J Biol Phys (2018) 44:345–360
https://doi.org/10.1007/s10867-018-9494-7

ORIGINAL PAPER

Laser interferometry of the hydrolytic changes in protein solutions: the refractive index and hydration shells

R. M. Sarimov[1] · T. A. Matveyeva[1] · V. N. Binhi[1,2]

Received: 3 May 2017 / Accepted: 29 March 2018 / Published online: 11 May 2018
© Springer Science+Business Media B.V., part of Springer Nature 2018

Abstract Using an original laser interferometer of enhanced sensitivity, an increase in the refractive index of a protein solution was observed during the reaction of proteolysis catalyzed by pepsin. The increase in the refractive index of the protein solution at a concentration of 4 mg/ml was 9×10^{-6} for bovine serum albumin and 2.4×10^{-6} for lysozyme. The observed effect disproves the existing idea that the refractive index of protein solutions is determined only by their amino acid composition and concentration. It is shown that the refractive index also depends on the state of protein fragmentation. A mathematical model of proteolysis and a real-time method for estimating the state of protein hydration based on the measurement of refractive index during the reaction are proposed. A good agreement between the experimental and calculated time dependences of the refractive index shows that the growth of the surface of protein fragments and the change in the number of hydration cavities during proteolysis can be responsible for the observed effect.

Keywords Enzymatic proteolysis · Optical methods in biochemistry · Mathematical model · Computer simulation

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10867-018-9494-7) contains supplementary material, which is available to authorized users.

✉ R. M. Sarimov
rusa@kapella.gpi.ru

[1] Prokhorov General Physics Institute of the Russian Academy of Sciences, GPI RAS, 38 Vavilov str., 119991 Moscow, Russia

[2] Faculty of Biology, Lomonosov Moscow State University, 1-12 Leninskie Gory, 119991 Moscow, Russia

1 Introduction

At present, there is a perception that the refractive index n of a protein solution depends on the concentration and amino acid composition of the protein, but does not depend on the fragmentation state of the protein [1–3]. This statement is based mainly on the experimental data of [4], where no changes in the refractive index were observed during hydrolysis of proteins. Further studies, e.g., [1], showed that the constant of refraction R of a protein is equal to the sum of the refraction constants of its amino acids. In other words, the constant of refraction and the refractive index of a protein, connected by the Lorentz–Lorenz formula, indeed do not depend on the state of fragmentation of the protein in solution.

However, this statement, like the theoretical estimates [3, 5] of the refractive index of protein solutions, is based on experimental data obtained mainly with the use of refractometry with a measurement accuracy on the order of 10^{-5}. Significant also is that the estimates did not take into account the interaction of the protein with the solvent.

Experimentally, this interaction leads to changes by an order of magnitude smaller; so it could not be detected in early experiments. However, in later works, it has already been found that the increment of the refractive index of a protein solution $\mathrm{d}n/\mathrm{d}c$ depends on the nature of the solvent. Solutions of the same protein of equal concentration, prepared in different solvents, had different increments [6]. These differences are obviously not related to the amino acid composition of proteins. They might be caused only by the interaction of the proteins with solvent, in particular by changing the optical properties of the solvent near the surface of the protein.

In recent decades, there has been a significant development of optical methods in biochemistry and biophysics. New works have appeared, in which the refractive index of solutions of amino acids and proteins is measured. Studies can be conditionally divided into those in which the volumes of the solutions studied are of the order of 1 ml and more, and those where the volumes are less than a microliter or even much smaller—correspond to monomolecular layers of substances.

In the first case, measurements of n are carried out mainly to control the equilibrium concentration of substances [7] or for dynamic control during crystal growth [8]. As a rule, the accuracy of measuring n in such studies is low, of the order of 10^{-4}–10^{-3}.

In studies with microvolumes of solutions, the interactions of proteins and protein complexes with surfaces or immobilized substances are usually studied. For example, the work [9] evaluated the adsorption of various proteins on hydrophilic and hydrophobic surfaces. In works [10, 11], the refractive index of a mixture of proteins was measured by the method of backscattering interferometry. The volumes of solutions that are used in backscattering interferometry are hundreds of pl, and the accuracy of determining n can reach 2×10^{-7}. The resolution of the refractive index meters that are based on surface plasmon resonance—where tens of pl of the analyte is also enough—reaches 4×10^{-8} [12]. However, these most sensitive sensors, like other thin-film refractometers, are not suitable for measuring the refractive index of solutions during proteolysis of proteins, since such reactions cannot pass completely on the sensor surface.

At present, the refractive index of solutions is not used to evaluate the kinetics of enzymatic reactions or parameters of hydration protein shells in solution. This is due to the fact that measuring n in large volumes with the necessary resolution of the order of 10^{-6} is a technically difficult problem. In our work, an original laser interferometer is used for differential measurement of n of two solutions, control and experimental, in 5-cm cells, which is convenient in the study of some biochemical reactions. The sensitivity threshold to a refractive index variation was about 5×10^{-7}, which is close to the usual refractive

index resolution of the surface plasmon resonance sensors with angular modulation [13]. Such a resolution has been achieved in our large-base interferometer by using a 400-kg floating stone base that served also as a thermostat for the samples, the selection of a single-mode laser beam, and precise electronics. Real-time measurements using this interferometer allowed one to monitor changes Δn during the proteolysis reaction.

The purpose of this work was (1) to reveal changes in the refractive index of solutions of two different proteins during their pepsin hydrolysis, (2) to develop a model linking the observed effect with the change in the surface area of hydration shells and the number of hydration protein cavities, and (3) to compare calculations with experiment for obtaining information about the state of protein hydration.

2 Materials and methods

2.1 Laser interferometer

The laser interferometer was constructed according to the Mach-Zehnder scheme, Fig. 1. A He-Ne laser LGN-215 with a wavelength of 633 nm was used. The radiation power at the input of the interferometer was about 1 mW. The optical system consists of two translucent mirrors, two opaque mirrors, and a lens for scaling the interference pattern. Changes in the interference pattern during the experiment are due to the difference in the refractive indices of the solutions in the test and control cuvettes. Four photo diode sensors are used to measure the intensity of the interference pattern in its different sections. Signals from photo diodes are digitized on a 16-bit ADC-DAC NI 6251 (National Instruments) and are normalized to the intensity of the laser radiation.

The interference pattern is sensitive to fluctuations in temperature in the cuvettes. To reduce the temperature fluctuations, the interferometer is closed by a thermostatic screen seen as black panels in Fig. 1, and the cuvettes are placed directly on parts of a massive base of the interferometer that acts as a thermostat.

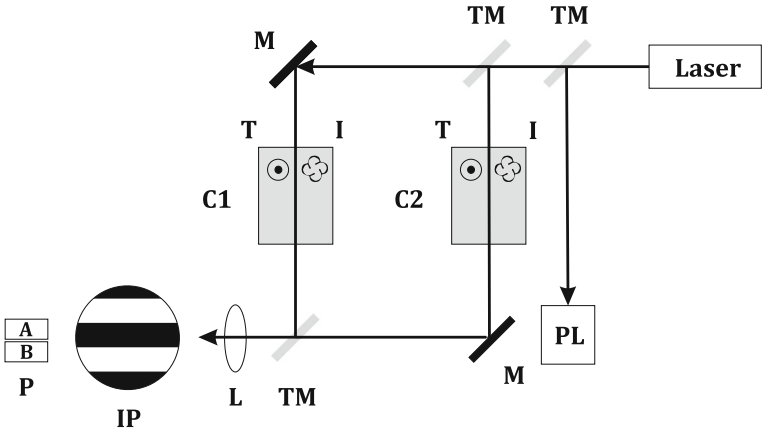

Fig. 1 Schematic diagram of the interferometer. *M* mirror, *TM* translucent mirror, *P* photodiode sensors A and B, *IP* interference pattern, *PL* photodiode for emission intensity normalization, *L* lens, *C1* and *C2* exposure and control cuvettes, *T* thermo sensor, *I* plastic impeller

In the course of the experiment, gradients of density and temperature inside the cuvettes appear. For their elimination, plastic impellers rotating at a frequency of 1 r/s were placed in the cuvettes. For temperature measurements, Honeywell HEL-705 submersible platinum sensors are used. The temperature difference between the test and control cells did not exceed 0.05 °C during the hour-long experiment. The temperature trends of the solutions were automatically taken into account when calculating the refractive index. A detailed description of the interferometer is available in [14].

The accuracy of measuring Δn of the protein solutions in this refractometer was determined mainly by fluctuations in the refractive index of the solutions themselves. The standard deviation of these fluctuations in hour-long measurements was about 2×10^{-7}, which determined the resolution of the refractometer. Fluctuations in the protein solutions exceeded the analogous fluctuations in water (SD 4×10^{-8}), in water with addition of HCl (pH = 1.5, SD 8×10^{-8}), as well as the intrinsic noise of the temperature sensors, recalculated into the noise of n (SD 5×10^{-8}).

2.2 Proteolysis of the protein

Pepsin-catalyzed proteolysis of the proteins was carried out as follows. On the day of the experiment, aqueous protein solutions were prepared from powdered bovine serum albumin (Merck), lysozyme (EC 3.2.1.17, Amresco) and pepsin (EC 3.4.23.1, Sigma, protein content 5.1%, activity ~1100 units/mg). The choice of water as a solvent is due to the minimum absorption of laser radiation. Water for solutions was obtained by distillation and deionization; electrical resistivity of water exceeded 5MΩ/cm. The average temperature of the solutions in the experiments with BSA was 22.9 ± 0.7 °C, and in experiments with lysozyme 21.3 ± 0.9 °C. HCl was added to the control and test cuvettes to shift the acidity of the solution to pH ~ 1.5, which increased the reaction rate and made it convenient for observation. The acidity control was performed with an error of pH about 0.1. Then, early dissolved BSA or lysozyme (LZM) was placed in both cuvettes. Before starting the measurements, the solution of pepsin was added to the test cuvette, and water to the control one, after which solutions were stirred. The volume of solutions was 18 ml. The concentration of BSA $c_B = 4$ mg/ml and LZM $c_L = 4$ mg/ml in the control and test cuvettes were the same, the concentration of pepsin in the experimental cuvette was $c_P = 2.2$ mg/ml. After the addition of pepsin, Δn measurement was started based on the shift of the interference pattern. The measurement of Δn was carried out for an hour; during this time, a significant amount of the protein was divided into peptide fragments.

2.3 The control of protease activity of pepsin

The control of protease activity of pepsin was carried out by electrophoresis of the products of proteolysis in polyacrylamide gel, dynamic light scattering, and UV absorption. Data on electrophoresis and dynamic light scattering are presented in [15]. UV absorption in the range 240–320 nm with a step of 1 nm was measured with a Shimadzu UV-1800 spectrophotometer for 50 min every 30 s. The measurements were carried out in a standard cuvette 10 × 10 × 45 mm. Pepsin solution of the volume 3.2 ml at concentration 2.2 mg/ml and pH = 1.5 was added to the test protein at concentration of BSA 4 mg/ml. The concentration of LZM was reduced to 1 mg/ml because of too high absorption. The products of the proteolysis reaction went directly in a spectrophotometer at a temperature of 22.4 °C.

3 Results

3.1 Proteolysis-induced changes in the refractive index of BSA- and LZM-pepsin solutions

Figure 2 shows the change in the refractive index of the albumin solution after the addition of pepsin. Kinetic curves are the average of the results of ten experiments; the standard errors of the mean are indicated. It can be seen from the figure that in the first 15–20 min there is a sharp increase in the refractive index in the solution of BSA with pepsin, $\Delta n \sim 9 \times 10^{-6}$ (curve 1). Further growth slows down, and Δn reaches 12×10^{-6} after 1 h of hydrolysis. After 20–30 min from the beginning of proteolysis, the rate of change in Δn is practically the same for both the solution of pepsin with BSA and separately for the solution of pepsin without BSA (curve 2). The increase in the refractive index in the latter case is most likely due to the hydrolysis of products entering the enzyme preparation, whose content is about $\sim$ 95% from the mass of the preparation. For the sake of brevity, this process is referred to as "self-hydrolysis".

The maximum proteolytic activity of pepsin is known to occur at pH = 1.5–2. To test the proteolytic activity of pepsin, the experiments were repeated with neutral acidity, i.e., at the same concentrations of protein and enzyme, but without the addition of HCl. Under these conditions, the proteolysis reaction practically did not go: the refractive index in one hour increased by less than 2×10^{-6}.

To highlight changes in Δn caused by the hydrolysis of only BSA, but not of the enzyme preparation, the curve 2 of self-hydrolysis was subtracted from the kinetic curve 1 of BSA hydrolysis. The result is shown in Fig. 3, curve 1. Thus, with the addition of pepsin to the BSA solution, the refractive index increased by 9×10^{-6} in an hour, and 95% of the effect was achieved 30 min after addition of pepsin.

Figure 3 also shows the change in Δn during hydrolysis of another protein, lysozyme, obtained in similar experiments and with the same self-hydrolysis of the enzyme preparation. For lysozyme, the kinetic curve averages six experiments.

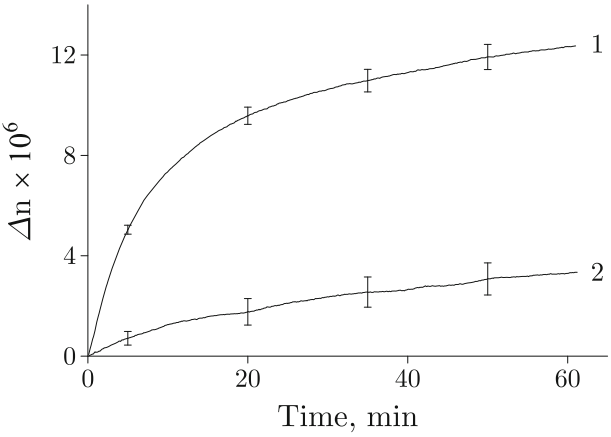

Fig. 2 Changes in the refractive index in the hydrolysis reaction of albumin with an enzyme preparation at pH = 1.5: (1) BSA 4 mg/ml and pepsin 2.2 mg/ml; (2) only pepsin 2.2 mg/ml. Data are shown as means ± SEM

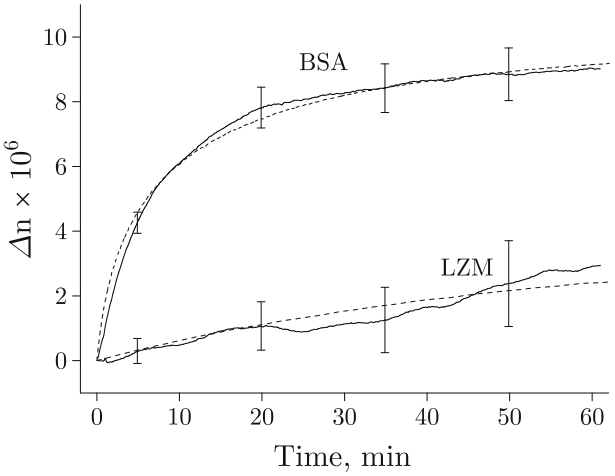

Fig. 3 The change in the refractive index as mean ± SEM. during proteolysis of the protein (4 mg/ml) with pepsin (2.2 mg/ml), taking into account the self-hydrolysis of the enzyme preparation: *solid lines* experimental data, *dash lines* dependences calculated from the hydrolytic fragmentation model according Eq. (8) with $u = 2.7 \times 10^{-3}$ nm, $\tau_{BSA} = 0.33$ s, $\tau_{LZM} = 6.1$ s

3.2 Theoretical refractive index dependence on the properties of protein hydration

In the above-described experimental model of proteolysis, the enzyme was able to cut the protein molecule in many points. In other words, the substrate represented not one substance, but a mixture of the original protein and peptides—the fragments arising during the reaction. For this reason, the use of the Michaelis–Menten equation to describe the kinetics of enzymatic reactions is difficult, although it is noted in many studies that the shape of the kinetic curve for substrate reduction or for product increase during the enzymatic catalysis coincides with the form of a simple one-step enzymatic reaction [16]. The models that are proposed for the description of this phenomenon are based on simplifications, in which the number of intermediates is limited, as in the "one-by-one mechanism" or they are completely absent, as in the "zipper mechanism" [16–18]. This is done in order that one could operate with concentration dependencies and use the Michaelis–Menten equation.

It should be noted that the kinetic curve of the growth of the refractive index during BSA proteolysis is a consequence of the integral contribution of many intermediates and 89 products. In this case, the description of the processes in which they participate in terms of concentrations and equations of chemical kinetics becomes ineffective. Therefore, when describing the kinetics of this enzymatic reaction, we used an original approach that consisted in a computer simulation of proteolysis without using the Michaelis–Menten equation, and in using the Lorentz–Lorenz equation to estimate the change in the refractive index of the solution.

The Lorentz–Lorenz equation [19] links the refractive index n of a substance of volume V with the amount q of particles of polarizability α that make up the substance:

$$\frac{n^2 - 1}{n^2 + 2} = \frac{4\pi}{3V} q\alpha$$

For a solution consisting of water molecules and various fragments of protein molecules, one can write

$$\frac{n^2-1}{n^2+2} = R,\ R = \frac{4\pi}{3V}\left(N\alpha' + \sum_k q_k\alpha_k\right) \tag{1}$$

where R is the constant of refraction, N and α' are the number and polarizability of water molecules, q_k and α_k are those of the fragments of type k. Since we are only interested in the changes of n, we do not take into account the presence of enzyme molecules, the number of which is assumed to be constant.

Changes in n are small, less than 10^{-5}, and are called by small changes in R. Calculating the derivative $\mathrm{d}n/\mathrm{d}R = (3/2)n^{-1}(1-R)^{-2}$, one can find

$$\Delta n(R) = \frac{3(R-R_0)}{2n_0(1-R_0)^2},\ n_0 \equiv \left(\frac{1+2R_0}{1-R_0}\right)^{1/2} \tag{2}$$

where n_0 and R_0 are the corresponding quantities at the initial moment of time $t = 0$, $\Delta n \equiv n - n_0$ is the change in the refractive index during hydrolysis.

As shown below, the change in the refractive index of a protein solution is mainly due to the fact that the local density of water near and inside the protein molecule may differ from the average density of water in the bulk volume.

On the one hand, there are more hydrophilic polar groups on the protein surface, the non-uniform electric field of which attracts water dipoles. In addition, the acidity of the solutions used, pH $\sim$ 1.5, is below the isoelectric point of many peptides. With such acidity, the protein fragments are generally electrically charged, which creates an additional heterogeneity of the electric field near the surface and is another factor in the attraction of water dipoles to the surface. For this reason, in the near-surface layer, water density ρ_h is on average higher than that ρ_w in the bulk volume and approaches to ρ_w with the distance from surface. During protein hydrolysis, the total surface area of the fragments increases. Hence, the amount of water with a higher density increases, and the refractive index of the solution should increase.

On the other hand, protein molecules are not compact and may contain cavities with a small number of water molecules, or even with their complete absence, since, for example, hydrophobic cavities of the order of 1 nm or less repel water [20]. Hydrolysis of the protein is accompanied by the disappearance of such cavities. Consequently, the amount of water in a less dense state, with density ρ_c, decreases, and the refractive index of the solution should increase.

The total effect of these two factors determines the change in the density of the solution and, consequently, its refractive index during hydrolysis.

Below we consider an idealized model of the protein solution, in which the protein molecules and their fragments are assumed to be spheres surrounded by hydration shells of thickness d, and hydration cavities are spheres of radius r.

To estimate the volume V of the solution, we write

$$V = \theta + V_h + V_c + V_w \tag{3}$$

where $\theta = m_q q/\rho_q$ is the total volume of protein fragments that is constant during proteolysis, m_q is the mass of a protein molecule, q is the initial number of protein molecules in solution, ρ_q is the density of the protein substance. The volumes V_h, V_c, and V_w—water volumes—are, respectively, the volume of water with density ρ_h in the hydration shells, the volume of water with density ρ_c in the cavities, and the volume of water with normal density ρ_w outside hydration shells and cavities.

The volume of water in the hydration shells is

$$V_h = dS$$

where $S \equiv \sum_k q_k s_k$ is the area of the hydration shells, q_k and s_k are the number and surface area of the fragments of k-th type.

Since the cavities are of finite size, the number of cavities during hydrolysis is not conserved, but decreases in proportion to the surface area of the fragments. If all the fragments are less than r in their size, then there are no cavities at all. The results of studies on the dynamic light scattering of protein solutions in an acid medium show that at pH = 2 the BSA molecule swells and its size increases by approximately 40% compared to the size at neutral pH [21]. This is the initial state of the protein in our experiments. It can be shown that the volume of the cavities of a molecule that disappear during hydrolysis is $zr\sigma$ in the order of magnitude, where r is the radius of the cavity, σ is the area of the added surface, and z is a coefficient of the order of unity. Therefore, if v is the initial volume of the cavities of q molecules, i.e., $v = (1.4^3 - 1)\theta$, then during the process of hydrolysis the volume of cavities decreases with the area S of the hydration shells as

$$V_c = v - zr(S - qs)$$

where s is the surface area of a single molecule of BSA before hydrolysis.

To find the volume of water V_w outside the cavities and hydration shells, one should take into account water densities ρ_c and ρ_h in these states. Defining the number of water molecules in the hydration shells and cavities, respectively, as $h \equiv \rho_h V_h/m$ and $c \equiv \rho_c V_c/m$, where m is the mass of a water molecule, we find the number of molecules w of water outside these regions: $w = N - h - c$. Hence,

$$V_w = mw/\rho_w = W - \rho_h V_h/\rho_w - \rho_c V_c/\rho_w$$

where $W = mN/\rho_w$ is the virtual volume of water, as if it were all a bulk water.

We define relative changes in the density of water δ_h in the hydration shells and δ_c in the cavities, respectively, according to definitions $\rho_h \equiv \rho_w(1 + \delta_h)$ and $\rho_c \equiv \rho_w(1 - \delta_c)$. Then increased water density in the hydration shells and reduced water density in the cavities correspond to the positive values of δ_h and δ_c. Further, substituting everything in (3), we get

$$V = \theta + W - \delta_h V_h + \delta_c V_c = V_0 - (\delta_h d + \delta_c zr)S \tag{4}$$

where $V_0 = \theta + W + \delta_c(v + zrqs)$ is the coefficient that does not change during hydrolysis. Since $zrqs \ll v$ and $v \sim \theta \ll W$, then we assume $V_0 = W$. With a designation

$$u \equiv \delta_h d + \delta_c zr \tag{5}$$

combining (1) and (4), we obtain for the refraction constant the expression

$$R(t) = \frac{4\pi}{3}\frac{N\alpha' + A(t)}{W - uS(t)} \tag{6}$$

where the notations $A(t) \equiv \sum_k q_k \alpha_k$ and $S(t) \equiv \sum_k q_k s_k$ are introduced for sums that depend on time. At the initial moment of time, there is only one sort of "fragments"—the intact protein molecules in the amount of $q \equiv q_0$ with a surface area $s \equiv s_0$ and polarizability $\alpha \equiv \alpha_0$. Obviously, $A(0) = q\alpha$, $S(0) = qs$ and

$$R_0 \equiv R(0) = \frac{4\pi}{3}\frac{N\alpha' + q\alpha}{W - uqs} \tag{7}$$

It is shown below that the function $A(t)$ is equal to constant $q\alpha$ that is small in comparison with $N\alpha'$. Thus, the relations (2), (6), and (7) determine the change in the refractive index n during proteolysis.

Due to the uniqueness of the protein fragments, the analytic derivation of the time-dependent density distribution of fragment sizes and, correspondingly, derivation of the function $S(t)$ is impossible. A computer simulation of proteolysis is used below, which allows finding the form of this function and estimating the change in the refractive index.

We assume that the state of fragmentation of a large ensemble of original q independent protein molecules at some point in time can be represented by the averaged result of sequential fragmentation of a single molecule. Averaging is performed over a large number of copies of the random process of fragmentation of one molecule.

The amino acid residue sequences of BSA (UniProt P02769) and LZM (UniProt P00698) and of their fragments obtained from the reaction with pepsin at pH < 2 are well known [22, 23]. Pepsin produces preferential cleavage of the protein chain at peptide bonds following the amino acid residues of phenylalanine (F) and leucine (L). This fact allows one to simulate the process of cutting protein molecules. For example, the BSA chain consists of 583 amino acids, and the elements F and L occupy certain positions. An iterative process of fragmentation of a molecule that is represented as such polypeptide chain is considered. At each step of the iteration, a uniformly distributed integer random number $z \in (1, L)$ is generated, where L is the number of amino acids in the chain. If the generated number corresponds to the number of any F- or L-element, then the "molecule" is considered to be cut after the element with the number z. As a consequence, fragments of different lengths appear.

The fragmentation state after some iteration step i is represented by a chain divided into segments at points z_k and, respectively, by a set of fragments with lengths $l_k = z_{k+1} - z_k$. The next step of the iteration generates the next random number z. If it coincides with one of the undivided points F or L, a new cut appears and, correspondingly, a pair of new fragments appear instead of the previous one. Then, the next step of the iteration occurs. If there is no coincidence, then the next step occurs without adding a cut. The process of 'proteolysis' is completed after a certain number of iterations, when there are no free points for cutting.

In this model, (1) there are q identical protein molecules, each of which consists of k fragments, (2) all fragments of the molecule are unique, and (3) the polarizability of the fragments are proportional to their masses [2, 3], or to the lengths of the corresponding intervals (z_k, z_{k+1}). For these reasons, A from (6) is reduced to a constant. Really,

$$A(t) \equiv \sum_k q_k \alpha_k \to q \sum_k \alpha_k = q\alpha \sum_k l_k = q\alpha$$

where the arrow denotes the transition from an ensemble of fragments of q original molecules to q copies of fragments of a single initial molecule.

Function $S(t)$ is not reducible to constant and, in the simulation procedure, it depends on the number i of the iteration step,

$$S(i) : S(t) \equiv \sum_k q_k s_k \to q \sum_k s_k = qs \sum_k s'_k$$

where $s'_k \equiv s_k/s$ is the relative surface area of the fragments that depends on i, and s is the surface area of the intact protein molecule:

$$s \equiv 4\pi \left(\frac{3m_q}{4\pi\rho_q} \right)^{2/3} \tag{8}$$

We specify that it is the state of fragmentation that is represented by a chain of elements with "cuts" at points z_k. The physical state of the fragments is the secondary structure of the chain, or a folded chain. To estimate the areas of the surface of fragments, we assume them to be folded into globules of spherical shape, with masses proportional to their lengths, or the intervals $l_k \equiv z_{k+1} - z_k$. It is not difficult to show that the relative surfaces are connected with the lengths l_k by the relation $s'_k = l_k^{2/3}$ and, therefore, $S(t) = qs \sum_k l_k^{2/3}$. Computer simulation allows one, at each step of the iteration, to find the average value of this sum over the set of realizations,

$$\sigma(i) \equiv \sum_k l_k^{2/3}(i) \tag{9}$$

and, thus, determine the dependence $S(i) = qs\sigma(i)$, and hence $\Delta n(i)$.

This, however, is not enough for comparison of Δn with the experimental curve. The latter is a dependence on time t rather than on the iteration number i. Note that the iteration number in this model is not only a simulation variable, but also a time function. Indeed, the amount of iterations, or the number of the current iteration, is incremented by one as a result of the next cutting. However, the probability of cutting is monotonously decreasing due to the reduction in the number of fragments suitable for cutting. Thus, it is necessary to find the dependence $i(t)$, which is feasible analytically.

The change in the number of fragments per unit time is determined by the probability p of cutting, which is obviously equal to the number of available points for cutting divided by the number of elements in the chain. In the continuum limit, if i is assumed to be a continuous quantity, $\mathrm{d}i/\mathrm{d}t = p(i)/\tau$, where τ is the time scale at the beginning of the process, when the factor of reducing the number of places suitable for cutting is not yet significant. The solution to this differential equation is the implicit function $t(i) = \tau \int_0^i p^{-1}(j)\mathrm{d}j$. For discrete time instants $t_i \equiv t(i)$

$$t(i) = \tau T(i), T(i) \equiv \sum_{j=0}^{i} p^{-1}(j) \tag{10}$$

It is easy to show by mathematical induction that the probability of cutting at i-th step of the iteration is given by expression $p(i) = p(0)\,[1 - 1/L]^i$, where L is the number of amino acids in the chain, and $p(0) \equiv F/L$, where F is the final number of fragments. Then

$$T(i) = \frac{1}{p(0)} \left\{ 1 - L \left[1 - \left(1 - \frac{1}{L} \right)^{1-i} \right] \right\} \tag{11}$$

Function $\sigma(i)$ (9) is completely determined by the idealized fragmentation procedure described above. This function is obtained by averaging over 10^3 realizations with a standard error generally not exceeding a fraction of a percent of the corresponding means. Functions $\sigma(i)$ (9) and $T(i)$ (11) are shown in Fig. 4. Since the final set of fragments of BSA and LZM hydrolysis is different, consists of 89 and 11 unique peptides, respectively, these functions for these proteins are different. It can be seen that in the process of proteolysis, the surface area of BSA fragments increases by 3.5 times, and of LZM—by two times. The dependence of T on the iteration step number is close to exponential function, where $i \gtrsim 100$.

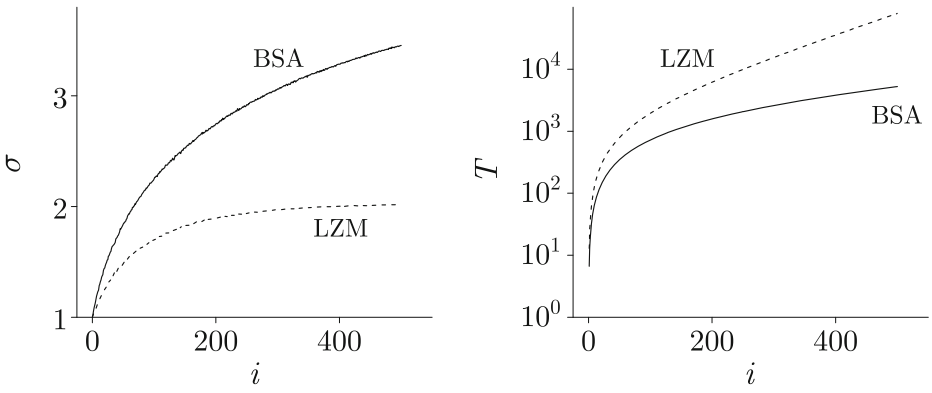

Fig. 4 Functions $\sigma(i)$ and $T(i)$ for BSA and LZM

Functions $t(i)$ and $\Delta n(i)$ together provide a parametrically defined dependence $\Delta n(t)$ that is suitable for comparison with experiment. Taking into account all definitions (2, 6, 9, 10), one can write formulas for calculating $\Delta n(t)$ as follows,

$$\Delta n(i) = \frac{3[R(i) - R_0]}{2n_0(1 - R_0)^2}, R(i) = \frac{4\pi}{3}\frac{N\alpha'}{W - uqs\,\sigma(i)}, t(i) = \tau T(i) \tag{12}$$

where term $A(t) = q\alpha$ is omitted in $R(i)$ in view of inequality $q\alpha \ll N\alpha'$, and the quantities n_0 and R_0 are given by (2) and (7), respectively. Recall that the parameter u, determined by (5), depends on the size of cavities and hydration shells and on the density of water in these areas.

To estimate the change $\Delta n(t)$ in the refractive index of the BSA-pepsin solution, the following values are used: of polarizability—water molecules $\alpha' = 1.468\times10^{-24}$ cm^3, BSA molecules $\alpha = 6.884 \times 10^{-21}$ cm^3; of masses—water molecules $m = 2.99 \times 10^{-23}$ g, BSA molecules $m_q = 1.1 \times 10^{-19}$ g; of density (20 °C)—bulk water $\rho_w = 0.9983$ g/cm^3, BSA $\rho_q = 1.32$ g/cm^3. The number of molecules in 1 ml of solution: of water $N = 3.334 \times 10^{22}$, of BSA molecules in the initial concentration 4 mg/ml $q = 3.626 \times 10^{16}$. For LZM-pepsin solution: the mass of a LZM molecule is 2.3757×10^{-20} g, LZM density is 1.237 g/cm^3, the number of LZM molecules in the initial concentration 4 mg/ml is 1.684×10^{17}.

There are three parameters for fitting the calculated and experimental curves: the characteristic times of BSA and LZM proteolysis and the parameter u that is the same for both proteins. The values ensuring the best correspondence are: $\tau_{\mathrm{BSA}} = 0.33$ s, $\tau_{\mathrm{LZM}} = 6.1$ s, $u = 2.7 \times 10^{-3}$ nm, Fig. 3. The accuracy of estimating these parameters is about 10%.

4 Discussion

Did proteolysis take place after the addition of pepsin to protein solutions in our experiments? Is the change in the refractive index related to proteolysis? The following facts allow one to answer these questions in the affirmative.

Previously, we investigated pepsin-catalyzed BSA proteolysis interferometrically and, in parallel, by the methods of electrophoresis, colloid chemistry, and dynamic light scattering (DLS) [15]. The electrophoresis data have shown that there were no intact BSA molecules

in the solution as early as the first minute after the addition of pepsin, and there were no fragments more than 10 kDa a few minutes later.

Using the DLS method, it was found that in the first 30 min after the addition of pepsin, the amount of scattered photons decreased, and the polydispersity index of the BSA solution, a parameter that shows the presence of particles of different sizes, increased. After 30 min, the observed parameters remained practically unchanged. The kinetics of changes in the quantities observed by the DLS method coincides with that in the refractive index.

In the present work, a spectrophotometric measurements were carried out to measure the absorption of solutions and confirm the kinetics of the proteolysis reaction, Fig. 5. In the course of reaction, the spectra shift to the "blue" region, Fig. 5a. The shift at wavelength range 270–295 is well known and is due to changes in the absorption of tryptophan and tyrosine [24]. To compare the kinetics of BSA and LZM proteolysis, measurements at the wavelength of 293 nm have been performed, Fig. 5b. The wavelength of 293 nm corresponds approximately to the half-width of the absorption maxima near 280 nm. During proteolysis, where the spectra shift to the blue region, the changes at this wavelength are close to the maximal values for both BSA and LZM.

The difference $OD_{0.5} - OD$ between the initial absorption and absorption in the process of proteolysis was normalized to the initial absorption. Measurements of the UV absorption difference between the initial and final stages of the reaction are often used to study the activity of enzymes. For example, in the work [25], also at a wavelength of 293 nm, the pepsin-catalyzed proteolysis of a LZM–substrate complex was studied.

The change in absorption of solutions during proteolysis, Fig. 5b, was similar to the change in the refractive index, Fig. 3. As in the case of the refractive index for the BSA–pepsin solution, a sharp increase occurs in the first 10–15 min of reaction, and a slow growth after about 20 min. The reaction time is consistent with pepsin activity indicated by the manufacturer.

It is interesting that both methods have shown an insignificant rate of the LZM proteolysis. This may be due to preservation of the protein structure close to native even at low pH values, and as a result, to the lack of availability of phenylalanine and leucine, where pepsin cuts the protein. Differences in the rate of pepsin protease activity with regard to different proteins were previously observed by many authors [26, 27]. Some proteins, such as β-lactoglobulin, are not practically hydrolyzed by pepsin, and the process of pepsin-catalyzed hydrolysis can be observed only when the protein is denatured either with the addition of ethanol [26] or by heating or destroying disulfide links [28].

Could factors such as the inaccuracy of temperature measurements, the change in pH, and the amount of dissolved gases affect the difference in the refractive index of solutions in the control and test cuvettes?

As noted above, temperature trends of a few hundredth degree in the course of the experiment were taken into account in the automatic calculation of the refractive index based on the shift of the interference pattern. For this conversion, the well-known temperature dependence of n was used [29]. The fluctuations in Δn that could be related to the inaccuracy of temperature measurements, did not exceed 6.7×10^{-7}, in terms of the standard deviation. This is more than an order of magnitude smaller than the maximum effect 9×10^{-6} observed in this paper. Consequently, the observed effect could not be a consequence of inaccurate temperature measurements.

The change in Δn could not be attributed to a change in pH, since pH did not change within an hour after the addition of pepsin to the BSA solution.

The effect of soluble gases on the refractive index of water was studied in [30]. Theoretical estimates and experimental measurements have shown that the maximum refractive index difference at the wavelength of 633 nm between the samples of degassed water and

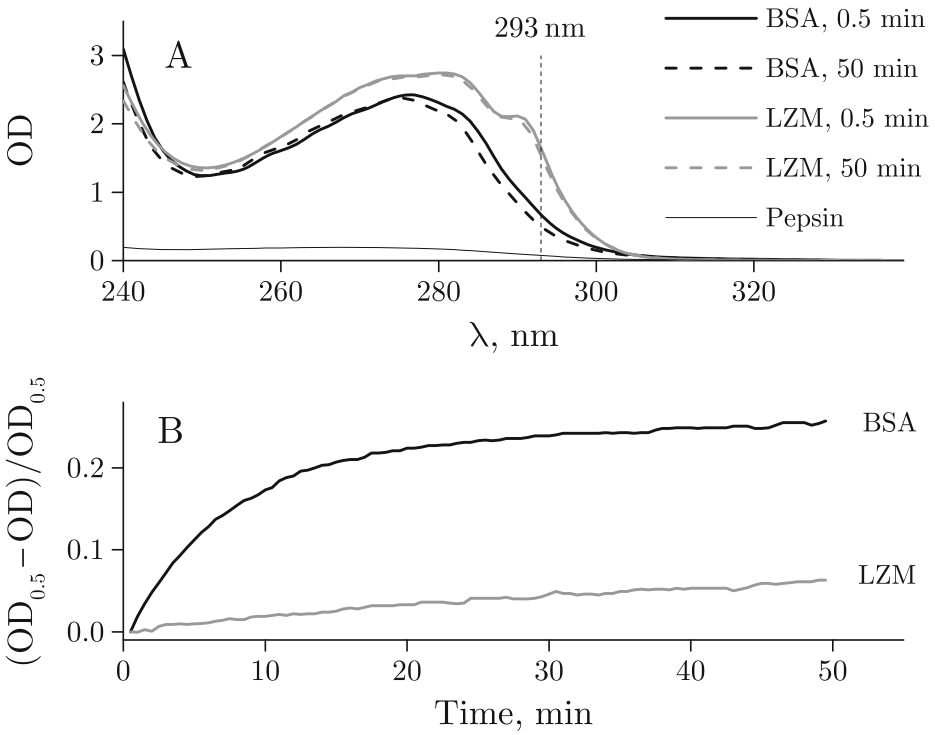

Fig. 5 **a** Absorption spectra at pH = 1.5 of the solution of BSA (4 mg/ml) and pepsin (2.2 mg/ml), LZM (1 mg/ml) and pepsin (2.2 mg/ml), only pepsin (2.2 mg/ml), 30 s and 50 min after the addition of pepsin. **b** Change in the relative absorption of BSA and LZM solutions during pepsin proteolysis at 293 nm: *OD* absorption in the process of proteolysis, $OD_{0.5}$ absorption at 30 s after the addition of pepsin

of water saturated with atmospheric gases, is less than 5×10^{-6}. Consequently, the effect observed in our experiment under conditions of approximately constant composition of dissolved gases and reaching $\Delta n = 1.2 \times 10^{-5}$, Fig. 2, is not associated with a change in the gas concentration.

All these facts confirm that in our experiments there was indeed a change in the refractive index of solutions due to the enzymatic pepsin–protein reaction.

The change in n during proteolysis indicates that the refractive index of a protein solution depends not only on the concentration and amino acid composition of the protein but also on its interaction with the solvent. Such an assumption was made earlier in [6] on the basis of indirect data. It was shown in this work that increment $\mathrm{d}n/\mathrm{d}c$ of a lysozyme solution depends on the nature of the solvent and varies from 0.153 cm^3/g for H_2O to 0.272 cm^3/g for buffer solution NaSCN 10 mM, HEPES 10 mM—that is almost double the initial value. Such a significant change in the increment, in comparison with the expected change in the interval 0.18–0.19 cm^3/g [3], indicates the participation of solvent molecules in the formation of the optical properties of the protein in solution. In other words, the solvent molecules influence the refractive index not only directly, but also indirectly, due to the protein–solvent interaction.

In our work, it is shown by a direct interferometric measurement that the refractive index of a protein solution varies during proteolysis and, hence, depends on the state of

fragmentation of molecules. The likely cause of this effect is the change in the amount of hydration water in the protein fragments, or peptides.

The model that is in good agreement with the experimental data confirms that the observed increase in the refractive index during hydrolysis is a consequence of an increase in the surface area of peptides and a reduction in the number of cavities in them compared to those of the intact protein molecules. The density of water in the hydration shells is greater, and inside the vanishing cavities is less than the density of bulk water. Due to both these factors, the relative amount of denser water in the process of hydrolysis increases, and consequently the refractive index increases.

Limitations of the model are related to the following. One of the idealizations is a spherical approximation for the shape of proteins, peptides, shells, and cavities. Another idealization is the equality of water densities in the hydration shells of intact protein molecules and peptides, while actually they could differ. In [31], the density of water in hydration shells was measured by means of neutron scattering and X-ray diffraction analysis, for proteins with different masses. An increase in the density of 8% relative to bulk water was observed for a protein with a mass of 14.5 kDa, and 16% for a protein with a mass of 68 kDa. The equality of the thickness of the hydration shells of intact protein molecules and of peptides is one more idealization. According to the data obtained using THz spectroscopy, the thickness of the hydration shell of BSA (66.5 kDa) is one and a half times greater than that of LZM (14.5 kDa) [32]. At the same time, these data on the density and thickness are not absolute and depend on the definition of these quantities.

At the same time, as Fig. 3 demonstrates, the model is applicable. From the comparison of the experiment and the theory, we found that the parameter of the model $u = \delta_h d + \delta_c z r$ has the value $(2.7 \pm 0.27) \times 10^{-3}$ nm. The relative variations in the water densities $\delta_h = \rho_h/\rho_w - 1$ and $\delta_c = 1 - \rho_c/\rho_w$ should be understood as values averaged over the volume of a hydration shell of thickness d and over the volume of a cavity of radius r, respectively. A possible refinement of the model is associated with the use of distribution functions for water densities, shell thicknesses, and cavity sizes instead of their average values, as well as a more realistic form of peptides.

It follows from (12) that the time dependence of the refraction constant R of the solution has the following structure

$$R(t) = \frac{k_1}{k_2 + k_3 u \sigma(t)}$$

where the coefficients k_1–k_3 do not depend on the state of fragmentation σ and, consequently, on the properties of hydration. This allows one to consider u an integral parameter of protein hydration.

Comparison of the interferometric data and the theory, Fig. 3, showed that the ratio of the time constants $\tau_{\mathrm{BSA}}/\tau_{\mathrm{LZM}}$ of the hydrolysis-conditioned changes in n of the BSA and LZM solutions is about 0.05, which does not contradict the spectrophotometric measurements Fig. 5b.

In general, the real-time laser interferometry of proteolytic changes in the refractive index of protein solutions represents a new method for studying the hydration of proteins.

5 Conclusions

The following statements are the result of this study.

An increase of $(9 \pm 0.8) \times 10^{-6}$ and $(2.4 \pm 1.3) \times 10^{-6}$ in the refractive index of BSA and LZM solutions, respectively, at a concentration of 4 mg/ml has been detected during the hour-long process of pepsin-catalyzed proteolysis.

By a direct interferometric measurements, the widespread assumption is refuted that the refractive index of protein solutions is only determined by the protein concentration and amino acid composition and is not dependent on the state of fragmentation of proteins.

A computational model is developed to describe the change in the refractive index of a protein solution during proteolysis. The model takes into account the change in the total area of the surface layer of protein fragments with altered density of water.

The observed correspondence between the experimental and computed time dependences of Δn shows that the growth of the surface of protein fragments during proteolysis can be responsible for the change in the refractive index.

Compliance with Ethical Standards

Declaration of interest The authors report no conflicts of interest.

References

1. McMeekin, T.L., Wilensky, M., Groves, M.L.: Refractive indices of proteins in relation to amino acid composition and specific volume. Biochem. Biophys. Res. Commun. **7**(2), 151 (1962). https://doi.org/10.1016/0006-291x(62)90165-1
2. McMeekin, T.L., Groves, M.L., Hipp, N.J.: Refractive indices of amino acids, proteins, and related substances. Adv. Chem. Ser. **1**, 54 (1964). https://doi.org/10.1021/ba-1964-0044.ch004
3. Zhao, H., Brown, P.H., Schuck, P.: On the distribution of protein refractive index increments. Biophys. J. **100**(9), 2309 (2011). https://doi.org/10.1016/j.bpj.2011.03.004
4. Barer, R., Joseph, S.: Refractometry of living cells. J. Cell Sci. **3**(32), 399 (1954)
5. Zhao, H., Brown, P.H., Magone, M.T., Schuck, P.: The molecular refractive function of lens γ-crystallins. J. Mol. Biol. **411**(3), 680–699 (2011). https://doi.org/10.1016/j.jmb.2011.06.007
6. Ball, V., Ramsden, J.J.: Buffer dependence of refractive index increments of protein solutions. Biopolymers **46**(7), 489 (1998). https://doi.org/10.1002/(sici)1097-0282(199812)46:7<489::aid-bip6>3.0.co;2-e
7. Cole, T., Kathman, A., Koszelak, S., McPherson, A.: Determination of local refractive index for protein and virus crystals in solution by mach-zehnder interferometry. Anal. Biochem. **231**(1), 92 (1995). https://doi.org/10.1006/abio.1995.1507
8. Yin, D.C., Inatomi, Y., Luo, H.M., Li, H.S., Lu, H.M., Ye, Y.J., Wakayama, N.I.: Interferometry measurement of protein concentration evolution during crystallization and dissolution with improved reliability and versatility. Meas. Sci. Technol. **19**(4), 045303 (2008)
9. Voros, J.: The density and refractive index of adsorbing protein layers. Biophys. J. **87**(1), 553 (2004). https://doi.org/10.1529/biophysj.103.030072
10. Markov, D.A., Swinney, K., Bornhop, D.J.: Label-free molecular interaction determinations with nanoscale interferometry. J. Am. Chem. Soc. **126**(50), 16659 (2004). https://doi.org/10.1021/ja047820m
11. Jepsen, S.T., Jorgensen, T.M., Zong, W., Trydal, T., Kristensen, S.R., Sorensen, H.S.: Evaluation of back scatter interferometry, a method for detecting protein binding in solution. The Analyst (Royal Society of Chemistry) **140**(3), 895 (2015). https://doi.org/10.1039/C4AN01129E
12. Kabashin, A.V., Nikitin, P.I.: Surface plasmon resonance interferometer for bio- and chemical-sensors. Opt. Commun. **150**(1–6), 5 (1998). https://doi.org/10.1016/s0030-4018(97)00726-8
13. Homola, J.: Surface plasmon resonance sensors for detection of chemical and biological species. Chem. Rev. **108**(2), 462 (2008). https://doi.org/10.1021/cr068107d
14. Binhi, V.N., Sarimov, R.M.: Relaxation of liquid water states with altered stoichiometry. Biophysics **59**(4), 515 (2014). https://doi.org/10.1134/S0006350914040058
15. Sarimov, R.M., Matveyeva, T.A., Vasin, A.L., Binhi, V.N.: Changes in the refractive index of a solution during proteolysis of bovine serum albumin with pepsin. Biophysics **62**(2), 177 (2017). https://doi.org/10.1134/S0006350917020221

16. Srividhya, J., Schnell, S.: Why substrate depletion has apparent first-order kinetics in enzymatic digestion. Comput. Biol. Chem. **30**(3), 209 (2006). https://doi.org/10.1016/j.compbiolchem.2006.03.003
17. Bull, H.B., Currie, B.T.: Peptic hydrolysis of egg albumin. I. Kinetic studies. J. Am. Chem. Soc. **71**(8), 2758 (1949)
18. Sachdev, G.P., Fruton, J.S.: Kinetics of action of pepsin on fluorescent peptide substrates. Proc. Natl. Acad. Sci. U.S.A. **72**(9), 3424 (1975)
19. Born, M., Wolf, E. Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light, 7th edn. Cambridge University Press, Cambridge (1999)
20. Zangi, R., Hagen, M., Berne, B.J.: Effect of ions on the hydrophobic interaction between two plates. J. Am. Chem. Soc. **129**(15), 4678 (2007). https://doi.org/10.1021/ja068305m
21. Tanford, C., Buzzell, J.G., Rands, D.G., Swanson, S.A.: The reversible expansion of bovine serum albumin in acid solutions. J. Am. Chem. Soc. **77**(24), 6421 (1955). https://doi.org/10.1021/ja01629a003
22. Keil, B.: Specificity of Proteolysis. Springer-Verlag, Berlin (1992)
23. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., Bairoch, A.: Protein identification and analysis tools on the ExPASy server. In: Walker, J.M. (ed.) The Proteomics Protocols Handbook, pp. 571–607. Humana Press, New Jersey (2005)
24. Beaven, G.H., Holiday, E.R.: Ultraviolet absorption spectra of proteins and amino acids. Adv. Protein Chem. **7**, 319 (1952). https://doi.org/10.1016/S0065-3233(08)60022-4
25. Hayashi, K., Imoto, T., Funatsu, M.: Proteolysis of lysozyme-substrate complex. J. Fac. Agric. Kyushu Univ. **15**, 387 (1969)
26. Dalgalarrondo, M., Dufour, E., Chobert, J.M., Bertrand-Harb, C., Haertle, T.: Proteolysis of β-lactoglobulin and β-casein by pepsin in ethanolic media. Int. Dairy J. **5**(1), 1 (1995)
27. Tam, J.J., Whitaker, J.R.: Rates and extents of hydrolysis of several caseins by pepsin, rennin, endothia parasitica protease and mucor pusillus protease. J. Dairy Sci. **55**(11), 1523 (1972). https://doi.org/10.3168/jds.S0022-0302(72)85714-X
28. Reddy, I.M., Kella, N.K., Kinsella, J.E.: Structural and conformational basis of the resistance of β-lactoglobulin to pectic and chymotryptic digestion. J. Agric. Food Chem. **36**(4), 737 (1988). https://doi.org/10.1021/jf00082a015
29. Schiebener, P., Straub, J., Levelt Sengers, J.M.H., Gallagher, J.S.: Refractive index of water and steam as function of wavelength, temperature and density. J. Phys. Chem. Ref. Data **19**(3), 677 (1990). https://doi.org/10.1063/1.555859
30. Harvey, A.H., Kaplan, S.G., Burnett, J.H.: Effect of dissolved air on the density and refractive index of water. Int. J. Thermophys. **26**(5), 1495 (2005). https://doi.org/10.1007/s10765-005-8099-0
31. Svergun, D.I., Richard, S., Koch, M.H.J., Sayers, Z., Kuprin, S., Zaccai, G.: Protein hydration in solution: Experimental observation by x-ray and neutron scattering. Proc. Natl. Acad. Sci. U.S.A. **95**(5), 2267 (1998). https://doi.org/10.2307/44039
32. Sushko, O., Dubrovka, R., Donnan, R.S.: Sub-terahertz spectroscopy reveals that proteins influence the properties of water at greater distances than previously detected. J. Chem. Phys. **142**(5), 055101 (2015)