Article

Changes of conformation in albumin protein with temperature

Piotr Weber [1,‡] 0000-0002-9405-5461, Piotr Bełdowski [2,‡] 0000-0002-7505-6063, Krzysztof Domino [3,‡*] 0000-0001-7386-5441, and Adam Gadomski [2,‡] 0000-0002-8201-1736

1 Atomic and Optical Physics Division, Department of Atomic, Molecular and Optical Physics, Faculty of Applied Physics and Mathematics, Gdańsk University of Technology, ul. G. Narutowicza 11/12, 80-233 Gdańsk, Poland; piotr.weber@pg.edu.pl

2 Institute of Mathematics and Physics, UTP University of Science and Technology, Kaliskiego 7, 85-796 Bydgoszcz, Poland

3 Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, Bałtycka 5, 44-100 Gliwice; kdomino@iitis.pl

* Correspondence: Affiliation 3

‡ These authors contributed equally to this work.

Abstract: We study a conformation of an albumin protein in the temperature range of $300K$-$315K$, i.e. in the physiological range of temperature. Using simulations we calculate values of two backbone angles, that carry most of information about positioning of the protein chain in a conformational space. Given these, we calculate energy components of such protein. Further, using the Flory theory we determine the temperature in which investigated albumin chain model is closest to the free joined chain model. Near the Flory temperature, we study energy components and the conformational entropy, both derived from two angles that reflect most of the chain dynamics in a conformational space. We show that the conformational entropy is an oscillating function of time in considered range of temperature. Our finding is that, the only regular oscillation pattern appears near the Flory temperature.

Keywords: conformation of protein; albumin protein; non-gaussian chain.

1. Introduction

Dynamics of proteins in solution can be modelled by a complex physical system [1] yielding many interesting effects. Biopolymers such as proteins consist of monomers (amino acids) connected linearly. Thus, a chain seams to be the most natural model of such polymers. However, depending on the properties of the polymer molecule, modelling it as a chain may give a better or worse outcome [2].The simplest approach to the polymer is the freely jointed chain, where each monomer moves independently. The excluded volume chain effect, which prevents to overlap polymer's segments is one of the reasons why physical polymer chain differs from an ideal model [3]. Diffusive dynamics of biopolymers is discussed further in [4] and references within. For a given combination of the polymer and the solvent, excluded volume varies with temperature and can be zero at the specific temperature, named the Flory temperature [5]. In other words, at the specific Flory temperature, the polymer chain becomes nearly ideal [6].

Given such polymer dynamics, our work concentrates on an analysis of the specific protein molecule and an investigation of the conformation of such protein referring to its Flory temperature. This specific protein is an albumin (shown in Figure 1), one of the essential components of the synovial fluid (SF). The concentration of the albumin is approximately 45% of all proteins present in the synovial fluid. The synovial fluid is a complex fluid which serves as a lubricant of natural joints systems [7]. From the biological point of view, this fluid is a mixture of water, proteins, lipids, and other biopolymers [8]. In general, proteins are present in the synovial fluid at concentrations inversely proportional to their molecular size. The critical protein in this fluid (of nonlinear viscoelastic rheopexy

involving character [11]) is the albumin since it plays an essential role in the process of articular cartilage lubrication [9]. As an active component of a mechanism of lubrication, albumin is affected by an increase of temperature in synovial fluid originating from friction. The temperature inside the articular cartilage is in the range $300K - 315K$ [10], and changes occurring inside the fluid can affect an outcome of the albumin impact on lubrication [11].

Let us move to the technical introduction of an albumin dynamics. For this reason consider chain that was formed by $N+1$ monomers, which have positions $\{\vec{r}_0, \vec{r}_1, \ldots, \vec{r}_N\}$. We note by $\vec{\tau}_j = \vec{r}_j - \vec{r}_{j-1}$ a vector of a distance between two neighbouring monomers. While considering the albumin we have $N+1 = 579$ amino-acids. For the length of the distance vector $\vec{\tau}_j$ we take the distance between two α-carbons in two consecutive amino-acids, hence we have $|\vec{\tau}_j| = \tau = 3.8$ [Å] [12]. Using such notation, we can define the end-to-end vector $\vec{R}$ by the following formula:

$$\vec{R} = \vec{r}_N - \vec{r}_0 = \sum_{j=1}^{N} \vec{\tau}_j. \tag{1}$$

In the case of an ideal chain, distance vectors are completely independent of each other. This property is expressed as a lack of correlation between any two different bonds, that is [13]:

$$\langle \vec{\tau}_i \cdot \vec{\tau}_j \rangle = \tau^2 \delta_{ij}, \tag{2}$$

where δ_{ij} is the Kronecker delta. Equations 1 and 2 imply that:

$$\langle \vec{R} \cdot \vec{R} \rangle = \langle R^2 \rangle = \sum_{i=1}^{N} \sum_{j=1}^{N} \langle \vec{\tau}_i \cdot \vec{\tau}_j \rangle = N\tau^2, \tag{3}$$

where R is the length of the vector $\vec{R}$. The value of R is measured by calculating the root mean square end-to-end distance of a polymer. The R'th value is assumed to depend on the polymer's length N in the following manner [14]:

$$R \sim \tau N^{\mu}, \tag{4}$$

where μ is called the size exponent. If a chain is the freely joined chain it appears that:

$$R \sim \tau N^{1/2}. \tag{5}$$

In the Flory temperature, the real polymer fulfils this relation relatively closely in comparison with other temperatures. Relation 5 brings the idea to consider a dynamics of the albumin protein in some range of temperature and looking for the temperature in which size exponent μ is closest to 0.5. The logarithm of Equation 4 yields:

$$\mu \simeq \frac{\log R - \log \tau}{\log N}. \tag{6}$$

This equation can be used to calculate the Flory temperature. In Equation 4 we use a $\sim$ symbol to emphasize that we use $\langle R^2 \rangle$ to obtain R. In the case discussed here both the parameter τ an N (the number of amino acids) are constant. Importantly, using molecular dynamics simulations one can obtain the root mean square end-to-end distance for various temperature.

For simulations to be tractable, further simplification leads to the chain that can be described by two backbone angles. One of them is an angle between three α-carbon occurring after each other and signed as ϕ. The second angle, noted as ψ, is a dihedral angle [15].

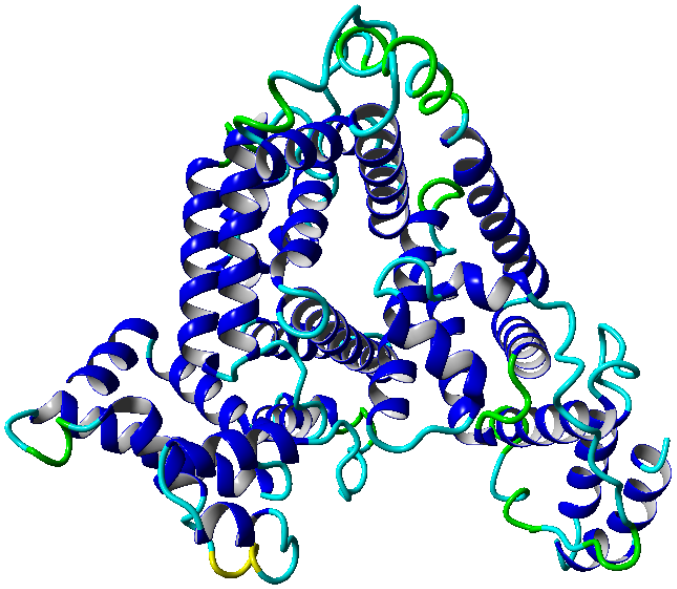

Figure 1. Structure of the albumin protein in a ribbon like form. The colours on the molecular surface indicate the the secondary structure: blue - $\alpha-$ helix, green - turn, yellow - 3-10 helix, cyan - coil.

2. Results

To estimate the Flory temperature, we collect outcomes from molecular dynamics simulation in the temperature range of $300K - 315K$. We use the temperature step of $3K$. The motivation for choosing such range is to cover the physiological range of temperatures. In the albumin structure we can distinguish 22 α-helices, which are presented in Figure 1. In Figure 2 we present a length of the α - helices as a function of time for temperature $300K$. We can see that the length of α - helices only fluctuates around some their mean value. For other values of temperature we have obtained very similar results. Our outcomes showed that during simulations these helices almost did not change their lengths.

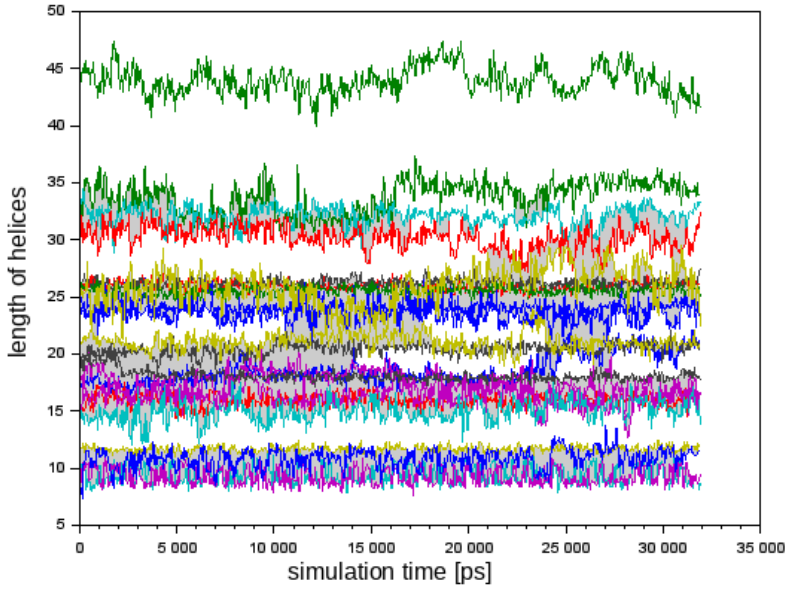

Figure 2. Lengths of helices vs the simulation time (in ps) for temperature $300K$. Each colour represents one α - helis.

For each value of temperature, we have calculated a polymer's end-to-end distance vector as a function of time. Next we have obtained the root mean square end-to-end distance of this quantity, which values are presented in Table 1:

Table 1. Root mean square length of the end-to-end vector for $300K$ - $315K$.

Parameter	$300K$	$303K$	$306K$	$309K$	$312K$	$315K$
$\langle R^2 \rangle^{1/2}$ [Å]	51.55	48.95	50.33	51.91	49.91	46.84

We can see that the root mean square length of the end-to-end vector for albumin varies with temperature. Using Equation 6 we estimate values of the Flory temperature for the given protein chain. Results are presented in Table 2:

Table 2. Size exponent obtained according to Formula 6.

Parameter	$300K$	$303K$	$306K$	$309K$	$312K$	$315K$
Size exponent (μ)	0.4102	0.4021	0.4065	0.4113	0.4051	0.3952

Analysing Table 2 the exponent value $\mu = 0.4113$ is nearest to $1/2$. This result also show that the biopolymer conformation expresses a local sub-diffusive environment, such as described in [16]. It indicates that the neighbourhood of $T = 309K$ can be considered as a region, where the Flory temperature can appear for this protein. Because the temperature grid has a relatively large step, we used the non-linear regression to estimate exponent values for immediate temperatures, see Figure 3.

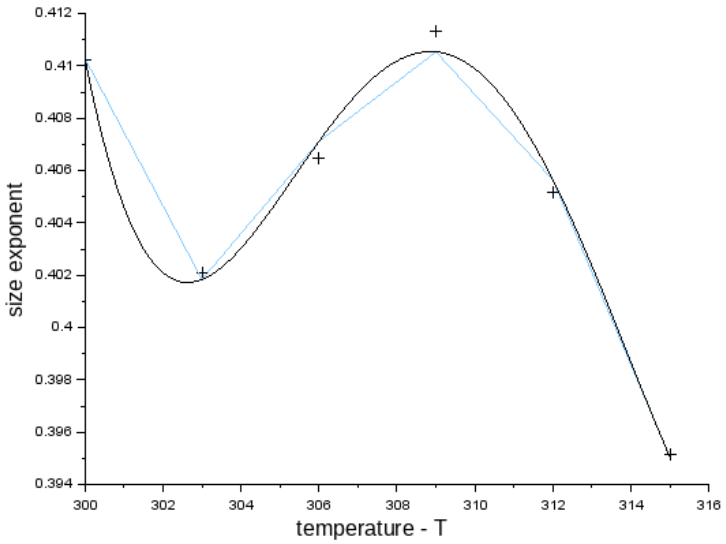

Figure 3. Values of μ for different temperatures and the fit of the polynomial of forth degree. Fit performed by the least square error procedure. Crosses represent μ values from Table 2 and regression polynomial is represented by solid curve. Regression line indicate that maximum is little bit smaller than $T = 309K$.

Further using simulations we can determine the angle ϕ and the angle ψ, see Introduction. Next we can determine energies that depends on these angles by means of Formula 7. In Figure 4, we present a logarithm of the component of energy tied directly to the angles ϕ (the angle energy component) versus simulation time.

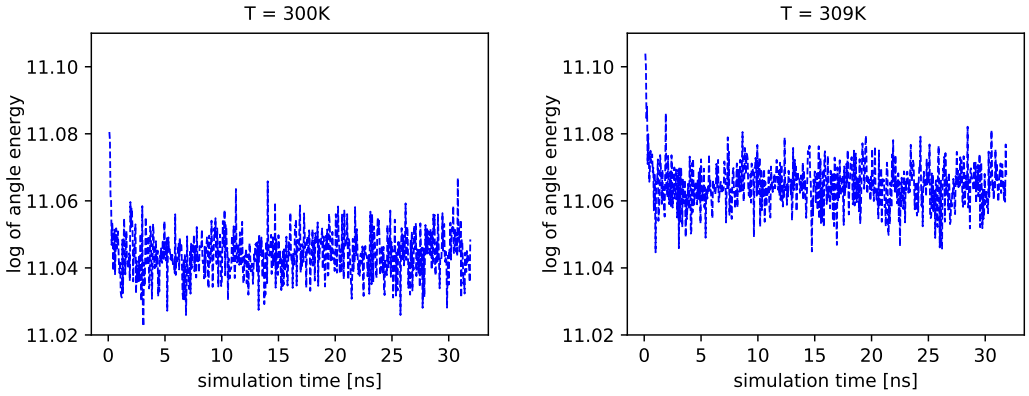

Figure 4. Logarithms of angle energies vs. the simulation time, exemplary outcome of $T = 300$ and $T = 309$.

The logarithm of the angle energy component fluctuates around the mean value. Such dynamics is similar for all temperatures, but the mean energy rises slightly with temperature.

Different outcomes comes from an analyse of the component of the energy connected with the dihedral angles ψ. See Figure 5 (blue lines) in double logarithmic scale, note linear regression lines.

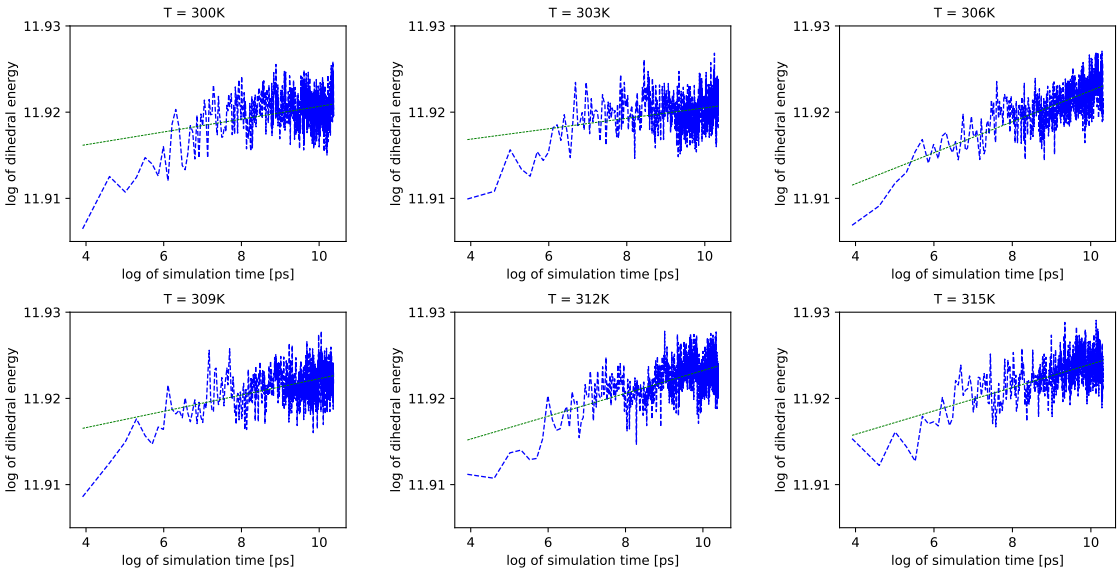

Figure 5. Dihedral energies for each temperature vs. simulation time. We use the double logarithmic scale. Green lines represent linear regression.

In Table 3 we present regression parameters from Figure 5. Value of the slope parameter a is biggest for $T = 306K$.

Table 3. Values of regression coefficients obtained according to formula $y = ax + b$ for dihedral energy in double logarithmic scale presented in the Figure 5.

Parameter	$300K$	$303K$	$306K$	$309K$	$312K$	$315K$
values of $a[\times 10^{-4}]$	7.4	6.0	18.0	9.5	13.2	13.6
standard error of $a[\times 10^{-4}]$	0.9	0.9	0.8	0.8	0.9	0.8
Standard deviation of the residual $[\times 10^{-3}]$	2.115	2.026	1.842	1.932	2.031	1.862

Dynamics of protein is governed mainly by hydrophobic and hydrophilic interactions [17]. Therefore it would be useful to study other components of energy, such as the Coulomb component

and the van der Waals component, see Equation 7 and following it discussion. The first oscillates around the almost constant trend is in every temperature. However, the mean value of these oscillations increase with temperature. The opposite situation appears for the van der Waals component. Similar to Coulomb component, for each selected temperature it follows the almost constant trend. However, when temperature increases the mean value of such oscillations decreases.

The Coulomb component and the van der Waals component obeys simple dynamics, hence we can move to further discussion of more interesting dihedral and angular components of the energy. Simulations of an albumin dynamics can be used to obtain frequency distributions of angles ϕ and ψ. Exemplary 2D histogram of such distribution is presented in Figure 6. We can see that most data are concentrated in the small area of space of angles. Similar behaviour can be observed for other temperatures and simulation times. It is because the structure of the investigated protein is quite rigid, and there appears only a little angular movement. Nevertheless, we can observe variations of the conformational entropy estimated form such histogram, see Figure 7.

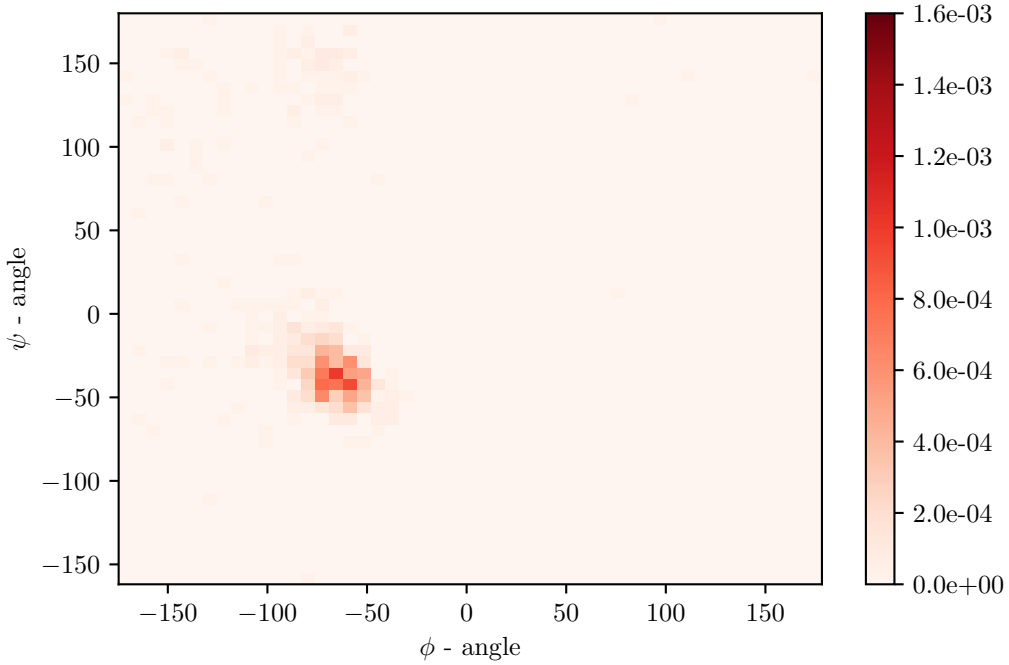

Figure 6. Numerically obtained distribution of angles ϕ and ψ for $T = 300K$ and $14ns$ of simulation.

According to formulas connecting probability distribution of angles and entropy [14] we can follow changes of entropy in time and temperature. Interesting properties of this entropy appear around Flory temperature. We can see that conformational entropy changes in time in an oscillatory way. Interestingly at $T = 309K$ that seams to be the Flory temperature, the oscillatory behaviour seams to be most regular.

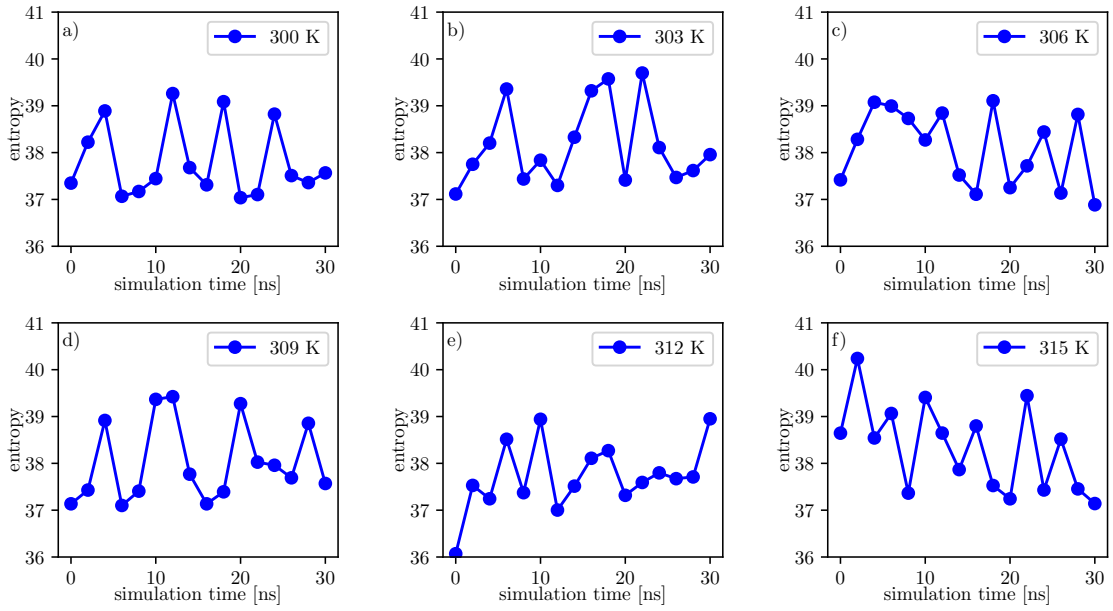

Figure 7. Value of conformational entropy for various temperatures and 0–30 ns of simulation time.

3. Discussion

Lubrication of synovial fluid is a complex mechanism originating from synergistic interactions between its components. Due to this complexity, understanding of the impact of all the molecular interaction is a great challenge [7]. Moreover, joints undergo several changes in lubrication regimes throughout their functioning, which change the physical and chemical conditions of the operating fluid. One of such examples is the increase in temperature resulting from friction. The complex sub-diffusive dynamics of the synovial fluid [16] has a positive impact on the lubrication. Thus, here we want to explore changes occurring in one of the essential components of the fluid - albumin and determine how the increase of temperature influences it.

From the general point of view, the protein of choice behaves stable regardless of the temperature. Figure 2 shows that lengths of the helices are very stable throughout the simulation. The angle energy of the protein fluctuates similarly regardless of the temperature, see Figure 4, however, its mean value rises monotonically with temperature. On the other hand, the size exponent μ fluctuates with temperature, see Figure 3 reaching the maximum of $\mu = 0.411$ at 309 K. This is the closest value to 0.5 suggesting that there may appear the Flory temperature.

Further, referring to Figure 5, the dihedral energy seams to obey (at least at some range of simulation times) the power law-like relation, although the scaling exponent is relatively small with values of order 10^{-3}. Interestingly, this scaling exponent is rather constant for lower and higher temperatures but behaves in more variable pattern for temperatures of range $306 - 309$ K. The exponent has a maximum at 306 K and a local minimum at 309 K, see Table 3. Such behaviour, for complex physical system, may suggest some crossover effect [5] in the temperature region of $306 - 309$ K.

To approve this outcome, we refer to the conformational entropy of the system. For its plot versus the simulation time, see Figure 7. At first, observe that for each temperature, the entropy value fluctuates in the same range. Nevertheless, for a temperature of 309 K, we can observe the most regular pattern of these fluctuations. We can observe this while comparing 309 K plot with neighbour 306 K and 312 K plots. Since one can regard the entropy as an information measure, such regular oscillations should comply with rather little 'information memory' of the complex physical system characteristic of an independent dynamics (each of monomers moves independently). Finally, we have rather regular fluctuations for 300 K as well. This example comply with relatively high size exponent, see Table 2. Nevertheless, 300 K is rather on the edge of the physiological range; hence, the 309 K outcome is much more interesting. Finally from Table 2 the scaling exponent value nearest to 0.5 appears at 309 K.

4. Materials and Methods

The structure of human albumin serum (code 1e78) has been downloaded from protein data bank (https://www.rcsb.org/structure/1E78) as a starting point to simulations. We use the YASARA Structure Software (Vienna, Austria) [18] to perform MD simulations. Besides this, three-site model (TIP3P) of water was used [19]. All-atom simulations were performed under the following conditions: temperature $310K$, $pH = 7.0$ and in 0.9% NaCl aqueous solution, with a time step of $2fs$. Berendsen barostat and thermostat with a relaxation time of $1fs$ were used to maintain constant temperature and pressure. All-atom molecular dynamics simulations were performed using the AMBER03 force field [20]. The AMBER03 potential function describing interactions among particles takes into account electrostatic, van der Waals, bond, bond angle, and dihedral terms:

$$E_{total} = \sum_{bonds} k_b \left(r - r_{eq}\right)^2 + \sum_{angle} k_\phi \left(\phi - \phi_{eq}\right)^2 + \sum_{dihedrals} \frac{V_n}{2}[1 + \cos\left(n\psi - \gamma\right)] + \sum_{i<j} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + \frac{q_i q_j}{\epsilon r_{ij}}\right] \quad (7)$$

where, k_b and k_ϕ are the force constants for the bond and bond angles, respectively; r and ϕ are bond length and bond angle; r_{eq} and ϕ_{eq} are the equilibrium bond length and bond angle; ψ is the dihedral angle and V_n is the corresponding force constant; the phase angle γ takes values of either 0° or 180°. The non-bonded part of the potential is represented by van der Waals (A_{ij}) and London dispersion terms (B_{ij}) and interactions between partial atomic charges (q_i and q_j). ϵ is the dielectric constant.
For each spherical angle distribution, we perform normalised histogram consisting of 50 bins. Such number a bin is a compromise between a resolution and smoothness of histograms. Taking the empirical probability p_i of data being in i-th bin, we estimate the entropy [15] [21]:

$$S = -R_0 \sum_i p_i \log(p_i), \quad (8)$$

where we use the gas constant $R_0 = 8.314 \frac{J}{K \cdot mol}$. The sum goes over the discredited Ramachandran space. Obviously, since entropy is an information measure, it depends on a bin size, hence we have here only an estimation applicable for entropy comparisons for different temperatures and simulation times, since bin sizes are constant for all simulations and temperatures.

5. Conclusions

We present several scenarios of an analysis of simulations of the albumin dynamics. The albumin is the linear protein, what make simulations and analysis straight forward. In particular, we get the end-to-end vector values, and we predict a Flory temperature of our polymer, where particular monomers dynamics are as independent of each other as possible. Our findings show that, near the Flory temperature we can observe the only regular oscillation like the behaviour of the conformational entropy. Such oscillation comply with rather 'memory less' dynamics of the complex physical system what is expected if each of monomers moves independently. Additionally, near the Flory temperature, there appears an unusual behaviour of the dihedral part of the energy of the system.

Author Contributions: Conceptualization, P.W., P.B., K.D. and A. G.; Investigation, P.W., P.B., K.D.; Writing—Original Draft Preparation, P.W. and P.B.; Writing—Review and Editing, P.B., K.D. and A.G.

Funding: This work is supported by UTP University of Science and Technology, Poland, grant BN-10/19 (PB and AG).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. M. Grimaldo, F. Roosen-Runge, F. Zhang, F. Schreiber, T. Seydel, Dynamics of proteins in solution, *Quarterly Reviews of Biophysics*, **2019**, 52, e7.
2. M. Rubinstein, R. H. Colby, *Polymer Physics Oxford University Press*, 2003.
3. P. G. De Gennes, Scaling Concepts in Polymer Physics *CORNELL UNIVERSITY PRESS*, Ithaca and London, 1979.
4. A. Gadomski, On (sub) mesoscopic scale peculiarities of diffusion driven growth in an active matter confined space, and related (bio) material realizations, *Biosystems*, **2019**, 176, 56-58.
5. P. Flory, Principles of Polymer Chemistry, *Cornell University Press*, 1953, ISBN 0-8014-0134-8.
6. M. Doi, S. F. Edwards, The Theory of The Polymer Dynamics. In *The International Series of Monographs on Physics*; Oxford University Press Inc.: New York, USA, 1986; 24–32.
7. Seror, J.; Zhu, L.; Goldberg, R.; Day, A.J.; Klein, J. Supramolecular synergy in the boundary lubrication of synovial joints. *Nat. Commun.* **2015**, 6, 6497, doi:10.1038/ncomms7497.
8. J. Katta, Z. Jin, E. Ingham and J. Fisher, *Biotribology of articular cartilage - a review of the recent advances, Med. Eng. Phys.* **2000**, 30, 1349-1363.
9. D. Rebenda, P. Čípek, D. Nečas, M. Vrbka, M. Hartl *24th International Conference ENGINEERING MECHANICS 2018. In The International Series of Monographs on Physics*; 709–712, doi: 10.21495/91-8-709
10. Moghadam, M. N.; Abdel-Sayed, P.; Camine, V. M.; Pioletti, D. P. Impact of synovial fluid flow on temperature regulation in knee cartilage. *Journal of Biomechanics* **2015**, 48, 370–374.
11. Oates, K.M.N.; Krause, W.E.; Jones, R.L.; Colby, R.H Rheopexy of synovial fluid and protein aggregation *J R Soc Interface* **2006**, 3, 167–174.
12. A. Liwo, S. Ołdziej, R. Kaźmierkiewicz, M. Groth, C. Czaplewski, Design of a knowledge-based force field for off-latice simulations of protein structure. *Acta Biochim. Pol.* **1997** *44*, No.3, 527–548.
13. S. Havlin, D. Ben-Avraham, Diffusion in disordered media, *Advances in physics* **2002**, *51*, No.1, 187–292.
14. Bhattacharjee S. M., Achille Giacometti A., Maritan A., Flory theory for polymers. *J. Phys.: Condens. Matter* **2013**, 25, 503101.
15. M. C. Baxa, E. J. Haddadian, J. M. Jumper, K. F. Freed, T. R. Sosnick, Loss of conformational entropy in protein folding calculated using realistic ensembles and its implications for NMR-based calculations. *P. Natl. Acad. Sci. USA* **2014** 111, No.43, 15396-15401.
16. P. Weber, P. Bełdowski, M. Bier, A. Gadomski Entropy Production Associated with Aggregation into Granules in a Subdiffusive Environment *Entropy* 2018, 20(9), 651 **2014**, *111*, 15396–15401.
17. S. R. Durell, A. Ben-Naim, Hydrophobic-Hydrophilic Forces in Protein Folding *Biopolymers* **2017**, 107(8).
18. Krieger, E.; Vriend, G. New ways to boost molecular dynamics simulations. *Journal of Computational Chemistry* **2015**, 36, 996–1007.
19. Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *The Journal of Physical Chemistry A* **2001**, 105, 9954–9960.
20. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of Computational Chemistry* **2003**, 24, 1999–2012.
21. A. Baruah, P. Rani, P. Biswas, Conformational Entropy of Intrinsically Disordered Proteins from Amino Acid Triads. *Scientific reports* **2015**, 5, 11740.

Sample Availability: Samples of the compounds ..... are available from the authors.