitted with an electrospray ion source. According to the average size of the generated peptides during the online digestion, collision-induced dissociation, electron capture/transfer dissociation, or exact mass determination with isotopic fine structure strategies could be envisioned for identification. The level of information in terms of 3D structure lies between HDX-MS and CD spectroscopy. Our methodology represents a starting point before the implementation of more advanced experimental MS-based methods such as HDX-MS or XL-MS. On top of that, there is virtually no upper limit to the size of the entire protein(s) being investigated. Nonetheless, the apparition rates of peptides can provide a good estimation of the local amplitude of k_{off}/k_{on} around the cleavage site exposed to the solvent (and the enzyme).

The most important asset is the wise choice of the proteolytic enzyme, which mainly relies and depends on the k_{off}/k_{on} ratio as well as $((k_{-1} + k_{cat})/k_1)$ in regard to the protein being probed in terms of structure and kinetics. Trypsin was successfully applied for the three investigated proteins during this study as its catalytic activity is relatively slow compared to other proteases (e.g., chymotrypsin, pepsin) according to the BRENDA database (https://www.brenda-enzymes.org/). The choice of other proteases will mainly depend on the protein(s) (e.g., single protein, multimers, ligand−receptor systems, supramolecular assemblies) to be investigated and the purpose

https://doi.org/10.1021/jasms.1c00274
J. Am. Soc. Mass Spectrom. 2022, 33, 284−295

of the experiments. In this context, the association of several enzymes is possible but each individual enzyme needs to be characterized separately prior to the experiment using the reactor. This option could be explored for enzymes with low probability of cleavage events (e.g., not enough cleavage sites in the protein sequence or inadequate k_{off}/k_{on} ratios). In addition, intrinsically disordered protein investigation is also possible.

In this study, we mainly focused on the times of appearance and apparition rates of the peptides generated after enzymatic digestion, which are impacted by the presence of secondary structures and/or proline. Nonetheless additional MS information is accessible in these data: the m/z ratios (exact mass determination and/or MS/MS) allow the identification of the detected peptides. The intensities of these peptides will provide a fair approximation for the mass balance control and the relative amount of the residual structure (i.e., the substrate before the enzymatic digestion) in regard to the substrate after the proteolysis. This information coupled to the potentially released peptides provides information about the sequential order of the enzymatic proteolysis (i.e., the order of the concurrent cleavage events which can be represented as a "digestion tree" in a global kinetic scheme) generating the detected peptides during the digestion. For larger proteins, the theoretical digestion tree should be established and defined if concurrent pathways are able to generate the same peptides. It is worth to mention that the apparition rates of the peptides could be normalized. This requires the knowledge of the response factor not only of the peptide alone but in competition with other. The latter is minimized for the first digested peptides. Using isotopically labeled standards would allow to estimate the response factor evolution during the online digestion if the primary structure of the proteins under investigation is available.

Further work is currently under investigation to elucidate the resilience of the residual structure after the ablation of one or several peptides and their respective roles in the preservation of the protein structures at the residues to domain levels.

■ ASSOCIATED CONTENT

Supporting Information

Tables S1−S2 Instruments default setup, Choice of the enzyme, Demonstration of the kinetic scheme, Figure S1 Enzymatic reactor, Figure S2 Numerical simulation, Figure S3 Sequence of β-lactoglobulin, Figure S4 Intensity as a function of time of β-lactoglobulin peptides, Table S3 β-lactoglobulin generated peptides, Figure S5 Influence of the temperature, Figure S6 β-lactoglobulin 3D representation, Table S4 Influence of the temperature, Figure S7 Cytochrome c 3D representation, Table S5 Cytochrome c generated peptides, Table S6 β-casein generated peptides. The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jasms.1c00274.

Instrument setups (Tables S1 and S2), enzyme choice, demonstration of the kinetic scheme, Figures S1−S7, and Tables S3−S6 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Johann Far − *Mass Spectrometry Laboratory, MolSys Research Unit, Quartier Agora, University of Liège, B-4000 Liège, Belgium;* orcid.org/0000-0003-1208-6262; Email: johann.far@uliege.be

Authors

Elodie Grifnée − *Mass Spectrometry Laboratory, MolSys Research Unit, Quartier Agora, University of Liège, B-4000 Liège, Belgium;* orcid.org/0000-0001-8005-4401

Christopher Kune − *Mass Spectrometry Laboratory, MolSys Research Unit, Quartier Agora, University of Liège, B-4000 Liège, Belgium;* orcid.org/0000-0002-3010-8173

Cédric Delvaux − *Mass Spectrometry Laboratory, MolSys Research Unit, Quartier Agora, University of Liège, B-4000 Liège, Belgium;* orcid.org/0000-0002-9575-0554

Loïc Quinton − *Mass Spectrometry Laboratory, MolSys Research Unit, Quartier Agora, University of Liège, B-4000 Liège, Belgium;* orcid.org/0000-0001-8153-9590

Gabriel Mazzucchelli − *Mass Spectrometry Laboratory, MolSys Research Unit, Quartier Agora, University of Liège, B-4000 Liège, Belgium*

Edwin De Pauw − *Mass Spectrometry Laboratory, MolSys Research Unit, Quartier Agora, University of Liège, B-4000 Liège, Belgium;* orcid.org/0000-0003-3475-1315

Complete contact information is available at: https://pubs.acs.org/10.1021/jasms.1c00274

Author Contributions

†J.F, G.M., and E.D.P. contributed equally to this paper.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge the support of FEDER BIOMED HUB Technology support and FNRS, GlycoCell project WALInov 2018 program. E.G. thanks L. Trzpiot, N. Rosière, and N. Tanteliarisoa Haingo for their support in sample preparations and M. Boulanger for the helpful reviews.

■ ABBREVIATIONS

ACN, acetonitrile; DTT, dithiothreitol; DT, dead time; ESI−MS, electrospray ionization−mass spectrometry; FA, formic acid; HRMS, high-resolution mass spectrometry; IAA, iodoacetamide; SASA, solvent-accessible surface area; TOF, time-of-flight

■ REFERENCES

(1) Han, X.; Aslanian, A.; Yates, J. R. Mass Spectrometry for Proteomics. *Curr. Opin. Chem. Biol.* **2008**, *12* (5), 483−490.

(2) Vitorino, R.; Guedes, S.; Trindade, F.; Correia, I.; Moura, G.; Carvalho, P.; Santos, M. A. S.; Amado, F. *De Novo* Sequencing of Proteins by Mass Spectrometry. *Expert Rev. Proteomics* **2020**, *17* (7−8), 595−607.

(3) Pan, S.; Aebersold, R.; Chen, R.; Rush, J.; Goodlett, D. R.; McIntosh, M. W.; Zhang, J.; Brentnall, T. A. Mass Spectrometry Based Targeted Protein Quantification: Methods and Applications. *J. Proteome Res.* **2009**, *8* (2), 787−797.

(4) Schaffer, L. V.; Millikin, R. J.; Miller, R. M.; Anderson, L. C.; Fellers, R. T.; Ge, Y.; Kelleher, N. L.; LeDuc, R. D.; Liu, X.; Payne, S. H.; Sun, L.; Thomas, P. M.; Tucholski, T.; Wang, Z.; Wu, S.; Wu, Z.; Yu, D.; Shortreed, M. R.; Smith, L. M. Identification and Quantification of Proteoforms by Mass Spectrometry. *Proteomics* **2019**, *19* (10), 1800361.

(5) Bantscheff, M.; Schirle, M.; Sweetman, G.; Rick, J.; Kuster, B. Quantitative Mass Spectrometry in Proteomics: A Critical Review. *Anal. Bioanal. Chem.* **2007**, *389* (4), 1017−1031.

(6) Wales, T. E.; Engen, J. R. Hydrogen Exchange Mass Spectrometry for the Analysis of Protein Dynamics. *Mass Spectrom. Rev.* **2006**, *25* (1), 158−170.
(7) Sinz, A. Chemical Cross-Linking and Mass Spectrometry to Map Three-Dimensional Protein Structures and Protein−Protein Interactions. *Mass Spectrom. Rev.* **2006**, *25* (4), 663−682.
(8) Lanucara, F.; Holman, S. W.; Gray, C. J.; Eyers, C. E. The Power of Ion Mobility-Mass Spectrometry for Structural Characterization and the Study of Conformational Dynamics. *Nat. Chem.* **2014**, *6* (4), 281−294.
(9) Sinz, A. Cross-Linking/Mass Spectrometry for Studying Protein Structures and Protein−Protein Interactions: Where Are We Now and Where Should We Go from Here? *Angew. Chem., Int. Ed.* **2018**, *57* (22), 6390−6396.
(10) Österlund, N.; Moons, R.; Ilag, L. L.; Sobott, F.; Gräslund, A. Native Ion Mobility-Mass Spectrometry Reveals the Formation of β-Barrel Shaped Amyloid-β Hexamers in a Membrane-Mimicking Environment. *J. Am. Chem. Soc.* **2019**, *141* (26), 10440−10450.
(11) Bleiholder, C.; Dupuis, N. F.; Wyttenbach, T.; Bowers, M. T. Ion Mobility−Mass Spectrometry Reveals a Conformational Conversion from Random Assembly to β-Sheet in Amyloid Fibril Formation. *Nat. Chem.* **2011**, *3* (2), 172−177.
(12) Ishima, R.; Torchia, D. A. Protein Dynamics from NMR. *Nature structural biology* **2000**, *7* (9), 740.
(13) McPherson, A. *In Situ* X-Ray Crystallography. *J. Appl. Crystallogr.* **2000**, *33* (2), 397−400.
(14) Johnson, W. C. Secondary Structure of Proteins Through Circular Dichroism Spectroscopy. *Annu. Rev. Biophys. Biophys. Chem.* **1988**, *17*, 145−166.
(15) Lerner, E.; Barth, A.; Hendrix, J.; Ambrose, B.; Birkedal, V.; Blanchard, S. C.; Börner, R.; Sung Chung, H.; Cordes, T.; Craggs, T. D.; Deniz, A. A.; Diao, J.; Fei, J.; Gonzalez, R. L.; Gopich, I. V.; Ha, T.; Hanke, C. A.; Haran, G.; Hatzakis, N. S.; Hohng, S.; Hong, S.-C.; Hugel, T.; Ingargiola, A.; Joo, C.; Kapanidis, A. N.; Kim, H. D.; Laurence, T.; Lee, N. K.; Lee, T.-H.; Lemke, E. A.; Margeat, E.; Michaelis, J.; Michalet, X.; Myong, S.; Nettels, D.; Peulen, T.-O.; Ploetz, E.; Razvag, Y.; Robb, N. C.; Schuler, B.; Soleimaninejad, H.; Tang, C.; Vafabakhsh, R.; Lamb, D. C.; Seidel, C. A.; Weiss, S. FRET-Based Dynamic Structural Biology: Challenges, Perspectives and an Appeal for Open-Science Practices. *eLife* **2021**, *10*, No. e60416.
(16) Morsa, D.; Baiwir, D.; La Rocca, R.; Zimmerman, T. A.; Hanozin, E.; Grifnée, E.; Longuespée, R.; Meuwis, M.-A.; Smargiasso, N.; Pauw, E. D.; Mazzucchelli, G. Multi-Enzymatic Limited Digestion: The Next-Generation Sequencing for Proteomics? *J. Proteome Res.* **2019**, *18* (6), 2501−2513.
(17) Pandeswari, P. B.; Sabareesh, V. Middle-down Approach: A Choice to Sequence and Characterize Proteins/Proteomes by Mass Spectrometry. *RSC Adv.* **2019**, *9* (1), 313−344.
(18) Fontana, A.; Polverino de Laureto, P.; Spolaore, B.; Frare, E.; Zambonin, M. Detecting Disordered Regions in Proteins by Limited Proteolysis. In *Instrumental Analysis of Intrinsically Disordered Proteins*; Uversky, V. N., Longhi, S., Eds.; John Wiley & Sons, Inc.: Hoboken, 2010; pp 569−626.
(19) Fontana, A.; de Laureto, P. P.; Spolaore, B.; Frare, E.; Picotti, P.; Zambonin, M. *Probing Protein Structure by Limited Proteolysis* **2019**, *51*, 299.
(20) Kuwata, K.; Era, S.; Hoshino, M.; Forge, V.; Goto, Y.; Batt, C. A. Solution Structure and Dynamics of Bovine β-Lactoglobulin A. *Protein Sci.* **1999**, *8* (11), 2541−2545.
(21) Banci, L.; Bertini, I.; Gray, H. B.; Luchinat, C.; Reddig, T.; Rosato, A.; Turano, P. Solution Structure of Oxidized Horse Heart Cytochrome c. *Biochemistry* **1997**, *36* (32), 9867−9877.
(22) Andrews, A. L.; Atkinson, D.; Evans, M. T. A.; Finer, E. G.; Green, J. P.; Phillips, M. C.; Robertson, R. N. The Conformation and Aggregation of Bovine?-Casein A. I. Molecular Aspects of Thermal Aggregation. *Biopolymers* **1979**, *18* (5), 1105−1121.
(23) Sipos, T.; Merkel, J. R. An Effect of Calcium Ions on the Activity, Heat Stability, and Structure of Trypsin. *Biochemistry* **1970**, *9*, 2766.
(24) Vandermarliere, E.; Mueller, M.; Martens, L. Getting Intimate with Trypsin, the Leading Protease in Proteomics: TRYPSIN IN PROTEOMICS. *Mass Spectrom. Rev.* **2013**, *32* (6), 453−465.
(25) Gardner, Q.-A. A.; Younas, H.; Akhtar, M. Studies on the Regioselectivity and Kinetics of the Action of Trypsin on Proinsulin and Its Derivatives Using Mass Spectrometry. *Biochim. Biophys. Acta, Proteins Proteomics* **2013**, *1834* (1), 182−190.
(26) Liigand, P.; Kaupmees, K.; Kruve, A. Influence of the Amino Acid Composition on the Ionization Efficiencies of Small Peptides. *J. Mass Spectrom.* **2019**, *54* (6), 481−487.
(27) Tolkach, A.; Kulozik, U. Reaction Kinetic Pathway of Reversible and Irreversible Thermal Denaturation of β-Lactoglobulin. *Lait* **2007**, *87* (4−5), 301−315.
(28) Lee, B.; Richards, F. M. The Interpretation of Protein Structures: Estimation of Static Accessibility. *J. Mol. Biol.* **1971**, *55* (3), 379−IN4.
(29) Shrake, A.; Rupley, J. A. Environment and Exposure to Solvent of Protein Atoms. Lysozyme and Insulin. *J. Mol. Biol.* **1973**, *79* (2), 351−371.
(30) Papaleo, E.; Saladino, G.; Lambrughi, M.; Lindorff-Larsen, K.; Gervasio, F. L.; Nussinov, R. The Role of Protein Loops and Linkers in Conformational Dynamics and Allostery. *Chem. Rev.* **2016**, *116* (11), 6391−6423.
(31) Nick Pace, C.; Martin Scholtz, J. A Helix Propensity Scale Based on Experimental Studies of Peptides and Proteins. *Biophys. J.* **1998**, *75* (1), 422−427.
(32) Fisher, W. R.; Taniuchi, H.; Anfinsen, C. B. On the Role of Heme in the Formation of the Structure of Cytochrome c. *J. Biol. Chem.* **1973**, *248* (9), 3188−3195.
(33) Farrell, H. M., Jr.; Wickham, E. D.; Unruh, J. J.; Qi, P. X.; Hoagland, P. D. Secondary Structural Studies of Bovine Caseins: Temperature Dependence of b-Casein Structure as Analyzed by Circular Dichroism and FTIR Spectroscopy and Correlation with Micellizationq. *Food Hydrocolloids* **2001**, *15*, 341.
(34) Livney, Y. D.; Schwan, A. L.; Dalgleish, D. G. A Study of β-Casein Tertiary Structure by Intramolecular Crosslinking and Mass Spectrometry. *J. Dairy Sci.* **2004**, *87* (11), 3638−3647.
(35) Kumosinski, T. F.; Brown, E. M.; Farrell, H. M., Jr. *Three-Dimensional Molecular Modeling of Bovine Caseins: An Energy-Minimized -Casein Structure1* **1993**, *76* (4), 931.

https://doi.org/10.1021/jasms.1c00274
J. Am. Soc. Mass Spectrom. 2022, 33, 284−295