Investigation of albumin-fullerenol interaction using laser correlation spectroscopy: the algorithm

Elina Nepomnyashchaya[*]**, Ekaterina Savchenko, Elena Velichko, and Evgenij Aksenov**

Peter the Great Saint-Petersburg Polytechnic University, 29 Polytechnicheskaya st., St.Petersburg 195251, Russia

*e-mail: elina.nep@gmail.com

Abstract. A new algorithm for the solution of the inverse problem of laser correlation spectroscopy is proposed. The algorithm allows the analysis of nanoparticle sizes in polydisperse solutions. Experimental results of the albumin-fullerenol interaction study demonstrating the efficiency of our approach are presented. © 2016 Journal of Biomedical Photonics & Engineering.

References

1. N. N. Pshenkina, "Serum albumin: structure and carrying function", Pharmacology 12, 1067–1091 (2011) [in Russian].
2. M. A. Orlova, T. P. Trofimova, A. P. Orlov, O. A. Shatalov, A. A. Svistunov, Yu. K. Napolov, and V. P. Chekhovin, "Fullerene derivatives as modulators for the cell proliferation and apoptosis processes", Oncohematology 4, 7–10 (2012) [in Russian].
3. L. B. Piotrovsky, M. Yu. Eropkin, E. M. Eropkina, M. A. Dumpis, and O. I. Kiselev, "Mechanisms of biologic action of fullerenes — dependence on aggregate state," Psychopharmacol. Biol. Narcol. 7(2), 1548–1554 (2007) [in Russian].
4. I. I. Grynyuk, S. V. Prylutska, N. S. Slobodyanik, O. Yu. Chunikhin, and O. P. Matyshevska, "The aggregate state of C60-fullerene in various media," Biotechnologia Acta 6(6), 71-76 (2013).
5. V. A. Shipelin, E. N. Trushina, L. I. Avren'eva, S. Kh. Soto, S. Yu. Batishcheva, G. Yu. Mal'tsev, I. V. Gmoshinski, S. A. Khotimchenko, and V. A. Tutel'yan, "Toxicological and Sanitary Characteristics of Fullerenol (Hydroxylated Fullerene C_{60}) in 28 Day In Vivo Experiment," Nanotechnologies in Russia 8(11-12), 799–809 (2013).
6. K. L. Linegar, Applications of Dynamic Light Scattering in Chemical Engineering: Polymers, Proteins, and Liquid Crystals, Thesis submitted to the Faculty of the Graduate School of the University of Maryland (2008).
7. J. Stetefeld, S. A. McKenna, and T. R. Patel, "Dynamic light scattering: a practical guide and applications in biomedical sciences," Biophys Rev. 8(4), 409–427 (2016).
8. R. Finsy, P. Groen, L. Deriemaeker, E. Geladé, and J. Joosten, "Data analysis of multiangle photon correlation measurements without and with prior knowledge," Part. Part. Syst. Charact. 9(1-4), 237–251 (1992).
9. R. Finsy, P. Groen, L. Deriemaeker, and M. van Laethem, "Singular value analysis and reconstruction of photon correlation data equidistant in time", J. Chem. Phys. 91(12), 7374–7383 (1989).
10. R. Finsy, N. Jaeger, R. Sneyers, and E. Gelade, "Particle sizing by photon correlation spectroscopy. part III: Mono and bimodal distributions and data analysis," Syst. Charact. 9(1-4), 125–137 (1992).
11. M. Shuai, J. Shen, J. C. Thomas, X. Zhu, W. Liu, and X. Sun, "Improved inversion procedure for particle size distribution determination by photon correlation spectroscopy," J. Appl. Opt. 51(25), 6220–6 (2012).
12. E. K. Nepomniashchaia, E. N. Velichko, and E. T. Aksenov, "Solution of Inverse Problem of Laser Correlation Spectroscopy by Regularization Method," Humanities & Science University Journal 13, 13–23 (2015).

1   Introduction

Albumin is the main blood protein, responsible for the transport of organic and inorganic substances (ligands), biologically active agents, fatty acids, ions, as well as medicinal preparations and their metabolites [1].

The binding function of serum albumin is caused by the structural features of this protein, the presence of active polar and hydrophobic areas on its surface. The albumin molecule, being not large in size, has two binding centres that determine the variety of its binding capacities. Under different environment conditions the protein can change its structure and lose the capability of binding with other molecules. For example, under strong variation of the solution pH or temperature the protein denaturation occurs, i.e., the globule unfolding and the transition from the tertiary structure to the primary one.

In relation with the active investigation of the protein transport function, the search is carried out for possible test molecules, the probes for activating this function. The probes are sensitive indicators of conformation transitions and other restructuring in the protein molecule. The methods using test probes are based on recording the variations of physical and chemical parameters, reflecting the process of interaction of the specific or test ligand with albumin. It is assumed that the fullerenol nanoparticles can be used for this purpose.

Fullerenol is one of the carbon modifications. It is a firm, hydrophilic, supramolecular complex, consisting of a fullerene molecule (in our case C_{60}), and the hydrate shell, enclosing the molecule [2]. Tis fullerene compound $C_{60}(OH)_{24}$ is well solvable in water, which makes it possible to use it in the study of protein solutions [3]. In the aqueous phase the fullerenol can be presented not only by free molecules, but also by multimolecular clusters, comprising 10--50 molecules, which can affect its biochemical properties.

To date the fullerenol is a promising object for nanobiotechnologies [4], particularly in the investigation of the albumin properties directly in the organism. However, the effect of fullerenols introduced into the organism on the human health is insufficiently studied, and further research is necessary to quantify these effects for complete risk estimation [5]. The characterisation of fullerenol toxic properties is of great importance, since it is expected to be used not only in laboratory diagnostics, but also in medicine, e.g., for efficient delivery of drugs, as nutrition and cosmetic supplements [5].

Thus, the study of albumin-fullerenol binding has two-fold purpose: the determination of albumin capacity of binding and transporting the substances under different conditions of environment, and the determination of the toxicity of the fullerenol itself for humans. In our paper, the solutions of albumin with fullerenol were studied by means of laser correlation spectroscopy. The developed algorithm of the experimental data analysis, based on the Tikhonov regularisation method, provides the required high accuracy of the results.

2   Experimental setup

The method of laser correlation spectroscopy is based on recording and analysing the light field scattered by the particles performing Brownian motion in a solution. The schematic diagram of measurements is presented in Fig. 1. As a source of radiation, we used the single-mode laser module with the power of continuous-wave radiation to 2.5 mW in the spectral range near 650 nm. The parasitic aureole was eliminated using the screen with a hole. The radiation is focused by the aspheric short-focus lens built in the module that allows the diameter of the beam to be 50 μm in the focus of the lens. The calculated caustic length in the solution amounts to 5 mm. The power supply of the radiation source was organised from storage cell, which provides the stability of power. The radiation scattered at the angle of 90° is detected using the photomultiplier. The diaphragm and the aperture of the single-mode optical fibre determine the angle of view for the scattered radiation detection. The length of the scattering volume, calculated for the scattering angle 90°, amounts to 4.8 mm. The signal for the photomultiplier is digitised by the ADC board at the frequency 50 MHz and was processed with a computer.
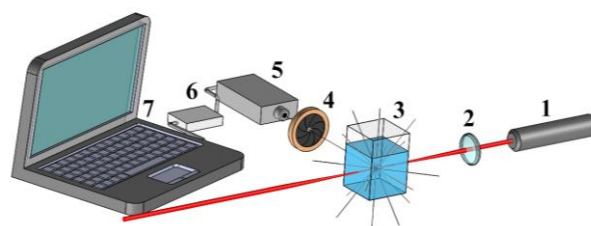

Fig. 1 Schematic diagram of laser correlation spectrometer. 1 – laser, 2 – lens, 3 – cuvette with the studied solution, 4 – diaphragm, 5 – photomultiplier, 6 – ADC, 7 – computer.

The recorded signal is a pseudorandom time dependence of the scattered radiation intensity. The spectrum of this signal is determined by the coefficients of diffusion of particles in the solution. To find the characteristic frequencies one can use the Fourier transform. However, due to the huge number of scatterers and random components in the signal, it is rather hard to select the characteristic frequencies from the spectrum. To reveal the regularities in the signal, the time-domain representation of the signal spectrum is used rather than the frequency-domain one, i.e., the correlation function of the signal is calculated.

3 Mathematical processing

According to the dynamic light scattering theory, the autocorrelation function can be described by the following expression [6]:

$$\left| g^{(1)}(\tau) \right| = \int_{0}^{\infty} F(\Gamma) e^{-\Gamma \tau}\, d\Gamma, \qquad (1)$$

where Γ is the diffuse spectrum broadening, $F(\Gamma)$ is the contribution from the particles of one size into the total intensity of scattering. The diffuse broadening is related to the diffusion coefficient D as:

$$\Gamma = Dq^2, \qquad (2)$$

where $q = (4\pi n/\lambda)\sin(\theta/2)$ is the scattering vector, n is the refractive index of the medium, λ s the wavelength, θ is the scattering detection angle.

Making use of the Stokes-Einstein formula

$$D = k_b T / 6\pi\eta R \qquad (3)$$

(η is the medium viscosity, k_b is the Boltzmann constant, T is the temperature, R is the radius), it is possible to calculate the hydrodynamic radius of the studies particles.

Thus, the solution of Eq. (1) is the main problem in the mathematical processing of the light scattering signal. This equation is an integral Fredholm equation of the first kind, and due to the presence of random noise in the experimental data, its solution belongs to the class of ill-posed problems. This means that the presence of even small error in the initial data will cause the accumulation of error in the process of common solution, and the result may strongly differ from the real one.

To date a variety of methods for solving the equations like Eq. (1) exists. Up to now, the most widespread in the commercial production is the method of cumulants. This method is rather simple in implementation, fast and convenient in application, thanks to which it is included into the International Standard ISO 13321. However, in the absence of a priori information this method allows the determination only of some mean values of the diffusion coefficients and their moments. Moreover, the method of cumulants provides reliable particle size only for monomodal distributions and for polydisperse solutions often yields wrong results [DLS: practical guide 2016]. In our case, the polydispersity of the mixture plays a crucial role in the study of the albumin binding properties, which makes it impossible to use simple monomodal methods.

To solve the inverse problem in the case of polydisperse solutions it is possible to use the methods of nonlinear optimisation (the Levenberg-Marquardt method). Unfortunately, they require rather exact a priori information about the shape and the number of

components in the distribution, and, therefore, are not always convenient in the processing of real signals.

The methods that do not require the knowledge of the distribution shape (with reasonable physical requirement of non-negativity of results) include the NNLS (Non-Negative Least) method and its different modifications based on regularisation, e.g., CONTIN (Constrained Regularization Method for Inverting Data) [DLS: practical guide 2016], as well as the SVR (Singular Value Analysis and Reconstruction) method and MEM (Maximum Entropy Method). In a number of papers, the results of applying the above methods to different polydisperse and monodisperse systems are compared [8 − 10]. As a rule, SVR yields higher precision of results than the other above methods do. However, the existing requirement of equidistant time intervals [10] between the points in the solution does not allow the study of solutions with large and small particles present simultaneously with similar accuracy.

The methods of regularisation and the entropy maximum allow the results close in accuracy, the regularisation methods operating almost twice as fast as the entropy maximum method [9]. At the same time, none of these methods yields "non-blurred" distributions in the case when many closely spaces peaks are present in the solution.

In this connection, the necessity appeared to develop an original algorithm for processing the data of laser correlation spectroscopy. As a base, we took the regularisation methods that possess greater stability against the errors and potentially can provide more information than the entropy maximum method [9].

Among the variety of regularisation methods, the CONTIN method was most popular until recent time. Unfortunately, it is very exacting with respect to the choice of regularisation parameter and, as already mentioned above, does not allow the resolution of narrow peaks. In this connection, we decided to use a more advanced modification, the Tikhonov regularisation, which was rarely used before because of complexity of algorithmisation. In recent years different modifications of Tikhonov regularisation were described in scientific literature, however, when the number of components in the signal is large, the accuracy of the solutions is low [11]. In spite of this fact, the potentialities of the method are undoubtful.

The numerical solution of Eq. (1) is presented as a set of N points. Thus, we arrive a system of N equations, which in the matrix form can be written as

$$A\boldsymbol{f} = \widetilde{\boldsymbol{g}}, \qquad (4)$$

where $A_{ij} = e^{-\Gamma_j \tau_i}$, $\boldsymbol{f} = F(\Gamma_j)$, $\widetilde{\boldsymbol{g}} = g^{(1)}(\tau)$.

When using the regularisation, certain a priori conditions are imposed on the solution of this system of equations, depending on the problem under solution. The commonly accepted conditions are that the solutions should be non-negative (bounded within the compact set $M \geq 0$), as well as the condition of smoothness, i.e., the absence of spikes.

To find the approximate solution of the system of equations (4) stable with respect to small variations of the right-hand side, in the Tikhonov methods the system of linear algebraic equations (SLAE) (4) is replaced with the minimisation problem with the addition of regularising term:

$$||Af - \widetilde{g}||^2 + \alpha \Omega(f) \to min, \qquad (5)$$

where $\alpha > 0$ is the smoothing parameter, $\Omega(f)$ is the stabilising functional, individually chosen for each problem. In order to get smooth solutions in laser correlation spectroscopy, the stabilising functional is chosen to reduce the jumps of zero derivative, i.e., $\Omega(f) = ||f||^2$. This is because the size distribution of particles as a rule is Gaussian. Such minimisation stabilises the solution of the system, improving its conditionality and increasing the consistency of real and desired solution.

The regularisation parameter α is chosen basing on the input data: too large α leads to the appearance of extra-smoothed solutions, while too small α makes the problem unstable. In our algorithm the parameter α was chosen to be 1% of the maximal diagonal element of the matrix A. After the first solution of the system (5) the discrepancy $||Af - \widetilde{g}||^2$ was calculated, the parameter α was reduced by 90% and the system was solved again. After the second iteration, the discrepancy was calculated again and compared with that obtained at the previous iteration. If the difference was greater than 10%, the parameter α was again reduced by 90%, and the next iteration was performed. In the proposed algorithm due to the comparison of discrepancies, it is not necessary to specify the noise level in the experiment, which is often not exactly known.

The system of equations (5) is solved using the upgraded Gauss method, in which the eigenvalues of the matrix are shifted towards greater values at the expense of introducing the regularising term, which makes the solution more stable with respect to noises, inevitably present in real experiments.

After completing the solution cycle and getting the final distribution f, the check for negative solution components is performed. The programme implies the possibility of using two different methods of excluding the negative solutions. In the first method if $f_j < 0$ we set $f_j = 0$ for all extreme values j in the present Gaussian and exclude these points from further calculations. Then we return to the specification of the initial values of α. In the second method f_j with the values $> 60\%$ of the minimal value are excluded from the calculations, rather than the extreme points. This calculation operates faster, but yields less precise results, therefore, it is more suitable for preliminary analysis of the size distribution of particles. The recalculation of (5) is continued until the elimination of all negative components.

Setting rather large α could lead to smoothing and broadening of narrow distributions. However, besides the elimination of negative values, the process of excluding positive values is introduced into the

algorithm. One at a time, the extreme points are excluded from the Gaussian until the given width is achieved, which allows the calculation of sizes in strongly polydisperse mixtures with the error ± 0.5 nm, P = 95%.

Finally, the used algorithm can be formalised as follows:

1. setting the initial (large enough) value of α;
2. solution of the system of equations (5) and calculation of the solution f;
3. calculation of discrepancy $||Af - \widetilde{g}||^2$ (after the first iteration reducing α by 90% and return to 1);
4. comparison: $||Af - \widetilde{g}||_i^2 < 0,1 \cdot ||Af - \widetilde{g}||_{i-1}^2$, if yes, then reduce α by 90% and go to 1, if not, then go to 5;
5. check for the presence of negative components in the solution f: if $f_i < 0$ then set $f_j = 0$ and return to 1. The corresponding component is excluded from further calculations;
6. if there are no components $f_i < 0$, then check the number of distribution points n > N (specified before starting the algorithm), if yes, then set $f_{min} = 0$ and return to 1, if no, then the end of calculation.

Before one begins the calculation, in order to accelerate the process one has to specify the domain of the expected solutions and the desirable accuracy. The search in the entire domain of admissible solutions is also possible, but requires much more time because of the cubic dependence of the time consumption upon the number of points in the Gaussian method. It is also necessary to specify the method of excluding negative solution components, as described above, and the number of points left in the peak (more points are taken in the presence of weak components). For better accuracy, the possibility is provided to perform the averaging of correlation, obtained from the experimental data.

The precision of experimental data processing was checked in model solutions, the mixtures of microspheres having different diameters. The results have shown that for single-component solutions the proposed algorithm does not yield in precision to the known commercial analogues with built-in processing (the spectrometers Zetasizer Nano ZS and Photocor) [12]. However, for three and more different sizes of particles the precision of the developed algorithm is by 30-60% higher. Thus, the programme of processing the results of laser correlation spectroscopy based on Tikhonov regularisation, described in the present paper, allows the study of polydisperse biological solutions.

4 Results

In the first series of experiments, we performed control measurements of the albumin molecule size in water and saline. The albumin concentration in the initial sample amounts to 100 g/l, which is close to physiological concentrations of this protein in the human blood plasma.

Under different diseases, the pH of the organism cam change both to the acidic and the alkalotic side. Therefore, we performed the study of the dependence of the protein hydrodynamic radius on the pH of the solution. The results are presented in Fig. 2. It is seen that at pH = 4.5 - 5 (the albumin isoelectric point) the hydrodynamic radius of the molecule is maximal, i.e., its spatial structure changes. The albumin molecule actively reacts to the smallest changes of the environment, including the pH variation. In particular, it can shrink, stretch, untwist, twist its spirals, or prepare the sites for a contact with ligands.
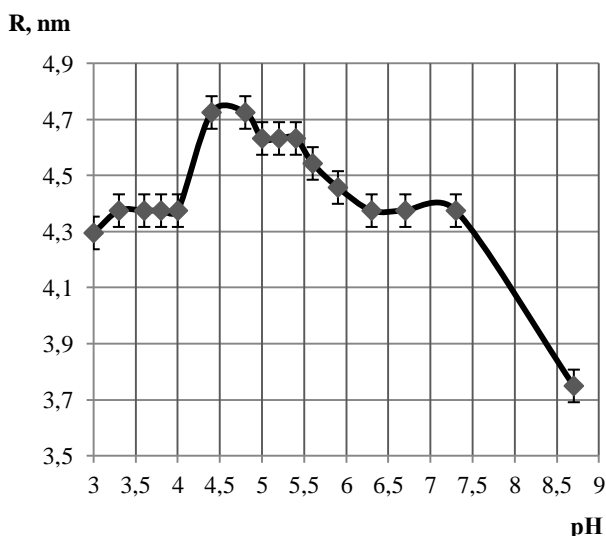

Fig. 2 Dependence of the size of a single albumin molecule on the pH of the solution.
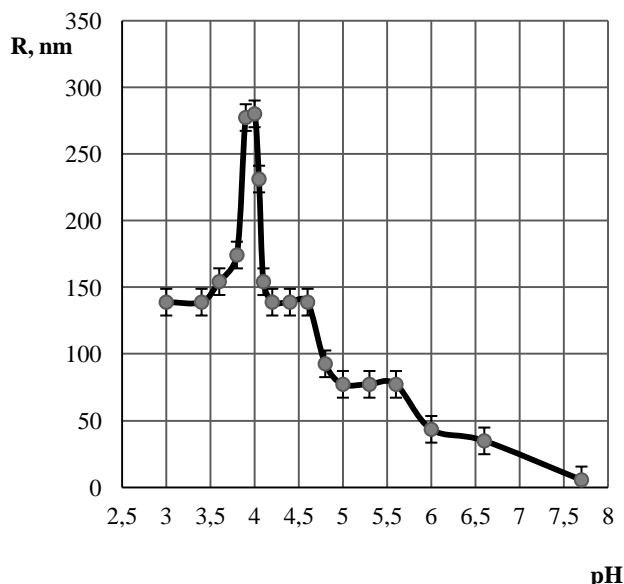

Fig. 3 Conglomerate radius versus pH for the binding of fullerenol and albumin.

In the second series of experiments, we studied the binding capacity of albumin with fullerenol under the variation of the solution pH. It was found that the optimal concentration ratio for binding two molecules is albumin : fullerenol=1:3. With growing acidity of the solution, the characteristic dimensions of the formed clusters increased. The obtained dependence of the albumin+fullerenol hydrodynamic radius on the solution pH is presented in Fig. 3.

From the plot it is seen that the maximal size of conglomerates is observed at pH = 4, which differs from the albumin isoelectric point. In the course of binding fullerenol, bilirubin, fatty acids, etc., the globule structure can change. In this case the conformation restructuring is possible both in the entire molecule and in the binding site at other pH values, too. As a result, the number of binding centres can change, and the binding of different number of molecules is possible. Hence, the obtained results may evidence in favour of both the swelling of the protein molecule itself at pH = 4, and the strong interaction with the fullerenol particles. Moreover, the fullerenol itself can lose stability in acidic solutions and form both own aggregates and the multilayer structures on the albumin surface.

Under the change of pH towards strong acidity (< 3 pH), the cluster size is decreased, which is due to the albumin denaturation. In this case, almost no aggregation is observed due to the global restructuring of the molecule and the reduction of the number of binding centres and their activity.

At pH = 7.5 the albumin molecule has its characteristic size R = 6 nm and binds the individual fullerenol particles with R = 2 nm without the formation of double layers, which is confirmed by the observation that the maximal size of the particles, found in the solution, amounted to 8 − 10 nm.

5 Conclusions

We proposed and implemented as a programme the algorithm for processing the information signals, obtained suing the method of laser correlation spectroscopy, based on Tikhonov regularisation. The algorithm allows the study of polydisperse solutions with high accuracy. The result of experimental approbation confirmed the expected accuracy of the calculations using the proposed algorithm in the analysis of biomolecular polydisperse solutions. By the example of albumin interacting with fullerenol, the specific features of the interaction of two molecular solutions under the action of different factors were shown. The experimental results have demonstrated that under normal conditions (physiological temperature and acidity of solutions) the interaction of albumin with fullerenol leads to the formation of bonds between the individual fullerenol molecules and albumin. The minor change of pH towards the greater acidity causes strong clustering of molecules. It is possible to suggest that the albumin becomes surrounded by the fullerenol particles and loses the binding capacity to other substances. Besides, a trend to the change of conglomerate sizes under the pH variation is observed, which allows

conclusions about the restructuring of albumin and the formation of multiple binding centres. Thus, fullerenol can be used as a probing substance for the analysis of structural changes in proteins. The observed effects require further studies aimed at specifying the mechanisms of interaction between albumin and fullerenol and assessing the possibility of using fullerenol as transport agent.